Simulated Annealing with Time dependent Energy Function via Sobolev inequalities

Matthias Löwe
Universität Bielefeld
Fakultät für Mathematik
Postfach 100131
33501 Bielefeld

Abstract

We analyze the Simulated Annealing Algorithm with an energy function U_t that depends on time. Assuming some regularity conditions on U_t (especially that U_t does not change too quickly in time), and choosing a logarithmic cooling schedule for the algorithm, we derive bounds on the Radon-Nikodym density of the distribution of the annealing algorithm at time t with respect to the invariant measure at time t. Moreover we estimate the entrance time of the algorithm into typical subsets V of the state space in terms of $\pi_t(V^c)$.

Keywords: Simulated Annealing, Sobolev inequalities, Spectral gap, Markov processes

1 Introduction

Let X be a finite set. The well known Simulated Annealing (SA) algorithm is an inhomogeneous Markov process Y_t on X with the aim to minimize a given function $U : X \to \mathbb{R}$. The idea behind SA is to think of U as an energy function and to choose the Markov process in such a way that the transition kernel at time t has at its invariant measure π_t, the Gibbs distribution with energy U at temperature $1/\beta_t$. So, if β_t diverges to infinity in the limit for large times the invariant measure becomes concentrated on the set of minima of U. On the other hand one has to "freeze" the process slowly enough to guarantee convergence of its distribution at time t towards equilibrium.

Since the introduction of SA by Kirkpatrick et al. in [K83] it has found a wide range of applications (see [GG84],[V93]), and the problems mentioned above, i.e. the optimal choice of β_t and the asymptotic behavior of SA have been studied by several authors. The problem of an asymptotically optimal "cooling schedule" (i.e. the dependence of β_t from t) has been settled by Hajek [Ha88], who showed the optimality of a logarithmic freezing schedule $\beta_t = \frac{1}{m}\log(A + Bt)$, where A and B are positive constants and m is, roughly speaking, the largest hill to be climbed to reach the global minimum.

The rate of convergence has been investigated by Gidas in [Gi85] and Mitra et al . in [M85] on the basis of Dobrushin coefficients for Markov chains. Cantoni has obtained bounds of the type $P(Y_t \geq E) \leq O(t^{-\alpha})$ in [C90] by using Freidlin-Vencel techniques (see [FV84]) on perturbed dynamical systems.

In [HS88] Holley and Stroock were able to give bounds for $P_x(Y_t = y)/\pi_t(y)$ for reversible SA processes based on Sobolev inequalities. Additionally they obtained sharp estimates for the spectral gap of the operators associated to the SA process. Based on refined spectral gap estimates Götze in [G92] proved effective bounds for the waiting time of escaping from an unlikely set (i.e. a set of high energy). These techniques were generalized by Deuschel and Mazza in [DM92] to Markov processes on finite sets similar to SA.

In this paper we are going to analyze that variant of the SA algorithm, where U depends on t, i.e. $U = U_t$. These processes have first been studied by Frigerio and Grillo in [FG93]. Not only they are a natural generalization of SA, but also they can be regarded as a discrete version of quantum diffusions (see [FG93], Section 4). Moreover these processes become interesting when thinking of SA as a sort of genetic algorithm (the reader is referred to the book of Holland [H75] for an introduction and a survey on these algorithms). Then the SA process can be regarded as an individual undergoing a mutation, the mutation being governed by the "fitness-function" U. In such a model the time dependence of U comes up naturally, if one e.g. thinks of the fitness of an individual being given by its environment, that usually changes in time.