Exact Distributed Sampling

Sriram V. Pemmaraju, Joshua Z. Sobel

University of Iowa

June 2023

Distributed Computing

- Focus on LOCAL and CONGEST models
- Every vertex is a machine
- Computation is done in rounds
- At the end of each round, vertices can communicate with neighbors
- In CONGEST model, message lengths are limited to $O(\log n)$
- In LOCAL model, message lengths are unbounded

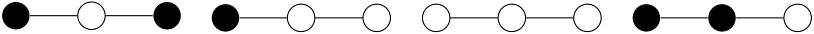
Weighted Local Constraint Satisfaction Problems

- Most distributed work focuses on arbitrary feasible solution
- Goal is to sample labelings from a weighted local CSP: $(G, L, \mathcal{C})$
  - Base graph $G = (V, E)$
  - Set of labels L
  - Set of constraints $\mathcal{C}$
- A labeling is a function from $V \to L$
- For $C \subseteq V$, a constraint on C is a function from $L^C \to \mathbb{R}^{\geq 0}$
- Weight of a labeling $\ell = \prod_{C \in \mathcal{C}} C(\ell_C)$
- Local, if every constraint has bounded diameter

Sampling Weighted Local CSPs

- We want to sample labelings proportional to their weights $\prod_{C \in \mathcal{C}} C(\ell_C)$
- Examples: uniform colorings, hardcore model, ...
- Hardcore model: independent sets with parameter $\lambda > 0$, where the weight of set I is $\lambda^{|I|}$
  - $L = \{0, 1\}$
  - $\mathcal{C}_1 = \left\{ C_v(\ell) = \begin{cases} 1, \text{if } \ell_v = 0 \\ \lambda, \text{if } \ell_v = 1 \end{cases} : v \in V \right\}$
  - $\mathcal{C}_2 = \left\{ C_{u,v}(\ell) = \begin{cases} 0, \text{if } \ell_u = \ell_v = 1 \\ 1, \text{otherwise} \end{cases} : \{u, v\} \in E \right\}$
  - $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$

Hardcore Model



weight $= \lambda^2$ | weight $= \lambda$ | weight $= 1$ | weight $= 0$

Local Metropolis

- Feng, Sun, and Yin (PODC '17) proposed a distributed Markov chain for sampling from many weighted local CSPs
- Every vertex proposes a new label at each step
- Each proposal is accepted or rejected
- Fischer and Ghaffari (DISC '18) as well as Feng, Hayes, and Yin (arXiv '18) improved algorithm by limiting proposals to marked vertices
- For many important weighted local CSPs, mixes in $O(\log n)$ CONGEST or LOCAL rounds

Markov Chain for Colorings

- Start from arbitrary coloring X
- Each round proceeds as follows
  - Each vertex is marked active with probability p
  - Each active vertex v randomly proposes a color, σ_v
  - Set $X_v = \sigma_v$, if for all $\{u, v\} \in E$
    - $\sigma_v \neq \sigma_u$
    - $\sigma_v \neq X_u$
    - $\sigma_u \neq X_v$

Our Results

- Markov chain simulation gives approximate sampling
- Exact markov chain sampling is possible in sequential setting
  - *Coupling from the past*: Propp, Wilson (Random Structures & Algorithms '96)
  - *Bounding chains*: Huber (STOC '98) & Häggström, Nelander (Scandinavian Journal of Statistics '99)
- We show exact distributed sampling is also possible in some cases
- We have a general condition for $O(\log n)$ exact sampling in the LOCAL model (stronger condition in CONGEST model)

Our Results (Hardcore Model)

- Our approach:
  $O(\log n)$ Certifiable ✓ CONGEST sampling when $\lambda < \frac{1}{\Delta}$
- Guo, Jerrum, Liu (STOC '17):
  $O(\log n)$ CONGEST sampling when $\lambda < \frac{1}{2\sqrt{e}\Delta - 1}$
- Feng, Yin (PODC '18):
  $O(\log^3 n)$ LOCAL sampling when $\lambda < \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}$
- Lower bounds (LOCAL model):
  - $\Omega(\log n)$: Guo, Jerrum, Liu (STOC '17)
  - $\Omega(n^{1/11})$ for $\lambda > \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}$: Feng, Sun, and Yin (PODC '17)

Open Questions

- Can we achieve LOCAL hardcore model lower bound in CONGEST?
- Can our approach be extended to other local weighted CSPs? Most importantly, colorings?
- Can our algorithm run faster in all-to-all models?