Abstract for GR-TR Conference on Statistical Mechanics and Dynamical Systems

Topic: Pattern Formation Preference: Oral

Hierarchical multifractal representation of symbolic sequences and application to human DNA

A. Provata*, P. Katsaloulis

Institute of Physical Chemistry, National Center for Scientific Research "Demokritos", 15310, Athens, Greece
* Electronic Address: aprovata@limnos.chem.demokritos.gr, aprovata@qmul.ac.uk

The 2-D Density Correlation Matrix (DCM) method is described for symbolic sequences of arbitrary size, based on segments of finite size. Viewed as a surface with random like local heights, this DCM can be characterised by its MultiFractal Spectrum (MFS) which indicates the presence of correlations in the corresponding symbolic sequences. This method is applied to symbolic DNA sequences in entire human chromosomes, reconstructed genomic sequences and artificial random sequences. It is shown that all human chromosomes have common characteristics in their MFS and deviate substantially from random and uncorrelated sequences of the same size. The correlations are shown to be crucial for the form of the multifractal spectra; surrogate shuffled chromosomes present random-like spectra, distinctly different from the actual chromosomes.

Hierarchical reconstruction of 2-D DCMs is further undertaken, based on superposition of tensor products and their MFS are analytically obtained. It is shown that retaining pair correlations when constructing the DCM via hierarchical superpositions, leads to a better representation of the genomic multifractal spectra. The selective presence of characteristic functional units of small sizes (such as the CG/GC combinations) is held responsible for the observed correlations and for the corresponding deviations in the spectra. Retaining higher order correlations in the construction of the tensor products is a way to approach closer the MFS of the actual genomic sequences. This hierarchical approach is generic and is applicable to correlated symbolic sequences in general.

---

[1] B.L. Hao,*Physica A*, **282** 225-246 (2000).

[2] A. Provata, Y. Oikonomou,*Phys. Rev. E*, **75**, Art. No. 056102 (2007).

[3] A. Provata, P. Katsaloulis,*Phys. Rev. E*, **81**, Art. No. 026102(2010).