Toward 4G IP-based Wireless Systems: A Proposal for the Uplink

Tony Ottosson[1], Anders Ahlén[2], Anna Brunstrom[3], Mikael Sternad[2] and Arne Svensson[1]
[1] Dept. Signals and Systems, Chalmers Univ. of Techn., SE-412 96 Göteborg, Sweden
[2] Signals and Systems, Uppsala University, Box 528, SE-75 120 Uppsala, Sweden
[3] Dept. Computer Science, Karlstad University, SE-651 88 Karlstad, Sweden

Abstract

A packet switched wireless cellular system with wide area coverage and high throughput is proposed. The system is designed to be cost effective and to provide high spectral efficiency. It makes use of a combination of tools and concepts: Smart antennas both at base stations and mobiles, provide antenna gain and improve the signal to interference ratio. The fast fading is predicted in both time and frequency and a slotted OFDM radio interface is used, in which time-frequency slots are allocated adaptively to different mobile users, based on their predicted channel quality. This enables efficient scheduling among sectors and users as well as fast adaptive modulation and power control. We here outline the uplink of the radio interface. Calculations based on simplifying assumptions illustrate how the channel capacity grows with the number of simultaneous users and the number of antenna elements. A high capacity can be attained already for moderate numbers of users and base station/terminal antennas.

1 Introduction

Current cellular networks, such as GPRS, provide high coverage and mobility but with low throughput. On the other hand, wireless local area networks (WLANS) such as IEEE 802.11b and Hiperlan/2 provide high throughputs over limited distances and with limited mobility. Since the upcoming third generation systems (WCDMA) will only allow a limited increase in throughput, the gap will remain between services and capacity offered by local area systems and by cellular wide area coverage systems.

The current trend in communications is toward integration of systems and services. Users of fixed networks will be accustomed to high throughputs, good connections and innovative services. They might therefore become less willing to pay for the expensive and relatively low performance services that current and future cellular systems may offer. The current crisis for the telecommunications industry may be seen as one indicator of these worries.

A remedy would be to develop new wide area coverage and high mobility systems with much increased throughput (at least ten-fold) at a cost that is not higher (but preferably lower) than todays cellular networks. However, even if all cutting edge techniques, such as iterative decoding, adaptive antennas, and space-time coding are used, the WCDMA system will provide only a moderate increase in capacity (spectral efficiency) at the expense of an increase in complexity and thus cost.

The Wireless IP project [1] within the Swedish research program for Personal Computing and Communications (PCC) studies problems that are important in the evolution of UMTS toward higher data rates, as well as in future 4G technologies for mobile systems. Our goal is to improve the spectral efficiency for packet data, in particular IP traffic, with sufficient quality of service for various traffic classes. In the downlink we aim at a ten-fold improvement over existing technologies. In the uplink, which is more demanding due to the asynchronous behavior of the terminals, we aim for a somewhat lower capacity increase. To achieve this goal we propose to take advantage of the fast time and frequency varying channel properties, and to use multiple (smart) antennas in both base stations and mobile terminals.

In [2, 3], these issues were discussed for downlink traffic. The present companion paper considers a radio interface for the uplink. In Section 2 we discuss various design trade-offs and present a proposal for the uplink where prediction, smart antennas and scheduling are discussed together with adaptive modulation and power control. Sections 3 provides a preliminary analysis of this system in a special yet important scenario.

2 The Radio Interface Uplink

We aim at a spectrally efficient and highly flexible packet switched system providing high throughput in both the downlink and the uplink. Our goal is to provide a capacity of $50-100$ Mbit/s per sector of each base station in the downlink and only slightly lower capacity in the uplink.

In a system for packet data aimed at wide area coverage, we expect delay insensitive data and streaming traffic classes to constitute a large part of the traffic. We also expect a significant fraction of the users to be on the move, and thus to encounter time-varying channel properties. This can be transformed from a problem into an advantage, if we introduce a feedback system in which the channel properties are predicted and resources are allocated to users according to these channel predictions. Channels from different users will fade independently. With many users in the system, a channel could ideally almost always be allocated to users who at that time encounter favorable conditions. This effect is sometimes denoted multiuser diversity. Fairness and delay sensitivity of course represent constraints that reduce the attainable efficiency.

In an ideal situation, without discrepancies in the local oscillators of the mobile stations, the uplink could be designed in the same way as we proposed for the downlink [2, 3]. A synchronous system would thus allow us to use the downlink approach. However, synchronization among terminals and base stations is a very difficult task that imposes tremendous requirements on the hardware and on the algorithms of the system. An asynchronous uplink would therefore be more desirable since only a rough synchronization would be needed.

To enable scheduling of resources to obtain multiuser diversity in an asynchronous system, we propose the use of a slotted system, in which frames of fixed length are allocated to one or several users. At the beginning of each frame, all users are assumed to transmit pilot sequences which are orthogonal to sequences from the other users, see Figure 1. This scheme eliminates the risk of co-channel interference due to training data during payload data transmission. Based on these training sequences, multiuser channel estimation is performed.

We furthermore advocate the use of slotted FDD, to avoid interference problems encountered in asynchronous TDD systems. Each slot is proposed to have a payload data transmission frame of duration 0.5ms, approximately the same as the slot length in UMTS-HSDPA [4, 5]. We assume the same channel bandwidths as in UMTS, 5MHz in both the uplink and the downlink. With the modulation considered in Section 3 below, one frame will accommodate an IP packet of maximal size.

Sharing of resources among users within a frame could be obtained by either CDMA or FDMA. As for the downlink case outlined in [2, 3], we propose the use of *OFDM in the uplink*, with a spacing between sub-carriers such that no inter-channel interference occurs for the worst scenario (low coherence bandwidth). This enables the adaptation of modulation and power to the instantaneous SNR of different sub-carriers.

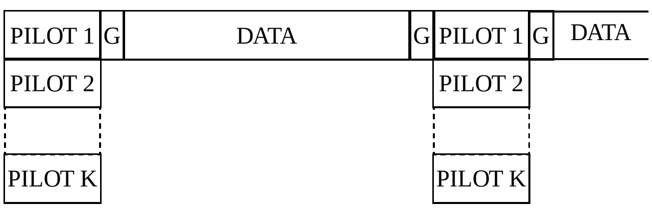

Figure 1: The slotted OFDM uplink. The data part of each frame has a duration of approximately 0.5ms. Orthogonal pilot sequences are transmitted simultaneously by K users, and are used by the base station for the estimation and prediction of channel quality and noise levels. After a guard time (G), one or several of the users are allowed to use the broadband data channel during the following data frame.

The use of OFDM also enables us to propose the use of *adaptive bandwidth allocation* among the users. In each slot, bins of sub-carriers are allocated to different simultaneous users (OFDMA), according to the predicted power on the different user sub-carriers. This will accommodate many simultaneous users with heterogenous and bursty traffic. However, different users have to be separated by significant guard-bands (unused sub-carriers). This will reduce the spectral efficiency, relative to that of the synchronously transmitting downlink.

The appropriate degree of partitioning of time-frequency frames between users is an interesting problem for further study. It is influenced by the granularity of the uplink packet traffic[1] and the efficiency loss due to guard-bands. If very bursty traffic, with many small packets, is handled by filling each time-frequency frame with data from many users, this will result in a significant loss of spectral efficiency.

[1]The uplink will frequently consist of bursty traffic. While browsing the web, for example, users often click on links which will generate short and infrequent packets. Such traffic will not fill a whole 0.5ms frame over the entire 5MHz bandwidth. However, some current trends indicate that file transfer and transmission of images are becoming frequent in the uplink. Since we do not know for sure what the future uplink traffic scenario would look like, it seems wise to outline a flexible system that can handle a wide variety of situations.

In order to handle such situations our preliminary suggestion is to partition the 5MHz uplink into a few narrowband (200 kHz) channels and one broadband channel (≥ 4 MHz). The narrowband channels use a statically allocated slot structure (as in e.g. GSM) and are thus well suited for control messages, real-time (especially voice) services and short data packets, while the broadband channel is more suited for best-effort and high speed services.

In order to perform a simplified analysis, see Section 3 below, we specialize to the case where users are given exclusive access to the broadband data channel, and where a single 200kHz narrowband channel is used.

A Channel Prediction

By predicting the short-term fading [6, 7], the transmission load among the users can be distributed in an optimal way, while fulfilling certain specified constraints on throughput and delay.

From extensive investigations on broadband measurement data, we have found that the received signal power distribution in the time domain as well as in the frequency domain can be predicted with reasonable accuracy for a time interval corresponding to a movement of up to half a carrier wavelength [6, 7].[2] Half a wavelength at 1900MHz would correspond to about 10 of our proposed 0.5ms slots at vehicular speeds of about 50 km/h. In such a situation we may allocate transmission loads among the users over the corresponding scheduling horizon. Interference-reducing schemes described in Sections B and C below improve the channel estimation accuracy and reduce the problem of predicting time-varying interference levels in packet data systems.

In our proposed uplink, a channel predictor located at the base station uses the pilot sequences of each frame to predict the time domain channel impulse responses for each user a few milliseconds into the future. FFTs of these predicted impulse responses provide predictions of the channel power for each sub-carrier and each user. The scheduling among users (Section C below) and the adaptive modulation (Section D) is based on the predicted channel and noise properties.

B Smart Antennas

We propose the use of smart antennas at the base station and preferably also in the terminals. We here assume only straightforward techniques of low computational complexity. In the base station, multiple antennas are used to form relatively narrow sectors (beams), and also to obtain diversity by Maximum Ratio Combining (MRC) using L diversity branches within each sector. In the mobile station simple lobe forming is used.

Transmitting in lobes from the mobile to the base station will increase the SNR and decrease the interference in other parts of the system. By using more narrow sectorization than commonly used today at base stations, scheduling among the users in each sector becomes less computationally complex as compared with scheduling of many users in a wide sector.

By using several antennas in each sector of the base station

[2]The achievable prediction performance depends on the available bandwidth [6, 7]. Here the discussion of predictors will assume a 4.8 MHz bandwidth for the broadband channel, according to the time-frequency structure proposed above.

the signal-to-noise-ratio can be increased further by the use of MRC, thus paving the way for the use of higher order modulation formats.

The antenna solutions presented here have been selected primarily for simplicity and robustness reasons. It is, of course, possible to introduce more sophisticated array signal processing schemes. The use of such schemes would contribute to further increase the capacity.

C Scheduling Among Users and Sectors

Scheduling is proposed to be used on different levels, as in the proposed downlink [2, 3]:

- *Among sectors:* Some pairs or groups of sectors at a base station or at nearby base stations may tend to create a large amount of mutual interference. In order to reduce this co-channel interference, time slots are allocated exclusively within the groups of such sectors, according to the traffic load in each sector.

  Information on the traffic load is exchanged infrequently via an inquiry procedure. After an inquiry to adjacent cells, the involved base stations determine the allocation of slots to be used by each base station in each sector.

- *Among users:* Based on the time slot allocation obtained from the inquiry process, the user scheduler will distribute time frames, and blocks of sub-carriers within frames, among the users of each sector based on their current average SNR predictions. Here different degrees of sophistication can be used to achieve different transmission goals [8, 9, 10, 14], such as throughput and user satisfaction with delay and QoS constraints.

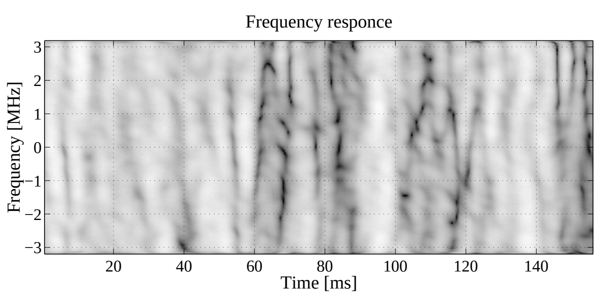

Figure 2: From [6]. Time-frequency representation of an estimated channel obtained from real measurement data on a 5MHz channel. White color denotes high power whereas dark color denotes low power. The dynamic range and the speed of the mobile is approximately 40 dB and 50 km/h, respectively. The coherence bandwidth is 4.9 MHz.

D Adaptive Modulation

The channel time-frequency pattern will depend on the scattering environment and on the velocity of the mobile terminal. Sometimes the whole frequency band will fade simultaneously (flat fading), see Fig. 2 whereas in other cases the channel will vary frequency selective as in Fig. 3. The fading patterns will be different for channels from different terminals.

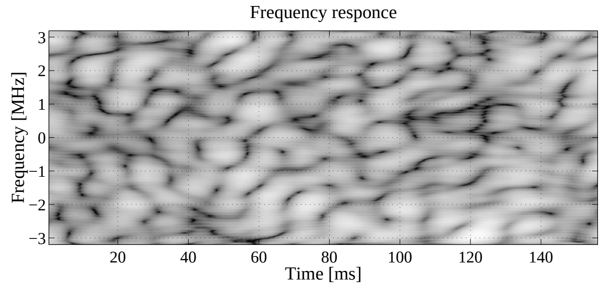

Figure 3: Conditions as in Figure 2, but the coherence bandwidth is here 0.6 MHz.

The modulation format used on a particular sub-carrier is selected according to the predicted user signal-to-noise-and-interference ratio (SNIR) on each sub-carrier or averaged over a group of sub-carriers.

We may also use an aggressive power control scheme, while keeping the interference on an acceptable level due to the above assumed use of lobe-formed transmission and the scheduling among sectors, which reduces the impact of the potentially worst co-channel interferers. For every time slot, the optimal power distribution among sub-carriers is obtained by waterfilling [11].

The performance evaluation below is based on a flat Rayleigh fading channel assumption, where the coherence bandwidth is large and the channel fades as in Fig. 2. The same modulation is therefore used on all sub-carriers within a slot interval.

3 Potential of the Proposed Uplink

We here estimate the spectral efficiency when K users transmit within one base station sector, under the following simplifying assumptions:

- A flat fading AWGN channel that is time-invariant within data frames, and Rayleigh fading between frames,
- accurate channel predictions are assumed,
- the broadband data channel is allocated exclusively to one user within each 0.5ms frame,
- the target service is reliable packet data transmission,
- fairness between users with different rates, QoS requirements, and delay constraints are neglected,
- some user does always have data to transmit,
- the allocated frames are fully utilized by payload bits,
- capacity loss due to guardtimes, frequency guardbands, cyclical prefixes, and training sequences are not taken into account.

The assumed scheduler works as a selection diversity scheme where the user with the best predicted SNR out of all K users will transmit its data in the whole frame and the whole frequency band, utilizing this resource fully.

In the base station receiver we assume MRC with L antennas within each sector. The resulting pdf after MRC and selection diversity of the received SNR (γ) can then be shown to be (see e.g. [12, pp. 186–194] for guidance in performing this calculation)

$$p(\gamma) = \frac{K e^{-\gamma/\bar{\gamma}} (\frac{\gamma}{\bar{\gamma}})^L}{\gamma \Gamma^K(L)} \left(\Gamma(L) - \Gamma(L, \tfrac{\gamma}{\bar{\gamma}}) \right)^{K-1}, \qquad (1)$$

where $\bar{\gamma}$ is the average SNR per receiver antenna, and

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t}\, dt \qquad (2)$$

$$\Gamma(a, z) = \int_z^\infty t^{a-1} e^{-t}\, dt \qquad (3)$$

are the gamma and incomplete gamma functions, respectively.

To exploit the improved SNR after diversity combining effectively, we use an adaptive modulation system (AMS). We assume $M = 2^{k_i}$-ary QAM using the $N = 5$ different modulation schemes: QPSK ($k_0 = 2$), 16-QAM ($k_1 = 4$), 64-QAM ($k_2 = 6$), 128-QAM ($k_3 = 7$), and 256-QAM ($k_4 = 8$). Further improvement could be made by using more modulation levels and by introducing coding (see e.g. [15, 14]). For simplicity we will, however, confine the discussion to the above mentioned assumptions.

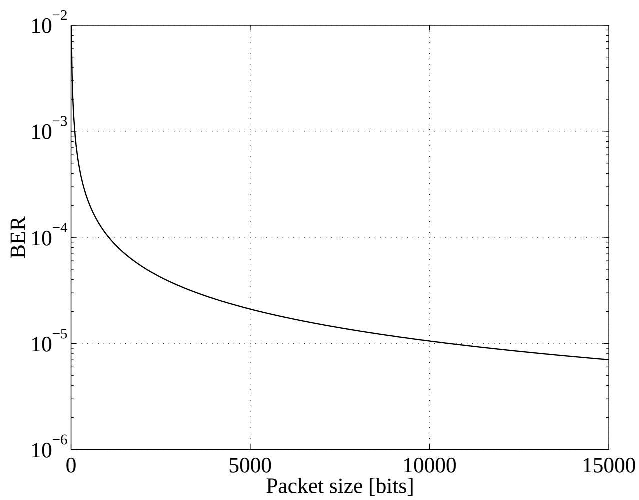

Figure 4: Required BER as a function of the packet size (in bits) to achieve a packet error rate of $P_E = 0.1$.

Since the target service is reliable packet transmission we need an ARQ scheme. For simplicity we will assume simple ARQ without any coding. The packet error rate can then be expressed as a function of the bit error rate (BER) as

$$P_E = 1 - (1 - \text{BER})^F, \qquad (4)$$

where F is the packet size in bits. The number of retransmissions is thus given by $1/(1 - P_E)$. For the ARQ to be useful, a reasonable value of the packet error rate would be $P_E = 0.1$. The actual throughput would then be 90% of the raw throughput. By inspecting Fig. 4 we conclude that a BER $= 8.7 \times 10^{-6}$ works for all packet sizes up to about 12000 bits. This corresponds to the maximal IP packet size (1500 bytes) and is a suitable packet size for a time-frequency bin of 0.5ms × 4.8 MHz. (In the downlink, shorter link level packets of about 1000bits are allocated to time-frequency bins of 0.5ms × 200kHz.)

The spectral efficiency using AMS and simple ARQ can thus be expressed as (see [13])

$$\eta \geq (1 - P_E) \sum_{i=0}^{N-1} k_i \int_{\gamma_i}^{\gamma_{i+1}} p(\gamma)\, d\gamma \qquad (5)$$

where η is measured in bits/s/Hz and where $k_i = \log_2(M)$ is the number of bits transmitted using uncoded M-ary QAM modulation. The γ_i values define the SNR regions in which to use the different modulation levels, such that 2^{k_i}-ary QAM is used in the SNR interval $[\gamma_i, \gamma_{i+1})$, with $\gamma_N = \infty$. No transmission is performed at SNRs below γ_0. These γ_i values are given in Table 1 for the target BER $= 8.7 \times 10^{-6}$ (see also Fig. 5). Compared to the downlink [2, 3], the required SNR is approximately 1 dB larger. The inequality in (5) comes from the fact that once the SNR is above the γ_i level, the BER (and thus the packet error rate) is lower than the target of 8.7×10^{-6}.

QPSK $k_0 = 2$	16-QAM $k_1 = 4$	64-QAM $k_2 = 6$	128-QAM $k_3 = 7$	256-QAM $k_4 = 8$
γ_0	γ_1	γ_2	γ_3	γ_4
9.65	13.50	17.85	20.17	22.57

Table 1: The γ_i function in dB. γ_i is calculated as the SNR required to achieve BER $= 8.7 \times 10^{-6}$ using $M = 2^{k_i}$ QAM.

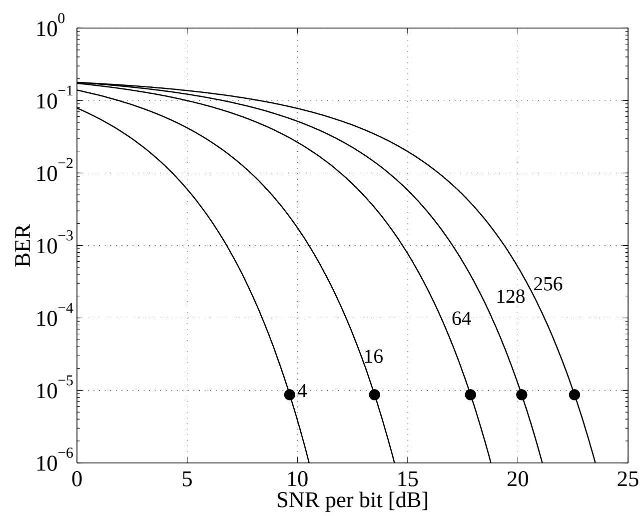

Figure 5: BER for M-ary QAM on an AWGN channel. The required SNRs (the γ_i) to achieve BER $= 8.7 \times 10^{-6}$ are shown as points.

The spectral efficiency in (5) is evaluated numerically for $\bar{\gamma} = 10$ dB per receiver antenna and information bit. The result is presented in Fig. 6 and shows that the spectral efficiency improves dramatically with an increasing multiuser selection diversity and with the number of receiving antennas in each sector of the base station. Already for two receiver antennas per base station sector, and a moderate number of users, the spectral efficiency is considerable. Compared to the downlink [2, 3], we loose, in this simplified example, around half a bit/sec/Hz.

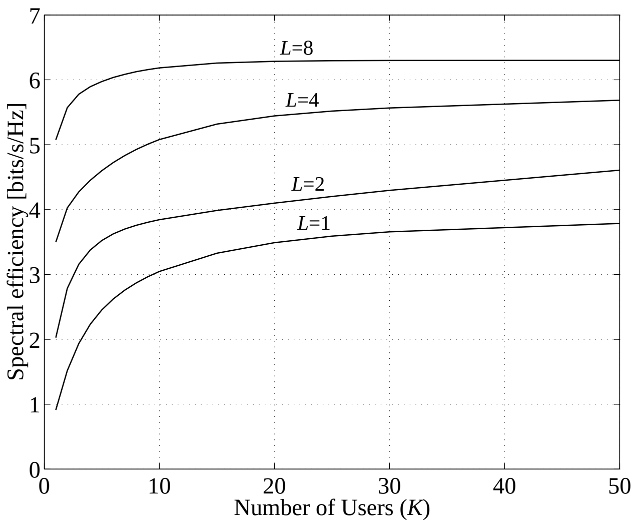

Figure 6: Spectral efficiency using AMS and simple ARQ with Lth order MRC diversity in the mobile and Kth order of selection diversity between the users. The SNR per receiver antenna and information bit is $\bar{\gamma} = 10$ dB. A 10% reduction of the raw bit error rate due to ARQ is taken into account.

In theory the spectral efficiency within one sector, η, should be multiplied by the number of sectors, S, used per cell. However, even if the interference is reduced by using many sectors, we cannot use all time-frequency bins in all sectors. Thus, we need to divide with a reuse factor which most likely is smaller than the number of sectors used. The spectral efficiency of each cell would therefore be larger than is indicated by Fig. 6, but smaller than ηS.

4 Conclusions

A radio interface uplink proposal for high throughput and wide area coverage has been presented. It makes use of the instantaneous channel conditions of several users and distributes the transmission loads by the use of adaptive bandwidth allocation so that an efficient channel usage is obtained. Calculations performed under certain simplifying assumptions indicated that a high spectral efficiency is attainable. Modifications on higher levels, and exchange of information between levels, provide ample possibilities for further improvements which are at present investigated within our project group.

References

[1] The Wireless IP Project within the SSF PCC (Personal Computing and Communication) program www.signal.uu.se/Research/PCCwirelessIP.html

[2] T. Ottosson, A. Ahlén, A. Brunström, M. Sternad and A. Svensson, "Towards 4G IP-based Wireless Systems," *Wireless World Research Forum Workshop*, Paris, Dec. 2001.

[3] T. Ottosson, A. Ahlén, A. Brunström, M. Sternad and A. Svensson, "Towards 4G IP-based Wireless Systems," Future Telecommunication Conference Beijing, China, Nov. 2001.

[4] S. Parkvall, E. Dahlman, P. Frenger, P. Beming, M. Persson, "The evolution of WCDMA towards higher speed downlink packet data access," *Nordic Radio Symposium 2001 (NRS 01)*, Nynäshamn, Sweden, April 3-5 2001.

[5] 3rd Generation Partnership Project (3GPP), Technical Specification Group Radio Access Network: "UTRA High Speed Downlink Packet Access (Release 4)," *3GPP TR 25.950 V2.0.0*, March 2001.

[6] T. Ekman, *Prediction of Mobile Radio Channels*. Licentiate Thesis, Signals and Systems, Uppsala University, Sweden, Dec. 2000.

[7] M. Sternad, T. Ekman and A. Ahlén, "Power prediction on broadband channels," *IEEE Vehicular Technology Conference, Spring 2001 (VTC2001S)*, Rhodes, Greece, May 6-9 2001.

[8] N. C. Ericsson, "Adaptive modulation and scheduling over fading channels," *GlobeCom99*, Rio de Janeiro, December 5-9, 1999. pp. 2668-2672.

[9] N. C. Ericsson, S. Falahati, A. Ahlén and A. Svensson, "Hybrid type-II ARQ/AMS supported by channel predictive scheduling in a multi-user scenario," *IEEE Vehicular Tech. Conf. (VTC2000-Fall)*, Boston, MA, USA, Sept. 24-28, 2000, pp. 1804-1811.

[10] N. C. Ericsson, S. Falahati, A. Ahlén and A. Svensson, "Scheduling and adaptive transmission for the downlink in 4G systems," *Proc. Future Telecommunication Conference*, Beijing, China, November 2001.

[11] D.N.Tse, "Optimal power allocation over parallel Gaussian channels," *ISIT*, Ulm, Germany, June 1997.

[12] M. D. Yacoub, *Foundations of Mobile Radio Engineering*, CRC Press, 1993.

[13] S. T. Chung and A. J. Goldsmith, "Degrees of freedom in adaptive modulation: a unified view," *IEEE Trans. Commun.*, Vol. 49, No. 9, pp. 1561–1571, September 2001.

[14] N. C. Ericsson, *On Scheduling and Adaptive Modulation in Wireless Communications*. Licentiate Thesis, Signals and Systems, Uppsala University, Sweden, June 2001.

[15] S. Falahati, *Convolutional Codes for Wireless Packet Data Systems*. Licentiate Thesis, Chalmers University of Technology, Feb. 2000.