In this study we propose a new information theoretic measure to quantify the complexity of biological systems based on time series data. We demonstrate the potential of our method using two distinct applications to human cardiac dynamics. Firstly, we show that the method clearly discriminates between segments of electrocardiogram records characterised by normal sinus rhythm, ventricular tachycardia and ventricular fibrillation respectively. Secondly, we investigate the multiscale complexity of cardiac dynamics with respect to age in healthy individuals using interbeat interval time series and compare our findings with a previous study which established a link between age and fractal-like long range correlations. The method we use is an extension of the symbolic mapping procedure originally proposed for permutation entropy. We build a Markov chain of the dynamics based on order patterns in the time series which we call an ordinal network, and from this model compute an intuitive entropic measure of transitional complexity. A discussion of the model parameter space in terms of traditional time delay embedding provides a theoretical basis for our multiscale approach. As an ancillary discussion, we address the practical issue of node aliasing and how this effects ordinal network models of continuous systems from discrete time sampled data, such as interbeat interval time series.
I. INTRODUCTION

Methods of ordinal time series analysis [1] seek to quantify, classify or construct models by mapping a measured signal to a set of symbols based on order patterns from the relative magnitude of observed values in short data segments of fixed length. This is distinct from traditional methods of nonlinear time series analysis which generally make use of absolute values in reconstructed phase space [2] or symbolic dynamics from binned data. Since its inception with the pioneering work of Bandt and Pompe [3], much of the literature pertaining to applied ordinal analysis has focused on the investigation of signals measured from complex biological systems. We refer readers to papers [1, 4] for a thorough review of these applications. More recent contributions in the field include ordinal based analysis of electrocardiograms and cardiac interbeat intervals [5–7], and electroencephalogram analysis for seizure onset detection, sleep staging, and anaesthesiology [6, 8].

Generally speaking, the theory underpinning the method is thus far incomplete, however, several key properties of ordinal analysis give merit to continuing work in the field. The first is that an ordinal partition — the partition which defines the map between the observed data and the symbolic dynamics of interest — has the generating property under certain conditions [9, 10], which implies topological conjugacy between phase space and the ordinal symbolic dynamics. Secondly, it has been shown with respect to discrete maps, that the distribution of ordinal symbols enumerated from a time series is governed by the dynamical properties of the underlying system [11]. The vast majority of the existing research essentially revolves around quantifying some aspect of the stationary distribution of ordinal patterns, and whilst it is true that each ordinal pattern contains information about the short range temporal structure of adjacent observations [1], questions pertaining to the relationship between the sequence of patterns encoded in the ordinal symbolic dynamics and the characteristic behaviour of the underlying system remain largely unexplored. However, recently a new measure denoted as the conditional entropy of ordinal patterns was introduced by Unakafov et. al. to quantify the transitional complexity of ordinal symbolic dynamics via the computation of entropy for successive patterns [12]. Independently of said study, we recently published preliminary work proposing and investigating the concept of ordinal networks [13] whereby we incorporated the theoretical and practical benefits of ordinal analysis into a framework for time series analysis using complex networks [14, 15], which
in turn enabled the quantification of aspects of the temporal structure of ordinal symbolic dynamics using simple measures from the field of network science.

In this study we refine our definition of ordinal networks, explain the nature of an ordinal partition under the assumption of determinism with regard to reconstructed phase space, and discuss the implications for parameter selection and practical application of the method (Section II). The latter discussion includes commentary on the phenomenon that we call node aliasing which can significantly effect measures of the connectivity or transitional structure on the networks, but can be minimised by suitably interpolating the data. Furthermore, we present an argument for ordinal networks as a generalised model for ordinal analysis and show that permutation entropy can be interpreted as the Shannon entropy of the stationary distribution of the Markov chain given by the network’s adjacency matrix. The central ideas we introduce in this work are two novel information theoretic measures of transitional complexity which we call local node out-link entropy and global node out-link entropy, and an accompanying scheme for multiscale time series analysis.

To validate our theoretical discussion we have undertaken a comparative test of permutation entropy, conditional entropy of ordinal patterns and global node out-link entropy using numerically generated time series from the Rössler system (Section IV). We then apply our novel measure in a multiscale ordinal network analysis for discrimination between short time electrocardiogram recordings characterised by normal sinus rhythm, ventricular tachycardia and ventricular fibrillation respectively (Section V), and for the investigation of age-related effects in cardiac dynamics using interbeat interval time series from long time electrocardiograms (Section V).

II. ORDINAL NETWORKS FROM TIME SERIES

A. Ordinal networks defined

In this section we state the procedure for mapping a time series to an ordinal network. Refer to Figure 1 for a simple illustration of the method which is intended to accompany the description that follows. We also discuss parameter selection in light of a phase space interpretation of ordinal symbols.

Consider a series of equi-spaced observations \( \{x_n\}_{n=1}^N \). To generate an ordinal network
FIG. 1: An illustration of the process for transforming a univariate time series into an ordinal network: (a) Take the first embedding vector with lag \( \tau \) and dimension \( m \), (b) find its ordinal pattern which is the amplitude rank of its elements, (c) repeat the process for all embedding vectors from the time series, and (d) define a network \( A \) with a node for each unique ordinal pattern, and edges assigned between nodes based on the temporal succession of ordinal patterns in the time series.

corresponding to \( x \) we first partition the data based on ordinal patterns of length \( m \) and time lag \( \tau \). To do this we take the series of embedding vectors \( \{z_n\}_{n=1}^{N-(m-1)\tau+1} \) where \( z_n = (x_n, x_{n+\tau}, x_{n+2\tau}, ..., x_{n+(m-1)\tau}) \). Each vector \( z_n \) is then mapped to a symbol \( (\pi_1, \pi_2, ..., \pi_m) \) where \( \pi_k \in \{1, 2, ..., m\} \) and \( \pi_k \neq \pi_l \iff k \neq l \) such that \( \pi_k < \pi_l \iff x_k > x_l, \forall x_k, x_l \in z_n \). In the case where two elements of \( z_n \) are equal in value, rank is assigned based on order of appearance in the vector, that is to say \( \pi_k < \pi_l \iff k < l \) for \( x_k = x_l \). We call \( s \) the unique set of all the ordinal symbols present in the time series. Thus far we have applied an ordinal partition to \( \{x_n\}_{n=1}^N \) and can write the symbol sequence \( S = \{s_n\}_{n=1}^{N-(m-1)\tau+1} \) for \( s_n \in s \) corresponding to the time series.

The next step is to construct a network \( G = \{V, E\} \) where \( V = |V| \) and \( E \) is the set of edges. We do this by applying the bijective operator \( \Gamma : s \to V \), hence each node represents an ordinal symbol from \( s \). We represent this network by the \( V \times V \) adjacency matrix \( A \) where an element \( a_{i,j} > 0 \) stipulates and edge between
node $i$ and node $j$. The weight of an edge is given by:

\[ a_{i,j} = |\{n \in \{1, 2, \ldots, (N-(m-1)\tau)\} : (\Gamma(s_n), \Gamma(s_{n+1})) = (i, j)\}|, \]

given $s_n \in s$ and $i, j \in V$, which is to say we place a directed edge from node $i$ to node $j$ with weight equal to the number of times this transition occurs in the symbol sequence $S$. The resulting ordinal network is therefore a Markov chain of the dynamics. It is important to note that the definition for an ordinal network presented here differs marginally from the one given in our previous work [13] in that here we retain self loops in $A$.

**B. Parameter selection**

As is widely understood in the field of nonlinear time series analysis, the selection of model parameters can have a significant effect on results and therefore must be treated with care [2]. In their seminal paper defining ordinal partitions, Bandt and Pompe asserted that a partition should arise “naturally” from the data with no additional model assumptions [3]. However, an ordinal partition is parameterised by an embedding dimension $m$ and an embedding lag $\tau$. Whilst it is advantageous that these parameters alone are sufficient to fully describe the partition, appropriate selection of $m$ and $\tau$ necessarily requires model assumptions. If it is assumed that some component of the time series dynamics is deterministic then a phase space interpretation of the ordinal partition, as first discussed by Groth [16], is arguably the most valid approach.

To apply an ordinal partition to a time series is to define a symbolic mapping where regions of the reconstructed phase space are uniquely mapped to the set of ordinal symbols. While these regions are strictly unbounded, they make up a set of disjoint subspaces that spans the complete phase space. The boundaries between the regions are given by the \( \binom{m}{2} \) inequalities between the dimensions of the reconstructed phase space that arise from the ordinal symbols [13, 16]. For example, the ordinal symbol $\{1, 2, 3\}$ corresponds to points embedded in $\mathbb{R}^3$ in the region with boundaries $x = y$, $x = z$ and $y = z$, and which satisfies $\{x > y > z\}$.

It should now be apparent that there exists a single fixed ordinal partition in $\mathbb{R}^m$ for each $m \in \mathbb{Z}^+$. Increasing $m$ serves to embed $\{x_n\}_{n=1}^N$ over these fixed partitions in successively higher dimensions thereby allowing trajectories to unfold over an increasing number of pos-
sible ordinal subspaces. In our previous work we noted that for ordinal network analysis it is important to select $m$ large enough such that different phase points are only mapped to the same symbol if they are genuine neighbours or else the resulting network will contain degeneracies [13]. While it may be possible to select $m$ for ordinal analysis using the traditional method of computing of false nearest neighbours [1, 17], the Euclidean metric used in this computation is not consistent with the topology of and ordinal partition. Rather, we opt for a scheme more akin to the asymptotic invariant method [17], which is to compute some measure on the ordinal network and observe how it changes with the embedding dimension. We shall discuss this approach in the context of the applications addressed in this study.

Embedding lag $\tau$ controls to what extent the dynamics are unfolded in $\mathbb{R}^m$. The general consensus in embedology is that $\tau$ should be selected such that the elements of the embedding vectors are as uncorrelated as possible to optimally unfold the attractor. In the context of ordinal analysis, an optimally unfolded attractor will occupy a maximal number of states over the partition [16] thereby maximising the useful dynamical information captured by the ordinal symbolic dynamics. This manifests as a relationship between measures on the ordinal symbolic dynamics and the intrinsic time scale(s) present in data when one considers a sensible parameter range [18, 19]. Furthermore, where data comprises both deterministic and stochastic components, selecting $\tau$ on an appropriate time scale can maximise the contribution of determinism relative the stochastic component in the resulting the ordinal symbolic dynamics [20]. Of particular relevance in to this study are the findings presented in [21] where short ordinal symbols were used as bio-markers in the study of human cardiac dynamics, and it was shown that the symbols were most effective as statistical classifiers when the embedding lag was set to a characteristic frequency traditionally used in heart rate variability analysis.

By extension of these findings, the relationship between ordinal measures and $\tau$ affords the possibility for multiscale analysis. This is achieved by computing ordinal measures from a single time series using several different lags to capture dynamics operating on different time scales. Such a scheme was applied to EEG signals in [8] and for the analysis of both EEG and HRV data in [6]. We implement a similar multiscale approach in our investigation of ECG data in Section V.
C. Node aliasing

The phenomenon which we call node aliasing is a type of artefact that can occur in ordinal networks when the sampling rate of the time series is too low relative to the rate of evolution of the trajectories and the length span of the elements of the ordinal partition in reconstructed phase space in the direction tangential to the flow. When node aliasing occurs, it falsely implies local uncertainty in the Markov chain which can give rise to misleading results.

For example, consider the trivial case of a continuous noiseless periodic time series oscillating with frequency \( \omega_0 \) and sampled at \( \omega \neq b\omega_0, \forall b \in \mathbb{N} \). It follows that there can exist an embedding vector \( z_k \) which lies on an intermediate point of the trajectory between any two pairs of temporally adjacent states \( z_n \) and \( z_{n+1} \). Since the edges of an ordinal network are allocated based on temporal succession, it is possible that the connectivity patterns in the resulting network may not be consistent with the true nature of the dynamics of the time series. This depends, of course, on the particular choice of embedding parameters, the topology of the trajectories and the sampling rate of the data. One example of an undesirable mapping would be that \( z_n \rightarrow s_1, z_{n+1} \rightarrow s_3, z_{k-1} \rightarrow s_1 \) and \( z_k \rightarrow s_2 \). The subgraph for nodes \( \{1, 2, 3\} \) is therefore:

\[
\begin{pmatrix}
1 & 2 & 3 \\
1 & 0 & 1 & 1 \\
2 & 0 & 0 & 0 \\
3 & 0 & 0 & 0
\end{pmatrix}
\]

We call this an instance of node aliasing because from node \( \{1\} \) is impossible to distinguish between the possible destination nodes \( \{2, 3\} \). To correctly characterise the periodic dynamics, which contain no such uncertainties, the correct subgraph should clearly be:

\[
\begin{pmatrix}
1 & 2 & 3 \\
1 & 0 & 1 & 0 \\
2 & 0 & 0 & 1 \\
3 & 0 & 0 & 0
\end{pmatrix}
\]

This can be obtained if our data is sufficiently over-sampled, but this creates two further problems. The first is that in many applications we do not have the luxury of over-sampled
data, however, we show that simple interpolation of the data can suffice for the case of continuous chaotic time series in Section IV. The second problem is that increasing the sampling rate will increase the number of self loops relative to the number of incoming and outgoing edges which distort certain local measures of transitional complexity. We address this issue in Section III C with our novel measure local node out-link entropy.

III. QUANTIFYING THE COMPLEXITY OF ORDINAL NETWORKS

A. Permutation entropy

Bandt and Pompe’s definition for the permutation entropy of a time series is the Shannon entropy of the corresponding set of ordinal symbols $s$:

$$h_{PE} = - \sum_i p_i \log p_i,$$

where the probability mass function $P(S = s_i) = p_i$ for $s_i \in s$ is estimated by counting the relative occurrence of each symbol in the symbolic dynamics $S$ [3]. Note that in this study we compute logarithms with base 2, therefore all entropy measures have units of bits. Perhaps the most interesting property of permutation entropy from the theoretical perspective is that it is an upper bound for Kolmogorov-Sinai entropy for piecewise monotone interval maps and directly corresponds under certain conditions [9, 10, 22].

To demonstrate the connection between permutation entropy and ordinal networks consider now that the probability mass function is the stationary distribution of the Markov chain of $S$. The stationary distribution can be computed by finding the stochastic matrix $P$ with elements:

$$p_{i,j} = \frac{a_{i,j}}{\sum_j a_{i,j}},$$

where $a_{i,j}$ are the elements of the adjacency matrix $A$ of the ordinal network, and then calculating vector $q$ with elements $q_i$, which is the left eigenvector of $P$ corresponding to the eigenvalue $\lambda = 1$. By normalising we obtain the probability mass function:

$$p_i = \frac{q_{i,\lambda=1}}{\sum_k q_{k,\lambda=1}}.$$

Readers familiar with complex network theory will recognise this as the is the eigenvector centrality of the network [23]. It is also possible to estimate the stationary distribution from
A by taking:

\[ p_i \approx \frac{\sum_j a_{i,j}}{\sum_k \sum_j a_{k,j}}, \tag{4} \]

where the sum of the edge weights for edges leaving node \( i \) is divided by the sum of edge weights over the entire network. This is a very close to but not exactly equivalent to counting the occurrences of ordinal symbols in \( S \) due to the fact that equation (4) counts occurrences based on the transition from one symbol to the next, and therefore does not count the final symbol in the sequence \( s_{n-m+1} \).

**B. Measures of transitional complexity**

An inherent property of any measure of the stationary distribution of a Markov chain, including permutation entropy, is that for an appropriate choice of partition the result will be the same for both periodic dynamics and uniformly distributed noise, and this may be undesirable in certain applications. Permutation entropy will only be minimised for fixed points and strictly monotonic signals. On the other hand, measures of transitional complexity quantify the local uncertainty of each state in the model and can therefore reveal additional information about the dynamics. For example, Froyland et. al. build Markov chains using a regular grid partition of phase space and, and by taking the Shannon entropy for each box, compute what they call *finite-time entropy* which measures the nonlinear growth of a small set of initial conditions as the system evolves \cite{Froyland2005}. This measure is shown to be useful for the detection of stable manifolds and local transport barriers as well as being closely related to finite-time Lyapunov exponents. Related measures of transitional complexity based on node centrality have been used to quantify divergence of trajectories in chaotic attractors \cite{Schreiber2000} and as an index of predictability for discrete processes \cite{Solero2019}.

Recently Unakafov et. al. proposed a direct and logical extension to permutation entropy which they called the conditional entropy of ordinal patterns, henceforth referred to in this paper as conditional permutation entropy. It has been shown analytically and numerically that conditional permutation entropy converges on Kolmogorov-Sinai entropy more quickly than permutation entropy \cite{Unakafov2015} and was applied to electroencephalograms for the classification of healthy subjects from those with epilepsy in both a normal state and during a
seizure \[8\]. Conditional permutation as defined in \[12\] is equivalent to:

\[ h_{CPE} = \sum_{i} \left( -p_{i} \sum_{j} p_{i,j} \log p_{i,j} \right) \]  

(5)

where \( p_{i} \) as per equation (1) and \( p_{i,j} \) is probability of a transition from \( s_{i} \) to \( s_{j} \) as estimated from \( S \). Conditional permutation entropy is therefore the expected value of the entropy of node \( i \) as averaged over the stationary distribution of the ordinal network where we are referring to the concept of node entropy as defined in \[25\] which is the inner summation in equation (5). For completeness we note that conditional permutation entropy can also be described as the global average of finite-time entropy \[24\] computed over an ordinal partition rather than a regular grid partition.

C. Local and global node out-link entropy of ordinal networks

Here we present the new measures central to this study. Consider a modified stochastic matrix \( P^{T} \) of an ordinal network that excludes the possibility of self loops and hence has elements:

\[ p_{i,j}^{T} = \begin{cases} 
0 & \text{if } i = j \\
\frac{a_{i,j}}{\sum_{j \neq i} a_{i,j}} & \text{if } i \neq j
\end{cases} \]  

(6)

Taking the Shannon entropy of row \( i \) gives the local node out-link entropy for node \( i \):

\[ h_{i}^{LNE} = -\sum_{j} p_{i,j}^{T} \log p_{i,j}^{T}, \]  

(7)

It is possible that the final node in the symbolic dynamics \( S \) has no outgoing edges in the ordinal network. If the final node has no out connections then we set \( h_{i}^{LNE} = 0 \) by definition to avoid singularities. The ordinal network mapping procedure guarantees that all other nodes will have nonzero out degree. By averaging over the network based on the stationary distribution as estimated by equation (4) we obtain the expected value of the transitional complexity of \( S \) which we call global node out-link entropy:

\[ h_{GNE} = \sum_{i} p_{i} h_{i}^{LNE}. \]  

(8)

The lower bound \( h_{GNE} = 0 \) implies that there is no uncertainty in the system. This corresponds to time series that are strictly monotonic or a symbolic dynamics \( S \) that is strictly
periodic. Note that periodic time series will only produce $h^{\text{GNE}} = 0$ if the ordinal map is free from degeneracies and aliasing effects. The maximum attainable entropy for node $i$ with degree $k_i$ is bounded at $h^{\text{LNE}}_i \leq \log(k_i)$. This corresponds to $p^T_{i,j} = 1/k_i$ for all out-connected edges and, hence, maximum uncertainty with respect to node degree. We choose not to normalise $h^{\text{LNE}}_i$ based on node degree so as to preserve the absolute amount of information generated by each node before we take the expected value in equation (8).

To derive an upper bound for $h^{\text{LNE}}_i$ which is independent of network topology, note that $V \leq m!$ which implies $k_i \leq (m! - 1)$ when self loops are removed. Maximum uncertainty for node $i$ arises when it is fully connected with $p^T_{i,j} = 1/(m! - 1)$ for all $j : j \neq i$, and therefore $h^{\text{LNE}}_i \leq \log(m! - 1)$. By extension, an ordinal network will have maximum uncertainty when it is fully connected with $V = m!$ and $p_i = 1/m!$ for all $i$. Assuming these conditions and substituting the upper bound for $h^{\text{LNE}}_i$ into equation 8, the global node out-link entropy is also bounded at $h^{\text{GNE}} \leq \log(m! - 1)$. This corresponds to an infinitely long time series of uniform IID noise. Note that the upper bounds for permutation entropy and conditional permutation entropy can be derived via similar reasoning and are both equal to $\log(m!)$[3, 8]. Upper bounds will be provided for reference in results from Sections IV and V.

The critical point of difference between our new measure $h^{\text{GNE}}$ and conditional permutation entropy is that we exclude self loops. In the case of a discrete sampled continuous dynamical system, the weights of the self loops give the rate of evolution through each node in the ordinal network. Specifically, for a given node, dividing the weight of the self loop by the total weight of the outgoing edges excluding the self loop gives the average time that the system spends in that node each time it is visited, with the respect to the sampling period of the data. It follows that by increasing the sampling frequency the weights of the self loops will increase. However, the weights of the outgoing edges should remain more or less the same if the base sampling frequency is high enough to minimise node aliasing. This relative change will skew the probability mass function for the transitional probabilities of a node in favour of the self loop and therefore cause conditional permutation entropy to converge to $h^{\text{CPE}} = 0$ as the sampling frequency tends to infinity. Using local and global node out-link entropy as computed from the modified stochastic matrix $P^T$ ensures that the resulting complexity measure is not sensitive to the sampling frequency because the we are separating information pertaining to transitional complexity from the rate of evolution and consider only the former.
IV. APPLICATION TO THE RÖSSLER SYSTEM

To demonstrate the effectiveness of global node out-link entropy $h_{GNE}$ in comparison with permutation entropy $h_{PE}$ and conditional permutation entropy $h_{CPE}$, we use numerically generated time series from a continuous chaotic Rössler system which is governed by the following set of equations:

\[
\begin{align*}
\frac{dx}{dt} &= -y - z \\
\frac{dy}{dt} &= x + \alpha y \\
\frac{dz}{dt} &= \beta + z(x - \gamma)
\end{align*}
\]  

(9)

where $\beta = 2$, $\gamma = 4$ and control parameter $\alpha \in [0.36, 0.43]$. Numerical solutions were computed using a fourth-fifth order Runge-Kutta algorithm with randomised initial conditions $\{x_0, y_0, z_0\} \in (0, 1)$. We take a highly sampled time series of the $x$-component with transients removed and fixed data length sufficiently long such there are a total of 160 cycles when the system is in the period-2 regime. To find a suitable embedding lag we take the first zero of the autocorrelation function which is $\tau = 8$ when the data is down-sampled to a sampling period of 0.2 time units. In the following analysis we use this embedding lag and scale it directly with the sampling frequency as necessary to maintain an equivalent embedding. To find an suitable embedding dimension we compute $h_{GNE}$ for a selection of periodic and chaotic time series for a range of $m$ as shown in Figure 2. We set $m = 14$ because it is the smallest embedding dimension for which $h_{GNE}$ begins to exhibit a degree of invariance for most of the time series, which is convergence to zero for periodic dynamics or convergence to a relatively stable constant for chaos.

Figure 3 shows the effect of oversampling on $h_{GNE}$ and $h_{CPE}$. In the case of $h_{GNE}$, increasing the sampling frequency reduces node aliasing, and values for periodic and chaotic dynamics converge to zero and fixed non-zero values respectively as the number of alias edges in the network becomes small. The only exception is the period-4 time series which converges to a value close to zero. This is because the ordinal partition is unable to separate the independent trajectories in this time series that are very close in reconstructed phase space as the system has just undergone a period doubling bifurcation. On the other hand, the measured value of $h_{CPE}$ is strongly related to the sampling frequency and converges to zero for oversampled time series due to the dominance of self loops in the ordinal network as
FIG. 2: Global node entropy $h^{GNE}$ plotted against the embedding dimension $m$ for a selection of time series generated by the Rössler system. The time series were highly sampled with sampling period $1 \times 10^{-3}$ to minimise the effects of node aliasing.

explained in Section III. Figure 3 also demonstrates the important result that data sampled at a lower rate and then interpolated with a simple cubic spline provides values of $h^{GNE}$ which are almost equal to those measured from the true highly sampled dynamics.

All three entropy measures under consideration have been computed from highly sampled time series for 1401 equi-spaced values of $\alpha \in [0.36, 0.43]$ and the results are plotted in Figure 4. A bifurcation plot and the largest Lyapunov exponent of the Rössler system with respect to the bifurcation parameter $\alpha$ are included for reference. Our new measure $h^{GNE}$ tracks the relative change in the largest Lyapunov exponent and detects periodic windows more effectively than both $h^{CPE}$ and $h^{PE}$. Periodic time series result in values of $h^{GNE}$ close to theoretically assured baseline value of zero entropy, and chaotic time series give consistently higher values which would facilitate thresholding for discrimination of dynamics if desired. On the other hand, $h^{CPE}$ and $h^{PE}$ provide arbitrary and inconsistent results for complexity in periodic time series. For example, note the step change in $h^{PE}$ after the first period doubling bifurcation. With regards to $h^{CPE}$ specifically, the total range of $h^{CPE}$ is small in comparison to the other measures which may imply that this measure could be less robust to degeneracies or other artefacts in the ordinal mapping process. Furthermore, $h^{CPE}$ trends down in the chaotic regime for $\alpha > 0.4$ which is inconsistent with the dynamics where the attractor is becoming more chaotic and hence less predictable as quantified by the largest Lyapunov exponent. We acknowledge that for the particular embedding parameters used in this study, $h^{PE}$ follows the positive trend in the largest Lyapunov exponent through the chaotic regimes more effectively than $h^{GNE}$, which remains relatively constant. However, it
FIG. 3: Global node out-link entropy $h^{GNE}$ and conditional permutation entropy $h^{CPE}$ plotted against the sampling period for a selection of time series generated by the Rössler system. The solid curves show results generated by taking highly sampled time series with sampling period $1 \times 10^{-4}$ (the rightmost datapoint on each curve) and then downsampling. The dotted curves show results generated by taking sparsely sampled time series with sampling period $2 \times 10^{-1}$ (the leftmost datapoint on each curve) and interpolating the data with a cubic spline. Maximum possible entropy for both measures at $m = 14$ is approximately 36.343.

is important to recall that $h^{GNE}$ and $h^{PE}$ are measuring fundamentally different properties of the dynamics (see Section III) and we therefore suggest computing both to ascertain which is more useful for a specific application. Finally, we report that the curve for $h^{GNE}$ against $\alpha$ is virtually identical for highly sampled data and for down-sampled data that has been re-interpolated back to the original sampling frequency (not shown) — a result that gives merit to our new measure despite the impracticality of obtaining significantly over-sampled data from experimental systems because interpolation can be sufficient for minimising node aliasing while also preserving the dynamics.
FIG. 4: The largest Lyapunov exponent $\lambda_1$ of the Rössler system plotted against the bifurcation parameter $\alpha$ and overlayed on the bifurcation diagram (computed from the $x$-component time series using the `lyap_k` and `extrema` functions from the TISEAN software package [26] respectively); and global node out-link entropy $h^{GNE}$, conditional permutation entropy $h^{CPE}$ and permutation entropy $h^{PE}$ computed using the $x$-component time series and plotted against the bifurcation parameter $\alpha$. These results are for highly sampled time series with sampling period $1 \times 10^{-3}$. Maximum possible entropy for all measures at $m = 14$ is approximately $36.343$. 
V. APPLICATION TO ELECTROCARDIOGRAMS

A. Classifying cardiac dynamics using ordinal network entropy

1. Description of data and methodology

In this section we apply ordinal network analysis, specifically the computation of global node out-link entropy, to a dataset of human electrocardiograms (ECGs) recorded at the Royal Infirmary of Edinburgh’s coronary care unit. Henceforth we refer to this as the RIE:CCU dataset. The dataset comprises 81 ECGs each 10000 points in length that have been sampled at 500 Hz with 10 bits resolution. The records come from 13 different patients and were originally measured to observe the evolution in cardiac dynamics from normal sinus rhythm (NSR) through to ventricular tachycardia (VT), and finally to ventricular fibrillation (VF). There are 31 records of NSR, 30 records of VT and 20 records of VF. For a thorough description of the data see [27]. In this study we use this dataset as a simple test of the ordinal network method for the classification of different dynamical behaviour. We do not apply any preprocessing to the data. Given the vast differences in the dynamics of ECG signals characterised by NSR, VT and VF, a method such as ours should provide statistically significant discrimination of the three pathological groups. The global node out-link entropy $h_{GNE}$ of each time series is computed for both a short and a long time embedding lag and the resulting two dimensional vector constitutes a measure of multiscale complexity.

To select the embedding lag $\tau$ we begin with the very simple assumption that the mean resting heart rate is 80 beats per minute, which equates to 375 samples per cycle. The short embedding lag is set at $\tau = 20$ which is, very approximately speaking, on the order of a quarter period of any of the individual PQRST components of the complete cycle. Hence, we have selected $\tau$ to unfold the individual wave components over the ordinal partition in reconstructed phase space, or in other words, each symbol captures the detailed shape of ECG wave in segments of approximately 0.2 to 0.6 cycles for $5 \leq m \leq 10$. The long embedding lag is chosen arbitrarily to be one order of magnitude larger at $\tau = 200$ which will capture dynamics over segments of approximately 2 to 6 cycles for $5 \leq m \leq 10$ and therefore encode information pertaining to the inter-cycle variability of the ECG signal in the ordinal network.
2. Results and discussion

The two dimensional multiscale plot of $h^{GNE}$ for $m = 8$ and the corresponding box plot are provided in Figure 5 showing clear three way discrimination between the pathological groups. While these results are as expected in terms of relative classification, the actual values of $h^{GNE}$ for each group raise questions about what aspect of complexity is being captured from the dynamics. For example, it may appear counter intuitive that an ECG record for VF is apparently less complex than for normal sinus rhythm. This should not be considered a flaw in the results but rather a limitation of the method with respect to this particular data set. It is necessary to recognise that what is being measured here is the transitional complexity of the ordinal network, which for this experimental investigation has been built from relatively short time series and therefore may not be a complete or accurate model of the underlying system. The dynamics of a heart attack may be highly unpredictable but, for specific parameters and finite data, may embed to produce an attractor of sparsely distributed independent trajectories and a corresponding ordinal network with low $h^{GNE}$.
(if it is appropriate to embed such a time series in the first place, theoretically speaking, given that the dynamics are likely to be predominantly stochastic). On the other hand, the prospect of inferring characteristics of a real system becomes more reasonable when long time series from controlled experimental conditions are available, as is the case for the data investigated in Section V B.

B. Heart rate variability analysis for different age groups using ordinal network entropy

1. Description of data and methodology

The second application of our ordinal network method which we include in this work is the investigation of age-related effects in interbeat interval dynamics from ECGs. We have used data from the Fantasia database which is publicly available on PhysioNet [28] and was original recorded for, and studied in [29]. The authors of this paper attempted to produce high quality ECG recordings and isolate age as the only experimental variable to the extent that it was practical to do so, such that they could investigate differences in fractal scaling between age groups via detrend fluctuation analysis. The data comprises long time ECG recordings taken under controlled conditions from 20 healthy young subjects aged 21 to 43, and 20 elderly subjects aged 68 to 85. The elderly subjects were rigorously screened for medical conditions and subsequently also deemed to be healthy. There are equal numbers of females and males in each group. The records range in length from 1558559 to 234258 observations and are sampled at 250Hz. The interbeat interval time series were computed from the ECG using automated beat annotations that had been confirmed by visual inspection. The reader should refer to [29] for a detailed description of the data.

Interbeat interval data poses two problems for ordinal analysis. The first is that equal values in the embedding vectors are encountered frequently due a low sampling rate. Some studies suggest the addition of small amounts of white noise to break tied values [18, 30], however, this approach treats ties inconsistently and the resulting symbolic dynamics will therefore contain an additional stochastic component. A more comprehensive method was proposed by Bian et. al. whereby ties were permitted in the ordinal symbols [31] (also implemented in [5]). The primary drawback with this approach is that symbols define subspaces
with different dimensions. For example, the symbol \{1, 2, 3\} defines a three dimensional volume whereas \{1, 2, 2\} defines a two dimensional plane. It is for this reason that we elect to follow the standard convention for dealing with tied elements by assigning rank based on temporal order and seek an alternative solution for the problem of frequent tied values encountered in this application.

The second problem is that an interbeat interval time series can be considered as a sparsely and unevenly sampled record of the instantaneous period of the signal which will cause unpredictable node aliasing and degrade the network model of the dynamics. We address both of these issues by taking the original time series \{y(n)\}_{n=1}^{N} with elements that are the time indices of each R wave in sequence, and transforming it into the new time series \{\hat{y}(\hat{n})\}_{n(1)}^{y(N)} , with elements:

\[
\hat{y}(y(n)) = \frac{1}{y(n+1) - y(n)}, \text{ for } n \in [1, N].
\] (10)

To estimate \(\hat{y}(\hat{n})\) for all other values of \(\hat{n} \in \{y(1), (y(1) + 1), (y(1) + 2), ..., y(N)\}\) that are not defined in equation (10) we use cubic spline interpolation so that the resulting sample rate is the same as the original ECG at 250Hz. The transformed time series \(\hat{y}(\hat{n})\) serves an approximation of the continuous heart rate signal with equi-spaced observations at a high sampling rate, thereby significantly reducing both the frequency of tied values and node aliasing. Since it makes sense to consider embedding lag in terms of number of heart beats, as given by the observed interbeat interval data, we define a new lag variable:

\[
\hat{\tau} = \frac{1}{N-1} \sum_{n=1}^{N-1} y(n+1) - y(n),
\] (11)

hence, our fundamental embedding lag is the mean period of one heart beat.

For multiscale analysis with global node out-link entropy \(h_{GNE}^{\tau}\) we use \(\hat{\tau} = 1\) as the short embedding lag to capture information about sequences of temporally adjacent beats. The long time scale embedding lag is set arbitrarily at \(\hat{\tau} = 10\) to be one order of magnitude larger. Figure 6 shows that the growth and decay of global node out-link entropy \(h_{GNE}^{\tau}\) is consistent between the age groups and has a clear maximum for both short and long \(\hat{\tau}\) respectively. It is therefore a simple task to select the embedding dimension corresponding to the peak values of \(h_{GNE}^{\tau}\) to maximise the transitional information encoded in the ordinal network, hence we have \(m = 7\) for the \(\hat{\tau} = 1\) and \(m = 6\) for the \(\hat{\tau} = 10\).
FIG. 6: Global node out-link entropy $h^{GNE}$ for the interbeat interval data from the Physionet Fantasia ECG dataset plotted against the embedding dimension $m$ for equivalent embedding lag (a) $\hat{\tau} = 1$, and (b) $\hat{\tau} = 10$ (see equation (10) for the definition of $\hat{\tau}$). Each datum is the result for a single time series at each value of $m$.

2. Results and discussion

It has been shown previously that standard permutation entropy is unable to discriminate between the age groups based on the interbeat interval time series from the Fantasia database [31]. We report the same result and, furthermore, that conditional permutation entropy does not effectively discriminate based on the interbeat interval time series nor with the approximation of the continuous heart rate (figures omitted for brevity). On the other hand, Figure 7a shows distinct clustering of the two age groups using $h^{GNE}$. Despite small overlap between the clusters, results from a one way ANOVA test confirm statistically significant discrimination of the means of each group ($p$-value $< 5 \times 10^{-04}$).

The original investigation of the Fantasia interbeat interval data by Iyengar et al. found characteristic differences in the fractal scaling exponents of the time series based on subject age [29]. It was shown that the fractal scaling exponent for records from elderly subjects comprised two distinct scaling regions with a bisection at approximately 40 beats. A two dimensional plot of the short and long range correlations as quantified by the fractal scaling exponent demonstrated that records from elderly subjects generally had higher short range correlation and lower long range correlation. The authors also reported that the variability
in the fractal scaling exponent was significantly higher in the elderly subject group. Our choice of parameters results in ordinal symbols which span approximately 7 beats and 60 beats for short and long time scales respectively, hence straddling the formerly reported division of time scales. In turn, Figure 7a shows that the transitional complexity of the ordinal networks, as quantified by our measure $h^{GNE}$, is generally higher for elderly subjects on short time scales, lower on long times scales, and has significantly greater variability than can be observed for young subjects, just as per the fractal scaling exponent as reported in the original study. While we refrain from any claim that $h^{GNE}$ is somehow equivalent to the fractal scaling exponent, it is by no means implausible that fractal correlation could manifest as complex transitional structures in ordinal networks. It is therefore reasonable to conclude that the results of our analysis support the findings from [29].

FIG. 7: (a) Multiscale scatter plot of global node out-link entropy $h^{GNE}$ for the interbeat interval data from the Physionet Fantasia ECG dataset. The x and y axes correspond to $h^{GNE}$ computed with equivalent embedding lag $\hat{\tau} = 1$ and $\hat{\tau} = 10$ respectively (see equation (10) for the definition of $\hat{\tau}$). Each datum is the result for a single time series. Maximum possible $h^{GNE}$ is approximately 9.4898 and 12.299 for $m = 6$ and $m = 7$ respectively. (b) Box plot for a 1D projection of the multiscale scatter plot onto the vector between the means of each subject group.
VI. CONCLUSIONS

In this paper we have introduced local and global node out-link entropy of ordinal networks as new measures to quantify the complexity of temporal structure in ordinal symbolic dynamics from time series. By assuming a deterministic component in the data and building on the existing literature, we have established a framework for conceptualising the ordinal mapping procedure in terms of reconstructed phase space. We also presented a case for ordinal networks as a general model for ordinal analysis, specifically with respect to the widely used permutation entropy measure, which we have shown is the stationary distribution of the Markov chain defined by the ordinal network. Moreover, we note here that the information required for the majority of approaches listed in Section I could also be extracted from these networks. Our comparative investigation using continuous chaotic time series from the Rössler system has demonstrated that global node out-link entropy tracks dynamical change through period doubling and periodic windows over a range of the bifurcation parameter. Qualitative assessment of the results confirm that our measure approaches a theoretical and intuitive lower bound of zero for periodic dynamics which makes it arguably more useful than the existing and closely related measure of conditional permutation entropy which exhibits a sensitivity to the sampling rate, and complimentary if not more effective than traditional permutation entropy for certain types of data, given that a fundamentally different property of the dynamics is being captured.

Considering previous studies relating to the relationships between embedding lag and ordinal analysis metrics, we proposed a simple scheme for multiscale analysis whereby we compute global node out-link entropy for different time scales by varying the lag parameter. In application, our method proved effective for discriminating between short time electrocardiogram recordings of normal sinus rhythm, ventricular tachycardia and ventricular fibrillation, and served to characterise differences in the transitional complexity of ordinal networks from interbeat interval time series on short and long time scales with respect to subject age. In the latter of these experimental investigations, we found that global node out-link entropy for healthy elderly subjects was higher on short time scales and lower for long time scales. This may suggest that normal ageing processes cause a change in the complexity of interbeat interval dynamics. We stress here the important distinction between measures of the model built from data and reality itself. Complexity in the ordinal network
may not always be a true reflection of dynamical complexity in the underlying system. The amount of data available, quality of the data, and most importantly, the appropriateness of applying an ordinal partition to symbolise the data must all be taken into consideration when interpreting results from ordinal analysis. However, in this particular application the database under examination contained long time records taken under controlled conditions, the assumption of some deterministic process governing the underlying dynamics (heart rate) is reasonable, and our results appear to align the findings from the original study of the data. On the latter point, the seemingly analogous behaviour of global node out-link entropy and the fractal scaling exponent that we observed compels more a thorough investigation in future studies.

The primary limitations of ordinal network analysis are degeneracies in the mapping process and node aliasing. These effects are evident when comparing the curves for global node out-link entropy with the largest Lyapunov exponent, the latter of which still produces superior results for well behaved noiseless data. Degeneracies can be minimised to some extent by choosing sufficiently high embedding dimensions, however this will make the ordinal patterns less robust to noise, increase the detrimental effect of node aliasing and can make computation impractical. The topology of an ordinal partition is simply not suitable for all types of data. With regard to node aliasing, we have shown for both toy data and experimental data that the number of alias links in the network can be reduced using interpolation to obtain oversampled data and vastly improve results, but this solution is not universally applicable either. On the other hand, it is important to recognise that while invariant measures such as Lyapunov exponents and Kolmogorov-Sinai entropy (recall from prior discussion that permutation entropy and conditional permutation entropy are related to this invariant [9, 10, 12, 22]) can theoretically be used to classify dynamics, in practical applications they are usually employed for change point detection or relative discrimination due to the practical difficulties associated with estimation from experimental data [2]. Taking this into consideration, although due diligence should be observed when applying ordinal analysis with respect to our best current understanding of the theory and practical pitfalls, these methods serve as a form of data compression, filtering the data and retaining information of interest [1]. The findings of this study have demonstrated that our novel method of ordinal network analysis can yield useful results in investigations of complex biological systems based on univariate time series data.
ACKNOWLEDGMENTS

MS is supported by the Australian Research Council Discovery Project DP 140100203.


