

Lemma 1 (Boltzmann Equilibrium). *The Boltzmann equilibrium distribution $p(x) \propto \exp(-2\phi(x)/\sigma^2)$ minimizes the free energy,*

$$F(p) = - \int p(x)\phi(x)dx - \frac{\sigma^2}{2} \int p(x) \log p(x)dx. \quad (88)$$

□

Lemma 3 (Score Matching). *Let p be the pdf of the rv X and q a pdf, such that*

$$\lim_{u \rightarrow -\infty} E\left[\frac{\partial \log q(X)}{\partial x_i} \mid X_i = u\right] = \lim_{u \rightarrow \infty} E\left[\frac{\partial \log q(X)}{\partial x_i} \mid X_i = u\right] = 0 \quad \square$$

for $i = 1 \dots n$. ==>

$$\begin{aligned} \frac{1}{2}E[\|\nabla_x \log p(X) - \nabla_x \log q(X)\|^2] &= \frac{1}{2}E[\|\nabla_x \log q(X)\|^2] \\ &+ E[\nabla_x^2 \cdot \log q(X)] + \frac{1}{2}E[\|\nabla_x \log p(X)\|^2]. \end{aligned}$$

Corollary 1.

$$\begin{aligned} \frac{1}{2}E[\|\nabla_x \log p_\infty^\theta(X_t^\theta) - \nabla_x \log p_t^\theta(X_t^\theta)\|^2] &= \frac{1}{2}E[\|\nabla_x \log p_t^\theta(X_t^\theta)\|^2] \\ &+ \frac{1}{2}E[\|\nabla_x \frac{2}{\sigma^2} \phi(X_t^\theta, \theta)\|^2] + \frac{2}{\sigma^2}E[\nabla_x^2 \cdot \phi(X_t^\theta, \theta)]. \end{aligned}$$

Corollary 2 (Probability Velocity).

$$\begin{aligned} \frac{1}{2}E[\|V_t^\theta(X_t^\theta)\|^2] &= \frac{1}{2}E[\|\frac{\sigma^2}{2}\nabla_x \log p_t^\theta(X_t^\theta)\|^2] \\ &+ \frac{1}{2}E[\|\nabla_x \phi(X_t^\theta, \theta)\|^2] + \frac{\sigma^2}{2}E[\nabla_x^2 \cdot \phi(X_t^\theta, \theta)]. \end{aligned}$$

where

$$V_t^\theta(x) \stackrel{\text{def}}{=} \nabla_x \left(\phi(x, \theta) - \frac{\sigma^2}{2} \log p_t^\theta(x) \right)$$

is the probability velocity at state x , and time t , on a network with parameter θ .

Theorem 1 (Velocity of the Average Potential).

$$\frac{dE[\phi(X_t^\theta, \theta)]}{dt} = \frac{1}{2} \left(E[\|V_t^\theta(X_t^\theta)\|^2] + E[\|\nabla_x \phi(X_t^\theta, \theta)\|^2] - E[\|\nabla_x \frac{\sigma^2}{2} \log p_t^\theta(X_t^\theta)\|^2] \right).$$

Theorem 2 (Velocity of the Entropy). Let ϕ, p_t^θ such

that

$$\lim_{t \rightarrow -\infty} \phi(x, \theta) \frac{\partial}{\partial x_i} p_t^\theta(x) = \lim_{t \rightarrow \infty} \phi(x, \theta) \frac{\partial}{\partial x_i} p_t^\theta(x) = 0$$

for $i = 1, \dots, n$. Then

$$\frac{dH(X_t^\theta)}{dt} = E[\nabla_x^2 \cdot \phi(X_t^\theta, \theta)] - \frac{\sigma^2}{2} E[\nabla_x^2 \cdot \log p_t^\theta(X_t^\theta)],$$

where $H(X_t^\theta)$ is the entropy of X_t^θ .

□

Theorem 3 (Velocity of the Free Energy).

$$\frac{dF_t^\theta(X_t^\theta)}{dt} = -E[\|V_t^\theta(X_t^\theta)\|^2],$$

where F_t^θ is the free energy of X_t^θ .

Theorem 4 (Exact CD). Exact CD is a Minimum Velocity Algorithm.

Theorem 5 (Standard CD). As $t \rightarrow 0$ standard CD \rightarrow exact CD.

Given a fixed network parameter θ'

$$\nabla_\theta \left. \frac{dE[\phi(X_t^{\theta'}, \theta)]}{dt} \right|_{t=0, \theta=\theta'} = \nabla_\theta \frac{1}{2} E[\|V_0^\theta(\xi)\|^2]$$

□