Speech Coding

Jari Hagqvist
Research Manager
Speech Coding
Nokia Research Center
Speech and Audio Systems Lab
Speech Coding in GSM and 3G WCDMA

Jari Hagqvist

Nokia Research Center
Outline

• GSM Air Interface Overview
• Speech Processing in GSM Mobiles
• GSM Full Rate
• GSM Half Rate
• GSM Enhanced Full Rate
• Adaptive Multi Rate Codec for GSM
• AMR Speech Codec in WCDMA
GSM Air Interface Overview

• GSM is a digital mobile system standard operating on 900 and/or 1800 MHz bands in Europe, Asia and Africa. The North American version is a single frequency standard operating on 1900 MHz.

• GSM is a combination of **FDMA** and **TDMA**:
  - FDMA structure:
    • 124 radio carriers with 200 kHz channels
      Eg. on 900 MHz:
      Uplink: 890 - 915 MHz (MS -> BTS)
      Downlink: 935 - 960 MHz (BTS - > MS)
  - TDMA structure:
    • 8 bursts (users) per radio carrier

  - Efficient channel allocation 200 kHz / 8 = 25 kHz
GSM Air Interface Overview (cont.)

- Consists of optimized (traffic) channels for speech and data as well as a number of control channels.

- Interleaving is used to randomise the effect of transmission errors. Bits will be spread over several TDMA frames.

- Slow frequency hopping (transmitting consecutive bursts on different frequencies) to protect against multipath fading.

- Gaussian minimum shift keying (GMSK) is used for modulation into carrier.

- Discontinuous Transmission (DTX) for lower power consumption and decreased interference.
Speech Processing in GSM Mobiles

UPLINK:

- A/D
- Speech Coding
  - VAD/DTX
- Echo Cancellation
- D/A
- Bad Frame Handling
- Speech Decoding
- Comfort Noise
- Channel Coding
- Interleaving
- Burst Building
- Modulation (GMSK)
- Demodulation
- De-interleaving
- Channel Decoding
- RF

DOWNLINK:
GSM Speech Traffic Channels

- GSM has four (five) standardized speech traffic channels:
  - Full Rate (TCH/FS),
  - Half Rate (TCH/HS), gross rate 11.4 kbit/s
  - Enhanced Full Rate (TCH/EFS), gross rate 22.8 kbit/s
  - Adaptive Multi Rate (AMR) at full rate and half rate (TCH/AFS, TCH/AHS), gross rate 22.8 / 11.4 kbit/s
  - AMR Wideband gross rate 22.8 kbit/s

- Each channel has an optimized channel coding (error protection) for the corresponding speech codec(s) used in that channel

- Unequal error protection: not all the speech bits are protected by channel coding

- Bad frame compensation by repeating and muting previous good frame
Speech Codecs in GSM

- Four narrow band (300-3400 Hz input bandwidth) speech codecs has been standardized for the GSM system:
  - **GSM Full Rate** (FR, standardized in 1987)
    - RPE-LTP (Regular Pulse Excitation-Long Term Prediction)
  - **GSM Half Rate** (HR, 1994)
    - VSELP (Vector Sum Excited Linear Prediction)
  - **GSM Enhanced Full Rate** (EFR, 1995)
    - ACELP (Algebraic Code Excited Linear Prediction)
  - **GSM Adaptive Multi Rate** (AMR, 1999)
    - ACELP (Algebraic Code Excited Linear Prediction)
  - A **Wideband Adaptive Multi Rate** (50-7000 Hz input bandwidth) codec was finalized 12/2000
    - ACELP
GSM Full Rate Codec

- **13.0 kbit/s** RPE-LTP (Regular Pulse Excitation - Long Term Prediction)
- Basic methodology: residual excited linear prediction
- Short-term modelling/filtering with 8th order Linear Predictive Coding (LPC). Inverse filtering employed in the encoder to form prediction error (residual).
- Long-term modelling/filtering with 1st order Long Term Prediction (LTP)
- Modelling of prediction error by Regular Pulse Excitation (RPE)
- RPE-LTP operates on 20 ms speech frames; for each frame 260 speech parameter bits are produced (13.0 kbit/s)
## Bit Allocation of GSM FR

<table>
<thead>
<tr>
<th></th>
<th>Bits per 5 ms</th>
<th>Bits per 20 ms</th>
<th>Bit-rate kbit/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPC filter</td>
<td>8 coefficients</td>
<td>36</td>
<td>1.8</td>
</tr>
<tr>
<td>LTP filter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lag</td>
<td>7</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>gain</td>
<td>2</td>
<td>8</td>
<td>1.8</td>
</tr>
<tr>
<td>Excitation signal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gain</td>
<td>6</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>pulse amplitudes</td>
<td>39</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td>phase</td>
<td>2</td>
<td>8</td>
<td>9.4</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>260</td>
<td>13.0</td>
</tr>
</tbody>
</table>

- The output bits are classified according to their subjective importance into three classes: 1A (50 bits), 1B (132 bits) and class 2 (78 bits)
GSM FR Channel Coding

- For error detection a 3-bit CRC is calculated over the 50 most important bits (Class 1A).

- Error correction coding of 9.8 kbit/s is added using a 1/2-rate convolutional coding for Class 1A and 1B bits (182 most sensitive bits).

- Interleaving over 8 TDMA frames.

```
G(D) = D^3 + D + 1
G0 = 1 + D^3 + D^4
G1 = 1 + D + D^3 + D^4
3 bit CRC
1/2-rate convolutional encoding 378 bits
Interleaving for 456 bits
```

| S P E E C H   | 3 bit CRC | 1/2-rate convolutional encoding 378 bits |
| E N C O D E R  |           |                                          |
| 260 bits/20 ms|           |                                          |
| 1A | 1B | 132 | 2 | 78 | 0000 tail bits |

Bit ordering according to subjective importance.
GSM FR Interleaving

Channel coded speech frame N+2

Channel coded speech frame N+1

Channel coded speech frame N

Bursts

Coded speech

tail

stealing flags, training sequence

tail

1 2 3 4 5 6 7 8
GSM Half Rate Codec

- **5.6 kbit/s** VSELP (Vector Sum Excited Linear Prediction)
- Analysis-by-synthesis CELP (Code Excited Linear Prediction) operating on 20 ms speech frames
- Short-term modelling/filtering with 10th order LPC
- Long term modelling/filtering (1st order) is employed as an adaptive codebook using the analysis-by-synthesis approach with fractional lag-values and non-uniform resolution
- Excitation formed by analysis-by-synthesis approach from a linear combination of basis vectors (⇒ Vector Sum Excitation)
- Four modes used: 1 unvoiced mode and 3 voiced modes
- Adaptive post enhancement filtering is applied in the decoder
# GSM HR Bit Allocation

<table>
<thead>
<tr>
<th>VOICED MODES 1,2,3</th>
<th>Bits per 5 ms</th>
<th>Bits per 20 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPC filter</td>
<td>10 coefficients</td>
<td>28</td>
</tr>
<tr>
<td>Frame energy</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Soft interpolation bit</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Voicing mode</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>LTP filter (adaptive codebook)</td>
<td>lag 8,4,4,4</td>
<td>20</td>
</tr>
<tr>
<td>Excitation signal (fixed codebook)</td>
<td>gain 5 (joint codebook index)</td>
<td>7 28</td>
</tr>
<tr>
<td>Excitation signal (fixed codebook)</td>
<td>gain (joint quantiser)</td>
<td>5 20</td>
</tr>
<tr>
<td>Total</td>
<td>112</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UNVOICED MODE</th>
<th>Bits per 5 ms</th>
<th>Bits per 20 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPC filter</td>
<td>10 coefficients</td>
<td>28</td>
</tr>
<tr>
<td>Frame energy</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Soft interpolation bit</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Voicing mode</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Excitation signal</td>
<td>codebook index</td>
<td>7 28</td>
</tr>
<tr>
<td>Excitation signal</td>
<td>gain (joint quantiser)</td>
<td>5 20</td>
</tr>
<tr>
<td>Excitation signal</td>
<td>codebook index</td>
<td>7 28</td>
</tr>
<tr>
<td>Total</td>
<td>112</td>
<td></td>
</tr>
</tbody>
</table>
GSM Half Rate Channel Coding

• 5.8 kbit/s using 1/2-rate convolutional coding with 3-bit CRC (CRC is protected by 1/3-rate convolutional code)

• Class 1A: 22 bits (convolutional code + 3 bit CRC)
  Class 1B: 73 bits (convolutional code)
  Class II: 17 bits (unprotected)

• Note: more bits are used for error protection than for source coding!

• Polynomials:
  \[1+D^2+D^3+D^5+D^6,\]
  \[1+D+D^4+D^6,\]
  \[1+D+D^2+D^3+D^4+D^6\]
  CRC: \[X^3+X+1\] (same as in full rate)

• Interleaving depth: 4 (block length 57 bits)
GSM Enhanced Full Rate Codec

- First codec with quality comparable to landline phone quality (better than 32 kbit/s ADPCM)
- **12.2 kbit/s** ACELP (Algebraic Code Excited Linear Predictive Coding)
- CELP type coding with a fixed pulse codebook enabling a fast excitation search procedure
- The same codec as was earlier chosen as the US1 codec for the EFR channels in the PCS1900 system in North America
GSM EFR Details

- **Short-term analysis (Linear Prediction)**
  - 10th order LP analysis twice for each 20 ms frame
  - Two 30 ms asymmetric windows (no lookahead)
  - Represented in LSPs (Line Spectral Pairs)
  - Joint predictive split matrix quantisation with 38 bits / 20 ms-frame

- **Adaptive codebook (ACB)**
  - Combined open-loop/closed-loop search
    - open-loop lag search every 10 ms
    - closed-loop lag search every 5 ms
  - Fractional lag with 1/6th resolution
  - Residual extended virtual lag for high pitch voices
GSM EFR Details (cont.)

• Fixed codebook (FCB) excitation search
  - Algebraic codebook with 10 pulses / 5 ms subframe
  - Predefined interlaced sets of pulse positions
  - Non-exhaustive search with low complexity
  - Pitch sharpening to improve coding of high pitch voices

• Bad frame handling
  - Reliable bad frame detection with 8-bit CRC
  - Partial replacement for parameters of received bad frames
  - Muting adjusted dynamically according to channel error conditions
### GSM EFR Bit Allocation

<table>
<thead>
<tr>
<th></th>
<th>Bits per 20 ms</th>
<th>Bit-rate kbit/s</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LPC filter</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 coefficients</td>
<td>38</td>
<td>1.9</td>
</tr>
<tr>
<td><strong>Adaptive</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lag</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>codebook</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gain</td>
<td>16</td>
<td>2.3</td>
</tr>
<tr>
<td><strong>Fixed</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gain</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>codebook</td>
<td></td>
<td></td>
</tr>
<tr>
<td>index</td>
<td>140</td>
<td>8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>244</td>
<td>12.2</td>
</tr>
</tbody>
</table>

- An additional 0.8 kbit/s used for internal error protection (8-bit CRC and repetition coding) -> 13 kbit/s
EFR Channel Coding

- The EFR codec was designed to use the GSM full rate channel coding as is, with an additional 8-bit CRC for improved error detection.
  -> Only the speech codec needs to be updated in upgrading from full rate to enhanced full rate networks.

- This provides a fast time-to-market for the EFR system, with implementation advantages and minimal additional costs.

- Some internal error protection (repetition) is added for more robust operation in channel errors.
EFR Channel Error Performance

Channel Error Test (ACR, MOS)

- C/I = 13 dB: 2% channel bit error-rate (well inside a cell)
- C/I = 10 dB: 5% channel bit error-rate (inside a cell)
- C/I = 7 dB: 8% channel bit error-rate (at a cell boundary)
Adaptive Multi Rate Codec for GSM and 3G

- ETSI SMG11 and 3GPP TSG-SA WG4 have recently standardized the Adaptive Multi Rate (AMR) codec for GSM and 3G WCDMA systems
- AMR is a versatile codec "toolbox" operating at several bit rates for robust operation in mobile channels
- AMR has been selected as the mandatory speech codec for the 3G WCDMA system by 3GPP
- AMR is also the mandatory speech codec for 3G-H.324M video telephony
- For more information and specifications see:
  - www.3gpp.org
  - ftp://ftp.3gpp.org/TSG_SA/WG4_CODEC/
What is AMR?

• AMR contains a set of fixed-rate speech codec modes each of which have a different error protection level (amount of channel coding)

• The codec dynamically adapts its error protection level to the channel error and traffic conditions (link adaptation)
  » Uses lower speech coding bit rate and more error protection in bad channel conditions

• This gives substantial improvements to the robustness against channel errors (especially in GSM)

• Also capacity benefits (eg. using the GSM half rate channel)

• Due to the fast power control in the 3G WCDMA system, the AMR link adaptation is not as useful in WCDMA as in GSM
AMR Operation

- Link adaptation "switches" the to the best curve (bit rate) in A, B and C
AMR Codec Bit Rates

- The AMR codec contains **8 source codecs**:

  - **12.2 kbit/s** (=GSM EFR) FR channel
  - **10.2 kbit/s** FR
  - **7.95 kbit/s** FR+HR
  - **7.40 kbit/s** (=US-TDMA IS-641 EFR) FR+HR
  - **6.70 kbit/s** (=PDC EFR) FR+HR
  - **5.90 kbit/s** FR+HR
  - **5.15 kbit/s** FR+HR
  - **4.75 kbit/s** FR+HR

- 8 codec modes operate in the GSM FR channel and 6 in the HR channel

- All modes are used in 3G WCDMA
Channel Coding

- In GSM each AMR mode has an **optimized channel codec** for operation in Full Rate (22.8 kbit) and/or Half Rate (11.4 kbit) channels

- Recursive Systematic Convolutional (RSC) codes and 6-bit CRC used

- Mode bits are protected by a separate block code

- 3G WCDMA is based on **generic channels** (Layer 1 "toolbox"), hence mode specific channel coding cannot be used for AMR in 3G like in GSM
  - Individual **rate matching** for the modes (by puncturing and repetition)
  - Unequal error protection is implemented using different Radio Access Bearers (RABs) with different QoS requirements
Ratio of Speech and Channel Coding (GSM)

Channel Mode: FS   FS   FS   FS   FS   FS   FS   HS   HS   HS   HS   HS
Codec Mode: 12.2  10.2  7.95  7.40  6.70  5.90  5.15  4.75  7.95  7.40  6.70  5.90
            10.2  7.95  7.40  6.70  5.90  5.15  4.75  7.95  7.40  6.70  5.90
            7.95  7.40  6.70  5.90  5.15  4.75  7.95  7.40  6.70  5.90  5.15
            7.40  6.70  5.90  5.15  4.75  7.95  7.40  6.70  5.90  5.15  4.75
            6.70  5.90  5.15  4.75  7.95  7.40  6.70  5.90  5.15  4.75  7.95
            5.90  5.15  4.75  7.95  7.40  6.70  5.90  5.15  4.75  7.95  7.40
            5.15  4.75  7.95  7.40  6.70  5.90  5.15  4.75  7.95  7.40  6.70

Channel Gross Bit Rate [kbit/s]

FS: Full-Rate
HS: Half-Rate

Channel Mode [FS/HS] / Codec Mode [kbit/s]
Codec Mode Adaptation (GSM)

- The **codec mode adaptation** chooses the optimum codec mode as a function of channel quality (or eg. network load).

- The most robust mode is chosen in poor channel conditions, while the codec mode providing best clean-channel quality is chosen in good error conditions.

- Codec mode adaptation based on **channel quality measurements** (C/I estimates) done in the mobile and network.

- Based on the measurements a Codec Mode Command (over downlink to MS) or Codec Mode Request (over uplink to network) is sent over the air interface (using **inband signaling**).

- In GSM, the inband signaling supports a set of up to 4 codec modes (2 bits) selected at call set-up or handover. In 3G all modes can be used.
AMR Adaptation Example

C/I [dB] vs Time [s]

- Red line: C/I
- Blue line: AMR Mode

Time range: 0.0 to 12.5 seconds
C/I range: 0 to 25 dB
AMR Mode range: 5.90 to 12.2 kbit/s

AMR Mode: GSM EFR
AMR System Block Diagram (GSM)

- **S**: Speech Data
- **Q**: Channel Quality
- **MC**: Mode Command
- **MR**: Mode Request
- **MI**: Mode Indicator

**In-Band Signalling:**
- **MC**: Mode Command
- **MR**: Mode Request
- **MI**: Mode Indicator

**Mobile System (MS)**
- Multi-Rate Speech Encoder
- Multi-Rate Channel Encoder
- Downlink Quality Estimator
- Link Adaptation

**Radio Channel (UPLINK)**
- Multi-Rate Channel Decoder

**Base Transceiver Station (BTS)**
- Uplink Mode Adaptor
- Uplink Quality Estimator
- Link Adaptation
- Multi-Rate Channel Encoder

**Radio Channel (DOWNLINK)**
- Multi-Rate Speech Decoder

**TransCorder (TC)**
- Multi-Rate Speech Decoder
- Speech Out
AMR Operation (UPLINK)

1. Call Setup Signalling:
   - Mode sub-set pre-selection with selection constraints:
     - max. 4 modes out of 8
     - Channel mode (FS/HS)
     - Modes supported by the network

2. Example mode sub-set (FS-channel):
   - 12.2, 7.95, 6.7 and 4.75 kbit/s
   - Selected out of 8 possible modes [kbit/s]:
     - 12.2, 10.2, 7.95, 7.4, 6.7, 5.9, 5.15, 4.75

3. Normal Multi Rate Codec Operation Without Adaptation:
   - 22.8 kbit/s (FS)
   - 11.4 kbit/s (HS)

4. Adaption of the Mode based on:
   - Channel Quality
   - Network Constraints
AMR Speech Codec Key Features

- **Frame length:** 20 ms with four subframes in all modes, 5 ms lookahead (except in 12.2 kbit/s mode)

- **LPC analysis:** Four different LSF-tables; 38 bits (EFR), 27 bits, 26 bits (IS-641) and 23 bits

- **Adaptive codebook:** GSM EFR and IS-641 based open-loop/closed-loop search in all modes

- **Fixed codebook:** ACELP codebooks in all modes using 10, 8, 4, 3 and 2 pulses

- **Quantisation:** GSM EFR (scalar) or IS-641 (joint VQ) based codebook gain quantisation in all modes

- **Post-processing:** Formant postfilter with tilt compensation and 60 Hz HP filter in all modes (also EFR-type post processing of excitation elements included). Anti-sparseness processing for 7.95 kbit/s and lower modes.
## Comparison of AMR source codecs

<table>
<thead>
<tr>
<th>Mode</th>
<th>Pre-processing (20 ms)</th>
<th>LPC (20 ms)</th>
<th>Open-loop pitch analysis (10 ms)</th>
<th>Adaptive codebook (5 ms)</th>
<th>Fixed codebook (5 ms)</th>
<th>Codebook gain quantisation (5 ms)</th>
<th>Post-processing (20 ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>12.2 (GSM EFR)</strong></td>
<td>Two asymmetric windows. Split matrix quantisation (LSFs with 38 bits). No lookahead (5 ms &quot;dummy lookahead&quot;)</td>
<td>EFR-type with 3 lag ranges.</td>
<td>Resolution 1/6 [17 3/6 - 94 3/6] and 1 [95-143]; 2nd and 4th subframe 1/6 always, searched around previous lag.</td>
<td>10 pulses searched in 5 tracks. Quantised to 35 bits.</td>
<td>Scalar quantisation with 5 bits (fixed codebook) and 4 bits (adaptive codebook) bits</td>
<td></td>
<td>Short-term with tilt compensation and 60 Hz HP.</td>
</tr>
<tr>
<td><strong>10.2</strong></td>
<td>HP filter (80 Hz) in all modes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>7.95</strong></td>
<td>One asymmetric window Vector quantisation (LSFs with 38, 26, 27, 26, 26, 23 and 23 bits) 5 ms lookahead</td>
<td></td>
<td>Resolution 1/3 [19 1/3 – 84 2/3] and 1 [85-143]; 2nd and 4th subframe 1/3 always, searched around previous lag.</td>
<td>8 pulses searched in 4 tracks. Quantised to 31 bits.</td>
<td></td>
<td></td>
<td>+ anti-sparseness processing</td>
</tr>
<tr>
<td><strong>7.4 (DAMPS EFR)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>6.7</strong></td>
<td></td>
<td></td>
<td>Resolution 1/3 [19 1/3 – 84 2/3] and 1 [85-143]; 2nd and 4th subframe integer or 1/3, searched around previous lag.</td>
<td>3 pulses searched in 3 tracks. Quantised to 14 bits.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>5.9</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>5.15</strong></td>
<td></td>
<td></td>
<td>Resolution 1/3 [19 1/3 – 84 2/3] and 1 [85-143]; 2nd 3rd and 4th subframe integer or 1/3, searched around previous lag.</td>
<td>2 pulses searched in 2 tracks. Quantised to 11 bits.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>4.75</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Nokia
### AMR source codec complexity: WMOPS

#### Table I: Theoretical Worst Case complexity in wMOPS

<table>
<thead>
<tr>
<th></th>
<th>TWC** AMR</th>
<th>TWC** EFR</th>
<th>TWC** FR</th>
<th>TWC** HR</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMR source encoder</td>
<td>14.08</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AMR source decoder</td>
<td>2.53</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total AMR source codec</strong></td>
<td>16.61</td>
<td>15.21</td>
<td>2.95</td>
<td>18.47</td>
</tr>
</tbody>
</table>

(Calculated from C-code implemented with ETSI TCH-HS basic operations.)


## AMR source codec complexity: RAM, ROM

### Table II: RAM

<table>
<thead>
<tr>
<th></th>
<th>Static (16 bits words)</th>
<th>Dynamic (16 bits words)</th>
<th>Total (16 bits words)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMR source encoder *</td>
<td>1429</td>
<td>3039</td>
<td>4468</td>
</tr>
<tr>
<td>AMR source decoder *</td>
<td>811</td>
<td>946</td>
<td>1757</td>
</tr>
<tr>
<td>Total AMR source codec *</td>
<td>2240</td>
<td>3039</td>
<td>5819</td>
</tr>
<tr>
<td>Total EFR source codec **</td>
<td>-</td>
<td>-</td>
<td>4711</td>
</tr>
<tr>
<td>Total FR source codec **</td>
<td>-</td>
<td>-</td>
<td>1201</td>
</tr>
<tr>
<td>Total HR source codec **</td>
<td>-</td>
<td>-</td>
<td>4636</td>
</tr>
</tbody>
</table>

### Table III: ROM

<table>
<thead>
<tr>
<th></th>
<th>ROM (tables) (16 bits words)</th>
<th>Program ROM (source code size) (# of operators)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total AMR source codec</td>
<td>14343</td>
<td>4830</td>
</tr>
<tr>
<td>Total EFR source codec **</td>
<td>5267</td>
<td>-</td>
</tr>
<tr>
<td>Total FR source codec **</td>
<td>80</td>
<td>-</td>
</tr>
<tr>
<td>Total HR source codec **</td>
<td>7881</td>
<td>-</td>
</tr>
</tbody>
</table>

*) “Complexity verification report of the AMR codec, Version: 2.0” Source: Alcatel, Philips, ST Microelectronics, Texas Instruments, Tdoc ETSI SMG11 117/99 (SMG11#10, Sophia Antipolis, 3-7 May 1999)

AMR Performance

Performance in GSM FR channel

✦ At low errors ($C/I \geq 13\text{dB}$): equivalent to EFR no errors
✦ At 4 dB C/I:
  → Still equivalent or close to EFR at 10 dB C/I (i.e. about 2 MOS improvement over EFR), and
  → In background noise, equivalent or close to FR at 10 dB C/I

Performance in GSM HR channel

✦ At low errors ($C/I \geq 16\text{dB}$):
  → Equivalent to G.728 (“wireline”), and
  → In background noise, equivalent to G.729/FR
✦ At high errors
  → Equivalent to FR, and
  → In background noise, FR or somewhat lower
AMR Codec Performance

Clean speech - static channel

Experiment 1a - Combined Test Result
Performance in Full Rate in Clean Speech

FR channel
AMR Codec Performance

Street noise - static channel

Experiment 2a - Combined Test Results
Performance in Full Rate with Street Noise

Experiment 2c - Combined Test Results
Performance in Half Rate with Street Noise

FR channel

HR channel
AMR Codec Performance

Dynamic channel

- Dynamic channel test designed to evaluate AMR performance in realistic radio environment with codec adaptation turned on.

Experiment 4a - Combined Test Results
Performance in Full Rate and Dynamic C/I Conditions

Experiment 4b - Combined Test Results
Performance in Half Rate and Dynamic C/I Conditions

FR channel

HR channel
• Voice activity in a normal conversation is about 50%

• Discontinuous Transmission (DTX): RF transmission is cut during speech pauses
  - longer battery life
  - less interference

• The Voice Activity Detector (VAD) controls the DTX by detecting presence of speech and background noise

• When the transmission is cut the speech decoder in the receiver generates locally a 'comfort noise'

• Silence Descriptor (SID) frames are sent periodically to the receiver to update the comfort noise

• SID update rate: 480 ms for GSM FR, EFR, 240 ms for GSM HR, 160 ms for GSM AMR
AMR speech codec in WCDMA
AMR in WCDMA

• Generic channel coding 'toolbox'
  - Flexibility in design
  - Unequal or equal error protection

• Fast power control
  - No need for fast link adaptation
  - Adaptation for optimising the cell capacity
Generic channel coding toolbox

• Radio network operator has great flexibility to design the error protection scheme and QoS parameters

• Unequal or equal error protection is available
  - QoS, speech quality and capacity issues drive the selection between UEP and EEP
  - UEP enables lower transmission power (higher capacity) with same QoS and speech quality to EEP

• Same fixed channel coding algorithm for each AMR mode
  - Lower AMR modes provide higher capacity, not error protection gain

• Rate matching through repetition/puncturing
Error protection performance

3GPP DL performance (vehicular A, 120 km/hr, SF=128)

Es/No vs FER

- Comparison of AMR 12.2 mode performance using EEP and different UEP schemes
Fast power control

• Inner power control loop
  - Controls the transmission power based on the channel quality
  - Objective is to maintain the QoS (FER) parameters

• Outer power control adjusts the AMR mode
  - AMR mode is decreased if FER target is not achieved and transmission power is saturating
  - Objective is to maintain the target capacity

• Fast AMR link adaptation is not needed
  - Power control minimises the effect of channel fading etc.
  - Mode requests using out of band signalling
Annex A: List of 3G AMR Speech Coding Specifications

- Available from www.3gpp.org (also the GSM specifications)

<table>
<thead>
<tr>
<th>TS</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS 26.071</td>
<td>AMR speech Codec; General description</td>
</tr>
<tr>
<td>TS 26.073</td>
<td>AMR speech Codec; C-source code</td>
</tr>
<tr>
<td>TS 26.074</td>
<td>AMR speech Codec; Test sequences</td>
</tr>
<tr>
<td>TS 26.090</td>
<td>AMR speech Codec; Transcoding Functions</td>
</tr>
<tr>
<td>TS 26.091</td>
<td>AMR speech Codec; Error concealment of lost frames</td>
</tr>
<tr>
<td>TS 26.092</td>
<td>AMR speech Codec; comfort noise for AMR Speech Traffic Channels</td>
</tr>
<tr>
<td>TS 26.093</td>
<td>AMR speech Codec; Source Controlled Rate operation</td>
</tr>
<tr>
<td>TS 26.094</td>
<td>AMR Speech Codec; Voice Activity Detector for AMR Speech Traffic Channels</td>
</tr>
<tr>
<td>TS 26.101</td>
<td>AMR speech Codec; Frame Structure</td>
</tr>
<tr>
<td>TS 26.102</td>
<td>AMR speech Codec; Interface to Iu and Uu</td>
</tr>
<tr>
<td>TS 26.103</td>
<td>Codec lists</td>
</tr>
<tr>
<td>TS 26.104</td>
<td>AMR speech Codec; Floating point C-Code</td>
</tr>
</tbody>
</table>