

**Fakultät für Physik and Astronomie
Universität Heidelberg**

Bachelorarbeit in Physik
eingereicht von

Simon Dario Tobias Ziegler

geboren in Hersbruck

2013

Optimierung der physikalischen Signalübertragung auf neuromorpher Hardware

Diese Bachelorarbeit wurde von Simon Ziegler angefertigt am
Kirchhoff-Institut für Physik in Heidelberg
unter der Betreuung von
Prof. Dr. Karlheinz Meier

Zusammenfassung:

Die korrekte Signalübertragung in der neuromorphen Hardware ist von zentraler Bedeutung für die Aussagekraft der durchgeführten Emulationen. In dieser Arbeit werden die optimierten Parameter für die Hardware durch direkte Messungen bestimmt. Ziel ist es eine möglichst fehlerfreie Übertragung der digitalen Aktionspotenziale durch Einstellung der Kommunikationsspannungen zu erreichen. Darüber hinaus werden die zur Weiterleitung benötigten Repeater und deren Verbindungen zueinander untersucht.

Abstract:

The correct signal transfer in neuromorphic hardware is essential for the quality of the realised emulation. In this thesis the optimum parameters for the hardware are determined via direct measurements. The aim is to reach a transmission of digital spike signals free of errors by modulation of the communication voltages. In addition the repeaters and their connections, which are required for transmission, are analyzed.

Inhaltsverzeichnis

1	Einleitung	7
1.1	Forschungsprojekt	7
1.2	Zielsetzung	7
2	Technische Grundlagen	8
2.1	HICANN Chip	8
2.2	Signalübertragung	8
2.3	Spannungen der Kommunikationsstruktur L1	9
2.4	Repeater	10
2.5	Intra- und Inter-HICANN Kommunikation	10
2.6	Vertikal-Setup und Wafer	11
3	Methoden der Intra-HICANN-Kommunikation	13
3.1	Aufbau der Kommunikationsstruktur	13
3.2	Messroutine	13
3.3	Messung der Hardware Parameter	14
4	Resultate	15
4.1	Einstellungen der DLL-Schaltung	15
4.2	Optimierung der Repeater Spannung	16
4.3	Unregelmäßige Wechselwirkung der L1 Komponenten	17
4.4	Optimierung der L1 Spannungen am Vertikal-Setup	17
4.5	Optimierung der L1 Spannungen am Wafers	19
4.6	Verbindungen der einzelnen Repeaterpaare	21
4.7	Fehlerrate in Abhängigkeit der Neuronadressen	22
4.8	Frequenzabhängigkeit der Übertragungsqualität	24
4.9	Fehlerquelle durch empfangene Null-Events	26
5	Ausblick und Diskussion	27

1 Einleitung

Der US Autor und Satiriker Ambrose Bierce erklärte das Gehirn mit folgenden Worten:

”Brain: an apparatus with which we think we think.”

Mit den Arbeiten von Alan Lloyd Hodgkin und Andrew F. Huxley und vielen folgenden Forschern in den Neurowissenschaften konnte das physikalische Verhalten von einzelnen Neuronen und Synapsen erfasst werden. Eine hinreichende Erklärung für das allgemein bekannte ”Denken”, wie es Bierce wahrnahm, lieferte es jedoch nicht.

Es scheint, das Gehirn mit seinen etwa 100 Milliarden Neuronen und 1000 Billionen Synapsen entfaltet seine Wirkung erst bei größeren Skalen.

1.1 Forschungsprojekt

Das EU Projekt BrainScaleS [1] wurde 2010 gegründet mit dem Ziel größere neuronale Netzwerke zu erforschen. Im Gegensatz zu ihrem Projektpartner Blue Brain in Lausanne [2], der Modelle neuronaler Netzwerke in numerischen Simulation mit Supercomputern untersucht, verfolgt die Forschungsgruppe Electronic Visions [3] in Heidelberg einen anderen Ansatz, die Emulation.

Da bei Simulationen die Rechenzeit und der Energieverbrauch nicht linear mit der Größe des Netzwerk zunehmen, ist die Emulation von neuronalen Netzwerken mit Silizium-Chips eine Alternative. Diese sogenannte neuromorphe Hardware erhält wegen Chip-internen Elektronenströme gegenüber den biologischen Ionenströmen des Gehirns einen Beschleunigungsfaktor in der Verarbeitungsdauer.

1.2 Zielsetzung

In dieser Arbeit werden die Spannungsparameter (siehe Kapitel 4.4) für die Chip-interne Kommunikation für eine möglichst fehlerfreie Signalübertragung optimiert. Dazu werden direkte Messungen an der neuromorphen Hardware HICANN [4] durchgeführt und ausgewertet.

2 Technische Grundlagen

2.1 HICANN Chip

Der *High Input Count Analog Neural Network Chip* (HICANN) enthält 512 Neuronen, circa 115.000 Synapsen, Speicherelemente (*Floating Gates*) und Kommunikationsstrukturen, deren Lokalisation in Abbildung [1] zu erkennen ist. Die Neuronen sind analoge Schaltungen [5], die sich für unterschiedliches Verhalten mit den in den *Floating Gates* gespeicherten Parametern konfigurieren lassen. Vergleichbar mit biologischen Netzwerken können diese Neuronen feuern und somit weitere Neuronen anregen. In dieser Arbeit wird nur die Kommunikationsinfrastruktur Layer 1 (L1) von Bedeutung sein, für Details zu den anderen Komponenten siehe [6].

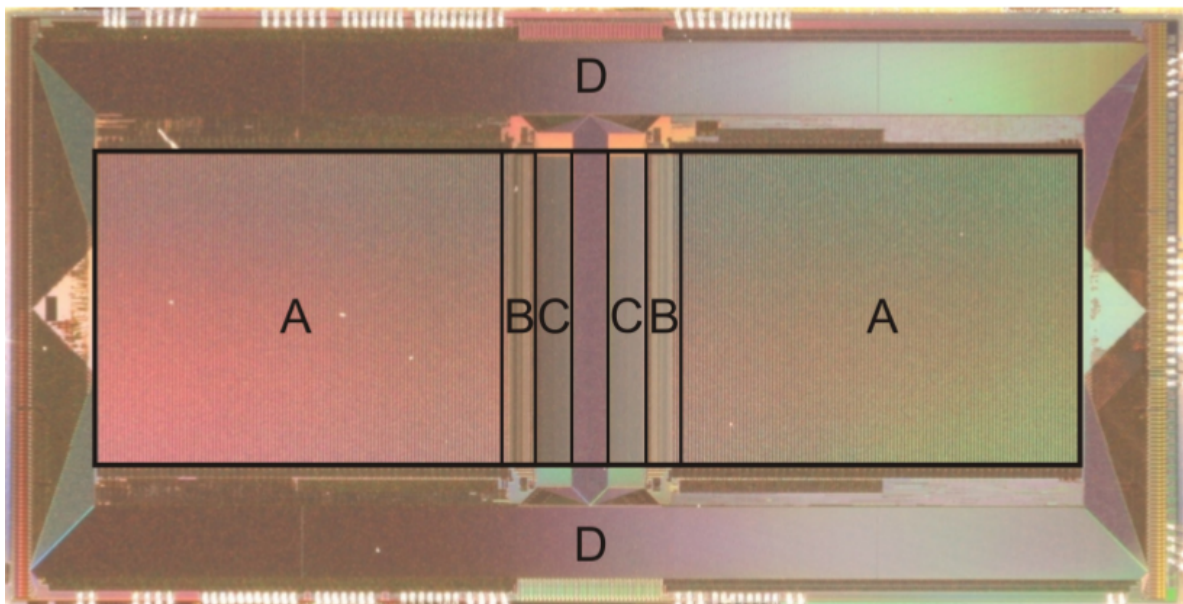


Abbildung 1: Fotografie eines HICANN Chip: (A) Synapsen-Feld (B) Neuronen (C) Floating Gate Speicher (D) L1-Netzwerk. Aus [7]

Das L1 Kommunikationsnetzwerk verbindet feuernde Neuronen (bzw. die externe Anregung) über Synapsen mit anderen Neuronen. Dafür benutzt es eine asynchrone serielle Übertragung von digitale Paketen.

2.2 Signalübertragung

Um Verbindungen zwischen Neuronen in der gemeinsamen Kommunikationsinfrastruktur L1 herzustellen, werden Feuersignale als 6 Bit-Events verschickt, die zu einer von 64 Neuronenadressen dekodiert werden können [8]. Da ein asynchrones Protokoll verwendet wird, muss der Code noch mit jeweils einem 0-Bit als Start- und Endbit versehen werden. Die *Sending Repeater* leiten diese Signale in einen von 64 horizontalen Bussen des Layer 1 Netzwerks, der über ein *CrossbarSwitch* mit einem von 256 vertikalen Bussen

verbunden werden kann. Über ein *SynapseDriverSwitch* können die Signale vom Bus den *Synapse Driver* erreichen, der die ersten beiden Bits der Adresse dekodiert und die restlichen 4-Bit in eine der vier Leitungen zu einer Zeile Synapsen schickt. Die Synapse mit der entsprechenden 4 Bitadresse schickt das Signal weiter zum Neuron.[9]

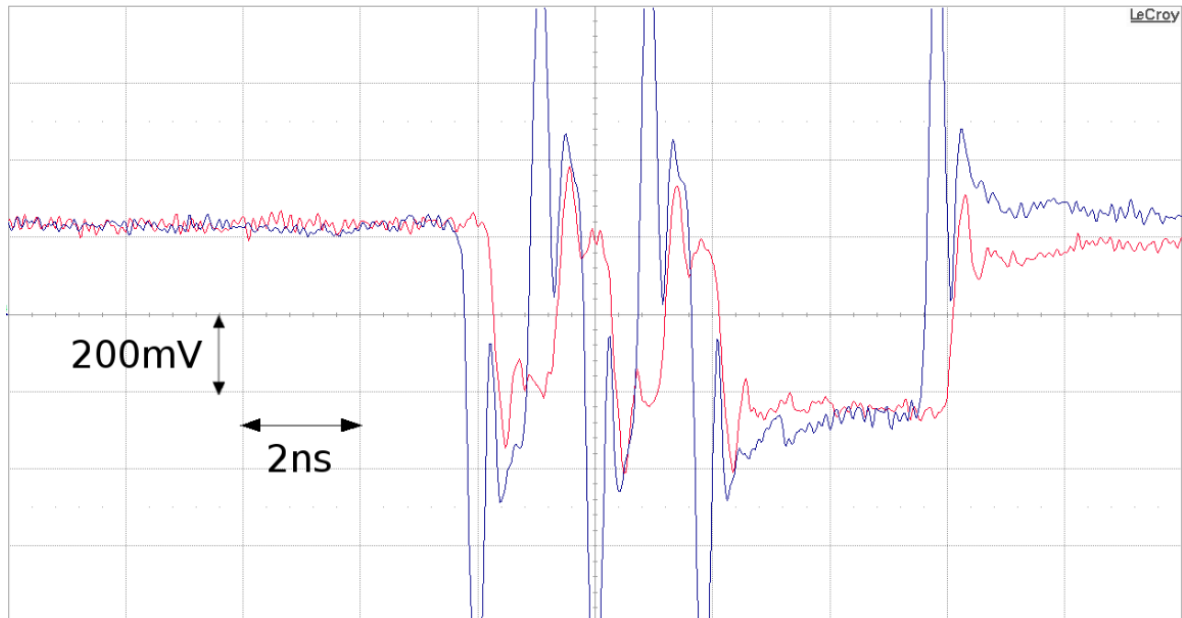


Abbildung 2: Analoge Aufnahme des digitalen 6-Bit-Signals 101000 mit Start- und Endbit: Vergleich des Spannungsverlauf beim *Sending Repeater* (blau) und nach 1cm Übertragungsweite in L1 (rot). Die Präemphase (lokales Maximum bzw. Minimum) ist am Anfang eines Bit beim Sender klar zu erkennen, beim Empfänger hingegen deutlich abgeschwächt. Aus [10]

Bits werden idealerweise als rechteckige Spannungspulse übertragen. Diese verändern ihre Form mit der Übertragungsweite. Deshalb müssen sie auf ihrer Übertragungsstrecke mit *Repeatern* gelesen und neu als digitale Impulse erzeugt werden. In der Abbildung [2] erkennt man die Abnahme der Amplitude und die Formveränderung des ankommenden Signals. Dadurch kann es in extremerer Form zu *Bitflips* kommen: Ein *Repeater* interpretiert ein oder mehrere Bits falsch und schickt eine inkorrekte Adresse weiter. Die Reduzierung der *Bitflips* bei der L1-Übertragung stellt die zentrale Aufgabe dieser Arbeit dar.

2.3 Spannungen der Kommunikationsstruktur L1

Für einen möglichst geringen Energieverbrauch wurde für das L1 Bussystem das *low voltage differential signaling* (LVDS)[10] implementiert. Das (für einen Bus) aus zwei Drähten bestehende System überträgt komplementäre Signale mit kleinen Spannungsamplituden.

Es wird nur der Spannungsunterschied zwischen den Drähten detektiert. Dieser wird

kaum von außen gestört, da die Störfaktoren gleichermaßen auf beide Drahtspannungen einwirken. Damit ist eine deutlich kleinere Spannung zur Signalübertragung nötig, als bei herkömmlichen Leitungen. Die zwei Drahtspannungen V_{OH} und V_{OL} sind an der Hardware einstellbar. Dabei gilt:

$$\Delta V = V_{OH} - V_{OL} \quad (1)$$

$$V_{CM} = \frac{V_{OH} - V_{OL}}{2} + V_{OL} \quad (2)$$

Wobei der Energieverbrauch bei Signalübertragung proportional zu ΔV^2 ist. Die *Common Mode* Spannung V_{CM} entspricht der mittleren Spannung der beiden Drähte ohne übertragene Signale.

2.4 Repeater

Der HICANN Chip enthält 6 Repeaterblöcke. 4 Blöcke befinden sich mit jeweils 64 *Repeatern* bei den vertikalen Bussen am oberen und unteren Chiprand. 2 Blöcke sind mit jeweils 32 *Repeatern* bei den horizontalen Busse mittig am linken und rechten Rand gelegen. In Abbildung [3] wird deutlich, dass somit jeder zweite Bus in jedem Block mit einem *Repeater* versehen ist. Da die *Repeater* im gegenüberliegenden Block jeweils auf dem freien Bus sitzen, werden durch benachbarte HICANN Chips die Signale aller Busse am Rand der Chips erneuert.

Nicht zu erkennen in der schematischen Abbildung [3] ist, dass die Busse zwischen zwei HICANN Chips um 2 Busse permutiert werden um die Einschränkung der *CrossbarSwitches* (siehe Kapitel 2.5) auszugleichen.

Damit die *Repeater* die Signale empfangen und senden können, benötigen sie eine Zeitreferenz, die ihnen die Periodendauer eines Bits vorgibt. Da sie nicht an einer globalen *Clock* angeschlossen sind, muss die Clockperiode aus den Signalen selbst mittels DLL (*Delay Locked Loop*) Schaltung eingestellt werden. Durch Senden langer Rechteckpulse mit 8-Bit Länge (= Neuronenadresse Null) werden die *Repeater* gelockt, das heißt eine Zeitreferenz für die Bitperiode vorgegeben. Da dieser Wert aufgrund von Leckströmen sich mit der Zeit verändert, werden nach der Lockzeit kontinuierlich Null Adressen in Layer 1 geschickt um die erreichbaren *Repeater* weiter einzustellen. Damit die DLL Schaltung eingestellt werden kann, wird sie vor der Signalübertragung auf die Spannung V_{allres} resetet.

Jeder *Repeater* enthält außerdem einen auslesbaren Speicher für drei empfangene Events (6-Bit Adresse und Ankunftszeit).

2.5 Intra- und Inter-HICANN Kommunikation

Die externe Anregung des HICANN Chip wird ermöglicht durch den *Background Generator* [7], der 6 Bit Adresssignale generiert und diese zu einem von den 8 *Sending Repeater* im linken mittleren Block leitet. Diese *Sending Repeater* sind wie die anderen *Repeater* fest mit einem Bus verbunden, in das sie das Signal einspeisen. Damit

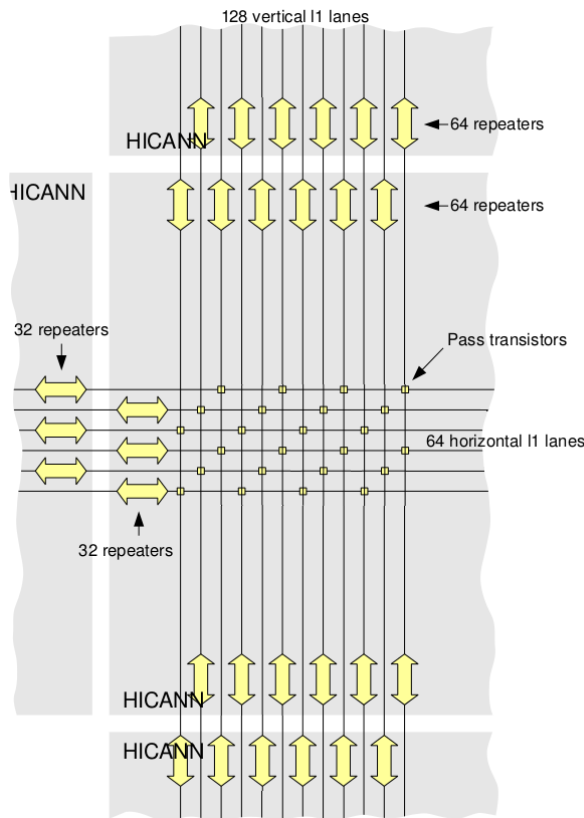


Abbildung 3: Schematische Darstellung des linken Bussystem eines HICANN Chip mit 3 Repeaterblocks und 3 angrenzenden Chips: Jeder der vertikalen (bzw horizontalen) Busse besitzt einen *Repeater*, sodass propagierende Signale regelmäßig erneuert werden. Aus [10]

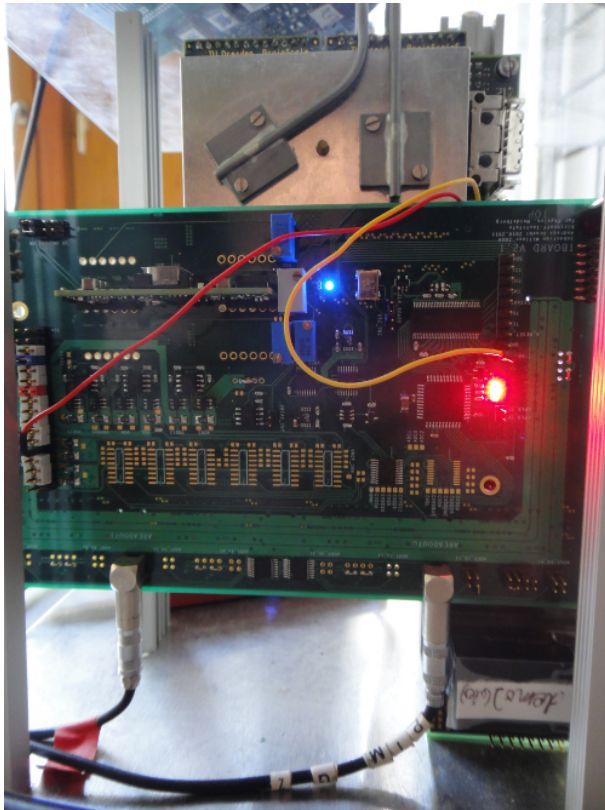
die L1 Kommunikationsstruktur keine zu große Kapazität durch Transistoren erhält, kann ein horizontaler Bus durch *CrossbarSwitches* nur mit 8 von 256 vertikalen Busse verbunden werden. Nach der Chip-Architektur sind die *Sending Repeater* so angeordnet, dass alle über L1 nur mit den selben 8 *Repeatern* verbunden werden können. Diese Empfänger-*Repeater* liegen im linken und rechten unteren Block. Somit sind für Intra-Chip-Kommunikation insgesamt 64 Verbindungen möglich und es werden 16 von 320 *Repeatern* benutzt.

Da an jeder Seite eines HICANN Chips weitere Chips angeschlossen werden können und die Busse permutiert werden, ergibt sich für die Inter-Chip-Kommunikation eine Vielzahl von Verbindungsmöglichkeiten über mehrere Repeater.

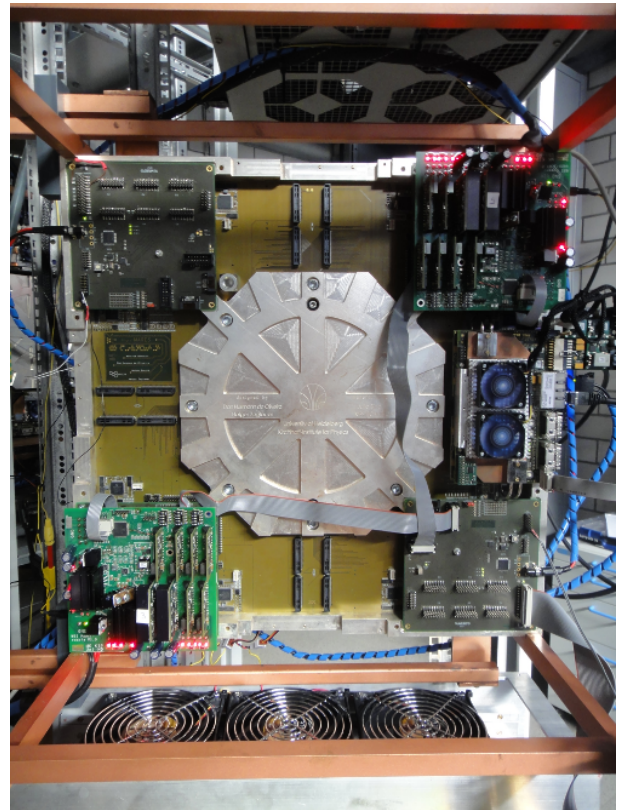
2.6 Vertikal-Setup und Wafer

Die neuromorphe Hardware unterscheidet sich fundamental von gewöhnlicher digitaler Hardware. Ihre Arbeitsweise ist dezentral und höchst parallel in den einzelnen Neuronen, deren Signale zeit-kontinuierlich übertragen werden. Um ihr Verhalten untersuchen zu

können, muss eine Schnittstelle zwischen neuromorpher Hardware und lesbarer Ausgabe geschaffen werden. Des Weiteren benötigt der Chip einstellbare Strom- und Spannungsquellen und Eingänge zur Konfiguration der neuromorphen Hardware.



(a) Aufbau des Vertikal-Setup



(b) Aufbau des Wafers

Abbildung 4: Fotografien der verwendeten Aufbauten mit ihren Versorgungselementen. Die HICANN Chips liegen verdeckt von (a) der rechten unteren schwarzen Box mit Aufschrift (b) der 8-eckigen Metallplatte

Das Vertikal-Setup ist eine Versorgungs- und Kommunikationsplattform für mehrere HICANN Chips, die unterschiedliche Messungen, d.h. Ein- und Ausgaben, ermöglicht. In dieser Arbeit wurde am Setup mit nur einem HICANN Chip gearbeitet.

Der Silizium Wafer [11] besteht aus 48 Retikeln, die jeweils 8 HICANN Chips enthalten und somit Inter-Chip Kommunikation ermöglichen. Die 8 HICANN Chips auf dem Retikel werden während der Herstellung miteinander verbunden. Die verschiedenen Retikeln hingegen werden erst in einem *post-processing* Verfahren verschalten [10]. Somit kann man Inter-HICANN Kommunikation über maximal 384 Chips betreiben.

Ein weiteres Ziel der Forschungsgruppe ist das Verbinden von mehreren Wafern zu größeren Netzwerken.

3 Methoden der Intra-HICANN-Kommunikation

3.1 Aufbau der Kommunikationsstruktur

Um die Übertragungsqualität einer Verbindung zwischen zwei *Repeatern* zu messen, muss zuerst die Verbindung hergestellt und die Repeater konfiguriert werden. Dazu wird der Bus des *Sending Repeaters* durch ein *CrossbarSwitch* mit einem vertikalen Bus des Empfänger-*Repeater* verbunden. Danach wird die DLL-Schaltung dieses *Repeaters* auf die Spannung V_{dllres} resettet, wozu die Zeit t_{reset} benötigt wird. Durch gesendete Null-Events wird dann der Empfänger-*Repeater* gelockt. Für einen stabileren Lockvorgang wird die Zeit t_{lock} bis zum Start der Messung abgewartet.

3.2 Messroutine

Zur Bestimmung der Übertragungsqualität werden die *broken events*, der prozentuale Anteil der inkorrekt empfangenen Adressen, gemessen. Dies entspricht der Fehlerrate der Verbindung.

Dazu schickt der *Sending Repeater* periodisch eine bestimmte Adresse in den Layer 1 Bus, der zu einem *Repeater* führt. Dieser Empfänger-*Repeater* nimmt in einer vorgegebenen Zeit drei Events auf und speichert sie. Diese werden ausgelesen und mit der Sendeadresse verglichen, bei Übereinstimmung wird x_{right} erhöht, ansonsten x_{wrong} . Dieser Vorgang wird 5-10 mal wiederholt für eine Einzelmessung, d.h. 15-30 mal wird eine Adresse getestet aus deren Ergebnissen sich der Prozentsatz der *broken events* berechnet.

$$broken\ events = \frac{x_{wrong} * 100}{x_{right} + x_{wrong}} \quad (3)$$

Dabei werden ankommende Nullen nicht berücksichtigt (außer bei Sendeadresse Null), da diese für das DLL-Locking benötigt werden und deshalb nicht als fehlerhaft betrachtet werden.

Die Messung für Intra-HICANN-Kommunikation verläuft über mehrere Stufen, bei denen die unteren Stufen in einer Schleife wiederholt nacheinander ausgeführt werden.

- Messung der *broken events* einer spezifischen Adresse
- Messung über Adressen 0 bis 63
- Messung über 64 Verbindungsmöglichkeiten der 8 *Sending Repeater* und 8 Receiver Repeater
- Messung über 8 HICANN Chips eines Retikels auf dem Wafer

Die vollständige Messroutine entspricht somit 4096 Messungen einer spezifischen Adresse für einen HICANN Chip (Setup) und 32768 Messungen für 8 HICANN Chips (Wafer). Um Messzeit auf dem Wafer zu sparen, wurde teilweise statt über alle 64 Intra-Verbindungen nur über 8 Intra-Verbindungen gemessen. Jedoch wurden alle 8 *Sending* und Empfänger-*Repeater* angesteuert und alle HICANN Chips eines Retikels vermessen.

3.3 Messung der Hardware Parameter

Um eine zuverlässige Übertragung zu gewährleisten, werden die optimalen Parameter-einstellungen gesucht. Dazu werden die Parameter variiert und mit der Messroutine die durchschnittlichen $\langle broken\ events \rangle$ gemessen. Damit wird das Optimum der folgenden Parameter bestimmt:

- die DLL Resetspannung V_{dllres}
- die Repeaterspannung V_{ccas}
- die L1 Spannungen V_{OH} und V_{OL}

Die Repeaterspannung V_{ccas} bestimmt mit welcher Verstärkung die Spannungsunterschiede aus L1 aufgenommen werden und wie der Repeater sich auf V_{CM} (siehe Gleichung 2) einstellt.

Da die Messroutine viele Einzelmessungen benötigt, kann die Vermessung der Hardware Parameter bis zu mehreren Tagen dauern.

4 Resultate

Die Resultate der Messungen am Vertikal-Setup und am Wafer werden im Folgenden präsentiert. Da die Ergebnisse der Messungen am Wafers von zentralerer Bedeutung sind, jedoch die Messzeit dafür zeitlich stark beschränkt war, müssen die Messungen am Wafer teilweise weitergeführt werden.

4.1 Einstellungen der DLL-Schaltung

Die Einstellung der DLL-Schaltung ist bedeutend für das korrekte Empfangen der Signale (siehe Kapitel 2.4) Für ein genaueres *Locken* der Repeater benötigt die DLL eine längere Resetzeit und eine längere Lockzeit als bisher angenommen.

Für zeiteffizientere Messungen wurden die Zeiten t_{reset} , t_{lock} am Setup vermessen. Jedoch ergibt sich, dass die HICANN Chips sehr unterschiedlich auf diese Zeiten reagieren. Zwei HICANN Chips auf dem untersuchten Wafer-Retikel benötigen eine deutlich längere Zeit als der Chip am Setup. Dies könnte an der Spannungsversorgung des Wafers liegen, die sich an die neue Konfiguration anpassen muss. Die Zeiten für den Wafer werden auf den hohen Wert $t_{lock} = t_{reset} = 0.1$ s gesetzt, bei denen ein sicheres Locken der beiden HICANN Chips auf dem Retikel gewährleistet ist.

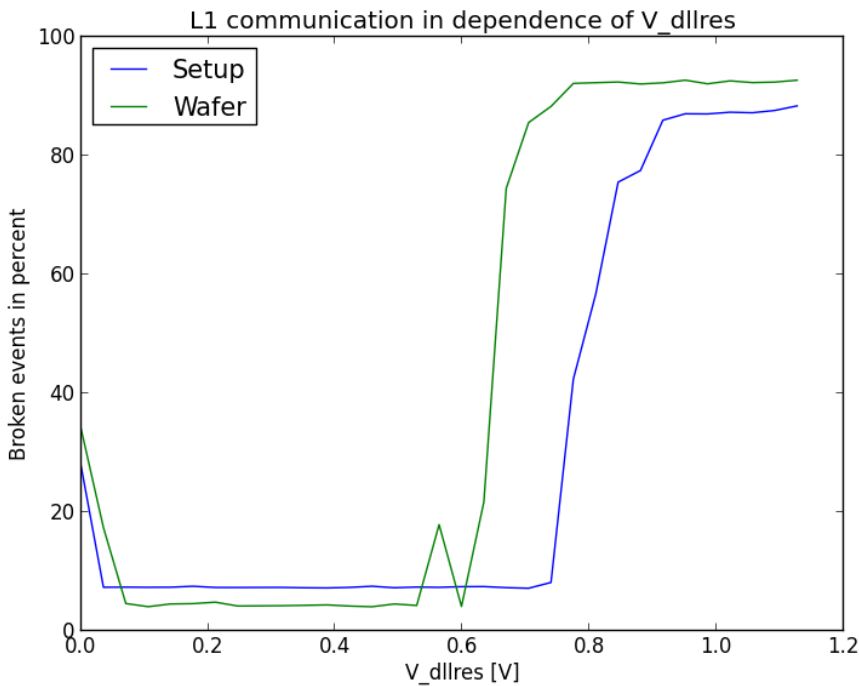


Abbildung 5: Messung der Fehlerrate in Abhängigkeit der DLL-Resetspannung V_{dllres} in 0.035 V Schritten mit eingestelltem $V_{OH} = 1.05$ V, $V_{OL} = 0.75$. Am Setup wird mit Frequenz = 100 MHz und am Wafer mit 150 MHz gemessen

Für die Bestimmung der DLL-Resetspannung V_{dllres} wurde auf dem Wafer und Setup

die Abhängigkeit der $\langle broken\ events \rangle$ von der Resetspannung V_{dllres} gemessen. In Abbildung [5] wurden auf dem Wafer jeweils 8 Intra-Verbindungen der 8 HICANN Chips auf einem Retikle und auf dem Setup 64 Intra-Verbindungen auf einem HICANN Chip vermessen. Das optimale Plateau des Setup ist deutlich kleiner als das des Wafers. Als optimaler Wert für V_{dllres} wird die Mitte des Plateaus des Wafers mit $V_{dllres} = 0.39\text{ V}$ gewählt.

4.2 Optimierung der Repeaterspannung

In Abbildung [6] ist die Repeaterspannung V_{ccas} gegen die Fehlerrate aufgetragen, wobei die Fehlerbalken durch die Standardabweichung von 4 Messungen auf dem Setup bestimmt wird.

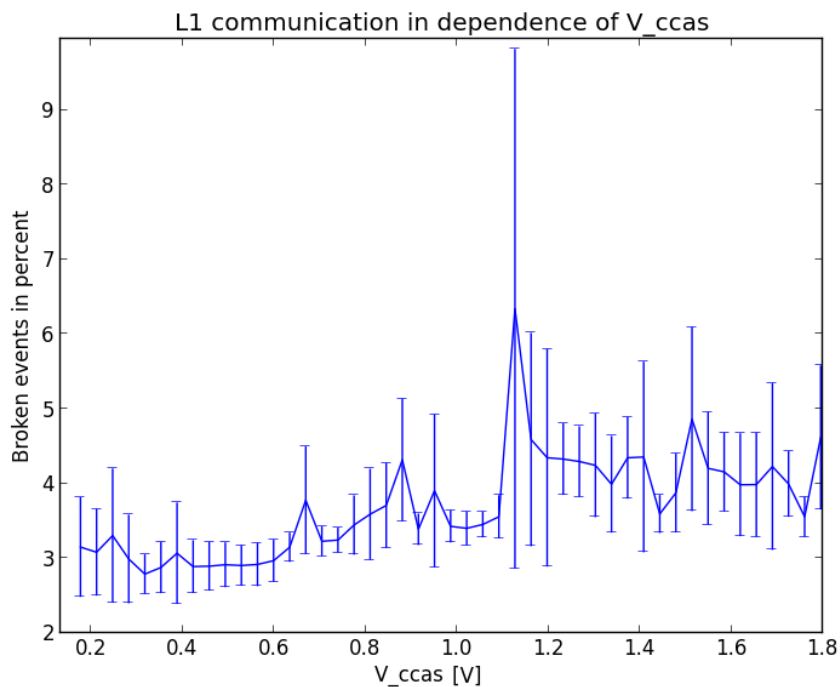


Abbildung 6: Vier Messungen der Fehlerrate am Setup in Abhängigkeit der Repeaterspannung V_{ccas} in 0.035 Volt Schritten. Die Messungen schwanken stark in höheren Spannungsbereichen. Der Anstieg bei $< 0.1\text{ V}$ auf 100 % wurde nicht vermessen

Die einzelnen Messpunkte schwanken stark und zeigen einen unerwarteten Verlauf. Dies könnte an den Komponenten des Vertikal-Setup liegen. Um diesen Effekt besser zu verstehen sollte mit verschiedenen HICANN Chips V_{ccas} am Setup untersucht werden. Der optimale Bereich von V_{ccas} für das Setup wird mit 0.4 V identifiziert und für weitere Messungen übernommen. Bei 0 Volt erreicht die Fehlerrate 100 %.

Die Untersuchung am Wafer ergab eine klarere Abhängigkeit von der Spannung V_{ccas} , siehe Abbildung [7]. Die Fehlerrate schwankt ab einer Spannung von 0.3 V um ein Plateau. Bei kleineren Spannungen zeigt sich eine Abweichung zu schlechteren Übertragungsraten.

Dieser Verlauf wird erwartet. Der *Repeater* kann ab einer bestimmten Spannung die empfangenen Daten mit der Spannungsverstärkerung decodieren.

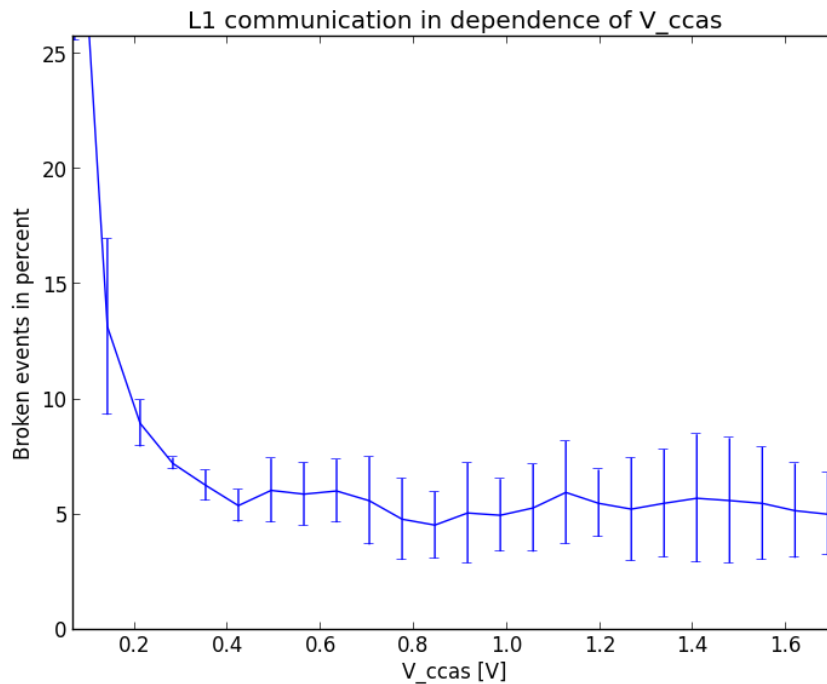


Abbildung 7: Drei Messungen der Fehlerrate am Wafer mit 8 HICANN Chips in Abhängigkeit der Repeaterspannung V_{ccas} in 0.07 Volt Schritten und $V_{OH} = 0.9$ V und $V_{OL} = 0.65$ V

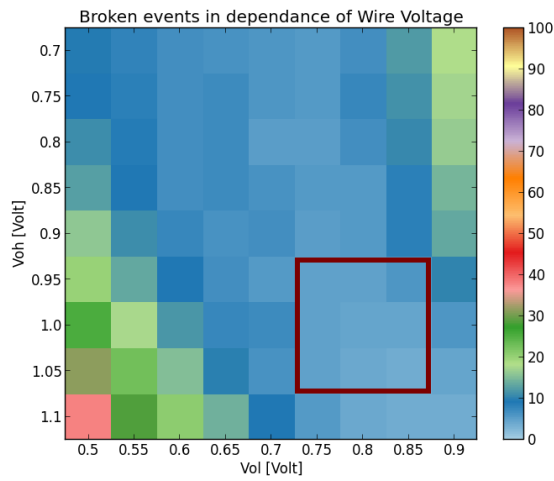
Für V_{ccas} wird der hohe Wert 1,7 V gewählt um mögliche Fehler auszuschließen. Dieser Wert kann im Laufe dieses Projekt zur Energieeinsparung herabgesetzt werden.

4.3 Unregelmäßige Wechselwirkung der L1 Komponenten

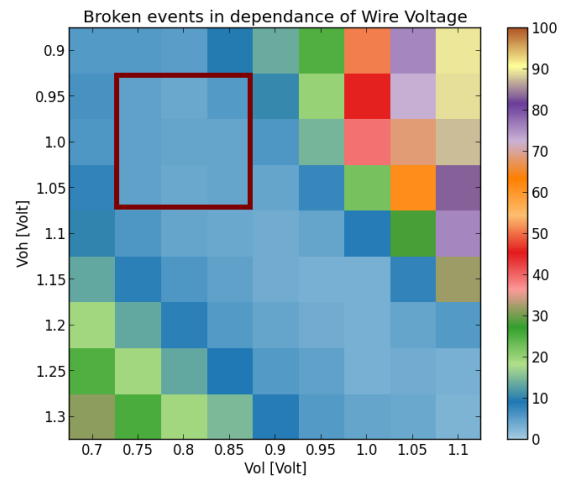
Im Laufe der Arbeit wurde entdeckt, dass die Übertragungsqualität der Verbindungen sich signifikant verbessert, falls in der Messroutine vor jeder Einzelmessung die *CrossbarSwitches* und der Background Generator auf Null zurückgesetzt werden. Jedoch wird dies schon beim Start der Messung ausgeführt, sodass weitere Ausführungen nicht nötig sein sollten. Da außer Energieverbrauch kein Zusammenhang erkennbar ist, ist dieses Phänomen unverstanden und sollte weiter untersucht werden.

4.4 Optimierung der L1 Spannungen am Vertikal-Setup

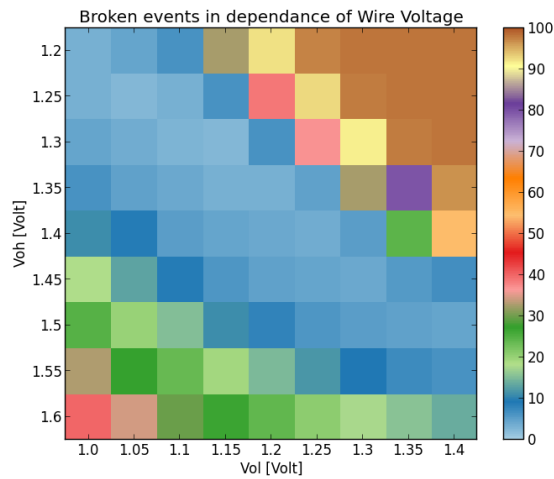
Mit verschiedenen Spannungen V_{OH} und V_{OL} , wie in Kapitel 3 beschrieben, wird der Chip auf dem Vertikal-Setup vermessen und die Ergebnisse in Abbildung [8] dargestellt.



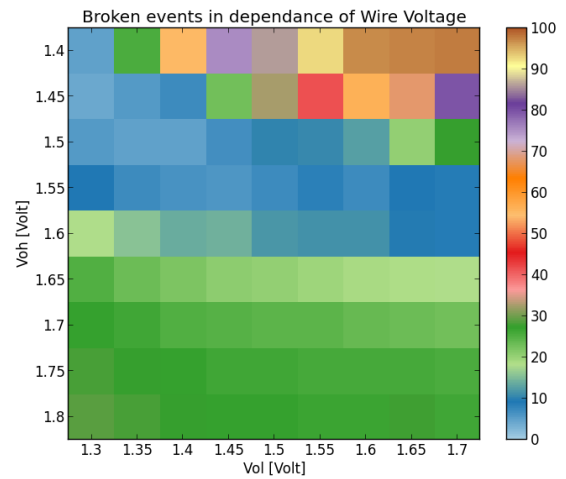
(a) $V_{OL} = 0.5 \text{ V}$ bis 0.9 V
 $V_{OH} = 0.7 \text{ V}$ bis 1.1 V



(b) $V_{OL} = 0.7 \text{ V}$ bis 1.1 V
 $V_{OH} = 0.9 \text{ V}$ bis 1.3 V



(c) $V_{OL} = 1.0 \text{ V}$ bis 1.4 V
 $V_{OH} = 1.2 \text{ V}$ bis 1.6 V



(d) $V_{OL} = 1.2 \text{ V}$ bis 1.7 V
 $V_{OH} = 1.4 \text{ V}$ bis 1.8 V

Abbildung 8: Übersicht der durchschnittlichen Fehlerrate bei Intra-HICANN Kommunikation in Abhängigkeit der Spannungen V_{OL} und V_{OH} am Vertikal-Setup mit der Frequenz = 100 MHz. Der rot markierte Bereich ist die Vorauswahl für weitere Messungen auf dem Wafer.

Gesucht wird die Spannungscombination mit der besten Übertragungsqualität, d.h. mit minimaler durchschnittlicher Fehlerrate. In Abbildung [8] sind mehrere Spannungscombinationen mit minimaler Fehlerrate (erkennbar an den *hellblauen Feldern* der Abbildung) bis zu größeren Spannungen möglich. Es scheint, dass der Spannungsunterschied ΔV die Fehlerrate bestimmt und dass weder zu kleine, noch zu große Unterschiede der Signalübertragung förderlich sind. Dieser optimale Unterschied verändert sich mit steigender V_{CM} . Erstaunlich bei kleineren Spannungswerten ist, dass auch die Einstellung $V_{OH} = V_{OL}$ eine gute Übertragung gewährleistet. Diese Einstellung sollte theoretisch keine Signalübertragung ermöglichen. Jedoch kann anscheinend die Präemphase, die den einzelnen Bits vorausgeht, für den Empfänger detektierbar sein und so eine Übertragung gewährleisten.

Da auf dem jetzigen Wafer keine Spannungen oberhalb von 1.1 V einstellbar sind, wird nun ein optimaler Bereich rot markiert, der nochmals mit mehr Statistik auf dem Wafer vermessen wird.

4.5 Optimierung der L1 Spannungen am Wafers

Die Resultate der Messungen am Wafers sind von zentralerer Bedeutung, da dieser bei späteren Emulationen von neuronalen Netzwerken benutzt wird. Weil der optimale Spannungsbereich am Vertikal-Setup sich nicht mit denen des Wafers deckt, muss ein neuer Spannungsbereich gesucht werden.

Durch die längere benötigte Reset- und Lockzeit (siehe Kapitel 4.1) und die höhere Statistik durch 8 HICANN Chips, konnte im Rahmen dieser Arbeit leider nicht der komplette optimale Spannungsbereich vermessen werden.

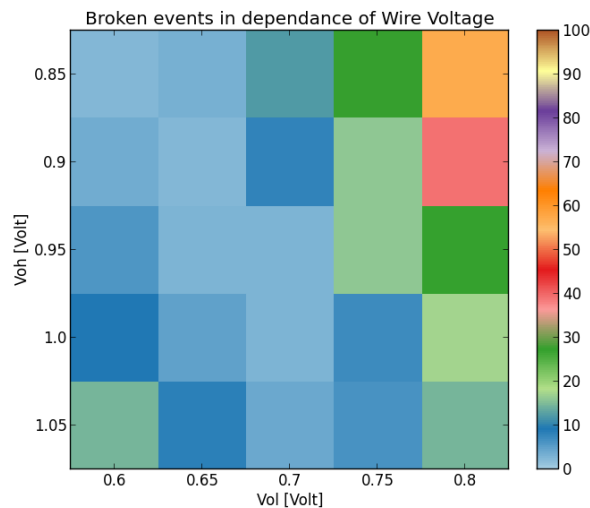


Abbildung 9: Übersicht der durchschnittlichen Fehlerrate bei Intra-HICANN Kommunikation in Abhängigkeit der Spannungen V_{OL} und V_{OH} auf dem Wafer

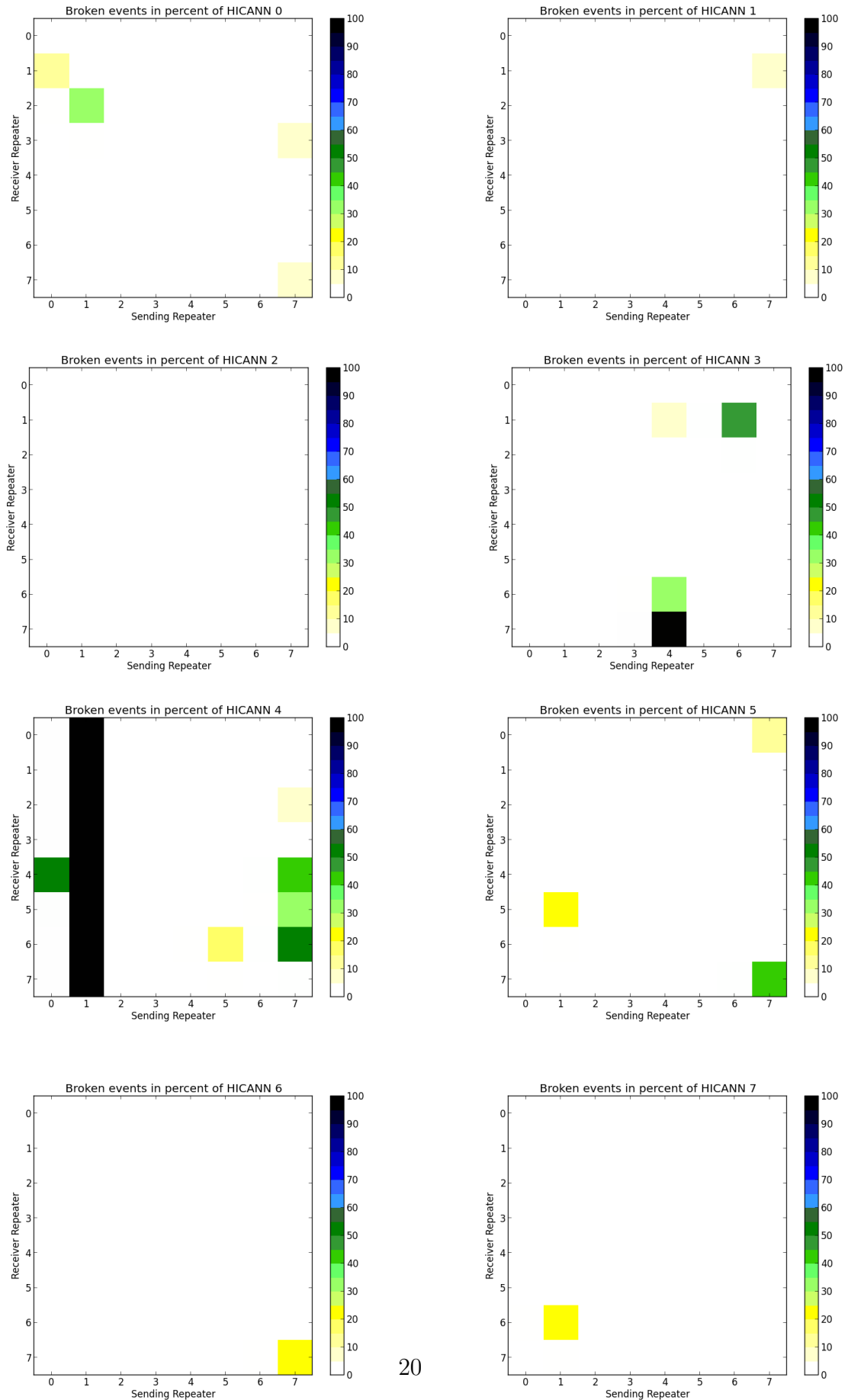


Abbildung 10: Übersicht der Fehlerrate der Intra-HICANN Verbindungen mit $V_{OL} = 0.9 \text{ V}$ und $V_{OH} = 0.65 \text{ V}$ für alle 8 HICANN Chips auf einem Retikel mit Frequenz 150 MHz.

Die Messung am Wafer wird mit der Frequenz 150 MHz durchgeführt, da Voruntersuchungen ergaben, dass die Fehlerrate sich gegenüber denen mit 100 MHz verbessern (siehe Kapitel 4.8). In Abbildung [9] sind die Ergebnisse der Spannungsmessung auf dem Retikel des Wafers aufgetragen. Diese Messung benötigte circa 70 Stunden. Der Spannungsbereich ist nicht komplett, da an seinem Rand weitere gute Einstellungen möglich sind. Die Spannungskombination $V_{OL} = 0.9$ V und $V_{OH} = 0.65$ V mit $\langle broken\ events \rangle = 2.702$ % ist das optimale Ergebnis.

Mit diesen Spannungen ist bei den einzelnen Intra-HICANN Verbindungen der 8 HICANN Chips in Abbildung [10] erkennbar, dass die meisten Verbindungen eine Fehlerrate von 0.0 % aufweisen. Darüber hinaus besitzt der HICANN 2 eine perfekte Übertragungsqualität mit durchschnittlicher Fehlerrate von 0.0 %. HICANN 4 dagegen besitzt einen nicht funktionalen *Sending Repeater*. Falls dieser *Repeater* nicht benutzt wird, verbessert sich die durchschnittliche $\langle broken\ events \rangle$ des Retikels deutlich auf 1.13 %. Mit dieser Einstellung ist eine weitgehend fehlerfreie Übertragung möglich.

4.6 Verbindungen der einzelnen Repeaterpaare

Zur Überprüfung des Messverfahrens und zur Untersuchung der L1-Kommunikation wurden bei einigen Messungen zusätzliche Plots angefertigt, die die einzelnen Verbindungsqualitäten in unterschiedlichen Stufen der Messroutine (siehe Kapitel 3) beschreiben. Exemplarisch zeigt die Abbildung [11] die Fehlerrate der 64 Intra-HICANN Verbindungen für die Adresse 11.

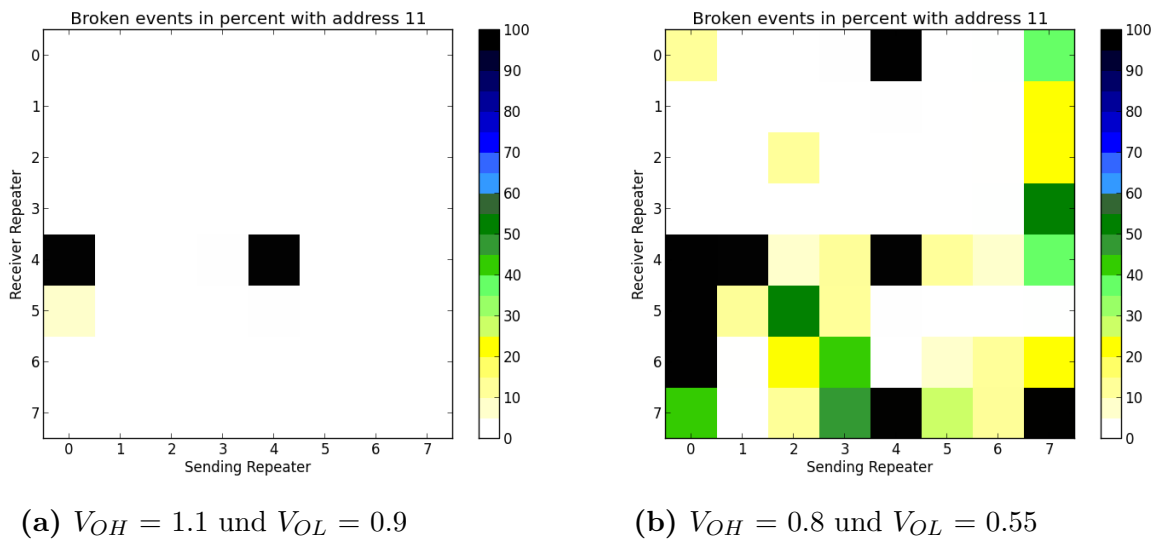


Abbildung 11: Übersicht der Fehlerrate der Intra-HICANN Verbindungen mit gesendeter Adresse 11 für zwei unterschiedliche Spannungseinstellungen auf dem Setup mit Frequenz = 100 MHz

Der Plot 11a enthält deutlich weniger fehlerhafte Events als Plot 11b aufgrund einer anderen Spannungseinstellung. Die Verbindung 0-4 und 4-4 haben bei allen Adressen

und Spannungseinstellungen eine Fehlerquote von 100 %. Da die beteiligten *Repeater* ansonsten funktional sind, (erkennbar an den anderen Verbindungskombinationen) sind anscheinend die *CrossbarSwitches* oder der Bus auf dem Chip beschädigt. Dies bedeutet, dass auch bei einer optimalen Einstellung der minimale Wert der *broken events* für diesen HICANN Chips bestenfalls 3.13 % betragen kann.

Des Weiteren erkennt man in Plot 11b, dass die Verbindungen mit den Empfänger-*Repeatern* 4 bis 7 deutlich schlechtere Ergebnisse aufweisen. Dies liegt an der räumlichen Lage der Empfänger 0 bis 3, die sich im linken unteren Block befinden und der Empfänger 4 bis 7, die im rechten unteren Block liegen. Somit muss das Signal von den *Sending Repeatern* zum rechten unteren Block eine deutlich längere Strecke zurücklegen, wodurch die Pulsform des Signals sich stärker verändert und für *Bitflips* anfälliger wird.

Die Untersuchung der Verbindungen ergibt, dass oftmals die gleichen Repeaterverbindungen eine höhere Fehlerrate aufweisen. Diese fehleranfälligen Verbindungen verbessern sich nicht durch eine Variation der Spannungen V_{OH} und V_{OL} . Jedoch sinkt ihre Fehlerrate signifikant durch die später durchgeführte DLL Optimierung.

4.7 Fehlerrate in Abhängigkeit der Neuronadressen

Die 6-Bit-Neuronadressen entsprechen je einem anderem Spannungssignal und sind somit unterschiedlich anfällig für *Bitflips*.

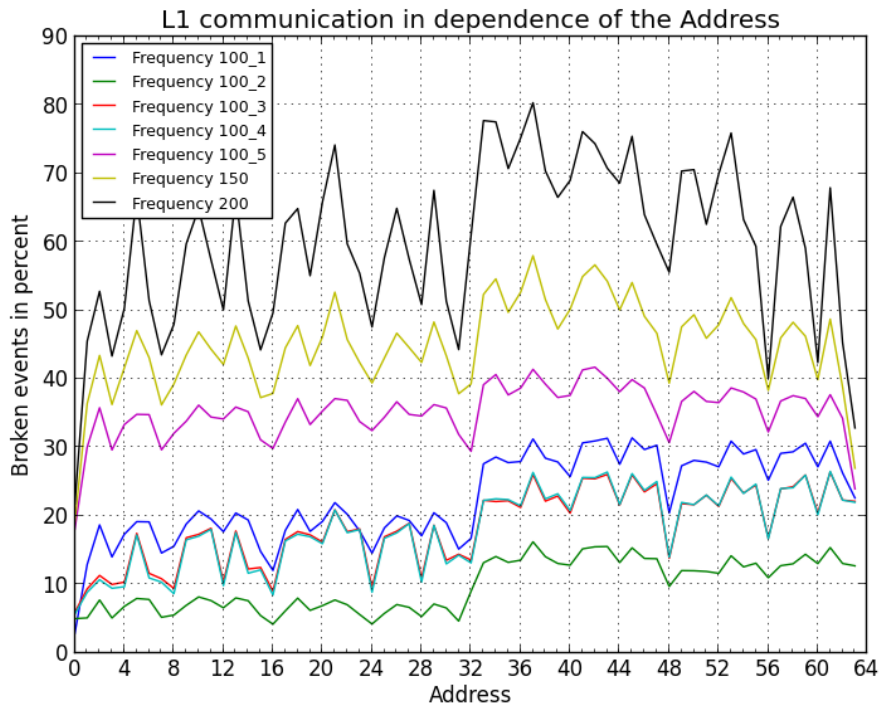


Abbildung 12: Darstellung der Fehleranfälligkeit der verschiedenen Neuronadressen am Setup mit unterschiedlichen Parametern/Einstellungen

In Abbildung [12] ist die Abhängigkeit der Fehlerrate von den Neuronadressen aus verschiedenen Messungen dargestellt.

Die verschiedenen Messungen unterscheiden sich deutlich aufgrund verwendeter Einstellungen, die zur Voruntersuchung der Hardware benutzt wurden. Hier ist nicht der absolute Wert von Bedeutung, sondern seine Schwankungen aufgrund der Adresse. Außerdem haben diese Messungen eine hohe Statistik mit circa 330.000 Einzelmessungen pro Kurve. Es ergibt sich eine Regelmäßigkeit welche Adressen tendenziell besser übertragen werden als andere. Dabei sind besonders Adressen mit wenig wechselnden Bits stabiler, wie 3 (000011), 15 (001111), 24 (011000), 56 (111000), 63 (111111). Hingegen sind Adressen mit wechselnden Bits, wie 5 (0000101), 21 (0010101), 37 (100101), 53 (110101), 61 (111101) deutlich instabiler. Hinzu kommt, dass alle Adressen größer 31 systematisch fehleranfälliger sind. Dies liegt am letzten Bit, das die 32 repräsentiert und als letztes vor dem Endbit decodiert wird. Dieses ist deutlich instabiler, da die Zeitschwankung der einzelnen Bitperioden sich akkumulieren und beim letzten Bit am größten ist.

Die Adresse Null ist die stabilste Adresse, wobei sie auch zu inkorrekten Adressen flippen kann. Dies ist vor allem relevant für die *Lock-Signale*, die dadurch zu Neuronenadressen werden könnten.

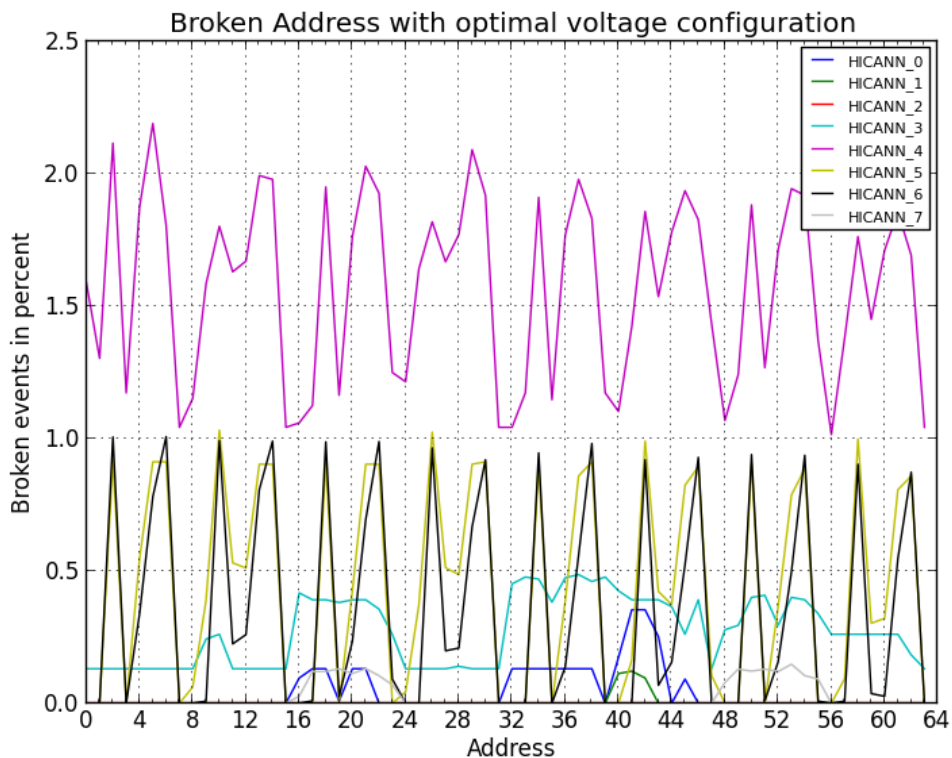


Abbildung 13: Darstellung der durchschnittlichen Fehleranfälligkeit der verschiedenen Bitadressen am Wafer im Spannungsbereich: $V_{OH} = 0.85$ bis 1.05 V und $V_{OL} = 0.6$ bis 0.7 V mit Frequenz = 100 MHz der 8 HICANN Chips

Bei den durchgeführten Messungen für die L1 Spannungen auf dem Wafer sind die Fehlerraten zusätzlich nach Adressen gespeichert worden. Aus den Werten des optimalen Spannungsbereichs $V_{OH} = 0.85$ bis 1.05 V und $V_{OL} = 0.6$ bis 0.7 V (vergleiche Abbildung [9]) werden die durchschnittlichen Abhängigkeit der Adressen in Abbildung [13] für die HICANN Chips aufgetragen. Dies entspricht circa 61.000 Einzelmessungen pro Kurve.

Dabei sind bei HICANN 4, 5 und 6 die gleichen Adressen stabil bzw. instabil wie bei den Untersuchungen am Setup. Jedoch ist keine Abhängigkeit mehr vom letzten Bit zu erkennen, d.h. kein systematischer Anstieg ab Adresse 32. Dies kann auf die längere DLL-Lockzeit am Wafer zurückgeführt werden, die eine genauere Bitperiode einstellt. Der HICANN 3 hingegen besitzt eine deutlich andere Abhängigkeit von den Neuroneadressen. Ähnlich zu dem Verhalten der HICANN Chips 0, 1 und 7, bei denen nur einzelne Adressebereiche fehleranfällig sind und die restlichen fehlerfrei übertragen werden. Es scheint, dass bei verschiedenen HICANN Chips die Übertragung unterschiedlich gestört wird.

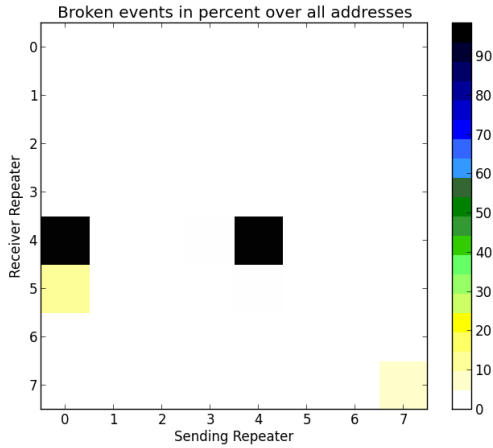
4.8 Frequenzabhängigkeit der Übertragungsqualität

Die Messung für einen bestimmten Spannungsbereich wurde für unterschiedliche Frequenzen der L1 Übertragung, d.h. unterschiedliche Sendedauer der einzelnen Bits, am Vertikal-Setup durchgeführt. In der Abbildung [12] erkennt man auch, dass am Setup die Fehlerrate mit der Frequenz zunimmt. Die durchschnittliche Fehlerrate bei den optimalen Spannungswerten $V_{OH} = 1.05$ V, $V_{OL} = 0.85$ V, ohne optimale DLL Einstellung, ist in Tabelle 1 aufgetragen. Die Fehlerrate erhöht sich stark mit der Frequenz. Exemplarisch werden für die verschiedenen Frequenzen einzelne Verbindungsplots mit dieser Spannungseinstellung in Abbildung [14] (a)-(c) dargestellt.

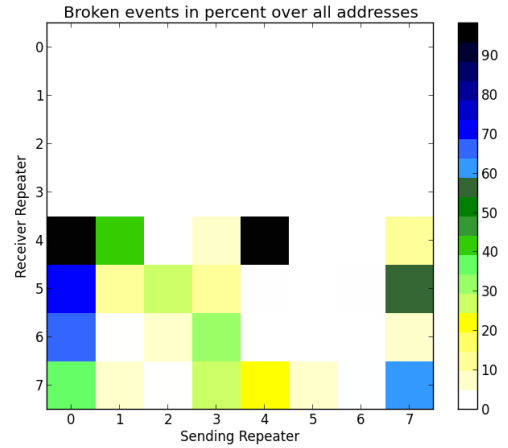
Es zeigt sich, dass durch die kürzere Lesezeit der einzelnen Bits die Verbindungen, vor allem mit den Empfängern im rechten unteren Block (wie in Kapitel 4.2 beschrieben), störanfälliger werden.

Frequenz [MHz]	Setup:	HICANN 0	HICANN 1	HICANN 2	HICANN 3
100	3.50%	1.43 %	0.00 %	9.15 %	4.04 %
150	11.31%	0.88 %	0.11 %	0.00 %	2.94 %
200	32.30%	6.14 %	2.55 %	1.03 %	9.24 %
Frequenz [MHz]		HICANN 4	HICANN 5	HICANN 6	HICANN 7
100		15.3%	0.92 %	3.65 %	1.18 %
150		15.2%	0.68 %	0.39 %	0.28 %
200		21.2%	7.01 %	4.92 %	1.47 %

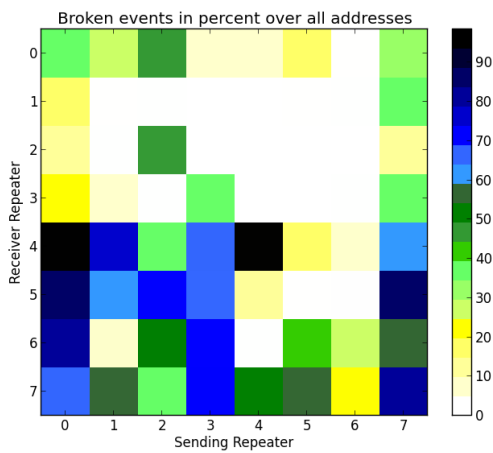
Tabelle 1: Übersicht der Übertragungsqualität bei unterschiedlichen Frequenzen in $\langle broken\ events \rangle$ am Vertikal-Setup und am Wafer.



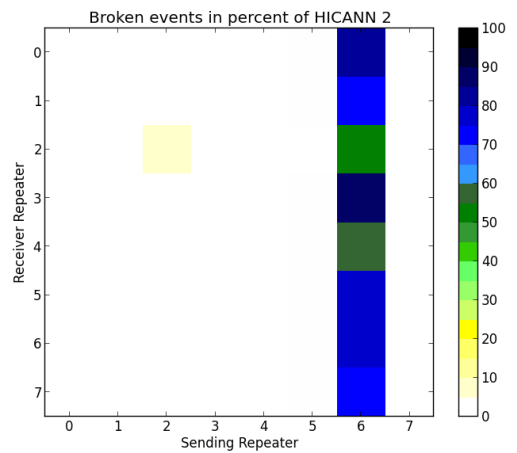
(a) Setup: Verbindungsplot mit 100 MHz



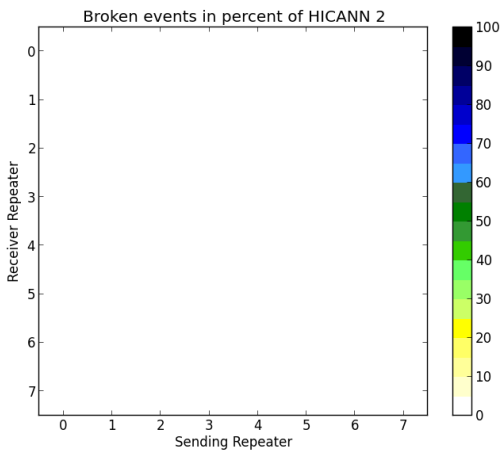
(b) Setup: Verbindungsplot mit 150 MHz



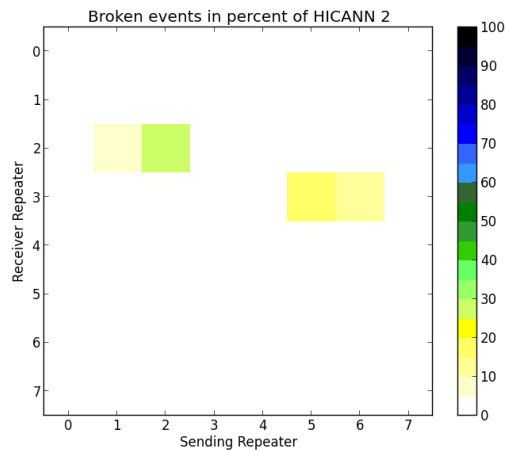
(c) Setup: Verbindungsplot mit 200 MHz



(d) Wafer: Verbindungsplot mit 100 MHz



(e) Wafer: Verbindungsplot mit 150 MHz



(f) Wafer: Verbindungsplot mit 200 MHz

Abbildung 14: Übersicht der Intra-HICANN Fehlerrate mit verschiedenen Frequenzen:

(a)-(c): **Vertikal-Setup** mit $V_{OH} = 1.05 \text{ V}$, $V_{OL} = 0.85 \text{ V}$.

(d)-(f): **Wafer HICANN 2** mit $V_{OH} = 0.9 \text{ V}$, $V_{OL} = 0.65 \text{ V}$

Die 8 HICANN Chips auf dem Retikel (siehe Tabelle 1) mit 150 MHz ergeben ungefähr gleiche oder verbesserte Fehlerraten gegenüber der Einstellung mit 100 MHz. Dies weicht von den Ergebnissen des Vertikal-Setup deutlich ab. Bei der Frequenz 200 verschlechtert sich wieder die Übertragung bei allen HICANN Chips.

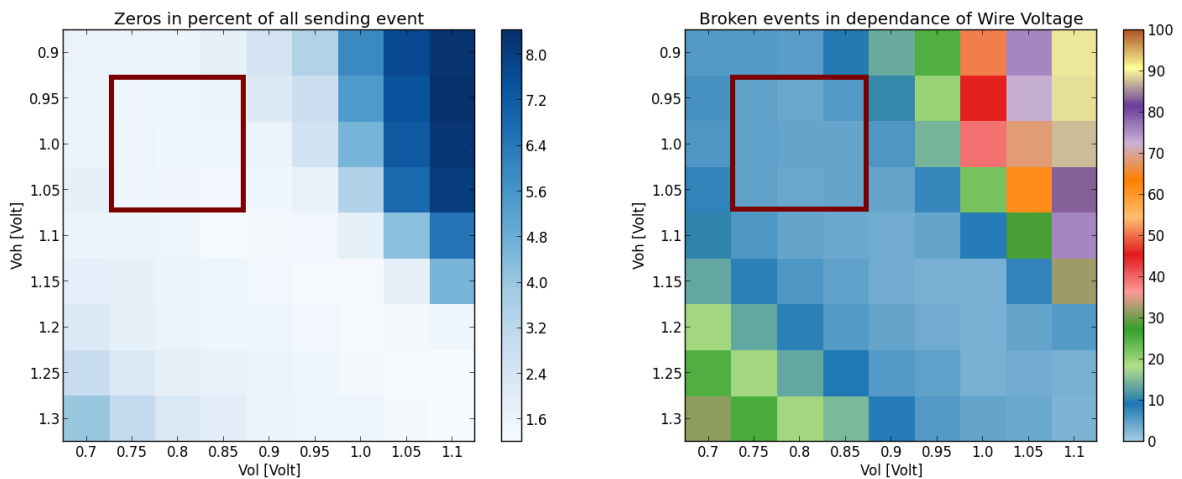
Dieser Effekt ist bisher unzureichend untersucht. Die erhöhte Fehlerrate bei 100 MHz beruht wahrscheinlich auf der einem Empfängerfehler. Es wurden fehlerhafte Adressen im Speicher der Repeater ausgelesen, deren Ankunftszeit sich nur um einen Sendezyklus unterschied. Da die Sendeadressen nur alle 40 Zyklen geschickt werden ist dies ein fehlerhaftes Verhalten der Repeater. Dieser Empfängerfehler wurde bei 150 MHz nicht beobachtet.

In Abbildung [14] (d)-(f) sind zusätzlich die Verbindungsplots des HICANN 2 bei unterschiedlichen Frequenzen aufgetragen. Dieser Chip zeigt gegenüber den anderen ein besonderes Verhalten, da bei 100 MHz der *Sending Repeater* 6 durchwegs eine hohe Fehlerrate erzielt. Dies ist bei steigender Frequenz nicht der Fall, sodass auch bei 200 MHz eine deutlich bessere Übertragung gewährleistet ist als bei 100 MHz.

Um die optimale Frequenz für den Wafer festzulegen sollten weitere Retikel untersucht werden.

4.9 Fehlerquelle durch empfangene Null-Events

Bei einigen Messung erreichten den Empfänger-*Repeater* ausschließlich Null Adressen und der Algorithmus versagte bei der Berechnung der *broken events*. Der Algorithmus wurde angepasst und dieses Phänomen untersucht.



(a) Prozentualer Anteil der empfangenen Nullen in Abhängigkeit der L1 Spannungen

(b) Prozentualer Anteil der empfangenen *broken events* in Abhängigkeit der L1 Spannungen

Abbildung 15: Vergleich der Null-Events mit der Übertragungsqualität

Bei den durchgeführten Messungen sollten aufgrund des ständigen DLL-Lock-Hinter-

grunds circa 1.6 % der empfangenen Events Null sein. Jedoch gibt es bei einigen Spannungen und Verbindungen eine deutlich erhöhte Anzahl.

Dies kann aufgrund von drei möglichen Ursachen auftreten. Erstens, es kommen an den Empfänger-*Repeatern* keine Signale an und alte gespeicherten Events werden ausgelesen, die meist Nullen sind. Dies wurde erkannt, da bei mehreren Übertragungsversuchen sich die Ankunftszeiten der Events nicht verändern.

Oder neue Null-Events erreichen den *Repeater*, deren Ankunftszeiten sich stetig ändern. Dies geschieht weil sie entweder stabiler sind als die anderen gesendeten Adressen (siehe Abbildung [12]) oder Adressen werden vom *Repeater* als Null falsch interpretiert. Die ankommenden Nullen wurden bei der Messung aus Kapitel 4.4 zusätzlich gespeichert und im Verhältnis zu allen gesendeten Adressen in Prozent in der Abbildung [15a] aufgetragen. Dieser kann mit der *broken events* der Spannungsmessung in Abbildung 15b verglichen werden.

Man erkennt, dass diese Ereignisse sich vor allem auf die schlechteren Spannungsbereiche beschränken und die optimalen Spannungen davon nicht beeinflusst werden.

5 Ausblick und Diskussion

Durch die in dieser Arbeit bestimmten Einstellungen, allen voran den L1 Spannungswerten V_{OH} und V_{OL} , ist es nun möglich die ersten neuronalen Netzwerke durch Layer 1 zu erstellen. Es gibt eine Vielzahl von verlässlichen Verbindungen mit Fehlerraten von 0.0 %. Jedoch zeigen die Ergebnisse auch, dass die einzelnen Verbindungsmöglichkeiten zuerst auf ihre Übertragungsqualität getestet werden müssen. Aber es ist nicht auszuschließen, dass sich die fehlerhaften Verbindungen weiter verbessern lassen.

Vor allem die Abhängigkeit von der Übertragungsfrequenzen und deren Wirkung muss tiefergehend analysiert werden. Ein weiterer zentraler Punkt zur Optimierung ist die Untersuchung der Spannungsversorgung am Wafer und deren Einfluss auf Layer 1. Auch die weitere Suche nach optimalen Parameter wie dem Spannungsbereich von V_{OH} und V_{OL} sollte fortgesetzt werden. Darüber hinaus kann die Abhängigkeit von weiteren Parameter und Einstellungen bei diesem komplexen Aufbau nicht ausgeschlossen werden.

Auch ohne Untersuchung bzw. Ausschließen von Verbindungen sollten die ersten Netzwerke modelliert werden können, da die Bitflips sich auf einen niedrigen einstelligen Prozentsatz beschränken. Dies hängt natürlich von der Fehlertoleranz des erstellten Netzwerks ab.

Der nächste Schritt ist die Untersuchung der Übertragungsqualität der Inter-HICANN und Inter-Retikel Kommunikation, sowie das Testen der Signalqualität über weite Strecken mit mehreren *Repeatern*.

Abschließend kann gesagt werden, dass die Kommunikationsstruktur Layer 1 des Wafers weitgehend funktioniert und weiterführende Messungen in diesem Bereich bzw. Experimente ermöglicht.

Literatur

- [1] **BrainScaleS Projekt** <http://brainscales.kip.uni-heidelberg.de/>
- [2] **Blue Brain Project** <http://bluebrain.epfl.ch/>
- [3] **Electronic Vision(s) Group** <http://www.kip.uni-heidelberg.de/cms/groups/vision>
- [4] **Sebastian Millner, Andreas Grübel, Karlheinz Meier, Johannes Schemmel und Marc-Olivier Schwartz** A VLSI Implementation of the Adaptive Exponential Integrate-and-Fire Neuron Model
- [5] **Sebastian Millner** Development of a Multi-Compartment Neuron Model Emulation
- [6] **Johannes Schemmel, Andreas Grübl, Sebastian Millner** Specification of the HICANN Microchip
- [7] **Alexander Kononov** Testing of an Analog Neuromorphic Network Chip
- [8] **Johannes Schemmel, Johannes Fieries und Karlheinz Meier** Wafer-Scale Integration of Analog Neural Networks
- [9] **Johannes Schemmel, Daniel Brüderle, Andreas Grübl, Matthias Hock, Karlheinz Meier and Sebastian Millner** A Wafer-Scale Neuromorphic Hardware System for Large-Scale Neural Modeling
- [10] **Matthias Hock** Test of Components for a Wafer-Scale Neuromorphic Hardware System
- [11] **Johannes Fieries, Johannes Schemmel, Karlheinz Meier** Realizing Biological Spiking Network Models in a Configurable Wafer-Scale Hardware System

Danksagung

Hiermit möchte ich mich recht herzlich bei Alexander Kononov bedanken, der mich in die komplexe Hardware eingearbeitet und während meiner Zeit in der Gruppe betreut hat. Des Weiteren möchte ich mich bei Eric Müller und Christoph Koke bedanken, die mir bei offenen Fragen oder Software Problemen weiter geholfen haben. Außerdem möchte ich meinem geschätzten Mitbewohner Constantin Pape danken, für seine Korrekturlesung.

Darüber hinaus möchte ich noch alle Leute aus der Electronic Vision(s) Gruppe würdigen, die diese neuromorphe Hardware erfunden, realisiert und in Betrieb genommen haben. Und weiterer Dank möchte ich zum Ausdruck bringen an diejenigen, die die Grundlagen und Möglichkeiten für dieses zukunftsträchtiges Projekt geschaffen haben, von dem ich ein Teil sein dürfte.

Erklärung

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 12 August 2013

.....