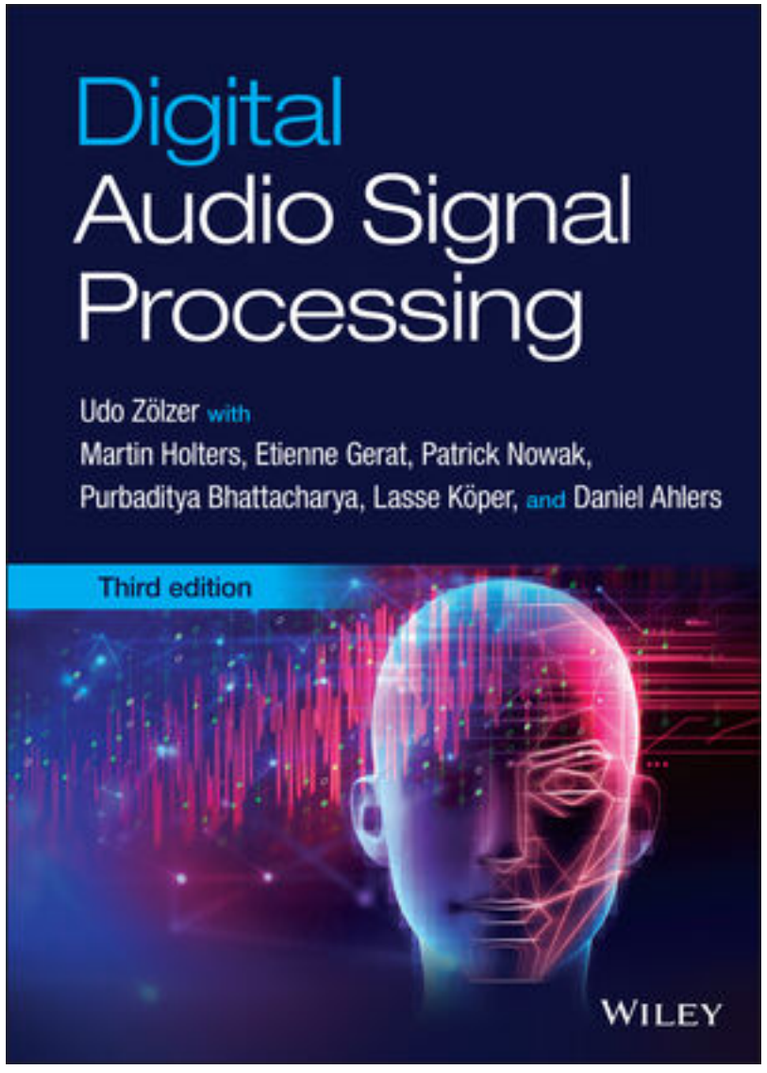


Lecture by Udo Zölzer

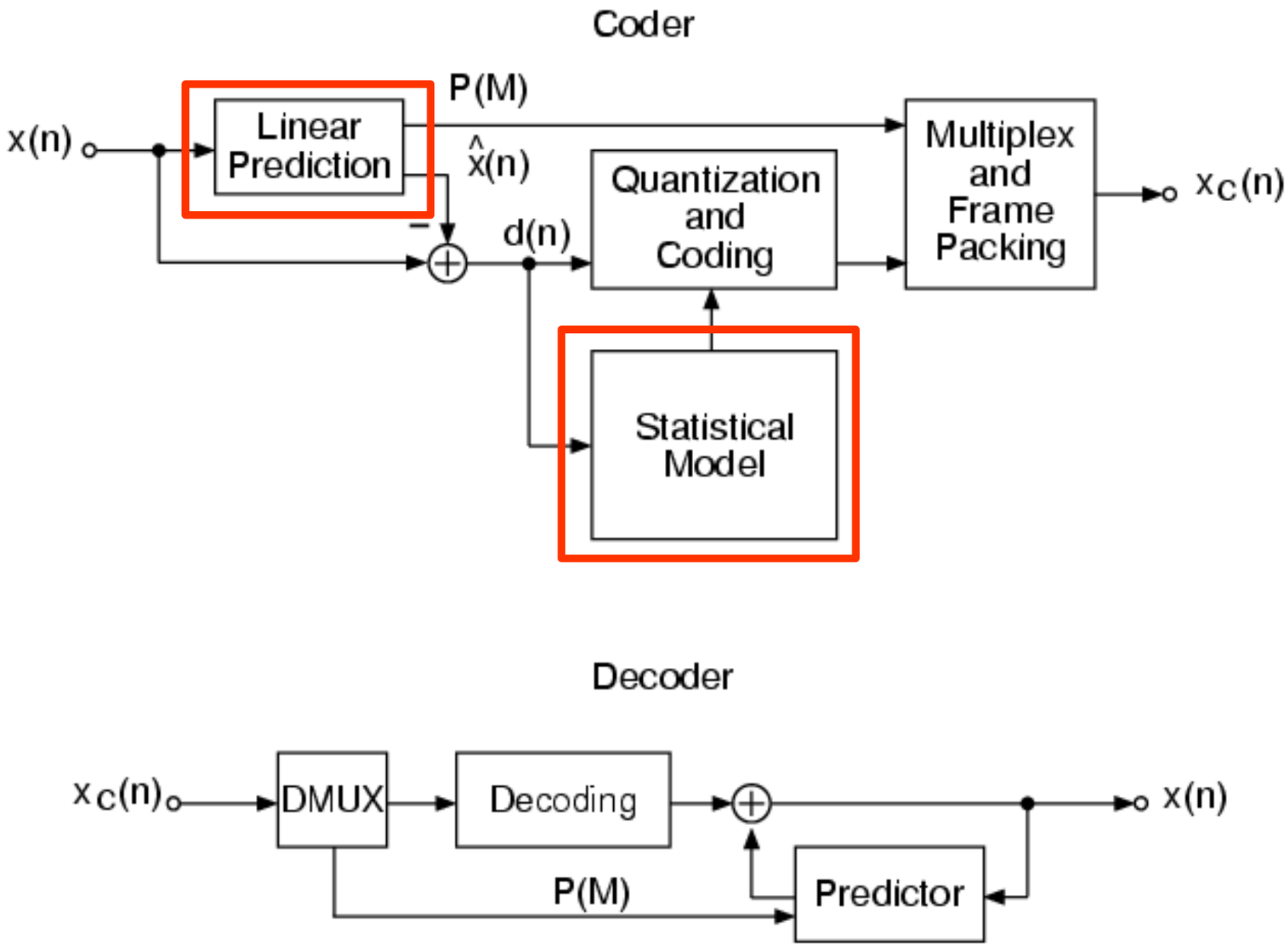
- Introduction
- Quantization
- Sampling Rate Conversion
- AD/DA Conversion
- Equalizers
- Room Simulation
- Dynamic Range Control
- **Audio Coding**
- Nonlinear Processing
- Machine Learning for Audio

AUDIO CODING

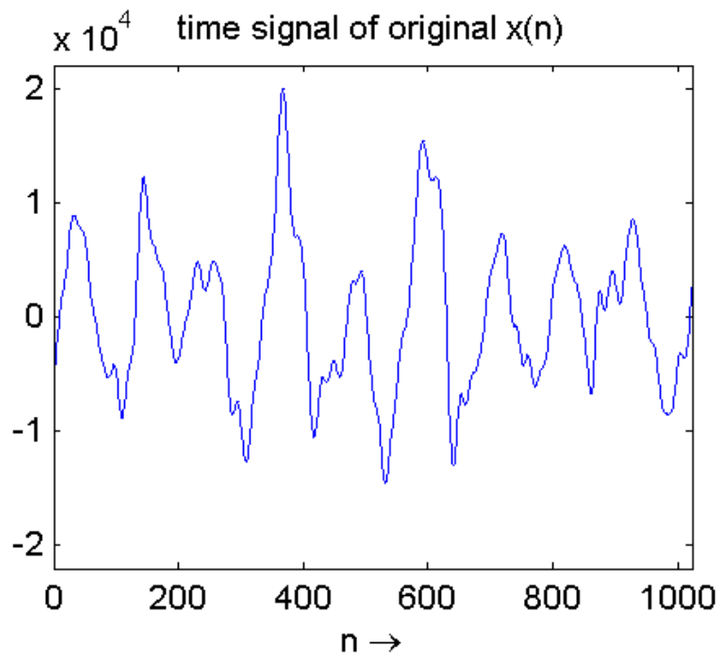
OUTLINE

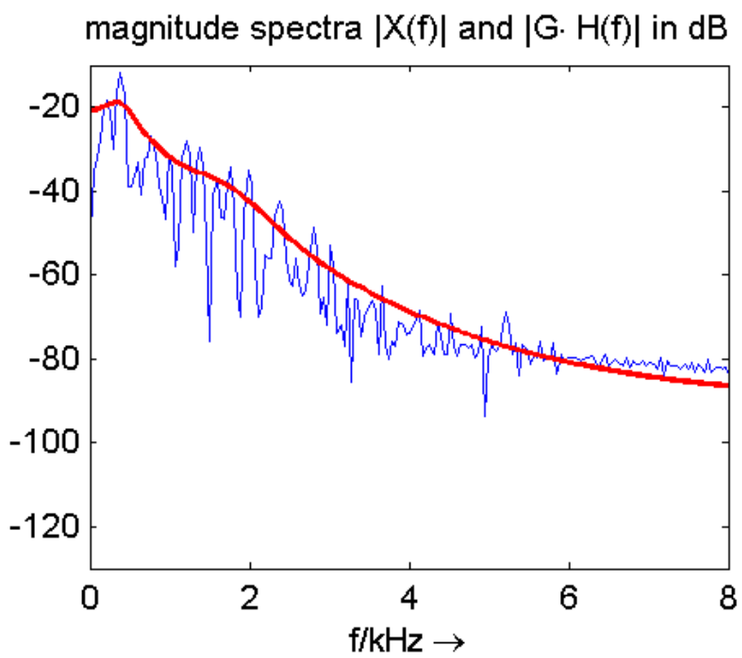
- Lossless Data Compression
- Lossy Data Compression
- Psychoacoustics
  - Critical Bands and Absolute Threshold
  - Masking
- ISO-MPEG1 Audio Coding

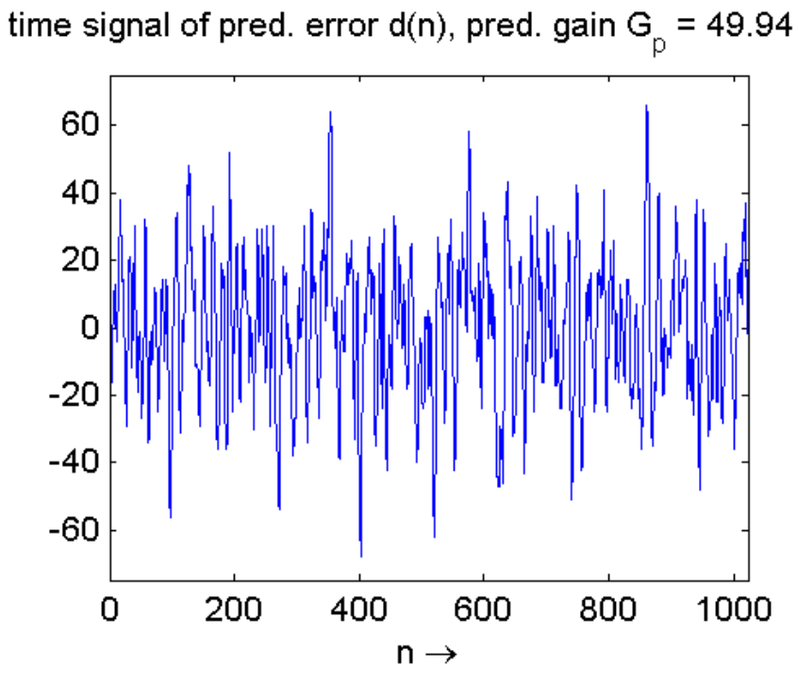
LOSSLESS DATA COMPRESSION

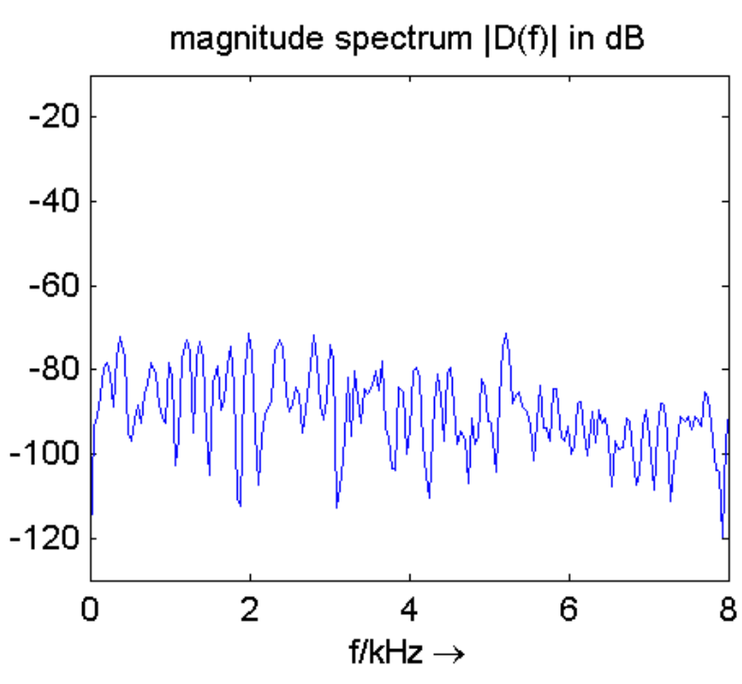


LINEAR PREDICTION EXAMPLE OF ORDER M=8

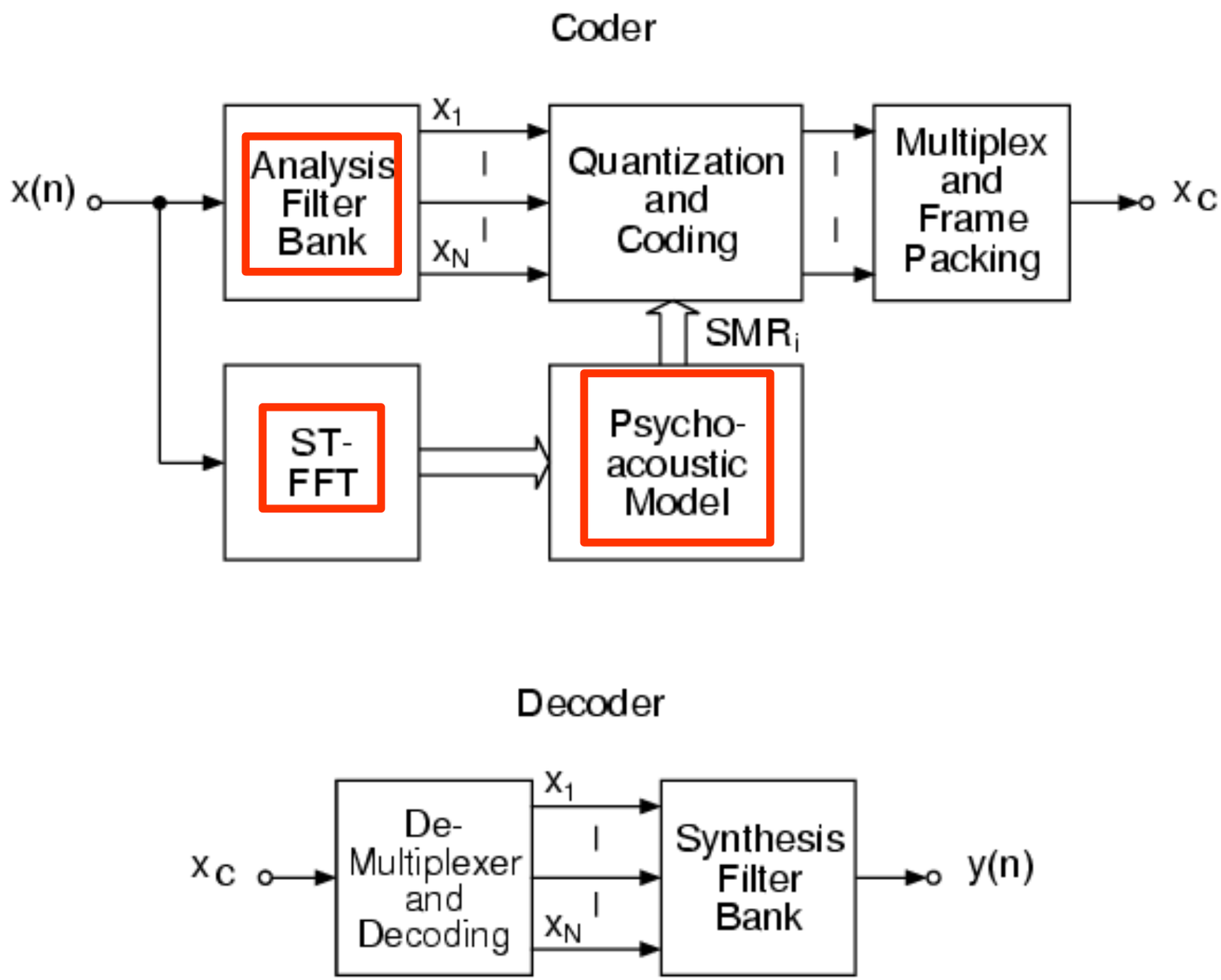






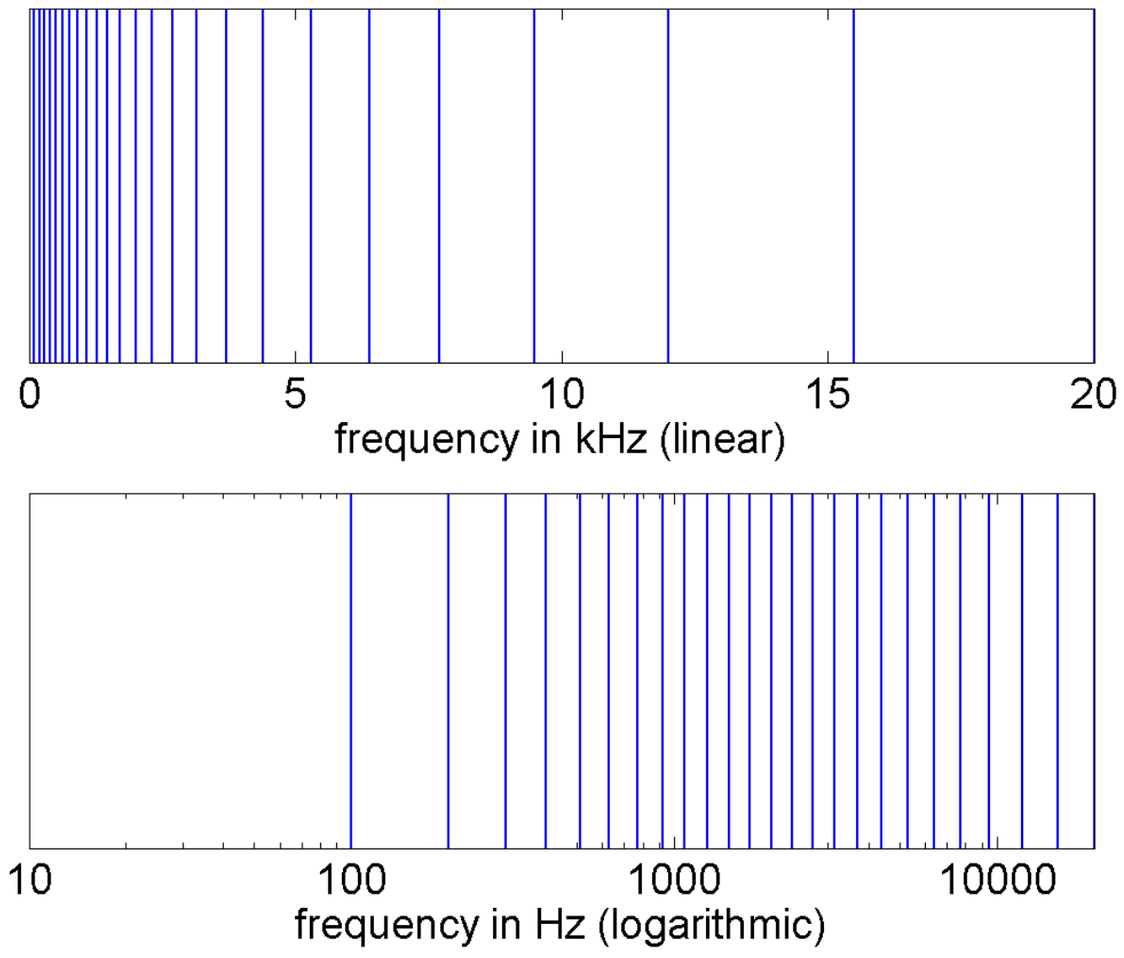


LOSSY DATA COMPRESSION



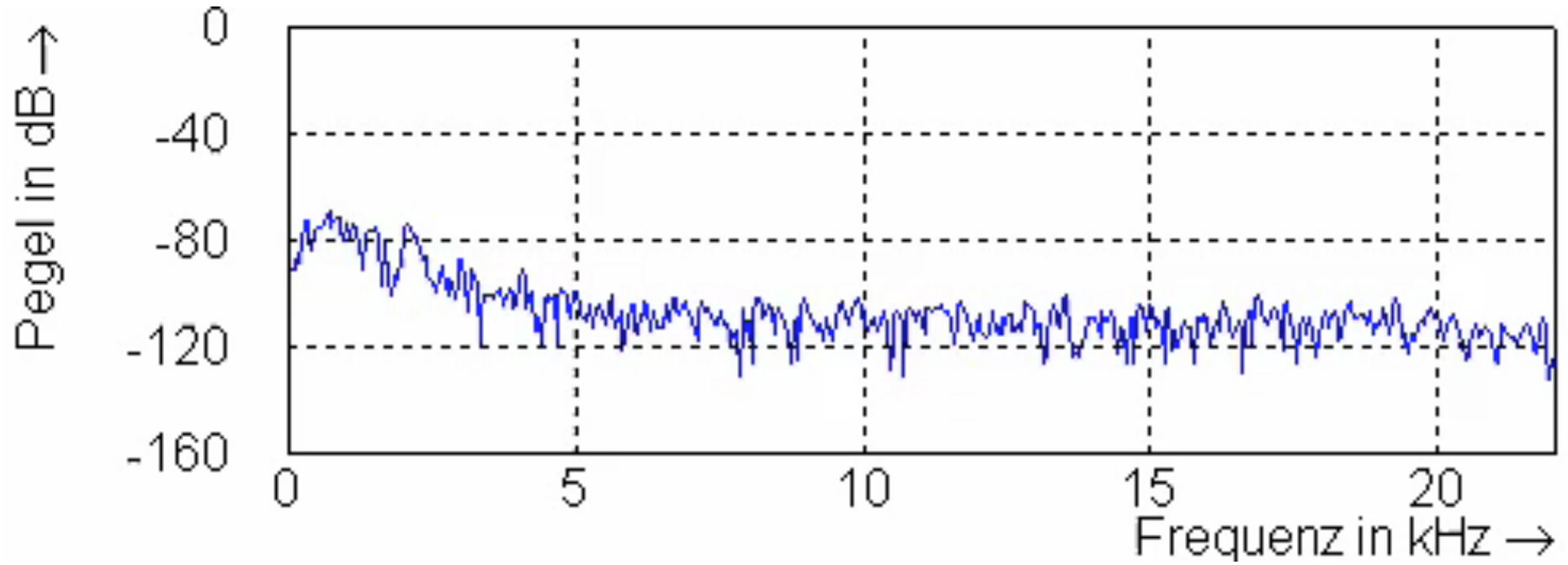
PSYCHOACOUSTICS – CRITICAL OR BARK BANDS

z/Bark	f_l/Hz	f_u/Hz	Δf_G/Hz	f_c/Hz
0	0	100	100	50
1	100	200	100	150
2	200	300	100	250
3	300	400	100	350
4	400	510	110	450
5	510	630	120	570
6	630	770	140	700
7	770	920	150	840
8	920	1080	160	1000
9	1080	1270	190	1170
10	1270	1480	210	1370
11	1480	1720	240	1600
12	1720	2000	280	1850
13	2000	2320	320	2150
14	2320	2700	380	2500
15	2700	3150	450	2900
16	3150	3700	550	3400
17	3700	4400	700	4000
18	4400	5300	900	4800
19	5300	6400	1100	5800
20	6400	7700	1300	7000
21	7700	9500	1800	8500
22	9500	1200	2500	10500
23	12000	15500	3500	13500
24	15500			

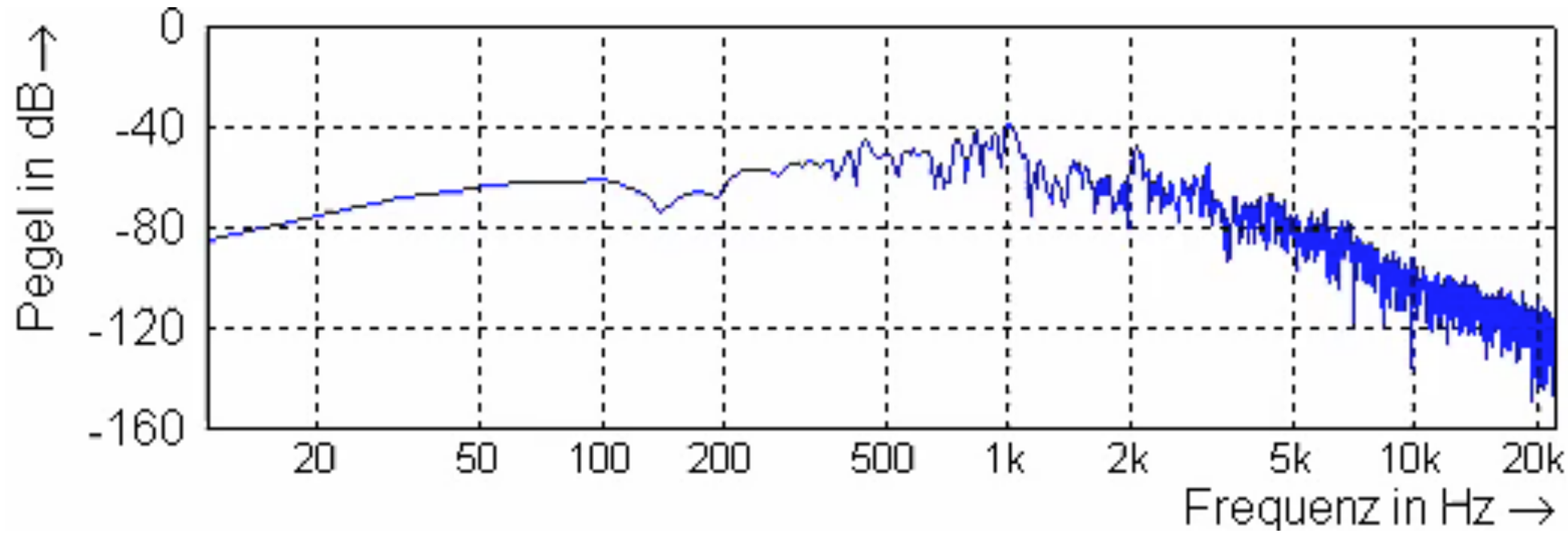


FREQUENCY ANALYSIS DEMO

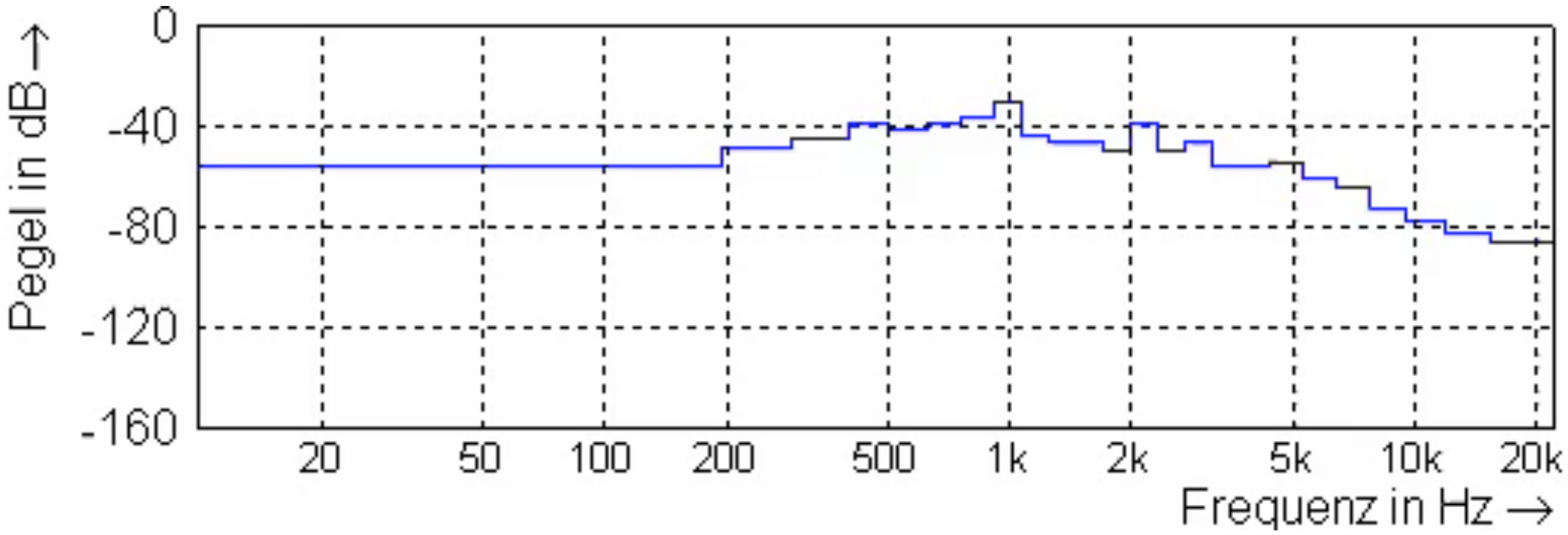
Linear analysis



Logarithmic analysis

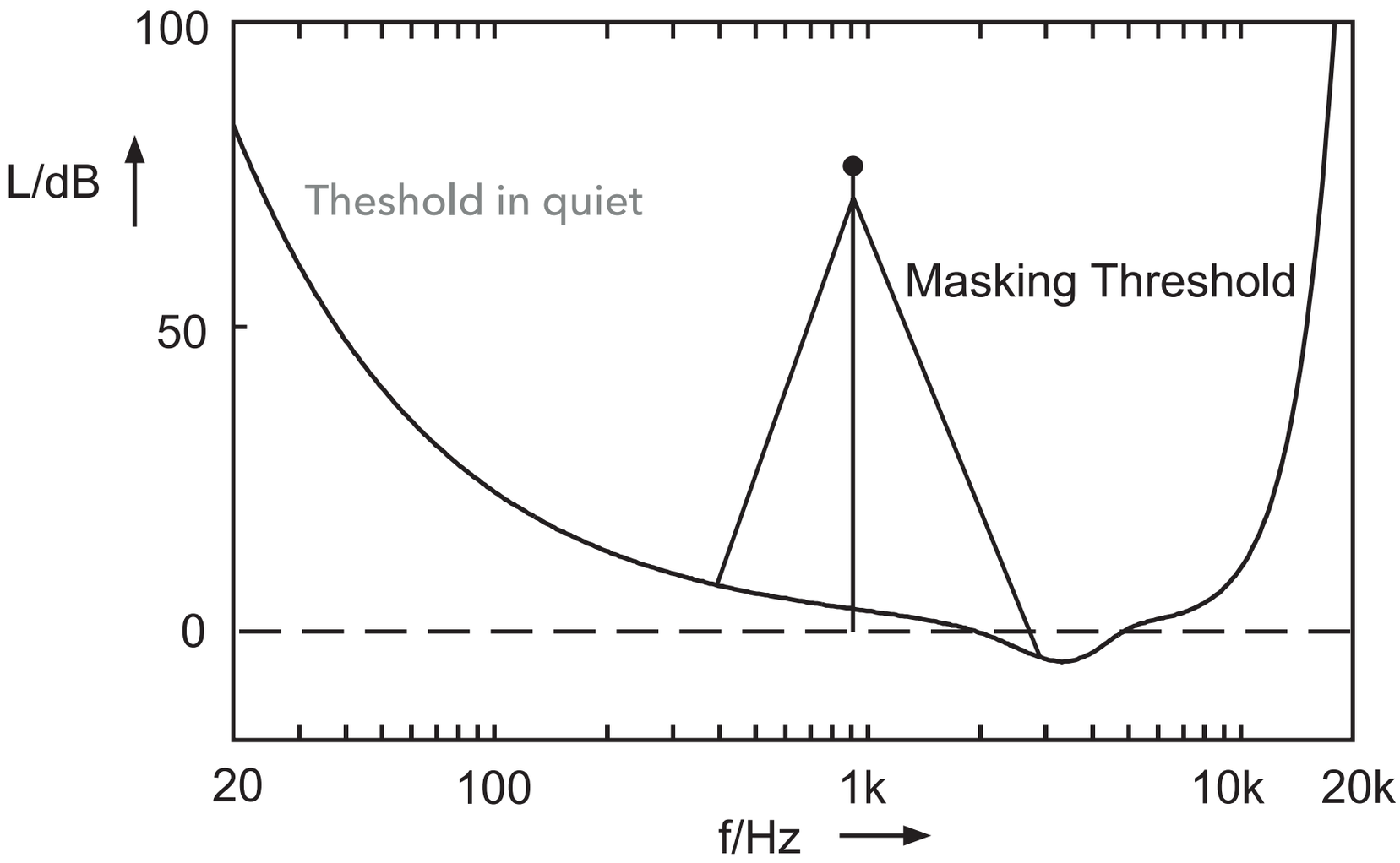


FREQUENCY BAND ANALYSIS DEMO

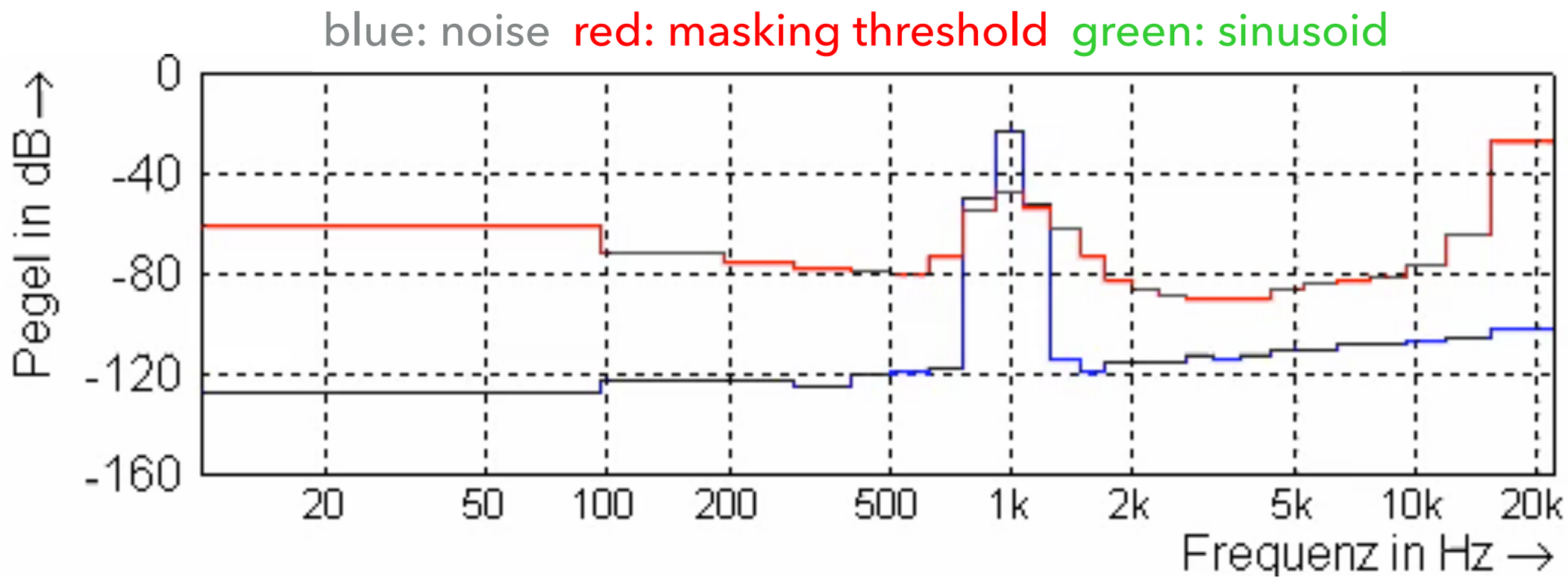


Frequency analysis in critical bands by accumulating the signal energy inside each critical band (also called “bark bands“).

ABSOLUTE THRESHOLD IN QUITE AND MASKING THRESHOLD

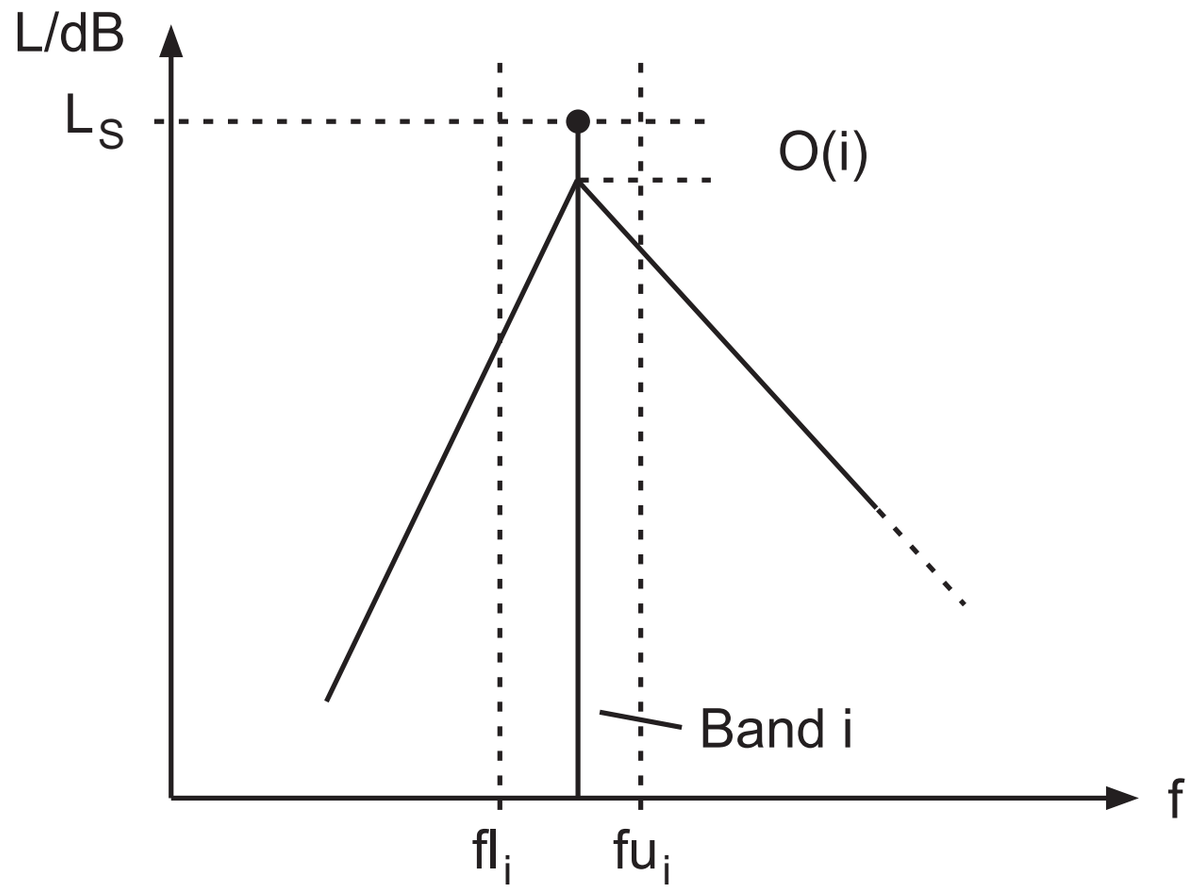


MASKING DEMO



- Calculation of masking threshold from bandlimited noise.
- Addition of sinusoid of increasing amplitude.
- Sinusoid has higher frequency than the noise signal.
- Sine amplitudes: -50, -40, -30, -20, -10 dB.

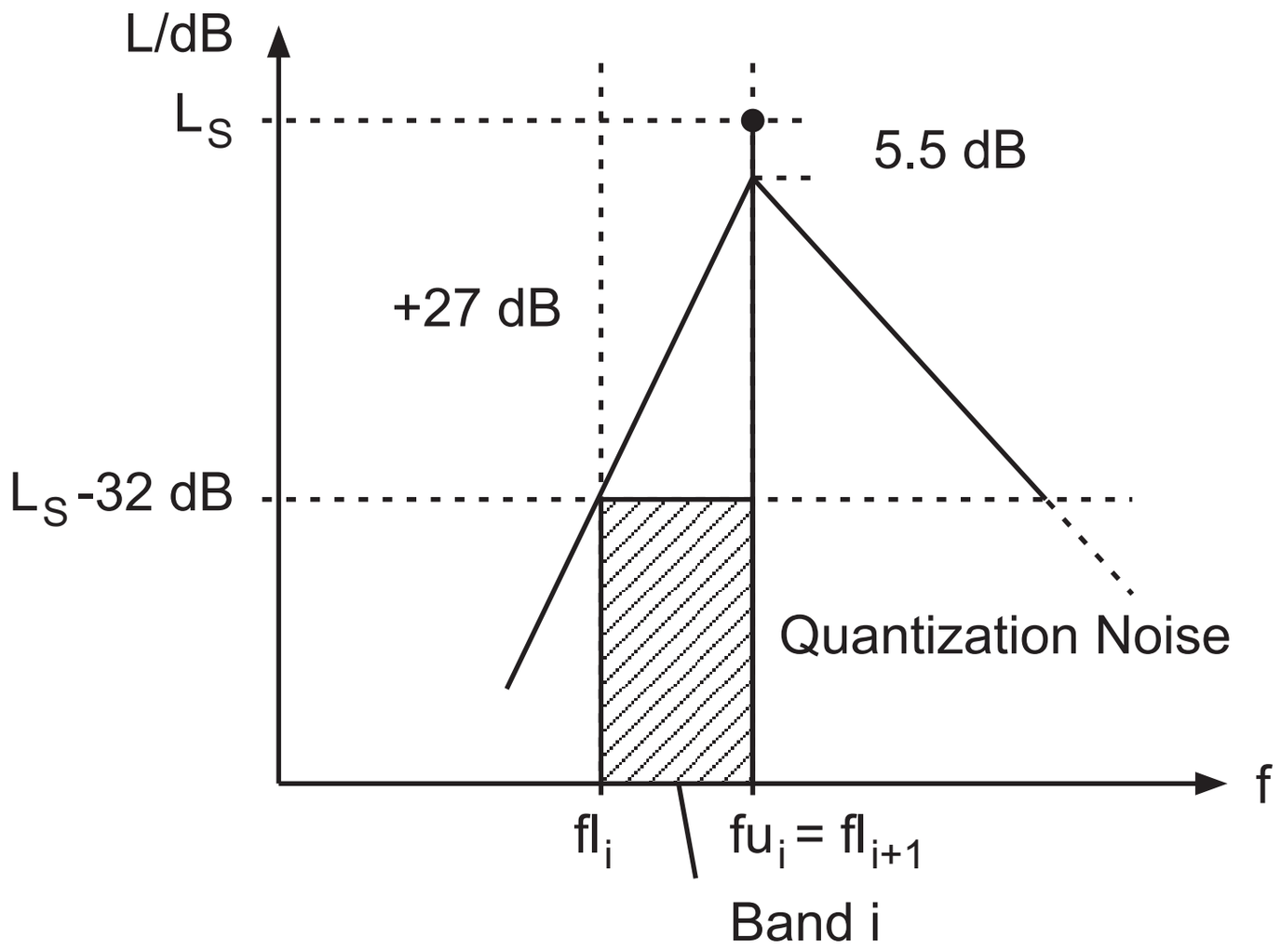
MASKING INSIDE CRITICAL BANDS



Tonality Offset $O(i)=\alpha(14{,}5+i)+(1-\alpha)5{,}5$

Tonality measure $0<\alpha<1$

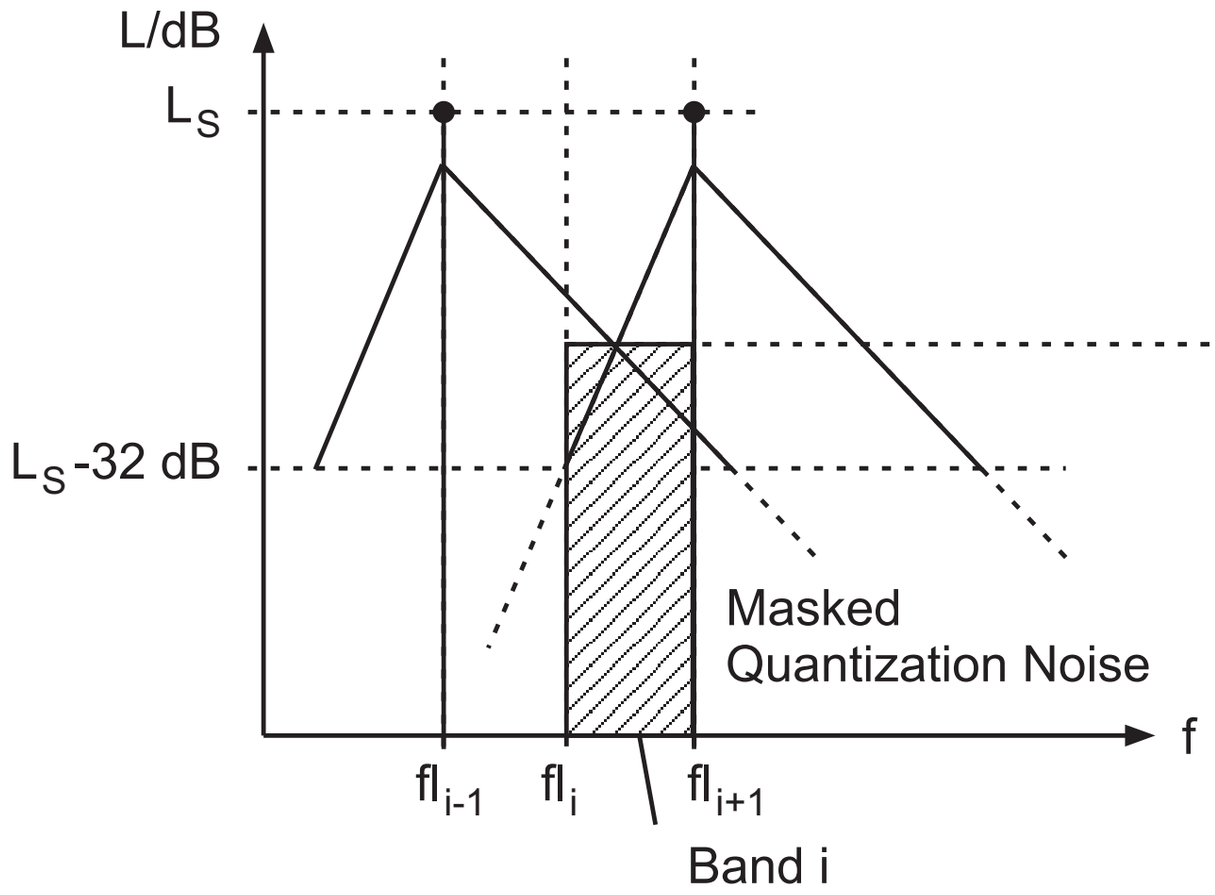
MASKING INSIDE CRITICAL BANDS

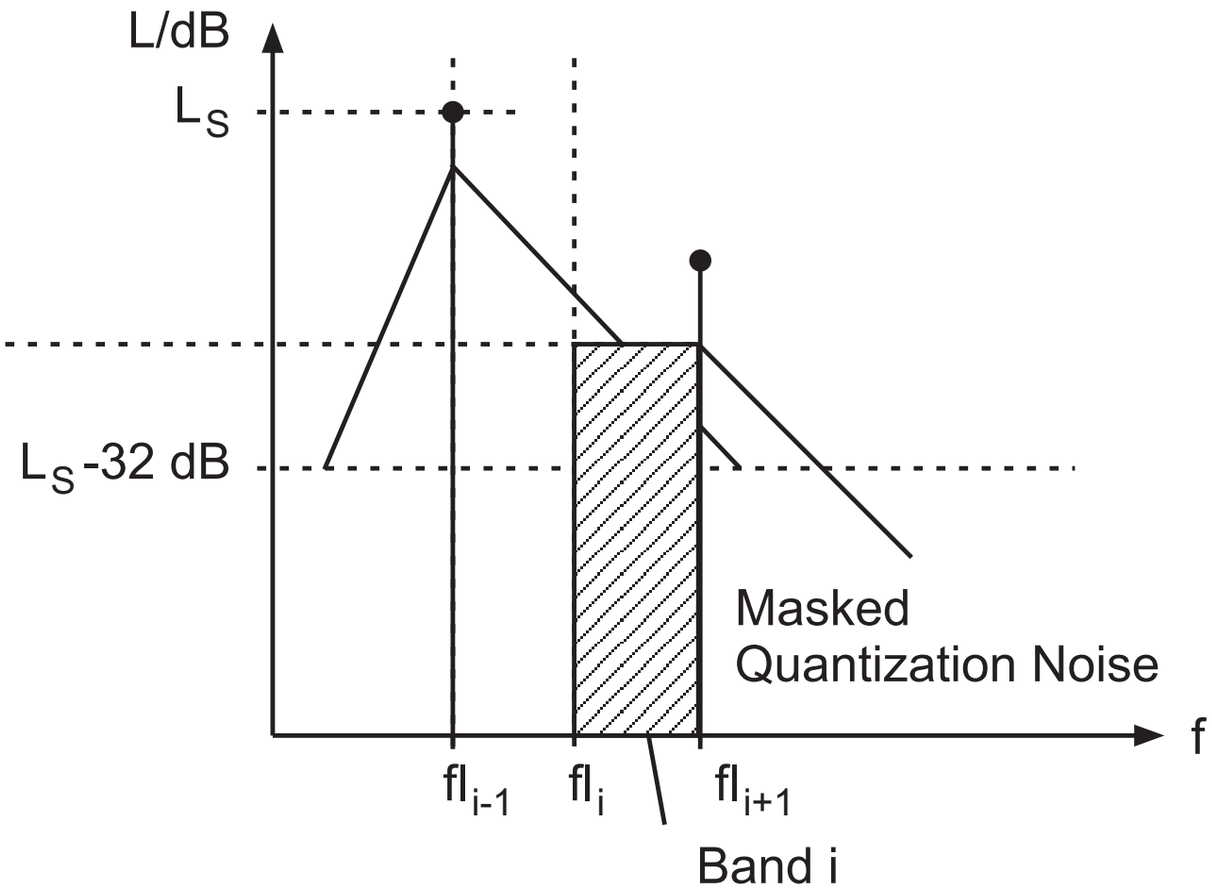


$$S_1 = 27 \quad \text{dB/Bark}$$

$$S_2 = 24 + 0.23(\frac{f_{c_i}}{\text{kHz}})^{-1} - 0.2\frac{L_S(i)}{\text{dB}} \quad \text{dB/Bark.}$$

MASKING ACROSS CRITICAL BANDS

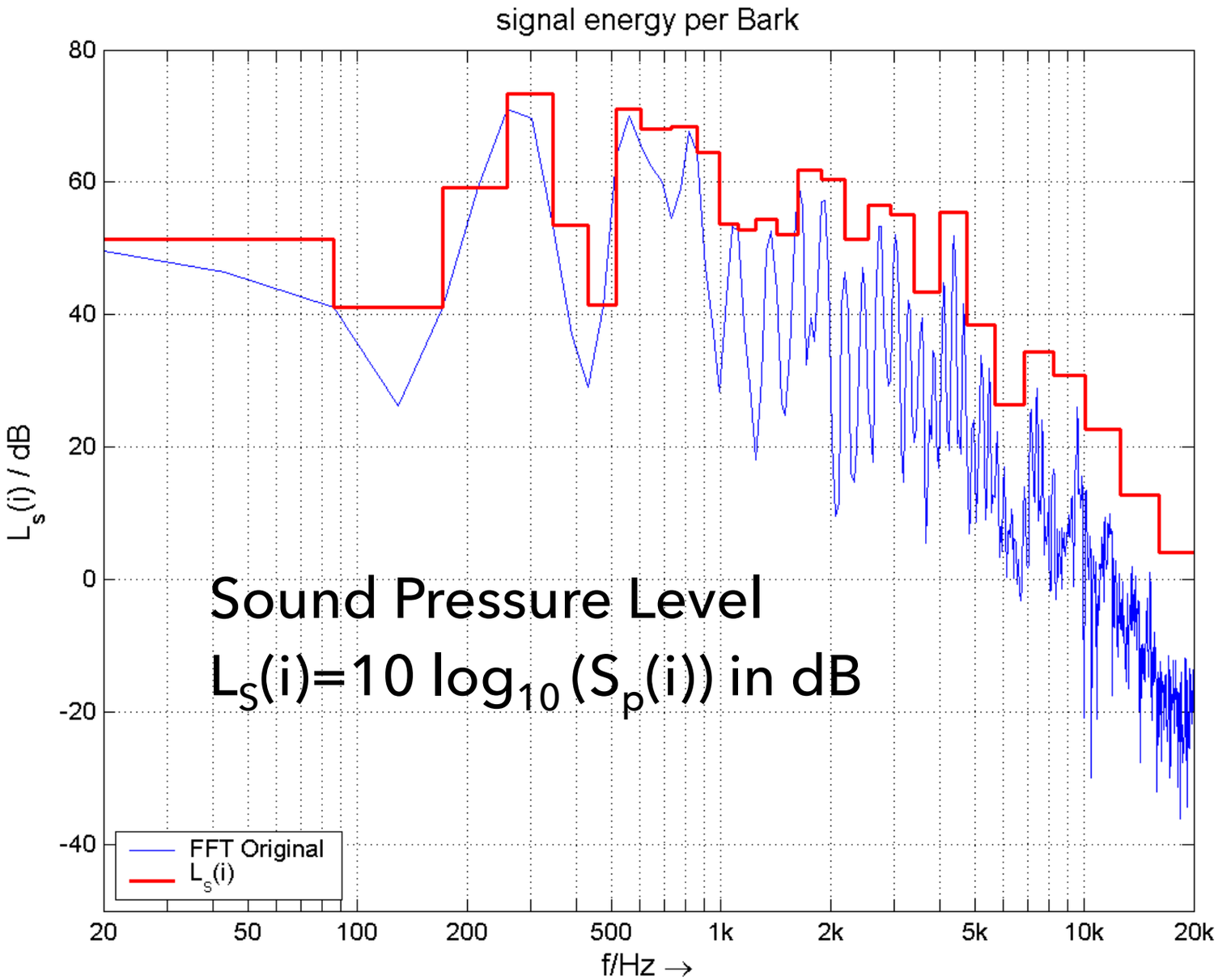




MASKING CALCULATION

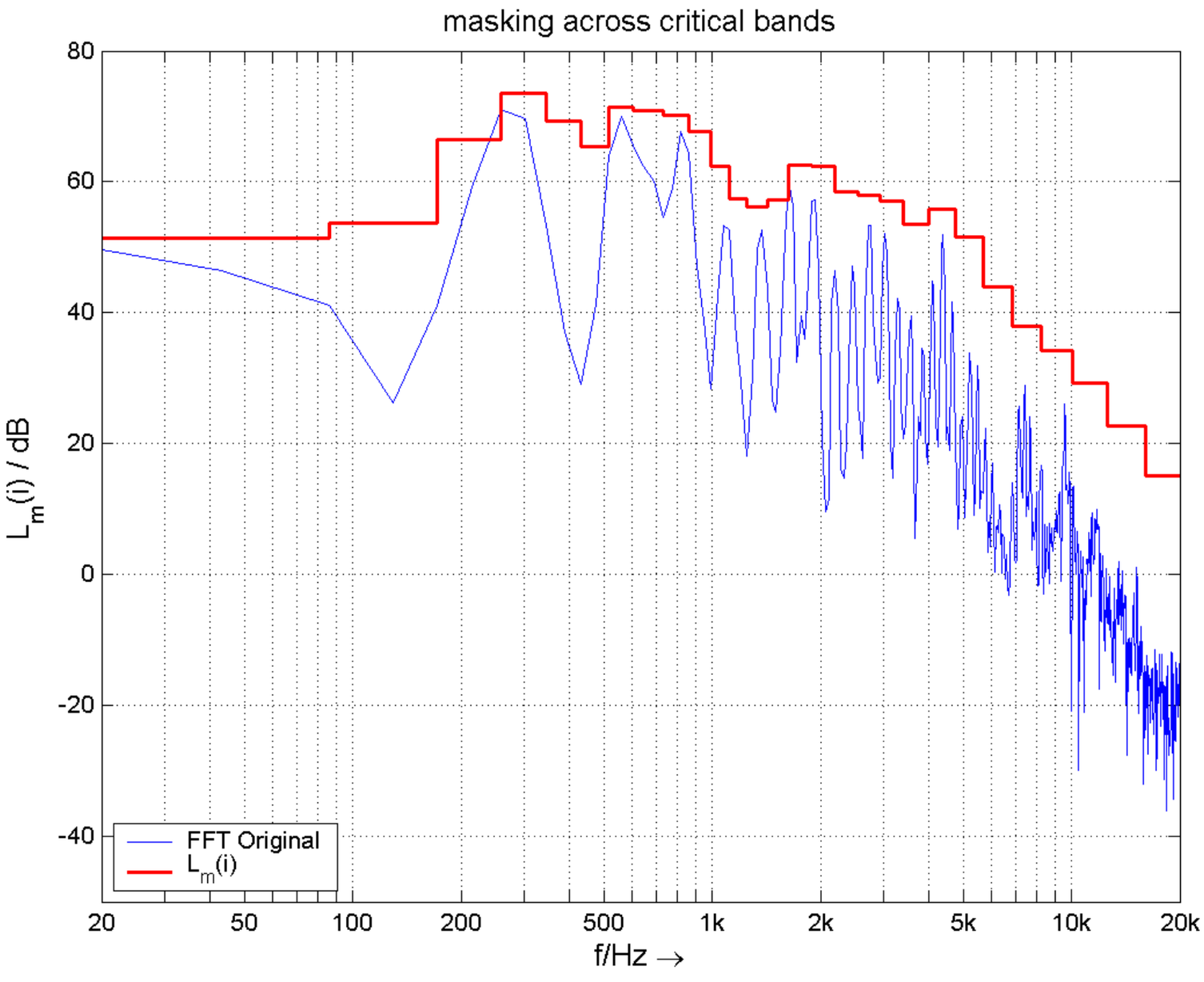
$$S_p(e^{j\Omega}) = S_p(e^{j(2\pi k/N)}) = X_R^2(e^{j(2\pi k/N)}) + X_I^2(e^{j(2\pi k/N)}),$$
$$S_p(k) = X_R^2(k) + X_I^2(k), \quad 0 \le k \le N-1,$$

Step 1: Calculation of Signal Power in Bark Band i $S_p(i) = \sum_{\Omega=\Omega_{li}}^{\Omega_{ui}} S_p(k)$



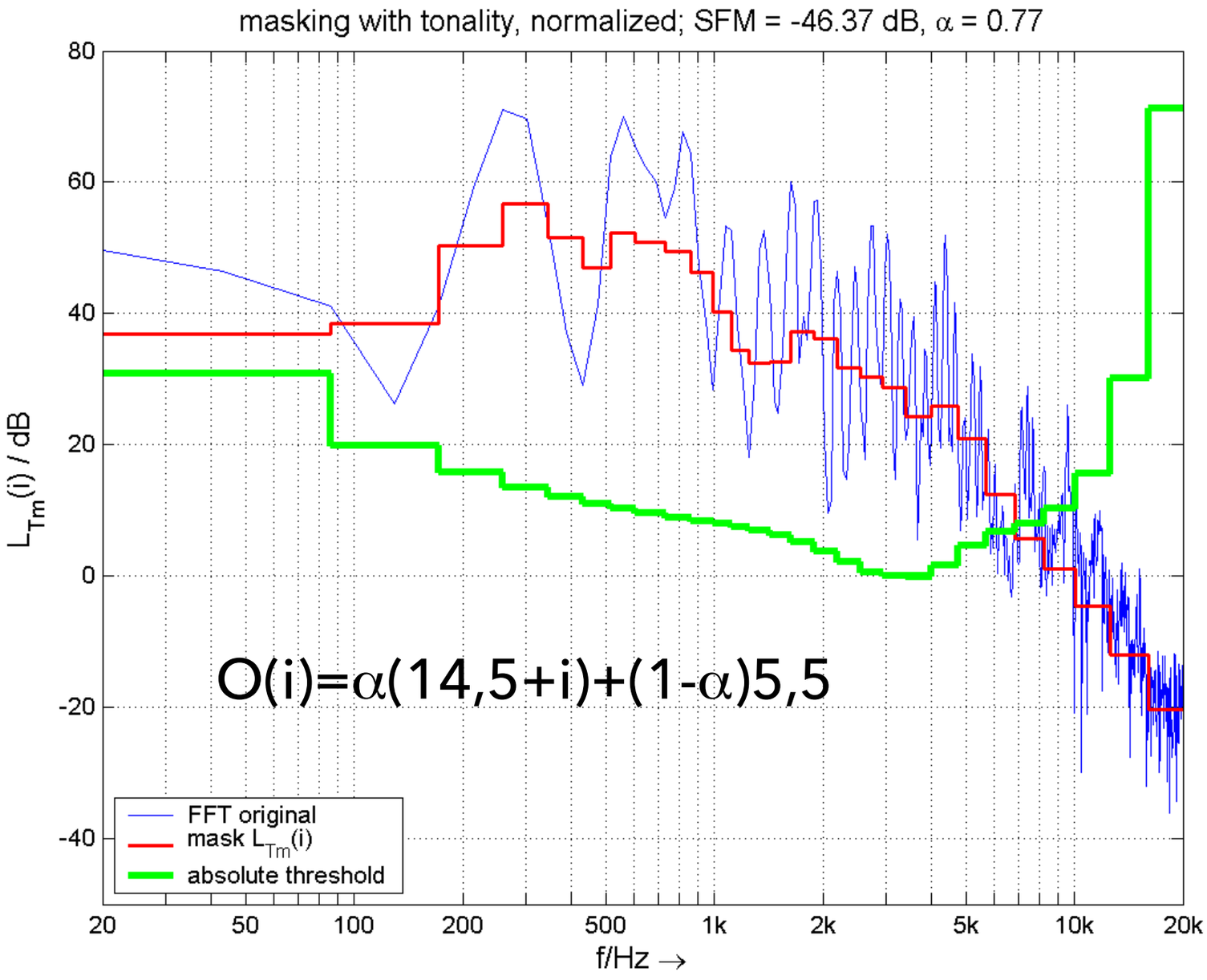
MASKING CALCULATION

Step 2: Masking across Critical Bands $L_m(i)=10\log_{10}(S_m(i))$ with $\mathbf{S}_m = \mathbf{B}\,\mathbf{S}_p$



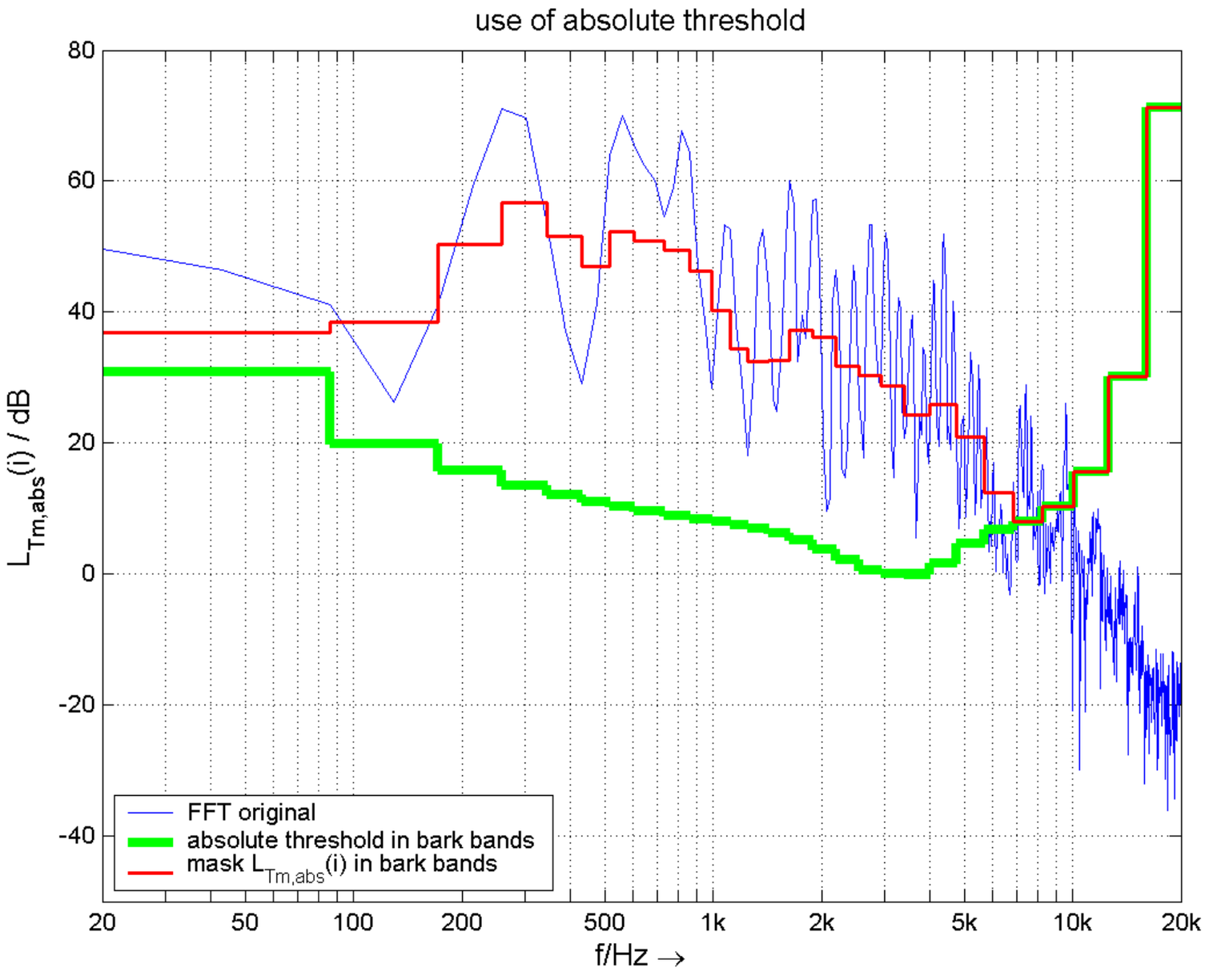
MASKING CALCULATION

Step 3: Spread Masking Threshold $L_{Tm}(i) = L_m(i) - O(i)$ in dB



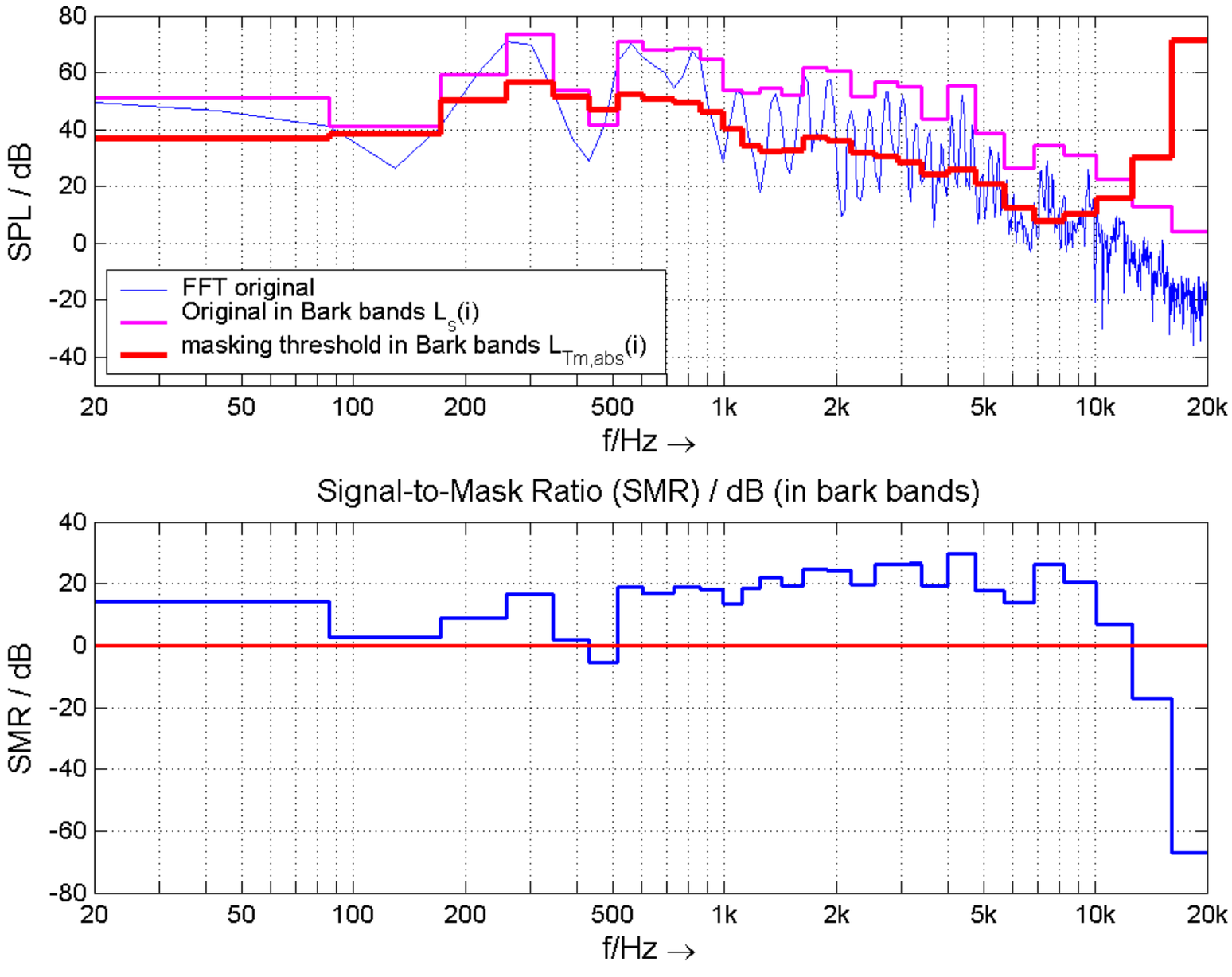
MASKING CALCULATION

Step 4: Global Masking Threshold $L_{Tm,abs}(i) = \max [L_{Tm}(i), L_{Tq}(i)]$

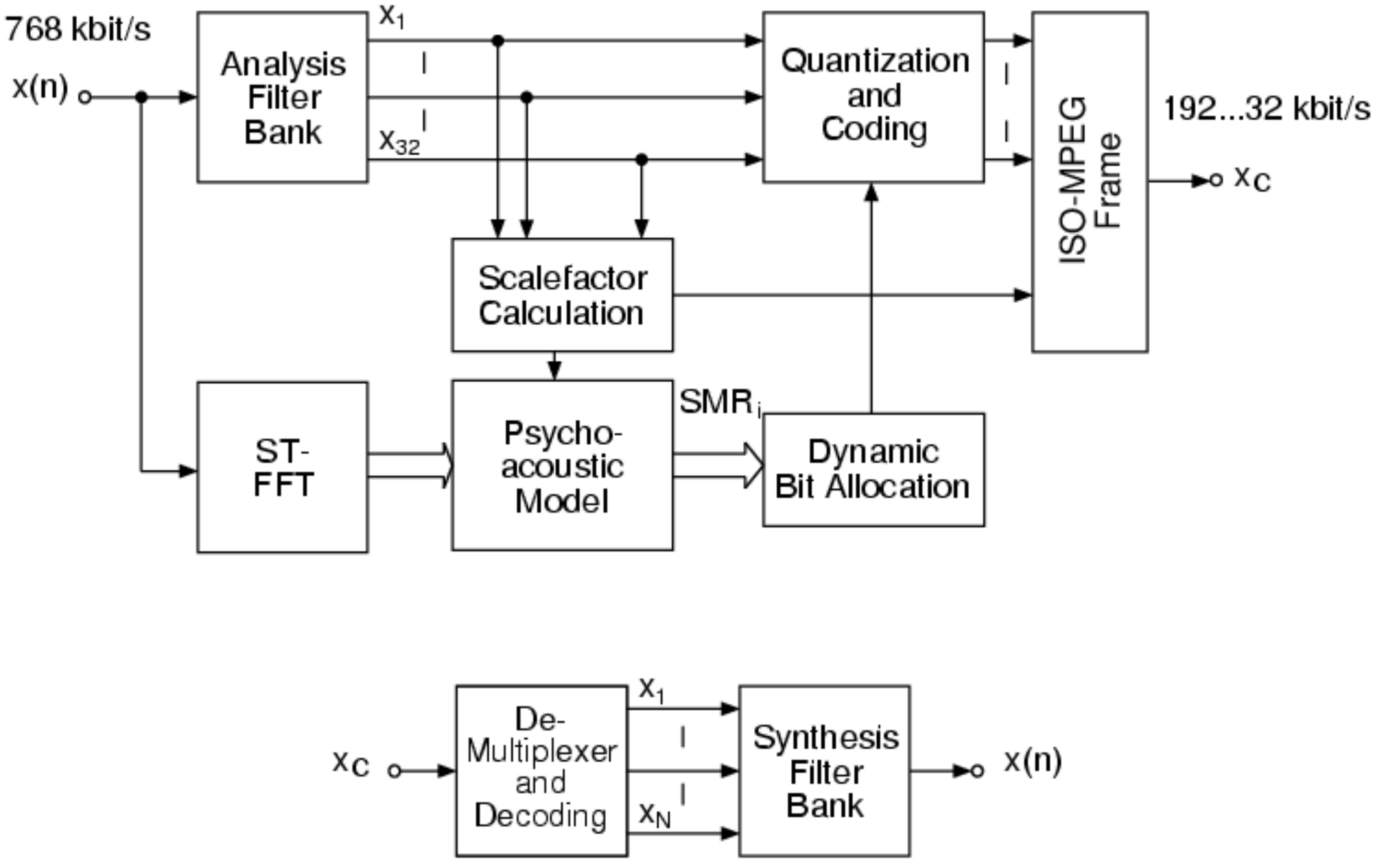


MASKING CALCULATION

Step 5: Calculation of Signal-to-Mask Ratio SMR(i) = $L_S(i) - L_{Tm,abs}(i)$



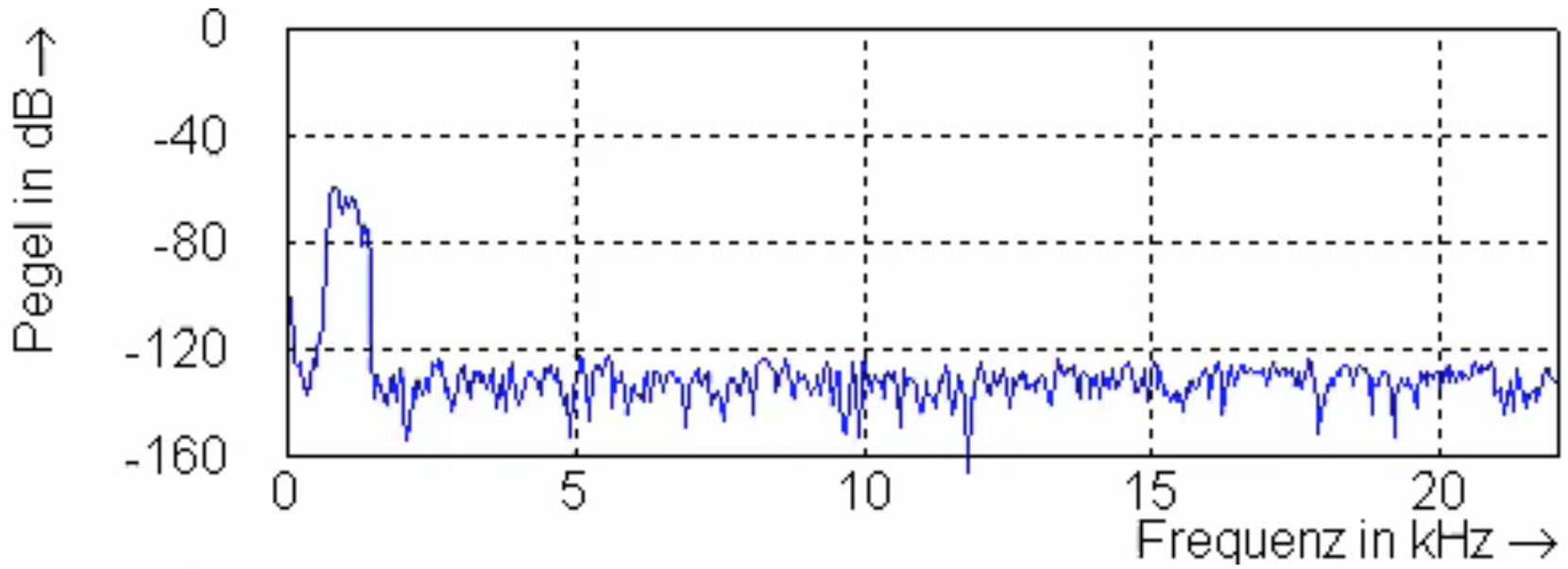
ISO-MPEG1 AUDIO CODING



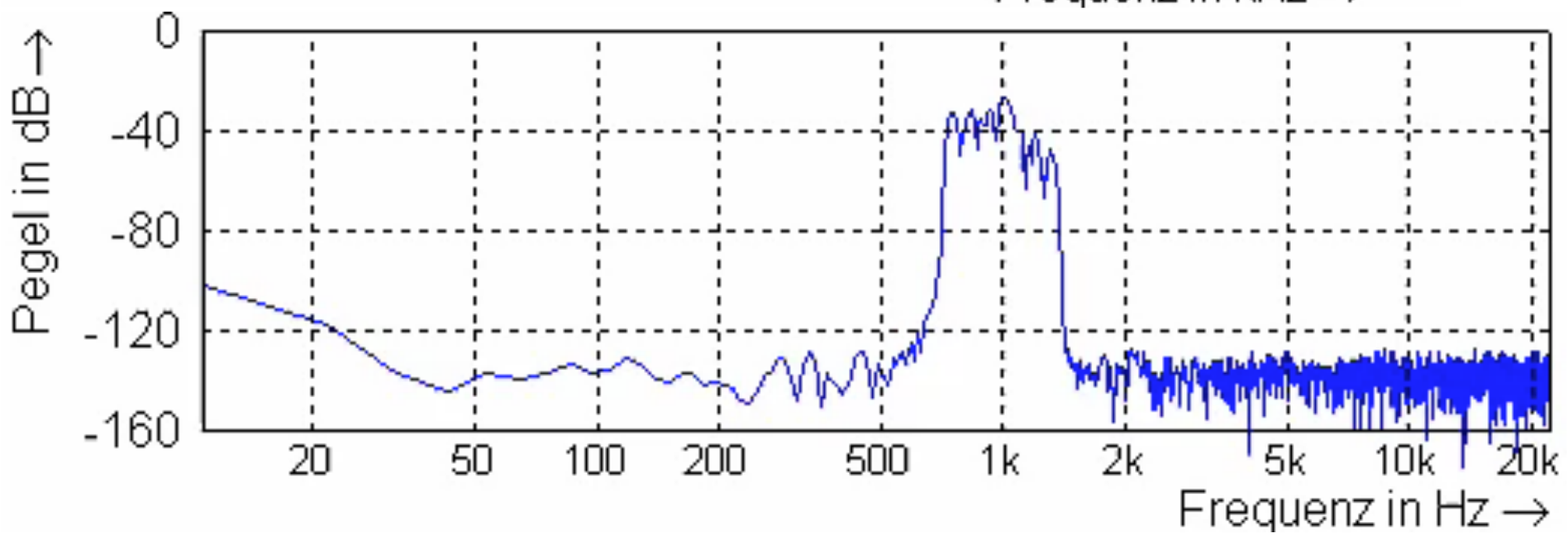
SUBBAND SIGNAL

Only second subband with frequency components from 700 to 1400 Hz, amplified by 12 dB

Linear analysis



Logarithmic analysis



ISO-MPEG1 AUDIO CODING – LAYER 3 FILTERS