Notes on Sampling Independent Sets

Chihao Zhang (scribed by Ms. Yongqing Xu)

Last updated: May 24, 2020

Let $G = (V, E)$ be an undirected graph. For every set of vertices $S \subseteq V$, we use $N(S)$ to denote its neighbors, namely $N(S) = \{v \in V \setminus S : \exists u \in S, \{u, v\} \in E\}$. We use $G[S]$ to denote the subgraph of G induced by S. Let I_S be the set of independent sets of $G[S]$. We use an assignment $\sigma \in \{0,1\}^V$ to encode a set of "occupied" vertices where $\sigma(v) = 1$ means the vertex v is occupied and $\sigma(v) = 0$ means v is unoccupied. We say an edge is violated (by σ) if both of its ends are occupied.

The Gibbs measure of independent sets $\mu(\cdot)$ is the uniform distribution over I_V. For every $S \subseteq V$, we use $\mu_S(\cdot)$ to denote the marginal of μ on S, namely

$$\forall \sigma_S \in \{0,1\}^S, \quad \mu_S(\sigma_S) = \sum_{\sigma \in \{0,1\}^V : \sigma|_S = \sigma_S} \mu(\sigma).$$

For every $\tau \in \{0,1\}^{V \setminus S}$, define $\mu_S^\tau(\cdot)$ as

$$\forall \sigma_S \in \{0,1\}^S, \quad \mu_S^\tau(\sigma_S) \sim \mathbf{1}\Big[\sigma_S \in I_S \wedge \bigwedge_{e=\{u,v\} : u \in S, v \in N(S)} (\sigma_S(u) = 0 \vee \pi(v) = 0)\Big].$$

That is, we fix τ as the assignment of the boundary of S, and $\mu_S^\tau(\sigma_S)$ is nonzero iff σ_S is an independent set and none of edges across the boundary are violated. It is fine that π itself contains violating edges.

The partial rejection sampling algorithm for sampling independent sets is described Algorithm 1, which first appeared in [GJL19].

Algorithm 1 Partial rejection sampling independent sets

Input: *An undirected graph $G = (V, E)$.*
Output: *A random independent set of G.*

Randomly choose $\sigma \in \{0,1\}^V$
$\mathsf{Res} \leftarrow V$
while $\mathsf{Res} \neq \emptyset$ **do**
    Resample the assignment of vertices in Res and update σ accordingly
    $\mathsf{Bad} \leftarrow \bigcup_{\substack{e=\{u,v\} \in E: \\ \sigma(u)=\sigma(v)=1}} e$
    $\mathsf{Res} \leftarrow \mathsf{Bad} \cup N(\mathsf{Bad})$
end while
Output σ

We shall prove the correctness of the algorithm, namely the output of Algorithm 1 is a uniform independent set. The proof is adapted from a more general one in [FVY19].

Consider the i-th iteration of the while loop in Algorithm 1. We use $X^{(i)}$ and $R^{(i)}$ to denote the assignment σ and the resampling set Res at the end of the loop. Moreover, we let $X^{(0)}$ be the assignment σ obtained at line 1 and let $R^{(0)} = V$. Therefore, the execution of the algorithm can be viewed as the transitions:

$$\left(X^{(0)}, R^{(0)}\right) \to \left(X^{(1)}, R^{(1)}\right) \to \cdots \to \left(X^{(t)}, R^{(t)}\right)$$

In the following, we shall prove that each $\left(X^{(i)}, R^{(i)}\right)$ satisfies certain property, which can imply that $X^{(t)}$ is a uniform independent set. The property is called *conditional Gibbs* in [FVY19].

Lemma 1. *For every $i = 0, \ldots, t$, every set of vertices $r \subseteq V$, and every assignment $x \in \{0,1\}^V$ such that $\mathbf{Pr}\left[R^{(i)} = r, X^{(i)}_{R^{(i)}} = x_r\right] > 0$, it holds that*

$$\mathbf{Pr}\left[X^{(i)}_S = x_s \,\middle|\, R^{(i)} = r, X^{(i)}_{R^{(i)}} = x_r\right] = \mu_S^{x_r}(x_s), \tag{1}$$

where $s \triangleq V \setminus r$ and for any set $S \subseteq V$, X_S denotes the restriction of X on S.

The condition (1) guarantees that the algorithm is well-behaved. It is clear that if Lemma 1 holds, $X^{(t)}$ is uniform since $R^{(t)} = \varnothing$ and therefore eq. (1) becomes to

$$\mathbf{Pr}\left[X^{(t)} = x\right] = \mu(x).$$

We prove Lemma 1 by applying induction on i. The case $i = 0$ holds trivially, so we consider a transition $(X, R_X) \to (Y, R_Y)$ where the lemma holds for (X, R_X) via induction hypothesis. It remains to show that for every $r_y \subseteq V$ and every $y \in \{0,1\}^V$ such that $\mathbf{Pr}\left[R_Y = r_y, Y_{R_Y} = y_{r_y}\right] > 0$, it holds that

$$\mathbf{Pr}\left[Y_S = y_{s_y} \,\middle|\, R_Y = r_y, Y_{R_Y} = y_{r_y}\right] = \mu_{s_y}^{y_{r_y}}(y_{s_y})$$

where $s_y \triangleq V \setminus r_y$.

We first look at those r_y and y satisfying $\mathbf{Pr}\left[R_Y = r_y, Y_{R_Y} = y_{r_y}\right] > 0$. It must be the case that for some $\tilde{y}_{s_y} \in \{0,1\}^{s_y}$, the resampling set of $\tilde{y}_{s_y} \cup y_{r_y}$ [1] is r_y. This means that y_{r_y} consists of vertices in violated edges and their neighbors. Therefore, $\mu_{s_y}^{y_{r_y}}(y_{s_y}) = \frac{\mathbf{1}[y_{s_y} \in I_{s_y}]}{\left|I_{s_y}\right|}$ and we are going to show that

$$\mathbf{Pr}\left[Y_S = y_{s_y} \,\middle|\, R_Y = r_y, Y_{R_Y} = y_{r_y}\right] = \frac{\mathbf{1}[y_{s_y} \in I_{s_y}]}{\left|I_{s_y}\right|} \tag{2}$$

Instead of directly proving the equality, we show that the LHS of eq. (2) has following properties:

(1) If $y_{s_y} \notin I_{s_y}$, then LHS= 0;

(2) Otherwise, if we replace y_{s_y} by another $y'_{s_y} \in I_{s_y}$, the LHS is invariant.

This two properties together imply Equation (2).

It is easy to see that (1) holds since the complement of the resampling set must be an independent set, otherwise the vertices in violating edges would have been added into the resampling set. So now we assume y_{s_y} is an independent set. Applying the total probability rule, we obtain

[1] $\tilde{y}_{s_y} \cup y_{r_y}$ denotes the assignment $\tilde{y} \in \{0,1\}^V$ such that $\tilde{y}(u) = \tilde{y}_{s_y}(u)$ if $u \in s_y$ and $\tilde{y}(u) = y_{r_y}(u)$ if $u \in r_y$.

$$
\begin{aligned}
&\mathbf{Pr}\left[Y_{S_Y} = y_{s_y} \,\middle|\, R_Y = r_y, Y_{R_Y} = y_{r_y}\right] \\
&= \frac{\mathbf{Pr}\left[(Y, R_Y) = (y, r_y)\right]}{\mathbf{Pr}\left[R_Y = r_y \wedge Y_{R_Y} = y_{r_y}\right]} \\
&= \frac{\sum_{\substack{(x, r_x): \\ \mathbf{Pr}[X=x, R_X=r_x]>0}} \mathbf{Pr}\left[X = x, R_X = r_x\right] \cdot \mathbf{Pr}\left[(Y, R_Y) = (y, r_y) \,\middle|\, (X, R_X) = (x, r_x)\right]}{\mathbf{Pr}\left[R_Y = r_y \wedge Y_R = y_{r_y}\right]} \\
&= \mathbf{Pr}\left[R_Y = r_y \wedge Y_R = y_{r_y}\right]^{-1} \cdot \Bigg(\sum_{\substack{r_x \subseteq V, x_{r_x} \in \{0,1\}^{r_x}: \\ \mathbf{Pr}\left[X_R = x_{r_x}, R_X = r_x\right] > 0}} \mathbf{Pr}\left[X_R = x_{r_x}, R_X = r_x\right] \cdot \\
&\qquad \sum_{x_{s_x} \in \{0,1\}^{s_x}} \mathbf{Pr}\left[X_S = x_{s_x} \,\middle|\, X_R = x_{r_x}, R_X = r_x\right] \cdot \mathbf{Pr}\left[(Y, R_Y) = (y, r_y) \,\middle|\, (X, R_X) = (x, r_x)\right] \Bigg) \\
&= \mathbf{Pr}\left[R_Y = r_y \wedge Y_R = y_{r_y}\right]^{-1} \cdot \Bigg(\sum_{\substack{r_x \subseteq V, x_{r_x} \in \{0,1\}^{r_x}: \\ \mathbf{Pr}\left[X_R = x_{r_x}, R_X = r_x\right] > 0}} \mathbf{Pr}\left[X_R = x_{r_x}, R_X = r_x\right] \cdot \\
&\qquad \sum_{x_{s_x} \in \{0,1\}^{s_x}} \frac{\mathbf{1}[x_{s_x} \in I_{s_x}]}{\left|I_{s_x}\right|} \cdot \mathbf{Pr}\left[(Y, R_Y) = (y, r_y) \,\middle|\, (X, R_X) = (x, r_x)\right] \Bigg), \qquad (3)
\end{aligned}
$$

where $s_x = V \backslash r_x$. The last equality above is due to the induction hypothesis $\mathbf{Pr}\left[X_S = x_{s_x} \,\middle|\, X_R = x_{r_x}, R_X = r_x\right] = \mu_{s_x}^{x_{r_x}}(x_{s_x}) = \frac{\mathbf{1}[x_{s_x} \in I_{s_x}]}{|I_{s_x}|}$.

Remember that we want to show that eq. (3) is invariant for $y_{s_y} \in I_{s_y}$. It is instructive to examine each term of eq. (3). It is clear that the term $\mathbf{Pr}\left[R_Y = r_y \wedge Y_R = y_{r_y}\right]^{-1}$ and $\mathbf{Pr}\left[X_R = x_{r_x}, R_X = r_x\right]$ are invariant for any independent set y_{s_y}. Therefore, we can fix a pair (r_x, x_{r_x}) such that $\mathbf{Pr}\left[X_R = x_{r_x}, R_X = r_x\right] > 0$ and examine

$$
\sum_{x_{s_x} \in \{0,1\}^{s_x}} \frac{\mathbf{1}[x_{s_x} \in I_{s_x}]}{\left|I_{s_x}\right|} \cdot \mathbf{Pr}\left[(Y, R_Y) = (y, r_y) \,\middle|\, (X, R_X) = (x, r_x)\right] \qquad (4)
$$

Since our partial rejection sampling algorithm only resamples vertices in r_x, $\mathbf{Pr}\left[(Y, R_Y) = (y, r_y) \,\middle|\, (X, R_X) = (x, r_x)\right] > 0$ only if $x_{s_x} = y_{s_x}$. Therefore, we can define an assignment $\tilde{x} \in \{0,1\}^V$ as

$$
\tilde{x}(u) = \begin{cases} y(u) & u \in s_x; \\ x_{r_x}(u) & u \in r_x, \end{cases}
$$

and eq. (4) becomes to

$$
\frac{\mathbf{1}[\tilde{x}_{s_x} \in I_{s_x}]}{\left|I_{s_x}\right|} \cdot \mathbf{Pr}\left[(Y, R_Y) = (y, r_y) \,\middle|\, (X, R_X) = (\tilde{x}, r_x)\right].
$$

If $\tilde{x}_{s_x} \notin I_{s_x}$, the ends of violating edges must belong to r_y, so the whole term is 0 and invariant on y_{s_y}. Otherwise, the term is the constant $\frac{1}{|I_{s_x}| 2^{|r_x|}}$.

References

[FVY19] Weiming Feng, Nisheeth K Vishnoi, and Yitong Yin. Dynamic sampling from graphical models. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 1070–1081, 2019. 1, 2

[GJL19] Heng Guo, Mark Jerrum, and Jingcheng Liu. Uniform sampling through the lovász local lemma. *Journal of the ACM (JACM)*, 66(3):1–31, 2019. 1