Flexible Multiple Description Coding of Audio

Janusz Klejsa

FlexCode Public Seminar
Tampere, June 16, 2008

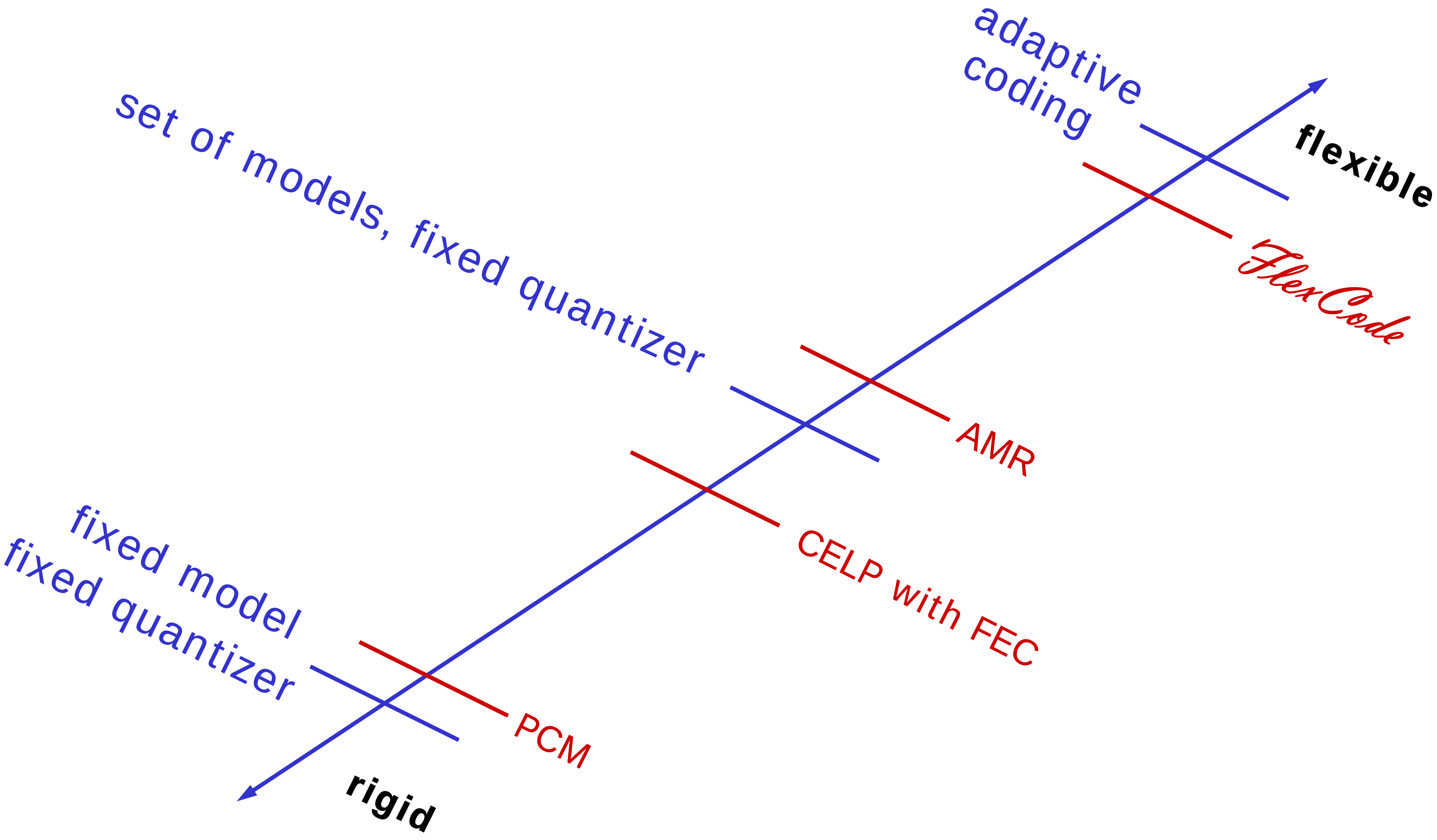
Outline

- Flexible Audio Coder
- Multiple Description Coding (MDC) in a Nutshell
- Application to Audio Coding
- Conclusions

Problem & Background

- Networks:
  - Heterogeneity increasing
  - Inherent variability (mobile users)
  - Layered structure well established (>20 years of OSI)
- Coders:
  - Designed for a specific environment
  - Inflexible schemes involved (trained codebooks, FEC…)
  - Feedback information underutilized

Adaptation and Coding



FlexCode

Flexible Audio Coder: Tools

- Tools
  - models of source, channel, receiver
  - high-rate quantization theory
  - multiple-description coding (MDC)
  - iterative source-channel decoding
  - distortion measures using the sensitivity matrix

Outline

- Flexible Audio Coder
- Multiple-Description Coding (MDC) in a Nutshell
  – Problem Statement
  – Notion of Multiple Descriptions
  – Illustrative Example
- Application to Audio Coding
- Conclusions

Problem Statement

- Goal:
  - robust transmission of audio stream over network with packet erasures
- Problem:
  - combating packet losses
- Solution:
  - robustness via redundancy
- Design trade-off:
  - bit-rate vs. quality

Forward Error Correction

- FEC can provide optimal performance, when
  - entropy rate of the source < channel capacity
  - no constraints on delay and complexity
  - system is time-invariant

  Context: separate source and channel coding
- Typical case:
  - real-time constraint: finite delay, reasonable complexity
  - feedback: packet-loss rate estimate available
- An alternative to FEC is needed!

  Context: joint source-channel coding

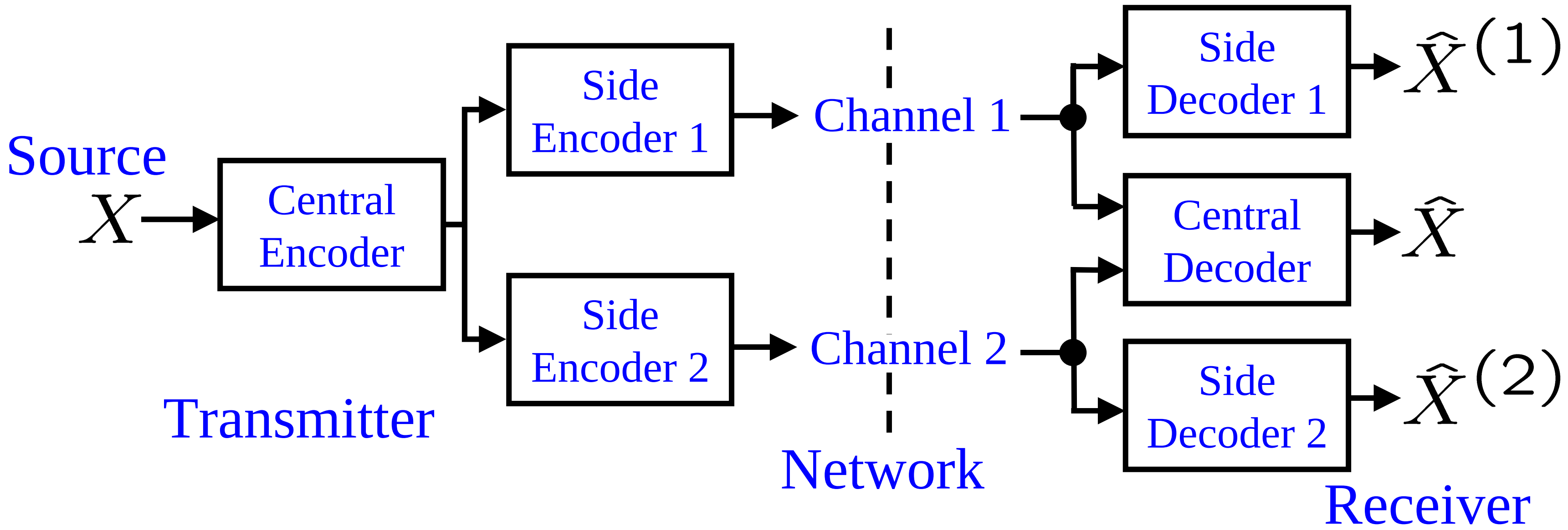
Problem Revisited

- Goal:
  - combating packet losses
- Constraints & Requirements:
  - finite (low) delay required
  - reasonable complexity
  - scalable in rate
  - scalable in redundancy

  } FlexCode

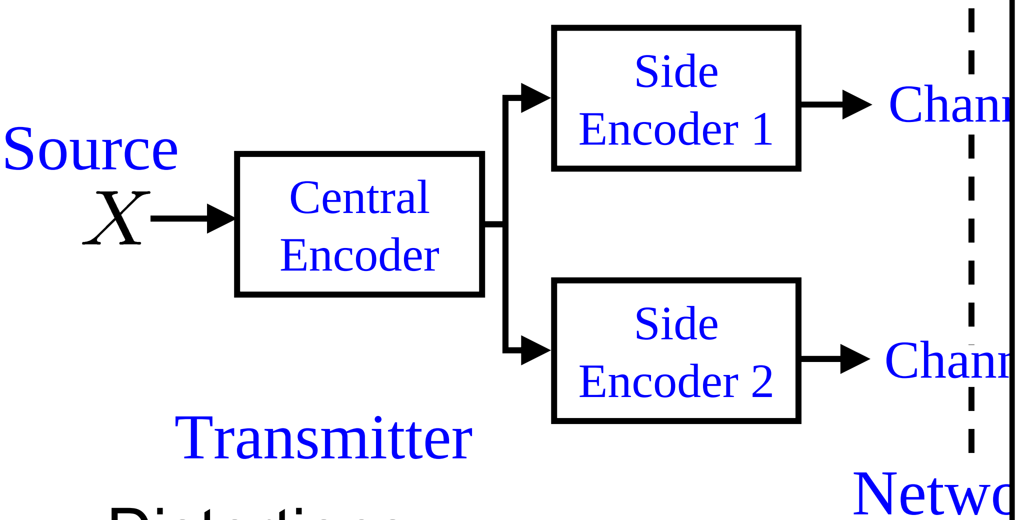
- Means to achieve the goal:
  - diversity of the network
  - source coding techniques
  - perception: *graceful decay of quality can be accepted*

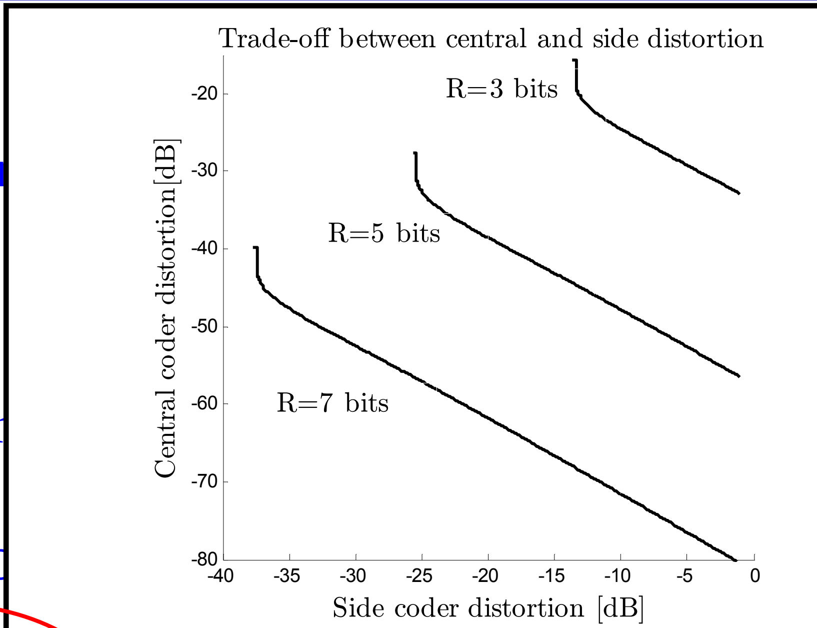
Notion of Multiple Descriptions (1)



- Create multiple descriptions of a single source
  - each description can reconstruct the source
  - the quality *gradually* improves with increasing number of received descriptions.
- Exploit network diversity
  - use setup that guarantees *independent losses of descriptions*

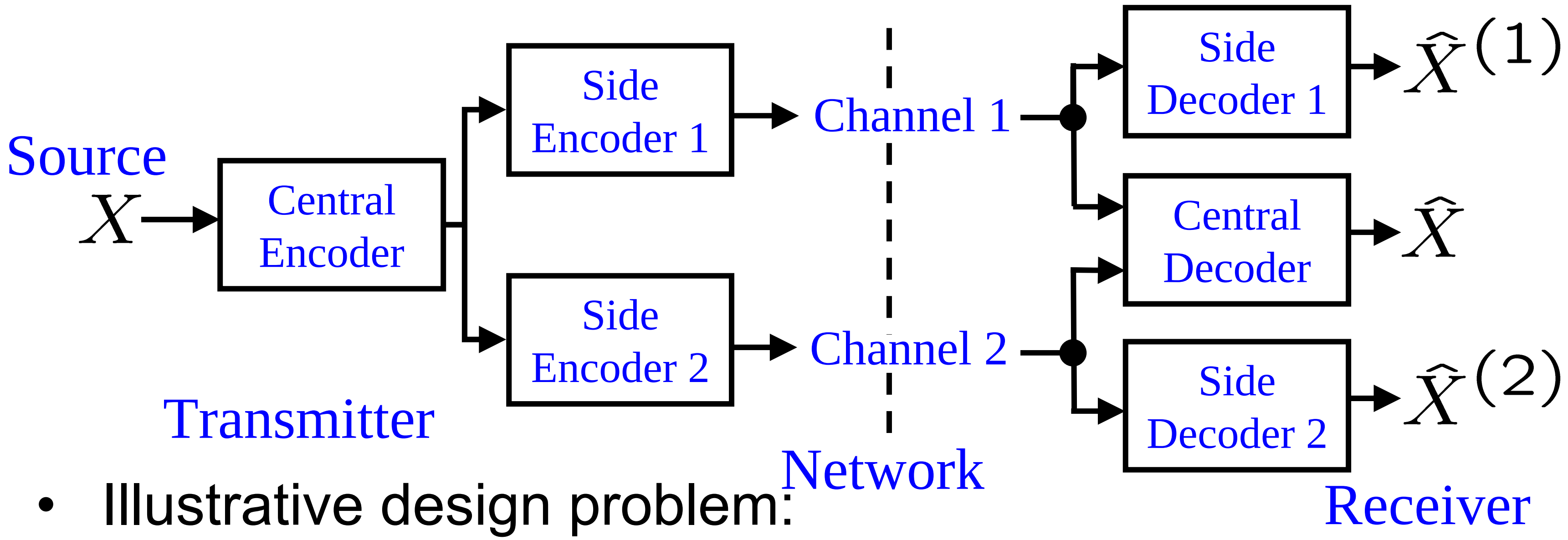
Notion of Multiple Descriptions (2)





- Distortions:
  - central distortion (MSE) $D_C(R_1, R_2) = \mathbf{E}\{(X - \hat{X})^2\}$ trade-off
  - side distortions (MSE) $D_S^{(j)}(R_j) = \mathbf{E}\{(X - \hat{X}^{(j)})^2\}, \; j = 1, 2$
- Constraints and channel properties:
  - fixed rate (constrained resolution) or fixed average rate (constrained entropy) $R_T = R_1 + R_2$
  - probabilities of description erasure p_1 and p_2

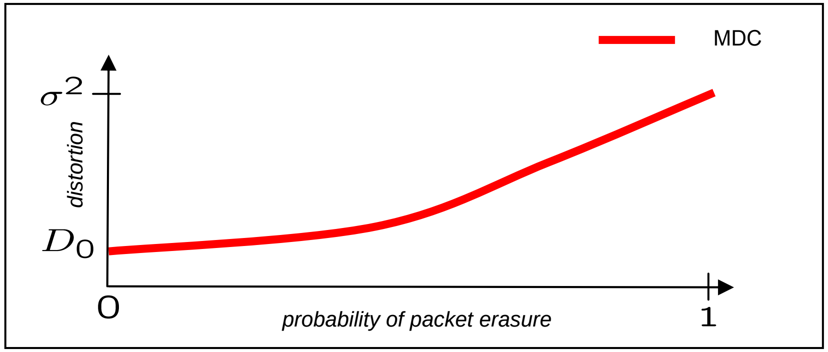
Notion of Multiple Descriptions (3)

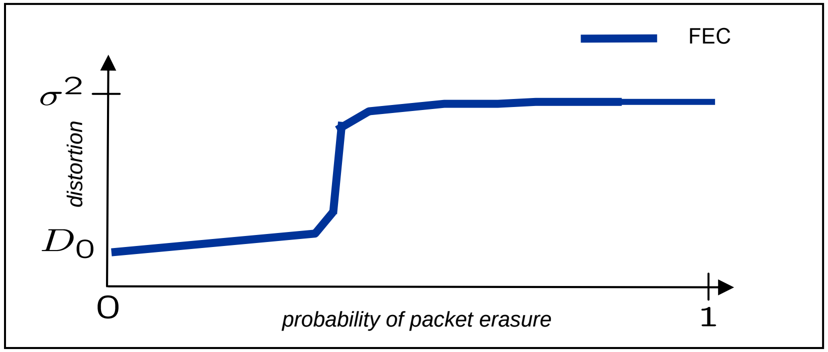


- Illustrative design problem:
  - Assume symmetrical case (most relevant), i.e.
    $p_1 = p_2 = p$, $R_1 = R_2 = R \Rightarrow D_S^{(1)}(R_1) = D_S^{(2)}(R_2) = D_S(R)$
  - Define total distortion (for simplicity assuming a scalar case)
    $$D_T = (1-p)^2 D_C(R) + 2p(1-p)D_S(R) + p^2\sigma^2$$
  - Formulate optimization problem
    $\min D_T$ subject to $R = R^*$ given p

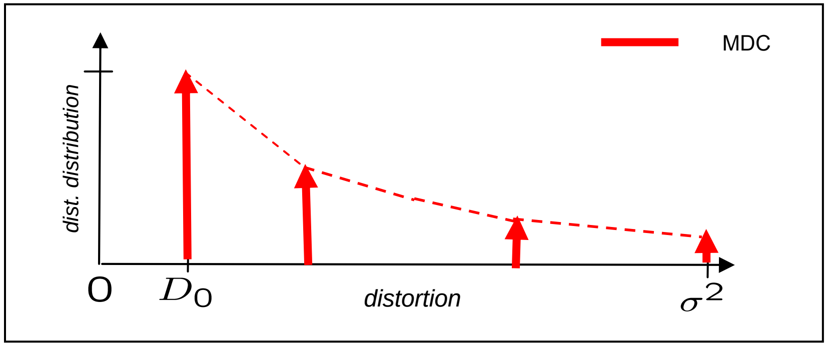
Comparison: MDC vs. FEC

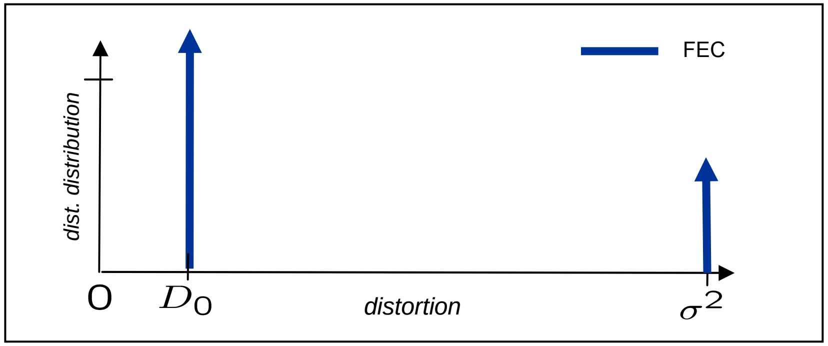
- Performance & properties
  - graceful (MDC) vs. fall-off-the-cliff (FEC) quality degradation





  - multi-level distortion distribution vs. two-level distortion distribution

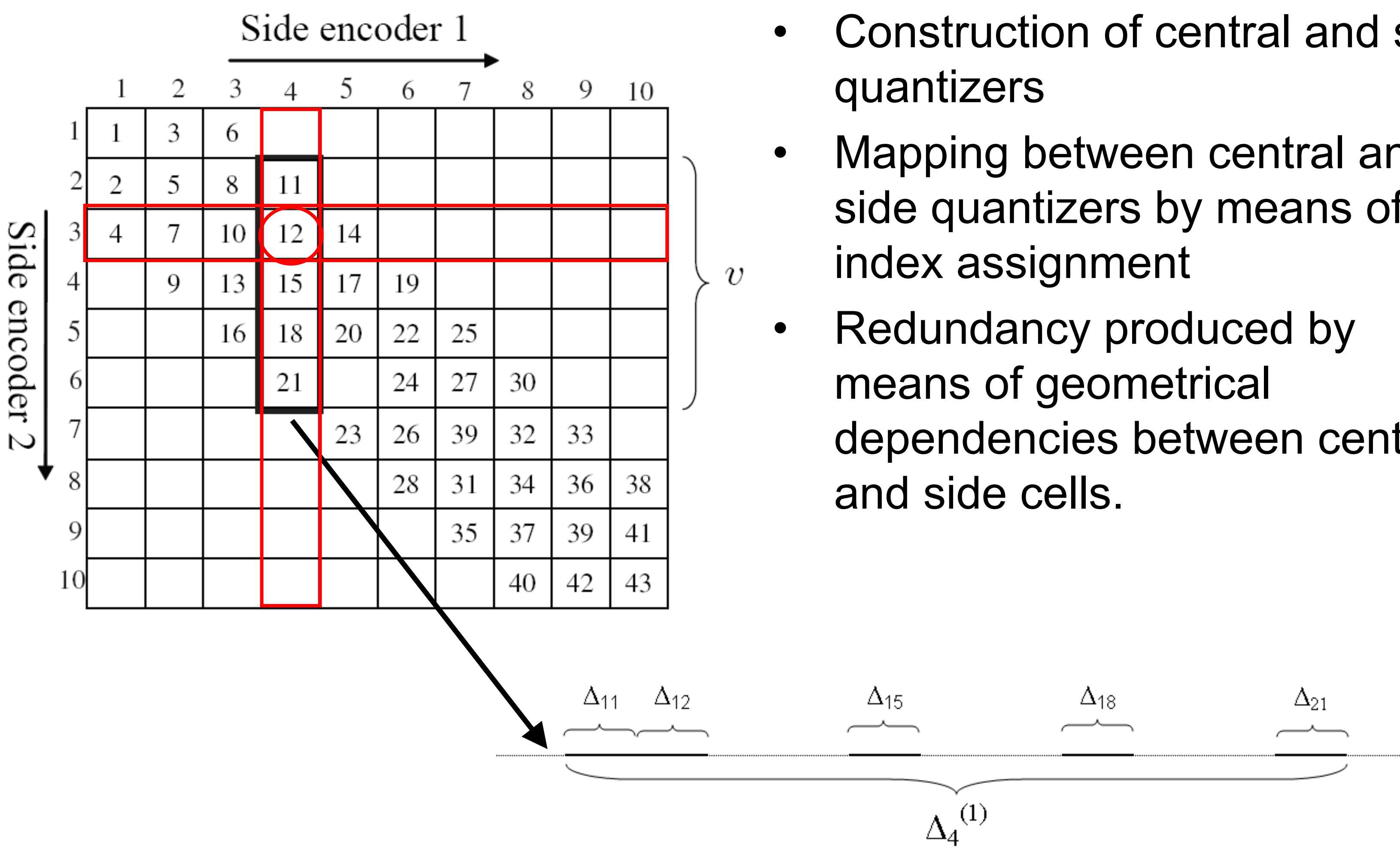




Context: joint source-channel coding | separate source and channel coding

Illustrative Example

- Quantization-based scalar MDC [Vaishampayan93,…]



- Construction of central and side quantizers
- Mapping between central and side quantizers by means of index assignment
- Redundancy produced by means of geometrical dependencies between central and side cells.

Developments of FlexCode

- Scalable scalar MDC (two-channel case)
  - uses predefined index assignment algorithms
  - quantizers are *designed analytically* for CE and CR cases
  - low complexity
- Lattice-based scalable scalar MDC (k-channel case)
  - index assignment found by means of geometrical construction of lattices
  - *analytical design* for both cases (CE and CR)
  - low complexity
- Rate allocation schemes obtained analytically

Scalability of MDC

- Scalability of FlexCode MDC:
  - Scalability of redundancy:
    - » optimal redundancy designed analytically (no training)
  - Scalability of rate:
    - » quantizers designed analytically (no training, no iterative procedures)
    - » no need to store the codebooks
- Example of scalability (scalar, two-channel MDC)
  - Published work:

Janusz Klejsa, Marcin Kuropatwinski, and W. Bastiaan Kleijn, "**Adaptive resolution-constrained scalar multiple-description coding**," in Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing, Mar. 2008, pp. 2945-2948.

Scalar Two-Channel MDC

- Total distortion

$$D_{total} = (1-p)^2 D_0 + 2(1-p)pD_s + p^2\sigma^2$$

- Design problems:
  - minimize D_{total} subject to constrained average bit-rate (entropy, CE)
  - minimize D_{total} subject to constrained bit-rate (resolution, CR)
- What can be designed?
  - Central and side quantizers
  - Index assignment matrix (index assignment algorithm, redundancy – related to number of diagonals υ)
- Goals:
  - Analytical formulas to design the quantizers
  - Analytical formulas to compute the optimal redundancy (or υ)

Analytical Design Method

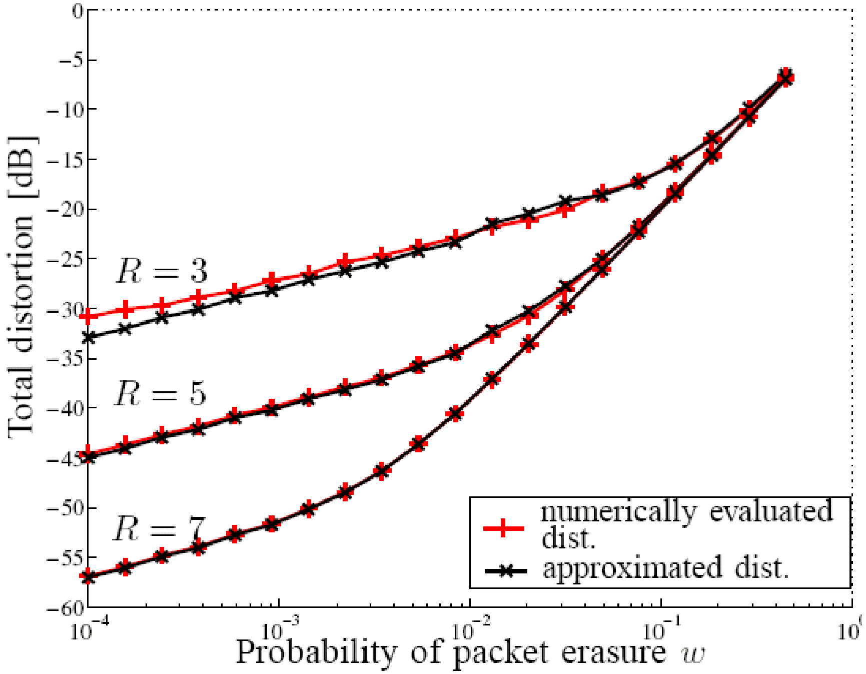
- Optimal redundancy minimizes the total distortion

$$D_{total}(R,v) = (1-p)^2 D_C(R,v) + 2(1-p)p D_S(R,v) + p^2\sigma^2$$

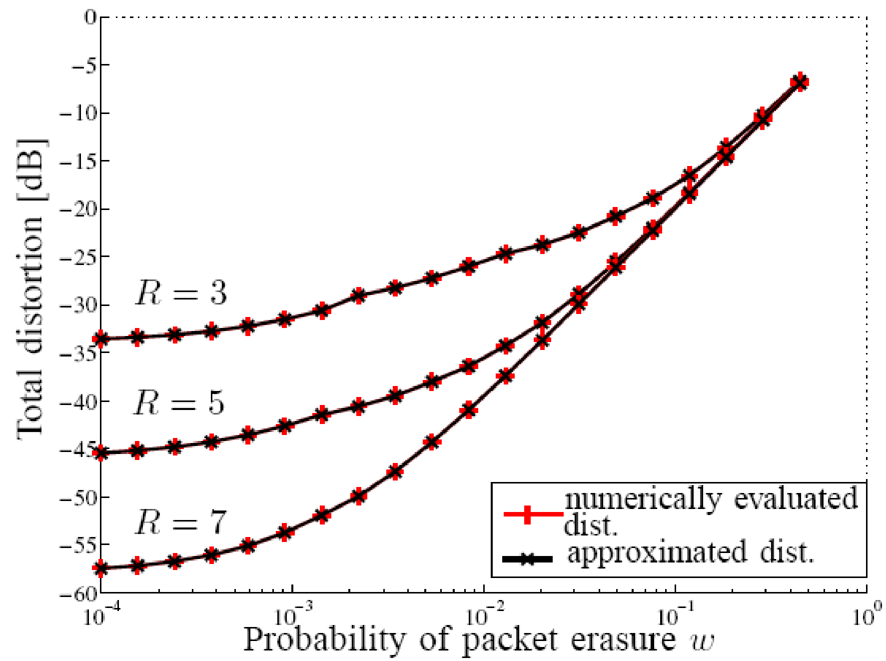
- Key point: find $D_C(R,v)$ and $D_S(R,v)$
- Solution:
  - high-rate quantization theory
  - analyzing geometry of the side quantization cell
  - parameterization of index assignment algorithms
- Result:
  - analytical expression for the total distortion parameterized in terms of redundancy
    - Analytical optimization of the redundancy
    - Optimal scaling of the quantizers

Results: Flexibility

- Adaptive quantization scheme (scalable rate&redundancy)



Resolution constrained case

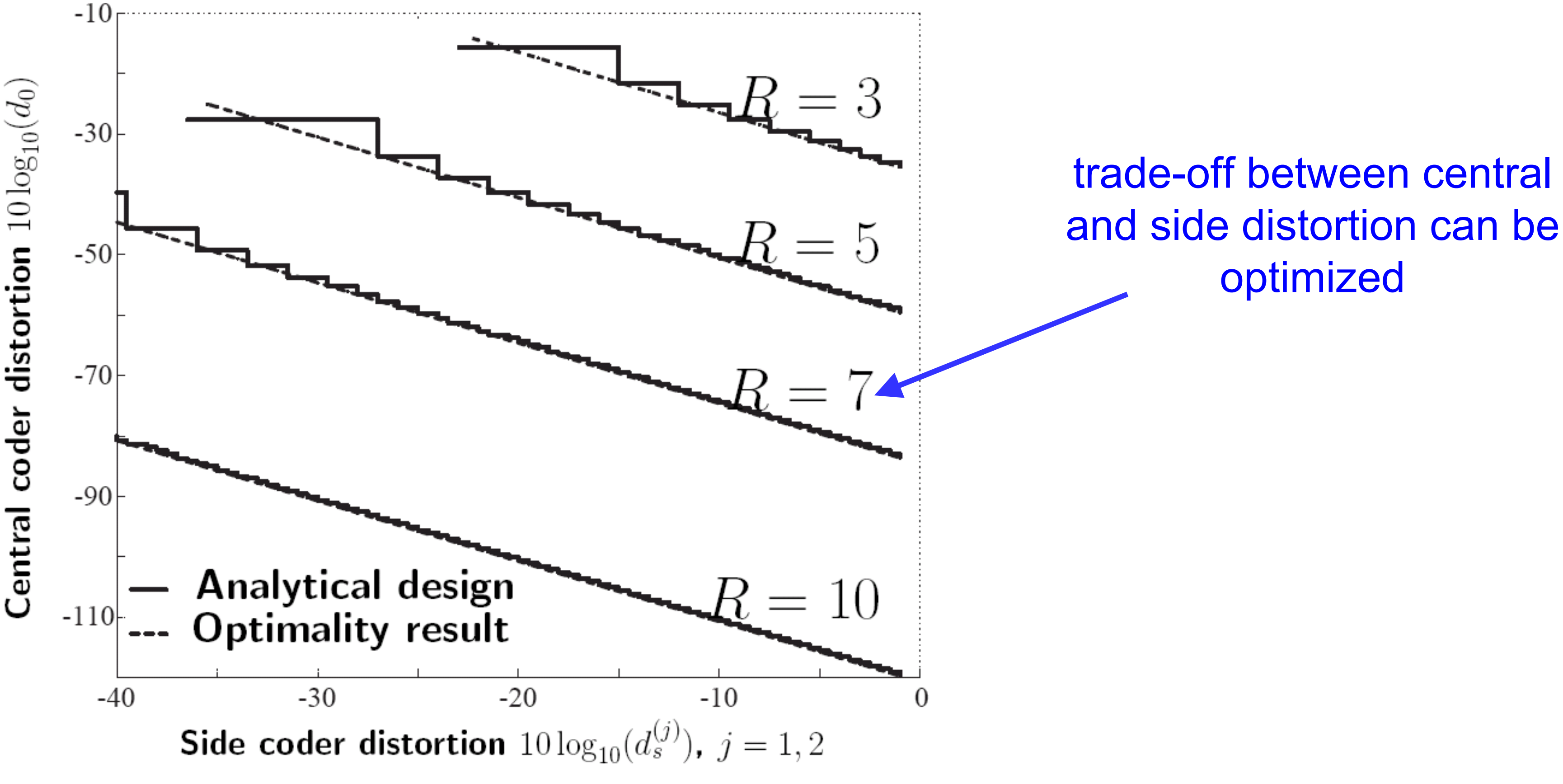


Entropy constrained case

Flexibility in terms of channel properties

Results: Scalability

- Adaptive quantization scheme (scalable rate&redundancy)



Outline

- FlexCode Project
- Multiple Description Coding (MDC) in a Nutshell
- Application to Audio Coding
  - Flexible Audio Coder
  - Exemplar Scenario
  - Optimal Rate Allocation
- Conclusions

Flexible Audio Coder (1)

- Source modeling and High-rate theory:
  - computable codebooks
  - scalable quantizers (inc. MDC) (CE, CR)
  - perceptual model derived from the signal model
- AR and transform modeling converged
  - KLT-based coder: transform derived from the model
  - quantization in the weighted domain
- Model and transform coefficients need to be transmitted
  - How to allocate the rate between model and signal?
    - Published work:

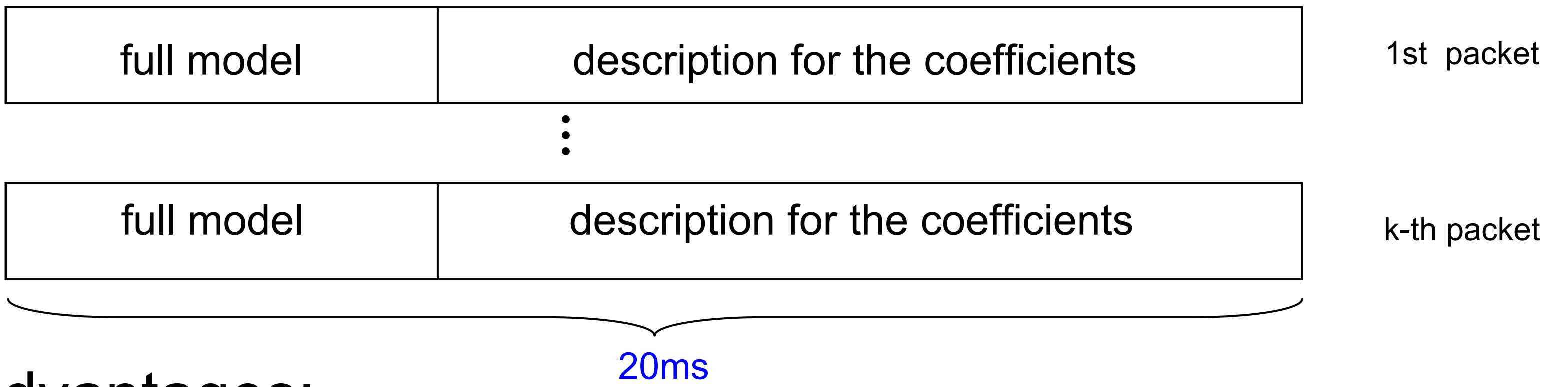
      W. Bastiaan Kleijn, and Alexey Ozerov, "**Rate distribution between model and signal**," In IEEE Worksh. on Apps. of Signal Processing to Audio and Acoustics (WASPAA), Mohonk, NY, Oct. 2007, pp. 243-246.

Flexible Audio Coder (2)

- Application of MDC
  - Transform coefficients
    - Major part of the bit-stream
    - Multi-level quality works fine with transform coefficients
      - Reverse water-filling appears naturally
  - Model parameters?
    - Problems:
      - Mismatch between encoder and decoder
      - Is a degraded quality of the model acceptable?
      - May result in unreasonable complexity of the coder
      - Rate spent on the model is low.
        Is it worth to consider MDC there?

Exemplar Scenario (1)

- Each description contains full information about the model, *k* descriptions are created.



- Disadvantages:
  - leads to large rate overhead (for large *k*)
- Advantages
  - Descriptions equally important (symmetrical balanced case);
  - Any subset of descriptions may be used for reconstruction;
  - In practice *k* is always low (≈two)=> the rate overhead acceptable
  - MDC does not introduce additional delay.

Exemplar Scenario (2)

- Setup:
  - a symmetrical, EC, two-channel MDC used for transform coefficients / signal
  - a model must be always received to decode the signal
- Rate spent for the description:

$$L_{X|\Theta} = -E\{\log(f_{X|\Theta}(X|\theta(X))V\}$$

Side cell volume

- Optimal criterion for selecting the model:

$$\hat{\theta} = \arg\max_{\theta} f_{X|\Theta}(x|\theta(x))$$

Quantized model

- Rate spent for the model

$$L_{\bar{\Theta}} = -E\{\log(f_{\bar{\Theta}}(\bar{\theta}(X))\}$$

- What is the optimal rate allocation between the model and the descriptions?

Exemplar Scenario (3)

- Optimal rate for the model minimizes total rate required to transmit the signal at certain distortion

$$\underbrace{L_X}_{\text{total rate}} = -\overset{\text{no. of descriptions}}{k}E\{\underbrace{\log(f_{\bar{\Theta}}(\bar{\Theta}(X))}_{\text{model rate}}+\underbrace{\log(f_{X|\bar{\Theta}}(X|\theta(X))V)}_{\text{rate per description}}\}$$

$$V = f(Distortion,\ Redundancy,\ Cell\ Geometry)$$

- Model selection and signal quantization decoupled by the index of resolvability

$$L_X = -kE\Big\{\underbrace{\log(f_{\bar{\Theta}}(\bar{\Theta}(X)) + \log\frac{f_{X|\bar{\Theta}}(X|\theta(X))}{f_{X|\hat{\Theta}}(X|\theta(X))}}_{= \Psi(X, \hat{\Theta}(X), \bar{\Theta}(X))} + \log(f_{X|\hat{\Theta}}(X|\theta(X))V)\Big\}$$

Exemplar Scenario (4)

- Optimal rate allocation for the model within mean of the index of resolvability

$$E\left\{\Psi(X,\hat{\Theta}(X),\bar{\Theta}(X))\right\} = \underbrace{E\left\{\log(f_{\bar{\Theta}}(\bar{\Theta}(X)) + \log\frac{f_{X|\bar{\Theta}}(X|\theta(X))}{f_{X|\hat{\Theta}}(X|\theta(X))}\right\}}$$

•Trade-off: rate spent on the model vs. penalty on using imperfect model

•Solvable analytically!

•Does not depend on the rate constraint!

•Does not depend on the redundancy!

Exemplar Scenario - Summary

- Optimal rate for the model obtained within the index of resolvability
  - The optimal rate is constant
  - Does not depend on the total rate and the redundancy
- Optimal redundancy obtained during designing MDC
  - Depends on the geometry of the quantizers
  - Depends on the channel (erasure probability)
  - Does not depend on the rate (constraint)
- Optimal scaling of the quantizers
  - Depends on the optimal redundancy and rate constraint

- Optimal rate allocation depends on the scenario.

Conclusions

- FlexCode approach to MDC
  - Usage of quantization-based multiple-description schemes to facilitate scalability
  - Codebooks computed on-line allow for building adaptive coder
  - Optimal rate allocation schemes can be derived for specified scenarios
  - Low-complexity (practical point of view)
  - Optimality for a finite delay

»Questions?

Flexible Multiple Description Coding of Audio

Janusz Klejsa

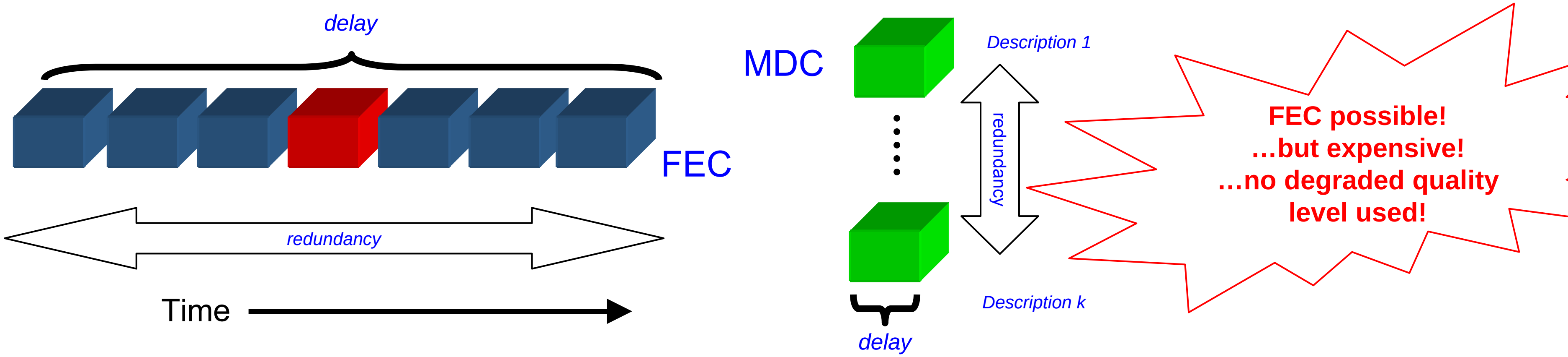
FlexCode Public Seminar
Tampere, June 16, 2008

MDC vs. FEC (1)

Four good reasons to use MDC (instead of FEC):

1) graceful degradation of quality (not possible with FEC)

2) good performance with low delay (expensive in case of FEC)



3) redundancy is easily scalable (difficult to scale the redundancy in FEC)

4) MDC is a joint source-channel coding without cross-layer optimization (FEC is used in separate source and channel coding setup)

- Finite delay, reasonable complexity!

MDC vs. FEC (2)

Three tricky points about using MDC:

1) Strong assumptions about the channel (…also an implementation issue)
2) Gracious-degradation of the performance vs. fall-of-the-cliff performance
3) Delay constraints are important

The Redundancy

- Allocating the redundancy

	1	2	3	4	5	6	7	8	9	10
1	1									
2		2								
3			3							
4				4						
5					5					
6						6				
7							7			
8								8		
9									9	
10										10

High redundancy

	1	2	3	4	5	6	7	8	9	10
1	1	3								
2	2	4	5							
3		6	7	9						
4			8	10	11					
5				12	13	14				
6						15	17			
7						16	18	19		
8							20	21	23	
9								22	24	25
10									26	27

	1	2	3	4	5	6	7	8	9	10
1	1	2	5	10	17	26	36	44	50	54
2	3	4	6	11	18	27	37	45	51	57
3	7	8	9	12	19	28	38	46	60	61
4	13	14	15	16	20	29	39	65	66	67
5	21	22	23	24	25	30	72	73	74	75
6	31	32	33	34	35	76	81	82	83	84
7	40	41	42	43	68	77	85	89	90	91
8	47	48	49	62	69	78	86	92	95	96
9	52	53	58	63	70	79	87	93	97	99
10	55	56	59	64	71	80	88	94	98	100

No redundancy

- How to obtain the quantizers *analytically*?
- How to allocate *optimal* redundancy?

  » High-rate theory!

FlexCode Details - Analytical Design Method (1)

- Central distortion (CR)

$$D_C = \frac{1}{12}\int_{t_1}^{t_r+1} f_X(x)\Delta(x,v)^2 dx = \frac{1}{12M^2v^2}\left(\int_{t_1}^{t_r+1}(f_X(x))^{\frac{1}{3}}dx\right)^3$$

- Central distortion (CE)

$$D_C = \frac{1}{12}\int_{t_1}^{t_r+1} f_X(x)\Delta(x,v)^2 dx$$

- Side distortion (CR)

$$D_S = \frac{f(v)}{v}\int_{t_1}^{t_r+1} f_X(x)\Delta(x,v)^2 dx = \frac{f(v)}{M^2v^3}\left(\int_{t_1}^{t_r+1}(f_X(x))^{\frac{1}{3}}dx\right)^3$$

- Side distortion (CE)

$$D_S = \frac{f(v)}{v}\int_{t_1}^{t_r+1} f_X(x)\Delta(x,v)^2 dx$$

Coefficients of quantization

- Result: total distortion can now be optimized analytically

Details - Analytical Design Method (2)

- Design of the quantizers (for optimal v)
  - central quantizer

$$\Delta(v) = \frac{1}{v} 2^{h(X)-R}$$

CE

$$\Delta(x,v) = \frac{1}{Mv} \frac{\int_{t_1}^{t_r+1} (f_X(x))^{\frac{1}{3}} dx}{(f_X(x))^{\frac{1}{3}}}$$

CR

  - side quantizer obtained by index assignment mapping

- Corollaries:
  - optimal redundancy does not depend on the rate
  - optimal scaling of the quantizers depends on the redundancy