A complex network based measure for the analysis of time series data

K. P. Harikrishnan* and Rinku Jacob†
Department of Physics, The Cochin College, Cochin-682 002, India

R. Misra‡
Inter University Centre for Astronomy and Astrophysics, Pune-411 007, India

G. Ambika§
Indian Institute of Science Education and Research, Pune-411 008, India

Methods of nonlinear time series analysis require that the information regarding the underlying system has to be obtained from the scalar time series of a single variable. For this, the generally accepted procedure is to construct a higher dimensional system of dimension M from the time series using the delay embedding method with a suitable time delay τ. For proper computation of nonlinear measures, such as, dimension, entropy and Lyapunov exponent, one has to use an embedding dimension greater than the actual dimension of the system. For real world data, the information regarding the dimension of the system is absent. The conventional nonlinear time series analysis uses the method of false nearest neighbor or saturated correlation dimension to get information regarding the dimension from an arbitrary time series.

Recently, complex network based measures are increasingly being applied in nonlinear time series analysis. The advantage of these measures is that they are effective even for short and non stationary time series where conventional methods are found to be inaccurate. Here also, the proper embedding dimension is essential for the accurate computation of the network measures. A network based measure to determine the underlying dimension from a time series has not been proposed so far. Here we propose such a measure and show that it is efficient even for short time series with number of data points $N < 5000$. The measure is based on the well known Kullback-Leibler (K-L) divergence used to differentiate between two probability distributions. Our method is to compute the probability distributions of the local clustering coefficients of the complex networks generated from the time series for increasing embedding dimensions M. The K-L measure is then computed for successive M values and the value of M for which the measure saturates is taken as the dimension of the system. The method is illustrated using a synthetic time series and a real world time series.

The paper is part of the Complex Networks Mini Symposium

*Electronic address: kp_hk2002@yahoo.co.in
†Electronic address: rinku.jacob.vallanat@gmail.com
‡Electronic address: rmisra@iucaa.in
§Electronic address: g.ambika@iiserpune.ac.in