On the Derandomization of Space-Bounded Computations

Jiří Šíma, Stanislav Žák

Institute of Computer Science
Academy of Sciences of the Czech Republic

Probabilistic (Monte Carlo) Algorithms

- the next step can be chosen randomly from a set of possibilities (behavior can vary even on a fixed input)
- an output may be incorrect with a certain (typically small) probability

Undirected Graph $S - T$ Connectivity: Given an undirected graph $G = (V, E)$ on $n = |V|$ vertices and $s, t \in V$, is there a path from s to t in G?

linear-time deterministic algorithm using breadth-first or depth-first search requires linear space (queue or stack implementation, respectively)

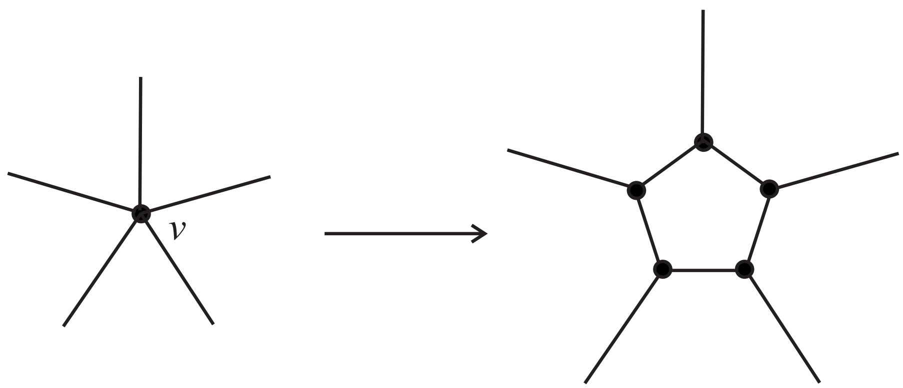
Random-Walk Algorithm

1. Let $v = s$.
2. Repeat up to $100n^4$ times:
   (a) If $v = t$, halt and accept.
   (b) Else v choose randomly from $\{w \in V \mid \{v, w\} \in E\}$.
3. Reject (if we haven't visited t yet).

- never accepts when there isn't path from s to t
- only requires space $O(\log n)$: current vertex v, a counter for the number of steps

Theorem 1 *For every d-regular undirected graph $G = (V, E)$ on n vertices and for any vertices $s, t \in V$ from the same connected component of G, the expected number of steps for a random walk started at s to visit t is $O(d^2 n^3 \log n)$.*

- 3-regular graph preserving $s - t$ connectivity:

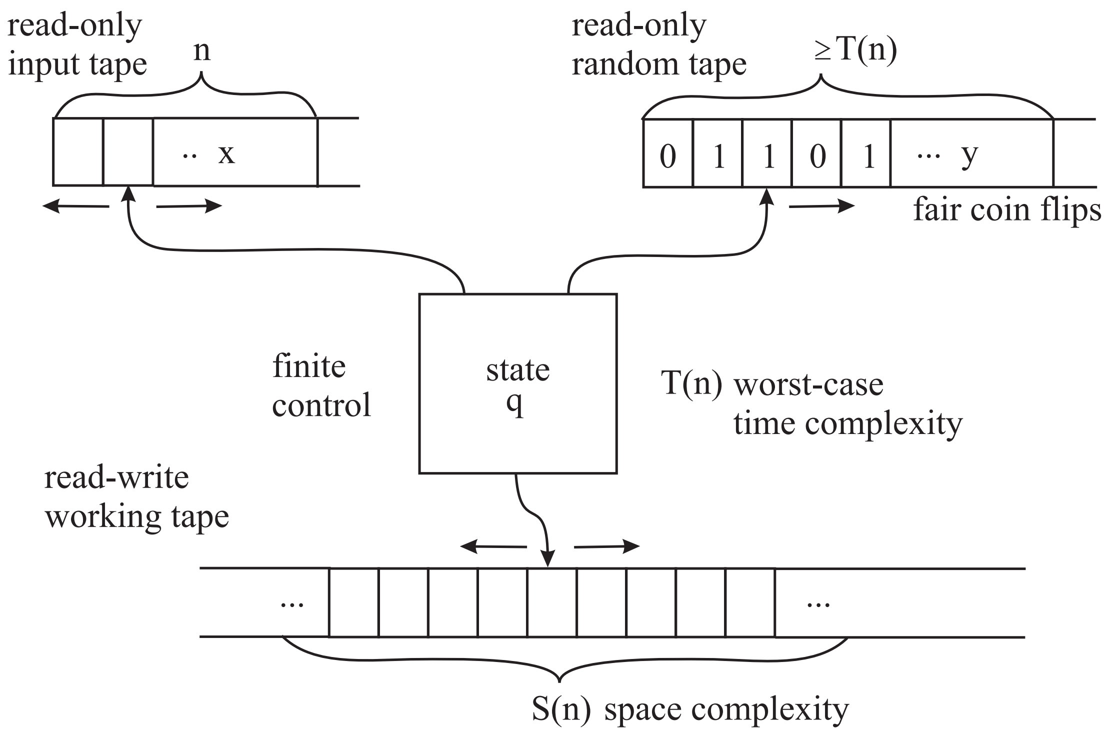


- a random walk from s visits t in $100n^4$ steps with probability at least $3/4$

undirected graph $s - t$ connectivity is solvable deterministically in space $O(\log n)$
(O. Reingold, 2005 --- Grace Murray Hopper ACM Award 2005, Gödel Prize 2009)

Can any $O(\log n)$-space probabilistic algorithm be derandomized while preserving its space complexity?

Probabilistic Turing Machine (PTM)



RL (Randomized Logarithmic-space)

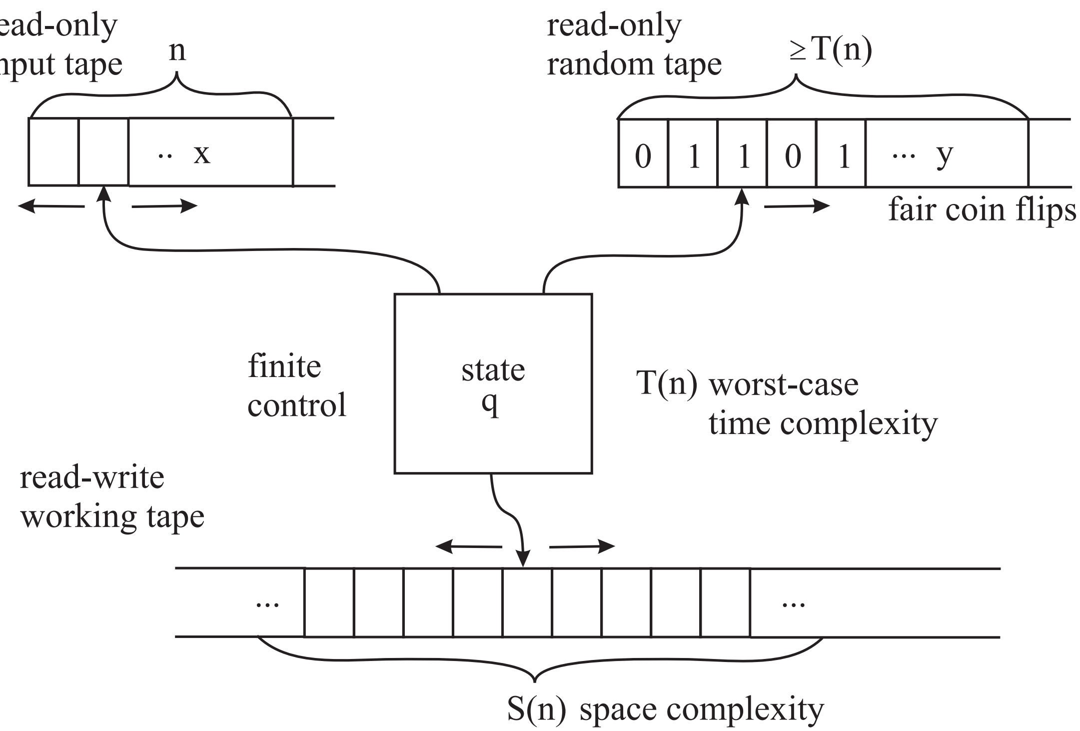
class of problems $L = L(M)$ solvable by PTMs M with one-sided error $0 < \delta < 1$ in logarithmic space $S(n) = O(\log n)$ and polynomial time $T(n) = O(n^c)$:

if $x \in L$, then $Pr_{y \sim U_{T(n)}}[M(x,y) = 1] \geq 1 - \delta$

if $x \notin L$, then $Pr_{y \sim U_{T(n)}}[M(x,y) = 1] = 0$, i.e. $M(x,y) = 0$ for every y

$\longrightarrow$ if $M(x,y) = 1$, then $x \in L$ (U_m is the uniform distribution on $\{0,1\}^m$)

Probabilistic Turing Machine (PTM)



BPL (Bounded-error Probabilistic Logarithmic-space)
class of problems $L = L(M)$ solvable by PTMs M with two-sided error $0 \leq \delta < \frac{1}{2}$
in logarithmic space $S(n) = O(\log n)$ and polynomial time $T(n) = O(n^c)$:

if $x \in L$, then $Pr_{y \sim U_{T(n)}}[M(x,y) = 1] \geq 1 - \delta$

if $x \notin L$, then $Pr_{y \sim U_{T(n)}}[M(x,y) = 1] \leq \delta$

(U_m is the uniform distribution on $\{0,1\}^m$)

Derandomization of Space-Bounded Computation

deterministic simulation of PTM performs $M(x,y)$ for every fixed setting of random input $y \in \{0,1\}^m$ (where $m = T(n)$) and computes the probability of accepting computations

$$Pr_{y \sim U_m}[M(x,y)=1] = \frac{\sum_{y \in \{0,1\}^m} M(x,y)}{2^m} = \begin{cases} \geq 1-\delta & \longrightarrow \text{ accepts } x \\ \leq \delta & \longrightarrow \text{ rejects } x \end{cases}$$

$\longrightarrow$ the simulation time is exponential in $T(n)$

Is there an efficient simulation of PTM? Does randomness add power?

$$\mathbf{BPL} \stackrel{?}{=} \mathbf{L}, \quad \mathbf{RL} \stackrel{?}{=} \mathbf{L}$$

Pseudorandom Generator (PRG)

$$g : \{0,1\}^s \longrightarrow \{0,1\}^m , \quad s \ll m$$

stretches a short uniformly random seed of s bits into m bits that cannot be distinguished from uniform ones by small space machines M:

$$\left| Pr_{y \sim U_m} [M(y) = 1] - Pr_{z \sim U_s} [M(g(z)) = 1] \right| \leq \varepsilon$$

where $\varepsilon > 0$ is the error

deterministic simulation of PTM performs $M(x, g(z))$ for every fixed setting of seed $z \in \{0,1\}^s$ and approximates the probability of accepting computations

$$Pr_{y \sim U_m} [M(x,y) = 1] \ \doteq \ \frac{\sum_{z \in \{0,1\}^s} M(x, g(z))}{2^s}$$

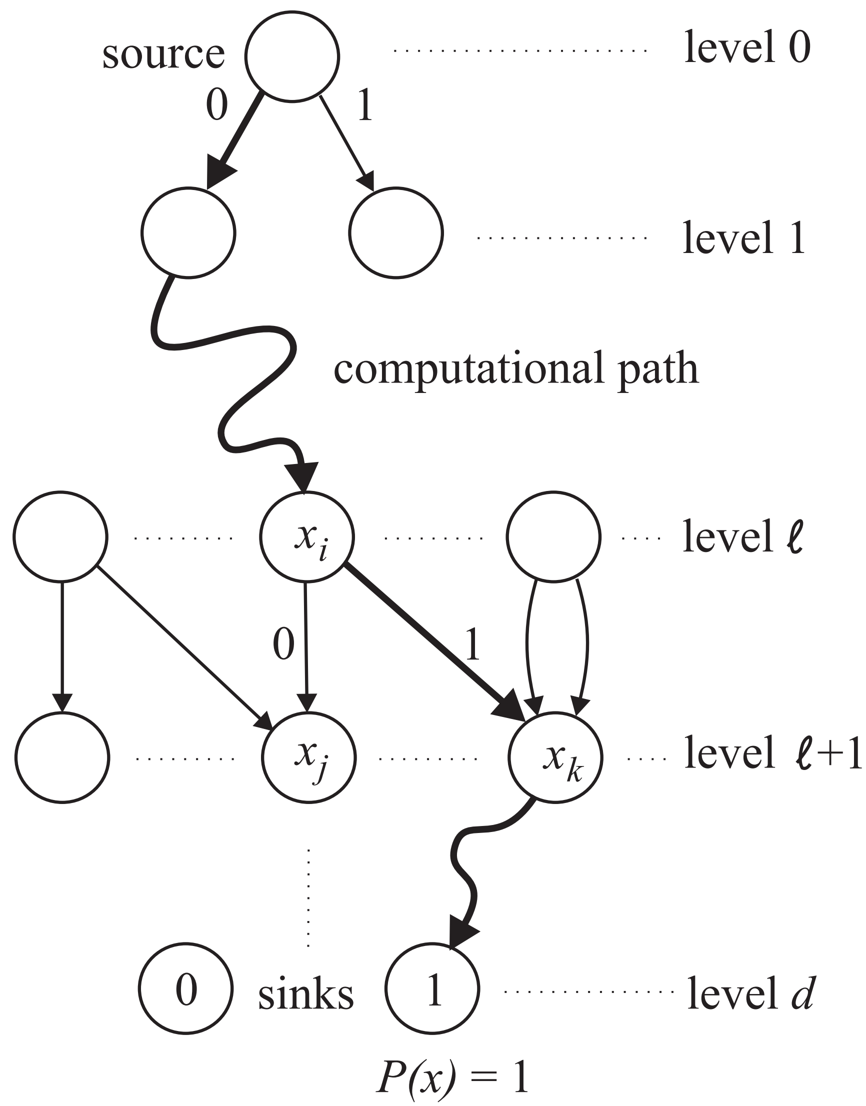
efficient derandomization (BPL=L): an explicit PRG with seed length $s = O(\log n)$ and sufficiently small error ε, computable in logarithmic space that fools logarithmic space machines M

Branching Program P

a leveled directed acyclic multi-graph $G = (V, E)$:

- one source $s \in V$ of zero in-degree at level 0
- sinks of zero out-degree at the last level d (=depth)
- every inner (=non-sink) node has out-degree 2
- the inner nodes are labeled with input Boolean variables $x_1, \ldots, x_n$
- the two edges outgoing from any inner node at level $\ell < d$ lead to nodes at the next level $\ell + 1$ and are labeled 0 and 1
- the sinks are labeled 0 and 1

width $=$ the maximum number of nodes in one level branching program P computes Boolean function $P : \{0,1\}^n \longrightarrow \{0,1\}$:



Branching Programs (BPs)

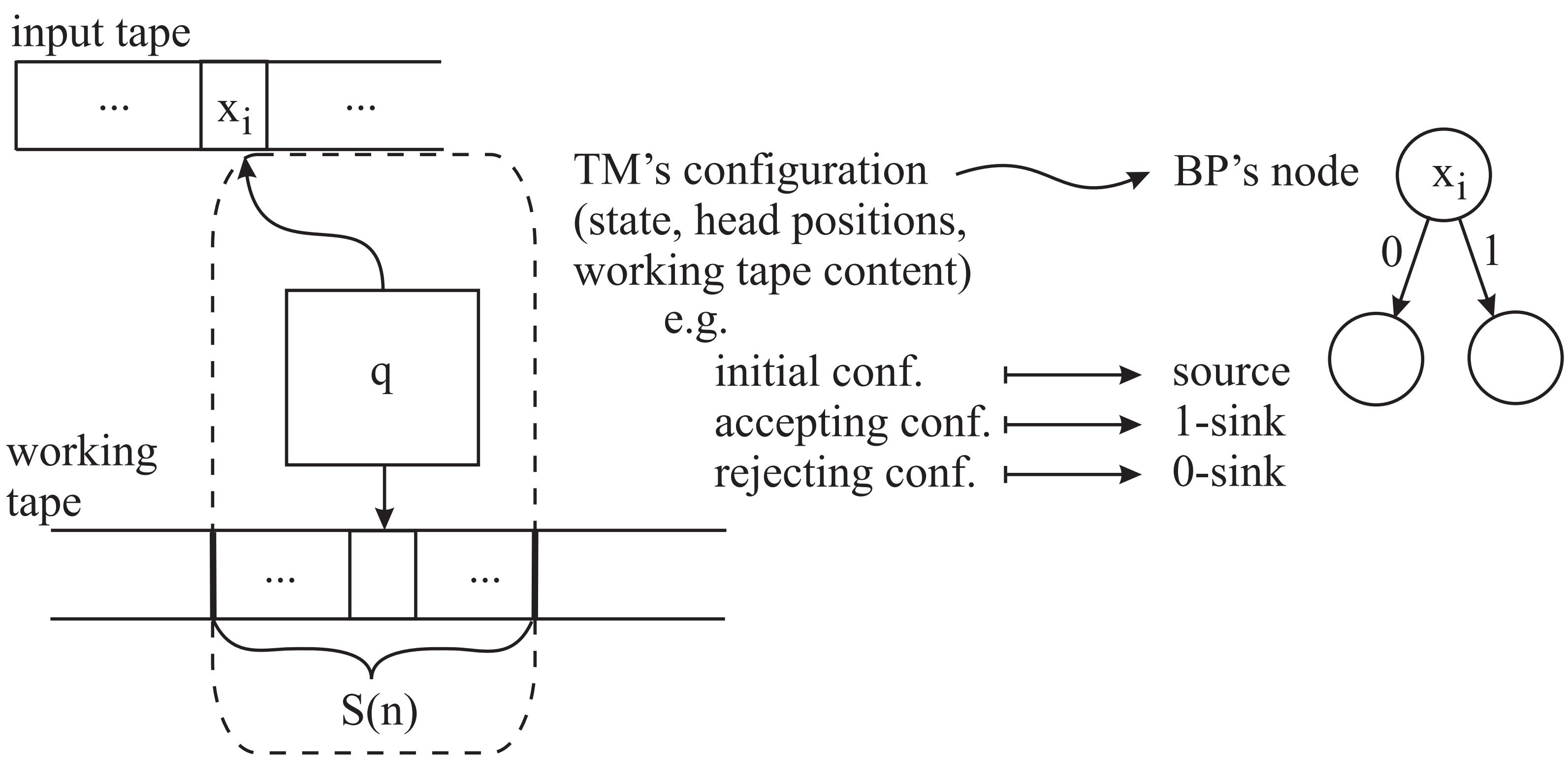
a non-uniform model of space-bounded computation:

infinite family of branching programs $\{P_n\}$, one P_n for each input length $n \geq 1$

Turing machine M that uses space $s(n)$ and runs in time $t(n)$

is modeled by

branching program P_n of width $2^{s(n)}$ and depth $t(n)$



Restrictions

Read-Once BPs (1-BPs): every input variable is tested at most once along each computational path

Oblivious BPs: at each level only one variable is queried
$\longrightarrow$ provably less efficient model (Beame, Machmouchi, CCC 2011)

an efficient construction of PRG for 1-BPs of polynomial size suffices to derandomize BPL

Explicit Pseudorandom Generators for 1-BPs

polynomial width: PRG with seed length $O(\log^2 n)$ (Nisan, 1992)

width $w = 2$: PRG with seed length $O(\frac{1}{\varepsilon} \log n)$ (Saks, Zuckerman, 1999)

width $w = 3$: known techniques fail to improve the seed length $O(\log^2 n)$ from Nisan's result (RANDOM 2009, STOC 2010, 2011, FOCS 2010, CCC 2011)

More Restrictions

regular 1-BP: every non-source node has in-degree 2

permutation 1-BP: regular 1-BP where the two edges leading to any non-source node are labeled 0 and 1 (i.e. edges between levels labeled with 0 respectively 1 create a permutation)

Recent Results on PRGs for regular 1-BPs

oblivious permutation 1-BPs of constant width: PRG with seed length $O\left(\log\frac{1}{\varepsilon}\log n\right)$
(Koucký, Nimbhorkar, Pudlák, STOC 2011)

oblivious regular 1-BPs of constant width:

- two constructions of PRG with seed length $O\left(\log n\left(\log\log n + \log\frac{1}{\varepsilon}\right)\right)$
  (Braverman, Rao, Raz, Yehudoff, FOCS 2010; Brody, Verbin, FOCS 2010)
- PRG with seed length $O\left(\log\frac{1}{\varepsilon}\log n\right)$ (De, CCC 2011)

$\times$ regular 1-BPs of constant width cannot even evaluate read-once conjunctions of non-constant number of literals (e.g. read-once DNF or CNF)

Hitting Set Generator

the one-sided error version of pseudo-random generator

Hitting Set:

Let $\varepsilon > 0$ and $\mathcal{P}_n$ be a class of BPs with n inputs. A set $H_n \subseteq \{0,1\}^n$ is an ε-hitting set for $\mathcal{P}_n$ if for every $P \in \mathcal{P}_n$,

$$Pr_{x \sim U_n}\left[P(x) = 1\right] = \frac{\left|P^{-1}(1)\right|}{2^n} \geq \varepsilon \quad \text{implies} \quad (\exists\, a \in H_n)\;\; P(a) = 1\,.$$

For every n (given in unary), the hitting set generator (HSG) for a class of families of BPs produces hitting set H_n.

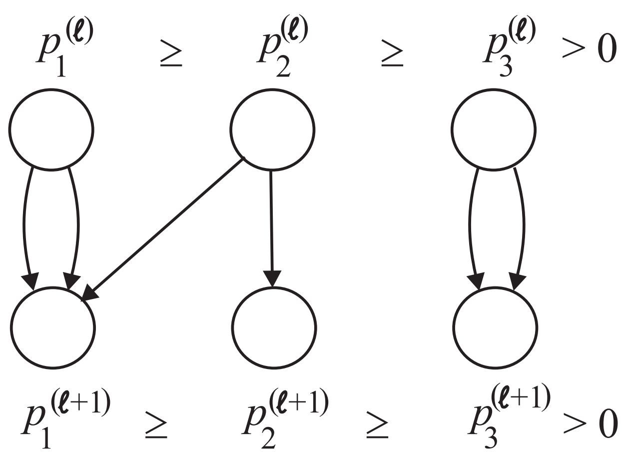
deterministic simulation of a randomized algorithm with one-sided error performs the computation for every fixed setting of random input from the hitting set and accepts if there is at least one accepting computation

Hitting Set Generator for 1-BPs of Width 3

a normalized form of BP: the probability distribution of inputs on the three nodes at each level is ordered as

$$p_1 \geq p_2 \geq p_3 > 0 \qquad (p_1 + p_2 + p_3 = 1)$$

a simple 1-BP of width 3 excludes one special level-to-level transition pattern in its normalized form (about 40 possible patterns in normalized width-3 1-BPs):



$\longrightarrow$ any regular width-3 1-BP is simple

a polynomial-time construction of $\left(\frac{191}{192}\right)$-hitting set for simple 1-BPs of width 3 which need not be oblivious (Šíma, Žák, SOFSEM 2007)

The Weak Richness Condition

A set $A \subseteq \{0,1\}^n$ is weakly ε-rich if for any index set $I \subseteq \{1,\ldots,n\}$ and for any partition $\{Q_1,\ldots,Q_q,R_1,\ldots,R_r\}$ of I ($q \geq 0$, $r \geq 0$) satisfying

$$\left(1 - \prod_{j=1}^{q}\left(1 - \frac{1}{2^{|Q_j|}}\right)\right) \times \prod_{j=1}^{r}\left(1 - \frac{1}{2^{|R_j|}}\right) \geq \varepsilon\,, \tag{1}$$

for any $c \in \{0,1\}^n$ there exists $a \in A$ that meets

$$(\exists\, j \in \{1,\ldots,q\})\,(\forall\, i \in Q_j)\, a_i = c_i \quad \text{and}$$
$$(\forall\, j \in \{1,\ldots,r\})\,(\exists\, i \in R_j)\, a_i \neq c_i\,. \tag{2}$$

Equivalent to ε-Hitting Sets for Read-Once DNF & CNF:

The product on the left-hand side of inequality in (1) expresses the probability that a random $a \in \{0,1\}^n$ (not necessarily in A) satisfies condition (2) which can be interpreted as a read-once conjunction of DNF and CNF

$$\bigvee_{j=1}^{q} \bigwedge_{i \in Q_j} \ell(x_i) \wedge \bigwedge_{j=1}^{r} \bigvee_{i \in R_j} \neg\ell(x_i) \quad \text{where} \quad \ell(x_i) = \begin{cases} x_i & \text{for } c_i = 1 \\ \neg x_i & \text{for } c_i = 0\,. \end{cases}$$

The Weak Richness Condition Is Necessary

Theorem 2 *Any ε-hitting set for the class of 1-BPs of width 3 is weakly ε-rich.*

Idea of Proof:

- 1-BPs of width 3 can implement read-once conjunctions of DNF and CNF

- a hitting set for a class of functions hits any of its subclass

The Weak Richness Condition

A set $A \subseteq \{0,1\}^n$ is weakly ε-rich if for any index set $I \subseteq \{1,\ldots,n\}$ and for any partition $\{Q_1,\ldots,Q_q,R_1,\ldots,R_r\}$ of I ($q \geq 0$, $r \geq 0$) satisfying

$$\left(1 - \prod_{j=1}^{q}\left(1 - \frac{1}{2^{|Q_j|}}\right)\right) \times \prod_{j=1}^{r}\left(1 - \frac{1}{2^{|R_j|}}\right) \geq \varepsilon\,, \tag{1}$$

for any $c \in \{0,1\}^n$ there exists $a \in A$ that meets

$$(\exists\, j \in \{1,\ldots,q\})\,(\forall\, i \in Q_j)\, a_i = c_i \quad \text{and}$$
$$(\forall\, j \in \{1,\ldots,r\})\,(\exists\, i \in R_j)\, a_i \neq c_i\,. \tag{2}$$

Observation:

Condition (1) implies that there is $j \in \{1,\ldots,q\}$ such that $|Q_j| \leq \log n$.

$\longrightarrow$ The (Full) Richness Condition:

Replace $Q_1,\ldots,Q_q$ by Q such that $|Q| \leq \log n$
and remove the blue text from the definition above.

The Richness Condition

A set $A \subseteq \{0,1\}^n$ is ε-rich if for any index set $I \subseteq \{1,\ldots,n\}$, for any subset $Q \subseteq I$ and partition $\{R_1,\ldots,R_r\}$ of $I \backslash Q$ $(r \geq 0)$ satisfying $|Q| \leq \log n$ and

$$\prod_{j=1}^{r}\left(1-\frac{1}{2^{|R_j|}}\right) \geq \varepsilon\,, \tag{3}$$

for any $c \in \{0,1\}^n$ there exists $a \in A$ that meets

$$(\forall\, i \in Q)\, a_i = c_i \quad \text{and} \quad (\forall\, j \in \{1,\ldots,r\})\,(\exists\, i \in R_j)\, a_i \neq c_i\,. \tag{4}$$

Comments:

- Any ε-rich set is weakly ε-rich.
- Condition (4) can be interpreted as a read-once CNF with $O(\log n)$ single literals and clauses whose sizes satisfy (3):

$$\bigwedge_{i\in Q}\ell(x_i) \wedge \bigwedge_{j=1}^{r}\bigvee_{i\in R_j}\neg\ell(x_i) \quad \text{where} \quad \ell(x_i)=\begin{cases} x_i & \text{for } c_i=1\\ \neg x_i & \text{for } c_i=0\,.\end{cases}$$

The Richness Condition Is 'Sufficient'

Theorem 3 (Šíma, Žák, SOFSEM 2012) *Let $\varepsilon > \frac{5}{6}$. If A is ε'^{11}-rich for some $\varepsilon' < \varepsilon$, then $H = \Omega_3(A)$ which contains all the vectors within the Hamming distance of 3 from any $a \in A$, is an ε-hitting set for the class of 1-BPs of width 3.*

The richness condition expresses an essential property of hitting sets for 1-BPs of width 3 while being independent of a rather technical formalism of BPs.

$\times$

Does a rich set exist?

Can it be constructed efficiently?

Almost $O(\log n)$-Wise Independent Sets Are ε-Rich

$A \subseteq \{0,1\}^n$ is (k,β)-wise independent set if for any index set $S \subseteq \{1,\ldots,n\}$ of size $|S| \le k$, the probability distribution on the bit locations from S is almost uniform, i.e. for any $c \in \{0,1\}^n$

$$\left| \frac{|A^S(c)|}{|A|} - \frac{1}{2^{|S|}} \right| \le \beta$$

where $A^S(c) = \{a \in A \,|\, (\forall i \in S)\, a_i = c_i\}$.

Alon, Goldreich, Håstad, Peralta, 1992: for any $\beta > 0$ and $k = O(\log n)$, (k,β)-wise independent set A can be constructed in time polynomial in $\frac{n}{\beta}$

Theorem 4 (Šíma, Žák, CSR 2011) *Let $\varepsilon > 0$, C be the least odd integer greater than $(\frac{2}{\varepsilon}\ln\frac{1}{\varepsilon})^2$, and $0 < \beta < \frac{1}{n^{C+3}}$. Then any $(\lceil (C+2)\log n \rceil, \beta)$-wise independent set is ε-rich.*

Corollary 1 *Almost $O(\log n)$-wise independent sets are hitting sets for the read-once conjunctions of DNF and CNF.*

previously known for read-once DNFs resp. read-once CNFs (De et al., RANDOM 2010)

The Hitting Set for 1-BPs of width 3

Corollary: Any almost $O(\log n)$-wise independent set extended with all the vectors within the Hamming distance of 3 is a polynomial-time constructible ε-hitting set for 1-BPs of width 3 with acceptance probability $\varepsilon > 5/6$.

Conclusion & Open Problems

- a breakthrough in the effort to construct HSGs for 1-BPs of bounded width (De, CCC 2011)

  $\times$

  Such constructions were known only for width 2 or for oblivious regular/permutation 1-BPs of bounded width.

- Can the result be achieved for any acceptance probability $\varepsilon > 0$?
  ($\times$ our result holds for $\varepsilon > 5/6$)

- Can the technique be extended to width 4 or to bounded width ?