Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 88 (2010)

A Polynomial Time Construction of a Hitting Set for Read-Once Branching Programs of Width 3*

Jiří Šíma sima@cs.cas.cz
Stanislav Žák stan@cs.cas.cz

Institute of Computer Science,
Academy of Sciences of the Czech Republic,
P.O. Box 5, 182 07 Prague 8, Czech Republic

Abstract

The relationship between deterministic and probabilistic computations is one of the central issues in complexity theory. This problem can be tackled by constructing polynomial time hitting set generators which, however, belongs to the hardest problems in computer science even for severely restricted computational models. We consider read-once branching (1-branching) programs of polynomial size for which such constructions have been known only in the case of width 2 and in very restricted cases of bounded width (e.g. permutation or regular oblivious read-once branching programs). In this paper, we characterize the hitting sets for read-once branching programs of width 3 by a necessary so-called richness condition, which is independent of a rather technical formalism of branching programs. This condition proves to be sufficient in a sense that any rich set extended with all strings within Hamming distance of 3 is a hitting set for width-3 1-branching programs that are weakly oblivious (i.e. at each level where only one node branches to its two different successors, all nodes are labeled with the same variable). Then, we prove that any almost $O(\log n)$-wise independent set satisfies the richness condition. By using such a set due to Alon et al. (1992) our result provides an explicit polynomial time construction of a hitting set for (weakly oblivious) read-once branching programs of width 3 with acceptance probability $\varepsilon > \sqrt{12/13}$.

1 Introduction

An *ε-hitting set* for a class of Boolean functions of n variables is a set $H \subseteq \{0,1\}^n$ such that for every function f in the class, the following is satisfied: If a random input is accepted

*The authors would like to thank Pavel Pudlák for pointing out the problem of hitting sets for width-3 1-branching programs. This research was partially supported by projects GA ČR P202/10/1333, MŠMT ČR 1M0545, and AV0Z10300504.

ISSN 1433-8092

by f with probability at least ε, then there is also an input in H that is accepted by f. An efficiently constructible sequence of hitting sets for increasing n is a straightforward generalization of the *hitting set generator* introduced in [5], which is a weaker (one-sided error) version of *pseudorandom generator* [9]. For the class of Boolean functions of polynomial complexity in any reasonable model, it is easy to prove the existence of ε-hitting set of polynomial size, if $\varepsilon > 1/n^c$ for a constant c and n is the number of variables. The proof is nonconstructive, since it uses a counting argument. An important problem in complexity theory is to find polynomial time constructible hitting sets for functions of polynomial complexity in different standard models like circuits, formulas, branching programs etc. Such constructions would have consequences for the relationship between deterministic and probabilistic computations in the respective models.

Looking for polynomial time constructions of hitting sets for unrestricted models belongs to the hardest problems in computer science. Hence, restricted models are investigated. We consider *read-once branching (1-branching) programs* of polynomial size, which is a restricted model of space-bounded computations [11] for which pseudorandom generators with seed length $O(\log^2 n)$ have been known for a long time through a result of Nisan [8]. Recently, considerable attention has been paid to improving this to $O(\log n)$ in the constant-width case, which is a fundamental problem with many applications in circuit lower bounds and derandomization [7]. The problem has been resolved for width 2 but the known techniques provably fail for width 3 [2, 7], which applies even to hitting set generators [4].

In the case of width 3, we do not know of any significant improvement over Nisan's result except for severely restricted so-called *regular* (oblivious) read-once branching programs of constant width having the in-degree of every vertex equal to 2, for which pseudorandom generators have recently been constructed with seed length $O(\log n \log\log n)$ [3, 4]. There has also been some recent progress in the case of *permutation* (oblivious) read-once branching programs of bounded width whose edges labeled with 0 (respectively 1) define a one to one mapping for each level-to-level transition [7], for which a pseudorandom generator has been constructed with seed length $O(\log n)$ [6]. In our previous work [10], we have made the first step for finding a polynomial time constructible hitting set for width 3. Using the result due to Alon et al. [1] we achieved such a construction if an additional, rather technical restriction is imposed on the program structure. For example, this restriction is met if one special pattern of level-to-level transitions in a normalized form of so-called *simple* width-3 1-branching programs is excluded, which covers the regular and permutation cases (see [10] for further details).

In the present paper, we provide a polynomial time construction of a hitting set for (weakly oblivious) read-once branching programs of width 3 with acceptance probability $\varepsilon > \sqrt{12/13}$, which is an important step in the effort of constructing hitting set generators for the model of read-once branching programs of bounded width. For this purpose, we first characterize the hitting sets for width-3 1-branching programs by a necessary so-called *richness* condition which is independent of the notion of branching programs. This richness condition proves to be 'sufficient' in a sense that any rich set extended with all strings within

Hamming distance of 3 is a hitting set for *weakly oblivious*[1] read-once branching programs of width 3. Our approach is based on a detailed analysis of structural properties of width-3 1-branching programs, which differs from the known techniques. Then, we prove that any almost $(C \log n)$-wise independent set which can be constructed in polynomial time by the result due to Alon et al. [1] satisfies this richness condition for a suitable constant C, which implies our result.

The paper is organized as follows. After a brief review of basic definitions regarding branching programs in Section 2 (see [11] for more information), the richness condition is formulated and proven to be necessary in Section 3 while its sufficiency is presented in Section 4. The subsequent four Sections 5–8 are devoted to the technical proof of this proposition. In addition, the richness of any almost $O(\log n)$-wise independent set is shown in Section 9. Finally, our result is summarized in Section 10.

2 Normalized Width-w 1-Branching Programs

A *branching program* P on the set of input Boolean variables $X_n = \{x_1, \ldots, x_n\}$ is a directed acyclic multi-graph $G = (V, E)$ that has one *source* $s \in V$ of zero in-degree and, except for *sinks* of zero out-degree, all the *inner* (non-sink) nodes have out-degree 2. In addition, the inner nodes get labels from X_n and the sinks get labels from $\{0, 1\}$. For each inner node, one of the outgoing edges gets the label 0 and the other one gets the label 1. The branching program P computes Boolean function $P : \{0,1\}^n \longrightarrow \{0,1\}$ as follows. The computational path of P for an input $\mathbf{a} = (a_1, \ldots, a_n) \in \{0,1\}^n$ starts at source s. At any inner node labeled by $x_i \in X_n$, input variable x_i is tested and this path continues with the outgoing edge labeled by a_i to the next node, which is repeated until the path reaches the sink whose label gives the output value $P(\mathbf{a})$. Denote by $P^{-1}(a) = \{\mathbf{a} \in \{0,1\}^n \mid P(\mathbf{a}) = a\}$ the set of inputs for which P outputs $a \in \{0, 1\}$. For inputs of arbitrary lengths, infinite families $\{P_n\}$ of branching programs, each P_n for one input length $n \geq 1$, are used.

A branching program P is called *read-once* (or shortly *1-branching* program) if every input variable from X_n is tested at most once along each computational path. Here we consider *leveled* branching programs in which each node belongs to a level, and edges lead from level $k \geq 0$ only to the next level $k + 1$. We assume that the source of P creates level 0, whereas the last level is composed of all sinks. The number of levels decreased by 1 equals the *depth* of P which is the length of its longest path, and the maximum number of nodes on one level is called the *width* of P. In addition, P is called *oblivious* if at each level only one variable is queried. We also say that P is *weakly oblivious* if at each level where only one node branches to two different nodes (i.e. the remaining nodes at this level have outgoing double edges), all nodes are labeled with the same variable.

For a 1-branching program P of width w define a $w \times w$ *transition matrix* T_k on level $k \geq 1$ such that $t_{ij}^{(k)} \in \{0, \frac{1}{2}, 1\}$ is the half of the number of edges leading from node $v_j^{(k-1)}$

[1] In fact, this is a technical assumption used in our sufficiency proof only for resolving one rather specific case[2] which, as we believe, could be removed.

$(1 \le j \le w)$ on level $k-1$ of P to node $v_i^{(k)}$ $(1 \le i \le w)$ on level k. For example, $t_{ij}^{(k)} = 1$ implies there is a *double edge* from $v_j^{(k-1)}$ to $v_i^{(k)}$. Clearly, $\sum_{i=1}^{w} t_{ij}^{(k)} = 1$ since this sum equals the half of the out-degree of inner node $v_j^{(k-1)}$, and $2 \cdot \sum_{j=1}^{w} t_{ij}^{(k)}$ is the in-degree of node $v_i^{(k)}$. Denote by a column vector $\mathbf{p}^{(k)} = (p_1^{(k)}, \ldots, p_w^{(k)})^{\mathsf{T}}$ the *distribution* of inputs among w nodes on level k of P, that is, $p_i^{(k)}$ is the probability that a random input is tested at node $v_i^{(k)}$, which equals the ratio of the number of inputs from $M(v_i^{(k)}) \subseteq \{0,1\}^n$ that are tested at $v_i^{(k)}$ to all 2^n possible inputs. It follows $\bigcup_{i=1}^{w} M(v_i^{(k)}) = \{0,1\}^n$ and $\sum_{i=1}^{w} p_i^{(k)} = 1$ for every level $k \ge 0$. Given the distribution $\mathbf{p}^{(k-1)}$ on level $k-1$, the distribution on the subsequent level k can be computed using transition matrix T_k as

$$\mathbf{p}^{(k)} = T_k \cdot \mathbf{p}^{(k-1)} . \tag{1}$$

It is because the ratio of inputs coming to node $v_i^{(k)}$ from previous-level nodes equals $p_i^{(k)} = \sum_{j=1}^{w} t_{ij}^{(k)} p_j^{(k-1)}$ since each of the two edges outgoing from node $v_j^{(k-1)}$ distributes exactly the half of the inputs tested at $v_j^{(k-1)}$.

We say that a 1-branching program P of width w is *normalized* if P has the minimum depth among the programs computing the same function (e.g. P does not contain the identity transition T_k) and P satisfies

$$1 > p_1^{(k)} \ge p_2^{(k)} \ge \cdots \ge p_w^{(k)} > 0 \tag{2}$$

for every $k \ge \log w$ (hereafter, log denotes the binary logarithm). Obviously, condition (2) can always be met mainly by possible splitting and permuting the nodes at each level of P:

Lemma 1 ([10]) *Any width-w 1-branching program can be normalized.*

In the sequel, we confine ourselves to the 1-branching programs of width $w = 3$. Any such normalized program P satisfies $p_1^{(k)} + p_2^{(k)} + p_3^{(k)} = 1$ and $1 > p_1^{(k)} \ge p_2^{(k)} \ge p_3^{(k)} > 0$, which implies

$$p_1^{(k)} > \frac{1}{3} , \qquad p_2^{(k)} < \frac{1}{2} , \qquad p_3^{(k)} < \frac{1}{3} \tag{3}$$

for every level $2 \le k \le d$ where $d \le n$ is the depth of P.

3 A Necessary Condition

Let $\mathcal{P}$ be a class of branching programs and $\varepsilon > 0$ be a real constant. A set of input strings $H \subseteq \{0,1\}^*$ is called an *ε-hitting set* for class $\mathcal{P}$ if for sufficiently large n, for every branching program $P \in \mathcal{P}$ with n input variables

$$\frac{|P^{-1}(1)|}{2^n} \ge \varepsilon \quad \text{implies} \quad (\exists\, \mathbf{a} \in H \cap \{0,1\}^n)\, P(\mathbf{a}) = 1 . \tag{4}$$

Furthermore, we say that a set $A \subseteq \{0,1\}^*$ is *ε-rich* if for sufficiently large n, 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 where $q \geq 0$ and $r \geq 0$ the following implication holds: If

$$\left(1-\prod_{j=1}^{q}\left(1-\frac{1}{2^{|Q_j|}}\right)\right)\times\prod_{j=1}^{r}\left(1-\frac{1}{2^{|R_j|}}\right)\geq\varepsilon\,, \tag{5}$$

then for any $\mathbf{c}\in\{0,1\}^n$ there exists $\mathbf{a}\in A\cap\{0,1\}^n$ such that

$$(\exists\, j\in\{1,\ldots,q\})\,(\forall\, i\in Q_j)\,a_i=c_i \tag{6}$$

$$\text{and}\quad (\forall\, j\in\{1,\ldots,r\})\,(\exists\, i\in R_j)\,a_i\neq c_i\,. \tag{7}$$

Particularly for $q=0$ inequality (5) reads

$$\prod_{j=1}^{r}\left(1-\frac{1}{2^{|R_j|}}\right)\geq\varepsilon \tag{8}$$

and conjunction (6) and (7) reduces to the second conjunct (7), while for $r=0$ inequality (5) reads

$$1-\prod_{j=1}^{q}\left(1-\frac{1}{2^{|Q_j|}}\right)\geq\varepsilon \tag{9}$$

and conjunction (6) and (7) reduces to the first conjunct (6).

Theorem 1 *Every ε-hitting set for the class of read-once branching programs of width 3 is ε-rich.*

Proof: We proceed by transposition. Assume a set $H\subseteq\{0,1\}^*$ is not ε-rich which means that for infinitely many n there is an index set $I\subseteq\{1,\ldots,n\}$, a partition $\{Q_1,\ldots,Q_q,R_1,\ldots,R_r\}$ of I satisfying (5), and a string $\mathbf{c}\in\{0,1\}^n$ such that every $\mathbf{a}\in H\cap\{0,1\}^n$ meets

$$(\forall\, j\in\{1,\ldots,q\})\,(\exists\, i\in Q_j)\,a_i\neq c_i \tag{10}$$

$$\text{or}\quad (\exists\, j\in\{1,\ldots,r\})\,(\forall i\in R_j)\,a_i=c_i\,. \tag{11}$$

We will use this partition and $\mathbf{c}$ for constructing a (non-normalized oblivious) width-3 1-branching program $P\in\mathcal{P}$ such that

$$\frac{|P^{-1}(1)|}{2^n}\geq\varepsilon\quad\text{and}\quad(\forall\,\mathbf{a}\in H\cap\{0,1\}^n)\,P(\mathbf{a})=0\,, \tag{12}$$

which negates that H is an ε-hitting set for $\mathcal{P}$ according to (4).

We assume $q\geq 1$, $r\geq 1$, and $|Q_q|>1$, while the proof for $q=0$ or $r=0$ or $|Q_q|=1$ is similar. As depicted in Figure 1, branching program P is composed of $q+r$ consecutive blocks corresponding to the partition classes $Q_1,\ldots,Q_q,R_1,\ldots,R_r$ which determine the indices of variables that are tested within these blocks. The block associated with Q_j

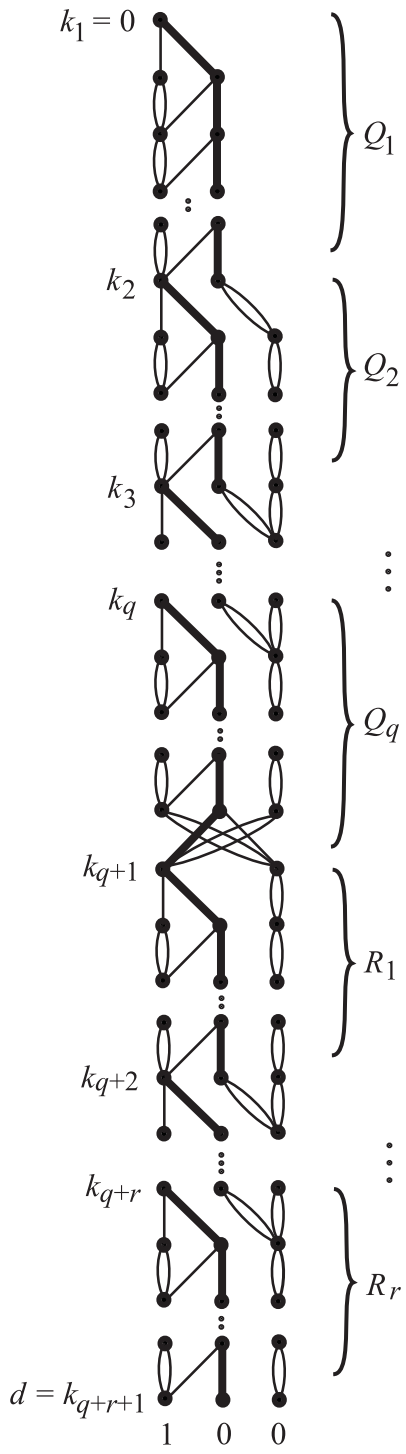

Figure 1: The Necessary Condition.

for $j \in \{1, \ldots, q\}$ starts on level $k_j = \sum_{\ell=1}^{j-1} |Q_\ell|$ of P (e.g. $k_1 = 0$) with a transition satisfying $t_{11}^{(k_j+1)} = t_{21}^{(k_j+1)} = \frac{1}{2}$, followed by a sequence of transitions that meet $t_{11}^{(k)} = 1$ and $t_{12}^{(k)} = t_{22}^{(k)} = \frac{1}{2}$ for every $k = k_j + 2, \ldots, k_j + |Q_j|$, except for the boundary level $k_q + |Q_q| = k_{q+1}$, which is defined below. In addition, there is a parallel double-edge path leading from the node $v_3^{(k_2+1)}$ on level $k_2 + 1$ up to node $v_3^{(k_{q+1}-1)}$, and thus $t_{33}^{(k)} = 1$ for every $k = k_2 + 2, k_2 + 3, \ldots, k_{q+1} - 1$. This path is wired up by $q - 1$ double edges coming from nodes $v_2^{(k_j)}$, that is, $t_{32}^{(k_j+1)} = 1$ for every $j = 2, \ldots, q$. Finally, a special boundary transition is defined on level k_{q+1} as $t_{31}^{(k_{q+1})} = t_{13}^{(k_{q+1})} = 1$ and $t_{12}^{(k_{q+1})} = t_{32}^{(k_{q+1})} = \frac{1}{2}$. Note that there are only two nodes $v_1^{(k_{q+1})}, v_3^{(k_{q+1})}$ on the boundary level k_{q+1}. Furthermore, P continues analogously with blocks corresponding to R_j for $j = 1, \ldots, r$, each starting on level $k_{q+j} = k_{q+1} + \sum_{\ell=1}^{j-1} |R_\ell|$ (e.g. $k_{q+r+1} = d$ is the depth of P) with the transition satisfying $t_{11}^{(k_{q+j}+1)} = t_{21}^{(k_{q+j}+1)} = \frac{1}{2}$, followed by $t_{11}^{(k)} = 1$ and $t_{12}^{(k)} = t_{22}^{(k)} = \frac{1}{2}$ for every $k = k_{q+j} + 2, \ldots, k_{q+j} + |R_j|$, including the parallel double-edge path, that is, $t_{33}^{(k)} = 1$ for every $k = k_{q+1}+1, \ldots, d$ and $t_{32}^{(k_{q+j}+1)} = 1$ for every $j = 2, \ldots, r$. The branching program P then queries the value of each variable x_i such that $i \in Q_j$ for some $j \in \{1, \ldots, q\}$ or $i \in R_j$ for some $j \in \{1, \ldots, r\}$ only on one level $k \in \{k_j, \ldots, k_{j+1}-1\}$ or $k \in \{k_{q+j}, \ldots, k_{q+j+1}-1\}$, respectively (i.e. the nodes on level k are labeled with x_i), while the single edge leading to $v_2^{(k+1)}$ (or to $v_1^{(k_{q+1})}$ for $k = k_{q+1} - 1$) on the subsequent level $k + 1$ (indicated by a bold line in Figure 1) gets label c_i. Finally, the sink $v_1^{(d)}$ gets label 1, whereas the sinks $v_2^{(d)}$, $v_3^{(d)}$ are labeled with the output 0, which completes the construction of P.

Clearly, P is an (oblivious) read-once branching program of width 3. The probability that an input reaches the node $v_3^{(k_{q+1})}$ on the boundary level k_{q+1} can simply be computed as

$$p_3^{(k_{q+1})} = \prod_{j=1}^{q} \left(1 - \frac{1}{2^{|Q_j|}}\right), \tag{13}$$

while the probability of the complementary event that an input reaches $v_1^{(k_{q+1})}$ equals $p_1^{(k_{q+1})} = 1 - p_3^{(k_{q+1})}$. Therefore, the probability that P outputs 1 can be expressed and lower bounded by (5):

$$\frac{|P^{-1}(1)|}{2^n} = p_1^{(d)} = \left(1 - \prod_{j=1}^{q} \left(1 - \frac{1}{2^{|Q_j|}}\right)\right) \times \prod_{j=1}^{r} \left(1 - \frac{1}{2^{|R_j|}}\right) \geq \varepsilon\,. \tag{14}$$

Furthermore, we split $H \cap \{0,1\}^n = A_1 \cup A_2$ into two parts so that every $\mathbf{a} \in A_1$ satisfies the first term (10) of the underlying disjunction, whereas every $\mathbf{a} \in A_2 = H \setminus A_1$ meets the second term (11). Thus, for any input $\mathbf{a} \in A_1$ and for every $j \in \{1, \ldots, q\}$ the block of P corresponding to Q_j contains a level $k \in \{k_j, \ldots, k_{j+1}-1\}$ where variable x_i is tested such that $a_i \neq c_i$. This ensures that the computational path for $\mathbf{a} \in A_1$ reaches $v_3^{(k_{q+1})}$ and further continues through $v_3^{(k_{q+1}+1)}, \ldots, v_3^{(d)}$, which gives $P(\mathbf{a}) = 0$ for every $\mathbf{a} \in A_1$. Similarly, for any input $\mathbf{a} \in A_2$ there exists a block of P corresponding to R_j for some $j \in \{1, \ldots, r\}$ such

that the computational path for $\mathbf{a}$ traverses nodes $v_1^{(k_{q+j})}, v_2^{(k_{q+j}+1)}, v_2^{(k_{q+j}+2)}, \ldots, v_2^{(k_{q+j}+|R_j|)}$. For $j < r$ this path continues through $v_3^{(k_{q+j+1}+1)}, \ldots, v_3^{(d)}$, whereas for $j = r$ it terminates at $v_2^{(d)}$, which gives $P(\mathbf{a}) = 0$ in both cases. Hence, P satisfies (12), which completes the proof. □

4 A Sufficient Condition

In the following theorem, we formulate the sufficiency of the richness condition introduced in Section 3. For an input $\mathbf{a} \in \{0,1\}^n$ and an integer constant $c \geq 0$, denote by $\Omega_c(\mathbf{a}) = \{\mathbf{a}' \in \{0,1\}^n \mid h(\mathbf{a}, \mathbf{a}') \leq c\}$ the set of so-called *h-neighbors* of $\mathbf{a}$, where $h(\mathbf{a}, \mathbf{a}')$ is the Hamming distance between $\mathbf{a}$ and $\mathbf{a}'$ (i.e. the number bits in which $\mathbf{a}$ and $\mathbf{a}'$ differ). We also define $\Omega_c(A) = \bigcup_{\mathbf{a} \in A} \Omega_c(\mathbf{a})$ for a given set $A \subseteq \{0,1\}^*$.

Theorem 2 *Denote $\delta = \sqrt{\frac{12}{13}}$. If A is $(\delta^{11} - \delta^{12})\varepsilon^{12}$-rich for $\varepsilon > \delta$ then $H = \Omega_3(A)$ is an ε-hitting set for the class of weakly oblivious read-once branching programs of width 3.*

Proof: After using Lemma 1, suppose a normalized weakly oblivious read-once branching program P of width 3 with sufficiently many input variables n meets

$$\frac{|P^{-1}(1)|}{2^n} \geq \varepsilon > \delta > \frac{11}{12}\,. \tag{15}$$

We will prove that there exists $\mathbf{a} \in H$ such that $P(\mathbf{a}) = 1$. On the contrary, we assume that $P(\mathbf{a}) = 0$ for every $\mathbf{a} \in H$. The main idea of the proof lies in using this assumption first for constraining the structure of branching program P so that the richness of A can eventually be employed to disprove this assumption.

4.1 The Plan of Proof

We start the underlying analysis of the structure of P from its last level d containing the sinks and we go backwards block after block to lower levels. In particular, we inspect the structure of a block whose *last* **level m** ($m = d$ at the beginning) satisfies the following four so-called m-conditions:

1. $t_{11}^{(m)} = t_{21}^{(m)} = \frac{1}{2}$,
2. $t_{32}^{(m)} > 0$,
3. $p_3^{(m)} < \frac{1}{12}$,
4. there is $\mathbf{a}^{(m)} \in A$ such that if we put $\mathbf{a}^{(m)}$ at node $v_1^{(m)}$ or $v_2^{(m)}$, then its onward computational path arrives to the sink labeled with 1.

Using the knowledge of the block structure (Sections 4 and 5), we define the partition classes $R, Q_1, \ldots, Q_q$ associated with this block and corresponding bits from input $\mathbf{c} \in \{0,1\}^n$ (Section 5). These are then used in the richness condition (6) and (7) either to find $\mathbf{a} \in H$ such that $P(\mathbf{a}) = 1$ (Section 8) if the complete partition (generated by all blocks) satisfies (5), or to ensure that the m'-conditions are also met for the first level m' of the block (Sections 6 and Section 7 particularly for m'-condition 4). In the latter case, the block analysis including the definition of associated partition classes is applied recursively for m replaced with m' etc. (Section 7).

4.2 The Initial Case of $m = d$

We will first observe that the four $m-$conditions can be met for $m = d$. Clearly, both edges outgoing from $v_1^{(d-1)}$ lead to the sink(s) labeled with 1 since $p_1^{(d-1)} > \frac{1}{3}$ due to (3) and $|P^{-1}(0)|/2^n < \frac{1}{12}$ according to (15). Hence, we will assume without loss of generality that $t_{11}^{(d)} = t_{21}^{(d)} = \frac{1}{2}$ (m-condition 1) while the remaining edges that originally led to the sinks labeled with 1 or 0 are possibly redirected to $v_1^{(d)}$ or $v_3^{(d)}$, respectively, so that the normalization condition $p_1^{(d)} \geq p_2^{(d)} > \frac{1}{6} > \frac{1}{12} > p_3^{(d)}$ (m-condition 3) is preserved by (15). Thus, sinks $v_1^{(d)}$ and $v_2^{(d)}$ are labeled with 1 (m-condition 4) whereas sink $v_3^{(d)}$ gets label 0. Finally, we show that $t_{32}^{(d)} > 0$ (m-condition 2). On the contrary, suppose $t_{32}^{(d)} = 0$, which implies $t_{33}^{(d)} > 0$ and $H \subseteq P^{-1}(0) \subseteq M(v_3^{(d-1)})$ due to $t_{31}^{(d)} = 0$. In the case of $t_{13}^{(d)} + t_{23}^{(d)} > 0$, the computational path for an h-neighbor $\mathbf{a}' \in \Omega_1(\mathbf{a})$ of $\mathbf{a} \in A \subseteq H \subseteq M(v_3^{(d-1)})$ that differs from $\mathbf{a}$ in the ith bit that is tested at node $v_3^{(d-1)}$ (i.e. $v_3^{(d-1)}$ is labeled with x_i), would reach the sink labeled with 1, and hence $P(\mathbf{a}') = 1$ which contradicts the assumption $H \subseteq P^{-1}(0)$. For $t_{33}^{(d)} = 1$, on the other hand, we could shorten P by removing the last level d while preserving its function and condition (15), which is in contradiction with the normalization of P. This completes the proof that m-conditions 1–4 can be assumed for $m = d$ without loss of generality.

4.3 A Technical Lemma

The following lemma represents a technical tool which will be used for the analysis of the block from **level μ** through m where $2 \leq \mu < m$ denotes the least level of P such that $t_{11}^{(\ell)} = 1$ for every $\ell = \mu + 1, \ldots, m-1$. For this purpose, define a *switching* path starting from $v \in \{v_2^{(k)}, v_3^{(k)}\}$ at level $\mu \leq k < m$ to be a computational path of length at most 3 edges leading from v to $v_1^{(\ell)}$ for some $k < \ell \leq \min(k+3, m)$ or to $v_2^{(m)}$ for $m \leq k+3$.

Lemma 2

(i) $3 < \mu < m - 1$.

(ii) *There are no two simultaneous switching paths starting from $v_2^{(k)}$ and from $v_3^{(k)}$, respectively, at any level $\mu \leq k < m$.*

(iii) *If* $t_{12}^{(k+1)} > 0$ *for some* $\mu \leq k < m$*, then* $t_{11}^{(\ell)} = t_{33}^{(\ell)} = 1$*,* $t_{12}^{(\ell)} = t_{22}^{(\ell)} = \frac{1}{2}$ *for every* $\ell = \mu + 1, \ldots, k$*, and* $t_{12}^{(k+1)} = \frac{1}{2}$ *(see Figure 2).*

(iv) *If* $t_{13}^{(k+1)} > 0$ *for some* $\mu < k < m$*, then one of the following four cases appears:*

1. $t_{11}^{(k)} = t_{23}^{(k)} = 1$ *and* $t_{12}^{(k)} = t_{32}^{(k)} = \frac{1}{2}$,
2. $t_{11}^{(k)} = t_{23}^{(k)} = 1$ *and* $t_{22}^{(k)} = t_{32}^{(k)} = \frac{1}{2}$,
3. $t_{11}^{(k)} = t_{22}^{(k)} = 1$ *and* $t_{13}^{(k)} = t_{33}^{(k)} = \frac{1}{2}$,
4. $t_{11}^{(k)} = t_{22}^{(k)} = 1$ *and* $t_{23}^{(k)} = t_{33}^{(k)} = \frac{1}{2}$.

In addition, if $t_{23}^{(k)} = 1$ *(case 1 or 2), then* $t_{11}^{(\ell)} = t_{33}^{(\ell)} = 1$ *and* $t_{12}^{(\ell)} = t_{22}^{(\ell)} = \frac{1}{2}$ *for every* $\ell = \mu + 1, \ldots, k - 1$ *(see Figure 2).*

Proof:

(i) For $\mu \leq 3$ there would be an h-neighbor $\mathbf{a}' \in \Omega_3(\mathbf{a})$ of input $\mathbf{a}^{(m)} \in A$ from m-condition 4, whose computational path starting from the source $v_1^{(0)}$ reaches $v_1^{(\mu)}$. Hence, $P(\mathbf{a}') = 1$ for $\mathbf{a}' \in H$ follows from $M(v_1^{(\mu)}) \subseteq M(v_1^{(m)}) \cup M(v_2^{(m)})$ and m-condition 4, which is a contradiction, and thus $\mu > 3$.

In addition, $t_{11}^{(m-1)} = 1$ because $t_{21}^{(m-1)} + t_{31}^{(m-1)} > 0$ implies $p_2^{(m-1)} > \frac{1}{6}$ and by m-condition 2 we get $p_3^{(m)} > \frac{1}{12}$, which contradicts m-condition 3.

(ii) Suppose there are two simultaneous switching paths starting from $v_2^{(k)}$ and from $v_3^{(k)}$, respectively, at some level $\mu \leq k < m$, and let $\mathbf{a}^{(m)} \in A$ be the input satisfying m-condition 4. Clearly, $\mathbf{a}^{(m)} \notin M(v_1^{(k)}) \subseteq M(v_1^{(m)}) \cup M(v_2^{(m)})$ since otherwise $P(\mathbf{a}^{(m)}) = 1$ for $\mathbf{a}^{(m)} \in H$. Thus, assume $\mathbf{a}^{(m)} \in M(v)$ for $v \in \{v_2^{(k)}, v_3^{(k)}\}$. Then there is an h-neighbor $\mathbf{a}' \in \Omega_3(\mathbf{a}^{(m)}) \cap M(v)$ of $\mathbf{a}^{(m)}$ whose computational path follows the switching path starting from v. Hence, $\mathbf{a}' \in M(v_1^{(m)}) \cup M(v_2^{(m)})$ implying $P(\mathbf{a}') = 1$ for $\mathbf{a}' \in H$ due to P is read-once. This completes the proof of (ii).

As depicted in Figure 2, denote by $v \in \{v_2^{(k)}, v_3^{(k)}\}$ a node at level $\mu < k < m$ with the edge outgoing to $v_1^{(k+1)}$, and let u be a node on level $k - 1$ from which an edge leads to v, while $v' \in \{v_2^{(k)}, v_3^{(k)}\} \setminus \{v\}$ and $u' \in \{v_2^{(k-1)}, v_3^{(k-1)}\} \setminus \{u\}$ denote the other nodes. It follows from (ii) there is no edge from u' to v nor to $v_1^{(k)}$, which would establish two simultaneous switching paths starting from $v_2^{(k-1)}$ and from $v_3^{(k-1)}$, respectively. Hence, there must be a double edge from u' to v'. Since P is normalized, $u' = v_2^{(k-1)}$ and $v' = v_3^{(k)}$ cannot happen simultaneously. Moreover, the second edge from u may lead either to $v_1^{(k)}$ or to v' if $v' \neq v_3^{(k)}$. Now, the possible cases can be summarized:

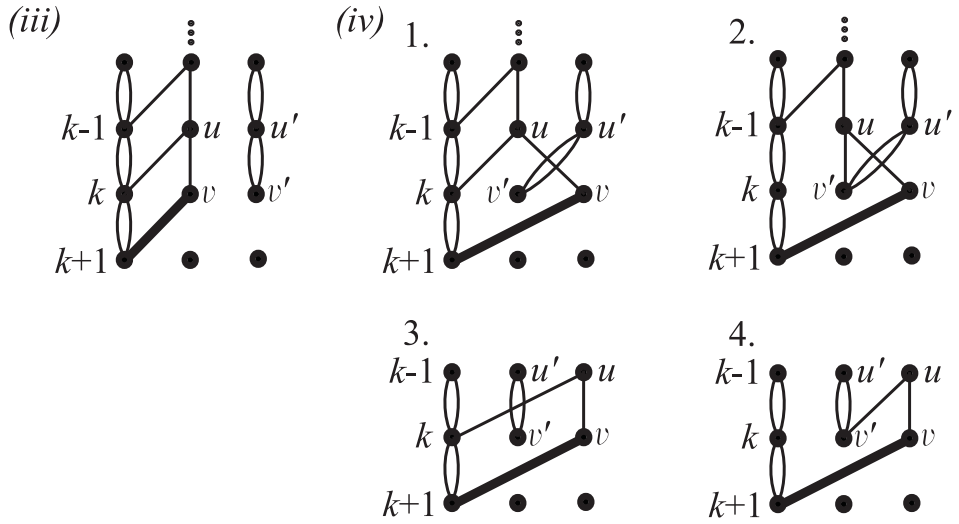

Figure 2: Lemma 2.iii and iv.

(iii) For $t_{12}^{(k+1)} > 0$ we know $v = v_2^{(k)}$ and $v' = v_3^{(k)}$, which implies $t_{11}^{(k)} = t_{33}^{(k)} = 1$ and $t_{12}^{(k)} = t_{22}^{(k)} = \frac{1}{2}$. The proposition follows when this argument is applied recursively for k replaced with $k-1$ etc. In addition, we will prove that $t_{12}^{(k+1)} < 1$ for $\mu \leq k < m$. Clearly, $t_{12}^{(m)} < 1$ from m-condition 2, and hence assume $k < m-1$. On the contrary, suppose $t_{12}^{(k+1)} = 1$, which implies $t_{23}^{(k+1)} = t_{33}^{(k+1)} = \frac{1}{2}$. For $k > \mu$ one could shorten P by identifying level k with μ without changing its function. For $k = \mu > 3$, on the other hand, there are at least two edges leading to $v_3^{(\mu)}$ because otherwise if only one edge leads to $v_3^{(\mu)}$ from $u \in \{v_1^{(\mu-1)}, v_2^{(\mu-1)}, v_3^{(\mu-1)}\}$, then either $\mathbf{a}^{(m)} \notin M(u)$, which means $\mathbf{a}^{(m)} \in M(v_1^{(\mu)}) \cup M(v_2^{(\mu)}) = M(v_1^{(\mu+1)}) \subseteq M(v_1^{(m)}) \cup M(v_2^{(m)})$ implying $P(\mathbf{a}^{(m)}) = 1$ according to m-condition 4, or $\mathbf{a}^{(m)} \in M(u)$ providing an h-neighbor $\mathbf{a}' \in \Omega_1(\mathbf{a}^{(m)}) \cap M(u) \subseteq H$ of $\mathbf{a}^{(m)}$ that differs from $\mathbf{a}^{(m)}$ in the variable that is tested at u so that $\mathbf{a}' \in M(v_1^{(\mu)}) \cup M(v_2^{(\mu)})$ implying $P(\mathbf{a}') = 1$. Hence, we could split $v_3^{(\mu)}$ into two nodes and merge $v_1^{(\mu)}$ and $v_2^{(\mu)}$ while preserving the function of P.

(iv) For $t_{13}^{(k+1)} > 0$ we know $v = v_3^{(k)}$ and $v' = v_2^{(k)}$ and the four cases listed in the proposition are obtained when the choice of $u \in \{v_2^{(k-1)}, v_3^{(k-1)}\}$ is combined with whether the second edge from u leads to $v_1^{(k)}$ or to v'. In addition, the remaining part for case 1 and 2 follows from (iii) when $k+1$ is replaced with k. In particular, we know $t_{12}^{(k)} > 0$ in case 1, while there is a switching path from $v_2^{(k-1)}$ to $v_1^{(k+1)}$ via $v_3^{(k)}$ (substituting for $t_{12}^{(k)} > 0$) in case 2 when a similar analysis applies to $v = v_2^{(k-1)}$ excluding two switching paths starting from $v_2^{(k-2)}$ and $v_3^{(k-2)}$, respectively. □

5 Definition of Partition Classes

5.1 The Block Structure from μ to ν (Definition of R)

In the following corollary, we summarize the block structure from level μ through **level $\boldsymbol{\nu}$** by using Lemma 2, where $\mu \leq \nu \leq m$ is the greatest level such that $t_{12}^{(\ell)} + t_{13}^{(\ell)} > 0$ for every $\ell = \mu + 1, \ldots, \nu$. In addition, let **level $\boldsymbol{\gamma}$** be the greatest level such that $\mu \leq \gamma \leq \nu$ and $t_{12}^{(\gamma)} > 0$ (for $\gamma > \mu$).

Corollary 1

1. $t_{11}^{(\ell)} = t_{33}^{(\ell)} = 1$ *and* $t_{12}^{(\ell)} = t_{22}^{(\ell)} = \frac{1}{2}$ *for* $\ell = \mu + 1, \ldots, \gamma - 1$ *(Lemma 2.iii),*
2. $t_{11}^{(\gamma)} = t_{23}^{(\gamma)} = 1$ *and* $t_{32}^{(\gamma)} = \frac{1}{2}$ *if* $\mu < \gamma < \nu$ *(case 1 of Lemma 2.iv),*
3. $t_{11}^{(\ell)} = t_{22}^{(\ell)} = 1$ *and* $t_{33}^{(\ell)} = \frac{1}{2}$ *for* $\ell = \gamma + 1, \ldots, \nu - 1$ *(case 3 of Lemma 2.iv),*
4. *if* $\nu > \mu$, *then* $t_{12}^{(\nu)} < 1$ *(Lemma 2.iii) and* $t_{13}^{(\nu)} < 1$ *for* $\nu < m$ *(similarly),*
5. $t_{12}^{(\ell)} = 0$ *for* $\ell = \nu + 1, \ldots, m$ *(Lemma 2.iii).*

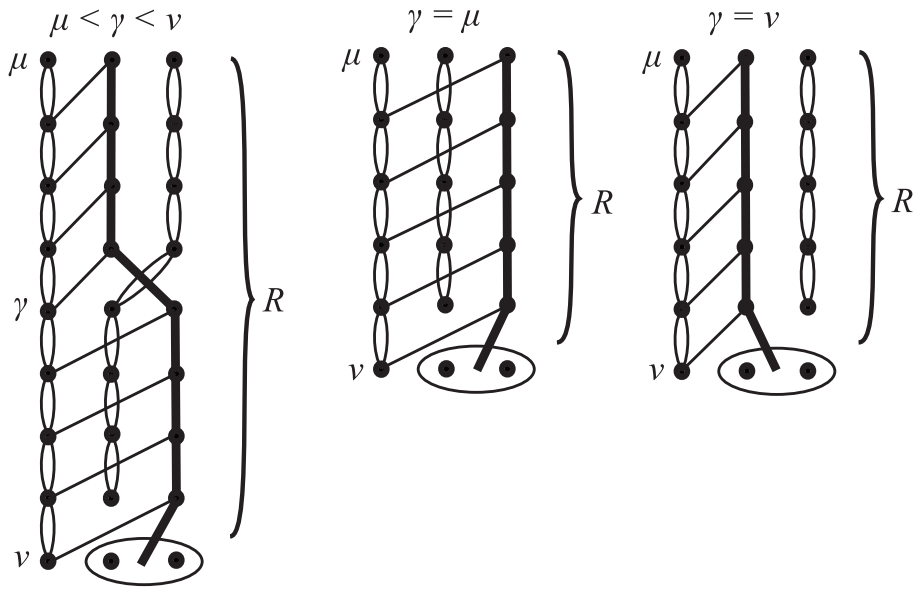

Figure 3: The block structure from level μ through $\nu < m$ according to Corollary 1.

Corollary 1 will be used for the definition of partition class R associated with the underlying block, which is illustrated in Figure 3 for $\nu < m$. Nevertheless, we will first exclude a special pathological case from this definition that occurs when $\nu = m$ and $t_{33}^{(m)} = 0$, that is, $t_{13}^{(m)} + t_{23}^{(m)} = 1$ which implies $t_{32}^{(m)} = 1$ according to Lemma 2.ii. In this case, no partition class is associated with the underlying block which is called an *empty block*. Moreover,

class R is neither defined for $\nu = \mu$ when only partition classes $Q_1, \ldots, Q_q$ are associated with the block (see Paragraph 5.2 and Lemma 3 in particular).

For a non-empty block and $\nu > \mu$, we define the partition class R to be a set of indices of the variables that are tested on the single-edge computational path $v_2^{(\mu)}, v_2^{(\mu+1)}, \ldots, v_2^{(\gamma-1)}$, $v_3^{(\gamma)}, v_3^{(\gamma+1)}, \ldots, v_3^{(\nu-1)}$ (or $v_3^{(\mu)}, v_3^{(\mu+1)}, \ldots, v_3^{(\nu-1)}$ if $\gamma = \mu$ or $v_2^{(\mu)}, v_2^{(\mu+1)}, \ldots, v_2^{(\nu-1)}$ if $\gamma = \nu$). For the future use of condition (6) and (7) we also define relevant bits of string $\mathbf{c} \in \{0,1\}^n$. Thus, let c_i be the corresponding labels of the edges creating this computational path (indicated by a bold line in Figure 3) including the edge outgoing from the last node $v_3^{(\nu-1)}$ (or $v_2^{(\nu-1)}$ if $\gamma = \nu$) that leads to $v_2^{(\nu)}$ or to $v_3^{(\nu)}$.

5.2 The Block Structure from ω to m (Definition of $Q_1, \ldots, Q_q$)

Furthermore, we define **level $\boldsymbol{\omega}$** to be the greatest level such that $\max(\nu - 1, \mu) \leq \omega \leq m$ and the double-edge path from Corollary 1 (see Figure 3) leading from $v_2^{(\mu)}$ to $v_2^{(\nu-1)}$ (for $\gamma = \mu < \nu$) or from $v_3^{(\mu)}$ to $v_2^{(\nu-1)}$ (for $\mu < \gamma < \nu$) or from $v_3^{(\mu)}$ to $v_3^{(\nu-1)}$ (for $\gamma = \nu > \mu$), or starting from $v_2^{(\mu)}$ or from $v_3^{(\mu)}$ if $\nu = \mu$, further continues up to level ω containing only nodes $v_\ell \in \{v_2^{(\ell)}, v_3^{(\ell)}\}$ for every $\ell = \mu, \ldots, \omega$. For the special case of $\omega = m$ (including the empty block) when this double-edge path reaches level m, we set $q = 0$ which means no partition classes $Q_1, \ldots, Q_q$ are associated with the underlying block. We will observe in the following lemma that $\nu > \mu$ in this case, which ensures that at least class R is defined in Paragraph 5.1 for the non-empty block when $\omega = m$.

Lemma 3 *If $\nu = \mu$, then $\omega < m$.*

Proof: On the contrary, suppose $\nu = \mu$ and $\omega = m$. Thus, $t_{12}^{(\mu+1)} = t_{13}^{(\mu+1)} = 0$ by the definition of ν. Since P is normalized, we know $t_{22}^{(\mu+1)} > 0$ and either $t_{22}^{(\mu+1)} = 1$ or $t_{23}^{(\mu+1)} = 1$ due to $\omega = m > \mu$, which implies $t_{22}^{(\ell)} = 1$ for $\ell = \mu + 2, \ldots, m - 1$. In addition, $t_{12}^{(m)} = 0$ according to Corollary 1.5, and $t_{22}^{(m)} = 0$ since $t_{22}^{(m)} = \frac{1}{2}$ would require $t_{13}^{(m)} > 0$ by the normalization of P, which contradicts Lemma 2.ii, and hence, $t_{32}^{(m)} = 1$. This gives a contradiction $\frac{1}{12} > p_3^{(m)} \geq p_2^{(\mu+1)} \geq p_2^{(\mu)}/2 > \frac{1}{12}$ according to m-condition 3 and the definition of μ. □

Thus, we will further assume $\boldsymbol{\omega} < \mathbf{m}$ throughout this Section 5. This implies $t_{12}^{(m)} = 0$ since otherwise $t_{12}^{(m)} = t_{32}^{(m)} = \frac{1}{2}$ (m-condition 2) forces $t_{33}^{(m)} = 1$ by Lemma 2.ii which would prolong the double-edge path from $v_3^{(\mu)}$ up to $v_3^{(m)}$ according to Lemma 2.iii. We will show that one can assume $t_{13}^{(m)} > 0$ without loss of generality. Suppose that $t_{13}^{(m)} = 0$, which implies $t_{22}^{(m)} = t_{23}^{(m)} = 0$ due to P is normalized, and hence $t_{32}^{(m)} = t_{33}^{(m)} = 1$. Moreover, we know $t_{11}^{(m-1)} = 1$ from Lemma 2.i and $t_{11}^{(m)} = t_{21}^{(m)} = \frac{1}{2}$ by m-condition 1. If $t_{12}^{(m-1)} = t_{13}^{(m-1)} = 0$, then $v_2^{(m-1)}$ and $v_3^{(m-1)}$ can be merged and replaced by $v_3^{(m)}$, while $v_1^{(m-1)}$ replaces $v_1^{(m-2)}$, which shortens P without changing its function. Hence, either $t_{12}^{(m-1)} > 0$ or $t_{13}^{(m-1)} > 0$ by Lemma 2.ii. In fact, $t_{12}^{(m-1)} > 0$ contradicts $\omega < m$ according to Lemma 2.iii since $t_{23}^{(m-1)} + t_{33}^{(m-1)} = t_{32}^{(m)} = t_{33}^{(m)} = 1$ can, without loss of generality, prolong the double-edge

path from $v_3^{(\mu)}$ through $v_3^{(m-2)}$ up to $v_3^{(m)}$. For $t_{13}^{(m-1)} > 0$, on the other hand, $v_2^{(m-1)}$ and $v_3^{(m-1)}$ can be merged while $v_1^{(m-1)}$ is split into two its copies, which produces $t_{11}^{(m-1)} = t_{21}^{(m-1)} = \frac{1}{2}$, $t_{32}^{(m-1)} = 1$, and $t_{11}^{(m)} = t_{21}^{(m)} = t_{12}^{(m)} = t_{22}^{(m)} = \frac{1}{2}$, $t_{33}^{(m)} = 1$. After this modification, level $m-1$ satisfies the four m-conditions 1–4 (see Paragraph 4.1) and thus, it can serve as a new level m while the original level $m > d$ (for $m = d$ program P could be shortened by removing its last level) is included in the upper-level neighboring block, which is consistent with its structure (see Paragraph 6.2 and Figure 5 in particular). Thus, we assume $t_{13}^{(m)} > 0$ without loss of generality, which implies $t_{32}^{(m)} = 1$ by Lemma 2.ii. Then Lemma 2.iv can be employed for $k = m-1$ where only case 3 and 4 may occur due to $\omega < m$ is assumed, which even implies $\omega < m-1$. In case 3, $t_{13}^{(m-1)} > 0$ and Lemma 2.iv can again be applied recursively to $k = m-2$ etc.

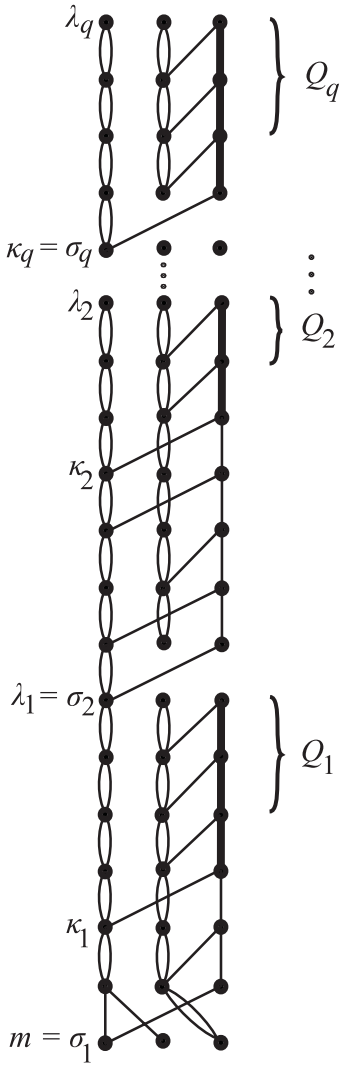

Figure 4: The definition of $Q_1, \ldots, Q_q$.

In general, starting with level $\boldsymbol{\sigma_1} = \mathbf{m}$ that meets $t_{13}^{(\sigma_j)} > 0$ for $j = 1$, we proceed to lower levels and inspect recursively the structure of subblocks indexed as j from **level $\boldsymbol{\lambda_j}$** through σ_j where $\omega \leq \lambda_j < \sigma_j - 1$ is the least level such that the transitions from case 3 or 4 of Lemma 2.iv, i.e. $t_{11}^{(\ell)} = t_{22}^{(\ell)} = 1$ and $t_{33}^{(\ell)} = \frac{1}{2}$, occur for all levels $\ell = \lambda_j + 1, \ldots, \sigma_j - 1$, as depicted in Figure 4. Note that $\lambda_j > \mu$ because $\lambda_j = \mu$ ensures $t_{22}^{(\mu+1)} = 1$ implying $\omega > \mu = \lambda_j$ by the definition of ω, which contradicts $\omega \leq \lambda_j$. In addition, we will observe that case 4 from Lemma 2.iv occurs at level $\lambda_j + 1$, that is $t_{23}^{(\lambda_j+1)} = \frac{1}{2}$. On the contrary, suppose that $t_{13}^{(\lambda_j+1)} = \frac{1}{2}$ (case 3). For $\lambda_j > \omega$, this means case 1 or 2 occurs at level $\lambda_j < \mu$ by the definition of λ_j, which would be in contradiction to $\omega \leq \lambda_j$ according to Lemma 2.iv. For $\lambda_j = \omega$, on the other hand, $t_{13}^{(\omega+1)} = \frac{1}{2}$ contradicts the definition of ω by Lemma 2.iv. This completes the argument for $t_{23}^{(\lambda_j+1)} = \frac{1}{2}$.

Furthermore, let **level $\boldsymbol{\kappa_j}$** be the least level such that $\lambda_j + 1 < \kappa_j \leq \sigma_j$ and $t_{13}^{(\kappa_j)} > 0$, which exists since at least $t_{13}^{(\sigma_j)} > 0$. Now we can define the corresponding partition class Q_j to be a set of indices of the variables that are tested on the computational path $v_3^{(\lambda_j)}, v_3^{(\lambda_j+1)}, \ldots, v_3^{(\kappa_j-2)}$, and let c_i be the corresponding labels of the edges creating this path including the edge outgoing from the last node $v_3^{(\kappa_j-2)}$ to $v_3^{(\kappa_j-1)}$ (indicated by a bold line in Figure 4), which correctly extends the definition of $\mathbf{c} \in \{0,1\}^n$ associated with R (Paragraph 5.1) and Q_k for $1 \leq k < j$ since $Q_j \cap R = \emptyset$ and $Q_j \cap Q_k = \emptyset$ due to P is weakly oblivious[2] and read-once, respectively. Finally, define next **level $\boldsymbol{\sigma_{j+1}}$** to be the greatest level such that $\omega + 1 < \sigma_{j+1} \leq \lambda_j$ and $t_{13}^{(\sigma_{j+1})} > 0$, and continue in the recursive definition of $\lambda_{j+1}, \kappa_{j+1}, Q_{j+1}$ with j replaced by $j + 1$ etc. if such σ_{j+1} exists, otherwise set $q = j$ and the definition of partition classes $Q_1, \ldots, Q_q$ associated with the underlying block is complete.

5.3 An Upper Bound on $p_1^{(m)} + p_2^{(m)}$ in Terms of $p_1^{(\omega+1)}$

In this paragraph, we will upper bound $p_1^{(m)} + p_2^{(m)}$ in terms of $p_1^{(\omega+1)}$ which will later be used for verifying the condition (5). For any $1 \leq j \leq q$, we know that $t_{11}^{(\ell)} = t_{22}^{(\ell)} = 1$ and $t_{23}^{(\ell)} = t_{33}^{(\ell)} = \frac{1}{2}$ for every $\ell = \lambda_j + 1, \ldots, \kappa_j - 1$ (see Figure 4), which gives

$$p_2^{(\kappa_j-1)} + p_3^{(\kappa_j-1)} = p_2^{(\lambda_j)} + p_3^{(\lambda_j)}, \tag{16}$$

$$p_3^{(\kappa_j-1)} = \frac{p_3^{(\lambda_j)}}{2^{|Q_j|}} \leq \frac{p_2^{(\lambda_j)} + p_3^{(\lambda_j)}}{2^{|Q_j|}}. \tag{17}$$

It follows from the definition of $\sigma_{j+1} > \omega + 1$ and equation (16) that

$$p_2^{(\sigma_{j+1})} + p_3^{(\sigma_{j+1})} = p_2^{(\lambda_j)} + p_3^{(\lambda_j)} = p_2^{(\kappa_j-1)} + p_3^{(\kappa_j-1)} \tag{18}$$

[2] The assumption that P is weakly oblivious is actually used here only for verifying $Q_q \cap R = \emptyset$ when $\lambda_q = \omega + 1$, while for $\lambda_q > \omega + 1$, the fact that P is read-once suffices.

for $1 \leq j < q$, and

$$p_2^{(\omega+1)} + p_3^{(\omega+1)} = p_2^{(\lambda_q)} + p_3^{(\lambda_q)} = p_2^{(\kappa_q-1)} + p_3^{(\kappa_q-1)} \tag{19}$$

since $t_{12}^{(\ell)} = 0$ for every $\ell = \omega+2, \ldots, m$ by Corollary 1.5 where $\nu + 1 \leq \omega + 2$ from the definition of ω. In particular, note that equation (19) is valid for the special case of $\lambda_q = \omega$ (recall $\lambda_q \geq \omega$ from the definition of λ_j) because $t_{11}^{(\lambda_q+1)} = t_{22}^{(\lambda_q+1)} = 1$ and $t_{23}^{(\lambda_q+1)} = t_{33}^{(\lambda_q+1)} = \frac{1}{2}$ (case 4 of Lemma 2.iv). Moreover, we know $t_{22}^{(\ell)} = 1$ for every $\ell = \kappa_j, \ldots, \sigma_j - 1$ and $t_{12}^{(\sigma_j)} = 0$, which implies

$$\begin{aligned} p_2^{(\sigma_j)} + p_3^{(\sigma_j)} &\geq p_2^{(\kappa_j-1)} + p_3^{(\kappa_j-1)} - p_3^{(\kappa_j-1)} \geq p_2^{(\kappa_j-1)} + p_3^{(\kappa_j-1)} - \frac{p_2^{(\lambda_j)} + p_3^{(\lambda_j)}}{2^{|Q_j|}} \\ &= \left(p_2^{(\sigma_{j+1})} + p_3^{(\sigma_{j+1})}\right)\left(1 - \frac{1}{2^{|Q_j|}}\right) \end{aligned} \tag{20}$$

for $1 < j < q$ according to (17) and (18), while formula (20) reads

$$p_3^{(m)} = p_3^{(\sigma_1)} \geq \left(p_2^{(\sigma_2)} + p_3^{(\sigma_2)}\right)\left(1 - \frac{1}{2^{|Q_1|}}\right) \tag{21}$$

for $j = 1 < q$ due to $t_{32}^{(m)} = 1$, whereas (20) is rewritten as

$$p_2^{(\sigma_q)} + p_3^{(\sigma_q)} \geq \left(p_2^{(\omega+1)} + p_3^{(\omega+1)}\right)\left(1 - \frac{1}{2^{|Q_q|}}\right) \tag{22}$$

for $j = q > 1$ according to (19). Thus starting with (21), inequality (20) is applied recursively for $j = 2, \ldots, q-1$, and, in the end, formula (22) is employed, leading to

$$p_3^{(m)} \geq \left(p_2^{(\omega+1)} + p_3^{(\omega+1)}\right)\prod_{j=1}^{q}\left(1 - \frac{1}{2^{|Q_j|}}\right) \tag{23}$$

which is also obviously valid for the special case of $q = 1$. This can be rewritten as

$$p_1^{(m)} + p_2^{(m)} \leq 1 - \left(1 - p_1^{(\omega+1)}\right)\prod_{j=1}^{q}\left(1 - \frac{1}{2^{|Q_j|}}\right) \tag{24}$$

which represents the desired upper bound on $p_1^{(m)} + p_2^{(m)}$ in terms of $p_1^{(\omega+1)}$.

6 The Conditional Block Structure below μ

6.1 Assumptions and Level $\mu + 1$

Throughout this Section 6, we will assume

$$p_3^{(\mu)} < \frac{1}{12}, \tag{25}$$

$$\prod_{j=1}^{q}\left(1 - \frac{1}{2^{|Q_j|}}\right) > \frac{4}{5} \tag{26}$$

where the product in (26) equals 1 for $q = 0$. Based on these assumption, we will further analyze the block structure below level μ in order to satisfy the m'-conditions 1–4 (see Paragraph 4.1) also for the first block level m' so that the underlying analysis can be applied recursively when inequalities (25) and (26) hold (Section 7). For this purpose, we still analyze level $\mu + 1$ in the following lemma which implies $\nu > \mu$ and thus guarantees that partition class R is defined for the underlying block if not empty.

Lemma 4 $t_{12}^{(\mu+1)} = \frac{1}{2}$.

Proof: Assumption (26) together with m-condition 3 ensures

$$p_2^{(\omega+1)} + p_3^{(\omega+1)} < \frac{5}{48} \tag{27}$$

for $\omega < m$ (implying $\omega < m-1$) according to (23). It follows from (25) that $t_{31}^{(\mu)} = 0$ which implies $t_{21}^{(\mu)} > 0$ by the definition of μ. In addition, $p_3^{(\mu-1)} < \frac{1}{6}$ due to $p_3^{(\mu)} \geq p_3^{(\mu-1)}/2$, which gives $p_1^{(\mu-1)} + p_2^{(\mu-1)} > \frac{5}{6}$. Hence,

$$p_2^{(\mu)} \geq \frac{p_1^{(\mu-1)}}{2} \geq \frac{p_1^{(\mu-1)} + p_2^{(\mu-1)}}{4} > \frac{5}{24}\,. \tag{28}$$

Consider first the case of $\omega = \mu$ ($\mu < m$ according to Lemma 2.i). If $t_{12}^{(\mu+1)} = 0$, then $p_2^{(\omega+1)} + p_3^{(\omega+1)} \geq p_2^{(\mu)} > \frac{5}{24}$ according to (28), which contradicts (27). Hence, $t_{12}^{(\mu+1)} > 0$ which implies $t_{12}^{(\mu+1)} = \frac{1}{2}$ by Lemma 2.iii for $k = \mu$.

For $\omega > \mu$, on the other hand, we know by the definition of ω that there is a double-edge path starting from $v_2^{(\mu)}$ or $v_3^{(\mu)}$ and traversing $v \in \{v_2^{(\mu+1)}, v_3^{(\mu+1)}\}$ which ends at level ω. Suppose first that $v = v_2^{(\mu+1)}$ which means either $t_{22}^{(\mu+1)} = 1$ or $t_{23}^{(\mu+1)} = 1$. In the latter case, we have $t_{32}^{(\mu+1)} = \frac{1}{2}$ which implies $t_{22}^{(\mu+1)} = \frac{1}{2}$ since $t_{12}^{(\mu+1)} = \frac{1}{2}$ would give a contradiction $\frac{1}{12} > p_3^{(\mu)} = p_2^{(\mu+1)} \geq p_3^{(\mu+1)} = p_2^{(\mu)}/2 > \frac{5}{48}$ according to (25) and (28). Thus $t_{22}^{(\mu+1)} > 0$ in both cases. For $\omega < m$, we have $t_{22}^{(\ell)} = 1$ for $\ell = \mu + 2, \ldots, \omega$, and $t_{12}^{(\omega+1)} = 0$ according to Lemma 2.iii, and hence, $p_2^{(\omega+1)} + p_3^{(\omega+1)} \geq p_2^{(\mu)}/2 > \frac{5}{48}$ according to (28), which contradicts (27). An analogous contradiction $\frac{1}{12} > p_3^{(m)} \geq p_2^{(\mu)}/2 > \frac{5}{48}$ is obtained for $\omega = m$. It follows that $v = v_3^{(\mu+1)}$ which implies $t_{33}^{(\mu+1)} = 1$ and $t_{12}^{(\mu+1)} = t_{22}^{(\mu+1)} = \frac{1}{2}$ by the normalization of P. This completes the proof that $t_{12}^{(\mu+1)} = \frac{1}{2}$. □

6.2 The Block Structure from m' to μ (m'-Conditions 1–3)

We define the first **level** $\boldsymbol{m'}$ of the underlying block to be the greatest level such that $2 \leq m' \leq \mu$ and $t_{32}^{(m')} > 0$ (m'-condition 2), which exists since at least $t_{32}^{(2)} > 0$. In the following lemma, we will analyze the initial block structure from the level m' through μ, which is illustrated in Figure 5 (where the dashed line shows that there is no edge from $v_1^{(k-1)}$ or from $v_2^{(k-1)}$ to $v_3^{(k)}$ for any $m' < k \leq \mu$).

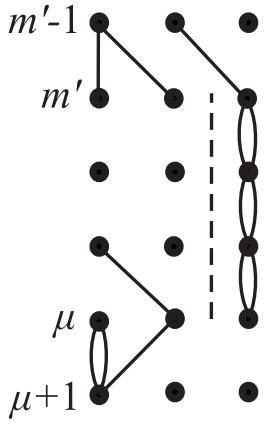

Figure 5: The block structure from m' to μ.

Lemma 5 *$t_{31}^{(k)} = t_{32}^{(k)} = 0$ and $t_{33}^{(k)} = 1$ for $k = m'+1, \ldots, \mu$.*

Proof: On the contrary, let $m' < k \leq \mu$ be the greatest level such that $t_{33}^{(k)} < 1$, that is $t_{33}^{(\ell)} = 1$ for $\ell = k+1, \ldots, \mu$. Obviously, $t_{33}^{(k)} > 0$ because $t_{32}^{(\ell)} = 0$ for every $\ell = m'+1, \ldots, k, \ldots, \mu$ by the definition of m', and $t_{31}^{(\ell)} = 0$ for every $\ell = k, \ldots, \mu$ since otherwise $p_3^{(\mu)} \geq p_3^{(\ell)} > \frac{1}{6}$, which contradicts (25). Hence, $t_{33}^{(k)} = \frac{1}{2}$ and the edge from $v_3^{(k-1)}$ to $v_3^{(k)}$ is the only edge that leads to $v_3^{(k)}$ due to $t_{31}^{(k)} = t_{32}^{(k)} = 0$, while the other edge from $v_3^{(k-1)}$ goes either to $v_1^{(k)}$ or to $v_2^{(k)}$. Thus, either $\mathbf{a}^{(m)} \in M(v_1^{(k)}) \cup M(v_2^{(k)})$ for $\mathbf{a}^{(m)}$ satisfying m-condition 4 (Paragraph 4.1), or an h-neighbor $\mathbf{a}' \in \Omega_1(\mathbf{a}^{(m)}) \cap M(v_3^{(k-1)})$ of $\mathbf{a}^{(m)}$ exists that differs from $\mathbf{a}^{(m)}$ in the variable that is tested at $v_3^{(k-1)}$ so that also $\mathbf{a}' \in M(v_1^{(k)}) \cup M(v_2^{(k)})$. Since $M(v_1^{(k)}) \cup M(v_2^{(k)}) = M(v_1^{(\mu)}) \cup M(v_2^{(\mu)})$ and $t_{12}^{(\mu+1)} = \frac{1}{2}$ by Lemma 4, there is an h-neighbor $\mathbf{a}'' \in \Omega_2(\mathbf{a}^{(m)}) \cap M(v_1^{(\mu+1)}) \subseteq H$ of $\mathbf{a}^{(m)}$ such that $P(\mathbf{a}'') = 1$ by m-condition 4 since $M(v_1^{(\mu+1)}) \subseteq M(v_1^{(m)}) \cup M(v_2^{(m)})$, which is a contradiction. Thus $t_{33}^{(k)} = 1$ for $k = m'+1, \ldots, \mu$. □

Lemma 5 together with assumption (25) gives

$$p_1^{(m')} + p_2^{(m')} = p_1^{(\mu)} + p_2^{(\mu)}, \tag{29}$$

$$p_3^{(m')} = p_3^{(\mu)} < \frac{1}{12} \tag{30}$$

which verifies m'-condition 3 for the first block level m'. Note that inequality (30) ensures $m' \geq 4$ due to $p_3^{(3)} \geq 1/2^3$. Finally, the following lemma shows m'-condition 1.

Lemma 6 *$t_{11}^{(m')} = t_{21}^{(m')} = \frac{1}{2}$ (m'-condition 1).*

Proof: Obviously, $t_{31}^{(m')} = 0$ since otherwise $p_3^{(m')} > \frac{1}{6}$, which contradicts (30). Similarly, $t_{21}^{(m')} = 1$ together with m'-condition 2 would imply $p_3^{(m')} \geq p_2^{(m'-1)}/2 \geq (p_2^{(m'-1)} +$

$p_3^{(m'-1)})/4 \geq p_1^{(m')}/4 > \frac{1}{12}$ violating (30). Finally, suppose that $t_{11}^{(m')} = 1$ which implies $t_{32}^{(m')} = \frac{1}{2}$ and $t_{33}^{(m')} < 1$ due to P is normalized. Hence, $t_{12}^{(m')} + t_{22}^{(m')} > 0$ and $t_{13}^{(m')} + t_{23}^{(m')} > 0$. Thus either $\mathbf{a}^{(m)} \in M(v_1^{(m'-1)}) \subseteq M(v_1^{(m')})$ or an h-neighbor $\mathbf{a}' \in \Omega_1(\mathbf{a}^{(m)}) \cap (M(v_2^{(m'-1)}) \cup M(v_3^{(m'-1)}))$ of $\mathbf{a}^{(m)}$ exists such that $\mathbf{a}' \in M(v_1^{(m')}) \cup M(v_2^{(m')})$. Since $M(v_1^{(m')}) \cup M(v_2^{(m')}) = M(v_1^{(\mu)}) \cup M(v_2^{(\mu)})$ and $t_{12}^{(\mu+1)} = \frac{1}{2}$ by Lemma 4, there is an h-neighbor $\mathbf{a}'' \in \Omega_2(\mathbf{a}^{(m)}) \cap M(v_1^{(\mu+1)}) \subseteq H$ of $\mathbf{a}^{(m)}$ such that $P(\mathbf{a}'') = 1$ which is a contradiction. The last possibility $t_{11}^{(m')} = t_{21}^{(m')} = \frac{1}{2}$ follows. □

6.3 An Upper Bound on $p_1^{(\omega+1)}$ in Terms of $p_1^{(m')} + p_2^{(m')}$

In Paragraph 5.3, we have upper bounded $p_1^{(m)} + p_2^{(m)}$ at the last block level in terms of $p_1^{(\omega+1)}$ provided that $\omega < m$. In this paragraph, we will extend this estimate by upper bounding $p_1^{(\omega+1)}$ (or $p_1^{(m)} + p_2^{(m)}$ for $\omega = m$) in terms of $p_1^{(m')} + p_2^{(m')}$ from the first block level. Putting these two bounds together, we will obtain a recursive formula for an upper bound on $p_1^{(m)} + p_2^{(m)}$ in terms of $p_1^{(m')} + p_2^{(m')}$ which will be used in Section 7 for verifying condition (5).

We first resolve the case of the empty block when $\nu = m = \omega$, $t_{33}^{(m)} = 0$, $t_{13}^{(m)} + t_{23}^{(m)} = 1$, and $t_{32}^{(m)} = 1$. It follows from Corollary 1 and Lemma 5 (see Figures 3 and 5, respectively) that $M(v_1^{(m')}) \cup M(v_2^{(m')}) = M(v_1^{(m)}) \cup M(v_2^{(m)})$ which ensures m'-condition 4 (m'-conditions 1–3 have already been checked in Paragraph 6.2) and $p_1^{(m')} + p_2^{(m')} = p_1^{(m)} + p_2^{(m)}$. Hence, the empty block can be skipped in our analysis by replacing m' with m, and we will further consider only the non-empty blocks.

It follows from the definition of partition class R (see Figure 3) and Lemma 4 that

$$p_1^{(\nu)} = p_1^{(\mu)} + p_2^{(\mu)} \left(1 - \frac{1}{2^{|R|}}\right) \qquad \text{for } \nu < m\,. \tag{31}$$

For $\nu = m$, we know $t_{33}^{(m)} > 0$ because we assume a non-empty block, and hence, either $t_{12}^{(m)} = t_{32}^{(m)} = \frac{1}{2}$ and $t_{33}^{(m)} = 1$, or $t_{13}^{(m)} = t_{33}^{(m)} = \frac{1}{2}$ and $t_{32}^{(m)} = 1$ by the definition of ν, Lemma 2.ii, and m-conditions 1 and 2, which also ensures $\omega = m$ in both cases. Thus,

$$p_1^{(m)} + p_2^{(m)} = p_1^{(\mu)} + p_2^{(\mu)} \left(1 - \frac{1}{2^{|R|}}\right) \qquad \text{for } \nu = m = \omega \tag{32}$$

For $\nu = m - 1$ we know $t_{12}^{(m)} = t_{13}^{(m)} = 0$ leading to $t_{32}^{(m)} = t_{33}^{(m)} = 1$, for which $\omega = m$ can be assumed without loss of generality.

Further assume $\nu < m - 1$, while the resulting formula for $\nu < m$ will also be verified for the case of $\nu = m - 1$ (when $\omega = m$) below in (34). We know by the definition of ν that $t_{12}^{(\nu+1)} = t_{13}^{(\nu+1)} = 0$, which excludes $t_{32}^{(\nu+1)} = 1$ and $t_{33}^{(\nu+1)} = 1$ since P is normalized. First consider the case of $\omega > \nu$ excluding $\omega = \nu - 1 \geq \mu$ and $\omega = \nu$ for now (cf. the definition of ω). Then the double-edge path from the definition of ω passes through a double edge from

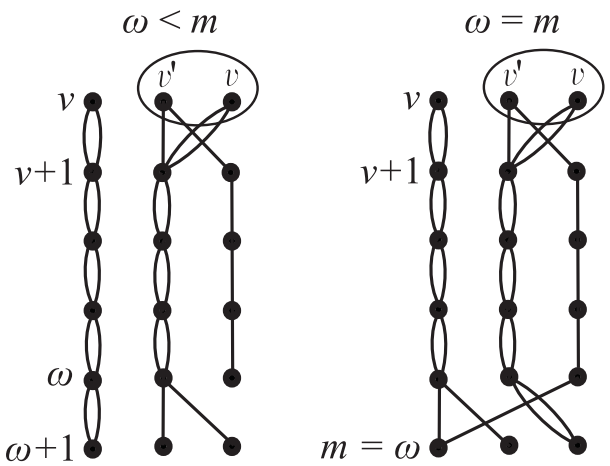

Figure 6: The block structure from $\nu < \omega$ to $\omega + 1$ (or to m if $\omega = m$).

$v \in \{v_2^{(\nu)}, v_3^{(\nu)}\}$ to $v_2^{(\nu+1)}$, while the two edges from the other node $v' \in \{v_2^{(\nu)}, v_3^{(\nu)}\} \setminus \{v\}$ lead to $v_2^{(\nu+1)}$ and to $v_3^{(\nu+1)}$, respectively, as depicted in Figure 6. For $\ell = \nu + 2, \ldots, \omega$, we have either $t_{22}^{(\ell)} = 1$ implying $t_{33}^{(\ell)} = \frac{1}{2}$ if $\ell < m$, or $t_{32}^{(\ell)} = 1$ if $\ell = m$. Moreover, $t_{12}^{(\omega+1)} = 0$ for $\omega < m$ by Corollary 1.5. Hence, $p_3^{(\nu+1)} = p_2^{(\mu)}/2^{|R|+1}$ (cf. Figure 3 and Lemma 4) upper bounds the fraction of all the inputs whose computational path traverses nodes $v', v_3^{(\nu+1)}, v_3^{(\nu+2)}, \ldots, v_3^{(\ell)}, v_1^{(\ell+1)}$ for some $\nu + 1 \le \ell \le \min(\omega, m-1)$. It follows that

$$p_1^{(\omega+1)} \le p_1^{(\nu)} + \frac{p_2^{(\mu)}}{2^{|R|+1}} \qquad \text{for } \omega < m \tag{33}$$

which is even valid for any $\max(\nu - 1, \mu) \le \omega < m$ since obviously $p_1^{(\omega+1)} = p_1^{(\nu)}$ for $\omega = \nu - 1 \ge \mu$ as well as for $\omega = \nu < m$, while

$$p_1^{(m)} + p_2^{(m)} \le p_1^{(\nu)} + \frac{p_2^{(\mu)}}{2^{|R|+1}} \qquad \text{for } \omega = m \tag{34}$$

which also holds for $\nu = m - 1$ because $p_1^{(m)} + p_2^{(m)} = p_1^{(\nu)}$ in this case.

In addition, we observe that

$$p_1^{(\mu)} + p_2^{(\mu)} \le 4p_2^{(\mu)}\,. \tag{35}$$

For $\mu > m'$, we have $p_1^{(\mu)} + p_2^{(\mu)} = p_1^{(\mu-1)} + p_2^{(\mu-1)} \le 2p_1^{(\mu-1)} \le 4p_2^{(\mu)}$ according to Lemma 5 and by $t_{21}^{(\mu)} > 0$ which follows from the definition of μ. For $\mu = m'$, on the other hand, we know $4p_2^{(\mu)} = 4p_2^{(m')} \ge 2p_1^{(m'-1)} = 2(1 - (p_2^{(m'-1)} + p_3^{(m'-1)})) \ge 2(1 - 2p_2^{(m'-1)}) \ge 2(1 - 4p_3^{(m')}) > 1 > p_1^{(\mu)} + p_2^{(\mu)}$ according to m'-conditions 1–3. This completes the proof of (35).

For $\nu < m$, equation (31) is plugged into (33) if $\omega < m$ or into (34) if $\omega = m$, while equation (32) is considered for $\nu = m$ (implying $\omega = m$). Then equations (35) and (29)

are employed, which results in

$$\begin{aligned} p_1^{(\omega+1)} &\leq p_1^{(\mu)} + p_2^{(\mu)} \left(1 - \frac{1}{2^{|R|}}\right) + \frac{p_2^{(\mu)}}{2^{|R|+1}} = p_1^{(\mu)} + p_2^{(\mu)} \left(1 - \frac{1}{2^{|R|+1}}\right) \\ &\leq \left(p_1^{(m')} + p_2^{(m')}\right) \left(1 - \frac{1}{2^{|R|+3}}\right) \qquad \text{for } \omega < m\,, \end{aligned} \tag{36}$$

$$p_1^{(m)} + p_2^{(m)} \leq \left(p_1^{(m')} + p_2^{(m')}\right) \left(1 - \frac{1}{2^{|R|+3}}\right) \qquad \text{for } \omega = m\,. \tag{37}$$

Formula (36) can further be plugged into (24) giving

$$p_1^{(m)} + p_2^{(m)} \leq 1 - \left(1 - \left(p_1^{(m')} + p_2^{(m')}\right) \left(1 - \frac{1}{2^{|R|+3}}\right)\right) \prod_{j=1}^{q} \left(1 - \frac{1}{2^{|Q_j|}}\right) \tag{38}$$

which is even valid for $\omega = m$ (i.e. $q = 0$) since equation (38) coincides with (37) in this case.

7 The Recursion

In the previous Sections 4–6, we have analyzed the structure of the block of P from level m' through m. We will now employ this block analysis recursively so that $m = m_r$ is replaced by $m' = m_{r+1}$. For this purpose, we introduce additional index $b = 1, \ldots, r$ to the underlying objects in order to differentiate among respective blocks. For example, the partition classes $R, Q_1, \ldots, Q_q$, defined in Section 5, corresponding to the bth block are denoted as $R_b, Q_{b1}, \ldots, Q_{bq_b}$, respectively.

7.1 Inductive Assumptions

In particular, we will proceed by induction on r, starting with $r = 0$ and $m_0 = d$. In the induction step for $r+1$, we assume that the four m_r-conditions from Paragraph 4.1 are met for the last block level $m = m_r$, and let the assumption (25) be satisfied for the previous blocks, that is,

$$p_3^{(\mu_b)} < \frac{1}{12} \tag{39}$$

for every $b = 1, \ldots, r$. In addition, assume

$$\varrho_r > \delta\,\varepsilon\,, \tag{40}$$

$$1 - \Pi_r < (1 - \delta)\,\varepsilon \tag{41}$$

where

$$\varrho_k = \prod_{b=1}^{k} \alpha_b\,, \qquad \alpha_b = \left(1 - \frac{1}{2^{|R_b|+3}}\right)\,, \tag{42}$$

$$\Pi_k = \prod_{b=1}^{k} \pi_b \,, \qquad \pi_b = \prod_{j=1}^{q_b} \left(1 - \frac{1}{2^{|Q_{bj}|}}\right) \tag{43}$$

for $k = 1, \ldots, r$, $\varrho_0 = \Pi_0 = 1$, and $\pi_b = 1$ for $q_b = 0$. It follows from (43) and (41) that

$$\pi_b \geq \Pi_r > 1 - (1 - \delta)\,\varepsilon \geq \delta > \frac{4}{5} \tag{44}$$

which verifies assumption (26) for every $b = 1, \ldots, r$. Hence, we can employ recursive inequality (38) from Section 6 which is rewritten as

$$p_{b-1} \leq 1 - (1 - p_b\alpha_b)\pi_b = 1 - \pi_b + p_b\alpha_b\pi_b \tag{45}$$

for $b = 1, \ldots, r$ where notation $p_b = p_1^{(m_b)} + p_2^{(m_b)}$ is introduced. Starting with

$$p_0 = p_1^{(d)} + p_2^{(d)} \geq \varepsilon \tag{46}$$

which follows from (15), recurrence (45) can be solved as

$$\varepsilon \leq \sum_{k=1}^{r}(1-\pi_k)\prod_{b=1}^{k-1}\alpha_b\pi_b + p_r\prod_{b=1}^{r}\alpha_b\pi_b < \sum_{k=1}^{r}(1-\pi_k)\Pi_{k-1} + p_r\varrho_r\Pi_r = 1 - \Pi_r + p_r\varrho_r\Pi_r \,. \tag{47}$$

7.2 Recursive Step

Throughout this paragraph, we will consider the case when

$$1 - \Pi_{r+1} < (1 - \delta)\,\varepsilon \tag{48}$$

(cf. assumption (41)), while the case complementary to (48), which concludes the induction, will be resolved below in Section 8. We will prove under assumption (48) that inductive assumptions (39)–(41) are met for r replaced with $r + 1$ together with the four m_{r+1}-conditions for the first block level m_{r+1} so that we can further proceed in the recursion.

By analogy to (44), inequality (48) implies

$$\pi_{r+1} > \delta > \frac{4}{5} \,. \tag{49}$$

For $\omega_{r+1} < m_r$, we know

$$p_r \leq 1 - \left(p_2^{(\omega_{r+1}+1)} + p_3^{(\omega_{r+1}+1)}\right)\pi_{r+1} \tag{50}$$

according to (24), and

$$p_2^{(\omega_{r+1}+1)} + p_3^{(\omega_{r+1}+1)} \geq p_3^{(\mu_{r+1})} \tag{51}$$

by the definition of ω_{r+1} and Lemma 2.iii–iv (for $k = \omega_{r+1}$), which altogether gives

$$\varepsilon < 1 - \Pi_r + \left(1 - p_3^{(\mu_{r+1})}\pi_{r+1}\right)\varrho_r\Pi_r \tag{52}$$

according to (47). Hence,

$$\delta\,\varepsilon < \left(1 - p_3^{(\mu_{r+1})}\pi_{r+1}\right)\varrho_r\Pi_r < 1 - p_3^{(\mu_{r+1})}\pi_{r+1} \tag{53}$$

follows from (41), which gives

$$p_3^{(\mu_{r+1})}\pi_{r+1} < 1 - \delta^2 \tag{54}$$

by the assumption of $\varepsilon > \delta$, implying

$$p_3^{(\mu_{r+1})} < \frac{1-\delta^2}{\delta} < \frac{1}{12} \qquad \text{for } \omega_{r+1} < m_r \tag{55}$$

due to (49). Inequality (55) is even valid for $\omega_{r+1} = m_r$ since

$$p_3^{(\mu_{r+1})} \le p_3^{(m_r)} < \frac{1}{12} \qquad \text{for } \omega_{r+1} = m_r \tag{56}$$

according to m_r-condition 3. Therefore, assumptions (25) and (26) of the analysis in Section 6 are also met for the $(r+1)$st block according to (55)–(56) and (49), respectively, which justifies recurrence inequality (45) for $b = r+1$ leading to the solution

$$\varepsilon < 1 - \Pi_{r+1} + p_{r+1}\varrho_{r+1}\Pi_{r+1} \tag{57}$$

by analogy to (47) where r is replaced with $r+1$. By combining (57) with (48), we obtain

$$\varrho_{r+1} > p_{r+1}\varrho_{r+1}\Pi_{r+1} > \delta\,\varepsilon\,. \tag{58}$$

Thus, inductive assumptions (39)–(41) are valid for r replaced by $r+1$ according to (55)–(56), (58), and (48), respectively.

In order to proceed in the next induction step, we still need to verify the four m_{r+1}-conditions from Paragraph 4.1 for m_{r+1}. In Paragraph 6.2, m_{r+1}-conditions 1–3 have been shown, and thus, it suffices to validate m_{r+1}-condition 4. For this purpose, we exploit the fact that A is $(\delta^{11} - \delta^{12})\varepsilon^{12}$-rich after we show corresponding condition (8) for partition $\{R_1, \ldots, R_{r+1}\}$ of $I = \bigcup_{b=1}^{r+1} R_b$. In particular,

$$(\delta^{11} - \delta^{12})\varepsilon^{12} < (\delta\,\varepsilon)^{11} < \prod_{b=1}^{r+1}\left(1 - \frac{1}{2^{|R_b|}}\right) \tag{59}$$

follows from (58) since for any $1 \le b \le r+1$,

$$\left(1 - \frac{1}{2^{|R_b|+3}}\right)^{11} < 1 - \frac{1}{2^{|R_b|}} \tag{60}$$

for $|R_b| \ge 1$ because $f(x) = \ln(1-\frac{1}{x})/\ln(1-\frac{1}{8x})$ is a decreasing function for $x = 2^{|R_b|} \ge 2$, and $f(2) < 11$. This provides required $\mathbf{a}^{(m_{r+1})} \in A$ such that for every $b = 1, \ldots, r+1$ there exists $i \in R_b$ that meets $a_i^{(m_{r+1})} \ne c_i$ according to (7). Obviously, the computational path for this $\mathbf{a}^{(m_{r+1})}$ ends up in sink $v_1^{(d)}$ or $v_2^{(d)}$ labeled with 1 when we put $\mathbf{a}^{(m_{r+1})}$ at node $v_1^{(m_{r+1})}$ or $v_2^{(m_{r+1})}$ by the definition of R_b, c_i and by the structure of branching program P (see Figure 3), which proves m_{r+1}-condition 4. Thus, the inductive assumptions are met for $r+1$ and we can proceed recursively for r replaced with $r+1$ etc. until condition (48) is broken.

8 The End of Recursion

In this section, we will consider the case of

$$1 - \Pi_{r+1} \geq (1-\delta)\,\varepsilon \tag{61}$$

complementary to (48), which concludes the recursion from Section 7 as follows. We will again employ the fact that A is $(\delta^{11}-\delta^{12})\varepsilon^{12}$-rich. First condition (5) for partition $\{Q_{11},\ldots,Q_{1q_1},Q_{21},\ldots,Q_{2q_2},\ldots,Q_{r+1,1},\ldots,Q_{r+1,q_{r+1}},\,R_1,\ldots,R_r\}$ of $I=\bigcup_{b=1}^{r+1}\bigcup_{j=1}^{q_b}Q_{bj}\cup\bigcup_{b=1}^{r}R_b$ is verified as

$$\begin{aligned}\left(1-\prod_{b=1}^{r+1}\prod_{j=1}^{q_b}\left(1-\frac{1}{2^{|Q_{bj}|}}\right)\right)\prod_{b=1}^{r}\left(1-\frac{1}{2^{|R_b|}}\right) &> (1-\Pi_{r+1})\varrho_r^{11}\\ &> (1-\delta)\varepsilon\,(\delta\,\varepsilon)^{11} = (\delta^{11}-\delta^{12})\varepsilon^{12}\end{aligned} \tag{62}$$

according to (60), (61), and (40). This provides $\mathbf{a}^*\in A$ such that there exists block $1\leq b^*\leq r+1$ and $1\leq j^*\leq q_{b^*}$ satisfying $a_i^*=c_i$ for every $i\in Q_{b^*j^*}$, and simultaneously, for every $b=1,\ldots,r$ there exists $i\in R_b$ that meets $a_i^*\neq c_i$ according to (6) and (7).

Lemma 7 *Denote $\lambda=\lambda_{b^*j^*}$. There are two generalized 'switching' paths (cf. Lemma 2.ii) starting from $v_2^{(k)}$ and from $v_3^{(k)}$, respectively, at some level $3<\max(\lambda-2,\mu_{b^*})\leq k<\lambda$, which may also lead to $v_3^{(\lambda)}$ in addition to $v_1^{(\lambda-1)}$ or $v_1^{(\lambda)}$.*

Proof: For the notation simplicity, we will omit the block index b^* in this proof. We know $\omega<m$ due to $q>0$, and $\lambda>\mu$ from Paragraph 5.2. Consider first the case when $t_{12}^{(\lambda)}=t_{13}^{(\lambda)}=0$. Obviously, $t_{22}^{(\lambda)}<1$ follows from the definition of λ for $\lambda>\omega$ and from the definition of ω for $\lambda=\omega$, which gives $t_{22}^{(\lambda)}=t_{32}^{(\lambda)}=\frac{1}{2}$ and $t_{23}^{(\lambda)}>0$ by the normalization of P. For $t_{33}^{(\lambda)}=\frac{1}{2}$, we obtain two switching paths $v_2^{(\lambda-1)},v_3^{(\lambda)}$ and $v_3^{(\lambda-1)},v_3^{(\lambda)}$. Thus assume $t_{33}^{(\lambda)}=0$ which ensures $t_{23}^{(\lambda)}=1$ and $\lambda>\mu+1$ since $\lambda=\mu+1$ would give $\omega>\lambda$. Consider first the case when $t_{12}^{(\lambda-1)}=t_{13}^{(\lambda-1)}=0$, which implies $t_{22}^{(\lambda-1)}>0$ and $t_{23}^{(\lambda-1)}>0$ by $t_{11}^{(\lambda-1)}=1$ and the normalization of P, providing two switching paths $v_2^{(\lambda-2)},v_2^{(\lambda-1)},v_3^{(\lambda)}$ and $v_3^{(\lambda-2)},v_2^{(\lambda-1)},v_3^{(\lambda)}$. Two switching paths $v_2^{(\lambda-2)},v_1^{(\lambda-1)}$ and $v_3^{(\lambda-2)},v_1^{(\lambda-1)}$ are also guaranteed when $t_{12}^{(\lambda-1)}>0$ and $t_{13}^{(\lambda-1)}>0$ appear simultaneously. For $t_{12}^{(\lambda-1)}=0$ and $t_{13}^{(\lambda-1)}>0$, we have $t_{22}^{(\lambda-1)}>0$ by the normalization of P, which together with $t_{32}^{(\lambda)}=\frac{1}{2}$ produces two switching paths $v_2^{(\lambda-2)},v_2^{(\lambda-1)},v_3^{(\lambda)}$ and $v_3^{(\lambda-2)},v_1^{(\lambda-1)}$. For $t_{12}^{(\lambda-1)}>0$ and $t_{13}^{(\lambda-1)}=0$, the case of $t_{23}^{(\lambda-1)}>0$ ensures two switching paths $v_2^{(\lambda-2)},v_1^{(\lambda-1)}$ and $v_3^{(\lambda-2)},v_2^{(\lambda-1)},v_3^{(\lambda)}$, while for $t_{23}^{(\lambda-1)}=0$ we obtain $t_{12}^{(\lambda-1)}=t_{22}^{(\lambda-1)}=\frac{1}{2}$ and $t_{33}^{(\lambda-1)}=1$, which implies $\lambda=\nu+1$ and $\omega>\lambda$ by Lemma 2.iii contradicting the definition of $\lambda\geq\omega\geq\nu-1$. This completes the argument for $t_{12}^{(\lambda)}=t_{13}^{(\lambda)}=0$.

The case of $t_{13}^{(\lambda)}>0$ and $t_{12}^{(\lambda)}>0$ produces two switching paths $v_2^{(\lambda-1)},v_1^{(\lambda)}$ and $v_3^{(\lambda-1)},v_1^{(\lambda)}$. Further consider the case when $t_{13}^{(\lambda)}>0$ and $t_{12}^{(\lambda)}=0$. Obviously, $t_{22}^{(\lambda)}<1$

follows from the definition of λ for $\lambda > \omega$ and from the definition of ω for $\lambda = \omega$. Hence, $t_{32}^{(\lambda)} > 0$ which provides two switching paths $v_2^{(\lambda-1)}, v_3^{(\lambda)}$ and $v_3^{(\lambda-1)}, v_1^{(\lambda)}$. Finally, consider the case when $t_{12}^{(\lambda)} > 0$ and $t_{13}^{(\lambda)} = 0$, for which $t_{33}^{(\lambda)} > 0$ generates two switching $v_2^{(\lambda-1)}, v_1^{(\lambda)}$ and $v_3^{(\lambda-1)}, v_3^{(\lambda)}$, while for $t_{33}^{(\lambda)} = 0$ we obtain $t_{32}^{(\lambda)} = \frac{1}{2}$ and $t_{23}^{(\lambda)} = 1$, which implies $\lambda = \nu$ and $\omega > \lambda$ by Lemma 2.iii contradicting the definition of $\lambda \geq \omega \geq \nu - 1$. □

By a similar argument to Lemma 2.ii, Lemma 7 gives an h-neighbor $\mathbf{a}' \in \Omega_2(\mathbf{a}^*) \subseteq H$ of input $\mathbf{a}^* \in A$ such that $\mathbf{a}' \in M(v_1^{(\lambda)}) \cup M(v_3^{(\lambda)})$. Thus, either $\mathbf{a}'' = \mathbf{a}' \in M(v_1^{(\lambda)}) \subseteq M(v_1^{(m_{b^*-1})}) \cup M(v_2^{(m_{b^*-1})})$, or $\mathbf{a}' \in M(v_3^{(\lambda)})$ which implies $\mathbf{a}' \in M(v_3^{(\kappa_{b^*j^*}-1)})$ since $a'_i = a_i^* = c_i$ for every $i \in Q_{b^*j^*}$ according to (6) (see Figure 4), and an h-neighbor $\mathbf{a}'' \in \Omega_3(\mathbf{a}^*)$ of $\mathbf{a}^*$ exists such that $\mathbf{a}'' \in M(v_1^{(\kappa_{b^*j^*})}) \subseteq M(v_1^{(m_{b^*-1})}) \cup M(v_2^{(m_{b^*-1})})$. Hence, $P(\mathbf{a}'') = 1$ because for every $b = 1, \ldots, b^* - 1 \leq r$ there exists $i \in R_b$ that meets $a'_i = a_i^* \neq c_i$ by condition (7) (see Figure 3). This completes the proof of Theorem 2. □

9 The Richness of Almost k-wise Independent Sets

In order to achieve an explicit polynomial time construction of a hitting set for read-once branching programs of width 3 we will combine Theorem 2 with the result due to Alon et al. [1] who provided simple efficient constructions of almost k-wise independent sets. In particular, for $\beta > 0$ and $k = O(\log n)$ it is possible to construct a (k, β)-*wise independent set* $\mathcal{A} \subseteq \{0,1\}^*$ in time polynomial in $\frac{n}{\beta}$ such that for sufficiently large n and any index set $S \subseteq \{1, \ldots, n\}$ of size $|S| \leq k$, the probability that a given $\mathbf{c} \in \{0,1\}^n$ coincides with a string $\mathbf{a} \in \mathcal{A}_n = \mathcal{A} \cap \{0,1\}^n$ on the bit locations from S is almost uniform, that is

$$\left| \frac{\left|\mathcal{A}_n^S(\mathbf{c})\right|}{|\mathcal{A}_n|} - \frac{1}{2^{|S|}} \right| \leq \beta \tag{63}$$

where $\mathcal{A}_n^S(\mathbf{c}) = \{\mathbf{a} \in \mathcal{A}_n \,|\, (\forall i \in S)\, a_i = c_i\}$. We will prove that any almost k-wise independent set is ε-rich for suitable k.

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.*

Proof: Let $\mathcal{A} \subseteq \{0,1\}^*$ be a $(\lceil (C+2) \log n \rceil, \beta)$-wise independent set. We will show that $\mathcal{A}$ is ε-rich. Given a partition $\{Q_1, \ldots, Q_q, R_1, \ldots, R_r\}$ of index set $I \subseteq \{1, \ldots, n\}$ satisfying condition (5) we will first properly select and modify the partition classes in order to upper bound their cardinalities by the logarithmic function so that the assumption concerning the almost $\lceil (C+2) \log n \rceil$-wise independence of $\mathcal{A}$ can be applied. Thus, observe there must be $1 \leq \ell \leq q$ such that $|Q_\ell| \leq \log n$, and denote $Q = Q_\ell$ for this ℓ. It is because, if $|Q_j| > \log n$ for every $j = 1, \ldots, q$, then we would have

$$\prod_{j=1}^{q} \left(1 - \frac{1}{2^{|Q_j|}}\right) \geq \left(1 - \frac{1}{2^{\log n}}\right)^{\frac{n}{\log n}} > 1 - \frac{1}{n} \cdot \frac{n}{\log n} = 1 - \frac{1}{\log n} \tag{64}$$

which breaks (5) for $n > 2^{1/\varepsilon}$. Furthermore, we confine ourselves to at most logarithmic-size subsets R'_j of partition classes R_j, that is

$$R'_j \begin{cases} = R_j & \text{if } |R_j| \le \log n \\ \subset R_j \text{ so that } |R'_j| = \lfloor \log n \rfloor & \text{otherwise} \end{cases} \tag{65}$$

which ensures $R'_j \subseteq R_j$ and $|R'_j| \le \log n$ for every $j = 1, \ldots, r$. For these new classes, assumption (5) can be rewritten as

$$\prod_{j=1}^{r} \left(1 - \frac{1}{2^{|R'_j|}}\right) > \left(1 - \frac{1}{2^{\log n}}\right)^{\frac{n}{\log n}} \prod_{|R_j| \le \log n} \left(1 - \frac{1}{2^{|R_j|}}\right) > \left(1 - \frac{1}{\log n}\right) \varepsilon = \varepsilon' \tag{66}$$

where $\varepsilon' > 0$ is arbitrarily close to ε for sufficiently large n.

Without loss of generality, assume $\log n \ge |R'_1| \ge |R'_2| \ge \cdots \ge |R'_r| \ge 1$. Denote by $\{s_1 > s_2 > \cdots > s_m\} = \{|R'_1|, \ldots, |R'_r|\}$ the set of all cardinalities $1 \le s_i \le \log n$ of classes $R'_1, \ldots, R'_r$, and let $r_i = |\{j \,|\, |R'_j| = s_i\}|$ be the number of classes R'_j of cardinality s_i for $i = 1, \ldots, m$, that is, $r = \sum_{i=1}^{m} r_i$. Moreover, we define

$$t_i = \frac{r_i}{2^{s_i}} > 0 \quad \text{for } i = 1, \ldots, m\,. \tag{67}$$

It follows from (66) that

$$0 < \varepsilon' < \prod_{j=1}^{r} \left(1 - \frac{1}{2^{|R'_j|}}\right) = \prod_{i=1}^{m} \left(1 - \frac{1}{2^{s_i}}\right)^{r_i} = \prod_{i=1}^{m} \left(\left(1 - \frac{1}{2^{s_i}}\right)^{2^{s_i}}\right)^{t_i} < e^{-\sum_{i=1}^{m} t_i} \tag{68}$$

implying

$$\sum_{i=1}^{m} t_i < \ln \frac{1}{\varepsilon'}\,. \tag{69}$$

In addition, we will further confine ourselves to the first $m' \ge 0$ cardinalities s_i satisfying

$$s_i > \log\left(\frac{2}{\varepsilon'^2} - 2\right) \quad \text{for } i = 1, \ldots, m' \tag{70}$$

whereas $s_i \le \log(2/\varepsilon'^2 - 2)$ for $i = m' + 1, \ldots, m$. Let r' be the corresponding maximal index j of class R'_j such that $|R'_{r'}| = s_{m'}$, that is,

$$r' = \sum_{i=1}^{m'} r_i = \sum_{i=1}^{m'} t_i 2^{s_i} > \left(\frac{2}{\varepsilon'^2} - 2\right) \sum_{i=1}^{m'} t_i \tag{71}$$

according to (67) and (70). We include the remaining constant-size classes R'_j for $j = r' + 1, \ldots, r$ into Q, that is,

$$Q' = Q \cup \bigcup_{j=r'+1}^{r} R'_j \tag{72}$$

whose size can be upper bounded as

$$|Q'| \leq \log n + \sum_{i=m'+1}^{m} r_i \log\left(\frac{2}{\varepsilon'^2} - 2\right) < 2\log n \tag{73}$$

for sufficiently large n, since

$$\sum_{i=m'+1}^{m} r_i = \sum_{i=m'+1}^{m} t_i 2^{s_i} < \left(\frac{2}{\varepsilon'^2} - 2\right) \ln \frac{1}{\varepsilon'} \tag{74}$$

according to (67) and (69). This completes the definition of new classes $Q', R'_1, \ldots, R'_{r'}$ which will be used for the argument below.

Now we turn to the proof of the proposition. In order to show for a given $\mathbf{c} \in \{0,1\}^n$ that there is $\mathbf{a} \in \mathcal{A}_n$ that meets the conjunction (6) and (7) for the underlying partition $\{Q_1, \ldots, Q_q, R_1, \ldots, R_r\}$ we will prove that the probability

$$\mathcal{P} = \frac{\left|\mathcal{A}_n^Q(\mathbf{c}) \setminus \bigcup_{j=1}^{r} \mathcal{A}_n^{R_j}(\mathbf{c})\right|}{|\mathcal{A}_n|} \tag{75}$$

of the event that $\mathbf{a} \in \mathcal{A}_n$ chosen uniformly at random satisfies $\mathbf{a} \in \mathcal{A}_n^Q(\mathbf{c})$ and $\mathbf{a} \notin \mathcal{A}_n^{R_j}(\mathbf{c})$ for every $j = 1, \ldots, r$, is strictly positive. This probability can be lower bounded by using new classes $Q', R'_1, \ldots, R'_{r'}$ as follows. First we define $\mathbf{c}' \in \{0,1\}^n$ that differs from $\mathbf{c}$ exactly on the constant number of bit locations from $R'_{r'+1}, \ldots, R'_r$, e.g.

$$c'_i = \begin{cases} 1 - c_i & \text{if } i \in \bigcup_{j=r'+1}^{r} R'_j \\ c_i & \text{otherwise,} \end{cases} \tag{76}$$

and observe that $\mathcal{A}_n^{Q'}(\mathbf{c}') \setminus \bigcup_{j=1}^{r'} \mathcal{A}_n^{R'_j}(\mathbf{c}') \subseteq \mathcal{A}_n^Q(\mathbf{c}) \setminus \bigcup_{j=1}^{r} \mathcal{A}_n^{R_j}(\mathbf{c})$. Let $\mathbf{a} \in \mathcal{A}_n^{Q'}(\mathbf{c}') \setminus \bigcup_{j=1}^{r'} \mathcal{A}_n^{R'_j}(\mathbf{c}')$ which means $\mathbf{a} \in \mathcal{A}_n^{Q'}(\mathbf{c}') \subseteq \mathcal{A}_n^Q(\mathbf{c}') = \mathcal{A}_n^Q(\mathbf{c})$ and $\mathbf{a} \notin \mathcal{A}_n^{R'_j}(\mathbf{c}') = \mathcal{A}_n^{R'_j}(\mathbf{c}) \supseteq \mathcal{A}_n^{R_j}(\mathbf{c})$ for every $j = 1, \ldots, r'$ by definitions (65), (72), and the fact that $S_1 \subseteq S_2$ implies $\mathcal{A}_n^{S_2}(\mathbf{c}) \subseteq \mathcal{A}_n^{S_1}(\mathbf{c})$. In addition, $\mathbf{a} \in \mathcal{A}_n^{Q'}(\mathbf{c}')$ implies $\mathbf{a} \notin \mathcal{A}_n^{R_j}(\mathbf{c})$ for every $j = r'+1, \ldots, r$ according to (76), and hence, $\mathbf{a} \in \mathcal{A}_n^Q(\mathbf{c}) \setminus \bigcup_{j=1}^{r} \mathcal{A}_n^{R_j}(\mathbf{c})$. It follows that

$$\mathcal{P} \geq \frac{\left|\mathcal{A}_n^{Q'}(\mathbf{c}') \setminus \bigcup_{j=1}^{r'} \mathcal{A}_n^{R'_j}(\mathbf{c}')\right|}{|\mathcal{A}_n|} = \frac{\left|\mathcal{A}_n^{Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|} - \frac{\left|\bigcup_{j=1}^{r'} \mathcal{A}_n^{R'_j \cup Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|}\,. \tag{77}$$

Furthermore, we will upper bound the probability of the finite union of events appearing in formula (77) by using Bonferroni inequality for constant number $C' = \min(C, r')$ of terms, which gives

$$\mathcal{P} \geq \frac{\left|\mathcal{A}_n^{Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|} - \sum_{k=1}^{C'} (-1)^{k+1} \sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \frac{\left|\bigcap_{i=1}^{k} \mathcal{A}_n^{R'_{j_i} \cup Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|} \tag{78}$$

$$= \sum_{k=0}^{C'} (-1)^{k} \sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \frac{\left|\mathcal{A}_n^{\bigcup_{i=1}^{k} R'_{j_i} \cup Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|}\,. \tag{79}$$

Note that C' is odd for $C < r'$ while equality holds in (78) for $C' = r'$, which is the probabilistic inclusion-exclusion principle. For any $0 \leq k \leq C' \leq C$, we know $\left|\bigcup_{i=1}^{k} R'_{j_i} \cup Q'\right| \leq \lceil (C+2)\log n \rceil$ according to (65) and (73), and hence,

$$\frac{\left|\mathcal{A}_n^{\bigcup_{i=1}^{k} R'_{j_i} \cup Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|} \geq \frac{1}{2^{|Q'|+\sum_{i=1}^{k}|R'_{j_i}|}} - \beta = \frac{1}{2^{|Q'|}} \prod_{i=1}^{k} \frac{1}{2^{|R'_{j_i}|}} - \beta \tag{80}$$

and similarly,

$$-\frac{\left|\mathcal{A}_n^{\bigcup_{i=1}^{k} R'_{j_i} \cup Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|} \geq -\frac{1}{2^{|Q'|}} \prod_{i=1}^{k} \frac{1}{2^{|R'_{j_i}|}} - \beta \tag{81}$$

according to (63) since $\mathcal{A}$ is $(\lceil (C+2)\log n \rceil, \beta)$-wise independent. We plug these inequalities into (79), which leads to

$$\begin{aligned} \mathcal{P} &\geq \sum_{k=0}^{C'} (-1)^k \sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \frac{1}{2^{|Q'|}} \prod_{i=1}^{k} \frac{1}{2^{|R'_{j_i}|}} - \beta \sum_{k=0}^{C'} \binom{r'}{k} \\ &\geq \frac{1}{2^{|Q'|}} \left(\sum_{k=0}^{C'} (-1)^k \sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \prod_{i=1}^{k} \frac{1}{2^{|R'_{j_i}|}} - \beta\, 2^{|Q'|}\, (r'+1)^{C'} \right) \end{aligned} \tag{82}$$

where

$$\beta\, 2^{|Q'|}\, (r'+1)^{C'} < \frac{1}{n^{C+3}}\, n^2\, n^C = \frac{1}{n} < \frac{\varepsilon'}{4\, C^C} \tag{83}$$

for sufficiently large $n > 4\, C^C/\varepsilon'$ by using the assumption on β, inequality (73), $r' < n$ (e.g., $r' = n$ would break (70)), and $C' \leq C$, which implies

$$\mathcal{P} \geq \frac{1}{n^2} \left(\sum_{k=0}^{C'} (-1)^k \sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \prod_{i=1}^{k} \frac{1}{2^{|R'_{j_i}|}} - \frac{\varepsilon'}{4\, C^C} \right). \tag{84}$$

By grouping the classes of the same cardinality together, the inner sum for $0 \leq k \leq C'$ at the right-hand side of inequality (84) can further be rewritten as

$$\sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \prod_{i=1}^{k} \frac{1}{2^{|R'_{j_i}|}} = \sum_{\substack{k_1 + k_2 + \cdots + k_{m'} = k \\ 0 \leq k_1 \leq r_1, \ldots, 0 \leq k_{m'} \leq r_{m'}}} \prod_{i=1}^{m'} \binom{r_i}{k_i} \left(\frac{1}{2^{s_i}}\right)^{k_i} \tag{85}$$

where $k_1, \ldots, k_{m'}$ denote the numbers of classes of corresponding cardinalities $s_1, \ldots, s_{m'}$ considered in a current summand, and

$$\binom{r_i}{k_i} \left(\frac{1}{2^{s_i}}\right)^{k_i} = \frac{r_i\, (r_i - 1) \cdots (r_i - k_i + 1)}{k_i!} \left(\frac{t_i}{r_i}\right)^{k_i} = \frac{t_i^{k_i}}{k_i!} \prod_{j=1}^{k_i - 1} \left(1 - \frac{j}{r_i}\right) \tag{86}$$

according to (67). Since $k_i \leq r_i$ and $k_i \leq k \leq C' \leq C$, we have

$$1 \geq \prod_{j=1}^{k_i-1} \left(1 - \frac{j}{r_i}\right) \geq \left(1 - \frac{\min(r_i, C') - 1}{r_i}\right)^{k_i-1} \geq \frac{1}{{C'}^{k_i-1}} > \frac{1}{C^{k_i}} \tag{87}$$

for $k_i > 0$, which implies

$$\prod_{i=1}^{m'} \binom{r_i}{k_i} \left(\frac{1}{2^{s_i}}\right)^{k_i} \geq \frac{1}{C^k} \prod_{i=1}^{m'} \frac{t_i^{k_i}}{k_i!} \geq \frac{1}{C^C} \prod_{i=1}^{m'} \frac{t_i^{k_i}}{k_i!} \tag{88}$$

according to (86). Hence,

$$\sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \prod_{i=1}^{k} \frac{1}{2^{|R'_{j_i}|}} \geq \frac{1}{C^C} \sum_{\substack{k_1 + \cdots + k_{m'} = k \\ 0 \leq k_1 \leq r_1, \ldots, 0 \leq k_{m'} \leq r_{m'}}} \prod_{i=1}^{m'} \frac{t_i^{k_i}}{k_i!} \tag{89}$$

follows from (85). We plug this inequality into (84) and obtain

$$\mathcal{P} \geq \frac{1}{C^C n^2} \left(\sum_{k=0}^{C'} (-1)^k \sum_{\substack{k_1 + \cdots + k_{m'} = k \\ 0 \leq k_1 \leq r_1, \ldots, 0 \leq k_{m'} \leq r_{m'}}} \prod_{i=1}^{m'} \frac{t_i^{k_i}}{k_i!} - \frac{\varepsilon'}{4} \right). \tag{90}$$

In order to apply the multinomial theorem, we remove the upper bounds that are set on indices $k_1 \leq r_1, \ldots, k_{m'} \leq r_{m'}$ in the inner sum of formula (90), that is,

$$\mathcal{P} \geq \frac{1}{C^C n^2} \left(\sum_{k=0}^{C'} (-1)^k \sum_{\substack{k_1 + \cdots + k_{m'} = k \\ k_1 \geq 0, \ldots, k_{m'} \geq 0}} \prod_{i=1}^{m'} \frac{t_i^{k_i}}{k_i!} - T - \frac{\varepsilon'}{4} \right), \tag{91}$$

which is corrected by introducing additional term

$$T = \sum_{k=0}^{C'} (-1)^k \sum_{\substack{k_1 + \cdots + k_{m'} = k \\ k_1 \geq 0, \ldots, k_{m'} \geq 0 \\ (\exists 1 \leq \ell \leq m')\, k_\ell > r_\ell}} \prod_{i=1}^{m'} \frac{t_i^{k_i}}{k_i!}. \tag{92}$$

We will show that this term T can be upper bounded by $\varepsilon'/4$. For this purpose, we take only the summands for even $k \geq 2$ into account since the summands for odd k are not positive while for $k = 0$ there is no $1 \leq \ell \leq m'$ such that $0 = k \geq k_\ell > r_\ell \geq 1$, which gives

$$\begin{aligned} T &\leq \sum_{k=2,4,\ldots}^{C'} \sum_{\substack{k_1 + \cdots + k_{m'} = k \\ k_1 \geq 0, \ldots, k_{m'} \geq 0 \\ (\exists 1 \leq \ell \leq m')\, k_\ell > r_\ell}} \frac{1}{2^{s_\ell}} \frac{r_\ell}{k_\ell} \frac{t_\ell^{k_\ell - 1}}{(k_\ell - 1)!} \prod_{\substack{i=1 \\ i \neq \ell}}^{m'} \frac{t_i^{k_i}}{k_i!} \\ &< \frac{\varepsilon'^2}{2\,(1 - \varepsilon'^2)} \sum_{k=2,4,\ldots}^{C'} \sum_{\substack{k_1 + \cdots + k_{m'} = k \\ k_1 \geq 0, \ldots, k_{m'} \geq 0 \\ (\exists 1 \leq \ell \leq m')\, k_\ell > r_\ell}} \frac{t_\ell^{k_\ell - 1}}{(k_\ell - 1)!} \prod_{\substack{i=1 \\ i \neq \ell}}^{m'} \frac{t_i^{k_i}}{k_i!} \end{aligned} \tag{93}$$

using (67) and (70). Formula (93) is rewritten by replacing indices $k_\ell - 1$ and $k - 1$ with k_ℓ and k, respectively, which is further upper bounded by omitting the condition on parameters $k_1, \ldots, k_{m'}$ of the inner sum concerning the existence of special index ℓ, as follows:

$$T \leq \frac{\varepsilon'^2}{2\left(1-\varepsilon'^2\right)} \sum_{k=1,3,\ldots}^{C'-1} \sum_{\substack{k_1+\cdots+k_{m'}=k \\ k_1\geq 0,\ldots,k_{m'}\geq 0}} \prod_{i=1}^{m'} \frac{t_i^{k_i}}{k_i!} = \frac{\varepsilon'^2}{2\left(1-\varepsilon'^2\right)} \sum_{k=1,3,\ldots}^{C'-1} \frac{\left(\sum_{i=1}^{m'} t_i\right)^k}{k!} \tag{94}$$

where the multinomial theorem is employed. Notice the sum on the right-hand side of equation (94) represents the first few terms of Taylor series of the hyperbolic sine at point $\sum_{i=1}^{m'} t_i$, which implies

$$T \leq \frac{\varepsilon'^2}{2\left(1-\varepsilon'^2\right)} \sinh\left(\sum_{i=1}^{m'} t_i\right) < \frac{\varepsilon'^2}{2\left(1-\varepsilon'^2\right)} \cdot \frac{\frac{1}{\varepsilon'} - \varepsilon'}{2} = \frac{\varepsilon'}{4} \tag{95}$$

according to (69) since the hyperbolic sine is an increasing function.

Now, the upper bound (95) on T is plugged into (91) and the multinomial theorem gives

$$\begin{aligned} \mathcal{P} &\geq \frac{1}{C^C n^2}\left(\sum_{k=0}^{C'} \frac{\left(-\sum_{i=1}^{m'} t_i\right)^k}{k!} - \frac{\varepsilon'}{2}\right) \\ &= \frac{1}{C^C n^2}\left(e^{-\sum_{i=1}^{m'} t_i} - \mathcal{R}_{C'+1}\left(-\sum_{i=1}^{m'} t_i\right) - \frac{\varepsilon'}{2}\right) \end{aligned} \tag{96}$$

where Taylor's theorem is employed for the exponential function at point $(-\sum_{i=1}^{m'} t_i)$ producing the Lagrange remainder

$$\mathcal{R}_{C'+1}\left(-\sum_{i=1}^{m'} t_i\right) = \frac{\left(-\sum_{i=1}^{m'} t_i\right)^{C'+1}}{(C'+1)!} e^{-\vartheta \sum_{i=1}^{m'} t_i} < \left(\frac{\sum_{i=1}^{m'} t_i}{\sqrt{C'}}\right)^{C'+1} \tag{97}$$

with parameter $0 < \vartheta < 1$. It follows that for $C' = C$, this remainder can be upper bounded as

$$\mathcal{R}_{C'+1}\left(-\sum_{i=1}^{m'} t_i\right) < \left(\frac{\ln \frac{1}{\varepsilon'}}{\sqrt{C}}\right)^{C+1} < \left(\frac{\varepsilon'}{2}\right)^{C+1} \leq \frac{\varepsilon'}{4} \tag{98}$$

for sufficiently large n by using (69) and the definition of C, while for $C' = r' < C$, the underlying upper bound

$$\mathcal{R}_{C'+1}\left(-\sum_{i=1}^{m'} t_i\right) < \left(\frac{\sum_{i=1}^{m'} t_i}{\frac{2}{\varepsilon'^2} - 2}\right)^{\frac{r'+1}{2}} < \frac{\ln \frac{1}{\varepsilon'}}{\frac{2}{\varepsilon'^2} - 2} \leq \frac{\varepsilon'}{4} \tag{99}$$

can be obtained from (71) and (69). Finally, inequality (68) together with the upper bound (98) or (99) on the Lagrange remainder is plugged into (96) which leads to

$$\mathcal{P} > \frac{\varepsilon'}{4\,C^C n^2} = \frac{\varepsilon\left(1-\frac{1}{\log n}\right)}{4\,C^C n^2} > 0 \tag{100}$$

according to (66). Thus, we have proven that for any $\mathbf{c} \in \{0,1\}^n$ the probability that there is $\mathbf{a} \in \mathcal{A}_n$ satisfying the conjunction (6) and (7) for partition $\{Q_1, \ldots, Q_q, R_1, \ldots, R_r\}$ is strictly positive, which means such $\mathbf{a}$ does exist. This completes the proof that $\mathcal{A}$ is ε-rich. □

10 Conclusion

In the present paper, we have made an important step in the effort of constructing hitting set generators for the model of read-once branching programs of bounded width. Such constructions have so far been known only in the case of width 2 and in very restricted cases of bounded width (e.g. permutation or regular oblivious read-once branching programs). We have now provided an explicit polynomial-time construction of a hitting set for (weakly oblivious) read-once branching programs of width 3 with acceptance probability $\varepsilon > \sqrt{12/13}$. Although this model seems to be relatively weak, the presented proof is far from being trivial. In particular, we have characterized the hitting sets for read-once width-3 branching programs by an elegant richness condition which is independent of the notion of branching programs. In addition, we have proven for a suitable constant C that any almost $(C \log n)$-wise independent set which can be constructed in polynomial time due to Alon et al. [1], satisfies this richness condition, which implies our result.

From the point of view of derandomization of unrestricted models, our result still appears to be unsatisfactory but it is the best we know so far. The issue of whether our technique based on the richness condition can be extended to the case of width 4 or to bounded width represents an open problem for further research. Another challenge for improving our result is to optimize parameter ε, e.g. to achieve the result for $\varepsilon \leq \frac{1}{n}$, which would be important for practical derandomizations.

References

[1] Alon, N., Goldreich, O., Håstad, J., and Peralta, R.: Simple constructions of almost k-wise independent random variables. Random Structures and Algorithms **3** (3) (1992) 289–304

[2] Bogdanov, A., Dvir, Z., Verbin, E., Yehudayoff, A.: Pseudorandomness for width 2 branching programs. Electronic Colloquium on Computational Complexity, Report No. **70** (2009)

[3] Braverman, M., Rao, A., Raz, R., Yehudayoff, A.: Pseudorandom generators for regular branching programs. Proceedings of the FOCS 2010 Fifty-First Annual IEEE Symposium on Foundations of Computer Science, (2010) 41–50

[4] Brody, J., Verbin, E.: The coin problem, and pseudorandomness for branching programs. Proceedings of the FOCS 2010 Fifty-First Annual IEEE Symposium on Foundations of Computer Science, (2010) 30–39

[5] Goldreich, O., Wigderson, A.: Improved derandomization of BPP using a hitting set generator. Proceedings of the RANDOM'99 Third International Workshop on Randomization and Approximation Techniques in Computer Science, LNCS **1671**, Springer-Verlag, Berlin (1999) 131–137

[6] Koucký, M., Nimbhorkar, P., Pudlák, P: Pseudorandom generators for group products. Electronic Colloquium on Computational Complexity, Report No. **133** (2010)

[7] Meka, R., Zuckerman, D.: Pseudorandom generators for polynomial threshold functions. Proceedings of the STOC 2010 Forty-Second ACM Symposium on Theory of Computing, ACM, New York, NY (2010) 427–436

[8] Nisan, N.: Pseudorandom generators for space-bounded computation. Combinatorica **12** (4) (1992) 449–461

[9] Nisan, N., Wigderson, A.: Hardness vs. randomness. Journal of Computer and System Sciences **49** (2) (1994) 149–167

[10] Šíma, J., Žák, S.: A polynomial time constructible hitting set for restricted 1-branching programs of width 3. Proceedings of the SOFSEM 2007 Thirty-Third International Conference on Current Trends in Theory and Practice of Informatics, LNCS **4362**, Springer-Verlag, Berlin (2007) 522–531

[11] Wegener, I.: Branching Programs and Binary Decision Diagrams—Theory and Applications. SIAM Monographs on Discrete Mathematics and Its Applications, SIAM, Philadelphia, PA (2000)

ECCC ISSN 1433-8092

http://eccc.hpi-web.de