Size and Structure of Random Ordered Binary Decision Diagrams

(Extended Abstract)

Clemens Gröpl*, Hans Jürgen Prömel**, and Anand Srivastav**

Humboldt-Universität zu Berlin
Institut für Informatik
Lehrstuhl Algorithmen und Komplexität

Abstract. We investigate the size and structure of ordered binary decision diagrams (OBDDs) for random Boolean functions. Wegener [Weg94] proved that for "most" values of n, the expected OBDD-size of a random Boolean function with n variables equals the worst-case size up to terms of lower order. Our main result is that this phenomenon, also known as strong Shannon effect, shows a threshold behaviour: The strong Shannon effect *does not* hold within intervals of constant width around the values $n = 2^h + h$, but it *does* hold outside these intervals. Also, the oscillation of the expected and the worst-case size is described. Methodical innovations of our approach are a functional equation to locate "critical levels" in OBDDs and the use of Azuma's martingale inequality and Chvátal's large deviation inequality for the hypergeometric distribution. This leads to significant improvements over Wegener's probability bounds.

1 Introduction

A Boolean function is a mapping $f : \{0,1\}^n \to \{0,1\}$ depending on Boolean variables. Efficient representation and manipulation of Boolean functions is an important issue in many 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 [Bry86,Bry92]). For a given variable ordering the OBDD is uniquely determined, and its size usually depends heavily on the chosen variable ordering. While there exist Boolean functions that have exponential OBDD-size for all variable orderings, many functions encountered in practice have polynomially sized OBDDs for *some* variable orderings. Of course, we would like this to be the typical, i. e. average case. Therefore, a thorough investigation of the relation between average-case and worst-case is recommended by theoretical and practical reasons. To do so, we have to make assumptions on the underlying probability distribution. From a structural point of view, the uniform distribution is most natural. In the following, we will briefly call a Boolean function chosen from the uniform distribution a *random* Boolean function.

* Graduate school "Algorithmische Diskrete Mathematik", supported by Deutsche Forschungsgemeinschaft, grant GRK 219/2-97.

** supported by Deutsche Forschungsgemeinschaft, grant Pr 296/3-2.

Previous Work. Recall that the ordered binary decision diagram (OBDD) of a Boolean function can be defined as the result of the application of the merging *and* the deletion rule to its binary decision tree. In the same way, the quasi-reduced ordered binary decision diagram (qOBDD) is the result of applying only the merging rule to its binary decision tree. (See Fig. 1 and 2.) The variables are tested according to some ordering $(x_1, \ldots, x_n)$ along each computation path. A *level* is the set of all nodes testing a particular variable.

The phenomenon that almost all functions have the same OBBD size as the hardest functions (up to a factor of $1+o(1)$) is called *strong Shannon effect* for random Boolean functions and OBBDs [Weg94]. If almost all functions have the same OBBD size (up to a factor of $1+o(1)$), but possibly smaller than the worst-case size, we say that the *weak Shannon effect* holds. Wegener proved that the weak Shannon effect holds for random Boolean functions and OBDDs: For arbitrary constants $\epsilon > 0$ the probability that the optimal OBDD size of a random Boolean function differs more than $O(n2^{2n/3})$ from the expected size is at most $O(2^{-n/3+\epsilon n})$. He also observed that the strong Shannon effect holds with probability tending to 1 as $h \nearrow +\infty$ for qOBDDs with a uniformly distributed random number $n \in [2^h \mathinner{..} 2^{h+1}-1]$ of levels. Since for almost all functions the size of qOBDDs and OBDDs is the same (up to a factor of $1+o(1)$), this result extends also to OBBDs. Interestingly, these results carry over to read-once branching programs, also called FBDDs.

The proof of Wegener is carried out in two steps. First, the results are proved for an arbitrary, but fixed variable ordering, and in a second step the generalization to arbitrary variable orderings is done using the second moment method. An important methodological innovation in the work of Wegener is the use of *urn experiments* (see [KSC78]) for the estimation of the expectation and the variance of the number of nodes on each OBDD-level.

Results. In Section 3 we show a threshold behaviour for the strong Shannon effect for random Boolean functions and OBDDs with respect to an arbitrary, but fixed variable ordering: The strong Shannon effect *does not* hold within intervals of constant width around the values $n = 2^h + h$, $h \in \mathbb{N}$, but it *does* hold outside these intervals.

In Section 4 we generalize our result to arbitrary variable orderings. This is done by computing large deviations. Unfortunately, the results of [KSC78] on the limit distributions of certain urn experiments cannot be applied in this context. Instead, we derive a special purpose inequality, using Azuma's martingale inequality and invoke Chvatal's bound on the hypergeometric distribution. We show that the probability of "large" de-

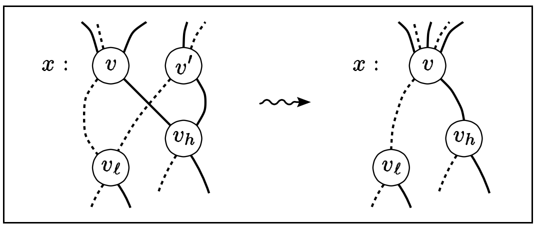

Fig. 1. Merging v and v'

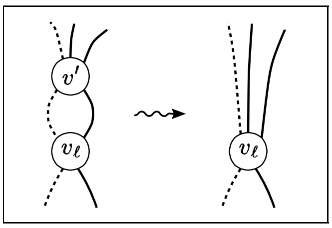

Fig. 2. Deleting v

viations is doubly exponentially small, which is a substantial improvement over Wegener's probability bound.

As a by-product of our proof we show an oscillation of the worst-case and the expected OBDD size. This generalizes and improves result of Heap and Mercer [HM94] and Liaw and Lin [LL92].

In Section 5, we identify those n for which the gap between the expected and the worst-case qOBDD-size is minimal.

Several important proofs are too complicated, technical and lengthy for an extended abstract. Due to space limitations we had to omit them. A full version of this paper is available electronically [GPS97].

2 Preliminaries

It is straightforward to show that the qOBDD of a Boolean function is uniquely determined for each variable ordering, as is the OBDD. Let f be a Boolean function and let qOBDD(f) be its qOBDD with respect to the variable ordering $(x_1, \ldots, x_n)$. 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 $c_1, \ldots, c_{i-1}$. Let Y_i resp. Z_i denote the number of nodes on level i of the qOBDD resp. OBDD of f. Define $k_i := 2^{i-1}$, $m_i := 2^{2^{n-i+1}}$, $m'_i := m_i - m_{i+1}$ $w_i := \min\{k_i, m_i\}$ and $w'_i := \min\{k_i, m'_i\}$. Then the inequalities $Y_i \leqslant w_i$ and $Z_i \leqslant w'_i$ are valid (see [Weg94]).

Let us denote the worst-case size of the whole qOBDD by $W(n) := \sum_{i=1}^{n} w_i$. In Section 3.1 we will derive the precise asymptotic values of $W(n)$ for suitable parametrizations of n. For the moment, we only mention that $W(n) = \Theta(2^n/n)$.

We will need the following inequalities, valid for $x \leqslant 0$.

$$\left.\begin{array}{r} 1+x \\ 1+x+\dfrac{x^2}{2}+\dfrac{x^3}{6} \end{array}\right\} \leqslant e^x \leqslant 1+x+\frac{x^2}{2}\,. \tag{2.1}$$

Finally, let us fix some notation. Intervals of integers will be denoted as $[a\,..\,b] := [a,b] \cap \mathbb{Z}$ and $[a] := [1\,..\,a]$. We write ab/cd for $\frac{ab}{cd}$. The notation $f \sim g$ is equivalent to $f = \big(1+o(1)\big)g$.

3 Strong Shannon Effect – Fixed Variable Orderings

In this section f is a random Boolean function. We consider an arbitrary, but *fixed* variable ordering $(x_1, \ldots, x_n)$ and state all results for this case. The extension to arbitrary variable orderings and the proof of the full statement will be carried out in Section 4 via large deviation inequalities.

Define $X_i := w_i - Y_i$. Y_i is the number of nodes at the i-th level of qOBDD(f) while X_i is the number of nodes that are "missing" at the i-th level compared with the worst-case size w_i. Put $X := \sum_i X_i$, $Y := \sum_i Y_i$ and let $E(X)$ resp. $E(Y)$ denote the expectations.

By Wegener's work we already know that the weak Shannon effect holds for random Boolean functions and OBDDs. More precisely, the OBDD-size of almost all functions *is* the expected qOBDD-size $E(Y)$, up to a factor of $1 + o(1)$. (See also Section 4 of this paper.) Thus for the study of the strong Shannon effect it suffices to compute the ratio $E(Y)/W(n)$ or $E(X)/W(n)$.

Theorem 3.1 (Main Theorem). *Let* $B := \bigcup_{h\in\mathbb{N}}[2^h + h - d(h) \mathinner{..} 2^h + h + d(h)]$ *and* $A := \mathbb{N} \setminus B$.

(i) If $d(h) \nearrow +\infty$ *and* $n \in A$*, then* $E(X)/W(n) = o(1)$*, i. e. the strong Shannon effect holds for random Boolean functions and OBDDs with respect to the variable ordering* $(x_1, \ldots, x_n)$.
(ii) If $d(h) = O(1)$ *and* $n \in B$*, then* $E(X)/W(n) = \Omega(1)$*, i. e. the strong Shannon effect does not hold for random Boolean functions and OBDDs with respect to the variable ordering* $(x_1, \ldots, x_n)$.

In order to prove Theorem 3.1 we compute the worst-case size $W(n)$ and the expectation $E(X)$. The computation of $W(n)$ is done in the next subsection. After this, we analyse $E(X)$, which is much more difficult, and in fact requires new methods, especially the notion of critical levels.

The point i where the two upper bounds k_i and m_i meet turns out to be crucial for all results in this paper. Therefore, let us introduce a concise notation for it. We define the function L by the equation

$$L(n) + \log L(n) = n\,, \tag{3.1}$$

and set

$$i_\delta := L(n) + \delta + 1\,. \tag{3.2}$$

Then

$$k_{i_\delta} = 2^{\delta+L} \quad \text{and} \quad m_{i_\delta} = 2^{2^{-\delta}L}\,. \tag{3.3}$$

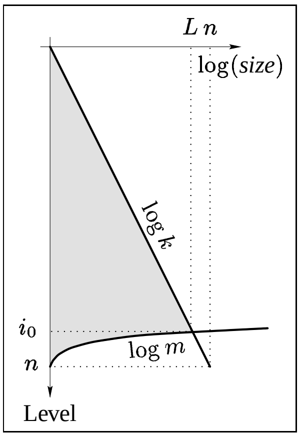

Fig. 3. The worst-case shape of an OBDD

It follows that the critical point is at $i = i_0$, where $k_{i_0} = m_{i_0}$. However, in general i_0 is not integral. It only marks the borderline where the worst-case level width turns from growing exponentially to shrinking doubly exponentially (see Fig. 3). A critical level is an $i_\delta \in \mathbb{N}$, where δ is sufficiently small. We remark that by Definition (3.1), the inverse function of L is $L^{-1}(i) = i + \log i$, and that $L(n) \sim n$.

3.1 The Worst-Case Size of qOBDDs

The behaviour of the worst-case bound $W(n)$ for various parametrizations of n is best expressed in terms of the ratio $W/2^L$. (Previous investigations [LL92,HM94] focused on the ratio $W/\binom{2^n}{n}$, which does have the same asymptotic behaviour, but leads to more cumbersome computations.) The proofs of the results in this section can be carried out without much difficulties using the functional equation for L, but due to space limitations we have to omit them.

The first theorem gives the asymptotic value of $W/2^L$ for parametrizations of n "close" to $2^h + h$. (See Fig. 4.)

Theorem 3.2. *Assume that $a = o(2^h)$. Then*

$$\left(\frac{W}{2^L}\right)\left(2^h + h + a\right) \sim \begin{cases} 2\,, & a \leqslant 0\,; \\ 1 + 2^{-a}\,, & a \geqslant 0\,. \end{cases}$$

Theorem 3.2 is complemented by Theorem 3.3, which describes how $W/2^L$ develops between $2^h + h$ and $2^{h+1} + h + 1$. (See Fig. 5.)

Theorem 3.3. *Assume that $n \nearrow +\infty$ and let h and $b = b(h)$ be chosen such that $n = b2^h + h$ and $b \in [1, 2]$. Then*

$$\left(\frac{W}{2^L}\right)\left(b2^h + h\right) \;=\; b\left(1 + 2^{(1-b)2^h}\left(1 + o(1)\right)\right)\,,$$

which is $\sim b$, if b converges to real number in $]1, 2]$. The convergence of $W/2^L$ is uniform on each interval $[1 + \varepsilon, 2] \ni b$, where $\varepsilon > 0$ is a constant.

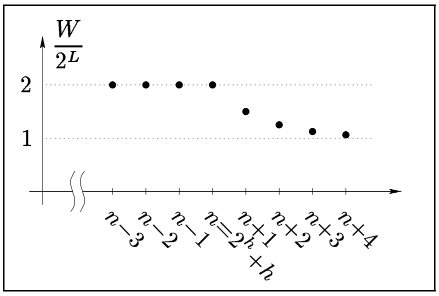

Fig. 4. The worst-case size of OBDDs near $n = 2^h + h$

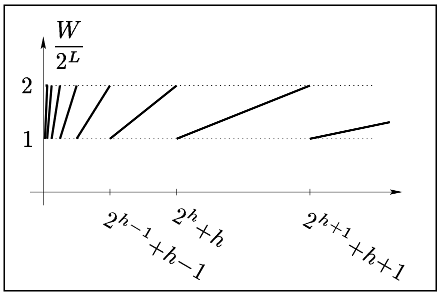

Fig. 5. The oscillation of the worst-case size of OBDDs

Our refined analysis of the oscillation of the worst-case size also leads to improvements over the upper bound of Liaw and Lin [LL92] and the lower bound of Heap and Mercer [HM94], which are valid for all n.[1]

Corollary 3.4. *Let $\varepsilon > 0$ be an arbitrarily small constant. Then*

$$1 \;\leqslant\; \left(\frac{W}{2^L}\right)(n) \;\leqslant\; 2 + O\left(2^{-\frac{1-\varepsilon}{2}n}\right)\,.$$

The lower bound is asymptotically tight for any sequence of the form $n = 2^h + h + a$, where $a \nearrow +\infty$, and $a = o\left(2^h\right)$. The upper bound is attained e. g. for $n = 2^h + h$.

[1] Essentially, the lower bound from [HM94] is $1/2$, and the upper bound from [LL92] is $2 + o(1)$.

3.2 Critical Levels and $E(X)$

Typically, not all nodes that are (numerically) possible at some level will exist. Recall that $X_i = w_i - Y_i$ is the number of nodes that are "missing" at the i-th level of a qOBDD. For each level, $E(X_i)$ can be computed by the following urn experiment (see [KSC78,Weg94]). We 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*. At level i, we are throwing k_i balls into m_i urns, so the expected number of non-empty urns is

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

where $q_i := \Pr(\text{ first urn empty }) = \left(1 - \frac{1}{m_i}\right)^{k_i}$. From this we see that

$$E(X_i) = \begin{cases} k_i - m_i\,(1-q_i)\ , & i \leqslant i_0\,; \\ m_i\,q_i\,, & i \geqslant i_0\,. \end{cases} \tag{3.4}$$

The following lemma contains asymptotics for $E(X_i)$. It is a key for the proof of the main theorem.

Lemma 3.5. *(i) If $\delta \leqslant 0$, then $E\left(X_{i_\delta}\right) \leqslant 2^{\left(2-2^{-\delta}\right)L}$.*
(ii) Denote the "middle" part of an qOBDD with n levels by

$$M_\epsilon(n) := \left[L(n) - \varepsilon\ ..\ L(n) + 1 + \left(\log L(n) - \log\log e + \varepsilon\right)\log e/L(n)\right],$$

where $\varepsilon > 0$ is an arbitrarily small constant. Then $\sum_{i\in[n]\setminus M(n)} E\left(X_i\right) = o(1)$.

So only two levels will be of particular interest for the analysis of the strong Shannon effect.

3.3 Proof of the Main Theorem

There seems to be no way to compute $E(X)/W$ directly, but in view of the results of Section 3.1, we can look at the ratio $E(X)/2^L$ instead. We will show that there is at most *one* "critical" level whose expected width differs significantly from its worst-case width, and that this level must be an $i_\delta \in \mathbb{N}$ for some sufficiently small δ (depending on n). Its existence decides upon whether or not the strong Shannon effect holds. This analysis is summarized in Theorems 3.8 and 3.9, which then lead to the proof of the main theorem.

But before we go into the technicalities, let us demonstrate the idea of the proof in a special case. The next proposition says that the strong Shannon effect does not hold for n of the form $n = 2^h + h$. The ratio of expected and worst-case size is a factor $\leqslant 1 - \frac{1}{2e}$. The tightness of the bounds will be shown in Theorem 3.8.

Proposition 3.6. *It holds*

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

Proof. Recall that $L(2^h + h) = 2^h$, which implies that $i_0 = 2^h + 1 \in \mathbb{N}$ is a level. Observe that $w_{i_0} = k_{i_0} = m_{i_0} = 2^L = 2^{2^h}$. By (3.4), we have

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

and by Theorem 3.2, $W\big(2^h + h\big) = 2 \cdot 2^{2^h}$. □

The proof of Proposition 3.6 was based on the fact that $L\big(2^h + h\big) = 2^h$ is an integer. We will see 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.

Definition 3.7. Let $\delta'(n)$ denote the gap between $L(n)$ and the next natural number, i. e.

$$\delta'(n) := \ell - L(n)\,, \qquad \text{where } \ell \in \mathbb{N} \text{ is such that } \big[\ell - \tfrac{1}{2}, \ell + \tfrac{1}{2}\big[\ni L(n)\,.$$

So $i_{\delta'(n)} = i_0 + \delta'(n) = \ell + 1$ is the integer nearest to i_0. In case of a tie, we round up. Observe that $\delta'(2^h + h) = 0$. As $|\delta'|$ gets larger, $E\big(X_{i_{\delta'}}\big)$ becomes more and more negligible compared to 2^L.

Theorem 3.8. *Let $n \in \mathbb{N}$ and choose $h(n) \in \mathbb{N}$ and $a = a(n) \in \mathbb{Z}$ such that $n = 2^h + h + a$ and $|a|$ is minimal.*

(i) For sequences of n such that $|a(n)| \nearrow +\infty$, $\lim_n \dfrac{E\big(X_{i_{\delta'}}\big)}{2^L} = 0\,.$

(ii) For sequences of n such that $a(n) = a \in \mathbb{Z}$ is a constant,

$$\lim_n \frac{E\big(X_{i_{\delta'}}\big)}{2^L} = \begin{cases} 2^{-a}\left(e^{-2^a} - 1\right) + 1\,, & a \leqslant 0\,; \\ 2^{-a}\, e^{-2^a}\,, & a \geqslant 0\,. \end{cases}$$

Note that the numbers $h(n)$ and $a(n)$ are well-defined, because $2^h + h + a = 2^{h+1} + h + 1 - a$ would imply that $(2^h + 1)/2 = a \in \mathbb{N}$, a contradiction.

Theorem 3.9. *For all sequences of n,*

$$\lim_n \frac{E(X) - E\big(X_{i_{\delta'}}\big)}{2^L} = 0\,.$$

Proof. Take $\varepsilon = \frac{1}{5}$ and observe that for large enough n the middle part $M_\varepsilon(n)$ can only contain the levels $i_{\delta'}$ and $i_{\delta'-1}$. By Lemma 3.5 (ii) we have

$$\lim_n \frac{E(X) - E\big(X_{i_{\delta'}}\big) - E\big(X_{i_{\delta'-1}}\big)}{2^L} = 0\,.$$

Since $\delta' - 1 \leqslant -\frac{1}{2}$, Lemma 3.5 (i) gives

$$0 \leqslant \frac{E\big(X_{i_{\delta'-1}}\big)}{2^L} \leqslant \frac{2^{(2-\sqrt{2})L}}{2^L} = o(1),$$

and the theorem is proved. □

Proof of the main theorem. We write $E(X)/W = \frac{E(X)/2^L}{W/2^L}$. By Corollary 3.4, $1 \leqslant W/2^L \leqslant 3$, so $E(X)/2^L$ gives the asymptotic of $E(X)/W$. By Theorem 3.9, all levels except $i_{\delta'(n)}$ are negligible. We will apply the definition of $a(n)$ from Theorem 3.8.

Assertion (i): Since $d(h) \nearrow +\infty$ and $n \in A$, we have $|a(n)| \nearrow +\infty$. Therefore, $\lim_n E\left(X_{i_{\delta'}}\right)/2^L = 0$, and we are done.

Assertion (ii): Since $d(h) = O(1)$ and $n \in B$, we have $a(n) = O(1)$. By partitioning the sequence of n into subsequences, we may assume that $a(n) = a$ is a constant. The subsequences may have finite or infinite length, but only a finite number of subsequences can be infinitely long, since $a(n) = O(1)$. For each infinite subsequence, we have proved in Theorem 3.8, Assertion (ii) that $\lim_n E\left(X_{i_{\delta'}}\right)/2^L > 0$. So the original sequence satisfies $\liminf_n E\left(X_{i_{\delta'}}\right)/2^L > 0$, i. e., $E\left(X_{i_{\delta'}}\right)/2^L = \Omega(1)$. □

The main theorem is illustrated by Fig. 6.

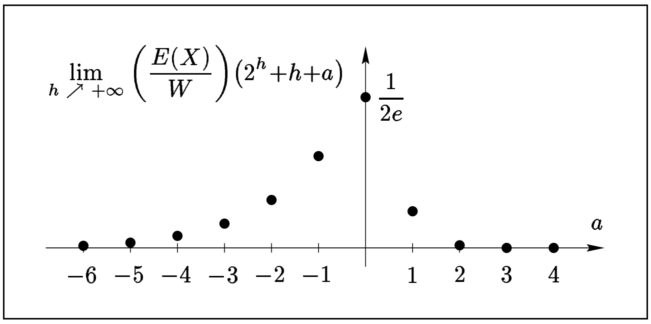

Fig. 6. When the strong Shannon effect does not hold.

4 Strong Shannon Effect – Arbitrary Variable Orderings

For the generalization of Theorem 3.1 to arbitrary variable orderings, we must show that for *almost all* Boolean functions *all* variable orderings lead to an OBDD size which is equal to the expected qOBDD size $E(Y)$ up to a factor of $1+o(1)$. This will be proved in two steps. In Section 4.1 we show that with (very) high probability the qOBDD size for all variable orderings is $E(Y)(1+o(1))$. In Section 4.2 the effect of the deletion rule is taken into account. We show here that with (very) high probability, the size of qOBDDs and OBDDs for all variable orderings is the same (up to a factor of $1+o(1)$). Altogether we obtain the desired generalization of Theorem 3.1 to arbitrary variable orderings.

4.1 The Reduction Effect of the Merging Rule

In this subsection, we study large deviations from the expected qOBDD-size which was computed in Section 3. Define $Y_\pi(f)$ to be the qOBDD size with respect to the variable ordering π. Note that the expected size $E(Y_\pi)$ does not depend on π, as we consider the uniform distribution. From Theorems 3.8 and 3.9, a similar argument as in the proof of the main theorem shows that $E(Y) \geqslant \left(1 - \frac{1}{2e} + o(1)\right)2^L = \Omega(2^n/n)$. Wegener [Weg94] proved that

$$\Pr\left(\exists\pi \; \left|Y_\pi - E(Y)\right| \geqslant O\left(2^{2n/3}/n\right)\right) = O\left(2^{-n/3+\varepsilon n}\right),$$

where $\varepsilon > 0$ is an arbitrary constant. Using Azuma's martingale inequality, we will derive the following result.

Theorem 4.1. *For every constant $c > 0$ and large n,*

$$\Pr\left(\exists\pi \; \left|Y_\pi - E(Y)\right| \geqslant n2^{\frac{1+c}{2}n}\right) \leqslant e^{-2^{cn}}.$$

Proof. First let us consider the variable ordering $\pi = (x_1, \dots x_n)$. At each level i, we have an urn experiment where k_i balls are thrown randomly into m_i urns. Y_i is the number of nodes on the x_i-level as well as the number of non-empty urns. Using Azuma's inequality (see [AS91, Chapter 7]), we find that $\Pr\left(\left|Y_i - E(Y_i)\right| > \lambda\sqrt{k_i}\right) < 2e^{-\lambda^2/2}$. We apply this bound to all levels $i \leqslant n-1$ (level n is negligible). At level i, there are $k_i \leqslant 2^{n-2}$ balls. If we set $\lambda := 2^{\frac{c}{2}n+1}$, we get

$$\Pr\left(\left|Y_i - E(Y_i)\right| \geqslant 2^{\frac{1+c}{2}n}\right) \leqslant \Pr\left(\left|Y_i - E(Y_i)\right| \geqslant 2^{\frac{c}{2}n+1}\sqrt{k_i}\right) \leqslant 2e^{-2^{cn+2}}.$$

Summing over all levels gives the result for any fixed variable ordering. Now we allow arbitrary variable orderings. Summing the probability bound over all $n!$ variable orderings, we get the claimed result, because the additional factor $n!$ is amply absorbed by the factor $e^{-2^{cn+2}}$. □

In other words, we have shown that the weak Shannon effect holds for arbitrary variable orders and qOBDDs, too.

4.2 The Reduction Effect of the Deletion Rule

Wegener already proved that the deletion rule does not yield much reduction when applied to a random qOBDD. In this section we give a more refined analysis and better probability bounds. Let us denote the difference between $|\text{qOBDD}(f)|$ and $|\text{OBDD}(f)|$ using the canonical variable ordering by $X'(f) := |\text{qOBDD}(f)| - |\text{OBDD}(f)|$. Of course, $X' = \sum_{i=1}^n X'_i$, where $X'_i := Y_i - Z_i$ denotes the number of nodes deleted on level i. With the techniques used in Section 3, the expected reduction by the deletion rule can be estimated as follows.

Theorem 4.2. *We have*

$$E\left(X'_{i_\delta}\right) \leqslant \begin{cases} 2^{\delta + \left(1 - 2^{-\delta-1}\right)L}, & \delta \leqslant 0; \\ 2^{2^{-\delta-1}L}, & \delta \geqslant 0. \end{cases}$$

Let $\delta_a := \frac{a \log e}{L(n)}$ and $a^2 = o(L)$. Then $E\left(X'_{i_{\delta_a}}\right) \leqslant \sqrt{2^{L-|a|+o(1)}}$.

As an example for the above theorem, note that $E\left(X'_{i_{-1}}\right) \sim 1/2$ and $E\left(X'_{i_1}\right) \sim 2^{L/4}$. (The tightness of the above bounds requires an extra argument.) The amount of reduction achieved by the deletion rule changes in a periodic way, too. However, the oscillations are not as distinct as for the merging rule.

Define $X'_*(f) := \max_\pi X'_\pi(f)$. (Again, the additional index π indicates the variable order.) Wegener [Weg94] used the second moment method to prove

$$\Pr\left(X'_* \geqslant O\left(2^{2n/3}n\right)\right) = O\left(2^{-n/3+\varepsilon n}\right)$$

for every constant $\varepsilon > 0$. Using Chvátal's inequality [Chv79], we can improve Wegener's result in the following way.

Theorem 4.3. *For every constant $c > 0$ and large n,*

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

5 Minimizing the Gap

Using the tools developed so far, we can also fix those n for which the gap between expected and worst-case size is *minimal* and show how it develops between both extremes. Somewhat surprisingly, a "typical" random qOBDD with $n = 2^h + 2h$ variables has only $E(X) = O\left(n^2\right)$ less than in the worst-case (which is $\Omega(2^n/n)$). But for qOBDD with only *one* level less, the following theorem implies that $E(X) = \Omega\left(2^{0.2786\, n}\right)$, asymptotically. The reason for this jump is that level $i_{\delta'(2^h+2h-1)}$ belongs to the middle part $M(2^h + 2h - 1)$, but $i_{\delta'(2^h+2h)}$ is *not* contained in $M(2^h + 2h)$.

Theorem 5.1.

(i) *Consider OBDDs of random Boolean functions with $n = 2^h + ch$ levels, where $h \nearrow +\infty$ and $c = c(h)$ converges to a real number in $[2, +\infty[$. Then $E(X) = n^{\tilde{c}}$, where $\tilde{c} \sim 2(c-1)$.*

(ii) *Consider OBDDs of random Boolean functions having $n = 2^h + 2h - c$ levels, where Boolean functions $h \nearrow +\infty$ and $c \in \mathbb{N}$ is a positive constant. Then $E(X) = 2^{c'n}$, where $c' = 1 - 2^{-c} \log e + o(1)$.*

6 Discussion

It has been observed by Löbbing, Schröer, and Wegener [SW94,LSW95] that zero-suppressed BDDs, also known as ZBDDs behave quite similar to OBDDs, if random Boolean functions are considered. The analyses of Liaw and Lin [LL92] and Wegener [Weg94] can be carried over without major changes. This is true for our results, too. Recall that ZBDDs can be defined as the result of the application of the merging rule and a modified deletion rule to the decision tree for a given Boolean function. In OBDDs, a node is deleted if both successors represent the same function. In ZBDDs, a node is deleted if its 1-sucessor is the 0-sink. It has been observed [Min93] that ZBDDs are particularly efficient for functions with an on-set consisting mostly of inputs for which many variables are assigned the value 0. However, if random Boolean functions are to be considered, the modified elimination rule for ZBDDs leads to the same probability distribution of X_i'. So the results of Section 4.2 also hold for ZBDDs.

The situation is more complicated with FBDDs. As yet, we know that for $n = 2^h + h$ the minimal FBDD-size is less than $\big(1 - 1/2e + o(1)\big)\, W(n)$ with high probability. It is still unknown whether this is true for an even greater gap than $1/2e$. We conjecture that an analogue to Theorem 3.1 holds for FBDDs, too.

References

[AS91] Noga Alon, Joel H. Spencer: The Probabilistic Method; John Wiley & Sons, New York, 1991.

[Bry86] Randal E. Bryant: Graph-Based Algorithms for Boolean Function Manipulation; IEEE Transactions on Computers, vol. C-35, no. 8, 1986, 677 – 691.

[Bry92] Randal E. Bryant: Symbolic Boolean Manipulation with Ordered Binary-Decision Diagrams; ACM Computing Surveys, vol. 24, no. 3, 1992, 293 – 318.

[Chv79] Vašek Chvátal: The Tail of the Hypergeometric Distribution; Discrete Mathematics, vol. 25, 1979, 285 – 287.

[GPS97] Clemens Gröpl, Anand Srivastav, Hans Jürgen Prömel: Full version of this extended abstract, available via http://www.informatik.hu-berlin.de/Institut/struktur/algorithmen/forschung/veri/shannon.ps.gz .

[HM94] Mark A. Heap, Melvin Ray Mercer: Least Upper Bounds on OBDD Sizes; IEEE Transactions on Computers, vol. 43, no. 6, 1994, 764 – 767.

[KSC78] Valentin F. Kolchin, Boris A. Sevast'i͡anov, Vladimir P. Chisti͡akov: Random Allocations; John Wiley & Sons, 1978, Chapter 1.

[LL90] Heh-Tyan Liaw, Chen-Shang Lin: On the OBDD-Representation of General Boolean Functions; NSC Rep., NSC79-0404-E002-35, 1990.

[LL92] Heh-Tyan Liaw, Chen-Shang Lin: On the OBDD-Representation of General Boolean Functions; IEEE Transactions on Computers, vol. 41, no. 6, 1992, 661 – 664.

[LSW95] Martin Löbbing, Olaf Schröer, Ingo Wegener: The Theory of Zero-Suppressed BDDs and the Number of Knights Tours; Proceedings of the IFIP WG 10.5 Workshop on Applications of the Reed-Muller Expansion in Circuit Design (Reed-Muller '95), August 27. – 29. 1995, Makuhari, Chiba, Japan, 1995, 38 – 45.

[Min93] Shin-Ichi Minato: Zero-Suppressed BDDs for Set Manipulation in Combinatorial Problems; Proceedings of the 30$^{\text{th}}$ ACM/IEEE Design Automation Conference, 1993, 272 – 277.

[SW94] Olaf Schröer, Ingo Wegener: The Theory of Zero-Suppressed BDDs and the Number of Knights Tours; Forschungsbericht Nr. 552/1994, Universität Dortmund, Fachbereich Informatik, 1994, 25 pages.

[Weg94] Ingo Wegener: The Size of Reduced OBDDs and Optimal Read-Once Branching Programs for Almost All Boolean Functions; IEEE Transactions on Computers, vol. 43, no. 11, 1994, 1262 – 1269.