TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

D I P L O M A R B E I T

Random Boolean Functions Induced by Random Boolean Expressions

Ausgeführt am Institut für
Diskrete Mathematik und Geometrie
der Technischen Universität Wien

unter der Anleitung von

Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Bernhard Gittenberger

durch

Clemens Haim
Neumayrgasse 22/3
1160 Wien

Datum

Unterschrift

D I P L O M A T H E S I S

Random Boolean Functions Induced by Random Boolean Expressions

Author:
Clemens Haim

Supervisor:
Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Bernhard Gittenberger

Abstract

We consider a set of Boolean expressions with a probability measure on it and call this our model. This model induces a probability measure on the Boolean functions. The induced probability depends strongly on the underlying Boolean expressions, so we consider several different sets of Boolean expressions, i.e. different models, distinguished by the connectors they are built with and/or by the structure of the expressions. We examine the relation between the underlying class of Boolean expressions and the induced probability and investigate the shape and its properties. We study the question if this probability measure has a limit when the size of the underlying Boolean expressions tends to infinity. The aim of this thesis is to give a broad overview of several different set-ups for such models considered in the literature with a slight focus on models that allow interesting conclusions regarding the number of satisfiable assignments of read-one expressions. Furthermore we develop an abstract model that unifies the models studied in the literature.

Zusammenfassung

In dieser Arbeit interessieren wir uns für Wahrscheinlichkeitsverteilungen auf Booleschen Funktionen in n Variablen, welche durch eine Wahrscheinlichkeitsverteilung auf Booleschen Ausdrücken induziert werden. Dabei sind die Eigenschaften der induzierten Wahrscheinlichkeitsverteilung von den zu Grunde liegenden Booleschen Ausdrücken abhängig. Daher betrachten wir verschiedene Klassen von Booleschen Ausdrücken, die wir bezüglich der vorkommenden logischen Operationen beziehungsweise bezüglich ihrer Struktur einteilen. Wir untersuchen den Zusammenhang zwischen den Eigenschaften der induzierten Wahrscheinlichkeitsverteilung und den verwendeten Booleschen Ausdrücken und untersuchen wann diese Wahrscheinlichkeitsverteilung eine Grenzverteilung hat. Das Ziel dieser Arbeit ist es dem Leser einen guten Überblick über bestehende Literatur und darin untersuchte Modelle zu geben, wobei wir dabei einen leichten Fokus auf Modelle legen, welche interessante Schlussfolgerungen bezüglich des Erfüllbarkeitsproblem im Kontext von Read-Once-Funktionen erlauben. Außerdem wird auch ein abstraktes Modell entwickelt, welches die in der Literatur studierten Modelle vereinheitlicht.

Acknowledgments

First of all I wish to thank my supervisor, Dr. Bernhard Gittenberger, most sincerely for an excellent academic as well as organizational support during the whole process of writing this thesis and finishing my studies.

I also want to thank my whole family, especially my parents, Daniela and Andreas, and my sisters, who supported me my whole live and throughout my studies and who aroused my great interest for natural sciences in my childhood.

I am also grateful to my girlfriend, Smaralda Ursu. With her I shared all joys and sorrows a student can have during his studies. Moreover she was my most precise proofreader and supported me whenever it was possible.

Table of Contents

Chapter 1

Introduction

Consider a model of random Boolean expressions. Every Boolean expression computes a Boolean function and therefore the model induces a distribution on the Boolean functions. Such models of Boolean functions often occur in the field of theoretic computer science; at first as a helpful tool for deriving upper bounds on the complexity of special Boolean functions and graphs and later as an autonomous research field: The first to utilize such models was Valiant [Val84] who used it to prove the existence of small monotone Boolean expressions computing the majority function. The proof mainly consists of approximating a Boolean function by iteratively constructing a Boolean expression with a single connector. This method was later called the amplification method, see [Bop85]. Razborov [Raz88] used random expressions for showing an upper bound on the formula size complexity of special graphs. Savický was the first to study models of Boolean expressions per se [Sav87, Sav90]. Paris et al. [PVW94] used random Boolean expressions with connectives $\wedge$, $\vee$ and literals $\{x_1, \bar{x_1}, ..., x_n, \bar{x_n}\}$, called Catalan-And/Or tree model, to construct a natural prior probability distribution on the interval $[0, 1]$ by considering the expected values of large Boolean expressions. They were the first who investigated the Catalan tree model although their motivation for examining this model came mainly from the field of uncertain reasoning.

The area of research of random Boolean expressions received heightened interest since Woods' paper on coloring rules for trees [Woo97]. He proved, as a consequence of a far more general statement, the existence of the limit distribution on Boolean functions when the size of the underlying Catalan

trees approaches infinity. His paper inspired the ongoing research in two ways: He showed that analytic combinatorics lend themselves for this kind of considerations and he formulated in [Woo97, Problem 6.6.] the question whether and how the probability of a Boolean function occurring in the limit distribution is related to the formula size complexity[1] of the function. This question was communicated to Savický who found together with Lefmann an affirmative answer to it in [LS95]. Moreover, since then such questions have been investigated in the literature on numerous occasions in relation to a different range of models of Boolean functions induced by Boolean expressions. See [CFGG04, GW05, Koz08] for the chronological development regarding the And/Or-Catalan model and [GGKM15] that takes commutativity and/or associativity into account as well as the early survey [Gar06] for an informative overview. Similar models for expressions allowing only the implication were investigated in the same context [FGGG12, GG10, GGKM12].

Around the year 2000 a group of Polish researchers started to systematically analyze the density of tautologies in different logical systems with a view to comparing classical and intuitionistic logic quantitatively, see e.g. [Zai03, FGGZ07, GK09] and again the survey of Gardy [Gar06] for a first overview on this topic and more references. The structures that were found alongside the main concepts and methods used in their research are very similar to those in the study of random Boolean expressions. Both areas did profit from each others' results which manifested itself in the several works that were jointly undertaken by leading scientists in both fields. In [FGGZ07] it is proven that in the implication model (Propositional formulas with the single connector being the implication) asymptotically all tautologies are intuitionistic ones and $\frac{5}{8}$ for the full propositional system, see [GK09].

The aim of this thesis is to provide an overview of the at first glance very manageable seeming area of random Boolean functions induced by (large) random expressions. It aims to demonstrate the most important results achieved in this area. Upon closer inspection there are two different lines taken in the literature. One was initiated by the search for good upper bounds for special Boolean functions, the so called amplification method. The models of random functions/expressions used there are built from a few functions that are successively composite with themselves so that they amplify themselves

[1]Here this is the smallest Catalan tree that expresses the function

to a special function with high probability. The depth of the expressions constructed in this case is growing continuously in contrast to the models reviewed in the other line of research, where all expressions of a given size (that is not the depth) built from a distinguished set of connectors and variables were considered. The two models have very different behaviors and the differences are going to be demonstrated and captured. It shows where the models have been applied or where they might be applicable.

Going into further detail, the structure of the thesis is as follows. In the second chapter the methodological background is given and the basic objects of interest are defined. The most elementary notions can be found in the appendix for completeness. Subsequently we will specify in Chapter 3 the notion of random Boolean function models induced by random expressions in a very abstract way, so one can see everything that follows from this point of view. Moreover, the questions that may arise concerning these models and which have been studied in the literature are stated on an abstract level so that it is easier for the reader to maintain an overview of the numerous different models that are being considered. In the fourth chapter we will examine the first class of models, namely random Boolean functions induced by expressions built from a set of connectors and a set of variables[2]. There one looks at expressions of a given size and then examines the behavior for big expressions, meaning that the size of the expressions tends to ∞. The size function has to satisfy some regularity conditions, in the moment we can think of the size as the number of connectors. This model class will prove to have a very regular behavior and the general requirements for falling under that regime are elaborated. After this universal approach some concrete models that have been investigated in the literature are presented and the most important results are stated. A model introduced by Yashunskii [Yas05], that considers the Boolean values of random expressions for random assignments, is investigated at the end of that chapter and a very interesting connection in relation to satisfiability of read-once formulas (c.f. Definition 2.8) is pointed out. In Chapter 5 the second model class of our interest will be introduced. In contrast to the first class, the size function of the expressions is the depth. So a large Boolean expression in this context is a Boolean expression of large depth. As a consequence such models exhibit

[2]In the general setting we will not consider only variables as the inputs of the expressions but also Boolean functions.

a far different behavior. Also the methods that are used to analyze such models are of different shape as we will see. After introducing the set-up for this model-class and demonstration of the structural differences by means of some simple examples, we will see where the origin of the research of such models comes from and then we will investigate some more recent results. In this thesis we will lay the focus on the first model class and investigate it in more detail. Nevertheless, we will also provide a good overview and thorough treatment of the second class. In the last part of the thesis we will very briefly present two more models that are strongly connected to the other models under consideration and two models that show far less connections, but are still related. After that we will conclude the thesis and additionally will state some open questions.

Chapter 2

Basic definitions, notions and used concepts

In this chapter the main definitions and notions that will be used throughout this work are provided. These are the definitions of Boolean functions, Boolean expressions and Catalan-And/Or trees. Moreover, we will present the modus operandi in enumerative combinatorics, the symbolic method and singularity analysis, which refer to some basic concepts that are at the very core of analytic combinatorics. This will be essential for the analysis of the questions arising within this work, especially for the first model under consideration, and for the treatment of the structures that we are going to investigate. The last section will briefly introduce the complexity of a Boolean function. Most of the readers will be very familiar with the following, nevertheless, it is presented here for the sake of completeness and exactness. The first part reviews Boolean functions and Boolean expressions and should not be skipped because they are the main objects used in this thesis and some of the definition might be unusual and have its subtlety. Moreover, one needs to take into consideration:

> **"The advanced reader who skips parts that appear too elementary may miss more than the less advanced reader who skips parts that appear too complex."**-G.Pólya [Pó54]

2.1 Boolean functions and Boolean expressions

Let us start with the notion of a Boolean function, where a brief discussion and a table of the most important Boolean functions is also presented in A.1. Let $\mathbb{N}$ denote the natural numbers starting from 1 and $\mathbb{N}_0$ the natural numbers with 0.

Definition 2.1 (Boolean function)
A Boolean function is a function from $\{0,1\}^{\mathbb{N}}$ *to* $\{0,1\}$ *and is denoted by* f. $\mathbb{B}$ *is the set of all Boolean functions.*

For the input variable of a Boolean function we use the symbol x and $x_1, x_2, ...$ for its entries. So $\vec{x} = (x_1, x_2, ...)$ and the $x_i, i \in \mathbb{N}$, are called the variables of the function. Let f be a Boolean function and $j \in \{0,1\}$; with $f_{[x_i \mapsto j]}$ we mean the Boolean function that is defined pointwise according to: $f_{[x_i \mapsto j]}((x_1, x_2, ..., x_{i-1}, x_i, x_{i+1}, ...)) = f_{[x_i \mapsto j]}((x_1, x_2, ..., x_{i-1}, j, x_{i+1}, ...))$.

Definition 2.2 (Essential variable)
A variable $x_i, i \in \mathbb{N}$, *of a Boolean function* f *is an essential variable when the function value depends on* x_i, *i.e. when there are two inputs differing only in variable* x_i *such that the images of these inputs are different. This is equivalent to:* $f_{[x_i \mapsto 0]}$ *is not the same function as* $f_{[x_i \mapsto 1]}$.

A Boolean function f, having all its essential variables in the set $\{x_1, x_2, ..., x_n\}$, can be trivially seen as a function $\hat{f}$:$\{0,1\}^n \to \{0,1\}$ with variables $\{x_1, x_2, ..., x_n\}$: $\hat{f}(x_1, x_2, ..., x_n) = f(x_1, x_2, ..., x_n, 0, 0, 0, ...)$. Let us call such a function f finite Boolean function or also just Boolean function if it is clear from the context and define the arity of it as n:

Definition 2.3 (Finite Boolean function)
A finite Boolean function f *with arity* n, *(arity*$(f) = n$*) is a function from* $\{0,1\}^n$ *to* $\{0,1\}$ *and is denoted by* f. $\mathbb{B}_n$ *is the set of all Boolean functions of arity* n.

Observe that for example $x_1 \wedge x_2$ is not a Boolean function as long as we do not specify the arity of it. And that $\top : (x_1, x_2, x_3) \mapsto 1$, the constant function TRUE, is a Boolean function with arity 3. A finite Boolean function $\hat{f}$ in $\mathbb{B}_n$ can be injectively identified with a(n) (infinite) Boolean function f:

$$f(x_1, x_2, ..., x_n, x_{n+1}, ...) := \hat{f}(x_1, x_2, ..., x_n)$$

or with a Boolean function of $\mathbb{B}_k, k \geq n$ in a similar way. A variable x_i, $i \in \{1, 2, ..., n\}$ is called essential variable of $\hat{f} \in \mathbb{B}_n$ when the corresponding $f \in \mathbb{B}$ has x_i as essential variable. Most of the time the considered Boolean function space will be clear from the context, but we should keep these identifications and embeddings in mind. In Appendix A.1 there is also a discussion on this topic and a table with the most important Boolean functions that we take into consideration.

As the title of this thesis suggests, we are interested in Boolean expressions and extensions of these. So let us define inductively what we understand by a Boolean expression:

Definition 2.4 (Connectors $\mathbb{K}$, base functions $\mathbb{F}$)
Let $\mathbb{K}$ be a finite set of finite Boolean functions with arities > 0 called the connectors. Let $\mathbb{F}$ be a finite set of finite Boolean functions, called the basis.

A Boolean expression with connectors in $\mathbb{K}$ over the Basis $\mathbb{F}$ can be defined in several ways.

Definition 2.5 (Boolean expression, inductive)
A Boolean expression with connectors in $\mathbb{K}$ over the Basis $\mathbb{F}$, is either:

- *an element of $\mathbb{F}$.*
- *or $(c, e_1, ..., e_i)$[1], $i \in \mathbb{N}$, if $e_1, e_2, ..., e_i$ are Boolean expressions with connectors in $\mathbb{K}$ over the Basis $\mathbb{F}$ and c is a connector in $\mathbb{K}$ with* arity(c) = i.

[1] In this context the tuple $(c, e_1, ..., e_i)$ is also written as $c(e_1, e_2, ..., e_i)$ when no misunderstanding occurs, if $c(e_1, e_2, ..., e_i)$ is either an expression or a function.

Definition 2.6 (Boolean expression, graph definition)
A Boolean expression with connectors in $\mathbb{K}$ over the Basis $\mathbb{F}$ is a plane(c.f. Definition A.3) tree with leaves labeled by elements from $\mathbb{F}$ and inner nodes labeled by elements from $\mathbb{K}$ such that the in-degree of every inner node equals the arity of the connector used for labeling this node. Due to this definition we call Boolean expressions also Boolean trees.

The set of Boolean expressions with connectors in $\mathbb{K}$ over the Basis $\mathbb{F}$ will be denoted by $\mathbb{E}_{\mathbb{K},\mathbb{F}}$ and the set of all such Boolean expressions by $\mathbb{E}$, so $\mathbb{E} := \bigcup_{\mathbb{K},\mathbb{F}} \mathbb{E}_{\mathbb{K},\mathbb{F}}$. We will use the terms (Boolean) expression, tree and formula synonymously depending on the actual point of view. In Figure 2.1 there is an example of two different Boolean expressions. The left has connectors in $\mathbb{K} = \{\vee, \neg, \oplus, \wedge\}$ and base functions in $\mathbb{F} = \{\top, \bot, \overline{x_1}, x_2, {}_1\wedge_3\}$, whereas ${}_1\wedge_3$ is the function represented by the expression $x_1 \wedge x_3$, the AND on x_1, x_3. The right expression has connectors in $\mathbb{K} = \{\vee, \neg, \oplus, \wedge\}$ and base functions in $\mathbb{F} = \{\top, \bot, x_1, \overline{x_1}, x_2, x_3\}$. Both expressions indeed represent the same Boolean function but in our setting they are different expressions.

Our first definition of expressions is similar to the definition of a Boolean expression in propositional logic and conforms more with the intuitive view of an expression and also shows one way to identify a Boolean expression e with a Boolean function: Let n be the number of different variables used in the expression. If the expression e is a function from the basis then we identify the expression with this function. In the other case the expression $c(e_1, e_2, ..., e_i)$ is recursively evaluated by

$$c(e_1, e_2, ..., e_i)((x_1, x_2, ..., x_n)) = c(e_1(x_1, x_2, ..., x_n), e_2(x_1, x_2, ..., x_n), ..., e_i(x_1, x_2, ..., x_n)),$$

whereas all e_j belong to $\mathbb{B}_n$ (maybe after embedding). When e is an expression, we denote the unique function that is computed by e by f_e.

The second definition will be our point of view of Boolean expressions because it seems that for our purpose this descriptive definition is more accessible[2]. Such a Boolean tree computes a Boolean function by evaluating all nodes of

[2]The concepts that are used to analyze Boolean expressions regarding questions arising in this work use some kind of extensions, expansions and pruning and terms like sprouting and growing which are all very descriptive when taking the point of view of trees

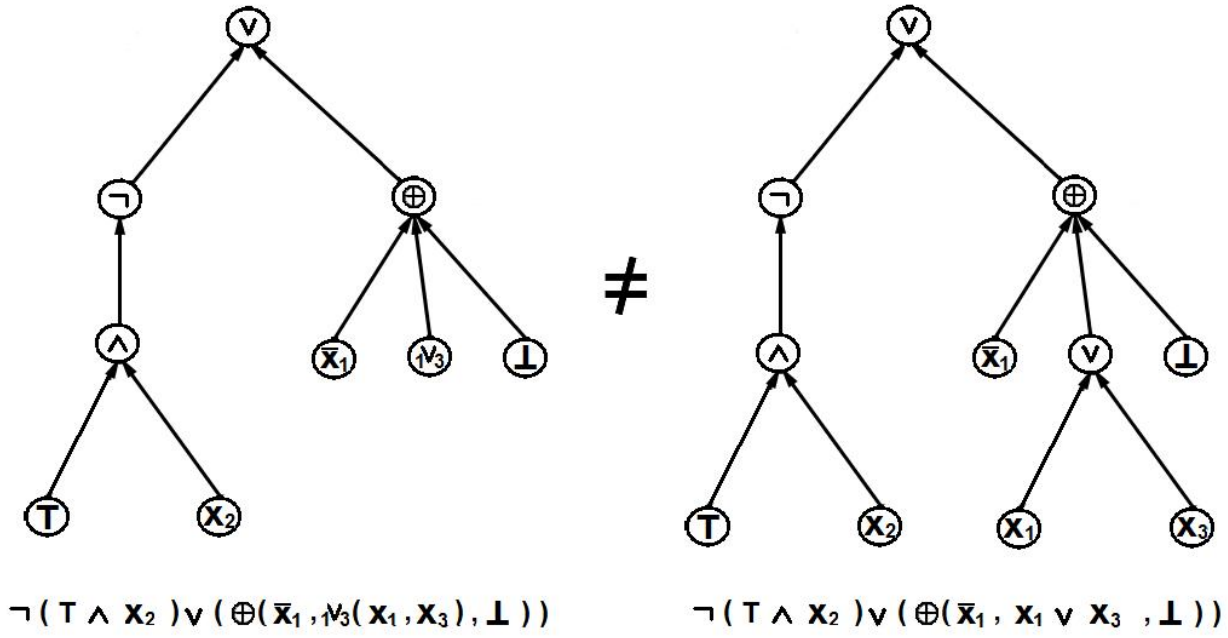

Figure 2.1: Two different expressions.

the tree starting with the leaves by plugging in the input values. Both definitions are trivially seen to be combinatorially equivalent[3] and corresponding expressions also represent the same Boolean function. The inductive character of the first definition is maintained in the inductive description of trees, which is worked with and analyzed throughout this thesis. Moreover, the definition with trees easily gives rise to a more general definition of Boolean 'expressions', namely the one of Boolean circuits (c.f. Definition A.5) which will be briefly discussed later in this work. Before we continue employing more main concepts we should define the Catalan-And/Or trees, the expressions that are most intensively studied in the literature as well as the concept of read-once formulas resp. functions.

[3]That means that there is a size preserving bijection, which is stated more precisely in the next pages.

Definition 2.7 (Catalan-And/Or trees)
The set of Catalan-And/Or expressions or trees with n variables is the set of Boolean expression with connectors in $\mathbb{K} = \{\wedge, \vee\}$ over the basis $\mathbb{F} = \{x_1, x_2, ..., x_n, \overline{x_1}, \overline{x_2}, ..., \overline{x_n}\}$, where $\wedge$ and $\vee$ are the logical AND respective OR functions from $\{0,1\}^2$ to $\{0,1\}$, as usual.[4] *In other words: Catalan-And/Or expressions are binary plane trees with nodes labeled by $\wedge$ and $\vee$ and leaves labeled by literals $\{x_1, x_2, ..., x_n, \overline{x_1}, \overline{x_2}, ..., \overline{x_n}\}$.*

Definition 2.8 (Read-once formula/expression and function)
Let $\mathbb{K}$ be an arbitrary set of connectors and $\mathbb{F} = \{x_1, ..., x_n\}$. An expression $e \in \mathbb{E}_{\mathbb{K},\mathbb{F}}$ is called read-once formula/expression when every variable x_i occurs at most once as a leaf of e. If only positive literals occur in the formula then we call such a formula a monotone read-once formula. A function f is called read-once function (w.r.t $(\mathbb{K}, \mathbb{F})$) if there is a read-once expression $e \in \mathbb{E}_{\mathbb{K},\mathbb{F}}$ with $f_e = f$.

2.2 Enumerative combinatorics

In this section we will introduce the main tools that we are going to use for our treatment of the objects reviewed above. At a later stage it will become clearer that we are interested in the number of expressions of a given shape and size. We, therefore, want to find the answer to questions like: How many Boolean expressions with connectors in $\mathbb{K}$ with m inner nodes ($\hat{=}$ connectors) are there or how fast/slow does this number grow with m, if it grows at all? Related questions for arbitrary structures are investigated in enumerative combinatorics. In the following section a brief introduction to the concepts used in this field will be given.

2.2.1 Counting and the symbolic method

As previously explained we are interested in the behavior of Boolean functions coming from large Boolean expressions respectively Boolean trees. So we need a way to measure the size. There are a lot of different possibilities

[4] Here is a small subtlety. Due to the associativity of $\wedge$ and $\vee$ one could interpret both operations as functions from $\{0,1\}^k$ to $\{0,1\}$ for any $k \geq 2$. For a fixed k this would lead to k-ary And/Or trees.

to feasibly define the size. The size measure needs to meet some very basic properties to make it accessible to the so called symbolic method. The following section is strongly based on the book 'Analytic Combinatorics' by Philippe Flajolet and Robert Sedgewick [FS09, I.1] but can also be found in the introductory chapters of many other textbook on combinatorics.

Definition 2.9 (Combinatorial class)
A finite or countable infinite set $\mathcal{A}$ with a size function $\mu_{\mathcal{A}} : \mathcal{A} \mapsto \mathbb{N}_0$, with property

- $|\mu^{-1}(m)| \in \mathbb{N}$,

meaning that the number of elements $a \in \mathcal{A}$ having size $m \in \mathbb{N}$ is finite, is a combinatorial class, or just class.

Given a combinatorial class, we define the number of elements $a \in \mathcal{A}$ with size m as a_m (resp. b_m if the class is $\mathcal{B}$ etc.). Throughout this work it will be our goal to examine these numbers a_m; at least for combinatorial classes having some structure that allows us to do so. The numbers are either tried to be evaluated exactly or their asymptotic behavior is elaborated. We are, therefore, going to count the number of objects in a combinatorial class of a given size m. For this counting it might help to count an 'equivalent' class that might be more descriptive:

Definition 2.10 (Combinatorial equivalence, counting)
A combinatorial class $\mathcal{A}$ is said to be combinatorially equivalent to class $\mathcal{B}$ when $a_m = b_m \ \forall m \in \mathbb{N}_0$. This is equivalent to the existence of a size preserving bijection between the classes.

Now we can define the ordinary generating function, short generating function or GF, of a combinatorial class as a formal power series:

Definition 2.11 (Ordinary generating function)
Given a combinatorial class $\mathcal{A}$. The ordinary generating function of $\mathcal{A}$ is the

formal power series

$$A(z) = \sum_{m=0}^{\infty} a_m z^m$$

We say that z is marking the size in the generating function. The number of elements from $\mathcal{A}$ with size m is the coefficient of z^m in $A(z)$, written $[z^m]A(z) = a_m$. Regarding combinatorial equivalence (counting) it is sufficient to know the GF of a combinatorial class: Two combinatorial classes are combinatorial equivalent iff their generating functions are identical. This means that we are not losing any information in the combinatorial sense when working with GFs instead of the classes. The same concept works also when one is interested in marking multiple parameters by introducing additional variables. We call such GFs multidimensional GFs, whereas we only need GFs that take trace of one additional parameter so that such two-dimensional GFs have the form

$$A(z) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{m,n} z^m y^n.$$

Our later analysis requires the following property of GFs:

Definition 2.12 (Aperiodicity)
The period of a formal power series $g(z)$ is the biggest number $d \in \mathbb{N}$ such that $g(z) = z^r h(z^d)$, with $h(z)$ a power series and $r \in \mathbb{N}_0$. A power series is said to be aperiodic if the period is 1. With $E := \{n \in \mathbb{N}_0 | [z^n]g(z) \neq 0\}$ and $r := \min(E)$ the period can be computed as the greatest common divisor of $E - r := \{n - r | n \in E\}$. A power series $g(z)$ is aperiodic iff there $\exists i < j < k$ with $g_i := [z^i]g(z), g_j := [z^j]g(z), g_k := [z^k]g(z) \neq 0$ and the greatest common divisor of $j - i, k - i$ is 1.

Our objects of interest are combinatorial classes:

Example 2.13 (Boolean expressions)
The class of Boolean expressions according to Definition 2.5 with connectors in $\mathbb{K}$ over the basis $\mathbb{F}$ with the size being the number of occurrences of elements in $\mathbb{K}$ is a combinatorial class. The class of Boolean expressions according to Definition 2.6 with connectors in $\mathbb{K}$ over the basis $\mathbb{F}$ with the

size being the number of internal nodes is a combinatorial class. These two classes are combinatorially equivalent. Other possible size measures, which satisfy the properties needed for combinatorial classes, are: number of leaves, total number of nodes (internal nodes plus leaves), number of occurrences of functions from a distinguished subset of $\mathbb{K}$ or $\mathbb{F}$ (or weighted versions), depth etc. $\triangle$

The vast advantage of using GF is that operations on combinatorial classes translate directly to simple computations with their GFs, see [Wil94, chapter 2] for an introduction to formal power series and manipulations of them. This important ingredient to combinatorics is called the symbolic method. First define the sum of two formal power series, $A(z) = \sum_{m=0}^{\infty} a_m z^m$, $B(z) = \sum_{m=0}^{\infty} b_m z^m$ as:

$$A(z) + B(z) := \sum_{m=0}^{\infty} (a_m + b_m) z^m.$$

and the product as:

$$A(z)B(z) := \sum_{m=0}^{\infty} \sum_{i=0}^{m} (a_i b_{m-i}) z^m.$$

The first combinatorial operation is the union:

Fact 2.14 (Combinatorial union)
Given two disjoint combinatorial classes $\mathcal{A}$, $\mathcal{B}$ and a combinatorial class $\mathcal{C}$. Suppose that $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$ and

$$\mu_{\mathcal{C}}(c) = \begin{cases} \mu_{\mathcal{A}}(c) & \text{if } c \in \mathcal{A} \\ \mu_{\mathcal{B}}(c) & \text{if } c \in \mathcal{B} \end{cases}$$

Then the GF of $\mathcal{C}$ is the sum of the GFs of $\mathcal{A}$ and $\mathcal{B}$, i.e.

$$C(z) = A(z) + B(z).$$

Now the Cartesian product of combinatorial classes:

Fact 2.15 (Cartesian product)
Given three combinatorial classes $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. Suppose that $\mathcal{C} = \mathcal{A} \times \mathcal{B} = \{(a,b)|a \in \mathcal{A}, b \in \mathcal{B}\}$ and

$$\mu_{\mathcal{C}}(c) = \mu_{\mathcal{C}}((a,b)) = \mu_{\mathcal{A}}(a) + \mu_{\mathcal{B}}(b).$$

Then the GF of $\mathcal{C}$ is the product of the GFs of $\mathcal{A}$ and $\mathcal{B}$, i.e.

$$C(z) = A(z)B(z).$$

With this formal background we should be able to count the number of Catalan-And/Or trees with m leaves, so we are reviewing the combinatorial class of Catalan-And/Or trees with size function being the number of leaves.

Example 2.16 (Catalan-And/Or trees, exact)
Let $\mathcal{T}$ be the combinatorial class of Catalan-And/Or trees and $\mathcal{L}$ the set of leaves labeled by literals from $\{x_1, x_2, ..., x_n, \overline{x_1}, \overline{x_2}, ..., \overline{x_n}\}$, both equipped with the size being the number of leaves[5]. A Catalan-And/Or tree with i variables is either a literal from $\{x_1, x_2, ..., x_n, \overline{x_1}, \overline{x_2}, ..., \overline{x_n}\}$, so an element of $\mathcal{L}$, or a root labeled by an $\vee$ ($\hat{=}$ ⓥ) with left and right subtree being Catalan-And/Or trees or a root labeled by an $\wedge$ ($\hat{=}$ ⓐ) with left and right subtree being Catalan-And/Or trees. This translates to the symbolic description of Catalan-And/Or trees:

$$\mathcal{T} = \mathcal{L} \cup \mathcal{T} \times \{ⓥ\} \times \mathcal{T} \cup \mathcal{T} \times \{ⓐ\} \times \mathcal{T}.$$

Observe that $L(z) = 2nz$, because all the $2n$ Objects in $\mathcal{L}$ have one leaf, and that the generating function from ⓥ and ⓐ are both $1 = 1z^0$ because they are internal nodes and no leaves. With Fact 2.14 and Fact 2.15 this leads to an equation for the generating functions:

$$T(z) = 2nz + T(z)1T(z) + T(z)1T(z).$$

[5]One could also take the size to be the number of internal nodes or as the sum of leaves and internal nodes to get the same results with shifts $n \mapsto n-1$ respective $n \mapsto 2n-1$ because a binary tree with n leaves has $n-1$ internal nodes. Also for k-ary trees, $k \geq 2$ (so for trees with internal nodes having exactly k children) such a shift is applicable and the size measures above can be transformed easily, but for general trees these size measures are substantially different.

Using manipulation rules for formal power series this leads to the two solutions $T(z)_{1,2} = \frac{1 \pm \sqrt{1-16nz}}{4}$. Because $T(0) = t_0$ is the number of Catalan-And/Or trees of size 0, which is 0, $T(z)$ finally evaluates to

$$T(z) = \frac{1 - \sqrt{1-16nz}}{4}.$$

Looking in a table of generating functions coefficients as in [Wil94, chapter 2.5] and elementary calculations lead to the coefficients

$$t_0 = 0,\ t_m = [z^m]T(z) = 2^{m-1}(2n)^m C_{m-1},\ m \geq 1,$$

with the Catalan number $C_m = \frac{(2m)!}{m!(m+1)!}$. This formula was first given in [PVW94]. △

2.2.2 Analytic combinatorics - Singularity analysis

The next step from combinatorics to analytic combinatorics is to interpret a GF, primary a formal power series, as an analytic function around 0.[6] This is not always possible because the coefficients of the power series might grow too fast[7] for convergence around 0 as it is the case for e.g. the number of directed graphs (cf. A.2) with m nodes, of which there are 2^{4m^2}. If the investigated GFs are analytic functions, then the very powerful approaches of analytic combinatorics can be used to analyze the asymptotic behavior of the coefficients of the GFs and therefore to count the combinatorial classes. P. Flajolet and R. Sedgewick describe the general approach of analytic combinatorics in two principles [FS09, p.227]:

- **First Principle of Coefficient Asymptotics.**
  The location of a function's singularities dictates the exponential growth of its coefficients.
- **Second Principle of Coefficient Asymptotics.**
  The nature of a function's singularities determines the associated subexponential factor.

[6] The readers who are not familiar with elementary complex analysis find the necessary background in [Rud87], but also [FS09] offers the basic concepts of elementary complex analysis needed for analytic combinatorics.

[7] Too fast here means superexponential.

A singularity is a point in the complex plane where the function ceases to be analytic and a function $\theta(m)$ from $\mathbb{N}_0$ to $\mathbb{R}$ is called subexponential when $\limsup |\theta(m)|^{\frac{1}{m}} = 1$. The principles say that the coefficients of an analytic GF $A(z)$ are a composition of an exponential growing factor A^m and a subexponential factor $\theta(m)$, i.e. $[z^m]A(z) = A^m\theta(m)$, and is determined by the location and type of the singularities. The first principle is specified with the following:

Fact 2.17 (Exponential growth, [FS09, Theorem IV.7])
If $A(z)$ is a GF (non-negative coefficients) analytic at 0 and R is the radius of the singularity nearest to the origin:

$$R := \sup\{r \geq 0 | A \text{ is analytic at all points of } 0 \leq z < r\}^8$$

then the coefficient $a_m = [z^m]A(z)$ satisfies

$$a_m = R^{-m}\theta(m)$$

with $\theta(m)$ being subexponential.

The real singularity nearest to the origin is also called (real) dominant singularity. For extracting the subexponential factor one substitutes z with $\frac{z}{R}$ so that we can w.l.o.g. assume that the dominant singularity is at 1. Regarding growth rates and the asymptotic behavior of functions related to coefficients of GFs, the $\mathcal{O}$-notation is convenient. For readers who are not familiar with this notation a short summary can be found in the Appendix A.3. The second principle claims that the subexponential factor of the coefficients depends only on the nature of the singularity. So it suggests that if we know the behavior of the GF near its singularity, then we know the associate subexponential factor. More concrete if $f(z)$ admits an expansion around its dominant singularity σ of the form

$$f(z) \underset{z\to\sigma}{\sim} g(z),$$

then

$$[z^m]f(z) \underset{m\to\infty}{\sim} [z^m]g(z)$$

[8] Pringsheim's Theorem guaranties that if a GF has a singularity then it has a real singularity with no other singularity nearer to the origin. Nevertheless, it might be the case that there are several (complex) singularities with the same distance to the origin.

or similarly

$$f(z) \underset{z\to\sigma}{=} o(g(z)) \quad \longrightarrow \quad [z^m]f(z) \underset{m\to\infty}{=} o([z^m]g(z)),$$

$$f(z) \underset{z\to\sigma}{=} \mathcal{O}(g(z)) \quad \longrightarrow \quad [z^m]f(z) \underset{m\to\infty}{=} \mathcal{O}([z^m]g(z)).$$

These suggestive transfers of the behavior of a function near its dominant singularity to the behavior of the coefficients manifest themselves in the next three results due to Flajolet and Odlyzko [FO90] with an expansion and good presentation in [FS09]. Before we can state the precise conditions, on which the asymptotic approximation of a GF near its singularity transfers to an asymptotic approximation for the coefficients, we need the definition of a Δ-domain

Definition 2.18 (Δ-domain, [FS09, Definition VI.1])
Given two numbers ϕ, R with $R > 1$ and $0 < \phi < \frac{\pi}{2}$, the domain $\Delta(\phi, R)$ is defined as

$$\Delta(\phi, R) := \{z \mid |z| < R,\ z \neq 1,\ |arg(z-1)| > \phi\}$$

and is called a Δ-domain. A function is Δ-analytic if it is analytic in some Δ-domain.

Fact 2.19 (Transfers: big-Oh, little-Oh and $\sim$, [FS09, Section VI. 3])
Let α be an arbitrary real number, not a negative integer and 0; and let $f(z)$ be a Δ-analytic function with Δ-domain Δ.

(i) Assume that $f(z) = \mathcal{O}((1-z)^{-\alpha})$, as $z \to 1$, $z \in \Delta$. Then one has:

$$[z^m]f(z) = \mathcal{O}(m^{\alpha-1}).$$

(ii) Assume that $f(z) = o((1-z)^{-\alpha})$, as $z \to 1$, $z \in \Delta$. Then one has:

$$[z^m]f(z) = o(m^{\alpha-1}).$$

(iii) Assume that $f(z) \sim (1-z)^{-\alpha}$, as $z \to 1$, $z \in \Delta$. Then one has:

$$[z^m]f(z) \sim \frac{m^{\alpha-1}}{\Gamma(\alpha)}.$$

All three statements together show that the transfer of the behavior of a GF near its singularity to its coefficients as suggested holds for the class of functions above. The authors prove these results by means of Cauchy's integration formula for the coefficients from complex analysis and a special integration contour called Hankel contour. It is also shown that a similar result holds for a more general class of functions, namely for functions of the form $(1-z)^{-\alpha}(\log\frac{1}{1-z})^{\beta}$, but we will not encounter such functions. In this thesis we will encounter the case where $\alpha = \frac{1}{2}$ and we will say that a function $T(z)$ has a square-root singularity at ρ (or square root singular expansion around ρ) if

$$f(z) = a - b\sqrt{1-\frac{z}{\rho}} + \mathcal{O}(1-\frac{z}{\rho})$$

around ρ. A direct application of the above transfers is the following result:

Lemma 2.20 ([GGKM15, Lemma 3.4])
Let $T(z)$, $S(z)$ be GF with the same unique dominant singularity ρ and with square-root singular expansions around ρ. Then, $\lim_{m\to\infty}\frac{S_m}{T_m} = \lim_{z\to\rho}\frac{S'(z)}{T'(z)}$.

With these results we are very well prepared for our further analysis of the Catalan-And/Or model and models encountered in the fourth chapter, and the reader will agree with the validity of the quote:

"Analytic methods are extremely powerful and when they apply, they often yield estimates of unparalleled precision."-A.Odlyzko [Odl95]

Example 2.21 (Catalan-And/Or trees, asymptotics)
In 2.16 we calculated the GF of Catalan-And/Or trees as

$$T(z) = \frac{1-\sqrt{1-16nz}}{4}.$$

This function is analytic for $|z| < \frac{1}{16n}$ and has an algebraic singularity at $z = \frac{1}{16n}$. Fact 2.17 gives $(16n)^m$ for the exponential growth factor of the

coefficients. Substituting $z = \frac{x}{16n}$ in $\frac{1-\sqrt{1-16nz}}{4}$ gives

$$T(x) = \frac{1-(1-x)^{\frac{1}{2}}}{4}.$$

With the $\sim$-transfer above one finally gets

$$t_m = [z^m]T(z) = (16n)^m[x^m]T(x) \sim (16n)^m \frac{m^{-\frac{3}{2}}}{2\sqrt{\pi}} \frac{1}{4}$$

as an asymptotically equivalent expression for the number of Catalan-And/Or trees. We will use that

$$t_m = \mathcal{O}((16n)^m) m^{-\frac{3}{2}}.$$

$\triangle$

The square-root singularity and the associated subexponential factor of form $(\frac{1}{m})^{\frac{3}{2}}$ are very often encountered in combinatorics, especially when counting tree structures. In the words of Bell, Burris, and Yeats [BBY06]

"almost any family of trees defined by a recursive equation that is nonlinear [. . .] lead[s] to an asymptotic law of the Pólya form $\mathbf{t(n) \sim C\rho^{-m} m^{-\frac{3}{2}}}$"

There is a generic underlying principle why these square-root singularities emerge in a lot of cases:

Fact 2.22 (Implicit functions, [FS09])
Suppose that $G(z, y)$ is an analytic function

$$G(z, y) = \sum_{m,n \geq 0} g_{m,n} z^m y^n$$

in a domain $|z| < R$ and $|y| < S$, for some $R, S > 0$ such that

$$g_{0,0} = 0, \quad g_{m,n} \geq 0, \quad g_{0,1} \neq 1, \quad \exists m \exists n \geq 2 : g_{m,n} > 0 \tag{2.1}$$

and that the characteristic system

$$\begin{aligned} y &= G(z, y) \\ 1 &= G_y(z, y) \end{aligned} \tag{2.2}$$

has a solution (z_0, y_0) in the domain of analyticity of G(z,y) and $z_0, y_0 > 0$. Then, there is an unique solution y(z), analytic for $|z| < z_0$ and convergent at $z = z_0$, where it has a square-root singular expansion:

$$y(z) \underset{z\to z_0}{=} y_0 - \gamma\sqrt{1-\frac{z}{z_0}} + \mathcal{O}(1-\frac{z}{z_0}), \quad \gamma := \sqrt{\frac{2z_0 G_z(z_0,y_0)}{G_{ww}(z_0,y_0)}}$$

that is valid in a Δ-domain. Additionally, if $y(z)$ is aperiodic, then we have the asymptotic expansion for the coefficients of $y(z)$ of the form

$$[z^m]y(z) \underset{m\to\infty}{=} \frac{\gamma}{2\sqrt{\pi m^3}}\left(\frac{1}{z_0}\right)^m (1+\mathcal{O}(m^{-1})).$$

The so called Drmota-Lalley-Woods theorem is one facet of the 'generality' of the square-root singularity. Before we study the theorem in the fourth chapter we will finish this chapter with a brief discussion of the basic concepts of combinatorial complexity theory.

2.3 Combinatorial complexity theory

In computer science a main focus of studies is to perform tasks with as little resource as possible, e.g perform a computation, store an object etc. Imagine you have a set of abstract objects/tasks $\mathcal{T}$ and a set of possible representations/solutions $\mathcal{S}$ for this task with a function $\mu : \mathcal{S} \mapsto \mathbb{N}$ that realizes the notion of resources needed for a representation/solution in $\mathcal{S}$. Moreover, you have a function $\tau : \mathcal{S} \mapsto \mathcal{T}$ that maps a representation/solution to the object/task it represents/solves. That means every representation/solution represents/solves exactly one task, but maybe there are multiple representations/solutions for one object/task or even none. With this the complexity $\mathcal{L}$ is defined as a function from $\mathcal{T}$ to $\mathbb{R} \cup \{+\infty\}$ as

$$\mathcal{L}(t) := \min\{\mu(s) | \tau(s) = t, s \in \mathcal{S}\}$$

and we say that t has complexity $\mathcal{L}(t)$ (in $\mathcal{S}$). If there is no r with $\tau(r) = t$ then $\mathcal{L}(t) = \infty$ meaning that t cannot be represented/solved in $\mathcal{S}$. We are going to use this notion for $\mathcal{S}$ being a combinatorial class and μ being its size measure. More concrete we are representing Boolean functions

f by Boolean expression from a combinatorial class $\mathcal{C}$ of Boolean expressions. The complexity $\mathcal{L}(f)$ of f is then the smallest size (according to the size measure μ) of an expression e from $\mathcal{C}$ that represents this function, so $\mathcal{L}(f) = \min\{\mu(e)|f_e = f, e \in \mathcal{C}\}$. In one single sentence: The complexity of a Boolean function with respect to $\mathcal{C}$ is the smallest size of a Boolean expression in $\mathcal{C}$ representing it. The complexity of Boolean functions is a broad field of research, the standard reference in this field [Weg87] is a recommended starting point to explore this area in more depth.

Chapter 3

Random Boolean functions induced by random Boolean expressions

In this chapter the general object of interest that we are going to study is defined, namely models of Boolean functions induced by random Boolean expressions. A huge number of such models has been defined and analyzed since the 90s, all of them having small differences and their own subtlety. We introduce an abstract model, including the models examined in the literature, and pose questions that may arise.

3.1 General model and the Shannon effect

In this section we define a general model of Boolean functions that offers a framework in which almost all models regarded in this thesis can be reviewed. Moreover, by means of this framework we will demonstrate what properties we are going to investigate and what different points of view one can take.

Definition 3.1 (Random Boolean function model)
Let $\mathbb{P}_{\mathbb{S}} : \mathbb{S} \mapsto [0,1]$ *be a probability function on* $\mathbb{S} \subseteq \mathbb{E}$. *We define a probability function* $\mathbb{P}$ *on the set of all Boolean functions* $\mathbb{B}$ *as the image probability of the function* $e \mapsto f_e$*:*

$$\mathbb{P}(f) := \mathbb{P}_{\mathbb{S}}(\{e \in \mathbb{S} | f_e = f\}) = \sum_{e \in \mathbb{S}: f_e = f} \mathbb{P}_{\mathbb{S}}(e)$$

and call the new probability space model of random Boolean functions or short random functions and denote it by $(\mathbb{S}, \mathbb{P}_{\mathbb{S}})$. *The set* $\mathbb{S}$ *is called the domain of the model and* $\mathbb{P}_{\mathbb{S}}$ *its distribution.*

W.l.o.g. we assume that $\mathbb{P}_{\mathbb{S}}$ is positive, meaning that every expression in $\mathbb{S}$ really occurs (has probability > 0) in the model. Observe that every model in our definition has the same domain for the induced probability function, namely $\mathbb{B}$. Since an expressions is a finite object, there are only finitely many variables occurring in an expression. This means that in our model only Boolean functions f with a finite number of essential variables have a positive probability $\mathbb{P}(f)$, so we could take the set of Boolean functions having only finitely many essential variables, $\mathbb{B}_{\leq\infty}$, as our domain for $\mathbb{P}$. Moreover, in most cases we are going to investigate, there is only a finite number of variables $\{x_1, ..., x_n\}$ used in all expressions from $\mathbb{S}$, then we can also identify $\mathbb{P}$ as a probability function on $\mathbb{B}_n$. We might also have a sequence of models $(\mathbb{S}_m, \mathbb{P}_{\mathbb{S}_m})_{m\in\mathbb{N}}$ or want to implement the notion of size in our model, meaning that we have a size function μ on $\mathbb{S}$, making $\mathbb{S}$ a combinatorial class. Such a model will be indicated by $(\mathbb{S}, \mathbb{P}_{\mathbb{S}}, \mu)$. Consequently a couple of questions arise concerning such models:

- Which Boolean functions have positive probability in the model $(\mathbb{S}, \mathbb{P}_{\mathbb{S}})$?
- What is the 'typical' behavior of a random Boolean function.
- What is the average or 'typical' complexity of a function in the model $(\mathbb{S}, \mathbb{P}_{\mathbb{S}}, \mu)$?
- What is the 'typical' shape of an expression computing f.
- Does the sequence $(\mathbb{S}_m, \mathbb{P}_{\mathbb{S}_m})_{m\in\mathbb{N}}$ converge to a limit distribution $\mathbb{P}_\infty$?
- If so, what is the shape of it, which functions have positive probability, can we compute it explicitly, does it converge and how fast?
- If it does not converge, can we find out something about its asymptotic behavior? For example:
  - What is the asymptotic behavior of the probability of a fixed function?

  - What is the asymptotic behavior of the average or maximal complexity of a function?
  - Are most/all functions of high or low complexity?
  - etc.
- etc.

The aim of this work is to investigate different random Boolean function models that have been examined in the literature and to answer at least some of the questions posed above. Which functions have positive probability, thus really occur in the model, depends on the used expressions. When we consider the expressions $\mathbb{E}_{\mathbb{K},\mathbb{F}}$ with connectors in $\mathbb{K}$ and base functions $\mathbb{F}$ as $\mathbb{S}$, then the functions with positive probability are exactly the ones that are expressible by means of $\mathbb{K}$ and $\mathbb{F}$. This gives rise to the definition of the set of Boolean functions induced by $(\mathbb{K}, \mathbb{F})$ and the notion of completeness:

Definition 3.2 ($\mathbb{B}_{\mathbb{K},\mathbb{F}}$, Completeness)
The set of all functions induced by the set of connectors $\mathbb{K}$ *and the set of base functions* $\mathbb{F}$ *is*

$$\mathbb{B}_{\mathbb{K},\mathbb{F}} := \{f \in \mathbb{B} | \exists e \in \mathbb{E}_{\mathbb{K},\mathbb{F}} : f_e = f\}$$

and we call the tuple $(\mathbb{K}, \mathbb{F})$ *complete iff* $\mathbb{B}_{\mathbb{K},\mathbb{F}} = \mathbb{B}_n$*, with* n *being the number of different variables used in* $\mathbb{F}$*. We also say that the set of expressions* $\mathbb{E}_{\mathbb{K},\mathbb{F}}$ *is complete.*

That the tuple $(\{\wedge, \vee\}, \{x_1, x_2, ..., x_n, \overline{x_1}, \overline{x_2}, ..., \overline{x_n}\})$ is complete is folklore[1], so for every function on n variables there is a Catalan-And/Or expression computing this function. Before we proceed with introducing subclasses of random Boolean function models, let us consider the most natural model, the uniform distribution on $\mathbb{B}_n$, with the complexity of a function being the size of the smallest And/Or formula expressing the function and examine on it the so called Shannon effect for Boolean formula size complexity:

Example 3.3 (Uniform distribution and the Shannon effect)
In terms of our definition above we consider $(\mathbb{S}_n, \mathbb{P}_{\mathbb{S}_n}, \mu)$, with $\mathbb{S}_n$ being the

[1] Consider the disjunctive or conjunctive normal form of a function.

Catalan-And/Or trees with n variables, $\mathbb{P}_{\mathbb{S}_n}$ an appropriate probability function that induces the uniform distribution on $\mathbb{B}_n$[2] and with $\mu(t)$ being the number of leaves of a tree t. Let us denote the probability induced on the Boolean functions as $\mathbb{P}_n$. Since there are 2^{2^n} functions in $\mathbb{B}_n$ and $\mathbb{P}_n$ is the uniform distribution on this set, $\mathbb{P}_n(f) = \frac{1}{2^{2^n}}$ and $\mathbb{P}_n(f) \underset{n\to\infty}{\to} 0$, so the function converges in distribution to the limit function $\mathbb{P}_\infty = 0$ and is, therefore, not a probability measure. Concerning the tree size complexity of a formula we now establish an upper bound for this complexity and examine the question of how many functions in $\mathbb{B}_n$ approximately adopt this value. We follow the presentation of [FS09, page 77]: An upper bound for the complexity of a Boolean function f in n variables is $2^{n+1} - 2$. To see this, consider the representation of f

$$f(x_1, ..., x_{n-1}, x_n) = (\overline{x_n} \wedge f(x_1, ..., x_{n-1}, 0)) \vee (x_n \wedge f(x_1, ..., x_{n-1}, 1)),$$

used for Boolean decision trees. With this and the four functions

$$\top \equiv (\overline{x_1} \vee x_1), \quad \bot \equiv (\overline{x_1} \wedge x_1), \quad \overline{x_1}, \quad x_1,$$

having complexity at most 2 we get the (upper bound) recursion

$$c_{n+1} \leq 2c_n + 2,\ c_1 \leq 2$$

for c_n being the complexity of a function in n variables. Solving this recursion we get the claimed upper bound. Lupanov [Lup60] showed a stronger upper bound for the complexity: $\frac{2^n}{\log_2(n)}(1 + o(1))$. For examining a lower bound for the complexity of most functions, i.e. of a fraction tending to 1 when n tends to ∞, we compare the number of Catalan-And/Or expressions of size at most m, say $T_{\leq m}$, (this is an upper bound of the number of functions in $\mathbb{B}_n$ with complexity at most m) to the total number of functions in $\mathbb{B}_n$, which is 2^{2^n}. If we find a(n) (hopefully big) $m(n)$ depending on n such that $T_{\leq m(n)}$ is little-Oh of $|\mathbb{B}_n|$, i.e. $\frac{T_{\leq m(n)}}{|\mathbb{B}_n|}$ tends to 0, then we have established a lower bound $m(n)$ on the number M that fulfills: A fraction tending to 1 when n tends to ∞ has at least complexity M. Let us make the corresponding calculations: According to 2.21 the number of Catalan-And/Or expressions

[2]This is possible because $\mathbb{K} = \{\wedge, \vee\}$ with $\mathbb{F} = \{x_1, x_2, ..., x_n, \overline{x_1}, \overline{x_2}, ..., \overline{x_n}\}$ is a complete tuple.

of size m is $T_m = \mathcal{O}(16^m n^m m^{-\frac{3}{2}})$ and easy computations show that also $T_{\leq m}$ as a function of m is of that order:

$$T_{\leq m} := \sum_{i=1}^{m} T_i = \mathcal{O}(16^m m^n m^{-\frac{3}{2}}).$$

When we choose $m(n)$ to be

$$m(n) := \frac{2^n}{4 + \log_2(n)}$$

then $T_{\leq m(n)}$ is indeed little-Oh of 2^{2^n}. So we established the following result due to Riordan and Shannon [RS42] $\triangle$

Fact 3.4 (Shannon effect for the uniform distribution)
Let $\mathbb{P}_n$ be the uniform distribution on $\mathbb{B}_n$. There is a function $g(n) = o(1)$ so that

$$\mathbb{P}_n(\{f \in \mathbb{B}_n | \mathcal{L}(f) \geq \frac{2^n}{\log_2(n)}(1 - g(n))\}) \longrightarrow 1$$

when n tends to ∞.

Regarding the upper bound of Lupanov this means that the probability/ratio of functions having almost maximal complexity tends to 1. This gives rise to the following definition:

Definition 3.5 (Shannon effect, [Weg87, Section 4.1])
Let $\mathcal{A}_n$ be a sequence of finite sets with probability measures $\mathbb{P}_n$ on it and complexity functions $\mathcal{L}_n : \mathcal{A}_n \to \mathcal{R} \cup \infty$ and let $\mathcal{L}_{max}(n)$ be the maximal complexity of an element in $\mathcal{A}_n$, i.e. $\mathcal{L}_{max}(n) := \max\{\mathcal{L}_n(\mathcal{A}_n)\}$, then we say that this system exhibits the Shannon effect iff there is a function $g(n) = o(1)$

$$\mathbb{P}_n(\{e \in \mathcal{A}_n | \mathcal{L}_n(e) \geq \mathcal{L}_{max}(n)(1 - g(n))\}) \longrightarrow 1$$

when n tends to ∞.

There are also stronger and weaker definitions of the Shannon effect occurring in the literature, see [Weg87] for the uniform distribution and [GGM14,

Definition 6]. A similar result holds for circuits with the quantity $\frac{2^n}{n}$ instead of $\frac{2^n}{\log_2(n)}$ and over $\mathbb{K} = \{\wedge, \vee, \neg\}$ and $\mathbb{F} = \{x_1, ..., x_n\}$. That was again established by Lupanov [Lup58](general upper bound) and Shannon [Sha49](lower bound for most functions). At this point it should be mentioned that the very broad field of Boolean function complexity works on establishing exactly such general lower and upper bounds as above and also for special functions as the majority- and parity-function. Mainly this happens for circuit size complexity, which is the smallest size of a Boolean circuit of a given class representing the function, but also for formulas. For a very up-to-date book in Boolean function complexity that concentrates on lower bounds see [Juk12].

In the next tree chapters, which represent the main part of this thesis, we are am introducing subclasses of random Boolean function models that we are going to investigate in detail as specifications of the general model.

Chapter 4

DLW Models

This chapter will review the first subclass of the random Boolean function model that we are considering. This class will fulfill the structural requirements for being accessible for singularity analysis, especially for the so called Drmota-Lalley-Woods Theorem (DLW Theorem) which will be introduced later. The DLW Theorem requires models of a certain structure and therefore we call the models regarded in this chapter DLW models. These models will turn out to exhibit a lot of regularity, smoothness and robustness.

4.1 The model

In this part we will restrict ourselves to special domains $\mathbb{S}$ of the model and size functions μ so that the structure of the combinatorial class is well suited. As domain of this subclass we allow expressions with fixed connectors $\mathbb{K}$ and base functions $\mathbb{F}$ of a given size $m \in \mathbb{N}_0$, so

$$\mathbb{S}_m := \{e \in \mathbb{E}_{\mathbb{K},\mathbb{F}} | \mu(e) = m\}.$$

To start with we will consider the probability on the expressions to be the uniform distribution:

$$\mathbb{P}_{\mathbb{S}_m}(e) := \frac{1}{|\mathbb{S}_m|}.$$

Then the probability of a Boolean function f is the ratio of expressions of size m computing f, to the total number of expressions:

$$\mathbb{P}_m(f) := \mathbb{P}_{\mathbb{S}_m}(\{e \in \mathbb{S}_m : f_e = f\} = \frac{|\{e \in \mathbb{S}_m : f_e = f\}|}{|\mathbb{S}_m|}. \tag{4.1}$$

With $S(z)$ being the GF of all expressions from $\mathbb{E}_{\mathbb{K},\mathbb{F}}$ and $f(z)$ being the GF of the expressions computing f, this can be written as

$$\mathbb{P}_m(f) = \frac{[z^m]f(z)}{[z^m]S(z)}. \tag{4.2}$$

Later in this section we will look at a generalization of the uniform distribution on the expressions. In this model every function f which has an expression of size m in $\mathbb{S}_m$ has positive probability. Since the base functions are fixed, only a finite number of variables occur in this model, say $\{x_1, ..., x_n\}$. Therefore, only functions on these variables have positive probability and we can restrict the domain of $\mathbb{P}_m$ to $\mathbb{B}_n$. We are interested in the shape of the model for large expressions, i.e. large m. It is not a priori clear if this question has an answer, because we have no guarantee that for two different, large numbers m_1, m_2 the models $(\mathbb{S}_{m_1}, \mathbb{P}_{m_1})$, $(\mathbb{S}_{m_2}, \mathbb{P}_{m_2})$ have a common shape. However, as we will see in this section, the question is well posed and permits an answer for a very general set-up. The answer will be the following: The probability measure $\mathbb{P}_m$, on the Boolean functions $\mathbb{B}_n$, converges in distribution to a probability measure $\mathbb{P}_\infty$, called limit distribution. Because the domain of $\mathbb{P}_m$ is $\mathbb{B}_n$ which has only finitely many elements this convergence is uniform. In this sense the shape of the model for large m is $\mathbb{P}_\infty$. So, this would answer the question how the induced probability of large Boolean expressions built with internal nodes from $\mathbb{K}$ and leaves from $\mathbb{F}$ looks like. As mentioned in [CFGG04] one could also be interested in the ratio of expressions computing f among all expressions of size $\leq m$. The probability of a function f can then be computed as

$$\mathbb{P}_{\leq m}(f) = \frac{\sum_{i=0}^{m}[z^i]f(z)}{\sum_{i=0}^{m}[z^i]S(z)} = \frac{[z^m]f(z)/(1-z)}{[z^m]S(z)/(1-z)}. \tag{4.3}$$

$\frac{1}{1-z}$ introduces a singularity at 1, but since both functions $f(z)$, $S(z)$ have the same singularity $\sigma < 1$ (see later), the asymptotic expansion of the coefficients does not change due to singularity analysis and it holds that both probability distributions are the same in the limit, whereas it is clear that $\mathbb{P}_{\leq m}$ and $\mathbb{P}_m$ are different.

The number of variables, occurring in functions of $\mathbb{F}$, is finite and not growing. So another dimension that we wish to include in our considerations, to achieve the goal to say something about models induced by large expressions, is to

consider somehow growing or developing base functions and/or connectors. Assume we have a sequence of sets of expressions

$$\mathbb{E}_{\mathbb{K}_n,\mathbb{F}_n}$$

with $\mathbb{F}_n$ and $\mathbb{K}_n$ sequences of connectors resp. base functions such that for all n the limit distribution $\mathbb{P}_{\infty,n}$ exists. Then, one can be interested in the question of the behavior of the limit probabilities of a function:

"What is the behavior of $\mathbb{P}_{\infty,n}(f)$ for large n."

The above considerations for the Catalan And/Or tree model (see below) have been extensively discussed in the literature. See [LS95] for the start of the research and [CFGG04, GW05, Koz08, GGM14, GM15] for a consecutive development and improvement on results in this area. Variants of this model including associativity/commutativity and/or a different size measure were investigated in [GGKM15, GGKM14]. To make this abstract procedure more concrete we demonstrate it by means of the Catalan And/Or model.

Example 4.1 (Catalan And/Or model, uniform distribution)
Consider the Catalan And/Or trees on n variable: $C_n := \mathbb{E}_{\mathbb{K},\mathbb{F}_n}$ with $\mathbb{K} := \{\wedge, \vee\}$ and $\mathbb{F}_\ltimes = \{x_1, x_2, ..., x_n, \overline{x_1}, \overline{x_2}, ..., \overline{x_n}\}$ and with size μ being the number of leaves. With that the probability $\mathbb{P}_{m,n}$ is

$$\mathbb{P}_{m,n} := \frac{|\{e \in C_n : \mu(e) = m, f_e = f\}|}{|\{e \in C_n : \mu(e) = m\}|}$$

and the limit distribution on n variables, if it exists, is

$$\mathbb{P}_{\infty,n} := \lim_{m\to\infty} \mathbb{P}_{m,n}.$$

At a later point in this section we investigate the results in the literature, considering the behavior of $\mathbb{P}_{\infty,n}(f)$ for $n \to \infty$ for a fixed function. $\triangle$

A second model that was intensively studied in the literature is the analogue of the Catalan And/Or model with the difference that internal nodes are labeled only by the implication and the leaves with positive literals; that means $\mathbb{K} = \{\to\}$ and $\mathbb{F} = \{x_1, ..., x_n\}$. This model was analyzed in

[FGGG12, GG10] with the extension of associativity in [GGKM12]. Also the Polish school around Zaionc investigated this model extensively with a different goal in mind, see [Zai03]. The survey of Gardy [Gar06] gives a very good overview on the class of models reviewed in this part. In the example above we let the size of the expressions grow first ($m \to \infty$) and then we let the number of variables grow ($n \to \infty$). To let the number of variables grow to infinity at first and then let the size m grow next, is in our cases not interesting because when the number of variables tends to infinity, the probability of a fixed function tends to 0 (see later). But to grow m and n simultaneously can lead to interesting results. So one can investigate the behavior of $\mathbb{P}_{m(k),n(k)}$ for functions $m(k), n(k)$ tending to ∞ when $k \to \infty$. This was done for the first time and until now the only time in [GM15]. The different limits are visualized in 4.1.

The above considerations were primarily done for the Catalan And/Or model and the model with implication. We now make some of the conclusions made there, but on a more general and somehow extended fashion. Let us now gather the requirements so that our general analysis works. Let $\mathbb{K} = \{k_1, k_2, ..., k_l\}$ and $\mathbb{F}$ arbitrary. Define $\mathbb{S} := \mathbb{E}_{\mathbb{K},\mathbb{F}}$. According to the definition of the expressions in Definition 2.5, this set is closed under the following construction:

- Take a connector k_i with arity($\mathrm{k_i}$) = j as root and append j expressions/trees $e_1, ..., e_j$ in an ordered way to it.

So the constructed tree $t := (k_i, e_1, ..., e_j)$ is also in $\mathbb{S}$. We now can define the property of the size function $\mu : \mathbb{S} \mapsto \mathbb{N}_0$ that we need.

Definition 4.2 (Admissible size function)
Regard the construction above. The size function μ is called admissible if:

- $\mu(t) = \mu((k_i, e_1, ..., e_j)) = w_i + \mu(e_1) + ... + \mu(e_j), w_i \in \mathbb{N}_0$.
- *When at least one w_i is 0, then $\mu(f) \neq 0, \forall f \in \mathbb{F}$.*
- *When $w_i = 0$, then* arity($\mathrm{k_i}$) $\neq 1$.
- *There is a $w_i > 0$ with* arity($\mathrm{k_i}$) ≥ 2.

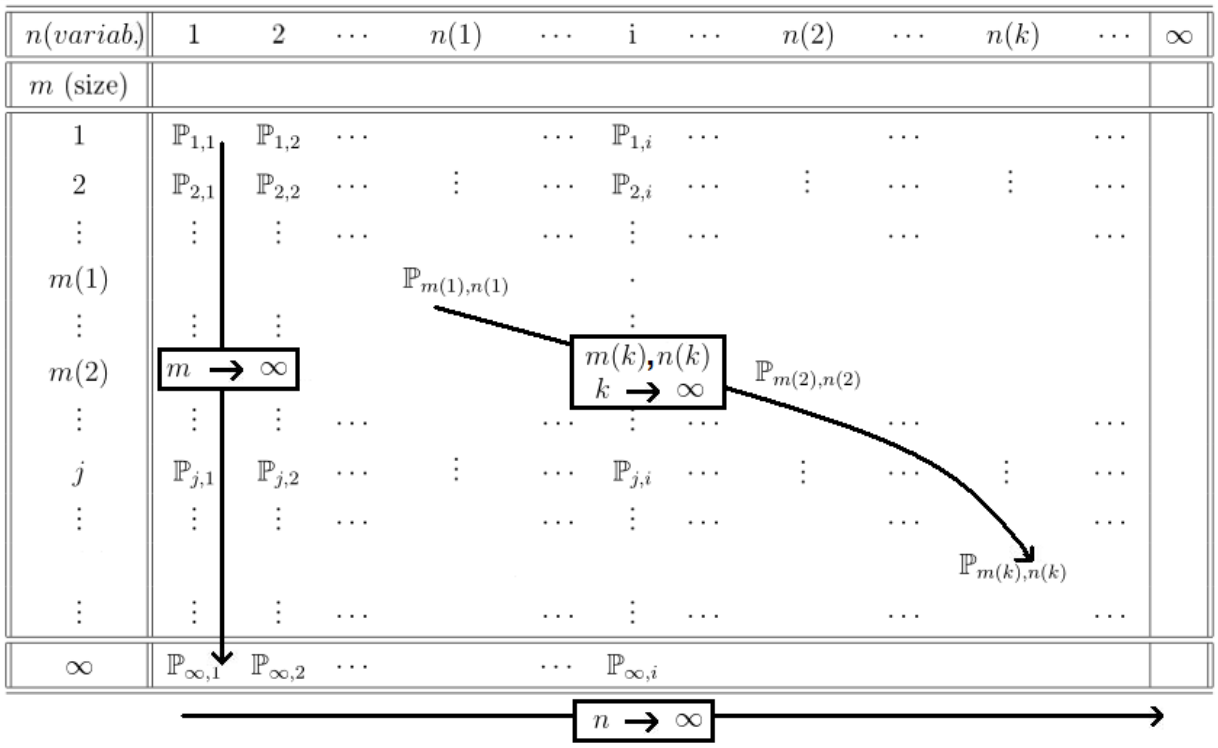

Figure 4.1: Different limits.

The first property is necessary because then the symbolic method with the Cartesian product construction in Fact 2.15 is applicable. The second and the third property ensure that $\mathbb{S}$ with $\mu(e_j)$ is a combinatorial class by forbidding the situation that there are base functions and connectors with size 0 or a connector with arity 1 and weight 0, which both would lead to an infinite number of expressions of fixed size. Finally, the fourth property says that there is at least one connector with arity greater equal 2 and positive weight. With such size functions the construction above leads to larger and larger expressions. The following observation characterizes the admissible size functions and also gives an example of a size function that is not admissible:

Observation 4.3 (Admissible size function)
The admissible size functions can be seen as weighting the base functions and the connectors and define the size of an expression as the sum of the weights of the base functions and connectors occurring as labels of the nodes. The weights of the connectors are denoted by w_i and the weight of the base functions by a_i. For example size functions that use the number of internal nodes, external nodes or the total number of nodes are admissible size functions: All base functions get the weight 0 and the connectors the weight $w_i = 1$ (respective 1 and $w_i = 0$ or 1 and $w_i = 1$). The depth of an expression is not admissible: The depth of a tree is 1 plus the maximum of the depths of the subtrees of the root.

In the entire section we will only consider models with admissible size functions and therefore we will omit the term admissible by chance. Regarding e.g. Fact 2.15, these models are perfectly suited for the symbolic method and also for singularity analysis as the following observation demonstrates.[1]

Proposition 4.4 (Square-root singularity)
Let $\mathbb{S} := \mathbb{E}_{\mathbb{K},\mathbb{F}}$ be a model with admissible size function μ and $\mathbb{K} = \{k_1, ..., k_l\}$. Let s_0 be the number of base functions of size 0 and $F(z)$ the GF of $\mathbb{F}$. Then, the GF $S(z)$ of this model has a square-root singularity

$$S(z) \underset{z \to z_0}{=} s_0 + y_0 - \gamma\sqrt{1 - \frac{z}{z_0}} + \mathcal{O}(1 - \frac{z}{z_0})$$

with γ, z_0, y_0, evaluated as in Fact 2.22 with

$$G(z, y) := F(z) + \sum_{i=1}^{l} z^{w_i}(y - s_0)^{\mathrm{arity(k_i)}} - s_0$$

that is valid in a Δ-domain. Additionally, if $S(z)$ is aperiodic, then we have the asymptotic expansion for the coefficients of $S(z)$ of the form

$$[z^m]S(z) \underset{m \to \infty}{=} \frac{\gamma}{2\sqrt{\pi m^3}} \left(\frac{1}{z_0}\right)^m (1 + \mathcal{O}(m^{-1})).$$

[1] This circumstance results from the fact that the underlying trees used for the expressions are of a special type, called simple variety of trees in the literature. Such trees are perfectly suited for singularity analysis and a lot of interesting properties of such varieties have been established see e.g. [FS09] and [MM78].

Proof. Let $\mathbb{K} = \{k_1, k_2, ..., k_l\}$ be the set of connectors. With this the description of $\mathbb{S}$ is

$$\mathbb{S} = \mathbb{F} \cup \bigcup_{i=1}^{l} (k_i \times \mathbb{S}^{\mathrm{arity(k_i)}}).$$

According to the symbolic method, which is applicable because of the first property of μ, this translates to the equation for the generating functions

$$S(z) = F(z) + \sum_{i=1}^{l} z^{w_i} S(z)^{\mathrm{arity(k_i)}}. \tag{4.4}$$

We consider two cases: In the first case assume that all base functions $f \in \mathbb{F}$ have weight > 0 and define $G(z, y)$ as

$$G(z, y) := F(z) + \sum_{i=1}^{l} z^{w_i} y^{\mathrm{arity(k_i)}}$$

so that in this case all requirements of 2.22 for the coefficients 2.1 are satisfied: $g_{0,0} = 0$ because there are no base functions of size 0 and by definition connectors have positive arity. That $g_{m,n} \geq 0$ is clear, $g_{0,1} \neq 1$ because of the third property of admissible size functions. Because of the fourth property of admissible size functions $\exists m \exists n \geq 2 : g_{m,n} > 0$. In the second case there is a base function with weight 0 but no connector with weight $w_i = 0$ due to the second property of admissible size functions. To apply Fact 2.22 we normalize the system $S(z) = G(z, S(z))$ to $S(z) - s_0 = G(z, S(z)) - s_0$ and substitute $\hat{S}(z) := S(z) - s_0$:

$$\hat{S}(z) = \hat{G}(z, \hat{S}(z)), \quad \hat{G}(z, y) := G(z, y + s_0) - s_0.$$

Again all conditions (2.1) in Fact 2.22 are satisfied because of the normalization and the fact that all w_i are greater 0. In both cases the function $S(z)$ (resp. $\hat{S}(z)$) is an analytic function around zero [2] with finite radius of convergence. The function $S(z)$ (resp. $\hat{S}(z)$) ceases to be analytic at z_0 obtained from the system 2.2 according to [FS09, Lemma VII.3]. So the solution of the system 2.2 is indeed in the domain of analyticity of $G(z, y)$ (resp. $\hat{G}(z, y)$). So all assertions follow directly from Fact 2.22. □

[2] The number of expressions in $\mathbb{S}$ with a given size can be upper bounded with a family of simple trees [FS09] labeled by finitely many labels which leads to an analytic function.

The above observation is applicable for the Catalan And/Or model, where in Example 2.21 we were also able to directly derive that the GF of the Catalan And/Or trees exhibit a square-root singularity. So it seems that we are able to count (asymptotically) the total number of expressions of a given size m in our models. But how can we count the number of expressions of size m calculating a given function so that we can quantify the probability of this function. We will not be able to do so explicitly but in a recursive way:

Example 4.5 (Generic equations)
Take again the Catalan And/Or model as in 4.1 for an illustrative example. Imagine that there are a_{m_1} expressions of size m_1 computing f_1 and a_{m_2} expressions of size m_2 computing f_2 and that $f_1 \wedge f_2 = f$. Then, there are $a_{m_1}a_{m_2}$ expressions of size $m_1 + m_2$ with root labeled by $\wedge$ and left subtree computing f_1 and right subtree computing f_2. Let f_i denote the i-th Boolean function in $\mathbb{B}_n$ in any order, as well as the class of Catalan And/Or expressions computing f_i with associated GF $f_i(z)$. Since the expressions are complete and the number of variables occurring in the base functions is n, all functions from $\mathbb{B}_n$ are occurring in the model, so $i \in \{1, ..., 2^{2^n}\}$. The recursive relation between the classes f_i leads to the generic equations for the generating functions: For $1 \leq i \leq 2^{2^n}$:

$$f_i(z) = z\mathbb{1}_{\{f_i \text{ is literal}\}} + \sum_{\substack{(f_j,f_l):\\ f_j \wedge f_l = f_i}} f_j(z)f_l(z) + \sum_{\substack{(f_j,f_l):\\ f_j \vee f_l = f_i}} f_j(z)f_l(z), \tag{4.5}$$

whereas $\mathbb{1}_{\{f_i \text{ is literal}\}}$ is the indicator function of f_i being a literal. $\triangle$

4.2 Existence and properties of the limit probability

These generic equations for the GFs of the expressions computing a function can be used directly to recursively calculate the number of expressions of a given size m computing a function f. So theoretically one is able to calculate the probability $\mathbb{P}_m(f)$ for every function f, which is just the number of expressions of size m computing f, divided by all expressions of size m. Practically, for large m and n, this is not possible and it would not reveal more than a clue if the limit probability $\mathbb{P}_\infty(f) := \lim_{m\to\infty} \mathbb{P}_m(f)$ exists for every

f.

The very important Drmota-Lalley-Woods Theorem (DLW Theorem) in connection with the transfers in Fact 2.19 give a dignified answer to the question of existence of the limit distribution. In a nutshell the DLW Theorem states that GFs that satisfy a system of functional equations as in equation 4.5 have a square-root singularity at a common point ρ. With that one shows the existence of the limit with a transfer. The DLW Theorem was independently found by Drmota [Drm97], Lalley [Lal93] and Woods [Woo97] approximately at the same time. The paper by Woods is of additional interest because he had a similar goal in mind. Woods investigated coloring rules for finite trees, which are rules on how the colors of the leaves propagate to the root, and the limit distribution of the colors of the roots. In our set-up one can see the function computed by a node as the color of this node and this color is determined according to the colors (functions) of the children-nodes and the coloring rules that are simply the composition 'rules' for Boolean functions with connectors. The version of the DLW Theorem due to [FS09, Theorem VII.6] is provide. To start with some definitions which are found together with the following explanations in [FS09]:

Definition 4.6 (Formal power series: valuation and distance)
The valuation $\mathrm{val}(\cdot)$ *of a power series* $f(z)$ *is the minimal* $m \in \mathbb{N}_0$ *so that* $[z^m]f(z) \neq 0$. *The valuation of a vector of power series* $\vec{y}(z) = (y_1(z), ..., y_k(z))$ *is the minimum over the individual valuations of* $y_i(z)$*:*

$$\mathrm{val}(\tilde{\mathrm{y}}) := \min\{\mathrm{val}(\mathrm{y_i})|1 \leq \mathrm{i} \leq \mathrm{k}\}.$$

The distance between two vectors of power series is

$$d(\vec{y}, \vec{w}) := 2^{-\mathrm{val}(\tilde{\mathrm{y}}-\tilde{\mathrm{w}})}.$$

With this distance measure the set of formal power series or vectors of them is a complete metric space and we call the convergence in this space formal convergence.

Definition 4.7 (Polynomial system, positive, proper, irreducible)

Let

$$\begin{aligned} y_1 &= \Phi_1(z, y_1, ..., y_k) \\ y_2 &= \Phi_2(z, y_1, ..., y_k) \\ &\vdots \\ y_k &= \Phi_k(z, y_1, ..., y_k) \end{aligned} \tag{4.6}$$

be a polynomial system of equations abbreviated with

$$\vec{y} = \vec{\Phi}(z, \vec{y})$$

that is nonlinear, i.e. at least one polynomial Φ_i is nonlinear in some of the indeterminates $y_1, ..., y_k$.

(i) A polynomial system is said to be positive if all components Φ_i have non-negative coefficients.

(ii) A polynomial system is said to be proper if

$$d(\vec{\Phi}(z, \vec{y}), \vec{\Phi}(z, \vec{w})) < K d(\vec{y}, \vec{w})$$

for some $K < 1$[3] and for $\vec{y}$, $\vec{w}$ with $[z^0]\vec{y} = [z^0]\vec{w} = [z^0 y^0]\vec{\Phi}(z, y)$ componentwise.

(iii) A polynomial system is said to be irreducible if the graph $G = (V, E)$ with $V := \{1, ..., k\}$ and $E := \{(i, j) | y_j \text{ occurs in } \Phi_i(\vec{y})\}$, its dependency graph, is strongly connected.

(iv) A polynomial system is said to be aperiodic if all solutions are aperiodic power series.

Since we are working with GFs, positivity is a natural condition. Properness says that the system is a contraction provided that the regarded functions coincide in the zeroth coefficient (the constant).[4] Algebraic irreducibility literally means that no subsystem of 4.6 can be solved before the whole system is solved. This property is crucial so that all component solutions adopt the same singularity:

[3] This is equivalent to the fact that the valuation of the difference of two vectors is larger after the mapping.

[4] In [FS09] it is assumed that the system satisfies a Lipschitz condition as in (ii) for all functions so that $\vec{\Phi}$ is a contraction on the complete metric space. Then, by iterating the

Theorem 4.8 (DLW Theorem, [FS09, Theorem VII.6])
Let $\vec{y} = \vec{\Phi}(z, \vec{y})$ be a nonlinear polynomial system as in 4.7 which is positive, proper and irreducible. Then, all component solutions y_i have the same radius of convergence $\rho < \infty$ and a square-root singularity:

$$y_i(z) = \alpha_i - \beta_i \sqrt{1 - \frac{z}{\rho}} + \mathcal{O}(1 - \frac{z}{\rho}),$$

with positive numbers α_i, β_i. If in addition one component y_i is aperiodic[5] *then the expansion holds in a Δ-domain and a transfer yields that the coefficients of y_i satisfy*

$$[z^m]y_i \sim \frac{\beta_i}{2\sqrt{\pi m^3}} \left(\frac{1}{\rho}\right)^m (1 + \mathcal{O}(m^{-1})).$$

Remark. *This is the polynomial version of the DLW Theorem. There also exists a general version of it, see [Drm97].*

Investigating the proof of this theorem shows how to compute the numbers ρ, α_i and β_i: One has to find the smallest positive ρ and $\alpha_1, ..., \alpha_k$ which solve the characteristic system

$$\begin{aligned} \alpha_1 &= \Phi_1(\rho, \alpha_1, ..., \alpha_k) \\ \alpha_2 &= \Phi_2(\rho, \alpha_1, ..., \alpha_k) \\ &\vdots \\ \alpha_k &= \Phi_k(\rho, \alpha_1, ..., \alpha_k) \\ 0 &= \det(\mathbf{I} - \vec{\Phi}_{\vec{y}}(\rho, \alpha_1, ..., \alpha_k)) \end{aligned} \tag{4.7}$$

with $\mathbf{I}$ being the identity matrix and $\vec{\Phi}_{\vec{y}}$ the Jacobi matrix with respect to $y_1, ..., y_k$ (compare Fact 2.22). For the numbers β_i Drmota [Drm97] proved that $(\beta_i)_{i=1,...,k}$ is an eigenvector to the eigenvalue 1 of the Jacobi matrix

function scheme $\vec{s}_i := \vec{\Phi}(z, \vec{s}_{i-1})$, one obtains a sequence of functions converging formal to the unique limit for an arbitrary start function. Our definition of properness is sufficient for the DLW Theorem as can be seen from its proof in [FS09]. In this case, the convergence of the scheme is guaranteed if we start with a function that coincides in the constant with the unique solution.

[5]For an irreducible system this is equivalent to the fact that all components are aperiodic.

$\vec{\Phi}_{\vec{y}}(z, \vec{y}(z))$ at the singularity ρ, i.e. at the point $(\rho, \alpha_1, ..., \alpha_k)$. Let us see how the DLW Theorem applies to the Catalan And/Or model.

Example 4.9 (Existence of the limit probability of the And/Or model) Consider the generic equations of the Catalan And/Or model in 4.5:

$$f_i(z) = z\mathbb{1}_{\{f_i \text{is literal}\}} + \sum_{\substack{(f_j,f_l):\\ f_j \wedge f_l = f_i}} f_j(z)f_i(z) + \sum_{\substack{(f_j,f_l):\\ f_j \vee f_l = f_i}} f_j(z)f_i(z). \tag{4.8}$$

They form a polynomial system in z and the components $y_i \triangleq f_i$. It is non-linear because we have connectors with arity greater one and positive weights and it is trivially positive. The properness follows because all base functions have weight > 0, see the proof of the next theorem. It is irreducible because TRUE is 'connected'[6] to every Boolean function g; and g is 'connected' to the function TRUE:

$$g = \top \wedge g, \quad \top = \top \vee g.$$

The GF for TRUE is aperiodic because there are expressions of every size greater equal 2 for TRUE:

$$\top = (x_1 \vee \overline{x_1}) = (x_1 \vee \overline{x_1}) \vee x_1 = \cdots .$$

All requirements of the DLW Theorem are satisfied so that all components have coefficients

$$[z^m]y_i \sim \frac{\beta_i}{2\sqrt{\pi m^3}} \left(\frac{1}{\rho}\right)^m (1 + \mathcal{O}(m^{-1})).$$

Let $S(z)$ be the GF of all expressions and f_i the Boolean function belonging to y_i. The total number of expressions of a given size m is the sum of these coefficients of index m

$$[z^m]S(z) = \sum_{1 \le j \le 2^{2^n}} [z^m]y_i(z)$$

so that the probability of a function f_i occurring among the expressions of size m is

$$\mathbb{P}_m(f_i) = \frac{[z^m]y_i(z)}{[z^m]S(z)} = \frac{[z^m]y_i(z)}{\sum_{1 \le j \le 2^{2^n}} [z^m]y_j(z)}.$$

[6]The GF for the function TRUE appears on the RHS of the generic equation for g.

Using the asymptotic expansions for the coefficients this leads to:

$$\mathbb{P}_\infty(f_i) = \lim_{m\to\infty} \mathbb{P}_m(f_i) = \frac{\beta_i}{\sum_{1\leq j\leq 2^{2^n}} \beta_j}.$$

This shows that the limit probability exists for all n and that all Boolean functions in $\mathbb{B}_n$ have positive probability. $\triangle$

The existence of the limit probability of the Catalan And/Or model was first shown in [LS95]. In [CFGG04] the DLW Theorem was first applied on this model as stated above. There one can also find numerical results for $n = 1, 2, 3$. We now want to follow the same approach as for the Catalan And/Or model but for a more general set-up. For this, observe that one can also be interested in models where the base functions and/or the connectors occur with different frequencies: Let $\mathbb{F} = \{b_1, ..., b_j\}$ be the base functions and $\mathbb{K} = \{k_1, ..., k_l\}$ the connectors. Furthermore, let p_{b_i} be the probability that the base function b_i occurs as a leaf of an expression and p_{k_i} is the probability that a connector k_i occurs as the label of an inner node. As previously stated we define the probability $\mathbb{P}_m(f)$ as the probability that an expression of size m computes f. This model is a random Boolean function model as defined in Definition 3.1. In contrast to the above model, where all expressions were equally likely (because we used the uniform distribution on expressions of size m), this model has a probability distribution on expressions that is induced by the probability of the connectors and base functions. For example in the Catalan And/Or model one could wish that x_1 occurs twice as often as the other literals and the connector $\wedge$ three times as often as $\vee$. This is easy to implement in the model by taking as base functions $\mathbb{F} = \{x_1, \tilde{x}_1, x_2, ...x_n, \overline{x_1}, \overline{x_2}, ..., \overline{x_n}\}$ and as connectors $\mathbb{K} = \{\wedge, \wedge_1, \wedge_2, \vee\}$. This changes the generic equations in the way that 3 occurs as a multiplicative factor in front of the sum responsible for the $\wedge$ and in the equation for the function x_1 the constant 2 instead of 1 occurs. In that way every rational ratio between the connectors resp. base functions can be realized. To implement arbitrary real valued probabilities in the model so that every inner node is labeled by connector k_i with probability p_{k_i} and every leaf with a base function b_i with given probability p_{b_i} one has to include this weighting in the generic equations:

Definition 4.10 (Generic equations for the general DLW model)
Let $\mathbb{F} = \{b_1, ..., b_j\}$ be the set of base functions with probabilities p_{b_i} and let $\mathbb{K} = \{k_1, ..., k_l\}$ be the connectors with probabilities p_{k_i}, with $\sum_{i=1}^{j} p_{b_i} = \sum_{i=1}^{l} p_{k_i} = 1$. Let μ be an admissible size function with weights of the connectors w_i and weights of the base functions a_i. Let n be the number of different variables occurring in all functions of $\mathbb{F}$ and s the number of different functions that are expressible in this model. W.l.o.g. assume that the used variables are $x_1, ..., x_n$ and let $(f_i)_{i=1,...,s}$ be an ordering of the expressible functions. Then, the system of the s equations, $i = 1, ..., s$

$$y_i(z) = \sum_{h=1}^{j} z^{a_h} p_{b_h} \mathbb{1}_{\{f_i = b_h\}} + \sum_{h=1}^{l} z^{w_h} p_{k_h} \sum_{\substack{(f_{i_1},...,f_{i_{\mathrm{arity}(k_h)}}): \\ k_h(f_{i_1},...,f_{i_{\mathrm{arity}(k_h)}}) = f_i}} y_{i_1} ... y_{i_{\mathrm{arity}(k_h)}}, \quad (4.9)$$

abbreviated with

$$\vec{y} = \vec{\Phi}(z, \vec{y}),$$

is called the system of generic equations.

Remark. *With this definition the probability that an expression of size m computes f_i is*

$$\mathbb{P}_m(f_i) = \frac{[z^m] y_i(z)}{[z^m] \sum_{j=1}^{s} y_j(z)}.$$

To the author's knowledge the only papers which consider such an introduction of probability (at least for a simpler model and only for the base functions) are by Yashunskii [e.g. [Yas05]], at which we will take a more closer look later.[7] Without any further conditions the equations in (4.9) satisfy the following requirements of the DLW Theorem.

Lemma 4.11 (Generic equations: nonlinear, positive, proper)
Let the model $\mathbb{S}$ be as in Definition (4.10) with admissible size function μ and positive probabilities $p_{b_i} > 0$ and $p_{k_i} > 0$. Then, the system of generic equations (4.9) is a nonlinear polynomial system that is positive and proper.

[7] In the next section we will investigate a different class of models for which such probability distributions on the base functions and connectors have been investigated in the literature.

Proof. The facts that it is positive and polynomial follow immediately. It is nonlinear because μ is an admissible size function. For properness one has to show that the valuation of $\vec{\Phi}(z,\vec{g}(z))-\vec{\Phi}(z,\vec{w}(z))$ is larger than the valuation of $\vec{g}(z)-\vec{w}(z)$ for vectors of power series $\vec{g}$ and $\vec{w}$ with $[z^0]\vec{g}=[z^0]\vec{w}=[z^0y^0]\vec{\Phi}(z,y)$. Equivalently one shows that the valuation of the difference of each component $\mathrm{val}(\Phi_\mathrm{i}(\mathrm{z},\tilde{\mathrm{g}})-\Phi_\mathrm{i}(\mathrm{z},\tilde{\mathrm{w}}))$ is increasing. When all connectors have sizes $w_i>0$ then 4.9 can be written as

$$y_i(z)=\sum_{h=1}^{j} z^{a_h}p_{b_h}\mathbb{1}_{\{f_i=b_h\}}+z\sum_{h=1}^{l} z^{w_h-1}p_{k_h}\sum_{\substack{(f_{i_1},\ldots,f_{i_{\mathrm{arity}(\mathrm{k_h})}}):\\ k_h(f_{i_1},\ldots,f_{i_{\mathrm{arity}(\mathrm{k_h})}})=f_i}} y_{i_1}\ldots y_{i_{\mathrm{arity}(\mathrm{k_h})}}, \tag{4.10}$$

with $w_h-1\geq 0$. So $\Phi_i(z,\vec{g}(z))-\Phi_i(z,\vec{w}(z))=z[P_i(z,\vec{g}(z))-P_i(z,\vec{w}(z))]$, with P_i a polynomial. The valuation of $\Phi_i(z,\vec{g}(z))-\Phi_i(z,\vec{w}(z))$ thus is indeed at least by one greater than the valuation of $\vec{g}(z)-\vec{w}(z)$. For the case that there is a weight $w_i=0$, all base functions have at least size 1 due to admissibility of μ. Then, $[z^0y^0]\vec{\Phi}(z,y)=0$ causes that both $\vec{g}(z)$ and $\vec{w}(z)$ start with z^1 in every component. Since all monomials of $\Phi_i(z,\vec{y})$ in $\vec{y}$ are at least quadratic except for the parts coming from unary connectors, which have in turn a multiplicative z in front of them because unary connectors have weight >0 for admissible size functions, it again holds that the valuation of the difference after the mapping is at least 1 larger than before the mapping: $\mathrm{val}(\mathrm{P_i}(\mathrm{z},\tilde{\mathrm{g}}(\mathrm{z}))-\mathrm{P_i}(\mathrm{z},\tilde{\mathrm{w}}(\mathrm{z})))>\mathrm{val}(\tilde{\mathrm{g}}(\mathrm{z})-\tilde{\mathrm{w}}(\mathrm{z}))$. □

Observation 4.12 (Generic equations: Irreducibility, Aperiodicity)
The irreducibility and aperiodicity of the system of generic equations (4.9) *for the generalized model with probabilities for the connectors and base functions does not depend on the probabilities:*
The system of generic equations is irreducible resp. aperiodic for fixed $p_{b_i}>0$ and fixed $p_{k_i}>0$ with $\sum_{i=1}^{j}p_{b_i}=\sum_{i=1}^{l}p_{k_i}=1$, if and only if this holds for all possible probability distributions on the connectors and base functions with $p_{b_i}>0$ and $p_{k_i}>0$.
This is easy to see and it is also clear that the system can fail to be irreducible if there is an i with $p_{b_i}=0$ or $p_{k_i}=0$; because then the whole structure of the equations can change, since the functions that can be computed might change, as well as the functions occurring on the RHS of a generic equation because

the connectors where it occurred have probabilities zero. So in this sense the irreducibility depends only on the used connectors, base functions and μ and so with aperiodicity.

We are now turning our attention to one of the main results in this section concerning the existence of the limit probability $\mathbb{P}_\infty$.

Theorem 4.13 (Limit probability, existence, positivity and continuity)
Let the model $\mathbb{S}$ be the set of Boolean expressions with connectors in $\mathbb{K} = \{k_1, ..., k_l\}$ and base functions in $\mathbb{F} = \{b_1, ..., b_j\}$. The base functions occur with probabilities $p_{b_i} > 0$ and the connectors with probabilities $p_{k_i} > 0$, with $\sum_{i=1}^{j} p_{b_i} = \sum_{i=1}^{l} p_{k_i} = 1$. Let μ be an admissible size function. Let s be the number of different functions that are expressible in this model. If the system of the generic equations forms an irreducible and aperiodic system, then the limit distribution $\mathbb{P}_\infty$ exists for all probability distributions $p_{b_i} > 0$, $p_{k_i} > 0$ and is positive on all expressible functions f_i, $i = 1, ..., s$. Moreover, the mapping

$$\begin{aligned} Q : D &\to \mathbb{R}^s \\ (p_{b_1}, ..., p_{b_j}, p_{k_1}, ..., p_{k_l}) &\mapsto (\mathbb{P}_\infty(f_1), ..., \mathbb{P}_\infty(f_s)) \end{aligned} \tag{4.11}$$

with $D := \{(p_{b_1}, ..., p_{b_j}, p_{k_1}, ..., p_{k_l}) | p_{b_i} > 0, p_{k_i} > 0, \sum_{i=1}^{j} p_{b_i} = \sum_{i=1}^{l} p_{k_i} = 1\}$ is a continuous[8] function.

Proof. This result is a consequence of the DLW Theorem and the observations made above. The conditions of this theorem are posed in such a way that (with Lemma 4.11 and Observation 4.12) the conditions of the DLW Theorem are satisfied for all probability distributions with positive probabilities. This gives:

$$\mathbb{P}_\infty(f_i) = \lim_{m\to\infty} \frac{[z^m] y_i}{[z^m] \sum_{j=1}^{s} y_j} = \lim_{m\to\infty} \frac{\frac{\beta_i}{2\sqrt{\pi m^3}} \left(\frac{1}{\rho}\right)^m (1 + \mathcal{O}(m^{-1}))}{\sum_{j=1}^{s} \frac{\beta_j}{2\sqrt{\pi m^3}} \left(\frac{1}{\rho}\right)^m (1 + \mathcal{O}(m^{-1}))} = \frac{\beta_i}{\sum_{j=1}^{s} \beta_j},$$

[8] As usual, the domain and range are equipped by the subspace topology induced by the Euclidean topology.

with $\beta_i > 0$, so the first assertion follows.

Now we prove the continuity of the function Q on the set

$$\tilde{D} := \{(p_{b_1}, ..., p_{b_j}, p_{k_1}, ..., p_{k_l}) | p_{b_i} > 0, p_{k_i} > 0\} \supset D.$$

Let $\vec{p}_n \underset{n\to\infty}{\to} \vec{p}_0 := (p_{b_1}, ..., p_{b_j}, p_{k_1}, ..., p_{k_l})$ and let $\vec{y} = (y_1, ..., y_s)$ be the vector of component-solutions for $\vec{p}_0$ and σ the radius of convergence of the components[9] and consider the generic equations

$$\begin{aligned} \alpha_1 &= \Phi_1(\rho, \alpha_1, ..., \alpha_s, \vec{p}) \\ \alpha_2 &= \Phi_2(\rho, \alpha_1, ..., \alpha_s, \vec{p}) \\ &\vdots \\ \alpha_s &= \Phi_s(\rho, \alpha_1, ..., \alpha_s, \vec{p}) \\ 0 &= \det(\mathbf{I} - \vec{\Phi}_{\vec{y}}(\rho, \alpha_1, ..., \alpha_s, \vec{p})). \end{aligned} \tag{4.12}$$

Let α_i^n $(i = 1, ..., s)$ denote the solutions of the system for parameters $\vec{p}_n$, ρ_n their radius of convergence and $\mathbb{P}_n$ the corresponding probability function. We show that the radius of convergence ρ_n as well as the values of the components at this point $\alpha_i^n(\rho_n)$ $(i = 1, ..., s)$ all converge to the corresponding values σ and $y_i(\sigma)$ $(i = 1, ..., s)$ for $n \to \infty$. This then leads to the convergence of the probabilities $\mathbb{P}_n(f_i)$ $(i = 1, ..., s)$ to $\mathbb{P}(f_i)$ $(i = 1, ..., s)$ for $n \to \infty$ because the vector $(\mathbb{P}_n(f_i))_{i=1,...,s}$ is an eigenvector of $\vec{\Phi}_{\vec{y}}(\rho_n, \alpha_1^n(\rho_n), ..., \alpha_s^n(\rho_n), \vec{p}_n)$ to the eigenvalue 1. The matrix is nonnegative so the multiplicity of the eigenvalue is 1 and the vector $(\mathbb{P}_n(f_i))_{i=1,...,s}$ is the unique positive normed eigenvector of $\vec{\Phi}_{\vec{y}}(\rho_n, \alpha_1^n(\rho_n), ..., \alpha_s^n(\rho_n), \vec{p}_n)$. Therefore, the eigenvector depends continuously on the coefficients of the matrix that themselves depend continuously on $\vec{p}_n$, ρ_n and $\alpha_i^n(\rho_n)$ $(i = 1, ..., s)$. Since these values all converge to the corresponding values for $\vec{p}_0$, the convergence of $(\mathbb{P}_n(f_i))_{i=1,...,s}$ to $(\mathbb{P}(f_i))_{i=1,...,s}$ is established.

So it remains to show that $\rho_n \to \sigma$ and $\alpha_i^n(\rho_n) \to y_i(\sigma)$ $(i = 1, ..., s)$ for $n \to \infty$. We first show that $\rho_n \to \sigma$ and that $\sum_{j=1}^s \alpha_j^n(\rho_n) \to \sum_{j=1}^s y_j(\sigma) =: Y(\sigma)$. For this we consider a simpler problem namely the single equation for the sum of the components instead of the system of equations for all

[9] Due to DLW theorem all components have the same radius of convergence

components. So we consider the equation $Y(z) = \Psi(z, Y(z), \vec{p})$ for the sum of the components with $\Psi := \sum_{j=1}^{s} \Phi_j$. For fixed parameter $\vec{p}$ this is a single functional equation for $Y(z)$ with a polynomial with nonnegative coefficients which depend continuously on $\vec{p}$ and z. For parameters $\vec{p}_n$ the radius of convergence of $Y_n(z)$[10] is ρ_n and is evaluated as the smallest positive value such that

$$\begin{aligned} Y_n(\rho_n) &= \Psi(\rho_n, Y_n(\rho_n), \vec{p}_n) \\ 1 &= \Psi_y(\rho_n, Y_n(\rho_n), \vec{p}_n). \end{aligned} \tag{4.13}$$

The finitely many coefficients of $\Psi(\rho, y, \vec{p})$ depend continuously on $\vec{p}$ and ρ; so if ρ_n converges to a value a, then there is an $N \in \mathbb{N}$ such that $\sup_{y\in[0,b]} \|\Psi(\rho_n, y, \vec{p}_n) - \Psi(a, y, \vec{p}_0)\| + \sup_{y\in[0,b]} \|\Psi_y(\rho_n, y, \vec{p}_n) - \Psi_y(\sigma, y, \vec{p}_0)\| < \epsilon$ for $n > N$ and $b > 0$.[11] Define the parameter p^* as $p_0 - \frac{1}{2}p_0$ and ρ^* as the radius of convergence of the solutions for the parameter p^*. All radii ρ_n are elements of $[0, \rho^*]$ for large enough n[12]. So the radii ρ_n have an accumulation point. Assume $a < \sigma$. Then, $\Psi_y(a, Y(a), \vec{p}_0) < 1$[13] and so there is an $N_1 \in \mathbb{N}$ with $\Psi_y(\rho_n, Y_n(\rho_n), \vec{p}_n) < 1$ for $n = N_1$ because there exists a subsequence of $Y_n(\rho_n)$ that converges to $Y(a)$, see Figure 4.2 for an illustration. So there is an n such that $(\rho_n, Y_n(\rho_n))$ does not solve the characteristic equations, a contradiction. The case where $a > \sigma$ works analogously. So ρ_n converges necessarily to σ and therefore also $Y_n(\rho_n)$ converges to $Y(\sigma)$.

Now in an intermediate step we consider a sequence $\vec{p}_n$ that converges from 'below' to $\vec{p}_0$, meaning that for $n_1 < n_2$ $\vec{p}_{n_1} \leq \vec{p}_{n_2} \leq \vec{p}_0$ componentwise. Let $\alpha_i^n(z) = \sum_{m=0}^{\infty} a_{i\,m}^n z^m$, $y_i(z) = \sum_{m=0}^{\infty} a_{im} z^m$. The fixed-point iteration $\vec{w}_{u+1}^n = \vec{\Phi}(\vec{w}_u^n), w_0 := ([z^0\vec{x}^0]\Phi_i(z, \vec{x}))_{i=1,\dots,s}$[14] gives a sequence of components converging formally to $\alpha_i^n(z)$ $(i = 1, ..., s)$ and shows together with the positivity of the system that for $n_1 < n_2$ all coefficients of $\alpha_i^{n_2}(z)$ are larger (or equal) than (as) the coefficients of $\alpha_i^{n_1}(z)$ and both are dominated by the coefficients of $y_i(z)$: $[z^m]y_i(z) \geq [z^m]\alpha_i^{n_2}(z) \geq [z^m]\alpha_i^{n_1}(z)$. So the sequence of

[10] $Y_n(z) := \sum_{i=1}^{s} \alpha_i^n(z)$

[11] This means that the sequence of the functions as well as their derivatives converge uniformly on the compact set $[0, b]$.

[12] See below.

[13] This holds because $\Psi_y(a, Y(a), \vec{p}_0)$ as a function of a is (strict) increasing and $a < \sigma$ and $\Psi_y(\sigma, Y(\sigma), \vec{p}_0) = 1$

[14] Compare with properness in 4.7

the radii is a nonincreasing sequence $\rho_{n_1} \geq \rho_{n_2} \geq \sigma$. Moreover, since $y_i(z)$ is an analytic function which is finite at its radius of convergence σ[15] and α_i^n is dominated by $y_i(z)$ on $[0,\sigma]$ there is an $M \in \mathbb{N}$ such that for all n the sum of terms of $\alpha_i^n(z)$ resp. $y_i(z)$ of degree M or greater is smaller than $\frac{\epsilon}{3}$ for $z \in [0,\sigma]$: $\sup_{z\in[0,\sigma]} \| \sum_{m=M}^{\infty} a^n_{i\,m} z^m \| < \frac{\epsilon}{3}$ resp. $\sup_{z\in[0,\sigma]} \| \sum_{m=M}^{\infty} a_{im} z^m \| < \frac{\epsilon}{3}$. So $\sup_{z\in[0,\sigma]} \|y_i(z) - \alpha_i^n(z)\| \leq \sup_{z\in[0,\sigma]} \| \sum_{m=M}^{\infty} a^n_{i\,m} z^m \| + \sup_{z\in[0,\sigma]} \| \sum_{m=M}^{\infty} a_{im} z^m \| + \sup_{z\in[0,\sigma]} \| \sum_{m=0}^{M-1} (a_{im} - a^n_{i\,m}) z^m \|$. The coefficients $a^n_{i\,m} (m = 1, ..., M-1)$ depend continuously on the parameters $\vec{p}_n$[16] and so there is an N such that the (finitely many) differences $a_{im} - a^n_{i\,m}$ $(m = 1, ..., M-1)$ are sufficient small such that $\sup_{z\in[0,\sigma]} \| \sum_{m=0}^{M-1} (a_{im} - a^{n_2}_{i\,m}) z^m \| < \frac{\epsilon}{3}$. Hence $\sup_{z\in[0,\sigma]} \|y_i(z) - \alpha_i^n(z)\| < \epsilon$ for all $n \geq N (i = 1, ..., s)$.

Let now $\vec{p}_n$ be a general sequence of parameters converging to $\vec{p}_0$ and let us again denote the component solutions for parameters $\vec{p}_n$ by α_i^n, the sum of these components by Y_n and the radius of convergence by ρ_n. We already know that ρ_n converges to σ and also that $Y_n(\rho_n)$ converges to $Y(\sigma)$. Moreover, there is a sequence $\vec{l_n}$ with the property that $\vec{p}_n \geq \vec{l_n}$ for all n, with the property that $\vec{l_n}$ converges from below to $\vec{p}_0$. Denote the component solutions for $\vec{l_n}$ by l_i^n. Since $\alpha_i^n(z) \geq l_i^n(z)$ on the interval $[0, \rho_n]$ it is not possible that the sequence $(\alpha_i^n(\rho_n))_{n\in\mathbb{N}}$ has an accumulation point that is smaller than $y_i(\sigma)$ because this would lead to the existence of an n_1 with $l_i^{n_1}(\rho_{n_1}) > \alpha_i^{n_1}(\rho_{n_1})$ due to the uniform convergence of $l_i^{n_1}$ on $[0,\sigma]$ as is illustrated in 4.3. So the smallest accumulation point of $\alpha_i^n(\rho_n)$ is $\geq y_i(\sigma)$ for $i = 1, ..., s$. From above we know that $\sum_{j=1}^{s} \alpha_j^n(\rho_n) = Y_n(\rho_n) \to Y(\sigma) = \sum_{j=1}^{s} y_j(\sigma)$ and therefore the largest accumulation point of $\alpha_i^n(\rho_n)$ cannot be larger than $y_i(\sigma)$. So $\alpha_i^n(\rho_n) \to y_i(\sigma)$ which eventually finishes the proof.

□

The proof of the continuity is very technical because we used a reduction to a one dimensional problem for the sum of the components for that we can easily show at hand the continuity of the radius of convergence and the

[15] This holds because it has a square-root singularity.

[16] Again with the fixed-point iteration and formal convergence one sees that the coefficients are polynomials in $\vec{p}_n$.

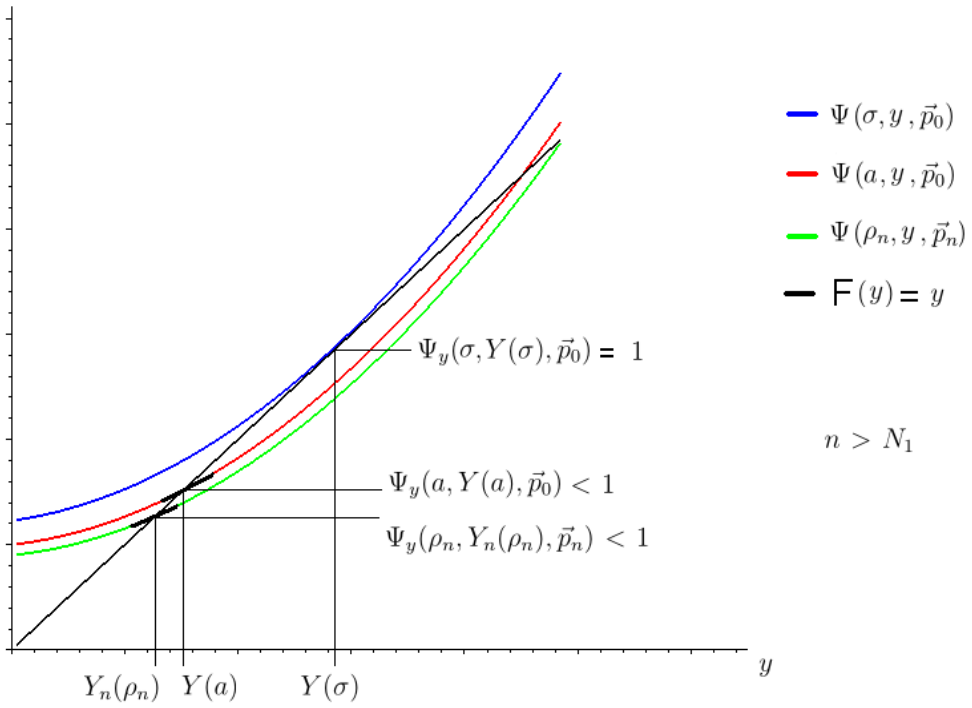

Figure 4.2

values of the component solutions at this point. Using results which consider the dependence of the solution(s) of positive polynomial fixed-point systems to the input parameters, the proof will presumably be much shorter and one probably can prove that the map D is infinitely often differentiable. A paper that might be interesting in this context is [EGK10] and [EKL10] where the least fixed points of positive polynomial systems are considered.

Nevertheless, the theorem above shows that if the size function is admissible and the irreducibility and aperiodicity conditions are satisfied, then all limit distributions exist, are positive and depend continuously on the probabilities of the base functions and connectors in the domain of positive probabilities D, where the irreducibility and aperiodicity are not changing. The question

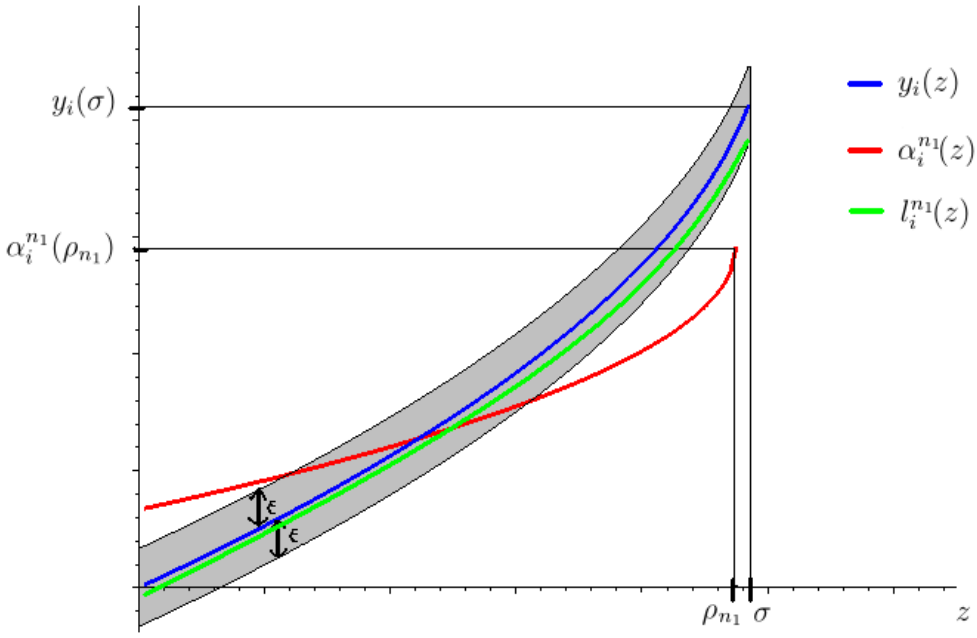

Figure 4.3

now is: What are sufficient conditions for irreducibility and aperiodicity? At first glance a model with connectors in $\mathbb{K}$ and base functions in $\mathbb{F}$ is irreducible and aperiodic iff the dual model built from the sets $\tilde{\mathbb{K}}$ and $\tilde{\mathbb{F}}$, the dual[17] connectors and base functions, is irreducible and aperiodic. This follows directly from the fact that if $f = c(g, h)$, then $\hat{f} = \hat{c}(\hat{g}, \hat{h})$. Moreover, the probability of a function in the initial model is the same as the probability of the dual function in the dual model. For a complete model the properties aperiodic and irreducible are monotone: Adding more connectors or/and base functions can not change these properties for a model in which all functions are expressible. An easy sufficient result that is stated because it shows the way how one could try to prove the irreducibility and aperiodicity for a given model is

[17] The dual function $\hat{f}$ to function f is defined pointwise as $\hat{f}(x_1, ..., x_n) := \overline{f(\overline{x_1}, ..., \overline{x_n})}$.

Lemma 4.14 (Regarding irreducibility, aperiodicity)
Let $\mathbb{K}$ and $\mathbb{F}$ be sets of connectors resp. base functions and n the number of variables of the functions in $\mathbb{F}$. Let f_1 be a Boolean function so that there exists a path from f_1 to every function and from every function to f_1 in the dependency graph of the generic equations, then the model is irreducible. Let $f_2 \in \mathbb{B}_n$ be a Boolean function for that there exist expressions e_1, e_2 in $\mathbb{E}_{\mathbb{K},\mathbb{F}}$ with $f_{e_i} = f_2$. Moreover, assume that the expression e_i ($i = 1, 2$) has (at least) k_i leaves labeled by $b_i \in \mathbb{F}$ with the property that these leaves are inessential for the expression[18] with $\gcd(\mathrm{k}_1, \mathrm{k}_2) = 1$ and that there is a connector c with weight 1 (resp. 0 and arity 2) and a base function b with weight 0 (resp. 1), then the model is in addition aperiodic.

Proof. Irreducibility: Every function f is connected to every function via f_1. Aperiodicity: There exist expressions of every size in the model: $c(b, ..., b)$ has size 1, $c(c(b, ..., b), ..., c(b, ..., b))$ has size 2 and so on. Substituting the k_i leaves labeled by b_i by an expression of size l gives an expression of size $s_i + lk_i$[19] computing f_2. So for all $l \in \mathbb{N}_0$ there exist expressions of size $s_i + lk_i$ expressing f_2. Now one shows that there are two expressions for f_2 that differ in size by one and therefore the system is aperiodic: Since $\gcd(\mathrm{k}_1, \mathrm{k}_2) = 1$ there exist numbers l_i with $l_1k_1 - l_2k_2 = 1$, so $((s_2 - s_1 + 1)l_1)k_1 - (s_2 - s_1 + 1)l_2k_2 = (s_2 - s_1 + 1)$ leading to $s_1 - s_2 + l_1^*k_1 - l_2^*k_2 = 1$. □

The above conditions are very restrictive and there are irreducible systems of a far different and more complicated connectivity. Nevertheless, the structure of the two most investigated models, the Catalan And/Or model and the model with implication, with size being the number of leaves (or internal nodes), is exactly of the nature above with $f_1 = f_2 = \mathrm{TRUE}$ and $k_1 = 1$: For the And/Or model see Example 4.9 and for the implication observe

$$f = \top \to f, \quad \top = f \to \top, \quad \top = x_1 \to \top = ((x_1 \to x_1) \to \top) = ... \quad .$$

Moreover, the proof of the Corollary below shows that also all models with connectors being a complete set of functions are of this structure.

[18] This means that the k_i leaves can be exchanged by any other function (for all leaves the same function) so that the expression still computes f_2.

[19] s_i is the size of the expression e_i where all leaves labeled by k_i are substituted by $c(b, ..., b)$

Corollary 4.15 (Complete Connectors)
Let $\mathbb{K}$ be a complete set of connectors and $\mathbb{F} = \{x_1, ..., x_n\}$. Let the size function μ be the number of connectors, then the generic equations form an irreducible and aperiodic system, hence Theorem 4.13 is applicable.

Proof. Since $\mathbb{K}$ is a basis there exists a tree $e(x_1, x_2)$ representing the function $x_1 \to x_2$. So every function f is connected to $\top$ and vice versa: $f = e(\top, f), \quad \top = e(f, \top)$. Moreover, every function is expressible. So the irreducibility follows from Lemma 4.14. For the aperiodicity observe that the only functions $g(x_1, ..., x_n)$, so that every partially evaluated function $g_{[x_{i_1} \mapsto v_1, ..., x_{i_{n-1}} \mapsto v_{n-1}]}\ (v_j)_{j=1,...,n-1} \in \{0,1\}^{n-1}$ is not $\top$ or $\bot$, are the functions $\oplus_n$ and $\neg(\oplus_n)$.[20] The set of all these functions $\{\oplus_i, \neg(\oplus_i) | 0 \leq i \leq n\}$ with $\oplus_0 := \bot$ and $\neg(\oplus_0) := \top$ is not a complete set of connectors: Indeed every expression built from these connectors is again such a function but there are Boolean functions that are not of this form. So a complete set $\mathbb{K}$ has to contain at least one connector that is TRUE resp. FALSE for a partial evaluation: $g_{[x_{i_1} \mapsto v_1, ..., x_{i_{n-1}} \mapsto v_{n-1}]} \in \{\top, \bot\}$ for a $(v_j)_{j=1,...,n-1} \in \{0,1\}^{n-1}$. This single connector as root, with all but one input substituted by expressions computing $\top, \bot$ according to the partial evaluation, is the required expression with 1 remaining leaf. So all requirements of 4.14 are fulfilled and the assertion follows. □

We will not further investigate the at first glance very complex seeming question of precise conditions for irreducibility and aperiodicity. It should also be mentioned that the limit probability can also exist without irreducibility; but then the powerful DLW Theorem is not applicable. In the next section the most important results regarding special models of the DLW-type and extensions of it are presented and summarized.

4.3 Instances of the DLW model

In this section the different instances of the DLW model that have been studied in the literature are introduced and the results are presented. These

[20] These are the functions that are true iff an odd (resp. even) number of variables is 1. It can be seen by induction (or by playing a kind of Sudoku in the truth table) that these two functions are the only ones who have to be evaluated until the last bit of every input in order to compute the whole function.

instances can be divided up in three classes according to the used connectors: The first class includes the Catalan And/Or expressions and extensions of them. The second class covers the Catalan trees with the implication as single connector and extensions of them. The third class involves arbitrary connectors but restricts the base functions to $\{\top, \bot\}$. This section is arranged according to this classification, whereas we will see that the first two classes have very similar behaviors.

4.3.1 $(\wedge, \vee)$-Expressions

All 5 models introduced here do not incorporate the extension of probabilities of the connectors or base functions. Hence all base functions and connectors have the same probability to occur. So we are looking at models $(\mathbb{S}_i, U_{\mathbb{S}_i}, \mu_i)$, $i = 1, ..., 5$. The $\mathbb{S}_i$'s are distinguished sets of Boolean expressions, whereas all of them have in common that the leaves are labeled by functions from $\mathbb{F} = \{x_1, x_2, ..., x_n, \overline{x_1}, \overline{x_2}, ..., \overline{x_n}\}$, $U_{\mathbb{S}_i}$ is the uniform distribution on subsets of $\mathbb{S}_i$ of fixed size. For the first 4 models the size function μ_i, $i = 1...4$, is the number of the leaves and μ_5 is the number of all nodes (leaves plus internal nodes) which is also called tree size complexity. The $\mathbb{S}_i$'s are:

- $\mathbb{S}_1$ is the set of all binary trees with inner nodes labeled by $\wedge$, $\vee$.
- $\mathbb{S}_2$ is the set of all binary non-plane[21] trees with inner nodes labeled by $\wedge$, $\vee$.
- $\mathbb{S}_3$ is the set of all trees with nodes labeled by connectors $\wedge$, $\vee$ of arbitrary arity ≥ 2 according to the in-degree of the labeled node such that an inner node and its children do not have the same labels.
- $\mathbb{S}_4$ is the set of all non-plane trees with nodes labeled by connectors $\wedge$, $\vee$ of arbitrary arity ≥ 2 according to the in-degree of the labeled node such that an inner node and its children do not have the same labels.
- $\mathbb{S}_5$ is $\mathbb{S}_3$.

$(\mathbb{S}_1, U_{\mathbb{S}_1}, \mu_1)$ is just the Catalan And/Or model also called binary plane model. The model $(\mathbb{S}_2, U_{\mathbb{S}_2}, \mu_2)$ incorporates the commutativity of the connectors:

[21] Non-plane trees are trees where the ordering of the children nodes is irrelevant. So they are trees as defined in A.3 without the property (plane).

$x_1 \wedge x_2 = x_2 \wedge x_1$, $x_1 \vee x_2 = x_2 \vee x_1$ and hence is called commutative-And/Or model.[22] The model $(\mathbb{S}_3, U_{\mathbb{S}_3}, \mu_3)$ incorporates the associativity of the connectors: $(x_1 \wedge x_2) \wedge x_3 = x_1 \wedge (x_2 \wedge x_3)$ and same for $\vee$; and is called associative-And/Or model. The condition "so that an inner node and its children do not have the same labels" means that internal nodes are labeled by $\wedge$ and $\vee$ in a stratified way. The model $(\mathbb{S}_4, U_{\mathbb{S}_4}, \mu_4)$ includes both commutativity and associativity of the connectors and is called commutative-associative-And/Or model. Finally, $(\mathbb{S}_5, U_{\mathbb{S}_5}, \mu_5)$ is the associative-And/Or model with the tree size complexity as size measure. For the associative-And/Or model this notion of size is indeed different from counting only leaves as the results show. For not associative models (i.e. constant arity of the connectors) it does not matter which size function one uses as explained earlier in this thesis. We will abbreviate these models in the following by $\mathbb{S}_i$, $i = 1...5$

For all 5 models the limit distributions exist for arbitrary n. For the first model we showed this by means of the DLW Theorem in the section above. For the models with associativity the generic equations are no longer polynomial in $\vec{y}$ but more general analytic functions. For irreducible and aperiodic systems of such a form there exists a general version of the DLW Theorem see [Drm97]. In the case of commutativity also terms in $y_i(z^2)$, $y_i(z^3)$, ... occur and one needs another extension which is found in [Kra11, Lemma 1.13]. In any case these extensions ensure the existence of the limit distribution and we will indicate them unified by $\mathbb{P}_n$ because it will be clear from the context which model we are actually referring to.

The binary plane model was first investigated in [LS95] and then with large improvements on results in [CFGG04, GW05, Koz08, GGM14, GM15]. The commutative and associative models and as well the binary plane model are investigated in [GGKM15]. [GGKM14] treats the model with tree size complexity. Finally, the early survey of Gardy [Gar06] gives an introductory overview on all these models and some more.

Let us first see what results have been established in the last 20 years in

[22] Observe that formally this model is not a random Boolean function model as defined in 3.1 because it uses non-plane trees. But there is a somehow unnatural random Boolean function model (of the strict definition) that coincides with $(\mathbb{S}_2, U_{\mathbb{S}_2}, \mu_2)$: Consider all binary trees (plane) and define the probability measure $\mathbb{P}_{\mathbb{S}_2}$ in such a way that the models coincide.

the case of our role model, the plane binary Catalan And/Or model $\mathbb{S}_1$. In [Woo97, Problem 6.6.] Woods stated the question whether and how the probability of a Boolean function occurring in the limit distribution is related to the formula size complexity. This question found a first answer in the result obtained by Lefmann and Savický in [LS95]. The result was improved in the following in [CFGG04] (upper bound) and in [GW05] (lower bound) to the final result:

Theorem 4.16
Consider the model $\mathbb{S}_1$. There exists a constant $c > 0$ such that for every positive integer n the following is valid: Let f be a Boolean function on n variables, then

$$(\frac{1}{8n})^{\mathcal{L}(f)} \leq \mathbb{P}_n(f) \leq (1 + \mathcal{O}(\frac{1}{n}))e^{-c\frac{\mathcal{L}(f)}{n^2}}. \tag{4.14}$$

The lower bound is established by counting the number of elements of a subset of expressions that compute a function f. This subset is sufficiently large to get a good lower bound and is easy to recursively describe by means of the symbolic method and they are able to derive an expansion around the single dominant singularity. For the upper bound the Markov inequality is used:

$$\begin{aligned}\mathbb{P}_n(f) = \mathbb{P}_n(e \text{ computes } f) \leq \mathbb{P}_n((1+\epsilon)^{\mu(e)} \geq (1+\epsilon)^{\mathcal{L}(f)}) \\ \frac{\mathbb{E}[(1+\epsilon)^{\mu(e)}]}{(1+\epsilon)^{\mathcal{L}(f)}},\end{aligned} \tag{4.15}$$

where the first inequality arises because the probability is taken over all expressions of size being larger than the complexity of f and the second is the Markov inequality. The largest ϵ for which $\mathbb{E}[(1+\epsilon)^{\mu(e)}]$ exists and is bounded would lead to the best upper bound that can be derived by this technique. The authors of [CFGG04] were able to derive $\epsilon = \frac{C}{n^2}$ leading to the above upper bound. The approach is as follows. The limit distribution is identified in a different way by means of infinite expressions. To an infinite expression pruning rules are applied to obtain a small expression computing the same function. The distribution of the infinite expressions is described with the help of a growing process which leads together with the pruning rules to

recursive relations of quantities related to $\mathbb{E}[(1+\epsilon)^{\mu(e)}]$. These recursions have to be analyzed to eventually obtain $\epsilon = \frac{C}{n^2}$ with $\mathbb{E}[(1+\epsilon)^{\mu(e)}] = (1+\mathcal{O}(\frac{1}{n}))$. After these results Kozik [Koz08] proved a very strong relation between the limit probability of a given function and its complexity for $\mathbb{S}_1$. This result was further refined and extended to the commutative and/or associative models in [GGKM15]. In the following let us define the complexity of the two constant functions $\top, \bot$ ad hoc as 0 so that the following theorem can be stated in an uniform way. The main results in [Koz08] and [GGKM15] are: Fix a model $\mathbb{S}_i$, $i = 1...4$.

Theorem 4.17 ([Koz08])
Let f be a Boolean function. Then, there is a constant λ_f[23] depending only on f so that

$$\mathbb{P}_n(f) \underset{n\to\infty}{\sim} \frac{\lambda_f}{n^{\mathcal{L}(f)+1}}. \tag{4.16}$$

In [Koz08] it is proven that $\mathbb{P}_n(f) \underset{n\to\infty}{=} \Theta(\frac{1}{n^{\mathcal{L}(f)+1}})$ in model $\mathbb{S}_1$ and the existence of the constant is suggested, which is verified in [GGKM15]. In [GGKM15] the exact values of the constants for the functions $\top$ and $\bot$ and for literals are computed and they give bounds for the constants for general functions. These bounds depend only on the number of trees of smallest size computing f (in the following called minimal trees) and on the complexity of f. For the model $\mathbb{S}_5$ a similar result is established in [GGKM14] with the difference that the function TRUE resp. FALSE has a probability bounded from below. In 4.4 is listed an overview of these results with the numerical values of the constants (if available) and associated asymptotic orders of the probabilities $\mathbb{P}_n(f)$. The column labeled 'Shannon effect' indicates whether the Shannon effect is disproved or not for the different models.

These results definitely answer the question 'What is the asymptotic behavior of the probability of a fixed function?' posed in section 3.1 for all 5 models. In the following section we will present some of the approaches that are used to establish these results.

In [Koz08] Kozik developed an approach, called 'pattern theory', which allows

[23] This constant is different for all 4 models

Model	$\mathbb{P}_n(f), n \to \infty$	$\mathbb{P}_n(x)$ $, n \to \infty$	$\mathbb{P}_n(\top)$ $, n \to \infty$	Shannon effect
binary plane: S_1	$\frac{\lambda_1}{n^{\mathcal{L}(f)+1}}$	$\frac{5}{16n^2}$	$\frac{3}{4n}$	disproved
associative: S_2	$\frac{\lambda_2}{n^{\mathcal{L}(f)+1}}$	$\frac{546-386\sqrt{2}}{n^2}$	$\frac{51-36\sqrt{2}}{n}$	
commutative: S_3	$\frac{\lambda_3}{n^{\mathcal{L}(f)+1}}$	$\frac{641}{2048n^2}$	$\frac{385}{512n}$	
ass.+com.: S_4	$\frac{\lambda_4}{n^{\mathcal{L}(f)+1}}$	$\frac{(2\ln(2)-1)^2(2\ln(2)+1)}{4n^2}$	$\frac{(2\ln(2)-1)^2}{4n}$	
ass.+tree-size: S_5	$\Theta(\frac{1}{n^{\mathcal{L}(f)+1}})$	$\Theta(\frac{1}{n^2})$	$0 < \alpha \leq \mathbb{P}(\top) \leq \beta < \frac{1}{2}$	disproved

Figure 4.4: Overview of the asymptotic behavior of the probabilities of fixed functions for the constant function, literals and general functions.

in the cases of models 1 to 4 to asymptotically count expressions which suffice structural constraints. He used this theory for the plane binary model before it was further extended and applied to the commutative and associative extensions of the Catalan And/Or model in [GGKM15]. The approaches for the models 1 to 4 have large similarities and are developed parallel in [GGKM15] therefore we will only demonstrate the approach for the simplest case, the binary plane model $\mathbb{S}_1$.

Let us first clarify the notation for this part and define pattern languages. The following definitions are found in [Koz08]: Let $\mathcal{F}_n$ be the set of Catalan And/Or trees with leaves $\{x_1, \overline{x}_1 ..., x_n, \overline{x}_n\}$ and $F_n(z)$ its GF with singularity ρ_n $(= \frac{1}{16n})$. We call the elements of $\mathcal{F}_n$ in the following expressions. A tree structure is a binary plane tree with inner nodes labeled by $\wedge$, $\vee$ and leaves 'labeled' by $\bullet$. The set of all tree structures is denoted by $\mathcal{T}$ with GF $\mathcal{T}(z)$. The tree structure of an expression e is defined as the tree which is

obtained by substituting every leaf of e by $\bullet$. A pattern language P is a set of binary plane trees with internal nodes labeled by $\wedge$, $\vee$ and leaves labeled by $\{\bullet, \square\}$ and the elements are called patterns. Leaves labeled by $\square$ are called placeholders and leaves labeled by $\bullet$ are called pattern leaves. Let $T \subset \mathcal{T}$ be a set of tree structures and define $P[T]$ as the set of all trees that can be obtained by substituting every placeholder of an element of P by elements of T. A pattern language P is said to be unambiguous if for every set of tree structures T' there is only one way to construct an element (structure) of $P[T']$ as above. For unambiguous pattern language P and tree structure t (or for expression e with tree structure t) in $P[T]$ one can distinguish the leaves of t that correspond to the pattern leaves of the (unique) pattern used to construct t, we call these leaves P-pattern leaves or just pattern leaves if it is clear from the context which pattern language we mean. We also need the composition of pattern languages: If P and S are pattern languages and $P[S]$ defined as above, then $P[S]$ is a pattern language with pattern leaves coming from both patterns. Let $P(x, y)$ be the GF of the pattern language P with x counting the pattern leaves and y the placeholders and $S(z)$ the GF of a set of tree structures. Unambiguity guarantees that the GF of $P[S]$ is $P(z, S(z))$. To make pattern theory amenable to singularity analysis we need the concept of subcriticality:

Definition 4.18 (Subcriticality, [Koz08, Definition 2.2])
We say that function $P(x, y)$ is subcritical for $T(z)$ if $T(z)$ is a GF having unique dominant singularity of the square-root type in $\rho \in \mathbb{R}_+$ and $P(x, y)$ is analytic in some set $\{(x, y) : |x| \leq \rho + \epsilon, |y| \leq t(\rho) + \epsilon\}$ for some $\epsilon \in \mathbb{R}_+$. We say that an unambiguous pattern language P is subcritical for a set of tree structures $T \subset \mathcal{T}$ if the GF $P(x, y)$ of P is subcritical for $T(z)$.

The pattern languages that we will use are:

$$N = N \vee N | \square \wedge N | \bullet, \qquad P = P \vee P | \square \wedge P | \bullet . \tag{4.17}$$

It is easy to verify that both are unambiguous and subcritical for the set of all structures $\mathcal{T}(z)$. The last definition to gather, before we can demonstrate the approach of pattern theory, is the definition of repetitions and restrictions.

Definition 4.19 (Repetitions, Restrictions)

Let e be an expression, P an unambiguous pattern language and $E \subset \{x_1, ..., x_n\}$. The number of P-repetitions of e is the number of P-pattern leaves of e minus the number of different variables (not literals !) occurring among the P-pattern leaves of e. The number of (P, E)-restrictions of e is the number of its P-repetitions plus the number of different variables in E that occur among the P-pattern leaves of e. A leaf that counts to the restrictions resp. repetitions will be called a restriction resp. a repetition.

If E is the set of the essential variables of e, then the leaves of e that are not (P, E)-restrictions, are exactly the P-pattern leaves that can be evaluated to $TRUE$ resp. $FALSE$ independently of the other P-pattern leaves without changing the function computed by e. The core of pattern theory is the fact that the number of restrictions seems to simultaneously be able to 'classify' the set of expressions according to structural aspects as well as aspects concerning the density of them. We will see that in the following.

Lemma 4.20 ([Koz08, Lemma 2.7])
Let T be a set of tree structures whose generating function $T(z)$ has unique dominant singularity in $\rho \in \mathbb{R}^+$ of the square-root type. Let P be an unambiguous pattern language, which is subcritical for T. Let $P[T](m, d)$ denote the number of trees from $P[T]$ of size m containing exactly d pattern leaves, and $w(d)$ be a nonzero polynomial of degree γ. Then,

$$\lim_{m\to\infty} \frac{\sum_{d\in\mathbb{N}} P[T](m, d)w(d)}{T(m)} = c_w$$

for some nonnegative real c_w. If additionally $w(d)$ has nonnegative values for all elements of $\mathbb{N}$ and there exists integer $r \geq \gamma$ for which $w(r) > 0$ and P contains a pattern with r regular leaves and at least one placeholder, then $c_w \neq 0$.

Proof. (Sketch) The function $\sum_{d\in\mathbb{N},h\in\mathbb{N}} x^d y^h P(d, h)w(d)$ can be represented as a sum of partial derivatives of the GF $p(x, y) = \sum_{d\in\mathbb{N},h\in\mathbb{N}} x^d y^h P(d, h)$ of P with respect to x. The subcriticality of $P(x, y)$ transfers to the subcriticality of $\sum_{d\in\mathbb{N},h\in\mathbb{N}} x^d y^h P(d, h)w(d)$ for T. With subcriticality one can show that $P_w(z) := P_w(x, y)|_{(z,T(z))}$ has unique dominant square-root singularity in ρ in the case with the additional assumptions (or radius of convergence strictly

greater than ρ in the general case). Moreover, $P_w(z)$ is the GF of the sequence $\sum_{d\in\mathbb{N}} P[T](m,d)w(d)$, therefore, $\lim_{m\to\infty} \frac{\sum_{d\in\mathbb{N}} P[T](m,d)w(d)}{T(m)} = c_w$. □

The next result is the main tool for the later analysis, it counts the expressions with a fixed number of restrictions among the pattern leaves. This is the first facet of the simultaneous 'classification' of pattern theory and the main tool of further analysis.

Theorem 4.21 ([Koz08, Lemma 2.8])
Let P be an unambiguous pattern language, which is subcritical for $\mathcal{T}$ and let $E \subset \{x_1, ..., x_n\}$ have cardinality l. We denote by $\mathcal{F}_n^{[k]}(P[T])(m)$ (resp. $\mathcal{F}_n^{[\geq k]}(P[T])(m)$) the number of expressions from $\mathcal{F}_n$ of size m whose structure belongs to $P[T]$ and which have k (resp. at least k) (P,E)-restrictions. For every $k \in \mathbb{N}$ for which P contains a pattern with at least $k+1$ pattern leaves and at least one placeholder, we have

$$\lim_{m\to\infty} \frac{\mathcal{F}_n^{[\geq k]}(P[T])(m)}{[z^m]F_n} \underset{n\to\infty}{\sim} \lim_{m\to\infty} \frac{\mathcal{F}_n^{[k]}(P[T])(m)}{[z^m]F_n} \underset{n\to\infty}{\sim} \frac{c_{k,l}}{n^k},$$

for some $c_{k,l} \in \mathbb{R}_+$.

Proof. (Sketch) $\left\{ {a \atop b} \right\}$ will denote the number of partitions of a set of size a into b nonempty subsets (Stirling number of the second kind) and $n^{\underline{k}}$ the falling factorial: $n^{\underline{k}} := n \cdot (n-1) \cdot ... \cdot (n-k+1)$.

Let $P[\mathcal{T}](m,d)$ be the number of structures from $P[\mathcal{T}]$ of size m that have exactly d pattern leaves as in the lemma above. For fixed number of variables n the number of possible leaf labels of such a structure so that the constructed expression has exactly k restrictions among the d P-pattern leaves is

$$\sum_{r=0,...,k} \left\{ {d \atop d-r} \right\} \cdot \binom{l}{k-r} \cdot (d-r)^{\underline{k-r}} \cdot (n-l)^{\underline{d-r-(k-r)}} \cdot n^{m-d} \cdot 2^m,$$

whereas the sum is taken over the number of repetitions which can be at most the number of restrictions (k). So the number of expressions from $\mathcal{F}_n$ of size m with d pattern leaves and n restrictions is $P[\mathcal{T}](m,d) \cdot w_{k,l}(d) \cdot (n-l)^{\underline{d-k}} \cdot n^{m-d} \cdot 2^m$ with $w_{k,l}(d) := \sum_{r=0,...,k} \left\{ {d \atop d-r} \right\} \cdot \binom{l}{k-r} \cdot (d-r)^{\underline{k-r}}$ being a

polynomial in d (for fixed k, l) (See [Koz08, Observation 2.5 and Observation 1.2]). So we get

$$\frac{\mathcal{F}_n^{[k]}(P[T])(m)}{[z^m]F_n} = \frac{2^m \sum_{d\in\mathbb{N}}(P[\mathcal{T}](m,d)\cdot w_{k,l}(d)\cdot (n-l)^{\underline{d-k}}\cdot n^{m-d})}{(2n)^m[z^m]T_n}$$

and an upper bound

$$\frac{\mathcal{F}_n^{[k]}(P[T])(m)}{[z^m]F_n} \leq \frac{\sum_{d\in\mathbb{N}}(P[\mathcal{T}](m,d)\cdot w_{k,l}(d)\cdot n^{d-k}\cdot n^{m-d})}{n^m[z^m]T_n}. \tag{4.18}$$

From Theorem 4.20 we know that this bound is asymptotically equivalent to $\frac{c_{k,l}}{n^k}$. A lower bound estimate of $(n-l)^{\underline{d-k}}$ ([Koz08, Lemma 2.6]) which shows that $(n-l)^{\underline{d-k}}$ is well estimated by n^{d-k} for big k eventually completes the case for exactly k restrictions. For at least k restrictions observe that the upper bound estimate of the expressions of size m with exactly k restrictions and d pattern leaves that leads to the numerator in (4.18) is also an upper bound for the expressions with at least k restrictions.

□

The first result that is established within this framework is that all tautologies are simple tautologies, whereas a simple tautology is a tautology that has a leaf labeled by x and a leaf labeled by $\overline{x}$, for a variable x, which are connected to the root by nodes labeled only by $\vee$. Such a path is called $\vee$-only path.

Theorem 4.22 (Tautologies,(cf. [Woo05],[Koz08, Theorem 3.2],[GGKM15, Theorem 3.6.]))
The density of tautologies among Catalan And/Or trees with n variables asymptotically (with n) equals the density of simple tautologies and

$$\mathbb{P}_n(TRUE) = \frac{3}{4n} + \mathcal{O}(\frac{1}{n^2}).$$

Proof. (Sketch) The proof in [Koz08] that 'all' tautologies are simple makes uses of the pattern $N[N]$ with similar considerations as in the following proposition 4.24 and therefore will be omitted. In [GGKM15] it is shown with Theorem 4.21 that the simple tautologies are asymptotically equivalent to simple tautologies realized by just one variable. The authors are then able to compute the constant in the asymptotic form of the probability of this set of 'simple' expressions directly from its GF. □

The analogous statement also holds for contradictions due to symmetry reasons. Before we state the main result from this section of the thesis we need another definition.

Definition 4.23 (Expansions, [GGKM15, Definition 5.2])
Let t be an And/Or tree computing f, v one of its nodes and t_v the subtree rooted at v. An expansion of t in v is a tree obtained by replacing the subtree t_v rooted at v by a tree $t_v \Diamond t_e$ or $t_e \Diamond t_v$, where $\Diamond \in \{\wedge, \vee\}$ and where t_e is an And/Or tree. We will say that such an expansion is valid when the expanded tree still computes f.

Proposition 4.24 (cf. [Koz08, Lemma 3.4] and [GGKM15, Proposition 5.5])
The set of non-negligible trees computing a non constant Boolean function f is the set of trees obtained by expanding a minimal tree of f once. Moreover, the only non-negligible valid expansions are:

- *The T-expansions: a valid expansion is a T-expansion if the inserted subtree t_e is a simple tautology (resp. a simple contradiction) and if the new label of v is $\wedge$ (resp. $\vee$).*
- *The X-expansions: a valid expansion is an X-expansion if the inserted subtree t_e is (up to commutativity and associativity) of the shape $x \vee \ldots$ (resp. $x \wedge \ldots$) where x is an essential variable of f and if the label of the father of t_e is $\wedge$ (resp. $\vee$).*[24].

Proof. Define $r := \mathcal{L}(f)$ and consider the pattern language $R = N^{(r+1)}[P \oplus N]$. The pattern $P \oplus N$ is the pattern so that a leaf of a structure is a $P \oplus N$-pattern leaf iff it is a P-pattern leaf or an N-pattern leave. It is straightforward that R is unambiguous and subcritical for $\mathcal{T}$. We say that an R-pattern leaf is on level i if it is an $N^{(i)}$-pattern leaf, but no $N^{(i-1)}$-pattern leaf. The index i reaches from 1 to $r+1$ and there are no N^0 pattern leaves. An R-pattern leaf that is not an $N^{(i+1)}$ pattern leaf is said to be on level $r+2$.

Let t be a minimal tree, i.e. a tree of size $\mathcal{L}(f)$ computing f and E the set of essential variables of f. We will only consider (P, E)-restrictions, with P a

[24] See 4.5 for an illustration of the subtree t_e for an X-expansions

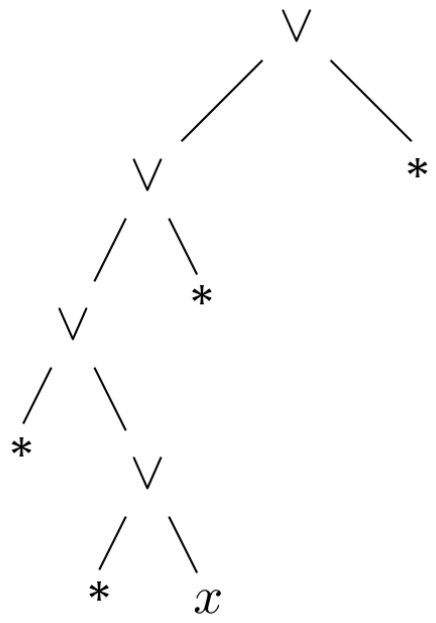

Figure 4.5: The shape of tree t_e that is used for an X-expansion. There has to be a path from the root to variable x where inner nodes are labeled only by $\vee$ (resp. $\wedge$). Such a path will be called $\vee$-only-path (resp. $\wedge$-only-path).

pattern, so for convenience we will call them P-restrictions. At first we show that

$$\mathbb{P}_n(f) \geq \frac{a}{n^{\mathcal{L}(f)+1}}.$$

For this consider all expressions of the form $s \wedge t$ with s a simple tautology. All these expressions compute f and let us denote the GF of simple tautologies by $S(z)$. The function $z^{\mathcal{L}(f)}$ is the GF of the single tree t so $S(z)z^{\mathcal{L}(f)}$ is the GF of expressions of the above form. With 2.20, 4.22 and the identity $(S(z)z^{\mathcal{L}(f)})' = S'(z)z^{\mathcal{L}(f)} + S(z)\mathcal{L}(f)z^{\mathcal{L}(f)-1}$ we get $\mathbb{P}_n(f) \geq \frac{d}{n}\rho_n^{\mathcal{L}(f)} = \frac{a}{n^{\mathcal{L}(f)+1}}$. From Theorem 4.21 we know that trees with at least $\mathcal{L}(f)+2$ R-restrictions can be neglected for computing the constant λ. The trees that have no leaves on level $r+2$ can also be neglected because due to subcriticality of R we can show that the GF of such trees has a radius of convergence that is larger than

ρ_n and therefore the contribution to the probability is zero. We now show that such trees computing f have at least $r+1$ restrictions among R-pattern leaves.

Consider the contrary: Let t be a term with at most r R-restrictions and at least one leaf on the level $r+2$. Define i as the first level of t in which there is no occurrence of a restriction, i.e. the smallest number so that t has the same number of $N^{(i-1)}$-restrictions that of $N^{(i)}$-restrictions. Since there is at least one restriction on the first level of t[25], i is not greater than $r+1$. We consider two cases:

First case: t has at most $r-1$ $N^{(i)}$-restrictions. Then, there are no restrictions on level i and so we can substitute all leaves in level i with FALSE without changing the function. Using the property of pattern language N we can replace all subtrees rooted at level i by $FALSE$. Finally, substituting $FALSE$ for all remaining non essential variables[26] and simplifying the tree to get rid of all leaves labeled by a constant leads to a tree, with at most $r-1$ restrictions among all leaves, still computing f. Since all remaining leaves are labeled by essential variables this means that t has size at most $r-1$, a contradiction.

Second case: t has exactly r $N^{(i)}$-restrictions. Since t has at most r R-restrictions there are no R-restrictions below level i and $i \leq r+1$. So all leaves on level $r+2$ (these are the leaves which are P-pattern leaves but not $N^{(r+1)}$-pattern leaves) are labeled by variables that are not essential and occur only once among all P-pattern leaves. Since there is at least one P-pattern leaf on level $r+2$ there is at least one node v on level $r+2$ with parent on level $r+1$. The subtree t_v rooted at v has at least one $P \oplus N$-pattern leaf and with the property of the pattern language $P \oplus N$ we know that we can substitute the $P \oplus N$-pattern leaves of t_v by $TRUE$ and $FALSE$ so that the subtree t_v valuates to the value $TRUE$ or $FALSE$ we wish. So we can substitute every node v on level $r+2$ with parent on level $r+1$ by a constant we wish, without changing the function the tree computes because all these variables are inessential. The same holds for all P-pattern leaves that are

[25] Otherwise one can derive with the property of pattern N the contradiction that f is constant.

[26] Now variables are substituted and not nodes, because there might be repetitions.

not labeled by a variable counting to the restrictions. Because at this stage we do not know which values we should substitute, we substitute all these nodes by a wildcard $*$ to obtain a tree t^*. Then, we use the following rules and symmetric variants of them

$$\begin{array}{rr} *\vee *\triangleright * & *\wedge *\triangleright * \\ *\vee \varphi\triangleright TRUE & *\wedge \varphi\triangleright FALSE \\ TRUE\vee \varphi\triangleright TRUE & FALSE\wedge \varphi\triangleright FALSE \\ FALSE\vee \varphi\triangleright \varphi & TRUE\wedge \varphi\triangleright \varphi \end{array} \tag{4.19}$$

to trim the tree t^* to finally obtain an expression (so there are no wildcards left) t' that still computes function f. The symbol φ denotes an expression which does not contain any wildcard. All remaining variables in t^* are either essential variables or repetitions. During the simplification process a rule from the second row has to be used at least once because all $*$ are eliminated during the process. Applying such a rule eliminates at least one restriction, so the expression t' contains at most $r-1$ restrictions and therefore its size is $r-1$, a contradiction.

So the non-negligible trees computing f have at least one leaf on level $r+2$ and exactly $r+1$ restrictions among R-pattern leaves. Let t be such a tree and consider the pattern language $R' = N^{(r+1)}[(P\oplus N)^2]$. It is not possible that t has an R'-restriction that is not an R-restriction (so a restriction on level $(r+3)$). This is true because on the one hand there cannot be a restriction on both level $(r+2)$ and $(r+3)$ because this would lead to a contradiction as in the first case above. On the other hand if there would be a restriction on level $(r+3)$ but none on level $(r+2)$, then node v on level $(r+2)$ with parent on level $(r+1)$ can be substituted by $*$. This wildcard now eliminates at least two R'-restrictions, the one on level $(r+3)$ and one due to using at least once a simplification rule of the second row which leads to a contradiction as in the second case. So there is no restriction on level $(r+3)$ and the non-negligible trees have at least one leaf on level $(r+3)$[27].

We now show that these non-negligible trees t' are valid expansions of minimal trees t. As in the second case above the node on level $(r+3)$ with parent on level $(r+2)$ and all the other variables that are no restrictions

[27] See above for the same consideration for level $(r+2)$.

can be substituted by a wildcard $*$ to obtain tree t^*. So there is at least one wildcard in t^* and applying the rules in (4.19) to t^* gives a minimal tree t of f, so all $*$ disappeared. The $*$'s can only decrease by applying a rule of the first or second row. So a rule from second row is applied exactly once and the other times a rule of the first row is applied. This means that there is a node v of t^* so that the subtree rooted at v has only $*$'s as leaf labels and there are no other $*$ anywhere else in t^*. Therefore, v evaluates to $*$ according to the rules. The parent of the corresponding node of v in t', lets say v' is the root of the tree t_e used in the expansion and the parent of v' in t is the node where the expansion takes place. So all non-negligible trees are indeed expansions of minimal trees.

The last step in the proof is to show that T-expansions and X-expansions are the only non-negligible expansions. For this we consider the pattern language $(P \oplus N)^2$ and all commutative variations of them. From Theorem 4.21 we know that expansions with trees t_e that have at least two $(P \oplus N)^2$-restrictions in a variation are negligible. Therefore, t_e has exactly one $(P \oplus N)^2$-restriction in every variation[28]. We distinguish two cases.

There is a variation of $(P \oplus N)^2$ for that the restriction is not an essential variable of f: The minimal tree that is expanded contains only essential variables. So if t_e does not compute $TRUE$ or $FALSE$ we can valuate t_e independently from the rest of the tree to a value we wish. This would give a tree smaller than the minimal tree still computing f. So t_e computes $TRUE$ or $FALSE$ and therefore is a tautology resp. contradiction and with Theorem 4.22 we conclude that we can assume that t_e is a simple tautology resp. simple contradiction. Since the expanded tree still computes f the new node for the expansion is an $\wedge$ if t_e is $TRUE$ and an $\vee$ if t_e is $FALSE$.
For every variation of $(P \oplus N)^2$ the restriction is an essential variable of f: For every variation of $(P \oplus N)^2$ its restriction is on level one[29]. Take a fixed variation of $(P \oplus N)^2$ and call the essential variable of the restriction x. If there is an $\wedge$ and an $\vee$ on the path from the root of t_e to x then there is a variation of $(P \oplus N)^2$ so that x is on level two and therefore this variation has its restriction on level two, a contradiction. $\square$

[28] Otherwise one can evaluate this tree to $FALSE$ or $TRUE$ independently from the rest of the tree, to obtain a tree still computing f with complexity smaller r.

[29] Again the contrary would lead to a tree computing f smaller than the minimal tree.

Theorem 4.25 (cf. [Koz08, Theorem 3.3] and [GGKM15, Proposition 5.1]) *Let f be a Boolean function, whose complexity is denoted by $\mathcal{L}(f)$. Then,*

$$\mathbb{P}_n(f) \underset{n\to\infty}{\sim} \frac{\lambda_f}{n^{\mathcal{L}(f)+1}},$$

where λ_f is depending on the number of possible expansions of minimal trees of f. We have

$$\frac{8\mathcal{L}(f)-3+l}{16^{\mathcal{L}(f)}}M_f \leq \lambda_f \leq \frac{4\mathcal{L}(f)^2+4\mathcal{L}(f)-3}{16^{\mathcal{L}(f)}}M_f$$

where M_f is the number of minimal trees representing f and

$$l = \begin{cases} \lceil \frac{\mathcal{L}(f)}{2} \rceil & \text{for } \mathcal{L}(f) > 1 \\ 0 & \text{for } \mathcal{L}(f) = 1 \end{cases}$$

Proof. (Sketch) Let M_f be the number of minimal trees computing f with the corresponding trees t_1,...,t_{M_f}. Let $\lambda_i^T(f)$ denote the number of possible T-Expansions of tree t_i and ω_1 the limit probability of simple tautologies (resp. simple contradictions) and $\lambda_i^X(f)$ resp. ω_2 the corresponding numbers for X-Expansions. Theorem 4.22 shows that ω_1 is $\frac{3}{4n}$ and direct computations with GFs in [GGKM15] show that $\omega_2 = \frac{1}{2n}$. From Propasition 4.24 we know that the probability of f is asymptotically equivalent to the limit probability of trees that are X- resp. T-Expansions of minimal trees of f. So with Lemma 2.20 this gives

$$\mathbb{P}_n(f) \underset{n\to\infty}{\sim} \sum_{j=1}^{M_f} \rho_n^{\mathcal{L}(f)}(\lambda_j^T(f)\omega_1 + \lambda_j^X(f)\omega_2) = \frac{\lambda_f}{n^{\mathcal{L}(f)+1}}.$$

Upper and lower bounds on the numbers $\lambda_i^T(f)$ and $\lambda_i^X(f)$ eventually give the bounds in the theorem. □

The above approach is generalized in [GGKM15] to handle also the proofs for the corresponding results for the commutative and/or associative models $(\mathbb{S}_2, ..., \mathbb{S}_4)$ to obtain the results listed in 4.4.

The Shannon effect is disproved in [GGM14] for model $\mathbb{S}_1$. With Theorem 4.21 it is easy to show that the limit probability of the set of functions of

constant complexity is $\mathcal{O}(1/n)$ and therefore converges to 0 as n tends to infinity (cf. [GGM14, Corollary 7]). So to disprove the Shannon effect one has to consider a family of functions with complexity tending to infinity as n tends to infinity. Such considerations are very different from the ones above because there the probability of a fixed function hence constant (i.e. small) complexity is investigated. This question requires another approach as above. With a different point of view of Catalan And/Or trees the authors prove that the limit probability of functions that have complexity of order $\Theta(n^{2+\epsilon})$ for an $\epsilon > 0$, is bounded from below by a positive constant for n tending to infinity, which disproves the Shannon effect. They consider Catalan trees as a tree of $\wedge$-labels (resp. $\vee$-labels) whose leaves are substituted by variables or by a Catalan And/Or tree rooted by an $\vee$ (resp. $\wedge$).

Observe that in the above considerations we let the size of the expressions tend to infinity ($m \to \infty$) at first (for constant n) to obtain the limit probability $\mathbb{P}_n$. Subsequently we let n tend to infinity to obtain the asymptotic behavior of $\mathbb{P}_n(f)$. In [GM15] the authors note that this leads to the fact that the considered expressions have a lot of repetitions in their leaves. They ask the question if this biases the induced distribution on Boolean functions. For this they consider the probability distribution $\mathbb{P}_m := \mathbb{P}_{m,n(m)}$ for the Catalan And/Or model with a sequence $(n(m))_{m\geq 1}$ that tends to infinity with m. So the size of the expressions and the number of variables tend together to infinity. The two different approaches are pictured in Figure 4.1[30]. As in the approach from above, where the limit for the size and the number of variables is separated, the asymptotic behavior of $\mathbb{P}_m(f)$ for a fixed function f satisfies:

Theorem 4.26 ([GM15, Theorem 1])
Let $(n(m))_{m\geq 1}$ be an increasing sequence of integers tending to infinity when m tends to infinity. Let f be a Boolean function f with complexity $\mathcal{L}(f)$, then there exists a positive constant λ_f such that

$$\mathbb{P}_m(f) \underset{m\to\infty}{\sim} \frac{\lambda_f}{n(m)^{\mathcal{L}(f)+1}}.$$

This result is proven with pattern theory whereas a modification of Theorem

[30] For us $m(k)$ equals k.

4.21 is needed. With this modification and some extra effort the line of the proof follows the argumentation considered above. The theorem allows to solve the Catalan satisfiability problem:

Corollary 4.27 ((Catalan- SAT)[GM15, Corollary 1])
Let $(n(m))_{m\geq 1}$ be an increasing sequence of integers tending to infinity when m tends to infinity. Pick up uniformly at random a Catalan And/Or expression of size n with leaf-labels in $\{x_1, \overline{x_1}, ..., x_{n(m)}, \overline{x_{n(m)}}\}$. This expression is satisfiable with probability tending to 1 when m tends to infinity.

With the result Theorem 4.17 in the previous model one can conclude the above theorem and corollary for a sequence $(n(m))_{m\geq 1}$ that grows sufficiently slow. But the two results above guarantee the asymptotic behavior resp. the satisfiability observation for every (probably very quickly) increasing sequence $(n(m))_{m\geq 1}$ tending to infinity.

They also consider a model which they call quotient model and derive analogue results in parallel. In the quotient model they consider equivalence classes of Boolean expressions and the induced equivalence relation for Boolean functions. Two expressions e_1, e_2 are considered equivalent when their underlying labeled trees without the leaves are equal, leaves labeled by literals of the same variable in e_1 are labeled by literals of the same variable in e_2 and leaves labeled by the same literals in e_1 are labeled by the same literals in e_2. So e_1, e_2 are equivalent when the variables in e_1 can be 'consistently' substituted to get e_2. This induces an equivalence relation on the Boolean functions: Let f be a Boolean function, then define $< f >$ as the class of all Boolean functions which have an expression that is equivalent to an expression of f. The probability of a class of equivalent functions $< f >$ is then defined as the number of equivalence classes of expressions that compute $< f >$[31]. This model exhibits an interesting threshold or rather saturation phenomena:

Theorem 4.28 ([GM15, Theorem 1])

[31]More exactly: An equivalence class of expressions $< E >$ computes an equivalence class of functions $< f >$ iff there is an representative e of $< E >$ that computes a function g that is a representative of $< f >$.

Let $(n(m))_{m\geq 1}$ be an increasing sequence of integers tend to infinity when m tends to infinity. There exists a sequence $(M_m)_{m\geq 1}$ such that $M_m \underset{m\to\infty}{\sim} \frac{m}{\ln m}$ and such that for all fixed equivalence classes of Boolean functions $< f >$ with $R(< f >)$ being the complexity of $< f >$ minus the number of essential variables of $< f >$[32], there exists a positive constant $\lambda_{<f>}$ satisfying:

(i) If (for all sufficiently large m) $n(m) \leq M_m$, then

$$\mathbb{P}_m(< f >) \underset{m\to\infty}{\sim} \frac{\lambda_{<f>}}{n(m)^{\mathcal{R}(<f>)+1}}.$$

(ii) If (for all sufficiently large m) $n(m) \geq M_m$, then

$$\mathbb{P}_m(< f >) \underset{m\to\infty}{\sim} \frac{\lambda_{<f>}}{(\frac{m}{\ln m})^{\mathcal{R}(<f>)+1}}.$$

When we want to compare this to the the result Theorem 4.17 and then observe that there are $\binom{n}{e}2^e$ different functions in the equivalence class of a function f with e essential variables. It holds that $\binom{n}{e}2^e \sim n^e$ so the probability of the set of equivalent functions denoted by F evaluates to $\mathbb{P}_n(F) \underset{n\to\infty}{\sim} \frac{\lambda_f}{n^{\mathcal{L}(f)+1}} n^e = \frac{\lambda_f}{n^{\mathcal{R}(F)+1}}$.

It is clear that the system is saturated when the number of variables $n(m)$ is m because a tree of size m can be labeled with at most m variables and in the quotient model it is not important which (at most) m-element subset of the set of variables is used for the labeling. The result above shows that this saturation happens earlier (for smaller growing $m(n)$). Indeed the system is saturated when the number of variables grows faster than $\frac{m}{\ln m}$. This saturation comes from the saturation of the multiplicative factor with which the number of different equivalent labellings of a structure of size $m+1$ grows compared to the number for m. Technical calculations show that this multiplicative factor grows as $2n(m)$ for $n(m) \leq M_m$ and as $\frac{2m}{\ln m}$ for faster growing $n(m)$ (cf. [GM15, Proposition 3]). Let us now finish this section with a brief discussion of the tree size model $(\mathbb{S}_5)$.

The model $(\mathbb{S}_5)$, which is the model of associative trees with tree-size complexity, is in detail investigated in [GGKM14], where the authors find a lot of

[32]It is easily seen that this number is well defined.

similarities as well as differences between the formula size model(s) and the tree size model. In words of the authors: **"We did not expect much difference, but first experiments indicated a rather different behavior. Indeed, when we tried the method which worked for And/Or trees to prove the results we expected, we failed. [. . .] Though it will eventually turn out that the trees in the tree size model very well fall under the same paradigms we encountered in many formula size models, it requires a technically very different treatment to obtain those results. [. . .] However, there is also a clear difference: Whereas the limiting probabilities occurring in the formula size models differ among themselves by constant factor, the tree size model gives a strong bias to the constant functions True and False."-[GGKM14]**

With 'the method' the authors mean pattern theory. The same paradigms they encounter is at one hand that there is a strong relation between the limit probability and the complexity. Moreover, the relation is of the same asymptotic form as for the formula size model for not constant functions. On the other hand most trees that compute a given function are a minimal tree expanded[33] once. The constant functions have a probability that is bounded from below which immediately implies that the Shannon effect does not hold for this model.

As the part 'in many formula size models' of the quote suggests, similar behavior concerning expansions and asymptotic behavior is observed in other tree size models, namely in the model with implication that we discuss very briefly in the next section.

4.3.2 $(\rightarrow)$-Expressions

In this section we consider two different models. The first is the model of binary trees with inner nodes labeled by $\rightarrow$ and leaves labeled by (positive) variables $\{x_1, ..., x_n\}$, all with the same probability, so we are again considering the uniform distribution on the expressions. The second model incorporates the property $x \rightarrow (y \rightarrow z) = y \rightarrow (x \rightarrow z)$ of the implication

[33] Clearly the expansions in this model are not the same as the expansions considered above.

into the model. The property shows that the order of the premisses of a formula of the form $A_1 \to (A_2 \to (A_3 \to (...A_{k-1} \to (A_k \to \alpha))))$, is irrelevant for A_i being arbitrary expressions. So an expression of the second model is either a single variable from $\{x_1, ..., x_n\}$ or an expression of the form $\{A_1, ..., A_k\} \to \alpha$, with A_i expressions. This can be seen as a tree labeled by α with (unordered) children $A_1, ..., A_k$. Again the uniform distribution on such trees is considered. The size function in both models is the number of leaves (formula size). We will refer to the first model as the binary model and to the second the generalized model.

In both models the same Boolean functions are expressible and it is clear that not every function is expressible. In fact, the set of functions that are expressible is the Post class S_0 [RW00] (or T_0^∞ in other resources). This is the set of functions $f \in \mathbb{B}_n$ so that there is an $i \in \{1, ..., n\}$ so that $f(x_1, ..., x_n) \leq x_i$ for all $(x_1, ..., x_n)$, so $f = x_i \vee g$ with an arbitrary function g; such functions are called 0-separating functions.

As the authors of [FGGG12] note, the models with implication are interesting for different fields of research for a couple of reasons: From the point of view of logic the implicational fragment is interesting because of the importance of modus ponens in propositional calculus and because it plays an important role in intuitionistic logic. Moreover, from the satisfiability point of view the relation to Horn formulas is interesting. For the scientists working in the field of quantitative logic it is also of special interest because it is somehow the easiest model[34] that is still rich enough to incorporate interesting properties and therefore it was also kind of a starting point in their studies.

The first who was interested in a quantitatively study of systems with implication was the Polish school around Zaionc who started a systematic research of the density of truth in several logical systems. Systems with implication are considered in [MTZ00, Kos03, Zai05]. In [Zai03, KZ04, FGGZ07, GK09] intuitionistic logic is quantitatively compared with classical logic by comparing the number of tautologies in both systems whereas [FGGZ07] is of special interest for us because it is shown that most[35] tautologies are of simple shape and the asymptotic behavior of the limit probability (in exactly our sense)

[34] In the way that there is only one connector.

[35] In the sense of the previous section.

of tautologies is precisely quantified as $\frac{1}{n}$.

The first who investigated the probability of general functions (not only the function TRUE) in the (binary) model with implication were the authors of [FGGG08] (a preliminary version of [FGGG12])and some of the authors then disproved the Shannon effect in [GG10]. Subsequently the results were extended to the generalized model in [GGKM12]. Let us state in the following the most important results of these papers.

The existence of the limit probability in both models is proven with the DLW theorem for polynomial systems resp. with the extended version for general analytic functions. The irreducibility and aperiodicity follow from Lemma 4.14 (and the considerations below). The limit probability of a fixed function f has the same asymptotic form as for the And/Or models. Let us ad hoc define the complexity of the function TRUE as 0. In the binary model:

Theorem 4.29 ([FGGG12, Theorem 7])
Let f be a Boolean function with complexity $\mathcal{L}(f)$. Then,

$$\mathbb{P}_n(f) \underset{n\to\infty}{=} \frac{\lambda_f}{4^{\mathcal{L}(f)} n^{\mathcal{L}(f)+1}} + \mathcal{O}(\frac{1}{n^{\mathcal{L}(f)+2}}),$$

with λ_f a constant independent of n.

In the generalized model:

Theorem 4.30 ([GGKM12, Theorem 1])
Let f be a Boolean function with complexity $\mathcal{L}(f)$. Then,

$$\mathbb{P}_n(f) \underset{n\to\infty}{=} \frac{\lambda_f}{n^{\mathcal{L}(f)+1}} + \mathcal{O}(\frac{1}{n^{\mathcal{L}(f)+2}}),$$

with λ_f a constant independent of n.

Both results are proven by means of valid expansions of minimal trees: It is shown that in both models the non-negligible trees computing a given function f are minimal trees expanded once. Expansions in these models

with implication are of different shape than the expansion considered in the And/Or model(s) (cf. [FGGG12, Definition 4] resp. [GGKM12, Definition 14]). Both models do not exhibit the Shannon effect. In the normal model the following holds:

Theorem 4.31 ([GG10, Theorem 2])
Let $R = 9\pi k^2/16$. Then, the probability of all functions of complexity at most R is larger than or equal to $9/64$, when the number of variables k tends to infinity.

In the generalized model a positive bound is given:

Theorem 4.32 ([GGKM12, Theorem 3])
If $g(n)$ is a function in n growing faster than n^2, i.e. $n^2 = o(g(n))$, then

$$\lim_{n\to\infty} \mathbb{P}_n(\{f|\mathcal{L}(f) \leq g(n)\}) \geq \alpha > 0.$$

Both results show that there is no Shannon effect in the corresponding models because the expressible functions are exactly the functions of the form $x \vee g$ and there are at least $2^{2^{n-1}}$ such functions which is large enough to show exactly as in Example 3.3 that the maximal complexity of an expressible function is (at least) exponential in n.

4.3.3 $(\top, \bot)$-Expressions

Let us finally consider in this section a very simple[36] model investigated by Yashunskii in [Yas05, Yas06, Yas07]: The model has base functions $\{\top, \bot\}$ and an arbitrary connector set $\mathbb{K}$. The size function is the number of connectors. The leaf $\top$ occurs with probability p and $\bot$ with $1-p$, the connectors occur all with the same probability.[37] Clearly the only functions that are expressible in this model are $\top$ and $\bot$, so the probability distribution induced

[36]Simple in the sense that there are only two functions occurring in the model, but nevertheless, not simple to analyze.

[37]Nevertheless, we allow the set $\mathbb{K}$ to contain connectors more than once, i.e. we are considering a multiset which then introduces a probability distribution on the (distinct) connectors.

on the expressible functions is fully characterized by

$$P(p) := \mathbb{P}_p(\top) := \lim_{m\to\infty} \mathbb{P}_{m,p}(\top),$$

the limit probability of TRUE when the base function TRUE has probability p. In [Yas05] the author shows that $P(p)$ exists for arbitrary $\mathbb{K}$ and $p \in (0,1)$ and is continuous on $(0,1)$ as a function of p ($\mathbb{K}$ is fixed).[38]. He also gives an explicit formula for the value $P(p)$ that depends only on two polynomials directly related to the arities of the connectors and the number of 1's in their truth tables. Let c_i be the number of connectors with arity i and let r be the largest arity of connectors in $\mathbb{K}$. Define

$$B(x) := c_1 x + c_2 x^2 + ... + c_r x^r$$

and

$$A(t,f) := \sum_{i=1}^{r} \sum_{j=0}^{i} a_{ij} t^j f^{i-j},$$

with a_{ij} the total number of 1's in the truth tables of the connectors with arity i on assignments with j 1's (i.e. with weight j). With these polynomials the generic equations are:

$$\begin{aligned} T(z) &= p + zA(T(z), F(z)) \\ F(z) &= (1-p) + z(B(T(z) + F(z)) - A(T(z), F(z))) \end{aligned} \tag{4.20}$$

with $T(z)$ (resp. $F(z)$) the GF of the expressions computing the function TRUE (resp. FALSE). The following theorem is directly proven with the help of the characteristic equations from the DLW Theorem for the general case when the system is irreducible and a careful analysis of the cases when the system is not irreducible.

Theorem 4.33 ([Yas06])
For any fixed p, $0 < p < 1$, there exists the limit $P(p)$ with

$$P(p) = \frac{A_y(\tau, \sigma - \tau)}{\omega^{-1} - A_z(\tau, \sigma - \tau) + A_y(\tau, \sigma - \tau)},$$

[38] For not degenerated sets of connectors this is an easy consequence of 4.13.

where ω and σ are real numbers, which form the unique solution of the system of equations:

$$\begin{aligned} \sigma &= 1 + \omega B(\sigma) \\ 1 &= \omega B'(\sigma) \end{aligned} \tag{4.21}$$

under the additional claim that the value of $|\omega|$ is minimal, while A_z and A_y are the partial derivatives of $A(z, y)$ and $\tau = \tau(p)$ is the uniquely determined algebraic function satisfying the equation

$$\tau(p) = p + \omega A(\tau(p), \sigma - \tau(p)).$$

In [Yas06] he also shows that the function $P(p)$ is infinitely often continuously differentiable on $(0, 1)$ and he also gives exact conditions when the limit probability on the boundary points $0, 1$ exists and when the function $P(p)$ is continuous on the whole interval $[0, 1]$. Moreover, he gives explicit expressions for a couple of different sets of connectors $\mathbb{K}$. His motivation for investigating such models was the following (cf.[Yas05]): Imagine a 'black box' that chooses a large random expression with connectors in $\mathbb{K}$ and with leaves labeled by TRUE with probability p and by FALSE with probability $1 - p$. The 'black box' then returns the value of the expression for the assignment. There arise two questions (cf.[Yas05]):

(1) Imagine $\mathbb{K}$ is known and p is unknown but fixed. Is it possible to obtain p from the outputs[39] of the 'black box'?

(2) Imagine $\mathbb{K}$ is unknown, but perhaps some properties are known (e.g. number of connectors, arities etc.). Is it possible to acquire some more information about $\mathbb{K}$, through the knowledge of the function $P(p)$[40]?

The answer of the first question in general is 'No' because there exist sets of connectors of maximal arity 3 for which $P(p)$ is not injective, see Figure 4.6. Moreover, the following holds:

Theorem 4.34 (Uniform approximation [Yas07, Theorem 4])
Let $f(p)$ be a continuous function mapping the segment $[0, 1]$ into itself. Then,

[39] These outputs are nothing else than draws from $\{0, 1\}$ according to $P(p)$.

[40] This function can be approximated with the help of a lot of outputs of the 'black box' for different p.

for any $\epsilon > 0$ there exist connectors $\mathbb{K}_\epsilon$ whose functions are all different and essentially depend on all their variables so that the function $P(p)$ is such that for all $p \in [0,1]$ the relation $|f(p) - P(p)| < \epsilon$ holds.

This means that every continuous function $f(p) : [0,1] \mapsto [0,1]$ can be uniformly approximated with functions $P(p)$. So in general the function $P(p)$ is not injective. For binary bases he shows that $P(p)$ is either strictly monotonous or constant for $0 < p < 1$ and also derives conditions for both cases. So in the case of binary bases with $P(p)$ not constant the answer is 'Yes'.

The answer of the second question for the 'black box' model is 'No' in the general case because there are different sets of connectors which have the same function $P(p)$. This is clear because there are different sets of connectors with the same polynomials $B_1 = B_2$ and $A_1 = A_2$, and there are even cases where $P(p)$ is the same for sets of connectors having different polynomials A, see [Yas05].

When trying to apply the results in this area to other areas of research then Yashunskii's model has definitely something to offer. The 'black box' considerations of Yashunskii are very interesting and somehow answer what is not possible to recognize when we only know the function $P(p)$. On closer inspection it directly offers an excellent statement regarding satisfiability of monotone read-once expressions: When one steps back and reconsiders the whole construction of the function $P(p)$, then $P(p)$ is nothing else than the probability that an expression, chosen uniformly from expressions built with connectors $\mathbb{K}$ of 'large' size m and leaves labeled independently by TRUE (resp. FALSE) with probability p (resp. $1-p$), is TRUE. The fact that all leaves are independently labeled allows to put it like this: In a random monotone read-once formula $e_{\mathbb{K}}$ all variables are independently substituted by TRUE or FALSE with probability p resp. $1-p$. This random substitution is nothing else than evaluating the expression for random assignments α_p. So the theory of Yashunskii provides for EVERY set of connectors an explicit formula for the probability that a random monotone read-once formula $e_{\mathbb{K}}$ is satisfiable for a random assignment α_p. Especially for $p = \frac{1}{2}$, $P(\frac{1}{2})$ approximately gives the expectation of the ratio of satisfying assignments of a random monotone read-once formula drawn uniformly from all

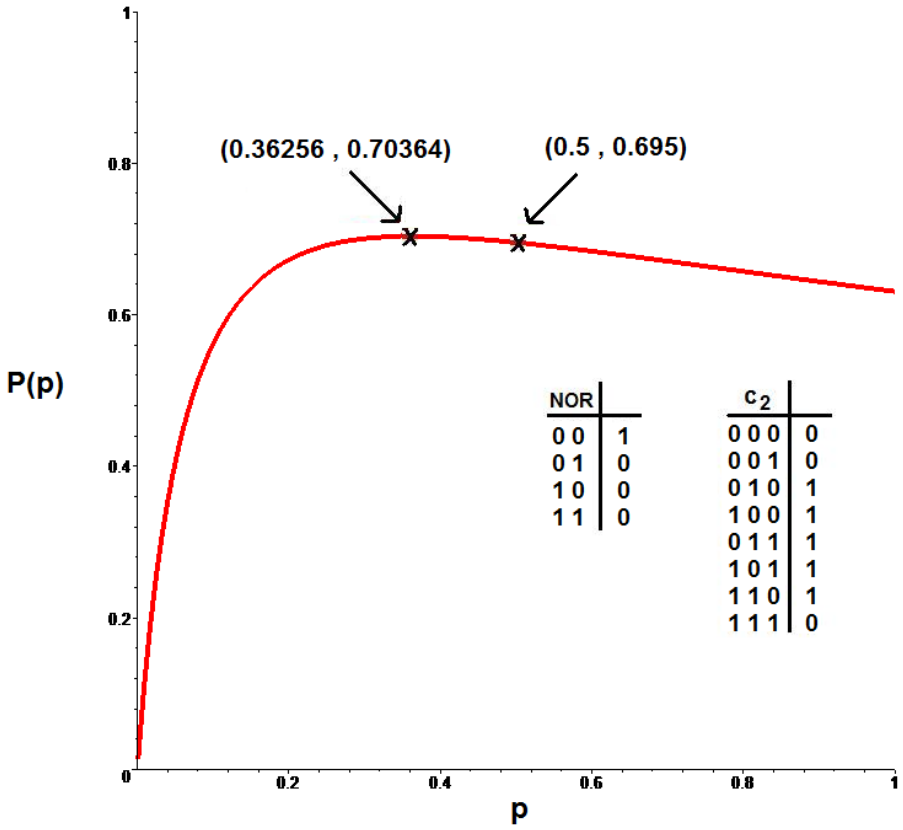

Figure 4.6: Non monotonous function $P(p)$ with two connectors explained by their truth tables in the figure. The maximum of $P(p)$ is at $p^* \approx 0.36256$

monotone read-once formulas with m[41] connectors. Another application is the following: If one wants to know, how to choose $p^* \in (0,1)$ (resp. $[0,1]$) so that a random assignment α_p^* satisfies a random formula $e_{\mathbb{K}}$ with maximal probability, then the answer is to choose

$$p^* = \operatorname*{argmax}_{p\in(0,1)}(P(p)).^{42}$$

[41] Or with $\leq m$ connectors as explained in and below 4.3.

[42] This is true for sufficient large random formulas $e_{\mathbb{K}}$ in the following sense: By private

Consider for example 4.6, with the plot of $P(p)$ for the two connectors NOR and c_2 pictured there. Then, approximately 0.695... of all assignments ($p = \frac{1}{2}$) satisfy a large random expression $e_{\mathbb{K}}$. The best choice for p is 0.36256... for which the probability of getting a satisfying assignment increases to 0.70364.... This is not a dramatic improvement but, as mentioned above, the function $P(p)$ can have arbitrary shape so that there exists a set of connectors $\mathbb{K}$ so that the expectation of $P(p)$ for p, chosen uniformly at random in $[0,1]$, is $< \epsilon$ and $P(p^*) > 1 - \epsilon$ for every $\epsilon > 0$. Based on the author's knowledge similar conclusions as above have not yet been carried out and similar results on the ratio of satisfying assignments of monotone read-once expressions do not exist. For sets of connectors with constant arity the conclusion can even be widened: A k-ary tree[43] with m inner nodes has $(k-1)(m-1)+k$ leaves, so a formula with m k-ary connectors has exactly $(k-1)(m-1)+k$ variables. So for a set of k-ary connectors $\mathbb{K}$ one can expect approximately $P(\frac{1}{2})2^{(k-1)(n-1)+k}$ satisfying assignments for a random monotone read-once formula with m connectors from $\mathbb{K}$. A similar statement clearly also holds for $p \neq \frac{1}{2}$. If one wants a similar statement for arbitrary connectors, not necessarily all with the same arity, one has to use as size function the number of leaves instead of the number of internal nodes, which just changes the generic equations to

$$\begin{aligned} T(z) &= zp + A(T(z), F(z)) \\ F(z) &= z(1-p) + B(T(z) + F(z)) - A(T(z), F(z)). \end{aligned} \tag{4.22}$$

It might turn out that similar results will also hold for this notion of size and it should be possible to use ideas similar to Yashunskii's to prove them. Another possibility would be to use relations between the number of variables (leaves) and internal nodes for simple varieties of trees similar to those derived in [FS09, Proposition VII.2]. Let $\mathbb{E}_{\mathbb{K},m}$ be the set of monotone read-once expressions with connectors from $\mathbb{K}$ of size m. Then, it might also be interesting to investigate the function

$$p \mapsto P_{\alpha_p}([e \text{ is satisfied by } \alpha_p]) =: P_e(p)$$

communication with Yashunskii [Yas15] it was figured out that the named showed in his Ph.D. thesis (in Russian) that $\mathbb{P}_{m,p}(\top)$ converges uniformly to $P(p)$ on any interval $[a,b]$ contained within the interval $(0,1)$. For continuous $P(p)$ on $[0,1]$ it might be possible to show this uniform convergence on the whole interval $[0,1]$.

[43] This is a tree where all inner nodes have in-degree k.

for a single sample expression $e \in \mathbb{E}_{\mathbb{K},m}$; this is the probability that a random assignment satisfies the expression e. Since

$$\mathbb{P}_{m,p}(\top) = \mathbb{E}_e(P_e(p)) = \sum_{e \in \mathbb{E}_{\mathbb{K},m}} \frac{P_e(p)}{|\mathbb{E}_{\mathbb{K},m}|}$$

one could ask if $P_e(p)$ is close to $\mathbb{P}_{m,p}(\top)$ for any expression e or maybe with high probability so that the shape of $P_e(p)$ is close to the shape of $\mathbb{P}_{m,p}(\top)$ for a lot of expressions e. This investigation would give a hint of how much of the macroscopic behavior of the model is found in the microscopic behavior of single expressions.

In the next chapter we are investigating a different branch of research, where the set of expressions is of different kind and the induced probability distribution is of quite different shape.

Chapter 5

Amplified expressions

In this chapter we present another class of models, let us call them amplified models, that, as we will see, are of essentially different shape as the models seen before and as a consequence have quite different properties. At first we will define the model and discuss briefly the structural differences. Then, by means of some simple examples we will point out the different behavior compared to the DLW models. After this short introductory part we will present the literature that is at the origin of the research of such models. After that we will investigate some recent results. These models play an important role in circuit design and Boolean complexity theory, so we will sometimes use the terminology of this field and refer to internal nodes labeled by Boolean connectors as (logical, Boolean) gates.

5.1 Definition and first investigations

In the very beginning let us define the models we are investigating in this chapter: Let μ_0 be a probability distribution on a set of Boolean functions $\mathbb{F} =: \mathbb{H}_0$ and $\mathbb{K}$ a set of k connectors with probability distribution P. Define $\mathbb{H}_m$ as the set of expressions that are built with connectors of $\mathbb{K}$ and expressions from $\mathbb{H}_{m-1}$, so

$$\mathbb{H}_m := \{c(e_1, e_2, ..., e_{\text{arity(c)}}) | c \in \mathbb{K}, e_i \in \mathbb{H}_{m-1} \text{ for } 1 \leq i \leq \text{arity(c)}\}, \text{ m} \geq 1. \tag{5.1}$$

Then define a probability distribution μ_m on the set of expressions $\mathbb{H}_m$ according to the probability distributions μ_{m-1} and P: Let $e \in \mathbb{H}_m$ be the expression $c(e_1, e_2, ..., e_{\text{arity(c)}})$, then define

$$\mu_m(e) := P(c) \cdot \mu_{m-1}(e_1) \cdot ... \cdot \mu_{m-1}(e_{\text{arity(c)}}). \tag{5.2}$$

As usual, this probability distribution on $\mathbb{H}_m$ defines a probability $\mathbb{P}_m$ on Boolean functions. We are again interested in the limit distribution $\mathbb{P}_\infty := \lim_{m\to\infty} \mathbb{P}_m$ if it exists.

The expressions resp. trees in $\mathbb{H}_m$ are balanced, meaning that all leaves ($b \in \mathbb{F}$) have the same depth, i.e. the path from every leaf to the root has the same length, which is exactly m. Such trees have the property that the smallest depth of a leaf is maximal possible (linear in m) in contrast to the expressions we investigated in the previous chapter where the typical path length is $\mathcal{O}(\sqrt{n})$, for n being the number of nodes (cf.[MM78]). Moreover, a balanced binary tree of depth m has 2^m leaves and $2^m - 1$ internal nodes. When there are two or more elements in $\mathbb{H}_0$ or $\mathbb{K}$[1] then the total number of expressions in $\mathbb{H}_m$ is at least 2^{2^m} so in general we can not hope to use analytic combinatorics to analyze such models. Indeed in the literature one tackles such models by directly analyzing the recursive relations coming from Identity 5.1 and Identity 5.2 or by means of a discrete Fourier transform.

Example 5.1
To demonstrate some of the above, let us consider a simple model, namely the model with the 3-ary majority function maj_3 (c.f. Appendix A.1) as single connector ($P(\text{maj}_3) = 1$) and leaves labeled by TRUE with probability p or by FALSE with probability $(1-p)$, so $\mathbb{H}_0 = \{\top, \bot\}$, $\mu_0(\top) = p$ and $\mu_0(\bot) = 1 - p$. So $\mathbb{H}_m$ is the set of all balanced 3-ary trees of depth m with internal nodes labeled by maj_3 and leaves $\top$ or $\bot$. Let X_1 , X_2 and X_3 be three independent Bernoulli variables, each TRUE with probability p_1, p_2 and p_3, respectively. Then, $\text{maj}_3(\text{X}_1, \text{X}_2, \text{X}_3)$ is TRUE with probability $F(p_1, p_2, p_3) = p_1p_2 + p_1p_3 + p_2p_3 - 2p_1p_2p_3$ (cf. [GMS93]). Let us denote $\mathbb{P}_m(\top)$ as a_m and $\mathbb{P}_\infty(\top)$ as a. One computes $a_{m+1} = 3a_m^2 - 2a_m^3$ with

[1]The case where there is only one element in both sets is not interesting.

$a_0 = p$; this is a simple recursion of the form $a_{m+1} = f(a_m)$ with $f(a_m) := F(a_m, a_m, a_m)$. It has 3 fixed points 0, 1/2 and 1, whereas 1/2 is an unstable fixed point the others are stable. So $a = 0$ for $p < 1/2$, $a = 1/2$ for $p = 1/2$ and $a = 1$ for $p > 1/2$. Let $P(p)$ be the limit probability of function TRUE as in the considerations of the model of Yashunskii so that we have

$$P(p) = \begin{cases} 0 & p < 1/2 \\ 1/2 & p = 1/2 \\ 1 & p > 1/2 \end{cases}$$

which is obviously not continuous. △

The discontinuity of $P(p)$ is in sharp contrast to the DLW models for which we proved in Theorem 4.13 that this function is continuous. Regarding the nature of maj_3 this behavior is predictable but as it will turn out such a behavior is typical for this kind of models. The majority function has the property that it amplifies the values to the majority value of the inputs. This amplification of the majority function was the starting point in the investigation of such models, as will be describe later and it is also exploited to exactly identify special classes of read-once formulas [GMS93]. Let us investigate an even more delicate situation.

Example 5.2
We consider again the majority function maj_3, but in contrast to the previous example we define

$$\mathbb{H}_m := \{\mathrm{maj}_3(\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3) | (\mathrm{e_i} \in \mathbb{H}_{\mathrm{m-1}},\ \mathrm{i} = 1, 2, 3) \text{ or } (\mathrm{e}_1 \in \mathbb{H}_{\mathrm{m-1}}, \mathrm{e}_2, \mathrm{e}_3 \in \mathbb{H}_0)\}$$

and

$$\mu_m(e) := \begin{cases} q \cdot \mu_{m-1}(e_1) \cdot \mu_{m-1}(e_2) \cdot \mu_{m-1}(e_3) & \text{if } e_i \in \mathbb{H}_{m-1},\ i = 1, 2, 3 \\ (1-q) \cdot \mu_{m-1}(e_1) \cdot \mu_0(e_2) \cdot \mu_0(e_3) & \text{if } e_1 \in \mathbb{H}_{m-1},\ e_2, e_3 \in \mathbb{H}_0 \end{cases}$$

So we are looking at expressions so that for growing m either all 3 variables of a majority gate in the last layer are substituted by majority trees $\mathrm{maj}_3(\mathrm{e}_1, \mathrm{e}_2, \mathrm{e}_3)$, $e_i \in \mathbb{H}_0$ (with probability q) or only the left one (with probability $1 - q$). This set-up is more general than our definition in 5.1; the

occurring trees are not necessarily fully balanced but they have the property that the subtrees of internal nodes of the same layer have the same depth. For $q = 1$ the previous example is retrieved. For $q = 0$ every majority node except for the last has exactly one majority gate as child. In this case we get the linear recursion

$$a_{m+1} = F(a_m, p, p) = 2a_m p + p^2 - 2a_m p^2 = a_m 2(p - p^2) + p^2$$

for the probability of TRUE. Solving this recursion and let m tend to infinity gives the solution $P(p) = \frac{p^2}{1-2(p-p^2)}$. Not surprisingly this is a continuous function of p. So $P(p)$ is continuous for $q = 0$ (1 majority node as child) and not continuous for $q = 1$ (3 majority nodes as children). Is there any $0 < q < 1$ so that $P(p)$ is not continuous? The answer is yes; $P(p)$ is infinitely often differentiable for $q \leq 0.5$, for $q = 0.5$ it is continuous but has unbounded derivative at $p = 0.5$ and for $q \geq 0.5$ it is not continuous. This result can be obtained by analyzing the recursion for a_m:

$$\begin{aligned} a_{m+1} &= (1-q)F(a_m, p, p) + qF(a_m, a_m, a_m) \\ &= (1-q)(a_m 2(p - p^2) + p^2) + q(3a_m^2 - 2a_m^3) =: f(a_m, p, q), \\ a_0 &= p. \end{aligned}$$

In 5.1 $P(p)$ is plotted for different models and parameters. The yellow curve belongs to the model with one or tree majority gates as children and $q = 0.6$. The function has a single discontinuity at $p = 0.5$ where it has of course the value 0.5. The limit from left (and right) $\lim_{p^- \to 0.5} P(p)$ exists and can be evaluated as the smallest solution (fixed point) of the cubic equation $a = f(a, 0.5, 0.6)$ which gives 0.21132..., and analogous for other values of $q > 0.5$. The orange curve is for $q = 0.5$ and gives a clue that for this parameter $P(p)$ is not differentiable at $p = 0.5$. The green curve belongs to $q = 0.4$ and the black one to $q = 0$. The blue curve is $P(p)$ for the model of Yashunskii and the red one finally is $F(p, p, p)$ which is clearly closest to the identity function $p \mapsto p$ because it has the least amplification incorporated, namely none.

As in the model of Yashunskii one can relate $P(p)$ to the ratio of satisfying assignment of a monotone read-once function with majority operations. But one has to take care with the conclusions that can be drawn because opposed to the model of Yashunskii, $\mathbb{P}_m(\top)$ definitely does not converge uniformly to

$\mathbb{P}_\infty(\top)$. For example if $q = 0.6$ it is not valid that a random assignment α_p, with p the probability for a 1 and $p < 0.5$, satisfies a large random monotone read-once functions from $\hat{\mathbb{H}}_m$ (This are all trees from $\mathbb{H}_m$ with distinct variables as leaves) with probability smaller than $0.21132 + \epsilon$. This only holds for all p smaller than a fixed value $c < 0.5$. Of course there also exists a sequence $(p_m)_{m\in\mathbb{N}}$ converging (maybe very slow) to 0.5 from below so that α_{p_m} satisfies a random monotone read-once function from $\hat{\mathbb{H}}_m$ with probability converging to 0.21132 for m tending to infinity; but there also exists such a sequence $(p_m)_{m\in\mathbb{N}}$ (converging sufficient fast to 0.5) for which this probability converges to 0.5. $\triangle$

The above shows that amplification models are of quite different shape than DLW models. The interest for them also comes from a different direction as we will see next.

5.2 Amplification models and origin of their study

The first who used ideas of amplifying was Shannon in [MS56]. He amplified the error probability of unreliable components to 0 by using more unreliable components in contact networks, whereas he did not investigated any kind of growth processes as above. The first who investigated such a model was Valiant. He investigated the model with 3-ary majority gates as above to obtain an upper bound on the size of monotone formulas computing the majority function [Val84]. A monotone formula in this context is a formula with $\wedge$ and $\vee$ gates and positive literals. Because of its importance in the literature and because it was one of the first times random Boolean expressions occurred in the literature we will sketch the approach. Moreover, it shows an interesting way of using random Boolean expressions as a tool.

Example 5.3 (cf. [Val84])
Let $\mathrm{maj_n}(\mathrm{x_1}, ..., \mathrm{x_n})$ be the majority function on n variables. For convenience assume that n is odd, so $\mathrm{maj_n}(\mathrm{x_1}, ..., \mathrm{x_n})$ is TRUE iff at least $(n+1)/2$ inputs are TRUE. Consider the full 3-ary tree of depth $m := c \log n$ with $\mathrm{maj_3}$

gates and label the leaves by the (positive) variables $\{x_1, ..., x_n\}$ u.a.r. with replacement. So we are actually considering $\mathbb{H}_m$ with $\mathbb{H}_0 := \{x_1, ..., x_n\}$ and $\mu_0(x_i) = \frac{1}{n}$ and $\mathbb{K} := \{\mathrm{maj}_3\}$. Such a random expression $t \in \mathbb{H}_m$ has depth $\mathcal{O}(\log n)$ and polynomial tree size complexity. We now claim that for $m = c \log n$ and for fixed input $x = (x_1, ..., x_n)$ the probability that a random formula t computes the majority function on x is larger than $1 - 2^{-n}$, so $\mathrm{Prob}([\mathrm{t(x)} = \mathrm{maj_n(x)}]) > 1 - 2^{-\mathrm{n}}$. This guarantees the existence of a formula of logarithmic depth and polynomial size computing the majority function $\mathrm{maj_n}$ because $\mathrm{Prob}([\mathrm{t} \neq \mathrm{maj_n}]) < 1$:

$$\begin{aligned}\mathrm{Prob}([\mathrm{t} \neq \mathrm{maj_n}]) &= \mathrm{Prob}([\exists \mathrm{x} : \mathrm{t(x)} \neq \mathrm{maj_n(x)}]) = \mathrm{Prob}(\bigcup_{\mathrm{x}} [\mathrm{t(x)} \neq \mathrm{maj_n(x)}]) \\ &\leq \sum_x \mathrm{Prob}([\mathrm{t(x)} \neq \mathrm{maj_n(x)}]) < 2^n \cdot 2^{-n} = 1.\end{aligned}$$

So let us assume that $\mathrm{maj_n(x)} = 1$ (the other case works similarly), so the ratio p_0 of 1's is at least $\frac{1}{2} + \frac{1}{2n}$. The inputs of the random tree are chosen u.a.r from $x_1, ..., x_n$ meaning that they are Bernoulli distributed with parameter p_0. From the examples above we know that probabilities larger than $\frac{1}{2}$ are amplified towards 1. Analyzing again the recursion $p_{m+1} = F(p_m, p_m, p_m),\ p_0 \geq \frac{1}{2} + \frac{1}{2n}$ shows that this amplification is fast enough to derive a constant c with $p_{c \log n} \geq 1 - 2^{-n}$. So there is a formula t with depth $c \log n$ and polynomial size with only $\mathrm{maj_3}$ gates computing the function $\mathrm{maj_n}$. Since a $\mathrm{maj_3}$ gate can be realized by a formula e with $\wedge$ or $\vee$ gates of constant size (4) we can substitute every $\mathrm{maj_3}$ gate by e to obtain a monotone formula of logarithmic depth and polynomial size computing the majority function. $\triangle$

The example above is of probabilistic nature and only proves the existence of such a formula. The bound of Valiant is $\mathcal{O}(n^{5.3})$ for the size. An improvement of this ($\mathcal{O}(n^{4.9})$) is established among some more general results in [GM97]. A partial derandomization of the construction above can be found in [HMP06] with a significant improvement on size ($\mathcal{O}(n^2)$). The only deterministic construction known for a monotone formula of polynomial size for $\mathrm{maj_n}$ is found in [AKS83], unfortunately the polynomial is of very high order.

Boppana, who coined the term 'amplification method', further investigated this amplification process on a more general level in [Bop85]. He showed that the result of Valiant is in a sense optimal and showed a similar result for general threshold function $tr(k, n)$. Savický was not interested in constructing

small formulas for given functions but he investigated the underlying construction resp. growth process per se [Sav90]. He investigated the model with μ_0 the uniform distribution on $\mathbb{H}_0 := \{\top, \bot, x_1, ..., x_n, \overline{x_1}, ..., \overline{x_n}\}$ with a single connector α that is balanced and nonlinear. Balanced means that α is TRUE for exactly half of the inputs. It is nonlinear if $\alpha(x) \neq c_0 \oplus c_1 x_1 \oplus ... \oplus c_n x_n$ any $c_0, ..., c_n \in \{0, 1\}$. He showed the following:

Theorem 5.4 ([Sav90, Theorem 5.4])
If $n \geq 2$ then the following two statements are equivalent:

(i) α is balanced and nonlinear

(ii) for all $f \in \mathbb{B}_n$ the following holds:

$$\mathbb{P}_\infty(f) = (\frac{1}{2})^{2^n}.$$

This means that in the models with balanced and nonlinear connector every function f is computed by the same probability for large depth m. This is shown by means of a discrete Fourier transform technique of the probability function. Further investigation of the approach would go too far but the interested reader can take [dW08] as a starting point from which one can delve further into the method of discrete Fourier transform on the Boolean cube. With the same technique Savický generalized the result above to more general leaf sets $\mathbb{H}_0$ and distributions μ_0 satisfying a condition of 'unbiasness' [Sav95]. Moreover, he proved a statement about the convergence speed of $\mathbb{P}_m(f)$ to $(\frac{1}{2})^{2^n}$ for fixed f, which says roughly speaking that for a function f that is 'very' balanced the convergence is fast. Brodsky and Pippenger [BP05] systematically investigated the growth process when $\mathbb{H}_0$ is a subset of $\{\top, \bot, x_1, ..., x_n, \overline{x_1}, ..., \overline{x_n}\}$ and μ_0 is the uniform distribution on this set with a single connector α. Several restrictions on α so that the limit distribution is the uniform distribution on distinguished sets of Boolean functions (e.g. linear function, monotone functions, self-dual functions and threshold functions etc.) were considered reaching from α being linear, self-dual nonlinear or monotone. The methods of choice for proving the results are once again the Fourier transform as well as the direct study of recurrence relations.

All these results have one in common: In every case the limit distribution is the uniform distribution on either a (small) set of Boolean functions, all

Boolean functions or a single Boolean function. We will observe this regime also in the models we are considering in the next section that will have $\wedge$ and $\vee$ as allowed connectors.

5.3 $(\wedge, \vee)$-Amplification models

The first paper which investigates amplification models with more than one connector is [PW06] where the authors consider full binary trees of depth m labeled by $\wedge$ and $\vee$ both with probability 1/2 and leaves labeled by 2^m distinct variables x_1,...,x_{2^m} from left to right. Let us denote this random tree by t_m. So the randomness of a tree comes from the random labeling of the internal nodes and not from the (fixed) labeling of the leaves. They are then interested in the random variable X_m defined as the mean output of the random formula:

$$X_m := \frac{1}{2^{2^m}} \sum_{x \in \{0,1\}^{2^m}} t_m(x).$$

By considering the random tree t_m as a rooted subtree of the random infinite binary tree t_∞ whose internal nodes are labeled by $\wedge$ and $\vee$ u.a.r. all variables X_m are defined on the same probability space for $m \geq 0$.[2] With this set-up we are ready to state the main result of there investigations:

Theorem 5.5 ([PW06, Theorem 2.1.])
X_m converges almost surely to X with $\mathrm{Prob}([\mathrm{X} = 0]) = \mathrm{Prob}([\mathrm{X} = 0]) = 1/2$.
Moreover, the first two moments of X_m are

$$\begin{aligned} E(X_m) &= \tfrac{1}{2} \\ E(X_m^2) &= \tfrac{1}{2} - \tfrac{1}{m} + \mathcal{O}(\tfrac{\log m}{m^2}). \end{aligned}$$

Proof. (Sketch)
The first step in the proof is to establish that the sequence $\{X_m\}$ is a martingale with respect to the filtration $\{\mathcal{F}_m\} := \sigma(G(v) : v \text{ is an internal node of } t_m)$. Readers who are not familiar with these basic concepts of probability theory

[2] To be exact the random tree t_m can be seen as a vector drawn u.a.r. from $\{0,1\}^{2^{m-1}}$, whereas a 0 (resp. a 1) as i-th entry means that the i-th node (in level order) of t_m is labeled by $\wedge$ (resp. $\vee$). So it is a vector of independent and identically $0-1$-random variables $(G(i))_{i=1,\ldots,2^{m-1}}$. An infinite tree is a(n) (infinite) sequence of such random variables $(G(i))_{i \in \mathbb{N}}$ and the tree t_m is the vector of the first 2^{m-1} entries.

can find the necessary background in [Dur05]; but we will not further need them except for this proof. Martingales $\{\hat{X}_m\}$ with bounded expectation of the absolute value (i.e. $E(|\hat{X}_m|) \leq c \in \mathbb{R}$) converge almost surely (Martingale Convergence Theorem). Because $0 < X_m < 1$, this requirement is satisfied so that X_m converges almost surely. That $E(X_m) = \frac{1}{2}$ can be seen from symmetry arguments or direct from the recursion for the expectation

$$E(X_{m+1}) = \frac{1}{2}(E(X_m))^2 + \frac{1}{2}(1 - (1 - E(X_m))^2) = E(X_m). \tag{5.3}$$

Let X_m^l and X_m^r be the mean outputs of the right resp. left subtree of the random tree t_m. Then, X_{m+1} is $X_m^l \cdot X_m^r$ when the root is labeled by $\wedge$ or $1 - (1 - X_m^l) \cdot (1 - X_m^r)$ when the root is labeled by $\vee$. So one establishes the recursive description of the second moment X_m^2

$$E(X_{m+1}^2) = \frac{1}{2}E((X_m^l \cdot X_m^r)^2) + \frac{1}{2}E((1 - (1 - X_m^l) \cdot (1 - X_m^r))^2).$$

X_m^l and X_m^r are independent and X_m^l, X_m^r, $1-X_m^l$, $1-X_m^r$ are all distributed as X_m, so the recursion simplifies to

$$E(X_{m+1}^2) = E(X_m^2)^2 + \frac{1}{4}.$$

This equation has a single attracting fixed point $\frac{1}{2}$. Further investigation of the equation with help of the transformations $Z_m := X_m(1 - X_m)$ and $Z_m \mapsto \frac{1}{Z_m}$ leads to the stated rate of convergence of $E(X_m^2)$ to $\frac{1}{2}$. Finally, observe that $E(X^2) = \lim_{m\to\infty} E(X_m^2) = 1/2$, where the first identity follows from the Bounded Convergence Theorem. The random variable X is symmetric around $1/2$ so $E(X^2) = 1/2$ is only possible for Prob([X = 0]) = Prob([X = 1]) = 1/2. □

Remark. *The limit of X_m does not necessarily exist. Indeed one can consider an infinite tree consisting of layers of sufficiently large levels labeled only by $\wedge$ resp. $\vee$ so that X_m is amplified toward 0 and 1 alternately.*

The authors of [PW06] offer the following intuitive explanation for the convergence of X_m to 0 and 1: There is either a slight predominance of nodes with $\wedge$ that forces X_m to 0 or a predominance of $\vee$ pushing X_m to 1. That sounds plausible but it is a folklore fact from 1-dimensional random walks

that for a(n) (unbiased) random $0-1$-sequence ($\hat{=}\ \wedge - \vee$-sequence) there are infinitely many numbers k so that there are k 0's and k 1's among the first $2k$ entries of the sequence with probability 1. It should be possible to follow that the ratio of nodes of a random balanced tree of depth m that are labeled by $\wedge$ is for infinitely many m greater than $1/2$. So the suggestive explanation should be refined.

Similarly as above one can also derive recurrences for higher-order moments of X_m to establish their asymptotic rates of convergence. In the paper this is done up to the fourth moment and there is also an investigation of the random variable $Z_m = X_m(1 - X_m)$ with tail estimates. In addition to these theoretical considerations they also tried to obtain a better understanding of X_m by means of experiments. The problem with experiments is that for even small values of m, say $m = 20$, it is computationally intractable to compute X_m for all possible trees, and for $m \geq 40$ it is computationally impossible to compute even one sample of X_m. So they introduced the notion of sensitivity of a leaf of an And/Or tree to quantify the potential effect of the value of this node to the value of X_m as the difference of the mean output when substituting the leaf by 1 or 0 respectively. The idea then is to approximate the value of samples of X_m by considering the value of X_m for a growing tree starting with a single node. Then, in every step the leaf with the highest sensitivity is chosen to be revealed to compute a better approximation of X_m for this tree until the approximation of the sample is 'good' enough. Their goal was to show that the number of nodes that have to be revealed to get a good approximation of one sample is far smaller than evaluating all nodes from the full balanced tree. This would disclose a lot about the distribution of X_m but unfortunately they are not able to rigorously prove such a result.

Before we investigate another model, let us draw similar conclusions as we did for the models of Yashunskii. We can interpret X_m as the mean output of a balanced monotone read-once formula of depth m with connectors $\wedge$ and $\vee$ that is drawn u.a.r. from all such expressions. Putting it in this way the result above informally gives that such a monotone read-once expression (of great depth) has either lots of satisfying assignments or very few. The case that such a formula has an intermediate number of satisfying assignments becomes very rare for increasing depth. The next result quantifies this vague formulation with the help of the results for X_m.

Proposition 5.6
For every $\epsilon > 0$ there exists an $M(\epsilon)$ such that for every $m > M(\epsilon)$ the following holds: Define $c := \sqrt{\frac{m}{1+\epsilon}}$. Take a random monotone read-once formula obtained from labeling the internal nodes of the full binary tree of depth m with $\wedge$ and $\vee$ independently with probability $1/2$. This formula has at most $(\frac{1}{c}) \cdot 2^{2^m}$ satisfying assignments with probability at least $\frac{1}{2} - \frac{1}{2c}$ and it has at least $(1-\frac{1}{c}) \cdot 2^{2^m}$ satisfying assignments with probability at least $\frac{1}{2} - \frac{1}{2c}$.

Proof. For fixed $\epsilon > 0$ and $c := \sqrt{\frac{m}{1+\epsilon}}$ assume that there are arbitrarily large m with $P(\frac{1}{c} \leq X_m \leq 1 - \frac{1}{c}) \geq \frac{1}{c}$. For such m we can upper bound $E(X_m^2)$: Let $a_i, i = 1, ..., s$ denote the finitely many values X_m can take and define $P(a_i) := P(X_m = a_i)$ and

$$\tilde{P}(a_i) := \begin{cases} P(a_i) & \text{if } a_i \neq \frac{1}{2} \\ \frac{P(a_i)}{2} & \text{if } a_i = \frac{1}{2} \end{cases}$$

to compute

$$\begin{aligned} E(X_m^2) &= \sum_{i:a_i < \frac{1}{c}} P(a_i) \cdot a_i^2 + \sum_{i:\frac{1}{c} \leq a_i \leq 1-\frac{1}{c}} P(a_i) \cdot a_i^2 + \sum_{i:\frac{1}{c} < a_i \leq 1} P(a_i) \cdot a_i^2 \\ &= \sum_{i:a_i < \frac{1}{c}} P(a_i) \cdot (a_i^2 + (1-a_i)^2) + \sum_{i:\frac{1}{c} \leq a_i < \frac{1}{2}} P(a_i) \cdot (a_i^2 + (1-a_i)^2) \\ &\quad + \sum_{i:a_i = \frac{1}{2}} \tfrac{P(a_i)}{2} \cdot (a_i^2 + (1-a_i)^2) \\ &= \sum_{i:a_i < \frac{1}{c}} P(a_i) \cdot (a_i^2 + (1-a_i)^2) + \sum_{i:\frac{1}{c} \leq a_i \leq \frac{1}{2}} \tilde{P}(a_i) \cdot (a_i^2 + (1-a_i)^2) \end{aligned}$$

The second equality holds due to the symmetry of X_m. The function $a \mapsto a^2 + (1-a)^2$ is monotonically decreasing in $[0, 1/2]$ so we get the upper bound

$$\begin{aligned} E(X_m^2) &\leq \sum_{i:a_i < \frac{1}{c}} P(a_i) \cdot 1^2 + \sum_{i:\frac{1}{c} \leq a_i \leq \frac{1}{2}} \tilde{P}(a_i) \cdot ((\tfrac{1}{c})^2 + (1-\tfrac{1}{c})^2) \\ &= P(0 < X_m < \tfrac{1}{c}) + ((\tfrac{1}{c})^2 + (1-\tfrac{1}{c})^2) P(\tfrac{1}{c} \leq X_m \leq \tfrac{1}{2}) \\ &\leq \max_{\frac{1}{c} \leq p \leq 1} \{\tfrac{1-p}{2} + ((\tfrac{1}{c})^2 + (1-\tfrac{1}{c})^2)\tfrac{p}{2}\} \\ &= \tfrac{1-\frac{1}{c}}{2} + ((\tfrac{1}{c})^2 + (1-\tfrac{1}{c})^2)\tfrac{1}{2c} \\ &= ... = \tfrac{1}{2} - \tfrac{1+\epsilon}{m} + \tfrac{(1+\epsilon)^{\frac{3}{2}}}{m^{\frac{3}{2}}}. \end{aligned}$$

From 5.5 we know that

$$E(X_m^2) = \frac{1}{2} - \frac{1}{m} + \mathcal{O}(\frac{\log m}{m^2}) \geq \frac{1}{2} - \frac{1}{m} - d\frac{\log m}{m^2}$$

for a constant d sufficient large. So we derived that there exist arbitrarily large m with

$$\frac{1}{2} - \frac{1}{m} - d\frac{\log m}{m^2} \leq \frac{1}{2} - \frac{1+\epsilon}{m} + \frac{(1+\epsilon)^{\frac{3}{2}}}{m^{\frac{3}{2}}},$$

a contradiction.

So there exists an $M(\epsilon)$ such that $P(\frac{1}{c} \leq X_m \leq 1 - \frac{1}{c}) \leq \frac{1}{c}$ for $m \geq M(\epsilon)$ and $c = \sqrt{\frac{m}{1+\epsilon}}$. So $P(X_m < \frac{1}{c}) \geq \frac{1}{2} - \frac{1}{2c}$ for $m \geq M(\epsilon)$ which means that $2^{2^m} X_m < (\frac{1}{c}) \cdot 2^{2^m}$ with probability at least $\frac{1}{2} - \frac{1}{2c}$. Since $2^{2^m} X_m$ is the number of satisfying assignments the first part of the assertion follows and the second part follows in the same way. □

Remark. *This shows that the probability that such a monotone read-once function f has an intermediate number A_f of satisfying assignments, i.e. $\frac{1}{c} \cdot 2^{2^m} \leq A_f \leq (1 - \frac{1}{c}) \cdot 2^{2^m}$, is at most $\frac{1}{c}$, which converges to 0 for m tending to infinity. So $P(\frac{1}{c} \cdot 2^{2^m} \leq A_f \leq (1 - \frac{1}{c}) \cdot 2^{2^m}) \leq \frac{1}{c} \to 0$*

Let us now finish this section with the discussion of a model that is the first and to the author's knowledge only amplification model in its full generality as defined at the beginning of the chapter that is investigated in the literature, namely the model of balanced And/Or trees in [FGG09]. In one sentence one looks at balanced trees of depth m with inner nodes independently labeled by $\wedge$ with probability p and with $\vee$ with probability $1-p$. The leaves are then labeled by variables $\{x_1, ..., x_n\}$ according to a distribution. In the notation from above this means $\mathbb{H}_0 = \{x_1, ..., x_n\}$ with μ_0 a probability distribution on $\mathbb{H}_0$ and $\mathbb{K} = \{\wedge, \vee\}$ with $P(\wedge) = p$ and $P(\vee) = 1 - p$. Before we can state the main result of [FGG09] we need some definitions. The weight of an assignment $\alpha \in \{0,1\}^n$ according to μ_0 is defined as

$$w(\alpha) := \sum_{i=1}^{n} \mu_0(x_i)\alpha_i.$$

Let $\beta = (\beta_1, ..., \beta_n)$ be a vector in $\mathbb{R}^n$ and $\theta \in \mathbb{R}$ and define the linear threshold function $T_{\beta,\theta}$ pointwise as

$$T_{\beta,\theta}(\alpha) = 1 \text{ iff } \beta_1\alpha_1 + ... + \beta_n\alpha_n \geq \theta.$$

The main results now read as follows:

Theorem 5.7 ([FGG09, Theorem 2.1.])
Let μ_0 be a positive probability distribution on $\{x_1, ..., x_n\}$ and $p \in [0,1]$, then the limit distribution $\mathbb{P}_\infty$ exists and is fully describable as:

- *If $P(\wedge) = p > 1/2$, then the support of $\mathbb{P}_\infty$ reduces to the single function $x_1 \wedge x_2 \wedge ... \wedge x_n$.*
- *If $P(\wedge) = p < 1/2$, then the support of $\mathbb{P}_\infty$ reduces to the single function $x_1 \vee x_2 \vee ... \vee x_n$.*
- *If $P(\wedge) = 1/2$, then the limit distribution $\mathbb{P}_\infty$ concentrates on the linear threshold functions: Let $\theta_0 = 0 < \theta_1 < ... < \theta_s = 1$ be the different weights of all assignments in $\{0,1\}^n$, then for $i \in \{1, ..., s\}$ we have $\mathbb{P}_\infty(T_{\mu_0,\theta_i}) = \theta_i - \theta_{i-1}$.*

Proof. (Sketch) The proof of this theorem in [FGG09] is divided into a couple of lemmas. We will follow mainly this presentation and sketch the most important steps. Observe that due to the duality of the problem for p and $1-p$ we can assume that $p \geq 1/2$.

Let $p \in [0,1]$ and fix an $\alpha \in \{0,1\}^n$. Define u_m as the probability that α satisfies a random function f, that is distributed according to $\mathbb{P}_m$:

$$u_m := \mathbb{P}_m([f(\alpha) = 1]).$$

P_0 equals μ_0 so $u_0 = \sum_{i=1}^n \mu_0(x_i)\alpha_i = w(\alpha)$. By considering the label of the root of an expression from $\mathbb{H}_{m+1}$ we get the recursive description

$$u_{m+1} = (2p-1)u_m^2 + 2(1-p)u_m.$$

0 is an attractive fixed point of this recursion for $p > 1/2$, so for $\alpha \neq (1, ..., 1)$ u_m converges to 0. For $p = 1/2$ it is easily seen that u_m is constant, so $u_m = w(\alpha)$.

This already allows to prove the case for $p \neq 1/2$: Assume that there is a function f different from $x_1 \wedge ... \wedge x_n$ with $\mathbb{P}_m(f) \not\to 0$. This function has a satisfying assignment $\alpha \neq (1, ..., 1)$ for which $\mathbb{P}_m([f(\alpha) = 1]) \not\to 0$ which is a contradiction. So $\mathbb{P}_m(x_1 \wedge ... \wedge x_n) \to 1$.

To prove the case $p = 1/2$ similar ideas apply: Let α, β be two different assignments with $w(\alpha) \leq w(\beta)$. Define $v_m := \mathbb{P}_m([f(\alpha) = 1 \text{ and } f(\beta) = 0])$. With the analogue quantities $v_m^{(a,b)}\mathbb{P}_m([f(\alpha) = a \text{ and } f(\beta) = b])$, for $a, b \in \{0, 1\}$ one can derive[3] a system of recursions of dimension 4 for the four sequences, from which it can be followed that $v_m \to 0$. So consequently a function f with $\mathbb{P}_m \not\to 0$ satisfies that for all assignments α, β with $w(\alpha) \leq w(\beta)$ it holds that $f(\alpha) \leq f(\beta)$. Functions fulfilling this condition are exactly the linear threshold functions of the shape $T_{\mu_0,\theta}$. Let $\theta_0 = 0 < \theta_1 < ... < \theta_s = 1$ be all different weights of assignments and α_i, $i = 1, ..., s$ corresponding assignments. From above we know that $\mathbb{P}_m([f(\alpha_i) = 1]) = w(\alpha_i) = \theta_i$. Exactly the functions T_{μ_0,θ_0},...,T_{μ_0,θ_i} are 1 on α_i so we have $\mathbb{P}_m(\{T_{\mu_0,\theta_0}, ..., T_{\mu_0,\theta_i}\}) \to \theta_i$ and by induction one follows that $\mathbb{P}_m(T_{\mu_0,\theta_i}) = \theta_i - \theta_{i-1}$. Which completes the proof. □

Direct consequences of the theorem above are [FGG09, Corollary 2.1. and Corollary 2.2]:

(i) If $p = 1/2$ and μ_0 is the uniform distribution on $\{x_1, ..., x_n\}$, then the limit distribution is uniform on the n threshold functions $x_1 + x_2 + ... + x_n \geq i$ for $i = 1, ..., n$.

(ii) If $p = 1/2$ and μ_0 is the uniform distribution on $\{x_1, \overline{x_1}, ..., x_n, \overline{x_n}\}$, then the limit distribution 'collapses' to the constant function TRUE and FALSE both with probability 1/2.

The result with negative literals is derived by considering (i) with $2n$ variables $\{x_1, y_1..., x_n, y_n\}$ and substituting y_i by $\overline{x_i}$. In the same manner one can derive the limit distribution for arbitrary functions as leaf sets. As in the other amplification models the distribution concentrates on a small number of functions and is highly discontinuous in the parameter p for the probability of the connectors but it is continuous in the parameters of the distribution on the variables (or more general functions) μ_0.

The authors are also interested in the speed of convergence by investigating the quantity $\|\mathbb{P}_m - \mathbb{P}_\infty\|_\infty := \max_{f \in \mathbb{B}_n} |\mathbb{P}_m(f) - \mathbb{P}_\infty(f)|$. They are able to evaluate the speed of convergence for fixed number of variables n and fixed μ_0 and p. For example by a careful analysis of the recurrence relations they found:

[3]This works as above by distinguishing the label of the root.

Proposition 5.8 ([FGG09, Proposition 4.2.])
For $p = 1/2$, the convergence speed of $(\mathbb{P}_m)$ is

$$\|\mathbb{P}_m - \mathbb{P}_\infty\|_\infty = 2^{-\Theta(m)}$$

if all assignments in $\{0,1\}^n$ have distinct weights; otherwise it is

$$\|\mathbb{P}_m - \mathbb{P}_\infty\|_\infty = \Theta(1/m).$$

They also note that it would be interesting to investigate the rate of convergence with respect to the number of variables m, but they do not expect it to be fast. A sufficiently fast convergence would lead to monotone polynomial size formulas for linear threshold functions which they are not known to have.

Since we have the limit distribution at hand in all cases, it is easy to derive statements regarding the number of satisfying assignments of expressions drawn according to the considered distribution. For example:

(i) If $p = 1/2$ with $\mathbb{H}_0 = \{x_1, ..., x_n\}$. Assume that there exists an x_i with $\mu_0(x_i) = a$, then an expression drawn from $\mathbb{H}_m$ according to μ_m has more than $\frac{1}{2}2^{2^m}$ satisfying assignments with probability converging to a for m tending to infinity.

Or with the help of the convergence rate in the case of $1/2 < p < 1$, which is $2^{-\Theta(m)}$:

(ii) If $1/2 < p < 1$, then the probability that an expression drawn from $\mathbb{H}_m$ according to μ_m has more than 1 satisfying assignment[4] is exponentially decreasing: $\mathbb{P}_m(\{f|\exists\alpha \neq \beta, f(\alpha) = f(\beta) = 1]) = 2^{-\Theta(m)}$. Informally speaking such a formula can only be satisfied if all constraints (x_i) are fulfilled.

(ii) is in sharp contrast with the related DLW model where for every p this probability $\mathbb{P}_m(\{f|\exists\alpha \neq \beta, f(\alpha) = f(\beta) = 1])$ converges to a positive constant.

[4]The assignment (1,...,1) is always a satisfying assignment for the formulas considered here.

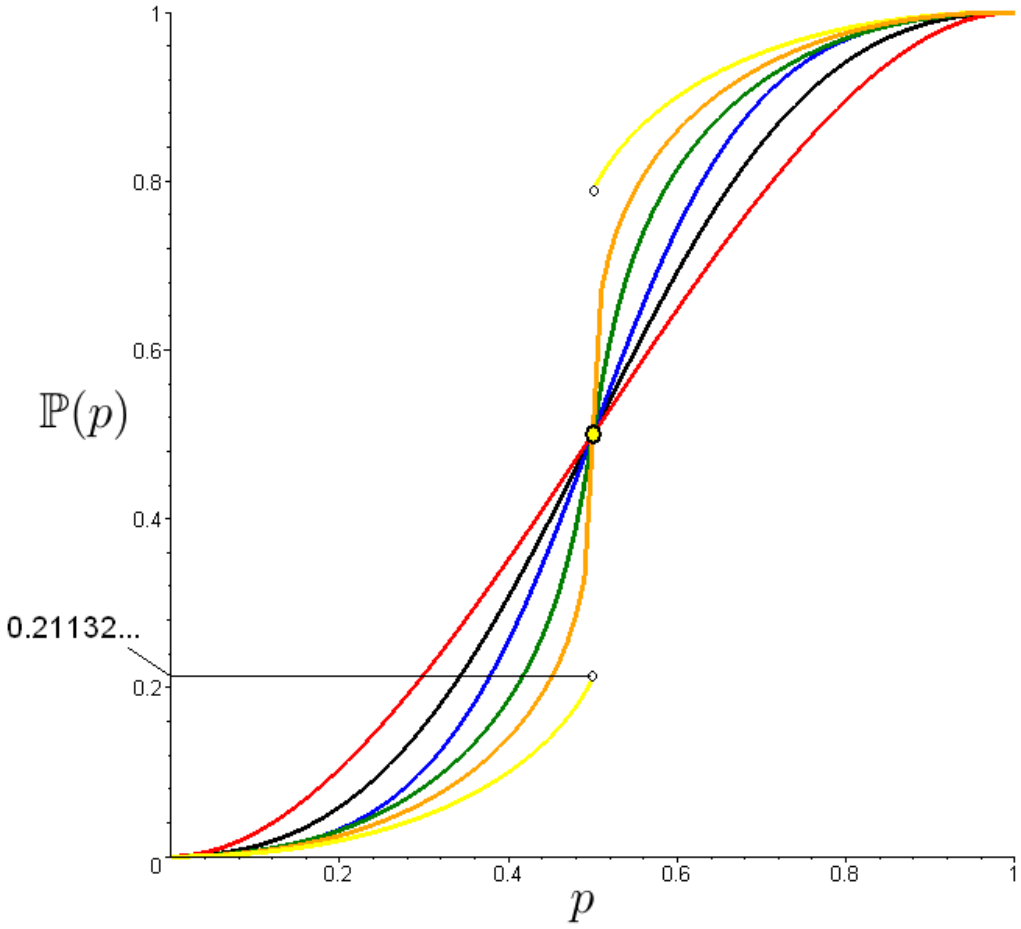

Figure 5.1: $P(p)$ for different models: yellow: $q = 0.6$, orange: $q = 0.5$, green: $q = 0.4$, black: $q = 0$, blue: corresponding model of Yashunskii, red: $F(p, p, p)$. The function $P(p)$ is approximated on the uniform grid with mesh-width 1/1000 where for each point 1000 iterations are performed (the blue and the red curves are exact).

Chapter 6

Other models, Conclusion and Open Problems

In the last part of the thesis we want to present in a nutshell several models that have not found their way into the main part of this work although they perfectly fit to the concept. After that we will very briefly present some papers that investigate problems of quite different nature (nonetheless there are connections). We will do so in an enumeration fashion and will not provide the necessary background and all definitions that would be necessary to fully capture the essence of the models. Nevertheless, these models are broached for sake of taking a view on what else can be done. At the end we will conclude the thesis and state open problems.

6.1 Other models

The first model that perfectly fits is the model of so called decorated critical Galton-Watson trees. Such a random tree is (binary case) constructed as follows: Start with a single node. Every leaf in the tree becomes with probability 1/2 an internal node with two children and with probability 1/2 the branching process stops at a leaf. Such a random tree is almost surely finite [AN72]. As usual, a random Boolean expression is then obtained by labeling the internal nodes by connectives and the leaves by literals (resp. (positive) variables). There is a strong relation between these models and the related DLW models, see [CFGG04, Proposition 1] and [Gar06]; that is why it was investigated almost parallel to the DLW models, see

[CFGG04, GW05, Gar06, FGGG12, GG10]. This relation also gives that our result 4.13 holds for these models when the conditions of the statement hold for the corresponding DLW model[1]. Apposed to the DLW models there is no existence question of the distribution, so one can investigate arbitrary connectors and base functions. The most important result is that for arbitrary binary connectors and variables $\{x_1, ..., x_n\}$ the probability of all read-once functions tends to 1 when the number n of variables tends to infinity [GG10, Theorem 10]. So this system does not exhibit the Shannon effect. The same statement holds for arbitrary leaf set $\mathbb{F}$ (especially for $\{x_1, \overline{x_1}, ..., x_n, \overline{x_n}\}$) when one generalizes the notion of read-once function so that a function f is called read-once when there is a tree with distinct leaves expressing f.

The second model uses random growing trees that are defined as follows [CGM11]: A growing tree of size 0 is a single leaf with probability 1, so $\mathbb{H}_0 = \{\bullet\}$ and $\mu_0(\bullet) = 1$. A growing tree of size m is a growing tree of size $m-1$, where a leaf is chosen uniformly at random to transform to an internal node with two children (the authors call this process 'sprouting'). Or the other way round: If one has a growing tree of size m and cuts uniformly at random two leaves having the same parent then one obtains a growing tree of size $m-1$. Such a tree is then labeled as usual. The relation to the amplification model with balanced And/Or trees is apparent: In the amplification model all leaves of the (in this case single) growing tree sprout, whereas in the growing tree model only one randomly picked leaf sprouts. [CGM11] investigates this model for several set-ups namely for connectors $\wedge$ and $\vee$ with a Bernoulli distribution with parameter p and a distribution μ_0 on either only positive literals or all literals[2] and the model with the single connective $\rightarrow$ and positive literals only. The results considering the limit distribution for the And/Or-model are found to be literally similar to the results from the balanced And/Or-trees. The only differences are related to the speed of convergence. In the model with implication the limit distribution collapses to TRUE. The authors offer two different approaches for the results, one with analytic combinatorics and the other with probabilistic methods con-

[1]The addressed relation is the following: Let $y_i(z)$ be the GF for the expressions computing f_i in the DLW model and $Y(z)$ the GF of all expressions and σ their singularity, then the probability of f_i in the Galton-Watson model is $\frac{y_i(\sigma)}{T(\sigma)}$, which we proved to be continuous in 4.13.

[2]In the case of none uniform distribution μ_0 on all literals the restriction $\mu_0(x_i) = \mu_0(\overline{x_i})$ is made.

sidering the discrete growth process as a continuous process with exponential clocks. Their intuitive explanation for the similarities between the balanced And/Or model and this model is that this model exhibits the property that the smallest path length from the root to a leaf is of order $\ln m$ with probability tending to one (cf.[Pit84]), and this quantity tends to infinity.

The next model is a slight modification of the amplification models considered above. [MS12] considers a single balanced connector α with arity k and arbitrary initial set $\mathbb{H}_0 \subset \mathbb{B}_n$ with arbitrary distribution μ_0. Then, as in the amplification model the balanced expressions built from α and $\mathbb{H}_0$ are considered with the difference that they introduce a noise parameter ϵ which gives the probability that a gate in an expression flips the computed value. So every single computation is false with probability ϵ. For a Boolean function $f \in \mathbb{B}_n$ let f^i $(i = 1, ..., 2^n)$ denote the i-th entry of f represented as a vector[3] and by $\delta(a; b)$ the Kronecker delta. By that the probability of a fixed function f is recursively described by:

$$\begin{aligned}&\mathbb{P}_{m+1}(f)\\ &= \sum_{f_1,...,f_k \in \mathbb{B}_n} \left\{ \prod_{j=1}^{k} \mathbb{P}_m(f_j) \right\} \left\{ \prod_{i=1}^{2^n} \left[(1-\epsilon)\delta(f^i; \alpha(f_1^i, ..., f_k^i)) + \epsilon\delta(1-f^i; \alpha(f_1^i, ..., f_k^i)) \right] \right\}.\end{aligned} \tag{6.1}$$

The authors do not investigate this process but a simplified and averaged process:

$$\tilde{\mathbb{P}}_{m+1}(f) = \sum_{f_1,...,f_k \in \mathbb{B}_n} \left\{ \prod_{j=1}^{k} \tilde{\mathbb{P}}_m(f_j) \right\} \left\{ \prod_{i=1}^{2^n} \frac{\exp^{\beta f^i \alpha(f_1^i,...,f_k^i)}}{2\cosh(\beta\alpha(f_1^i, ..., f_k^i))} \right\}. \tag{6.2}$$

β is related to ϵ by $\epsilon = (1-\tanh(\beta))/2$ and is called the inverse 'temperature' parameter. The case $\beta = 0$ ($\hat{=}$ $\epsilon = 1/2$) corresponds to the completely random case and $\beta \to \infty$ ($\hat{=}$ $\epsilon = 0$) to the deterministic case. A subtlety here is that the maximal unreliability or unpredictability of the system is not obtained for $\epsilon = 1$ but for $\epsilon = 1/2$. When $\epsilon = 1$ that means flipping all computations, which thus is then fully deterministic. This process is analyzed by means of methods from statistical physics, which we can not provide here, to obtain the following interesting result (cf. [MS12]):

[3]So f^i is the entry of the truth table of the i-th line of f,

Theorem 6.1
For any initial distribution μ_0[4] *and balanced gate* α *with* $k \geq 3$*; the uniform distribution* $\mathbb{P}_\infty(f) = \frac{1}{2^{2^n}}$ *is the unique and stable solution of the process defined by Equation* (6.2) *when* $\epsilon > \epsilon(k) = \frac{1-b(k)}{2}$*, where* $b(k)$ *is* $\frac{2^{k-1}}{\binom{k-1}{(k-1)/2}}$ *for* k *odd and* $\frac{2^{k-2}}{\binom{k-2}{(k-2)/2}}$ *else.*

The quantity $\epsilon(k)$ approaches $1/2$ from below for $k \to \infty$. Informally speaking this means that reliable computation for ϵ in this regime is not possible for error probability greater than $\epsilon(k)$ and smaller $1 - \epsilon(k)$ because then the process defined by Equation (6.2) has maximal possible entropy for $m \to \infty$. The also prove that, in the cases when the noiseless process ($\epsilon = 0$) concentrates on a single function, $\mathbb{P}_m$ concentrates on the same function with any desired probability for increasing depth m when $\epsilon < \epsilon(k)$. The authors also investigate the related model for layered $K \times m$ Boolean circuits, which are circuits with depth m where the inputs of a layer are only gates from the layer below and every layer has exactly K nodes, see [MSR10]. The literature of reliable computation in different regimes for computation are broad; the interested reader can find more literature on this, especially for similar models of computation as here, in the two cited papers.

The last model we want to touch is a model investigated in [MSDM11]. The authors consider a random layered $K \times m$ Boolean circuit of increasing depth m, whereas the randomness of the circuit comes from the random labeling with connectives (e.g. $\wedge$ and $\vee$ or NOR, etc.) and the random connectivity (i.e. the random inputs of every gate from the gates from the layer below). The are interested in the behavior of the mean output of the nodes in the last layer, which is a random variable. They examine cases where the mean output converges for $m \to \infty$ very quickly or where it flows to a stable period-two orbit. Moreover, they investigate if chaotic behavior is possible (it is for a model with connectives of arity ≥ 5) and construct a connective where the model has chaotic behavior.

[4] This especially means that $\mathbb{H}_0$ is arbitrary.

6.2 Conclusion

We saw a great variety of different random Boolean function models. We classified them according to the behavior of the induced probability distribution on Boolean function. The first class, the DLW models together with the decorated Galton-Watson trees, exhibits a very regular behavior. Every (expressible) function has non vanishing probability to occur and the probability depends continuously (or probably infinitely differentiable as in the model of Yashunskii) on the parameters for the probabilities of the connectors and base functions. For the DLW model this is formulated in Theorem 4.13. The second model class, the amplification models and the growing tree model, exhibit a quite different behavior. Informally speaking they are more sensitive. They are either uniform distributed on all Boolean functions when everything is well 'balanced' and 'unbiased' or they collapse to a small number of Boolean functions. Moreover, they are uncontinuous with respect to the parameter for the probability of connectors and sometimes also with respect to the distribution on the base functions. These differences and the corresponding classification are formulated on a 'macroscopic' level after we have observed their behavior; but it seems that there is an underlying 'microscopic' structure that causes these two different regimes, namely the structure of the underlying trees. More concrete as already mentioned the trees in the second regime have all the property that a leaf is far away from its root as m tends to infinity, whereas the trees from the first class have the contrary property. So the question is: Can we somehow predict the 'macroscopic' behavior of the induced distribution when we know the 'microscopic' properties of the trees? Or in what cases (connectors) can we do so? Does this reveal anything about the structure of Boolean expressions and Boolean functions?

Another fact that runs like a common thread through the thesis is that all models we have seen (except for the uniform distribution) either do not exhibit the Shannon effect or it is conjectured that they do not exhibit the Shannon effect (commutative and associative And/Or-DLW-models) or lower resp. upper bounds of the functions the distribution concentrates on (e.g. linear threshold functions) that would prove or disprove the Shannon effect are not known. So where is a 'natural' non uniform probability distribution that does exhibit the Shannon effect? It is a matter of fact that in Boolean complexity theory one is not able to find functions with large lower bounds

on the complexity (with AND, OR, NOT gates) or is not able to prove them. Or as S. Jukna formulates it

"we know that almost all Boolean functions are complex, but we cannot exhibit any single example of a complex function!"-S.Jukna [Juk12]

All random Boolean function models in the sense of this thesis and that are found in the literature investigate large random Boolean expressions ($\hat{=}$ trees). But there is also a natural extension of these models: Boolean functions can not only be represented as Boolean expressions but also in a more 'powerful' way as Boolean circuits. So theoretically it is possible to consider random Boolean circuits and the induced distribution on Boolean functions but the analysis of such models needs a different approach and difficulties might arise. In the most general sense whenever one has a set of representations of Boolean functions one can define a distribution on this set to obtain a distribution on Boolean functions that might be interesting.

Another question that arises is if the results from pattern theory can be generalized to prove a similar asymptotic behavior of the probability of a fixed function in the general DLW model with non uniform probabilities for the connectors and literals. For models where only the connectors are non uniform this is straightforward because Theorem 4.21 holds without any changes: The theorem says that the probability of expressions whose structure belongs to $P[T]$ and which have k restrictions is asymptotically $\frac{c_{k,l}}{n^k}$. Because the connectors are irrelevant here the theorem also holds for non uniform distributions on the connectors. To generalize this theorem for probabilities on the literals one has to analyze the probability that a tree of size m with d pattern leaves is labeled by literals so that the tree has n restrictions. In Theorem 4.21 this is done for the uniform distribution on the literals by counting the number of such expressions.
As mentioned already in the section where we considered the model of Yashunskii, it would be interesting to study the same model with respect to the number of leaves or total number of nodes or other size measures. This would lead to similar conclusions for monotone read-once functions as we draw for the number of internal nodes. Moreover, it would be appreciable to extend the results of infinite differentiability of the function $\mathbb{P}_p(\top)$ (with respect to p) and of uniform convergence of $\mathbb{P}_{m,p}(\top)$ to $\mathbb{P}_p(\top)$ to related results for gen-

eral DLW models.

In the tree size model of And/Or expressions the limit probability for TRUE and FALSE is bounded from below and above by constants (see 4.4) but the authors actually expect that this probability converge for n tending to infinity.

The apparent similarities of the DLW formula size models raise the question of a unification:

"Would it be possible to prove a meta-theorem that would give a relation between the probability distributions induced by logical models taking some properties of the connectives into account, or not?"-[GGKM12]

Appendix

Auxiliary Elementary Notions

This appendix covers the elementary notions used in this thesis, which have been excluded from the main text to improve its legibility. Most readers will be very familiar with these notions and notations and can omit them without any doubt. We are commencing with a brief discussion on Boolean functions and list the most important ones used in this thesis. Then, we will present the definitions of graphs, trees and circuits. In the last part of the appendix the $\mathcal{O}$-notation for the asymptotic behavior of functions will be presented.

A.1 Special Boolean functions

The definition of Boolean function is stated in 2.1. In summary: A Boolean function is a function from $\{0,1\}^{\mathbb{N}}$ to $\{0,1\}$, a finite Boolean function maps from $\{0,1\}^n$ to $\{0,1\}$ for an $n \in \mathbb{N}_0$ and we say the functions have arity n. Every finite Boolean function from $\{0,1\}^n$ to $\{0,1\}$ can be identified with a Boolean function mapping from $\{0,1\}^l$ to $\{0,1\}$ for $l \geq n$ and also with a(n) (infinite) Boolean function; and every Boolean function that depends only on a finite number of variables can be identified with a finite Boolean function. With this identifications in mind one should pay attention to the fact that when using a suggestive symbol as $\vee$ or $\oplus$ for a Boolean function it is also important to specify the variables of this functions (if necessary) because $\vee(x_1, x_2) := x_1 \vee x_2$ and $\vee'(x_1, x_3) := x_1 \vee x_3$ are different functions even after identifying them with infinite Boolean functions. For functions for which associative connectors as $\vee$ and $\oplus$ are used as symbol, there is also the subtlety that the number of variables the functions are defined on is

Symbol	Arity	Values				Name, Explanatory note
x_1		0	0	1	1	Values for the first variable
x_2		0	1	0	1	Values for the second variable
$\top \hat{=} TRUE$	0	1	1	1	1	TRUE, is always 1.
$\bot \hat{=} FALSE$	0	0	0	0	0	FALSE, is always 0.
$x_1 \hat{=} Id$	1	0	0	1	1	Identity, is 1 iff $x_1 = 1$.
$\overline{x_1} \hat{=} \neg$	1	1	1	0	0	Negation, is 1 iff $x_1 = 0$.
$\wedge \hat{=} AND$	2	0	0	0	1	AND, is 1 iff both variables are 1.
$\vee \hat{=} OR$	2	0	1	1	1	OR, is 1 iff (at least) one variable is 1.
$\uparrow \hat{=} NAND$	2	1	1	1	0	NAND, is 1 iff not both variables are 1.
$\downarrow \hat{=} NOR$	2	1	0	0	0	NOR, is 1 iff both variables are 0.
$\oplus$	2	0	1	1	0	Parity, is 1 iff exactly one variable is 1.
$\rightarrow$	2	1	1	0	1	Implication, is 1 iff $x_1 = 0$ or $x_2 = 1 (\hat{=} (\neg x_1) \vee (x_2))$.
$\leftrightarrow$	2	1	0	0	1	Equivalence, is 1 iff the variables have same value.
$tr_{(k,n)}$	n	-	-	-	-	Threshold, is 1 iff at least k from n variables are 1.
$\oplus_n$	n	-	-	-	-	Parity, is 1 iff an odd number of variables is 1.

Figure A.1: Table of important Boolean functions.

not clear. In such a case it is important to clarify which $\vee$ we are actually referring to. In A.1 all functions of arity 0, all functions with arity 1 with respect to variable substitution and the most important functions of arity 2 and n are listed. For the threshold function $tr_{(k,n)}$ the n-ary AND($k = n$), the n-ary OR ($k = 1$) and the n-ary TRUE ($k = 0$) (resp. FALSE ($k > n$)) are included. The threshold function $tr_{(n+1,2n+1)}$ is also called majority function maj_{2n+1}.

A.2 Graphs, trees, circuits

Graphs and specializations of them play an essential role in discrete mathematics. Also in this thesis special graphs, called trees, are the main object of interest.

Definition A.2 (Directed graph)
A directed graph, or just graph, $G = (V, E)$ is a tuple of vertex (or node set V) and edge set E, both finite, whereas E contains only (ordered) pairs of elements in V but no loops: $(v, v) \notin E, v \in V$.

One can think of a graph as the vertices being points in the plane and two different vertices v_1, v_2 being connected by an edge reaching from v_1 to v_2 iff $e = (v_1, v_2) \in E$ and we say that v_2 is a parent of v_1 and v_1 is a child of v_2. The number of parents of a vertex is called the out-degree of the vertex and the number of children is called in-degree. A sequence of nodes $(n_1, ..., n_i), n_j \in V$ is called a path from n_1 to n_i iff $(n_j, n_{j+1}) \in E, \forall 1 \leq j \leq i-1$ and the path length from n_1 to n_i is then defined as $i - 1$. A graph is said to be strongly connected if for every $v_1 \neq v_2$ there is a path from v_1 to v_2. There are a lot of other properties a directed graph can have, all of which give rise to the definition of special families of graphs and there are also extensions of graphs, e.g. multigraphs, weighted graphs, labeled graphs etc. Further information can be found in books on graph theory, but the definitions are often inconsistent in the literature. An extension and at the same time restriction of graphs are rooted plane trees, called trees in this work:

Definition A.3 (Rooted plane trees, short: trees)
A (rooted plane) tree is a graph $G = (V, E)$ with the properties:

- *(rooted) Exactly one vertex, called root, has out-degree* 0.
- *(tree) All nodes except for the root have out-degree exactly* 1.
- *(plane) The children of every node are ordered.*

According to this ordering we call the child-nodes of a node the first child, the second child and so on or in the case of two children, left and right child.

The depth of a tree is the maximal path length occurring in the tree. The subtree rooted at a node v is the tree that consists of all nodes v_1 for which there exists a path from v_1 to v and associated edges. Nodes with in-degree > 0 are called inner or internal nodes and else input nodes or leaves. There is an isomorphic recursive definition of trees that will be most of the time the way of looking at things:

Definition A.4 (Rooted plane trees, short: trees)
A tree is either a single node or a tuple of trees: $(t_1, t_2, ..., t_i), i \in \mathbb{N}$, t_j *a tree.*

In Figure A.2 one finds a given graph and tree, represented in the plane, whereas in the case of the graph the order of the children is irrelevant in contrast to the tree-case. Next we will introduce a generalization of Boolean expressions defined in Definition 2.5 resp. Definition 2.6, the circuit. In Boolean complexity theory circuits are the main objects of interest, they are designed to compute a special Boolean function. In circuit design it is of great interest to find small circuits for computing a given function, because this can lead to small/effective calculation units in computers, which abstractly and idealized can be seen as circuits. For defining it, one needs to refer to the definition of Boolean function from 2.1 or from A.1. Let $\mathbb{K}$ and $\mathbb{F}$ be sets of finite Boolean functions, i.e. they have only finitely many input variables.

Definition A.5 (Boolean circuits)
A Boolean circuit with connectors $\mathbb{K}$ *over the basis* $\mathbb{F}$*, is a graph with the properties:*

- *Exactly one vertex, called output, has out-degree* 0*.*
- *The other nodes have out-degree greater or equal to* 1*.*
- *The children of every node are ordered.*
- *There is no path from any node to itself, i.e. acyclic.*
- *Nodes with in-degree* 0 *are labeled by elements from* $\mathbb{F}$ *and all other nodes are labeled by a Boolean function from* $\mathbb{K}$ *with arity being equal to the in-degree of the node.*

The definitions of children nodes, parent nodes, inner nodes, input nodes etc. for trees are similarly used for circuits. A circuit computes a Boolean function: For an input $\vec{x} = (x_1, ..., x_n)$ start evaluating the input nodes as the values of the function labeling these nodes. These values are the new inputs for the functions that label the parent nodes. A demonstrative example of a circuit and the evaluation of an input is pictured in Figure A.3. In the

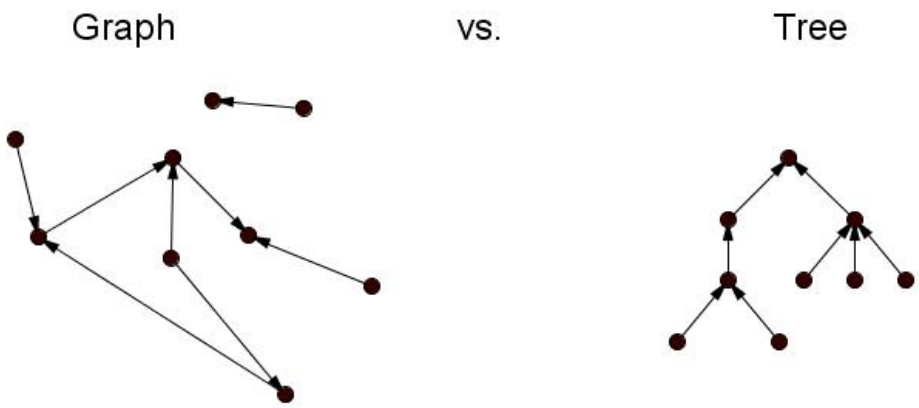

Figure A.2: Graph and Tree.

context of Boolean circuits one calls the in-degree of a node also fan-in and the out-degree fan-out.

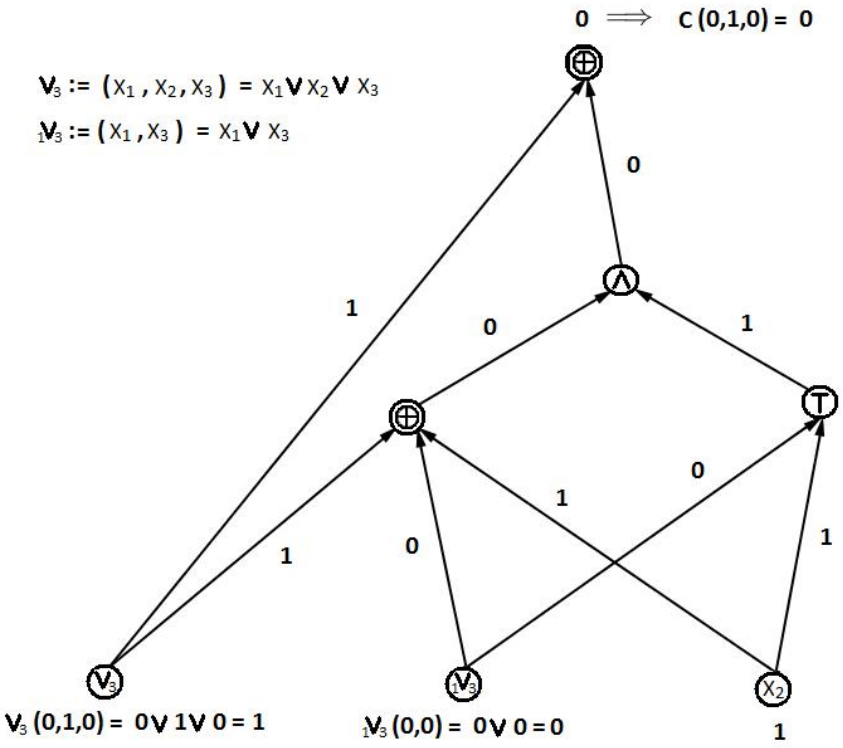

Figure A.3: Circuit and evaluation for input $(0, 1, 0)$. For the definition of the used functions see Appendix A.1

A.3 Oh-notation

A useful notation that we are going to utilize for specifying the asymptotic behavior of a function at a point s or at ∞ and that we are going to use extensively is the so called $\mathcal{O}$-notation:

Definition A.6 (Asymptotics)
Given two real valued functions from a subset of the real numbers, say $g(x)$ and $f(x)$ and $s \in \mathbb{R} \cup \{-\infty, \infty\}$.

Little-Oh: *We write $f(x) \underset{x \to s}{=} o(g(x))$ iff for all positive real numbers C and for all sequences $x_n \to s$ there is a number N with*

$$|f(x_n)| \leq C|g(x_n)|, \forall n \geq N$$

and say f is little-Oh of g as $x \to s$.

Big-Oh: *We write $f(x) \underset{x \to s}{=} \mathcal{O}(g(x))$ iff there exists a positive real number C so that for all sequences $x_n \to s$ there is a number N with*

$$|f(x_n)| \leq C|g(x_n)|, \forall n \geq N$$

and say f is big-Oh of g as $x \to s$.

Θ: *We write $f(x) \underset{x \to s}{=} \Theta(g(x))$ iff $f(x) \underset{x \to s}{=} \mathcal{O}(g(x))$ and $g(x) \underset{x \to s}{=} \mathcal{O}(f(x))$ and say f is of the same asymptotic order as g.*

$\sim$: *We write $f(x) \underset{x \to s}{\sim} g(x)$ iff $\lim_{x \to s} \frac{f(x)}{g(x)} = 1$ and say f and g are asymptotically equivalent as $x \to s$.*

These notations alongside the most important rules for manipulations as well as a range of examples can be found in [GKP68, Section 9.2]. We will use them mostly for $s = \infty$ and for functions defined on $\mathbb{N}_0 \subset \mathbb{R}$, hence we will not specify s when it is clear from the context. Then, we write $f(n) = \mathcal{O}(g(n))$ instead of $f(n) \underset{n \to \infty}{=} \mathcal{O}(g(n))$ for convenience.

Bibliography

[AKS83] M. Ajtai, J. Komlós, and E. Szemerédi. An O(n log n) sorting network. Proceedings of 15th ACM Symposium on Theory of Computing, pages. 1-9., 1983.

[AN72] K. Athreya and P. Ney. Branching Processes. Springer, 1972.

[BBY06] J. P. Bell, S. N. Burris, and K. A Yeats. Counting rooted trees: The universal law $t(n) \sim C\rho^{-n}n^{-3/2}$. Electronic Journal of Combinatorics 13, R63, 1–64., 2006.

[Bop85] R. B. Boppana. Amplification of probabilistic Boolean formulas. In Proceedings of the 26th IEEE Symposium on Foundations of Computer Science, pages 20–29, 1985.

[BP05] A. Brodsky and N. Pippenger. The boolean functions computed by random boolean formulas or how to grow the right function. Random Structures and Algorithms, 27:490–519, 2005.

[CFGG04] B. Chauvin, P. Flajolet, D. Gardy, and B. Gittenberger. And/Or trees revisited. Combinatorics, Probability and Computing, 13(4-5):475–497, July-September 2004.

[CGM11] B. Chauvin, D. Gardy, and C. Mailler. The growing tree distribution on Boolean functions. In Proceedings of the 8th SIAM Workshop on Analytic and Combinatorics (ANALCO), pages 45–56, 2011.

[Drm97] M. Drmota. Systems of functional equations. Random Structures and Algorithms. Random Structures and Algorithms, 1997.

[Dur05] R. Durrett. Probability: Theory and Examples. Duxbury, Belmont, CA, 3rd edition, 2005.

[dW08] R. de Wolf. A Brief Introduction to Fourier Analysis on the Boolean Cube. Theory of Computing Libary, Graduate Surveys, TCGS 1 ,pages 1–20, 2008.

[EGK10] J. Esparza, A. Gaiser, and S. Kiefer. Computing Least Fixed Points of Probabilistic Systems of Polynomials. In Proceedings of STACS, pages 359–370, 2010.

[EKL10] J. Esparza, S. Kiefer, and M. Luttenberger. Computing the Least Fixed Point of Positive Polynomial Systems. SIAM Journal on Computing, v.39 n.6, p.2282-2335, 2010.

[FGG09] H. Fournier, D. Gardy, and A. Genitrini. Balanced And/Or trees and linear threshold functions. In 6th SIAM Workshop on Analytic and Combinatorics (ANALCO), pages 51–57, 2009.

[FGGG08] H. Fournier, D. Gardy, A. Genitrini, and B. Gittenberger. Complexity and limiting ratio of boolean functions over implication. In MFCS, pages 347–362, 2008.

[FGGG12] H. Fournier, D. Gardy, A. Genitrini, and B. Gittenberger. The fraction of large random trees representing a given Boolean function in implicational logic. Random Structures and Algorithms, 40(3):317–349, 2012.

[FGGZ07] H. Fournier, D. Gardy, A. Genitrini, and M. Zaionc. Classical and intuitionnistic logic are asymptotically identical. In Springer-Verlag, editor, Annual Conference on Computer Science Logic (CSL'07), pages 177–193, Lausanne, Suisse, 2007.

[FO90] P. Flajolet and A. Odlyzko. Singularity analysis of generating functions. SIAM Journal on discrete mathematics, Vol.3 (1990), no.2, 216-240, 1990.

[FS09] P. Flajolet and R. Sedgewick. Analytic Combinatorics. Cambridge U.P.,Cambridge, 2009.

[Gar06] D. Gardy. Random Boolean expressions. In Colloquium on Computational Logic and Applications, volume AF, pages 1–36. DMTCS Proceedings, 2006.

[GG10] A. Genitrini and B. Gittenberger. No Shannon effect on probability distributions on Boolean functions induced by random expressions. In 21st International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, Vienna, Austria. DMTCS Proceedings, 2010.

[GGKM12] A. Genitrini, B. Gittenberger, V. Kraus, and C. Mailler. Probabilities of boolean functions given by random implicational formulas. Electronic Journal of Combinatorics , 19(2):P37, 20 pages, (electronic), 2012.

[GGKM14] A. Genitrini, B. Gittenberger, V. Kraus, and C. Mailler. The relation between tree size complexity and probability for Boolean functions generated by uniform random trees. arXiv:1407.0501v1, submitted: 2014.

[GGKM15] A. Genitrini, B. Gittenberger, V. Kraus, and C. Mailler. Associative and commutative tree representations for boolean functions. Theoretical Computer Science 570 , 70–101, 2015.

[GGM14] A. Genitrin, B. Gittenberger, and C. Mailler. No Shannon-effect induced by And/Or trees. DMTCS Proceedings BA, pages 109-120, 2014.

[GK09] A. Genitrini and J. Koziki. Quantitative comparison of intuitionistic and classical logics - full propositional system. In LFCS, pages 280–294, 2009.

[GKP68] R. L. Graham, D. E. Knuth, and O. Patashnik. Concrete Mathematics. Addison-Wesley, 1968.

[GM97] A. Gupta and S. Mahajan. Using amplification to compute majority with small majority gates. Computational Complexity, 6(1):46–63, 1997.

[GM15] A. Genitrini and C. Mailler. Generalised and Quotient Models for Random And/Or Trees and Application to Satisfiability. arXiv:1507.08448v1, submitted 2015.

[GMS93] S.A. Goldman, M.J.Kearns, and R.E. Schapire. Exact identifiation of read-once formulas using fixed points of

amplification functions. SIAM Journal on Computing, 22(4):705-726, 1993.

[GW05] D. Gardy and A. Woods. And/or tree probabilities of Boolean functions. First International Conference on the Analysis of Algorithms, pages 139–146, Barcelona (Spain), DMTCS Proceedings AD, June 2005.

[HMP06] S. Hoory, A. Magen, and T. Pitassi. Monotone circuits for the majority function. 10th International Workshop on Randomization and Computation (RANDOM), 2006.

[Juk12] S. Jukna. Boolean Function Complexity: Advances and Frontiers. Springer, 2012.

[Kos03] Z. Kostrzycka. On the density of truth of implicational parts of intuitionistic and classical logics. Journal of Applied Non-Classical Logics, 13(2), 2003.

[Koz08] J. Kozik. Subcritical pattern languages for And/Or trees. In Fifth Colloquium on Mathematics and Computer Science, Blaubeuren, Germany. DMTCS Proceedings, September 2008.

[Kra11] V. Kraus. Asymptotic study of families of unlabelled trees and other unlabelled graph structures. Ph.D. Thesis, Institute of Discrete Mathematics and Geometry, Vienna University of Technology, 2011.

[KZ04] Z. Kostrzycka and M. Zaionc. Statistics of intuitionistic versus classical logics. Studia Logica, 76(3):307–328, 2004.

[Lal93] S.P. Lalley. Finite range random walk on free groups and homogeneous trees. Ann. Probab., 21(4):2087–2130, 1993.

[LS95] H. Lefmann and P. Savický. Some typical properties of large AND/OR Boolean formulas. Proc. MFCS'95, Lecture Notes in Computer Science 969, pages 237 - 246, 1995.

[Lup58] O. B. Lupanov. On a method of circuit synthesis. Izvestia VUZ Radiofizika, 1(1):120–140, (in Russian), 1958.

[Lup60] O. B. Lupanov. Complexity of formula realization of functions of logical algebra. Kibernetiki, 3:61-80. English translation: Problems of Cybernetics, Pergamon Press, 3:782–811, 1962, 1960.

[MM78] A. MEIR and J. W. MOON. On the altitude of nodes in random trees. Canadian Journal of Mathematics 30, 997–1015, 1978.

[MS56] E. F. Moore and C.E. Shannon. Reliable circuits using less reliable relays. Journal of Franklin Institute, 262(3):191–208, 1956.

[MS12] A. Mozeika and D. Saad. On reliable computation by noisy random Boolean formulas. IEEE Transactions on Information Theory, VOL. X, NO. X, 2012.

[MSDM11] J. Machta, S. Mertens S. DeDeo, and C. Moore. Parallel complexity of random Boolean circuits. Journal of Statistical Mechanics: Theory and Experiment, Volume 2011, 2011.

[MSR10] A. Mozeika, D. Saad, and J. Raymond. Noisy random Boolean formulae: A statistical physics perspective. Phys. Rev. E, 82(4):041112, 2010.

[MTZ00] M. Moczurad, J. Tyszkiewicz, and M. Zaionc. Statistical properties of simple types. Mathematical Structures in Computer Science, 10(5):575–594, 2000.

[Odl95] A. Odlyzko. Asymptotic enumeration methods. In Handbook of Combinatorics, R. Graham, M. Grötschel, and L. Lovász, Eds., vol. II. Elsevier, pages 1063–1229, 1995.

[Pó54] G. Pólya. Induction and Analogy in Mathematics. Princeton University Press, 1954.

[Pit84] B. Pittel. On growing random binary trees. Journal of Mathematical Analysis and Applications, 103(2):461–480, 1984.

[PVW94] J. B. Paris, A. Vencovská, and G. M. Wilmers. A natural prior probability distribution derived from the propositional calculus. Annals of Pure and Applied Logic, 70:243–285, 1994.

[PW06] R. Pemantle and M. Ward. Exploring the average values of Boolean functions via asymptotics and experimentation. In The Proceedings of the Third Workshop on Analytic Algorithmic and Combinatorics (ANALCO'06), pages 253–262, 2006.

[Raz88] A. A. Razborov. Formulas of bounded depth in basis $\{\wedge, \oplus\}$ and some combinatorial problems (in Russian). Voprosy Kibernttiky, SloZnosi vycislenij i prikladnaja mattmaticeskaja logika 149-166, Moscow, 1988.

[RS42] J. Riordan and C.E. Shannon. The number of two-terminal series-parallel networks. Journal of Mathematics and Physics, 21:83–93, 1942.

[Rud87] W. Rudin. Real and complex analysis. 3rd ed. McGraw-Hill Book Co., 1987.

[RW00] S. Reith and K. W. Wagner. The complexity of problems defined by Boolean circuits. In The Mathematical Foundation of Informatics pages 141–156, 2000.

[Sav87] P. Savický. Boolean functions represented by random formulas. announcement, Commentationes Mathematicae Universitatis Carolinae 28(2), 1987.

[Sav90] P. Savický. Random Boolean formulas representing any Boolean function with asymptotically equal probability. Discrete Mathematics, 83:95–103, 1990.

[Sav95] P. Savický. Bent functions and random Boolean formulas. Discrete Mathematics 147, pages 211-234, 1995.

[Sha49] C.E. Shannon. The synthesis of two-terminal switching circuits. Bell System Technical Journal, 28(28):59–98, 1949.

[Val84] L. Valiant. Short monotone formulae for the majority function. Journal of Algorithms, 5:363–366, 1984.

[Weg87] I. Wegener. The complexity of Boolean functions. Wiley and Sons Inc., 1987.

[Wil94] H. S. Wilf. generatingfunctionology. Academic Press Inc., Boston, MA, second edition, 1994.

[Woo97] A. R. Woods. Coloring rules for finite trees, and probabilities of monadic second order sentences. Random Structures and Algorithms, 10(4):453–485, 1997.

[Woo05] A. Woods. On the probability of absolute truth for And/Or formulas. Bulletin of Symbolic Logic, 12(3), 2005.

[Yas05] A. D. Yashunskii. On the Properties of Asymptotic Probability for Random Boolean Expression Values in Binary Bases. Lecture Notes in Compututer Science, : Proceedings of the Third International Symposium SAGA 2005, Moscow, Russia (Springer, Berlin, 2005), pages 202–212, 2005.

[Yas06] A. D. Yashunskii. On the Asymptotic Probability for Values of Random Boolean Expressions. Diskret. Analiz i Issled. Operatsii, Ser. 1, 13 (2), 66, 2006.

[Yas07] A. D. Yashunskii. Uniform Approximation of Continuous Functions by Probability Functions of Boolean Bases. Moscow University Mathematics Bulletin, Vol. 62, No. 2, pages 78–84, 2007.

[Yas15] A. D. Yashunskii. Private Communication, September-October 2015.

[Zai03] M. Zaionc. Statistics of implicational logic. Electronic Notes in Theoretical Computer Science, 84, 2003.

[Zai05] M. Zaionc. On the asymptotic density of tautologies in logic of implication and negation. Reports on Mathematical Logic, 39:67–87, 2005.