Dual representations of classical and quantum entropies: Theory and Perspectives

 Olivier Rioul

The MAXENT (maximum entropy) principle gives a natural correspondance between (classical and quantum) statistical mechanics and information theory. Its general formulation leverages on log-concavity (or Klein's inequality) and the celebrated information inequality, which is equivalent to Gibbs' inequality. The resulting convex optimization allows one to relate entropy to the log-canonical partition function by Legendre-Fenchel duality. This can be further generalized to dual variational characterization of relative entropy or Kullback-Leibler divergence, relating Gibbs variationnal principle to Donsker-Vardhan variational representations.

From an information theoretical point of view, the Gibbs principle and its associated evidence lower bound serves as the basis of modern machine learning algorithms, such as the expectation-maximization (EM) algorithm and its restriction to parameters output by a neural network. Interestingly, a large class of alternating optimization algorithms including EM and Blahut-Arimoto for computing channel capacity and source rate-distortion functions can be cast in a generic information geometric alternating projections on convex sets of measures. On the other hand, Donsker-Vardhan representations, when restricted to the output of a neural network, allows one to efficiently estimate information measures, and when applied to parametric estimators with quadratic risk, provide HCR-like and CRB-like statistical lower bounds. An alternative derivation on such lower bounds in the Bayesian setting makes use of the Weyl-Heisenberg uncertainty principle, which opens up new perspectives.

In this keynote, I will select and review several of the aforementioned topics, and discuss the interactions between classical and quantum interpretations.