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

Jiří Šíma[a,*], Stanislav Žák[a]

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

Abstract

Recently, an interest in constructing pseudorandom or hitting set generators for restricted branching programs has increased, which is motivated by the fundamental issue of derandomizing space-bounded computations. Such constructions have been known only in the case of width 2 and in very restricted cases of bounded width. In this paper, we characterize the hitting sets for read-once branching programs of width 3 by a so-called richness condition. In particular, we show that such sets hit the class of read-once conjunctions of DNF and CNF (i.e. the weak richness). Moreover, we prove that any rich set extended with all strings within Hamming distance of 3 is a hitting set for read-once branching programs of width 3. Then, we show 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 read-once branching programs of width 3 with acceptance probability $\varepsilon > 5/6$.

Keywords: Derandomization, Hitting set, Read-once branching program, Bounded width

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 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 [9], which is a weaker (one-sided error) version of pseudorandom generator [13]. Recall that an *ε-pseudorandom generator* for a class of Boolean functions of n variables is a function $\mathbf{g} : \{0,1\}^s \longrightarrow \{0,1\}^n$ which stretches a short uniformly random *seed* of *length* s bits into n bits ($s \ll n$) that cannot be distinguished from uniform ones. In particular, for every function

*Corresponding author

Email addresses: sima@cs.cas.cz (Jiří Šíma), stan@cs.cas.cz (Stanislav Žák)

f in the class, condition $|Pr_{\mathbf{x}\sim U_n}[f(\mathbf{x})=1] - Pr_{\mathbf{y}\sim U_s}[f(\mathbf{g}(\mathbf{y}))=1]| \leq \varepsilon$ holds where $\mathbf{x} \sim U_n$ means that $\mathbf{x}$ is uniformly distributed in $\{0,1\}^n$.

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 where 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 [17] for which pseudorandom generators with seed length $O(\log^2 n)$ have been known for a long time through a result of Nisan [12]. Note that an explicit pseudorandom generator for this model which is computable in logarithmic space and has seed length $O(\log n)$ would suffice to derandomize the complexity class BPL (Bounded-error Probabilistic Logarithmic-space). Recently, considerable attention has been paid to improving the seed length to $O(\log n)$ in the constant-width case, which is a fundamental problem with many applications in circuit lower bounds and derandomization [11]. The problem has been resolved for width 2 but the known techniques provably fail for width 3 [3, 5, 6, 8, 11], which applies even to hitting set generators [5].

In the case of width 3, we do not know of any significant improvement over Nisan's result except for some recent progress in the severely restricted case of so-called regular oblivious read-once branching programs. Recall that an *oblivious* branching program queries the input variables in a fixed order, which represents a provably weaker computational model [2]. For constant-width *regular* oblivious 1-branching programs which have the in-degree of all nodes equal to 2 (or 0), three independent constructions of ε-pseudorandom generator with seed length $O(\log n(\log\log n + \log(1/\varepsilon)))$ were achieved [4, 5, 6]. This seed length has later been improved to $O(\log n \log(1/\varepsilon))$ for constant-width *permutation* oblivious 1-branching programs [10, 6] which are regular programs with the two edges incoming to any node labeled 0 and 1, i.e. edges labeled with 0 respectively 1 create a permutation for each level-to-level transition [11].

In the constant-width regular 1-branching programs the fraction of inputs that are queried at any node is always lower bounded by a positive constant. This excludes the fundamental capability of general (non-regular) branching programs to recognize the inputs that contain a given substring on a non-constant number of selected positions. In our approach, we manage the analysis also for this essential case. In particular, we identify two types of convergence of the number of inputs along a computational path towards zero which implement read-once DNFs and CNFs, respectively. Thus, we achieve the construction of a hitting set generator for general width-3 1-branching programs which need

not be regular nor oblivious. In our previous work [14], we constructed the hitting set for so-called *simple* width-3 1-branching programs which exclude one specific pattern of level-to-level transition in their normalized form and cover the width-3 regular case.

In the present paper, we provide a polynomial time construction of a hitting set for read-once branching programs of width 3 with acceptance probability $\varepsilon > 5/6$, which need not be oblivious. This represents 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 formulate a so-called *richness* condition which is independent of a rather technical definition of branching programs. In fact, the (full) richness condition implies its weaker version which is equivalent to the definition of hitting sets for read-once conjunctions of DNF and CNF. Thus, a related line of study concerns pseudorandom generators for read-once formulas, such as read-once DNFs [7].

We show that the richness property characterizes in a certain sense the hitting sets for width-3 1-branching programs. In particular, its weaker version proves to be necessary for such hitting sets, while the sufficiency of richness represents the main result of this paper. More precisely, we show that any rich set extended with all strings within Hamming distance of 3 is a hitting set for 1-branching programs of width 3 with the acceptance probability greater than 5/6. The proof is based on a detailed analysis of structural properties of the width-3 1-branching programs that reject all the inputs from the candidate hitting set. Then, we prove that for a suitable constant C, 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 the richness condition, which implies our result. In addition, it follows from the latter result that almost $O(\log n)$-wise independent sets are weakly rich and hence, they hit the class of read-once conjunctions of DNF and CNF which is a generalization of the earlier result from [7]. A preliminary version of this article appeared as extended abstracts [16, 15] where our result is formulated for acceptance probability $\varepsilon > 11/12$.

The paper is organized as follows. After a brief review of basic definitions regarding branching programs in Section 2 (see [17] for more information), the weak richness condition is formulated and proved to be necessary in Section 3. The richness condition and its sufficiency is presented in Section 4 including the intuition behind the proof. The subsequent four Sections 5–8 are devoted to the technical proof of this proposition. Furthermore, our theorem that any almost $O(\log n)$-wise independent set is rich is presented in Section 9 where also the main steps of the technical proof occupying the subsequent four Sections 10–13 are introduced. Finally, our result is summarized in Section 14.

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 $P^{-1}(a) = \{\mathbf{a} \in \{0,1\}^n \,|\, P(\mathbf{a}) = a\}$ 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 queried 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 all nodes at each level 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 $v_j^{(k-1)}$ $(1 \leq j \leq w)$ on level $k-1$ of P to node $v_i^{(k)}$ $(1 \leq i \leq w)$ 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 $\mathbf{p}^{(k)} = (p_1^{(k)}, \ldots, p_w^{(k)})^\intercal$ 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 $k \geq 0$. 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

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

for every $k \geq \log w$ (hereafter, log denotes the binary logarithm). Obviously,

condition (2) can always be met by possible splitting (if $p_w^{(k)} = 0$) and permuting the nodes at each level of P:

Lemma 1 ([14]). *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 $1 > p_1^{(k)} \geq p_2^{(k)} \geq p_3^{(k)} > 0$, 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 $2 \leq k \leq d$ where $d \leq n$ is the depth of P. Note that the strict inequalities for $p_1^{(k)}$ and $p_3^{(k)}$ in (3) hold since $p_i^{(k)} \neq \frac{1}{3}$ according to (1) and $t_{ij}^{(k)} \in \{0, \frac{1}{2}, 1\}$.

3. The Weak Richness Condition Is Necessary

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

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

Furthermore, we say that a set $A \subseteq \{0,1\}^*$ is *weakly 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 $r \geq 0$, and for any $\mathbf{c} \in \{0,1\}^n$ 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 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).

Note that the product on the left-hand side of inequality (5) expresses the probability that a random string $\mathbf{a} \in \{0,1\}^n$ (not necessarily in A) satisfies conjunction (6) and (7). Moreover, this formula can be interpreted as a read-once conjunction of a DNF and a CNF (each variable occurs at most once)

$$\bigvee_{j=1}^{q} \bigwedge_{i \in Q_j} \ell(x_i) \wedge \bigwedge_{j=1}^{r} \bigvee_{i \in R_j} \neg\ell(x_i)\,, \quad \text{where} \quad \ell(x_i) = \begin{cases} x_i & \text{for } c_i = 1 \\ \neg x_i & \text{for } c_i = 0 \end{cases} \tag{10}$$

which accepts a random input with probability at least ε according to (5). Hence, the weak richness condition is, in fact, equivalent to the definition of a hitting set for read-once conjunctions of DNF and CNF. The following theorem shows that the weak richness condition is necessary for any set to be a hitting set for width-3 1-branching programs. It is because the 1-branching programs of width 3 can implement any read-once conjunction of DNF and CNF and a hitting set for a class of functions hits any of its subclass.

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

Proof. We proceed by transposition. Assume a set $H \subseteq \{0,1\}^*$ is not weakly 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

$$(\forall\, j \in \{1,\ldots,q\})\,(\exists\, i \in Q_j)\; a_i \neq c_i \tag{11}$$

$$\text{or} \quad (\exists\, j \in \{1,\ldots,r\})\,(\,\forall i \in R_j)\; a_i = c_i\,. \tag{12}$$

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

$$\frac{\left|P^{-1}(1)\right|}{2^n} \geq \varepsilon \quad \text{and} \quad (\forall\, \mathbf{a} \in H \cap \{0,1\}^n)\, P(\mathbf{a}) = 0\,, \tag{13}$$

which negates that H is an ε-hitting set for 1-branching programs of width 3 according to (4). In fact, P implements the corresponding conjunction of DNF and CNF (10).

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 queried within these blocks. The block associated with Q_j 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 $v_3^{(k_{q+1}-1)}$,

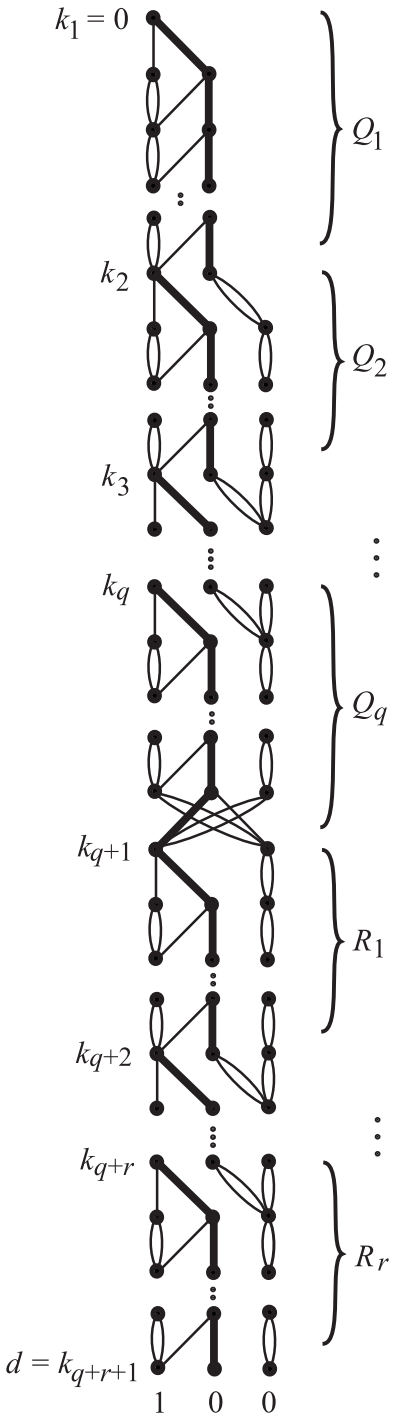

Figure 1: The Necessary Condition.

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

$$p_3^{(k_{q+1})} = \prod_{j=1}^{q} \left(1 - \frac{1}{2^{|Q_j|}}\right), \tag{14}$$

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):

$$\frac{\left|P^{-1}(1)\right|}{2^n} = p_1^{(d)} = \left(1 - \prod_{j=1}^{q} \left(1 - \frac{1}{2^{|Q_j|}}\right)\right) \times \prod_{j=1}^{r} \left(1 - \frac{1}{2^{|R_j|}}\right) \geq \varepsilon\,. \tag{15}$$

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 (11) of the underlying disjunction, whereas every $\mathbf{a} \in A_2 = H \setminus A_1$ meets the second term (12). 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 (13), which completes the proof. □

4. The Richness Condition Is Sufficient

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

$$\prod_{j=1}^{r} \left(1 - \frac{1}{2^{|R_j|}}\right) \geq \varepsilon\,, \tag{16}$$

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

$$(\forall\, i \in Q)\, a_i = c_i \text{ and } (\forall\, j \in \{1, \ldots, r\})\, (\exists\, i \in R_j)\, a_i \neq c_i\,. \tag{17}$$

One can observe that an ε-rich set is weakly ε-rich (see Section 3) since inequality (5) implies (16) and ensures that there is index $j \in \{1, \ldots, q\}$ of $Q_j = Q$ such that $|Q| \leq \log n$. In particular, if $|Q_j| > \log n$ for every $j = 1, \ldots, q$, then inequality (5) would give

$$1 - \varepsilon \geq \prod_{j=1}^{q} \left(1 - \frac{1}{2^{|Q_j|}}\right) \geq \left(1 - \frac{1}{2^{\log n}}\right)^{\frac{n}{\log n}} > 1 - \frac{1}{n} \cdot \frac{n}{\log n} = 1 - \frac{1}{\log n} \tag{18}$$

which is a contradiction for $n > 2^{1/\varepsilon}$. Thus, we have (17) which validates the conjunction of (6) and (7) completing the argument.

It follows that any rich set is a hitting set for read-once conjunctions of DNF and CNF. Also note that formula (17) can be interpreted as a read-once CNF (cf. 10)

$$\bigwedge_{i \in Q} \ell(x_i) \wedge \bigwedge_{j=1}^{r} \bigvee_{i \in R_j} \neg \ell(x_i)\,, \quad \text{where} \quad \ell(x_i) = \begin{cases} x_i & \text{for } c_i = 1 \\ \neg x_i & \text{for } c_i = 0 \end{cases} \tag{19}$$

which contains at most logarithmic number of single literals together with clauses whose sizes satisfy (16). Moreover, Theorem 3 in Section 9 proves that any almost $O(\log n)$-wise independent set satisfies the richness condition.

The following theorem shows that the richness condition is, in a certain sense, sufficient for a set to be a hitting set for 3-width 1-branching programs. 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. *Let $\varepsilon > \frac{5}{6}$. If A is ε'^{11}-rich for some $\varepsilon' < \varepsilon$, then $H = \Omega_3(A)$ is an ε-hitting set for the class of read-once branching programs of width 3.*

PROOF. Suppose a read-once branching program P of width 3 with sufficiently many input variables n meets

$$\frac{\left|P^{-1}(1)\right|}{2^n} \geq \varepsilon > \frac{5}{6}\,. \tag{20}$$

We will prove that there exists $\mathbf{a} \in H$ such that $P(\mathbf{a}) = 1$. On the contrary, we assume that

$$P(\mathbf{a}) = 0 \quad \text{for every} \quad \mathbf{a} \in H\,, \tag{21}$$

which will lead to a contradiction. Without loss of generality[1], we assume that P is normalized according to Lemma 1.

4.1. The Plan of the Proof

In this paragraph we will informally explain the main ideas of the proof with the pointers to the subsequent paragraphs and sections where the precise and detailed argument is given. The assumption that branching program P accepts large fraction of inputs and rejects all the inputs from candidate hitting set H constrains the structure of P severely. In particular, we inspect the structure of P with respect to (20) and (21) from its last level d (containing the sinks) and we proceed in the analysis step by step backwards to lower levels[2]. For this purpose, various parameters denoting certain levels in P are defined which are used to describe the structure of P. These definitions of levels, indicated in boldface, are scattered in the following proof since the definition of a level often builds on the previous analysis of P.

The underlying inspection reveals that the structure at the end of branching program P can be split into *blocks* whose typical shape is schematically depicted in Figure 2 while the subsequent Figures 3–8 focus on particular parts of the block. Figure 2 also summarizes the definitions of levels in the block having the form of "$a \leq \mathbf{b} \uparrow \leq c : \ C(b)$" which means b is the *greatest* level such that $a \leq b \leq c$ and condition $C(b)$ is satisfied (similarly, $\downarrow$ denotes the *least* such level). In addition, there are main equations listed in Figure 2 concerning the distribution $p_1^{(k)}, p_2^{(k)}, p_3^{(k)}$ of inputs among three nodes at important levels of the block.

The *last* level of the block is denoted by m and this **level m** satisfies the following four so-called *m-conditions*:

[1] More precisely, the logical argument goes as follows. Branching program P is transformed to an equivalent branching program P_1 which computes the same function as P (i.e. P_1 preserves (20) and (21)) and has some additional property (e.g. P_1 is normalized). In the following proof, several equivalent transformations are employed one after the other in order to achieve various extra properties, which generates a sequence of branching programs $P, P_1, P_2, \ldots, P_c$. After showing that the existence of the last program P_c eventually leads to a contradiction one can conclude that the original program P cannot exist.

[2] Recall that we number the levels of P from the zero level containing the source up to the last level d which is composed of sinks. This means that, in figures, the lower levels are situated on the top of branching program whereas the upper levels are located at the bottom.

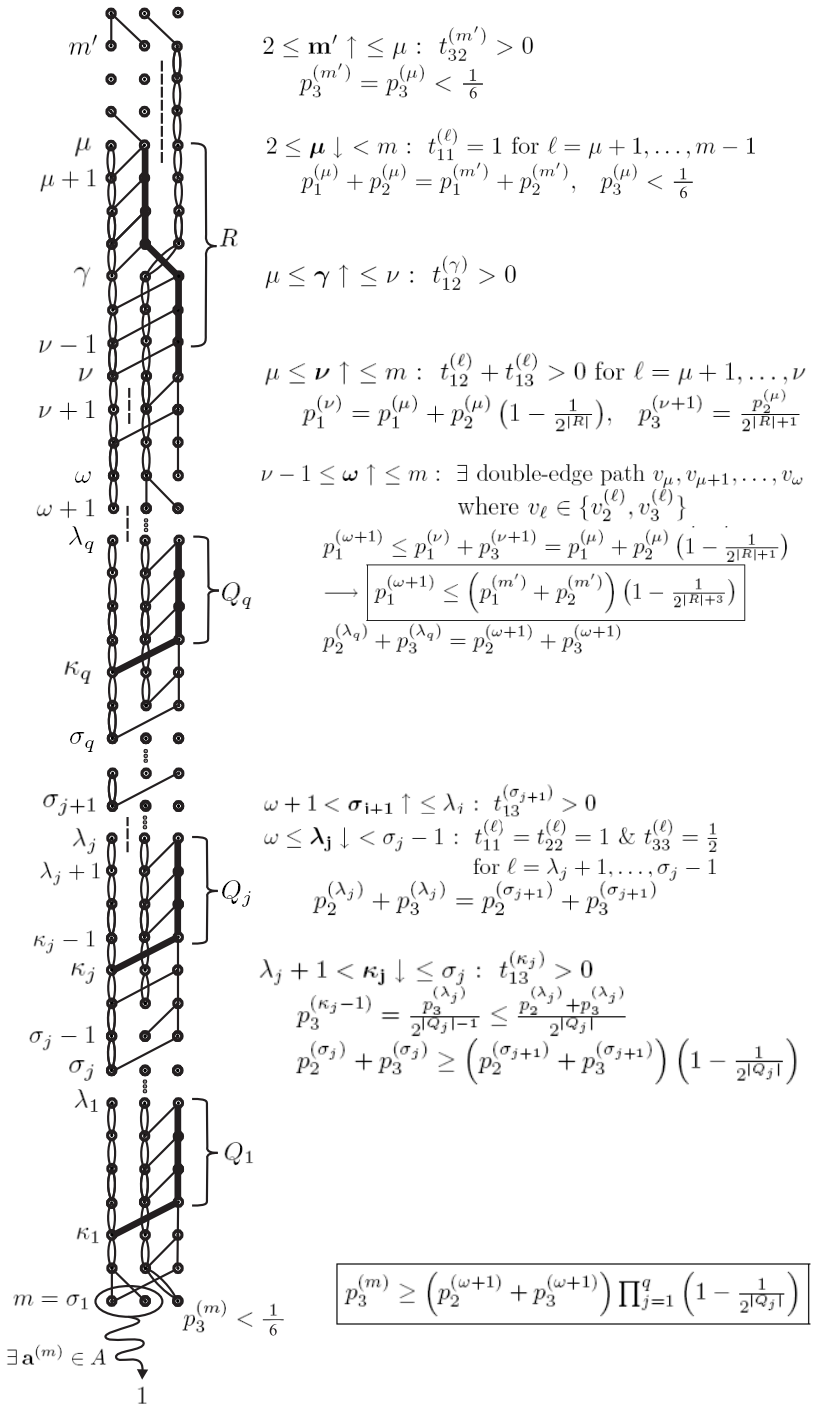

Figure 2: The structure of a typical block

1. $t_{11}^{(m)} = t_{21}^{(m)} = \frac{1}{2}$,
2. $t_{32}^{(m)} > 0$,
3. $p_3^{(m)} < \frac{1}{6}$,
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.

Without loss of generality[1], these m-conditions can also be met for $m = d$ (Paragraph 4.2) which is the last level of P. In particular, the sinks $v_1^{(d)}$ and $v_2^{(d)}$ are labeled with 1 according to m-condition 4. Thus, the inspection of the structure of P starts with the analysis of the first[3] block which constitutes the tail of P.

The *first* level of the block is denoted by m' (its formal definition can be found in Paragraph 6.2) which, in a typical case, proves also to satisfy the four m'-conditions 1–4. Thus the block is delimited by levels m' and m. The shape of the block is being revealed step by step by the case analysis (Sections 5 and 6) which starts from m and proceeds towards lower levels down to m'. We will now shortly outline the structure of a typical block as depicted in Figure 2 which results from this analysis. From the first level m' through μ, there is no edge between the first two columns on the one hand, and the third column on the other hand, which means there is a double-edge path in the third column from m' through μ (Paragraph 6.2). Moreover, there is a double-edge path in the first column starting at level μ which leads up to level $m-1$ where it is split into vertices $v_1^{(m)}$ and $v_2^{(m)}$ at the next level m (cf. m-condition 1).

On the top of Figure 2, a single-edge path from μ to ν is indicated in boldface which is used to define the partition class R associated with this block (Paragraph 5.1). In particular, class R contains all the indices of the variables that are queried on this computational path up to level $\nu - 1$. Moreover, the edge labels on this path define relevant bits of $\mathbf{c} \in \{0,1\}^n$ so that any input passing through this path that differs from $\mathbf{c}$ in at least one bit location from R turns to the double-edge path in the first column and consequently comes in node $v_1^{(m)}$ or $v_2^{(m)}$. This implements one CNF clause $\vee_{i \in R} \neg \ell(x_i)$ from (19). Similarly, sets Q_j for $j = 1, \ldots, q$ associated with this block are defined (Paragraph 5.2) using the single-edge paths from λ_j to κ_j which are also highlighted in Figure 2 so that any input that passes through $v_3^{(\lambda_j)}$ and agrees with $\mathbf{c}$ on all the bit locations from Q_j reaches the double-edge path in the first column coming in $v_1^{(m)}$ or $v_2^{(m)}$. This implements DNF monomials $\wedge_{i \in Q_j} \ell(x_i)$ in (10) which are candidates for the monomial $\wedge_{i \in Q} \ell(x_i)$ in (19).

Under certain assumptions ((34) and (35)), one can show that level m' satisfies m'-condition 1–3 (Paragraph 6.2). This opens the possibility that the first level $m' = m_r$ of the current rth block at the same time represents the last level of the next lower-level $(r+1)$st block to which the structural analysis could

[3]The blocks are numbered from the bottom to the top of branching program P in the order reverse to that of levels.

recursively be applied (Section 7). It suffices to show that level m' also meets m'-condition 4. For this purpose, the richness condition (17) is employed for $Q = \emptyset$ and for the partition classes $R_1, \ldots, R_r$ associated with the first r blocks (that have been analyzed so far), provided that this partition satisfies (16). This gives an input $\mathbf{a}^{(m')} \in A$ such that for every block $j = 1, \ldots, r$ there is $i \in R_j$ such that $a_i^{(m')} \neq c_i$ according to (17), that is, $\mathbf{a}^{(m')}$ satisfies $\wedge_{j=1}^{r} \vee_{i \in R_j} \neg \ell(x_i)$ from (19). Hence, if we put this $\mathbf{a}^{(m')}$ at node $v_1^{(m')}$ or $v_2^{(m')}$ ($m' = m_r$), then the block structure in Figure 2 ensures that $\mathbf{a}^{(m')}$ also traverses $v_1^{(\mu)}$ or $v_2^{(\mu)}$ and reaches the double-edge path in the first column coming in $v_1^{(m)}$ or $v_2^{(m)}$ ($m = m_{r-1}$), by the definition of R and c_i for $i \in R$. This argument is applied recursively to each block $j = r, r-1, \ldots, 1$ which implies that $\mathbf{a}^{(m')}$ eventually arrives to the sink $v_1^{(d)}$ or $v_2^{(d)}$ ($m_0 = d$) labeled with 1. This proves the m'-condition 4 also for level m'. Thus the analysis including the definition of an associated partition class $R = R_{r+1}$ and sets $Q_j = Q_{r+1,j}$ is applied recursively to the next $(r+1)$st block for m replaced with m' etc.

If, on the other hand, the underlying partition does not satisfy (16), then one can prove that there is a set $Q = Q_{bj}$ associated with the bth block among the first r blocks (that have been analyzed so far) such that $|Q| \leq \log n$, and the recursive analysis ends (Section 8). In this case, the richness condition (17) for this set Q and for partition $R_1, \ldots, R_{b-1}$ provides $\mathbf{a} \in A$ such that $a_i = c_i$ for each $i \in Q$ and for every block $j = 1, \ldots, b-1$ there is $i \in R_j$ such that $a_i \neq c_i$, that is, $\mathbf{a} \in A$ satisfies $\wedge_{i \in Q} \ell(x_i) \wedge \wedge_{j=1}^{b-1} \vee_{i \in R_j} \neg \ell(x_i)$ according to (19). Moreover, one can show (Lemma 8) that there is an h-neighbor $\mathbf{a}' \in \Omega_2(\mathbf{a}) \subseteq H$ that differs from this $\mathbf{a}$ in at most two bits so that these bits guarantee that the computational path for $\mathbf{a}'$ in the bth block either reaches the double-edge path in the first column, or comes in node $v_3^{(\lambda_j)}$ (see Figure 2). In the latter case, $\mathbf{a}'$ further traverses the path corresponding to Q which reaches the double-edge path in the first column anyway by the definition of Q_j and c_i for $i \in Q_j$. In both cases, input $\mathbf{a}'$ traverses node $v_1^{(m_{b-1})}$ or $v_2^{(m_{b-1})}$, and by the above-mentioned recursive argument it eventually arrives to the sink $v_1^{(d)}$ or $v_2^{(d)}$ labeled with 1. This provides the desired contradiction $P(\mathbf{a}') = 1$ for $\mathbf{a}' \in H$.

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 $|P^{-1}(0)|/2^n < \frac{1}{6}$ according to (20). 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 $p_1^{(d)} \geq p_2^{(d)} > \frac{1}{6} > p_3^{(d)}$ (m-condition 3) is preserved by (20). 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 (20), 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[1].

4.3. A Technical Lemma

Let **level μ'** be the least level of P such that $2 \leq \mu' < m$ and $t_{11}^{(\ell)} = 1$ for every $\ell = \mu' + 1, \ldots, m - 1$. We define **level μ** as

$$\mu = \begin{cases} \mu' - 1 & \text{if } t_{12}^{(\mu')} = 1 \text{ and } t_{11}^{(\mu')} = t_{21}^{(\mu')} = \frac{1}{2} \\ \mu' & \text{otherwise.} \end{cases} \tag{22}$$

For the analysis of a single block structure (Sections 4–6, 8), we swap $v_1^{(\mu)}$ and $v_2^{(\mu)}$ if $\mu = \mu' - 1$ for the notation simplicity so that $t_{11}^{(\ell)} = 1$ for every $\ell = \mu + 1, \ldots, m - 1$ at the cost of violating condition $p_1^{(\mu)} \geq p_2^{(\mu)}$ given by (2). Thus, for $\mu = \mu' - 1$, assume $p_1^{(\mu)} < p_2^{(\mu)}$, $t_{11}^{(\mu+1)} = 1$, and $t_{12}^{(\mu+1)} = t_{22}^{(\mu+1)} = \frac{1}{2}$. For the recursion (Section 7) when the last level m in the next (lower-level) block may coincide with level μ of the current block we will nevertheless assume the original node order and $p_1^{(\mu)} \geq p_2^{(\mu)}$.

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

Lemma 2.

(i) $\mu > 3$.

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

(iii) *If $t_{12}^{(k+1)} > 0$ for some level k such that $\mu \leq k < m$, 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 3).*

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

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 3).*

PROOF.

(i) Suppose $\mu \leq 3$ and let $\mathbf{a}^{(m)} \in A$ be the input from m-condition 4. Then there exists an h-neighbor $\mathbf{a}' \in \Omega_3(\mathbf{a}^{(m)})$ of $\mathbf{a}^{(m)}$ whose computational path starting from 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 $\mu > 3$.

(ii) Suppose there are two simultaneous switching paths starting from $v_2^{(k)}$ and $v_3^{(k)}$, respectively, at some level k such that $\mu \leq k < m$, 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).

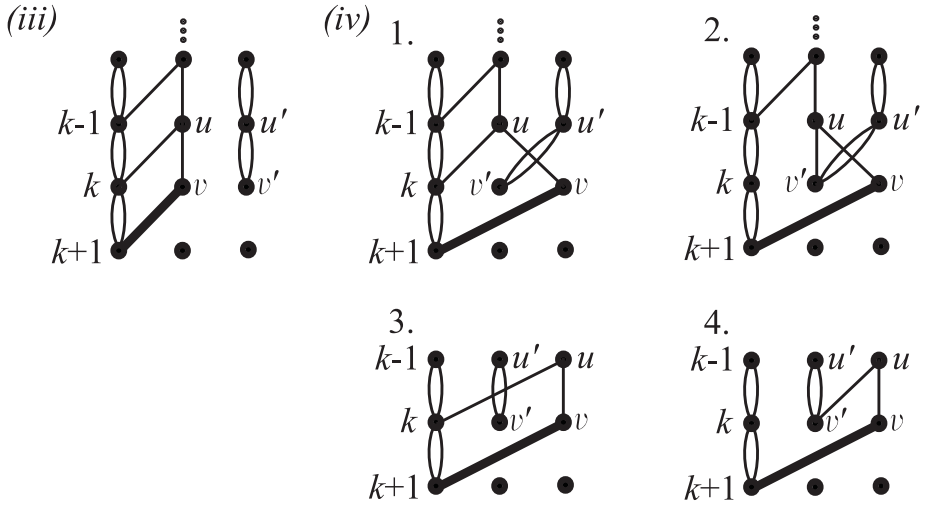

Figure 3: Lemma 2.iii and iv.

As depicted in Figure 3, at level k such that $\mu < k < m$, denote by $v \in \{v_2^{(k)}, v_3^{(k)}\}$ a node 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 $v_3^{(k-1)}$, 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:

(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 suppose $k < m-1$. Also for $k = \mu = \mu' - 1$ we know $t_{12}^{(\mu+1)} = \frac{1}{2}$ and thus we further assume $k \geq \mu'$. 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, which is a contradiction with the normalization of P.
For $k = \mu$, on the other hand, we know that $\mu = \mu' > 3$ from the definition of μ and we will first observe that there are at least two edges leading to $v_3^{(\mu)}$. Suppose that only one edge leads to $v_3^{(\mu)}$ from $u \in \{v_1^{(\mu-1)}, v_2^{(\mu-1)}, v_3^{(\mu-1)}\}$. If $\mathbf{a}^{(m)} \notin M(u)$, then $\mathbf{a}^{(m)} \in M(v_1^{(\mu)}) \cup M(v_2^{(\mu)}) = M(v_1^{(\mu+1)}) \subseteq M(v_1^{(m)}) \cup M(v_2^{(m)})$ implying $P(\mathbf{a}^{(m)}) = 1$ according to m-condition 4. If $\mathbf{a}^{(m)} \in M(u)$, then an h-neighbor $\mathbf{a}' \in \Omega_1(\mathbf{a}^{(m)}) \cap M(u) \subseteq H$ of $\mathbf{a}^{(m)}$ exists which differs from $\mathbf{a}^{(m)}$ in the variable that is tested at u and thus $\mathbf{a}' \in M(v_1^{(\mu)}) \cup M(v_2^{(\mu)})$ implying $P(\mathbf{a}') = 1$. Now, with the two edges leading to $v_3^{(\mu)}$, we could split $v_3^{(\mu)}$ into two nodes and merge $v_1^{(\mu)}$ and $v_2^{(\mu)}$ while preserving the function of P. Thus, for $t_{12}^{(\mu+1)} = 1$ we can construct an equivalent branching program with $t_{12}^{(\mu+1)} = 0$[1].

(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 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 in case 2 there is a switching path from $v_2^{(k-1)}$ to $v_1^{(k+1)}$ via $v_3^{(k)}$ (substituting for $t_{12}^{(k)} > 0$) and 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 Class R and Sets $Q_1, \ldots, Q_q$

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 ν** by using Lemma 2, where ν is the greatest level such that $\mu \leq \nu \leq m$ and $t_{12}^{(\ell)} + t_{13}^{(\ell)} > 0$ for every $\ell = \mu+1, \ldots, \nu$. Note that $\nu = \mu$ for $t_{12}^{(\mu+1)} = t_{13}^{(\mu+1)} = 0$. In addition, let **level γ** be the greatest level such that $\mu \leq \gamma \leq \nu$ and $t_{12}^{(\gamma)} > 0$ if such γ exists, otherwise set $\gamma = \mu$.

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* $t_{12}^{(\gamma)} = t_{32}^{(\gamma)} = \frac{1}{2}$ *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).*

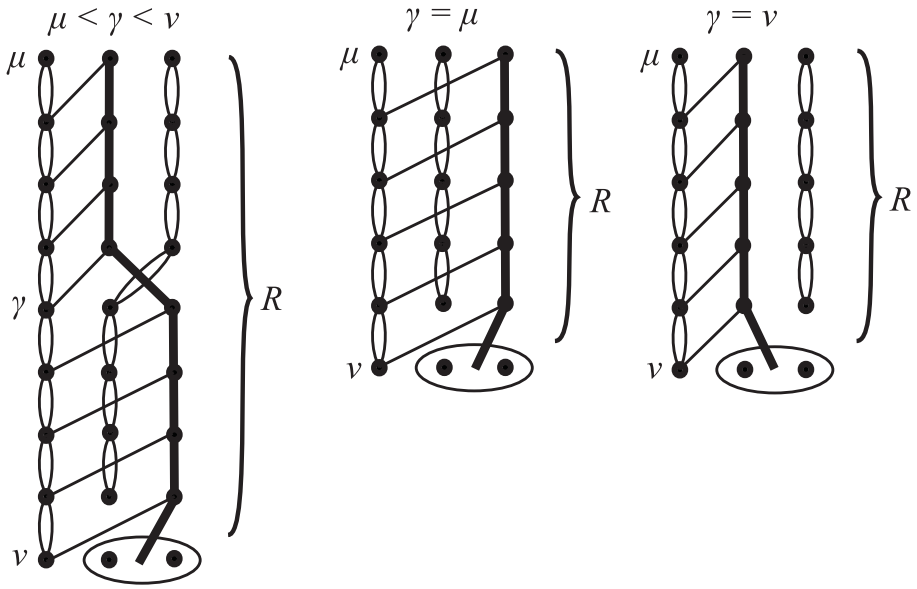

Figure 4: The block structure from level μ through $\nu < m$ according to Corollary 1.

Figure 4 shows a typical structure of the block from level μ through ν for the case of $\nu < m$, which comes out of Corollary 1. In particular, there are two disjoint double-edge paths starting from level μ. One follows the first column from $v_1^{(\mu)}$ through $v_1^{(\nu)}$. For $\nu < \gamma < \nu$, the other double-edge path starts from $v_3^{(\mu)}$, follows the third column and turns to $v_2^{(\gamma)}$ on level γ, and further continues through the second column up to $v_2^{(\nu-1)}$. For $\gamma = \mu$, this double-edge path follows only the second column leading from $v_2^{(\mu)}$ through $v_2^{(\nu-1)}$, whereas for $\gamma = \nu$, it follows the third column from $v_3^{(\mu)}$ through $v_3^{(\nu-1)}$. In addition, there is a node left on each level from μ through $\nu - 1$ that does not lay on the underlying two disjoint double-edge paths. These remaining nodes are connected in a single-edge path from level μ through $\nu - 1$ extended with an edge to $v_2^{(\nu)}$ or $v_3^{(\nu)}$. For each node on this single-edge path the other outgoing edge leads to the double-edge path in the first column.

Furthermore, we shortly analyze level m for the special case of $\nu = m$ as depicted in Figure 5. Recall that $t_{11}^{(m)} = t_{21}^{(m)} = \frac{1}{2}$ and $t_{32}^{(m)} > 0$ by m-condition 1

and 2, respectively. Moreover, either $t_{12}^{(m)} = \frac{1}{2}$ (i.e. $\nu = \gamma$) or $t_{13}^{(m)} > 0$ (i.e. $\nu > \gamma$) by the definition of ν. It follows from Lemma 2.ii that either $t_{33}^{(m)} = 1$ for $\nu = \gamma$ or $t_{32}^{(m)} = 1$ for $\nu > \gamma$. In the latter case of $\nu > \gamma$, the other edge from $v_3^{(m-1)}$ may lead either to $v_3^{(m)}$ (i.e. $t_{13}^{(m)} = t_{33}^{(m)} = \frac{1}{2}$) or to $v_1^{(m)}$ (i.e. $t_{13}^{(m)} = 1$) or $v_2^{(m)}$ (i.e. $t_{13}^{(m)} = t_{23}^{(m)} = \frac{1}{2}$). This completes the analysis of level $\nu = m$. We say that the underlying block is an *empty block* if $\nu = m$ and $t_{33}^{(m)} = 0$ (i.e. $t_{13}^{(m)} + t_{23}^{(m)} = 1$ and $t_{32}^{(m)} = 1$).

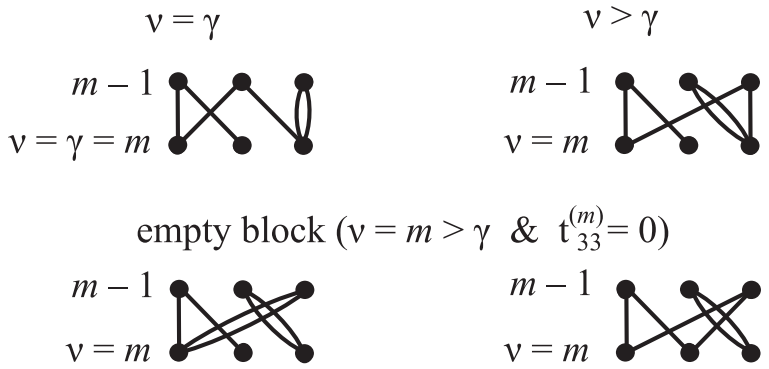

Figure 5: Level m for $\nu = m$.

Corollary 1 will be used for the definition of partition class R associated with the current block, if this block is not empty, which is illustrated in Figure 4. Moreover, class R is neither defined for $\nu = \mu$ when only sets $Q_1, \ldots, Q_q$ are associated with the block (see Paragraph 5.2 and Lemma 3 in particular). Thus, 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 $\gamma = \nu$) where **level $\boldsymbol{\nu'}$** is defined as

$$\nu' = \min(\nu, m-1)\,. \tag{23}$$

For the future use of condition (17) we also define relevant bits of string $\mathbf{c} \in \{0,1\}^n$. Thus, let c_i^R be the corresponding labels of the edges creating this computational path (indicated by a bold line in Figure 4) 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 $v_3^{(\nu')}$.

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 $\mu \le \omega \le m$ and there is a double-edge path from μ through ω containing only nodes $v_\ell \in \{v_2^{(\ell)}, v_3^{(\ell)}\}$ for every $\ell = \mu, \ldots, \omega$. Note that this path possibly extends the double-edge path from Corollary 1 (see Figure 4) 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 $\gamma = \nu > \mu$). Hence,

$$\omega \geq \max(\nu - 1, \mu)\,. \tag{24}$$

For the special case of $\omega = m$ (including the empty block) when this double-edge path reaches level m, no sets $Q_1, \ldots, Q_q$ are associated with the current block and we set $q = 0$. In this case, we will observe in the following lemma that $\nu > \mu$, which ensures that at least class R is defined for a non-empty block (Paragraph 5.1) when $\omega = m$.

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

Proof. On the contrary, suppose $\omega = m$ and $\nu = \mu$. It follows from Corollary 1.5 that $t_{12}^{(m)} = 0$. Moreover, $t_{22}^{(m)} = 0$ since $t_{22}^{(m)} > 0$ would require $t_{13}^{(m)} > 0$ by the normalization of P, which contradicts Lemma 2.ii, and hence, $t_{32}^{(m)} = 1$. If $t_{21}^{(m-1)} + t_{31}^{(m-1)} > 0$, then $p_3^{(m)} \geq p_2^{(m-1)} \geq p_1^{(m-2)}/2 > \frac{1}{6}$ due to (3) which is in contradiction to m-condition 3. Hence, $t_{11}^{(m-1)} = 1$ which means $\mu' < m - 1$. Furthermore, $t_{12}^{(\mu+1)} = t_{13}^{(\mu+1)} = 0$ by the definition of ν implying $\mu = \mu'$. 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 $\omega = m$, which implies $t_{22}^{(\ell)} = 1$ for $\ell = \mu + 2, \ldots, m - 1$. It follows that $p_2^{(\mu+1)} \leq p_3^{(m)} < \frac{1}{6}$ according to m-condition 3.

On the other hand, we know $t_{21}^{(\mu)} + t_{31}^{(\mu)} > 0$ by the definition of μ', which implies $p_2^{(\mu)} > \frac{1}{6}$ due to (3). Hence, $t_{22}^{(\mu+1)} = t_{32}^{(\mu+1)} = \frac{1}{2}$ and $t_{23}^{(\mu+1)} = 1$ because of $p_2^{(\mu+1)} < \frac{1}{6}$. This ensures $t_{11}^{(\mu)} = t_{21}^{(\mu)} = \frac{1}{2}$ since $p_1^{(\mu-1)} > \frac{1}{3}$. Thus, $\frac{1}{6} > p_2^{(\mu+1)} > p_1^{(\mu-1)}/4$ which rewrites as $p_1^{(\mu-1)} < \frac{2}{3}$ implying $p_2^{(\mu-1)} + p_3^{(\mu-1)} > \frac{1}{3}$, and hence $p_2^{(\mu-1)} > \frac{1}{6}$ due to $p_2^{(\mu-1)} \geq p_3^{(\mu-1)}$. Clearly, $t_{32}^{(\mu)} = 0$ since otherwise we get a contradiction $\frac{1}{6} > p_2^{(\mu+1)} \geq \frac{1}{4}p_1^{(\mu-1)} + \frac{1}{2}p_2^{(\mu-1)} > \frac{1}{4} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{6}$. Similarly, $t_{22}^{(\mu)} = 1$ produces a contradiction $\frac{1}{6} > p_2^{(\mu+1)} \geq \frac{1}{4}p_1^{(\mu-1)} + \frac{1}{2}p_2^{(\mu-1)} > \frac{1}{6}$. It follows that $t_{12}^{(\mu)} > 0$ whereas $t_{12}^{(\mu)} = 1$ contradicts $\mu = \mu'$ according to (22), and hence $t_{12}^{(\mu)} = t_{22}^{(\mu)} = \frac{1}{2}$ and $t_{33}^{(\mu)} > 0$. This gives a contradiction $\frac{1}{6} > p_2^{(\mu+1)} \geq \frac{1}{4}(p_1^{(\mu-1)} + p_2^{(\mu-1)}) + \frac{1}{2}p_3^{(\mu-1)} > \frac{1}{4}(p_1^{(\mu-1)} + p_2^{(\mu-1)} + p_3^{(\mu-1)}) = \frac{1}{4}$. □

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[1]. 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 $t_{11}^{(m)} = t_{21}^{(m)} = \frac{1}{2}$ by m-condition 1 and m-condition 3 ensures $t_{11}^{(m-1)} = 1$. 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 previous upper-level neighboring block, which is consistent with its structure (see Paragraph 6.2 and Figure 7 in particular).

Thus, we assume $t_{13}^{(m)} > 0$ without loss of generality, which implies $t_{32}^{(m)} = 1$ by Lemma 2.ii and $t_{11}^{(m-1)} = 1$ according to m-condition 3. 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.

In general, we start 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 λ_j is the least level such that $\omega \leq \lambda_j < \sigma_j - 1$ and the transitions from case 3 or 4 of Lemma 2.iv occur for all levels $\ell = \lambda_j + 1, \ldots, \sigma_j - 1$ as depicted in Figure 6. This means $t_{11}^{(\ell)} = t_{22}^{(\ell)} = 1$ and $t_{33}^{(\ell)} = \frac{1}{2}$ for every $\ell = \lambda_j + 1, \ldots, \sigma_j - 1$. 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 Q_j associated with the current block (a candidate for Q in the richness condition (17)) to be a set of indices of the variables that are tested on the computational path $v_3^{(\lambda_j)}, v_3^{(\lambda_j+1)}, \ldots, v_3^{(\kappa_j-1)}$, and let $c_i^{Q_j}$ be the corresponding labels of the edges creating this path including the edge from $v_3^{(\kappa_j-1)}$ to $v_1^{(\kappa_j)}$ (indicated by a bold line in Figure 6). This extends the definition of $\mathbf{c} \in \{0,1\}^n$ associated with R and Q_k for $1 \leq k < j$, which are usually pairwise disjoint due to P is read-once. Nevertheless, the definition of $\mathbf{c}$ may not be unique for indices from their nonempty intersections in some very special cases (including those corresponding to neighboring blocks) but the richness condition will only be used for provably disjoint sets (see Section 7).

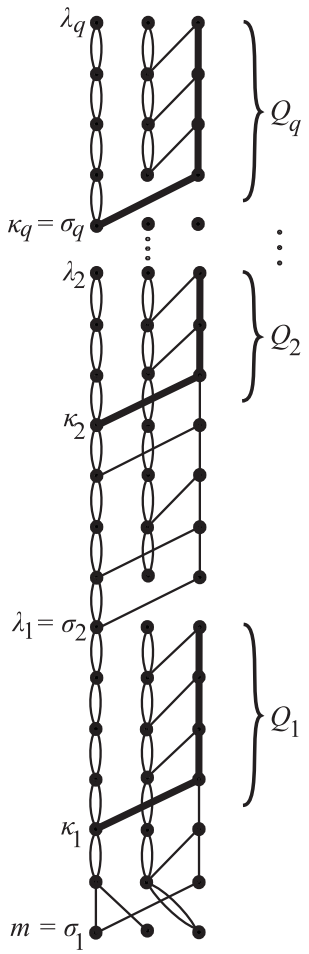

Figure 6: The definition of $Q_1, \ldots, Q_q$.

Finally, define next **level $\boldsymbol{\sigma_{j+1}}$** to be the greatest level such that $\omega + 1 < \sigma_{j+1} \leq \lambda_j$ and $t_{13}^{(\sigma_{j+1})} > 0$, if such σ_{j+1} exists, 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} does not exist, then set $q = j$ and the definition of sets $Q_1, \ldots, Q_q$ associated with the current 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 (16). 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 6), which gives

$$p_2^{(\kappa_j-1)} + p_3^{(\kappa_j-1)} = p_2^{(\lambda_j)} + p_3^{(\lambda_j)}, \tag{25}$$

$$p_3^{(\kappa_j-1)} = \frac{p_3^{(\lambda_j)}}{2^{|Q_j|-1}} \leq \frac{p_2^{(\lambda_j)} + p_3^{(\lambda_j)}}{2^{|Q_j|}} \tag{26}$$

because $p_3^{(\lambda_j)} \leq \frac{1}{2}(p_2^{(\lambda_j)} + p_3^{(\lambda_j)})$. We know $t_{12}^{(\ell)} = 0$ for every $\ell = \omega+2, \ldots, m$ according to Corollary 1.5 where $\nu + 1 \leq \omega + 2$ by (24). Moreover, it follows from the definition of $\sigma_{j+1} > \omega + 1$ that $t_{13}^{(\ell)} = 0$ for every $\ell = \sigma_{j+1}+1, \ldots, \lambda_j$ for any $1 \leq j < q$, and $t_{13}^{(\ell)} = 0$ for every $\ell = \omega+2, \ldots, \lambda_q$. Hence,

$$p_2^{(\sigma_{j+1})} + p_3^{(\sigma_{j+1})} = p_2^{(\lambda_j)} + p_3^{(\lambda_j)} = p_2^{(\kappa_j-1)} + p_3^{(\kappa_j-1)} \tag{27}$$

for $1 \leq j < q$ and

$$p_2^{(\omega+1)} + p_3^{(\omega+1)} = p_2^{(\lambda_q)} + p_3^{(\lambda_q)} = p_2^{(\kappa_q-1)} + p_3^{(\kappa_q-1)} \tag{28}$$

according to (25). Note that equation (28) holds trivially for $\lambda_q = \omega+1$ and it is also valid for $\lambda_q = \omega$ (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). Furthermore, we know $t_{22}^{(\ell)} = 1$ for every $\ell = \kappa_j, \ldots, \sigma_j - 1$ and $t_{12}^{(\sigma_j)} = 0$, which implies

$$\begin{aligned} p_2^{(\sigma_j)} + p_3^{(\sigma_j)} &\geq p_2^{(\kappa_j-1)} + p_3^{(\kappa_j-1)} - p_3^{(\kappa_j-1)} \\ &\geq p_2^{(\kappa_j-1)} + p_3^{(\kappa_j-1)} - \frac{p_2^{(\lambda_j)} + p_3^{(\lambda_j)}}{2^{|Q_j|}} \\ &= \left(p_2^{(\sigma_{j+1})} + p_3^{(\sigma_{j+1})}\right)\left(1 - \frac{1}{2^{|Q_j|}}\right) \end{aligned} \tag{29}$$

for $1 < j < q$ according to (26) and (27), while formula (29) reads

$$p_3^{(m)} = p_3^{(\sigma_1)} \geq \left(p_2^{(\sigma_2)} + p_3^{(\sigma_2)}\right)\left(1 - \frac{1}{2^{|Q_1|}}\right) \tag{30}$$

for $j = 1 < q$ due to $t_{32}^{(m)} = 1$, whereas (29) is rewritten as

$$p_2^{(\sigma_q)} + p_3^{(\sigma_q)} \geq \left(p_2^{(\omega+1)} + p_3^{(\omega+1)}\right)\left(1 - \frac{1}{2^{|Q_q|}}\right) \tag{31}$$

for $j = q > 1$ according to (28). Thus starting with (30), inequality (29) is applied recursively for $j = 2, \ldots, q-1$, and, in the end, formula (31) is employed, leading to

$$p_3^{(m)} \geq \left(p_2^{(\omega+1)} + p_3^{(\omega+1)}\right)\prod_{j=1}^{q}\left(1 - \frac{1}{2^{|Q_j|}}\right) \tag{32}$$

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

$$p_1^{(m)} + p_2^{(m)} \leq 1 - \left(1 - p_1^{(\omega+1)}\right) \prod_{j=1}^{q} \left(1 - \frac{1}{2^{|Q_j|}}\right) \tag{33}$$

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

$$p_3^{(\mu)} < \frac{1}{6}, \tag{34}$$

$$\prod_{j=1}^{q} \left(1 - \frac{1}{2^{|Q_j|}}\right) > \frac{2}{3} \tag{35}$$

where the product in (35) equals 1 for $q = 0$. Based on these assumption, we will further analyze the block structure below[2] level μ in order to satisfy the m'-conditions 1–4 (see Paragraph 4.1) also for the first block level m' (the formal definition of m' appears at the beginning of Paragraph 6.2) so that the analysis can be applied recursively when inequalities (34) and (35) 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. For $\mu = \mu' - 1$, the proposition follows from the definition of μ, and thus assume $\mu = \mu'$. Consider first the special case of $\mu + 1 = m$ for which $t_{12}^{(m)} = 0$ implies $t_{32}^{(m)} = 1$ by using m-condition 2, Lemma 2.ii, and the normalization of P. Hence, we get a contradiction $\frac{1}{6} > p_3^{(m)} \geq p_2^{(m-1)} \geq \frac{1}{2} p_1^{(m-2)} > \frac{1}{6}$ by m-condition 3 and the definition of μ'. Thus further assume $\mu < m - 1$. Clearly, $t_{32}^{(\mu+1)} < 1$ by the normalization of P whereas $t_{33}^{(\mu+1)} = 1$ implies $t_{12}^{(\mu+1)} = \frac{1}{2}$, and thus, further consider the case when no double edge leads to $v_3^{(\mu+1)}$. If $t_{12}^{(\mu+1)} > 0$, then $t_{12}^{(\mu+1)} = \frac{1}{2}$ by Lemma 2.iii for $k = \mu$. On the contrary, suppose $t_{12}^{(\mu+1)} = 0$, which gives $t_{22}^{(\mu+1)} > 0$ due to $t_{32}^{(\mu+1)} < 1$. Assumption (34) ensures $t_{31}^{(\mu)} = 0$ which implies $t_{21}^{(\mu)} > 0$ by the definition of μ'.

We will first show that

$$p_2^{(\mu+1)} < \frac{1}{4}. \tag{36}$$

For $\omega < m$, assumption (35) together with m-condition 3 ensures

$$p_2^{(\omega+1)} + p_3^{(\omega+1)} < \frac{1}{4} \tag{37}$$

according to (32), which gives (36) for $\omega = \mu$. For $\omega > \mu$, 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 $v_2^{(\mu+1)}$ as we assume no double edge to $v_3^{(\mu+1)}$. 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, $p_2^{(\mu+1)} \le p_2^{(\omega+1)} + p_3^{(\omega+1)} < \frac{1}{4}$ due to (37). Similarly, $p_2^{(\mu+1)} \le p_3^{(m)} < \frac{1}{6}$ for $\omega = m$ by m-condition 3, which completes the argument for (36).

Suppose first that $t_{21}^{(\mu)} = 1$, which together with $p_1^{(\mu-1)} > \frac{1}{3}$ implies $t_{22}^{(\mu+1)} = t_{32}^{(\mu+1)} = \frac{1}{2}$ according to (36). Obviously, $\frac{1}{2} < t_{12}^{(\mu)} + t_{13}^{(\mu)} < 2$ by the normalization of P. For $t_{12}^{(\mu)} + t_{13}^{(\mu)} = 1$, either $t_{12}^{(\mu)} = t_{33}^{(\mu)} = 1$ or $t_{32}^{(\mu)} = t_{13}^{(\mu)} = 1$ when P could be shortened without changing its function, or $t_{12}^{(\mu)} = t_{13}^{(\mu)} = t_{32}^{(\mu)} = t_{33}^{(\mu)} = \frac{1}{2}$ implying $p_1^{(\mu)} = p_2^{(\mu)} = p_3^{(\mu)} = \frac{1}{3}$ which contradicts (3). Hence, $t_{12}^{(\mu)} + t_{13}^{(\mu)} = \frac{3}{2}$. Denote $i \in \{2,3\}$ so that $t_{1i}^{(\mu)} = 1$ whereas $j \in \{2,3\}$ satisfies $t_{1j}^{(\mu)} = t_{3j}^{(\mu)} = \frac{1}{2}$. If $t_{13}^{(\mu+1)} = 1$, then we could shorten P while preserving its function, and hence $t_{23}^{(\mu+1)} > 0$ due to $t_{33}^{(\mu+1)} < 1$. It follows that $p_2^{(\mu+1)} \ge \frac{1}{2}p_1^{(\mu-1)} + \frac{1}{4}p_j^{(\mu-1)} = \frac{1}{4}(2p_1^{(\mu-1)} + p_j^{(\mu-1)}) = \frac{1}{4}(1 - p_i^{(\mu-1)} + p_1^{(\mu-1)}) \ge \frac{1}{4}$ which contradicts (36). Hence, $t_{11}^{(\mu)} = t_{21}^{(\mu)} = \frac{1}{2}$ due to $t_{11}^{(\mu)} < 1$ and $t_{31}^{(\mu)} = 0$, which implies $t_{12}^{(\mu)} < 1$ since $\mu = \mu'$.

In addition, there are no 'switching paths' (cf. Lemma 2.ii) starting simultaneously from all three vertices $v_1^{(\mu-1)}, v_2^{(\mu-1)}, v_3^{(\mu-1)}$ and leading to $v_1^{(\mu)}$ or $v_1^{(\mu+1)}$ since otherwise an h-neighbor $\mathbf{a}' \in \Omega_2(\mathbf{a}^{(m)}) \cap M(v_i^{(\mu-1)}) \subseteq H$ of $\mathbf{a}^{(m)} \in M(v_i^{(\mu-1)})$ from m-condition 4 would exist for some $i \in \{1,2,3\}$ such that $\mathbf{a}' \in M(v_1^{(\mu+1)})$ implying $P(\mathbf{a}') = 1$. Recall we still need to contradict $t_{12}^{(\mu+1)} = 0$, provided that $t_{11}^{(\mu)} = t_{21}^{(\mu)} = \frac{1}{2}$, $t_{12}^{(\mu)} < 1$, $t_{11}^{(\mu+1)} = 1$, $t_{22}^{(\mu+1)} > 0$, and $t_{33}^{(\mu+1)} < 1$.

We will first consider the case of $t_{12}^{(\mu)} = \frac{1}{2}$ which implies $t_{13}^{(\mu)} = 0$ since three parallel switching paths starting from level $\mu - 1$ are excluded. Suppose that $t_{33}^{(\mu)} > 0$ which also gives $t_{13}^{(\mu+1)} = 0$ because of ruling out the three switching paths, and hence $t_{23}^{(\mu+1)} > 0$ due to $t_{33}^{(\mu+1)} < 1$. In addition, we know $t_{22}^{(\mu)} + t_{32}^{(\mu)} = \frac{1}{2}$ since we assume $t_{12}^{(\mu)} = \frac{1}{2}$. It follows that $p_2^{(\mu+1)} \ge \frac{1}{4}p_1^{(\mu-1)} + \frac{1}{4}p_2^{(\mu-1)} + \frac{1}{4}p_3^{(\mu-1)} = \frac{1}{4}$ which contradicts (36). Hence, $t_{33}^{(\mu)} = 0$ implying $t_{23}^{(\mu)} = 1$ due to $t_{13}^{(\mu)} = 0$, which gives $t_{32}^{(\mu)} = \frac{1}{2}$. For $t_{23}^{(\mu+1)} > 0$, we would again get a contradiction $p_2^{(\mu+1)} \ge \frac{1}{4}p_1^{(\mu-1)} + \frac{1}{4}p_2^{(\mu-1)} + \frac{1}{2}p_3^{(\mu-1)} \ge \frac{1}{4}$, and hence we have $t_{23}^{(\mu+1)} = 0$ and $t_{13}^{(\mu+1)} > 0$ because of $t_{33}^{(\mu+1)} < 1$. We can assume without loss of generality[1] that $t_{13}^{(\mu+1)} = \frac{1}{2}$ since otherwise $t_{12}^{(\mu)} = t_{32}^{(\mu)} = \frac{1}{2}$ and $t_{13}^{(\mu+1)} = 1$ (implying $t_{22}^{(\mu+1)} = t_{32}^{(\mu+1)} = \frac{1}{2}$) could be replaced with $t_{12}^{(\mu)} = 1$ while $t_{23}^{(\mu)} = 1$ is replaced with $t_{22}^{(\mu)} = t_{32}^{(\mu)} = t_{23}^{(\mu)} = t_{33}^{(\mu)} = \frac{1}{2}$ and $t_{23}^{(\mu+1)} = t_{33}^{(\mu+1)} = \frac{1}{2}$ where $v_3^{(\mu)}$ is a copy of $v_2^{(\mu)}$, which redefines level μ. Thus, it follows from $t_{13}^{(\mu+1)} = \frac{1}{2}$ and $t_{23}^{(\mu+1)} = 0$ that $t_{33}^{(\mu+1)} = \frac{1}{2}$ and $t_{22}^{(\mu+1)} = 1$ by the normalization of P.

Recall once more we have $t_{11}^{(\mu)} = t_{21}^{(\mu)} = t_{12}^{(\mu)} = t_{32}^{(\mu)} = \frac{1}{2}$, $t_{23}^{(\mu)} = 1$, $t_{11}^{(\mu+1)} =$

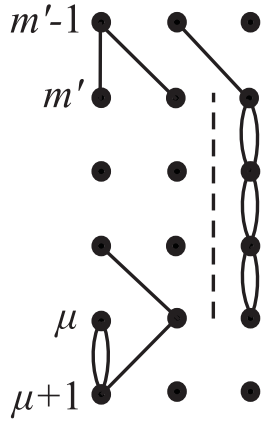

Figure 7: The block structure from m' to μ.

$t_{22}^{(\mu+1)} = 1$, and $t_{13}^{(\mu+1)} = t_{33}^{(\mu+1)} = \frac{1}{2}$. We know $p_1^{(\mu-1)} \leq 2p_2^{(\mu+1)} < \frac{1}{2}$ due to (36) and $p_2^{(\mu-1)} = 2p_3^{(\mu)} < \frac{1}{3}$ by (34), which implies $p_1^{(\mu+1)} = \frac{1}{2}p_1^{(\mu-1)} + \frac{3}{4}p_2^{(\mu-1)} < \frac{1}{2}$. This gives a contradiction $p_2^{(\mu+1)} \geq \frac{1}{2}(p_2^{(\mu+1)} + p_3^{(\mu+1)}) = \frac{1}{2}(1 - p_1^{(\mu+1)}) > \frac{1}{4}$ according to (36), which completes the argument for $t_{12}^{(\mu)} = \frac{1}{2}$.

Further consider the case of $t_{13}^{(\mu)} > 0$ which ensures $t_{12}^{(\mu)} = 0$ or equivalently $t_{22}^{(\mu)} + t_{32}^{(\mu)} = 1$. We know $t_{22}^{(\mu)} < 1$ by the normalization of P, and hence $t_{32}^{(\mu)} > 0$, which also ensures $t_{13}^{(\mu+1)} = 0$ since the three parallel switching paths starting from level $\mu - 1$ are excluded. It follows that $t_{23}^{(\mu+1)} > 0$ due to $t_{33}^{(\mu+1)} < 1$. Thus, we get a contradiction $p_2^{(\mu+1)} \geq \frac{1}{4}p_1^{(\mu-1)} + \frac{1}{2}p_2^{(\mu-1)} \geq \frac{1}{4}p_1^{(\mu-1)} + \frac{1}{4}p_2^{(\mu-1)} + \frac{1}{4}p_3^{(\mu-1)} = \frac{1}{4}$ according to (36).

Similarly, for the remaining case of $t_{12}^{(\mu)} = t_{13}^{(\mu)} = 0$ we obtain $t_{32}^{(\mu)} = t_{33}^{(\mu)} = 1$ by the normalization of P, which again ensures $t_{13}^{(\mu+1)} = 0$ implying $t_{23}^{(\mu+1)} > 0$. Hence, $p_2^{(\mu+1)} \geq \frac{1}{4}p_1^{(\mu-1)} + \frac{1}{2}p_2^{(\mu-1)} + \frac{1}{2}p_3^{(\mu-1)} \geq \frac{1}{4}$, which contradicts (36). This completes the proof of the lemma. □

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

We define the first **level 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 level m' through μ, which is illustrated in Figure 7 (where the dashed line shows that there is no edge from $v_1^{(k-1)}$ or $v_2^{(k-1)}$ to $v_3^{(k)}$ for any $m' < k \leq \mu$).

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

Proof. On the contrary, let k be the greatest level such that $m' < k \leq \mu$ and $t_{33}^{(k)} < 1$, 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 (34). 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 (34) gives

$$p_1^{(m')} + p_2^{(m')} = p_1^{(\mu)} + p_2^{(\mu)}, \tag{38}$$

$$p_3^{(m')} = p_3^{(\mu)} < \frac{1}{6} \tag{39}$$

which verifies m'-condition 3 for the first block level m'. Note that inequality (39) ensures $m' \geq 3$ due to $p_3^{(2)} \geq \frac{1}{4}$. 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 (39). For $t_{11}^{(m')} = 1$ or $t_{21}^{(m')} = 1$ we obtain $t_{12}^{(m')} + t_{22}^{(m')} > 0$ and $t_{13}^{(m')} + t_{23}^{(m')} > 0$ by the normalization of P. Thus either $\mathbf{a}^{(m)} \in M(v_1^{(m'-1)}) \subseteq M(v_1^{(m')}) \cup M(v_2^{(m')})$ 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 m 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 m'. 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 (16).

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 $t_{32}^{(m)} = 1$ (see Figure 5). It follows from Corollary 1 and Lemma 5

(see Figures 4 and 7, 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 4) and Lemma 4 that

$$p_1^{(\nu)} = p_1^{(\mu)} + p_2^{(\mu)} \left(1 - \frac{1}{2^{|R|}}\right) \qquad \text{for } \nu < m\,. \tag{40}$$

For $\nu = m$ when $\nu' = \nu - 1$, 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 $t_{32}^{(m)} = 1$ (see Figure 5) by the definition of ν, Lemma 2.ii, and m-conditions 1 and 2, which also ensures $\omega = m$ in both cases. Thus,

$$p_1^{(m)} + p_2^{(m)} = p_1^{(\mu)} + p_2^{(\mu)} \left(1 - \frac{1}{2^{|R|+1}}\right) \qquad \text{for } \nu = m = \omega \tag{41}$$

according to (23). 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[1].

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 (43). 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. (24). Then the double-edge path from the definition of ω passes through a double edge from $v \in \{v_2^{(\nu)}, v_3^{(\nu)}\}$ 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 $v_3^{(\nu+1)}$, respectively, as depicted in Figure 8. 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 4 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 $\nu + 1 \leq \ell \leq \min(\omega, m-1)$. It follows that

$$p_1^{(\omega+1)} \leq p_1^{(\nu)} + \frac{p_2^{(\mu)}}{2^{|R|+1}} \qquad \text{for } \omega < m \tag{42}$$

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

$$p_1^{(m)} + p_2^{(m)} \leq p_1^{(\nu)} + \frac{p_2^{(\mu)}}{2^{|R|+1}} \qquad \text{for } \omega = m \tag{43}$$

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

In addition, we will prove the following lemma:

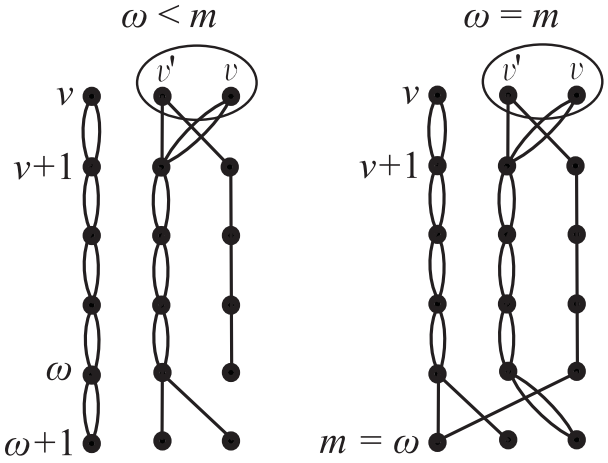

Figure 8: The block structure from $\nu < \omega$ to $\omega + 1$ (or to m if $\omega = m$).

Lemma 7.

$$p_1^{(\mu)} + p_2^{(\mu)} \leq 4p_2^{(\mu)} \,. \tag{44}$$

Proof. First consider the case of $\mu > m'$. Clearly, $t_{21}^{(\mu)} > 0$ follows from the definition of μ' for $\mu = \mu'$, while for $\mu = \mu' - 1$, the case of $t_{21}^{(\mu)} = 0$ translates to original $t_{11}^{(\mu)} = 0$ (before $v_1^{(\mu)}$ and $v_2^{(\mu)}$ were swapped) which contradicts the normalization of P by Lemma 5. Hence, we have $p_1^{(\mu)} + p_2^{(\mu)} = p_1^{(\mu-1)} + p_2^{(\mu-1)} \leq 2p_1^{(\mu-1)} \leq 4p_2^{(\mu)}$ according to Lemma 5. For $\mu = m'$, on the other hand, we will distinguish three cases. For $t_{32}^{(\mu)} = t_{33}^{(\mu)} = 1$, we know $p_1^{(\mu)} = p_2^{(\mu)}$ by Lemma 6, which implies (44). For $t_{12}^{(\mu)} + t_{22}^{(\mu)} = \frac{1}{2}$, we have $t_{33}^{(\mu)} = 1$ by Lemma 2.ii, which gives $p_1^{(\mu)} \leq \frac{1}{2}p_1^{(\mu-1)} + \frac{1}{2}p_2^{(\mu-1)} < \frac{1}{2} + \frac{1}{6} = \frac{2}{3}$ since $p_2^{(\mu-1)} < \frac{1}{3}$ by m'-conditions 2 and (39). In addition, $p_3^{(\mu-1)} < p_3^{(\mu)} < \frac{1}{6}$ implying $p_2^{(\mu-1)} + p_3^{(\mu-1)} < \frac{1}{2}$ which means $p_2^{(\mu)} \geq \frac{1}{2}p_1^{(\mu-1)} > \frac{1}{4}$ by Lemma 6. It follows that $p_1^{(\mu)} < \frac{2}{3} < \frac{3}{4} < 3p_2^{(\mu)}$ which gives (44). Similarly for $t_{13}^{(\mu)} + t_{23}^{(\mu)} > 0$, we have $t_{32}^{(\mu)} = 1$ implying $\frac{1}{6} > p_3^{(\mu)} \geq p_2^{(\mu-1)} \geq p_3^{(\mu-1)}$ due to (39), and hence $p_1^{(\mu-1)} > \frac{2}{3}$ which ensures $p_2^{(\mu)} > \frac{1}{3}$ by Lemma 6, while $p_1^{(\mu)} \leq \frac{1}{2}p_1^{(\mu-1)} + p_3^{(\mu-1)} < \frac{1}{2} + \frac{1}{6} = \frac{2}{3} < 1 < 3p_2^{(\mu)}$ which completes the proof of the lemma. □

For $\nu < m$, equation (40) is plugged into (42) if $\omega < m$ or into (43) if $\omega = m$, while equation (41) is considered for $\nu = m$ (implying $\omega = m$). Then Lemma 7 and equation (38) are employed, which results in

$$\begin{aligned} p_1^{(\omega+1)} &\leq p_1^{(\mu)} + p_2^{(\mu)} \left(1 - \frac{1}{2^{|R|}}\right) + \frac{p_2^{(\mu)}}{2^{|R|+1}} = p_1^{(\mu)} + p_2^{(\mu)} \left(1 - \frac{1}{2^{|R|+1}}\right) \\ &\leq \left(p_1^{(m')} + p_2^{(m')}\right) \left(1 - \frac{1}{2^{|R|+3}}\right) \qquad \text{for } \omega < m \,, \end{aligned} \tag{45}$$

$$p_1^{(m)} + p_2^{(m)} \le \left(p_1^{(m')} + p_2^{(m')}\right)\left(1 - \frac{1}{2^{|R|+3}}\right) \qquad \text{for } \omega = m\,. \tag{46}$$

Formula (45) can further be plugged into (33) giving

$$p_1^{(m)} + p_2^{(m)} \le 1 - \left(1 - \left(p_1^{(m')} + p_2^{(m')}\right)\left(1 - \frac{1}{2^{|R|+3}}\right)\right)\prod_{j=1}^{q}\left(1 - \frac{1}{2^{|Q_j|}}\right) \tag{47}$$

which is even valid for $\omega = m$ (i.e. $q = 0$) since equation (47) coincides with (46) 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 (see Figure 2). 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 sets $R, Q_1, \ldots, Q_q$, defined in Section 5, corresponding to the bth block are denoted as $R_b, Q_{b1}, \ldots, Q_{bq_b}$, respectively.

It follows from the definition of partition class in Paragraph 5.1 that, for any $b > 1$, the nodes labeled with the variables whose indices are in R_b are connected with the nodes corresponding to R_{b-1} through a computational path which traverses nodes $v_1^{(\nu'_b)}, v_1^{(\nu'_b+1)}, \ldots, v_1^{(m_b-1)}$ since $\nu'_b \le m_b - 1$ according to (23). Hence, sets $R_1, \ldots, R_r$ are pair-wise disjoint because P is read-once, and thus they create a partition.

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 level $m = m_r$ of the $(r + 1)$st block (see Paragraph 4.2 for $r = 0$), and let the assumption (34) be satisfied for the previous blocks, that is,

$$p_3^{(\mu_b)} < \frac{1}{6} \tag{48}$$

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

$$1 - \Pi_r < \delta = \min\left(\varepsilon - \varepsilon', \frac{6\varepsilon - 5}{7}\right) < \frac{1}{7} \tag{49}$$

where $\varepsilon > \frac{5}{6}$ and $\varepsilon' < \varepsilon$ are the parameters of Theorem 2 and denote

$$\Pi_k = \prod_{b=1}^{k} \pi_b\,, \qquad \pi_b = \prod_{j=1}^{q_b}\left(1 - \frac{1}{2^{|Q_{bj}|}}\right)\,, \tag{50}$$

$$\varrho_k = \prod_{b=1}^{k} \alpha_b \,, \qquad \alpha_b = \left(1 - \frac{1}{2^{|R_b|+3}}\right) \tag{51}$$

for $k = 1, \ldots, r$, $\varrho_0 = \Pi_0 = 1$, and $\pi_b = 1$ for $q_b = 0$. It follows from (50) and (49) that

$$\pi_b \geq \Pi_r > 1 - \delta > \frac{2}{3} \tag{52}$$

which verifies assumption (35) for every $b = 1, \ldots, r$. Hence, we can employ recursive inequality (47) from Section 6 which is rewritten as

$$p_{b-1} \leq 1 - (1 - p_b\alpha_b)\pi_b = 1 - \pi_b + p_b\alpha_b\pi_b \tag{53}$$

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

$$p_0 = p_1^{(d)} + p_2^{(d)} \geq \varepsilon \tag{54}$$

which follows from (20), recurrence (53) can be solved as

$$\begin{aligned} \varepsilon &\leq \sum_{k=1}^{r}(1 - \pi_k) \prod_{b=1}^{k-1} \alpha_b\pi_b + p_r \prod_{b=1}^{r} \alpha_b\pi_b < \sum_{k=1}^{r}(1 - \pi_k)\Pi_{k-1} + p_r\varrho_r\Pi_r \\ &= 1 - \Pi_r + p_r\varrho_r\Pi_r \,. \end{aligned} \tag{55}$$

In addition,

$$\varrho_r > p_r\varrho_r\Pi_r > \varepsilon - \delta \geq \varepsilon' \tag{56}$$

follows from (55) and (49).

7.2. Recursive Step

Throughout this paragraph, we will consider the case when

$$1 - \Pi_{r+1} < \delta \tag{57}$$

(cf. assumption (49)), while the case complementary to (57), which concludes the recursion, will be resolved below in Section 8. Assuming condition (57), we will prove that inductive assumptions (48) and (49) are met for r replaced with $r+1$ together with the four m_{r+1}-conditions for the first level m_{r+1} of the $(r+1)$st block so that we can further proceed in the recursion.

By analogy to (52), inequality (57) implies

$$\pi_{r+1} > 1 - \delta > \frac{2}{3} \,. \tag{58}$$

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

$$p_r \leq 1 - \left(p_2^{(\omega_{r+1}+1)} + p_3^{(\omega_{r+1}+1)}\right)\pi_{r+1} \tag{59}$$

according to (33), and

$$p_2^{(\omega_{r+1}+1)} + p_3^{(\omega_{r+1}+1)} \geq p_3^{(\mu_{r+1})} \tag{60}$$

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

$$\varepsilon < 1 - \Pi_r + \left(1 - p_3^{(\mu_{r+1})}\pi_{r+1}\right) \varrho_r \Pi_r \tag{61}$$

according to (55). Hence,

$$\varepsilon - \delta < \left(1 - p_3^{(\mu_{r+1})}\pi_{r+1}\right) \varrho_r \Pi_r < 1 - p_3^{(\mu_{r+1})}\pi_{r+1} \tag{62}$$

follows from (49), which gives

$$p_3^{(\mu_{r+1})} < \frac{1 - \varepsilon + \delta}{1 - \delta} \leq \frac{1}{6} \qquad \text{for } \omega_{r+1} < m_r \tag{63}$$

by (58) and (49). Inequality (63) is even valid for $\omega_{r+1} = m_r$ since

$$p_3^{(\mu_{r+1})} \leq p_3^{(m_r)} < \frac{1}{6} \qquad \text{for } \omega_{r+1} = m_r \tag{64}$$

according to m_r-condition 3. Therefore, assumptions (34) and (35) of the analysis in Section 6 are also met for the $(r+1)$st block according to (63)–(64) and (58), respectively, which justifies recurrence inequality (53) for $b = r+1$ leading to the solution

$$\varepsilon < 1 - \Pi_{r+1} + p_{r+1}\varrho_{r+1}\Pi_{r+1} \tag{65}$$

by analogy to (55) where r is replaced with $r+1$. Similarly to (56), we obtain

$$\varrho_{r+1} > \varepsilon' \tag{66}$$

by combining (65) with (57). Thus, inductive assumptions (48) and (49) are valid for r replaced by $r+1$ according to (63)–(64) and (57), 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 level m_{r+1}. In Paragraph 6.2, m_{r+1}-conditions 1–3 have been proven, and thus, it suffices to validate m_{r+1}-condition 4. For this purpose, we exploit the fact that A is ε'^{11}-rich after we show corresponding condition (16) for partition $\{R_1, \ldots, R_{r+1}\}$ of $I = \bigcup_{b=1}^{r+1} R_b$. In particular,

$$\varepsilon'^{11} < \varrho_{r+1}^{11} < \prod_{b=1}^{r+1} \left(1 - \frac{1}{2^{|R_b|}}\right) \tag{67}$$

follows from (66) since for any $1 \leq b \leq r+1$,

$$\left(1 - \frac{1}{2^{|R_b|+3}}\right)^{11} < 1 - \frac{1}{2^{|R_b|}} \tag{68}$$

for $|R_b| \geq 1$ because $f(x) = \ln(1 - \frac{1}{x}) / \ln(1 - \frac{1}{8x})$ is a decreasing function for $x = 2^{|R_b|} \geq 2$, 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 $a_i^{(m_{r+1})} \neq c_i$ according to (17) for $Q = \emptyset$. 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 4), 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 (57) is broken.

8. The End of Recursion

In this section, we will consider the case of

$$1 - \Pi_{r+1} \geq \delta \tag{69}$$

complementary to (57), which concludes the recursion from Section 7 as follows. Suppose $|Q_{bj}| > \log n$ for every $b = 1, \ldots, r+1$ and $j = 1, \ldots, q_b$, then we would have

$$\Pi_{r+1} = \prod_{b=1}^{r+1} \prod_{j=1}^{q_b} \left(1 - \frac{1}{2^{|Q_{bj}|}}\right) \geq \left(1 - \frac{1}{2^{\log n}}\right)^{\frac{n}{\log n}} \tag{70}$$

$$> 1 - \frac{1}{n} \cdot \frac{n}{\log n} = 1 - \frac{1}{\log n}\,, \tag{71}$$

which breaks (69) for sufficiently large n. Hence, there must be $1 \leq b^* \leq r+1$ and $1 \leq j^* \leq q_{b^*}$ such that $|Q_{b^*j^*}| \leq \log n$, and we denote $Q = Q_{b^*j^*}$. Clearly, $Q \cap R_b = \emptyset$ for $b = 1, \ldots, b^* - 2$ due to P is read-once while it may happen that $Q \cap R_{b^*-1} \neq \emptyset$ for $j^* = 1$, $\kappa_{b^*1} = \sigma_{b^*1} = m_{b^*-1}$, and $t_{23}^{(m_{b^*-1})} = 0$. Thus, let r^* be the maximum of b^*-2 and b^*-1 such that $Q \cap R_{r^*} = \emptyset$. We will again employ the fact that A is ε'^{11}-rich. First condition (16) for partition $\{R_1, \ldots, R_{r^*}\}$ of $I = \bigcup_{b=1}^{r^*} R_b$ is verified as

$$\prod_{b=1}^{r^*} \left(1 - \frac{1}{2^{|R_b|}}\right) > \varrho_r^{11} > \varepsilon'^{11} \tag{72}$$

according to (68) and (56). This provides $\mathbf{a}^* \in A$ such that $a_i^* = c_i^Q$ for every $i \in Q$ and at the same time, for every $b = 1, \ldots, r^*$ there exists $i \in R_b$ that meets $a_i^* \neq c_i^{R_b}$ according to (17).

Lemma 8. *Denote $\lambda = \lambda_{b^*j^*}$. There are two generalized 'switching' paths (cf. Lemma 2.ii) starting from $v_2^{(k)}$ and $v_3^{(k)}$, respectively, at level k satisfying $3 < \max(\lambda - 2, \mu_{b^*}) \leq k < \lambda$, which may 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 8 gives an h-neighbor $\mathbf{a}' \in \Omega_2(\mathbf{a}^*) \subseteq H$ of $\mathbf{a}^* \in A$ such that $\mathbf{a}' \in M(v_1^{(\lambda)}) \cup M(v_3^{(\lambda)})$. Thus, either $\mathbf{a}' \in M(v_1^{(\lambda)}) \subseteq M(v_1^{(m_{b^*-1})}) \cup M(v_2^{(m_{b^*-1})})$ or $\mathbf{a}' \in M(v_3^{(\lambda)})$ which implies $\mathbf{a}' \in M(v_1^{(\kappa_{b^* j^*})}) \subseteq M(v_1^{(m_{b^*-1})}) \cup M(v_2^{(m_{b^*-1})})$ since $a'_i = a^*_i = c_i^Q$ for every $i \in Q$ according to (17) (see Figure 2 and 6). Note that $M(v_1^{(\kappa_{b^* j^*})}) = M(v_1^{(m_{b^*-1})})$ for $r^* = b^* - 2$. Hence, $P(\mathbf{a}') = 1$ because for every $b = 1, \ldots, r^*$ there exists $i \in R_b$ that meets $a'_i = a^*_i \neq c_i^{R_b}$ due to (17) (see Figure 2 and 4). 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 distribution on S is almost uniform, i.e. the probability that a given $\mathbf{c} \in \{0,1\}^n$ coincides with the strings from $\mathcal{A}_n = \mathcal{A} \cap \{0,1\}^n$ on

the bit locations from S can be approximated as

$$\left| \frac{\left|\mathcal{A}_n^S(\mathbf{c})\right|}{|\mathcal{A}_n|} - \frac{1}{2^{|S|}} \right| \leq \beta\,, \tag{73}$$

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, for suitable k, any almost k-wise independent set is ε-rich. It follows that almost $O(\log n)$-wise independent sets are hitting sets for the class of read-once conjunctions of DNF and CNF (cf. [7]).

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. Assume $\{R_1, \ldots, R_r\}$ is a partition of index set $I \subseteq \{1, \ldots, n\}$ satisfying condition (16), and $Q \subseteq \{1, \ldots, n\} \setminus I$ such that $|Q| \leq \log n$. In order to show for a given $\mathbf{c} \in \{0,1\}^n$ that there is $\mathbf{a} \in \mathcal{A}_n$ that meets (17) for Q and partition $\{R_1, \ldots, R_r\}$, we will prove that the probability

$$p = p(\mathcal{A}_n) = \frac{\left|\mathcal{A}_n^Q(\mathbf{c}) \setminus \bigcup_{j=1}^r \mathcal{A}_n^{R_j}(\mathbf{c})\right|}{|\mathcal{A}_n|} \tag{74}$$

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*.

The main idea of the proof lies in lower bounding the probability (74). By using the assumption that $\mathcal{A}$ is almost $O(\log n)$-wise independent this probability can be approximated by the probability that any $\mathbf{a} \in \{0,1\}^n$ (not necessarily in $\mathcal{A}_n$) satisfies (17) which can be expressed and lower bounded as

$$p(\{0,1\}^n) = \frac{1}{2^{|Q|}} \prod_{j=1}^{r} \left(1 - \frac{1}{2^{|R_j|}}\right) \geq \frac{\varepsilon}{n} > 0 \tag{75}$$

according to (16) and $|Q| \leq \log n$. In particular, we briefly comment on the main steps of the proof which are schematically depicted in Figure 9 including references to corresponding sections, lemmas, and equations. In Section 10, we will first modify the partition classes R_j so that their cardinalities are at most logarithmic whereas the classes of small constant cardinalities are merged with Q and also $\mathbf{c}$ is adjusted correspondingly. Lemma 9 then ensures that the probability p from (74) is lower bounded when using these modified classes. Furthermore, Bonferroni inequality (the inclusion-exclusion principle) and the assumption concerning the almost k-wise independence are employed in Section 11 where also the classes of the same cardinality are grouped. In Section 12, we will further reduce the underlying lower bound on p only to a sum over frequent cardinalities of partition classes to which Taylor's theorem is applied in Section 13, whereas a corresponding Lagrange remainder is bounded using the assumption on constant C.

Modifications of Partition Classes (Section 10)

- superlogarithmic cardinalities:
  $R'_j \subseteq R_j$ so that $|R'_j| \leq \log n$ (76)
- small constant cardinalities:
  $R_{\leq} = \bigcup_{|R'_j| \leq \sigma} R'_j$ where σ is a constant (82) & (85)
  $\longrightarrow Q' = Q \cup R_{\leq}$ (89), $c'_i = 1 - c_i$ for $i \in R_{\leq}$ (92)

Lemma 9: $p \geq \frac{\left|\mathcal{A}_n^{Q'}(\mathbf{c}') \setminus \bigcup_{j=1}^{r'} \mathcal{A}_n^{R'_j}(\mathbf{c}')\right|}{|\mathcal{A}_n|}$ (93)

↓ **Bonferroni inequality**

$$p \geq \sum_{k=0}^{C'} (-1)^k \sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \frac{\left|\mathcal{A}_n^{\bigcup_{i=1}^{k} R'_{j_i} \cup Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|} \quad (96)$$

↓ **almost k-wise independence** (Section 11)

$$p \geq \frac{1}{2^{|Q'|}} \left(\sum_{k=0}^{C'} (-1)^k \sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \prod_{i=1}^{k} \frac{1}{2^{|R'_{j_i}|}} - \frac{\varepsilon'}{8} \right) \quad (99) \text{ \& } (100)$$

Grouping the Same Cardinalities (Lemma 10)

$\sigma < s_1, \ldots, s_{m'} \leq \log n$ ... cardinalities of R'_j
$r_i = \left|\left\{ j \,|, |R'_j| = s_i \right\}\right|$ ... the number of classes of cardinality s_i

$$p > \frac{1}{n^2} \left(\sum_{k=0}^{C'} (-1)^k \sum_{\substack{k_1 + \cdots + k_{m'} = k \\ 0 \leq k_1 \leq r_1, \ldots, 0 \leq k_{m'} \leq r_{m'}}} \prod_{i=1}^{m'} \frac{t_i^{k_i}}{k_i!} \prod_{j=1}^{k_i - 1} \left(1 - \frac{j}{r_i}\right) - \frac{\varepsilon'}{8} \right) \quad (104)$$

where $t_i = \frac{r_i}{2^{s_i}}$ (78)

Frequent Cardinalities (Section 12 & Lemma 11)

$r_1 > r_2 > \cdots > r_{m''} > \varrho$ where ϱ is a constant (81) & (83)

$$p > \frac{1}{n^2} \left(\sum_{k=0}^{C'} (-1)^k \sum_{\substack{k_1 + \cdots + k_{m''} = k \\ k_1 \geq 0, \ldots, k_{m''} \geq 0}} \prod_{i=1}^{m''} \frac{t_i^{k_i}}{k_i!} - \frac{\varepsilon'}{2} \right) \quad (117) \text{ \& Lemma 12.i}$$

↓ **multinomial theorem**

$$p > \frac{1}{n^2} \left(\sum_{k=0}^{C'} \frac{\left(-\sum_{i=1}^{m''} t_i\right)^k}{k!} - \frac{\varepsilon'}{2} \right) \quad (119)$$

↓ **Taylor's theorem**

$$p > \frac{1}{n^2} \left(e^{-\sum_{i=1}^{m''} t_i} - \mathcal{R}_{C'+1}\left(-\sum_{i=1}^{m''} t_i\right) - \frac{\varepsilon'}{2} \right) \quad (120)$$

↓ (16) $\longrightarrow \sum_{i=1}^{m} t_i < \ln \frac{1}{\varepsilon'}$ (80)
↓ Lagrange remainder $\mathcal{R}_{C'+1}\left(-\sum_{i=1}^{m''} t_i\right) < \frac{\varepsilon'}{4}$ (Lemma 12.ii)

$p > \frac{\varepsilon'}{4n^2} > 0$ (127)

Figure 9: The main steps of the proof of Theorem 3

10. Modifications of Partition Classes

We properly modify the underlying partition classes in order to further upper bound their cardinalities by the logarithmic function so that the assumption concerning almost $\lceil (C+2)\log n \rceil$-wise independence of $\mathcal{A}$ can be applied in the following Section 11. Thus, we confine ourselves to at most logarithmic-size arbitrary subsets R'_j of partition classes R_j, that is

$$R'_j \begin{cases} = R_j & \text{if } |R_j| \leq \log n \\ \subset R_j \text{ so that } |R'_j| = \lfloor \log n \rfloor & \text{otherwise}\,, \end{cases} \tag{76}$$

which ensures $R'_j \subseteq R_j$ and $|R'_j| \leq \log n$ for every $j = 1, \ldots, r$. For these new classes, assumption (16) can be rewritten as

$$\begin{aligned} \prod_{j=1}^{r} \left(1 - \frac{1}{2^{|R'_j|}}\right) &> \left(1 - \frac{1}{2^{\log n}}\right)^{\frac{n}{\log n}} \prod_{|R_j| \leq \log n} \left(1 - \frac{1}{2^{|R_j|}}\right) \\ &> \left(1 - \frac{1}{n} \cdot \frac{n}{\log n}\right) \varepsilon = \left(1 - \frac{1}{\log n}\right) \varepsilon = \varepsilon'\,, \end{aligned} \tag{77}$$

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

Denote by $\{s_1, s_2, \ldots, s_m\} = \{|R'_1|, \ldots, |R'_r|\}$ the set of all cardinalities $1 \leq s_i \leq \log n$ of classes $R'_1, \ldots, R'_r$, and for every $i = 1, \ldots, m$, let $r_i = |\{j \,|\, |R'_j| = s_i\}|$ be the number of classes R'_j having cardinality s_i, that is, $r = \sum_{i=1}^{m} r_i$. Furthermore, we define

$$t_i = \frac{r_i}{2^{s_i}} > 0 \quad \text{for } i = 1, \ldots, m\,. \tag{78}$$

It follows from (77) and (78) that

$$\begin{aligned} 0 < \varepsilon' &< \prod_{j=1}^{r} \left(1 - \frac{1}{2^{|R'_j|}}\right) = \prod_{i=1}^{m} \left(1 - \frac{1}{2^{s_i}}\right)^{r_i} \\ &= \prod_{i=1}^{m} \left(\left(1 - \frac{1}{2^{s_i}}\right)^{2^{s_i}}\right)^{t_i} < e^{-\sum_{i=1}^{m} t_i} \end{aligned} \tag{79}$$

implying

$$\sum_{i=1}^{m} t_i < \ln \frac{1}{\varepsilon'}\,. \tag{80}$$

Moreover, we define constants

$$\varrho = \frac{C}{1 - \left(1 - \frac{\varepsilon'^2}{4(1+\varepsilon'^2)}\right)^{\frac{1}{C}}} > C \geq 1\,, \tag{81}$$

$$\sigma = \log\left(\frac{4\varrho\,(1+\varepsilon'^2)}{\varepsilon'^2}\right) \tag{82}$$

which are used for sorting the cardinalities $s_1, \ldots, s_m$ so that

$$r_i > \varrho \text{ and } s_i > \sigma \quad \text{for} \quad i = 1, \ldots, m'' \tag{83}$$

$$r_i \leq \varrho \text{ and } s_i > \sigma \quad \text{for} \quad i = m'' + 1, \ldots, m' \tag{84}$$

$$s_i \leq \sigma \quad \text{for} \quad i = m' + 1, \ldots, m\,. \tag{85}$$

We will further confine ourselves to the first $m' \geq 0$ cardinalities satisfying $s_i > \sigma$ for $i = 1, \ldots, m'$. Without loss of generality, we can also sort the corresponding partition classes so that

$$|R'_j| > \sigma \quad \text{for} \quad j = 1, \ldots, r' \tag{86}$$

$$|R'_j| \leq \sigma \quad \text{for} \quad j = r' + 1, \ldots, r\,, \tag{87}$$

which implies

$$r' = \sum_{i=1}^{m'} r_i = \sum_{i=1}^{m'} t_i 2^{s_i} > \frac{4\varrho\,(1+\varepsilon'^2)}{\varepsilon'^2} \sum_{i=1}^{m'} t_i \tag{88}$$

according to (78), (83)–(84), and (82). We include the remaining constant-size classes R'_j for $j = r' + 1, \ldots, r$ into Q, that is,

$$Q' = Q \cup \bigcup_{j=r'+1}^{r} R'_j \tag{89}$$

whose size can be upper bounded as

$$|Q'| \leq \log n + \sum_{i=m'+1}^{m} r_i \log\left(\frac{4\varrho\,(1+\varepsilon'^2)}{\varepsilon'^2}\right) < 2\log n \tag{90}$$

for sufficiently large n, since

$$\sum_{i=m'+1}^{m} r_i = \sum_{i=m'+1}^{m} t_i 2^{s_i} < \frac{4\varrho\,(1+\varepsilon'^2)}{\varepsilon'^2}\,\ln\frac{1}{\varepsilon'} \tag{91}$$

according to (78), (80), (85), and (82). This completes the definition of new classes $Q', R'_1, \ldots, R'_{r'}$. In addition, 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.

$$c'_i = \begin{cases} 1 - c_i & \text{if } i \in \bigcup_{j=r'+1}^{r} R'_j \\ c_i & \text{otherwise.} \end{cases} \tag{92}$$

The modified $Q', R'_1, \ldots, R'_{r'}$ and $\mathbf{c}'$ are used in the following lemma for lower bounding the probability (74).

Lemma 9.

$$p \geq \frac{\left|\mathcal{A}_n^{Q'}(\mathbf{c}') \setminus \bigcup_{j=1}^{r'} \mathcal{A}_n^{R'_j}(\mathbf{c}')\right|}{|\mathcal{A}_n|} = \frac{\left|\mathcal{A}_n^{Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|} - \frac{\left|\bigcup_{j=1}^{r'} \mathcal{A}_n^{R'_j \cup Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|}\,. \tag{93}$$

PROOF. For verifying the lower bound in (93) it suffices to show that

$$\mathcal{A}_n^{Q'}(\mathbf{c}') \setminus \bigcup_{j=1}^{r'} \mathcal{A}_n^{R'_j}(\mathbf{c}') \subseteq \mathcal{A}_n^{Q}(\mathbf{c}) \setminus \bigcup_{j=1}^{r} \mathcal{A}_n^{R_j}(\mathbf{c}) \tag{94}$$

according to (74). Assume $\mathbf{a} \in \mathcal{A}_n^{Q'}(\mathbf{c}') \setminus \bigcup_{j=1}^{r'} \mathcal{A}_n^{R'_j}(\mathbf{c}')$, 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 (76), (89), (92), 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 (92), and hence, $\mathbf{a} \in \mathcal{A}_n^{Q}(\mathbf{c}) \setminus \bigcup_{j=1}^{r} \mathcal{A}_n^{R_j}(\mathbf{c})$. This completes the proof of the lower bound, while the equality in (93) follows from $\mathcal{A}_n^{R'_j \cup Q'}(\mathbf{c}') \subseteq \mathcal{A}_n^{Q'}(\mathbf{c}')$ for every $j = 1, \ldots, r'$. □

11. Almost k-Wise Independence

Furthermore, we will upper bound the probability of the finite union of events appearing in formula (93) by using Bonferroni inequality for constant number $C' = \min(C, r')$ of terms, which gives

$$p \geq \frac{\left|\mathcal{A}_n^{Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|} - \sum_{k=1}^{C'} (-1)^{k+1} \sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \frac{\left|\bigcap_{i=1}^{k} \mathcal{A}_n^{R'_{j_i} \cup Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|} \tag{95}$$

$$= \sum_{k=0}^{C'} (-1)^{k} \sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \frac{\left|\mathcal{A}_n^{\bigcup_{i=1}^{k} R'_{j_i} \cup Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|} \tag{96}$$

according to Lemma 9. For notational simplicity, the inner sum in (96) over $1 \leq j_1 < j_2 < \cdots < j_k \leq r'$ for $k = 0$ reads formally as it includes one summand $|\mathcal{A}_n^{Q'}(\mathbf{c}')|/|\mathcal{A}_n|$. Note that C' is odd for $C < r'$, while equality holds in (95) 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 (76) and (90), and hence,

$$\frac{\left|\mathcal{A}_n^{\bigcup_{i=1}^{k} R'_{j_i} \cup Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|} \geq \frac{1}{2^{|Q'| + \sum_{i=1}^{k} \left|R'_{j_i}\right|}} - \beta = \frac{1}{2^{|Q'|}} \prod_{i=1}^{k} \frac{1}{2^{\left|R'_{j_i}\right|}} - \beta \tag{97}$$

(where the product in (97) equals formally 1 for $k = 0$) and similarly,

$$-\frac{\left|\mathcal{A}_n^{\bigcup_{i=1}^{k} R'_{j_i} \cup Q'}(\mathbf{c}')\right|}{|\mathcal{A}_n|} \geq -\frac{1}{2^{|Q'|}} \prod_{i=1}^{k} \frac{1}{2^{\left|R'_{j_i}\right|}} - \beta \tag{98}$$

according to (73) since $\mathcal{A}$ is $(\lceil (C+2)\log n \rceil, \beta)$-wise independent. We plug these inequalities into (96), which leads to

$$\begin{aligned} p &\geq \sum_{k=0}^{C'} (-1)^k \sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \frac{1}{2^{|Q'|}} \prod_{i=1}^{k} \frac{1}{2^{\left|R'_{j_i}\right|}} - \beta \sum_{k=0}^{C'} \binom{r'}{k} \\ &\geq \frac{1}{2^{|Q'|}} \left(\sum_{k=0}^{C'} (-1)^k \sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \prod_{i=1}^{k} \frac{1}{2^{\left|R'_{j_i}\right|}} - \beta\, 2^{|Q'|}\, (r'+1)^{C'} \right), \quad (99) \end{aligned}$$

where

$$\beta\, 2^{|Q'|}\, (r'+1)^{C'} < \frac{1}{n^{C+3}}\, n^2\, n^C = \frac{1}{n} < \frac{\varepsilon'}{8} \quad (100)$$

for sufficiently large $n > 8/\varepsilon'$ by using the assumption on β, inequality (90), $r' < n$ (e.g., $r' = n$ would break (86) and (82)), and $C' \leq C$. The following lemma rewrites the inner sum in formula (99).

Lemma 10. *For* $0 \leq k \leq C'$,

$$\sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \prod_{i=1}^{k} \frac{1}{2^{|R'_{j_i}|}} = \sum_{\substack{k_1 + \cdots + k_{m'} = k \\ 0 \leq k_1 \leq r_1, \ldots, 0 \leq k_{m'} \leq r_{m'}}} \prod_{i=1}^{m'} \frac{t_i^{k_i}}{k_i!} \prod_{j=1}^{k_i - 1} \left(1 - \frac{j}{r_i}\right). \quad (101)$$

Proof. By grouping the classes of the same cardinality together, the left-hand side of inequality (101) can be rewritten as

$$\sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq r'} \prod_{i=1}^{k} \frac{1}{2^{|R'_{j_i}|}} = \sum_{\substack{k_1 + k_2 + \cdots + k_{m'} = k \\ 0 \leq k_1 \leq r_1, \ldots, 0 \leq k_{m'} \leq r_{m'}}} \prod_{i=1}^{m'} \binom{r_i}{k_i} \left(\frac{1}{2^{s_i}}\right)^{k_i}, \quad (102)$$

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

$$\binom{r_i}{k_i} \left(\frac{1}{2^{s_i}}\right)^{k_i} = \frac{r_i\,(r_i - 1) \cdots (r_i - k_i + 1)}{k_i!} \left(\frac{t_i}{r_i}\right)^{k_i} = \frac{t_i^{k_i}}{k_i!} \prod_{j=1}^{k_i - 1} \left(1 - \frac{j}{r_i}\right) \quad (103)$$

according to (78). □

Thus, we plug equations (100) and (101) into (99) and obtain

$$p > \frac{1}{n^2} \left(\sum_{k=0}^{C'} (-1)^k \sum_{\substack{k_1 + \cdots + k_{m'} = k \\ 0 \leq k_1 \leq r_1, \ldots, 0 \leq k_{m'} \leq r_{m'}}} \prod_{i=1}^{m'} \frac{t_i^{k_i}}{k_i!} \prod_{j=1}^{k_i - 1} \left(1 - \frac{j}{r_i}\right) - \frac{\varepsilon'}{8} \right). \quad (104)$$

Note that for $m' = 0$ (implying $r' = C' = 0$), the inner sum in (104) equals 1.

12. Frequent Cardinalities

We sort out the terms with frequent cardinalities (83) from the sum in formula (104), that is,

$$p > \frac{1}{n^2}\left(\sum_{k=0}^{C'}(-1)^k \sum_{\substack{k_1+\cdots+k_{m''}=k\\ 0\le k_1\le r_1,\ldots,0\le k_{m''}\le r_{m''}}} \prod_{i=1}^{m''}\frac{t_i^{k_i}}{k_i!}\prod_{j=1}^{k_i-1}\left(1-\frac{j}{r_i}\right) - T_1 - \frac{\varepsilon'}{8}\right), \tag{105}$$

where the inner sum in (105) equals zero for $k > r'' = \sum_{i=1}^{m''} r_i$, and

$$T_1 = \sum_{k=0}^{C'}(-1)^{k+1} \sum_{\substack{k_1+\cdots+k_{m'}=k\\ 0\le k_1\le r_1,\ldots,0\le k_{m'}\le r_{m'}\\ (\exists\, m''+1\le \ell\le m')\, k_\ell>0}} \prod_{i=1}^{m'}\frac{t_i^{k_i}}{k_i!}\prod_{j=1}^{k_i-1}\left(1-\frac{j}{r_i}\right) \tag{106}$$

sums up the terms including rare cardinalities (84). In addition, we know

$$1 \ge \prod_{i=1}^{m''}\prod_{j=1}^{k_i-1}\left(1-\frac{j}{r_i}\right) \quad \ge \tag{107}$$

$$\prod_{i=1}^{m''}\left(1-\frac{C-1}{\varrho}\right)^{k_i-1} \quad > \quad \left(1-\frac{C}{\varrho}\right)^{C} = 1-\frac{\varepsilon'^2}{4(1+\varepsilon'^2)} \tag{108}$$

according to (83), (81), and $k_i \le k = \sum_{i=1}^{m''} k_i \le C' \le C < \varrho$. The upper bound (107) and lower bound (108) on the underlying product are used to lower bound the negative terms of (105) for odd k and the positive ones for even k, respectively, that is,

$$p > \frac{1}{n^2}\left(\sum_{k=0}^{C'}(-1)^k \sum_{\substack{k_1+\cdots+k_{m''}=k\\ 0\le k_1\le r_1,\ldots,0\le k_{m''}\le r_{m''}}} \prod_{i=1}^{m''}\frac{t_i^{k_i}}{k_i!} - \frac{\varepsilon'^2}{4(1+\varepsilon'^2)}\,T_2 - T_1 - \frac{\varepsilon'}{8}\right) \tag{109}$$

where

$$T_2 = \sum_{k=0,2,4,\ldots}^{C'} \sum_{\substack{k_1+\cdots+k_{m''}=k\\ 0\le k_1\le r_1,\ldots,0\le k_{m''}\le r_{m''}}} \prod_{i=1}^{m''}\frac{t_i^{k_i}}{k_i!}\,. \tag{110}$$

The following lemma upper bounds the above-introduced terms T_1 and T_2.

Lemma 11.

(i) $T_1 < \frac{\varepsilon'}{8}$.

(ii) $T_2 < \frac{1+\varepsilon'^2}{2\,\varepsilon'}$.

PROOF.

(i) We can only take the terms of (106) for odd $k = 1, 3, 5, \ldots$ into account since those for even k are nonpositive (e.g. the term for $k = 0$ equals zero because there is no $m'' + 1 \leq \ell \leq m'$ such that $k_\ell > 0$ in this case). Thus,

$$\begin{aligned} T_1 &\leq \sum_{k=1,3,5,\ldots}^{C'} \sum_{\substack{k_1+\cdots+k_{m'}=k \\ 0\leq k_1\leq r_1,\ldots,0\leq k_{m'}\leq r_{m'} \\ (\exists\, m''+1\leq\ell\leq m')\, k_\ell>0}} \frac{r_\ell}{2^{s_\ell}} \frac{1}{k_\ell} \frac{t_\ell^{k_\ell-1}}{(k_\ell-1)!} \prod_{\substack{i=1 \\ i\neq\ell}}^{m'} \frac{t_i^{k_i}}{k_i!} \\ &\leq \frac{\varrho}{2^\sigma} \sum_{k=1,3,5,\ldots}^{C'} \sum_{\substack{k_1+\cdots+k_{m'}=k \\ 0\leq k_1\leq r_1,\ldots,0\leq k_{m'}\leq r_{m'} \\ (\exists\, m''+1\leq\ell\leq m')\, k_\ell>0}} \frac{t_\ell^{k_\ell-1}}{(k_\ell-1)!} \prod_{\substack{i=1 \\ i\neq\ell}}^{m'} \frac{t_i^{k_i}}{k_i!} \end{aligned} \tag{111}$$

according to (78) and (84). Formula (111) is rewritten by replacing indices $k_\ell - 1$ and $k-1$ with k_ℓ and k, respectively, which is further upper bounded by removing the upper bounds that are set on indices $k_1, \ldots, k_{m'}$ and by omitting the condition concerning the existence of special index ℓ, as follows:

$$T_1 \leq \frac{\varrho}{2^\sigma} \sum_{k=0,2,4,\ldots}^{C'-1} \sum_{\substack{k_1+\cdots+k_{m'}=k \\ k_1\geq 0,\ldots,k_{m'}\geq 0}} \prod_{i=1}^{m'} \frac{t_i^{k_i}}{k_i!} = \frac{\varrho}{2^\sigma} \sum_{k=0,2,4,\ldots}^{C'-1} \frac{\left(\sum_{i=1}^{m'} t_i\right)^k}{k!}, \tag{112}$$

where the multinomial theorem is employed. Notice that the sum on the right-hand side of equation (112) represents the first few terms of Taylor series of the hyperbolic cosine at point $\sum_{i=1}^{m'} t_i \geq 0$, which implies

$$T_1 < \frac{\varrho}{2^\sigma} \cosh\left(\sum_{i=1}^{m'} t_i\right) < \frac{\varepsilon'^2}{4(1+\varepsilon'^2)} \cdot \frac{\frac{1}{\varepsilon'} + \varepsilon'}{2} = \frac{\varepsilon'}{8} \tag{113}$$

according to (80) and (82) since the hyperbolic cosine is an increasing function for nonnegative arguments.

(ii) Similarly as in the proof of (i), we apply the multinomial theorem (cf. (112)) and the Taylor series of the hyperbolic cosine (cf. (113)) to (110), which gives

$$T_2 \leq \sum_{k=0,2,4,\ldots}^{C'} \sum_{\substack{k_1+\cdots+k_{m''}=k \\ k_1\geq 0,\ldots,k_{m''}\geq 0}} \prod_{i=1}^{m''} \frac{t_i^{k_i}}{k_i!} = \sum_{k=0,2,4,\ldots}^{C'} \frac{\left(\sum_{i=1}^{m''} t_i\right)^k}{k!} \tag{114}$$

$$\leq \cosh\left(\sum_{i=1}^{m''} t_i\right) < \frac{1+\varepsilon'^2}{2\,\varepsilon'}. \tag{115}$$

□

We plug the bounds from Lemma 11 into (109) and obtain

$$p > \frac{1}{n^2}\left(\sum_{k=0}^{C'}(-1)^k \sum_{\substack{k_1+\cdots+k_{m''}=k \\ 0\le k_1\le r_1,\ldots,0\le k_{m''}\le r_{m''}}} \prod_{i=1}^{m''}\frac{t_i^{k_i}}{k_i!} - \frac{3\,\varepsilon'}{8}\right). \tag{116}$$

13. Taylor's Theorem

In order to apply the multinomial theorem again, we remove the upper bounds that are set on indices in the inner sum of formula (116), that is,

$$p > \frac{1}{n^2}\left(\sum_{k=0}^{C'}(-1)^k \sum_{\substack{k_1+\cdots+k_{m''}=k \\ k_1\ge 0,\ldots,k_{m''}\ge 0}} \prod_{i=1}^{m''}\frac{t_i^{k_i}}{k_i!} - T - \frac{3\,\varepsilon'}{8}\right), \tag{117}$$

which is corrected by introducing additional term

$$T = \sum_{k=0}^{C'}(-1)^k \sum_{\substack{k_1+\cdots+k_{m''}=k \\ k_1\ge 0,\ldots,k_{m''}\ge 0 \\ (\exists 1\le \ell\le m'')\, k_\ell > r_\ell}} \prod_{i=1}^{m''}\frac{t_i^{k_i}}{k_i!}\,. \tag{118}$$

Thus, inequality (117) can be further rewritten as

$$p > \frac{1}{n^2}\left(\sum_{k=0}^{C'}\frac{\left(-\sum_{i=1}^{m''}t_i\right)^k}{k!} - T - \frac{3\,\varepsilon'}{8}\right) \tag{119}$$

$$= \frac{1}{n^2}\left(e^{-\sum_{i=1}^{m''}t_i} - \mathcal{R}_{C'+1}\left(-\sum_{i=1}^{m''}t_i\right) - T - \frac{3\,\varepsilon'}{8}\right), \tag{120}$$

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

$$\mathcal{R}_{C'+1}\left(-\sum_{i=1}^{m''}t_i\right) = \frac{\left(-\sum_{i=1}^{m''}t_i\right)^{C'+1}}{(C'+1)!}\,e^{-\vartheta\sum_{i=1}^{m''}t_i} < \left(\frac{\sum_{i=1}^{m''}t_i}{\sqrt{C'}}\right)^{C'+1} \tag{121}$$

with parameter $0 < \vartheta < 1$. Note that the upper bound in (121) assumes $C' > 0$, whereas for $C' = r' = 0$ implying $m'' = m' = 0$, we know $\mathcal{R}_1(0) = 0$. This remainder and term T are upper bounded in the following lemma.

Lemma 12.

(i) $T < \frac{\varepsilon'}{8}$.

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

PROOF.

(i) We take only the summands of (118) 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

$$T \leq \sum_{k=2,4,6,\ldots}^{C'} \sum_{\substack{k_1+\cdots+k_{m''}=k \\ k_1\geq 0,\ldots,k_{m''}\geq 0 \\ (\exists 1\leq \ell\leq m'')\, k_\ell > r_\ell}} \frac{1}{2^{s_\ell}} \frac{r_\ell}{k_\ell} \frac{t_\ell^{k_\ell-1}}{(k_\ell-1)!} \prod_{\substack{i=1 \\ i\neq\ell}}^{m''} \frac{t_i^{k_i}}{k_i!}$$

$$\leq \frac{1}{2^\sigma} \sum_{k=2,4,6,\ldots}^{C'} \sum_{\substack{k_1+\cdots+k_{m''}=k \\ k_1\geq 0,\ldots,k_{m''}\geq 0 \\ (\exists 1\leq \ell\leq m'')\, k_\ell > r_\ell}} \frac{t_\ell^{k_\ell-1}}{(k_\ell-1)!} \prod_{\substack{i=1 \\ i\neq\ell}}^{m''} \frac{t_i^{k_i}}{k_i!} \tag{122}$$

using (78) and (83). Formula (122) 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 concerning the existence of special index ℓ, as follows:

$$T \leq \frac{1}{2^\sigma} \sum_{k=1,3,5,\ldots}^{C'-1} \sum_{\substack{k_1+\cdots+k_{m''}=k \\ k_1\geq 0,\ldots,k_{m''}\geq 0}} \prod_{i=1}^{m''} \frac{t_i^{k_i}}{k_i!} = \frac{1}{2^\sigma} \sum_{k=1,3,5,\ldots}^{C'-1} \frac{\left(\sum_{i=1}^{m''} t_i\right)^k}{k!}, \tag{123}$$

where the multinomial theorem is employed. Notice that the sum on the right-hand side of equation (123) represents the first few terms of Taylor series of the hyperbolic sine at point $\sum_{i=1}^{m''} t_i$, which implies

$$T \leq \frac{1}{2^\sigma} \sinh\left(\sum_{i=1}^{m''} t_i\right) < \frac{\varepsilon'^2}{4\varrho\,(1+\varepsilon'^2)} \cdot \frac{\frac{1}{\varepsilon'} - \varepsilon'}{2} < \frac{\varepsilon'}{8} \tag{124}$$

according to (80) and (82) since the hyperbolic sine is an increasing function.

(ii) For $C' = C \geq 1$, Lagrange remainder (121) can further be upper bounded as

$$\mathcal{R}_{C'+1}\left(-\sum_{i=1}^{m''} t_i\right) < \left(\frac{\ln\frac{1}{\varepsilon'}}{\sqrt{C}}\right)^{C+1} < \left(\frac{\varepsilon'}{2}\right)^{C+1} < \frac{\varepsilon'}{4} \tag{125}$$

for sufficiently large n by using (80) and the definition of C, while for $C' = r' < C$, the underlying upper bound

$$\mathcal{R}_{C'+1}\left(-\sum_{i=1}^{m''} t_i\right) \leq \left(\frac{\sum_{i=1}^{m'} t_i}{\frac{4\varrho\,(1+\varepsilon'^2)}{\varepsilon'^2}}\right)^{\frac{r'+1}{2}} < \frac{\ln\frac{1}{\varepsilon'}}{\frac{4\varrho\,(1+\varepsilon'^2)}{\varepsilon'^2}} < \frac{\varepsilon'}{4} \tag{126}$$

can be obtained from (88) and (80). □

Finally, inequality (79) together with the upper bounds from Lemma 12 are plugged into (120), which leads to

$$p > \frac{\varepsilon'}{4n^2} = \frac{\varepsilon}{4n^2}\left(1 - \frac{1}{\log n}\right) > 0 \tag{127}$$

according to (77). 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 (17) for Q and partition $\{R_1, \ldots, R_r\}$ is strictly positive, which means such $\mathbf{a}$ does exist. This completes the proof that $\mathcal{A}$ is ε-rich. □

14. 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. regular oblivious read-once branching programs). We have now provided an explicit polynomial-time construction of a hitting set for read-once branching programs of width 3 with acceptance probability $\varepsilon > \frac{5}{6}$. Although this model seems to be relatively weak, the presented proof is far from being trivial. In particular, we have formulated a so-called richness condition which is independent of the notion of branching programs. This condition characterizes the hitting sets for read-once branching programs of width 3. We have shown that such a hitting set hits read-once conjunctions of DNF and CNF, which corresponds to the weak richness condition. On the other hand, the richness condition proves to be sufficient for a set extended with all strings within Hamming distance of 3 to be a hitting set for width-3 1-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. It also follows that almost $O(\log n)$-wise independent sets are hitting sets for read-once conjunctions of DNF and CNF.

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.

Acknowledgements

The authors would like to thank Pavel Pudlák for pointing out the problem of hitting sets for width-3 1-branching programs. J.Š.'s research was partially supported by project GA ČR P202/12/G061 and RVO: 67985807. S.Ž.'s research was partially supported by project GA ČR P202/10/1333 and RVO: 67985807.

References

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

[2] P. Beame, W. Machmouchi, Making branching programs oblivious requires superpolynomial overhead, in: Proceedings of the CCC 2011 Twenty-Sixth Annual IEEE Conference on Computational Complexity, 2011, pp. 12–22.

[3] A. Bogdanov, Z. Dvir, E. Verbin, A. Yehudayoff, Pseudorandomness for width 2 branching programs, Electron. Colloq. Computat. Complex. (ECCC) TR09-070 (2009).

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

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

[6] A., De, Pseudorandomness for permutation and regular branching programs, in: Proceedings of the CCC 2011 Twenty-Sixth Annual IEEE Conference on Computational Complexity, 2011, pp. 221–231.

[7] A. De, O. Etesami, L. Trevisan, M. Tulsiani, Improved pseudorandom generators for depth 2 circuits, in: Proceedings of the RANDOM 2010 Fourteenth International Workshop on Randomization and Computation, in: LNCS, vol. 6302, Springer-Verlag, 2010, pp. 504–517.

[8] B. Fefferman, R. Shaltiel, Ch. Umans, E. Viola, On beating the hybrid argument, in: Proceedings of the ITCS 2012 Third ACM Conference on Innovations in Theoretical Computer Science, 2012, pp. 468–483.

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

[10] M. Koucký, P. Nimbhorkar, P. Pudlák, Pseudorandom generators for group products, in: Proceedings of the STOC 2011 Forty-Third ACM Symposium on Theory of Computing, ACM Press, 2011, pp. 263–272.

[11] R. Meka, D. Zuckerman, Pseudorandom generators for polynomial threshold functions, in: Proceedings of the STOC 2010 Forty-Second ACM Symposium on Theory of Computing, ACM Press, 2010, pp. 427–436.

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

[13] N. Nisan, A. Wigderson, Hardness vs. randomness, J. Comput. Syst. Sci. 49 (2) (1994) 149–167.

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

[15] J. Šíma, S. Žák, Almost k-wise independent sets establish hitting sets for width-3 1-branching programs, in: Proceedings of the CSR 2011 6th International Computer Science Symposium in Russia, in: LNCS, vol. 6651, Springer-Verlag, 2011, pp. 120–133.

[16] J. Šíma, S. Žák, A sufficient condition for sets hitting the class of read-once branching programs of width 3, in: Proceedings of the SOFSEM 2012 Thirty-Eighth International Conference on Current Trends in Theory and Practice of Informatics, in: LNCS, Springer-Verlag, 2012 (to appear).

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