Quantenoptische Zufallsgeneratoren
Methoden und Analysen

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Physik
der Johann Wolfgang Goethe-Universität
Frankfurt am Main

von
Eric Hildebrandt
aus Wiesbaden

Frankfurt 2002
D F 1

vom Fachbereich .....der
Johann Wolfgang Goethe-Unversität als Dissertation angenommen.

Dekan: .....

Gutachter: .....

Datum der Disputation: .....

Für Ingrid, Hertha und Gertrude

Inhaltsverzeichnis

1 Einleitung **9**
1.1 Eigene Zielsetzung 11

2 Stand der Forschung **15**
2.1 Pseudozufallszahlengeneratoren 15
2.2 Physikalische Zufallsgeneratoren 18
2.2.1 Zufallserzeugung mit Hilfe der Dynamik makroskopischer Systeme 19
2.2.2 Mikroskopische Vielteilchensysteme 22
2.2.3 Quantenmechanische Zufallsgeneratoren 25

3 Theoretische Grundlagen **31**
3.1 Quantenoptische Zufallsgeneratoren 31
3.1.1 Der Zufallsprozeß beim quantenoptischen Zufallsgenerator 31
3.1.2 Erzeugung einzelner Photonen 34
3.1.3 Von-Neumann-Regularisierung 40
3.2 Statistische Tests der Zufallszahlen 41
3.2.1 Methodisches Vorgehen bei statistischen Tests 42
3.2.2 Tests nach FIPS 140-1 46
3.2.3 Test auf Gleichverteilung der Nullen und Einsen 48
3.2.4 χ^2-Test 49
3.2.5 Kontingenz-Test 50
3.2.6 Autokorrelationstest 51
3.2.7 Universelle Tests nach Maurer und Coron 53
3.2.8 Tests aus der nichtlinearen Zeitreihen-Analyse 55

4 Aufbau quantenoptischer Zufallsgeneratoren **59**
4.1 Prinzipieller Aufbau quantenoptischer Zufallsgeneratoren 59
4.2 Die Einphotonenquelle auf Basis der parametrischen Fluoreszenz 60
4.2.1 Der Pump-Laser 60
4.2.2 Der optisch nichtlineare Kristall 61
4.3 Die Photonenquelle auf Basis abgeschwächter Lichtpulse 63
4.4 Das Zufall generierende Element 65
4.4.1 Aufbau in Freistrahloptik 65
4.4.2 Aufbau in Faseroptik 71
4.5 Die Signalverarbeitung und Datenaufnahme 73
4.5.1 Auf NIM-Einschüben basierende Signalverarbeitungselektronik . . 73

4.5.2 Kompakte Signalverarbeitungs- und Datenaufnahme-Elektronik . . 75
4.5.3 Hybride Signalverarbeitungs- und Datenaufnahme-Elektronik . . . 80

5 Durchführung und Ergebnisse der Experimente 83
5.1 Experimente mit der Einphotonenquelle auf Basis der parametrischen Fluoreszenz . . . 83
5.1.1 Justage der Einphotonenquelle . . . 84
5.1.2 Freistrahloptik . . . 85
5.1.3 Faseroptik . . . 96
5.1.4 Abschätzung der Photonenpaarrate . . . 97
5.2 Experimente mit der Photonenquelle auf Basis abgeschwächter Pulse . . . 97
5.2.1 Justage der Photonenquelle . . . 98
5.2.2 Freistrahloptik . . . 100
5.2.3 Faseraufbau . . . 101

6 Diskussion 103
6.1 Analyse der Zufallssequenzen . . . 103
6.1.1 Gründe für das Auftreten der Antikorrelationen . . . 105
6.1.2 Zusammenfassende Bewertung der Analysen . . . 108
6.2 Vergleich der verschiedenen Aufbauten . . . 108
6.2.1 Vor- und Nachteile der verwendeten Einphotonen-Quellen . . . 108
6.2.2 Faseroptischer versus freistrahloptischer Aufbau . . . 114
6.2.3 Mittelwertschwankungen im Teilungsverhältnis . . . 115
6.2.4 Zusammenfassende Bewertung der Aufbauten . . . 122
6.3 Vorschläge weiterer Varianten quantenoptischer Zufallsgeneratoren . . . 123
6.3.1 Mehrfachgeneratoren . . . 123
6.3.2 Quantenoptische Zufallsgeneratoren mit Hong-Ou-Mandel-Aufbau 125
6.3.3 „Integrierter“ Aufbau quantenoptischer Zufallsgeneratoren . . . 129
6.3.4 Zusammenfassende Bewertung der Aufbauvarianten . . . 131
6.4 Mögliche Optimierungen für den technischen Einsatz . . . 131
6.4.1 Schneller Triggerdetektor . . . 131
6.4.2 Optimierte Filterung von parasitärem Licht . . . 132
6.4.3 Der Pumplaser . . . 133
6.4.4 Kompakte Einphotonenquellen . . . 136
6.4.5 Reduktion der Anzahl der benötigten Detektoren . . . 137
6.4.6 Zusammenfassende Bewertung der Optimierungsvorschläge . . . 138
6.5 Resistenz der Zufallsgeneratoren gegen Beeinflussungsversuche . . . 138
6.5.1 Destruktive Angriffe . . . 139
6.5.2 Nichtdestruktive Angriffe . . . 139
6.5.3 Maßnahmen zur Schadensbegrenzung bei Angriffen . . . 141
6.5.4 Signatur eines Zufallsgenerators . . . 142
6.5.5 Zusammenfassende Bewertung zur Resistenz und zu Signaturen . . 146

7 Zusammenfassung 147

Anhang 149

A Läufe der quantenoptischen Zufallsgeneratoren **149**

B Analyse der Zufallssequenzen **153**
B.1 Vergleichs-Pseudozufallszahlengeneratoren . . . 153
B.2 Häufigkeitsverteilung der Bitwerte . . . 154
B.3 Resultate der Tests nach FIPS 140-1 . . . 158
B.4 χ^2-Tests . . . 160
B.4.1 Ergebnisse der χ^2-Tests . . . 161
B.5 Kontingenztests . . . 163
B.6 Die Mächtigkeit von χ^2- und Kontingenz-Tests . . . 164
B.7 Universelle Tests nach Maurer und Coron . . . 164
B.8 Komplexitätstests . . . 170
B.8.1 Lineare Komplexitätstests . . . 171
B.8.2 Kompressionstests . . . 172
B.9 Autokorrelationstests . . . 173
B.9.1 Einfache Autokorrelationsfunktionstests . . . 173
B.9.2 Einfache Autokorrelationstests zur Analyse von Spitzen im Verlauf der Einswahrscheinlichkeit des Laufes *LH5* . . . 175
B.9.3 Zweistufige Autokorrelationskoeffiziententests . . . 178
B.10 Analyse der teilweise auftretenden Antikorrelationen . . . 181
B.10.1 Läufe mit reduzierter Laserpumpleistung . . . 182
B.10.2 Läufe mit künstlich eingefügter Zeitverzögerung . . . 184
B.10.3 Dreistufige Autokorrelationstests . . . 186
B.10.4 Autokorrelationstests der Läufe mit hybrider Datenaufnahme-Elektronik . . . 186
B.11 Die Mächtigkeit der Autokorrelationstests . . . 192

C Berechnung der Zählraten beim HOM-Zufallsgenerator **193**
C.1 Fall 1: Ununterscheidbare Photonen . . . 194
C.2 Fall 2: Unterscheidbare Photonen . . . 196

D Approximative Verteilungsfunktionen **199**
D.1 Die Güte der Normalapproximation . . . 199
D.2 Die Güte der χ^2 Approximation . . . 200

E Theorie und Details der universellen Tests **203**
E.1 Theorie des universellen Tests . . . 203
E.2 Durchführung des universellen Tests . . . 204
E.3 Verbesserte Variante des universellen Tests von Coron . . . 206

F Darstellung der Datenaufnahme-Parameter **209**

G Mathematische Verfahren zur (Pseudo-)Regularisierung **211**
G.1 Regularisierungsmethoden . . . 211
G.1.1 Von-Neumann-Regularisierung . . . 213
G.1.2 Von-Neumann-Regularisierung antikorrelierter Rohdaten . . . 215
G.1.3 Markovketten . . . 217

G.1.4 Von-Neumann-Regularisierung für statistisch abhängige Bits . . . 218
G.1.5 Regularisierung nach Elias . . . 219
G.1.6 Elias-Regularisierung bei statistisch abhängigen Bits . . . 221
G.1.7 Iterierte Von-Neumann-Regularisierung . . . 222
G.1.8 Iterierte Von-Neumann-Regularisierung bei statistisch abhängigen Bits . . . 224
G.2 Pseudoregularisierungsverfahren . . . 224
G.2.1 Die XOR-Transformation . . . 225
G.2.2 XOR-Verknüpfung des Rohdatenstroms mit dem Ausgabestrom eines Pseudozufallsgenerators . . . 226
G.2.3 Transformation mit kryptographisch starken Blockchiffren . . . 226
G.3 Fazit: Welches Regularisierungsverfahren sollte man verwenden? . . . 228

Literaturverzeichnis **238**

Danksagung **239**

Lebenslauf **241**

Kapitel 1

Einleitung

Im Zentrum der Naturwissenschaften steht die Suche nach Gesetzmäßigkeiten, die eine Strukturierung der Welt, wie wir sie in den mannigfaltigen Phänomenen erkennen, ermöglichen. Daher läßt sich die Tätigkeit des Wissenschaftlers auch als eine Form der Datenkompression [118, 119] verstehen: Die unendliche Menge der Phänomene wird auf eine endliche, überschaubare Anzahl von Naturgesetzen zurückgeführt, die jene erklären. Neben dem reinen Erkenntnisinteresse eröffnet diese Reduktion der Phänomene auf Naturgesetze die praktisch wichtige Möglichkeit, die Zukunft in beschränktem Maße vorherzusagen.
Echter Zufall widersetzt sich aufgrund seiner impliziten Regellosigkeit[1] gerade solch einer Kompression, daher läßt er sich in einfacher Form nur in seiner typischen, kollektiven Form wahrscheinlichkeitstheoretisch fassen; die Beschreibung einer individuellen, zufällig erzeugten Sequenz hingegen kann im Mittel nicht wesentlich kürzer sein als die Nennung der Sequenz [21, 82, 117].
Gerade diese Unmöglichkeit der indivuellen Beschreibung und Vorhersage hat schon früh zum Bau und Einsatz von Zufallsgeneratoren geführt. So datiert die künstliche Erzeugung von Zufall unter Verwendung von Zufallsgeneratoren, wie z.B. Würfeln, Orakelknochen usw., mindestens bis in die Zeiten der frühen Hochkulturen zurück, wo diese ersten Zufallsgeneratoren zu divinatorischen Zwecken[2] und beim Glücksspiel verwendet wurden. Der letztgenannte Anwendungszweck erfreut sich nach wie vor großer Beliebtheit und ist heute sicherlich das *kommerziell* bedeutendste Einsatzgebiet von Zufallsgeneratoren, während der erstgenannte doch etwas an Bedeutung verloren hat und zudem immer noch mit traditionellen, wenig technischen Mitteln (Kaffeesatz, Tarotkarten...) betrieben wird.
Im letzten Jahrhundert sind allerdings eine ganze Reihe von Anwendungen für Zufallszahlen in Wissenschaft und Technik neu hinzugekommen: Beispielhaft seien die Monte-Carlo-Verfahren in der numerischen Mathematik [106], die Realisierung von Sequenzen von Einzelereignissen gemäß einer vorgegebenen Wahrscheinlichkeitsverteilung für Simu-

[1] Wie sehr Zufall hiermit auch der Grunddisposition des menschlichen Verstandes zuwiderläuft, zeigt sich sowohl in der Wahrnehmung, die in zufälligen Mustern immer noch eine Struktur zu erkennen sucht, als auch in dem menschlichem Unvermögen „per Hand“ wirkliche Zufallsmuster zu generieren [7]; unser Gehirn verfügt also durchaus nicht über einen „eingebauten“ Zufallsgenerator.

[2] Aus Sicht der Gläubigen sind die Ergebnisse natürlich nicht zufällig, sondern durch einen verborgenen, göttlichen Determinismus verursacht.

lationen [14] und probabilistische Algorithmen, z.B. zur Primzahlsuche [107], genannt. Weiter werden Zufallszahlen gern in der Informatik bei der Evaluation der Effizienz von Algorithmen [68] verwendet, man denke z. B. an das Laufzeitverhalten von Sortieralgorithmen. In der Kryptographie schließlich verwenden Stromchiffrierer (pseudo-)zufällige Bitströme, um Nachrichten bitweise zu verschlüsseln [86, 91].
Die verschiedenen Methoden, die heute zur Erzeugung von Zufallszahlen verwendet werden, lassen sich im wesentlichen in zwei große Gruppen einordnen:

Algorithmische Methoden: Hier wird aus einer oder mehreren vorhergehenden Zahlen die nächste Zahl *berechnet*, wobei am Anfang ein Startwert (entweder eine einzelne Zahl oder eine Zahlenfolge) stehen muß.
Da in solchen Ansätzen alles streng deterministisch ist, wird bei gleichem Startwert auch immer dieselbe Sequenz durchlaufen. Solche Methoden sind also vom Prinzip her eher *Zufalls-Strecker* [72] (für einen zufällig gewählten Startwert) als Zufallsgeneratoren. Allerdings bedeutet dies nicht notwendigerweise, daß aus Teilen der Sequenz leicht auf den Startwert oder den weiteren Fortgang der Folge geschlossen werden kann.

Physikalische Methoden: Es wird ein physikalischer Prozeß verwendet, der glaubhaft zufällig ist – also weder vorhersagbar noch deterministisch – und eine Entscheidung zwischen mindestens zwei Alternativen erlaubt (*Laplace-Mechanismus*). Die Zahlen, i. a. einzelne Bitwerte, werden dann durch Messungen an diesem physikalischen Prozeß erzeugt. Dieser Weg wird bei den in dieser Arbeit präsentierten quantenoptischen Zufallsgeneratoren beschritten.

Für viele Verwendungszwecke sind Zufallszahlen, die mit Hilfe algorithmischer Methoden erzeugt wurden, völlig ausreichend oder sogar vorteilhaft:
So geht es bei Anwendungen von Zufallszahlen in der Statistik, bei Simulationen oder in der numerischen Mathematik oft lediglich darum, Zahlen parat zu haben, die in einem bestimmten Intervall gleichverteilt sind. Unvorhersagbarkeit wird hierbei nicht benötigt, nur Strukturbildung sollte vermieden werden; ja es kann sogar vorteilhaft sein, Zahlen zu generieren, die gleichmäßiger verteilt sind als wirkliche Zufallszahlen [106].
Ist die Erzeugung der Zahlen deterministisch, so handelt es sich selbstverständlich immer nur um *Pseudozufall*, da eine wesentliche Eigenschaft des „echten“ Zufalls fehlt: seine *prinzipielle* Unvorhersagbarkeit – oder wie es John von Neumann drastisch ausdrückt: *Any one who considers arithmetical methods of producing random digits is, of course, in a state of sin.* [96]. In der vorliegenden Arbeit soll es hauptsächlich um die – hoffentlich – „sündenfreien“ Methoden gehen, mit denen sich Zufallszahlen erzeugen lassen. Dies nicht zuletzt deshalb, weil es durchaus Anwendungen von Zufallszahlen gibt, bei denen es tatsächlich wesentlich auf die Eigenschaft der prinzipiellen Unvorhersagbarkeit ankommt, nämlich in der *Kryptographie* und bei *Glücksspielen*.
Bei kryptographischen Anwendungen wird immer die *praktische* Unvorhersagbarkeit der Zufallszahlen durch Dritte gefordert und für Hochsicherheitsbelange, z.B. bei der Schlüsselpaar-Generierung in Trustcentern, ist sogar *prinzipielle* Unvorhersagbarkeit notwendig. Die Effizienz des Zufallszahlengenerators steht häufig nicht im Vordergrund; es sind Sicherheit, Stabilität und Unbeeinflußbarkeit, auf denen das Hauptaugenmerk liegt.

Praktische Unvorhersagbarkeit läßt sich durchaus auch mit kryptographisch starken, algorithmischen Pseudozufallsgeneratoren, s. Abschnitt 2.1, erreichen, aber auch sie brauchen dennoch einen zufälligen Startwert, den sie natürlich nicht selbst produzieren können – d.h. auch bei ihnen kommt der echte Zufall wieder ins Spiel.
Zufälligkeit und prinzipielle Unvorhersagbarkeit wird zudem für die Erzeugung von Schlüsseln für die Verschlüsselung mit Einmalschlüsseln[3] (*OTP, One Time Pad*) gefordert, da sich nur dann die prinzipielle Unbrechbarkeit dieser Verschlüsselungsmethode mathematisch beweisen läßt.

1.1 Eigene Zielsetzung

Trotz des lebhaften Interesses an der Quanteninformationstheorie [99] und der intensiven, weltweiten Forschungsaktivitäten auf diesem Gebiet ist die praxisrelevante, skalierbare Realisierung von Quantencomputern oder auch nur kaskadierbaren Quantengattern aufgrund der hiermit verbundenen großen, experimentellen Schwierigkeiten immer noch ein weitgehend offenes Problem. Die einzige, hinreichend praxisrelevante Nutzanwendung im Bereich der Quanteninformationstheorie stellt nach wie vor die Quantenkryptographie[4] dar; allerdings muß sie mit anderen kryptographischen Verfahren konkurrieren, die erheblich kostengünstiger und flexibler einsetzbar[5] sind.
Im Zuge der Zusammenarbeit der Quantenoptik-Gruppe des Physikalischen Instituts der Johann-Wolfgang-Goethe-Universität mit dem Technologie-Zentrum der Deutschen Telekom AG (heute: Technologie-Zentrum in der T-Systems) entstand daher die Idee, im Rahmen der Suche nach weiteren praxisrelevanten Nutzanwendungen der Quanteninformationstheorie verschiedene Möglichkeiten zur Realisierung quantenoptischer Zufallszahlengeneratoren experimentell zu untersuchen.
Da physikalische Zufallsgeneratoren im Vergleich zu softwarebasierten Pseudozufallsgeneratoren immer aufwendiger sind, werden sie in der Praxis primär für kryptographische Anwendungen[6] interessant sein. Bei ihnen ist die Nichtvorhersagbarkeit der Zufallssequenzen entscheidend, gewünscht ist eine Art Orakel, das einzelne Bits ausgibt, die mit einer Wahrscheinlichkeit von 1/2 den Wert Eins bzw. Null haben, wobei völlig unvorhersagbar sein sollte, wann welcher Wert ausgegeben wird. Dies läßt sich mit Hilfe eine Reihe von Anforderungen an physikalische Zufallsgeneratoren zusammenfassend formulieren:

[3] Bei dieser Verschlüsselung wird der zu verschlüsselnde Klartext und ein gleich langer Einmalschlüssel mit Hilfe einer XOR-Operation bitweise zum Chiffrentext umgewandelt. Zum Dechiffrieren wird der Chiffrentext einfach nochmals mit demselben Einmalschlüssel XOR-verknüpft. Ist der Einmalschlüssel wirklich zufällig und wird er tatsächlich nur einmal verwendet, so läßt sich das Chiffrat ohne den Einmalschlüssel nicht entschlüsseln, da das zufällige Bitmuster des Schlüssels den Klartext vollständig maskiert.

[4] Im Rahmen des EU-Projektes EQCSPOT, European Quantum Cryptography and Single Photon Optical Technologies wurde versucht die Quantenkryptographie und die bei ihr benötigten Komponenten bis zur industriellen Einsatzreife weiterzuentwickeln. Seit Anfang 2002 bietet das aus dem Projekt hervorgegangene Unternehmen *id Quantique* auch die ersten kommerziellen Quantenkryptographie-Komponenten an.

[5] Es handelt sich meist um Software-Lösungen auf standardisierter Hardware. Kritisch für den praktischen Einsatz der Quantenkryptographie sind ihre begrenzte Reichweite und die Anforderungen, welche sie an den Übertragungsweg (dedizierte Glasfaser mit geringer Streckendämpfung) stellt.

[6] Überdies ist gerade in der „Kryptographengemeinde“ das Interesse an möglichst fundamentalen, physikalischen Zufallsgeneratoren am größten.

- Ein genuin zufälliger Prozeß, der nicht vorhersagbar ist, soll zugrunde gelegt werden. Dieser Prozeß soll weiter möglichst einfach, „intuitiv einsichtig“ und aufgrund wissenschaftlicher Erkenntnisse glaubhaft zufällig sein, um mit einer größtmöglichen Akzeptanz potentieller Anwender rechnen zu können.
- Der Prozeß soll möglichst resistent gegen äußere Beeinflussung sein; dies gilt insbesondere für das Einstrahlen elektromagnetischer Wellen.
- Aus Gründen der Betriebssicherheit, des Arbeits- und Brandschutzes kommt die Verwendung von radioaktiven Quellen nicht in Frage.
- Elektronische Rauschquellen sind nicht akzeptabel, da bei ihnen nur schwer auszuschließen ist, daß sich äußere Einflüsse auf den Rauschprozeß auswirken[7].
- Hinsichtlich der Geschwindigkeit, mit der die Zufallsbits erzeugt werden, sind Bitraten im Bereich von kBit/s meist ausreichend; höhere Datenraten wären bei gleicher Qualität der Zufallzahlen natürlich durchaus wünschenswert.
- Der Generator soll hinreichend wartungsarm, kompakt, leicht zu transportieren und kostengünstig sein.
- Es wäre wünschenswert, daß das Funktionsprinzip des Generators, eventuell mit gewissen Modifikationen, eine Miniaturisierung des Generators nicht grundsätzlich ausschlösse.

In der vorliegenden Arbeit wird untersucht, inwieweit quantenoptische Zufallsgeneratoren diese Bedingungen erfüllen und für einen Einsatz im Rahmen kryptographischer Anwendungen geeignet sind. Hierbei werden im Hinblick auf einen Einsatz in der Praxis sowohl freistrahl- als auch faseroptische Realisierungen aufgebaut und getestet.
Außerdem wird untersucht, welche Vor- und Nachteile Einphotonen-Anzahlzustände als „Betriebsmedium“ des Zufallsgenerators gegenüber in der Amplitude stark abgeschwächten Lichtfeldern[8] mit einer Poissonstatistik der Photonenanzahl haben.
Zur Beurteilung der Güte der Zufallszahlen werden die statistischen Eigenschaften der generierten Zufallsbitströme mit Hilfe statistischer Tests, die der Problemstellung angemessenen sind, analysiert und mit der Güte von Zufallszahlen aus (kryptographisch starken) Pseudozufallszahlengeneratoren verglichen.
Schließlich werden noch Entwürfe für alternative quantenoptische Zufallsgeneratoren und für die physikalische Generierung von „Signaturen“ in den erzeugten Binärsequenzen präsentiert.
Im folgenden zweiten Kapitel der Arbeit wird der Stand der Forschung anhand der verschiedenen Funktionsprinzipien physikalischer Zufallsgeneratoren dargelegt. In Kapitel 3 werden die theoretischen Grundlagen erläutert, auch für die nachher verwendeten, empirischen Tests der Zufallszahlen. Kapitel 4 ist dann der Beschreibung des Aufbaus der verschiedenen Varianten quantenoptischer Zufallsgeneratoren gewidmet. In Kapitel 5 werden die durchgeführten Experimente mit den verschiedenen Aufbauvarianten

[7] Insbesondere besteht bei ihnen die Gefahr, daß mit Hilfe von elektromagnetischen Störfeldern der Rauschprozeß *gezielt* beeinflußt werden könnte.

[8] Experimentell werden gepulste Lichtfelder verwendet, Fragestellungen, die beim Einsatz kontinuierlicher Lichtfelder auftreten, werden aber in der Diskussion ebenfalls berücksichtigt.

beschrieben. Nach einer kurzen Zusammenfassung der wesentlichen Ergebnisse der Analysen der Zufallssequenzen werden im Diskussionskapitel die Vor- und Nachteile der verschiedenen Aufbauvarianten quantenoptischer Zufallsgeneratoren erörtert und weitergehende Ausgestaltungs- und Optimierungsmöglichkeiten dargelegt. Der Haupttext schließt mit der Zusammenfassung und dem Literaturverzeichnis.
Im Anhang finden sich dann noch eine Reihe von Detailinformationen und Ergänzungen:

1. eine Auflistung der durchgeführten Generatorläufe,
2. eine ausführliche Darstellung der Ergebnisse der Analysen der Zufallssequenzen mit Wertung der statistischen Tests auf ihre Eignung zum Test physikalischer Zufallsgeneratoren,
3. einige Detailrechnungen für den vorgeschlagenen HOM-Zufallsgenerator,
4. ein Abschnitt zur Güte der in einigen Tests verwendeten Normalapproximation,
5. Theorie und Durchführungsdetails der universellen Tests,
6. zusätzliche Informationen zur Darstellung der Datenaufnahme-Parameter bei der Datenaufnahme-Software und
7. eine weitgehend in sich abgeschlossenen Darstellung mathematischer Verfahren zur Regularisierung tendenzbehafteter Rohdaten.

Kapitel 2

Stand der Forschung

Aufgrund der großen Bedeutung, die Zufallszahlen in der angewandten Mathematik, Statistik, Nachrichtentechnik und Kryptographie haben, beschäftigt man sich schon seit langem mit der Erzeugung und Prüfung von Zufallszahlen. Hierbei liegt der Schwerpunkt eindeutig auf der Entwicklung und Untersuchung von Pseudozufallszahlengeneratoren, also der algorithmischen und somit[1] deterministischen Erzeugung von Zahlen, deren statistische Eigenschaften genuinen Zufallszahlen möglichst nahe kommen sollen.

2.1 Pseudozufallszahlengeneratoren

Bei der Entwicklung von Pseudozufallsgeneratoren spielt das Einsatzgebiet, für das die Generatoren vorgesehen sind, eine große Rolle, da es die gewünschten Eigenschaften der Generatoren bestimmt:
Pseudozufallsgeneratoren für die Anwendung in der Statistik, numerischen Mathematik, Informatik und bei Simulationen sollten folgende Eigenschaften haben [68, 76]:

- Die erzeugten Zufallzahlen sollten gute statistische Eigenschaften aufweisen, d.h. die Zahlen sollten praktikable[2] Tests auf Abweichung vom ideal zufälligen Verhalten bestehen.

- Der Generator sollte sich auf eine mathematische Theorie stützen können, damit eine theoretische Analyse zumindest prinzipiell möglich ist.

- Der Generator sollte eine lange Periode[3] haben, d.h. bei fortlaufender Erzeugung

[1] Algorithmen müssen nicht deterministisch sein, es gibt auch probabilistische Algorithmen, allerdings setzen diese bereits Zufallszahlen voraus und können folglich nicht zur Erzeugung solcher eingesetzt werden.

[2] Als praktikabel gelten hierbei solche Tests, die ein polynomial von der Stichprobengröße abhängiges Laufzeitverhalten aufweisen, wobei der maximale Exponent natürlich möglichst klein sein sollte. Solche Tests lassen sich auf herkömmlichen Rechnern effizient durchführen.

[3] Gibt der verwendete Algorithmus eine Zahl aus, die bereits einmal generiert wurde, so befindet er sich danach in einem inneren Zustand, der schon einmal auftrat und alle nachfolgenden Zahlen sind dementsprechend ebenfalls bereits aufgetreten. Beim Entwurf eines Pseudozufallsgenerators muß daher dafür Sorge getragen werden, daß dieser Zyklus maximal lang ist und nicht abhängig von den Startwerten kürzere Zyklen, s. a. [40], auftreten können.

von Zufallszahlen sollte es hinreichend lange[4] dauern, bis sich die erzeugten Zahlen wiederholen.

- Die Erzeugung der Zahlen sollte möglichst effizient und schnell sein.
- Der verwendete Algorithmus sollte gut portierbar sein, d.h. in einer verbreiteten Programmiersprache ohne explizite Maschinenabhängigkeit (z.B. von der Registerlänge) programmierbar sein.
- Die erzeugten Zufallszahlen sollten reproduzierbar sein, d.h. bei gegebenem Startwert bzw. gegebenen Startwerten wird immer dieselbe Zufallszahlensequenz erzeugt.
- Gern wird auch die Möglichkeit gewünscht, an eine bestimmte Stelle innerhalb einer Zufallssequenz springen zu können, ohne alle vorhergehenden Zahlen generieren zu müssen.

Kryptographisch starke Pseudozufallsgeneratoren werden für kryptographische Zwecke eingesetzt und müssen daher neben einer möglichst effizienten Zahlengenerierung und guten statistischen Eigenschaften der Zufallszahlen noch weitere Eigenschaften besitzen:

- Der Generierungsalgorithmus verwendet neben den Startwerten noch zusätzliche, flexibel wählbare Parameter, die es erlauben, trotz eines einheitlichen Grundalgorithmus eine Vielzahl *unterschiedlicher* Pseudozufallsgeneratoren gleicher Güte zu erzeugen.
- Es muß prinzipiell oder zumindest praktisch unmöglich[5] sein, aus der Kenntnis eines Teiles der Zufallssequenz auf die ihm vorhergehenden oder nachfolgenden Zahlen schließen zu können.

Für beide Gruppen von Pseudozufallsgeneratoren existieren eine ganze Reihe von Realisierungsvorschlägen und zugehörige Implementationen (s. z. B. [68, 75, 101, 106]), wobei Pseudozufallszahlengeneratoren für numerische Anwendungen den Löwenanteil stellen. Da bei numerischen Anwendungen häufig große Mengen von Zufallszahlen generiert werden müssen, dürfen die i. a. in Software implementierten Generatoren nicht viel Rechenleistung verbrauchen und somit auch nicht zu komplex sein. Einige typische Generatoren, die später auch bei den empirischen Tests zum Vergleich verwendet werden, werden in Abschnitt B.1 näher beschrieben. Hier seien nur summarisch die wichtigsten kryptographisch starken Generatortypen aufgeführt.
Für kryptographische Anwendungen wurden insbesondere folgende zwei Arten von Zufallsgeneratoren vorgeschlagen:

[4] „Hinreichend lang“ bemißt sich hierbei an der Menge der benötigten Zahlen: Als Faustregel gilt, daß die Periode mindestens doppelt so groß sein sollte, wie die Menge der verwendeten Zahlen.

[5] *Prinzipiell unmöglich* bedeutet hierbei, daß sich Abhängigkeiten zwischen den Teilen einer Sequenz nur durch Tests ermitteln ließen, deren Zeit bzw. Resourcenverbrauch so hoch wäre, daß sich prinzipiell kein Rechner konstruieren ließe, auf dem diese Tests erfolgreich durchgeführt werden könnten. *Praktisch unmöglich* schwächt dies dahin gehend ab, daß sich kein solcher Rechner mit vertretbarem Aufwand bauen ließe.

1. *Pseudozufallszahlengeneratoren, deren praktische Unvorhersagbarkeit auf der Schwierigkeit zahlentheoretischer Probleme basiert* [72, 86], erlauben es, ihre Unvorhersagbarkeit mathematisch zu beweisen, wobei allerdings immer angenommen werden muß, daß die ihnen zugrundeliegenden zahlentheoretischen Probleme tatsächlich nicht leichter zu lösen sind, als heute angenommen wird. Es gibt immer wieder neue Vorschläge für solche Zufallsgeneratoren, die bekanntesten Beispiele für diesen Typ von Generator sind:

   - der *RSA-Zufallsgenerator*[6] [3, 6], dessen Sicherheit auf der Schwierigkeit des Faktorisierung großer Zahlen basiert.

     Dieser Generator benötigt eine Zahl $N = p \cdot q$, die das Produkt zwei großer[7], nicht identischer Primzahlen p und q darstellt und eine Zahl d, die teilerfremd zu dem Produkt $(p-1) \cdot (q-1)$ ist, d.h. der größte gemeinsame Teiler dieses Produktes und der Zahl d ist lediglich die Eins. Zufallszahlen werden dann gemäß der Vorschrift:

     $$x_{n+1} = x_n^d \bmod N$$

     erzeugt. Allerdings dürfen für kryptographische Anwendungen lediglich $\log_2(k)$ Bits, für den Zufallsbitstroms verwendet werden [6], wenn k für die Anzahl der Binärstellen von N steht. Eine verbesserte Variante des Generators von Micali und Schnorr [92] ermöglicht allerdings die Nutzung aller Bits für die Ausgabe des Zufallsbitstroms.

   - der *BBS-Generator*[8] [13], der das Quadrieren modulo N verwendet:

     $$x_{n+1} = x_n^2 \bmod N.$$

     Hierbei ist N wieder das Produkt zwei großer, nicht identischer Primzahlen p und q, wobei allerdings diese Primzahlen noch eine Reihe weiterer Bedingungen erfüllen müssen, außerdem kann der Anfangswert x_0 ebenfalls nicht beliebig gewählt werden, da sonst die Gefahr besteht, daß der Generator sehr schnell in einen Zyklus gerät, s. [13]. Als Ausgabe liefert der BBS-Generator in seiner ursprünglich vorgeschlagenen Form[9] pro Generierungsschritt jeweils das niederwertigste Bit der Zahl x_{n+1}.

2. *Stromchiffren*: Hierbei handelt es sich um Chiffrierverfahren, die einen praktisch nicht vorhersagbaren, von einem geheimen Schlüssel deterministisch abhängigen Bitstrom erzeugen. Stromchiffren lassen sich auf verschiedenste Weise realisieren [91, 113], gern wird dabei auf bereits vorhandene Primitive zurückgegriffen, z.B. :

[6] Die Abkürzung RSA steht für die Initialen der Erfinder (innerhalb der freien Wissenschaft) des asymmetrischen Verschlüsselungsverfahrens, das auf der Faktorisierung großer Zahlen basiert: Ron Rivest, Adi Shamir und Ron Adleman.

[7] Die Länge der Primzahlen ist ein Sicherheitsparameter des Generators; eine typische, für die nächsten Jahre noch als sicher angesehene Länge sind 1024 Bits. Bei Fortschritten in der algorithmischen Zahlentheorie bzw. Rechnertechnik verschiebt sich dieser Wert zu größeren Längen hin.

[8] Die Abkürzung BBS steht für die Initialen der Erfinder des Generators: M. Blum, L. Blum und M. Shub.

[9] Die Effizienz des Generators läßt sich steigern indem mehrere niederwertige Bits entnommen werden, die Anzahl hängt von der Größe der Zahl N und einer Konstante ab.

- Mehrere, eventuell asynchron laufende, rückgekoppelte[10] Schieberegister werden miteinander nichtlinear verknüpft [113, 111].
- Es werden Blockchiffren[11] in folgenden Modi zur Zufallsgenerierung verwendet:

  – In einem Rückkopplungsmodus, bei welchem der verschlüsselte Ausgabeblock als Eingabeblock für den nächsten Verschlüsselungsschritt dient[12]; die aneinandergefügten Ausgabeblöcke ergeben den Ausgangsbitstrom der Stromchiffre. Selbstverständlich kann man die Ausgabeblöcke auch einzeln als Zufallszahlen verwenden.
  – Im sogenannten *Zählermodus*, bei dem ein typischerweise 64-Bit langer Zählerwert als zu verschlüsselnder Klartextblock dient und die chiffrierten Ausgangsblöcke wieder die Zufallszahlen bzw. den Zufallsstrom darstellen. Nach jedem Verschlüsselungsschritt wird der Zähler um Eins inkrementiert. Dieser Modus weist den großen Vorteil auf, daß es bei ihm möglich ist, an eine beliebige Stelle des Zufallsstromes zu springen, da außer dem frei wählbaren Zählerwert und dem (konstanten) Schlüssel keine weiteren Werte in die Berechnung eingehen.

In der Praxis werden meist Stromchiffren zur Erzeugung kryptographisch starker Pseudozufallszahlensequenzen verwendet, da sich die oben aufgeführten Vorgehensweisen in Hardware oder in Software[13] hinreichend effizient realisieren lassen. Pseudozufallszahlengeneratoren, die auf zahlentheoretischen Problemen basieren, sind aufgrund der hierfür notwendigen sehr großen ganzen Zahlen (> 512 Bit Darstellungslänge) weniger effizient auf herkömmlichen Prozessoren zu implementieren und daher oft zu langsam; Abhilfe können spezielle Langzahlprozessoren bringen, die allerdings meist nur in dedizierter kryptographischer Hardware, wie z.B. Sicherheits-Chipkarten, zu finden sind.

2.2 Physikalische Zufallsgeneratoren

Physikalische Zufallsgeneratoren lassen sich grob nach ihrem *grundlegenden Funktionsprinzip* in drei Klassen unterteilen:

[10] Hierbei werden nach dem „Schieben“ die Bitwerte *bestimmter* Schieberegister-Zellen über eine adäquate Rückkopplungsfunktion verknüpft und das Ergebnis in die höchste Zelle geschrieben. Als Ausgabewert dient der beim Schieben aus der niedrigsten Zelle geschobene Wert. Im einfachsten Fall handelt es sich um *lineare* Schieberegister, bei denen die Rückkopplungsfunktion einfach eine XOR-Verknüpfung der Bitwerte ist.

[11] Blockchiffren sind schlüsselabhängige, stark nichtlineare Funktionen, die einen n Bit langen Eingabeblock auf einen gleichlangen Ausgabeblock eineindeutig abbilden. Auch für die Zufallsgenerierung ist die Eineindeutigkeit durchaus von Bedeutung, da hierdurch bei Verwendung des OFB-Modus (s.u.) verhindert wird, daß sich Zyklen bilden, welche die maximale Periode reduzieren.

[12] Dies ist der sogenannte OFB-Modus, *Output Feedback Mode*, bei ihm sollten alle (typisch 64) Bits eines Ausgabeblockes rückgekoppelt werden, da eine Rückkopplung von weniger Bits die Wahrscheinlichkeit drastisch erhöht, daß solch ein Generator in einen Zyklus gerät, s. a. [91], Sec. 7.25

[13] Schieberegister lassen sich natürlich effizienter in Hardware als in Software realisieren. Es gibt allerdings auch spezielle für die Implementation in Software optimierte Stromchiffren, wie z.B. RC4 und SEAL.

- Zufallsgeneratoren, welche die *Unvorhersagbarkeit von Ereignissen aufgrund der (komplexen) Dynamik eines makroskopischen Systems* verwenden.
- Zufallsgeneratoren, die eine *fluktuierende makroskopische Meßgröße eines mikroskopischen Vielteilchensystem* zur Zufallserzeugung benutzen.
- Zufallsgeneratoren, welche die *Unbestimmtheit einer Messung mit mehreren möglichen Meßresultaten an einem quantenmechanischen Objekt* nutzen.

Grundsätzlich können natürlich auch mehrere Funktionsprinzipien gleichzeitig bei einem physikalischen Zufallsgenerator zum Tragen kommen, allerdings wird meist eines der Prinzipien überwiegen.

2.2.1 Zufallserzeugung mit Hilfe der Dynamik makroskopischer Systeme

Generatoren dieser Klasse verwenden ein im Grunde genommen deterministisches, makroskopisches System, das allerdings dennoch unter bestimmten Umständen nicht vorhersagbares Verhalten aufweisen kann. Es lassen sich hierbei mehrere Möglichkeiten unterscheiden:

1. Mikroskopisches, nicht kontrollierbares[14] „Rauschen“ wird durch die Systemdynamik, die in diesem Fall durchaus auch linear sein kann, auf eine makroskopische Skala gebracht.
2. Unabhängige Ereignisketten bzw. Systeme mit einer an sich deterministischen Dynamik, aber nicht vollständig bekannten Startbedingungen bzw. mathematischen Modellparametern werden, via Koinzidenz oder Kombination, zu einem (vermeintlich) stochastischen Gesamtprozeß vereinigt[15].
3. Die Systemdynamik ist nichtlinear und zeigt eine sensitive Abhängigkeit von den Startbedingungen gekoppelt mit starken Mischungseigenschaften im Phasenraum. Durch Festlegung der Steuerparameter der Systemdynamik auf bestimmte Werte kann man erreichen, daß Prognosen nur noch bis zu einem Vorhersagehorizont möglich sind. Mißt man nun eine durch die Systemdynamik bestimmte Meßgröße zu Zeitenpunkten, zwischen denen eine Zeitspanne liegt, die größer als dieser Vorhersagehorizont ist, so erhält man unkorrelierte Werte.

Bekannte Beispiele für „klassische“ Zufallsgeneratoren, wie Würfel, Glücksrad, Roulette und ähnliche bei Glücksspielen[16] eingesetzte Zufallsgeneratoren, basieren weitgehend nicht auf einer komplexen[17] Dynamik, sondern bei ihnen rührt die Zufälligkeit des Endresultates daher, daß bei einer entsprechenden Wahl der Parameter sich kleine unkontrollierbare Abweichungen in den Anfangsbedingungen stark auf den (diskreten) Ausgang

[14] Das muß nicht unbedingt heißen, daß es *grundsätzlich* nicht kontrollier- oder unterdrückbar ist, sondern es reicht aus, wenn es beim betrachteten Aufbau vorhanden ist.

[15] Man denke z.B. an ein Roulette mit einem sich drehenden Kessel und einer gegenläufig umlaufenden Kugel.

[16] Die heutigen Geldspielautomaten basieren nicht auf oben genannten Prinzipien, bei ihnen werden meist Pseudozufallszahlengeneratoren verwendet!

[17] Das Auftreffen und Abprallen von Würfen oder Münzen auf einer Oberfläche wird hierbei nicht berücksichtigt.

des Zufallsexperiments auswirken. Ein schönes Beispiel hierfür ist der Wurf einer fairen Münze mit Auffangen [29, 37, 44, 66], dies ist ein Vorgang, der eine recht einfache Dynamik besitzt und daher auch nur unter bestimmten Bedingungen, nämlich möglichst hoher Steighöhe und schneller Drehung der Münze, als hinreichend zufällig angesehen werden kann: Sind diese Bedingungen nicht erfüllt, ist auch keine Zufälligkeit gegeben, so gibt es z.B. Zauberer, die in der Lage sind, gezielt Kopf oder Zahl zu werfen [29].
Ähnlich ist die Lage beim Roulette; auch hier läßt sich das Ergebnis eines Laufes bei entsprechender Kenntnis der Anfangsbedingungen und zusätzlicher Einflußgrößen, insbesondere den Modellparametern, in Grenzen[18] vorhersagen [8, 9]. Solche Systeme sind also eher „Zufallsverstärker“, d.h. sie bringen eine implizit in den Startbedingungen oder Modellparametern vorhandene (mikroskopische) Ungenauigkeit, im Extremfall sogar quantenmechanisch bedingte Unbestimmtheit, auf eine makroskopische Skala. Dementsprechend nimmt ihre Eignung als Zufallsgenerator auch in dem Maße ab, in dem die Anfangsbedingungen bekannt sind bzw. konstant gehalten werden.
Entschärfen lassen sich solche Probleme, indem man Systeme mit einer komplizierteren Dynamik verwendet; ein bekanntes Beispiel für solch eine Vorgehensweise ist das Ziehungsgerät der Lottozahlen, bei dem ebenfalls ein makroskopisches Vielteilchensystem verwendet wird.
Beim Entwurf von Zufallsgeneratoren, die mit makroskopischen Systemen arbeiten, kann man sich auch die Erkenntnisse der nichtlinearen Dynamik zu Nutze machen und eine Systemdynamik auswählen, die bei geeigneter Parameterwahl chaotisches Verhalten zeigt und sich bei hinreichend großen Intervallen zwischen der Zufallsdatengenerierung nicht vorhersagen läßt. T. Kuusela hat genau dies getan [71], indem er einen Zufallsgenerator auf Basis eines chaotischen Taktgenerators konstruierte. Für diesen chaotischen Taktgenerator werden zwei identische LC-Oszillatoren, deren nichtlineares Verhalten von der nichtlinearen Kennlinie der Kapazitätsdioden herrührt, mit Hilfe eines periodischen unipolaren Rechteck-Taktsignals (ca. 500 kHz) zu erzwungenen Schwingungen angeregt, die aufgrund einer geeigneten Wahl der Parameter nicht periodisch, sondern chaotisch sind. Die an den beiden Kapazitätsdioden anliegenden chaotischen Spannungssignale werden von zwei schnellen Komparatoren in Rechtecksignale mit Logikpegeln umgewandelt. Wobei eines der beiden Signale das Übernahmesignal für ein D-Flipflop generiert und das andere als Eingangssignal für dieses Flipflop dient[19]. Der Q-Ausgang des D-Flipflop dient dann als Ausgang des chaotischen Taktgenerators. Dieses Vorgehensweise muß verwendet werden, da ein einzelnes chaotisches Logik-Signal nicht zufällig genug ist, d.h. es ist nicht in der Lage, alle möglichen Bitsequenzen zu erzeugen. So aber gelingt es, Zufallszahlen mit der recht hohen Rate von 60.000 Zahlen à 8 Bit Länge (entsprechend einer Bitrate von 480.000 Bits/s) zu erzeugen. Die so erzeugten Rohdaten zeigen zwar keine offensichtlichen Korrelationen auf, allerdings lassen sich in der Wiederholungsabbildung[20] noch Strukturen erkennen, die von einer Ungleichverteilung der generierten

[18] Wenn die Anfangsbedingungen hinreichend genau bekannt sind, läßt sich z.B. die Wahrscheinlichkeit berechnen, mit der die Kugel in eines der Viertel des Kessels fällt; diese Wahrscheinlichkeit ist nicht für alle Viertel gleich. Wählt man aber eine binäre Zufallsvariable wie Gerade-Ungerade oder Rot-Schwarz, lassen sich keine Vorhersagen machen, die bessere Ergebnisse liefern als bloßes Raten.

[19] Eine ähnliche Vorgehensweise wird gerne verwendet, um die statistischen Eigenschaften von Pseudozufallszahlengeneratoren zu verbessern, allerdings im Rahmen der Programmlogik. Es empfiehlt sich hierbei, möglichst Generatoren unterschiedlicher Struktur zu kombinieren.

[20] Hierzu wurden jeweils 8 Bit zu einer Zahl zusammengefaßt.

Zahlen herrühren. Um die Verteilung zu glätten, werden daher jeweils zwei aufeinanderfolgende 8 Bit lange Zahlen mit Hilfe einer bitweisen XOR-Verknüpfung, d.h zu einer Ausgabezahl zusammengefaßt.
Eine besonders schöne Eigenschaft[21] dieser Art von Generator sei hier noch erwähnt: Alle relevanten Spannungen sind recht groß (mehrere Volt), so daß etwaige elektromagnetische Störsignale die Dynamik des Generators nur unwesentlich beeinflussen.
Solch ein System ist natürlich immer noch deterministisch, insofern kann man es als eine „physikalische Realisierung“ eines Pseudozufallsgenerators, nämlich der mathematisch modellierten Dynamik des Systems, bezeichnen. Es unterscheidet sich allerdings von einer (einfachen) Rechnersimulation der Dynamik dadurch, daß in einem realen physikalischen System immer kleine Rauschanteile die Variablen geringfügig verändern. Dies ist also ein Fall, in dem zwei unterschiedliche Funktionsprinzipien der Zufallserzeugung gleichzeitig eine Rolle spielen. Ob der Zufallsgenerator trotz seiner zufriedenstellenden statistischen Eigenschaften, aufgrund seines grundsätzlich deterministischen Charakters für kryptographische Anwendungen ungeeignet sein könnte[22], müßte vor einem Einsatz in der Praxis allerdings noch genauer geprüft werden.
Es ist interessant, diesen Generator mit Zufallsgeneratoren zu vergleichen, die zwar ebenfalls mit Oszillatoren arbeiten, allerdings nicht mit chaotischen [30, 41, 67]. Auch in diesem Fall wird die Übernahme eines Bitwertes eines Oszillators durch einen anderen gesteuert; hierbei handelt es sich bei dem Abtast-Oszillator, der die Übernahme steuert, um einen niederfrequenten Oszillator geringer Güte mit höherem Frequenz- und Phasenrauschen, während es sich bei dem anderen um einen stabilen hochfrequenten (8 MHz) Oszillator handelt. Die Zufälligkeit des Oszillators hängt kritisch von den Phasenschwankungen oder Frequenzfluktuationen des niederfrequenten Oszillators ab. Da die Fluktuationen eines einfachen Oszillator-Ansatzes nicht stark genug sind, verbessern Fairfield et al. [41] die statistischen Eigenschaften mit Hilfe eines zusätzlichen Verwürflers, außerdem wird eine Pseudoregularisierungslogik (s. S. 224) verwendet, um die relativen Häufigkeiten von Einsen und Nullen anzugleichen. Sowohl die Oszillatoren als auch der zusätzliche Verwürfler und die Pseudoregularisierungslogik werden digital ausgeführt, so dass sich diese Zufallsgeneratoren als integrierte Schaltkreise fertigen lassen [41]; lediglich das RC-Glied für den niederfrequenten Oszillator muß extern angeschlossen werden. Die Bitrate dieser Art von Zufallsgenerator hängt natürlich stark von den verwendeten Oszillatoren ab, für den oben beschriebenen Generator betrug sie allerdings lediglich 27 Bits/s.
Bei einem aktuellen Typ dieser Art von Generator[23] [30] werden die Frequenzfluktuationen durch Modulation des niederfrequenten Oszillators mit Hilfe einer Rauschspannung[24] erzeugt. Um den Einfluß von externen Störungen zu minimieren, wird als Rauschspannung die Spannungsdifferenz des verstärkten Johnson-Rauschens zweier nahe beieinander liegender Widerstände verwendet. Die statistischen Eigenschaften des Generators

[21] Man vergleiche dies mit Eigenschaften der in Abschnitt 2.2.2.1 aufgeführten Zufallsgeneratoren, die elektrisches Rauschen verwenden.

[22] Es wurden schon Verschlüsselungssysteme, die chaotische Systeme verwendeten, von Kryptanalytikern gebrochen, s. [11].

[23] Er ist in den Intel *82802 Firmware Hub* integriert, es handelt sich bei diesem um einen integrierten Baustein eines Computer-Hauptplatinen-Chip-Sets.

[24] Der interne Ursprung des Zufalls ist in diesem Fall also ein Typ von Zufallsgenerator, wie er in Abschnitt 2.2.2.1 besprochen wird, allerdings mit analogem Ausgangsignal.

werden verbessert durch eine auf dem Chip befindliche Signalverarbeitungslogik, die eine Variante[25] der von Neumann-Regularisierung (s. S. 40) durchführt. Die Bitraten des Zufallsgenerators liegen bei bis zu 75 kBit/s.
Abschließend sei noch kurz auf einen recht ausgefallenen Typ von Zufallsgenerator hingewiesen, der sich die komplexe Dynamik eines Systems zu Nutze macht: Der Lavalampen-Zufallsgenerator von SGI [115]. Bei diesem werden sechs Lavalampen[26] von einer CCD-Kamera aufgenommen. Bis auf die Blasen-Muster der unter konstanten Bedingungen betriebenen Lavalampen ändert sich das Bild nicht. Von Zeit zu Zeit wird aus jedem siebten[27] der 921.600 Byte Daten eines Einzelbildes mit Hilfe einer kryptographischen Hash-Funktion[28] eine 160 Bit lange Zahl erzeugt, die als Startwert für einen ebenfalls kryptographisch starken Pseudozufallszahlengenerator (BBS-Generator) verwendet wird.
Diese Vorgehensweise – eine (große) Menge strukturierter Daten werden mit Hilfe einer Hash-Funktion zu einer nicht aus den ursprünglichen Daten erschließbaren, kurzen Bitsequenz verlustbehaftet „komprimiert“ – ist eine beliebte Methode, um kurze, zufällige Bitsequenzen aus verhältnismäßig stark strukturierten Daten zu erzeugen. Die meisten E-Mail-Verschlüsselungsprogramme verwenden diese Vorgehensweise bei der Generierung von zufälligen Sitzungsschlüsseln, z.B. aus Mausbewegungen oder dem Tippverhalten des Nutzers.

2.2.2 Mikroskopische Vielteilchensysteme

Aufgrund der großen Anzahl innerer Freiheitsgrade zeigen mikroskopische Vielteilchensysteme bei der Messung makroskopischer Zustandsgrößen Schwankungen um einen Mittelwert. Dies läßt sich ebenfalls als Funktionsprinzip für einen Zufallsgenerator nutzen, indem man die makroskopische Größe periodisch mißt und den erhaltenen Meßwert mit einem Schwellwert, bevorzugt den Median der Meßgröße, vergleicht und anschließend auf eine der folgenden Weisen eine binäre Zufallsgröße generiert:

1. Liegt ein Meßwert über dem Schwellwert, so wird eine Eins generiert, liegt er darunter, wird eine Null ausgegeben.

2. Bleibt der Meßwert über oder unter dem Schwellwert, wird eine Null generiert; findet im Vergleich zum vorhergehenden Wert ein Unter- bzw. Überschreiten des Schwellwertes statt, wird eine Eins ausgegeben.

3. Die in Punkt 2 aufgeführte Vorgehensweise läßt sich noch erweitern, indem man die auf diese Weise generierten Einsen in einem durch einen externen Takt vorgegebenen Zeitintervall „zählt“ und anschließend die Parität des Zählwertes als binäre

[25] Die Details des Verfahrens sind aus „patentrechtlichen“ Gründen noch nicht offengelegt, aber die Unterdrückung von möglichen Korrelationen aufeinanderfolgender Bits der Rohdaten soll dabei verbessert worden sein.

[26] Es handelt sich dabei tatsächlich um die in letzter Zeit wieder in Mode gekommenen Design-Ungetüme, bei denen von einer Lampe aufgeheizte und angestrahlte Wachstropfen in einer Flüssigkeit noch oben schweben.

[27] Man nimmt diesen Wert, damit nicht immer dasselbe Byte eines Pixels, z.B. immer der Rot-Wert, verwendet wird.

[28] Es wird der vom National Institute of Standards standardisierte *Secure Hash Standard*-Algorithmus verwendet.

Zufallsvariable verwendet. Dies läßt sich effizient mit Hilfe eines als Modulo-2-Zähler fungierendem Flipflops[29] erledigen [110].

Die Zeit zwischen zwei Messungen am System darf bei allen drei Vorgehensweisen nicht zu kurz sein, da sonst die Gefahr besteht, daß die mikroskopischen Fluktuationen den Meßwert innerhalb der Zeitspanne prinzipiell nicht stark genug ändern können und somit Korrelationen zwischen aufeinanderfolgenden Meßwerten unvermeidlich werden. Überdies muß sichergestellt werden, daß es nicht möglich ist, durch eine (z.B. periodische) Störung des Systems von außen das Vielteilchensystem so zu beeinflussen, daß es zu einseitig veränderten Schwankungen kommt.

2.2.2.1 Elektronische Rauschgeneratoren

Zufallsgeneratoren, die auf diesem Prinzip basieren, werden in der Praxis besonders gern mit elektronischen Rauschquellen[30] realisiert, da dies den Aufbau vereinfacht und kleine Abmessungen ermöglicht, bis hin zu integrierten Schaltkreisen [30, 125] oder auch Chipkarten (z.B. [100], viele moderne Krypto-Chipkarten enthalten physikalische Zufallsgeneratoren). Der größte Teil der kommerziell erhältlichen Zufallsgeneratoren basiert ganz oder teilweise auf diesem Prinzip [1, 30, 59, 45, 58, 125, 35]! Die elektronischen Rauschquellen verwenden i. A. entweder das Johnson-Rauschen in Widerständen [24, 30, 125] oder das Rauschen von Halbleiter-Bauelementen [1, 59, 50, 58, 110]; hierbei werden besonders gern Zenerdioden eingesetzt, deren Rauschspannung bei Betrieb im Durchbruchsgebiet gemessen wird. Das Vielteilchensystem sind hierbei die Ladungsträger der Diode unter dem Einfluß der statistischen Wärmebewegung. Für eine ausführliche Darstellung der (Schaltungs-)Technik sei insbesondere auf die Dissertation von Manfred Richter [110] verwiesen, hier sei lediglich das generelle Vorgehen erläutert.
Typischerweise wird die Rauschspannung des Widerstandes oder der Zenerdiode entweder direkt oder nach einer Filterung[31] [1, 59] mit Hilfe eines Verstärkers verstärkt und anschließend dessen Ausgangssignal digitalisiert. Eine Variante besteht darin, daß man mit zwei Rauschquellen arbeitet und nach Verstärkung und Digitalisierung die beiden digitalen Signale logisch kombiniert [58]. Typische Bitraten[32] von Zufallsgeneratoren, die auf elektronischem Rauschen basieren, liegen zwischen 1 kBit/s [59] und 50 kBit/s [110]; höhere Bitraten lassen sich natürlich durch parallelen Betrieb mehrerer Generatoren erreichen [45, 95].
Abschließend sei noch ein interessanter Vorschlag von G.B. Agnew [4] für einen elektronischen Zufallsgenerator erwähnt, der sich vermutlich auch gut für eine Integration in VLSI-Chips eignet. Er schlägt hierzu vor, zwei Metall-Isolator-Halbleiter-Kondensatoren als Grundelemente des Generators zu verwenden. Diese Kondensatoren bestehen aus

[29] Zu Beginn des nächsten Intervalls wird der Wert des Flipflops in ein nachfolgendes Übernahme-Flipflop geschrieben und das Zähl-Flipflop wieder zurückgesetzt.

[30] Prinzipiell lassen sich natürlich auch anders geartete Rauschquellen verwenden, wie z.B. akustische.

[31] Meist wird das Rauschsignal wechselspannungsgekoppelt an den Verstärker gegeben. Werden keine allzu hohen Bitraten angestrebt, so ist es vorteilhaft das Rauschspektrum z.B. bei ca. 100 kHz abzuschneiden, um Probleme mit dem Ankoppeln von Radiofrequenz-Störsignalen (Rundfunksender!) zu minimieren.

[32] Eine ganze Reihe von Zufallsgeneratoren [1, 24, 58, 35] wird an die RS232-Schnittstelle angeschlossen, dies allein beschränkt bereits die maximal mögliche Bitrate. Durch die Verwendung schnellerer, serieller Schnittstellen, wie z.B. USB oder Firewire, ließe sich das Problem allerdings umgehen.

einem p-dotierten Halbleitersubstrat, auf das eine isolierende Oxidschicht mit Metallisierung aufgebracht wurde. Legt man eine positive Spannung zwischen Metallisierung und Substrat an, so bildet sich im Substrat ein Potentialtopf aus. Allerdings werden im Laufe der Zeit auch thermisch generierte Elektronen als Dunkelstrom[33] in das Gebiet fließen; dieser Prozeß gehorcht einer Poisson-Statistik und scheint auch für benachbarte Kondensatoren unkorreliert zu sein. Nach Abschalten der Spannung wird mit Hilfe eines ladungssensitiven Differenzverstärkers die Ladungsdifferenz zwischen den zwei Kondensatoren gemessen und aus dem Ergebnis ein Zufallsbit generiert, anschließend wird die Spannung wieder für eine gewisse Periode angelegt, um genügend Rauschelektronen für Generierung des nächsten Bits zu sammeln.
Aufgrund der relativ niedrigen Rauschspannungspegel müssen alle auf elektronischem Rauschen basierenden Zufallsgeneratoren sehr gut gegen äußere Einwirkungen, welche die statistischen Eigenschaften des Rauschens verändern, abgeschirmt werden. Angriffsszenarien auf Rauschgeneratoren reichen von gezielten Temperaturänderungen[34], Manipulationen an der Versorgungsspannung bis hin zur Beeinflussung des Generators durch Einstrahlen von Radiofrequenzen.

2.2.2.2 Optische Rauschgeneratoren

Bei optischen Rauschgeneratoren ist die Grenzziehung zu den in den nächsten Abschnitten folgenden Zufallsgeneratoren, die auf Messungen an quantenmechanischen Objekten basieren, nicht sehr trennscharf. Senkt man nämlich die Lichtintensität, so wird sich die Quantennatur des Lichts bemerkbar machen.
Dennoch sollen in diesem Abschnitt nur Generatoren erwähnt werden, welche die Quantennatur des Lichtes nicht direkt als Mechanismus der Zufallserzeugung verwenden, d.h. Generatoren, die sich prinzipiell auch noch im rein klassischen Wellenbild erklären lassen. Licht hat gegenüber elektrischen Phänomenen einen großen Vorteil, der auch bei der rein optischen Informationsverarbeitung genutzt wird, es eignet sich sehr gut zu Parallelisierung von Operationen. Diese Eigenschaft machen sich Marron et al. [87] in ihrem Zufallsmatrix-Generator (Random Array Generator) auf der Basis von Laser-Speckle-Mustern zu Nutze. Dieser Generator liefert gleich eine zweidimensionale binäre Matrix aus Null- und Eins-Werten, die überdies einer vom Anwender spezifizierten räumlichen Wahrscheinlichkeitsverteilung folgen. Das Muster wird hierzu von einem 2-dimensional räumlich auflösenden Detektor abgetastet und der digitalisierte Ausgabewert jedes einzelnen Detektor-Pixels anhand eines Schwellwertes, der aufgrund der gewünschten räumlichen Wahrscheinlichkeitsverteilung festgesetzt wird, auf die Werte Null bzw. Eins rechnerisch abgebildet. Wenn die Abbildungsoptik zwischen Laser und Detektor so justiert wird, daß die Specklegröße ungefähr der Pixelgröße des Detektors entspricht, sind die Werte benachbarter Pixel statistisch voneinander unabhängig.
In einer Folgearbeit derselben Forschungsgruppe [88] wird ein neuartiger Zufallsgenerator für zweidimensionale Zufallsvektoren präsentiert, der anders als der Zufallsmatrix-Generator keine nachfolgende rechnerische Bearbeitung der Detektorwerte benötigt, da die Wahrscheinlichkeitsverteilung bereits beim optischen Aufbau mitberücksichtigt wird.

[33] Licht kann ebenfalls freie Elektronen erzeugen, die den Potentialtopf füllen, dieses Prinzip wird bei den CCD-Kameras verwendet.
[34] Ein in der Praxis durchaus wirksamer Angriff: Kältespray.

Um dies zu erreichen, wird die räumliche Wahrscheinlichkeitsverteilung als Kontrollmuster auf ein Objekt[35] „aufgebracht“ und dieses mit Hilfe einer Optik auf ein Detektionssystem[36] abgebildet. Durch ein stark abschwächendes Grauglas vor dem Detektor wird die Leuchtdichte des Bildes auf dem Detektor so stark abgesenkt, daß räumlich aufgelöste Einzelquantendetektion sinnvoll möglich ist. Als Detektionssystem dient ein positionsempfindlicher, photonenzählender Detektor, der bis zu 10^5 Detektionsereignisse pro Sekunde noch räumlich aufgelöst detektieren kann. Jedes Detektionsereignis liefert einen zweidimensionalen Zufallsvektor.
Neben der Möglichkeit räumliche Wahrscheinlichkeitsverteilungen des Nutzers vorzugeben, erlaubt das Einbringen eines Kontrollobjektes „in“[37] den Strahlengang es auch, Abbildungsfehler der Optik und Abweichungen der räumlichen Detektionseffizienz von der Gleichverteilung beim Detektor auszugleichen.
Eine ähnliche Art von optisch basiertem Zufallsgenerator für zweidimensionale Zufallsvektoren haben unabhängig von obigen Arbeiten Tang et al. [124] entwickelt.

2.2.3 Quantenmechanische Zufallsgeneratoren

Bei quantenmechanischen Objekten ist der Zufall elementar, d.h. bereits ein einzelnes Objekt trägt ihn in sich. Befindet sich ein quantenmechanisches Objekt nicht gerade in einem bestimmten Eigenzustand einer zu messenden Größe, so ist der Wert dieser Meßgröße unbestimmt, solange man keine Messung vornimmt. Dies ist nicht lediglich eine Folge des fehlenden Wissens des Experimentators um den exakten Wert der Größe, sondern vielmehr ist die Größe *an sich* unbestimmt. War dies lange Zeit heiß umstritten, so kann doch nach den zahlreichen Messungen zum EPR-Paradoxon und der Verletzung der Bellschen Ungleichung (s. Abschnitt 3.1.1) davon ausgegangen werden, daß die Unbestimmtheit tatsächlich eine intrinsische Eigenschaft quantenmechanischer Objekte ist:
Es lassen sich nur Aussagen über Wahrscheinlichkeitsverteilungen machen; welches der möglichen Resultate bei einer einzelnen Messung realisiert wird, ist prinzipiell nicht sicher vorhersagbar. Es gibt keine *verborgenen Parameter*, die auf deterministische Weise den Ausgang einer Messung vorherbestimmen.
Will man ein quantenmechanisches Objekt als Zufallsquelle für einen Zufallsgenerator verwenden, so lassen sich verschiedene Wege einschlagen:

- Man verwendet die Unbestimmtheit eines Entstehungszeitpunktes[38]; z. B. den Emissionszeitpunkt eines Teilchens bzw. γ-Quants beim radioaktiven Zerfall oder den Emissionszeitpunkt eines Lichtquants bei einer stark abgeschwächten, aber kontinuierlich betriebenen Lichtquelle.

[35] In der Arbeit wurden zwei Varianten erprobt: ein von einer LED beleuchtetes Schwarzweiß-Negativ und ein Videomonitor, der vom selben Computer angesteuert wurde, der auch die Zufallsmatrizen abspeicherte.

[36] Innerhalb der Klassifizierung der Zufallsgeneratoren gehört dieser Generator eigentlich in die Kategorie der „Messung an quantenmechanischen Objekten“, aber aufgrund seiner Ähnlichkeit mit dem klassischen Vorläufer, wird er hier auch im Anschluß an diesen aufgeführt.

[37] Im Falle des Videomonitors ist das Videobild natürlich der Ausgangspunkt des Strahlenganges.

[38] Natürlich läßt sich alternativ auch die Länge des Zeitintervalles zwischen zwei zufälligen Emissionen als Zufallsvariable verwenden.

- Man nutzt die Unbestimmtheit einer Emissionsrichtung, indem man z.B. den Detektionsort eines Teilchens oder Quants auf einem zweidimensional räumlich auflösenden Detektor als Zufallsvariable verwendet.

- Man verwendet eine *Welcher-Weg-Entscheidung* einzelner quantenmechanischer Objekte; dies geschieht bei dem weiter unten ausführlich vorgestellten quantenoptischen Zufallsgenerator.

Natürlich sind auch Kombinationen dieser Zufallsmechanismen möglich.

2.2.3.1 Zufallsgeneratoren auf Basis des radioaktiven Zerfalls

Bei Zufallsgeneratoren, die den radioaktiven Zerfall als Zufallsmechanismus verwenden, nutzt man den Umstand, daß es unmöglich ist, zu wissen, wann ein einzelner Atomkern zerfällt bzw. ein Teilchen emittiert. Bei Kernreaktionen, wie z.B. der Emission eines α-Teilchens, spielt bei der Überwindung der Aktivierungsenergie der „Tunnel-Effekt“ eine große Rolle. Bei ihm handelt es sich um einen quantenmechanischen, nichtdeterministischen Effekt. Daher lassen sich auch lediglich Wahrscheinlichkeitsaussagen machen, wann ein Kern ein Teilchen aussendet.
Experimentelle Realisierungen von Zufallsgeneratoren, die radioaktive Quellen[39] verwenden [5, 50, 60], benutzen meist β-Strahler (insbesondere ^{137}Cs). Detektiert wird hierbei die γ-Strahlung bei Niveauübergängen im Tochterkern [60, 50].
Wurden in älteren Arbeiten von Inoue et al. [60] und M. Gude [50] mit flüssigem Stickstoff gekühlte Ge-Li-Detektoren verwendet, was für einen praktischen Einsatz außerhalb eines Labors zu aufwendig ist, so kann man neueren Arbeiten von John Walker [126] und Ricardo Aguayo et. al. [5] aber entnehmen, daß dies keineswegs unbedingt notwendig ist; handelsübliche Geiger-Müller-Zählrohre sind vollkommen ausreichend[40].
Die Generierung der Zufallsbits erfolgt aus den beiden Zeitintervallen zwischen drei Zählereignissen[41]. Ist das erste Intervall größer als das zweite, wird ein Zufallsbit mit dem Wert Null erzeugt und im umgekehrten Fall eines mit dem Wert Eins. Im relativ unwahrscheinlichen Fall, daß die beiden Intervalle gleich lang sind, wird gar kein Bit ausgegeben. Der große Vorteil dieser Vorgehensweise besteht darin, daß man keine zusätzliche, nachfolgende Regularisierung der Zufallsbits benötigt, da bei hinreichend konstanten Emissionsraten die Wahrscheinlichkeit für den Bitwert Null gleich der für den Bitwert Eins ist!
Ein Nachteil von Zufallsgeneratoren auf Basis des radioaktiven Zerfalls besteht natürlich darin, daß die verwendete Quelle im Laufe der Zeit schwächer wird; allerdings wirkt sich dies primär auf die Bitrate des Generators aus, die statistischen Eigenschaften der

[39] Im Prinzip kann man natürlich auch die kosmische Hintergrundstrahlung oder andere natürliche Quellen der Radioaktivität verwenden, allerdings kommt man in diesem Fall (hoffentlich!) nur auf sehr geringe Bitraten (< 1 Bit/s).

[40] Aware Electronics Corporation bietet daher zusätzlich zu ihren Strahlungsmeßgeräten auch gleich ein PC-basiertes Programm zur Zufallszahlen-Erzeugung aus den Meßdaten an, s. [27].

[41] In [60] wird ein alternatives Verfahren angewandt: Die Zeitwerte (entspr. Anzahl der Clock-Ticks) von drei Intervallen werden zusammengezählt und die niederwertigste Dezimalstelle wird direkt als Zufallszahl ausgegeben. Dieses Vorgehen ist effizienter, führt aber zu (sehr) geringen Abweichungen von der Idealverteilung.

Zufallsbits selbst werden nicht merklich verändert[42], aufgrund der Kürze der Intervalle zwischen zwei Zählereignissen im Verhältnis zur Halbwertszeit der Quelle [5, 126].
Auch wenn die von diesen Zufallsgeneratoren erzeugten Zufallszahlen sehr gut den Erwartungen entsprechen, so werden die Generatoren in der Praxis doch recht selten eingesetzt. Der Hauptgrund[43] hierfür ist in den zusätzlichen Schutzmaßnahmen zu sehen, die auf Grund der Verwendung von radioaktiven Quellen getroffen werden müssen. Überdies ergeben sich gerade im kommerziellen Bereich noch zusätzliche versicherungsrechtliche Fragen[44] und damit verbundene Kosten[45]. Zudem ist die Akzeptanz dieser Technik bei den Beschäftigten nicht besonders hoch, zumal es Alternativen beim Aufbau physikalischer Zufallsgeneratoren gibt.

2.2.3.2 Zufallsgeneratoren auf Basis einer unbestimmten Emissionsrichtung

Da bei diesem Typ von Zufallsgenerator die Unbestimmtheit der Emissionsrichtung die entscheidende Voraussetzung für die Zufälligkeit der generierten Bits ist, muß unbedingt sichergestellt werden, daß weder Schwankungserscheinungen bei der physikalischen Quelle noch äußere Einflüsse die Unbestimmtheit der Emissionsrichtung zerstören. Auch hier lassen sich wieder die verschiedensten Emissionsprozesse verwenden, so schlagen z.B. Edelkind et al. [33] vor, die Unbestimmtheit der Emissionsrichtung einer radioaktiven Quelle zu verwenden.
Einen optischen Zufallsgenerator, der sich die unbestimmte Emissionsrichtung zunutze macht, wurde von Á. Stefanov et al. [120] entwickelt; dies wird auf sehr geschickte Weise erreicht: Das Licht einer gepulsten LED ($\lambda = 830$ nm) wird in eine 2 m lange Einmoden-Glasfaser eingekoppelt; am Ausgang dieser Glasfaser befindet sich das Licht in einer einzigen räumlichen Mode, die aufgrund des kleinen Kerndurchmessers der Glasfaser als Kugelwelle abgestrahlt wird. Positioniert man nun im Abstand einiger Millimeter gegenüber der Austrittsfläche der Einmoden-Glasfaser zwei dicht parallel nebeneinander liegende Mehrmoden-Glasfasern, so ist gewährleistet, daß nicht vorherbestimmt werden kann, in welche dieser beiden Glasfasern ein abgestrahltes Photon[46] eingekoppelt wird. Die Ausgangs-Lichtfelder der beiden unterschiedlich langen Mehrmoden-Glasfasern fallen auf *einen* Detektor, wobei aufgrund der unterschiedlichen Längen der beiden Glasfasern, die eine Laufzeitdifferenz von ca. 60 ns bedingen, durch Koinzidenz-Messung mit dem elektrischen Ansteuerpuls[47] für die LED leicht unterschieden werden kann, aus welcher der Glasfasern ein Photon auf den Detektor fällt. Wenn die LED mit einer Puls-

[42] Selbst wenn sich dies bemerkbar machte, führte es lediglich zu ungleichen Wahrscheinlichkeiten für Einsen und Nullen; diese ließen sich mit den in Abschnitt G besprochenen Regularisierungsmethoden leicht angleichen.

[43] Hinzu kommt noch, daß die Bitraten bei der (üblichen) Verwendung von radioaktiven Standard-Eichquellen bei weniger als ca. 2000 Bits/s liegt; dies ließe sich aber vermutlich noch etwas verbessern.

[44] Insbesondere die Brandschutzversicherung ist kritisch, da sich die Folgekosten im Brandfalle durch eine mögliche Kontaminierung erhöhen.

[45] Es handelt sich bei dem Einsatzort i.a. *nicht* um ein Physiklabor!

[46] Durch das Einkoppeln in die Einmoden-Glasfaser wird das Lichtfeld bereits stark abgeschwächt, über die Entfernung der beiden Mehrmoden-Glasfasern von der Austrittsöffnung der Einmoden-Glasfaser und die elektrische Ansteuerung der LED läßt sich erreichen, daß die mittlere Intensität lediglich 0.1 Photon pro Lichtpuls beträgt, s. a. Abschnitt 3.1.2.

[47] Hierbei müssen natürlich zusätzliche Verzögerungen berücksichtigt und noch ein zusätzlicher 60 ns verzögerter Puls generiert werden.

rate von 1 MHz angesteuert wird, können auf diese Weise Zufallsbits mit einer Rate bis zu 100 kBit/s erzeugt werden, wobei es sich natürlich um Rohdaten handelt, die noch regularisiert werden müssen.

2.2.3.3 Zufallsgeneratoren auf Basis stark abgeschwächter Lichtfelder

Takeuchi et al. [122] haben für Anwendungen in der Kernphysik[48] einen Zufalls-Pulser entwickelt. Bei diesem Pulser emittiert eine kontinuierlich betriebene LED Photonen hin zu einem (miniaturisierten) Photomultiplier, der sie detektiert. Die Ausgangspulse des Photomultipliers werden verstärkt, diskriminiert und auf eine kurze elektrische Pulsbreite gebracht.
Damit der Zufalls-Pulser möglichst dem Ideal der Poisson-Statistik entspricht,wird mit Hilfe eines auf Auf-Ab-Zählern und einem Digital-Analog-Wandler basierenden Regler die LED gezielt angesteuert. Denn ähnlich wie bei einer radioaktiven Quelle muß man dafür sorgen, daß die mittlere Rate möglichst konstant bleibt[49]. Zwar wurde dieser Zufalls-Pulser nicht zum Generieren von Zufallsbits entwickelt, aber seine elektrischen Ausgangspulse lassen sich natürlich mit Hilfe der gleichen Methodik wie bei den auf Radioaktivität basierenden Zufallsgeneratoren zum Erzeugen von Zufallsbits verwenden.
Beim historisch ersten Experiment zur Quantenkryptographie verwenden Bennett et al. [10] Daten, die mit einem optischen Zufallsgenerator generiert wurden, zum zufälligen Ansteuern von Pockelszellen. Der Generator, mit dem diese Daten erzeugt werden, besteht aus einer LED, die stark abgeschwächte Lichtpulse abstrahlt, und einem direkt im Ausgangslichtfeld der LED stehenden Photomultiplier. Die Intensität des Lichtfeldes wird so gewählt, daß nur in der Hälfte aller Detektionszeitfenster Ereignisse registriert werden, wobei die Fenster mit Ereignis einen Bitwert von Eins generieren und die ohne Ereignis eine Bitwert von Null. Die so generierten Rohdaten werden noch einer von-Neumann-Regularisierung (s. Abschnitt 3.1.3) unterzogen und zusätzlich mit einer Pseudozufallszahlenfolge XOR-verknüpft, s. a. Abschnitt G.2.2, bevor sie verwendet werden.
Als Komponente für Experimente zum Test der Bellschen Ungleichung [127] bei strikter raumartiger Trennung der Detektionen wurde von T. Jennewein und U. Achleitner [63] von der Universität Innsbruck ein Zufallsgenerator gebaut, der eine kontinuierlich betriebene LED verwendet. Als Zufall generierendes Element (s. a. Abschnitt 3.1.1) dient ein Strahlteiler mit zwei Photomultipliern zur Detektion der Photonen in den Ausgangsarmen. Je nachdem, welcher der beiden Photomultiplier ein Lichtquant[50] detektiert, wird ein Bitwert Eins bzw. Null als Zufallsbit erzeugt. Sprechen beide Photomultiplier in einem Zeitintervall von $\Delta t \leq 2$ ns an, werden beide Ereignisse ignoriert. Die Quellenintensität der LED wurde so einreguliert, daß das mittlere Zeitintervall zwischen zwei Detektionsereignissen ca. 10 ns betrug, um eine möglichst hohe Rate von Zufallsbits für die Experimente zur Verfügung zu haben.

[48] Die von dem Zufalls-Pulser erzeugten elektrischen Pulse sollen möglichst genau der Poisson-Statistik bei einer bestimmten mittleren Pulsrate entsprechen, und daher als Modell für einfache Zerfälle dienen können.

[49] Bei einer radioaktiven Quelle ist dies bei entsprechend langer Halbwertszeit, i. A. in sehr guter Näherung automatisch erfüllt.

[50] Es können bei der Verwendung von Poisson-Lichtfeldern natürlich auch mehrere sein, s. Abschnitt 3.1.2.

2.2.3.4 Zufallsgeneratoren und Quantenkryptographie

Unter die Bezeichnung Quantenkryptographie [57, 104] werden alle Verfahren subsumiert, die quantenmechanische Zustände verwenden, um eine abhörsichere[51] Übertragung von Daten zu erreichen. Hierbei findet allerdings i. a. keine Übertragung von Nutzdaten statt, sondern eine Generierung von korrelierten Zufallssequenzen an den beiden Endpunkten der Quantenkryptographie-Strecke. Den korrelierten Zufallssequenzen können anschließend die geheimen Schlüssel entnommen werden für die algorithmische Chiffrierung der Daten einer herkömmlicher Datenübertragung. Die Quantenkryptographie dient also zum Schlüssel-„Austausch“ zwischen den beiden Endpunkten.
Betrachtet man, was Abhören im üblichen, d.h. klassischen Sinne, bedeutet, so erkennt man schnell, daß es entweder ein Abzapfen eines Teils des Übertragungssignales oder ein Kopieren des Signals (ohne merkbare Beeinträchtigung des Original-Signals) voraussetzt. Verwendet man allerdings für die Übertragung quantenmechanische Zustände, so lassen sich durch die adäquate Wahl der Zustände, die oben genannten Manipulationen bemerken.
So erlaubt z.B. die Verwendung von einzelnen Photonen als Informationsträger, das Anzapfen des Signals gut zu erkennen, da man von einem einzelnen Photon nichts „Abzapfen“ kann, ohne daß dies dem legitimen Empfänger auffällt. Allerdings schützt die Verwendung von einzelnen Photonen *allein* nicht vor dem Kopieren des Signals, hierzu bedarf es der Wahl bestimmter Einphotonen-Anzahlzustände, z.B. von Überlagerungszuständen aus den Zuständen zweier nichtorthogonaler Basen[52], die sich nicht unverfälscht kopieren lassen [129]. Es ist gerade die Heisenbergsche Unbestimmtheitsrelation, die das *exakte* Kopieren, das einer unverfälschten Messung zweier nichtkommutierender Variablen entspräche, verbietet.
Will man Daten[53] mit Hilfe solcher Zustände übertragen, so wird jedem Bitwert jeweils ein Zustand von i. a. zwei Zuständen der beiden nichtorthogonalen Basissysteme zugeordnet. Bei der Übertragung wählt dann der Sender die Basis für jedes zu übertragende Bit *zufällig* aus und schickt den entsprechenden Zustand über eine freistrahl- oder faseroptische Übertragungsstrecke zum Empfänger, der ebenfalls *zufällig* und unabhängig vom Sender eine Basis wählt, in der er den Zustand mißt. Meist wird die zufällige Wahl der Basis mit Hilfe eines externen Zufallsgenerators und elektrooptischen Komponenten realisiert [10]. In einer Veröffentlichung von J. G. Rarity et al. [108] wird hingegen die Verwendung von Zufallsgeneratoren auf Basis von Strahlteilerwürfeln oder Faserkopplern vorgeschlagen, die keinen[54] externen Zufallsgenerator benötigen, da sie selbst bereits passiv Zufall generieren. Der Vorschlag ist besonders elegant, da weitgehend auf zusätzliche elektronische Komponenten verzichten werden kann.

[51] Strenggenommen handelt es sich nicht um eine *abhörsichere*, sondern um eine *abhörsensitive* Übertragung, d.h. man kann mit einer hohen Wahrscheinlichkeit erkennen, ob abgehört wird oder nicht.

[52] Bei entsprechendem Entwurf des Übertragungssystems [47] lassen sich vermutlich [102, 48] auch orthogonale Zustände verwenden.

[53] Nutzdaten überträgt man auf diese Weise besser nicht, da der weitaus größte Teil der Quanten verloren geht, mithin also eine sehr große Redundanz eingebaut werden müßte.

[54] Dies gilt allerdings nur für die Quantenkryptographie mit Photonenpaaren (s. z. B. [34]), hier kann nämlich sowohl bei Sender als auch bei Empfänger *passive* Randomisierung eingesetzt werden, bei der immer ein Photon vernichtet wird. Die auf schwachen Lichtpulsen basierende Quantenkryptographie erlaubt dies nicht, da immer nur jeweils ein Lichtpuls zur Verfügung steht; daher muß auf der Sendeseite die herkömmliche Vorgehensweise verwendet werden, bei der *aktiv* und ohne Verlust randomisiert wird.

Kapitel 3

Theoretische Grundlagen

3.1 Quantenoptische Zufallsgeneratoren

Photonen werden durch die Parameter Energie, Impuls und Polarisation beschrieben. Photonen haben keine innere Struktur, sie sind im wahrsten Sinne des Wortes *elementar*: Das Photon läßt sich nicht weiter in Untereinheiten auflösen und analysieren. Auch gibt es bei Photonen keine internen Parameter, die lediglich nicht zugänglich sind und eventuell die quantenmechanischen Eigenschaften subtil beeinflussen könnten. Da sich Photonen überdies leicht erzeugen und hinreichend gut detektieren lassen, andererseits aber ohne direkte Eingriffe in den Strahlengang nicht zu beeinflussen sind, bieten sie sich in idealer Weise als Bestandteil eines Zufallsprozesses an.

3.1.1 Der Zufallsprozeß beim quantenoptischen Zufallsgenerator

Die verschiedenen, im Rahmen der vorliegenden realisierten quantenoptischen Zufallsgeneratoren nutzen alle einen *genuin* zufälligen, elementar[1] quantenmechanischen Prozeß, um Zufall zu generieren: Die *Welcher-Weg*-Entscheidung eines Photons am (50:50) Strahlteiler [49, 108] mit anschließender Detektion des Photons in den Ausgängen des Strahlteilers, s. Abb. 3.1. Das Teilungsverhältnis von 50:50 bezieht sich dabei immer auf eine feste Polarisations- und Ausbreitungsrichtung des Photons und eine feste Zentralwellenlänge[2]. Unter diesen Voraussetzungen ist der Prozeß ideal für einen Zufallsgenerator geeignet, da das Photon elementar und seine Wechselwirkung mit dem Strahlteiler nicht deterministisch ist, d.h. man kann bei jedem einzelnen der Photonen nicht vorhersagen, ob man es im Transmissions- oder im Reflektionsarm des 50:50-Strahlteiler detektieren wird. Es gibt keine verborgenen Parameter, die dies in irgendeiner Weise festlegen.
Daß dies so ist und nicht nur aufgrund fehlender Kenntnisse einer wie auch immer gearteten inneren Struktur des Photons so erscheint, wird durch eine große Anzahl im-

[1] Mit „elementar“ ist hierbei gemeint, daß die Wechselwirkung der Photonen mit dem Strahlteiler *formal* elementar beschreibbar [42, 105] ist; der Strahlteiler selbst ist als makroskopisches Gebilde natürlich alles andere als elementar.

[2] Inwiefern dies notwendig ist, inwieweit es bei den praktischen Aufbauten erreicht wird und welche Verbesserungsmöglichkeiten vorhanden sind, s. Abbschnitte 4.4 und 6.2.3.

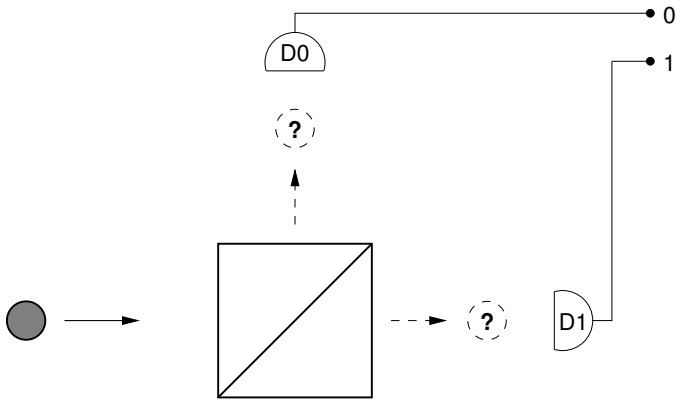


Abbildung 3.1: *Welcher-Weg*-Entscheidung am Strahlteiler

mer weiter verfeinerter Experimente[3] zum Einstein-Podolsky-Rosen-„Paradoxon“ und zur Verletzung der Bellschen Ungleichung nahegelegt. Bei diesen Experimenten werden die beiden hinsichtlich nichtkommutierender Observablen miteinander verschränkten Photonen eines Photonenpaares an zwei raumartig voneinander getrennten Orten einer Messung dieser Observablen unterzogen, wobei an beiden Orten zufällig gewählt wird, welche der beiden Observablen gemessen wird. Bisher haben die auf diese Weise gewonnen Meßreihen immer wieder die These bestätigt, daß lokale Theorien verborgener Parameter[4] mit den Ergebnissen nicht in Einklang zu bringen sind, wobei allerdings eingestanden werden muß, daß aufgrund gewisser experimenteller Randbedingungen, wie z.B. der Detektionseffizienz der Detektoren, noch gewisse „Schlupflöcher“ existieren, die ein vollständiges Verwerfen lokaler Theorien verborgener Parameter bisher (noch) nicht erlauben. Die oben erwähnten Experimente schließen *nichtlokale* Theorien verborgender Parameter zwar *nicht* aus, diese würden aber implizieren, daß sich die raumartig voneinander getrennten Photonen dennoch instantan beeinflussen könnten, was ziemlich fragwürdig erscheint und den Begriff des Determinismus bis zur Unkenntlichkeit verändert.

Es sei an dieser Stelle noch einmal ausdrücklich betont: Es wird in der vorliegenden Arbeit davon ausgegangen, daß Photonen keine verborgenen Parameter besitzen. Es wird also *vorausgesetzt*, daß sich die einzelnen Photonen an einem Strahlteiler tatsächlich zufällig verhalten und nicht in vorausbestimmter Weise. Die Frage inwieweit Messungen, die mehrere mögliche Ergebnisse haben können, an einem Quantensystem tatsächlich gemäß der vorhergesagten Wahrscheinlichkeiten zufällig sind, wird in der vorliegenden Arbeit also nicht erörtert, für diese Frage sei auf die Arbeiten von Erber et. al. [38, 39] zu Quantensprüngen in Einzelatomen verwiesen.

Unabhängig von der Frage der verborgenen Parameter läßt sich allerdings eine hypothetische Beeinflussung der Photonen untereinander durch Betriebsbedingungen ausschlie-

[3] Stellvertretend für die große Zahl von Experimenten sei hier nur auf zwei neuere Experimente zum Test der Bellschen Ungleichung [127, 133] und einen Übersichtsartikel [54] verwiesen.

[4] „Anschaulich“ stelle man sich hierbei das Photon als ein Teilchen mit einem (komplizierten) inneren „Mechanismus“ vor, der bei der Emission (durchaus mit jeweils unterschiedlichen Anfangswerten) in Gang gesetzt würde und zu jedem Zeitpunkt das „Verhalten“ des Photons bei den verschiedenen möglichen Messungen determinierte.

ßen, die dafür sorgen, daß während der Zeitspanne, in der das aktuelle Photon mit dem Strahlteiler wechselwirkt und anschließend von einem der Quantendetektoren registriert wird, das nachfolgende Photon noch gar nicht erzeugt wurde.
Als weiterer Vorteil des erwähnten Prozesses kommt noch hinzu, daß Photonen sich in linearen Medien nicht durch äußere elektromagnetische Felder beeinflussen lassen.
Dieser Prozeß erfüllt daher bereits wesentliche Anforderungen an physikalische Zufallsgeneratoren, wie sie in der Einleitung, Abschnitt 1.1 aufgeführt wurden und läßt sich überdies mit verfügbarer Technik gut realisieren. Man hat hier also eine Art von „Quanten-Münzwurf“ vorliegen, der allerdings sehr viel besser, „zufälliger“, als der makroskopische Wurf einer Münze ist und sich daher *im Prinzip* gut als glaubwürdiges „Zufalls-Standardnormal“ eignet.
Wenn allerdings von einem Teilungsverhältnis von 50:50 die Rede ist, so bezieht sich dies nur im theoretischen Idealfall allein auf den Stahlteiler[5]. Im Experiment muß man einerseits noch die spektrale Breite des Lichtes und andererseits die Eigenschaften der Detektoren mitberücksichtigen; denn erst, wenn ein Photon detektiert wurde, läßt sich sagen, welchen Weg es genommen hat! Detektierte man jedes Photon in einem der Ausgänge des Strahlteilers unabhängig von der Wellenlänge des Photons mit einer Quanteneffizienz von 100% (d.h., wenn ein Photon sich im Ausgang befindet, so wird es auch detektiert), dann hinge die Wahrscheinlichkeit, ein einzelnes Photon in dem einen oder dem anderen Ausgang zu finden, tatsächlich nur vom Transmissions- bzw. Reflektionskoeffizienten des Strahlteilers ab, die bei Verwendung eines Breitband-Strahlteilers auch für unterschiedliche Wellenlängen des spezifizierten Bereiches recht nahe beim idealen Teilungverhältnis liegen.
In der Realität ist die Detektion eines Photons leider nicht so effizient (s. Abschnitt 4.4.1.2) und außerdem stark von der Wellenlänge des Photons abhängig. Zudem können die Eigenschaften der beiden in den Ausgängen stehenden Detektoren voneinander abweichen, insbesondere werden i. a. die Quanteneffizienzen der Detektoren nicht genau gleich sein[6]. Daher wird sich in der Praxis das Teilungsverhältnis immer auf die Gesamtheit von Strahlteiler und Detektoren beziehen. Aufgrund der in das Teilungsverhältnis eingehenden Detektionseffizienzen läßt sich auch nicht ausschließen, daß Temperatureffekte eine Rolle spielen, da diese die Detektionseffizienz der Detektoren beeinflussen können. Somit kann selbst bei einer festen Wellenlänge der Photonen i. a. *nicht* von einem konstanten Teilungsverhältnis ausgegangen werden, auch wenn die Veränderungen nach einer gewissen Aufwärmphase nur gering sein werden, s. S. 154.
Welche Auswirkungen die spektrale Breite der Photonen im Zusammenspiel mit der Quanteneffizienz der Detektoren hat, wird im Abschnitt 6.2.3 diskutiert.

[5] Hierbei ist überdies zu bedenken, daß das Teilungsverhältnis eines Strahlteilers, selbst wenn nur eine Wellenlänge zu berücksichtigen ist und nicht eine ganzer Wellenlängenbereich immer nur bis zu einer gewissen Toleranz spezifiziert ist.

[6] In der Tat war dies beim Laboraufbau auch so.

3.1.2 Erzeugung einzelner Photonen

Die meisten Lichtquellen emittieren Photonen gemäß einer Bose-Einstein-Verteilung[7] oder einer Poisson-Verteilung. So sendet der bei quantenoptischen Experimenten gern verwendete Laser (idealisiert) sogenannte „kohärente Zustände“ mit einer mittleren Photonenanzahl $\bar{n}$, einer Varianz $\bar{n}$ und somit einer Standardabweichung $\sqrt{\bar{n}}$ aus; die Wahrscheinlichkeit, n Photonen in einem Zeitintervall zu emittieren, ist gegeben durch:

$$P_{Poisson}(n) = \frac{\bar{n}^n}{n!} \cdot e^{-\bar{n}}$$

Es läßt sich also nicht *sicher* vorhersagen, ob eine kontinuierliche Lichtquelle während eines betrachteten Zeitintervalles oder eine gepulste Quelle in einem ausgesandten Lichtpuls gerade *genau* ein Photon emittiert oder gar keines bzw. mehrere auf einmal. Zur Illustration sind in den Abbildungen 3.2 und 3.3 die Wahrscheinlichkeiten aufgetragen, n Photonen im Puls eines idealen Lasers mit einer mittleren Photonenzahl $\bar{n}$ von 1 bzw. 0,1 zu finden. Man kann hierbei gut erkennen, daß man eine sehr niedrige mittlere Photonenrate – gern wird der Wert $\bar{n} = 0{,}1$ verwendet – braucht, um die Wahrscheinlichkeit für das gleichzeitige Auftreten von mehreren Photonen gering zu halten.
Eine ideale Einphotonenquelle hingegen hätte eine Verteilung, bei der lediglich für $n = 1$ die Wahrscheinlichkeit $P(n)$ gleich Eins wäre und bei allen anderen Werten gleich Null. Eine *reale* Einphotonenquelle auf Basis der parametrischen Fluoreszenz, s. Abschnitt 3.1.2.1, kommt diesem Ideal schon recht nahe, da sich eine Wahrscheinlichkeit von $P(1) > 0{,}9$ erreichen läßt, wenn man von einem Anteil von zufälligen Koinzidenzen[8] von 5–10% ausgeht, s. a. [55].
In vielen Forschungsinstituten wird bereits seit geraumer Zeit an der Entwicklung von Lichtquellen gearbeitet, die Anzahlzustände abstrahlen können. Diese ohnehin schon schwierige Aufgabe wird noch zusätzlich dadurch erschwert, daß für die meisten Experimente und praktischen Anwendungen ein möglichst gerichtetes Aussenden der Anzahlzustände wichtig ist. Lediglich der einfachste Fall – einzelne Photonen – läßt sich bisher verhältnismäßig gut realisieren.
Die verschiedenen Methoden, um gesteuert oder zumindest gesichert[9] Lichtfelder mit möglichst geringer, im Idealfall verschwindender, Varianz um die mittlere Photonenanzahl zu erzeugen, lassen sich nach der Vorgehensweise gliedern:

Verringerung der Varianz durch parametrische Verstärkung Dieser Vorschlag von D. Stoler [121] war einer der ersten, in dem beschrieben wurde, wie man die Schwankungen der Photonenanzahl um die mittlere Photonenzahl herkömmlicher Laser-Lichtfelder verringern könnte. Erreicht werden sollte dies mit Hilfe eines entarteten optischen parametrischen Verstärkers. Er sollte hierzu mit einer frequenzverdoppelten Lichtwelle gepumpt werden, die vorher mit Hilfe eines nichtlinear optischen Kristalls aus der

[7] Die Kohärenzzeit thermischer Lichtquelle ist sehr viel kleiner als die bestmögliche zeitliche Auflösung der verwendeten Detektoren, deshalb geht die Statistik der gezählten Photonen in den Grenzfall einer Poisson-Verteilung über, s. [84], S. 233.

[8] Diese machen sich dann in einer von Null verschiedenen Wahrscheinlichkeit für $P(0)$ bemerkbar, d.h. $P(0) = 1 - P(1)$.

[9] Der Unterschied ist folgender: Bei der gesteuerten Erzeugung kann man den Emissionszeitpunkt des Photons selbst bestimmen, bei der gesicherten Erzeugung, weiß man ihn lediglich.

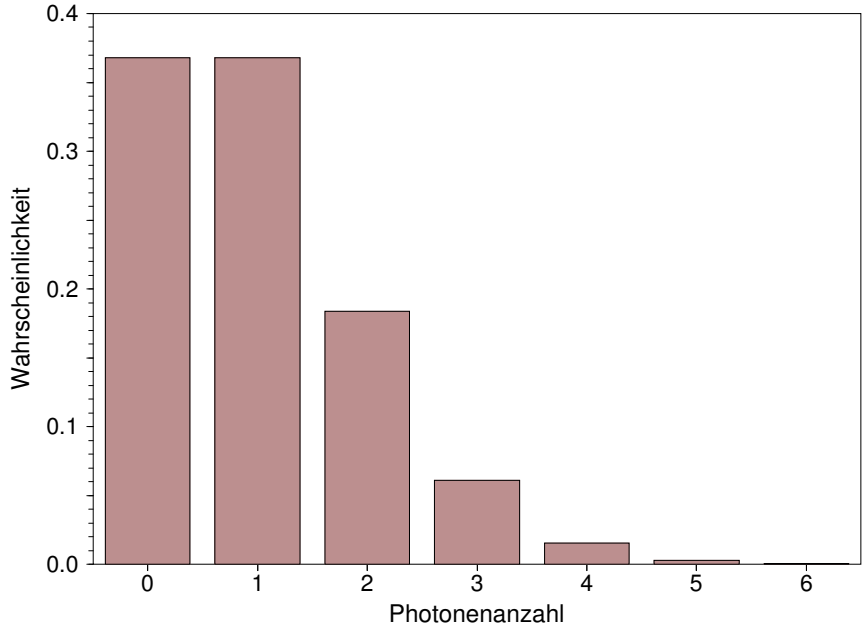


Abbildung 3.2: Wahrscheinlichkeit für eine Anzahl von n Photonen in einem kohärenten Puls mit einer mittleren Photonenanzahl $\bar{n} = 1$

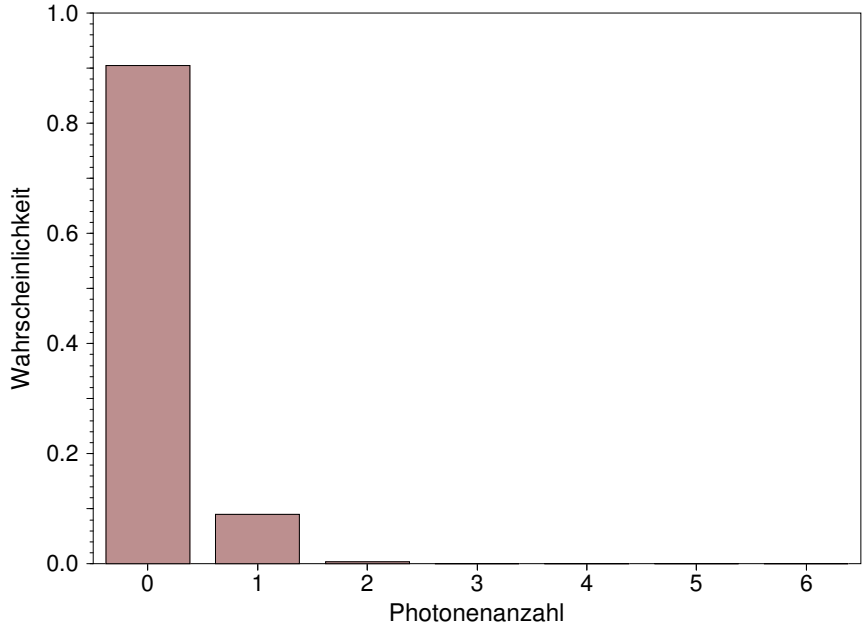


Abbildung 3.3: Wahrscheinlichkeit für eine Anzahl von n Photonen in einem kohärenten Puls mit einer mittleren Photonenanzahl $\bar{n} = 0{,}1$

Fundamentalen durch Frequenzverdopplung erzeugt wird. Die fundamentale Welle dient dann beim entarteten Fall des parametrischen Verstärkers, für beide, in diesem Spezialfall kollineare, Eingangswellen. Durch entsprechende Wahl der Parameter des parametrischen Verstärkers sollte sich theoretisch die Photonenstatistik auf der Fundamentalen dadurch beeinflussen lassen, daß durch Umwandlung je zweier Photonen der Fundamen-

talen in ein Photon der Pumpwellenlänge die Tendenz der Photonen zum „Klumpen“ (*Bunching*), d.h. dem verstärkten Auftreten von mehreren Photonen gleichzeitig, verringert wird. Man bezeichnet dies auch als „Anti-Bunching“. Da die Umwandlung bei hohen Photonenzahlen wahrscheinlicher ist als bei geringen, wirkt der parametrische Verstärker also wie ein nichtlinearer Absorber.

Verringerung der Varianz durch Rauschunterdrückung direkt bei der Erzeugung Hierfür werden Leucht- oder Laserdioden, die einen hohen Wirkungsgrad besitzen müssen, mit einem sehr rauscharmen Elektronenstrom betrieben, um diese Eigenschaft auf das Lichtfeld zu übertragen. Auf diese Weise[10] gelingt es allerdings wie beim vorangegangenen Vorschlag lediglich, die Varianz um den Mittelwert $\bar{n}$ etwas zu reduzieren [130] (sogenanntes „*Amplituden-Squeezing*“). Eine Realisierung von Photonen-Anzahlzuständen, die eine exakte Anzahl von Photonen haben, d.h. bei denen die Varianz verschwindet, ist auf diesem Wege noch nicht gelungen.

Verwendung selektiv getriebener singulärer Emitter Bei diesen Verfahren werden keine makroskopischen Quellen eingesetzt, sondern vereinzelte, quantenmechanische Systeme, wobei bisher einzelne Farbstoffmoleküle, Quantenpunkte und NV-Zentren verwendet wurden. Die Anti-Bunching-Eigenschaften des Lichtfeldes kommen bei diesem Ansatz dadurch zustande, daß der singuläre Emitter nach Emission eines Photon erst nach einer gewissen Zeitspanne in der Lage ist, ein weiteres Photon zu emittieren.
Ein Ansatz auf diesem Wege eine Einphotonenquelle zu bauen, stammt von Brunel et. al. [19]: Sie verwenden die kontrollierte Fluoreszenz *einzelner* organischer Farbstoffmoleküle bei tiefen Temperaturen (1,8 K); die Eigenschaften des ausgesandte Lichtes entsprechen tatsächlich denen einer Einphotonenquelle. Ein ähnliches Experiment mit einem einzelnen organischen Farbstoffmolekül als Lichtquelle *bei Zimmertemperatur* wurde von B. Lounis und W. E. Moerner [85] durchgeführt. Allerdings wurde hierbei kein reiner Einphotonenzustand erreicht, sondern in einem Fünftel der Fälle wurden zwei Photonen ausgesandt.
Bei Zimmertemperatur besteht allerdings immer noch das große Problem, daß die verwendeten einzelnen Farbstoffmoleküle aufgrund der starken Belastungen durch das optische Pumpen mit einem Laser nur eine sehr begrenzte Haltbarkeit aufweisen.
Dieses Problem wird bei zwei neuen, festkörperphysikalischen Ansätzen zur Realisierung einer Einphotonenquelle durch Verwendung stabilerer Systeme als Emitter vermieden. P. Michler et al. [93] realisieren solch eine Quelle, indem sie die Photonenerzeugung durch Verwendung des anharmonischen Verhaltens eines Exzitonen-Mehrfachüberganges eines *einzelnen* Quantenpunktes regeln. Die Anregung der Quantenpunkte erfolgt mit 250 fs langen Pulsen eines Titan-Saphir-Lasers. Es läßt sich durch Koinzidenzmessungen zeigen, daß bei einer Emission tatsächlich niemals mehr als ein Photon abgestrahlt wird. Überdies erfolgt die Emission der einzelnen Photonen innerhalb weniger Nanosekunden, in enger zeitlicher Korrelation mit dem Pumplaser und mit hoher Quanteneffizienz ($\eta \approx 1$), so daß die Quelle dem Ideal einer gepulsten Einphotonenquelle sehr nahe kommt. Leider hat diese Erzeugungsmethode (noch) zwei Nachteile, die sich aber in Zukunft zumindest teilweise beseitigen lassen werden: Die einzelnen Photonen werden nicht in

[10] Für eine Übersicht über diesen und ähnliche Ansätze, s. [131].

eine Richtung abgestrahlt und die Proben mit den Quantenpunkten müssen auf unter 4 K gekühlt werden.
Der zweite festkörperphysikalische Ansatz, um eine Einphotonenquelle zu realisieren, verwendet die Fluoreszenz von NV-Zentren, d.h. einem System aus einer Fehlstelle in der direkten Nachbarschaft eines *einzelnen* Atoms einer Stickstoff-Verunreinigung und dem Stickstoffatom, in Diamant bei Zimmertemperatur. Die hohe Quanteneffizienz der Abstrahlung, die kurze Zerfallszeit des angeregten Zustandes und die Stabilität des Systems gestatten hierbei die Realisierung einer Einphotonenquelle. Angeregt wird das System mit Hilfe eines frequenzverdoppelten Neodym-YAG-Laser, die Fluoreszenz erfolgt im roten bis nah-infraroten Teil des Spektrums mit einer Bandbreite von ca. 120 nm. Das emittierte Licht zeigt deutliche Antibunching-Eigenschaften, ein vorhandener Fluoreszenz-Untergrund des Diamant bedingt allerdings, daß kein perfektes Antibunching gemessen wird. Der große Vorteil dieser Einphotonenquelle besteht darin, daß sie einen kompakten Aufbau, eine gerichtete Abstrahlung[11] und einen Betrieb bei Zimmertemperatur ohne Degradationserscheinungen erlaubt.
Bei den in der vorliegenden Arbeit beschriebenen Experimenten werden die Photonen-Anzahlzustände nicht direkt realisiert, sondern sie werden *indirekt* erzeugt. Dies ist für einen Einphotonen-Zustand gut möglich [55], man geht dabei folgendermaßen vor: Man nehme ein Paar zeitlich und räumlich korrelierter Photonen und detektiere eines der Photonen mit einem (Trigger-)Detektor, ohne das andere Photon des Paares zu vernichten. Aufgrund der Korrelation der beiden Photonen kann dann davon ausgegangen werden, daß bei einem Ansprechen des (Trigger-)Detektors (im Idealfall) noch ein Photon in dem korrespondierenden Raumbereich innerhalb des betrachteten Zeitintervalles vorhanden sein muß.
Es bleiben bei dieser Vorgehensweise noch zwei Probleme zu lösen:

- Man muß ein Verfahren finden, um einzelne Photonenpaare zu erzeugen.
- Die beiden Photonen eines Paares müssen ohne größere Verluste räumlich getrennt werden können, damit die Detektion des einen Photons das andere nicht ebenfalls vernichtet.

Diese beiden Probleme lassen sich elegant lösen, indem man die sogenannte *parametrische Fluoreszenz* zur Erzeugung von Photonenpaaren nutzt.

3.1.2.1 Parametrische Fluoreszenz

Unter *parametrischer Fluoreszenz* [20, 123, 132] versteht man die nichtlineare Wechselwirkung eines einfallenden Pumplichtfeldes mit einem optisch anisotropen Kristall, bei der Paare von zeitlich[12], energetisch und räumlich korrelierten Photonen ausgesandt werden. Es gibt eine ganze Reihe von optisch anisotropen Kristallen, die zur Erzeugung parametrischer Fluoreszenz benutzt werden, am gebräuchlichsten sind hierbei KDP (Kaliumdihydrogenphosphat KH_2PO_4), ADP (Ammoniumdihydrogenphosphat $NH_4H_2PO_4$,

[11] Das abgestrahlte Licht wurde direkt in eine Glasfaser eingekoppelt.
[12] Die beiden Photonen werden gleichzeitig innerhalb eines Zeitintervalles von 50–100 Femtosekunden abgestrahlt, s. [56]; mit den verwendeten Photodetektoren, die eine Zeitauflösung von 500 ps haben und zudem noch eine Totzeit von typ. 50 ns besitzen, läßt sich das Photonenpaar daher nicht mit einem einzelnen Detektor zeitlich in die beiden einzelnen Photonen auflösen.

Lithiumniobat ($LiNbO_3$), Lithiumjodat ($LiIO_3$), Kaliumniobat (KN abgekürzt, $KNbO_3$) und BBO (β–Bariumborat BaB_2O_4). Die parametrische Fluoreszenz ist eine experimentell gut realisierbare Methode, Photonenpaare zu erzeugen, die außerdem noch den entscheidenden Vorteil hat, die Photonen eines Paares in zwei relativ kleine Raumwinkelbereiche[13] abzustrahlen, so daß sie sich mit herkömmlichen Detektoren effizient detektieren lassen.
Die parametrische Fluoreszenz läßt folgendermaßen anschaulich beschreiben: Ein einfallendes Pump-Photon wird aufgrund der Nichtlinearität des Mediums spontan in ein Photonenpaar umgewandelt, das aus einem *„Signal"-Photon* der Frequenz ω_s und einem *„Idler"-Photon* der Frequenz ω_i besteht. Da hierbei sowohl Energie als auch Impuls erhalten bleiben müssen, die Energie eines Photons durch $\hbar\omega$ und der Impuls durch $\hbar\vec{k}$ gegeben ist, hat man als Beziehung zwischen dem Pump-Photon und den Photonen des erzeugten Paares:

$$\omega_s + \omega_i = \omega_p \quad \text{und} \quad \vec{k}_s + \vec{k}_i = \vec{k}_p. \tag{3.1}$$

Die Wahrscheinlichkeit, daß ein Pumpphoton in ein Photonenpaar umgewandelt wird, ist abhängig von der Stärke der Nichtlinearität, typischerweise[14] aber leider sehr klein: $P_{konv.} \approx 10^{-11}$. Somit ist die Wahrscheinlichkeit, zwei Photonenpaare gleichzeitig zu erhalten, entsprechend gering; man kann also von einzelnen Photonenpaaren ausgehen. Dies bezieht sich insbesondere auf den Zeitraum zwischen Emission des Paares und Detektion der Photonen des Paares; ist es doch gerade für einen quantenoptischen Zufallsgenerator sehr wünschenswert, immer nur einzelne Photonen für den Zufallsprozeß zu verwenden, da sich so eine *direkte* Beeinflussung[15] aufeinanderfolgender Zufallsereignisse grundsätzlich ausschließen läßt. Wurde nämlich das nachfolgende Photonenpaar noch gar nicht von der Quelle emittiert, kann es natürlich nicht direkt von dem vorhergehenden Paar bzw. von einem der Photonen des Paares beeinflußt werden. Allerdings gehorcht die Emission der Photonenpaare einer Poisson-Statistik [74] und somit existiert eine nicht verschwindende Wahrscheinlichkeit dafür, daß ein Photonenpaar emittiert wird, während sich das vorangehende noch „im Flug" durch den experimentellen Aufbau befindet; allerdings ist die Wahrscheinlichkeit hierfür bei den gegebenen experimentellen Randbedingungen recht klein, s. Diskussion, Abschnitt 6.2.3.1.
Mit der parametrischen Fluoreszenz läßt sich somit zwar keine Lichtquelle realisieren, die zu *bestimmten* Zeiten genau ein Photonenpaar emittiert[16] und somit auch keine entsprechende Einphotonenquelle, aber durch das Signal des Triggerdetektors weiß man im nachhinein, wann das Signalphoton emittiert wurde. Für die meisten Experimente ist dies völlig ausreichend, so auch für die mit einzelnen Photonen arbeitenden quantenoptischen Zufallsgeneratoren.

[13] Für den sogenannten *kollinearen Fall*, s.u., sind die beiden Raumwinkelbereiche deckungsgleich.
[14] In Abschnitt 5.1.4 wird diese Konversionseffizienz für die durchgeführten Experimente berechnet.
[15] Eine *indirekte* Beeinflussung ist sehr wohl möglich, z. B. über Totzeiteffekte der Detektoren.
[16] Verwendet man einen Kurzpuls-Laser zum Pumpen des nichtlinearen Kristalls, so kann man aufgrund der instantanen Natur des parametrischen Prozesses sicher sein, daß ein auslaufendes Photonenpaar während des Zeitintervalles der Wechselwirkung zwischen Lichtpuls und Kristall entstanden sein muß. Ob allerdings überhaupt ein Photonenpaar bei der Wechselwirkung entsteht, ist natürlich immer noch ein stochastischer Vorgang: Nur bei hoher Pulsintensität kann man hinreichend sicher sein, daß ein Paar entsteht. Die bei der parametrischen Fluoreszenz entstehenden Paare sind aber Poisson-verteilt, so daß bei solch hoher Intensität des Pumplichtes häufig mehrere Photonenpaare pro Pump-Puls entstehen würden.

Bei der experimentellen Realisierung der parametrischen Fluoreszenz muß man allerdings noch einen wichtigen, weiteren Aspekt berücksichtigen. Soll möglichst intensives Fluoreszenzlicht erzeugt werden, d.h. möglichst viele Photonenpaare, muß man erreichen, daß an verschiedenen Stellen eines Kristalls erzeugtes Fluoreszenzlicht sich *phasenrichtig* überlagert, da sonst die Intensität des Fluoreszenzlichtes ab- und nicht wie gewünscht zunehmen könnte. Dies wird durch die Dispersion des Mediums erschwert, denn für eine gegebene Frequenzanpassung $\omega_p = \omega_s + \omega_i$ ist im allgemeinen der Betrag des Wellenvektors der Pumpwelle größer als die Summe der Signal- und Idler-Wellenvektoren: $|\vec{k}_p| > |\vec{k}_s + \vec{k}_i|$, so daß die Phasenlage zwischen Pump- und dem Fluoreszenzlicht nicht konstant bleibt.
Methoden um den Wellenvektor der Pumpwelle so abzustimmen, daß es zu einer phasenrichtigen Überlagerung des entstehenden Fluoreszenzlichtes kommt, nennt man *Phasenanpassungen*. Meist nutzt man bei ihnen die doppelbrechenden Eigenschaften des optisch nichtlinearen Kristalls aus, d.h. den Umstand, daß der Brechungsindex (bzw. die Phasengeschwindigkeit des Lichtes im Kristall) für verschiedene Polarisations- und Ausbreitungsrichtungen im allgemeinen unterschiedlich ist.
Im Falle eines optisch negativen, uniaxialen Kristalls, wie z.B. Lithiumjodat, schickt man die Pumpwelle als außerordentliche Welle durch den Kristall, und Signal-und Idler-Welle laufen entweder beide als ordentliche Welle (*Typ-I-Phasenanpassung*, s. Abb. 3.4) oder eine läuft als ordentliche und eine als außerordentliche Welle (*Typ-II-Phasenanpassung*) durch den Kristall[17].

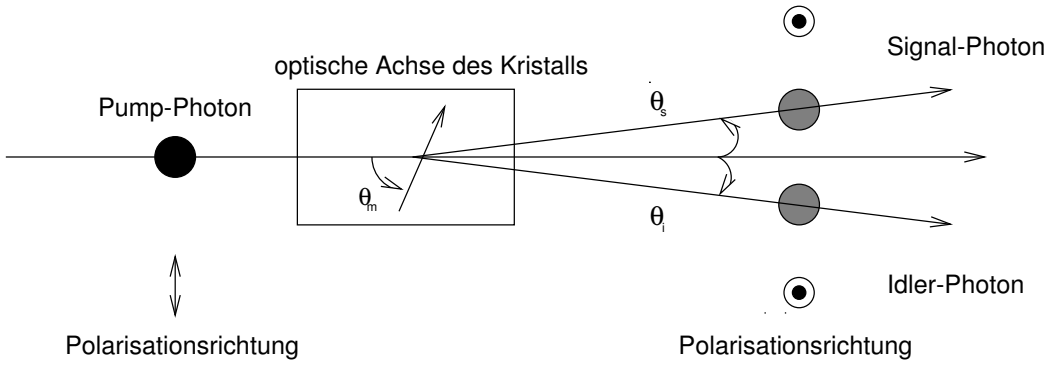


Abbildung 3.4: Typ-I-Phasenanpassung, nichtkollinearer Fall

Welcher der beiden Fälle dabei realisiert wird, hängt vom Winkel zwischen der optischen Achse und der Ausbreitungsrichtung der Pumpwelle ab. Durch ihn wird auch bestimmt, ob die Photonenpaare kollinear oder nichtkollinear, zum Pumpstrahl emittiert werden. Die Impulserhaltung fordert im nichtkollinearen Fall, daß die Signal- bzw. Idler-Photonen in Winkeln Θ_s und Θ_i relativ zur Strahlachse des Pumplasers emittiert[18] werden:

[17] Zur besseren Übersicht ist die Brechung der Signal- und Idler-Strahlen am Übergang Kristall-Luft in den Abbildungen nicht dargestellt.

[18] Der Winkel zwischen Pumpstrahl und der Richtung der Photonen eines Paares ist in Abb. 3.4 zur Verdeutlichung stark übertrieben dargestellt, im Experiment werden die Winkel so klein wie möglich gewählt und liegen unter 1°.

$$\frac{n_p}{\lambda_p} = \frac{n_s \cos\Theta_s}{\lambda_s} + \frac{n_i \cos\Theta_i}{\lambda_i} \quad \text{und}$$

$$\frac{n_s \sin\Theta_s}{\lambda_s} = \frac{n_i \sin\Theta_i}{\lambda_i},$$

wobei n_p, n_s und n_i die Brechungsindices des Kristalls für die Wellenlängen λ_p, λ_s und λ_i sind. Die Winkel Θ_s und Θ_i können hierbei auf einem Kegelmantel um die Strahlachse des Pumplasers liegen. Im kollinearen Fall gilt $\Theta_s = \Theta_i = 0$. Neben der eben beschriebenen Winkel-Phasenanpassung bieten Kristalle. wie z.B. Kaliumniobat, bei denen die Temperaturabhängigkeit der Brechungsindices entlang der Kristallhauptachsen unterschiedlich ist, noch zusätzlich die Möglichkeit einer Phasenanpassung mit Hilfe der Temperatur. Der Vorteil dieser Methode liegt darin, daß es nicht nötig ist, den Kristall wie bei der Winkel-Phasenanpassung mit Hilfe einer Winkelverstelleinheit zu drehen und sich hiermit die Justage des Gesamtaufbaus vereinfacht. Allerdings erkauft man sich dies mit leichten Schwankungen in der Wellenlängenverteilung der Signal- und Idler-Strahlen, da sich bereits kleine Temperaturschwankungen von 0,1° C auf die Phasenanpassung auswirken, s. Abschnitt 5.1.2.3.

3.1.3 Von-Neumann-Regularisierung

Bei nahezu allen[19] physikalischen Zufallsgeneratoren taucht ein grundsätzliches Problem auf: Wie nahe das Verhältnis von Nullen und Einsen auch beim idealen Teilungsverhältnis von 50:50 liegen mag, wenn die zufällige Binärsequenz nur lang genug ist, wird man auch kleinste Abweichungen bemerken, s. Abschnitt 3.2.3.
Dieses Problem läßt sich aber direkt bei der Datenaufnahme oder auch im nachhinein beheben, indem man ein einfaches mathematisches Verfahren anwendet, das auf John von Neumann [96] zurückgeht und als einzige Voraussetzung die statistische Unabhängigkeit aufeinanderfolgender Ereignisse verlangt.
Bei der Von-Neuman-Regularisierung werden jeweils zwei aufeinanderfolgende Eingangsbits überlappungsfrei zusammengefaßt und aus ihnen ein Ausgangsbit nach folgender Vorschrift generiert:

Bit 1	Bit 2	Ausgangs-Bit
1	1	–
1	0	1
0	1	0
0	0	–

Dieses Verfahren bringt zwar den Nachteil einer mindestens 75-prozentigen Reduktion der maximal erreichbaren Bitrate mit sich, garantiert dafür aber auch, daß keine systematischen Fehler bei der 50:50 Verteilung der relativen Häufigkeit der Nullen und Einsen auftreten.

[19] Bei Zufallsgeneratoren, die radioaktive Zerfälle in der in Abschnitt 2.2.3.1 erläuterten Weise zur Zufallsgenerierung verwenden, tritt dieses Problem nicht auf.

Neben der Von-Neumann-Regularisierung, die durch ihre Einfachheit besticht, gibt es auch noch andere, ausgefeiltere Methoden, um das Verhältnis von Nullen und Einsen im Mittel auf den Idealwert von 50:50 zu bringen. Diese Methoden sind komplizierter, dafür ermöglichen einige von ihnen aber selbst dann eine Regularisierung, wenn statistische Abhängigkeiten zwischen den Zufallsbits vorliegen, während andere effizienter als die Von-Neumann-Regularisierung sind und eine höhere Netto-Bitrate erlauben. Eine ausführliche Darstellung der Von-Neumann-Regularisierung und anderer, praktisch anwendbarer Regularisierungsmethoden findet sich im Anhang, Abschnitt G.

3.2 Statistische Tests der Zufallszahlen

Aufgrund der großen Bedeutung, die Zufallszahlen für Simulationen, Monte-Carlo-Methoden und in der Kryptographie haben, wurden im Laufe der Zeit eine große Anzahl von statistischen Tests entwickelt (s. z.B. [25, 28, 51, 68, 77, 78, 90]), um insbesondere algorithmisch arbeitende Zufallszahlengeneratoren auf statistische Defekte zu testen.
Im Prinzip lassen sich sogar beliebig viele weitere Tests ersinnen, und in der Tat werden auch immer wieder neue Tests vorgeschlagen. Dies hat allerdings auch einen triftigen Grund: Selbst bei einem Pseudozufallszahlengenerator, der bereits eine Reihe von Tests bestanden hat, kann man sich nie sicher sein, daß er einen neu entworfenen Test ebenfalls bestehen wird. Daher hat man keine andere Wahl, als vor dem Einsatz eines Pseudozufallszahlengenerators, einen Test durchzuführen, der möglichst diejenigen statistischen Eigenschaften der Zufallszahlen testet, die bei der jeweiligen Anwendung benötigt werden und bei denen eine Abweichung vom statistischen Idealverhalten zu Fehlern führen könnte.
So viele Tests für Pseudozufallszahlengeneratoren es auch geben mag – es stellt sich dennoch die Frage, welche dieser Tests man am besten zur Untersuchung *physikalischer Zufallsgeneratoren* verwenden sollte. So wurden z.B. einige der Standardtests [78] für algorithmische Zufallsgeneratoren speziell entwickelt, um statistische Defekte wie langreichweitige Korrelationen aufzudecken, die bei physikalischen Zufallsgeneratoren i.a. ohnehin keine Rolle spielen, s. Abschnitt 3.2.6.
Natürlich sollte ein physikalischer Zufallsgenerator dennoch beliebige statistische Tests auf Zufälligkeit – also auch die oben erwähnten – bestehen, allerdings zeigt sich hierbei ein spezifisches Problem der meisten physikalischen Zufallsgeneratoren:
Die „Rohdaten“ eines physikalischen Zufallsgenerators, also die Daten, welche keiner Regularisierung unterzogen wurden, besitzen eine relative Häufigkeitsverteilung der Nullen und Einsen, die i. a. *nicht* dem theoretischen Idealfall einer Binomialverteilung mit gleicher Wahrscheinlichkeit für das Auftreten einer Eins oder einer Null entspricht. Läßt sich die sogenannte *Nullhypothese* (s.u.), die das rein zufallsgesteuerte statistische Verhalten beschreibt, *nur* für eine ideale Gleichverteilung von Nullen und Einsen analytisch[20] berechnen, ist ein Test der bekanntermaßen tendenzbehafteten Rohdaten wenig sinnvoll, da nur die ohnehin bereits bekannten Abweichungen von der idealen Gleichverteilung

[20] Numerische Berechnungen, die sich in solchen Fällen oft auf Simulationen stützen, bei denen wiederum häufig Pseudozufallszahlengeneratoren benutzt werden, sind natürlich etwas unbefriedigend. Derart erhaltene Ergebnisse lassen sich zwar durch Cross-Validierung mit mehreren Pseudozufallsgeneratoren absichern, aber ein unangenehmer Beigeschmack bleibt.

„aufgedeckt“ würden. Eventuell vorhandene, tieferliegende statistische Defekte würden vollständig von diesem Umstand überdeckt.
Um dennoch weitergehende statistische Tests durchführen zu können, bieten sich zwei mögliche Vorgehensweisen an:

1. Man betrachtet nicht die Rohdaten, sondern testet nur *regularisierte Daten*, bei denen die Häufigkeitsverteilung der Nullen und Einsen ideal sein sollte.

2. Man benutzt nur Tests, deren Statistik als Parameter die Wahrscheinlichkeit für eine Eins $P(Bit = 1) = p$ erlauben. Hierbei wird der Wert p aus der für den Test jeweils verwendeten Stichprobe geschätzt. Allerdings erlauben nicht alle Tests dieses Vorgehen, so daß interessante Tests, wie z.B. der *binäre Ableitungstest* [28], nicht durchgeführt werden können.

In der vorliegenden Arbeit wird bei den meisten Tests der zweite Weg eingeschlagen. Lediglich zum Vergleich und um die Auswirkungen von Regularisierungsverfahren darzulegen, werden vereinzelt Tests an regularisierten Daten durchgeführt.

3.2.1 Methodisches Vorgehen bei statistischen Tests

Im folgenden wird die Vorgehensweise bei den statistischen Tests der Zufallsgeneratoren beschrieben, wobei in diesem Abschnitt allerdings nur auf die grundlegende Methodik von *einstufigen* und *zweistufigen statistischen Tests* eingegangen wird. Die Details der von Test zu Test verschiedenen Statistiken werden in nachfolgenden Unterkapiteln besprochen. Lediglich der Kolmogorov-Smirnov-Test, der einen integralen Bestandteil der in dieser Arbeit verwendeten zweistufigen, statistischen Tests darstellt, wird bereits hier erläutert.
Statistische Tests werden allgemein verwendet, um zu entscheiden, mit welcher Wahrscheinlichkeit ein experimentell gewonnener Datensatz mit einer i. a. vor dem Experiment aufgestellten Hypothese vereinbar ist. Formal wird die Hypothese hierbei durch eine theoretische, meist parametrisierte Wahrscheinlichkeitsverteilung einer Zufallsvariablen dargestellt, wobei diese Zufallvariable mit Hilfe einer Funktion, der sogenannten *Statistik*, aus den Daten berechnet wird.
Es wird allerdings nicht die Wahrscheinlichkeit für die Gültigkeit der Hypothese ermittelt, sondern umgekehrt die Wahrscheinlichkeit für das Ablehnen der sogenannten *Nullhypothese*, d.h. der Hypothese, daß bloße Zufallsschwankungen für die Ergebnisse verantwortlich sind, berechnet. Hierbei sollte die Wahrscheinlichkeit, daß man die Nullhypothese zu unrecht ablehnt[21], d.h. Effekte erkennt, wo eigentlich nur der Zufall regiert, hinreichend klein sein. Typische Werte für diese *Signifikanzniveau* genannte Wahrscheinlichkeit sind z.B. $\alpha = 0{,}02$, $\alpha = 0{,}05$ oder bisweilen auch $\alpha = 0{,}1$. Entsprechend ist $1-\alpha$ das *Sicherheitsniveau*, auf dem man „ausschließen kann“, daß Zufallsschwankungen lediglich einen vermeintlichen Effekt vorgetäuscht haben, d.h. man die Hypothese als gültig akzeptiert, obwohl sie es nicht ist. Wählt man α relativ groß und somit das Sicherheitsniveau entsprechend klein, besteht die Gefahr, daß eine Hypothese von den Daten bestätigt zu werden scheint, obwohl dies eigentlich nicht der Fall ist. Wählt man andererseits α sehr klein, so besteht die Gefahr, daß die Nullhypothese nicht abgelehnt wird, obwohl

[21] Man nennt dies auch *Fehler erster Art*.

sie eigentlich verworfen werden sollte[22], da tatsächlich ein Effekt vorliegt, der sich allerdings nicht so deutlich in den Daten zeigt, daß er auf einem so hohen Sicherheitsniveau bestätigt werden könnte.
Die Wahl des Sicherheitsniveaus bzw. von α hängt auch von der Stichprobengröße ab, da Zufallsschwankungen mit zunehmender Stichprobengröße geringer ausfallen, so daß sich höhere Sicherheitsniveaus erreichen lassen.
Hat man ein Sicherheitsniveau gewählt, so muß noch berücksichtigt werden, ob es sich um einen einseitigen oder einen zweiseitigen Test handelt, d.h. ob die Region der für die Nullhypothese akzeptablen Statistik-Werte nach einer oder nach zwei Seiten abgegrenzt ist. Im ersten Fall läßt sich aus der Wahrscheinlichkeitsverteilung der Statistik direkt der Wert der Statistik ablesen, ab dem die Nullhypothese verworfen wird, im zweiten Fall werden zwei Werte der Statistik so gewählt, daß die Wahrscheinlichkeit für die Werte der Statistik außerhalb des von ihnen begrenzten Akzeptanz-Gebietes gerade gleich α ist. Bei Verteilungen mit symmetrischer Dichtefunktion, wie z.B. der Normalverteilung, wird man die beiden Werte so wählen, daß die Wahrscheinlichkeit für die Werte außerhalb des Intervalls auf jeder Seite jeweils $\alpha/2$ beträgt.
Will man Zufallgeneratoren auf statistische Defekte testen, so hat man es mit einer etwas ungewöhnlichen Situation zu tun: Das Idealverhalten des Zufallgenerators sollte gerade das zufällige Verhalten sein, d. h. die Gültigkeit der Nullhypothese ist in diesem Fall das, was man erwartet. *Jeder „Effekt“* im obigen Sinne ist ein *Defekt*, wobei allerdings i.a. keine spezifische Hypothese betrachtet wird, sondern alle möglichen anderen Hypothesen, d.h. Defekte werden zusammengefaßt zu einer Alternative.

3.2.1.1 Einstufige Tests

Bei einstufigen statistischen Tests geht man folgendermaßen vor:

1. Man wähle eine hinreichend große[23] Stichprobe aufeinanderfolgender[24], Bits aus dem Zufallsbitstrom.

2. Man berechne aus den Daten $(D_1, \ldots, D_n)$ der Stichprobe mit Hilfe einer für den Test spezifischen Funktion, der *Statistik* $X(D_1, \ldots, D_n)$, einen einzelnen Wert x.

Der auf diese Weise erhaltene Wert x wird nun danach beurteilt, wie wahrscheinlich es ist, daß ein idealer Zufallsgenerator ihn erzeugt haben könnte. Dies kann auf zwei Arten geschehen:

[22] Man bezeichnet dies als *Fehler zweiter Art*.

[23] Wie groß „hinreichend groß“ ist, hängt von den Details der jeweiligen Test-Statistik ab. Wird z.B. eine Approximation für die tatsächliche, theoretische Wahrscheinlichkeitsverteilung verwendet, so muß die Stichprobe natürlich so groß sein, daß die Bedingungen für eine gute Approximation erfüllt sind, s.a. Abschnitt D.

[24] Bei physikalischen Zufallsgeneratoren ist dieses Vorgehen angemessen, bei Pseudozufallsgeneratoren kann es aber durchaus geboten sein, auch unzusammenhängende Teilmengen zu testen. Aufgrund der *Berechnung* der jeweils folgenden Zahl aus ihren Vorgängern stellt die maximale Länge der Zahlen, die i.a. der Registerlänge des verwendeten Prozessors entspricht, eine natürliche Einteilung des Bitstroms dar. Es ist daher auch durchaus möglich, daß die niederwertigen Bits der Zahlzahlen für sich betrachtet andere statistische Eigenschaften haben als die höherwertigen.

1. *Es wird ein Signifikanzniveau α von vornherein vorgeben*, das beim Test von Zufallsgeneratoren festlegt, wie hoch die Wahrscheinlichkeit ist, daß auch ein idealer Zufallsgenerator den Test aufgrund starker Zufallsschwankungen nicht besteht.

   Anschließend berechnet man durch Inversion der Wahrscheinlichkeitsverteilung der Nullhypothese den Wert oder die Werte der Statistik, die diesem Signifikanzniveau entsprechen.

   Liegt der aus den empirischen Daten errechnete Wert x der Statistik innerhalb des spezifizierten Bereiches, wird der Test als auf dem Sicherheitsniveau $(1 - \alpha)$ bestanden angesehen; ansonsten gilt der Generator als defektbehaftet[25].

2. *Es wird keine feste Wahrscheinlichkeit α vorgegeben*, sondern erst anhand des empirisch erhaltenen Wertes x der Statistik und der Wahrscheinlichkeitsverteilung wird berechnet, wie hoch die Wahrscheinlichkeit α dafür wäre, daß die Stichprobe eines idealen Zufallsgenerators außerhalb eines Bereiches läge, dessen (eine) Grenze x ist. Hierbei wird ähnlich wie im vorhergehenden Fall bei Schritt 2 vorgegangen; natürlich muß auch hierbei beachtet werden, ob es sich um einen einseitigen oder einen zweiseitigen Test handelt. Ist der auf diese Weise erhaltene Wert α relativ groß, so gilt der Generator als nicht defektbehaftet.

Beide Vorgehensweisen zeigen ein Problem auf, das sich bei einstufigen Tests nicht vermeiden läßt: Führt man einen einstufigen Test lediglich *einmal* durch, so läßt sich i.a. die Güte eines Zufallsgenerators nicht hinreichend genau beurteilen. Nur bei sehr großen Abweichungen vom Idealverhalten kann man einen Generator guten Gewissens als „nicht zufällig“ ablehnen. Dieses Problem läßt sich etwas abmildern, indem man mehrere einstufige Tests durchführt; günstiger ist es allerdings, gleich zweistufige Tests durchzuführen, die präzisere Aussagen hinsichtlich der Güte des Generators ermöglichen.

3.2.1.2 Zweistufige Tests

Die erste Stufe eines zweistufigen Tests unterscheidet sich lediglich in einem Punkt von einem einstufigen Test: Es wird nicht nur einmal eine Statistik berechnet, sondern der Test wird mehrere Male auf unterschiedlichen Teilsequenzen der Daten durchgeführt, so daß man entsprechend viele Realisierungen der Statistik erhält. Erhält man im Falle des einstufigen Tests also lediglich einen einzigen Wert für die Statistik, so läßt sich bei zweistufigen Tests aus der Menge aller Testergebnisse der ersten Stufe eine empirische Häufigkeitsverteilung für die Statistik erstellen. Diese empirische Häufigkeitsverteilung ist es, die man in einer zweiten Teststufe mit der theoretischen Wahrscheinlichkeitsverteilung vergleicht. Auf diese Weise erhaltene Ergebnisse sind wesentlich aussagekräftiger als ein einzelner Testwert, bei dem es sich durchaus um einen Ausreißer oder – schlimmer noch – um einen zufälligerweise guten Wert eines ansonsten schlechten Generators handeln könnte.
Starke Abweichungen der empirischen von der theoretischen Verteilung erkennt man natürlich leicht durch einen visuellen Vergleich der Graphen der beiden Verteilungen.

[25] Nach dem oben Gesagten ist natürlich klar, daß auch ein idealer Generator durch den Test fallen kann, daher darf α bei einstufigen Tests nicht zu groß gewählt werden, für typische Werte vgl. z.B. FIPS-Tests in Abschnitt 3.2.2.

Ein objektiveres Maß für die Größe der Abweichung läßt sich mit Hilfe der Kolmogorov-Smirnov-Statistik berechnen.

3.2.1.3 Kolmogorov-Smirnov-Test

Will man eine empirische mit einer theoretischen Verteilung auch quantitativ vergleichen, braucht man ein Maß für den Abstand zwischen den beiden Verteilungen. Der Kolmogorov-Smirnov-Test verwendet hierfür die maximalen Abweichungen der beiden Verteilungen in positiver wie negativer Richtung. Bezeichnet man die empirische Verteilung mit $F_n(x)$ und die theoretische mit $F(x)$, so sind diese Abstände[26] gegeben durch:

$$K_n^+ = \sqrt{n} \max_{-\infty<x<\infty} (F_n(x) - F(x))$$
$$K_n^- = \sqrt{n} \max_{-\infty<x<\infty} (F(x) - F_n(x)) .$$

Ist die theoretische Verteilung analytisch gegeben, lassen sich K_n^+ und K_n^- aus den empirischen Daten leicht ermitteln. Will man einschätzen, wie wahrscheinlich es ist, daß es sich bei der empirischen Verteilung um eine Realisierung der theoretischen Verteilung handelt, muß man wiederum ein Signifikanzniveau vorgeben bzw. berechnen. Das Signifikanzniveau ist hierbei die Wahrscheinlichkeit, daß der Abstand K_n^+ (bzw. der Abstand K_n^-) kleiner oder gleich einem Schwellwert s ist, für den Fall, daß es sich bei der empirischen Häufigkeitsverteilung tatsächlich um eine Realisierung der theoretischen Wahrscheinlichkeitsverteilung handelt.
Für den Zusammenhang zwischen dem Schwellwert s und der Signifikanz gilt folgende Approximation[27] [68]:

$$\lim_{n\to\infty} P(K_n^+ \leq s) = 1 - e^{-2s^2} \left(1 - \frac{2}{3} \cdot \frac{s}{\sqrt{n}} + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

Für die Werte von K_n^- gilt die gleiche Approximation. Es empfiehlt sich hierbei wieder, direkt den empirisch erhaltenen Wert für K_n^+ als Schwellwert s zu wählen und das zugehörige Signifikanzniveau mit Hilfe der oben stehenden Formel zu berechnen.
Typischerweise wird eine empirische Verteilung an manchen Stellen größer und an anderen wiederum kleiner als die theoretische Verteilung sein. Liegt die empirische Verteilung *nahe* bei der theoretischen, erhält man entsprechend *kleine Werte* für K_n^+ bzw. K_n^-, und die entsprechenden Wahrscheinlichkeiten sind dementsprechend ebenfalls recht gering. Zwar kann dies grundsätzlich auch vorkommen, doch sind solche sehr niedrigen Werte meist ein Zeichen dafür, daß der Zufallsgenerator in dem Sinne „zu zufällig ist“, als daß seine Fluktuationen zu gering sind. Ein solches Verhalten ist allerdings höchstens von Pseudozufallsgeneratoren zu erwarten.

[26] Beide Abstände getrennt zu betrachten, ist die „Smirnov-Variante“ des Tests. Beim Kolmogorov-Smirnov-Test im strengen Sinne wird nur der größere der beiden Abstände betrachtet. Die hier verwendete Variante weist den Vorteil auf, daß sich für die Signifikanz der maximalen Abweichungen sowohl ein endlicher analytischer Ausdruck als auch eine sehr gute Approximation angeben läßt.

[27] Die Approximation ist bei mehr als 100 Tests der ersten Stufe sehr gut erfüllt, erst in der vierten Nachkommastelle zeigen sich Abweichungen vom exakten Ergebnis; mehr als zwei Nachkommastellen sind für eine Interpretation der Ergebnisse ohnehin unnötig.

Sind umgekehrt die Werte für K_n^+ und K_n^- recht groß – ein Fall, der bei physikalischen Generatoren eher anzutreffen ist – dann sind die Wahrscheinlichkeiten dafür, daß die Werte für K_n^+ und K_n^- idealer Zufallsgeneratoren kleiner als die empirisch erhaltenen wären, sehr nah bei Eins, da es fast sicher ist, daß die empirische Verteilung näher an der theoretischen liegen sollte, als dies der Fall ist.
Häufig kommt es auch vor, daß einer der beiden Werte groß, der andere aber klein ist: Auch dies ist ein Zeichen für einen statistischen Defekt des Generators.
„Gute Werte“ der Signifikanz werden also weder nahe bei Eins noch nahe bei Null liegen, sondern bei 1/2. Man sieht also, daß die Signifikanz im Falle des Kolmogorov-Smirnov-Tests eine etwas andere Bedeutung hat als bei den oben erwähnten einfachen Tests.
Grundsätzlich besteht natürlich die Möglichkeit, in der statistischen Auswertung noch eine Stufe weiter zu gehen, d.h. man führt eine möglichst große Zahl von zweistufigen Tests durch, berechnet jeweils die Werte für K_n^+ und K_n^- und ermittelt auf diese Weise empirische Verteilungen für K_n^+ und K_n^-, die sich nun wiederum mit ihrer theoretischen Verteilung vergleichen lassen. Allerdings muß man bei diesem Vorgehen mit zwei Problemen rechnen:

- Die benötigten Datenmengen nehmen stark zu, da man sehr viele Tests der ersten Stufe machen muß.
- Häufig wird für die theoretische Verteilung, zu der man in der zweiten Stufe den Abstand berechnet, eine Approximation verwendet[28], so daß die Gefahr besteht, in einer dritten Teststufe von Approximationsfehlern in die Irre geleitet zu werden.

Aus diesen beiden Gründen wird bis auf die dreistufige Auswertung der Autokorrelationskoeffizienten, s. Abschnitt B.10.3, darauf verzichtet, quantitative dreistufige statistische Tests durchzuführen. Wie gut die Approximationen der Verteilungsfunktionen sind, wird im Anhang, Abschnitt D ausgeführt.

3.2.2 Tests nach FIPS 140-1

FIPS 140-1 steht für *Federal Information Processing Standards Publication 140-1*, *Security Requirements for Cryptographic Modules*, herausgegeben vom U.S. Department of Commerce und dem National Institute of Standards and Technology[29] (NIST). Innerhalb des Dokumentes, das diesen Standard beschreibt, findet man in Kapitel 4 unter „4.11.1 Power-Up Tests“ einen Unterabschnitt „Statistical random number generator test“. In diesem Abschnitt werden vier Selbsttests dargelegt, die ein Kryptomodul bei jedem Anschalten an einem eingebauten Zufallsgenerator durchführen soll. Als Stichprobe schreibt der FIPS-Standard eine 20.000 Bit-Sequenz vor, an der alle vier Tests durchgeführt werden sollen. Falls ein Zufallsgenerator einen der vier Tests nicht besteht, soll das Kryptomodul in einen Defektmodus gehen. Bei den vier Tests handelt es sich um den Monobit-Test, den Poker-Test, den Lauf-Test und den Test auf lange Läufe.
Die genaue Vorgehensweise bei diesen vier Tests wird im folgenden näher erläutert, vorher sei aber noch angemerkt, daß die Tests nach FIPS 140-1 nur deshalb hier erwähnt

[28] Dies ist z.B. beim Autokorrelationskoeffizententest der Fall.
[29] Der Text läßt sich über das Internet beziehen, man findet ihn auf den Seiten des NIST, s. bei http://www.nist.gov.

und zum Testen der quantenoptischen Zufallsgeneratoren verwendet werden, weil sie einen gewissen offiziellen Charakter[30] haben. Die Tests stellen nur einen Minimalstandard dar, denn aufgrund der kleinen Stichprobengröße sind sie lediglich dazu geeignet, recht große Abweichungen vom Idealverhalten zu erkennen.

3.2.2.1 Der Monobit-Test

1. Man zähle die Anzahl der Einsen (X) innerhalb der 20.000 Bit-Sequenz.

2. Der Generator hat den Test[31] bestanden, wenn gilt:
   $9\,654 < X < 10\,346$, bei einem idealen Generator beträgt die Wahrscheinlichkeit dafür, daß die Zufallsvariable X außerhalb dieses Bereiches liegt, lediglich 10^{-6} .

3.2.2.2 Der Poker-Test

1. Man teile die 20.000 Bit lange Sequenz in 5.000 aufeinanderfolgende, nichtüberlappende 4-Bit-Segmente ein und zähle, wie oft jeder der 16 möglichen, verschiedenen Segment-Werte vorkommt. Man bezeichne mit $f(i)$ die Anzahl, wie oft ein 4 Bit-Wert i mit $0 \leq i \leq 15$ auftritt.

2. Man berechne den Wert der χ^2-Statistik (s. u. Abschnitt 3.2.4), gemäß:

$$\chi^2 = \frac{16}{5000} \cdot \left(\sum_{i=0}^{15} f(i)^2 \right) - 5000$$

3. Der Test gilt als bestanden, falls gilt: $1.03 < \chi^2 < 57.4$, die Wahrscheinlichkeit, daß der χ^2-Wert bei einem idealen Generator außerhalb dieser Grenzen liegt, beträgt 10^{-6}.

3.2.2.3 Der Lauflängen-Test

1. Ein *Lauf* ist die Anzahl von Einsen bzw. Nullen zwischen zwei Nullen bzw. Einsen. Beim Lauf-Test wird die Häufigkeitsverteilung der Läufe der Einsen und Nullen in der Stichprobe ermittelt.

2. Der Test gilt als bestanden, wenn sowohl die Häufigkeiten der Läufe der Einsen als auch der Nullen mit Längen 1 bis 6 innerhalb der in Tabelle 3.1 angegebenen Grenzen liegen. Läufe, die länger als 6 aufeinanderfolgende Nullen bzw. Einsen sind, werden hierbei unter Länge 6 mitgezählt[32].

[30] Weitergehende statistische Tests sind für den sich momentan in der Standardisierung befindenden Nachfolgestandard vorgesehen.

[31] Eine allgemeinere und ausführlichere Erläuterung solcher Tests findet sich in Abschnitt 3.2.3.

[32] Daher wird das Intervall auch größer gewählt.

Länge des Laufes	Intervall
1	2 267 – 2 733
2	1 079 – 1 421
3	502 – 748
4	223 – 402
5	90 – 223
6 und mehr	90 – 223

Tabelle 3.1: Akzeptanz-Intervalle für die Auftrittshäufigkeiten der Läufe beim FIPS-Lauftest

3.2.2.4 Der Test auf lange Läufe

1. Als lange Läufe gelten solche, die länger als 34 aufeinanderfolgende Nullen oder Einsen lang sind.

2. Der Generator besteht den Test, wenn es in der Stichprobe keinen langen Lauf gibt.

3.2.3 Test auf Gleichverteilung der Nullen und Einsen

Bei einem idealen Zufallsgenerator sollten keine statistischen Abhängigkeiten zwischen den generierten Bits bestehen und die Wahrscheinlichkeit p, eine Eins zu erhalten, sollte gleich der Wahrscheinlichkeit q sein, eine Null zu erhalten. Die Anzahl der Einsen[33] x in einer n Bit großen Stichprobe sollte einer Binomialverteilung $b(x; n, p)$ folgen. Betrachtet man große Stichproben, was beim Test von Zufallsgeneratoren meist der Fall ist, so läßt sich die Binomialverteilung sehr gut (s. S. 199) durch die Dichte der Normalverteilung approximieren:

$$b(n,k,p) = \binom{n}{k} p^k (1-p)^{n-k} \approx \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\cdot\left(\frac{x-\overline{x}}{\sigma}\right)^2},$$

hierbei ist $\overline{x} = n \cdot p$ der Erwartungswert und $\sigma = \sqrt{n \cdot p \cdot (1-p)}$ die Standardabweichung der Binomialverteilung. Ein naheliegender, einfacher erster Test besteht darin, eine Stichprobe von n aufeinanderfolgenden Bits aus dem Strom der Zufallsbits zu entnehmen und die Anzahl der Einsen x in ihr zu ermitteln. Die quantitative Auswertung erfolgt hierbei nach der in Abschnitt 3.2.1.1 beschriebenen, allgemeinen Vorgehensweise bei einstufigen Tests, unter der Voraussetzung, daß die Zufallsbits von einem „idealen“ Zufallsgenerator erzeugt wurden.
Wie allerdings bereits in Abschnitt 3.2 erwähnt, sind bei physikalischen Zufallsgeneratoren die Wahrscheinlichkeiten p und $q = (1-p)$ für Einsen und Nullen selten wirklich gleich. Will man dennoch unregularisierte Rohdaten eines Generators testen, so muß man die Wahrscheinlichkeit $P(1) = \hat{p}$ für das Auftreten einer Eins aus den Daten[34]

[33] Das gilt natürlich analog für die Anzahl der Nullen.

[34] $P(1)$ läßt sich entweder aus der Gesamtsequenz oder anhand der jeweiligen einzelnen Stichprobe schätzen. Als Schätzwert $\hat{p}$ für $P(1)$ wird aber meist die relative Häufigkeit der Einsen innerhalb der gesamten Stichprobe verwendet, da sich bei der Untersuchung der empirischen Häufigkeitsverteilung, lokale Schwankungen der Wahrscheinlichkeit $\hat{p}$ dann stärker bemerkbar machen.

schätzen. Für den Test wird entweder die ganze Sequenz oder eine hinreichend, z.B. 10 MByte, große Stichprobe genommen und in Teilsequenzen einheitlicher Länge N unterteilt, für die jeweils die Anzahl der Einsen n_E bestimmt wird. Aus den so erhaltenen Werten wird nach Standardnormierung

$$X_E = \frac{n_E - N \cdot \hat{p}}{\sqrt{N \cdot \hat{p} \cdot (1 - \hat{p})}}$$

eine empirische Häufigkeitsverteilung erstellt und diese mit der theoretischen Verteilung (Normalverteilung) verglichen. Um die Abweichung der empirischen von der theoretischen Verteilung auch quantitativ zu erfassen, wird anschließend in einer zweiten Teststufe ein Kolmogorov-Smirnov-Tests durchgeführt, s. a. Abschnitt 3.2.1.2.

3.2.4 χ^2-Test

χ^2-Tests werden immer dann durchgeführt, wenn es gilt, eine diskrete Zufallsvariable X mit k möglichen Werten, daraufhin zu untersuchen, ob ihre empirische Häufigkeitsverteilung mit einer vermuteten, theoretischen Wahrscheinlichkeitsverteilung $(p_i, i = 1 \ldots k)$ verträglich ist. Hierzu nimmt man eine möglichst große Stichprobe von n unabhängigen Realisierungen[35] der Zufallsvariable X und berechnet als Maß für die „Verträglichkeit“ den sogenannten χ^2-Wert auf folgende Weise:
Für jede Kategorie i der k möglichen Werte wird das Abstandsquadrat zwischen der empirisch erhaltenen Anzahl x_i und dem Erwartungswert $n \cdot p_i$ gebildet und auf den Erwartungswert normiert. Die so erhaltenen Werte aller Kategorien werden anschließend aufsummiert:

$$\begin{aligned} \chi^2 &= \sum_{i=1}^{k} \frac{(x_i - n \cdot p_i)^2}{n \cdot p_i} \\ &= \frac{1}{n} \sum_{i=1}^{k} \frac{x_i^2}{p_i} - n. \end{aligned}$$

Die Wahrscheinlichkeitsverteilung der χ^2-Werte ist approximativ[36] bekannt, es ist die sogenannte χ^2-Verteilung mit $k-1$ Freiheitsgraden, welche die Wahrscheinlichkeit dafür angibt, daß ein empirisch ermittelter χ^2-Wert kleiner oder gleich einem vorgegebenen Wert v ist:

$$P(\chi^2 \leq v) = \frac{\gamma\left(\frac{k-1}{2}, \frac{v}{2}\right)}{\Gamma\left(\frac{k-1}{2}\right)}, \text{ wobei}$$

$$\gamma(a, x) = \int_0^x e^{-t} t^{a-1} dt \text{ und}$$

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

[35] Als Daumenregel gilt, daß mindestens $x_i > n \cdot \min p_i$ sein sollte, wobei es allerdings empfehlenswert ist, eine größere Stichprobe zu wählen, s. [68].

[36] Das ist auch ein Grund für die Mindestgröße n, welche die Stichprobe haben sollte, s. vorangehende Fußnote.

für die unvollständige bzw. vollständige Gammafunktion stehen. Soll ein Zufallsgenerator mit Hilfe eines χ^2-Tests auf Musterbildung oder Korrelationen untersucht werden, so wird zuerst eine Stichprobe aus dem Strom der Zufallsbits entnommen, indem n nichtüberlappende Blöcke fester Länge m herausgeschnitten werden. Die Bitfolgen der Blöcke werden als Binärkodierung ganzer Zahlen interpretiert; auf diesen n Zahlenwerten wird dann der χ^2-Test durchgeführt, wobei die Gleichverteilung der Zahlen ($p_i = (1/2)^m,\ \forall i$) als theoretische Verteilung gefordert wird.
Werden Rohdaten eines physikalischen Zufallsgenerators getestet, so ist es meist nicht sinnvoll einen χ^2-Test auf Gleichverteilung durchzuführen, da die Wahrscheinlichkeiten von Nullen und Einsen oft *bekanntermaßen* unterschiedlich sind, s. Abschnitt B.4. Diesem Umstand trägt man am besten Rechnung, indem die theoretischen Wahrscheinlichkeiten p_i für die Zahlenwerte der Bit-Blöcke gemäß der Anzahl der Nullen und Einsen im Block mit Hilfe der Wahrscheinlichkeiten $\hat{p}$ und $q = (1 - \hat{p})$ berechnet werden, wobei $\hat{p}$ aus der Stichprobe geschätzt wird, d.h. $\hat{p}_i = \hat{p}^{\#1} \cdot (1-\hat{p})^{m-\#1}$, wobei #1 für die Anzahl der Einsen im betrachteten m Bit langen Block steht.
Beim einstufigen χ^2-Test handelt es sich um einen einseitigen Test, dessen Auswertung nach der in Abschnitt 3.2.1.1 beschriebenen, allgemeinen Vorgehensweise bei einstufigen Tests erfolgt.
Allerdings sind einzelne χ^2-Tests für sich genommen noch nicht sehr aussagekräftig, es empfiehlt sich daher, zweistufige Tests durchzuführen; hierbei entnimmt man eine hinreichend große Anzahl von nichtüberlappenden Testsequenzen aus m Bit langen Blöcken, berechnet die einzelnen χ^2-Werte, sortiert sie nach aufsteigender Reihenfolge, ermittelt die empirische Verteilungsfunktion und vergleicht[37] sie mit der theoretischen χ^2-Verteilung mit $2^m - 1$ Freiheitsgraden.

3.2.5 Kontingenz-Test

Ein Kontingenz-Test [109] läßt sich immer dann gut einsetzen, wenn untersucht werden soll, ob innerhalb einer Gesamtheit von Elementen, die Ausprägungen zweier Merkmale dieser Elemente unabhängig voneinander[38] auftreten. Die beiden Merkmale, die untersucht werden, können im allgemeinen Fall eine Anzahl von möglichen diskreten Werten annehmen, die für die beiden Merkmale auch durchaus nicht gleich groß sein müssen. Da im folgenden einzelne Bits betrachtet werden, wird die Darstellung im weiteren auf den Spezialfall von jeweils zwei Werten pro Merkmal, nämlich **0** und **1**, beschränkt. Der Kontingenz-Test wird in diesem Fall dazu verwendet, Abhängigkeiten zwischen den Bits einer Sequenz aufzudecken. Hierbei stellen die nicht überlappenden Paare der Bitsequenz die betrachteten Elemente der Gesamtheit dar und die beiden Bits eines Paares entsprechen den beiden „Merkmalen“. Für physikalische Zufallsgeneratoren, bei denen besonders Korrelationen unmittelbar aufeinanderfolgender Bits interessieren, werden die

[37] Dies kann entweder visuell an Hand eines Graphen oder mit Hilfe der Kolmogorov-Smirnov-Statistik geschehen.

[38] Üblicherweise wird dieser Test natürlich genau gegenteilig angewendet, nämlich um zu prüfen, inwieweit zwei Merkmale gemeinsam auftreten. Hierzu muß man zuerst die Nullhypothese, d.h. den Umstand, daß Merkmalspaare lediglich *zufällig* gehäuft auftreten, auf einem (möglichst hohen) Sicherheitsniveau ablehnen. NB: Ist dies möglich, so bedeutet dies allerdings lediglich, daß es *sehr wahrscheinlich* ist, daß ein Zusammenhang zwischen den beiden Merkmalen besteht, zwingend ist dies nicht!

Paare entsprechend aus aufeinanderfolgenden Bits der Sequenz gebildet[39].
Beim *einstufigen* Kontingenz-Tests geht man folgendermaßen vor:

1. Man nehme eine Stichprobe mit n Elementen (Bit-Paaren).

2. Man zähle die Anzahl der Elemente n_{ik}, die zur i-ten Gruppe hinsichtlich des ersten Merkmals (erstes Bit) und zur k-ten Gruppe hinsichtlich des zweiten Merkmals (zweites Bit) gehören, wobei hier $i, k \in \{0, 1\}$ sind.

3. Weiter berechne man die Größen:

$$n_{i\cdot} = \sum_{k=0}^{1} n_{ik} \quad \text{und} \quad n_{\cdot k} = \sum_{i=0}^{1} n_{ik}, \text{ wobei gelten muß:}$$

$$n = \sum_{i=0}^{1} \sum_{k=0}^{1} n_{ik}.$$

4. Man überprüfe die Nullhypothese mit Hilfe eines χ^2-Tests.

Der χ^2-Wert wird hierbei gemäß:

$$\begin{aligned} \chi^2 &= n \cdot \sum_{i=0}^{1} \sum_{k=0}^{1} \frac{\left(n_{ik} - \frac{n_{i\cdot} n_{\cdot k}}{n}\right)^2}{n_{i\cdot} n_{\cdot k}}, \text{ oder explizit ausgeschrieben:} \\ \chi^2 &= \frac{n \cdot (n_{00} n_{11} - n_{01} n_{10})^2}{n_{0\cdot}\ n_{1\cdot}\ n_{\cdot 0}\ n_{\cdot 1}}, \text{ berechnet.} \end{aligned}$$

Es läßt sich zeigen [109], daß diese Größe einer χ^2-Verteilung mit einem Freiheitsgrad folgt.
Will man kleinste Korrelationen aufdecken, empfiehlt es sich, wieder einen *zweistufigen* Test durchzuführen; dies geschieht analog zu dem Vorgehen bei den zweistufigen χ^2-Tests.

3.2.6 Autokorrelationstest

Die Berechnung einer Autokorrelationsfunktion stellt allgemein eine gute Methode dar, um Strukturen möglicher Abhängigkeiten innerhalb eines (diskreten) stochastischen Signals – in unserem Fall einer diskreten, binären Zufallssequenz – aufzuspüren. Berechnet man hierbei lediglich *eine* Autokorrelationsfunktion anhand einer relativ kleinen Teilmenge der Daten, so lassen sich auf diese Weise zwar starke Abweichungen einer Sequenz von der Korrelationsfreiheit erkennen, aber geringe Abweichungen vom theoretischen Verhalten wird man auf diese Weise nicht aufdecken können. Will man diese ebenfalls erkennen können, bedarf es wieder eines zweistufigen Vorgehens.
Es werden beim Test der quantenoptischen Zufallsgeneratoren daher auch zwei verschiedene Arten von Autokorrelationstests verwendet:

[39] Grundsätzlich können aber auch Paare aus nicht aufeinanderfolgenden Bits untersucht werden; wie bereits erwähnt, kann dies bei der Untersuchung von Pseudozufallszahlengeneratoren sinnvoll sein.

1. Ein Autokorrelationstest, bei dem die standardnormierte Variable $x(d)$ für aufeinanderfolgende Verschiebungen d mit $1 \leq d \leq k$ berechnet wird, wobei typische Werte für k zwischen 32 und 128 liegen. Es ist zwar kein Problem, die Autokorrelationsfunktion für größere Verschiebungswerte zu berechnen, da aber bei physikalischen Prozessen, die zur Zufallsgenerierung verwendet werden, die Autokorrelationswerte zu größeren Verschiebungen hin generell abnehmen[40], ist dies unnötig. Dieser Test dient nur dazu, sich einen ersten Überblick[41] über etwaige statistische Defekte zu verschaffen; so wird er z.B. dazu benutzt, die in Abschnitt 5.1.2.5 geschilderten anfänglichen Probleme mit der neuen Datenaufnahme-Elektronik aufzudecken bzw. zu illustrieren. Selbstverständlich kann man auch mehrere solcher Tests in einer Grafik zusammenfassen, um etwaige Tendenzen bei bestimmten Autokorrelationskoeffizienten deutlicher zu erkennen.

2. Ein Autokorrelationskoeffizienten-Verteilungstest[42], bei dem für eine feste Verschiebung d, wobei insbesondere $d = 1$ gewählt wird, auf einer großen Anzahl von Stichproben (typ. 128) gleicher Länge (typ. 8 kB) die jeweilige standardnormierte Variable $x(d)$ berechnet wird. Aus den Realisierungen der standardnormierten Variable wird anschließend eine empirische Verteilungsfunktion ermittelt, graphisch dargestellt und mit Hilfe eines Kolmogorov-Smirnov-Tests mit der theoretisch zu erwartenden Normalverteilung verglichen. Da sich statistische Defekte physikalischer Zufallsgeneratoren am stärksten beim ersten Autokorrelationskoeffizienten bemerkbar machen, werden die Tests hauptsächlich auf Stichproben mit einer festen Verschiebung[43] von $d = 1$ durchgeführt. Teilweise wird auch noch eine dritte Teststufe verwendet, um die Abweichungen einer größeren Zahl dieser Autokorrelationskoeffiziententests von der theoretischen Verteilung zu erfassen, s. Abschnitt B.10.3.

3.2.6.1 Theoretische Verteilung der Autokorrelationskoeffizienten

Die Berechnung der Autokorrelationskoeffizienten geschieht bei den Tests auf eine etwas andere Weise [91] als sonst üblich: Anstatt der Multiplikation wird eine XOR-Operation verwendet. Hierdurch wird von vornherein die binäre Natur der Zufallssequenzen und die Symmetrie der Nullen und Einsen bei der theoretischen Idealverteilung berücksichtigt. Folgende Formel dient zur Berechnung der Autokorrelationskoeffizienten:

$$A(d) = \sum_{i=0}^{n-d-1} b_i \oplus b_{i+d}.$$

[40] Ausnahme: Bei periodischen Prozessen können sie auch wieder zunehmen, aber bei solchen Prozessen haben die Autokorrelationskoeffizienten i. a. bereits bei kleinen Verschiebungen Werte, die sich stark von den Werten idealer Zufallsprozesse unterscheiden.

[41] So müßte bei einer quantitativen Auswertung z.B. auch berücksichtigt werden, daß die einzelnen Autokorrelationskoeffizienten nicht voneinander unabhängig sind.

[42] Der Kürze wegen, wird im weiteren für diesen Verteilungstests der Ausdruck Autokorrelationskoeffiziententest verwendet.

[43] Angesichts einer Strichprobenlänge von typ. 8 kB und kleinen Verschiebungen sind Randeffekte natürlich vernachlässigbar.

Hierbei ist d wieder die Verschiebung[44], n die Länge der Sequenz, auf der die Autokorrelation berechnet wird, b_i der Wert des Bits[45] an der Stelle i und $\oplus$ die XOR-Operation, d.h. die Addition Modulo 2.
Für die Nullhypothese unabhängiger Bits mit den Wahrscheinlichkeiten $P(b_i = 1) := p$ für eine Eins und $P(b_i = 0) := q = 1 - p$ für eine Null, gilt:

$$P\left(b_i \oplus b_j = 1\right) \;=\; (p \cdot q) + (q \cdot p) = 2 \cdot p \cdot (1 - p) := P,$$

so daß $A(d)$ binomialverteilt ist:

$$b\left(A(d), n - d, P\right) = \binom{n-d}{A(d)} \cdot P^{A(d)} \cdot (1 - P)^{(n-d)-A(d)}.$$

Für die theoretische Verteilungsfunktion müßte man dementsprechend die kumulierte Binomialverteilung verwenden. Für so große n, wie sie bei den Tests auftreten, ist dies allerdings nicht sehr praktikabel, es wird daher die Normalapproximation[46] mit standardnormierter Variable verwendet, die hier berechnet wird gemäß:

$$x = \frac{A(d) - (n - d) \cdot P}{\sqrt{(n - d) \cdot P \cdot (1 - P)}}.$$

Da der Wert von P nur im Idealfall ($p = q = 1/2$) bekannt ist, muß er für bekanntermaßen tendenzbehaftete Zufallsbits aus der Stichprobe geschätzt werden. Hierzu schätzt man die Wahrscheinlichkeit p, daß ein Bit gleich Eins ist, mit Hilfe der relativen Häufigkeit der Einsen ab. Hierfür lassen sich entweder die relativen Häufigkeiten in der Teilstichprobe oder in der Gesamtstichprobe verwenden. Angesichts der i. a. recht großen Werte der Teilstichprobenlänge n spielt es aber keine Rolle, welche Schätzung man für die Einswahrscheinlichkeit benutzt. Bei den später in Abschnitt B.9 aufgeführten Tests wird daher immer die aus der Gesamtstichprobe geschätzte Wahrscheinlichkeit für eine Eins verwendet[47].

3.2.7 Universelle Tests nach Maurer und Coron

Der von Ueli Maurer entwickelte universelle statistische Test für Zufallsgeneratoren [90] ist der einzige Test, der von vornherein für die Untersuchung von Hardware-Zufallsgeneratoren entworfen wurde. Der Test ist in der Lage bei entsprechender Länge der Bit-Sequenz, einen idealen Zufallsgenerator von einem Generator zu unterscheiden,

[44] Die Verschiebung d sollte hierbei möglichst im Bereich $1 \leq d \leq \lfloor n/2 \rfloor$ liegen, um Randeffekte zu vermeiden; dies ist bei den späteren Tests, die typische Werte von $n = 65536$ und $d < 128$ verwenden, immer der Fall.

[45] Um die Laufzeit des Tests zu minimieren, arbeitet die Programm-Routine natürlich direkt mit Bytes und nicht mit einzelnen Bits. Weitergehende Optimierungen, wie z.B. der Einsatz von FFT-Algorithmen, wurden aber nicht weiter verfolgt.

[46] Es gäbe sonst ohnehin ein Problem bei der Anwendung des Kolmogorov-Smirnov-Tests, da dieser stetige Verteilungen für die theoretische Verteilung voraussetzt.

[47] Dies ist ein restriktives Vorgehen, da etwaige systematische Abweichungen der empirischen Verteilung der x gegenüber der theoretischen Standardnormalverteilung bei einer Schätzung aus der betrachteten Teilstichprobe kleiner wären.

bei dem Einsen und Nullen nicht gleich wahrscheinlich sind oder Korrelationen zwischen aufeinanderfolgenden Bits bestehen[48]. Die Grundidee des Tests besteht darin, daß sich ein statistischer Test mit Hilfe eines universellen Quellkodierers realisieren läßt, d.h. man verwendet die „Komprimierbarkeit“ einer Bitsequenz als Maß für ihre „Zufälligkeit“, s. a. Abschnitt B.8. Allerdings ist es nicht unbedingt notwendig, auch tatsächlich die Bitsequenz zu komprimieren, es reicht aus, eine Größe zu berechnen, die ein Maß für die Länge der komprimierten Sequenz darstellt.
Ein Vorteil dieses Tests besteht darin, daß er eine kryptographisch relevante Größe ermittelt: die *effektive Schlüssellänge*. Hierbei nimmt man an, daß nicht überlappende, L Bit lange Blöcke aus der vom Generator erzeugten Bitsequenz als Schlüssel[49] für ein Chiffriersystem fortlaufend entnommen werden. Maurer definiert hierbei die effektive Schlüssellänge als den Logarithmus[50] zur Basis 2 aus der minimalen Anzahl der Schlüssel, bezeichnet mit $\mu_Q(L, 1/2)$, die ein Angreifer bei optimaler Suchstrategie testen muß, um den richtigen Schlüssel mit einer Wahrscheinlichkeit von $1/2$ zu finden[51], wenn ein L Bit langer Schlüssel verwendet wurde. Für einen *idealen* Zufallsgenerator ist die effektive Schlüssellänge gerade gleich $\log_2 \mu_Q(L, 1/2) = n - 1$.
Weiter zeigt er, daß die effektive Schlüssellänge für eine *binäre Quelle ohne Gedächtnis* ($BQoG_p$), die den Bitwert Eins mit einer Wahrscheinlichkeit p ausgibt, gleich:

$$\log_2 \mu_{BQoG_p}(L, 1/2) \approx \log_2 \sum_{i=0}^{pL} \binom{L}{i} \text{ ist.}$$

Die effektive Schlüssellänge ist überdies nicht nur für kryptanalytische Zwecke interessant, sondern sie stellt auch ein generell interessantes Charakterisierungsmerkmal für eine Binärsequenz dar. Dies rührt daher, daß ein enger Zusammenhang zwischen ihr und der Entropie pro Bit, definiert durch:

$$H(x) = -p \log_2 p - (1-p) \log_2 (1-p), \text{ mit } 0 < p < 1 \text{ und}$$

$$H(0) = H(1) = 0,$$

besteht. Die auf die Blocklänge L normierte, effektive Schlüssellänge erreicht nämlich im Grenzwert großer Blocklängen asymptotisch die Entropie pro Bit, d.h.:

$$\lim_{L \to \infty} \frac{\log_2 \mu_{BQoG_p}(L, 1/2)}{n} = H(p).$$

Die binäre Entropie $H(p)$ entspricht daher gerade dem Faktor, um den die effektive Schlüssellänge verringert wird, wenn es sich nicht um eine symmetrische Quelle handelt.

[48] Präziser ausgedrückt: Er kann eine *binäre, gedächtnislose, symmetrische Quelle* von *binären, gedächtnislosen, unsymmetrischen Quellen* und *binären, (stationären), symmetrischen oder unsymmetrischen Quellen mit endlichem Gedächtnis* unterscheiden.

[49] Bei einem guten Schlüssel-Generator sind alle 2^L möglichen Schlüssel gleich wahrscheinlich, bei einem schlechten Generator hingegen könnte ein Angreifer versuchen, die Zeit, die eine erfolgreiche Schlüsselsuche benötigte, dadurch zu reduzieren, daß er mit den wahrscheinlichsten Schlüsseln anfängt; in diesem Fall wäre die *effektive* Schlüssellänge kleiner, als angesichts der Blocklänge zu erwarten.

[50] Logarithmus zur Basis 2: Liefert direkt die Anzahl der Bits, die zur Darstellung des Arguments notwendig sind.

[51] Die Wahl des Wertes $1/2$ für die Erfolgswahrscheinlichkeit ist etwas willkürlich, aber nicht unüblich.

Eine analoge Relation zwischen effektiver Schlüssellänge und Entropie pro Bit[52] gilt auch im allgemeinen Fall einer ergodischen, stationären Quelle.
Die Nachteile des universellen Tests seien aber in dieser Stelle auch nicht verschwiegen:

- Die notwendige Länge der Testsequenz vergrößert sich exponentiell mit der Blocklänge, so daß der Test schon bald nicht mehr angemessen durchführbar ist [51]. Dies ist insbesondere deswegen problematisch, weil zu kurze Testsequenzen immer zu einer Ablehnung des getesteten Generators führen, quasi unabhängig von der tatsächlichen Qualität.

- Der Test ist nicht in der Lage Defekte zu *klassifizieren*, d.h. er zeigt lediglich, daß ein Generator mit hoher Wahrscheinlichkeit nicht ideal ist, aber nicht aufgrund *welchen* Defektes. Dies ist insofern bedauerlich, als dies das Testen von Rohdaten eines physikalischen Zufallsgenerators quasi ausschließt, da bei Rohdaten die Wahrscheinlichkeiten für Nullen und Einsen nicht exakt gleich sind und der universelle Test dies natürlich als Defekt erkennt. Eventuell vorhandene, tieferliegende statistische Defekte des Generators werden also hiervon bereits maskiert und lassen sich daher mit Hilfe des universellen Tests nicht aufdecken.

Die Theorie des universellen Tests und die bei einer Durchführung des Tests zu beachtenden Details sind im Anhang, Abschnitt E beschrieben. Zusätzlich wird dort auch noch eine verbesserte Variante des universellen Tests von Jean-Sébastian Coron dargelegt. Diese Variante verwendet eine leicht modifizierte Teststatistik, die gegenüber der ursprünglichen Teststatistik den Vorteil hat, daß ihr Erwartungswert bis auf statistische Schwankungen für *alle* Blocklängen der Entropie der Quelle entspricht und nicht erst im Grenzwert der Blocklänge gegen unendlich.

3.2.8 Tests aus der nichtlinearen Zeitreihen-Analyse

Im Zuge der immer weitergehenden Modellierung komplexer Systeme wurde in den letzten Jahrzehnten festgestellt, daß in der Natur dynamische Systeme existieren, die eine gewisse Strukturähnlichkeit mit Pseudozufallsgeneratoren aufweisen: Ähnlich wie ein Pseudozufallszahlengenerator trotz eines relativ einfachen Algorithmus eine sehr lange Sequenz von Pseudozufallszahlen erzeugen kann, die sich in ihren Eigenschaften nur schwer oder sogar überhaupt nicht von genuin zufälligen Sequenzen unterscheiden läßt, kann in der Natur eine scheinbar stochastische Zeitreihe einer Meßgröße dennoch von einem deterministischen System mit einer mathematisch verhältnismäßig einfach zu beschreibenden Dynamik erzeugt worden sein. Solche Systeme weisen eine nichtlineare Dynamik [114] mit starken Mischungseigenschaften[53] im Phasenraum auf.
Lag am Anfang das Augenmerk der Wissenschaft hauptsächlich auf der Untersuchung der Dynamik von nichtlinearen Modell-Gleichungen, wie z.B. der logistischen Abbildung, gewann später auch die Untersuchung von empirisch gewonnenen Zeitreihen auf Vorliegen von niedrigdimensionalem Chaos zunehmend an Bedeutung. Hierbei wird ver-

[52] In diesem Fall handelt es sich nicht um die binäre Entropie. So gehen z.B. für den Fall einer Quelle mit endlichem Gedächtnis sowohl die stationären Zustandswahrscheinlichkeiten als auch die Übergangswahrscheinlichkeiten in die Entropie mit ein, s. [90].

[53] Hierin ähneln sie gängigen Blockchiffren.

sucht, ein niedrigdimensionales[54], deterministisches, dynamisches System zu finden, das direkt oder indirekt als „Generator“ der Zeitreihe angesehen werden kann und eine tiefere Erkenntnis verbunden mit einer besseren Vorhersagbarkeit der Zeitreihe erlaubt. Die im Rahmen der Analyse von Zeitreihen niedrigdimensionaler chaotischer Systeme entwickelten Methoden wurden dann für die Untersuchung von beliebigen Zeitreihen verallgemeinert und dem Reigen traditioneller Analysemethoden hinzugefügt, s. [65] für eine ausführliche und kritische Darstellung dieser Analysemethoden.
Allerdings ist es natürlich nicht immer möglich, eine Zeitreihe auf ein einfaches, dynamisches System zurückzuführen, da es selbstverständlich auch häufig vorkommt, daß Zeitreihen von komplexen, nur durch eine große Zahl von Parametern zu beschreibenden Systemen generiert werden (hochdimensionales Chaos) oder sich überhaupt nur stochastisch beschreiben lassen, wie dies z.B. der Fall ist, wenn sie durch das Zusammenwirken einer Vielzahl von (quasi-) unabhängigen Faktoren zustande kommen.
Dementsprechend ist eine Anwendung von Analysemethoden aus der nichtlinearen Dynamik bei der Untersuchung von Zeitreihen auch nur dann erfolgversprechend, wenn eine einfache, der Zeitreihe unterliegende Dynamik theoretisch plausibel erscheint und überdies nicht durch zusätzliche, systemexterne Faktoren, wie z.B. Rauschen, stark maskiert wird. Das Anwenden von Methoden aus der nichtlinearen Dynamik bei der Analyse von Zeitreihen, bei denen diese Voraussetzungen nicht gegeben sind, ist hingegen ein Vabanquespiel, da es leicht zu einer Mißinterpretation der Ergebnisse kommen kann[55].
Lassen sich nun die Analysemethoden aus der nichtlinearen Dynamik für die Untersuchung der Ausgabesequenzen von Zufallsgeneratoren verwenden?
Diese Frage läßt sich nicht pauschal beantworten, sondern es müssen zwei Fälle unterschieden werden:

1. Die Binärsequenzen sind (indirekt) von einem Generationsmechanismus mit einer mathematisch einfachen, (teilweise) deterministischen Dynamik erzeugt worden.

2. Jedes Bit der Sequenz wird von einem Zufallsmechanismus erzeugt, der stochastisch ist, d.h. höchstens sehr geringe „Determinismus-Verunreinigungen“ durch externe Einflüsse aufweist.

Im ersten Fall kann eine Untersuchung mit den Methoden der nichtlinearen Dynamik sinnvoll sein, insbesondere natürlich dann, wenn sogar bekanntermaßen ein deterministisches System eingesetzt wird, wie im Falle des Generators von T. Kuusela, s. S. 20.
Im zweiten Fall lassen sich viele Methoden aus der nichtlinearen Dynamik, wie z.B. die diversen Phasenraummethoden[56], nicht sinnvoll einsetzen, da in diesem Fall die prinzipiellen Voraussetzungen für ihren Einsatz nicht gegeben sind.
Eine Ausnahme stellen Methoden dar, die kein niedrigdimensionales Modell unterstellen, sondern eine informationstheoretische Charakterisierung der Sequenz erlauben.

[54] Die Dimensionalität bezieht sich hierbei auf den Phasenraum des dynamischen Systems und entspricht der Anzahl seiner Freiheitsgrade.

[55] Kapitel 7 des Buches über nichtlineare Zeitreihenanalyse [65] von H.Kantz und T. Schneider widmet sich ausführlich der Problematik der Anwendung nichtlinearer Methoden, wenn die deterministischen Anteile einer Zeitreihe nur schwach ausgeprägt sind.

[56] Alle phasenraumbasierten Methoden haben Probleme bei hochdimensionalen Systemen, da eine enorme Datenanzahl notwendig ist, um z.B. bei hohen Dimensionen eine genügend hohe Dichte von Meßpunkten zu erreichen. Hinzu kommt noch, daß die notwendige Phasenraumrekonstruktion durch zusätzliches Rauschen extrem erschwert wird.

3.2.8.1 Die Transinformation

Insbesondere das informationstheoretische Maß *Transinformation* [32] ist ein interessantes Analyse-Instrument für die Untersuchung stationärer Symbolsequenzen. Sie erlaubt es, Korrelationen innerhalb von Symbolsequenzen aufzuspüren. Betrachtet man Symbole c_1 und c_2 im Abstand m voneinander, so berechnet sich die Transinformation T^m als Summe über alle möglichen Symbolpaare, gemäß:

$$T^m = k \cdot \sum_{(c_1,c_2)\in A^2} p^{(m)}\left(c_1, c_2\right) \cdot \ln \frac{p^{(m)}\left(c_1, c_2\right)}{p(c_1) \cdot p(c_2)},$$

wobei A für das Alphabet der Symbole steht, $p(c)$ die relative Häufigkeit des jeweiligen Einzelsymbols ist und $p(m)\left(c_1, c_2\right)$ die relative Häufigkeit der jeweiligen Symbolpaare. Wählt man statt des natürlichen Logarithmus den Logarithmus zur Basis Zwei, so ergibt sich für die Normierungskonstante k gerade der Wert $k = 1$ und die Transinformation gibt dann die Information in Bits an, die man über das im Abstand m folgende zweite Symbol bereits erhält, wenn man das erste Symbol „empfängt“. Die Verwendung der Transinformation als Analyse-Instrument hat eine Reihe von Vorteilen [53], da sie:

- *jede* Art von Abhängigkeit zwischen den untersuchten Symbolen einer Sequenz feststellt und nicht nur die linearen Abhängigkeiten,
- invariant bezüglich Koordinaten-Transformationen ist,
- keine Zuweisung von Zahlen (und Verwendung von algebraischen Operationen) für die Symbole notwendig macht und keine Struktur des Zustandsraumes unterstellt,
- dann und nur dann verschwindet, wenn die Symbole c_1 und c_2 im Abstand m voneinander unabhängig sind und
- aufgrund der Auswertung von Symbolpaaren, auch noch bei Stichprobengrößen verwendet werden kann, bei denen Methoden, welche die relative Häufigkeiten von Symbol-Blöcken auswerten, bereits zu starke statistische Schwankungen zeigen.

Allerdings hat die Transinformation auch Nachteile:

- Sie ist nicht so spezifisch wie die Autokorrelation: Alle Abweichungen gehen in eine Zahl ein, die nur positiv oder gleich Null sein kann.
- Für endliche Sequenzen gibt es einen systematischen Schätzfehler [83], der zu einem zu großen Wert der Transinformation führt, sich allerdings korrigieren läßt.

So interessant die Transinformation für die Untersuchung von Abhängigkeiten innerhalb von Symbolsequenzen auch sein mag, H. Herzel et al. [53] haben in Fortführung der Arbeit von Wentian Li [83] gezeigt, daß die Transinformation just für unseren Spezialfall binärer Sequenzen gegenüber der Autokorrelationsfunktion keine zusätzlichen Informationen erbringt; erst ab ternären Symbolsequenzen liefert sie zusätzliche Informationen. Deshalb wird bei den Untersuchung der Bitsequenzen der Generatoren auch auf diese Untersuchungsmethode verzichtet.

Kapitel 4

Aufbau quantenoptischer Zufallsgeneratoren

In diesem Kapitel werden die experimentellen Details der verschiedenen, realisierten Aufbauvarianten quantenoptischer Zufallsgeneratoren beschrieben. Mögliche Veränderungen und Varianten, die nicht experimentell realisiert wurden, werden in der Diskussion, Abschnitt 6.3 dargelegt.

4.1 Prinzipieller Aufbau quantenoptischer Zufallsgeneratoren

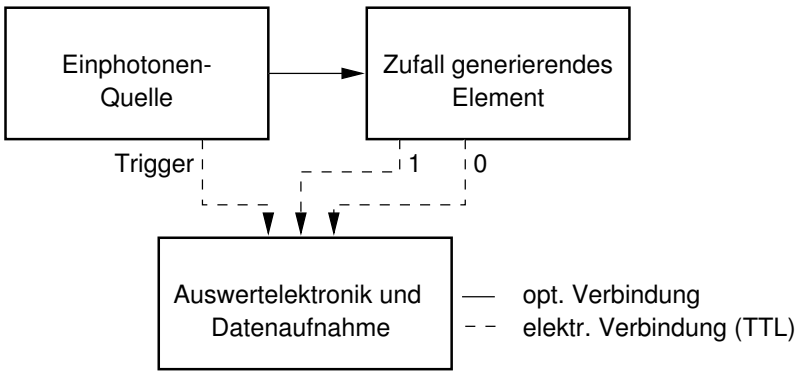


Abbildung 4.1: Prinzipieller Aufbau quantenoptischen Zufallszahlengeneratoren

Beim prinzipiellen Aufbau eines quantenoptischen Zufallsgenerators, s. Abb. 4.1, lassen sich drei grundlegende Bestandteile unterscheiden:

- eine Photonenquelle, sei es eine Quelle, die approximative Photonen-Anzahlzustände aussendet, oder eine Quelle, die stark abgeschwächten Poisson-Lichtfelder emittiert,
- das eigentliche, Zufall generierende Element, wobei die Detektion des Photons ebenfalls mit dazu gehört, und

- die Elektronik für die Signalverarbeitung, Datenaufnahme und Übertragung zum Meßrechner.

Die Varianten quantenoptischer Zufallsgeneratoren unterscheiden sich hierbei durch die verschiedenen Photonenquellen und Zufall generierenden Elemente.
In den folgenden Abschnitten werden die einzelnen Komponenten der realisierten quantenoptischen Zufallsgeneratoren detailliert beschrieben.

4.2 Die Einphotonenquelle auf Basis der parametrischen Fluoreszenz

Zur Erzeugung einzelner Photonen wird die parametrische Fluoreszenz (s. Abschnitt 3.1.2.1) verwendet. Die Einphotonenquelle selbst setzt sich aus drei Bestandteilen zusammen:

- dem Pumplaser mit zwei weiteren optischen Komponenten (Filter und Linearpolarisator),
- dem optisch nichtlinearen Kristall (entweder temperaturgeregelt oder drehbar mit Hilfe einer Winkelverstelleinheit) und
- einem Triggerdetektor.

4.2.1 Der Pump-Laser

Die Zentralwellenlängen der einzelnen Photonen eines Paares entsprechen der doppelten Pumpwellenlänge. Es werden Einzelquantendetektoren auf Si-Basis benutzt, da sie die zur Zeit höchste Quanteneffizienz aufweisen, s. Abschnitt 4.4.1.2. Damit sich diese Detektoren hinreichend effizient einsetzen lassen, dürfen allerdings weder Signal- noch Idlerwellenlänge über 950 nm liegen, da die Quanteneffizienz der Detektoren ihr Maximum bei ca. 700 nm hat und zu größeren Wellenlängen hin quasi linear abnimmt, s. Abb. 4.6.
Am günstigsten wäre es natürlich, wenn die Zentralwellenlängen in der Nähe des Maximums der Detektionseffizienz der Detektoren, d.h. bei ca. 700 nm, lägen. Dies setzte allerdings voraus, daß man einen UV-Laser mit einer Wellenlänge von 350 nm als Pumpe verwendete. Da kein kleiner UV-Laser zur Verfügung steht und die Verwendung eines leistungsstarken (teureren) UV-Argonionen-Laser für diese Aufgabe, gerade auch im Hinblick auf den späteren Einsatz des Generators, nicht sinnvoll erscheint, wird ein im Physikalischen Institut bereits vorhandener Helium-Cadmium-Laser mit einer Wellenlänge von 442 nm für die Experimente verwendet (Omnichrome Serie 56X, Model 4056R-S-A01 mit einer spezifizierten Ausgangsleistung von 16 mW bei 442 nm). Dieser Laser ist relativ kompakt und braucht keine Wasserkühlung; zudem liegt seine Betriebswellenlänge mit 442 nm unterhalb der sichtbaren Wellenlängen eines Argon-Ionen-Lasers[1].
Der Laser steht getrennt vom restlichen Aufbau des quantenoptischen Zufallsgenerators, der sich im Inneren eines lichtdichten, schwarzen Holzkastens befindet. Dies empfiehlt

[1] Wassergekühlte Laser scheiden aus praktischen Gründen aus und luftgekühlte Argon-Ionen-Laser haben Wellenlängen über 450 nm.

sich, da der eingebaute Lüfter des Lasers und das durch die Lüftungsschlitze dringende Licht der Plasmaröhre den optischen Aufbau nur stören würden. Der Pumplaserstrahl wird mit Hilfe eines Spiegels durch eine kleine Öffnung hindurch in den Kasten gelenkt. Da es sich bei dem verwendeten Laser um einen Gas-Metalldampf-Laser handelt, der neben der gewünschten Laseremission im blauen auch rein thermische Strahlung im infraroten Spektralbereich abgibt, bedarf es noch eines Blauglases (Typ Schott BG 39) als Filter, um zu verhindern, daß diese Strahlung in den nachfolgenden optischen Aufbau eindringt und das Signal-Rauschverhältnis beeinträchtigt.
Leider emittiert der verwendete Laser zufällig polarisierte Wellenzüge, die nur innerhalb ihrer Kohärenzzeit eine feste Polarisation aufweisen, daher muß zusätzlich noch ein Linearpolarisator (Glan-Thompson-Polarisationsprisma) verwendet werden, um das Pumplicht zu polarisieren; dies bedeutet allerdings, daß nur die Hälfte der Ausgangsleistung des Lasers zum Pumpen des Kristalls zur Verfügung steht. Bei der noch neuen Laserröhre sind dies 10 mW; im Laufe der Zeit wird die Laserleistung auf den spezifizierten Wert abnehmen, so daß dann nur noch ca. 8 mW Pumpleistung zur Verfügung stehen werden. Durch Drehen des Linearpolarisators läßt sich die effektive Pumpleistung bei Bedarf auch weiter senken, da nur der Anteil des Lichtes mit der für die Phasenanpassung richtigen Polarisationsrichtung zum Prozeß der parametrischen Fluoreszenz beiträgt, s. Abschnitt 3.1.2.1.

4.2.2 Der optisch nichtlineare Kristall

Bei allen Experimenten wird die parametrische Fluoreszenz mit Typ-I-Phasenanpassung verwendet. Die beiden Photonen eines Paares haben hierbei die gleiche Polarisationsrichtung. Will man sie effizient trennen, muß man dafür sorgen, daß sie vom Kristall in verschiedene Richtungen emittiert werden. Dies erreicht man durch eine Phasenanpassung für den nichtkollinearen Fall, s. Abb. 3.4 (auf S. 39) bzw. Abb. 4.4 (auf S. 67).
Beim ersten Aufbau der Einphotonenquelle, s. Abb. 4.2 auf S. 65, wurde ein mit Hilfe der Temperatur phasenangepaßter Kaliumniobat-Kristall[2] verwendet.
Kaliumniobat, $KNbO_3$, ist eines der effizientesten optisch nichtlinearen Medien, das insbesondere gern bei der Frequenzverdopplung von Laserdioden benutzt wird. Die Temperaturphasenanpassung ermöglicht es, allein durch Verändern der Temperatur einen kontinuierlichen Übergang zwischen kollinearem und nichtkollinearem Fall zu erreichen. Dies bedeutet allerdings auch, daß selbst kleine Temperaturveränderungen, wie sie bei der Temperaturregelung des Kristalls durchaus vorkommen, bereits die Abstrahlcharakteristik des Kristalls verändern. Da überdies aufgrund der verwendeten Pumpwellenlänge eine relativ hohe Phasenanpasstemperatur (80,4°C) eingestellt werden muß, die laut Hersteller auf die Dauer zu einer Beeinträchtigung der optischen Eigenschaften[3] des Kristalls führen kann, wurde im Hinblick auf einen Dauerbetrieb des Generators der Kaliumniobat-Kristall bei späteren Experimenten gegen einen Lithiumjodat-Kristall ausgetauscht.

[2] Der Kristall (HBL: $3 \times 3 \times 5\ mm^3$) ist mit Hilfe zweier Plastikschrauben an die Wände der Öffnung eines durchbohrten, temperierbaren Kupferblocks gedrückt. Der Kupferblock wird zur thermischen Isolierung von einem Styroformblock umgeben, der zwei einander gegenüberliegende Öffnungen für das Pumplicht bzw. das aus dem Kristall auslaufende Lichtfeld hat.

[3] In der Tat war dies stellenweise auch schon am verwendeten Kristall zu sehen.

Lithiumjodat, $LiIO_3$, ist ähnlich effizient wie Kaliumniobat und hat überdies den Vorteil, auch im ultravioletten Wellenlängenbereich transparent zu sein[4], d.h. der Spielraum bei der Wahl des Pumplasers ist größer. Als Nachteil hat zu gelten, daß Lithiumjodat hygroskopisch ist. Allerdings wird dies durch die ohnehin empfehlenswerte Antireflexbeschichtung stark abgemildert. Für den Dauerbetrieb empfiehlt es sich dennoch, einen komplett von der Luft isolierten Lithiumjodat-Kristall zu verwenden[5].
Beim Laboraufbau werden die Auswirkungen der Luftfeuchtigkeit auf den Lithiumjodat-Kristall durch folgende Maßnahmen gering gehalten: Der Kristall befindet sich in einer Fassung, deren zum Pumplaser gerichtete Seite von einem Deckglas luftdicht abgeschlossen wird. Die gegenüberliegende Seite bleibt zwar offen[6], an die Kristallfassung schließt sich aber direkt ein mit Silika-Gel-Perlen gefüllter Metallbehälter an, bei dem in der Mitte ein rechteckiger Ausgangskanal für das Fluoreszenzlicht ausgespart ist. Die seitlichen Wände des Ausgangskanals sind dabei aus Drahtnetz, so daß das Silika-Gel die Luftfeuchtigkeit der von außen in den Kanal eintretenden Luft senken kann. Zusätzlich wird der Kristall in der Fassung und mit dem daran befestigten Metallbehälter außerhalb der Meßzeiten in einem Vorratsbehälter mit Silika-Gel aufbewahrt.
Da der Lithiumjodat-Kristall winkel-phasenangepaßt wird und mittlere Temperaturschwankungen (d.h. $< \pm 30°C$) die Brechungsindices nicht maßgeblich ändern, ist der Aufbau nach einmaliger Einstellung des gewünschten Phasenanpasswinkels sehr stabil – Schwankungen wie beim Einsatz des Kaliumniobat-Kristalls treten nicht auf, s. Abschnitt 5.1.2.3.
Für den Laboraufbau wurde ein antireflexbeschichteter Lithiumjodat-Kristall mit einer Länge von 9 mm (10 mm Breite, 7 mm Höhe) verwendet, der vom Hersteller bereits vorkonfektioniert wurde (Schnittwinkel von 39° zur optischen Achse). Da der Schnittwinkel vom Hersteller nur auf 0,5° Genauigkeit spezifiziert wird, ist der gefaßte Kristall samt Luftfeuchtigkeitsschutzvorrichtung mit einer Halterung auf einer Winkelverstelleinheit befestigt, die es erlaubt, den Kristall um eine Achse senkrecht zum Pumpstrahl zu drehen. Durch Verändern des Drehwinkels lassen sich die Abstrahlwinkel von Signal- und Idler-Strahl beeinflussen, wobei es sich empfiehlt, möglichst kleine Abstrahlwinkel zu wählen, da der Prozeß der parametrischen Fluoreszenz für diese am effizientesten ist. Allerdings hat dies auch Nachteile: Es ist ein relativ langer optischer Weg nötig, um eine gute räumliche Trennung von Pump-, Signal- und Idler-Strahl zu erreichen. Im aktuellen Aufbau wird die Trennung der Strahlen durch zwei hochreflektierende Silberspiegel (12 mm Durchmesser) erreicht. Sie sind im Abstand von 67 cm hinter der Quelle derart aufgestellt, daß der eine den Idler-Strahl auf einen Photonendetektor lenkt, der als Triggerdetektor fungiert, während der andere den Signal-Strahl auf das zufallgenerierende Element schickt. Die beiden Spiegel stehen eng nebeneinander und sind dabei so angeordnet, daß der Pumpstrahl gerade zwischen ihnen hindurchläuft und auf einen Strahlabsorber trifft, so daß der Streulichtanteil im Aufbau gering gehalten werden kann. Der Winkel zwischen den Hauptausbreitungsrichtungen der Photonen der Photonenpaare und dem Pumpstrahl beträgt ca. 0,3°.

[4] Kaliumniobat ist für Wellenlängen unter 420 nm stark absorbierend.

[5] Die Kristalle sind kommerziell auch in durchsichtigen Behältern erhältlich, die mit einem Schutzgas oder einer im Brechungsindex angepaßten Schutzflüssigkeit gefüllt sind.

[6] Dies geschieht, um eine möglichst hohe Effizienz zu erreichen; ein Vorgehen wie bei der Eingangs-Seite würde aufgrund von Reflexion zu Verlusten bei den Photonenpaaren führen.

Bei dem Trigger-Photonendetektor handelt es sich um ein passiv gequenchtes Detektormodul der Firma EG&G, s. Abschnitt 4.4.1.2.

4.3 Die Photonenquelle auf Basis abgeschwächter Lichtpulse

Einphotonenquellen auf Basis abgeschwächter Lichtpulse lassen sich mit verhältnismäßig geringem technischen Aufwand realisieren, da ihr Grundaufbau recht einfach ist: Man nehme eine herkömmliche Lichtquelle, z.B. eine Leuchtdiode oder eine Laserdiode, und schwäche diese mit Hilfe von Filtern so stark ab, daß die Wahrscheinlichkeit für das gleichzeitige[7] Auftreten mehrerer Photonen hinreichend[8] klein ist. Diese Photonenquellen lassen sich weiter unterteilen in gepulste und kontinuierliche Quellen. Bei den im folgenden vorgestellten Aufbauten kommen nur gepulste Quellen zum Einsatz, da der Grundaufbau (s. Abb. 4.1) eine Photonenquelle mit Trigger-Ausgang vorsieht. Dies empfiehlt sich, da hierdurch Auswirkungen der Detektor-Dunkelzählrate oder möglicher, äußerer Manipulationsversuche minimiert werden. Anders als bei der Einphotonenquelle auf Basis der parametrischen Fluoreszenz, die einen Triggerdetektor benötigt, ist dies bei der Einphotonenquelle, die stark abgeschwächte Pulse emittiert, nicht der Fall, denn das elektrische Pulssignal, das die Lichtquelle ansteuert, kann gleichzeitig als Triggersignal verwendet werden.
Bei der konkreten Wahl der Lichtquelle für die statistische Einphotonenquelle sollten außerdem noch folgende Zusatzbedingungen erfüllt sein:

- Die Lichtquelle sollte sich ohne aufwendige Beschaltung mit einem herkömmlichen Pulsgenerator ansteuern lassen.
- Die von der Lichtquelle ausgesandten Pulse sollten eine Zeitdauer von ca. 10 ns oder weniger haben.
- Das Spektralprofil sollte in demselben Wellenlängenbereich wie das der auf der parametrischen Fluoreszenz basierenden Einphotonenquelle (s.o.) liegen.
- Die Lichtquelle sollte preisgünstig sein und einen kompakten Aufbau der Quelle erlauben.

Die Motivation hinter der Aufstellung dieser Bedingungen dürfte – bis auf die Forderung hinsichtlich des Spektralprofils – offensichtlich sein. Wie bereits in Abschnitt 4.2.1 erwähnt, ist es am günstigsten, wenn die Photonen, die bei der Zufallsgenerierung verwendet werden, in einem Wellenlängenbereich liegen, für den die Quanteneffizienz der Detektoren möglichst hoch ist. Anders als bei der echten Einphotonenquelle ist dies im Falle einer statistischen Einphotonenquelle auch leicht zu erfüllen, da es für den

[7] „Gleichzeitig“ bezieht sich dabei auf die Zeitauflösung der Detektoren bzw. der Signalverarbeitung; bei den vorgestellten Experimenten wird sie durch das Koinzidenzfenster von 10 ns gegeben.

[8] Welche Wahrscheinlichkeit man als hinreichend klein ansieht, ist natürlich Ermessenssache. Bei den Experimenten wird der Wert $P(n \geq 2) \approx 0{,}01$, der bei Experimenten zur Quantenkryptographie mit abgeschwächten Lichtpulsen [57, 104] auch gern verwendet wird, als Richtwert genommen. Dies entspricht einer mittleren Photonenrate von 0,1 Photon pro Puls.

Wellenlängenbereich in dem die Detektoren eine sehr gute Quanteneffizienz besitzen, s. Abb. 4.6, Laserdioden bzw. LEDs gibt.
Bei den in der vorliegenden Arbeit durchgeführten Experimenten mit statistischen Einphotonenquellen wird allerdings auf eine möglichst gute Vergleichbarkeit zwischen der statistischen Einphotonenquelle und der echten Einphotonenquelle Wert gelegt und dabei eine schlechtere Quanteneffizienz und eine somit niedrigere Bitrate in Kauf genommen, da *nicht* im Wellenlängenbereich maximaler Quanteneffizienz der Detektoren gearbeitet werden kann.
Für die Lichtquelle fiel die Wahl auf eine schnelle Signal-Leuchtdiode Typ Honeywell High Speed Fiber Optic LED HFE 4023-323 mit 150 MHz Bandbreite und einem Spektralprofil von 850 nm ± 50 nm. Die Diode hat eine Kugellinse direkt über dem Chip zur effizienten Einkopplung in Mehrmodenfasern, so daß aufgrund der geringen Strahldivergenz keine zusätzliche Kollimationslinse notwendig ist. Ein Interferenzfilter (Zentralwellenlänge 882 nm, Halbwertsbreite 8,6 nm) dient dazu, das Spektralprofil des Ausgangslichtfeldes auf den Zentralbereich der Wellenlängenverteilung bei der Einphotonenquelle einzuschränken.
Die Abschwächung der Lichtpulse erfolgt durch den Interferenzfilter und zwei zusätzlich hinter ihm in den Strahlengang eingebrachte Graugläser. Die LED und der Interferenzfilter befinden sich in einem einseitig abgeschlossenen Tubus, der durch die direkt am Ausgang angebrachten Graugläser abgeschlossen wird; die Quelle ist daher recht kompakt (ca. 11 cm Länge) und könnte überdies auch noch *erheblich* weiter verkleinert werden. Die Graugläser werden so gewählt, daß die mittlere Photonenzahl pro Puls nicht größer als 0,1 Photon pro Puls ist, zur genauen Vorgehensweise, s. Abschnitt 5.2.1. Das ist eine sehr restriktive Wahl, da die nachfolgenden optischen Komponenten des Zufall generierenden Elementes nochmals abschwächend wirken und die Detektoren, aufgrund ihrer nicht 100% betragenden Quanteneffizienz, die gemessene mittlere Photonenzahl weiter senken. Wegen ihres kompakten Aufbaus kann die statistische Einphotonenquelle direkt vor dem Zufall generierenden Element plaziert werden, so daß der gesamte Generatoraufbau kompakter gestaltet werden kann als im Fall der Einphotonenquelle auf Basis der parametrischen Fluoreszenz.
Die LED ist sehr einfach beschaltet: Über einen Vorwiderstand[9] werden Vor- und Pulsspannung angelegt, das einzige weitere Bauteil ist eine zur LED antiparallel geschaltete, schnelle Signaldiode, die dazu dient, Überschwinger in den positiven Spannungsbereich, die schädlich für die LED sein könnten, kurzzuschließen.
Die Ansteuerung der LED erfolgt mit einem handelsüblichen Pulsgenerator (Hewlett Packard, Typ 8007B), wobei die Parameter so gewählt werden, daß die von der LED ausgesandten Lichtpulse eine Länge von $\leq$ 10 ns haben[10]. Hierzu wird die LED mit -1 V vorgespannt und mit elektrischen Pulsen aus dem Pulsgenerator angesteuert, die eine Länge von 8,2 nm (FWHM) haben und eine Amplitude von -2 V. Die Vorspannung ist notwendig, da nur auf diese Weise hinreichend kurze Pulse erzeugt werden können. Negative Spannungen werden verwendet, damit das Signal des Pulsgenerators

[9] Es wurde anstatt 50 Ω ein Wert von 27 Ω für den Widerstand gewählt, um mit Hilfe eines höheren Stroms (50 mA bei einer Betriebsspannung der LED von 1,7 V) kürzere Lichtpulse zu erreichen.

[10] Typischer Wert für die Länge der Lichtpulse: 8,6 ns FWHM, kürzere Pulse sind prinzipiell möglich (bis ca. 5 ns), da die verwendete Koinzidenzelektronik allerdings ein minimales Koinzidenzfenster von 10 ns hat, wäre dies beim Laboraufbau nicht sinnvoll.

auch als Triggersignal[11] dienen kann, wobei allerdings ein Kondensator mit der Kapazität ($C_{AC} = 1$ nF) zur AC-Kopplung an die nachfolgende Signalverarbeitung verwendet werden muß, damit die Vorspannung nicht ebenfalls am Eingang anliegt. Mit Hilfe von Verzögerungsleitungen wird dafür gesorgt, daß das Triggersignal bei den Experimenten koinzident zu den elektrischen Ausgangspulsen der Signaldetektoren ist. Die Verzögerungen müssen hierbei aufgrund der unterschiedlichen optischen Wege bei Freistrahl- und Faseroptik jeweils für den entsprechenden Aufbau dimensioniert werden.

4.4 Das Zufall generierende Element

4.4.1 Aufbau in Freistrahloptik

Das Zufall generierende Element besteht beim Aufbau in Freistrahloptik, s. Abb. 4.2, Abb. 4.3, Abb. 4.4 und Abb. 4.5 aus drei Komponenten: einem Strahlteilerwürfel (falls polarisierend, mit einem zusätzlichen $\lambda/2$-Plättchen davor) und den beiden Quantendetektoren in den Ausgängen des Strahlteilers. Die Grundidee wurde schon in Abschnitt 3.1.1 erläutert, daher wird hier nur auf die experimentelle Realisierung eingegangen.

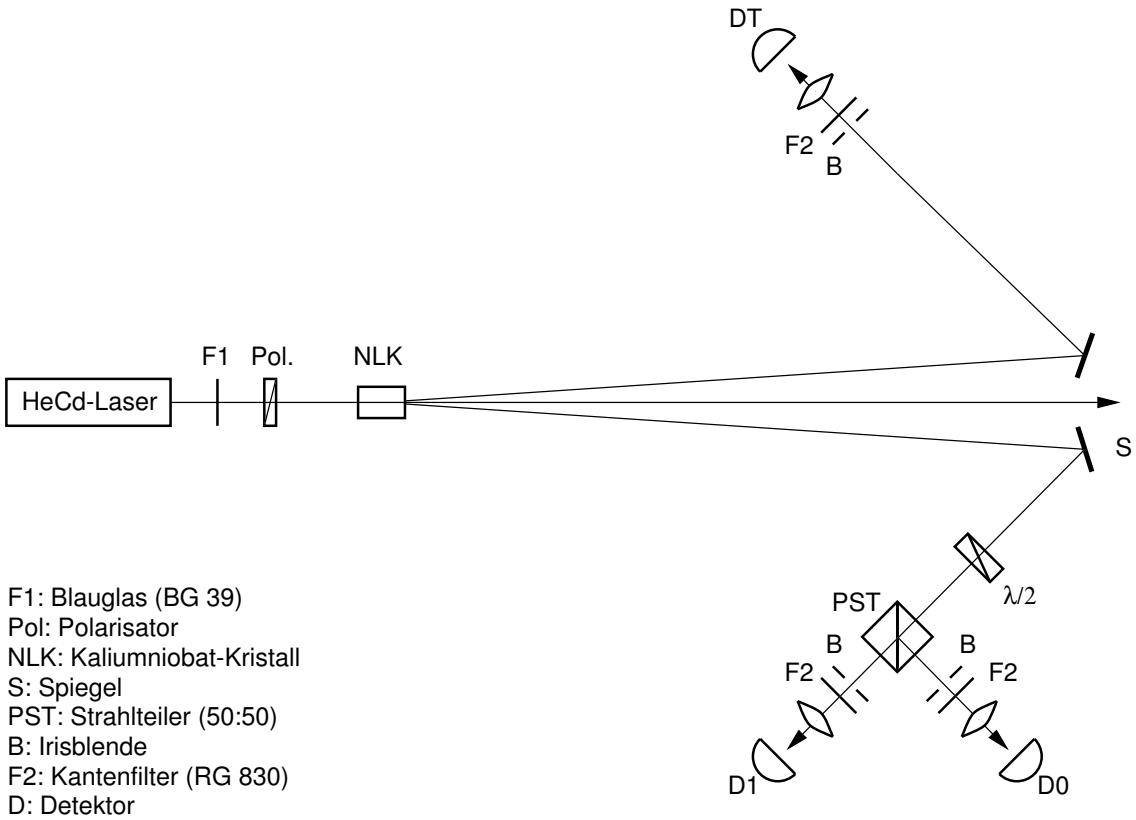


Abbildung 4.2: Freistrahloptischer Aufbau eines quantenoptischen Zufallsgenerators mit Kaliumniobat-Kristall, $\lambda/2$-Plättchen und polarisierenden Strahlteiler

[11] Bei den Experimenten mit der gepulsten, statistischen Einphotonenquelle werden für die Signalvorverarbeitung NIM-Einschübe benutzt, diese setzen negative Spannungen bei den Signalen voraus.

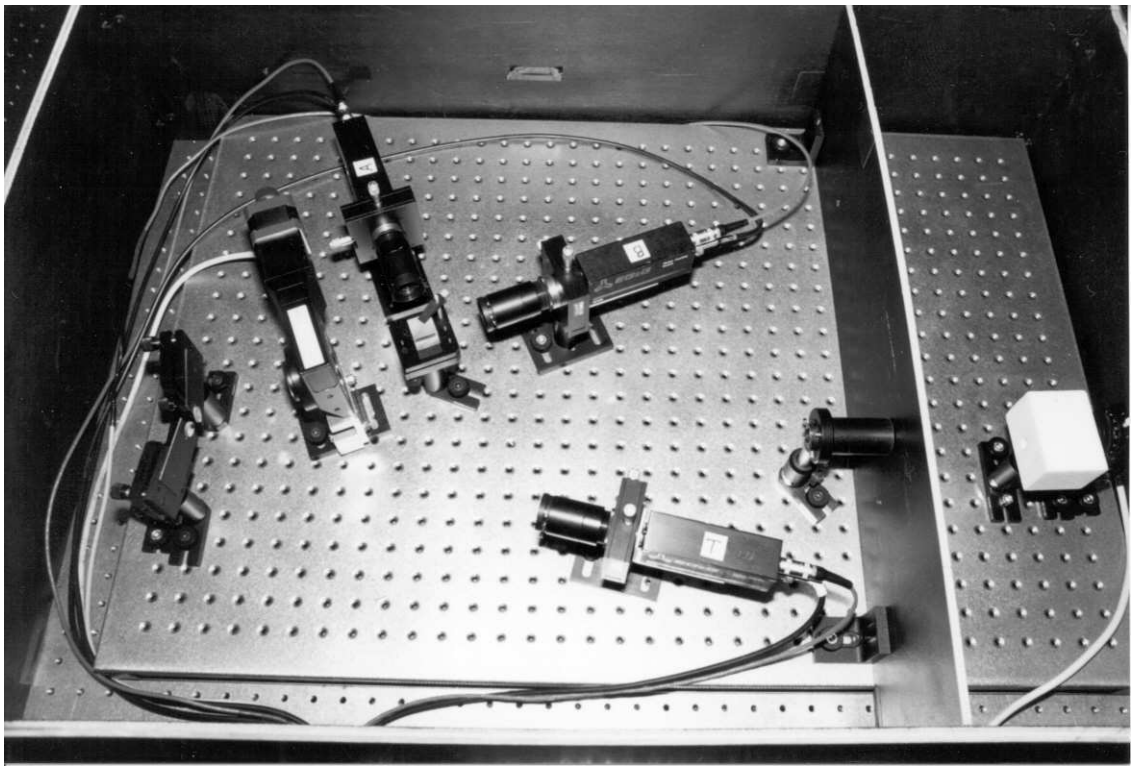


Abbildung 4.3: Photographie des freistrahloptischen Aufbaus eines quantenoptischen Zufallsgenerators mit Kaliumniobat-Kristall, $\lambda/2$-Plättchen und polarisierenden Strahlteiler

4.4.1.1 Der Strahlteilerwürfel

Der Strahlteilerwürfel ist der integrale Bestandteil des Zufall generierenden Elementes. Anfangs wurde für das Experiment ein polarisierender Strahlteiler (entspiegelter dielektrischer Breitband-Strahlteilerwürfel (650-900 nm) mit Absorption $< 0{,}5\%$, Melles Griot) benutzt. Polarisierende Strahlteiler haben eine hohe Transmission für parallel zur Einfallsebene polarisiertes Licht und eine niedrige für senkrecht zur Einfallsebene polarisiertes Licht. Umgekehrt wird das senkrecht polarisierte Licht sehr gut reflektiert, während das parallel polarisierte nahezu gar nicht reflektiert wird. Das Licht aus der Einphotonenquelle auf Basis der parametrischen Fluoreszenz ist linear polarisiert. Mit Hilfe eines drehbar gelagerten $\lambda/2$-Plättchens (Doppelplatte aus Quarz, für eine Lichtwellenlänge von $\lambda = 884$ nm gefertigt, breitbandentspiegelt, 25 mm Durchmesser, Spindler & Hoyer) läßt sich die lineare Polarisationsrichtung des Lichtes drehen. Dies ermöglicht es, die lineare Polarisationsrichtung des einfallenden Lichtes so zu drehen, daß das Strahlteilungsverhältnis genau 50:50 beträgt.
Auch mit dem später verwendeten unpolarisierenden Strahlteiler läßt sich aber ohne zusätzliche Einstellmöglichkeit ein Teilungsverhältnis von 48:52 erreichen. Die Abweichungen vom idealen Teilungsverhältnis rühren hierbei nicht nur von geringen wellenlängenbedingten Abweichungen im Teilungsverhältnis des Strahlteilerwürfels her, sondern auch von leichten Unterschieden in der Quanteneffizienz der Detektoren, des Durchmessers der Irisblenden vor den Detektoren und von minimalen Fehlstellungen der Detektoren. Diese

Fehlstellungen lassen sich bei den freistrahloptischen Aufbauten nur schwer vermeiden, da auf eine sehr kleine Detektorfläche (s.u.) fokussiert werden muß. Beim faseroptischen Aufbau hingegen, werden solche Fehlstellungen vermieden, da der Aufbau symmetrischer ist, allerdings weicht nun das Teilungsverhältnis des Faserkopplers selbst stärker von 50:50 ab, s. Abschnitt 4.4.2.

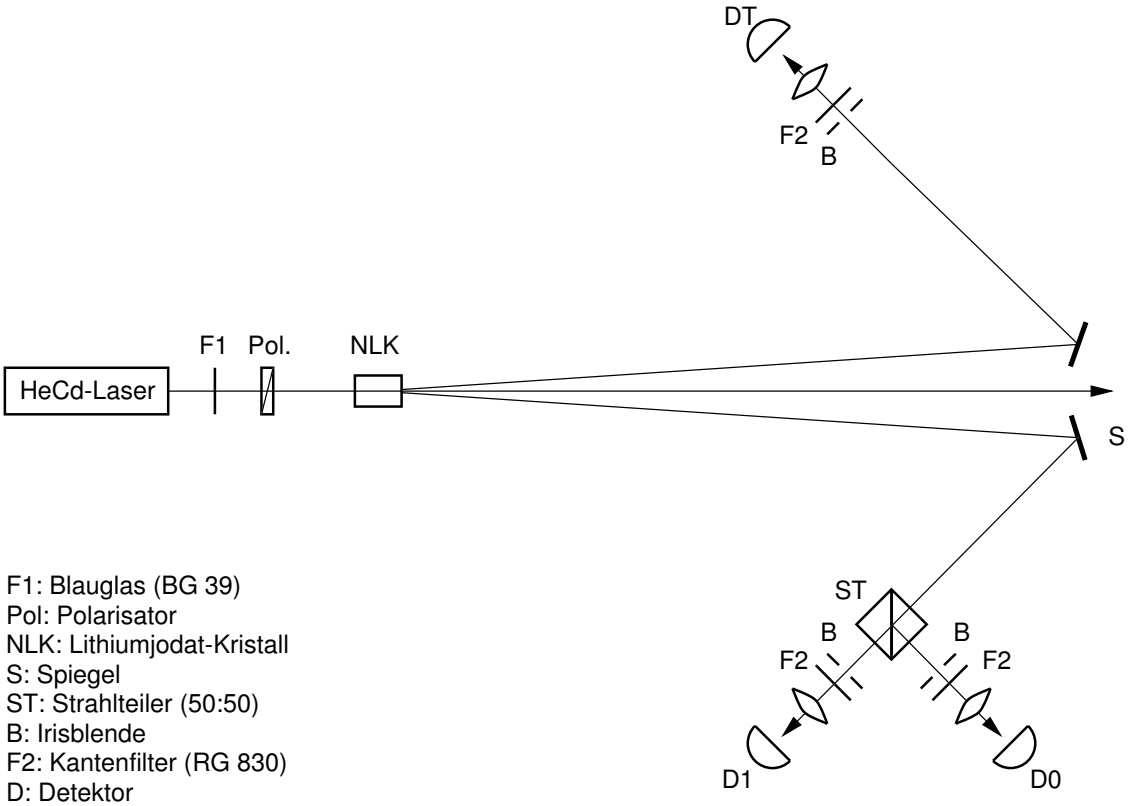
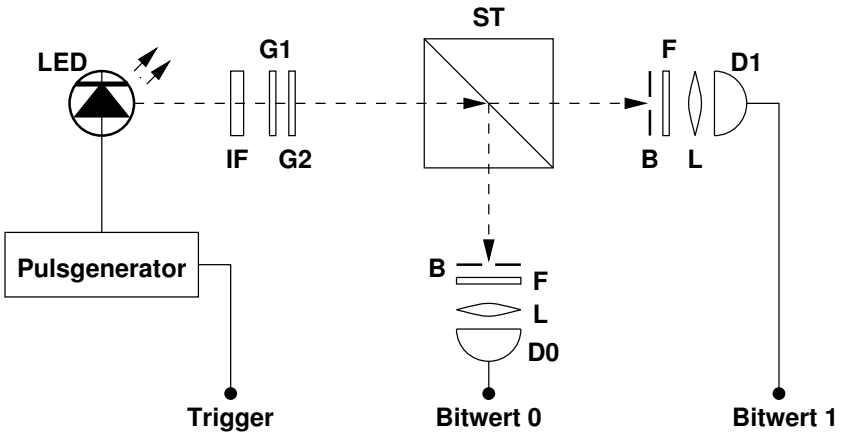


Abbildung 4.4: Freistrahloptischer Aufbau eines quantenoptischen Zufallsgenerators mit Lithiumjodat-Kristall und unpolarisierenden Strahlteiler

4.4.1.2 Die Detektoren

Die Detektoren befinden sich in den beiden Ausgängen des Strahlteilerwürfels. Ihr Abstand zu den Glasflächen des Strahlteilerwürfels beträgt ca. 4 cm. Die Detektoren bestehen aus einem Tubus mit optischen Komponenten und dahinterliegenden kommerziell erhältlichen Detektormodulen, die sich mit Hilfe einer XY-Verschiebeeinheit senkrecht zur Tubusachse verschieben lassen. Bei den optischen Komponenten, die sich im Tubus befinden, handelt es sich jeweils um eine Irisblende (Öffnung: 0,5 – 12 mm), einen Kantenfilter (RG 830, SCHOTT, entspiegelt)und eine kurzbrennweitige, achromatische Sammellinse (f = 30 mm, ebenfalls entspiegelt). Die Irisblende dient dazu den Raumwinkel, den die verschiedenen Detektoren sehen, so einzugrenzen, daß er für alle drei Detektoren gleich ist, auch wenn die Detektoren in unterschiedlichen Entfernungen vom



IF: Interferenzfilter
G: Grauglas
ST: Strahlteiler
B: Irisblende
F: Farbglasfilter (RG 830)
L: Linse
D: Detektor

Abbildung 4.5: Freistrahloptischer Aufbau eines quantenoptischen Zufallsgenerators mit gepulster, statistischer Einphotonenquelle

Kristall stehen[12]. Der Kantenfilter unterdrückt alles Licht mit einer Wellenlänge unter 830 nm, so z. B. Streulicht des Pumplasers (442 nm); er ist aufgrund der hohen Empfindlichkeit der Detektoren unumgänglich.
Der Abschnitt des Tubus, der sich zwischen Linse und Detektorhalterung befindet, ist in seiner Länge veränderlich und arretierbar, so daß sich auf diese Weise der Fokus der Linse longitudinal verstellen läßt. Die Möglichkeit, den Fokus zu verändern und das Detektormodul senkrecht zu Tubusachse zu verschieben, erlaubt es, das Licht auf die recht kleine Detektorfläche[13] (Durchmesser: 200 μm) zu fokussieren.

EG&G-Detektormodule Für die Funktion des quantenoptischen Zufallszahlengenerators sind Einzelquantendetektoren von größter Bedeutung. Ohne effiziente Detektoren ist die technische Realisierung eines auf quantenoptischen Prinzipien beruhenden Generators ausgeschlossen. Glücklicherweise gibt es seit einigen Jahren ausgesprochen effiziente Einzelquanten-Detektormodule der Firma EG&G. Diese Module zeichnen sich bei Betriebstemperaturen im Bereich von 10°–35° C durch hohe Quanteneffizienz, Rauscharmut (Dunkelzählrate < 100 Hz), gute Zeitauflösung, kompakte Bauweise und Robustheit aus. Überdies haben sie den großen Vorteil, daß sie einen zweistufigen On-Chip-Peltierkühler besitzen, eine externe aktive Kühlung also überflüssig ist, sofern für gute Luftzufuhr bzw. Wärmeableitung über das Gehäuse gesorgt wird, s. a. Abschnitt 3.2.3.

[12] Der Triggerdetektor befindet sich in einer geringfügig kleineren Entfernung vom Kristall als die beiden Signaldetektoren.

[13] Bei den neuen Detektoren wird eigentlich nicht auf die Detektorfläche, sondern auf die Faserzuleitung zum Chip fokussiert, diese hat allerdings einen noch kleineren Kerndurchmesser (100 μm).

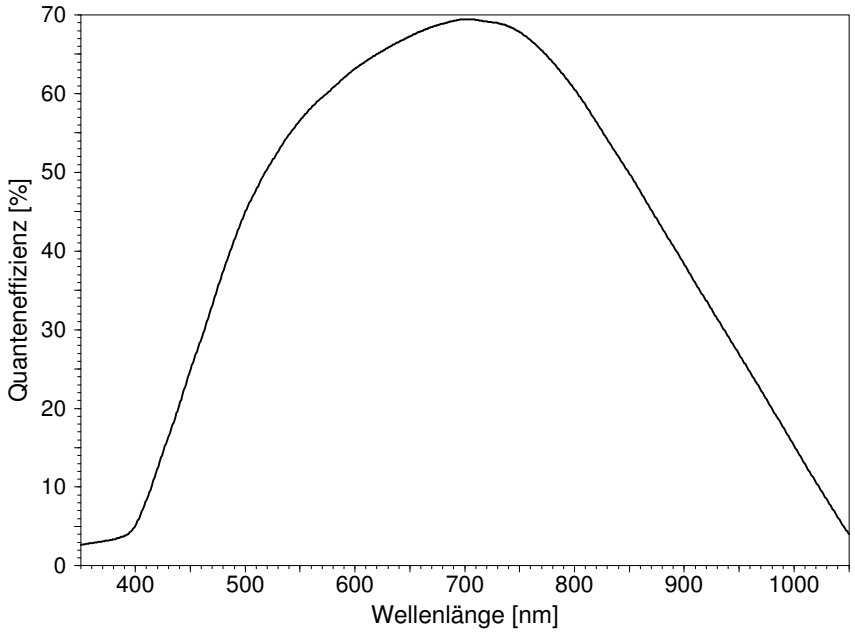


Abbildung 4.6: Quanteneffizienz der EG&G-Detektormodule, Typ: SPCM-AQ, in Abhängigkeit von der Detektionswellenlänge

Da es sich bei dem eigentlichen Detektorchip um eine Si-Avalanchephotodiode (APD) handelt, folgt die Quanteneffizienz den typischen Kurven für Si-Photodioden, einzig aufgrund der Kühlung ist eine Verschiebung des Quanteneffizienzmaximums zu kürzeren Wellenlängen hin zu beobachten, s. Abb. 4.6). In dem für die folgenden Experimente relevanten Wellenlängenbereich um 884 nm läßt sich die Wellenlängenabhängigkeit der Quanteneffizienz sehr gut durch eine abfallende Gerade approximieren.
Die verwendeten Avalanchedioden sind gezielt für die Detektion einzelner Photonen entwickelt worden und werden bei der Herstellung nach möglichst niedriger Dunkelzählrate selektiert. Zur Detektion einzelner Photonen werden die Avalanchedioden in einem speziellen Zustand, dem „Geiger-Modus“, betrieben, der sich dadurch auszeichnet, daß bei ihm die APDs mit einer Spannung, die 20 bis 40 Volt über ihrer Durchbruchsspannung liegt, vorgespannt werden. Ein innerhalb der Sperrschicht der APD absorbiertes Photon erzeugt ein Elektron-Loch-Paar und aufgrund des hohen elektrischen Feldes wird das Elektron[14] des Paares in einer Generationsschicht beschleunigt und kann weitere Elektronen ins Leitungsband anregen, die wiederum beschleunigt werden und weitere Elektronen anregen können. Auf diese Weise bildet sich eine Elektronenlawine aus, die aber anders als im herkömmlichen Betriebsmodus einer APD (Vorspannung $<$ Durchbruchsspannung) nicht von allein wieder abklingt, sondern erst dann, wenn die Spannung an der Diode unter die Durchbruchsspannung gefallen ist. Dies läßt sich auf zwei Weisen erreichen: *passiv*, durch einen in Reihe zur APD geschalteten, strombegrenzenden hochohmigen Widerstand R, über den sich die Sperrschichtkapazität C_s entlädt oder *aktiv*, indem gleich bei Einsetzen des Lawinenstroms die Vorspannung sofort unter die Durch-

[14] Die Löcher tragen wenig zum Aufbau der Lawine bei, da die Dotierungsprofile der APDs so angelegt sind, daß ihre Ionisationskoeffizienten sehr viel kleiner als die der Elektronen sind.

bruchsspannung gesenkt wird. Die aktive Vorgehensweise ist empfehlenswert, da man anders als im passiven Fall nicht mehrere Zeitkonstanten $\tau = C_s \cdot R$ lang warten muß, bevor die Sperrschichtkapazität wieder aufgeladen werden kann. Daher weisen Detektoren, die eine Beschaltung für dieses „active quenching“ genannte Vorgehen aufweisen, eine kürzere Totzeit auf, so daß ihre Detektionseffizienz höher ist und sich Sättigungseffekte erst bei höheren Zählraten bemerkbar machen.
Die Detektor-Module enthalten bereits die komplette Elektronik, die für die Ansteuerung der APDs nötig ist. Aufgrund ihres Metallgehäuses und der hohen Ausgangspegel (TTL-Pegel) sind sie sehr störunempfindlich gegenüber elektromagnetischer Beeinflussung von außen.
Bei den experimentellen Aufbauten der quantenoptischen Zufallsgeneratoren finden zwei verschiedene Modultypen ihre Anwendung[15]. Für den Triggerdetektor wird ein Detektor Typ EG&G SPCM-200-PQ verwendet, seine Dunkelzählrate beträgt weniger als 100 Zählereignisse pro Sekunde, die Zeitauflösung beträgt 500 ps FWHM, und die elektrischen Ausgangspulse haben eine Länge von 300 ns.
Bei diesem Modul handelt sich um ein älteres Modell, bei dem die APD „passiv gequencht“ wird und das daher eine Detektor-Totzeit von 200 ns aufweist. Aufgrund der relativ langen Totzeit geht der Detektor bei hohen Photonenraten in die Sättigung, d.h. er zählt einen beträchtlichen Anteil der Photonen, die absorbiert werden, nicht mehr, da sie während der Totzeit auf den Detektor auftreffen. Die längere Totzeit ist neben der kleinen Pumpleistung und der mit ihr einhergehenden niedrigen Photonenpaar-Rate ein limitierendes Element für die maximal mögliche Bitrate des Zufallsgenerators. Die Quanteneffizienz dieses Detektors beträgt bei 884 nm ca. $\eta_{pq} = 0{,}23$.
Die beiden Detektoren in den Ausgängen des Strahlteilers sind Detektoren des Typs EG&G SPCM-AQ-131-FL. Sie haben eine Dunkelzählrate von weniger als 100 Zählereignisse pro Sekunde, eine Zeitauflösung von 300 ps FWHM, eine Ausgangspulslänge von 10 ns und eine Totzeit[16] von 50 ns. Aufgrund dieser kürzeren Totzeit, die durch ein „aktives Quenching“ der Elektronen-Lawine erreicht wird, zeigen sich Sättigungserscheinungen erst bei höheren Photonenraten.
Eine negative Eigenschaft der Detektoren (alte wie neue Modelle) ist ihre Tendenz, daß mit einer gewissen Wahrscheinlichkeit nach einem Detektionsereignis zusätzliche Nachpulse auftreten. Allerdings ist die mittlere Nachpulswahrscheinlichkeit recht klein, sie beträgt für die verwendeten Detektoren $P_{Nachpuls} = 3 \cdot 10^{-3}$ bzw. $P_{Nachpuls} = 2 \cdot 10^{-3}$. Zudem ist die Verteilung der Nachpulse stark zeitabhängig; sie hat einen nahezu exponentiell abfallenden Verlauf, wobei die weitaus meisten Nachpulse im Zeitraum bis 200 ns nach der Detektion auftreten. Da für den Zufallsgenerator nur *Koinzidenzereignisse* zwischen zwei Detektoren[17] relevant sind, wird der Einfluß des Nachpulsens noch weiter reduziert. Beim Aufbau mit einer Einphotonenquelle ist der Einfluß noch geringer, da der Triggerdetektor mit seiner längeren Totzeit (200 ns) während der Zeitspanne, in der die Nachpulswahrscheinlichkeit der Signaldetektoren am größten ist, ohnehin keine Pulse liefern und es somit zu keinem Koinzidenzereignis kommen kann.

[15] Es standen nur zwei der neueren Detektoren des Typs EG&G SPCM-AQ-131-F zur Verfügung.

[16] Die in den Meßprotokollen der verwendeten Signaldetektoren für die Totzeit aufgeführten Werte sind mit 32 und 27 ns sogar noch kürzer.

[17] Im Falle der Photonenquelle, die stark abgeschwächte Pulse aussendet, handelt es sich um Koinzidenzereignisse zwischen dem elektrischen Ansteuerpuls der LED und einem Detektorsignal.

Die neuen Detektor-Modelle haben jeweils eine FC-Faseranschlußbuchse, was Experimente mit faseroptischen Bauteilen (z.B. Faser-3dB-Koppler statt eines Strahlteilerwürfels, s. Abschnitt 4.4.2) stark vereinfacht[18] und diverse Komponenten spart, wie z.B. die x-y-Verstelleinheiten. Allerdings wird durch die notwendig Ankopplung des FC-Faseranschlusses an die Detektoroberfläche die Quanteneffizienz etwas vermindert, sie beträgt bei 884 nm für die verwendeten Detektoren $\eta_{aq} = 0{,}31$ bzw. 0,29.

4.4.2 Aufbau in Faseroptik

Der Aufbau mit faseroptischen Komponenten, s. Abb. 4.7 bzw. 4.8, unterscheidet sich vom Aufbau in Freistrahloptik lediglich durch ein anderes Zufall generierendes Element: Anstatt eines Strahlteilerwürfels mit Detektoren in den Ausgangsarmen wird ein Mehrmoden-Faserkoppler mit an den Ausgangsfasern angeschlossenen Detektoren verwendet. Die Wahl fiel dabei auf einen Mehrmoden- und nicht auf einen Einmoden-Faserkoppler[19], da aufgrund des wesentlich größeren Kerndurchmessers von Mehrmodenfasern das Einkoppeln des Signalphotons in die Faser erheblich effizienter möglich ist als bei einer Einmoden-Glasfaser [128].

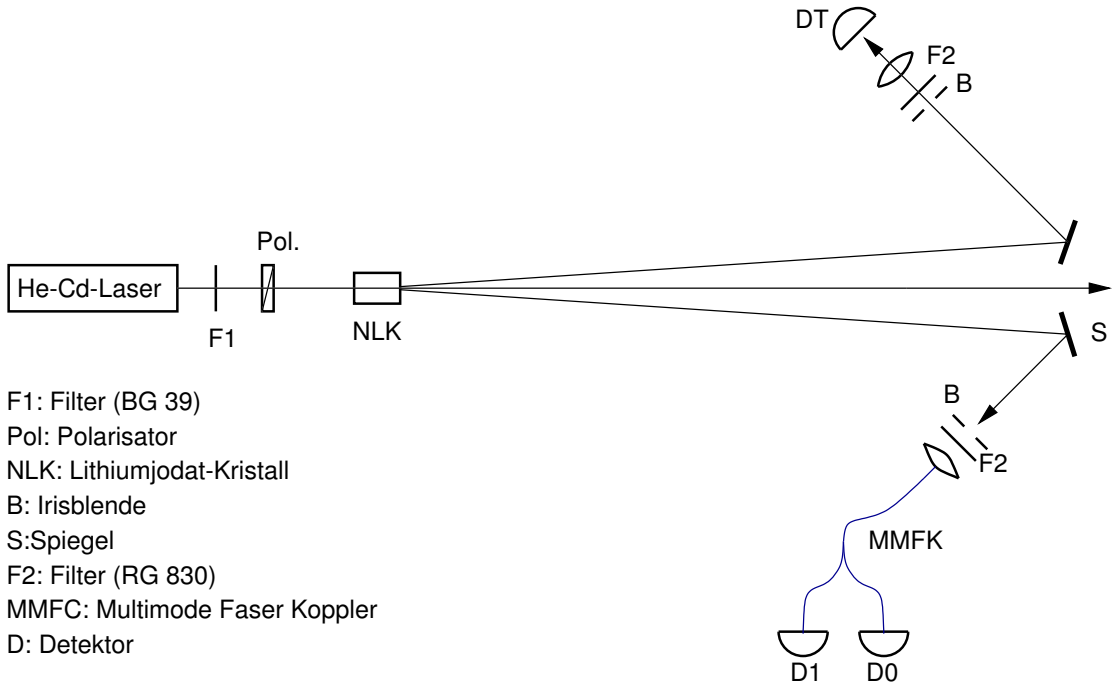
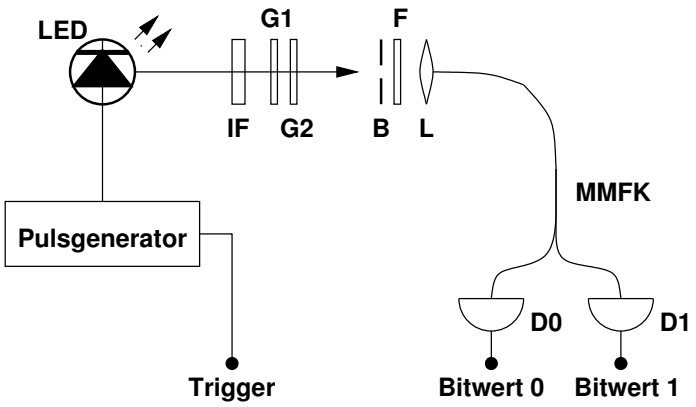


Abbildung 4.7: Faseroptischer Aufbau eines quantenoptischen Zufallsgenerators mit Einphotonenquelle

[18] Allerdings erschweren die Anschlüsse die Justage des Aufbaus mit Freistrahloptik erheblich, da sie für solch eine Verwendung eigentlich nicht vorgesehen sind und das Fokussieren in die Faseranschlüsse relativ schwierig ist. Selbstverständlich gibt es diese Detektoren auch als Modell ohne Faseranschlüsse, solche Detektoren standen aber nicht zur Verfügung.

[19] Einen Einmoden-Faserkoppler hätte man für die ungewöhnliche (Zentral-) Wellenlänge speziell fertigen lassen müssen.



IF: Interferenzfilter
G: Grauglas
ST: Strahlteiler
B: Irisblende
F: Farbglasfilter (RG 830)
L: Linse
D: Detektor
MMFK:Multimode Faserkoppler

Abbildung 4.8: Faseroptischer Aufbau eines quantenoptischen Zufallsgenerators mit gepulster, statistischer Einphotonenquelle

Bei dem verwendeten Mehrmoden-Faserkoppler handelt es sich um einen 1 auf 2 Koppler[20] Model 5202 der Firma FOD, mit einer 1 m langen Eingangsfaser und zwei 30 cm langen Ausgangsfasern, die alle mit FC-Steckern konfektioniert sind. Der Koppler ist für den Wellenlängenbereich von 850 bis 1300 nm konstruiert, wobei allerdings das mit 3 ± 1 dB angegebene Teilungsverhältnis im Bereich um 884 nm bei typisch 40:60 liegt, was leider erheblich[21] vom idealen Teilungsverhältnis von 50:50 abweicht. Die Zusatzverluste des Kopplers werden von FOD mit kleiner als 1 dB spezifiziert.
Zum Einkoppeln des Signalphotons bzw. der Lichtpulse der LED in die Eingangsfaser des Faserkopplers dient ein Fokussierkollimator mit achromatischem Objektiv (Spindler & Hoyer, MB 04), der eine ganze Reihe von Freiheitsgraden bietet, um die Einkopplungseffizienz zu erhöhen. Neben der longitudinalen Verstellbarkeit des Fokus, kann der Fokussierkollimator senkrecht zur Strahleinfallsrichtung verkippt und mit Hilfe einer zusätzlichen x-y-Verschiebeeinheit überdies senkrecht zur Einfallsrichtung bewegt werden.
Die Ausgangsfasern des Faserkopplers werden direkt an die mit FC-Faseranschlußbuchsen versehenen Detektoren angeschlossen. Gegenüber dem Aufbau in Freistrahloptik reduziert dies den Einfluß des ohnehin geringen Streulichtes noch weiter. Hierdurch verbessert sich auch die Symmetrie zwischen den Signaldetektoren, da nicht mehr auf

[20] Natürlich handelt es sich eigentlich um einen 2 auf 2 Koppler, während aber beide Ausgangsfasern aus dem Gehäuse des Kopplers herausgeführt sind, gilt dies nur für eine der Eingangsfasern; die andere ist intern abgeschlossen.

[21] Es ließe sich im Prinzip natürlich auch ein speziell für den gewünschten Wellenlängenbereich optimierter Faserkoppler mit einem ausgewogenen Teilungsverhältnis fertigen.

jede Detektorfläche einzeln justiert werden muß, was ebenfalls eine große Zeitersparnis darstellt.
Die Fasern zeigen in Bezug auf beim Aufbau auftretende Biegungen (Krümmungsradien > 10 cm) kein die Detektorzählraten beeinflussendes Verhalten.
Da die Ausgangsfasern des Faserkopplers direkt an die Detektoren angeschlossen sind, müssen die Irisblende und der nachfolgende Kantenfilter, die sich beim Aufbau in Freistrahloptik im Tubus vor dem jeweiligen Detektor befinden (s. Abschnitt 4.4.1.2), jetzt in einem Tubus direkt vor der Fasereinkopplungsoptik plaziert werden. Dies bietet einen großen Vorteil: Es läßt sich nämlich damit sicherstellen, daß beide Detektoren auf jeden Fall denselben Raumwinkel „sehen“. Insbesondere für den mit abgeschwächten Lichtpulsen arbeitenden Zufallsgenerator ist dies wichtig, s. Abschnitt 5.2.3.

4.5 Die Signalverarbeitung und Datenaufnahme

Bei den Laboraufbauten wurden im Laufe der Arbeit drei verschiedene Varianten der Signalverarbeitungs- und Datenaufnahme-Elektronik verwendet:

1. Eine auf NIM-Einschüben (von EG&G) basierende Elektronik, gefolgt von einer einfachen Schnittstellen-Elektronik zur Datenaufnahme mit Hilfe eines Meßrechners.

2. Eine kompakte, kostengünstige Elektronik, die Signalverarbeitung, programmierbare Datenaufnahmemöglichkeiten und einen Zwischenspeicher kombiniert und eine effiziente Datenübertragung zum Meßrechner erlaubt.

3. Eine hybride Elektronik, bei der die Pulsformung und die Koinzidenzbildung mit Hilfe von NIM-Einschüben erfolgt, aber die Bitübernahme, Zwischenspeicherung und Datenübertragung zum Rechner durch Teile der bei Punkt 2 genannten Elektronik.

Hierbei stellt die kompakte Elektronik eine Eigenentwicklung dar, die im Hinblick auf die praktische Anwendung quantenoptischer Zufallsgeneratoren entworfen wurde. Die Gemeinsamkeiten und Unterschiede der drei Varianten der Datenaufnahme-Elektronik werden weiter unten besprochen.

4.5.1 Auf NIM-Einschüben basierende Signalverarbeitungselektronik

In Abb. 4.9 ist die für die Verarbeitung der Signale notwendige Elektronik im Prinzipschaltbild dargestellt.
Die NIM-Einschübe verwenden negative Logikpegel (kompatibel zu ECL-Logikpegeln), um möglichst kurze Verarbeitungszeiten zu erreichen. Dies bedingt eine Umwandlung der Detektorsignale mit ihren TTL-Pegeln in Signale mit ECL-Pegeln. Beim Generator mit Einphotonenquelle muß das Signal des Triggerdetektors (s. Abschnitt 4.4.1.2) auf 10 ns verkürzt[22] werden, damit man eine sinnvolle Koinzidenz zwischen dem Triggerdetektor und den Signaldetektoren erhält. Versäumte man dies, nähme die Rate der zufälligen Koinzidenzen unnötigerweise beträchtlich zu.

[22] Hierzu wird einer der vier Diskriminatoren aus einem vierfach Diskriminator, EG&G Typ 924 verwendet.

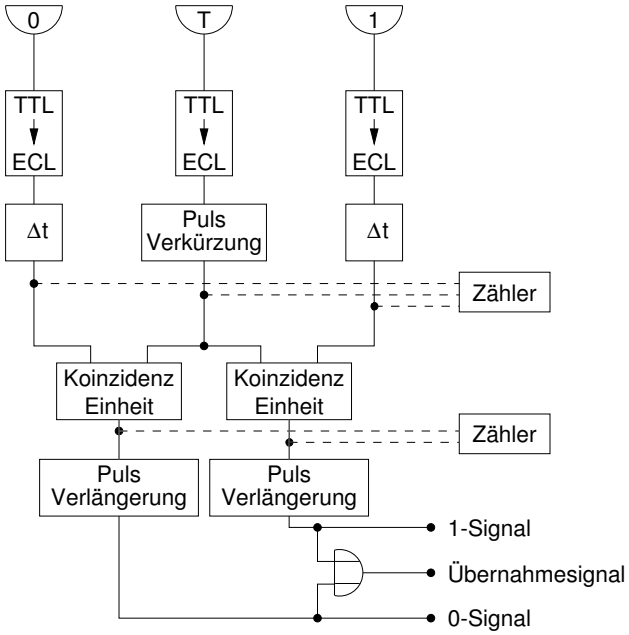


Abbildung 4.9: Prinzipschaltbild der Signalverarbeitungselektronik

Aufgrund der unterschiedlichen internen Verschaltung der alten und der neuen Detektormodule und der zusätzlichen Verzögerung, die der Triggerdetektorpuls bei der Verkürzung erfährt, ist es außerdem notwendig, die Signaldetektorsignale über kompensierende Verzögerungsleitungen den Koinzidenzeinheiten (EG&G, Modell C314/NL) zuzuführen. Bei den experimentellen Aufbauten mit der statistischen Einphotonenquelle wird kein Triggerdetektor, sondern nur ein bereits hinreichend kurzes elektronisches Triggersignal verwendet, so daß in diesem Fall nur das Triggersignal mit Hilfe einer Verzögerungsleitungen in Koinzidenz mit den Signalen der Detektoren gebracht wird.
Die Ausgangssignale der Koinzidenzeinheiten stellen die eigentlichen Meßsignale dar, d.h. je nachdem, welche der Koinzidenzeinheiten einen Ausgangspuls liefert, generiert der Zufallsgenerator einen Bitwert Eins oder Null. Da die Ausgangspulse aus den Koinzidenzen lediglich 10 ns lang sind, ist es nötig, sie für die direkte Datenübertragung zum Meßrechner wieder zu verlängern. Als Meßrechner wurde anfangs ein relativ langsamer Standard-PC (Prozessor: Intel 80386, 20 MHz Takt) genutzt[23], so daß aufgrund der interruptgesteuerten Datenaufnahme-Software (s. u. Abschnitt 4.5.1.1) allein dieser Umstand[24] bereits längere Pulse verlangte.
Die Ausgänge der Koinzidenzeinheiten können außerdem über ECL-TTL-Wandlungsgatter an eine im Meßrechner steckende Zählerkarte angeschlossen werden, so daß sowohl die Einzelzählraten der Detektoren als auch die Koinzidenzzählraten ermittelt werden

[23]Bei den Messungen mit der weiter unten beschriebenen kompakten Datenaufnahme-Elektronik wurde ein leistungsfähigerer Meßrechner (Prozessor Intel 80486, 50 MHz Takt) eingesetzt. Auch bei späteren Kontrollmessungen wurde dieser Rechner verwendet.

[24]Das Problem liegt hierbei nicht beim Übergabesignal, das den Hardware-Interrupt auslöst, sondern bei der Dauer der Signalpulse, denn ihre HIGH-Pegel müssen auch dann noch am Eingang des Rechners anliegen, wenn der Rechner den Zustand der Eingangsports schließlich einliest.

können. Dies ist insbesondere für die Justage wichtig; so optimiert man z.B. die Detektorpositionen durch eine Maximierung der jeweiligen Koinzidenzzählraten (unter Abzug der zufälligen Koinzidenzen).
Die Ausgangssignale der Koinzidenzeinheiten werden zuerst jeweils mit einer Gate&Delay-Einheit[25] (EG&G, Typ 416 A) auf eine Länge von 4 μs gebracht[26]. Hierbei werden gleichzeitig die negativen Pegel in positive (TTL-)Logikpegel umwandelt. Mit einem Monoflop (74HCT221) je Signalleitung werden die Pulse weiter verlängert auf 100 μs[27]. Aus diesen beiden Signalen wird zusätzlich mit Hilfe eines Oder-Gatters (74HCT32) ein Übernahmesignal generiert, das dem Meßrechner anzeigt, wann er ein Bit einlesen soll[28].
Die Bits werden über den Parallel-Port eingelesen, wobei das Signal für eine logische Eins auf Pin 13 (Printer Select) und das für eine logische Null auf Pin 12 (Paper End) gelegt wird. Das Übernahmesignal liegt an Pin 10 (Acknowledge) und kann einen Hardware-Interrupt (Interrupt 7, der „Printer-Interrupt“) auslösen. Diese Art des Einlesens wurde gewählt, damit der Generator an einen beliebigen PC-Meßrechner angeschlossen werden kann. Auch die Stromversorgung für die Monoflops und das Oder-Gatter geschieht dabei über die parallele Schnittstelle.

4.5.1.1 Datenaufnahme-Software

Der Meß-PC läuft unter dem „Betriebssystem“ MSDOS und die Datenerfassungssoftware beschränkt sich auf ein Minimum. Das Erfassungsprogramm besteht im wesentlichen aus drei Interrupt-Routinen (eine für jeden Aufnahmemodus, s. Anhang F), Routinen für Datenspeicherung und Dateibenennung und einer Routine für die (abschaltbare) Bildschirmdarstellung der wichtigsten Parameter. Alle Routinen wurden in C++ programmiert. Auf die Programm-Interna soll hier nicht detailliert eingegangen werden, es sei nur angemerkt, daß die Datenaufnahme interruptgesteuert geschieht, wobei das Programm hierzu in eine Interrupt-Routine springt, in der das Zufallsbit an die nächstfolgende Bitposition in einen 32 kB großen Pufferspeicher geschrieben wird.
Durch eine Reihe von Kommandozeilen-Parametern kann der Datenaufnahme-Modus, die Dateilänge und die Anzahl der aufzunehmenden Zufallsdateien angegeben werden; auch kann man festgelegen, ob eine (alphanumerische) Bildschirmdarstellung der elementaren statistischen Eigenschaften des Laufes und der Parameter erfolgen soll, für eine genaue Beschreibung der Parameter, s. Anhang F.

4.5.2 Kompakte Signalverarbeitungs- und Datenaufnahme-Elektronik

Die im vorhergehenden Abschnitt beschriebene Signalverarbeitungselektronik bewährt sich in den Experimenten zwar gut, aber sie hat auch einige Nachteile: NIM-Einschübe sind relativ teuer, groß und leistungszehrend. Gerade im Hinblick auf einen Einsatz des quantenoptischen Zufallsgenerators in der Praxis wurde daher eine kleinere, preisgün-

[25] Verwendet wird nur die Möglichkeit die Signaldauer zu verlängern.
[26] Der für die Pulsverlängerung verwendete Einschub kann nur Pulse mit einer maximalen Länge ca. 4,6 μs erzeugen.
[27] In der Abb. 4.9 sind die beiden Pulsverlängerungsstufen nur zusammengefaßt dargestellt.
[28] Wenn mindestens eines der Signale auf HIGH ist, ist das Übernahme-Signal ebenfalls HIGH. Das Oder-Gatter hat nichts mit der Von-Neumann-Regulierung zu tun!

stigere und energiesparende Signalverarbeitungs- und Datenaufnahme-Elektronik entwickelt und gebaut.
Auch die kompakte Elektronik muß in der Lage sein, die Pulslängen der Detektorpulse einander anzugleichen und Koinzidenzen zwischen Trigger- und dem jeweiligen Signaldetektor registrieren zu können. Zusätzlich wurde noch eine Zwischenspeicherungslogik eingebaut, die den Datentransfer zwischen Zufallsgenerator und Meßrechner vereinfachen und so den Meßrechner entlasten soll. Dies ist insbesondere für den praktischen Einsatz des Zufallsgenerators wichtig, da der Zufallsgenerator nur eine „Zulieferfunktion“ für einen Rechner hat, der innerhalb seiner Programmbearbeitung Zufallsdaten benötigt. Doch selbst wenn ein spezieller Rechner lediglich für das Aufnehmen von Zufallsdaten abgestellt würde, wäre es dennoch vorteilhaft, die Datenübertragung zwischen Zufallsgenerator und Rechner möglichst effizient zu gestalten, da der Rechner dann z.B. auch mehr Zeit für den Einsatz effizienter Regularisierungsverfahren hätte, s. Anhang G.
An die kompakte Datenaufnahmeelektronik werden folgende Anforderungen gestellt:

- Sie soll ohne schwer zu beschaffende und teure Spezialbauteile aufgebaut sein, damit im Falle eines Defektes eine schnelle und preisgünstige Reparatur möglich ist. Überdies sollte die Stromversorgung im gleichen Gehäuse untergebracht sein, um auf ein zusätzliches Netzteil verzichten zu können.

- Die einlaufenden Pulse sollen alle auf gleiche Länge gebracht werden und die beiden Koinzidenzeinheiten sollen möglichst kleine (< 10 ns) Koinzidenzfenster aufweisen, um den Prozentsatz der zufälligen Koinzidenzen möglichst niedrig zu halten.

- Nur wenn *genau* einer der beiden Signaldetektoren mit dem Triggerdetektor in Koinzidenz anspricht, soll ein Zufallsbit weitergegeben werden.

- Um den nachgeschalteten Rechner zu entlasten, sollen die Zufallsbits in einem Puffer zwischengespeichert werden, um bei Bedarf vom Rechner byteweise in größeren Blöcken ausgelesen werden zu können. Der Rechner muß dann nicht mehr auf jedes Ereignis interruptgesteuert reagieren.

- Der Pufferspeicher soll mit der Pulsformungs- und Koinzidenzelektronik über einen programmierbaren Baustein verbunden sein, der auch die Implementierung *einfacher* Regularisierungsalgorithmen erlaubt.

Das entsprechende Blockschaltbild für eine solche Datenaufnahmeelektronik ist in Abb. 4.10 dargestellt, in den folgenden Unterkapiteln wird der genaue Aufbau näher besprochen.

4.5.2.1 Pulsformung und Koinzidenz

Da die verwendeten Detektoren Pulse mit TTL-Pegeln ausgeben, wird für die integrierten Bausteine, die für Pulsformung und Koinzidenz verwendet werden, eine TTL-kompatible Logikfamilie verwendet, da sonst noch eine zusätzliche Umwandlungsstufe notwendig wäre; dies hat überdies den Vorteil, daß sich die (Standard-)Bauteile leichter beschaffen lassen. Die Wahl fiel auf die Advanced Speed CMOS Logikfamilie, da

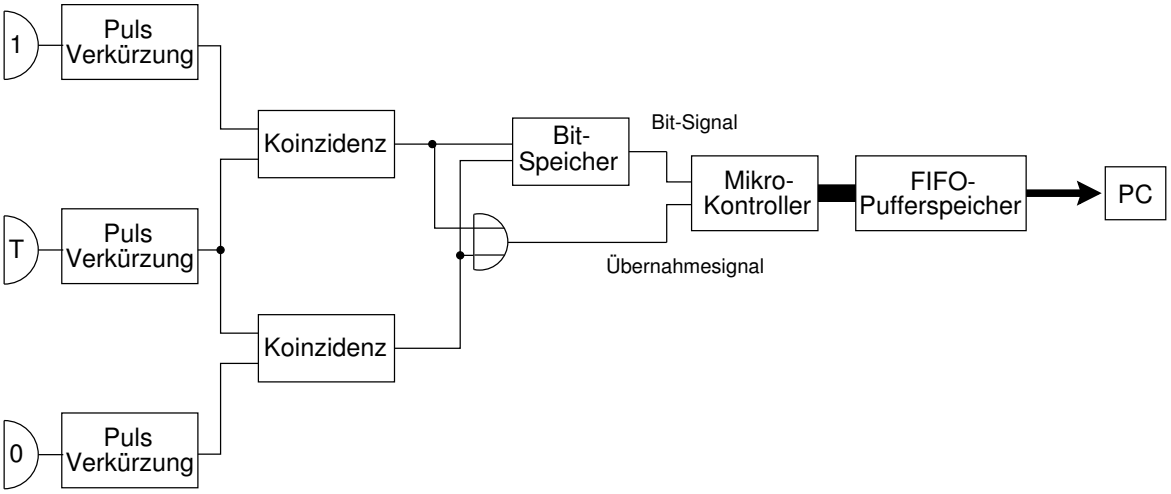


Abbildung 4.10: Blockschaltbild der kompakten Datenaufnahmeelektronik

sie besagte TTL-Pegel-Kompatibilität mit niedriger Leistungsaufnahme[29] und schnellem Schaltverhalten vereint.
Wie bereits erwähnt wird beim Einphotonen-Generator für den Triggerdetektor ein etwas älterer Typ von Detektor-Modul verwendet, dessen TTL-Ausgangspulse eine Länge von 300 ns haben und daher für eine sinnvolle Koinzidenz mit den nur 10 ns langen Pulsen der beiden Signaldetektormodule verkürzt werden müssen. Um Laufzeitunterschiede zwischen dem Trigger- und den Signaldetektorsignalen zu vermeiden, durchlaufen die Pulse der Signaldetektoren die gleiche Pulsverkürzung[30]. Diese Pulsverkürzung (s. Abb.4.11) wird schaltungstechnisch mit einer Kombination aus zwei Negationsgattern und einem Differenzierer realisiert. Das erste Gatter dient der Pulsformung, ein R-C-Glied differenziert den Puls, so daß nach Durchlaufen nur kurze Nadelpulse an den Pulsflanken übrig bleiben. Der positive Nadelpuls wird über eine Signaldiode abgeleitet, um ein Überschreiten des HIGH-Pegels am Eingang des zweiten Negationsgatters zu vermeiden; der negative wird nach Durchlaufen des zweiten Gatters zu einem positiven Puls kurzer Länge. Nachdem die verschiedenen Detektorsignale auf gleiche Länge (ca. 10 ns) gebracht sind, wird mit Hilfe von NAND-Gattern auf Koinzidenz getestet. Hierbei wird das Triggersignal (einlaufend an Buchse Bu2) an jeweils einen der Eingänge der beiden NAND-Gatter gelegt, und an den jeweils anderen das entsprechende Signaldetektor-Signal. An den Ausgängen der NAND-Gatter liegt somit immer ein HIGH-Pegel, es sei denn die Eingänge liegen (kurzzeitig) beide auf HIGH, so daß der Ausgangspegel des Gatters auf LOW liegt. Die Ausgänge dieser einfachen Koinzidenzeinheiten sind zum einen mit den Eingängen eines aus zwei NAND-Gattern aufgebauten Flipflops verbunden, und zum anderen mit den Eingängen eines XOR-Gatters. Einer der beiden Ausgänge des Flipflops ist mit dem Eingangs-Port PB1 eines Mikrocontollers *Programmable Interrupt Controller, PIC* verbunden. Das Flipflop dient zur Zwischenspeicherung der kurzen Koinzidenzsignale, da der nachfolgende PIC zu langsam (s.u.) ist, um Signale,

[29] Im Gegensatz zu Fast-TTL-Bausteinen, die eine verhältnismäßig hohe Leistungsaufnahme haben.
[30] Diese führt allerdings zu keiner merklichen Verkürzung, da die Pulse mit 10 ns ohnehin schon kurz genug sind.

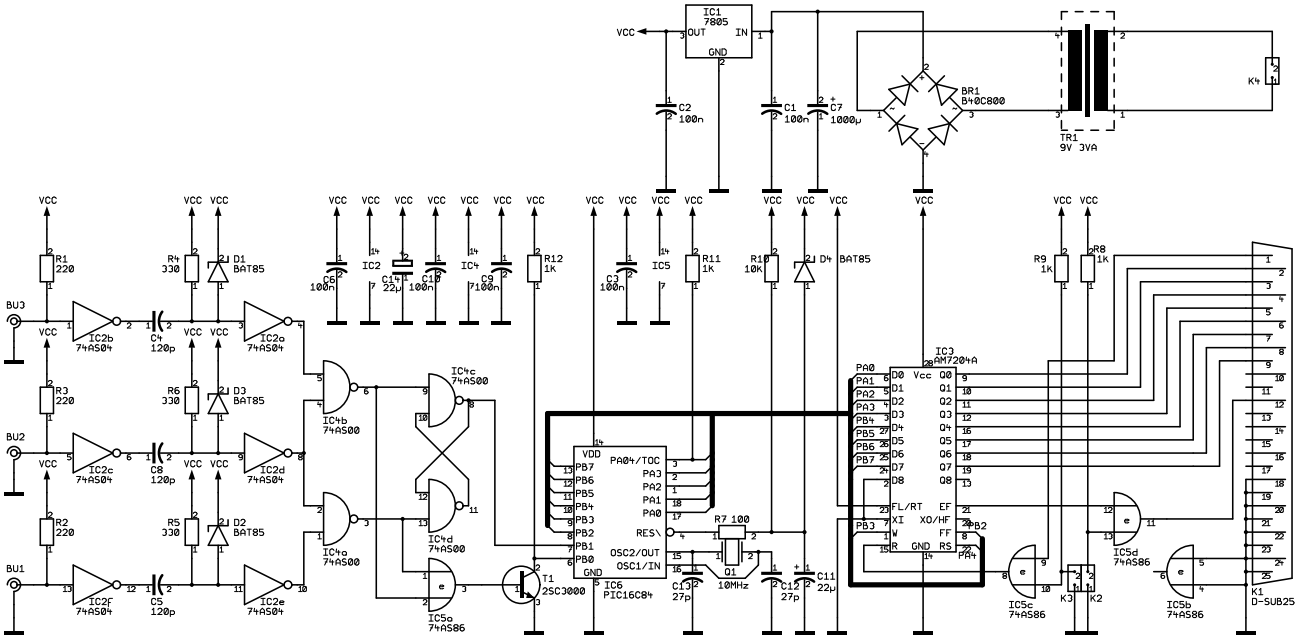


Abbildung 4.11: Schaltplan der kompakten Datenaufnahme-Elektronik

die nur 10 ns lang sind, an seinen Eingängen zu registrieren.
Das XOR-Gatter dient zu Generierung eines Interrupt-Signals für den PIC, da es nur dann einen HIGH-Pegel an seinem Ausgang liefert, wenn *genau* eines der beiden Koinzidenzsignale anliegt, ansonsten liegt sein Ausgang auf LOW. Da der Ausgangssignalpegel des XOR-Gatters aufgrund der kurzen Eingangspulslängen nicht ausreicht, um den Interrupt[31] am PIC direkt auszulösen, wird ein zusätzlicher Transistor benutzt, der durch den Puls kurzzeitig in Sättigung geschaltet wird und so am Porteingang PB0 des PICs eine HIGH-LOW-Flanke erzeugt.

4.5.2.2 Bitaufnahme, Zwischenspeicherung und Datentransfer

Ein wesentliches Problem der NIM-Einschübe verwendenden einfachen Datenaufnahmeelektronik besteht in der ineffizienten Datenübergabe an den Meßrechner. Der Meßrechner muß interruptgesteuert jedes einzelne vom Generator erzeugte Bit einzeln einlesen, so daß er einen nicht unerheblichen Teil seiner Rechenzeit für den reinen Datentransfer benötigt. Läßt sich dies bei einem reinen Meß- und Datenaufnahmerechner noch verschmerzen, ist es bei einem Rechner, der Zufallsbits lediglich als Input für eigene Programme braucht, völlig untragbar.
Deshalb wird bei der kompakten Datenaufnahmeelektronik ein 8-Bit CMOS EEPROM Microcontroller (PIC), Typ Microchip 16C84 verwendet, um die einzelnen Bits zu Bytes zusammenzufassen, und diese immer, wenn ein Byte vollständig ist, an einen nachfolgenden High Density First-In First-Out (FIFO) 4096 × 9-Bit CMOS Speicher[32] wei-

[31] Es wird allerdings aufgrund der großen Interrupt-Latenz von 4,75 Instruktionszyklen keine Interrupt-Routine verwendet, sondern lediglich in einer Endlosschleife abgefragt, ob das Interrupt-Flag gesetzt ist.
[32] Das neunte Bit der FIFO-Blöcke dient üblicherweise als Parität und wird hier nicht verwendet.

terzureichen. Die Verwendung eines PIC bietet sich an, da sich PICs aufgrund ihrer relativ leichten Programmierbarkeit und einfachen Beschaltung sehr gut dafür eignen, kompliziertere Digitalschaltungen durch ein Programm zu ersetzen. Dies bringt überdies den Vorteil mit sich, daß sich auch nach Aufbau der Schaltung noch leicht Änderungen in der Verarbeitung der Bits vornehmen lassen. So können durch Programmänderung zusätzliche Verzögerungen oder auch einfache Regularisierungsalgorithmen (z.B. Von-Neumann-Regularisierung) implementiert werden.
Da der verwendete PIC mit 10 MHz getaktet wird und ein Instruktionszyklus vier Taktzyklen dauert, braucht er für die Abarbeitung eines Befehls[33] 400 ns, angesichts der relativ niedrigen Bitraten der Zufallsbits, ist dies aber ausreichend. Allerdings ist es notwendig, das generierte Zufallsbit in einem Flipflop zwischenzuspeichern, da mindestens zwei Instruktionszyklen vergehen, bis der PIC nach einem Signal an seinem Eingangs-Port PB0 das Bit über den Eingangs-Port PB1 einliest.
Der als Eingang programmierte Port PB2 wird dazu verwendet, das Full-Flag des FIFO abzufragen; ist das Full-Flag gesetzt (active low), überträgt der PIC das Byte nicht in den FIFO, sondern wartet, bis im FIFO wieder Speicherplatz vorhanden ist; währenddessen werden allerdings auch keine Zufallsbits mehr vom PIC angenommen.
Der FIFO faßt 4 kB und hat die ausgesprochen praktische Eigenschaft, daß das Schreiben in ihn und das Lesen aus ihm entkoppelt sind. Daher ist es möglich, daß der PIC Daten in den FIFO überträgt, während der Computer Daten aus dem FIFO liest. Für die mit dem momentanen Aufbau erreichbaren Bitraten unter 10 kBit pro Sekunde ist es somit völlig ausreichend, einmal pro Sekunde Daten auszulesen[34].
Der Meßrechner liest die Daten aus dem FIFO über eine bidirektionale Parallelschnittstelle, die Druckerschnittstelle moderner PCs, ein. Um den Generator auch problemlos an eine herkömmliche Schnittstelle anschließen zu können, wurde für den Datenausgang eine 25-polige Sub-D-Buchse mit entsprechender Anschlußbelegung gewählt: An Pin 1 (Strobe) liegt der Read-Clock für den FIFO, an den Pins 2 bis 9 liegen die Datenausgänge und an Pin 12 (Paper Empty) ist das Empty-Flag des FIFO angeschlossen. Mit Hilfe der Jumper K1 und K2 läßt sich überdies die Signalflanke von „Strobe“ und „Paper Empty“ wählen. Ist das Empty-Flag gesetzt (active low), können keine Daten mehr aus dem FIFO gelesen werden, da keine mehr im Speicher sind. Das auf dem Computer laufende Datenübertragungsprogramm darf dann keine Daten mehr anfordern, da sonst immer nur Nullen gesendet werden!
Ist hingegen das Full-Flag des FIFO gesetzt, so werden keine Bytes mehr vom PIC in den FIFO geschrieben, und auch keine neuen Zufallsbits vom PIC übernommen. Erst wenn wieder Platz im FIFO ist, d.h. wenn der angeschlossene Rechner Daten aus dem FIFO ausgelesen hat, geht das Abspeichern der Zufallsbits weiter. Bei den Experimenten wird der FIFO aber hinreichend oft ausgelesen, so daß dieses unfreiwillige „Anhalten“ des Zufallsgenerators nicht auftritt.

[33] Die Befehle des PICs benötigen i.a. einen Instruktionszyklus, lediglich bedingte Programmverzweigungen benötigen zwei.

[34] Bei den Messungen wurde allerdings der FIFO in einer Programmierschleife ständig vom Rechner ausgelesen, nur unterbrochen von Festplattenzugriffen und Bildschirmausgaben.

4.5.2.3 Datenaufnahme-Software

Die Datenaufnahme-Software für die kompakte Datenaufnahme-Elektronik ist eine vereinfachte Version des Datenaufnahme-Programmes der auf NIM-Einschüben basierenden Datenaufnahme-Elektronik. Benutzerschnittstelle und Kommandozeilen-Optionen wurden beibehalten, lediglich der innere Aufbau des Programmes wurde leicht verändert: Da aufgrund des Zwischenspeichers innerhalb der kompakten Datenaufnahmeelektronik nun keine Notwendigkeit mehr besteht, interruptgesteuert zu arbeiten, wird nur noch in einem Abfragemodus gearbeitet. Hierbei wird der mit Daten gefüllte Teil des FIFO-Pufferspeichers über die Parallelschnittstelle byteweise vollständig ausgelesen.

4.5.3 Hybride Signalverarbeitungs- und Datenaufnahme-Elektronik

Bei der hybriden Signalverarbeitungs- und Datenaufnahme-Elektronik wird die Pulsformung und Koinzidenz wieder mit Hilfe von NIM-Einschubmodulen realisiert, s. Abschnitt 4.5.1; allerdings wird das Übernahmesignal anders erzeugt und die nachfolgende Datenaufnahme-Elektronik verwendet den größten Teil der Elektronik der kompakten Datenaufnahme-Elektronik. Mit Hilfe eines NIM-Logikmoduls (EG&G, 4fold logic, Model CO4010) wird aus den Koinzidenzsignalen ein Übernahmesignal generiert. Hierzu wird der Einfachheit halber aber keine XOR, sondern lediglich eine ODER-Verknüpfung der beiden Signale gewählt. Dies mag auf den ersten Blick etwas inkonsequent erscheinen, da auch dann ein Bit[35] generiert wird, wenn auf *beiden* Signal-Eingängen ein HIGH-Pegel liegt, dies tritt allerdings aufgrund der relativ niedrigen Koinzidenzzählraten selten auf. *Eine* der beiden Koinzidenz-Ausgangsleitungen dient als Signal für den Bitwert Eins und wird über eine Gate&Delay-Einheit[36] an den Bitwert-Eingang des PIC (Eingangs-Port PB1, s. Abschnitt 4.5.2.2) innerhalb der kompakten Datenaufnahme-Elektronik angeschlosssen. Der Bitwert Null wird durch einen LOW-Pegel auf der Signal-Leitung bei gleichzeitig vorhandenem Übernahme-Signal generiert. Das Übernahme-Signal[37] wird mit Hilfe einer Gate&Delay-Einheit auf die Dauer eines Instruktionszyklus des PIC (400 ns) verlängert und an den Interrupt-fähigen Eingang des PICs (Eingangs-Port PB0) gelegt[38].
Wie in Abschnitt 4.5.2.2 erwähnt, muß bei der Ansteuerung des PIC dafür gesorgt werden, daß das Bitsignal nach Generierung des Interrupts lange genug am Eingangs-Port PB0 des PIC anliegt. Wählte man nämlich die Pulslänge zu kurz[39], so läge trotz eines Eins-Signals kein HIGH-Pegel mehr am Eingang des PIC beim Lesen des Eingangszustand, so daß nur Nullwerte aufgenommen würden. Eine zu lange Pulslänge ist andererseits auch nicht vorteilhaft, da in diesem Fall die Möglichkeit besteht, daß ein weiteres Übernahmesignal generiert wird, obwohl noch der vorhergehende Signal-Wert anliegt.
Da während der insgesamt 2 μs dauernden Interrupt-Bearbeitung kein neuer Interrupt

[35] Das in diesem Fall generierte Bit hätte immer den Wert Eins, s.u..
[36] Die Gate&Delay-Einheit gibt positive Pulse aus, die zur nachfolgenden TTL-Logik kompatibel sind.
[37] Die vom Logik-Modul verursachte Laufzeitverzögerung beträgt lediglich 7 ns und kann daher vernachlässigt werden, s.u..
[38] Genau genommen wird das Signal wieder an die Basis des Transistors T1 gelegt, vgl. Abschnitt 4.5.2.1.
[39] In der Tat führt z.B. eine Bitpulslänge von 1,2 μs zu einem erheblich reduzierten Anteil von Einsen.

bearbeitet[40] wird, tritt das Problem insbesondere bei Pulsen auf, die länger als diese Zeitspanne sind, so daß schließlich[41] für die Länge des Bitwertpulses eine Zeitdauer von 2 μs gewählt wird.

[40] Das Interrupt-Flag wird nur am Anfang der Aufnahme-Routine abgefragt und erst am Ende wird das Flag wieder zurückgesetzt für die nächste Abfrage.

[41] Bei zwei Läufen, *LH1* und *LH4* wurde eine Bitpuls-Länge von 4,6 μs verwendet, was allerdings starke Antikorrelationen zur Folge hatte, s. Abschnitt B.10.4 und 6.1.1.2.

Kapitel 5

Durchführung und Ergebnisse der Experimente

In diesem Kapitel wird die Durchführung der Experimente mit den beiden verschiedenen Typen von quantenoptischen Zufallsgeneratoren und ihren Aufbauvarianten beschrieben. Neben der Vorgehensweise bei der Justage der verwendeten Komponenten werden die Gründe für die verschiedenen Modifikationen am experimentellen Aufbau dargelegt und die typischen Leistungsdaten der Generatorvarianten aufgeführt. Eine Zusammenfassung der Ergebnisse der Analysen der erhaltenen Zufallsdaten mit Hilfe statistischer Tests findet sich in der Diskussion, eine ausführliche Darstellung im Anhang.
Um das Kapitel nicht mit einer länglichen Tabelle zu überfrachten, wurde die ausführliche Auflistung der durchgeführten Läufe, mit Bitraten und zusätzlichen Bemerkungen zu den einzelnen Läufen in den Abschnitt A des Anhangs verschoben.
Zur Vereinfachung der Verweise in den folgenden Abschnitten ist jeder Lauf mit einem Kürzel versehen. Läufe mit einer Einphotonenquelle auf Basis der parametrischen Fluoreszenz werden durch ein Kürzel, das mit K oder L anfängt, gekennzeichnet, je nachdem ob der Kaliumniobat- oder der Lithiumjodat-Kristall verwendet wird. Die Läufe mit der statistischen Einphotonenquelle besitzen Kürzel, die mit P beginnen. Ein F im Kürzel steht dafür, daß der Lauf mit einem faseroptischen Aufbau durchgeführt wurde. Die nach den Buchstaben eines Kürzels folgende Zahl zeigt an, der wievielte Lauf es in einer Reihe von Läufen einer bestimmten Aufbauvariante ist.

5.1 Experimente mit der Einphotonenquelle auf Basis der parametrischen Fluoreszenz

Allen verschiedenen Aufbauvarianten des quantenoptischen Zufallsgenerators mit einer Einphotonenquelle auf Basis der parametrischen Fluoreszenz ist eigen, daß als erstes die Einphotonenquelle optimal einjustiert werden muß. Da sich das Verhalten der Quelle, wenn sie einmal einjustiert ist, sehr gut reproduzieren läßt, ist eine vollständige Neujustage nur am Anfang der Experimente oder bei größeren Änderungen notwendig, wie z.B. der Verwendung eines anderen optisch nichtlinearen Kristalls. Auch bei längeren Stand-

zeiten des Aufbaus oder dem Entfernen des nichtlinearen Kristalls aus seiner Halterung[1] zeigt es sich, daß lediglich eine geringfügige Justage nötig ist.

5.1.1 Justage der Einphotonenquelle

Bevor mit der eigentlichen Justage begonnen wird, berechnet man anhand des bekannten Schnittwinkels des Kristalls und der verwendeten Pumpwellenlänge, welche Temperatur (im Falle des Kaliumniobat-Kristalls) bzw. welcher Winkel (im Falle des Lithiumjodat-Kristalls) für eine Typ-I-Phasenanpassung für den kollinearen Fall notwendig ist. Nachdem man den entsprechenden Parameter auf den berechneten Wert gebracht hat, wird nachgeprüft ob auch tatsächlich Photonen mit der doppelten Wellenlänge der Pumplasers erzeugt werden. Hierzu wird geht man in zwei Schritten vor:

1. Man stellt hinter einem RG 830-Farbglas und einem dielektrischen Spiegel für die Pumpwellenlänge – beide dienen als Filter für das um Größenordnungen intensivere Licht des Pumplasers – einen Quanten-Detektor und kontrolliert ob der Detektor bei richtiger Justage auf die optische Achse[2] erheblich mehr Quanten detektiert als bei einer Fehlstellung. Zusätzlich muß noch die richtige Position der Linse zur Fokussierung des Lichtes auf die Detektorfläche bzw. den Faseranschluß[3] gefunden werden.

2. Man fügt vor dem Detektor einen 50:50 unpolarisierenden Strahlteiler ein und orientiert ihn so, daß der in unveränderter Stellung stehende Detektor eine maximale Zählrate zeigt. In den anderen Ausgangsarm des Strahlteilers stellt man ebenfalls einen Quantendetektor und justiert diesen auf optimale Zählrate.

3. Zur Überprüfung, ob auch tatsächlich Photonenpaare entstehen, verwendet man zwei Koinzidenzeinheiten. An eine werden die elektrischen Detektorsignale *direkt* angelegt, so daß auf diese Weise die *echten Koinzidenzen* durch einen Zähler am Ausgang der Koinzidenzeinheit gezählt werden können. An die andere wird das Signal des einen Detektors direkt und das des anderen über eine zusätzlichen Verzögerungsleitung angelegt. Hierbei wird die Verzögerung größer als das Koinzidenzfenster[4], gewählt, so daß sich auf diese Weise die *zufälligen Koinzidenzen* bestimmen lassen.

   Bei maximaler Einzelzählrate der beiden Detektoren ist die Koinzidenz-Zählrate der *echten* Koinzidenzen durchaus nicht maximal, da die Quanteneffizienz der Detektoren ihr Maximum bei einer niedrigeren Wellenlänge hat, als der Zentralwellenlänge der Photonen eines Paares, s. Abb. 4.6 auf S. 69. Daher ist es notwendig, die beiden Detektoren abwechselnd solange in ihrer Position zu justieren, bis die Rate der *echten* Koinzidenzen maximal wird.

[1] Dies ist beim Lithiumjodat-Kristall aufgrund seines hygroskopischen Charakters notwendig.

[2] Beim kollinearen Fall wird die Justage gerade deswegen stark vereinfacht, weil der Pumpstrahl die optische Achse „markiert“.

[3] Gerade das Fokussieren in den Faseranschluß der neueren Detektoren des Typs EG&G SPCM-AQ-131-F ist recht zeitaufwendig; eine Neujustage ist allerdings glücklicherweise nur bei Veränderungen am Aufbau notwendig.

[4] Bei einem Koinzidenzfenster von Δ_{koinz} =10 ns wurde für die Messung der zufälligen Koinzidenzen eine Verzögerung von $\Delta_{zuf} = 40{,}5$ ns verwendet

Nachdem auf diese Weise sichergestellt ist, daß tatsächlich Photonenpaare erzeugt werden, kann man anschließend dazu übergehen, den nichtkollinearen Fall einzujustieren. Hierzu wird zuerst der entsprechende Phasenanpass-Parameter geringfügig so verändert, daß die Hauptausbreitungsrichtungen der Photonen eines Paares in einem kleinen Winkel (ca. 0,3°) links und rechts von der durch den Pumplaserstrahl vorgegebenen optischen Achse emittiert werden. Da die Photonen des Paares jetzt räumlich getrennt sind, kann man die beiden Quantendetektoren jeweils direkt[5] in die beiden Strahlengänge stellen. Das weitere Vorgehen entspricht dem beim kollinearen Fall: Die Position der Detektoren wird auf die maximale Rate der echten Koinzidenzen hin optimiert.
Neben der Position der Detektoren beeinflußt auch noch der Durchmesser der Irisblenden vor den Detektoren die Koinzidenzraten. Ist es bei der anfänglichen Justage noch sinnvoll den Blendendurchmesser möglichst groß (≈1 cm) zu wählen, um überhaupt Photonen zu detektieren, empfiehlt es sich bei der Justage auf die maximale Koinzidenzrate, die Blendendurchmesser zu verkleinern, da sonst der Anteil der zufälligen Koinzidenzen zu hoch liegt. Verkleinert man den Blendendurchmesser, nimmt die Rate der zufälligen Koinzidenzen ab, wobei allerdings festzustellen ist, daß Durchmesser unter ca. 2 mm das Verhältnis der echten zu den zufälligen Koinzidenzen nicht mehr verbessern[6]. Durchmesser über 4 mm verbieten sich allerdings fast von allein, da die Notwendigkeit, daß alle Detektoren den gleichen Raumwinkel sehen, im Falle des („langsameren“) Triggerdetektors schon zu einem Betrieb mit nichtlinearem (Sättigungs-) Verhalten führen würde. Typische Blendendurchmesser[7] liegen daher zwischen 2 und 4 mm.
Wenn dies alles erledigt ist, kann die Einphotonenquelle als justiert angesehen werden.

5.1.2 Freistrahloptik

Für die Messungen am Aufbau des quantenoptischen Zufallsgenerators mit Einphotonenquelle wird der Detektor, der bei der Justage der Quelle zur Detektion des Signalphotons dient, ausgetauscht gegen das Zufall generierende Element, bestehend aus Strahlteiler und zwei Detektoren.
Im Zuge der Experimente wurden zwei verschiedene Varianten des freistrahloptischen Aufbaus untersucht:

1. Die ersten Messungen (Läufe *K1* bis *K10*) wurden mit einer Einphotonenquelle, die einen Temperatur phasenangepaßten Kaliumniobat-Kristall verwendet, durchgeführt. Bei diesem ersten Meß-Aufbau bestand das Zufall generierende Element aus einem drehbarem $\lambda/2$-Plättchen, einem polarisierendem Strahlteiler und zwei Detektoren in den Ausgangsarmen des Strahlteilers. Für die Datenaufnahme wurde die in Abschnitt 4.5.1 beschriebene auf NIM-Einschüben basierende Elektronik verwendet.

2. Bei der zweiten Aufbau-Variante wurde der Kaliumniobat-Kristall gegen einen

[5] Um einen übermäßig langen Aufbau zu vermeiden werden allerdings noch zwei Spiegel verwendet, um die Strahlen umzulenken, s. Abschnitt 4.2.2.

[6] Dieser Wert hängt natürlich auch von der Entfernung des jeweiligen Detektors vom nichtlinearen Kristall ab.

[7] Bei Wahl des Durchmessers hat man einen gewissen Spielraum, die konkrete Wahl hängt davon ab, ob man minimale zufällige Koinzidenzen oder eine möglichst hohe Bitrate haben möchte, bei leicht erhöhtem Anteil der zufälligen Koinzidenzen.

über den Winkel phasenangepaßten Lithiumjodat-Kristall ausgetauscht und der polarisierende Strahlteiler des Zufall generierenden Elementes gegen einen unpolarisierenden ersetzt; dementsprechend wurde auch das $\lambda/2$-Plättchen aus dem Aufbau entfernt. Später wurde bei dieser Aufbau-Variante, die auf NIM-Einschüben basierende Elektronik durch eine kompakte aus Standard-Bauteilen aufgebauten Signalverarbeitungs- und Datenaufnahme-Elektronik (s. Abschnitt 4.5.2) ersetzt (bei den Läufen *LN2,LN4* bis *LN7* und *LD1* bis *LD9*). Bei den letzten Läufen, *LH1* bis *LH5*, wurde eine hybride Elektronik aus NIM-Einschüben und Teilen der neu entwickelten Datenaufnahme-Elektronik verwendet.

Es sind zwei Gründe, die zum Austausch des polarisierenden Strahlteilers und des $\lambda/2$-Plättchen führten:

1. Es hat sich herausgestellt, daß es aufgrund der Verwendung dielektrischer Schichten in Transmissionsrichtung des Strahlteilers immer einen gewissen Untergrund gibt, selbst wenn aufgrund der Eingangspolarisation des Lichtes eine völlige Reflexion zu erwarten wäre. Dies ist eine Abweichung vom Verhalten des idealen, polarisierenden Strahlteilers, die sich auch nur in Transmission zeigt; in Reflexion findet man die theoretischen Erwartungen bestätigt. Zwar sieht es nicht so aus, als ob es nach Einstellung eines Teilungsverhältnisses von 50:50 mit Hilfe des $\lambda/2$-Plättchens Probleme gäbe, aber da man in diesem Fall ein zusätzliches Stellelement hat, das sich dejustieren kann, erscheint es sinnvoller, einen unpolarisierenden Strahlteiler zu verwenden.

   Mit anderen Arten von polarisierenden Strahlteilern, die nicht dielektrische Schichten benutzen, sondern beispielsweise kristalloptische Eigenschaften, wie z.B. einem Wollaston-Prisma, sollte sich dieses Asymmetrieproblem allerdings vermeiden lassen.

2. Das $\lambda/2$-Plättchen dreht die lineare Polarisationsrichtung nur für eine Wellenlänge exakt, nämlich für die Wellenlänge, für die es gefertigt wurde. Diese Wellenlänge wurde beim Laboraufbau gleich der Zentralwellenlänge (884 nm) der Photonen der Paare gewählt. Allerdings haben die Photonen eines Paares eine Bandbreite von ca. 100 nm um die Zentralwellenlänge, was dazu führt, daß für Wellenlängen, die von der Zentralwellenlänge abweichen, das $\lambda/2$-Plättchen nicht mehr nur die lineare Polarisationsrichtung dreht, sondern zusätzlich eine elliptische Polarisationskomponente einführt. Dies erscheint gerade im Zusammenspiel mit dem oben aufgeführten Asymmetrieproblem als nicht unproblematisch. Dies auch deswegen, weil der anfangs verwendete temperatur-phasenangepaßte Kaliumniobat-Kristall ein leicht schwankendes Abstrahlprofil besitzt, s. Abschnitte 4.2.2 und 5.1.2.3). Bei diesen Schwankungen verändert sich der Winkel, in dem die Zentralwellenlänge gegenüber dem Pumpstrahl emittiert wird, mit der Temperatur des Kristalls; dies führt dazu, daß die Detektoren einen anderen, leicht verschobenen Wellenlängenbereich sehen. Ein zeitlich schwankendes, eventuell sogar mit der Temperaturregelung korreliertes Teilungsverhältnis läßt sich in diesem Fall nicht ausschließen.

Bei allen Aufbauvarianten zeigen sich allerdings geringe, zeitliche Schwankungen des Teilungsverhältnisses auf großen Zeitskalen, s. Anhang B.2. Diese Schwankungen wer-

den aufgrund des leicht unterschiedlichen thermischen Verhaltens der beiden Signaldetektoren hervorgerufen; mit einer zusätzlichen aktiven Kühlung läßt sich dieser Effekt minimieren.

5.1.2.1 Zählraten beim Freistrahlaufbau

Die Höhe der Zählraten ist stark abhängig vom jeweiligen Aufbau, da z. B. die Blendenöffnung der Irisblenden oder die Einkoppeleffizienz der Lichtfelder in die Faseranschlüsse der Signaldetektoren einen starken Einfluß auf die Zählraten haben. In Tabelle 5.1 sind die Einzel- und Koinzidenzzählraten für jeweils einen typischen Lauf mit dem Kaliumniobat- bzw. Lithiumjodat-Kristall angegeben[8]. Die Koinzidenzen beziehen sich immer auf die Koinzidenz zwischen dem jeweiligen Signaldetektor und dem Triggerdetektor.

Lauf	Detektor	Einzelzählrate $[1/s]$	echte Koinz. $[1/s]$	zufällige Koinz. $[1/s]$
K10	T	121.921 ± 7052		
	A	108.890 ± 4863	7727 ± 246	352 ± 28
	B	109.021 ± 3025	7985 ± 314	374 ± 33
LH 5	T	140.898 ± 766		
	A	96.650 ± 982	3490 ± 57	146 ± 13
	B	85.834 ± 853	2934 ± 56	106 ± 11

Tabelle 5.1: Tabelle mit den Zählraten pro Sekunde der Läufe *K10* und *LH5*

Es fällt auf, daß die Einzelzählraten des Triggerdetektors zu niedrig sind, sollte der Triggerdetektor doch eigentlich die doppelte Photonenrate sehen wie die Signaldetektoren, da vor ihm kein Strahlteiler steht wie bei diesen. Der Grund hierfür liegt in der niedrigeren Quanteneffizienz und längeren Totzeit, die der Triggerdetektor – ein älteres Modell – aufweist[9].
Mit den verwendeten Koinzidenzeinheiten (10 ns Fenster) liegt das Verhältnis zwischen echten und zufälligen Koinzidenzen bei typischerweise 20:1 oder besser. Bei Kontrollmessungen vor dem Lauf *LH5* stellte sich überraschenderweise heraus, daß die elektrischen Signale der beiden Signaldetektoren, die eigentlich kürzer als das gewählte Koinzidenzfenster sind, durch die TTL-ECL-Wandlung auf ca. 30 ns und damit auf die dreifache Breite des Koinzidenzfensters verlängert werden; daher müssen sie bei Verwendung von NIM-Einschüben für die Koinzidenz vorher wieder verkürzt werden. Dieses Problem betrifft alle Läufe vor dem Lauf *LH5*, welche *nicht* die kompakte Datenaufnahme-Elektronik verwendeten, d.h. insbesondere alle Läufe mit der einfachen Datenaufnahme-Elektronik. Außerdem gilt es für die entsprechenden Kontrollmessungen zur Ermittlung der Koinzidenzraten vor und nach einem Lauf. Die betroffenen Läufe weisen daher einen höheren Anteil von zufälligen Koinzidenzen auf. Dementsprechend besser ist auch das

[8] Es handelt sich hierbei um die mittleren Raten mit Standardabweichung aus Kontrollmessungen vor dem Start des jeweiligen Laufes. Die Werte sind aufgrund einer geringeren effektiven Quanteneffizienz der Detektoren bei höheren Chiptemperaturen i.a. vor Beginn eines Laufes höher als nach Ende des Laufes, s. a. Abschnitt 5.1.2.3.

[9] Eines der neuen Modelle, an derselben Stelle wie der Triggerdetektor aufgestellt, sieht erheblich mehr.

Verhältnis von echten zu zufälligen Koinzidenzen nach Beseitigung dieser Verlängerung bzw. den Läufen mit der kompakten Datenaufnahmeelektronik: Bei der Kontrollmessung zu Lauf *LH5* ergibt sich ohne Pulsverkürzung ein Verhältnis von nur 12:1, mit Pulsverkürzung bei ansonsten unveränderten Rahmenbedingungen hingegen ein Verhältnis von 32:1. Aus oben erwähnten Kontrollmessungen mit und ohne zusätzliche Pulsverkürzung der Signaldetektor-Pulse ergibt sich, daß die Rate der „echten" Koinzidenzen bei den früheren Kontrollmessungen um ca. 12 % zu hoch liegt, während die zufälligen Koinzidenzen um einen Faktor drei zu hoch sind. Dennoch läßt sich festhalten, daß der elementare Zufall des Zufall generierenden Elementes nur verhältnismäßig geringfügig durch andere Zufallsmechanismen „verunreinigt" wird.
Die Koinzidenzzählraten – und damit auch die Bitraten[10] des Generators – lassen sich im Prinzip noch weiter erhöhen, indem man die Irisblenden vor den drei Detektoren weiter öffnet, allerdings steigt dann auch der Anteil der zufälligen Koinzidenzen. Sie lassen sich bei einer Koinzidenzdetektion nie ganz vermeiden, sondern nur durch ein kleineres Koinzidenz-Fenster minimieren[11].

5.1.2.2 Messungen mit der einfachen Datenaufnahme-Elektronik

Bevor längere Messungen mit den verschiedenen Typen der quantenoptischen Generatoren gemacht werden, muß kontrolliert werden, inwieweit die Elektronik ihren Zweck erfüllt. Wie in Abschnitt 4.5.1 bereits erwähnt, mußten aufgrund des anfangs verwendeten Rechners (80386 Prozessor, 20 MHz) und dem Einlesen der einzelnen Bits über den Parallelport des Rechners die elektrischen Signale für die Bitwerte leider beträchtlich verlängert werden. Es empfiehlt sich daher eine Messung mit dem Test-Datenaufnahmeformat durchzuführen, bei welchem der Status beider Signalleitungen aufgezeichnet wird, s. Anhang, S.209. Es zeigt sich hierbei, daß die langen elektrischen Pulsdauern von 100 μs einen bis zu sechsprozentigen[12] Anteil von „Fehlern" erzeugen, bei denen auf beiden Eingangsleitungen ein HIGH-Pegel anliegt[13], also scheinbar ein Null- und ein Einswert gleichzeitig generiert werden. Solche Ereignisse werden von der Datenaufnahmesoftware bei den Standard-Datenaufnahmeformaten nicht aufgezeichnet, sondern lediglich gezählt und automatisch verworfen, da es sich bei ihnen um einen von den langen Pulsdauern hervorgerufenen Artefakt handelt.
Weiter kann man feststellen, daß es auch nicht möglich ist, mit der einfachen Elektronik die Bitraten zu erreichen, die man aufgrund der Koinzidenzzählraten erwarten würde. Wegen der interruptgesteuerten Übernahme jedes einzelnen Bits durch den Steuerrechner wird die maximale Bitrate auf ca. 4700 Bits/s limitiert, unabhängig von den tatsächlichen Koinzidenz-Zählraten, deren Summe i. a. höher liegt. Dieses Problem läßt sich nur durch eine effizientere Datenaufnahme lösen, die in der Lage ist, *kurze* Pulse

[10] Vorausgesetzt die Datenaufnahme-Elektronik ist schnell genug, was z.B. anfangs nicht der Fall war, s. Abschnitt 4.5.1.

[11] Grundsätzlich ließe sich das auch erreichen, da die Zeitauflösung der Detektoren unter 1 ns liegt. Mit schnellen GaAs-Logik-Bausteinen ließen sich für eine kompakte Datenaufnahmeelektronik eventuell Koinzidenzeinheiten mit 1–2 ns Koinzidenz-Fenstern bauen.

[12] Der Anteil ist abhängig von der Bitrate, bei niedrigeren Raten werden auch entsprechend weniger „Fehler" erzeugt.

[13] Alle Fehler sind von dieser Art, aufgrund der langen der Signalpulse gab es keine Fehler, bei denen ein Übergabesignal generiert wurde, aber gar kein HIGH-Pegel (mehr) an der Signalleitungen anliegt.

zu verarbeiten, eine Zwischenspeicherung für die Bits besitzt und über eine effiziente Datenübertragung der Zufallsbits zum Meßrechner verfügt. Deshalb wurde auch eine neue, kompakte Signalverarbeitungs- und Datenaufnahme-Elektronik entwickelt.

5.1.2.3 Messungen mit dem temperaturgeregelten Kristall

Die Möglichkeit das Teilungsverhältnis des Zufall generierenden Elementes mit Hilfe des $\lambda/2$-Plättchens einzustellen, funktioniert wie erwartet gut. Allerdings zeigt sich bei der jeweils vor und nach einem Lauf durchgeführten Messung der Einzeldetektions- und Koinzidenzzählraten, daß die Zählraten insbesondere nach einem längeren Lauf stark schwanken. Zur Illustration sind in Abbildung 5.1 die Koinzidenzzählraten zwischen einem der Signaldetektoren[14] und dem Triggerdetektor vor und nach einem 59 Stunden dauernden Lauf[15] (*K8*) aufgetragen. Zum Vergleich sind in Abb. 5.2 die ent-

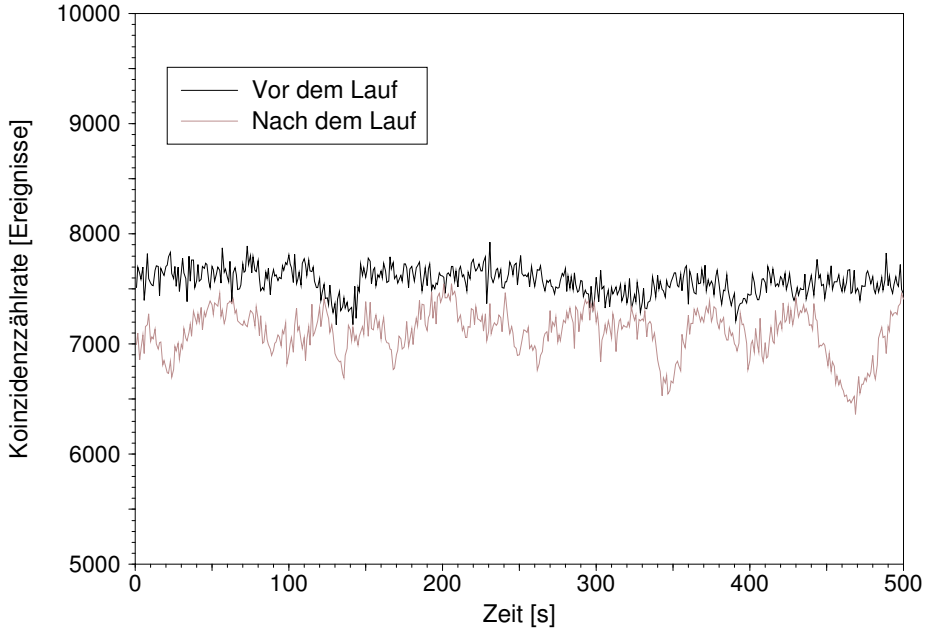


Abbildung 5.1: Koinzidenzrate vor und nach dem Lauf *K8* mit dem temperaturphasenangepaßten Kaliumniobatkristall

sprechenden Zählraten für einen 52 Stunden dauernden Lauf (*LO2*) mit dem winkelphasenangepaßten Lithiumjodat-Kristalls dargestellt. Man erkennt in Abb. 5.1 bei der Messung nach dem Lauf einen deutlichen Rückgang der Koinzidenzzählrate bei gleichzeitiger starker Zunahme der Schwankungen im Vergleich zur Messung vor dem Start des Laufes. Der Rückgang der Koinzidenzzählrate, der sich bereits bei den Einzelzählraten feststellen läßt, rührt daher, daß sich die Detektionseffizienz der Detektoren bei zunehmender Temperatur verschlechtert; dementsprechend tritt auch bei Verwendung des winkel-phasenangepaßten Lithiumjodat-Kristalls solch ein Rückgang auf, s. Abb. 5.2.

[14] Es handelt sich um den Detektor, der die Nullen generiert.

[15] Bei anderen Läufen mit dem Kaliumniobat-Kristall waren die Schwankungen anfangs sogar noch stärker.

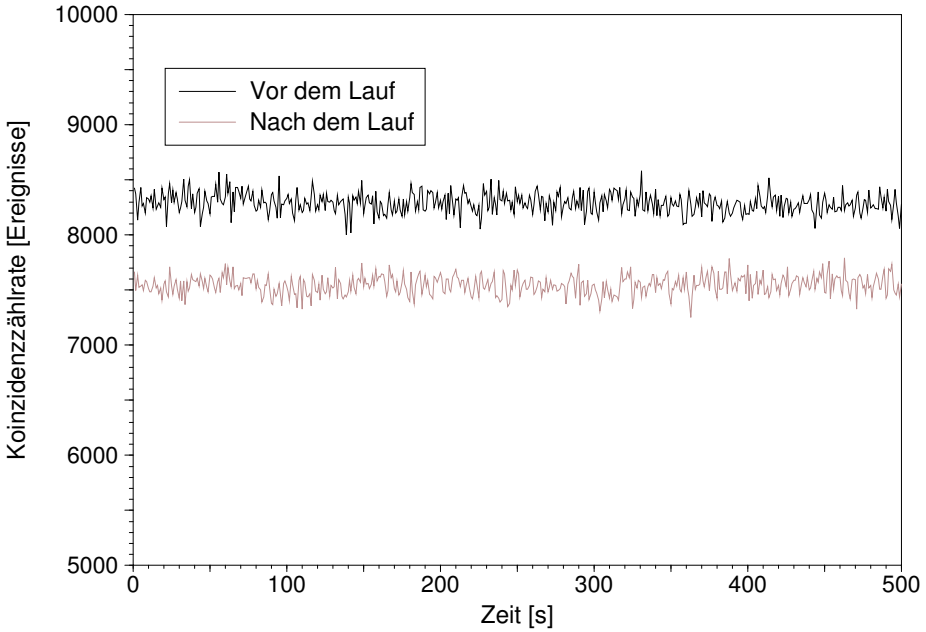


Abbildung 5.2: Koinzidenzrate vor und nach dem Lauf *LO2* mit dem winkel-phasenangepaßten Lithiumjodatkristall

Die starken Schwankungen in der Koinzidenzzählrate werden allerdings von thermischen Effekten[16] im temperatur-phasenangepaßten Kaliumniobat-Kristall verursacht, die zu einer Fehlanpassung führen; hierbei werden die zentralen Ausbreitungsrichtungen längs derer das Signal- bzw. das Triggerphoton propagieren, im Winkel verändert, so daß weniger Photonen auf die Detektoren fallen. Daher schwanken die beiden Koinzidenzzählraten zwischen den Signaldetektoren und dem Triggerdetektoren auch gleichartig. Wie in Abb. 5.2 zu sehen ist, treten diese Schwankungen bei Messungen mit dem winkelphasenangepaßten Lithiumjodat-Kristall nicht auf.
Nach den anfänglichen Testmessungen im beide Signalwerte abspeichernden Format und einer weiteren Messung wurden die meisten[17] längeren Läufe (*K1*,..., *K9*) mit einer in Software realisierten (s. S. 209) Von-Neumann-Regularisierung durchgeführt. Dies geschah, um gleich möglichst ideale Zufallssequenzen zur Verfügung zu haben und nicht erst große Datenmengen[18] mit tendenzbehafteten Bits abspeichern zu müssen. Es stellte sich aber heraus, daß dies keine glückliche Entscheidung war, da die Regularisierung zwar in der Tat ihren Zweck erfüllt, dies aber so gründlich, daß etwaige tieferliegende, physikalisch bedingte Eigenheiten des Generators von ihr verschleiert werden. Ab der letzten Messung mit dem Kaliumniobat-Kristall wurde deswegen bei allen weiteren Messungen auf eine direkt bei der Datenaufnahme stattfindende Regularisierung verzichtet.

[16] Hierbei spielt auch die aufgrund der verwendeten Pumpwellenlänge notwendige, verhältnismäßig hohe Phasenanpassungstemperatur von 80,4°C eine Rolle.

[17] Lediglich beim letzten Lauf (*K10*) mit dem Kaliumniobatkristall wurden Rohdaten abgespeichert.

[18] Bei einer Abspeicherung der Rohdaten ist mit einer mindestens vierfachen Datenmenge zu rechnen und der für die Datenerfassung zur Verfügung stehende Festplattenplatz im Laborrechner war anfangs recht eingeschränkt.

5.1.2.4 Messungen mit dem winkel-phasenangepaßten Kristall

Das Auswechseln des Kaliumniobat-Kristalls gegen einen über den Winkel phasenanpassbaren Lithiumjodat-Kristall beseitigt das Problem der schwankenden Koinzidenz-Zählraten, sie bleiben jetzt weitgehend konstant (s. Abb. 5.2). Dies war auch zu erwarten, da sich der Winkel, einmal eingestellt, nicht mehr verändert. Bei allen Messungen mit dem Lithiumjodat-Kristall wurde im Hinblick auf einen möglichst einfachen Aufbau ein unpolarisierender Strahlteiler verwendet und so das $\lambda/2$-Plättchen ganz eingespart. Es zeigt sich, daß auch bei diesem einfacheren Aufbau das Teilungsverhältnis des Zufall generierende Elementes noch nahe genug bei 50:50 liegt, s. Abschnitt B.2.
Grundsätzlich besteht aber auch bei Verwendung eines unpolarisierenden Strahlteilers die Möglichkeit in beschränktem, aber durchaus ausreichendem Maße, das Teilungsverhältnis des Zufall generierenden Elementes zu verändern: Lassen sich die Blenden vor den Detektoren doch leicht unterschiedlich einstellen[19], so daß auf einen der Detektoren weniger Photonen fallen, als vom Strahlteiler in diese Richtung gelenkt werden. Bei den einzelnen Läufen wurde allerdings immer versucht, die Blendenöffnungen der beiden Signaldetektoren möglichst gleich weit zu öffnen. Lediglich die Blende vor dem Triggerdetektor wird ca. 10% größer als die Blende vor den Signaldetektoren gewählt, da er sich in etwas weiterem Abstand vom Kristall (bzw. den Umlenkspiegeln) befindet.
Bei den Läufen *LO1* bis *LO3* und *LN3* des Generators wurde noch die einfache Datenaufnahme-Elektronik verwendet, ab dem Lauf *LN1* wurde dann die kompakte Signalverarbeitungs- und Datenaufnahme-Elektronik eingesetzt. Hierbei gab es zwei große Probleme; das eine ließ sich weitgehend beseitigen (s. Abschnitt 5.1.2.5 und Abschnitt B.9), das andere (s. Abschnitt B.9) wurde durch Einsatz der hybriden Elektronik (s. Abschnitt 4.5.3) sehr stark abgeschwächt. Die hybride Elektronik wurde für alle Läufe ab Lauf *LH1* verwendet, d.h. bei den letzten Läufen des freistrahloptischen Aufbaus mit Einphotonenquelle, allen Läufen des Generators mit dem entsprechenden faseroptischen Aufbau und allen Läufen des Generators mit einer statistischen Einphotonenquelle.

5.1.2.5 Messungen mit der kompakten Signalverarbeitungs- und Datenaufnahme-Elektronik

Wie in Abschnitt 4.5.2 beschrieben, besteht die kompakte Elektronik aus einem Schaltungsteil für die Puls-Verkürzung und Formung, einer Koinzidenz-Logik, einem programmierbaren Kontrollbaustein (PIC) und einem FIFO-Speicher. Für die weiteren Messungen am Zufallsgenerator, ist es natürlich wichtig, inwieweit die einzelnen Bestandteile der Elektronik auch tatsächlich ihre Funktionen erfüllen.
Nach Dimensionierung des Differenzierglieds in der Pulsformung und Pulsverkürzung funktioniert dieser Teil der Schaltung gut. Die verschiedenen Signale sind nach dieser Stufe alle etwa 10 ns lang. Die als Koinzidenzeinheiten arbeitenden NAND-Gatter sind ebenfalls in der Lage, ihre Funktion zufriedenstellend zu erfüllen, und die Koinzidenzfenster sind auch nicht größer als bei den bisher verwendeten NIM-Einschüben. Was noch etwas zu wünschen übrig läßt, ist allerdings der Umstand, daß man sich bei den doch recht kurzen Pulsen im Hinblick auf Anstiegssteilheit und Amplitude der Ausgangssi-

[19] Da Irisblenden verwendet werden, ist eine *exakt* gleich große Öffnung der beiden Blenden vor den Signaldetektoren ohnehin nur schwer zu erreichen.

gnale schon an den Grenzen der verwendeten Logik befindet. Eine schnellere Logik, die allerdings nicht mehr Standardbauteile verwenden würde[20], wäre hier eventuell wünschenswert.
Nach anfänglichen Problemen mit der Programmierung des PICs und der Generierung des Interruptsignals durch das XOR-Gatter[21], funktioniert die Übernahme der einzelnen Bits in den PIC und die Übergabe der Bytes an den FIFO. Als sehr kritisch erweist sich allerdings die Zwischenspeicherung der Bitwerte in dem aus zwei NAND-Gattern aufgebautem Flipflop, s. Abschnitte 6.1.1.1 und B.9.3, im Zusammenspiel mit der Bitübernahme durch den PIC.
Der PIC hat keine Probleme, die relativ niedrigen[22] anfallenden Bitraten im kHz-Bereich zu verarbeiten. Bei Tests mit künstlichen, von einem Signalgenerator produzierten Pulsen, traten erst oberhalb von 100 kHz Sättigungserscheinungen auf (bei 125 kHz liegt die Bitrate um ca. 5%, bei 150 kHz um ca. 10% zu niedrig), so daß die Datenaufnahme-Elektronik die erreichbare Bitrate, insbesondere beim Einsatz einer Einphotonenquelle auf Basis der parametrischen Fluoreszenz, nicht merklich schmälert[23].
Will man ganz sicher gehen, daß Sättigungserscheinungen erst bei noch höheren Raten auftreten, kann man den PIC durch einen schnelleren, diskreten Aufbau mit einzelnen Digitalbausteinen ersetzen; dies wird bei einem mit Photonenpaaren arbeitenden Generator allerdings kaum nötig sein. Bei Generatoren hingegen, die mit abgeschwächten Pulsen arbeiten, könnte dies bei hohen Pulsrepetitionsraten aber durchaus notwendig werden.
Leider hat sich die Zwischenspeicherung der Bytes im FIFO anfangs als außerordentlich problemträchtig erwiesen. Aus nicht geklärten Gründen kam es bei dem zuerst verwendeten Typ von FIFO-Baustein, nach einer gewissen Laufzeit der Messung vereinzelt und nur zwei Bytes betreffend, zu Wiederholungen in den an den Computer gesandten Daten. Nach den ersten paar Megabyte Daten (bisweilen auch früher) mehrten sich diese Wiederholungen und es gab eine Tendenz zu immer stärkeren Wiederholungen, auch von mehr als zwei Bytes, unterbrochen von Passagen, in denen wieder alles in Ordnung zu sein schien. Im weiteren Verlauf häuften sich diese Wiederholungen und wurden immer stärker, bis das Auslesen der Daten offensichtlich gar nichts mehr mit dem eigentlichen Zufallsgenerierung zu tun hatte. Schließlich wurden nur noch in schneller Folge die immer gleichen Werte aus dem FIFO gelesen, was sich leicht an extrem hohen Datenraten erkennen ließ, welche der Generator mit Einphotonen-Quelle gar nicht hätte erreichen können.
Zur Illustration seien hier Graphen der standardnormierten Autokorrelationskoeffizienten[24] in Abhängigkeit von der Verschiebung aus den oben erwähnten verschiedenen

[20] Dies kann Probleme bei der Bauteilbeschaffung geben, falls bei einem praktischen Einsatz eine Reparatur nötig werden sollte.

[21] Das Signal direkt aus dem XOR-Gatter hat eine zu niedrige Amplitude, um einen Interrupt am Port PB0 auszulösen, s. S. 78.

[22] Bei der Verwendung von stark abgeschwächten Lichtpulsen mit hoher Pulsfrequenz könnte es allerdings Probleme geben, s.u..

[23] Die Rechteck-Pulse aus dem Signalgenerator kommen allerdings in regelmäßigen Abständen, was bei den Detektorsignalen nicht der Fall sein muß, insofern spielen Sättigungserscheinungen beim praktischen Betrieb bereits bei etwas niedrigeren Raten eine (kleine) Rolle.

[24] Zur besseren Übersicht ist in den folgenden Graphen die Ordinate gleich mit Prozentangaben versehen worden, welche die Wahrscheinlichkeit angeben, daß ein Korrelationskoeffizient bei statistisch

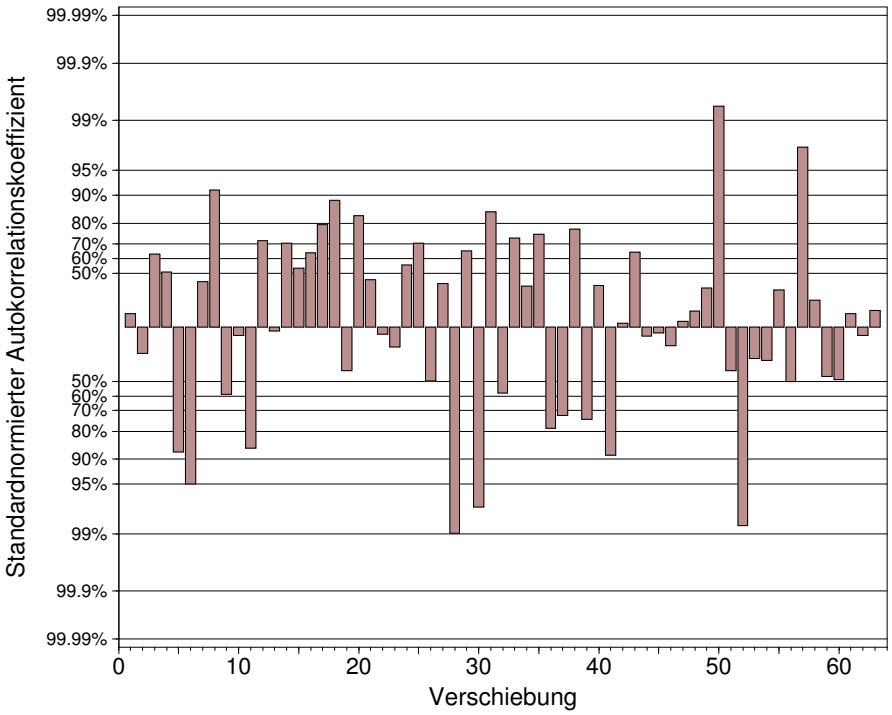


Abbildung 5.3: Standardnormierte, binäre Autokorrelationsfunktion einer Teilsequenz zu Beginn der Datenaufnahme

Phasen gezeigt. Hierbei beträgt die Stichprobenlänge 16 kB, die maximale Verschiebung 63 und alle Stichproben sind einem ersten Probelauf *LN1* des Generators mit der kompakten Datenaufnahme-Elektronik entnommen.
In Abb. 5.3 sieht die Autokorrelationsfunktion noch akzeptabel aus[25]. Etwas weiter in der Bitsequenz (s. Abb. 5.4, S. 94) erkennt man aber bereits einen sehr großen, negativen[26] Autokorrelationskoeffizienten bei einer Verschiebung von 16, also gerade um zwei Bytes. Kurz danach (s. Abb. 5.5, S. 95) ist das Wiederholungsproblem nicht mehr zu übersehen: Die meisten Autokorrelationskoeffizienten haben Werte, die bei einer echten Zufallsfolge extrem unwahrscheinlich sind.
Da sowohl der PIC als auch der Pufferspeicher im Computer die Daten byteweise schreiben bzw. lesen und immer nur maximal ein Byte zwischenspeichern, ist klar, daß es sich um ein Problem des FIFOs bzw. des Zusammenspiels zwischen FIFO und Computer[27] handeln muß.

unabhängigen Bitsequenzen in dem entsprechenden Intervall, d.h. zwischen den beiden Markierungen mit gleichen Prozentangaben, liegt.

[25] Das ist natürlich keine sehr genaue Aussage. Wie man signifikante Abweichungen bei Autokorrelationskoeffizienten präziser nachweisen kann, wird in Abschnitt B.9 ausführlich dargelegt.

[26] Der Koeffizient ist negativ, da die binäre Autokorrelationfunktion eine XOR-Verknüpfung verwendet, die bei stark korrelierten Werten mehr Nullen als Einsen liefert; durch die Bildung der Differenz mit dem theoretischen Erwartungswert ergibt sich somit ein negativer Koeffizient, sein Wert beträgt: −15,33.

[27] Nähere Betrachtung der aufgenommenen Daten zeigt, daß sich teilweise längere Byte-Sequenzen wiederholen. Das kann natürlich nicht vom PIC herrühren, da Fehler in der Datenübergabe von PIC zu FIFO höchstens zu einer Wiederholung von einem einzelnen Byte, gegebenenfalls mehrmals hintereinander, führen könnten.

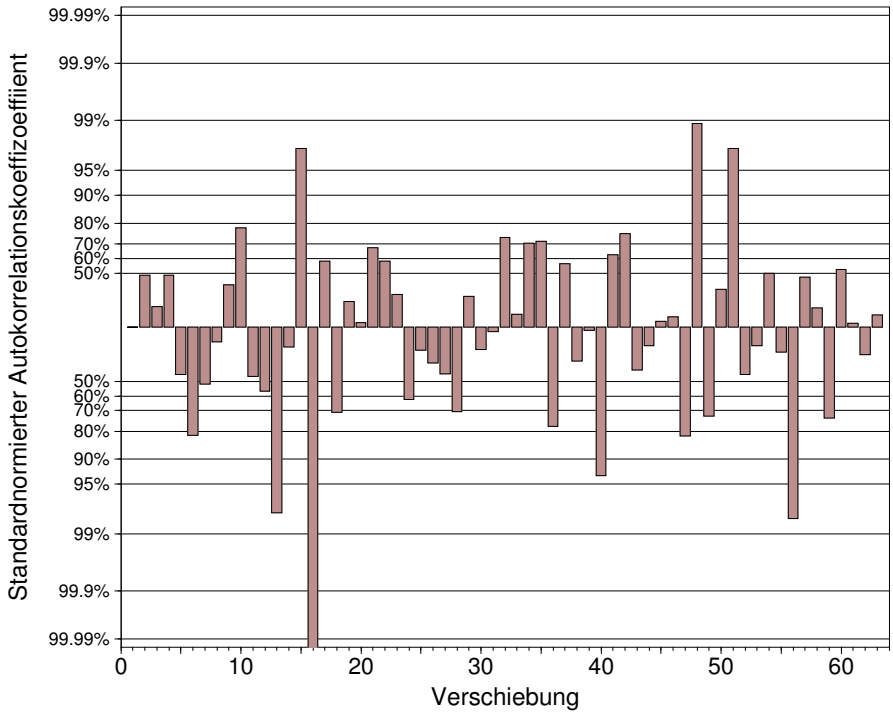


Abbildung 5.4: Standardnormierte, binäre Autokorrelationsfunktion einer „problematischen“ Teilsequenz

Zur Fehlereinkreisung wurden folgende Kontrollmessung vorgenommen: Der FIFO wurde aus der neuen Datenaufnahme-Elektronik entfernt, und die Ausgänge des PICs wurden über die bidirektionale Druckerschnittstelle direkt vom Computer ausgelesen. Hierzu wird der Ausgangs-Port PB3 mit Pin 10 des Druckerports (Acknowledge-Eingang, interruptfähig) verbunden und so umprogrammiert, daß er eine Interruptflanke für den Hardware-Interrupt 7 (Druckerinterrupt) des Computers generiert. Bei Auslösen dieses Interrupt springt der Computer in eine Interruptroutine, die das Byte vom PIC übernimmt. Bei der Untersuchung der auf diese Weise aufgenommenen Daten zeigten sich die oben dargelegten Effekte nicht. Dies zeigte, daß nur die Datenübergabe vom FIFO zum Computer die Ursache für das Problem darstellen konnte.
Um einen Defekt des verwendeten Exemplars konnte es sich ebenfalls nicht handeln, da das Problem auch nach Auswechseln des Bausteins gegen einen neuen FIFO gleichen Typs weiterhin auftrat. Eine mögliche Ursache für die seltsame Fehlfunktion hätte ein starkes Übersprechen sein können; der FIFO-Baustein verfügt nämlich über einen Eingang, der bei entsprechender Ansteuerung eine *Retransmit*-Funktion auslöst, d.h. die gelesenen Daten können noch einmal gelesen werden; das wäre natürlich ein „guter Kandidat“ für die Ursache dieser Effekte gewesen. Allerdings ist der Pin in der Schaltung deaktiviert, indem er direkt an der Versorgungsspannung, d.h. auf HIGH-Pegel, liegt. Ein Übersprechen, das diesen Pin kurzzeitig auf LOW-Pegel ziehen könnte, sollte aufgrund des Schaltungslayouts nicht erfolgen können und ließ sich auch nicht bei Messungen mit dem Oszilloskop entdecken. Erstaunlicherweise verschwanden die Probleme sofort, nachdem ein pin-kompatibler FIFO-Baustein von einem anderen Hersteller eingesetzt wurde.

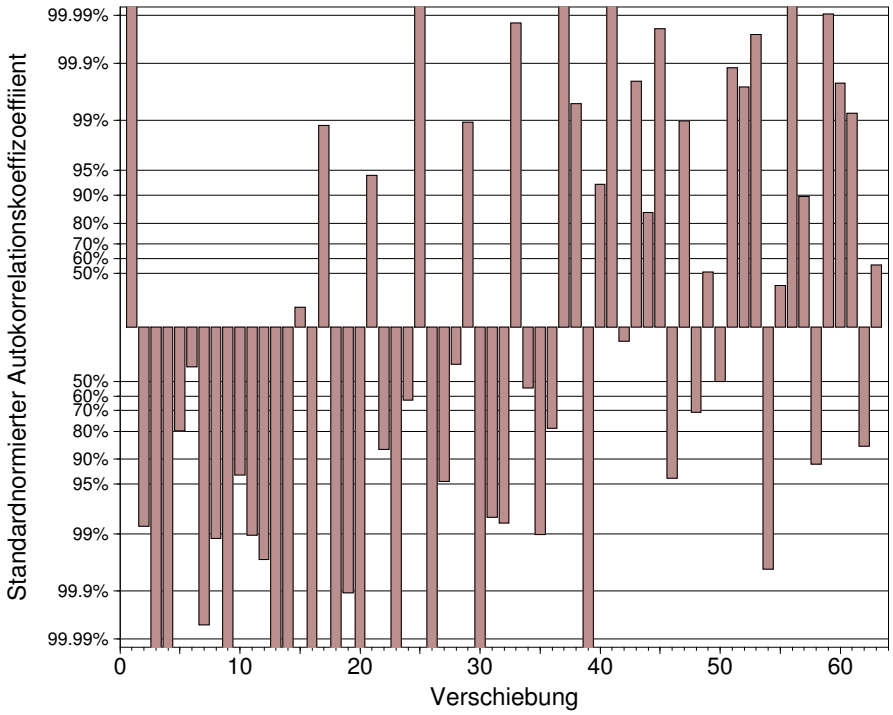


Abbildung 5.5: Standardnormierte, binäre Autokorrelationsfunktion einer ganz und gar nicht zufälligen Teilsequenz

Es wird hier trotzdem so ausführlich auf dieses Problem eingegangen, um einerseits zu illustrieren, daß Probleme auch an Stellen auftreten können, an denen man sie eigentlich gar nicht erwartet und andererseits, um zu zeigen, wie sich Abweichungen vom zufälligen Verhalten eines Zufallsgenerators mit Hilfe von Tests entdecken lassen.
Bedauerlicherweise trat später bei *einem* der letzten Läufe (*LH5*) allerdings doch noch einmal kurzzeitig ein ähnlich geartetes Problem auf, das wahrscheinlich vom FIFO verursacht wurde, s. Abschnitt B.9.2.

5.1.2.6 Manipulationsversuche durch Lichteinstrahlung

Die freistrahloptischen Laboraufbauten sind – anders als man dies bei einem prototypischen Aufbau eines quantenoptischen Zufallsgenerators für den industriellen Einsatz täte – nicht gegen absichtlichen Einfluß von außen geschützt. Zwar ist eine Absicherung gegen optische Beeinflussung von außen kein grundsätzliches Problem, da hiermit aber eine verringerte Flexibilität und eine erschwerte Justage einhergehen, unterblieb dies bei den Laboraufbauten. Zur Überprüfung, wie kritisch sich optische Störversuche auf die Zufallsgenerierung auswirken, wurde probeweise mit einer handelsüblichen Taschenlampe[28] Licht durch den unbenutzten Eingang des Strahlteilerwürfels in Richtung auf einen

[28] Das ist ein relativ brachiales Vorgehen, da die Detektoren für das Zählen von einzelnen Lichtquanten optimiert sind. Sie werden bei solch „intensiver“ Bestrahlung recht stark belastet, deswegen wurde auch nur *einmal* solch ein optischer Manipulationsversuch versucht.

der Signaldetektoren gestrahlt[29]. Durch eine leichte Variation der Einstrahlrichtung der Taschenlampe ist es möglich, gezielt die Bitgenerationsrate *eines* Signaldetektors zu beeinflussen. Je nachdem, wieviel Licht auf ihn fällt, steigt die Rate[30] bis zu einem Faktor vier an. Interessanterweise führt eine noch stärke Bestrahlung zu einem gegenläufigen Effekt, die Rate nimmt wieder ab. Wahrscheinlich wird in diesem Fall der Signaldetektor so stark bestrahlt, daß auf den Detektor auftreffende Photonen bereits vor Ende des Aufladens der Detektordiodenkapazität für eine Entladung sorgen und so die effektive Totzeit des Signaldetektors erhöhen.
Durch die starke Zunahme der Rate, bei lediglich einem der beiden Signaldetektoren weicht das Verhältnis der relativen Häufigkeiten der Einsen und Nullen im Ausgangsbitstrom der Rohdaten natürlich stark vom idealen bzw. dem Verhältnis ohne Beeinflussung ab. Verwendet man eine nachgeschaltete Von-Neumann-Regularisierung wird diese Abweichung selbstverständlich korrigiert; während der Manipulationsversuche zeigte die Teilungsverhältnisanzeige bei einer Datenaufnahme mit gleichzeitiger Von-Neumann-Regularisierung keine markanten Abweichungen vom idealen Verhältnis.

5.1.3 Faseroptik

Der Aufbau in Faseroptik verwendet den winkelangepaßten Lithiumjodat-Kristall. Der Umbau auf ein Zufall generierendes Element, das einen Faserkoppler benutzt, stellte sich als erfreulich einfach heraus. Als erstes wurde der Fokussierkollimator einjustiert. Hierzu wurde einer der Signaldetektoren mit einer Mehrmoden-Glasfaser direkt an den Fokussierkollimator angeschlossen, und dieser anschließend auf optimale Koinzidenz-Zählraten justiert. Dabei ist zu beachten, daß sowohl bei dieser Justage als auch nach Austausch der einfachen Mehrmoden-Glasfaser durch den Faserkoppler, der längere optische Weg aufgrund der jeweiligen Glasfaser-Strecke durch eine entsprechende elektronische Verzögerung des Trigger-Signals ausgeglichen wird.
Ist der Fokussierkollimator optimal einjustiert und stabil fixiert, ist das Zufall generierende Element bereits vollständig justiert, da der Anschluß des Faserkopplers und der Detektoren an dessen Ausgängen mit ein paar Handgriffen erledigt ist. Diese einfache Handhabung erlaubt es auch, die beiden Signaldetektoren sehr gut miteinander zu vergleichen, da sie leicht über eine Mehrmoden-Glasfaser direkt an den Fokussierkollimator angeschlossen werden können.
Hierzu werden die Zählraten der beiden Detektoren bei ansonsten identischen Bedingungen ermittelt, man erhält für die Einzelzählrate[31] bei Detektor A: 319.816 ± 1643 [1/s] und bei Detektor B: 288.363 ± 1493 [1/s]. Die effektive Quanteneffizienz für den Wellenlängenbereich der statistischen Einphotonenquelle, ist also bei Detektor A relativ zu Detektor B um ca. 11% höher, was auch mit den Angaben in den Meßprotokollen des Herstellers übereinstimmt.

[29] Dies geschah beim freistrahloptischen Aufbau mit Kaliumniobat-Kristall, polarisierendem Strahlteilerwürfel und alter Datenaufnahme-Elektronik in der Zeit zwischen den Läufen *K9* und *K10*.

[30] Dies ist die Koinzidenzdetektionsrate zwischen Triggerdetektor und dem Signaldetektor, die Einzeldetektionsrate des Signaldetektors steigt natürlich noch viel extremer an.

[31] Es handelt sich hierbei um die Werte der Kontrollmessungen vor dem Lauf *LF1*. Die Zählraten waren bei diesem Lauf aufgrund der Irisblendenöffnung von 2,1 mm vor dem Faserkollimator und 2,5 mm vor dem Triggerdetektor relativ hoch. Der Triggerdetektor registrierte eine Einzelzählrate von: 167.265 ± 722 [1/s].

Koinzidenzen [1/s]	direkt am Faserkollimator	an den Ausgängen des Faserkopplers
echte (AT)	14169 ± 169	5960 ± 75
zufällige (AT)	693 ± 26	215 ± 15
echte (BT)	14353 ± 114	3468 ± 61
zufällige (BT)	407 ± 20	128 ± 10

Tabelle 5.2: Koinzidenzzählraten pro Sekunde zwischen Signaldetektoren und Triggerdetektor mit und ohne Faserkoppler, Kontrollmessungen vor Lauf *LF1*

Ziemlich drastisch wirkt sich das Einfügen des Faserkopplers aus: Die Einzelzählraten betragen dann nur noch 121.448 ± 1051 [1/s] bei Detektor A und 83.713 ± 886 [1/s] bei Detektor B, wobei die ungleichen Zählraten durch das ungleiche Teilungsverhältnis des Faserkopplers dominiert werden. Die entsprechenden Werte für die Koinzidenzzählraten sind in Tabelle 5.2 aufgeführt.
Trotz des Betriebs des Triggerdetektors in einem Zählratenbereich, wo er bereits starke Sättigungserscheinungen zeigt, erhält man ein gutes Verhältnis zwischen echten und zufälligen Koinzidenzen von ca. 27:1 für den Lauf *LF1*. Bei beiden Läufen mit dem faseroptischen Aufbau des Zufallsgenerators auf Basis einer Einphotonenquelle wird wie bei den letzten Läufen mit dem freistrahloptischen Aufbau die hybride Signalverarbeitungs- und Datenaufnahme-Elektronik verwendet.

5.1.4 Abschätzung der Photonenpaarrate

Aus den Einzelzählraten des Laufes *LF1*[32] läßt sich auch gut die Konversionseffizienz η_{Konv} abschätzen, mit der die blauen Pumpphotonen durch parametrische Fluoreszenz in Photonenpaare umgewandelt werden. Da ohnehin nur eine grobe Berechnung möglich ist, werden Verluste an den Umlenkspiegeln und dem Faser-Einkoppler nicht berücksichtigt; die recht kleinen Dunkelzählraten werden ebenfalls vernachlässigt. Berücksichtigt werden selbstverständlich die effektive[33] Quanteneffizienz des Detektors[34] in Abhängigkeit von der Wellenlänge und die Durchlaßcharakteristik des Farbglases. Aus der detektierten Einzelzählrate läßt sich so die Photonenrate berechnen; die Konversionseffizienz ergibt sich als Quotient aus der Photonenrate und der Anzahl der Pumpphotonen pro Sekunde bei einer Pumpleistung von 10 mW. Man erhält für sie einen Wert von $\eta_{Konv} = 5{,}3 \cdot 10^{-11}$.

5.2 Experimente mit der Photonenquelle auf Basis abgeschwächter Pulse

Die Messungen mit der statistischen Einphotonenquelle unterscheiden sich nur durch die verwendete Lichtquelle von den Messungen mit der Einphotonenquelle auf Basis

[32] Die Einzelzählraten bei den freistrahloptischen Aufbauten sind hierfür nicht gut geeignet, da bei ihnen die Einkopplung in die Faseranschlüsse der Detektoren schwer einschätzbare, stark justageabhängige, zusätzliche Verluste verursacht.

[33] In sie gehen neben der Quanteneffizienz auch noch die zusätzlichen Verluste durch die Ankopplung von Faseranschluß und Chip ein.

[34] Totzeitkorrekturen sind bei den niedrigen Zählraten noch nicht notwendig.

der parametrischen Fluoreszenz. Allerdings stellt sich der Versuchsaufbau nun erheblich reduziert dar, denn kein Laser, kein Kristall und auch kein Triggerdetektor werden benötigt; lediglich das Zufall generierende Element bleibt gleich. Gleichzeitig bedeutet dies auch, daß der Streulichtanteil sehr stark abnimmt, da kein Pumplaserstrahl mehr den Aufbau durchquert[35]. Daher werden die Irisblenden vor den Detektoren bzw. dem Fokussierkollimator weit[36] geöffnet, damit möglichst viele von der Quelle abgestrahlte Photonen auf den Detektor fallen. Dies ist auch noch aus zwei anderen Gründen möglich:

1. Anders als bei der parametrischen Fluoreszenz (s. S. 84) muß nicht darauf geachtet werden, daß bestimmte Bereiche des abgestrahlten Lichtfeldes nicht auf den Detektor fallen; alle Photonen sind bei der statistischen Einphotonenquelle „gleichberechtigt".

2. Es werden nur die beiden „schnellen" Signaldetektoren verwendet, d.h. es muß keine Rücksicht auf die Sättigung des Triggerdetektors bei hohen Zählraten[37] genommen werden.

Eine hinreichend große Blendenöffnung ist außerdem notwendig, weil bei der statistischen Einphotonenquelle explizit sichergestellt werden muß, daß beide Signaldetektoren tatsächlich denselben Raumwinkel sehen.
Als erstes muß auch bei den Experimenten mit einer statistischen Einphotonenquelle, die Quelle justiert werden.

5.2.1 Justage der Photonenquelle

Der Aufbau der Quelle selbst ist recht einfach, daher besteht die Hauptaufgabe darin, die optische Dichte der Graugläser, die zur Abschwächung der Lichtintensität dienen[38], so zu wählen, daß die angestrebte Lichtintensität von ca. 0,1 Photon pro Puls erreicht wird.
Strenggenommen wird allerdings nicht die Intensität des von der Quelle abgestrahlten Lichtfeldes auf diesen Wert gebracht, sondern nur der Anteil, der auch tatsächlich auf die Fläche des zum Einmessen verwendeten Quantendetektors gelangt. Auch bei vollständig geöffneter Irisblende gelangt gerade beim freistrahloptischen Aufbau, aufgrund der schwierigen Fokussierung des Lichtes auf die Faseranschlüsse der Detektoren, nicht der gesamte Anteil des abgestrahlten Lichtes auch tatsächlich auf die Detektoroberfläche. Beim faseroptischen Aufbau wird ein dedizierter Einkoppelkollimator verwendet,

[35] Aufgrund der sehr hohen Empfindlichkeit der Detektoren konnte allerdings dennoch nicht auf die Kantenfilter vor den Detektoren bzw. dem Fokussierkollimator verzichtet werden.

[36] Die Photonenquelle auf Basis abgeschwächter Pulse sendet aufgrund der ins Gehäuse der LED eingebauten Kugellinse ein gut kollimiertes Lichtfeld aus. Anfangs wurden die Irisblenden maximal (12 mm) geöffnet, bei den Läufen allerdings auf 5 mm Öffnung begrenzt, denn ab einer gewissen Irisblendenöffnung stellt man keine Erhöhung der Zählraten mehr fest, was als Zeichen gelten kann, daß man den gesamten Abstrahlraumwinkel erfaßt.

[37] Allerdings sind die Zählraten aufgrund der starken Abschwächung des LED-Lichtfeldes durch die Graugläser ohnehin recht niedrig.

[38] Insgesamt tragen natürlich auch noch das Farbglas vor dem jeweiligen Detektor bzw. dem Fasereinkoppler und insbesondere der Interferenzfilter stark zur Abschwächung bei; diese optischen Elemente sind allerdings nicht variabel.

der sehr viel effizienter in die Faser einkoppelt, so daß bei diesem Aufbau die Graugläser vor der LED unterschiedlich gewählt werden müssen; dementsprechend bedarf es bei der Umstellung vom freistrahloptischen auf den Faseraufbau auch eines erneuten Einmessens.
Die geeignete optische Dichte der Graugläser wird anhand der Einzählrate eines Detektors bestimmt, der optimal einjustiert direkt[39] vor der Quelle steht.
Hierbei gilt es allerdings zu beachten, daß die verwendeten Quantendetektoren lediglich detektieren, *ob* zu einer bestimmten Zeit Photonen absorbiert werden und eine Elektronen-Lawine auslösen. Es kann also anhand des Ausgangssignals *nicht zwischen einem oder mehreren Photonen unterschieden werden.* Überdies muß berücksichtigt werden, daß nur ein relativ kleiner Anteil der Photonen überhaupt eine Lawine auslöst, da die Quanteneffizienz η der Detektoren für die betrachtete Wellenlänge von 884 nm lediglich bei ca. 30 % ($\eta = 0{,}3$) liegt. Somit berechnet sich die Wahrscheinlichkeit, daß beim Auftreffen von n Photonen auf die Detektorfläche ein Detektorsignal ausgelöst wird, gemäß:

$$P_{Det}(n, \eta) = 1 - (1 - \eta)^n,$$

wobei der Ausdruck in der Klammer gerade die Wahrscheinlichkeit ist, daß kein elektrisches Ausgangssignal erzeugt wird[40]. Kombiniert man dies mit der Wahrscheinlichkeit, daß ein Puls bei einer mittleren Intensität von $\bar{n}$ genau n Photonen „enthält“:

$$P_{Poisson}(n, \bar{n}) = \frac{\bar{n}^n}{n!} e^{-\bar{n}},$$

so erhält man die (mittlere) Wahrscheinlichkeit für ein Detektorsignal:

$$P_{Sig}(\bar{n}, \eta) = \sum_{n=0}^{\infty} P_{Poisson}(n, \bar{n}) \cdot P_{Det}(n, \eta),$$

wobei die Summation aufgrund der stark abfallenden Wahrscheinlichkeit von $P_{Poisson}(0.1, n)$ problemlos bei $n = 100$ abgebrochen werden kann. Die erwartete Einzelzählrate beträgt also für eine Pulsfrequenz f im Mittel:

$$R(\bar{n}, \eta, f) = P_{Sig} \cdot f.$$

Die Graugläser werden nun so gewählt[41], daß die Einzelzählrate des Detektors nicht größer ist als die gemäß dieser Formel berechnete Rate.
Das Vorgehen bei der Wahl der optischen Dichte der Graugläser ist recht restriktiv, sie werden so gewählt, daß die mittlere Photonenzahl pro Puls, die auf die Detektorfläche *auftrifft*, nicht größer als 0,1 Photon pro Puls ist. Damit liegt die Wahrscheinlichkeit, daß mehr als ein Photon bei einem Puls auf den Detektor trifft, auch nur bei $P(n > 1) = 0{,}01$, s. o. Abschnitt 3.1.2. Leider fallen durch dieses Vorgehen die Zählraten bei vorgegebener

[39] Beim faseroptischen Aufbau ist der Detektor über eine Mehrmoden-Glasfaser an den direkt vor der Quelle stehenden Faserkollimator angeschlossen.

[40] Es wird hier statistische Unabhängigkeit unterstellt.

[41] Für den Aufbau in Freistrahloptik werden Graugläser mit einer optischen Dichte von 4,6 benötigt und beim Aufbau mit Faseroptik ist eine optische Dichte von 5,6 erforderlich. Die Bestimmung fand immer ohne das Zufall generierende Element statt, aber mit der entsprechenden Einkoppeloptik, direkt hinter der Quelle.

Pulsfrequenz recht niedrig aus, da sie durch die nichtideale Quanteneffizienz der Detektoren gegenüber der Photonenrate reduziert werden, in Abschnitt 6.2.1.3 der Diskussion werden daher Vorschläge diskutiert, wie sich die Zählraten erhöhen ließen.

5.2.2 Freistrahloptik

Es werden zwei Läufe mit dem freistrahloptischen Aufbau und der statistischen Einphotonenquelle durchgeführt: ein kurzer Testlauf (*PB1*) mit einer Pulsrate von 250 kHz und ein längerer Lauf (*PB2*) mit einer Pulsrate von 500 kHz.
Zufällige Koinzidenzen zwischen Trigger(-puls) und Signaldetektoren spielen keine Rolle, ihre Zählraten sind quasi Null und damit vernachlässigbar.
Daher ist es auch aussagekräftiger, die Koinzidenzzählrate zwischen den beiden Signaldetektoren zu betrachten, da sie direkt angibt, wie häufig es vorkommt, daß beide Signaldetektoren gleichzeitig ansprechen, was sie aufgrund des Lichtfeldes der stark abgeschwächten Quelle nur selten[42] tun sollten. Diese Rate liegt im Verhältnis zu den Bitzählraten ebenfalls sehr niedrig, s. Tab. 5.3.

Lauf	Pulsrate	Zählrate A [1/s]	Zählrate B [1/s]	Koinzidenzen AB [1/s]
PB1	250 kHz	2211 ± 44	1792 ± 41	13 ± 4
PB2	500 kHz	4989 ± 71	3989 ± 63	35 ± 6
PF	500 kHz	2050 ± 48	1654 ± 45	5 ± 2

Tabelle 5.3: Koinzidenz- und Einzel- Zählraten der Signaldetektoren für die Läufe mit der statistischen Einphotonenquelle

Beim längeren der beiden Läufe (*PB2*) zeigen sich zu Anfang des Laufes (im ersten und im dritten MB der Daten) drei Spitzen[43] im Verlauf der Einswahrscheinlichkeit, s. Abb. 5.6, die sich nicht mit statistischen Fluktuationen erklären lassen. Eine Autokorrelationsanalyse zeigt, daß es sich nicht um ein Problem der Datenaufnahme oder Speicherung handelt, anders als dies z.B. bei Problemen mit dem FIFO der Fall ist. Da es sich um Störungen am Anfang des Laufes handelt, lassen sich auch etwaige Überhitzungseffekte ausschließen, da diese erst nach längerem Betrieb auftreten sollten. Ein starker Anstieg des Umgebungsstörlichtes zu den entsprechenden Zeiten – z.B. durch Anschalten des Raumlichtes – wäre zwar bei den ersten beiden Spitzen durchaus denkbar, aber aufgrund des weniger abrupten Anstiegs und der längeren Dauer bei der dritten Spitze eher unwahrscheinlich.
Wahrscheinlich handelte es sich um eine elektrische (Kontakt)-Störung in einer der verschiedenen, meist aus mehreren Teilstücken zusammengesetzten, elektrischen Verzögerungsleitungen.

[42] Die Wahrscheinlichkeit für zwei Photonen im Lichtpuls ist um einen Faktor 10 kleiner als die für ein Photon, s. Abschnitt 3.1.2.

[43] Die erste Spitze erstreckt sich über eine Datenaufnahmezeitraum von ca. 31 s, die zweite über ca. 145 s und die dritte über ca. 415 Sekunden.

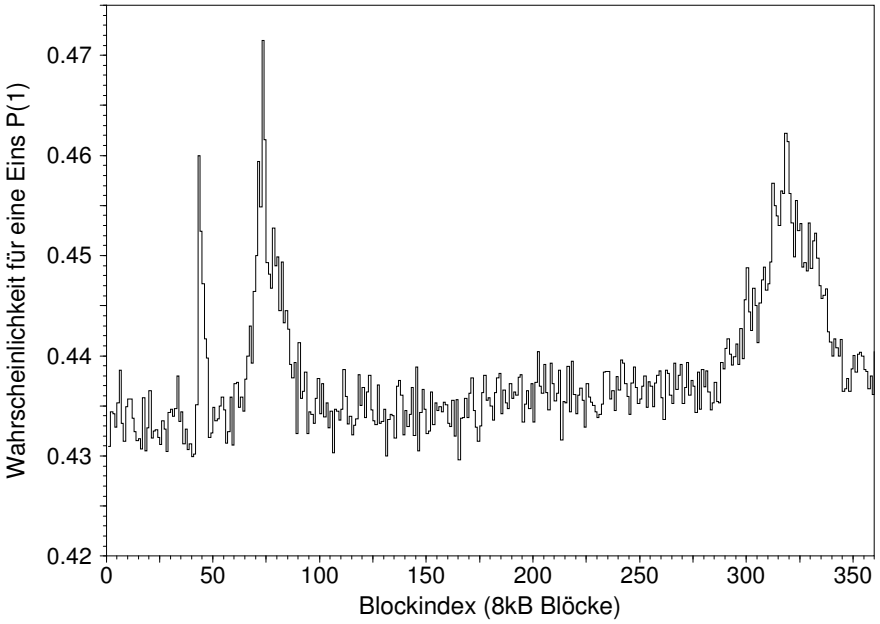


Abbildung 5.6: Verlauf der Wahrscheinlichkeit für eine Eins berechnet auf 8 kB Blöcken für den Anfang des Laufes *PB2*

5.2.3 Faseraufbau

Es wurde ein Lauf (*PF*) durchgeführt, die grundlegenden Daten finden sich in der Tabelle 5.3 des vorangehenden Abschnitts. Die niedrigere Bitrate beim faseroptischen Aufbau mit statistischer Einphotonenquelle gegenüber dem freistrahloptischen Aufbau wird durch die im Vergleich zum Strahlteilerwürfel stärkere Dämpfung des Faserkopplers verursacht, werden doch die zur Abschwächung des Lichtfeldes notwendigen Graugläser *ohne Faserkoppler* bestimmt.

Kapitel 6

Diskussion

In diesem Kapitel werden die wesentlichen Ergebnisse der Analyse der Zufallssequenzen summarisch dargelegt und diskutiert. Weiter werden die Vor- und Nachteile der verschiedenen Aufbauvarianten besprochen, Optimierungen für den Wirkbetrieb vorgeschlagen und alternative Aufbaumöglichkeiten für quantenoptische Zufallsgeneratoren aufgezeigt. Nach einem Abschnitt über die Resistenz der Zufallsgeneratoren gegen äußere Beeinflussungsversuche schließt das Kapitel mit der Darlegung verschiedener Arten von Signaturen bei quantenoptischen Zufallsgeneratoren.

6.1 Analyse der Zufallssequenzen

In diesem Abschnitt der Diskussion werden nur die wichtigsten Ergebnisse aus den statistischen Analysen der Zufallssequenzen präsentiert, für eine ausführliche Darstellung der Testdurchführung, der Ergebnisse und für eine Wertung der einzelnen Tests sei auf den Anhang, Abschnitt B verwiesen.
Bis auf wenige Ausnahmen werden alle Tests auf Rohdaten der quantenoptischen Zufallsgeneratoren durchgeführt. Dies hat zur Konsequenz, daß man bereits bei den einfachsten der Tests, nämlich dem Test auf die Häufigkeitsverteilung der Bitwerte, Abweichungen von der theoretisch erwarteten Normalverteilung findet. Dies verwundert nicht, da das nicht ideale Teilungsverhältnis des Strahlteilers bzw. Faserkopplers zusammen mit den unterschiedlichen Quanteneffizienzen der verwendeten Signaldetektoren erwartungsgemäß zu einer Wahrscheinlichkeit für den Bitwert Eins führt, die nicht gleich $p = 0{,}5$ ist. Der Test der Häufigkeitsverteilung der Bitwerte läßt sich glücklicherweise mit der aus den Rohdaten ermittelten Wahrscheinlichkeit parametrisieren. Interessanterweise finden sich aber auch dann noch Abweichungen von der theoretisch zu erwartenden Verteilung. Bei näherer Untersuchung stellt man fest, daß leicht unterschiedliche thermische Drifften der Quanteneffizienz der Detektoren hierfür verantwortlich sind. Reduziert man die Schwankungen der Labor- und damit auch der Detektortemperatur mit Hilfe einer Klimaanlage, nimmt dieser Effekt entsprechend ab. Er zeigt sich überdies nur bei Rohdaten; eine einfache Regularisierung beseitigt ihn. Für eine ausführlichere und auch graphische Darstellung sei auf Abschnitt B.2 verwiesen.
Die Tests nach FIPS 140-1 bestehen die quantenoptischen Zufallsgeneratoren mit freistrahloptischen Aufbau alle ohne Probleme, was angesichts der relativ weiten Akzeptanz-

schranken der Tests nicht überrascht. Lediglich Sequenzen der Zufallsgeneratoren mit faseroptischen Aufbau fallen aufgrund des Teilungsverhältnis des Faserkopplers (40:60 statt dem Idealwert von 50:50) erwartungsgemäß durch diese nicht parametrisierten Tests.
Die beiden sich von der Methodik her ähnelnden zweistufigen Kontingenz- und χ^2-Tests auf Bitpaaren liefern auch ähnliche Ergebnisse, wobei der Kontingenztest den Vorteil aufweist, daß bei ihm von vornherein nur aus der Stichprobe geschätzte Wahrscheinlichkeiten in die Berechnung der Statistik eingehen und somit eine Abweichung der Einswahrscheinlichkeit vom idealen Wert nicht extra berücksichtigt werden muß. Beide Tests sind trotz großer Stichprobengröße nicht dazu in der Lage, schwache Antikorrelationseffekte aufzudecken. Für eine detaillierte Darstellung wird auf die Abschnitte B.4, B.5 und B.6 des Anhangs verwiesen.
Überraschend sind die Ergebnisse der universellen Tests nach Maurer und Coron. Vor dem Test der Daten der quantenoptischen Zufallsgeneratoren werden beide Tests auf Datensätzen von kryptographisch starken Pseudozufallsgeneratoren getestet, hierbei fallen starke Fluktuationen der Testergebnisse auf, wenn für die Stichprobengrößen die vorgeschlagenen Minimalwerte verwendet werden. Es werden daher wesentlich größere Stichproben verwendet. Da der Verlauf der Teststatistik bei beiden Tests weitgehend gleichartig ist, kann man sich auf den universellen Test nach Coron beschränken.
Aufgrund der in Abschnitt 3.2.7 beschriebenen Einschränkungen der universellen Tests ist eigentlich nicht zu erwarten, daß sich Rohdaten physikalischer Zufallsgeneratoren damit gut testen lassen. In der Tat ist dies bei Rohdaten mit stark vom idealen Teilungsverhältnis abweichenden relativen Häufigkeiten von Einsen und Nullen auch so; bei nur geringen Abweichungen (ca. 2%) wirken sich diese nur gering auf die Testergebnisse aus, so daß erst bei großen Stichproben der Lauf als nichtideal abgelehnt wird.
Zur Evaluation, ob universelle Tests bei der gängigen Blockgröße von 8 Bit in der Lage sind, tieferliegende Defekte zu erkennen, wurde ein Lauf mit Antikorrelationen und nahe bei 50:50 liegendem Teilungsverhältnis getestet. Der Lauf wird von den universellen Tests *nicht* als problematisch erkannt; dies rührt offensichtlich daher, daß die Testergebnisse stärker durch die relativen Häufigkeiten der Bits als durch ihre statistischen Abhängigkeiten innerhalb der Bitblöcke bestimmt werden.
Ein ähnlich dominanter Einfluß der relativen Häufigkeiten auf die Ergebnisse statistischer Test besteht auch bei Komplexitätstests, s. Abschnitt B.8. Diese an sich sehr interessanten Tests, bei denen gerade in letzter Zeit wesentliche Fortschritte erzielt wurden [22, 79, 97, 98], weisen überdies noch das Problem auf, daß wenig über Erwartungswerte und Varianz der zugehörigen Komplexitätsmaße für *endliche* Stichproben bekannt ist. Zur Illustration der Problematik ist der Einfluß der relativen Häufigkeiten am Beispiel einfacher Komplexitätstests mit Hilfe eines vielverwendeten Kompressionsprogramms in Abschnitt B.8.2 dargelegt.
Als die am besten geeigneten Tests, um auch noch schwache Defekte in den Zufallssequenzen zu erkennen, haben sich ein- bis dreistufige Autokorrelationskoeffiziententests erwiesen. Mit jeder Stufe nimmt bei ihnen die Erkennungsempfindlichkeit für schwache (globale) Defekte zu.
Mit einstufigen Tests, die eine Visualisierung der Testergebnisse in Abhängigkeit von der Verschiebung erlauben, lassen sich stärkere Defekte gut erkennen, wie sie z.B. durch fehlerhaft arbeitende, elektronische Bauteile verursacht werden, s. Abschnitte 5.1.2.5,

B.9.1 und B.9.2.
Zweistufige Tests erlauben es, auch noch schwache globale Defekte aufzudecken, wie z.B. Antikorrelationen bei aufeinanderfolgenden Bits, s. Abschnitt B.9.3. Mit zweistufigen Autokorrelationskoeffiziententests lassen sich die Antikorrelationen erkennen, die keiner der oben erwähnten, anderen Tests nachweist. Zweistufige Tests haben sich auch als außerordentlich nützlich bei der Klärung der Ursachen der Antikorrelationen, die bei den Läufen mit der kompakten Datenaufnahme-Elektronik auftreten, s. Abschnitte B.10.1 und B.10.2, erwiesen.
Dreistufige Tests werden allgemein nur selten eingesetzt, da die Gefahr besteht, daß die Testergebnisse der dritten Stufe von Approximationsartefakten der ersten Stufe stark beeinflußt werden könnten. Bei den Autokorrelationskoeffiziententests der ersten Stufe ist die verwendete Approximation allerdings sehr gut, s. Abschnitt D und für die zweite Teststufe ist eine exakte Verteilung berechenbar, so daß dreistufige Tests gefahrlos durchgeführt werden können. Dies ist sogar mit relativ „kleinen“ Stichproben möglich, wie sich anhand von Testdatensätzen kryptographisch starker Pseudozufallszahlengeneratoren[1], verifizieren läßt, s. Abschnitt B.10.3.
Mit Hilfe dreistufiger Autokorrelationskoeffiziententests können auch noch sehr schwache Antikorrelationen in Läufen entdeckt werden, die mit Tests der zweiten Stufe nicht zu erkennen sind, s. Abschnitt B.10.4.

6.1.1 Gründe für das Auftreten der Antikorrelationen

In den folgenden Unterabschnitten werden die Gründe für das Auftreten unterschiedlich starker Antikorrelationen bei verschiedenen Läufen aufgeführt und ergriffene oder mögliche Gegenmaßnahmen erläutert.

6.1.1.1 Antikorrelationen bei den Läufen mit kompakter Datenaufnahme-Elektronik

Die in zweistufigen Autokorrelationskoeffiziententests deutlich zu erkennenden Antikorrelationen, s. Abschnitt B.9.3, treten in den Läufen mit der kompakten Datenaufnahme-Elektronik auch noch auf Zeitskalen auf, die lang im Verhältnis zur Detektionszeit, zu den Koinzidenzintervallen und zu den Detektortotzeiten sind. Daher ist nicht davon auszugehen, daß sie von den eigentlichen physikalischen Zufallsmechanismen, optischen Komponenten oder Detektoren herrühren, da sich z.B. Probleme mit Detektortotzeit lediglich auf relativ kurzen Zeitskalen auswirken können. Probleme innerhalb der Datenaufnahme-Elektronik bleiben deshalb als die wahrscheinlichste Ursache.
Da sich die Antikorrelationen auch nur auf Bitebene auswirken und abhängig von der Zeitspanne zwischen der Verarbeitung zweier Ereignisse sind, s. Abschnitt B.10.1 und B.10.2, lassen sich der FIFO-Baustein oder die PIC-interne Verarbeitung als Ursachen ausschließen. Da die Pulsverkürzung und die aus NAND-Gattern aufgebaute Koinzidenzschaltung aufgrund der Zeitskalen, auf denen sich die Probleme noch zeigen, ebenfalls nicht in Frage kommen, ist sehr wahrscheinlich die Zwischenspeicherung in dem aus zwei NAND-Gattern aufgebauten Flipflop und die Übergabe an den PIC, s. S. 78, dafür verantwortlich:

[1] In Abschnitt B.1 ist beschrieben, welche Pseudozufallsgeneratoren verwendet werden.

1. Die Entscheidung, das Flipflop aus zwei noch unbenutzten NAND-Gattern eines für die Koinzidenzsschaltung verwendeten integrierten Baustein aufzubauen, war zwar preiswert, aber leider kontraproduktiv, da die hierdurch verlängerten Signallaufzeiten zu länger andauernden instabilen Zuständen führen können.

2. Die Ansteuerung eines Flipflops durch ein stochastisches Signal führt ohne zusätzliche schaltungstechnische Maßnahmen häufig zu Problemen [31].

Um diese Probleme zu umgehen, wurde in den späteren Läufen auch die Elektronik, welche dem PIC vorgeschaltet ist, durch NIM-Einschübe ersetzt, vgl. Abschnitt 4.5.3. Wie in Abschnitt B.10.4 gezeigt, reduziert dies bei entsprechender Wahl der Signallängen auch tatsächlich die Antikorrelationen drastisch.

6.1.1.2 Antikorrelationen bei den Läufen mit hybrider Datenaufnahme-Elektronik

Wie in Abschnitt B.10.4 ausführlich dargelegt, weisen die beiden Läufe mit hybrider Datenaufnahme-Elektronik *LH1* und *LH4* in zweistufigen Autokorrelationskoeffiziententests Antikorrelationen von ähnlicher Größe auf wie die Läufe mit der kompakten Datenaufnahme-Elektronik. Bei den späteren Läufen mit kürzer gewählten Signal-Pulsen hingegen, lassen sich erst bei Einsatz von dreistufigen Tests schwache Spuren von Antikorrelationen aufdecken und bei den Läufen mit einer gepulsten statistischen Einphotonenquelle sind keine Antikorrelationen feststellbar.
Betrachtet man die Unterschiede der verschiedenen Läufe und berücksichtigt auch noch die Ergebnisse der Läufe mit verminderter Laserleistung und solcher mit verzögerter Datenaufnahme, s. Abschnitt B.10.1 bzw. B.10.2, so erkennt man, daß die Antikorrelationen von zeitlich kurz aufeinanderfolgenden Ereignissen erzeugt werden.
Bei den Läufen mit der *gepulsten* statistischen Einphotonenquelle treten daher schon allein deshalb keine Antikorrelationen auf, weil die Abstände zwischen den Pulsen 2 μs lang sind, 90 % der Pulse gar kein Photon enthalten und zudem nur jedes dritte Photon überhaupt registriert wird, da die Quanteneffizienz der Detektoren lediglich bei $\eta \approx 0{,}3$ liegt, s. Abschnitt 131.
Die gut erkennbaren Antikorrelationen bei den Läufen *LH1* bzw. *LH2*, sind durch die Länge des Bitwertsignals von 4,6 μs bedingt. Aufgrund der minimalen Dauer von 2 μs, welche der PIC benötigt um einen Bitwert einzulesen und in einem internen Register abzuspeichern, wirken sich Übernahmepulse eines nachfolgenden Ereignisses innerhalb der ersten 2 μs nur dann auf die Datenaufnahme aus, wenn ein Eins-Signal weniger als 800 ns (2 Befehlszyklen des PIC) nach einem Null-Signal kommt, dann wird statt einer Null eine Eins eingelesen[2]. In diesem Fall wird aber lediglich ein Bit mit dem Wert Null ausgelassen, was sich zwar auf die relativen Häufigkeiten (geringfügig) auswirkt, aber keine Korrelationen verursacht.
In den verbleibenden 2,6 μs ist allerdings sehr wohl eine Antikorrelationen erzeugende Beeinflussung der Datenaufnahme möglich. Folgende Fälle lassen sich unterscheiden:

[2]Der umgekehrte Fall kann natürlich nicht auftreten, da ein Null-Signal lediglich durch das Fehlen eines HIGH-Pegels auf der Signal-Leitung bei vorhandenem Übernahme-Signal erzeugt wird. Bei den anderen Kombinationen Eins-Eins und Null-Null, wird der richtige Wert für das Bit eingelesen, aber nur ein Bit generiert.

1. Der nachfolgende Übernahmepuls kommt so knapp nach dem vorangehenden Einlesevorgang, daß der PIC innerhalb der Länge des Bitwertsignals (des ersten Ereignisses) den Eingangsport noch abfragen kann.

2. Der nachfolgende Übernahmepuls kommt zwar noch innerhalb der Dauer des vorangehenden Bitwertsignals, aber so spät, daß bei Abfrage des Ports der Signalpegel des vorangehenden Signals nicht mehr anliegt.

Da die Abstände zwischen zwei Ereignissen abfallend exponentialverteilt sind[3], ist der zweite Fall wahrscheinlicher als der erste. Da nur einer der Detektoren das Bitwertsignal erzeugt und der andere Wert durch einen fehlenden HIGH-Pegel auf der Eingangsleitung repräsentiert wird, muß unterschieden werden, welche Bitwerte aufeinander folgen:

- Zeitlich eng aufeinanderfolgenden Null-Werte werden in beiden Fällen vom PIC als zwei Bits mit dem Wert Null eingelesen. Ist das erste Ereignis ein Null-Wert und es folgt ein Eins-Wert, so wird dies ebenfalls in beiden Fällen auch so vom PIC eingelesen. Ist es umgekehrt, ein Eins-Wert wird von einem Null-Wert gefolgt, werden aufgrund der Länge des Bitsignals im ersten Fall zwei Einsen generiert und im zweiten Fall erst eine Eins und dann eine Null generiert.

- Sind die beiden Ereignisse Eins-Werte, so wird der PIC im ersten Fall tatsächlich zwei Eins-Werte einlesen (allerdings aufgrund von *einem* Bitwertsignal), im zweiten Fall hingegen liest er ein Bit mit dem Wert Eins gefolgt von einem Bit mit dem Wert Null ein, da bei Abfrage des Eingangs-Ports der HIGH-Pegel des vorangehenden Bitwert-Signals nicht mehr anliegt und das zweite Ereignis aufgrund der Totzeit[4] des Gate&Delay-Generators keinen Bitwertpuls erzeugt.

Betrachtet man die dargelegten, verschiedenen Fälle zusammen, so sieht man, daß es bei eng aufeinanderfolgenden Ereignissen eine Tendenz zu alternierenden Werten gibt, die sich in den Ergebnissen der Autokorrelationskoeffizienztentests niederschlägt.
Bei den Läufen, die eine Länge des Bitwertsignal von $2\,\mu s$ verwenden, sollte diese Tendenz eigentlich nicht auftreten, entspricht diese Zeitspanne doch gerade der minimalen Bitfolgezeit, die noch vom PIC bewältigt werden kann. Dennoch treten Antikorrelationen auf, die allerdings *erheblich* schwächer ausgeprägt sind als im obigen Fall. Durch eine künstlich eingeführte Totzeit bei der Datenaufnahme, die sich, ähnlich wie in Abschnitt B.10.2 beschrieben, durch Umprogrammieren des PICs erzeugen ließe, sollten sich diese schwachen Antikorrelationen beseitigen lassen. Wählte man eine Länge von $2\,\mu s$ für die künstliche Totzeit, so sollten Einflüsse vom vorhergehenden Ereignis vollständig ausgeschlossen sein. Die geringe zusätzlichen Zeitverzögerung würde sich auch nur wenig auf die Bitrate des Generators auswirken, vgl. Bitraten von Lauf *LD1* und *LD2* in der Auflistung der Läufe in Anhang A.

[3] Bedingt durch die Totzeiten der Detektoren gilt dies allerdings erst für Abstände oberhalb der Totzeit (des Triggerdetektors) von 200 ns.

[4] Die Totzeit des Gate&Delay-Generators (EG&G, Typ 416 A) entspricht der eingestellten Pulsbreite zuzüglich $0{,}2\,\mu s$.

6.1.2 Zusammenfassende Bewertung der Analysen

Bei den statistischen Tests zur Analyse der erzeugten Zufallsbitströme zeigt sich, daß für den Test von Rohdaten physikalischer Zufallsgeneratoren Autokorrelationskoeffiziententests besonders gut geeignet sind. Insbesondere mehrstufige Autokorrelationskoeffiziententests eigenen sich hervorragend dazu, tieferliegende Defekte sowohl anschaulich aufzuzeigen als auch quantitativ nachzuweisen. Universelle Tests, die speziell für den Test physikalischer Zufallsgeneratoren vorgeschlagenen wurden, sind zumindest für die Tests von Rohdaten auf tieferliegende Defekte nicht geeignet, da ihre Ergebnisse zu stark von den relativen Häufigkeiten dominiert werden.
Bei den Läufen der quantenoptischen Zufallsgeneratoren mit einer kontinuierlich arbeitenden Einphotonenquelle auf Basis der parametrischen Fluoreszenz ließen sich mit Hilfe von zweistufigen Autokorrelationskoeffiziententests Antikorrelationen zwischen direkt aufeinanderfolgenden Bits aufdecken.
Die Antikorrelationen stammen allerdings nicht aus dem eigentlichen Zufallsprozeß – quantenoptische Zufallsgeneratoren haben also nicht per se Defekte – sondern es handelt sich um Artefakte, die von Problemen innerhalb der Datenaufnahme-Elektronik erzeugt werden. Ändert man die Datenaufnahme, was durch Einsatz der hybriden Datenaufnahme-Elektronik geschieht, so werden die bei eng aufeinanderfolgenden Detektionsereignissen auftretenden Antikorrelationen sehr stark unterdrückt. Durch Einführen einer künstlichen Totzeit innerhalb der Elektronik ließen sich die verbleibenden, allerdings nur noch sehr schwachen Defekte, vollständig eliminieren. Bei der gepulst betriebenen Photonenquelle, die stark abgeschwächte Pulse verwendet, zeigen sich keine Antikorrelationen, aufgrund der Größe der minimalen Abstände, die durch die Pulsfrequenz vorgegebenen sind. Bei Einsatz einer kontinuierlich betriebenen Quelle hätten sich dieselben Probleme mit der Elektronik gezeigt.
Für den praktischen Betrieb empfiehlt sich ein Neuentwurf der kompakten Datenaufnahme-Elektronik, speziell der Zwischenspeicherung der Bitwerte, zumal der Einsatz gleichartiger, schneller Detektoren, die sich scharf schalten lassen, möglichst kleine Koinzidenzfenster und die Zufallsgenerierung aus zusätzlichen Zufallsprozessen, s. Abschnitt 6.3.1, ohnehin Modifikationen verlangen.

6.2 Vergleich der verschiedenen Aufbauten

In den folgenden Unterabschnitten wird diskutiert, welche Vor- und Nachteile die verschiedenen Aufbauten der quantenoptischen Zufallsgeneratoren aufweisen.

6.2.1 Vor- und Nachteile der verwendeten Einphotonen-Quellen

Angesichts des relativ großen Aufwandes für die Erzeugung von möglichst guten Approximationen von Einphotonen-Anzahlzuständen ist natürlich die Frage zu stellen, ob der Aufwand gerechtfertigt ist oder eine Verwendung von abgeschwächten kohärenten Zuständen eventuell völlig ausreicht. Die Vorteile der quantenoptischen Lichtfelder, die mit Hilfe der parametrischen Fluoreszenz erzeugt werden, liegen auf der Hand:

- Es handelt sich bei ihnen um eine sehr gute Näherung für einen Einphotonen-Anzahlzustand.

- Diese approximative Einphotonen-Anzahlzustände sind relativ anschaulich und ermöglichen in Verbindung mit der Welcher-Weg-Entscheidung am Strahlteiler eine plausible, konzeptionell einfache und reine Realisierung eines elementaren Zufallsprozesses, quasi als eine Art „Quanten-Münzwurf“.

Leider bringt die Erzeugung der Einphotonen-Anzahlzustände mit Hilfe der parametrischen Fluoreszenz auch eine Reihe von Nachteilen mit sich:

- Ihre Anzahl ist aufgrund der niedrigen Erzeugungsrate bei kleiner Pumpleistungen recht begrenzt. Mit leistungsstarken Pumplasern läßt sich dieses Problem zwar umgehen, aber man erkauft es sich mit erhöhten Kosten, stärkeren Kühlanforderungen und einem vergrößerten Aufbau. Vorschläge, wie man diese Probleme mit Hilfe einer noch zu entwickelnden, kompakten Photonenpaarquelle umgehen könnte, werden in Abschnitt 6.4.4 dargelegt.

- Die erzeugten Photonen sind bei der von uns verwendeten Typ-I-Phasenanpassung recht breitbandig[5] und eine Beschränkung der Bandbreite, z.B. durch Interferenzfilter, senkt die Rate der Photonenpaare, die für die Experimente zur Verfügung stehen, leider drastisch.

- Photonenpaare mit Photonen, deren Zentralwellenlängen im Maximum der Detektionseffizienz liegen, lassen sich nur durch Einsatz von Pumplasern im UV-Bereich erzeugen.

- Verlustprozesse im Aufbau und insbesondere die Detektion mit nichtidealer Quanteneffizienz reduzieren die Rate der Signalphotonen, die bei Koinzidenz mit dem Triggerphoton registriert werden, stark.

Demgegenüber haben Photonenquellen auf Basis von stark abgeschwächten Lichtfeldern mit Poisson-Statistik folgende Vorteile:

- Sie sind sehr kostengünstig, kompakt, verbrauchen wenig Strom und lassen sich sowohl gepulst als auch kontinuierlich betreiben.

- Die Anzahl der zur Verfügung stehenden Photonen ist sehr groß.

- Es stehen Quellen zur Verfügung, die bei Wellenlängen um 700 nm abstrahlen, wo die Quanteneffizienz der Einzelphotonendetektoren ihr Maximum von $\eta_{max} = 0{,}7$ hat[6]. Überdies lassen sich die Verluste im Aufbau und nichtideale Quanteneffizienzen der Detektoren bei der Abschwächung des Lichtfeldes mitberücksichtigen, s. Abschnitt 6.2.1.3.

[5] Aus einer Divergenz des Pumplasers von 1 mrad um den Phasenanpassungswinkel des entarteten Falles $\lambda_s = \lambda_i$ resultiert eine spektrale Bandbreite des Fluoreszenzlichts von ca. 100 nm. Bei Verwendung eines nichtlinear optischen Kristalls für Typ-II-Phasenanpassung lassen sich durchaus geringere Bandbreiten von nur ca. 5 nm erreichen, s. [18], allerdings ist dies meist mit einer erheblich geringeren Photonenpaarrate verbunden.

[6] Zum Vergleich: In den Experimenten hatten die Detektoren bei der Zentralwellenlänge von 884 nm Quanteneffizienzen um $\eta = 0{,}3$, die Bitrate würde sich also bei Lichtfeldern um 700 nm – bei sonst gleichen Parametern – mehr als verdoppeln.

- Die Bandbreite des abgestrahlten Lichtfeldes kann sehr gering gewählt werden, was vorteilhaft für ein konstantes Teilungsverhältnis und eine gleichbleibende Detektionseffizienz ist.
- Das Lichtfeld kann gut kollimiert werden, was z.B. das Einkoppeln in Glasfasern vereinfacht, insbesondere wenn es sich um Einmodenfasern handelt.

Neben diesen Vorteilen weisen die stark abgeschwächten Poisson- Lichtfelder allerdings auch eine Reihe von Nachteilen auf:

- Es handelt sich nicht um *echte* Einphotonenanzahlzustände, das Auftreten von mehreren Photonen (innerhalb eines Pulses) läßt sich nicht ausschließen, sondern nur durch die Wahl einer hinreichend kleinen mittleren Photonenzahl unter einer (willkürlich) gewählten Wahrscheinlichkeit halten.
- Durch die notwendig starke Abschwächung weisen sie einen sehr hohen Vakuumanteil auf, d.h. der weitaus größte Teil der Pulse enthält überhaupt kein Photon, s. Abschnitt 3.1.2.
- Alterungseffekte[7] und Schwankungen im Ansteuerstrom der Quelle können zu einer Veränderung der Photonenstatistik der Quelle führen, was insbesondere bei Pulsen schwer zu erkennen ist.

Wägt man die oben aufgeführten Vor- und Nachteile der verschiedenen Arten von Einphotonenquellen gegeneinander ab, so spricht in erster Linie die konzeptionelle Reinheit für die Einphotonenquelle auf Basis der parametrischen Fluoreszenz, da sie dem Einphotonen-Anzahlzustand sehr viel näher kommt als die statistische Einphotonenquelle. In Verbindung mit dem Zufall generierenden Element erlaubt sie daher dem Ideal einer Folge von einzelnen „Quantenmünzwürfen“ möglichst nahe zu kommen.
Legt man allerdings den Schwerpunkt stärker auf die Einfachheit der experimentellen Realisierung, wie dies gerade bei einem kompakten, kostengünstigen, wartungsarmen Aufbau der Fall wäre, so ist der statistischen Einphotonenquelle der Vorzug zu geben.
Sicherlich könnte eine kompakte Einphotonenquelle, sei es auf Basis der parametrischen Fluoreszenz oder einem der neueren Ansätze, wie sie auf S. 34 erwähnt wurden, auch den experimentellen Aufwand (und die Kosten) beträchtlich senken, allerdings fordert die gute Approximation des Einphotonen-Anzahlzustandes *immer* den Einsatz möglichst effizienter Detektoren, während dies bei der statistischen Einphotonenquelle eine weitaus geringere Rolle spielt.

6.2.1.1 Beurteilung der Quellen für Einphotonen-Anzahlzustände

Wie in Abschnitt 5.1.2.3 dargelegt, weist die Einphotonenquelle auf Basis der parametrischen Fluoreszenz, die einen temperatur-phasenangepaßten Kaliumniobat-Kristall verwendet, stärkere Schwankungen in den Zählraten auf als die Quelle, mit dem winkelphasenangepaßten Lithiumjodat-Kristall. Auch wenn durch eine weitergehende thermische Isolierung[8] des Kaliumniobat-Kristalls diese Schwankungen noch etwas reduziert

[7] Sie sind weniger kritisch, da durch sie lediglich die Photonenrate und damit die Bitrate des Zufallsgenerators gesenkt wird.

[8] Hierbei müßten allerdings auch zusätzliche Glasfenster in den optischen Weg eingebracht werden, was insbesondere am Ausgangsfenster zwingend eine sehr gute Entspiegelung erforderte.

werden könnten, so ist doch insgesamt festzuhalten, daß der Aufbau mit einem winkelangepaßten Kristall für einen kontinuierlichen, produktiven Einsatz geeigneter ist, da nach einer einmaligen Fixierung der Komponenten im Aufbau keinerlei weitere Justage mehr notwendig wäre.
Bei allen experimentellen Aufbauten wurde für die Erzeugung von Photonenpaaren stets der nichtkollineare Fall der parametrischen Fluoreszenz mit Typ-I-Phasenanpassung verwendet. Um die beiden Photonen eines Paares räumlich zu trennen, mußte daher ein größerer Abstand zwischen Photonenpaarquelle und den Detektoren in Kauf genommen werden.
Eine Möglichkeit, den Aufbau kompakter zu gestalten, besteht darin, statt des nichtkollinearen Falles den kollinearen Fall bei der Erzeugung der Photonenpaare zu verwenden. Damit sich die Photonen eines Paares aber immer noch effizient trennen lassen, muß dann die parametrische Fluoreszenz mit Typ-II-Phasenanpassung eingesetzt werden, da bei ihr die Photonen eines Paares unterschiedlich polarisiert sind und sich somit durch polarisationsoptische Komponenten gut trennen lassen. Allerdings muß auch dafür gesorgt werden, daß das Pumplaserlicht, das nun dieselbe Ausbreitungsrichtung wie die Photonenpaare hat, von dem weiteren Strahlengang ferngehalten wird, was sich durch Filter, dichroitische Spiegel oder Prismen bzw. einer Kombination mehrerer dieser Komponenten erreichen läßt[9]. Diese Filterelemente sollten für das Photonenpaarlicht natürlich maximal durchlässig sein, da sich jeder Verlust der Photonen eines Paares quadratisch[10] auf die Koinzidenzzählrate auswirkt. Somit hat man hier eine Abwägung zwischen der Gesamtgröße des Zufallsgeneratoraufbaus und der erhöhten Anforderung an die Güte der optischen Komponenten zu treffen.

6.2.1.2 Echte versus zufällige Koinzidenzen

Anders als bei Experimenten, bei denen nur die Gesamtzahl von Koinzidenzereignissen in einem bestimmten Zeitintervall von Interesse ist, weshalb die Anzahl der zufälligen Koinzidenzen von der Zahl der insgesamt gemessenen abgezogen werden kann, ist dies bei den von uns untersuchten quantenoptischen Zufallsgeneratoren nicht möglich. Bei ihnen zählt jedes einzelne Ereignis und für ein einzelnes Bit läßt sich grundsätzlich nicht entscheiden, ob es durch eine echte oder eine zufällige Koinzidenz erzeugt wurde. Eine Messung der zufälligen Koinzidenzrate ist in diesem Fall lediglich eine Zusatzinformation, welche die „Güte“ des Quantenzufalls angibt und es erlaubt, abzuschätzen, wie hoch der Anteil derjenigen Bits ist, die durch andere Ursachen erzeugt wurden. Bei diesen Ursachen kann es sich um systembedingte handeln, wie z.B. thermisches Rauschen der Detektoren oder Umgebungslicht, oder auch um Manipulationen, wie z.B. gezieltes Einstrahlen von Licht, s. Abschnitt 5.1.2.6.
Der weitaus größte Anteil der zufälligen Koinzidenzen wird allerdings dadurch verursacht, daß aufgrund der nichtidealen Quanteneffizienzen oder sonstiger Verlustmechanismen nur einer der beiden beteiligten Detektoren ein Photon eines Paares registriert, während der andere Detektor das zugehörige Photon nicht registriert und daher auch

[9] Man bedenke, daß das Pumplaserlicht auch hinter dem Kristall immer noch mehr als 10 Größenordnungen intensiver ist als das Photonenpaarlicht.

[10] So senkt bereits eine nicht entspiegelte Glasplatte, die bei senkrechter Inzidenz 4% des einfallenden Lichtes reflektiert, die Koinzidenzrate um 16%.

noch für Photonen direkt[11] folgender Photonenpaare ansprechbar ist, was er sonst aufgrund der Detektortotzeit nicht wäre. Dies ist ein weiterer Grund, warum es bei quantenoptischen Zufallsgeneratoren die Einphotonenquellen auf Basis der parametrischen Fluoreszenz verwenden, besonders wichtig ist, Detektoren mit möglichst hoher Quanteneffizienz einzusetzen. Beim Einsatz einer Photonenquelle, die abgeschwächte Pulse emittiert und bei der das Triggersignal elektronisch generiert wird, sind die Auswirkungen einer nichtidealen Quanteneffizienz nicht so gravierend, s. Abschnitt 6.2.1.3.

6.2.1.3 Stark abgeschwächte Lichtquellen: Gepulst oder kontinuierlich?

Die Frage, welche Art von statistischer Einphotonenquelle besser für einen quantenoptischen Zufallsgenerator geeignet ist, gepulste oder kontinuierliche, hängt stark von der Art der verwendeten Detektoren ab, speziell von ihrer Totzeit.
Generell ist eine gepulste Quelle im Hinblick auf Totzeiteffekte unproblematischer, da man die Abstände der Pulse so wählen kann, daß das nächste Photon überhaupt erst nach einer Zeit, die der Summe der Gesamtlaufzeit des Photons und der Totzeit des Detektors entspricht, emittiert wird. Hiermit wird die Gefahr vermieden, daß sich (schwache) Antikorrelationseffekte aufgrund des Detektorsverhaltens einstellen.
Diese Wahl limitiert die Pulsrate allerdings auf $1/t_{tot}$, d.h. für die Totzeit[12] der Signaldetektoren von typ. 50 ns auf eine maximale Pulsrate von 20 MHz. Geht man von einer mittleren Photonenzahl pro Puls von 0,1 Photon aus, so bedeutet dies eine Photonenrate von max. 2 MHz. Bei nichtidealen Einzelquantendetektoren läge dann die entsprechende Bitrate je nach Wellenlänge des Lichtes zwischen 30% bis 70% der Photonenrate, was allerdings für viele Zwecke immer noch mehr als ausreichend wäre.
Höhere Zählraten lassen sich bei einer Photonenquelle auf Basis stark abgeschwächter Pulse dadurch erreichen, daß nicht das von der Quelle emittierte Lichtfeld auf eine mittlere Photonenanzahl von 0,1 Photon pro Puls abschwächt wird, sondern das *effektive* Lichtfeld, wie es der Detektor registriert. Es wird also die vom Idealwert abweichende Quanteneffizienz des Detektors quasi als ein Art „zusätzliches Grauglas“ betrachtet[13] und die Zählrate dementsprechend auf einen ideal gedachten Detektor angepaßt, d.h. man setzt $\eta = 1$. Dies würde bei den Wellenlängen, die in den vorgestellten Experimenten verwendet wurden, eine mehr als dreimal so hohe Zählrate ermöglichen.
Mit Blick auf die nicht wirklich identischen Quanteneffizienzen der beiden Detektoren wurde bei den Experimenten mit der Photonenquelle auf Basis abgeschwächter Pulse allerdings stets das restriktive Vorgehen gewählt. Will man aber eine möglichst hohe Zählrate erreichen und daher die Quanteneffizienz als Abschwächungsfaktor direkt berücksichtigen, so empfiehlt es sich, den größeren der beiden, i. a. leicht unterschiedlichen, Quanteneffizienzwerte der Signaldetektoren zu verwenden. Hierdurch wird in beiden Fällen das letztendlich im Detektor registrierte Lichtfeld unterhalb des gewählten Wertes für die mittlere Photonenzahl liegen.

[11] Dies bezieht sich auf die Verzögerungszeit, die bei der Bestimmung der zufälligen Koinzidenzen verwendet wird.

[12] Das bezieht sich auf die „aktiv gelöschten“ Detektoren, „passiv gelöschte“ mit einer Totzeit von 200 ns erlauben lediglich 5 MHz.

[13] Desgleichen lassen sich natürlich auch etwaige andere Verluste in dieser Form mit in den Abschwächungsfaktor einrechnen.

Die üblichen Sättigungsprobleme, die der Hersteller für große Detektionsraten angibt, treten bei gepulsten Quellen auch nicht in gravierender Form auf, da sich die Angaben immer auf kontinuierliche Lichtfelder mit Poissonstatistik beziehen und sich überdies die Photonen auf zwei Signaldetektoren verteilen. Allerdings muß bei hohen Zählraten mit zunehmenden Kühlanforderungen gerechnet werden, s. S. 154.
Für einen Aufbau mit gepulsten Poisson-Lichtfeldern empfiehlt es sich, scharf schaltbare Detektoren[14] einzusetzen, da ohnehin keine wirkliche Koinzidenzdetektion wie bei den Einphotonen-Anzahlzuständen auf Basis der parametrischen Fluoreszenz stattfindet. Durch die Verwendung solcher Detektoren ließe sich die externe Koinzidenz-Einheit einsparen.
Ein zusätzlicher Vorteil dieser Vorgehensweise besteht darin, daß sich hierdurch die Pulsraten noch über den oben angegebenen restriktiven Wert steigern lassen. Erzeugt man nämlich das Triggersignal für die Detektoren nicht direkt aus dem elektrischen Ansteuerpuls für die Lichtquelle, sondern immer nur dann, wenn ein hinreichend langer Abstand zum letzten Detektionsereignis eines der beiden Signaldetektoren gewährleistet ist, läßt sich ebenfalls vermeiden, daß totzeitbedingte Antikorrelationen bei der Bitgenerierung auftauchen, s.u.. Dies läßt sich z.B. durch eine Hemmung des Pulssignals durch das Ausgangssignal eines Monoflops, das vom Detektionsereignis gestartet wird, erreichen.
Kann man Einzelquanten-Detektoren mit einer geringen Totzeit, wie z.B. Photomultiplier einsetzen, wie dies bei Betriebswellenlängen der statistischen Einphotonenquelle bei Wellenlängen unterhalb von 650 nm gut möglich ist, so lassen sich auch kontinuierliche Quellen einsetzen, s. [63]. In diesem Fall, empfiehlt es sich allerdings dennoch, eine nachgeschaltete Elektronik zu verwenden, welche die Zählereignisse, die in einer Zeitspanne von der Größenordnung der Detektortotzeit registriert werden, verwirft, um (geringe) totzeitbedingte Antikorrelationseffekte zu vermeiden.
Wenn man kontinuierliche Quellen mit den bei unseren Experimenten verwendeten Avalanche-Photodiodendetektoren kombinieren wollte, muß auch noch berücksichtigt werden, daß bei diesen Detektoren Nachpulse auftreten können, vgl. S. 70 und Abschnitt 6.4.1.
Solche Nachpulse könnten sich durch schwache *Korrelationen* im Ausgangsbitstrom bemerkbar machen, da bei solch einem Aufbau alle Detektionsereignisse, bei denen einer der Detektoren anspricht, zur Generierung eines Bits führen. Dies war neben der Verminderung des Einflusses der Dunkelzählereignisse auch der Grund für den Einsatz einer gepulsten[15] statistischen Einphotonenquelle.
Die Nachpulsverteilung fällt glücklicherweise exponentiell mit der Zeit ab, daher würden Nachpulse kaum zu langreichweitigen Korrelationen im Ausgangsbitstrom führen, sondern nur das nächstfolgende Bit und eventuell in noch schwächerem Maße das übernächste betreffen.
Da sich durch Nachpulse generierte Ereignisse allerdings nicht von Ereignissen, die von in kurzem Abstand folgenden Photonen erzeugt wurden, unterscheiden lassen, empfiehlt

[14] Das aktuelle Modell der Detektoren von EG& G wird jetzt ohnehin standardmäßig mit einem zusätzlichem „Gate“-Eingang geliefert, der eine Beschränkung der Messung auf einen kurzen Zeitschlitz erlaubt.

[15] Zusätzlich wurden die Abstände zwischen den einzelnen Pulsen viel größer gewählt als die Zeit, in der nach einer Detektion Nachpulse auftreten können.

sich deshalb wieder eine künstlich eingeführte Totzeit[16] nach einer Detektion; wählt man für sie 500 ns, läßt sich hiermit die Nachpulswahrscheinlichkeit auf unter 10^{-4} senken.

6.2.2 Faseroptischer versus freistrahloptischer Aufbau

Wie bereits im vorangehenden Kapitel erwähnt, ist der faseroptische Aufbau für beide Arten von quantenoptischen Zufallsgeneratoren erheblich justagefreundlicher, insbesondere wegen der faseroptischen Anschlüsse der Signaldetektoren. Überdies kommt als weiterer Vorteil hinzu, daß die faseroptischen Zuleitungen einen kompakteren Aufbau und eine – gerade im Hinblick auf die beim längeren Betrieb anfallenden Kühlerfordernisse, s. S. 154, flexible Positionierung der Detektoren erlauben.
Als weiterer großer Vorteil eines faseroptischen Aufbaus ist der bessere Schutz gegen Streulicht, Licht von außen und optische Manipulationsversuche (s. Abschnitt 5.1.2.6) zu nennen, zumal die Faseranschlüsse an den Detektoren auch versiegelt werden können.
Für den produktiven Betrieb eines quantenoptischen Zufallsgenerators in Faserbauweise empfiehlt es sich, nicht einen Standard-Mehrmoden-Faserkoppler zu verwenden, sondern einen Mehrmoden-Faserkoppler in Auftrag zu geben, der so gefertigt wird, daß er in einem Bereich von ± 50 nm der Zentralwellenlänge der Photonen der Paare ein möglichst gleichmäßiges und nah bei 50:50 liegendes Teilungsverhältnis besitzt.
Bei einem Faseraufbau mit einer Photonenquelle auf Basis abgeschwächter Pulse ist die Fertigung solch eines Faserkopplers noch unproblematischer, da in diesem Fall durch Einsatz einer schmalbandigen Quelle das Teilungsverhältnis nur für einen verhältnismäßig kleinen Wellenlängenbereich bei 50:50 liegen muß. Da etwaige Verluste auch mit in die Abschwächung des Lichtfeldes eingerechnet werden können, ist es in diesem Fall auch gut möglich, einen Einmoden-Faserkoppler[17] zu verwenden, was im Fall einer auf Photonenpaaren basierenden Quelle nicht effizient möglich ist, s. a. Abschnitt 4.4.2.
Ein Nachteil der faseroptischen Aufbauten sind sicherlich die etwas höheren Einkoppelverluste[18] bzw. niedrigeren effektiven Quanteneffizienzen[19], bedingt durch die detektorinterne Ankopplung der Faseranschlüsse an den Detektorchip.
Diese zusätzlichen Verluste lassen sich beim Aufbau mit einer Photonenquelle auf Basis abgeschwächter Poisson-Lichtfelder wieder als zusätzlichen Abschwächungsfaktor des Lichtfeldes behandeln, s. S. 112, und sich somit quasi kompensieren; dies ist bei Verwendung einer Einphotonenquelle auf Basis der parametrischen Fluoreszenz natürlich nicht möglich.
Für einen Aufbau, der abgeschwächte Poisson-Lichtfelder als Photonenquelle verwendet, kann man daher den faseroptischen Aufbau klar als den besseren empfehlen, zumal bei ihm auch das Problem, wie man sicherstellt, daß beide Signaldetektoren auch tatsächlich den gleichen Raumwinkel sehen, gar nicht erst auftritt, s. Abschnitt 5.2.

[16] Gegenüber der normalen Totzeit bezieht sich die künstliche Totzeit auf *beide* Detektoren, da sonst wiederum Antikorrelationsprobleme auftauchen würden.

[17] Solche Einmoden-Faserkoppler werden von vielen Herstellern für Standardwellenlängen mit einem nach Kundenwunsch wählbarem Teilungsverhältnis gefertigt.

[18] Das bezieht sich auf freistrahloptische Aufbauten mit Detektoren ohne Faseranschlüsse und nicht auf unseren Aufbau, bei dem die Detektoren mit Faseranschlüssen für den Freistrahlaufbau „zweckentfremdet“ wurden.

[19] Gegenüber Detektoren ohne Faserstecker sind die Quanteneffizienzen um absolut 10% verringert, d.h. statt 56% beträgt die Quanteneffizienz bei der Referenzwellenlänge 820 nm nur 46%.

Überdies ließe sich durch den Einsatz von faseroptischen Abschwächungsgliedern in Verbindung mit einer flexiblen Stromregelung der Leucht- bzw. Laserdiode auch ein vollständig faseroptischer Aufbau von der Quelle bis zu den Detektoren realisieren, was sehr wünschenswert wäre, da dies eine bessere Abschirmung gegen äußere Einflüsse und einen noch kompakteren Aufbau ermöglichen würde.
Für den Fall der Einphotonenquelle auf Basis der parametrischen Fluoreszenz muß zwischen den Vorteilen des einfacheren, flexibleren Aufbaus und den höheren Verlusten abgewogen werden. Aufgrund der sehr positiven Erfahrungen mit dem faseroptischen Aufbau bei den Experimenten und der besseren Resistenz, die er gegen Beeinflussungsversuche von außen bietet, s. Abschnitt 6.5.2.1, wird aber auch hier eine Empfehlung für den Faseraufbau ausgesprochen. Falls eine noch zu entwickelnde, kompakte Einphotonenquelle, s. Abschnitt 6.4.4, Faseranschlüsse hätte, wäre ein Justage sogar völlig unnötig und ein quantenoptischer Zufallsgenerator könnte innerhalb kürzester Zeit aufgebaut werden.

6.2.3 Mittelwertschwankungen im Teilungsverhältnis

Wie in Abschnitt B.2 dargelegt, zeigen quantenoptische Zufallsgeneratoren, wie nahezu alle physikalischen Zufallsgeneratoren, keine gleichen relativen Häufigkeiten von Einsen und Nullen bei den unregularisierten Rohdaten. Dies war zu erwarten und stört bei vielen Anwendungen auch nicht.
Die relativen Häufigkeiten sind bei quantenoptischen Zufallsgeneratoren, die als Zufall generierendes Element einen Strahlteiler bzw. Faserkoppler verwenden und bei denen das Photon in den Ausgangsarmen mit Einzelquanten-Detektoren registriert wird, abhängig von den Eigenschaften der Photonen, des Strahlteilers, der Detektoren und teilweise auch der nachfolgenden Datenaufnahme und Verarbeitungselektronik[20].
Im folgenden werden die Ursachen und Auswirkungen der Eigenschaften der verschiedenen Bestandteile des Zufall generierenden Elementes für die verschiedenen Aufbauten diskutiert.

6.2.3.1 Photonen

Bei ihnen spielen vier Eigenschaften eine Rolle: Anzahl, Polarisation, Spektrum und die Ausbreitungsrichtung.
Die Anzahl der Photonen, die sich gleichzeitig, d.h. im Rahmen des Koinzidenzfensters, im[21] Zufall generierenden Element befinden, ist bei der abgeschwächte Pulse emittierenden Photonenquelle aufgrund der mittleren Anzahl von 0,1 Photon pro Puls nur alle 100 Pulse größer als eins, s. Abschnitt 3.1.2. Dementsprechend hat man auch eine geringe Rate von Ereignissen, bei denen beide Detektoren ansprechen, s. Abschnitt 5.2.2.
Bei der Einphotonenquelle auf Basis der parametrischen Fluoreszenz, läßt sich die Wahrscheinlichkeit für mehr als zwei Paare während dieses Intervalls, und somit zwei Si-

[20] Letzteres sollte möglichst vermieden oder falls dies nicht gelingt, zumindest minimiert werden. Es handelt es sich dabei allerdings nicht um ein Spezifikum quantenoptischer Zufallsgeneratoren.

[21] „Im“ bezieht sich hierbei auf das Koinzidenzintervall von 10 ns im Rahmen dessen Photonen als gleichzeitig registriert werden, die tatsächliche „Aufenthaltsdauer“ im Zufall generierenden Element ist beim freistrahloptischen Aufbau unter 1 ns und beim Faseraufbau unter 1,3 ns.

gnalphotonen im Zufall generierenden Element, aus der Anzahl der Photonen[22] die auf das Element fallen und der Poissonverteilung für diesen Fall berechnen:

$$P_{n=2}(\Delta_{koinz}) = \frac{(\rho t)^2}{2} \cdot e^{-\rho t},$$

hierbei ist ρ die mittlere Photonendichte, in Photonen pro Sekunde. Selbst für die relativ hohen Einzelzählraten des Laufes *LF1* erhält man $P_{n=2}(\Delta_{koinz}) = 6{,}9 \cdot 10^{-5}$. Die Wahrscheinlichkeit ist also auch bei hohen Zählraten recht klein[23]. Daher sind es also tatsächlich nahezu nur Einphotonen-Anzahlzustände, die beim Prozeß der Zufallsgenerierung verwendet werden.
Überdies muß noch berücksichtigt werden, daß aufgrund der verschiedenen Verlustmechanismen nur ein geringer Bruchteil der Photonen, die in das Zufall generierende Element gelangen, auch koinzident zu einem Detektionsereignis des Triggerdetektors sind und somit tatsächlich zu einem registrierten Zufallsereignis führen.
Die *Polarisationsrichtung* liegt bei der Einphotonenquelle auf Basis der parametrischen Fluoreszenz prinzipbedingt fest, s. Abschnitt 3.1.2.1, daher wird sich die Polarisationsrichtung bei den Aufbauten mit einem unpolarisierten Strahlteiler bzw. Faserkoppler nicht in Schwankungen bemerkbar machen können, sondern lediglich als konstanter Faktor in das Teilungsverhältnis eingehen. Eine Ausnahme hierzu stellen lediglich die anfänglichen Experimente mit dem über die Temperatur phasenangepaßten Kaliumniobat-Kristall und der Kombination aus $\lambda/2$-Plättchen und polarisierten Strahlteiler dar. Die bei dieser Variante auftretenden Probleme mit der veränderlichen Polarisation[24] wurden bereits in Abschnitt 5.1.2 ausgeführt und waren auch der Grund, warum die weiteren Experimente mit einem unpolarisierten Strahlteiler und einem über den Winkel phasenangepaßten, optisch nichtlinearen Kristall durchgeführt wurden.
Bei der Photonenquelle, die stark abgeschwächte Pulse aussendet, liegt die Polarisationsrichtung nicht fest. Bei den durchgeführten Experimenten wurde aufgrund des Einsatzes eines unpolarisierenden Strahlteilerwürfels bzw. eines Faserkopplers bei allen Experimenten mit dieser Photonenquelle auf den Einsatz eines Polarisators verzichtet. Die schwankende Polarisationsrichtung kann sich bei einer (geringen) Abhängigkeit der Teilungsrate von der Polarisationsrichtung grundsätzlich in einem dementsprechend variierenden Teilungsverhältnis bemerkbar machen. Bei einer zeitlich fluktuierenden, zufälligen Polarisation, wie sie bei der breitbandigen, gepulsten LED vorhanden ist, werden sich solche Fluktuationen allerdings wegmitteln, lediglich wenn die Polarisation des Lichtfeldes auf langsamen Zeitskalen eine Tendenz aufweisen würde, könnte sich dies minimal[25] auf die relativen Häufigkeiten im Ausgangsbitstrom auswirken.
Beim Faseraufbau wären die Auswirkungen der Polarisationseffekte noch weniger bemerkbar, da die verwendeten Mehrmodenfasern nicht polarisationserhaltend sind und der Faserkoppler auch nicht polarisationsempfindlich ist, was sich darin zeigt, daß sich die Teilungsraten bei leichtem Biegen der Zuleitungsfasern nicht ändern.

[22] Die Berechnung erfolgt wie in Abschnitt 5.1.4 beschrieben.

[23] Zumal noch hinzukommt, daß viele der Photonen durch die nichtideale Quanteneffizienz der Detektoren und die Einfügedämpfung des Faserkopplers nicht registriert werden.

[24] In diesem Fall verändert sich sowohl der von den Detektoren gesehene Anteil des auf das $\lambda/2$-Plättchen fallenden Bereich des Lichtspektrums als auch in Folge davon die Polarisation.

[25] Allerdings würden solche Effekte von den thermisch bedingten relativen Quanteneffizienzschwankungen der Detektoren überlagert.

Bei der *spektralen Charakteristik* sind die Photonen der abgeschwächte Pulse emittierenden Photonenquelle aufgrund des verwendeten Interferenzfilters auf einen relativ kleinen Wellenlängenbereich beschränkt, so daß sich Abhängigkeiten der Transmissions- und Reflexionskoeffizienten oder der Quanteneffizienz der Detektoren von der Wellenlänge nur geringfügig bemerkbar machen. Eine Abhängigkeit zwischen Ausbreitungsrichtung und abgestrahltem Spektrum besteht nicht.
Dies ist bei Photonenpaaren anders. Bei den beiden Photonen eines Paares sind aufgrund der Phasenanpassung, s. Abschnitt 3.1.2.1, S. 38, die jeweiligen Ausbreitungsrichtungen und Spektren miteinander gekoppelt. Überdies sind die Spektren, die bei guter Justage der Blenden und Detektoren gleiche Zentralwellenlängen und -breiten haben, relativ breitbandig, s. S. 109; die Frage inwieweit sich dies in Verbindung mit einem wellenlängenabhängigen Teilungsverhalten des Strahlteilers auf den Ausgabebitstrom auswirkt, wird in Abschnitt 6.2.3.4 erörtert.

6.2.3.2 Strahlteiler

Der erste Aufbau des quantenoptischen Zufallsgenerators, mit einem polarisierenden Strahlteiler und davor im Strahlengang befindlichem, drehbaren $\lambda/2$-Plättchen, erlaubt zwar ein anfängliches Einstellen des Teilungsverhältnisses auf 50:50, ist aber unnötig kompliziert, ohne wesentliche Vorteile mit sich zu bringen. Neben den bereits in Abschnitt 4.4.1.1 erwähnten Problemen, schwankt das Teilungsverhältnis allein aufgrund des Einflusses des thermischen Verhaltens der Detektoren ohnehin immer leicht, so daß ein einmaliges, festes Einstellen des Teilungsverhältnisses nicht sinnvoll ist. Da sich das $\lambda/2$-Plättchen auch mit Hilfe eines Schrittschaltmotors rechnergesteuert verdrehen läßt, könnte man natürlich das Teilungsverhältnis auch zeitlich variabel einstellen, um es nahe dem Idealwert zu halten. Dies ist allerdings aufwendiger, als nachträglich die zur Verfügung stehenden, effektiven Regularisierungsmethoden, s. Anhang, Abschnitt G anzuwenden und führt auch nicht zu besseren Ergebnissen. Überdies besteht die Gefahr[26], daß die so generierten Zufallsströme „verschlimmbessert“ werden, indem man auch legitime Abweichungen vom idealen Teilungsverhältnis „herausregelt“. Da sich auch beim Einsatz des unpolarisierenden Breitband-Strahlteilers hinreichend gute Teilungsverhältnisse erzielen lassen, s. Abschnitt B.2, sind die aufwendigeren Lösungen nicht notwendig. Das Teilungsverhältnis des Faserkopplers ist, s. a. Abschnitt 4.4.2, S. 72, erheblich stärker wellenlängenabhängig. Wieder betrifft dies die breitbandige Einphotonenquelle auf Basis der parametrischen Fluoreszenz stärker als die schmalbandigere statistische Einphotonenquelle.

6.2.3.3 Detektoren

Die Detektoren haben in dem Wellenlängenbereich, der für die Experimente gewählt wurde, eine linear mit der Wellenlänge stark abfallende Quanteneffizienz[27], s. Abschnitt 131, Abb. 4.6. Grundsätzlich unterscheidet sich das Detektionsverhalten der beiden Signaldetektoren in Abhängigkeit von der Wellenlänge nur durch eine konstante Differenz.

[26] Durch hinreichend große Stichproben zur Berechnung des Ist-Wertes des Teilungsverhältnisses sollte sich diese Gefahr allerdings weitgehend vermeiden lassen.

[27] Die Wellenlängenabhängigkeit der Detektoren läßt sich daher durch parallele Geraden, mit gleicher Steigung approximieren, wobei lediglich der y-Abschnitt für den jeweiligen Detektor spezifisch ist.

Daher geht es bei hinreichend[28] genauer Justage auch nur statisch in die Teilungsrate ein. Allerdings machen sich aufgrund von unterschiedlichem thermischen Verhalten der Detektoren auf Zeitskalen im Minutenbereich langsame, schwache Drifteffekte im Teilungsverhältnis bemerkbar, s. Abschnitt B.2. Diese Unterschiede lassen sich durch Aufbauvarianten vermeiden, die nicht zwei, sondern nur einen Signaldetektor verwenden, s. Abschnitt 6.4.5.

6.2.3.4 Auswirkungen der Abhängigkeiten

Die oben aufgeführten Abhängigkeiten des Teilungsverhältnisses wirken sich nicht in gleichem Maße auf den Strom der Zufallsbits aus. Sieht man von den anfänglichen Experimenten[29] mit dem polarisierenden Strahlteiler ab, spielt die Polarisiationsrichtung des Lichtes keine Rolle.
Die Wellenlängenabhängigkeit des Teilungsverhältnisses macht sich allerdings durch die Wellenlängenabhängigkeit des Strahlteilers bzw. insbesondere des Faserkopplers bei den Aufbauten mit einer Einphotonenquelle auf Basis der parametrischen Fluoreszenz bemerkbar. Da sich zusätzlich noch die langsamen Fluktuationen durch das unterschiedliche thermische Verhalten der Signaldetektoren bemerkbar[30] machen, lassen sich die Auswirkungen der Bandbreite der Photonen auf das Teilungsverhältnis nur ermitteln, indem man Stichproben aus dem Zufallsstrom entnimmt, bei denen die thermisch bedingten Effekte nicht so stark ausgeprägt sind.
Um den Einfluß der spektralen Verteilung der Photonen auf das Teilungsverhältnis zu veranschaulichen, empfiehlt sich das folgende einfache Modell:
Jede einzelne Zufallsentscheidung werde durch einen Wurf einer Münze[31] dargestellt, wobei die Wahrscheinlichkeiten für Kopf oder Zahl, bzw. Null oder Eins, nicht gleich sein müssen.
Die Zufallsentscheidung eines breitbandigen Photons an einem Zufall generierenden Element mit wellenlängenabhängigen Teilungsverhältnis läßt sich dann durch die *zufällige* Wahl und den anschließenden Wurf einer Münze aus einem Sack mit einer großen Menge[32] unterschiedlicher Münzen mit Zurücklegen modellieren; hierbei wird davon ausgegangen, daß die einzelnen Münzwürfe unabhängig voneinander geschehen.
Bei einer statischen Verteilung[33] der Münzen wird das Verhältnis der Einsen und Nullen in einer längeren Ausgabesequenz gerade durch den Mittelwert über die Wahrscheinlichkeiten aller Münzen gegeben sein. Da die Varianz wesentlich von der Verteilung der

[28] Die beiden Signaldetektoren sollten einen identischen Raumwinkel sehen, da dieser beim parametrischen Fluoreszenzlicht den Spektralbereich des Lichtes bestimmt, der auf den Detektor fällt.

[29] Da bei diesen Experimenten bereits bei der Datenaufnahme eine Regularisierung durchgeführt wurde, sind etwaige Fluktuationen nicht in den Datensätzen erkennbar.

[30] Hierbei handelt es sich allerdings nicht um einen wellenlängenabhängigen Effekt.

[31] „Randeffekte“ werden vernachlässigt.

[32] Da das Spektrum kontinuierlich ist, müßte man rein formal eine unendliche Zahl von Münzen vorsehen, da sich allerdings das Teilungsverhältnis nur mit endlicher Genauigkeit bestimmen läßt, ist eine endliche Zahl von Münzen ausreichend.

[33] Bei den ersten Experimenten mit dem temperatur-phasenangepaßten Kaliumniobat-Kristall wäre genau dies bereits auf kleinen Zeitskalen im Sekundenbereich nicht gegeben gewesen, was der Grund für die Verwendung des Lithiumjodat-Kristalls bei den nachfolgenden Experimenten war. Bei diesen führt lediglich das bereits erwähnte unterschiedliche Temperaturverhalten der beiden Signaldetektoren zu einer Drift im Teilungsverhältnis auf längeren Zeitskalen.

verschiedenen Münzen abhängt, lassen sich unterschiedliche „Münzsäcke“ durch sie voneinander unterscheiden.
Für eine genaue Analyse wie sich die spektralen Bandbreiten der beiden verschiedenen Lichtfelder im Zusammenspiel mit dem Strahlteiler bzw. Faserkoppler und den Detektoren bei der Zufallsgenerierung bemerkbar machen, muß eine quantitative Modellrechnung durchgeführt werden. Hierbei werden folgende vereinfachende Annahmen gemacht:

- Es wird angenommen, daß die Signaldetektoren und der Triggerdetektor (bei der parametrischen Fluoreszenz) den gleichen Raumwinkel und das gleiche[34] Wellenlängenspektrum sehen, was bei idealer Justage der Fall sein sollte.

- Verluste durch Absorption u. ä. werden nicht berücksichtigt, die wellenlängenabhängige Quanteneffizienz der Detektoren und das Filterverhalten des RG 830 Farbglases vor den Detektoren hingegen schon. Sättigungseffekte beim Triggerdetektor werden nicht berücksichtigt.

- Die Wellenlänge des Pumplasers wird als monochromatisch angesehen, so daß sich aus der Wellenlänge des Triggerphotons exakt die Wellenlänge des Signalphotons berechnen läßt.

- Zufällige Koinzidenzen werden nicht berücksichtigt.

In diesem Fall ist die Wellenlängenverteilung der beiden korrelierten Photonen jeweils durch eine Gaußverteilung gegeben, wobei die beiden Wellenlängen aufgrund der Phasenanpassung miteinander korreliert sind.
Bei der Photonenquelle auf Basis abgeschwächter Pulse sind die oben aufgeführten Annahmen natürlich unnötig, bei ihr wird eine Gaußsche Spektralverteilung angenommen, mit gleicher Zentralwellenlänge (884 nm) aber einer unterschiedlichen spektralen Breite, die bei ihr durch den Interferenzfilter gegeben ist.
In Abb. 6.1 sind die Spektralverteilungen der Signalphotonen der gepulsten Quelle und der Quelle auf Basis der parametrischen Fluoreszenz dargestellt; man erkennt deutlich die erheblich größere Bandbreite[35] der Photonen bei der parametrischen Fluoreszenz.
Der bei der Zufallsgenerierung auf Basis der parametrischen Fluoreszenz verwendete approximative Einphotonenanzahlzustand wird durch die Detektion eines der Photonen des Paares hergestellt. Diese Detektion durch den Triggerdetektor führt aufgrund der Energiekorrelation der beiden Photonen eines Paares, s. Abschnitt 3.1.2.1, und der wellenlängenabhängigen Quanteneffizienz des Triggerdetektors zu einer entsprechend verschobenen spektralen Verteilung des Signalphotons. Zur Illustration sind in Abb. 6.2 die errechneten Spektren des Signalphotons ohne und mit (idealer) Koinzidenzdetektion des Triggerphotons dargestellt. Zusätzlich ist auch noch die Verteilung der Detektorereignisse bei der Detektion des Signalphotons in Abhängigkeit von der Wellenlänge dargestellt.

[34] Die genaue räumliche Korrelation der Photonen eines Paares (aufgrund der Impulserhaltung bei der Phasenanpassung) innerhalb dieses Raumwinkels wird also nicht weiter berücksichtigt.
[35] Es wird von einer typischen Bandbreite von 100 nm ($1/e$-Wert) ausgegangen.

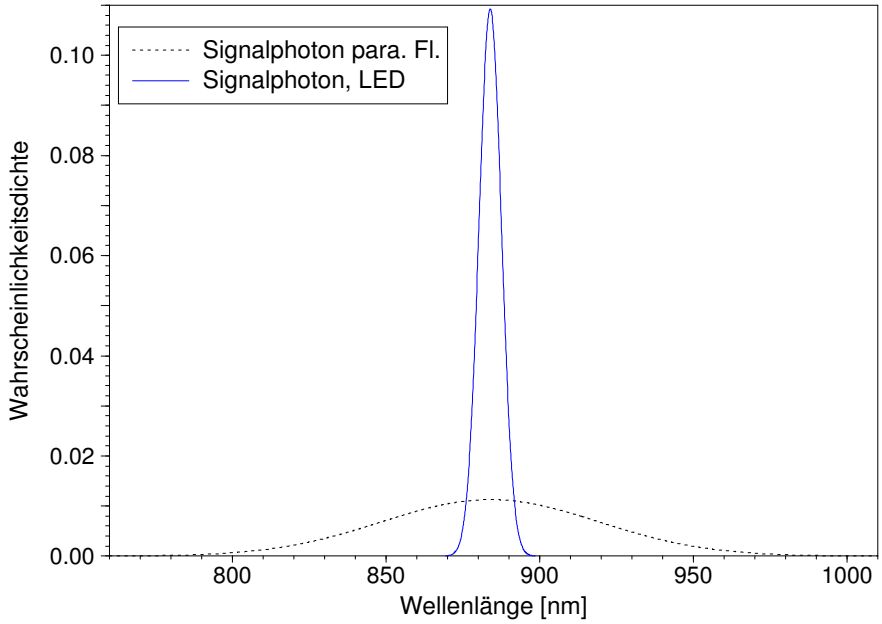


Abbildung 6.1: Spektrale Verteilung der Signalphotonen für die beiden verwendeten Einphotonenquellen

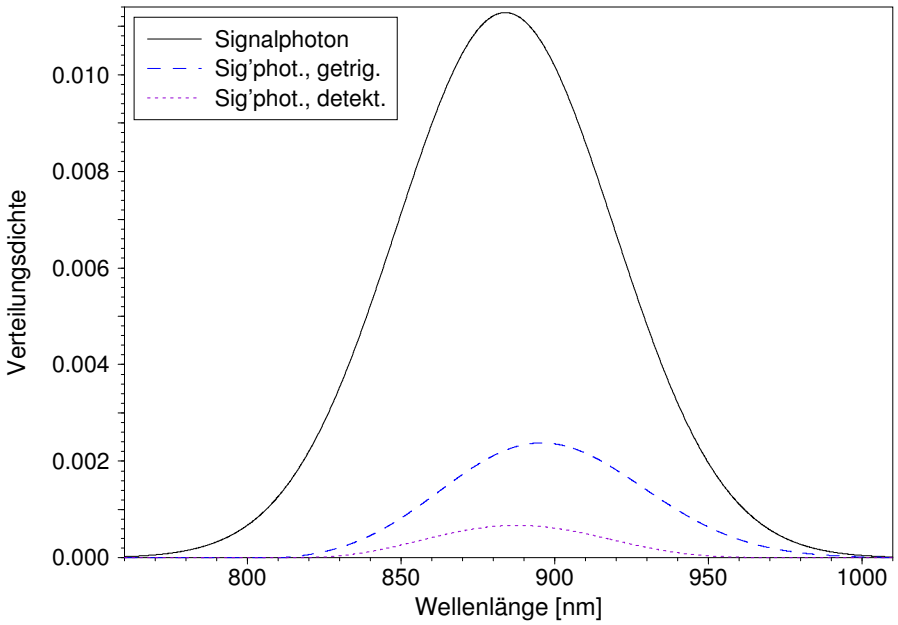


Abbildung 6.2: Signalphotonen-Verteilung, Verteilung der Signalphotonen mit koinzidentem Triggerdetektor-Signal, Verteilung der Detektionsereignisse (ohne Strahlteiler)

Man sieht hier neben dem starken Einfluß einer nichtidealen Quanteneffizienz auf die Detektion auch noch gut die Wellenlängenabhängigkeit der Quanteneffizienz, die sich in einer Verschiebung der Verteilung des Spektrums[36] der „registrierbaren“ Einphotonen-

[36]Nur die Verteilung der Signalphotonen ist in Abb.6.2 auf Eins normiert, die anderen beiden Verteilungen sind es aufgrund der in die Berechnung eingehenden Quanteneffizienzen der Detektoren nicht.

anzahlzustände bemerkbar macht.
Betrachtet man allerdings die Verteilung der Detektorereignisse nach einer Detektion des Signalphotons, so macht sich die gleichartige Wellenlängenabhängigkeit der verwendeten Detektoren bemerkbar, indem die Verteilung wieder zur Zentralwellenlänge hin verschoben ist. Aufgrund der wellenlängenabhängigen Quanteneffizienz und der RG 830 Farbgläsern vor den Detektoren hat diese Verteilung aber eine geringere Breite als das Signalphotonenspektrum. Von den verwendeten Zufall generierenden Elementen ist beim Faserkoppler das Teilungsverhältnisses am stärksten wellenlängenabhängig, daher sollten sich bei ihm mögliche Auswirkungen der unterschiedlichen Bandbreiten der Photonenquellen am deutlichsten zeigen. In der Tat zeigt sich im Vergleich der beiden Läufe[37] mit Faseraufbau und Klimaanlage im Raum *LF2* und *PF*, daß die Wahrscheinlichkeit für eine Eins beim Lauf *PF* im Mittel um drei Prozentpunkte höher lag als beim Lauf *LF2* mit der Quelle auf Basis der parametrischen Fluoreszenz. Dies bedeutet allerdings lediglich, daß sich die über die verschiedenen Teilungsverhältnisse je Wellenlänge gemittelten Gesamtteilungsverhältnisse unterscheiden; da dieser Wert weitgehend[38] statisch ist, ist dies unproblematisch. Interessanter ist die Frage, ob die Anzahl der Einsen aufgrund ei-

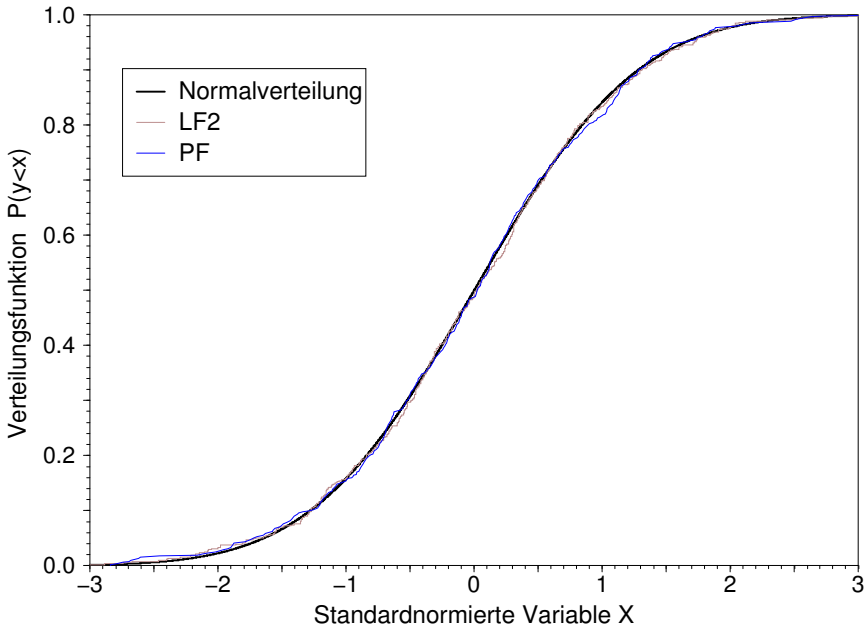


Abbildung 6.3: Verteilung der standardnormierten Werte der Einsen in 8 kB Stichproben innerhalb der letzten 4 MB der Läufe *LF2* und *PF*

nes wellenlängenabhängigen Teilungsverhältnisses stärker um den Mittelwert schwankt, als bei einem festen Teilungsverhältnis. Auf den ersten Blick scheint dies der Fall zu sein, blättert man auf die Seite 157 und betrachtet sich in Abb. B.4 die empirische Ver-

[37] Die beiden Läufe sind gut miteinander vergleichbar, weil am Zufall generierenden Element selbst keine Änderungen vorgenommen werden müssen; dies ist ein Vorteil des faseroptischen Aufbaus, bei dem die Detektoren direkt an den Faserkoppler angeschlossen sind. Die thermisch bedingten Schwankungen der Detektoren wirken sich hier auch nur gering aus, überdies werden bei den Stichproben für die weiter unten erwähnten Tests möglichst stationäre Teile der Läufe untersucht.

[38] Die leichten Änderungen rühren wieder vom Temperaturverhalten der Detektoren her.

teilungsfunktion der Einsen in einem „weitgehend stationären“ Abschnittes des Laufes *LO2* mit einem Strahlteiler, sieht man, daß die Varianz tatsächlich größer als theoretisch erwartet ist. Allerdings liegt dies lediglich daran, daß dieser Abschnitt eben doch nicht wirklich stationär ist, was man im zeitlichen Verlauf der Einswahrscheinlichkeit, s. Abb. B.2 auf S. 155, an einem leichten Anstieg erkennen kann. Es ist dieser Anstieg, der zu dieser Abweichung bei der Verteilung führt. In Abb. 6.3 sind für die jeweils letzten 4 MB der beiden Läufe *LF2* und *PF* die Verteilungsfunktionen[39] der standardnormierten Anzahl der Einsen in Stichproben von 8 kB dargestellt.
Wie man deutlich erkennt, sind beide Verteilungen sehr nah bei der theoretischen Verteilung. Dies bedeutet, daß sich die differierenden Bandbreiten nicht in unterschiedlich starken Schwankungen um den jeweiligen Mittelwert bemerkbar machen, sondern in beiden Fällen die jeweilige Verteilung dem von der theoretischen Kurve beschriebenen Verhalten eines Strahlteilers mit festem Teilungsverhältnis folgt, wobei es sich hierbei, wie oben erwähnt, um einen Mittelwert handelt.

6.2.4 Zusammenfassende Bewertung der Aufbauten

Alle in der vorliegenden Arbeit untersuchten Aufbauten quantenoptischer Zufallsgeneratoren eignen sich prinzipiell gut zur Generierung von Zufallsbitströmen.
Bei allen wirken sich allerdings die leicht unterschiedlichen thermischen Eigenschaften der Einzelquantendetektoren in Form einer Drift der relativen Häufigkeiten auf langen Zeitskalen aus. Durch eine aktive Kühlung lassen sich diese Effekte jedoch sehr stark abmildern, überdies werden sie von einer beim produktiven Einsatz i. a. nachfolgenden mathematischen Regularisierung der Rohdaten ohnehin eliminiert. Innerhalb der Rohdaten ließen sie sich lediglich dann ganz vermeiden, wenn für beide möglichen Ereignisse ein gemeinsamer Signaldetektor verwendet würde, s. Abschnitt 6.4.5.
Der Einfluß der unterschiedlichen Bandbreite der verwendeten Lichtfelder zeigt sich auch beim Einsatz des relativ stark wellenlängenabhängigen Faserkopplers nur im Teilungsverhältnis und nicht in stärkeren Fluktuationen.
Für einen möglichst kompakten, kostengünstigen Aufbau empfehlen sich die Aufbauten mit einer stark abgeschwächten LED, s. a. Abschnitt 6.3.3, während für eine konzeptionell nah am Ideal der kontrollierten Zufallsentscheidung liegenden Realisierung einem Zufallsgenerator auf Basis der parametrischen Fluoreszenz der Vorzug gegeben werden sollte.
Für Zufallsgeneratoren, die mit Photonen-Anzahlzuständen arbeiten, sind alle Verlustprozesse sehr kritisch, daher sollten für den praktischen Einsatz von quantenoptischen Zufallsgeneratoren auf Basis der parametrischen Fluoreszenz die Zentralwellenlänge der Photonen noch näher beim Quanteneffizienzmaximum der Detektoren liegen; dies ließe sich mit einer noch zu entwickelnden kompakten Photonenpaarquelle erreichen, s. Abschnitt 6.4.4. Weiter hat es sich bei diesen Generatoren gezeigt, daß eine Winkel-Phasenanpassung nicht nur einfacher, sondern auch aufgrund der besseren Stabilitätseigenschaften einer thermischen Phasenanpassung überlegen ist.
Unabhängig von der verwendeten Photonenquelle sprechen die gegenüber den freistrahloptischen Aufbauten erheblich einfachere Justage, die besseren Schutzmöglichkeiten gegen äußere Beeinflussung, die größere Flexibilität in der geometrischen Gestaltung des

[39] Die Berechnung geschieht wie in Abschnitt B.2.

Aufbaus und die damit einhergehende Erleichterung der Kühlung der Detektoren im Dauerbetrieb grundsätzlich für einen Aufbau in Faseroptik.

6.3 Vorschläge weiterer Varianten quantenoptischer Zufallsgeneratoren

Angesichts der verhältnismäßig niedrigen Photonenraten, die sich mit Einphotonenquellen auf Basis der parametrischen Fluoreszenz erreichen lassen, ist es lohnenswert, zu überlegen, ob sich die Rate, mit welcher der Zufallsgenerator Bits generiert, auch anderweitig steigern läßt als durch eine bloße Erhöhung der Rate, mit der die Quelle Einphotonenzustände aussendet. In den folgenden Unterabschnitten werden daher eine Reihe von Aufbauvarianten quantenoptischer Zufallgeneratoren vorgeschlagen, die eine Erhöhung der Rate erlauben, und ihre Vor- und Nachteile erörtert.
Grundsätzlich stehen folgende Möglichkeiten zur Verfügung, die von der Quelle ausgesandten Photonen möglichst effizient zu nutzen:

1. Das Triggerphoton wird ebenfalls zur Generierung von Zufallsbits verwendet, s. Abschnitt 6.3.1.1.

2. Die negativ-exponentialverteilten Abstände zwischen den Photonenpaaren werden zusätzlich zur Generierung von Zufallsbits verwendet, s. Abschnitt 6.3.1.2.

3. Durch die Verwendung von Photonenpaaren bei der *Welcher-Weg-Entscheidung* wird die effektive Detektionswahrscheinlichkeit der Detektoren erhöht, s. Abschnitt 6.3.2.

Überdies lassen sich auch jeweils die erste und die zweite bzw. die zweite und die dritte Möglichkeit kombinieren.

6.3.1 Mehrfachgeneratoren

Neben dem Minimieren von Absorption und ungewollter Reflektion an optischen Komponenten, die bei den vorgestellten Aufbauten erreicht werden, z.B. durch den nichtkollinearen Aufbau, empfiehlt es sich auch, möglichst *alle* physikalisch vorhandenen Zufallsmechanismen zur Generierung von Zufallsbits zu verwenden. Aus den verschiedenen Zufallsmechanismen können dann in solch einem Mehrfachgenerator mehrere zufällige Bitströme gleichzeitig erzeugt werden.
Neben der reinen Erhöhung der Bitrate gegenüber einem einfachen Zufallsgenerator lassen sich mit Hilfe der zusätzlichen Bitströme, die i. a. unterschiedlich große Raten aufweisen werden, auch noch weitere Verbesserungen der Eigenschaften des Zufallsgenerators erzielen, z.B. durch Verwendung der verschiedenen Ströme innerhalb eines mathematischen Regularisierungsverfahrens. Auch für das Einfügen von Signaturen in den Ausgangsbitstrom eines Zufallsgenerators, s. Abschnitt 6.5.4, ließen sich solche Mehrfachgeneratoren gut einsetzen.

6.3.1.1 Zufallsgenerierung mit dem Triggerphoton

Beim Zufallsgenerator auf Basis der parametrischen Fluoreszenz verhalten sich die beiden Photonen eines Paares symmetrisch zueinander, d.h. im Prinzip ist es beliebig, welches der beiden Photonen man auf den Strahlteiler schickt und welches man auf den Triggerdetektor fallen läßt. Daher läßt sich das bisher nur zum Triggern benutzte Photon auch noch zusätzlich zur Zufallsgenerierung nutzen, indem man es ebenfalls auf einen Strahlteiler mit zwei in den Ausgängen stehenden Signaldetektoren schickt, s. Abb. 6.4. Seine Funktion als Triggerphoton kann es immer noch erfüllen, nur daß jetzt zwei „Triggerdetektoren“ vorhanden sind, was überdies von Vorteil ist, da Sättigungseffekte abgeschwächt werden und somit höhere Zählraten möglich sind.

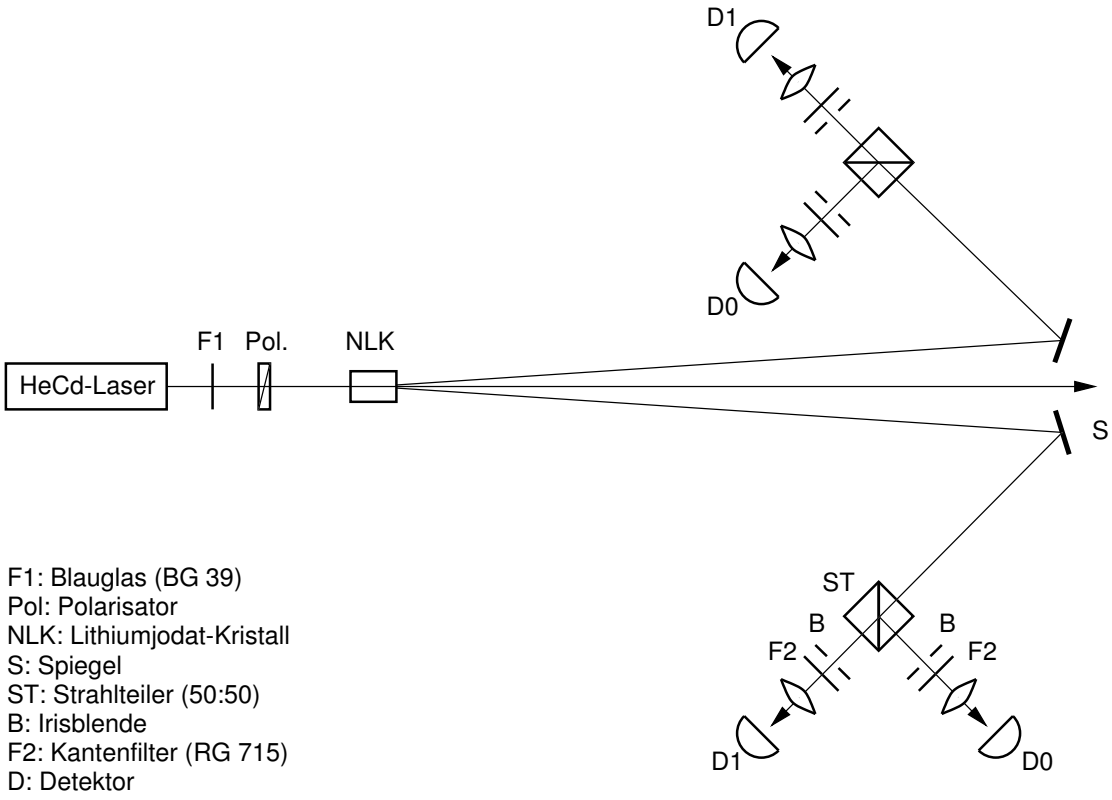


Abbildung 6.4: Der quantenoptische „Doppel“-Zufallsgenerator

Aufgrund der vollständigen Symmetrie, kann natürlich umgekehrt das bisher als Signalphoton benutzte Photon als Triggerphoton für das andere dienen. Auf diese Weise hat man den Generator quasi verdoppelt und erzielt dementsprechend auch eine doppelt[40] so hohe Bitrate. Allerdings wird die Erhöhung der Bitrate solch eines „Doppel-Zufallsgenerators“ gegenüber der einfachen Version durch einen zusätzlich notwendigen Detektor erkauft.
Bei einer Version für den praktischen Einsatz wird man daher versuchen, die Kosten

[40]Bei einer Variante mit vier Detektoren ist aufgrund von geringeren Sättigungseffekten auch noch mit einer leichten Erhöhung der Rate zu rechnen; bei Einsatz von nur jeweils einem Detektor bei den beiden Zufall generierenden Elementen geht dieser Vorteil natürlich wieder verloren.

möglichst gering zu halten und nur jeweils einen Detektor für „Signal-“ und „Triggerphoton“ verwenden, s. Abschnitt 6.4.5.
Anders als bei den bisher vorgestellten Aufbauten ist allerdings darauf zu achten, daß eine möglichst weitgehende Symmetrie der beiden Zufall generierenden Prozesse gegeben ist, so sollten in diesem Fall auch die Detektoren alle gleichen Typs sein, auch wenn dies bedeutet, daß Totzeiteffekte der Detektoren nun stärker berücksichtigt werden müssen, s. a. Abschnitt 6.4.1. Die Verschränkung der beiden Photonen eines Paares bzgl. Energie (d.h. Wellenlänge) und Impuls(richtung) bringt es außerdem mit sich, daß die zwei Bits, die im Falle eines Koinzidenzereignisses generiert werden, nicht statistisch unabhängig voneinander sind. Durch einen möglichst symmetrischen Aufbau lassen sich diese Effekte zwar minimieren, aber ganz beseitigen lassen sie sich nicht. Andererseits läßt sich diese – teilweise steuerbare – statistische Abhängigkeit gerade dazu nutzen, Korrelationssignaturen, s. Abschnitt 6.5.4.2, in den kombinierten Ausgangsbitstrom einzufügen.

6.3.1.2 Zufallsgenerierung aus den zeitlichen Abständen

Im Idealfall sind die Intervalle zwischen aufeinanderfolgenden Photonenpaaren aufgrund der Poissonverteilung der Paare abfallend exponentialverteilt [74]. Allerdings verursachen Totzeiteffekte der Detektoren bei kleinen Intervallen eine Veränderung dieser Statistik. Da sich diese Totzeiteffekte in gleicher Weise bei allen Detektionsereignissen bemerkbar machen, läßt sich ihr Einfluß beseitigen, indem man das bereits auf S. 26 bei den radioaktiven Zufallsgeneratoren beschriebene Verfahren zur Generierung eines Zufallsbits aus den Abständen zwischen Ereignissen verwendet: Je nachdem, welches der beiden Intervalle zwischen drei aufeinanderfolgenden Ereignissen größer ist, wird eine Null oder eine Eins generiert. Durch dieses Vorgehen, mit dem die unendlich vielen Möglichkeiten der Abstände auf eine binäre Entscheidung abgebildet werden, wird auch gleich der Einfluß von langsamen Schwankungen der Pumpleistung des Lasers oder von Drifteffekten bei den Detektoren, wie sie durchaus in der Praxis auftreten, s. S. 154, auf die relativen Häufigkeiten von Einsen und Nullen eliminiert.
Diese Methode der Erzeugung zufälliger Bits läßt sich mit der bereits bestehenden kombinieren. Die Ausgangsbitrate[41] des Generators erhöht sich gegenüber der einfachen Variante um nahezu 50%. Auch ein Einsatz bei einem Generator auf Basis abgeschwächter Lichtfelder als zusätzlicher oder alleiniger Zufallgenerierungsmechanismus, s. Abschnitt 6.3.3, ist möglich.

6.3.2 Quantenoptische Zufallsgeneratoren mit Hong-Ou-Mandel-Aufbau

Bisher wurden für die Zufallsgenerierung immer nur einzelne Photonen am Strahlteiler betrachtet, interessant für die Zufallgenerierung ist allerdings auch das Verhalten von Photonenpaaren am Strahlteiler [15, 17, 73]. Im Jahre 1987 haben C. K. Hong, Ou

[41] Es ließe sich auch aus der Größe jedes *einzelnen* Intervalls zwischen zwei aufeinanderfolgenden Ereignissen direkt ein Zufallsbit generieren, indem man es mit dem Median der Verteilung der Intervalle (unter Berücksichtigung der Totzeit) vergleicht und je nachdem, ob die Länge des Intervalls kleiner oder größer als dieser Wert ist, eine Null oder eine Eins generierte. Auf diese Weise ließe sich sogar die Ausgangsbitrate um nahezu 100 % steigern, wobei sich allerdings Schwankungen des Medians, z.B. durch eine leicht veränderliche Laserleistung, direkt auf die relativen Häufigkeiten auswirken.

und L. Mandel [56] erstmals experimentell untersucht, was passiert, wenn man die Photonen eines Photonenpaares, die mit Hilfe des nichtkollinearen Falles der parametrischen Fluoreszenz räumlich getrennt erzeugt wurden, auf die beiden Eingänge eines Strahlteilers lenkt und an diesem überlagert[42], s. a. Abb. 6.5. Hiermit gelang es ihnen, den von H. Fearn und R. Loudon [42] vorhergesagten Effekt nachzuweisen, daß die beiden Photonen, wenn sie ununterscheidbar[43] voneinander sind, im verlustfreien Idealfall *immer zusammen als Paar* aus einem der beiden Ausgänge austreten. In den beiden Ausgängen befindliche Detektoren werden daher bei exakter Überlagerung der Photonen eines Paares am Strahlteiler eine minimale – im Idealfall[44] sogar verschwindende – Koinzidenzrate ausweisen.
Bei einer Laufzeitdifferenz[45] zwischen den Photonen, die größer als die Kohärenzzeit der Einzelphotonen ist, wird die Koinzidenzrate hingegen nicht verschwinden, sondern der Hälfte der Photonenpaarrate entsprechen, da in diesem Fall die Photonen zufällig und unabhängig voneinander in einen der beiden Ausgänge laufen und vier der acht möglichen Fälle[46] zu einem Koinzidenzereignis führen. Indem man also die optischen Weglängen, welche die Photonen bis zum Strahlteil zurücklegen, so einstellt, daß die Koinzidenzrate minimiert wird, läßt sich der Aufbau einjustieren.
Der Aufbau von Hong et al. läßt sich aufgrund der zufälligen Welcher-Weg-Entscheidung der Photonen am Strahlteiler ebenfalls als Zufallsgenerator, im folgenden „HOM-Zufallsgenerator“ genannt, verwenden; bei ihm ist es nicht die Entscheidung eines einzelnen Photons am Strahlteiler, sondern eines Photonenpaares, die zur Zufallgenerierung[47] verwendet wird. Die Berechnungen zu einigen der weiter unten erwähnten Eigenschaften des HOM-Generators finden sich im Anhang, Abschnitt C.
In Abb. 6.5 ist eine platzsparende Aufbauvariante[48] dieses Zufallsgenerators mit gefaltetem Strahlengang dargestellt. Die Verschiebeeinheit an einem der Spiegel ist für den Feinabgleich der optischen Weglängen notwendig. Damit durch sie nicht die Überlagerung der Moden am Strahlteiler[49] verschlechtert wird, muß bereits beim Aufbau auf möglichst exakte Symmetrie geachtet werden, so daß keine großen Verschiebungswege notwendig sind.

[42] C. K. Hong et al. vermaßen hiermit die Subpikosekunden-Zeitintervalle zwischen den Photonen des Paares.

[43] Hierzu müssen die Photonen dieselbe Polarisationsrichtung aufweisen, die gleiche Wellenlängenverteilung besitzen, die Moden am Strahlteiler exakt überlagert werden und die beiden Photonen zeitgleich am Strahlteiler ankommen.

[44] Idealfall bedeutet: perfekte Überlagerung der Moden, 100 % Quanteneffizienz, keine Detektortotzeit und keine Verluste bei den Photonen eines Paares durch Absorption, Streuung usw..

[45] Die Ununterscheidbarkeit der Photonen läßt sich natürlich auch auf andere Weise aufheben, z.B. durch eine Änderung der Polarisationrichtung eines der Photonen um 90°.

[46] Es gibt zwei Möglichkeiten, die beiden unterscheidbaren Photonen auf die Eingänge zu verteilen und jeweils vier Möglichkeiten, wie sie in die Ausgänge laufen können, s. Anhang, Abschnitt C.

[47] Bei einem realen Aufbau tragen allerdings aufgrund von geringer Modenfehlanpassung oder Verlusten durchaus auch einzelne Photonen zur Zufallsgenerierung bei.

[48] Die Abbildung ist natürlich nicht maßstäblich, auch sind die Winkel zwischen Pumpstrahl und den Ausbreitungsrichtungen der Photonen übertrieben groß gezeichnet.

[49] Um dies zu vermeiden, kann man z.B. auch trompetenzug-artige Spiegelkonstruktionen im Strahlengang verwenden, bei denen eine Verschiebung die Ausfallrichtung des Strahls nicht ändert. Allerdings sind die ausfallenden Strahlenbündel bei der parametrischen Fluoreszenz ohnehin divergent, daher erscheinen solche Konstruktionen hier nicht vorteilhaft, zumal die bei ihnen in beiden Armen zusätzlich notwendigen Spiegel die Verluste erhöhen.

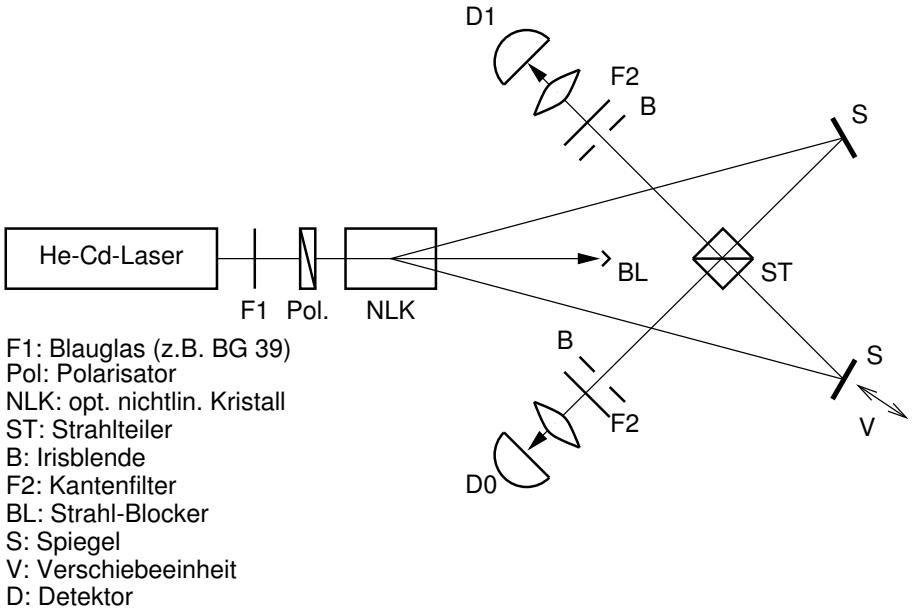


Abbildung 6.5: Eine Aufbauvariante des „HOM-Zufallsgenerators“ mit gefaltetem Strahlengang

Der HOM-Zufallsgenerator bietet folgende Vorteile:

- Das Teilungsverhältnis ist im verlustlosen Idealfall symmetrisch[50], auch wenn der Strahlteiler unterschiedlich große und von der Wellenlänge abhängige Transmissions- und Reflexionskoeffizienten besitzt.

- Die effektive Detektionseffizienz[51] wird auf $P = 1-(1-\eta)^2 = 2\eta-\eta^2 = \eta\cdot(2-\eta)$ erhöht, d.h. für $\eta = 0{,}3$ erhöht sich die effektive Quanteneffizienz um 70 % auf einen Wert von $\eta_{eff} = 0{,}51$.

- Es wird kein Triggerdetektor mehr verwendet, so daß sich Sättigungseffekte aufgrund hoher Einzelphotonenraten weniger stark bemerkbar machen, als beim Zufallsgenerator, der mit Einphotonen-Anzahlzuständen arbeitet [52].

- Durch Einführen einer hinreichend großen Weglängendifferenz zwischen den Wegen, welche die beiden Photonen bis zum Strahlteiler zurücklegen, läßt sich kontinuierlich zwischen unterscheidbaren und ununterscheidbaren Photonen wechseln, so daß sich also zwischen der Zufallsentscheidung eines Photonenpaares und zweier voneinander unabhängiger Zufallsentscheidungen der einzelnen Photonen eines

[50] Abweichungen vom idealen Teilungsverhältnis durch unterschiedliche Detektoreffizienzen werden hierdurch allerdings nicht berührt.

[51] Hierbei wird davon ausgegangen, daß beim gleichzeitigen Auftreffen von zwei Photonen auf einen Detektorchip die Detektionswahrscheinlichkeit für jedes einzelne Photon sich nicht verändert. Dann entspricht die effektive Detektionswahrscheinlichkeit gerade der komplementären Wahrscheinlichkeit zu dem Fall, daß beide Photonen kein Ereignis auslösen (Term $(1-\eta)^2$).

[52] Bei diesem sieht der Triggerdetektor prinzipbedingt eine doppelt so hohe Photonenrate wie ein jeder der Signaldetektoren; bei einem gut einjustierten HOM-Zufallsgenerator sehen beide Signaldetektoren die gleiche Photonenpaarrate.

Paares am Strahlteiler wählen läßt[53]. Dies ließe sich eventuell zum Einfügen zeitlich veränderbarer Signaturen, s. Abschnitt 6.5.4, nutzen.

Wenn der HOM-Zufallsgenerator auf Ununterscheidbarkeit der Photonen einjustiert ist, sollte die Rate der Koinzidenzen zwischen den beiden Detektoren minimal sein, im Idealfall gleich Null. Lediglich etwaige Verluste, wenn z.B. eines der beiden Photonen eines Paares absorbiert wurde, und Streulicht sorgen für einen „Untergrund“ von einzelnen Photonen, der sich in einer geringen Koinzidenzrate zwischen den beiden Signaldetektoren bemerkbar macht, s. Abschnitt C. Ereignisse, bei denen beide Detektoren ansprechen, lassen sich aber immer von einer nachfolgenden Datenaufnahmeelektronik verwerfen.
Zusätzliche Lichteinstrahlung, z.B. bei einem Manipulationsversuch, s. Abschnitt 5.1.2.6, würde zu einer Zunahme dieses „Untergrundes“ führen und wäre somit sowohl an einer erhöhten Einzelzählrate also auch an einer erhöhten Koinzidenzrate zu erkennen.
Auch beim HOM-Zufallsgenerator lassen sich die zeitlichen Abstände zwischen aufeinanderfolgenden Photonenpaaren zum Generieren eines zusätzlichen Zufallsstromes verwenden. Eine Verwendung beider Photonen zur Generierung je eines Zufallstromes ist hingegen nicht sinnvoll, weil hierdurch ein wesentlicher Vorteil des HOM-Zufallsgenerators – die bessere effektive Quanteneffizienz bei der Detektion – wieder verloren ginge.
Leider hat der HOM-Zufallsgenerator auch eine Reihe von Nachteilen:

- Da die Generierung der Photonenpaare einer Poisson-Statistik folgt, beim HOM-Generator die beiden Photonen eines Paares auch als Paar gemeinsam an der Welcher-Weg-Entscheidung teilnehmen und somit keine Triggermöglichkeit mehr zur Verfügung steht, wird de facto eine kontinuierlich arbeitende, sehr schwache Poisson-Photonenquelle verwendet. Daher muß bei HOM-Generator darauf geachtet werden, daß sich die Detektortotzeiten nicht im Zufallsstrom bemerkbar machen[54].

- Die notwendige Überlagerung der beiden Eingangsmoden am Strahlteiler bzw. Faserkoppler bedingt einen präzisen Abgleich der optischen Wege[55] der beiden Photonen eines Paares. Insbesondere ein freistrahloptischer Aufbau des HOM-Zufallsgenerators wird daher empfindlich auf Erschütterungen reagieren. Zwar ist der HOM-Aufbau bei weitem nicht so empfindlich wie ein Interferometer, da bei ihm nicht die Wellenlänge, sondern die Kohärenzlänge, die für den Abgleich der Weglängen relevante Größe ist, aber eine Einstellmöglichkeit der optischen Wege mit Hilfe einer elektronisch ansteuerbaren Verschiebeeinheit und einer Regel-Software wird dennoch notwendig sein, um die beiden optischen Wege während des Betriebs gleich zu halten.

- Ein faseroptischer Aufbau ist zwar auch beim HOM-Zufallsgenerator grundsätzlich möglich, aufgrund der notwendigen, möglichst guten Überlagerung der beiden

[53] Die beiden Extremfälle liegen bei gleichen optischen Weglängen bzw. bei Weglängen, die sich um eine Kohärenzlänge oder mehr unterscheiden, dazwischen hat man eine Überlagerung beider Fälle.
[54] Dies ließe sich wieder durch Einführen einer künstlichen Totzeit erreichen, in der keinerlei Detektionsereignisse aufgenommen werden.
[55] Der Abgleich muß auf ca. 1/10 der Kohärenzlänge der Einzelphotonen, typ. auf 1 μm, erfolgen.

Moden am Faserkoppler ist aber eine Verwendung von Einmodenfasern und Polarisationskontrollern zwingend erforderlich[56]. Einmodenfasern bringen allerdings erheblich höhere Einkoppelverluste der Photonen eines Paares in die jeweiligen Fasern mit sich, vgl. [128], so daß die effektive Photonen(paar)rate gesenkt wird, was den Vorteil einer erhöhten effektiven Quanteneffizienz leicht wieder zunichte macht.

Bei einem realen Aufbau werden sich gegenüber dem Idealfall noch zusätzlich folgende Effekte bemerkbar machen:

- Absorption, Streuung und Restreflektion an Oberflächen können dazu führen, daß ein Photon des Paares fehlt, und daher nur eine einfache Welcher-Weg-Entscheidung als Zufallsprozeß dient, die allerdings auch nicht minder zufällig ist.

- Fehlüberlagerung der Moden am Strahlteiler, führen zu einer Beimischung der einfachen Welcher-Weg-Entscheidung einzelner Photonen.

- Thermische Drifteffekte der Quanteneffizienz der Detektoren führen zu einer langsamen Drift im Teilungsverhältnis, die sich stärker auswirkt als beim einfachen Zufallsgenerator, s. Anhang, Abschnitt C.

6.3.3 „Integrierter“ Aufbau quantenoptischer Zufallsgeneratoren

Für einen weitgehend „integrierten“ Aufbau[57] eines quantenoptischen Generators eignet sich eine stark abgeschwächte Poisson-Photonenquelle besser als eine Einphotonenquelle auf Basis der parametrischen Fluoreszenz, und zwar sowohl aufgrund der erheblich niedrigeren Kosten als auch der einfacheren Justage.
Bei solch einem „integrierten“ Aufbau würde sich insbesondere ein abgewandelter, vereinfachter Aufbau des Zufallsgenerators von Takeuchi et al. [122] anbieten. Der Zufall würde direkt aus der Poisson-Statistik des abgeschwächten Lichtfeldes generiert werden, was den Vorteil mit sich brächte, daß kein zusätzliches Zufall generierendes Element und auch nur ein Detektor notwendig wäre und somit die Integration vereinfacht würde. Anders als bei Takeuchi et al., die mit ihrem Zufalls-Pulser nach der Detektion des Lichtes elektrische Pulse erzeugen, die möglichst genau der Poisson-Statistik bei einer bestimmten mittleren Pulsrate entsprechen sollten, empfiehlt es sich, hier lediglich Zufallbits aus den Intervallen von jeweils drei Detektionsereignissen zu erzeugen, vgl. S. 26 bzw. Abschnitt 6.3.1.2, da dies die Anforderungen an die Stabilisierung auf mittleren und langen Zeitskalen der Lichtquelle senkt und keine explizite Berücksichtigung von Totzeiteffekten verlangt. Der grundsätzliche Aufbau solch eines integrierten quantenoptischen Zufallsgenerators ist in Abb. 6.6 skizziert.

[56] Zudem empfiehlt es sich, auch die Fasern möglichst kurz zu halten, zu fixieren und thermisch durch eine entsprechende Isolierung des Aufbaus zu isolieren, damit unterschiedliche Einwirkungen auf die beiden optischen Wege durch äußere Einflüsse möglichst gering gehalten werden.

[57] Das „integriert“ steht in Anführungszeichen, weil hierbei in erster Linie an eine Integration in ein makroskopisches Modul gedacht ist und nicht in einem Chip, da dies angesichts der momentanen Einzelquantendetektor-Technik (im Geiger-Modus betriebene APDs) noch nicht praktisch realisierbar ist.

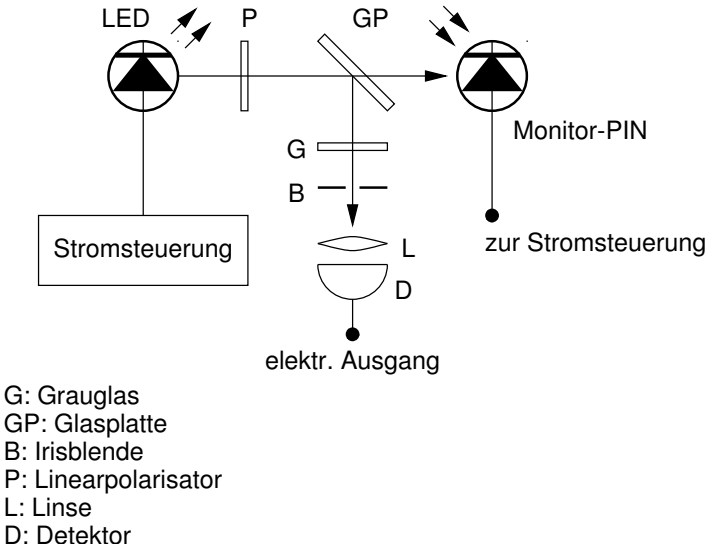


Abbildung 6.6: Vorschlag für einen einfachen, kostengünstigen „integrierten“ Aufbau eines quantenoptischen Abstand-Zufallsgenerators

Als Lichtquelle empfiehlt sich eine LED oder eine Laserdiode, deren Lichtfeld wieder durch ein Grauglas stark abgeschwächt wird. Das Grauglas dient zur groben Grundabschwächung, die Feinregulierung der Intensität auf einen Zustand mit niedriger mittlerer Photonenzahl erfolgt über die Reflektion an der Glasplatte[58] und die den Diodenstrom regelnde Monitor-PIN-Diode. Die Verwendung einer Monitor-PIN-Diode[59] zur nachführenden Stromregelung empfiehlt sich aus folgenden Gründen:

- Bauteilstreuungen bei den LEDs können leicht ohne größeren Justageaufwand, rein elektronisch, ausgeglichen werden.

- Es lassen sich hiermit Defekte leichter lokalisieren; bei einem Abbruch der Zufallsgenerierung läßt sich so erkennen, ob die LED oder der Signaldetektor eine Funktionsstörung hat.

- Alterungseffekte der LED im laufenden Betrieb, die zu einer Abnahme der Lichtleistung und damit der Bitrate führen, können kompensiert werden.

- Zur Erkennung etwaiger Manipulationsversuche können sowohl die Bitrate als auch die Meßwerte der Monitordiode herangezogen werden.

Die Wahl des Diodenstromes würde in gleicher Weise geschehen wie bei der Justage der abgeschwächten Photonenquelle bei unseren Experimenten, s. Abschnitt 5.2.1.

[58] Da die Glasplatte polarisationsabhängig reflektiert, empfiehlt es sich, mit Hilfe eines Linearpolarisators das Lichtfeld auf eine Polarisationsrichtung zu beschränken.

[59] Insbesondere Laserdioden verfügen bisweilen bereits über eine in das Gehäuse eingebaute Monitor-Diode, diese könnte natürlich anstelle einer externen Monitor-Diode verwendet werden.

6.3.4 Zusammenfassende Bewertung der Aufbauvarianten

Bei quantenoptischen Zufallsgeneratoren, welche die parametrische Fluoreszenz verwenden, empfiehlt sich die Verdopplung der Bitrate durch Einsatz eines zusätzlich Zufall generierenden Elementes im Strahlengang des Triggerphotons. Hiermit werden überdies Sättigungseffekte bei den Detektoren verringert, so daß mit einer weiteren leichten Erhöhung der Bitrate zu rechnen ist. Etwaige Korrelationseffekte zwischen den beiden Bitströmen, aufgrund der Phasenanpass-Bedingung, s. Abschnitt 3.1.2.1, können entweder durch sorgfältige Wahl der optischen Elemente und Einsatz von entsprechenden Regularisierungsverfahren, Anhang, Abschnitt G, verhindert oder gezielt für Signaturen, s. Abschnitt 6.5.4.2, ausgenutzt werden.
Angesichts der Möglichkeit ohne größeren Aufwand einen weiteren Zufallsstrom zu generieren, sollte die Zufallsgenerierung aus den Abständen zwischen den Detektionserreignissen beim praktischen Einsatz der Generatoren immer vorgesehen werden.
Als Fazit zum HOM-Generator kann man festhalten, daß er zwar aufgrund der höheren Zählraten, die man ohne Einsatz zusätzlicher Detektoren erzielen kann, und der bei ihm automatisch gegebenen Symmetrie der Welcher-Weg-Entscheidung zunächst besticht, aber aufgrund des erheblich aufwendigeren Aufbaus, der Schwierigkeiten, ihn in Faseroptik zu realisieren, und der Notwendigkeit, ihn aktiv mechanisch zu stabilisieren, für den praktischen Einsatz weniger gut geeignet erscheint.
Die integrierte Aufbauvariante empfiehlt sich für einen möglichst kleinen und kostengünstigen Aufbau, da kein Pumplaser notwendig ist und aufgrund der Einbeziehung der Detektorquanteneffizienz in die Abschwächung der Lichtquelle, s. S. 214, keine kostspieligen Detektormodule verwendet werden müssen, sondern kostengünstigere APDs[60] eingesetzt werden können.

6.4 Mögliche Optimierungen für den technischen Einsatz

Bei den experimentell realisierten, quantenoptischen Zufallsgeneratoren handelt es sich um Laboraufbauten, die in erster Linie dazu dienen sollen, die grundsätzlichen Möglichkeiten und etwaigen Probleme beim Aufbau von quantenoptischen Zufallsgeneratoren aufzuzeigen. Deshalb finden auch nur in relativ beschränktem Maße Optimierungen statt, wie sie für den praktischen Einsatz natürlich durchaus wichtig sind. In den folgenden Unterabschnitten werden daher eine Reihe von Verbesserungen und Optimierungsmöglichkeiten für den quantenoptischen Zufallsgenerator dargestellt, der eine Einphotonenquelle auf Basis der parametrischen Fluoreszenz verwendet.

6.4.1 Schneller Triggerdetektor

Der Einsatz eines schnelleren Triggerdetektors wird bei einer technischen Realisierung für den kontinuierlichen Betrieb schon allein deshalb nahegelegt, weil der Triggerdetektor die doppelte Photonenrate der Signaldetektoren sieht und daher der Einsatz eines aktiv gelöschten, hocheffizienten Detektors wünschenswert ist. Im Falle eines Mehrfachgenerators, s. Abschnitt 6.3.1.1, ist der Einsatz von Detektoren des gleichen Typs wie

[60] Selbstverständlich muß es sich dennoch um rauscharme und für den Betrieb im Geigermodus geeignete Avalanchephotodioden handeln, wie z.B. die APD von EG&G, Typ C30902E

bei den „Signaldetektoren“ schon allein aus Symmetriegründen geboten. Wird für den Triggerdetektor ebenfalls ein aktiv gelöschter Einzelquantendetektor mit kurzer Totzeit verwendet, empfiehlt es sich, nach jedem registrierten Generatorereignis durch eine künstlich *für alle Detektoren* eingeführte Totzeit dafür zu sorgen, daß kein weiteres, kurz darauf folgendes Ereignis registriert werden kann.
Erst wenn die Ansprechwahrscheinlichkeiten für die Signaldetektoren wieder gleich groß sind, sollte wieder detektiert werden, da andernfalls eine leichte Tendenz zu Antikorrelationen auftreten wird. Insbesondere mit Detektoren, die einen Gate-Eingang besitzen, über den sich der Detektor „scharf“ schalten läßt, lassen sich solche künstlichen Totzeiten gut realisieren.
Wird der Triggerdetektor durch ein aktiv gelöschtes Modell ersetzt, könnte es im Prinzip verstärkte Probleme mit Nachpulsen geben; daß dies aber nicht der Fall sein wird, zeigt folgendes: Aufgrund der Koinzidenzdetektion und der geringen Breite der Ausgangspulse (< 10 ns) ist die Wahrscheinlichkeit relativ klein, daß ein Nachpuls des Signaldetektors[61] einen korrelierten Bitwert generiert. Legt man nämlich die mittlere Nachpulswahrscheinlichkeit von $P_{Nachpuls} = 3 \cdot 10^{-3}$ eines Nachpulses im Intervall von 50 ns bis 500 ns nach der ursprünglichen Detektion zugrunde, ist die Wahrscheinlichkeit für die Generierung eines korrelierten Bitwertes gerade gegeben durch das Produkt aus der Nachpulswahrscheinlichkeit und der Wahrscheinlichkeit, daß ein (reguläres) Detektionsereignis des Triggerdetektors[62] im Koinzidenzintervall von 10 ns liegt: $P_{Nachpuls} \cdot P_t(\Delta t_{koinz})$. Geht man von einer mittleren Einzelzählrate des Triggerdetektors von ca. 300.000 Ereignissen[63] aus, so ist $P_t \approx 3 \cdot 10^{-3}$, daher liegt die Gesamtwahrscheinlichkeit in der Größenordnung von 10^{-6} und kann vernachlässigt werden.
Tatsächlich wird die Wahrscheinlichkeit für ein auf diese Weise generiertes, korreliertes Bit sogar noch niedriger liegen, da zu der regulären Detektion des Triggerdetektors noch eine Detektion an einem der Signaldetektoren gehören sollte. Fällt das Signalphoton auf den anderen Signaldetektor und wird detektiert, wird ohnehin kein Bit generiert[64], fällt es aber auf den Signaldetektor mit dem Nachpuls, kann man schlecht von einer Korrelation reden. Nur wenn es auf den anderen Signaldetektor fiele und nicht detektiert würde, wird überhaupt ein korreliertes Bit durch den Nachpuls generiert.

6.4.2 Optimierte Filterung von parasitärem Licht

Einphotonen-Anzahlzustände sind sehr fragil. Dies gilt in noch stärkerem Maße für ihre Erzeugung aus Photonenpaaren, da alle Dämpfungseffekte – wie Absorption, Restreflexion an optischen Bauteilen oder nichtideale Quanteneffizienzen und Sättigungseffekte bei den Detektoren – quadratisch die effektive Einphotonenrate, d.h. die Koinzidenzrate

[61] Wenn ein Nachpuls des Triggers koinzident zu einem echten Signalphoton-Detektionsereignis auftritt, ist dies natürlich unproblematisch, da er keine Korrelationen verursacht, es sei denn, es handelt sich um einen Mehrfachgenerator, s. Abschnitt 6.3.1.1.

[62] Die Wahrscheinlichkeit, daß bei beiden Detektoren Nachpulse koinzident auftreten ist auf jeden Fall kleiner als $P^2_{Nachpuls} = 9 \cdot 10^{-6}$. Da die Größe des Koinzidenzintervalles nur von der Elektronik bestimmt wird – die Zeitauflösung der Detektoren ist erheblich besser – ließe sich durch eine Reduzierung des Intervalls z.B. auf 1 ns diese Wahrscheinlichkeit noch weiter senken.

[63] Hier wird also für den Triggerdetektor eine ähnliche hohe Zählrate angenommen wie für den Signaldetektor direkt am Ausgang des Fasereinkopplers, s. Abschnitt 5.1.3.

[64] Dies gilt natürlich nur im Falle der XOR-Verknüpfung der beiden Signale der Signaldetektoren.

zwischen dem Triggerdetektor und den Signaldetektoren, verringern. Dementsprechend empfiehlt es sich, Dämpfungseffekte möglichst zu minimieren, was in den Laboraufbauten durch die Verwendung des nichtkollinearen Falls bei der parametrischen Fluoreszenz und eine Entspiegelung der Bauteile erreicht wird.
Auch sollten Filter und Farbgläser so gewählt werden, daß sie maximal durchlässig für das Photonenpaarlicht sind, bei sehr guter Filterwirkung gegenüber dem parasitären Streulicht auf der Pumpwellenlänge. Die für die Unterdrückung des Streulichtes verwendeten RG 830 Farbgläser verringern zwar die parasitäre Detektion des Pumplaser-Streulichtes um fünf Größenordnungen, haben aber leider nur einen Reintransmissionsgrad von 0,95 bei der Zentralwellenlänge der Photonenpaare.
Nimmt man eine erschwerte Justage[65] in Kauf, kann man bei einem Neuaufbau RG 715 Farbgläser für diesen Zweck verwenden, die einen Reintransmissionsgrad von 1,0 im Wellenlängenbereich des Photonenpaarlichtes haben, bei gleicher Filterwirkung auf das Pumplicht. Der Aufbau sollte in diesem Fall allerdings optisch sehr gut abgedichtet werden, da diese Farbgläser Umgebungslicht im roten Spektralbereich weniger stark filtern als die in den Experimenten eingesetzten RG 830 Farbgläser.
Alternativ ließe sich statt des Farbglases eventuell auch eine Kombination aus Prisma, Blende und Konvexlinse zur Unterdrückung des Streulichtes direkt vor den Detektoren bzw. dem Fasereinkoppler einsetzen; allerdings erschwert auch dies die Justage.

6.4.3 Der Pumplaser

Der Laser stellt hinsichtlich Erstinvestition und laufenden Kosten den größten Kostenfaktor dar. Grundsätzlich hat man bei der Wahl des Pumplasers folgende Möglichkeiten:

- *Verwendung eines HeCd-Gaslasers:* Dies entspricht dem momentanen Aufbau. Um eine möglichst hohe Bitrate zu erreichen, wäre es allerdings wünschenswert, wenn die Zentralwellenlänge der Photonen eines Paares möglichst der Wellenlänge entspräche, bei der die Quanteneffizienz der Detektoren maximal ist (um 700 nm). Da die Pumpwellenlänge bei der Hälfte der Zentralwellenlänge liegt, bedeutet dies, daß man hierzu einen Laser mit einer Emissionswellenlänge von ca. 350 nm brauchte, also einen UV-Laser. Mit HeCd-Lasern[66], mit einer UV-Emission bei 325 nm, steht ein verhältnismäßig kostengünstiger Lasertyp hierfür zur Verfügung. Leider bringt die Verwendung von UV-Lasern aufgrund von erhöhter Streuung und Fluoreszenzanregung eine Erhöhung des Streulichtes mit sich. Daher ist es wichtig, den Pumpstrahl so weit wie möglich vom Zufall erzeugenden Element fernzuhalten.

  Es sei an dieser Stelle nicht verschwiegen, daß ein genereller, zusätzlicher Nachteil der Gaslaser in der relativ kurzen Lebensdauer der Laserröhren und den damit verbundenen höheren Betriebskosten liegt. Nachfragen bei verschiedenen Herstellern von HeCd-Lasern ergaben, daß mittlere Lebensdauern der Röhren von 6000 Stunden garantiert werden (Hersteller: Kimmon Ltd.) und für Industriekunden die Möglichkeit eines flexiblen Röhrenaustauschprogramms[67] besteht, mit dem sich ein kontinuierlicher Betrieb ohne längere Ausfallzeiten sicherstellen läßt.

[65] Das Labor und insbesondere der Aufbau muß noch stärker abgedunkelt werden.
[66] UV-Argon-Ionen-Laser scheiden trotz höherer Leistung wegen ihrer hohen Kosten aus.
[67] Es setzt allerdings einen zweiten Reserve-Laser voraus: Ein Laser mit verbrauchter Röhre bleibt in der deutschen Filiale und wird erst zum Einbau einer neuen Röhre in die Fabrik geschickt, wenn der

- *Verwendung eines diodengepumpten, frequenzverdoppelten Festkörperlasers:* Dies wäre besonders interessant aufgrund der langen Lebensdauern der zum Pumpen des Festkörperlasers verwendeten Halbleiterlaserdioden. Da für den Zufallsgenerator im Vergleich zu sonstigen Anwendungen nur relativ niedrige Leistungen (d.h. < 100 mW) benötigt werden, sind mittlere Lebensdauern über 20.000 Stunden zu erwarten. Überdies sind diese Laser auch in der Anschaffung preisgünstiger als HeCd-Laser. Allerdings schlägt negativ zu Buche, daß ihre Betriebswellenlänge mit $\lambda = 457$ nm verhältnismäßig hoch liegt, so daß auch bei höherer Leistung gegenüber dem jetzt verwendeten HeCd-Laser nicht mit einer entsprechend höheren Bitrate zu rechnen ist.

- *Verwendung einer frequenzverdoppelten Halbleiterlaserdiode:* Seit kurzer Zeit sind frequenzverdoppelte Laserdioden als kommerzielle Komplettlösung für den Wellenlängenbereich um 450 nm erhältlich. Sie zeichnen sich durch längere Lebensdauer (mind. 10.000 Stunden, bei niedrigerer Leistung läßt sich die Dauer eventuell sogar verdreifachen), einen kompakten Aufbau, hohe Leistung (im Vergleich zu HeCd-Lasern) und einen relativ günstigen Preis aus. Ihr Nachteil ist allerdings die relativ lange Wellenlänge; könnte man solch einen Laser bei niedrigeren Wellenlängen (350-440 nm) bekommen, wäre dies ideal. Wichtig ist in diesem Zusammenhang auch, daß für die Verwendung solch eines Lasers als Pumpe für die parametrische Fluoreszenz, die für andere Anwendungen wichtige Leistungsstablilität nicht so wichtig ist, s. a. Abschnitt 6.4.4.

- *Verwendung einer „blauen“ Halbleiterlaserdiode:* Dies wäre vermutlich die beste Lösung. Nachdem nun „blaue“Laserdioden[68] erhältlich sind und die Lebenszeit (bei Laserdioden der Firma NICHIA) bereits über 5.000 Stunden (unter Standardbedingungen) im kontinuierlichen Betrieb beträgt, scheinen geschätzte 10.000 Stunden im Dauerbetrieb nicht unrealistisch. Es gibt jetzt sogar schon solche Laserdioden mit 30 mW Ausgangsleistung, so daß sie einen HeCd-Laser ersetzen können, zumal ihre niedrigere Wellenlänge von $\lambda = 405$ nm noch zusätzlich den Vorteil mit sich bringt, daß die Quanteneffizienzen[69] für die Detektion der Photonen eines Paares (bei 810 nm) fast doppelt so hoch sind wie die Werte bei der in den Laboraufbauten verwendeten Wellenlänge, s. S. 69, Abb. 4.6.

  Leider sind diese Laserdioden zur Zeit noch unverhältnismäßig teuer im Vergleich zu Laserdioden im (infra)roten Spektralbereich; aber aufgrund ihres bald anstehenden Einsatzes in optischen Laufwerken für den Massenmarkt ist mittelfristig mit einer ähnlichen Entwicklung wie bei den Laserdioden für CD-Player zu rechnen, d.h. große Stückzahlen und damit einhergehender Preisverfall. Hinsichtlich Preis, Betriebssicherheit, Größe, Integrationsfähigkeit in Module, Lärmentwicklung, Leistungsverbrauch und Lebensdauer werden sie mit Sicherheit die beiden oben genannten Alternativen in den Schatten stellen. Die kleinen Abmessungen der Laser-

Betriebslaser an die minimale garantierte Röhrenbetriebszeit herankommt. Ist die neue Röhre in den Reserve-Laser eingebaut, wird der Betriebslaser durch ihn ersetzt und zur Filiale geschickt.

[68] Mit ihrer Wellenlänge von $\lambda = 405$ nm sind sie natürlich (im Gegensatz zum Sprachgebrauch) violett.

[69] Im Gegensatz zur Pumpleistung gehen die Quanteneffizienzen der Detektoren quadratisch in die Koinzidenz- und damit schlußendlich in die Bitrate ein!

diode sind natürlich auch interessant, wenn es darum geht, ein Zufallsgenerator-Modul zu bauen.

- *Direkte Verwendung einer „roten" Halbleiterlaserdiode:* Eine weitere Alternative bestünde darin, eine kostengünstige und leistungsstarke rote Halbleiter-Laserdiode (bei 650 nm) als Pumpe zu verwenden und Photonenpaaren im nahen Infrarotbereich bei 1300 nm zu generieren. Ein großer Vorteil hiervon bestünde in der guten Verfügbarkeit von faseroptischen Standardkomponenten, wie sie für die Telekommunikation verwendet werden. Allerdings stehen in diesem Wellenlängenbereich leider noch keine dedizierten Einzelquantendetektoren zur Verfügung, dementsprechend haben selbst die im Hinblick auf die Sensitivität bei schwachen Lichtfeldern am besten geeigneten Ge- oder InGaAs-Avalanche-Dioden bei der Detektion einzelner Photonen nur eine sehr niedrige Quanteneffizienz. Ein weiterer – insbesondere für den praktischen Betrieb schwerwiegender – Nachteil besteht darin, daß diese Detektoren, um das elektronische Detektorrauschen zu reduzieren, entweder sehr stark gekühlt (möglichst mit flüssigem Stickstoff) oder in einem speziellen Triggermodus betrieben werden müssen, der dementsprechend auch eine gepulste Pumpe[70] voraussetzt.

Der Faktor Lebensdauer darf bei einem praxistauglichen Aufbau nicht auf die leichte Schulter genommen werden, da die Installation einer neuen Röhre bzw. einer neuen Laserdiodenpumpeinheit meist beim Hersteller (oft in den USA) vorgenommen wird und bis zu sechs Wochen dauern kann. Will man einen kontinuierlichen Betrieb sicherstellen, muß man immer einen Ersatzlaser vorhalten.
Der Pumplaser sollte daher möglichst nicht in den Generatoraufbau integriert werden, sondern getrennt vom Aufbau stehen, da die mittlere Lebensdauer von Laserröhren bzw. Laserdioden einen Austausch alle ein bis zwei Jahre bei kontinuierlichem Betrieb bedingt. Wird das Laserlicht des Pumplasers z.B. über eine Lichtleitfaser[71] in den restlichen Generatoraufbau geleitet, ließe sich ein schneller, justagearmer Austausch des Lasers ohne längere Betriebsbeeinträchtigung gewährleisten, da auf diese Weise die Eigenschaften des Pumplichtfeldes im Aufbau hinreichend gleich gehalten werden könnten. Außerdem wäre solch ein Aufbau auch unempfindlicher gegenüber mechanischen Störungen. Damit die volle Pumpleistung in der beim nichtlinearen Prozeß für die Pumpe vorgesehenen Polarisationsrichtung vorliegt, muß allerdings entweder eine polarisationserhaltende Faser verwendet oder ein zusätzliches faseroptisches Polarisationsstellglied vorgesehen werden.

[70] Alternativ kann man auch den nichtentarteten Fall der parametrischen Fluoreszenz verwenden und das Triggerphoton bei einer anderen Wellenlänge als das Signalphoton generieren, die sich mit Si-Einzelquantendetektoren noch hinreichend effektiv detektieren läßt. Entsprechend verschiebt sich dann die Wellenlänge des Signalphotons zu noch längeren Wellenlängen hin. Beispielsweise hätte für eine Pumpe bei 600 nm und einem Signalphoton bei der Telekommunikationswellenlänge von 1550 nm das Triggerphoton eine Wellenlänge von 979 nm. Die Wellenlänge des Triggerphotons liegt in diesem Fall bereits in einem Bereich, in dem ein Si-Einzelquantendetektor „nur" noch eine Quanteneffizienz von 20% hat. Dennoch wird man aufgrund der stark unterschiedlichen Quanteneffizienzen der beiden eingesetzten Detektortypen, bei den angestrebten hohen Photonenpaarraten mit Sättigungseffekten beim Triggerdetektor rechnen müssen.

[71] Da die Rayleigh-Streuung mit der vierten Potenz der Frequenz zunimmt, muß man allerdings bei kurzwelligen Pumpwellenlängen mit stärkeren Verlusten in dieser Lichtleitfaser rechnen; die Faser sollte daher möglichst kurz sein. Bei Verwendung eines UV-Lasers ist die Faser auch stärkeren Strahlungsbelastungen ausgesetzt, so daß sie vermutlich turnusmäßig durch eine neue ersetzt werden müßte.

6.4.4 Kompakte Einphotonenquellen

Wie im vorhergehenden Abschnitt schon erwähnt, wäre die Ideallösung für die Laserquelle ein frequenzverdoppelter Diodenlaser im Bereich von 350–440 nm. Einer der Hauptgründe, warum es zur Zeit der Durchführung der Experimente noch keinen günstigen, halbwegs leistungsstarken (mind. 10 mW) frequenzverdoppelten Diodenlaser in diesem Bereich gibt, liegt darin, daß die Hersteller mit massiven Schwierigkeiten kämpfen, die Intensitätsfluktuationen des Lasers unter die Grenze von 10% zu bringen.
Für die Verwendung solch eines Lasers als Pumplaser bei der parametrischen Fluoreszenz sind diese Intensitätsfluktuationen hingegen nahezu[72] irrelevant, denn sie würden sich lediglich in einer schwankenden Photonenpaarrate bemerkbar machen.
Gelänge es, den Pumpstrahl, der beim Durchlaufen des Kristalls nur unwesentlich geschwächt wird, zu „recyclen“, käme man bereits mit relativ niedrigen Pumpleistungen aus.
Eine kompakte Photonenpaarquelle bestünde also aus einer frequenzverdoppelten Laserdiode, wobei das frequenzverdoppelte Licht gar nicht ausgekoppelt würde, sondern im Resonator bliebe. Innerhalb dieses Resonators befände sich der nichtlineare Kristall zur Erzeugung der Photonenpaare; auf diese Weise könnte es gelingen, zu einer kompakten, sehr preiswerten und leistungsstarken Photonenpaarquelle zu kommen. Es sei hier allerdings nicht verschwiegen, daß noch eine ganze Reihe von Problemen zu lösen wären, bevor solch eine Photonenpaarquelle einsatzfähig wäre.
Eine etwas leistungsschwächere, dafür aber leichter zu realisierende Variante einer kompakten Einphotonenquelle ließe sich auf der Basis einer kommerziell erhältlichen, entspiegelten blauen Laserdiode und eines breitbandentspiegelten optisch nichtlinearen Kristalls entwickeln. In diesem Fall würde die Laserdiode mit einem externen Resonator betrieben, der den optisch nichtlinearen Kristall für die Erzeugung der Photonenpaare im optischen Weg enthält. Da die Photonenpaarerzeugung den Pumpstrahl nur vernachlässigbar schwächt, sollte es bei Minimierung sonstiger optischer Verluste möglich sein, den Laser zum Anschwingen zu bringen, bei gleichzeitiger Photonenpaarerzeugung durch den Kristall. Durch die Photonenpaarerzeugung im Laserresonator hätte man auch eine erheblich größere Pumpleistung für den nichtlinearen Prozeß zur Verfügung als beim herkömmlichen Pumpen mit der ausgekoppelten Leistung.
Gegenüber einer kompakten Quelle mit frequenzverdoppelter Laserdiode vermeidet man bei dieser Variante auch das Problem, daß etwaiges Streulicht des Pumplasers, der in jenem Fall dieselbe Wellenlänge hätte wie die Zentralwellenlänge der Photonen eines Paares, das Signal-Rausch-Verhältnis verschlechtern könnte. Lediglich eine möglichst absorbtionsarme räumliche Trennung von Photonenpaar-Lichtfeld und Pumplicht ist auch hier notwendig.
Bei beiden Varianten wäre es im Hinblick auf einen möglichst einfachen Einsatz der kompakten Quelle sehr wünschenswert, die Photonen eines Paares gleich getrennt in Glasfasern einzukoppeln.

[72] Bei einer zusätzlichen Zufallsgenerierung aus den Abständen zwischen den Paaren, s. Abschnitt 6.3.1.2 , könnte sich dies – je nachdem wie man die Bits generiert – bemerkbar machen.

6.4.5 Reduktion der Anzahl der benötigten Detektoren

Die für den quantenoptischen Zufallsgenerator auf Basis der parametrischen Fluoreszenz notwendigen Detektoren sind für einen Einsatz des Generators in der Praxis leider nahezu prohibitiv teuer.
Daher ist es es wünschenswert, zu überlegen, inwieweit sich bei der Detektion des „Signalphotons“ ein Detektor einsparen läßt, indem man „einfach“ beide Ausgänge des Strahlteilers bzw. Faserkopplers in unterscheidbarer Weise auf lediglich *einen* Detektor lenkt. Die Entscheidung, über welchen Ausgang das Photon auf den Detektor fällt, läßt sich nur durch eine zusätzliche optische Verzögerung erreichen, die so groß gewählt werden muß, daß bei der Koinzidenzdetektion mit dem Triggerphoton zwischen dem direkten und dem verzögerten Eintreffen des Signalphotons unterschieden werden kann. Angesichts einer Zeitauflösung der Detektoren von 300 ps wäre bei Aufbauten mit der bisher verwendeten Elektronik die minimale Größe des Koinzidenzfensters (10 ns) die ausschlaggebende Größe für die Dimensionierung. Legt man die verwendete Fenstergröße zugrunde, so ist bereits ein optischer Verzögerungsweg von ca. 3 m bei einem freistrahloptischen und ca. 2 m bei einem faseroptischen Aufbau notwendig. Während dies bei einem faseroptischen Aufbau unproblematisch ist – Glasfasern lassen sich gut aufrollen – kompliziert ein 3 m langer optischer Verzögerungsweg einen freistrahloptischen Aufbau beträchtlich. Der für den praktischen Einsatz notwendige, kompakte Aufbau ließe sich dann nur durch eine Faltung der optischen Verzögerungsstrecke erreichen; die hierfür notwendigen zusätzlichen optischen Elemente würden aber wiederum die Absorptions- und Reflexionsverluste erhöhen.
Durch den Einsatz eines hochauflösenden Zeitpulshöhenkonverters und zwei[73] nachfolgenden Diskriminatoren ließe sich allerdings eine bessere zeitliche Auflösung, s. [16], bis hinunter auf die minimal mögliche Auflösung der Detektoren erzielen. In diesem Fall wäre zwar eine freistrahloptische Variante im Prinzip realisierbar, allerdings bestünde bei ihr das Problem, wie man die verzögerte der beiden räumlichen Ausgangsmoden des Lichtfeldes möglichst verlustfrei ebenfalls auf die Detektorfläche abbildet. Durch Einsatz eines polarisierenden Strahlteilers und zusätzlichen $\lambda/2$-Plättchen im Verzögerungsweg wäre zwar solch eine Abbildung grundsätzlich möglich, aber aufgrund der breiten Wellenlängenverteilung des Lichtfeldes nur mit recht großen Verlusten, was somit wiederum zu einem stark vom idealen Wert abweichenden Teilungsverhältnis führen würde. Somit läßt sich als Fazit festhalten, daß ein faseroptischer Aufbau auch hier wieder die bessere Realisierungsvariante darstellt.
Bei der Gestaltung der faseroptischen Variante mit nur einem Detektor ist ein ähnliches Vorgehen, wie es Á. Stefanov et al. [120] bei ihrem Zufallsgenerator verwenden, empfehlenswert, d.h. die beiden unterschiedlich langen Ausgangsfasern werden beide auf dieselbe Detektorfläche abgebildet. Allerdings sollten aufgrund des stark verlustbehafteten Einkoppelns von Einphotonen-Anzahlzuständen in Einmodenglasfasern besser Mehrmodenglasfasern mit entsprechenden Faserkopplern verwendet werden. Die beiden unterschiedlichen langen Ausgangsfasern des jeweiligen Faserkopplers müssen hierbei entweder durch Schmelz- oder Kernanschliff-Technik zu einem an den Detektor anzuschließenden Ausgang „zusammengefaßt“ werden, wobei diese Stelle möglichst nah am

[73] Bei Mehrfachgeneratoren, die auch auf das Triggerphoton zu Zufallsgenerierung verwenden, würde man dementsprechend drei oder vier Diskriminatoren einsetzen.

Detektor liegen sollte, um die Koppelverluste gering zu halten. Dies setzt natürlich den Zugriff auf entsprechende Methoden der Faserbearbeitung voraus[74].

6.4.6 Zusammenfassende Bewertung der Optimierungsvorschläge

Für einen praktischen Betrieb des Generators sollten gleichartige, schnelle Detektoren, die sich scharf schalten lassen, und möglichst kleine Koinzidenzfenster verwendet werden. Etwaige Korrelationen zwischen aufeinanderfolgenden Bits aufgrund von Detektor-Totzeiteffekten lassen sich durch die Einführung einer künstlichen Totzeit verhindern.
Ein wesentlicher Nachteil des quantenoptischen Zufallsgenerators auf Basis der parametrischen Fluoreszenz ist die verhältnismäßig hohe Pumpleistung, die für hinreichend große Bitraten notwendig ist. Dieses Problem ließe sich durch eine kompakte Photonenpaarquelle wesentlich entschärfen; daher sollte für den praktischen Einsatz solch eine Quelle entwickelt werden, und zwar möglichst mit Faserausgängen.
Eine Reduktion der benötigten Detektoren ohne wesentliche Verschlechterung der Detektion, ist aufgrund der hohen Detektorkosten wünschenswert. Sie setzt allerdings bei dem praxisrelevanten faseroptischen Aufbau mit Mehrmodenfasern eine entsprechende Modifikation der Detektormodule voraus, die aufgrund der Verbindung zwischen Detektorchip und Faser nur vom Hersteller durchgeführt werden könnte. Lediglich wenn es gelänge, die Photonen eines Paares in der oben erwähnten kompakten Quelle hocheffizient in zwei Einmodenfasern einzukoppeln, ließen sich die bisher verwendeten Detektormodule verwenden[75].

6.5 Resistenz der Zufallsgeneratoren gegen Beeinflussungsversuche

Grundsätzlich kann man zwischen zwei verschiedenen Angriffsarten auf den Zufallsgenerator unterscheiden:

- *Destruktive Angriffe*, die darauf abzielen, die Zufallsgenerierung komplett zum Erliegen zu bringen.

- *Nichtdestruktive Angriffe*, bei denen versucht wird, die Funktion des Generators (möglichst unbemerkt) zu beeinflussen, so daß der Angreifer daraus einen Nutzen ziehen kann oder dem legitimen Anwender des Generators ein Schaden daraus entsteht.

[74] Anders als bei Á. Stefanov et al., die Einmodenfasern verwenden und daher die Ausgangslichtfelder der beiden Fasern sehr gut auf die Detektorfläche abbilden können, liegt der Kerndurchmesser von Mehrmodenfasern sehr viel näher an den Abmessungen der aktiven Detektorfläche ($< 200\ \mu$m). Bei Detektoren mit FC-Faseranschluß entspricht der Anschluß natürlich gerade einem Faserdurchmesser. Daher ist diese Vorgehensweise beim Einsatz von Mehrmodenfasern technisch aufwendiger.

[75] In diesem Fall würden im ganzen Aufbau Einmodenfasern verwendet und die Einkopplung zweier Einmodenfasern in den Mehrmodenfaseranschluß des Detektors ließe sich durch einen speziell gefertigten Stecker erreichen.

6.5.1 Destruktive Angriffe

Gegen rohe Gewalt sind natürlich auch quantenoptische Zufallsgeneratoren nur zu schützen, indem man sie entsprechend „panzert“. Auch für andere (äußere) destruktive Angriffe wie:

- die Stromversorgung des Lasers bzw. der LED, der Detektoren, der Signalverarbeitungselektronik oder der Datenaufnahmeelektronik zu unterbinden,
- den optischen Weg zwischen Pumplaser bzw. LED und eigentlichem Zufallsgenerator zu blockieren (z.B. indem man eine Lichtleitfaser entfernt bzw. zerstört),
- die Datenleitung zwischen Zufallsgenerator und Computer zu entfernen bzw. zu zerstören,

gilt, daß sie zu einem sofortigem Erliegen der Zufallsgenerierung führen; mit internen Manipulationen an der Optik oder Elektronik kann dies natürlich ebenfalls erreicht werden. Solche Angriffe lassen sich nur mit angemessenen isolierenden Mitteln verhindern.

6.5.2 Nichtdestruktive Angriffe

Anders als bei den destruktiven Angriffen besteht bei nichtdestruktiven Angriffen die Gefahr, daß sie unentdeckt bleiben und hierdurch die Sicherheit nachfolgender kryptographischer Verfahren und Anwendungen gefährdet wird. Daher ist es wichtig, solche Angriffe zu erschweren, zu erkennen und ihre Auswirkungen möglichst gering zu halten oder ganz zu neutralisieren.
Nichtdestruktive Manipulationsversuche z.B. durch Einstrahlen von Licht lassen sich auf zwei verschiedene Arten leicht entdecken: durch die Kontrolle der Signaldetektorraten und durch eine Überwachung des Verhältnisses zwischen echten und zufälligen Koinzidenzen.
Während die erste Variante sich auch für die Manipulationsüberwachung von quantenoptischen Zufallsgeneratoren mit gepulsten, statistischen Einphotonenquellen eignet, gilt dies für die zweite Variante nicht. Der Grund hierfür besteht darin, daß ein Angreifer in diesem Fall ein ebenfalls gepulstes Störsignal synchron zur Taktung des legitimen Signals[76] einstrahlen könnte.
Da bei einem einmal einjustierten Aufbau die Zählraten bis auf leichte Schwankungen (s. Abschnitt B.2) nicht zunehmen, ist eine Überwachung der Signalraten, bzw. des Teilungsverhältnisses zwischen Nullen und Einsen, eine einfache und allgemein anwendbare Methode, die sich überdies leicht in Software implementieren läßt, s. Abschnitt 4.5.1.1.
Eine Überwachung des Verhältnisses zwischen echten und zufälligen Koinzidenzen ist nur sinnvoll bei einem quantenoptischen Zufallsgenerator, der eine Einphotonen-Anzahlzustände aussendende Quelle (z.B. auf Basis der parametrischen Fluoreszenz) verwendet.

[76] Die Synchronisation auf die Taktung des Generators ist zwar nicht ganz so einfach, da hierfür zuerst einmal das elektrische Ansteuersignal für die Einphotonenquelle, z.B. über seine elektromagnetischen Abstrahlungen, abgehört werden muß, aber wenn keine entsprechenden Gegenmaßnahmen getroffen werden, ist dies durchaus machbar. Lediglich wenn die Taktung selbst zufällig geschehen würde, wäre dies ausgeschlossen.

In diesem Fall wird nicht die Statistik der erhaltenen Zufallszahlen untersucht, sondern es werden die zusätzlichen Informationen genutzt, die man aus der Koninzidenzdetektion gewinnt. Der Vorteil dieser Methode liegt darin, daß man nicht tatsächlich zufällig erzeugte Muster versehentlich als Fehlfunktionen des Generators interpretiert; eine Gefahr, die z.B. bei der Anwendung des Selbsttests nach FIPS-140-1-Standard durchaus besteht. Ein Nachteil des Verfahrens besteht allerdings darin, daß man zusätzliche Hardware vorsehen muß.
Die Idee, die hinter der laufenden Kontrolle des Generators steht, ist relativ einfach: Wenn der Generator ordnungsgemäß funktioniert, dann wird die Rate der echten Koinzidenzen zwischen dem Triggerdetektor und den jeweiligen Signaldetektoren wesentlich höher sein, als die der zufälligen Koinzidenzen. Typischerweise liegt die Rate der echten Koinzidenzen um einen Faktor 20 bis 30 über der Rate der zufälligen. Weicht das Verhältnis von echten zu zufälligen Koinzidenzen stärker von diesem Wert ab, dann liegt entweder ein Defekt im Generator oder ein Angriff auf den Generator vor[77].
Bei quantenoptischen Zufallsgeneratoren, welche die parametrische Fluoreszenz ausnutzen, führen insbesondere Manipulationen am Kristall oder der Optik sofort zu einem drastischen Einbruch der Rate der echten Koinzidenzen[78].

6.5.2.1 Angriffe von außen

Gegen Angriffe von außen ist ein quantenoptischer Zufallsgenerator, welcher die parametrische Fluoreszenz verwendet, besser geschützt als ein Zufallsgenerator auf Basis abgeschwächter Pulse. Alle Angriffe von außen, die eine Veränderung in der Statistik der Zufallsbits erreichen wollen, müssen genau zwei Detektoren beeinflussen, nämlich jeweils den Triggerdetektor und genau einen der Signaldetektoren. Angenommen, ein Angreifer strahlte extrem helle Lichtblitze über den eigentlich für den Pumplaser vorgesehenen Anschluß ein, dann würden mit hoher Wahrscheinlichkeit alle Detektoren gleichzeitig ansprechen. Werden aber beide Signaldetektoren gleichzeitig beeinflußt, so hat dies entweder gar keine Folgen (nämlich dann wenn der Triggerdetektor nicht ebenfalls gleichzeitig anspricht) oder es führt zu einem „Fehler“, d.h. einem Ereignis bei dem beide Koinzidenzeinheiten ein Signal liefern. Diese Fehler werden allerdings von der Datenaufnahmelogik nicht[79] mit abgespeichert; somit hätte der Angreifer in beiden Fällen sein Ziel nicht erreicht.
Dies gilt nicht nur für intensive Lichtblitze, sondern für alle Angriffe, bei denen der Angreifer quasi den ganzen Generator am Stück, bzw. große Teile des Generators, auf einmal zu beeinflussen sucht. Für den Aufbau des Zufallsgenerators folgt daraus, daß die beiden Signaldetektoren möglichst nahe beieinanderstehen sollten, damit jeder Beeinflussungsversuch von außen immer beide betrifft. Die läßt sich insbesondere beim faseroptischen Aufbau auch sehr gut erreichen.

[77] Ab welcher Abweichung man Alarm schlägt, ist natürlich Ermessenssache.

[78] Stärkere Manipulationen würden sich allerdings ohnehin in der Datenrate bemerkbar machen; so führt das Verdrehen des Kristalls um nur 0,5° bereits zu einem totalen Zusammenbruch der Bitrate. Bei einem Aufbau mit Freistrahloptik gilt dies quasi für jede Fehlstellung einer optischen Komponente, dies ist auch einer der Hauptgründe, warum sich ein möglichst kompakter Aufbau oder – besser noch – faseroptischer Aufbau empfiehlt.

[79] Dies gilt nur für die Verwendung einer XOR-Verknüpfung für das Datenübernahmesignal, bei der später verwendeten Oder-Verknüpfung gilt es nicht.

Gelingt es einem Angreifer, mit welchen Mitteln auch immer, selektiv immer nur einen der Signaldetektoren und gleichzeitig[80] auch den Triggerdetektor zu beeinflussen, so kann er auf diese Weise selektiv Bits einschleusen. Allerdings darf er es nicht übertreiben, da sich dies in einer höheren Bitrate zeigen würde und durch ein Ansteigen der echten gegenüber den zufälligen Koinzidenzen zu bemerken wäre. Wenn er allerdings nur mäßig viele Bits unbemerkt einschleust, kann er damit eigentlich nichts anfangen, da er nicht wissen kann, wieviele echte Zufallsbits zwischen den von ihm eingeschleusten liegen.

6.5.2.2 Angriffe durch Veränderung des inneren Aufbaus

Natürlich hilft alle Koinzidenztechnik nichts, wenn ein Angriff auf das Innenleben des Generators derart durchgeführt wird, daß der optische Teil des Generators samt Detektoren von der Elektronik getrennt und anstelle der Detektoren ein fremder Pseudozufallszahlengenerator an die Eingänge der Datenaufnahmeelektronik angeschlossen wird. Ebenfalls sehr anfällig in dieser Hinsicht sind die elektronischen Komponenten der Signalverarbeitungselektronik, was sich bereits bei den diversen Problemen mit dem FIFO zeigte, s. Abschnitt 5.1.2.5.
Gegen solche und ähnlich geartete Attacken, die direkte Eingriffe in den inneren Aufbau des Generators darstellen, helfen nur mechanische Abwehrmaßnahmen[81].

6.5.3 Maßnahmen zur Schadensbegrenzung bei Angriffen

Will man die Auswirkungen von Defekten bzw. Angriffen auf nachfolgende Anwendungen minimieren, so empfiehlt es sich, noch zwei nachfolgende Bearbeitungsstufen vorzusehen:

1. Eine Regularisierungsroutine, die einerseits dafür sorgt, daß einseitige Beeinflussungen, die sich auf das Teilungsverhältnis auswirken, weitgehend neutralisiert werden und andererseits bei massiven Defekten des Generators[82] keine Daten mehr ausgibt, was dem Prinzip entspricht: „Besser gar keine Zufallsdaten als schlechte“.

2. Eine bitweise XOR-Verknüpfung[83] der regularisierten Zufallsdaten mit dem Datenstrom eines kryptographisch starken Pseudozufallszahlengenerators oder einer Stromchiffre. Hierbei wird mit Hilfe einer logischen Exklusiv-Oder-Operation aus je einem Bit des Rohdatenstromes und des Chiffrenstroms der Stromchiffre ein neues Ausgabe-Bit erzeugt. Auf diese Weise kann ein Angreifer selbst bei extremer Manipulation des Generators nicht mehr wissen, wie sich seine Beeinflussungen auf nachfolgende Anwendungen auswirken, da er hierfür nicht nur den Ausgabestroms des physikalischen Zufallsgenerators (nach der Regularisierung) kennen müßte, sondern auch den Chiffrenstrom.

[80] In Bezug auf das verwendete Koinzidenzfenster von 10 ns.

[81] Außerdem sollte der Generator aus Sicherheitsgründen nur von autorisiertem Personal geöffnet und gewartet werden.

[82] Dies wäre z. B. bei Ausfall eines der Signaldetektoren der Fall.

[83] Gegenüber einer einfachen blockweisen Verschlüsselung der Zufallsdaten besitzt dies den Vorteil, daß auch bei extremen Defekten, die zu einer Wiederholung im „Zufalls“-Datenstrom führen, trotzdem keine Wiederholungen im Ausgabedatenstrom auftreten. Allerdings kann man dies durch Verwendung entsprechender Chiffriermodi auch mit einer Blockchiffre erreichen, s. Abschnitt G.2.3.

Die Regularisierung kann zwar im Zufallsgenerator durchgeführt werden, muß es aber nicht, zumal die größere Rechenleistung in einem externen Rechner auch effizientere Regularisierungsverfahren erlaubt, s. Anhang, Abschnitt G. Die XOR-Verknüpfung sollte aus Sicherheitsgründen sogar immer *nach* der Übertragung der Zufallsdaten zu einem Rechner erfolgen, da hiermit auch gleichzeitig ein etwaiges Abhören des Zufallsdatenstroms unwirksam[84] gemacht werden kann. Da eine bitweise XOR-Verknüpfung von Zufallsdaten mit einer Stromchiffre[85] auch zur Regularisierung der Zufallsdaten dienen kann, s. Abschnitt G.2.2, ließe sich zur Vereinfachung im Prinzip auch auf die vorangehende Regularisierung verzichten, allerdings sind dann sowohl die Regularisierung als auch die Verschlüsselung des Zufallsstroms von der Sicherheit der Stromchiffre bzw. ihres Schlüssels abhängig.

6.5.4 Signatur eines Zufallsgenerators

Bei den empirischen Tests von Zufallszahlen wird meist ein Black-Box-Standpunkt eingenommen, d.h. man betrachtet lediglich die Bits, wie sie von einer Black-Box ausgegeben werden, so daß es unerheblich ist, ob sie tatsächlich in Echtzeit von einem Zufallsprozeß erzeugt werden oder nur als eine Aufzeichnung der Ausgabe eines Zufallsprozesses „abgespielt“ werden.
Unterscheiden kann man diese beiden Fälle nur dadurch, daß man den Zufallsprozeß beeinflußt und damit die ausgegebene Bitsequenz verändert, was beim reinen Abspielen von bereits vorher generierten Daten nichts ändern würde. Jede Störung, die eine *merkliche* Auswirkung auf die Ausgabe der Black-Box hat, somit also ein Merkmal, eine Struktur erzeugt, stört (zumindest lokal) wiederum die Zufälligkeit des Bitstroms. Bei der Unterscheidung zwischen echtem Zufall von Pseudozufall oder „gespeichertem“ oder „wiederholtem“ Zufall steht man also vor dem Paradox, daß man die Zufälligkeit stören muß, um sie zu erkennen. Dieser Gedanke ist es, der in direkter Linie zur *Signatur eines Zufallsgenerators* – dem bewußt bei der Konzeption gewählten Abweichen vom „reinen Zufall“ – führt.
Im Abschnitt 6.5.2 wurde bereits betont, daß Angriffe durch Veränderung des inneren Aufbaus des Zufallsgenerators am bedrohlichsten sind und unbedingt verhindert werden müssen. Wie man das Erkennen von Defekten oder Manipulationen am inneren Aufbau eines Zufallsgenerators ermöglicht, ist eng mit der Frage verbunden, ob man Zufallsgeneratoren anhand einer Signatur, d.h. anhand charakteristischer Eigenschaften, identifizieren kann, um so beim Fehlen dieser Signatur auf ein Problem schließen zu können.
Folgende Arten möglicher Signaturen lassen sich unterscheiden:

- *Genuine Signaturen im Zufallsprozeß*: Hierbei handelt es sich um Abweichungen vom idealen Zufall, die vom physikalischen Zufallsprozeß selbst herrühren.

[84] Dies gilt natürlich nur, wenn ein Angreifer den Schlüssel der Stromchiffre bzw. des kryptographischen starken Pseudozufallsgenerators nicht kennt.
[85] bzw. der Ausgabesequenz eines kryptographisch starken Pseudozufallszahlengenerators.

- *Künstlich eingeführte Signaturen*: Hier lassen sich nochmals zwei verschiedene Arten unterscheiden:
  - Signaturen, die durch reale, nicht ideal arbeitende Bauteile verursacht werden. Hierzu zählen z.B. statistische Defekte, die durch Totzeiten von Detektoren oder Abweichungen vom 50:50-Teilungsverhältnis des Strahlteilers verursacht werden.
  - Zusätzlich aufgeprägte Signaturen: So läßt sich z.B. durch Modulation des Pumplasers bzw. der Leuchtdiode die Bitrate beeinflussen, oder durch Verwendung einer Kombination aus verstellbarem Polarisationsdreher ($\lambda/2$-Platte) und polarisierendem Strahlteiler die Teilungsrate des Strahlteilers ändern.

Die oben genannten Signaturen lassen sich noch unterscheiden in *veränderliche* und *unveränderliche Signaturen*, wobei natürlich bei beiden immer statistische Schwankungen berücksichtigt werden müssen. Veränderliche Signaturen sollen es hierbei dem Angreifer noch schwerer machen, überhaupt zu erkennen, daß eine Signatur vorhanden ist.

6.5.4.1 Bitraten- und Teilungsverhältnis-Signaturen

Die einfachste Möglichkeit, eine unveränderliche Signatur zu realisieren, besteht darin, einen Strahlteiler zu verwenden, der kein 50:50-Teilungsverhältnis aufweist, sondern ein signifikant davon abweichendes Teilungsverhältnis, z.B. 60:40. Da die relativen Häufigkeiten von Nullen und Einsen in großen Stichproben statistisch nur sehr wenig, und durch Driftphänomene nur geringfügig und langsam (s. Abschnitt B.2) schwanken, läßt sich diese Signatur leicht erkennen.
Dies gilt allerdings auch für den Angreifer, und kennt er sie, so kann er sie leicht nachbilden, indem er z.B. die Bitfolge eines Pseudozufallsgenerators entsprechend anpaßt. Generell empfiehlt es sich daher, entweder eine schwer erkennbare unveränderliche oder eine veränderliche Signatur zu verwenden.
Eine naheliegende Abwandlung des obigen Verfahrens besteht darin, die Teilungsrate des Strahlteilers veränderbar zu machen, indem man einen polarisierenden Strahlteiler verwendet und mit Hilfe einer rotierbaren $\lambda/2$-Platte den linearen Polarisationszustand des Signalphotons so dreht, daß sich das gewünschte Teilungsverhältnis einstellt. Dies ermöglicht es, die Teilungsrate über einen weiten Bereich einzustellen. Indem man den Drehwinkel der $\lambda/2$-Platte rechnergesteuert verändert, kann man durch Kontrolle des Verhältnisses von Nullen und Einsen anschließend testen, ob die generierten Bits tatsächlich vom Generator erzeugt wurden.
Wenn der Angreifer allerdings weiß, wie die $\lambda/2$-Platte eingebaut ist und wie sie angesteuert wird, kann er selbstverständlich den Bitstrom seines deterministischen Zufallsgenerators entsprechend dem zu erwartenden Teilungsverhältnis anpassen[86].
Eine weitere Möglichkeit, eine veränderliche Signatur zu erzeugen, besteht im Modulieren des Pumplasers: Dreht man die Polarisationsrichtung[87] des Pumplaserlichtes um

[86] Natürlich reicht auch ein Angriff auf die Ausgangsleitungen der Koinzidenzeinheiten, da sich an ihnen direkt die Nullen und Einsen abgreifen lassen und sich somit das Teilungsverhältnis hinreichend genau rekonstruieren läßt.
[87] Das läßt sich z.B. mit elektrooptischen Modulatoren erreichen.

90°, so werden die Photonen eines Paares im nichtkollinearen Fall in andere Richtungen abgelenkt und treffen nicht mehr auf die Detektoren[88], so daß die Bitrate sofort auf Null zurückgeht. Diese Methode hat den Vorteil, daß sie nur die Rate und nicht die Statistik der Zufallsbits verändert. Sie ist allerdings nur anwendbar, wenn sichergestellt werden kann, daß der Pumplaser auch wirklich moduliert wird, die Modulation nicht regelmäßig ist, ein Angreifer nicht einfach die Detektorsignale zur Detektion der Modulation verwenden kann und überdies die Bits direkt eingelesen werden. Da dies doch eine ganze Reihe von Voraussetzungen sind, ist die Methode für das Feststellen von Angriffen auf den inneren Aufbau des Generators weniger geeignet; allerdings kann sie dazu benutzt werden, die Resistenz des Generators gegen äußere Angriffe noch weiter zu erhöhen.

6.5.4.2 Korrelationssignaturen

Die im vorhergehenden Abschnitt vorgestellten Signaturverfahren lassen sich leicht von einem Angreifer erkennen und somit auch einfach außer Kraft setzen. Korrelationssignaturen hingegen sind nicht auf Anhieb, sondern erst durch das Testen größerer Datenmengen hinreichend sicher zu erkennen. So lassen sich z.B. die in Abschnitt B.9.3 dargelegten Abweichungen der empirischen Verteilung des ersten Korrelationskoeffizienten von seiner theoretischen Verteilung sogar hierfür zu Nutze machen. Die Größe der Abweichung wird bei ihnen maßgeblich durch den minimalen zeitlichen Abstand zwischen zwei aufeinanderfolgenden Bits bestimmt. Die Datenaufnahme-Elektronik läßt sich mit Hilfe von Verzögerungsroutinen innerhalb des PIC-Programmes auf einen bestimmten Mindestabstand, den die Bits haben müssen, einstellen; bei Bits, die in kürzeren Abständen aufeinander folgen, wird das zweite dann einfach nicht registriert. Auf diese Weise lassen sich die oben genannten Abweichungen gezielt beeinflussen, und daher als Signatur verwenden. Allerdings sind die oben erwähnten, elektronisch bedingten Abweichungen bei einem Generator für den produktiven Einsatz sicherlich nicht wünschenswert.
Korrelationssignaturen lassen sich bei quantenoptischen Zufallsgeneratoren aber auch auf fundamentalere Weise erzeugen. Insbesondere Mehrfachgeneratoren sind hierfür gut geeignet, wenn auch das Triggerphoton zur Zufallsgenerierung verwendet wird, s. Abschnitt 6.3.1.1, da Signal- und Triggerphoton bereits aufgrund der Phasenanpass-Bedingung korreliert sind, s. Abschnitt 3.1.2.1. Sorgt man nun dafür, daß die Zufall generierenden Elemente auf diese Korrelation reagieren, z.B. indem ihr Teilungsverhältis wellenlängenabhängig ist, werden auch die Werte der beiden Zufallsbits, die bei jedem Ereignis generiert werden, entsprechend korreliert sein. Da die Anordnung der Bits der beiden Zufallsströme bei ihrer Zusammenfügung zu *einem* Ausgangsdatenstrom prinzipiell frei gewählt werden kann, lassen sich auf vielfältigste Weise Korrelationssignaturen in den Ausgangsbitstrom einfügen. Gerade diese frei wählbare Anordnung, die im PIC vorgenommen werden kann, ist besonders gut geeignet für das Einfügen von veränderlichen oder auch interaktiven Signaturen in einen Zufallsstrom.

[88] Ein ähnlicher Effekt ließe sich auch durch Verdrehen des Kristalls erreichen, dies ginge allerdings nicht annähernd so schnell und ist auch nicht ganz unproblematisch hinsichtlich der Rückführgenauigkeit.

6.5.4.3 Mehrfach-Signaturen

Die Sicherheit von Signatur-Verfahren läßt sich im Prinzip durch Kombination mehrerer verschiedenartiger Signaturen erhöhen. Insbesondere bei der Verwendung eines mehrere Zufallsströme produzierenden Generators bietet sich diese Vorgehensweise an. Für einen Angreifer ist es in diesem Fall schwierig, die verschiedenen Signaturen alle auf einmal zu detektieren und zu imitieren.

6.5.4.4 Detektion, Stabilität und Entfernung von Signaturen

Die einfachen Signaturen, wie z.B. bei verändertem Teilungsverhältnis oder bei Pumplichtmodulation, lassen sich mit wenig Aufwand detektieren. Kompliziertere Signaturen, wie z.B. bei den Korrelations-Signaturen, bedürfen aufwendigerer Tests. Sind größere Datenmengen zum Testen notwendig, kann es sinnvoll sein, die Zufallsdaten erst zu testen, bevor man sie weiter verwendet, da man sonst bei einem Angriff Gefahr läuft, die mit ihnen durchgeführten Berechnungen als kompromittiert verwerfen zu müssen.
Die Stabilität der Signaturen ist bei unveränderlichen Signaturen leicht dadurch zu erhöhen, daß man größere Stichproben bei den Tests verwendet. Als Beispiel sind in Abb. 6.7 die empirischen Verteilungen[89] des ersten Autokorrelationskoeffizienten in Abhängigkeit von der Anzahl der Tests[90] erster Stufe dargestellt.

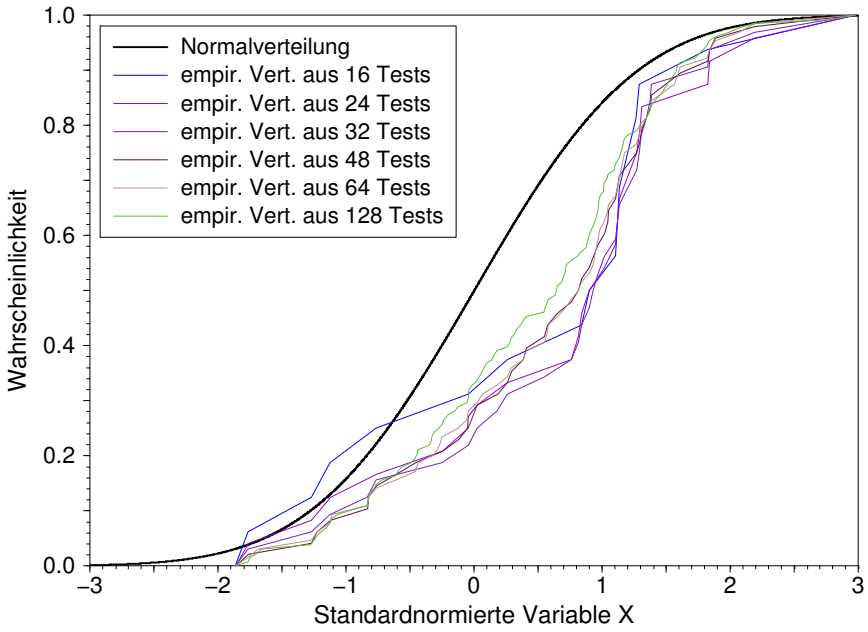


Abbildung 6.7: Empirische Verteilungsfunktionen des ersten Autokorrelationskoeffizienten bei unterschiedlicher Anzahl der Tests erster Stufe, Testdaten des Laufes *LN4*

Bei veränderlichen Signaturen ist dies natürlich nicht so leicht möglich, sondern man muß im Zweifelsfall mehrere Tests kombinieren. Generell gilt für beide Typen von Si-

[89] Die üblichen Treppenkurven für die empirischen Verteilungen wurden ausnahmsweise durch direkte Verbindungen ersetzt, da sonst der Graph zu unübersichtlich geworden wäre.

[90] Die Stichprobenlänge bei diesen Tests war 16 kB, getestet wurde der Anfang des antikorrelationsbehafteten Laufes *LN4*.

gnaturen, daß man bei stärkeren Abweichen vom Idealfall stabilere Signaturen hat und auch stärkere Drifteffekte keine Probleme bereiten[91].
Bevor man die Zufallsbits verwendet, sollte man die Signaturen aus dem Bitstrom entfernen; zu diesem Zweck bieten sich wieder Regularisierungsverfahren an. Je nach verwendeter Signatur bedarf es auch des jeweils angemessenen Regularisierungsverfahrens. Während bei einfachen Veränderungen der relativen Häufigkeiten von Nullen und Einsen Regularisierungsverfahren für unabhängige Bits angewendet werden können, müssen bei Korrelationssignaturen Verfahren für statistisch abhängige Bits verwendet werden. Hierfür geeignete Regularisierungsmethoden sind im Anhang G ausführlich beschrieben.

6.5.5 Zusammenfassende Bewertung zur Resistenz und zu Signaturen

Die besonders kritischen, nichtdestruktiven Beeinflussungsversuche lassen sich durch Überwachung der verschiedenen Kenngrößen des Generators erkennen. Ihre Auswirkungen auf die Anwendungen können durch ein nachfolgendes Regularisierungsverfahren und eine Chiffrierung des Zufallsstroms mit einer Stromchiffre ausgeschlossen werden.
Als ein weitergehendes Mittel, um nachträgliche Manipulationen an einem physikalischen Zufallgenerator zu erkennen, bieten sich Signaturen an. Da sich Signaturen auf einem herkömmlichen PC durch statistische Tests erkennen und mit Hilfe entsprechender Regularisierungsverfahren entfernen lassen, sollten sie im praktischen Einsatz eines Zufallsgenerators als zusätzliche Absicherung gegen Manipulationen eingesetzt werden. Ihr Einfügen kann durch das Ändern physikalischer Parameter oder programmtechnisch in der Datenaufnahme-Elektronik des Zufallsgenerators geschehen; gerade im Falle von Zufallsgeneratoren mit mehreren Zufallsströmen empfiehlt sich die zweite Vorgehensweise, zumal sie auch interaktiv veränderbare Signaturen erlaubt.

[91] Allerdings reduziert dies natürlich auch die effektiv nutzbare Bitrate, da beim Entfernen der Signatur mit Hilfe von angepaßten Regularisierungsverfahren auch mehr Bits verworfen werden müssen.

Kapitel 7

Zusammenfassung

In der vorliegenden Arbeit wird untersucht, inwieweit sich quantenoptische Zufallsgeneratoren, bei denen die „Welcher-Weg-Entscheidung“ einzelner Photonen am Strahlteiler bzw. Faserkoppler zur Zufallsgenerierung verwendet wird, zur Erzeugung von Zufallsbitströmen eignen.
Es werden hierbei im wesentlichen vier verschiedene Varianten aufgebaut, die sich durch die eingesetzte Lichtquelle und die Realisierung des optischen Aufbaus unterscheiden, um zu erkennen, welche Detailprobleme sich beim Aufbau solcher Generatoren zeigen. Als Lichtquellen werden eine Einphotonenquelle auf Basis der parametrischen Fluoreszenz und eine Quelle, die stark abgeschwächte, gepulste Poisson-Lichtfelder abstrahlt, eingesetzt. Bei der optischen Realisierung wird jeweils einmal Freistrahl- und einmal Faseroptik für das Zufall generierende Element verwendet.
Die Rohdaten-Bitströme der verschiedenen Varianten werden mit Hilfe von statistischen Verfahren untersucht, die für Tests von physikalischen Zufallsgeneratoren geeignet sind. In der Diskussion werden die verschiedenen Testverfahren hinsichtlich ihrer Eignung zum Aufdecken tieferliegender Defekte bewertet.
Thermische Einflüsse auf die Rohdaten-Ströme werden dargelegt, Methoden zur Verringerung der Einflüsse angegeben und gezeigt, wie mit Hilfe von mathematischen Regularisierungsverfahren ideale Bitströme aus den Rohdaten erzeugt werden können.
Anhand von (mehrstufigen) Autokorrelationskoeffizententests werden die Auswirkungen von Problemen mit verschiedenen Datenaufnahme-Elektroniken auf die Rohdaten-Ströme analysiert. Die Ursachen der Probleme werden diskutiert, mögliche Lösungen, wie sich die Probleme stark verringern bzw. vermeiden lassen, werden vorgeschlagen und experimentell untersucht.
Die Einflüsse der Eigenschaften der verwendeten Photonenquellen im Zusammenspiel mit den verwendeten optischen Komponenten und Detektoren werden analysiert und ihre Auswirkungen auf die Zufallsgenerierung diskutiert.
Zur Erhöhung der Ausgangbitrate quantenoptischer Zufallsgeneratoren werden verschiedene Ausführungen von Mehrfachzufallsgeneratoren vorgeschlagen, insbesondere für den quantenoptischen Zufallsgenerator auf Basis der parametrischen Fluoreszenz. Als weitere, interessante Variante eines quantenoptischen Zufallsgenerators wird das theoretische Konzept für den „HOM-Generator“ präsentiert, bei dem beide Photonen eines Photonenpaares bei einer gemeinsamen „Welcher-Weg-Entscheidung“ zur Zufallsgenerierung

verwendet werden. Die vorgeschlagenen Varianten quantenoptischer Zufallsgeneratoren werden hinsichtlich ihrer Eignung für einen praktischen Einsatz diskutiert und bewertet. Für den Dauereinsatz quantenoptischer Zufallsgeneratoren als Komponente in Sicherheitsinfrastrukturen, wie z.B. Trustcentern, werden Optimierungen, Möglichkeiten der Kostenreduzierung und weitere Aufbauvarianten vorgeschlagen. Die Optimierungen werden hinsichtlich ihrer Praxistauglichkeit diskutiert und gewertet.
Mögliche Angriffe auf quantenoptische Zufallsgeneratoren werden diskutiert und zur Erkennung von Manipulationen an physikalischen Zufallsgeneratoren werden verschiedene Möglichkeiten vorgestellt, um künstliche Signaturen einzufügen, sie vor Verwendung der Zufallsdaten zu verifizieren und aus dem Zufallsstrom zu entfernen.

Anhang A

Läufe der quantenoptischen Zufallsgeneratoren

In der folgenden Auflistung der Läufe sind bis auf die allerersten, kurzen Testläufe alle späteren Läufe enthalten; zur Vereinfachung der weiteren Verweise im Text werden die Läufe mit einem Kürzel versehen. Läufe mit einer Einphotonenquelle auf Basis der parametrischen Fluoreszenz werden durch ein Kürzel, das mit *K* oder *L* anfängt, gekennzeichnet, je nachdem, ob der Kaliumniobat- oder der Lithiumjodat-Kristall verwendet wird. Die Läufe mit der Photonenquelle auf Basis abgeschwächter Lichtpulse besitzen Kürzel, die mit *P* beginnen. Neben Startzeitpunkt und Zeitdauer des Laufes sind auch noch der Datenaufnahmetyp (von-Neumann-regularisierte oder Rohdaten, abgekürzt durch VNR bzw. BIN), die mittlere Bitrate pro Sekunde und eine kurze Beschreibung der Spezifika des Laufes in der Tabelle aufgeführt. Die mittlere Bitrate bezieht sich immer auf den gesamten Lauf, typischerweise ist die Bitrate zu Anfang eines Laufes größer als gegen Ende, s. Abschnitt 5.1.2.3.
Die durchaus großen Unterschiede in der Bitrate rühren von den unterschiedlichen Rahmenbedingungen bei den verschiedenen Läufen her. Bei Läufen mit abgeschwächter Pumpleistung bzw. zusätzlichen Verzögerungen bei der Datenaufnahme ist dies natürlich intuitiv klar, desgleichen bei den Läufen mit Von-Neumann-Regularisierung[1] bei der Datenaufnahme, die eine ungefähr um den Faktor vier verringerte Bitrate aufweisen.
Bei Läufen, welche die einfache Datenaufnahme-Elektronik verwenden, sorgt diese für eine Beschränkung der maximalen Bitgenerationsrate, s. Abschnitt 5.1.2.2. Die relativ hohen Bitraten bei den Läufen mit den faseroptischen Aufbauten und der Einphotonenquelle auf Basis der parametrischen Fluoreszenz erklären sich aus der höheren effektiven Detektionseffizienz der Detektoren aufgrund der besseren Einkopplung der Lichtfelder, s. 4.4.1.2.

K1 *Start*: 10.01.1997, 14:27, *Dauer*: 24 h 39 m 32 s, *Datenvol.*:13 MB, VNR, *Bitrate*: 1133 ± 7
Spezifika: Einphotonenquelle mit $KNbO_3$, Aufbau in Freistrahloptik mit $\lambda/2$-Plättchen, nachfolgendem polarisierenden Strahlteiler und alter Datenaufnahme-Elektronik

K2 *Start*: 14.01.1997, 10:53, *Dauer*: 25 h 41 m 46 s, *Datenvol.*:12 MB, VNR, *Bitrate*: 1088 ± 22
Spezifika: wie *K1*

K3 *Start*: 15.01.1997, 14:54, *Dauer*: 22 h 53 m 0 s, *Datenvol.*: 11 MB, VNR, *Bitrate*: 1119 ± 3
Spezifika: wie *K1*

K4 *Start*: 16.01.1997, 15:19, *Dauer*: 22 h 49 m 48 s, *Datenvol.*: 11 MB, VNR, *Bitrate*: 1122 ± 11
Spezifika: wie *K1*

[1] Die Von-Neumann-Regularisierung ist in diesem Fall auch für die geringeren Schwankungen der Bitrate verantwortlich.

K5 *Start*: 20.06.1997, 16:11, *Dauer*: 60 h 4 m 16 s, *Datenvol.*: 30 MB, VNR, *Bitrate*: 1163 ± 3
Spezifika: weitgehend wie *K1*, aber aufgrund leicht veränderter Strahlcharakteristika des neuen HeCd-Pumplasers mit geänderter Entfernung vom Pumplaser zum Kristall

K6 *Start*: 24.06.1997, 17:09,*Dauer*: 59 h 23 m 25 s, *Datenvol.*: 30 MB, VNR, *Bitrate*: 1176 ± 1
Spezifika: Neujustage des Aufbaus gegenüber *K5* führte zwar zu höheren Koinzidenzzählraten, Bitrate ist aber nicht erhöht

K7 *Start*: 27.06.1997, 19:24, *Dauer*: 59 h 37 m 30 s, *Datenvol.*: 30 MB, VNR, *Bitrate*: 1172 ± 3
Spezifika: wie *K6*

K8 *Start*: 30.06.1997, 18:38, *Dauer*: 59 h 0 m 23 s, *Datenvol.*: 30 MB, VNR, *Bitrate*: 1184 ± 1
Spezifika: Neujustage der Signaldetektoren, hohe Koinzidenzraten

K9 *Start*: 04.07.1997, 18:38, *Dauer*: 59 h 0 m 23 s, *Datenvol.*: 30 MB, VNR, *Bitrate*: 1184 ± 1
Spezifika: –

K10 *Start*: 18.07.1997, 17:29, *Dauer*: 53 h 47 m 26 s, *Datenvol.*: 106 MB, BIN, *Bitrate*: 4593 ± 66
Spezifika: einziger nicht regularisierter Lauf mit dem Kaliumniobat-Kristall

LO1 *Start*: 29.08.1997, 16:46, *Dauer*: 51 h 49 m 29 s, *Datenvol.*: 100 MB, BIN, *Bitrate*: 4766 ± 8
Spezifika: Freistrahlaufbau mit $LiIO_3$-Kristall und unpolarisierendem Strahlteiler, Pumplaserleistung verringert, um Sättigungseffekte beim Triggerdetektor zu verringern

LO2 *Start*: 8.09.1997, 14:38, *Dauer*: 51 h 53 m 37 s, *Datenvol.*: 100 MB, BIN, *Bitrate*: 4760 ± 9
Spezifika: –

LO3 *Start*: 13.09.1997, 19:21, *Dauer*: 54 h 53 m 34 s, *Datenvol.*: 100 MB, BIN *Bitrate*: 4245 ± 24
Spezifika: –

LN1 *Start*: 2.02.1998, *Dauer*: abgebrochen, *Datenvol.*: 2 MB, BIN, *Bitrate*: -
Spezifika: Testlauf mit kompakter Datenaufnahme-Elektronik, starke Probleme mit dem FIFO, Lauf bis auf den Anfang unbrauchbar

LN2 *Start*: 24.03.1998, 18:36, *Dauer*: 3 h 43 m 19 s, *Datenvol.*: 5 MB, BIN, *Bitrate*: 3130 ± 7
Spezifika: Test des Zufallsgenerators ohne den FIFO, d.h. PIC liefert die Bytes direkt an den IO-Ausgang, Daten werden byteweise übernommen, Antikorrelationen sind vorhanden

LN3 *Start*: 25.03.1998, 20:17, *Dauer*: 3 h 1 m 5 s, *Datenvol.*: 5 MB, BIN, *Bitrate*: 3860 ± 4
Spezifika:Vergleichsmessung mit der einfachen Datenaufnahme-Elektronik

LN4 *Start*: 26.03.1998, 13:54, *Dauer*: 5 h 4 m 2 s, *Datenvol.*: 10 MB, BIN, *Bitrate*: 4599 ± 17
Spezifika: Messung mit kompakter Datenaufnahme-Elektronik, ohne FIFO

LN5 *Start*: 26.03.1998, 19:59, *Dauer*: 15 h 24 m 12 s, *Datenvol.*: 20 MB, BIN, *Bitrate*: 3026 ± 25
Spezifika: Messung mit kompakter Datenaufnahme-Elektronik, ohne FIFO und mit halber Pumpleistung

LN6 *Start*: 27.03.1998, 15:48, *Dauer*: 26 h 37 m 22 s, *Datenvol.*: 20 MB, BIN, *Bitrate*: 1750 ± 7
Spezifika: Messung mit kompakter Datenaufnahme-Elektronik, ohne FIFO und mit einem Viertel der vollen Pumpleistung

LN7 *Start*: 29.04.1998, 21:36, *Dauer*: 11 h 56 m 43 s, *Datenvol.*: 18 MB, BIN, *Bitrate*: 3396 ± 35
Spezifika: Messung mit kompakter Datenaufnahme-Elektronik, ohne FIFO

LD1 *Start*: 4.05.1998, 19:38, *Dauer*: 13 h 45 m 35 s, *Datenvol.*: 30 MB, BIN, *Bitrate*: 5084 ± 95
Spezifika: Messung mit kompakter Datenaufnahme-Elektronik und neuem FIFO-Chip

LD2 *Start*: 6.05.1998, 19:09, *Dauer*: 14 h 15 m 32 s, *Datenvol.*: 30 MB, BIN, *Bitrate*: 4907 ± 118
Spezifika: wie *LD1*, aber mit Gesamtverzögerung bei der Bitübernahme von 4 μs

LD3 *Start*: 12.05.1998, 14:06, *Dauer*: 4 h 5 m 44 s, *Datenvol.*: 8 MB, BIN, *Bitrate*: 4558 ± 164
Spezifika: wie *LD1*, aber mit Gesamtverzögerung bei der Bitübernahme von 6 μs

LD4 *Start*: 12.05.1998, 18:25, *Dauer*: 18 h 10 m 52 s, *Datenvol.*: 30 MB, BIN, *Bitrate*: 3857 ± 221
Spezifika: wie *LD1*, aber mit Gesamtverzögerung bei der Bitübernahme von 10 μs

LD5 *Start*: 20.05.1998, 11:33, *Dauer*: 5 h 37 m 44 s, *Datenvol.*: 10 MB, BIN, *Bitrate*: 4144 ± 123
Spezifika: wie *LD1*, aber mit Gesamtverzögerung bei der Bitübernahme von 18 μs

LD6 *Start*: 25.05.1998, 20:01, *Dauer*: 17 h 4 m 0 s, *Datenvol.*: 21 MB, BIN, *Bitrate*: 2886 ± 247
Spezifika: wie *LD1*, aber mit Gesamtverzögerung bei der Bitübernahme von 32 μs

LD7 *Start*: 27.05.1998, 21:04, *Dauer*: 12 h 28 m 31 s, *Datenvol.*: 20 MB, BIN, *Bitrate*: 3737 ± 57
Spezifika: wie *LD1*, aber mit Gesamtverzögerung bei der Bitübernahme von 50 μs

LD8 *Start*: 2.06.1998, 19:32, *Dauer*: 16 h 22 m 50 s, *Datenvol.*: 21 MB, BIN, *Bitrate*: 2993 ± 138
Spezifika: wie *LD1*, aber mit Gesamtverzögerung bei der Bitübernahme von 92 μs

LD9 *Start*: 24.07.1998, 17:07, *Dauer*: 3 h 57 m 25 s, *Datenvol.*: 11 MB, BIN, *Bitrate*: 6482 ± 108
Spezifika: wie *LD1*

LH1 *Start*: 27.07.1998, 18:40, *Dauer*: 15 h 46 m 2 s, *Datenvol.*: 52 MB, BIN, *Bitrate*: 7693 ± 149
Spezifika: parametrische Fluoreszenz, Freistrahloptik, Lauf mit hybrider Datenaufnahme-Elektronik, 400 ns Übernahmepuls, 100 ns vor Bitwert-Puls von 4,6 μs Länge, starke Antikorrelationen

LH2 *Start*: 30.07.1998, 00:19, *Dauer*: 9 h 54 m 27 s, *Datenvol.*: 33 MB, BIN, *Bitrate*: 7769 ± 97
Spezifika: parametrische Fluoreszenz, Freistrahloptik, Lauf mit hybrider Datenaufnahme-Elektronik, 400 ns Übernahmepuls, keine Verzögerung, Bitwert-Puls von 2 μs

LH3 *Start*: 4.08.1998, 21:06, *Dauer*: 12 h 17 m 57 s, *Datenvol.*: 33 MB, BIN, *Bitrate*: 6303 ± 571
Spezifika: parametrische Fluoreszenz, Freistrahloptik, Lauf mit hybrider Datenaufnahme-Elektronik, 400 ns Übernahmepuls, keine Verzögerung, Bitwert-Puls von 2 μs

LH4 *Start*: 5.08.1998, 10:34, *Dauer*: 6 h 31 m 28 s, *Datenvol.*: 15 MB, BIN, *Bitrate*: 5361 ± 112
Spezifika: parametrische Fluoreszenz, Freistrahloptik, Lauf mit hybrider Datenaufnahme-Elektronik, 400 ns Übernahmepuls, keine Verzögerung, Bitwert-Puls von 4.6 μs Länge; ähnlich *LH1*, zur Überprüfung der Antikorrelationsprobleme, zeigt ebenfalls starke Antikorrelationen

LH5 *Start*: 8.08.1998, 18:38, *Dauer*: 18 h 41 m 9 s, *Datenvol.*: 53 MB, BIN, *Bitrate*: 6615 ± 124
Spezifika: parametrische Fluoreszenz, Freistrahloptik, Lauf mit hybrider Datenaufnahme-Elektronik (Parameter wie *LH3*), FIFO-bedingte Störung zu Anfang des Laufs

LF1 *Start*: 20.08.1998, 21:36, *Dauer*: 13 h 3 m 40 s, *Datenvol.*: 55 MB, BIN, *Bitrate*: 9823 ± 102
Spezifika: parametrische Fluoreszenz, hybride Datenaufnahme-Elektronik (Parameter wie *LH3*), Faseraufbau

LF2 *Start*: 21.08.1998, 22:43, *Dauer*: 59 h 16 m 49 s, *Datenvol.*: 251 MB, BIN, *Bitrate*: 9876 ± 38
Spezifika: parametrische Fluoreszenz, hybride Datenaufnahme-Elektronik (Parameter wie *LH3*), Faseraufbau, mit Klimaanlage im Raum

PB1 *Start*: 13.08.98, 18:35, *Dauer*: 5 h 51 m 19 s, *Datenvol.*: 10 MB, BIN, *Bitrate*: 3980 ± 20
Spezifika: abgeschwächte, gepulste Photonenquelle, 250 kHz Pulsfrequenz, freistrahloptischer Aufbau, hybride Datenaufnahme-Elektronik (Parameter wie *LH3*)

PB2 *Start*: 17.08.98, 18:25, *Dauer*: 15 h 1 m 0 s, *Datenvol.*: 55 MB, BIN, *Bitrate*: 8542 ± 38
Spezifika: abgeschwächte, gepulste Photonenquelle, 500 kHz Pulsfrequenz, freistrahloptischer Aufbau, hybride Datenaufnahme-Elektronik (Parameter wie *LH3*); elektrische Störungen zu Anfang des Laufes

PF *Start*: 25.08.98, 23:03, *Dauer*: 18 h 44 m 28 s, *Datenvol.*: 30 MB, BIN, *Bitrate*: 3731 ± 27
Spezifika: abgeschwächte, gepulste Photonenquelle, 500 kHz Pulsfrequenz, Faseraufbau, hybride Datenaufnahme-Elektronik (Parameter wie *LH3*), mit Klimaanlage im Raum

Anhang B

Analyse der Zufallssequenzen

Wenn nicht gesondert hervorgehoben, werden die statistischen Tests zur Analyse der Zufallssequenzen mit Rohdaten der quantenoptischen Zufallsgeneratoren durchgeführt, d.h. mit Sequenzen, die *nicht regularisiert* wurden und somit neben ungleichen Häufigkeiten der Null- und Eins-Werte auch Drifteffekte bei den relativen Häufigkeiten aufweisen können, s. Abschnitt B.2.

B.1 Vergleichs-Pseudozufallszahlengeneratoren

Will man die Güte physikalischer Zufallsgeneratoren mit Hilfe von statistischen Tests einschätzen, empfiehlt es sich, die verwendeten Testmethoden vorher an algorithmischen Pseudozufallsgeneratoren zu evaluieren. Hierfür werden eine Reihe von Referenzgeneratoren unter folgenden zwei Gesichtspunkten ausgewählt: Sie sollen hinreichend repräsentativ für gängige Generatortypen sein, und es sollen sowohl statistisch und kryptographisch schwache als auch starke[1] Pseudozufallsgeneratoren dabei sein.
Es mag vielleicht erstaunen, warum statistisch bekanntermaßen schwache Generatoren überhaupt zum Vergleich herangezogen werden, aber dies hat durchaus einen Grund: Die statistischen Tests lassen sich auf diese Weise hinsichtlich ihrer „Mächtigkeit“ beurteilen, sollte ein statistischer Test doch i. a. statistisch schwache Zufallsgeneratoren auch als solche erkennen.
Die Referenzgeneratoren werden aus folgenden Arten von Pseudozufallsgeneratoren gewählt:

- *lineare Kongruenzen verwendende Zufallszahlengeneratoren*, da deren Eigenschaften gut untersucht wurden, s. z. B. [68]. Bei ihnen wird die nächste Zufallszahl aus ihrer Vorgängerzahl gemäß:

  $$X_{n+1} = a \cdot X_n \pmod{m}$$

  berechnet. Für alle linearen Kongruenz-Generatoren wird mit dem hexadezimalen Startwert X_0: `5ED4BA6D` (dezimal: 1590999661) begonnen. Folgende, teilweise etwas modifizierte, lineare Kongruenz-Generatoren werden als Referenz verwendet:

  - `RANDU`, ein Generator mit bekanntermaßen [68] schlechten statistischen Eigenschaften und den Parameterwerten $a = 56539$, $m = 2^{31}$ und X_0 ungerade.

[1] Pseudozufallsgeneratoren mit guten statistischen Eigenschaften müssen nicht kryptographisch stark sein, da es bei der kryptographischen Stärke um die Frage der Vorhersagbarkeit und nur indirekt um eine Gleichverteilung der Werte geht; umgekehrt gilt allerdings, daß kryptographisch starke Pseudozufallsgeneratoren auch sehr gute statistische Eigenschaften haben. Gute Pseudozufallsgeneratoren dienen insbesondere dazu, die korrekte Implementation der statistischen Tests in Software zu kontrollieren.

– Der „Minimal-Standard-Generator“ von Park und Miller [101], mit den Parameterwerten $a = 7^5 = 16807$ und $m = 2^{31} - 1 = 2147483647$; der Generator entspricht[2] dem `ran0` aus den *Numerical Recipes* [106].

– Der `ran1` genannte Generator aus den *Numerical Recipes*, der gegenüber dem `ran0` noch eine zusätzliche Verwürfel-Tabelle verwendet, die dafür sorgt, daß die generierten Zahlen in variabler, anderer Reihenfolge ausgegeben werden.

– Der `ran2` genannte Generator aus den *Numerical Recipes*, ein Kombinationsgenerator, bei dem zwei lineare Kongruenz-Generatoren (mit den Parametern $a_1 = 40014$ und $m_1 = 2147483563$ bzw. $a_2 = 40692$ und $m_2 = 2147483399$) durch Addition modulo m_1 miteinander kombiniert werden; zusätzlich kommt wieder eine Verwürfel-Tabelle zum Einsatz.

- *kryptographisch starke Zufallsgeneratoren*, welche die symmetrischen Blockchiffrier-Algorithmen DES[3] bzw. IDEA[4] im Output-Feedback-Modus und im Counter-Modus (s. S. 18) zur Erzeugung von Zufallszahlen benutzen.

B.2 Häufigkeitsverteilung der Bitwerte

Einer der elementarsten, statistischen Tests, den man auf den generierten Bitsequenzen durchführen kann, ist die Untersuchung der Häufigkeitsverteilung der Einsen (bzw. der Nullen) durch einen Vergleich der empirisch erhaltenen Verteilungsfunktion mit der theoretisch erwarteten. Bereits bei diesem einfachen Test zeigen sich einige interessante Aspekte quantenoptischer Zufallsgeneratoren.
Hierzu wird für den typischen, langen Lauf *LO2* zuerst der empirische Schätzwert für die Wahrscheinlichkeit $\hat{P}(1) = \hat{p}$ ermittelt, daß ein Bit den Wert Eins hat. Anschließend wird eine empirische Häufigkeitsverteilung der Anzahl der Einsen n_E in nicht überlappenden Blöcken mit der Länge N Bits (typ. $N = 128$ kB $= 1.048.576$ Bits) erstellt, wobei die Werte n_E vorher noch standardnormiert (X_E) werden. Im Idealfall sollten die standardnormierten Werte X_E in sehr guter Näherung, s. a. Abschnitt D.1, einer Normalverteilung folgen. Betrachtet man sich allerdings die empirische Verteilungsfunktion und vergleicht sie mit der theoretischen Kurve (s. Abb. B.1), so sieht man auch ohne Anwendung eines Kolmogorov-Smirnov-Tests sofort, daß die empirische Verteilungsfunktion sehr stark von einer Normalverteilung abweicht. Diese Abweichung ist ein Zeichen dafür, daß die Wahrscheinlichkeit $P(1)$ zeitlich stärker schwankt, als im theoretischen Idealfall zu erwarten wäre. Noch deutlicher läßt sich dies erkennen, wenn man die lokale Wahrscheinlichkeit p_{Block} für eine Eins auf nicht überlappenden Blöcken gleicher Länge[5] schätzt und graphisch in Abhängigkeit vom Blockindex aufträgt, s. Abb. B.2. Zwar sind die Schwankungen recht gering, würde man sie auf einer Skala von Null bis Eins betrachten, aber in der vergrößerten Darstellung von Abb. B.2 kann man eine deutliche Tendenz erkennen.

[2] Lediglich der letzte Schritt, in dem aus den Ganzzahlwerten Fließkommazahlen errechnet werden, unterbleibt hier natürlich, dies gilt auch für die folgenden Generatoren.

[3] DES steht für *Data Encryption Standard* und bezeichnet einen vom National Institute of Standards, NIST im *FIPS 46-2* standardisierten, stark nichtlinearen Verschlüsselungsalgorithmus, der 64 Bit lange Klartext-Eingangsblöcke in Abhängigkeit von einem 56 Bit langen, geheimen Schlüssel eineindeutig auf ebenfalls 64 Bit lange Chiffrat-Blöcke abbildet. Aufgrund von je einem Paritätsbit pro 7-Schlüsselbits ist der eingegebene Schlüsselformat schließlich 64 Bit lang.

[4] IDEA steht für *International Data Encryption Algorithm*. IDEA verschlüsselt 64 Bit lange Eingabeblöcke in Abhängigkeit von einem 128 Bit langen Schlüssel und erzeugt dabei ebenfalls 64 Bit lange Chiffrat-Blöcke; für eine nähere Beschreibung s. z.B. [91], Abschnitt 7.6.

[5] Damit die graphische Darstellung nicht durch zu viele Punkte überladen und durch rein statistische Schwankungen der Verlauf zu undeutlich wird, werden hierfür ebenfalls Blöcke der Länge $N = 128$ kB verwendet.

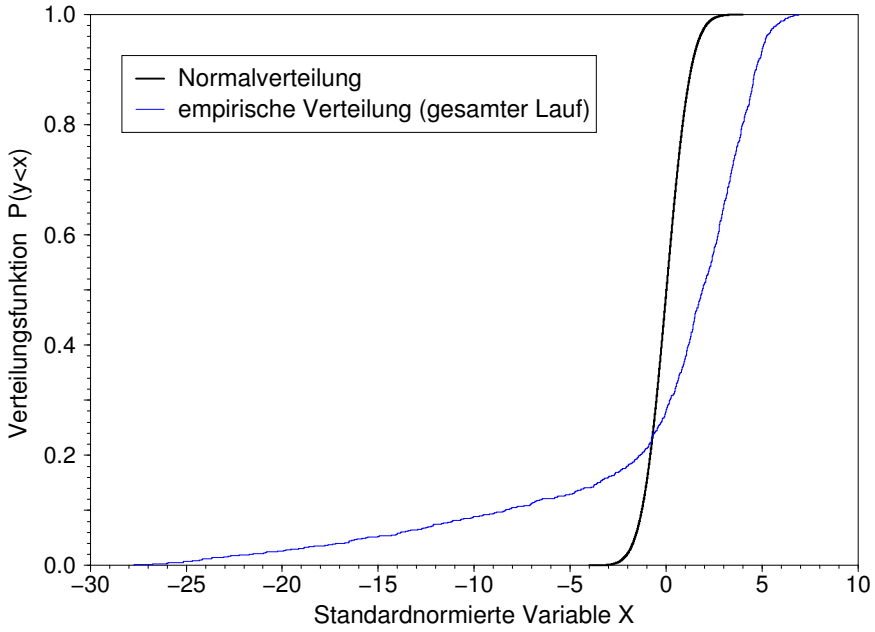


Abbildung B.1: Empirische und theoretische Verteilungsfunktion der standardnormierten Anzahl der Einsen in Blöcken der Länge N =128 kB für den typischen Lauf *LO2*

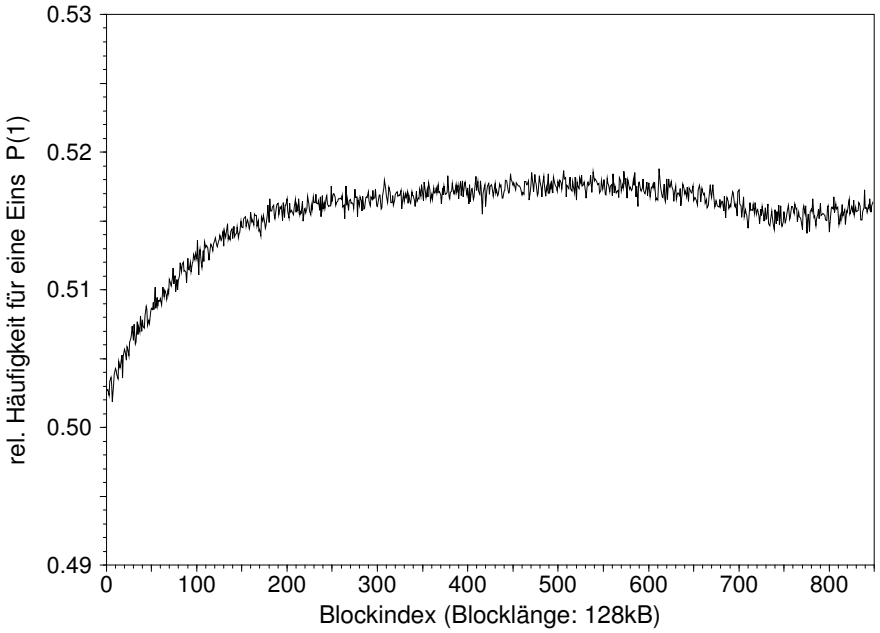


Abbildung B.2: Verlauf der geschätzten Wahrscheinlichkeit p_{Block} für eine Eins bei einer Blockgröße von N = 128 kB für den Lauf *LO2*

Diese Tendenz findet sich auch bei anderen Meßläufen zu Anfang der Zufallsgenerierung, sie kommt aufgrund von Aufwärmeffekten der Detektoren zustande und läßt sich gut durch folgende

Formel[6] beschreiben:

$$p(t) = p_{stat} - (p_0 - p_{stat}) \cdot e^{-\frac{\ln 2 \cdot t}{t_h}},$$

wobei p_0 für die Anfangswahrscheinlichkeit, p_{stat} für die „stationäre“ Endwahrscheinlichkeit für eine Eins und t_h für die „Halbwertszeit“ des Anstiegs[7] steht.
Ebenfalls gut zu erkennen ist der weitgehend stationäre Zustand, der sich nach dieser „Aufwärmphase“ ausbildet. Die oben dargestellte, stark von der Normalverteilung abweichende, empirische Häufigkeitsverteilung wird also hauptsächlich durch die Aufwärmphase verursacht. Dies sieht man besonders gut, wenn man eine empirische Häufigkeitsverteilung, die aus den Daten der ersten 32 MB des Laufes erstellt wurde (s. Abb. B.3), mit einer Verteilungsfunktion vergleicht (s. Abb. B.4), die aus den darauf folgenden 32 MB des „weitgehend stationären“ Abschnittes erstellt wurde. Die beiden Signifikanzwerte $P(K_n^+) = 0{,}85$ und $P(K_n^-) = 0{,}98$, welche die Kolmogorov-Smirnov-Auswertung für den zweiten Abschnitt liefert, sind zwar besser als die Werte des ersten Abschnitts $P(K_n^+) = 1{,}00$ und $P(K_n^-) = 1{,}00$, aber dennoch zeigt insbesondere der Wert für die negative Abweichung deutlich, daß der Test die langsamen Schwankungen[8] in der relativen Häufigkeit „aufdecken“ kann.

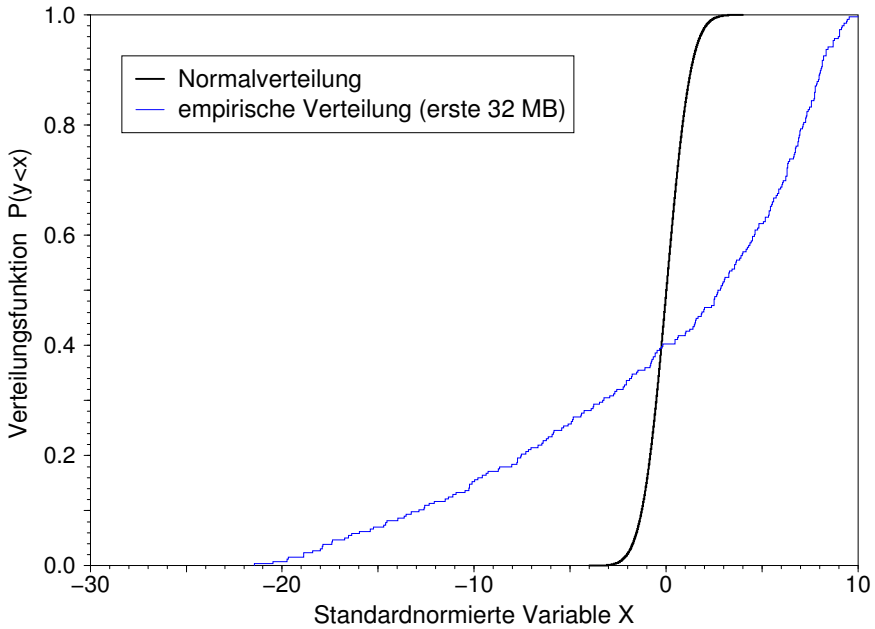


Abbildung B.3: Empirische und theoretische Verteilungsfunktion der standardnormierten Anzahl der Einsen in N =128 kB Blöcken für die ersten 32 MB des Laufes *LO2* („Aufwärmphase“)

Betrachtet man sich die Darstellung des Verlaufes der Wahrscheinlichkeit für eine Eins aus Abb. B.2 genauer, so sieht man gegen Ende der Messung einen leichten Abfall der Wahrscheinlichkeit.

[6] Je nachdem welcher der Detektoren die Einsen generiert, muß entweder eine abklingende oder eine zunehmende Exponentialfunktion verwendet werden, i. a. nahm die Einswahrscheinlichkeit aber meist im Laufe der Zeit zu.

[7] Je nach Wahl der „Einheit“ für t_h läßt sich der Wahrscheinlichkeitverlauf auf Block- oder Bitebene an eine empirische Kurve anpassen. Unter Verwendung der Anpasskurven lassen sich auch Surrogat-Daten, die ebenfalls „Aufwärmeffekte“ zeigen, mit Hilfe von Pseudozufallsgeneratoren erzeugen.

[8] Im Verlauf der Einswahrscheinlichkeiten dieser Stichprobe, sieht man auch beim zweiten 32 MB großen Abschnitt eine langsame Zunahme der Einswahrscheinlichkeit, die sich in der Verteilungsfunktion in Abb. B.4 durch eine erhöhte Varianz bemerkbar macht.

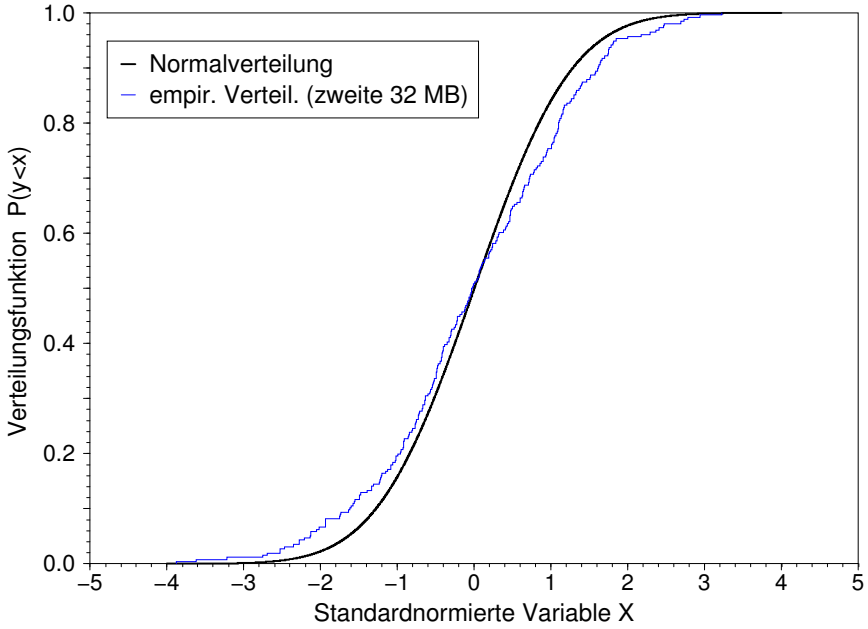


Abbildung B.4: Empirische und theoretische Verteilungsfunktion der standardnormierten Anzahl der Einsen in N =128 kB Blöcken für die zweiten 32 MB des Laufes *LO2* („weitgehend stationäre Phase“)

Solche langsamen Schwankungen werden im wesentlichen durch Temperaturschwankungen aufgrund von Wetterumschwüngen bzw. Tag-Nacht-Wechseln[9] erzeugt.
Zur Illustration, wie stark sich die Temperaturschwankungen auswirken können, sind in Abb. B.5 Verläufe der Wahrscheinlichkeit für eine Eins für zwei Läufe[10] (*LF1* und *LF2*) dargestellt, die sich neben ihrer unterschiedlichen Länge darin unterscheiden, daß bei einem von ihnen (*LF2*) das Labor mit Hilfe einer transportablen Klimaanlage durch Kühlung in der Temperatur konstant gehalten wurde (26° ± 0,3°) und auch die Detektoren im Luftstrom der Klimaanlage standen.
Wie man sehr gut erkennen kann, führt die Temperierung durch die Klimaanlage dazu, daß die Aufwärmeffekte wesentlich moderater ausfallen als ohne Klimaanlage. Da beim faseroptischen Aufbau sich überdies etwaige mechanische Schlupfeffekte der optischen Justageeinrichtungen meist auf die Zählraten von Nullen und Einsen auswirken müßten, kann man daraus schließen, daß die Ursache für die Veränderung der relativen Häufigkeit der Einsen tatsächlich im unterschiedlichen thermischen Verhalten der beiden Signaldetektoren zu sehen ist. Dies wird auch noch dadurch nahegelegt, daß beim Lauf mit Klimaanlage nicht nur die Umgebungstemperatur im Labor niedriger (26° C statt 29° bis 31,7° im Falle ohne Klimaanlage) ist, sondern auch die Gehäusetemperatur der Detektoren (32° statt 41,9° C). Bedenkt man, daß die Detektoren nur bis 35° C Betriebstemperatur spezifiziert sind, sind die Abweichungen im Verhalten der Detektoren also durchaus verständlich.
Im Prinzip könnte natürlich auch eine thermische Abhängigkeit des Teilungsverhältnisses[11] des Faserkopplers dafür verantwortlich sein; aufgrund der sehr viel geringeren Schwankungen der Raumtemperatur gegenüber der Detektortemperatur und des Auftretens ähnlicher Schwankungen bei den Aufbauten mit Freistrahloptik erscheint es aber wahrscheinlicher, daß die Schwan-

[9] Der oben dargestellte Lauf erstreckte sich über mehr als zwei Tage.
[10] Bei beiden wird ein faseroptischer Aufbau und eine Einphotonenquelle auf Basis der parametrischen Fluoreszenz verwendet.
[11] Die Spezifikation des Faserkopplers gibt tatsächlich eine Schwankung von ± 2% für das Teilungsverhältnis an, allerdings ohne den Temperaturbereich anzugeben.

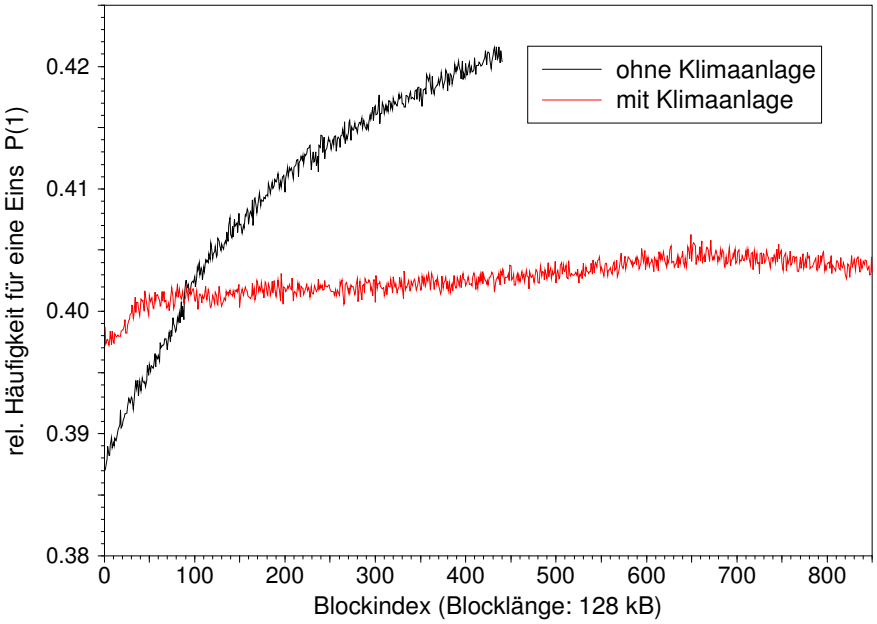


Abbildung B.5: Verlauf der geschätzten Wahrscheinlichkeit p_{Block} für eine Eins bei einer Blockgröße von $N = 128$ kB für einen Lauf mit *LF2* und einen ohne *LF1* Klimaanlage

kungen der relativen Häufigkeiten tatsächlich vom Detektorverhalten herrühren. Generell gilt aber auch bei Einsatz einer Klimaanlage, daß sich geringe Unterschiede im thermischen Verhalten der Signaldetektoren in einer Schwankungsbreite der relativen Häufigkeit von einem halben Prozent bemerkbar machen.
Abweichungen der relativen Häufigkeiten vom Idealwert $p = 0{,}5$ und *langsame* Schwankungen der Häufigkeiten, wie die oben dargestellten, lassen sich allerdings leicht durch die im praktischen Betrieb i.a. nachfolgende Regularisierung der Rohdaten vollständig beseitigen, s. Abschnitt G. Zur Illustration werden die Rohdaten des Laufes ohne Klimaanlage einer Von-Neumann-Regularisierung unterzogen und aus diesen Daten der Verlauf der geschätzten Wahrscheinlichkeit p_{Block} berechnet, s. Abb. B.6.
Wie man sieht, fluktuiert nun die relative Häufigkeit tatsächlich statistisch um den Idealwert $p = 0{,}5$, man beachte auch die feinere Ordinaten-Teilung. In Abb. B.7 ist die dazugehörige empirische Verteilungsfunktion der Einsen in 128 kB Blöcken dargestellt; auch sie liegt jetzt nah bei der theoretisch zu erwartenden Normalverteilung, was sich auch in entsprechenden Wahrscheinlichkeiten des Kolmogorov-Smirnov-Tests zeigt: $P(K_n^+) = 0{,}44$ und $P(K_n^-) = 0{,}72$. Man sieht deutlich, wie jetzt die empirische Verteilungsfunktion die Normalverteilung quasi „umspielt“ und nicht einseitig verschoben neben ihr liegt.

B.3 Resultate der Tests nach FIPS 140-1

Zuerst wird hier wieder der Lauf *LO2* getestet. Es werden 100 FIPS-Tests an 100 aufeinanderfolgenden, nicht überlappenden Stichproben[12] durchgeführt. Der Generator besteht dennoch jeden der Tests. Weicht allerdings bereits die Wahrscheinlichkeit für eine Eins stark vom Idealwert ab, wie dies z.B. aufgrund des nicht idealen Teilungsverhältnisses des verwendeten Mehrmoden-Faserkopplers bei allen Faseraufbauten der Fall ist, so ist natürlich auch nicht zu erwarten, daß

[12] Man beachte, daß es sich um Stichproben mit nicht regularisierten Rohdaten handelt und sich die Wahrscheinlichkeit für eine Eins, die allerdings nahe bei $p = 0{,}5$ liegt, gerade am Anfang relativ stark ändert, s. Abb. B.2 auf S. 155.

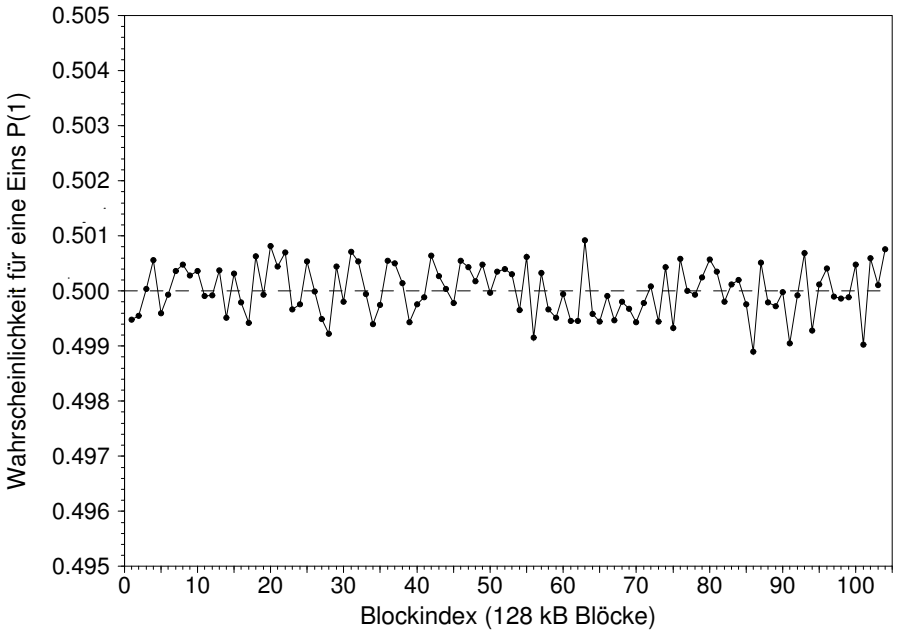


Abbildung B.6: Verlauf der geschätzten Wahrscheinlichkeit p_{Block} für eine Eins bei einer Blockgröße von $N = 128$ kB für die von-Neumann-regularisierten Daten des Laufes ohne Klimaanlage (s. Abb. B.5)

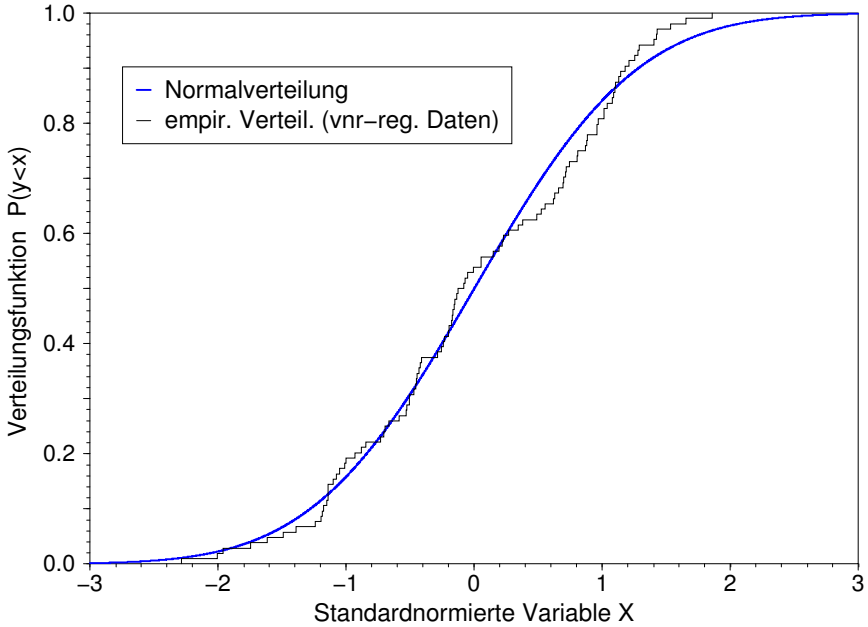


Abbildung B.7: Empirische und theoretische Verteilungsfunktion der standardnormierten Anzahl der Einsen in Blöcken der Länge N =128 kB für die von-Neumann-regularisierten Daten des Laufes ohne Klimaanlage (s. Abb. B.5)

die Bitsequenzen der Rohdaten die FIPS-Tests bestehen. Bei Durchführung von 100 FIPS-Tests am Anfang des Laufes *LF1* des faseroptischen Aufbaus des Generators wird erwartungsgemäß keiner der 100 Tests bestanden.
Um einen Vergleich zu haben und einschätzen zu können, „wie stark“ die FIPS-Tests sind, werden sie auch an Sequenzen der Vergleichs-Pseudozufallszahlengeneratoren durchgeführt. Es zeigt sich dabei, daß nur Sequenzen bekanntermaßen schlechter Generatoren, wie z.B. `RANDU`, die Tests nicht bestehen, alle besseren Pseudozufallszahlengeneratoren (aus den *Numerical Recipes*) bestehen die Tests problemlos, die auf starken Blockchiffren basierenden Zufallszahlengeneratoren selbstverständlich auch.
Somit kann man feststellen, daß die FIPS-Tests für die grundsätzliche Prüfung eines Zufallsgenerators nicht geeignet sind, sondern tatsächlich nur für ihren Hauptanwendungszweck: die Überprüfung eines Generators auf Hardware-Defekte beim Anschalten einer Chiffriereinheit. Da zu diesen Defekten allerdings auch stärkere Abweichungen von der idealen 50:50-Verteilung der Einsen und Nullen gehören, fallen erwartungsgemäß all diejenigen physikalischen Zufallsgeneratoren durch diese Tests, die stärkere Abweichungen vom idealen Teilungsverhältnis aufweisen, wie z.B. die faseroptischen Aufbauvarianten der quantenoptischen Zufallsgeneratoren. Andererseits bestehen physikalische Generatoren mit geringeren Abweichungen oder die Datensätze stark abweichender Generatoren nach einer Von-Neumann-Regularisierung die FIPS-Tests problemlos, selbst dann, wenn sie tieferliegende Defekte besitzen.
Zur Zeit wird daher auch im NIST an einem Nachfolge-Standard für den FIPS-140-1 gearbeitet; in diesem sollen dann auch stärkere Tests enthalten sein.

B.4 χ^2-Tests

Bei der Durchführung der χ^2-Tests lassen sich eine Reihe von Parametern frei wählen: die Länge der Blöcke, ihre Position im Zufallsstrom, die Stichprobengröße pro (einstufigem) Test und die Anzahl der Tests erster Stufe bei Durchführung zweistufiger Tests. Wie im Abschnitt zur Theorie bereits ausgeführt, ist es beim Test physikalisch erzeugter Zufallszahlen i. a. sinnvoll lückenlos aufeinanderfolgende Blöcke zu testen. Die Größe der Blöcke läßt sich zwar prinzipiell frei wählen, da sich etwaige Defekte aber bei Generatoren, die einzelne Bits generieren[13], besonders in Korrelationen zum nächsten Nachbarn bemerkbar machen sollten, werden die χ^2-Tests nicht auf Blöcken der Länge $n = 4$ wie beim Poker-Test aus den Tests nach FIPS 140-1 (s. Abschn. B.3) durchgeführt, sondern auf Bitpaaren. Wegen der besseren und willkürfreieren Interpretierbarkeit der Testresultate, s. Abschnitt 3.2.1.2, werden grundsätzlich zweistufige Tests durchgeführt. Die Stichprobengröße der Tests erster Stufe wird hierbei relativ groß gewählt, um etwaige statistische Probleme aufgrund zu kleiner Stichproben von vornherein auszuschließen. Da die Zufallsdaten in Dateien à einem Megabyte abgespeichert sind, erweist es sich als praktisch, bei der Durchführung von zweistufigen Tests 128 Tests mit einer Stichprobengröße von jeweils 8 kB, entsprechend einer Anzahl von 32.768 Bitpaaren, für jede Datei durchzuführen. Die Auswertung der zweiten Teststufe erfolgt mit einem nachfolgenden Kolmogorov-Smirnov-Test, wobei eine χ^2-Verteilung mit drei Freiheitsgraden für die theoretische Verteilung verwendet werden muß. Die Interpretation der Ergebnisse der Kolmogorov-Smirnov-Tests, d. h. was man genau als eine „starke“ Abweichung vom idealen Verhalten eines Zufallsgenerators ansieht, ist natürlich eine Ermessenssache, aber Werte für die Wahrscheinlichkeiten $P(K_n^+)$ und $P(K_n^-)$ kleiner 0,05 oder größer als 0,95 lassen sich zumindest als „verdächtig“ (s. a. [68], S. 46 ff) ansehen.
Wie bereits im Abschnitt 3.2.4 erwähnt, sollten bei Tests von Rohdaten die aus der Stichprobe geschätzten Werte für die Wahrscheinlichkeit für eine Eins verwendet werden; tut man dies nicht, führt bereits eine geringe Abweichung vom idealen Wert $p = 1/2$ zu einer massiven Abweichung der empirischen Verteilungsfunktion von der theoretischen. Zur Illustration sind in Abb.

[13] Bei Generatoren, die z.B. ein physikalisches Rauschsignal in einen mehrere Bits lange Zahl digitalisieren, wäre natürlich eine entsprechende Blockgröße angeraten.

B.8 zwei Tests auf identischen Stichproben des Laufs *LO2* des Generators mit einer durchaus nah am idealen Wert liegenden, geschätzten Einswahrscheinlichkeit von $\hat{p} = 0{,}50287$ dargestellt. Sie unterscheiden sich nur darin, daß beim einen die geschätzte Einswahrscheinlichkeit zur Ermittlung der empirischen Verteilungsfunktion verwendet wird und beim anderen die ideale mit $p = 1/2$.

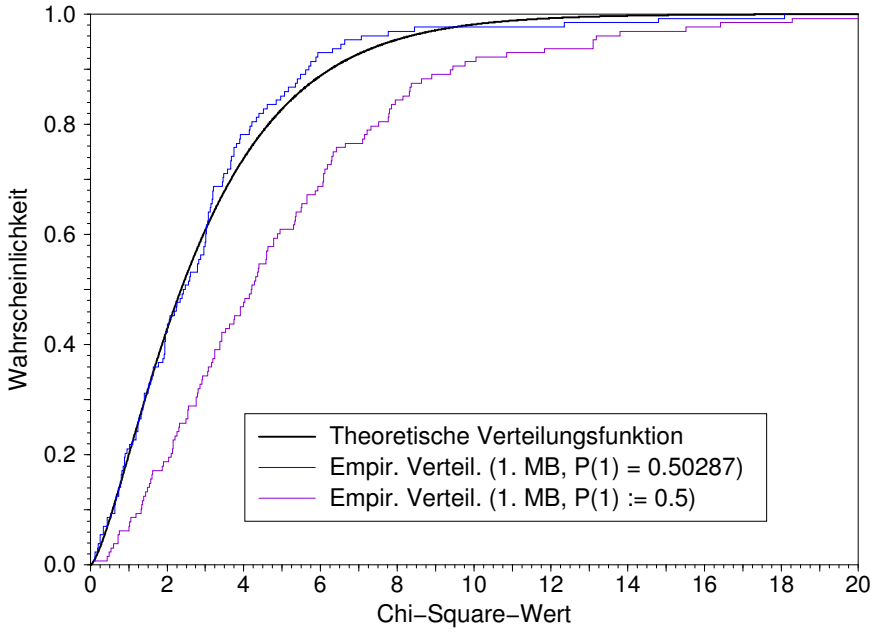


Abbildung B.8: Empirische Verteilungsfunktionen zweier zweistufiger χ^2-Tests einer identischen Stichprobe aus Lauf *LO2* im Vergleich zu theoretischen Verteilungsfunktion

Die Ergebnisse der Kolmogorov-Smirnov-Tests fassen den visuellen Eindruck in Zahlen: Für den Test mit der geschätzten Einswahrscheinlichkeit sind die Werte für die Wahrscheinlichkeit, daß die maximale positive bzw. negative Abweichung von der theoretischen Kurve kleiner als der aus der Stichprobe berechnete Wert ist, mit $P(K^+_{128}) = 0{,}53046$ bzw. $P(K^-_{128}) = 0{,}41618$ nicht zu beanstanden, die entsprechenden Wahrscheinlichkeiten hingegen für die mit der idealen Einswahrscheinlichkeit ermittelten Verteilungsfunktion berechnen sich zu $P(K^+_{128}) = 0{,}0000$ bzw. $P(K^-_{128}) = 1{,}0000$. Es ist also ausgeschlossen, daß die maximale positive Abweichung noch kleiner ist, aber dafür sicher, daß es die negative ist. Für Läufe mit noch stärkerer Abweichung vom idealen Wert (z.B. 48:52 Verhältnis zwischen Nullen und Einsen) weichen die empirischen Verteilungsfunktionen natürlich noch stärker von der theoretischen Verteilungsfunktion ab; daher muß beim Test von Rohdaten immer die Einswahrscheinlichkeit aus der Stichprobe geschätzt werden.

B.4.1 Ergebnisse der χ^2-Tests

In Abb. B.9 sind die empirischen Verteilungsfunktionen von 3 aufeinanderfolgenden zweistufigen χ^2-Tests mit den oben aufgeführten Testparametern auf aufeinanderfolgenden Stichproben von je 1 MB dargestellt; in der Tabelle B.1 finden sich die dazugehörigen Ergebnisse der Kolmogorov-Smirnov-Tests. Für diese Tests wurde bewußt der Lauf *LD1* des Generators gewählt, bei dem aufgrund von Problemen mit der kompakten Datenaufnahme-Elektronik Antikorrelationen zwischen aufeinanderfolgenden Bits bestehen. Diese Korrelationen sollten sich eigentlich auch in den Ergebnissen der χ^2-Tests bemerkbar machen.
Wie man allerdings sieht, sind keine starken Abweichungen der empirischen Verteilungsfunk-

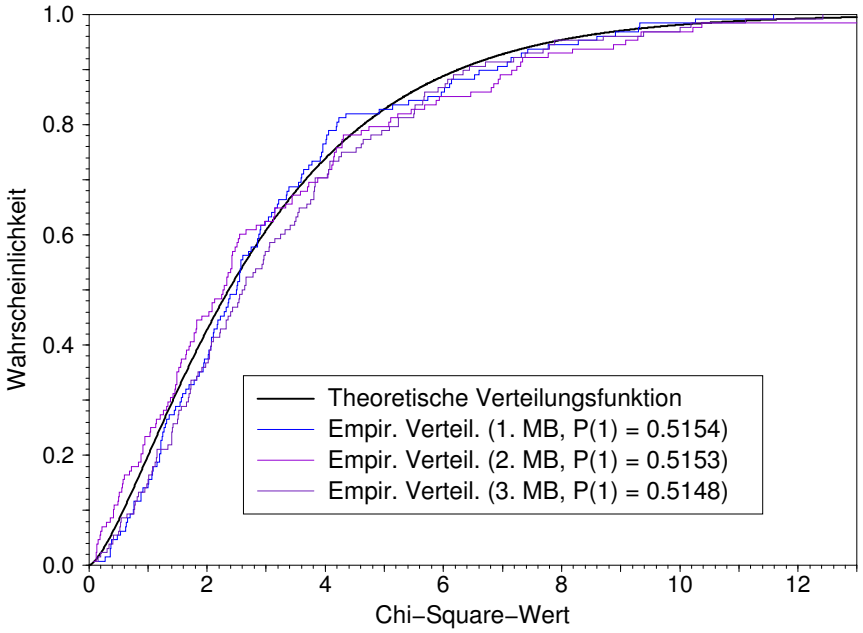


Abbildung B.9: Empirische Verteilungsfunktionen für drei zweistufige χ^2-Tests auf korrelationsbehafteten Rohdaten des Laufes *LD1* im Vergleich zur theoretischen Verteilungsfunktion

tionen von der theoretischen Verteilungsfunktion zu erkennen; auch die Kolmogorov-Smirnov-Auswertung der Chi-Square-Tests in Tabelle B.1 liefert keine überzeugenden[14] Hinweise für eine starke Abweichung.

Stichprobe	K^+_{128}	$P(K^+_{128})$	K^-_{128}	$P(K^-_{128})$
1. MB	0.571167	0.49677	0.691579	0.63145
2. MB	0.763990	0.70282	0.706115	0.64644
3. MB	0.068718	0.01341	0.853171	0.77851
4. MB	0.576971	0.50360	0.341586	0.22407
5. MB	1.085264	0.91123	0.419530	0.31411

Tabelle B.1: Ergebnisse der Kolmogorov-Smirnov-Auswertung der ersten fünf zweistufigen χ^2-Tests auf korrelationsbehafteten Rohdaten des Laufes *LD1*

Um die in diesen Daten bekanntermaßen (s. Abschnitt B.9.3) vorhandenen Antikorrelationen zu entdecken, ist ein zweistufiger χ^2-Tests bei gleicher Stichprobengröße offensichtlich nicht geeignet, deshalb wird auch darauf verzichtet, andere Läufe der verschiedenen Generatorvarianten auf diese Weise zu testen.
Abschließend sei noch kurz erwähnt, wie die Vergleichspseudozufallsgeneratoren bei zweistufigen χ^2-Tests[15] abschneiden. Lediglich der bekanntermaßen schwache Pseudozufallsgenerators `RANDU` besteht den Test nicht, alle Tests der anderen Vergleichsgeneratoren zeigen keine all-

[14] Lediglich die positive Abweichung beim dritten Test ist recht niedrig, aber solch ein einzelner Wert kann auch bei ideal zufälligen Generatoren vorkommen.

[15] Es werden dieselben Testparameter wie bei den Tests der Daten des quantenoptischen Zufallsgenerators verwendet, wobei allerdings für die Einswahrscheinlichkeit der ideale Wert von $p = 0{,}5$ genommen wird. Es wird aber nicht versucht, die Tests gezielt auf bestimmte Schwächen der Pseudozufallsgeneratoren hin zu optimieren!

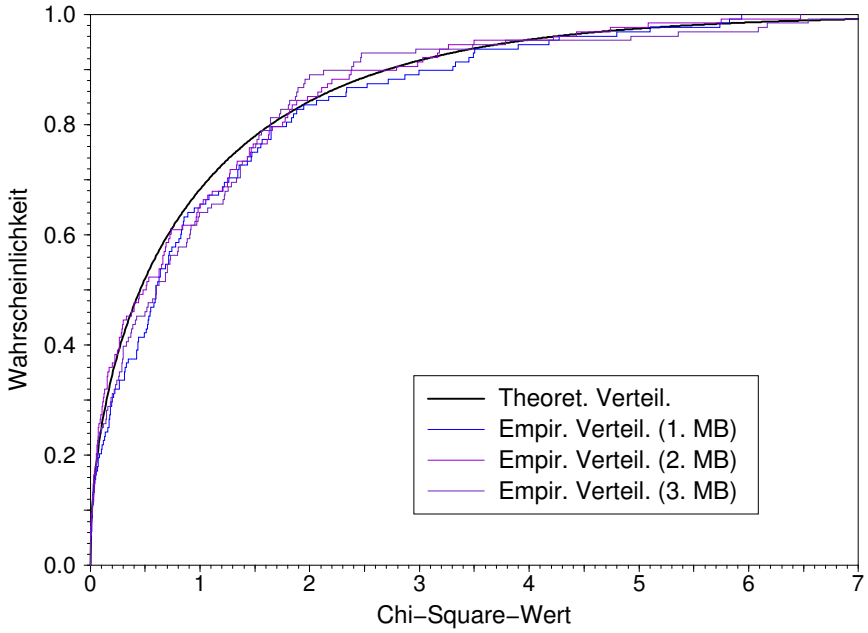


Abbildung B.10: Empirische Verteilungsfunktionen für drei zweistufige Kontingenz-Tests auf korrelationsbehafteten Rohdaten des Laufes *LD1* im Vergleich zur theoretischen Verteilungsfunktion(χ^2-Verteilung mit einem Freiheitsgrad)

zu „verdächtigen“ Abweichungen[16] der empirischen Verteilungsfunktionen von der theoretischen Verteilungsfunktion.

B.5 Kontingenztests

Für die zweistufigen Kontingenztests werden ähnliche Parameter gewählt, wie bei den zweistufigen χ^2-Tests, d.h. es werden jeweils 8 kB große Stichproben (entsprechend 32.768 Bitpaaren) für die 128 Tests der ersten Stufe verwendet.
Abb. B.10 zeigt die Ergebnisse der Kontingenztests auf der gleichen Stichprobe (Anfang des Laufs *LD1*) wie die χ^2-Tests, Tab. B.2 listet die entsprechenden Werte der Kolmogorov-Smirnov-Tests auf.

Stichprobe	K^+_{128}	$P(K^+_{128})$	K^-_{128}	$P(K^-_{128})$
1. MB	0.167678	0.063959	1.223238	0.953462
2. MB	0.453671	0.354537	0.616895	0.549162
3. MB	0.541573	0.460863	0.871372	0.791816

Tabelle B.2: Ergebnisse der Kolmogorov-Smirnov-Auswertung der ersten drei zweistufigen Kontingenz-Tests auf korrelationsbehafteten Rohdaten des Laufes *LD1*

Lediglich der erste Test sieht bei beiden Abweichungen etwas „verdächtig“ aus, die anderen sind nicht zu beanstanden. Man erkennt also, daß man gegenüber den χ^2-Tests etwas an Mächtigkeit gewinnt, aber nicht genug, um die innerhalb dieses Laufes vorhandenen Korrelationen klar

[16] `RANO` zeigt bei zwei von fünf Tests relativ kleine Werte für $P(K^+_{128})$, insgesamt ist bei ihm eine gewisse Tendenz zu größeren positiven Abweichungen hin zu erkennen.

aufzudecken. Deshalb wurde auch bei den Tests anderer Läufe auf Kontingenztests verzichtet.
Die Ergebnisse bei den Vergleichspseudozufallsgeneratoren liegen ähnlich. Interessant ist, daß `RANDU` zwar immer noch als statistisch schlecht erkannt wird, aber nicht in dem Maße wie bei den χ^2-Tests, so daß sein schlechtes Abschneiden dort offensichtlich auf eine stärkere Abweichung von den idealen Werten für die Einswahrscheinlichkeit und nicht so sehr auf andere Defekte, wie Korrelationen, zurückzuführen ist, s. folgender Abschnitt.

B.6 Die Mächtigkeit von χ^2- und Kontingenz-Tests

Aufgrund des ähnlichen grundlegenden Aufbaus des Kontingenztest (bei zwei Merkmalen) und dem χ^2-Test auf Bitpaaren, ist es nicht verwunderlich, daß sich die beiden Tests in ihrer Aussagekraft nur wenig voneinander unterscheiden. Ein Vorteil des Kontingenztests sei allerdings an dieser Stelle erwähnt: Anders als beim χ^2-Test gehen bei ihm von vornherein nur aus der Stichprobe geschätzte Wahrscheinlichkeiten in die Berechnung der Statistik ein, so daß ein Abweichen von idealen Werten nicht extra berücksichtigt werden muß.
Insgesamt betrachtet ist das schlechte Abschneiden der (zweistufigen) χ^2- und Kontingenz-Tests bei bekanntermaßen (anti-)korrelationsbehafteten Bitsequenzen erstaunlich, gerade aufgrund der im Vergleich zu üblichen statistischen Anwendungen recht großen Stichproben. Dies legt zumindest die Vermutung nahe, daß die weiter unten beschriebenen, zweistufigen Autokorrelationskoeffiziententests auch bei anderen statistischen Untersuchungen, bei denen nach (schwachen) Korrelationseffekten bzw. dem gemeinsamen Auftreten von Merkmalen gesucht wird, eventuell von Nutzen sein könnten.

B.7 Universelle Tests nach Maurer und Coron

Die universellen Tests nach Maurer bzw. Coron werden im folgenden immer auf 8 Bit langen Blöcken durchgeführt, da diese einerseits keine übermäßig großen Stichproben verlangen und die Tests sich mit ein Byte großen Blöcken besonders effizient programmieren lassen.
Zuerst werden die beiden Varianten des universellen Tests auf einem Lauf eines kryptographisch starken Pseudozufallszahlengenerators[17]. Hierbei wird die jeweils berechnete Teststatistik nicht nur für den gesamten Lauf, sondern auch für Zwischenwerte in festen Intervallen ermittelt, um überprüfen zu können, wie sich die Stichprobenlänge auf sie auswirkt.
In Abb. B.11 und B.12 sind die Werte der Teststatistik für einen universellen Test nach Maurer auf 8 Bit langen Blöcken dieses Laufes dargestellt; hierbei sind in Abb. B.11 die Ergebnisse der Teststatistik der ersten 5 MB des Laufes in Intervallen von 1024 Blöcken abgebildet und in Abb. B.12 die Werte für den gesamten Lauf mit einer Länge von 20 MB in Intervallen von 65536 Blöcken. Die entsprechenden Ergebnisse für einen universellen Test nach Coron sind in Abb. B.13 und B.14 dargestellt. In allen Abbildungen sind neben den empirischen Werten für die Teststatistik noch ihr Erwartungswert und die Standardabweichungen in negative und positive Richtung als Kurven eingezeichnet. Die Werte der Teststatistik für den gesamten Lauf betragen $\hat{f}_U(R^N) = 7{,}184048$ für die Maurer-Variante und $\hat{H}_{Block} = 8{,}000393$ für die Coron-Variante, die zugehörigen Werte für die Ablehnungsquote, betragen $\rho_{Mau} = 0{,}115$ und $\rho_{Cor} = 0{,}112$, so daß beide Tests den DES-Generator als hinreichend zufällig akzeptieren. Vergleicht man die Ergebnisse der Tests nach Maurer bzw. Coron miteinander, so fällt auf, daß sich der grundlegende Verlauf der Teststatistiken gleicht. Es ist daher vollkommen ausreichend, nur jeweils eine der beiden Varianten des universellen Tests zu verwenden; aufgrund seiner Vorteile, s. Abschnitt E.3, wird für die weiteren Untersuchungen die Testvariante von Coron gewählt.
Betrachtet man die Testergebnisse, so fallen starke Fluktuationen zu Anfang der Tests auf,

[17] Der Zufallsgenerator verwendet DES im Counter-Modus.

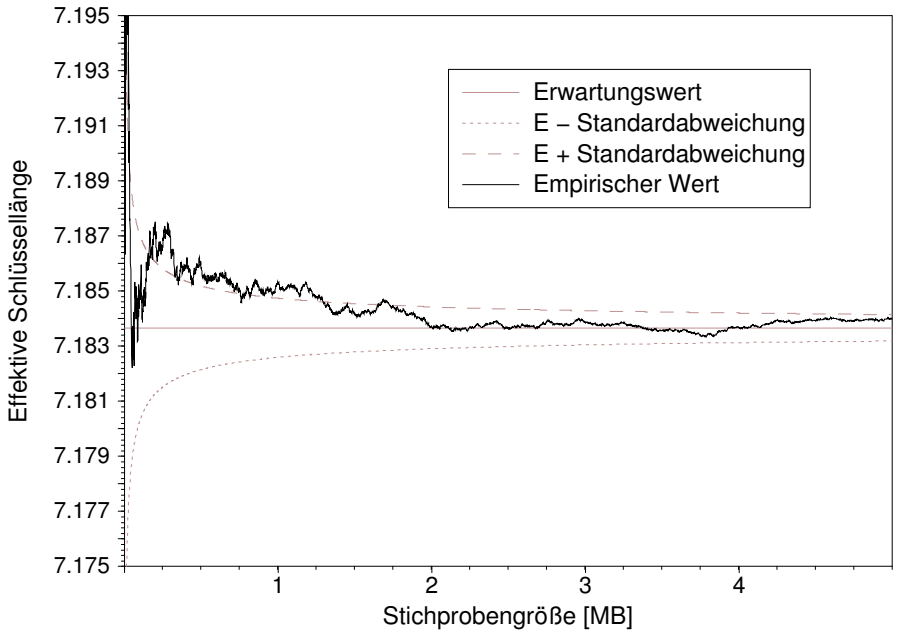


Abbildung B.11: Teststatistik des universellen Tests nach MAURER auf 8 Bit langen Blöcken in Abhängigkeit von der Stichprobenlänge für die ersten 5 MB eines DES-Pseudozufallsgenerator-Laufes, Ausgabeintervall: alle 1024 Blöcke

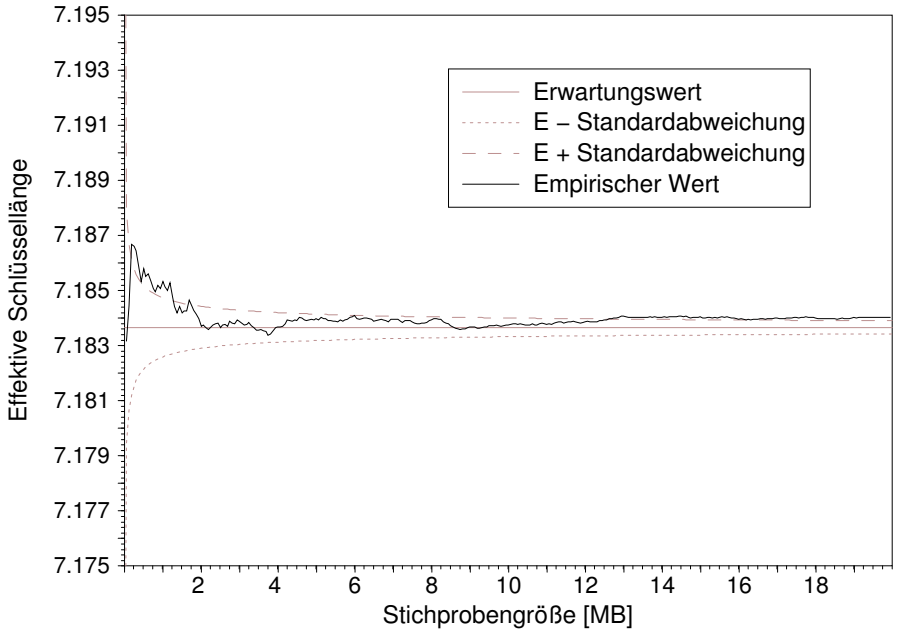


Abbildung B.12: Teststatistik des universellen Tests nach MAURER auf 8 Bit langen Blöcken in Abhängigkeit von der Stichprobenlänge für 20 MB eines DES-Pseudozufallsgenerator-Laufes, Ausgabeintervall: alle 65536 Blöcke

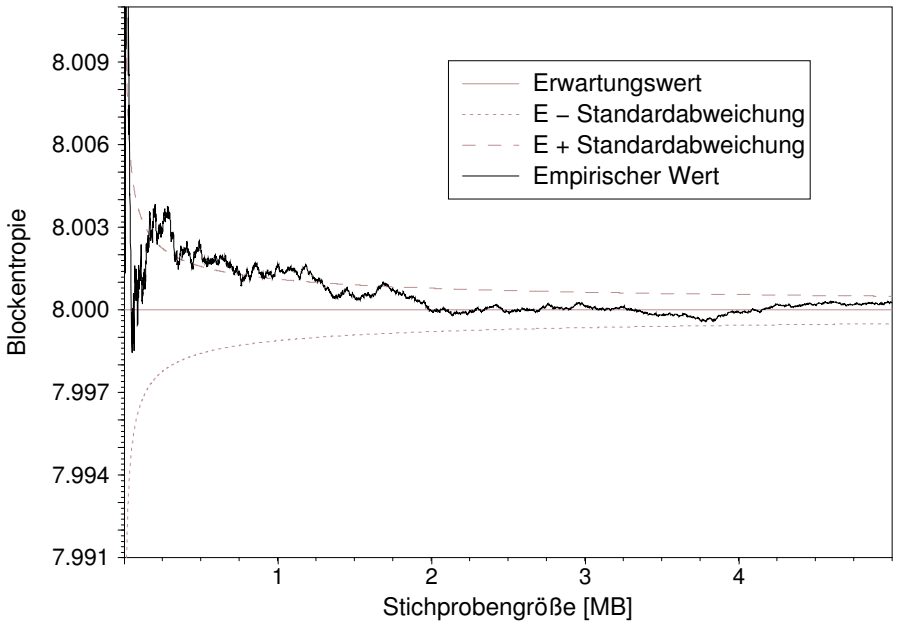


Abbildung B.13: Teststatistik des universellen Tests nach Coron auf 8 Bit langen Blöcken in Abhängigkeit von der Stichprobenlänge für die ersten 5 MB eines DES-Pseudozufallsgenerator-Laufes, Ausgabeintervall: alle 1024 Blöcke

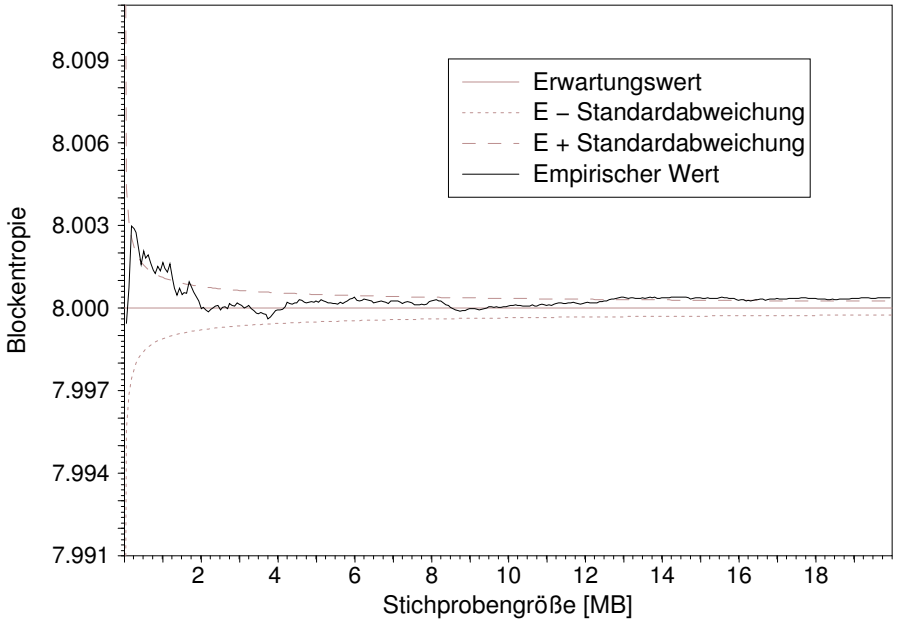


Abbildung B.14: Teststatistik des universellen Tests nach Maurer auf 8 Bit langen Blöcken in Abhängigkeit von der Stichprobenlänge für 20 MB eines DES-Pseudozufallsgenerator-Laufes, Ausgabeintervall: alle 65536 Blöcke

die durchaus noch weiter reichen als bis zu den minimalen Werten[18] für K. Es ist also bei der praktischen Durchführung eines universellen Tests unbedingt darauf zu achten, daß entsprechend große Stichproben[19], z. B. $K = 10000 \cdot 2^L$, zur Verfügung stehen.
Wie im Abschnitt 3.2.7 zur Theorie der universellen Tests erwähnt, lassen sich die universellen Tests schlecht auf Rohdaten durchführen. Zur Illustration und um zu prüfen, wie sich der Coron-Test bei unsymmetrischen Quellen verhält, werden allerdings trotzdem mehrere Tests mit Rohdaten-Läufen durchgeführt. Hierbei müssen allerdings Läufe gewählt werden, die ein Teilungsverhältnis in der Nähe von 50:50 haben, d.h. die Daten, welche von einem quantenoptischen Zufallsgenerator mit faseroptischen Aufbau erzeugt wurden, kommen aufgrund ihres in der Nähe 40:60 liegenden Teilungsverhältnisses gar nicht erst in Frage.
In Abb. B.15 ist der Verlauf der Coron-Teststatistik für die ersten 20 MB des Laufes *LO3* dargestellt. Wie wegen des nicht idealen Teilungsverhältnisses[20] zu erwarten ist, wird der maximale Wert der (Block-) Entropie nicht erreicht, sondern lediglich ein Wert von $\hat{H}_{Block} = 7{,}994638$ und der Test liefert eine Ablehnungsquote[21] von $\rho_{LO3} = 0{,}000000$. Dies bedeutet, daß nur bei einer extrem niedrigen Ablehnungsquote der Lauf noch als akzeptabler Lauf eines idealen Zufallsgenerators akzeptiert würde.

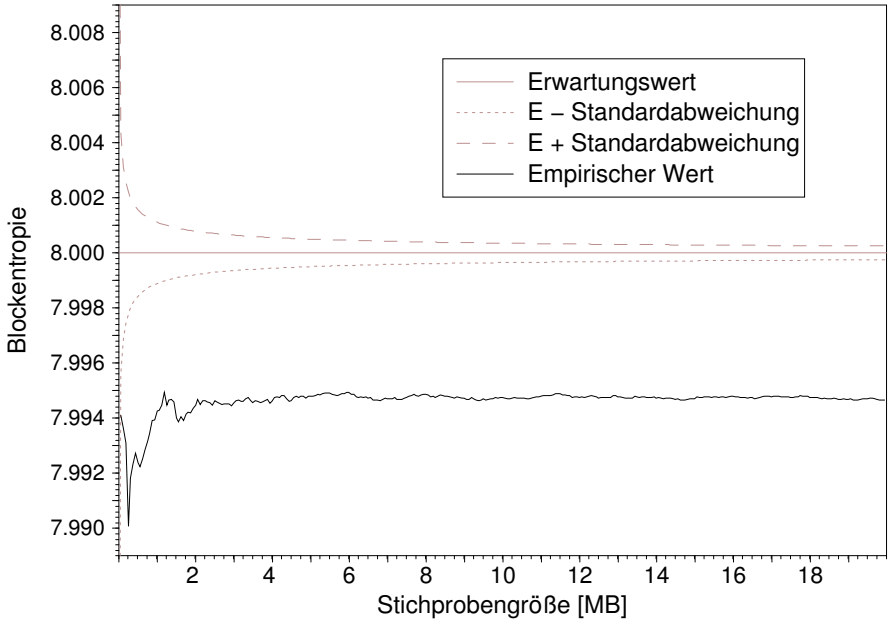


Abbildung B.15: Teststatistik des universellen Test nach Coron auf 8 Bit langen Blöcken der ersten 20 MB des Laufes *LO3*, Ausgabeintervall: alle 65536 Blöcke

Interessanterweise ist es aber durchaus nicht immer so, daß der universelle Tests einen Lauf einer unsymmetrischen Quelle gleich deutlich erkennt. In Abb. B.16 sind die empirischen Werte für die ersten 5 MB des Laufes *LO2* dargestellt, diese Testwerte befinden sich noch vollständig innerhalb

[18] Typischerweise, vgl. Abschnitt 3.2.7, wird für einen universellen Test eine Vorlaufsequenz der Länge $Q = 10 \cdot 2^L$ und eine Testsequenz mit einer Mindestlänge $K = 1000 \cdot 2^L$ benötigt; bei der hier verwendeten Blocklänge von 8 Bit bedeutet dies, daß $Q = 2560$ und $K > 256000$ Blöcke (Bytes) lang sein sollten.

[19] Daher nehmen die praktischen Einsatzmöglichkeiten der universellen Tests, s. a. [51], mit zunehmender Blocklänge noch stärker ab, als mit Blick auf die Mindestwerte für K zu erwarten wäre.

[20] Die Wahrscheinlichkeit für eine Eins beträgt für die ersten 20 MB: $p = 0{,}4846 \pm 0{,}0021$, wobei die Standardabweichung auf 8 kB-Blöcken berechnet wurde.

[21] Die Ablehnungsquote wird nur auf sechs Nachkommastellen genau berechnet, mehr wäre auch nicht sinnvoll.

des Intervalls einer Standardabweichung um den Mittelwert. Dies ist umso erstaunlicher, als es sich um Daten aus einem Lauf mit einer deutlichen Veränderung der Wahrscheinlichkeit für den Bitwert Eins handelt, s. Abschnitt B.2, Abb. B.2.

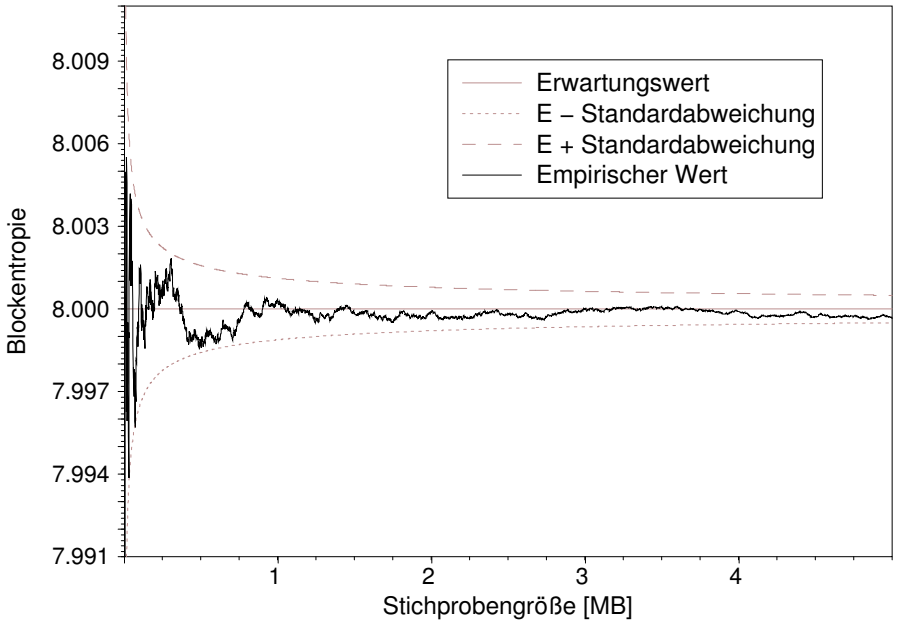


Abbildung B.16: Teststatistik des universellen Test nach Coron auf 8 Bit langen Blöcken der ersten 5 MB des Laufes *LO2*, Ausgabeintervall: alle 1024 Blöcke

Betrachtet man sich allerdings den Verlauf der Teststatistik für die ersten 20 MB des Laufes, s. Abb. B.17, so sieht man deutliche Abweichungen in den späteren Abschnitten des Laufes. Ab etwa dem achten MB machen sich diese in deutlichen, einseitig zunehmenden Abweichungen vom Erwartungswert bemerkbar. Dementsprechend ist die Ablehnungsquote, die zu dem stark abweichenden Wert der Teststatistik nach 20 MB von $\hat{H}_{Block} = 7{,}997335$ gehört, ähnlich klein wie bei dem Lauf *LO3*. Interessanterweise findet sich an der Stelle, an der die Abweichungen zunehmen, im Verlauf der Einswahrscheinlichkeit dieses Laufes, s. Abb. B.2, keine markante Änderung.
Vergleicht man aber die mittlere Wahrscheinlichkeit für eine Eins innerhalb der ersten 5 MB des Laufes *LO2*, die $p_{5MB} = 0{,}5052 \pm 0{,}0025$ beträgt, mit der mittleren Wahrscheinlichkeit für die ersten 20 MB, die bei einem Wert $p_{20MB} = 0{,}5103 \pm 0{,}0041$ liegt, so stellt man fest, daß es anscheinend nicht die dynamische Veränderung der Wahrscheinlichkeit ist, auf die der Coron-Test anspricht, sondern die Abweichung der Wahrscheinlichkeit vom idealen Wert $p = 0{,}5$.
Es stellt sich daher die Frage, inwieweit der Coron-Test in der Lage ist, Korrelationen in einem Lauf aufzudecken. Glücklicherweise liegt bei dem Testlauf *LN2* die Wahrscheinlichkeit für eine Eins recht nah[22] beim idealen Wert. Da dieser Lauf mit dem Autokorrelationskoeffizententest deutlich erkennbare Antikorrelationen aufweist, ist er somit ein guter Kandidat, um die Eignung des Coron-Tests zum Aufdecken von Korrelationen zu prüfen.
In Abb. B.18 ist der Verlauf der Teststatistik für den nur 5 MB umfassenden Lauf dargestellt. Man sieht: Der Coron-Test ist nicht in der Lage, die Antikorrelationen aufzudecken. Mit einem Wert für die Teststatistik von $\hat{H}_{Block} = 7{,}999884$ und einer entsprechend günstigen Ablehnungsquote von $\rho_{LN2} = 0{,}816$ wird der Lauf nicht als fehlerbehaftet erkannt.
Der Coron-Test wurde auch noch auf von-Neumann-regularisierten Daten durchgeführt, insbe-

[22] Die mittlere Wahrscheinlichkeit für eine Eins beträgt $p = 0{,}4995 \pm 0{,}002$.

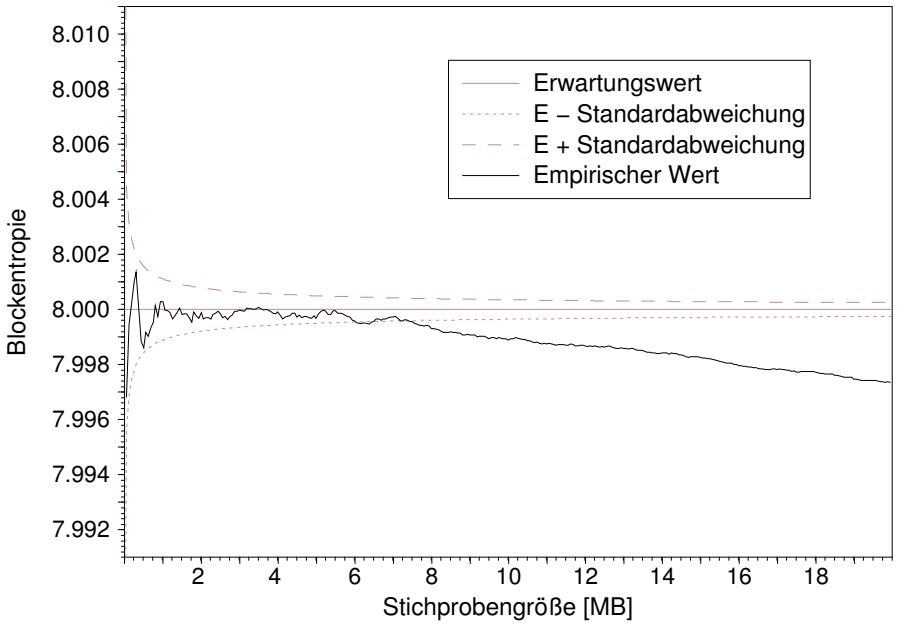


Abbildung B.17: Teststatistik des universellen Test nach Coron auf 8 Bit langen Blöcken der ersten 20 MB des Laufes *LO2*, Ausgabeintervall: alle 65536 Blöcke

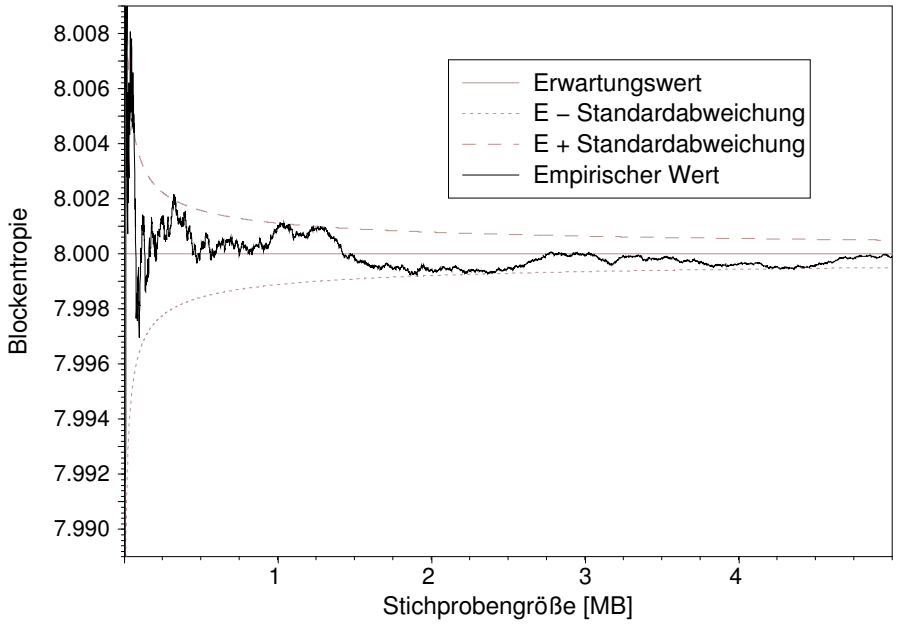


Abbildung B.18: Teststatistik des universellen Test nach Coron auf 8 Bit langen Blöcken des (anti)-korrelationsbehafteten Laufes *LN2*, Ausgabeintervall: alle 1024 Blöcke

sondere auf den bereits bei der Datenaufnahme regularisierten frühen Läufen[23]. Wie aufgrund der vorangehenden Ergebnisse allerdings nicht anderes zu vermuten, führen die idealen Wahrscheinlichkeiten nach der Regularisierung auch zu entsprechend guten Testergebnissen, daher wird hier auch auf eine Darstellung verzichtet.
Als Fazit läßt sich festhalten, daß universelle Tests sowohl in der Maurer- als auch der Coron-Variante, nicht dazu geeignet sind, Antikorrelationen aufzudecken, die sich selbst bei einer erheblich kleineren Stichprobe mit einem Autokorrelationskoeffizententests problemlos aufdecken lassen.
Betrachtet man die Ergebnisse der universellen Tests, so fällt nicht nur auf, daß sie verhältnismäßig große Stichproben benötigen, wie bereits in der Veröffentlichung von Gustafson et al. [51] für große Blocklängen bemängelt wurde, sondern daß ihre Ergebnisse in weitaus stärkerem Maße durch die relativen Häufigkeiten der Bits als durch die statistischen Abhängigkeiten innerhalb der Bitblöcke bestimmt werden.

B.8 Komplexitätstests

Wie kann eine einzelne Bitsequenz zufällig genannt werden, wenn doch jede einzelne Sequenz die gleiche Erzeugungswahrscheinlichkeit hat? Dieses Paradoxon läßt sich mit Hilfe der *algorithmischen Komplexität*[24] „auflösen".
Die algorithmische Komplexität formalisiert hierbei den intuitiven Begriff vom Zufall als regellosem, nicht genauer beschreibbaren Verhalten, indem sie die Zufälligkeit durch die Länge des kürzesten Programmes auf einer *universellen Turingmaschine* mißt, welches die zu untersuchende Bitsequenz erzeugt. Die „wirklich" und nicht nur pseudozufälligen Sequenzen sind hierbei diejenigen, deren algorithmische Komplexität gerade ihrer Länge entspricht, d.h. sie sind so regellos, daß sie nur genannt, aber nicht effizienter generiert werden können. Umgekehrt bedeutet dies nun aber auch, daß sich alle nicht „wirklich" zufälligen Sequenzen komprimieren lassen, nämlich gerade auf die Länge des kürzesten Programmes, das sie generiert.
Der Begriff der algorithmischen Komplexität ist recht radikal [23], so weist er natürlich *allen* von Pseudozufallsgeneratoren erzeugten Bitsequenzen eine relativ kleine Komplexität zu, da sie sich offensichtlich recht kompakt, nämlich durch Angabe des Programmes des Generators und einigen wenigen zusätzlichen Parametern, beschreiben lassen. Betrachtet man nun aber eine bestimmte Klasse von Zufallsgeneratoren, wie z.B. die linearen Kongruenzgeneratoren, s. Abschnitt B.1, so weisen die von ihnen erzeugten Sequenzen bei empirischen Tests der Zufallszahlen starke „Qualitätsunterschiede" auf, je nachdem welche Parameter man für den Generator wählt. Ihre algorithmische Komplexität wird sich dennoch kaum unterscheiden. Erschwerend kommt noch hinzu, daß die algorithmische Komplexität beim Durchlaufen[25] einer Sequenz stark schwanken kann [82].
So elegant und intuitiv der Begriff der algorithmischen Komplexität auch ist, so unbrauchbar ist er für praktische Tests auf Zufälligkeit, denn leider läßt sich die algorithmische Komplexität grundsätzlich nicht berechnen [82]. Praktikable, d.h. effizient berechenbare, Maße für „die"[26]

[23] Eigentlich war noch geplant zu prüfen, inwieweit nachträglich von-Neumann-regularisierte, korrelationsbehaftete Läufe beim Coron-Test als (leicht) fehlerbehaftet aufgedeckt werden können; aufgrund der oben aufgeführten Ergebnisse der Coron-Tests von korrelationsbehafteten Läufen mit Wahrscheinlichkeiten in der Nähe der idealen Werte erschien dies aber nicht mehr sinnvoll.

[24] Sie wird nach den beiden Entdeckern [21, 117] auch gern *Kolmogorov-Chaitin-Komplexität* genannt.

[25] Man spricht dann vom *Komplexitätsprofil*; bei ihm wird für jede Subsequenz, die vom Anfang bis zu einer fortlaufenden Position der zu untersuchenden Bitsequenz reicht, ein Komplexitätsmaß berechnet. Bei der algorithmischen Komplexität werden die Schwankungen allerdings theoretisch hergeleitet.

[26] *Die* algorithmische Komplexität in ihrem ursprünglichen Sinn ist gerade deswegen ein eindeutiges Maß, weil sie auf einer *universellen Turingmaschine*, die jeden anderen Computer emulieren kann, definiert ist.

algorithmische Komplexität kann man nur dadurch erreichen, daß man algorithmische Beschreibungen auf einer weniger mächtigen Maschine betrachtet, so daß eine Berechnung möglich wird. Auf diese Weise lassen sich Komplexitätsmaße *relativ zu bestimmten Automatentypen* definieren[27]:

- Die *Lineare Komplexität* [111], bei der die Länge des kürzesten linear rückgekoppelten Schieberegisters, das die untersuchte Folge erzeugt, als Maß für die Komplexität verwendet wird.
- Die *Maximal-Ordnungs-Komplexität* [61, 62], bei der die Länge des kürzesten Schieberegisters mit allgemeiner, möglicherweise nichtlinearer, aber gedächtnisloser Rückkopplung als Maß dient.
- Die *Lempel-Ziv* oder *Sequenzkomplexität*[28] [80, 94, 134], bei der ein endlicher Zustandsautomat zugrunde gelegt wird und die man als ein Maß für die Rate ansehen kann, mit der neue Bitmuster beim Durchlaufen der Testsequenz auftreten.
- Die „*grammatikalische Komplexität*“[29] [22, 79, 97, 98], bei der die Länge der Regeln der kürzesten kontextfreien Grammatik, welche die untersuchte Folge beschreibt, als Komplexitätsmaß dient. Die exakte Berechnung der „grammatikalischen Komplexität“ ist ein NP-vollständiges Problem [116], aber glücklicherweise gibt es seit kurzem effiziente Algorithmen, die es erlauben, dieses Maß in guter Näherung zu berechnen [79].

Leider können die so definierten Maße nur von beschränktem Nutzen sein, da sie teilweise so speziell sind, daß sie zum Teil selbst eindeutig nicht zufälligen Sequenzen dennoch eine hohe Komplexität zuweisen, lediglich weil sie sich *auf der entsprechenden Referenzmaschine* nicht kompakter als durch die vollständige Wiedergabe der Originalsequenz reproduzieren lassen.
Überdies ist bei grammatikalischen Komplexitätsmaßen zu bedenken, daß ihnen Ersetzungsregeln zugrunde liegen, mit deren Hilfe sich wiederholende Muster in Sequenzen beschreiben lassen. Betrachtet man aber z.B. die Ziffernfolge von π bis zu einer bestimmten Länge, so wird diese vermutlich dennoch ein große Komplexität haben, obwohl es kompakte Formeln zur Berechnung gibt.

B.8.1 Lineare Komplexitätstests

Bei diesen Tests wird mit Hilfe des Berlekamp-Massey-Algorithmus, s. z.B. [89, 91, 111], für eine gegebene Bitsequenz das kürzeste, linear rückgekoppelte Schieberegister ermittelt, das mit entsprechenden Startwerten diese Sequenz generieren könnte. Die Länge dieses Schieberegisters wird dann als *lineare Komplexität* bezeichnet und ist ein relatives Maß für die Komplexität der Sequenz [111]. Interessanterweise gibt es eine Veröffentlichung von Dai et al. [12] in der bewiesen wird, daß die lineare Komplexität im Limes unendlich langer Sequenzen im Mittel halb so groß wie die algorithmische Komplexität ist.
Aussagekräftiger, aber auch aufwendiger in der Berechnung als die lineare Komplexität, ist das *lineare Komplexitätsprofil* [111], bei dem die lineare Komplexität nicht nur für die gesamte Sequenz bestimmt wird, sondern sukzessive für alle Positionen beim Durchlaufen der Bitsequenz.
Die linearen Komplexitätstests wurden zwar programmiert, es stellte sich aber bereits bei den Voruntersuchungen mit bekanntermaßen schlechten Pseudozufallsgeneratoren heraus, daß diese bereits eine große lineare Komplexität besitzen und somit im Bezug auf dieses Maß als gute Zufallsgeneratoren anzusehen gewesen wären. Auch eine Untersuchung des Komplexitätsprofiles

[27] Die Aufzählung erfolgt in aufsteigender Mächtigkeit der Automaten.
[28] Die Bezeichnung *Sequenzkomplexität* bzw. englisch *sequency* ist etwas unglücklich, da sich Komplexitätsmaße für Zufallssequenzen immer auf eine Sequenz beziehen.
[29] Eine präzisere, aber dafür etwas unhandlichere Bezeichnung wäre: „approximative kontextfreie Grammatik-Komplexität“; der ihr zugrunde liegende Automat ist ein sogenannter „Kellerautomat“.

lieferte keine deutlichen Hinweise auf die schlechten Eigenschaften dieser Generatoren. Daher wurden die linearen Komplexitätstests für die Untersuchung der Datensätze als untauglich verworfen[30], sinnvoll einsetzbar sind sie natürlich dennoch immer dann, wenn tatsächlich linear rückgekoppelte Schieberegister als Grundkomponenten für einen Pseudozufallsgenerator eingesetzt werden.
Aufgrund der schlechten Erfahrungen mit der linearen Komplexität wurden dann ähnlich geartete Tests, die auf verallgemeinerten, aber ebenfalls auf Schieberegistern basierten Komplexitätsmaßen gründen, wie die Maximal-Ordnungs-Komplexität[31] [61, 62], auch nicht programmiert und durchgeführt.

B.8.2 Kompressionstests

Ein sehr einfacher, leicht durchzuführender Test auf Zufälligkeit, welcher die in Abschnitt B.8 erwähnte Idee der schlechten Komprimierbarkeit von zufälligen Sequenzen verwendet, besteht darin, daß man versucht, eine Sequenz mit einem Kompressionsverfahren zu komprimieren und anschließend das Kompressionsverhältnis K_V ermittelt, d.h. den Quotienten aus der Länge der komprimierten Sequenz L_k zur Länge der Originalsequenz L_o bildet: $K_V = L_k/L_o$. Für eine Zufallssequenz sollte dieser Wert in der Nähe von Eins liegen, während er für komprimierbare Sequenzen kleiner als Eins sein sollte. Betrachtet man die in Tab. B.3 aufgeführten Ergebnisse

Datensatz	L_o [Byte]	L_k [Byte]	K_V
LD1, 1. MB	1048576	1048767	1.00018
LD1, regularisiert, 1. MB	1048576	1048773	1.00019
LF1, 1. MB	1048576	1016799	0.96952

Tabelle B.3: Kompressionsverhalten von 1 MB großen Zufallsdatensätzen bei Kompression mit `gzip -9`

der mit `gzip -9` durchgeführten „Kompressionstests“[32], fällt folgendes auf: Datensätze mit einer Einswahrscheinlichkeit, die nahe beim Ideal 1/2 liegen, wie *LD1* (mit $P(1) \approx 0{,}515$) oder die 1 MB große Stichprobe des Laufes *LD1* nach Von-Neumann-Regularisierung, lassen sich nicht komprimieren. Hingegen lassen sich die Daten des Laufes *LF1*, (mit $P(1) \approx 0{,}387$) geringfügig komprimieren. Andererseits zeigen sich beim Lauf *LD1* bereits bei zweistufigen Autokorrelationskoeffizententests Antikorrelationen, s. u. Abschnitt B.9.3 und dementsprechend nach Von-Neumann-Regularisierung Korrelationen, s. Anhang, Abschnitt G.1.2, während sich bei Lauf *LF1* erst durch dreistufige Tests geringe Spuren von Antikorrelationen nachweisen lassen. Mit solch einfachen „Kompressionstests“ lassen sich also tiefer liegende Defekte nicht erkennen.
Ein massiver Nachteil solch einfacher Ad-hoc-Tests besteht zudem darin, daß die genaue Verteilung des Kompressionsverhältnisses für endliche, zufällige Sequenzen für das jeweilige Kompressionsverfahren i. a. nicht bekannt ist und somit ein exakter, statistischer Test wie bei den anderen Testverfahren nicht möglich[33] ist. Überdies würde sie auch noch von verschiedenen Parametern

[30] Dementsprechend werden sie sowohl was Theorie als auch Testergebnisse angeht, aus Platzgründen in dieser Arbeit nicht näher besprochen.

[31] Wie bei der linearen Komplexität kann auch bei der Maximal-Ordungs-Komplexität ein Profil für die Testsequenz erstellt werden.

[32] Bei der Kompression mit `gzip` wird eine Variation des LZ77 (Sliding Window)-Algorithmus [135] kombiniert mit statischer Huffmann-Kodierung verwendet.

[33] Lediglich für die Tests von Maurer bzw. Coron, die ein von Kompressionsverfahren abgeleitetes Komplexitätsmaß verwenden, ist dies bisher gelungen, s. Abschnitt 3.2.7.

des Kompressionsverfahrens, wie z.B. der Fenstergröße[34] beim `gzip`-Kompressionsverfahren, abhängen.
Besser als solch ein Ad-hoc-Vorgehen ist daher eine Berechnung der Sequenzkomplexität, da sie präzise definiert [80, 134] ist und nicht von zusätzlichen, wählbaren Parametern abhängt.
Eine positive Eigenschaft der Sequenzkomplexität ist ihr Verhalten im Grenzwert unendlich langer Bitsequenzen, die von einer stationären Quelle erzeugt wurden; in diesem Fall läßt sich zeigen [134], daß ihr Erwartungswert gegen die Entropie der Quelle geht.
Bedauerlicherweise sind keine solch exakten Aussagen für den Fall endlicher Bitsequenzen bekannt [81], so daß die Durchführung exakter statistischer Tests auch bei der Sequenzkomplexität problematisch ist.
Zudem würde beim Test von Rohdaten eines physikalischen Zufallsgenerators wieder das bereits in Abschnitt 3.2 erwähnte Problem auftreten, daß die Abweichung vom idealen Teilungsverhältnis alle anderen Defekte kaschiert, wie dies bereits beim einfachen Kompressionstest mit `gzip -9` zu sehen war. Für einen aussagekräftigen Test von Rohdaten müßte also nicht nur die genaue Verteilung der Teststatistik bekannt sein, sondern die Verteilung sollte sich auch mit der Wahrscheinlichkeit für eine Eins parametrisieren lassen.
Als ein Komplexitätsmaß in Bezug auf einen bestimmten Maschinentyp hat die Sequenzkomplexität zudem noch einige unintuitive Eigenschaften:

- Die Sequenzkomplexität einer Bitsequenz kann sich stark unterscheiden, abhängig davon, ob man sie in der Originalreihenfolge oder in umgekehrter Reihenfolge durchläuft [46].
- Es gibt Sequenzen, die eine hohe Sequenzkomplexität haben, obwohl sie keineswegs zufällig sind [134]. Dies gilt z.B. für eine Sequenz, welche aus der Aneinanderreihung der natürlichen Zahlen in ihrem Binärformat besteht, d.h. `11011100101110111....`

Aus diesen Gründen wird auch davon abgesehen, Tests mit der Sequenzkomplexität durchzuführen.

B.9 Autokorrelationstests

Die Autokorrelationsfunktionstests verwenden wie die vorangehenden χ^2- bzw. Kontingenztests eine Stichprobengröße von 8 kB für die Tests der ersten Stufe. Bei der Berechnung der Autokorrelationskoeffizienten werden immer nur die Bits der Stichprobe in ihrer gegebenen Reihenfolge[35] berücksichtigt, dementsprechend treten mit zunehmender Verschiebung stärker werdende Randeffekte auf. Bei der verwendeten maximalen Verschiebung von $l = 128$ spielen diese Randeffekte aber angesichts einer Sequenzlänge von 65.536 Bits keine Rolle, zumal die Autokorrelationskoeffizienten bei kleinen Verschiebungen (typ. $l = 1$) im Zentrum der Untersuchung stehen.

B.9.1 Einfache Autokorrelationsfunktionstests

Die bereits in Abschnitt 5.1.2.5 zur Aufdeckung der anfänglichen Probleme mit dem FIFO verwendeten einfachen Autokorrelationsfunktionstests erlauben einen ersten Test auf Antikorrelationen in einer Sequenz. Allerdings können mit ihnen nur verhältnismäßig starke Abweichungen vom ideal zufälligen, korrelationsfreien Verhalten aufgedeckt werden. Dies sei zur Illustration an zwei Autokorrelationsfunktionstests aufeinander folgender 8 kB langen Teilsequenzen vom Anfang einer antikorrelationsbehafteten Sequenz[36] verdeutlicht, s. Abb. B.19 und Abb. B.20.
Auch wenn einige der Autokorrelationskoeffizienten betragsmäßig große Werte aufweisen, so ist dies doch nicht gleich ein Zeichen für einen Defekt, da solche Werte durchaus auch bei idealen

[34] In der Tat wird eine Bitsequenz, die aus zwei identischen Hälften besteht, die erheblich größer sind als das entsprechende Fenster, nicht merklich besser komprimiert als eine der Hälften.
[35] Es wird also keine zyklische Verschiebung verwendet.
[36] Wieder die bereits bei den χ^2-Tests verwendete Sequenz *LD1*.

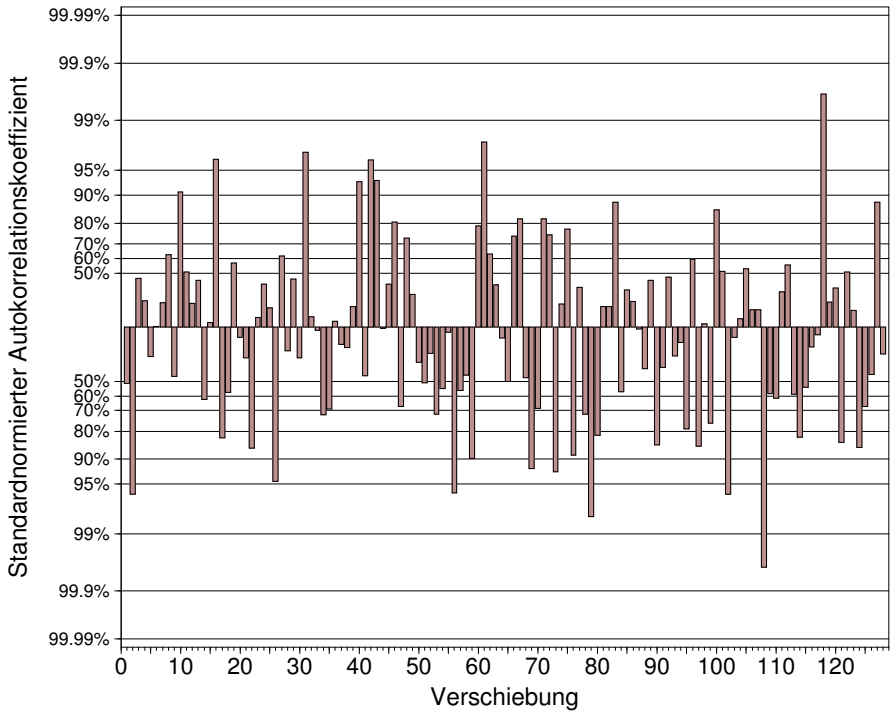


Abbildung B.19: Standardnormierte, binäre Autokorrelationsfunktion einer antikorrelationsbehafteten Sequenz des Laufes *LD1*, 8 kB Stichprobe

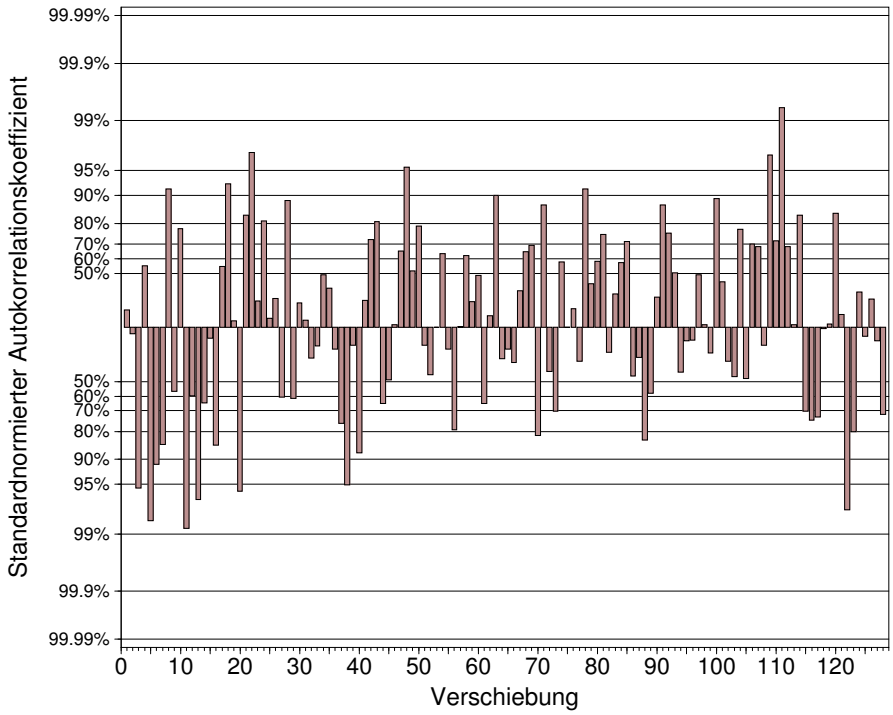


Abbildung B.20: Standardnormierte, binäre Autokorrelationsfunktion einer antikorrelationsbehafteten Sequenz, direkt auf die erste Stichprobe folgend

Zufallsgeneratoren vorkommen können. So findet man betragsmäßig große Autokorrelationskoeffizienten durchaus auch bei kryptographisch starken Pseudozufallsgeneratoren s. Abb. B.21 und natürlich auch bei kryptographisch schwachen, s. Abb. B.22.

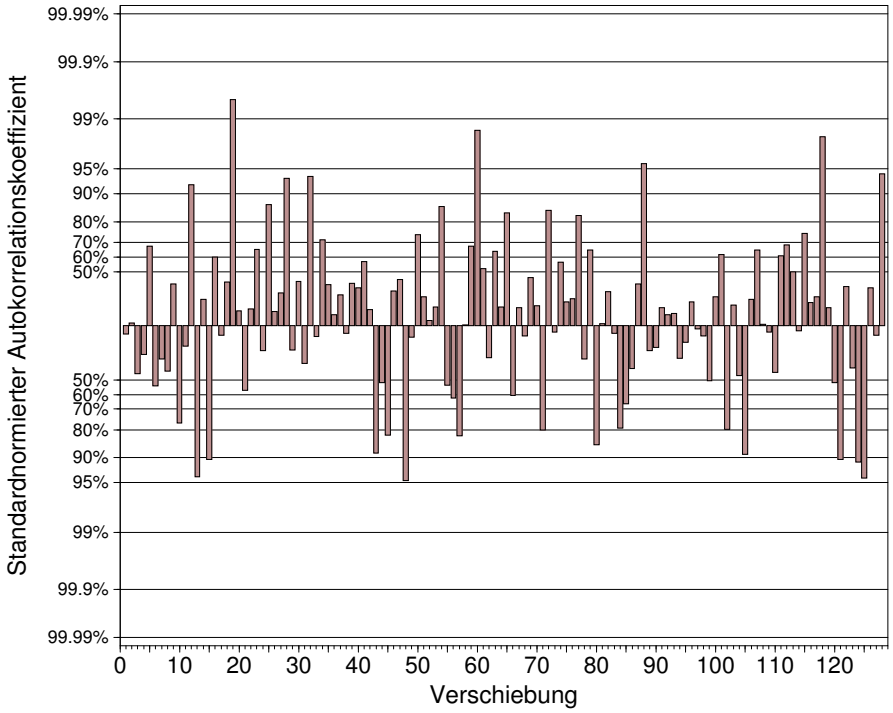


Abbildung B.21: Standardnormierte, binäre Autokorrelationsfunktion einer Sequenz, die von einem DES-Countermodus-Pseudozufallsgenerator erzeugt wurde

Eindeutige, sofort erkennbare Abweichungen vom zufälligen Verhalten lassen sich mit diesem Test nur bei sehr schlechten Zufallsgeneratoren erkennen, wie z.B. bei Pseudozufallsgenerator `RANDU`. Man erkennt also, für eine weitergehende Untersuchungen sind unbedingt zweistufige Autokorrelationskoeffizententests notwendig.

B.9.2 Einfache Autokorrelationstests zur Analyse von Spitzen im Verlauf der Einswahrscheinlichkeit des Laufes *LH5*

Einfache Autokorrelationstests sind allerdings durchaus gut geeignet, um stärkere Defekte zu analysieren. So zeigen sich bei Lauf *LH5* zu Anfang des insgesamt 53 MB umfassenden Laufes massive Störungen, die sich deutlich in einer schlagartig veränderten Wahrscheinlichkeit für eine Eins bemerkbar machen, s. Abb. B.23, in welcher der Verlauf der auf 8 kB großen Blöcken geschätzten Wahrscheinlichkeit für eine Eins aufgetragen ist.
Bei der ersten Abweichung (Blockindices 272 bis 317) ist offensichtlich, daß es sich keinesfalls um eine zufällige Fluktuation handeln kann, da es sich nicht nur um eine sehr starke Abweichung vom Mittelwert handelt, sondern auch längere Zeit *exakt* gleiche Wahrscheinlichkeiten auftreten. Die Konstanz des Wahrscheinlichkeitswertes schließt auch eine elektrische (Kontakt-)Störung aus, wie sie bei Lauf *PB2* auftrat, s. Abschnitt 5.2.2.
Bei der zweiten Abweichung, die sich lediglich in einem Block, nämlich bei Blockindex 333, massiv bemerkbar macht, ist der Typ der Störung nicht offensichtlich, eine Autokorrelationsanalyse[37]

[37] Aufgrund der sehr großen Absolutwerte für die Autokorrelationskoeffizienten werden in Abb. B.24 keine Prozentintervalle für die Ordinatenbeschriftung verwendet, sondern direkt die standardnormierten

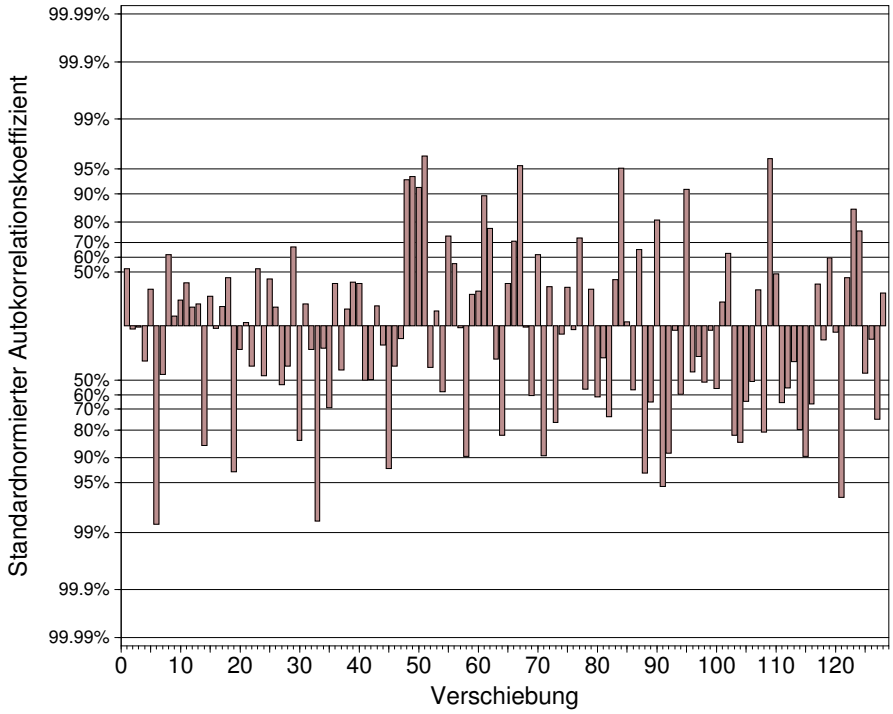


Abbildung B.22: Standardnormierte, binäre Autokorrelationsfunktion einer mit dem linearen Kongruenzgenerator `ran0` erzeugten Sequenz

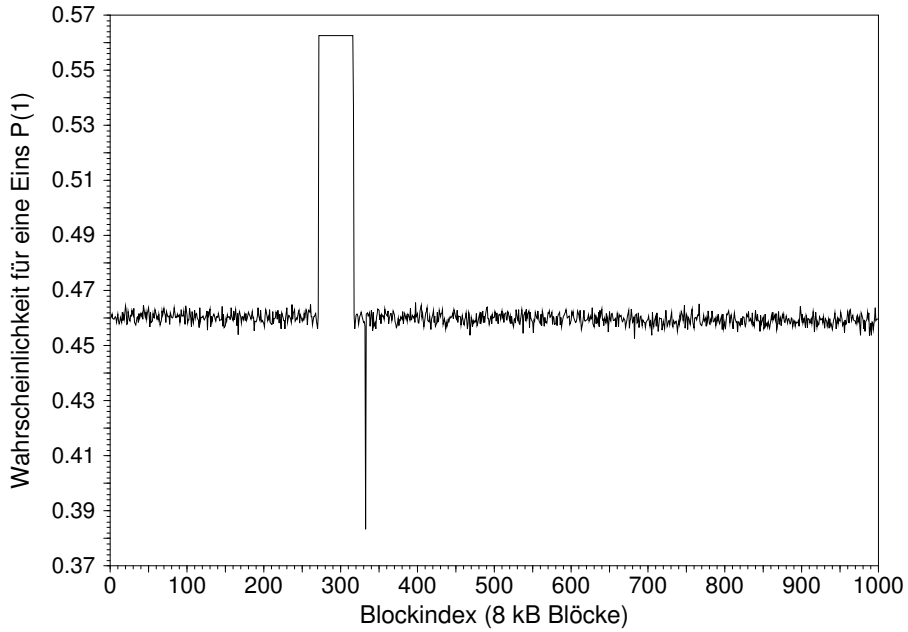


Abbildung B.23: Ausschnitt aus dem Verlauf der geschätzten Wahrscheinlichkeit p_{Block} für eine Eins bei einer Blockgröße von $N = 8$ kB für den Lauf *LH5*

des betreffenden Blockes zeigt aber sofort, daß es sich keinesfalls um eine elektrische Störung oder gar eine zufällige Fluktuation handeln kann, s. Abb. B.24.

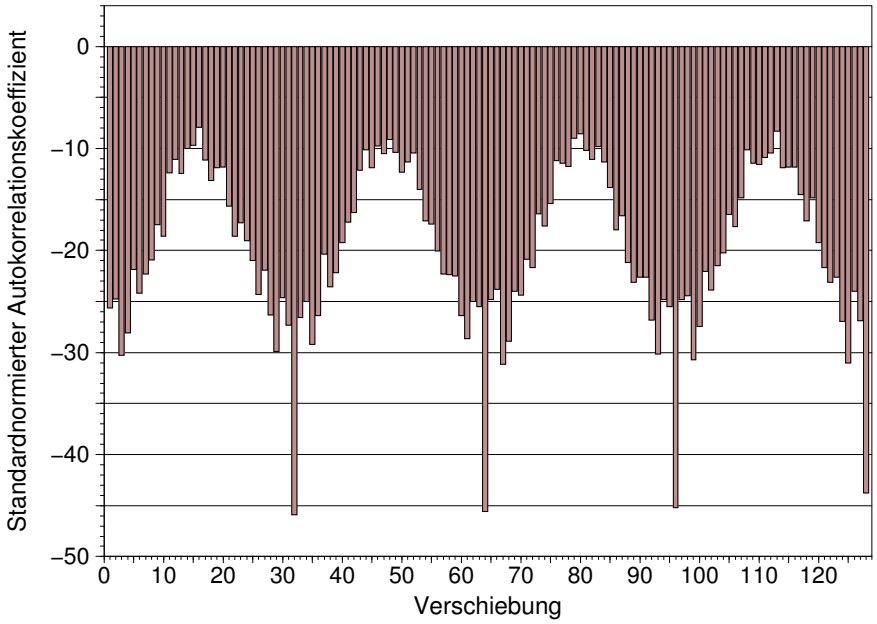


Abbildung B.24: Standardnormierte, binäre Autokorrelationsfunktion des 8 kB großen Blockes bei dem die zweite Abweichung auftritt

Die stark negativen Werte der Autokorrelationsfunktion deuten darauf hin, daß lange Folgen von Bits widerholt werden; auffällig sind zudem die Minima der Autokorrelationsfunktion bei Vielfachen von 32.
Eine Inspektion der entsprechenden Bitblöcke mit einem Hex-Editor zeigt dann auch sehr schnell, daß beide Abweichungen offensichtlich von einem selten auftretenden Fehler in der Datenaufnahme-Elektronik/Software nach dem programmierbaren Interrupt-Controller herrühren. Bei der ersten Abweichung wird nämlich, beginnend bei Adresse: `001E000`, der zwei Byte[38] lange Hexadezimalwert `23cf` für die nächsten 366 kB wiederholt; besonders verdächtig ist hierbei, daß der Anfang der Störung innerhalb der Datei genau auf einer 8 kB-Grenze liegt.
Die zweite 2048 Byte lange Abweichung fängt ebenfalls an einer 8 KB Grenze an; das Problem beginnt bei Adresse:`0098000` und zerfällt in zwei Blöcke mit unterschiedlicher Struktur: Im ersten 1336 Bytes langen Block wird ein „Zähler“, beginnend mit dem Hexadezimalwert `b1aa 0000` bis zum Wert `b2f7 0000` hochgezählt; anschließend folgen in einem zweiten Block nur noch Nullen bis zum Ende der Abweichung.
Betrachtet man die früheren Probleme mit dem FIFO, s. Abschnitt 5.1.2.5, scheint insbesondere die erste Abweichung ähnliche Ursachen zu haben. Bei der zweiten Abweichung ist dies weniger klar, hier könnte auch ein nur selten auftretender Fehler in der Datenaufnahmeroutine des PCs die Ursache sein. Abweichungen, wie die oben dokumentierten, kamen allerdings nur in diesem einen Lauf vor.

Zahlenwerte; zur Orientierung: das 99.99%-Intervall geht von −3,895 bis +3,895.

[38] Dies ist wiederum der Grund, warum es sich nicht um ein Problem des PICs handeln kann, das PIC-Programm kann nämlich nur ein Byte abspeichern.

B.9.3 Zweistufige Autokorrelationskoeffizententests

Für eine weitergehende Analyse des in Abbschnitt B.9.1 mit einstufigen Tests untersuchten Laufes *LD1*[39] empfehlen sich zweistufige Autokorrelationskoeffizententests[40].
Bei physikalischen Zufallsgeneratoren ist natürlich insbesondere die Verschiebung um ein Bit interessant; in Abb. B.25 sind daher die empirischen Verteilungsfunktionen für den ersten Autokorrelationskoeffizienten von jeweils fünf zweistufigen Tests und die Normalverteilung, der sie sich annähern sollten, abgebildet. In Tabelle B.4 sind die dazugehörigen quantitativen Ergebnisse[41] der Kolmogorov-Smirnov-Tests aufgeführt.

Test Nr.	K^+_{128}	$P(K^+_{128})$	K^-_{128}	$P(K^-_{128})$
1	0.053316	0.00879	1.271988	0.96362
2	0.001406	0.00009	1.742447	0.99793
3	0.107308	0.02895	1.499345	0.98983
4	0.020187	0.00200	1.830187	0.99890
5	0.260495	0.14031	1.981805	0.99966

Tabelle B.4: Kolmogorov-Smirnov-Auswertung der empirischen Verteilungsfunktionen von fünf zweistufigen Tests zu Beginn des Laufes *LD1* für den ersten Autokorrelationskoeffizienten

Die Ergebnisse des Kolmogorov-Smirnov-Tests bestätigen den ersten Eindruck, den man beim Betrachten der Graphen hat: Die Wahrscheinlichkeitswerte, daß die jeweiligen maximalen Abweichungen in positiver oder negativer Richtung kleiner sind, liegen bei den fünf Tests sehr nahe bei ihren Extremwerten. Offensichtlich liegt bei diesem Lauf ein statistischer Defekt des Generators vor, der vom Autokorrelationskoeffizententest – anders als bei allen vorangehenden Tests – auch eindeutig erkannt wird.
Die empirischen Verteilungsfunktionen der standardnormierten Variable $X(l = 1)$ sind nach rechts zu größeren Werten hin verschoben. Das bedeutet, daß die Summe über die einzelnen XOR-Terme $b_i \oplus b_{i+1}$ größer ist, als erwartet. Größer kann die Summe aber nur werden, wenn die Anzahl der Summanden, die eine Eins liefern, größer ist als die Anzahl der Summanden, die eine Null liefern. Einen Wahrheitswert Eins liefern die XOR-Terme aber genau dann, wenn sie ungleichwertige Bits (also $0 \oplus 1$ oder $1 \oplus 0$) enthalten. Man sieht also, daß es bei diesem Lauf eine Tendenz zu Antikorrelationen zwischen direkt aufeinanderfolgenden Bits gibt.
Um zu überprüfen, ob die Antikorrelationen weiter reichen, wird auch noch für den zweiten Autokorrelationskoeffizienten ein Verteilungstest durchgeführt; in Abb. B.26 sind die empirischen Verteilungsfunktionen des zweiten Autokorrelationskoeffizienten für fünf Tests[42] dargestellt, die Ergebnisse der zugehörigen Kolmogorov-Smirnov-Tests sind in Tab. B.5 aufgeführt:
Man kann in Abb. B.26 deutlich erkennen, daß die empirischen Verteilungsfunktionen des zweiten Autokorrelationskoeffizienten wesentlich näher an der Normalverteilung liegen, als dies bei den entsprechenden Verteilungsfunktionen des ersten Autokorrelationskoeffizienten der Fall ist.

[39] Bei ihm diente ein freistrahloptischer Aufbau mit Einphotonenquelle auf Basis der parametrischen Fluoreszenz und mit kompakter Signalverarbeitungs- und Datenaufnahme-Elektronik zur Zufallsgenerierung.

[40] Wie bereits in Abschnitt 3.2.6 erwähnt, werden in der ersten Stufe 128 Tests jeweils auf einer 8 kB großen Stichprobe durchgeführt und aus den 128 Autokorrelationskoeffizienten für *eine* feste Verschiebung eine empirische Verteilungsfunktion ermittelt und diese mit der theoretischen Verteilungsfunktion verglichen.

[41] Die Wahrscheinlichkeitswerte geben an, wie wahrscheinlich es ist, daß ein Wert für die maximale Abweichung in positiver oder negativer Richtung kleiner als der empirisch ermittelte Wert ist.

[42] Es werden die gleichen Stichproben wie beim Test des ersten Autokorrelationskoeffizienten verwendet.

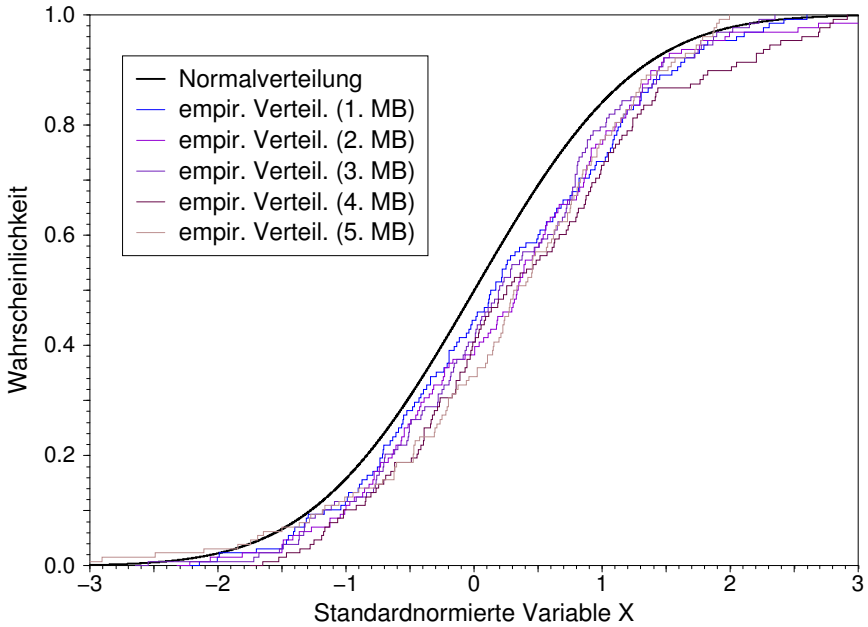


Abbildung B.25: Empirische Verteilungsfunktionen des ersten Autokorrelationskoeffizienten von fünf zweistufigen Tests zu Beginn des Laufes *LD1*

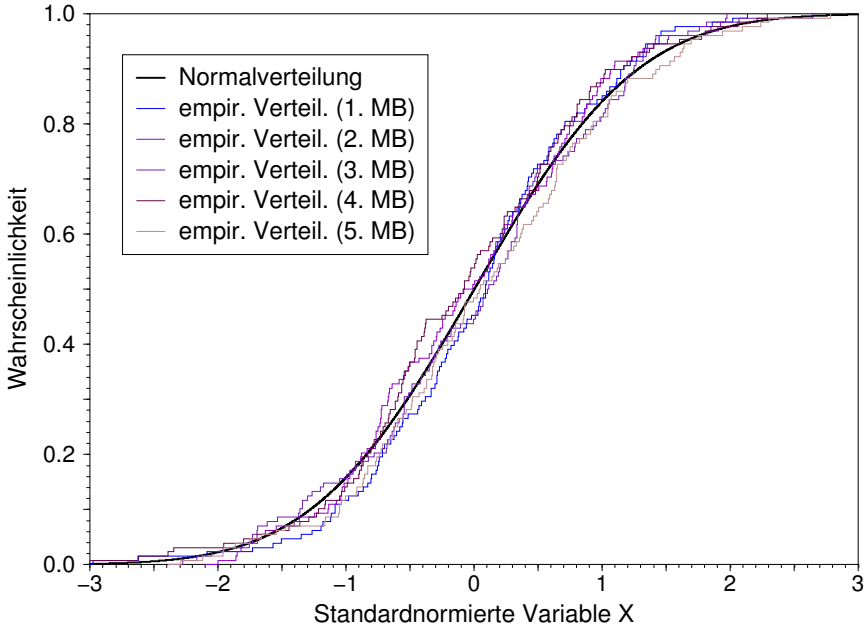


Abbildung B.26: Empirische Verteilungsfunktionen des zweiten Autokorrelationskoeffizienten von fünf zweistufigen Tests zu Beginn des Laufes *LD1*

Test Nr.	K_{128}^{+}	$P(K_{128}^{+})$	K_{128}^{-}	$P(K_{128}^{-})$
1	0.508084	0.42114	0.705120	0.64542
2	0.767022	0.70562	0.373716	0.26036
3	0.382520	0.27053	0.678761	0.61797
4	1.036892	0.89066	0.328706	0.20995
5	0.091330	0.02184	0.699570	0.63973

Tabelle B.5: Kolmogorov-Smirnov-Auswertung der empirischen Verteilungsfunktionen von fünf zweistufigen Tests zu Beginn des Laufes *LD1* für den zweiten Autokorrelationskoeffizienten

Dementsprechend findet sich in den Ergebnissen der Kolmogorov-Smirnov-Tests lediglich in einem Fall (K_{128}^{+} und die zugehörige Wahrscheinlichkeit $P(K_{128}^{+})$ in Test Nr. 5) ein bedenklicher Wert; solche Werte können aber vereinzelt auch bei idealen Sequenzen auftreten.
Da es nicht sinnvoll wäre, jetzt für alle weiteren Koeffizienten eine graphische Darstellung der empirischen Verteilungsfunktion und eine tabellarische Auflistung der Ergebnisse der zugehörigen Kolmogorov-Smirnov-Tests zu präsentieren, sind in den beiden Abbildungen B.27 und B.28 lediglich die Werte für $P(K_{128}^{+})$ und $P(K_{128}^{-})$ graphisch in Abhängigkeit von der Verschiebung für die ersten zehn zweistufigen Tests dargestellt.

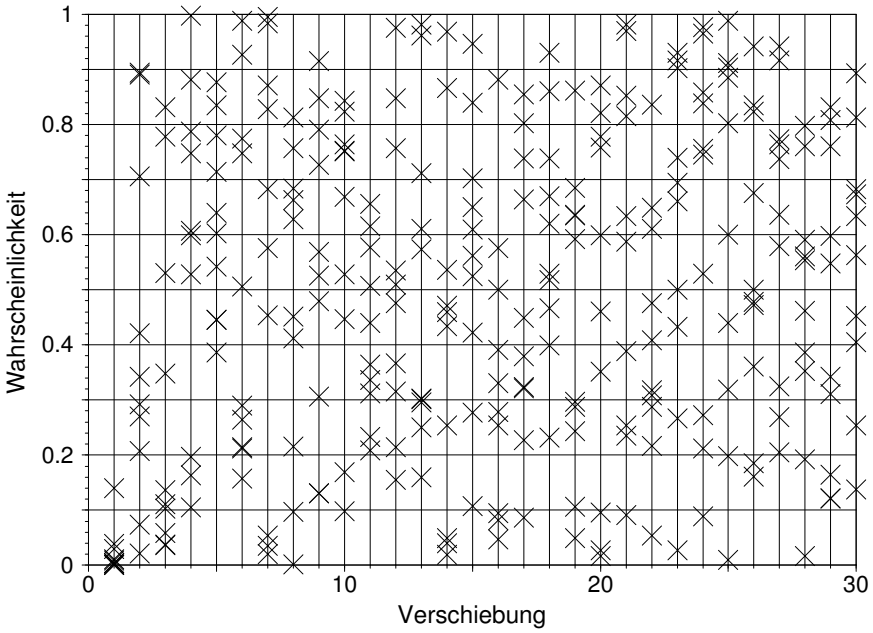


Abbildung B.27: $P(K_{128}^{+})$-Werte der ersten 30 Verschiebungswerte für 10 aufeinander folgende Tests des Laufes *LD1*

Man erkennt in beiden Abbildungen deutlich, daß die Verteilung der Wahrscheinlichkeitswerte der Kolmogorov-Smirnov-Tests bei den größeren Verschiebungen sehr viel ausgeglichener ist als bei der Verschiebung um ein Bit. Eine weitergehende Interpretation der Verteilungen verbietet sich allerdings, da die Autokorrelationskoeffizienten einer Stichprobe nicht unabhängig voneinander sind und das menschliche Auge überdies auch gern Strukturen erkennt, wo gar keine sind.

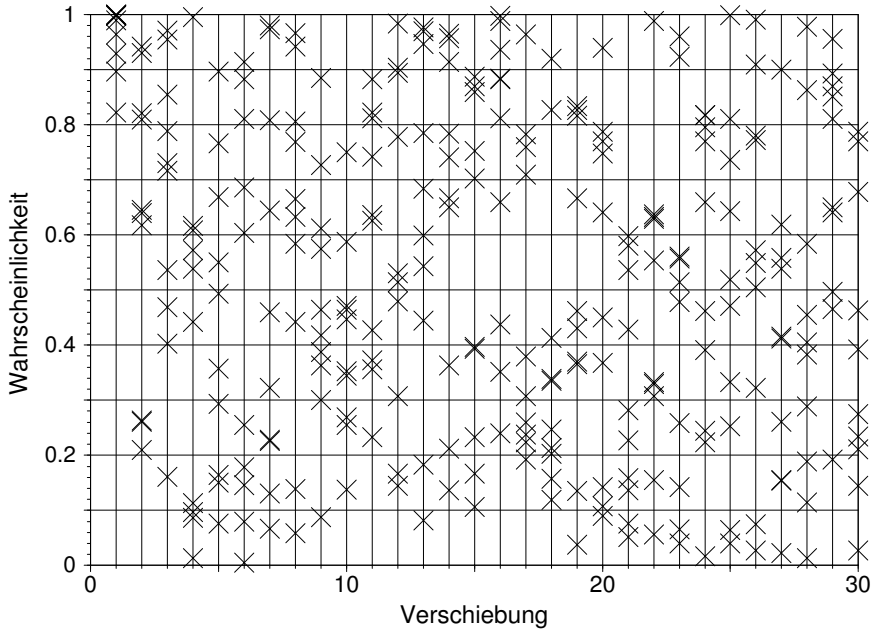


Abbildung B.28: $P(K^-_{128})$-Werte der ersten 30 Verschiebungswerte (neue Datenaufnahme-Elektronik) für 10 aufeinanderfolgende Tests des Laufes *LD1*

B.10 Analyse der teilweise auftretenden Antikorrelationen

Im vorangegangenen Abschnitt wurde beschrieben, wie sich mit Hilfe des Autokorrelationskoeffizententests Korrelationen in Zufallssequenzen der Generatoren feststellen lassen; und tatsächlich wurden auch Antikorrelationen in Läufen (z.B. *LD1*) festgestellt.
Vergleicht man zur Kontrolle den Lauf *LD1* mit einem Lauf (*LN3*) des Generators, der sich von diesem lediglich durch die Verwendung der vorher verwendeten auf NIM-Einschüben basierenden Signalverarbeitungs- und Datenaufnahme-Elektronik unterscheidet, so stellt man fest, daß sich durch den Autokorrelationskoeffizententest bei diesem keine Antikorrelationen feststellen lassen, s. Abb. B.29 bzw. Tab. B.6 für die quantitativen Ergebnisse des Kolmogorov-Smirnov-Tests.

Test Nr.	K^+_{128}	$P(K^+_{128})$	K^-_{128}	$P(K^-_{128})$
1	0.581807	0.50928	0.421826	0.31685
2	0.385909	0.27447	0.733661	0.67395
3	0.654662	0.59201	0.803775	0.73832
4	1.280816	0.96525	0.302971	0.18258
5	0.849389	0.77559	0.252799	0.13309

Tabelle B.6: Kolmogorov-Smirnov-Auswertung der empirischen Verteilungsfunktionen von fünf Tests zu Beginn des Laufes *LN3* für den ersten Autokorrelationskoeffizienten

Daher liegt der Verdacht nahe, daß die Antikorrelationen entweder von Problemen innerhalb der kompakten Signalverarbeitungs- und Datenaufnahme-Elektronik herrühren, oder anderen Ursprungs (z.B. Auswirkungen von Detektor-Totzeiten) sind und dies von der vorher verwendeten Datenaufnahme-Elektronik lediglich kaschiert wurde[43].

[43] In der Tat ist es bei der alten Datenaufnahme-Elektronik notwendig, die Signale aus den Koinzi-

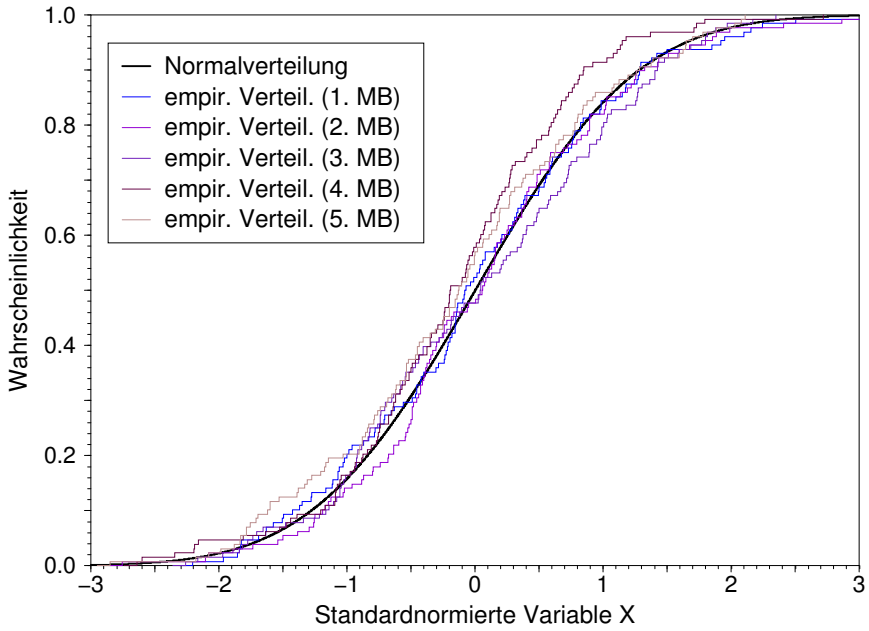


Abbildung B.29: Empirische Verteilungsfunktionen des ersten Autokorrelationskoeffizienten für einen Lauf (*LN3*) mit der „alten“ Datenaufnahme-Elektronik

Zur Einkreisung der Ursache der Antikorrelationen werden noch weitere Messungen durchgeführt:

1. Läufe mit reduzierter Laserpumpleistung und
2. Läufe mit künstlich eingefügter Zeitverzögerung bei der Registrierung der Bits durch die kompakte Datenaufnahme-Elektronik.

B.10.1 Läufe mit reduzierter Laserpumpleistung

Bei diesen Messungen wird die Leistung der Pumplasers mit Hilfe von Neutralglasfiltern gesenkt, was sich in einer entsprechend niedrigeren Rate der Paarerzeugung und damit auch in einem größeren mittleren Abstand der Photonenpaare zueinander auswirkt. Bei niedrigerer Pumpleistung ist also zu erwarten, daß die Abweichungen von der theoretischen Verteilung geringer ausfallen, da nun die Anzahl der kurz hintereinander folgenden Paare niedriger ist und sich daher indirekte Einflüsse des vorhergehenden Bits auf das nachfolgende weniger bemerkbar machen.
In Abb. B.30 sind für den Anfang eines Laufes (*LN5*) mit halber und in Abb. B.31 eines Laufes (*LN6*) mit einem Viertel der ursprünglichen Laserpumpleistung die ersten fünf empirischen Verteilungsfunktionen für den ersten Autokorrelationskoeffizienten aufgetragen.
In Tabelle B.7 sind die zugehörigen Ergebnisse der Kolmogorov-Smirnov-Tests für den Lauf *LN5* mit halber Pumpleistung aufgeführt und in Tabelle B.8 die Ergebnisse für den Lauf *LN6* mit einem Viertel der Laserleistung.

denzeinheiten auf ca. 100 μs zu verlängern. Dies führt dazu, daß ein bestimmter Anteil von „Fehlern“ (ca. 12%) auftritt, d.h. Ereignisse, bei denen auf beiden Signalleitungen – der für Einsen und der für Nullen – ein HIGH-Pegel vorliegt. Diese Fehler werden nicht als Bits registriert. Aufgrund der starken Verlängerung der Pulse (die Ausgangspulse der Koinzidenzeinheiten sind lediglich ca. 10 ns lang) werden also alle Ereignisse, die innerhalb eines Zeitintervalls von 100 μs aufeinanderfolgen – oft fälschlicherweise – als Fehler verworfen.

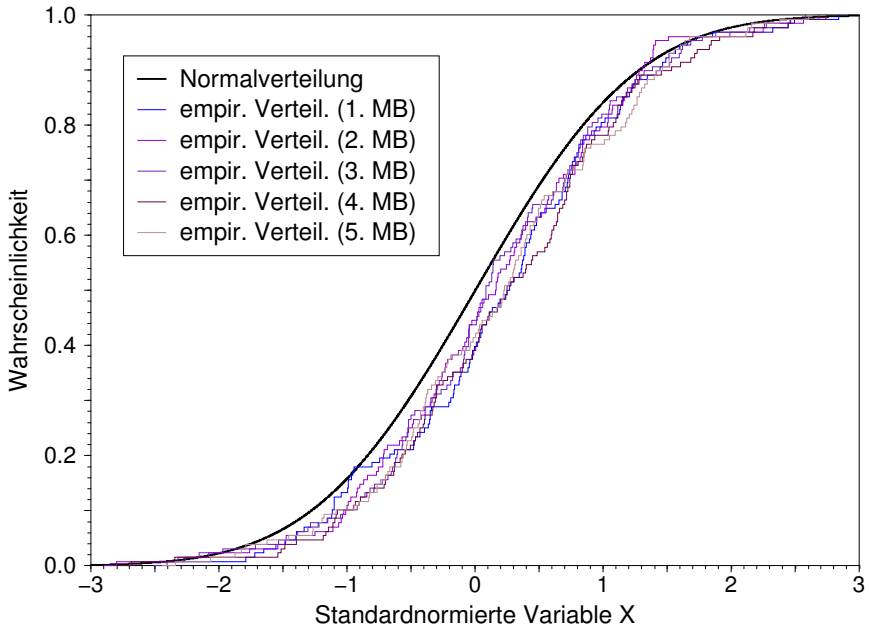


Abbildung B.30: Empirische Verteilungsfunktionen des ersten Autokorrelationskoeffizienten mit neuer Datenaufnahme-Elektronik bei auf die Hälfte reduzierter Laserleistung, Lauf *LN5*

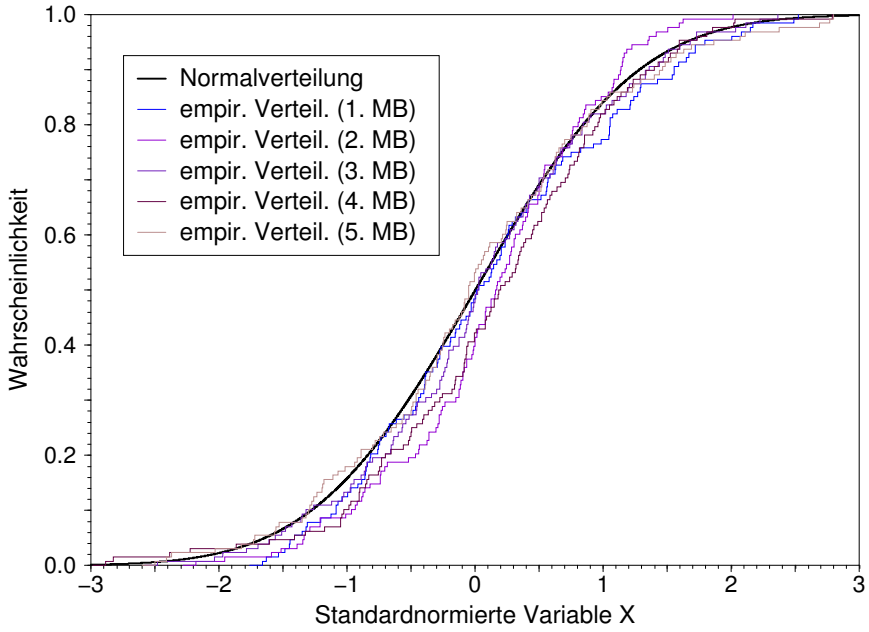


Abbildung B.31: Empirische Verteilungsfunktionen des ersten Autokorrelationskoeffizienten mit neuer Datenaufnahme-Elektronik bei auf ein Viertel reduzierter Laserleistung, Lauf *LN6*

Wie man sieht, nehmen die Abweichungen der empirischen Verteilungsfunktionen von der theoretischen in der Tat bei verminderter Rate der Paare ab. Die Ergebnisse der Kolmogorov-Smirnov-Tests tendieren aber auch bei reduzierter Paarrate zu den Extremwerten hin, so daß sich immer noch Antikorrelationen nachweisen lassen.

Test Nr.	K_{128}^{+}	$P(K_{128}^{+})$	K_{128}^{-}	$P(K_{128}^{-})$
1	0.095709	0.02369	1.453369	0.98662
2	0.371445	0.25775	0.839986	0.76821
3	0.089141	0.02094	1.166954	0.93887
4	0.069645	0.01372	1.543571	0.99225
5	0.090586	0.02153	1.252841	0.95988

Tabelle B.7: Ergebnisse der Kolmogorov-Smirnov-Tests für 5 Stichproben des Laufes *LN5* mit halber Pumplaserleistung)

Test Nr.	K_{128}^{+}	$P(K_{128}^{+})$	K_{128}^{-}	$P(K_{128}^{-})$
1	0.157456	0.05720	0.878029	0.79709
2	0.634679	0.56991	1.630836	0.99557
3	0.200871	0.08845	0.798527	0.73379
4	0.207805	0.09397	1.337387	0.97425
5	0.473880	0.37963	0.411123	0.30411

Tabelle B.8: Ergebnisse der Kolmogorov-Smirnov-Tests für 5 Stichproben des Laufes *LN6* mit einem Viertel der üblichen Pumplaserleistung)

B.10.2 Läufe mit künstlich eingefügter Zeitverzögerung

Will man die Zeitskala bestimmen, auf der sich die Antikorrelationen auswirken, empfiehlt es sich, die Datenaufnahme-Routine im PIC zu modifizieren. Die Modifikationen bestehen darin, zu der minimal notwendigen Laufzeit von 2 μs, welche die Interrupt-Routine im PIC benötigt, um den anliegenden Bitwert einzulesen, noch eine zusätzliche, programmtechnisch realisierte Verzögerung hinzuzufügen. Da während dieser verlängerten Interrupt-Bearbeitung keine weiteren Bitwerte aufgenommen werden, läßt sich durch die Erhöhung der Verzögerung von Lauf zu Lauf des Generators, die minimale Zeit zwischen zwei aufeinanderfolgenden Bits bei der Zufallsgenerierung ebenfalls schrittweise vergrößern.
Führt man anschließend für jeden der Meßläufe eine Reihe von zweistufigen Verteilungstests für den ersten Autokorrelationskoeffizienten durch, so läßt sich an den Auswertungsergebnissen der Kolmogorov-Smirnov-Tests feststellen, ob sich noch Antikorrelationen zeigen.
Insgesamt werden die Tests für 8 Läufe mit verschiedenen Verzögerungszeiten durchgeführt. Die entsprechenden Verzögerungszeiten[44] betragen entsprechend der Gesamtdauer der Interrupt-Routine 2, 4, 6, 10, 18, 32, 50 und 92 μs. Die generierte Datenmenge bzw. Dauer der Läufe, ist hierbei nicht immer gleich gewählt[45], so daß auch die Anzahl der erhaltenen Ergebnisse der Kolmogorov-Smirnov-Tests unterschiedlich ist.
In Abb. B.32 und B.33 sind die Ergebnisse zweistufiger Autokorrelationskoeffizententests[46] für den ersten Autokorrelationskoeffizienten dargestellt. Jedes einzelne Kreuz steht für einen Wahrscheinlichkeitswert $P(K_{128}^{-})$ bzw. $P(K_{128}^{+})$ eines Kolmogorov-Smirnov-Tests.

[44] Der Grund für die uneinheitlichen Abstände zwischen den verschiedenen Verzögerungszeiten erklärt sich daraus, daß bedingt durch die relativ lange Zeitdauer für einen Lauf versucht wurde, durch „Verdoppeln“ der Verzögerungszeit mit einer möglichst geringen Zahl von Läufen auszukommen.

[45] Bei Läufen mit geringer Verzögerung zeigten die Ergebnisse der Kolmogorov-Smirnov-Tests eine so starke Tendenz, daß die Läufe kürzer gehalten wurden, erst bei den längeren Verzögerungen wurden größere Datenmengen aufgenommen.

[46] Es werden dieselben Stichprobengrößen wie in den vorangehenden Abschnitten verwendet, d.h. ein zweistufiger Test umfaßt eine Gesamtstichprobengröße von 1 MB.

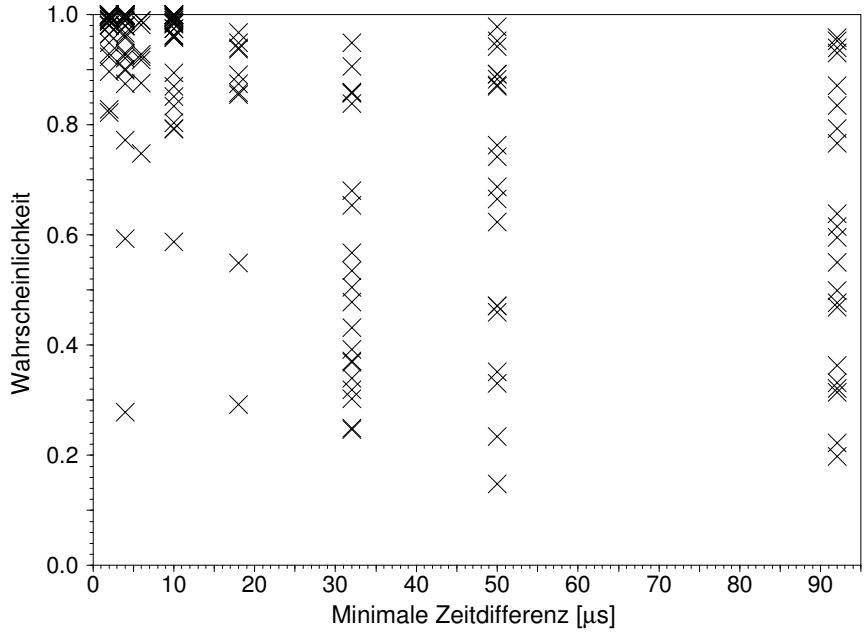


Abbildung B.32: $P(K_{128}^{-})$-Werte zweistufiger Autokorrelationskoeffiziententests des ersten Autokorrelationskoeffizienten für verschiedene Läufe mit verzögerter Bitübernahme

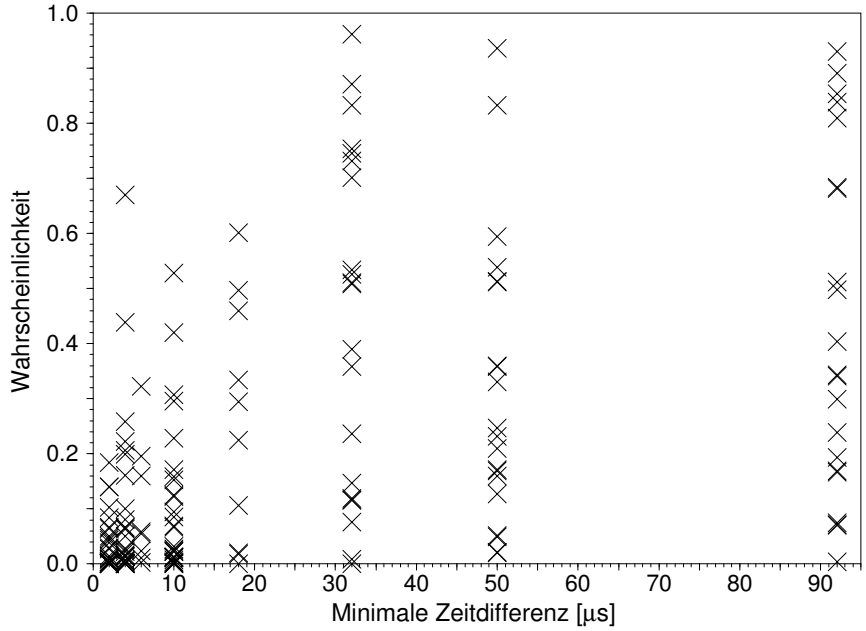


Abbildung B.33: $P(K_{128}^{+})$-Werte zweistufiger Autokorrelationskoeffiziententests des ersten Autokorrelationskoeffizienten für verschiedene Läufe mit verzögerter Bitübernahme

Man sieht insbesondere bei den Werten für $P(K_{128}^{-})$ in Abb. B.32 sehr gut, wie die Antikorrelationen mit zunehmender Verzögerung bei der Bitaufnahme abnehmen. Bei der längsten Verzögerung ist keine so offensichtliche Tendenz der Verteilung der Werte wie bei den kürzeren Verzögerungen zu erkennen, was auch der Grund dafür ist, daß keine Messungen mit noch längerer Verzögerung durchgeführt werden. Lediglich das Ausbleiben von niedrigen Wahrscheinlichkeitswerten für die

$P(K_{128}^-)$ fällt auf. Dies ist gleichbedeutend damit, daß es selten Autokorrelationskoeffizientenverteilungen für den ersten Autokorrelationskoeffizienten gibt, die sehr nah an der theoretischen Verteilung liegen. Betrachtet man die zugehörigen $P(K_{128}^+)$-Werte, so stellt man fest, daß bei diesen sowohl große als auch kleine Werte auftreten, was die Entscheidung, keine Läufe mit längeren Verzögerungen durchzuführen, zusätzlich unterstützt.

B.10.3 Dreistufige Autokorrelationstests

Die im vorangehenden Abschnitt in Abb. B.32 und B.33 dargestellten Verteilungen der $P(K_{128}^-)$ bzw. $P(K_{128}^+)$ sind zwar für kurze Verzögerungen deutlich tendenzbehaftet, aber bei den längeren Verzögerungen ist eine Beurteilung schwieriger. Es erscheint daher nachträglich sinnvoll, die Verteilung der Testwerte noch genauer zu untersuchen; dies geschieht durch eine weitere Teststufe.
Zwar ist die Anzahl der zweistufigen Tests bei der längsten Verzögerung (Lauf *LD8*) nicht allzu groß, aber aufgrund der guten Approximation auf der ersten und zweiten Teststufe ist es dennoch sinnvoll, aus den 21 Werten der K_{128}^- bzw. K_{128}^+ wiederum eine Verteilungsfunktion[47] zu erstellen und diese mit der theoretischen Kolmogorov-Smirnov-Verteilung, s. S. 45, für die Abweichungen K_{21}^- bzw. K_{21}^+ zu vergleichen.
Zur Kontrolle, inwieweit solch ein Test überhaupt aussagekräftig sein kann oder ob die kleinen Stichproben der dritten Stufe eventuell die Ergebnisse stark verfälschen, wird zuerst auf einer lediglich 10 MB umfassenden Stichprobe eines kryptographisch starken Pseudozufallsgenerators[48] ein dreistufiger Referenztest durchgeführt. Der graphischen Darstellung der Ergebnisse für die Verteilung der K_{10}^- in Abb. B.34 und der K_{10}^+ in Abb. B.35 entnimmt man, daß auch bei relativ kleinen Stichproben die Verteilung eines idealen Zufallsgenerators die theoretische Verteilung „umspielt“ und somit eine Durchführung dreistufiger Tests durchaus sinnvolle Ergebnisse liefert. Eine quantitative Auswertung erscheint allerdings aufgrund der doch relativ großen Sprünge bei kleinen Stichproben nicht sinnvoll, daher wird nur ein graphischer Vergleich der empirischen mit der theoretischen Verteilung durchgeführt, der aber dennoch sehr aufschlußreich sein kann.
Die entsprechenden Ergebnisse solch eines dreistufigen Tests sind für den Lauf *LD8* in den Abb. B.36 (Verteilung der K_{128}^-) und B.37 (Verteilung der K_{128}^+) dargestellt. Während die empirische Verteilungsfunktion der K_{128}^+ recht nah bei der theoretischen Verteilungsfunktion liegt, gilt dies für die Verteilungsfunktion der K_{128}^- weniger gut, so daß auch noch bei gegenüber der normalen Zeit für die Bitübernahme relativ langen Verzögerungen mit Hilfe eines dreistufigen Autokorrelationskoeffizententests – wenn auch sehr geringe – Spuren von Antikorrelationen zu entdecken sind.

B.10.4 Autokorrelationstests der Läufe mit hybrider Datenaufnahme-Elektronik

Wie in Abschnitt 4.5.3 erwähnt, lassen sich bei der hybriden Elektronik drei Parameter einstellen:die Länge des Übernahme-Signals, die Länge des Bit-Signals und die Zeitverzögerung zwischen Bit-Signal und Übernahme-Signal.
Beim ersten Lauf (*LH1*) wurde noch eine Zeitverzögerung von 100 ns zwischen den beiden Pulsen eingestellt, dies geschah um sicherzustellen, daß auf jeden Fall ein stabiles Bit-Signal am Eingang-Port des PIC anliegt, bevor die Interrupt-Bearbeitung startet. Bei den nachfolgenden Läufen wurde allerdings auf diese Verzögerung verzichtet, da angesichts einer minimalen Dauer der Interrupt-Bearbeitung von 400 ns auch ohne Zeitverzögerung ein stabiles Signal anliegt.

[47] Es wird also nicht eine Verteilung aus den Wahrscheinlichkeitswerten gebildet, sondern direkt aus den Werten für die maximale Abweichung der einzelnen Tests der zweiten Stufe, diese Verteilung sollte einer Kolmogorov-Smirnov-Verteilung folgen.
[48] Der Zufallsgenerator verwendet DES im OFB-Modus, s. Abschnitt B.1.

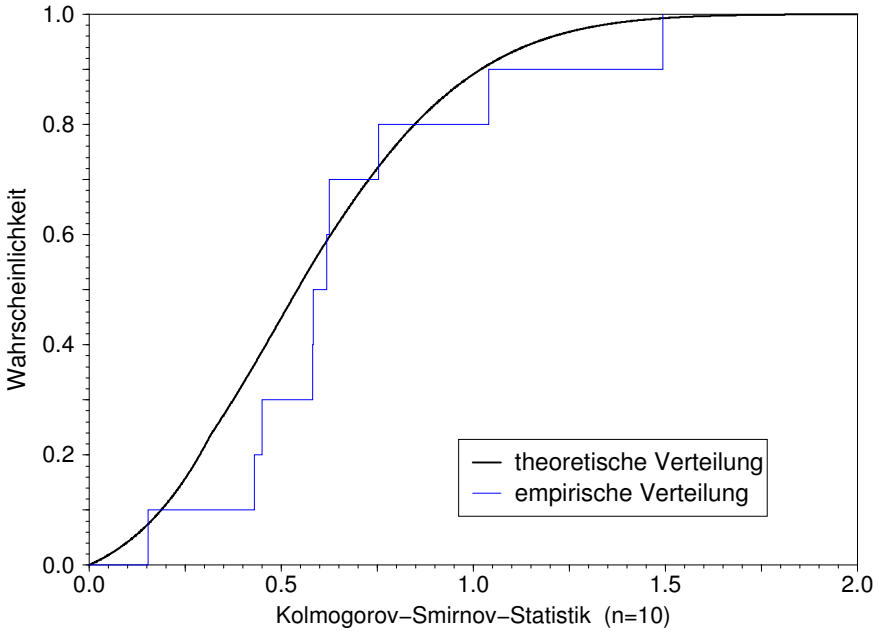


Abbildung B.34: Vergleich der empirischen Verteilungsfunktion aus den K_{128}^{-}-Werten von 10 zweistufigen Tests für den ersten Autokorrelationskoeffizienten mit der theoretischen Verteilungsfunktion für eine 10 MB umfassende Stichprobe eines Pseudozufallsgenerators

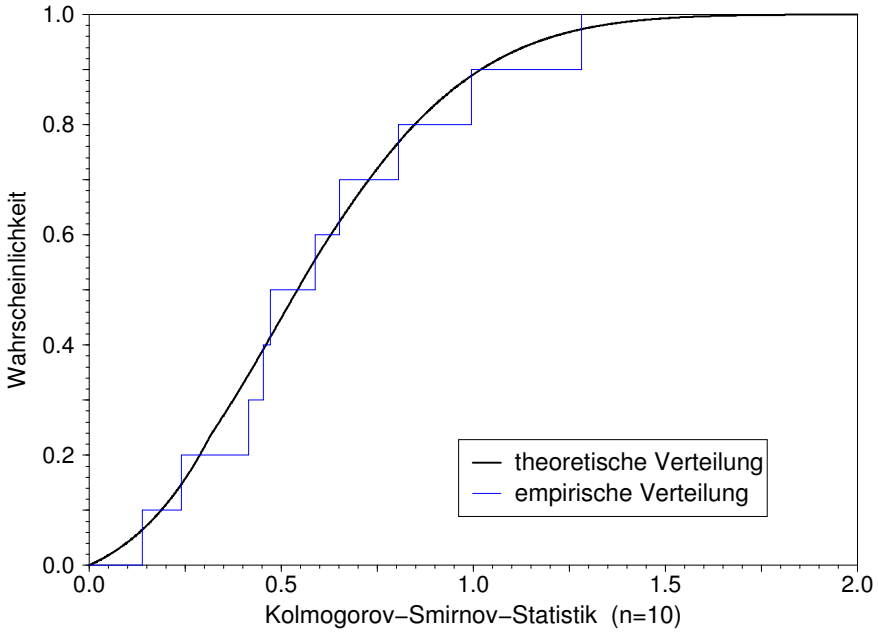


Abbildung B.35: Vergleich der empirischen Verteilungsfunktion aus den K_{128}^{+}-Werten von 10 zweistufigen Tests für den ersten Autokorrelationskoeffizienten mit der theoretischen Verteilungsfunktion für eine 10 MB umfassende Stichprobe eines Pseudozufallsgenerators

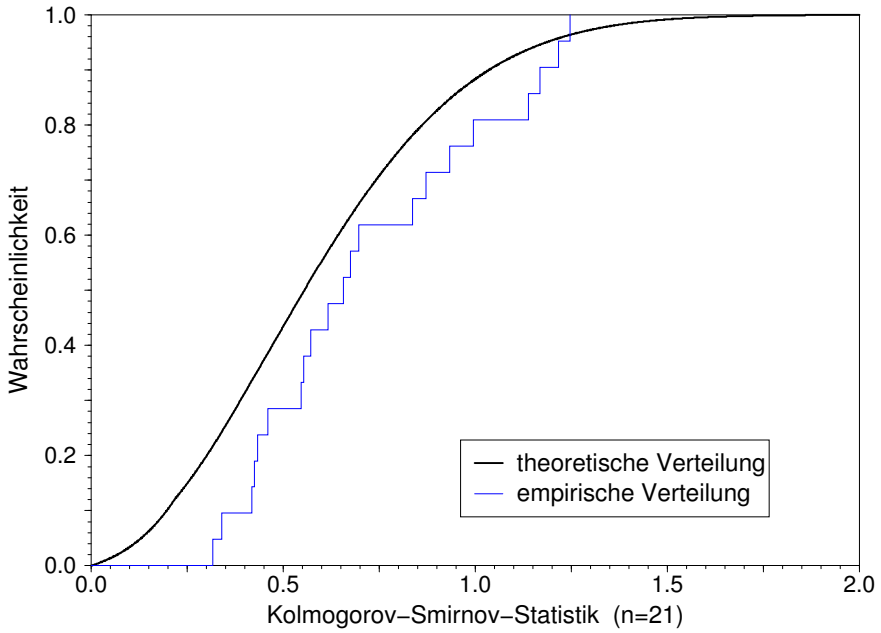


Abbildung B.36: Darstellung der empirischen Verteilungsfunktion aus den K_{128}^{-}-Werten von 21 zweistufigen Autokorrelationskoeffiziententests auf dem Lauf *LD8* für den ersten Autokorrelationskoeffizienten

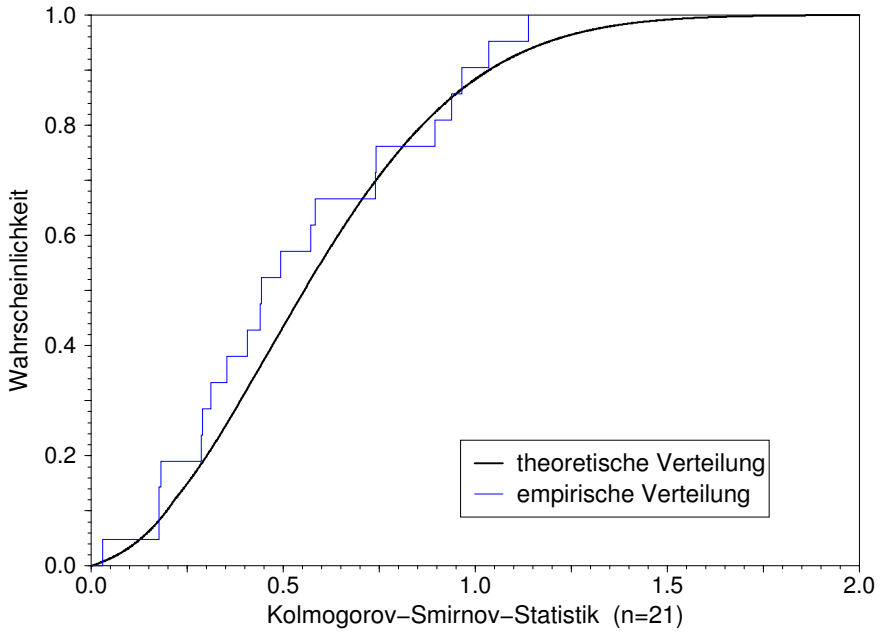


Abbildung B.37: Darstellung der empirischen Verteilungsfunktion aus den K_{128}^{+}-Werten von 21 zweistufigen Autokorrelationskoeffiziententests auf dem Lauf *LD8* für den ersten Autokorrelationskoeffizienten

Die beim Lauf *LH1* mit hybrider Datenaufnahme-Elektronik und dem später nochmals zur Kontrolle durchgeführten ähnlichen Lauf *LH4* gewählte Bit-Signallänge von 4,6 μs hat sich als zu lang herausgestellt; bei der Kontrolle der Signale mit einem Oszilloskop war zu erkennen, daß während dieser Zeit weitere Übernahme-Signale eintreffen, die auch noch einen weiteren Einlesevorgang beim PIC auslösen können.
Sowohl bei Lauf *LH1* als auch bei Lauf *LH4* treten sehr starke Antikorrelationen auf, die sich insbesondere in den Wahrscheinlichkeiten für $P(K_{128}^-)$ der Kolmogorov-Smirnov-Auswertung niederschlagen: Alle 52 zweistufigen Autokorrelationskoeffizententests des ersten Autokorrelationskoeffizienten liegen bei Lauf *LH1* beim maximalen Wert $P(K_{128}^-) = 1{,}0$ und bei Lauf *LH4* ist lediglich einer der 15 Testwerte ungleich[49] diesem Maximalwert. Die Gründe für diese starken Antikorrelationen werden in Abschnitt 6.1.1.2 erörtert.
Bei allen anderen Läufen, welche die hybride Elektronik verwenden, wird daher für die Bit-Signallänge eine Zeitdauer von 2 μs gewählt, die gerade der minimalen Zeit entspricht, die zur Aufnahme eines Bits durch den PIC benötigt wird.
Betrachtet man sich die in Abb. B.38 abgebildeten ersten fünf zweistufigen Autokorrelationskoeffizententests für den ersten Autokorrelationskoeffizienten des Laufes *LH5*, so fallen bis auf die bereits weiter oben auf S. 175 dargelegten FIFO-Probleme bei der dritten 1 MB umfassenden Stichprobe keine weiteren markanten Abweichungen auf.

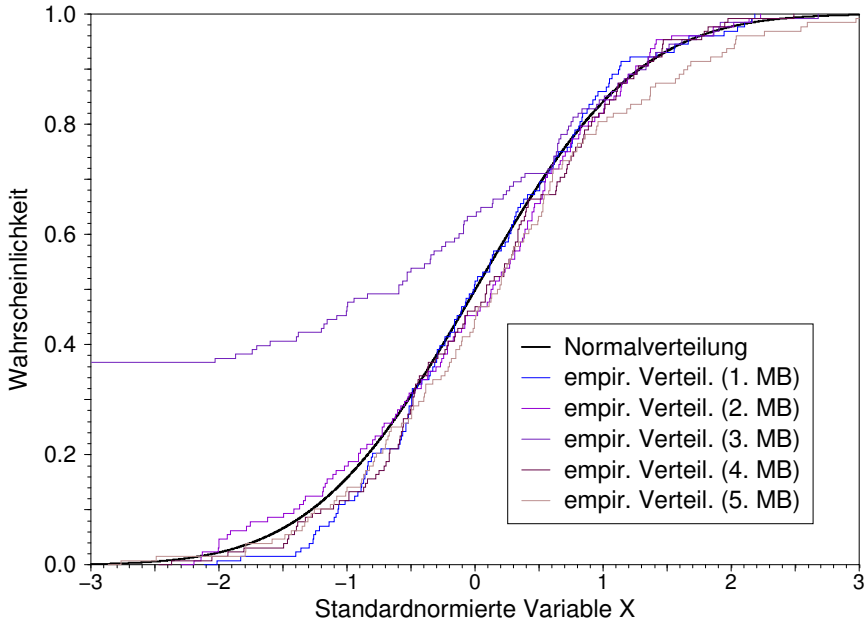


Abbildung B.38: Empirische Verteilungsfunktionen des ersten Autokorrelationskoeffizienten von fünf zweistufigen Tests auf Daten des Laufes *LH5*

Untersucht man allerdings die Verteilung der Werte K_{128}^- der insgesamt 53 zweistufigen Autokorrelationskoeffizententests in einer weiteren Teststufe und vergleicht sie mit der entsprechenden theoretischen Kolmogorov-Smirnov-Verteilung, ergibt sich ein etwas anderes Bild, s. Abb. B.39. Man erkennt, daß die Verteilung[50] der K_{128}^- zu größeren Werten hin verschoben ist, d.h. die maximalen Abstände der Verteilungen der zweiten Teststufe liegen zu weit von der theoretischen Verteilung entfernt, was genau dann der Fall ist, wenn Antikorrelationen auftreten. Diese sind

[49] Der Wert beträgt $P(K_{128}^-) = 0{,}99985$, was natürlich auch nicht wesentlich besser ist.

[50] Auf eine graphische Darstellung der Verteilung der K_{128}^+ wird hier verzichtet, sie zeigt – wie zu erwarten – genau das entgegengesetzte Verhalten.

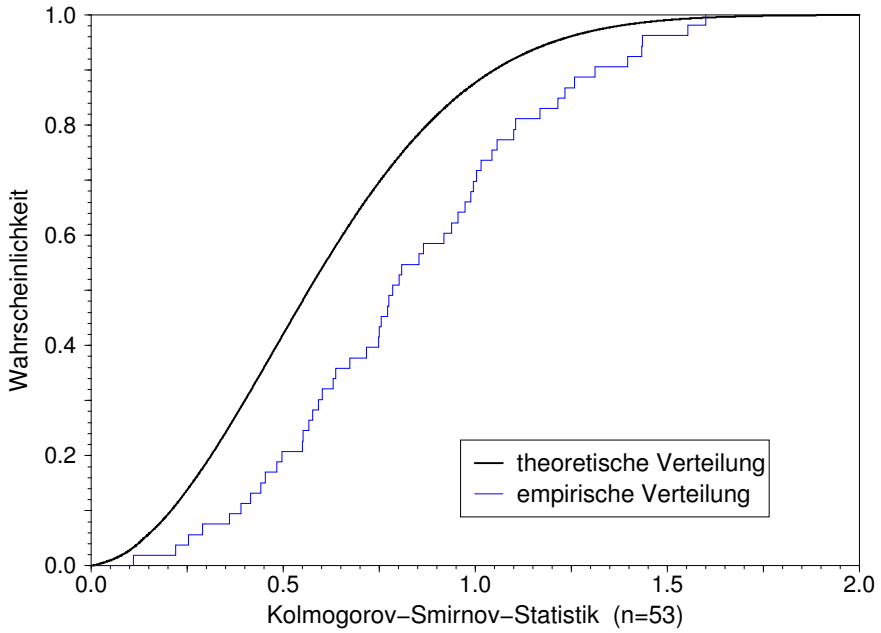


Abbildung B.39: Vergleich der empirischen Verteilungsfunktion aus den Ergebnissen K_{128}^{-} von 53 zweistufigen Autokorrelationskoeffiziententests für den ersten Autokorrelationskoeffizienten mit der theoretischen Verteilungsfunktion für den Lauf *LH5*

allerdings erheblich schwächer ausgeprägt als vorher, was sich darin zeigt, daß sie sich nicht direkt auf der zweiten Teststufe erkennen lassen, sondern erst mit einem dreistufigen Test.
Bei den folgenden Untersuchungen werden keine zweistufigen Autokorrelationskoeffiziententests mehr verwendet, sondern nur noch dreistufige, da diese offensichtlich eine höhere Nachweiskraft haben.
In Abb. B.40 ist das Ergebnis eines dreistufigen Autokorrelationskoeffiziententests für die Werte der negativen Abweichungen K_{128}^{-} von 55 zweistufigen Autokorrelationskoeffiziententests des Laufes[51] *LF1* eines Zufallsgenerators mit faseroptischem Aufbau und einer Einphotonenquelle auf Basis der parametrischen Fluoreszenz dargestellt. Auch hier sieht man ähnliche Abweichungen wie beim dreistufigen Test des Laufes *LH5*.
Zum Vergleich ist in Abb. B.41 ein dreistufiger Test auf den Werten der negativen Abweichungen K_{128}^{-} von 55 zweistufigen Autokorrelationskoeffiziententests des Laufes[52] *PB2* dargestellt.
Man erkennt deutlich, daß bei dem dreistufigen Test des Laufes *PB2* keine markanten Abweichungen der empirischen von der theoretischen Verteilungsfunktion zu erkennen sind. Vergleicht man die entsprechende Verteilung für die positiven Abweichungen K_{128}^{+}, die in Abb.B.42 dargestellt ist, mit der theoretischen Verteilung, so sind auch bei ihr keine markanten Abweichungen zu erkennen.
Auch die Auswertungen[53] der Läufe *LF2* und *PF* fügen sich in das bisherige Bild ein: Bei Lauf *LF2* zeigen sich ebenfalls noch schwache Antikorrelationen, wohingegen Lauf *PF* keine markanten Abweichungen[54] von der theoretischen Verteilung aufweist.

[51] Bei diesem Lauf wurden dieselben Werte für die Längen von Bit-Signal und Übernahme-Puls verwendet wie bei Lauf *LH5*.

[52] Dies ist ein Lauf eines Zufallsgenerators in freistrahloptischer Ausführung mit einer gepulsten, stark abgeschwächten Poisson-Lichtquelle, bei dem die hybride Datenaufnahme-Elektronik mit den gleichen Parametern wie bei Lauf *LF1* verwendet wird.

[53] Der kurze Testslauf *PB1* umfaßt nur 10 MB und wird daher nicht für dreistufige Tests herangezogen.

[54] Da „nur“ 30 MB Stichprobengröße zur Verfügung stehen, ist die empirische Verteilungsfunktion in

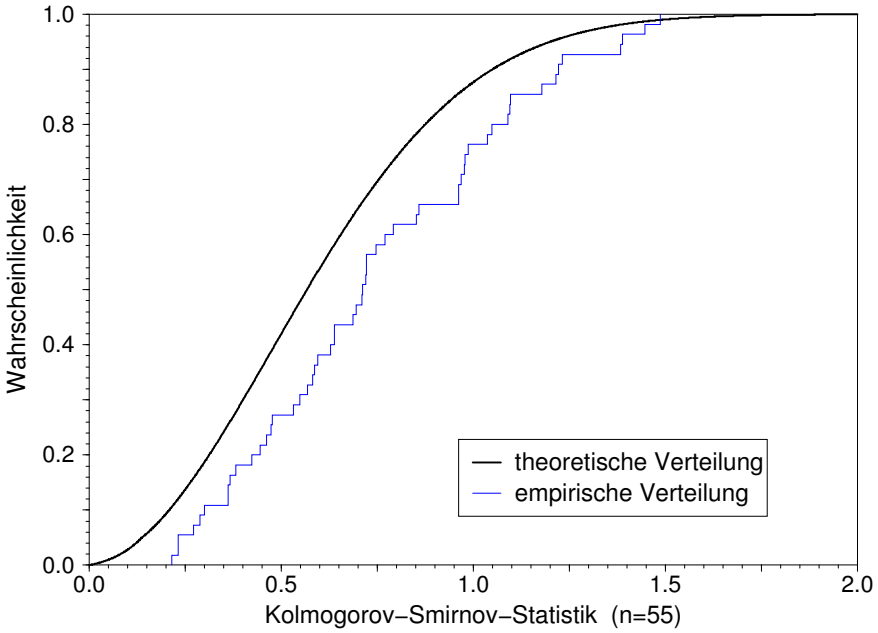


Abbildung B.40: Vergleich der empirischen Verteilungsfunktion aus den Ergebnissen K_{128}^{-} von 55 zweistufigen Autokorrelationskoeffizententests für den ersten Autokorrelationskoeffizienten mit der theoretischen Verteilungsfunktion für den Lauf *LF1*

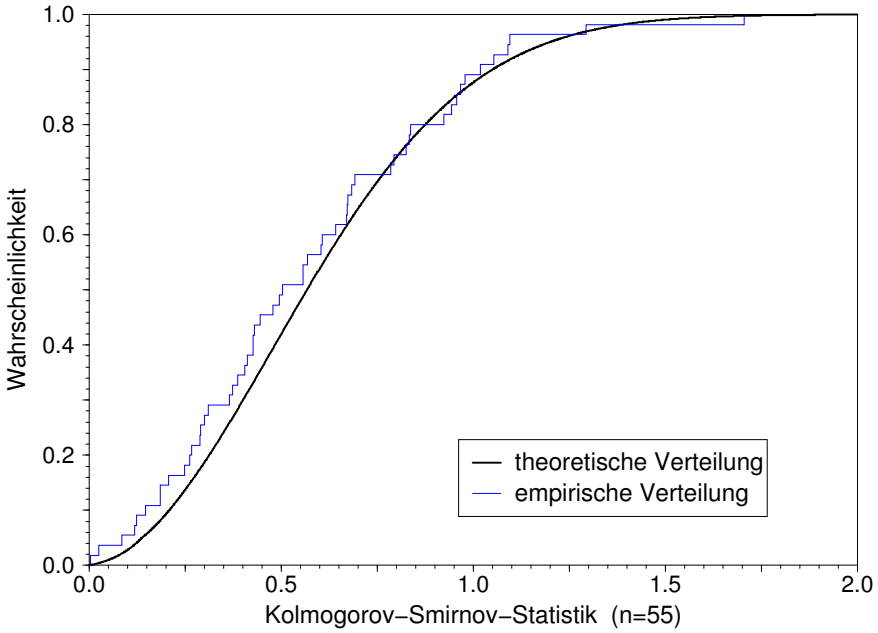


Abbildung B.41: Vergleich der empirischen Verteilungsfunktion aus den Ergebnissen K_{128}^{-} von 55 zweistufigen Autokorrelationskoeffizententests für den ersten Autokorrelationskoeffizienten mit der theoretischen Verteilungsfunktion für den Lauf *PB2*

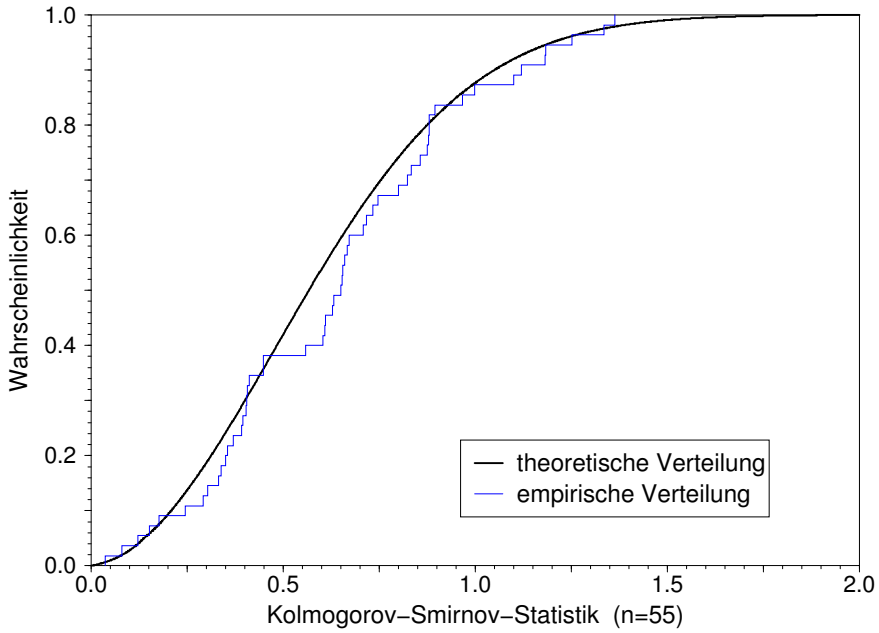


Abbildung B.42: Vergleich der empirischen Verteilungsfunktion aus den Ergebnissen K^{+}_{128} von 55 zweistufigen Autokorrelationskoeffiziententests für den ersten Autokorrelationskoeffizienten mit der theoretischen Verteilungsfunktion für den Lauf *PB2*

B.11 Die Mächtigkeit der Autokorrelationstests

Wie in den letzten Abschnitten deutlich geworden sein dürfte, sind Autokorrelationskoeffiziententests ein sehr gutes Werkzeug zur Untersuchung von Abhängigkeiten in Bitsequenzen, wobei ein großer Vorteil dieser Tests in ihrer guten Interpretierbarkeit liegt, die direkte Rückschlüsse auf die Art der Defekte erlaubt. Anderes als die vorher dargelegten Tests erlauben es die Autokorrelationskoeffiziententests auch noch schwache (Anti-)Korrelationen aufzudecken, wobei sie mit jeder weiteren Teststufe das Aufdecken feinerer Abweichungen von der Korrelationsfreiheit erlauben. Allerdings muß einschränkend darauf hingewiesen werden, daß dies nur dann gilt, wenn es sich um einen globalen Defekt handelt, der sich auch in der gesamten Stichprobe bemerkbar macht; lokale Abweichungen würden durch die Vergrößerung der Gesamtstichprobe bei den höheren Teststufen eher verdeckt werden. Dies erkennt man auch an den Ergebnissen der ersten fünf zweistufigen Tests des Laufes *LH5*, s. Abb. B.38, bei denen ein Autokorrelationskoeffiziententest aufgrund eines lokalen Defektes, s. a. Abschnitt B.9.2, sehr stark von der theoretischen Verteilung abweicht. Im dreistufigen Test des gesamten Laufs hingegen, s. Abb. B.39, fällt dieser einzelne Test nicht auf und würde bei einer quantitativen Auswertung mit Hilfe der Kolmogorov-Smirnov-Statistik auch nicht erkannt. Somit haben also auch die niedrigeren Teststufen durchaus ihre Berechtigung.

diesem Fall natürlich etwas grober.

Anhang C

Berechnung der Zählraten beim HOM-Zufallsgenerator

Bei den folgenden Berechnungen wird zwar ein von 50:50 abweichendes Teilungsverhältnis des Strahlteilers zugelassen und es werden auch ungleiche, vom Idealwert abweichende und unterschiedliche Quanteneffizienzen der beiden Detektoren[1] berücksichtigt, aber es wird idealisierend davon ausgegangen, daß die Photonen eines Paares ideal am Strahlteiler überlappen und weder vorher noch am Strahlteiler Verluste durch Absorption, Reflektion an Oberflächen oder Streuung vorhanden sind. Der betrachtete Idealfall entspräche also einer Messung mit maximaler Sichtbarkeit.

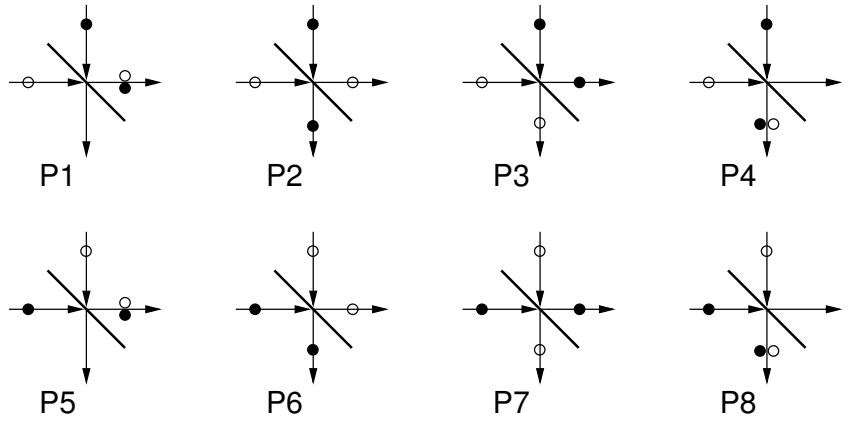


Abbildung C.1: Überlagerungsprozesse am Strahlteiler beim HOM-Aufbau

Bei den folgenden Berechnungen werden Wahrscheinlichkeitsamplituden, s. [18], für die verschiedenen Möglichkeiten der Überlagerung der Photonen am Strahlteiler, s. Abb. C.1, verwendet. Die zu den Überlagerungsprozessen gehörenden Werte der Wahrscheinlichkeitsamplituden sind in Tabelle C.1 aufgeführt[2]. Hierbei stehen r und t für die komplexen Reflexions- bzw. Transmis-

[1] Überdies werden keine Totzeiteffekte der Detektoren berücksichtigt, da davon ausgegangen wird, daß diese bei einem Zufallsgenerator im HOM-Aufbau durch die nachfolgende Datenaufnahmeelektronik neutralisiert werden.

[2] Der besseren Lesbarkeit wegen werden im weiteren die Wahrscheinlichkeitsamplituden aus Tabelle C.1 von rechts nach links und zeilenweise von oben nach unten mit A_1 bis A_8 abgekürzt.

sionskoeffizienten, für die bei Vernachlässigung von Verlusten gilt:

$$\begin{aligned} |r|^2 + |t|^2 &= 1 \\ r^* t + r t^* &= 0. \end{aligned}$$

Der relative Phasensprung um $\pi/2$ zwischen Transmission und Reflexion wird bei den Wahrscheinlichkeitsamplituden explizit durch einen Faktor i vor dem Reflexionskoeffizienten berücksichtigt, so daß also r und t im weiteren reellwertig sind.

P	P1	P2	P3	P4
A	$A_1 = \frac{1}{\sqrt{2}} irt$	$A_2 = \frac{1}{\sqrt{2}} t^2$	$A_3 = \frac{-1}{\sqrt{2}} r^2$	$A_4 = \frac{1}{\sqrt{2}} irt$
P'	P5	P6	P7	P8
A'	$A_5 = \frac{1}{\sqrt{2}} irt$	$A_6 = \frac{-1}{\sqrt{2}} r^2$	$A_7 = \frac{1}{\sqrt{2}} t^2$	$A_8 = \frac{1}{\sqrt{2}} irt$

Tabelle C.1: Wahrscheinlichkeitsamplituden für die Überlagerungsprozesse der beiden Photonen eines Photonenpaares am Strahlteiler

Im Fall ununterscheidbarer Photonen, s. Abschnitt C.1, müssen die zu nicht unterscheidbaren Prozessen gehörenden Amplituden A und A' erst summiert werden, bevor die Norm gebildet und quadriert wird, um die Gesamtwahrscheinlichkeit zu erhalten. Anders bei den unterscheidbaren Photonen, s. Abschnitt C.2, hier wird zuerst für jede Wahrscheinlichkeitsamplitude die quadrierte Norm gebildet und anschließend werden die resultierenden Einzelwahrscheinlichkeiten zur Gesamtwahrscheinlichkeit aufsummiert.
Neben den Wahrscheinlichkeitsamplituden gehen noch folgende Faktoren in die Formeln für die Einzelzählraten der Detektoren ein:

- N_p, die mittlere Rate der auf den Strahlteiler treffenden Photonenpaare, die mittlere Rate der Photonen ist also $2N_p$,
- η_{p0} und η_{p1}, die effektiven Quanteneffizienzen der jeweiligen Signaldetektoren, wenn ein Photonenpaar auf die Detektorfläche fällt und
- η_{e0} und η_{e1}, die Quanteneffizienz der jeweiligen Signaldetektoren für die Detektion eines einzelnen Photons.

Hierbei gilt für die effektive Quanteneffizienz η_p, s. Abschnitt 6.3.2 :

$$\eta_p = 1 - (1 - \eta_e)^2 = 2 \cdot \eta_e - \eta_e^2 = \eta_e \cdot (2 - \eta_e)$$

C.1 Fall 1: Ununterscheidbare Photonen

Bei ununterscheidbaren Photonen haben wir zwei nicht zu unterscheidende Prozesse, bei denen ein Photonenpaar auf den jeweiligen Detektor fällt und vier ebenfalls voneinander nicht zu unterscheidende Prozesse, bei denen nur jeweils ein Photon auf den Detektor fällt. Während sich die Wahrscheinlichkeitsamplituden für die erstgenannten Prozesse, je nach betrachtetem Ausgang unterscheiden, ist die Summe der Wahrscheinlichkeitsamplituden der letztgenannten Prozesse für beide Ausgänge gleich, da in dem von uns betrachteten, verlustfreien Fall dann in jedem Ausgang ein Photon vorhanden ist.
Zur Berechnung der Zählraten der Signaldetektoren müssen die Ereignisse, bei denen ein Photonenpaar auf den Detektor fällt, mit der effektiven Quanteneffizienz gewichtet werden, während die Ereignisse, bei denen nur ein einzelnes Photon auf den Detektor fällt, lediglich mit der einfachen Quanteneffizienz eingehen.

Für die Zählrate eines Detektors im „horizontalen Ausgang“ von Abb. C.1 setzt man daher an:

$$N_0 = N_p \cdot \left(\eta_{p0} \cdot |A_1 + A_5|^2 + \eta_{e0} \cdot |A_2 + A_3 + A_6 + A_7|^2\right),$$

bzw. für den anderen Ausgang:

$$N_1 = N_p \cdot \left(\eta_{p1} \cdot |A_4 + A_8|^2 + \eta_{e1} \cdot |A_2 + A_3 + A_6 + A_7|^2\right).$$

Setzt man nun die Wahrscheinlichkeitsamplituden aus Tabelle C.1 in die einzelnen Terme ein, so erhält man nach kurzer Rechnung:

$$\begin{aligned} |A_1 + A_5|^2 &= 2 \cdot r^2 \cdot t^2 \\ |A_4 + A_8|^2 &= 2 \cdot r^2 \cdot t^2 \\ |A_2 + A_3 + A_6 + A_7|^2 &= 2 \cdot (t^2 - r^2). \end{aligned}$$

Hieraus ersieht man sofort, daß die Formeln für die Zählraten der Detektoren in den Ausgängen des Strahlteilers sich nur durch die unterschiedlichen (effektiven) Quanteneffizienzen unterscheiden:

$$N_{0/1} = N_p \cdot \left(\eta_{p0/1} \cdot 2 \cdot r^2 \cdot t^2 + \eta_{e0/1} \cdot 2 \cdot (t^2 - r^2)\right).$$

Handelt es sich um einen Strahlteiler mit nicht idealem Teilungsverhältnis, bei dem also Transmission t^2 und Reflexion r^2 nicht gleich sind, so spielt dies – anders im Fall einzelner Photonen am Strahlteiler – hier demnach keine Rolle. Dies ist angesichts der Symmetrie der Prozesse, welche in die Berechnung der Zählrate eingehen, auch anschaulich klar.
Sind Transmission und Reflexion gleich, d.h. $r^2 = t^2 = 1/2$, so ist die Differenz im zweiten Term gerade identisch Null und die Photonen eines Paares laufen immer nur gemeinsam in einem Ausgang; in diesem Fall erhält man für die mittleren Zählraten $N_{0/1} = \eta_{p0/1} \cdot N_p/2$.
Auch im Fall eines mäßig vom Ideal abweichenden Teilungsverhältnisses wird der zweite Term aufgrund der Differenzbildung nur wenig beitragen. Somit geht die Quanteneffizienz der Detektoren in erster Linie durch die effektive Quanteneffizienz in die Zählrate ein. Unterschiedliche Quanteneffizienzen der Detektoren werden somit beim HOM-Zufallsgenerator die Hauptursache für Abweichungen vom idealen Fall gleicher relativer Häufigkeiten der Nullen und Einsen sein.
Da es hauptsächlich die effektive Quanteneffizienz ist, die in die Zählraten eingeht, ist es interessant zu betrachten, wie sich eine geringe Differenz $\Delta\eta_e = \eta_{e1} - \eta_{e0}$ zwischen den Quanteneffizienzen der Signaldetektoren auf die Differenz der effektiven Quanteneffizienzen, $\Delta\eta_p = \eta_{e1} - \eta_{e0}$ auswirkt. Setzt man $\eta_{e1} = \eta_{e0} + \Delta\eta_e$ in die Formel für die effektive Quanteneffizienz ein[3] und formt etwas um, so erhält man für diese Differenz:

$$\Delta\eta_p = \Delta\eta_e \cdot (2 - 2\eta_0 - \Delta\eta_e).$$

Je nachdem wie groß der Ausdruck in der Klammer ist, kann die Abweichung bei der effektiven Quanteneffizienz größer oder kleiner als die Differenz der einfachen Quanteneffizienzen sein. Da $\Delta\eta_e$ meist relativ klein ist und typischerweise in der Größenordnung von $\eta/10$ liegt, wird die Summe von den ersten beiden Summanden dominiert. Bei einer Quanteneffizienz von $\eta_e \approx 0{,}5$ liegt der Klammerterm bei Eins und dementsprechend ist $\Delta\eta_p \approx \Delta\eta_e$. Liegt die Quanteneffizienz η_e unter diesem Schwellwert, ist die Differenz $\Delta\eta_p > \Delta\eta_e$; für einen Wert von $\eta_e \approx 0{,}3$ wäre z.B. $\Delta\eta_p \approx 1{,}4 \cdot \Delta\eta_e$; unterschiedliche Quanteneffizienzen wirken sich in diesem Fall also stärker aus als beim Zufallsgenerator mit einzelnen Photonen.

[3] Hierbei wird oBdA. angenommen, daß $\eta_{e1} > \eta_{e0}$ ist.

C.2 Fall 2: Unterscheidbare Photonen

In diesem Fall sind alle Prozesse am Strahlteiler voneinander unterscheidbar und man erhält folgenden Ansatz für die Zählraten:

$$\begin{aligned} N_0 &= N_p \cdot \left(\eta_{p0} \cdot \left(|A_1|^2 + |A_5|^2 \right) + \eta_{e0} \cdot \left(|A_2|^2 + |A_3|^2 + |A_6|^2 + |A_7|^2 \right) \right), \text{ bzw.} \\ N_1 &= N_p \cdot \left(\eta_{p1} \cdot \left(|A_4|^2 + |A_8|^2 \right) + \eta_{e1} \cdot \left(|A_2|^2 + |A_3|^2 + |A_6|^2 + |A_7|^2 \right) \right). \end{aligned}$$

Auch hier sind in den beiden Gleichungen wieder die Terme gleich, deren Amplituden zu Prozessen gehören, bei denen die Photonen in unterschiedliche Ausgangsarme laufen. Setzt man die Wahrscheinlichkeitsamplituden ein, so erhält man unter Berücksichtigung von:

$$\begin{aligned} |A_1|^2 &= \frac{1}{2} r^2 t^2 = |A_4|^2 \\ |A_5|^2 &= \frac{1}{2} r^2 t^2 = |A_8|^2 \\ |A_2|^2 &= \frac{1}{2} t^2 t^2 = |A_7|^2 \\ |A_3|^2 &= \frac{1}{2} r^2 r^2 = |A_6|^2, \text{ gerade:} \end{aligned}$$

$$N_{0/1} = N_p \cdot \left(\eta_{p0/1} \cdot r^2 \cdot t^2 + \eta_{e0/1} \cdot \left(r^4 + t^4 \right) \right).$$

Die Formeln für die mittleren Einzelzählraten unterscheiden sich also auch in diesem Fall wieder nur in den unterschiedlichen (effektiven) Quanteneffizienzen; Abweichungen des Strahlteilers vom idealen Teilungsverhältnis spielen keine Rolle. Allerdings sind nun die Entscheidungsprozesse der beiden Photonen am Strahlteiler unabhängig voneinander; daher fällt auch im Idealfall $r^2 = t^2 = 1/2$ der Anteil der Prozesse, bei denen die Photonen in unterschiedliche Arme gehen, nicht weg, sondern es ergibt sich:

$$N_{0/1} = N_p \cdot \left(\eta_{p0/1} \cdot \frac{1}{4} + \eta_{e0/1} \cdot \frac{1}{2} \right).$$

Man erkennt, daß im Falle unterscheidbarer Photonen der Anteil der einzelnen Photonen, die auf den Detektor fallen, vorherrscht. Bei den Einzelzählraten im Falle unterscheidbarer Photonen muß natürlich berücksichtigt werden, daß alle Ereignisse verworfen werden, bei denen beide Detektoren gleichzeitig ansprechen. Somit trägt der zweite Term im Falle idealer Quanteneffizienzen gar nicht und unter realen Bedingungen in verminderter Form zur Zufallsgenerierung bei.

Bei realen Quanteneffizienzen muß aber noch berücksichtigt werden, daß zu kleiner Quanteneffizienz hin die Wahrscheinlichkeit zunimmt, daß nur einer der beiden Detektoren anspricht, auch wenn auf beide ein Photon fällt. Wie stark dieser Effekt ist, läßt sich berechnen, indem man jeweils die beiden Terme zusammenfaßt, die zu den einzelnen Photonen in einem Ausgangsarm gehören, und mit der Wahrscheinlichkeit wichtet, daß genau einer der beiden Detektoren anspricht.

Insgesamt fallen $2 \cdot N_p$ Photonen auf den Strahlteiler, und mit einer Wahrscheinlichkeit von $P_{01} = r^4 + t^4$ tritt der Fall auf, daß sie getrennt in die beiden Ausgänge des Strahlteilers laufen. Die Wahrscheinlichkeit, daß darauf nur einer der beiden Detektoren anspricht und somit ein Zufallsbit generiert wird, beträgt:

$$\begin{aligned} P(1Bit|2Photonen) &= \eta_0 \cdot (1 - \eta_1) + \eta_1 \cdot (1 - \eta_0) \\ &= \eta_0 + \eta_1 - 2\eta_0\eta_1 \end{aligned}$$

Somit erhält man also für den Anteil N_e der Zufallsereignisse, die nicht aufgrund eines Photonenpaares in einem der Ausgänge des Strahlteilers generiert wurden, sondern durch ein einzelnes Photon:

$$N_e = 2 \cdot N_p \cdot (r^4 + t^4) \cdot (\eta_0 + \eta_1 - 2\eta_0\eta_1).$$

Betrachtet man den Fall mit idealem Teilungsverhältnis $r^2 = t^2 = 1/2$ und gleichen Quanteneffizienzen, so erhält man für eine typische Quanteneffizienz von $\eta = 0{,}3$:

$$\begin{aligned} N_e &= 2 \cdot N_p \cdot \left(\frac{1}{4} + \frac{1}{4}\right) \cdot \left(2\eta - 2\eta^2\right) \\ &= N_p \cdot 0{,}42 \ . \end{aligned}$$

Man sieht also, wie stark die nichtidealen Quanteneffizienzen dazu beitragen, daß die im Idealfall aufgrund des Ansprechens beider Detektoren vollständig verworfenen Ereignisse doch mit in die Zufallsgenerierung eingehen.

Anhang D

Approximative Verteilungsfunktionen

Bei zweistufigen, statistischen Tests wird eine empirisch ermittelte Verteilungsfunktion mit einer theoretischen verglichen. Oft handelt es sich bei dieser theoretischen Verteilungsfunktion allerdings „nur“ um eine Approximation der exakten Verteilungsfunktion. So wird zum Beispiel beim Test der Verteilung der Autokorrelationskoeffizienten als theoretische Verteilung die Normalverteilung verwendet, die exakte Verteilungsfunktion ist hingegen durch die (kumulierte) Binomialverteilung[1] gegeben. Es stellt sich daher die Frage, ob die Approximation gut genug ist oder die Gefahr besteht, daß sie das Resultat des Tests verfälscht.

D.1 Die Güte der Normalapproximation

Einer der Sätze von Berry-Esséen liefert eine Abschätzung für die Güte der Approximation der Normalverteilung $N(x)$:

$$N(x) = \int_{-\infty}^{x} \phi(y)dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-y^2}{2}} dy$$

an die (kumulative) Binomialverteilung $B(k; n, p)$:

$$B(k; n, p) = \sum_{j=0}^{k} \binom{n}{j} p^j (1-p)^{n-j},$$

s. z.B. [70], S. 107. Unter Verwendung von standardisierten Zufallsvariablen

$$X^* = \frac{X - np}{\sqrt{np(1-p)}},$$

erhält man für $\forall a, b, n, p$, wobei $-\infty \leq a \leq b \leq +\infty$ und $0 \leq p \leq 1$:

$$\left| P\left\{a \leq X^* \leq b\right\} - \int_a^b \phi(x)dx \right| \leq \frac{0.8}{\sqrt{np(1-p)}}.$$

[1] Im Prinzip läßt sich die Binomialverteilung heutzutage natürlich per Computer auch für große n berechnen, allerdings wäre sie ohnehin nicht als theoretische Verteilungsfunktion für einen Kolmogorov-Smirnov-Test geeignet, da bei diesem eine *stetige* theoretische Verteilungsfunktion vorausgesetzt wird.

Bei Zufallsgeneratoren weicht p in der Regel nicht stark vom Idealwert $p = 1/2$ ab, so daß die Abweichung der Approximation von der exakten Verteilung kleiner als $\approx 0{,}4/\sqrt{n}$ ist.
Beim Autokorrelationskoeffizienten-Test (s. S. 51) werden Stichproben mit 8 kB Daten genommen, d.h. $n = 65536$ Bits; in diesem Fall weicht die Approximation also höchstens um 0,0015625 vom exakten Wahrscheinlichkeitswert ab. Kritisch könnte solch eine Abweichung natürlich besonders dann werden, wenn sie an einer Stelle der Größe eines Wahrscheinlichkeitwertes der Verteilung gleichkommt. Um abzuschätzen, inwieweit dies tatsächlich ein Problem bei Kolmogorov-Smirnov-Tests darstellen könnte, empfiehlt es sich, die Normalverteilung besser doch einmal explizit mit der kumulierten Binomialverteilung zu vergleichen. Hierzu wird eine Berechnungsroutine für die kumulierte Binomialverteilung verwendet, die auch mit den sehr großen Binomialkoeffizienten $\binom{65536}{k}$ umgehen kann. Die auftretenden Binomialkoeffizienten mit $k = 1 \ldots 1000$ bzw. $k = 64536 \ldots 65536$ werden exakt berechnet, während bei der Berechnung der übrigen Koeffizienten eine Stirling-Approximation[2] verwendet wird. Diese Approximation ist sehr gut: Der relative Fehler der approximierten Binomialkoeffizienten gegenüber den exakten[3] ist hierbei kleiner als 10^{-17}. Dies führt zu einer relativen Abweichung der approximativ berechneten, kumulierten Binomialverteilung von der exakten Verteilung, die kleiner als 10^{-11} ist. Solch eine Abweichung spielt höchstens in den „Außenbereichen“ der Verteilung eine Rolle; diese Bereiche haben allerdings für die praktische Durchführung der Tests keine Relevanz[4], so tragen sie insbesondere nicht zu den Abweichungen zwischen theoretischer und empirischer Verteilung bei, die mit Hilfe des Kolmogorov-Smirnov-Tests untersucht werden.
In Abb. D.1 sind sowohl die nach oben erwähnter Vorgehensweise berechnete Binomialverteilung $B(X)$ für die Parameter $n = 65536$ und $p = 0{,}5$ als auch die entsprechende Normalverteilung $N(x)$ aufgetragen; wie man sieht, liegen die beiden Graphen außerordentlich nah „beieinander“.
Damit sich die geringen Unterschiede besser erkennen lassen, ist in Abb. D.2 die Differenz $B(X) - N(X)$ zwischen der kumulierten Binomialverteilung $B(X)$ und der Normalverteilung $N(X)$ dargestellt, wobei zusätzlich noch die oben erwähnte obere Grenze für die Abweichung eingezeichnet ist. Wie man deutlich erkennt, stellt sie in der Tat eine gute Abschätzung für die *maximale* Abweichung der Normal-Approximation dar. Da es sich bei ihr aber nur um einen konstanten Wert handelt, gibt sie die lokalen Abweichungen abseits des Mittelwertes nur schlecht wieder.
Abschließend läßt sich zusammenfassen, daß die Normalverteilung für die im weiteren verwendeten Stichprobengrößen eine exzellente Approximation der kumulierten Binomialverteilung darstellt und daher keinerlei Beeinträchtigungen der Kolmogorov-Smirnov-Tests durch Approximations-Artefakte zu befürchten sind.

D.2 Die Güte der χ^2 Approximation

Leider läßt sich die Güte der χ^2-Approximation nur mit einem *erheblich* größeren Rechenaufwand durch einen Vergleich mit der exakten Multinomialverteilung ermitteln; daher wird hier auch darauf verzichtet. Allerdings stellt sich bei einem solchen Vergleich heraus [31], daß in der Tat an den Rändern der Verteilung eine merkliche Abweichung zwischen der Approximation und der exakten Multinomialverteilung besteht. Dies kann dazu führen, daß bei sehr kleinen Ablehnungswerten die Generatoren häufiger verworfen werden, als man bei Verwendung der Approximation unter der Nullhypothese erwarten würde.

[2] Bis zur vierten Ordnung: $n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} + \frac{1}{288n^2} + \frac{139}{51840n^3} + \frac{571}{2488320n^4}\right)$, s. [2].

[3] Diesen Wert für den relativen Fehler erhält man aus dem Quotienten des exakten und des approximierten Koeffizienten für $\binom{65536}{1000}$, für $k > 1000$ ist der relative Fehler sogar noch kleiner.

[4] Die Abweichung spielt in den Außenbereichen gerade deswegen eine Rolle, weil sie in die Größenordnung der Wahrscheinlichkeit selbst kommt; das heißt aber gleichzeitig auch, daß solche Werte der standardnormierten Variable X nur sehr selten auftreten.

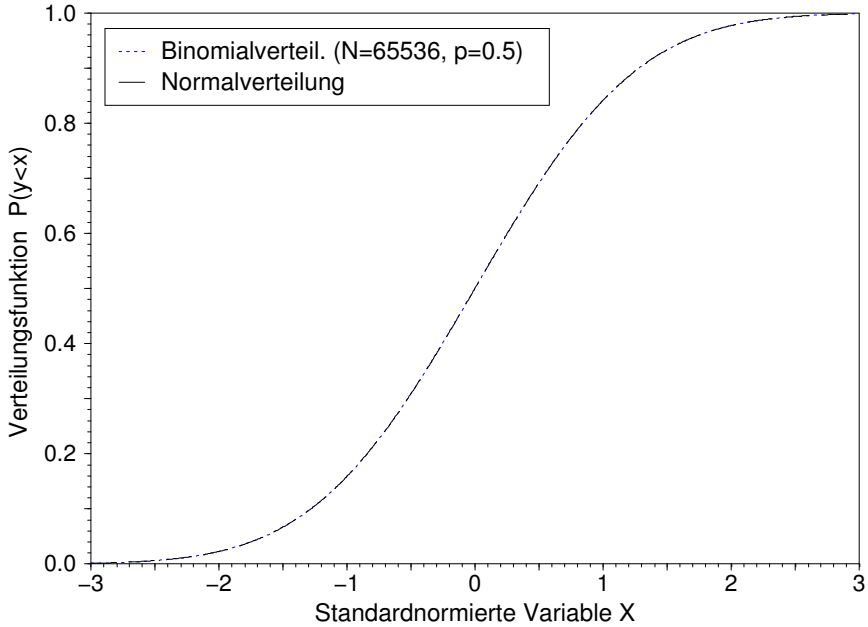


Abbildung D.1: Kumulierte, standardnormierte Binomialverteilung $B(X)$ und Normalverteilung $N(X)$

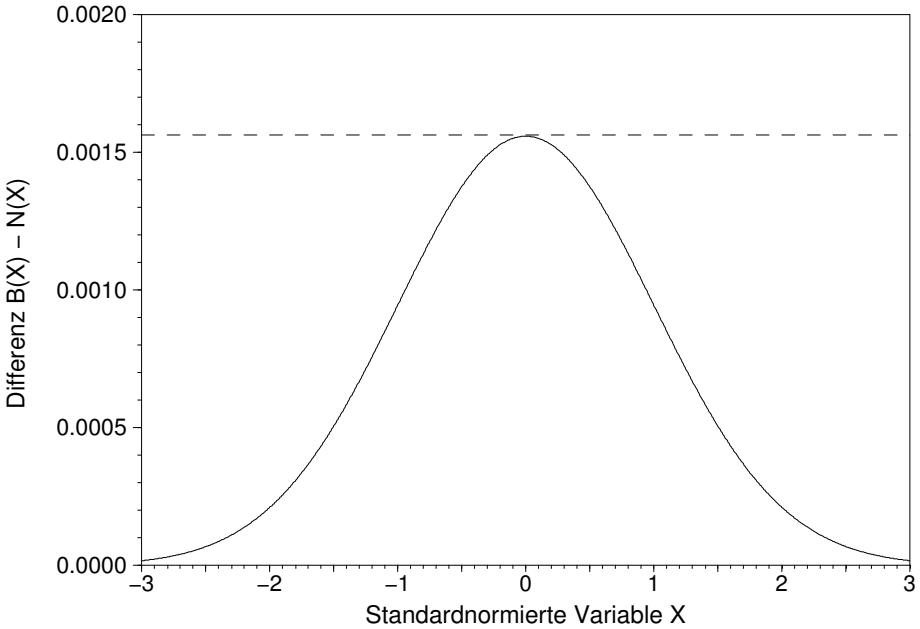


Abbildung D.2: Differenz zwischen der kumulierten, standardnormierten Binomialverteilung $B(X)$ und der Normalverteilung $N(X)$, gestrichelte Linie: $0{,}4/\sqrt{(n)}$ für $n = 65536$

Anhang E

Theorie und Details der universellen Tests

E.1 Theorie des universellen Tests

Die effektive Schlüssellänge für eine *binäre Quelle ohne Gedächtnis* ($BQoG_p$), die den Bitwert Eins mit einer Wahrscheinlichkeit p ausgibt, ist, wie in Abschnitt 3.2.7 bereits aufgeführt, gleich:

$$\log_2 \mu_{BQoG_p}(L, 1/2) \approx \log_2 \sum_{i=0}^{pL} \binom{L}{i},$$

aber als Berechnungsvorschrift für die effektive Schlüssellänge läßt sich dieser Ausdruck nur schlecht verwenden. Deshalb wird bei der Durchführung des universellen Tests anders vorgegangen und eine äquivalente Teststatistik[1] benutzt. Bei ihrer Berechnung werden die Abstände verwendet, in denen sich identische Bit-Blöcke innerhalb einer in aufeinanderfolgende, nichtüberlappende Bit-Blöcke aufgeteilten Binärsequenz wiederholen. Teilt man die Bitsequenz in Blöcke der Länge L auf, d.h. $b_n(S^N) = [s_{L(n-1)+1}, \ldots, s_{L_n}]$, so bedeutet dies, daß für den Fall identischer Blöcke gilt: $b_n(S^N) = b_{n-i}(S^N)$. Der Abstand $A_n(S^N)$ zwischen zwei Blöcken entspricht dann gerade gleich i, falls der Block $b_n(S^N)$ bereits einmal aufgetreten ist oder gleich n, falls dies noch nicht der Fall war, was natürlich am Anfang der Sequenz häufig vorkommen wird. Daher teilt man die zu untersuchende Sequenz in einen Initialisierungsabschnitt mit Q Blöcken und einem Testabschnitt mit K Blöcken. Hierbei muß Q so groß gewählt werden, daß jeder Block möglichst bereits im Initialisierungsabschnitt auftritt; es empfiehlt sich ein Wert $Q \geq 10 \cdot 2^L$. Der Testabschnitt sollte so groß wie möglich gewählt werden, d.h. $K \geq 1000 \cdot 2^L$, wobei dies natürlich mit zunehmender Blockgröße L immer schwieriger wird, da der Testabschnitt exponentiell an Länge zunimmt, daher wird in der Praxis meist nur bis zu einem Wert von $L = 16$ getestet.
Die Teststatistik $f_T(s^N)$ wird dann definiert als der Mittelwert des Logarithmus zur Basis Zwei aus den K Abständen $A_{Q+1}(S^N), A_{Q+2}(S^N), \ldots, A_{Q+K}(S^N)$, d.h.:

$$f_U(s^N) = \frac{1}{K} \sum_{n=Q+1}^{Q+K} \log_2 A_n(S^N).$$

Die Verteilungsfunktion der Teststatistik $f_U(s^N)$ läßt sich für die Nullhypothese einer Bitsequenz

[1] Strenggenommen ist diese Teststatistik allerdings erst im Grenzwert großer Blocklängen wirklich äquivalent, was hinsichtlich der exponentiell mit der Blockgröße zunehmenden Länge der benötigten Bitsequenz nicht unproblematisch ist! Die Variante des universellen Tests von Coron behebt dieses Problem, s.u.

R^N aus einer symmetrischen, binären Quelle ohne Gedächtnis durch eine Normalverteilung approximieren mit dem Erwartungswert:

$$E[f_U(R^N)] = E[\log_2 A_n(R^N)] = 2^{-L} \sum_{i=1}^{\infty} (1 - 2^{-L})^{i-1} \log_2 i,$$

und der Standardabweichung:

$$\sigma = c(L, K) \sqrt{\frac{\mathrm{Var}[\log_2 A_n(R^N)]}{K}}.$$

Mit wachsender Länge K der Testsequenz nimmt σ ab, dementsprechend lassen sich Abweichungen eines Generators vom idealen Verhalten mit zunehmendem K besser aufdecken.
Die Varianz $\mathrm{Var}[\log_2 A_n(R^N)]$ in obiger Formel ist gegeben durch:

$$\begin{aligned} \mathrm{Var}[\log_2 A_n(R^N)] &= E[(\log_2 A_n(R^N))^2] - (E[\log_2 A_n(R^N)])^2 \\ &= 2^{-L} \sum_{i=1}^{\infty} (1 - 2^{-L})^{i-1} (\log_2 i)^2 - (E[f_U(R^N)])^2. \end{aligned}$$

Der Korrekturfaktor $c(L, K)$ rührt daher, daß die Abstände $A_n(R^N)$ nicht voneinander unabhängig sind; er verringert die Standardabweichung gegenüber dem Wert, den sie bei voneinander unabhängigen $A_n(R^N)$ hätte. In der Originalveröffentlichung von Maurer wurde $c(L, K)$ noch mit Hilfe einer Simulation bestimmt. Eine exakte, aber leider etwas unhandliche, analytische Formel findet sich in einer Arbeit von Jean-Sébastien Coron und David Naccache [26]; dort findet man glücklicherweise aber auch eine gute Approximation für $c(L, K)$:

$$c(L, K)^2 = d(L) + \frac{e(L) \times 2^L}{K};$$

die Werte für $E[f_U(R^N)]$, $\mathrm{Var}[\log_2 A_n(R^N)]$, $d(L)$ und $e(L)$ sind in der ebenfalls aus [26] entnommenen Tabelle E.1 für verschiedene Blocklängen L zusammengefaßt[2].

E.2 Durchführung des universellen Tests

Der universelle Test verwendet drei ganzzahlige Parameter L, Q, K und eine Bitsequenz des zu testenden Zufallsgenerators als Eingabe. Die drei Eingabe-Parameter L, Q, K stehen dabei für folgende Größen:

- L: die Block-Länge in Bits, typische Werte: $3 \leq L \leq 16$, für eine effiziente Implementierung eignen sich die Werte $L = 8$ bzw. $L = 16$ natürlich besonders gut,
- Q: die Anzahl der Initialisierungsschritte, wobei Q so groß gewählt wird, daß jeder Block mit hoher Wahrscheinlichkeit schon einmal aufgetreten sein sollte, d.h. $Q \geq 10 \cdot 2^L$ und
- K: die Anzahl der Testschritte; K sollte möglichst groß gewählt werden, d.h. möglichst $K \geq 1000 \cdot 2^L$.

Dementsprechend muß eine Stichprobe $s^N = [s_1 \ldots s_N]$ mit $(Q + K) \times L = N$ Bits zum Testen verfügbar sein. Bei großen Blocklängen nimmt die für den Test notwendige Länge der Bitsequenz exponentiell zu: So benötigt man für $Q = 10$ und $K = 1000$ bei einem Test mit 8 Bit Blocklänge lediglich 258.560 Byte, also ca. 258 kB, wohingegen man bei einem Test mit einer Bocklänge von $L = 16$ bereits 129.280 kB, d.h. ca. 126 MB benötigt.
Durchgeführt wird der Test folgendermaßen:

[2] $e(L)$ und $d(L)$ sind analytisch durch schnell konvergierende Potenzreihen gegeben, für Details s. [26].

L	$E[f_U(R^N)]$	$\mathrm{Var}[\log_2 A_n(R^N)]$	$d(L)$	$e(L)$
3	2.4016068	1.9013347	0.2732725	0.4890883
4	3.3112247	2.3577369	0.3045101	0.4435381
5	4.2534266	2.7045528	0.3296587	0.4137196
6	5.2177052	2.9540324	0.3489769	0.3941338
7	6.1962507	3.1253919	0.3631815	0.3813210
8	7.1836656	3.2386622	0.3732189	0.3730195
9	8.1764248	3.3112009	0.3800637	0.3677118
10	9.1723243	3.3564569	0.3845867	0.3643695
11	10.1700323	3.3840870	0.3874942	0.3622979
12	11.1687649	3.4006541	0.3893189	0.3610336
13	12.1680703	3.4104380	0.3904405	0.3602731
14	13.1676926	3.4161418	0.3911178	0.3598216
15	14.1674884	3.4194304	0.3915202	0.3595571
16	15.1673788	3.4213083	0.3917561	0.3594040
∞	$L - 0.8327462$	3.4237147	0.3920729	0.3592016

Tabelle E.1: $E[f_U(R^N)]$, $\mathrm{Var}[\log_2 A_n(R^N)]$, $d(L)$ und $e(L)$ für $3 \leq L \leq 16$ und $L \to \infty$

1. Die Bitsequenz s^N werde in n jeweils L Bit lange Blöcke unterteilt[3], d.h. für alle $1 \leq n \leq Q + K$ ist der n-te Block gegeben durch $b_n(S^N) = [s_{L(n-1)+1}, \ldots, s_{L_n}]$.

2. **Initialisierung**:

   - Man erstelle eine Tabelle mit 2^L Platzhaltern, indiziert durch die möglichen Werte eines Blockes $0, \ldots, 2^{L-1}$. Jeder Platzhalter muß groß genug sein, um einen Blockindex n aufnehmen zu können; hierzu reichen 2 Byte Speicherplatz pro Eintrag aus.
   - Man durchlaufe die ersten Q Blöcke der Blocksequenz und speichere für jeden Bitblock b_k den Index k an der entsprechenden Stelle der Tabelle. Nach dieser Initialisierung sind mit hoher Wahrscheinlichkeit alle Zellen der Tabelle mit Blockindices ausgefüllt.

3. **Testphase**: Man durchlaufe nun die restlichen K Blöcke der Sequenz und berechne:

$$f_U(s^N) = \frac{1}{K} \sum_{n=Q+1}^{Q+K} \log_2 A_n(S^N).$$

   Hierbei sind die Abstände A_n gegeben durch die Differenz zwischen dem Index n des aktuellen Blocks und dem Tabelleneintrag[4] zu diesem Blockwert. Nach Berechnung von A_n wird der Wert n in die entsprechende Zelle der Index-Tabelle geschrieben. Um die Laufzeit des Algorithmus zu verkürzen, summiert man die Werte für den Logarithmus $\log_2 A_n(S^N)$ auf und teilt erst nach Durchlaufen der gesamten Testsequenz durch die Anzahl der Testblöcke K.

[3] Bei einer Implementierung des Tests wird die „Unterteilung“ natürlich direkt beim eigentlichen Test-Schritt vollzogen.

[4] Der Tabelleneintrag entspricht dem Index des letzten Auftretens des Blockwertes in der Sequenz. Sollte trotz Initialisierungssequenz einmal ein Block noch nicht aufgetreten und somit seine Tabellenzelle leer sein, setzt man: $A_n = n$.

Der erhaltene empirische Wert für $f_U(R^N)$ läßt sich nun leicht quantitativ bewerten, indem man berechnet, wie wahrscheinlich es bei gültiger Nullhypothese[5] ist, daß sich dieser Wert ergibt. Ein Generator wird dann abgelehnt, wenn der empirisch gewonnene Wert für $f_U(R^N)$ außerhalb eines symmetrischen um den Erwartungswert liegenden Intervalls $[t_1, t_2]$ liegt, wobei die Schwellwerte t_1 und t_2 definiert werden durch:

$$t_1 = E[f_U(R^N)] - y \cdot \sigma \quad \text{und} \quad t_2 = E[f_U(R^N)] + y \cdot \sigma.$$

Hierbei ist y die Anzahl der Standardabweichungen σ die $f_U(R^N)$ maximal vom Erwartungswert entfernt liegen darf. Da es sich um einen zweiseitigen Test handelt, ist der Zusammenhang zwischen y und der Ablehnungsquote ρ gegeben durch $\mathcal{N}(-y) = \rho/2$, wobei $\mathcal{N}(x)$ für die Normalverteilung steht:

$$\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\eta^2/2} d\eta.$$

Als Werte für die Ablehnungsquote ρ empfiehlt Maurer $\rho \approx 0{,}001 \ldots 0{,}01$, je nach Anwendung des Generators, d.h. nur in einem Promille bzw. einem Prozent aller Testfälle, wird ein guter Generator zu Unrecht verworfen. Bei den später durchgeführten Tests wird eine leicht verändertes Vorgehen gewählt. Anstatt eine feste Ablehnungsquote von vornherein vorzugeben, wird für den jeweils erhaltenen empirischen Wert der Teststatistik berechnet, mit welcher Ablehnungsquote[6] er sich noch verwerfen ließe. Je größer der Wert ist, umso näher kommt ein Generator dem ideal zufälligen Verhalten.

E.3 Verbesserte Variante des universellen Tests von Coron

Jean-Sébastian Coron schlägt eine modifizierte Teststatistik [25] für den universellen Test vor, die gegenüber der ursprünglichen Teststatistik den Vorteil hat, daß ihr Erwartungswert bis auf statistische Schwankungen für *alle* Blocklängen der Entropie der Quelle entspricht und nicht erst im Grenzwert der Blocklänge gegen unendlich.
An der Testdurchführung selbst ändert sich nur wenig, lediglich die Software-Routinen für die Berechnung der Teststatistik müssen geringfügig modifiziert werden. Bei der Auswertung des Tests sind dementsprechend natürlich der Erwartungswert, die Varianz bzw. die modifizierte Varianz (s.o.) der neuen Teststatistik zu verwenden.
Die modifizierte Teststatistik unterscheidet sich von der ursprünglichen Maurerschen darin, daß an Stelle des Logarithmus zur Basis Zwei eine andere Funktion $g(j)$ verwendet wird:

$$f_U^g(s^N) = \frac{1}{K} \sum_{n=Q+1}^{Q+K} g\left(A_n(S^N)\right).$$

Diese Funktion $g(j)$ wird hierbei von Coron gerade so bestimmt, daß der Erwartungswert der neuen Teststatistik für ergodische, stationäre Quellen gerade gleich der Entropie eines L Bit langen Blockes ist, für $g(j)$ erhält man dann:

$$g(j) = \frac{1}{\ln(2)} \sum_{k=1}^{j-1} \frac{1}{k}.$$

[5] d.h. der Annahme, daß die Testsequenz von einer binären, symmetrischen Quelle ohne Gedächtnis erzeugt wurde.

[6] Es wird in diesem Fall also die Wahrscheinlichkeit berechnet, daß ein Wert der Teststatistik außerhalb eines symmetrischen *Intervalls* um den Erwartungswert der Statistik liegt, wobei die eine Grenze des Intervalls der erhaltene Testwert ist.

L	$\mathrm{Var}[g(A_(R^N))]$	$d(L)$	$e(L)$
3	2.5769918	0.3313257	0.4381809
4	2.9191004	0.3516506	0.4050170
5	3.1291382	0.3660832	0.3856668
6	3.2547450	0.3758725	0.3743782
7	3.3282150	0.3822459	0.3678269
8	3.3704039	0.3862500	0.3640569
9	3.3942629	0.3886906	0.3619091
10	3.4075860	0.3901408	0.3606982
11	3.4149476	0.3909846	0.3600222
12	3.4189794	0.3914671	0.3596484
13	3.4211711	0.3917390	0.3594433
14	3.4223549	0.3918905	0.3593316
15	3.4229908	0.3919740	0.3592712
16	3.4233308	0.3920198	0.3592384
∞	3.4237147	0.3920729	0.3592016

Tabelle E.2: $\mathrm{Var}[g(A_n(R^N)]$, $d(L)$ und $e(L)$ für $3 \leq L \leq 16$ und $L \to \infty$

Da die Berechnung von $g(j)$ für große j immer langwieriger wird, verwendet man für große j die sehr gute Approximation[7]:

$$\sum_{k=1}^{n} \frac{1}{k} = \ln(n) + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \mathcal{O}\left(\frac{1}{n^4}\right),$$

hierbei steht γ für die Eulersche Konstante:

$$\gamma = -\int_0^\infty e^{-x} \ln x \, dx \simeq 0.577216.$$

Durch diese Wahl für $g(j)$ erreicht man, daß der Erwartungswert der Teststatistik für eine binäre, symmetrische Quelle ohne Gedächtnis gerade den Wert $E[f_U^g(R^N)] = L$, also der beim Test verwendeten Blocklänge, annimmt; dieser Wert entspricht genau der Entropie einer ideal zufälligen Sequenz für L Bit lange Blöcke. Für eine binäre, *unsymmetrische* Quelle ohne Gedächtnis (BUQ) erhält man[8]:

$$E[f_U^g(R_{BUQ}^N)] = L \times H(p),$$

wobei $H(p)$ die Entropie pro Bit ist.
Um die Varianz der Teststatistik für die Nullhypothese einer binären, symmetrischen Quelle ohne Gedächtnis zu berechnen, muß wie beim ursprünglichen universellen Test auch hier die Varianz $\mathrm{Var}[f_U^g(R^N)]$ durch einen Wert $c_g(L,K)$ korrigiert werden:

$$\mathrm{Var}[f_U^g(R^N)] = c_g(L,K)^2 \times \frac{\mathrm{Var}\left[g\left(A_n(R^N)\right)\right]}{K}.$$

Bei einer binären, symmetrischen Quelle ohne Gedächtnis gilt:

$$\Pr[A_n(R^N) = j] = 2^{-L}(1-2^{-L})^{j-1}, \text{ für } j \geq 1,$$

[7] Bereits für $j \geq 23$ ist der Approximationsfehler kleiner als 10^{-8}.
[8] Für die Details der Herleitung, s. [25].

so daß man für die Varianz $\mathrm{Var}[g(A_n(R^N))]$ erhält:

$$\begin{aligned}\mathrm{Var}\left[g\left(A_n(R^N)\right)\right] &= E\left[\left(g\left(A_n(R^N)\right)\right)^2\right] - \left(E\left[g\left(A_n(R^N)\right)\right]\right)^2 \\ &= 2^{-L}\sum_{j=2}^{\infty}(1-2^{-L})^{j-1}\left(\sum_{k=1}^{j-1}\frac{1}{k\ln(2)}\right)^2 - L^2.\end{aligned}$$

Der Korrekturfaktor $c_g(L,K)$ läßt sich wiederum gut approximieren durch:

$$c_g(L,K)^2 = d(L) + \frac{e(L)\times 2^L}{K};$$

die entsprechenden angepaßten Werte für $\mathrm{Var}[g(A_n(R^N))]$, $d(L)$ und $e(L)$ sind in der [25] entnommenen Tabelle E.2 für verschiedene Blocklängen L aufgeführt. Die Berechnung der Schwellwerte, ab denen ein Generator verworfen wird, erfolgt wieder auf die in Abschnitt E.2 beschriebene Weise.
Zwar erlaubt die verbesserte Variante des universellen Tests die Berechnung des Erwartungswertes der Teststatistik für *unsymmetrische* Quellen, aber aufgrund der fehlenden Berechnungsmöglichkeit für die zugehörige Varianz ist es dennoch nicht möglich, diese Testvariante direkt zum Testen von Rohdaten zu verwenden.

Anhang F

Darstellung der Datenaufnahme-Parameter

Wird eine Bildschirmdarstellung der Datenaufnahme-Parameter gewählt, werden folgende Größen numerisch dargestellt:

- Anzahl der aufzunehmenden Dateien (typ. 1 MB groß, Dateigröße und Dateianzahl werden als Kommandozeilenoption beim Programmstart übergeben.)
- Anzahl der bisher aufgenommenen Dateien,
- Startdatum und Startzeit (bezogen auf die aktuelle Datei),
- Anzahl der Meßereignisse[1],
- Anzahl der bisher aufgenommenen Bits ohne Regularisierung (Meßereignisse abzüglich der Fehler),
- Anzahl der Fehler[2],
- Art der Datenerfassung[3]:
  - Aufnahme aller ankommenden Signale in einem Zweibitformat, d.h. für jedes Meßereignis werden zwei Bits verwendet, von denen das jeweilige Bit dann gesetzt wird, wenn der Pegel der zugehörigen Signalleitung auf HIGH liegt; es werden also auch Ereignisse mitaufgenommen, bei denen auf beiden Leitungen ein HIGH-Pegel vorliegt[4],
  - Aufnahme der Bits ohne Fehler, d.h. nur wenn *genau eine* der Signalleitungen HIGH ist, wird der entsprechende Wert (Null oder Eins) übernommen,
  - Aufnahme der Bits mit Von-Neumann-Regularisierung (s. 3.1.3)
- Verhältnis zwischen regularisierten und unregularisierten Bits
- Anzahl der in die Datei übernommenen Bits,

[1] Alle nun folgenden Größen beziehen sich immer auf die aktuelle Datei.

[2] Fehler sind die Ereignisse, bei denen ein Interrupt ausgelöst wurde, aber entweder auf der 1- und der 0-Leitung ein HIGH-Pegel oder auf beiden ein LOW-Pegel lag. Während der Verwendung dieser Datenaufnahmelektronik stellte es sich heraus, daß alle Fehler solche sind, bei denen zwei HIGH-Pegel auf beiden Leitungen registriert werden.

[3] Sie läßt sich beim Programmstart durch eine Kommandozeilenoption auswählen.

[4] Dieses Format wird nur zum Test der Signalverarbeitungselektronik verwendet.

- Anzahl der Nullen und der Einsen und das daraus berechnete Verhältnis der Nullen zu den Einsen,
- Anzahl der bisher gespeicherten Werte in Prozent der Dateigröße.

Anhang G

Mathematische Verfahren zur (Pseudo-)Regularisierung

G.1 Regularisierungsmethoden

Physikalische Zufallsgeneratoren erzeugen meist Binärfolgen, die von sich aus keine ideale Gleichverteilung der Nullen und Einsen aufweisen; würde eine ideale Gleichverteilung doch verlangen, daß die Statistik der Nullen und Einsen einer Binomialverteilung mit den Parametern $p = P(Bit = 1) = 0{,}5$ und $q = P(Bit = 0) = 1 - p$ mit beliebiger Genauigkeit folgte. Natürlich kann man versuchen, durch entsprechendes Einstellen der experimentellen Parameter des Zufallsgenerators möglichst nahe an das ideale Verhältnis der Nullen und Einsen von 50:50 heranzukommen. Wenn aber die zufällige Binärfolge nur lang genug ist, wird dennoch selbst die kleinste Abweichung von diesem Verhältnis feststellbar (s. Abschnitt 3.2.3). Daher ist es auch nicht sinnvoll, mit viel Aufwand ein möglichst nahe bei 50:50 liegendes Verhältnis erreichen zu wollen.
Benötigt man dennoch eine ideale Gleichverteilung von Nullen und Einsen, so geht man am besten folgendermaßen vor:

1. Man wählt die Parameter des Zufallsgenerators derart, daß man hinreichend nahe beim idealen Verhältnis der Nullen und Einsen von 50:50 liegt; meist ist dies mit vertretbarem Aufwand möglich.
2. Mit Hilfe einer sogenannten *Regularisierungsmethode* wird aus der vom Zufallsgenerator erzeugten Binärfolge eine ideale Ausgangsbitsequenz erzeugt.

Unter einer *Regularisierungsmethode* versteht man hierbei eine mathematische Prozedur, die es unter bestimmten Voraussetzungen erlaubt, aus einer tendenzbehafteten Bitfolge, die bei einigen Verfahren sogar Korrelationen zwischen den einzelnen Bits enthalten kann, eine Folge unabhängiger, gleichverteilter Bits[1] zu erzeugen. Regularisierungsmethoden werden meist in Software realisiert, wobei, je nach programmiertechnischem Aufwand, die entsprechende Regularisierungsroutine in einem nachgeschalteten Computer oder direkt in einem Mikrocontroller ausgeführt wird, der im Zufallsgenerator miteingebaut ist und auch gleichzeitig für andere Steueraufgaben verwendet werden kann. Einfache Regularisierungsmethoden lassen sich auch direkt in Hardware realisieren.
Regularisierungsmethoden lassen sich in echte Regularisierungsverfahren und in Pseudoregularisierungsverfahren (s. Abschnitt G.2) einteilen. Die echten Regularisierungsmethoden lassen sich

[1] Regularisierungsverfahren, die aus der Eingangsbitsequenz nicht gleichverteilte Zufallszahlen erzeugen, sondern solche, die einer anderen Statistik (z.B. einer Normalverteilung) gehorchen, werden hier nicht behandelt, s. dazu z.B. [52].

weiter grob nach den Binärfolgen klassifizieren, auf die sie erfolgreich angewandt werden können:

- Regularisierungsverfahren für tendenzbehaftete, aber *korrelationsfreie* Binärfolgen,
- Regularisierungsverfahren für tendenzbehaftete Binärfolgen *mit Korrelationen* zwischen aufeinander folgenden Bits.

Im folgenden werden die gängigsten Regularisierungsmethoden dargelegt und ihre Vor- und Nachteile erörtert. Hierbei wird der Schwerpunkt auf Regularisierungsmethoden für tendenzbehaftete, aber korrelationsfreie Binärfolgen gelegt, da diese für quantenoptische Zufallgeneratoren primär von Bedeutung sind. Verallgemeinerungen des jeweiligen Verfahrens für die Anwendung auf Binärfolgen mit (kurzreichweitigen) Korrelationen werden ebenfalls aufgeführt.
Auch wenn die einfachsten der weiter unten erläuterten Verfahren schon Jahrzehnte alt sind, so sind doch andere wiederum erst vor wenigen Jahren entwickelt worden und auch heute werden noch neue Regularisierungsalgorithmen vorgeschlagen. Als ein Beispiel für eine weitere, neue Variante sei auf die Arbeit von Juels et al. [64] verwiesen.
Bevor die einzelnen Verfahren besprochen werden, sei hier noch eine wichtige Kenngröße für die Effizienz eines Regularisierungsverfahrens definiert: Unter der *Effizienz* η eines Regularisierungsverfahrens versteht man den Quotienten aus dem Erwartungswert für die Anzahl der Ausgangsbits und der Anzahl der Eingangsbits im Grenzwert großer Eingangsbitmengen:

$$\eta = \lim_{n\to\infty} \frac{E[\#Ausgangsbits]}{\#Eingangsbits}.$$

Hierbei ist die obere Schranke für die Effizienz eines wie auch immer gearteten Regularisierungsverfahrens durch die Entropie pro Bit gegeben:

$$H_p = -p\log_2(p) - (1-p)\log_2(1-p), \text{ wobei wieder } p = P(Bit = 1) \text{ ist.}$$

Die folgenden Regularisierungsmethoden werden vorgestellt:

- *Von-Neumann-Regularisierung*[2] [96]: Dies ist das älteste und einfachste Regularisierungsverfahren, das aufgrund seines leicht zu implementierenden Algorithmus immer noch sehr gern verwendet wird. Es setzt voraus, daß die Bits des Eingangsbitstrom statistisch unabhängig voneinander sind. Für die Regularisierung wird der Eingangsbitstrom in nichtüberlappende, direkt aufeinanderfolgende Bitpaare eingeteilt, aus denen die Bits des Ausgangsbitstroms generiert werden.
- *Von-Neumann-Regularisierung für Eingangsbitfolgen mit zusätzlichen statistischen Abhängigkeiten* [112]: Diese Verallgemeinerung des Von-Neumann-Verfahrens erlaubt es, aus einem Eingangsbitstrom mit ungleichen relativen Häufigkeiten von Nullen und Einsen und statistischen Abhängigkeiten zwischen aufeinanderfolgenden Bits, eine Ausgangsbitfolge mit unabhängigen, gleichverteilten Bits zu erzeugen.
- *Verallgemeinerte Von-Neumann-Regularisierung nach Elias* [36]: Sie stellt eine Verallgemeinerung des Von-Neumann-Verfahrens dar, bei der nicht Bitpaare, sondern größere Bitblöcke betrachtet und diese nicht auf einzelne Bits, sondern auf Ausgangsbitblöcke abgebildet werden.
- *Verallgemeinerte Von-Neumann-Regularisierung nach Elias für Eingangsfolgen mit statistischen Abhängigkeiten*, dies ist die Verallgemeinerung des obengenannten Verfahrens auf Eingangsbitfolgen mit statistischen Abhängigkeiten zwischen aufeinanderfolgenden Bits.

[2] Damit der Anhang über Regularisierungsverfahren in sich abgeschlossen ist, wird die Ausgangsproblematik kurz erläutert und das Von-Neumann-Verfahren ebenfalls noch einmal dargelegt.

- *Iterierte Von-Neumann-Regularisierung nach Peres* [103]: Bei dieser Verallgemeinerung der Von-Neumann-Regularisierung wird die Effizienz der Regularisierung gegenüber der einfachen Von-Neumann-Regularisierung erhöht, indem nicht nur die Bitpaare des Eingangsstromes selbst, sondern zusätzlich aus diesen abgeleitete Bitströme einer Regularisierung unterworfen werden. Durch eine iterierte Vorgehensweise kann die Regularisierungseffizienz noch weiter gesteigert werden.

- *Iterierte Von-Neumann-Regularisierung nach Peres für Eingangsfolgen mit statistischen Abhängigkeiten* ist die Verallgemeinerung des oben genannten Verfahrens auf Eingangsbitfolgen mit statistischen Abhängigkeiten der Bits untereinander.

Bei den Verfahren, die eine Eliminierung der statistischen Abhängigkeiten der Bits erlauben, wird immer nur der für die Praxis besonders relevante Fall der Abhängigkeit vom direkten Vorgänger betrachtet, modelliert durch eine Markovkette; im Prinzip lassen sich auch kompliziertere Abhängigkeiten entfernen, allerdings nimmt die Regularisierungsrate hierbei ab und der Aufwand zu.

G.1.1 Von-Neumann-Regularisierung

Dieses Verfahren geht auf John von Neumann [96] zurück, der es im Rahmen seiner Arbeiten zur Erzeugung von Zufallszahlen bereits in den 50er Jahren entwickelte. Es ist das älteste Regularisierungsverfahren und damit quasi der „Urahn“[3] der meisten später entwickelten Verfahren. Da es gleichzeitig das einfachste aller Regularisierungsverfahren ist und sich daher mit geringem Aufwand sowohl in Soft- als auch in Hardware implementieren läßt, wird es auch heute noch häufig eingesetzt, um physikalisch erzeugte Binärfolgen zu regularisieren, s. z.B. [50, 110]. Diese Regularisierungsmethode läßt sich immer dann anwenden, wenn sichergestellt werden kann, daß aufeinanderfolgende Bits des zu regularisierenden Bitstroms statistisch unabhängig voneinander sind.
Bei der Von-Neumann-Regularisierung geht man folgendermaßen vor: Der Eingangsbitstrom wird überlappungsfrei und lückenlos in Paare aufeinanderfolgender Bits aufgeteilt. Aus diesen Paaren werden die Ausgangsbits gemäß folgender Abbildungsregel generiert:

Bit-Paar	Ausgangsbit
[11]	–
[10]	1
[01]	0
[00]	–

Aufgrund dieser Abbildungsvorschrift wird also nur dann ein Bit ausgegeben, wenn die Bits des Eingangsbitpaares ungleiche Werte haben. Man sieht leicht, daß die Null- und Eins-Werte bei den auf diese Weise erzeugten Bits tatsächlich gleichverteilt sind: Ist nämlich $p_E = P(Bit = 1)$ die Wahrscheinlichkeit, daß ein Bit des Eingangsbitstroms den Wert Eins hat und $q_E = P(Bit = 0) = 1 - p_E$ die Wahrscheinlichkeit, daß der Wert des Bits gleich Null ist, so gilt für die entsprechenden Wahrscheinlichkeiten der aus den Bitpaaren generierten Ausgangsbits:

$$p_A = P([10]) = p_E \cdot q_E = q_E \cdot p_E = P[01] = q_A$$

Somit ist die Wahrscheinlichkeit, mit der eine Eins ausgeben wird, gleich der Wahrscheinlichkeit, mit der eine Null erzeugt wird; eine ideale Gleichverteilung im Ausgangsbitstrom ist daher gegeben. Ein zusätzlicher Vorteil der Von-Neumann-Regularisierung besteht darin, daß es nicht

[3] In der Tat sind die meisten Regularisierungsverfahren Abwandlungen oder Verallgemeinerungen der Von-Neumann-Regularisierung.

notwendig ist, die Wahrscheinlichkeiten p bzw. q zu kennen, da sie an keiner Stelle des Verfahrens explizit benötigt werden. Lediglich schnelle Schwankungen der Wahrscheinlichkeiten auf Ebene der Bitpaare könnten dazu führen, daß die Regularisierung nicht ihren Zweck erfüllt; in der Praxis treten solch schnellen Schwankungen allerdings nicht auf.
Leider hat die Von-Neumann-Regularisierung neben dem Vorteil ihrer großen Einfachheit auch zwei Nachteile:

1. Die verwendete Abbildung geht sehr verschwenderisch mit den Zufallsbits um. Selbst wenn das Verhältnis der relativen Häufigkeiten von Nullen und Einsen in der Eingangsbitfolge bereits 50:50 beträgt, werden dennoch 75% der Eingangsbits verworfen! Weicht das Verhältnis vom Idealwert ab, so erhält man noch weniger Ausgangsbits. Die Regularisierungseffizienz beträgt bei einem Eingangsbitstrom mit der Wahrscheinlichkeit für eine Eins p_E lediglich:

   $$\eta_{vnr} = p_E \cdot q_E = p_E \cdot (1 - p_E),$$

   s. Abb. G.1, in der neben η_{vnr} zum Vergleich auch noch die Kurve für die Entropie pro Bit in Abhängigkeit von p_E eingezeichnet ist.

2. Es wird bei der Von-Neumann-Regularisierung vorausgesetzt, daß die zu verarbeitende Eingangsbitfolge aus statistisch unabhängigen Bits besteht. Falls Abhängigkeiten zwischen dem gerade generierten Bit und den ihm vorausgehenden Bits bestehen, wird keine ideale Zufallsfolge mehr generiert. Allerdings vermindert die Regularisierung dennoch die Korrelationen[4] zwischen den Bits, da im Schnitt (s.o.) drei von vier Bits des Eingangsbitstroms „verworfen“ werden; sie verbessert also auch in diesem Fall die Statistik.

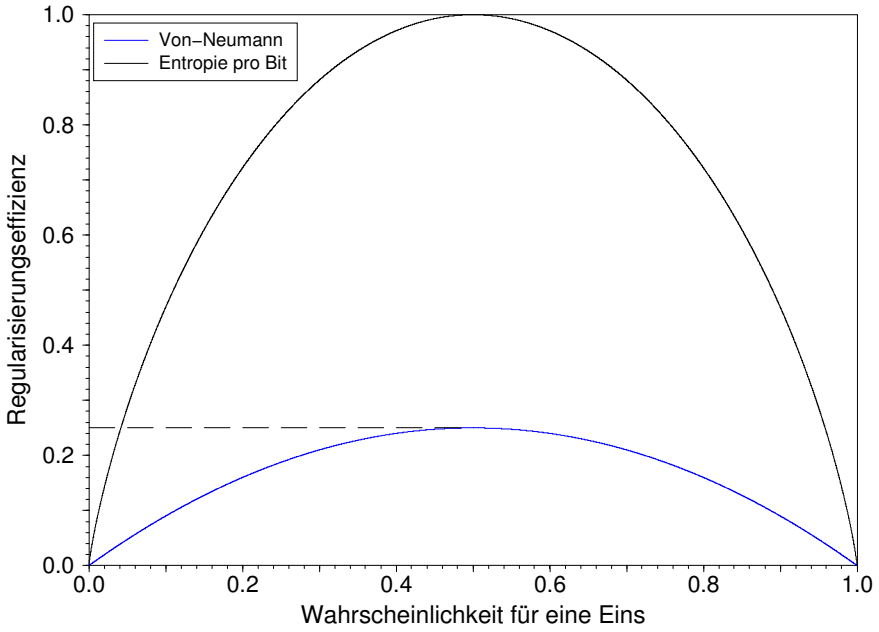


Abbildung G.1: Regularisierungseffizienz der Von-Neumann-Regularisierung in Abhängigkeit von der Wahrscheinlichkeit p_e für eine Eins im Eingangsbitstrom

Der zweite Nachteil spielt bei den meisten physikalischen Zufallsgeneratoren keine große Rolle, da sich durch entsprechende Einstellung der Generatorparameter Korrelationen zwischen auf-

[4] Wie in Abschnitt G.1.2 ausführlich erläutert wird, gilt dies für Korrelationen in stärkerem Maße als für Antikorrelationen.

einander folgenden Bits vermeiden lassen. Dennoch ist es im Hinblick auf alternative Ausführungen quantenoptischer oder sonstiger physikalischer Zufallsgeneratoren wünschenswert, eine Regularisierungsmethode zur Verfügung zu haben, die nicht nur Null- und Einshäufigkeiten einander angleicht, sondern überdies noch statistische Abhängigkeiten zwischen den generierten Bits zerstört. Solche Regularisierungsverfahren sind auch notwendig, um Korrelationssignaturen, s. Abschnitt 6.5.4.2, aus dem Zufallsstrom zu entfernen.
Das unter Punkt eins aufgeführte Effizienzproblem betrifft hingegen *jeden* physikalischen Zufallsgenerator, bei dem die Von-Neumann-Regularisierung verwendet wird. Das größtenteils unnötige Verwerfen von mindestens 75% der ursprünglich generierten Bits ist gerade bei physikalischen Zufallsgeneratoren mit ihren nicht allzu hohen Bitraten besonders ärgerlich, da es ihre Verwendung bei Aufgaben, die größere Mengen von Zufallsbits benötigen, erschwert und somit ihren möglichen Einsatzbereich weiter einschränkt.

G.1.2 Von-Neumann-Regularisierung antikorrelierter Rohdaten

Was passiert, wenn man die Von-Neumann-Regularisierung, verwendet, um (anti-) korrelationsbehaftete Rohdaten zu regularisieren? Zwar ist eine Voraussetzung für die Anwendung der Von-Neumann-Regularisierung die statistische Unabhängigkeit der Bits des zu regularisierenden Eingangsbitstroms, dennoch ist es durchaus praxisrelevant zu betrachten, wie sich die Von-Neumann-Regularisierung auf Eingangsbitströme auswirkt, bei denen diese Voraussetzung verletzt wird.

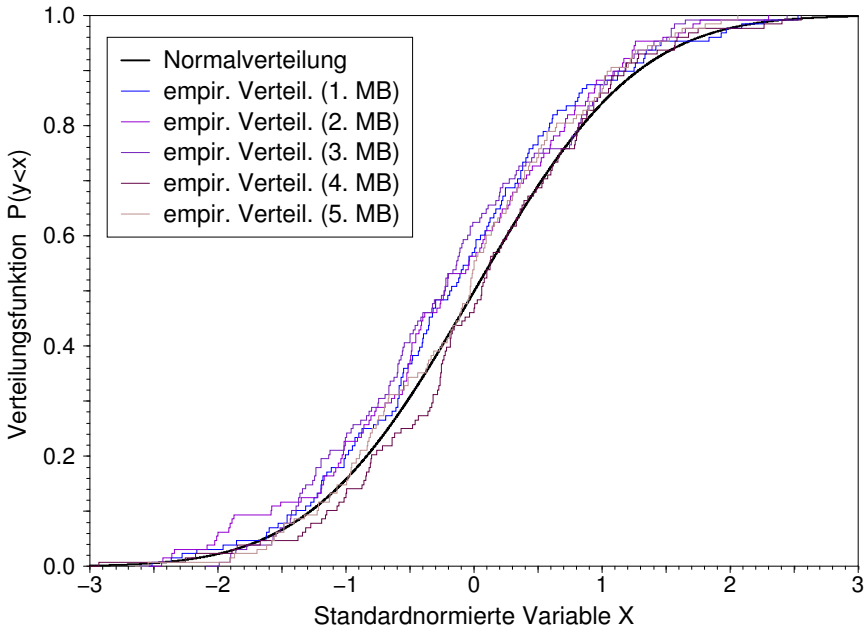


Abbildung G.2: Empirische Verteilungsfunktionen des ersten Autokorrelationskoeffizienten von fünf zweistufigen Tests auf von-Neumann-regularisierten Daten des Laufes *LD1*

Natürlich ist nicht zu erwarten, daß die regularisierten Daten ideale Eigenschaften haben werden, da hierfür die Korrelationsfreiheit eine notwendige Voraussetzung bei der einfachen Von-Neumann-Regulierung ist. Dennoch erwartet man auch in diesem Fall, daß das Verwerfen von 75 % der Bits sich in besseren statistischen Eigenschaften bemerkbar machen sollte. Im Folgenden wird exemplarisch untersucht, welche Ergebnisse zweistufige Autokorrelationskoeffizententests der von-Neumann-regularisierten Daten eines antikorrelationsbehafteten Laufes (*LD1*) liefern.
In Abb. G.2 sind die empirischen Verteilungen der ersten fünf zweistufigen Autokorrelationsko-

effiziententests für den ersten Autokorrelationskoeffizienten dargestellt. Die zugehörigen Werte[5] der Kolmogorov-Smirnov-Auswertung finden sich in Tabelle G.1.

Test Nr.	K_{128}^{+}	$P(K_{128}^{+})$	K_{128}^{-}	$P(K_{128}^{-})$
1	1.166488	0.93874	0.152902	0.05428
2	1.280734	0.96523	0.082885	0.01846
3	1.482878	0.98877	0.261144	0.14093
4	0.368717	0.25463	0.969320	0.85601
5	0.834415	0.76376	0.240357	0.12174
6	0.906193	0.81681	0.170724	0.06612
7	0.955336	0.84791	0.780419	0.71781

Tabelle G.1: Kolmogorov-Smirnov-Auswertung der empirischen Verteilungsfunktionen des ersten Autokorrelationskoeffizienten von sieben zweistufigen Tests der regularisierten Daten des Laufes *LD1*

Vergleicht man diese Kurven der Verteilungsfunktionen mit den Kurven für die Rohdaten in Abb. B.25 auf S. 179, so fallen zwei Merkmale auf:

1. Die Kurven liegen – wie erwartet – näher bei der theoretischen Verteilung. Insbesondere sind nicht *alle* Kurven zu einer Seite hin verschoben wie bei den Verteilungfunktionen aus den Tests der Rohdaten, s. Abschnitt B.9.3.

2. Die Verteilungsfunktionen sind gegenüber der theoretischen Verteilung meist zu negativen Abszissenwerten hin verschoben, also genau in entgegengesetzter Richtung wie die Verteilungsfunktionen der Rohdaten. Dies bedeutet, daß es eine Tendenz zu Korrelationen gibt.

Die beiden Merkmale spiegeln sich natürlich auch in den Werten der Kolmogorov-Smirnov-Auswertung wieder, wobei allerdings die Wahrscheinlichkeitswerte der einzelnen Verteilungen für sich genommen nur in zwei Fällen so groß sind, daß sie direkt „verdächtig“ erscheinen. Betrachtet man allerdings alle Werte zusammen, so fällt doch die in Abb. G.2 ebenfalls gut erkennbare Häufung der Kurven auf einer Seite der theoretischen Verteilung auf.
Wie läßt sich die „Umwandlung“ der Antikorrelationen innerhalb der Rohdaten in Korrelationen bei den regularisierten Daten erklären?
Hierzu ist es interessant zu bemerken, daß sich die Von-Neumann-Regularisierung durchaus nicht symmetrisch gegenüber den verschiedenen Abweichungen verhält. Dies sieht man besonders deutlich, wenn man sich die beiden möglichen Extremfälle der Abweichung – vollständige Korrelation bzw. Antikorrelation der Eingangsbits – betrachtet:

- *Bei vollständiger Korrelation* bestehen alle Eingangspaare der Von-Neumann-Regularisierung nur aus Nullen oder Einsen, was dazu führt, daß der Algorithmus keinerlei Bits ausgibt.

- *Bei vollständiger Antikorrelation*, d.h. einer alternierenden Sequenz von Nullen und Einsen, bestehen die Eingangspaare jeweils aus gleichartigen Paaren mit zwei unterschiedlichen Bitwerten. Je nach Anfangsstelle, an der Regularisierung beginnt, wird eine Ausgangsfolge generiert, die nur aus Nullen oder nur aus Einsen besteht, d.h. die Ausgangssequenz ist vollständig korreliert.

[5] Die 30 MB Rohdaten liefern lediglich 7 MB regularisierte Daten, in der Tabelle werden daher gleich die Werte aller sieben Tests aufgelistet.

Der erste Fall ist nicht so abwegig, wie man auf den ersten Blick meinen könnte, so kann er durchaus bei einem massiven Defekt des Generators, wie z.B. dem vollständigen Ausfall eines der Signaldetektoren, auftreten. In diesem Fall verhindert aber die Von-Neumann-Regularisierung die Ausgabe eines fehlerhaften Zufallsstroms, was natürlich sehr wünschenswert ist.
Der zweite Fall ist in diesem Extrem allerdings tatsächlich sehr unwahrscheinlich. Dennoch verdeutlicht er sehr gut, warum sich bei dem zweistufigen Autokorrelationskoeffizententests von-Neumann-regularisierter Stichproben schwach antikorrelationsbehafteter Daten Korrelationen zeigen.
Als Fazit kann man somit festhalten, daß Antikorrelationen des Eingangsbitstromes sehr viel kritischer für eine nachfolgende Von-Neumann-Regularisierung sind als Korrelationen. Während man sich darauf verlassen kann, daß eine nachfolgende Von-Neumann-Regularisierung Korrelationen grundsätzlich stark abschwächt, schwächt sie zwar Korrelationseffekte bei Antikorrelationen der Rohdaten in den regularisierten Daten ab, aber in geringerem Maße. Will man den Einfluß von Antikorrelationen im Eingangsbitstrom auf die statistischen Eigenschaften des Ausgangsbitstroms eliminieren, so sollten dediziert für (anti-)korrelationsbehaftete Eingangsbitstöme entworfende Regularisierungsverfahren, s. z.B. Abschnitt G.1.4, eingesetzt werden.

G.1.3 Markovketten

Markovketten stellen die einfachsten Modelle für einen Zufallsprozeß dar, bei dem die generierten Bits nicht unabhängig voneinander sind. Bei Markovketten hängt der Ausgang des Zufallsexperimentes zur Generierung eines Bits statistisch noch zusätzlich vom Ergebnis des vorhergehenden Experimentes ab, allerdings auch nur von diesem.
Für die folgenden Verfahren wird ein Spezialfall der Markovkette betrachtet: die zeit- und zustandsdiskrete, homogene Markovkette. *Zustandsdiskret*, da ein Bit nur die beiden Zustände „**0**“ und „**1**“ hat, *zeitdiskret*[6], da lediglich die Position des Bits in der generierten Sequenz betrachtet wird und nicht der genaue Zeitpunkt der Bitgenerierung und *homogen*, da die Übergangswahrscheinlichkeiten von Zustand zu Zustand nicht von der absoluten Position innerhalb der Zufallssequenz abhängen (sollen).
Natürlich stellen diese zusätzlichen Modellannahmen eine Vereinfachung der tatsächlichen Verhältnisse dar. Vollkommen unproblematisch ist bei einem quantenoptischen Zufallsgenerator die Diskretisierung der Zustände, da sowohl Bits als auch Photonen tatsächlich diskret sind. Die Annahme einer zeitdiskreten Reihenfolge der Zustände bzw. Bits scheint zwar auf den ersten Blick unbedenklich, da die generierten Bitfolgen keinerlei Information über die zeitliche Abfolge der Bits enthalten, für den eigentlichen Generationsvorgang muß dies allerdings nicht gelten, da sich z.B. Detektortotzeiten auf kleinen Zeitskalen auf die Übergangswahrscheinlichkeiten auswirken könnten.
Homogenität ist dagegen eher auf sehr großen Skalen problematisch, da z.B. temperaturbedingte Drifterscheinungen die Zustandswahrscheinlichkeiten für Einsen und Nullen ändern können. Auf kleinen Skalen hingegen ist Homogenität gegeben und genau diese Homogenität auf kleinen Skalen ist es, die bei Regularisierungsverfahren meist gefordert wird.

[6] Zeitdiskret trifft es eigentlich nicht ganz, da die Zeit bis auf die Reihenfolge vollständig vernachlässigt wird, besser: „zeitvernachlässigend“.

Die relevanten Größen für die betrachteten zweigliedrigen Markovketten sind in Abb. G.3 aufgeführt, hierbei werden die Zustandswahrscheinlichkeiten mit Großbuchstaben und die Übergangswahrscheinlichkeiten mit Kleinbuchstaben bezeichnet.

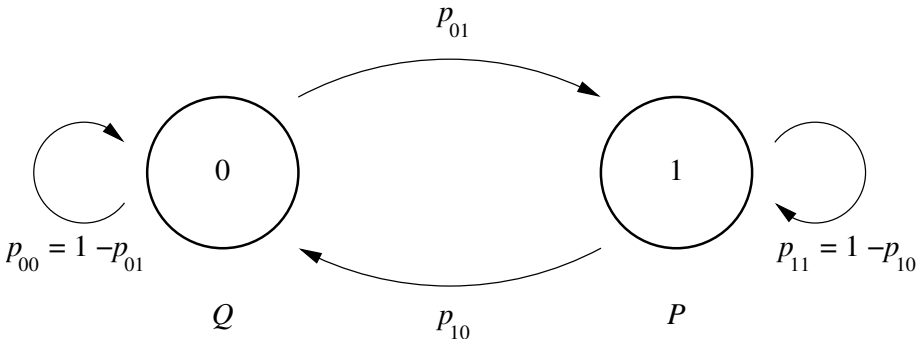


Abbildung G.3: Wahrscheinlichkeiten bei der Markovkette

G.1.4 Von-Neumann-Regularisierung für statistisch abhängige Bits

Eine von Samuelson [112] und Pratt[7] vorgeschlagene Verallgemeinerung des Von-Neumann-Verfahrens erlaubt die Regularisierung von Zufallssequenzen mit statistischen Abhängigkeiten. In erster Linie[8] ist das Verfahren dazu geeignet, einen häufig auftretenden Defekt physikalischer Zufallsgeneratoren zu beheben: die Abhängigkeit der Generierungswahrscheinlichkeit vom Wert des vorangehenden Bits. Da das Verfahren sich auf die Von-Neumann-Regularisierung stützt, teilt es auch dessen Vor- und Nachteile: gute Implementierbarkeit, bei geringer Regularisierungseffizienz.
Bei der hier vorgestellten Variante der Regularisierungsmethode wird davon ausgegangen, daß die Werte der Eingangsbits vom Wert des jeweils vorangehenden Bits abhängig sein können, d.h.:

$$\begin{aligned} P(B_i = 1|B_{i-1} = 1) = p_{11} \quad &\neq \quad p_{01} = P(B_i = 1|B_{i-1} = 0), \text{ bzw.} \\ P(B_i = 0|B_{i-1} = 1) = p_{10} \quad &\neq \quad p_{00} = P(B_i = 0|B_{i-1} = 0). \end{aligned}$$

Die Grundidee des Verfahrens ist recht einfach: Man wende die Von-Neuman-Regularisierung einfach zweimal hintereinander an. Dies geschieht hierbei auf folgende Weise: Zuerst teilt man wieder die Eingangsbitfolge in nicht überlappende Paare und erzeugt mit Hilfe folgender Abbildung:

Paare	vorläufiges Ausgangsbit
[10]	0
[11]	1
[01]	–
[00]	–

eine Zwischen-Bitsequenz. Diese Sequenz besteht bereits aus unabhängigen Bits, da bei einer Markovkette ein Bit nur von seinem direkten Vorgänger beeinflußt wird und die Paare, welche

[7] Die an dieser Stelle dargestellte Variante geht auf einen in [112] zitierten Vorschlag von John W. Pratt zurück.

[8] Das Verfahren läßt sich im Prinzip derart modifizieren, daß auch statistische Abhängigkeiten über größere Reichweiten bei der Regularisierung aufgehoben werden, allerdings nimmt dann die Regularisierungsrate drastisch ab.

die Ausgangsbits erzeugen, mit demselben Startzustand (einer Eins) anfangen. Die Übergangswahrscheinlichkeiten vom ersten zum zweiten Bit (das bestimmt, ob eine Eins oder eine Null ausgegeben wird) haben in diesem Fall somit die festen Werte p_{10} für eine Null bzw. p_{11} für eine Eins.
Da i.a. $p_{10} \neq p_{11}$ sein wird, sind die relativen Häufigkeiten der Einsen und Nullen dieser Zwischen-Sequenz allerdings noch nicht gleich, so daß noch eine herkömmliche Von-Neumann-Regularisierung der Bits durchgeführt werden muß, um schlußendlich eine Sequenz unabhängiger und gleichverteilter Bits zu erhalten.
Aufgrund der sequentiellen Verarbeitung der Bits und der einfachen Regularisierungsvorschriften, läßt sich dieses Verfahren gut in Soft- oder Hardware implementieren. Allerdings geht es mit den Eingangsbits noch verschwenderischer als das Von-Neumann-Verfahren um: Selbst im günstigsten Fall bleibt nach jeder der Abbildungen höchstens 1/4 der Eingangsbitrate übrig, so daß die letztendliche Ausgangsbitrate auf 1/16 oder noch weniger der ursprünglichen Rate reduziert wird.
Abschließend sei aber noch erwähnt, daß sich das Verfahren effektiver gestalten läßt: Verwendet man statt der ersten Von-Neumann-Regularisierung eine etwas trickreichere und nur wenig aufwendigere Vorgehensweise, s. Abschnitt G.1.6, so kann man wieder die Werte für die Regularisierungsrate der einfachen Von-Neumann-Regularisierung erzielen!

G.1.5 Regularisierung nach Elias

Eine einfachere Regularisierung als die Von-Neumann-Methode läßt sich nicht finden, allerdings kann man sich die Frage stellen, ob es nicht eine Regularisierungsmethode gibt, die effizienter ist als sie – wenn auch auf Kosten einer höheren Komplexität des Verfahrens.
Peter Elias [36] hat genau dies getan und eine ganze Familie von Regularisierungsverfahren ersonnen, um aus einer Sequenz unabhängiger, aber nicht gleichverteilter Bits eine Sequenz mit gleichverteilten Bits zu erzeugen. Hierbei stellt die herkömmliche Von-Neumann-Regularisierung einen Spezialfall der von ihm entwickelten allgemeineren Prozeduren dar.
Die Grundidee der von Elias vorgeschlagene Verfahren besteht darin, daß man sich nicht mehr wie bei der Von-Neumann-Regularisierung auf Paare beschränkt, sondern größere Bitblöcke betrachtet und diesen Blöcken auch nicht einzelne Bits, sondern gleich ganze Bitblöcke zuweist. Elias gelingt es zu zeigen, daß das Verfahren im Limes großer Blockgrößen den theoretischen Grenzwert für die Effizienz, die Entropie pro Bit, erreicht. Überdies gibt er für den praktisch relevanten Fall fester Blocklänge N explizite Ausdrücke für die obere und untere Schranke der Effizienz des Verfahrens an. Generell gilt für die Effizienz η_N eines auf Blöcken arbeitenden Regularisierungsverfahrens, daß:

$$\eta_N \leq \frac{H(X^N)}{N} \text{ ist,}$$

wobei $H()$ die Berechnungsfunktion für die Entropie ist und X^N für die N-Bit langen Eingangsblöcke steht. Bestehen die Blöcke aus statistisch unabhängigen Bits, so gilt $H(X^N) = N \cdot H(X^1)$ und es folgt, daß die Effizienz höchstens so groß sein kann wie die Entropie pro Bit, die nur im Falle einer Gleichverteilung der Bitwerte $P(X = 0) := q = P(X = 1) := p = 0{,}5$ ihren Maximalwert von Eins annimmt.
Elias gibt nun für die Effizienz seiner blockbasierten Regularisierungsverfahren folgende Abschätzung an:

$$\sum_{k=0}^{N} \binom{N}{k} p^k q^{N-k} \frac{\log_2 \binom{N}{k}}{N} \geq \eta_N(p) \geq \sum_{k=0}^{N} \binom{N}{k} p^k q^{N-k} \frac{\log_2 \binom{N}{k}}{N} - \frac{3}{N}.$$

In Abb. G.4 sind für verschiedene Wahrscheinlichkeiten eine Eins zu erhalten ($p = 0.5, 0.6, 0.7$),

die unteren[9] und oberen Schranken der Effizienz des Verfahrens in Abhängigkeit von der Blocklänge aufgetragen[10].

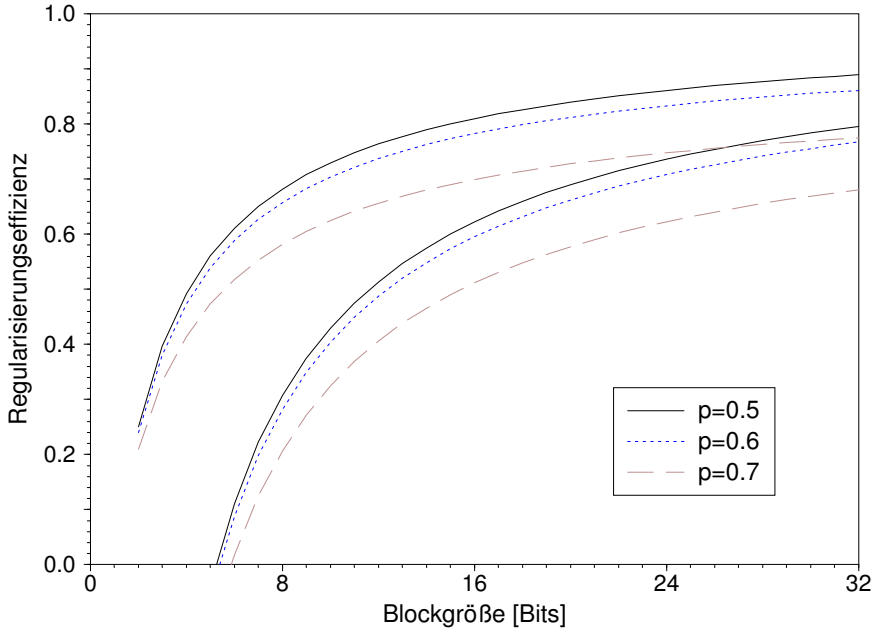


Abbildung G.4: Effizienz der Elias-Regularisierung in Abhängigkeit von der Blocklänge für verschiedene Einswahrscheinlichkeiten p

Den Gewinn gegenüber der Von-Neumann-Regularisierung sieht man sofort, indem man ihre Effizienz, d.h. die Effizienz bei der Blockgröße $N = 2$, mit der Effizienz bei größeren Blocklängen vergleicht.
Das eigentliche Verfahren funktioniert nun folgendermaßen:
Ähnlich wie beim Von-Neumann-Verfahren wird die Eingangssequenz in nicht überlappende, aufeinander folgende Blöcke eingeteilt, nur daß diese nun eine Größe $N \geq 2$ haben, wobei N möglichst groß gewählt werden sollte. Allerdings ist die Zuweisung der Ausgangs- zu den Eingangsblöcken nicht mehr so einfach wie beim Von-Neumann-Verfahren, sondern bedarf eines größeren Aufwandes:

1. Die Menge aller möglichen Blöcke der Länge N wird nach der Anzahl der Einsen (k) innerhalb der Blöcke in Klassen eingeteilt. Hierbei enthält jede der $(N + 1)$ Klassen S_k genau $\binom{N}{k}$ Blöcke.
2. Für jede dieser Klassen wird nun die Binärdarstellung des zugehörigen Binomialkoeffizienten $\binom{N}{k}$ dazu benutzt, um eine Zuordnung zwischen Eingangs- und Ausgangsblöcken aufzustellen:

$$\binom{N}{k} = \alpha_n 2^n + \alpha_{n-1} 2^{n-1} \cdots + \alpha_0 2^0,$$

$$\text{wobei} \quad n_{(k)} = \lfloor \log_2 \binom{N}{k} \rfloor \text{ ist.}$$

[9] Die untere Schranke ist für kleine Blocklängen negativ und daher natürlich nicht zutreffend. Das Von-Neumann-Verfahren liegt in der Praxis immer in der Nähe der oberen Schranke.

[10] Die Kurven sind hierbei lediglich der Übersichtlichkeit wegen durchgezogen, die Effizienz ist natürlich nur für ganze Blöcke definiert. Die Werte für p liegen bei physikalischen Zufallsgeneratoren i.a. nahe bei 1/2, die zugehörigen Schranken liegen in diesen Fällen so nahe bei denen für $p = 0{,}5$, daß sie in Abb. G.4 nicht extraeingezeichnet werden.

Für jeden nichtverschwindenden Koeffizienten $\alpha_j \in \{\alpha_n \cdots \alpha_0\}$, mit $0 \leq j \leq n$, werden die 2^j möglichen Ausgabeblöcke der Länge j den 2^j unterschiedlichen Sequenzen aus S_k zugewiesen, die noch nicht angewiesen sind. Hierbei wird jeweils einem Block der Klasse S_k für $\alpha_0 = 1$ der „leere Ausgabeblock“ zugewiesen, d.h. in diesem Fall wird kein Block ausgegeben. Die Klassen S_0 und S_N haben nur jeweils ein Element; diesem wird immer der „leere Ausgabeblock“ zugewiesen.

3. Indem man die obigen Zuordnungen für alle Klassen S_k mit $0 \leq k \leq N$ vornimmt, hat man die vollständige Regularisierungsprozedur angegeben.

Zur Illustration dieser etwas unübersichtlichen Regularisierungsvorschrift sei hier die Zuordnung anhand einer Blockgröße von $N = 4$ verdeutlicht:

Klasse	Anzahl der Einsen	Anzahl der Elemente: $\binom{N}{k}$	Binärdarstellung von $\binom{N}{k}$	$n_{(k)}$
S_0	0	1	1	0
S_1	1	4	100	2
S_2	2	6	110	2
S_3	3	4	100	2
S_4	4	1	1	0

In diesem Fall hat man $2^4 = 16$ mögliche Eingangsblöcke, die sich nach der Anzahl ihrer Einsen in $(N + 1 = 5)$ Klassen S_k einteilen lassen.
Die Klassen S_0 und S_4 werden jeweils auf den leeren Ausgabeblock abgebildet, d.h. die Eingangsblöcke [0000] und [1111] liefern keine Ausgabe.
Die Zuordnung der Eingangsblöcke zu den Ausgangsblöcken sei am Beispiel der Blöcke der Klasse S_1 dargestellt, die folgende Eingangsblöcke umfaßt (in lexikalischer Ordnung): [0001],[0010],[0100],[1000]. In der Binärdarstellung von $\binom{N}{k}$ ist nur für $j = 2$ der Koeffizient $\alpha_j \neq 0$, so daß die vier Eingangsblöcke den $2^j = 2^2 = 4$ möglichen Ausgangsblöcken [00],[01],[10],[11] zugeordnet werden. Analog geht man für die anderen Klassen vor. So hat man z.B. für S_2 eine Zuordnung der 6 möglichen Eingangsblöcke zu den Ausgangsblöcken [00],[01],[10],[11],[0],[1].
Die Zuweisung der Eingangsblöcke zu den Ausgangsblöcken ist wie bei der Von-Neumann-Regularisierung deterministisch; welchen der Eingangsblöcke man welchem Ausgangsblock zuweist, das ist beliebig und kann daher mit Blick auf Laufzeit- oder Speicherplatzoptimierung geschehen.

G.1.6 Elias-Regularisierung bei statistisch abhängigen Bits

Das Regularisierungsverfahren von Elias läßt sich im Prinzip auch auf abhängige Bits anwenden[11]. Wie schon beim Verfahren von Samuelson und Pratt besteht die Grundidee darin, in einem ersten Schritt aus der Eingangsbitfolge mit statistischen Abhängigkeiten eine Bitsequenz zu generieren, die zwar noch tendenzbehaftet ist, aber keine statistischen Abhängigkeiten zwischen den Bits mehr aufweist. Aus ihr wird anschließend in einem zweiten Schritt eine gleichverteilte Ausgabesequenz erzeugt; für diesen zweiten Schritt wird jetzt allerdings das Regularisierungsverfahren von Elias verwendet.
Für den ersten Schritt schlägt Elias folgende Vorgehensweise vor:

1. Man nimmt das erste Bit x_1 der Eingangssequenz, um festzulegen, aus welchem Zustand heraus das nächste Bit x_2 erzeugt wird.

2. Die restliche Eingangssequenz $x_2x_3 \ldots$ wird nach folgender Regel in zwei Teilsequenzen S_0 und S_1 zerlegt:

[11] Auch hier wird wieder nur die statistische Abhängigkeit des generierten Bits vom direkten Vorgänger behandelt.

S_0**:** enthält alle x_m, für die $x_{m-1} = 0$ ist, aufgereiht in aufsteigender Reihenfolge,
S_1**:** enthält alle x_m, für die $x_{m-1} = 1$ ist, aufgereiht in aufsteigender Reihenfolge.

Diese beiden Teilsequenzen sind nun Sequenzen mit unabhängigen Bits, da die Bits des jeweiligen Prozesses immer aus nur einem Zustand heraus generiert wurden, d.h. die Wahrscheinlichkeiten, eine Null bzw. eine Eins zu erhalten, haben im Falle von S_0 gerade die festen Werte p_{00} und p_{01} und bei S_1 die Werte p_{10} und p_{11}.
Nach Anwendung des Regularisierungsverfahren auf die beiden Teilsequenzen werden dann einfach die beiden zugehörigen Teilausgangssequenzen zu einer gemeinsamen Ausgangssequenz verknüpft, ob dies durch einfache Konkatenation der Teilausgangssequenzen oder durch bitweises Abwechseln zwischen den beiden Sequenzen geschieht, ist dabei unerheblich.

G.1.7 Iterierte Von-Neumann-Regularisierung

Bei dieser von Yuval Peres [103] vorgeschlagenen Regularisierungsmethode handelt es sich ebenfalls um eine Verallgemeinerung der Von-Neumann-Regularisierung. Der Grundgedanke des Verfahrens besteht darin, nicht nur die Bitpaare des Eingangsbitstroms mit unterschiedlichen Bitwerten zur Erzeugung der Ausgangsbits zu benutzen, sondern auch noch zwei weitere Bitströme, die aus den bei der Von-Neumann-Regularisierung verworfenen Paaren bzw. aus der Paarstatistik abgeleitet werden. Diese Grundidee läßt sich überdies iterativ anwenden und liefert somit eine ganze Familie von Regularisierungsverfahren.
Wie bei der Von-Neumann-Regularisierung wird auch hier vorausgesetzt, daß es keine Korrelationen zwischen den Bits des Eingangsstromes gibt. Peres beweist für den Fall, daß die m Bits eines Eingangsbitblocks, auf den das Verfahren angewandt wird, vertauschbare Zufallsvariablen[12] sind, die Gleichverteilung des regularisierten, kürzeren, k Bit langen Ausgangsbitblocks im Raum $[0,1]^k$.
Die iterierten Von-Neumann-Regularisierungen erhält man nun auf folgende Weise: Die herkömmliche Von-Neumann-Regularisierung stellt die iterierte erster Ordnung Ψ_1 dar und wird hier allerdings folgendermaßen definiert:

$$\Psi_1(x_1, x_2, \ldots, x_{2n+1}) = \Psi_1(x_1, x_2, \ldots, x_{2n}) = (y_1, y_2, \ldots, y_k), \text{ wobei}$$

$y_i = x_{2m_i}$ und $m_1 < m_2 < \ldots < m_k$ all diejenigen Indices $m \leq n$ sind, für die gilt $x_{2m} \neq x_{2m-1}$; bei einer ungeraden Anzahl von Bits im Eingangsbitstrom wird das letzte Bit einfach weggelassen. Bei dieser Definition wird außerdem jeweils der Wert des zweiten Bits[13] eines Bitpaares mit ungleichen Bitwerten als Ausgangsbitwert verwendet, was natürlich keinen prinzipiellen Unterschied zur bisherigen Definition darstellt.
Die oben bereits erwähnten abgeleiteten Bitströme beschafft man sich folgendermaßen: Sind die zu bearbeitenden Eingangsbits gegeben durch $x_1, \ldots, x_{2n}$ so erzeugt man einen zusätzlichen Bitstrom halber Länge durch eine XOR-Verknüpfung aufeinander folgender Bits (ohne Überlapp): $u_j = x_{2j-1} \oplus x_{2j}$. Der andere Bitstrom wird aus den bei der einfachen Von-Neumann-Regularisierung verworfenen Bitpaaren $v_j = x_{2i_j}$ gebildet, wobei $i_1 < i_2 < \ldots < i_{n-k}$ diejenigen Indices $i \leq n$ sind, bei denen die zugehörigen Bitpaare gleiche Bitwerte haben, d.h. $x_{2i} = x_{2i-1}$ ist, und k die Anzahl der Indices mit $m \leq n$ darstellt, für die gilt: $x_{2i} \neq x_{2i-1}$. Für ein Bitpaar [11] wird also ein Wert 1 in diesen Bitstrom geschrieben und für ein Bitpaar [00] ein Wert 0. Die Länge dieses Bitstroms ist natürlich variabel, da sie von der Anzahl der „verworfenen“ Bitpaare abhängt; sie ist aber immer kleiner als die halbe Länge des Eingangsbitstroms.
Die ν-te iterierte Von-Neumann-Regularisierung wird nun auf rekursive Weise folgendermaßen definiert:

$$\Psi_\nu(x_1, x_2, \ldots, x_{2n}) = \Psi_1(x_1, x_2, \ldots, x_{2n}) * \Psi_{\nu-1}(u_1, \ldots, u_n) * \Psi_{\nu-1}(v_1, \ldots, v_{n-k}),$$

[12] Eine Anzahl von n Zufallsvariablen heißt vertauschbar, wenn ihre $n!$ Permutationen alle dieselbe n-dimensionale Wahrscheinlichkeitsdichte haben, s. [43], VII.4

[13] Bei der übliche Definition der Von-Neumann-Regularisierung wird es i.a. umgekehrt gehandhabt.

wobei $*$ für die Konkatenationsoperation für Bitströme steht. Für Bitströme mit einer ungeraden Anzahl von Bits, wird hierbei ebenfalls einfach das letzte Bit nicht berücksichtigt, d.h. die ν-te iterierte Von-Neumann-Regularisierung wird definiert gemäß:

$$\Psi_\nu(x_1, x_2, \ldots, x_{2n+1}) = \Psi_\nu(x_1, x_2, \ldots, x_{2n}).$$

Zur Illustration sei die iterierte Von-Neumann-Regularisierung anhand der nach der einfachen Von-Neumann-Regularisierung $\Psi_1(x_1, x_2, \ldots, x_{2n})$ folgenden zweiten Iterierten $\Psi_2(x_1, x_2, \ldots, x_{2n})$ kurz erläutert. Hierbei wird der Gesamtausgangsbitstrom durch die Konkatenation der Ausgangsbitströme von drei Von-Neumann-Regularisierungen erzeugt, nämlich: der einfachen Von-Neumann-Regularisierung auf den Bitpaaren des Eingangsbitstroms (erster Term), einer Von-Neumann-Regularisierung auf den durch Exklusiv-Oder-Verknüpfung erzeugten Bits (zweiter Term) und einer Von-Neumann-Regularisierung auf den einfachen Bitwerten der verworfenen Bitpaare der ersten Von-Neumann-Regularisierung (dritter Term):

$$\Psi_2(x_1, x_2, \ldots, x_{2n}) = \Psi_1(x_1, x_2, \ldots, x_{2n}) * \Psi_1(u_1, \ldots, u_n) * \Psi_1(v_1, \ldots, v_{n-k}).$$

Peres gibt die Regularisierungseffizienz $\eta_\nu(p)$ der ν-ten iterierten Von-Neumann-Regularisierung in Abhängigkeit von der Wahrscheinlichkeit p für eine Eins an; die ebenfalls rekursive Formel[14] lautet:

$$\eta_\nu(p) = p \cdot q + \frac{1}{2}\,\eta_{\nu-1}(p^2 + q^2) + \frac{1}{2}\,(p^2 + q^2)\,\eta_{\nu-1}\left(\frac{p^2}{p^2 + q^2}\right),$$

wobei $\eta_1(p) = p \cdot q$ die übliche Rate für die einfache Von-Neumann-Regularisierung ist. In Abb. G.5 ist die jeweilige Regularisierungseffizienz für verschiedene Werte von p in Abhängigkeit von der Iterationsordnung aufgetragen[15].
Wieviel effizienter iterierte Von-Neumann-Regularisierungen im Vergleich zur einfachen Von-Neumann-Regularisierung sind, läßt sich bereits aus dem Vergleich der Effizienz $\eta_1(p)$ der herkömmlichen Von-Neumann-Regularisierung mit der Effizienz $\eta_2(p)$ der zweiten iterierten Von-Neumann-Regularisierung ersehen: Bei einer Wahrscheinlichkeit $p = 0{,}5$ für eine Eins im Eingangsbitstrom beträgt die (maximale) Regularisierungseffizienz $\eta_1(0{,}5) = 0{,}25$, hingegen hat $\eta_2(p)$ bereits den Wert $\eta_2(0{,}5) = 0{,}4375$. Somit läßt sich selbst mit relativ geringem programmtechnischen Aufwand die mittlere Regularisierungsausbeute um 75% gegenüber der einfachen Von-Neumann-Regularisierung steigern!
Durch Einsatz iterierter Von-Neumann-Regularisierungen höherer Ordnung läßt sich diese Regularisierungsausbeute sogar noch weiter erhöhen, wobei allerdings der zusätzliche Gewinn mit steigender Ordnung abnimmt. Für den Grenzfall, daß die Ordnung gegen unendlich geht, kann Peres sogar beweisen, daß die Regularisierungseffizienz gleichmäßig gegen die theoretische Obergrenze, die Entropie pro Bit, konvergiert. Glücklicherweise konvergiert die Regularisierungseffizienz mit steigender Iterationsordnung ν recht schnell gegen diesen Wert: Für $p = 0{,}5$ beträgt die Regularisierungseffizienz der achten Iterierten bereits $\eta_8(p) \approx 0{,}9$, was dem theoretischen Maximalwert der Entropie pro Bit von $H(p = 0{,}5) = 1{,}0$ bereits sehr nahe kommt.
In der Praxis gilt es abzuwägen, wieviel Aufwand man treiben will, d.h. welche Iterationsordnung man wählt, um eine hohe Regularisierungsrate zu erreichen. Da die Regularisierung typischerweise in Echtzeit erfolgt und der Prozessor auch noch andere Aufgaben ausführen muß, wie z. B. die Datenübernahme vom Zufallsgenerator, ist bei der Wahl der Iterationsordnung, sowohl die Rechenleistung des Prozessors als auch seine Belastung durch andere Prozesse zu berücksichtigen. Soll die Regularisierung durch einen im Zufallsgenerator eingebauten Mikrocontroller erfolgen, so muß man sich aufgrund der niedrigeren Rechenleistung und des geringeren zur Verfügung stehenden Speichers auf niedrige Iterationsordnungen beschränken.

[14] Die Formel für die Effizienz gilt im Grenzwert einer unendlichen Anzahl von Eingangsbit.

[15] Die Kurven sind hierbei lediglich der Übersichtlichkeit wegen durchgezogen, die Regularisierungsrate ist natürlich wieder nur für ganze Iterationsstufen definiert.

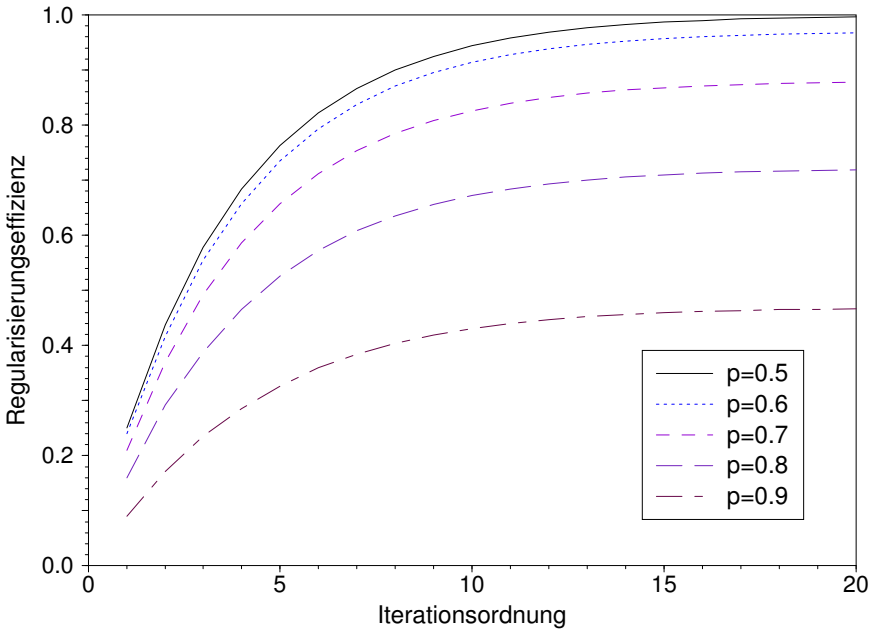


Abbildung G.5: Regularisierungseffizienz der iterierten Von-Neumann-Regularisierung in Abhängigkeit von der Iterationsordnung ν für verschiedene Werte der Einswahrscheinlichkeit p im Eingangsbitstrom

Bei der Implementation einer iterierten Von-Neumann-Regularisierung empfiehlt es sich, den Eingangsbitstrom in Bitblöcke fester Länge b zu unterteilen, die dann nacheinander durch die entsprechende Regularisierungsvorschrift Ψ bearbeitet werden und deren Ausgangsbitblöcke aneinander gehängt werden:

$$\Phi(x_1, \ldots, x_b) * \Phi(x_{b+1}, \ldots, x_{2b}) * \ldots .$$

Hierbei darf die Länge b der Blöcke einerseits nicht zu kurz sein, da sonst die Regularisierungsrate nur unwesentlich größer als bei der einfachen Von-Neumann-Regularisierung ist, anderseits aber auch nicht zu lang, denn sonst könnte die bei dieser Regularisierungsmethode vorausgesetzte Vertauschbarkeit für die Bits eines Blockes möglicherweise nicht mehr gewährleistet sein. So können z.B. Drifteffekte, welche die Wahrscheinlichkeit p über die Länge des Blockes verändern, zu einer Verletzung dieser Voraussetzung führen, und ohne diese Voraussetzung kann keine Gleichverteilung der Ausgangsbits garantiert werden.

G.1.8 Iterierte Von-Neumann-Regularisierung bei statistisch abhängigen Bits

Die iterierte Von-Neumann-Regularisierung läßt sich auch auf den Fall verallgemeinern, daß es sich bei dem Eingangsbitstrom um eine endliche Markovkette handelt, wobei der praxisrelevanteste Fall wieder die zweigliedrige Markovkette ist. Man geht hierbei genauso vor, wie in Abschnitt G.1.6 für die Elias-Regularisierung beschrieben, nur daß anstatt der Elias-Regularisierung die iterierte Von-Neumann-Regularisierung nach Peres verwendet wird.

G.2 Pseudoregularisierungsverfahren

Die bisher besprochenen Regularisierungsverfahren zeichnen sich dadurch aus, daß sie garantieren, aus einem Eingangsbitstrom mit einer Entropie pro Bit kleiner Eins einen kürzeren Aus-

gangsbitstrom zu erzeugen, der den *Maximalwert der Entropie* als Erwartungswert erreicht. Erkauft wird diese ideale Eigenschaft des Ausgangsbitstroms durch seine kürzere Länge. Diese Reduktion gegenüber der Länge des Eingangsbitstroms ist unvermeidlich, da nur auf diese Weise die Entropie erhöht werden kann. Jede *deterministische* Transformation, die einen Bitstrom von n Bits wiederum in einen gleich langen Bitstrom abbildet, ist nämlich höchstens Entropie erhaltend, erhöht werden kann die Entropie nicht, so daß sich auf diese Weise *im Prinzip* kein Bitstrom mit *idealen* Eigenschaften erhalten läßt.
Für manche Anwendung ist allerdings ein Bitstrom mit idealen Eigenschaften gar nicht unbedingt notwendig. Oft ist ein Bitstrom ohne *feststellbare* statistische Defekte oder sogar mit geringen statistischen Defekten ausreichend, die unterhalb einer gewählten kritischen Schwelle liegen. Gibt man sich mit diesen etwas geringeren Anforderungen zufrieden, so kommen eine ganze Reihe von mathematischen Verfahren zur „*Pseudoregularisierung*“ in Betracht:

- die XOR-Transformation des Eingangsbitstroms,
- die XOR-Verknüpfung des Rohdatenstromes mit dem Ausgangsbitstrom eines guten Pseudozufallszahlengenerators, mit gleicher relativer Häufigkeit der Nullen und der Einsen und
- die Transformation des Eingangsbitstroms mit kryptographisch starken Blockchiffren.

G.2.1 Die XOR-Transformation

Die XOR-Transformation des Eingangsbitstroms stellt eine sehr einfache Art der Pseudoregularisierung dar, um den Ausgangsbitstrom näherungsweise zu regularisieren.
Bei der einfachsten Vorgehensweise wird der Eingangsbitstrom wie bei der Von-Neumann-Regularisierung in nicht überlappende Bitpaare aufgeteilt. Das Ausgangsbit ist dann gegeben durch den Ausgangswert der XOR-Verknüpfung der logischen Werte der beiden Bits des Paares:

Bit-Paar	Ausgangsbit	Wahrscheinlichkeit
[11]	0	p^2
[10]	1	$p \cdot (1-p)$
[01]	1	$(1-p) \cdot p$
[00]	0	$(1-p)^2$

Die Wahrscheinlichkeit, daß das Ausgangsbit den Wert Eins bzw. Null hat, beträgt in Abhängigkeit von der Wahrscheinlichkeit p_E für eine Eins im Eingangsstrom:

$$\begin{aligned} P(B_{aus}=1) &= 2 \cdot p_E \cdot (1-p_E) = 2p_E - 2p_E^2, \text{ bzw.} \\ P(B_{aus}=0) &= p_E^2 + (1-p_E)^2 = 1 - (2p_E - 2p_E^2) = 1 - P(Bit_{aus}=1) \end{aligned}$$

Beträgt nun die Wahrscheinlichkeit für eine Eins im Eingangsstrom $p_E = 0.5 + \epsilon$, so erhält man nach der XOR-Transformation die Wahrscheinlichkeit für eine Eins im Ausgangsbitstrom:

$$P(B_{aus}=1 \,|\, p_E) = 2 \cdot (0.5+\epsilon) - 2 \cdot (0.5+\epsilon)^2 = 0.5 - 2 \cdot \epsilon^2.$$

Durch die XOR-Transformation *reduziert* sich also die systematische Abweichung vom Idealwert. Allerdings wird dies mit einer konstant 50-prozentigen Reduktion der Bitrate erkauft.
Grundsätzlich läßt sich die systematische Abweichung vom idealen Verhältnis noch weiter reduzieren, indem man auf den derart erzeugten Ausgangsbitstrom wiederum eine XOR-Transformation anwendet. Allerdings führt dies bereits zu einer konstanten Reduktion um insgesamt 75% gegenüber der ursprünglichen Eingangsbitrate. In diesem Falle kann man natürlich auch gleich die Von-Neumann-Regularisierung verwenden[16], die eine *ideale* Verteilung der Ausgangsbitfolge erzeugt.

[16] Die Von-Neumann-Regularisierung führt natürlich zu einer Reduktion von *mindestens* 75%, allerdings befindet man sich bei den praxisrelevanten Fällen meist nur wenig über diesem Wert.

Der Vorteil der XOR-Transformation gegenüber der Von-Neumann-Regularisierung besteht lediglich darin, daß immer ein fester und bei einfacher XOR-Transformation höherer Prozentsatz der Eingangsbitrate ausgegeben wird. Im Falle eines Defektes kann ersteres allerdings von Nachteil sein, da z.B. lange Läufe von Nullen bzw. Einsen in der Eingangssequenz auf entsprechende, wenn auch nur halb so lange, Null-Läufe in der Ausgangssequenz abgebildet werden. Bei der Von-Neumann-Regularisierung hingegen fallen solche Läufe quasi ganz[17] heraus.
Zusammenfassend läßt sich die XOR-Transformation als eine Pseudoregularisierungsmethode charakterisieren, die sich zwar leicht in Hardware oder Software implementieren läßt, aber nur in einfacher Form und für geringe Abweichungen vom Idealverhältnis sinnvoll anwendbar ist. Echte Regularisierungsverfahren sind ihr auf jeden Fall vorzuziehen.

G.2.2 XOR-Verknüpfung des Rohdatenstroms mit dem Ausgabestrom eines Pseudozufallsgenerators

In diesem Fall kombiniert man die Vorteile eines physikalischen Zufallsgenerators – Korrelationsfreiheit und Unvorhersagbarkeit – mit dem Vorteil eines (guten) Pseudozufallsgenerators, gleiche relative Häufigkeiten von Nullen und Einsen zu besitzen. Erreicht wird dies durch eine XOR-Verknüpfung der Rohdatenbits des physikalischen Zufallsgenerators mit den entsprechenden Ausgangsbits des Pseudozufallsgenerators. Seien die Wahrscheinlichkeiten für eine Eins bzw. eine Null innerhalb der Rohdaten $P(B_{rnd} = 1) = p$ und $P(B_{rnd} = 0) = (1-p) = q$ mit $p \neq q$ und die als ideal angenommen Wahrscheinlichkeiten im Ausgangsstrom des Pseudozufallsgenerators $P(B_{prng} = 1) = P(B_{prng} = 0) = 1/2$, dann erhält man für die entsprechenden Wahrscheinlichkeiten des schlußendlichen Ausgangsbitstroms nach Durchführung der XOR-Operation:

$$\begin{aligned} P(B_{aus} = 0) &= P(B_{rnd} = 0 \oplus B_{prng} = 0) + P(B_{rnd} = 1 \oplus B_{prng} = 1) \\ &= q \cdot \frac{1}{2} + p \cdot \frac{1}{2} = \frac{1}{2}\text{ , bzw.} \\ P(B_{aus} = 1) &= P(B_{rnd} = 0 \oplus B_{prng} = 1) + P(B_{rnd} = 1 \oplus B_{prng} = 0) \\ &= q \cdot \frac{1}{2} + p \cdot \frac{1}{2} = \frac{1}{2}. \end{aligned}$$

Die Wahrscheinlichkeiten für Nullen und Einsen sind also nach der XOR-Verknüpfung der beiden Zufallsströme im Ausgangsstrom gleich. Als Pseudozufallszahlengeneratoren empfehlen sich natürlich wieder insbesondere die in Abschnitt B.1 aufgeführten kryptographisch starken Pseudozufallsgeneratoren; alternativ kann hierfür auch der Chiffrenstrom von dedizierten Stromchiffrierern verwendet werden.

G.2.3 Transformation mit kryptographisch starken Blockchiffren

Kryptographisch starke Blockchiffren lassen sich sehr leicht zur Pseudoregularisierung verwenden. Generell transformiert eine Blockchiffre einen n Bit langen Eingangsblock[18] mit Hilfe eines schlüsselabhängigen, stark nichtlinearen[19] Algorithmus in einen ebenfalls n Bit langen Ausga-

[17] Selbstverständlich heißt dies nicht, daß es keine legitimen langen Läufe in einer Zufallssequenz geben könnte! Bei einer Von-Neumann-Regularisierung werden zwar diese ursprünglichen Läufe zerstört, dafür werden aber innerhalb der Ausgangssequenz neue erzeugt, die von Eingangsteilsequenzen mit alternierenden Bitwerten verursacht werden. Defekte an einem Generator wirken sich bei einer nachfolgenden Von-Neumann-Regularisierung deswegen weniger aus, weil sie alternierende Bitfolgen erzeugen müßten, um zu langen Läufen in der regularisierten Ausgabesequenz zu führen.

[18] Bei vielen Blockchiffren (z.B. bei DES und IDEA) ist $n = 64$, erst bei neueren Blockchiffren ist n größer, meist $n = 128$.

[19] Beim Entwurf kryptographisch starker Blockchiffren wird i.a. gefordert, daß das *Lawinen-Kriterium* [69, 91] gilt: Die Änderung eines einzigen Bits des Schlüssels oder des Eingangsblockes führt zur Änderung der Bitwerte bei mindestens 50% der Bits des Ausgangsblocks.

beblock. Es bietet sich daher an, eine Blockchiffre zur effizienten Pseudoregularisierung zu verwenden, da anders als bei den meisten echten Regularisierungsmethoden keine Bits „verworfen“ werden; hierbei geht man folgendermaßen vor:

1. Man wählt einen geheimen, zufälligen[20] Schlüssel für die Blockchiffre. Da es im Prinzip ausreicht, nur einmal einen Schlüssel zu wählen und er auch nur eine geringe Anzahl von Bits umfaßt, kann man die Bits für diesen Schlüssel aus dem Zufallsstrom entnehmen und sie mit Hilfe eines (ineffizienten) Regularisierungsalgorithmus regularisieren.
2. Der restliche Zufallstrom wird in n Bit lange Blöcke aufgeteilt, die anschließend nacheinander von der Blockchiffre mit dem im ersten Schritt erzeugten Schlüssel verschlüsselt werden.
3. Optional kann in unregelmäßigen Abständen gemäß Schritt Eins ein neuer Schlüssel generiert werden.

Die oben genannte Vorgehensweise ist aus folgenden Gründen ebenfalls „nur“ als Pseudoregularisierung anzusehen:

- Die Bits eines Ausgangsblocks sind, wenn auch in komplizierter, nichtlinearer Weise, von den Bits des Schlüssels abhängig.
- Die Bits eines Ausgangsblocks sind *nicht* voneinander unabhängig, allerdings sind diese nichtlinearen Abhängigkeiten viel zu schwach, um festgestellt werden zu können[21].
- Bei Verwendung einfacher Chiffriermodi, wie z.B. dem *Electronic Code Book-Modus, (ECB)* werden gleiche Eingabe-Blöcke auf ebenfalls gleiche Ausgabe-Blöcke abgebildet.

Selbst wenn man also den Schlüssel häufig wechselt, z.B. indem man den jeweils vorhergehenden Eingangsblock als Schlüssel für den nächsten Eingangsblock verwendet, fügt man doch immer Abhängigkeiten in den Ausgangsbitstrom ein. Etwas polemisch ausgedrückt, kann man sagen, daß diese verlustfreien Pseudoregularisierungsverfahren den Zufalls-Bitstrom „verschlimmbessern“. Wobei man der Fairneß halber darauf hinweisen muß, daß die durch sie eingefügten, komplizierten Abhängigkeiten in empirischen Tests nicht zu entdecken sind, wie die guten statistischen Eigenschaften der auf starken Blockchiffren beruhenden Pseudozufallszahlengeneratoren zeigen. Überdies haben starke Blockchiffren die wünschenswerte Eigenschaft, daß sie etwaige Abhängigkeiten im Eingangsstrom, wie z.B. leichte Korrelationen zwischen den Bits, im Ausgangsstrom vollständig unkenntlich machen.
Will man allerdings einen langen Einmalschlüssel (*One Time Pad*) mit Hilfe eines physikalischen Zufallsgenerators erzeugen, sollten *keine* Pseudoregularisierungverfahren, sondern nur exakte Regularisierungsverfahren verwendet werden, da nur sie die ideal zufälligen Eigenschaften des Einmalschlüssels garantieren, die für den mathematischen Beweis der Sicherheit des zugehörigen Verschlüsselungsverfahrens notwendig sind. Dies mag vielleicht auf den ersten Blick etwas übertrieben erscheinen, aber nur auf diese Weise läßt sich der absolute Sicherheitsanspruch der Verschlüsselungsverfahren auf Basis von langen Einmalschlüsseln aufrechterhalten. Nach heutigem (öffentlichen) Wissensstand ist allerdings nicht davon auszugehen, daß sich die durch Pseudoregularisierungsverfahren auf Basis von Blockchiffren eventuell eingeführten Abhängigkeiten

[20] Der Schlüssel muß nicht unbedingt zufällig gewählt sein, da die Pseudoregularisierung mit starken Verschlüsselungsalgorithmen auch bei einem sehr stark strukturbehafteten Schlüssel nicht erkennbar schlechter ist. Zufällig gewählte Schlüssel lassen sich lediglich schlechter erraten, was hier allerdings kaum eine Rolle spielt.

[21] Das Problem aus einem Eingangsbitstrom mit niedriger Entropie auf deterministische und überdies umkehrbare Weise einen Ausgangsbitstrom mit scheinbar hoher Entropie zu erzeugen, stellt das Grundproblem der Kryptographie dar. Denn durch eine Verschlüsselung soll ein Bitstrom mit Korrelationen und einer (starken) Ungleichverteilung der einzelnen Symbole, nämlich der *Klartext*, mit Hilfe eines schlüsselabhängigen Algorithmus zu einem *Chiffrentext* verschlüsselt werden, der keine erkennbare Struktur aufweist, so daß ein Kryptanalytiker keinen Ansatzpunkt für einen Angriff findet.

zwischen den Bits eines Ausgangsblockes tatsächlich für eine praktische Attacke gegen einen den Ausgangsbitstrom verwendenden kryptographischen Algorithmus ausnutzen ließen; hierzu sind die nichtlinearen Abhängigkeiten zwischen den Bits zu kompliziert und die Entropie pro Bit des Eingangsbitstrom i.a. doch zu groß.
Sehr viel kritischer ist das Problem, das bei einem Defekt im Zufallsgenerator auftreten kann. Starke Musterbildung[22] des defekten Generators wird durch eine nachfolgende, einfache Transformation der Daten mit einer Blockchiffre lediglich verschleiert. Zwar sind die einzelnen Blöcke nach der Transformation nicht mehr musterbehaftet, aber eine Wiederholung des Eingangsmusters führt auch wieder zu einem gleichen Ausgabeblock, was sich allerdings nicht immer leicht erkennen läßt, da es sich aufgrund der Blocklänge (typ. 64 oder 128 Bit) nur in entsprechend langreichweitigen Korrelationen zeigt. Umgehen läßt sich dieses Problem, indem man nicht allein den Rohdatenbitstrom des Zufallsgenerators verschlüsselt, sondern zusätzlich eine Rückkopplung des Ausgangsbitstroms vorsieht, ähnlich wie dies bei der Verwendung einer Blockchiffre im OFB-Modus als Pseudozufallsgenerator geschieht, s. S. 18.
Abschließend sei noch kurz erwähnt, daß sich natürlich auch kryptographisch starke Hashfunktionen, s. z.B. [91], Kap. 9, gut für eine Pseudoregularisierung verwenden lassen, s. Abschnitt 36. Allerdings werden hierbei in starkem Maße Bits verworfen, so daß sich Hashfunktionen eher für die „Zufallsdestillierung“ relativ kurzer Zufallssequenzen aus einer großen Menge von schwach zufälligen Rohdaten eignen[23]. Für die Regularisierung der Rohdaten herkömmlicher, physikalischer Zufallsgeneratoren sind sie hingegen weniger geeignet, da sie weder im Hinblick auf das Laufzeitverhalten noch die Regularisierungseffizienz guten Regularisierungsverfahren überlegen sind. Ein kleiner Vorteil der Pseudoregularisierung mit Hilfe von kryptographisch starken Hashfunktionen liegt lediglich darin, daß sie – wie die Pseudoregularisierung mit Blockchiffren – eine konstante Regularisierungseffizienz garantieren. Wenn man allerdings Pseudoregularisierungsverfahren auf Basis kryptographisch starker Algorithmen einsetzen möchte, wird man mit Blick auf die ideale Regularisierungseffizienz wohl eher gleich Blockchiffren oder die XOR-Verknüpfung mit einer Stromchiffre verwenden.

G.3 Fazit: Welches Regularisierungsverfahren sollte man verwenden?

Es stehen mehrere Regularisierungsalgorithmen zur Lösung der beiden Hauptprobleme physikalischer Zufallsgeneratoren – eine Tendenz in den relativen Häufigkeiten von Nullen und Einsen und statistische Abhängigkeiten zwischen aufeinanderfolgenden Bits – zur Verfügung.
Generell sind effizientere Algorithmen natürlich zu bevorzugen, allerdings setzt dies z.B. bei den Verfahren von Elias oder Peres fast zwangsläufig eine Software-Implementierung voraus. Will man Regularisierungsalgorithmen in Hardware oder softwarebasiert auf einem Mikrocontroller implementieren, wird man auf die einfacheren Algorithmen von Samuelson und Pratt bzw. von John von Neumann zurückgreifen.
Als Endpunkt der Überlegungen in diesem Abschnitt wird empfohlen, bei Verwendung eines physikalischen Zufallsgenerators zumindest Software-Implementationen effizienter Regularisierungsalgorithmen zu nutzen, da sie es ermöglichen, einen Bitstrom zu erzeugen, der tatsächlich *maximale Entropie* hat. Zwar mag dies im Hinblick auf die beim Einsatz in der Praxis ohnehin meist nachfolgende Transformation des Bitstroms durch starke Blockchiffren oder durch eine XOR-Verknüpfung mit einer Stromchiffre übertrieben wirken, man sollte hier allerdings bedenken, daß diese Transformationen die Entropie „lediglich“ auf eine größere Anzahl von Bits verteilen, nicht aber erhöhen können. Insofern bedeutet eine vorgeschaltete Regularisierungsstufe

[22] Am realistischsten sind lange Läufe mit nur einem Bitwert im Ausgangsbitstrom, aber es können auch komplizierte Defekte auftreten, s. Abschnitt 5.1.2.5.

[23] Wie dies bei dem auf Seite 22 erwähnten „Lavalampen-Zufallsgenerator“ [115] geschieht.

einen – zugegebenermaßen kleinen – zusätzlichen Gewinn an Sicherheit.
Überdies lassen sich mit Hilfe von Regularisierungsverfahren die Auswirkungen geglückter Angriffe auf einen physikalischen Zufallsgenerator minimieren: Regularisierungsalgorithmen verändern immer die Anordnung der Bits *und* reduzieren ihre Anzahl, somit wird die Vorhersagbarkeit des Ausgangsbitstroms bei teilweise bekanntem Eingangsbitstrom erschwert, was insbesondere im Zusammenwirken mit einer nachfolgenden Verschlüsselung die Sicherheit zusätzlich erhöht. Auch im Falle eines Defektes des Generators ist dies von Vorteil, da eine alleinige Verschlüsselung keine Bits verwirft.
Für die nachfolgende Chiffrierung der regularisierten Zufallsströme empfiehlt sich die Verwendung einer Stromchiffre, da diese zusätzlich regularisierend wirkt und auch im Falle eines massiven Defektes des Generators immer noch einen guten Pseudozufallsstrom liefert.

Literaturverzeichnis

[1] AB, Protego Information: *SG100*. http://www.protego.se.

[2] Abramowitz, Milton und Irene A. Stegun (Herausgeber): *Pocketbook of Mathematical Functions*. Verlag Harr Deutsch, Frankfurt a. M., Frankfurt a. M., 1984.

[3] Adleman, R.L., L. Rivest und A. Shamir: *A Method for Obtaining Digital Signatures and Public-key Cryptosystems*. Communications of the ACM, 21(2):120–126, 1978.

[4] Agnew, G.B.: *Random Sources for Cryptographic Systems*. In: *Advances in Cryptology – EUROCRYPT '87 (LNCS 304)*, Band 304 der Reihe *LNCS*, S. 77–81, Berlin, 1988. Springer Verlag.

[5] Aguayo, Ricardo, Geoff Simms und P.B. Siegel: *Throwing nature´s dice*. American Journal of Physics, 64(6):752–758, 1996.

[6] Alexi, W., B. Chor, O. Goldreich und C. P. Schnorr: *RSA and Rabin Functions: Certain Parts are as Hard as the Whole*. SIAM Journal on Computing, 17(2):194–209, 1988.

[7] Bar-Hillel, Maya und Willem A. Wagenaar: *The Perception of Randomness*. Advances in Applied Mathematics, 12:428–454, 1991.

[8] Basieux, Pierre: *Die Welt als Roulette*. Rowohlt, Reinbek bei Hamburg, 1995.

[9] Bass, Thomas A.: *Der Las Vegas-Coup*. Birkhäuser Verlag, Basel, 1991. orig. Titel: The Newtonian Casino.

[10] Bennett, C. H., F. Bessette, G. Brassard, L. Salvail und J. Smolin: *Experimental Quantum Cryptography*. Journal of Cryptology, 5:3–28, 1992.

[11] Beth, T., D. E. Lazic und A. Mathias: *Cryptanalysis of cryptosystems based on remote chaos replication*. In: Desmedt, Yvo (Herausgeber): *Advances in Cryptology, Crypto '94*, Band 839 in *Lecture Notes in Computer Science*, S. 318–331, Berlin, 1994. Springer.

[12] Beth, Thomas und Zong Duo Dai: *On the Complexity of Pseudo-Random Sequences - or: If You Can Describe a Sequence it Can´t be Random*. In: Quisquater, Jean Jacques und Joos Vandewalle (Herausgeber): *Advances in Cryptology, EuroCrypt ´89*, Band 434 der Reihe *Lecture Notes in Computer Science*, S. 533 – 543, Berlin, 1989. Springer-Verlag.

[13] Blum, L., M. Blum und M. Shub: *A simple unpredictable pseudo-random number generator*. SIAM Journal on Computing, **15**:364–383, 1986.

[14] Bratley, P., B.L. Fox und L.E. Schrage: *A Guide to Simulation*. Springer, New York, 2nd Auflage, 1987.

[15] Brendel, J., R. Lange, E. Mohler und W. Martienssen: *Strahlteilungs- und Interferenzexperimente mit Photonenpaaren*. Annalen der Physik, 7:26–40, 1991.

[16] Brendel, J., E. Mohler und W. Martienssen: *Time-Resolved Dual-Beam Two-Photon Interferences with High Visibility*. Phys. Rev. Lett., 66:1142–1145, 1991.

[17] Brendel, J., S. Schütrumpf, R. Lange und W. Martienssen: *A Beam Splitting Experiment with Correlated Photons*. Europhys. Lett., 5:223–228, 1988.

[18] Brendel, Jürgen: *Quantenphänomene in der Welt des Lichtes*. Harri Deutsch Verlag, Thun; Frankfurt a. M., 1994. Dissertation am Fachbereich Physik der Johann-Wolfgang-Goethe-Universität.

[19] Brunel, Christian, Brahim Lounis, Philippe Tamarat und Michel Orrit: *Triggered Source of Single Photons based on Controlled Single Molecule Fluorescence*. Phys. Rev. Lett., 83(14):2722–2725, 1999.

[20] Burnham, D. C. und D. L. Weinberg: *Observation of Simultaneity in Parametric Production of Optical Photon Pairs*. Phys. Rev. Lett., 25:84–87, 1970.

[21] Chaitin, G. J.: *Information, Randomness & Incompleteness*, Band 8 der Reihe *Series in Computer Science*. World Scientific, Singapore, 2nd Auflage, 1990.

[22] Charikar, Moses, Eric Lehman, Ding Liu, Rina Panigrahy, Manoj Prabhakaran, April Rasala, Amit Sahai und Abhi Shelat: *Approximating the Smallest Grammar: Kolmogorov Complexity in Natural Models*. In: *STOC´02*, Montreal, Quebec, Canada, 19-21 Mai 2002. ACM.

[23] Compagner, Aaldert: *Definitions of randomness*. American Journal of Physics, 59(8):254–309, 1991.

[24] ComScire: *ComScire QNG, True Random Numbers*. http://shell.rmi.net/ comscire.

[25] Coron, Jean-Sébastien: *On the Security of Random Sources*. In: Imai, H. und Y. Zheng (Herausgeber): *Proceedings of PKC '99*, Band 1560 der Reihe *Lecture Notes in Computer Science*, S 29 ff, Berlin, 1999. Springer.
http://www.eleves.ens.fr/home/coron/science.html.

[26] Coron, Jean-Sébastien und David Naccache: *An Accurate Evaluation of Maurer's Universal Test*. In: Tavares, S. und H. Meijer (Herausgeber): *Proceedings of SAC '98*, Band 1556 der Reihe *Lecture Notes in Computer Science*, S. 57–71, Berlin, 1998. Springer.
http://www.eleves.ens.fr/home/coron/science.html.

[27] Corp., Aware Electronics: *AW-RAND*. http://www.aw-el.com.

[28] Davies, Neville, Ed Dawson, Helen Gustafson und A. N. Pettitt: *Testing for Randomness in Stream Ciphers Using the Binary Derivative*. Statistics and Computing, 5:307–310, 1995.

[29] DeGroot, Morris H.: *A Conversation with Persi Diaconis*. Statistical Science, 1(3):319–334, 1986.

[30] Division, Intel Platform Security: *The Intel Random Number Generator*. Intel Corporation, http://developer.intel.com/design/security/rng/rng_v3.htm, Order Number: 298029-001, Dezember 1999.

[31] Dr. Schabhüser, Bundesamt für Sicherheit in der Informationstechnik: *Persönliche Mitteilung*.

[32] Ebeling, Werner, Jan Freund und Frank Schweitzer: *Komplexe Strukturen: Entropie und Information*. B.G. Teubner, Stuttgart, 1998.

[33] Edelkind, Jamie, Ilya M. Vitebskiy, Alexander Figotion und Vadim Popovich: *Random number generator based on directional randomness associated with naturally occuring random events, and method therefor.* United States Patent, Nr. 5.987.483, November 1999.

[34] Ekert, A. K.: *Quantum Cryptography Based on Bell's Theorem.* Phys. Rev. Lett., 67:661–663, 1991.

[35] Elektronic Westphal: *Z5000, Zufallsgenerator für echte Zufallszahlen.* http://home.t-online.de/home/p.westphal/zz.htm.

[36] Elias, Peter: *The Efficient Construction of an Unbiased Random Sequence.* Annals of Mathematical Statistics, 43(3):865–870, 1972.

[37] Engel, Eduardo M.R.A.: *A Road to Randomness in Physical Systems*, Band 71 der Reihe *Lecture Notes in Statistics.* Springer Verlag, Berlin, 1992.

[38] Erber, T., P. Hammerling, G. Hockney, M. Porrati und S. Putterman: *Resonance Fluorescence and Quantum Jumps in Single Atoms: Testing the Randomness of Quantum Mechanics.* Annals of Physics, 190:254–309, 1989.

[39] Erber, T. und S. Putterman: *Randomness in quantum mechanics – nature's ultimate cryptogram.* Nature, 318:41–43, November 1985.

[40] Erber, T., T. M. Rynne, W.F. Darsow und M.J.Frank: *The Simulation of Random Processes on Digital Computers: Unavoidable Order.* Journal of Computational Physics, 49:394–419, 1983.

[41] Fairfield, R.C., R.L. Mortenson und K. B. Coulthart: *An LSI Random Number Generator (RNG).* In: *Advances in Cryptology – Proceedings of CRYPTO 84*, Band 196 der Reihe *LNCS*, S. 203–230, Heidelberg, 1985. Springer Verlag.

[42] Fearn, H. und R. Loudon: *Quantum Theory of the Lossless Beam Splitter.* Opt. Comm., 64:485–490, 1987.

[43] Feller, William: *An Introduction to Probability Theory and Its Applications*, Band 2. John Wiley & Sons, New York, 1966.

[44] Ford, Joseph: *How random is a coin toss.* Physics Today, S. 40–47, April 1983.

[45] Freitag, Rolf: *Zuverlässiger Zufall.* Linux Journal, S. 39–41, Juni 1999.

[46] Gilbert, E. N. und T. T. Kadota: *The Lempel-Ziv Algorithm and Message Complexity.* IEEE Trans. Inf. Theory, 38(6):1839–1945, 1992.

[47] Goldenberg, Lior und Lev Vaidman: *Quantum Cryptography Based on Orthogonal States.* Phys. Rev. Letters, 75(7):1239 ff, 1995.

[48] Goldenberg, Lior und Lev Vaidman: *Goldenberg and Vaidman Reply.* Phys. Rev. Letters, 77(15):3265, 1996.

[49] Grangier, P., G. Roger und A. Aspect: *Experimental Evidence for a Photon Anticorrelation Effect on a Beam Splitter: A New Light on Single-Photon Interferences.* Europhys. Lett., 1:173–179, 1986.

[50] Gude, Martin: *Ein quasi-idealer Gleichverteilungsgenerator basierend auf physikalischen Zufallsphänomenen.* Doktorarbeit, RWTH Aachen, 1987.

[51] Gustafson, H. M., E. P. Dawson und J. Dj. Golić: *Randomness Measures Related to Subset Occurrence.* In: Dawson, E. und J. Dj. Golić (Herausgeber): *Cryptography: Policy and Algorithms*, Band 1029 der Reihe *Lecture Notes in Computer Science*, S. 132–143, Berlin, 1996. Springer-Verlag.

[52] Han, Te Sun und Mamoru Hoshi: *Interval Algorithm for Random Number Generation.* IEEE Trans. Inf. Theory, 43(2):599–611, März 1997.

[53] Herzel, Hanspeter und Ivo Grosse: *Measuring correlations in symbol sequences.* Physica A, 216:518–542, 1995.

[54] Home, D. und F. Selleri: *Bells´s Theorem and the EPR Paradox.* La Rivista del Nouvo Cimento, 14(9):1–93, 1991.

[55] Hong, C. K. und L. Mandel: *Experimental Realization of a Localized One-Photon State.* Phys. Rev. Lett., 56:58–60, 1986.

[56] Hong, C. K., Z. Y. Ou und L. Mandel: *Measurement of Subpicosecond Time Intervall between Two Photons by Interference.* Phys. Rev. Lett., 59:2044–2046, 1987.

[57] Hughes, Richar J., D. M. Alde, P. Dyer, G. G. Luther, G. L. Morgan und M. Schauer: *Quantum Cryptography.* Contemporary Physics, 36(3):149–163, 1995.

[58] ICATT: *ORION´s Random Number Generator.* http://www.valley.ineract.nl/av/com/rng/home.html.

[59] Informationstechnik mbH, SIT Gesellschaft für Systeme: *Der Zufallsgenerator FZG 100.* Datenblatt.

[60] Inoue, H., H. Kumahora, Y. Yoshizawa, M. Ichimura und O. Miyatake: *Random Numbers Generated by a Physical Device.* Applied Statistics, 32(2):115–120, 1983.

[61] Jansen, Cees J.A.: *Investigations On Nonlinear Steamcipher Systems: Construction and Evaluation Methods.* Doktorarbeit, Technical University of Delft, Delft, 1989.

[62] Jansen, Cees J. A. und Dick E. Boekee: *The Shortest Feedback Shift Register That Can Generate A Given Sequence.* In: Brassard, G. (Herausgeber): *Advances in Cryptology – Proceedings of CRYPTO 89*, Band 435 der Reihe *LNCS*, S. 90–98, Heidelberg, 1990. Springer Verlag.

[63] Jennewein, Thomas, Ulrich Achleitner, Gregor Weihs, Harald Weinfurter und Anton Zeilinger: *A fast and compact quantum random number generator.* Review of Scientific Instruments, 71(4):1675–1680, 2000.

[64] Juels, Air, Markus Jakobson, Elizabeth Shriver und Bruce L. Hillyer: *How to Turn Loaded Dice into Fair Coins.* IEEE Trans. Inf. Theory, 46(3):911–921, Mai 2000.

[65] Kantz, Holger und Thomas Schreiber: *Nonlinear time series analysis.* Cambridge University Press, Cambridge, 1997.

[66] Keller, Joseph H.: *The Probability of Heads.* American Mathematical Monthly, 93:191–197, 1986.

[67] KG, I.Q.uality GmbH & Co.: *Zufallszahlengenerator Random Master.* http://iqualitiy.de.

[68] Knuth, D.E.: *The Art of Computer Programming*, Band 2: Seminumerical Algorithms. Addison-Wesley, Reading, MA, 3. Auflage, 1998.

[69] Konheim, Alan G.: *Cryptography, A Primer.* John Wiley & Sons, New York, 1981.

[70] Krickeberg, Klaus und Herbert Ziezold: *Stochastische Methoden.* Springer-Verlag, Berlin, 4. Auflage, 1995.

[71] Kuusela, T.: *Random Number Generation Using a Chaotic Circuit.* Journal of Nonlinear Science, 3:445–457, 1993.

[72] Lagarias, J.C.: *Pseudorandom Numbers.* Statistical Science, **8**:31–39, 1993.

[73] Lange, R., J. Brendel, E. Mohler und W. Martienssen: *Beam Splitting Experiments with Classical and Quantum Particles.* Europhys. Lett., 5:619–622, 1988.

[74] Larchuk, T. S., M. C. Teich und B. E. A. Saleh: *Statistics of Entangled-Photon Coincidences in Parametric Downconversion.* In: *Annals of the New York Academy of Sciences*, Band 755, S. 681–686, 1995.

[75] L'Ecuyer, P.: *Efficient and Portable Combined Random Number Generators.* Comm. ACM, **31**:742–774, 1988.

[76] L'Ecuyer, Pierre: *Random Number Generators.* In: Gass, S. I. und C. M. Harris (Herausgeber): *Encyclopedia of Operations Research and Management Science*, S. 571–578. Kluwer Academic Publishers, 1996.

[77] Leeb, H.: pLab – *a system for testing random numbers.* In: Vajteršic, M. und P. Zinterhof (Herausgeber): *Proceedings of the International Workshop Parallel Numerics '94, Smolenice, Sept. 19–21*, S. 89–99, Slovakia, 1994. Slovak Academy of Sciences, Institute for Informatics. http://random.mat.sbg.ac.at.

[78] Leeb, H.: *On the digit test.* In: Hellekalek, P., G. Larcher und P. Zinterhof (Herausgeber): *Proceedings of the 1st Salzburg Minisymposium on Pseudorandom Number Generation and Quasi-Monte Carlo Methods, Salzburg, Nov 18, 1994*, Band ACPC/TR 95-4 der Reihe *Technical Report Series*, S. 109–121. ACPC – Austrian Center for Parallel Computation, Universität Wien, Austria, 1995.

[79] Lehman, Eric: *Approximation Algorithms for Grammar-Based Data Compression.* Doktorarbeit, Massachusetts Institute of Technology, Februar 2002.

[80] Lempel, Abraham und Jacob Ziv: *On the Complexity of Finite Sequences.* IEEE Trans. Inf. Theory, 22(1):75–81, 1976.

[81] Leung, A. K. und S. E. Tavares: *Sequence Complexity as a Test for Cryptographic Systems.* In: Blakley, G. R. und D. Chaum (Herausgeber): *Advances in Cryptology – CRYPTO 85*, Band 196 der Reihe *LNCS*, S. 468–474, Berlin, 1985. Springer Verlag.

[82] Li, Ming und Paul Vitányi: *An Introduction to Kolmogorov Complexity and it's Applications.* Graduate Texts in Computer Science. Springer-Verlag, Berlin, 1997.

[83] Li, Wentian: *Mutual Information Functions versus Correlation Functions.* Journal of Statistical Physics, 60(5/6):823–837, 1990.

[84] Loudon, R.: *The quantum theory of light.* Clarendon Press, Oxford, 1983.

[85] Lounis, B. und W. E. Moerner: *Single Photons on Demand from a Single Molecule at Room Temperature.* Nature, 407:491, 2000.

[86] Luby, Michael: *Pseudorandomness and Cryptographic Applications.* Princeton University Press, Princeton, NJ, 1996.

[87] Marron, J., A. J. Martino und G. M. Morris: *Generation of Random Arrays Using Clipped Laser Speckle.* Applied Optics, 25:26–30, 1986.

[88] Martino, Anthony J. und G. Michael Morris: *Optical random number generator based on photoevent locations.* Applied Optics, 30(8):981–989, 1991.

[89] Massey, J. L.: *Shift-register synthesis and BCH decoding.* IEEE Trans. Inf. Theory, 15:122–127, 1969.

[90] Maurer, U.M.: *A universal statistical test for random bit generators.* Journal of Cryptology, **5**:89–105, 1992.

[91] Menezes, Alfred J., Paul C. van Oorschot und Scott A. Vanstone: *Handbook of Applied Cryptography.* CRC Press, Boca Raton, Florida, 1997.

[92] Micali, S. und C.P. Schnorr: *Efficient, Perfect Random Number Generators.* In: Goldwasser, S. (Herausgeber): *Advances in Cryptology - Crypto '88*, Band 403 der Reihe *Lecture Notes in Computer Science*, S. 173–197. Springer Verlag, Berlin, 1988.

[93] Michler, P., A. Kiraz, C. Becher, W. V. Schoenfeld, P. M. Petroff, Lidong Zhang, E. Hu und A. Imamoğlu: *A Quantum Dot Single-Photon Turnstile Device.* Science, 290:2282–2285, December 2000.

[94] Mund, Sibylle: *Ziv-Lempel Complexity for Periodic Sequences and its Cryptographic Application.* In: Davies, D. W. (Herausgeber): *Advances in Cryptology – EUROCRYPT 91*, Band 547 der Reihe *LNCS*, S. 114–126, Berlin, 1991. Springer Verlag.

[95] Murry, Herschell F.: *A General Approach for Generating Natural Random Variables.* IEEE Transactions on Computers, S. 1210–1213, December 1970.

[96] Neumann, John von: *Various Techniques Used in Connection with Random Digits. Notes by G. E. Forsythe.* Nat. Bur. Stand. Applied Math. Series, 12:36–38, 1951. nachgedruckt in „Collected Works“, Vol. 5, S. 769-770.

[97] Nevill-Manning, Craig G.: *Inferring Sequential Structure.* Doktorarbeit, University of Waikato, May 1996.

[98] Nevill-Manning, Craig G. und Ian H. Witten: *Identifying Hierarchical Structure in Sequences: A linear-time algorithm.* Journal of Artificial Intelligence Research, 7:67–82, 9 1997.

[99] Nielsen, Michael A. und Isaac L. Chuang: *Quantum Computation and Quantum Information.* Cambridge University Press, Cambridge, 2002.

[100] N.N.: *Hitachi II, Vervielfachter Speicher.* à la Card, S. 252–253, Juli 1999.

[101] Park, S.K. und K.W. Miller: *Random number generators: good ones are hard to find.* Comm. ACM, **31**:1192–1201, 1988.

[102] Peres, Asher: *Quantum Cryptography with Orthogonal States?* Phys. Rev. Letters, 77(15):3264, 1996.

[103] Peres, Yuval: *Iterating von Neumann´s Procedure for Extracting Random Bits.* The Annals of Statistics, 20(1):590–597, 1992.

[104] Phoenix, Simon J. D. und Paul D. Townsend: *Quantum Cryptography: how to beat the code breakers using quantum mechanics.* Contemporary Physics, 36(3):165–195, 1995.

[105] Prasad, S., M. O. Scully und W. Martienssen: *A Quantum Description of the Beam Splitter.* Opt. Comm., 62:139–145, 1987.

[106] Press, W.H., W.T. Vetterling, S.A. Teukolsky und B.P.Flannery: *Numerical Recipes in C.* Cambridge University Press, New York, 1992.

[107] Rabin, M. O.: *Probabilistic algorithm for testing primality.* Journal of Number Theory, 12:128–138, 1980.

[108] Rarity, J.G., C.M. Owens und P.R.Tapster: *Quantum random-number generation and key sharing.* Journal of Modern Optics, 41(12):2435–2444, 1994.

[109] Rényi, A.: *Wahrscheinlichkeitsrechnung.* VEB - Deutscher Verlag der Wissenschaften, Berlin, 1971.

[110] Richter, Manfred: *Ein Rauschgenerator zur Gewinnung von quasi-idealen Zufallszahlen für die stochastische Simulation.* Doktorarbeit, RWTH Aachen, 1992.

[111] Rueppel, Rainer A.: *Analysis and Design of Stream Ciphers.* Communications and Control Engineering Series. Springer-Verlag, Berlin, 1986.

[112] Samuelson, Paul A.: *Constucting an Unbiased Random Sequence.* Journal of the American Statistical Society, 63(2):1526–1527, 1968.

[113] Schneier, Bruce: *Angewandte Kryptographie.* Addision-Wesley, Bonn, 1. Auflage, 1996.

[114] Schuster, Heinz Georg: *Deterministisches Chaos.* VCH, Weinheim, 1994.

[115] SGI. http://lavarand.sgi.com.

[116] Shelat, Abhi: *Evaluating Grammar-Based Data Compression Algorithms.* Diplomarbeit, Massachusetts Institute of Technology, August 2001.

[117] Shiryayev, A. N. (Herausgeber): *Selected Works of A. N. Kolmogorov*, Band 3 der Reihe *Mathematics and its Applications (Soviet Series), Vol. 27.* Kluver, Dordrecht, 1987.

[118] Solomonoff, R. J.: *A Formal Theory of Inductive Inference. Part I.* Information and Control, 7:1–22, 1964.

[119] Solomonoff, R. J.: *A Formal Theory of Inductive Inference. Part II.* Information and Control, 7:224–254, 1964.

[120] Stefanov, A., N. Gisin, O. Guinnard, L. Guinnard und H. Zbinden: *Optical Quantum Random Number Generator.* quant-ph/9907006, 1999.

[121] Stoler, David: *Photon Antibunching and Possible Ways to Observe It.* Phys. Rev. Lett., 33(23):1397–1400, December 1974.

[122] Takeuchi, S., T. Nagai, K. Hasegawa und Y. Hosono: *Hight Performance Random Pulser Based on Photon Counting.* IEEE Transactions on Nuclear Science, 33(1), 1986.

[123] Tang, C. L.: *Spontaneous and Stimulated Parametric Process.* In: Rabin, H. und C. L. Tang (Herausgeber): *Quantum Electronics.* Academic, New York, 1975.

[124] Tang, Qing und Xianyu Su: *Monte Carlo calculation for random numbers produced by an optical method.* Wuli, 16(6):349–352, 1987.

[125] Tundra: *RBG 1210.* http://www.tundra.com, unter der Rubrik *Data Security.*

[126] Walker, John: *HotBits: Genuine random numbers, generated by radioactive decay.* http://www.fourmilab.ch/hotbits/.

[127] Weihs, Gregor, Thomas Jennewein, Christoph Simon, Harald Weinfurter und Anton Zeilinger: *Violation of Bells's Inequality unter Strict Einstein Locality Conditions.* Physical Review Letters, 81(23):5039–5042, 1998.

[128] Weihs, Gregor, Michael Reck, Harald Weinfurter und Anton Zeilinger: *Two-photon interference in optical fiber multiports.* Phys. Rev. A, 54(1):893–897, 1996.

[129] Wooters, W. K. und W. H. Zurek: *A single quantum cannot be cloned.* Nature, 299:802–803, 1982.

[130] Yamamoto, Y., N. Imoto und S. Machida: *Amplitude squeezing in a semiconductor laser using quantum nondemolition measurement and negative feedback.* Phys. Rev. A, 33(5):3243 ff, Mai 1986.

[131] Yamamoto, Y., S. Machida, S. Saito, N. Imoto, T. Yanagawa, M. Kitagawa und G. Björk: *Quantum Mechanical Limit in Optical Precision Measurement and Communication*. In: Wolf, E. (Herausgeber): *Progress in Optics*, Band XXVIII, S. 87–179. Elsevier Science Publishers B.V., Amsterdam, 1990.

[132] Yariv, A.: *Quantum Electronics*. Wiley, New York, 1967.

[133] Zbinden, H., J. Brendel, N. Gisin und W. Tittel: *Experimental test of nonlocal quantum correlations in relativistic configurations*. Phys. Rev. A, 63(2):022111–1 –10, Januar 2001.

[134] Ziv, Jacob: *Coding Theorems for Individual Sequences*. IEEE Trans. Inf. Theory, 24(4):405–412, 1978.

[135] Ziv, Jacob und Abraham Lempel: *A Universal Algorithm for Sequential Data Compression*. IEEE Trans. Inf. Theory, 23(3):337–343, 1977.

Danksagung

Meinem hochgeschätzten Lehrer und Doktorvater Herrn Prof. Dr. Dr. Werner Martienssen gebührt mein besonderer Dank. Gab er mir doch die Gelegenheit dazu, mich in seiner Arbeitsgruppe mit dem interessanten Gebiet der Quantenoptik intensiver zu befassen. Für die Diskussionen bei der Anfertigung der Arbeit, seinen Rat und Zuspruch und seine Geduld, als sich der Abschluß berufsbedingt stark in die Länge zog, danke ich ihm herzlich.

Mein herzlicher Dank gilt auch Herrn Prof. Dr. Dr. Werner Dultz, der unsere Arbeitsgruppe über Jahre hinweg förderte und so viele Experimente und Arbeiten überhaupt erst ermöglicht hat. Seinem Engagement und seiner Vermittlung ist es auch zu verdanken, daß diese Arbeit wesentlich von der der Deutschen Telekom AG finanziell gefördert wurde. Für diese Unterstützung geht mein Dank an Herrn Dr. Gunter Laßmann und Herrn Dipl. Math. Gerhard Zesch, deren Interesse für quantenoptische Zufallsgeneratoren zu einem entsprechenden Forschungsvertrag führte, ohne den die experimentelle Realisierung kaum möglich gewesen wäre.

Meinen beiden Kollegen Herrn Zesch und Herrn Dr. Friedrich Fiedler möchte ich für ihren steten Ansporn danken, der wesentlich dazu beitrug, daß diese Dissertation fertig wurde.

Herrn Prof. Hermann Dinges danke ich für hilfreiche Diskussionen zu Fragen der Teststatistik.

Herrn Dr. Jürgen Brendel, Herrn Dr. Helmar Becker und Herrn Dr. Karsten Siebert danke ich für die kollegiale und angenehme Atmosphäre in der Arbeitsgruppe für Quantenoptik besonders herzlich. Herrn Dipl. Phys. Thomas Baumgart möchte ich für seine Hilfe bei mannigfaltigen elektronischen Problemen danken.

Meinem Kommilitonen und Kollegen Herrn Dipl. Phys. Ulf Linketscher gilt mein besonderer Dank für die Diskussionen zum Einsatz von Methoden der nichtlinearen Dynamik zur Analyse von Zeitreihen.

Der feinmechanischen Werkstatt unter Leitung von Herrn Herbert Hassenpflug sei ebenfalls gedankt. Die sorgfältige Anfertigung der diversen mechanischen Komponenten des experimentellen Aufbaus hatte nicht geringen Anteil am Gelingen der Versuche.

Frau Ingeborg Derlien möchte ich herzlichst dafür danken, daß ich, auch als ich nicht mehr am Physikalischen Institut beschäftigt war, die Max- Born- Bibliothek außerhalb der üblichen Öffnungszeiten weiterhin nutzen konnte.

Schlußendlich danke ich Frau Hannah Döring für ihre Hilfsbereitschaft bei verwaltungstechnischen Fragen in der Endphase der Promotion.

Eric Hildebrandt

Lebenslauf

Name:	Eric Hildebrandt
Geburtsdatum:	5.6.1964
Geburtsort:	Wiesbaden
Staatsangehörigkeit:	deutsch

Schulischer Werdegang

1970–1974	Johannes-Maas-Grundschule, Wiesbaden
1974–1983	Elly-Heuss-Gymnasium, Wiesbaden
31.05.1983	Abitur (Leistungskurse: Mathematik, Physik)

Ersatzdienst

7.6.1983	Verpflichtung zum achtjährigen Katastrophenschutzdienst beim Arbeiter Samariter Bund e.V. (Ortsverband Wiesbaden)

Studium

10/83	Immatrikulation an den Fachbereichen Physik und Ost- und Außereuropäische Sprach- und Kulturwissenschaften der Johann-Wolfgang-Goethe-Universität, Frankfurt a. M. Doppelstudium (Hauptfächer: Physik, Sinologie, Nebenfach: Chemie)
12.10.1985	Vordiplom in Physik
09/86	Exmatrikulation
10/86–07/88	Studium der chinesischen Philosophie, Sprache und Geschichte an der Nanjing University, Nanjing, Volksrepublik China (DAAD-Stipendium)
10/88	Neuimmatrikulation an der J.-W.-Goethe-Universität (Hauptfächer: s.o., Nebenfächer: Mathematik, Japanologie)
11.6.1993	Diplom in Physik, Diplomarbeit bei Prof. Dr. Dr. W. Martienssen Thema der Diplomarbeit: „Nichtklassische Lichtfelder im Mach-Zehnder-Interferometer“
7/93 bis 8/98	Promotionsstudium auf dem Gebiet der Quantenoptik im Physikalischen Institut der J.-W.-Goethe-Universität, Frankfurt am Main bei Herrn Prof. Dr. Dr. W. Martienssen
seit 9/98	Wissenschaftlicher Mitarbeiter im Technologiezentrum der Deutschen Telekom AG, (heute: Technologiezentrum, T-Systems) Bereich Sicherheit, Abteilung für technische Sicherheit