Approximation algorithms for partition functions of edge-coloring models

Guus Regts

29 March 2016

The Classification Program of Counting Complexity, Simons Institute

Introduction: FPTAS for partition functions

Correlation decay method (assuming $\Delta = \Delta(G)$ is constant) yields a (deterministic) FPTAS for:

- The number of weighted independent sets with weight $\lambda \geq 0$ for λ small enough (Weitz, 2006)
- The number of matchings in a graph (Bayati, Gamarnik, Katz, Nair and Tetali, 2007)
- The number of k-colorings of a graph for $k > \alpha\Delta + 1$ for α large enough (Lu and Yin, 2013)
- Partition function of real symmetric matrices A with $|A_{i,j} - 1| \leq c/\Delta$ (for some constant c) (Lu and Yin, 2013)

Introduction: QPTAS for partition functions

Low order Taylor approximations yield a (deterministic) QPTAS for:

- Permanent of complex matrices Z with $|Z_{i,j} - 1| \leq 0.5$ (Barvinok, 2016+)
- Permanent of real matrices Z with $\delta \leq Z_{i,j} \leq 1$ (Barvinok, 2016+)
- The partition function of complex symmetric matrices A with $|A_{i,j} - 1| \leq 0.34/\Delta(G)$ (Barvinok and Sobéron, 2016)
- The partition function of a complex-valued Boolean polynomial (Barvinok, 2015+)
- The Tutte polynomial, $Z(u,v)(G) := \sum_{A \subseteq E} u^{k(A)} v^{|A|}$ for any fixed $v \in \mathbb{C}$ and $u \in \mathbb{C}$ with $|u|$ large enough (depending on v and $\Delta(G)$). (R., 2015+)

Contents of the presentation

1. What are edge-coloring models and their partition functions?
2. Main result
3. Algorithm and proof sketch
4. Discussion and concluding remarks

Partition functions of edge-coloring models

Definition

A k-color edge-coloring model h is a collection of symmetric tensors $\{h^d\}_{d \in \mathbb{N}}$ with $h^d \in (\mathbb{C}^k)^{\otimes d}$.

Partition functions of edge-coloring models

Definition

A k-color edge-coloring model h is a collection of symmetric tensors $\{h^d\}_{d\in\mathbb{N}}$ with $h^d \in (\mathbb{C}^k)^{\otimes d}$.
Let $e_1, \ldots, e_k$ be the standard basis for $\mathbb{C}^k$. Then we can write

$$h^d = \sum_{\phi:[d]\to[k]} h^d(\phi) e_{\phi(1)} \otimes \cdots \otimes e_{\phi(d)}.$$

Partition functions of edge-coloring models

Definition

A k-color edge-coloring model h is a collection of symmetric tensors $\{h^d\}_{d\in\mathbb{N}}$ with $h^d \in (\mathbb{C}^k)^{\otimes d}$.
Let $e_1, \ldots, e_k$ be the standard basis for $\mathbb{C}^k$. Then we can write

$$h^d = \sum_{\phi:[d]\to[k]} h^d(\phi) e_{\phi(1)} \otimes \cdots \otimes e_{\phi(d)}.$$

Definition

The partition function of h is the graph parameter defined by $G = (V, E) \mapsto p(G)(h)$ with

$$p(G)(h) = \sum_{\phi:E\to[k]} \prod_{v\in V} (h^{\deg(v)}(\phi(\delta(v))).$$

Partition functions of edge-coloring models

Definition

A tensor network is a pair (G, h) where h is a collection of symmetric tensors $\{h^v\}_{v \in V(G)}$ with $h^v \in (\mathbb{C}^k)^{\otimes \deg(v)}$.
Let $e_1, \ldots, e_k$ be the standard basis for $\mathbb{C}^k$. Then we can write

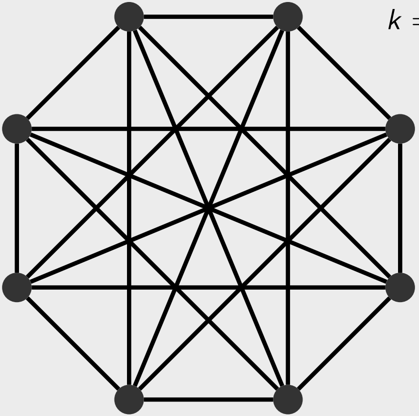
$$h^v = \sum_{\phi:[d]\to[k]} h^v(\phi) e_{\phi(1)} \otimes \cdots \otimes e_{\phi(d)}.$$

Definition

The contraction of (G, h) is defined by $G = (V, E) \mapsto p(G)(h)$ with

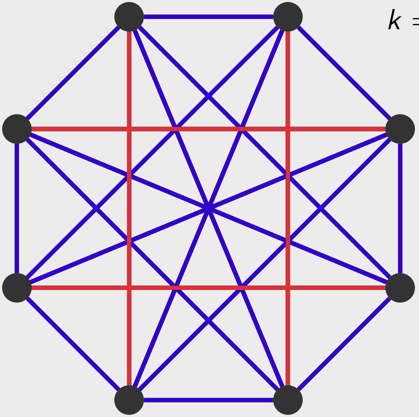
$$p(G)(h) = \sum_{\phi:E\to[k]} \prod_{v \in V} (h^v(\phi(\delta(v))).$$

An example: counting perfect matchings



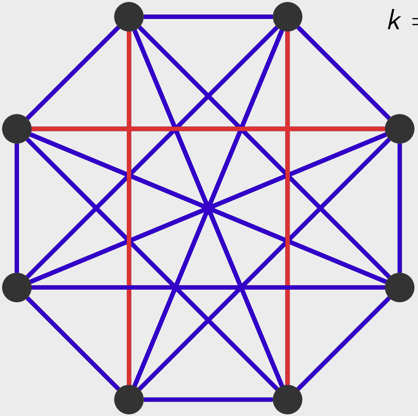
$$k = 2;\ h^d(\phi) := \begin{cases} 1 \text{ if } \sum_{i=1}^{d} \phi(i) = d+1 \\ 0 \text{ otherwise} \end{cases}$$

An example: counting perfect matchings



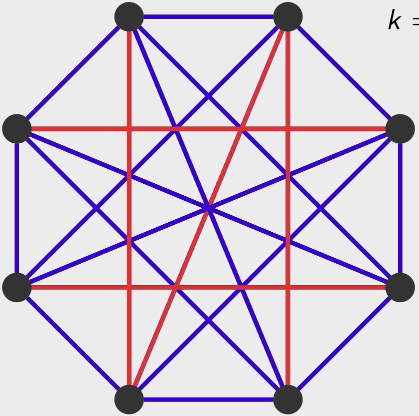
$$k = 2;\ h^d(\phi) := \begin{cases} 1 \text{ if } \sum_{i=1}^{d} \phi(i) = d + 1 \\ 0 \text{ otherwise} \end{cases}$$

An example: counting perfect matchings



$$k = 2;\ h^d(\phi) := \begin{cases} 1 \text{ if } \sum_{i=1}^{d} \phi(i) = d + 1 \\ 0 \text{ otherwise} \end{cases}$$

An example: counting perfect matchings



$$k = 2;\ h^d(\phi) := \begin{cases} 1 \text{ if } \sum_{i=1}^{d} \phi(i) = d + 1 \\ 0 \text{ otherwise} \end{cases}$$

Where do you find edge-coloring models?

$$p(G)(h) = \sum_{\phi:E(G)\to[k]} \prod_{v\in V(G)} h^{\deg(v)}(\phi(\delta(v)))$$

Where do you find edge-coloring models?

$$p(G)(h) = \sum_{\phi:E(G)\to[k]} \prod_{v\in V(G)} h^{\deg(v)}(\phi(\delta(v)))$$

- Theoretical computer science: holant problems, tensor networks
- Statistical physics: partition functions of the vertex model (de la Harpe and Jones 1993)
- Knot theory: Lie algebra weight systems
- Invariant theory: invariants of the (complex) orthogonal group $O_k = \{g \in \mathbb{C}^{k\times k} \mid gg^T = I\} \quad p(G)(gh) = p(G)(h)$.
- Combinatorics: counting perfect matchings, counting graph homomorphisms

Counting graph homomorphisms

Definition

Let A be a symmetric $n \times n$ matrix. The partition function of A is the graph invariant given by $G = (V, E) \mapsto p(G)(A)$ with

$$p(G)(A) := \sum_{\phi: V \to [n]} \prod_{uv \in E} A_{\phi(u),\phi(v)}.$$

Counting graph homomorphisms

Definition

Let A be a symmetric $n \times n$ matrix. The partition function of A is the graph invariant given by $G = (V, E) \mapsto p(G)(A)$ with

$$p(G)(A) := \sum_{\phi : V \to [n]} \prod_{uv \in E} A_{\phi(u),\phi(v)}.$$

Lemma (Szegedy 2007)

Write $A = B^T B$ (for some complex matrix B). Let $b_1, \ldots, b_n$ be the columns of B. Define h^d by $h^d = \sum_{i=1}^n b_i^{\otimes d}$. Then

$$p(G)(h) = p(G)(A).$$

Main result

Theorem (R. 2015)

Let (G, h) be a tensor network, where G has n vertices. If $|h^v(\phi) - 1| \leq \frac{0.35}{\Delta(G)+1}$ for each $\phi : [\deg(v)] \to [k]$ and each v, then for each $\varepsilon > 0$ we can (deterministically) compute in time $n^{O(\log n/\varepsilon)}$ a number ξ such that

$$e^{-\varepsilon} \leq \left| \frac{p(G)(h)}{\xi} \right| \leq e^{\varepsilon}.$$

Main result

Let $x \in \mathbb{C}^k$ be such that $x^T x \neq 0$ and let $X^d := x^{\otimes d}$.

Theorem (R. 2015)

Let (G, h) be a tensor network, where G has n vertices. If $|h^v(\phi) - X^d(\phi)|$ is sufficiently small for each $\phi : [\deg(v)] \to [k]$ and each v, then for each $\varepsilon > 0$ we can (deterministically) compute in time $n^{O(\log n/\varepsilon)}$ a number ξ such that

$$e^{-\varepsilon} \leq \left| \frac{p(G)(h)}{\xi} \right| \leq e^{\varepsilon}.$$

Proof: high level overview

- Let (G, h) be given. Let $\mathbb{1}$ be the all-ones model. Define

$$p(z) = p(G)(\mathbb{1} + z(h - \mathbb{1})).$$

Then $p(0) = k^{|E(G)|}$ and $p(1) = p(G)(h)$.

Proof: high level overview

- Let (G, h) be given. Let $\mathbb{1}$ be the all-ones model. Define

$$p(z) = p(G)(\mathbb{1} + z(h - \mathbb{1})).$$

Then $p(0) = k^{|E(G)|}$ and $p(1) = p(G)(h)$.

- If $p(z) \neq 0$ for all $|z| \leq q$ with $q > 1$, then $\ln(p(1))$ is well approximated by a low-order Taylor polynomial around 0.

Proof: high level overview

- Let (G, h) be given. Let $\mathbb{1}$ be the all-ones model. Define
$$p(z) = p(G)(\mathbb{1} + z(h - \mathbb{1})).$$
Then $p(0) = k^{|E(G)|}$ and $p(1) = p(G)(h)$.
- If $p(z) \neq 0$ for all $|z| \leq q$ with $q > 1$, then $\ln(p(1))$ is well approximated by a low-order Taylor polynomial around 0.
- The Taylor polynomial can be expressed in terms of the derivatives of p at 0.

Proof: high level overview

- Let (G, h) be given. Let $\mathbb{1}$ be the all-ones model. Define

$$p(z) = p(G)(\mathbb{1} + z(h - \mathbb{1})).$$

  Then $p(0) = k^{|E(G)|}$ and $p(1) = p(G)(h)$.
- If $p(z) \neq 0$ for all $|z| \leq q$ with $q > 1$, then $\ln(p(1))$ is well approximated by a low-order Taylor polynomial around 0.
- The Taylor polynomial can be expressed in terms of the derivatives of p at 0.
- $p(G)(h) \neq 0$ for all bounded degree graphs G and all h close enough to $\mathbb{1}$.

Approximating polynomials I

Lemma (Barvinok 2015)

Let p be a polynomial of degree n such that $p(z) \neq 0$ for all $|z| \leq q$ with $q > 1$. Let $f(z) = \ln p(z)$ and let $T_m(z) = \sum_{k=0}^{m} f^{(k)}(0)\frac{z^k}{k!}$. Then for $m = O(\ln(n/\varepsilon))$ we have that

$$e^{-\varepsilon} \leq \left| \frac{p(1)}{e^{T_m(1)}} \right| \leq e^{\varepsilon}.$$

Approximating polynomials I

Lemma (Barvinok 2015)

Let p be a polynomial of degree n such that $p(z) \neq 0$ for all $|z| \leq q$ with $q > 1$. Let $f(z) = \ln p(z)$ and let $T_m(z) = \sum_{k=0}^{m} f^{(k)}(0)\frac{z^k}{k!}$. Then for $m = O(\ln(n/\varepsilon))$ we have that

$$e^{-\varepsilon} \leq \left| \frac{p(1)}{e^{T_m(1)}} \right| \leq e^{\varepsilon}.$$

Proof sketch:
Write

$$p(z) = p(0) \prod_{i=1}^{n} (1 - z/\alpha_i).$$

Approximating polynomials I

Lemma (Barvinok 2015)

Let p be a polynomial of degree n such that $p(z) \neq 0$ for all $|z| \leq q$ with $q > 1$. Let $f(z) = \ln p(z)$ and let $T_m(z) = \sum_{k=0}^{m} f^{(k)}(0)\frac{z^k}{k!}$. Then for $m = O(\ln(n/\varepsilon))$ we have that

$$e^{-\varepsilon} \leq \left| \frac{p(1)}{e^{T_m(1)}} \right| \leq e^{\varepsilon}.$$

Proof sketch:
Write

$$p(z) = p(0) \prod_{i=1}^{n} (1 - z/\alpha_i).$$

Then

$$f(z) = \ln(p(z)) = \ln(p(0)) + \sum_{i=1}^{n} \ln(1 - z/\alpha_i).$$

Approximating polynomials II

Using the standard Taylor approximation of the natural logarithm, $\ln(1+x) = -\sum_{i=1}^{\infty} \frac{1}{i}(-x)^i$, we find that

$$\ln(1 - 1/\alpha_i) =$$

Approximating polynomials II

Using the standard Taylor approximation of the natural logarithm, $\ln(1+x) = -\sum_{i=1}^{\infty} \frac{1}{i}(-x)^i$, we find that

$$\ln(1 - 1/\alpha_i) = -\sum_{j=1}^{m} \frac{1}{j}\left(\frac{1}{\alpha_i}\right)^j + R_m,$$

with

$$|R_m| \leq \sum_{j>m} \frac{1}{m+1}\left(\frac{1}{q}\right)^j \leq \frac{1}{m+1}\frac{1}{(1-1/q)q^{m+1}}.$$

Approximating polynomials II

Using the standard Taylor approximation of the natural logarithm, $\ln(1+x) = -\sum_{i=1}^{\infty} \frac{1}{i}(-x)^i$, we find that

$$\ln(1 - 1/\alpha_i) = -\sum_{j=1}^{m} \frac{1}{j}\left(\frac{1}{\alpha_i}\right)^j + R_m,$$

with

$$|R_m| \leq \sum_{j>m} \frac{1}{m+1}\left(\frac{1}{q}\right)^j \leq \frac{1}{m+1}\frac{1}{(1-1/q)q^{m+1}}.$$

This implies that

$$|f(1) - T_m(1)| \leq \frac{n}{m+1}\frac{1}{(1-1/q)q^{m+1}}.$$

Approximating polynomials II

Using the standard Taylor approximation of the natural logarithm, $\ln(1+x) = -\sum_{i=1}^{\infty} \frac{1}{i}(-x)^i$, we find that

$$\ln(1 - 1/\alpha_i) = -\sum_{j=1}^{m} \frac{1}{j}\left(\frac{1}{\alpha_i}\right)^j + R_m,$$

with

$$|R_m| \leq \sum_{j>m} \frac{1}{m+1}\left(\frac{1}{q}\right)^j \leq \frac{1}{m+1}\frac{1}{(1-1/q)q^{m+1}}.$$

This implies that

$$|f(1) - T_m(1)| \leq \frac{n}{m+1}\frac{1}{(1-1/q)q^{m+1}}.$$

Taking $m = O(\log(n/\varepsilon))$ we get $|f(1) - T_m(1)| \leq \varepsilon$ and applying exp to both sides we have the lemma.

Proof: high level overview

- Let (G, h) be given. Let $\mathbb{1}$ be the all-ones model. Define
$$p(z) = p(G)(\mathbb{1} + z(h - \mathbb{1})).$$
Then $p(1) = k^{|E(G)|}$ and $p(1) = p(G)(h)$.
- If $p(z) \neq 0$ for all $|z| \leq q$ with $q > 1$, then $\ln(p(1))$ is well approximated by order $\ln(|V|)$ Taylor polynomial around 0.
- The Taylor polynomial of $\ln(p(1))$ can be expressed in terms of the derivatives of p at 0.
- $p(G)(h) \neq 0$ for all bounded degree graphs G and all h close enough to $\mathbb{1}$.

Computing derivatives I

Recall $f(z) = \ln p(z)$. So

$$f'(z) = \frac{p'(z)}{p(z)}$$

Computing derivatives I

Recall $f(z) = \ln p(z)$. So

$$f'(z) = \frac{p'(z)}{p(z)} \text{ that is } p'(z) = p(z)f'(z)$$

Recall $f(z) = \ln p(z)$. So

$$f'(z) = \frac{p'(z)}{p(z)} \text{ that is } p'(z) = p(z)f'(z)$$

So

$$p^{(2)}(z) = p(z)f^{(2)}(z) + p'(z)f'(z).$$

Computing derivatives I

Recall $f(z) = \ln p(z)$. So

$$f'(z) = \frac{p'(z)}{p(z)} \text{ that is } p'(z) = p(z)f'(z)$$

So

$$p^{(2)}(z) = p(z)f^{(2)}(z) + p'(z)f'(z).$$

More generally,

$$p^{(m)}(z) = \sum_{i=0}^{m-1} \binom{m-1}{i} p^{(i)}(z)f^{(m-i)}(z).$$

Computing derivatives I

Recall $f(z) = \ln p(z)$. So

$$f'(z) = \frac{p'(z)}{p(z)} \text{ that is } p'(z) = p(z)f'(z)$$

So

$$p^{(2)}(z) = p(z)f^{(2)}(z) + p'(z)f'(z).$$

More generally,

$$p^{(m)}(z) = \sum_{i=0}^{m-1} \binom{m-1}{i} p^{(i)}(z) f^{(m-i)}(z).$$

As $p(0) \neq 0$ this yields a nondegerate triangular system to compute $f^{(m)}(0)$ in terms of the $p^{(k)}(0)$ in $O(m^2)$ time.

Computing derivatives II

Recall that $p(z) = p(G)(\mathbb{1} + z(h - \mathbb{1})$. So

$$p(z) = \sum_{\phi:E\to[k]} \prod_{v\in V} (1 + z(h^v(\phi(\delta(v))) - 1))$$

Computing derivatives II

Recall that $p(z) = p(G)(\mathbb{1} + z(h - \mathbb{1})$. So

$$p(z) = \sum_{\phi:E\to[k]} \prod_{v\in V} (1 + z(h^v(\phi(\delta(v))) - 1))$$

$$= \sum_{i=0}^{n} z^i \sum_{\substack{U\subseteq V\\ |U|=i}} \sum_{\phi:E\to[k]} \prod_{v\in U} (h^v(\phi(\delta(v))) - 1).$$

Computing derivatives II

Recall that $p(z) = p(G)(\mathbb{1} + z(h - \mathbb{1}))$. So

$$p(z) = \sum_{\phi:E\to[k]} \prod_{v\in V} (1 + z(h^v(\phi(\delta(v))) - 1))$$

$$= \sum_{i=0}^{n} z^i \sum_{\substack{U\subseteq V\\ |U|=i}} \sum_{\phi:E\to[k]} \prod_{v\in U} (h^v(\phi(\delta(v))) - 1).$$

So $p^{(m)}(0)$ can be computed in time $O(|V|^m)$.

Proof: high level overview

- Let (G, h) be given. Let $\mathbb{1}$ be the all-ones model. Define

$$p(z) = p(G)(\mathbb{1} + z(h - \mathbb{1})).$$

Then $p(1) = k^{|E(G)|}$ and $p(1) = p(G)(h)$.

- If $p(z) \neq 0$ for all $|z| \leq q$ with $q > 1$, then $\ln(p(1))$ is well approximated by order $\log(|V|)$ Taylor polynomial at 0.
- The Taylor polynomial of $\ln(p(1))$ can be expressed in terms of the derivatives of p at 0 at cost $|V|^m$ for order m
- $p(G)(h) \neq 0$ for all bounded degree graphs G and all h close enough to $\mathbb{1}$.

Zero-free regions

Theorem (R. 2015)

Let (G, h) be a tensor network. If $|h^v(\phi) - 1| \leq \frac{0.355}{\Delta(G)+1}$ for each $\phi : [\deg(v)] \to [k]$ and each v, then

$$p(G)(h) \neq 0.$$

Proof is by sophisticated induction.

Proof: high level overview

- Let (G, h) be given. Let $\mathbb{1}$ be the all-ones model. Define

$$p(z) = p(G)(\mathbb{1} + z(h - \mathbb{1})).$$

Then $p(1) = k^{|E(G)|}$ and $p(1) = p(G)(h)$.

- If $p(z) \neq 0$ for all $|z| \leq q$ with $q > 1$, then $\ln(p(1))$ is well approximated by order $\log(|V|)$ Taylor polynomial at 0.
- The Taylor polynomial of $\ln(p(1))$ can be expressed in terms of the derivatives of p at 0 at cost $|V|^m$ for order m
- $p(G)(h) \neq 0$ for all graphs G and all h such that

$$|h^v(\phi) - 1| \leq \frac{0.355}{\Delta(G) + 1}$$

for all v and ϕ.

Phase transitions

- Correlation decay method, for counting independent sets, matchings, graph coloring etc., is based on absence of phase transition, i.e., uniqueness of Gibss measure.

Phase transitions

- Correlation decay method, for counting independent sets, matchings, graph coloring etc., is based on absence of phase transition, i.e., uniqueness of Gibss measure.
- The method presented here is also based on absence of phase transition, i.e., via the Lee-Yang theorem no complex zeros $\Rightarrow$ no phase transition.

Comparison

- Correlation decay method yields an FPTAS, but currently only seems to work for positive real numbers, i.e., # weighted independent sets with weight $\lambda > 0$, partition function of symmetric matrices A with $A_{i,j} > 0$, the chromatic polynomial at positive integers.

Comparison

- Correlation decay method yields an FPTAS, but currently only seems to work for positive real numbers, i.e., # weighted independent sets with weight $\lambda > 0$, partition function of symmetric matrices A with $A_{i,j} > 0$, the chromatic polynomial at positive integers.
- The method presented here only seems to yield a QPTAS, but works for complex numbers: partition function of complex valued symmetric matrices/edge-coloring models, the Tutte/chromatic polynomial at a complex number, etc.

Comparison

- Correlation decay method yields an FPTAS, but currently only seems to work for positive real numbers, i.e., # weighted independent sets with weight $\lambda > 0$, partition function of symmetric matrices A with $A_{i,j} > 0$, the chromatic polynomial at positive integers.
- The method presented here only seems to yield a QPTAS, but works for complex numbers: partition function of complex valued symmetric matrices/edge-coloring models, the Tutte/chromatic polynomial at a complex number, etc.
- Partition functions of complex edge-coloring models make sense!

Further work

Together with Alexander Barvinok and Viresh Patel:

- Try to push QPTAS to FPTAS: faster computation of derivatives indicates that this can be done in certain cases: partition functions of complex edge-coloring models/symmetric matrices and Tutte polynomial on bounded degree graphs!
- Try to find larger zero-free regions. Also other shapes than disks.

Thank you for your attention!