Asynchronous Simultaneous Small Packet Transmission in Cellular Wireless System

Chandra S. Bontu*, Jagadish Ghimire†, Shalini Periyalwar†, Mark Pecen†
Advanced Technology, BlackBerry
*4000 Innovation Drive, Ottawa, Canada K2K 3K1, E-Mail: cbontu@blackberry.com
†176 Columbia Street West, Waterloo, Canada N2L 3L3, E-Mail: jghimire, shperiyalwar, mpecen@blackberry.com

Abstract — **Existing cellular wireless systems operate with a predominantly network controlled, connection oriented frame based protocol architecture. A User Terminal (UT) is required to signal the network frequently to maintain its connectivity, and makes requests for uplink bandwidth to transmit packets. In such systems, for emerging applications, such as machine-to-machine (M2M), where small packets are transmitted sporadically, the bandwidth required for signaling is much more than the actual bandwidth to transmit the data. In this paper, we propose a novel connectionless uplink simultaneous accessing technique without the need of strict synchronization. UTs pick an uplink (UL) resource which is advertised by the network as a common radio resource for small packet transmissions, and transmit the data packets. The data packet includes a preamble sequence and network assigned identity. The transmitted packets from each UT may not be in perfect synchronization with the UL system timing. The base station (BS) uses novel multi-user detection technique to separate these transmissions. Simulation results of the proposed method for an OFDM based system are shown be promising.**

Keywords: Simultaneous access, Multi-user detection, Connectionless radio access.

I. Introduction

A variety of applications are envisioned to be supported by future cellular systems. For some of these applications, the bandwidth used for signaling (to set up and maintain the connection, requesting radio resources etc.) is more than the bandwidth used for transmitting the actual data. With the increasing proliferation of smart phones and the corresponding dearth of spectrum, there is a need to consider efficient means of allowing multiple users to access the same resource, without imposing stringent connection oriented network control.

Applications such as Machine-to-Machine (M2M) can cause severe congestion while using cellular radio systems [1][2]. The reason for congestion can be attributed to the overwhelming number of low-bit rate User Terminals (UTs) trying to maintain network connectivity and further sending scheduling requests (SRs) for transmitting small packets. To circumvent this congestion, one can consider contention based radio access mechanisms. However, a contention based access mechanism, if not properly configured, may cause many collisions.

In [3], a coded expanded random access method has been proposed. In this method, a contention window is used, during which the UTs contend for the resources that are available in a non-contention window. A new contention based access (CBA) for machine type communications for 3GPP LTE has been presented in [4], wherein the base station (BS) can broadcast a resource grant for multiple UTs to contend. The UTs will also include their temporary identity assigned by the network when the data is transmitted over the resource. The temporary identity is protected by robust coding for reliable detection at the BS. If the BS detects a collision, a dedicated resource grant is given to the UTs in the next instant. Small and sporadic packet transmissions from many UTs can overload the physical resource dedicated for the random access channel. In [2], a novel method to eliminate the congestion on Random Access Channel (RACH) has been presented. A joint estimation and contention resolution protocol for random access has been presented in [5]. The proposed contention protocol operates in rounds that are divided into equal duration slots. The emphasis is that the initial round is used to reliably estimate the number of users and the subsequent rounds are used to resolve the transmissions. In [6], a contention mechanism has been proposed which relies on the BS broadcasting the rate at which the UTs can initiate or reinitiate the random access procedure.

In this paper, we propose an approach where multiple UTs can simultaneously transmit their data packets without strict UL synchronization. The BS broadcasts an available set of radio resources for applications which transmit small and sporadic data packets. Further, to distinguish various transmissions, the BS broadcasts a set of preamble sequences, which are inserted in to the packets transmitted by the UTs. The position of these preamble sequences with the data packets is also optionally specified by the BS and may be UT specific. The BS can use novel iterative multi-user (MU) techniques to detect the individual bursts. The complexity of the proposed scheme is at the BS, further making it very attractive for applications such as M2M communications.

The paper is organized as follows. Section II contains the system description for the proposed asynchronous simultaneous access scheme for small packet transmission. The transmission model employed at the UT is described in Section III. Section IV provides details on the receiver processing at the BS. Section V contains the performance results for the proposed scheme, and Section VI provides the conclusions.

II. System Description

The category of UTs considered in this paper may be infrequently uploading small packets to the network. This is a typical scenario with many new applications, such as in machine to machine communications. The UTs register with the network via any cellular network, such as LTE. We assume that the UTs are not sending any extra signaling to the network for maintaining connectivity, and particularly may not be maintaining UL synchronization. The BS assigns some of the available UL radio resources to small and sporadic packet transmissions from the UTs. Whenever a UT wants to send a packet, it listens to the DL for a broadcast message from the BS, with the descriptor of the common radio resources for small packets. It then forms a packet and transmits over the common resources. The UL timing is adjusted based on the DL receive timing.

As illustrated in Fig. 1, for example, in a 10 MHz LTE system, some of the frequency-time resources can be allocated to small packet transmissions. A bandwidth of 180 kHz in a subframe of 1 ms may form a basic unit of radio resource for this purpose. Subframes in which these resources are available may be defined in the system information broadcast message. Guard time Δ_{GT} and a guard band Δ_{GB} are provisioned based on the deployment scenario.

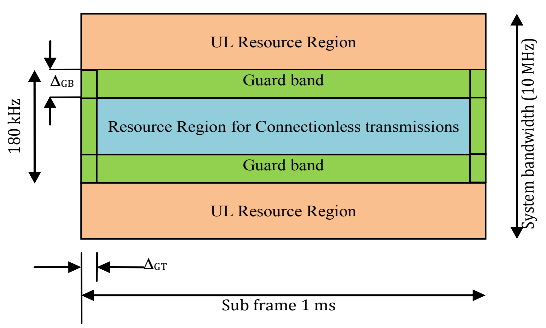

Fig. 1. Radio resource grants for connectionless transmissions in a 10 MHz LTE uplink.

III. Transmission Model

UT-ℓ forms its transmission payload as illustrated in Fig. 2. The Medium Access Control (MAC) payload consists of the Radio Link Control (RLC) Protocol Data Unit (PDU) along with a network assigned UT ID. A Cyclic Redundancy Check (CRC) is calculated and appended to the MAC PDU. A physical layer PDU is formed by applying forward error correction and subsequent symbol mapping. A network assigned set of preamble sequence symbols, $\{P_{k1}^{\ell}, k1 = 0,1,\dots,C-1\}$ are inserted into the physical layer PDU every η symbols starting from the symbol-α^{ℓ}. The transmitted symbol vector can be expressed as follows:

$$S_k^{\ell} = \begin{cases} P_{k1}^{\ell} & \text{for } mod(k,\eta) = \alpha^{\ell} \\ D_{k2}^{\ell} & \text{otherwise} \end{cases} \quad \text{for } k = 0,\dots,L-1 \tag{1}$$

where $k1 = \lfloor k/\eta \rfloor$ and $k2 = \lfloor k/\eta \rfloor \eta + mod(k,\eta) - 1 + \alpha^{\ell}$. An L point Inverse Discrete Fourier Transform (IDFT) is then performed to obtain $\boldsymbol{s}^{\ell} = \{s_m^{\ell}, m = 0,1,\dots,L-1\}$, as

$$\boldsymbol{s}^{\ell} \overset{\mathcal{F}_L}{\leftrightarrow} \boldsymbol{S}^{\ell}$$

where $\mathcal{F}_L$ represents an L point Discrete Fourier Transform (DFT) operation.

$\boldsymbol{s}^{\ell}$ is transmitted after prepending it with a cyclic prefix of L_{CP} symbols.

The transmit power of the payload is adjusted such that the received power at the BS meets the target receive power level as broadcasted by the BS. This can be performed by measuring the path loss on the downlink.

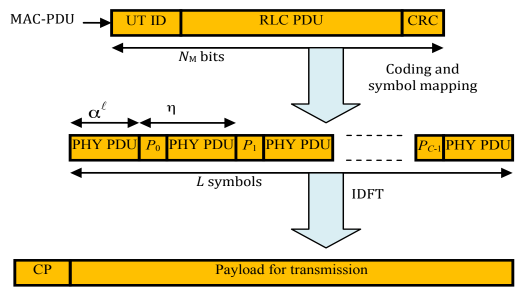

Fig. 2. Packet structure for small packet transmission

UT-ℓ adjusts its uplink transmission timing and transmits at suitable power level by listening to the DL transmission from the BS (e.g., by using the cell specific reference symbols transmitted by the BS in LTE systems).

Preamble assignment:

UT-ℓ is assigned a preamble sequence $\boldsymbol{P}^{\ell}$ selected from a set of available preamble sequences, $\boldsymbol{P}(v), v = 0,1,\dots,N_P - 1$. The UTs can be differentiated by the preamble sequence at the receiver.

IV. Receiver Processing

As described in the previous sections, since the UL transmissions from the UTs are not perfectly synchronized in reaching the BS, the received signal can be expressed as a sum of time shifted packets from each UT as

$$r_m(n) = \sum_{\ell=0}^{N-1} \sum_{i=0}^{D-1} h_i^{\ell}(n) s_{m-i+\Delta_{\ell}}^{\ell} + n_m(n) \quad \text{for } m = 0,1,\dots,L-1+L_{CP} \tag{2}$$

where $r_m(n)$ and $n_m(n)$ represent the mth sample of the received signal and the thermal noise plus other interference received at the nth receive antenna of the BS, respectively: N is the number of simultaneously transmitting UTs; s_i^{ℓ} represents the ith transmitted sample from the ℓth UT; $\{h_i^{\ell}(n), i = 0,\dots,D-1\}$ are the channel coefficients of the channel between the UT-ℓ and the nth receive antenna of the BS; and Δ_ℓ represents the relative receive time difference between the UT-ℓ and the UT-0 transmissions (i.e. $\Delta_0 = 0$). Without loss of generality, we assume the UTs are ordered such that UT-0's transmissions arrive first and so on.

The corresponding frequency domain version of the received signal, $\boldsymbol{r} \overset{\mathcal{F}}{\leftrightarrow} \boldsymbol{R}$, can be expressed as follows ($\mathcal{F}$ represents the DFT operation):

$$R_k(n) = \sum_{\ell=0}^{N-1} H_k^\ell(n) S_k^\ell e^{\frac{j2\pi\Delta_\ell k}{L}} + N_k\ (n) \quad \text{for } k = 0,1,\dots,L-1 \tag{3}$$

Here we assume that the DFT window is synchronized with the received packet from UT-0. The symbols in upper-case and lower case represent the frequency-domain and time-domain representations, respectively.

Fig. 3 illustrates the various steps involved at the BS in detecting the transmitted packets from the UTs. As indicated in equation (3), the first step in the detection process is to estimate the relative time offsets, Δ_ℓ between the UT packet transmissions. Subsequently, based on these estimated candidate offsets, the channel weights for each of these transmissions are estimated. Finally the channel weights and the corresponding time offsets are used to detect the data.

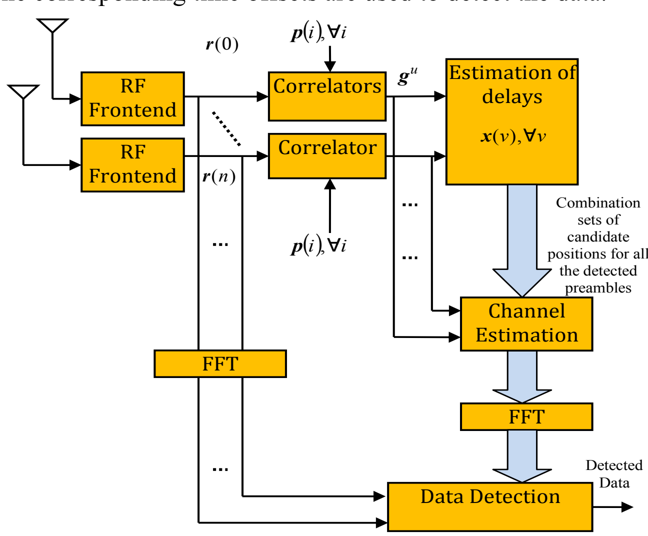

Fig. 3. Block diagram of the receiver processing at the BS

As shown in Fig. 3, the received base band signal is correlated with the set of available preambles to find the approximate timing of the data packet from each of the UTs. This correlation function can be performed either in time domain or frequency domain. The received signal in the time domain can be correlated with an L point IDFT of each preamble sequence. For example, the correlator outputs for the nth receive antenna can be expressed as follows:

$$g_m^u(n) = \sum_{i=0}^{C-1} r_{m-i}(n) p_i^*\ (u) \quad \text{for } m = 0,1,\dots,L+L_{CP}-1 \tag{4}$$

where $g_m^u(n)$ is the mth sample of the output of the uth correlator, and $\{p_i(u), i = 0,1,\dots,C-1\}$ is the C point IDFT of the uth preamble sequence, i.e., $\boldsymbol{p}(u) \overset{\mathcal{F}_C}{\leftrightarrow} \boldsymbol{P}(u)$. The superscript * indicates complex conjugate operation.

When $\left|g_{m_0}^u(n)\right|^2 > \nu E[\|r_m(n)\|^2]$, m_0 is declared as one of the candidate positions for the preamble sequence, $\boldsymbol{P}(u)$. The scaling factor ν is a design parameter which is selected such that the performance and computational complexity are optimized. These candidate positions, $\boldsymbol{m}_0(u)$, are collected over the received signals from all the receive antennas.

As depicted in Fig. 2, the preamble symbols are separated by η data symbols. Therefore, in the time domain if the preamble is found at position m_0, the positions $m_0 + C, m_0 + 2C, \dots, m_0 + (\eta-1)C$ are also expected to have time domain response of the preamble sequence. Thus, the candidate positions can be reduced to $\widetilde{\boldsymbol{m}}_0(u) = mod(\boldsymbol{m}_0(u), C)$. For all these candidate timing positions, the following metric is calculated:

$$t_m^u(n) = \left|\sum_{l=1}^{\eta}\sum_{i=0}^{C-1} r_{m+lC-i}(n) p_i^*\ (u)\right|^2 \quad \text{for } m \in \widetilde{\boldsymbol{m}}_0(u) \tag{5}$$

Based on the metric $t_m^u(n)$, the candidate positions, $\widetilde{\widetilde{\boldsymbol{m}}}_0(u)$, are reduced further by comparing $t_m^u(n)$ with another threshold $\nu_1 E[\|r_m(n)\|^2]$. For the multiple receive antennas, the candidate positions for the same preamble are obtained as an *intersection* of the two candidate position sets on each antenna.

From all the candidate positions estimated for all the N_P preambles, N_u combination sets of candidate positions, $\{\boldsymbol{x}(v), v = 0,1,\dots,N_u-1\}$, are formed by picking one candidate position for each preamble. For example, if the length of $\widetilde{\widetilde{\boldsymbol{m}}}_0(u)$ is ρ_u, $N_u = \prod_{i=0}^{N_P-1}\rho_i$.

In this paper, we assume that channel is constant over the radio resources allotted for one packet, i.e. 180KHz in 1 ms. The channel weights for each preamble are estimated for each candidate set $\boldsymbol{x}(v)$ as follows:

$$\hat{h}^u \cong \frac{1}{\eta}\sum_{l=1}^{\eta}\sum_{i=0}^{C-1} r_{m+lC-i} p_i^*\ (u) \quad \text{for } m \in \boldsymbol{x}(u) \text{ and } u = 0,1,\dots,N_P-1 \tag{6}$$

Ideally, preamble sequences are expected to have very small cross-correlations with each other both in frequency domain and time domain. However, practical sequences (for example, Zadoff-Chu Sequences) often have non-negligible cross-correlation terms. These cross-correlation terms can cause the quality of channel estimation to degrade significantly. Since the transmitted preambles are *a priori* known at the receiver, this effect can be compensated as follows:

$$\hat{\hat{\boldsymbol{h}}} = \boldsymbol{\Gamma}^{-1}\hat{\boldsymbol{h}} \tag{7}$$

where $\hat{\boldsymbol{h}}$ a column vector consisting of the channel estimates for all the preambles computed according to equation (6); $\hat{\hat{\boldsymbol{h}}}$ a column vector consisting of the cross-correlation-compensated channel estimates and $\boldsymbol{\Gamma} = \boldsymbol{\phi}\boldsymbol{\phi}^H$; $\boldsymbol{\phi} = [\boldsymbol{p}(0)\ \boldsymbol{p}(1) \ldots \boldsymbol{p}(N_P - 1)]^T$; $\boldsymbol{p}(u)$is a row vector consisting of all the symbols from the u th preamble sequence, $\boldsymbol{p}(u) = [p_0^u\ p_1^u \ldots p_{C-1}^u]$. The matrix $\boldsymbol{\Gamma}^{-1}$ can be pre-calculated at the receiver.

For data detection, the receiver picks the lowest index, w, of $\boldsymbol{x}(v)$, and computes the frequency domain sample of the received vector as follows:

$$\boldsymbol{R}(n) = \mathcal{F}_L(\{r_w(n), r_{w+1}(n), \ldots r_{w+L-1}(n)\}) \quad \text{for } n = 0, 1, \ldots, N-1 \tag{8}$$

The data symbols can be estimated using the maximum likelihood rule. We assume that the information data and the interfering data are QPSK modulated, for simplicity. It is straight forward to generalize the above estimate for other modulation schemes. An ML data estimator evaluates the following expression:

$$\hat{S}_k^l = \frac{\sum_i A_i p\left(R_k / S_k^l = \mathcal{A}_i\right)}{\sum_i p\left(R_{km}/S_k^l = \mathcal{A}_i\right)} \tag{9}$$

where $\mathcal{A}_i \in \left\{\pm\frac{1}{\sqrt{2}} \pm \frac{j}{\sqrt{2}}\right\}$ and …

$$p\left(R_k / S_k^l\right) = \sum_{S_k^0}\sum_{S_k^1} \cdots \sum_{S_k^{l-1}}\sum_{S_k^{l+1}} \cdots \sum_{S_k^{N-1}} p\left(R_k / S_k^l, \tilde{\boldsymbol{S}}_k^l\right) \tag{10}$$

where $\tilde{\boldsymbol{S}}_k^l = \left\{S_k^0, S_k^1, \ldots, S_k^{l-1}, S_k^{l+1}, \ldots S_k^{N-1}\right\}$.

$p\left(R_k / S_k^l = \mathcal{A}_i\right)$ represents the conditional probability of R_k. To detect the symbol $\hat{S}_k^l$, transmitted by the UT-l, the conditional propability $p\left(R_k / S_k^l, \tilde{\boldsymbol{S}}_k^l\right)$ is averaged over all the possible transmitted symbols from the other UEs as shown (10). The conditional probability is calculated over the receiver signals over all the receive antennas as follows:

$$p\left(R_k / S_k^l, \tilde{\boldsymbol{S}}_k^l\right) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left\{\frac{\sum_n \left|R_k(n) - \sum_{i-0}^{N_P-1} \widehat{H}(n)_l S_k^l\right|^2}{2\sigma_n^2}\right\}, \tag{11}$$

where σ_n^2 is the variance of the thermal noise as represented in equation (2), i.e. $\sigma_n^2 = E[n_m^2]$; $E[\ldots]$ represents the expected value. The detected data symbol vectors corresponding to the minimum residual mean squared error (MSE), or whichever of them pass CRC among all $\boldsymbol{x}(v)$, are selected as the recovered data.

V. Simulation Results

The performance of the proposed system is evaluated through computer simulations. In the simulations, we assume a packet of length, 96 symbols: 24 symbols of preamble and 72 symbols of data. The modulation scheme used by all the user terminals is QPSK. The preamble is a 24 symbol Zadoff-Chu sequence [7] which is created from 29 length Zadoff-Chu sequence truncated to 24 symbols. The roots used to generate these Zadoff-Chu sequences are 5, 7, 13 and 19. The roots are selected such that the roots and differences of the roots are prime to the length of the sequence. The transmit power of the preamble and data symbols is assumed to be equal in our evaluation.

In the simulations, we assume two receive antennas at the BS. The number N of UTs that are simultaneous transmitting are fixed for each simulation run and are varied across $N = 2, 3$ and 4. For simplicity, the channel is assumed to be constant over one packet transmission, i.e. 180kHz in 1 ms. The proposed method can be easily extended to the frequency selective channel. Further the channel model is assumed to quasi-static, i.e. the channel is constant for each packet transmission and an independent channel weight is generated for each packet transmission. The average received power at the BS from each UT is assumed to be the same. Since the UT listens to the DL transmission before transmitting the data packet on the UL, it can adjust the transmit power such that the path loss is compensated to meet the required received power level at the BS.

The data and preambles are arranged as shown in Fig. 4 below. The value of η is equal to 4. The pilots used by each UT are assumed be pre-assigned by the BS. The time-domain signal of each OFDM symbol is generated by IFFT operation, which is further appended with a Cyclic Prefix (CP) of length equal to 1/4*th* of the OFDM symbol. The transmission timing of the data packets is randomly selected for each packet or each packet burst from 0 to the duration of the CP. This allows us to simulate a large range of timing offset among the packets from different users. The packet burst consists of multiple packets transmitted consecutively by the UTs. For example if the expected PHY payload size is 54 octets (which fits in 3 data packets), then three consecutive resource units are assigned to the UTs.

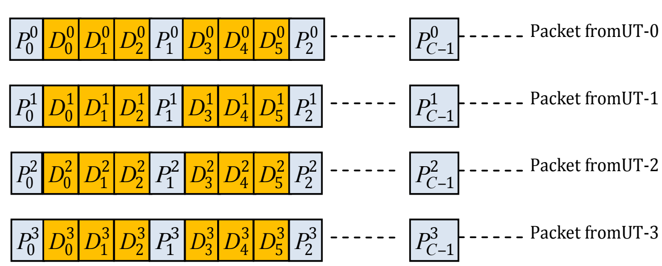

Fig. 4. Simulation scenario

Fig. 5 depicts the average bit error rate (BER) as a function of average signal to noise ratio (SNR). The SNR is defined as the average received power at the BS for each UT transmission to the receiver's thermal noise power level. The BER results when the channel weights are perfectly known are also shown for comparison. As shown there is very small loss resulting from the channel estimation. The channel estimation using the time-domain correlation is refined by the use of the **Γ**matrix as stated earlier, to cancel the non-negligible cross-correlation between these sequences. These simulations are assuming perfect timing offset recovery As long as the pilot detection scheme is able to find the correct peak as one of the candidate peaks, our scheme would eventually be able to choose it at the end of data detection stage, as discussed earlier. Perfect timing recovery is reasonable to assume when the UTs are transmitting the data packets continuously. In this situation, the BS can estimate the packet timing across consecutive data packet transmissions.

Fig. 6 shows the performance results for two simultaneous users with the effects of packet transmission timing recovery mechanism as described in Section IV. When the timing is recovered across multiple consecutive subframes, the BER performance improves. Here the number of candidate timing positions is limited to a maximum of 4. Better performance is expected by increasing the complexity of the timing recovery circuit. Moreover, the peak detection performance would also improve if the transmit power of the preamble symbols are higher than transmit power of the data symbols.

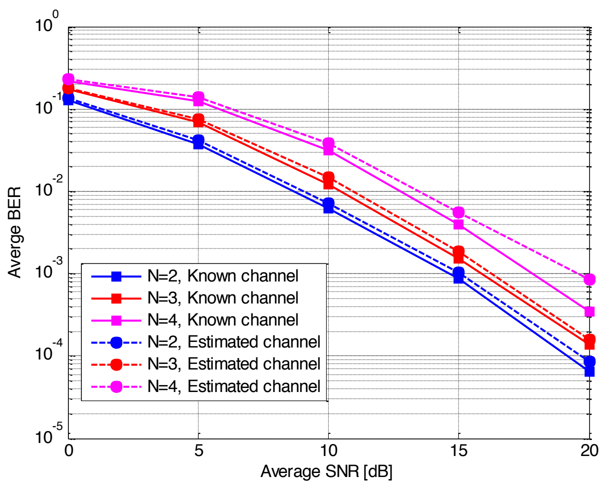

Fig. 5. BER performance for simultaneous transmissions

VI. Conclusions

A novel simultaneous radio access mechanism is proposed for efficient transmission of asynchronous small data packet. The mechanism relies on the use of orthogonal preambles to differentiate simultaneous user transmissions. At the BS, the received overlapping transmissions are separated by novel ML processing. The simulation results indicate that such a scheme can viably support two or more UTs. Thus, such a scheme may be a suitable candidate for the evolution of macro-cellular wireless networks supporting large numbers of UTs that are likely to be transmitting small packets infrequently, without the need to maintain connection with the network all the time. Further improvements are expected with enhanced detection techniques, better preamble sequences and increased receiver complexity. The main advantage with the proposed technique is that most of computation complexity is at the BS, allowing for a quasi connectionless operation on the uplink.

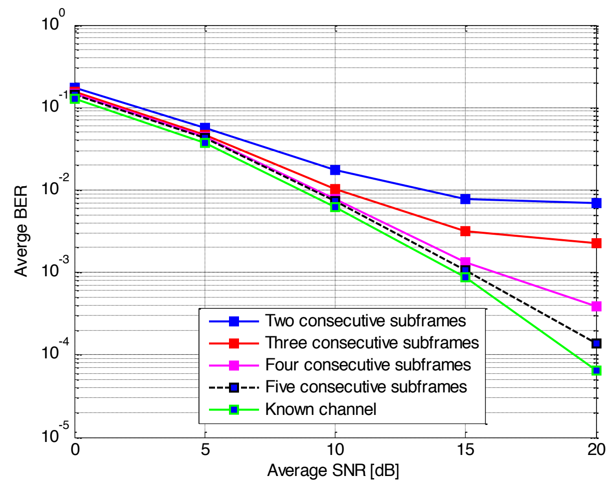

Fig. 6. BER performance with timing recovery over multiple subframes (N=2)

References

[1] A. Maeder, et.al., "The Challenge of M2M Communications for the Cellular Radio Access Network", *11th Würzburg Workshop on IP: Joint ITG and Euro-NF Workshop: Visions of Future Generation Networks, EuroView 2011,* August 1-2 2011, Würzburg, Germany.

[2] A. Lo, et.al., "Enhanced LTE-Advanced Random-Access Mechanism for Massive Machine-to-Machine (M2M) Communications", *27th Wireless World Research Forum, WWRF27-WG4-08, 'A new convergence framework: CLOUD, PIPE and DEVICE over Mobile Broadband,* 18-20 October 2011, Düsseldorf, Germany.

[3] N. K. Pratas, et.al., "Code-Expanded Random Access for Machine-Type Communications", IEEE GLOBECOM 2012 *Workshop: Second International Workshop on Machine-to-Machine Communications 'Key' to the Future Internet of Things,* Anaheim, CA, Dec 3-7, 2012, pp.1681-6.

[4] K. Zhou, et.al., "Contention Based Access for Machine-type Communications over LTE", *IEEE 75th Vehicular Technology Conference (VTC Spring),* 2012, 6-9 May 2012, Yokohama, Japan, pp. 1-5.

[5] C. Stefanovic, et.al., "Joint Estimation and Contention-Resoulution Protocol for Wireless Random Access", arXiv:1210.0693v1 [cs.IT], October 2, 2012, http://arxiv.org/abs/1210.0693v1.

[6] Z. Jiang and X. Zhong, "Fast Retrial and Dynamic Access Control Algorithm for LTE Advanced Based M2M Network", *AICT 2012:The Eighth Advanced International Conference on Telecommunications*, Stuttgart, Germany, May 27, 2012 to June 1, 2012, pp. 24-28.

[7] B. M. Popovic, "Generalized Chirp-Like Polyphase Sequences with Optimum Correlation Properties", *IEEE Transctions on Information Theory*, Vol. 38, No. 4, July 1992, pp. 1406-1409.