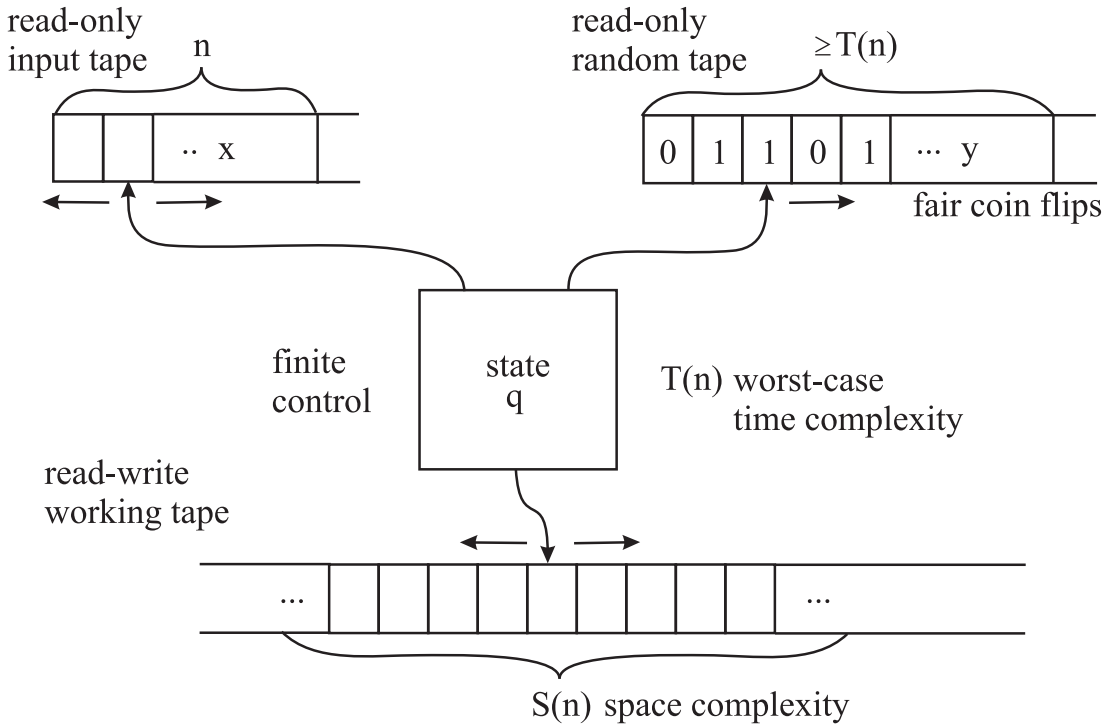
A Sufficient Condition for Sets Hitting the Class of Read-Once Branching Programs of Width 3

Jiří Šíma, Stanislav Žák

Institute of Computer Science
Academy of Sciences of the Czech Republic

Probabilistic Turing Machine (PTM)



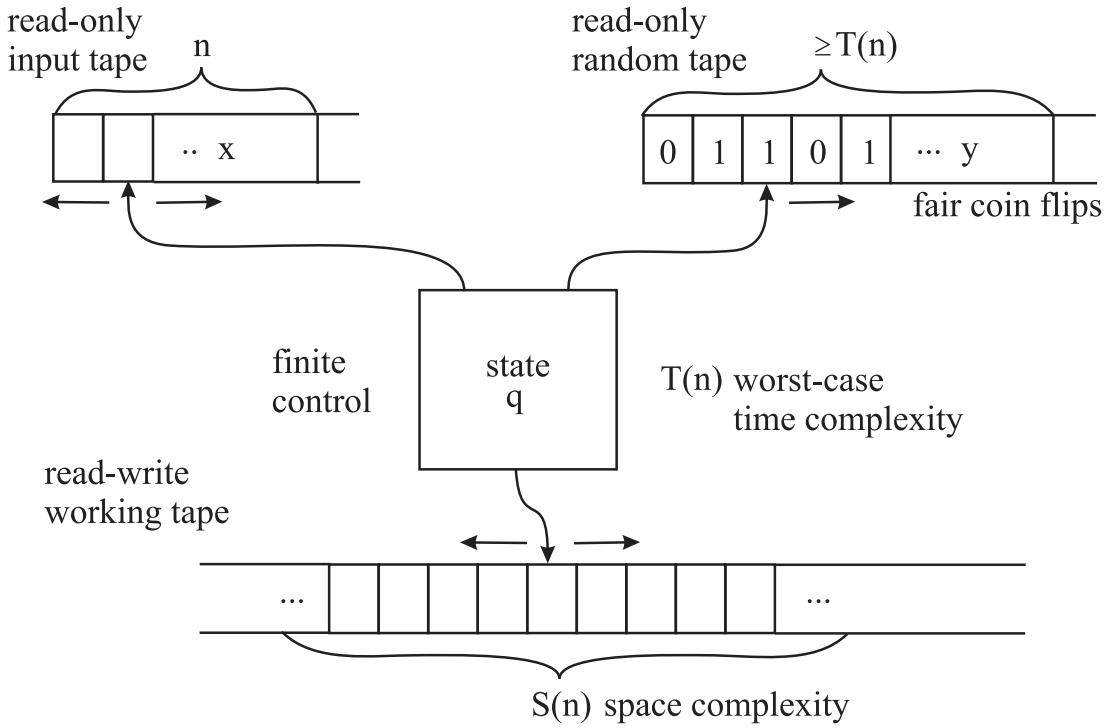
BPL (Bounded-error Probabilistic Logarithmic-space)
class of problems $L = L(M)$ solvable by PTMs M with two-sided error $0 \leq \delta < \frac{1}{2}$
in logarithmic space $S(n) = O(\log n)$ and polynomial time $T(n) = O(n^c)$:

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 $\{0, 1\}^m$)

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 $x \in L$

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

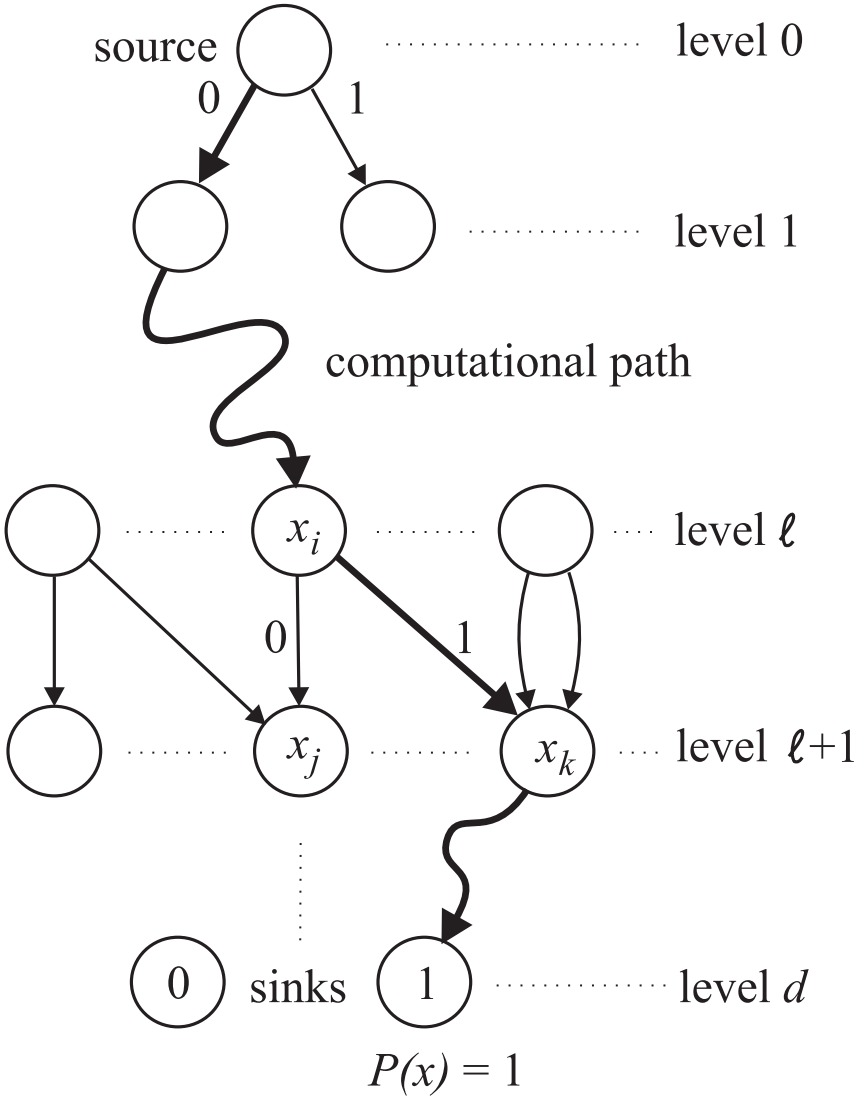
Branching Program P

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)

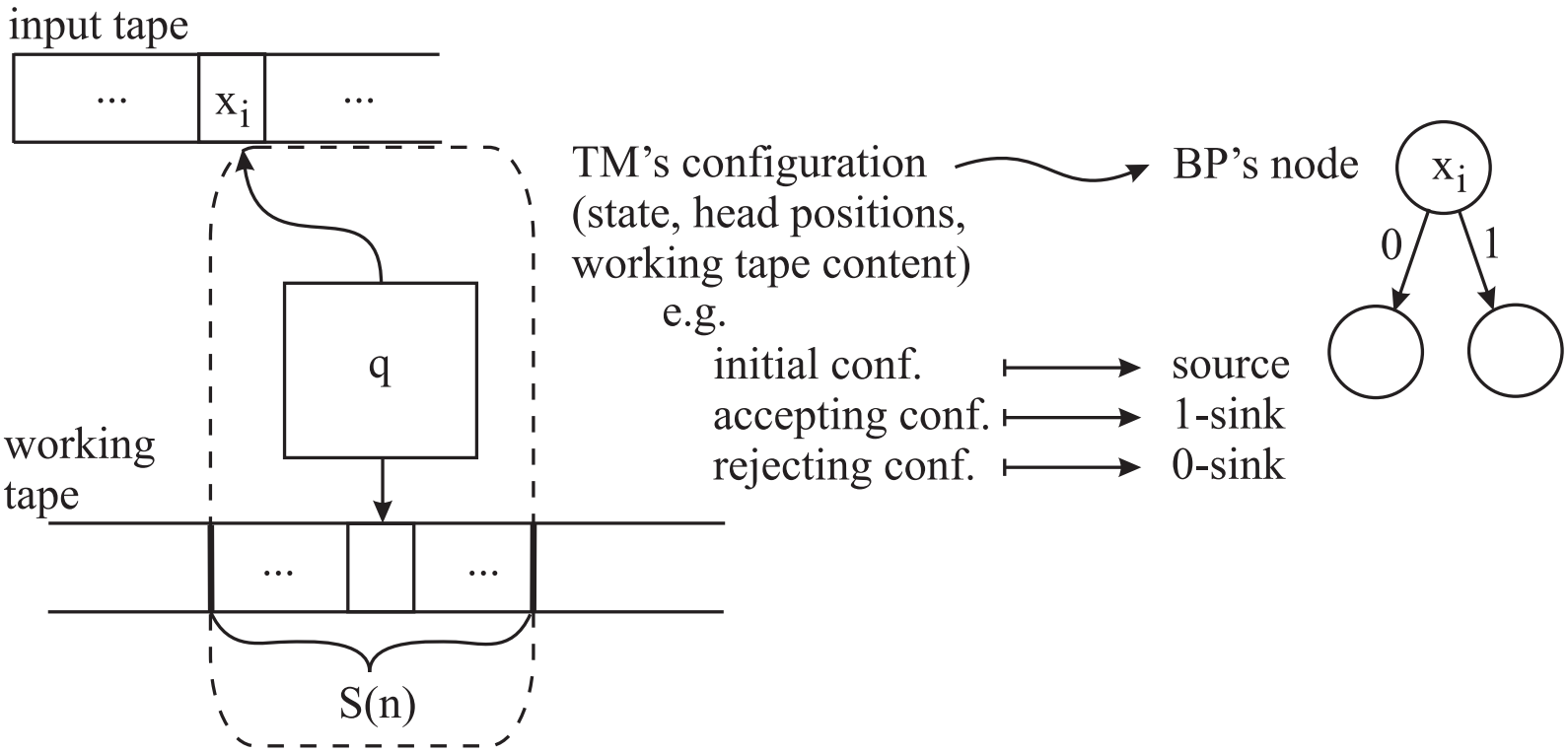
a non-uniform model of space-bounded computation:

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 inner non-source node has in-degree 2

permutation 1-BP: regular 1-BP where the two edges leading to any inner 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. DNF, 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 $P \in \mathcal{P}_n$,

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

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

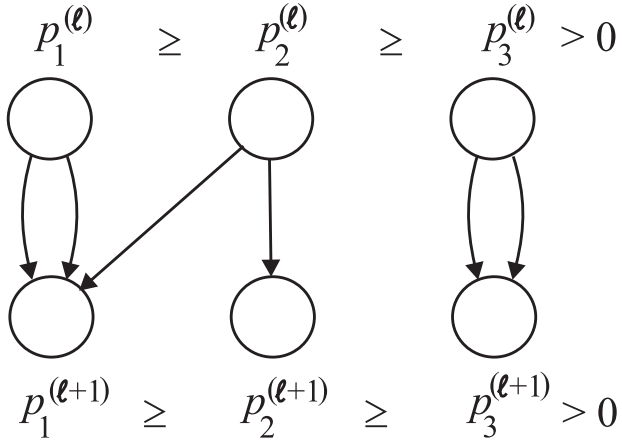
deterministic simulation of a randomized algorithm with one-sided error performs the computation for every fixed setting of random input from the hitting set and accepts if there is at least one accepting computation

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 1 *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 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):

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

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 $|S| \leq k$, the probability distribution on the bit locations from S is almost uniform, i.e. for any $c \in \{0,1\}^n$

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

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

`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 2 `(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 conjunctions of DNF and CNF.*

previously known for DNFs resp. CNFs `(De,Etesami,Trevisan,Tulsiani,RANDOM 2010)`

Intuition behind the Proof

Let A be a $(\lceil (C+2)\log n \rceil, \beta)$-wise independent set. We will show A is ε-rich:

Assume subset $Q \subseteq I \subseteq \{1,\ldots,n\}$ and partition $\{R_1,\ldots,R_r\}$ of $I \setminus Q$ satisfy $|Q| \leq \log n$ and $\prod_{j=1}^{r}(1-1/2^{|R_j|}) \geq \varepsilon$.

For a given $c \in \{0,1\}^n$ we want to show that there is $a \in A$ that meets

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

We will prove that the probability

$$p(A) = \frac{\left|A^Q(c) \setminus \bigcup_{j=1}^{r} A^{R_j}(c)\right|}{|A|} > 0\,.$$

By using the assumption that A is almost $O(\log n)$-wise independent one can approximate

$$p(A) \sim p\left(\{0,1\}^n\right) = \frac{1}{2^{|Q|}} \prod_{j=1}^{r}\left(1 - \frac{1}{2^{|R_j|}}\right) \geq \frac{\varepsilon}{n} > 0\,.$$

The Main Result: The Richness Condition Is 'Sufficient'

Theorem 3 *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.*

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 $\varepsilon > 5/6$.

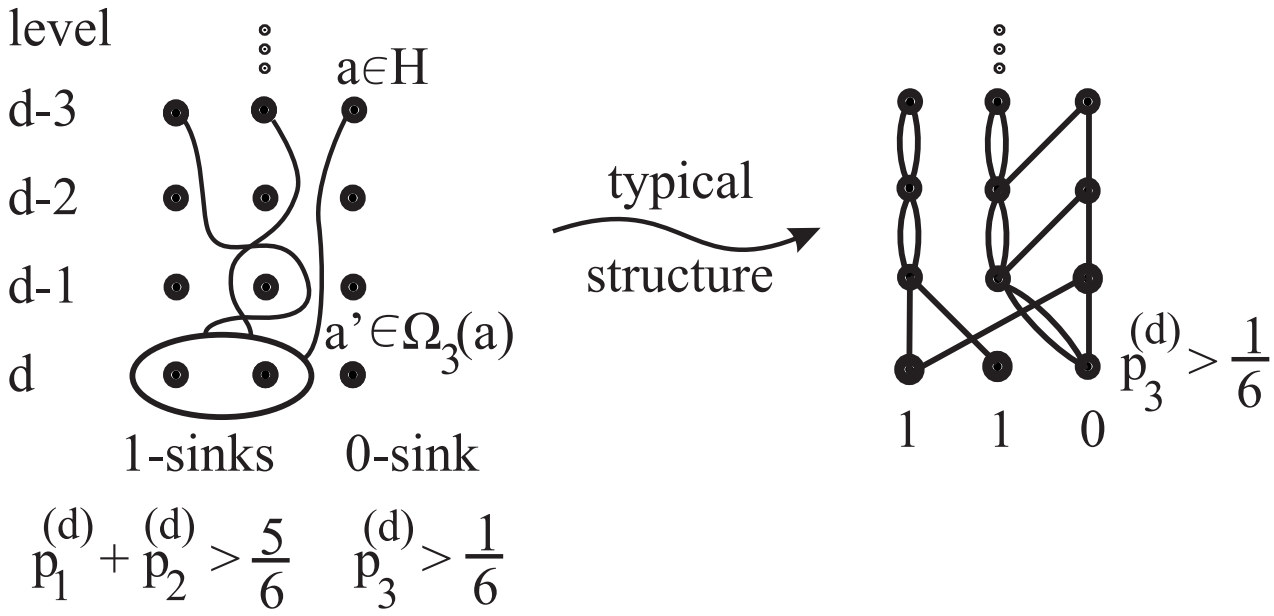
The richness condition expresses an essential property of hitting sets for 1-BPs of width 3 while being independent of a rather technical formalism of BPs.

Idea of Proof

Let P be a normalized 1-BP of width 3 such that $\frac{\left|P^{-1}(1)\right|}{2^n} \geq \varepsilon > \frac{5}{6}$.

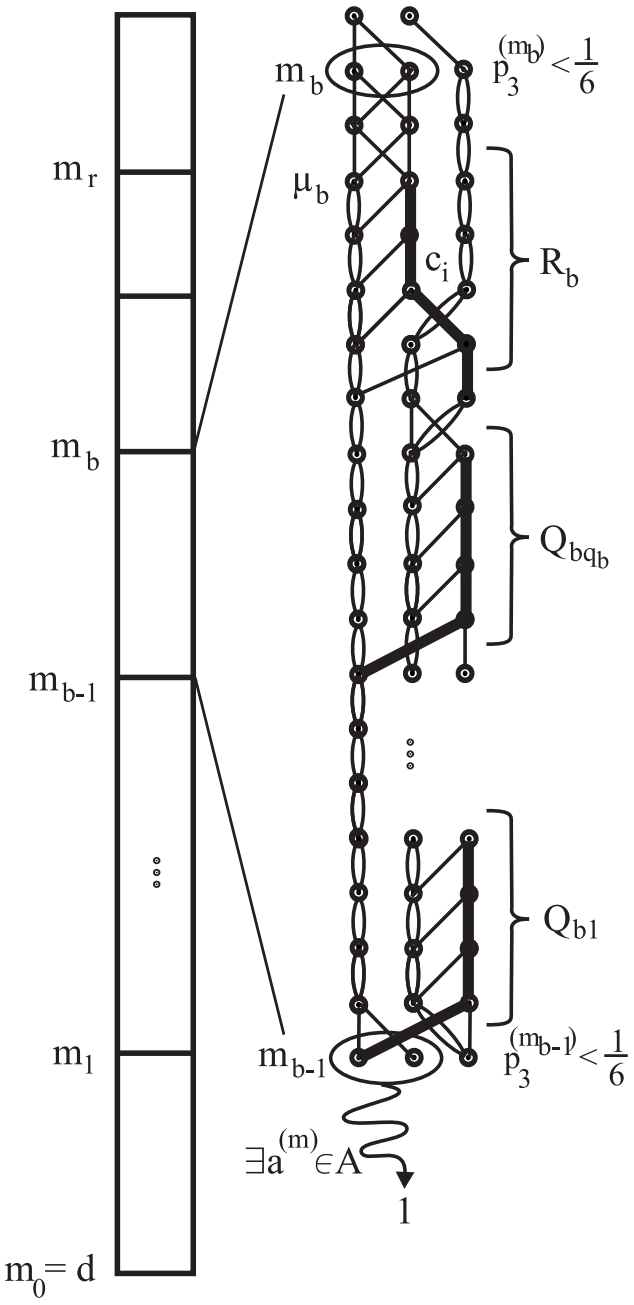
On the contrary assume that $P(a) = 0$ for every $a \in H = \Omega_3(A)$, which constrains the structure of P:

A Simple Example: There are no three paths of length at most 3 starting in the three different nodes at one level and leading to 1-sinks at level d (= depth).



Starting from the last level d, the structure of P is analyzed inductively block after block to lower levels (very complex case analysis).

A Block-Like Structure of BP P



R_b (resp. $Q_{b1}, \ldots, Q_{bq_b}$) contains the indices of the variables that are queried on the corresponding 'boldface' computational path whose edge labels define relevant bits $c_i \in \{0,1\}$ $(i \in R_b)$ so that any input passing through this path that differs from c (resp. agrees with c) reaches the double-edge path in the first column

$$\prod_{b=1}^{r} \left(1 - \frac{1}{2^{|R_b|}}\right) \geq \varepsilon$$

$$\longrightarrow \quad (\forall\, b \in \{1, \ldots, r\})\,(\exists\, i \in R_b)\; a_i \neq c_i$$

Recursive Step:

1. Either $\prod_{b=1}^{r+1} \left(1 - \frac{1}{2^{|R_b|}}\right) \geq \varepsilon$

   $\longrightarrow$ continue in the analysis with block $r+1$,

2. or there is Q such that $|Q| \leq \log n$ among

   $Q_{11}, \ldots, Q_{1q_1}, \ldots, Q_{r+1,1}, \ldots, Q_{r+1,q_{r+1}}$

   $\longrightarrow \quad (\forall\, i \in Q)\; a_i = c_i$

Conclusion & Open Problems

- the explicit polynomial-time construction of a hitting set for 1-BPs of width 3
- 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 ?