ITU-T Workshop on
"From Speech to Audio: bandwidth extension, binaural perception"
Lannion, France, 10-12 September 2008

An Open-Source Softphone for Musicians Playing over IP

Mathias Bartl, Dr. Christian Hoene

University of Tübingen

Content

- **Motivation**
  for an open source, ultra low delay audio coding
- Extending the packet loss concealment algorithm
  ITU G.711 Appendix 1 to full bandwidth
- Comparing ultra low delay audio codecs
- Summary

From speech to ultra-low delay audio

- Telephone usage has long traditions but will they prevail?

- Some new usage scenarios for the telephone
  - luxury and high definition telephone calls
  - sharing iPOD listening experiences
  - playing ego shooter games together
  - performing and playing instruments over the network.

- Musicians playing over the network have the highest demands regarding latency and quality.

Latency requirements: Faster than sound

- Musicians need to hear and synchronize to each other
- Musicians sit less than 8 meters apart
  - Speed of sound: 344 m/s.
  - Delay at 8 meters: 25 ms.
- If latency >25ms, the orchestra needs a conductor
- Physical limit: No distances greater than 4000km

Algorithmic latency of the codec must be extremly small!

Do we need open source?

- Arguments
  - Most VoIP soft-phones can be downloaded for free (Skype, Netmeeting, Ekiga, …)
  - One can setup a small SIP-based telephone switching center with open source software.
  - Many line based telephone calls are for free
- Do I have to pay the IPR license fee of my audio codec?

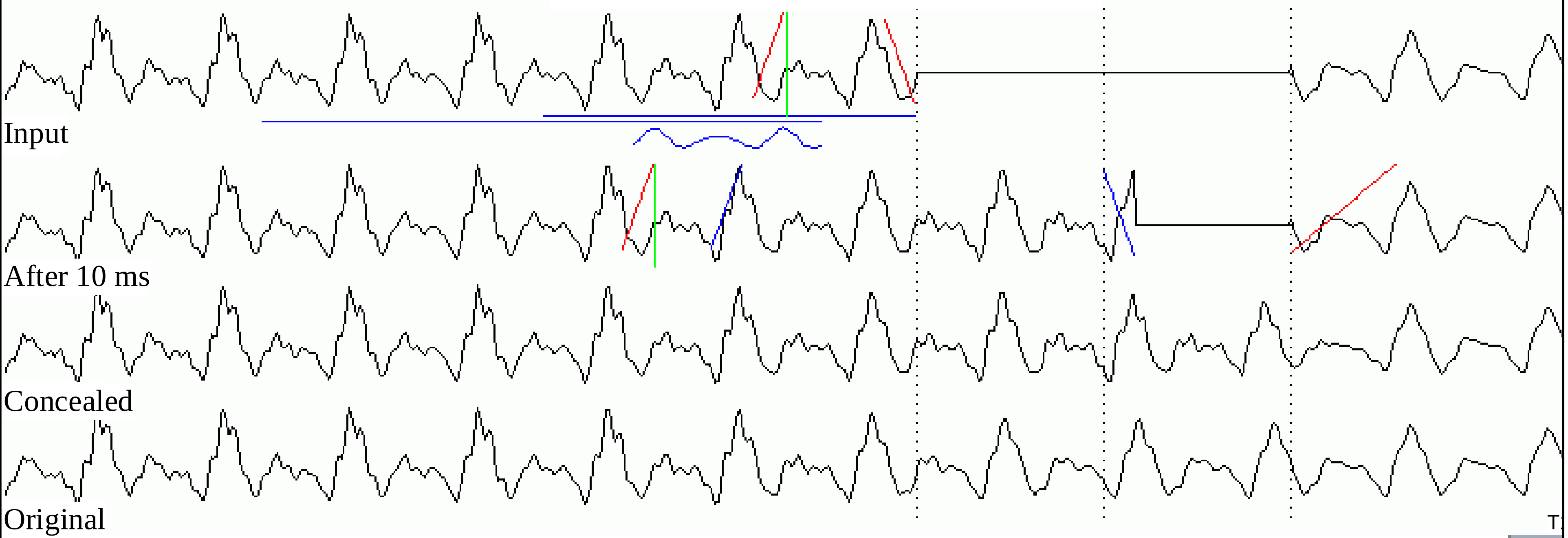
We shall consider open source codecs without IPR fees.

Content

- Motivation
  for an open source, ultra low delay audio coding
- **Extending the packet loss concealment algorithm ITU G.711 Appendix I to full bandwidth**
- Comparing ultra low delay audio codecs
- Summary

Full-bandwidth and IP support for ITU-T G.711 Appendix I Packet Loss Concealment

- G.711 Appendix I implements „a high quality low-complexity algorithm for packet loss concealment".
- Indented for speech compressed with µ-law and A-law at 8000 Hz received in blocks of 10 ms.
- Uses an extrapolation based on Reverse Order Replicated Pitch Periods (ROPRPP) algorithm.



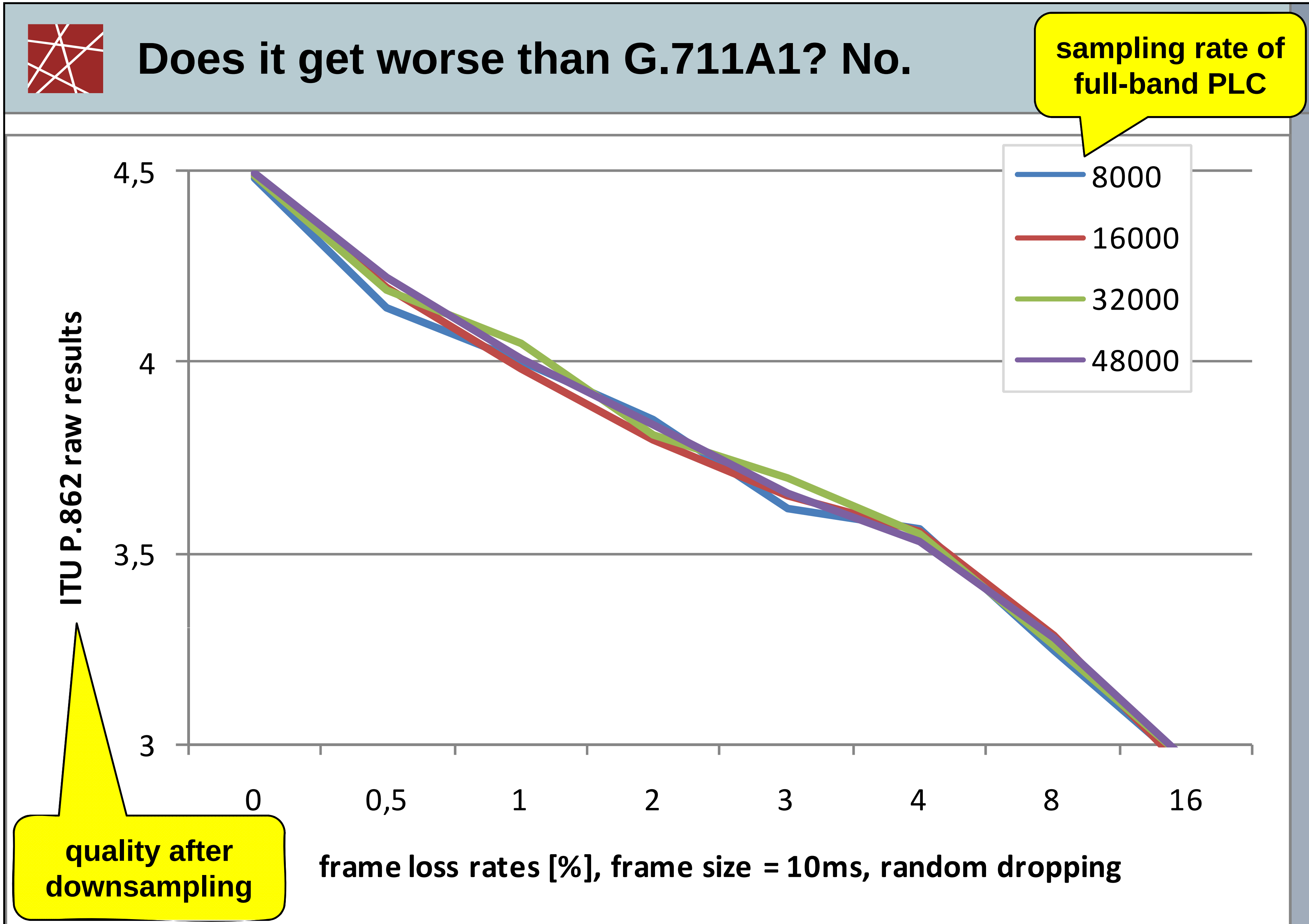
Our Enhancements

1. Support of any sampling rate
   - Pitch detection is done in two steps:
   1. The pitch is found with a crosscorrelation over a decimated signal having a sampling frequency of about 4000 Hz
   2. Second, the fine grade pitch period is done at the full sampling rate.
2. Support of arbitrary block sizes
   - The receiver cannot control the packet size, thus it has to cope with any size.
   - We changed the algorithm to work on samples instead of frames.
3. Support of variable playout jitter
   1. Delaying playout by repeating the last pitch period
   2. Faster playout by dropping the next pitch period
   3. Faster playout after frame loss
   4. Variable algorithmic delay (even smaller than 3,75 ms).

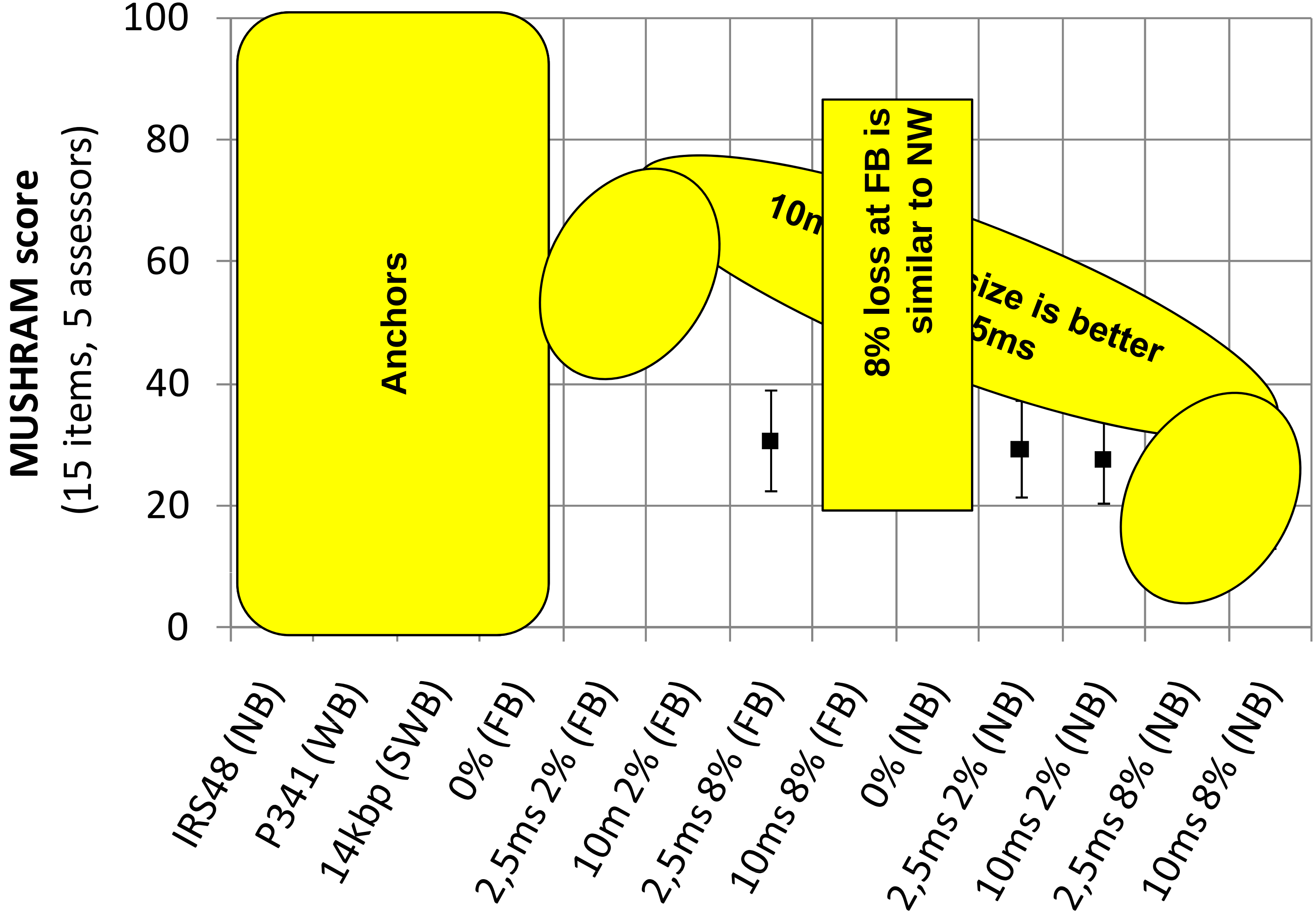
Subjective and Objective Evaluation

- Samples taken from ITU BS.1387 and „Kiel Corpus Vol. 1“
- Objective assessment:
  - TU P.862 (NB/WB) and ITU BS.1387
- Subjective tests: following Mushra test description (ITU-R BS.1534-1)
  - Headphone Sennheiser HD 280 PRO
  - Software “MUSHRAM 1.0“ (by E. Vincent)
  - calculated ANOVA and 95% confidence interval
  - Anchors: NB-IRS48, WB-P341, SWB-14kbps made with G.191 software

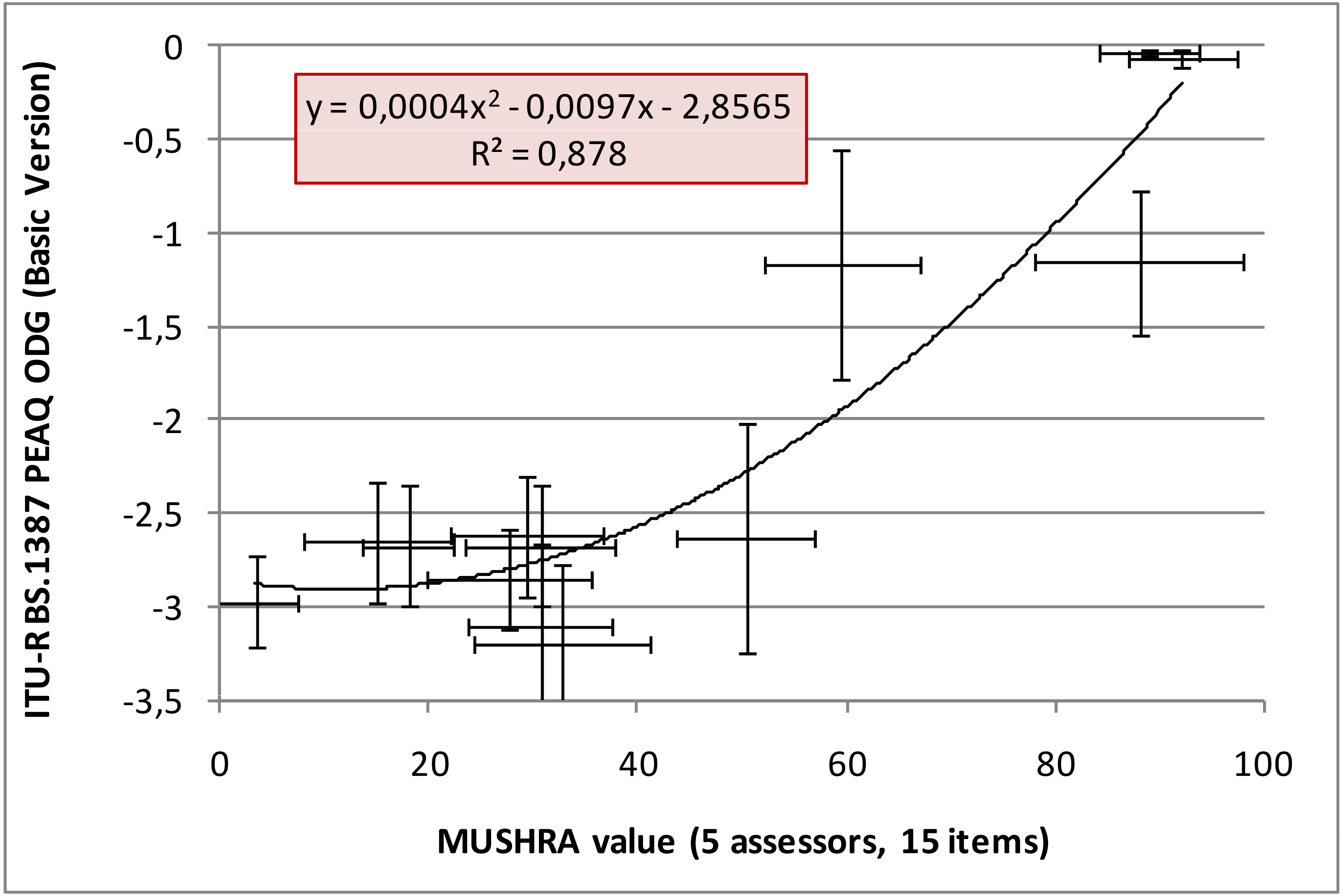
Does it get worse than G.711A1? No.



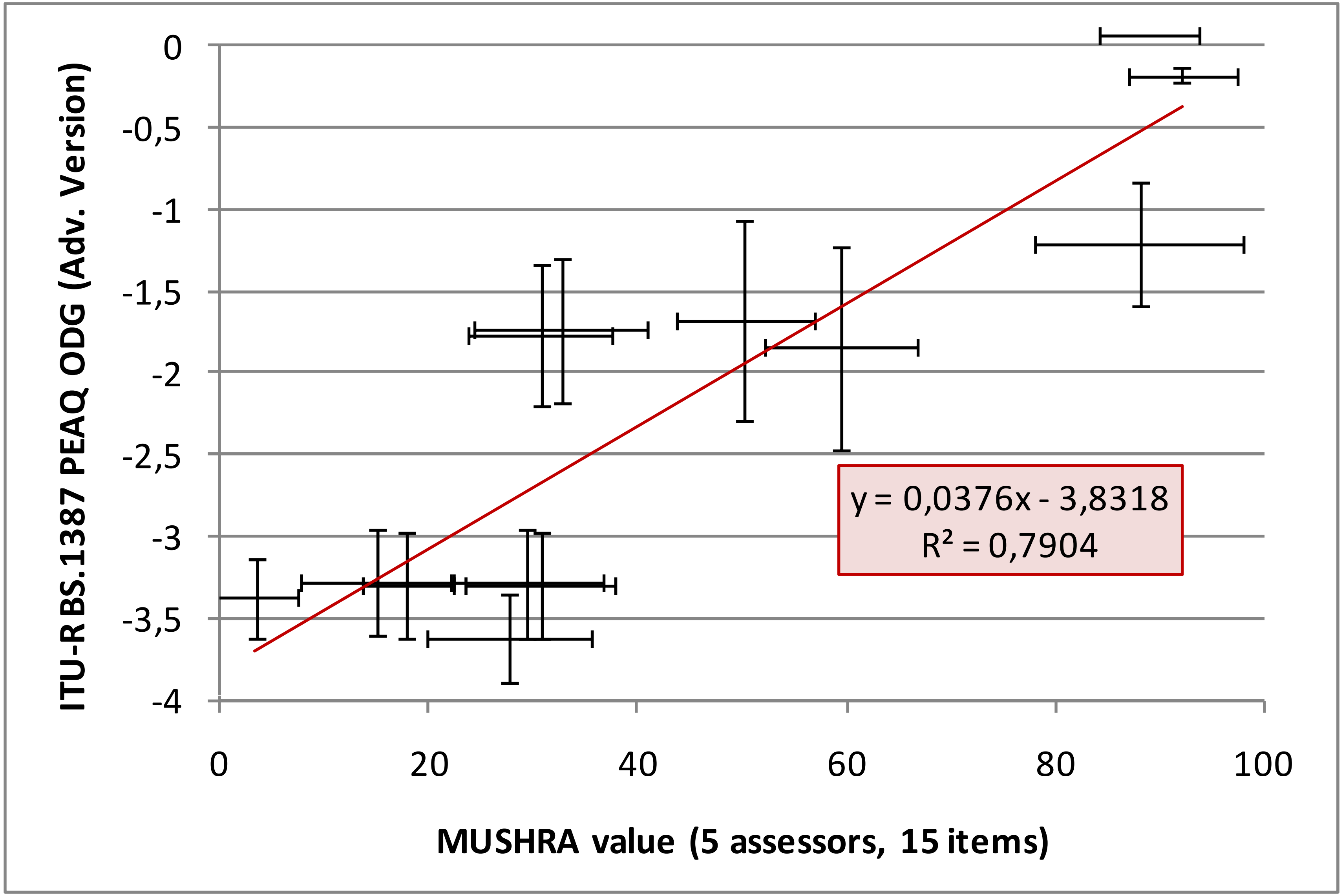
MUSHRA Listening Results



Comparing MUSHRA with PEAQ (BV) based on Samples with Concealed Packet Losses



Comparing MUSHRA with PEAQ (AV)



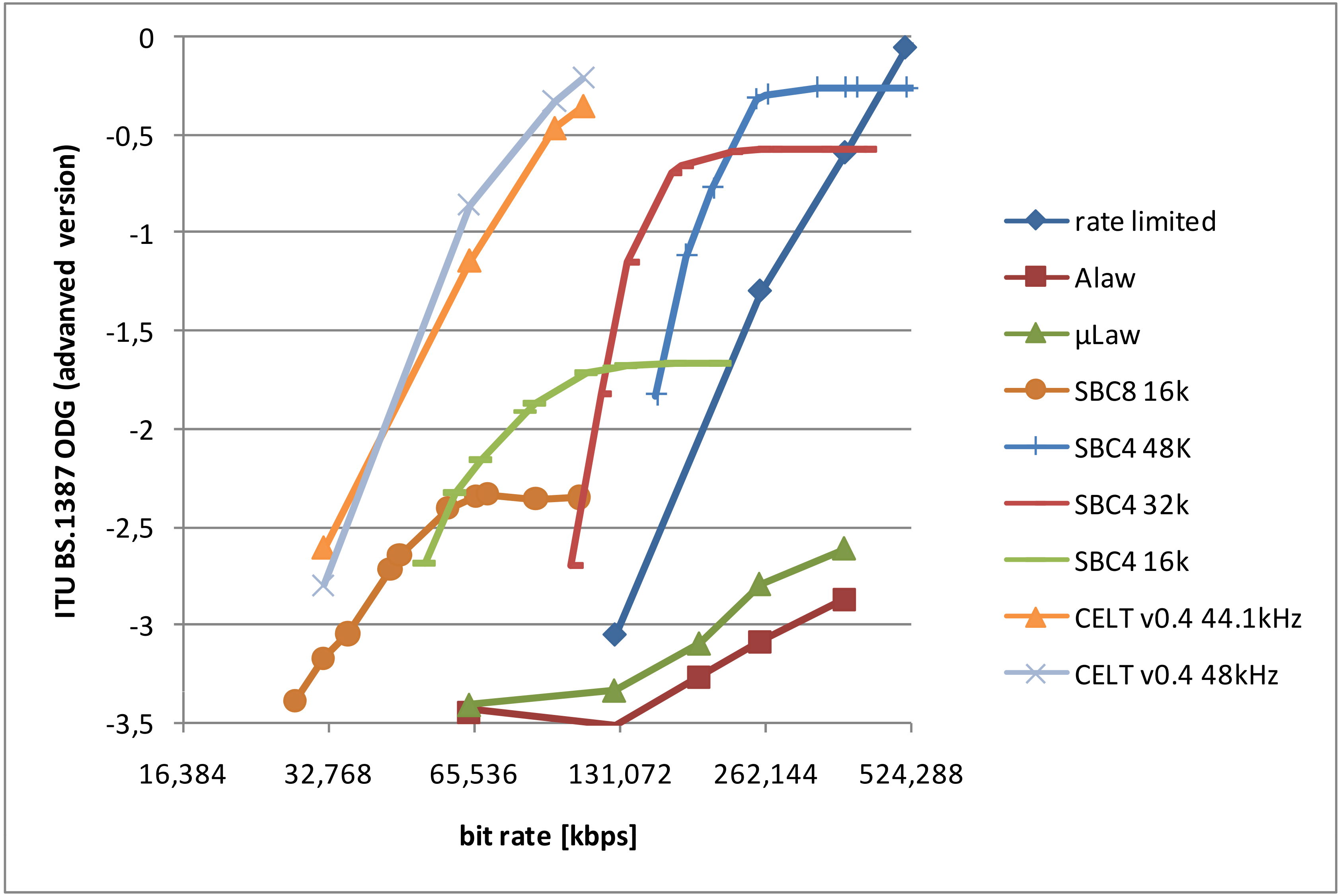
Content

- Motivation
  for an open source, ultra low delay audio coding
- Extending the packet loss concealment algorithm
  ITU G.711 Appendix 1 to full bandwidth
- **Comparing ultra low delay audio codecs**
- Summary

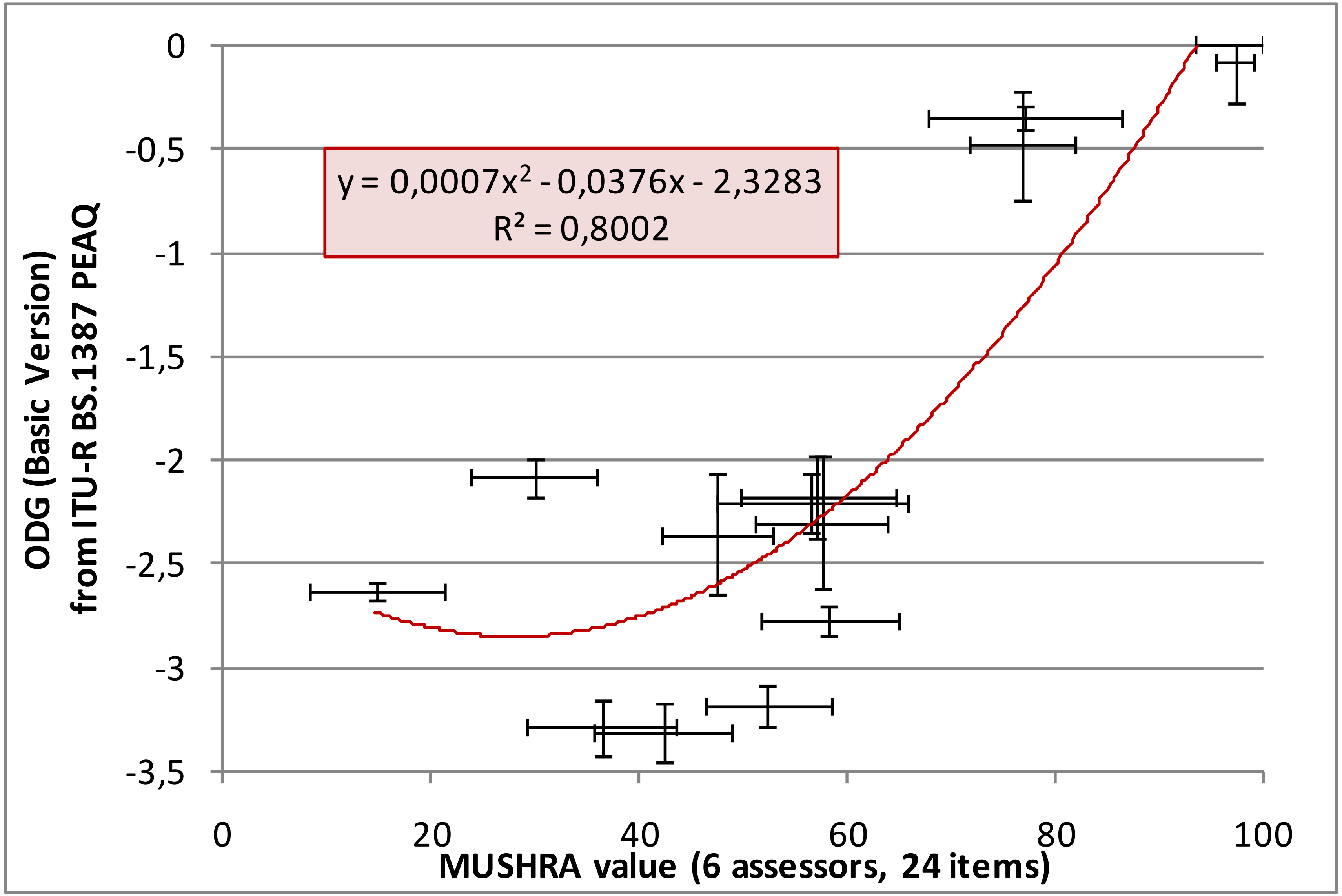
Overview on ultra-low delay audio codecs

1. PCM at different bandwidths
2. µ- and A law
   - by Jayant and Noll (original intended also for video)
3. Ultra-low delay audio codec (ULD)
   - by Gerald Schuller, Fraunhofer IDMT
   - Coverted by patents, currently not available
4. Bluetooth Subband Codec (SBC)
   - by Frans de Bont, Philipps
   - Intended for Bluetooth headphones
   - Covered by patents but free to use for Bluetooth applications
5. Constrained-Energy Lapped Transform (CELT)
   - by Jean-Marc Valin, Xiph.Org Foundation
   - http://www.celt-codec.org/
   - currently in development

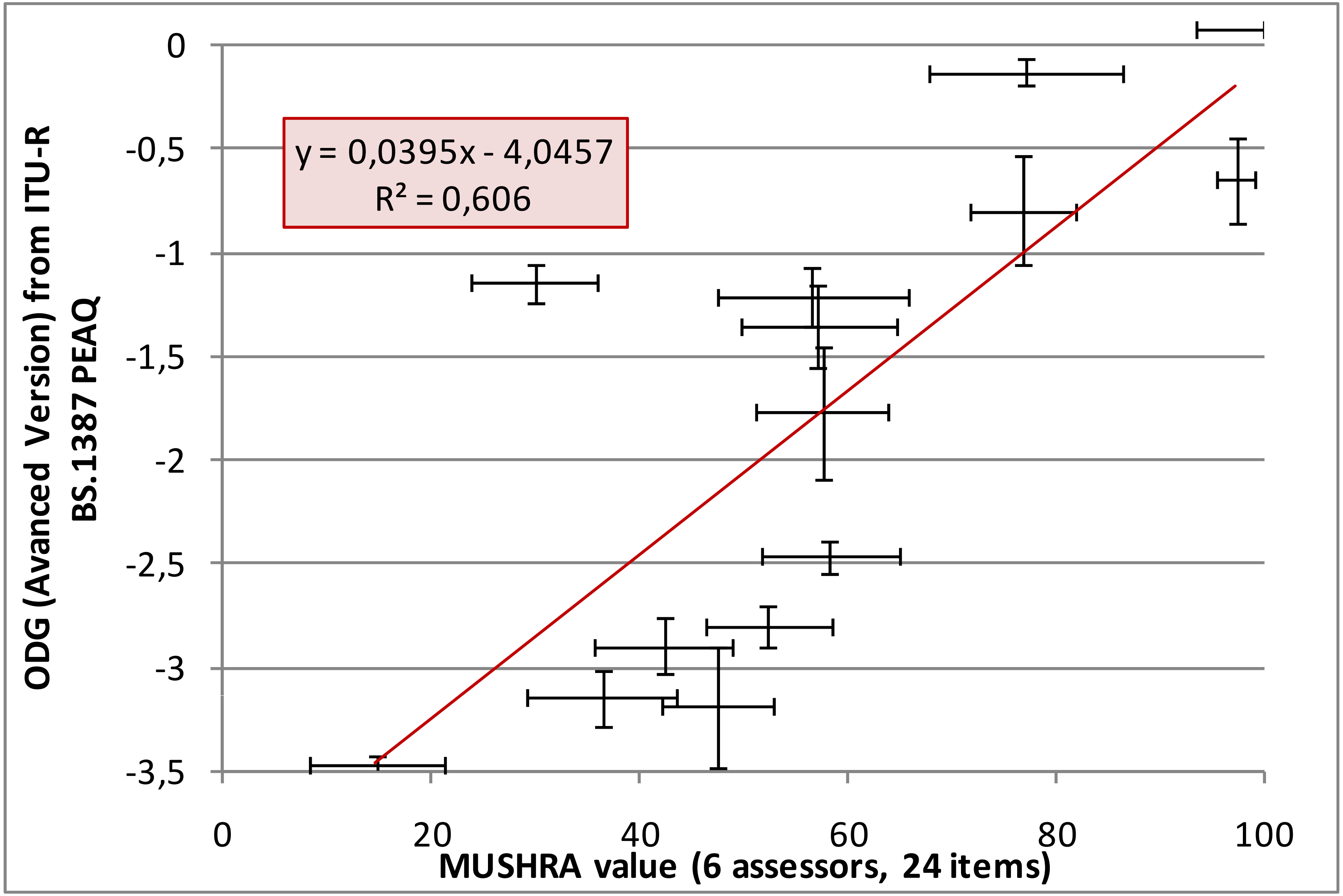
Codec Contest



Comparing MUSHRA with PEAQ (BV) based on Samples Coded with SBC



Comparing MUSHRA with PEAQ (AV)



Summary

- G.711 Appendix I extended to full bandwidth
  - Already included in open source soft-phone Ekiga
- Ultra low delay audio codecs
  - Proposed Bluetooth SBC for Internet usage
  - Done performance comparison
  - CELT v0.4 by Jean-Marc Valin won
- Comparing MUSHRA with PEAQ
  - The PEAQ basic version does have a better correlation after applying polymeric mapping function
- Start IETF and ITU standardization now?