How to count when you can sample

Daniel Štefankovič
(University of Rochester)


Approximation algorithms for the normalizing constant of Gibbs distributions, Huber, AAP, 2015.

Outline

1. Counting problems
2. Monte Carlo, SCV
3. Product method
4. Temperature
5. Cooling schedules
6. More…
Counting

- independent sets
- spanning trees
- matchings
- perfect matchings
- k-colorings
Counting

- independent sets
- spanning trees
- matchings
- perfect matchings
- k-colorings
Compute the number of spanning trees
Compute the number of spanning trees

Kirchhoff’s Matrix Tree Theorem: \( \det(D - A)_{vv} \)
Compute the number of spanning trees of $G$ using a polynomial-time algorithm.
Counting

- independent sets
- spanning trees
- matchings
- perfect matchings
- k-colorings
Compute the number of independent sets (hard-core gas model)

Independent set of a graph = subset $S$ of vertices, no two in $S$ are neighbors
# independent sets = 7

independent set = subset S of vertices
no two in S are neighbors
Compute the number of independent sets of \( G \) using a polynomial-time algorithm.
Compute the number of independent sets of \( G \) via a polynomial-time algorithm (unlikely)
graph \ G \mapsto \# \text{ independent sets in } \ G

\#P-complete

\#P-complete even for 3-regular graphs

(Dyer, Greenhill, 1997)
graph $G \mapsto \# \text{ independent sets in } G$

approximation

randomization
We would like to know $Q$

Goal: random variable $Y$ such that

$$P\left( (1-\varepsilon)Q \leq Y \leq (1+\varepsilon)Q \right) \geq 1-\delta$$

(fully polynomial randomized approximation scheme):

FPRAS:

$$G, \varepsilon, \delta \rightarrow \text{polynomial-time algorithm} \rightarrow Y$$
Outline

1. Counting problems
2. Monte Carlo, SCV
3. Product method
4. Temperature
5. Cooling schedules
6. More…
Theorem (Monte Carlo)

Using $n$ samples from $X$ we can construct $Y$ such that

$$P\left( (1-\varepsilon)E[X] \leq Y \leq (1+\varepsilon)E[X] \right) \geq 1-\delta$$

$$n = \Theta\left( \frac{V[X]}{E[X]^2} \cdot \frac{1}{\varepsilon^2} \cdot \ln \left( \frac{1}{\delta} \right) \right)$$

squared coefficient of variation (SCV)
Squared Coefficient of Variation

$$SCV(X) = \frac{V[X]}{E[X]^2} = \frac{E[X^2]}{E[X]^2} - 1$$
Squared Coefficient of Variation

\[ X, X_1, \ldots, X_t \text{ i.i.d.} \]

\[ Y = \frac{X_1 + \cdots + X_t}{t} \]

\[ \text{SCV}(Y) = \frac{\text{SCV}(X)}{t} \]

\[ \text{SCV}(X) = \frac{\text{Var}(X)}{\text{E}[X]^2} = \frac{\text{E}[X^2]}{\text{E}[X]^2} - 1 \]
Squared Coefficient of Variation

\[ X, X_1, \ldots, X_t \text{ i.i.d.} \]

\[ Y = X_1. \ldots X_t \]

\[ \text{SCV}(Y) = (1 + \text{SCV}(X_1)) \ldots (1 + \text{SCV}(X_t)) - 1 \]

\[ \text{SCV}(X) = \frac{\text{Var}[X]}{E[X]^2} = \frac{E[X^2]}{E[X]^2} - 1 \]
Squared Coefficient of Variation

\[ X, X_1, \ldots, X_t \text{ i.i.d.} \]

\[ Y = X_1, \ldots, X_t \]

\[ \text{SCV}(Y) = (1 + \text{SCV}(X_1)) \ldots (1 + \text{SCV}(X_t)) - 1 \]

\[ \text{SCV}(X_i) \leq \frac{\varepsilon^2}{4t} \Rightarrow \text{SCV}(Y) \leq \frac{\varepsilon^2}{4} \]
Theorem (Monte Carlo)

Using $n$ samples from $X$ we can construct $Y$ such that

$$P\left( (1-\varepsilon)E[X] \leq Y \leq (1+\varepsilon)E[X] \right) \geq 1-\delta$$

$$n = \Omega\left( \frac{V[X]}{E[X]^2} \frac{1}{\varepsilon^2 \ln(1/\delta)} \right)$$

The Bienaymé-Chebyshev inequality

Let $X_1,\ldots,X_n$ be independent, identically distributed random variables, $Q = E[X]$. Let

$$Y = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

Then

$$P( Y \text{ gives (1±ε)-estimate of } Q ) \geq 1 - \frac{V[X]}{n E[X]^2} \frac{1}{\varepsilon^2}$$

Chernoff’s bound

Let $X_1,\ldots,X_n$ be independent, identically distributed random variables, $0 \leq X \leq 1$, $Q = E[X]$. Let

$$Y = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

Then

$$P( Y \text{ gives (1±ε)-estimate of } Q ) \geq 1 - e^{-\varepsilon^2 \cdot n \cdot E[X] / 3}$$

Median “boosting trick”

$$n = \frac{1}{E[X]} \frac{4}{\varepsilon^2}$$

BY BIENAYME-CHEBYSHEV:

$$P( \bullet \in \left[ (1-\varepsilon)Q, (1+\varepsilon)Q \right] ) \geq 3/4$$

Median trick – repeat 2T times

BY BIENAYME-CHEBYSHEV:

$$P( \bullet \in \left[ (1-\varepsilon)Q, (1+\varepsilon)Q \right] ) \geq 3/4$$

BY CHERNOFF:

$$P( > T \text{ out of } 2T \bullet ) \geq 1 - e^{-T/4}$$

$$P( \text{median is in } \bullet ) \geq 1 - e^{-T/4}$$
Outline

1. Counting problems
2. Monte Carlo, SCV
3. **Product method**
4. Temperature
5. Cooling schedules
6. More…
GOAL: given a graph $G$, estimate the number of independent sets of $G$

Example $G$: \[ \begin{array}{c}
\end{array} \]

\# independent sets = \[ \frac{1}{P(\text{graph})} \]
P( ) =

\[ P( ) \times P( ) \times P( ) \times P( ) \]

\[ X_1 \quad X_2 \quad X_3 \quad X_4 \]

\[ X_i \in [0,1] \text{ and } E[X_i] \geq \frac{1}{2} \Rightarrow \frac{V[X_i]}{E[X_i]^2} = O(1) \]

\[ P(\text{A} \cap \text{B}) = P(\text{A})P(\text{B}|\text{A}) \]
FPRAS using $O(n^2)$ samples.
Outline

1. Counting problems
2. Monte Carlo, SCV
3. Product method
4. Temperature
5. Cooling schedules
6. More…
easy = hot

"permissive IS"

hard = cold
Hamiltonian

“permissive IS”
hard = hot

easy = cold

“hardcore model”
Hamiltonian

hard = hot

easy = cold

“hardcore model”

0  1  2
Hamiltonian

\[ H : \Omega \rightarrow \{0, \ldots, n\} \]

Goal: estimate \( |H^{-1}(0)| \)

using the product method, that is,

\[ |H^{-1}(0)| = E[X_1] \ldots E[X_t] \]
Gibbs distributions

Distributions between hot and cold

\[ \beta = \text{inverse temperature} \]

\[ \beta = 0 \quad \Rightarrow \quad \text{hot} \quad \Rightarrow \quad \text{uniform on } \Omega \]

\[ \beta = \infty \quad \Rightarrow \quad \text{cold} \quad \Rightarrow \quad \text{uniform on } H^{-1}(0) \]

\[ \mu_{\beta}(x) \propto \exp(-H(x) \beta) \]
Distributions between hot and cold

$$\mu_\beta(x) \propto \exp(-H(x)\beta)$$

$$\mu_\beta(x) = \frac{\exp(-H(x)\beta)}{Z(\beta)}$$

Normalizing factor = partition function

$$Z(\beta) = \sum_{x \in \Omega} \exp(-H(x)\beta)$$
Partition function

\[ Z(\beta) = \sum_{x \in \Omega} \exp(-H(x)\beta) \]

have: \[ Z(0) = |\Omega| \]
want: \[ Z(\infty) = |H^{-1}(0)| \]
Partition function - example

\[ Z(\beta) = \sum_{x \in \Omega} \exp(-H(x)\beta) \]

have: \[ Z(0) = |\Omega| \]
want: \[ Z(\infty) = |H^{-1}(0)| \]

\[ Z(\beta) = \]
\[ + 4 e^{-2.\beta} \]
\[ + 7 e^{-0.\beta} \]

\[ Z(0) = 16 \]
\[ Z(\infty) = 7 \]
Assumption:
we have a sampler oracle for $\mu_\beta$

$$\mu_\beta (x) = \frac{\exp(-H(x)\beta)}{Z(\beta)}$$

graph $G_\beta$ \[\xrightarrow{\text{SAMPLER ORACLE}}\] subset of $V$ from $\mu_\beta$
Outline

5. Cooling schedules

5a) ratio estimator
5b) cooling schedule (focus on individual SCVs)
5c) SCV of the ratio estimator
5d) paired product estimator
5e) static schedules
5f) existence of good schedules from 5b)

6. More…
Assumption: we have a sampler oracle for $\mu_\beta$

\[ \mu_\beta (x) = \frac{\exp(-H(x)\beta)}{Z(\beta)} \]

$W \sim \mu_\beta$
Assumption: we have a sampler oracle for $\mu_\beta$

$$\mu_\beta(x) = \frac{\exp(-H(x)\beta)}{Z(\beta)}$$

$W \sim \mu_\beta \rightarrow X = \exp(H(W)(\beta - \alpha))$
Assumption:
we have a sampler oracle for $\mu_\beta$

$$\mu_\beta (x) = \frac{\exp(-H(x)\beta)}{Z(\beta)}$$

$W \sim \mu_\beta \rightarrow X = \exp(H(W)(\beta - \alpha))$

can obtain the following ratio:

$$E[X] = \sum_{s \in \Omega} \mu_\beta(s) X(s) = \frac{Z(\alpha)}{Z(\beta)}$$
Our goal restated

Partition function

\[ Z(\beta) = \sum_{x \in \Omega} \exp(-H(x)\beta) \]

Goal: estimate \( Z(\infty) = |H^{-1}(0)| \)

\[ Z(\infty) = \frac{Z(\beta_1)}{Z(\beta_0)} \cdot \frac{Z(\beta_2)}{Z(\beta_1)} \cdots \frac{Z(\beta_t)}{Z(\beta_{t-1})} \cdot Z(0) \]

\( \beta_0 = 0 < \beta_1 < \beta_2 < \ldots < \beta_t = \infty \)
Outline

5. Cooling schedules

5a) ratio estimator
5b) cooling schedule (focus on individual SCVs)
5c) SCV of the ratio estimator
5d) paired product estimator
5e) static schedules
5f) existence of good schedules from 5b)

6. More…
Our goal restated

$$Z(\infty) = \frac{Z(\beta_1)}{Z(\beta_0)} \frac{Z(\beta_2)}{Z(\beta_1)} \ldots \frac{Z(\beta_t)}{Z(\beta_{t-1})} Z(0)$$

Cooling schedule:

$$\beta_0 = 0 < \beta_1 < \beta_2 < \ldots < \beta_t = \infty$$

How to choose the cooling schedule?

minimize length, while satisfying

$$\frac{V[X_i]}{E[X_i]^2} = O(1) \quad \frac{E[X_i]}{Z(\beta_i)} = \frac{Z(\beta_i)}{Z(\beta_{i-1})}$$
Our goal restated

\[
Z(\infty) = \frac{Z(\beta_1)}{Z(\beta_0)} \frac{Z(\beta_2)}{Z(\beta_1)} \cdots \frac{Z(\beta_t)}{Z(\beta_{t-1})} Z(\beta_0) Z(\beta_1) \cdots Z(\beta_t)
\]

Assumption:
we have a sampler oracle for \( \mu_\beta \)

\[
\mu_\beta (x) = \frac{\exp(-H(x)\beta)}{Z(\beta)}
\]

Cooling schedule:

\[
\beta_0 = 0 < \beta_1 < \beta_2 < \ldots < \beta_t = \infty
\]

How to choose the cooling schedule:
minimize length

\[
\frac{V[X_i]}{E[X_i]^2} = O(1) \quad E[X_i] = \frac{Z(\beta_i)}{Z(\beta_{i-1})}
\]

can obtain the following ratio:

\[
E[X] = \sum_{s \in \Omega} \mu_\beta(s) x(s) = \frac{Z(\alpha)}{Z(\beta)}
\]
Outline

5. Cooling schedules

5a) ratio estimator
5b) cooling schedule (focus on individual SCVs)
5c) SCV of the ratio estimator
5d) paired product estimator
5e) static schedules
5f) existence of good schedules from 5b)

6. More…
Express SCV using partition function (going from $\beta$ to $\alpha$)

$$ E[X] = \frac{Z(\alpha)}{Z(\beta)} $$

$$ W \sim \mu_\beta \rightarrow X = \exp(H(W)(\beta - \alpha)) $$

$$ \frac{E[X^2]}{E[X]^2} = \frac{Z(2\alpha-\beta)\ Z(\beta)}{Z(\alpha)^2} $$

$$ \frac{V[X]}{E[X]^2} + 1 $$
Outline

5. Cooling schedules

5a) ratio estimator
5b) cooling schedule (focus on individual SCVs)
5c) SCV of the ratio estimator
5d) paired product estimator
5e) static schedules
5f) existence of good schedules from 5b)

6. More…
Instead of
\[ \frac{Z(\alpha)}{Z(\beta)} \]

Do
\[ \frac{Z(\gamma)}{Z(\beta)} \frac{Z(\gamma)}{Z(\alpha)} \]
IDEA: Paired Product Estimator

(going from $\beta$ to $\gamma$) \[ E[X] = \frac{Z(\gamma)}{Z(\beta)} \]

(going from $\alpha$ to $\gamma$) \[ E[Y] = \frac{Z(\gamma)}{Z(\alpha)} \]

$W \sim \mu_{\beta} \rightarrow X = \exp(H(W)(\gamma - \alpha))$

$U \sim \mu_{\alpha} \rightarrow Y = \exp(H(U)(\gamma - \beta))$

\[
\frac{E[X^2]}{E[X]^2} \quad = \quad \frac{Z(\alpha) Z(\beta)}{Z(\gamma)^2} \quad = \quad \frac{E[Y^2]}{E[Y]^2}
\]
Outline

5. Cooling schedules

5a) ratio estimator
5b) cooling schedule (focus on individual SCVs)
5c) SCV of the ratio estimator
5d) paired product estimator
5e) static schedules
5f) existence of good schedules from 5b)

6. More…
Instead of

\[ \beta_0 < \beta_1 < \beta_2 < \ldots < \beta_t = \infty \]

Do

\[ E[X_0 X_1 \ldots X_t] / E[Y_0 Y_1 \ldots Y_t] \]
Outline

5. Cooling schedules

5a) ratio estimator
5b) cooling schedule (focus on individual SCVs)
5c) SCV of the ratio estimator
5d) paired product estimator
5e) static schedules
5f) existence of good schedules from 5b)

6. More…
Parameters: $A$ and $n$

\[ Z(\beta) = \sum_{x \in \Omega} \exp(-H(x)\beta) \]

\[ Z(0) = A \]

\[ H: \Omega \rightarrow \{0, \ldots, n\} \]

\[ Z(\beta) = \sum_{k=0}^{n} a_k e^{-\beta k} \]

\[ a_k = |H^{-1}(k)| \]
## Parameters

\[ Z(0) = A \]

\[ H: \Omega \rightarrow \{0, \ldots, n\} \]

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>independent sets</td>
<td>(2^V)</td>
<td>(E)</td>
</tr>
<tr>
<td>matchings</td>
<td>(\approx V!)</td>
<td>(V)</td>
</tr>
<tr>
<td>perfect matchings</td>
<td>(V!)</td>
<td>(V)</td>
</tr>
<tr>
<td>k-colorings</td>
<td>(k^V)</td>
<td>(E)</td>
</tr>
</tbody>
</table>
Previous cooling schedules

**Z(0) = A**

**H:Ω → {0,...,n}**

\[ \beta_0 = 0 < \beta_1 < \beta_2 < \ldots < \beta_t = \infty \]

"Safe steps"

\[ \beta \rightarrow \beta + 1/n \]

\[ \beta \rightarrow \beta (1 + 1/\ln A) \]

\[ \ln A \rightarrow \infty \]

Cooling schedules of length

\[ O( n \ln A) \]

\[ O( (\ln n) (\ln A) ) \]
Previous cooling schedules

\[ Z(0) = A \]

\[ \beta_0 = 0 < \beta_1 < \beta \]

“Safe steps”

\[ \beta \rightarrow \beta + 1/n \]

\[ \beta \rightarrow \beta (1 + 1/\ln A) \]

\[ \ln A \rightarrow \infty \]

Cooling schedules of length

\[ O(n \ln A) \]

\[ O( (\ln n) (\ln A) ) \]

Our goal restated

\[ Z(\infty) = \frac{Z(\beta_1)}{Z(\beta_0)} \frac{Z(\beta_2)}{Z(\beta_1)} \ldots \frac{Z(\beta_t)}{Z(\beta_{t-1})} Z(0) \]

Cooling schedule:

\[ \beta_0 = 0 < \beta_1 < \beta_2 < \ldots < \beta_t = \infty \]

How to choose the cooling schedule?

minimize length, while satisfying

\[ \frac{V[X_i]}{E[X_i]^2} = O(1) \quad E[X_i] = \frac{Z(\beta_i)}{Z(\beta_{i-1})} \]

(Bezáková, Štefankovič, Vigoda, V. Vazirani’06)
"Safe steps"

\[ \beta \rightarrow \beta + \frac{1}{n} \]
\[ \beta \rightarrow \beta \left(1 + \frac{1}{\ln A}\right) \]
\[ \ln A \rightarrow \infty \]

(Bezáková, Štefankovič, Vigoda, V.Vazirani’06)

\[ W \sim \mu_\beta \rightarrow X = \exp(H(W)(\beta - \alpha)) \]

\[ Z(\beta) = \sum_{k=0}^{n} a_k e^{-\beta k} \]

\[ \frac{1}{e} \leq X \leq 1 \]

\[ \frac{V[X]}{E[X]^2} \leq \frac{1}{E[X]} \leq e \]
“Safe steps”

\[ \beta \rightarrow \beta + \frac{1}{n} \]

\[ \beta \rightarrow \beta \ (1 + \frac{1}{\ln A}) \]

\[ \ln A \rightarrow \infty \]

(Bezáková, Štefankovič, Vigoda, V. Vazirani’06)

\[ W \sim \mu_\beta \rightarrow X = \exp(H(W)(\beta - \alpha)) \]

\[ Z(\beta) = \sum_{k=0}^{n} a_k e^{-\beta^k} \]

\[ Z(\infty) = a_0 \geq 1 \]

\[ Z(\ln A) \leq a_0 + 1 \]

\[ \mathbb{E}[X] \geq 1/2 \]
“Safe steps”

\[ \beta \to \beta + \frac{1}{n} \]
\[ \beta \to \beta \left(1 + \frac{1}{\ln A}\right) \]
\[ \ln A \to \infty \]

(Bezáková, Štefankovič, Vigoda, V. Vazirani’06)

\[ W \sim \mu_\beta \rightarrow X = \exp(H(W)(\beta - \alpha)) \]

\[ Z(\beta) = \sum_{k=0}^{n} a_k \ e^{-\beta^k} \]

\[ E[X] \geq \frac{1}{2e} \]
No better fixed schedule possible

\[ Z(0) = A \quad H : \Omega \rightarrow \{0, \ldots, n\} \]

**THEOREM:**

A schedule that works for all (with \( a \in [0, A-1] \))

\[ Z_a(\beta) = \frac{A}{1+a} \left( 1 + a e^{-\beta n} \right) \]

has LENGTH \( \geq \Omega( (\ln n)(\ln A) ) \)
5. Cooling schedules

5a) ratio estimator
5b) cooling schedule (focus on individual SCVs)
5c) SCV of the ratio estimator
5d) paired product estimator
5e) static schedules
5f) existence of good schedules from 5b)

6. More…

length $O^* \left( \ln A \right)^{1/2}$
Cooling schedule (definition refresh)

\[ Z(\infty) = \frac{Z(\beta_1)}{Z(\beta_0)} \frac{Z(\beta_2)}{Z(\beta_1)} \ldots \frac{Z(\beta_t)}{Z(\beta_{t-1})} Z(0) \]

Cooling schedule:

\[ \beta_0 = 0 < \beta_1 < \beta_2 < \ldots < \beta_t = \infty \]

How to choose the cooling schedule?

minimize length, while satisfying

\[ \frac{V[X_i]}{E[X_i]^2} = O(1) \quad E[X_i] = \frac{Z(\beta_i)}{Z(\beta_{i-1})} \]
\[
\frac{E[X^2]}{E[X]^2} = \frac{Z(\alpha) \ Z(\beta)}{Z(\gamma)^2} \leq C
\]

\[
f(\gamma) = \ln Z(\gamma)
\]

\[
(f(\alpha) + f(\beta))/2 \leq (\ln C)/2 + f(\gamma)
\]

\[
\leq C' = (\ln C)/2
\]
Properties of *partition functions*

\[ f(\gamma) = \ln Z(\gamma) \]

- \( f \) is decreasing
- \( f \) is convex
- \( f'(0) \geq -n \)
- \( f(0) \leq \ln A \)
Properties of partition functions

\[ f(\gamma) = \ln Z(\gamma) \]

- \( f \) is decreasing
- \( f \) is convex
- \( f'(0) \geq -n \)
- \( f(0) \leq \ln A \)

\[ f(\beta) = \ln \sum_{k=0}^{n} a_k e^{-\beta k} \]

\[ (\ln f)' = \frac{f'}{f} \]

\[ f'(\beta) = \frac{-\sum_{k=0}^{n} a_k k e^{-\beta k}}{\sum_{k=0}^{n} a_k e^{-\beta k}} \]
GOAL: proving Lemma:
for every partition function there exists a cooling schedule of length $O^*((\ln A)^{1/2})$

\[ f(\gamma) = \ln Z(\gamma) \]

- $f$ is decreasing
- $f$ is convex
- $f'(0) \geq -n$
- $f(0) \leq \ln A$

Proof:

Let $K := \Delta f$

Then

\[ \Delta (\ln |f'|) \geq \frac{1}{K} \]
Proof:

Let $K := \Delta f$

Then

$$\Delta (\ln |f'|) \geq \frac{1}{K}$$

$c := (a+b)/2, \quad \Delta := b-a$

have

$f(c) = (f(a)+f(b))/2 - 1$

$f$ is convex

$$(f(a) - f(c)) / \Delta \leq f'(a)$$

$$(f(c) - f(b)) / \Delta \geq f'(b)$$
Let $K := \Delta f$

Then

$$\Delta (\ln |f'|) \geq \frac{1}{K}$$

$c := (a+b)/2$, $\Delta := b-a$

have $f(c) = (f(a)+f(b))/2 - 1$

$f$ is convex

$$\frac{f''(b)}{f''(a)} \leq 1 - \frac{1}{\Delta f} \leq e^{-\Delta f}$$

$$(f(a) - f(c)) / \Delta \leq f'(a)$$

$$(f(c) - f(b)) / \Delta \geq f'(b)$$
$f: [a,b] \rightarrow \mathbb{R}$, convex, decreasing can be “approximated” using $f'(a) \frac{f'(b)}{f'(b)} (f(a)-f(b))$ segments
Outline

6. More

6a) **cooling schedule** (focus on product SCVs)
6b) constructing the schedule
6c) warm starts
\[ \beta_0 = 0 < \beta_1 < \beta_2 < ... < \beta_t = \infty \]

\[ f(\gamma) = \ln Z(\gamma) \]

Do

\[ \beta_0 \rightarrow \gamma_1 \leftarrow \beta_1 \rightarrow \gamma_2 \leftarrow \beta_2 \rightarrow \gamma_3 \leftarrow \beta_3 \rightarrow \gamma_4 \leftarrow ... \]

\[ E[X_0 X_1 \ldots X_t] / E[Y_0 Y_1 \ldots Y_t] \]

\[ SCV(X_0 X_1 \ldots X_t) = -1 + \exp(2 (\delta_0 + \delta_1 + ... + \delta_t)) \]

\[ \delta_i = (f(\beta_{i+1}) + f(\beta_i))/2 - f(\gamma_{i+1}) \]

Squared Coefficient of Variation

\[ X, X_1, \ldots, X_t \text{ i.i.d.} \]
\[ Y = X_1, \ldots, X_t \]

\[ SCV(Y) = (1 + SCV(X_i)) \ldots (1 + SCV(X_i)) - 1 \]

\[ SCV(X) = \frac{V[X]}{E[X]^2} = \frac{E[X^2]}{E[X]^2} - 1 \]

\[ \frac{E[X^2]}{E[X]^2} = \frac{Z(\alpha) Z(\beta)}{Z(\gamma)^2} \leq C \]

\[ f(\gamma) = \ln Z(\gamma) \]

\[ \frac{(f(\alpha) + f(\beta))/2 \leq (\ln C)/2 + f(\gamma)}{\leq C' = (\ln C)/2} \]
\[ \text{SCV}(X_0X_1\ldots X_t) = -1 + \exp(2 (\delta_0 + \delta_1 + \ldots + \delta_t)) \]

\[ \delta_i = \frac{(f(\beta_{i+1}) + f(\beta_i))}{2} - f(\gamma_{i+1}) \]

Claim (Huber’15):

If \( f(\beta_i) - f(\beta_{i+1}) \leq 1/(\ln Cn) \) then

\[ \text{SCV}(X_0X_1\ldots X_t) \leq 2e \]
Outline

6. More

6a) cooling schedule  (focus on product SCVs)
6b) constructing the schedule
6c) warm starts
TPA (Huber-Schott’2010)

\[ \beta \leftarrow 0 \]
repeat

sample \( W \sim \mu_{\beta} \)

sample \( S \sim \exp(1) \)

\[ \beta \leftarrow \beta + S/H(W) \]

until \( \beta = \infty \)

\[ \beta_0 = 0 < \beta_1 < \beta_2 < \ldots < \beta_t = \infty \]
Claim: 

\[ f(\beta_1), f(\beta_2), \ldots f(\beta_{t-1}) \text{ is a Poisson point process (PPP) on } [f(0), f(\infty)] \]

To generate the schedule 
* take union of several PPPs 
* leave out every k-th point

TPA (Huber-Schott’2010)

\[ \beta \leftarrow 0 \]
repeat
- sample \( W \sim \mu_\beta \)
- sample \( S \sim \exp(1) \)
- \( \beta \leftarrow \beta + S/H(W) \)
until \( \beta = \infty \)

\[ Z(\beta) = \sum_{x \in \Omega} \exp(-H(x)\beta) \]

\[ f(\beta) = \ln Z(\beta) \]
Claim:

\[ f(\beta_1), f(\beta_2), \ldots, f(\beta_{t-1}) \text{ is PPP}(1) \text{ on } [f(0), f(\infty)] \]

\[
P(\beta + S/H(W) > \alpha) = \frac{\sum_{x \in \Omega} \exp(-H(x)\beta)}{Z(\beta)} P(S > (\alpha-\beta) H(W)) = \frac{Z(\alpha)}{Z(\beta)}
\]

\[
P(S > (\alpha-\beta) H(W)) = \sum_{x \in \Omega} \exp(-H(x)\beta)
\]

TPA (Huber-Schott’2010)

\[
\beta \leftarrow 0 \\
\text{repeat} \\
\text{sample } W \sim \mu_\beta \\
\text{sample } S \sim \exp(1) \\
\beta \leftarrow \beta + S/H(W) \\
\text{until } \beta = \infty
\]

\[\beta_0 < \beta_1 < \beta_2 < \ldots < \beta_t = \infty\]
Outline

6. More

6a) cooling schedule (focus on product SCVs)
6b) constructing the schedule
6c) warm starts
Mixing time:
\[ \tau_{\text{mix}} = \text{smallest } t \text{ such that } |\mu_t - \pi|_{TV} \leq 1/e \]

Relaxation time:
\[ \tau_{\text{rel}} = 1/(1-\lambda_2) \]

\[ \tau_{\text{rel}} \leq \tau_{\text{mix}} \leq \tau_{\text{rel}} \ln \left(1/\pi_{\text{min}}\right) \]

(discrepancy may be substantially bigger for, e.g., matchings)
Mixing time:
\[ \tau_{\text{mix}} = \text{smallest } t \text{ such that } |\mu_t - \pi|_{TV} \leq 1/e \]

Relaxation time:
\[ \tau_{\text{rel}} = 1/(1-\lambda_2) \]

Estimating \( \pi(S) \)
\[ X \sim \pi \]
\[ Y = \begin{cases} 
1 & \text{if } X \in S \\
0 & \text{otherwise} 
\end{cases} \]
\[ E[Y] = \pi(S) \]

METHOD 1

\[ X_1 \]
\[ X_2 \]
\[ X_3 \]
\[ \vdots \]
\[ X_s \]
Mixing time: 
\[ \tau_{\text{mix}} = \text{smallest } t \text{ such that } |\mu_t - \pi|_{TV} \leq 1/e \]

Relaxation time: 
\[ \tau_{\text{rel}} = \frac{1}{1-\lambda_2} \]

Estimating \( \pi(S) \)

- \( X \sim \pi \)
- \( Y = \begin{cases} 1 & \text{if } X \in S \\ 0 & \text{otherwise} \end{cases} \)
- \( E[Y] = \pi(S) \)

Method 1

- \( X_1 \)
- \( X_2 \)
- \( X_3 \)
- \( \vdots \)
- \( X_s \)

Method 2

- \( X_1 \)
- \( X_2 \)
- \( X_3 \)
- \( \ldots \)
- \( X_s \)

(Gillman’98, Kahale’96, ...)

...
Further speed-up

Mixing time:
\[ \tau_{\text{mix}} = \text{smallest } t \text{ such that } \left| \mu_t - \pi \right|_{TV} \leq \frac{1}{e} \]

Relaxation time:
\[ \tau_{\text{rel}} = \frac{1}{1 - \lambda_2} \]

\[ \left| \mu_t - \pi \right|_{TV} \leq \exp\left(-t/\tau_{\text{rel}}\right) \text{Var}_{\pi}(\mu_0/\pi) \]

\[ \left( \sum \pi(x)(\mu_0(x)/\pi(x)-1)^2 \right)^{1/2} \]

small \( \Rightarrow \) called warm start

METHOD 2
(Gillman’98, Kahale’96, …)
Mixing time:

\[ \tau_{\text{mix}} = \text{smallest } t \text{ such that } |\mu_t - \pi|_{TV} \leq \frac{1}{e} \]

Relaxation time:

\[ \tau_{\text{rel}} = \frac{1}{1 - \lambda^2} \]

Further speed-up

Sample at \( \beta \) can be used as a warm start for \( \beta' \)

\[ \iff \]

Cooling schedule can step from \( \beta' \) to \( \beta \)

\[ |\mu_t - \pi|_{TV} \leq \exp\left(-\frac{t}{\tau_{\text{rel}}}\right) \text{Var}_\pi\left(\frac{\mu_0}{\pi}\right) \]

\[ \left(\sum \pi(x)(\frac{\mu_0(x)}{\pi(x)}-1)^2\right)^{1/2} \]

Small \( \Rightarrow \) called warm start

METHOD 2

(Gillman’98, Kahale’96, ...)

\[ X_1 \quad X_2 \quad X_3 \quad \cdots \quad X_s \]
sample at $\beta$ can be used as a warm start for $\beta'$

$\iff$

cooling schedule can step from $\beta'$ to $\beta$

$m = O\left( (\ln n)(\ln A) \right)$

\[
Z(0) = A \quad H : \Omega \rightarrow \{0, \ldots, n\}
\]

$\beta_0 = 0 < \beta_1 < \beta_2 < \ldots < \beta_t = \infty$

"Safe steps"

$\beta \rightarrow \beta + 1/n$

$\beta \rightarrow \beta \left( 1 + \frac{1}{\ln A} \right)$

$\ln A \rightarrow \infty$

Cooling schedules of length

$O\left( n \ln A \right)$

$O\left( (\ln n)(\ln A) \right)$

(Bezáková, Štefankovič, Vigoda, V.Vazirani'06)

\[
= \text{“well mixed” states}
\]
run the our cooling-schedule algorithm with METHOD 2 using “well mixed” states as starting points
Output of our algorithm:

$\beta_0 \rightarrow \beta_1 \rightarrow \beta_2 \rightarrow \beta_3 \rightarrow \cdots \rightarrow \beta_m$

Small augmentation (so that we can use sample from current $\beta$ as a warm start at next) still $O^*(\ln A^{1/2})$

$k = O^*(\ln A^{1/2})$

Use analogue of Frieze-Dyer for independent samples from vector variables with slightly dependent coordinates.