Visualizing the clustering transition using unsupervised learning

Alexandre G.R. Day[1,*]

[1]Department of Physics, Boston University, 590 Commonwealth Ave., Boston, MA 02215, USA

(Dated: March 24, 2018)

We study the solutions of two random boolean satisfiability problems 3-SAT and XORSAT using recently developed machine learning techniques. We use non-linear embedding techniques to learn the local manifolds in which the solutions organize. In particular we provide, to the best of our knowledge, the first visualizing of the so-called clustering transitions that are known to occur in those problems. Finally, using unsupervised clustering methods we are able to automatically extract quantities such as the entropy of clusters.

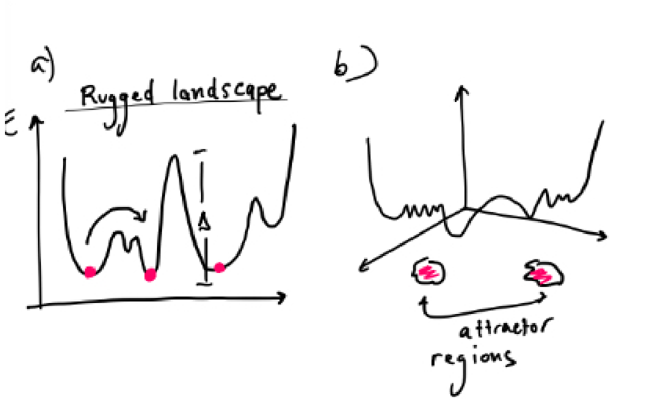

FIG. 1: Problem setting. (a) Rugged landscape cartoon picture. (b) In 3 dimensions we project attractors to a 2D space

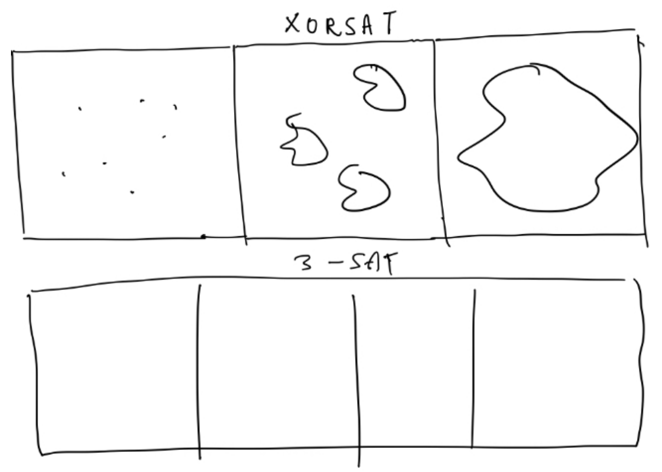

FIG. 2: t-SNE visualization of the solution space as a function of the constraint parameter α.

Landscape induces complexity, but not always : [1]

I. INTRODUCTION

Chris Moore discussion, etc. cite Mzard, stat. phys. and yourself.

II. MODEL STUDIED

XORSAT, 3-SAT.

A. Method used

BP, SP and Gaussian elimination.

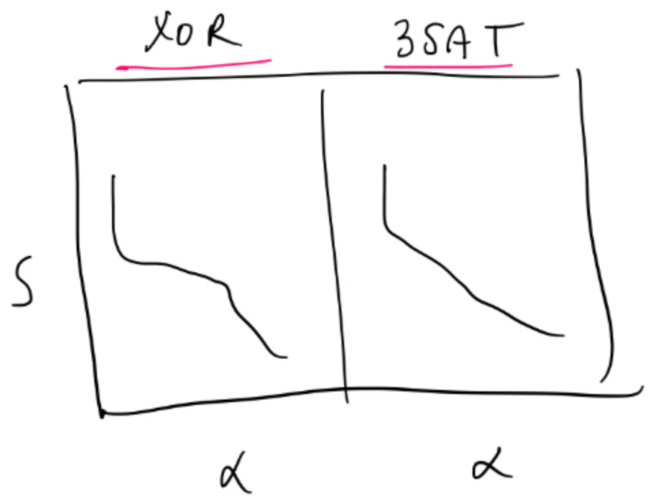

FIG. 3: Entropy of clusters vs. α.

---

*Electronic address: agrday@bu.edu

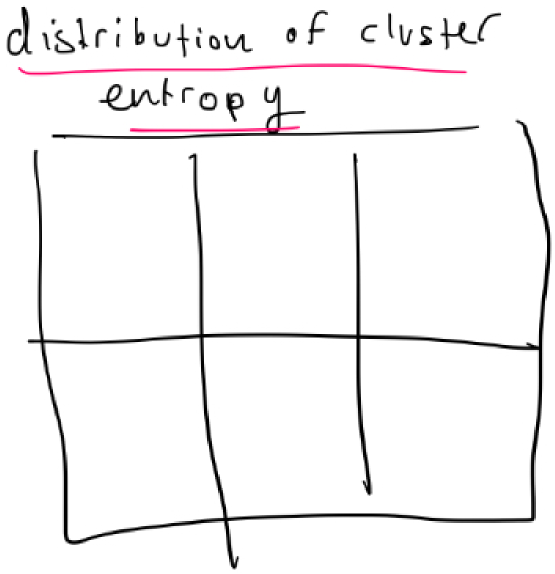

FIG. 4: Distribution of cluster entropies at different α.

III. RESULTS

A. Density clustering

B. Results and Discussion

Spin glass models such as the Sherrington-Kirkpatrick model or the Anderson model are known to be NP-complete [2]. Consequently it is believed that no algorithm can be devised in order to compute the ground-state in a time scaling polynomially with the system's size. Much work has been made in trying understanding what makes a problem hard from a statistical mechanics point of view. In particular, in random satisfiability problems such k-SAT and XORSAT the onset of a glass transition has been associated with the appearance of frozen variables and clustering in the solution space, which has been conjectured to induce failure of local search algorithms [3]. Note that glassiness does not necessarily imply hardness to solve. Phenomenologically, it appears that for glassy problems, any local search/stochastic method will fail, i.e. finding the ground-state is exponential in the system size. However, if one is able to devise some global method based on non-local updates, then some glassy problems are known to be in P. This is the case for instance of XORSAT (equivalent to solving a linear system mod 2) which can be efficiently solved using gaussian eliminination [1].

a. Clustering XORSAT, 3-SAT

b. Entropy via density clustering Relation to complexity and broader applications in stat. mechanics.

IV. CONCLUSION

Packages are available for t-SNE and efficient clustering available at alexandreday github

V. ACKNOWLEDGEMENTS

Acknowlegements.—

[1] F. Ricci-Tersenghi, Science **330**, 1639 (2010).
[2] D. Venturelli, S. Mandra, S. Knysh, B. O'Gorman, R. Biswas, and V. Smelyanskiy, Physical Review X **5**, 031040 (2015).
[3] C. Moore and S. Mertens, The nature of computation (OUP Oxford, 2011).

Supplemental Material

VI. DENSITY CLUSTERING VIA KERNEL DENSITY ESTIMATES