Optimal Codec for Wireless Communication Links

W. Bastiaan Kleijn and Minyue Li
School of Electrical Engineering
KTH - Royal Institute of Technology
Stockholm, Sweden

HEARCOM Context

services:

recognition server

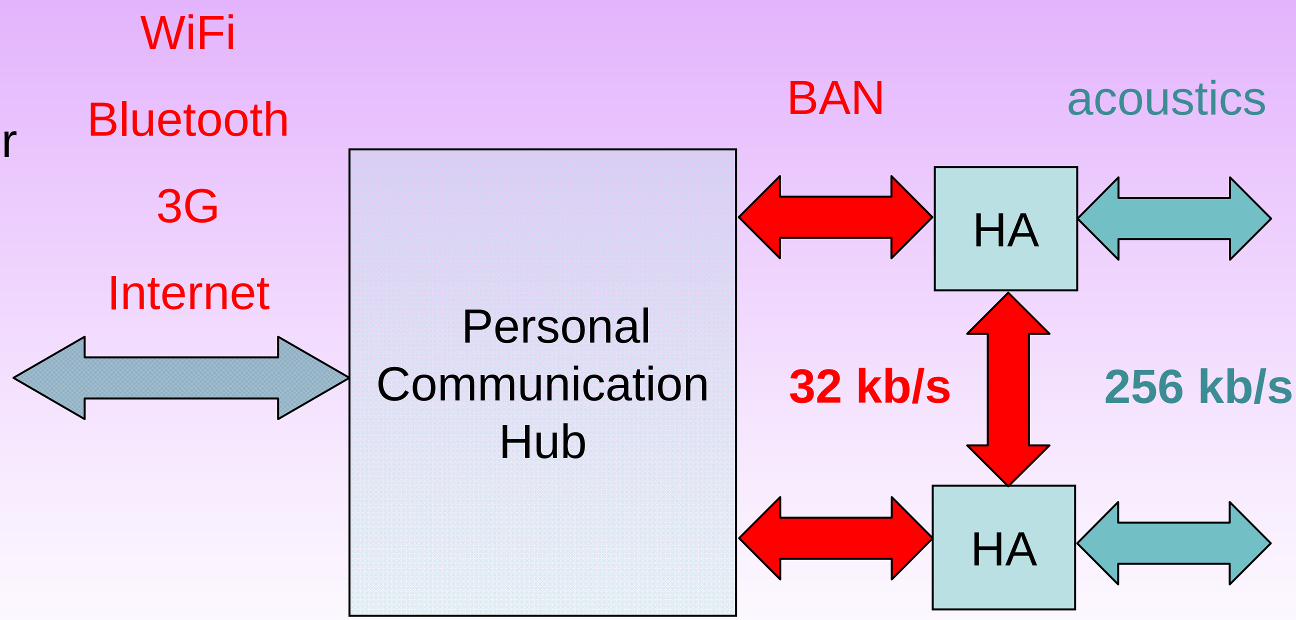
announcements

telephone

VoIP

radio/television

theater/church audio

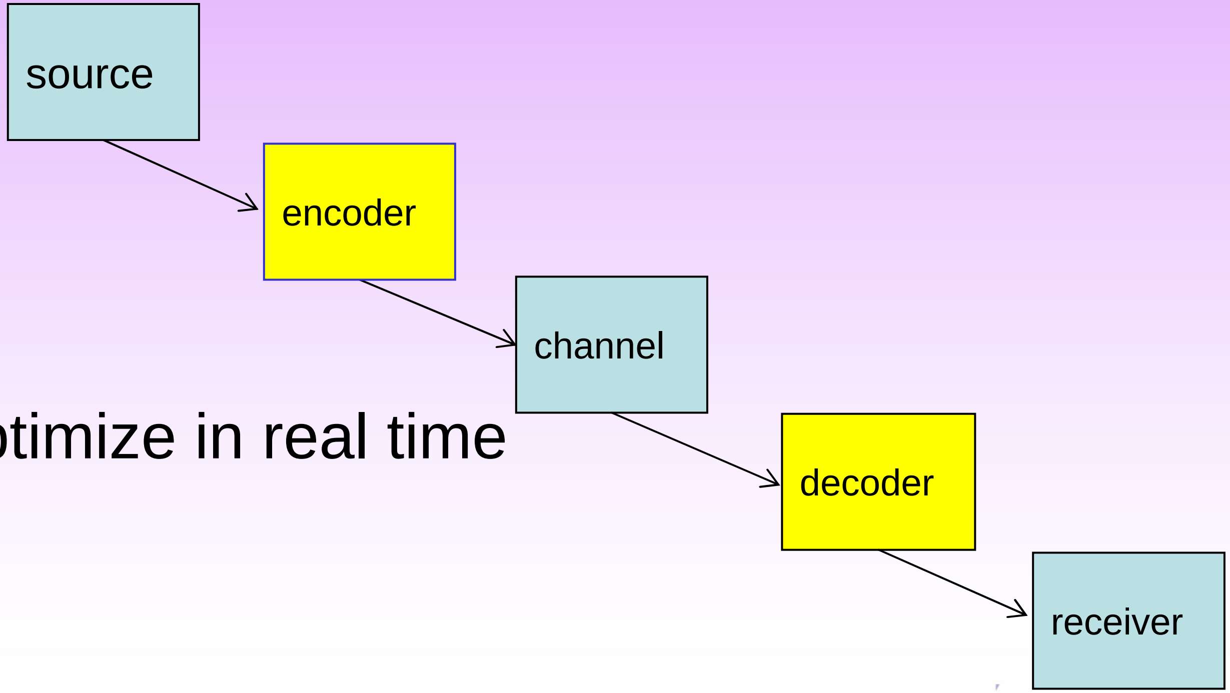


Objectives

- Audio coder for hearing instruments
  - Real-time scalable
  - Delay < 1 ms; facilitates inter-ear communication
  - Sampling-rate 16 kHz
  - “Design” Rate 32 kb/s
  - Low complexity (processor dependent)
  - Quality better than ITU-T G.722 at 64 kb/s
    - Delay 1.625 ms.

Design Philosophy: *Model Everything*

- Statistical models of
  - Source
  - Channel
  - Receiver
  - Encoder
  - Decoder
- Estimate / optimize in real time



Real-Time Design Method Based on High-Rate Theory

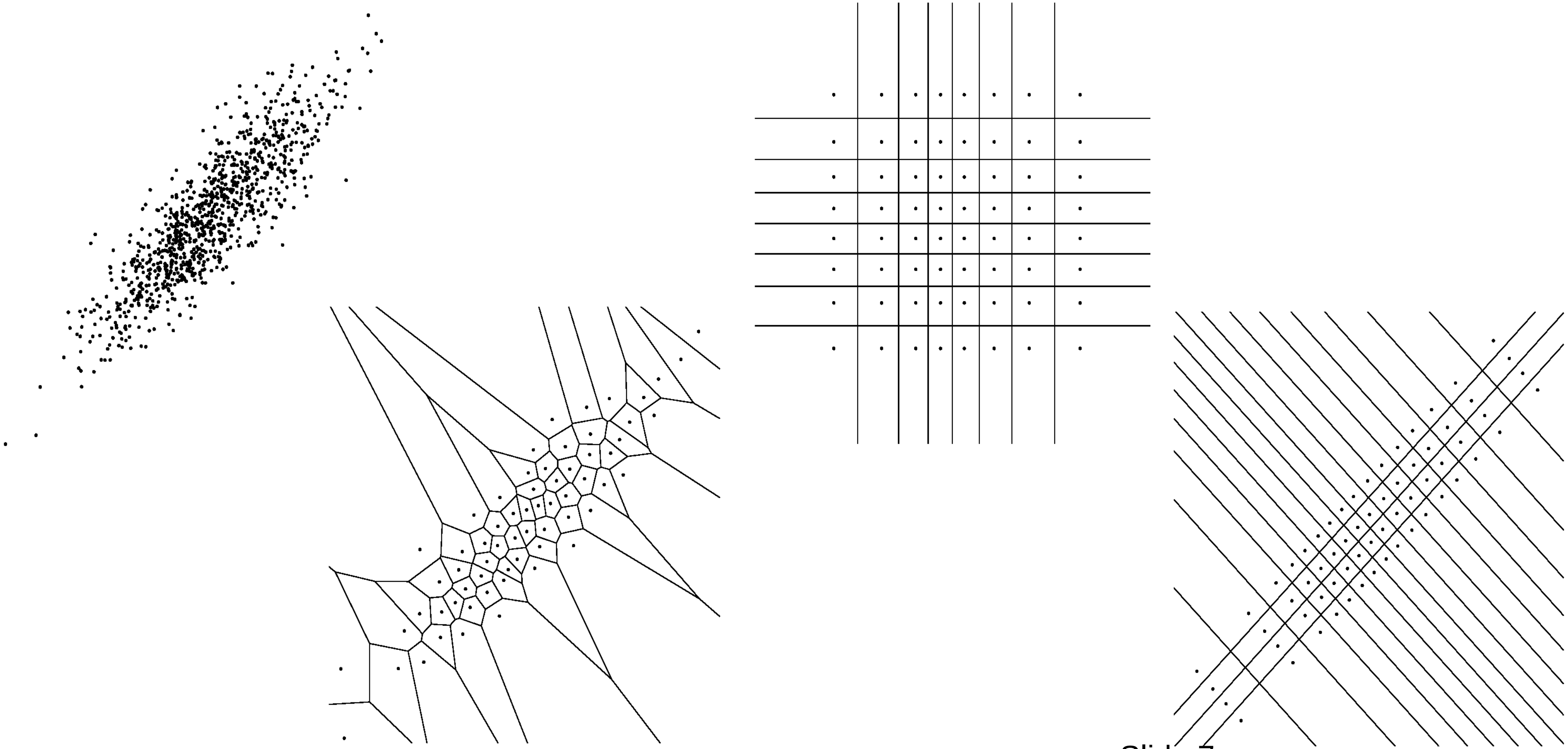
- Rate-distortion theory (Shannon, 1959)
  - Needs densities; bounds for simple densities only
  - Variable-rate only
  - No direct relation to practical systems
- Lloyd algorithm (Lloyd, 1958)
  - Not a model; leads directly to quantizer
  - Iterative / results in codebooks / not scalable
  - Locally optimal / no need density function
- High-rate theory (Bennett, 1948)
  - Assumes signal density constant in quantization cell
  - Asymptotically optimal
  - Fixed and variable rate
  - Analytic solutions / scalable
  - Provides centroid density / requires additional step

Coder Design Choices

- Signal modeling
  - Predictive coding
  - Transform coding
- Adaptation of the model
  - Backward adaptation
  - Forward adaptation
- Rate or distortion variation
  - Constrained-entropy coding
  - Constrained-resolution coding

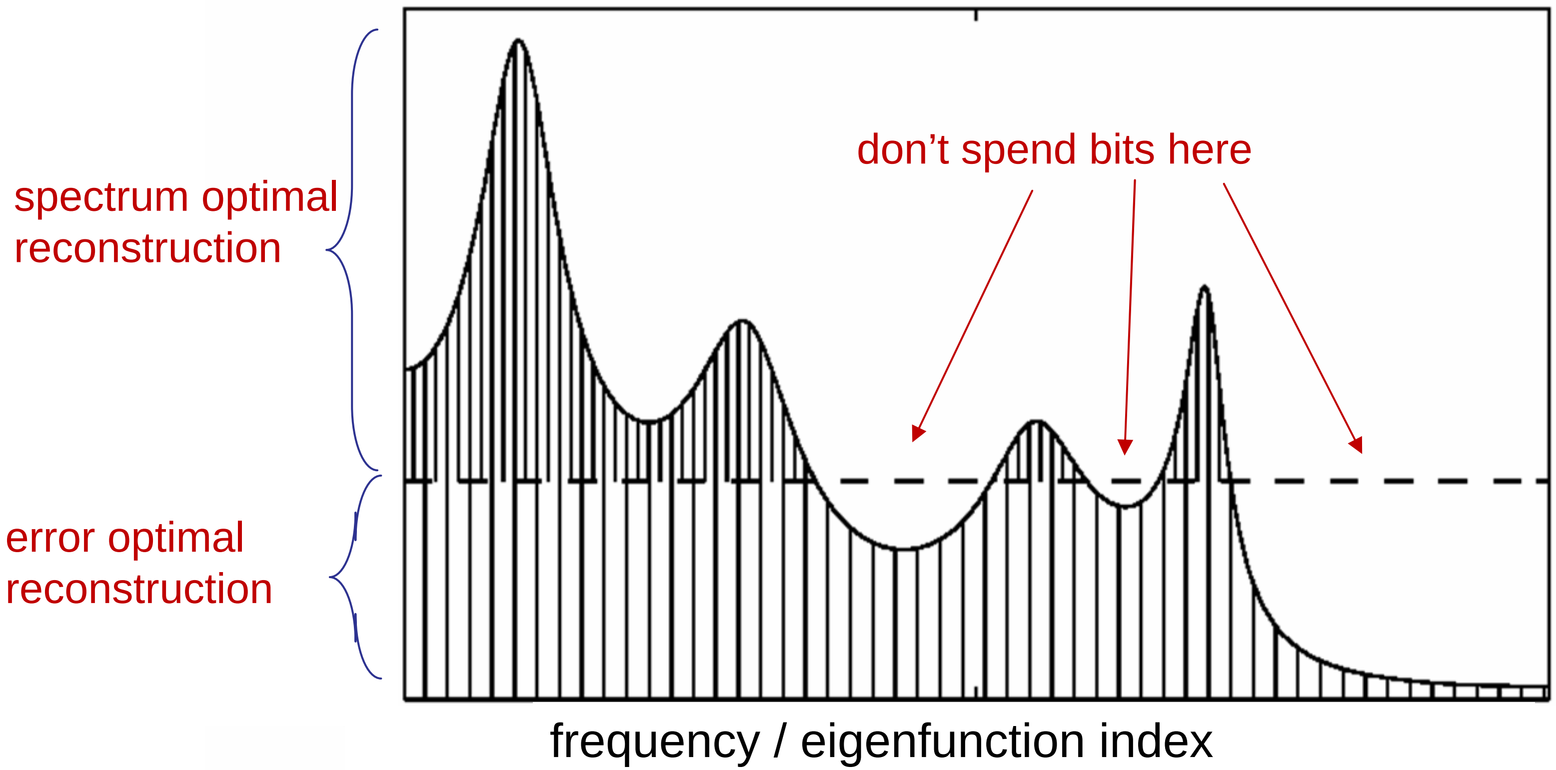
Choices: Prediction vs Transform

- 3-bit/dimension constrained-resolution quantizer



Reverse Waterfilling

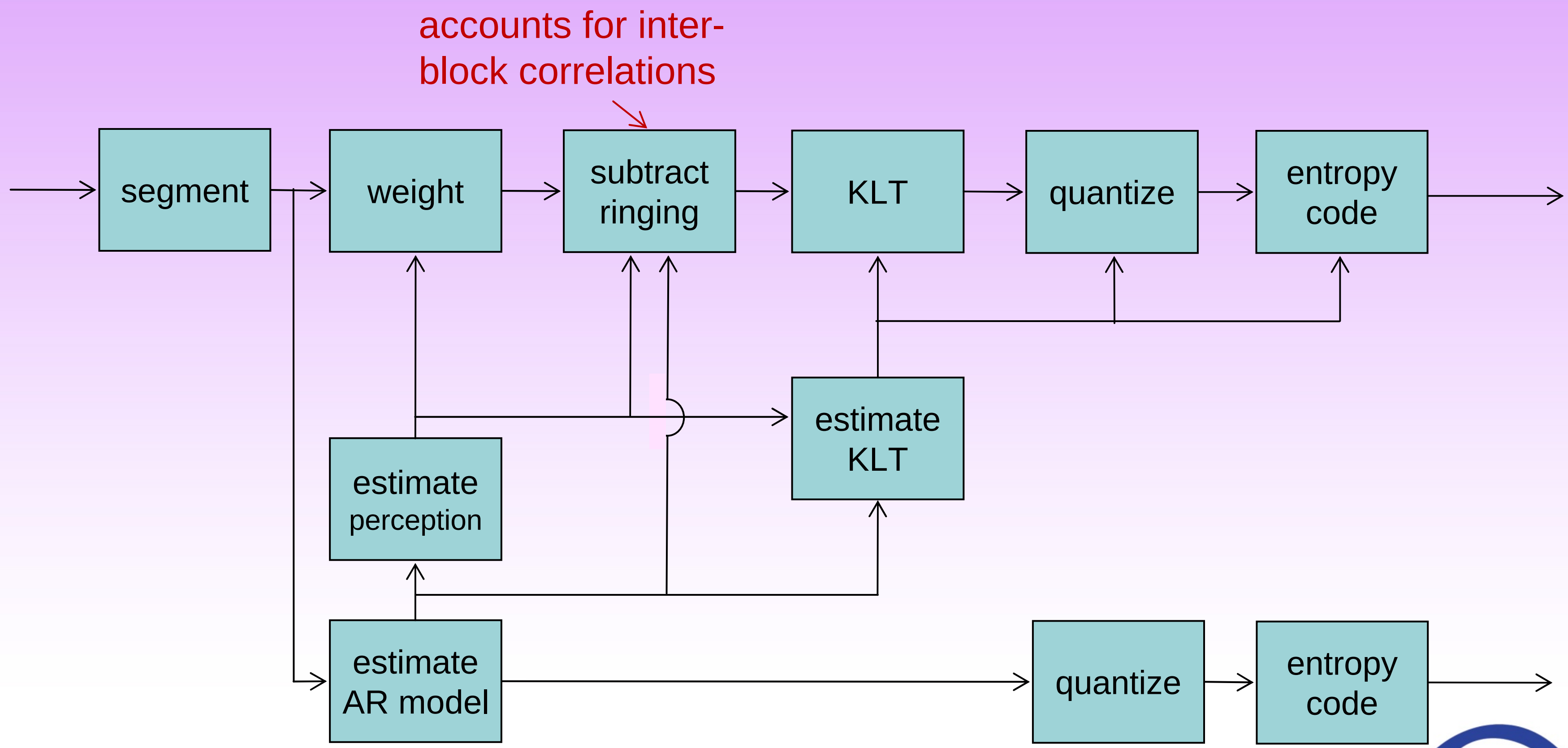
- Code only where needed



Choices: Modeling

- DCT-based coding: unitary transform
  - Does not affect space-filling
  - Reverse water filling
  - Imperfect decorrelation for fixed transform
  - *Long delay*
- CELP (analysis-by-synthesis AR coder)
  - AR model functions well
  - Inefficient space filling
  - No inherent reverse water filling (requires *post-filter*)
  - Nightmare for adaptive coding

Modeling: Forward Adaptive



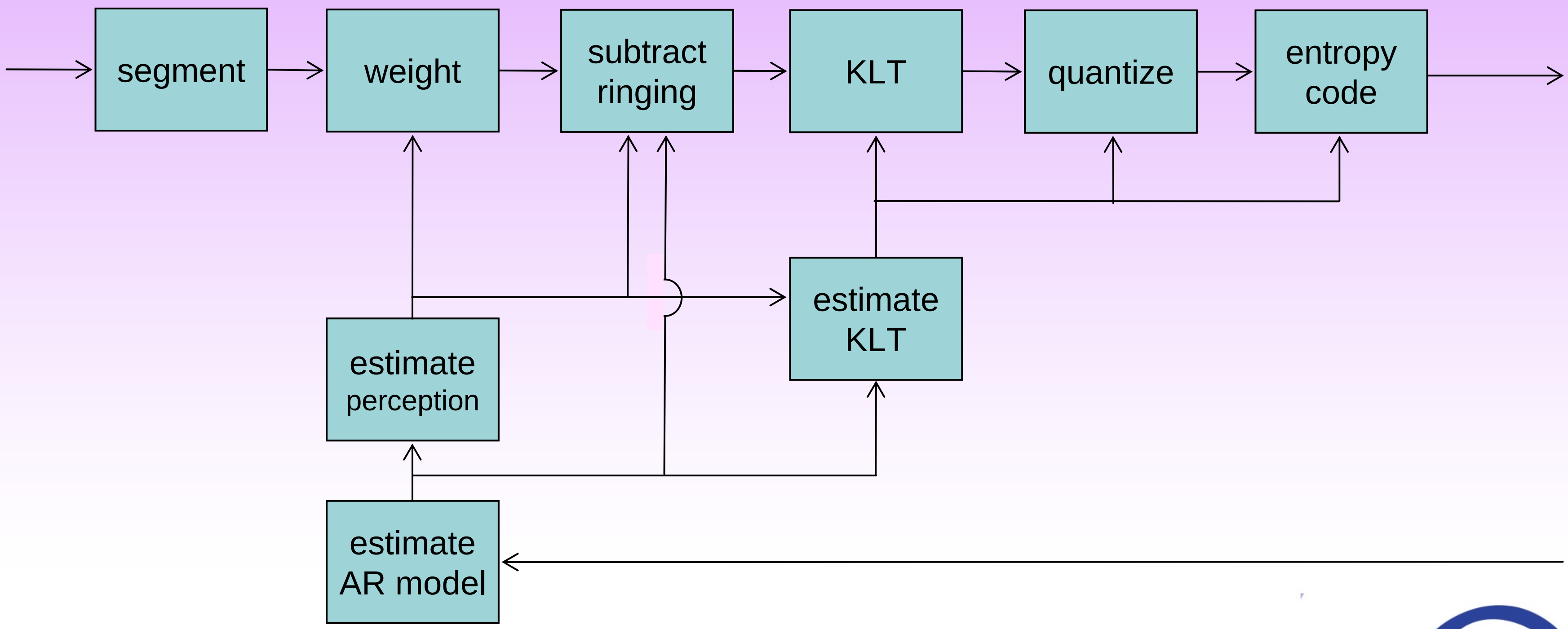
Choices: Adaptation

- Forward adaptation
  - Estimate model parameters on original signal
  - Transmit model parameters
  - Requires long blocks

long delay (or estimated on past signal)

- Backward adaptation
  - Estimate model parameters on reconstructed signal
  - No transmission required
  - Frequent update natural -> short delay
  - Estimate not accurate after transition

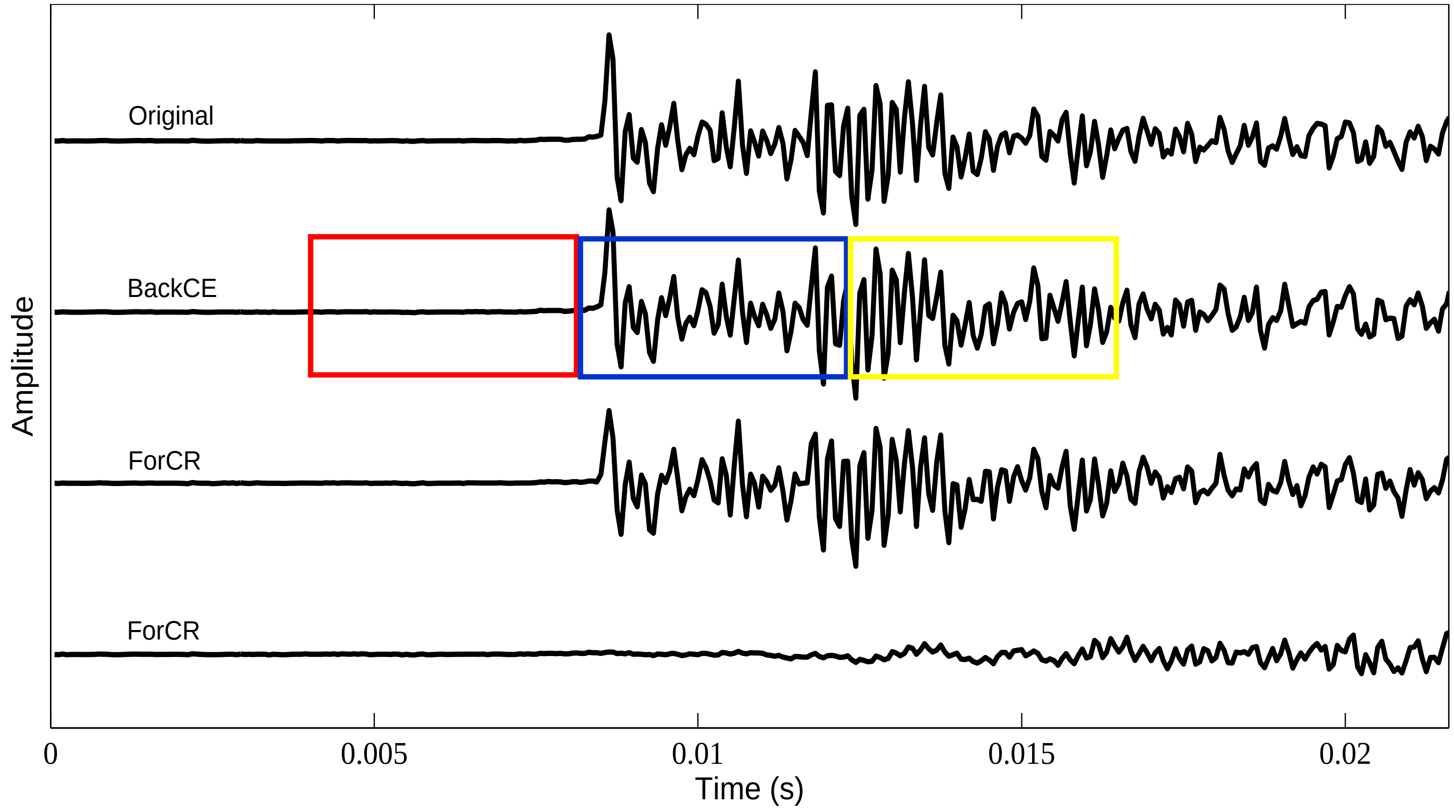
HEARCOM Coder Architecture



Constrained-Entropy Coding

- Constrained-resolution coding
  - Fixed rate
  - Outliers in distortion → leads to error feedback in backward adaptation
- Constrained-entropy coding
  - "Fixed" distortion → No error feedback in backward adaptation!
  - Outliers in rate

Illustration: Backward Adapt and CE

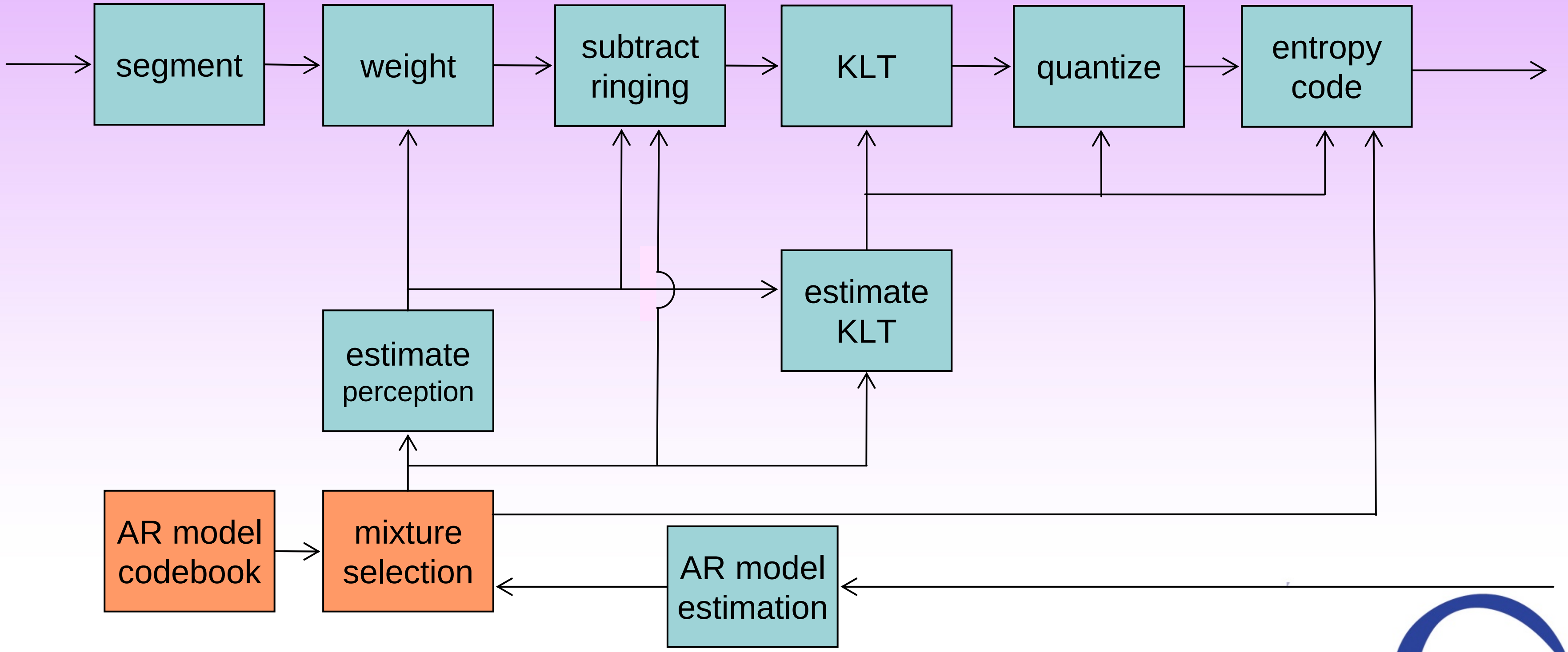


- Constrained entropy = fixed distortion
- Same average rate

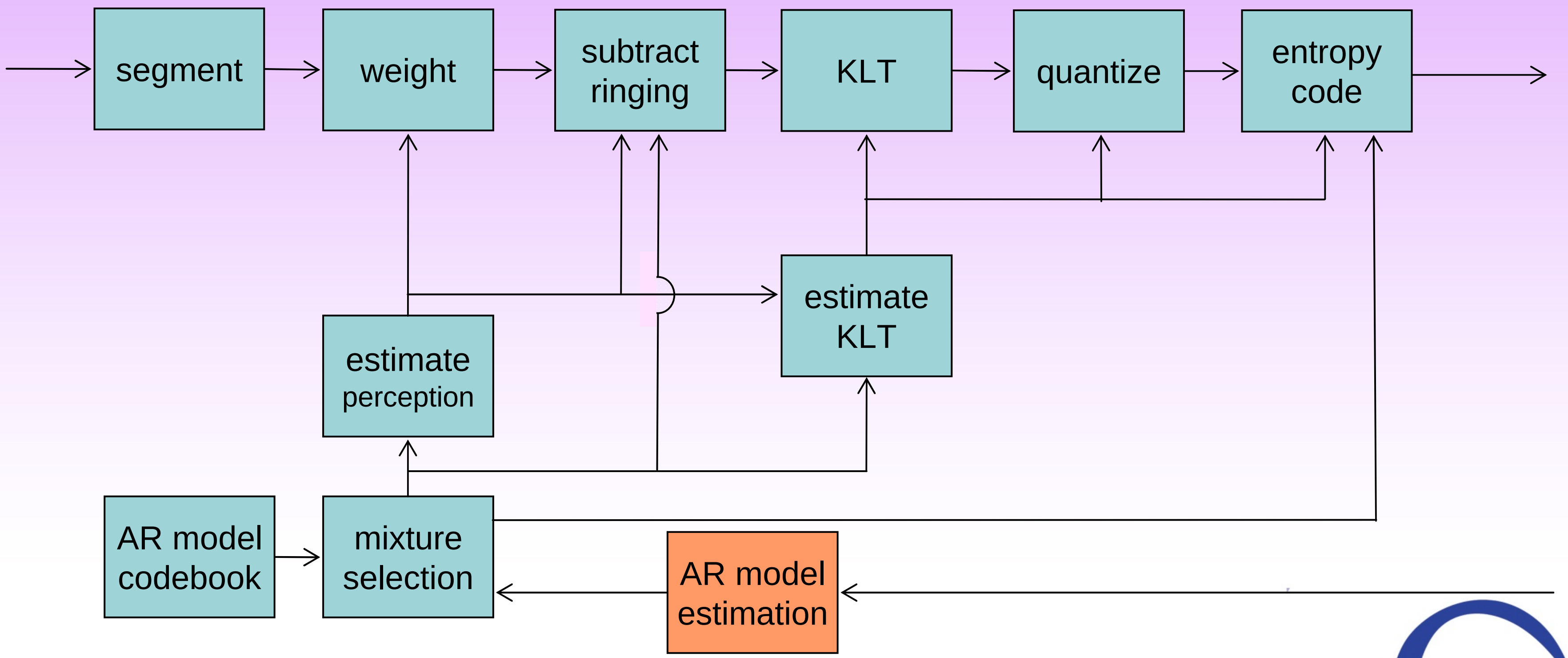
Constrained-Entropy Coding

- Constrained resolution coding
  - Fixed rate
  - Outliers in distortion → leads to error feedback in backward adaptation
- Constrained entropy coding
  - "Fixed" distortion → No error feedback in backward adaptation!
  - Outliers in rate → generally very bad rate variation

Removal Rate Outliers



Improved Backward Model Estimation



HEARCOM Coder Architecture

- KLT modeling:
  - Better than CELP
  - Better than MDCT
- Backward-adaptive, variable rate
  - Very short delay
  - Constant quality
  - No feedback of distortion
  - Limited rate variation
  - Enhanced model estimation

Comparative Quality and Rate

- MUSHRA (MUltiple Stimuli with Hidden Reference and Anchor) test
- HEARCOM coder is *scalable*!

coder	Rate (kb/s)	Delay (ms)	MUSHRA score
ITU-T G.722	64	1.625	69.6
ITU-T-G.722	48	1.625	56.9
HEARCOM	32*	0.625	71.3

Conclusions

- Architecture determines effectiveness low-D coding
  - CELP not naturally scalable
  - MDCT has long delay, imperfect decorrelation
  - KLT-based architecture performs best
- HEARCOM coder attributes:
  - Scalable: can redesign coder in real time at any time
  - Given quality: reduces rate of G.722 by factor two
  - Very low delay (0.625 ms)
  - Low computational complexity
  - Architecture can also be used for forward-adaptive approach
    - Facilitates multiple-description coding
  - Coder is programmed in C