Abstract—We consider synchronization techniques required to enhance the cellular network capacity using base station cooperation. In the physical layer, local oscillators are disciplined by the global positioning system (GPS) and over the backbone network for outdoor and indoor base stations, respectively. In the medium access control (MAC) layer, the data flow can be synchronized by two approaches. The first approach uses so-called time stamps. The data flow through the user plane and through copies of it in each cooperative base station is synchronized using a timing protocol on the interconnects between the base stations. The second approach adds mapping information to the data after the user plane processing is almost finalized. Each forward-error encoded transport block, its modulation and coding scheme and the resources where it will be transmitted are multicast over the interconnect network. Interconnect latency is reduced below 1 ms to enable coherent interference reduction for mobile radio channels.

I. INTRODUCTION

It has been realized in the early research on multiple-input multiple-output (MIMO) radio systems that the MIMO concept can be extended to remove the crucial interference in future cellular networks, see [1] [2] [3]. Interference reduction is reached using a joint beam-forming of the signals transmitted by several cooperative base stations. Fundamental effects as the channel rank advantage and the enhanced capacity have been demonstrated by simulations [4] and in a multi-cell measurement campaign [5]. For an overview of the cooperative base station concept we refer to [6] and references therein.

Point-to-point MIMO technology has been standardized, e.g. in indoor wireless LANs as in IEEE 802.11n and in cellular systems, e.g. 3G Long Term Evolution (3G-LTE). A cellular trial system with 2 transmit and 2 receive antennas in 20 MHz bandwidth has been implemented [7] and tested [8]. Although the basic MIMO technology is quite similar, the numbers of in- and outputs for cooperative base stations and hence the complexity are further increased.

Joint beam-forming can be implemented as a matrix-vector product in hardware and scaled easily to larger numbers of antennas and higher bandwidths [9]. Note that the computation of beam-forming weights scales cubically with the number of antennas. An implementation on a multi-core processor\(^1\) recently demonstrated that the beam-forming matrices for a cellular 12x12 MIMO-OFDM system can be computed in real-time [10]. If in each cell two antennas are used, e.g. to exploit cross-polarization multiplexing, the desired signals in the cell can be maximized and the signals from the five most interfering cells can be suppressed.

As this new technology is getting mature, this paper is concerned with the synchronization of cooperative base stations. Using a simple model, we show that frequency offset requirements are within the short-term phase drift of the primary reference clocks in commercial base stations. For outdoor base stations, global positioning system (GPS) synchronization can hence be reused from terrestrial digital video broadcasting (DVB-T) to compensate the frequency offset. For indoor base stations, a precise network synchronization protocol is required. An appropriate protocol based on the packet delay time is specified in the IEEE 1588 standard.

Distributed implementation is a fundamental requirement of current cellular radio design. In recent work a central unit is widely used for the joint beam-forming. If the data signals are made synchronously available at each base station, the beam-forming can be implemented in a decentralized manner. Here, we show how the medium access control (MAC) layers of distant base stations can be synchronized to enable distributed beam-forming.

The paper is organized as follows. In Section II, we consider synchronization requirements for cooperative base stations. Section III shows experimental results for GPS synchronization and how it may be combined with network synchronization. Two concepts for MAC synchronization are described in section IV.

II. SYNCHRONIZATION REQUIREMENTS

A. System model for MIMO-OFDM downlink

The downlink multiple-input multiple-output (MIMO)-orthogonal frequency division multiplexing (OFDM) transmission via \(N_T\) transmit and \(N_R\) receive antennas for each subcarrier is described by

\[
y = HCx + n ,
\]

where \(H\) is the \(N_R \times N_T\) channel matrix and \(C\) the unitary \(N_T \times N_R\) codebook matrix; \(x\) denotes the \(N_T \times 1\) vector of transmit symbols; \(y\) and \(n\) denote the \(N_R \times 1\) vectors of the received signals and of the additive white Gaussian noise (AWGN) samples.

\(^1\)The IBM cell processor is also used in the third generation of SONY’s low-cost play stations.
B. Impact of frequency offset between two base stations

The carrier frequency offset between two commercial base stations is much smaller than the sub-carrier spacing. We may ignore the inter-carrier interference and use a simple flat-fading model to evaluate the impact of synchronization errors. As an example, we consider joint zero forcing (ZF) transmission from two single-antenna base stations (BSs) to two single-antenna terminals.

We assume that both base stations are not perfectly synchronized. The frequency offset results in independent time-continuous phase shifts imposed on the effective channel coefficients seen between the antennas of BSs and terminals. Let \( \Phi(t) \) denote the phase rotation matrix which contains the individual phase factors on its diagonal with \( \varphi_1 \) and \( \varphi_2 \) being the carrier frequencies of both base stations and

\[
\Phi(t) = \text{diag}(\exp(j\varphi_1 t), \exp(j\varphi_2 t)),
\]

the transmission equation then modifies to

\[
y = H\Phi(t)Cx + n
\]

In the following we will neglect the noise for simplicity. We specify

\[
H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \begin{bmatrix} h_1^T \\ h_2^T \end{bmatrix}
\]

where \( h_i^T \) are the row vectors of matrix \( H \). Assume that both BS obtain ideal channel knowledge at time instant \( t = 0 \), then they form the ZF beam-forming matrix according to

\[
C = H^{-1} = \frac{\epsilon}{h_{11}h_{22} - h_{12}h_{21}} \begin{bmatrix} h_{22} - h_{12} & -h_{12} \\ h_{21} & h_{11} \end{bmatrix} \quad \text{[c_1,c_2]}
\]

where \( \epsilon \) is a constant factor that results from the transmit power constraint and \( c_i \) are the column vectors of matrix \( C \). Now, terminal 1 receives the signal

\[
y_1 = h_1^T \Phi(t)Cx + n
\]

Note that \( h_1^T Cx = \epsilon x_1 + 0 \cdot x_2 \). Thus, the received signal falls into two parts, \( y_1 = y_u + y_i \), namely the useful signal component \( y_u \) and the interference component \( y_i \), which are given as

\[
y_u = \exp(j\varphi_1 t) (\epsilon - (1 - \exp(j\theta t))h_{12}^*c_{12}) x_1 \quad \text{(7)}
\]

\[
y_i = \exp(j\varphi_2 t) (0 - (1 - \exp(j\theta t))h_{12}^*c_{22}) x_2 \quad \text{(8)}
\]

with \( \theta = \varphi_2 - \varphi_1 \). With these components, we can determine the mean signal to interference ratio (SIR) by calculating

\[
\text{SIR} = \left| \frac{y_u}{y_i} \right|^2 = \frac{|(h_{11}h_{22} - h_{12}h_{21})\epsilon + (1 - \exp(j\theta t))|h_{12}|^2|^2}{(1 - \exp(j\theta t))|h_{12}^*c_{22}|^2}
\]

where we assumed that both signals \( x_i \) have identical power. With Jensen’s inequality, we can lower bound the mean SIR according to

\[
\text{SIR} = \frac{E\{|y_u|^2\}}{E\{|y_i|^2\}} \geq \frac{E\{|h_{11}h_{22} - h_{12}h_{21}|^2\} \epsilon^2}{|1 - \exp(j\theta t)|^2 E\{|h_{12}|^2|\} h_{11}|^2 + E\{|h_{12}|^2|\} h_{11}|^2}
\]

Assuming time-invariant Rayleigh-fading, i.e. all channel coefficients \( h_{ij} \) are independent complex Gaussian variables with zero mean and unit variance, we yield a lower bound for the mean SIR

\[
\text{SIR} \geq \frac{\epsilon^2}{1 - \cos(\theta t)} + 2 \approx 2 \left( \frac{\epsilon^2}{(\theta t)^2} + 1 \right) \quad \text{(11)}
\]

where we used the Taylor expansion of cosine for small angles, \( \cos(\alpha) \approx 1 - 0.5\alpha^2 \), to obtain the approximation on the right hand side.

For a static channel, we assume \( \text{SIR} > 1000, \epsilon = 1 \) and \( t = N \text{ ms} \), where \( N \) is the delay for the channel state information (CSI) in units of the transmission time interval (TTI) of \( 1 \text{ ms} \) in 3G-LTE. In this way we get the rule that \( \theta \leq 45 \text{ Hz}/N \). Assuming \( N = 5 \), the frequency offset should be less than \( 3 \cdot 10^{-5} \) at \( 2.6 \text{ GHz} \). This is more than an order of magnitude smaller than the common frequency offset spread of standard factory models of commercial base stations.

III. Physical layer synchronization

A. Primary clock reference

In order to facilitate coherent cooperation, the primary clock boards of cooperative base stations must hence be synchronized to a common external reference clock. A well established solution for DVB-T is based on the GPS. There are commercial Rubidium- and crystal-based reference clocks which can be phase-locked to the GPS.

Each GPS satellite transmits a specific Gold sequence with \( 1 \text{ ms} \) period from which the phase information is recovered at the receiver. For carrier phase tracking, it is important to correct the individual Doppler shift due to the fast motion of each satellite. Commercial GPS receivers cover the corresponding signal processing, i.e. the phase is not directly accessible. Instead, GPS-receivers provide a one pulse per second (1PPS) output reference signal. The 1PPS signal is accompanied by a string defined by the National Marine Electronics Association in the NMEA-0183 protocol [11] providing precise time and location among other information. Next, the clock device uses an internal reference oscillator, e.g. Rubidium or oven-controlled crystal oscillators (OCXO), that is phase-locked to the 1PPS signal. A high Q-factor and low phase noise is needed to stabilize the clock between the 1PPS pulses.

The timing offset between two commercial AccuBeat AR83A [12] GPS-locked Rubidium references is shown in Fig. 1. After 15 minutes warm-up time, the mutual timing offset between two clocks averages to \( 100 \text{ ns} \). The measured frequency offset is \( 6 \cdot 10^{-10} \). If both clocks are operated for one day with free LOS and wide field-of-view (FOV) to the sky, as on top of a fixed antenna mast, the mutual timing offset on average reduces to \( 10 \text{ ns} \) and the frequency offset to \( 5 \cdot 10^{-12} \).

Timing and frequency offset of such commercial Rubidium reference clocks are hence much better than needed in practice, see Section II. Note that base stations already use one part of
these reference clocks. Costly crystal oscillators with high Q-factors and low phase noise are commonly used as a primary reference clock in each base station.

In free-running mode, two exemplary primary reference clocks revealed a relative large frequency offset of $10^{-7}$. The short-term phase drift of one reference is only about $\pi/12$ in $10\,\text{ms}$ at $2.6\,\text{GHz}$. The offset from the mean carrier frequency is hence less than $2 \cdot 10^{-9}$ which is better than needed for base station cooperation, see Section II. Hence, integrating the GPS synchronization into the primary clock board may be a promising option for cooperative base stations. It would be sufficient to lock the mean carrier frequency of the primary reference clock to the GPS as in a commercial GPS-disciplined clock. The high costs of commercial GPS-disciplined reference clocks can be significantly reduced in this way.

Indoor base stations may have no free line-of-sight (LOS) and wide FOV to the sky. Moreover, the infrastructure in a flat cellular network architecture consists of a power supply and a wired or wireless network connection technology typically based on Ethernet over copper and replacements as optical fiber, microwave or free-space optics.

In order to provide primary frequency synchronization for indoor base stations as well, the only option is network synchronization. There are several timing protocols, out of which the precise timing protocol (PTP) specified in the IEEE 1588 standard [13] is the most promising. PTP operates over standard Ethernet and defines a ping-pong protocol [14] between a master and a slave. As a result, the network propagation time between the master and the slave is precisely measured every 2 seconds and at even higher rates in the more recent version V2 of the standard. The primary clock at the slave may be similarly stabilized as for GPS synchronization using local oscillators with high Q-factor and low phase noise.

Notice that the PTP protocol performance strongly depends on the network infrastructure [15]. While a simple cross-over cable has negligible timing jitter (4.5\,ns), switches (70\,ns), multilayer switches (76\,ns) and in particular routers (20\,\mu s) have a wider spread of the measured packet delay. These values are valid for extremely low bandwidth utilization in the backbone network. By averaging over longer periods, precise timing can be achieved similar as for GPS synchronization.

However, if the bandwidth utilization in the backbone increases above 5\%, the packet delay spread grows significantly. Switches and routers could measure and correct the load-specific switching and routing times internally, but this is not supported by the current infrastructure. A potential solution is to prioritize PTP packets used for network synchronization. If all other queues are halted in a switch or router when a PTP packet is passed through, the packet delay spread can be significantly reduced and end-to-end packet delay measurements using PTP become more reliable.

The proposed primary synchronization architecture for cooperative base stations includes both GPS and network synchronization. Selected outdoor base stations are synchronized via the GPS and in addition they act as PTP servers. This combination of GPS and PTP server is called a grandmaster. Other base stations, in particular indoors, act as PTP slaves. They are linked to the next PTP server over a fixed or wireless multihop network in which PTP packets are handled with the highest priority.

B. Frame synchronization

So far, we have considered the synchronization of the carrier and sampling frequencies, which are both phase-locked to the primary clock in a base stations. Synchronization of the frame structure is also needed for cooperative BSs. It can be based on the 1PPS signal. A frame length in 3G-LTE is 10 ms, i.e. after 100 frames there is a new 1PPS pulse. Based on a synchronized primary clock, the intermediate 99 frame starts can be derived in each base station (BS) independently using a counter which is reset by the 1PPS pulse.

As described below, resource assignment is related to the frame number which is synchronized as well. Frame number synchronization is derived from combining the 1PPS signal with the NMEA-0183 string provided each second by the GPS. The string contains the coordinated universal time (UTC). At the BS, the frame number is encoded as the UTC time in the most significant bits (MSB) and the frame numbers 0...99 within a second as the 7 least significant bits (LSB). With a 32 bit frame number one could precisely identify each frame transmitted in more than a year using the remaining 25 MSB. The frame number is transmitted over the down-link control channel, and a precise time synchronization becomes available at each mobile terminal. This might be of interest for future real-time protocols and services.

IV. MAC LAYER SYNCHRONIZATION

A. From centralized to distributed cooperation

Two forms of cooperation are discussed in the literature: centralized and distributed cooperation. In the classical realization of base station cooperation, cooperative beam-forming is performed at a central processing unit, where the data streams are synchronous. After the beam-forming, the transmitted waveforms are distributed with propagation times much...
smaller than the cyclic prefix length to each cooperative base station. Such a centralized cooperation may be applicable to legacy cellular networks, such as global system for mobile communications (GSM) and wideband code division multiple access (WCDMA) in which the network is coordinated by a remote network controller (RNC).

In more recent cellular systems such as High Speed Packet Access (HSPA) and 3G-LTE, any lower layer processing, i.e. transmit and receive processing and even the scheduling in the MAC layer, is performed independently at each base station. There is no RNC which could coordinate the transmission of multiple base stations. Adjacent base stations exchange information in order to coordinate their work. Consider the wireless access network architecture used by the 3G-LTE. It is accompanied by a service architecture evolution (SAE). The new architecture is described in [16] and shown in Fig. 2.

Data is distributed over a so-called S1 interface from an advanced gateway (aGW) to the base station. For information exchange between different BS, e.g. for coordinated radio resource management (RRM), there is a point-to-point interface denoted as X2. Note that this is a logical network structure. S1 and X2 links may be multiplexed over the same physical network, using e.g. Gigabit Ethernet. Current latency on X2 is specified as 20 ms.

Therefore, we need a distributed realization in order to integrate base station cooperation into future cellular systems. The cooperative beam-forming has then to be performed in the physical layer of each base station separately, and it has to be organized over the backbone network. In order to make the consequences of the distributed approach more transparent, we use a mathematical model for the synchronous transmission of two base stations serving two terminals in different cells

\[
\begin{bmatrix}
    x_1 \\
    x_2
\end{bmatrix}
= \begin{bmatrix}
    W_{11} & W_{12} \\
    W_{21} & W_{22}
\end{bmatrix}
\begin{bmatrix}
    d_1 \\
    d_2
\end{bmatrix},
\]

(12)

where \(d_1\) and \(d_2\) are the data signals intended for terminals 1 and 2, respectively, \(W_{ij}\) are the cooperative beam-forming weights, and \(x_1\) and \(x_2\) the waveforms transmitted at both base stations.

In the centralized approach, the data signals \(d_1\) and \(d_2\) are demultiplexed at the central station and thus they are naturally synchronous. The transmit waveforms \(x_1\) and \(x_2\) are distributed timely over a synchronous network technology such as synchronous digital hierarchy (SDH) or open base station architecture initiative (OBSAI) to the cooperative base stations.

In a distributed realization, however, the data streams \(d_1\) and \(d_2\) and not the waveforms are distributed over the network. The CSI needed to calculate the beam-forming weights is exchanged frequently between the base stations involved in the cooperation. Weights are obtained by running in each base station redundantly the same algorithm as for centralized cooperation. In our example shown in Fig. 3, each base station 1 and 2 compute their own transmit waveform \(x_1\) and \(x_2\) independently from the data streams \(d_1\) and \(d_2\) and the locally computed weight vectors \((W_{11}, W_{12})\) and \((W_{21}, W_{22})\), respectively. In practice, Eq. 12 implies that both data streams \(d_1\) and \(d_2\) have to be synchronized so that the correct data is available when the cooperative beam-forming is carried out locally.

This is a challenging problem which has not been previously addressed, to the best of our knowledge. It has been proposed [17] to use existing multicast protocols to distribute the data more efficiently to the set of cooperative base stations forming a multicast group. But it has not been recognized that the data packets arrive asynchronously at different nodes after passing through an unknown network. Consequently, one has to synchronize the data streams again in each cooperative base station prior to performing cooperative beam-forming.

Consider the enormous complexity of MAC processors for 3G-LTE, see e.g. [8] for an initial implementation example. In the MAC processor, packets are classified according to the internet protocol (IP) packet number, destination address and type of service (TOS) fields extracted from the IP header [18]. Next, the IP packet is filled into a particular user queue. Note
that packets can be disordered in these queues, since IP packets may be received over different routes in the network. In 3G-LTE, there is a convention that a transport block is terminated after 1 ms forming a TTI. According to the assigned user bandwidth in this time slot an IP packet may be segmented into multiple transport blocks having variable lengths. Retransmissions are organized in multiple parallel hybrid automatic repeat request (HARQ) processes. Each transport block is encoded and mapped onto the space-time-frequency resource grid. Resource assignment and modulation and coding scheme (MCS) are adaptive and steered by a scheduling algorithm in the control plane. In this way, multipath as well as multiuser diversity gains in the wireless channel and statistical multiplexing gains for multiuser traffic can be realized.

All these complex user plane processes and the control plane algorithms behind finally result in the two data streams $d_1$ and $d_2$, refer to Eq. 12. The streams need to be provided synchronously at the interface to the physical layer prior to the cooperative beam-forming. As a consequence, the MAC layers of distributed cooperative base stations need tight synchronization. In the following we describe two proposals. The essential difference is the processing stage in the user plane where the data are replicated and distributed.

**B. Multicast on S1**

In this proposal, data of terminals in a cooperative set are multicast from the aGW to all BSs involved in the cooperation. Since the data arrive asynchronously and possibly disordered at the BS, MAC processing is synchronized over X2, as detailed in Fig. 4.

There is a copy of the entire user plane of the master BS in any cooperating BS involved in the cooperation as a slave. The control plane of the master BS remotely controls the synchronous data flow in the cooperative user plane replicas with the same process in the master BS. In more detail, each IP packet received from the network is classified and buffered in the corresponding user queue. The user plane operates on a 1 ms clock basis which is the length of 1 TTI. We have to make sure that the same data are read out of the queues in each cooperative base station as in the master station. But the position of the requested data in these queues may be different, due to the potential disorder of the IP packets. Moreover, there is an address offset in each IP packet due to the accumulated length of the transport blocks belonging to the same IP packet but being processed already in previous TTls. In order to synchronize the read-out of data from the queues, we need to know the time-to-send (TtS) which is given by the TTI number and the radio frame. Furthermore we need the IP packet number and the address offset inside the IP packet. The mapping information for the transport block is created by the scheduling algorithm running in the master base station. This compound information is denoted as the time stamp in the following.

Only the stamp but no data are sent for each TTI from the master to all slave base stations. Each slave creates a pointer to the right data in its queue, and performs for the same transport block redundantly segmentation, channel coding and mapping as in the master base station. The only exception concerns retransmitted data where timely exchange of control information over the network may become critical. As a way out, retransmissions may be realized on exclusive resources and only by the master base station so that interference from other base stations is avoided.

Characterizing this approach, there is a remaining central functionality in the aGW where the multicast group of the cooperation is hosted. The data rate on S1 is multiplied due to the multicasting but in turn the X2 interface is exclusively used for control information. Copies of the user plane processing require additional hardware effort at each cooperative base station. Formatting of the time stamp has to be specified.

**C. Unicast on S1**

This proposal is shown in Fig. 5. IP packets are transmitted from the aGW over S1 to the serving BS only, where a unique instance of the MAC processor is situated, as in non-cooperative base stations. The master station organizes the

\[\text{Eq. 12}\]
multicast group for the transport of scheduled and encoded data over the X2 interfaces, i.e., after passing the data through the entire user plane at the master station. Mapping information for each TTI is forwarded together with those ready-to-send data sequences to the slave BS via the X2 interface.

With a broad-band X2 interface, implementation of the second approach may be simpler and requires less hardware effort. There is no redundant user plane processing which needs remote control. Scheduled data are distributed and only the mapping functionality is replicated in each cooperative base station. This has the advantage that data and non IQ samples are transmitted over the X2 network which reduces the data rate. A striking advantage is that retransmissions are possible in the cooperative mode as well. There is less data load on S1, since IP packets are sent only to the master BS. But now we have significant data load on the X2 interface. An additional data transport protocol on X2 has to be specified, accordingly. The sum of the data rates on S1 and X2 may be slightly higher in the second approach.

D. Latency

In both approaches, cooperating base stations exchange the in-cell and out-of-cell channel information extracted from the received pilot signals at each terminal. In this way, the weight calculation can be performed redundantly in each BS. Since the channel must not change until the weights are calculated and applied to the data, the latency of the X2 is critical. It should be below 1 ms, accordingly. This can be reached by several networking technologies as long as the propagation distance is limited by the latency of control information on X2. Data can only be transmitted in cooperative mode after the control information has been exchanged and the data flow in the user plane is synchronized. The longer we have to wait for this information, the longer the data transmission will be delayed.

CONCLUSIONS

We have studied synchronization requirements to enhance the capacity in future cellular networks using base station cooperation. A combination of GPS and network synchronization based on the PTP protocol may be used for synchronizing the primary clock as well as the frame structure in distant base stations. The costs of network-wide synchronization can be significantly reduced by integrating GPS or PTP receiver and a phase-locked loop in the primary clock boards of each base station. Cooperative beam-forming can be integrated in the current cellular system architecture in a distributed form. This implies that we need to synchronize also the flow of data received asynchronously in each cooperative base station over the backbone. Two possible solutions have been described and compared.

ACKNOWLEDGMENTS

The authors are grateful for financial support from the German Ministry of Education and research (BMBF) in the national collaborative projects ScaleNet and EASY-C. Thanks also to C. von Hemolt (HHI), R. Halfmann (NSN) and R. Irmer (Vodafone) for inspiring discussions.

REFERENCES