LOSS CONCEALMENTS FOR LOW-BIT-RATE PACKET VOICE IN VOIP

Benjamin W. Wah

Department of Electrical and Computer Engineering
and the Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA

Plenary Presentation
IEEE Int'l Conf. on Computer Networks and Mobile Computing

October 22, 2003

Voice over IP

- Real-time:
  - Interactive communication
  - ITU G.114: end-to-end delay less than 400 msec acceptable
- Loss:
  - Long-burst or frequent short-burst intolerable
  - Some degradations on voice samples tolerable
- Packet network:
  - Packet size: less than MTU to avoid fragmentation
  - Packet rate: 20 - 30 packets per second
- Low bit-rate coded speech:
  - Error propagation
- Importance
  - Better loss-concealment schemes will allow seamless integration of wireless networks and the Internet for voice delivery

Environments

- Network-layer protocol — Loss unavoidable
  - IPv4: best-effort, no real-time support
  - IPv6: best-effort, may support real-time traffic
  - Wireless: future IP-based
- Transport-layer protocol — Loss of real-time voice not handled
  - TCP: reliable but not suitable for real-time
  - UDP: unreliable
- Application-layer protocol — Loss of real-time voice not handled
  - RTP: no loss recovery scheme
  - H.323: umbrella standard for interoperability
- Packet losses in real-time voice communications left for end-point applications

Problem Addressed in this Talk

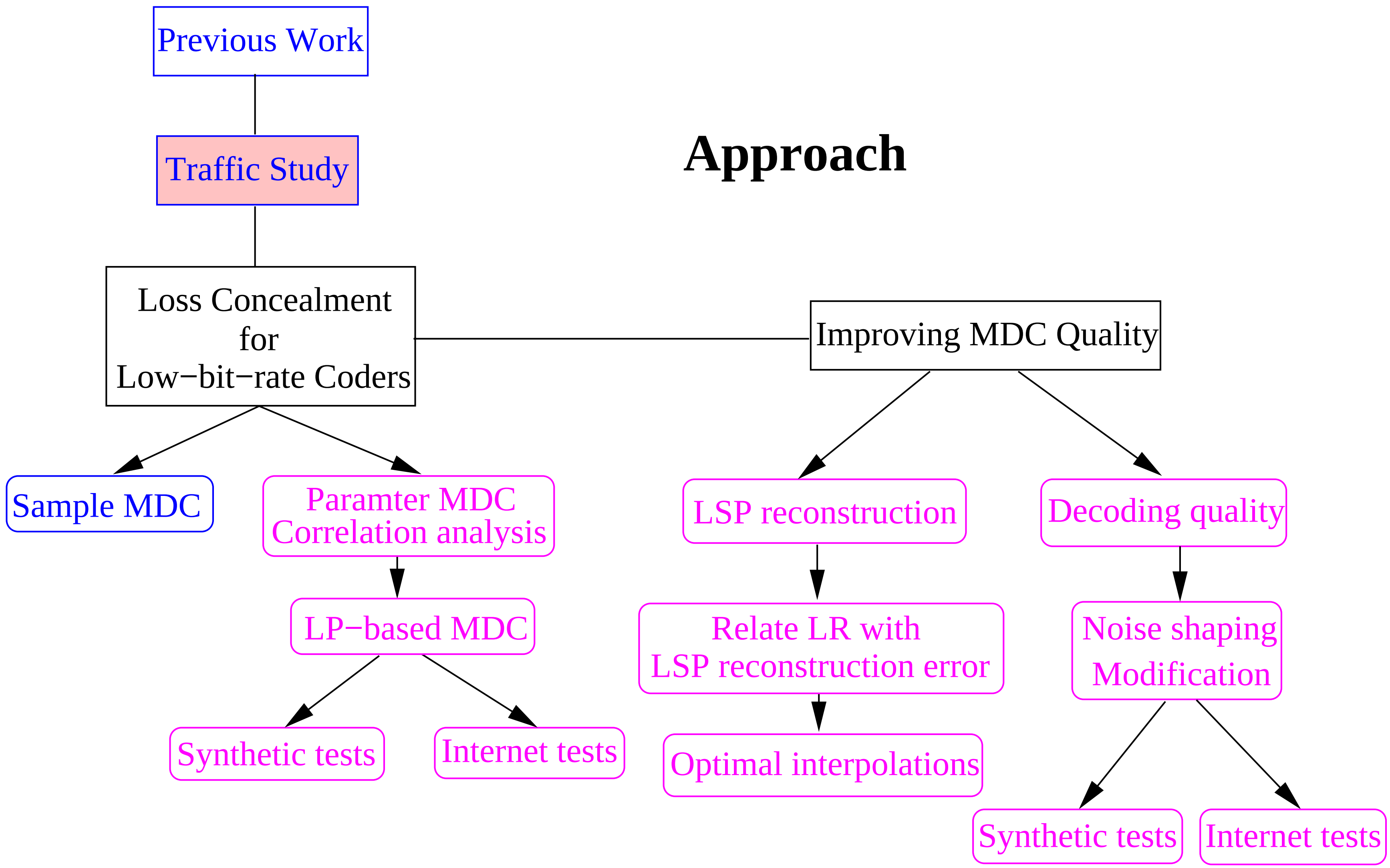
Design, analyze and evaluate robust end-to-end loss-concealment schemes

- Allow reliable and real-time low bit-rate voice transmissions
- Unreliable IP networks, like the Internet and wireless wide area networks

Applications:

- Internet telephony
- Teleconferencing
- Wireless communications

Approach

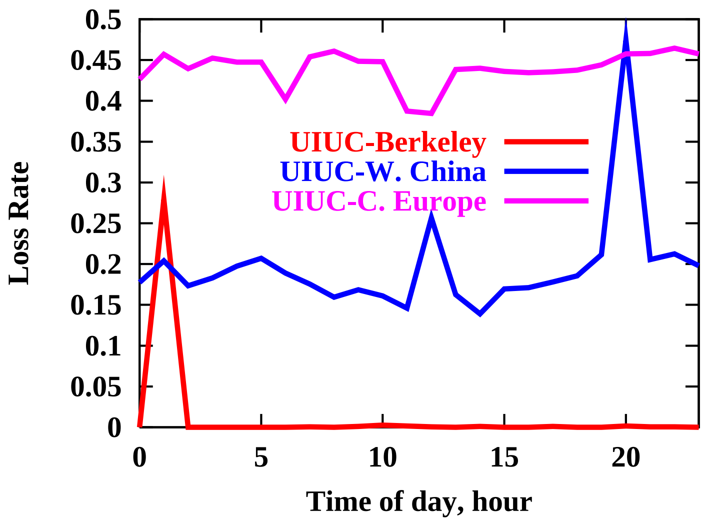

Previous Work
Traffic Study
Loss Concealment for Low−bit−rate Coders
Improving MDC Quality
Sample MDC
Paramter MDC Correlation analysis
LP−based MDC
Synthetic tests
Internet tests
LSP reconstruction
Relate LR with LSP reconstruction error
Optimal interpolations
Decoding quality
Noise shaping Modification
Synthetic tests
Internet tests

IP Voice Traffic Loss Characteristics

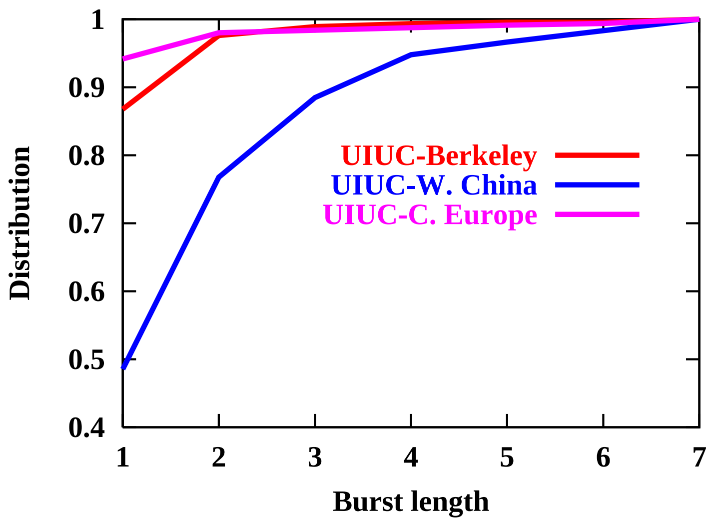
- Example connections

Connection	Loss rate
UIUC-Berkeley	low-medium
UIUC-Western China	medium-high
UIUC-Central Europe	high

- Loss behavior



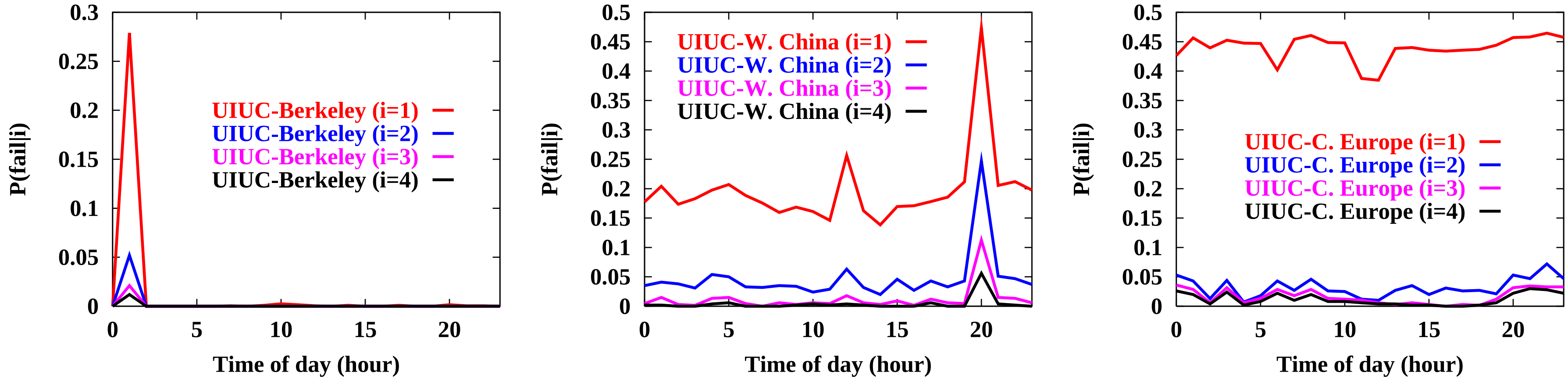
Loss rate can go up to 50%



Most losses have short burst lengths

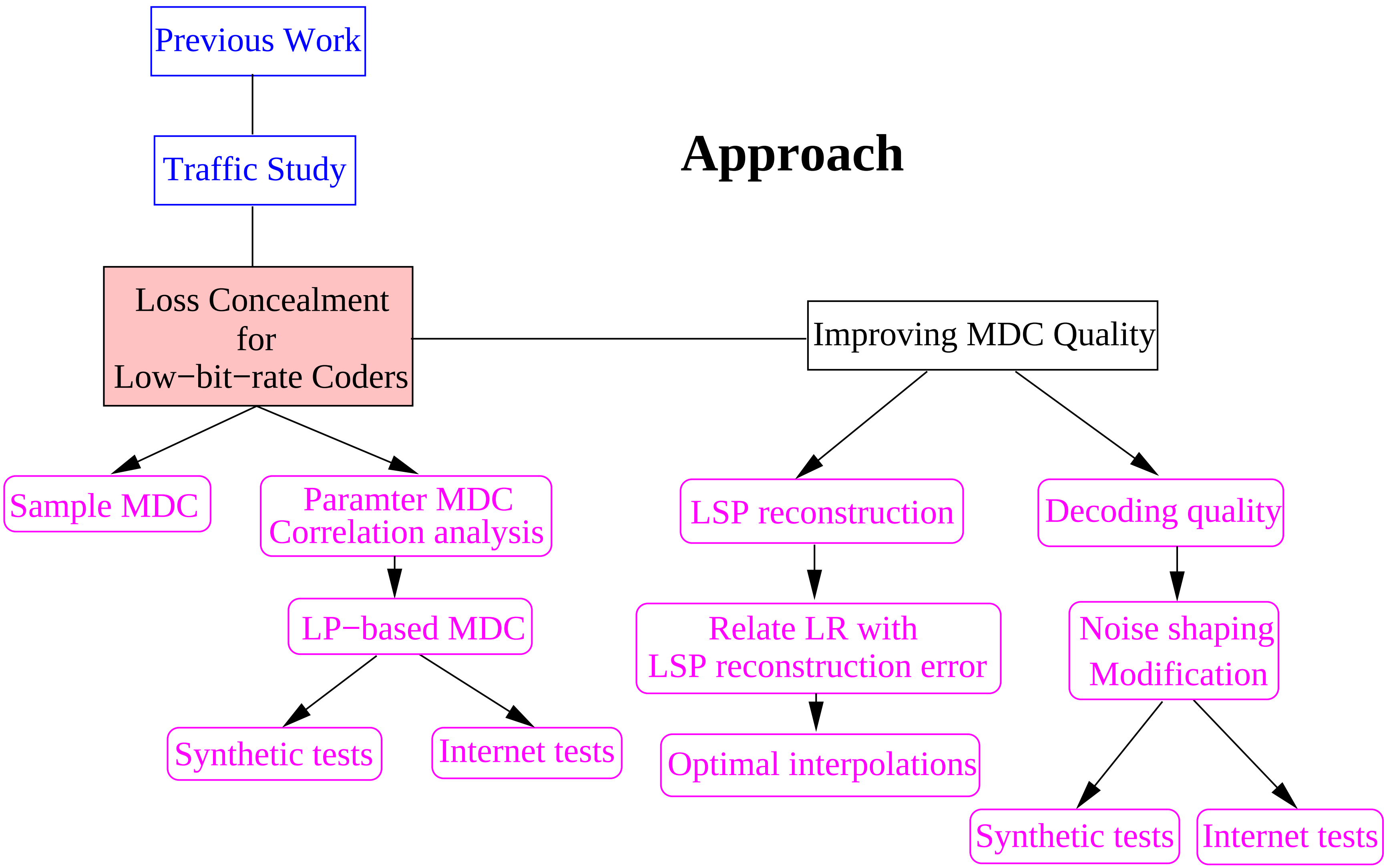
Reducing Unrecoverable Loss by Interleaving

- Bursty losses are difficult to handle
- Interleaving: disperse burst losses to isolated losses
- P(fail|i): prob. of losses that cannot be recovered under interleaving factor i



- Small interleaving factor 2 – 4 is enough
- Multiple-description coding is promising

Approach


Previous Work
Traffic Study
Loss Concealment for Low−bit−rate Coders
Improving MDC Quality
Sample MDC
Paramter MDC Correlation analysis
LP−based MDC
Synthetic tests
Internet tests
LSP reconstruction
Relate LR with LSP reconstruction error
Optimal interpolations
Decoding quality
Noise shaping Modification
Synthetic tests
Internet tests

Low Bit-Rate Coders Tested

	Bit rate (bps)	Quantization of LSP	Excitation
FS CELP	4800	scalar	stochastic code/adaptive code
ITU G.723.1 (I)	5300	predictive-split VQ	algebraic code/adaptive code
ITU G.723.1 (II)	6300	predictive-split VQ	multi pulse/adaptive code
FS MELP	2400	multi-stage VQ	mixed pulse- and noise-like
ITU G.729	8000	predictive-split VQ	algebraic code/adaptive code

- G.729 is popular in video conferencing, audiovisual communications, VoIP, and wireless communications
  - Acceptable bandwidth: 16k in G.728, 32k in G.726 (ADPCM), and 64k in G.711 (PCM) may be too high for IP or wireless applications
  - Lower computation: complexity reduced version G.729A
  - Shorter delay: 10ms, compared to 30ms of G.723.1
  - High quality: MOS 3.9, comparable to 32k ADPCM

Voice Streams Tested

Index	Length (ms)	Speakers
1	21432	2 male, 1 female
2	22560	2 male, 1 female
3	4424	1 female
4	5091	1 female
5	4160	1 male
6	4082	1 male
7	4867	1 male, 1 female
8	73615	1 male, 1 female

Objective Measures

- Itakura-Saito likelihood ratio

$$LR = \frac{a_r R_o a_r^T}{a_o R_o a_o^T}$$

  - a_r: vector of LP coefficients of reconstructed speech
  - a_0: vector of LP coefficients of original speech
  - R_0: correlation matrix derived from original speech

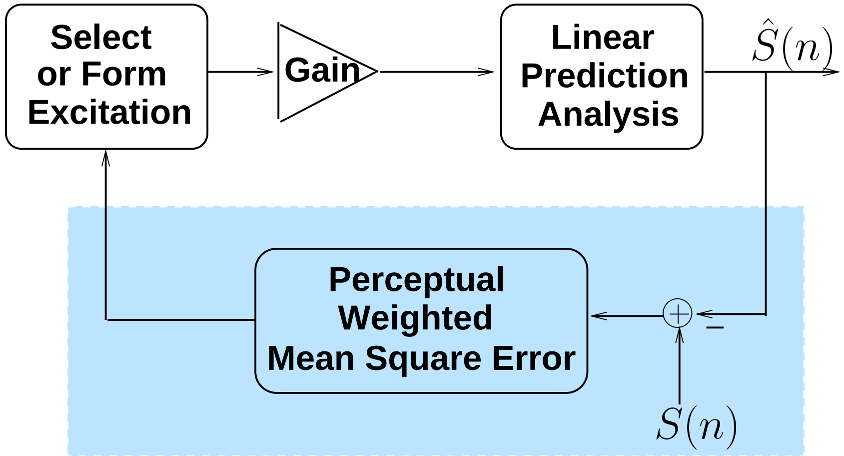
- Cepstral distance:

$$CD = 4.34[(c_{o,\,0} - c_{r,\,0})^2 + 2\sum_{i=1}^{\infty}(c_{o,\,i} - c_{r,\,i})^2]^{\frac{1}{2}} \text{ [dB]}$$

  - $c_{o,i}$: cepstra of original sample i
  - $c_{r,i}$: cepstra of reconstructed sample i

Typical Linear Predictive Coder

$$H(w) = \frac{1}{A(w)} = \frac{1}{1 - \sum_{k=1}^{10} a_k e^{jwk}}$$

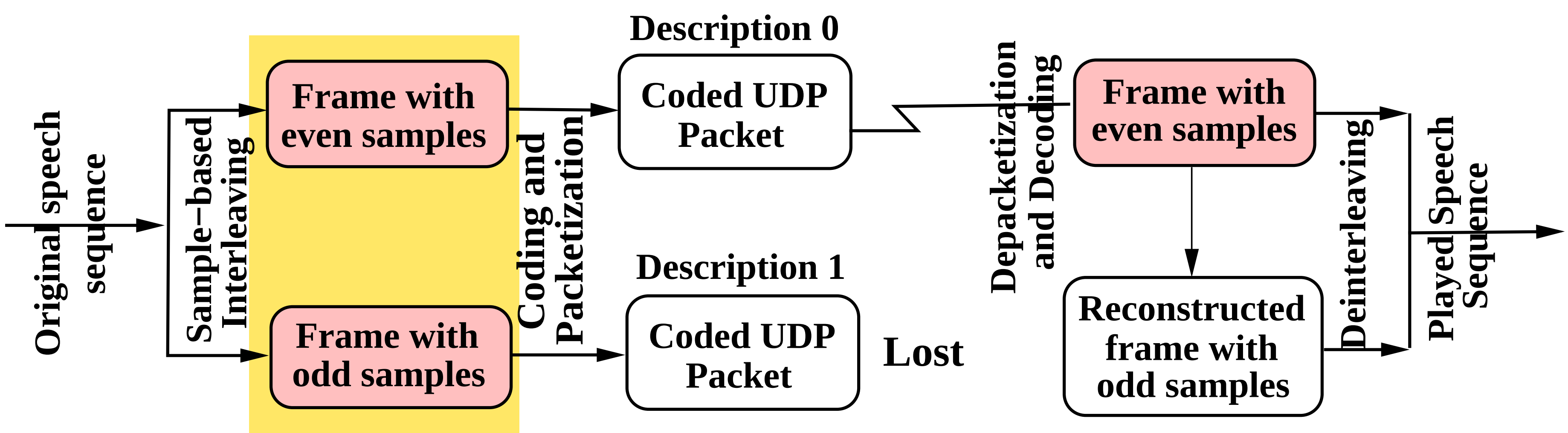


Major techniques:

- Frame-oriented
- Linear prediction analysis, coefficients generally represented by LSP
- Excitations: pitch information and random noise
- Can be open-loop or closed-loop

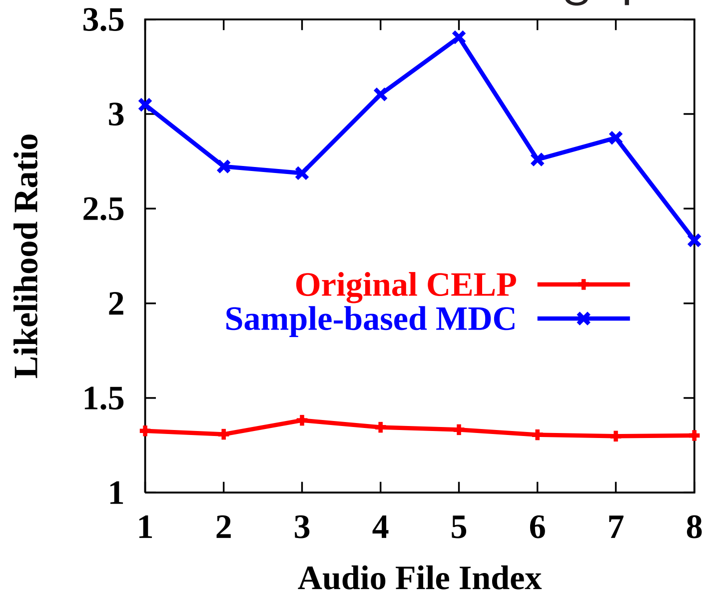
- FS CELP$\star$, ITU G.723.1 ACELP, ITU G.723.1 MP-MLQ, MELP, and ITU G.729

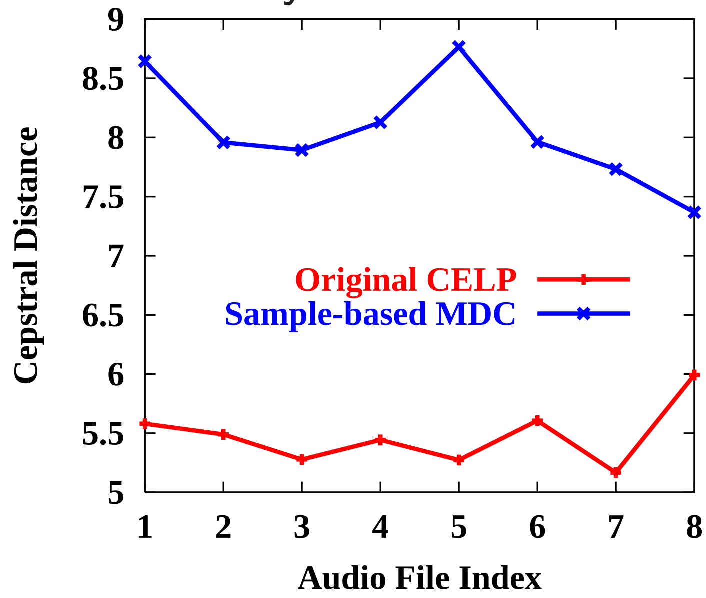
Coder-Independent Sample-Based MDC



Performance of Coder-Independent Sample-Based MDC

Coding quality degrades dramatically

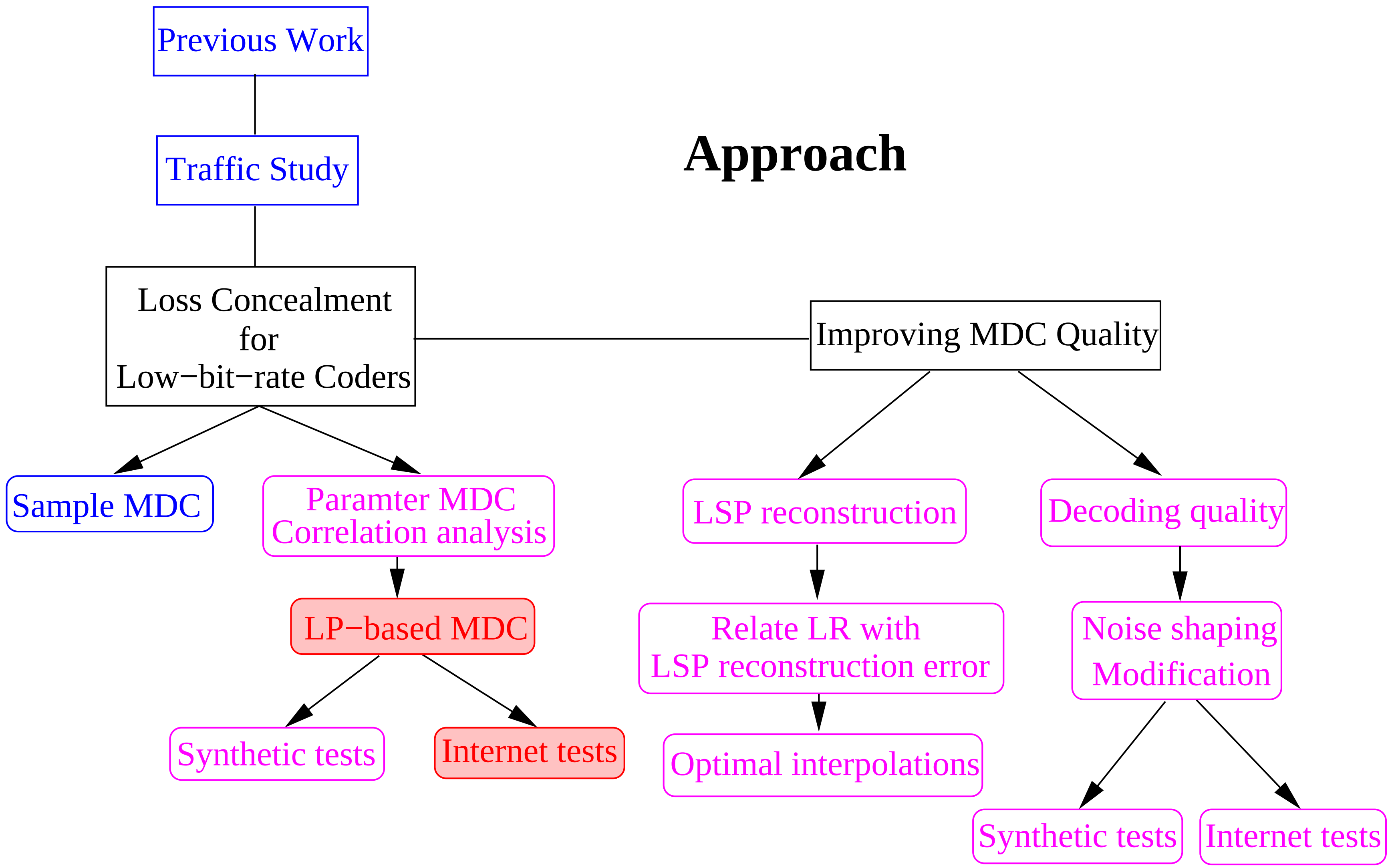




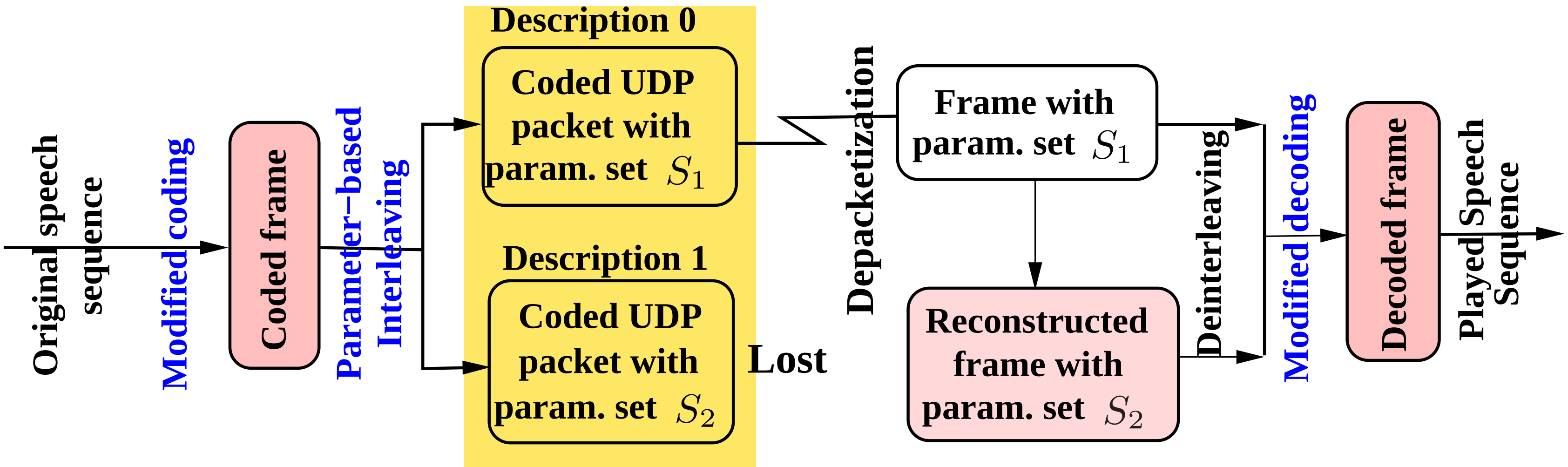
Drawbacks:

- Aliasing: caused by down sampling
- Coding-frame time span lengthened

Approach


Previous Work
Traffic Study
Loss Concealment for Low−bit−rate Coders
Improving MDC Quality
Sample MDC
Paramter MDC Correlation analysis
LP−based MDC
Synthetic tests
Internet tests
LSP reconstruction
Relate LR with LSP reconstruction error
Optimal interpolations
Decoding quality
Noise shaping Modification
Synthetic tests
Internet tests

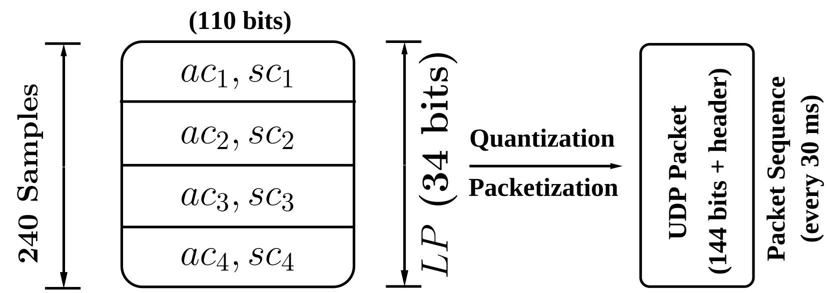
Coder-Dependent Parameter-Based MDC



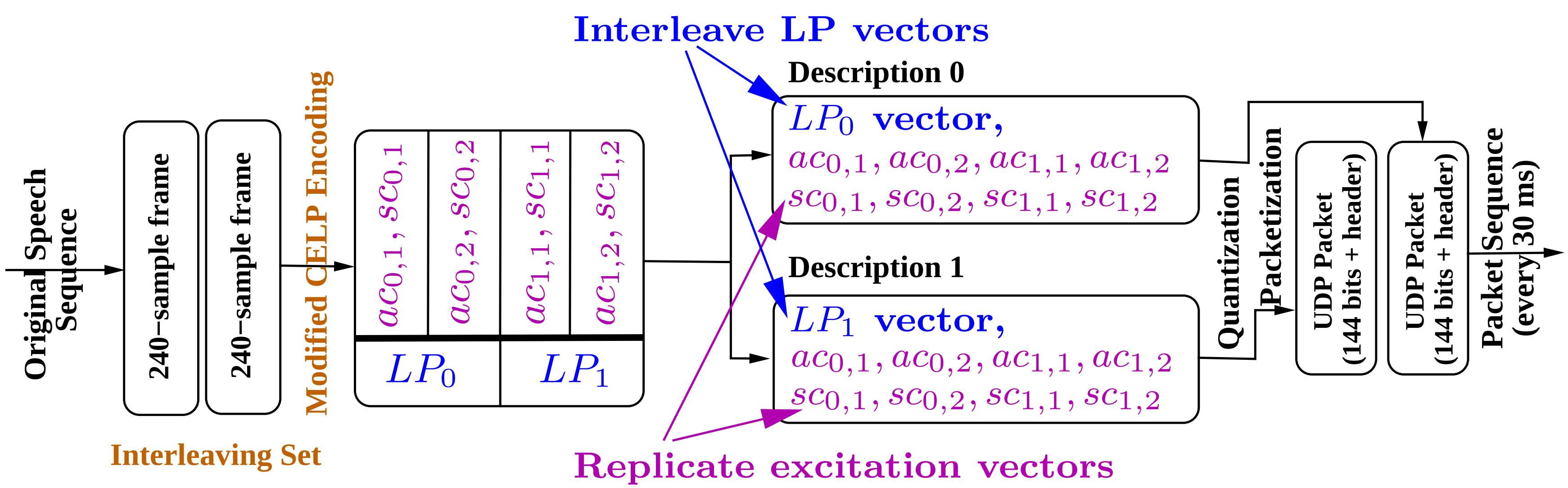
- Parameters of linear predictive coders:
  - Linear predictor equivalent representations:
    - Reflection coefficient (RF), Log area ratio (LAR), LSP
  - Excitation
- MDC design by correlation analysis

FS CELP SDC and LP-Based Two-Way MDC

- FS CELP SDC:



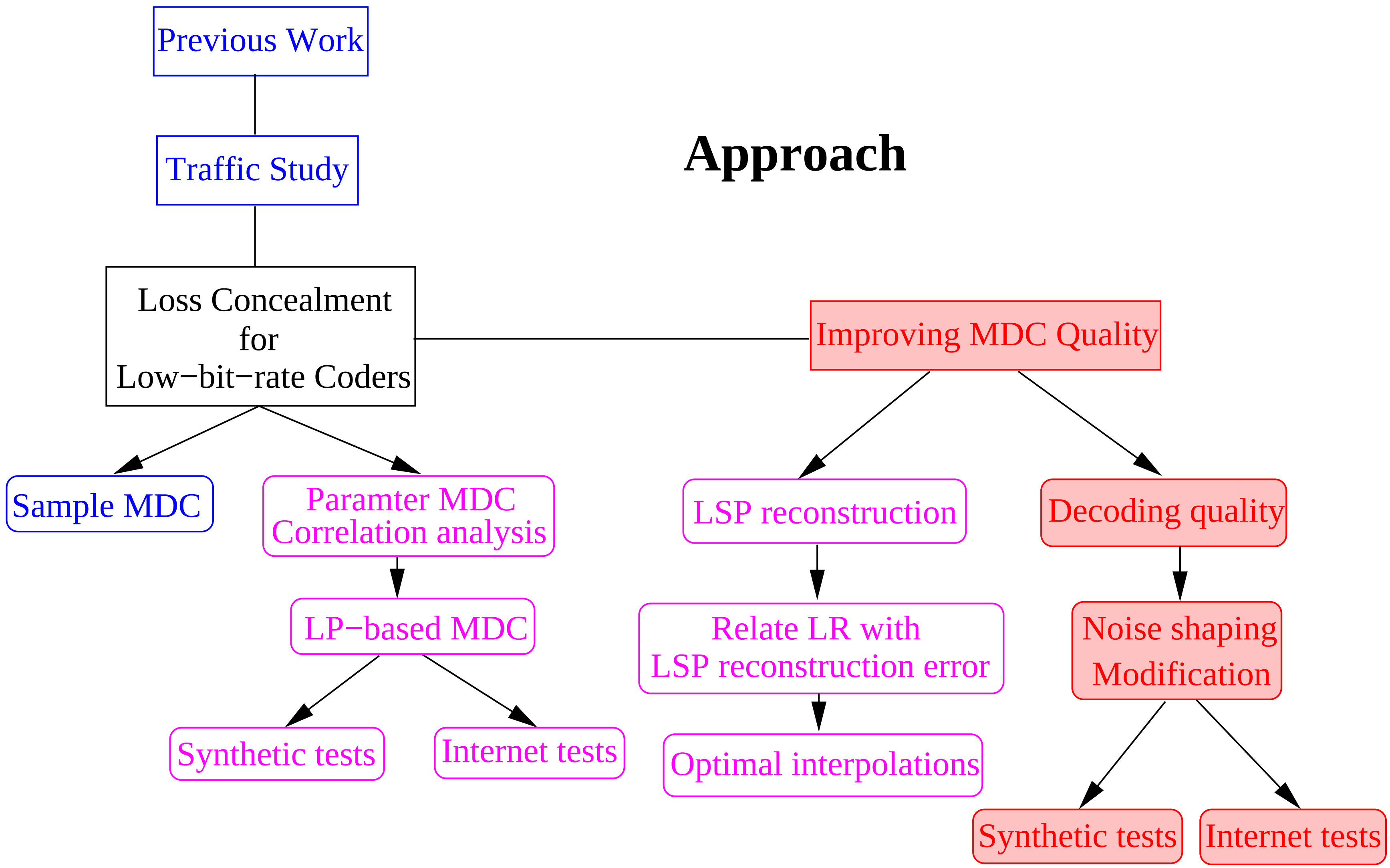
- Construction of two-way MDC:



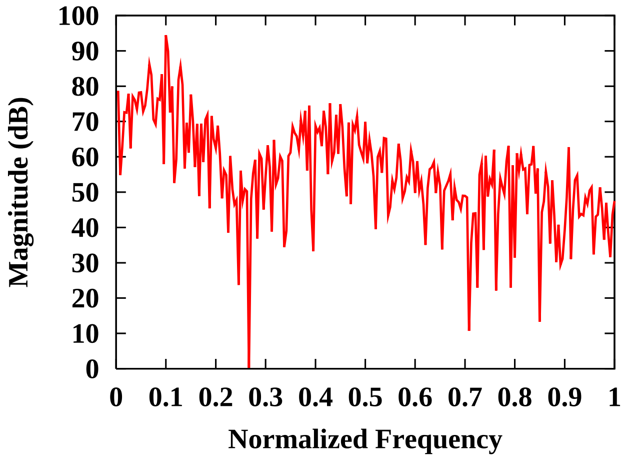
G.729 Built-in Loss Concealment

- LSPs are duplicated from previous frame
- Both adaptive/fixed codebook gains are duplicated from previous frame, but are attenuated to gradually reduce their impact
- Excitation reconstruction depends on the property of the previous frame
  - If voiced: fixed codebook contribution is set to 0 and pitch delay is duplicated from previous frame.
  - If unvoiced: adaptive codebook contribution is set to 0 and fixed codebook contribution is generated randomly.

Approach


Previous Work
Traffic Study
Loss Concealment
for
Low−bit−rate Coders
Improving MDC Quality
Sample MDC
Paramter MDC
Correlation analysis
LP−based MDC
Synthetic tests
Internet tests
LSP reconstruction
Relate LR with
LSP reconstruction error
Optimal interpolations
Decoding quality
Noise shaping
Modification
Synthetic tests
Internet tests

Causes for Quality Degradation in FS-CELP



Speech perception:

- Valley noise more noticeable
- Formant important

- Significant higher coding-noise inside formant regions due to MDC

a) Two-way MDC

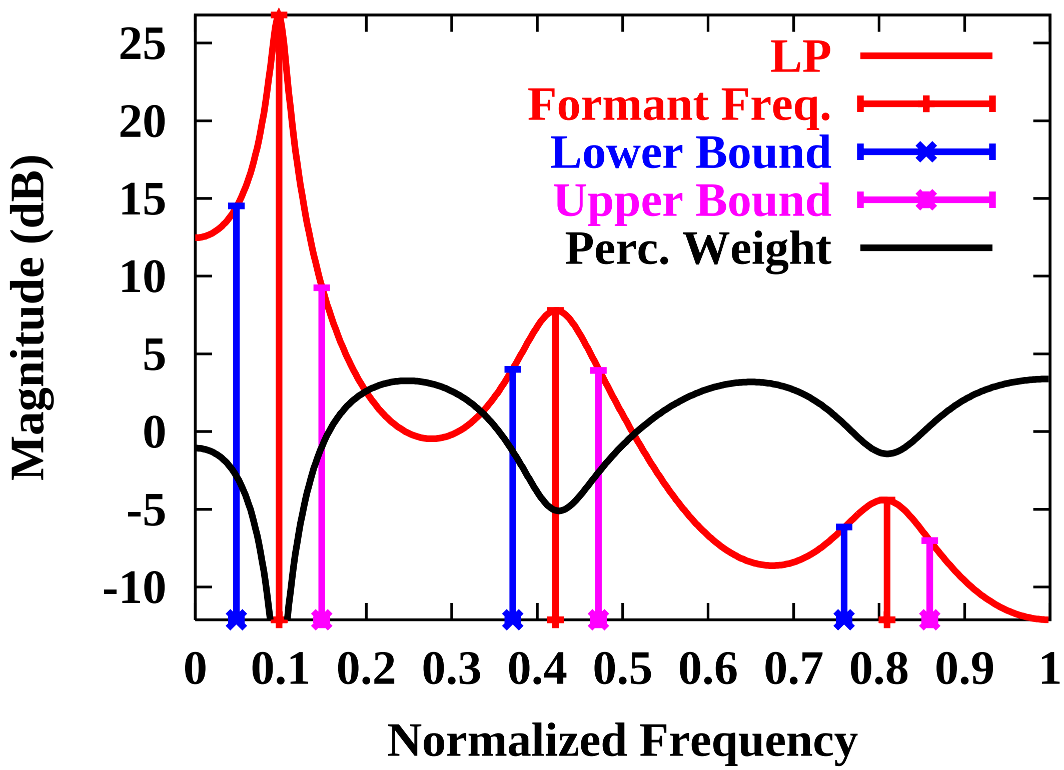
	E_{R_F}	$E_{\overline{R_F}}$
SDC	1.1591e+8	2.3976e+7
Two-way MDC	2.3049e+8	4.4340e+7
Ratio	1.99	1.85

b) Four-way MDC

	E_{R_F}	$E_{\overline{R_F}}$
SDC	2.3507e+8	4.4775e+7
Four-way MDC	8.6771e+8	1.2685e+8
Ratio	4.69	3.83

Perceptual Weighting Filter

- Goal: noise shaping
  - De-emphasize coding noise (distortions) inside formant regions

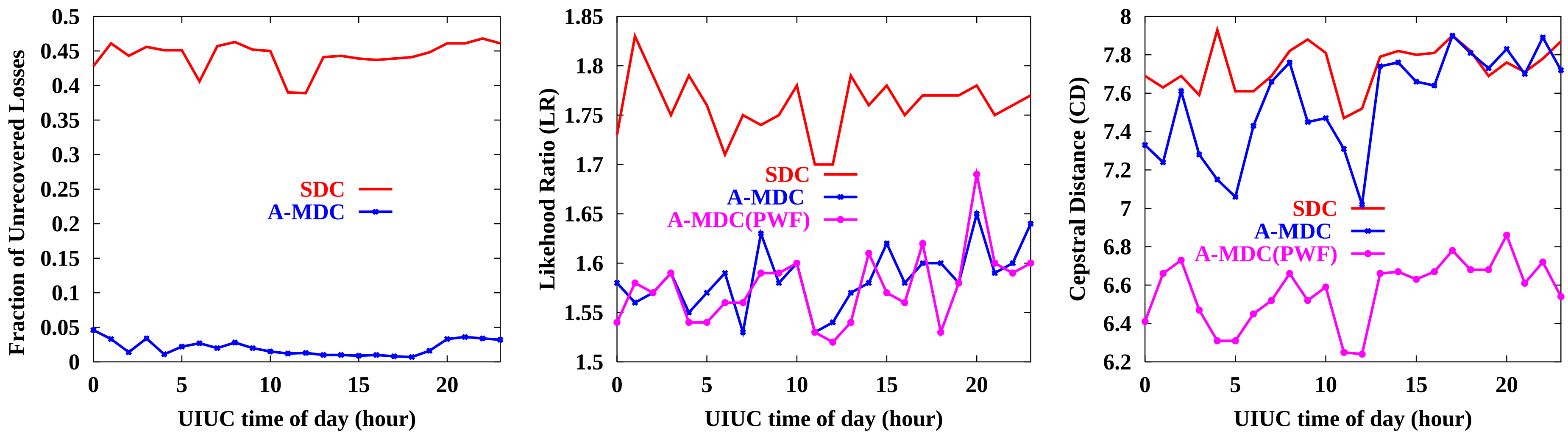


Internet Test Setup

- Components:
  - Sender
  - Receiver: 200 msec jitter buffer, start clock when first packet arrives
  - Internet simulator: delay and drop packet according to traffic traces
- Comparison between:
  - SDC
  - Adaptive MDC: dynamically switch between two-way and four-way MDC depending on loss conditions
- Comparison metrics:
  - Quality in LR and CD
  - Fractions of unrecoverable losses

Internet Tests on FS-CELP

UIUC-Central Europe



Summary of adaptive MDC:

- Recovering the decoding state and effective in reducing unrecovered losses
- SDC with no loss: $LR = 1.33$, $CD = 5.55$
- Improved CD