Sublinear-Time Algorithms for Monomer-Dimer Systems on Bounded Degree Graphs

Marc Lelarge[1,2] and Hang Zhou[1]

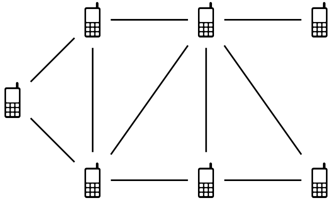
[1]École Normale Supérieure de Paris, France

[2]INRIA, France

December 16, 2013

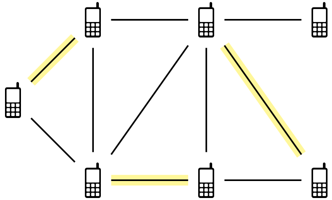
Background

- Sublinear-time algorithms for graph problems
- Optimization vs. Counting and Statistics



Background

- Sublinear-time algorithms for graph problems
- Optimization vs. Counting and Statistics



Monomer-Dimer systems

- $G = (V, E)$ undirected graph with $|V| = n$
- Maximum degree Δ
- Monomer $\leftrightarrow$ one vertex, Dimer $\leftrightarrow$ two adjacent vertices
- Monomer-Dimer arrangement $\leftrightarrow$ matching

Definition

- $\mathbb{M}$: set of all matchings of G
- Partition function: $Z(G, \lambda) = \sum_{M \in \mathbb{M}} \lambda^{|M|}$
- Gibbs distribution:

$$\pi_{G,\lambda}(M) = \frac{\lambda^{|M|}}{Z(G, \lambda)}, \quad \forall\, M \in \mathbb{M}$$

- Marginal probability:

$$p_{G,\lambda}(v) := \sum_{M \not\ni v} \pi_{G,\lambda}(M), \quad \forall\, v \in V$$

Previous Results

- Partition function:
  - $\#P$-complete (Valiant 1979, Vadhan 2002)
  - Randomized polynomial-time approximation scheme (Sinclair 1993, Jerrum Sinclair 1997)
  - Deterministic polynomial-time approximation scheme (Bayati *et al.* 2007)
- Matching statistics
  - $\#P$-hard (Sinclair Srivastava 2013)
- Permanent of expander graphs
  - Randomized polynomial-time approximation scheme (Jerrum Sinclair Vigoda 2004)
  - Deterministic polynomial-time approximation algorithm (Gamarnik Katz 2010)

Our Main Results

Local computations for marginal probability

- Approximation algorithm
- Complexity lower bound

Randomized sublinear-time approximation algorithms for

- Partition function
- Matching statistics
- Permanent of expander graphs

Notations

- ϵ-approximation solution : within additive error ϵ
- ϵ-approximation algorithm : outputs an ϵ-approximation solution
  with probability at least $2/3$

- Oracles: $\mathcal{D}(v)$ and $\mathcal{N}(v, i)$
- Our focus: Query complexity

Marginal probability

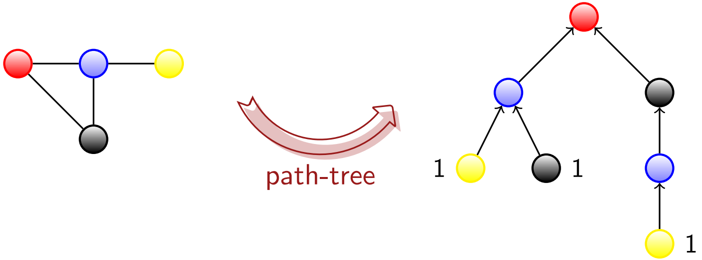
$$p_{G,\lambda}(v) = \frac{1}{1 + \lambda \sum_{u \in N(G,v)} p_{G \setminus \{v\},\lambda}(u)}$$

What is the probability that ● is matched? $(\lambda = 1)$

Marginal probability

$$p_{G,\lambda}(v) = \frac{1}{1 + \lambda \sum_{u \in N(G,v)} p_{G\setminus\{v\},\lambda}(u)}$$

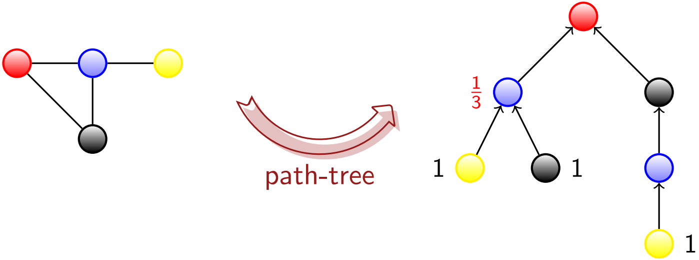
What is the probability that  is matched? $(\lambda = 1)$



Marginal probability

$$p_{G,\lambda}(v) = \frac{1}{1 + \lambda \sum_{u \in N(G,v)} p_{G \setminus \{v\},\lambda}(u)}$$

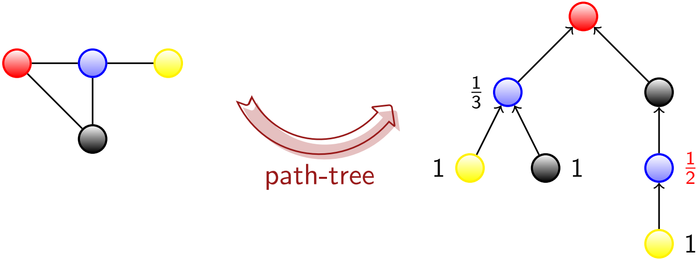
What is the probability that  is matched? $(\lambda = 1)$



Marginal probability

$$p_{G,\lambda}(v) = \frac{1}{1 + \lambda \sum_{u \in N(G,v)} p_{G \setminus \{v\},\lambda}(u)}$$

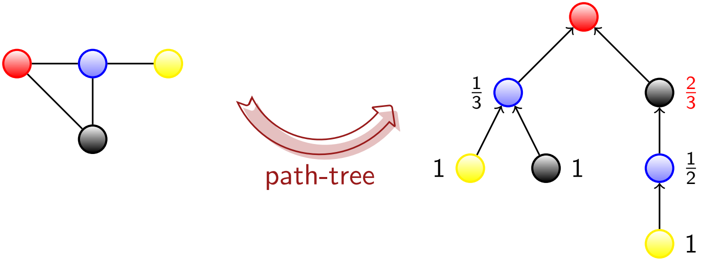
What is the probability that  is matched? $(\lambda = 1)$



Marginal probability

$$p_{G,\lambda}(v) = \frac{1}{1 + \lambda \sum_{u \in N(G,v)} p_{G \setminus \{v\},\lambda}(u)}$$

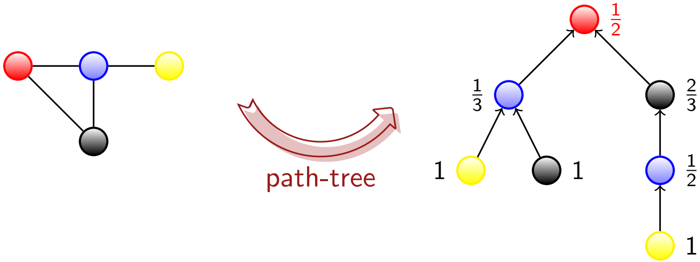
What is the probability that  is matched? ($\lambda = 1$)



Marginal probability

$$p_{G,\lambda}(v) = \frac{1}{1 + \lambda \sum_{u \in N(G,v)} p_{G \setminus \{v\},\lambda}(u)}$$

What is the probability that  is matched? ($\lambda = 1$)



Path-tree

- $T(v)$: the path-tree rooted at v
- $x_u(v) := \dfrac{1}{1 + \lambda \sum_{w \succ u} x_w(v)}$, for every $u \in T(v)$
- $x_v(v) = p_{G,\lambda}(v)$

Truncated path-tree

- $T^h(v)$: $T(v)$ truncated at depth h
- $x_u^h(v)$: the solution of the recursions in $T^h(v)$

Correlation decay property (Bayati *et al.* 2007)

$|\log x_v^h(v) - \log p_{G,\lambda}(v)| \le \epsilon$, for any $h \ge h(\epsilon, \Delta) = \tilde{O}\left(\sqrt{\lambda\Delta}\log(1/\epsilon)\right)$.

Optimality

Proposition

We can compute an ϵ-approximation of $p_{G,\lambda}(v)$ using $O\left(\Delta^{h(\epsilon,\Delta)}\right)$ queries.

Proposition (lower bound)

Any approximation algorithm requires at least the above query complexity.

Optimality

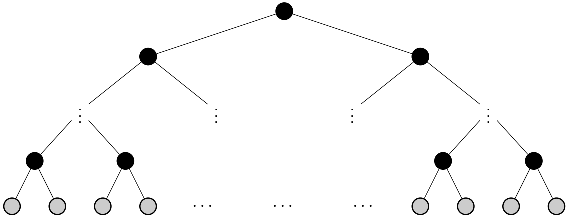
Proposition

We can compute an ϵ-approximation of $p_{G,\lambda}(v)$ using $O\left(\Delta^{h(\epsilon,\Delta)}\right)$ queries.

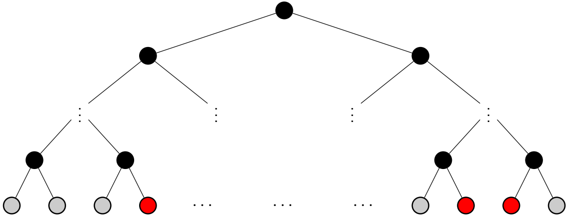
Proposition (lower bound)

Any approximation algorithm requires at least the above query complexity.

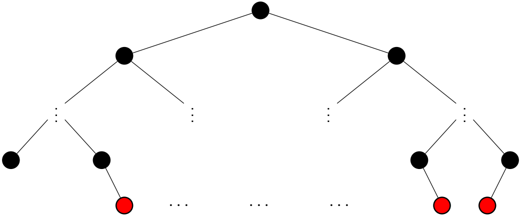
Idea of the lower bound



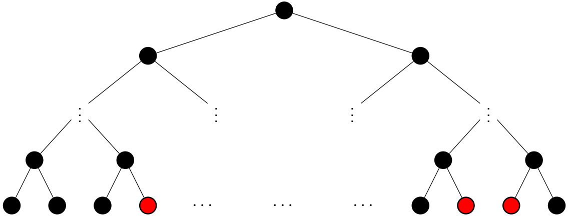
Idea of the lower bound



Idea of the lower bound



Idea of the lower bound



Partition function

$$Z(G,\lambda) = \prod_{1\le k\le n} p^{-1}_{G_k,\lambda}(v_k)$$

Algorithm for $\log Z(G,\lambda)$

- Take $\Theta(1/\epsilon^2)$ samples uniformly at random from $[1,\ldots,n]$;
- Compute an ϵ-approximation of $\log p^{-1}_{G_k,\lambda}(v_k)$ for every sample k;
- Return $n\cdot$(the average of the estimates).

Partition function

Main Theorem

We have an ϵn-approximation algorithm for the logarithm of the partition function with $\tilde{O}\left((1/\epsilon)^{\tilde{O}(\sqrt{\Delta})}\right)$ queries. In addition, any ϵn-approximation algorithm needs $\Omega(1/\epsilon^2)$ queries.

Average matching size

$$
\begin{aligned}
E(G,\lambda) &:= \sum_{M\in\mathbb{M}} |M| \cdot \pi_{G,\lambda}(M) \\
&= n - \frac{1}{2} \sum_{1\le k\le n} p_{G,\lambda}(v_k)
\end{aligned}
$$

Theorem

We have an ϵn-approximation algorithm for the average matching size with $\tilde{O}\left((1/\epsilon)^{\tilde{O}(\sqrt{\Delta})}\right)$ queries. In addition, any ϵn-approximation algorithm needs $\Omega(1/\epsilon^2)$ queries.

Entropy of a matching

$$\begin{aligned} S(G,\lambda) &:= -\sum_{M\in\mathbb{M}} \pi_{G,\lambda}(M)\log \pi_{G,\lambda}(M) \\ &= \log Z(G,\lambda) + \log\lambda \cdot E(G,\lambda) \end{aligned}$$

Corollary

We have an ϵn-approximation algorithm for the entropy of a matching with $\tilde{O}\left((1/\epsilon)^{\tilde{O}(\sqrt{\Delta})}\right)$ queries. In addition, any ϵn-approximation algorithm needs $\Omega(1/\epsilon^2)$ queries.

Application: Permanent of expander graphs

- Bi-partite graph G with vertices $X \cup Y$, $|X| = |Y| = n$
- α-expander graph:
  $|N(S)| \geq (1+\alpha)|S|$, for $S \subset X$ or $S \subset Y$ with $|S| \leq n/2$
- PERM : Permanent of the adjacency matrix of G

Lemma (Gamarnik Katz 2010)

$$1 \leq \frac{Z(G,\lambda)}{\lambda^n \cdot \mathrm{PERM}} \leq e^{O(n\lambda^{-1}\log^{-1}(1+\alpha)\log\Delta)}.$$

Corollary

We have an ϵn-approximation algorithm for $\log \mathrm{PERM}$ with query complexity $\tilde{O}\left((1/\epsilon)^{\tilde{O}(\sqrt{\Delta/(\epsilon\alpha)})}\right)$.

Conclusion

- Marginal probability *(correlation decay property)*
- Sublinear-time approximation for monomer-dimer systems
- Extension to the partition function of independent sets
- Experiments on large real-world graphs

Thank you!