

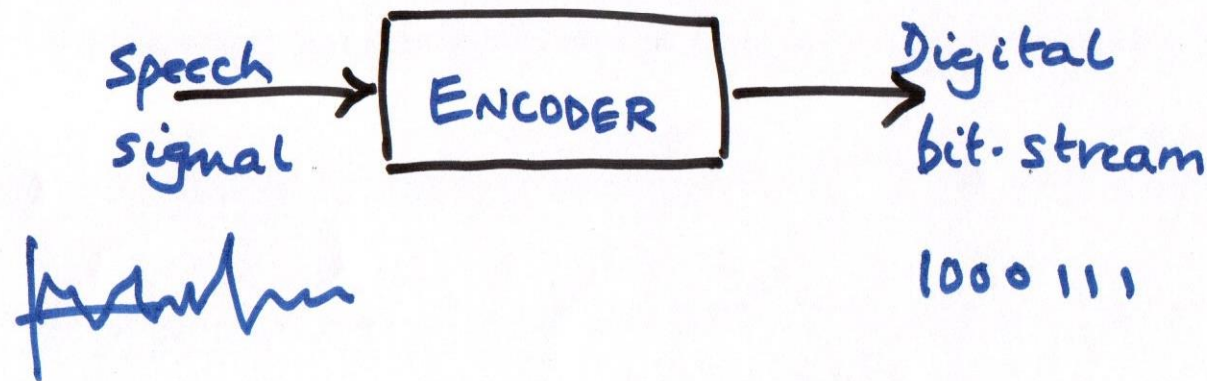
# Speech Compression / coding

Finding an efficient digital representation of speech for storage / transmission.



CDEEP  
IIT Bombay

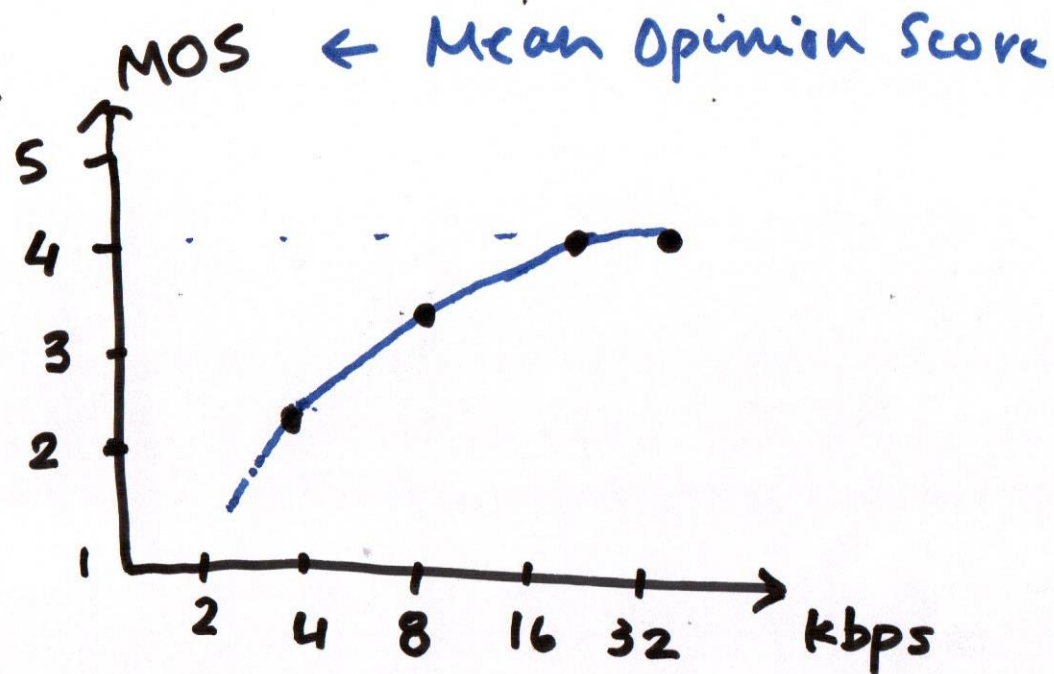
EE 679 L 24 / Slide 6



Engg trade-offs

- Speech quality
- Bit-rate
- Comput. complexity
- Delay

# Speech quality vs bit-rate



CDEEP  
IIT Bombay

EE 679 L 24 / Slide 7

Obj. measures

SNR

log S.D.

Percep. measures  
(PESQ)

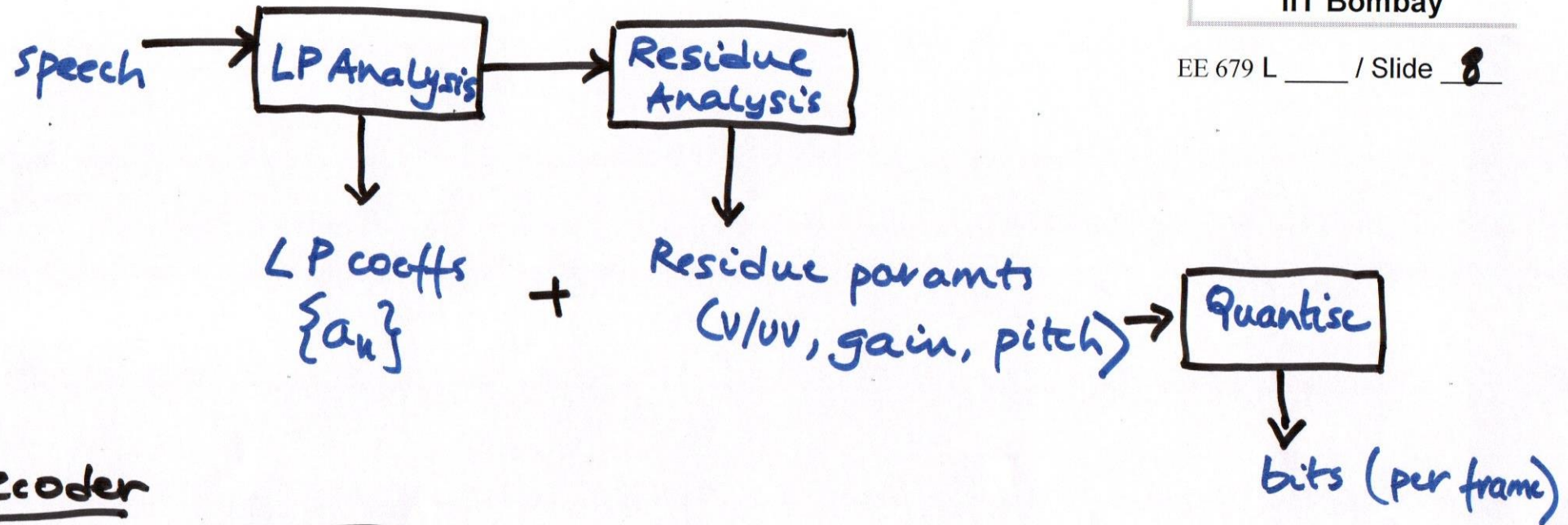
Waveform Coders →  $f_s = 8\text{kHz} \times 16\text{b/samples}$   
= 128 kbps

Model-based Coders



# LPC Vocoder (frame-based)

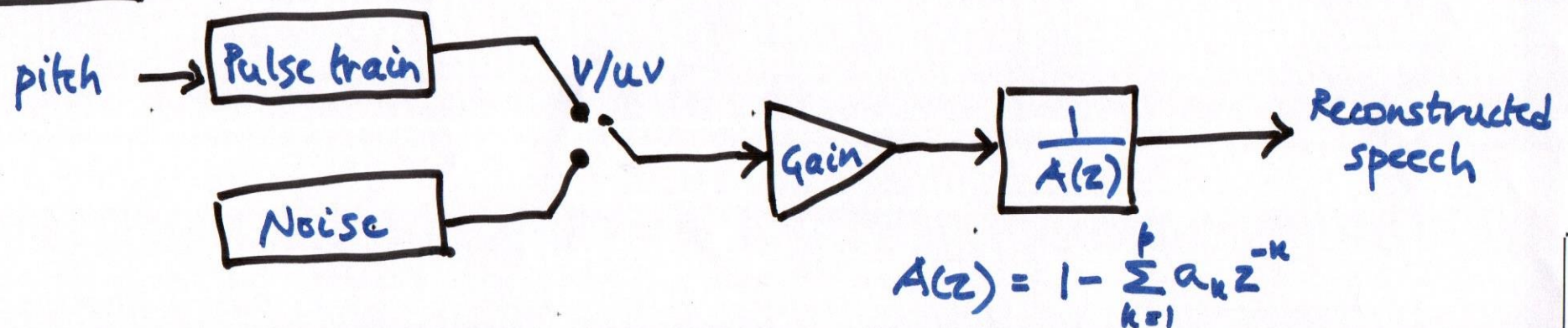
## Encoder



CDEEP  
IIT Bombay

EE 679 L \_\_\_\_ / Slide 8

## Decoder



# Parameter quantization in LPC<sup>vo</sup> coder



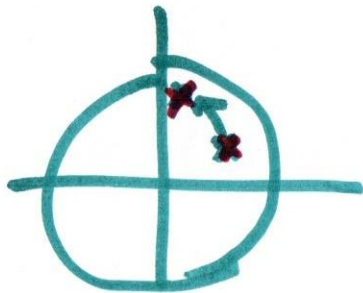
CDEEP  
IIT Bombay

EE 679 L \_\_\_\_ / Slide 9

1. LP coeffs :  $\{a_n\}$

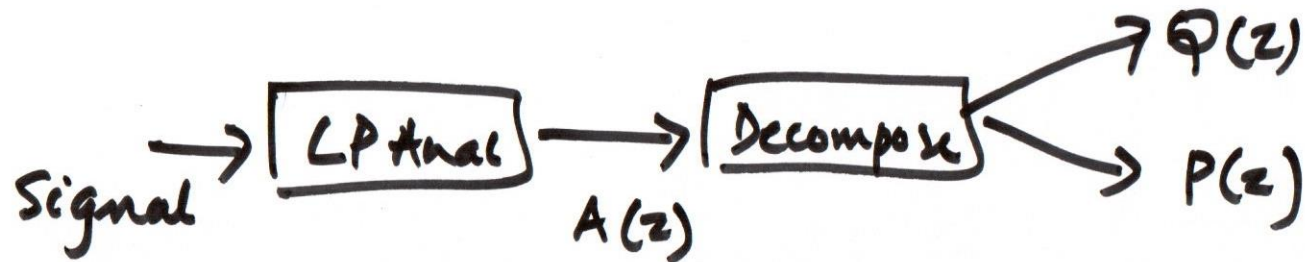
Considerations:

- i) how sensitive are  $\{a_n\}$  to quantiz<sup>n</sup>.
- ii) how amenable are  $\{a_n\}$  to simple interpolation for good synthesis quality.



$a_n \rightarrow \hat{a}_n$   
↓  
leads to shifts  
in pole loc<sup>n</sup>s

## Using LSFs to represent LPCs



$$A(z) = 1 - \sum_{i=1}^p a_i z^{-i}$$

is decomposed:

$$P(z) = A(z) + z^{-(p+1)} A(z^{-1})$$

$$Q(z) = A(z) - z^{-(p+1)} A(z^{-1})$$

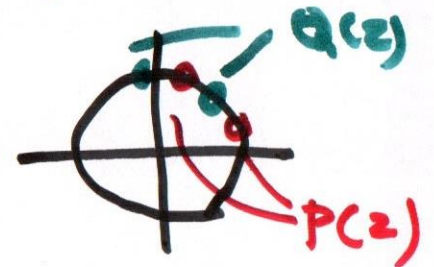
$$\Rightarrow A(z) = \frac{P(z) + Q(z)}{2}$$

LSFs are the zeros of  $P(z)$  &  $Q(z)$



CDEEP  
IIT Bombay

EE 679 L \_\_\_\_ / Slide 10

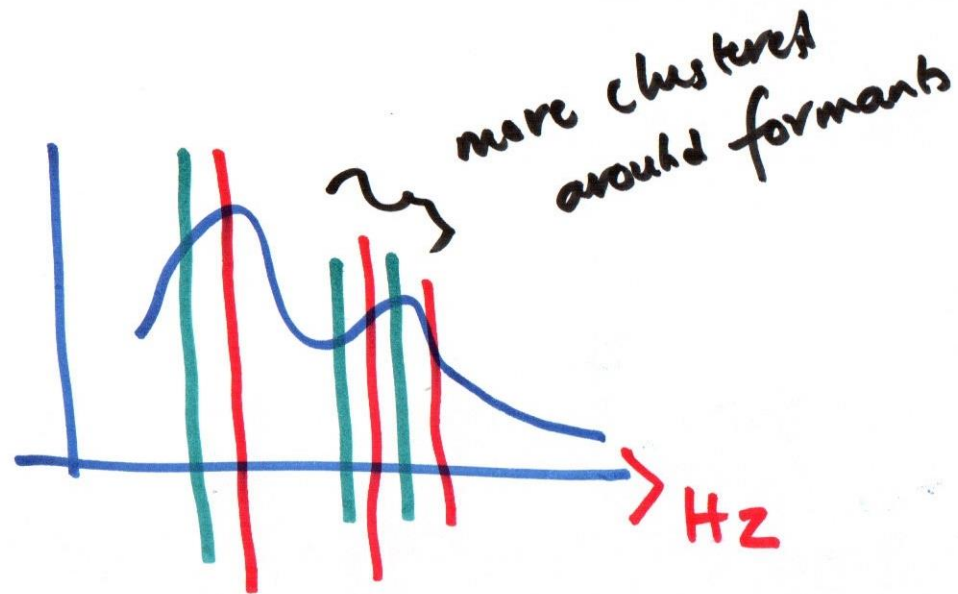






CDEEP  
IIT Bombay

EE 679 L \_\_\_\_ / Slide 11



LSFs  $\leftarrow$  superior quantiz<sup>n</sup>  
& interpol<sup>n</sup> properties.

Quantiz<sup>n</sup> of Pitch ( $F_0$ )

Pitch JND  $\approx 1\%$ .

Consider  $\log(F_0)$  for unif quantiz<sup>n</sup>  
in the range 70 Hz to 350 Hz

$\Rightarrow$  8-bits / frame

Quant of  $G_{dc}$

Use dB  
with 6 bits /  
frame