



University of Tennessee, Knoxville  
**TRACE: Tennessee Research and Creative  
Exchange**

---

Doctoral Dissertations

Graduate School

---

12-2001

## Language extraction from ZnS

Dowman Parks Varn  
*University of Tennessee*

Follow this and additional works at: [https://trace.tennessee.edu/utk\\_graddiss](https://trace.tennessee.edu/utk_graddiss)

---

### Recommended Citation

Varn, Dowman Parks, "Language extraction from ZnS. " PhD diss., University of Tennessee, 2001.  
[https://trace.tennessee.edu/utk\\_graddiss/6402](https://trace.tennessee.edu/utk_graddiss/6402)

This Dissertation is brought to you for free and open access by the Graduate School at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact [trace@utk.edu](mailto:trace@utk.edu).

To the Graduate Council:

I am submitting herewith a dissertation written by Dowman Parks Varn entitled "Language extraction from ZnS." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Physics.

Chia C. Shih, Major Professor

We have read this dissertation and recommend its acceptance:

Geoffrey S. Canright, John J. Quinn, Charles R. Collins

Accepted for the Council:

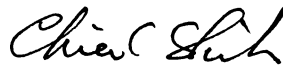
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

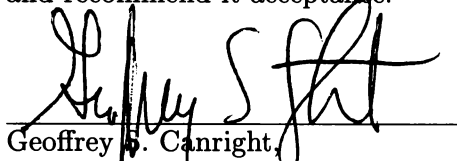
I am submitting herewith a dissertation written by Dowman Parks Varn entitled "Language Extraction from ZnS." I have examined the final paper copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Physics.



---

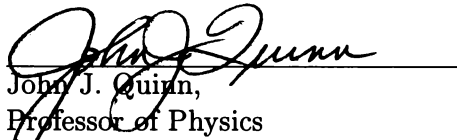
Chia C. Shih,  
Professor of Physics

We have read this dissertation  
and recommend its acceptance:



---

Geoffrey S. Canright,  
Associate Professor of Physics



---

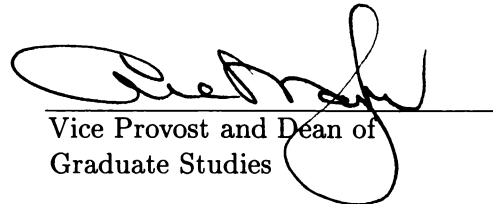
John J. Quinn,  
Professor of Physics



---

Charles R. Collins,  
Associate Professor of Mathematics

Accepted for the Council:



---

Vice Provost and Dean of  
Graduate Studies

# Language Extraction from ZnS

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Dowman Parks Varn

December, 2001

*for my parents,  
Jim and Pat*

and have immensely enjoyed our collaboration. My parents have been an unending source of support my entire life, and to them I dedicate this thesis. Finally, without the love, patience, and fortitude of my wife, Mary, I could not have flourished.

I acknowledge financial support for my studies at the University of Tennessee as a teaching assistant and later as a research assistant under grant DMR-9820816. I also acknowledge support under a grant to the Santa Fe Institute from the National Science Foundation, award PHY-9910217, "Pilot Program for NSF Physics Graduate Student Fellowships."

# Contents

<b>1</b>	<b>Prelude</b>	<b>1</b>
<b>2</b>	<b>A Brief Introduction to Computational Mechanics</b>	<b>5</b>
2.1	The Intellectual Precursors of Computation Mechanics . . . . .	5
2.1.1	Symbolic Dynamics . . . . .	5
2.1.2	The Formal Theory of Languages . . . . .	6
2.1.3	Information Theory . . . . .	10
2.2	Computational Mechanics . . . . .	11
2.3	De Bruijn Graphs . . . . .	15
<b>3</b>	<b>Some Practical Details regarding Polytypes and Diffraction Patterns</b>	<b>17</b>
3.1	Notational Matters concerning Polytypes . . . . .	17
3.2	A Brief Description of the Experiment . . . . .	18
3.3	Some Kinematic Details . . . . .	18
3.4	How to fix Diffraction Patterns . . . . .	19
3.5	Some Assumptions . . . . .	21
<b>4</b>	<b>All about Two-Layer Correlation Functions and Diffraction Patterns</b>	<b>23</b>
4.1	$Q$ -extraction . . . . .	24
4.2	Figures of Merit . . . . .	25
4.3	Asymptotic Behavior of the Correlation Functions . . . . .	26
4.3.1	$Q$ s decay to an asymptotic value of $\frac{1}{3}$ . . . . .	27
4.3.2	Periodic Oscillation in the $Q$ s . . . . .	28
4.3.3	Aperiodic Oscillation in the $Q$ s . . . . .	28
4.4	Measures of Correlation and Diffraction . . . . .	28
<b>5</b>	<b>Examples of Processes, Correlation Functions and their Diffraction Patterns</b>	<b>31</b>
5.1	The Fair Coin Toss . . . . .	32
5.2	The Period One or 3C Process . . . . .	36
5.3	The Period Two or 2H Process . . . . .	36
5.4	The Golden Mean Process . . . . .	39
5.5	The 3C/2H Process . . . . .	40
5.6	The 4H Process . . . . .	44
5.7	The 3C/2H/4H Process . . . . .	49
5.8	The 6H Process . . . . .	52
5.9	The 3C/2H/6H Process . . . . .	52
5.10	A Period 14 $D$ -pair . . . . .	57
5.11	The Noisy Period Two Process . . . . .	58
5.12	The Even System . . . . .	61
5.13	The Sum Zero Process . . . . .	63

# List of Figures

2.1	An example of a finite automaton. . . . .	8
2.2	The finite automaton for the golden mean language . . . . .	9
2.3	A pictorial representation of causal states. . . . .	12
2.4	Examples of de Bruijn Graphs. . . . .	15
3.1	Correction Factor $\mathcal{C}(l)$ for ZnS diffraction patterns. . . . .	21
5.1	The $\epsilon$ -machine for an unbiased RNG or a fair coin toss. . . . .	32
5.2	$Q_s(n)$ for the fair coin toss as a function of $n$ . . . . .	34
5.3	The $\log_2 \Psi_q(n)$ vs. $n$ for the fair coin toss. . . . .	34
5.4	Diffraction pattern for a randomly stacked two dimensional hexagonal lattice using asymptotic values for the $Q_s$ . . . . .	35
5.5	Diffraction pattern for a randomly stacked two dimensional hexagonal lattice using sequence calculated values for the $Q_s$ . . . . .	35
5.6	The $\epsilon$ -machine for the period one process. . . . .	36
5.7	Diffraction pattern for the period one or 3C process. . . . .	37
5.8	The recurrent portion of $\epsilon$ -machine for the period two process. . . . .	38
5.9	Diffraction pattern for a lattice stacked according to the period two rule. . . . .	39
5.10	The recurrent portion of the $\epsilon$ -machine for the golden mean process. . . . .	40
5.11	$Q_s(n)$ for the golden mean process as a function of $n$ . . . . .	41
5.12	The logarithm of $\Psi_q(n)$ for the golden mean process as a function of $n$ . . . . .	41
5.13	The diffraction pattern for the golden mean process. . . . .	42
5.14	The recurrent portion of the $\epsilon$ -machine for the 3C/2H process. . . . .	43
5.15	$Q_s(n)$ vs. $n$ for the 3C/2H process with $q = 0.01$ . . . . .	44
5.16	The logarithm of $\Psi_q(n)$ for the 3C/2H process as a function of $n$ , with $q = 0.01$ . . . . .	45
5.17	The diffraction pattern for the 3C/2H process with $q = 0.01$ . . . . .	45
5.18	The diffraction pattern for the 3C/2H process with $q = 0.05$ . . . . .	46
5.19	The diffraction pattern for the 3C/2H process with $q = 0.10$ . . . . .	46
5.20	The diffraction pattern for the 3C/2H process with $q = 0.20$ . . . . .	47
5.21	The diffraction pattern for the 3C/2H process with $q = 0.30$ . . . . .	47
5.22	The diffraction pattern for the 3C/2H process with $q = 0.40$ . . . . .	48
5.23	The recurrent portion of the $\epsilon$ -machine for the 4H process. . . . .	48
5.24	The diffraction pattern for the 4H process. . . . .	49
5.25	The recurrent portion of the $\epsilon$ -machine for the 3C/2H/4H process. . . . .	50
5.26	$Q_s(n)$ vs. $n$ for 3C/2H/4H process. . . . .	51
5.27	The diffraction pattern for the 3C/2H/4H process. . . . .	52
5.28	The $\epsilon$ -machine for the 6H process. . . . .	53
5.29	The diffraction pattern the 6H process. . . . .	54
5.30	The recurrent portion of the $\epsilon$ -machine for the 3C/2H/6H process. . . . .	55
5.31	$Q_s(n)$ vs. $n$ for 3C/2H/6H process. . . . .	56



7.25	The diffraction pattern for the even process and the $r = 2$ approximation. . . . .	108
7.26	The $Q_s(n)$ vs. $n$ for the even process and the $r = 3$ approximation. . . . .	108
7.27	The diffraction pattern for the even process and the $r = 3$ approximation. . . . .	109
8.1	Growth faults in the 2H structure of ZnS depicted on an $r = 3$ de Bruijn graph. . . .	113
8.2	Growth faults in the 3C structure of ZnS depicted on an $r = 3$ de Bruijn graph. . . .	114
8.3	Deformation faults in the 2H structure of ZnS depicted on an $r = 3$ de Bruijn graph.	116
8.4	Deformation faults in the 3C structure of ZnS depicted on an $r = 3$ de Bruijn graph.	117
8.5	Layer displacement faults in the 2H structure of ZnS depicted on an $r = 3$ de Bruijn graph. . . . .	118
8.6	Layer displacement faults in the 3C structure of ZnS depicted on an $r = 3$ de Bruijn graph. . . . .	119
8.7	Double deformation faults in the 3C structure of ZnS depicted on an $r = 3$ de Bruijn graph. . . . .	120
9.1	Uncorrected diffractometer pattern for Experimental Data SK229. . . . .	122
9.2	Corrected diffraction pattern for Experimental Data SK229. . . . .	123
9.3	The $Q_s(n)$ vs. $n$ for Experimental Data SK229. . . . .	123
9.4	The $r = 3$ machine for diffraction pattern SK229. . . . .	125
9.5	The $Q_s(n)$ vs. $n$ for Experimental Data SK229 and the $r = 3$ approximation. . . . .	126
9.6	The diffraction pattern for Experimental Data SK229 and the $r = 3$ approximation.	126
9.7	Uncorrected diffraction pattern for Experimental Data SK230. . . . .	127
9.8	Diffraction pattern for Experimental Data SK230 corrected for $C(l)$ . . . . .	128
9.9	The $Q_s(n)$ vs. $n$ for Experimental Data SK230. . . . .	128
9.10	The $r = 3$ machine for diffraction pattern SK230. . . . .	130
9.11	The $Q_s(n)$ vs. $n$ for Experimental Data SK230 and the $r = 3$ approximation. . . . .	131
9.12	The diffraction pattern for Experimental Data SK230 and the $r = 3$ approximation.	131
9.13	Uncorrected diffraction pattern for Experimental Data SK231. . . . .	133
9.14	The corrected diffraction pattern for Experimental Data SK231. . . . .	134
9.15	The $Q_s(n)$ vs. $n$ for Experimental Data SK231. . . . .	135
9.16	The $r = 3$ machine for diffraction pattern SK231. . . . .	136
9.17	The $Q_s(n)$ vs. $n$ for Experimental Data SK231 and the $r = 3$ approximation. . . . .	137
9.18	The diffraction pattern for Experimental Data SK231 and the $r = 3$ approximation.	137
9.19	Uncorrected diffraction pattern for Experimental Data SK232. . . . .	138
9.20	The corrected diffraction pattern for Experimental Data SK232. . . . .	139
9.21	The $Q_s(n)$ vs. $n$ for Experimental Data SK232. . . . .	140
9.22	The $r = 3$ machine for diffraction pattern SK232. . . . .	141
9.23	The $Q_s(n)$ vs. $n$ for Experimental Data SK232 and the $r = 3$ approximation. . . . .	142
9.24	The diffraction pattern for Experimental Data SK232 and the $r = 3$ approximation.	143
9.25	Uncorrected diffraction pattern for Experimental Data SK134. . . . .	144
9.26	The corrected diffraction pattern for Experimental Data SK134. . . . .	144
9.27	The $Q_s(n)$ vs. $n$ for Experimental Data SK134. . . . .	145
9.28	The $r = 3$ machine for diffraction pattern SK134. . . . .	147
9.29	The $Q_s(n)$ vs. $n$ for Experimental Data SK134 and the $r = 3$ approximation. . . . .	148
9.30	The $Q_s(n)$ vs. $n$ for Experimental Data SK134 and the fault model. . . . .	148
9.31	The diffraction pattern for Experimental Data SK134 and the $r = 3$ approximation.	149
9.32	The diffraction pattern for Experimental Data SK134 and the fault model.	149
9.33	Uncorrected diffraction pattern for Experimental Data SK135. . . . .	150
9.34	Corrected diffraction pattern for Experimental Data SK135. . . . .	151
9.35	The $Q_s(n)$ vs. $n$ for Experimental Data SK135. . . . .	152
9.36	The $r = 3$ machine for diffraction pattern SK135. . . . .	154

# List of Tables

2.1	The Chomsky Hierarchy . . . . .	6
2.2	The number of simple cycles in a de Bruijn graph of range $r$ . . . . .	16
5.1	The first few values of the correlation functions for the fair coin toss. . . . .	33
5.2	The first few values of the correlation functions for the period one process. . . . .	37
5.3	The first few values of the correlation functions for the period two process. . . . .	38
5.4	The first few values of the correlation functions for the golden mean process. . . . .	40
5.5	The first few values of the correlation functions for the 3C/2H process, with $q = 0.01$ . . . . .	43
5.6	The first few values of the correlation functions for the 4H process. . . . .	49
5.7	The first few values of the correlation functions for the 3C/2H/4H process. . . . .	51
5.8	The first few values of the correlation functions for the 6H process. . . . .	53
5.9	The first few values of the correlation functions for the 3C/2H/6H process. . . . .	54
5.10	The first few values of the correlation functions for period 14 $D$ -pair. . . . .	58
5.11	The first few values of the correlation functions for the noisy period two process. . . . .	60
5.12	The first few values of the correlation functions for the even system. . . . .	63
5.13	The first few values of the correlation functions for the sum zero process. . . . .	66
5.14	The first few values of the correlation functions for the Thue-Morse process. . . . .	68
5.15	Computational Results . . . . .	72
5.16	Correlative and diffractive results . . . . .	73
6.1	Exact and noisy correlation functions for the random number process. . . . .	85
6.2	Solution at $r = 3$ for the random number generator. . . . .	85
6.3	Exact and noisy correlation functions for the second process. . . . .	85
6.4	Solution at $r = 3$ for for a second process. . . . .	86
6.5	Mathematical requirements for general solution of equations as a function of $r$ . . . . .	88
7.1	Machine reconstruction results for the 3C/2H/6H Process . . . . .	90
7.2	Computational results for 3C/2H/6H process and the $r = 0, 1$ and 2 approximations. . . . .	92
7.3	Machine reconstruction results for the 3C/6H process. . . . .	95
7.4	Computational results for 3C/6H process and the $r = 0, 1, 2$ and 3 approximations. . . . .	98
7.5	Machine reconstruction results for the noisy period two system. . . . .	99
7.6	Computational results for The Noisy Period Two System and the $r = 0, 1, 2$ and 3 approximations. . . . .	104
7.7	Machine reconstruction results for the even system. . . . .	104
7.8	Computational results for the even process and the $r = 0, 1, 2$ and 3 approximations. . . . .	109
9.1	The first few correlation functions for experimental data SK229. . . . .	122
9.2	Machine reconstruction results for the experimental diffraction pattern SK229. . . . .	124
9.3	Computational Results for the $r = 0, 1, 2$ and 3 approximations to experimental data SK229. . . . .	124
9.4	The first few $Q$ s for experimental data SK230. . . . .	127

# Chapter 1

## Prelude

It is fascinating that Nature should produce such complicated structures with compounds such as SiC, CdI<sub>2</sub>, and ZnS, all of which have a very simple proportion of atoms in them. If one just picks up at random one crystal of any of these materials, out of an as grown bunch of crystals, that tiny speck of the material can have one of several hundred structures - some ordered, some randomly faulted and some with non-random faulting. Which one will it have and why?

*M.T. Sebastian and P. Krishna [60]*

Perhaps the most fundamental questions we can ask about a solid are “What is it made of?” and “How are the constituent parts assembled?” This is so elementary, and yet so basic to any detailed understanding of the thermal, electrical, magnetic, optical, and elastic properties of materials. At the beginning of the twenty-first century, concern over the placement of the atoms in a solid seems quaint and anachronistic, more suited to the dawn of the twentieth century. X-ray diffraction, electron diffraction, optical microscopy, x-ray diffraction tomography, to name a few, are powerful techniques to uncover structure in solids. With this arsenal of tools, and the efforts of many researchers, surely we can have nothing novel to say about the discovery and description of structure in solids, save perhaps the refinement of well-worn techniques or the analysis of particularly obstinate cases. But careful examination of present technology reveals that while we are quite good at finding and describing periodic order in nature, cases that lack such order are much more difficult. Certainly in the complete absence of structural order, as in a gas, statistical methods exist that permit a satisfying understanding of the properties of the system without knowing (or even wanting to know) the details of the microscopic placement of the constituents. But it is the in-between cases, where order and disorder coexist, that has proven so elusive to both analyze and describe. In this thesis, we will tackle these in-between cases for a special type of layered material, called polytypes. They exhibit disorder in one dimension only, making the analysis more tractable. We will give a method for determining the structure of these solids from experimental data and demonstrate how this structure, both the random and the non-random part, can be compactly expressed. From our solution, we will be able to calculate the effective range of the inter-layer interactions, as well as the configurational energies of the disordered stacking sequences.

Zinc sulfide would seem a rather unpretentious material. Its stoichiometry is simple enough, the proportion of zinc to sulphur atoms is 1:1. Each zinc atom is tetrahedrally coordinated to four sulphur atoms and vice versa. It is known to crystallize into the face-centered cubic (fcc) structure - alternately called cubic zinc sulfide, zinc blende or the sphalerite structure - at low temperatures. In this configuration, one can think of the zinc atoms as forming an fcc lattice and the sulphur atoms forming an inter-penetrating fcc lattice displaced from the latter one by one-quarter of a body diagonal [45]. For our purposes, it is useful to look along the [111] direction of the conventional unit cell. Taking a zinc atom at the origin of the this cell, we see that the zinc atoms are arranged hexagonally in the (111) plane with sulphur atoms residing a quarter body

process, called the  $\epsilon$ -machine, directly from the data itself. This model is the optimal, minimal and unique description of the process. So far, computational mechanics has been applied to both artificial systems such as cellular automata [36], the logistic map [17] [80], and the one-dimensional Ising model [22] [15], as well as more physical systems like the dripping faucet [29], atmospheric turbulence [50], and geomagnetic data [12].

It is worth mentioning that our technique for  $\epsilon$ -machine reconstruction is novel. Instead of the usual situation of analyzing a data stream or long sequence of symbols, our information about the underlying process comes in the form of two-layer correlation functions. We will find that they provide only impoverished information about the process and impose severe restrictions on the classes of processes we will be able to detect. We adopt this method not for its superiority to previous methods, but out of necessity. Nonetheless, we will find that it gives a reasonable answer to our problem. While our method is specifically directed towards machine reconstruction in a particular polytypic material, ZnS, we hope that concepts and techniques explored here will prove useful in machine reconstruction from other kinds of power spectra.

The plan for the rest of this thesis is as follows. Since computational mechanics is not a common tool in most physicists' tool chest, we will devote chapter 2 to explaining the basic ideas and providing definitions for some of the quantities we will use subsequently. In chapter 3 we will very briefly discuss the experimental details of the data we will use to infer the underlying process. In chapter 4, we show how correlation information can be extracted from the experimental data. In chapter 5, in order to build intuition, we give examples of the different diffraction patterns arising from different, fake processes. In chapter 6, we discuss our technique for  $\epsilon$ -machine reconstruction from two-layer correlation functions. In chapter 7, we provide examples of  $\epsilon$ -machine reconstruction from diffraction spectra of known processes. In chapter 8, we give a discussion of previous models of disorder in layered solids and show how they relate to  $\epsilon$ -machines. In chapter 9, we employ this machinery on real experimental data for ZnS and give the models for the underlying process. In chapter 10, we present our conclusions and possible directions for future work.

## Chapter 2

# A Brief Introduction to Computational Mechanics

Computational mechanics is not familiar to most physicists. We will therefore give a brief introduction to the ideas leading up to computational mechanics as well as an overview of the theory itself. There are several good references available on computational mechanics and the interested reader is urged to consult these for a much more detailed exposition. These references include Crutchfield and Feldman [15] [16], Feldman [22] [23], Feldman and Crutchfield [24] Shalizi [61], Shalizi and Crutchfield [62], Hanson [36] and Young [80].

### 2.1 The Intellectual Precursors of Computation Mechanics

Computational mechanics arises from the marriage of three distinct lines of thought: symbolic dynamics, language theory and information theory. Since we borrow terminology and concepts from each to formulate computational mechanics, let us consider each separately.

#### 2.1.1 Symbolic Dynamics

In the study of nonlinear dynamical systems, discretizing the formalism can greatly facilitate the analysis. While some systems are naturally discrete, most are not and it is therefore necessary to introduce a some kind of discretization process. One way to do this is to describe a dynamical system by using a map [77] [66] [47]. Maps of course can be use to describe other systems, but the motivation for their study in physics is the connection to dynamical systems. Maps have the convenient property of being discrete in time, and we can additionally impose a partition  $\mathbf{B} = \{B_1, B_2, \dots, B_b\}$  on the other dynamical variables. Doing so, we associate some  $B_i \in \mathbf{B}$  with the state of the system with each iteration of the map. In this way, we can build up a sequence of symbols describing the time evolution of the system. Symbolic dynamics [6] [8] [37] [44] is nothing more than the study of such a symbol sequence. From the sequence generated from an appropriate choice of  $\mathbf{B}$ , it is possible to determine if the orbit of the trajectory is periodic or not, and put a lower bound on the entropy production of the system.

The slicing of a continuum into finite number of wedges may seem severe, and one may think that this is artificial in some way. Much of our experience in the macroscopic domain points to a continuous world. Our measuring instruments - our basic tools for empirical discovery - however are not. In the process of measurement, we must always face the fact that there is some finite resolution and we can not distinguish structure on a scale smaller than this. So partitioning the world into cells is really not so far from our practice.

the derivation of the word. Finally,  $F$  are the actual production rules we use make words. We write them as  $P \rightarrow Q$ , and this simply means we replace  $P$  with  $Q$ . We consider several examples.

**Example 2.1** Let us construct the grammar for the language that allows every word over an alphabet  $\mathcal{A} = \{0, 1\}$ . So let us take  $\mathcal{A}_T = \{0, 1\}$  and  $\mathcal{A}_N = \{S\}$ . The production rules  $F$  are just

$$\begin{aligned} S &\rightarrow \lambda \\ S &\rightarrow 0S \\ S &\rightarrow 1S \end{aligned}$$

To make a word, one starts with  $S$  and can use any of the substitution rules. We can quit when there are no more symbols from  $\mathcal{A}_N$  in the word. So, to get the empty string  $\lambda$  we use the first rule and stop. To generate the string 11, we say  $S \rightarrow 1S \rightarrow 11S \rightarrow 11\lambda \rightarrow 11$ . In this fashion, these three rules allow us make all possible words composed from the alphabet  $\mathcal{A}_T = \{0, 1\}$ .

**Example 2.2** Let us examine a more difficult example. Suppose we want to generate the language such that there are no two adjacent 0s, i.e., '00' is a forbidden word. Let us take  $\mathcal{A}_T = \{0, 1\}$  and  $\mathcal{A}_N = \{S\}$  as before. We can accomplish this with the following production rules

$$\begin{aligned} S &\rightarrow \lambda \\ S &\rightarrow 01S \\ S &\rightarrow 1S \end{aligned}$$

The second rule requires that we always follow a 0 with a 1. This language is called the golden mean language, and we will discuss it in more detail in a later chapter.

**Example 2.3** As a final example, let us consider the language such that between any two 0s there is an even number of 1s. Another way of saying this is that the set of forbidden words can be expressed as  $\mathcal{F}_L = \{01^{2k+1}0\}$  with  $k$  a non-negative integer. Let us once again take  $\mathcal{A}_T = \{0, 1\}$  and  $\mathcal{A}_N = \{S\}$  as before. We generate this language with the following production rules

$$\begin{aligned} S &\rightarrow \lambda \\ S &\rightarrow 11S \\ S &\rightarrow 0S \end{aligned}$$

The second rule insures that we always make 1s in pairs. This language is called the even language, and we will also discuss it in more detail in a later chapter.

We can now define a regular language:

**Definition 2.2** A grammar  $G = (V_N, V_T, S, F)$  is said to be regular if every rule in  $F$  has the form either  $A \rightarrow PB$  or  $A \rightarrow P$  where  $A, B \in V_N$  and  $P \in V_T^*$ .

$V_T^*$  is the set of all words over  $V_T$ . From the definition, we can see that the three examples of languages introduced above are all regular languages. Higher level languages are defined by relaxing the restrictions on word production.

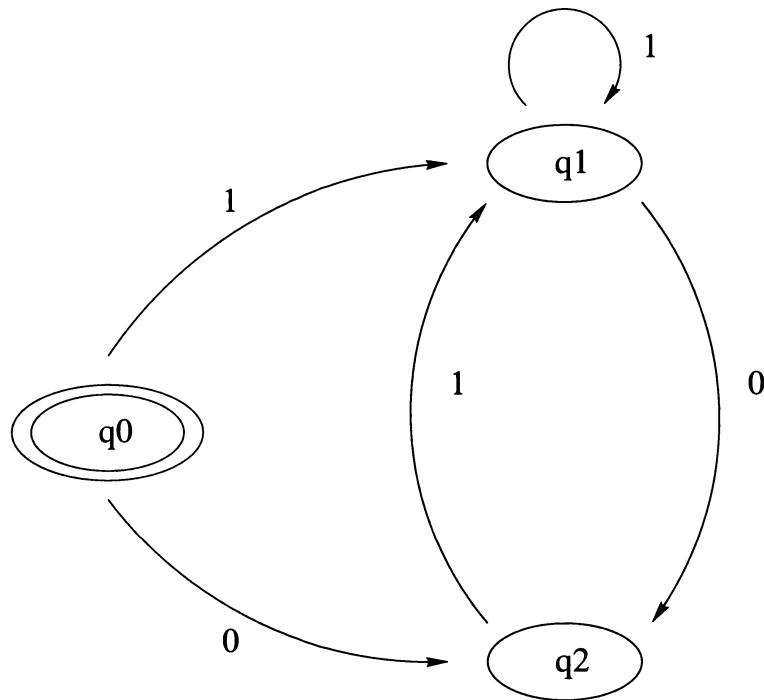


Figure 2.2: The finite automaton for the golden mean language. The golden mean language is the set of all words over  $\mathcal{A} \in \{0, 1\}$  that do not contain two consecutive 0s. This automaton clearly does not generate two consecutive 0s, as emission of a 0 on the transition from state q1 to state q2 is always followed by the emission of a 1 back to state q2.

accepted. Otherwise, we reject it and determine that the proposed string is not part of the language. If, when the string is entirely read in, the final state is an accepting state, then we say that the string is in the language. For our purposes, we can treat all states as accepting states. Figure 2.2 shows the finite automaton for the golden mean language. Let us give a formal definition to a finite automaton.

**Definition 2.3** We define a finite automaton  $M$  as an ordered quintuple  $(\mathcal{N}, \mathcal{A}, \delta, S, \mathcal{N}_F)$ , where  $\mathcal{N}$  is a finite set of states,  $\mathcal{A}$  is a finite alphabet,  $\delta$  is transition function,  $S$  is a start state, and  $\mathcal{N}_F$  is a set of accepting states,  $\mathcal{N}_F \subseteq \mathcal{N}$ .

The transition function  $\delta$  just tells us how one state evolves into another on emission of a symbol  $a$ . That is,  $\delta(\mathcal{N}_i, a)$  is a state for each state  $\mathcal{N}_i \subseteq \mathcal{N}$ , and symbol  $a$ .  $\delta(\mathcal{N}_i, a)$  may not be defined for all  $\mathcal{N}_i$  and  $a$ , in which case the transition is not allowed.

It is helpful to distinguish between different kinds of graphs. Let us call a finite automaton *deterministic* if the symbol emitted at each edge takes the system to a definite state. For example, suppose that the transition from q3 to q0 in figure 2.1 is labeled with a 0 instead of a 1. Then if we are at state q3, emission of a zero does not uniquely define the next state. Figure 2.1 would still be perfectly fine finite automaton, but we would call it *nondeterministic*. Determinism then here means something a little different than what a physicist might think. Determinism does not imply that each state has a unique successor state, but rather that each state on emission of a particular symbol has a unique successor state, provided such a transition is allowed. It can be shown that any nondeterministic finite automaton can be written as an equivalent deterministic finite automaton, although in general the deterministic version has exponentially more states.

There are several ways one can interpret equation 2.1. One is that  $H$  gives the average number of yes-no questions that one needs to ask to determine the value of  $X$ , if the questions are chosen optimally. Another is that we can interpret  $H$  as the expectation value of the surprise, where the surprise is defined as  $(-\log \Pr(x))$ . This makes some sense, in the following way. For very common events,  $\Pr(x)$  is not too far from one, so  $(-\log \Pr(x))$  is small indicating that we are not so ‘surprised’ when we see  $x$ . However, for rare events,  $\Pr(x)$  is small making  $(-\log \Pr(x))$  large, indicating we are ‘surprised’ to see  $x$ . To get the expectation value of the surprise, we multiply the surprise by the probability of seeing each event,  $\Pr(x)$ , and sum over events. This is just  $H$ . Of course we should not overlook the similarity of equation 2.1 to the thermodynamic entropy  $S$  written in terms of canonical probabilities [52] as

$$S = -k \sum_i P_i \ln P_i \quad (2.2)$$

with  $P_i = \frac{\exp^{-\beta E_i}}{Z}$  and  $\sum_i P_i = 1$ . ( $E_i$ ,  $Z$ ,  $\beta$ , and  $k$  have their usual thermodynamic meanings here.)

There are several other entropies defined in information theory, and we list them now. They involve two distributions, let us call them  $X$  and  $Y$ , and take values from the finite sets  $\mathcal{A}$  and  $\mathcal{B}$  respectively. As usual, we denote the variable with capital letters and the particular value it assumes in lower-case. Let us first fix some notation. We define the joint probability  $\Pr(x, y)$  to be the probability that  $X = x$  and  $Y = y$ . We define the conditional probability  $\Pr(x|y) \equiv \frac{\Pr(x, y)}{\Pr(y)}$ . With these definitions in place, we define the *joint entropy* of two variables as

$$H[X, Y] \equiv - \sum_{(x, y) \in \mathcal{A} \times \mathcal{B}} \Pr(x, y) \log \Pr(x, y). \quad (2.3)$$

We can also define the *conditional entropy* of one variable on another in terms of their joint entropy as

$$H[X|Y] \equiv H[X, Y] - H[Y]. \quad (2.4)$$

The interpretation of  $H[X|Y]$  is simple enough. It represents the uncertainty remaining in  $X$  once we know  $Y$ .

Finally, let us define *mutual information* between two random variables  $X$  and  $Y$ .

$$I[X; Y] \equiv \sum_{(x, y) \in \mathcal{A} \times \mathcal{B}} \Pr(x, y) \log \frac{\Pr(x, y)}{\Pr(x) \Pr(y)}. \quad (2.5)$$

We can interpret  $I$  as the reduction in the uncertainty in one variable due to knowledge of another. We note that  $I$  is symmetric in its arguments, as  $I[X; Y] = I[Y; X]$ .

## 2.2 Computational Mechanics

Let us now give a brief account of computational mechanics, aiming at providing an intuitive understanding rather than formal discourse. For all of the technical details, proofs, and mathematical rigour, as well as a more detailed and complete disquisition, the reader is referred to [16] [62] [61]. The basic paradigm of computational mechanics is stated easily enough. We assume an observer has access to a one-dimensional data stream (often called a measurement channel) that is produced by some system. The data can be discrete, or we can apply some discretization process to make it



We should mention a few properties of causal states. First, by the construction given, we see that they are *minimal*. By this we mean that they have no unnecessary structure in them. To eliminate any structure would be to throw away important predictive information. We could always add more structure, more complexity, but this would not improve the predictive power of the model. It would only mean that we are carrying around more information than we need. Secondly, also by construction, they are *maximally prescient*. Since we have kept all information that is shown to have predictive value, no other formulation can have more predictive power. The causal states have as much predictive power as the underlying process will allow. Lastly, causal states are *unique*. Up to a trivial relabeling of states, the causal states admit no reformulation.

We can now imagine that we have found the states as given in, say, figure 2.3. We have been observing the process for a very long time and let us say we know what causal state we are in. We observe another symbol. This new symbol becomes part of the history, and this new history must belong to some causal state we have listed. (This can be true in the limit of having seen some infinite past. Theorists allow such luxuries.) So we say that, upon seeing a new symbol, we make a transition to another causal state. That is, we can think of transitions connecting causal states on the emission of a symbol. This is reminiscent of a finite automaton, and in fact we can treat causal states connected in this way as a finite automaton. When we do so, we have an  $\epsilon$ -machine. Graphically, an  $\epsilon$ -machine looks just a finite automaton, except that there are conditional probabilities attached to the arcs.

Let us mention a few properties of  $\epsilon$ -machines. First it turns out that they are deterministic. There is a unique successor state for each causal state on the emission of an allowed symbol. The machine is also *Markovian*. That is, knowledge of the current causal state is sufficient for optimal prediction. We do not need to know the history of past states. Finally,  $\epsilon$ -machines are the maximally accurate predictors with the minimal statistical complexity (we define statistical complexity below). They are the best one can do, and they invoke the least complexity to do it. It is tempting to think of  $\epsilon$ -machines as stochastic versions of finite automata, and indeed there are resemblances. But stochastic finite automata need not be minimal, maximally prescient, or deterministic. We do not interpret the nodes to be causal states.

Additionally,  $\epsilon$ -machines have a unique start state, as do finite automata. Often, the start state is *transient*, that is, in the limit of an infinite string, transient states are causal states that are visited with probability zero. Of course, other states can be transient also. We differentiate these from *recurrent* states, which are those causal states that are visited with a probability greater than zero in the limit of an infinite string. All states in an  $\epsilon$ -machine are accepting states. Also, we require that the recurrent states be strongly connected, that is, colloquially put, the graph is not allow split into two separate pieces.

The goal of computational mechanics is to reconstruct the  $\epsilon$ -machine from data. Once done, the  $\epsilon$ -machine is a model of the system that produced the data, with all of the nice aforementioned properties. It tells us how information is stored and processed by the system. From the  $\epsilon$ -machine, we can calculate a number of quantities that describe these information storage and use features.

*The Source Entropy Rate:* Known variously as the thermodynamic entropy density and the metric entropy,  $h_\mu$  specifies the irreducible randomness produced by a source after correlations are accounted for. It has units of *bits/symbol*. For a completely random source  $h_\mu = 1$  bits/symbol, while completely predictable processes have an  $h_\mu = 0$  bits/symbol. We can define  $h_\mu$  as,

$$h_\mu \equiv \lim_{L \rightarrow \infty} \frac{H[S^L]}{L}. \quad (2.7)$$

There are several alternate ways to find  $h_\mu$ . We can take the limit in equation 2.7, or, we can say

$$h_\mu = H[S_1 | \overleftarrow{S}]. \quad (2.8)$$

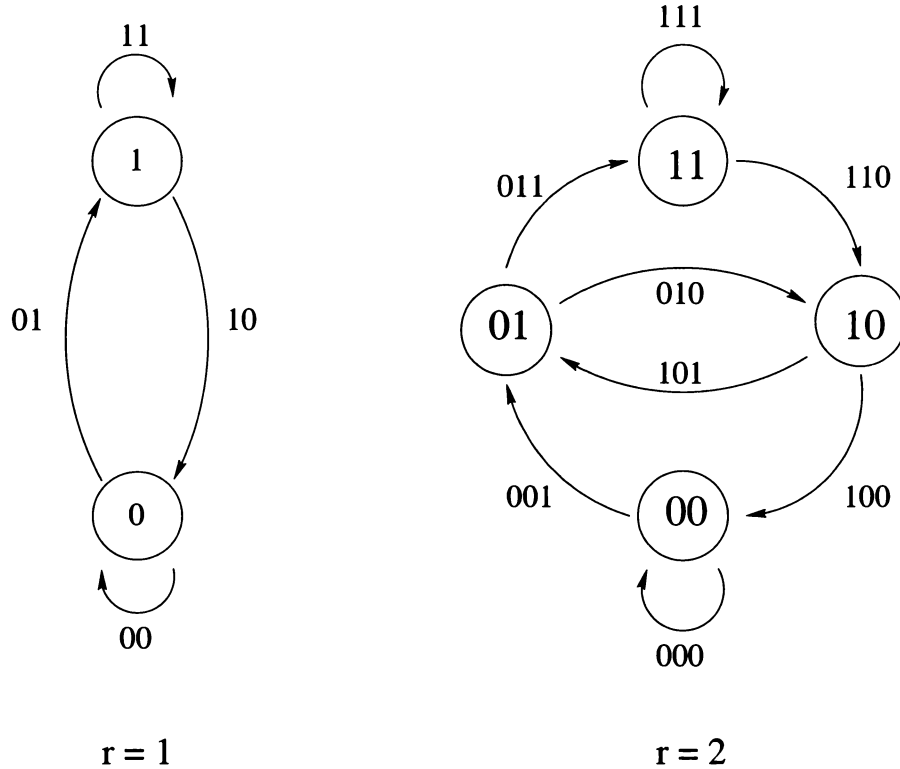


Figure 2.4: Examples of de Bruijn Graphs. The left most graph shows the  $r = 1$  de Bruijn graph. This graph explicitly postulates a memory of one, so we can label the nodes by the last symbol seen. Transitions between nodes are labeled with two symbols. The  $r = 2$  graph is shown on the left. Note that there are  $2^r$  nodes and  $2^{r+1}$  arcs.

*Transient Information:* We define  $\mathbf{T}$  as the transient information. Basically,  $\mathbf{T}$  is a measure of the amount of information an observer must extract from measurements to synchronize to a process. We can write the transient information as

$$\mathbf{T} = \sum_{L=1}^{\infty} L[h_{\mu}(L) - h_{\mu}]. \quad (2.17)$$

The units of  $\mathbf{T}$  are *bits*  $\times$  *symbol*.

Unless noted otherwise, we use equations 2.10, 2.12, 2.14, 2.16, and 2.17 to find the relevant computational quantities.

## 2.3 De Bruijn Graphs

De Bruijn graphs are **not** a part of computational mechanics, but they will form an important step in our method to reconstruct  $\epsilon$ -machines. So we shall give an introduction here.

Introduced to provide a useful representation for one-dimensional ground states, de Bruijn graphs are directed graphs with  $2^r$  nodes and  $2^{r+1}$  arcs connecting nodes. To construct a graph, we write down all possible sequences of  $r$  spins as nodes and then connect two nodes,  $\mathcal{N}_1$  and  $\mathcal{N}_2$ , with a

## Chapter 3

# Some Practical Details regarding Polytypes and Diffraction Patterns

Let us now abruptly switch gears and turn our attention to some experimental and notational matters concerning polytypes and diffraction patterns. In §3.1 we will discuss the various ways that the stacking sequences of close-packed structures are described; we will give a short account of the experiments we will analyze in §3.2; we will briefly provide some kinematic details in §3.3; we will show how to correct the experimental diffraction patterns for unwanted effects in §3.4; and finally in §3.5 we will list some assumptions necessary to analyze the data.

### 3.1 Notational Matters concerning Polytypes

There are many ways to express stacking sequences of closed-packed structures. These include Ramsdell notation, *ABC* notation, the Hägg or  $\Delta$ – $\nabla$  notation, Zhdanov notation, and the h-k notation [60]. To avoid confusion, we will minimize the notations we use to just few; but since our work is cross-disciplinary, it will be convenient to have several at our disposal. Let us list and discuss each.

*ABC notation.* An unambiguous way to specify the stacking sequence, the *ABC* notation is applicable for both ordered and disordered sequences. Simply put, this notation specifies the absolute position of each layer in the polytype. We will call the orientation of each layer its *spin* and since the *ABC* notation gives the absolute position for each layer, we will say that *A*, *B* or *C* gives the absolute spin of a particular layer.

*Hägg notation.* Due to stacking constraints, no two adjacent layers may have the same absolute spin, and we can take advantage of this by introducing relative spins. If the  $(n + 1)^{\text{th}}$ -layer is related to the  $(n)^{\text{th}}$ -layer by a clockwise rotation ( $A \rightarrow B \rightarrow C \rightarrow A$ ), Hägg denoted this interlayer spin by '+', and counterclockwise rotations ( $A \rightarrow C \rightarrow B \rightarrow A$ ) are labeled with a '-'. We will find this useful, except that we prefer to use '1' and '0' in place of '+' and '-'. Up to a trivial overall rotation of the crystal, the *ABC* notation and the Hägg notation are completely interchangeable.

*Ramsdell notation.* Applicable only to ordered structures, the Ramsdell notation is a convenient short-hand for particular polytypes. The format is  $xZ$ , where  $x$  is the repeat length along the stacking direction and  $Z$  specifies the the symmetry present, either hexagonal(H), rhombohedral(R), cubic(C), or unknown(L). Since we only discuss in detail three ordered structures, it seems easy enough to just list them. The 3C structure is ...*ABCABC*..., the 2H structure is ...*ABABAB*... and the 6H structure is ...*ABCACB*.... For longer period polytypes, this notation can be quite useful.

elastic scattering, we have  $|\mathbf{k}'| = |\mathbf{k}|$ . The de Broglie relation between the magnitude of the wave vector and the wave length is  $|\mathbf{k}| = 2\pi/\lambda$ . We write the change in the wave vector as

$$\Delta\mathbf{k} = \mathbf{k}' - \mathbf{k}. \quad (3.1)$$

We have the standard relation between the incident wave vector and the change in the wave vector, namely,

$$|\Delta\mathbf{k}| = 2|\mathbf{k}| \sin \theta. \quad (3.2)$$

where  $\theta$  is the angle of incidence. We take the standard hexagonal net [70] in the plane of the modular layers, with the usual primitive translation vectors,  $\mathbf{a}$  and  $\mathbf{b}$  and let  $a = |\mathbf{a}| = |\mathbf{b}|$ . We define  $\mathbf{c}$  to be perpendicular to the hexagonal net and take the magnitude of  $|\mathbf{c}| = c$  to be the spacing between modular layers. Let us then write the 'reciprocal lattice vector' as

$$\Delta\mathbf{k} = \mathbf{G} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^* \quad (3.3)$$

where  $\mathbf{a}^*$ ,  $\mathbf{b}^*$ , and  $\mathbf{c}^*$  are primitive translation vectors of the reciprocal lattice. Clearly we do not have a reciprocal lattice any more than we have a lattice structure in real space, but, as in real space, there is crystallinity in the 'reciprocal modular layers.' So, in reciprocal space, as in real space, the disorder is confined to the stacking direction. Since we are interested in the diffraction pattern along the  $10.l$  row we have

$$\Delta\mathbf{k} = \mathbf{G} = \mathbf{a}^* + l\mathbf{c}^*. \quad (3.4)$$

We find the magnitude of  $\Delta\mathbf{k}$  to be

$$|\Delta\mathbf{k}| = \frac{2\pi}{c} \left[ \frac{4}{3} \left( \frac{c}{a} \right)^2 + l^2 \right]^{\frac{1}{2}}. \quad (3.5)$$

### 3.4 How to fix Diffraction Patterns

We want to extract information about the two layer correlation functions from the diffraction pattern, but there are other factors which affect the measured intensity and for which we must account [4] [5] [9] [28] [48] [78]. Let us call the *uncorrected* diffraction pattern  $\mathcal{I}(l, Q)$  the raw data as reported by M.T. Sebastian and P. Krishna [60].  $Q$  represents the correlation functions which we are trying to find. We can write intensity of the uncorrected diffraction pattern as

$$\mathcal{I}(l, Q) = \mathcal{C}(l) \times I(l, Q). \quad (3.6)$$

where  $\mathcal{C}(l)$  are diffraction effects not dependent on the correlation functions, and  $I(l, Q)$  represents those factors which depend on the correlation functions. Let us call  $\mathcal{C}(l)$  *correction* factors and  $I(l, Q)$  the *corrected* diffraction pattern. We will detail the dependence of the diffraction pattern on the correlation functions in the next chapter. For now, let us concentrate on the correction factors. We have considered several factors, and we list them now.

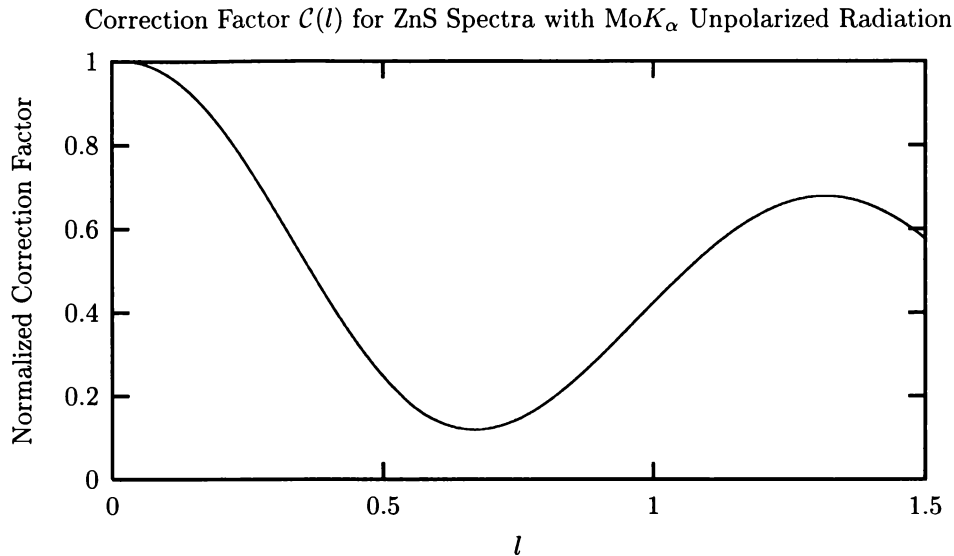


Figure 3.1: Correction Factor  $\mathcal{C}(l)$  for ZnS diffraction patterns assuming unpolarized radiation of wavelength  $\lambda = 0.7107 \text{ \AA}$ . The corrections included are for atomic form factors, structure factors, anomalous scattering factors and the polarization factor.

### 3.5 Some Assumptions

To make the analysis tractable, we employ the following assumptions.

*Perfect Modular Layers.* We assume the modular layers themselves are undefected. That is, each layer is crystalline in the strict sense, with no point defects, impurities, or distortions in the two dimensional lattice structure.

*Scattering Power same for all Layers.* We assume that each layer diffracts x-rays with the same intensity. There is no reason to believe that this is not so, unless absorption effects are important or the geometry of the crystal is such that each layer does not have the same cross-sectional area.

*Constant spacing between Modular Layers.* The spacing between layers can change slightly between polytypes, but this is known to be quite small [57], perhaps about 0.3% between the 2C and the 3H modifications of ZnS. We therefore treat the separation between different modular layers as constant, regardless of the local stacking arrangement.

*Entire Layer is Shifted with respect to Neighbors.* We assume that the entire modular layer is shifted with respect to its neighbors. This is reasonable, since the stacking fault energy is quite small. Another way to say this, is that we assume that the stacking faults extend all through the crystal.

*Spherical Atoms.* To calculate the atomic form factors, we make the assumption of neutral, spherical atoms. The bonding in ZnS is at least partially covalent, so we know this is not completely realistic. But since only two of the electrons in a ZnS pair are directly involved in the bonding, the other forty-four being more or less bound by their respective nuclei, we do not consider this to be a large source of error.

*Stationary Process.* The requirement of stationarity is necessitated by our use of computational mechanics. Simply put, a stationary process is one in which the probability of finding a word or

## Chapter 4

# All about Two-Layer Correlation Functions and Diffraction Patterns

Now that we have corrected the diffraction pattern for the form factor, the polarization factor, etc., we can now examine that part of the diffraction pattern which depends only on the the number of layers and their arrangement. Let us visualize the picture we have. There are  $N$  hexagonal, close-packed layers, with each layer occupying one of three positions,  $A, B$  or  $C$ . These layers are then stacked, and we wish to find the relationship between the stacking order and the diffraction pattern. Let us introduce three quantities,  $Q_c(n)$ ,  $Q_r(n)$ , and  $Q_s(n)$ , which we will call the two-layer correlation functions. Let us take  $c, r, s$  to stand for cyclic, reverse, and same respectively.  $Q_c(n)$  is defined as the probability that any two layers at a separation of  $n$  are cyclically related. By cyclic, we mean that if the  $i^{\text{th}}$  layer is, say  $A$ , then the  $(i+n)^{\text{th}}$  layer is  $B$ .  $Q_r(n)$  and  $Q_s(n)$  are defined in a similar fashion. Since these are probabilities,  $0 \leq Q_\alpha(n) \leq 1$ , where  $\alpha \in \{c, r, s\}$ . It is clear that

$$Q_c(n) + Q_r(n) + Q_s(n) = 1 \quad \forall n. \quad (4.1)$$

With these definitions, we can write an expression for the diffraction pattern, see Yi and Canright [79],

$$I(l) = \frac{I_0}{N} \left\{ \frac{\sin^2(\pi Nl)}{\sin^2(\pi l)} - 2\sqrt{3} \sum_{n=1}^N \left[ (N-n) \left[ Q_c(n) \cos\left(2\pi nl + \frac{\pi}{6}\right) + Q_r(n) \cos\left(2\pi nl - \frac{\pi}{6}\right) \right] \right] \right\}. \quad (4.2)$$

This expression for the diffraction pattern is only valid for a stacked sequence of two dimensional hexagonal layers. It is easy to see that  $I(l)$  is periodic in  $l$  with period one, so we need only examine  $l$  over the unit interval.

It is instructive to rewrite equation 4.2 by expanding the cosine terms. Doing so, we get,

$$I(l) = \frac{I_0}{N} \left\{ \frac{\sin^2(\pi Nl)}{\sin^2(\pi l)} - 3 \sum_{n=1}^N \left\{ (N-n) \left[ (Q_c(n) + Q_r(n)) \cos(2\pi nl) + \frac{1}{\sqrt{3}} (-Q_c(n) + Q_r(n)) \sin(2\pi nl) \right] \right\} \right\}. \quad (4.3)$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \oint \frac{\sin^2 N\pi l}{\sin^2 \pi l} \cos(2\pi n l) dl = 1$$

and

$$\lim_{N \rightarrow \infty} \frac{1}{N} \oint \frac{\sin^2 N\pi l}{\sin^2 \pi l} \sin(2\pi n l) dl = 0.$$

Carrying out the integrations in equations 4.7 and 4.8, we find

$$A_n = \frac{3I_0}{2} \left[ \frac{2}{3} - Q_c(n) - Q_r(n) \right] \quad (4.9)$$

and

$$B_n = \frac{3I_0}{2\sqrt{3}} \left[ Q_c(n) - Q_r(n) \right]. \quad (4.10)$$

We can solve these for  $Q_c(n)$  and  $Q_r(n)$  to get,

$$Q_c(n) = \frac{1}{3} - \frac{1}{3I_0} \left[ A_n - \sqrt{3}B_n \right] \quad (4.11)$$

and

$$Q_r(n) = \frac{1}{3} - \frac{1}{3I_0} \left[ A_n + \sqrt{3}B_n \right]. \quad (4.12)$$

Finally, we need to find an expression for  $I_0$ . We can do this by integrating over the diffraction pattern. We find

$$I_0 = \oint I(l) dl. \quad (4.13)$$

## 4.2 Figures of Merit

We can exploit the fact that, due to stacking constraints,  $Q_c(1) + Q_r(1) = 1$ . Adding equations 4.11 and 4.12 gives

$$A_1 = -\frac{I_0}{2}. \quad (4.14)$$

This is always true for error-free data, but we can use this as measure of how corrupted the data is over a particular interval. So let us define

$$\gamma = \frac{\oint I(l) \cos(2\pi l) dl}{\oint I(l) dl}. \quad (4.15)$$

We ignore terms that do not scale with  $N$ , as they do not contribute in the many layer limit. We note that there must be some term which grows at least linearly in  $N$ , so that the integrated intensity will likewise grow linearly with  $N$ . Additionally, no term grows faster than quadratically. It is then helpful to understand under what conditions the summation gives rise to quadratic terms and under what conditions the quadratic terms may cancel. It is likely that experiment will only see structure proportional to the highest power of  $N$ . Let us then treat three cases, two of which are of empirical relevance. The first is the case where the  $Q$ s decay to some asymptotic value, and the second is the case where they approach some limit cycle of period  $T_q$ . As a final case, we mention those processes in which the  $Q$ s neither approach an asymptotic limiting value or become cyclic.

### 4.3.1 $Q$ s decay to an asymptotic value of $\frac{1}{3}$

There seem to be many processes, mostly disordered ones, that can lead to  $Q$ s that decay to some limiting value. We prove in the appendix that if the correlation functions assume a constant asymptotic value, that value must be  $\frac{1}{3}$ . So let us specialize to cases where the  $Q$ s reach an asymptotic value of  $\frac{1}{3}$  at some  $n = n_c$ . We will provide a more precise definition of  $n_c$  in a later section. We assume that  $n_c \ll N$ . Typical values of  $n_c$  are less than one-hundred.

Let us treat the special case of  $l = 0$ . The first term in equation 4.5 goes as  $N^2$  and the sine term in the summation vanishes. The argument of the cosine term vanishes, giving a value of one. So we can write the  $N$ -layer diffraction intensity as

$$I'(0) = N^2 - 3 \sum_{n=1}^N (N - n) [(Q_c(n) + Q_r(n))]. \quad (4.17)$$

Since we are interested in how  $I'(0)$  scales with  $N$ , it is permissible to replace the  $Q$ s by their asymptotic values. Carrying out the summation and using the notation  $Q_c$  and  $Q_r$  for the asymptotic value of  $Q_c(n)$  and  $Q_r(n)$  for  $n$  large, we get to order  $N^2$

$$I'(0) = N^2 \left[ 1 - \frac{3}{2} (Q_c + Q_r) \right]. \quad (4.18)$$

So when the asymptotic values of  $Q_c(n)$  and  $Q_r(n)$  are  $\frac{1}{3}$ , there will be no Bragg peak at  $l = 0$ . Therefore knowing then the asymptotic behavior of the  $Q$ s tells us the whether we see a Bragg peak at integer  $l$ .

We can also make some statements about the possibility of Bragg peaks for non-integer  $l$ . These peaks must come from the summation term in equation 4.5. Let us again restrict our attention to the many layer limit and examine the case of  $l \neq$  integer. We can now approximately evaluate equation 4.5. The first term is of order one, so we neglect it. The sine term in the summation again cancels. We are then left with the cosine term. Setting  $Q_c(n)$  and  $Q_r(n)$  to their asymptotic values, we get,

$$\begin{aligned} I'(l \neq \text{integer}) &\approx -3 \sum_{n=1}^N (N - n) \left[ \frac{1}{3} + \frac{1}{3} \right] \cos(2\pi n l) \\ &\approx -3 \left( \frac{2}{3} \right) \left( -\frac{N}{2} \right) \\ &\approx N. \end{aligned} \quad (4.19)$$

So, if the  $Q$ s approach an asymptotic value of  $\frac{1}{3}$ , we see that  $I'(l) \propto N$  for all  $l$ , showing that there are no Bragg peaks in the spectrum at all. This then implies that it is necessary to have  $Q$ s that



where  $B$  is some constant. We then find  $\lambda_q$  from the slope of the line of  $\log_2 \Psi_q(n)$  vs.  $n$ . Due to the finite size of our sample, we find that we should only use those  $\log_2 \Psi_q(n) \geq -5.5$ . The statistics are no longer very reliable for  $\Psi_q(n)$  smaller than this.

*Period of Limit Cycle:* Applicable only in the case of periodic oscillation, this quantity gives the length in terms of  $n$ , that we must go for the  $Q$ s to complete an oscillation. We call this quantity  $T_q$ .

*Cutoff Length:* While not really an intrinsic measure of the  $Q$ s, this quantity, which we give the symbol  $n_c$ , is the highest  $n$  for which we use the calculated values of the  $Q_\alpha(n)$  in finding the diffraction pattern. We discuss and motivate a definition for this quantity in a later section.

*Scattering Type:* We classify diffraction patterns according to whether they exhibit Bragg, pure continuous, or singular continuous scattering. In the event of mixed scattering, we can determine the amount of energy diffracted into each type.

## Chapter 5

# Examples of Processes, Correlation Functions and their Diffraction Patterns

In this chapter, we attempt to expound on and flesh out the relationships between processes which generate binary sequences, the two-layer correlation functions they imply and the resulting diffraction patterns. All of the systems we consider can be represented as a language, or their probabilistic versions as finite  $\epsilon$ -machines, except the Thue-Morse process. We undertake this exercise to gain intuition into the interplay of these three. Ultimately, our task is to begin with a diffraction pattern and infer the underlying process. That is we want to go  $\text{diffraction pattern} \Rightarrow Qs \Rightarrow \text{process } (\epsilon\text{-machine})$ . This procedure, which we can think of as a ‘backwards procedure’, is composed of two parts. First one finds the  $Qs$  from diffraction data and second one finds the  $\epsilon$ -machine which can generate these  $Qs$ . The first is easy, while the second is more difficult. In contrast, the ‘forward procedure’,  $\text{process} \Rightarrow Qs \Rightarrow \text{diffraction pattern}$ , is easy and well defined. It is this forward procedure we address in this chapter. We then will consider about a dozen or so elementary processes and find the resulting correlation functions and diffraction patterns. These sample processes are chosen to give a natural cross-section of possible cases we might encounter in nature, as well as some more exotic cases of theoretical interest. They are, the fair coin toss, the 3C, the 2H, the golden mean process, the 4H, the 3C/2H/4H process, the 6H, the 3C/2H/6H process, a period 14  $D$ -pair, the noisy period two, the even system, the sum zero process, and the Thue-Morse sequence. Many have the convenient property of only requiring at most two recurrent causal states. Except for the last four, all are expressible as SFT. The Thue-Morse sequence can not be written as a finite state machine, but it does give us a chance to examine a pattern that implies an infinite correlation between symbols and has a zero entropy density, while being aperiodic. We will calculate a number of properties for these systems, so that we may compare them.

For all of the following examples, unless stated otherwise, we take a sample of the process 10,000 characters in length to find the diffraction pattern using equation 4.3. All diffraction patterns are normalized to one over a unit interval. To find the correlation functions, we calculate them directly from a sample of the process 400,000 characters long. We take such a large sample to minimize the statistical fluctuations inherent in using a pseudo-random number generator. For completely predictable processes, where we can find the correlation functions analytically, we do so.

It is also perhaps worthwhile to make a comment on terminology. We have repeatedly used the term ‘process’ when referring to a spacial pattern or arrangement of 0s and 1s. This may indicate that there are some dynamics going on. Computational mechanics has roots in dynamical systems theory, where indeed this is a reasonable implication. For our work, we have used this term

Table 5.1: The first few values of the correlation functions for the fair coin toss. We see that even though this is a random process, the correlation functions do not assume their asymptotic value immediately. This is due to the stacking constraints for close-packed layers. We can see that they approach an asymptotic value of  $\frac{1}{3}$ .

$n$	$Q_c(n)$	$Q_r(n)$	$Q_s(n)$
1	1/2	1/2	0
2	1/4	1/4	1/2
3	3/8	3/8	1/4
4	5/16	5/16	3/8
5	11/32	11/32	5/16
6	21/64	21/64	11/32

are no Bragg peaks in the spectrum, there is a broad diffuse concentration of scattered intensity centered at  $l = 0.5$ . The intensity at  $l = 0.5$  is nine times that at  $l = 0$ . We also note that the diffraction pattern is symmetric about  $l = 0.5$ , as we expect from the equality  $Q_c(n) = Q_r(n)$ .

We can also consider the artificial case of a completely disordered sequence that is not required to obey any stacking constraints. This is, we allow  $..AA..$ , etc. When we do so, we find that  $Q_\alpha(n) = \frac{1}{3}$  for all  $n$  and  $\alpha$ , where  $\alpha \in \{c, r, s\}$ . The diffraction pattern is just a flat line, that is  $I(l) = 1$ . So it is the stacking constraints that impose what feature we see in Figure 5.4.

We now make some computational remarks. It is found that the  $Q$ s do not settle down to their asymptotic value for large  $n$ . They seem to oscillate, and ‘fidget’ about. We might expect some noise superimposed on the asymptotic values of the  $Q$ s due to the finite size of the sample used to determine them. These fluctuations should be on the order of  $1/\sqrt{N_s}$  where  $N_s$  is the length of the symbol sequence used. For our case,  $N_s = 400,000$ , giving fluctuations of about 0.002 in the correlation functions. These fluctuations can inflict havoc in the calculated diffraction pattern. It is not hard to see why. Failure of the correlation functions to reach their asymptotic values exactly will cause the summation term in equation 4.3 to become ‘unbalanced’. The sum has the potential to be proportional to  $N^2$ , if a particular value of  $l$  favors  $Q$ s slightly higher/lower than  $\frac{1}{3}$ . Even a small bias in the sum can create terms that alter the ‘correct’ intensity noticeably. We show a diffraction pattern for the fair coin toss where the  $Q$ s have not been forced to their asymptotic value in Figure 5.5. In all subsequent diffraction patterns, where it is clear that the  $Q$ s tend to an asymptotic value, we impose that asymptotic value of  $\frac{1}{3}$  on all  $Q_\alpha(n)$  such that  $n \geq n_c$ . We define  $n_c$  as the smallest  $n$  for which  $|Q_\alpha(n) - \frac{1}{3}| \leq \delta$ . We take  $\delta$  to be approximately 0.002. The exact value may change slightly with each diffraction pattern, and is somewhat subjective. We also note that not only do our  $Q$ s show some noise, but they also seem to have a low frequency oscillation imposed on them. We believe that this is unphysical and somehow a by-product in our random number generator.

Previous researchers have studied the diffraction pattern for randomly stacked close-packed structures. An expression for  $I(l)$  in closed form is given by Guinier [33] and agrees with our Figure 5.4.

We make a few remarks concerning some properties of the fair coin toss. It is known to have a statistical complexity and an excess entropy of  $C_\mu = \mathbf{E} = 0$  bits and an entropy rate of  $h_\mu = 1$  bits/symbol. The total predictability is  $\mathbf{G} = 0$  bits/symbol. The transient information is  $\mathbf{T} = 0$  bit symbols, implying the an observer need extract no information to synchronize with the source. The  $Q$ s display asymptotic decay and have a cutoff length of approximately  $n_c = 10$ . We can calculate the correlation length exactly from equations 5.2, and obtain a value of 1. We can also determine  $\lambda_q$  from our computed  $Q$ s. Using the first six values of  $\Psi_q(n)$ , we find a correlation length of  $\lambda_q = 0.994 \pm 0.004$ . We see that our measure for the correlation length for the  $Q$ s is not zero, as one might expect for a random process. Indeed, the graph 5.4 implies a memoryless process. It does

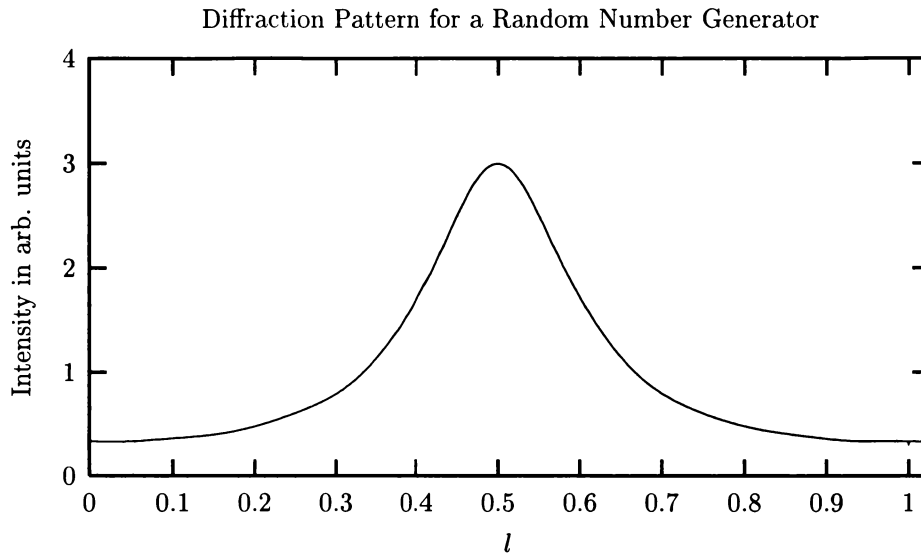


Figure 5.4: Diffraction pattern for a randomly stacked two dimensional hexagonal lattice using asymptotic values for the  $Q_s$ . We force the correlation functions to their asymptotic value for  $n \geq n_c$ . Doing so, we get a smooth diffraction pattern.

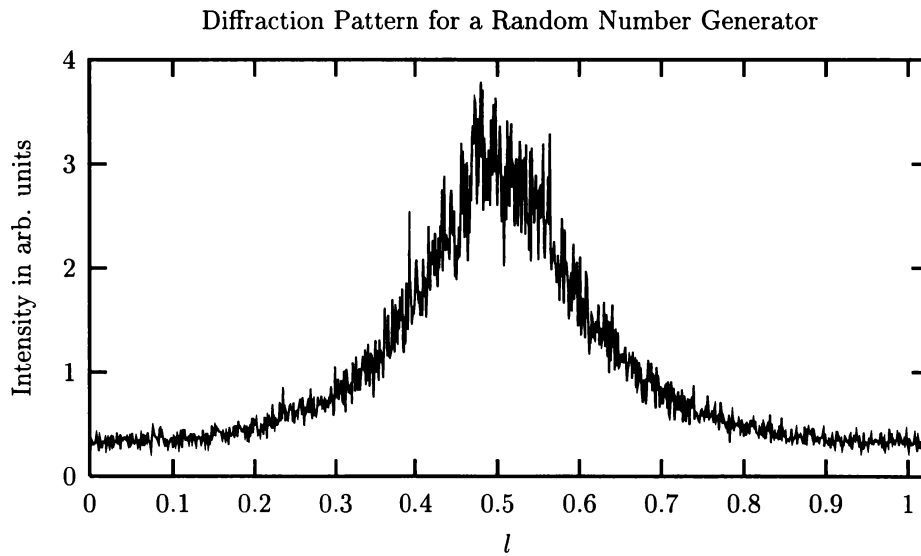


Figure 5.5: Diffraction pattern for a randomly stacked two dimensional hexagonal lattice using sequence calculated values for the  $Q_s$ . The diffraction pattern is not smooth, and exhibits what appear to be random fluctuations. These fluctuations depend on the length of the sequence used to find the  $Q_s$ . The longer the sequence, the smaller the fluctuations.

Table 5.2: The first few values of the correlation functions for the period one process. We see that correlation functions are periodic in  $n$  with a period of three.

$n$	$Q_c(n)$	$Q_r(n)$	$Q_s(n)$
1	1	0	0
2	0	1	0
3	0	0	1
4	1	0	0
5	0	1	0
6	0	0	1

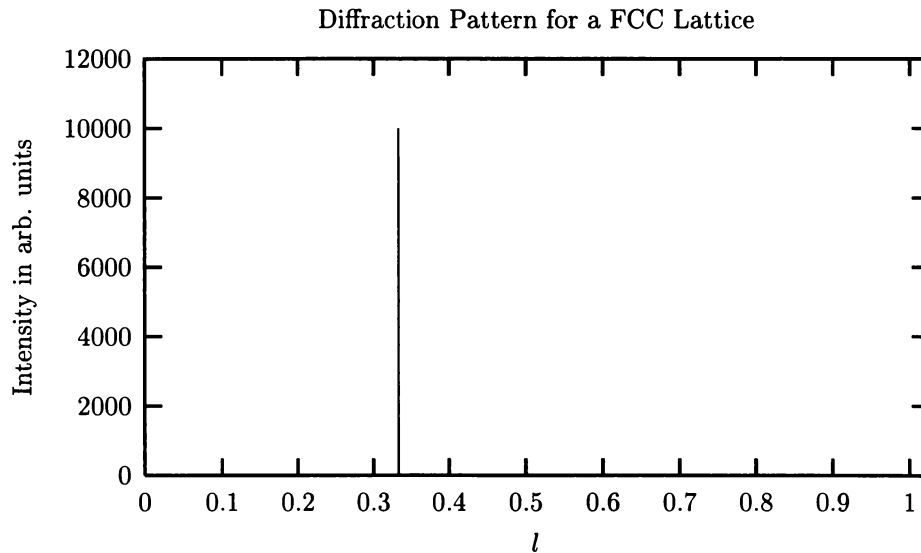


Figure 5.7: Diffraction pattern for the period one or 3C process. We observe a single Bragg peak at  $l = \frac{1}{3}$ . Everywhere else, including integer  $l$  the diffracted intensity is zero.

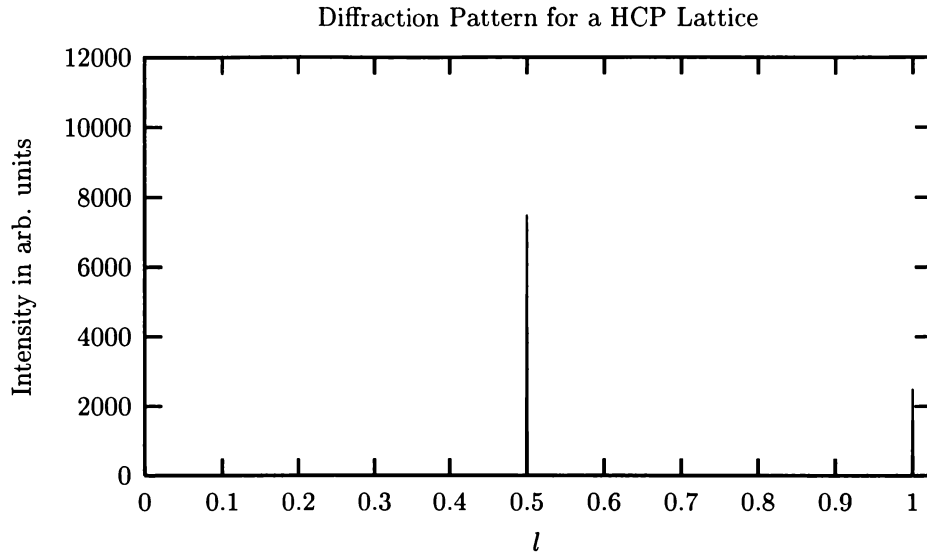


Figure 5.9: Diffraction pattern for a lattice stacked according to the period two rule. We see two Bragg peaks, one at  $l = \frac{1}{2}$  and one at  $l = 1$ .

## 5.4 The Golden Mean Process

We next consider the diffraction pattern generated by a lattice stacked according to the golden mean process. Stated simply, the golden mean system allows all sequences that do not contain two consecutive zeros, ie, '00' is an irreducible forbidden word (IFW). We call this the golden mean system because the logarithm of the total number of allowed sequences grows with the word length  $L$  at a rate given by the logarithm of the golden mean,  $\phi = \frac{1}{2}(1 + \sqrt{5})$ . See Crutchfield and Feldman [16]. We treat the probabilistic version of the system here. We allow 0s and 1s with equal probability, except when the previous symbol is a 0. We insist that 0 always be followed by a 1. The recurrent portion of the  $\epsilon$ -machine for the golden mean process is given in Figure 5.10. The regular expression for the corresponding language is  $\mathcal{R} = (1 + 01)^*$ . The first few  $Q$ s generated by this process are shown in Table 5.4. We note immediately that the  $Q_c(n) \neq Q_r(n)$ , suggesting that the diffraction pattern will not have reflection symmetry about  $l = \frac{1}{2}$ . We also see that  $Q_c(2) = 0$ . It can be demonstrated that this is a consequence of '00' being a forbidden word. An illustration of the  $\log_2 \Psi_q(n)$  versus  $n$  is shown in figure 5.12. One is tempted to think of the golden mean language as some sort of mixture between the period one and the period two languages. Somehow they are competing. We might expect then that the  $Q$ s are a compromise of the two. This interpretation seems validated when we examine the first few  $Q$ s. For  $n = 1, 2$ , indeed the  $Q$ s for the golden mean lie between those of the period one and period two. At  $n = 3$ , we see the first departure from this trend. For the golden mean,  $Q_c(3)$  is greater than that of either of the other two. The  $Q_r(4)$  for both the period one and the period two are zero, while the golden mean gives it a value of 0.500. So we conclude that the golden mean is not some sort of compromise between the two, at least not in such a simple-minded way.

When we examine the  $Q_s(n)$  for the golden mean process as shown in Figure 5.11, we see another interesting feature. There appear to be two frequencies superimposed on each other. Let us define the frequency in a natural way as the number of oscillations in  $Q_s(n)$  per  $n$ . We might expect then that there is a competition amongst frequencies, as the period one process has a natural frequency of one-third, while the period two has a natural frequency of one-half. We recall that the stacking

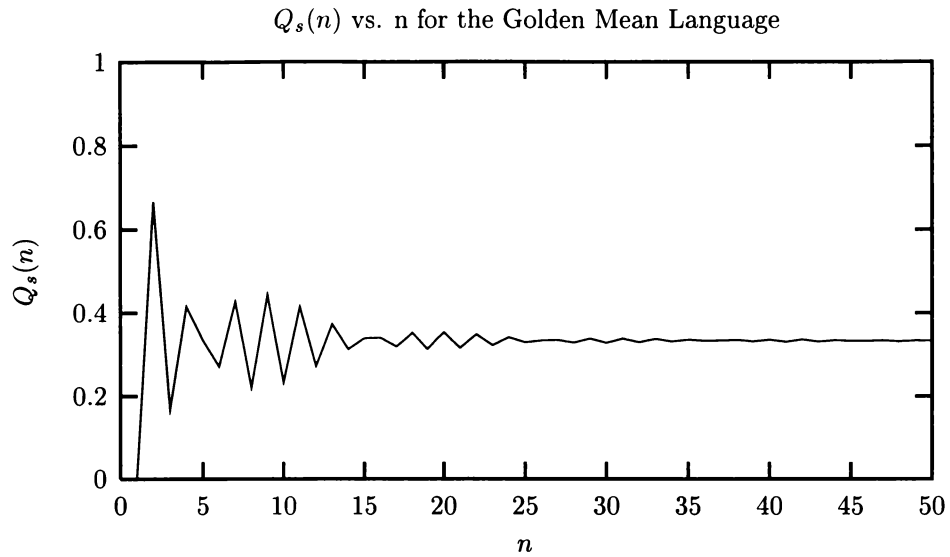


Figure 5.11: The  $Q_s(n)$  for the golden mean process a function of  $n$ . The correlation functions for the golden mean process decay to  $\frac{1}{3}$  with a correlation length of  $\lambda_q = 4.48 \pm 0.06$ .

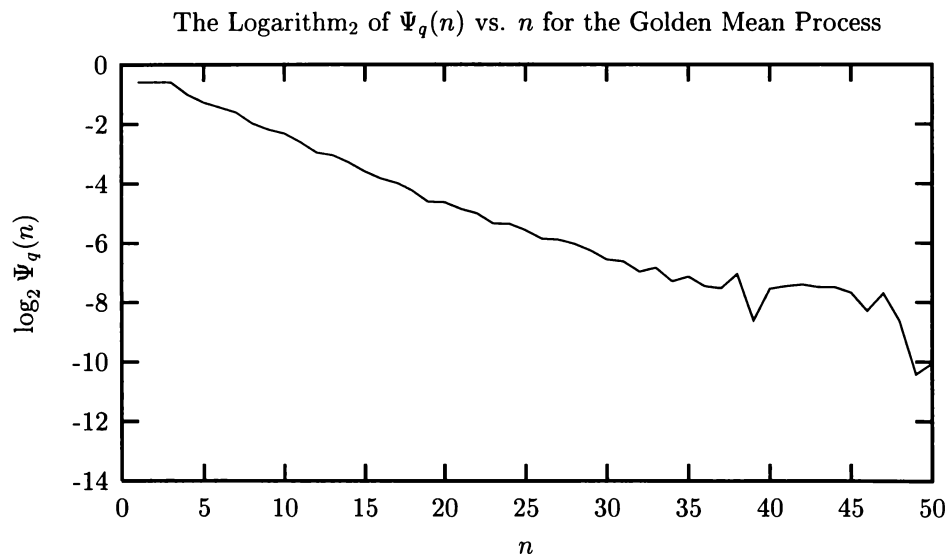


Figure 5.12: The logarithm of  $\Psi_q(n)$  for the golden mean process as a function of  $n$ . Using the first twenty-five values of  $\Psi_q(n)$ , we get  $\lambda_q = 4.48 \pm 0.06$ .

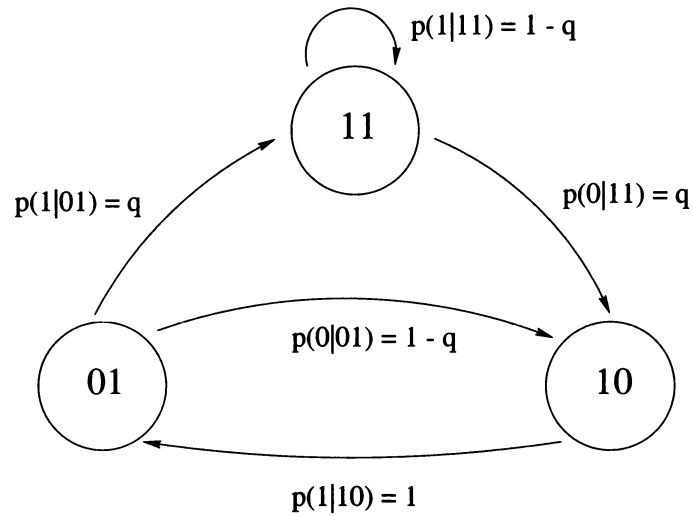


Figure 5.14: The recurrent portion of the  $\epsilon$ -machine that generates the 3C/2H process. This process has memory of two, so we can represent it with a portion of an  $r = 2$  de Bruijn graph. The causal states are labeled by the last two symbols seen.

Table 5.5: The first few  $Q$ s generated by the 3C/2H process, with  $q = 0.01$ . As for the golden mean process, a vanishing value of  $Q_c(2)$  implies that '00' is a forbidden word, as can be seen from the  $\epsilon$ -machine.

$n$	$Q_c(n)$	$Q_r(n)$	$Q_s(n)$
1	0.668	0.332	0
2	0	0.336	0.664
3	0.339	0.329	0.332
4	0.329	0.013	0.658
5	0.336	0.651	0.013
6	0.013	0.013	0.973



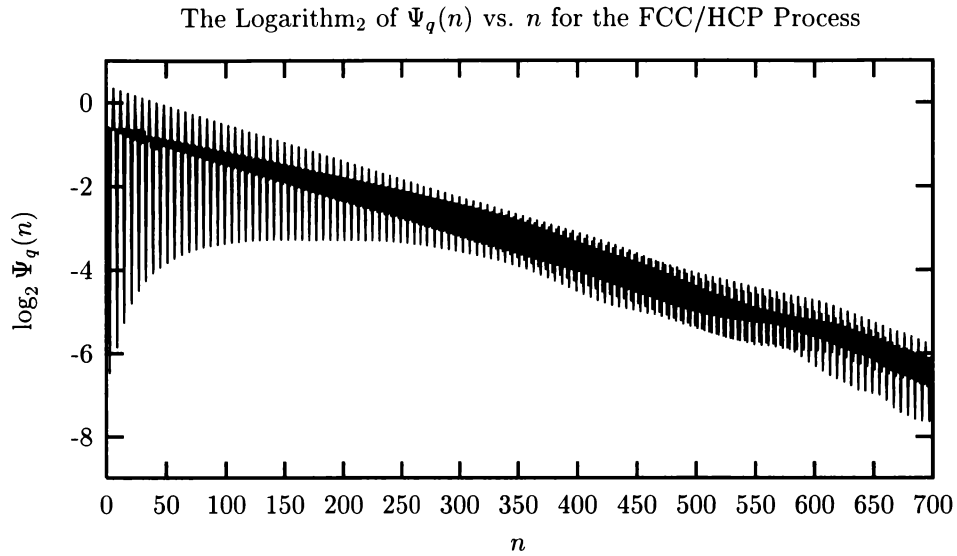


Figure 5.16: The logarithm of  $\Psi_q(n)$  for the 3C/2H process as a function of  $n$ , with  $q = 0.01$ . We find a correlation length of  $123 \pm 2$ .

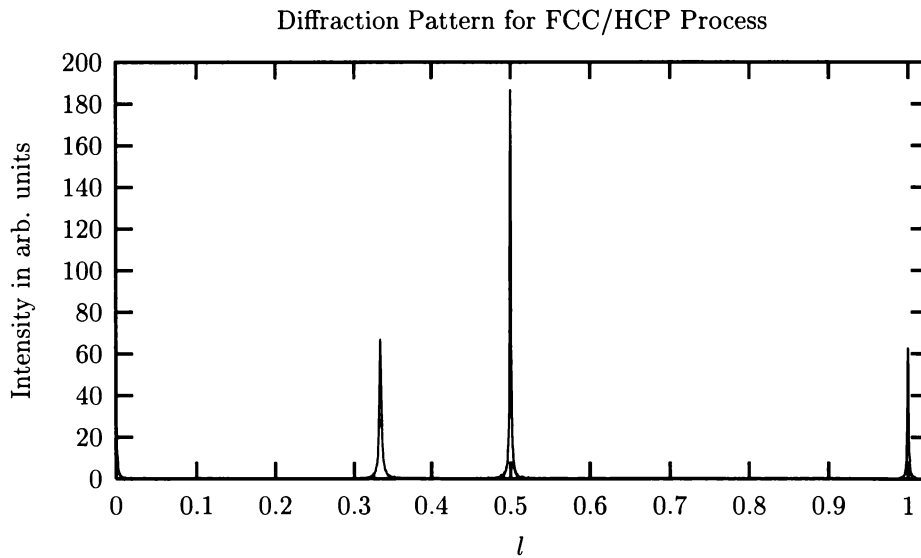


Figure 5.17: The diffraction pattern for a lattice stacked according to the 3C/2H process with  $q = 0.01$ . We see three sharp peaks in the spectrum, one at  $l = \frac{1}{3}$  corresponding to the 3C structure and the other two at  $l = \frac{1}{2}$  and 1 corresponding to the 2H structure.

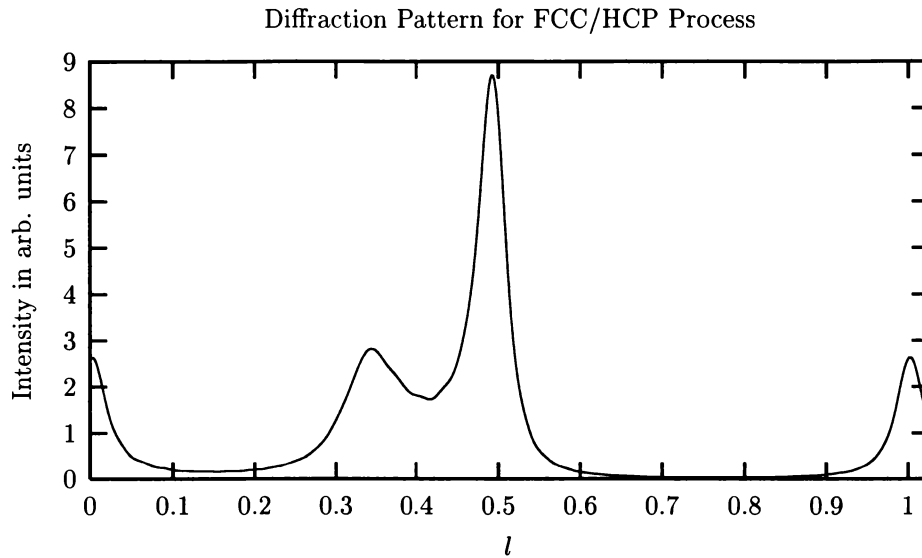


Figure 5.20: The diffraction pattern for a lattice stacked according to the 3C/2H process with  $q = 0.20$ . There is no longer clear separation between the two peaks at  $l = \frac{1}{3}$  and  $\frac{1}{2}$ . We also see the peaks beginning to shift.

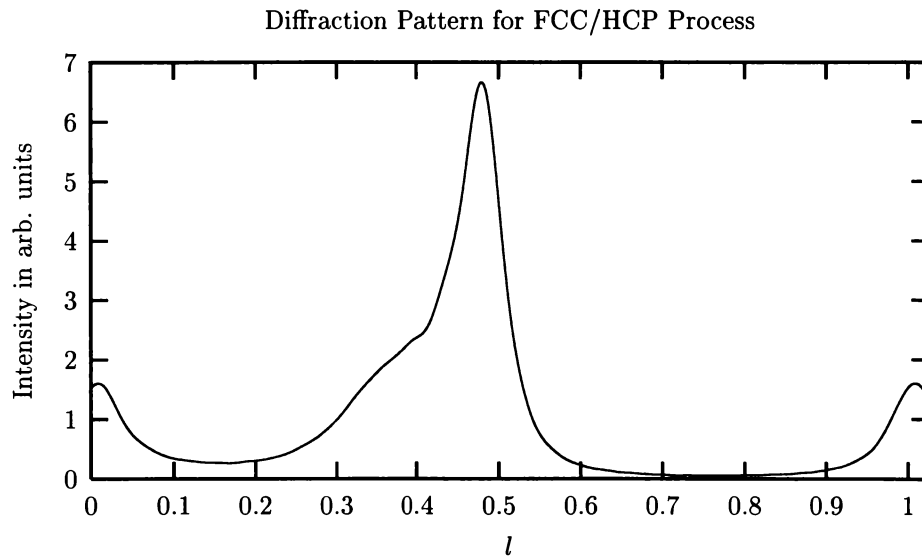


Figure 5.21: The diffraction pattern for a lattice stacked according to the 3C/2H process with  $q = 0.30$ . The peak corresponding to the 3C has almost completely disappeared, being absorbed into the peak at  $l \approx \frac{1}{2}$ , which has shifted noticeably to the left.

Table 5.6: The first few  $Q$ s generated by the 4H process. The correlation functions are periodic in  $n$  with period two. The correlation length is infinite.

$n$	$Q_c(n)$	$Q_r(n)$	$Q_s(n)$
1	0.500	0.500	0
2	0.250	0.250	0.500
3	0.500	0.500	0
4	0	0	1.000
5	0.500	0.500	0
6	0.250	0.250	0.500

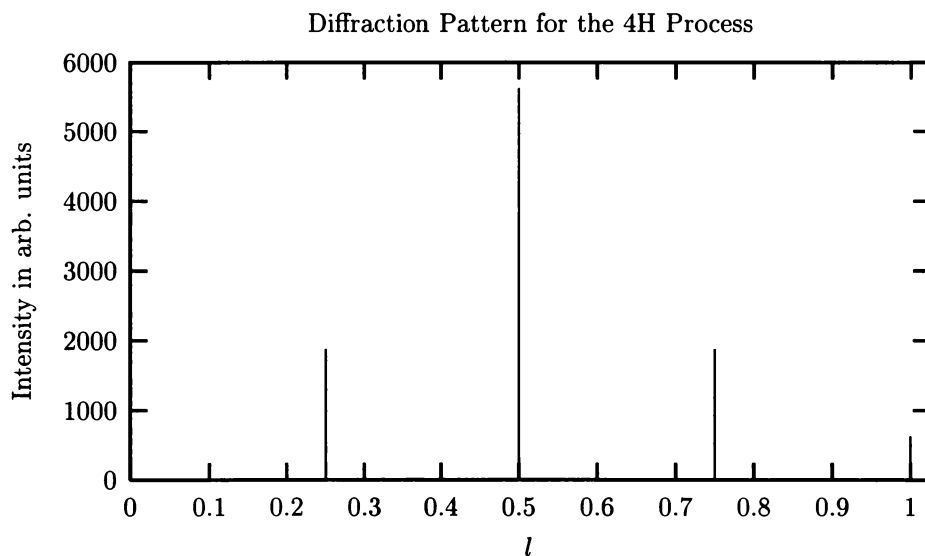


Figure 5.24: The diffraction pattern for a lattice stacked according to the 4H process. We see a completely point spectrum with four equally spaced Bragg peaks at  $l = \frac{m}{4}$  with  $m = 1, 2, 3, 4$ .

are periodic in  $n$ , with a period of  $T_q = 4$ . The diffraction pattern for this process is shown in figure 5.24. We see four Bragg peaks, at  $l = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ , and 1. Since this a crystal, we have  $h_\mu = 0$  and the predictability  $\mathbf{G} = 1$  bits/symbol. We find that the excess entropy and the statistical complexity are  $\mathbf{E} = C_\mu = 2$  bits. The transient information is  $\mathbf{T} = 3$  bit symbols. The correlation length is infinite.

## 5.7 The 3C/2H/4H Process

We now consider a process that represents, at least roughly, three crystal structures interspersed. We treat this case to find out what the diffraction pattern for say, a crystal in the midst of a transformation from 2H to 3C via an intermediary 4H structure might look like. There is some experimental evidence that this might be important in the transition between 3C and 2H structures on annealing at sufficiently high temperatures [27]. The conditional probabilities attached to the states are not necessarily intended to be realistic, but we hope that they are not so different from

Table 5.7: The first few  $Q$ s generated by the 3C/2H/4H process. These correlation functions decay asymptotically to  $\frac{1}{3}$  and have a correlation length  $\lambda_q = 19 \pm 1$ .

$n$	$Q_c(n)$	$Q_r(n)$	$Q_s(n)$
1	0.556	0.444	0
2	0.148	0.259	0.593
3	0.370	0.481	0.149
4	0.200	0.044	0.755
5	0.389	0.566	0.045
6	0.203	0.168	0.629

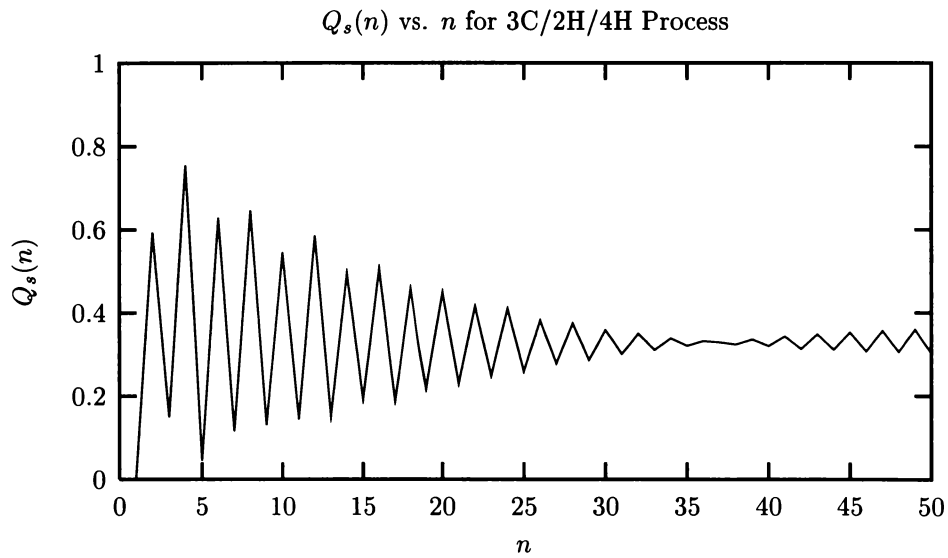


Figure 5.26:  $Q_s(n)$  vs.  $n$  for 3C/2H/4H process.

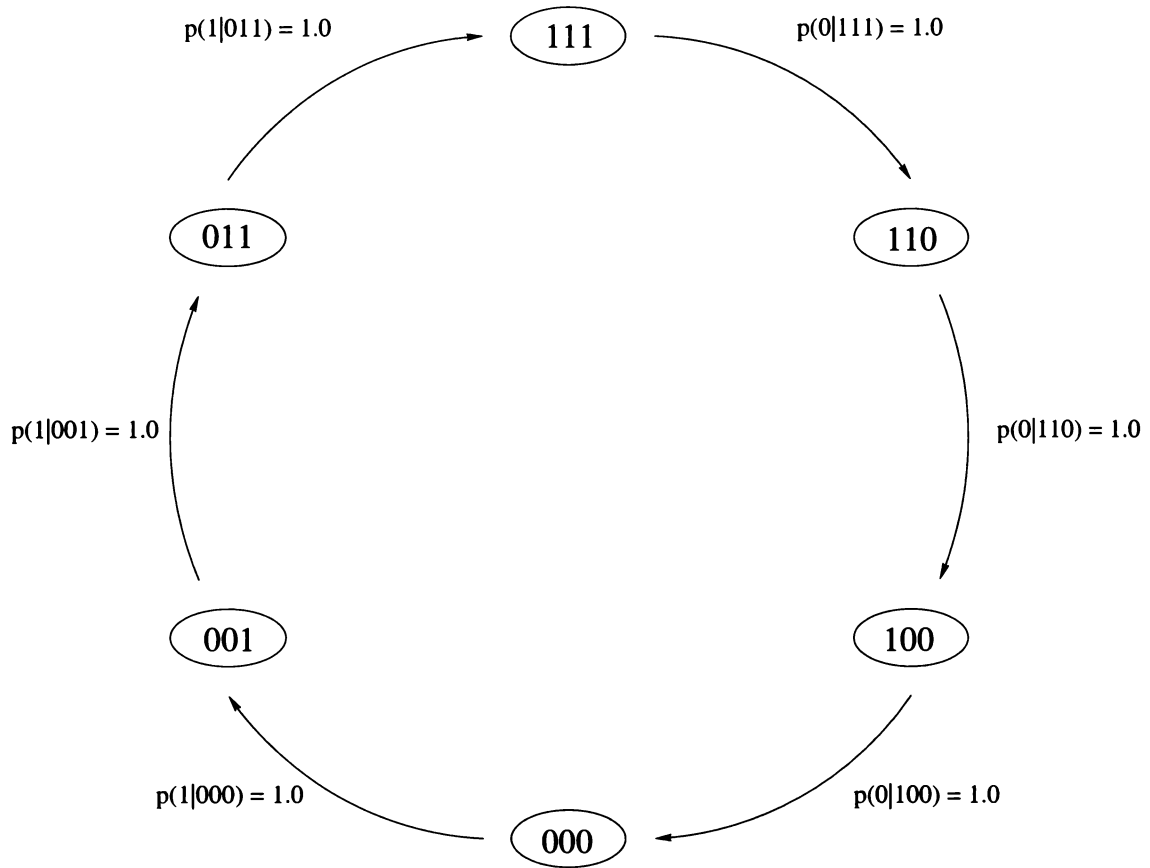


Figure 5.28: The recurrent portion of the  $\epsilon$ -machine for the 6H process. This is a portion of an  $r = 3$  de Bruijn graph, hence we label the causal states by the last three symbols seen.

Table 5.8: The first few  $Q$ s generated by the 6H process. The correlation functions are periodic in  $n$  with period six and have an infinite correlation length.

$n$	$Q_c(n)$	$Q_\tau(n)$	$Q_s(n)$
1	0.500	0.500	0
2	0.333	0.333	0.333
3	0.333	0.333	0.333
4	0.333	0.333	0.333
5	0.500	0.500	0.000
6	0.000	0.000	1.000

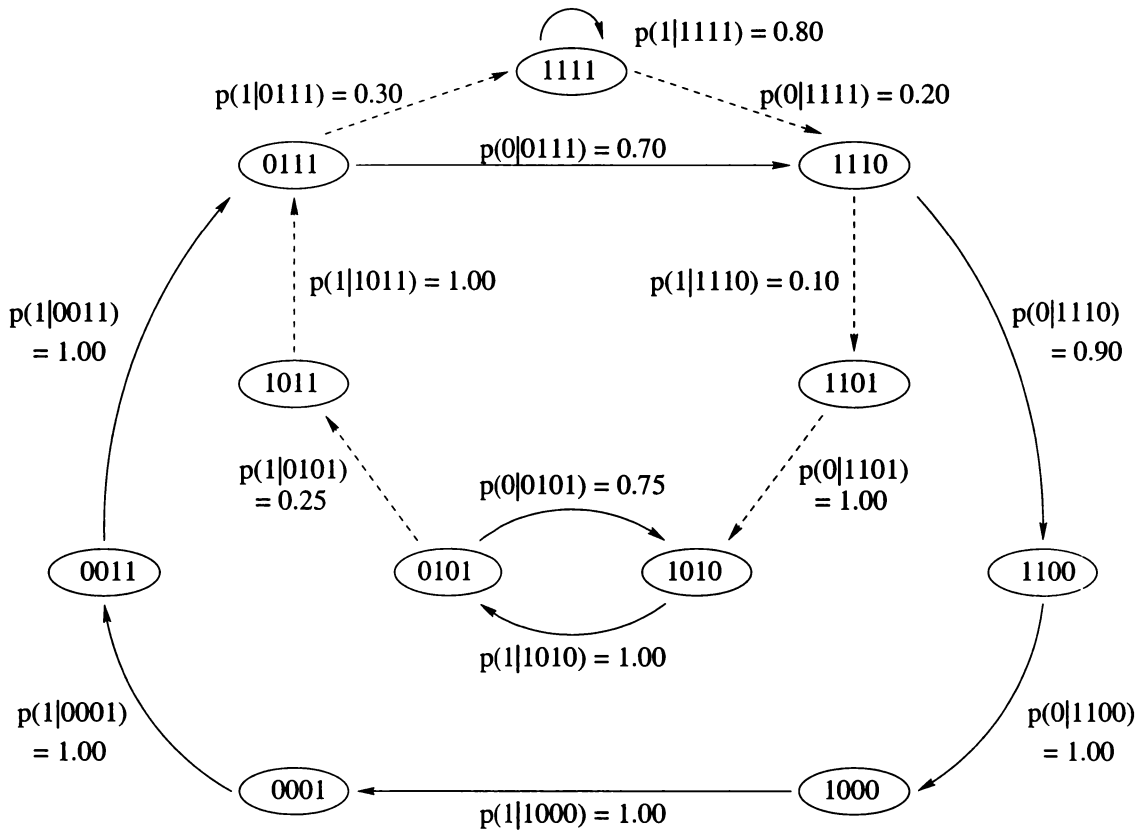


Figure 5.30: The recurrent portion of the  $\epsilon$ -machine for the 3C/2H/6H process. In order to have a barrier between each simple cycle corresponding to some of the crystal structure present, we require that no two such simple cycles share a state. We need an  $r = 4$  de Bruijn graph to accomplish this. The solid lines represent transitions between states associated with the simple cycles giving rise to some crystalline order, and the dotted lines indicate the relatively weak transitions between simple cycles.

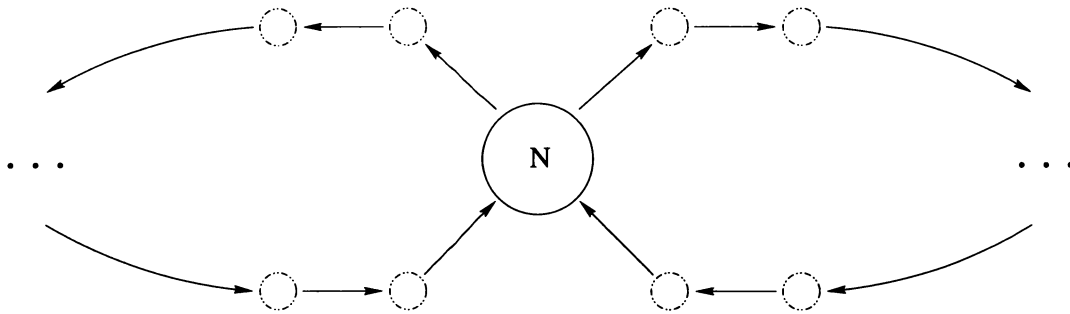


Figure 5.33: A pictorial representation of  $D$ -pair. Each cycle shares the node  $\mathcal{N}$ , which since this is a finite memory process with  $r = 8$ , can be labeled by the last eight symbols seen. So we label the node  $\mathcal{N} = 11010100$ . The two simple cycles which share the node  $\mathcal{N}$  have their nodes represented by the smaller circles. Since each cycle is period fourteen, there should be thirteen smaller nodes in each cycle; but this is onerous to draw, so we have only explicitly shown four nodes in each cycle, with the other nine represented by the  $\dots$ .

sequences does not seem to extinguish them. Again, it is difficult to draw general conclusions from just one sample, but it is suggestive that enhanced intensity at these  $l$  might indicate some 6H structure present.

For this process, we find a correlation length of  $\lambda_q = 49 \pm 1$ . The entropy density is  $h_\mu = 0.332$  bits/symbol and the statistical complexity is  $C_\mu = 3.17$  bits. We find the excess entropy to be  $\mathbf{E} = 1.84$  bits and the transient information to be  $\mathbf{T} = 3.71$  bit symbols. This process implies a memory of range four layers.

## 5.10 A Period 14 $D$ -pair

Canright and Watson [11], on the basis of elementary physical symmetries, proposed that certain simple cycles of a de Bruijn graph could be degenerate in energy and have a zero energy domain wall between them. They assumed a finite interaction between spins on a one-dimensional chain, and considered the case where each spin could assume only discrete values. They found it possible to find such pairs of symmetry related simple cycles that shared a node on a de Bruijn graph. The symmetry of the cycles insured that they had the same energy density (energy per spin) and the sharing of a node insured that there would be no energy cost in flipping from one cycle to the other. They called this pair a “ $D$ -pair” to indicate that a long string made from a series of these would be both disordered and degenerate. An illustration on this is shown in figure 5.33.

The particular  $D$ -pair we will examine has a period of fourteen and is found on an  $r = 8$  de Bruijn graph. The regular expression for this language is  $\mathcal{R} = (11010100 + (001110 + 100011))^*$ . The node that they share is  $\mathcal{N} = 11010100$ , which is invariant under simultaneous spacial and spin inversion. If the system is at the node  $\mathcal{N}$  it has two options to proceed. Taking, say, the left path the series of spins would be 00111011010100 and taking the the right path would give a series 10001111010100. On either path, the cycle returns to the node  $\mathcal{N}$  after visiting thirteen nodes in between. Each node in each cycle has a partner related by simultaneous spatial and spin inversion in the other cycle. Therefore, if the Hamiltonian that describes the system has these symmetries, then the two cycles will be degenerate and since they share a node, there is no energy cost to flip between them. If these simple cycles correspond to the ground state of the system, then we might see such disordered and degenerate states. The question then becomes, “What would the diffraction pattern for a lattice stacked according to this language look like?” This was the original motivation for treating this process. Yi and Canright [79] examined a number of these  $D$ -pairs, and we will

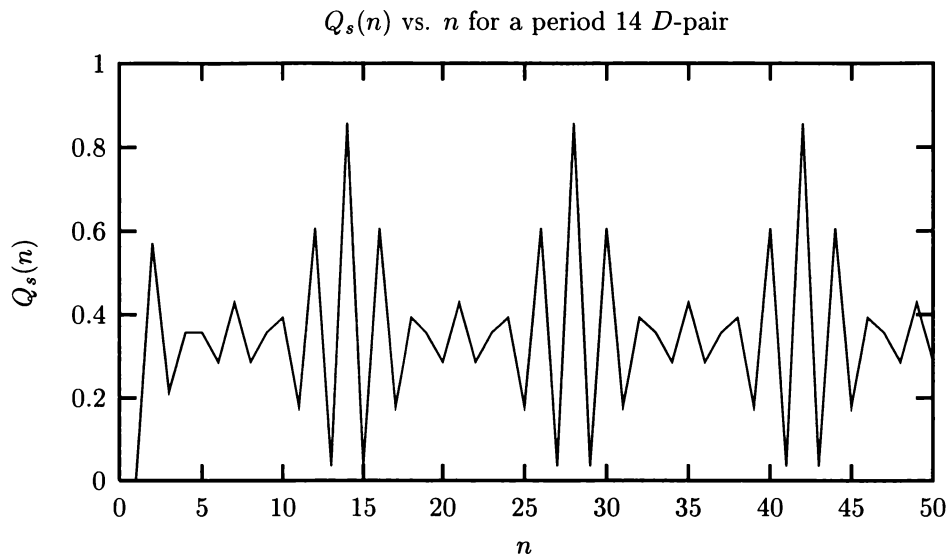


Figure 5.34: The  $Q_s(n)$  for period 14  $D$ -pair. Except for the first few  $n$ , we see periodic correlation functions with period fourteen. This seems odd considering that this process contains some randomness. See text for details.

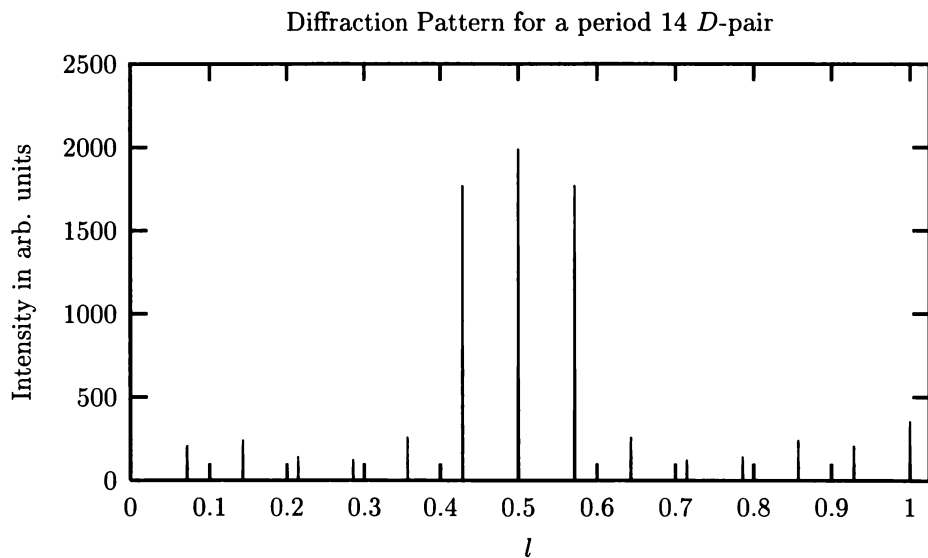


Figure 5.35: The diffraction pattern for a lattice stacked according to the period 14  $D$ -pair. We see a total of fourteen equally spaced Bragg Peaks in the spectrum. Not clear in this figure is the diffuse background scattering.



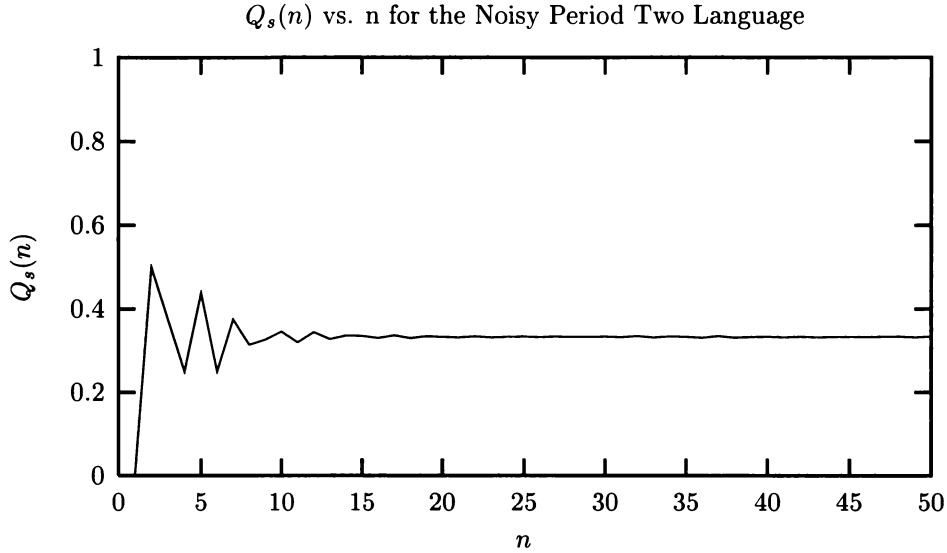


Figure 5.38:  $Q_s(n)$  vs.  $n$  for the noisy period two process. We see that the correlation functions rather quickly decay to their asymptotic value of  $\frac{1}{3}$ . We find a correlation length of  $\lambda = 2.04 \pm 0.03$ .

Another interesting feature of the diffraction pattern is the vanishing intensity at  $l = \frac{5}{6}$ . It is not clear why this is so.

The entropy production of this process is  $h_\mu = \frac{1}{2}$  bits/symbol, and it has both a statistical complexity and excess entropy of  $C_\mu = \mathbf{E} = 1$  bit. The transient information is  $\mathbf{T} = 3.3$  bit symbols. We find an entirely diffuse diffraction pattern and the two-layer correlation functions all decay asymptotically to  $\frac{1}{3}$ , as they must. A graph of  $\log_2 \Psi_q(n)$  vs.  $n$  is shown in figure 5.39. We find a correlation length of  $\lambda_q = 2.04 \pm 0.03$ .

## 5.12 The Even System

Like the noisy period two process, the even system is a process that can not be represented as a portion of a de Bruijn graph. The regular expression for the corresponding language is  $\mathcal{R} = (0+11)^*$ . We can think of this language as the set of words such that there are always an even number of 1s sandwiched between any two 0s. We can also give the forbidden words, which are  $\mathcal{F}_{\mathcal{L}} = \{01^{2k+1}0\}$  where  $k$  is a non-negative integer. In this sense, just as for the noisy period two, the language has an infinite memory, since for any string of 1s, however long, the language must remember whether there have been an even or odd number of 1s since the last 0. Another way of stating this is to say that there is no longest irreducible forbidden word. The recurrent part of the  $\epsilon$ -machine for this process is given in Figure 5.41.

The first few  $Q_s$  for this process are tabulated in Table 5.12 and a plot of  $Q_s(n)$  vs.  $n$  for the first fifty  $n$  is given in figure 5.42. We see that  $Q_s(n)$  quickly approaches its asymptotic value of  $\frac{1}{3}$ . Indeed,  $Q_s(n)$  shows little structure. A plot of  $\log_2 \Psi_q(n)$  vs.  $n$  is given in figure 5.43. We see that the approximation of exponential decay is not very good here. It is not at all clear why this is so. Using the first nine values of  $\log_2 \Psi_q(n)$  versus  $n$  we can calculate the correlation length and we find it to be  $\lambda_q = 1.7 \pm 0.2$ . The diffraction pattern for this process is shown in figure 5.44. We see completely diffuse pattern with two maxima, one at  $l = 0.389$  and a second smaller maximum at  $l = 0.772$ . We also see that the diffraction pattern has a zero at  $l = \frac{5}{6}$ . This is curiously at the

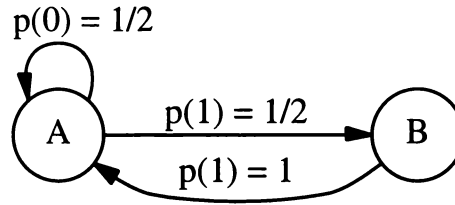


Figure 5.41: The recurrent portion of the  $\epsilon$ -machine that generates the even process. Like the noisy period two process, this process can not be represented by a finite memory machine. We have labeled the two causal states by A and B. Superficially the machine resembles that of the golden mean process. We see that if state A emits a 1 the machine advances state B where it must always emit another 1. Thus this process can never generate a sequence that has an odd number of 1s sandwiched between two 0s.

Table 5.12: The first few values of the correlation functions generated by the even system.

$n$	$Q_c(n)$	$Q_r(n)$	$Q_s(n)$
1	0.667	0.333	0
2	0.167	0.500	0.333
3	0.416	0.167	0.416
4	0.333	0.375	0.292
5	0.271	0.354	0.374
6	0.364	0.334	0.303

same  $l$  value as we saw a zero in the noisy period two.

Turning our attention to computational measures, we see that the entropy rate is  $h_\mu = \frac{2}{3}$  bits/symbol and the predictability  $\mathbf{G} = \frac{1}{3}$  bits/symbol. The statistical complexity  $C_\mu = 0.918$  bits and the excess entropy  $\mathbf{E} = 0.913$  bits. The transient information is  $\mathbf{T} = 3.09$  bit symbols.

### 5.13 The Sum Zero Process

We now consider a process that is similar in spirit to the period 14  $D$ -pair in spirit, but is strictly sofic. That is, we wish to design a process that has both long range correlations as reflected in the  $Q$ s but also has some randomness such that  $h_\mu \neq 0$ . We need long range correlations to see Bragg scattering. So let us imagine, however artificial, a process that does that. Consider the stacking rule such that every other layer has the same orientation but the layers sandwiched between can have a randomly chosen orientation, subject of course to the stacking constraints. The sequence would then look like  $\dots AxAxAxAx\dots$ , where  $x \in \{B, C\}$ . Clearly this will have long range correlations. In terms of a language, we can think of the sequence as being divided into doublets, such that  $\dots aaaaaaa\dots$  becomes  $\dots(aa)(aa)(aa)(aa)\dots$  with  $a \in \{0, 1\}$ . The rule then is that each doublet must contain exactly one 0 and one 1, but the order is arbitrary. If we remind ourselves of the physical meaning of each symbol, that is a 1 gives a relative rotation of sixty degrees between adjacent layers about the stacking direction while a 0 gives a rotation in just the opposite sense, then in terms of operators they are inverses. Hence one followed by the other produces no net rotation and is therefore a ‘sum zero’ operation. Hence the name sum zero process. We can easily translate this into an  $\epsilon$ -machine, the recurrent portion of which is shown in figure 5.45. The regular expression for this language is  $\mathcal{R} = (10 + 01)^*$ .

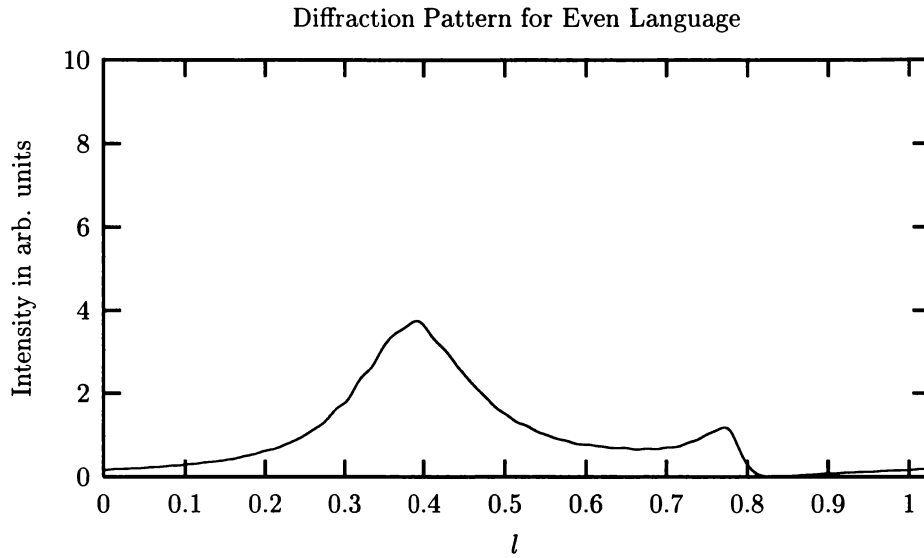


Figure 5.44: The diffraction pattern for a lattice stacked according to the even process. The spectrum is rather featureless, as with the noisy period two, and also has an isolated zero at  $l = \frac{5}{6}$ .

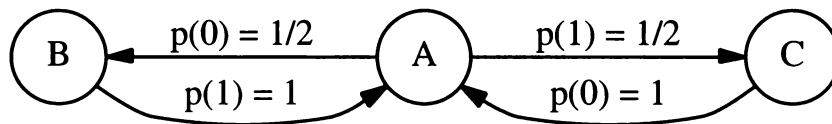


Figure 5.45: The recurrent portion of the  $\epsilon$ -machine that generates the sum zero process. Since this is a strictly sofic process, it can not be described by a finite memory process. We label the causal states by the letters A, B, and C.

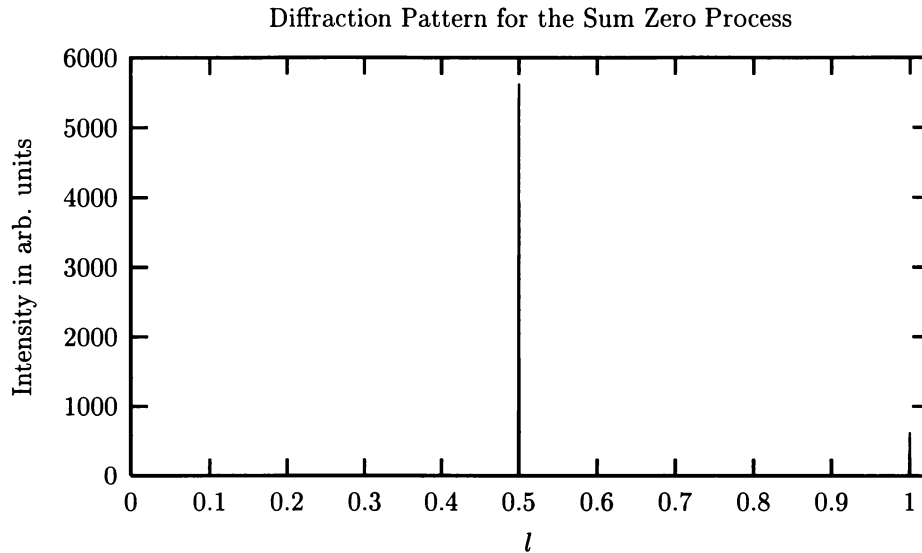


Figure 5.46: The diffraction pattern for a lattice stacked according to the sum zero process. Having two Bragg peaks in the spectrum, one at  $l = \frac{1}{2}$  and the other at  $l = 1$ , this diffraction pattern resembles that of the period two process. The most easily recognizable difference is the the ratio of the intensities is not the same. For the period two we have  $I(l = \frac{1}{2})/I(l = 1) = 3.00$  and for the sum zero process we have  $I(l = \frac{1}{2})/I(l = 1) = 9.00$ . Not seen in this plot is the constant background scattering.

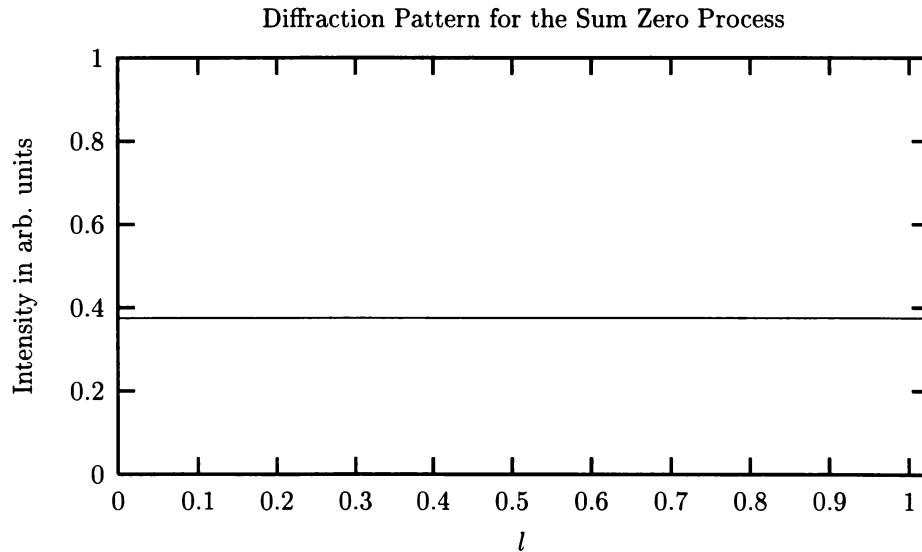


Figure 5.47: The background diffraction pattern for a lattice stacked according to the sum zero process. We see a constant background intensity. Approximately 37.5% of the diffracted intensity falls into this diffuse scattering.

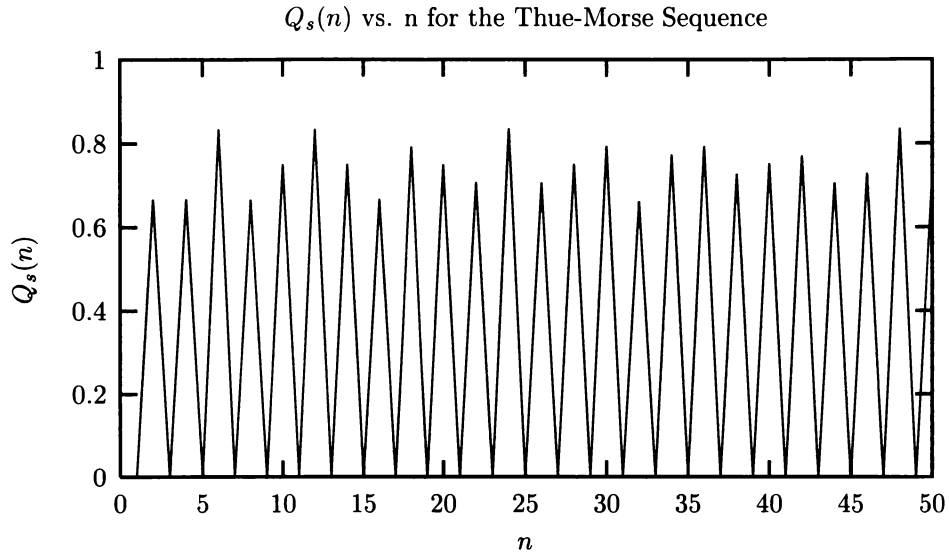


Figure 5.48:  $Q_s(n)$  vs.  $n$  for the Thue-Morse sequence. For odd  $n$ ,  $Q_s(n)$  is zero, but for the even values of  $n$ ,  $Q_s(n)$  do not approach an asymptotic values, but seem to oscillate in what appears to be a random fashion. Of course the Thue-Morse sequence is completely predictable, so the oscillations are not random.

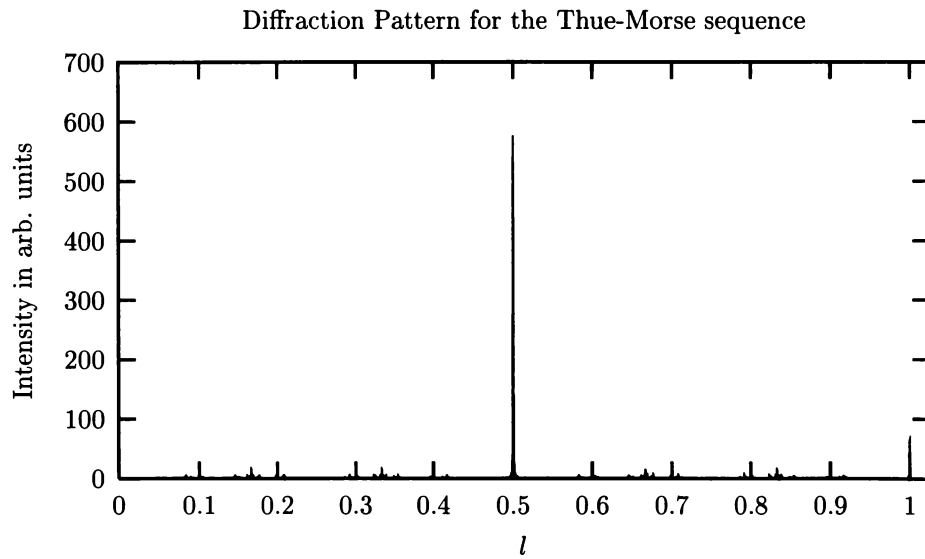


Figure 5.49: The diffraction pattern for a lattice stacked according to the Thue-Morse sequence. We use 1024 layers for this pattern and calculate the intensity at 10,000 equally spaced points. We see two Bragg peaks, one at  $l = \frac{1}{2}$  and the other at  $l = 1$ . This is similar to the period two process and the sum zero process. We find the ratio of the intensities of the two Bragg peaks to be  $I(l = \frac{1}{2})/I(l = 1) = 8.95$ .

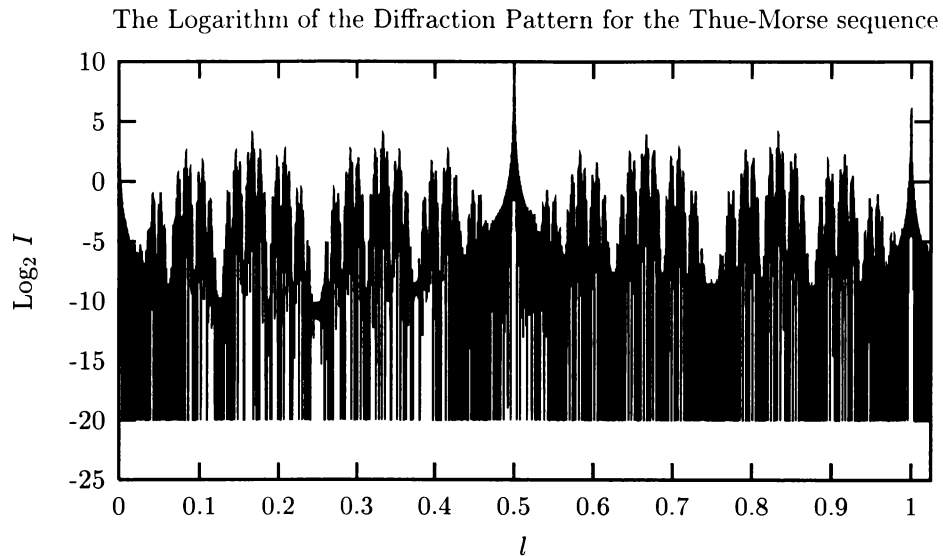


Figure 5.51: The logarithm of the diffraction pattern for a lattice stacked according to the Thue-Morse sequence. We find points in the spectrum where the scattered intensity is zero, and we have cut off the lower portion of the graph at -20.

an optimal prediction of the next spin. The period one does, however, have an infinite correlation length. Knowing one spin and the stacking rule provides information about spins infinitely far away. The Thue-Morse sequence is an example of a system with both an infinite correlation length and memory. The noisy period two has an infinite memory yet correlation information about the absolute positions of the stacking layers decays.

Another interesting point to consider is that diffuse scattering does not preclude the underlying process from being SS. Indeed, we found that two of three SS systems we examined had relatively featureless spectra. There is the interesting phenomenon of the isolated, vanishing intensity at  $l = \frac{5}{6}$  for these two spectra, but it is not known how general this phenomenon is.

We also see that Bragg peaks do not preclude some randomness in a process. We have two examples of this, in the period 14  $D$ -pair and the sum zero process. Also, we see that a completely predictable process, such as the Thue-Morse, can have at least part of the spectrum continuous.

Table 5.16: Correlative and diffractive results. In the second column, we give the asymptotic behavior of the correlations, see §4.3. In the third column, we give the correlation length,  $\lambda_q$ , for each process, see §4.4. The fourth column shows the period of the limit cycle for periodic processes, see §4.4. The last column shows the scattering type for the diffraction pattern. We see that there are three possibilities: absolute continuous (AC), singular continuous (SC), and pure point (PP) or for physicists,  $\delta$ -function or Bragg scattering.

System	$Q$ Behavior	$\lambda_q$	$T_q$	Scattering Type
Fair Coin Toss	<i>decays</i>	1	-	AC
Biased RNG, $q = 0.01$	<i>decays</i>	50.4	-	AC
Biased RNG, $q = 0.02$	<i>decays</i>	23.7	-	AC
Biased RNG, $q = 0.05$	<i>decays</i>	9.57	-	AC
Biased RNG, $q = 0.10$	<i>decays</i>	4.49	-	AC
Biased RNG, $q = 0.20$	<i>decays</i>	2.14	-	AC
Biased RNG, $q = 0.30$	<i>decays</i>	1.43	-	AC
3C	<i>periodic</i>	$\infty$	3	PP
2H	<i>periodic</i>	$\infty$	2	PP
Golden Mean	<i>decays</i>	4.48	-	AC
3C/2H, $q = 0.01$	<i>decays</i>	123	-	AC
3C/2H, $q = 0.05$	<i>decays</i>	25.7	-	AC
3C/2H, $q = 0.10$	<i>decays</i>	13.1	-	AC
3C/2H, $q = 0.20$	<i>decays</i>	5.5	-	AC
3C/2H, $q = 0.30$	<i>decays</i>	3.5	-	AC
3C/2H, $q = 0.40$	<i>decays</i>	3.0	-	AC
4H	<i>periodic</i>	$\infty$	4	PP
3C/2H/4H	<i>decays</i>	19.2	-	AC
6H	<i>periodic</i>	$\infty$	6	PP
3C/2H/6H	<i>decays</i>	49	-	AC
Period 14 $D$ -pair	<i>periodic</i>	$\infty$	14	PP/AC
Noisy Period Two	<i>decays</i>	2.04	-	AC
Even	<i>decays</i>	1.70 (?)	-	AC
Sum Zero	<i>periodic</i>	$\infty$	2	PP/AC
Thue-Morse	<i>aperiodic</i>	$\infty$	-	PP/SC

## Chapter 6

# The Finite $r$ Approximation to $\epsilon$ -Machine Reconstruction from Two-Layer Correlation Functions

Now that we have extracted the correlation functions or  $\{Q_\alpha(n)\}$  from the diffraction data, we can proceed to reconstruct the  $\epsilon$ -machine. We note that there has been work done in the area of relating correlations among symbols in a data stream to word probabilities [2], but our case is somewhat different. Our correlation information is with respect to the absolute spins of the stacking sequence and we want the word probabilities of the relative spin sequence. Our task then, is to relate  $\{Q_\alpha(n)\}$  to word probabilities  $\{p(\omega)\}$ , where  $\omega \in \mathcal{A}^L$ , and  $\mathcal{A}^L$  is the set of all words of length  $L$  over the alphabet  $\mathcal{A}$ , with  $\mathcal{A} \in \{0, 1\}$ . We do this in successive approximations, by considering a machine of finite memory  $r$  and then writing down equations which relate the  $\{Q_\alpha(n)\}$  to the  $\{p(\omega)\}$ . This is equivalent to approximating the process by an  $r^{\text{th}}$ -order Markov process, which we can graphically represent by an  $r^{\text{th}}$ -order de Bruijn graph. It is known that such a graph has  $2^r$  nodes and  $2^{r+1}$  arcs connecting nodes. Since each node has a memory of  $r$ , transitions between nodes are labeled by symbol sequences of length  $r + 1$ , or words of length  $L$ . An  $r^{\text{th}}$ -order Markov process can be completely specified by assigning a probability to each arc. We note, however, only  $2^r$  of these probabilities are independent. The other  $2^r$  probabilities are then constrained by conservation of probability.

We find the  $r^{\text{th}}$ -order approximation by writing down the de Bruijn graph that corresponds to the  $r^{\text{th}}$ -order process. We require conservation of probability at each node, which gives  $2^r$  equations, of which  $2^r - 1$  are independent. We can express this conservation principle mathematically as

$$p(0u) + p(1u) = p(u0) + p(u1) \quad \forall u \quad (6.1)$$

where  $u$  is a sequence of symbols corresponding to a particular node. We additionally require that the total probability to see a word of length  $L$  be unity, i.e.,

$$\sum_{\omega \in \mathcal{A}^L} p(\omega) = 1. \quad (6.2)$$

This then give  $2^r$  equations for  $2^{r+1}$  variables. We find the other  $2^r$  equations by relating  $\{Q_\alpha(n)\}$  to  $\{p(\omega)\}$ . First let us define  $\xi(\omega)$  as



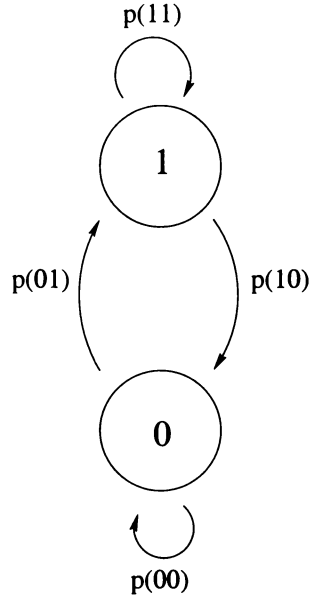


Figure 6.2: The most general  $r = 1$  machine. This graph implies a memory of one, so we label the states by the last symbol seen. This graph can be described by two independent parameters. Two of the four variables,  $p(11)$ ,  $p(10)$ ,  $p(01)$  and  $p(00)$  are constrained by conservation of probability, so we need to use two values from the correlation functions to fix the graph. We use  $Q_c(2)$  and  $Q_r(2)$  for this purpose.

## 6.2 The $r = 1$ approximation

We now work out the  $r = 1$  approximate machine. We need a total of four equations, two of which are constraints among  $\{p(\omega)\}$  and two of which relate  $\{p(\omega)\}$  and  $\{Q_\alpha(n)\}$ . Let us find the  $\{p(\omega)\}$  in terms of  $Q_c(2)$  and  $Q_r(2)$ . Using equation 6.1 we get  $p(10) + p(00) = p(01) + p(00)$ , giving  $p(01) = p(10)$ . Requiring the sum of the probabilities to be unity gives  $p(11) + 2p(01) + p(00) = 1$ . We now use equations 6.5 to get  $Q_c(2) = p(00)$  and  $Q_r(2) = p(11)$ . This is a system of linear equations which are easily solved to give

$$\begin{aligned} p(11) &= Q_r(2) \\ p(01) &= p(10) = \frac{1}{2}[1 - Q_c(2) - Q_r(2)] \\ p(00) &= Q_c(2). \end{aligned} \tag{6.7}$$

## 6.3 The $r = 2$ approximation

We proceed analogously to find the  $r = 2$  approximation to the process. The most general  $r = 2$  machine is shown in figure 6.3. There are a total of eight word probabilities at  $r = 2$ , so we need four equations relating constraints among the  $\{p(\omega)\}$  and four equations relating the  $\{Q_\alpha(n)\}$  to the  $\{p(\omega)\}$ . We can write the first three equations as

$$p(011) = p(110)$$

$$\begin{aligned}
p(001) &= \frac{1}{6}[+3Q_c(2) + 2Q_c(3) + 3Q_r(2) + 4Q_r(3) - 3] \\
p(010) &= \frac{1}{3}[-3Q_c(2) - 2Q_c(3) - 3Q_r(2) - Q_r(3) + 3] \\
p(011) &= \frac{1}{6}[+3Q_c(2) + 4Q_c(3) + 3Q_r(2) + 2Q_r(3) - 3] \\
p(100) &= \frac{1}{6}[+3Q_c(2) + 2Q_c(3) + 3Q_r(2) + 4Q_r(3) - 3] \\
p(101) &= \frac{1}{3}[-3Q_c(2) - Q_c(3) - 3Q_r(2) - 2Q_r(3) + 3] \\
p(110) &= \frac{1}{6}[+3Q_c(2) + 4Q_c(3) + 3Q_r(2) + 2Q_r(3) - 3] \\
p(111) &= \frac{1}{6}[-3Q_c(2) - 4Q_c(3) + 3Q_r(2) - 2Q_r(3) + 3].
\end{aligned} \tag{6.11}$$

## 6.4 The $r = 3$ approximation

Lastly we treat the case of  $r = 3$ . The most general  $r = 3$  machine is shown in figure 6.4. A total of 16 word probabilities dress the  $r = 3$  graph, so we need eight constraints among the  $\{p(\omega)\}$  and eight equations relating  $\{p(\omega)\}$  to  $\{Q_\alpha(n)\}$ . The first seven constraints between the  $\{p(\omega)\}$  can be written as

$$\begin{aligned}
p(0111) &= p(1110) \\
p(0001) &= p(1000) \\
p(0011) + p(1011) &= p(0111) + p(0110) \\
p(0101) + p(1101) &= p(1011) + p(1010) \\
p(0010) + p(1010) &= p(0101) + p(0100) \\
p(0001) + p(1001) &= p(0011) + p(0010) \\
p(0100) + p(1100) &= p(1001) + p(1000).
\end{aligned} \tag{6.12}$$

We still require the overall probability of seeing a word of length four to be unity, so we have,

$$\begin{aligned}
p(0000) + p(0001) + p(0010) + p(0011) + p(0100) \\
p(0101) + p(0110) + p(0111) + p(1000) + p(1001) \\
p(1010) + p(1011) + p(1100) + p(1101) + p(1110) \\
p(1111) = 1.
\end{aligned} \tag{6.13}$$

We now write out the equations which relate the  $\{Q_\alpha(n)\}$  and the  $\{p(\omega)\}$  and find,

$$\begin{aligned}
Q_c(2) &= p(0000) + p(0001) + p(0010) + p(0011) \\
Q_r(2) &= p(1100) + p(1101) + p(1110) + p(1111) \\
Q_c(3) &= p(0110) + p(0111) + p(1010) + p(1011) \\
&\quad + p(1100) + p(1101) \\
Q_r(3) &= p(0010) + p(0011) + p(0100) + p(0101) \\
&\quad + p(1000) + p(1001)
\end{aligned}$$

$$\begin{aligned}
Q_c(4) &= p(1111) + p(1000) + p(0100) + p(0010) \\
&\quad + p(0001) \\
Q_r(4) &= p(0000) + p(0111) + p(1011) + p(1101) \\
&\quad + p(1110) \\
Q_c(5) &= \frac{p^2(0000)}{p(0000) + p(0001)} + \frac{p(0011)p(0111)}{p(0111) + p(0110)} \\
&\quad + \frac{p(0101)p(1011)}{p(1011) + p(1010)} + \frac{p(0110)p(1101)}{p(1101) + p(1100)} \\
&\quad + \frac{p(0111)p(1110)}{p(1110) + p(1111)} + \frac{p(1001)p(0011)}{p(0011) + p(0010)} \\
&\quad + \frac{p(1010)p(0101)}{p(0101) + p(0100)} + \frac{p(1011)p(0110)}{p(0110) + p(0111)} \\
&\quad + \frac{p(1100)p(1001)}{p(1001) + p(1000)} + \frac{p(1101)p(1010)}{p(1010) + p(1011)} \\
&\quad + \frac{p(1110)p(1100)}{p(1100) + p(1101)} \\
Q_r(5) &= \frac{p^2(1111)}{p(1111) + p(1110)} + \frac{p(1100)p(1000)}{p(1000) + p(1001)} \\
&\quad + \frac{p(1010)p(0100)}{p(0100) + p(0101)} + \frac{p(1001)p(0010)}{p(0010) + p(0011)} \\
&\quad + \frac{p(1000)p(0001)}{p(0001) + p(0000)} + \frac{p(0110)p(1100)}{p(1100) + p(1101)} \\
&\quad + \frac{p(0101)p(1010)}{p(1010) + p(1011)} + \frac{p(0100)p(1001)}{p(1001) + p(1000)} \\
&\quad + \frac{p(0011)p(0110)}{p(0110) + p(0111)} + \frac{p(0010)p(0101)}{p(0101) + p(0100)} \\
&\quad + \frac{p(0001)p(0011)}{p(0011) + p(0010)}.
\end{aligned} \tag{6.14}$$

The last two relations of equations 6.14 require some explanation. At  $L = 5$ , a typical term in the sum to find  $Q_c(5)$  might look like  $p(00111)$ . We want to express this probability of seeing a length five word as some function of the probability of seeing length four words. So we say that  $p(00111) = p(0011)p(1|0011)$ , where  $p(1|0011)$  is the conditional probability of seeing a 1 having already seen a 0011. But looking at the figure 6.4 it is clear that having last seen a 0011 puts us at the node 011. The probability of seeing another 1 is just the branching ratio at node 011, which we can write as  $\frac{p(0111)}{p(0111) + p(0110)}$ . Therefore, we find that  $p(00111) = \frac{p(0011)p(0111)}{p(0110) + p(0111)}$ .

It is reasonable to ask why we need the de Bruijn graphs to write down these relations. Some seem to follow directly from probability theory. Indeed, the de Bruijn graphs are a convenient and pictorial way to describe  $r^{\text{th}}$ -order Markov processes. Up to  $r = 2$ , the de Bruijn graphs impose no additional constraints on the probabilities and one can derive the same  $r = 2$  equations from elementary probability theory. At  $r = 3$ , they do alter the equations slightly. By making the assumption of a finite range, the conditional probabilities are truncated to look only at a depth  $r$ . Consider the term  $p(00111)$ . We found this term to be equal to  $\frac{p(0011)p(0111)}{p(0110) + p(0111)}$  which can be rewritten

$$\begin{aligned}
f_7 &= p(1001) + p(1000) - p(0100) - p(1100) \\
f_8 &= p(0000) + p(0001) + p(0010) + p(0011) + p(0100) \\
&\quad + p(0101) + p(0110) + p(0111) + p(1000) + p(1001) \\
&\quad + p(1010) + p(1011) + p(1100) + p(1101) + p(1110) \\
&\quad + p(1111) - 1 \\
f_9 &= p(0000) + p(0001) + p(0010) + p(0011) - Q_c(2) \\
f_{10} &= p(1100) + p(1101) + p(1110) + p(1111) - Q_r(2) \\
f_{11} &= p(0110) + p(0111) + p(1010) + p(1011) \\
&\quad + p(1100) + p(1101) - Q_c(3) \\
f_{12} &= p(0010) + p(0011) + p(0100) + p(0101) \\
&\quad + p(1000) + p(1001) - Q_r(3) \\
f_{13} &= p(1111) + p(1000) + p(0100) + p(0010) \\
&\quad + p(0001) - Q_c(4) \\
f_{14} &= p(0000) + p(0111) + p(1011) + p(1101) \\
&\quad + p(1110) - Q_r(4) \tag{6.16} \\
f_{15} &= \frac{p^2(0000)}{p(0000) + p(0001)} + \frac{p(0011)p(0111)}{p(0111) + p(0110)} \\
&\quad + \frac{p(0101)p(1011)}{p(1011) + p(1010)} + \frac{p(0110)p(1101)}{p(1101) + p(1100)} \\
&\quad + \frac{p(0111)p(1110)}{p(1110) + p(1111)} + \frac{p(1001)p(0011)}{p(0011) + p(0010)} \\
&\quad + \frac{p(1010)p(0101)}{p(0101) + p(0100)} + \frac{p(1011)p(0110)}{p(0110) + p(0111)} \\
&\quad + \frac{p(1100)p(1001)}{p(1001) + p(1000)} + \frac{p(1101)p(1010)}{p(1010) + p(1011)} \\
&\quad + \frac{p(1110)p(1100)}{p(1100) + p(1101)} - Q_c(5) \\
f_{16} &= \frac{p^2(1111)}{p(1111) + p(1110)} + \frac{p(1100)p(1000)}{p(1000) + p(1001)} \\
&\quad + \frac{p(1010)p(0100)}{p(0100) + p(0101)} + \frac{p(1001)p(0010)}{p(0010) + p(0011)} \\
&\quad + \frac{p(1000)p(0001)}{p(0001) + p(0000)} + \frac{p(0110)p(1100)}{p(1100) + p(1101)} \\
&\quad + \frac{p(0101)p(1010)}{p(1010) + p(1011)} + \frac{p(0100)p(1001)}{p(1001) + p(1000)} \\
&\quad + \frac{p(0011)p(0110)}{p(0110) + p(0111)} + \frac{p(0010)p(0101)}{p(0101) + p(0100)} \\
&\quad + \frac{p(0001)p(0011)}{p(0011) + p(0010)} - Q_r(5).
\end{aligned}$$

Table 6.1: Exact and noisy correlation functions for the random number process. These are the correlation functions used in solving equations 6.16. The exact  $\{Q_\alpha(n)\}$  are found from an analytical solution for the random number generator, and the noisy version is found by considering a finite sample of the process.

$n$	$Q_c(n)$ exact	$Q_r(n)$ exact	$Q_c(n)$ noisy	$Q_r(n)$ noisy
2	0.25000	0.25000	0.2539	0.2489
3	0.37500	0.37500	0.3744	0.3732
4	0.31250	0.31250	0.3127	0.3142
5	0.34375	0.34375	0.3444	0.3407

Table 6.2: Solution at  $r = 3$  for the random number generator. Using the exact correlation functions, we get a fitness of  $\mathcal{F} = 1.26 \times 10^{-14}$ . With the noisy correlation functions, the fit is not nearly as good,  $\mathcal{F} = 1.72 \times 10^{-6}$ .

Word	Exact Answer	Numerical Solution with exact $Q$ 's	Numerical Solution with noisy $Q$ 's
$p(0000)$	0.0625	0.0626	0.0668
$p(0001)$	0.0625	0.0624	0.0619
$p(0010)$	0.0625	0.0625	0.0630
$p(0011)$	0.0625	0.0625	0.0622
$p(0100)$	0.0625	0.0626	0.0631
$p(0101)$	0.0625	0.0624	0.0599
$p(0110)$	0.0625	0.0626	0.0648
$p(0111)$	0.0625	0.0624	0.0607
$p(1000)$	0.0625	0.0624	0.0618
$p(1001)$	0.0625	0.0626	0.0634
$p(1010)$	0.0625	0.0625	0.0599
$p(1011)$	0.0625	0.0625	0.0633
$p(1100)$	0.0625	0.0624	0.0622
$p(1101)$	0.0625	0.0626	0.0633
$p(1110)$	0.0625	0.0624	0.0608
$p(1111)$	0.0625	0.0626	0.0625

Table 6.3: Exact and noisy correlation functions for the second process. These are the correlation functions used in solving equations 6.16. The exact  $\{Q_{\alpha}(n)\}$  are found from direct calculation of the process. Only four digits of accuracy are reported here for convenience. The noisy version is found by considering a finite sample of the process.

$n$	$Q_c(n)$ exact	$Q_r(n)$ exact	$Q_c(n)$ noisy	$Q_r(n)$ noisy
2	0.2300	0.1300	0.2279	0.1289
3	0.3600	0.4500	0.3611	0.4504
4	0.2900	0.1200	0.2891	0.1191
5	0.2969	0.5254	0.2981	0.5252

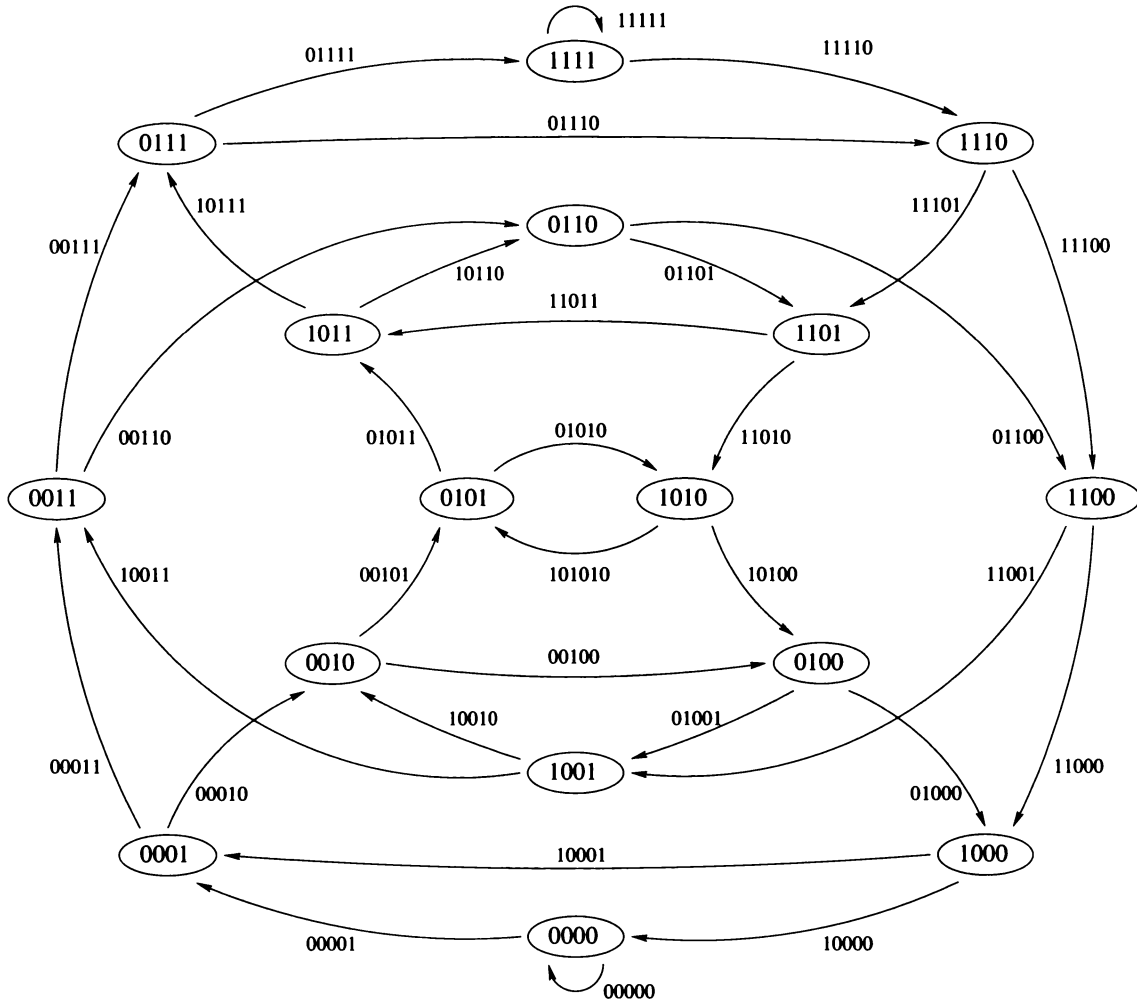


Figure 6.5: The most general  $r = 4$  machine. The number of nodes has grown to sixteen and now there are thirty-two arcs. Conservation of probability provides sixteen constraints among the arcs. The remaining sixteen constraints come from the correlation functions, and we would use  $Q_\alpha(n)$  with  $\alpha \in \{c, r\}$  and  $n \in \{2, 3, 4, 5, 6, 7, 8, 9\}$ . To solve for this system we would need to solve thirty-two simultaneous algebraic equations, eight of which would be non-linear. At  $n = 9$ , the two equations relating correlation functions to word probabilities would have 171 terms.

## Chapter 7

# Examples of $\epsilon$ -Machine Reconstruction from Known Processes

We will now consider four examples of machine reconstruction from processes that can not be represented on an  $r = 3$  de Bruijn graph. For any process that is describable in terms of a third-order Markov process, our reconstruction procedure will find the underlying process. It is for those cases where the structure is not third-order Markovian that we wish to determine how the reconstruction algorithm works. So, we will treat the  $r = 4$  process given in §5.9, another  $r = 4$  process not so far discussed, the noisy period two in §5.11, and the even system, §5.12.

### 7.1 Machine Reconstruction for the 3C/2H/6H Process

Let us begin with the 3C/2H/6H process described in §5.9. We see that this process has both a strong fcc component and 6H cycle. We might therefore expect that this will give our algorithm some difficulty as the simultaneous existence of these two is possible only on an  $r = 4$  graph. The machine reconstruction results are given in table 7.1. We also calculate the correlation functions and the diffraction patterns of the various  $r$  approximations. Since the  $r = 0$  approximation at best corresponds to a biased random number generator, and such a process clearly does not represent the structure seen in the 3C/2H/6H process, we do not calculate it here.

Figure 7.1 shows  $Q_s(n)$  versus  $n$  for both the 3C/2H/6H process and the  $r = 1$  approximation to it. We see that with exception of  $n = 1, 2$  and  $3$ , the correlations die out far too fast for the  $r = 1$  approximation. Indeed, there appears to be some significant long range structure that the  $r = 1$  is failing to capture. A comparison of the diffraction pattern of the  $r = 1$  approximation to that of the 3C/2H/6H process is shown in figure 7.2. We see that Bragg peak at  $l = \frac{1}{3}$ , save for a small bump, is missing in the  $r = 1$  approximation. The only structure in the diffraction pattern that the approximation models moderately well is that at  $l = \frac{2}{3}$ . The other features in the spectrum, at  $l = \frac{1}{6}, \frac{1}{2}$  and  $\frac{5}{6}$ , are totally absent in the  $r = 1$  approximation to the diffraction pattern.

Figure 7.3 compares the  $Q_s(n)$  for the  $r = 2$  approximation to that of the 3C/2H/6H process. We see that very little, if any, additional structure has been discovered by increasing  $r$ . Examining the diffraction patterns in figure 7.4 paints an even more dismal picture. We see a smooth diffuse background for the  $r = 2$  spectrum, giving scant notice to the diffraction maximum in the 3C/2H/6H process. It seems that no progress has been made in increasing  $r$  from 1 to 2. We can also compare the computation measures for the two approximations, as shown in table 7.2. We observe that the entropy density  $h_\mu$  only decreases slightly (from 0.910  $\rightarrow$  0.904) as  $r$  goes 1  $\rightarrow$  2. The excess entropy

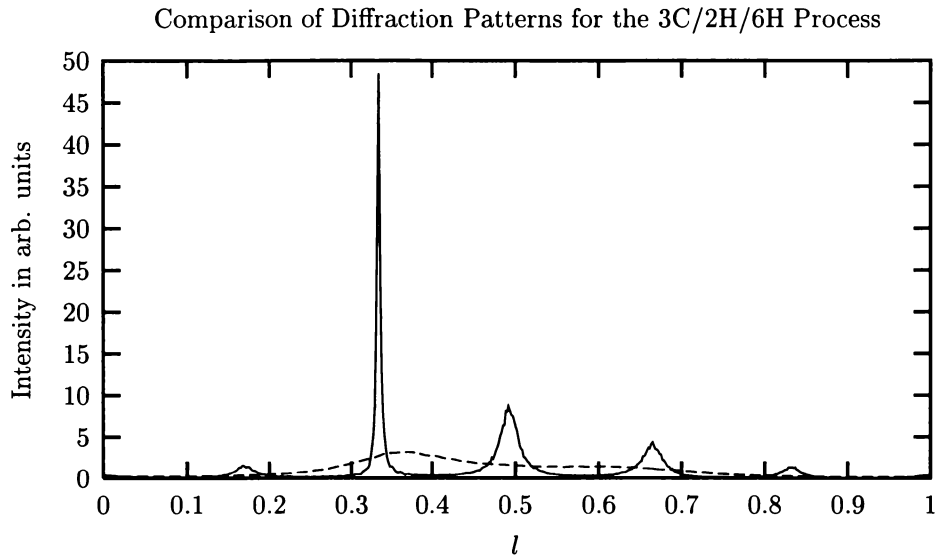


Figure 7.2: The diffraction pattern for a lattice stacked according to the 3C/2H/6H process (solid line) and the  $r = 1$  approximation (dashed line). The diffraction pattern for the  $r = 1$  approximation is clearly missing all of the structure present. Of particular note is the fact that the small rises at  $l = \frac{1}{6}$  and  $\frac{5}{6}$  are completely absent in the diffraction pattern for the  $r = 1$  approximation.

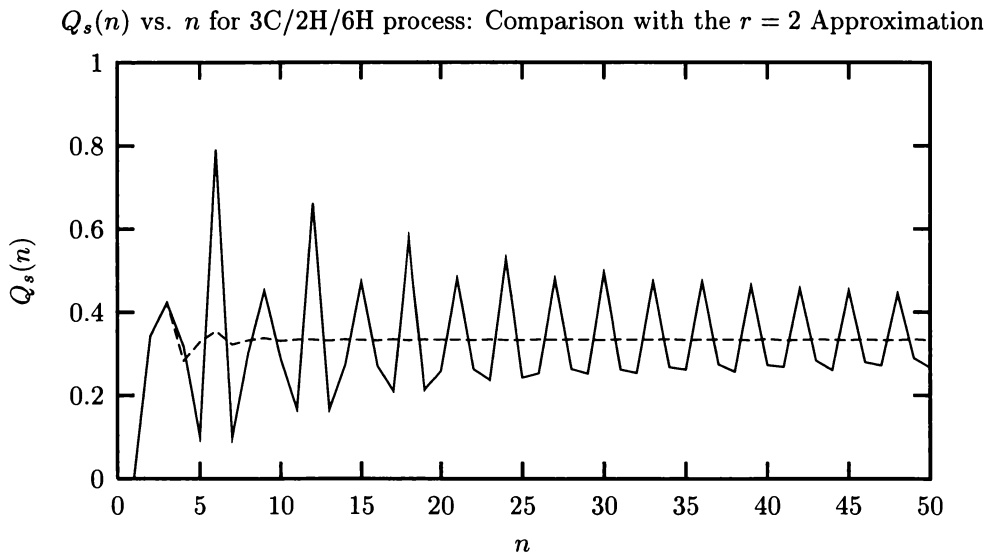


Figure 7.3: The  $Q_s(n)$  vs.  $n$  for the 3C/2H/6H Process (solid line) and the  $r = 2$  approximation to the process (dashed line). Again, as with the  $r = 1$  approximation, we see that the correlation functions die out too quickly for the  $r = 2$  approximation.



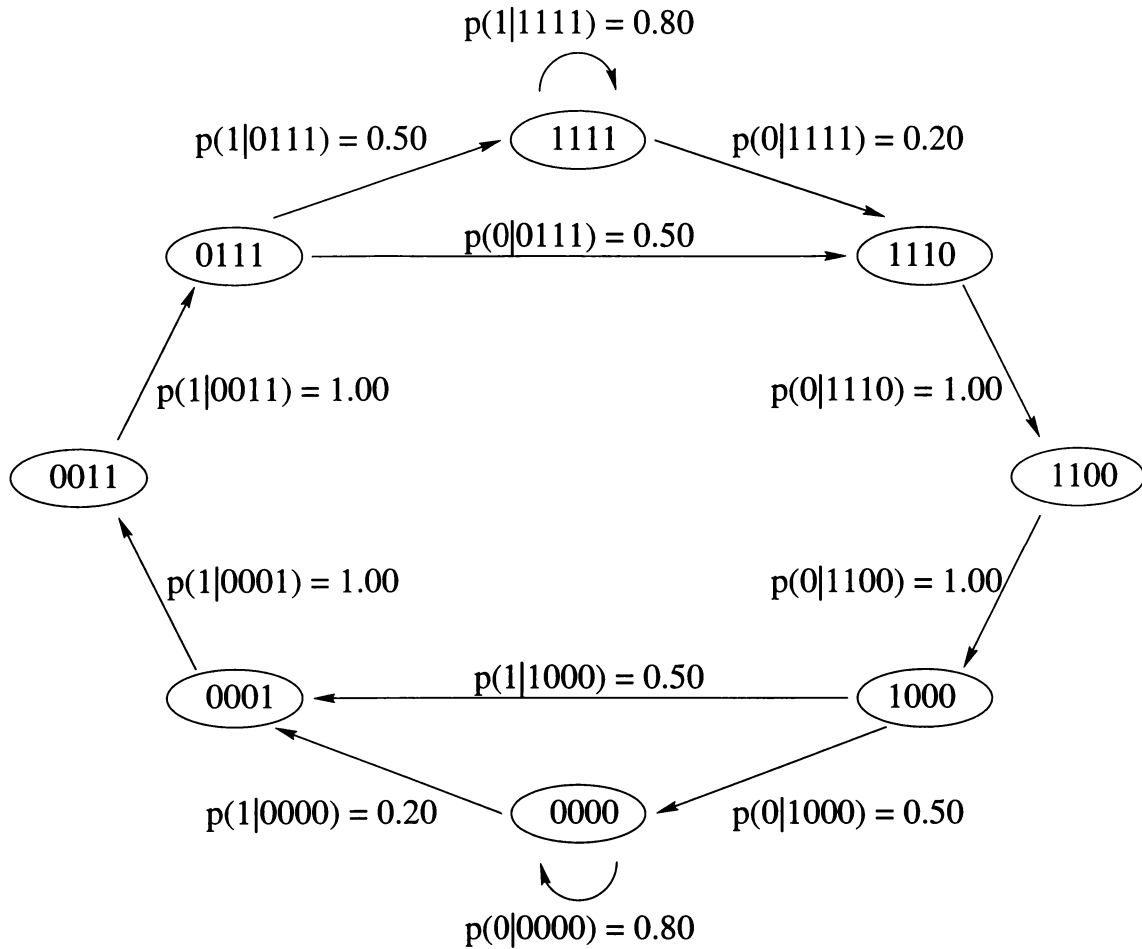


Figure 7.5: The recurrent portion of the  $\epsilon$ -machine for the 3C/6H process. Since we wish to have a barrier between the two cycles representing the 3C and the 6H, we must use a graph with an  $r = 4$  memory. SO this graph is a portion of a  $r = 4$  de Bruijn graph.

see that actually '0000' does not occur at all in the 3C/2H/6H process, and the reconstructed machine significantly overestimates  $p(1111)$ . Similarly, other word probabilities are not well-represented. We understand these results as stemming from the inability of an  $r = 3$  graph to simultaneously support both 3C and 6H structure.

## 7.2 Machine Reconstruction for the 3C/6H Process

We now apply machine reconstruction to another process which is only describable by a machine with a memory of at least  $r = 4$ . We expect that some ZnS crystals will incompletely transform to a twinned 3C structure and contain remnants of the 6H structure. We call this the 3C/6H process, and the recurrent portion of the  $\epsilon$ -machine is shown in figure 7.5.

The results for machine reconstruction up to  $r = 3$  are shown in table 7.3. We find the correlation functions for the  $r = 1$  approximation and the first fifty values of  $Q_s(n)$  from the  $r = 1$  approximation are compared with those of the 3C/6H process in figure 7.6. Perhaps not surprisingly, there is only good agreement for the first few  $n$ , after which the  $Q_s(n)$  for the  $r = 1$  approximation decay too

Table 7.3: Machine reconstruction results for the 3C/6H process. Comparing the last two columns, we see that the agreement is not too bad, but there is still a troublingly large discrepancy between the reconstructed word probabilities (WPs) and the those of the exact process.

r values	Words	Word Probabilities	r values	Words	Reconstructed WPs	Exact WPs
$r = 0$	$p(0)$	0.498	$r = 3$	$p(0000)$	0.296	0.227
	$p(1)$	0.502		$p(0001)$	0.026	0.091
$r = 1$	$p(00)$	0.407		$p(0010)$	0.030	0.000
	$p(01)$	0.091		$p(0011)$	0.046	0.091
	$p(10)$	0.091		$p(0100)$	0.026	0.000
	$p(11)$	0.410		$p(0101)$	0.000	0.000
$r = 2$	$p(000)$	0.316		$p(0110)$	0.045	0.000
	$p(001)$	0.091		$p(0111)$	0.025	0.091
	$p(010)$	0.000		$p(1000)$	0.027	0.091
	$p(011)$	0.091		$p(1001)$	0.049	0.000
	$p(100)$	0.091		$p(1010)$	0.000	0.000
	$p(101)$	0.000	$p(1011)$	0.027	0.000	
	$p(110)$	0.091	$p(1100)$	0.052	0.091	
	$p(111)$	0.319	$p(1101)$	0.029	0.000	
			$p(1110)$	0.024	0.091	
		$p(1111)$	0.300	0.227		

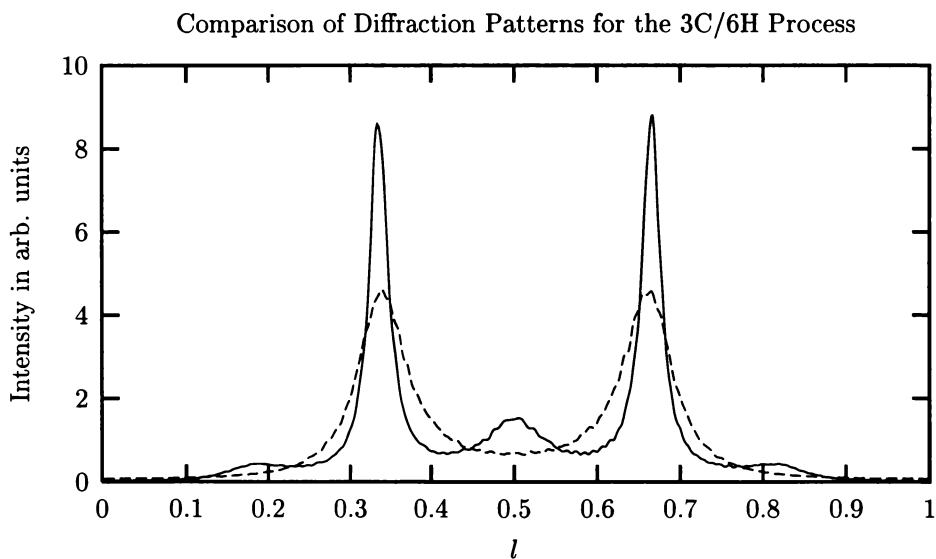


Figure 7.7: The diffraction pattern for a lattice stacked according to the 3C/6H process (solid line) and the  $r = 1$  approximation (dashed line). We see that the rise at  $l = \frac{1}{2}$  in the diffraction pattern for the 3C/6H process is completely absent in the  $r = 1$  approximation.

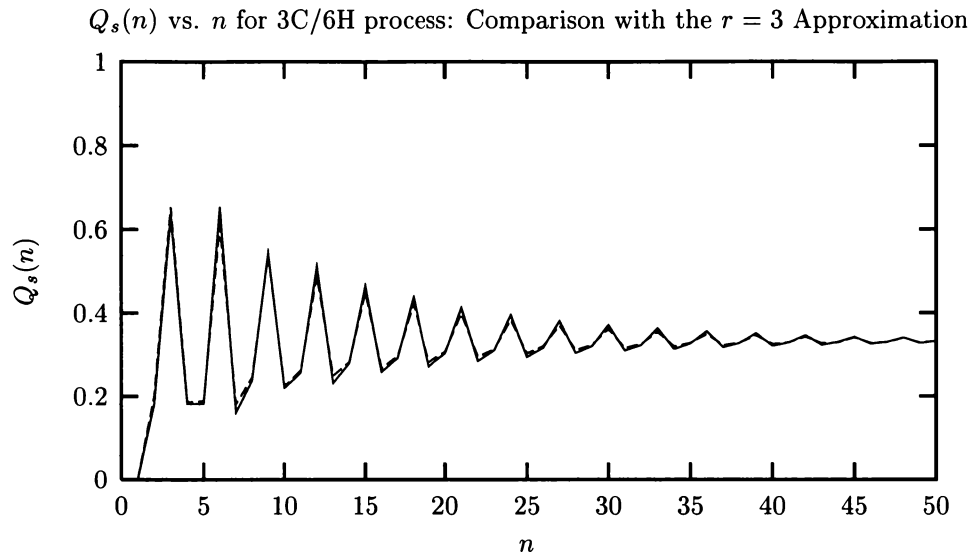


Figure 7.10: The  $Q_s(n)$  vs.  $n$  for the 3C/6H process (solid line) and the  $r = 3$  approximation to the process (dashed line). Here we see quite good agreement between the correlation functions for  $r = 3$  approximation to the process and those of the actual process.

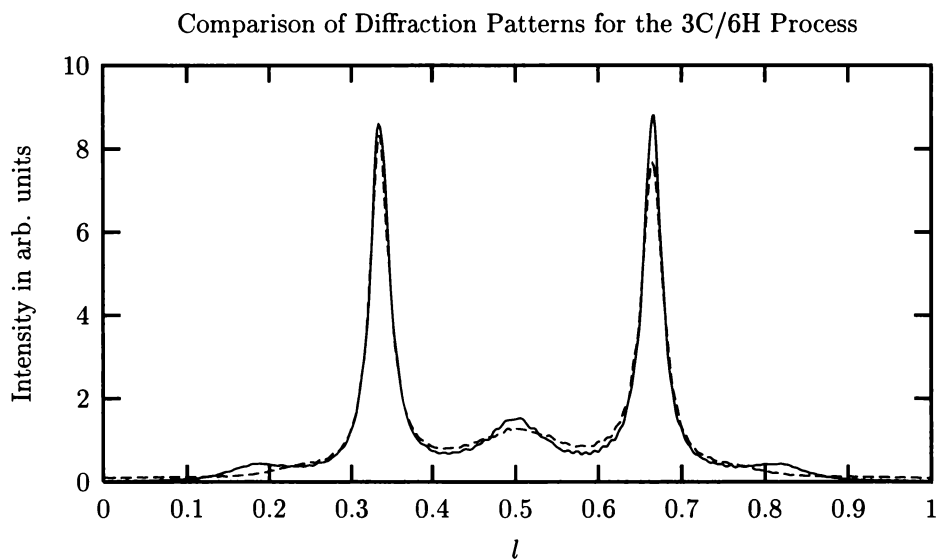


Figure 7.11: The diffraction pattern for a lattice stacked according to the 3C/6H process (solid line) and the  $r = 3$  approximation (dashed line). As with the  $r = 3$  correlation functions, the agreement between the  $r = 3$  approximation to the diffraction pattern and the diffraction pattern of the actual process is quite good, except for the small rises at  $l = \frac{1}{6}$  and  $\frac{5}{6}$ . We interpret this as resulting from the fact that at  $r = 3$  there is not a sufficient structure in the graph to simultaneously support both 3C and 6H simple cycles.

Table 7.5: Machine reconstruction results for the noisy period two system. Comparing the last two columns we see that the same difficulty that arose in the 3C/6H/2H process is present here, namely that the recurrent portion of the  $\epsilon$ -machine is not strongly connected since  $p(0000) \neq 0$  even though  $p(0001) = p(1000) = 0$ . The problem is less severe here due to the smallness of the weight attached to  $p(0000)$ . The other word probabilities are reasonably well-represented by the  $r = 3$  approximation.

r values	Words	Word Probabilities	r values	Words	Reconstructed WPs	Exact WPs
$r = 0$	$p(0)$	0.250	$r = 3$	$p(0000)$	0.007	0.000
	$p(1)$	0.750		$p(0001)$	0.000	0.000
$r = 1$	$p(00)$	0.000		$p(0010)$	0.000	0.000
	$p(01)$	0.250		$p(0011)$	0.000	0.000
	$p(10)$	0.250		$p(0100)$	0.005	0.000
	$p(11)$	0.500		$p(0101)$	0.116	0.125
$r = 2$	$p(000)$	0.000		$p(0110)$	0.010	0.000
	$p(001)$	0.000		$p(0111)$	0.115	0.125
	$p(010)$	0.125		$p(1000)$	0.000	0.000
	$p(011)$	0.125		$p(1001)$	0.005	0.000
	$p(100)$	0.000		$p(1010)$	0.122	0.125
	$p(101)$	0.250		$p(1011)$	0.124	0.125
	$p(110)$	0.125		$p(1100)$	0.002	0.000
	$p(111)$	0.375	$p(1101)$	0.132	0.125	
			$p(1110)$	0.116	0.125	
		$p(1111)$	0.249	0.250		

$Q_s(n)$  vs.  $n$  for the Noisy Period Two Process and The  $r = 0$  Approximation

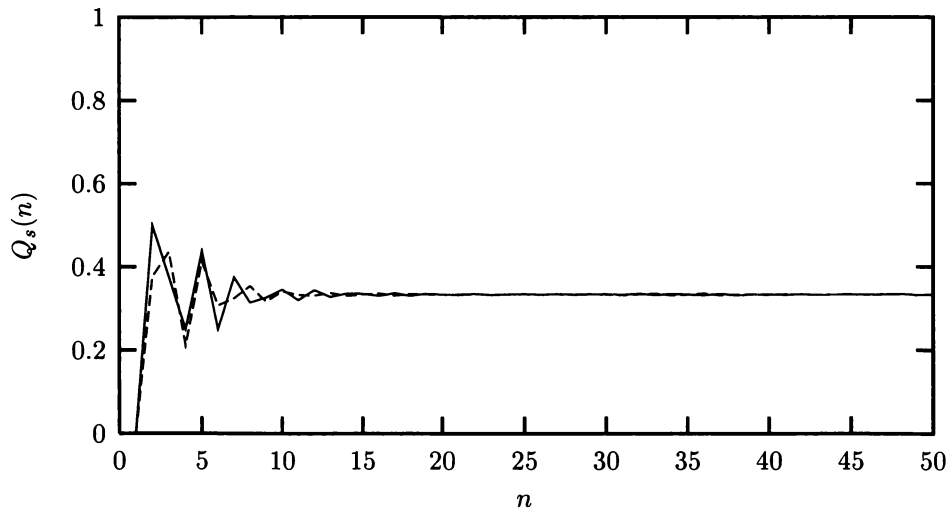


Figure 7.12: The  $Q_s(n)$  vs.  $n$  for the noisy period two process (solid line) and the  $r = 0$  approximation to the process (dashed line). We see that both decay quickly to the asymptotic value of  $\frac{1}{3}$ , but there is some small difference for small  $n$ .

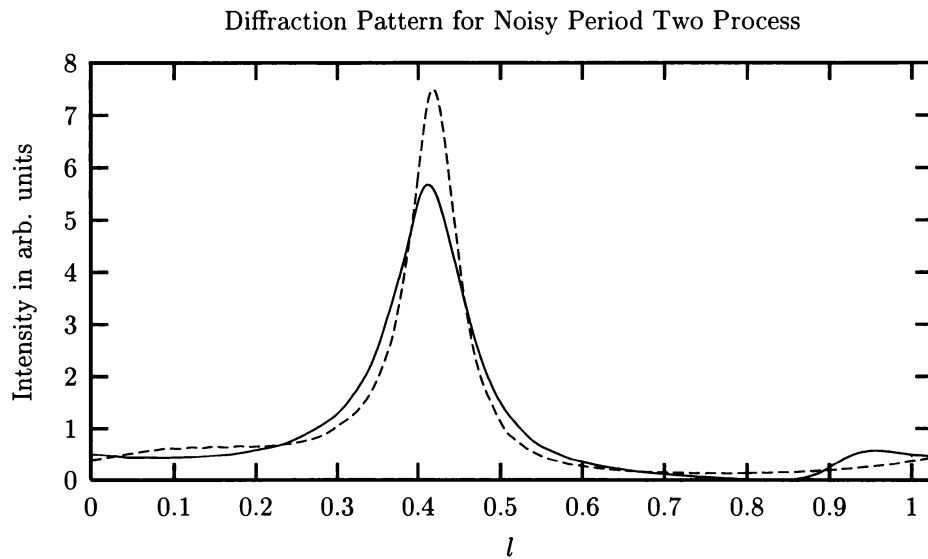


Figure 7.15: The diffraction pattern for a lattice stacked according to the noisy period two process (solid line) and the  $r = 1$  approximation (dashed line).

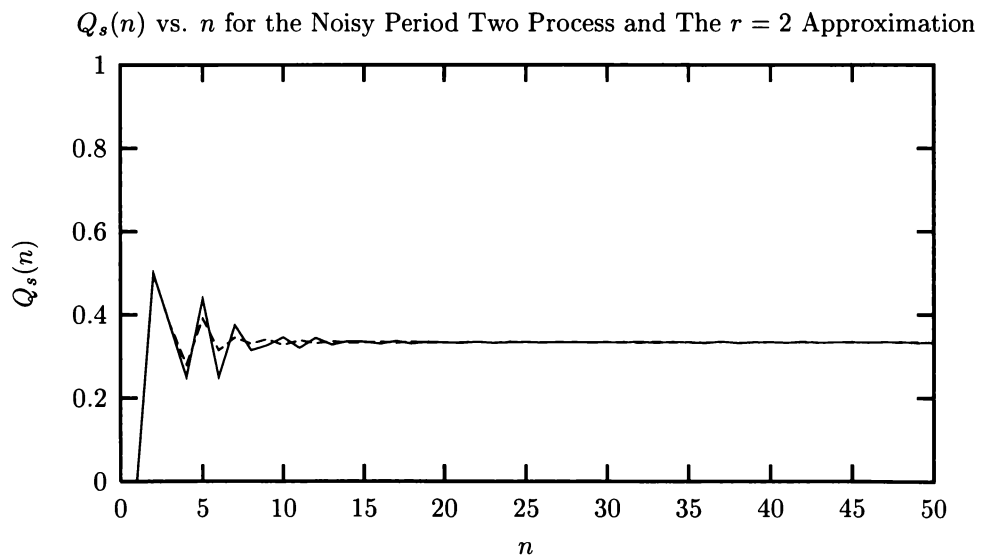


Figure 7.16: The  $Q_s(n)$  vs.  $n$  for the noisy period two process (solid line) and the  $r = 2$  approximation to the process (dashed line).

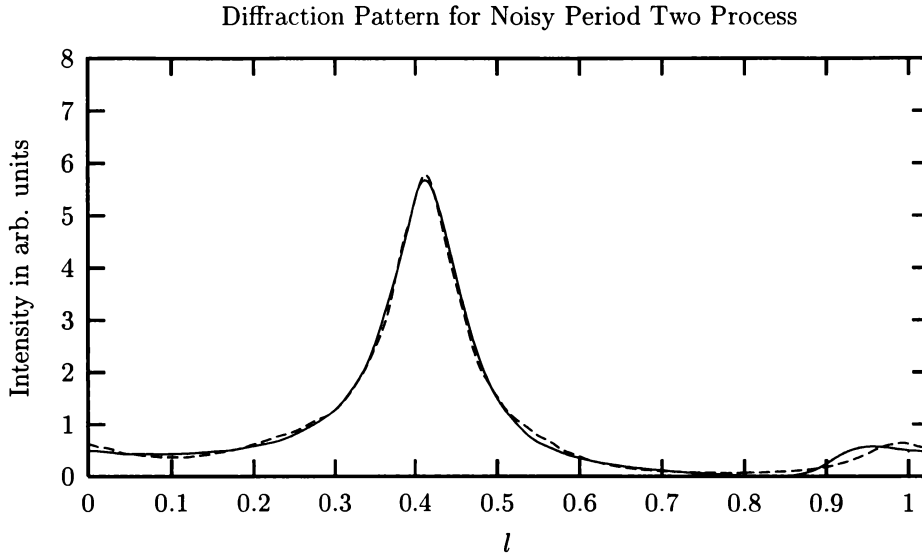


Figure 7.19: The diffraction pattern for a lattice stacked according to the noisy period two process (solid line) and the  $r = 3$  approximation (dashed line).

column of table 7.5, along with the word probabilities for the noisy period two in the last column. There are several curiosities. We found that the lower approximations were detecting the forbidden word ‘00’ and words constructed from it. At  $r = 3$ , some of these words have become ‘unforbidden.’ For example, ‘0100’, containing ‘00’ as a subword is forbidden by the noisy period two process, but  $r = 3$  approximation assigns it a small probability weight of 0.005. We do acquire a new irreducible forbidden word at  $r = 3$ , ‘0110’, but the  $r = 3$  approximation gives this a small probability weight of 0.010. Most troubling is the isolated node ‘0000’. While assigning only a small weight of 0.007, it is nonetheless isolated from the rest of the graph as  $p(0001) = p(1000) = 0.000$ . The same difficulty that plagued us in the 3C/2H/6H process arises here. Given the small weight, though, we chose to ignore it and proceed as if  $p(0000) = 0.000$ . A plot of  $Q_s(n)$  versus  $n$  for both the  $r = 3$  approximation and the noisy period two is given in figure 7.18. We see excellent agreement. There is similar excellent agreement between the diffraction patterns as shown in figure 7.19. The broad diffuse maximum is well-represented by the  $r = 3$  approximation, but the minimum at  $l = \frac{5}{6}$  is slightly off. There is a rise in the spectrum of the noisy period two that is not accounted for well in the  $r = 3$  approximation. From the example in §7.2, we see that the details of the diffuse background scattering are important, and the point is reinforced here.

Table 7.6 gives the computational results for the noisy period two and the four  $r$  approximations. As we might expect, the entropy density decreases with increasing  $r$ , except for the  $r = 2 \rightarrow 3$  step. There is a slight rise in  $h_\mu$ . This is counter-intuitive, and it is not known why  $h_\mu$  increases.

## 7.4 Machine Reconstruction for the Even Process

As a final example, let us consider another process that we will not be able to find using a finite  $r$  procedure. Instead, we expect to once again generate graphs of increasing size that approximate the process. The recurrent portion of the  $\epsilon$ -machine for this process is given in figure 5.41. The machine reconstruction results for the first four  $r$  approximations are shown in table 7.7.

Figure 7.20 shows the comparison between the  $r = 0$  approximation and the even system for

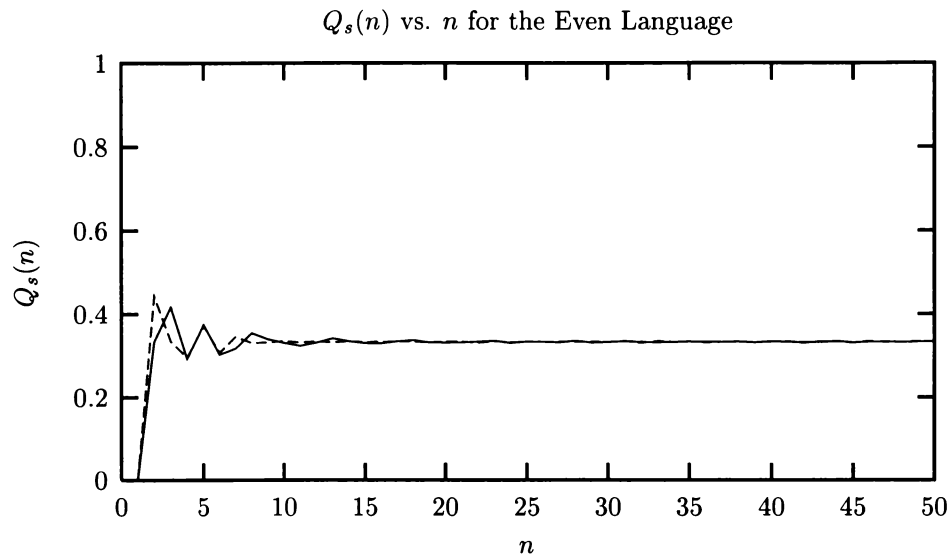


Figure 7.20: The  $Q_s(n)$  vs.  $n$  for the even process (solid line) and the  $r = 0$  approximation to the process (dashed line).

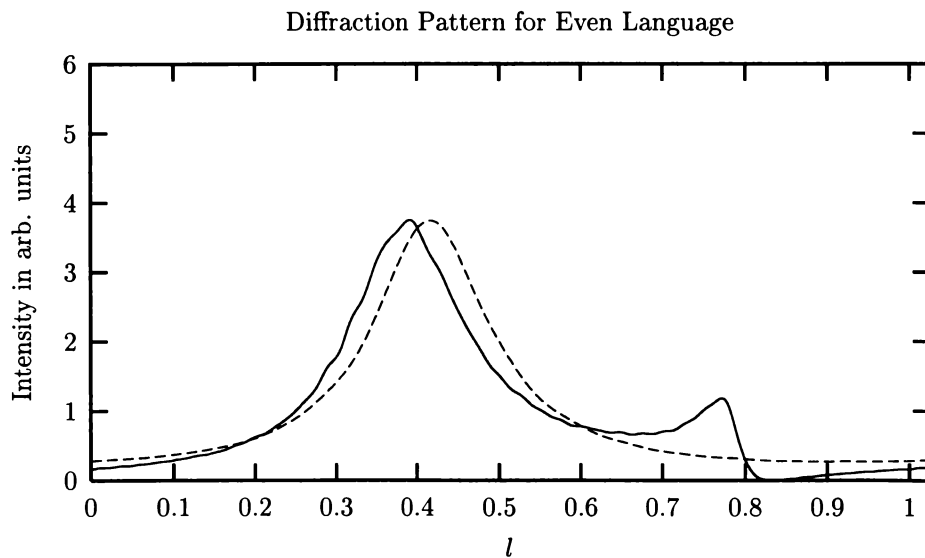


Figure 7.21: The diffraction pattern for a lattice stacked according to the even process (solid line) and the  $r = 0$  approximation (dashed line).

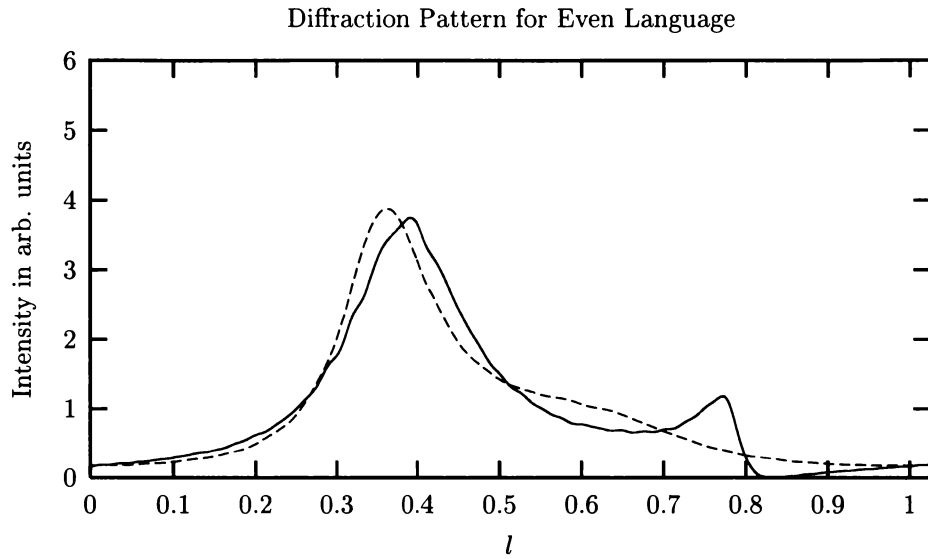


Figure 7.23: The diffraction pattern for a lattice stacked according to the even process (solid line) and the  $r = 1$  approximation (dashed line).

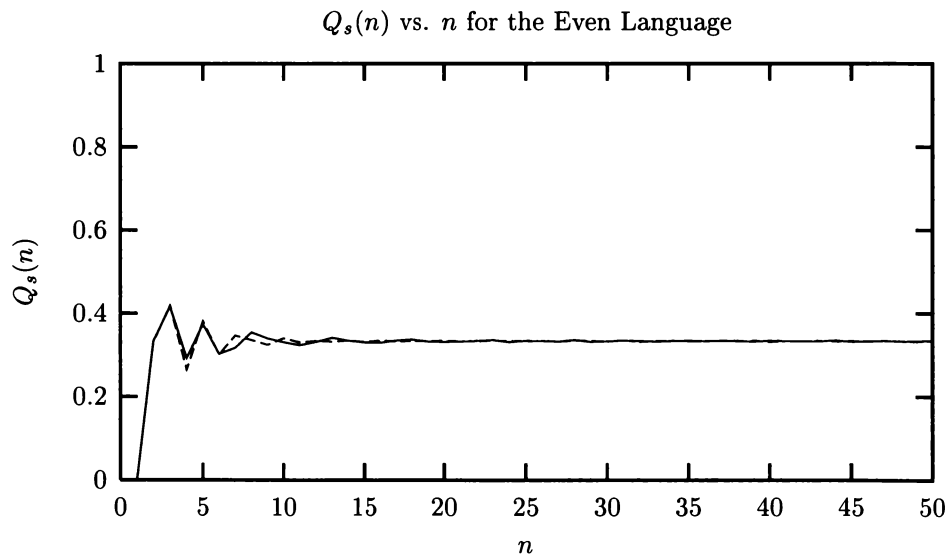


Figure 7.24: The  $Q_s(n)$  vs.  $n$  for the even process (solid line) and the  $r = 2$  approximation to the process (dashed line).



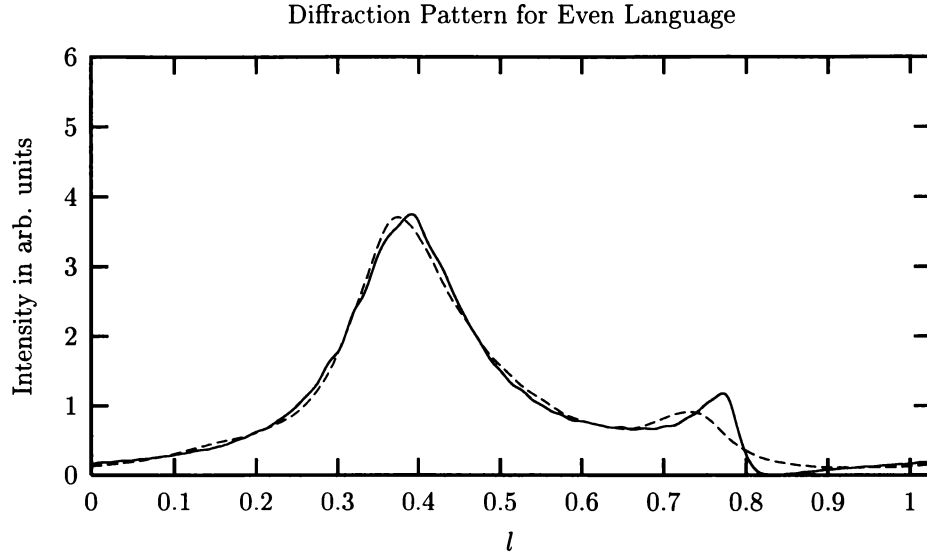


Figure 7.27: The diffraction pattern for a lattice stacked according to the even process (solid line) and the  $r = 3$  approximation (dashed line).

Table 7.8: Computational results for the even process and the  $r = 0, 1, 2$  and 3 approximations. As with the noisy period two process, there is a slight rise in the entropy density as  $r$  goes from 2 to 3.

System	Language Type	$r$	$h_\mu$	$\mathbf{G}$	$C_\mu$	$\mathbf{E}$	$\mathbf{T}$
Even	SS	$\infty$	0.667	0.333	0.918	0.913	3.09
$r = 0$ Approximation	SFT	0	0.918	0.082	0.0	0.0	0.000
$r = 1$ Approximation	SFT	1	0.874	0.126	0.918	0.044	0.044
$r = 2$ Approximation	SFT	2	0.792	0.208	1.79	0.208	0.290
$r = 3$ Approximation	SFT	3	0.803	0.197	2.63	0.222	0.351

results are given in table 7.8. We see the same decrease in entropy density as  $r$  increases except at  $r = 2 \rightarrow 3$ , where there is a slight increase in  $h_\mu$ .

## Chapter 8

# Previous Classifications of Disorder in Layered Materials

### 8.1 Overview

A zeroth order attempt to ‘classify’ disorder in physical systems is to acknowledge that the disorder exists, but give no further details. This is the approach taken in a recent paper on determination of the polytype distribution in SiC [26]. There the authors perform x-ray diffraction experiments on powder samples of SiC. For several samples, they discovered diffuse background scattering which they can not associate with any crystalline polytype. They can, however, determine the fraction of the scattered intensity diffracted into this background, and simply refer to this fraction as the percentage of disorder present. It is certainly an honest approach, akin to the admission of ignorance. We believe much more desirable though, is a statistical description of this disorder. Another approach with a long history is the assumption that there is a crystal structure present but that there are stacking errors, or faults present, which lead to the disorder. Typically, often guided by intuition, and ease of implementation, one assumes a certain number and kind of possible ways for the disorder to exist and calculates what effect this disorder can have on the diffraction pattern. Often this analysis is confined to considering only the effect of the assumed disorder on the Bragg peaks. We saw in §7.2 that this can be misleading. Proposed structures may account relatively well for the placement, intensity, shape, etc. of Bragg peaks but still not represent the underlying mechanism. It is important to take into account the intensity distribution over an entire unit interval. Recent work by Gosk [30] [31] does just this, but he still adheres to a picture of imposing *a priori* a select number and kind of possible faults. We have several difficulties with these approaches. The first is not specific enough, and the second requires assumptions that in general are not supportable and in fact may be misleading. A more serious objection of the second approach is our contention that the fundamental picture of faulting is deeply flawed. We discuss this in detail in §9.2. A final quibble is perhaps philosophical. Instead of needing to assume some underlying mechanism, we would rather make fewer initial assumptions and let the data more directly tell us about the underlying mechanism.

The zeroth order attempt at describing disorder is simple enough and needs no further exposition. We do wish, however, to examine the faulting picture more closely so that we may compare our results with previous work.

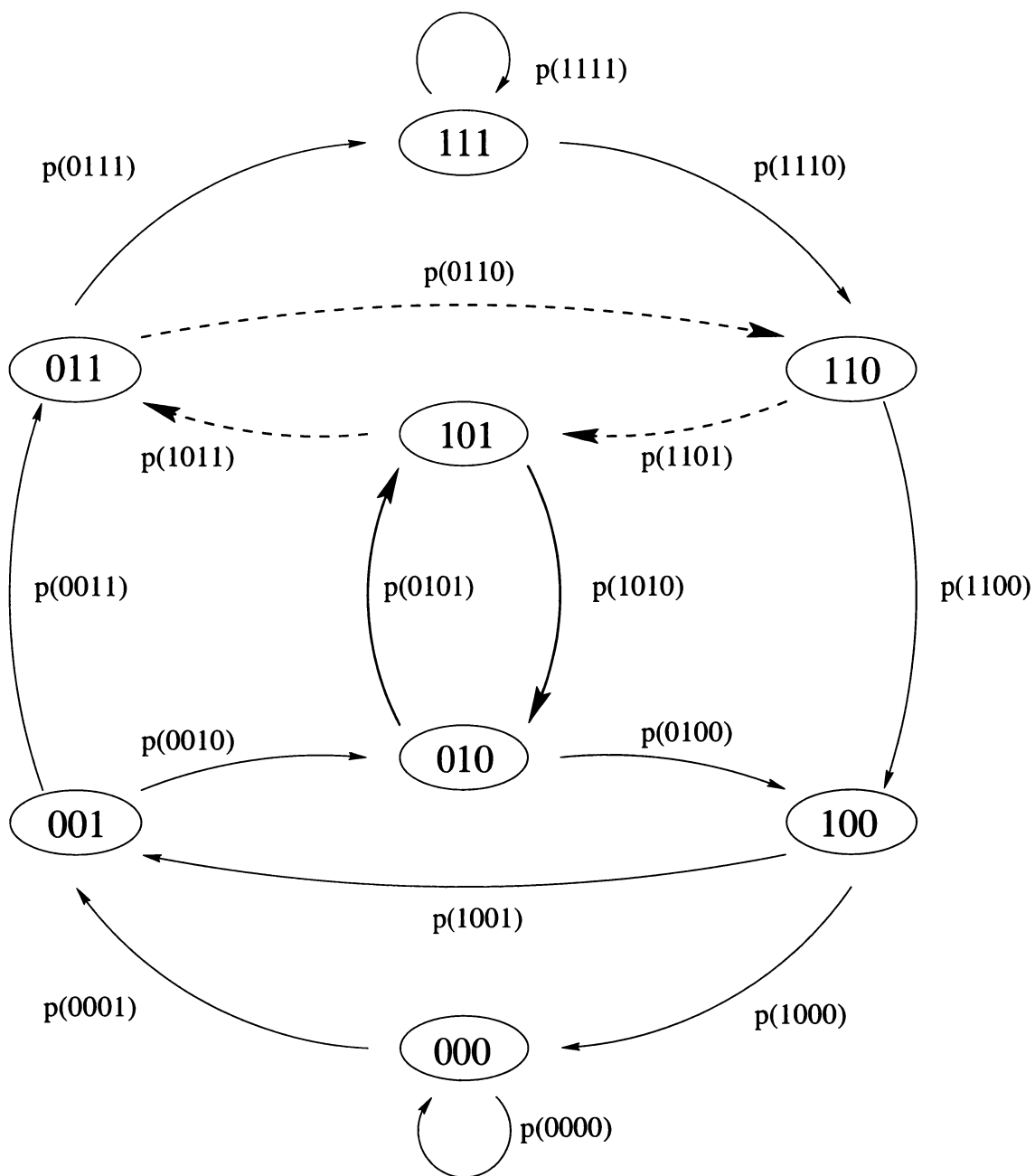


Figure 8.1: Growth faults in the 2H structure of ZnS depicted on an  $r = 3$  de Bruijn graph. The broad solid lines represent the 2H structure, the dashed lines are the growth faults, and the thin solid lines are the remainder of the graph. For convenience, only the faults in the upper portion of the graph are shown, corresponding to an insertion of a 1. In general, there are of course the spin inverse of these present (insertion of a 0), and these faults occupy the spin symmetric portion in the lower part of the graph. Growth faults can be seen on a  $r = 1$  graph.

$$\dots A B A B | C A C A C A \dots$$

The vertical bar indicates the plane across which the slip occurred. In terms of relative spins, a deformation fault in the 2H structure is realized by flipping a spin. In this example, the unfaulted sequence ...10101010... transforms to ...1011010..., where again the underlined character demarcates the flipped spin. The minimum size de Bruijn graph on which this fault can be demonstrated is  $r = 3$ . This is shown in figure 8.3.

In the 3C structure, deformation faults appear much the same. An example of a deformation fault in a 3C structure is

$$\dots A B C A B C | B C A B C A \dots$$

The vertical bar again indicates the slip plane. Expressed in relative spins, the unfaulted 3C crystal, ...11111111..., becomes ...11110111..., giving a single spin flip. This can be expressed on a  $r = 1$  de Bruijn graph; it is shown on a  $r = 3$  de Bruijn graph in figure 8.4.

### 8.2.3 Layer Displacement Faults in the 2H and 3C structure of ZnS

Layer displacement faults are characterized by a shifting of one or two layers in the crystal, while leaving the remainder of the crystal undisturbed. As such, these faults do not interrupt the long range order present in a structure. They are thought to be introduced at high temperatures by diffusion of the atoms through the crystal. Sebastian and Krishna [60] give a nice discussion of the possible mechanisms. In the 2H structure, an example of a layer displacement fault is:

$$\dots A B A B \underline{C} B A B A \dots$$

where the underlined layer is the faulted layer. Written as relative spins, ...10101010... becomes ...10110010..., the underlined characters indicating the the relative spins that have flipped. The minimal de Bruijn graph necessary to show this structure is  $r = 3$ , and this is displayed in figure 8.5.

Layer displacement faults in 3C structures are more difficult to realize, since each layer is sandwiched between two unlike layers and changing its orientation would violate stacking constraints. It is therefore necessary for two adjacent layers to shift. Consequently one might expect that these are more rare. An example of layer displacement is the following:

$$\dots A B C A B C \underline{B A} C A B C A \dots$$

where the underlined layers are faulted. The relative spin sequence changes from a series of all 1s to one where three consecutive spins have been flipped to 0. The minimal graph on which this can be demonstrated is  $r = 3$ , and this is shown in figure 8.6.

### 8.2.4 Additional Faulting Structures

The previous examples by no means exhaust all the possible faulting structures known or postulated to be important in polytypism of close-packed lattices. Additionally, one finds the double deformation fault for 3C structures, shown in figure 8.7. Another fault, unnamed and requiring a de Bruijn graph of  $r = 4$  is postulated in 2H structures. There are also faults that are believed to correspond to the removal or insertion of entire layers in the crystal. Called extrinsic faults, one mechanism for their creation is irradiation. They have high fault energies, and are thus rare. None of the samples considered in this work have been irradiated, so it is not discussed further here.



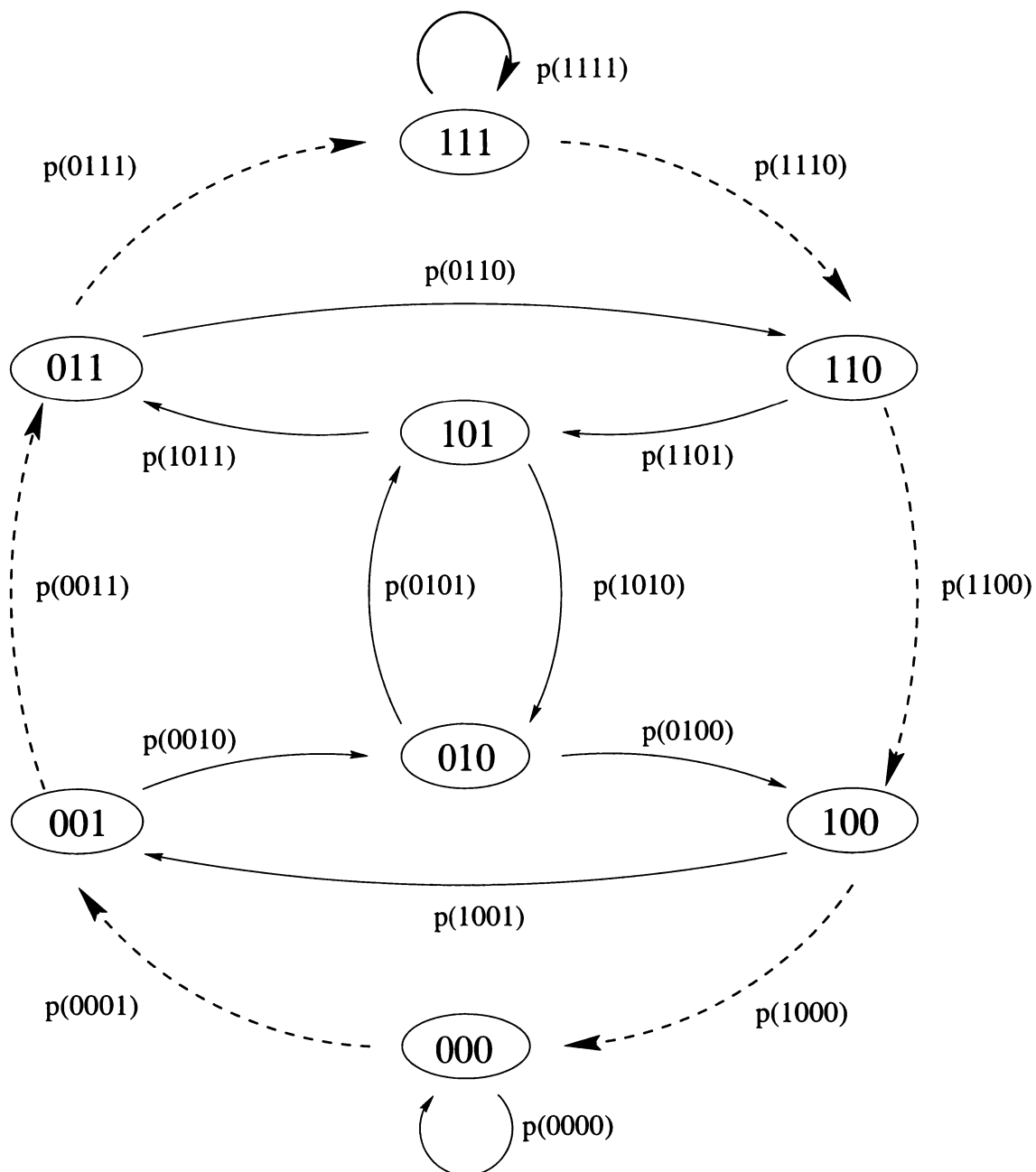


Figure 8.6: Layer displacement faults in the 3C structure of ZnS depicted on an  $r = 3$  de Bruijn graph. The broad solid lines represent the 3C structure (only the positive chirality structure (1)\* shown), the dashed lines are the layer displacement faults, and the thin solid lines are the remainder of the graph.

## Chapter 9

# $\epsilon$ -Machine Reconstruction from Experimental ZnS Diffraction Patterns

Now that the work of previous researchers has been discussed and new theoretical tools and procedures have been introduced, we are equipped to examine experimental spectra and discover the underlying process which gives rise to them. We will examine seven diffraction patterns for single crystal ZnS found in reference [60]. They will be referred to by the page number on which they appear in [60]. Even though the data was taken in the mid 1980s, it unfortunately is only available in graphical form [55]. It was therefore necessary to scan the graphs and digitize the data.

A brief word about significant figures in the following. While most quantities are calculated to a precision of three decimal places, this in no way implies that any quantity is this well known. Recall that the spirit of this work is exploratory, and we are attempting to expound a new technique for the characterization and discovery of patterns in the stacking order of ZnS. As such, we do not want the clarity and power of the exposition to be lost in too little precision. No error analysis has been attempted, but if one were performed, the error would undoubtedly be large.

### 9.1 Machine Reconstruction from Experimental Diffraction Pattern SK229

The diffraction pattern along the  $10.l$  row for an as grown 2H ZnS crystal is shown in figure 9.1 and the diffraction pattern corrected for  $\mathcal{C}(l)$  is displayed in figure 9.2. We immediately notice in figure 9.2 that the pattern is not periodic in  $l$ , but instead suffers from variations in the intensity. We see that the peaks at  $l = -\frac{1}{2}$  and  $\frac{1}{2}$ , are of similar intensity, but the peak at  $l = \frac{3}{2}$  seems to be have about roughly one-half their brightness. The peaks at  $l = 0$  and  $1$  also differ in their intensity. So we can be sure that this spectrum contains substantial error, and we will find this to plague the other spectra as well. As discussed in §4.2, there are criteria we can use to help select a unit interval over which the intensity seems not to vary too much. Looking at the spectrum in figure 9.2, we might expect this interval to be between  $l = -\frac{1}{2}, \frac{1}{2}$ , and we indeed find that choosing  $l_0 = -0.33$  gives reasonable figures of merit, namely  $Q_s(1) = 0.008$ ,  $\gamma = -0.489$  and  $\alpha = 1.004$ . The first few correlation functions found by integrating over this interval are shown in table 9.1. Values near one-half for the  $Q_c(n)$  and  $Q_r(n)$  with odd  $n$  are what we expect for a disordered 2H crystal. Calculating the correlation length for the  $Q_s$  up to  $n = 40$ , we find it to be  $\lambda_q = 19 \pm 2$ . In figure 9.3 a plot of  $Q_s(n)$  vs.  $n$  is given that shows the oscillatory behavior in  $Q_s(n)$  with period two over the

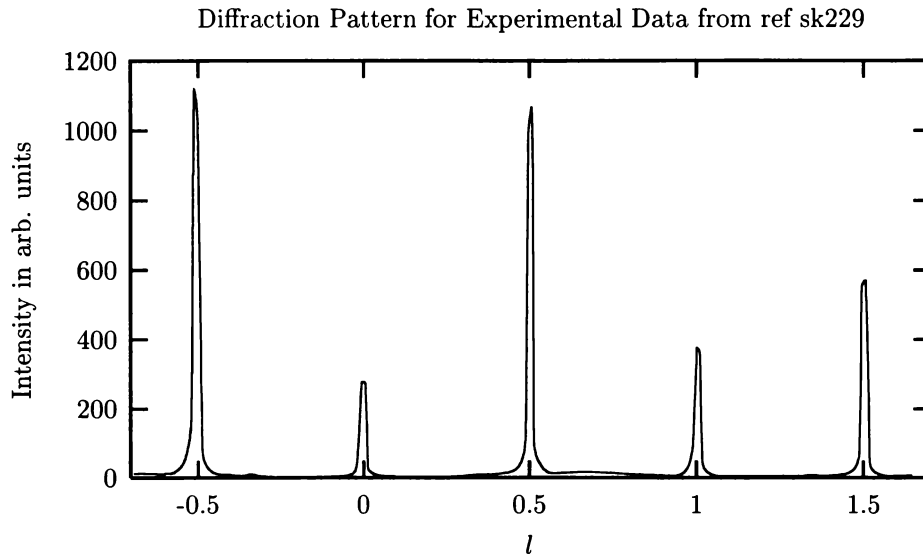


Figure 9.2: Diffraction pattern for Experimental Data SK229 corrected for  $\mathcal{C}(l)$ .

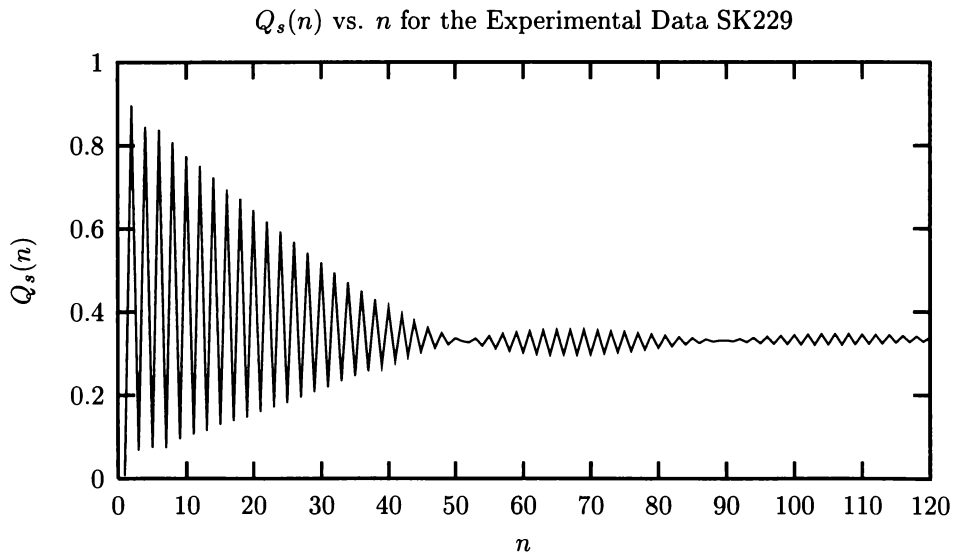


Figure 9.3: The  $Q_s(n)$  vs.  $n$  for Experimental Data SK229. We use  $l_0 = -0.330$  and get a value of  $\gamma = -0.489$  and  $\alpha = 1.004$ . We find a correlation length of  $\lambda_q = 19 \pm 2$  over the first forty layers.



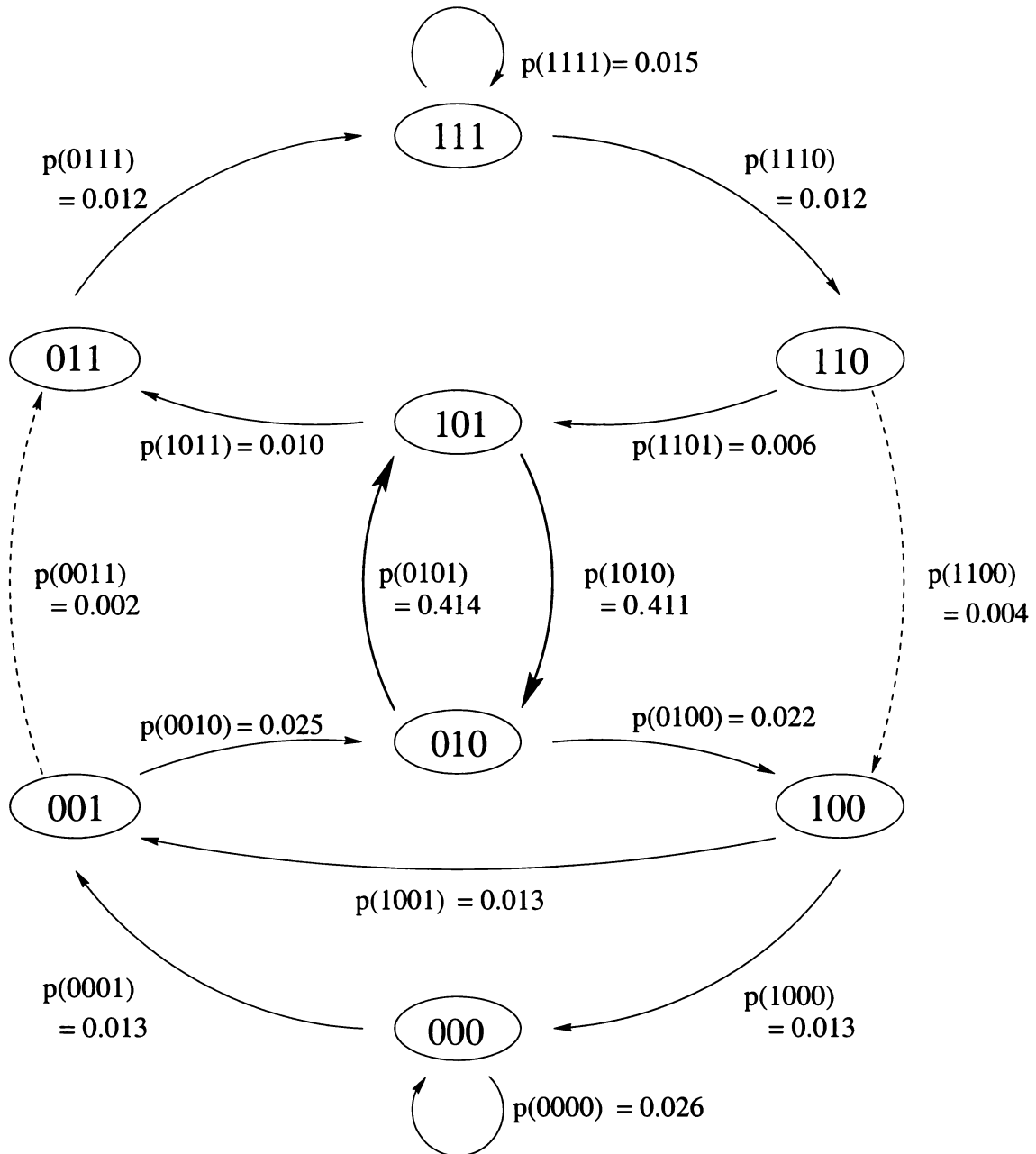


Figure 9.4: The  $r = 3$  machine for diffraction pattern SK229. The bold arcs correspond to the 2H crystal structure and the dashed arcs have so little probability that we can take them to be zero for the purpose of decomposing the graph into a crystal and faults. The missing 0110 arc indicates that this word is absent.

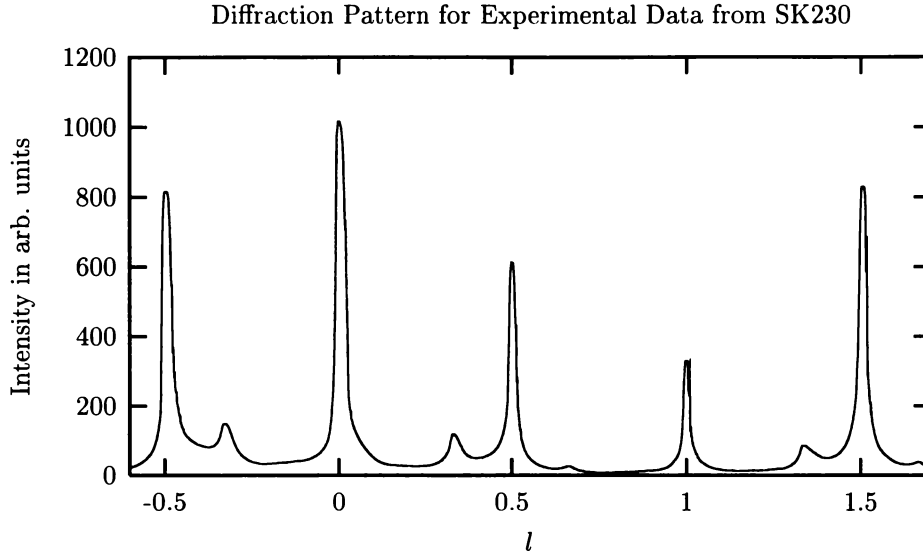


Figure 9.7: Uncorrected diffractometer pattern of intensity vs.  $l$  along the  $10.l$  row for a perfect 2H crystal of ZnS after annealing for one hour at 300 C. Experimental Data SK230.

Table 9.4: The first few  $Q$ s for experimental data SK230.

$n$	$Q_c(n)$	$Q_r(n)$	$Q_s(n)$
1	0.467	0.524	0.009
2	0.164	0.077	0.759
3	0.398	0.439	0.163
4	0.214	0.124	0.662
5	0.396	0.450	0.154
6	0.242	0.106	0.652

## 9.2 Machine Reconstruction from Experimental Diffraction Pattern SK230

Figure 9.7 shows the diffraction pattern along the  $10.l$  row of a perfect 2H ZnS crystal that has been annealed at 300 C for one hour. The same diffraction pattern corrected for  $C(l)$  is shown in figure 9.8. Again we notice that the intensity is not periodic in  $l$ , but appears to fall off slowly as  $l$  increases. The intensity seems to have been reduced by a factor of two on the interval  $-0.5 \leq l \leq 1.5$ . Integrating over the interval  $-0.63 \leq l \leq 0.37$  we find the figures of merit to be  $Q_s(n) = 0.009$ ,  $\gamma = -0.486$  and  $\alpha = 1.022$ . The first few correlation functions found from integrating over this interval are shown in table 9.4. We can see values of  $Q_c(n), Q_r(n)$  near one-half for  $n$  odd, suggesting that the original 2H structure is not too corrupted. This is reasonable considering strong peaks at integer and half integer  $l$ . Considering the first forty layers we get a correlation length of  $\lambda_q = 8.4 \pm 0.4$ .

The results for machine reconstruction appear in table 9.5 and figure 9.10. We notice immediately two missing arcs, 0011 and 1001. Again the probability weights on the arcs corresponding to the 2H cycle are large, the sum of  $p(0101)$  and  $p(1010)$  being 60.9%. The remaining 30% arc weight can be

Table 9.5: Machine reconstruction results for the experimental diffraction pattern SK230. The fitness for the  $r = 3$  solution is  $\mathcal{F} = 1.01 \times 10^{-5}$ .

$r = 0$	$p(0)$	0.528	$r = 3$	$p(0000)$	0.066
	$p(1)$	0.472		$p(0001)$	0.049
$r = 1$	$p(00)$	0.164	$p(0010)$	0.049	
	$p(01)$	0.380	$p(0011)$	0.000	
	$p(10)$	0.380	$p(0100)$	0.026	
	$p(11)$	0.077	$p(0101)$	0.316	
$r = 2$	$p(000)$	0.118	$p(0110)$	0.031	
	$p(001)$	0.046	$p(0111)$	0.005	
	$p(010)$	0.347	$p(1000)$	0.049	
	$p(011)$	0.032	$p(1001)$	0.000	
	$p(100)$	0.046	$p(1010)$	0.293	
	$p(101)$	0.334	$p(1011)$	0.036	
	$p(110)$	0.032	$p(1100)$	0.022	
	$p(111)$	0.045	$p(1101)$	0.011	
			$p(1110)$	0.005	
			$p(1111)$	0.040	

relegated to disorder. Again we can attempt to understand this disorder in terms of stacking faults. One possible break down is to say that the simple cycle  $1011 \rightarrow 0111 \rightarrow 1110 \rightarrow 1101$  is a deformation fault and assign to each arc the same weight. This can be done if each is given the weight 0.005. We might then want to identify the cycle  $1011 \rightarrow 0110 \rightarrow 1101$  as a growth fault, and assign the weight of 0.006 to each arc. If we then follow the cycle  $1011 \rightarrow 0110 \rightarrow 1100 \rightarrow 1000 \rightarrow 0001 \rightarrow 0010$  this appears to be a deformation fault, with each arc bearing a weight of approximately 0.022. We can finally round out the faults by taking the cycle  $0100 \rightarrow 1000 \rightarrow 0001 \rightarrow 0010$  as a deformation fault with each arc taking a weight of about 0.026. This is a consistent breakdown of the graph into crystalline and faulted cycles, but it is not the only possible such decomposition.

We could, for instance, take the position that there is no (single) deformation fault on the upper part of the graph by instead treating the cycle  $1011 \rightarrow 0111 \rightarrow 1110 \rightarrow 1100 \rightarrow 1000 \rightarrow 0001 \rightarrow 0010$  as a fault in its own right. One can find a logically consistent distribution of stacking faults under this assumption that differs from the previous analysis. We then are faced with a situation of a single graph giving rise to two different faulting configurations, and should well question the validity and usefulness of the fault picture. It seems rather that the graph is the thing, being concise and unique at each  $r$ . It is perhaps not as intuitive as thinking of a crystal permeated with a certain fraction of stacking ‘errors’, but it is unambiguous. We consider this no mere semantic quibble either. Where the faulting picture becomes less tenable as the fraction of stacking faults increases, the graphical picture seamlessly handles any amount of disorder, from a near perfect crystal to complete randomness and everything in between. Indeed, the graph, or  $\epsilon$ -machine, provides the minimum structure needed to specify the statistics of the stacking. In short, it is the answer. Our interpretation of the graph is then a matter of convenience and perhaps psychology, but a description of disordered structures based on ‘faultology’ is not fundamental.

This point can be made more quantitative by a careful treatment of our procedure for assigning stacking faults. It is always possible to break down an infinite sequence generated by de Bruijn graph into a sequence of simple cycles. This decomposition however is not unique. Since a simple cycle is just the indefinite repetition of a finite length sequence, we see that this just corresponds to some crystal structure. A graph, of course, may have more than one simple cycle. If there are two such cycles sharing, say, a node, and the conditional probability to branch at this node strongly

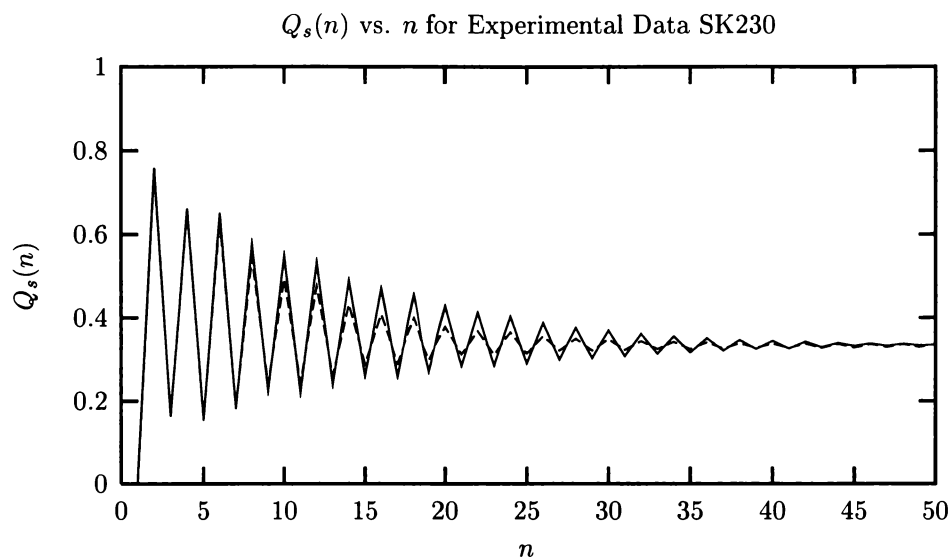


Figure 9.11: The  $Q_s(n)$  vs.  $n$  for Experimental Data SK230 (solid line) and the  $r = 3$  approximation (dashed line).

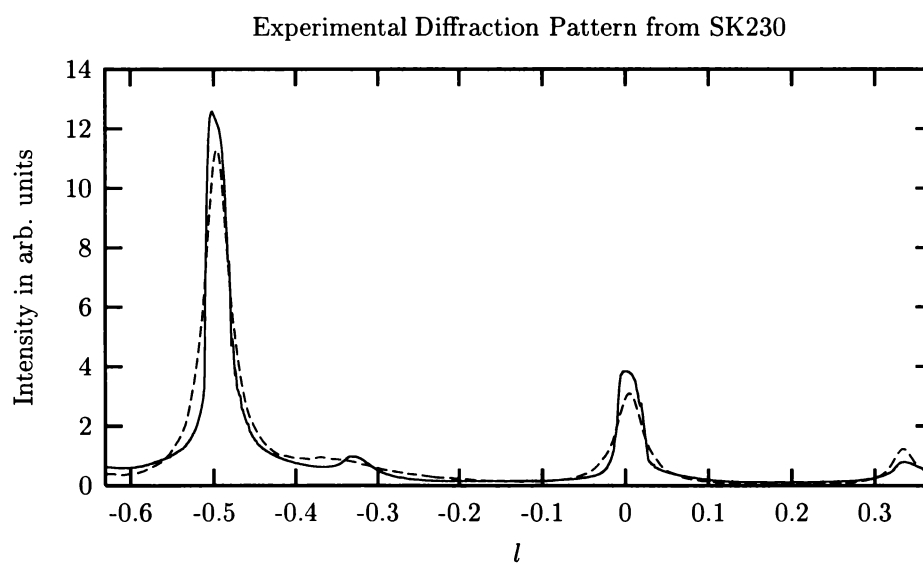


Figure 9.12: The diffraction pattern for Experimental Data SK230 and the  $r = 3$  approximation.

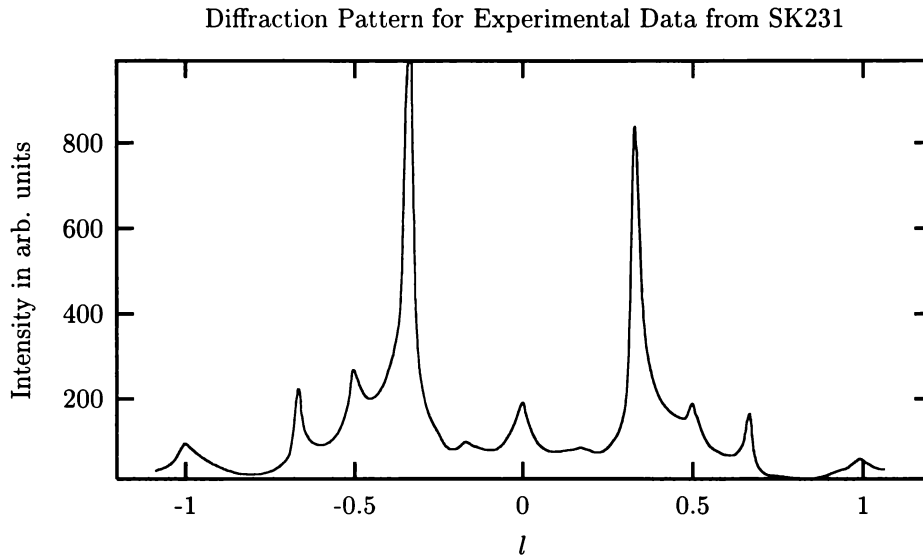


Figure 9.13: Uncorrected single crystal diffractometer reading along the  $10.l$  reciprocal lattice row after annealing a 2H ZnS crystal at 600 C for one hour. SK231 experimental data.

a likely cause is undiscovered computation in the process. Such order might be manifest in higher  $r$  diagrams. Indeed there is speculation that the transformation to the 6H structure (111000)\* is important in annealed crystals. In order to detect competition between the 3C and 6H structures it is important that there be a ‘barrier’ between them, that is their cycles must not share a node. For  $r = 3$ , they share the 111 node; so, by construction, the proposed process lacks the necessary richness to model this. There is such a barrier in  $r = 4$  graph, as seen in §7.1, 7.2, and we have hopes that the competition between these polytypes can be detected there. That is, however, beyond the scope of this current work. In figure 9.12, we observe that the agreement between the experimental diffraction pattern and the diffraction pattern from the  $r = 3$  approximation to it is less than satisfactory. Again, smeared out peaks is what we expect if the process does not contain sufficient structure to model the computation present.

### 9.3 Machine Reconstruction from Experimental Diffraction Pattern SK231

In figure 9.13 we see the diffraction pattern along the  $10.l$  row for a ZnS crystal annealed at 600 C for one hour and figure 9.14 shows this same pattern after correcting for  $\mathcal{C}(l)$ . This pattern also suffers from a lack of periodicity in  $l$  although the symmetry present leads us to hope that the gradual falling off of intensity is not too great over either the interval  $-1.0 \leq l \leq 0.0$  or  $0.0 \leq l \leq 1.0$ . Indeed, the figures of merit over the former interval are rather good, being  $Q_s(n) = 0.000$ ,  $\gamma = -0.500$ , and  $\alpha = 1.024$ . The first few correlation functions found by integration over this interval are given in table 9.8. We notice that these first few  $Q_s$ s, save those at  $n = 1$ , are confined to a rather narrow interval of  $0.244 \leq Q \leq 0.487$ . Compared with previous patterns, this one doesn’t seem to have as much variation, the  $Q_s$  having more in common with a random number generator. We might guess that the original 2H structure is largely obliterated, and in fact the total weight assigned to the 1010 and 0101 arc is 14.8%. This is not much more than one sees in a random number generator

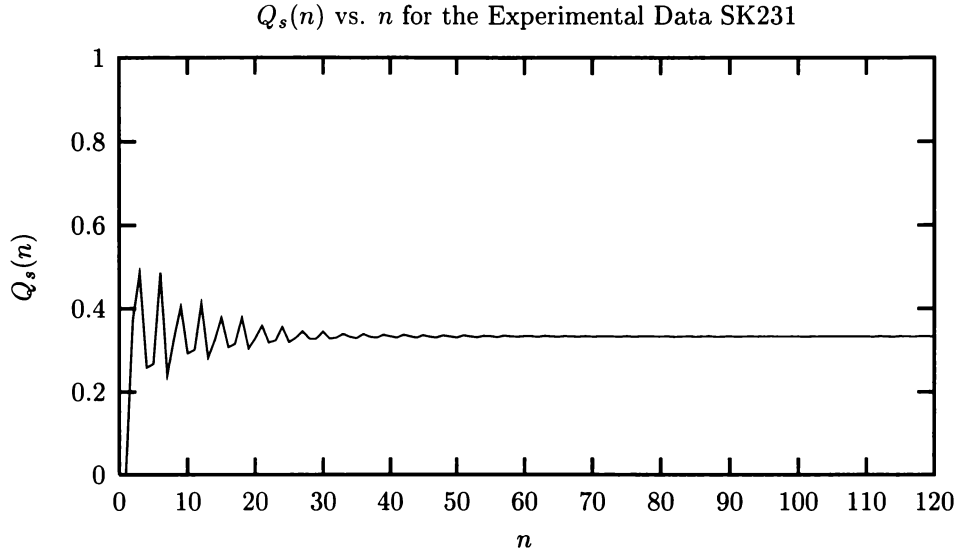


Figure 9.15: The  $Q_s(n)$  vs.  $n$  for Experimental Data SK231. We use  $l_0 = -1.000$  and get a value of  $\gamma = -0.500$  and  $\alpha = 1.024$ . We find a correlation length of  $\lambda_q = 6.8 \pm 1.7$  over the first twenty  $n$ .

Table 9.9: Machine reconstruction results for the experimental diffraction pattern SK231. We find a fitness of  $\mathcal{F} = 2.52 \times 10^{-4}$  for the  $r = 3$  solution.

$r = 0$	$p(0)$	0.498	$r = 3$	$p(0000)$	0.214
	$p(1)$	0.502		$p(0001)$	0.041
$r = 1$	$p(00)$	0.324	$p(0010)$	0.056	
	$p(01)$	0.186	$p(0011)$	0.011	
	$p(10)$	0.186	$p(0100)$	0.038	
	$p(11)$	0.304	$p(0101)$	0.084	
$r = 2$	$p(000)$	0.251	$p(0110)$	0.040	
	$p(001)$	0.072	$p(0111)$	0.028	
	$p(010)$	0.116	$p(1000)$	0.043	
	$p(011)$	0.071	$p(1001)$	0.024	
	$p(100)$	0.072	$p(1010)$	0.064	
	$p(101)$	0.114	$p(1011)$	0.057	
	$p(110)$	0.071	$p(1100)$	0.029	
	$p(111)$	0.233	$p(1101)$	0.040	
			$p(1110)$	0.036	
			$p(1111)$	0.193	

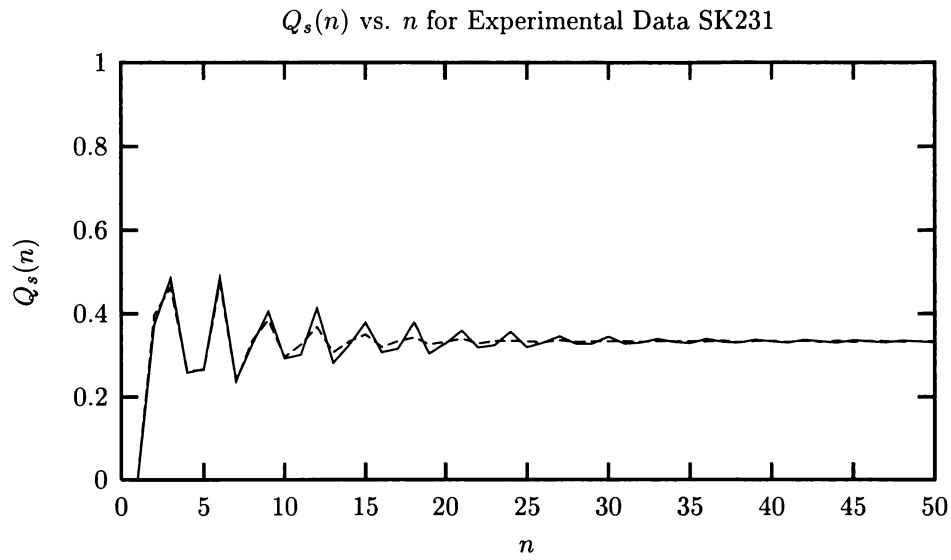


Figure 9.17: The  $Q_s(n)$  vs.  $n$  for Experimental Data SK231 (solid line) and the  $r = 3$  approximation (dashed line).

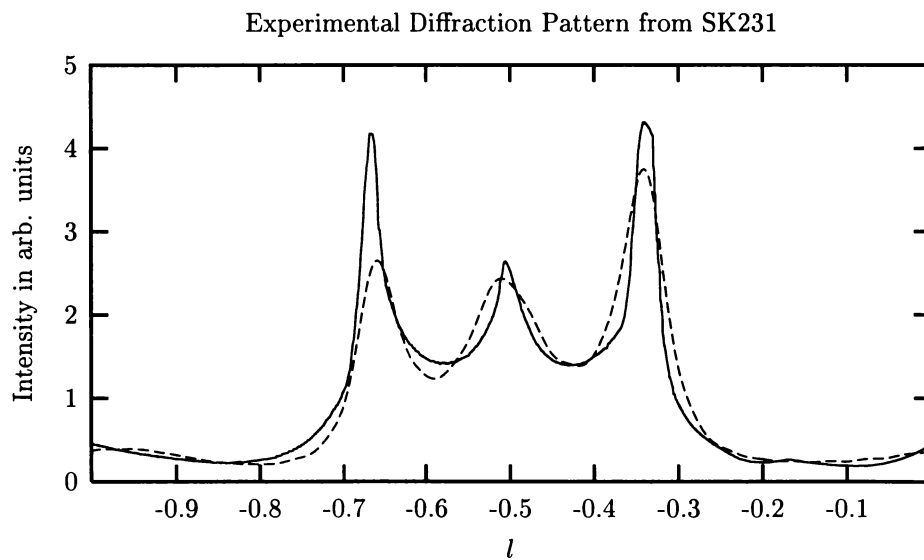


Figure 9.18: The diffraction pattern for Experimental Data SK231 and the  $r = 3$  approximation.

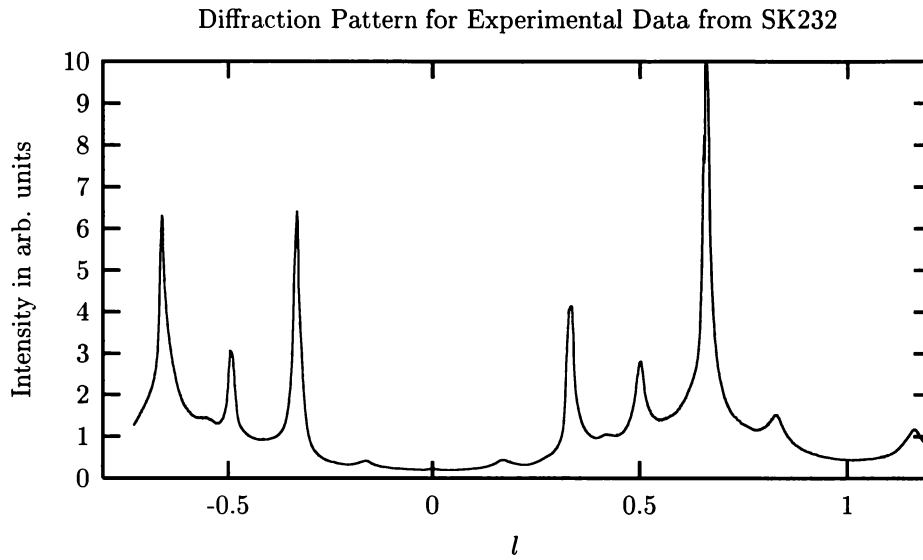


Figure 9.20: The diffraction pattern for Experimental Data SK232 corrected for  $C(l)$ .

Table 9.11: The first few  $Q$ s for experimental data SK232 found by integration over the interval  $-0.72 \leq l \leq 0.28$ .

$n$	$Q_c(n)$	$Q_r(n)$	$Q_s(n)$
1	0.542	0.445	0.014
2	0.289	0.406	0.305
3	0.274	0.211	0.515
4	0.388	0.387	0.225
5	0.358	0.397	0.245
6	0.260	0.214	0.526



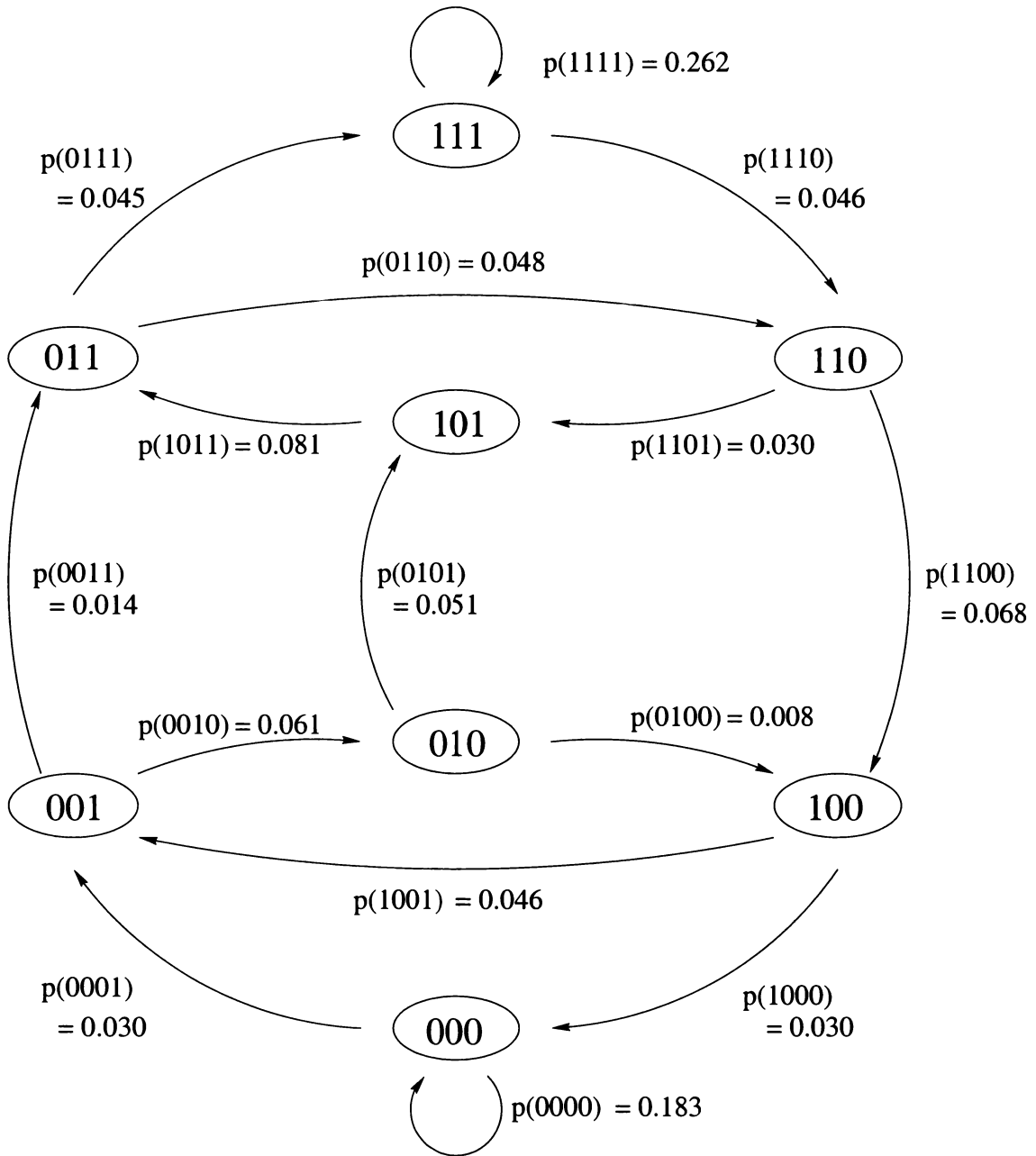


Figure 9.22: The  $r = 3$  machine for diffraction pattern SK232.

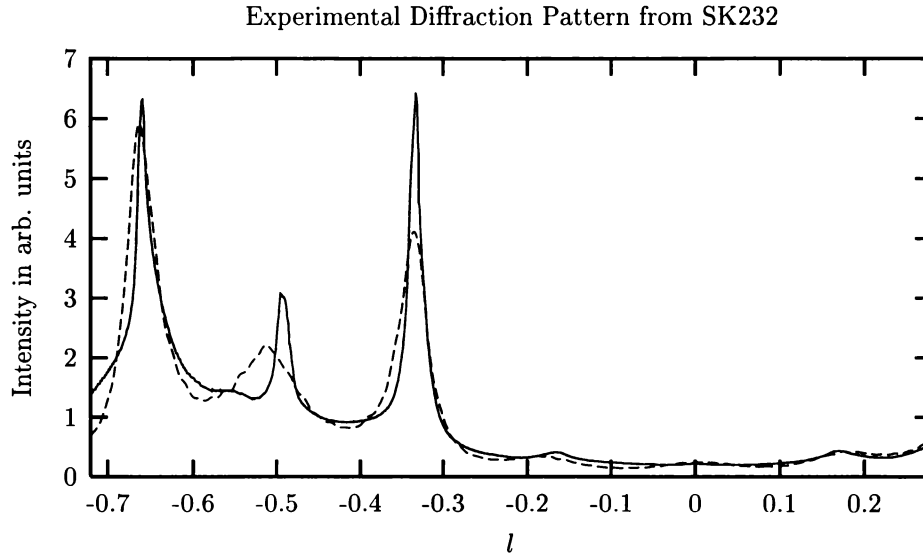


Figure 9.24: The diffraction pattern for Experimental Data SK232 and the  $r = 3$  approximation.

## 9.5 Machine Reconstruction from Experimental Diffraction Pattern SK134

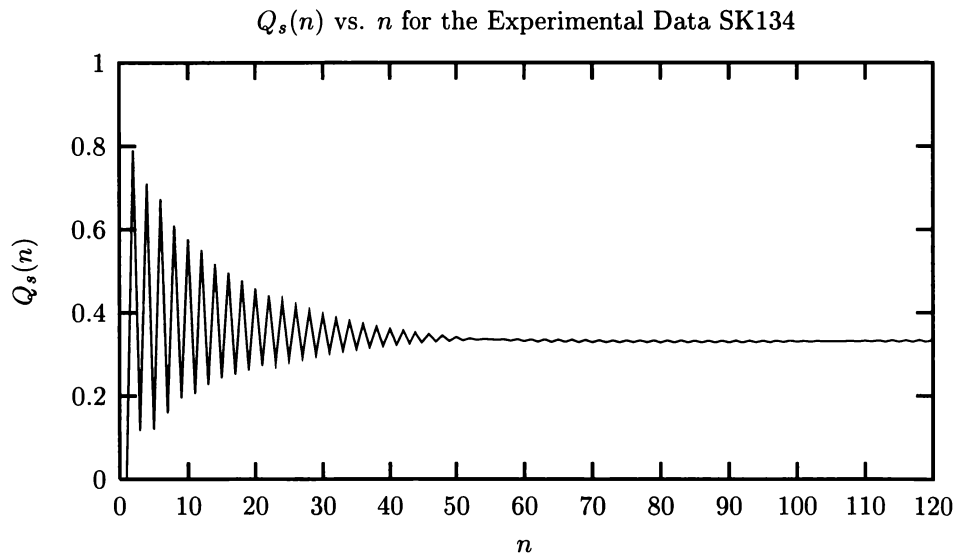
Figure 9.25 shows the diffraction pattern along the  $10.l$  row of a ZnS crystal obtained from annealing a perfect 2H crystal at 300 C for one hour. The same diffraction pattern is shown in figure 9.26 after correcting for  $\mathcal{C}(l)$ . Like the other diffraction patterns from this series, there is no strict periodicity in  $l$  but rather a slow degradation in the intensity as one moves from left to right across the spectrum. We use the same standard criteria to help select a unit interval to analyze, and we find that integration over the interval  $0.04 \leq l \leq 1.04$  gives the figures of merit to be  $Q_s(1) = -0.006$ ,  $\gamma = -0.509$ , and  $\alpha = 0.948$ . The first few correlation functions are shown in table 9.14 and a plot of  $Q_s(n)$  vs.  $n$  is shown in figure 9.27. The correlation functions show large and roughly equal values for the  $Q_c(n)$ ,  $Q_r(n)$  for  $n$  odd, suggesting that this crystal retains much of its original 2H character. Indeed, examining the machine reconstruction results in table 9.15 we see that arcs 0101 and 1010 together comprise 65.5.% of the probability weight for the total graph. The fitness for this machine reconstruction is  $\mathcal{F} = 5.25 \times 10^{-5}$ . We find a correlation length of  $\lambda_q = 9.5 \pm 0.5$  over the first forty layers.

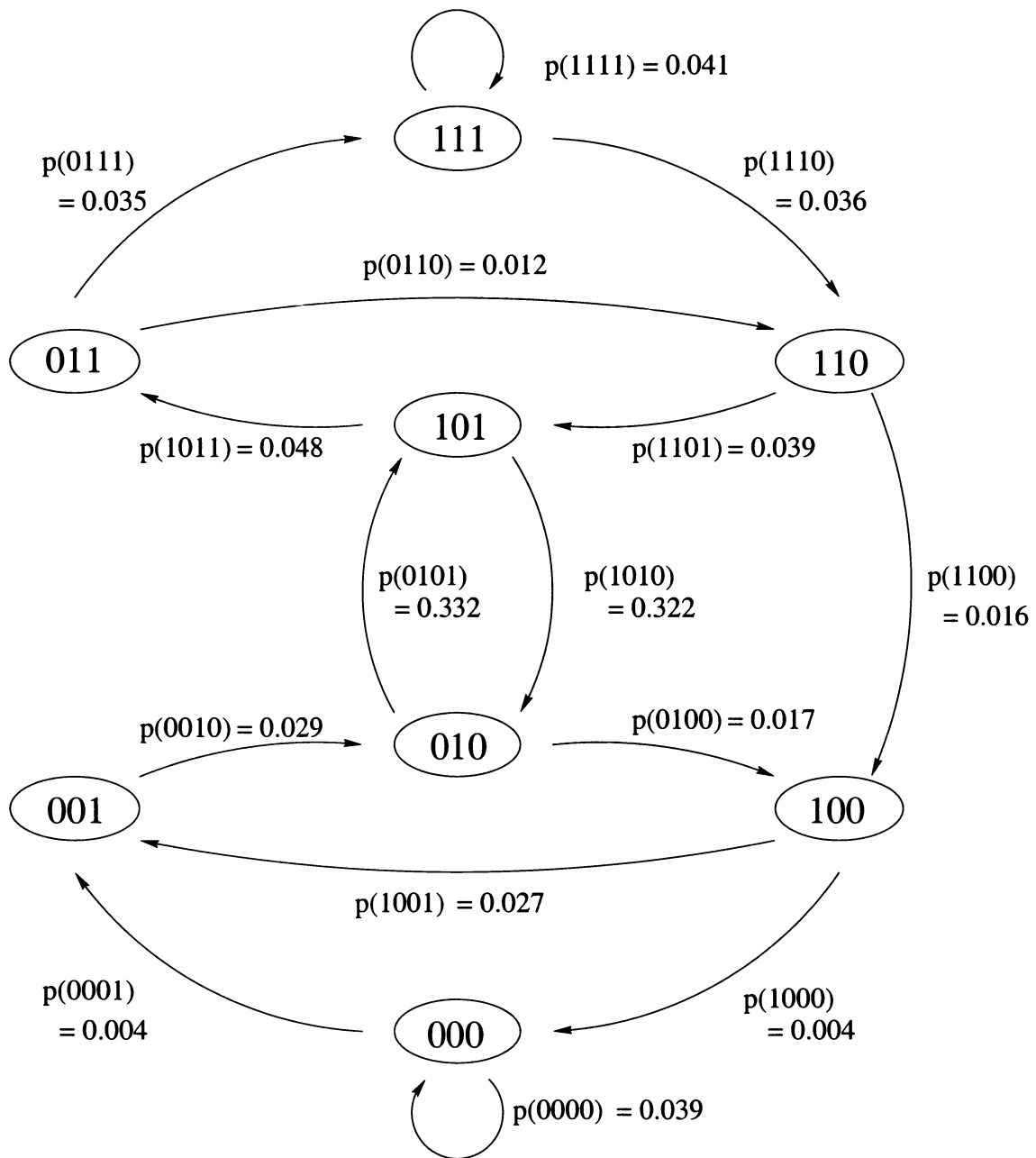
In most respects we see that this spectrum is much like SK230. Both began as perfect 2H crystals and have been annealed at 300 C for one hour. Both retain much of their much original 2H structure, (60.9% and 65.5% respectively) and have similar correlation lengths (8.4 and 9.5 respectively). The rate of entropy production per layer is similar ( $h_\mu = 0.487$  and 0.501 respectively). This speaks well of the consistency of our technique.

Figure 9.29 shows a comparison of the  $Q_s(n)$  obtained from the experimental diffraction pattern and that obtained from the  $r = 3$  approximation. Differences become apparent around  $n \approx 16$ , with the approximate machine underestimating the correlations for larger  $n$ . This is behavior similar to that which we saw in SK230, except there the disagreement began at  $n \approx 10$ . Sebastian and Krishna [60] attribute the disorder in both cases to deformation faulting, but in this present case they give a specific probability for this faulting, namely 5%. They arrive at this value by considering in some detail the change in the shape, placement, etc of the peaks. They examine several candidate

Table 9.14: The first few  $Q_s$  for experimental data SK134 found by integration over the interval  $0.04 \leq l \leq 1.04$ .

$n$	$Q_c(n)$	$Q_r(n)$	$Q_s(n)$
1	0.502	0.504	-0.006
2	0.077	0.133	0.790
3	0.475	0.408	0.117
4	0.093	0.197	0.710
5	0.478	0.402	0.120
6	0.116	0.211	0.673

Figure 9.27: The  $Q_s(n)$  vs.  $n$  for Experimental Data SK134. We use  $l_0 = 0.040$  and get a value of  $\gamma = -0.509$  and  $\alpha = 0.948$ . We find a correlation length of  $\lambda_q = 9.5 \pm 0.5$  over the first forty  $n$ .

Figure 9.28: The  $r = 3$  machine for diffraction pattern SK134.

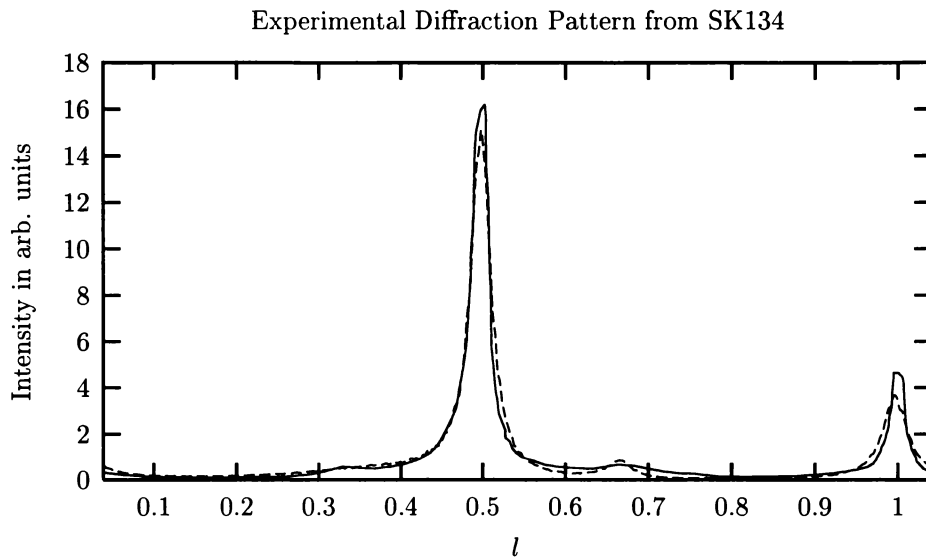


Figure 9.31: The diffraction pattern for Experimental Data SK134 (solid line) and the  $r = 3$  approximation (dashed line).

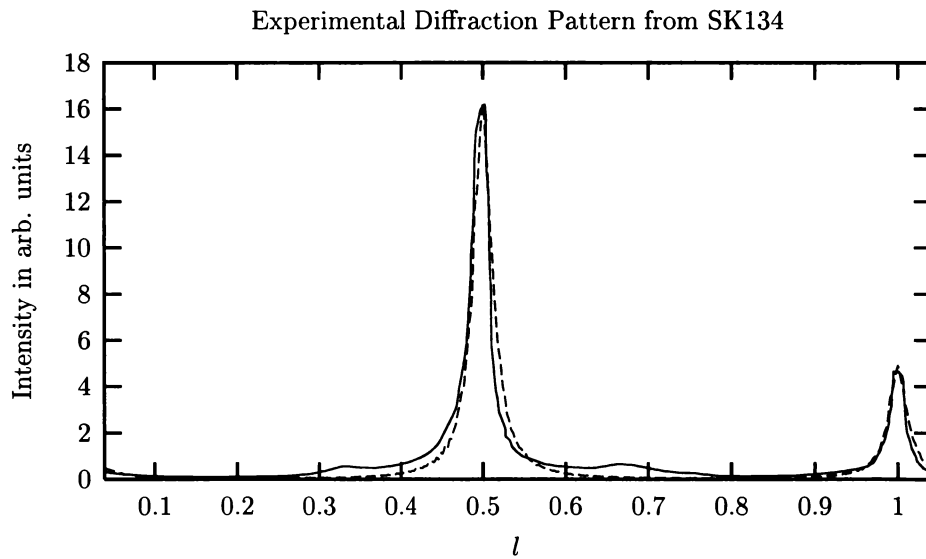


Figure 9.32: The diffraction pattern for Experimental Data SK134 (solid line) and the fault model (dashed line).

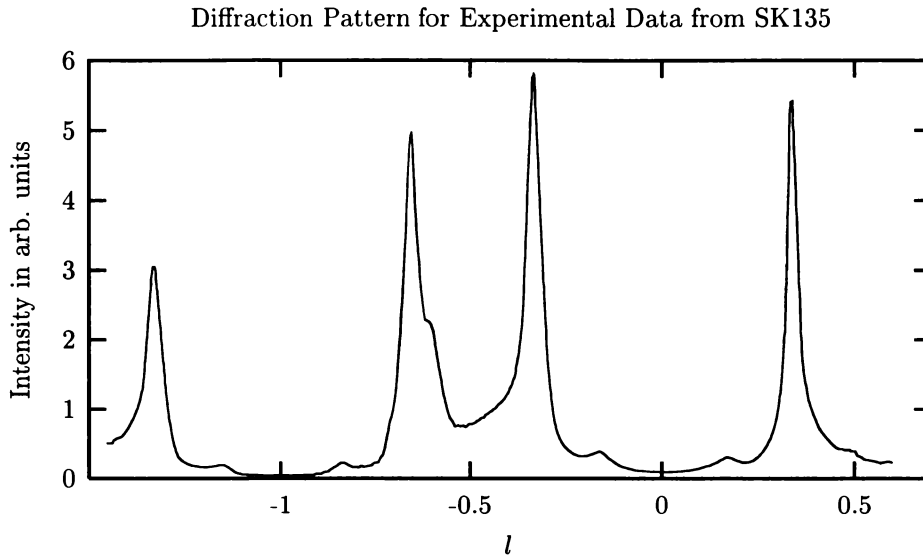


Figure 9.34: Corrected diffraction pattern for Experimental Data SK135.

Table 9.17: The first few  $Q_s$  for experimental data SK135.

$n$	$Q_c(n)$	$Q_r(n)$	$Q_s(n)$
1	0.475	0.523	0.002
2	0.376	0.395	0.229
3	0.228	0.139	0.633
4	0.401	0.480	0.120
5	0.355	0.354	0.291
6	0.283	0.198	0.519

as a disordered 3C crystal, due to the presence of asymmetrically broadened peaks and the absence of peak shifts. This same diffraction pattern corrected for  $C(l)$  is shown in figure 9.34. Again there is no strict periodicity in the spectrum which we attribute to experimental error. Employing our criteria for selecting a suitable interval to analyze, we find that the interval  $-0.80 \leq l \leq 0.20$  gives figures of merit to be  $Q_s(1) = 0.002$ ,  $\gamma = -0.498$  and  $\alpha = 0.932$ . Using this interval, we find the correlation functions and the first few of them are shown in table 9.17. Since  $Q_c(3) \approx 0.23 \neq 0.5$ , and  $Q_r(3) \approx 0.14 \neq 0.5$ , we expect that the original 2H structure has largely been eliminated. A plot of  $Q_s(n)$  vs.  $n$  for this data is given in figure 9.35. We find a correlation length for this crystal to be  $\lambda_q = 4.4 \pm 0.7$ .

Examining the  $r = 3$  machine reconstruction results for this process in table 9.18 we see that the antiferromagnetic arcs (0101 and 1010) have a relatively small combined weight of only about 4%. In fact, the probability weight for the 0101 arc is zero. So our original suspicion that the 2H structure has largely been eliminated proves correct. In its place we see large ferromagnetic arcs of nearly equal weight occupying a total of 54% of the weight on the graph. So we agree with Sebastian and Krishna that this is a disordered, twinned crystal. In addition to the 0101 arc, we also find the 1001 and 0010 arcs missing. Looking at the  $r = 3$  graph in figure 9.36, this implies that the twinning fault mechanism is important, as Sebastian and Krishna found, but also the remnant of the

Table 9.19: Computational results for the  $r = 0, 1, 2$  and 3 approximations to experimental data SK135.

System	Language Type	$r$	$h_\mu$	$\mathbf{G}$	$C_\mu$	$\mathbf{E}$	$\mathbf{T}$
$r = 0$ Approximation	SFT	0	0.998	0.002	0.000	0.000	0.000
$r = 1$ Approximation	SFT	1	0.775	0.225	0.999	0.224	0.224
$r = 2$ Approximation	SFT	2	0.727	0.273	1.775	0.320	0.367
$r = 3$ Approximation	SFT	3	0.590	0.410	2.478	0.705	1.112
Fault Approximation	SFT	1	0.529	0.471	1.000	0.471	0.471

1010 arc has some role. Instead of faulting ...1111|0000... where the vertical line indicates the fault plane, we see that the path ...1111|01000... has nearly twice as much probability weight associated with it. In the lower portion of the graph, we see that twinned faulting is largely responsible for the (0)\* fcc cycle converting to the (1)\* fcc cycle and we also observe that double deformation faulting is important. Perhaps it is interesting to mention that, while a modular layer of ZnS has spin inversion symmetry [71] and thus the one-dimensional Hamiltonian describing the energetics of the stacking is also spin invariant, in general these graphs are not spin inverse invariant. That is, the probability of seeing a word and its spin inverse is not the same. By spin inverse, of course we mean just flipping all the spins in a word, *i.e.* 1101  $\rightarrow$  0010. There is of course no reason why we should expect spin inversion; after all, then one could never have a crystal of purely one fcc structure or the other. We note that the fault picture always assumes spin inversion symmetry. Sebastian and Krishna (1994) attribute the faulting to the mechanism of twinned faulting and assign a probability of 12% for this to occur.

Examining the  $Q_s(n)$  derived from experiment with those found from the  $r = 3$  machine in figure 9.37, we find reasonable agreement up to  $n \approx 15$ , and slight deviations thereafter. Looking at the  $Q_s(n)$  found from the faulting picture and comparing it with experimentally derived ones, figure 9.38, we find that the fault picture reproduces the general form of the plot, but overestimates the magnitude of the oscillations. We can further examine the diffraction patterns. In figure 9.39, the diffraction pattern found from the  $r = 3$  approximation is compared with experiment. We see a reasonable fit, except perhaps at a shoulder in the experimental spectrum at  $l = -0.6$  and the small rise at  $l = -0.16$ . We can speculate that there is some minor competition between the 3C and 6H cycles that is not being well modeled here. Comparison of the fault derived diffraction pattern with that from experiment, figure 9.40, reveals good agreement with the peak at  $l = -\frac{1}{3}$  and fair agreement with the one at  $l = -\frac{2}{3}$ . However, the diffuse scattering between peaks is not at all well represented. Additionally, the small rise in diffuse scattering at  $l = \pm\frac{1}{6}$  is likewise absent in the fault model diffraction pattern.

The computational results for the various machine approximations to the process generating the diffraction pattern SK135 are shown in table 9.19. While the fault approximation has a similar but slightly lower rate of entropy production per layer as compared to the  $r = 3$  approximation, the other measures of computation are uniformly lower. This seems to indicate that the fault approximation is missing some important computational aspects of the stacking.

## 9.7 Machine Reconstruction from Experimental Diffraction Pattern SK137

The last experimental spectrum we analyze is shown in figure 9.41. The intensity versus  $l$  in reciprocal space for the 10. $l$  row of an as grown crystal is shown. This same diffraction pattern corrected for  $\mathcal{C}(l)$  is shown in figure 9.42. We see that the spectrum again is not strictly periodic in  $l$ , so we need to

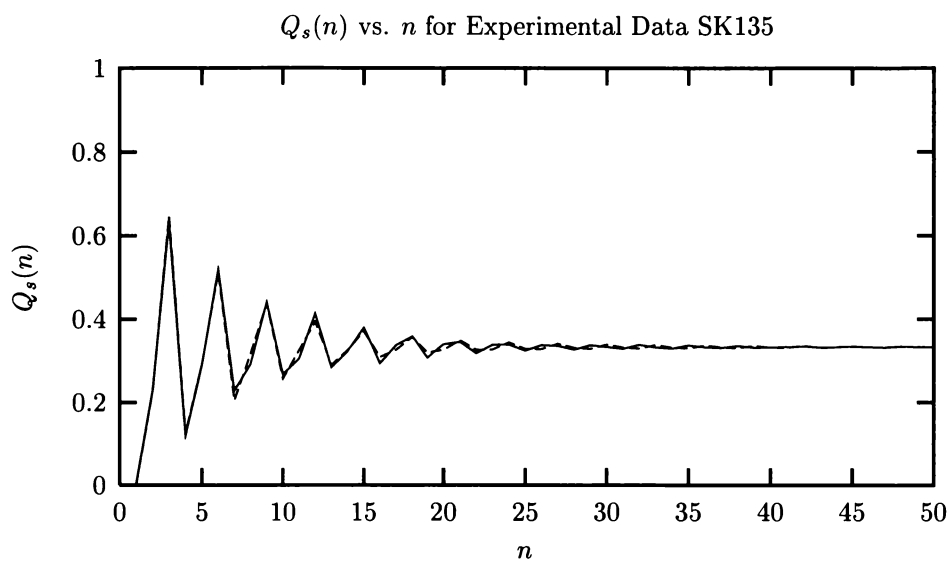


Figure 9.37: The  $Q_s(n)$  vs.  $n$  for Experimental Data SK135 (solid line) and the  $r = 3$  approximation (dashed line).

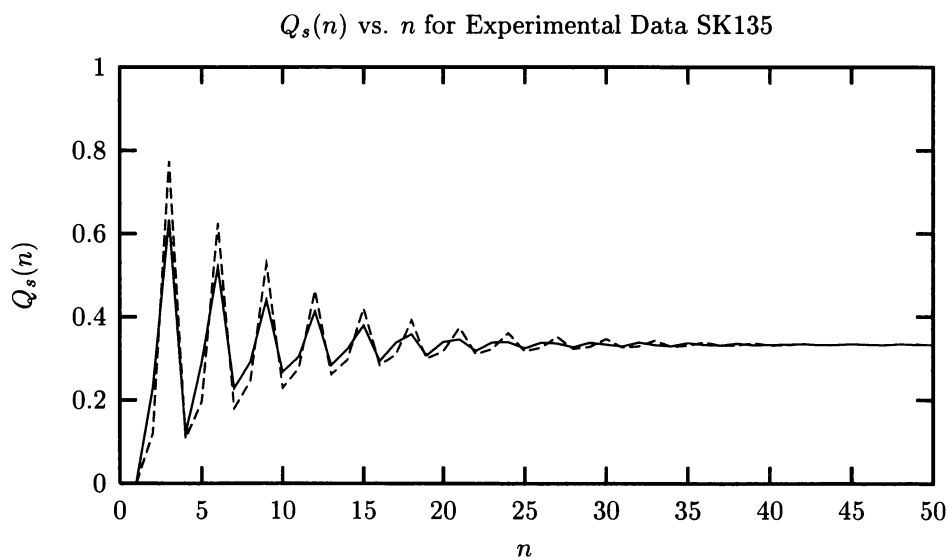


Figure 9.38: The  $Q_s(n)$  vs.  $n$  for Experimental Data SK135 (solid line) and the fault approximation (dashed line).



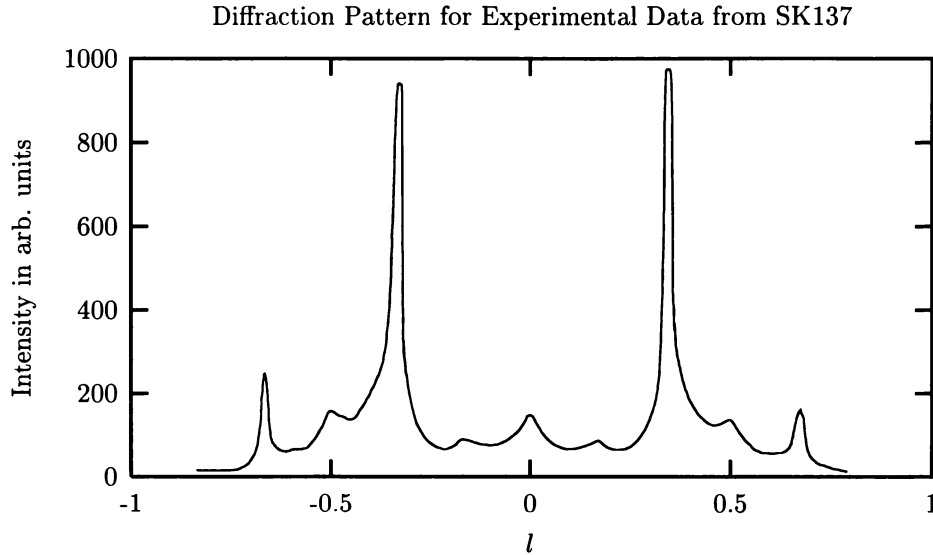


Figure 9.41: Uncorrected diffraction pattern for Experimental Data SK137.

Table 9.20: The first few  $Q_s$  for experimental data SK137.

$n$	$Q_c(n)$	$Q_r(n)$	$Q_s(n)$
1	0.484	0.506	0.010
2	0.338	0.324	0.339
3	0.230	0.245	0.525
4	0.386	0.386	0.228
5	0.377	0.367	0.255
6	0.231	0.246	0.523

choose an interval which gives the best figures of merit. We find this interval to be  $-0.80 \leq l \leq 0.20$ , and the figures of merit turn out to be  $Q_s(1) = 0.010$ ,  $\gamma = -0.485$  and  $\alpha = 0.982$ . The first few correlation functions are tabulated in table 9.20 and a graph of  $Q_s(n)$  vs.  $n$  is shown in table 9.43. We find a correlation length of  $\lambda_q = 12 \pm 3$ .

Machine reconstruction results can be found in table 9.21 and the the  $r = 3$  machine is displayed in figure 9.44. Sebastian and Krishna [60] report this as a disordered 3C crystal, which we can confirm, as the ferromagnetic arcs consume about 44% of the weight in the graph. The faulting mechanism is not so clear. There is only one forbidden word, 0011, and the remaining words, save 1101, all appear at about the 3% to 9% level. We will not attempt a fault analysis since it certainly is not unique and most faulting mechanisms seem to play at least some role. Sebastian and Krishna [60] report that this crystal is well described by a random distribution of twin faults with a 6.8% of occurrence.

A comparison of the  $Q_s(n)$  derived from experiment and the  $Q_s(n)$  obtained from the  $r = 3$  reconstructed machine is shown in figure 9.45. There is reasonable agreement until about  $n \approx 10$ , after which the reconstructed results fall off to the asymptotic value too soon. For  $n$  between 10 and 40, the  $Q_s(n)$  from the  $r = 3$  approximate machine only weakly echo the oscillations in  $Q_s(n)$  from experiment. This provides a hint that there is significant computation missing at  $r = 3$ . The fault

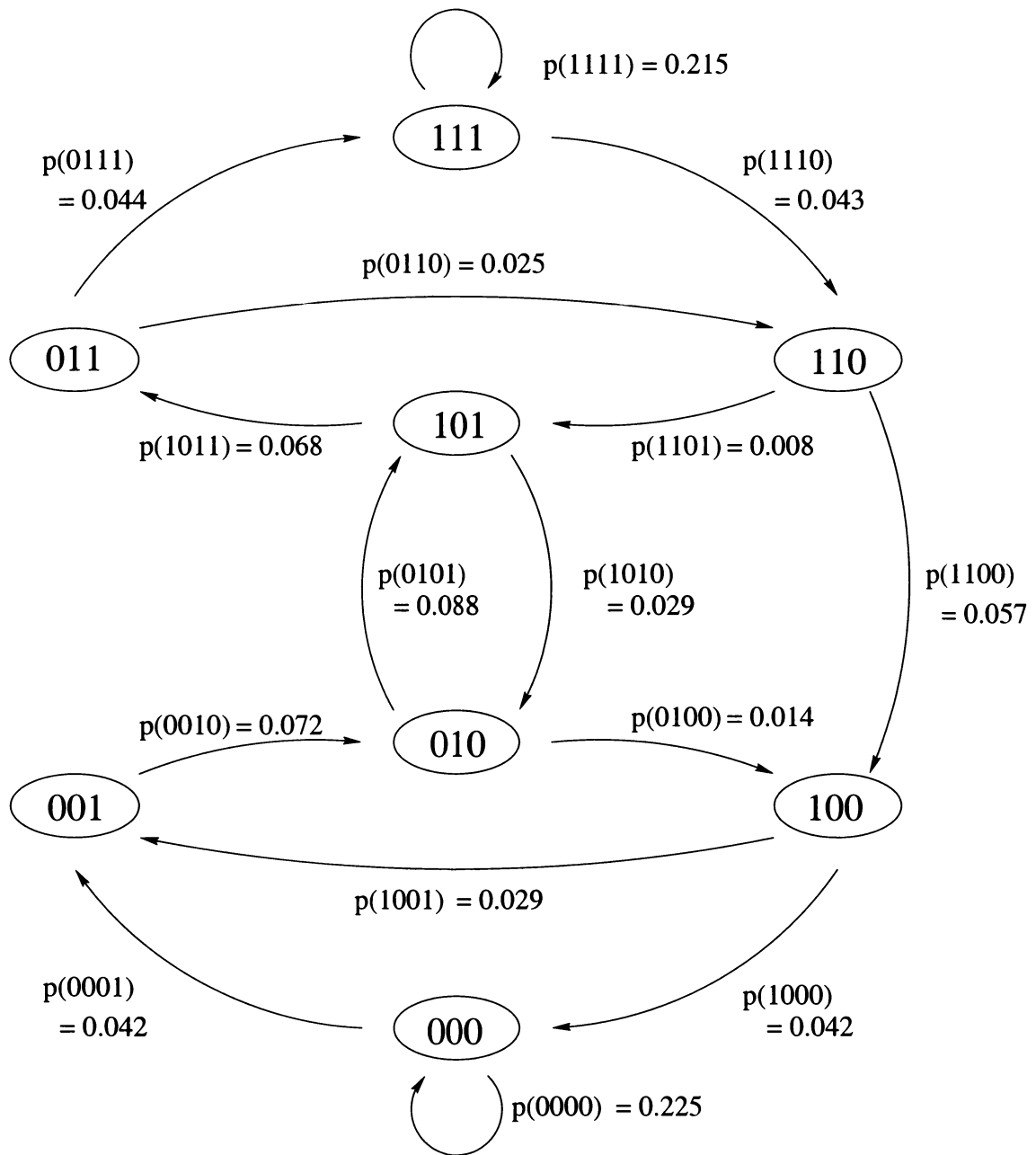


Figure 9.44: The  $r = 3$  machine for diffraction pattern SK137.

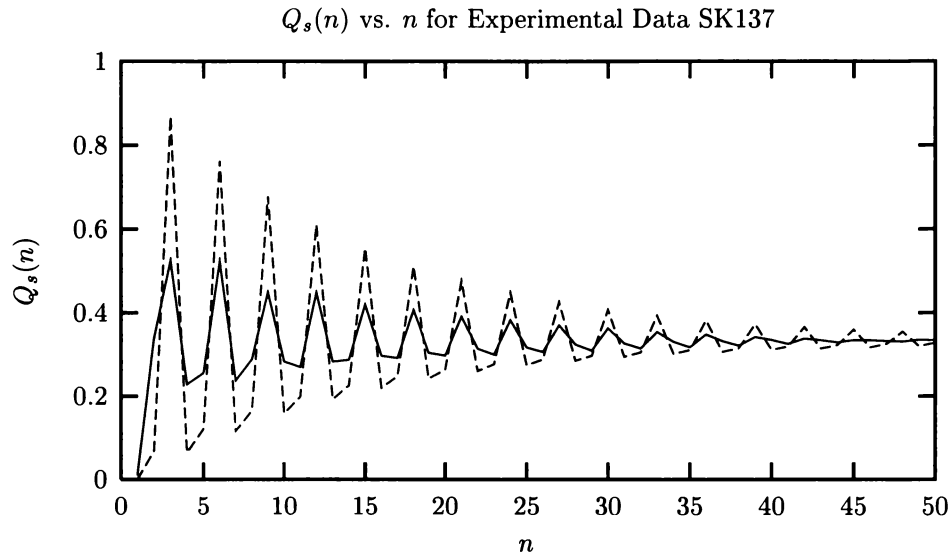


Figure 9.46: The  $Q_s(n)$  vs.  $n$  for Experimental Data SK137 (solid line) and the fault approximation (dashed line).

Table 9.22: Computational results for the  $r = 0, 1, 2$  and 3 approximations to experimental data SK137.

System	Language Type	$r$	$h_\mu$	$\mathbf{G}$	$C_\mu$	$\mathbf{E}$	$\mathbf{T}$
$r = 0$ Approximation	SFT	0	1.000	0.000	0.000	0.000	0.000
$r = 1$ Approximation	SFT	1	0.924	0.076	1.000	0.076	0.076
$r = 2$ Approximation	SFT	2	0.816	0.184	1.922	0.291	0.398
$r = 3$ Approximation	SFT	3	0.651	0.349	2.744	0.792	1.408
Fault Approximation	SFT	1	0.359	0.641	1.000	0.641	0.641

model with a 6.8% twinned fault probability fares worse. It over estimates the magnitude of the oscillations in the  $Q_s(n)$  significantly for  $n \leq 50$ . This simple model (it can after all be expressed as an  $r = 1$  machine), seems to insert too much correlation into the  $Q_s(n)$ . A comparison of diffraction patterns is also possible. For the  $r = 3$  reconstruction, figure 9.47 compares the diffraction pattern with the experimental one. The diffuse scattering is well represented, but the two peaks at  $l = -\frac{2}{3}$  and  $-\frac{1}{3}$  lack sharpness. Comparing the diffraction pattern generated from the fault model with experiment as shown in figure 9.48, we see that the peaks are reasonably well reproduced but the diffuse scattering is almost completely absent in the fault model. The rise in intensity at  $l = -\frac{1}{2}$  is also missing in the fault model.

Table 9.22 shows the computational quantities for the various  $r$  approximations and the fault model. The fault model misses much of the complexity present, as measured by  $C_\mu$  and  $\mathbf{E}$ . It is likely that neither the  $r = 3$  approximation or the fault approximation is modeling the computation present very well, but clearly the fault picture falls far shorter of representing reality here.

## 9.8 Configurational Energies of Polytypes

Now that we have a statistical model for the stacking of the layers, we can calculate physical parameters that depend on this stacking. One such quantity amenable to calculation is the difference in the configurational energies of the particular polytypes. Numerical calculations find that the configurational energy depend only the nearest and the next nearest neighbors in the stacking arrangement. Engel and Needs [21] have done a first-principles pseudopotential calculation of the total energy of five ZnS polytypes, from which they can determine the strength of the interactions up to the third nearest layer. The most general expression possible for inter-layer interactions up to the third nearest neighbors is given by [64]

$$E = E_0 - J_1 \sum_i s_i s_{i+1} - J_2 \sum_i s_i s_{i+2} - J_3 \sum_i s_i s_{i+3} - K \sum_i s_i s_{i+1} s_{i+2} s_{i+3}. \quad (9.1)$$

Terms with an odd number of spins do not appear because of symmetry considerations. We take the  $s_i = \pm 1$  here. Engel and Needs [21] have found that

$$\begin{array}{ll} J_1 = 0.00187 \text{ eV per ZnS pair} & J_3 \text{ negligible} \\ J_2 = -0.00008 \text{ eV per ZnS pair} & K \text{ negligible} \end{array}$$

Let us rewrite equation 9.1 in terms of the energy per ZnS pair, and take the zero of the energy such that  $E_0 = 0$ . We have then,

$$\tilde{E} = -J_1 \langle s_i s_{i+1} \rangle - J_2 \langle s_i s_{i+2} \rangle \quad (9.2)$$

where  $\langle \dots \rangle$  means the expectation value of ‘...’. We can find the expectation values directly from word probabilities,

$$\langle s_i s_{i+1} \rangle = p(11) + p(00) - 2p(01) \quad (9.3)$$

$$\langle s_i s_{i+2} \rangle = p(111) + p(101) + p(000) + p(010) - 2p(110) - 2p(100). \quad (9.4)$$

We show the configurational energy in terms of meV per ZnS pair in table 9.23 for both the crystalline structures considered by Engels and Needs as well as the seven disordered polytype structures on which we have performed machine reconstruction. We see that the two of the disordered samples, SK232 and SK135, have energies not too different from the lowest energy crystal, 3C. As we might expect from the relative magnitudes of  $J_1$  and  $J_2$ , the contribution from the  $J_1$  term completely dominates the energy.

This is one example of a quantity that can be calculated once the statistical nature of the stacking is known. It is obviously desirable to calculate other measurable, physical parameters from the reconstructed  $\epsilon$ -machine. Examples of such quantities would be the specific heat and the transmission of electrons through a disordered, layered sample.

# Chapter 10

## Finale

Understanding disorder in three dimensions is a difficult task [76]. Physicists often resort to descriptions that use correlation information, but this usually does not provide insight into the underlying mechanisms of the disorder or indeed, even provide a detailed picture of the disorder. Recent progress has been made, however, in the analysis and description of disorder in one-dimensional systems. These new methods provide a detailed way to describe, classify and quantify systems both simple, such as periodic structure, and complex. We can meaningfully discuss the entropy density in the one-dimensional system, as well as specify quantities that describe computation and memory. There have, however, been relatively few applications of these methods to physical systems. We bridge this gap for the case of polytypism. Since polytypes can be treated as a one-dimensional spin system and also have interesting physical properties not yet explained, they provide an ideal system to explore the usefulness of these theories. From the physical side, the phenomenon of polytypism has been known for nearly ninety years, but continues to defy theoretical explanation or, even in the case of disordered sequences, adequate description - until now. We have demonstrated that one can describe both the ordered and disordered stacking of polytypes using  $\epsilon$ -machines. Further, we have introduced a new technique that takes as input spectral information instead of the specific sequences for machine reconstruction. Since data from physical systems often comes in this form, we are hopeful that this will prove useful for pattern discovery elsewhere. In the process of applying these new ideas to the old mystery of polytypism, we have given a critical examination of the concept of faulting in polytypes. We have shown that, in the case where the faulting picture is meaningful,  $\epsilon$ -machines provide an equivalent description, as in SK229. But we have discredited the general fault picture as inadequate to describe the disorder in polytypes, and have shown that  $\epsilon$ -machines provide a unique description, at each range of memory  $r$ , of the underlying architecture. Additionally, we demonstrate the usefulness of the  $\epsilon$ -machine description, in that we can calculate physical quantities from it, such as the range of interaction between modular layers and the stacking energy. We could also consider other quantities, such as the local density of phonon states, the specific heat and the transmission coefficient through layered structures [34]. Indeed, any quantity that can be expressed in statistical terms is amenable to calculation once a model of the underlying process (the  $\epsilon$ -machine) is known. For those that are not, a specific sample of the language can be used to numerically calculate quantities. This has important applications in terms of, say, heterostructures. There one is interested in how physical quantities change as the stacking rules are changed. Indeed, such heterostructures can be made artificially in the laboratory and understanding how the band gap changes with the stacking rules has important consequences for semiconductor technology. It is known that the band gap in ZnS is sensitive to the period of the polytype [60] and that there is an anomalous photovoltaic effect of up to several hundred volts/cm found only in disordered ZnS [20] [56] [65].

Let us now discuss our machine reconstruction results for experimental data. From the  $\epsilon$ -machine reconstruction procedure, we see that  $r = 3$  captures most of the structure present. We can not rule

# Bibliography

# Bibliography

- [1] Ash, Robert B. (1990). *Information Theory*. New York: Dover Publications, Incorporated.
- [2] Attard, Phil, Owen G. Jepps and Stjepan Marčelja (1997). "Information Content of Signals using Correlation Function expansions of the Entropy." *Physical Review E*, **56**: 4052-4067.
- [3] Axel, Françoise and Hikaru Terauchi (1991). "High-Resolution X-Ray-Diffraction Spectra of Thue-Morse GaAs-AlAs Heterostructures: Towards a Novel Description of Disorder." *Physical Review Letters*, **66**: 2223-2226.
- [4] Azároff, Leonid V. (1968). *Elements of X-Ray Crystallography*. New York: McGraw-Hill Book Company.
- [5] Azároff, Leonid V., Roy Kaplow, N. Kato, Richard J. Weiss, A.J.C. Wilson and R.A. Young (1974). *X-Ray Diffraction*. New York: McGraw-Hill, Incorporated.
- [6] Badii, Remo and Antonio Politi (1997). *Complexity: Hierarchical Structures and Scaling and Physics*, volume 6 of *Cambridge Nonlinear Science Series*. Cambridge: Cambridge University Press.
- [7] Baumhauer, H. (1912). "Über die Kristalle des Carborundums." *Zeitschrift für Kristallographie*, **50**: 33-39.
- [8] Beck, Christian and Friedrich Schlögl (1993). *Thermodynamics of Chaotic Systems*, volume 4 of *Cambridge Nonlinear Science Series*. Cambridge: Cambridge University Press.
- [9] Buerger, M.J. (1942). *X-Ray Crystallography: An Introduction to the Investigation of Crystals by Their Diffraction of Monochromatic X-Radiation* New York: John Wiley & Sons, Incorporated.
- [10] Burden, Richard L. and J. Douglas Faires (1989). *Numerical Analysis, Fourth Edition* Boston: PWS-Kent Publishing.
- [11] Canright, Geoffrey S. and Greg Watson (1996). "Disordered Ground States for Classical Discrete-State Problems in One Dimension." *Journal of Statistical Physics*, **84**: 1095-1131.
- [12] Clarke, Richard W., Mervyn P. Freeman and Nicholas W. Watkins (2001). "The Application of Computational Mechanics to the Analysis of Geomagnetic Data." *Unpublished*.
- [13] Cover, Thomas M. and Joy A. Thomas (1991). *Elements of Information Theory*. New York: John Wiley & Sons, Incorporated.
- [14] Crutchfield, James P. (1994). "The Calculi of Emergence: Computation, Dynamics, and Induction." *Physica D*, **75**: 11-54.

- [32] Grassberger, Peter (1986). "Toward a Quantitative Theory of Self-Generated Complexity." *International Journal of Theoretical Physics*. **25**: 907-938.
- [33] Guinier, André (1963). *X-Ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies*. San Francisco: W.H. Freeman and Company.
- [34] Gumps, Godfrey, Girija Dubey, A. Salman, B.S. Mahmoud and Danhong Huang (1995). "Statistical and Transport Properties of Quasiperiodic Layered Structures: Thue-Morse and Fibonacci." *Physical Review B*, **52**: 210-219.
- [35] Hahn, Theo, A.J.C. Wilson and U. Shmueli (1992). *International Tables for Crystallography*, 3<sup>rd</sup>, revised edition. Boston: Kluwer Academic publishers.
- [36] Hansen, James E. (1993). *Computational Mechanics of Cellular Automata*. Ph.D. thesis, University of California, Berkeley.
- [37] Hao, Bai-Lin and Wei-Mou Zheng (1998). *Applied Symbolics Dynamics and Chaos*. Singapore: World Scientific Publishing Company.
- [38] Hendricks, Sterling, and Edward Teller (1942). "X-Ray Interference in Partially Ordered Layer Lattices." *Journal Of Chemical Physics*, **10**: 147-167.
- [39] Hopcroft, John E. and Jeffrey D. Ullman (1979). *Introduction to Automata Theory, Languages, and Computation*. Reading, Massachusetts: Addison-Wesley Publishing Company.
- [40] Jagodzinski, Heinz (1949). "Eindimensionale Fehlordnung in Kristallen und ihr Einfluss auf die Röntgeninterferenzen. I. Berechnung des Fehlordnungsgrades aus den Röntgenintensitäten." *Acta Crystallographica*, **2**: 201-207.
- [41] Jagodzinski, Heinz (1949). "Eindimensionale Fehlordnung in Kristallen und ihr Einfluss auf die Röntgeninterferenzen. II. Berechnung der fehlgeordneten dichtesten Kugelpackungen mit Wechselwirkungen der Reichweite 3." *Acta Crystallographica*, **2**: 208-214.
- [42] Jagodzinski, Heinz (1987). "Diffuse X-ray Scattering from Crystals." *Progress in Crystal Growth and Characterization of Materials*. **14**: 47-102.
- [43] Kabra, V.K. and Dhananjai Pandey (1988). "Long-Range Ordered Phases without Short-Range Correlations." *Physical Review Letters*, **61**: 1493-1496.
- [44] Kitchens, Bruce P. (1998). *Symbolic Dynamics One-sided, Two-sided and Countable State Markov Shifts*. Berlin: Springer-Verlag.
- [45] Kittel, Charles (1996). *Introduction to Solid State Physics*, 7<sup>th</sup> Edition. New York: John Wiley & Sons, Incorporated.
- [46] Klinkner, Kristina Lisa, Cosma Rohilla Shalizi and James P. Crutchfield (2001). "An Algorithm for Pattern Discovery in Time Series." *Journal of Machine Learning Research*, submitted.
- [47] McCauley, Joseph L. (1993). *Chaos, Dynamics and Fractals: an Algorithmic approach to Deterministic Chaos* volume 2 of *Cambridge Nonlinear Science Series*. Cambridge: Cambridge University Press.
- [48] Milburn, G.H.W. (1973). *X-ray Crystallography: An Introduction to the Theory and Practice of Single-crystal Structure Analysis*. London: Butterworth & Company Limited.
- [49] Packard, N.H., J.P. Crutchfield, J.D. Farmer and R.S. Shaw (1980). "Geometry from a Time Series." *Physical Review Letters*, **45**: 712-716.



- [69] Upper, Daniel Ray (1997). *Theory and Algorithms for Hidden Markov Models and Generalized Hidden Markov Models*. Ph.D. thesis, University of California, Berkeley.
- [70] Vainshtein B.K. (1981). *Modern Crystallography I: Symmetry of Crystals Methods of Structural Crystallography*. Berlin: Springer-Verlag.
- [71] Varn, Dowman P. and Geoffrey S. Canright (2001). "The Crystal Problem for Polytypes." *Acta Crystallographica A*. **57**: 4-19.
- [72] Verma, Ajit Ram and P. Krishna (1966). *Polymorphism and Polytypism in Crystals*. New York: John Wiley & Sons, Incorporated.
- [73] Warren, B.E. (1969). *X-Ray Diffraction*. Reading, Massachusetts: Addison-Wesley Publishing Company.
- [74] Wilson, A.J.C. (1942). "Imperfection in the Structure of Cobalt II. Mathematical Treatment of the Proposed Structure." *Proceedings of the Royal Society Series A: Mathematical and Physical Sciences*, **180**: 277-285.
- [75] Watson, Greg, Geoffrey S. Canright and Frank L. Somer, Jr. (1997). "Reasonable and Robust Hamiltonians violating the Third Law of Thermodynamics." *Physical Review E* **56**: 6459-6465.
- [76] Welberry, T.R. (1985). "Diffuse x-ray Scattering and Models of Disorder." *Reports on Progress in Physics*. **48**: 1543-1593.
- [77] Wiggins, Stephen (1990). *Introduction to Applied nonlinear Dynamical Systems and Chaos*. New York: Springer-Verlag.
- [78] Woolfson, Michael M. (1997). *An Introduction to X-ray Crystallography*. Cambridge: Cambridge University Press.
- [79] Yi, Jaichul, and Geoffrey S. Canright (1996). "Possible disordered ground states for layered solids and their Diffraction Patterns." *Physical Review B*. **53**: 5198-5210.
- [80] Young, Karl (1991). *The Grammar and Statistical Mechanics of Complex Physical Systems*. Ph.D. Thesis, University of California, Santa Cruz.

# Appendix

# Appendix A

## Conjecture

We state and prove a conjecture in this section.

**Conjecture 1** *When a process results in two-layer correlation functions that decay to an asymptotic value, that value is  $1/3$ .*

We begin by making some definitions. Suppose we have a sample consisting of  $N$  layers, where the orientation of the  $i^{\text{th}}$  layer is denoted by  $x_i \in \{A, B, C\}$ . Let us define a triad of quantities  $f_\alpha$  such that

$$f_c(x_i, x_j) = \begin{cases} 1 & \text{if } x_i \text{ is cyclically related to } x_j \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.1})$$

with  $f_r(x_i, x_j)$  and  $f_s(x_i, x_j)$  defined in an obviously similar way. It is clear that the following identity must hold,

$$f_c(x_i, x_j) + f_r(x_i, x_j) + f_s(x_i, x_j) = 1 \quad \forall \quad i, j. \quad (\text{A.2})$$

Using the  $f_s$ , we can define the  $Q_s$ .

$$Q_\alpha(n) = \frac{1}{(N-n)} \sum_{k=1}^{N-n} f_\alpha(x_{k+n}, x_k) \quad (\text{A.3})$$

where  $\alpha \in \{c, r, s\}$ . We now write the identity

$$\frac{1}{(N-n)} \sum_{k=1}^{N-n} \sum_{\alpha} f_\alpha(x_{k+n}, x_k) = 1. \quad (\text{A.4})$$

Recognizing the last term in the inner sum as  $Q_s(n)$ , and inserting an identity, we have

$$Q_s(n) + \frac{1}{(N-n-1)} \sum_{k=1}^{N-n-1} \{f_r(x_{k+n}, x_k)[f_s(x_{k+n+1}, x_k) + f_c(x_{k+n+1}, x_k)] \\ + f_c(x_{k+n}, x_k)[f_s(x_{k+n+1}, x_k) + f_r(x_{k+n+1}, x_k)]\} = 1.$$

### Vita

Dowman Parks Varn was born on 11 March 1964. He was graduated from Clemson University with a Bachelor of Science degree in Physics in 1986. He was graduated from the University of Virginia in 1994 With a Master of Arts in Physics. He is married to Mary Varn, and they are the proud parents of three dogs.