A Turbo-Coded Burst-by-Burst Adaptive Wide-Band Speech Transceiver

Thomas Keller, Matthias Münster, and Lajos Hanzo, *Senior Member, IEEE*

Abstract—Turbo-coded burst-by-burst adaptive orthogonal frequency division multiplex (AOFDM) wide-band speech transceivers are proposed. A constant throughput adaptive OFDM transceiver was designed and benchmarked against a time-variant rate scheme. The proposed joint adaptation of source-codec, channel-codec, and modulation regime results in attractive, robust, high-quality audio systems, capable of conveying near-unimpaired wide-band audio signals over fading dispersive channels for signal-to-noise ratios (SNR) in excess of about 5 dB.

Index Terms—Adaptive OFDM, audio compression, transmission, wide-band speech compression.

I. Background and Motivation

BURST-BY-BURST adaptive quadrature amplitude modulation (AQAM) transceivers [1] have recently generated substantial research interests in the wireless communications community [2]–[4]. The transceiver reconfigures itself on a burst-by-burst basis, depending on the instantaneous perceived wireless channel quality. More explicitly, the associated channel quality to be experienced by the next transmission burst is estimated on the basis of the current received burst and the specific modulation mode, which is expected to achieve the required performance target is then selected for the transmission of the next burst. In other words, modulation schemes of different robustness and of different data throughput are invoked [5], [6]. In the event of an anticipated error burst due to a low expected instantaneous channel quality, the transmitter can also be temporarily disabled, while the data are delayed and buffered, until the channel quality improves, provided that the associated delay is not excessive for the service supported. Due to this feature, the distribution of channel errors becomes typically less bursty, than in conjunction with nonadaptive modems. This is an attractive feature in conjunction with channel coding, resulting in potentially increased coding gains [5]. Furthermore, the soft-decision channel codec metrics can also be invoked in estimating the instantaneous channel quality [5]. Recently, block turbo-coded AQAM transceivers have also been proposed for dispersive wide-band channels in conjunction with conventional decision feedback equalizers (DFE) [5], where the mean squared error (MSE) at the DFE's output was used as the channel quality metric, controlling the choice of modem modes. An alternative neural-network radial basis function (RBF) DFE-based AQAM modem design was proposed in [6], where the RBF DFE provided the channel quality estimates for the modem mode switching regime.

Further recent work on combining various conventional channel coding schemes with adaptive modulation has been reported by Matsuoka *et al.* [7], Lau *et al.* [8], and Goldsmith *et al.* [9]. For data transmission systems, which do not necessarily require a low transmission delay, variable-throughput adaptive schemes can be devised, which operate efficiently in conjunction with powerful error correction codecs, such as long block length turbo codes [42]. By contrast, fixed rate burst-by-burst adaptive systems, which sacrifice a guaranteed bit error rate (BER) performance for the sake of maintaining a fixed data throughput, are more amenable to employment in the context of low-delay interactive speech and video communications systems. The above burst-by-burst adaptive principles can also be extended to adaptive orthogonal frequency division multiplexing (AOFDM) schemes [1], [10], and to adaptive joint-detection based code division multiple access (ACDMA) arrangements [11].

OFDM was first proposed by Chang in 1966 [12], and revived by Cimini's often-cited paper [13]; but it was developed to its full potential in the late 1990s, when a whole host of contributions appeared, for example, in [14]. Other developments were due to Rohling *et al.* [20] at the University of Hamburg; Huber *et al.* at Erlangen University [19]; Meyr *et al.* [15], [16] at Aachen University; Jones, Wilkinson, and Barton in the U.K. [17], [18]; and di Benedetto and Mandarini at the University of Rome [21], to name just a few of the key contributors. Further significant advances over more benign, slowly varying dispersive Gaussian fixed links are due to Chow, Cioffi, and Bingham [23] from the U.S., where OFDM or, equivalently, discrete multitone (DMT) modulation became the dominant solution for asymmetric digital subscriber loop (ADSL) applications, potentially up to a bit rate of 8 Mb/s. In Europe, OFDM has been favored for both digital audio broadcasting (DAB) and digital video broadcasting [27], [28] (DVB), as well as for high-rate wireless asynchronous transfer mode (WATM) systems due to its ability to combat the effects of highly dispersive channels [29]. The notion of adaptive bit allocation in the context of OFDM was proposed as early as 1989 by Kalet [22], which was further developed by Chow *et al.* [23], and was refined for duplex wireless links, for example, in [1] and [10]. Finally, various OFDM-based speech and video systems were proposed in [25] and [26]. The co-channel interference sensitivity of OFDM can be mitigated with the aid of multiple antenna assisted reception techniques [32]–[34].

Against this background, in this paper we propose a burst-by-burst adaptive 7-kHz bandwidth audio transceiver scheme, based on turbo-coded multimode constant throughput OFDM. The rationale behind proposing this system was
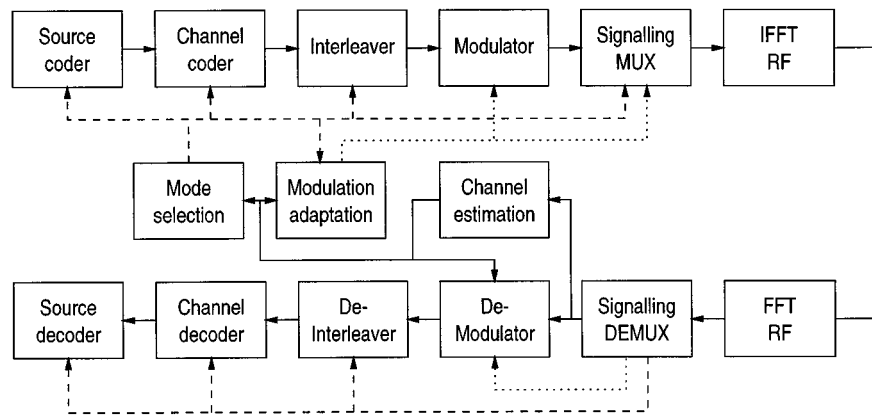
Fig. 1.   Schematic of the multimode adaptive OFDM transceiver.

that nonadaptive OFDM was also a contender [30] for the Pan-European Universal Mobile Telecommunications System [31] (UMTS), and hence it was beneficial to explore the potential of a substantially enhanced turbo coded fixed-rate AOFDM wide-band audio arrangement. First, OFDM provides a powerful framework for exploiting both the time- and frequency-domain channel properties by adapting the bit-allocation to subcarriers, as we will demonstrate below. Second, OFDM is amenable to powerful soft-decision based turbo coding [24], [26], [46]. Third, although our adaptive transceiver requires a programmable-rate speech or audio codec, to date only a limited number of such codecs has been proposed in the literature. Specific examples are the lower-quality 4 kHz bandwidth—i.e., narrow-band—advanced multirate (AMR) speech codec [36], which was designed for UMTS, and the higher-quality 7 kHz bandwidth PictureTel codec, which is the potential candidate for the new ITU-T G.722.1 audio-coding standard [35]. The recommended bit rates are 24 kb/s and 32 kb/s, but the codec can also be programmed for other bit rates, which we will capitalize on in the proposed system.

We will show that the AOFDM bit rate can be adaptively controlled in an effort to find the best compromise in terms of loading the AOFDM subcarriers more heavily, and hence to increase the available throughput bit rate for maintaining a higher speech coding rate and higher speech quality, while maintaining also a high robustness against transmission errors. A further tradeoff is that although the more heavily loaded, higher-throughput AOFDM modem is more vulnerable against transmission errors due to using more corrupted subcarriers, the longer turbo interleaving improves the turbo codec's performance.

The proposed AOFDM system is constituted by two adaptation loops, namely an inner constant throughput transmission regime, and an outer switching control regime, which jointly maintain the required target bit rate of the system, while employing a set of distinct operating modes. This system was contrived, in order to highlight the system design aspects of joint burst-by-burst adaptive modulation, channel coding, and source coding. The paper is structured as follows. Section II provides a brief system overview, listing also our experimental conditions, which is followed by Section III, detailing the philosophy of our constant throughput burst-by-burst adaptive OFDM modem. Section IV investigates the multimode modem adapta-

tion regime proposed, leading to a discussion on the adaptive audio source codec employed in Section V. Our system performance results are summarized in Section VI along with our future research endeavors.

II. SYSTEM OVERVIEW

The structure of the proposed adaptive OFDM transceiver is depicted in Fig. 1. The top half of the diagram is the transmitter section, which consists of the source and channel encoders, a channel interleaver for decorrelating the channel's frequency-domain fading, an adaptive modulator, a multiplexer adding signalling information to the transmitted data, and an inverse fast Fourier transform/radio frequency (IFFT/RF) OFDM stage. The receiver section, seen at the lower half of the graph, consists of an RF/FFT OFDM receiver, a demultiplexer extracting the signalling information, an adaptive demodulator, a deinterleaver/channel decoder, and the source decoder. The parameter adaptation linking the receiver and transmitter chain of the transceiver consists of a channel quality estimator and the mode selection, as well as the modulation adaptation blocks.

The open-loop control structure of the adaptation algorithms can be observed in the figure, where the local receiver's operation is controlled by the signalling information that is contained in the received OFDM symbol, while the channel quality information estimated by the local receiver is employed, in order to determine the parameter set to be employed by the local transmitter for the next timeslot on the reverse link. This requires channel reciprocity. If the uplink and downlink channel's frequency domain transfer functions are different, the system has to invoke closed-loop control. Explicitly, instead of the modem modes of the received OFDM subcarriers, the OFDM symbol has to carry the receiver's perception of the required modem modes to be used by the remote transmitter in order to maintain the target BPS of the receiver. In this case, however, the channel quality estimates will be associated with twice as high an estimation latency, as in case of the above open-loop control regime. The two distinct adaptation loops distinguished by the dotted and dashed lines are the inner and outer adaptation regimes, respectively. The outer adaptation loop controls the overall throughput of the system, so that a fixed-delay decoding of the received data packets becomes possible. This controls the

TABLE I
OFDM PARAMETERS OF THE PROPOSED ADAPTIVE SYSTEM

OFDM FFT Length	1024
Active subcarriers	928
Guard interval length	64 samples
Sampling rate	3.78 MHz
TDD frame duration	4.615 ms
TDD slot duration	$288\mu s$

packet size of the channel codec, the block length of the channel encoder and interleaver, as well as the target throughput of the inner adaptation loop. The operation of the AOFDM modulator, controlled by the inner loop, is transparent to the rest of the system. The operation of the adaptation loops is described in more detail below.

A. System Parameters

The transmission parameters have been adopted from the TDD-mode of the Pan-European UMTS system [31], having a carrier frequency of 1.9 GHz and a TDD-frame and time slot duration of 4.615 ms and 288 μs, respectively. In the envisaged slot-by-slot TDD mode, each uplink (UL) time slot is followed by a downlink (DL) time slot and vice versa. The sampling rate is assumed to be 3.78 MHz, leading to a 1024-subcarrier OFDM symbol, having a cyclic extension of 64 samples in each time slot. In order to assist in the spectral shaping of the OFDM signal, there are a total of 96 virtual subcarriers at the bandwidth boundaries. Table I gives an overview of the transmission parameters employed for this system.

The 7-kHz bandwidth PictureTel audio codec, which is the potential candidate for the new ITU-T G.722.1 audio coding standard [35], has been chosen for this system due to its good audio quality, robustness to packet dropping, and adjustable bit rate, which will be discussed in more depth below. A drawback of this audio codec is its relatively high overall encoding/decoding delay which totals about 60 ms. Hence, our objective was to minimize the delay inflicted by the other system components. As a result, we decided to assign each audio frame to a single AOFDM time slot.

The channel encoder/interleaver combination is a convolutional constituent coding based turbo encoder [42] employing block interleavers with a subsequent pseudorandom channel interleaver of a dimension identical to the number of OFDM subcarriers. The constituent recursive systematic convolutional (RSC) encoders are of constraint length 3, with octally represented generator polynomials of (7,5) and eight iterations are performed at the decoder, utilizing the so-called maximum *a posteriori* (MAP) algorithm [43] and the so-called log-likelihood ratio (LLR) soft inputs provided by the demodulator.

The channel model consists of a four path COST 207 typical urban channel impulse response (CIR) [44], where each impulse is subjected to independent Rayleigh fading having a normalized Doppler frequency of $2.25 \cdot 10^{-6}$, corresponding to a pedestrian scenario with a walking speed of 3 mph or, equivalently, 4.8 kph. The impact of a higher speed of motion on the

system performance will be alluded to at the end of Section IV. The unfaded CIR and the amplitude of the corresponding frequency domain channel transfer function are shown in Fig. 2. The gray shaded areas in Fig. 2(b) represent the virtual subcarriers, where it should be noted that according to the conventions of the N-point discrete Fourier transform (DFT) the subcarrier index 0 denotes the center frequency and increasing positive frequencies are found at indices $1 \leq i \leq N/2$, whereas negative frequencies are associated with indices $N/2 + 1 \leq i \leq N - 1$, and N represents the DFT length.

III. CONSTANT THROUGHPUT ADAPTIVE MODULATION

The constant throughput AOFDM algorithm attempts to allocate the required number of bits for transmission to the specific OFDM subcarriers exhibiting a low BER due to their unattenuated spectral envelope in Fig. 2(b), while the use of high BER subcarriers is minimized. We assume an open-loop adaptive system, basing the decision concerning the next transmitted OFDM symbol's modulation scheme allocation on the channel transfer function estimate gained at the reception of the most recent OFDM symbol by the receiver. Subband adaptive modulation [1], [10] where the modulation scheme is adapted not on a subcarrier-by-subcarrier basis but for blocks of adjacent subcarriers, is employed in order to simplify the AOFDM modem mode signalling requirements.

If the CIR $h(t, \tau)$ varies slowly compared to the OFDM symbol duration, then the Fourier transform $H(t, f)$ of the CIR can be defined, and the data symbols transmitted in the subcarriers $n \in [0, \cdots, N]$ are exposed to the frequency-domain fading determined by the instantaneous channel transfer function $H(t, n \cdot \Delta f) = H_n$.

The allocation of bits to subcarriers is based on the estimated frequency domain channel transfer function $\hat{H}_n$. On the basis of this and the overall signal-to-noise ratio (SNR) γ, the local SNR of each subcarrier n can be calculated as $\gamma_n = \gamma \cdot |\hat{H}_n|^2$. The predicted BER $p_e(\gamma_n, m)$ of each subcarrier n and for each of the possible modulation modes $m \in [0, \cdots, M]$ can now be computed and summed over the N_j subcarriers in subband j, in order to determine the expected number of bit errors for each subband and for each modulation mode, which is given by

$$e(j, m) = b(j, m) \sum_i p_e(\gamma_i, m)$$

for all subcarrier indices i in subband j. In our case, four modulation modes are employed for $m = 0, \cdots, 3$, which are "no transmission," BPSK, QPSK, and 16-QAM, respectively. Clearly, $p_e(j, 0) = 0$, and the other bit error probabilities can be evaluated using the Gaussian Q-function [45]. The number of bits transmitted in subband j, when using modulation mode m, is denoted by $b(j, m)$.

The bit allocation regime operates iteratively, allocating bits to subcarriers by choosing that specific subcarriers for transmitting the next bit to be assigned for transmission, which increases the system's BER by the smallest amount. In other words, the bits to be transmitted are allocated consecutively, commencing by assigning bits to the highest channel quality subcarriers, gradually involving the lower channel quality carriers. This strategy

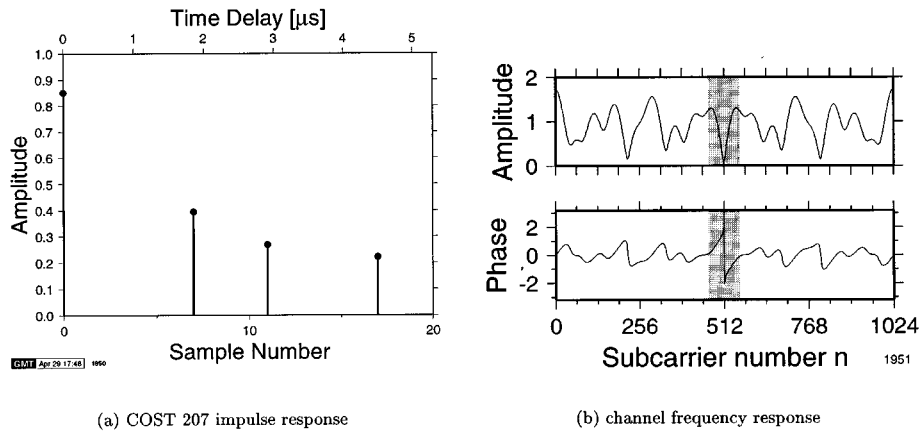(a) COST 207 impulse response   (b) channel frequency response

Fig. 2. Channel model: (a) COST 207 channel impulse response (CIR) and (b) unfaded frequency domain channel transfer function $H(n)$. The grey shaded area represents the virtual subcarriers.

is known as the so-called "greedy" approach, which is also the basis for the often cited Hughes-Hartogs algorithm [53].

More explicitly, for each subband, a state variable s_j is initialized to 0, and then the subband index j, for which the differential BER increment $(e_{s_j+1} - e_{s_j})/(b_{s_j+1} - b_{s_j})$ due to assigning the next bit to be transmitted is the lowest is found. The state variable s_j is incremented from 0, if it is not yet set to the index of the highest order modulation mode, i.e., to 16 QAM. This search for the lowest BER "cost" or BER penalty, when allocating additional bits is repeated, until the total number of bits allocated to the current OFDM symbol is equal to or higher than the target number of bits to be transmitted. Clearly, the higher the target number of bits to be transmitted by each OFDM symbol, the higher the BER, since gradually lower and lower channel quality subcarriers have to be involved.

The transmitter modulates the subcarriers using the specific modulation modes indexed by the state variables s_j, eventually padding the data with dummy bits, in order to maintain the required constant data throughput. The specific modulation modes chosen for the different subbands have to be signalled to the receiver for demodulation. By contrast, in a closed-loop controlled regime the receiver signals to the remote transmitter its requested AOFDM modes, and hence it is aware of the modes used. Alternatively, blind subband modem mode detection algorithms can be employed at the receiver [46]. For the scope of this paper, we assume 32 subbands of 32 subcarriers in each 1024-subcarrier OFDM symbol. Perfect channel estimation and subband modem mode signalling were assumed.

A. Performance

Fig. 3 shows an example of the fixed throughput AOFDM scheme's performance under the channel conditions characterized above, for a block length of 578 coded bits. As a comparison, the performance of a fixed BPSK modem transmitting the same number of bits in the same channel, employing 578 out of 1024 subcarriers, is depicted. The number of bits per AOFDM symbol is based on a 200-bit/20 ms useful data throughput, which corresponds to a 10 kb/s data rate, padded with 89 bits which can contain a check-sum for error detection and high-level AOFDM mode signalling information, as well as half-rate channel coding. Specifically, the signalling of the AOFDM mode assignment and
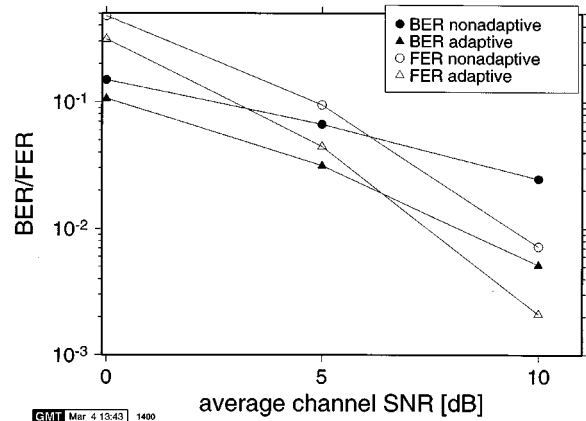

Fig. 3. FER and uncoded BER for fixed throughput adaptive and nonadaptive modulation in the fading time dispersive channel of Fig. 2 for a block length of 578 coded bits per 1024 subcarriers for the system of Table I.

the bit rate mode in the context of the Switch-I and Switch-II arrangements to be detailed in Section IV-A requires a strong error protection invoking a code associated with a low coding rate. In systems capitalizing on adaptive modulation, a mismatch between the actual modulation mode and the mode employed for demodulation is likely to severely increase the BER. Furthermore, an erroneous bit rate selection at the source decoder will result in a severe distortion of the reconstructed audio signal. The BER plotted in the figure is the hard-decision BER at the receiver before channel decoding. It can be seen that the adaptive modulation yields a significantly improved performance, which is reflected also in the AOFDM symbol error rate (SER) or, synonymously, the frame error rate (FER). This FER is the probability of a decoded AOFDM symbol containing errors, in which case it is unusable for the source decoder and hence it is dropped. This error event can be detected by using the check-sum of the AOFDM data symbol.

The modulation scheme allocation for the 578-bit AOFDM modem for an average channel SNR of 5 dB is given in Fig. 4(a) for 100 consecutive AOFDM symbols. The unused subbands with indices 15 and 16 contain the virtual carriers, and therefore do not transmit any data. It can be seen that the adaptation algorithm allocates data to the high-quality subcarriers on
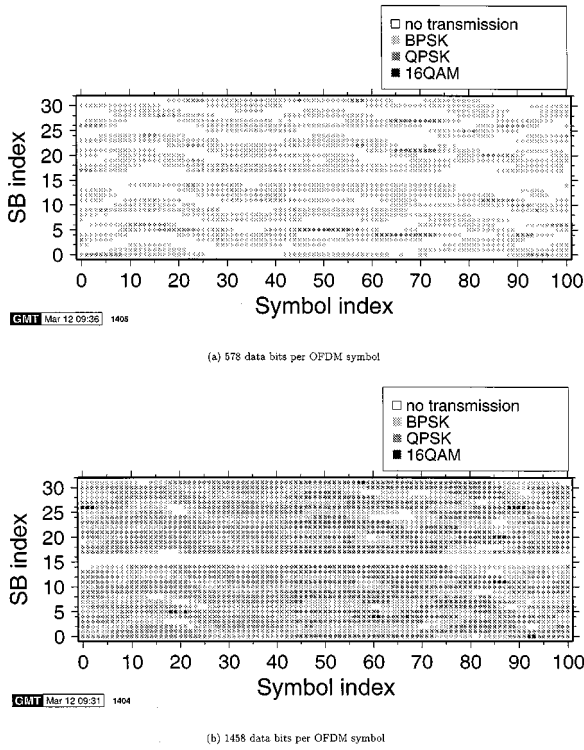
Fig. 4. Overview of modulation scheme allocation for the 578-bit fixed throughput adaptive modem over the fading time-dispersive channel of Fig. 2 at 5 dB average channel SNR.

an AOFDM symbol-by-symbol basis, while keeping the total number of bits per AOFDM symbol constant. As a comparison, Fig. 4(b) shows the equivalent overview of the modulation modes employed for the fixed bit rate of 1458 bits per AOFDM symbol. It can be seen that in order to hit the throughput target, hardly any subbands are in "no transmission" mode, and overall higher order modulation modes have to be employed.

Fig. 5 shows the subcarrier SNR for the first transmitted frame over the channel of Fig. 2 for a long-term SNR of 5 dB. It can be seen that the subcarrier SNR experienced by the modem varies significantly both across the overall AOFDM bandwidth, as well as within the subbands, delineated by the dotted vertical lines. The different-shade gray markers at the bottom of the graph indicate the modem mode employed for each subband, and the circular markers indicate the expected BER, averaged over the subcarriers of each subband. Fig. 5(a) gives the modem mode allocation and BER for the 10 kb/s mode, corresponding to the first column of Fig. 4(a), while Fig. 5(b) depicts the same information for the 32 kb/s mode, which corresponds to the first column of Fig. 4(b). Although in Fig. 5 the average estimated subband BER before turbo decoding seems to be unacceptably high, the powerful turbo decoding results in a sufficiently low FER and most importantly a breaking up of consecutive frame errors, which are especially detrimental for the audio signal reconstruction process in the source decoder.

IV. MULTIMODE ADAPTATION

While the fixed throughput adaptive algorithm described above copes well with the frequency-domain fading of the channel, there is also a medium-term time-domain variation of

the overall channel capacity. Hence, in addition to the previously proposed fixed-rate frequency-domain bit-allocation scheme, in this section we propose the employment of a time-variant bit rate scheme, in order to gauge its additional performance potential benchmarked against the fixed-rate schemes. We will then also contrive appropriate matching audio transceivers at a later stage that are capable of delivering the required number of audio bits. However, our experience demonstrated that it was an arduous task to employ powerful block-based turbo channel coding schemes in conjunction with variable throughput adaptive schemes for real-time applications, such as voice or video telephony. Nonetheless, a multimode adaptive system can be designed that allows us to switch between a set of different source—and channel coders as well as transmission parameters, depending on the overall instantaneous channel quality. We have investigated the employment of the estimated overall BER at the output of the receiver, which is determined on the basis of the sum of all the $e(j, s_j)$ subband error contributions after modem mode adaptation. On the basis of this expected input error rate of the channel decoder, the FER must be estimated, and compared to the expected FER of the other modem modes. Then, the mode having the lowest FER is selected; and the source encoder, the channel encoder, and the adaptive modem are set up accordingly.

We have defined four different operating modes, which correspond to the uncoded audio data rates of 10, 16, 24, and 32 kb/s at the source encoder's output. In conjunction with half-rate channel coding and allowing for check-sum and signalling overheads, the number of transmitted coded bits per AOFDM symbol is 578, 722, 1058, and 1458 for the four source-coded modes, respectively.
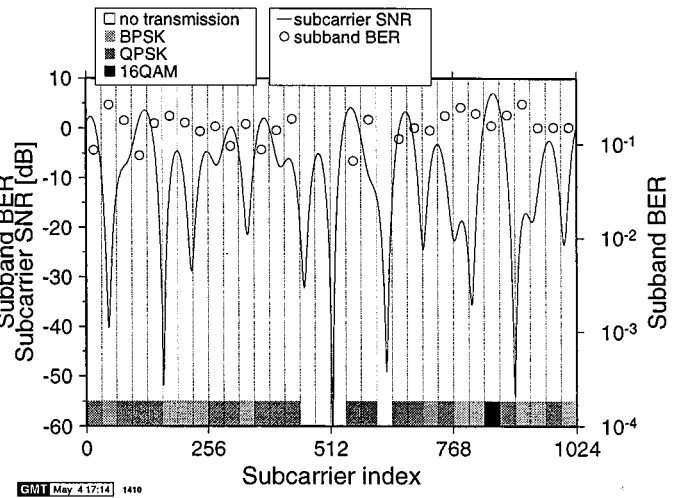
A. Mode Switching

Fig. 6 shows the observed FER for all four modes versus the uncoded BER over the channel of Fig. 2(a) that was predicted at the transmitter during the channel estimation and modem mode adaptation. The predicted BER was discretized into intervals of 1%, and the FER was averaged over these intervals. It can be seen that for estimated BER values below 5%, no frame errors were observed for any of the modes. Interestingly, for higher estimated BER values, the higher throughput modes exhibited a lower FER than the lower throughput modes, which was consistent with the turbo coder's performance increase for longer block lengths. An FER of 1% was observed for a 7% predicted input BER for the 10 kb/s mode, while BERs of 8% to 9% were allowed for the longer blocks.

In this paper, we assumed the best-case scenario of using the measured FER statistics of Fig. 6 for the mode switching algorithm. In this case, the FER corresponding to the predicted overall BER values for the different modes are compared, and the coded AOFDM mode with the lowest FER is chosen for transmission. The mode switching sequence for the first 500 AOFDM symbols at 5 dB channel SNR over the channel of Fig. 2(a) is depicted in Fig. 7. It can be seen that in this segment of the sequence, 32 kb/s transmission is the most frequently employed mode, followed by the 10 kb/s mode. The intermediate modes are mostly transitory, as the improving or deteriorating channel conditions render switching between the 10 and 32 kb/s modes necessary. This behavior is consistent
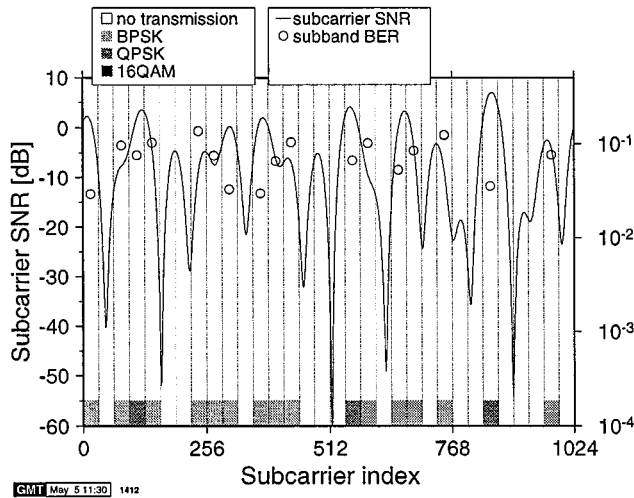
(a) 578 data bits per OFDM symbol   (b) 1458 data bits per OFDM symbol

Fig. 5. Subcarrier SNR versus subcarrier index for the first transmitted frame in the channel of Fig. 2 for a long-term SNR of 5 dB, with selected modem mode and average estimated subband BER before turbo-decoding for the 32 subbands. The two subbands around carrier 512 contain virtual carriers. (a) 10 kb/s mode and (b) 32 kb/s mode.
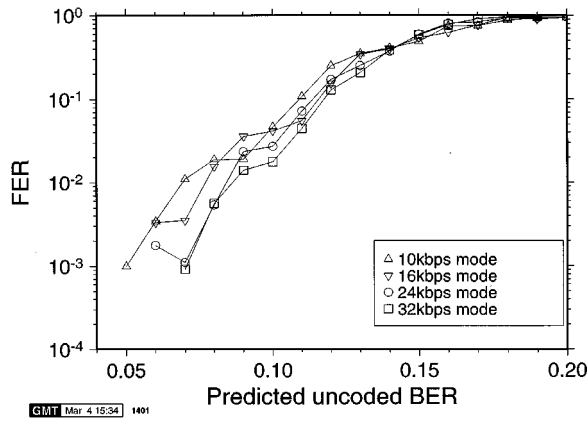

Fig. 6. Frame error rate versus the predicted uncoded BER for 10, 16, 24, and 32 kb/s modes.
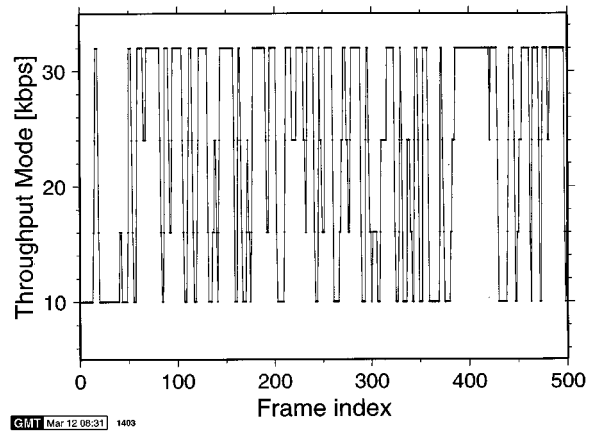

Fig. 7. Mode switching pattern at 5 dB channel SNR over the COST 207 channel of Fig. 2(b).

with the relative frequencies (Rel. Fr.) experienced and logged in Table I, for the system configuration denoted by "Switch-I," where the four available throughput modes—namely, 10, 16, 24, and 32 kb/s—have been employed in the context of the bit rate selection scheme outlined in Section IV. By contrast, the "Switch-II" configuration disabled the 10 kb/s mode. The impact of this on the system's performance, specifically on the quality of the reconstructed audio signal, will be discussed in Section VI. Before we embark on a brief portrayal of the 7 kHz bandwidth audio codec, let us briefly comment on the impact of an increasing speed of motion on the system's performance. Specifically, the degree of correlation between the channel transfer function estimated at the receiver—which is the basis for the modulation mode assignment to be employed by the local transmitter in the next time slot on the reverse link—and the channel transfer function actually incurred by an AOFDM symbol during transmission on the reverse link decreases. This reduces the performance gain of adaptive modulation.

V. THE PICTURETEL CODEC

In recent years, speech coding research has been focused on coding 7 kHz bandwidth, rather than 3.4 kHz bandwidth, speech signals in an effort to increase the perceived speech quality [37]–[40]. The challenge in this context has been the encoding of the speech spectral components above 3.4 kHz, which on average account for less than 1% of the speech energy, yet they substantially influence the perceived speech quality. A plausible approach is to separate these two bands using subband coding, which allows the designer to independently control the number of bits allocated to them. A more refined approach is to invoke frequency-domain coding techniques, such as transform coding [35], [40] which allows a more intelligent, finely grained distribution of the available coding bits to the most important audio signal frequencies. Furthermore, the bit rate can be adaptively controlled in an effort to find the best compromise in terms of loading the AOFDM subcarriers more

TABLE II
FER AND RELATIVE FREQUENCY OF DIFFERENT BITRATES IN THE FIXED BIT RATE- AND IN THE BURST-BY-BURST SWITCHING
SCHEMES (SUCCESSFULLY TRANSMITTED FRAMES) FOR AN SNR OF 5 dB

Scheme	FER[%]	Rel.Fr.:10 kbps [%]	Rel.Fr.:16 kbps [%]	Rel.Fr.:24 kbps [%]	Rel.Fr.:32 kbps [%]
Fixed-10 kbps	4.45	95.55	0.0	0.0	0.0
Fixed-16 kbps	5.58	0.0	94.42	0.0	0.0
Fixed-24 kbps	10.28	0.0	0.0	89.72	0.0
Fixed-32 kbps	18.65	0.0	0.0	0.0	81.35
Switch-I	4.44	21.87	13.90	11.59	48.20
Switch-II	5.58	0.0	34.63	11.59	48.20

heavily in order to increase the available bit rate for maintaining a higher speech coding rate and higher speech quality, while maintaining also a high robustness against transmission errors.

The current 32 kb/s G.722 ITU-T standard wide-band speech codec [41] is becoming antiquated, and the PictureTel transform coder (PTC) is the potential candidate for the new ITU-T G.722.1 wide-band audio coding standard [35]. It is based on the so-called modulated lapped transform (MLT) [50], followed by a quantization stage using a perceptually motivated psychoacoustic quantization model and Huffman coding for encoding the residual frequency domain coefficients.

At its input, the PTC expects frames of 320 pulse code modulated (PCM) audio samples, obtained by sampling the audio signal at a frequency of 16 kHz with a quantizer resolution of 14, 15, or 16 bits. Furthermore, the input samples are assumed to contain frequency components up to 7 kHz. The PictureTel standard proposal recommends operating the codec at output bit rates of 24 or 32 kb/s, generating output frame lengths of 480 or 640 bits per 20 ms, respectively, for which the codec was optimized. Due to its programmability, other bit rates are also feasible, a feature which we capitalized on in the context of the 10 and 16 kb/s system modes. The total delay encountered by an audio frame, when passing through the codec (consisting of encoder and decoder), is estimated to be on the order of about 60 ms, which is a result of the time domain frame overlapping technique associated with the MLT and the computational delay inherent in the codec.

Since the PTC employs Huffman coding for encoding the frequency domain coefficients, the decoding is very sensitive to bit errors. Hence, a single bit error can render the whole audio frame undecodable. The PTC's standard reaction to such a frame error is simply to repeat the previous frame of MLT coefficients, as long as this event occurs relatively rarely. For consecutive frame errors, the output signal is gradually muted after decoding the first erroneous frame. For further details concerning the PictureTel Transform Codec, the interested reader is referred to [35].

VI. SIMULATION RESULTS

Our discussions related to the associated system design tradeoffs and the impact of an automatic bit rate selection scheme on the audio quality of the system will be mainly based on measurements performed around channel SNR values of 5 dB over the channel model of Fig. 2(a), since for very low SNRs of around 0

dB, the frame dropping rate is excessive, yielding an unacceptable audio quality. By contrast, for high channel SNRs around 10 dB, the FER is too low to enable us to illustrate the tradeoffs between the audio quality and FER effectively. An estimate of the average quality of the reconstructed audio signal is provided by audio segmental SNR calculations, which provide an approximate measure of the subjectively perceived audio quality [52].

A. Frame Error Results

The basic tradeoff between the system's throughput and audio frame dropping-rate is illustrated for the 5 dB channel SNR scenario with the aid of our four fixed bit rate modes in Table II. The second column reflects for each bit rate the associated frame dropping rate that we will encounter.

As expected, by increasing the required AOFDM throughput bit rate, the FER will also increase, since a high proportion of reduced-quality AOFDM subcarriers has to be used for conveying the increased number of audio bits, although the performance of the turbo channel codec improves due to its increased interleaver depth. Experiments have shown that an FER of around 5% in conjunction with the 16 kb/s fixed bit rate mode is still sufficiently low, in order to provide a perceptually acceptable audio quality. In the third to sixth columns of Table II, the relative frequency of encountering error-free audio frames for the different audio bit rates is portrayed. Observe furthermore in the table that the same performance figures were also summarized for the two variable bit rate transmission arrangements denoted by *Switch-I* and *Switch-II* which are based on the four fixed bit rate schemes. These arrangements invoked a system philosophy controlling the audio source codec and channel codec on a time-variant basis, in order to take the channel quality fluctuations into account, as will be highlighted below.

Specifically, both of our experimental switching regimes, namely Switch-I and Switch-II, employed the same switching algorithm, as described in Section IV-A, with the only difference being that Switch-I incorporated in addition to the two standard bit rates of 24 and 32 kb/s proposed by the PictureTel company a 10 and 16 kb/s mode, with the intention of lowering the FER further due to the more modest "loading" of the OFDM symbols. For these switching schemes, the 5 dB SNR-related results in Table II underline that, for example, in comparison to the 16 kb/s fixed-rate mode, the system throughput was substantially improved, conveying (11.59 + 48.20) % of the audio frames in the 24 and 32 kb/s mode, rather than in the 16 kb/s mode, while maintaining the same FER of 5.58%, as the

TABLE III
FER AND RELATIVE FREQUENCY OF DIFFERENT BIT RATES IN THE SWITCH-II SCHEME (SUCCESSFULLY TRANSMITTED FRAMES) FOR CHANNEL SNRs OF 0, 5, AND 10 dB

Ch. SNR [dB]	FER [%]	Rel.Fr.:10 kbps [%]	Rel.Fr.:16 kbps [%]	Rel.Fr.:24 kbps [%]	Rel.Fr.:32 kbps [%]
0	37.69	0.0	37.79	14.42	10.10
5	5.58	0.0	34.63	11.59	48.20
10	0.34	0.0	7.81	5.61	86.24

16 kb/s mode. Although exhibiting a slightly lower FER, the Switch-I arrangement was shown to produce an audio quality inferior to that of the Switch-II arrangement. This was due to the employment of the 10 kb/s bit rate mode in the Switch-I arrangement, which produced a relatively low subjective audio quality. In this context, it is interesting to see that although the Switch-I arrangement assigns about 22% percent of all frames to the 10 kb/s transmission mode, the FER was increased only by about 1.1%, when disabling this subjectively low-quality but error resilient mode in the Switch-II arrangement. This is an indication of the conservative decision regime of our bit rate selector. The relative frequency of invoking the different bit rates in conjunction with the Switch-II arrangement has been evaluated additionally for channel SNRs of 0 dB and 10 dB, which characterizes the operation of the bit rate selector once again. The associated results are presented in Table III, which become plausible in the light of our previous discussions.

B. Audio Segmental SNR Performance and Discussions

In addition to our previous results, Fig. 8 displays the cumulative density function (CDF) of the segmental SNR (SEGSNR) of consecutive 20 ms duration audio segments obtained from the reconstructed signal of an audio test signal at the output of the PTC decoder for the schemes described above. These CDF's were recorded at a channel SNR of 5 dB over the channel model of Fig. 2(a). The step-function like CDF discontinuity at an SEGSNR of 0 dB corresponds to the frame dropping rate of the associated transmission scheme, which was summarized in Table III for the various systems. As expected, for any given SEGSNR value, it is desirable to maintain as low a proportion of the audio frames' SEGSNRs below a given abscissa value as possible. Hence, we concluded that the best SEGSNR CDF was attributable to the Switch-II scheme, while the worst to the fixed 10 kb/s arrangement, as suggested before. In the range of high audio SEGSNRs, the preference order of the various fixed bit rate schemes followed our expectations, i.e., the fixed-rate 32 kb/s scheme performed best in SEGSNR terms, when neglecting frame drops. The tradeoff was that due to its high audio bit rate of 32 kb/s, the scheme exhibited the inherently highest SEGSNR. Moreover, due to its high throughput requirement, this scheme was forced to invoke a substantial proportion of partially impaired, low-quality AOFDM subcarriers, which often resulted in corrupted and dropped audio frames. Since the fixed 10 kb/s scheme exhibited the lowest audio SEGSNR performance, this scheme was excluded from the Switch-II arrangement. However, FERs in excess of 10% resulted in distinctively audible artifacts, which—despite their high error-free SEGSNRs—virtually
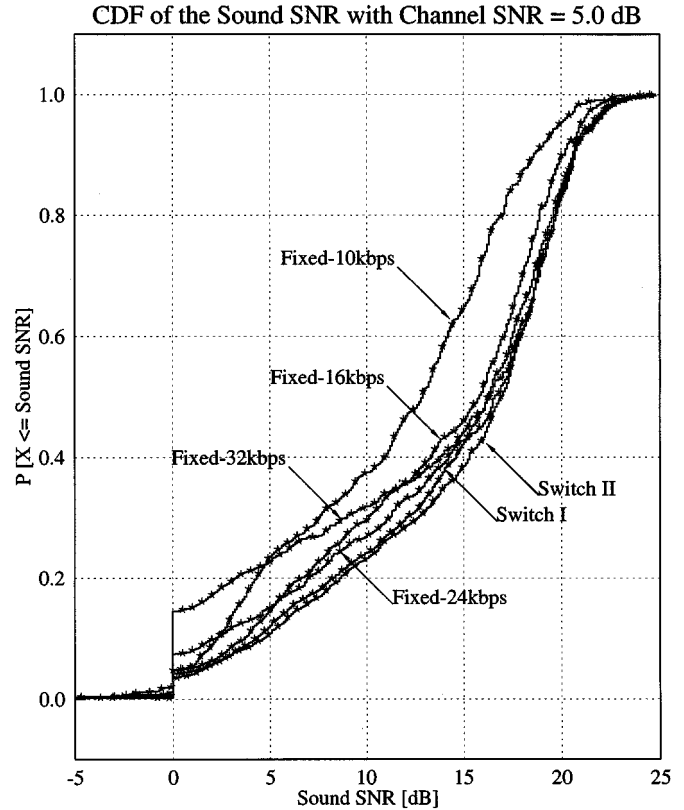

Fig. 8. Typical CDF of the segmental sound SNR of a reconstructed audio signal transmitted over the fading time dispersive channel.

rendered the fixed-rate 24 and 32 kb/s modes unacceptable. *Hence, our proposed switching scheme—Switch-II—which is based on the 16, 24, and 32 kb/s bit rates, achieved at medium SNRs the best compromise between average error-free audio quality and frame dropping rate, which has been verified by our informal listening tests. By allowing the system to temporarily reduce its throughput from 32 kb/s to 24 and 16 kb/s, we were able to avoid dropping corrupted 32 kbps coded audio frames, which was unavoidable in the context of the fixed-rate 32 kbps scheme.*

As outlined in Section IV-A, the mode switching algorithm operates on the basis of statistically evaluated experimental results for the prediction of the FER. A robust, channel-independent switching regime operating on the basis of the turbo coder's quality perceptions can overcome this dependence. Furthermore, a target-FER driven switching scheme instead of the minimal-FER algorithm employed for this series of experiments will be investigated in the future. Our further research will also consider the employment of a multiuser AOFDM audio system.

REFERENCES

[1] L. Hanzo, W. T. Webb, and T. Keller, *Single- and Multi-Carrier Quadrature Amplitude Modulation*, New York: Wiley, IEEE Press, 2000.

[2] W. Webb and R. Steele, "Variable rate QAM for mobile radio," *IEEE Trans. Commun.*, vol. 43, pp. 2223–2230, July 1995.

[3] A. J. Goldsmith and S. G. Chua, "Variable rate variable power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 1218–1230, Oct. 1997.

[4] S. Sampei, S. Komaki, and N. Morinaga, "Adaptive modulation/TDMA scheme for large capacity personal multimedia communications systems," *IEICE Trans. Commun.*, vol. E77-B, pp. 1096–1103, Sept. 1994.

[5] C. H. Wong, T. H. Liew, and L. Hanzo, "Blind modem mode detection aided block turbo coded burst-by-burst wideband adaptive modulation," in *Proc. ACTS'99*, Sorrento, Italy, June 8–11, 1999, pp. 303–108.

[6] M. S. Yee, T. H. Liew, and L. Hanzo, "Radial basis function decision feedback equalisation assisted block turbo coded burst-by-burst adaptive modems," in *Proc. VTC'99 Fall*, Amsterdam, The Netherlands, Sept. 19–22, 1999, pp. 1600–1604.

[7] H. Matsuako, S. Sampei, N. Morinaga, and Y. Kamio, "Adaptive modulation systems with variable coding rate concatenated code for high quality multi-media communication systems," in *Proc. IEEE Veh. Technol. Conf.*, 1996, pp. 487–491.

[8] V. K. N. Lau and M. D. Macleod, "Variable rate adaptive trellis coded QAM for high bandwidth efficiency applications in Rayleigh fading channels," in *Proc. IEEE Veh. Technol. Conf.*, 1998, pp. 348–352.

[9] S. G. Chua and A. Goldsmith, "Adaptive coded modulation for fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 595–602, May 1998.

[10] T. Keller and L. Hanzo, "Adaptive orthogonal frequency division multiplexing schemes," in *Proc. ACTS Summit*, Rhodes, Greece, June 1998, pp. 794–799.

[11] E. L. Kuan, C. H. Wong, and L. Hanzo, "Burst-by-burst adaptive joint detection CDMA," in *Proc. VTC'99*, Houston, TX, May 1999, pp. 1628–1632.

[12] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *Bell Syst. Tech. J.*, vol. 46, pp. 1775–1796, Dec. 1966.

[13] J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. COM-33, pp. 665–675, July 1985.

[14] K. Fazel and G. Fettweis, Eds., *Multi-Carrier Spread-Spectrum*. Norwell, MA: Kluwer, 1997, p. 260.

[15] F. Classen and H. Meyr, "Synchronization algorithms for an OFDM system for mobile communications," in *Codierung fuer Quelle, Kanal und Uebertragung*. ITG Fachbericht, Berlin: VDE-Verlag, 1994, pp. 105–113. no. 130.

[16] ——, "Frequency synchronization algorithms for OFDM systems suitable for communication over frequency selective fading channels," in *Proc. IEEE Veh. Technol. Conf.*, 1994, pp. 1655–1659.

[17] S. J. Shepherd, P. W. J. van Eetvelt, C. W. Wyatt-Millington, and S. K. Barton, "Simple coding scheme to reduce peak factor in QPSK multicarrier modulation," *Electron. Lett.*, vol. 31, pp. 1131–1132, July 1995.

[18] A. E. Jones, T. A. Wilkinson, and S. K. Barton, "Block coding scheme for reduction of peak to mean envelope power ratio of multicarrier transmission schemes," *Electron. Lett.*, vol. 30, pp. 2098–2099, 1994.

[19] S. H. Mueller and J. B. Huber, "Vergleich von OFDM-Verfahren mit reduzierter Spitzenleistung," *2. OFDM-Fachgespraech in Braunschweig*, 1997.

[20] T. May and H. Rohling, "Reduktion von Nachbarkanalstörungen in OFDM-Funkuebertragungssystemen," *2. OFDM-Fachgespraech in Braunschweig*, 1997.

[21] M. D. Benedetto and P. Mandarini, "An application of MMSE predistortion to OFDM systems," *IEEE Trans. Commun.*, vol. 44, pp. 1417–1420, Nov. 1996.

[22] I. Kalet, "The multitone channel," *IEEE Trans. Commun.*, vol. 37, pp. 119–124, Feb. 1989.

[23] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commun.*, vol. 48, pp. 772–775, 1995.

[24] T. Keller, J. Woodard, and L. Hanzo, "Turbo-coded parallel modem techniques for personal communications," in *Proc. 47th Veh. Technol. Conf.*. Phoenix, AZ, May 5–7, 1997, pp. 2158–2162.

[25] P. Cherriman, T. Keller, and L. Hanzo, "Constant-rate turbo-coded and block-coded orthogonal frequency division multiplex videophony over UMTS," in *Proc. Globecom'98*, Sydney, Australia, Nov. 1998, pp. 2848–2852.

[26] J. P. Woodard, T. Keller, and L. Hanzo, "Turbo-coded orthogonal frequency division multiplex transmission of 8 kbps encoded speech," in *Proc. ACTS'97*, Aalborg, Denmark, Oct. 1997, pp. 894–899.

[27] P. R. K. Fazel, S. Kaiser, and M. Ruf, "A concept of digital terrestrial television broadcasting," *Wireless Personal Commun.*, vol. 2, pp. 9–27, 1995.

[28] H. Sari, G. Karam, and I. Jean Claude, "Transmission techniques for digital terrestrial tv broadcasting," *IEEE Commun. Mag.*, pp. 100–109, Feb. 1995.

[29] J. Borowski, S. Zeisberg, J. Hübner, K. Koora, E. Bogenfeld, and B. Kull, "Performance of OFDM and comparable single carrier system in MEDIAN demonstrator 60GHz channel," in *Proc. ACTS Summit*, Aalborg, Denmark, Oct. 1997, pp. 653–658.

[30] B. Engstroem and C. Oesterberg, "A system for test of multiaccess meethods based on OFDM," in *Proc. VTC'94*, Stockholm, Sweden, pp. 1843–1847.

[31] E. Dahlman, B. Gudmundson, M. Nilsson, and J. Skold, "UMTS and IMT-2000 based on wideband CDMA," *IEEE Commun. Mag.*, vol. 36, pp. 70–80, Sept. 1998.

[32] Y. Li and N. R. Sollenberger, "Interference suppression in OFDM systems using adaptive antenna arrays," in *Proc. Globecom 98*, vol. 1, Sydney, Australia, Nov. 8–12, 1998, pp. 213–218.

[33] M. Münster, T. Keller, and L. Hanzo, "Co-channel interference suppression assisted adaptive OFDM in interference limited environments," in *Proc. IEEE Veh. Technol. Conf. Fall*, vol. 1, Amsterdam, Netherlands, Sept. 19–22, 1999, pp. 284–288.

[34] M. Münster and L. Hanzo, "Co-channel interference cancellation techniques for antenna array assisted multiuser OFDM systems," in *Proc. 3G-'2000 Conf.*, London, Mar. 27–29, 2000, pp. 256–260.

[35] , ftp://standard.pictel.com/sg16_q20/1999_09_Geneva/.

[36] GSM 06.90: Digital cellular telecommunications system (Phase 2+); Adaptive Multi-Rate (AMR) speech transcoding, version 7.0.0, Release, 1998.

[37] J. Paulus and J. Schnitzler, "16 kbit/s wideband speech coding based on unequal subbands," in *Proc. ICASSP*, 1996, pp. 255–258.

[38] J.-H. Chen and D. Wang, "Transform predictive coding of wideband speech signals," in *Proc. ICASSP*, 1996, pp. 275–278.

[39] A. Ubale and A. Gersho, "A multi-band CELP wideband speech coder," in *Proc. ICASSP*, 1997, pp. 1367–1370.

[40] P. Combescure, J. Schnitzler, K. Fischer, R. Kirchherr, C. Lamblin, A. Le Guyader, D. Massaloux, C. Quinquis, J. Stegmann, and P. Vary, "A 16, 24, 32 kbit/s wideband speech codec based on ATCELP," in *Proc. ICASSP*, 1999.

[41] CCITT Recommendation G.722, 7 kHz Audio Coding within 64 kbit/s, ITU-T.

[42] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. ICC*, 1993, pp. 1064–1070.

[43] L. Bahl *et al.*, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, pp. 284–287, 1974.

[44] M. Failli, "Digital land mobile radio communications COST 207," Tech. Rep., European Commission, 1989.

[45] J. G. Proakis, *Digital Communications*, 2nd ed. London: McGraw-Hill, 1989.

[46] T. Keller and L. Hanzo, "Blind-detection assisted sub-band adaptive turbo-coded OFDM schemes," in *Proc. IEEE Veh. Technol. Conf.*, Houston, TX, 1999, pp. 489–493.

[47] K. R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages and Applications*, London: Academic, 1990.

[48] H. S. Malvar and D. H. Staelin, "The LOT: Transform coding without blocking effects," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 553–559, Apr. 1989.

[49] H. S. Malvar, "Lapped transform for efficient transform/subband coding," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 969–978, June 1990.

[50] ——, *Signal Processing with Lapped Transforms*. Boston, MA: Artech House, 1992.

[51] S. Shlien, "The modulated lapped transform, Its time-varying forms, and its applications to audio coding standards," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 5, pp. 359–366, July 1997.

[52] N. Kitawaki, M. Honda, and K. Itoh, "Speech-quality assessment methods for speech-coding systems," *IEEE Commun. Mag.*, vol. 22, pp. 26–33, Oct. 1984.

[53] D. Hughes-Hartogs, "Ensemble modem structure for imperfect transmission media," U.S. Patents 4 679 222, July 1987; 4 731 816, Mar. 1988; and 4 833 796, May 1989.

Thomas Keller studied electrical engineering at the University of Karlsruhe, Ecole Superieure d'Ingenieurs en Electronique et Electrotechnique, Paris, and the University of Southampton. He graduated with a Dipl.-Ing. degree in 1995. Between 1995 and 1999 he had been with the Wireless Multimedia Communications Group at the University of Southampton, where he completed the Ph.D. degree in mobile communications. His current areas of interest include adaptive OFDM transmission, wideband channel estimation, and error correction coding. He recently joined Ubinetics, Cambridge, U.K.

Matthias Münster studied Electrical Engineering at Aachen University of Technology, Germany, where he graduated with a Dipl.-Ing. degree in 1998. Since then he has been with the Wireless Multimedia Communications Group at the University of Southampton, working toward a Ph.D. in mobile communications. His current areas of interest include channel estimation—as well as detection techniques for multi-user OFDM systems.

Lajos Hanzo (M'91–SM'92) graduated in electronics in 1976, and in 1983, he was conferred a Ph.D. During his 24-year career in telecommunications he has held various research and academic posts in Hungary, Germany and the U.K. Since 1986 he has been with the Department of Electronics and Computer Science, University of Southampton, U.K. and has been a Consultant to Multiple Access Communications Ltd., U.K. Currently he holds a Chair in telecommunications. He co-authored five books on mobile radio communications, published over 300 research papers, organized and chaired conference sessions, presented overview lectures, and was awarded a number of distinctions. Currently he is managing an academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored both by industry and under the auspices of the Engineering and Physical Sciences Research Council (EPSRC) U.K., the European IST Programme and the Mobile Virtual Centre of Excellence (VCE), U.K. For further information on research in progress and associated publications please refer to http://www-mobile.ecs.soton.ac.uk.