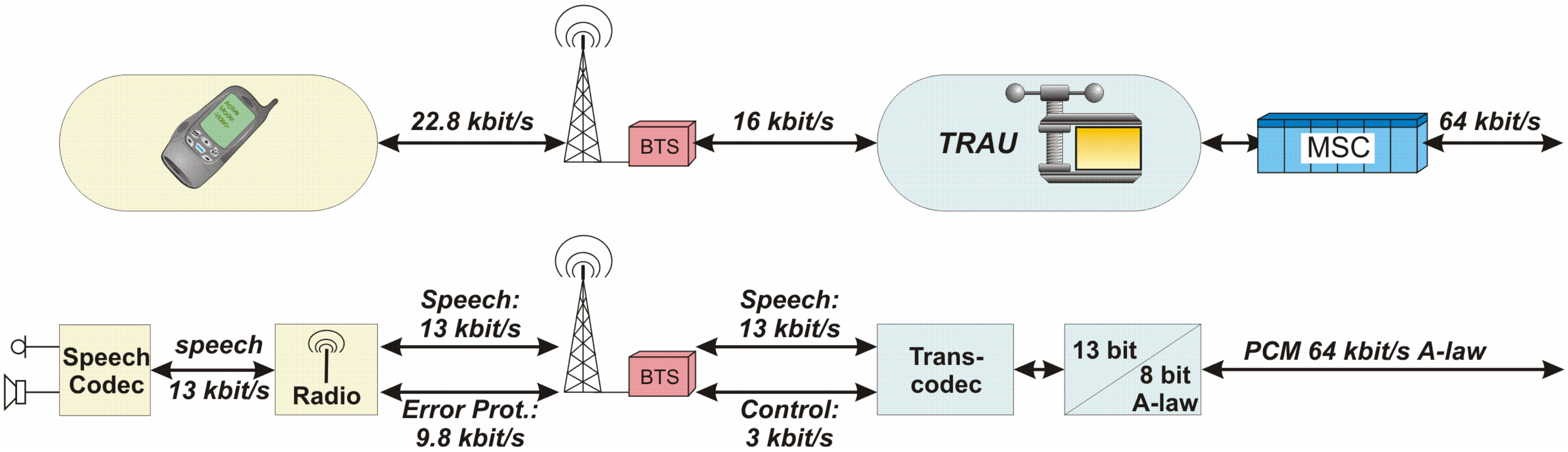
Speech Signal Transmission in GSM



Speech Signal Transmission in GSM

In a communication system service adaptations have to be performed according to the restrictions set by the transmission environment. In GSM like in any other radio communication system the scare frequency resource is the most limiting factor. The service signal bandwidth has to be adapted to the limited bandwidth on the air interface taking into account the modulation scheme and the requested robustness against transmission errors.

In a GSM traffic channel the available (gross) bit rate on the air interface is 22.8 kbit/s. This bit rate must be shared by the data signal and the necessary error protection add on as requested by the quality of service (QoS) class. On the other hand in the GSM core network (and in fixed networks too) the applied transmission is based on ISDN standards. The basic ISDN rate is 64 kbit/s.

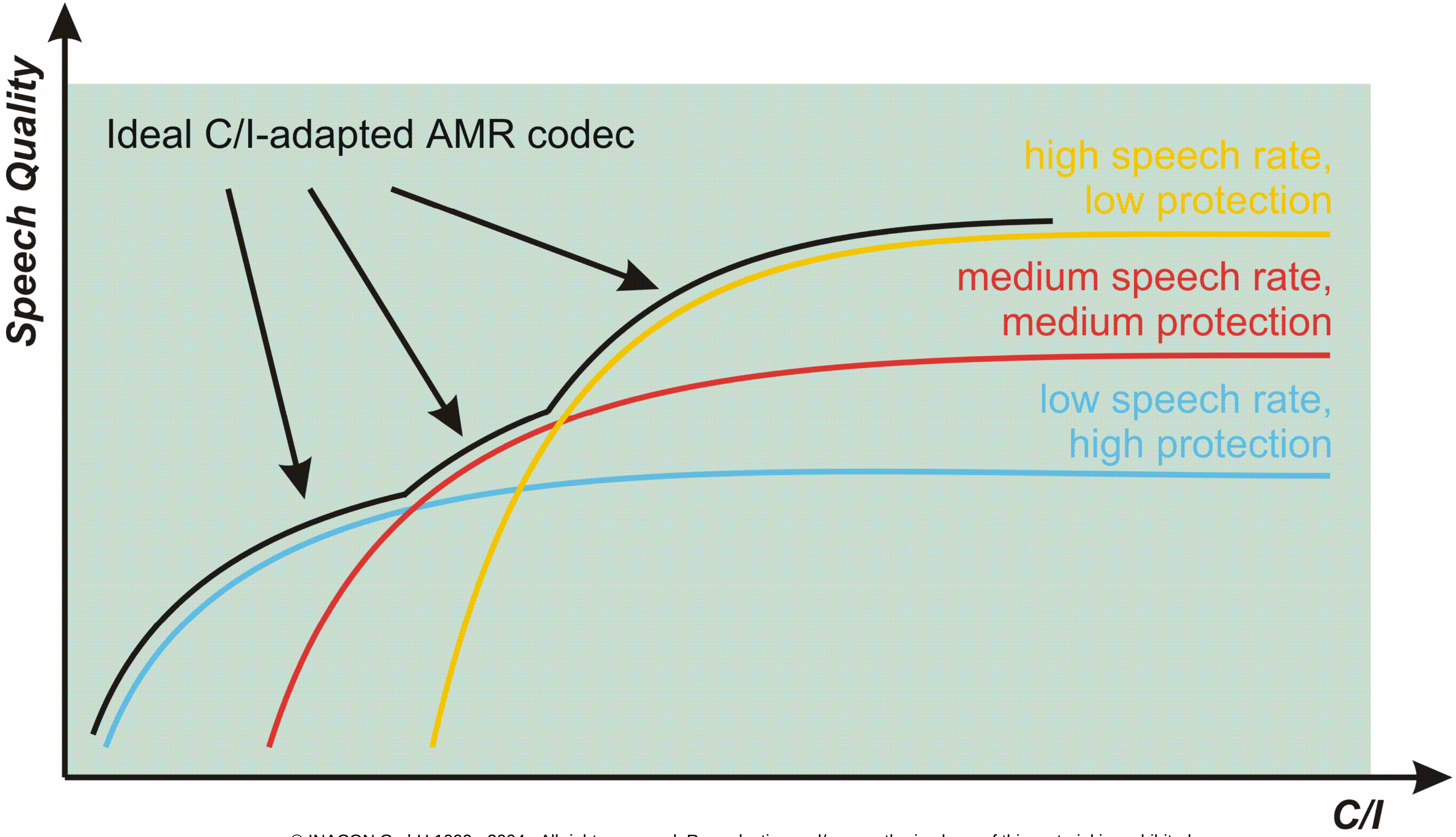
A speech signal in ISDN is coded by the ITU-T standard G.711 which encodes a 3.1 kHz speech signal – sampled with 8kHz and converted to 8 bit resolution (using the A- (or µ-) law companding scale). This results in a data rate of 64 kbit/s for a speech signal which matches the ISDN rate.

The codec function is realized in the mobile station and on the other side it is part of the TRAU unit (Transcoder Rate and Adaption Unit) which is normally implemented with the MSC (Mobile services Switching Center). The transcoder changes the A- (or µ-) law companded input signal into a 13 bit linear quantized signal. By applying redundancy and irrelevancy reducing coding techniques the output signal shows a bit rate of 13 kbit/s only. Together with control information this signal is transmitted inside the BSS (Base station SubSystem) using 16 kbit/s transmission links. The transmission frame on this interface is named TRAU frame. The base station adds specific Forward Error Correction (FEC) overhead to the speech data to cope with the critical radio conditions.

In the receiver after error correction and decoding the speech signal is delivered to the user.

[ITU-T G.711]

Channel Adaptive Speech Codec



Channel Adaptive Speech Coding

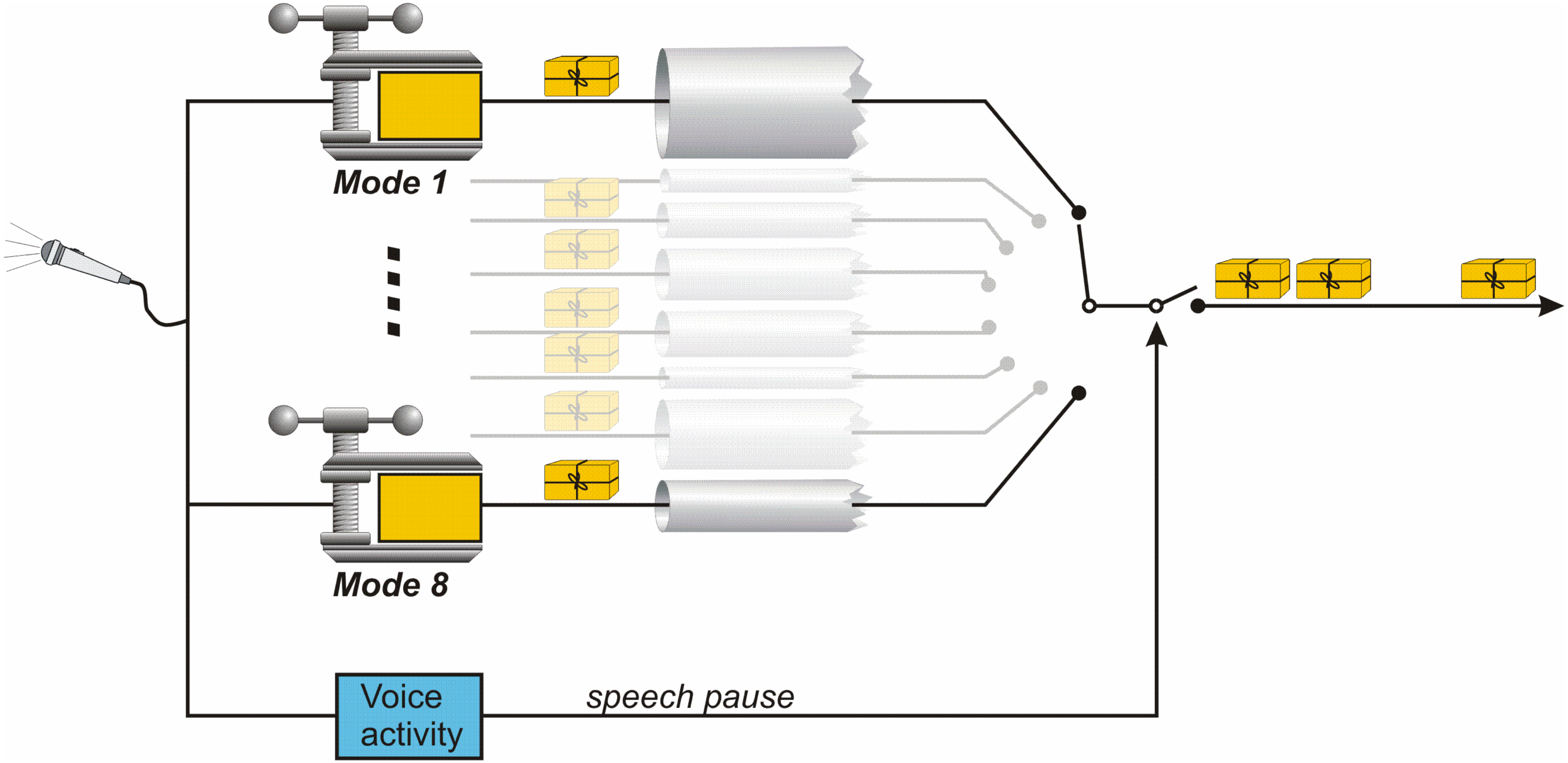
In digital mobile communication systems an almost error free operation can be achieved due to powerful error protection techniques against transmission errors. However when the carrier to interference ratio (C/I) drops below a certain threshold and too many errors occur, the protection mechanisms are no longer capable to cope with this high amount of errors. The residual errors will corrupt the decoding process and may lead to very annoying artifacts in the reconstructed speech signal. If there are too many errors and the BFI flag is set in 16 consecutive frames the connection is released.

In GSM the maximum bit rate of a traffic channel is 22.8 kbit/s. So for a full rate coder speech signal of 13 kbit/s rate additional 9.8 kbit/s are spent for error protection. For a lower rate speech signal the error protection effort can be increased up to the total of 22.8 kbit/s resulting into a higher protection level and more transmission errors can be corrected or - with other words – communication is still possible for lower C/I-values.

An adaptive codec may switch between different codec modes with different (e.g. decreasing) speech rates but also different (increasing) protection levels. In this way an adaptive codec switching between different codec modes can cope with changing radio conditions but still keeps the communication link.

It should be noted, that in addition, AMR offers the opportunity for rural coverage improvements and deeper in-building coverage because of the greater robustness of the full-rate channel. By building AMR into network build plans, operators can deliver capacity requirements with significantly less infrastructure, reducing capital investment and operating costs.

AMR Speech Codec



AMR Speech Codec

The Adaptive Multi Rate (AMR) speech coding scheme is a combination of new speech codec with adaptable output data rates and the discontinuous transmission scheme (DTX).

The AMR speech coder consists of the multi rate speech coder, a source controlled rate scheme including a voice activity detector and a comfort noise generation system, and an error concealment mechanism to combat the effects of transmission errors and lost packets.
The multi rate speech coder is a single integrated speech codec with eight source rates from 4.75 kbit/s to 12.2 kbit/s, and a low rate background noise encoding mode. The speech coder is capable (theoretically) of switching its bit-rate every 20 ms speech frame upon command.

During a normal telephone conversation, the participants alternate so that, on the average, each direction of transmission is occupied about 50% of the time. Discontinuous transmission is a mode of operation where the speech encoder encodes speech frames containing only background noise with a lower bit-rate than normally used for encoding speech. A network may adapt its transmission scheme to take advantage of the varying bit-rate. This may be done for the following two purposes:

⇒ In the MS, battery life will be prolonged or a smaller battery could be used for a given operational duration.
⇒ The average required bit-rate is reduced, leading to a more efficient transmission with decreased load and hence increased capacity.

The following functions are required for the source controlled rate operation:

⇒ a Voice Activity Detector (VAD) on the TX side;
⇒ evaluation of the background acoustic noise on the TX side, in order to transmit characteristic parameters to the RX side;
⇒ generation of comfort noise on the RX side during periods when no normal speech frames are received.

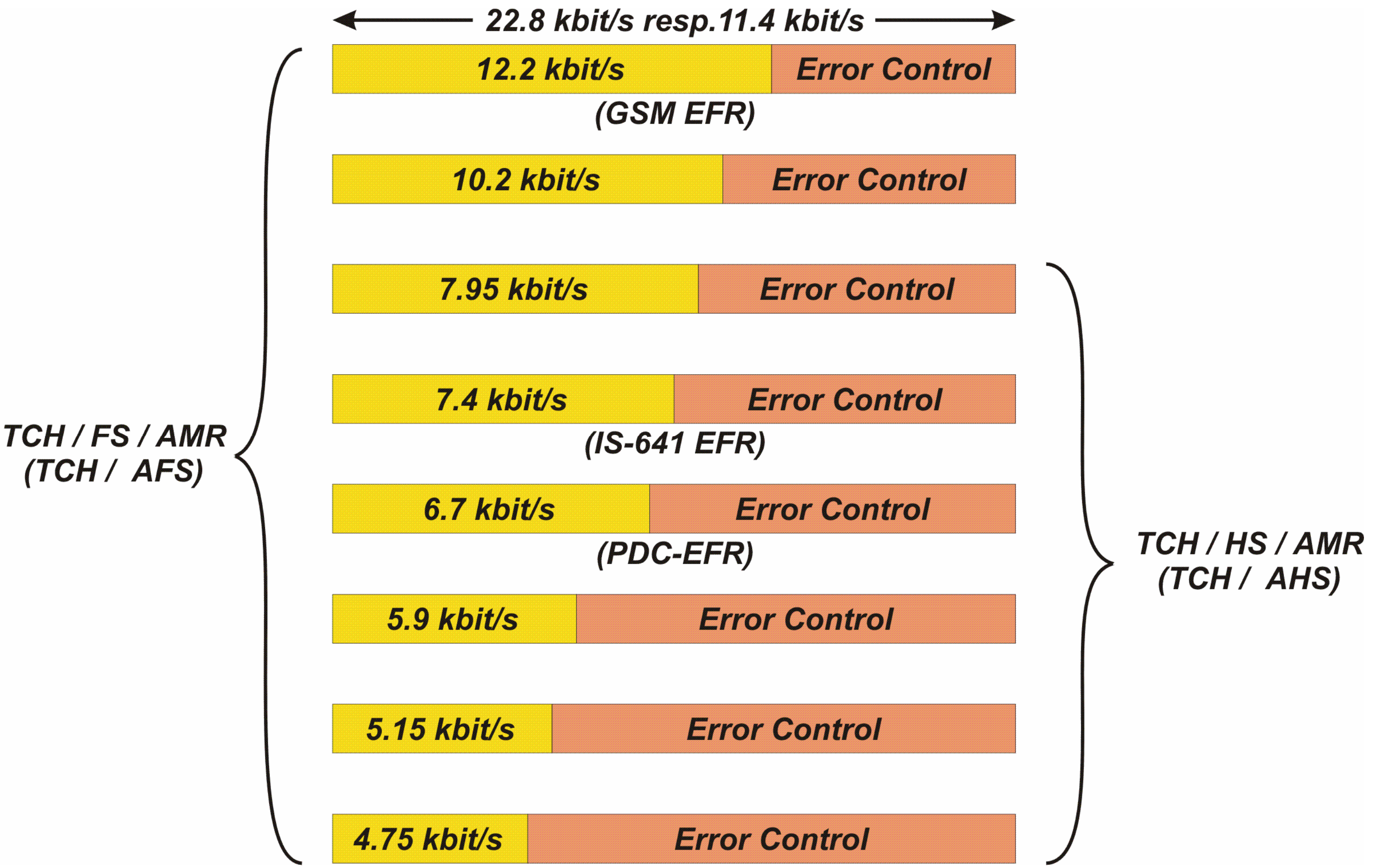
The transmission of comfort noise information to the RX side is achieved by means of a Silence Descriptor (SID) frame, which is sent at regular intervals.

AMR encoded speech signals may be transmitted using full- or half rate traffic channels.

Note: the AMR codec will not only be introduced in GSM, but also in UMTS.

[3GPP TS 26.071, 26.073]

AMR Coding Modes



AMR Coding Modes

The AMR codec offers 8 different source rates between 12.2 kbit/s and 4.75 kbit/s. The difference between the speech data rate and the GSM full rate channel of 22.8 kbit/s (respectively 11.4 kbit/s for the half rate channel) is used for error protection.
The 12.2 kbit/s mode complies with the Enhanced Full Rate codec of GSM. This mode offers near 64 kbit/s PCM quality. In the same way the 7.4 kbit/s mode is conform to the TIA/EIA IS-641 TDMA IS-136 Enhanced Full Rate Speech Codec (USA) and the 6.7 kbit/s mode complies to the ARIB 6.7 kbit/s Enhanced Full Rate Speech Codec (Japan).

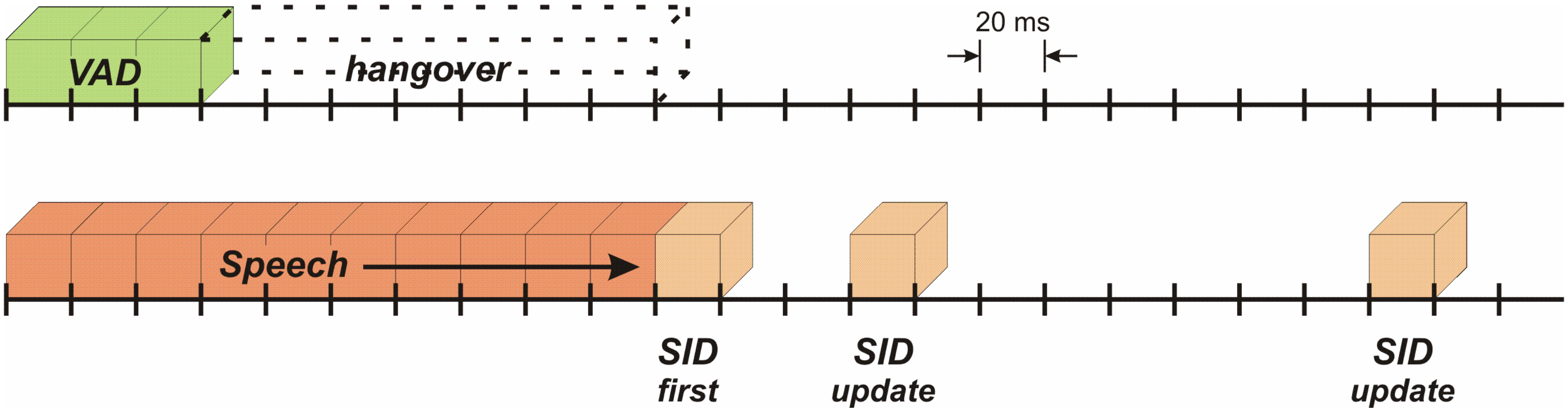
The Full Rate channel mode is directed for maximum robustness to channel errors. This additional robustness may be used to extend the coverage in marginal signal conditions, or to improve the capacity by using a tighter frequency re-use (assuming a high AMR MS penetration).
The Half Rate channel mode addresses maximum capacity. More than 100% capacity increase is expected relative to GSM Full Rate or EFR. Significant quality improvements relative to the existing Half Rate will be given for a large portion of mobiles as a result of the codec mode adaptation to the channel conditions and excellent (wire line like) speech quality in half rate mode for low error conditions.
Mixed Half/Full Rate channel mode allows a trade off between quality and capacity enhancements according to the radio and traffic conditions and operator priorities.

In Full Rate mode all eight codec modes are applicable, in Half Rate mode only a subset of the six lower rate codec modes are used. Not all codec modes must be offered at a time during one connection. A subset of up to four codec modes can be selected at call set up or handover.

TCH/AFS = Traffic CHannel / Adaptive Fullrate Speech
TCH/AHS = Traffic Channel / Adaptive Halfrate Speech

[TR 101.714 (4.3); 3GPP TR 26.975 (4.3)]

AMR Discontinuous Transmission (DTX)



AMR Discontinuous Transmission (DTX)

Discontinuous transmission (DTX) is a mechanism, which allows the radio transmitter to be switched off most of the time during speech pauses. There are two benefits out of this: power consumption is reduced in the MS resulting in longer operation time per battery load and the overall interference level over the air interface will be reduced.

Implementation of the DTX mode is mandatory in the MS and for the receiving path in the BSS. The network determines DTX operation in uplink direction. In downlink direction the MS shall handle DTX at any time, regardless, whether DTX in uplink is commanded or not.

With the Voice Activity Detector (VAD) transition from “1” to “0” a pause in the speech flow is detected. Because it needs eight consecutive frames to make a new update silence descriptor (SID) analysis available at receiver side a hangover period of seven frames is appended. During this period the data frames are still handled as “speech” frames (encoded and transmitted). After end of the hangover period a SID_FIRST frame is transmitted to the receiver indicating the begin of a speech pause. The first updated SID_UPDATE frame will follow as the third frame after SID_FIRST. The SID_UPDATE frame will then be repeated every 8th frame.
Whereas the SID_UPDATE frame always includes a new comfort noise parameter set, the SID_FIRST contains no information only an indication to mark the beginning of a speech pause.

When a SID_FIRST or SID_UPDATE is stolen by a FACCH or RATSCCH frame then the subsequent frame shall be scheduled for transmission for the stolen frame.

In case less then 24 speech frames have been transmitted since the last SID_UPDATE no hangover period is introduced but this last analysed SID_UPDATE frame shall repeatedly passed to the receiver whenever a SID_UPDATE frame is to be produced until a new updated SID analysis is available.

For the period between the SID_FIRST and the first SID_UPDATE frame the receiver will calculate the comfort noise parameters from the last seven speech frames.

[3GPP TS26.093 (Annex A)]; [3GPP TS 26.103 (5.4)]

AMR Modes Bit Rates

encoded bits / AFS		AMR modes		encoded bits / AHS		
class 1a	class 1b	kbit/s	bit / 20ms	class 1a	class 1b	class 2
81	163	12.2	244	-	-	-
65	139	10.2	204	-	-	-
75	84	7.95	159	67	56	36
61	87	7.4	148	61	59	28
55	79	6.7	134	55	55	24
55	63	5.9	118	55	47	16
49	54	5.15	103	49	42	12
39	56	4.75	95	39	44	12

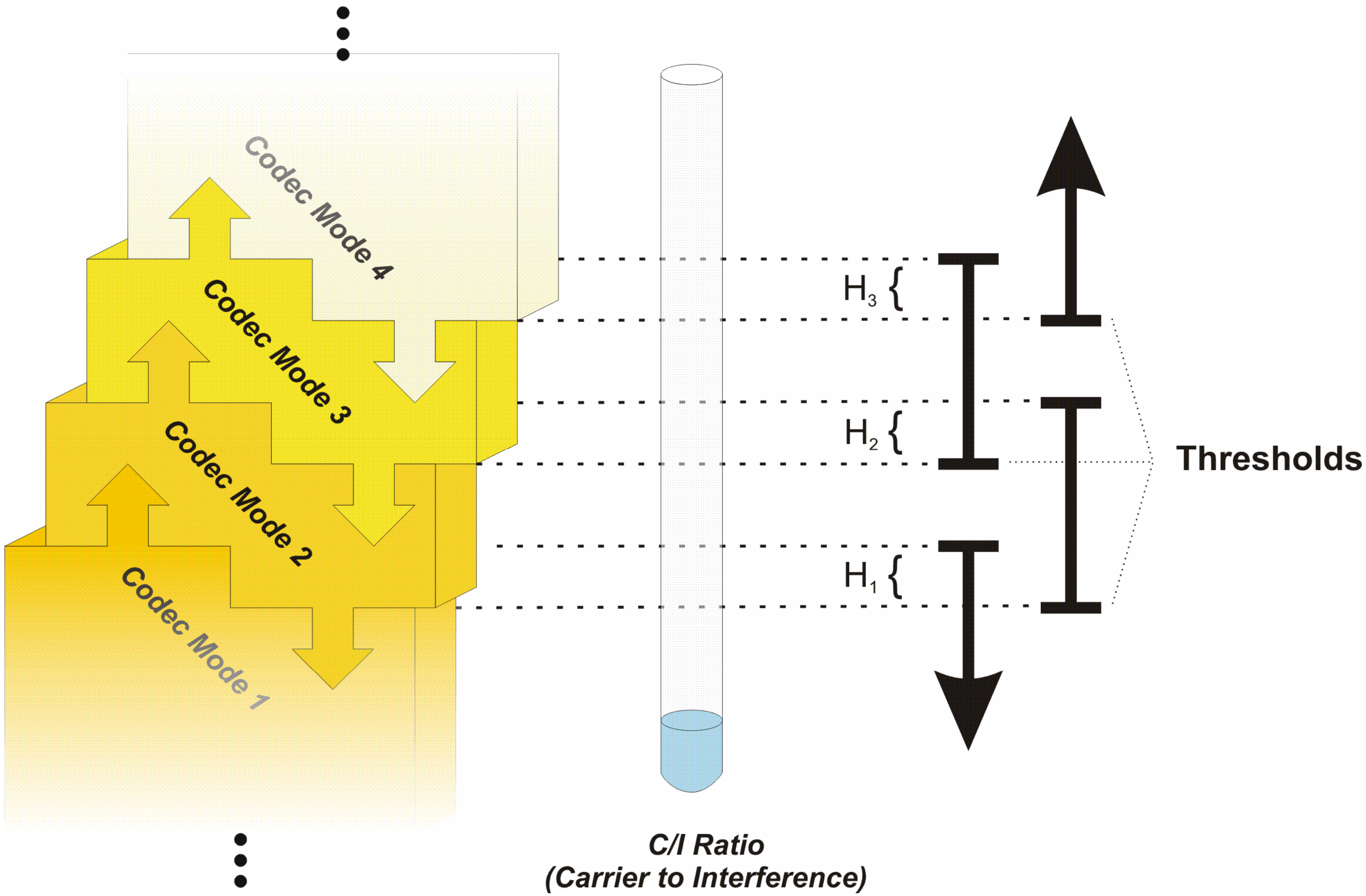
AMR Modes Bit Rates

At the speech coder output a distinct number of bits per 20 ms segment is delivered depending on the selected codec mode. These number of bits determine the bit rate. For channel coding these data streams will be regrouped in class 1a, class 1b and class 2 according to subjective importance and different error protection levels will be applied.

Class 1a and 1b bits are protected specific channel encoding schemes. Class 2 bits are of minor importance and are not protected during radio transmission. Note due to the higher channel capacity in the full rate channel all bits are protected (no class 2 bits)

[3GPP TS 05.03 (5.4)]

Thresholds and Hysteresis Values



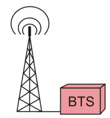
Thresholds and Hysteresis Values

The switching between different codec modes and hence the setting of the codec mode indicators and the command and request indications are aligned to the C/I value of the actual link. Based on the normalized C/I, value thresholds and hysteresis values are defined for switching between the modes. Hysteresis values are given to prevent toggling between neighbouring codec modes.

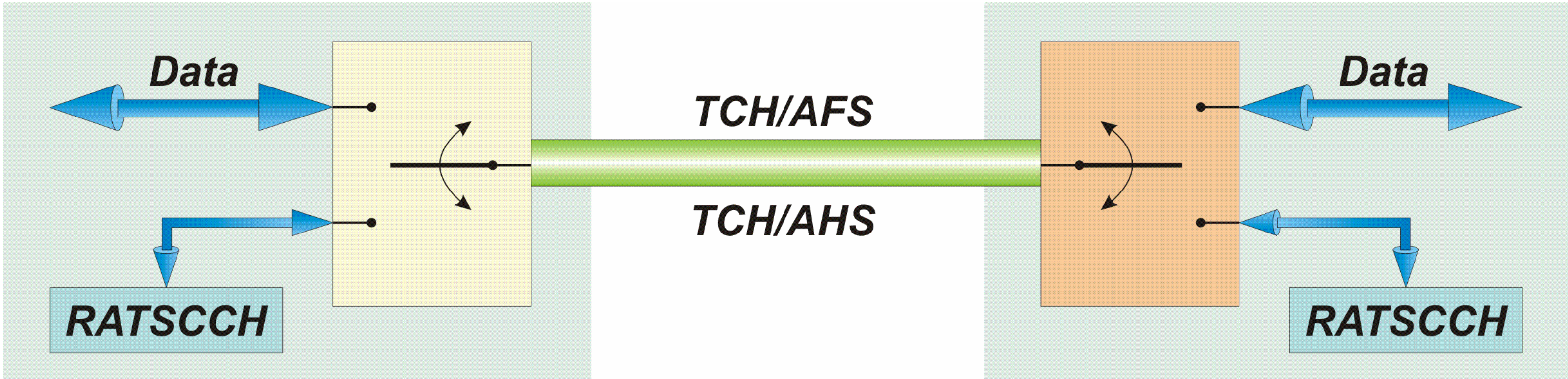
The range indications describe the codec mode to be used. For a signal change to lower C/I values the lower end indication of the individual ranges are taken as trigger for switching to the next lower codec mode. In the same way when the C/I value exceeds the upper indication of each range a switch to the next higher codec mode will be initiated.

[3GPP TS 05.09 (3.3.2 & 3.4.2)]

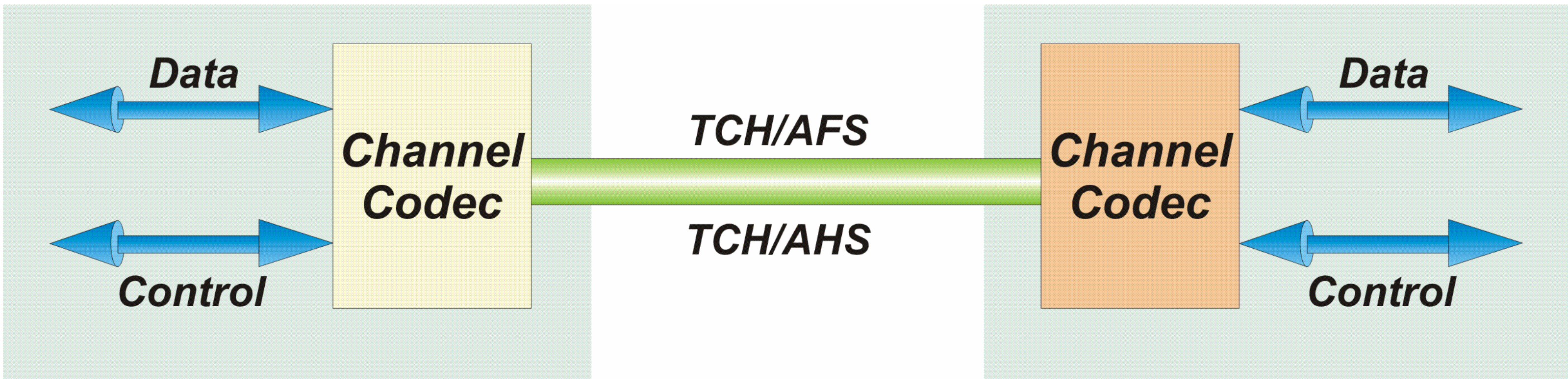
In Band Signaling



- AMR Reconfiguration



- Mode signaling



In Band Signaling

There are two different ways defined to exchange signaling information using in band procedures:

AMR reconfiguration

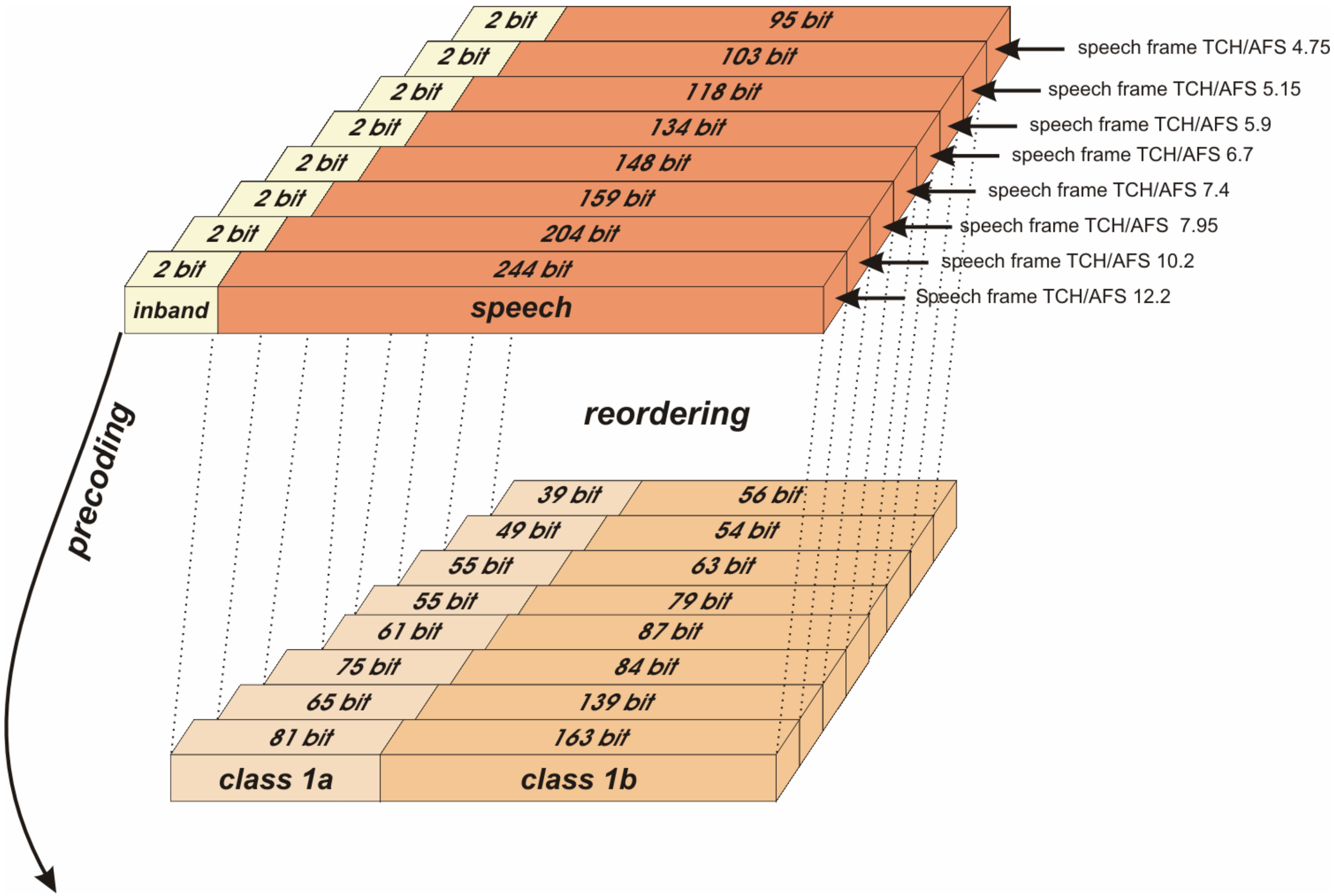
For less frequent signaling as for changing of the AMR configuration a more robust procedure is applied based on frame stealing. For this procedure a specific protocol is defined, the Robust AMR Traffic Synchronised Control Channel protocol (RATSCCH) .

Mode signaling

For frequent signaling as it is requested for codec mode indication and codec mode command or request exchange the related 2 bit code word of the codec mode is multiplexed together with the data signal in the channel encoder and demultiplexed at the receiving side.

[3GPP TS 05.09 (3.2]

(1) Speech Frame Channel Coding AFS



(1) Speech Frame Channel Coding AFS

The speech-frames are delivered by the speech coder in a sequence of blocks. The length of each speech frame depends on the actual used codec mode.

Additionally the 2 bits in-band signaling describing the codec modes are fed to the channel coder (codec mode indication or codec mode command/request depending on the frame number).

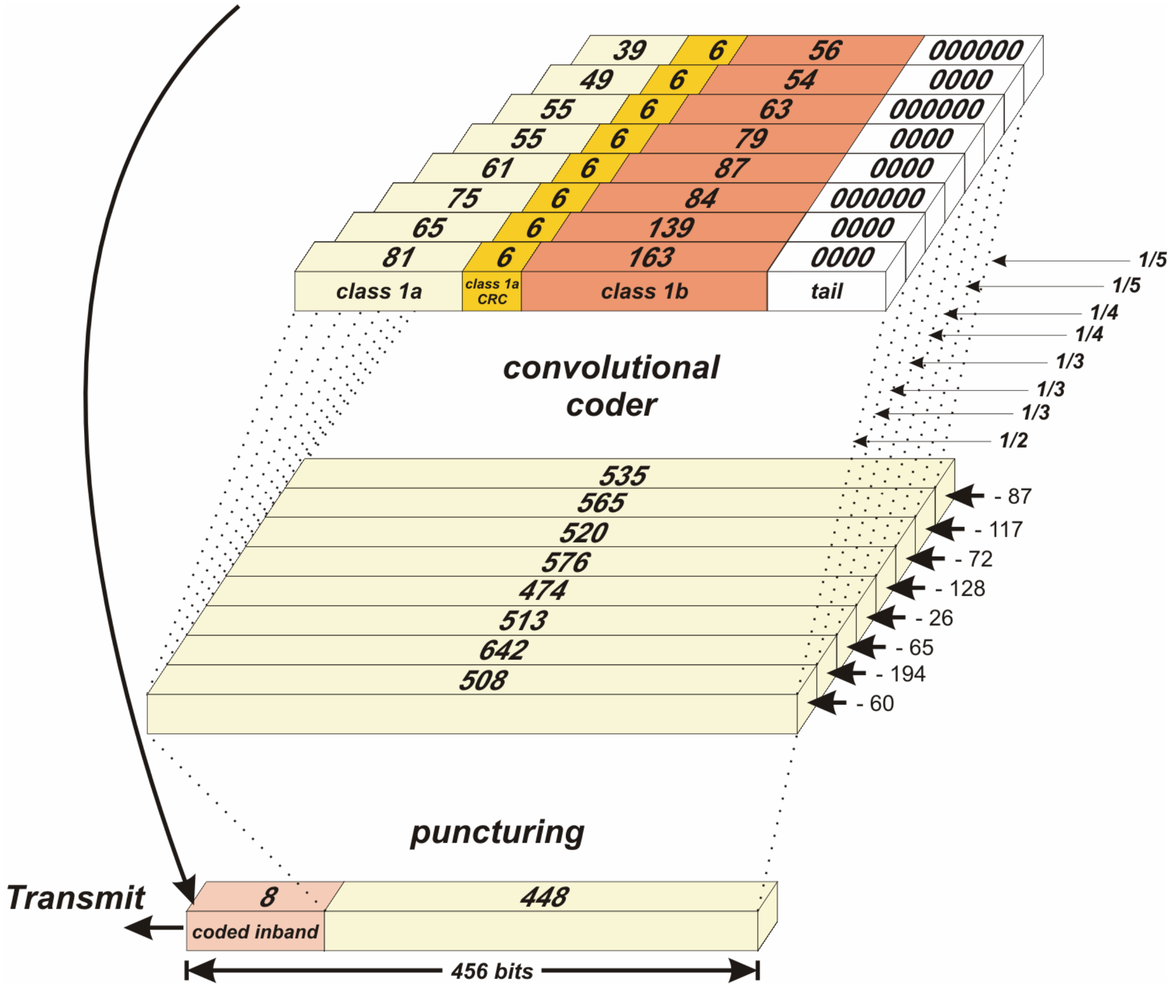
The two inputs are handled differently:

⇒ the inband signaling will be precoded to a length of 8 bit.

⇒ For the speech frame the first step is a reordering of the speech bits according to their subjective importance. The reordered speech bits are divided into two classes for which different protection levels are applicable.

[3GTS 05.03(3.9 and Table 1)]

(2) Speech Frame Channel Coding AFS



(2) Speech Frame Channel Coding AFS

Class 1a bits have a strong influence on speech quality therefore they will be additionally protected by a cyclic redundancy check (also called parity bits).

Next the class 1a bits plus its CRC, class 1b and the tail bits are coded by a convolutional coder. For the different codec modes also different coding rates between 1/2 and 1/5 are applied.

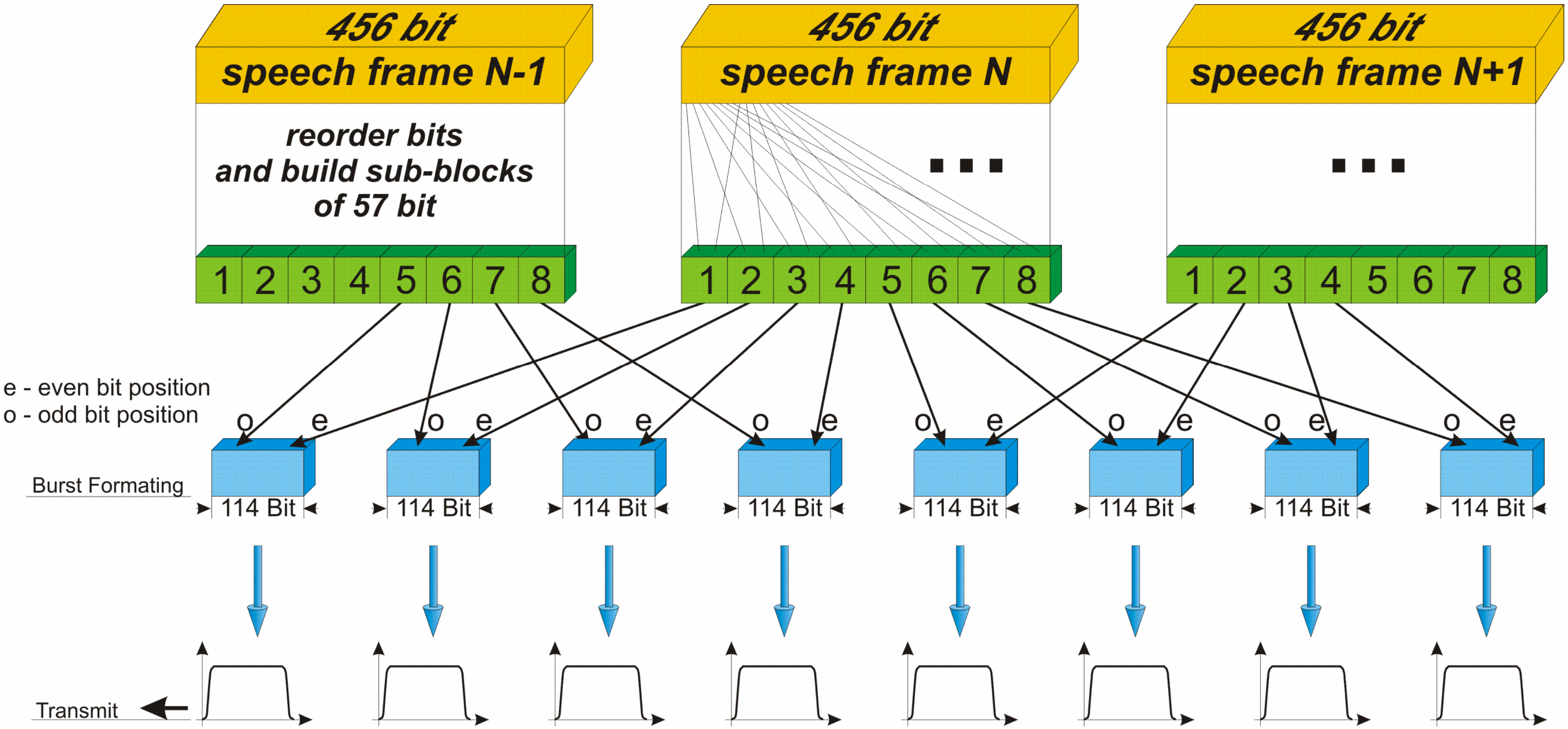
Note that the length of the tail bits to reset the convolutional coder differs between 4 bit and 6 bit.

After the coding process the output rate exceeds the maximum number of for four burst transmission. Puncturing has to be applied to fix the coded speech bits to a constant length of 448 bits. The number of bits to be punctured depend on the codec mode.

Last the coded and punctured speech data is appended to the precoded inband signal. The total length is now 456 bits that – after interleaving over 8 blocks – will be mapped normal bursts for transmission.

[3GTS 05.03 (3.9, 3.9.2 and Table 1); 06.93]

AFS Speech Frame Interleaving and Burst Generation



AFS Speech Frame Interleaving and Burst Generation

The 456 channel coded bits will be completely re-ordered and re-arranged to 8 sub-blocks of 57 bits each:

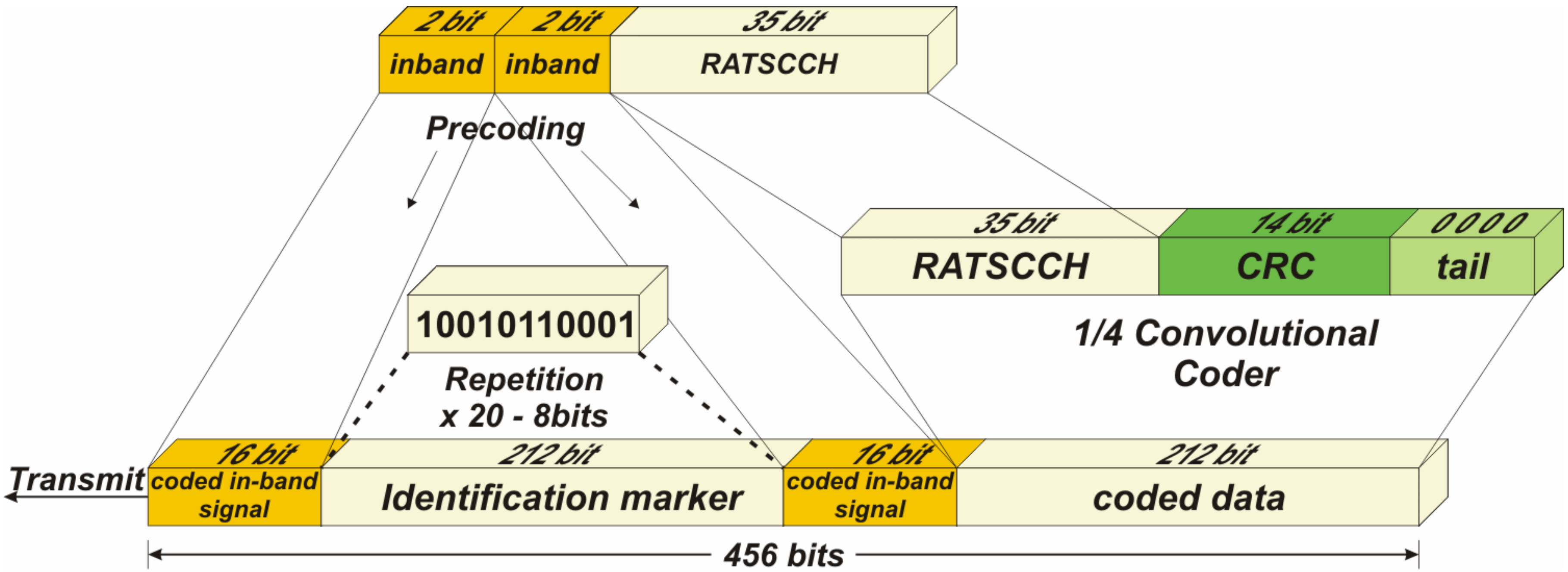
Coded bit 0 is put into sub-block 1 at position 0
Coded bit 1 is put into sub-block 2 at position 49
Coded bit 2 is put into sub-block 3 at position 41
Coded bit 3 is put into sub-block 4 at position 33
Coded bit 4 is put into sub-block 5 at position 25
Coded bit 5 is put into sub-block 6 at position 17
Coded bit 6 is put into sub-block 7 at position 9
Coded bit 7 is put into sub-block 8 at position 1
Coded bit 8 is put into sub-block 1 at position 50
Coded bit 9 is put into sub-block 2 at position 42
: : :
Coded bit 455 is put into sub-block 8 at position 8

The sub blocks are interleaved over 8 blocks (normal bursts). The first 4 sub-blocks (1-4) of frame N are mapped onto the even bits of 4 normal bursts. The sub-block 5-8 of frame N are mapped onto the odd bits of the following 4 normal bursts.

Note: a normal burst includes always bits from 2 different speech-frames. Even bits come from the higher frame number (N+1), odd bits from the lower frame (N). This adds a further delay.

[3GPP TS 05.03 (3.1.3 & table 1)

Channel Coding of RATSCCH Frame



Channel Coding of RATSCCH Frame

Any RATSCCH message has a fixed length of 35 bits, the same length as the comfort noise parameters in the SID_UPDATE frame. So a RATSCCH frame is handled in principle in the same way as a SID_UPDATE frame.

The 35 bit will be protected by a cyclic redundancy code of 14 bit, commonly called parity bits. With the 4 tail bits to reset the convolutional coder, the coding can commence with a rate of 1/4. That means the 35+14+4 bits become a channel coded sequence of 212 bit.

Both inband channels are precoded to 16 bit each (CMI or CMC/CMR depending on the direction).

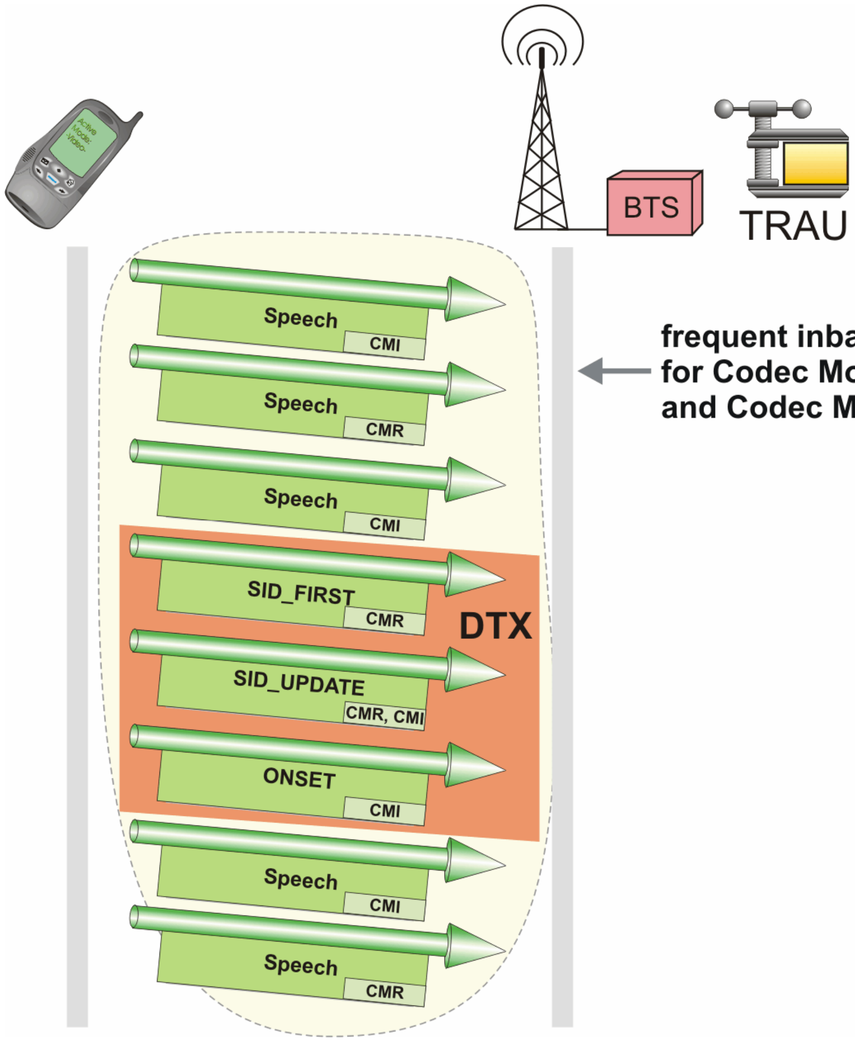
Also an identification marker which is a repetition of a fixed 11 bit value of “10010110001” will be added to the code. The identification marker will be repeated 20 times and the last 8 bits of the sequence will be discarded.

Between RATSCCH and SID_UPDATE frames there are no differences in the coding scheme. The positioning of the coded bits and of course the identification marker are different.

The RATSCCH frame will be interleaved over 8 blocks (normal interleaving scheme as for speech).

[3GTS 05.03 (3.9.5)]

(2) Overview of level 3 and inband signaling procedures



frequent inband signaling for Codec Mode Indication CMI and Codec Mode Command CMC/CMR

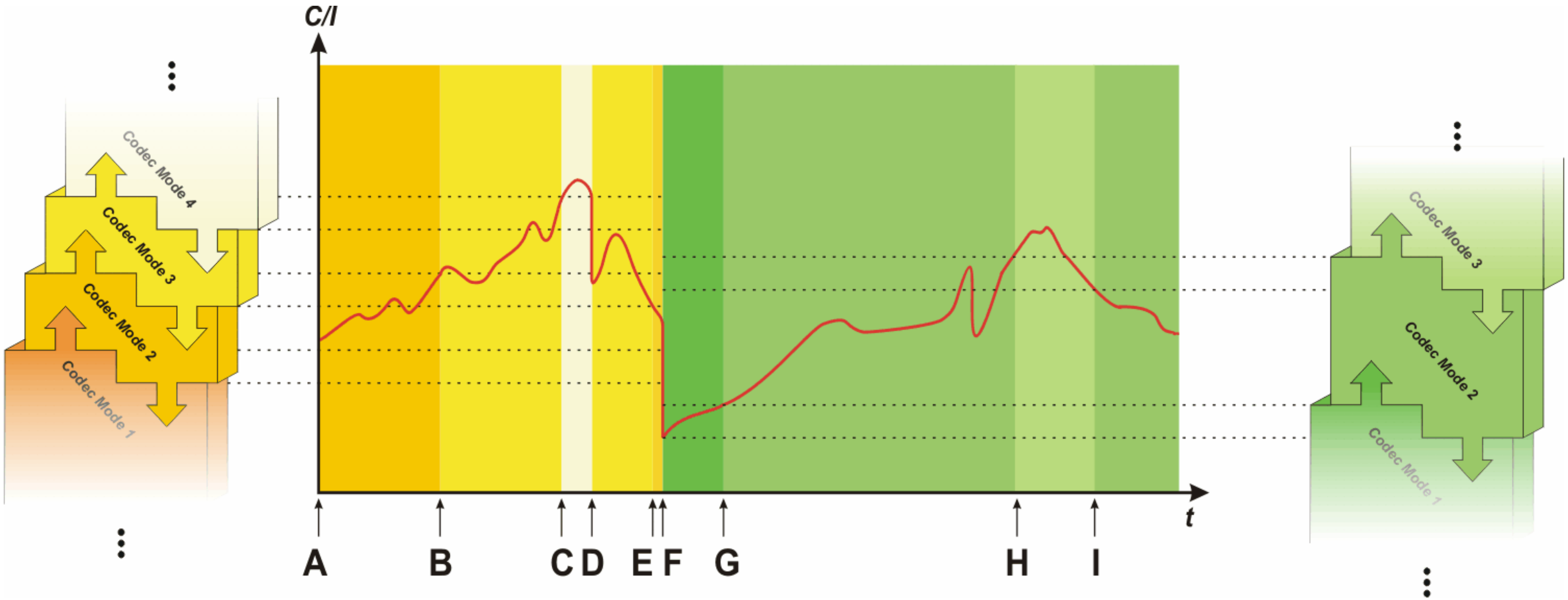
(2) Overview of level 3 and inband signaling procedures

Frequent inband signaling

In all RATSCCH-, DTX-control- and SPEECH-frames one or both inband channels can be found. The inband channels indicate the actual used codec mode or they are used to request or command the addressee to adapt the codec mode for the other direction. If the frame bears only one of the signaling channels the type alternates with every other frame.

[3GTS 05.09]

Example for Codec Mode transitions during Handover



Example for Codec Mode transitions during Handover

A The initial codec mode default adjust in an active codec set of 4 is the second lowest (codec mode_2)

B The C/I ratio passes the value of threshold 2 plus hysteresis 2, now codec mode_3 is used

C The C/I ratio passes the value of threshold 3 plus hysteresis 3, switch over to codec mode_4

D Threshold 3 is passed, codec mode_3 is to be used again

E Threshold 2 is passed, codec mode_2 is to be used again, the C/I ratio becomes lower and lower

F A handover to the “green” cell is performed, the new initial codec mode in this cell is codec mode_1 ((active codec set = 3)

G The C/I becomes better and passes the threshold 1 plus hysteresis 1 value, new codec mode is codec mode_2

H The C/I ratio passes the threshold 2 plus hysteresis 2 value, now codec mode_3 is used (highest mode in this active codec set)

I Use again codec mode_2 because the C/I ratio has fallen below of threshold 2

Trigger of a codec mode adaptation
The adaptation of the coding mode is triggered by passing a C/I threshold value. It is only allowed to change to the next higher or lower mode of the active codec set. Even if the C/I ratio changes rapidly the adaptation has to be done in two steps.

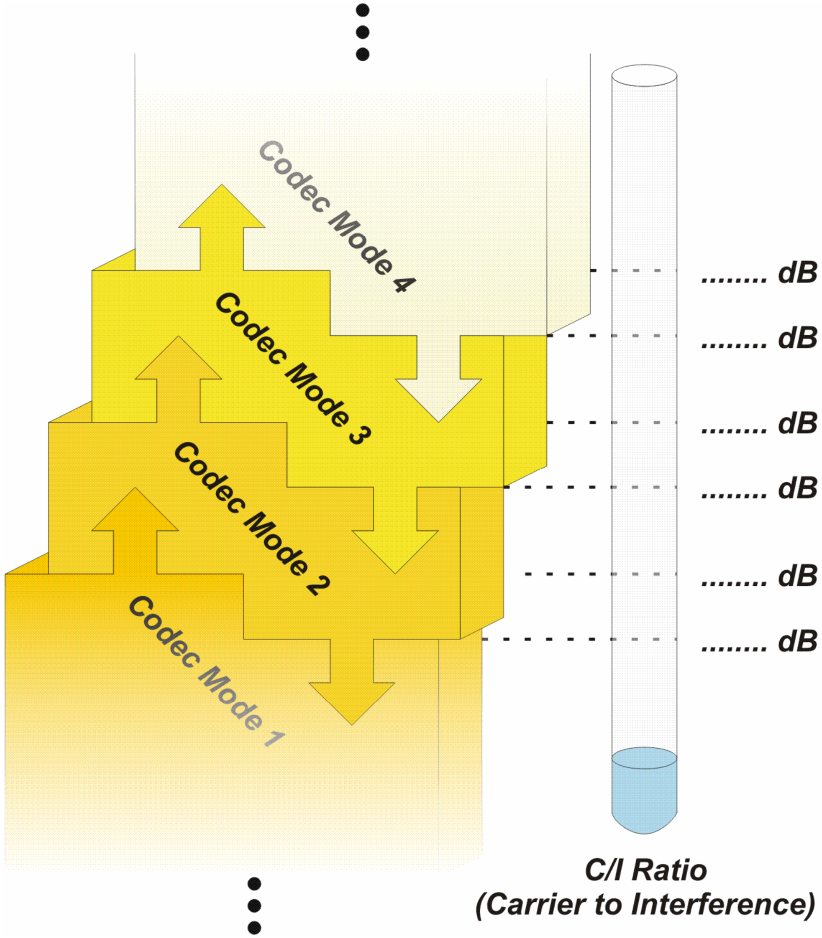
Only in case of an handover the coding mode transition may span over more than one step.

[3GTS 05.09 (3.3.2)]

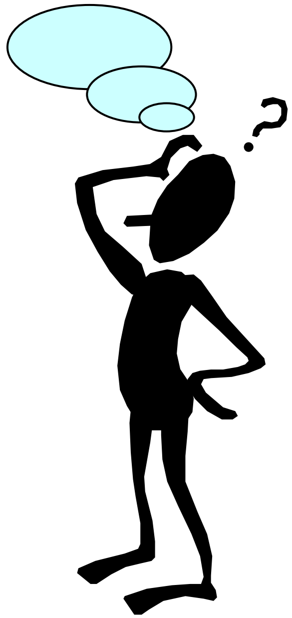
Practical Exercise

These parameters for multirate speech below have been received in an AMR configuration IE.

Please insert the switch over values between the codec modes.



1	1	1	0	0	1	0	0
X	X	0	0	1	1	0	1
0	1	0	0	0	1	0	0
1	1	0	1	0	0	0	1
1	0	1	0	0	1	0	0



(2) RATSCCH Messages

34–30	29–26	25–20	19–16	15–10	9–6	5–0
0 0 0 0 1	HYST3	THRESH3	HYST2	THRESH2	HYST1	THRESH1

(2) RATSCCH Messages

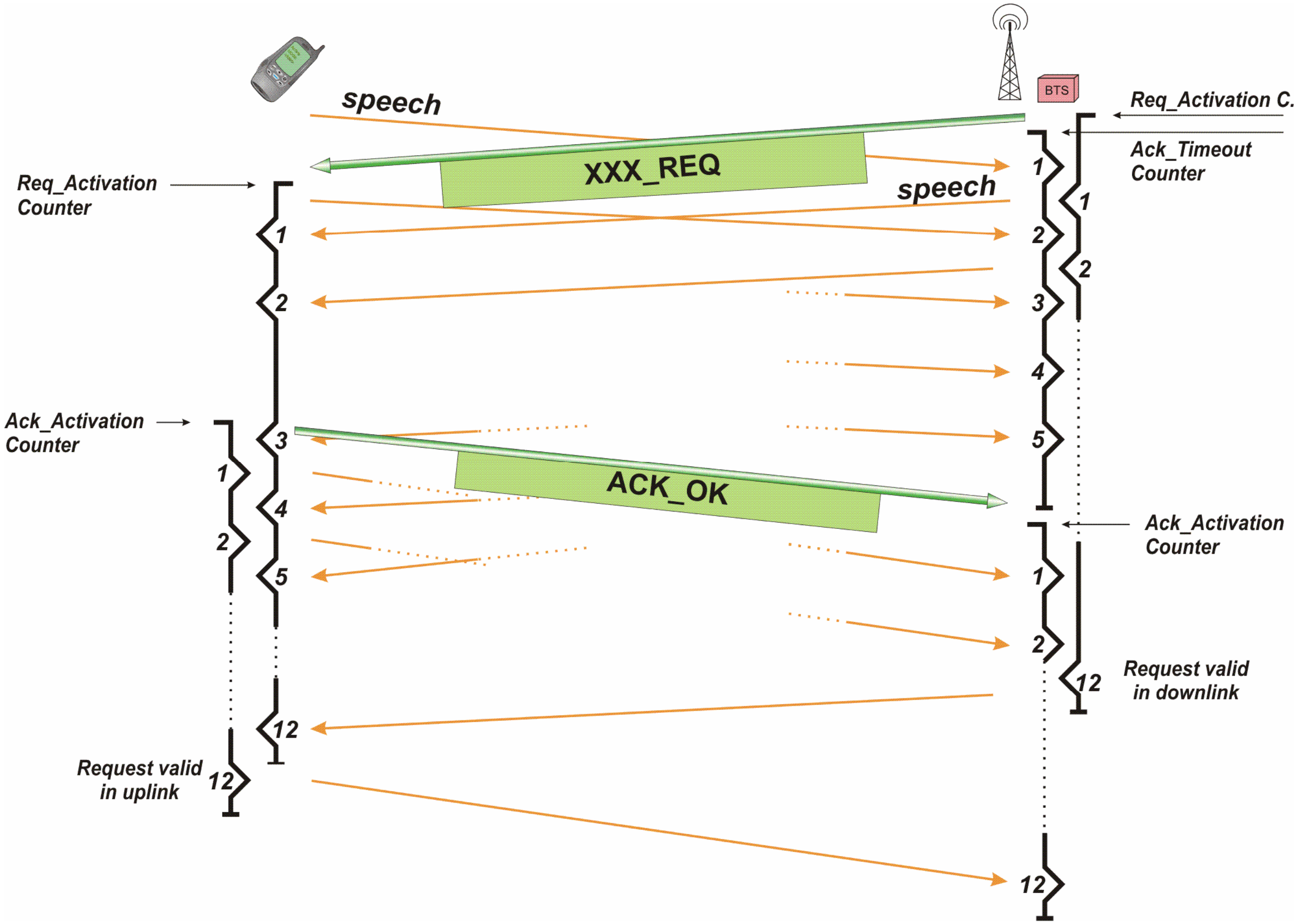
THRESH_REQ

The TRESH_REQ message allows the allocation of new threshold and hysteresis values for a given active codec set.

Again fields not used are set to “1” (less then 4 codec modes).

[3GPP TS 05.03 (3.2.2.3.6)

REQUEST – ACKNOWLEDGEMENT CYCLE



REQUEST – ACKNOWLEDGEMENT CYCLE

Information exchange between the BTS and the MS consists typically of a REQUEST – ACKNOWLEDGEMENT cycle. There shall be only one REQ – ACK cycle ongoing between the BTS and an MS and the next cycle shall only start after finalization of the previous one. To keep the cycle time short both sides shall continuously monitor the incoming signal for the RATSCCH pattern and decode the RATSCCH message. The ACK message shall be sent back latest within three frames.

With the transmit of a REQ message the BTS shall start two counters for received speech frames (ACK_Timeout) and for sent frames (REQ_Activation). At the receive side – after error free decoding of the received REQ message – the MS will start also two counters for received frames (REQ_Activation) and for sent frames after ACK message (ACK_Activation).

After the BTS has received the ACK message it will stop the ACK_Timeout counter and will start an ACK_Activation counter.

The content of the REQ message will become valid after the counters have received the value 12. In the direction to the MS the message will become valid when both REQ_Activation counters will reach the count number 12, in the direction towards the BTS when the ACK_Activation counters reach the 12. The activation will take place at four different points of time, but exactly synchronized and defined in both directions.

[3GPP TS 05.03 (3.2.2.2)

Index

A

B

C

D