Performance Evaluation of Proportional Fair Scheduling Algorithm with Measured Channels

Troels B. Sørensen and Manuel Rubió Pons
Department of Communication Technology
Aalborg University
Niels Jernes Vej 12, DK9220 Aalborg Øst, Denmark
{tbs}@kom.aau.dk

***Abstract*— Motivated by the promising performance results of the *proportional fair* (PF) packet scheduling algorithm, often quoted in connection with WCDMA High Speed Downlink Packet Access (HSDPA), we have performed an evaluation of the PF gain in comparison to the simpler *round robin* (RR) scheduler when subjected to measured channel traces. Specifically, we applied measured signal fading recorded from GSM cell phone users making calls on an indoor wireless office system. Different from reference channel models, these measured channels have much more irregular fading between users, which as we show by simulation leads to vanishing system throughput gain of PF over RR; the *maximum throughput* (Max TP) scheduler on the other hand, is much less influenced by the measured channels, but also less fair in giving equal throughput to the users, as is the RR scheduler in comparison to PF.**

Proportional Fair Scheduling; HSDPA; Channel Measurements

I. INTRODUCTION

Packet scheduling is one of the key components of the HSDPA concept [1]. Specially, the *proportional fair* (PF) scheduling algorithm has received much attention due to its favourable trade-off between total system throughput and fairness in throughput between scheduled users [2][3]. PF (and *maximum throughput*, Max TP) scheduling benefits from multi-user diversity [3][4] in which the scheduler tracks the channel fluctuations of the users and only schedules users when their instantaneous channel quality is near the peak. PF has been studied extensively under well-defined propagation channel conditions, such as flat fading channels with Rayleigh and/or Rician type of fading [5], or the ITU Vehicular and Pedestrian channels [7] which are typically applied in standardization work [6]. The concern addressed in this paper is how well the PF algorithm performs under measured channel conditions, representative of an indoor office environment [8], where the fading is typically less well defined. Particularly, in measurements we often see a mixture of fast and slow fading which is not easily described by an analytical model.

There are several good surveys of HSDPA with explanations of the basic operation and typical performance numbers, e.g. [1] and [2], and some also with discussions of aspects for practical implementation, e.g. [6]. In this paper we apply only the basic principles, which we briefly discuss, and refer the reader to the above mentioned references for more detailed HSDPA information. Specifically, in our evaluation we take into account the adaptation delay between measurements (channel quality indication) and link adaptation and scheduling, but otherwise make simplifications of real HSDPA operation.

In the following Section II, we first describe the operation of the three different schedulers. Next, in Section III we describe the channels with particular emphasis on the characteristics of the measured signal fading and how it was applied for the evaluation of scheduler performance. In continuation of this, Section IV explains about the simulation and evaluation procedure which was used to generate the results presented and discussed in Section V. Finally, we conclude our paper in Section VI.

II. SCHEDULER OPERATION

For our HSDPA implementation we assume capabilities corresponding to an HSDPA *user equipment* (UE) with data rate capability up to 3.6 Mbps; the highest rate is achieved per *transmission time interval* (TTI) sending data packets with 16QAM modulation and 5 spreading codes [2]. Each packet may be modulated and coded differently so that at any scheduling instant the optimum *modulation and coding scheme* (MCS) can be selected for the packet transmission; this selection at the base station is based on feedback information from the user equipment at previous scheduling instants, informing the base station about the instantaneous channel quality at the UE. By the instantaneous channel quality we understand the TTI averaged SINR at the UE, while in real HSDPA operation the UE itself estimates the supported data rate in correspondence of this SINR and informs the base station via the *Channel Quality Indicator* (CQI) [2]. The time between measurement of the channel quality and the scheduling instant is referred to as the *adaptation delay* and is counted in TTIs as scheduling and link adaptation decisions take place on TTI boundaries.

For the PF scheduler, we base the scheduling decision in TTI n+1 on the Relative Channel Quality Indicator (RCQI) defined in (1).

$$RCQI_k\,[n+1] = \frac{\tilde{R}_k\,[n+1]}{T_k\,[n]} \qquad (1)$$

$\tilde{R}_k\,[n+1]$ is the (estimated) supported throughput for the k'th user at TTI n+1.

$T_k[n]$ is the averaged throughput delivered to the UE in the past given by (2).

$$T_k[n] = T_k[n-1]\left(1 - \frac{1}{t_a}\right) + \frac{1}{t_a} R_k[n] \qquad (2)$$

As seen in (2) $T_k[n]$ at TTI *n* is the output of an autoregressive filter with the actual user throughput $R_k[n]$ in TTI *n* as the input, and with the averaging controlled by the time-constant t_a; by actual throughput we refer to the data successfully transmitted. By always selecting the user which maximises the RCQI this scheduling algorithm leads to proportional fairness between users as discussed in [3].

The Max TP scheduler [2] operates in a similar fashion except that the RCQI is directly proportional to the supported throughput, hence at any scheduling instant only the best user is served. Since throughput maps monotonically with the SINR this scheduler is equivalent to a maximum SINR scheduler.

In contrast to PF and Max TP, the RR scheduler is blind to the instantaneous channel quality. Instead, RR shares resources (time, code, and power) equally between users in a round robin fashion. In this case the channel quality feedback is used only for determining the optimum MCS for the packet transmission.

III. Radio Channels – Reference and Measured

A. *Reference channels*

For our work we initially decided to consider two different reference channels, which traditionally are used for simulation in low time-dispersive and low mobility radio channels, thus similar of the environment in which the measurements were obtained. One is the well-known flat Rayleigh fading channel with Jakes (Doppler) spectrum and the other, almost similar but with reduced fading dynamics, the ITU Pedestrian A (*Ped. A*) channel model [7], both considered at 3 km/h. The models serve both in comparison to the measurements and as a validation of our evaluation procedure. Whereas the fading depth in flat Rayleigh fading may go some 20 dB below the average level (20.0 dB at the 99th percentile), this happens seldom in the *Ped. A* channel (14.1 dB at the 99th percentile); here we have assumed ideal WCDMA rake receiver processing in which effectively two, chip-spaced, taps of the *Ped. A* channel are coherently combined (modified ITU *Ped. A*). We found in our evaluation that the two channels lead to only minor differences in the performance of the schedulers and we therefore include only Rayleigh as a reference in the following.

B. *Measured channels*

The measured channels are derived from uplink GSM signal strength measurements made in a multi-floored office building in Copenhagen, Denmark [8]. Due to the large coherence bandwidth typical of such indoor environments we assumed that these narrowband signal strength measurements completely determine the performance of the WCDMA HSDPA signal transmission in the wider 5 MHz bandwidth. Similarly, the application of the measurements to the downlink direction of transmission is justified in this case by the similar surroundings of user equipment and base station antennas. The signals originated from GSM cell phone users making calls on an indoor wireless office system, a setting which, although not typical, we see of some relevance for HSDPA terminal usage, e.g. voice over IP.

The measurements were conducted over a 6 weeks period in which almost 10000 calls were recorded. Not all however were eligible for analysis due to too low signal strength or the presence of interfering signals. From this raw material we selected calls from 4 randomly chosen weekdays; 178 of these calls had sufficient signal strength (noise discrimination) and sufficient duration to be useful for the evaluation.

Without too much detail on the measurement procedure we will simply mention that the signal sampling and detection left us with a sequence of samples representing the logarithmic envelope of the cell phone signals taken every GSM frame (60/13 ms). As will be explained further in the next section, we finally selected segments of 100 s from each call and normalised these to an average SINR for the evaluation. Fig. 1 shows two such examples of very different channels – one with almost constant fading and the other with highly variable fading.

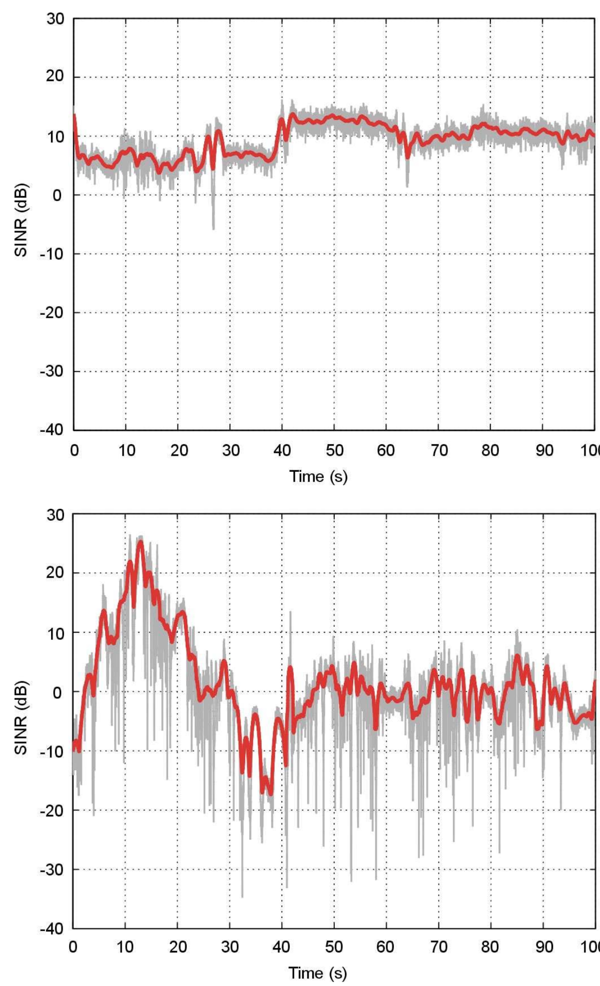

Figure 1 Examples of measured channels normalised to 10 dB average SINR. The curve shows the fading at a resolution of 60/13 ms with a 1 s moving average shown on top (solid black curve).

It is clearly evident that the measured fading is less regular as the fading produced by the reference channels; furthermore the measured channels contain a slow, or shadow, fading component.

On the curves we have shown also a 1 s moving average of the same data. From these averaged data we computed the 10^{th} and 90^{th} percentiles within each 100 s segment and plotted the percentiles against each other as in Fig. 2. There is a clear trend of increased variability below and above the average level (10 dB SINR) whenever one of them increases, but also some exceptions with increased variability in mostly one direction. The maximum variability measured between the 10^{th} and 90^{th} percentile levels is on the order of 40 dB, but many channels have considerably lower average variability – down to 1 dB!

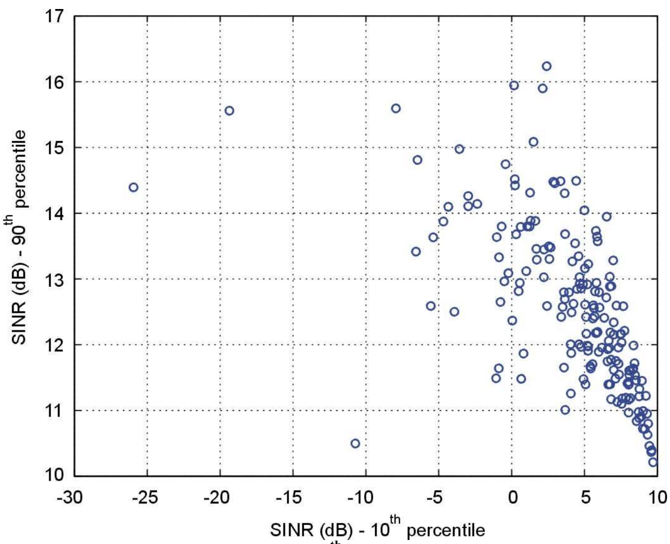

Figure 2 Computed 10 versus 90^{th} percentiles for the 1 s averaged channels of all 178 calls. Average SINR equals 10 dB.

For the faster variations, defined here as the original samples less the average variations, we analysed the rms power gradient [9] per call to find that the maximum observed gradients lie in the range from 0.12 to 0.16 dB/ms. For comparison, the rms power gradient for Jakes model at 3 km/h can be calculated to be 0.19 dB/ms at the measurement frequency [9]. However, many calls exhibit a considerably lower rms gradient (e.g. the first one in Fig. 1). Quite in accordance with the examples in Fig. 1 the faster variations exhibit fades some 35 dB below and 5 dB above the 1 s averaged level.

For the evaluation later on we need to point out the presence of a frequency discrimination induced noise component resulting from the narrowband detection technique used to obtain the measurements. After post-processing this noise component was reduced to an equivalent rms deviation of 0.54 dB with characteristics much alike a log-normal random process with *iid* (independent identically distributed) samples. We shall later discuss the influence to the performance results based on this model.

IV. EVALUATION PRCOEDURE

In this section we list the main assumptions for our evaluation. Our strategy was to evaluate the influence of the fading variability to the performance of PF. Given typically used values for the averaging time constant in (2) on the order of 1 s, and considering the finite time duration of the observed measured channels, we decided to base our evaluation on a 100 s long segment from each of the 178 calls. To be consistent with the format of the measured channels we arbitrarily assume a TTI equal to the GSM frame duration. Consequently, the 100 s maps into a sequence of 21666 TTIs over which we apply the scheduling and link adaptation decisions for the three different schedulers.

To define an appropriate operating point and to emphasise the fading variability between channels we normalised each to a specific average SINR within the link adaptation range of the considered HSDPA system. This average SINR, $\langle SINR \rangle_{100s}$, can be interpreted as in (3) where P_{HS_DSCH} is the power allocated to the *high speed downlink shared channel* (HS-DSCH), SF_{16} is the spreading factor of the HS-DSCH, and $P_{other} + P_{nosie}$ is the power of interference (other cell) plus noise which we assume Gaussian. The notation $\langle\ \rangle_{100s}$ is used to represent the linear time-average over 100 s.

$$\langle SINR \rangle_{100s} = SF_{16} \frac{P_{HS\_DSCH}}{P_{other} + P_{nosie}} \left\langle |E|^2 \right\rangle_{100s} \qquad (3)$$

Together these terms can be seen as a normalisation factor that adjusts the linear time-average of the measured signal strength, $\left\langle |E|^2 \right\rangle_{100s}$, to obtain the HS-DSCH SINR after despreading. By default the factor was adjusted to obtain $\langle SINR \rangle_{100s}$ = 10 dB.

We determine the throughput in the WCDMA/HSDPA system through mapping tables (link adaptation) for the low time-dispersive case that link the optimum MCS, and hence supported data rate and throughput, to the TTI averaged SINR at the UE (the TTI instantaneous values resulting after the normalisation in (3)). The actual throughput is determined similarly by linking the *block error probability* (BLEP) to the SINR, thereby conditioning the selected data rate to obtain the actual throughput in each TTI. Apart from the optimum MCS, the link adaptation also decides on the optimum number of multi-codes for transmission; in this paper we limit ourselves to a maximum of 5 multi-codes in correspondence of a UE with 3.6 Mbps data rate capability [2]. Fig. 3 shows the resulting optimum link adaptation function in comparison to the Shannon bound. For this curve HARQ has been assumed with a first transmission BLEP target of 10% and maximum two transmissions to successfully receive the data. Also shown overlayed on the link adaptation curve are selected mapping curves for the BLEP versus SINR. The curves are given for different MCS and different number of multi-codes; for the most robust MCS (QPSK, effective code rate ¼) we utilise the whole range of multi-codes starting with one code for the leftmost curve and ending with 5 for the rightmost, whereas on ly 4 and 5 codes, respectively, are used for 16QAM.

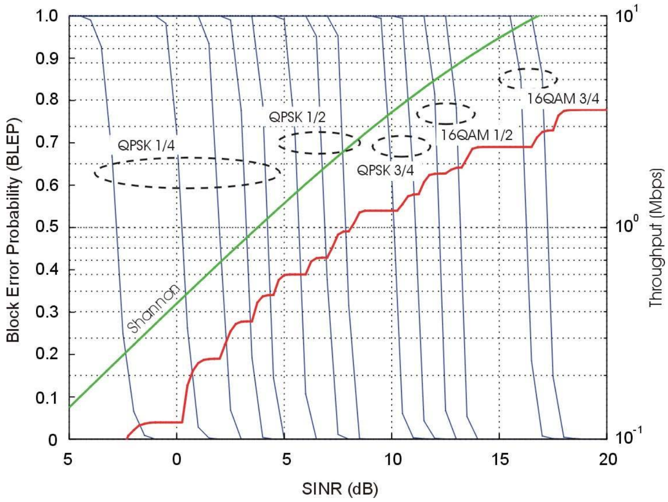

Figure 3 SINR to throughput mapping together with the BLEP performance of the corresponding MCS. The curves represent simulated WCDMA rake receiver performance in the ITU *Ped. A* channel which we apply for the low time-dispersive case.

From the set of BLEP mapping curves there is a clear correspondence with the transitions in the link adaptation curve. With 5 codes and 16 QAM, effective code rate ¾, we can confirm that the maximum instantaneous throughput is 3.6 Mbps. Furthermore, it can be observed that over most of the dynamic range the optimum link adaptation performs within 4-5 dB of the Shannon bound, but will deviate from this in the evaluation because of the link adaptation delay and different assumptions for the (H)ARQ mechanism. By default we use an *adaptation delay* of 2 TTIs (2 times 60/13 ms) which is in the realistic range for minimum link adaptation/packet scheduling delay [6].

Our comparison of the different schedulers is based on the total throughput each obtains over the duration of the respective channels (100 s), that is, summed over all users and all TTIs, and in each comparison with reference to RR. This is done for different number of simultaneously scheduled users, from 2 up to 12. In each of these cases we evaluate the linear gain as the ratio between total throughputs for a number of different combinations of users, or more explicitly, different combinations of channels. For the reference channels which are independent between users we run 50 different combinations - Monte Carlo (MC) simulations – while due to the more irregular fading we evaluate 100 combinations for the measured channels (out of a large number of possible combinations). What is shown in the results is the linear gain averaged over the different combinations, in comparison to the RR scheduler. At the beginning of the simulations each user scheduled by the PF scheduler is assumed to have an average throughput (T_k) in correspondence of the average SINR divided by the number of simultaneously scheduled users. We further assume that users have full buffers and that failed transmissions are handled by a simple ARQ protocol.

A. *Performance validation with reference channels*

To validate our evaluation procedure we first ran simulations with the reference channels and changed the average SINR, the *adaptation delay*, and the time constant t_a to confirm that results were in correspondence of known behaviour. Fig. 4 shows the linear gain of PF and Max TP over RR when we normalised the reference channels to different average SINR values. Similar to other published results, e.g. [3] and [6], we see the well-known multi-user diversity gain increasing with the number of simultaneously scheduled users, but reaching saturation for high number of users. Due to the more efficient SINR to throughput mapping in the lower SINR ranges (cf. Fig. 3) we see increased gain for lower average SINR values. Also we confirmed that the increase of adaptation delay from the default of 2 TTIs leads, as expected, to lower gain of PF in accordance with the decorrelation in the Jakes model. The influence of the time constant t_a which is default 1 s is discussed in the following section.

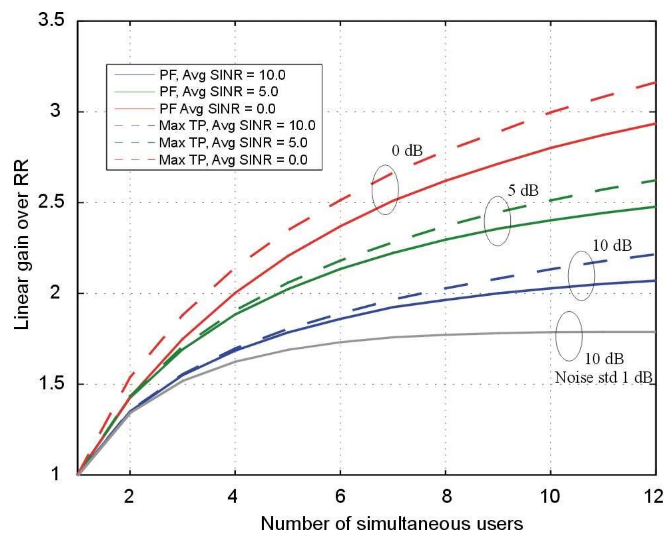

Figure 4 Scheduling gain of PF relative to RR in reference channels with results shown for different average SINR values (t_a = 1 s). Also, for PF at 10 dB average SINR we have shown the performance when uncertainty is introduced to the channel quality at the UE (1 dB std. - see Section V).

V. PERFORMANCE WITH MEASURED CHANNELS

A. *Results*

The first thing we notice when the schedulers are compared with the measured channels is the high variation in the gain of PF relative to RR, depending on the specific combination of channels considered. Moreover, as an average consideration, PF performs worse than RR with gain values below 1 as it is seen in Fig. 5; basically, for all the average SINR operating points the gain over RR is below 1, and there is no multi-user diversity gain. Quite interestingly, opposite to the results in Fig. 4, Fig. 5 shows that an increase of the average SINR leads to the better performance.

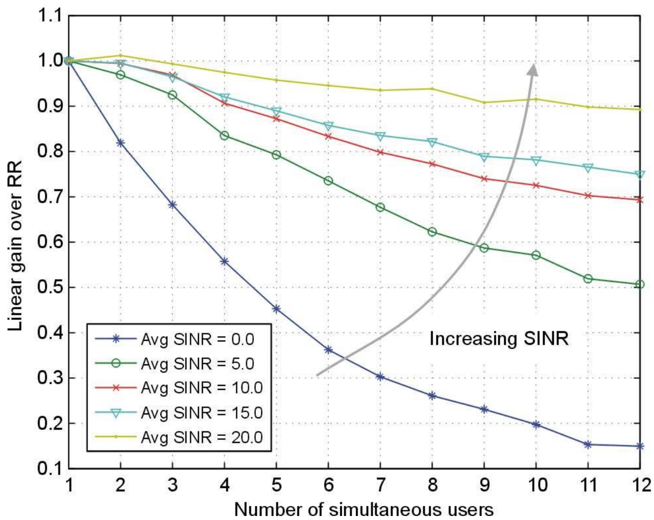

Figure 5 Scheduling gain of PF relative to RR in measured channels ($t_a = 1$ s). Similar to Fig. 4 results are shown for different average SINR values.

In Fig. 6 we have increased the time constant t_a from the default of 1 s. For very high values the performance of PF approaches that of Max TP since the longer averaging diminishes the influence of the denominator in (1).

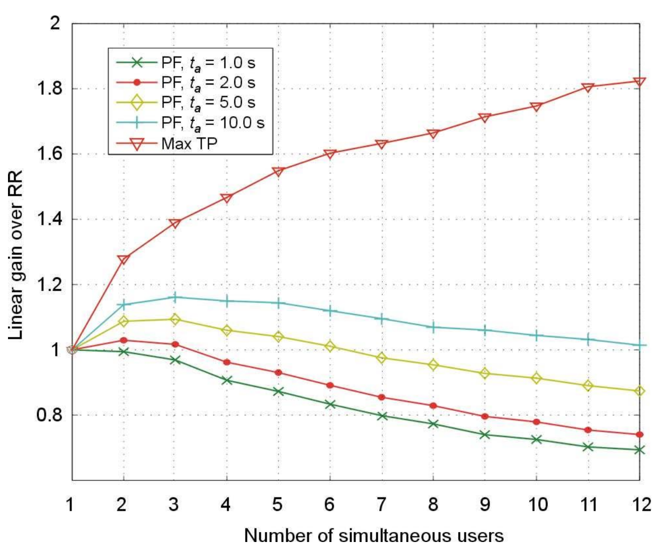

Figure 6 Scheduling gain of PF and Max TP relative to RR in measured channels (10 dB average SINR); results are shown for different time constants.

While Max TP approaches the gain achieved in the reference channel (2.2 at 12 users, cf. Fig. 4), PF has vanishing system throughput gain over RR in the measured channels, but is also considerably fairer in terms of user throughput. From the example in Fig. 7 showing accumulated throughput for one combination of 8 simultaneous users over the duration of the channel (100 s) it is clear that RR, and especially Max TP, give different throughput to users while PF provides almost the same to all users. Remember that all channels are at the same average SINR, hence in the reference channel conditions all three schedulers give equal throughput to all users.

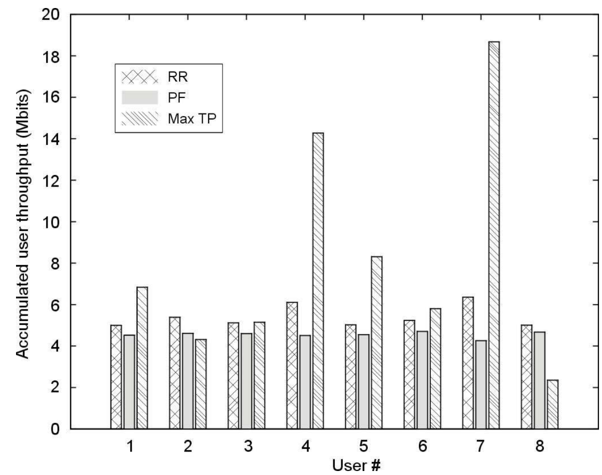

Figure 7 One example of throughput fairness between RR, PF and Max TP in the case of 8 simultaneous users.

B. Discussion

From the previous result the immediate conclusion is that the quite different performance of PF relative to RR in the measured channels can be explained by the difference in user fairness. While the RR scheduler is only fair in terms of resource usage, and hence "favourises" users in good average conditions (like Max TP) who can really make maximum utilisation of the resources, PF tries to compensate the throughput of users in temporarily poor average channel conditions: After some time the averaged throughput, T_k, of these users in (1) would become very low, hence increasing their priority in the PF RCQI despite the nominator is obviously also low in such a case. In fact, some of the measured channels have temporarily much worse average channel conditions than the examples shown in Fig. 1 and would cause RR to still provide throughput from other users in good average channel conditions while PF results in no or at least only low throughputs. Such behaviour could explain the improved performance with increasing average SINR operating point, and the deteriorating performance with increasing number of users since the probability of occasionally scheduling users in good average channel conditions increases with the number of users for RR. The results are not unexpected as such since from theory we know that the PF behaviour is influenced by the different fading characteristics [5] – the lower fading variability the lower gain. Clearly, the fading variability is different in the measured than the reference channels, but as it was mentioned in Section III.B the fast fading variability is to some extent much like in the Jakes model. What mostly make a difference is the average channel variation and in general the fading variability within and between the channels. Add to this the total dynamic range over the set of channels, which, with the example in Fig. 1, is in the order of 50 dB; in comparison, the link adaptation range in Fig. 3 is some 20 dB. Part of the dynamic range is covered by average channel behaviour as can be deduced from Fig. 2, though most channels stay within 8 dB based on the 10 and 90th percentiles. From examination of the data these are typically also the ones with limited fast fading dynamics.

To see if we could somehow improve the situation we changed the ARQ protocol to a simple model for HARQ in which every 2nd retransmission succeeded with probability 1, irrespective of the SINR. Also, we implemented a modification to the PF scheduler so that the RCQI was constrained by a limit on the BLEP: The scheduler prioritises the user with the highest RCQI satisfying a first transmission BLEP target lower than a specified threshold. Results for PF including these modifications are given in Fig. 8. If we compare to previous results we now see gain figures above 1 for low number of simultaneously scheduled users, and increasing with lower BLEP thresholds. We can conclude from this that adding robustness in the form of HARQ and a more conservative scheduler improves the situation, but is by itself insufficient to bring substantial gain to PF over RR.

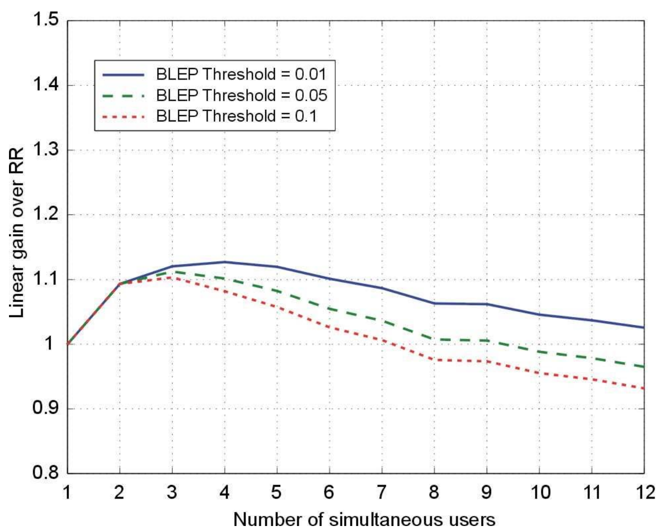

Figure 8 Scheduling gain of PF relative to RR in measured channels (10 dB average SINR, $t_a = 1$ s) when the scheduling metric (RCQI) is constrained by the BLEP.

As a final remark, using the reference channels we reproduced some of the behaviour seen with the measured channels. For example, Fig. 4 shows the effect when uncertainty is introduced to the channel quality at the UE. In this case, we added log-normally distributed *iid* noise samples to the SINR values given by the reference channels; the standard deviation of 1 dB corresponds approximately to the effect of the measurement noise mentioned at the end of Section III.B (0.54 dB standard deviation) if we assume that it applies independently for the channel quality measurement and the subsequent transmission ($\sqrt{2(0.54\,\text{dB})^2} = 0.76\,\text{dB}$). As we see in Fig. 4, the effect is to lower the point where the gain saturates and furthermore, reduce the achievable gain of PF over RR scheduling. Hence, this partly explains for the observed behaviour in the measured channels.

VI. Conclusions

The proportional fair scheduling algorithm is known to provide an appealing trade-off between system throughput and fairness between scheduled users. We have shown in this paper however, that when subjected to measured channels from an indoor scenario the PF algorithm is unable to gain in throughput over the simpler RR scheduling algorithm. The Max TP scheduler on the other hand shows approximately the same relative behaviour as in evaluation with reference channel models, here specifically flat Rayleigh fading. While our results do not have general application to WCDMA/HSDPA terminal usage, they are of some relevance for voice over IP applications.

We justified that the different performance of PF over RR in the measured channels is partly related to the uncertainty in channel quality information due to noise and, specifically, the more irregular fading variability within and between the measured channels in comparison to the fading in the reference channels; the measured channels show temporarily very bad average channel conditions. We speculate that a major cause for the difference observed in our evaluation is the limited dynamic range of the modelled HSDPA system in comparison to the measurements. This could indicate the importance of including more accurate HARQ modelling to extend the dynamic range downwards, adding support for more multi-codes to extend dynamic range upwards, and modifying the PF scheduler to pre-empt users in very bad average SINR conditions. This will be the topic of a future study.

References

[1] H. Holma and A. Toskala, WCDMA for UMTS, 3rd edition, Wiley, 2004.

[2] T. E. Kolding, K. I. Pedersen, J. Wigard, F. Frederiksen, and P. E. Mogensen, "High Speed Downlink Packet Access: WCDMA Evolution," IEEE Vehicular Technology Society News, February 2003, pp. 4-10.

[3] P. Viswanath, D. Tse, and R. Laroia , "Opportunistic beamforming using dumb antennas," IEEE Transactions on Information Theory, vol. 48, no. 6, June 2002.

[4] D. Tse, "Multiuser diversity in wireless networks," *Wireless Communication Seminar*, Stanford University, April 2001.

[5] J. M. Holtzman, "Asymptotic analysis of Proportional Fair algorithm," *IEEE Proc. Personal Indoor Mobile Radio Communications (PIMRC)*, September 2001, pp. 33-37.

[6] T. E. Kolding, "Link and system performance aspects of Proportional Fair scheduling in WCDMA/HSDPA," *Proceedings of 58th IEEE Vehicular Technology Conference (VTC)*, Florida USA, October 2003, pp. 1454-1458.

[7] "Guidelines for the evaluation of radio transmission technologies for IMT-2000," Recommendation ITU-R M.1225, 1997.

[8] T. B. Sørensen, "Intelligent Distributed Antenna Systems (IDAS); assessment by measurement and simulation," *Ph.D. Thesis*, Aalborg University, February 2002.

[9] P.C.F Eggers, C. Thommesen, J.B. Andersen, "On spatial power gradients", IEEE Transactions on Vehicular Technology, vol. 49, issue 1, January 2000, pp. 235 - 238