Ordered Binary Decision Diagrams and the Shannon Effect

Clemens Gröpl [1],

Freie Universität Berlin, Institut für Informatik, Takustraße 9, D-14195 Berlin, Germany

Hans Jürgen Prömel

Humboldt-Universität zu Berlin, Institut für Informatik, Unter den Linden 6, D-10099 Berlin, Germany

Anand Srivastav [2]

Christian-Albrechts-Universität zu Kiel, Ludewig-Meyn-Str. 4, Mathematisches Seminar II, D-24098 Kiel, Germany

Abstract

We investigate the size and structure of ordered binary decision diagrams (OBDDs) for random Boolean functions. It was known that for most values of n, the expected OBDD size of a random Boolean function with n variables is equal to the worst-case size up to terms of lower order. Such a phenomenon is generally called strong Shannon effect. Here we show that the strong Shannon effect is not valid for all n. Instead it undergoes a certain periodic 'phase transition': If n lies within intervals of constant width around the values $n = 2^h + h$, then the strong Shannon effect does not hold, whereas it does hold outside these intervals. Our analysis provides doubly exponential probability bounds and generalises to ordered Kronecker functional decision diagrams (OKFDDs).

Key words: Binary decision diagram, Boolean function, probabilistic analysis, Shannon effect.

[1] Corresponding author, email: groepl@inf.fu-berlin.de. C. Gröpl was supported by Deutsche Forschungsgemeinschaft, grants GRK 219/2-97 (graduate school "Algorithmische Diskrete Mathematik") and Pr 296/6-3.

[2] A. Srivastav was supported by Deutsche Forschungsgemeinschaft, grant Pr 296/3-2.

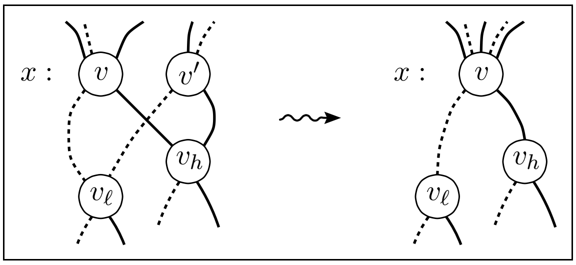

Fig. 1. Merging v and v'

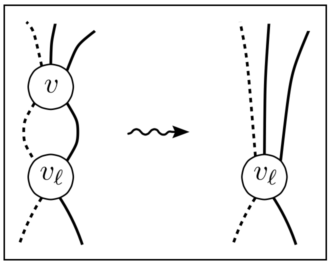

Fig. 2. Deleting v

1 Introduction

A Boolean function is a mapping $f\colon \{0,1\}^n \to \{0,1\}$ depending on Boolean variables. Efficient representation and manipulation of Boolean functions is an important issue in applications, e. g. formal verification. The state-of-the-art data structure for Boolean functions are *ordered binary decision diagrams*, abbreviated OBDD (see Bryant's articles [6,8]).

An OBDD is an acyclic directed graph with one root and two terminals 0 and 1. Every non-terminal node is labelled with a variable and has two outgoing edges pointing to succesor nodes. Which one of them is 'active' depends on the value of the variable. To compute the value of the function, one starts at the root node and follows the unique path of active edges to a terminal. Each OBDD has a global ordering in which the variables are tested. A *level* is the set of all nodes labeled with a given variable. An OBDD is called *reduced*, if it contains no nodes representing the same subfunction. The reduced ordered binary decision diagram of a Boolean function is uniquely determined for each variable ordering. It can be obtained (or even defined) as the result of the application of two reduction rules to its binary decision tree. The *merging rule* allows to identify nodes that test the same variable and have coinciding successor nodes (Fig. 1). By the *deletion rule*, nodes with both outgoing edges pointing to the same successor can be removed, because the function they represent does not depend on the variable that is about to be tested (Fig. 2). The *quasi-reduced* ordered binary decision diagram (qOBDD) results from applying (only) the merging rule to its binary decision tree. It will play an important role in our analysis. The quasi-reduced OBDD of a Boolean function is also uniquely determined for each variable ordering, and from the qOBDD of a function one can obtain its OBDD by applying the deletion rule. We consider reduced OBDDs unless stated otherwise.

Many tasks can be performed efficiently with OBDD, including equivalence checking, satisfiability test, satisfiability count, synthesis (computation of $f \otimes g$ from f and g, where $\otimes$ is a Boolean operation), replacement of variables by functions (substitution), redundancy test ("does $f(x_1,\ldots,x_n)$ depend on x_i?"), and more. Efficient implementation are e. g. [5,27]. The size of the OBDD for a given function generally depends on the choice of the variable ordering. The optimal variable ordering problem is hard to approximate up to constant factors [26]. For heuristics see e. g. [25]. While there exist Boolean functions having exponential OBDD size for all variable orderings (like the 'middle' bit of integer multiplication [7]), the importance of OBDDs stems from the fact that many functions encountered in practice have moderately sized OBDDs at least for some variable orderings.

This paper is devoted to theoretical investigations on the size of the OBDD of a Boolean function which has been chosen according to the uniform distribution. Using counting arguments, Liaw and Lin [22] showed that the expected optimal OBDD size in this model is never more than a constant factor apart from the worst-case OBDD size of any function on n variables. Later Wegener [30] showed that the two quantities in fact coincide up to terms of lower order, if the number of variables is uniformly distributed in $[2^h \,..\, 2^{h+1} - 1]$ and h goes to infinity. [3] The phenomenon that almost all functions have the same OBDD size as the hardest functions up to a factor of $1 + o(1)$ is called the *strong Shannon effect* for random Boolean functions and OBDDs. [4] If almost all functions have the same OBDD size up to a factor of $1 + o(1)$, which may be smaller than the worst-case size, then the *weak Shannon effect* holds. Wegener proved that the weak Shannon effect holds for all n. He also showed that the strong Shannon effect holds for 'most' n, in the sense just described.

These results led to the conjecture [30] that the strong Shannon effect also holds for all n. But this turns out not to be the case. We will uncover a periodic 'phase transition' in the development of the gap between the expected and the worst-case size of the OBDD of a random Boolean function. The strong Shannon effect does not hold within intervals of constant width around the values $n = 2^h + h$, where $h \in \mathbb{N}$, but it does hold outside these intervals.

Theorem 1.1 (Main Theorem)
Denote the minimal size of an OBDD for a Boolean function f by $Z_*(f)$ and the worst-case OBDD size by W. Let

$$B \;:=\; \bigcup_{h\in\mathbb{N}} \Big[2^h + h - d(h) \,..\, 2^h + h + d(h)\Big]\,,$$

where d is specified in (i) and (ii) below, and set $A := \mathbb{N} \setminus B$.

(i) If $n \to \infty$ in such a way that $n \in A$ for some $d(h) \to \infty$, then

$$\Pr\Big(Z_* = \big(1 - o(1)\big)W\Big) \;\sim\; 1\,,$$

i. e., the strong Shannon effect holds for the minimal OBDD size of random Boolean functions.

[3] Intervals of integers are denoted as $[a \,..\, b] := [a, b] \cap \mathbb{Z}$ and $[a] := [1 \,..\, a]$.
[4] In this paper, o- and O-terms are unsigned, whereas ω-, Ω-, and Θ-terms are nonnegative.

(ii) If $n \to \infty$ in such a way that $n \in B$ for some $d(h) = O(1)$, then

$$\Pr\Big(Z_* = \big(1 - \Omega(1)\big)W\Big) \sim 1\,,$$

and the strong Shannon effect does not *hold for the minimal OBDD size of random Boolean functions.*

The same conclusions hold for the minimal qOBDD size.

In the proof, we first investigate the case of a fixed variable ordering and then generalise to arbitrary (including the optimal) variable orderings. We pursue a refinement of Wegener's urn experiment approach. His key observation was that the size of each OBDD level is given by the classical urn occupancy experiment, whose expectation and variance are well-known [20]. He then handled the case of arbitrary variable orderings using the method of second moments. Instead, we will invoke specialised large deviation inequalities which yield stronger estimates and allow some simplifications in the proof. We cannot apply the results of Kolchin, Sevast'i͡anov and Chisti͡akov [20] on the limit distributions of the urn occupancy experiment, because they make no assertion about the convergency rate. Instead, we use a large deviation inequality that follows from Azuma's martingale inequality, and invoke Chvátal's bound on the hypergeometric distribution. The resulting probability bound is doubly exponential in n, which improves upon the exponential bound of [30]. Another methodical innovation of our approach is the use of a functional equation to locate a certain 'critical level' in the OBDD.

The new approach has the advantage that it enables us to extend Main Theorem 1.1 to a generalisation of OBDDs called ordered Kronecker functional decision diagram (OKFDD) [13]. OKFDDs include several other (earlier) extensions of the OBDD data structure as special cases. Such a result apparently cannot be obtained by the second moment method.

2 Preliminaries

Throughout this section we consider a random Boolean function f together with an arbitrary, but fixed variable ordering. Without loss of generality, we will assume this ordering to be $(x_1, \ldots, x_n)$. We denote the qOBDD for f with respect to the canonical variable ordering by qOBDD(f). The nodes on level i of qOBDD(f) represent the different subfunctions of f that can be obtained by substituting the first $i-1$ variables $x_i, \ldots, x_{i-1}$ by constants $c_1, \ldots, c_{i-1}$. In OBDDs, a node is present only if the subfunction really depends on the variable tested there. Let Y_i resp. Z_i denote the number of nodes on level i of qOBDD(f) resp. OBDD(f); so $Y_i \geqslant Z_i$. Upper bounds on the level sizes in qOBDDs follow from the growth rate of the decision tree, $k_i := 2^{i-1}$, and from the number of Boolean functions

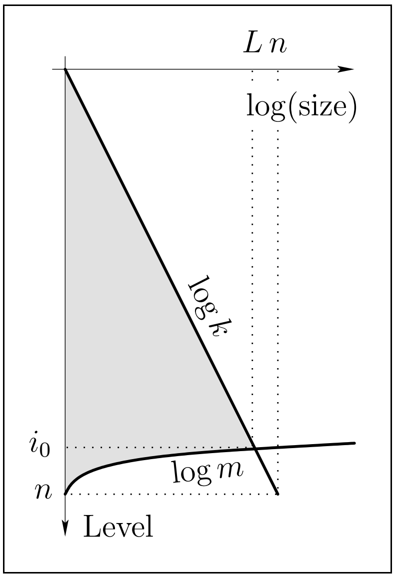

Fig. 3. The worst-case shape of a qOBDD

with $n - i + 1$ variables, $m_i := 2^{2^{n-i+1}}$ (see Fig. 3). For reduced OBDDs, we set $m'_i := m_i - m_{i+1}$. Clearly, $Y_i \leqslant w_i := \min\{k_i, m_i\}$ and $Z_i \leqslant w'_i := \min\{k_i, m'_i\}$. These bounds are tight, as was shown in [18, Lemma 2] by a direct construction. So the worst-case size of the whole qOBDD is $W(n) = \sum_{i=1}^{n} w_i$. Liaw and Lin [22] showed that $W(n) = \Theta(2^n/n)$ (see also [18]). We review results on the worst-case size (as far as they are applied in our analysis) in Section 3. More details about how $W(n)/(2^n/n)$ oscillates between certain critical parametrizations of n can be found in [17,15].

Recall that Y_i was defined as the number of nodes at the i-th level of qOBDD(f). It turns out that Y_i is almost always almost equal to Z_i, the corresponding number for OBDD(f). Thus, the merging rule alone is already sufficient to reduce the diagram size from 2^n (the size of the decision tree) to some value below $W(n) = \Theta(2^n/n)$, whereas the effect of the deletion rule is comparatively negligible for a random Boolean function. This fact was observed by Liaw and Lin [22] and rigorously proved by Wegener [30]. The consequence for our analysis is that we may focus on quasireduced OBDDs instead of reduced OBDDs.

Since we want to decide when the strong Shannon effect holds, what we are really interested in is $X_i := w_i - Y_i$, the number of nodes that are 'missing' at the i-th level of the qOBDD compared with the worst-case size w_i. We put $X := \sum_i X_i$, $Y := \sum_i Y_i$ and let $\mathrm{E}(X)$ resp. $\mathrm{E}(Y)$ denote the expectations. Then the strong Shannon effect for OBDDs with respect to a fixed variable ordering holds if and only if $\mathrm{E}(X)/W \to 0$ as $n \to +\infty$. We estimate the worst-case size $W(n)$ and the expectation $\mathrm{E}(X)$ separately. The difficult part is the analysis of $\mathrm{E}(X)$.

Wegener showed that random OBDDs are essentially worst-case shaped whenever n is such that for all levels i either $k_i \ll m_i$ or $k_i \gg m_i$ holds. His result is essentially Part (i) of our Theorem 1.1 (but here we determine the range of n for which the condition holds). In order to prove Part (ii), we must deal with those n for which an i exists such that k_i and m_i are of (about) the same size. We also take some care to let both parts (ranges of n) be complementary to each other.

We need a concise notation for the *critical point* i where the upper bounds k_i and m_i meet [5]. We define the function L by the functional equation

$$L(n) + \log L(n) = n \tag{2.1}$$

and set

$$i_\delta := L(n) + \delta + 1\,. \tag{2.2}$$

Then by definition,

$$k_{i_\delta} = 2^{\delta+L} \quad \text{and} \quad m_{i_\delta} = 2^{2^{-\delta}L}\,. \tag{2.3}$$

With this notation, the critical point is precisely at $i = i_0$, and $w_{i_0} = k_{i_0} = m_{i_0} = 2^{L(n)}$. But note that i_0 is not an integer in general. It only marks the borderline where the worst-case level width turns from growing exponentially to shrinking doubly exponentially (see Fig. 3). What we really have to deal with is the *critical level* (if it exists), which has the form $i_\delta \in \mathbb{N}$, where $|\delta|$ must be sufficiently small. (See Definition 4.3.) By definition of the function L, we have

$$w_i = \begin{cases} k_i\,, & i \leqslant i_0\,;\\ m_i\,, & i \geqslant i_0\,. \end{cases} \tag{2.4}$$

There is no closed formula for L, whereas by the defining functional equation (2.1), the inverse function of L is simply $L^{-1}(i) = i + \log i$. As an alternative definition, one can also obtain L as a pointwise limit of a sequence of functions L_r, $r \in \mathbb{N}_0$, defined by $L_{-1}(n) := 1$ and $L_{r+1}(n) := n - \log L_r(n)$. It is elementary to show that the pointwise limit $L(n) := \lim_{r\to\infty} L_r(n)$ exists iff $n \geqslant 1$. In fact, the inequalities

$$L_{2r}(n) > L_{2r+2}(n) > L(n) > L_{2r+3}(n) > L_{2r+1}(n) \tag{2.5}$$

are valid for $r \in \mathbb{N}_0$ and $n > 1$. The first approximations are

$$\begin{aligned} L_0(n) &= n\,,\\ L_1(n) &= n - \log n\,,\\ L_2(n) &= n - \log(n - \log n)\,. \end{aligned} \tag{2.6}$$

From these one can easily show that $L(n) \sim n$ and $2^{L(n)} \sim 2^n/n$. Also, $L_1(n) - L(n) = o(1)$, since $L_1(n) - L_2(n) = o(1)$. The inequalities (2.5) will be applied occasionally, but the asymptotics are more important.

[5] We consider m_i instead of m_i' for technical reasons, although our interest is mainly in reduced OBDDs.

Working with L simplifies the calculations. Earlier investigations dealt with L_1 in some way [30] or approximated L by other means [22]. In [18], $\lceil i_0 \rceil$ was implicit in the form of $\min\{i \in \mathbb{N} \mid k_i \geqslant m_i\}$, but no asymptotic for $\lceil i_0 \rceil - i_0$ was given. We decided to use L itself and apply the functional equation (2.1) in the calculations.

The following inequalities are valid for $x \leqslant 0$.

$$\left.\begin{array}{r} 1+x \\ 1+x+\dfrac{x^2}{2}+\dfrac{x^3}{6} \end{array}\right\} \leqslant e^x \leqslant 1+x+\frac{x^2}{2}\,. \tag{2.7}$$

3 Worst-Case Sizes

This section is devoted to worst-case bounds for $W(n)$ and $W'(n)$. Let us refer to the ratios $W/2^L$ and $W'/2^L$ as the *relative worst-case size* [6] . A detailed analysis of the *oscillation* of the relative worst-case size was given in [17,15]; here we only repeat the main result. In view of the oscillations, the global bounds are tight, and for every particular sequence of n, the relative worst-case size is known up to a factor of $1 + o(1)$. Our global lower and upper bounds slightly improve upon those of Liaw and Lin [22], Heap and Mercer [18], and Breitbart, Hunt, and Rosenkrantz [4]. However, our main results can be shown using only $W'(n) \sim W(n) = \Theta\Big(2^{L(n)}\Big)$, and this was already known from [22].

Theorem 3.1 ([17,15]**)**

(i) *For $n \to +\infty$,*

$$1 \;<\; \frac{W(n)}{2^{L(n)}} \;\leqslant\; 2 + O\left(2^{-L(n)/2}\right) \;=\; 2 + O\left(\sqrt{\frac{n}{2^n}}\right) .$$

Furthermore, if $n \in \mathbb{N} \setminus \{2^h + h \mid h \in \mathbb{N}\}$ is large enough, then the upper bound can be improved to $W/2^L < 2$.

(ii) *For large enough n,*

$$1 \;<\; \frac{W'(n)}{2^{L(n)}} \;<\; 2\,,$$

and both inequalities are asymptotically tight.

Essentially, the lower bound from [4] is $1 + o(1)$, and the upper bound from [22] is $2 + o(1)$.

We remark that oscillations do *not* occur for the relative worst-case size of general branching programs (where variables may be tested arbitrarily often, not neces-

[6] Remember that $2^L \sim 2^n/n$.

sarily respecting a global ordering) [4]. The results of [15] imply that the relative worst-case size of read-once branching programs oscillates in a similar way as that of OBDDs. (In read-once branching programs, variables can be tested at most once, but in arbitrary order.)

4 Expected qOBDD Size for a Fixed Variable Ordering

In this section, we determine the expected size of the levels of the qOBDD with a fixed variable ordering for a random Boolean function. In a typical qOBDD, not all levels are as large as in the worst-case examples. We defined $X_i = w_i - Y_i$ as the number of 'missing' nodes at the i-th level of a qOBDD. For each level, the expected value $\mathrm{E}(X_i)$ can be computed by the following urn experiment [20,30], a variation of the classical urn occupancy experiment.

Think of the subfunctions that result from constant assignments to the first $i-1$ variables as *balls* and of the possible subfunctions at level i as *urns*. Then at level i, there are k_i balls being thrown into m_i urns, so $\mathrm{E}(Y_i)$ is the expected number of non-empty urns.

Proposition 4.1

(i)

$$\mathrm{E}(Y_i) = \sum_{j\in[m_i]} \mathrm{Pr}(\text{ \textit{j-th urn is non-empty} }) = m_i\,(1-q_i)\,,$$

where

$$q_i := \mathrm{Pr}(\text{ \textit{first urn is empty} }) = \left(1-\frac{1}{m_i}\right)^{k_i}\,.$$

(ii)

$$\mathrm{E}(X_i) = \begin{cases} k_i - m_i\,(1-q_i)\,, & i \leqslant i_0\,; \\ m_i\,q_i\,, & i \geqslant i_0\,. \end{cases}$$

The proof is straightforward. For (ii), use (2.4).

As already mentioned, we have to be especially careful about those n for which an i exists such that k_i and m_i are of roughly the same size, because then $q_i \neq o(1)$ and $q_i \neq 1+o(1)$ (and we are in Part (ii) of the Main Theorem). These *critical levels* are what we have to look at in this section. The existence of critical levels is the reason why the strong Shannon effect breaks down occasionally. We say that a level j is *critical* if $\mathrm{E}(X_j)$ is bigger than a constant fraction of W. In our analysis, we show that such a j necessarily has the form $j = i_\delta \in \mathbb{N}$ for some sufficiently small $|\delta|$. Of course, for $|\delta| \to +\infty$, no level i_δ can be critical, since then $w_j = o(W)$ by (2.4). But the actual threshold for δ which we will determine in what follows is much smaller.

We cannot compute $\mathrm{E}(X_j)/W$ directly in order to decide whether the strong Shannon effect holds. Instead we look at the ratio $\mathrm{E}(X_j)/2^L$ and use the results from Section 3. We also show that there is at most *one* critical level.

The main technical contribution of this section is Theorem 4.8, which leads to an 'intermediate' result about the expected qOBDD size with respect to a fixed variable ordering of random Boolean functions, Theorem 4.12. The restriction to a fixed variable ordering can be dropped when we apply strong large deviation bounds. This is done for the merging rule in Section 5 and for the deletion rule in Section 6. The effect of the deletion rule is comparatively negligible.

Before we delve into the technicalities, we explain the idea of the proof in an easy special case.

Proposition 4.2 *For $n = 2^h + h$ and $h \nearrow +\infty$, $\mathrm{E}(X)/W = \Omega(1)$, so the strong Shannon effect does not hold. More precisely,*

$$\liminf_{\substack{n=2^h+h \\ h\nearrow+\infty}} \frac{\mathrm{E}\left(X_{i_0}\right)}{2^L} \geqslant \frac{1}{e}, \quad \textit{and} \quad \liminf_{\substack{n=2^h+h \\ h\nearrow+\infty}} \frac{\mathrm{E}(X)}{W} \geqslant \frac{1}{2e}.$$

Proof. Recall that $L(2^h+h) = 2^h$, which implies that the *critical point* $i_0 = 2^h+1 \in \mathbb{N}$ is also a *level*. Observe that $w_{i_0} = k_{i_0} = m_{i_0} = 2^L = 2^{2^h}$. By Proposition 4.1, we have

$$\mathrm{E}\left(X_{i_0}\right) \;=\; m_{i_0}\left(1 - \frac{1}{m_{i_0}}\right)^{k_{i_0}} \;=\; 2^{2^h}\left(1 - \frac{1}{2^{2^h}}\right)^{2^{2^h}} \;\sim\; \frac{2^{2^h}}{e},$$

and by (2.4),

$$W\Big(2^h + h\Big) \;=\; \sum_{i=1}^{i_0} k_i + \sum_{i=i_1}^{n} m_i \;=\; 2 \cdot 2^{2^h} - 1 + o\Big(2^{2^h}\Big) \;\sim\; 2 \cdot 2^{2^h}. \qquad \square$$

The tightness of these bounds will be shown in Theorem 4.8. Some prerequisites are developed in the next section.

4.1 Prerequisites

The proof of Proposition 4.2 was easy because $L\Big(2^h+h\Big) = 2^h$ is an integer. Things get more complicated for arbitrary n. We will show that taking more than a constant number of steps away from the 'bad' values $n = 2^h + h$ is enough to guarantee that the strong Shannon effect holds, and that any constant number is not sufficient.

Stated in another way, the difficulty that arises in the proof for general n is that the *critical point* i_0 does no longer coincide with the *critical level* of the qOBDD, which in fact can be $\lfloor i_0 \rfloor$ or $\lceil i_0 \rceil$, depending on n. Therefore, we introduce a parameter $\delta'(n) \in \mathbb{R}$ such that $i_{\delta'(n)}$ will be the critical level, if there is any.

Definition 4.3 *Let $\delta'(n)$ denote the difference between $L(n)$ and the next natural number, i. e.*

$$\delta'(n) := \ell - L(n)\,,$$

where $\ell \in \mathbb{N}$ is the unique element of $\left]L(n) - \frac{1}{2},\, L(n) + \frac{1}{2}\right] \cap \mathbb{N}$, and write

$$i' := i_{\delta'}\,.$$

By definition, $i_{\delta'(n)} = i_0 + \delta'(n) = \ell + 1$ is the integer nearest to i_0. In case of a tie, we round up, so $\delta'(n) \in \left]\frac{-1}{2}, \frac{1}{2}\right]$.

Obviously, $\delta'(2^h + h) = 0$. This is the case we considered in Proposition 4.2. As $|\delta'|$ gets larger, $\mathrm{E}\Big(X_{i_{\delta'}}\Big)$ becomes negligible compared to 2^L very soon, and hence there is always at most one critical level. We need to find out how large δ' is, depending on n. Therefore, we introduce two other parameters $h(n)$ and $a(n)$ such that $n = 2^{h(n)} + h(n) + a(n)$. Then a is closely related to δ'.

There is one little complication with this approach. As n grows from $2^{h'} + h'$ to $2^{h'+1} + h' + 1$, the parameter $\delta'(n)$ first goes up from 0 to about $1/2$, then jumps down to about $-1/2$ and finally becomes 0 again. (This is because $L(n)$ grows slightly slower than n). In the following definition we force that the jumps of $a(n)$ are at the same positions as those of $\delta'(n)$ by the requirement that a and δ' have the same sign. It is not necessary for us to determine the exact position of the jumps.

Definition 4.4 *For $n \in \mathbb{N}$, we define $h(n) \in \mathbb{N}$ and $a(n) \in \mathbb{Z}$ by the requirements that (i) $n = 2^{h(n)} + h(n) + a(n)$, (ii) $a(n) \cdot \delta'(n) \geqslant 0$, and (iii) $\Big|a(n)\Big|$ is minimal under conditions (i) and (ii).*

It is easy to check that the numbers $h(n)$ and $a(n)$ are well-defined. Next we note some immediate consequences of Definitions 4.3 and 4.4.

Proposition 4.5

(i) If $\delta'(n) > \delta'(n+1)$, then $n \sim \sqrt{2}\,2^{h'}$ for some $h' \in \mathbb{N}$.

(ii) If $(h'_t)_{t\in\mathbb{N}}$ and $(a'_t)_{t\in\mathbb{N}}$ are sequences such that $a'_t = o\Big(2^{h'_t}\Big)$ as $t \to +\infty$ and $n_t := 2^{h'_t} + h'_t + a'_t$, then $h(n_t) = h'_t$ and $a(n_t) = a'_t$.

(iii) For large n, we have $\Big|a(n)\Big| \leqslant 0.42 \cdot 2^{h(n)}$ and thus, $n = \Theta\Big(2^{h(n)}\Big)$.

Proof. Assertion (i): Using the notation of Definition 4.3, we have

$$\delta'(n) \;=\; \ell - L(n) \;=\; \ell - n + \log L(n)\,.$$

Since $L(i) \sim i$,

$$\frac{L(n+1)}{L(n)} \;=\; \frac{n+1-\log L(n+1)}{n-\log L(n)} \;=\; 1+\frac{1-\log\frac{L(n+1)}{L(n)}}{n-\log L(n)} \;=\; 1+\frac{1-o(1)}{n}$$

and hence,

$$\log\frac{L(n+1)}{L(n)} \;=\; \left|O\Big(\frac{1}{n}\Big)\right| .$$

So

$$\delta'(n+1) \;=\; (\ell+1)-(n+1)+\log L(n)+\left|O\Big(\frac{1}{n}\Big)\right| \;=\; \delta'(n)+\left|O\Big(\frac{1}{n}\Big)\right| \qquad (4.1)$$

unless

$$\delta'(n)+\log\frac{L(n+1)}{L(n)} \;>\; \frac{1}{2}\,. \qquad (4.2)$$

The latter can happen only if $\delta'(n)=\frac{1}{2}-\left|O\Big(\frac{1}{n}\Big)\right|$, so

$$\frac{1}{2}-\left|O\Big(\frac{1}{n}\Big)\right| \;=\; \delta'(n) \;=\; \log L(n)-h'$$

for some $h'\in\mathbb{N}$. This shows that

$$n \;\sim\; L(n) \;=\; 2^{h'+\frac{1}{2}-|O(\frac{1}{n})|} \;\sim\; \sqrt{2}\,2^{h'}\,.$$

Assertion (ii): Notice that $a'_t = o(n_t)$. First assume that $a'_t > 0$. Using Equation (4.1) for $n=2^{h'_t}+h'_t$, $n=2^{h'_t}+h'_t+1,\ldots$ up to $n=n_t-1$ and adding up the error terms, we find that

$$\delta'(n_t)-\underbrace{\delta'(2^{h'_t}+h'_t)}_{=0} \;=\; O\left(\frac{a_t}{2^{h'_t}+h'_t}\right) \;=\; o(1)\,.$$

In particular, inequality (4.2) does never hold for these n. It follows that $h(n_t)=h\Big(2^{h'_t}+h'_t\Big)=h'_t$, and hence $a(n_t)=a'_t$. A similar argument works for $a'_t<0$, but here we use the estimate $\delta'(n-1) \;=\; \delta'(n)-\left|O\Big(\frac{1}{n}\Big)\right|$, which can be proved in the same way as (4.1).

Assertion (iii): On each range of n where $h(n)$ is constant we have $a(n)=n-2^{h(n)}-h(n) = n-$ *const.*, so $n\mapsto a(n)$ is isotone there. Hence the maximal value of $\left|a(\cdot)\right|$ is attained for n or $n+1$ in the situation of Assertion (i). So let $h'\in\mathbb{N}$ be such that $n\sim\sqrt{2}\,2^{h'}$. A simple calculation shows that $h(n)=h'$, $a(n)\sim(\sqrt{2}-1)2^{h(n)}$ and $h(n+1)=h'+1$, $a(n+1)\sim(1/\sqrt{2}-1)2^{h(n+1)}$. Therefore,

$$\limsup_{n\in\mathbb{N}}\frac{\left|a(n)\right|}{2^{h(n)}} \;=\; \max\left\{\sqrt{2}-1,1-\frac{1}{\sqrt{2}}\right\} \;<\; 0.42\,. \qquad \square$$

How are the two parameters δ' and a related? Given an a, we would like to know (approximately) how big δ' is. One would expect that a and δ' are nearly proportional as long as we do not move away too far from the 'nice' values $n = 2^h + h$. We introduce another parameter, $\tilde{a}(n)$, to make this connection explicit and precise. The notation $\tilde{a}(n)$ emphasises that the new parameter has about the same size as $a(n)$.

Definition 4.6 *We define*

$$\delta_x := \frac{x \log e}{L(n)}$$

and

$$\tilde{a}(n) := \frac{\delta'(n)\, L(n)}{\log e} .$$

Obviously, $\delta' = \delta_{\tilde{a}}$.

Lemma 4.7

(i) If $a = o(n)$, then $\tilde{a} = a + O(a^2/n) \sim a$.
(ii) If $a \to \pm\infty$, then $\tilde{a} \to \pm\infty$.

Proof. Assertion (i): Let $h = h(n)$ and $a = a(n)$ as in Definition 4.4, and $L = L(n)$. To determine $\delta'(n)$, we expand L using the functional equation (2.1):

$$L = n - \log L = \underbrace{n - h}_{\in \mathbb{N}} - \log \frac{L}{2^h} . \tag{4.3}$$

We claim that $\delta'(n) = \frac{\log L}{2^h} = o(1)$. Using (2.1) once more, we find that

$$\log \frac{L}{2^h} = \log \frac{2^h + h + a - \log L}{2^h} = \log \left(1 + \frac{h - \log L + a}{2^h} \right) . \tag{4.4}$$

Since

$$\log(1 + x) = x \log e + O\Big(x^2\Big) \tag{4.5}$$

for $x = o(1)$, and $L(n) \sim n$,

$$h - \log L = h - \log \Big(2^h(1 + o(1))\Big) = -\log \Big(1 + o(1)\Big) = o(1) .$$

So

$$\log \frac{L}{2^h} = \log \left(1 + \frac{a + o(1)}{2^h} \right) = \frac{a \log e}{2^h} \left(1 + O\Big(\frac{a}{2^h}\Big) \right) = o(1)$$

is indeed the fractional part of L as was suggested in (4.3), and $\delta'(n) = \log(L/2^h) = a \log e/L + O(a^2/L^2) \sim a \log e/L$. Therefore, $\tilde{a} = L\,\delta'/\log e = a + O(a^2/n) \sim a$.

Assertion (ii): To prove the contrapositive, assume that there exists an infinite subsequence of n for which $\tilde{a} = O(1)$. By definition of δ',

$$L = \ell - \delta' = \ell - \delta_{\tilde{a}}\,, \tag{4.6}$$

where $\ell = \ell(n) \in \mathbb{N}$ and $\tilde{a} = \tilde{a}(n) = O(1)$. For notational convenience, assume that the subsequence is equal to the original one. Application of the mapping $L^{-1}\colon\ i \mapsto i + \log i$ to both sides of (4.6) gives

$$n = \ell - \delta_{\tilde{a}} + \log(\ell - \delta_{\tilde{a}})\,. \tag{4.7}$$

Observe that $|\delta_{\tilde{a}}|$ is rather small; we have

$$\delta_{\tilde{a}} = O(\tilde{a})/L = O(1/n)\,.$$

So we can rewrite (4.7) as

$$n = \ell + \log \ell - c\,,$$

with a small correction term c, whose size is only

$$c := \delta_{\tilde{a}} + \log \frac{\ell}{\ell - \delta_{\tilde{a}}} = O(1/n) \tag{4.8}$$

by (4.5). Therefore, we can guess that $h(n)$ is equal to

$$h' := \log \ell - c = n - \ell \in \mathbb{N}\,. \tag{4.9}$$

So far, we know that h' is an integer close to $\log \ell$. But what about $2^{h'}$ and ℓ? Expressing ℓ in terms of c and h', we find

$$\ell = 2^{\log \ell} = 2^{h'+c} = 2^{h'+O(1/n)} = 2^{h'}\Big(1 + O(1/n)\Big) = 2^{h'} + O\Big(2^{h'}/n\Big)\,,$$

and using (4.6) and (4.8), we have

$$2^{h'} = 2^{\log \ell - c} = 2^{\log(L+O(1))-O(1/n)} = \Big(L + O(1)\Big)2^{O(1/n)} \sim L = O(n)\,.$$

Hence $2^{h'} = \ell + O(1)$. So by (4.9)

$$a' := n - 2^{h'} - h' = n - \ell - h' + O(1) = O(1)\,.$$

But this implies that $h(n) = h'$ and $a(n) = a' = O(1)$ by Proposition 4.5 (ii). Going from the subsequence back to the original sequence, we have shown that $|\tilde{a}| \not\twoheadrightarrow +\infty$ implies $|a| \not\twoheadrightarrow +\infty$. By definition, a and $\tilde{a}$ have the same sign. $\square$

4.2 *Expected Size of the qOBDD Levels*

Now we are prepared to extend the idea of Proposition 4.2 to the case of general n. Assertions (i) and (ii) of Theorem 4.8 are concerned with the size of the (possibly)

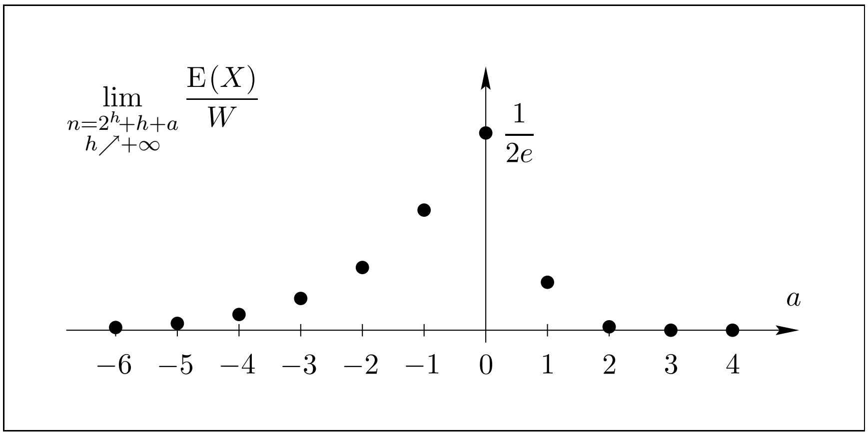

Fig. 4. When the strong Shannon effect does not hold (Corollary 4.11).

critical level i', while Assertion (iii) says that the other levels altogether contribute only $o(1)$ to $\mathrm{E}(X)/2^L$. Note that (i) implies the existence of a limit for constant a (which is not at all obvious). See Fig. 4 for an illustration.

Theorem 4.8 *Let i' be the critical level as in Definition 4.3.*

(i) For sequences of n such that $a(n) = o(\sqrt{n}\,)$,

$$\frac{\mathrm{E}\Big(X_{i'}\Big)}{2^L} = \begin{cases} 2^{-a+o(1)}\left(e^{-2^{a+o(1)}} - 1\right) + 1\,, & a \leqslant 0\,; \\ 2^{-a+o(1)}\, e^{-2^{a+o(1)}}\,, & a \geqslant 0\,. \end{cases}$$

(ii) For sequences of n such that $\Big|a(n)\Big| \to +\infty$, $\dfrac{\mathrm{E}\Big(X_{i'}\Big)}{2^L} = o(1)\,.$

(iii) For all sequences of n, $\dfrac{\mathrm{E}\Big(X - X_{i'}\Big)}{2^L} = 2^{-\Omega(\sqrt{n}\,)}$.

Proof. Let $k := k_{i'}$ and $m := m_{i'}$. Since we want to use Proposition 4.1, we need upper and lower bounds for $q := q_{i'}$.

Estimation of q:
Writing

$$q = \left(1 - \frac{1}{m}\right)^k = \left(1 - \frac{1}{m}\right)^{m\frac{k}{m}} = \left(1 - \frac{1}{m}\right)^{(m-1)\frac{k}{(m-1)}}$$

and using the inequalities $(1-1/x)^{x-1} > 1/e > (1-1/x)^x$, valid for $x \geqslant 2$, we see that

$$e^{-k/m} > q > e^{-k/(m-1)} \geqslant e^{-k/m}e^{-2k/m^2} \geqslant e^{-k/m}\left(1 - \frac{2k}{m^2}\right)\,. \tag{4.10}$$

Therefore,

$$q = e^{-k/m}\left(1 - \left|O\left(\frac{k}{m^2}\right)\right|\right)\,, \tag{4.11}$$

which is $e^{-k/m}\Big(1 - \Big|O(1/m)\Big|\Big)$, if $a \leqslant 0$ (because then $k \leqslant m$).

Proof of Assertion (i):
In view of the proof of Assertions (ii) and (iii), the estimations we derive to prove Assertion (i) rely on the weaker assumption $\tilde{a} = o(n)$ and do not require that $a = o(\sqrt{n}\,)$. By Lemma 4.7 (i), $\tilde{a} = a + o(1)$ for $a = o(\sqrt{n}\,)$. We investigate positive and negative values of $\tilde{a}$ separately.

Estimation of $k/2^L$ and $m/2^L$ for $\tilde{a} = o(n)$:
Clearly,

$$\frac{k}{2^L} = 2^{\delta'} = e^{\tilde{a}/L} = e^{o(1)} \sim 1\,, \tag{4.12}$$

because $n \sim L(n)$. — To estimate $m/2^L$, we apply the Taylor approximation $e^x = 1 + x + O(x^2)$ for $x \to 0$, which yields

$$\Big(1 - 2^{-\delta'}\Big)L = \Big(1 - e^{-\tilde{a}/L}\Big)L = \tilde{a} + r \sim \tilde{a} \tag{4.13}$$

for some

$$r = r(n) = O\Big(\tilde{a}^2/L\Big) = O\Big(\tilde{a}^2/n\Big) = o(\tilde{a})\,. \tag{4.14}$$

So

$$\frac{m}{2^L} = 2^{\left(2^{-\delta'}-1\right)L} = 2^{-\tilde{a}-r}\,. \tag{4.15}$$

The case $0 \geqslant \tilde{a} = o(n)$:
In this case, $\mathrm{E}\Big(X_{i'}\Big) = k - m(1-q)$ by Proposition 4.1, and the asymptotic (4.11) for q implies that

$$\begin{aligned} \mathrm{E}(X_{i_{\delta'}}) = k - m(1-q) &= k - m\Big(1 - e^{-k/m}(1 - O(k/m^2))\Big) \\ &= k - m\Big(1 - e^{-k/m}\Big) + O(k/m)\,. \end{aligned} \tag{4.16}$$

By substituting (4.12) and (4.15) into (4.16), we obtain

$$\frac{\mathrm{E}(X_{i'})}{2^L} = 2^{-\tilde{a}-r}\left(e^{-2^{\tilde{a}+r}e^{\tilde{a}/L}} - 1\right) + e^{\tilde{a}/L} + O\Big(2^{-L}\Big)\,. \tag{4.17}$$

In particular, (i) follows for $a = -\Big|o(\sqrt{n}\,)\Big|$.

The case $0 \leqslant \tilde{a} = o(n)$:
In this case, $\mathrm{E}\Big(X_{i'}\Big) = mq$ by Proposition 4.1 and $q = e^{-k/m}\Big(1 + O(k/m^2)\Big)$ by (4.11). Since $\delta' = \delta_{\tilde{a}} = o(n)\log e/L = o(1)$, we have

$$\frac{k}{m^2} = \frac{2^{\delta'+L}}{\left(2^{2^{-\delta'}L}\right)^2} = 2^{o(1)+L-2\cdot 2^{o(1)}L} = o(1)\,,$$

and the asymptotic for q simplifies to $q \sim e^{-k/m}$. Together with (4.12) and (4.15), we obtain

$$\frac{\mathrm{E}(X_{i'})}{2^L} = \frac{m\,q}{2^L} \sim \frac{m}{2^L}\,e^{-k/m} = \frac{2^{-\tilde{a}-r}}{e^{2^{\tilde{a}+r+o(1)}}}\,, \tag{4.18}$$

where $r = O(\tilde{a}^2/L)$ by (4.14). In particular, (i) follows for $a = \left|o(\sqrt{n}\,)\right|$.

Proof of Assertion (ii), first part:
We split the sequence of n into four subsequences depending on whether $\left|\tilde{a}(n)\right| \leqslant \sqrt{n}$ or $\left|\tilde{a}(n)\right| > \sqrt{n}$ and whether $\tilde{a}(n) \geqslant 0$ or $\tilde{a}(n) \leqslant 0$. Lemma 4.7 (ii) tells us that $\left|\tilde{a}(n)\right| \to +\infty$, because $\left|a(n)\right| \to +\infty$.

For the two subsequences satisfying $\left|a(n)\right| \leqslant \sqrt{n}$, the estimations from the proof of Assertion (i) can be applied, since we only used the premise $\tilde{a} = o(n)$ in the proof. Also, we still have $a \sim \tilde{a}$, so (4.14) implies $r = O(1)$.

The case $0 \leqslant \tilde{a}(n) \to +\infty \ \wedge\ \tilde{a}(n) \leqslant \sqrt{n}$*:*
Since $\tilde{a} + r \to +\infty$, $\mathrm{E}\big(X_{i'}\big)/2^L \to 0$ follows immediately from (4.18).

The case $0 \geqslant \tilde{a}(n) \to -\infty \ \wedge\ \tilde{a}(n) \geqslant -\sqrt{n}$*:*
In this case, $\tilde{a} + r \to -\infty$, so $2^{\tilde{a}+r} = o(1)$. Using (4.17) and the Taylor approximation $e^x = 1 + x + \big(1 + o(1)\big)x^2/2$ for $x \to 0$ we see that

$$\begin{aligned} \frac{\mathrm{E}(X_{i'})}{2^L} &= 2^{-\tilde{a}-r}\left(-2^{\tilde{a}+r}e^{\tilde{a}/L} + 2^{2(\tilde{a}+r)-1}e^{2\tilde{a}/L}\Big(1+o(1)\Big)\right) + e^{\tilde{a}/L} + O\Big(2^{-L}\Big) \\ &= 2^{\tilde{a}+r-1}e^{2\tilde{a}/L}\Big(1+o(1)\Big) + O\Big(2^{-L}\Big) \ = \ o(1)\,. \end{aligned} \tag{4.19}$$

So far we have proved Assertion (ii) for $|\tilde{a}| \leqslant \sqrt{n}$. To complete the proof of Theorem 4.8 (ii) and (iii), we need the following lemma. •••

Lemma 4.9 *Assume that $n \to +\infty$ and $j = j(n) \in [n]$ is a level such that*

$$|j - i_0| \geqslant \frac{\sqrt{n}\,\log e}{L(n)}\,,$$

where i_0 denotes the critical point. Let $a' := (j - i_0)L(n)/\log e$. Then

$$\frac{\mathrm{E}(X_j)}{2^L} \leqslant \frac{\mathrm{E}(X_j)}{w_j} = \begin{cases} 2^{a'+O(1)} = 2^{-\Omega(\sqrt{n}\,)}\,, & j < i_0\,; \\ e^{-2^{\Omega(\sqrt{n}\,)}}\,, & j > i_0\,. \end{cases}$$

Proof. Recall that the definition of $\tilde{a}(n)$ was devised such that $i_{\delta'} = i_{\delta_{\tilde{a}}}$. In this lemma, we are no longer concerned with $i_{\delta'}$, but an arbitrary level j. Nevertheless, we can define

$$a' := \frac{(j - i_0)\,L(n)}{\log e}\,. \tag{4.20}$$

Then $j = i_{\delta_{a'}}$ is satisfied and $|j - i_0| \geqslant \sqrt{n}\,\log e/L(n)$ is equivalent to $|a'| \geqslant \sqrt{n}$. We write $k := k_j$, $m := m_j$, $q := q_j$, and $L := L(n)$. We consider two cases:

$a' \leqslant -\sqrt{n}$ and $a' \geqslant \sqrt{n}$. Recall that w_j is given by (2.4).

First, assume that $a' \leqslant -\sqrt{n}$. Then $j < i_0$ and by (2.3), $k = 2^{\delta_{a'}+L} = 2^L 2^{j-i_0} \leqslant 2^L$ and $m = 2^{2^{-\delta_{a'}} L}$; hence

$$\frac{k}{m} \leqslant \frac{2^L}{m} = 2^{\left(1-2^{-\delta_{a'}}\right)L} = 2^{\left(1-e^{-a'/L}\right)L} \leqslant 2^{a'} = 2^{-\Omega(\sqrt{n}\,)}\,, \tag{4.21}$$

using (2.7) for the last inequality. In particular, we have by (4.11)

$$q = e^{-k/m}\Big(1 + O(k/m^2)\Big) = e^{-k/m}\Big(1 + O(2^{a'}/m)\Big)\,.$$

Therefore, analogously to (4.16), it holds

$$\begin{aligned} \mathrm{E}(X_j) = k - m(1-q) &= k - m\Big(1 - e^{-k/m}(1 - O(2^{a'}/m))\Big) \\ &= k - m\Big(1 - e^{-k/m}\Big) + O(2^{a'})\,. \end{aligned}$$

Using $e^x = 1 + x + \big(1 + o(1)\big)x^2/2$ for $x \to 0$, we get

$$\begin{aligned} \mathrm{E}(X_j) &= k - m\left(1 - \left(1 - \frac{k}{m} + \frac{k^2}{2m^2}\Big(1 + o(1)\Big)\right)\right) + O\Big(2^{a'}\Big) \\ &= \frac{k^2}{2m}\Big(1 + o(1)\Big) + O\Big(2^{a'}\Big)\,. \end{aligned} \tag{4.22}$$

Here the leading term is bounded by

$$\frac{k^2}{2m} = k \cdot \frac{k}{2m} \leqslant k \cdot 2^{a'-1}$$

because of (4.21). Since $j < i_0$, $w_j = k$ and we infer

$$\frac{\mathrm{E}(X_j)}{w_j} = \frac{\mathrm{E}(X_j)}{k} \leqslant 2^{a'+O(1)} = 2^{-\Omega(\sqrt{n}\,)} \tag{4.23}$$

as claimed.

Now for the second case, $a' \geqslant \sqrt{n}$. In this case $j > i_0$ and $k = 2^{j-i_0}2^L > 2^L$. By Proposition 4.1 (ii) and (4.10) we have

$$\mathrm{E}(X_j) = mq \leqslant me^{-k/m} \leqslant me^{-2^L/m}\,. \tag{4.24}$$

By (2.3),

$$\frac{2^L}{m} = 2^{\left(1-2^{-\delta_{a'}}\right)L}\,.$$

For $x \geqslant 0$, the inequality $1 - e^{-x} \geqslant \frac{x}{1+x}$ follows from (2.7) and the mapping $x \mapsto \frac{x}{1+x}$ is monotone. Using $a'/L \geqslant a'/n \geqslant \sqrt{n}$ in the last inequality, we see that

$$\Big(1 - 2^{-\delta_{a'}}\Big)L = \Big(1 - e^{-a'/L}\Big)L \geqslant \frac{a'/L}{1 + a'/L}L \geqslant \frac{\sqrt{n}}{1 + \sqrt{n}/L} = \Omega(\sqrt{n}\,)\,.$$

By (2.4) $w_j = m$ and therefore,

$$\frac{\mathrm{E}(X_j)}{w_j} = \frac{\mathrm{E}(X_j)}{m} \leqslant e^{-2^L/m} \leqslant 2^{-2^{\Omega(\sqrt{n})}} .$$

□(Lemma 4.9)

Proof of Theorem 4.8, continued:
Proof of Assertion (ii), conclusion:
In the remaining cases the subsequences satisfy $\left|\tilde{a}(n)\right| > \sqrt{n}$ and are therefore covered by Lemma 4.9 — just set $j := i'$.

Proof of Assertion (iii):
If $j \in [n] \setminus \{i_{\delta'}\}$, then $|j - i_0| \geqslant 1/2$ holds by definition of δ'. So by Lemma 4.9,

$$\frac{\mathrm{E}\big(X - X_{i'}\big)}{2^L} = \sum_{j \in [n] \setminus \{i'\}} \frac{\mathrm{E}(X_j)}{2^L} = n \cdot 2^{-\Omega(\sqrt{n})} = 2^{-\Omega(\sqrt{n})} .$$

□(Theorem 4.8)

Theorem 4.8 gives us a fairly complete overview of the 'expected' shape of a random qOBDD. Above and below $i_{\delta'}$, the levels are essentially full. If and only if δ' is sufficiently small, i. e., $\delta' = O(1/n)$, then the expected size of the critical level $i_{\delta'}$ is by a factor $1 - \Omega(1)$ smaller than its worst-case width. We summarize these observations in a corollary.

Corollary 4.10 *Assume that $n \to +\infty$ and $j = j(n) \in [n]$.*

(i) For all j,

$$1 - \frac{1}{e} \leqslant \frac{\mathrm{E}\big(Y_j\big)}{w_j} \leqslant 1 .$$

(ii) If $|j - i_0| = \omega(1/n)$, then

$$\frac{\mathrm{E}\big(Y_j\big)}{w_j} \sim 1 .$$

Proof. Assertion (i) follows from Theorem 4.8 (in particular, equations (4.17) and (4.18) from the proof of its Assertion (i)), since $\mathrm{E}(X_j)/w_j$ is maximized if $j = i' = i_0$ with value $\mathrm{E}\big(X_{i_0}\big) = w_{i_0}/e$. Hence,

$$\frac{\mathrm{E}(Y_j)}{w_j} = \frac{w_j - \mathrm{E}(X_j)}{w_j} \geqslant 1 - \frac{1}{e} .$$

Assertion (ii) is immediate from Theorem 4.8 (ii) (for small $|j - i_0|$) and Lemma 4.9 (for large $|j - i_0|$). □

For small $a(n)$, we can determine the ratio of the expected qOBDD size to its worst-case size very precisely. The special case where $a(n)$ is a constant is shown in Fig. 4 on page 14.

Corollary 4.11 *If $n \to +\infty$ is such that $a(n) = o(\sqrt{n}\,)$, then*

$$\frac{\mathrm{E}(Y)}{W} = \begin{cases} 1 - \dfrac{2^{-a+o(1)}\left(e^{-2^{a+o(1)}} - 1\right) + 1}{2 + o(1)}\,, & a \leqslant 0\,; \\[2ex] 1 - \dfrac{2^{-a+o(1)}\,e^{-2^{a+o(1)}}}{1 + 2^{-a} + o(1)}\,, & a \geqslant 0\,. \end{cases}$$

Proof. Direct plug-in from Theorem 4.8 and the results on the worst-case size of [17,15]. □

It is not hard to see that the decreasing rate of $\mathrm{E}(X)/W$ is doubly exponential for $a \to +\infty$ and exponential for $a \to -\infty$.

In [15, Chapter 6] the asymptotic size of $\mathrm{E}(Y)/W$ was investigated for general n. One also has to take the effect of the deletion rule into account. It turns out that the 'Shannon gap' $\big(W - \mathrm{E}(Z^*)\big)/W$ is *minimized* for some parametrisation $n = \big(\frac{5}{4} + o(1)\big)2^h$, with value $2^{\left(\frac{2}{5}+o(1)\right)n}$.

4.3 Strong Shannon Effect for qOBDDs with a Fixed Variable Ordering

Now we extract a qualitative result from the preceding quantitative analysis of the expected size of the qOBDD with a fixed variable ordering for a random Boolean function. For which n is the expected size $\mathrm{E}(Y)$ of the qOBDD (with a fixed variable ordering) equal to the worst-case size W up to terms of lower order? The answer is: if and only if n stays apart from $2^h + h$.

Theorem 4.12 *Let $B := \bigcup_{h\in\mathbb{N}}\big[2^h + h - d(h) \mathinner{..} 2^h + h + d(h)\big]$ and $A := \mathbb{N} \setminus B$.*

(i) If $n \to +\infty$ such that $n \in A$ for some sequence $d(h) \to +\infty$, then

$$\frac{\mathrm{E}(Y)}{W} = 1 - o(1)\,.$$

(ii) If $n \to +\infty$ such that $n \in B$ for some sequence $d(h) = O(1)$, then

$$\frac{\mathrm{E}(Y)}{W} = 1 - \Omega(1)\,.$$

Proof. We have $\mathrm{E}(X)/W = \Theta\Big(\mathrm{E}(X)/2^L\Big)$, since $W = \Theta(2^L)$ by Theorem 3.1. By Theorem 4.8 (iii), all levels except i' are negligible (they contribute only $o(1)$ to $\mathrm{E}(X)/2^L$).

Assertion (i): Since $d(h) \to +\infty$ and n is a sequence chosen from the set A, we have $\big|a(n)\big| \to +\infty$ as $n \to +\infty$. Therefore, $\lim_n \mathrm{E}\Big(X_{i'}\Big)\Big/2^L = 0$ by Theorem 4.8 (ii), and we are done.

Assertion (ii): Since $d(h) = O(1)$ and n is a sequence chosen from the set B, we have $a(n) = O(1)$. By partitioning the sequence of n into subsequences, we may assume that $a(n)$ is a constant. These subsequences may have finite or infinite length, but only a finite number of subsequences can be infinitely long, because the original sequence satisfied $a(n) = O(1)$. For each infinite subsequence of integers n where $a(n)$ is a constant Theorem 4.8 (i) implies that $\lim_n \mathrm{E}\Big(X_{i'}\Big)\Big/2^L > 0$. So the original sequence satisfies $\liminf_n \mathrm{E}\Big(X_{i'}\Big)\Big/2^L > 0$, i. e., $\mathrm{E}\Big(X_{i'}\Big)\Big/2^L = \Omega(1)$. □

Note that Markov's inequality implies (since $Y \leqslant W' \sim W$) that the strong Shannon effect for the qOBDD size for a *fixed* variable ordering of a random Boolean function *does* hold in Case (i) of Theorem 4.12, whereas it *does not* hold in Case (ii).

5 Strong Shannon Effect for Optimal qOBDDs

We have seen that the *expected* qOBDD size for a *fixed* variable ordering is approximately $W(n)$ if and only if $\big|a(n)\big| \to +\infty$ (Theorem 4.12). The next step is to consider qOBDDs with *optimal* variable orderings. Our approach is to prove that for a random Boolean function with high probability *all* variable orderings lead to *almost the same* qOBDD size. (Here we apply Azuma's inequality.) Then in particular, an optimal variable ordering does only a little better than the canonical one.

5.1 Azuma's Inequality

Azuma's martingale inequality is by now a standard method to prove strong concentration of random variables. Here we give a purely 'combinatorial' formulation.

Theorem 5.1 (Azuma's Inequality, see e. g. [1, Theorem 7.4.2]**)**
If B is a finite domain and $S\colon B^k \to \mathbb{R}$ is a function satisfying the 'Lipschitz

condition'

$$\forall\, b, b' \in B^k : \#\{j \mid b_j \neq b'_j\} \leqslant 1 \;\rightarrow\; \left|S(b) - S(b')\right| \leqslant 1 \tag{5.1}$$

and the coordinates of b are chosen independently at random, then

$$\Pr_b\left(\left|S(b) - \mathrm{E}(S)\right| \geqslant \lambda\sqrt{k}\right) \;\leqslant\; 2\,e^{-\lambda^2/2}\,.$$

The way in which we apply Azuma's inequality to the urn occupancy experiment is somewhat simple-minded, but we cannot expect a significantly stronger result to hold (see next section). For related work on urn occupancy, see [1,19,12].

Corollary 5.2 *Consider an urn experiment where k balls are thrown independently uniformly at random into m urns, and denote by y the number of non-empty urns. Then*

$$\Pr\left(\left|y - \mathrm{E}(y)\right| \geqslant \lambda\sqrt{k}\right) \;\leqslant\; 2\,e^{-\lambda^2/2}\,.$$

Proof. Denote a random assignment of balls to urns by $b\colon\, [k] \rightarrow [m]$ and let $y(b) := \#b[k]$ be the number of non-empty urns for this particular assignment. Clearly, y satisfies the Lipschitz condition (5.1), since the number of non-empty urns can only change by 1 if we move a ball from one urn to another. Therefore, Theorem 5.1 is applicable. □

5.2 Optimal qOBDDs

From Corollary 5.2 we obtain the following strong concentration result for the size Y of the quasireduced OBDD with respect to the canonical variable ordering of a random Boolean function. As it turns out, the probability that Y is somewhat more than $\sqrt{\mathrm{E}(Y)}$ apart from $\mathrm{E}(Y)$ is only doubly exponentially small in n.

Theorem 5.3 *For every $c > 0$,*

$$\Pr\left(\left|Y - \mathrm{E}(Y)\right| \geqslant n2^{\frac{1+c}{2}L}\right) \;\leqslant\; 2ne^{-2^{cL}/4}\,.$$

Proof. At each level j, we have an urn experiment where k_j balls are thrown into m_j urns, and Y_j is the number of nodes at level j of the qOBDD as well as the number of non-empty urns. So by Corollary 5.2,

$$\Pr\left(\left|Y_j - \mathrm{E}(Y_j)\right| \geqslant \lambda\sqrt{k_j}\right) \;\leqslant\; 2e^{-\lambda^2/2}\,. \tag{5.2}$$

We consider two cases. From Section 2 recall that $Y_j \leqslant w_j = \min\{k_j, m_j\}$ for all levels $j \in [n]$. For $j > i_1$, we have

$$Y_j \leqslant m_j \leqslant m_{i_1} = 2^{L/2}\,. \tag{5.3}$$

If $j \leqslant i_1$, then $Y_j \leqslant k_j \leqslant k_{i_1} \leqslant 2^{L+1}$. Using (5.2) with $\lambda := 2^{\frac{1+c}{2}L}\big/\sqrt{k_j}$, we get for $j \leqslant i_1$

$$\Pr\left(\left|Y_j - \mathrm{E}(Y_j)\right| \geqslant 2^{\frac{1+c}{2}L}\right) \;\leqslant\; 2e^{-2^{cL}/4}\,, \tag{5.4}$$

since $\lambda^2/2 \geqslant 2^{(1+c)L-L-1}\big/2 = 2^{cL}/4$. For $j > i_1$, (5.3) gives $\left|Y_j - \mathrm{E}(Y_j)\right| \leqslant 2^{L/2}$ and (5.4) trivially holds. So

$$\Pr\left(\exists\, j \in [n] :\; \left|Y_j - \mathrm{E}(Y_j)\right| \geqslant 2^{\frac{1+c}{2}L}\right) \;\leqslant\; n \cdot 2e^{-2^{cL}/4}\,, \tag{5.5}$$

and the theorem follows. □

We remark that using Azuma's inequality one cannot improve the point $n2^{\frac{1+c}{2}L}$ where Theorem 5.3 'cuts off' beyond $\omega(\sqrt{k_j}\,)$, and for $j = i'$ we have $k_j = \Omega(2^L)$. The question arises whether a weaker (maybe not doubly exponential) probability bound is provable for some cut-off point $O\Big(2^{\left(\frac{1}{2}-\Omega(1)\right)L}\Big)$. The answer is "at least in general: no", because $Y_{i'}$ is asymptotically normally distributed for certain parametrizations of n. For example, if $n = 2^h + h$, then $i' = 2^h + 1$ and $k_{i'} = m_{i'} = 2^h$, and the distribution of Y_{2^h+1} is asymptotically normal by Theorem I.3.1 of [20]. Their result can only be applied if the ratio k/m is a *constant*, because it makes no assertion on the convergency rate.

Now we turn to optimal variable orderings. Define $Y_*(f) := \min_\pi Y_\pi(f)$, where the index π runs over all variable orderings. $Y_*(f)$ is the minimal size of a qOBDD for f. Clearly $\mathrm{E}(Y_\pi)$ does not depend on π, since we consider the uniform distribution for f. Let us write $\mathrm{E}(Y) = \mathrm{E}(Y_\pi)$. For most Boolean functions, even choosing an optimal variable ordering gives little improvement.

Theorem 5.4 *For every $c > 0$,*

$$\Pr\left(\left|Y_* - \mathrm{E}(Y)\right| \geqslant n2^{\frac{1+c}{2}L}\right) \;\leqslant\; e^{-2^{cL}/4+O(n\log n)}\,.$$

Proof. If $\left|Y_* - \mathrm{E}(Y)\right|$ is large, then there exists a variable ordering π such that $\left|Y_\pi - \mathrm{E}(Y)\right|$ is large. For each variable ordering, there is only a doubly exponentially small fraction of exceptional Boolean functions. So we simply multiply the probability bound from Theorem 5.3 by $n!$, the number of all possible variable orderings, which satisfies $n! < n^n = e^{n\ln n}$. □

From a larger perspective, the important facts here are that $2^{\frac{1+c}{2}L} = o(W)$ and $W \cdot e^{-2^{cL}/4+O(n \log n)} = o(1)$. This implies that Theorem 4.12 (which is about the expected size) carries over to the case of optimal variable orderings; only the o- and Ω-terms change. But we also know how large Y_* is with high probability:

Theorem 5.5 *Let $n \to +\infty$. Then*

$$\Pr\Big(Y_* = (1 - o(1))W\Big) \sim 1 \quad \textit{iff} \quad \Big|a(n)\Big| = \omega(1)\,.$$

That is, the strong Shannon effect holds for qOBDDs with optimal variable orderings if and only if n is such that $\Big|a(n)\Big| \to +\infty$. □

In other words, the strong Shannon effect for the *optimal* qOBDD size of a random Boolean function *does* hold in Case (i) of Theorem 4.12, whereas it *does not* hold in Case (ii).

6 Deletion Rule

So far, we know that the minimal qOBDD size is approximately the worst-case size W if and only if $\Big|a(n)\Big| \to +\infty$ (Theorem 5.5). In this section, we show that the deletion rule gives only a comparatively small amount of reduction. Thus, the same statement is true for minimal OBDDs as well. This finishes the proof of Main Theorem 1.1.

Actually, our analysis of large deviations goes beyond just proving that the weak Shannon effect holds for all sequences of n, which was already shown by Wegener [30] using the second moment method. We obtain a doubly exponential probability bound that enables us to generalise Main Theorem 1.1 to OKFDDs.

6.1 Chvátal's Inequality

To estimate the probability of large deviations from the expected amount of reduction by the deletion rule, we apply a large deviation inequality for hypergeometrically distributed random variables due to Chvátal, cited here in slightly adapted form.

Theorem 6.1 (Chvátal [9]**)** *Consider an urn experiment where y balls are chosen without replacement from an urn containing white and black balls. Denote the fraction of black balls in the urn by p and let x' be the number of black balls chosen. Then x' is a hypergeometrically distributed random variable with parameters p and*

y, mean $\mathrm{E}(x') = py$, and for all $\varepsilon \geqslant 0$, we have

$$\Pr\Big(x' \geqslant (p+\varepsilon)y\Big) \leqslant e^{-2\varepsilon^2 y}\,.$$

6.2 *Bounding the Effect of the Deletion Rule*

We denote the amount of reduction achieved by the deletion rule at level i by $X'_i = Y_i - Z_i$ and put $X' := \sum_{i=1}^{n} X'_i = Z - Y$. We will show that the probability that X' is somewhat bigger than $\sqrt{W}$ or $\sqrt{\mathrm{E}(Y)}$ is only doubly exponentially small. First we consider qOBDDs with a fixed variable ordering. It turns out that the random variable X'_i is hypergeometrically distributed *if we condition* on a particular value of Y_i. Hence we can apply Chvátal's inequality.

Theorem 6.2 *For every $c > 0$,*

$$\Pr\Big(X' \geqslant n2^{\frac{1+c}{2}L}\Big) \leqslant e^{-(2+o(1))2^{cL}}\,.$$

Proof. The number of nodes deleted at some level j is given by the following urn experiment. Among the m_j subfunctions which are possible at level j of the qOBDD there are m_{j+1} functions that do not depend essentially on the variable x_j. Since the merging rule has already been applied, we have a situation in which Y_j balls are chosen *without replacement* from an urn containing $m_j - m_{j+1}$ 'white' and m_{j+1} 'black' balls. The black balls correspond to those nodes at level j which are deleted afterwards. Therefore, X'_j is hypergeometrically distributed with parameters Y_j and

$$p_j := \frac{m_{j+1}}{m_j} = 2^{\left(2^{-\delta-1}-2^{-\delta}\right)L} = 2^{-2^{-\delta-1}L} = \sqrt{2^{-2^{-\delta}L}} = \frac{1}{\sqrt{m_j}}\,. \qquad (6.1)$$

In our case Y_j is itself a random variable. Therefore we switch to conditional probabilities. For $y \leqslant w_j$

$$\Pr\Big(X'_j \geqslant (p_j+\varepsilon_j)w_j \,\Big|\, Y_j = y\Big) \;\leqslant\; \Pr\Big(X'_j \geqslant (p_j+\varepsilon_j)w_j \,\Big|\, Y_j = w_j\Big)$$

since choosing more balls makes it more likely to get more than a given amount of black balls. In this way, we can invoke Theorem 6.1 and estimate the situation at each level j as follows, using a parameter ε_j to be specified later.

$$\begin{aligned}\Pr\Big(X'_j \geqslant (p_j+\varepsilon_j)w_j\Big) &= \sum_{y=0}^{w_j} \Pr\Big(X'_j \geqslant (p_j+\varepsilon_j)w_j \;\Big|\; Y_j = y\Big)\Pr\Big(Y_j = y\Big)\\ &\leqslant \Pr\Big(X'_j \geqslant (p_j+\varepsilon_j)w_j \;\Big|\; Y_j = w_j\Big)\\ &\leqslant e^{-2\varepsilon_j^2 w_j}\,. \qquad (6.2)\end{aligned}$$

A sufficient condition for $X' \leqslant n2^{\frac{1+c}{2}L}$ is that the inequality

$$X'_j \leqslant 2^{\frac{1+c}{2}L} \tag{6.3}$$

is satisfied for all j. Let $\delta := j - i_0$ (so $j = i_\delta$). If $j > i_1$, then (6.3) holds trivially, because

$$X'_j \leqslant Y_j \leqslant w_j = m_j = 2^{2^{-\delta}L} \leqslant 2^{\frac{1}{2}L}\,.$$

For $j \leqslant i_1$, we want to show that (6.3) holds with high probability using (6.2). So we define

$$\varepsilon_j := \frac{2^{\frac{1+c}{2}L}}{w_j} - p_j\,, \tag{6.4}$$

which gives

$$(p_j + \varepsilon_j)w_j = 2^{\frac{1+c}{2}L}\,.$$

We claim that $\varepsilon_j \geqslant 0$. Note that $p_j = 1/\sqrt{m_j}$ by (6.1) and for $j \geqslant i_0$, $w_j = m_j$. Comparing the logarithms, we find that

$$\log \frac{2^{\frac{1+c}{2}L}}{w_j\,p_j} = \log\left(2^{\frac{1+c}{2}L}\,p_j\right) = \left(\tfrac{1+c}{2} - 2^{-\delta-1}\right)L \geqslant \tfrac{c}{2}\,L > 0\,, \tag{6.5}$$

which proves $\varepsilon_j \geqslant 0$ for $j \geqslant i_0$. For $j \leqslant i_0$, $w_j = k_j$ and since $\varepsilon_{i_0} \geqslant 0$ and the mappings $j \mapsto k_j$ and $j \mapsto p_j$ are isotone for every n,

$$\varepsilon_j \;=\; \frac{2^{\frac{1+c}{2}L}}{w_j} - p_j \;=\; \frac{2^{\frac{1+c}{2}L}}{k_j} - p_j \;=\; \frac{2^{\frac{1+c}{2}L}}{k_{i_0}} - p_{i_0} \;=\; \varepsilon_{i_0} \;\geqslant\; 0\,.$$

Hence $\varepsilon_j \geqslant 0$ for all j.

Using (6.2) for $j \leqslant i_1$ with the ε_j defined in (6.4), we get the estimate

$$\Pr\left(X'_j \geqslant 2^{\frac{1+c}{2}L}\right) \leqslant e^{-2\varepsilon_j^2 w_j}\,. \tag{6.6}$$

We need to lower bound $\varepsilon_j^2 w_j$. Again we consider two cases.

If $i_0 \leqslant j \leqslant i_1$, then by (6.1),

$$\varepsilon_j^2 w_j \;=\; \left(\frac{2^{\frac{1+c}{2}L}}{w_j} - p_j\right)^2 w_j \;=\; 2^{(1+c)L}p_j^2 - 2^{\frac{1+c}{2}L+1}p_j + 1 \;=\; t_j^2 - 2t_j + 1\,,$$

where $t_j := 2^{\frac{1+c}{2}L}\,p_j \geqslant 2^{\frac{c}{2}L}$ by (6.5). Therefore,

$$\varepsilon_j^2 w_j \geqslant 2^{cL}\Big(1 + o(1)\Big)\,. \tag{6.7}$$

If $j \leqslant i_0$, then $w_j = k_j \leqslant 2^L$ and

$$\begin{aligned}\varepsilon_j^2 w_j &= \left(\frac{2^{\frac{1+c}{2}L}}{w_j} - p_j\right)^2 w_j = \left(\frac{2^{\frac{1+c}{2}L}}{k_j} - \frac{1}{\sqrt{m_j}}\right)^2 k_j = \left(\sqrt{\frac{2^{(1+c)L}}{k_j}} - \sqrt{\frac{k_j}{m_j}}\right)^2 \\ &\geqslant \left(\sqrt{\frac{2^{(1+c)L}}{2^L}} + O(1)\right)^2 = 2^{cL}\left(1 + O\Big(\sqrt{2^{-cL}}\Big)\right)^2 = 2^{cL}\Big(1 + o(1)\Big).\end{aligned} \tag{6.8}$$

(6.7) and (6.8) together imply that $\varepsilon_j^2 w_j \geqslant 2^{cL}\Big(1 + o(1)\Big)$ for all levels $j \leqslant i_1$. So by (6.6) the probability that (6.3) fails for at least one level is bounded by

$$n \cdot e^{-2 \cdot 2^{cL}(1+o(1))} = e^{-(2+o(1))2^{cL}}, \tag{6.9}$$

and the theorem follows. □

Now we can show that the gap X' between the qOBDD size Y and the OBDD size Z is small for all variable orderings with overwhelming probability. We define $X'_*(f)$ to be the maximal number of nodes that can be deleted from qOBDD$_\pi(f)$ for any variable ordering π. Formally, $X'_*(f) := \max_\pi X'_\pi(f)$.

Theorem 6.3 *For every* $c > 0$*,*

$$\Pr\left(X'_* \geqslant n2^{\frac{1+c}{2}L}\right) \leqslant e^{-(2+o(1))2^{cL}}.$$

Proof. All we have to do is to multiply the probability bound from the fixed variable ordering case by the number of variable orders, which is $n! < n^n = e^{n\ln(n)}$. We get a probability bound of

$$e^{n\ln n} \cdot e^{-(2+o(1))2^{cL}} = e^{-(2+o(1))2^{cL}}, \tag{6.9'}$$

which proves the theorem. □

6.3 Optimal OBDDs

We say that the weak Shannon effect holds if almost all Boolean functions have almost the same size for a certain kind of representation. Combining the large deviation results Theorem 5.4 and Theorem 6.3, we obtain the following corollary.

Corollary 6.4 *Let $Z_*(f)$ denote the minimal OBDD size of a Boolean function f. Then*

$$\Pr\left(\left|Z_* - \mathrm{E}(Y)\right| \geqslant 2n2^{\frac{1+c}{2}L}\right) \leqslant e^{-2^{cL}/4+O(n\log n)},$$

and since $\mathrm{E}(Y) = \Omega(2^L)$, the weak Shannon effect *holds for OBDDs (and qOBDDs) with optimal variable orderings representing random Boolean functions.*

Proof. This follows from $\left|Z_* - \mathrm{E}(Y)\right| \leqslant \left|Y_* - \mathrm{E}(Y)\right| + \left|X'_*\right|$. □

Main Theorem 1.1 for OBDDs with optimal variable orderings now follows easily. In Theorem 4.12 we proved that the strong Shannon effect undergoes periodic phase transitions if we restrict ourselves to the special case of a fixed variable ordering. Since the weak Shanon effect holds – even with a doubly exponential probability bound – this result carries over to the case of optimal variable orderings as well.

7 Other Decision Diagrams

In this Section, we explain how our results carry over to some modifications of the OBDD data structure which have been proposed in the literature.

7.1 Zero-Suppressed Binary Decision Diagrams (ZBDDs)

Zero-suppressed binary decision diagrams (ZBDDs) are a variant of OBDDs with a modified deletion rule, which allows a node to be deleted if and only if its ‘high’ successor is the terminal 0 (hence the name zero-suppressed). ZBDDs were introduced by Minato [23] and have found applications in two-level logic minimisation [10] and various combinatorial problems [11,28,24].

Schröer and Wegener [28] observed that ZBDDs behave quite similar to OBDDs, if random Boolean functions are considered. The analyses of Liaw and Lin [22] and Wegener [30] carry over without major changes. The same is true for our results. Let us explain why.

First, observe that a quasireduced ZBDD is the same as a qOBDD, because both binary decision diagram types use the same merging rule. Therefore, the analysis of qOBDDs we gave in Sections 4 and 5 does not need to be modified.

The key observation is that the modified deletion rule in ZBDDs leads to the same probability distribution of X'_i. We quote from the proof of Theorem 6.2, page 24: “Among the m_j subfunctions which are possible at level j of the qOBDD there are m_{j+1} functions that do not depend essentially on the variable x_j.” For ZBDDs, this sentence should read: "Among the m_j subfunctions which are possible at level j of the qZBDD (= qOBDD) there are m_{j+1} functions g such that $g_{x_j=1} = 0$." For a random Boolean function $g = g(x_j, \ldots, x_n)$, the events $g(0, x_{j+1}, \ldots, x_n) = g(1, x_{j+1}, \ldots, x_n)$ and $g(1, x_{j+1}, \ldots, x_n) = 0$ have the same probabilities, since we consider the uniform distribution. Thus, the results of Section 6 hold for ZBDDs, too. In particular, we have a modified 'Main Theorem 1.1'.

7.2 Ordered Kronecker Functional Decision Diagrams (OKFDDs)

There is another modification which is called *ordered functional decision diagram* (OFDD). OFDDs were introduced by Kebschull, Schubert, and Rosenstiel in [21]. In OFDDs, the Reed-Muller expansion takes over the part of the Shannon expansion. Define $f_0 := f(0, x_2, \ldots, x_n)$ and $f_1 := f(1, x_2, \ldots, x_n)$ and $f_2 := f_0 \oplus f_1$. Then we can expand f as

$$f(\vec{x}) \; = \; f_0(\vec{x}) \, \oplus \, x_1 \wedge f_2(\vec{x}) \, . \tag{7.1}$$

The deletion rule for OFDDs is syntactically the same as for ZBDDs, but now it has a different meaning: A node can be eliminated if and only if the function it represents does not depend on the variable tested there. OFDDs are particularly useful for algorithms that deal with the ring sum expansion [29,14], although standard operations like $\wedge$ and $\vee$ can lead to an exponential blow-up [3].

To define *ordered Kronecker functional decision diagrams* (OKFDDs), we have to consider yet another possibility for functional decomposition, namely

$$f(\vec{x}) \; = \; f_1(\vec{x}) \, \oplus \, \bar{x}_1 \wedge f_2(\vec{x}) \, . \tag{7.2}$$

which leads to a data structure similar to OFDDs. Since there exist functions which have polynomial OBDD size and exponential OFDD size and vice versa, Drechsler, Sarabi, M. Theobald, Becker, and Perkowski [13] combined the three decomposition types into a hybrid data structure. In OKFDDs, each variable is assigned one of the decomposition types (7.1) and (7.2) or the 'usual' Shannon decomposition

$$f(\vec{x}) \; = \; \bar{x}_1 \wedge f_0(\vec{x}) \, \oplus \, x_1 \wedge f_1(\vec{x}) \, . \tag{7.3}$$

Still OKFDDs are a unique representation for each such *decomposition type list* and variable ordering and can be manipulated efficiently (but note the remark on OFDDs above). There exist functions for which OKFDDs are exponentially smaller than both OBDDs and OFDDs [2].

Returning to our results on the strong Shannon effect, the situation is slightly more complicated for OKFDDs, because the choice of the decomposition type list constitutes another potential for minimisation.

Again, quasireduced OKFDDs are the same as qOBDDs. Also, regardless which decomposition type is performed at a level, a node can be deleted if and only if the subfunction it represents does not depend on the variable tested there, so we do not even need to modify the proof of Theorem 6.2 as for ZBDDs.

If the OKFDD is in fact an OFDD, then the same conclusions as for ZBDDs can be made.

For arbitrary decomposition type lists, the key observation is that multipying the doubly exponential probability bounds with either $n!$ or $3^n\,n!$ does little harm, see (6.9') on page 26. Thus, the large deviation results from Section 6 hold for OKFDDs as well (only the o-terms change). This includes the *weak* Shannon effect (Corollary 6.4) and a 'Main Theorem 1.1' for OKFDDs (in which Z_* denotes the optimal OKFDD size).

We remark that there seems to be no way to overcome this difficulty using the second moment method as applied in [30] due to the weaker probability bounds it supplies.

7.3 Free Binary Decision Diagrams

Another interesting generalisation of OBDDs are read-once branching programs, which are also known as free binary decision diagrams (FBDDs). In FBDDs, variables can be tested in arbitrary order, but only once on each evaluation path. Wegener [30] has shown that the strong Shannon effect for FBDDs holds for 'most' values of n. Based upon work presented in this article, Gröpl [15] has shown that there exist certain ranges of n (whose lengths tend to infinity as $n \to \infty$) such that the optimal FBDD size of Boolean function depending on n variables is a constant factor smaller than the optimal OBDD size with high probability (i. e., the strong Shannon effect does not hold). Together with [30] this implies a result similar to Main Theorem 1.1 about the strong Shannon effect in FBDDs.

References

[1] Noga Alon, Joel H. Spencer: The Probabilistic Method; John Wiley & Sons, New York, 1991.

[2] Bernd Becker, Rolf Drechsler, Michael Theobald: OKFDDs versus OBDDs and OFDDs; Int. Colloq. on Automata, Languages and Programming, Springer LNCS 944, 475 – 486, 1995.

[3] Bernd Becker, Rolf Drechsler, Ralph Werchner: On the Relation Between BDDs and FDDs; Information and Computation, vol. 123, no. 2, 185 – 197, 1995.

[4] Y. Breitbart, H. Hunt III, D. Rosenkrantz: On the Size of Binary Decision Diagrams Representing Boolean Functions; Theoret. Computer Sci., vol. 145, 45 – 69, 1995.

[5] K. S. Brace, R. L. Rudell, R. E. Bryant: Efficient Implementation of a BDD Package; Proc. 27th Design Autom. Conf., 40 – 45, 1990.

[6] Randal E. Bryant: Graph-Based Algorithms for Boolean Function Manipulation; IEEE Trans. Computers, vol. C-35, no. 8, 677 – 691, 1986.

[7] Randal E. Bryant: On the Complexity of VLSI Implementations and Graph Representations of Boolean Functions with Application to Integer Multiplication; IEEE Trans. Computers, vol. C-40, 205 – 213, 1991.

[8] Randal E. Bryant: Symbolic Boolean Manipulation with Ordered Binary-Decision Diagrams; ACM Computing Surveys, vol. 24, no. 3, 293 – 318, 1992.

[9] Vašek Chvátal: The Tail of the Hypergeometric Distribution; Discrete Mathematics, vol. 25, 285 – 287, 1979.

[10] Olivier Coudert: Two-Level Logic Minimization: An Overview; Integration, vol. 17, 97 – 140, 1994.

[11] Olivier Coudert: Solving Graph Optimization Problems with ZBDDs; Proc. Europ. Design & Test Conf., 1997.

[12] Devdatt Dubhashi, Desh Ranjan: Balls and Bins: A Study in Negative Dependence; Random Structures and Algorithms, vol. 13, no. 2, 99–124, 1998.

[13] R. Drechsler, A. Sarabi, M. Theobald, B. Becker, M. A. Perkowski: Efficient Representation and Manipulation of Switching Functions Based on Ordered Kronecker Functional Decision Diagrams; Proc. 31st Design Autom. Conf., 415 – 419, 1994.

[14] Rolf Drechsler, Michael Theobald, Bernd Becker: Fast OFDD Based Minimization of Fixed Polarity Reed-Muller Expressions; Proc. Europ. Design Autom. Conf., 2 – 7, 1994.

[15] Clemens Gröpl: Binary Decision Diagrams for Random Boolean Functions; Dissertation, Humboldt Universität zu Berlin, available via http://edoc.hu-berlin.de, 126 pages, 1999.

[16] Clemens Gröpl, Anand Srivastav, Hans Jürgen Prömel: Size and Structure of Random Ordered Binary Decision Diagrams (Extended Abstract); Proc. 15th Symp. Theoret. Asp. Computer Sci., Springer LNCS 1373, 238 – 248, 1998.

[17] Clemens Gröpl, Anand Srivastav, Hans Jürgen Prömel: On the Evolution of the Worst-Case OBDD Size; Information Processing Letters, vol. 77, 1 – 7, 2001.

[18] Mark A. Heap, Melvin Ray Mercer: Least Upper Bounds on OBDD Sizes; IEEE Trans. Computers, vol. 43, no. 6, 764 – 767, 1994.

[19] Anil Kamath, Rajeev Motwani, Krishna Palem, Paul Spirakis: Tail Bounds for Occupancy and the Satisfiability Threshold Conjecture; Random Structures and Algorithms, vol. 7, no. 1, 59–80.

[20] Valentin F. Kolchin, Boris A. Sevast’i͡anov, Vladimir P. Chisti͡akov: Random Allocations; John Wiley & Sons, 1978.

[21] U. Kebschull, E. Schubert, W. Rosenstiel: Multilevel Logic Synthesis Based on Functional Decision Diagrams; Proc. Europ. Design Autom. Conf., 43 – 47, 1992.

[22] Heh-Tyan Liaw, Chen-Shang Lin: On the OBDD-Representation of General Boolean Functions; IEEE Trans. Computers, vol. 41, no. 6, 661 – 664, 1992.

[23] Shin-Ichi Minato: Zero-Suppressed BDDs for Set Manipulation in Combinatorial Problems; Proc. 30th Design Autom. Conf., 272 – 277, 1993.

[24] Shin-Ichi Minato: Binary Decision diagrams and Applications for VLSI CAD; Kluwer Academic Publishers, Boston, 141 pages, 1996.

[25] Christoph Meinel, Thorsten Theobald: Algorithms and Data Struktures in VLSI-Design; Springer Verlag, Berlin, 267 pages, 1998.

[26] Detlef Sieling: The Nonapproximability of OBDD Minimization; Information and Computation, vol. 172, no. 2, 103 – 138, 2002.

[27] Fabio Somenzi: CUDD: Colorado University Decision Diagram Package; available via http://vlsi.colorado.edu/~fabio/CUDD/, 2001.

[28] Olaf Schröer, Ingo Wegener: The Theory of Zero-Suppressed BDDs and the Number of Knight’s Tours; Formal Methods in System Design, vol. 13, 235 – 253, 1998.

[29] C.-C. Tsai, M. Marek-Sadowska: Boolean Matching Using Generalized Reed-Muller Forms; Proc. 31st Design Autom. Conf., 339 – 344, 1994.

[30] Ingo Wegener: The Size of Reduced OBDDs and Optimal Read-Once Branching Programs for Almost All Boolean Functions; IEEE Trans. Computers, vol. 43, no. 11, 1262 – 1269, 1994.