GSM Network and Services

Voice coding

From voice to radio waves

voice/source coding

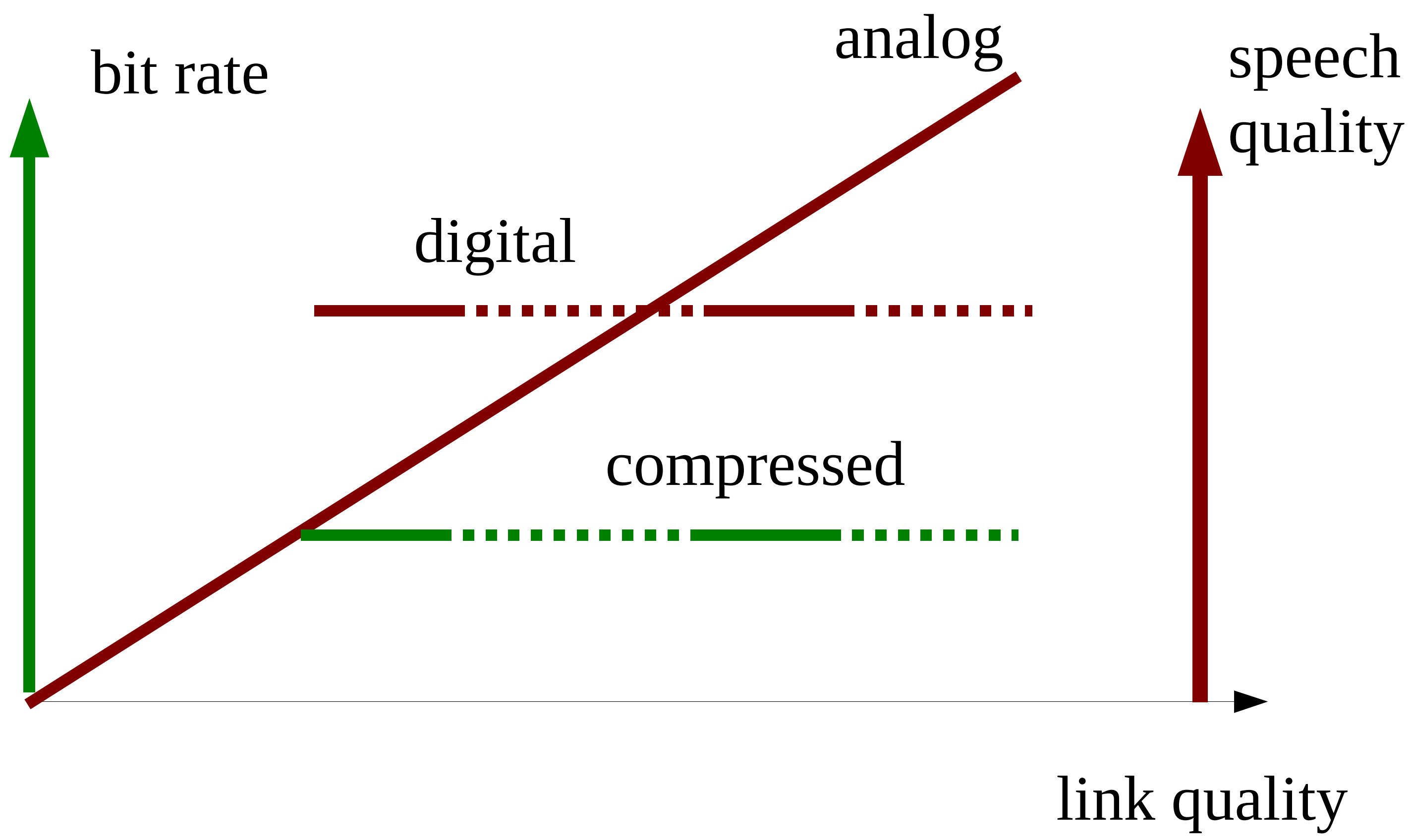
channel coding
- block coding
- convolutional coding
- interleaving

encryption

burst building

modulation
- diff encoding
- symbol coding
- carrier modulation

Why digital?


bit rate
analog
speech
quality
digital
compressed
link quality

Analog vs Digital

- If radio conditions are very good, an analog connection could do the job and in many cases a better job.
- Digital coding allows:
  - compression
  - error detection and error correction
  - half duplex implementation
  - adaption to radio conditions
- Digital drawback: processing, delay, voice only codecs.

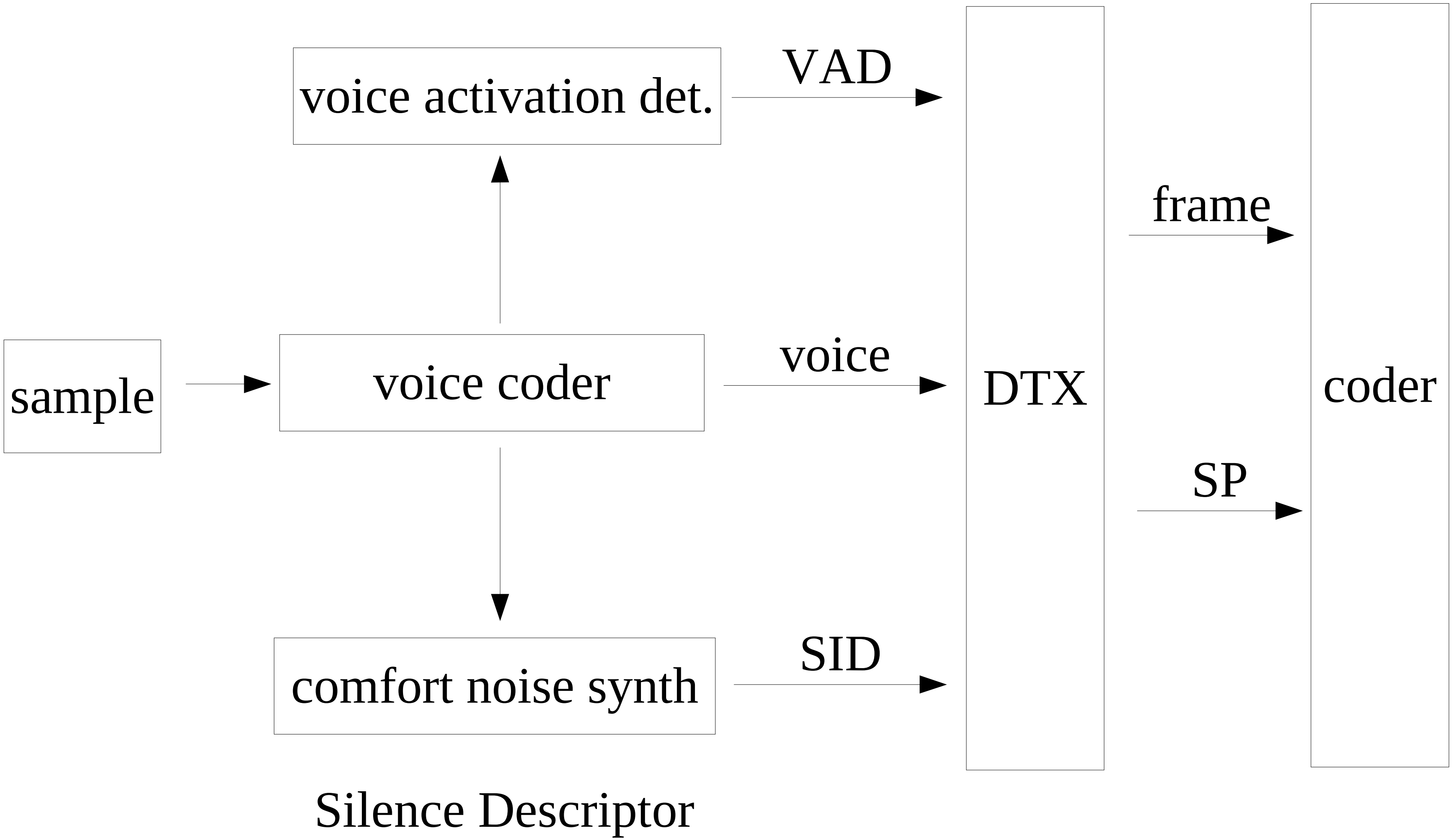
GSM voice codec

- Sampled at 8 kHz, 13 bits per sample. Compare this to regular A-law coding that samples at 8kHz using 8 bits per sample.
- Divides the sampled voice into 20 ms blocks for the codec to work on. Each block is coded in 260 bits resulting in 13 kb/s.
- The GSM codec is based on modeling of human speech organs. Tuned for males!
- Voice activation to allow discontinuous transmission.

Remember

- The 260 bits from the voice codec are encoded using 456 bits that are interleaved into eight bursts (half of each burst).
- A 26-multiframe last for 120 ms and holds 24 TCH burst, 1 SACCH burst and one idle frame.
- 24/4 = 6 voice samples every 26-multiframe.
- 26-multiframe takes 120 ms that is equal to 6 voice samples of 20 ms.
- One frame takes 120/26 = 4.6 ms.

GSM discontinuous transmission



DTX coder

- The coder will detect if the sample is voice or “silence”. When detected the Voice activity detection will set a VAD bit to 1.
- The voice coder will output a voice frame of 260 bits or a SID frame of 35 bits.
- Depending on the VAD bit the DTX will output voice frames or SID frames.
- The DTX will pass the voice frames or *in band encoded* SID frames to the channel coder.

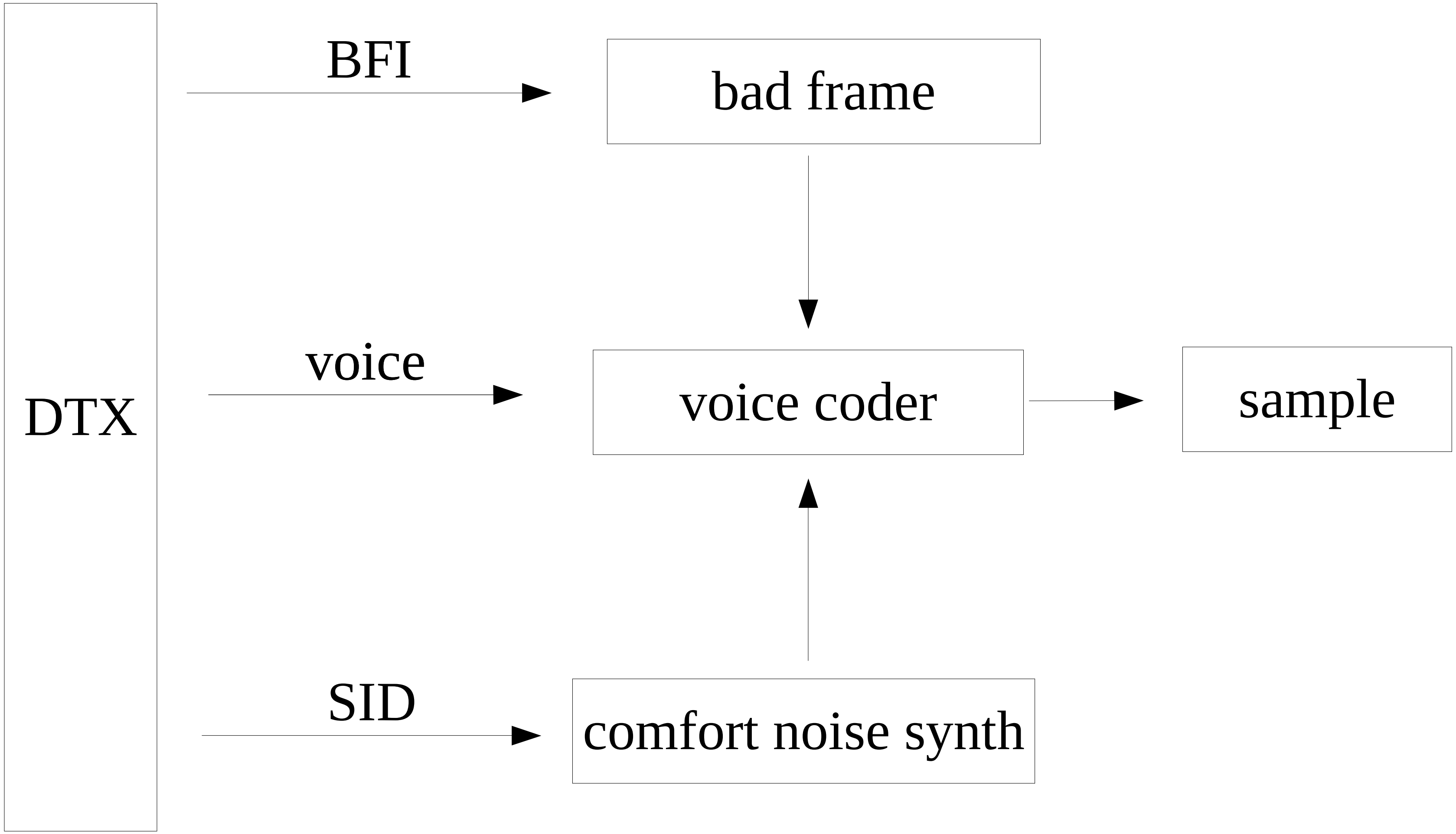
DTX encoder

- The channel coder will, based on the SP flag, control when data is actually sent over the air:
  - any frame that contains speech information
  - the first frame after a speech frame since it holds a SID frame
  - frames to keep the radio measurement active

Comfort noise

- If there is no noise in the background when you speak you will think that the connection is lost. This is extremely annoying
- Since phone conversations are almost always half duplex it is wasteful to keep a full duplex connection open.
- Why? Wasting what?

DTX decoder


DTX
BFI
voice
SID
bad frame
voice coder
comfort noise synth
sample

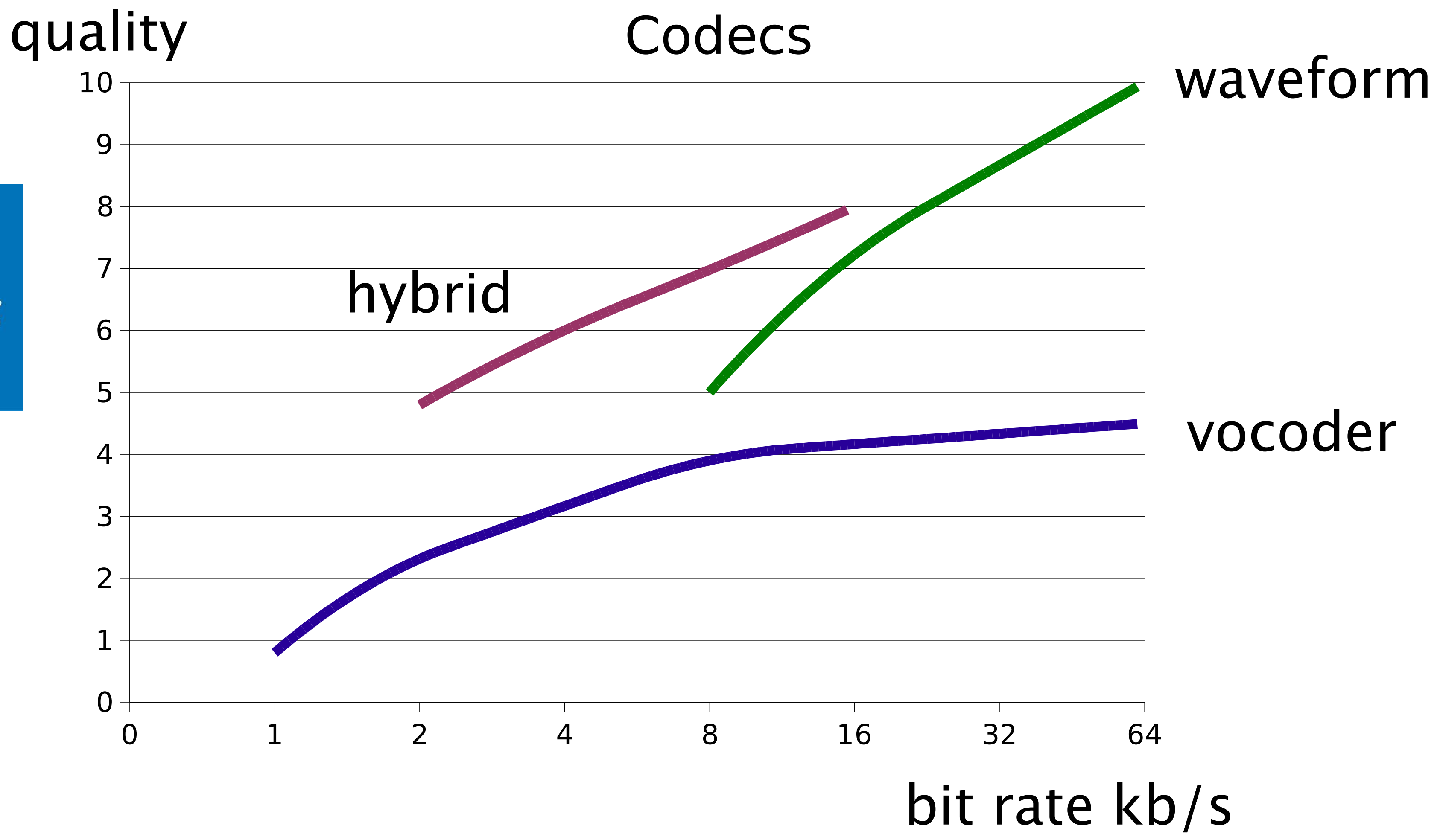
DTX decoder

- Correct voice frames are passed directly to the voice decoder.
- Correct SID frames are passed to the comfort noise synthesizer.
- If the Bad Frame Indicator (BFI) is set then
  - an incorrect voice frame is replaced by the previous
  - an incorrect SID frame is replaced by the last valid SID frame or last valid speech frame (that probably contains noise)

Waveform vs Vocoder

- Waveform codecs encode audio with no knowledge of the signal.
  - Pulse Code Modulation (PCM) is one example that is used in fixed line telephony.
- Vocoder, or source coders, try to represent the audio generator, in this case the vocal tract, and represent a sample as a the parameters that define the generator.
  - Vocoders can encode speech at a very low bit rate. Even if the result does not sound like the original audio signal.

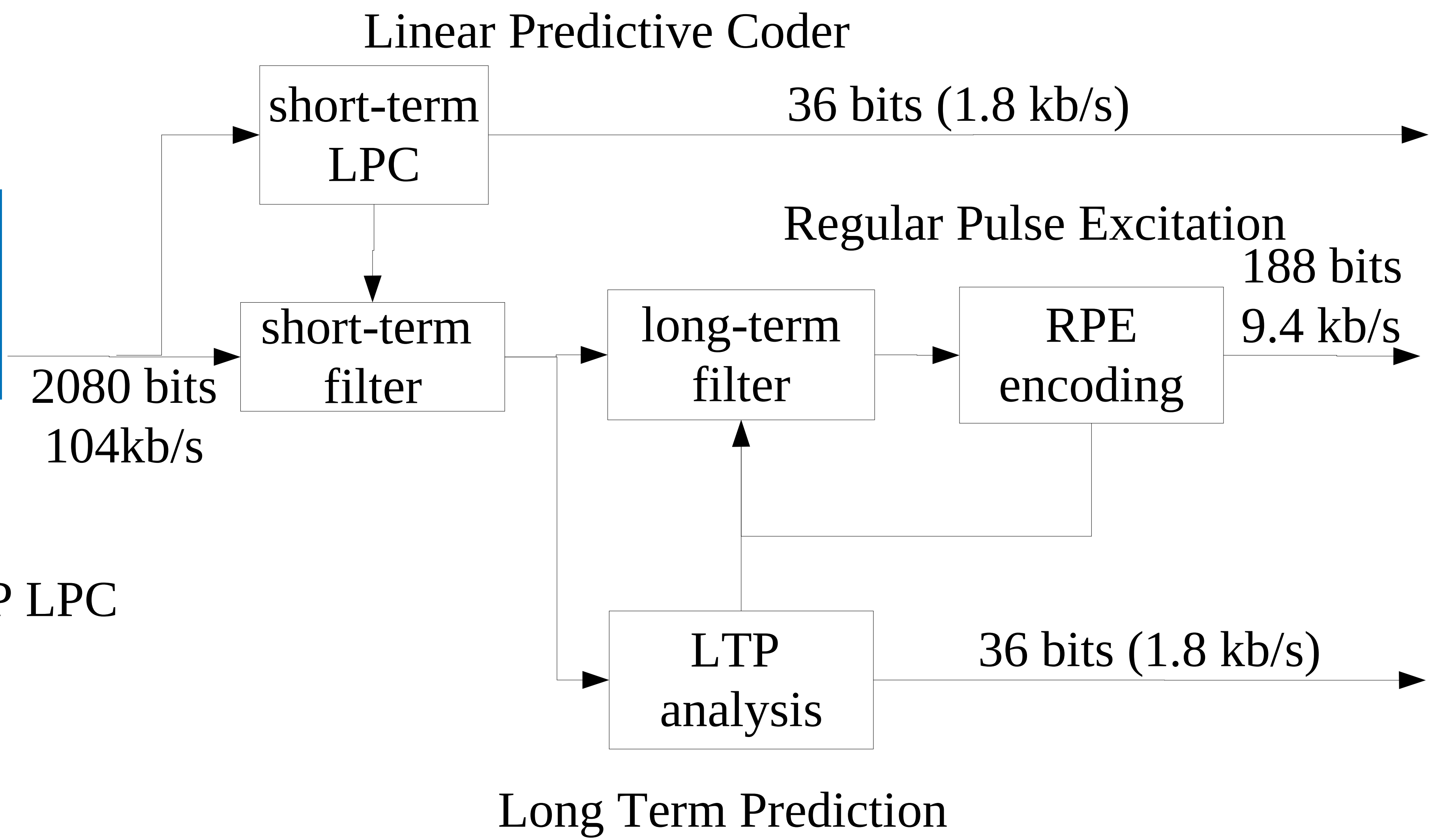
Voice codec


quality
Codecs
10
9
8
7
6
5
4
3
2
1
0
0
1
2
4
8
16
32
64
waveform
hybrid
vocoder
bit rate kb/s

Hybrid coder

- In a hybrid coder a vocoder is used to encode the voice signal.
- The decoded sample of the result is then compared to the original sample.
- The difference between the original and the decoded result is encoded using a waveform coder.
- The result is a high quality low bit rate codec.

GSM codec

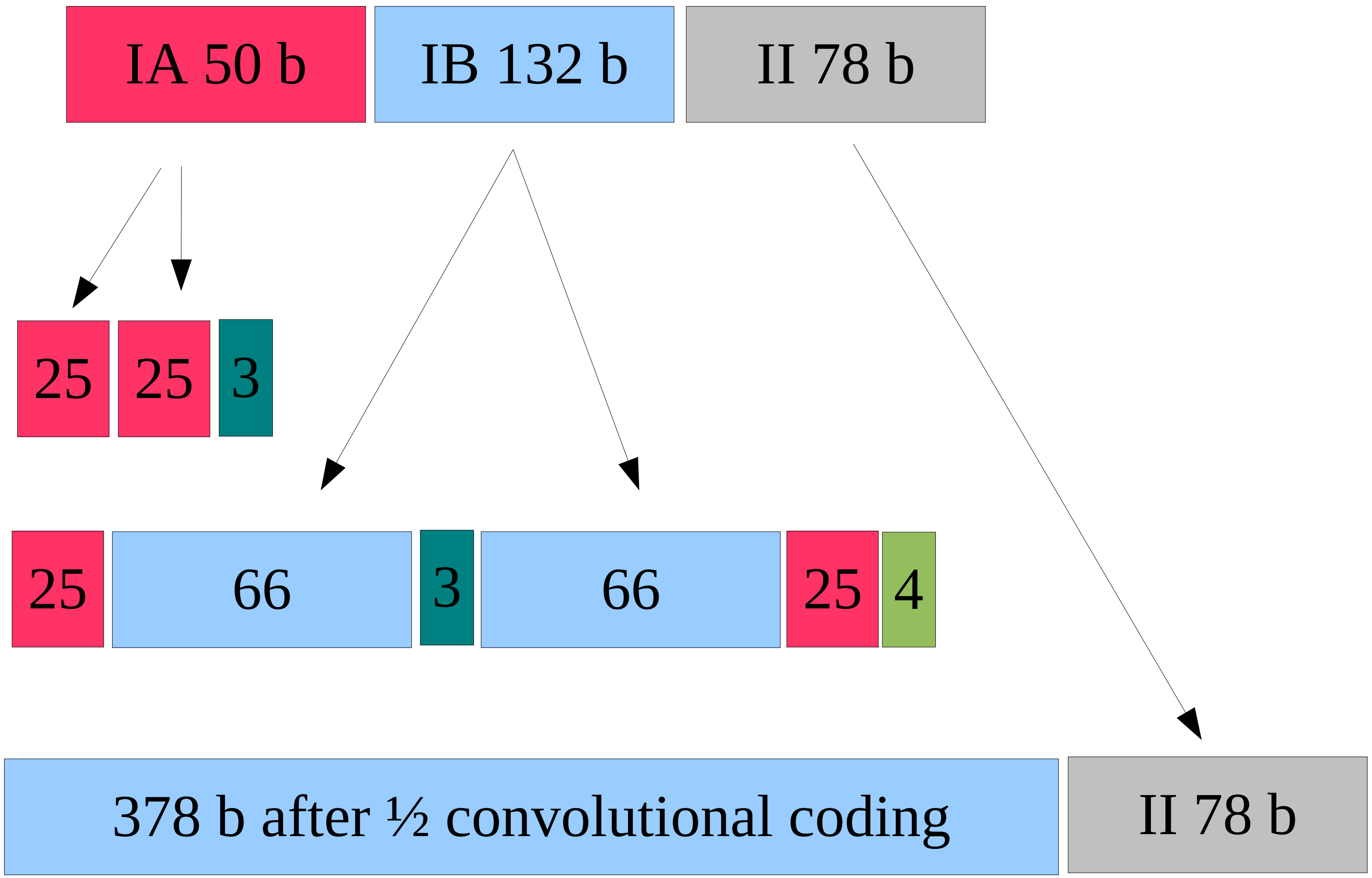


RPE-LTP LPC

GSM codec

- The resulting 260 bits are divided into:
  - Class 1A: 50 most significant bits
  - Class 1B: 132 significant bits
  - Class 2: 78 less significant bits
- Classing is done by subjective evaluation of the perceived speech quality (*Mean Opinion Score, MOS*).

Channel coding



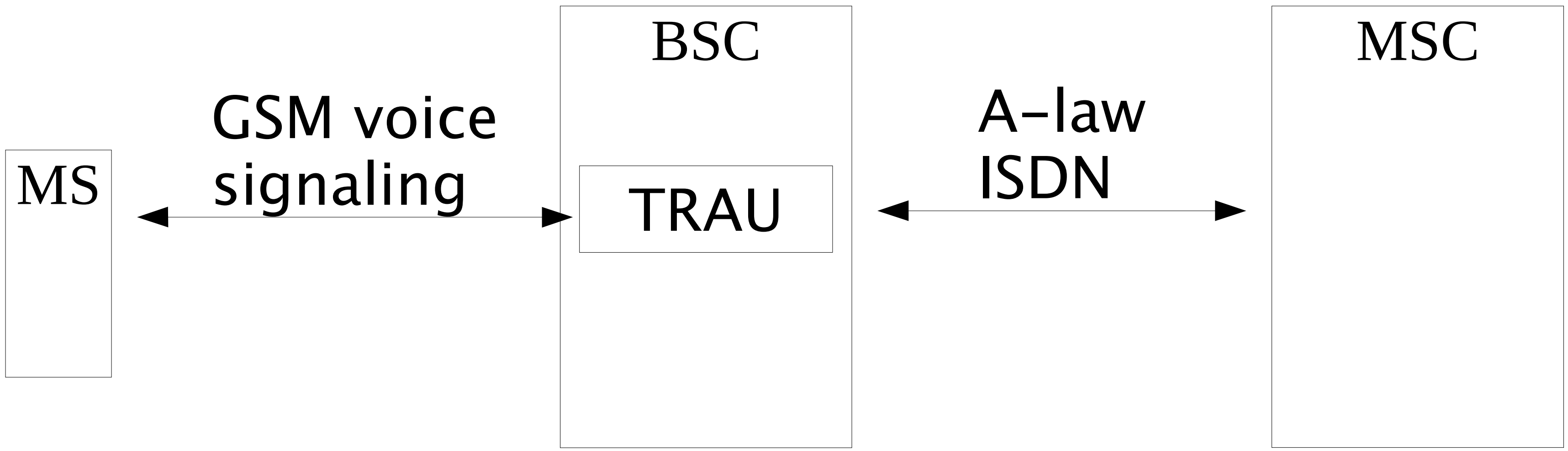
Other codecs

- Half rate (HR): 6.5 kb/s used to increase capacity since it only needs a half rate TCH. Speech quality is considerably less.
- Enhanced Full Rate (EFR): 12.2 kb/s but further protected by CRC in a resulting 13 kb/s codec. Better speech quality.
- Adaptive Multi Rate (AMR): the codec adapts to the radio conditions; high BER will use a low rate codec that is better protected .

AMR

kb/s	12.2	10.2	6.7	4.75
Block	244	204	134	95
Class 1A	81	65	55	39
Rate	“1/2”	“1/3”	“1/4”	“1/5”
Raw	508	642	576	535
Puncturing	60	194	128	87
Result	448	448	448	448
In band	8	8	8	8
Frame	456	456	456	456

BSS side coder



The TRAU needs not only the 13kb/s voice stream but also BFI information etc. 3 kb/s signaling and error detection is added resulting in a 16 kb/s stream.

4 streams will share a 64 kb/s ISDN channel