

# WHY DO COMPLEX SYSTEMS LOOK CRITICAL?

**MATTEO MARSILI**

INTERNATIONAL CENTER FOR THEORETICAL PHYSICS (ICTP), TRIESTE

The study of complex systems is limited by the fact that only few variables are accessible for modeling and sampling, which are not necessarily the most relevant ones to explain the systems behavior. In addition, empirical data typically under sample the space of possible states. We study a generic framework where a complex system is seen as a system of many interacting degrees of freedom, which are known only in part, that optimize a given function. We show that the underlying distribution with respect to the known variables has the Boltzmann form, with a temperature that depends on the number of unknown variables. In particular, when the influence of the unknown degrees of freedom on the known variables is not too irregular, the temperature decreases as the number of variables increases. This suggests that models can be predictable only when the number of relevant variables is less than a critical threshold. Concerning sampling, we argue that the information that a sample contains on the behavior of the system is quantified by the entropy of the frequency with which different states occur. This allows us to characterize the properties of "maximally informative samples": Within a simple approximation, the most informative frequency size distributions have power law behavior and Zipf's law emerges at the crossover between the under sampled regime and the regime where the sample contains enough statistics to make inferences about the behavior of the system. These ideas are illustrated in some applications, showing that they can be used to identify relevant variables or to select most informative representations of data, e.g. in data clustering.

MARCH 20

3:30PM

WATANABE 112

(REFRESHMENTS AT 3:15)