



Voice Coding with Opus

Koen Vos, Karsten Vandborg Sørensen,
Søren Skak Jensen, Jean-Marc Valin

Two Opus presentations

- This talk: Voice Mode (Koen)
 - Features
 - Technology
 - Listening test results
- Next talk: Audio Mode (Jean-Marc)

What is Opus?

- Flexible speech and audio codec
- Best-in-class performance across a wide range of applications
- IETF Standard RFC 6716 (Sep. 2012)
- Royalty free
- Open source

Flexible Indeed

- Bitrates from 6 to 510 kbps
- Frame sizes from 2.5 to 60 ms
- Narrowband to full-band (in 5 steps)
- Speech and music
- Mono and stereo
- Rate control
- Variable complexity

All changeable dynamically, signalled within the bitstream

Merging Two Codecs

1. SILK

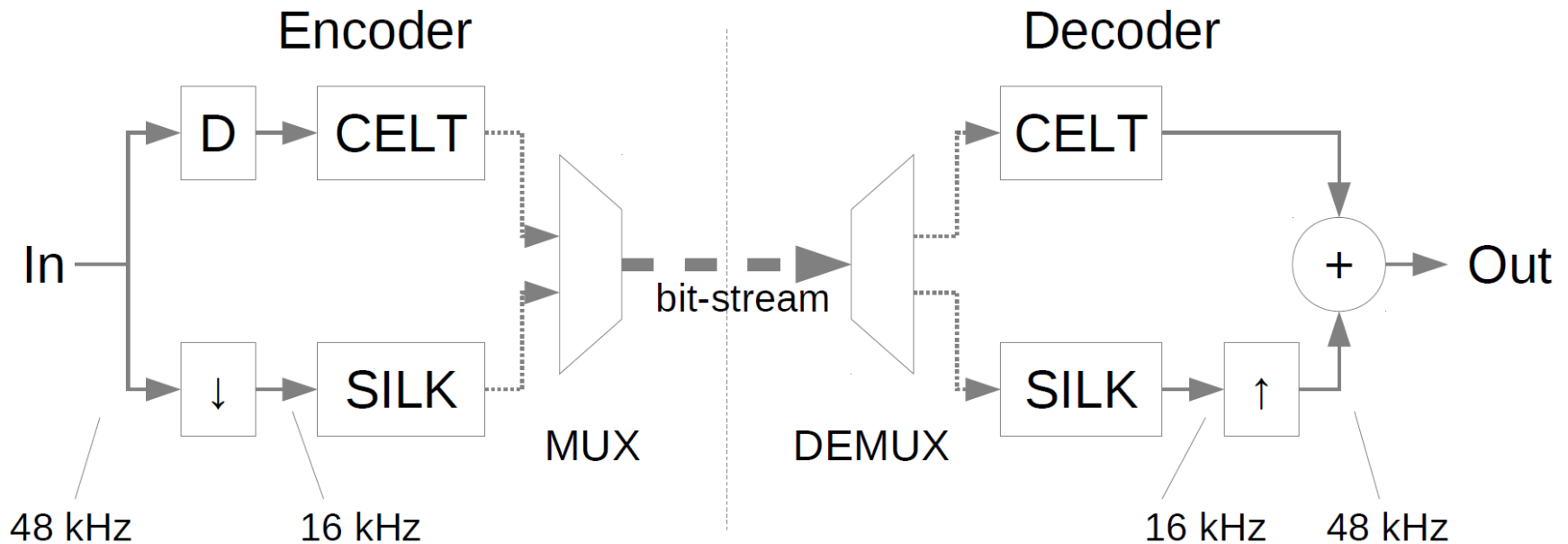
- Developed by Skype
- Based on Linear Prediction
- Efficient for voice
- Up to 8 kHz audio bandwidth

2. CELT

- Developed by Xiph.Org
- Based on MDCT
- Good for universal audio/music

Hybrid Mode

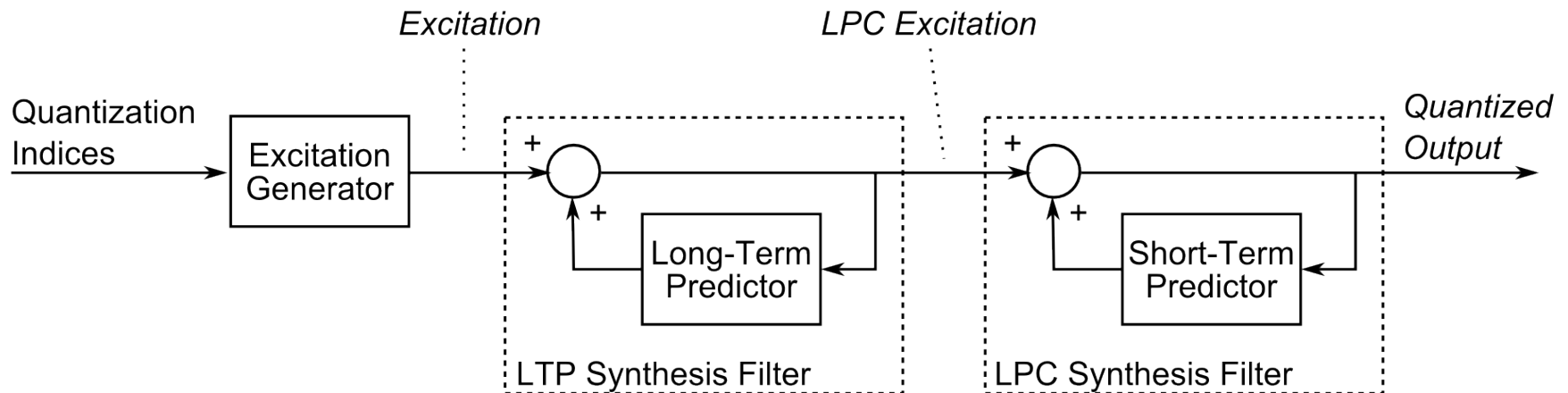
For super-wideband or full-band voice



SILK Decoder

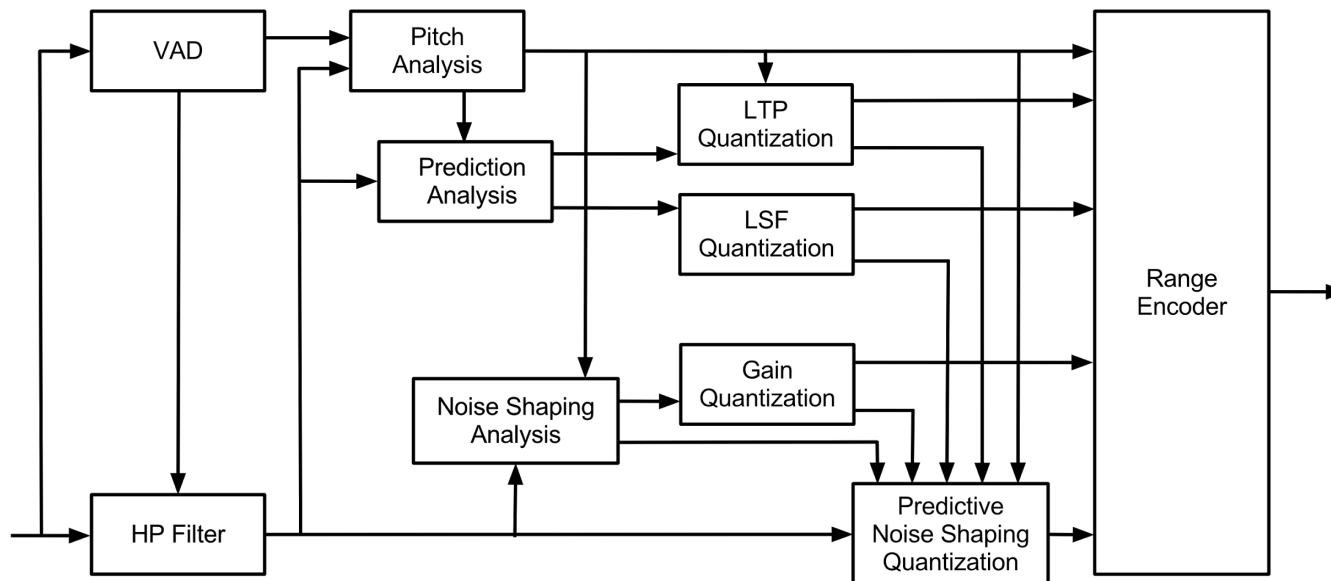
Standard defines only the decoder

- Doesn't get much simpler



SILK Encoder

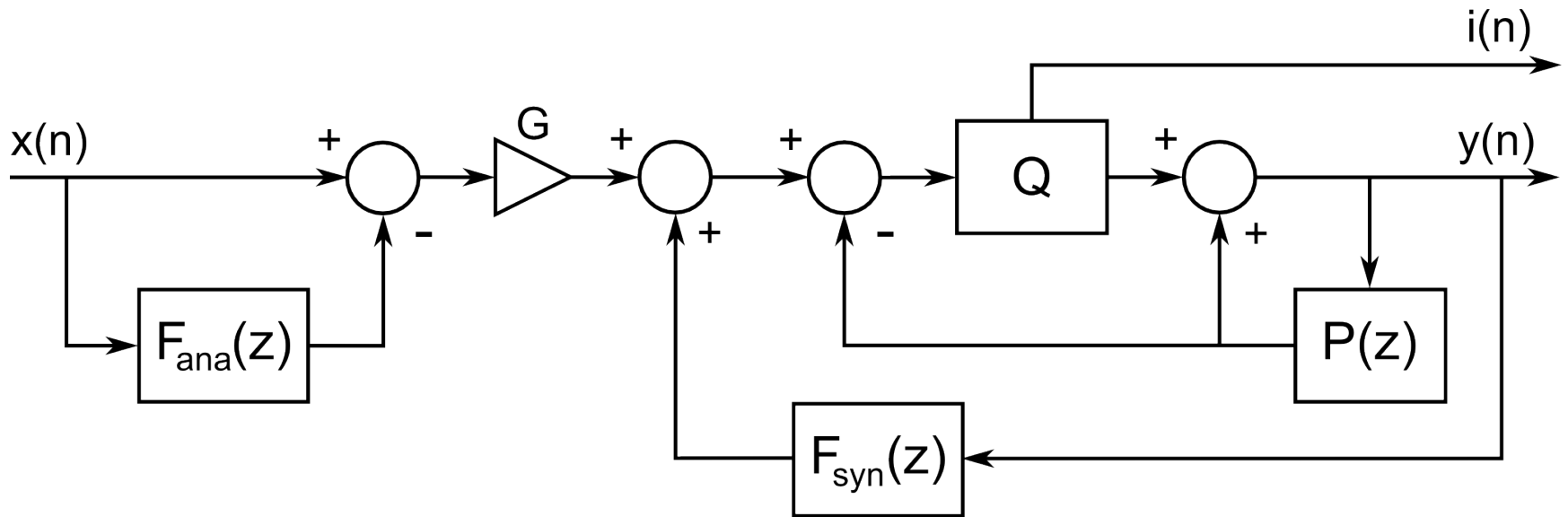
Standard includes high-quality reference implementation



Predictive Noise Shaping Quantization

- Linear short- and long-term prediction to model formants and harmonics
 - Reduce entropy of residual
- Short- and long-term emphasis filtering
 - Emphasize important spectral components
 - Reduce input noise
- Short- and long-term noise shaping
 - Mask quantization noise

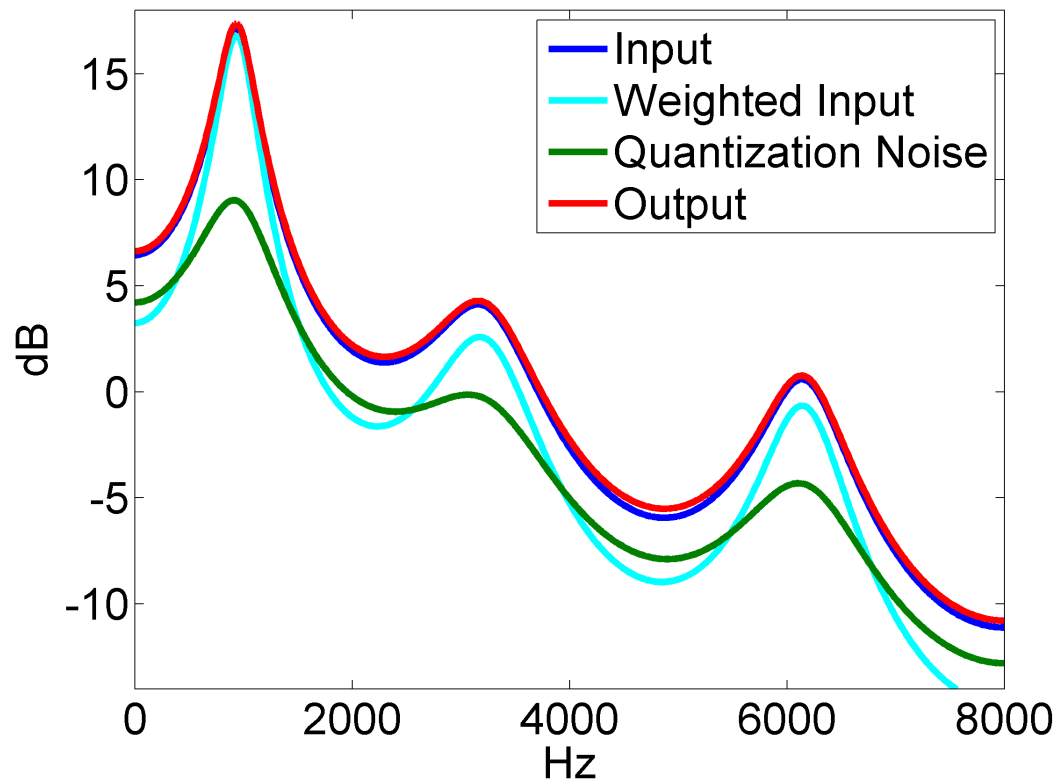
Predictive Noise Shaping Quant. II



$$Y(z) = G \cdot \frac{1 - F_{ana}(z)}{1 - F_{syn}(z)} \cdot X(z) + \frac{1}{1 - F_{syn}(z)} \cdot Q(z)$$

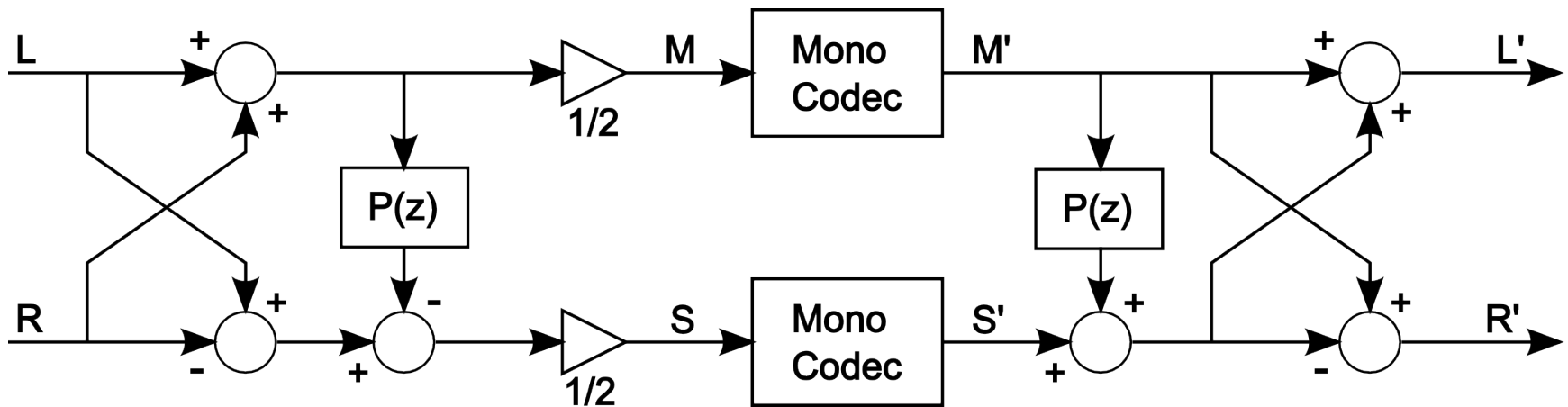
Predictive Noise Shaping Quant. III

Example (short-term shaping only)



Stereo

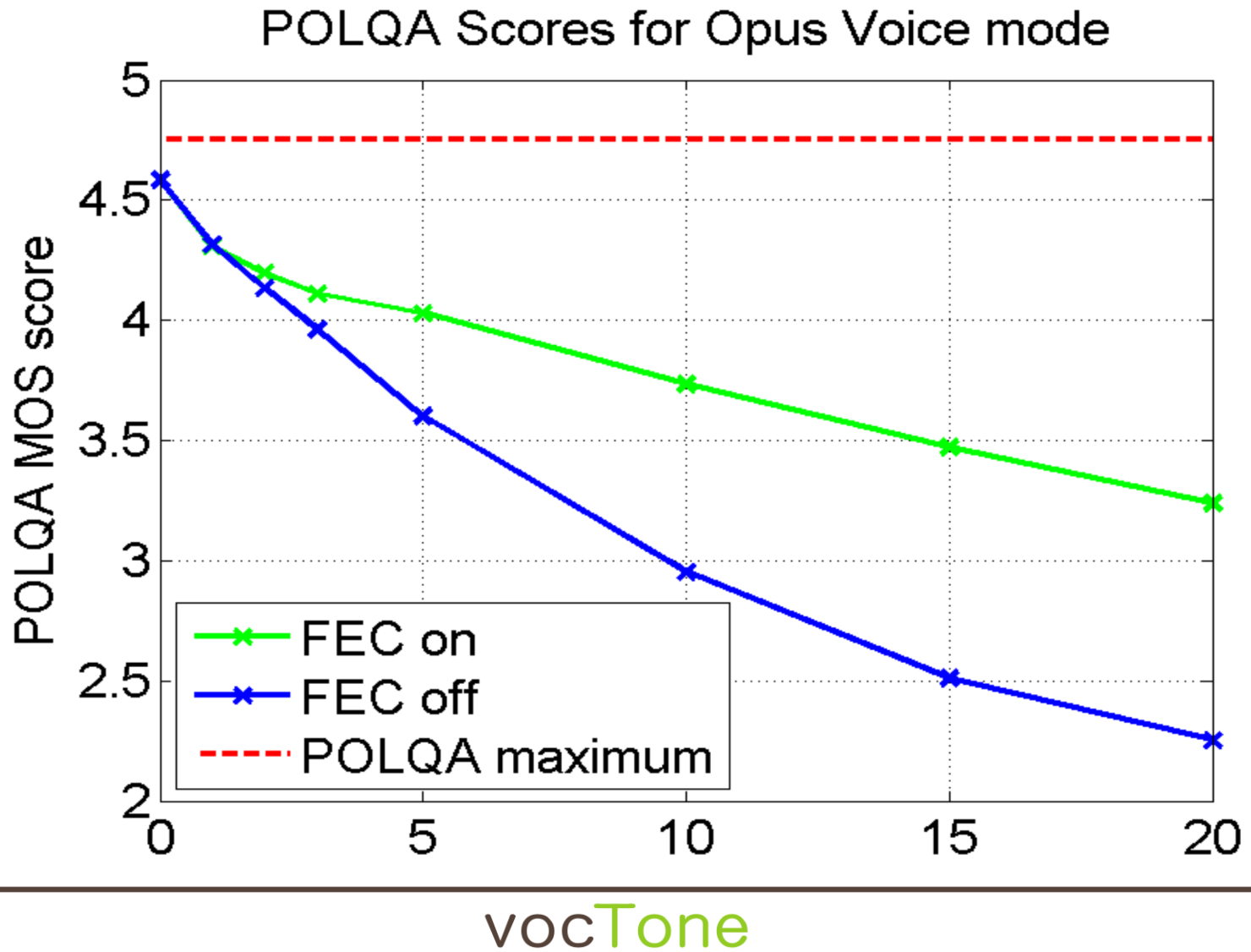
- Mid-Side representation
- Side is predicted from mid; residual coded



Internet Robustness

- **Forward Error Correction (FEC)**
 - Include coarse encoding of previous packet, for active speech
- **Flexible Error Propagation**
 - Code packets more independently for channels with packet loss
- **Discontinuous Transmission (DTX)**
 - Reduce packet rate during silence
- **Packet Loss Concealment (PLC)**
 - Decoder side
 - Fills in DTX blanks

FEC

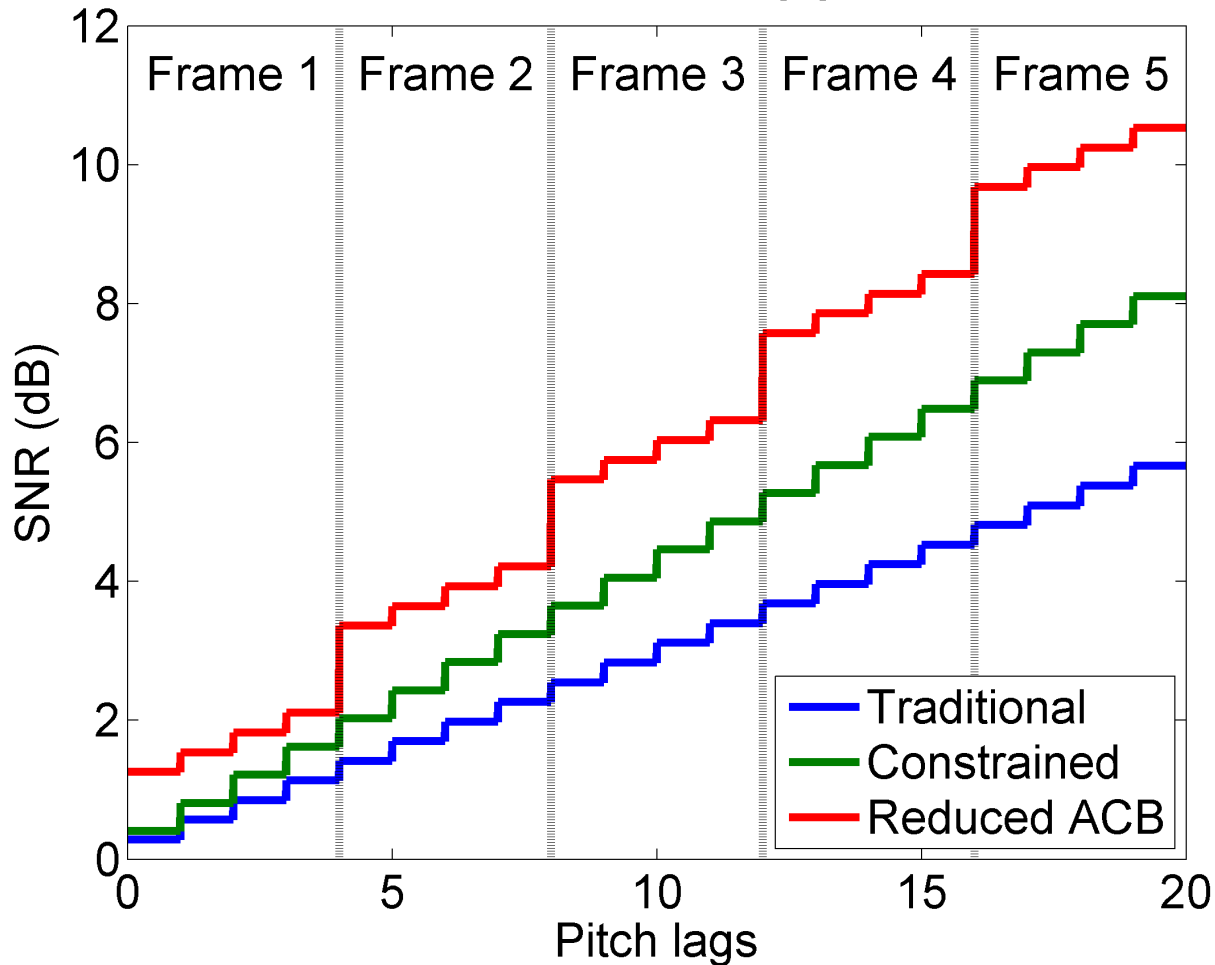


Flexible Error Propagation

- Reduce LTP filter state at beginning of a packet, in encoder and decoder
- Spend more bits only during first pitch period
- Other codecs constrain LTP filter coefficients and spend more bits throughout the packet

Effect of LTP scaling

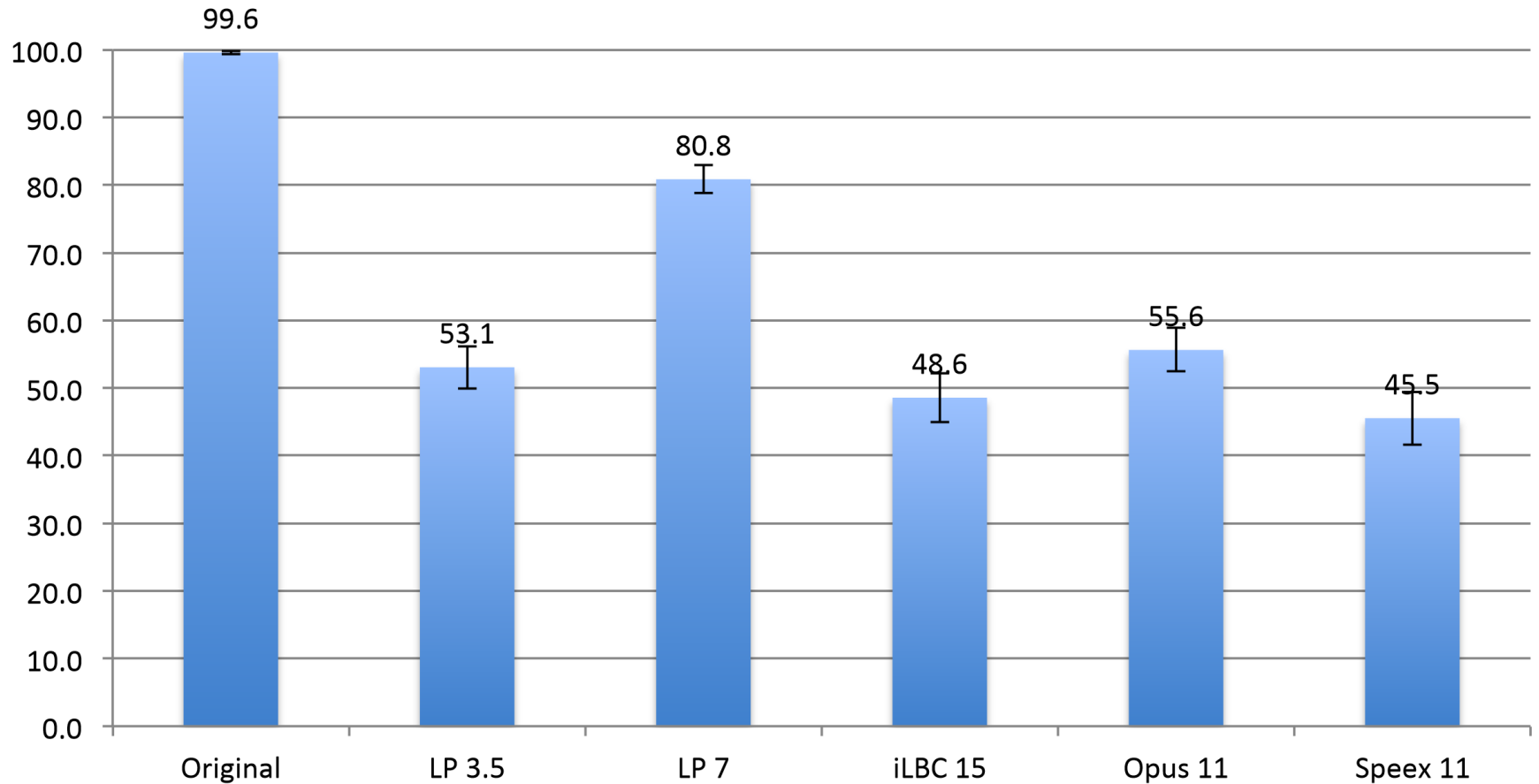
12 dB LTP Coding gain



Packet Loss Example

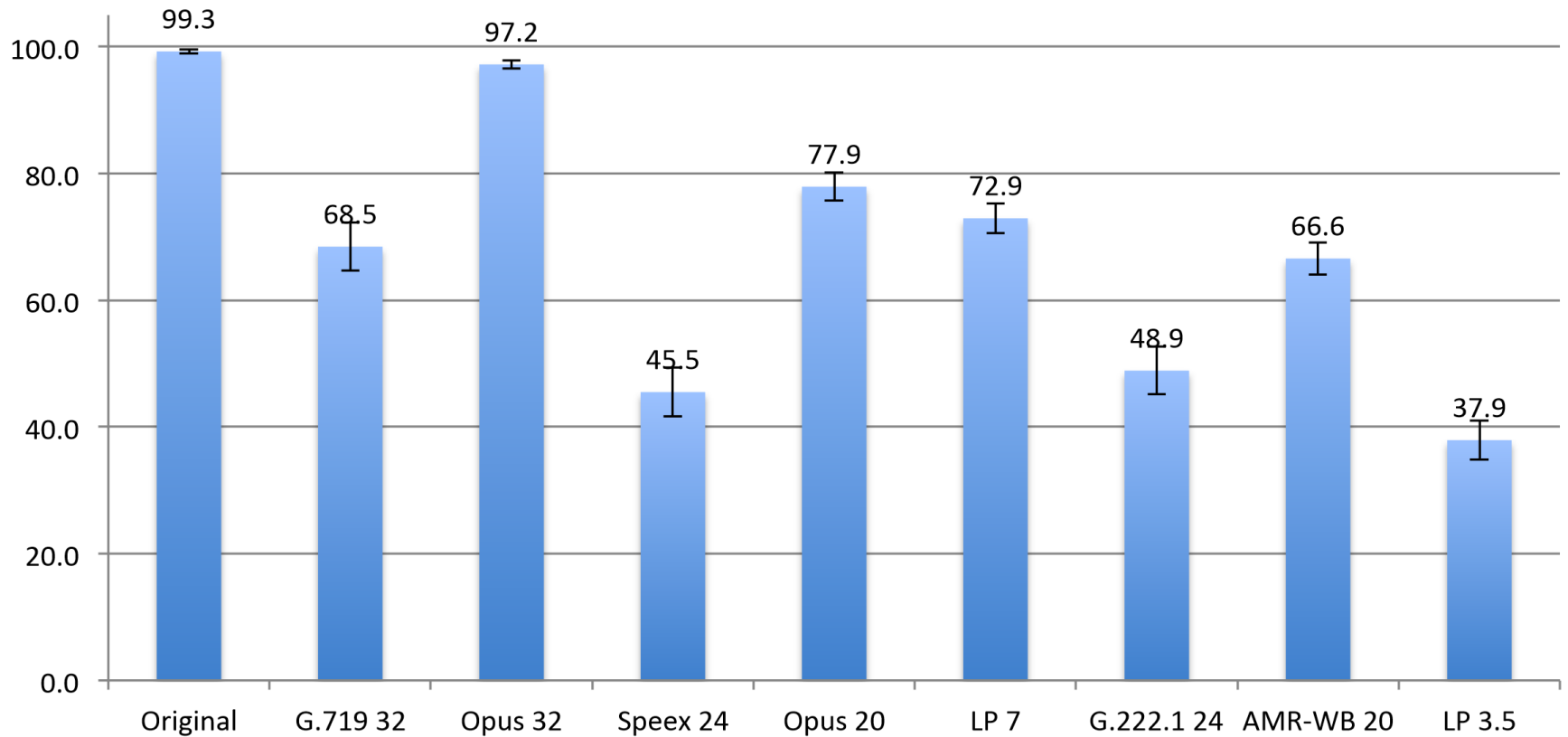
- Original
- AMR-WB, 30% packet loss
- Opus without FEC, 30% packet loss
- Opus with FEC, 30% packet loss

Listening Results: Narrowband



Google Mushra Test

Listening Results: Wide/Full-Band



Google Mushra Test

Questions?

Find all things Opus at <http://www.opus-codec.org>