GAUSSIAN-WIDTH GRADIENT COMPLEXITY, REVERSE LOG-SOBOLEV INEQUALITIES AND NONLINEAR LARGE DEVIATIONS

RONEN ELDAN

Abstract. We prove structure theorems for measures on the discrete cube and on Gaussian space, which provide sufficient conditions for mean-field behavior. These conditions rely on a new notion of complexity for such measures, namely the Gaussian-width of the gradient of the log-density. On the cube $\{-1,1\}^n$, we show that a measure $\nu$ which exhibits low complexity can be written as a mixture of measures $\{\nu_\theta\}_{\theta \in I}$ such that: (i) for each $\theta$, the measure $\nu_\theta$ is a small perturbation of $\nu$ such that $\log \frac{d\nu_\theta}{d\nu}$ is a linear function whose gradient is small and, (ii) $\nu_\theta$ is close to some product measure, in Wasserstein distance, for most $\theta$. Thus, our framework can be used to study the behavior of low-complexity measures beyond approximation of the partition function, showing that those measures are roughly mixtures of product measures whose entropy is close to that of the original measure. In particular, as a corollary of our theorems, we derive a bound for the naïve mean-field approximation of the log-partition function which improves the nonlinear large deviation framework of Chatterjee and Dembo (Adv Math, 319:313–347, 2017. ISSN 0001-8708. https://dx.doi.org/10.1016/j.aim.2017.08.003) in several ways: (1) It does not require any bounds on second derivatives. (2) The covering number is replaced by the weaker notion of Gaussian-width. (3) We obtain stronger asymptotics with respect to the dimension. Two other corollaries are decomposition theorems for exponential random graphs and large-degree Ising models. In the Gaussian case, we show that measures of low-complexity exhibit an almost-tight reverse log-Sobolev inequality.

Contents

1 Introduction ....................................... 1549
  1.1 Main structure theorems. ............................ 1550
  1.2 A large deviation framework for functions of low complexity. .......................... 1554
    1.2.1 An example application: triangles in $G(N,p)$ .......................... 1555
  1.3 Mean-field behavior of the Ising model with large degree. .......................... 1556
  1.4 A decomposition theorem for exponential random graphs. .......................... 1557
  1.5 Approach. ..................................... 1559

2 The Gaussian Case ................................... 1561
  2.1 Tilts are close to product measures. .......................... 1561
  2.2 A reverse log-Sobolev inequality. .......................... 1564

Partially supported by the Israel Science Foundation (Grant No. 715/16).
1 Introduction

Let $\mu$ be a measure on the discrete hypercube $C_n = \{-1, 1\}^n$. In this work, we are interested in the following question: Under what natural conditions does this measure admit an approximate decomposition into a mixture of product measures most of which having roughly the same entropy as the measure $\mu$? This form of simplicity is a strong manifestation of what is referred to in the statistical mechanics literature as mean-field behavior.

Our main theorem provides a sufficient condition for such behavior, using a new notion of complexity, namely Gaussian-width gradient complexity. We say that a measure has low-complexity if one has nontrivial bounds on the Gaussian width of the gradient of its log-density (this is made rigorous and quantitative below). Our definition is inspired by Chatterjee and Dembo [CD16], where covering numbers are considered.

Our main theorem (Theorem 3 below) shows that for a measure $\mu$ on $C_n$ with log-Lipschitz density, a low-complexity condition implies the existence of an approximate decomposition into product measures as described above. Additionally, these measures can be written as small tilts of the original measure, namely, they can be attained by applying a change of density with respect to some log-linear function whose gradient is small.

Perhaps the most studied manifestation of mean field behavior is the approximation of the partition function, up to first order, via a product measure. More precisely, defining $C_n = \{-1, 1\}^n$ equipped with the uniform measure $\mu$, the Gibbs variational principle states that

$$\log \int e^f d\mu = \sup_{\nu} \left( \int f d\nu - D_{KL}(\nu \| \mu) \right)$$

(1)

where the supremum is taken over all probability measures $\nu$ on $C_n$ and $D_{KL}$ denotes the Kullback–Leibler divergence (defined below). The naive mean-field approximation is said to hold true when the supremum is approximately saturated by the class of product measures.