**TD3 : RANDOM SAMPLING**

**Exercice 1 Metropolis chain**

Let \( \{X_n\}_{n \in \mathbb{N}} \) be a Markov chain on a finite state space \( E \) (for which the stationary distribution is unknown, and \( \pi \) a probability distribution). The goal of the exercise is to modify the transition probabilities of \( \{X_n\} \) so that the stationary distribution is \( \pi \) and the chain reversible.

To do that, we modify the transition from one state to another the following way: if at time \( n \), the chain is in state \( x \), if \( y \) is the next state drawn according to the transition probability of the chain, at time \( n+1 \), the state is \( y \) with probability \( a(x, y) \) and \( x \) with probability \( 1 - a(x, y) \).

1. Give the transition probabilities of the Markov chain described above. Remember that the chain must be reversible.
2. How to choose \( (a(x, y))_{x,y} \) so that the stationary distribution of this chain is \( \pi \)? Can we maximize the values of \( a(x, y) \)? Remember that the chain must be reversible.
3. Let \( G \) be an undirected connected graph, that is not known entirely, and consider a random walk on this graph. When visiting one state, the information available on this graph is only the set of neighbors and their degree. We want to sample a state uniformly at random. How to do?

**Exercice 2 Perfect sampling with a coin**

Let \( X \) be a finite set and \( P \) a probability distribution on \( X \). The goal of this exercise is to sample a random variable \( X \) according to \( P \) exactly with a non-biased. We also want to minimize average the number of time we throw the coin, that is \( \mathbb{E}[T] \), where \( T \) is the number of throws.

1. Show that all strategies can be represented by a binary tree (potentially infinite), where leaves are labeled by elements of \( X \). What is the condition so that the number of throws is almost surely finite and the distribution sampled \( P \)?

Let us denote \( H(X) = \sum_{x \in X} p(x) \log_2(1/p(x)) \) the entropy of \( X 

2. Show that \( \mathbb{E}[T] \geq H(X) \).
3. Give a necessary and sufficient condition on \( P \) so that this bound can be reached.

Let \( q \in [0, 1] \), and its dyadic decomposition \( q = \sum_{j \in \mathbb{N}} q_j \) avec \( q_j \in \{0, 2^{-j}\} \). We note \( T_q = \sum_{j \in \mathbb{N}} jq_j \), and first admit that \( T_q < -q \log q + 2q \).
4. In the general case, propose a strategy the guaranties \( H(X) \leq \mathbb{E}[T] \leq H(X) + 2 \).
5. Show that \( T_q < -q \log q + 2q \).

**Exercice 3 Independent sets of fixed size**

Let \( G = (S, A) \) be an undirected graph. The goal of this exercise is to sample an independent set of size \( k \) uniformly at random. His is still a problem that to not have a satisfactory solution yet.

Consider the Markov chain described by the Gibbs sampler:

1. choose \( v \in X_n \) and \( w \in S \) uniformly;
2. if \( X_n \cup \{w\} \setminus \{v\} \) is an independent set of size \( k \), then \( X_{n+1} = X_n \cup \{w\} \setminus \{v\} \). Else \( X_{n+1} = X_n \).

1. Is this chain aperiodic?

Let \( \Delta \) be the maximum degree of a vertex. We now assume that \( k \leq |S|/(3\Delta + 3) \).

2. Show that the chain is then irreducible.

3. Show that the stationary distribution is the uniform distribution on the independent sets of size \( k \).

Let \((X^1_n)\) and \((X^2_n)\) be two Markov chains on the independent sets of size \( k \), whose transition probabilities is according to the Gibbs sampler. We define the coupling as follows:

— Choose the same \( w \) for the two chains;
— choose \( v^1 \in X^1_n \) uniformly in \( X^1_n \). If \( v^1 \in X^2_n \), choose \( v^2 = v^1 \), otherwise, choose \( v^2 \) uniformly in \( X^2_n \setminus X^1_n \).

4. Show that the choice of \( v^2 \) is uniform.

We are now interested in the coupling time. For this, we focus on \( d_n = |X^1_n - X^2_n| \). There is coupling when \( d_n = 0 \).

5. Describe the possible possibilities for \( d_{n+1} \) in function of \( d_n \) (the possible values and their probabilities).

6. Show that
\[
\mathbb{E}[d_{n+1} \mid d_n = \ell] \leq \ell \left( 1 - \frac{|S| - (\Delta + 1)(3k - 3)}{|S|k} \right)
\]

7. Show that
\[
\mathbb{E}[d_{n+1}] \leq k \left( 1 - \frac{|S| - (\Delta + 1)(3k - 3)}{|S|k} \right)^n.
\]

8. Deduce an upper bound for the mixing time of the chain.