Statistical physics and information theory for systems with local constraints

Zhang, Q.

Citation

Zhang, Q. (2021, December 1). *Statistical physics and information theory for systems with local constraints. Casimir PhD Series*. Retrieved from https://hdl.handle.net/1887/3244220

Chapter 6

Conclusions

Statistical physics deals with the description of systems with many interacting microscopic constituents and develops tools to characterize the macroscopic properties emerging in the so-called thermodynamic limit, where the number of constituents goes to infinity. Formally, this is achieved by introducing the concept of statistical ensemble, i.e. a collection of (unobservable) microscopic configurations of the system, each of which is assigned a certain probability that is calculated from certain (observable) macroscopic constraints. Traditionally, the constraints considered in physics are global conserved quantities, such as the total energy and the total number of particles. There are different ways in which the constraints can be enforced, and different resulting ensembles. The two most important possibilities are the microcanonical ensemble, where the constraint is enforced in a hard fashion (i.e. every configuration in the ensemble matches the constraint exactly, as in an isolated system with a fixed total energy), and the canonical ensemble, where the constraint is enforced in a soft fashion (i.e. the constraint is met as an ensemble average, as in a system with fluctuating energy in a thermal bath at fixed temperature). Traditionally, these two ensembles are believed to become asymptotically the same in the thermodynamic limit, a notion that goes under the name of ensemble equivalence. Ensemble equivalence implies that, as the system becomes larger, it does not matter which of the two descriptions is adopted; in particular, the canonical and microcanonical entropies per particle become the same.

However, evidence has accumulated that ensemble equivalence can break down in certain circumstances. The most studied scenario where this can happen is the presence of phase transitions. In this case there are certain phases, or regions in parameter space, where ensemble equivalence breaks down (the canonical and microcanonical entropies per particle are different) and other phases where it is restored. A much more recent scenario for the breakdown of ensemble equivalence is the presence of local constraints in the system, i.e. constraints attached to each of the fundamental units, or at least a finite fraction of them: in other words, the number of constraints is extensive in the size of the system. An example is that of networks with a given number of connections (degree) for each node separately. In this case, it turns out that ensemble equivalence is broken throughout the entire parameter space, as in fact the simplest such models do not even have phase transitions.

In this thesis, we studied novel aspects of the breakdown of ensemble equivalence under local constraints and identified for the first time, to the best of our knowledge, their impact on certain key results in information theory.

In Chapter 1 we set the stage by introducing the main notions and establishing an informal analogy between ensembles in statistical physics and typical sequences in information theory. Systems with given macroscopic properties in statistical physics have a direct analogue in typical sets in information theory, which are in turn at the basis of several key results, including the maximum compressibility of data and the minimum rate of information transmission to preserve lossless communication.

In Chapter 2 we introduced the first model that studies ensemble nonequivalence in presence of both an extensive number of local constraints and a phase transition. The model is also the first one wherein ensemble nonequivalence is studied on weighted networks, i.e. networks where links can carry different weights. By making the constraints of each node of the network depend on a global temperature-like parameter, we showed that it is possible to induce a form of Bose-Einstein condensation whereby a finite fraction of the total link weights concentrates among a finite number of nodes. We also showed that the traditional criterion for ensemble equivalence, i.e. the vanishing of the relative fluctuations for the constraints in the canonical ensemble, becomes incorrect in the case of local constraints: nonvanishing relative fluctuations capture the onset of the condensation transition, but they do not capture the breakdown of ensemble equivalence. This is quite different from what people have been intuitively reasoning so far about relative fluctuations, and shows that the novel mechanism based on local constraints is quite subtle.

In Chapter 3 we significantly extended the framework of systems with local constraints, from the case of networks with constraints on each node to that of generic systems that can be described by rectangular matrices with constraints on the row and/or column sums. Such constrained matrices can represent a wide range of systems with spatial heterogeneity and/or temporal non-stationarity. We found that, in this more general setting, ensemble nonequivalence can occur in an even stronger way, with the difference between canonical and microcanonical entropies being of the same order as the entropies themselves. Such form of ensemble nonequivalence is as strong as the traditional one encountered in presence of phase transitions, while at the same time maintaining the property of being unrestricted in parameter space, as happens for the other known systems with local constraints. Therefore, it is the most robust form of nonequivalence documented so far. For many specific settings, we calculated explicitly the mathematical quantities distinguishing the two ensembles.

In Chapter 4 me made a major leap from statistical physics to information theory in order to investigate the consequences of ensemble nonequivalence for the compression of modern big data structures such as large networks or long multivariate time series. We first established a rigorous analogy between typical sets, i.e. the collectively most probable outcomes of an information source, and microcanonical ensembles as subsets of the canonical configurations of a physical system. We then showed that, when ensemble equivalence holds, the analogy is actually an identity: the microcanonical configurations coincide with the typical set of the canonical ensemble, implying no difference between sources subject to hard and soft constraints. However, when ensemble equivalence breaks down, the typical set of the canonical ensemble is irreducible to the microcanonical ensemble. In this case, we showed that, for hard constraints, standard information-theoretic results have to be generalized using the calculations of Chapter 3. We revised in particular the traditional information-theoretic bounds for data compression based on Shannon entropy. We found that microcanonical sources require less storage space but more computational costs, while the opposite is true for canonical sources. This highlights a novel trade-off between memory and computation.

Finally, in Chapter 5 we considered an even more general setting with multiple information sources (such as those generating multivariate time series in finance or neuroscience) where constraints can couple both the outputs of different sources at the same time (spatial constraints) and the outputs of the same source at different times (temporal constraints). We found that temporal constraints never break ensemble equivalence, while spatial constraints do break it if the number of sources is finite. Again, while for canonical sources the standard Shannon theory remains valid, for microcanonical sources various information-theoretic quantities have to be corrected using the calculations of Chapter 3. Moreover, we find that the normalized (per output) total correlation between all sources subject to hard constraints coincides with the normalized (per output) difference between canonical and microcanonical entropies, which in turn only requires the knowledge of the covariance matrix between the canonical constraints and can therefore be calculated explicitly. If spatial constraints are deactivated, ensemble equivalence is restored and the normalized total correlation vanishes, so the microcanonical outputs become asymptotically mutually independent just like the canonical outputs.