Randomness and Computation 2023
Tutorial 8 (week 10)
Issue date: Wednesday, 15th November, 2023.

1. **Gibbs sampling for the Ferromagnetic Ising model.** A spin system is a probability distribution on $\mathcal{X} = \{-1, +1\}^V$, where V vertices of graph $G = (V, E)$. The value $\gamma(v)$ is called the spin of v. The Ising model is used by physicist as a classical approximation to model magnetic properties of materials. The interpretation is that magnets have up or down orientation, encoded by the $+1$ and -1, and are placed in the vertices of the graph, where the edges will encode the interaction between magnets. The nearest-neighbor ferromagnetic Ising model is one of the most widely studied spin system. The energy of a configuration γ is defined to be

$$H(\gamma) = -\sum_{\substack{v,w \in V \\ v \sim w}} \sigma(v)\sigma(w), \tag{1}$$

where the energy increases with the number of pairs of neighbors whose spin disagree. The Gibbs distribution corresponding to energy H is the probability distribution μ on $\mathcal{X}$ defined by

$$\mu(\sigma) = \frac{1}{Z(\beta)} e^{-\beta H(\sigma)},$$

where $\beta \geq 0$ is a parameter related to the inverse temperature and $Z(\beta) = \sum_{\sigma \in \mathcal{X}} e^{-\beta H(\sigma)}$ is a normalization term called partition function, that plays an important role in the physical description of the system.

   (a) Show that for $\beta = 0$ the distribution μ is nothing else than the uniform distribution over all spin configurations.

   (b) Show that for $\beta = \infty$, μ is the uniform distribution over the set of configuration σ that minimize $H(\sigma)$, i.e., with probability $1/2$ all spins aligned to $+1$ and with probability $1/2$ all spins aligned to -1.

   (c) Show that the transition probability of the Glauber dynamics selecting uniformly at random a vertex w from V reads:

   $$P(\sigma_t, \sigma_{t+1}^{\pm 1_w}) = (1 + \sigma(w) \tanh(\beta S(\sigma, w)))/2, \tag{2}$$

   where $S(\sigma, w) = \sum_{u:u\sim w} \sigma(u)$, σ_t is the initial spin configuration and $\sigma_{t+1}^{\pm 1_w}$ is the updated configuration with $\sigma_{t+1}(x) = \sigma_t(x) \quad \forall x \neq w$ and $\sigma_{t+1}^{+1_w}(w) = +1$ or $\sigma_{t+1}^{-1_w}(w) = -1$.

   (d) Prove that the corresponding Markov chain is irreducible and aperiodic.

   (e) Show it satisfies detailed balance.

2. **Path coupling for the Ferromagnetic Ising model**.

   Let's define the distance ρ on $\mathcal{X}$ by

$$\rho(\sigma, \tau) = \frac{1}{2} \sum_{u \in V} |\sigma(u) - \tau(u)|,$$

   which is easy to see that quantifies how many spins are different.

   We are going to follow the standard technique for proving rapid mixing presented in the course, where we design a coupling consisting on applying the same update to both chains X_t and Y_t. Let's consider σ and τ such that $\rho(\sigma, \tau) = 1$, i.e., they agree everywhere except in vertex v. Let's define $N(v) = \{u : u \sim v\}$ to be the set of neighbors vertices of v.

   (a) Explain why when the vertex w selected during the update satisfies $w = v$ we have $\rho(X_{t+1}, Y_{t+1}) = 0$.

   (b) Justify why when $w \notin N(v) \cup \{v\}$, then the distance does not change, i.e., $\rho(X_{t+1}, Y_{t+1}) = 1$.

   (c) Show that the probability of $\rho(X_{t+1}, Y_{t+1}) = 2$, when $w \in N(v)$, satisfies the condition

$$P(\rho = 2 | w \in N(v)) = |p(\tau, w) - p(\sigma, w)|.$$

   (d) Using the relation $\tanh(\beta(x+1)) - \tanh(\beta(x-1)) \leq \tanh \beta$ prove the relation

$$E[\rho_{t+1} | \rho_t = 1] \leq 1 - \frac{1 - \Delta \tanh(\beta)}{n} = 1 - \frac{c(\beta)}{n} \leq e^{-c(\beta)/n}.$$

   (e) Discuss under which conditions of Δ and $\tanh(\beta)$ this Glauber dynamics for the Ferromagnetic Ising model shows rapid mixing.

3. **Path coupling for Gibbs distribution on independent sets.** Let $G = (V, E)$ be a graph with maximum degree Δ, Ω be the set of independent sets on G, and $x \in \{0, 1\}^V$ a binary encoding of the vertices composing a given independent set. Let $\pi(x)$ the Gibbs distribution on independent sets:

$$\pi(x) = \begin{cases} \frac{\lambda^{|x|}}{Z(\lambda)} & \text{if } x(v)x(w) = 0 \quad \forall \{v, w\} \in E \\ 0 & \text{Otherwise.} \end{cases} \tag{3}$$

   where $|x|$ is the Hamming weight of configuration x, i.e., $|x| = \sum_{v \in V} x(v)$ and $Z(\lambda) = \sum_{x \in \mathcal{X}} \lambda^{|x|}$ normalizes π.

   The Glauber dynamics updates configuration X_t to a new configuration X_{t+1} by first selecting a vertex $v \in V$ uniformly at random and then implementing an update.

   Using the general definition of a Glauber dynamic update, show that we obtain the following update rule. First, set $X_{t+1}(w) = X_t(w) \quad \forall w \neq v$. Then, if exist $w' \in N(v)$, where $N(v)$ is the neighborhood of v, such that $X_t(w') = 1$.

- we set (M1) [1] $X_{t+1}(v) = 0$
- otherwise (M2)

$$X_{t+1}(v) = \begin{cases} 1 & \text{with probability } \lambda/(1+\lambda), \\ 0 & \text{with probability } 1/(1+\lambda). \end{cases} \tag{4}$$

(a) Let's define two chains X_t and Y_t that we will couple via the application of the exact same update rule at each step. Let's define the distance $\rho(X, Y)$ between two independent sets by the amount of vertices that are different, i.e, $\rho(X, Y) = \sum_{i \in V} |x(i) - y(i)|$. Assume that X_t and Y_t differ in a single vertex v:

   i. Explain why when the vertex w selected during the update satisfies $w = v$ we have $\rho(X_{t+1}, Y_{t+1}) = 0$.

   ii. Justify why when $w \notin N(v) \cup \{v\}$, then the distance does not change, i.e., $\rho(X_{t+1}, Y_{t+1}) = 1$.

   iii. Show that the probability of $\rho(X_{t+1}, Y_{t+1}) = 2$, when $w \in N(v)$, is always smaller than $\lambda/(1+\lambda)$.

(b) For which condition between the parameter λ and maximum degree of the graph Δ we have $E(\rho(X_{t+1}, Y_{t+1})) \leq 1 - c(\lambda)/n \leq e^{-c(\lambda)/n}$ with $c(\lambda) > 0$ when $\rho(X_t, Y_t) = 1$?

(c) Prove that the mixing time will satisfy $t_{\text{mix}}(\epsilon) \leq \frac{n}{c(\lambda)}(\log n + \log \epsilon^{-1})$.

Raul Garcia-Patron Sanchez

[1] Remark that the labeling (M1) and (M2) is used to facilitate the discussion of the solution.