Almost k-Wise Independent Sets Establish Hitting Sets for Width-3 1-Branching Programs

Jiří Šíma, Stanislav Žák

Institute of Computer Science
Academy of Sciences of the Czech Republic

Derandomization of Space-Bounded Computation

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

pseudorandom generator $g : \{0,1\}^s \longrightarrow \{0,1\}^n\,, s \ll n$

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

$$|Pr_{x \sim U_n}[M(x) = 1] - Pr_{y \sim U_s}[M(g(y)) = 1]| \leq \varepsilon$$

where U_n is the uniform distribution on $\{0,1\}^n$ and $\varepsilon > 0$ is the error

deterministic simulation performs the computation for every fixed setting of the seed (which replaces the random string of a randomized algorithm) and approximates the probability of accepting/rejecting computations

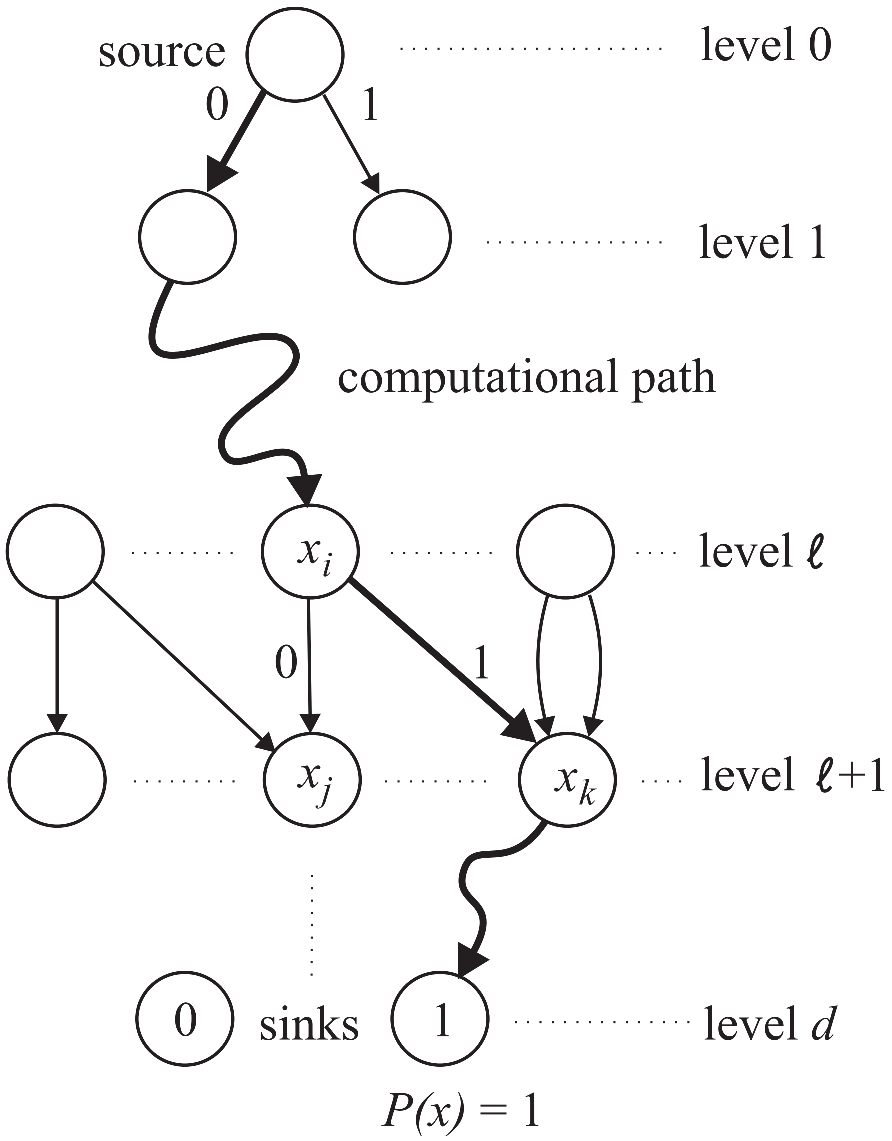
efficient derandomization (BPL=L) if there is an explicit pseudorandom generator computable in space $O(\log n)$ with seed length $O(\log n)$

Branching Program P

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

- one source $s \in V$ of zero in-degree at level 0
- two 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 two 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 $n \geq 1$

a computation that uses space $s(n)$ and runs in time $t(n)$ is modeled by P_n of width $2^{s(n)}$ and depth $t(n)$ (e.g. TM's configurations are represented by BP's nodes)

`Klivans, van Melkebeek, 1999:` if DSPACE(O(n)) requires branching programs of size $2^{\Omega(n)}$, then BPL=L.

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

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 $O(\log n)$ where $\varepsilon = O(1/n)$ (Saks, Zuckerman, 1999)

width $w = 3$: known techniques fail to improve the seed length $O(\log^2 n)$ from Nisan's result

$\longrightarrow$ **Additional Restrictions:**

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

oblivious regular 1-BPs of constant width: PRG with seed length $O(\log n \log \log n)$ where $\varepsilon = O(1/\log n)$
(Braverman, Rao, Raz, Yehudoff; Brody, Verbin, 2010)

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)

oblivious permutation 1-BPs of constant width: PRG with seed length $O\left(\log n \log \frac{1}{\varepsilon}\right)$
(Koucký, Nimbhorkar, Pudlák, 2010)

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 $P \in \mathcal{P}_n$,

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

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

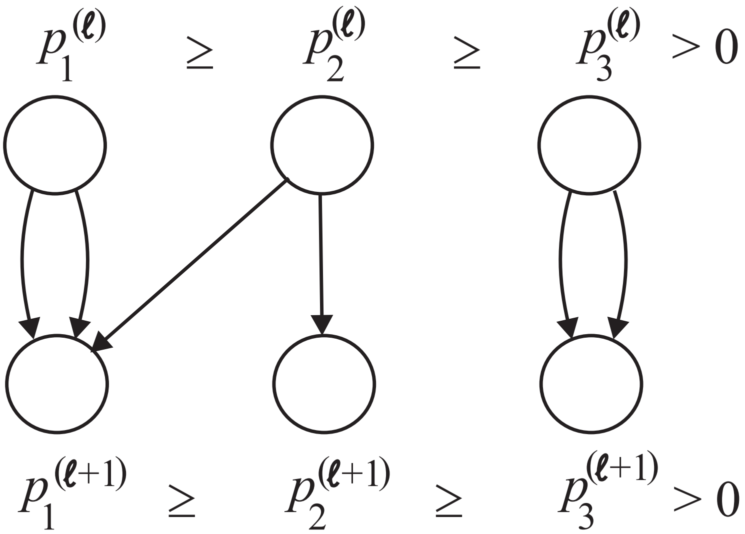
deterministic simulation of a randomized algorithm with one-sided error performs the computation for every string from the hitting set and accepts if there is at least one accepting computation

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):



a polynomial-time construction of $\left(\frac{191}{192}\right)$-hitting set for simple 1-BPs of width 3 (Šíma, Žák, 2007)

The Richness Condition

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

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

for any $Q \subseteq \{1,\ldots,n\} \setminus I$ such that $|Q| \leq \log n$, for any $c \in \{0,1\}^n$ there exists $a \in A$ that meets

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

formula (2) can be interpreted as a read-once CNF with $O(\log n)$ single literals and clauses whose sizes satisfy (1):

$$\bigwedge_{i\in Q} \ell(x_i) \wedge \bigwedge_{j=1}^{r} \bigvee_{i \in R_j} \neg\ell(x_i)$$

$$\text{where} \quad \ell(x_i) = \begin{cases} x_i & \text{for } c_i = 1 \\ \neg x_i & \text{for } c_i = 0 \end{cases}$$

for any such a read-once CNF formula, a rich set A contains at least one satisfying assignment
(i.e. A is a hitting set for this class of formulas)

Sufficiency of the Richness Condition

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 branching programs:

Theorem 1 *Let* $\varepsilon > \frac{5}{6}$. *If* A *is* ${\varepsilon'}^{11}$*-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.*

Idea of proof:

- on the contrary, a normalized 1-BP P of width 3 such that $\left|P^{-1}(1)\right|/2^n \geq \varepsilon$ and $P(a) = 0$ for every $a \in H$, is assumed
- starting from the last level, the structure of P is inductively analyzed block after block (corresponding to partition classes R_j) until a set Q ($|Q| \leq \log n$) suitable for the richness condition is found
- the richness condition is employed to achieve a contradiction
- the proof includes a rather tedious case analysis, e.g. decreasing the lower bound for ε from the original $\sqrt{12/13}$ to $5/6$ increases significantly the number of cases to be analyzed

The Necessary Condition

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 $\{R_1,\ldots,R_r,Q_1,\ldots,Q_q\}$ of I ($r \geq 0$, $q \geq 0$) satisfying

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

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

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

Any ε-rich set is weakly ε-rich: condition (3) implies that there is $j \in \{1,\ldots,q\}$ such that $|Q_j| \leq \log n$

formula (4) can be interpreted as a read-once conjunction of DNFs and CNFs whose sizes satisfy (3):

$$\bigvee_{j=1}^{q}\bigwedge_{i\in Q_j}\ell(x_i)\wedge\bigwedge_{j=1}^{r}\bigvee_{i\in R_j}\neg\ell(x_i)$$

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

The Main Result

Any almost $O(\log n)$-wise independent set is ε-rich.

(k,β)-wise independent set $A \subseteq \{0,1\}^n$: 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 $c \in \{0,1\}^n$

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

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

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

Theorem 3 *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: 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 $\varepsilon > 5/6$.

Idea of Proof

Let A be a $(\lceil (C+2)\log n\rceil, \beta)$-wise independent set.

We will show that A is ε-rich:

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

In order to show for a given $c \in \{0,1\}^n$ that there is $a \in A$ that meets

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

we will prove that the probability

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

Intuition:

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

The Main Steps of the Proof

Modifications of Partition Classes:

- superlogarithmic cardinalities:
  $R'_j \subseteq R_j$ so that $|R'_j| \leq \log n$
- small constant cardinalities:
  $R_{\leq\sigma} = \bigcup_{|R'_j|\leq\sigma} R'_j$ where σ is a suitable constant

$\longrightarrow \quad Q' = Q \cup R_{\leq\sigma}, \quad c'_i = 1 - c_i$ for $i \in R_{\leq\sigma}$

Lemma: $p \geq \dfrac{\left|A^{Q'}(c') \setminus \bigcup_{j=1}^{r'} A^{R'_j}(c')\right|}{|A|}$

$\downarrow$ Bonferroni inequality

$$p \geq \sum_{k=0}^{C'} (-1)^k \sum_{1\leq j_1<j_2<\cdots<j_k\leq r'} \frac{\left|A^{\bigcup_{i=1}^{k} R'_{j_i} \cup Q'}(c')\right|}{|A|}$$

$\downarrow$ Almost $O(\log n)$-wise independence

$$p \geq \frac{1}{2^{|Q'|}} \left(\sum_{k=0}^{C'} (-1)^k \sum_{1\leq j_1<j_2<\cdots<j_k\leq r'} \prod_{i=1}^{k} \frac{1}{2^{\left|R'_{j_i}\right|}} - \frac{\varepsilon'}{8} \right)$$

The Main Steps of the Proof II

Grouping the Classes of the Same Cardinalities

$\sigma < s_1, \ldots, s_{m'} \leq \log n$ ... cardinalities of R'_j

$r_i = \left|\{j \,|, \left|R'_j\right| = s_i\}\right|$ ... # classes of cardinality s_i

$$p > \frac{1}{n^2}\left(\sum_{k=0}^{C'}(-1)^k \sum_{\substack{k_1+\cdots+k_{m'}=k \\ 0\leq k_1\leq r_1,\ldots,0\leq k_{m'}\leq r_{m'}}} \prod_{i=1}^{m'} \frac{t_i^{k_i}}{k_i!} \prod_{j=1}^{k_i-1}\left(1-\frac{j}{r_i}\right) - \frac{\varepsilon'}{8}\right)$$

where $t_i = \dfrac{r_i}{2^{s_i}}$

Frequent Cardinalities

$r_1 > r_2 > \cdots > r_{m''} > \varrho$ where ϱ is a suitable constant

$$p > \frac{1}{n^2}\left(\sum_{k=0}^{C'}(-1)^k \sum_{\substack{k_1+\cdots+k_{m''}=k \\ k_1\geq 0,\ldots,k_{m''}\geq 0}} \prod_{i=1}^{m''} \frac{t_i^{k_i}}{k_i!} - \frac{\varepsilon'}{2}\right)$$

The Main Steps of the Proof III

$\big\downarrow$ Multinomial theorem

$$p > \frac{1}{n^2}\left(\sum_{k=0}^{C'} \frac{\left(-\sum_{i=1}^{m''} t_i\right)^k}{k!} - \frac{\varepsilon'}{2}\right)$$

$\big\downarrow$ Taylor's theorem

$$p > \frac{1}{n^2}\left(e^{-\sum_{i=1}^{m''} t_i} - \mathcal{R}_{C'+1}\left(-\sum_{i=1}^{m''} t_i\right) - \frac{\varepsilon'}{2}\right)$$

$\big\downarrow$ $\sum_{i=1}^{m} t_i < \ln \frac{1}{\varepsilon'}$

Lagrange remainder $\mathcal{R}_{C'+1}\left(-\sum_{i=1}^{m''} t_i\right) < \frac{\varepsilon'}{4}$

$$p > \frac{\varepsilon'}{4n^2} > 0 \qquad \square$$

Conclusion & Open Problems

- the explicit polynomial-time construction of a hitting set for 1-BPs of width 3
- an important step in the effort to construct HSGs for 1-BPs of bounded width

  $$\times$$

  such constructions were known only for width 2 and 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 ?