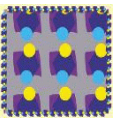# Listening to Perceptual Audio Coders:

Speech Coding and Music

Heiko Purnhagen, Bernd Edler
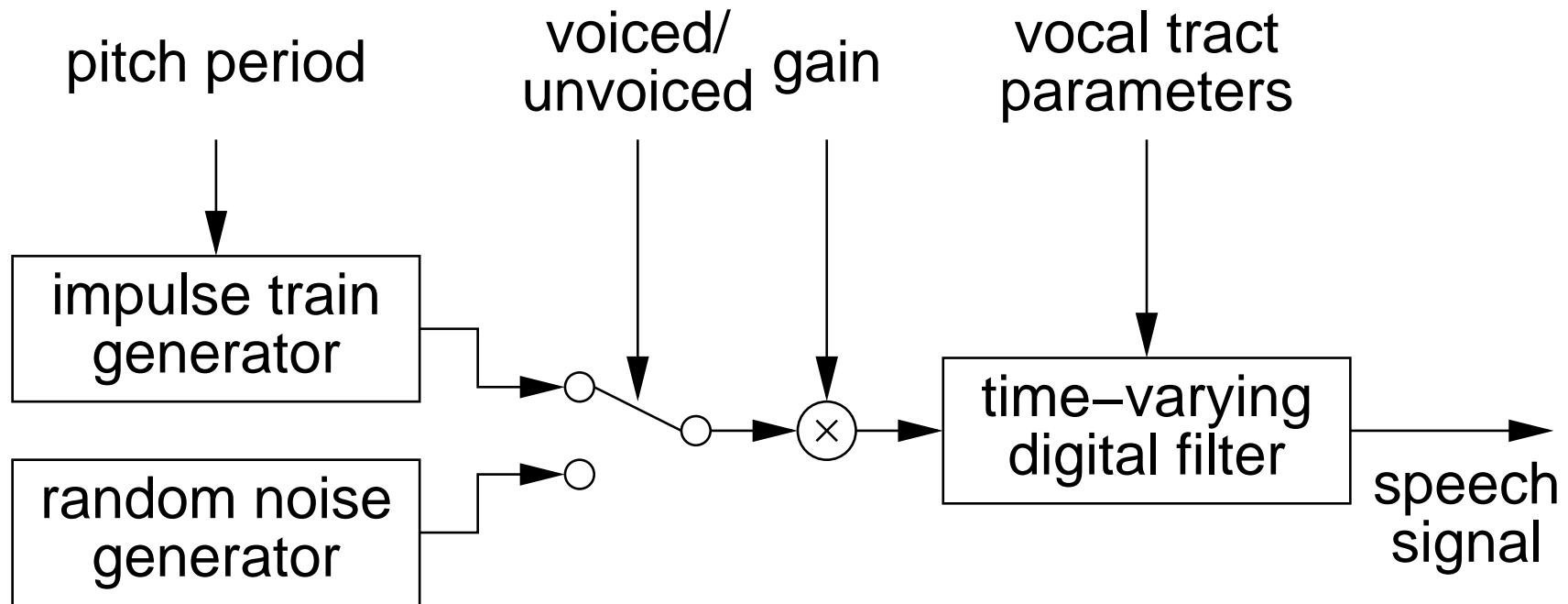
Laboratorium für Informationstechnologie
Universität Hannover

AES 112th Convention, München   11-May-2002

Heiko Purnhagen (purnhage@tnt.uni-hannover.de)
Laboratorium für Informationstechnologie
Abteilung Systemtechnik

lfi

Universität Hannover

Speech synthesis model based on linear predictive coding (LPC)

Heiko Purnhagen (purnhage@tnt.uni-hannover.de)
Laboratorium für Informationstechnologie
Abteilung Systemtechnik
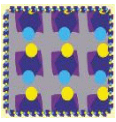
Universität Hannover
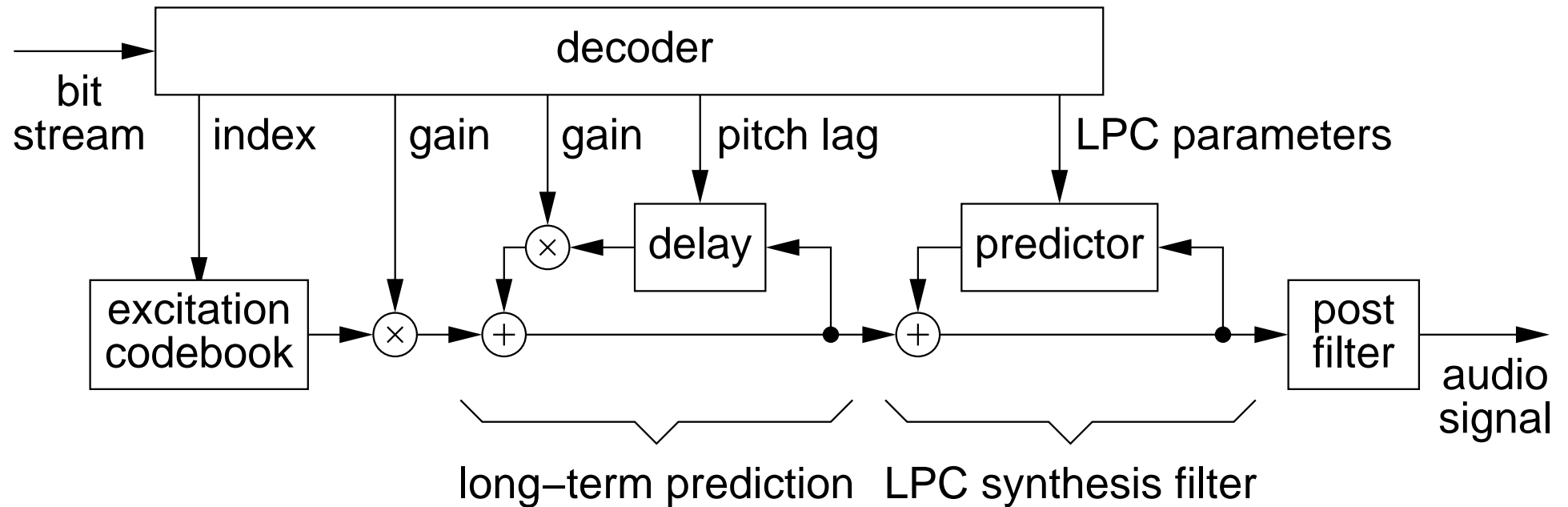
Code excited linear prediction (CELP) decoder
with long-term prediction (LTP)
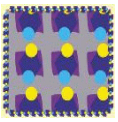
Universität Hannover
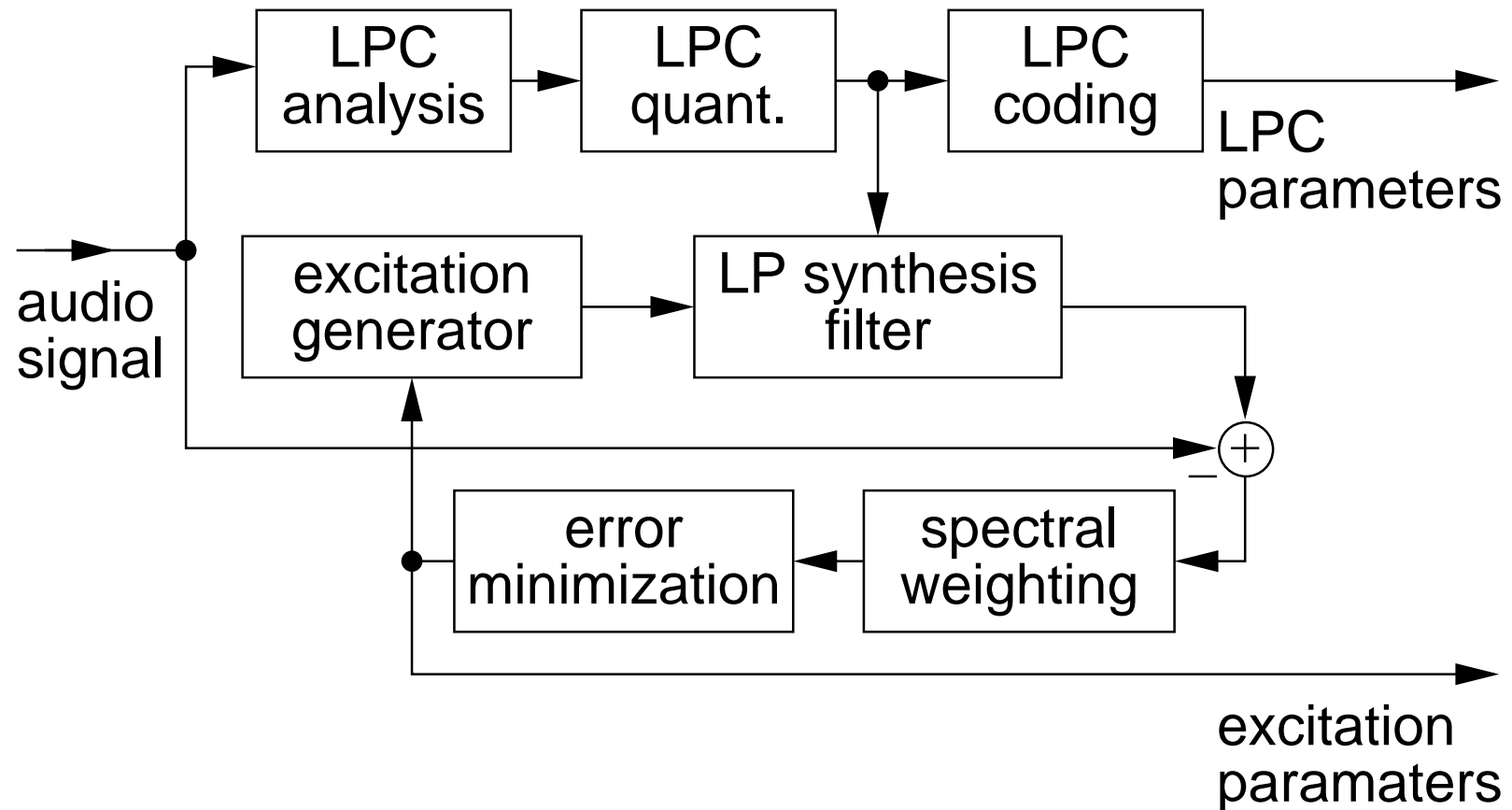
Basic structure of a CELP encoder
(analysis-by-synthesis paradigm)

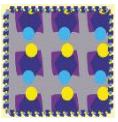Heiko Purnhagen (purnhage@tnt.uni-hannover.de)
Laboratorium für Informationstechnologie
Abteilung Systemtechnik

Universität Hannover

Reasons for characteristic artifacts

- specialized source model for speech

  - only one harmonic component
    - → bad modelling of polyphonic or complex signals
  - coding and interpolation of LPC parameters
    - → tuned for speech

- simplified perceptual model

  - spectral weighting filter in encoder controlled by LPC analysis
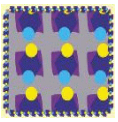   - → sub-optimal shaping of quantization noise

Heiko Purnhagen (purnhage@tnt.uni-hannover.de)
Laboratorium für Informationstechnologie
Abteilung Systemtechnik

Universität Hannover
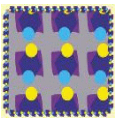
Sound Examples: Narrowband Signals

male speech	original	coded
female speech	original	coded
vocal quartet	original	coded
music A	original	coded
speech + music B	original	coded
music B	original	coded
radio	original	coded
vocal	original	coded
tonal A	original	coded
tonal B	original	coded
transients	original	coded

sampling frequency: 8 kHz, bit rate: 6 kbit/s

Heiko Purnhagen (purnhage@tnt.uni-hannover.de)
Laboratorium für Informationstechnologie
Abteilung Systemtechnik

lfi

Universität Hannover

male speech	original	coded
female speech	original	coded
vocal quartet	original	coded
music A	original	coded
speech + music B	original	coded
music B	original	coded
radio	original	coded
tonal A	original	coded
transients	original	coded

sampling frequency: 16 kHz, bit rate: 16 kbit/s

Heiko Purnhagen (purnhage@tnt.uni-hannover.de)
Laboratorium für Informationstechnologie
Abteilung Systemtechnik

Universität Hannover