POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

SP&L Homework 1

Tesi di Laurea Magistrale in
Telecommunication Engineering

Author: **Hong CHEN**

Student ID: 11024836
Advisor: Prof. Name Surname
Co-advisors: Name Surname, Name Surname
Academic Year: 2025-07

Contents

Contents **i**

1 Homework 1 **1**
1.1 Sinusoidal Signal Model 1
1.1.1 Noise Generation 2
1.1.2 Estimation and MSE Computation 2
1.2 Frequency Estimation 4
1.3 Frequency modulation 9
1.3.1 Maximum Likelihood Estimator 11

1 | Homework 1

1.1. Sinusoidal Signal Model

The given sinusoidal signal is defined as:

$$x[n] = a_0 \cos(\omega_0 n + \phi_0)$$

The signal is passed through a discrete-time filter with impulse response:

$$h[n] \quad \longleftrightarrow \quad H(z) = \frac{2}{1 + 0.9z^{-1}}$$

The filtered output is:

$$y[n] = x[n] * h[n] + w[n]$$

Using the z-transform representation, the output in the frequency domain is:

$$Y(z) = H(z)X(z) + W(z)$$

Noise $w[n]$ follows a zero-mean Gaussian distribution with a covariance matrix:

$$w \sim \mathcal{N}(0, C_{ww})$$

where the $N \times N$ covariance matrix is defined as:

$$[C_{ww}]_{i,j} = \sigma_w^2 \rho^{|i-j|}$$

Interpretation of Each Matrix Element:
Each element of the matrix $[C_{ww}]_{i,j}$ represents the following covariance:

$$\mathrm{Cov}(w_i, w_j) = \mathbb{E}[(w_i - \mu_i)(w_j - \mu_j)]$$

Given that the noise is zero-mean as specified in the problem (i.e., $\mu = 0$), the expression simplifies to:

$$[\mathbf{C}_{ww}]_{i,j} = \mathbb{E}[w_i \cdot w_j] = \sigma_w^2 \rho^{|i-j|}$$

- When $i = j$: The matrix element lies on the main diagonal and represents the variance at time index i, which quantifies the strength (or power) of the noise at that particular time point;
- In this case, the value is σ_w^2 When $i = j$: The element describes the degree of correlation between the noise at time indices i and j.

1.1.1. Noise Generation

Generation of Correlated Noise Samples:
For each set of parameters, correlated noise is generated by transforming standard white noise through eigenvalue decomposition of the covariance matrix $\mathbf{C}_{ww}$. The transformation is defined as:

$$\mathbf{w}_k = \mathbf{A} \cdot \mathbf{z}_k, \quad \mathbf{A} = \mathbf{Q}\mathbf{\Lambda}^{\mathbf{1/2}}$$

where:

- $\mathbf{z}_k$:is a realization of standard white noise;
- $\mathbf{Q}$ contains the eigenvectors of $\mathbf{C}_{ww}$;
- $\mathbf{\Lambda}$is the diagonal matrix of eigenvalues of $\mathbf{C}_{ww}$;
- $\mathbf{A}$ is the transformation matrix that imparts the desired covariance structure.

This method ensures that the synthesized noise possesses the target covariance properties defined by $\mathbf{C}_{ww}$.

1.1.2. Estimation and MSE Computation

Given a set of L realizations of the noise sequence $w_k[n]_{k=1}^{L}$, the sample covariance matrix is estimated as:

$$\hat{\mathbf{C}}_{ww} = \frac{1}{L}\sum_{k=1}^{L} \mathbf{w_k}\mathbf{w_k^T}$$

The MSE between the estimated sample covariance and the theoretical covariance is defined as:

$$MSE(N) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left([\hat{C}_{ww}]_{i,j} - [C_{ww}]_{i,j}\right)^2$$

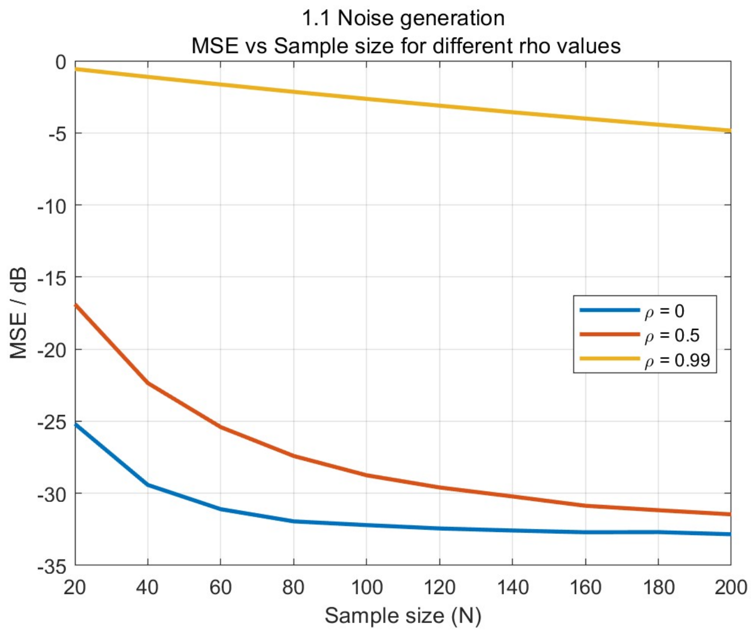

Figure 1.1: MSE vs N

We analyze the behavior of MSE as a function of N for different values of ρ (e.g., $\rho = 0, 0.5, 0.99$). Figure 1.1 illustrates the estimation error of the sample covariance matrix $hat\mathbf{C}$ relative to the theoretical covariance matrix $\mathbf{C_w w}$;

Overall Trend: For any value of ρ, the Mean Squared Error (MSE) decreases as the noise dimension N increases.

This indicates that, under a fixed number of samples L, the sample covariance estimation becomes relatively more accurate with increasing dimension. In other words, the estimation error tends to converge as N grows.

Analysis of Different Correlation Coefficients:

- $\rho = 0$ (White Noise): Represents completely uncorrelated samples. The covariance matrix is diagonal. Estimation is relatively straightforward and yields the lowest MSE.
- $\rho = 0.5$ (Moderate Correlation): Indicates moderate dependency between samples. Estimation becomes more difficult, and the MSE is significantly higher than in the white noise case.
- $\rho = 0.99$ (High Correlation): Represents strong temporal dependence in the noise sequence. The covariance matrix is nearly full and numerically close to being singular.

1.2. Frequency Estimation

The noise $w[n]$ is independent and identically distributed (i.i.d.) Gaussian noise, so the probability density function (PDF) of the observed signal $x[n]$ is:

$$p(\mathbf{x};\omega) = \prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left(-\frac{(x[n]-A\cos(\omega n+\phi))^2}{2\sigma_w^2}\right)$$

The log-likelihood function is:

$$\ln p(\mathbf{x};\omega) = -\frac{N}{2}\ln(2\pi\sigma_w^2) - \frac{1}{2\sigma_w^2}\sum_{n=0}^{N-1}(x[n]-A\cos(\omega n+\phi))^2$$

The Fisher information $I(\omega)$ is key to the Cramér-Rao bound (CRB) and is defined as the negative expectation of the second derivative of the log-likelihood function:

$$I(\omega) = -\mathbb{E}\left[\frac{\partial^2 \ln p(\mathbf{x};\omega)}{\partial\omega^2}\right]$$

First, compute the first-order derivative:

$$\frac{\partial \ln p(\mathbf{x};\omega)}{\partial\omega} = \frac{A}{\sigma_w^2}\sum_{n=0}^{N-1}(x[n]-A\cos(\omega n+\phi))\cdot n\sin(\omega n+\phi)$$

Next, compute the second-order derivative:

$$\frac{\partial^2 \ln p(\mathbf{x};\omega)}{\partial\omega^2} = -\frac{A}{\sigma_w^2}\sum_{n=0}^{N-1}\left[n^2 A\cos(\omega n+\phi)\sin(\omega n+\phi) - n(x[n]-A\cos(\omega n+\phi))\cos(\omega n+\phi)\right]$$

Since the noise $w[n]$ has a mean of 0, we have:

$$\mathbb{E}\left[x[n]-A\cos(\omega n+\phi)\right] = 0 \tag{1.1}$$

The cross term satisfies:

$$\mathbb{E}\left[w[n]\cdot\cos(\omega n+\phi)\right] = 0 \tag{1.2}$$

Thus, we finally obtain:

$$I(\omega) = \frac{A^2}{\sigma_w^2}\sum_{n=0}^{N-1} n^2\cos^2(\omega n+\phi)$$

For large N, the summation can be approximated by an integral, using trigonometric properties:

$$\sum_{n=0}^{N-1} n^2 \cos^2(\omega n + \phi) \approx \sum_{n=0}^{N-1} n^2 \cdot \frac{1}{2} = \frac{1}{2} \sum_{n=0}^{N-1} n^2$$

Using the summation formula:

$$\sum_{n=0}^{N-1} n^2 = \frac{(N-1)N(2N-1)}{6}$$

Thus,

$$I(\omega) \approx \frac{A^2}{\sigma_w^2} \cdot \frac{1}{2} \cdot \frac{(N-1)N(2N-1)}{6} = \frac{A^2(N-1)N(2N-1)}{12\sigma_w^2}$$

Cramér-Rao Bound: The Cramér-Rao Bound (CRB) is the inverse of the Fisher information:

$$\text{CRB}(\omega) = \frac{1}{I(\omega)} = \frac{12\sigma_w^2}{A^2(N-1)N(2N-1)}$$

Substitute the Signal-to-Noise Ratio $\text{SNR} = \frac{A^2}{2\sigma_w^2}$

$$\text{CRB}(\omega) = \frac{6}{\text{SNR} \cdot N \cdot (N-1) \cdot (2N-1)}$$

In this experiment, we employ Monte Carlo simulations to repeatedly estimate the target frequency ω_0 and assess the average deviation between the estimated values and the true frequency. The metric used to quantify the estimation error is the Mean Squared Error (MSE), defined as:

$$\text{MSE} = \frac{1}{L} \sum_{i=1}^{L} \left(\hat{\omega}_0^{(i)} - \omega_0 \right)^2 \tag{1.3}$$

where:

- L: Number of independent simulation runs (e.g., 100 iterations);
- $\hat{\omega}_0^{(i)}$: The estimated frequency obtained in the $i-th$ simulation;
- ω_0: The true (known) frequency.

In this experiment, we employ an approximate Maximum Likelihood Estimation (MLE) method to estimate the frequency ω_0 of a sinusoidal signal.
Given the complexity of implementing the exact analytical form of the MLE, we adopt a numerical approach that combines:

- Initial estimation via Fast Fourier Transform (FFT);

- Local search-based ptimization refinement;

This hybrid strategy offers a favorable trade-off between estimation accuracy and computational complexity, making it suitable for practical implementation.
The observed signal $x[n]$ is first zero-padded and then transformed using the Fast Fourier Transform (FFT). The frequency corresponding to the maximum magnitude in the resulting spectrum is selected as the initial frequency estimate, denoted as ω_{FFT}

$$\boldsymbol{\omega}_{\mathrm{FFT}} = \arg\max_{\omega} |\mathrm{FFT}(\boldsymbol{x}[\boldsymbol{n}])| \tag{1.4}$$

A finely spaced search interval is constructed in the neighborhood of ω_{FFT}:

$$\boldsymbol{\omega} \in \left[\boldsymbol{\omega}_{\mathrm{FFT}} - \frac{\pi}{\boldsymbol{N}_{\mathrm{FFT}}}, \boldsymbol{\omega}_{\mathrm{FFT}} + \frac{\pi}{\boldsymbol{N}_{\mathrm{FFT}}}\right] \tag{1.5}$$

For each candidate frequency ω within the search interval, the likelihood (or matching score) is computed as:

$$\text{likelihood}(\omega) = \sum_{n=0}^{N-1} x[n] \cdot \cos(\omega n) \tag{1.6}$$

The frequency estimate $\hat{\omega}_0$ is then determined as the frequency that maximizes this likelihood function.
A larger N indicates a longer observation window, meaning the signal contains more frequency information.
As a result:

- The frequency resolution of the FFT spectrum improves.
- The peak in the local search region becomes more pronounced and easier to identify.

As the number of samples N increases, the MSE of frequency estimation decreases, since more data provides higher frequency resolution and stronger statistical stability.
The frequency points in the FFT output are discrete and uniformly spaced, with a frequency interval given by:

$$\Delta \boldsymbol{f} = \frac{1}{N_{\mathrm{FFT}}} \quad \Rightarrow \quad \Delta \boldsymbol{\omega} = \frac{2\pi}{N_{\mathrm{FFT}}} \tag{1.7}$$

A larger N_{FFT} results in finer frequency resolution.
This leads to a denser search grid around the true frequency, allowing the identified peak to more closely approximate the actual frequency.

The system is defined by the transfer function:

$$\boldsymbol{H(z) = \frac{2}{1 + 0.9z^{-1}}} \tag{1.8}$$

This represents a first-order IIR filter. This is a high-pass filter, as its pole is located at $z = -0.9$, which lies near the unit circle in the direction of $\omega = -\pi$. As a result, the filter exhibits the following characteristics:

- Enhancement of high-frequency components;
- Attenuation of low-frequency components

In the time domain, the corresponding recursive implementation is as follows:

$$y\left[n\right] = -0.9y\left[n-1\right] + 2x\left[n\right] \quad \rightarrow \quad AR(1) \tag{1.9}$$

Observations from Simulation Results:

- Without filtering, while the MSE decreases significantly with increasing SNR and approaches the CRB, it worsens as the signal frequency increases. This is due to stronger spectral leakage and flatter peaks at higher frequencies under finite observation length, which reduce the estimator's frequency resolution.

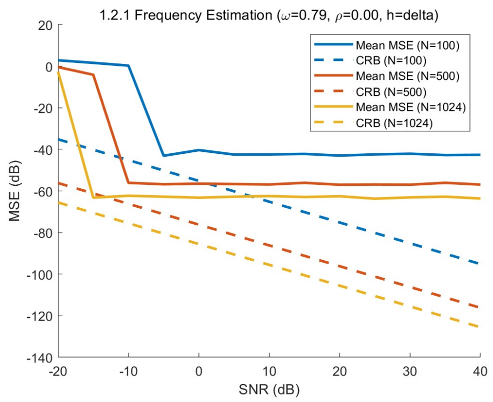

Figure 1.2: MSE/CRB vs. SNR ($\omega = \frac{\pi}{4}$)

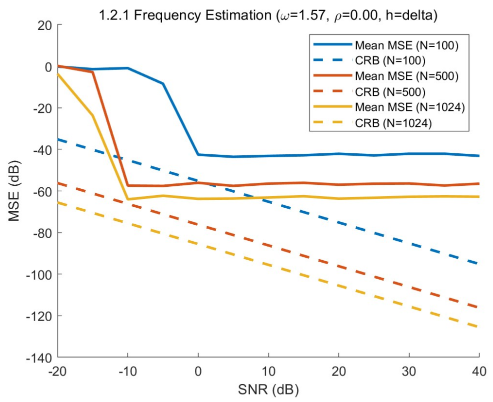

Figure 1.3: MSE/CRB vs. SNR ($\omega = \frac{\pi}{2}$)

- Without filtering ($h[n] = \delta[n]$), the signal structure remains intact, with undistorted spectral peaks. The estimator can reliably detect the true frequency, especially under low SNR, as δ avoids noise amplification. This leads to lower MSE and more robust performance.

  The filter $H(z) = \frac{2}{1+0.9z^{-1}}$ functions as a high-pass filter and significantly affects

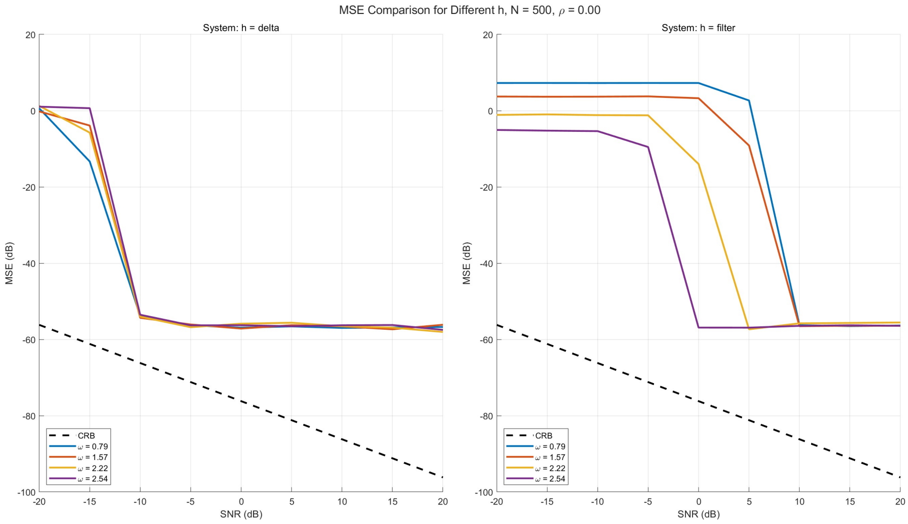

Figure 1.4: MSE vs. SNR in different filter

low-frequency signals. As shown in Figure 1.4, in the low-to-moderate SNR range, it requires higher SNR than the δ filter to achieve the same MSE performance.

- At high SNR levels, the MSE curve flattens due to inherent limits in the estimation algorithm, such as resolution limits, spectral leakage, and frequency quantization error. Increasing SNR further yields negligible improvement beyond this error floor.

As ρ increases, the correlation between noise samples strengthens, and the spectral energy concentrates toward lower frequencies, showing a "red noise" behavior. The power spectral density is:

$$S_w(\omega) = \frac{\sigma^2}{1 + \rho^2 - 2\rho\cos(\omega)} \tag{1.10}$$

It peaks at $\omega = 0$, showing that larger ρ concentrates noise energy in low frequencies, forming red noise.

In this study, we extend the noise model from white noise ($\rho = 0$) to correlated noise ($\rho = 0.9$), introducing a more complex covariance structure. Correlated noise exhibits temporal memory, which means that noise samples are no longer independent. This leads to longer-lasting interference and more severe distortion of the sinusoidal signal's frequency characteristics.

Experimental Findings:

Compared to the white noise scenario, the MSE increases significantly, especially in the low-to-moderate SNR range.

The Cramér–Rao Bound (CRB) also rises, indicating a reduction in the amount of available frequency information.

Even in high SNR conditions, the MSE remains far from the CRB, revealing that the

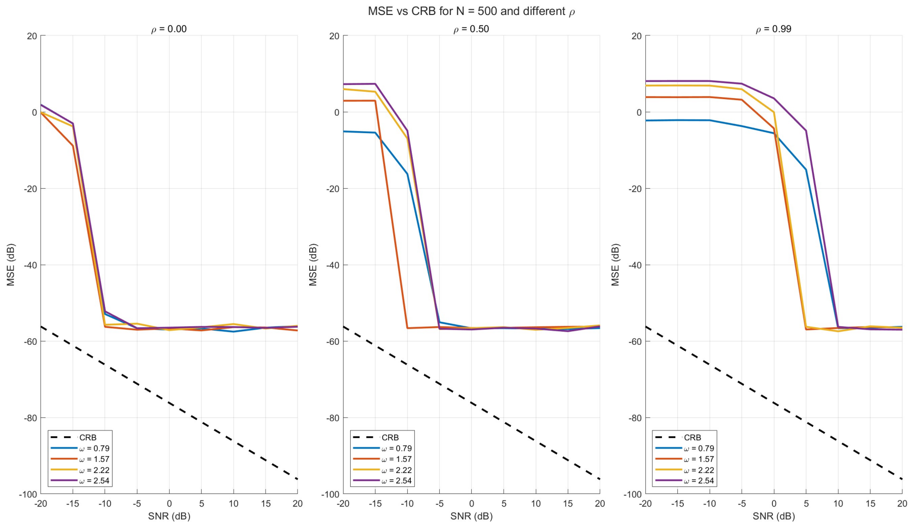

Figure 1.5: MSE vs. SNR in different ρ

estimator's efficiency is substantially limited by the presence of correlated noise.

1.3. Frequency modulation

The frequency-modulated signal is given by:

$$x[n] = a \cdot \cos\left(\gamma n^2 + \phi\right) + w[n] \tag{1.11}$$

Where:

- $w[n]$ is zero-mean Gaussian white noise with covariance $C_w w = \sigma_w^2 I$,
- Sampling frequency f_S =22kHz,
- Total samples $N = 22 \times 10^3$ (signal duration: 1 second).

$$\omega(n) = \frac{d}{dn}\left(\gamma n^2 + \phi\right) = 2\gamma n \tag{1.12}$$

Design requirements:

- At $n = 0$, $\omega(0) = \pi/8$,
- At $n = N = 22000$, $\omega(N) = 3\pi/4$.

So to compute γ:

$$\omega(N) - \omega(0) = 2\gamma(N - 0) = \frac{3\pi}{4} - \frac{\pi}{8} \tag{1.13}$$

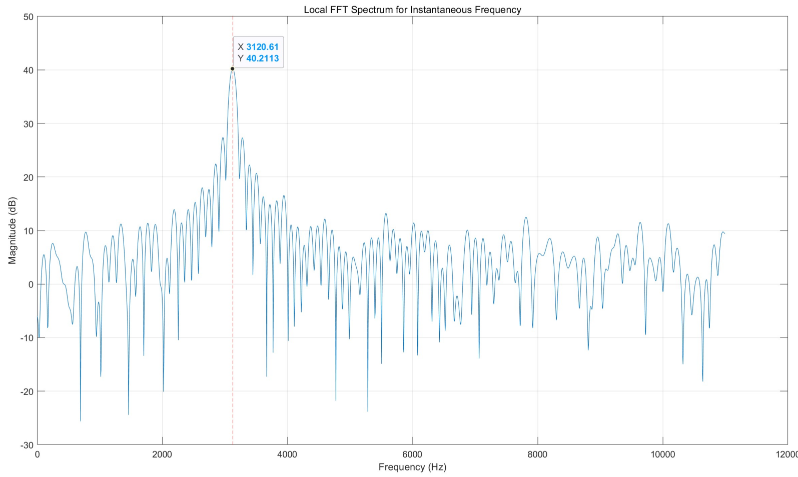

Figure 1.6: Instantaneous Frequency Estimation at $n_0 = 10000$

That leads to $\gamma = \frac{\frac{3\pi}{4} - \frac{\pi}{8}}{2N} = \frac{5\pi}{16N}$. Calling $\theta(n) = \gamma n^2 + \phi$, the instantaneous frequency for discrete signals computed as:

$$\boldsymbol{f_{\text{inst}}(n) \approx \frac{1}{2\pi}(\theta(n+1) - \theta(n)) \cdot f_s} \tag{1.14}$$

Spectral Analysis and Frequency Estimation: The windowed signal is zero-padded and transformed using the Fast Fourier Transform (FFT). The position of the dominant spectral peak is identified and used to estimate the instantaneous frequency.

Frequency Unit Conversion and Validation: The estimated frequency, originally in angular units (radians per sample), is converted to Hertz (Hz). This value is then compared against the theoretical instantaneous frequency defined by: $\omega(n_0) = 2\gamma n_0$ to verify estimation accuracy.

As shown in Figure 1.6, a window centered at $n_0 = 10000$ with a length of 200 samples was selected for analysis.
The simulation results indicate that the estimated instantaneous frequency is 3.12 kHz.

According to the theoretical model, the instantaneous frequency at this time point is given by:

$$f_{inst-theory}(n_0) = \frac{2\gamma n_0}{2\pi} \cdot f_s = 3.125kHz \tag{1.15}$$

Figure 1.6 presents the frequency estimation at a fixed time index n_0. Figure 1.7 shows the short-time frequency content across the full 1-second signal using the STFT spectrogram. The figure clearly reveals two time-varying frequency components: one sweeping

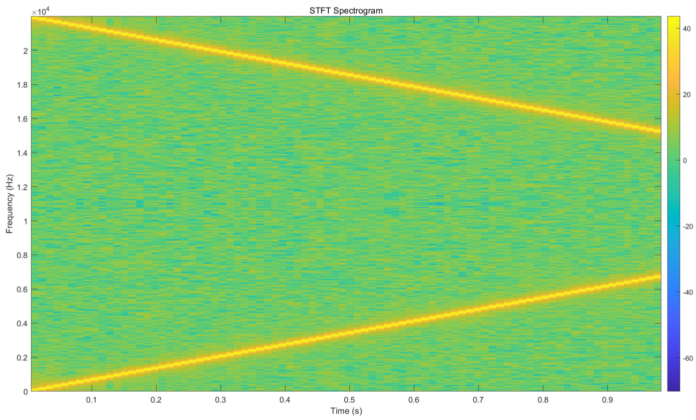

Figure 1.7: STFT Spectrogram of a Chirp Signal

downward and the other upward, forming a symmetric structure.

1.3.1. Maximum Likelihood Estimator

Since the noise $w[n]$ is assumed to be i.i.d. Gaussian, the probability density function of each observation $x[n]$ is given by:

$$p(x[n];\gamma) = \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left(-\frac{(x[n] - a\cos(\gamma n^2 + \phi))^2}{2\sigma_w^2}\right) \tag{1.16}$$

Joint Likelihood Function:
Because the noise samples are independent, the joint likelihood of the full observation vector $\mathbf{x} = [x[0], x[1], \&, x[N`1]]^T$ is the product of the individual likelihoods:

$$p\left(x;\gamma\right) = \prod_{n=0}^{N-1} p\left(x\left[n\right];\gamma\right) \tag{1.17}$$

Taking the logarithm of the joint likelihood function yields the log-likelihood:

$$\ln p(\mathbf{x};\gamma) = -\frac{N}{2}\ln(2\pi\sigma_w^2) - \frac{1}{2\sigma_w^2}\sum_{n=0}^{N-1}\left(x[n] - a\cos(\gamma n^2 + \phi)\right)^2 \tag{1.18}$$

Since the first term is constant with respect to γ, maximizing the likelihood is equivalent to minimizing the squared error term:

$$\hat{\gamma}_{\text{MLE}} = \arg\min_{\gamma} \sum_{n=0}^{N-1}\left(x[n] - a\cos(\gamma n^2 + \phi)\right)^2 \tag{1.19}$$

which is a non-linear least-squares problem and can be solved numerically.
The maximum likelihood estimation problem is thus transformed into a non-linear least squares estimation problem: Finding the value of γ that minimizes the fit error between the observed signal and a frequency-modulated sinusoidal model.