



Innovative R&D by NTT

A Packet Loss Recovery Technique with Line Spectral Frequency Modification in 3GPP EVS Codec

Kimitaka TSUTSUMI¹, Kei Kikui¹, Jérémie Lecomte²

¹NTT DOCOMO Inc., ²Fraunhofer IIS

Agenda



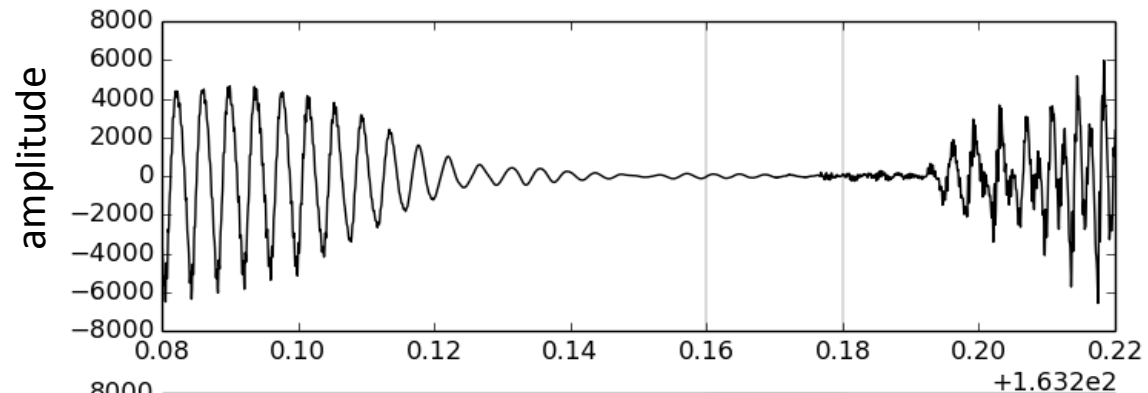
- Problem statement
- LSF quantization in EVS codec
- Analysis on speech overshoot
- Proposed method
- Experimental result
- Conclusion

Problem: Speech overshoot

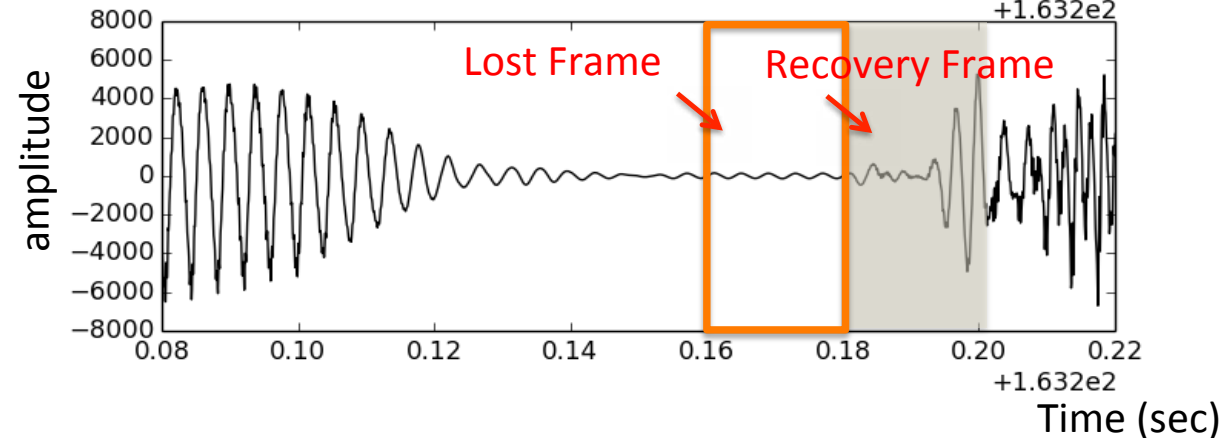


Observed power surge at the first frame recovered from packet loss at a frame before speech onset

Waveform under Clean channel

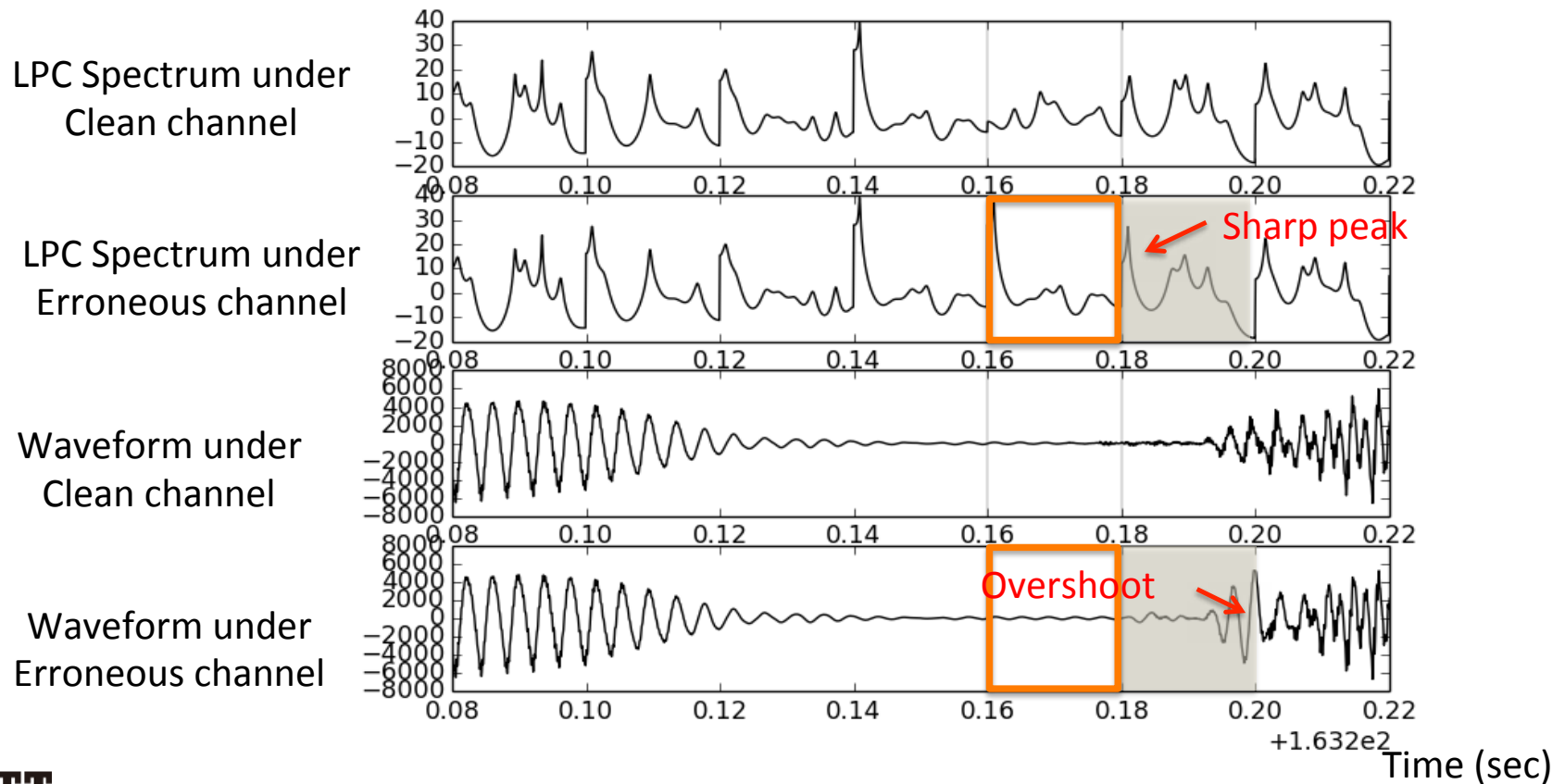


Waveform under Erroneous channel



Problem: LPC spectrum

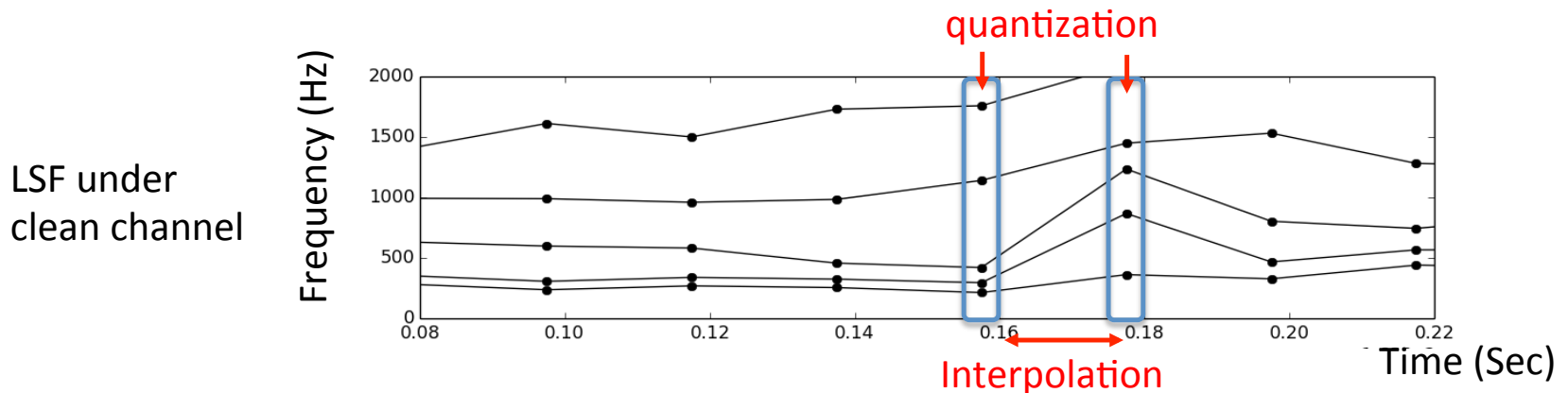
Sharp peak in LPC spectrum observed at the recovery frame suffering from speech overshoot



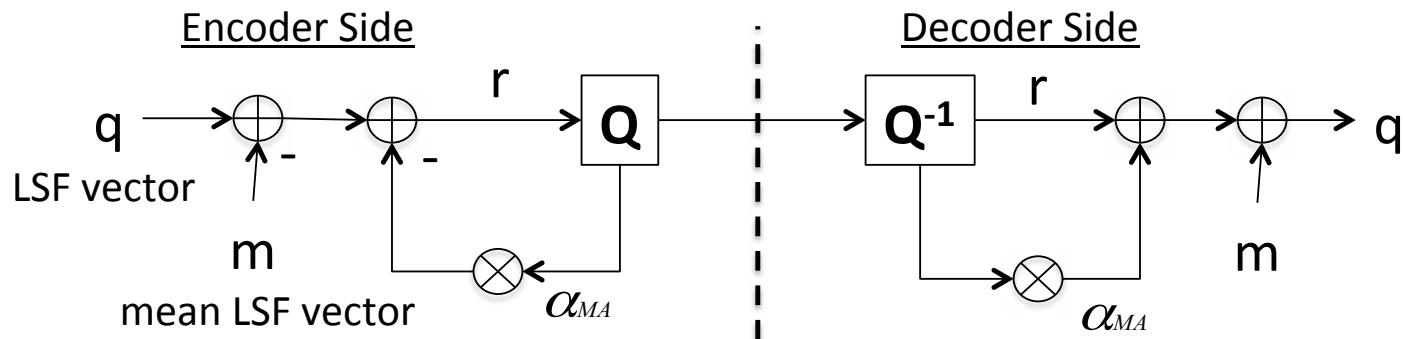
LSF quantization in EVS codec



- Quantization only at selected sub-frames
- Interpolation from available LSF vectors



- MA prediction for quantization at speech onset



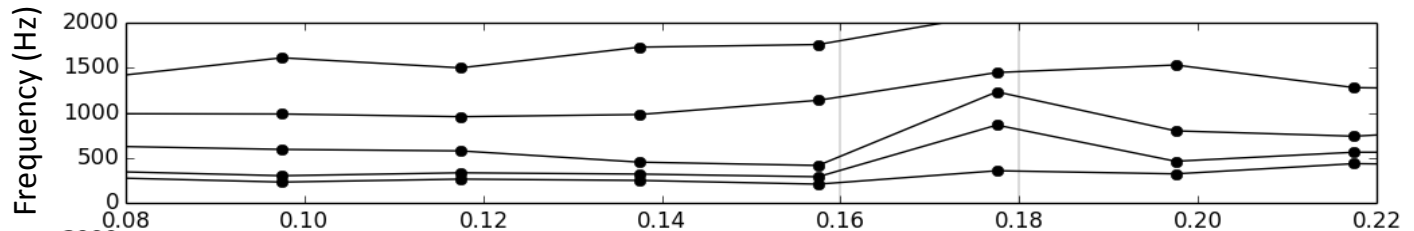
Recovery of LSF in EVS codec



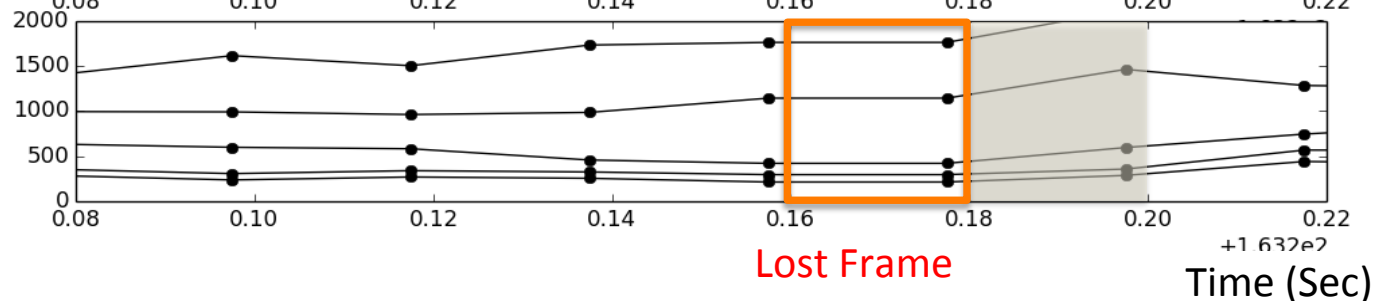
- Two frames required for complete recovery
- The first Recovery frame: affected by $r^{[-1]}$
- The Second recovery frame: affected by interpolation

$$z^{[0]} = r^{[0]} + \alpha_{MA} r^{[-1]}$$

LSF under
clean channel



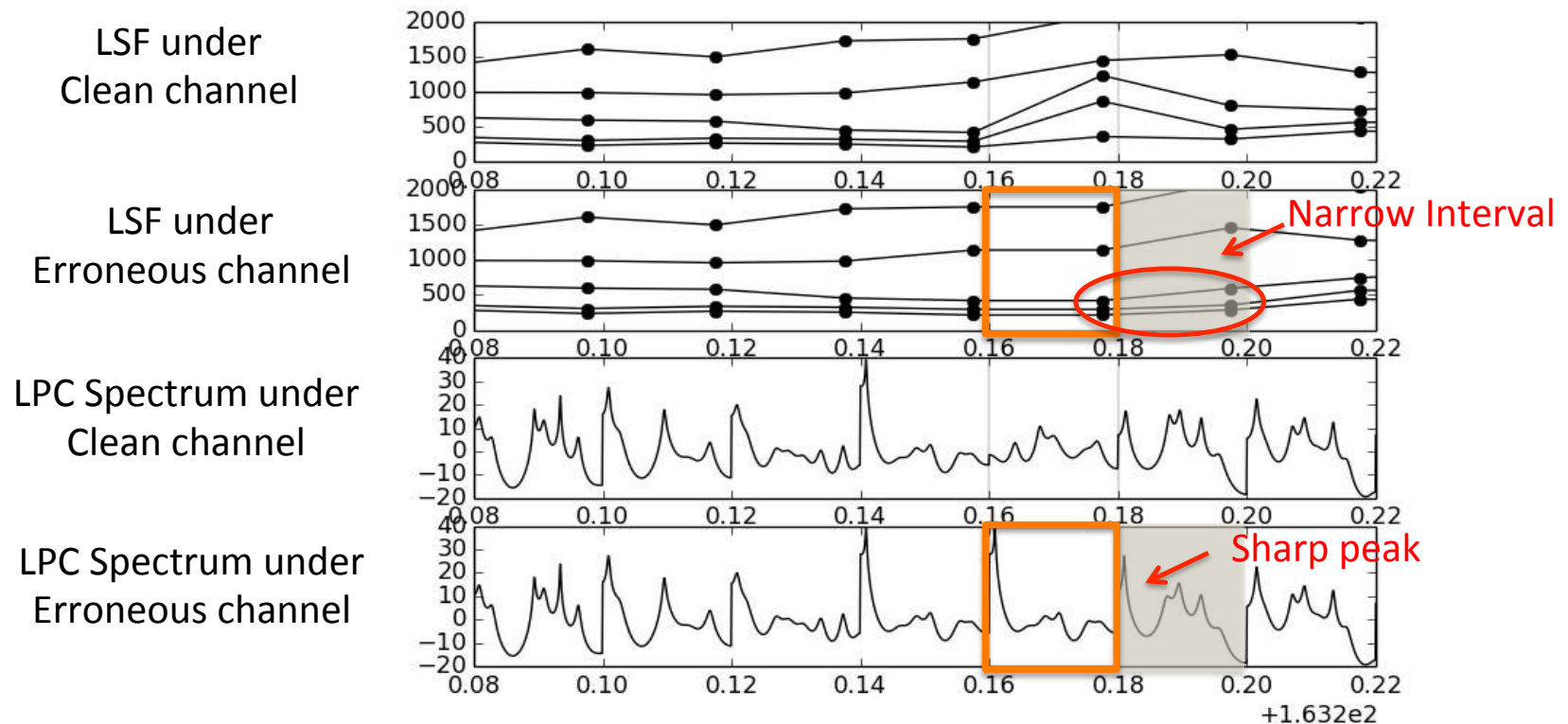
LSF under
erroneous channel



Analysis on speech overshoot



- Sharp peak in LPC spectrum caused by narrow interval between LSF parameters



Speech overshoot detection



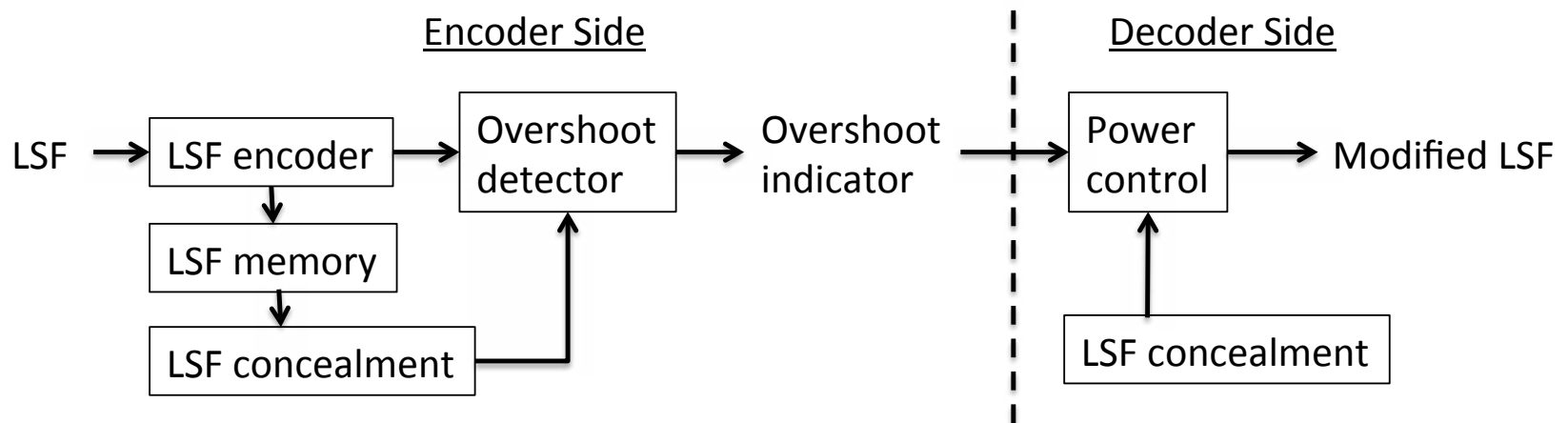
- Difficult to detect at the decoder side
 - ✓ Due to packet loss, or power surge of original?
- Frame class at overshoot not classified as “Onset”
 - ✓ Waveform at overshoot looks like speech onset

Frame class of overshoot frame	The number of frames
Unvoiced	183
Unvoiced transition	0
Voiced transition	1
Voiced	65
Onset	7
Total	256

Proposed method: overview



- Speech overshoot detection based on PLC simulation at encoder side
- Transmit the simulation result to the decoder

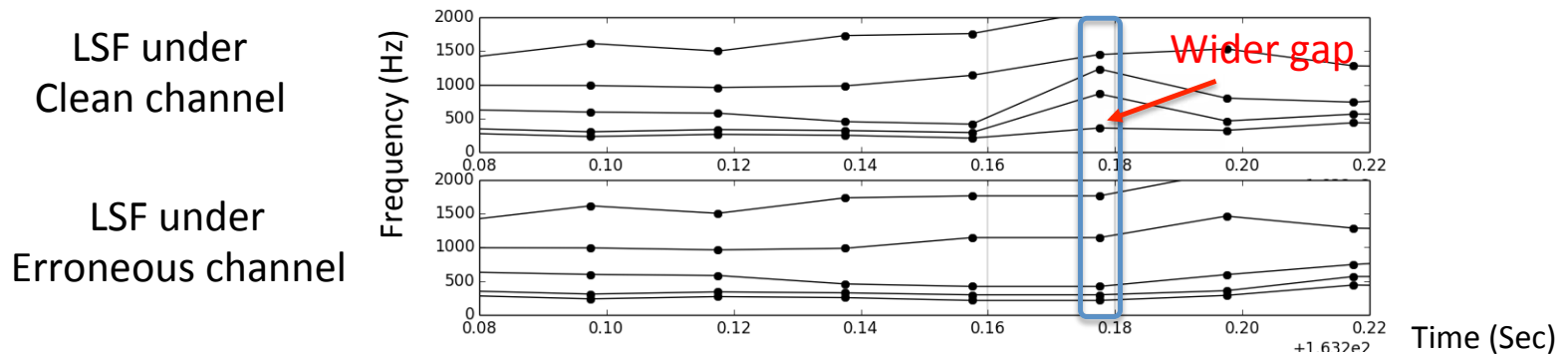


Block-diagram of proposed technology

Proposed method: power control



- Decoder side tool
- Narrow interval compared to clean channel



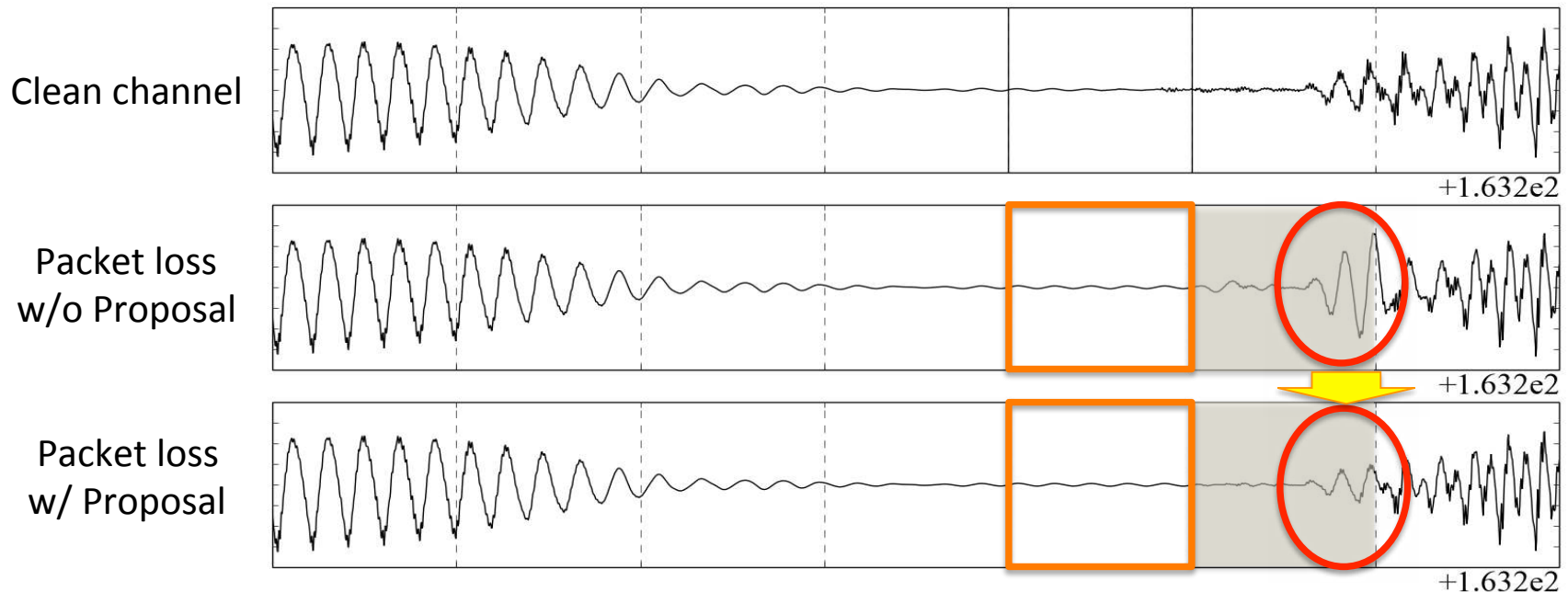
- Increase interval of the output vector of LSF concealment

$$\tilde{\omega}_j^{[-1]} = \begin{cases} j \cdot \delta & (1 \leq j < idx) \\ \dot{\omega}_j^{[-1]} & (idx \leq j < 16) \end{cases} \quad \delta = \omega_{idx-1}^{[-1]} / (idx - 1)$$

Proposed method: power control



- Sharp peak in LPC spectrum prevented
- Power surge at the recovery frame suppressed



Proposed method: overshoot detector



- Encoder side tool
- Test if modified LSF vector closer to LSF under clean channel than concealed LSF

$$\sum_{i=1}^{16} \overset{\text{concealed LSF}}{\left| \dot{\omega}_j^{[-1]} - \omega_j^{[-1]} \right|^2} > \alpha \cdot \sum_{i=1}^{16} \overset{\text{Modified LSF based on concealed LSF}}{\left| \tilde{\omega}_j^{[-1]} - \omega_j^{[-1]} \right|^2} \quad \overset{\text{LSF under clean channel}}{\omega_j^{[-1]}}$$

- Test if each of the elements in modified LSF vector closer to LSF under clean channel than concealed LSF

$$\left| \dot{\omega}_j^{[-1]} - \omega_j^{[-1]} \right|^2 > \left| \tilde{\omega}_j^{[-1]} - \omega_j^{[-1]} \right|^2$$

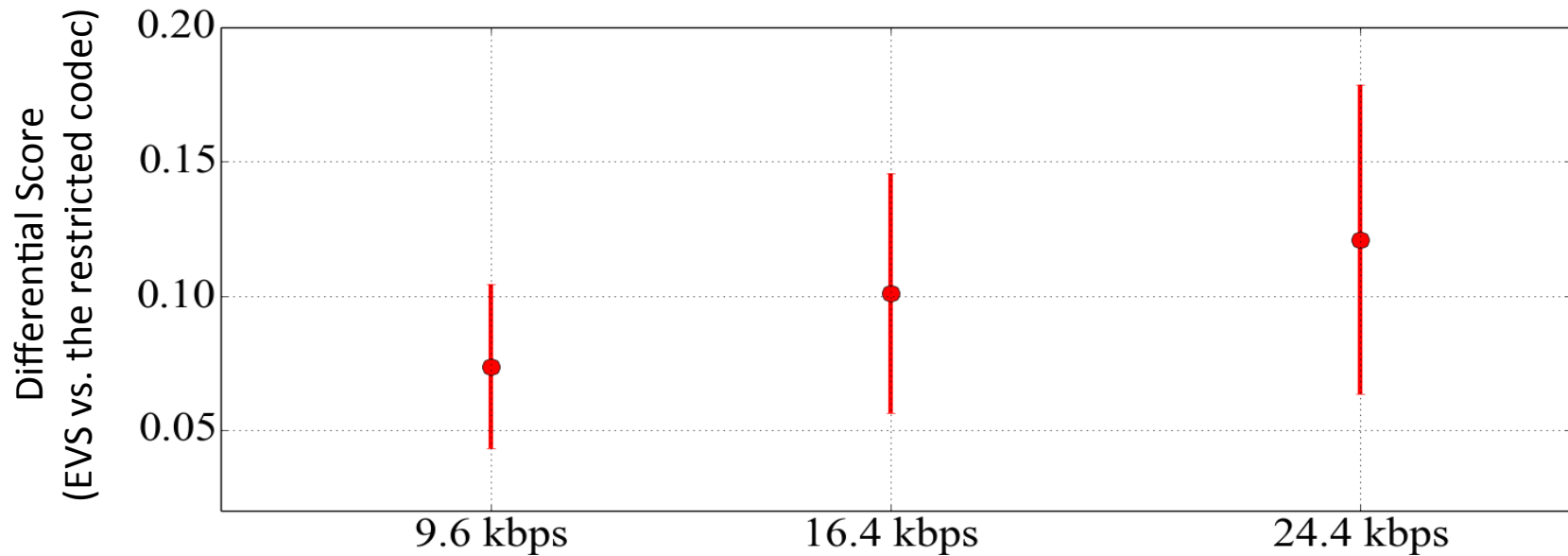
- Limit activation only at onset

$$\text{Minimum value of Algebraic CB gain} \longrightarrow g_{\min} > g_{th} \longleftarrow \text{Predetermined threshold}$$

Experimental result: objective quality



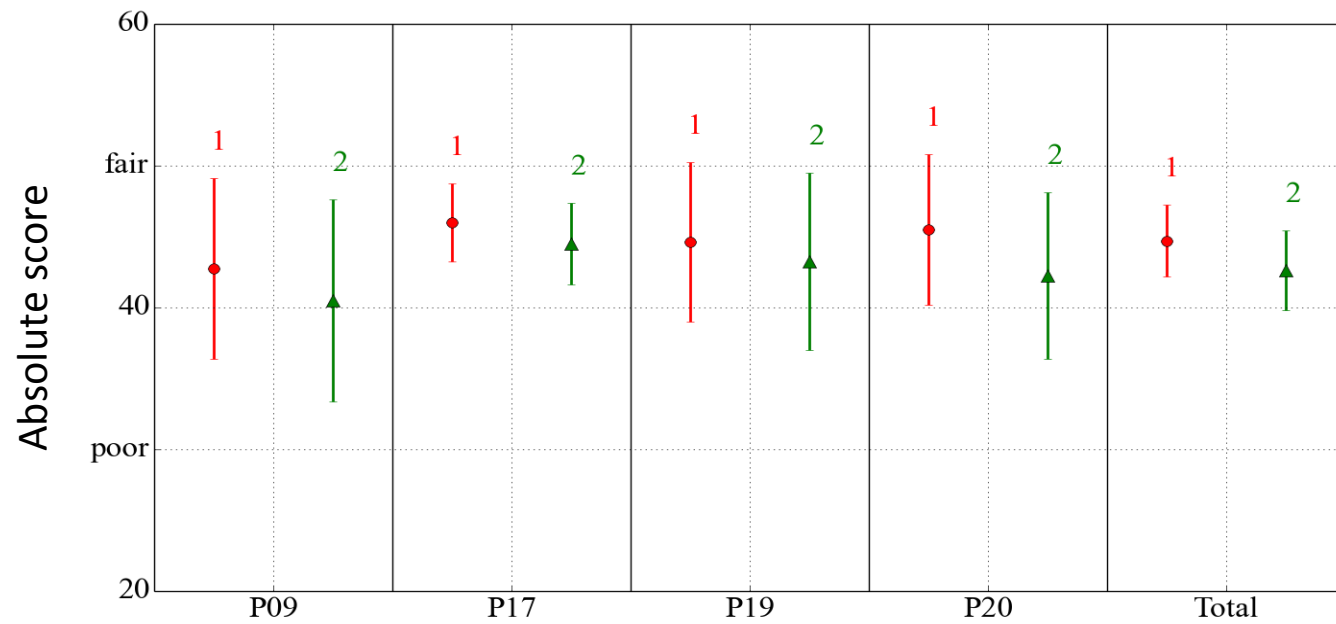
- Wideband PESQ for artificially created error pattern
- Speech materials (wideband) of 170 seconds in total
- Observed significant improvement for all the operating bit-rate
- Comparison with the restricted codec (EVS codec with the proposed method deactivated)



Experimental result: subjective quality



- Mushra test for 8 expert listeners
- 4 speech items with 6 % random error pattern
- Comparison between EVS codec and the restricted EVS codec
- Observed improvement in average score for all the items

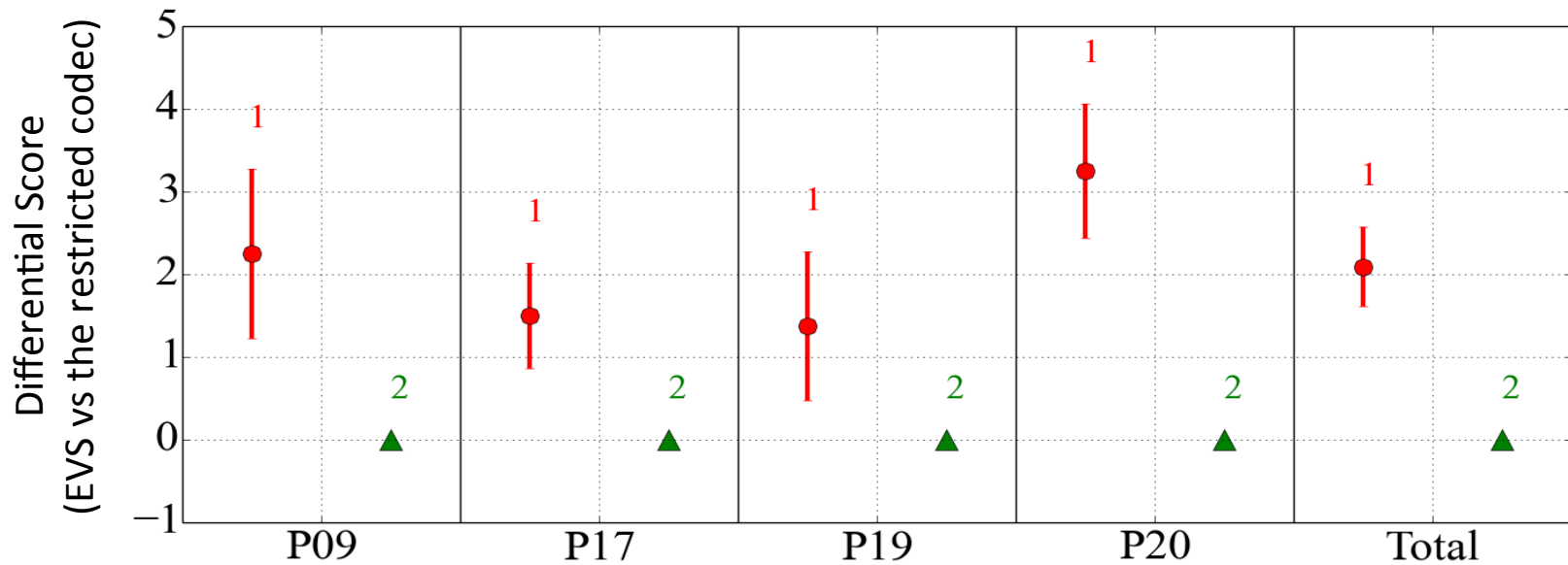


1. EVS, 2. the restricted codec (EVS codec with the proposed method deactivated)

Experimental result: additional analysis



Observed significant improvement with proposed method



1. EVS, 2. the restricted codec (EVS codec with the proposed method deactivated)

Conclusion



- Examined a behavior of speech overshoot at the recovery frame from packet loss around speech onset
- Power control method, and speech overshoot detection method developed
- Observed quality improvement by experimental results