DUAL DEGREE PROJECT

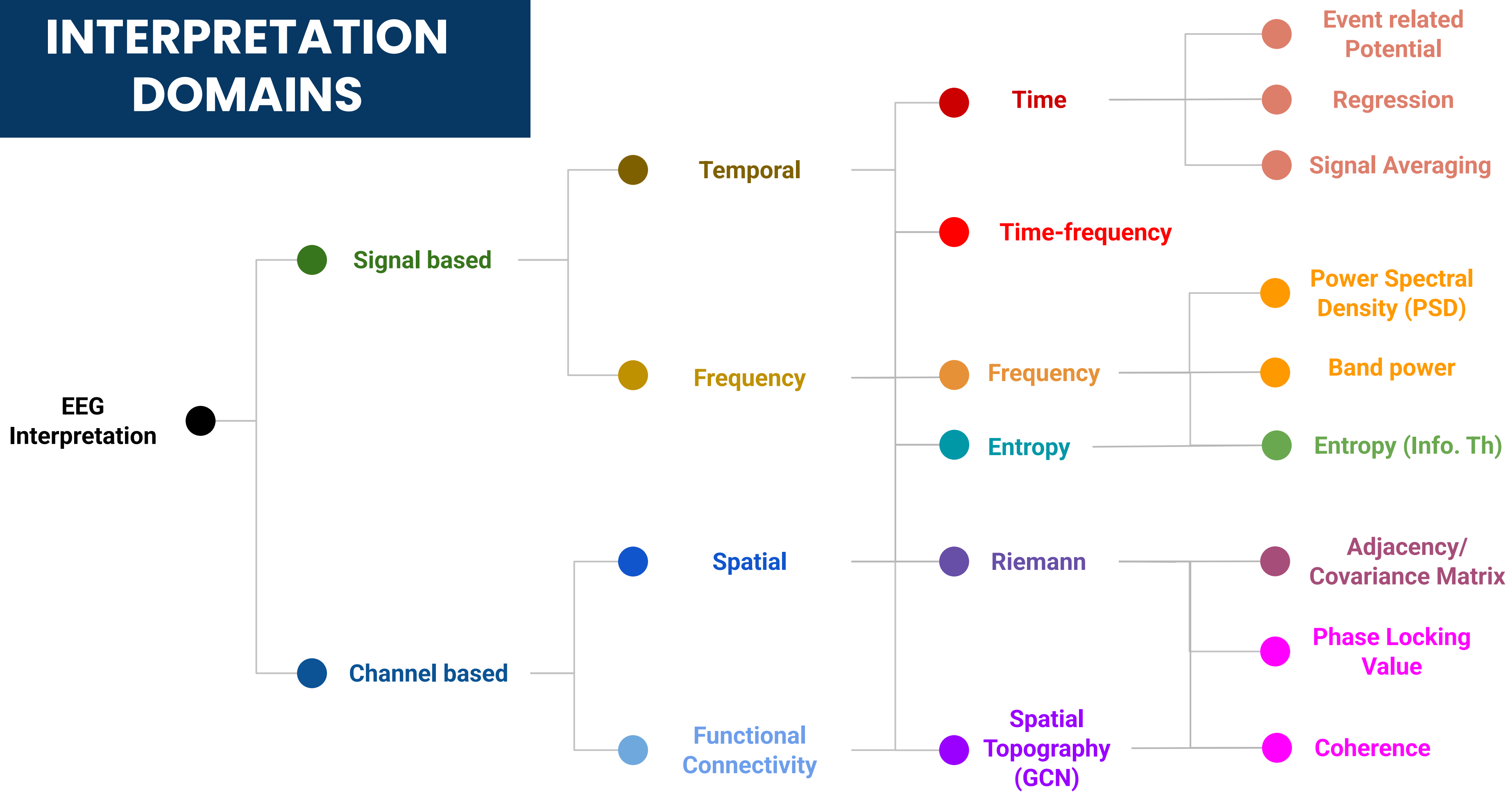
# Multifaceted Interpretation of EEG Signals

By Sanjeev M

Under guidance of

Prof. Chandrashekar Lakshminarayanan

# EEG SIGNAL INTERPRETATION DOMAINS



# Motivation

Real-world brain signals are inherently multidimensional, containing key information across domains:

- **Spatial** (across electrode channels)
- **Temporal** (waveform evolution - Event related potentials, signal averaging)
- **Frequency** (oscillatory rhythms)
- and advanced **domains**: **time-frequency** (how frequencies change over time), **entropy** (signal complexity), and **Riemann domain** (spatial covariance and geometry).

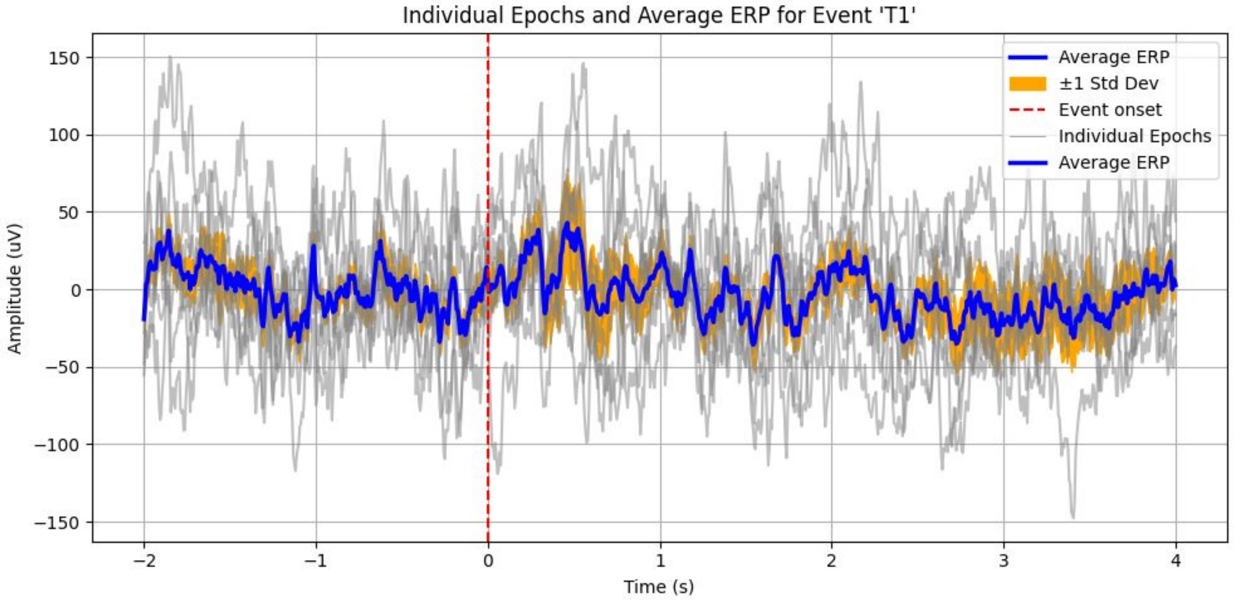
Traditional EEG analysis often focuses on a single domain, missing the complex interactions between these information sources, which circumscribes the information that we can use for modelling.

Our approach systematically extracts and integrates features from all these domains (Temporal, Frequency, Time-Frequency, Entropy, and Riemann), enabling a richer, more accurate understanding of neural dynamics.

This comprehensive strategy increases contextual information and, with careful attention to each feature, can significantly boost predictive accuracy. Each domain will be addressed in our methodology.

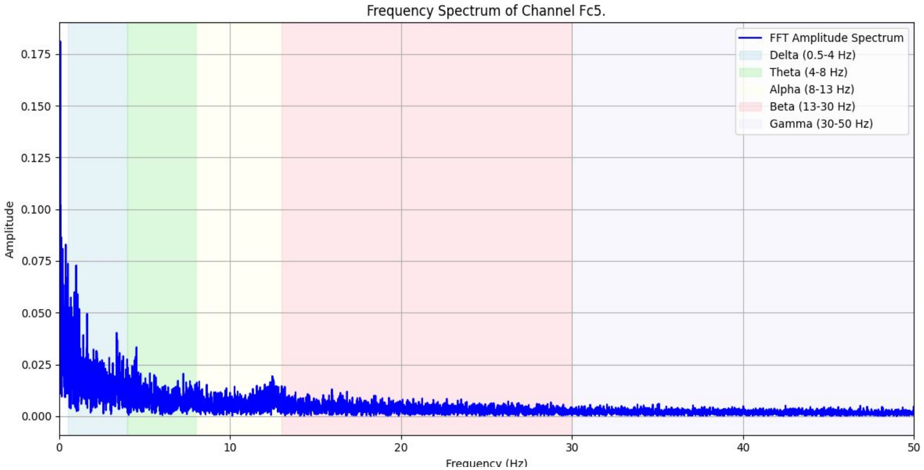
# Temporal domain

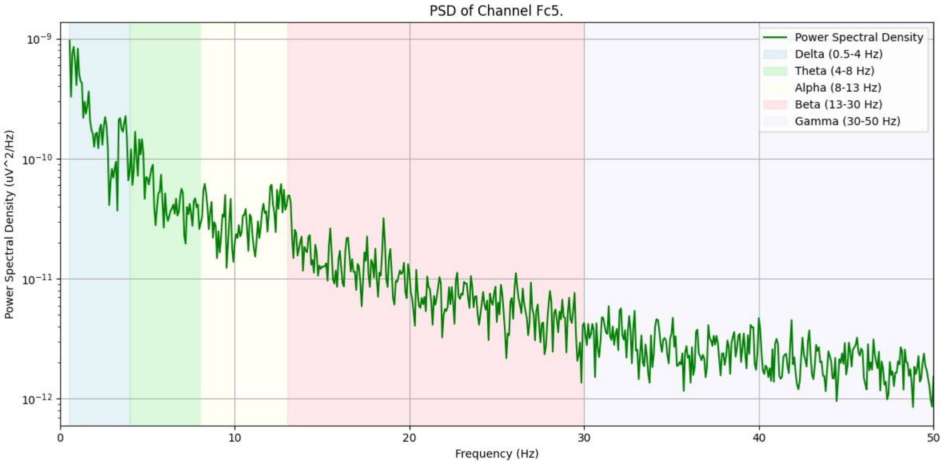
*x(t)=2sin(2π · 10t)+sin(2π · 20t)*




- **ERPs** - Event related potential
- Time domain **statistics** - mean, variance, peak latency
- **Signal Averaging** - set baseline at each event epoch

# Frequency domain




$$\mathrm{PSD(f)} = \lim_{T \to \infty} |X(f)|^2/T$$

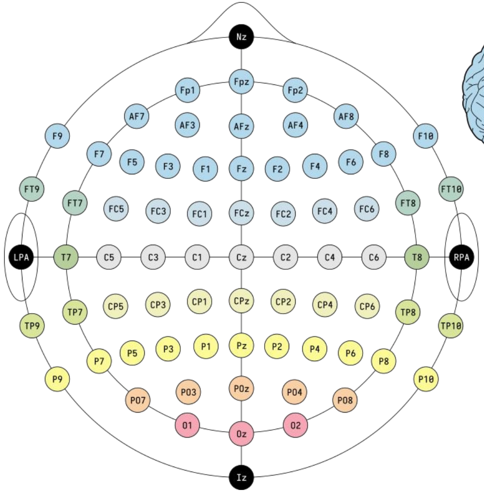
$$x[k] = \sum_{n=0}^{N-1} x[n] e^{\frac{-j2\pi kn}{N}}$$

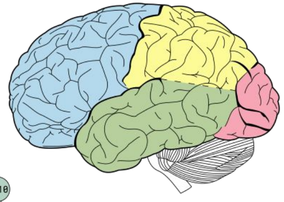
- **FFT** - Fast Fourier Transform
- **Frequency Bands:** delta, theta, alpha, beta, gamma in EEG
- **PSD**- power spectral density

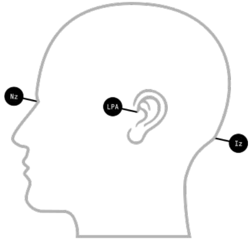
# International 10–20 / 10–10 System

**Definition**: Standardized method to place EEG electrodes on the scalp.
**10–20 system**: Electrodes are placed at 10% or 20% of the total distance between anatomical landmarks (nasion, inion, left/right preauricular points).
**10–10 system**: A higher-resolution extension with electrodes at 10% intervals → more electrodes, better spatial resolution.
**Purpose**:

- Ensures reproducibility across subjects/studies.
- Captures brain activity from **specific cortical regions** (frontal, temporal, parietal, occipital, central).
- Useful for **spatial localization** of signals.





# Entropy domain

1. Shannon entropy is the bedrock of information theory. You can introduce it by defining it as a measure of **uncertainty** in a discrete set of outcomes. The more unpredictable an event is, the higher its entropy.

$$H(X) = -\sum_{i=1}^{n} P(x_i)\log_2(P(x_i))$$

2. But **EEG signals** are **continuous** and Shannon entropy is used for discrete variables as probability can be estimated among them. So we shall use probability density which will give rise to **Differential Entropy.**

$$h(X) = -\int_{-\infty}^{\infty} f(x)\log_2(f(x))dx$$

3. **Assumption:** To calculate the probability density function f(x) we assume that a small section of EEG signals are from a Gaussian distribution. This will give us the probability density function as below:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad h(X) = \frac{1}{2}\log_2(2\pi e\sigma^2)$$

4. However **variance ($\sigma^2$)** of a frequency band is directly proportional to **Power Spectral Density (PSD)**

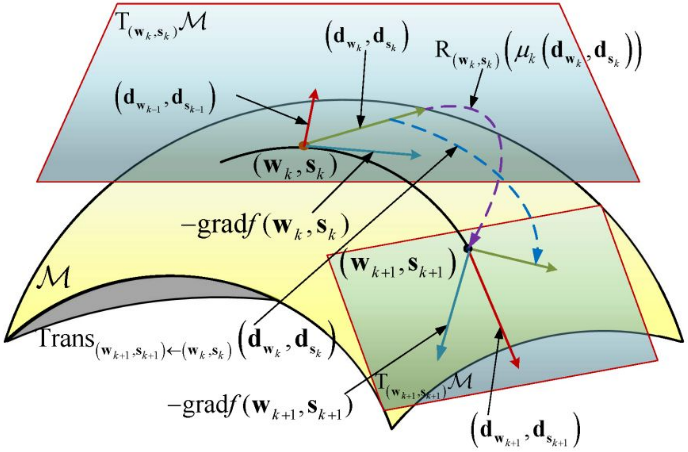
$$\sigma^2 \propto \text{PSD} \qquad h(X_i) = ½ \log_2(2\pi e.\text{PSD}_i)$$

5. This Differential entropy of channel X = H(X) = $\sum_i h(X_i)$ can be used to calculate the correlation between the channels using Mutual Information (MI) : I(X; Y)

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

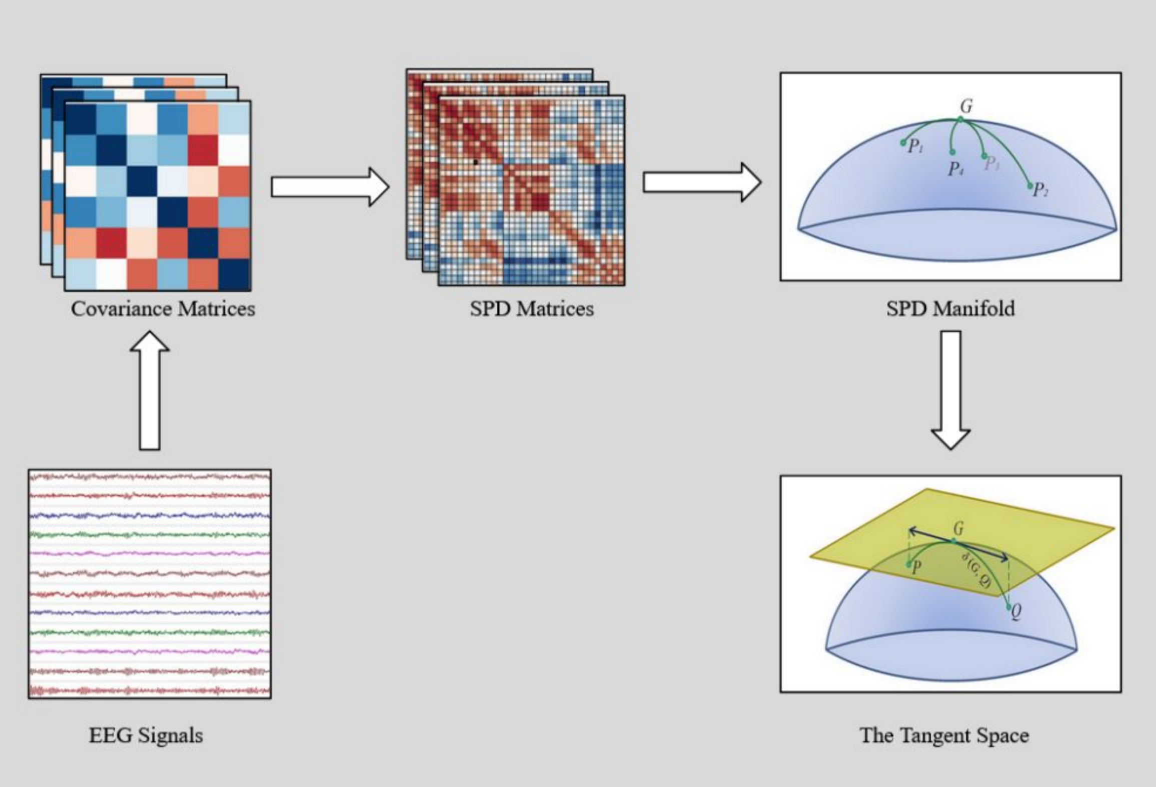
# Riemann domain

1. If EEG signals has N channels, covariance matrix will be a N x N matrix
2. The N x N matrix is a **SPD (Symmetric positive definite)** matrix. A matrix A ∈ R n×n is deemed Symmetric Positive Definite (SPD) if it holds that $A = A^T$ and $x^TAx > 0$ for all non-zero vectors x ∈ $R^n$. The eigenvalues of such a matrix A, denoted by $\lambda(A)$, are guaranteed to be positive.
3. The SPD manifold is a non-linear curved space not fit for Euclidean operations. Normally a manifold is **locally Euclidean** of dimension n, where each point's neighborhood resembles $R^n$. However, when a space such as the SPD manifold is non-linear and is equipped with a **differentiable structure**, then it is a differential manifold. Further, when it is aligned with Riemann geometry by being equipped with a metric, it becomes a **Riemannian manifold**.




*gradient descent algorithms on the Riemannian product manifold*

# Riemann domain




The SPD Manifold

# Riemann domain

1. **Covariance Matrix: (per epoch)**
   Consider $X \in R^{M \times T}$ , representing an EEG signal that has undergone band-pass filtering, where M denotes the number of channels and T represents the number of temporal samples. To analyze the statistical properties of the EEG signal, we construct the covariance matrix P as follows:

$$\mathbf{P} = \frac{1}{T-1}\mathbf{X}\mathbf{X}^T$$

   The covariance matrix P is Symmetric Positive definite (SPD) matrix.

2. **Riemann geodesic distance:**
   a. *Log Euclidean metric:* The Log-Euclidean Metric (LEM) offers an efficient and robust computational method for handling the manifold of SPD matrices. Under the LEM, the geodesic distance between two points P1 and P2 on the manifold of SPD matrices is defined as:

$$\delta_L(P_1, P_2) = \| \operatorname{Log}(P_1) - \operatorname{Log}(P_2) \|_F$$

   *b.* *Affine-Invariant Metric (AIM):* The AIM is another highly regarded metric for SPD manifolds, especially valued for its invariant properties under affine transformations. The geodesic distance under this metric between P1 and P2 is computed as:

\

$$\delta_A(P_1, P_2) = \left\| \operatorname{Log}(P_1^{-1/2} P_2 P_1^{-1/2}) \right\|_F$$

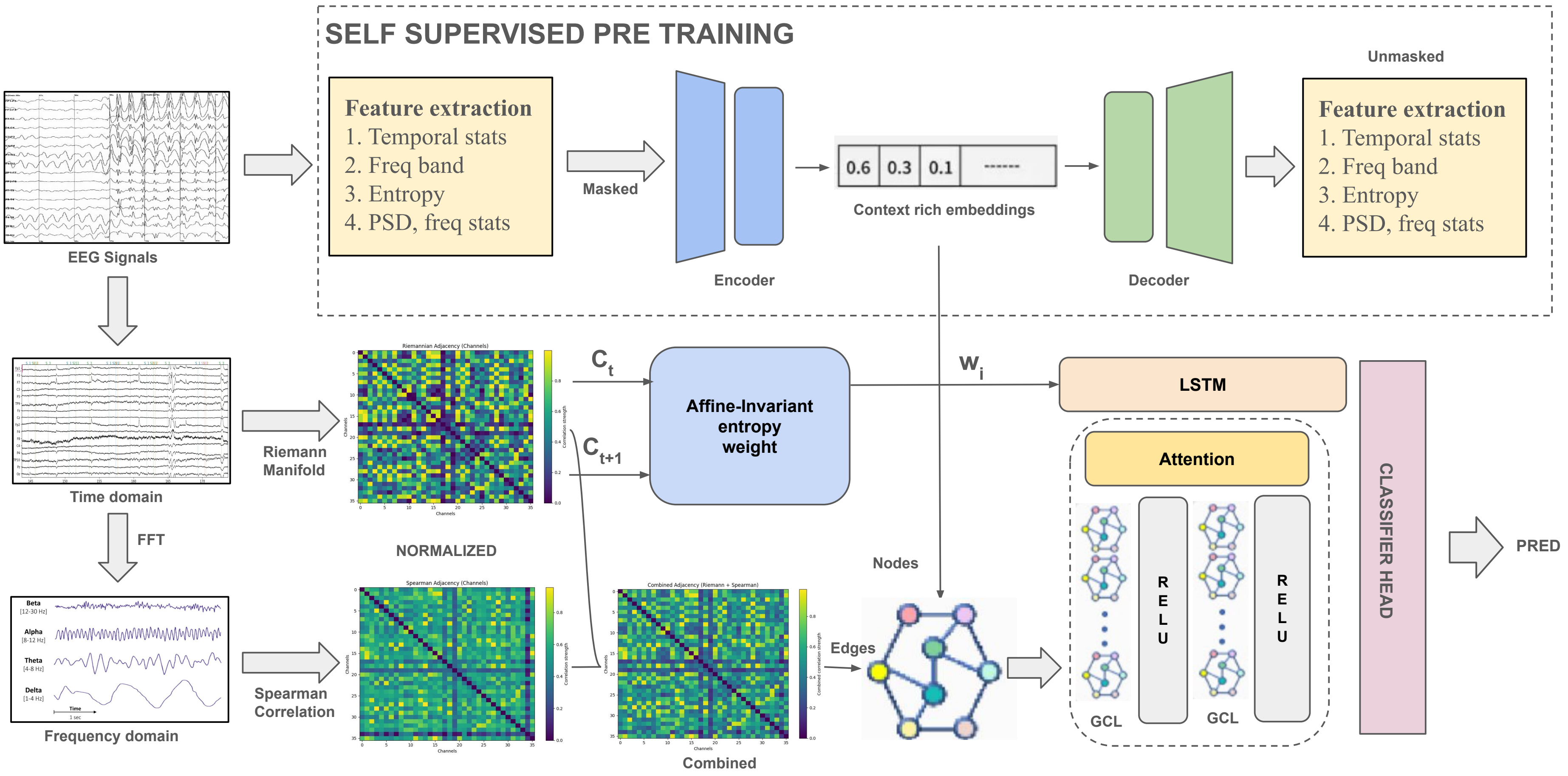
# Riemann domain

1. **Channel Correlation:**
   a. You can see that in the **Log Euclidean metric** - essentially when each element, take Cij for example, is converted to Log (Cij), then every euclidean operation on the matrix will replicate calculation Riemann geometry. Then we can compute the Adjacency matrix using the off diagonal entries, compute the normalized matrix, and diagonal entries are made 0.
   b. This gives us the **Riemann Adjacency matrix**: R, which can be used in transfer learning because the current adjacency matrix gives meaningful relations on euclidean operations. This allows the model to generalize across multiple subjects. However, this won't be the only metric to compute the adjacency weights. Additionally we would be computing spearman correlation values.
      i. *Riemann adjacency matrix:* Computed from M x T matrix per epoch - Temporal domain
      ii. *Spearman correlation matrix:* computed using PSD - Frequency domain
2. **Inter state or Inter epoch correlation:**
   a. This can be used to calculate how much weight to be assigned to the memory passed through LSTM.
   b. Consider adjacency matrix $C_t$ at time t and the result $h_t$ to be passed to $h_{t+1}$ through LSTM.
   c. We will be using **AIRM (Affine Invariant Riemann distance)** for this. This distance is proportional to the **KL divergence** between normal distributions, $N(0, C_t)$ & $N(0, C_{t+1})$. According to information theory when there is high KL divergence, high similarity, so less info shall be passed and if there is less **KL divergence,** then more info needs to be passed. That's the idea.
   d. I call the weight of the information to be transferred as *Affine Invariant Entropy weight*

$$w_i = e^{-(d^2_{AIRM}/\sigma^2)}$$

$$h_{t+1} = \mathrm{LSTM}(x_{t+1}, w_i \cdot h_t, c_t)$$

# Model Architecture



THANK YOU