TSG-RAN Working Group 1 meeting #17
Stockholm, Sweden
November 20-23 2000

TSGR1#17(00)1428

Agenda item: AH24: HSDPA

Source: Ericsson

Title: Performance Comparison of Chase Combining and Incremental Redundancy for HSDPA

Document for: Discussion

---

1. Introduction

In this paper we investigate and report on the performance of different fast HARQ schemes considered for HSDPA. The performance of HARQ type-II schemes with Chase combining and Incremental Redundancy (IR) is compared. It is shown that the basic performance difference between Chase and IR increases both with the code rate and the size of the modulation alphabet. For R = ¼ encoded QPSK the gain from using IR compared to Chase combining is very small while, for R = ¾ encoded 64-QAM the gain is much larger and may be significant, especially if there is large errors in the link adaptation.

The outline of the paper is as follows: In Section 2 we present the simulation assumptions and the transport block processing scheme used for the simulations. In Section 3 the performance differences of Chase combining and incremental redundancy (IR) are evaluated. Finally, the conclusions are presented in Section 4.

2. Transport-Block Processing and Assumptions used in the Simulations

The transport-block processing scheme used in the simulations is shown in Figure 1. At the left hand side of Figure 1 we multiplex together K_1 transport blocks, each having the size $N_{\mathrm{TrBlk}} = 320$, before adding $m_{\mathrm{CRC}} = 16$ CRC bits[1]. Thus, at the input of the Turbo encoder we have $N_{un\mathrm{coded}} = K_1 \times N_{\mathrm{TrBlk}} + m_{\mathrm{CRC}}$ bits. The turbo encoder then generates the N_{coded} coded bits.

---

[1] In this paper we have used $N_{\mathrm{TrBlk}} = 320$ and $m_{\mathrm{CRC}} = 16$. However, the scheme used for the evaluation allows for an arbitrary N_{TrBlk} and m_{CRC}.

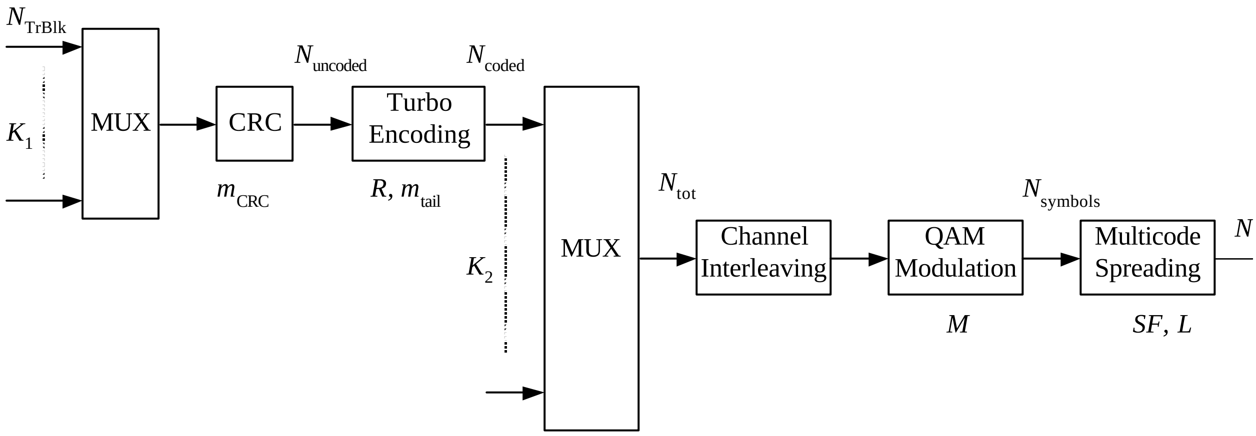

Figure 1: The transport block processing scheme used in the simulations.

In this study we assume an HSDPA TTI of 0.667 ms (i.e. one slot). With Spreading factor SF and L multi-codes, this corresponds to $N_{symbols} = (L / SF) \times 2560$ QAM symbols at the output of the modulator. The number of bits at the input of the QAM modulator must then equal $\log_2(M) \times N_{symbols}$. The channel interleaver does not add or remove any bits and therefore we shall have $N_{coded} = N_{symbols} / K_2$ bits at the output of each of the K_2 turbo encoders. In this study we have chosen to use a random channel interleaver with a interleaving depth corresponding to one slot. Since N_{coded} must be an integer number we must place some restriction on the selection of K_2. For example, we may select K_2 from the set {1, 2, 4, 5, 8, 10, 16} which are all factors of $N_{chip} = 2560 = 2^9 \times 5$. We may also set K_2 equal to L or $\log_2(M)$, which always will be factors of $N_{symbols}$.

One reason for having the second multiplexing after the channel encoder is that the current 3GPP turbo encoder is limited to block sizes less than 5110 bits. Another reason is that we may select a fixed value of K_1 (e.g. $K_1 = L$) and use different values of K_2 for different modulation and coding schemes (MCSs). The reason for having a fixed value of K_1 is that this makes it possible to perform the turbo encoding before the MCS is known. Only the rate matching (i.e. puncturing or repetition) will then depend on the MCS. In this report however we use $K_2 = 1$ as long as N_{coded} ? 5110.

In order to efficiently support MCSs containing only one, or a few number of transport block(s) we have extended the current 3GPP Turbo coder with an additional coding polynomial , see Figure 2.

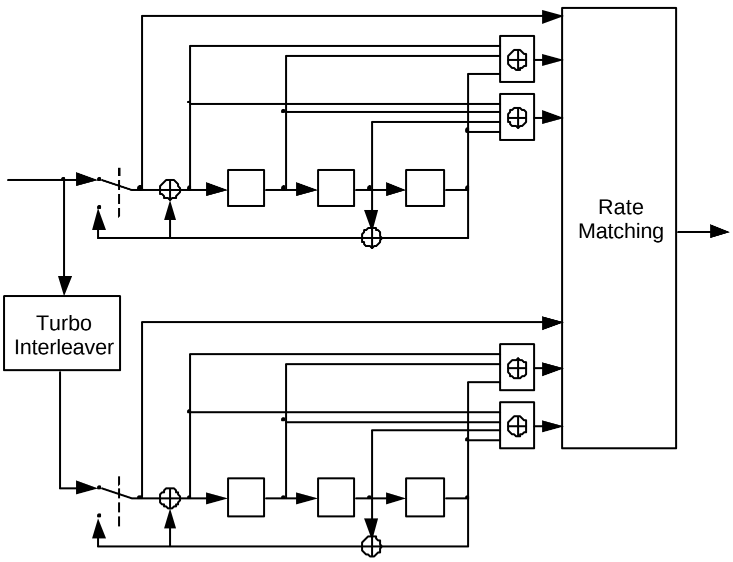

Figure 2: The rate R = 1 / 6 parent turbo encoder used in the simulations.

The total number of transport blocks transmitted in one slot is $K_{tot} = K_1 \times K_2$. For each K_{tot} we must select the best combination of K_1, K_2, and M. K_{tot} can be any value between 1 and 32 (assuming $L / SF = 3/4$). Simulations (not presented in this document) have shown that K_2 shall be selected as small as possible. There are two reasons for this: Firstly, this minimises the number of CRC and tail bits that wee need to transmit and hence we can instead use a slightly lower code rate. Secondly, the block size of the turbo code becomes larger when K_2 decreases which also results in better performance.

Note that the specific parameters used in the simulations are example values only. The purpose of this paper is to compare the performance differences between Chase combining and incremental redundancy. The conclusions drawn from these simulations will be valid also for other parameter values.

2.1 Simulation Assumptions

The simulation assumptions are listed in Table 1. Note that 50% of the maximum transmission power is allocated to other channels. Since these channels are orthogonal to the HS-DSCH (no time dispersion), the other channels are not transmitted in these simulations.

Table 1: Simulation Parameters

Parameter	Value	Comment
Spreading Factor of HS-DSCH	As defined	
Propagation conditions	AWGN	Stationary mobiles.
Overhead Power Allocation (CPICH)	10% (-10 dB)	
Allocated power for HS-DSCH	40% (-4 dB)	40% of NodeB´s maximum transmission power
Closed loop Power Control	OFF	
Frame Length	0.67 ms	
Ior/Ioc	Variable	
Channel Estimation	Ideal	
Number of CRC bits	16	Note: 24 bits is specified in the Link simulation assumptions document.
Tail bits	6	Parity bits for the tails are not transmitted.
Max no. of iterations for Turbo Codes	8	
Metric for Turbo Code	Max	
Turbo Code Rates	Variable	
Input to Turbo Decoder	Soft	As specified in the Link simulation assumptions document.
Turbo Interleaver	As in 3GPP	
Information Bit Rates Simulated (Kbps)		As defined
Number of Multicodes Simulated		As defined

2.2 MCS Candidates

In Table 2 we show the best possible combination of K_1, K_2, and M for all possible values of K_{tot}. The selection of these parameters is based on simulating different combinations and selecting the one that gives the best slot error rate performence. Note that for K_{tot} ? 16 the number of information bits per slot is larger than the maximum size of the internal turbo encoder interleaver (i.e. larger than 5110 bits) and hence $K_2 = 1$ is not possible. Also note that K_2 must be a factor in $N_{tot} = \log_2(M) \times N_{symbols}$ (see Figure 1). Therefore, it is not possible to find a valid combination of K_1, K_2, and M for K_{tot} = 17, 19, 23, 29, and 31 unless the restriction on the maximum turbo encoder interleaver size is removed. It is of course also possible to drop the assumption that the second multiplexing shall combine K_2 blocks of equal size. This will however require different turbo interleaver sequences for the different input blocks to the second multiplexer, and this option has not been investigated in this study. Given the assumed constraints we must select a set of MCSs from the MCS candidates in Table 2.

Table 2: MCS candidates for L / SF = 3/4, m_{crc} = 16, N_{TrBlk} = 320, N_{chip} = 2560.

K_{tot}	K_1	K_2	M	R	Info bits per slot	Info rate (Mbps)
1	1	1	4	0.09	320	0.48
2	2	1	4	0.17	640	0.96
3	3	1	4	0.25	960	1.44
4	4	1	4	0.34	1280	1.92
5	5	1	4	0.42	1600	2.40
6	6	1	4	0.50	1920	2.88
7	7	1	4	0.59	2240	3.36
8	8	1	4	0.67	2560	3.84
9	9	1	16	0.38	2880	4.32
10	10	1	16	0.42	3200	4.80
11	11	1	16	0.46	3520	5.28
12	12	1	16	0.50	3840	5.76
13	13	1	16	0.54	4160	6.24
14	14	1	16	0.58	4480	6.72
15	15	1	16	0.63	4800	7.20
16	8	2	16	0.67	5120	7.68
17	-	-	-	-	5440	8.16
18	9	2	16	0.75	5760	8.64
19	-	-	-	-	6080	9.12
20	10	2	64	0.56	6400	9.60
21	7	3	64	0.59	6720	10.08
22	11	2	64	0.61	7040	10.56
23	-	-	-	-	7360	11.04
24	12	2	64	0.67	7680	11.52
25	5	5	64	0.70	8000	12.00
26	13	2	64	0.72	8320	12.48
27	9	3	64	0.75	8640	12.96
28	14	2	64	0.78	8960	13.44
29	-	-	-	-	9280	13.92
30	15	2	64	0.84	9600	14.4
31	-	-	-	-	9920	14.88
32	8	4	64	0.89	10240	15.36

3. Performance with Soft Combining and Incremental Redundancy Combining

3.1 Link-Level Performance on AWGN Channels

We have chosen ten different modulation and coding schemes from Table 2 and labelled them MCS-1 to MCS-10. The MCS table was designed by choosing every third MCS candidate in Table 2 (i.e. K_{tot} = 3, 6, 9, …, 30). Note that these ten MCSs were chosen only for the purpose of comparison. The performances of these ten MCSs are shown in Figure 3.

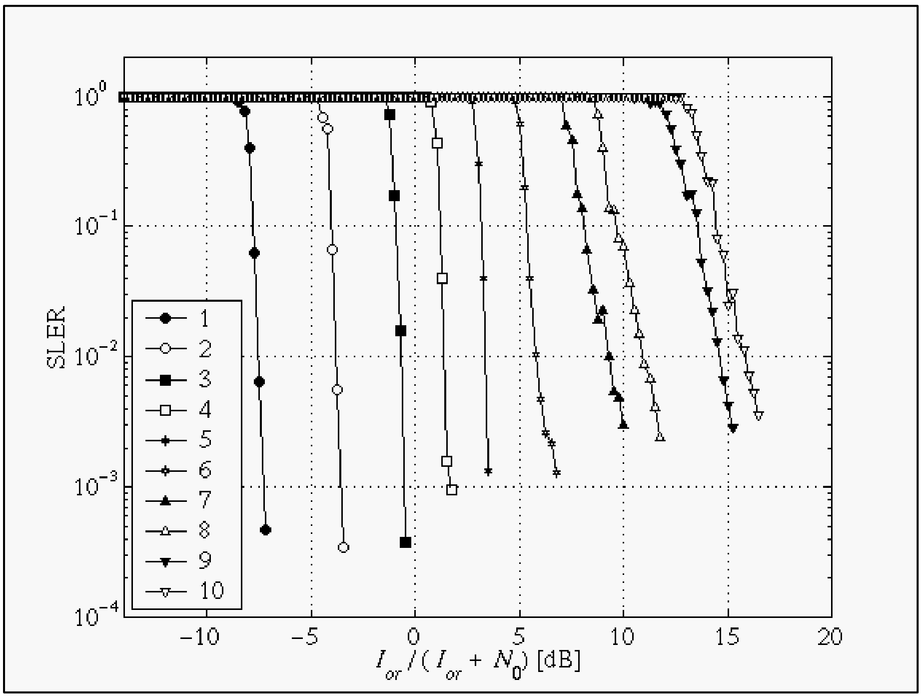

Figure 3: Static (AWGN) channel. Simulated slot error rate (sler) versus $\hat{I}_{or}/(I_{or} ? N_0)$ in dB for MCS-1 to MCS-10 (as defined in Table 3) m_{crc} = 16, *SF* = 4, *L* = 3, 8 iterations.

For each of the ten MCSs we have compared the performance obtained with Chase combining (i.e. identical retransmissions and soft combining) and Incremental Redundancy (IR). As an example we show the results for MCS-1 (K_{tot} = 3) in Figure 4, MCS-6 (K_{tot} = 18) in Figure 5, and MCS-9 (K_{tot} = 27) in Figure 6. Results are shown for the original transmission (first), second, third, and fourth transmission. We see that the difference is small for MCS-1 and much larger for MCS-6 and MCS-9. The performance gains obtained by using IR compared to Chase combining with these ten MCSs are summarised in Table 3.

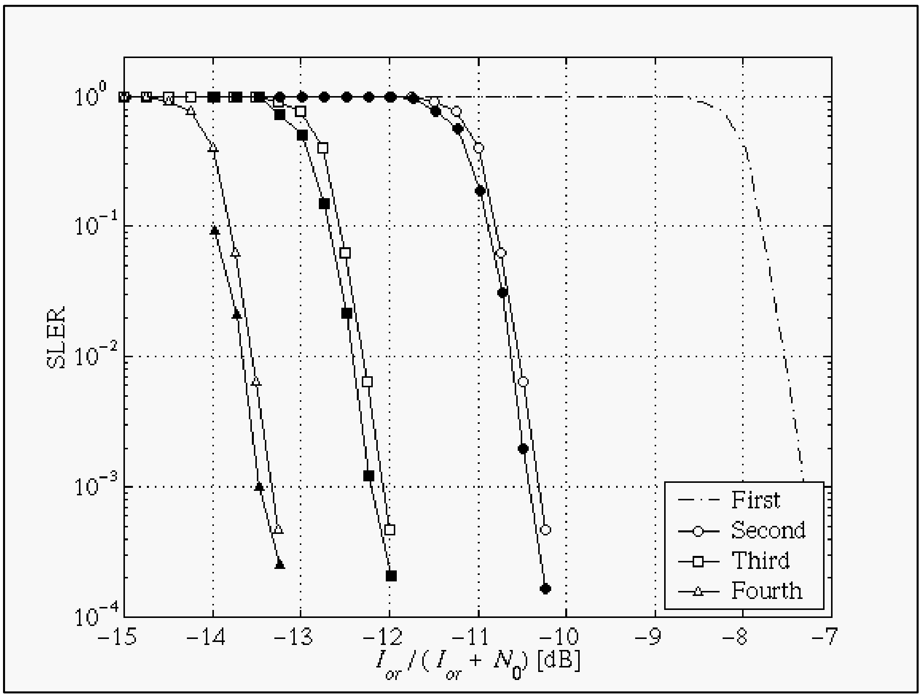

Figure 4: K_{tot} = 3. Static (AWGN) channel. Simulated slot error rate (sler) versus $\hat{I}_{or}/(I_{or}$? $N_0)$ in dB. K_2 = 1, M = 4, m_{crc} = 16, SF = 4, L = 3, 8 iterations, white and black markers are used for Chase combining and IR respectively.

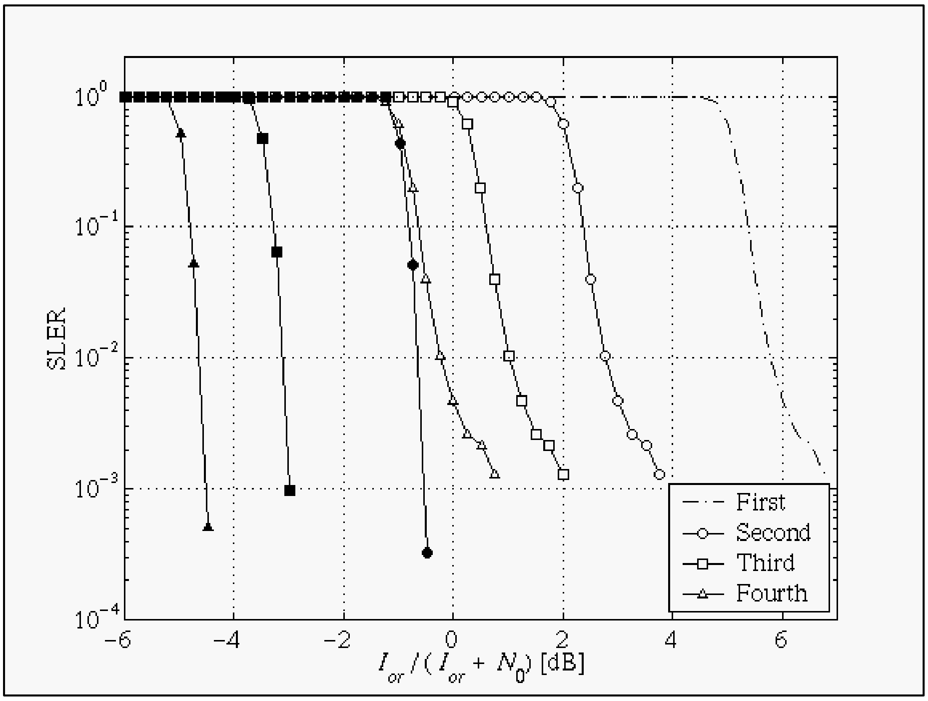

Figure 5: K_{tot} = 18. Static (AWGN) channel. Simulated slot error rate (sler) versus $\hat{I}_{or}/(I_{or}$? $N_0)$ in dB. K_2 = 2, M = 16, m_{crc} = 16, SF = 4, L = 3, 8 iterations, white and black markers are used for Chase combining and IR respectively.

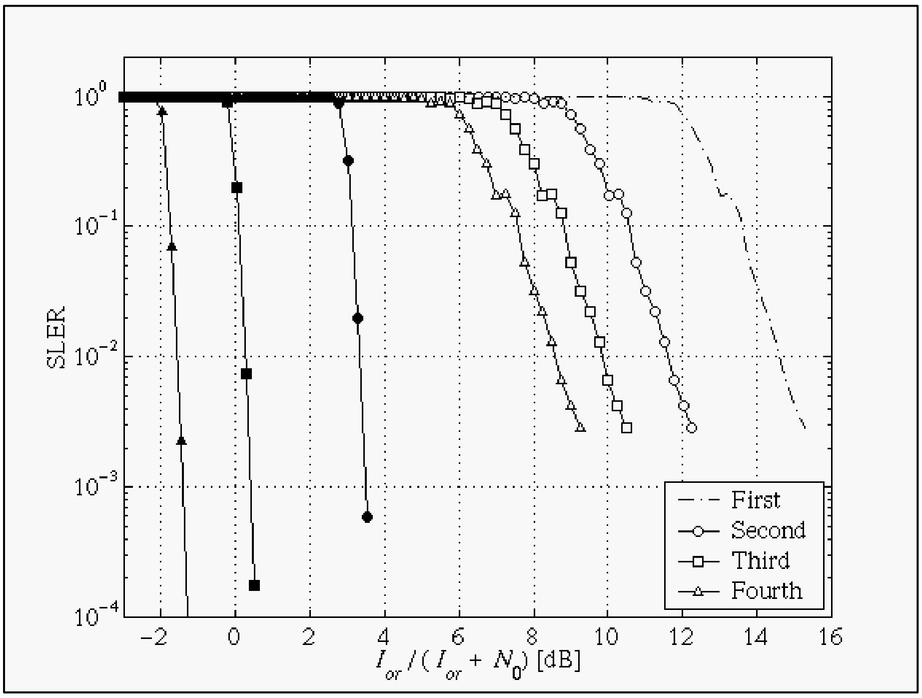

Figure 6: K_{tot} = 27. Static (AWGN) channel. Simulated slot error rate (sler) versus $\hat{I}_{or} /(I_{or} ? N_0)$ in dB. K_2 = 3, M = 64, m_{crc} = 16, SF = 4, L = 3, 8 iterations, white and black markers are used for Chase combining and IR respectively.

Table 3: Achievable gain with IR over Chase combining at BLER = 10^{-2}.

MCS #	K_{tot}	M	R	IR Gain 2nd Trans. [dB]	IR Gain 3rd Trans. [dB]	IR Gain 4th Trans. [dB]
1	3	4	0.25	0.1	0.1	0.1
2	6	4	0.50	0.9	1.0	1.0
3	9	16	0.38	1.0	1.0	1.0
4	12	16	0.50	1.6	1.8	1.8
5	15	16	0.63	2.3	2.9	2.9
6	18	16	0.75	3.4	4.1	4.4
7	21	64	0.59	3.0	5.8	6.0
8	24	64	0.67	5.7	6.8	7.1
9	27	64	0.75	8.3	9.6	10.2
10	30	64	0.84	8.8	9.0	10.9

From Table 3 we conclude that IR gives significantly better performance compared to Chase combining when the code rate is larger than 0.50. We also see that for the same code rate we obtain higher gains for higher modulation orders. An intuitive explanation for why this is the case is that the Gray encoding of the QAM constellation makes effectively one bit out of four an uncertain bit in the 16-QAM case. In the 64-QAM it is only one bit out of six that is an uncertain bit. The soft demodulation thus results in (relatively) fewer uncertain bits in the 64-QAM case than in the 16-QAM case. Since the turbo encoder uses soft metric it can easily correct the uncertain bits if other bits are more certain (i.e. they have larger soft values). But apparently a high rate code is not as effective in taking advantage of this as a more powerful low rate code is. The gains obtained for MCS-1 to MCS-4 are smaller than, or equal to 1.8 dB. Therefore, from a performance versus complexity point of view it may be a good idea to use IR only for the higher MCSs (i.e. MCS-5 to MCS-10). Note that the performance differences between Chase combining and IR are measured at a SLER of 1%. Since the curves with IR (in most cases) falls down steeper than those for Chase combining the relative difference will be smaller if the comparison is done at a higher SLER value.

3.2 Throughput and Delay Performance

In this subsection we evaluate the delay and throughput performances when using Chase combining and IR. The plots presented in this section uses C / I instead of $\hat{I}_{or}/(I_{or}\ ?\ N_0)$ on the x-axis. The relation between C / I and $\hat{I}_{or}/(I_{or}\ ?\ N_0)$ is (with the assumptions in Table 1) as follows: $C/I\ ?\ 4\ ?\ \hat{I}_{or}/(I_{or}\ ?\ N_0)$ [dB]. As a reference we show in Figure 7 the throughput without any type of soft value combining (i.e. HARQ Type-I) for the ten different MCSs in Table 3. We see that without soft combining the throughput falls rapidly to zero when the channel quality is insufficient.

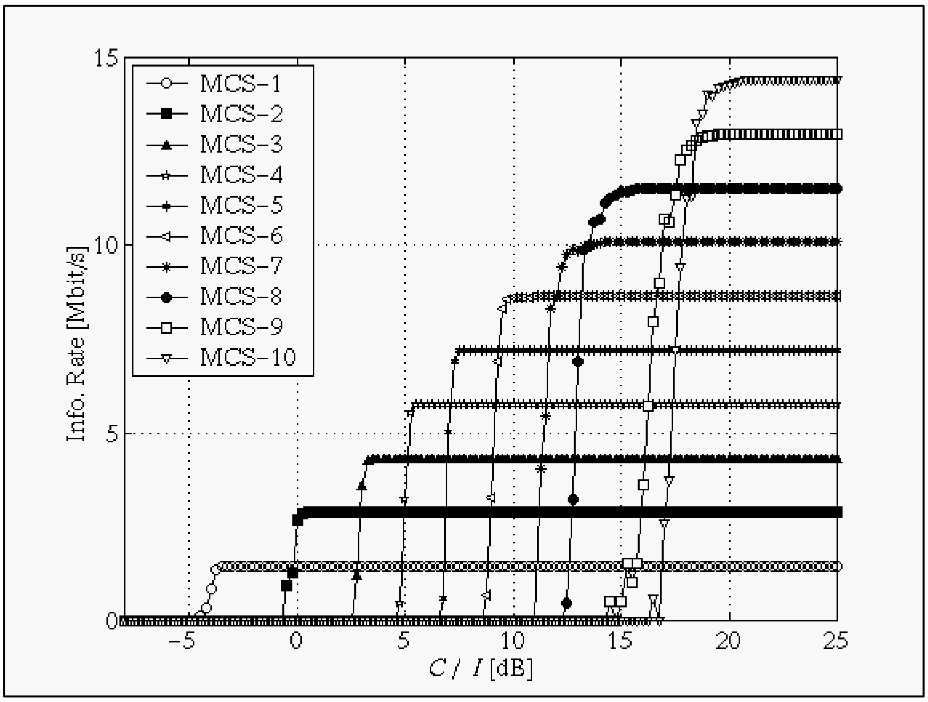

Figure 7: Information rate in Mbit/s versus *C / I* in dB for MCS-1, …, MCS-10. Selective repeat Type-I HARQ (no soft combining).

In Figure 8 we show the throughput results obtained when using Chase combining. We see that the throughput curves falls down smother compared to the case without soft combining. The throughput results with IR are shown in Figure 9. Comparing the IR and the Chase throughput results we see that the curves falls down to a first plateau in exactly the same way, but with IR the plateau lasts much longer than for Chase combining.

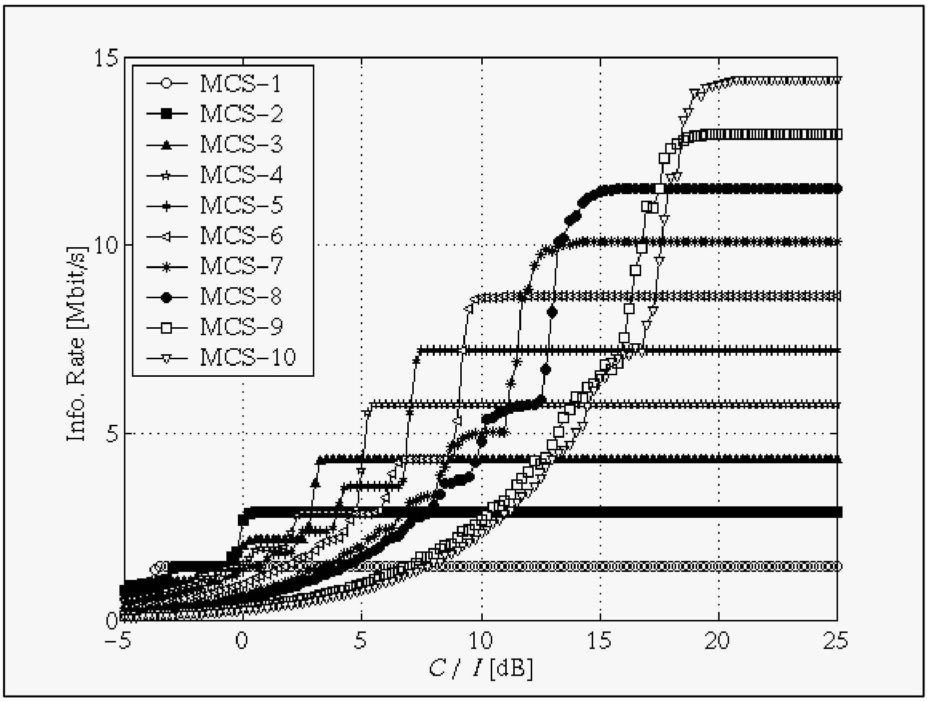

Figure 8: Information rate in Mbit/s versus *C* / *I* in dB for MCS-1, …, MCS-10. Selective repeat Type-II HARQ with Chase combining.

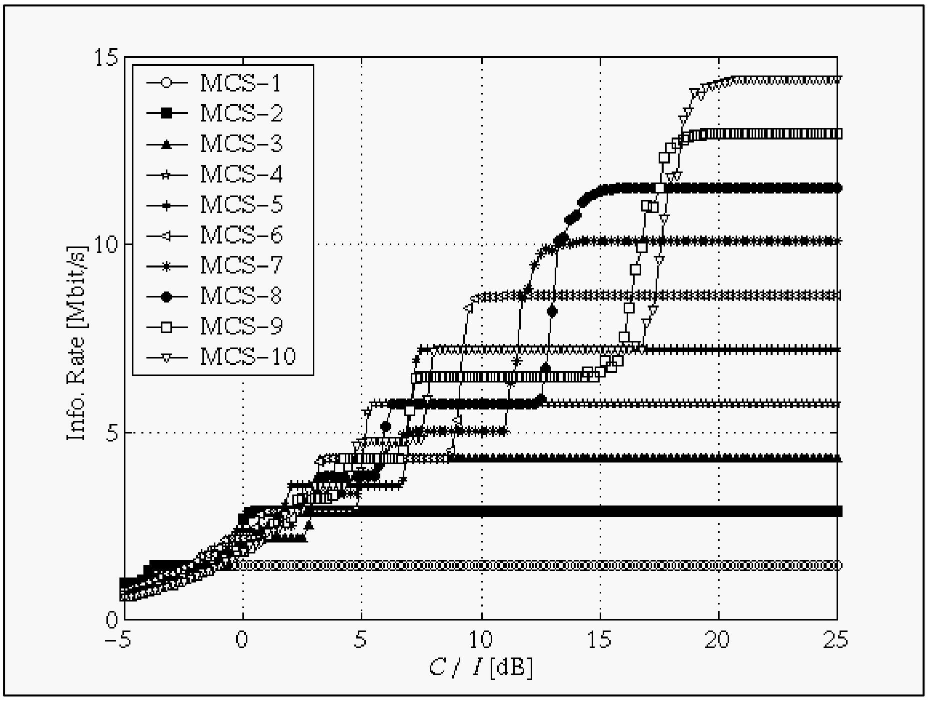

Figure 9: Information rate in Mbit/s versus *C* / *I* in dB for MCS-1, …, MCS-10. Selective repeat Type-II HARQ with incremental redundancy (IR).

The delay characteristics when using Chase combining are shown in Figure 10, and the corresponding results when using IR are shown in Figure 11. Comparing these two figures we clearly see a large performance advantage for the IR scheme when a high MCS is used.

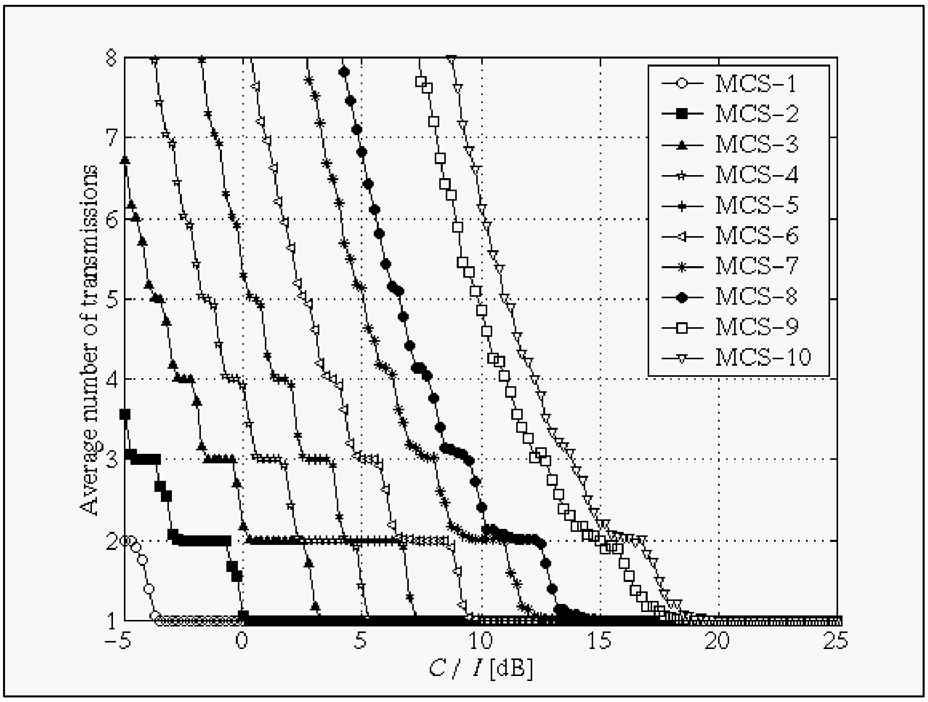

Figure 10: Average number of transmissions versus *C / I* in dB for MCS-1, …, MCS-10. Selective repeat Type-II HARQ with Chase combining.

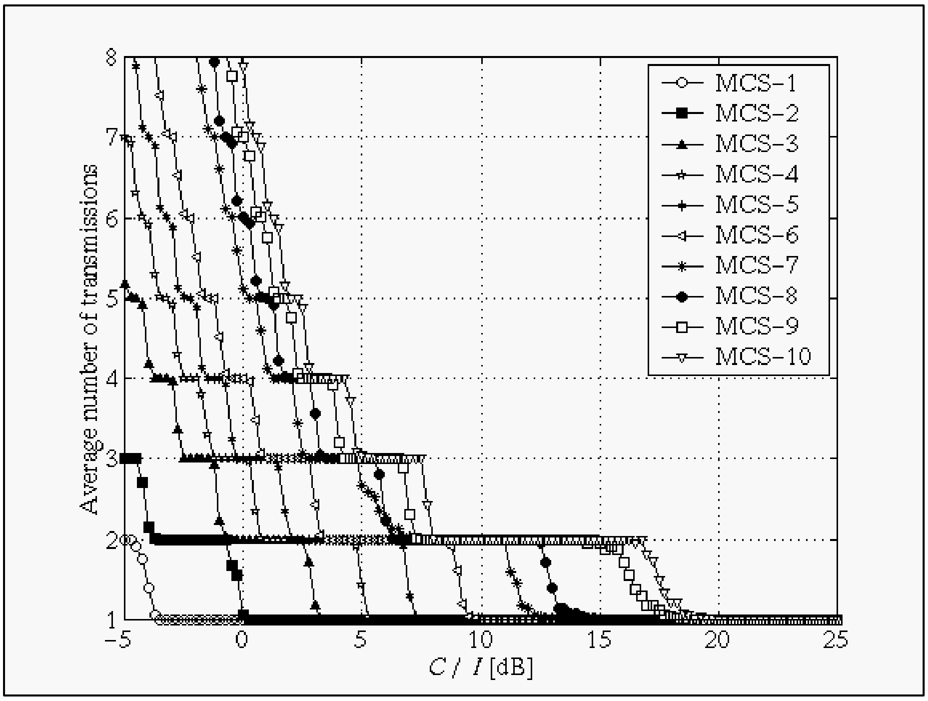

Figure 11: Average number of transmissions versus *C / I* in dB for MCS-1, …, MCS-10. Selective repeat Type-II HARQ with incremental redundancy (IR).

3.3 Results with Imperfect Channel Quality Estimates

In Figure 12 to Figure 15 we show the performance obtained when the channel quality estimates are noisy. The MCS is based on the C/I estimate and is selected from a lookup table containing the following switch points: [-0.75 2.83 5.05 7.04 9.25 11.81 13.23 17.5 18.4] dB. The C/I estimate is assumed to be normally distributed in a decibel scale with a mean value equal to the true C/I, and a standard deviation of ?. The troughput results are shown in Figure 12 and Figure 13 below.

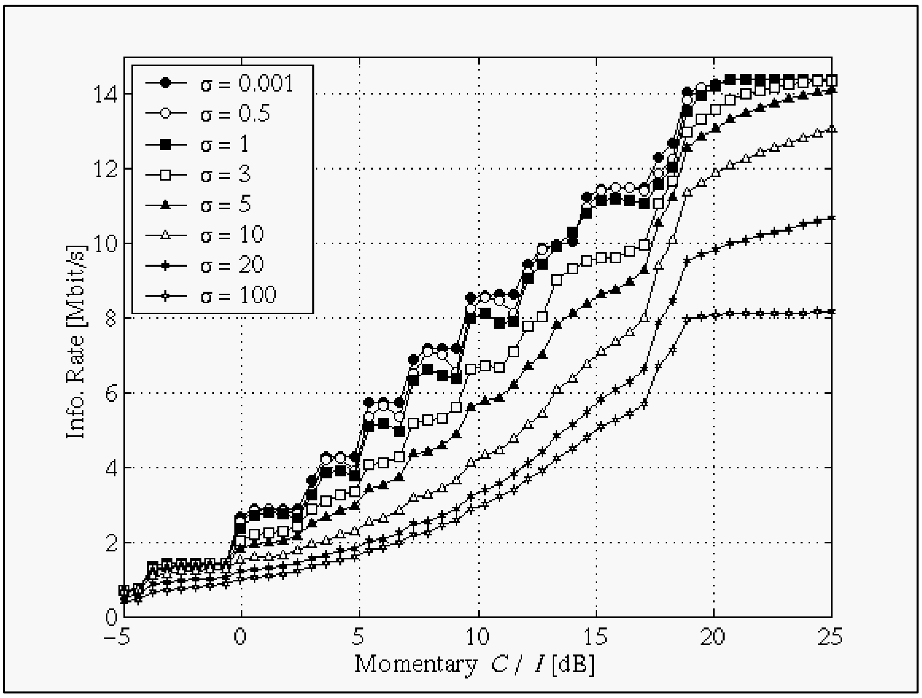

Figure 12: Information rate in Mbit/s versus momentary *C / I* in dB for ? = 0.001, 0.5, 1, 3, 5, 10, 20, and 100. Selective repeat Type-II HARQ with Chase combining.

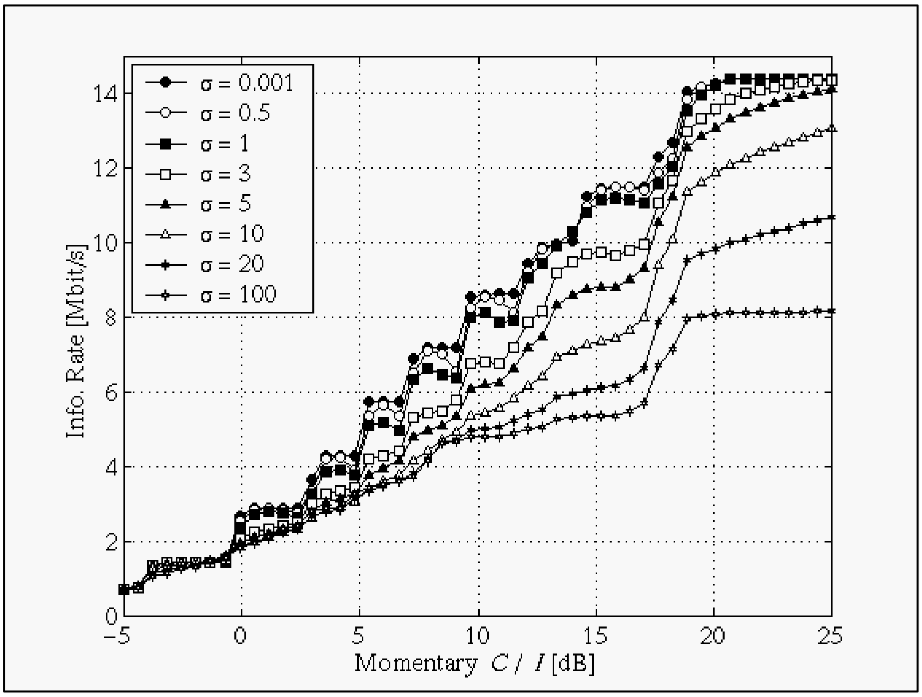

Figure 13: Information rate in Mbit/s versus momentary *C / I* in dB for ? = 0.001, 0.5, 1, 3, 5, 10, 20, and 100. Selective repeat Type-II HARQ with incremental redundancy (IR).

Comparing the throughput results in Figure 12 and Figure 13 we see that for large values of ?, the IR based scheme performs significantly better. For small values of ? the performance difference between the Chase combining and IR systems is small. The results from Figure 12 and Figure 13 are summarised in Table 4 below. For ? < 5 there is no significant difference between Chase combining and IR. From Table 4 we see that both the *C* / *I* distribution of the UEs as well as the variance of the channel quality estimate is important when evaluating the performance gains that can be obtained with IR.

Table 4: Relative throughput gain in percent with IR compared to Chase combining for ? **= 5, 10, 20, and 100.**

	? = 5	? = 10	? = 20	? = 100
C / *I* = -5 [dB]	5 %	14 %	36 %	69 %
C / *I* = 0 [dB]	8 %	23 %	49 %	79 %
C / *I* = 5 [dB]	10 %	33 %	69 %	100 %
C / *I* = 10 [dB]	8 %	28 %	55 %	63 %
C / *I* = 15 [dB]	3 %	5 %	6 %	7 %

Similar conclusions as was drawn from Figure 12 and Figure 13 can be made by comparing the delay characteristics shown in Figure 14 and Figure 15.

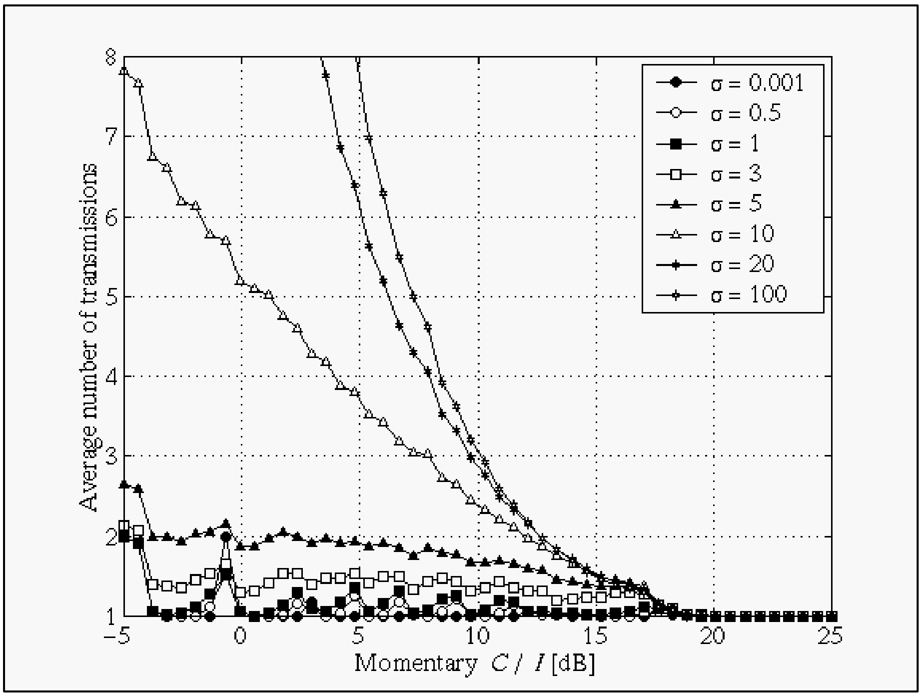

Figure 14: Average number of transmissions versus momentary *C* / *I* in dB for ? **= 0.001, 0.5, 1, 3, 5, 10, 20, and 100. Selective repeat Type-II HARQ with Chase combining.**

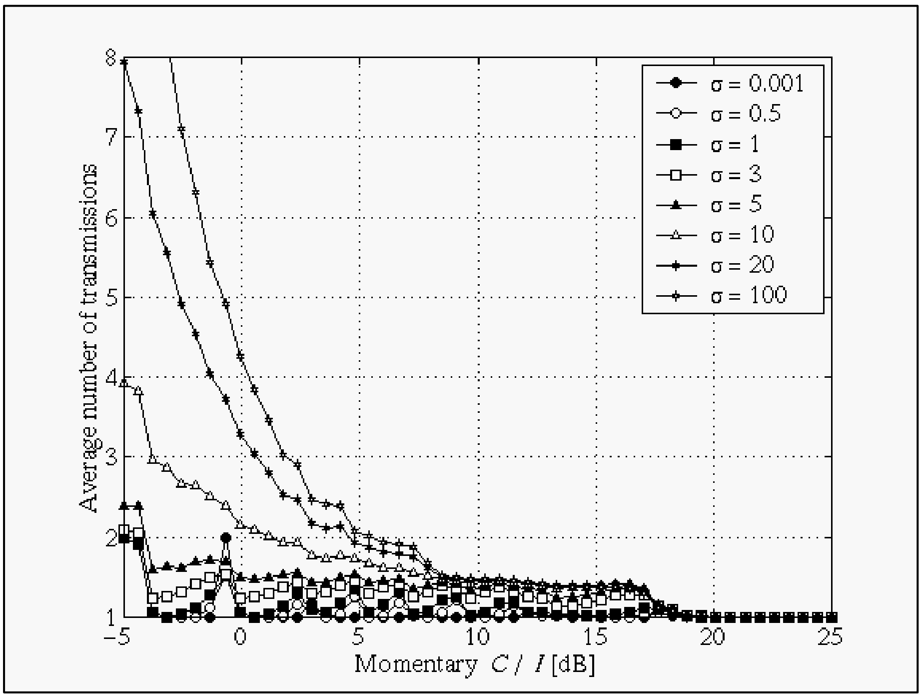

Figure 15: Average number of transmissions versus momentary *C / I* in dB for ? = 0.001, 0.5, 1, 3, 5, 10, 20, and 100. Selective repeat Type-II HARQ with incremental redundancy (IR).

4. Conclusions and Discussion

In this paper we have shown that the link-level performance of IR is significantly better than Chase combining for high channel coding rates and for large modulation sizes. We have also shown that for the same code rate we obtain higher gains with IR compared to Chase combining for higher modulation orders. It has also been shown that this link-level gain may be important in case the error in the channel quality estimate is large. There is no significant difference between IR and chase for low coding rates and small modulation sizes or when the error of the channel quality estimate is small.

Given these observations we conclude that IR needs to be studied further for HSDPA. It is necessary to compare the performance difference between IR and Chase when the channel quality changes from the first transmission and the retransmission. It is expected that the difference between the two type-II HARQ approaches will be smaller in this case.

It might be reasonable to consider a combination of IR and chase combining where Chase combining is allowed for the lower MCSs and IR is required only for the higher MCSs. Thus a UE that can only support the lowest MCSs does not have to implement IR since there is anyway no gain with IR for the low MCSs. Another alternative is to use IR only for the first retransmission and then repeat the bits from the first two transmissions in the following retransmissions. This would provide most of the gain that IR can give while still reducing the buffering requirement in the UE compared to a system that uses IR for all retransmissions.

The error of the channel quality estimate that is used to select the MCS might very well have large errors in scenarios when the UEs move at large speeds. Furthermore, in a system dominated by packet data users the interference level will be bursty and thus difficult to predict. IR will in scenarios like these be a more robust solution and thus it might be possible to use the high MCSs more often.