

# **An FPTAS for #Knapsack and Related Counting Problems**

**Parikshit Gopalan**

**Adam Klivans**

**Raghu Meka**

**Daniel Štefankovič**

**Santosh Vempala**

**Eric Vigoda**

# An FPTAS for #Knapsack and Related Counting Problems

Parikshit Gopalan, Adam Klivans, Raghu Meka  
Daniel Štefankovič, Santosh Vempala, Eric Vigoda

The image displays two screenshots of the arXiv.org website, side-by-side, showing the abstracts of the two papers mentioned in the title. Both screenshots are from the Cornell University Library and feature a search bar at the top with the text "arXiv.org > cs > arXiv:1008.3187" and "arXiv.org > cs > arXiv:1008.1687" respectively. The left screenshot shows the abstract for "Polynomial-Time Approximation Schemes for Knapsack and Related Counting Problems using Branching Programs" by Parikshit Gopalan, Adam Klivans, and Raghu Meka, submitted on 18 Aug 2010. The right screenshot shows the abstract for "A Deterministic Polynomial-time Approximation Scheme for Counting Knapsack Solutions" by Daniel Štefankovič, Santosh Vempala, and Eric Vigoda, submitted on 10 Aug 2010. Both abstracts describe algorithms for counting solutions to knapsack problems with a constant number of constraints and contingency tables. The right abstract also includes a "Submission history" section and a "Which authors of this paper are endorsers?" link.

# What **can** be counted?

(in **polynomial-time**)

**exactly?**

**very little...**

number of spanning trees (using determinant),  
Kirchoff'1847.

perfect matchings in planar graphs  
(using Pfaffians), Kasteleyn'1960.

(rest: usually #P-hard)

# What **can** be counted?

(in **polynomial-time**)

**approximately?      a little more...**

perfect matchings in bipartite graphs (permanent of non-negative matrices), Jerrum, Sinclair, Vigoda'2001.

Ferromagnetic Ising model, Jerrum, Sinclair'1989.

Independent sets ( $\Delta \leq 5$ ), Weitz'2004.

k-colorings ( $k \geq (11/6)\Delta$ ), Vigoda'1999.

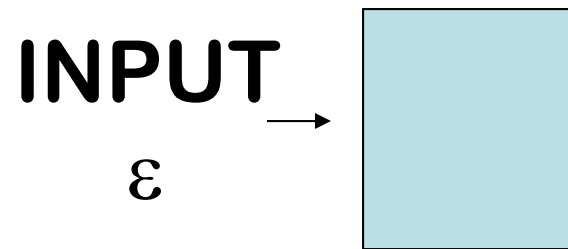
....

(approximate counting  $\approx$  random sampling,  
Jerrum, Valiant, Vazirani'1986)

# Approximate counting

(in polynomial-time)

**deterministic:**

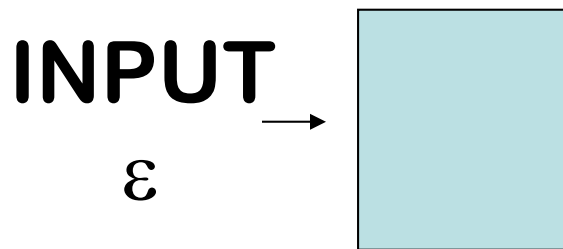


→ OUT

$$1 - \varepsilon \leq \frac{\text{OUT}}{Q} \leq 1 + \varepsilon$$

---

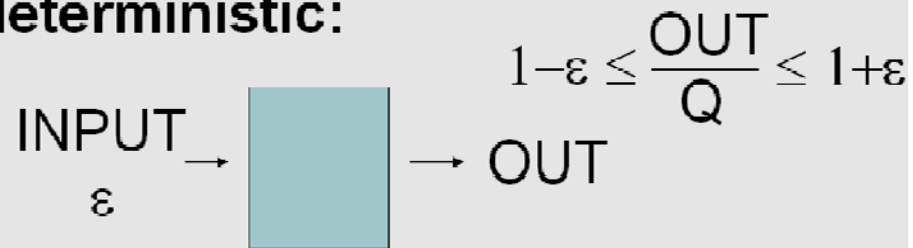
**randomized:**



→ OUT

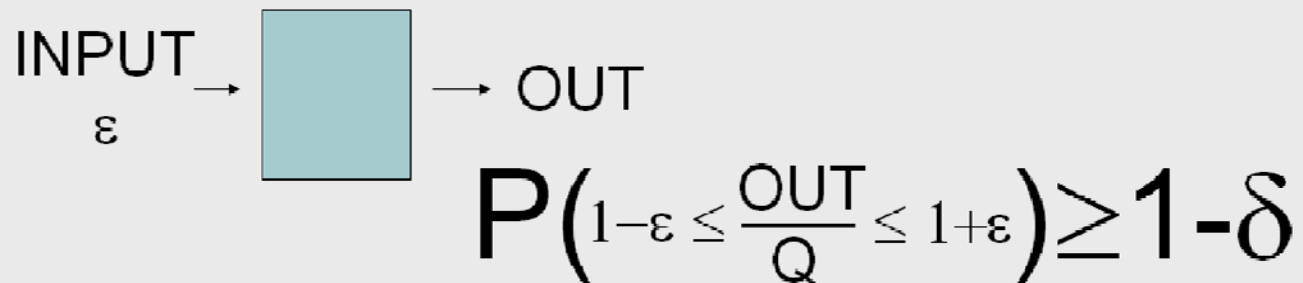
$$\mathbf{P}\left(1 - \varepsilon \leq \frac{\text{OUT}}{Q} \leq 1 + \varepsilon\right) \geq \mathbf{1 - \delta}$$

**deterministic:**

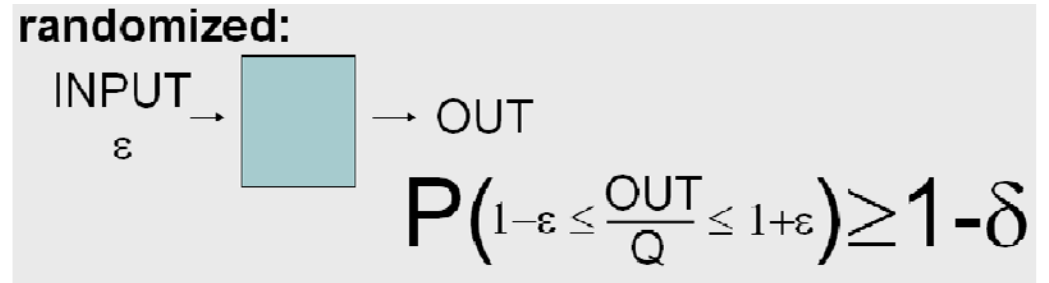
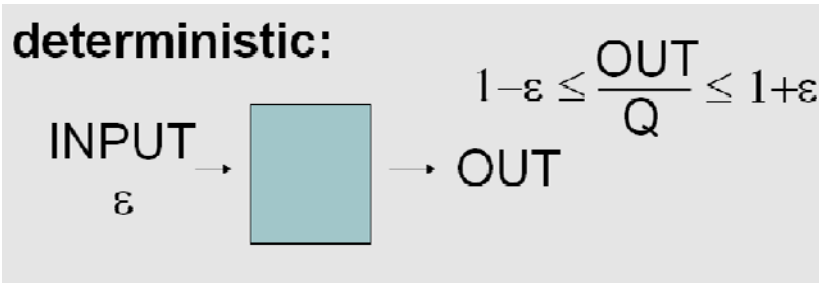


**not too many examples: independent sets in degree  $\leq 5$  graphs (Weitz'2004), matchings in bounded degree graphs (Bayati, Gamarnik, Katz, Nair, Tetali'2007), satisfying assignments of DNF formulas with terms of size  $\leq C$  (Ajtai, Wigderson'1985)**

**randomized:**



**more examples; Monte Carlo, usually using a Markov chain (dependence  $1/\varepsilon^2$ )**



**1) is randomness necessary ?**

**Is  $P = BPP$  ?**

**Primes  $\in P$  (Agarwal, Kayal, Saxena 2001)**

**2) dependence on  $\varepsilon$  ?**

**Monte Carlo  $\longrightarrow \Theta(1/\varepsilon^2)$**

# Knapsack (optimization)

INPUT:

$(w_1, v_1), \dots, (w_n, v_n), L$  (integers)

The diagram shows the word "weights" with two arrows pointing to  $w_1$  and  $w_n$  in the input tuple. Similarly, the word "values" has two arrows pointing to  $v_1$  and  $v_n$  in the input tuple.

OUTPUT:

$S \subseteq [n]$

$$\max \sum_{i \in S} v_i$$

$$\sum_{i \in S} w_i \leq L$$



# Dynamic program #1 (L is small)

$T[i, M]$

(optimal solution with  
items  $1, \dots, i$  and limit  $M$ )

$$T[i, M] = \max \begin{cases} T[i-1, M] \\ T[i-1, M-w_i] + v_i \end{cases}$$

## Knapsack (optimization)

INPUT:  $(w_1, v_1), \dots, (w_n, v_n), L$  (integers)

OUTPUT:

$S \subseteq [n]$

$$\max \sum_{i \in S} v_i$$

$$\sum_{i \in S} w_i \leq L$$

# Dynamic program #2 ( $v_i$ 's are small)

$$T[i, V]$$

(smallest weight of  
a subset of  $1, \dots, i$ ,  
with value  $\geq V$ )

$$T[i, V] = \min \begin{cases} T[i-1, V] \\ T[i-1, V - v_i] + w_i \end{cases}$$

 approximation algorithm

## Knapsack (optimization)

INPUT:  $(w_1, v_1), \dots, (w_n, v_n), L$  (integers)

OUTPUT:  $S \subseteq [n]$   $\max \sum_{i \in S} v_i$   
 $\sum_{i \in S} w_i \leq L$

# Counting knapsack

INPUT:

$$w_1, \dots, w_n, L$$

OUTPUT: How many  $S \subseteq [n]$

with

$$\sum_{i \in S} w_i \leq L$$

are there?

#P-hard

# Counting knapsack

Dyer, Frieze, Kannan, Kapoor, Perkovic,  
Vazirani'1993

$\exp(O^*(n^{1/2})) / \varepsilon^2$

randomized approximation algorithm

Morris, Sinclair'1999

$O(n^c / \varepsilon^2)$

randomized approximation algorithm  
(MCMC, canonical paths)

Dyer'2003

$O(n^{2.5} + n^2/\varepsilon^2)$

randomized approximation algorithm  
(dynamic programming)

**OURS:  $O^*(n^3/\varepsilon)$**

# Dyer'2003:

## $T[i, M]$

(number of solutions  
with items  $1, \dots, i$  and  
limit  $M$ )

### Counting knapsack

INPUT:

$w_1, \dots, w_n, L$

OUTPUT:

How many  $S \subseteq [n]$


with

$$\sum_{i \in S} w_i \leq L$$

are there?

#P-hard

$$T[i, M] = T[i-1, M] + T[i-1, M-w_i]$$

+ rejection sampling  approximate counter

+ rejection sampling



approximate counter

$$w_i' = \frac{n^2 w_i}{L} \quad L' = n^2 L$$

rounding:

$$w_i'' = \lfloor w_i' \rfloor$$

1) get more solutions,  $\Omega'' \supseteq \Omega'$

2) not too many more,  $|\Omega''| \leq (n+1)|\Omega'|$

Proof:

$S'' \in \Omega'' - \Omega'$ ,  $X$  heaviest in  $S''$ , then  $S'' - \{X\} \in \Omega'$

Dyer'2003:

$$T[i, M]$$

(number of solutions with items 1, ..., i and limit M)

$$T[i, M] = T[i-1, M] + T[i-1, M-w_i]$$

+ rejection sampling  approximate counter

Counting knapsack

INPUT:

$$w_1, \dots, w_n, L$$

OUTPUT:

How many  $S \subseteq [n]$

with

$$\sum_{i \in S} w_i \leq L$$

are there?

#P-hard

# Our dynamic program

deterministic approximation algorithm

$\tau(i, A) =$  smallest  $M$  such that knapsack with  $w_1, \dots, w_i, M$  has  $\geq A$  solutions

$$\tau(i, A) = \min_{\alpha \in [0, 1]} \max \begin{cases} \tau(i-1, \alpha A) \\ \tau(i-1, (1-\alpha) A) + w_i \end{cases}$$

$$Q = 1 + \varepsilon/(n+1)$$

$$s = \lceil n \log_Q 2 \rceil$$

$$T[0..n, 0..s]$$

## Our dynamic program

deterministic approximation algorithm

$\tau(i, A)$  = smallest  $M$  such that knapsack with  $w_1, \dots, w_i, M$  has  $\geq A$  solutions

$$\tau(i, A) = \min_{\alpha \in [0, 1]} \max \begin{cases} \tau(i-1, \alpha A) \\ \tau(i-1, (1-\alpha) A) + w_i \end{cases}$$

$$T(i, j) = \min_{\alpha \in [0, 1]} \max \begin{cases} T(i-1, \lfloor j + \ln_Q \alpha \rfloor) \\ T(i-1, \lfloor j + \ln_Q (1-\alpha) \rfloor) + w_i \end{cases}$$

**Lemma 1:**  $\tau(i, Q^{j-i}) \leq T[i, j] \leq \tau(i, Q^j)$



$$T(i,j) = \min_{\alpha \in [0,1]} \max \begin{cases} T(i-1, \lfloor j + \ln_Q \alpha \rfloor) \\ T(i-1, \lfloor j + \ln_Q (1-\alpha) \rfloor) + w_i \end{cases}$$

Lemma 2: can compute recursion efficiently

only few values of  $\alpha$  matter

$$Q^{-j}, \dots, Q^0, 1-Q^0, \dots, 1-Q^j$$

can use binary search

$$\text{TOTAL RUN TIME} = O\left(\frac{n^3}{\epsilon} \log(n/\epsilon)\right)$$

# How to deal with more constraints ?

(e.g., contingency tables, multi-dimensional knapsack, ...)

multi-dimensional knapsack:

How many  $S \subseteq [n]$

with

$$\sum_{i \in S} w_{j,i} \leq L_j \quad j \in \{1, \dots, k\}$$

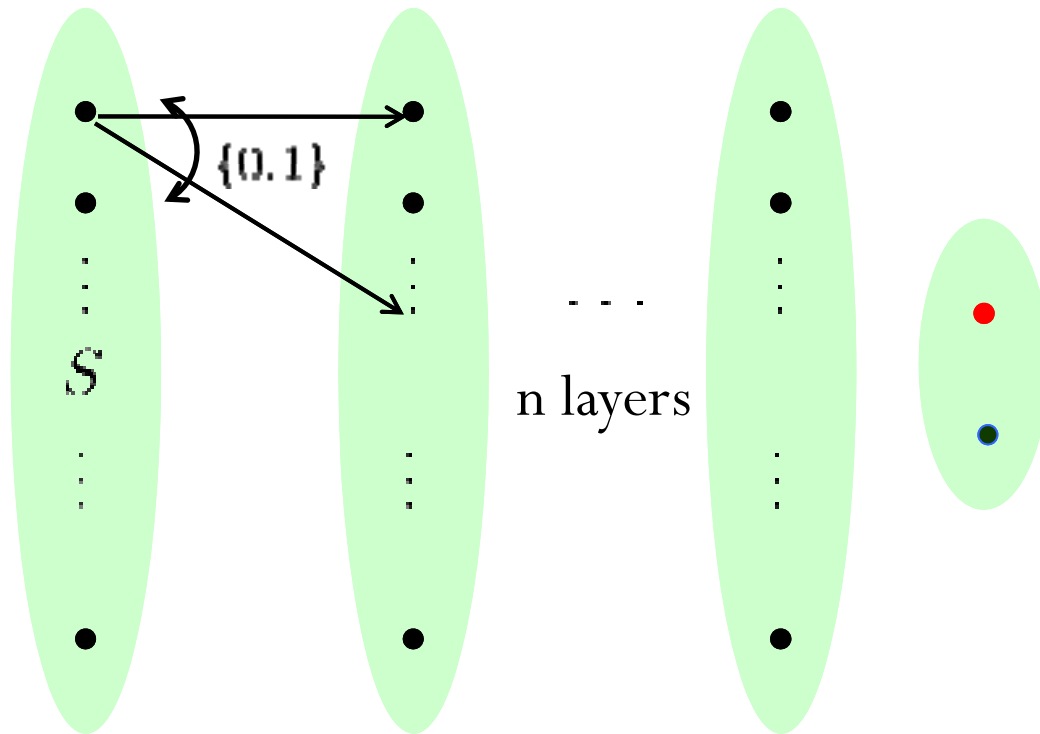
are there?

$O((n/\varepsilon)^{O(k^2)} \log W)$  algorithm

The screenshot shows the arXiv.org interface for the paper "Polynomial-Time Approximation Schemes for Knapsack and Related Counting Problems using Branching Programs" by Parikshit Gopalan, Adam Kivavans, and Raghu Meka. The page includes the Cornell University Library logo, a search bar, and navigation links. The abstract discusses a deterministic polynomial-time algorithm for counting solutions to knapsack problems with multiple constraints. It mentions that the algorithm produces a  $(1 + \epsilon)$ -multiplicative approximation in time  $\text{poly}(n, \log W, 1/\epsilon)$ . The paper also provides information on how to download the paper (PDF, PostScript, etc.) and lists references and citations.

# Read once branching programs

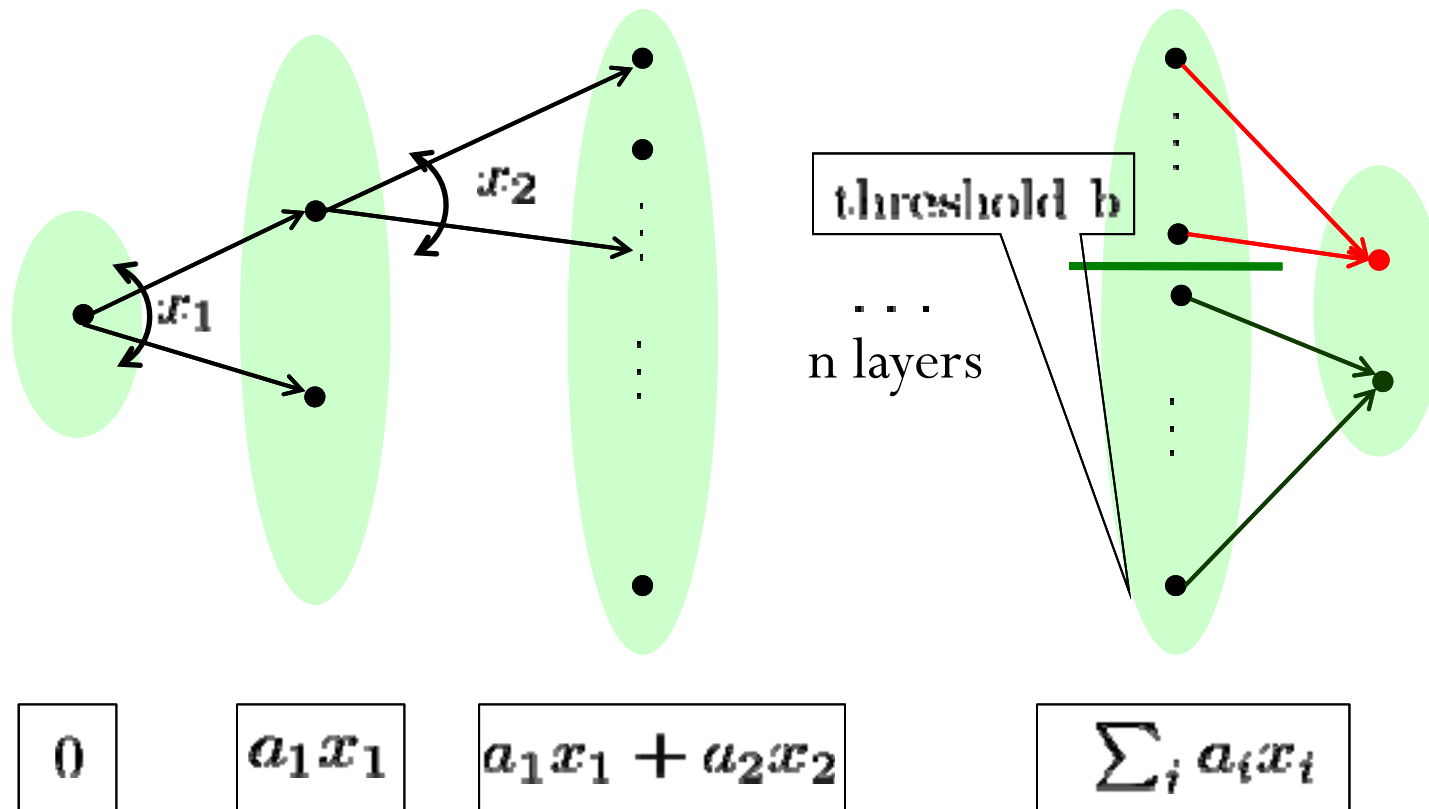
$(S, n)$ -ROBP



- Layered directed graph
- $S$  vertices per layer
- Edges between consecutive layers
- Edges labeled  $\{0, 1\}$
- **Input:**  $(x_1, \dots, x_n)$
- **Output:** Label of final vertex reached

**Counting the number of accepting paths ?**  
dynamic programming, time =  $O(nS)$

# ROBP for knapsack



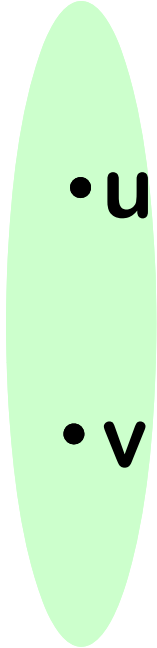
**Problem: width too large**

**Solution: reduce width by approximating**

# Monotone ROBPs

monotone:  $u \leq v \iff A(u) \subseteq A(v)$

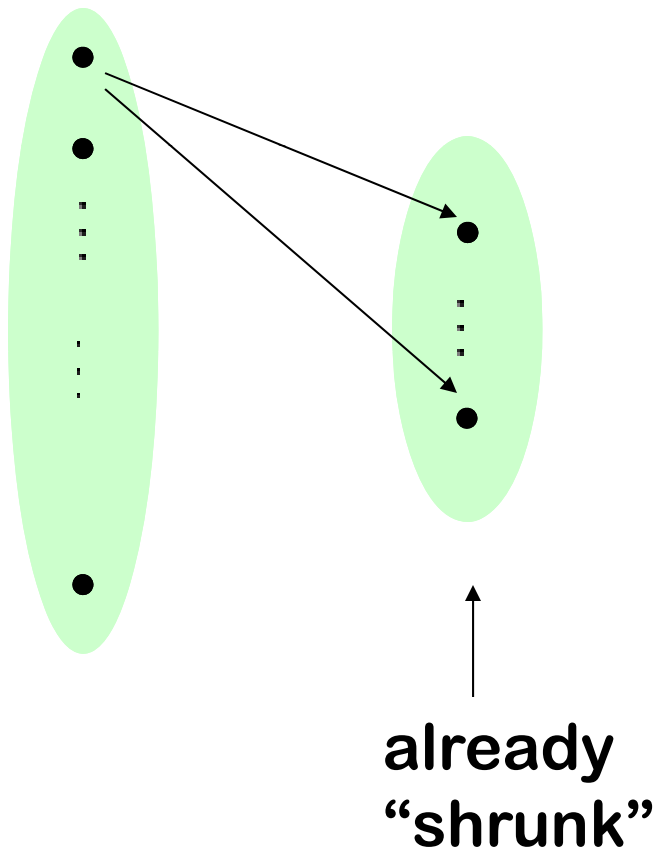
accepting paths from u



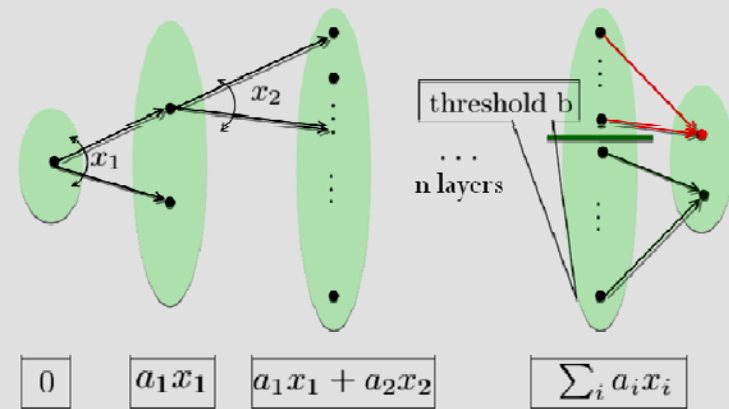
given implicitly

- ordering: given  $u, v$ , is  $u \leq v$  ?
- midpoint: given  $u, v$ , get  $w$  s.t.  
 $|\{x; u \leq x \leq w\}| = |\{x; w \leq x \leq v\}| \pm 1$
- transitions: given  $u$ , get the outneighbors of  $u$

group the vertices in the layers according to the rough number of accepting paths processing right-left



### ROBP for knapsack



**Problem:** width too large

**Solution:** reduce width by approximating

### Monotone ROBPs

monotone:  $u \leq v \iff A(u) \subseteq A(v)$  accepting paths from  $u$

given implicitly

- **ordering:** given  $u, v$ , is  $u \leq v$  ?

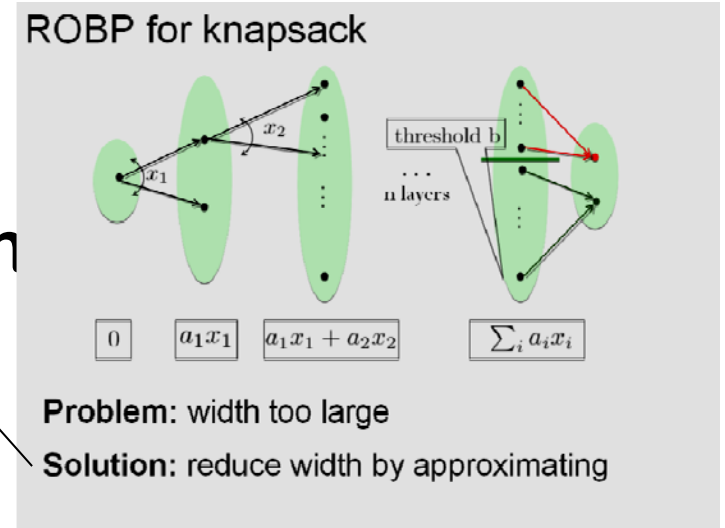
- **midpoint:** given  $u, v$ , get  $w$  s.t.  
 $|\{x; u \leq x \leq w\}| = |\{x; w \leq x \leq v\}| \pm 1$

- **transitions:** given  $u$ , get the outneighbors of  $u$



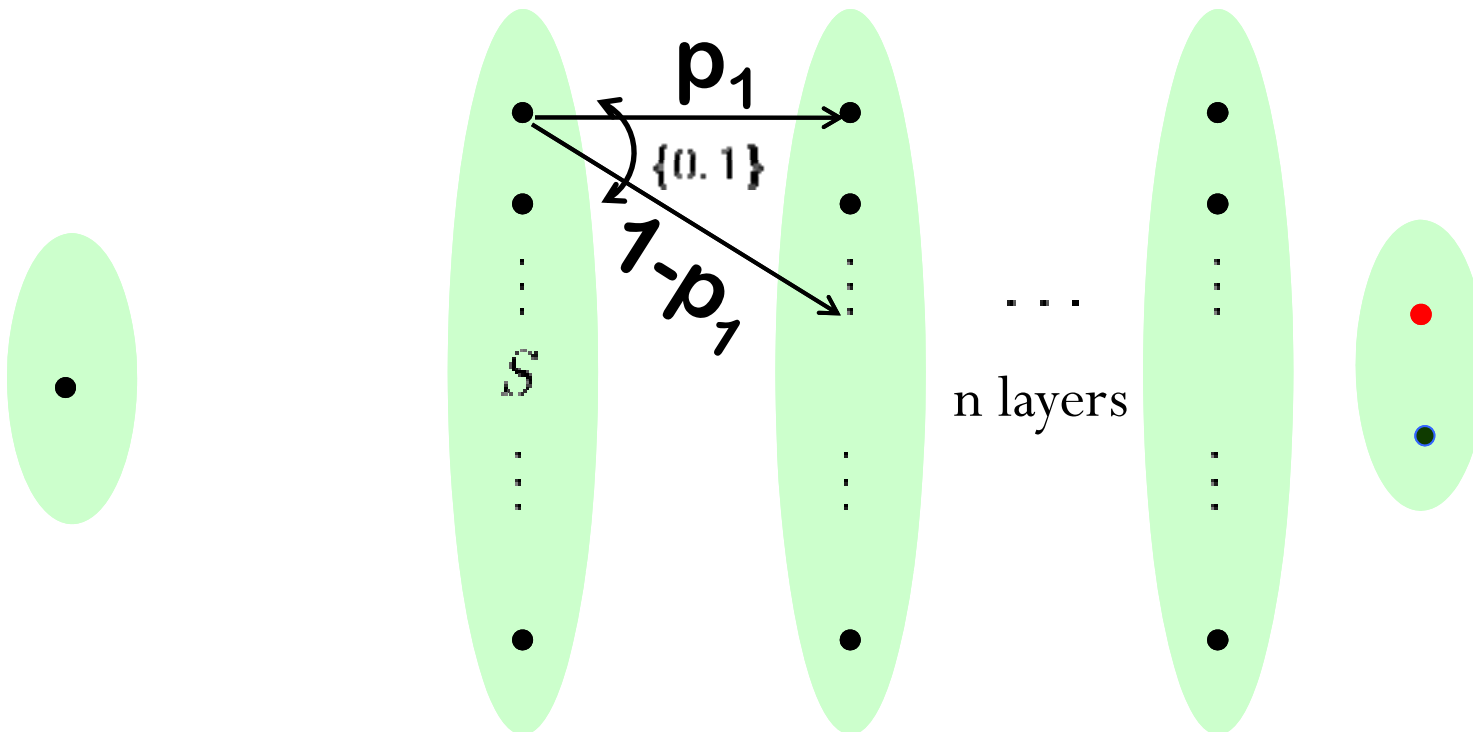
# More constraints?

can be generalized to distribution given by small space sources.

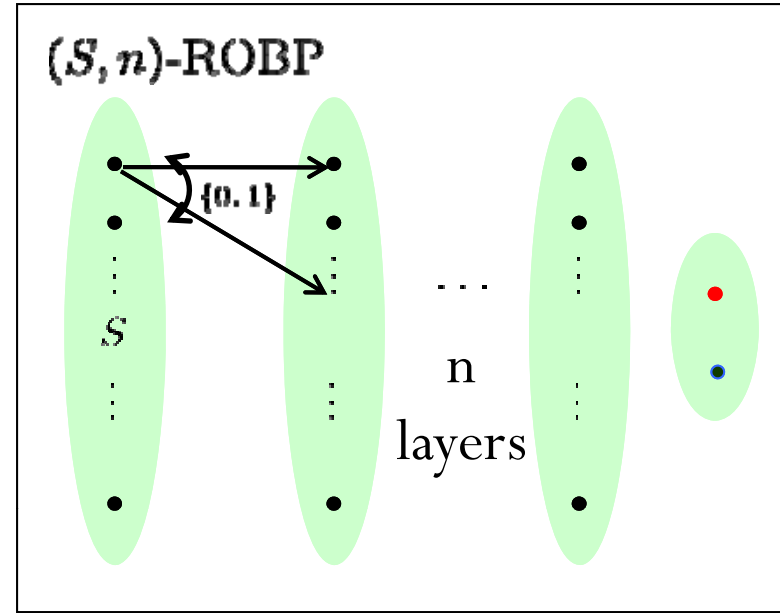
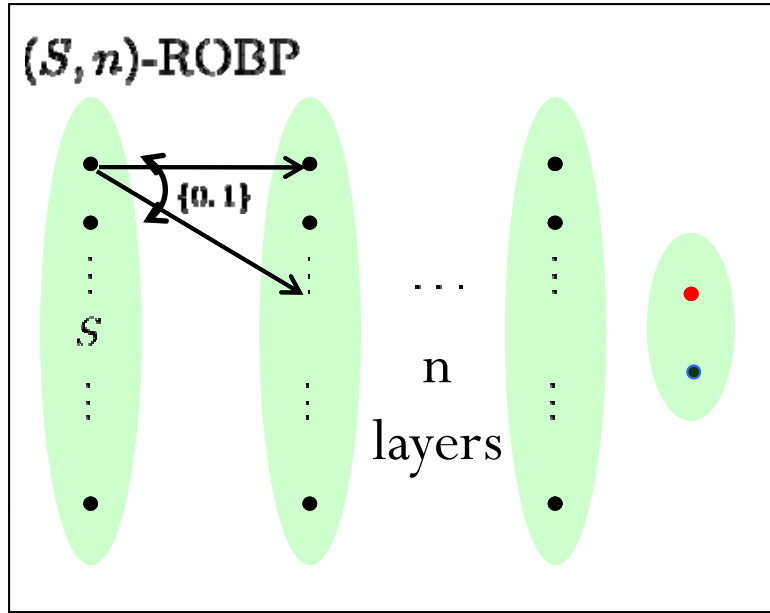


small space sources =

ROBP + probability distributions on outgoing edges



# More constraints?

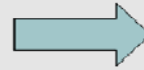


can be combined to get  $(S^2, n)$ -ROBP for intersection

additive approximation preserved



+ rejection sampling



approximate counter

$$w_i' = \frac{n^2 w_i}{L} \quad L' = n^2 L$$

rounding:

$$w_i'' = \lfloor w_i' \rfloor$$

- 1) get more solutions,  $\Omega'' \supseteq \Omega'$
- 2) not too many more,  $|\Omega''| \leq (n+1)|\Omega'|$

Proof:

$S'' \in \Omega'' - \Omega'$ ,  $X$  heaviest in  $S''$ , then  $S'' - \{X\} \in \Omega'$

Dyer'2003:

Counting knapsack

INPUT:

$w_1, \dots, w_n, L$

OUTPUT:

How many  $S \subseteq [n]$

with

$$\sum_{i \in S} w_i \leq L$$


are there?

#P-hard

$T[i, M]$

(number of solutions with items  $1, \dots, i$  and limit  $M$ )

$$T[i, M] = T[i-1, M] + T[i-1, M-w_i]$$

+ rejection sampling  approximate counter

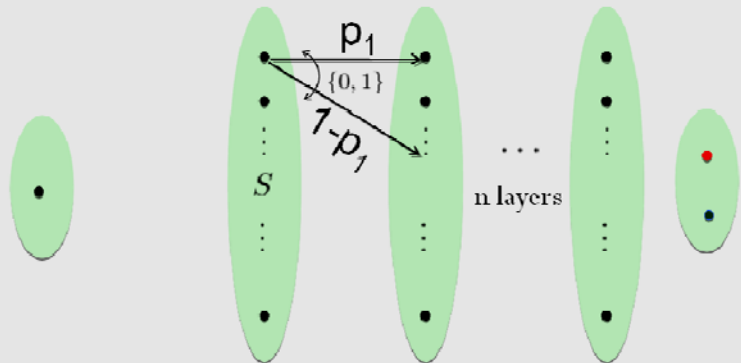
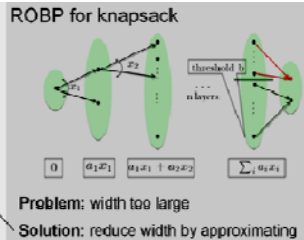
1) uniform distribution given by  $\Omega''$  can be given by small space source

2) additive approximation  $\Rightarrow$   
multiplicative approximation

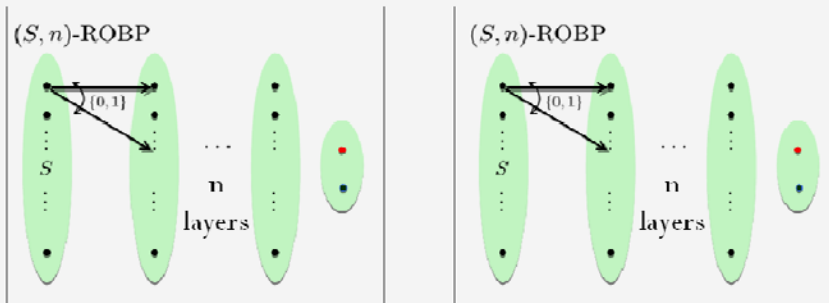
# More constraints?

can be generalized to distributions given by **small space sources**.

**small space sources =**  
ROBP + probability distributions on outgoing edges



# More constraints?



can be combined to get  $(S^2, n)$ -ROBP for intersection

**additive** approximation preserved

# How to deal with more constraints?

(e.g., contingency tables, multi-dimensional knapsack, ...)

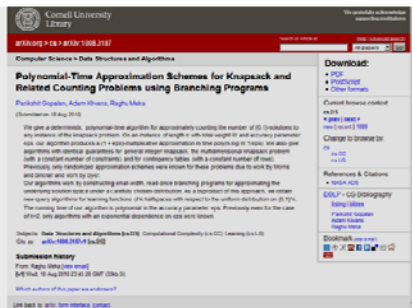
## multi-dimensional knapsack:

How many  $S \subseteq [n]$

$$\text{with } \sum_{i \in S} w_{j,i} \leq L_j \quad j \in \{1, \dots, k\}$$

are there?

$$O((n/\epsilon)^{O(k^2)} \log W) \text{ algorithm}$$



+ rejection sampling  $\rightarrow$  approximate counter

$$w_i' = \frac{n^2 w_i}{L} \quad L' = n^2$$

rounding:  
 $w_i'' = \lfloor w_i' \rfloor$

- 1) get more solutions,  $\Omega'' \supseteq \Omega'$
- 2) not too many more,  $|\Omega''| \leq (n+1)|\Omega'|$

Proof:  
 $S'' \in \Omega'' - \Omega'$ ,  $X$  heaviest in  $S''$ , then  $S'' - \{X\} \in \Omega'$

Dyer'2003: Counting knapsack  
INPUT:  $w_1, \dots, w_n, L$   
OUTPUT: number of solutions with items  $1, \dots, n$  and limit  $M$   
 $T[i, M] = T[i-1, M] + T[i-1, M-w_i]$

+ rejection sampling  $\rightarrow$  approximate counter

- 1) uniform distribution given by  $\Omega''$  can be given by small space source
- 2) additive approximation  $\Rightarrow$  multiplicative approximation

# Other problems:

contingency tables  
with constant  
number of rows

More constraints?  
can be generalized to distributions given by **small space sources**.

**small space sources** = ROBP + probability distributions on outgoing edges

More constraints?  
 $(S, n)$ -ROBP

can be combined to get  $(S^2, n)$ -ROBP for intersection

**additive** approximation preserved

How to deal with more constraints?  
(e.g., contingency tables, multi-dimensional knapsack, ...)

**multi-dimensional knapsack:**  
How many  $S \subseteq [n]$  with  $\sum_{i \in S} w_{i,j} \leq L_j \quad j \in \{1, \dots, k\}$  are there?

$O((n/\epsilon)^{O(k^2)} \log W)$  algorithm

+ rejection sampling → **approximate counter**

Dyer 2003: **Contingent knapsack**

$w_i = \frac{n^2 w_i}{L} \quad L = n^2$

rounding:  $w_i'' = \lfloor w_i' \rfloor$

1) get more solutions,  $|\Omega''| \geq |\Omega'|$   
2) not too many more,  $|\Omega''| \leq (n+1)|\Omega'|$

Proof:  
 $S'' \in \Omega'' - \Omega'$ , X heaviest in  $S'$ , then  $S'' - \{X\} \in \Omega'$

1) uniform distribution given by  $\Omega''$  can be given by small space source

2) additive approximation  $\Rightarrow$  multiplicative approximation

What other problems are solvable  
using the technique?

Thanks!