Copyright Notice

Institute of Communication Networks and Computer Engineering
University of Stuttgart
Pfaffenwaldring 47, D-70569 Stuttgart, Germany
Phone: ++49-711-685-68026, Fax: ++49-711-685-67983
Email: mail@ikr.uni-stuttgart.de, http://www.ikr.uni-stuttgart.de

IEEE ICC 2012 - Wireless Networks Symposium

Transmit mode selection schemes for distributed coordinated transmission of data traffic

Christian M. Mueller
University of Stuttgart, Institute of Communication Networks & Computer Engineering
Pfaffenwaldring 47, 70569 Stuttgart, Germany
Email:christian.mueller@ikr.uni-stuttgart.de

***Abstract*—Coordinated Multi-Point transmission schemes have a large potential to increase spectral efficiency in cellular networks. While the physical layer aspects of coordinated transmission received a lot of attention, its integration into higher layers yet remains to be investigated. In this work, we analyze the transmit mode selection for coordinated and uncoordinated transmission of data traffic. We propose two transmit mode selection schemes to decide whether to use coordinated transmission for an object or not. One scheme is based on object size, the other one is based on queue state. We derive the optimal parameter settings for our selective coordination schemes and show how they improve system performance.**

I. Introduction

Coordinated Multi-Point (CoMP) transmission schemes are a promising technique to increase spectral efficiency of cellular networks. 3GPP has put a focus on CoMP schemes in the development of LTE-Advanced [1] and distinguishes between *Joint Transmission*, *Joint Detection*, *Coordinated Beamforming* and *Coordinated Scheduling* (see [2], [3] for an overview). This publication is concerned with the latter two schemes in the downlink of a cellular network. Standardization currently favors a CoMP implementation in a single basestation with several remote radio heads, which requires costly investments in backhaul infrastructure. We instead focus on a distributed implementation where several macro basestations cooperate in a decentralized fashion. Most of research on CoMP dealt with physical layer aspects, whereas the integration into higher layers received little attention [3], [4]. This integration is important, given that only an appropriate higher layer functionality can ensure that applications actually benefit from higher net transmission rates at the physical layer. Part of this integration, namely the transmit mode selection for coordinated transmission of data traffic, is the topic of this paper.

In [5], [6], the authors showed the large potential of coordinated transmission, but indicated that coordination gains might turn into losses if data traffic is very bursty: With bursty traffic, a user's downlink queue at the basestation runs empty from time to time. It might happen that buffers run empty while coordination with neighbors is still ongoing. The constraints on resource allocation the basestations agreed upon might then not fit the new system state anymore and actually be harmful to the system's performance. This observation raises a number of questions: When shall we use the one or the other transmit mode? How and when, or for which kind of traffic, shall we trigger coordinated transmission? And, further on, how can we make bursty data traffic more suitable for coordinated transmission? This paper is concerned with the first two questions, leaving the latter question for future work.

As in [7], we consider two different ways to initiate a coordinated transmission. In [7], we analyzed spectral efficiency and transmission delay of both coordinated transmission modes by means of simulation. We did however not consider any transmit mode selection. In this work, we propose a parameter to decide whether to use coordinated transmission or not, depending on the size of the transferred object and depending on the queue state of a user's downlink buffer at the basestation. We derive the optimal parameter settings and show how selective coordination impacts spectral efficiency.

The remainder of this paper is organized as follows: Section II describes how coordinated transmission works and presents our abstract, high-level view on coordinated transmission. Section III gives a formal description of the transmission modes that we consider here. Section IV describes the remaining parts of our system model, i.e., the radio channel, the data traffic model and other wireless system aspects. It also presents the scenario parameters that we use for numerical evaluation throughout this work. Sections V and VI present our main contribution: Section V proposes a transmit mode selection scheme based on object size and presents results for a numerical example. Because the object size is usually not known before transmission, section VI proposes another transmit mode selection scheme which is based on queue state. In [8], we further extend our analysis by an evaluation of per-user throughput using selective coordination.

II. Coordinated transmission

In a cellular system with reuse one, i.e., all basestations use the same frequency resources, downlink transmissions of a basestation to its users cause interference to transmissions in neighbor cells. This inter-cell interference is the limiting factor to cellular network capacity in urban areas. An appropriate choice of time/frequency resources or antenna parameters of potentially interfering transmissions can reduce inter-cell interference. With *Coordinated Scheduling*, one tries to find orthogonal time/frequency resource allocations for otherwise interfering transmissions. With *Coordinated Beamforming*, one tries to steer the antenna beams in different directions, as depicted in the example in Figure 1. The challenge here lies in finding suitable resource allocations and transmission

978-1-4577-2053-6/12/$31.00 ©2012 IEEE

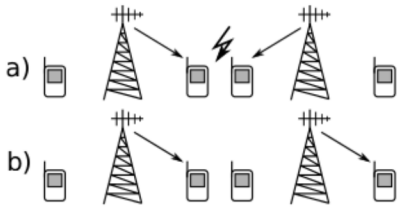

Figure 1: Coordinated Beamforming example: a) without, b) with coordination between basestations

parameters which reduce inter-cell interference, maximize system throughput or satisfy another objective function. For an overview of different objective functions, more general forms of interference coordination or management and many related physical layer aspects, see [3], [4].

To coordinate resources, basestations exchange signaling messages over a fast backhaul network. For the *Joint Transmission* CoMP scheme, basestations need to also exchange user data, which leads to even more demanding requirements on backhaul infrastructure. This exchange of signaling messages takes a certain time and makes resource allocation for distributed CoMP differ from resource allocation in other transmission modes, because resources have to be planned more far into the future. This requires channel prediction and estimation of the number of resources required to transmit an object. Another approach is to reserve a certain share of resources until the transmission of an object ends. The latter approach is beneficial if transmitted objects are large and if we want to avoid potentially negative impact from channel prediction errors. We here focus on the latter approach.

We model coordinated transmissions in an abstract way: The setup of a coordinated transmission consists of a measurement step and a coordination step. In the measurement step, user terminals perform channel measurements towards their serving cell and possibly also towards neighbor cells. In the coordination step, the basestations use these measurements to coordinate resource allocation for their users. We abstract from the coordination algorithm and characterize a coordinated transmission solely by the following three parameters:

Coordination setup time τ. Channel measurements and signaling for inter-basestation coordination take a certain time. This time is variable and depends on the channel measurement type, the propagation delay on backhaul links and the coordination algorithm. We denote this setup time as τ.

Coordination gain G. The coordination of the resource allocation yields a better SINR for coordinated transmissions. We model this gain by adding a constant offset G to a user's current SINR (in dB). Because of SINR clipping and the logarithmic Shannon equation, users with bad SINR benefit more from coordinated transmission than users with good SINR. This corresponds to a real system, where cell-edge users benefit more than cell-center users.

Coordination overhead factor η. The use of coordinated transmission comes at a certain cost. First, channel measurement results consume uplink resources. Second, signaling messages between basestations consume resources on backhaul links. Third, there is cost in terms of constraints imposed on resource allocation in serving and neighbor cells. These constraints are the outcome of the coordination algorithm and normally increase system throughput. However, in some cases, the system state changes too quickly and the coordinated resources cannot be used as intended. This happens for example if a user's buffer at the basestation ran empty while coordination among the basestations was ongoing.

Because it is unclear how to weight backhaul or uplink resources against downlink resources, we neglect uplink and backhaul overhead and measure the overhead in terms of downlink resources only. We assume that a basestation coordinates resources with its neighbors as long as the user's buffer is not empty. When it runs empty, no more resources are allocated to this user and the coordination stops. Due to non-zero backhaul delay, there is a certain time between the instant the buffer runs idle (and hence the end of the transmission) and the release of all constraints on resource allocation. We assume this time is the same time τ that is required to initiate a coordinated transmission. During this time, resource allocation in serving and neighbor cells is still constrained, which might lead to a loss in spectral efficiency compared to an unconstrained resource allocation. This requires to trade-off potential gains from coordinated transmission against potential losses caused by constraints imposed on resource allocation. For large objects the gain outweighs the overhead. This might not be true for small objects.

To what extent periods with constrained resource allocation degrade spectral efficiency depends on many factors. If there are many users, there might be no degradation at all because the basestations are able to reuse the constrained resources for other users. If there are few users and the constraints, e.g., largely restrict the number of allowed frequencies or antenna precodings, the constrained resources might not be used at all because the basestation cannot serve any user with the remaining set of resources. To model the variable influence of these constraints, we introduce the overhead factor $\eta \in [0; 1]$. An overhead factor $\eta = 0$ means that all resources can be fully reused and periods with constrained resource allocation have no negative impact. An overhead factor $\eta = 1$ means that all constrained resources cannot be used at all.

III. Transmission modes

We distinguish different transmission modes and characterize a transmission by the following two characteristic functions, where s is the size of the transmitted object and c is the channel quality expressed in terms of the long-term signal-to-interference-and-noise-ratio (SINR) in dB:

$\chi(s, c)^{(mode)}$ Number of resource units (in [Hz s]) to transfer an object of size s at SINR c.

$\omega(s, c)^{(mode)}$ Overhead in number of resource units that is caused in serving or neighbor cells.

A. Reference transmission modes

We consider two reference transmission modes. The overhead $\omega(s, c)$ is zero for both reference modes. Our baseline

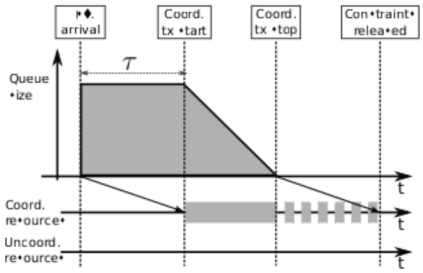

Figure 2: *Delayed* transmit mode

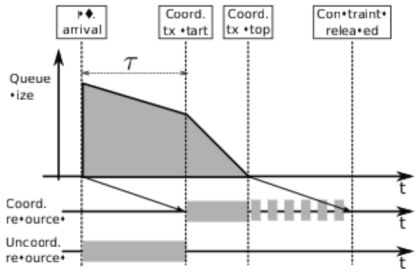

Figure 3: *Immediate* transmit mode

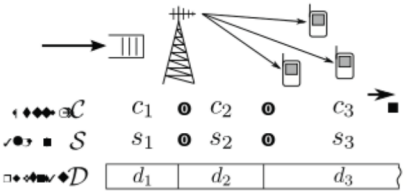

Figure 4: System model

is the *uncoordinated* transmission, which could be a transmit diversity scheme and does not require any coordination with neighbor basestations. The number of resource units required to transfer an object of size s at SINR c is:

$$\chi^{(\text{uncoord})}(s,c) = s/\gamma(c) \tag{1}$$

The function $\gamma(c)$ yields the spectral efficiency at SINR c. For simplification, we assume that the SINR is constant over the whole duration of the object transfer. We further treat interference as white noise and use the Shannon equation to determine the spectral efficiency, i.e. $\gamma(c) = \log_2(1 + c_{\text{linear}})$.

The *perfectly coordinated* transmission denotes the upper bound on spectral efficiency where coordination among basestations is ideal and happens in zero time:

$$\chi^{(\text{perfect})}(s,c) = s/\gamma(c+G) \tag{2}$$

B. Delayed transmission with precedent coordination

In the so-called *delayed* transmission mode, we assume the basestation buffers incoming data until the coordination with its neighbors is complete (Figure 2). It then starts the object transmission on the coordinated resources. Once the user's buffer runs empty, the basestation notifies its neighbors and all constraints on resource allocation are released. The number of resource units required for transmission and the overhead caused on resource allocation after the buffer ran empty are

$$\begin{aligned}\chi^{(\text{del})}(s,c) &= s/\gamma(c+G)\\ \omega^{(\text{del})}(s,c) &= \eta r\tau\end{aligned} \tag{3}$$

The symbol r is the (constant) bandwidth r a basestation allocates to a user. This bandwidth can also be interpreted as a rate of resource units or symbols.

C. Immediate transmission with concurrent coordination

In the so-called *immediate* transmission mode, we assume the basestation starts transmitting every object in uncoordinated mode. While the transmission is ongoing, the basestation coordinates resource allocation with its neighbors. Once coordination is complete (i.e. τ seconds after the object's arrival), the basestation continues object transmission on the coordinated resources. The characteristic functions then are

$$\begin{aligned}\chi^{(\text{im})}(s,c) &= \begin{cases} s/\gamma(c) & \text{for } s \le r\tau\gamma(c) \\ r\tau + \frac{s - r\tau\gamma(c)}{\gamma(c+G)} & \text{otherwise} \end{cases}\\ \omega^{(\text{im})}(s,c) &= \begin{cases} \eta s/\gamma(c) & \text{for } s \le r\tau\gamma(c) \\ \eta r\tau & \text{otherwise} \end{cases}\end{aligned} \tag{4}$$

IV. Evaluation Methodology

We use an analytical model for our evaluations. Albeit a simple model, it provides fundamental insight in system behavior and permits evaluations using heavy-tailed object size distributions, which otherwise quickly lead to prohibitively long simulation times when using them in system-level simulations. Similar models have been used in literature before, but we are not aware of previous work that applied such a model to the analysis of coordinated transmissions.

A. Wireless system model

We consider a single basestation which sequentially transfers objects to its users (Figure 4). The basestation serves only one user at a time. After the basestation has completed the transfer of an object to one user, it immediately starts the transfer of a new object to a random user. We assume that system bandwidth r is constant. The basestation always allocates the whole bandwidth. Please note that this model describes a wireless network with rate-fair resource allocation.

The size of an object s is random, with probability density function (pdf) $f_{\mathsf{S}}(x)$. The channel quality c is the long-term SINR of a user, which is assumed constant over the whole duration of an object transfer. We treat interference as noise and use the Shannon equation for the mapping of SINR to spectral efficiency. We assume an infinite number of users and model the channel quality as a random variable with pdf $f_{\mathsf{C}}(x)$.

The most important metric in our analysis is the spectral efficiency. Using the characteristic functions introduced in section III, the spectral efficiency ϱ of an object transfer is

$$\varrho(s,c) = s/\left(\chi^{(\text{mode})}(s,c) + \omega^{(\text{mode})}(s,c)\right) \tag{5}$$

By the law of the unconscious statistican, we get the expected value of the spectral efficiency as [9]:

$$\mathrm{E}[\varrho] = \int\int \varrho(s,c) f_{\mathsf{SC}}(s,c)\,ds\,dc \tag{6}$$

where $f_{\mathsf{SC}}(s,c)$ is the joint pdf of s and c. Because the random variables s and c are independent of each other, $\mathrm{E}[\varrho]$ becomes

$$\mathrm{E}[\varrho] = \int\int \varrho(s,c) f_{\mathsf{S}}(s) f_{\mathsf{C}}(c)\,ds\,dc \tag{7}$$

Equation (7) gives the average spectral efficiency of an object transfer. If we are however interested in the time-average of the spectral efficiency of our system, we have to derive an expression for the spectral efficiency at an arbitrary

instant t_0. If we pick an arbitrary t_0, we have a higher chance of observing object transfers with a long duration, e.g., the transfer of a large object at low SINR. This bias with the transfer duration is a result from renewal theory, known as the waiting time paradox [10]. Our system model constitutes a renewal process, with the recurrent events being the time instances at which the transmission of an object ends. The duration an object transfer occupies the wireless channel is $d(s,c) = \chi(s,c)/r$. The probability that an arbitrary instant t_0 falls in an object transfer of duration d is [10]:

$$f_{\mathsf{D}[\mathsf{t}]}(d) = K \cdot d \cdot f_{\mathsf{D}}(d) \tag{8}$$

where $K = 1/\mathrm{E}[d(s,c)]$, and pdf $f_{\mathsf{D}}(d)$ is the unbiased pdf of d. In [11], the same principle was used to calculate a biased SINR distribution to evaluate the performance of an IMT-Advanced system. In the following, we use the notation $[t]$ to distinguish between biased and unbiased values.

By application of equation (8) and the law of the unconscious statistician, we get the expected value of $d[t]$:

$$\mathrm{E}[d[t]] = \iint d(s,c) f_{\mathsf{S}[\mathsf{t}]\mathsf{C}[\mathsf{t}]}(s,c) ds\, dc \tag{9}$$

$$= \iint d(s,c) K d(s,c) f_{\mathsf{S}}(s) f_{\mathsf{C}}(c) ds\, dc \tag{10}$$

In the same way we get the following expression for the time-average of the spectral efficiency of our system:

$$\mathrm{E}[\varrho[t]] = \iint \varrho(s,c) f_{\mathsf{S}[\mathsf{t}]\mathsf{C}[\mathsf{t}]}(s,c) ds\, dc \tag{11}$$

$$= K \iint \varrho(s,c) d(s,c) f_{\mathsf{S}}(s) f_{\mathsf{C}}(c) ds\, dc \tag{12}$$

B. Data traffic model

Most applications on mobile devices use HTTP and HTTPS for data transfer [12], [13]. This comprises web surfing using a browser, so-called smartphone apps, the AppStore or Android Market and even mobile video applications. The most popular mobile video site which causes a significant share of mobile video traffic is YouTube [13]. YouTube servers use a rate-limiting and block-sending scheme at the application layer, which usually transmits videos in 64 kByte blocks [14]. In total, HTTP and HTTPS based applications account for more than 80% of the traffic volume on mobile devices [12], [13]. The distribution of HTTP object sizes is heavy-tailed and only slightly differs between fixed and wireless networks [13], [15].

We restrict our traffic model to objects sent over HTTP/S. Although the impact of the arrival process certainly is of interest, we limit ourselves to the transmission of individual objects and assume there is always an object available which has to be transmitted. We do not take a fragmented transmission of the objects into account and hence ignore the round-based transmission behavior of TCP. The results presented here can thus be regarded as the "good case" where a basestation receives the application layer data unit as a whole.

C. Scenario parameters and model validation

For our numerical results in sections V and VI, we use the following assumptions and parameters: We use an empirical distribution of SINR c, gathered from wideband SINR measurements in a system-level simulation of a 3GPP Case 1 scenario with 3D antenna patterns, 19 sites, pathloss and shadowing (no fast fading) according to [1]. The cdf of our SINR shows a good fit with Fig. A.2.2-2 in the system simulator calibration section of [1]. Because we use the Shannon equation instead of LTE transport formats, we clipped the SINR at 22 dB. To model the SINR improvement of coordinated transmissions using a CS/CP CoMP scheme, we add a constant offset of $G = 6\,\mathrm{dB}$ to a user's current SINR. This offset roughly corresponds to a cancellation of the first up to second most significant interferer in our 3GPP Case 1 3D scenario. For the coordination setup time τ, we consider $\tau = 10\,\mathrm{ms}$ as our minimum value. 3GPP documents specify an average backhaul delay of 10 ms between LTE macro cells as a realistic value [16]. For the object size distribution, we use the HTTP response size distribution published in [15], which provides a representation of the empirical object size distribution function as a mixture of three log-normal distribution functions (see Figure 10). We set the range of valid object sizes to 10 up to 10^9 bytes.

For validation of our model, we consider the special case of uncoordinated transmission. In this case, we use equations (5) and (1) in equation (12) and simplify the expression to get the expected value of the spectral efficiency:

$$\mathrm{E}[\varrho[t]^{(uncoord)}] = \mathrm{E}[\tfrac{1}{\mathsf{C}}]^{\,1} \leq \mathrm{E}[\gamma(c)] \tag{13}$$

This result is the average spectral efficiency of a system with rate-fair resource allocation. We can relate this result to the spectral efficiency of a system with resource-fair resource allocation by Jensen's inequality (see right hand side of equation (13)). In our setup, we get $\mathrm{E}[\varrho[t]^{(uncoord)}] \approx 2\,\mathrm{bps/Hz}$ for the rate-fair resource allocation and $\mathrm{E}[\gamma(c)] \approx 3\,\mathrm{bps/Hz}$ for a resource-fair resource allocation. Because the result is independent of the object size distribution, we can easily verify it with system simulation. Please note that these values are larger than the reference results for the 3GPP Case 1 3D scenario because of the Shannon capacity and because we do not take any control and signaling overhead into account. We also compared selected results of our numerical analysis with results from system-level simulation with truncated object size distributions, which also showed a good match.

V. Selective Coordination Based on Object Size

To illustrate the benefit of selective coordination we first investigate the spectral efficiency of the *immediate* and *delayed* transmit modes for different object sizes. We then derive the *selective coordination thresholds*, first user-individual thresholds and then a system-wide coordination threshold.

A. Spectral efficiency of immediate and delayed transmit mode

Figure 5 and Figure 6 depict the spectral efficiency of a single object transfer (using equation (5)) for the *immediate*

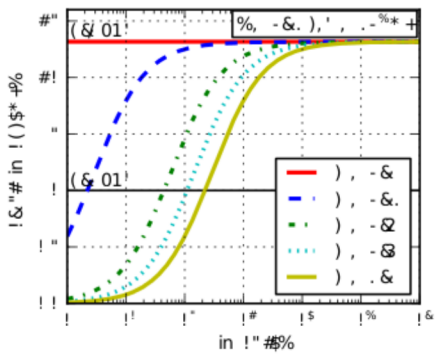

Figure 5: *Delayed* transmit mode

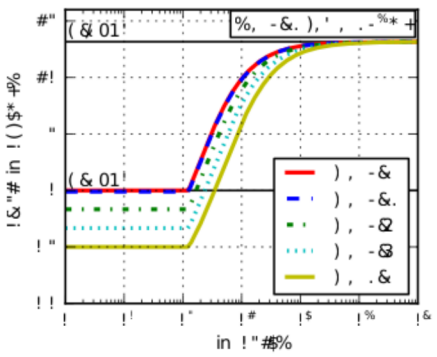

Figure 6: *Immediate* transmit mode

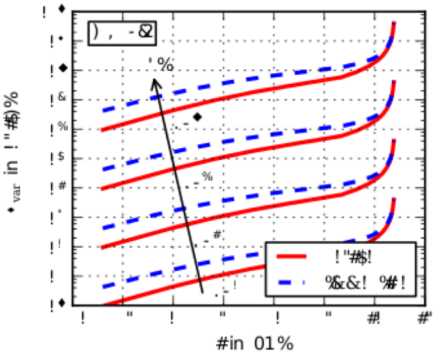

Figure 7: Burst size threshold S_{var}

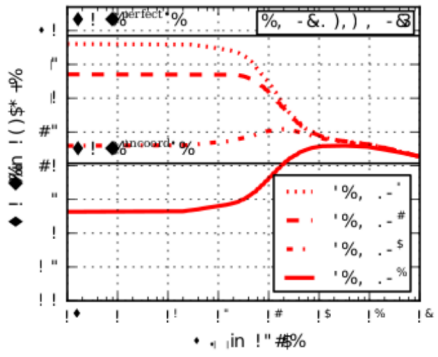

Figure 8: *Delayed* mode with S

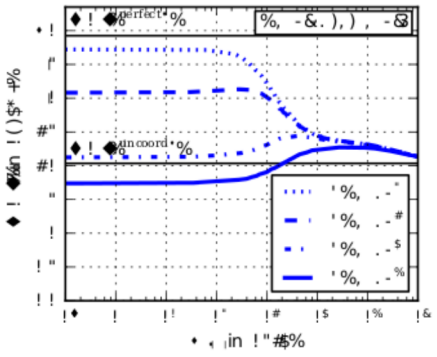

Figure 9: *Immediate* mode with S

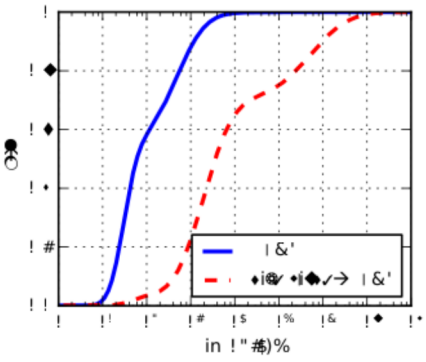

Figure 10: Object size

and *delayed* transmission modes at SINR $c = 0\,\text{dB}$, setup delay $\tau = 10\,\text{ms}$, bandwidth $r = 10\,\text{MHz}$ and variable overhead factor η. Both modes show a similar behavior: the larger the object size, the better the spectral efficiency. The *delayed* transmission mode is more efficient except for small object sizes up to a few KBytes. This is because the lag after the buffer runs empty in *delayed* mode always equals $\eta r\tau$. If the buffer runs empty before coordinated transmission starts, the lag in *immediate* mode is smaller. For larger objects, *immediate* mode is less efficient because of the uncoordinated transmission period at the beginning of an object transfer. The higher efficiency of the *delayed* mode comes at the cost of increased delay: all traffic has to wait for τ seconds before transmission starts. It thus trades throughput against delay. For a transfer delay analysis and its impact on TCP see [7].

B. User-individual object-size based coordination thresholds

In Figure 5 and Figure 6 we can observe that the spectral efficiency for both modes is above the one of uncoordinated transmission at $\gamma(0\,\text{dB})$, if the object's size is larger than a certain threshold value. We thus propose to use uncoordinated transmission for objects smaller than this threshold and coordinated transmission for larger objects.

The threshold object size S_{var} varies depending on transmission mode, channel quality, coordination gain and overhead factor. From Eq. (3) and Eq. (4), we can derive S_{var} for both transmission modes by setting $\chi^{(\text{mode})} + \omega^{(\text{mode})} < \chi^{(\text{uncoord})}$ and solving for s:

$$S_{\text{var}}^{(\text{del})} = \eta r\tau \left(\frac{1}{(c)} - \frac{1}{(c+G)} \right)^{-1} \tag{14}$$

$$S_{\text{var}}^{(\text{im})} = r\tau \left(\eta + 1 - \frac{(c)}{(c+G)} \right) \left(\frac{1}{(c)} - \frac{1}{(c+G)} \right)^{-1} \tag{15}$$

There is a linear relationship between S_{var} and the product of resource unit rate and setup time, $r\tau$, whereas the dependency on c and G is non-linear. The influence of η is linear for *delayed* mode, but only approximately linear for *immediate* mode. Figure 7 plots S_{var} over the relevant parameter range.

C. System-wide object-size based coordination threshold

Instead of a user-individual threshold S_{var}, we can also use a single, system-wide threshold S The advantage of a single threshold lies in an easier and more robust implementation: A system-wide threshold can be based on long-term statistics of the SINR and does not require per-user channel estimates. It can also be used as initial value as long as no good channel estimates for the current user are available.

Figure 8 and Figure 9 depict the average spectral efficiency $\mathrm{E}[\varrho[t]]$ (using eq. (12)) over a constant, system-wide coordination threshold S The figures also show the spectral efficiency of the reference transmit modes. We can observe that the higher the product of setup time and bandwidth, $r\tau$, the smaller the spectral efficiency. We can also observe that $\mathrm{E}[\varrho[t]]$ has a local maximum for a certain coordination threshold S which becomes even more pronounced the higher the rate r.

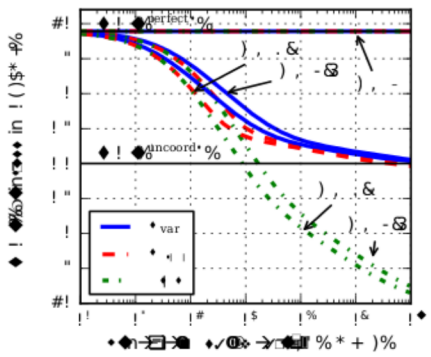

Figure 11: *Delayed* mode, S vs. S_{var}

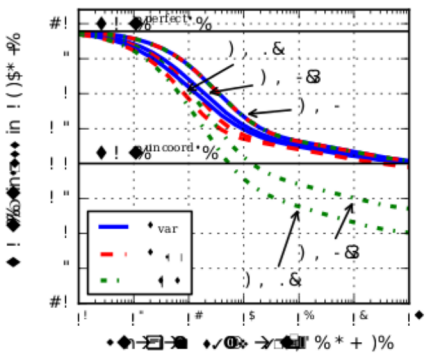

Figure 12: *Immediate* mode, S vs. S_{var}

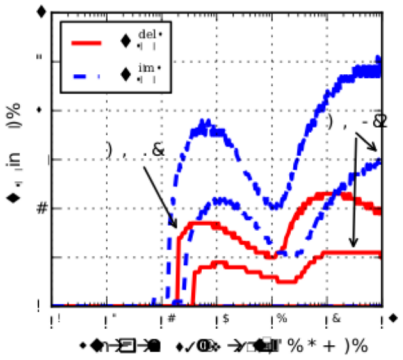

Figure 13: Optimal queue-state threshold

To determine this optimal threshold value, Eq. (12) has to be solved for its local maximum. This derivation is lengthy and a closed-form solution is not always possible. We hence only present results for numerically determined S values here.

D. Evaluation with object-size based coordination thresholds

Figure 11 and Figure 12 compare the spectral efficiency when using user-individual coordination thresholds to a system-wide coordination threshold. They depict the gain/loss of average spectral efficiency compared to an uncoordinated reference system for (i) user-individual coordination thresholds (S_{var}, blue curve), (ii) for a system-wide coordination threshold (S dashed red curve) and (iii) without any threshold, i. e., if coordinated transmission was used for all object transfers (dotted green line). We depict the gains over the product $r\tau$ and cover a wide range of values. The lower end could be a system with a small setup time $\tau = 10\,\text{ms}$ and a user allocated bandwith of 10 kHz ($r\tau = 10^2\,\text{Hz s}$). The upper end could be a system with a large setup time $\tau = 1\,\text{s}$ and a large user allocated bandwith of 100 MHz ($r\tau = 10^{\blacklozenge}\,\text{Hz s}$).

The amount of data that can be sent in τ seconds increases with bandwidth r. Consequently, the coordination overhead gets larger in relation to the objects' sizes. This is why the threshold object size above which coordinated transmission provides gains in spectral efficiency also increases with $r\tau$ (compare Figure 7). For large $r\tau$, S_{var} attains values of several 100 KBytes to few MBytes. Such a threshold would be larger than typical web sites and larger than the block size of YouTube transfers (see section IV-B). Most of mobile video traffic would thus not benefit from coordinated transmission. If we compare Figure 10, we see that coordinated transmission would only be used for less than 5% of the transferred objects, respectively 30% of the traffic volume.

At a typical point of operation with setup time $\tau = 10\,\text{ms}$ and user allocated bandwidth of 1 MHz ($r\tau \approx 10^{\}$), CoMP-enabled systems perform better than uncoordinated systems. If however τ is in the range of 100 ms ($r\tau \approx 10^{\bullet}$), coordination gains might have largely vanished or even turned into losses. Selective coordination ensures that spectral efficiency is always better or equal to an uncoordinated system. The spectral efficiency gap between an appropriately chosen threshold and no threshold at all is more pronounced for *delayed* transmit mode. Selective coordination is thus even more important in the *delayed* transmit mode compared to the *immediate* transmit mode except for very small overhead factors. For $\eta = 0$, the *delayed* mode is preferable over *immediate* mode, especially at large $r\tau$ values. The uncoordinated transmission at the beginning of an object transfer in *immediate* mode significantly lowers spectral efficiency for large $r\tau$ values.

Finally, Figure 11 and Figure 12 show that average spectral efficiency when using a system-wide threshold is only slightly worse compared to variable thresholds. If imperfect channel estimates were used for the computation of S_{var}, the gain from variable thresholds decreases further. We can thus conclude that a single, system-wide threshold S is sufficient.

VI. Selective Coordination Based on Queue State

Assuming the object size is known upon arrival of the first byte of an object is quite optimistic. We therefore make another more pessimistic assumption where the size of the object is not known, but coordinated transmission is triggered after the per-user buffer at the basestation did not run empty for a specified time T. Please note that although the size of an object is not known at the beginning of a transfer, a real basestation is not completely oblivious to the object size. Supposing that the wireless link is the bottleneck of the transfer, data will accumulate in the basestation's buffer and allow for an object size estimate.

A. System-wide queue-state based coordination threshold

In the *busy-time dependent delayed* mode, the basestation triggers coordination with neighbor basestations if the buffer did not run empty for time T. It then suspends the ongoing transmission until coordination of resource allocation is complete. In the *busy-time dependent immediate* mode, transmission starts on uncoordinated resources. If the buffer did not run empty for time T, coordination of resource allocation with neighbor basestations is initiated. After time τ, the basestation uses the coordinated resources to continue transmission.

The derivation of the optimal system-wide threshold T requires solving an integral over the object size distribution function. As for S exact solution is not always possible and we only present results for a numerically determined T

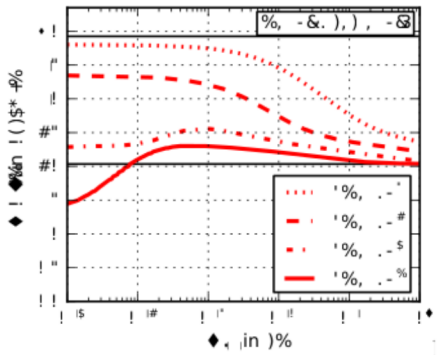

Figure 14: *Delayed* mode with T

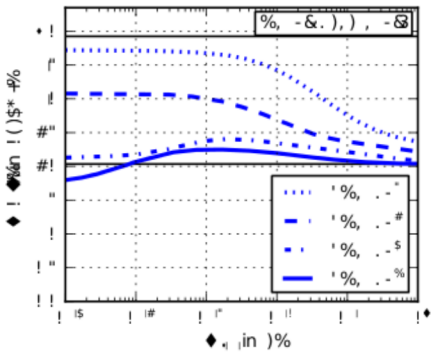

Figure 15: *Immediate* mode with T

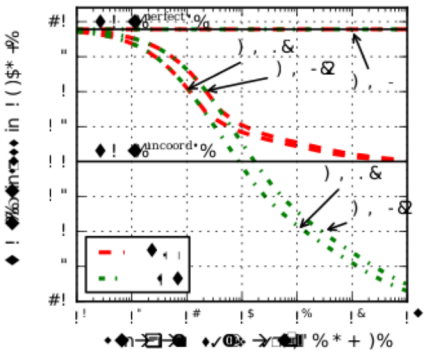

Figure 16: *Delayed* mode with T

B. Evaluation with queue-state based coordination thresholds

Figure 14 and Figure 15 depict the spectral efficiency over T with the same parameter settings as Figure 8 and Figure 9. The local maximum of $\mathrm{E}[\varrho[t]]$ also exists. Figure 13 shows the corresponding queue busy time thresholds that achieve this maximum spectral efficiency. As long as $r\tau$ is small, $T = 0$ achieves maximum spectral efficiency. For larger $r\tau$, optimal T values are in the range of 1 to 4 ms. In an LTE system, discretization of T to millisecond granularity to match the LTE transmission time interval length will not allow choosing the optimal waiting time parameter and thus spectral efficiency will be slightly lower. For *delayed* mode, Figure 16 shows the gain/loss in average spectral efficiency when using T compared to the uncoordinated system. If we compare with Figure 11, we can see that the same performance can be achieved with queue-state based thresholds as compared to object-size based thresholds. The same is true for *immediate* mode (the plot is left out for lack of space).

VII. Conclusion

We proposed and analyzed different approaches for selective coordinated transmission. Our results show that selective coordination is important, especially for large backhaul delays and large bandwidths. The appropriate choice of the selective coordination threshold can actually make the difference between spectral efficiency gains or losses of a coordinated system compared to an uncoordinated systems. We have presented results for a single, system-wide coordination threshold and for per-user thresholds, for which we derived the optimal threshold values. We used a novel lightweight analytical model which allows readers to easily verify our results and extend it to other parameters or specific coordination algorithms.

Our results show that coordination gains achieved at the physical layer cannot always be preserved at higher layers, in particular if setup time and bandwidth are large. If further the coordination overhead from constraints imposed on resource allocation is non-negligible, the presented *delayed* and *immediate* transmit modes do not perform well and other schemes trying to minimize this overhead might perform better. An investigation of such schemes is part of our future work.

Another important aspect is the influence of TCP. On the one hand, the round-based transmission behavior of TCP splits an object transfer into multiple smaller transfers, which might have a negative impact on spectral efficiency. On the other hand, often many TCP connections are active at the same time and the interleaving of bursts from these connections might cause the user's buffer to run empty less frequently, which might have a positive impact on spectral efficiency.

References

[1] 3GPP, *Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA Physical layer aspects*, 2010, no. TR36.814.

[2] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, and M. Tanno, "Coordinated multipoint transmission/reception techniques for lte-advanced," *Wireless Communications, IEEE*, vol. 17, no. 3, 2010.

[3] P. Marsch and G. P. Fettweis, Eds., *Coordinated Multi-Point in Mobile Communications: From Theory to Practice*. Cambridge University Press, July 2011.

[4] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Communications Magazine*, vol. 49, no. 2, Feb. 2011.

[5] M. Necker, "A novel algorithm for distributed dynamic interference coordination in cellular ofdma networks," Diss., Inst. of Comm. Networks and Computer Engineering, Universität Stuttgart, 2009.

[6] R. Madan, A. Sampath, N. Bhushan, A. Khandekar, J. Borran, and T. Ji, "Impact of coordination delay on distributed scheduling in lte-a femtocell networks," in *IEEE GLOBECOM*, 2010.

[7] C. Mueller, "Analysis of interactions between internet data traffic characteristics and coordinated multipoint transmission schemes," in *IEEE WCNC*, Cancun, Mexico, 2011.

[8] C. Müller, "When CoMP is beneficial - and when it is not. Selective coordination from a spectral efficiency and a users' throughput perspective," in *IEEE Wireless Communications and Networking Conference (WCNC 2012)*, April 2012.

[9] R. Goodman, *Introduction to stochastic models*. Benjamin/Cummings Pub. Co., 1988.

[10] W. Feller, *An Introduction to Probability Theory and Its Applications, Vol. 2*, 2nd ed. John Wiley & Sons, Ltd, 1971.

[11] K. Sambale and B. Walke, "Upper bound cell spectral efficiency of imt-advanced scenarios," in *17th European Wireless Conference*, 2011.

[12] H. Falaki, D. Lymberopoulos, R. Mahajan, S. Kandula, and D. Estrin, "A first look at traffic on smartphones," in *IMC*, 2010.

[13] G. Maier, F. Schneider, and A. Feldmann, "A first look at mobile hand-held device traffic," in *PAM*, April 2010.

[14] S. Alcock and R. Nelson, "Application flow control in youtube video streams," *SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 2, 2011.

[15] F. Hernandez-Campos, J. Marron, G. Samorodnitsky, and F. Smith, "Variable heavy tails in internet traffic," *Performance Evaluation*, vol. 58, no. 2-3, 2004.

[16] 3GPP R3-070702, "Reply LS to R3-070527/R1-071242 on Backhaul (X2 interface) Delay," 3GPP TSG-RAN-WG3 meeting 55bis, 2007.