Multiplex recurrence networks from multi-lead ECG data

Cite as: Chaos 30, 123106 (2020); doi: 10.1063/5.0026954
Submitted: 26 August 2020 · Accepted: 6 November 2020 · Published Online: 1 December 2020

Sneha Kachhara\textsuperscript{a}\textsuperscript{)} and G. Ambika\textsuperscript{b}\textsuperscript{)}

AFFILIATIONS
Indian Institute of Science Education and Research (IISER) Tirupati, Tirupati 517507, India
\textsuperscript{a)}Author to whom correspondence should be addressed: snehakachhara@students.iisertirupati.ac.in
\textsuperscript{b)}g.ambika@iisertirupati.ac.in

ABSTRACT
We present an integrated approach to analyze the multi-lead electrocardiogram (ECG) data using the framework of multiplex recurrence networks (MRNs). We explore how their intralayer and interlayer topological features can capture the subtle variations in the recurrence patterns of the underlying spatio-temporal dynamics of the cardiac system. We find that MRNs from ECG data of healthy cases are significantly more coherent with high mutual information and less divergence between respective degree distributions. In cases of diseases, significant differences in specific measures of similarity between layers are seen. The coherence is affected most in the cases of diseases associated with localized abnormality such as bundle branch block. We note that it is important to do a comprehensive analysis using all the measures to arrive at disease-specific patterns. Our approach is very general and as such can be applied in any other domain where multivariate or multi-channel data are available from highly complex systems.

Published under license by AIP Publishing. https://doi.org/10.1063/5.0026954

The electrocardiogram (ECG) is a record of the electrical activity of the heart in the form of a time series. The study of cardiac dynamics through ECG has gathered a lot of attention in the nonlinear dynamics community, with attempts to justify the degree of chaos in this system and to identify anomalies in the case of a disease. Nonlinear time series methods, sometimes coupled with machine learning approaches, have been employed to this end with reasonable success. However, the interpretation of feature-based classification studies in terms of underlying dynamics is not explored, except for some particular ailments such as arrhythmias and chronic heart failure. Moreover, the ECG comes as highly correlated multivariate data from 3, 5, or 12 leads. Most of the study from the dynamics point of view until now are on the average data over leads or on a single lead. In the present work, we aim to study multi-lead ECG within the framework of multiplex recurrence networks (MRNs), which highlights spatio-temporal features of the cardiac dynamics as reflected in ECG. To this end, we employ layer similarity/dissimilarity measures in addition to the standard complex network measures defined for multiplex complex networks. We include three levels of structural aspects of MRNs: the coarse structure in terms of links across layers, the interlayer features in degree distributions, and the local micro-structures using local clustering coefficients. We show that abnormalities in the cardiac dynamics in the case of a disease manifest in a multitude of ways and can be understood only by consolidated results from a set of measures.

I. INTRODUCTION
Research related to chaotic behavior in the cardiovascular system has seen consistent progress in the last two decades.\textsuperscript{1,2} While the initial attempts were aimed at establishing the presence of deterministic chaos in the cardiac system, techniques of nonlinear time series analysis and machine learning are now being applied to understand cardiac dynamics.\textsuperscript{3} The focus has shifted to the identification of signatures of altered dynamics (chaotic or not) in patients as compared to healthy,\textsuperscript{2} leading to a good progress in machine learning-based diagnostics and early-warning tools. However, this approach relies heavily on the availability of huge databases of quality data to train the algorithms and provides little insight into the underlying dynamics that reflects the intricacies related to cardiac malfunctions.

The physiologically relevant signal in the context of the heart is the electrocardiogram (ECG) that records the electrical activity of the heart.\textsuperscript{1} The data on heart rate variability (HRV), on the
other hand, are the number of cardiac cycles per minute: both ECG and HRV reflect related but different aspects of the cardiovascular dynamics.\(^7\) While HRV data have the advantage that it is easy to obtain, even with simple wearable devices, the full ECG waveform, on the other hand, is very sensitive to noise, requires specific devices to record, and is usually available only for a short duration. However, the ECG contains more information about the dynamics and also at different timescales. As such, analyzing ECG from a dynamical systems perspective can be quite rewarding and relevant.\(^8\) Studies in this direction indicate a reduction in complexity in some diseases such as chronic heart failure (CHF) and arrhythmias;\(^4,10\) however, the question of how exactly the dynamics is altered in the case of specific diseases, and measures that reflect these abnormalities from a dynamical point of view, remains relatively unexplored.

We note that limitations such as short duration and non-stationarity may have restricted the application of tools of non-linear time series analysis to ECG data. Recently, a few methods have been developed to transform a given time series into a complex network.\(^9\) Among them, the method of recurrence networks (RNs) proves to be useful in analyzing short and non-stationary data. It maps the recurrence pattern in the reconstructed dynamics from a given time series into a complex network.\(^11\) In this context, we have recently reported bimodality in the degree distribution and scaling of link density with recurrence threshold as characteristic features of RNs from ECG.\(^12\)

The multivariate data taken from a complex dynamical system have properties that can be uncovered only with a comprehensive approach of analysis. The cardiac dynamics by nature is spatio-temporal and can be understood better if data from different spatial locations like that from the multi-lead ECG are used in the analysis. There has been some progress in developing diagnostic tools based on a 3-lead vectorcardiogram (VCG), which is constructed from the 12-lead ECG and provides a succinct representation of phase relationships. However, the 12-lead ECG is preferred\(^13\) in the clinical practice by cardiologists to make a diagnosis based on some or all of the leads, depending on the nature of the disease. In the context of multivariate data, Eroglu et al.\(^14\) have proposed the framework of multiplex recurrence networks (MRNs) in which the layers of the network correspond to the different time series of the data. The patterns of connections inside each layer are governed by the dynamics as reflected in the corresponding time series of the original data. The framework is very recent but has been successful with different types of applications such as analysis of palaeobotanical data,\(^15\) climate forcings for Indian summer monsoon,\(^16\) multiphase flow dynamics,\(^17\) driver fatigue detection,\(^18\) linkage between carbon and energy markets,\(^19\) and even for self-reports of human experience (EMA and ESM data).\(^20\)

Recent approaches based on machine learning rely heavily on the nature and quantity of training data, and the set of features are not necessarily meaningful for understanding the dynamical aspect of the underlying cardiac conditions. Our method takes a different approach, built on the dynamical features associated with a cardiac condition and stressing on its similarity (or dissimilarity) between leads. Moreover, the embedding parameters can be easily tuned for data with different properties such as sampling frequency, gaps, etc., such that the resulting reconstructed dynamics would reflect the same dynamical features. Hence, the MRN framework is more flexible than current approaches and captures the spatio-temporal dynamics of the cardiac system effectively.

We start with the hypothesis that the lead-to-lead variations and the similarity underlying data between pairs of leads are to be studied to understand the complexity of the spatio-temporal dynamics of a normal heart. Even though the dynamics of a normal heart is established as chaotic, the variations in its spatio-temporal dynamics are not studied in detail. The present study tries to capture these features using data from the different leads that correspond to responses from different key positions of the heart. Moreover, such a study can give information on any anomaly appearing in a specific set of leads, which can be evidence of a specific disorder. We show how this is feasible by using the framework of MRNs. We use the ECG data from the six precordial leads placed closest to the heart as the multivariate data. The data from all leads represent the electrical activity of the heart as captured from different spatial points. They are ordered and synchronized in time and for the same interval of time. The corresponding networks are therefore multiplexed by identifying nodes that correspond to points at the same time on the reconstructed trajectory across the layers. The analysis of inter-layer similarities in such MRNs can uncover patterns not apparent in a single layer alone, as well as provide deeper insight into localized anomalies.

We begin by describing briefly the construction of RNs and hence MRNs from the six layers of RNs and the network quantifiers defined on them. We study the MRNs at different levels of complexity; from individual links in layers to the overall topological features. We base the analysis on three quantifiers: links, quantified by average edge overlap; degree distributions compared between pairs of different layers; and distribution of simple cliques, quantified by average local clustering coefficients in different layers. In this way, we can compare MRNs in two ways: one in terms of overall average values of quantifiers (one value for each subject) and the other in terms of measures for every pair of layers (15 values for each subject, since we have 15 unique pairs of layers from the six layers). We illustrate how specific diseases can cause differences in measures of similarity between particular pairs of layers, reflecting the intricate variations in the underlying spatio-temporal dynamics.

II. DATA AND PROCESSING

We use the PTB diagnostico\(^12,14\) database from Physionet.\(^21\) In total, data from 125 subjects are used in our analysis, 51 of which are healthy (HC) and the rest of them are from one of the four diseases: bundle branch block (BB, 15), cardiomyopathy/heart failure (CM, 16), dysrhythmia (DR, 14), and myocardial infarction with no secondary diagnosis (MI, 29). Each record originally consists of 12-lead ECG, mostly 60 s in duration, sampled at 1000 Hz (60,000 points). We choose data from the precordial leads \(v_1\) to \(v_6\) that are placed closest to the heart for analysis. Each data is pre-processed,\(^22\) filtered with 0.5–40 Hz, normalized in the range \((0,1)\), and binned to 5000 points.

III. CONSTRUCTION OF MULTIPLEX RECURRENCE NETWORKS

The construction of multiplex recurrence networks (MRNs) is done by first constructing the recurrence network from the
time series of each lead. This procedure involves two main steps: embedding the time series in a phase space based on Taken’s embedding theorem and identifying each point on the reconstructed phase space trajectory (or attractor) as a node of the network, making connections between them based on their proximity in recurrences. The reconstruction of the attractor in phase space requires a delay time \( \tau \) and embedding dimension \( m \) that are to be chosen appropriately for the data under analysis. For ECG data from each lead, we fix the embedding dimension as 4 using the method of false nearest neighbors (FNNs). The delay time \( \tau \), i.e., the time at which autocorrelation falls to 1/e, is chosen as the minimum value of such \( \tau \) among the time series from the six leads. Then, a recurrence threshold \( (\epsilon) \) is chosen to construct the links between nodes and to arrive at the adjacency matrix of the recurrence network. The variation of network measures with the recurrence threshold itself is an indicator of the complexity of the underlying attractor, as reported recently. In the present work, we set it to \( \epsilon = 0.1 \) for the sake of uniformity based on the previous study.

For the chosen \( \epsilon \), the recurrence matrix \( R \) is constructed such that if two points \( i \) and \( j \) on the reconstructed attractor lie within distance \( \epsilon \) of each other, we set the corresponding matrix element \( R_{ij} \) to be 1, and 0 otherwise, i.e.,

\[
R_{ij} = \Theta (\epsilon - \| \vec{v}_i - \vec{v}_j \| ),
\]

where \( \vec{v}_i \) and \( \vec{v}_j \) are the corresponding vectors of \( i \) and \( j \) in the phase space and \( \Theta \) is the Heaviside step function.

The adjacency matrix \( A \) of a recurrence network is obtained from \( R \) as

\[
A_{ij} = R_{ij} - \delta_{ij},
\]

where \( \delta_{ij} \) is the Kronecker delta function that is inserted to avoid self-links in the network. This construction results in an unweighted and undirected network of size \( N \), where \( N \) is the number of points on the reconstructed attractor.

We display in Fig. 1 the reconstructed attractors and the corresponding recurrence networks from two different leads for two typical datasets, one of healthy and one of BB. The differences between leads are obvious and more pronounced in the case of a disease. To capture such differences, we construct the MRN by treating RN from each lead as a single layer of a multilayer network, by connecting the nodes in different layers that correspond to the same instant of time in the reconstructed attractors. We keep the same ordering of layers as the lead numbers so that it is easy to correlate the results with locations of the leads. Thus constructed, the MRN can be represented by a supra-adjacency matrix \( M \), which consists of blocks of adjacency matrices from different layers as diagonal elements and identity matrices as off-diagonal elements.

\[
M = \begin{bmatrix}
A^{[1]} & I_N & \cdots & I_N \\
I_N & A^{[2]} & \cdots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
I_N & \cdots & I_N & A^{[m]} 
\end{bmatrix}
\]

Thus, it encapsulates the properties of the multivariate data effectively, with display of its complex features in the interlayer attributes.

For a typical dataset, we show ECG data from each lead, the supra-adjacency matrix, and the constructed MRN in Fig. 2. We analyze the MRNs with a primary focus on interlayer similarities and differences. The quantifiers used in the study are briefly described in Sec. IV and more details can be found elsewhere.

### IV. Measures from Multiplex Recurrence Networks

We expect a rich structure in multilayer networks because of both interlayer and intralayer connections. For multiplex networks, there is a node-to-node correspondence by construction that grants them an additional level of order. Keeping this ordered structure in mind, we select a few measures that can capture their properties.
For comparing RNs from two layers, we can consider features at two different levels: local, like degree distribution, and global, such as link density (LD). There are other measures such as the average clustering coefficient (CC) that incorporates both local and global structures. In addition to these, in the case of multiplex networks, we can effectively use a few other measures defined below which quantify the extent of similarity among the layers.

The average edge overlap, denoted by \( \omega \), is a consolidated quantifier of the overlap of links (or edges) considering all the network. First proposed by Lacasa et al., \( \omega \) gives an estimate of the expected fraction of layers containing a link defined as

\[
\omega = \frac{\sum_{m} \sum_{i} \sum_{j} A_{i,j}^{[l]}}{m \sum_{i} \sum_{j} (1 - \delta_{i,j} \sum_{l} A_{i,j}^{[l]})},
\]

where \( \delta_{i,j} \) stands for the Kronecker delta function, \( m \) is the number of layers, and \( A_{i,j}^{[l]} \) corresponds to the layer \( l \) as defined in Eq. (2). For a network with \( m \) layers, \( \omega \) by definition can take values in the range \([1/m,1]\); \( \omega \) being 1 only if links are identical in all layers. Thus, for MRNs from ECG with six layers, \( \omega \) will be in \([\frac{1}{6},1]\].

The cosine similarity (CS) is based on the inner product of two vectors and in the present context, in terms of vectors of the degrees of nodes \( D^{[l]} \) and \( D^{[l']} \) in two respective layers \( l \) and \( l' \), as

\[
CS = \frac{D^{[l]} \cdot D^{[l']}}{\|D^{[l]}\| \|D^{[l']}\|}.
\]

The index of dissimilarity (ID) calculates differences between the two distributions. The distributions considered here are that of local clustering coefficients. For two layers \( l \) and \( l' \), ID is defined as

\[
ID = \frac{1}{N} \sum_{i=1}^{N} \left| C_{i}^{[l]} - C_{i}^{[l']} \right|,
\]

where \( C_{i}^{[l]} \) and \( C_{i}^{[l']} \) are the local clustering coefficients for the respective layers and \( N \) is the total number of nodes.

The Jensen–Shannon divergence is a measure to discern two distributions based on the entropy of mixing. In the context of multiplex networks, we can use the concept of Jensen–Shannon distance (JSD) to assess the similarity of the probability degree distributions in different layers. For two layers \( l \) and \( l' \), with probability degree distributions \( P^{[l]} \) and \( P^{[l']} \), respectively, we have

\[
\text{JSD}(P^{[l]}||P^{[l']}) = \sqrt{D(P^{[l]}||M) + D(P^{[l']}||M)},
\]

where \( M = \frac{P^{[l]} + P^{[l']}}{2} \) is the point-wise mean and \( D \) represents Kullback–Leibler divergence.

The interlayer mutual information is calculated on the respective degree distributions as follows:

\[
I_{l,l'} = \sum_{k} \sum_{k'} P^{[l]}(k) P^{[l']}(k') \log \left( \frac{P^{[l]}(k) P^{[l']}(k')}{P^{[l]}(k) P^{[l']}^{n}} \right),
\]

where \( P^{[l]}(k) \) is the joint distribution of probability of degrees \( k^{[l]} \) in layer \( l \) and degree \( k^{[l']} \) in layer \( l' \). \( I_{l,l'} \) captures the topological similarity in the two layers.

In addition, we use the Pearson correlation coefficient to compare local clustering coefficients in the two layers,

\[
r(l,l') = \frac{\sum_{i=1}^{N} (C_{i}^{[l]} - \bar{C}^{[l]}) (C_{i}^{[l']} - \bar{C}^{[l']})}{\sqrt{\sum_{i=1}^{N} (C_{i}^{[l]} - \bar{C}^{[l]})^{2}} \sqrt{\sum_{i=1}^{N} (C_{i}^{[l']} - \bar{C}^{[l']}\bar{C}^{[l']})^{2}}},
\]

where \( C_{i}^{[l]} \) and \( \bar{C}^{[l]} \) are the local clustering coefficients for the respective layers and \( N \) is the total number of nodes.

The measure \( \omega \) is calculated over all layers (i.e., one value for each MRN), while other measures are calculated for every pair of layers (total 15 values for each MRN). Given two (or more) MRNs, we can compare them across every pair of layers in terms of measures such as JSD, CS, etc. We can also take average across all 15 pairs of layers for these measures in addition to \( \omega \) and average ID. Since the structure of MRN reflects the underlying spatio-temporal dynamics of the cardiac system, any specific variations in measures from two layers and statistically significant changes in a property across subjects of a class will help to understand how diseases can alter heart dynamics and functions.

V. AVERAGE MEASURES IN MRNs

The one-to-one correspondence of nodes between layers of multiplex networks enables us to compare structures across individual layers on a very basic and concrete level that of the individual links. In the case of MRNs, links across layers can be associated...
with the recurrences in the underlying dynamics that occur synchronously, making them even more relevant. We compute the average edge overlap, $\omega$ [Eq. (4)], for different datasets and the average link density for the six layers of MRN from every dataset. The results for all the datasets used in the study are shown in Fig. 3. Each circle represents one subject, and the size of the circle is proportional to the variance in link density (LD) across layers.

We note from Fig. 3 that there is an apparent positive correlation between $\omega$ and average LD. This correlation is explored further using the correlation coefficient for all datasets of a category, as shown in Table I. This positive correlation is not surprising since a network with a very high LD naturally has high $\omega$, merely because of an increased number of links in all layers. However, their exact relationship depends on how the links are distributed inside the layers. We note that the correlation is low for DR and high for MI and BB.

The results suggest that the MRNs from ECG, in general, have a high degree of association among layers, leading to high values for $\omega$. We note also that all the values for $\omega$ are higher than 0.4, much higher than the lowest bound of $\frac{1}{6} = 0.167$. In some cases, it reaches 0.8–0.9, which is very close to having all layers identical. This would mean that most of the links, or the recurrence points, are common across layers. The healthy cases mostly occupy the middle region, with no healthy subject having higher average LD than 0.6. However, in extreme cases of BB and CM, with either very low or very high average LD, the values of $\omega$ are high (as compared to a healthy person with the same average LD). This leads to the conjecture that there are disproportionate changes in some of the layers in these cases. We will explore this possibility further with the interlayer similarity measures in Sec. VI.

### VI. INTER-LAYER SIMILARITIES OF DEGREES

In this section, we discuss the measures to relate the degree distributions across layers. We first discuss how the degree of each node differs from layer to layer as captured by the cosine similarity or CS, and then, how the overall distribution of degrees differs from layer to layer through Jensen–Shannon distance, $\text{JSD}$, and mutual information, $I$.

The CS by definition is a local measure and captures node-to-node differences across layers. We compute the CS values taking leads pairwise as (1,2), (1,3) etc., for each dataset. We present the average values of CS computed pairwise for data of each category in Fig. 4(a) along with standard errors. We find that for layer pairs (1,2), (1,4), (2,4), (2,6), and (4,6), the values from healthy are the lowest and, in general, BB has the highest value in all layers.

Similarly, the values for $\text{JSD}$ are calculated pairwise and presented in Fig. 4(b). We observe that these values are consistently higher (from healthy) for BB and CM, while DR and MI are closer to healthy for most pair of layers.

Finally, the values for $I$, calculated in a similar fashion, are presented in Fig. 4(c). We find that these values for pairs of layers (1,2), (1,4), (2,4), (2,6), and (4,6) are highest for healthy.

All the three measures are averaged over all pairs in each dataset, and then, their means and standard deviations for all datasets in each category are tabulated in Table II. We see that $\text{JSD}$ is higher for BB and CM, while for DR and MI, it stays close to the corresponding range for healthy. As for $I$, healthy have a high value for most of the pairs of layers.

### VII. INTERLAYER SIMILARITY AS REFLECTED IN LOCAL CC

In this section, we compute the average value of local clustering coefficients (CCs) of each layer of the MRN in order to analyze the microstructure. We take the average CC for all the layers and plot them against average LD for every dataset, as shown in Fig. 5. Each circle in the figure represents a subject, and its size is proportional to the variance in average CC across layers for that subject. We note that the magnitude of variance is low for healthy subjects. Also, the average CC and LD are correlated, as reflected in the correlation coefficient summarized in Table III. Specifically, for DR, we observe that the values lie significantly away from the main diagonal and hence differ from healthy and MI.

Now to depict the dissimilarity in interlayer topology, we compute the index of dissimilarity (ID), which is a measure of the...
differences in the two distributions of local CC values in two layers. The results are shown in Fig. 6(a). We also calculate the Pearson correlation coefficient for the local CC values for each pair of layers. That is, for every subject, for a given pair of layers, we calculate the local clustering coefficients for both layers and then take the Pearson correlation between them. In this way, we obtain correlation between CC values for each dataset for every pair of layers. These values are depicted in Fig. 6(b), for each class, for each pair of layers. The average value of both these measures, over all pairs of layers, for each category is presented in Table III. We find that, in general, ID is higher than healthy for all diseases but most prominently so for CM. Correspondingly, the correlation is also low for CM.

### VIII. VARIATION IN MEASURES AMONG PAIRS OF LAYERS

The different measures computed for MRNs from multi-lead ECG data and presented in Secs. I–VII are further analyzed statistically and consolidated together in this section. For this, we use Welch’s t-test to compute a significance value for every measure for each category and pair of layers so that we can understand how significant the differences in computed measures are for each disease from healthy. The results are summarized in the form of

<table>
<thead>
<tr>
<th>Category</th>
<th>CS</th>
<th>JSD</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>0.1626 ± 0.0189</td>
<td>0.1293 ± 0.0104</td>
<td>0.0523 ± 0.0176</td>
</tr>
<tr>
<td>BB</td>
<td>0.1844 ± 0.0063</td>
<td>0.1549 ± 0.0145</td>
<td>0.0444 ± 0.0138</td>
</tr>
<tr>
<td>CM</td>
<td>0.1480 ± 0.0099</td>
<td>0.1523 ± 0.0152</td>
<td>0.0408 ± 0.0124</td>
</tr>
<tr>
<td>DR</td>
<td>0.1469 ± 0.0060</td>
<td>0.1287 ± 0.0121</td>
<td>0.0374 ± 0.0104</td>
</tr>
<tr>
<td>MI</td>
<td>0.1664 ± 0.0074</td>
<td>0.1324 ± 0.0121</td>
<td>0.0424 ± 0.0107</td>
</tr>
</tbody>
</table>
FIG. 5. Values of average clustering coefficient (CC) and average link density of different subjects. Each circle represents a subject and its size indicates variance in CC across layers for that subject. The colors are representative of the category the subject belongs to, with healthy in green.

significance matrices in Fig. 7. Each entry in a given matrix represents the significance value for that category as compared to the corresponding measure for healthy, for the pair of layers indicated. A p-value of <0.05 is color-coded (p > 0.05 is white) from blue to green in decreasing order such that green indicates the least p-value, corresponding to the most significant. Also, p-value of <0.005 is indicated in the figure.

From the figure, we can conclude that different classes of diseases have different types of variations in the cardiac dynamics and not all measures show differences for all diseases. For BB, we observe isolated elements in the matrix of significance for measures $CS$, $JSD$, and $ID$. For CM, there seems to be a gradual change across adjacent layers, reflected strongly in $JSD$. For DR, there are few elements in green in all measures except in $I$, and for MI, significant difference is seen for multiple elements in both $I$ and $ID$. No pair of layers show any significant difference in $JSD$ for MI; hence, the corresponding matrix is completely white. These results indicate that a disease like BB could be localized; hence, some layers are affected more than others, while DR and MI manifest in an integrated manner affecting all the leads. We also observe that a measure like $I$ has multiple pairs of layers for each disease that are significantly different from healthy, while the measure $JSD$ is more specific. Since each measure encapsulates similarity/dissimilarity of a different kind, we expect them together to bring out disease-specific features.

FIG. 6. (a) Index of dissimilarity ($ID$) and (b) Pearson correlation coefficient for local clustering coefficients among pair of layers of MRNs constructed from ECG data. Each point in the plot represents the average value of the corresponding measure for the pair of layers across different subjects of a category (color-coded by category).

### TABLE III. Interlayer similarity measures, index of dissimilarity (ID), and correlation coefficient, in local CC, averaged across every pair of layers of MRNs for each category. The errors indicate standard deviations across subjects of a category.

<table>
<thead>
<tr>
<th>Category</th>
<th>$ID$</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>0.0395 ± 0.0096</td>
<td>0.4982 ± 0.1579</td>
</tr>
<tr>
<td>BB</td>
<td>0.0437 ± 0.0080</td>
<td>0.4445 ± 0.1809</td>
</tr>
<tr>
<td>CM</td>
<td>0.0484 ± 0.0065</td>
<td>0.3847 ± 0.1554</td>
</tr>
<tr>
<td>DR</td>
<td>0.0451 ± 0.0144</td>
<td>0.5077 ± 0.1602</td>
</tr>
<tr>
<td>MI</td>
<td>0.0414 ± 0.0082</td>
<td>0.4661 ± 0.1540</td>
</tr>
</tbody>
</table>
IX. CONCLUSIONS

The spatio-temporal features of the electrical excitation patterns of the heart are complex and recorded sequentially in the multi-lead ECG data. These leads capture ECG from different locations on the body, and variations in them can indicate cardiac malfunctions. In this study, we analyze the multi-lead ECG data from 125 subjects by constructing multiplex recurrence networks (MRNs) in order to discern patterns of variations in the underlying cardiac dynamics. This includes 51 healthy subjects and disease cases like bundle branch block (BB), cardiomyopathy (CM), dysrhythmia (DR), and myocardial infarction (MI). We take data from the six precordial leads (chest leads) and construct multiplex recurrence networks (MRNs) with six layers corresponding to the six leads. The present approach is more detailed, inclusive, and captures the spatio-temporal nature of cardiac dynamics from the temporal responses reaching at different spatial locations.

The measures specific to multiplex networks, such as edge overlap ($\omega$) and mutual information ($I$), are highly relevant for ECG data as they can highlight features of cardiac dynamics obscured by inherent non-linearity and correlations in data. The measures such as cosine similarity (CS) and Jensen–Shannon distance (JSD) can provide additional insight into the topological differences across layers. Moreover, the differences in distributions of local clustering coefficients (CCs), captured by the index of dissimilarity ($ID$)
and correlation coefficient, can provide more information on the interlayer similarities in their micro-structures.

The results on $\omega$ and average link density establish that MRNs from ECG have high similarity from layer-to-layer at the most basic level. $\omega$ for healthy is found between 0.4 and 0.8 and is generally higher than patients for the same average link density. Moreover, most cases of patients with high $\omega$ show a large variation in link density from layer to layer. Our results illustrate that healthy cardiac dynamics has less range of variations across layers as compared to diseases. Most extreme cases are those of BB and CM, while DR and MI are mostly within the range for healthy. The extreme values observed could be either due to some of the layers being very similar or vice-a-versa. The value of $CS$, which measures variations in degrees of nodes across layers, has the highest value for BB, as compared to other diseases and healthy.

We extend the study to interlayer similarities as reflected in degree distributions and distributions of local CCs. We find that the JSD is consistently high for BB and CM among all pairs of layers, while for DR and MI, the values are closer to healthy. On the other hand, $I$ is highest for healthy on average across all pairs of layers. In particular, we observe that layer 4 differs most from other layers in the case of BB and CM. For DR and MI, no particular layer or pair of layers stand out in terms of differences in degree distributions. We note that the healthy do not differ much from layer to layer in terms of average local CCs. Thus, we can infer that there is an overall coherence in the healthy cardiac system that is hampered in the case of a disease. There are some specific differences in the way different abnormalities manifest in cardiac dynamics. A localized anomaly, such as that of BB, for example, will affect only some of the layers, while in the case of MI, all the layers show significant differences.

Our study is aimed at exploring the nature of variations in the underlying spatio-temporal dynamics due to any type of disease as revealed through the measures computed from the multi-lead ECG data. Such an understanding of the dynamics is basic to a proper knowledge of the cardiac system and can provide insight for smart algorithms. The proposed framework can work with different data that differ in specifications such as sampling frequency, duration, gaps, etc., because these details only require changing the embedding parameters for reconstruction of dynamics. It makes the MRN framework much more flexible than most of the existing approaches. Further studies in this direction can combine deep learning approaches with dynamical systems theory. In this case, the measures calculated from MRNs can serve as input for the deep learning approaches with dynamical systems theory. In this case, the MRN framework much more flexible than most of the existing approaches.

Moreover, the type of analysis presented here can be applied to multivariate data in other domains where traditional approaches of the data analysis are insufficient. Thus, astrophysical data, climate data, financial data, EEG etc., are some examples where multivariate time series data from a single system are readily available. Furthermore, the framework of MRNs is not limited to time series data as the measures can be employed to analyze similarities in any real-world multiplex network where interlayer correlations are non-trivial and meaningful. It will be interesting to see such applications in other data-based networks.

ACKNOWLEDGMENTS

One of the authors, Sneha Kachhara, acknowledges research fellowship received from the Council of Scientific and Industrial Research (CSIR), India as Senior Research Fellow (SRF).

DATA AVAILABILITY

The data that support the findings of this study are openly available in PTB diagnostics ECG database at https://doi.org/10.13026/C28C71, Refs. 23 and 24.

REFERENCES

1. L. Glass, "Introduction to controversial topics in nonlinear science: Is the normal heart chaotic?", Chaos 19, 028501 (2009).


