

Optimierung von Phasen- und Ratenparametern in einem stochastischen Modell neuronaler Feueraktivität

Masterarbeit

zur Erlangung des akademischen Grades
Master of Science
vorgelegt dem
Fachbereich 12 Mathematik
der Goethe-Universität Frankfurt am Main

vorgelegt von

Name: Straub Vorname: Benjamin

geb. am: 18.07.1990 in: Offenbach

Betreuerin: Prof. Dr. Gaby Schneider
Zweitgutachter: Prof. Dr. Anton Wakolbinger

Frankfurt am Main, den 17. Juli 2014

Danksagung

An dieser Stelle möchte ich mich ganz herzlich bei Frau Prof. Dr. Gaby Schneider für die sehr angenehme Betreuung, dem interessanten Thema und die hilfreichen Ratschläge und Rückmeldungen bedanken. Ich durfte mich bei Fragen sowie zum Einordnen von Ergebnissen jederzeit an Sie wenden.

Herrn Dr. Brooks Ferebee danke ich für Ratschläge und Ideen zur simulationsbasierten Optimierung und für seine hilfreichen Impulse.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Neurophysiologische Motivation	1
1.2	Mathematische Problemstellung	3
1.3	Aufbau der Arbeit	6
2	Optimale Parameterwahl getrennt für beide Stufen	8
2.1	Das Modell	8
2.2	Berechnen der Detektionswahrscheinlichkeit	10
2.2.1	Zwei Stimuli	10
2.2.2	Beliebige Anzahl Stimuli	12
2.3	Unterscheidung nur auf Basis von λ	15
2.3.1	Nullstimulus	15
2.3.2	Zwei Stimuli	17
2.3.3	Drei Stimuli	21
2.3.4	Beliebige Anzahl Stimuli	25
2.3.5	Asymptotisch beste Wahl	30
2.4	Unterscheidung nur auf Basis von τ	40
2.4.1	Zwei Stimuli, $N = 1$	41
2.4.2	Drei Stimuli, $N = 1$	42
2.4.3	Beliebige Anzahl Stimuli, $N = 1$	47
2.4.4	Zwei Stimuli, $N = n$	50
2.4.5	Drei Stimuli, $N = n$	51
2.4.6	Beliebige Anzahl Stimuli und zufällige erste Stufe	54
3	Optimale Parameterwahl auf Basis von λ und τ	55
3.1	Nullstimulus	55
3.2	Zwei Stimuli	58
3.3	Drei Stimuli	65
3.3.1	Zweite Stufe in Abhängigkeit von Stufe 1	66
3.3.2	Global beste Wahl für λ und τ	69
3.4	Vier Stimuli	73
3.5	Sieben Stimuli	76
4	Einführen von Ähnlichkeitsbeziehungen	79
4.1	Konstruktion Fehlerfunktion	79
4.1.1	Bedingte Fehlentscheidung	81
4.2	Stochastische Optimierung	83
4.3	Auswirkung Fehlerfunktion	85
4.3.1	Analyse der Entscheidungsschranken	86
4.3.2	Lineare Fehlerfunktion	87
4.3.3	Sensibilität gegenüber der Fehlerfunktion	90

INHALTSVERZEICHNIS

5 Schlussbetrachtung	92
5.1 Zusammenfassung	92
5.2 Ausblick	93
Abbildungsverzeichnis	94
Literaturverzeichnis	96

1 Einleitung

1.1 Neurophysiologische Motivation

Unsere Sinnesorgane und Sinnesrezeptoren, die über den gesamten Körper verteilt sind, empfangen die physikalischen Reize unserer Umwelt und übersetzen diese in Nervensignale. Dabei bilden die Nervenzellen (Neurone) den elementaren Baustein des Gehirns. Die Kernfrage der Hirnforschung ist hierbei, wie die zahlreichen psychologischen Phänomene der Wahrnehmung durch biophysikalische Vorgänge im Gehirn hervorgebracht werden.

Um dies zu erforschen, wird typischerweise einem Tier ein Sinnesreiz präsentiert und die Reaktion ausgewählter Neurone gemessen. Über die Dendriten empfängt das Neuron die Eingangssignale. Durch die Fortpflanzung von Spannungsveränderungen entlang der elektrisch geladenen Zellmembran werden die Signale innerhalb eines Neurons weitergeleitet. Der zentrale Zellkörper fasst diese zu einem Gesamtzustand zusammen und gibt diese über den Axonhügel weiter. Wird dort ein bestimmter Schwellenwert überstiegen, löst die Zellmembran eine Folge von Aktionspotentialen (Spikes) aus. Diese wandern am Axon entlang, welches über Synapsen mit den dendritischen Fasern anderer Neurone verbunden ist (für eine detaillierte Beschreibung siehe z.B. [Braitenberg u. Schüz (1991); Kandel u. a. (1996)]. Mit jeder Wiederholung des Stimulus fällt die Anzahl und die zeitliche Abfolge der Spikes verschieden aus.

Die physikalischen Eigenschaften der Spikes unterscheiden sich nicht, unabhängig davon ob z.B. ein visueller Reiz oder ein Klang diese ausgelöst haben.

Ob ein visueller Reiz oder ein Klang für den Spike verantwortlich war, wird anhand der neuronalen Bahn codiert, die vom Sinnesrezeptor über möglicherweise mehrere Zwischenstufen bis zu dem Neuron führt. Auf Grundlage der vier Stimuluseigenschaften Rate, Synchronizität, Oszillation und Phase kann ein Neuron aber noch mehr unterscheiden:

Rate 1958 fanden David Hubel und Thorsten Wiesel an der Harvard Medical School heraus, dass bestimmte Neuronen des visuellen Cortex besonders gut auf Lichtbalken einer bestimmten Orientierung ansprechen. Dazu wurde die Anzahl Spikes (Rate) einzelner Cortexneuronen auf bestimmte Test-Reize (Stimuli) gemessen. Vermutet wurde, dass die wesentliche Information in der Rate steckt. Zudem wurde entdeckt, dass Neurone, die auf ähnliche Orientierungen ansprechen, auch im Cortex räumlich nahe beieinander liegen [Hubel u. Wiesel (1962)].

Die Informationsverarbeitung in unserem Gehirn ist jedoch unzureichend beschrieben, solange man nur mit Hilfe eines einzelnen Neurons codiert. Neurone können nicht beliebig schnell aufeinanderfolgende Spikes hervorbringen, wodurch zum einen schnell verändernde Reize nicht zuverlässig übertragen werden können und zum anderen nur eine beschränkte Feuerrate zur Verfügung steht, mit der Änderungen der Reize nicht sicher unterschieden werden können [Bethge u. Pawelzik (2002)].

Typischerweise besitzt ein Neuron tausend bis zehntausend synaptische Eingänge [Larkman (1991)], womit es die Aktivitäten einer ganzen Population vorgeschalteter Neuronen empfängt. Doch wie lässt sich diese starke Vernetzungsstruktur für eine schnelle und genaue Codierung benutzen?

Ein erster Ansatz geht auf Barlow (1972) zurück, der durch Einzelzellbetrachtungen zeigen konnte, dass der Seheindruck im visuellen Cortex schrittweise entsteht. Die Neuronen zu Beginn entschlüsseln nur elementare Grundformen und geben diese Informationen weiter. In der späteren Verarbeitung antworten die Neuronen dann auf immer komplexere Formen und am Ende des visuellen Pfades steht der Gesamteindruck. So konnte die Existenz von Neuronen nachgewiesen werden, die nur auf komplexe Formen, wie z.B. Gesichter reagieren [Perrett u. a. (1982)].

Würde der Informationsprozess unseres Gehirns nach diesem Schema ablaufen, wäre jedoch eine unüberschaubare Anzahl an Neuronen notwendig, es müssten immer neue Zellen für unbekannte Objekte entstehen und wenn die zuständige Zelle absterben würde, würden wir das Objekt vergessen [Engel u. a. (1992); Singer (1993)]. Deshalb wurde vorgeschlagen, dass die gemeinsame Aktivität von verschiedenen Nervenzellverbänden zum Darstellen komplexer Objekte genutzt werden könnte [Gerstein u. Aertsen (1989)]. Auf diese Weise könnte das Problem einer begrenzten Anzahl an Neuronen gelöst werden, da sich die Neuronen dynamisch zu Verbänden zusammenfügen können; jedes Neuron kann zu verschiedenen Zeitpunkten an unterschiedlichen Verbänden teilhaben [Singer (1993)].

Das Bindungsproblem und Synchronizität Für eine derartige Informationsverarbeitung wird ein Mechanismus benötigt, mit dem sich gemeinsam-aktive Neuronen identifizieren lassen, was unter das sogenannte Bindungsproblem fällt [Gray u. a. (1992)]. Ein dazu wesentlicher Mechanismus wird in der neuronalen Synchronizität vermutet [Singer (1993)]. Darunter versteht man, dass die elektrische Aktivität verschiedener Neurone bzw. Neuronenverbände, die mehr oder weniger weit auseinanderliegen können, in hohem Maße korreliert sind. In den späten 1980er Jahre konnte Synchronizität im visuellen Cortex von anästhesierten Katzen und Affen, denen bewegte Balken oder Gitter gezeigt wurden, nachgewiesen werden [Eckhorn u. a. (1988), Gray u. Singer (1989)]. Ferner konnte Synchronizität auch im visuellen Cortex von wachen Katzen und Affen nachgewiesen werden [Frien u. a. (1994), Taylor u. a. (2005)]. Schließlich wurde Synchronizität auch im visuellen Cortex des Menschen aufgefunden [Adjamian u. a. (2004), Wyart u. Tallon-Baudry (2008)].

Oszillation Neben der Synchronizität der Spikeaktivität spielt ein weiteres Phänomen eine wichtige Rolle beim Formieren der Zellareale: Das synchrone Feuern mehrerer Neurone geht mit koordinierten rhythmischen Entladungen (Oszillationen) einher [Gray u. Singer (1989), Hoogenboom u. a. (2006)]. Gray u. Singer (1989) zeigten, dass sich die Aktionspotentiale im visuellen Cortex dem oszillatorischem Rhythmus anpassen, wodurch die Neuronen, die am selben Rhythmus teilhaben, ihre Entladungen mit einer hohen Genauigkeit synchronisieren. Insbesondere spielen Oszillationen im Beta- und Gamma-Band (20-80 Hz) eine bedeutende Rolle beim Einrichten genauer Synchronisation der Neuronenaktivität. Gamma-Band Synchronizität korrespondiert zu Zykluslängen von 25 ms bis zu 12 ms.

Die Gamma-synchronisierten Spikes treten etwa in einem halben Zyklus auf und befinden sich daher innerhalb etwa 12 ms bis zu 6 ms [Fries (2009)].

Phase Zusätzlich zu den Oszillationen weisen die synchronisierten Neuronen kleine Verschiebungen bis zu mehreren Millisekunden auf [Buzsáki u. Chrobak (1995); König. u. a. (1995); Roelfsema u. a. (1997); Schneider u. Nikolić (2006)]. Zu Beginn wurde vermutet, dass diese Verschiebungen nur das ungenaue Festsetzen der Feueraktivität an den Oszillationszyklus widerspiegeln [Buzsáki u. Chrobak (1995); Roelfsema u. a. (1997)]. Im visuellen Kortex von anästhesierten Katzen konnte gezeigt werden, dass die zeitlichen Feuermuster stimuluspezifisch sind und als Phasenparameter interpretiert werden können [Schneider u. a. (2006); Fries u. a. (2007); Havenith u. a. (2011)]. Die zeitlichen Feuermuster passen sich dabei systematisch an die Stimuluseigenschaften an und erlauben somit eine Schlussfolgerung auf den vorherrschenden Stimulus [Havenith u. a. (2011)].

Offen bleibt bisher jedoch die Frage, welchen Beitrag die Phase, die zeitliche Abfolge von Spikes, zur Informationsverarbeitung liefert. In welcher Beziehung steht dabei der Informationsgehalt der Feuerrate mit dem der zeitlichen Phase? Diese Zusammenhänge werden in dem folgenden mathematischen Modell analysiert:

1.2 Mathematische Problemstellung

Die vier erforschten Phänomene (Rate, Phase, Synchronizität und Oszillation) sind in dem folgenden stochastischen Spiketrain-Modell enthalten. Das ELO-Modell („exponential locking to a free oscillator“)[Schneider (2008)] zeigt hohe Ähnlichkeiten zu empirischen Spiketrains und kann Einzelspiketrains und Interaktionen zwischen Prozessen abbilden. Außerdem kann es auf Spiketrains anderer Hirnregionen verallgemeinert werden [Bingmer u. a. (2011), Schiemann u. a. (2012)].

In Abbildung 1 ist das ELO-Modell für zwei Neurone und ein Stimulus dargestellt. Das ELO-Modell ist in zwei Prozessstufen unterteilt: Im Hintergrund steht ein globaler Oszillationsprozess \mathbb{B} , der unabhängige und normal-verteilte Intervallabschnitte mit Parameter (μ, σ) hervorbringt (Oszillation), vgl. erste Zeile Abbildung 1.

$$\mathbb{B} = \dots, B_{-2}, B_{-1}, B_0, B_1, B_2, \dots; \quad (B_i - B_{i-1}) \sim \mathcal{N}(\mu, \sigma^2) \forall i \in \mathbb{Z}.$$

An den Intervallgrenzen starten unabhängige, inhomogene Poisson-Prozesse (Synchronizität) mit exponentiell abnehmender Feuerrate

$$\rho(x) = \frac{\lambda}{\tau} \exp(-x/\tau).$$

Jedes Neuron besitzt dabei eine stimuluspezifische Rate λ und Zeitkonstante (Phase) τ . Von nun an wird die Zeitkonstante τ im ELO-Modell als Phase bezeichnet. Mit Phase wird die zeitliche Verschiebung der Neurone gegeneinander beschrieben. Die Zeitkonstante

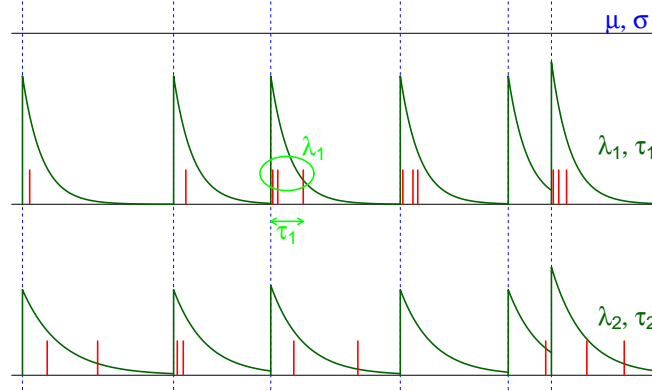


Abbildung 1: Das ELO-Modell für zwei Neurone. Im Hintergrund steht ein globaler Oszillationsprozess, der unabhängige und normal-verteilte Intervallabschnitte mit Parameter (μ, σ) hervorbringt (blau). An den Intervallgrenzen starten unabhängige, inhomogene Poisson-Prozesse mit exponentiell abnehmender Feuerrate $\rho(x) = \frac{\lambda}{\tau} \exp(-x/\tau)$ (grün). Neuron 1 (2) besitzt für den vorherrschenden Stimulus die Feuerrate λ_1 (λ_2) und die Phase τ_1 (τ_2).

τ steht dabei in direktem Zusammenhang zur Phase. Dazu betrachten wir die Zeitpunkte der Spikes bezüglich eines Oszillationszyklus: Sei nämlich für Neuron i

$$X_i := \frac{1}{N_i} \sum_{j=1}^{N_i} X_j^i, \quad \text{wobei } N_i \sim \text{Pois}(\lambda_i), X_j^i \sim \text{Exp}(\tau_i),$$

der mittlere Abstand vom Intervallbeginn der N_i aufgetretenen Spikes. Die zeitliche Verschiebung (Phase) zweier Neuronen i und j ist

$$\mathbb{E}[X_i - X_j] = \mathbb{E}[X_i] - \mathbb{E}[X_j] = \tau_i - \tau_j$$

und steht somit in direktem Zusammenhang zu der Zeitkonstante τ . In dieser Arbeit wird deshalb einfachheitshalber die Phase synonym zur Zeitkonstante τ verwendet.

Damit sich die Poisson-Prozesse benachbarter Intervalle größtenteils nicht überlagern, wird angenommen, dass im Vergleich zur mittleren Intervalllänge μ die Phase τ klein ist (z.B. $\tau < 0.5\mu$). Dies deckt sich mit empirischen Ergebnissen [Havenith u. a. (2011)]. Deshalb werden die Auswirkungen möglicher Überlagerungen vernachlässigt.

In der folgenden Arbeit wird ein einzelner Oszillationszyklus betrachtet, da die Informationsverarbeitung innerhalb des Gehirns sehr schnell abläuft und damit von Interesse ist,

welchen Beitrag bereits ein einzelner Oszillationszyklus liefern kann.

Vernachlässigt man, dass ein Überlagern der Prozesse auftreten kann, ist dies jedoch keine Einschränkung. Damit kann man zur Stimulusklassifizierung die Aktivitäten eines Neurons als Überlagerung unabhängiger Poisson-Prozesse zusammenfassen und die Fragestellung nach dem Zusammenspiel von λ und τ bleibt dieselbe. Betrachtet man nämlich z.B. zwei Oszillationszyklen entspricht dies im „Ein-Oszillations-Fall“ gerade dem Prozess mit doppelter Rate und selber Phase.

Im Folgenden wird nun für ein Neuron ein einzelner Oszillationszyklus innerhalb des ELO-Modells betrachtet. Untersucht wird dabei die Frage, wie die Stimuluserkennung innerhalb eines Oszillationszyklus in Abhängigkeit von Raten- und Phasenparametern optimiert werden kann und welchen Beitrag die Phase zur Informationsverarbeitung liefern kann. Es wird angenommen, dass jeder Stimulus mit derselben Wahrscheinlichkeit auftritt. Als Entscheidungsregel wird dabei die Bayessche Entscheidungsregel gewählt. Da alle Stimuli dieselbe Auftretswahrscheinlichkeit besitzen, entscheidet man sich also für den Stimulus, der für den beobachteten Spiketrain die höchste Dichte besitzt.

In Abbildung 2 ist die Problemstellung im Falle von zwei Stimuli - Stimulus 1 besitzt eine geringere Rate und Phase - illustriert.

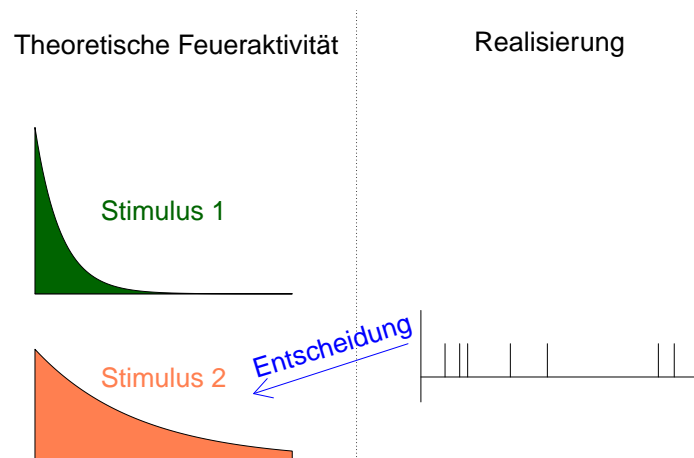


Abbildung 2: Entscheidungsaufgabe im Falle von 2 Stimuli und ein Neuron in einem Oszillationszyklus. Gegeben die stimuluspezifische Feueraktivität gemäß λ_i und τ_i , soll nach einer beobachteten Spike-Folge entschieden werden, welcher Stimulus diese hervorgebracht hat. Dabei wird der Stimulus gewählt, der eine höhere Dichte für diese Realisierung besitzt.

1.3 Aufbau der Arbeit

Hauptfrage dieser Arbeit ist, welchen Beitrag die Phase zum korrekten Detektieren der Stimuli (vgl. Abbildung 2) liefern kann. Um dies bewerten zu können, muss zuerst bestimmt werden, wie gut das Detektieren allein auf Basis der Rate funktioniert. Anhand dieser Baseline kann nun der andere Extremfall verglichen werden, wenn allein die Phase (alle Raten gleich) zum Detektieren genutzt wird. Die beiden Randfälle lassen sich dann mit simultanem Unterscheiden mit Hilfe von Rate und Phase vergleichen und die optimalen Kombinationen von Rate und Phase können analysiert werden.

Zu Beginn wird angegeben, wie sich die Detektionswahrscheinlichkeit, die Wahrscheinlichkeit, dass man den korrekten Stimulus detektiert, für eine beliebige, gegebene Parameterwahl bestimmen lässt. Die Frage nach der optimalen Parameterwahl, die also die Detektionswahrscheinlichkeit maximiert, wird für ein einzelnes Neuron und zunächst zwei Stimuli bearbeitet, und anschließend auf mehrere Stimuli übertragen. Dabei werden mathematisch sinnvolle Randbedingungen an Phase und Rate erarbeitet.

In Kapitel 2 wird zuerst untersucht, wie allein auf Basis von Rate (λ) die Stimuluserkennung optimiert werden kann. Dort wird ein Algorithmus entwickelt, mit dem die Berechnung der optimalen Raten (alle Phasen gleich) für eine beliebige Anzahl Stimuli annähernd in Echtzeit möglich ist. Weiterhin wird eine asymptotische Lösung bewiesen, die bereits für kleine λ eine außerordentlich brauchbare Lösung liefert.

Im zweiten Teil von Kapitel 2 wird analysiert, wie Stimuluserkennung allein auf Basis der Phase (τ) optimiert werden kann. In diesem Fall lässt sich für eine beliebige Anzahl Stimuli eine analytische Lösung der optimalen Parameterwahl beweisen.

In Kapitel 3 wird gleichzeitig auf Basis von Rate und Phase unterschieden. Dabei ist es eine große Herausforderung, die Interaktion von Phase und Rate in Bezug auf die optimale Parameterwahl zu verstehen. Hier lässt sich die optimale Parameterwahl nur noch numerisch bestimmen. Dabei wird die Bedeutung der Phase, selbst im Falle von zwei Stimuli, herausgearbeitet und das genauere Detektieren durch das Zusammenspiel von Rate und Phase ersichtlich.

Numerisch wird die optimale Parameterkombination für bis zu 7 Stimuli bestimmt. Für eine größere Anzahl Stimuli ist eine numerisch „exakte“ Berechnung aufgrund des hochdimensionalen Problems nicht mehr zu realisieren.

Alle bisherigen Berechnungen hatten das Ziel, die Detektionswahrscheinlichkeit zu maximieren. In empirischen Untersuchungen wurden die Aktionspotentiale für ähnliche Stimuli untersucht. Dabei wurden in verschiedenen Orientierungen gerichtete Balken gezeigt [Eckhorn u. a. (1988), Gray u. Singer (1989)]. Die geschätzten Parameter lassen sich dann in einem Kreisdiagramm (gemäß ihrer Orientierung), in sogenannten Tuning-Curves darstellen.

Innerhalb des ELO-Modells spielt die Zuordnung der Parameter zu den Stimuli keine Rolle. In der Realität reagieren Neurone auf ähnliche Stimuli auch auf „ähnliche Weise“. Um diese Beziehung innerhalb des ELO-Modells zu berücksichtigen, wird in Kapitel 4 eine Fehlerfunktion eingeführt, die Fehlentscheidungen unterschiedlich stark gewichtet. Dadurch wird der Rechenaufwand deutlich höher, da nun nicht mehr für jeden Stimulus nur die Wahrscheinlichkeit, sich für diesen korrekterweise zu entscheiden, bestimmt werden muss, sondern jegliche Fehlentscheidung zu berechnen ist.

Dort lassen sich strukturell bereits Trends vorhersagen. Zum expliziten Bestimmen bietet sich hier keine numerische Berechnung mehr an, weshalb stochastische Optimierung angewandt wird. Diese wird zuerst an die bekannten Ergebnisse aus der globalen Maximierung justiert. Anschließend wird diese auf die Fehlerfunktion angewendet und der Einfluss von verschiedenen Nachbarschaftsbeziehungen analysiert.

2 Optimale Parameterwahl getrennt für beide Stufen

In dieser Arbeit wird für ein Neuron ein einzelner Oszillationszyklus innerhalb des ELO-Modells betrachtet. Hierbei wird untersucht, welchen Beitrag die Phase zur Informationsverarbeitung leisten kann.

Zu Beginn dieses Kapitel wird das mathematische Modell, das einen Oszillationszyklus beschreibt, näher erläutert.

Anschließend wird erarbeitet, auf welche Weise die Stimuli detektiert werden und wie sich die Detektionswahrscheinlichkeit bestimmen lässt.

Mit diesem Maß, lassen sich dann verschiedene Parameterwahlen untersuchen und der Einfluss der Phase messen. Dazu wird zuerst die Detektion allein auf Basis der Rate (alle Phasen gleich) optimiert. Als Gegenspieler wird die Detektion allein auf Basis der Phase (alle Raten gleich) optimiert.

Mit diesen Ergebnissen lassen sich dann das Zusammenspiel von Rate und Phase bewerten und die Bedeutung der Phase herausarbeiten.

2.1 Das Modell

Betrachtet wird eine endliche Familie S_1, S_2, \dots, S_k von inhomogenen Poisson-Prozessen mit Rate

$$\rho_j(x) = \frac{\lambda_j}{\tau_j} \exp(-x/\tau_j).$$

Damit die einzelnen Prozesse wohldefiniert sind, ist $\lambda_j \geq 0$ und $\tau_j > 0$ für $j = 1, \dots, k$ gefordert.

Jeder Prozess S_j lässt sich durch ein zweistufiges Zufallsexperiment beschreiben. Dazu wird folgender Satz verwendet, siehe z.B [Cox u. Lewis (1966)]:

Satz 2.1 (Cox und Lewis, 1962). *Bezeichne X_1, X_2, \dots die zufälligen Spike-Zeitpunkte eines inhomogenen Poisson-Prozesses mit kontinuierlicher Erwartungsfunktion $\Lambda(t)$, und N_t repräsentiere die Anzahl Spikes, die bis zum Zeitpunkt t erschienen sind. Bedingt auf die Anzahl Spikes $N_{t_0} = n$, sind die Zeitpunkte X_1, X_2, \dots, X_n bezüglich ihrer Ordnungsstatistik aus einer Stichprobe mit Verteilungsfunktion $F(t) = \Lambda(t)/\Lambda(t_0)$ für $t \in [0, t_0]$.*

Mit diesem allgemeinen Resultat über die Verteilung von Auftritts-Zeitpunkten in einem inhomogenen Poisson-Prozess erhält man direkt:

Korollar 2.1. *Jeder Prozess S_j entsteht durch folgendes zweistufige Zufallsexperiment:*

- *In der ersten Stufe wird eine natürliche Zahl N gemäß einer Poissonverteilung mit Parameter λ_j gezogen.*
- *In der zweiten Stufe werden gemäß des Ausgangs N unabhängige Zeitpunkte X_i , $i = 1, \dots, N$, aus einer Exponentialverteilung mit Parameter τ_j gezogen.*

Beweis. Für $\Lambda(t)$ erhält man

$$\begin{aligned}\Lambda(t) &= \int_0^t \frac{\lambda}{\tau} \exp(-x/\tau) dx \\ &= \lambda(1 - \exp(-t/\tau)).\end{aligned}$$

Für $t_0 \rightarrow \infty$ ist $\Lambda(t_0) = \lambda$ und damit

$$F(t) = \frac{\Lambda(t)}{\Lambda(t_0)} = 1 - \exp(-t/\tau_j),$$

also $X_i \sim \text{Exp}(\tau)$. Weiterhin ist

$$N_{t_0} \sim \text{Pois}(\lambda) \quad \text{für } t_0 \rightarrow \infty.$$

□

Der Prozess S_j modelliert die Spike-Folge eines Neurons als Antwort auf Stimulus j . Es wird angenommen, dass jeder Stimulus mit gleicher Wahrscheinlichkeit erscheint. Der ausgewählte Stimulus ist nicht beobachtbar, nur eine Realisierung des zugehörigen Punktprozesses S_j . Gemäß dieser Information soll nun entschieden werden, welcher Stimulus zum Zuge gekommen ist.

Definition 2.1. *Das Ereignis, dass Stimulus i vorliegt, wird notiert als*

$$\{S = i\} := \{\text{„Stimulus } i \text{ liegt vor“}\}.$$

Dass man sich für Stimulus i entscheidet, wird bezeichnet durch

$$\{E = i\} := \{\text{„Entscheidung für Stimulus } i\text{“}\}.$$

In der Arbeit wird allein der Fall betrachtet, dass alle Stimuli dieselbe Auftrittswahrscheinlichkeit besitzen. Für k Stimuli gilt also

$$\mathbb{P}(\{S = i\}) = \frac{1}{k} \quad \forall i = 1, \dots, k.$$

Das grundlegende Ziel während des Entscheidungsprozesses ist, die Detektionswahrscheinlichkeit zu maximieren. Die Detektionswahrscheinlichkeit ist die Wahrscheinlichkeit, dass falls Stimulus i die Spike-Folge hervorgebracht hat, man sich auch für Stimulus i entscheidet.

Definition 2.2. *Für k Stimuli ist die von der Entscheidungsregel R abhängige Detektionswahrscheinlichkeit definiert als*

$$p_D^R := \frac{1}{k} \sum_{i=1}^k \mathbb{P}_{\mathbb{R}}(\{E = i\} | \{S = i\}).$$

Dies wirft folgende Fragen auf: Welche Entscheidungsregel R maximiert die Detektionswahrscheinlichkeit? Wie sind die freien Parameter λ_j und τ_j , $j = 1, \dots, k$, zu wählen, damit die Detektionswahrscheinlichkeit, die Wahrscheinlichkeit den korrekten Stimulus auszuwählen, maximiert wird. Welche Beschränkungen werden benötigt, um die Frage sinnvoll beantworten zu können?

Diesen Fragen wird erst für den Fall $k = 2$, zwei Stimuli, nachgegangen und anschließend wird die Fragestellung auf mehrere Stimuli übertragen.

2.2 Berechnen der Detektionswahrscheinlichkeit

2.2.1 Zwei Stimuli

Es seien nun zwei Stimuli S_1, S_2 gegeben mit $\lambda_1, \lambda_2, \tau_1, \tau_2 > 0$. Wie in Abschnitt 2.1 beschrieben, bezeichne N die Anzahl und X_1, \dots, X_n die Zeitpunkte der Spikes.

Sei

$$B := \{(N = n, X_1 = x_1, \dots, X_n = x_n)\}$$

der Ausgang der Realisierung. Die gesamte Information über λ und τ steckt dabei in der Anzahl Spikes und der Summe der Zeitpunkte $\sum X_i$.

Bemerkung 2.1. Das Paar $(N, \sum X_i)$ ist eine suffiziente Statistik für die Parameter (λ, τ) .

Wegen

$$\mathbb{P}(B | \{S = j\}) = \frac{\lambda_j}{n!} e^{-\lambda_j} \cdot \prod_{i=1}^n \frac{1}{\tau_j} e^{-x_i/\tau_j} = \frac{\lambda_j}{n!} e^{-\lambda_j} \frac{1}{\tau_j^n} e^{-\frac{\sum x_i}{\tau_j}}$$

ist nach dem Faktorisierungssatz von Neyman und Fisher N suffizient für λ und $(N, \sum X_i)$ suffizient für τ .

Die Bayessche Entscheidungsregel besagt (für einen Beweis siehe z.B. [Camastra u. Vinciarelli (2007)], Seite 95):

Satz 2.2 (Bayessche Entscheidungsregel). Die Detektionswahrscheinlichkeit p_D^R wird durch folgende Entscheidungsregel maximiert:

$$\begin{aligned} \text{Entscheide für } S_1, \text{ wenn } & \mathbb{P}(\{S = 1\} | B) > \mathbb{P}(\{S = 2\} | B), \\ \text{entscheide für } S_2, \text{ wenn } & \mathbb{P}(\{S = 2\} | B) > \mathbb{P}(\{S = 1\} | B). \end{aligned}$$

Anmerkung: Die Detektionswahrscheinlichkeit ist gerade die Gegenwahrscheinlichkeit der in der Bayesschen Entscheidungsregel üblichen Fehlerwahrscheinlichkeit.

Wendet man den Satz von Bayes doppelt an, erhält man

$$\frac{\mathbb{P}(\{S = 1\} | B)}{\mathbb{P}(\{S = 2\} | B)} = \frac{\mathbb{P}(B | \{S = 1\})}{\mathbb{P}(B | \{S = 2\})} \cdot \frac{\mathbb{P}(\{S = 1\})}{\mathbb{P}(\{S = 2\})}$$

Mit $\mathbb{P}(\{S = 1\}) = \mathbb{P}(\{S = 2\}) = 1/2$ ergibt dies

$$\frac{\mathbb{P}(\{S = 1\} | B)}{\mathbb{P}(\{S = 2\} | B)} = \frac{\mathbb{P}(B | \{S = 1\})}{\mathbb{P}(B | \{S = 2\})}.$$

Es reicht also aus, allein die Dichten für die Realisierung B unter den beiden Stimuli zu betrachten.

Man entscheidet sich demnach genau dann für Stimulus 1, falls der Ausgang B unter den Parametern λ_1, τ_1 eine höhere Dichte besitzt als unter λ_2, τ_2 .

Lemma 2.1. *Bei 2 Stimuli mit Ratenparametern $\lambda_1, \lambda_2 > 0$ und Phasenparametern $\tau_1, \tau_2 > 0$, entscheidet man sich gemäß der Bayesschen Entscheidungsregel für Stimulus 1, falls gilt*

$$n \log \frac{\lambda_1 \tau_2}{\lambda_2 \tau_1} + \left(\frac{1}{\tau_2} - \frac{1}{\tau_1} \right) \sum_{i=1}^n x_i > \lambda_1 - \lambda_2.$$

Diese Ungleichung stellt die Entscheidungsregel da: Analog entscheidet man sich für Stimulus 2, falls die linke Seite echt kleiner ist. Gilt Gleichheit, haben beide Stimuli dieselbe Dichte für die Realisierung B und man kann sich gemäß der Bayesschen Entscheidungsregel für einen beliebigen Stimulus entscheiden. Für $\tau_1 \neq \tau_2$ kommt Gleichheit mit Wahrscheinlichkeit 0 vor, für $\tau_1 = \tau_2$ aber $\lambda_1 \neq \lambda_2$ ist Gleichheit wegen der Ganzzahligkeit von n nicht möglich und für $\tau_1 = \tau_2$ und $\lambda_1 = \lambda_2$ ist die Dichte überall identisch und die Stimuli lassen sich nicht unterscheiden. In den interessanten Fällen ist für das Bestimmen der Detektionswahrscheinlichkeit nur der Fall „echt größer“ zu betrachten.

Beweis. Man entscheidet sich für Stimulus 1, falls

$$\frac{\mathbb{P}(\{S = 1\} | B)}{\mathbb{P}(\{S = 2\} | B)} \stackrel{\text{Satz 2.2}}{=} \frac{\mathbb{P}(B | \{S = 1\})}{\mathbb{P}(B | \{S = 2\})} \stackrel{!}{>} 1.$$

Eingesetzt liefert dies

$$\frac{\mathbb{P}(B | \{S = 1\})}{\mathbb{P}(B | \{S = 2\})} = \frac{\frac{\lambda_1}{n!} e^{-\lambda_1} \cdot \prod_{i=1}^n \frac{1}{\tau_1} e^{-x_i/\tau_1}}{\frac{\lambda_2}{n!} e^{-\lambda_2} \cdot \prod_{i=1}^n \frac{1}{\tau_2} e^{-x_i/\tau_2}} = \left(\frac{\lambda_1}{\lambda_2} \right)^n e^{\lambda_2 - \lambda_1} \left(\frac{\tau_2}{\tau_1} \right)^n e^{1/\tau_2 - 1/\tau_1} \sum_{i=1}^n x_i \stackrel{!}{>} 1.$$

Wendet man den natürlichen Logarithmus auf diese Gleichung an, erhält man den gewünschten Ausdruck. \square

Definition 2.3. *Für k Stimuli ist die Detektionswahrscheinlichkeit definiert als*

$$p_D := p_D^R,$$

wobei für R die Bayesschen Entscheidungsregel verwendet wird.

Eine Parameterwahl, die für gewisse Beschränkungen an Rate und Phase die Detektionswahrscheinlichkeit maximiert, wird als optimal bezeichnet.

Korollar 2.2. Für 2 Stimuli mit $\lambda_1, \lambda_2 > 0$ und $\tau_1, \tau_2 > 0$ lässt sich die Detektionswahrscheinlichkeit p_D bestimmen über

$$p_D = \frac{p_1 + p_2}{2},$$

wobei

$$p_1 := \mathbb{P} \left(N \log \frac{\lambda_1 \tau_2}{\lambda_2 \tau_1} + \left(\frac{1}{\tau_2} - \frac{1}{\tau_1} \right) \sum_{i=1}^n X_i > \lambda_1 - \lambda_2 \mid \{S = 1\} \right)$$

und analog für p_2 .

In Abhängigkeit von den verschiedenen Parameterkombinationen ergeben sich für p_1 die folgenden Fälle:

$$p_1 = \begin{cases} \sum_{n=0}^{\infty} \frac{\lambda_1}{n!} e^{-\lambda_1} \cdot \mathbb{P}_1 \left(\sum_{i=1}^n X_i > \frac{\lambda_1 - \lambda_2 - n \log \frac{\lambda_1 \tau_2}{\lambda_2 \tau_1}}{\frac{1}{\tau_2} - \frac{1}{\tau_1}} \right) & \text{für } \tau_2 < \tau_1, \\ \sum_{n=0}^{\infty} \frac{\lambda_1}{n!} e^{-\lambda_1} \cdot \mathbb{P}_1 \left(\sum_{i=1}^n X_i < \frac{\lambda_1 - \lambda_2 - n \log \frac{\lambda_1 \tau_2}{\lambda_2 \tau_1}}{\frac{1}{\tau_2} - \frac{1}{\tau_1}} \right) & \text{für } \tau_2 > \tau_1, \\ \mathbb{P}_1 \left(N > \frac{\lambda_1 - \lambda_2}{\log \lambda_1 / \lambda_2} \right) & \text{für } \tau_1 = \tau_2 \text{ und } \lambda_1 > \lambda_2, \\ \mathbb{P}_1 \left(N < \frac{\lambda_1 - \lambda_2}{\log \lambda_1 / \lambda_2} \right) & \text{für } \tau_1 = \tau_2 \text{ und } \lambda_1 < \lambda_2, \\ \frac{1}{2} & \text{für } \tau_1 = \tau_2 \text{ und } \lambda_1 = \lambda_2. \end{cases}$$

Für $\tau_1 = \tau_2$ und $\lambda_1 = \lambda_2$ ist es irrelevant, für welchen Stimulus man sich entscheidet. Der Wert $1/2$ bedeutet, dass man beide Stimuli mit Wahrscheinlichkeit $1/2$ auswählt, beide Stimuli also gleich behandelt. Jeder andere Wert wäre zwar auch denkbar, aber diese Wahl ist konsistenter zum Fall, in dem mehr als zwei Stimuli vorkommen.

2.2.2 Beliebige Anzahl Stimuli

Nun wird eine beliebige Anzahl Stimuli S_1, \dots, S_k betrachtet. Analog zum Fall $k = 2$ wird wieder die Bayessche Entscheidungsregel verwendet. Da für alle Stimuli gilt

$$\mathbb{P}(\{S = i\}) = \frac{1}{k} \quad \forall i = 1, \dots, k,$$

wird der Stimulus j ausgewählt, für den gilt:

$$\mathbb{P}(B \mid \{S = j\}) > \mathbb{P}(B \mid \{S = i\}) \quad \forall i = 1, \dots, k \text{ mit } i \neq j.$$

Aus Platzgründen wird die Schreibweise

$$\mathbb{P}_j(\cdot) := \mathbb{P}(\cdot \mid \{S = j\})$$

eingeführt.

Satz 2.3. *Bezeichne $k \in \mathbb{N}$ die Zahl an Stimuli. Gegebenen Ratenparameter $\lambda_1, \dots, \lambda_k > 0$ und Phasenparameter $\tau_1, \dots, \tau_k > 0$ ist $T_j := \{i \in \{1, \dots, k\} \setminus \{j\} \mid \tau_i = \tau_j \wedge \lambda_i \neq \lambda_j\}$ und $R_j := \{i \in \{1, \dots, k\} \setminus \{j\} \mid \tau_i = \tau_j \wedge \lambda_i = \lambda_j\}$.*

Die Detektionswahrscheinlichkeit lässt sich bestimmen über

$$p_D = \frac{1}{k} \sum_{i=1}^k p_i,$$

wobei

$$p_j = \frac{1}{|R_j|} \sum_{n=|l|}^{\lfloor u \rfloor} \frac{\lambda_j}{n!} e^{-\lambda_j} \cdot \mathbb{P}_j \left(\max_{\tau_r < \tau_j} \left(\frac{\lambda_j - \lambda_r - n \log \frac{\lambda_j \tau_r}{\lambda_r \tau_j}}{\frac{1}{\tau_r} - \frac{1}{\tau_j}} \right) < \sum_{i=1}^n X_i < \min_{\tau_r > \tau_j} \left(\frac{\lambda_j - \lambda_r - n \log \frac{\lambda_j \tau_r}{\lambda_r \tau_j}}{\frac{1}{\tau_r} - \frac{1}{\tau_j}} \right) \right)$$

für

$$l := \begin{cases} \max_{\lambda_r^{T_j} < \lambda_1} \left(\frac{\lambda_1 - \lambda_r^{T_j}}{\log \lambda_1 / \lambda_r^{T_j}} \right) & \text{falls } |\{\lambda_r^{T_j} < \lambda_1\}| > 0, \\ 0 & \text{sonst} \end{cases}$$

und

$$u := \begin{cases} \min_{\lambda_r^{T_j} > \lambda_1} \left(\frac{\lambda_1 - \lambda_r^{T_j}}{\log \lambda_1 / \lambda_r^{T_j}} \right) & \text{falls } |\{\lambda_r^{T_j} > \lambda_1\}| > 0, \\ \infty & \text{sonst.} \end{cases}$$

Beweis. Gemäß der Bayessche Entscheidungsregel und Lemma 2.1 wird genau dann Stimulus s ausgewählt, falls gilt:

$$n \log \frac{\lambda_s \tau_r}{\lambda_r \tau_s} + \left(\frac{1}{\tau_r} - \frac{1}{\tau_s} \right) \sum_{i=1}^n x_i > \lambda_s - \lambda_r \quad \forall r \neq s.$$

Analog zum Zwei-Stimuli-Fall muss bei der Bestimmung der bedingten Wahrscheinlichkeiten wieder dieselbe Fallunterscheidung durchgeführt werden. Zur besseren Lesbarkeit wird sich dabei auf p_1 beschränkt. Um p_1 zu bestimmen, muss unterschieden werden zwischen solchen Stimuli mit Parameter $\tau_r = \tau_1$ und solchen mit $\tau_r \neq \tau_1$.

Falls für alle $r \neq 1$ gilt: $\tau_r \neq \tau_1$, dann lässt sich p_1 schreiben als

$$p_1 = \sum_{n=0}^{\infty} \frac{\lambda_1}{n!} e^{-\lambda_1} \cdot \mathbb{P}_1 \left(\max_{\tau_r < \tau_1} \left(\frac{\lambda_1 - \lambda_r - n \log \frac{\lambda_1 \tau_r}{\lambda_r \tau_1}}{\frac{1}{\tau_r} - \frac{1}{\tau_1}} \right) < \sum_{i=1}^n X_i < \min_{\tau_r > \tau_1} \left(\frac{\lambda_1 - \lambda_r - n \log \frac{\lambda_1 \tau_r}{\lambda_r \tau_1}}{\frac{1}{\tau_r} - \frac{1}{\tau_1}} \right) \right).$$

Dies entspricht den ersten beiden Fällen im Zwei-Stimuli-Fall. Existiert aber neben Stimulus 1 ein weiterer Stimulus mit demselben τ , dann kann man diese nur aufgrund der Spike-Anzahl N unterscheiden. Falls ein weitere Stimulus mit sowohl demselben τ , als auch selbem λ existiert, dann lässt sich dieser gar nicht von Stimulus 1 unterscheiden. In diesem Fall wird einer aus den „gleichen“ Stimuli uniform gewählt.

Bezeichne $T \subset \{2, \dots, k\}$ die Menge der Stimuli, die dasselbe τ besitzen wie Stimulus 1, aber ein verschiedenes λ . Ist Stimulus r in T , wird λ_r durch λ_r^T bezeichnet. Unter diesen wird sich nur dann für Stimulus 1 entschieden, falls gilt

$$N_1 > \max_{\lambda_r^T < \lambda_1} \left(\frac{\lambda_1 - \lambda_r^T}{\log \lambda_1 / \lambda_r^T} \right) =: l \quad \text{und} \quad N_1 < \min_{\lambda_r^T > \lambda_1} \left(\frac{\lambda_1 - \lambda_r^T}{\log \lambda_1 / \lambda_r^T} \right) =: u.$$

Die Summationsgrenzen von p_1 ändern sich demnach zu

$$p_1 = \sum_{n=[l]}^{\lfloor u \rfloor} \frac{\lambda_1}{n!} e^{-\lambda_1} \cdot \mathbb{P}_1 \left(\max_{\tau_r < \tau_1} \left(\frac{\lambda_1 - \lambda_r - n \log \frac{\lambda_1 \tau_r}{\lambda_r \tau_1}}{\frac{1}{\tau_r} - \frac{1}{\tau_1}} \right) < \sum_{i=1}^n X_i < \min_{\tau_r > \tau_1} \left(\frac{\lambda_1 - \lambda_r - n \log \frac{\lambda_1 \tau_r}{\lambda_r \tau_1}}{\frac{1}{\tau_r} - \frac{1}{\tau_1}} \right) \right).$$

Bis zu diesem Schritt wurden noch nicht die Stimuli betrachtet, die dieselben Parameter τ und λ besitzen. Um das endgültige p_1 zu erhalten, muss dieses noch durch die Anzahl der „gleichen“ Stimuli geteilt werden. □

Bemerkung 2.2. *Existiert ein Stimulus mit $\lambda_j = 0$, gilt:*

$$\mathbb{P}(\{E = j\}) = \begin{cases} 1 & \text{falls } N = 0, \\ 0 & \text{sonst,} \end{cases}$$

da

$$\mathbb{P}(\{N = 0\} | \{S = j\}) > \mathbb{P}(\{N = 0\} | \{S = i\}) \quad \forall i = 1, \dots, k \text{ mit } i \neq j.$$

Bemerkung 2.3. *Existiert in Satz 2.3 ein Stimulus mit $\lambda_j = 0$, muss l ersetzt werden durch*

$$l := \begin{cases} \max_{\lambda_r^{T_j} < \lambda_1} \left(\frac{\lambda_1 - \lambda_r^{T_j}}{\log \lambda_1 / \lambda_r^{T_j}} \right) & \text{falls } |\{\lambda_r^{T_j} < \lambda_1\}| > 0, \\ 1 & \text{sonst,} \end{cases}$$

und die Detektionswahrscheinlichkeit ändert sich zu

$$p_D = \frac{1}{k+1} \left(1 + \sum_{i=1}^k p_i \right).$$

2.3 Unterscheidung nur auf Basis von λ

In diesem Abschnitt wird angenommen, dass alle Stimuli dasselbe τ besitzen. Eine Unterscheidung der Stimuli ist also nur auf Basis der ersten Stufe, der Spike-Anzahl möglich. O.B.d.A kann angenommen werden, dass die Stimuli geordnet sind durch $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$.

Die Frage ist nun, wie die Parameter λ_j , $j = 1, \dots, k$, zu wählen sind, damit sich die Stimuli bestmöglich unterscheiden lassen, die Detektionswahrscheinlichkeit also maximiert wird. Um diese Frage sinnvoll zu beantworten, muss man eine zusätzliche Bedingung einführen: Indem man λ_1 immer größer wählt, lässt sich die Detektionswahrscheinlichkeit offensichtlich beliebig nah an 1 bringen. Also muss man abhängig von λ_1 (fest) die restlichen $\lambda_i \leq \lambda_1$, $i = 2, \dots, k$ so wählen, dass die Detektionswahrscheinlichkeit maximiert wird.

Da nun alle Stimuli dieselbe Phase besitzen und der Rate nach geordnet sind, vereinfachen sich die Entscheidungsschranken aus Satz 2.3:

Korollar 2.3. Für $\tau_1 = \tau_2, \dots, \tau_k$ und $0 \leq \lambda_k \leq \lambda_{k-1} \leq \dots \leq \lambda_1$ lässt sich die Detektionswahrscheinlichkeit schreiben als

$$p_D = \frac{1}{k} \left(\mathbb{P}_{\lambda_k}(0 \leq N \leq l_{k-1}) + \sum_{i=2}^{k-1} \mathbb{P}_{\lambda_i}(l_i < N \leq l_{i-1}) + \mathbb{P}_{\lambda_1}(N > l_1) \right),$$

für

$$l_j = \frac{\lambda_j - \lambda_{j-1}}{\log(\lambda_j / \lambda_{j-1})} \quad \text{für } j = 1, \dots, k-1.$$

Im ersten Teilabschnitt wird der Zusammenhang zwischen Entscheidungsschranken und den Ratenparametern analysiert und dadurch ein Algorithmus entwickelt, mit dem sich für eine beliebige Anzahl Stimuli die optimalen Ratenparameter numerisch exakt berechnen lassen.

Im zweiten Abschnitt wird über die Normalapproximation asymptotisch eine analytische Lösung erarbeitet.

2.3.1 Nullstimulus

Gegeben sind zwei Stimuli S_1, S_2 mit $0 \leq \lambda_2 \leq \lambda_1 \leq \lambda_{\max}$ und $\tau_1 = \tau_2$. Gesucht ist das $\tilde{\lambda}_2$, sodass

$$p_D(\lambda_1, \tilde{\lambda}_2) \geq p_D(\lambda_1, \lambda_2) \quad \forall \lambda_2 \leq \lambda_1.$$

Diese Frage lässt sich auf folgende Problemstellung reduzieren. Gegeben eine $Pois(\lambda_1)$ -verteilte Zufallsvariable N_1 , wie ist der Parameter λ_2 einer weiteren poissonverteilten Zufallsvariable N_2 zu wählen, damit sich diese bestmöglich unterscheiden lassen, siehe Abbildung 3. Man erwartet, dass diese Parameter möglichst weit voneinander entfernt liegen, das optimale $\lambda_2 = 0$ ist. Auf diese Weise lassen sich die Dichten am stärksten „trennen“.

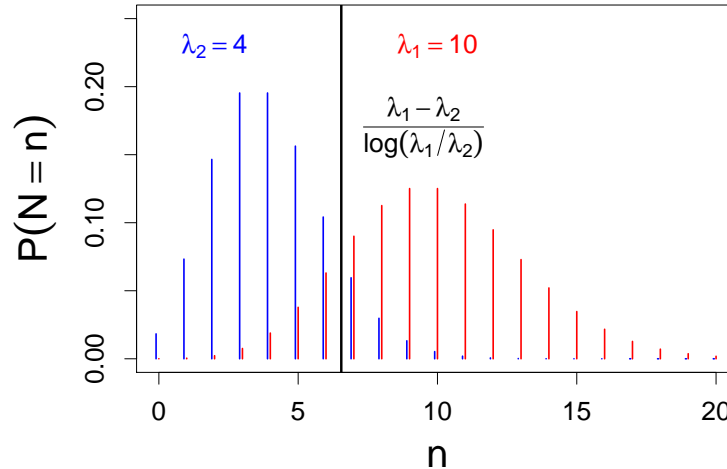


Abbildung 3: Entscheidungsschranke nur auf Basis von λ für zwei Stimuli. In rot (blau) ist die Dichte einer Poissonverteilung mit Parameter λ_1 (λ_2) dargestellt, welche die Spikes-Anzahl von Stimulus 1 (2) repräsentiert. Beobachtet man links von der Entscheidungsschranke viele Spikes, entscheidet man sich für Stimulus 2, rechts davon für Stimulus 1.

Lemma 2.2. Sei $\tau_1 = \tau_2$ und $0 \leq \lambda_2 \leq \lambda_1 \leq c$, $c \in \mathbb{R}^+$ fest. Dann wird die Detektionswahrscheinlichkeit maximiert durch

$$\lambda_2 = 0.$$

Beim Bestimmen der Detektionswahrscheinlichkeit haben wir gesehen, wie die optimalen Entscheidungsschranken abhängig von λ_1 und λ_2 sind. Nun ist λ_2 aber nicht bekannt. Der Trick hierbei ist, dass man den Weg rückwärts geht: Angenommen man kennt die optimale Schranke, wie lautet dazu das optimale λ_2 ?

Beweis. Sei also $l \in \mathbb{N}$ die optimale Schranke. Bis einschließlich l entscheidet man sich für Stimulus 2, oberhalb von l für Stimulus 1. Das dazu optimale λ_2 erhält man, indem man

$$p_2^l(\lambda_2) = \mathbb{P}(0 \leq N_2 \leq l) = \sum_{i=0}^l \frac{\lambda_2^i}{i!} e^{-\lambda_2}$$

nach λ_2 maximiert. Dies nach λ_2 abgeleitet ergibt

$$\begin{aligned} \frac{\partial p_2^l(\lambda_2)}{\partial \lambda_2} &= e^{-\lambda_2} \left(\sum_{i=1}^l \left(\frac{\lambda_2^{i-1}}{(i-1)!} - \frac{\lambda_2^i}{i!} \right) - 1 \right) \\ &= e^{-\lambda_2} \left(\sum_{i=0}^{l-1} \frac{\lambda_2^i}{(i)!} - \sum_{i=1}^l \frac{\lambda_2^i}{i!} - 1 \right) \\ &= e^{-\lambda_2} \left(1 - \frac{\lambda_2^l}{l!} - 1 \right) = e^{-\lambda_2} \frac{\lambda_2^l}{l!} \\ &\stackrel{!}{=} 0 \quad \text{für } \lambda_2 = 0, \text{ falls } l > 0. \end{aligned}$$

Falls $l = 0$ ist, bleibt offensichtlich $\lambda_2 = 0$ die optimale Wahl. Also ist unabhängig von der optimalen Schranke, wie erwartet, immer $\lambda_2 = 0$ die beste Wahl. \square

Damit ist der Zwei-Stimuli-Fall relativ uninteressant, da der zweite Stimulus deterministisch 0 ist und deshalb nun als Nullstimulus bezeichnet wird. Dieser kann als der Zustand gedeutet werden, in dem kein Reiz vorliegt.

Bemerkung 2.4. *Im Folgenden wird der Nullstimulus nicht zu der Anzahl der Stimuli mitgezählt und stets als vorhanden angenommen.*

2.3.2 Zwei Stimuli

Gegeben sind zwei Stimuli S_1, S_2 mit $0 < \lambda_2 \leq \lambda_1 \leq \lambda_{\max}$ und $\tau_1 = \tau_2$. Gesucht ist das $\tilde{\lambda}_2$, sodass

$$p_D(\lambda_1, \tilde{\lambda}_2, 0) \geq p_D(\lambda_1, \lambda_2, 0) \quad \forall \lambda_2 \leq \lambda_1.$$

Berechnet man für festes λ_1 die Detektionswahrscheinlichkeit für $\lambda_2 \leq \lambda_1$ und bestimmt numerisch das Maximum, stellt sich heraus, dass λ_2 , als eine Funktion von λ_1 nicht kontinuierlich wächst, sondern eine Treppenfunktion ist. In Abbildung 4 ist dies für $\lambda_1 = 2, \dots, 10$ umgesetzt. Dass sich dabei eine Treppenfunktion ergibt, ist auch nicht verwunderlich: Eine poisson-verteilte Zufallsvariable nimmt nur ganzzahlige Werte an und solange die Entscheidungsschranke nicht einen neuen ganzzahligen Wert erreicht, handelt es sich eigentlich noch um dieselbe Schranke.

Deshalb ist es naheliegend, wieder den Weg über die optimale Schranke einzuschlagen: Kennt man die obere Schranke l_1 , entscheidet man sich demzufolge für Stimulus 2, falls höchstens l_1 Spikes auftreten, aber mindestens einer, und für Stimulus 1, falls mehr als l_1 Spikes auftreten.

Lemma 2.3. *Für 2 Stimuli mit $\tau_1 = \tau_2$ und $0 < \lambda_2 < \lambda_1 \leq \lambda_{\max}$, $\lambda_{\max} \in \mathbb{R}^+$ fest. Dann ist in Abhängigkeit von der oberen Entscheidungsschranke $l_1 \in \mathbb{N}$ das optimale λ_2 gegeben durch*

$$\lambda_2 = l_1!^{1/l_1}.$$

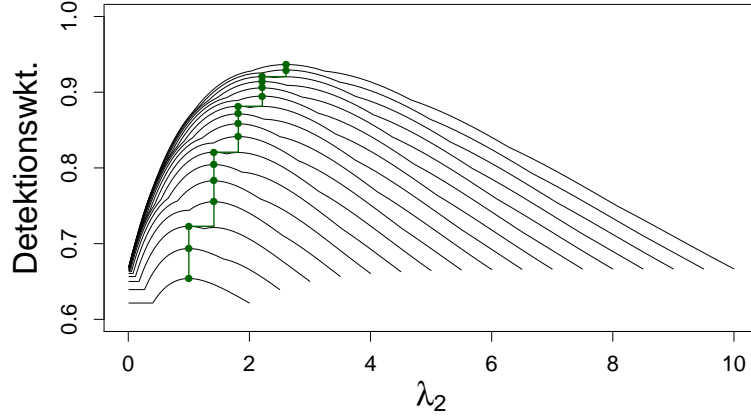


Abbildung 4: Das optimale λ_2 als Treppenfunktion von λ_1 . Die schwarzen Kurven entsprechen aufsteigend einem λ_1 von 2, 2.5, 3, ..., 10 und geben die Detektionswahrscheinlichkeit in Abhängigkeit von λ_2 an. Die grünen Punkte stellen die optimale Wahl für λ_2 dar.

Beweis. Unter Beachtung der Existenz des Nullstimulus erhält man das optimale λ_2 wieder analog, indem man

$$p_2^{l_1}(\lambda_2) = \mathbb{P}(0 < N_2 \leq l_1) = \sum_{i=0}^{l_1} \frac{\lambda_2^i}{i!} e^{-\lambda_2}$$

nach λ_2 maximiert. Nach λ_2 abgeleitet, ergibt sich für $l_1 > 0$

$$\begin{aligned} \frac{\partial p_2^{l_1}(\lambda_2)}{\partial \lambda_2} &= e^{-\lambda_2} \sum_{i=1}^{l_1} \left(\frac{\lambda_2^{i-1}}{(i-1)!} - \frac{\lambda_2^i}{i!} \right) \\ &= e^{-\lambda_2} \left(1 - \frac{\lambda_2^{l_1}}{l_1!} \right) \\ &\stackrel{!}{=} 0 \quad \text{für } \lambda_2 = l_1!^{1/l_1}. \end{aligned}$$

□

Mit der Stirling-Formel erhält man

$$\lambda_2(l_1) \approx \frac{l_1}{e} \cdot \sqrt{2\pi l_1}^{-1/l_1}$$

Ist $l_1 = 0$, fallen die obere und untere Schranke zusammen, und die genaue Wahl von λ_2 ist irrelevant. Um konsistent zum allgemeinerem Fall zu bleiben, wird dann $\lambda_2 = 0$ gesetzt und im Falle von Null Spikes wird Stimulus 2 oder der Nullstimulus mit derselben Wahrscheinlichkeit zufällig gewählt.

Kennt man die optimale Schranke, ist λ_2 damit eindeutig bestimmt. Kennt man die Werte für λ_1 , bei denen die optimale Schranke gerade um 1 nach oben springt, ist das optimale λ_2 in den gesamten Intervallen bestimmt.

Lemma 2.4. *Gegeben zwei Stimuli S_1, S_2 mit $\tau_1 = \tau_2$ und $\lambda_2 < \lambda_1$. Um herauszufinden, für welches λ_1 das optimale λ_2 von $\lambda_{(l_1)} = l_1!^{1/l_1}$ auf $\lambda_{(l_1+1)} = (l_1 + 1)!^{1/(l_1+1)}$ springt, ist das λ_1 gesucht, sodass gilt:*

$$\mathbb{P}_{\lambda_{(l_1)}}(0 < N \leq l_1) + \mathbb{P}_{\lambda_1}(N = l_1 + 1) = \mathbb{P}_{\lambda_{(l_1+1)}}(0 < N \leq l_1 + 1).$$

Dabei steht \mathbb{P}_λ für die bedingte Wahrscheinlichkeit, dass der echte Parameter von N gerade λ beträgt.

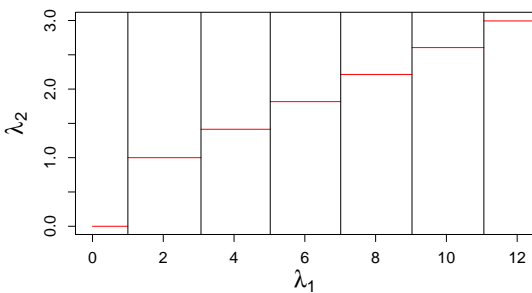
Beweis. Die Höhe der Schranke ist nur von λ_1 abhängig und monoton wachsend mit λ_1 . Mit höherem λ_1 wird die Wahrscheinlichkeit immer geringer, dass Stimulus 1 genau $l_1 + 1$ Spikes hervorbringt.

Die Sprungstellen sind gerade die Stellen, an den beiden Wahlen der oberen Schranken optimal sind. Gesucht ist also

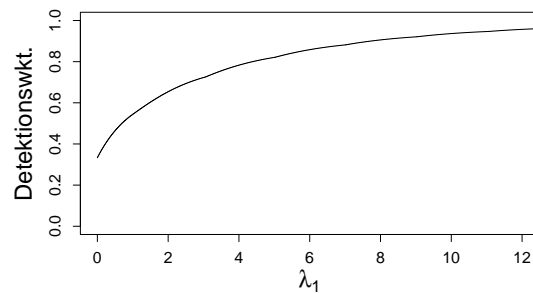
$$\mathbb{P}_{\lambda_{(l_1)}}(0 < N \leq l_1) + \mathbb{P}_{\lambda_1}(N > l_1) = \mathbb{P}_{\lambda_{(l_1+1)}}(0 < N \leq l_1 + 1) + \mathbb{P}_{\lambda_1}(N > l_1 + 1),$$

umgeformt erhält man das Behauptete. □

Für jedes l_1 lässt sich diese Gleichung numerisch lösen. Die Werte, bei denen das optimale λ_2 von 0 auf $\lambda_{(1)}$ auf $\lambda_{(2)}$ auf $\lambda_{(3)}$ auf $\lambda_{(4)}$ auf $\lambda_{(5)}$ springt, sind ungefähr $\lambda_1 = 1, 3.07, 5.02, 7.01, 9.03$ (Vertikalen Linien in Abbildung 5). In Abbildung 5 ist die optimale



(a) Optimales λ_2 (rot) als Stufenfunktion von λ_1 .



(b) Maximale Detektionswkt. für die Optimale Wahl von λ_2 .

Abbildung 5: Das optimale λ_2 und die maximale Detektionswahrscheinlichkeit in Abhängigkeit von λ_1 für 2 Stimuli.

Wahl für λ_2 und die zugehörige maximale Detektionswahrscheinlichkeit in Abhängigkeit von λ_1 dargestellt, die man über

$$p_D^{\lambda_1} = \frac{1}{3} \left(1 + \mathbb{P}_{\lambda_{(l_1)}}(0 < N \leq l_1) + \mathbb{P}_{\lambda_1}(N > l_1) \right)$$

erhält. Beginnend bei $1/3$ für $\lambda_1 = 0$, in diesem Fall lassen sich die Stimuli nicht unterscheiden, nähert sich die Detektionswahrscheinlichkeit monoton für große λ_1 der 1 an. In Abbildung 6 sind die Werte von λ_1 dargestellt, bei denen die Schranke von l_1 nach $l_1 + 1$ springt. Wie bereits in Abbildung 5 zu erkennen war, sind die Abstände zwischen zwei Sprüngen annähernd gleich groß.

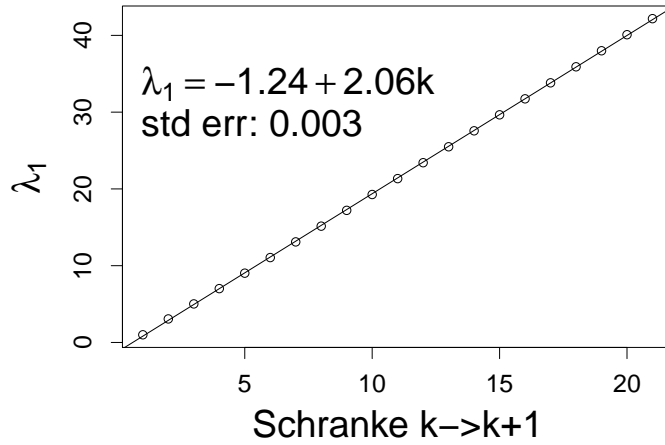


Abbildung 6: Die Werte von λ_1 , für die die optimale Schranke um 1 nach oben springt, dargestellt durch kleine Kreise. Durch diese Punkte ist eine lineare Regressionsgerade gelegt, welche einen annähernd linearen Zusammenhang aufzeigt.

Satz 2.4. *Bezeichne κ den Wert von λ_1 , für den die optimale Schranke von l_1 auf $l_1 + 1$ springt. Dann ist eine grobe Abschätzung für κ gegeben durch*

$$l_1!^{1+1/l_1} e^{l_1^{1/l_1}} \leq \kappa^{l_1+1} e^{-\kappa} \leq (l_1 + 1)! e^{-(l_1+1)^{1/(l_1+1)}}.$$

Beweis. „ \Leftarrow “ Sei $\lambda_2 = (l_1 + 1)^{1/(l_1+1)}$. Dann gilt auf Grund der Optimalität

$$\begin{aligned} \mathbb{P}_{\kappa}(N = l_1 + 1) &\leq \mathbb{P}_{\lambda_2}(N = l_1 + 1) \\ \Leftrightarrow \frac{\kappa^{l_1+1}}{(l_1 + 1)!} e^{-\kappa} &\leq \frac{\lambda_2^{l_1+1}}{(l_1 + 1)!} e^{-\lambda_2} \\ \Leftrightarrow \kappa^{l_1+1} e^{-\kappa} &\leq (l_1 + 1)! e^{-(l_1+1)^{1/(l_1+1)}} \end{aligned}$$

„ \Rightarrow “ Sei jetzt $\lambda_2 = l_1!^{1/l_1}$. Dann gilt

$$\begin{aligned} & \mathbb{P}_\kappa(N = l_1 + 1) \geq \mathbb{P}_{\lambda_2}(N = l_1 + 1) \\ \Leftrightarrow & \frac{\kappa^{l_1+1}}{(l_1 + 1)!} e^{-\kappa} \geq \frac{\lambda_2^{l_1+1}}{(l_1 + 1)!} e^{-\lambda_2} \\ \Leftrightarrow & \kappa^{l_1+1} e^{-\kappa} \geq l_1!^{1+1/l_1} e^{-l_1!^{1/l_1}} \end{aligned}$$

□

2.3.3 Drei Stimuli

Gegeben sind 3 Stimuli S_1, S_2, S_3 mit $0 < \lambda_3 < \lambda_2 < \lambda_1 \leq \lambda_{\max}$ und $\tau_1 = \tau_2 = \tau_3$. Dabei bezeichne $L = (0, l_2, l_1)$, $l_i \in \mathbb{N}$, die Entscheidungsschranken. Für festgehaltene Schranke l_2 wird Stimulus 3 detektiert, falls

$$0 < N \leq l_2$$

und nach Lemma 2.1 ist das optimale $\lambda_3 = l_2!^{1/l_2}$. Für Stimulus 2 entscheidet man sich, falls

$$l_2 < N \leq l_1.$$

Lemma 2.5. Sei $\tau_1 = \tau_2 = \tau_3$ und $0 < \lambda_3 < \lambda_2 < \lambda_1 \leq \lambda_{\max}$, $\lambda_{\max} \in \mathbb{R}^+$ fest. Dann ist in Abhängigkeit von den Entscheidungsschranken l_1 und l_2 das optimale λ_2 gegeben durch

$$\lambda_2 = \left(\frac{l_1!}{l_2!} \right)^{1/(l_1-l_2)}.$$

Beweis. Da man sich für Stimulus 2 entscheidet, falls $(l_2 + 1) \leq N \leq l_1$, muss

$$p_2^{l_1, l_2}(\lambda_2) = \sum_{i=l_2+1}^{l_1} \frac{\lambda_2^i}{i!} e^{-\lambda_2}$$

nach λ_2 maximiert werden. Nach λ_2 abgeleitet, ergibt sich für $l_1 > l_2$

$$\begin{aligned} \frac{\partial p_2^{l_1, l_2}(\lambda_2)}{\partial \lambda_2} &= \sum_{i=l_2+1}^{l_1} \left(\frac{i \lambda_2^{i-1}}{i!} e^{-\lambda_2} - \frac{\lambda_2^i}{i!} e^{-\lambda_2} \right) \\ &= e^{-\lambda_2} \left(\sum_{i=l_2}^{l_1-1} \frac{\lambda_2^i}{i!} - \sum_{i=l_2+1}^{l_1} \frac{\lambda_2^i}{i!} \right) \\ &= e^{-\lambda_2} \left(\frac{\lambda_2^{l_2}}{l_2!} - \frac{\lambda_2^{l_1}}{l_1!} \right) \\ &= e^{-\lambda_2} \lambda_2^{l_2} \left(\frac{1}{l_2!} - \frac{\lambda_2^{l_1-l_2}}{l_1!} \right). \end{aligned}$$

Das Maximum wird angenommen, falls die Klammer Null wird, also für

$$\lambda_2^{l_1-l_2} = \frac{l_1!}{l_2!}.$$

□

Mit der Stirling-Formel ergibt sich approximativ

$$\lambda_2(l_1, l_2) \approx \left(\frac{\sqrt{l_1} \left(\frac{l_1}{e}\right)^{l_1}}{\sqrt{l_2} \left(\frac{l_2}{e}\right)^{l_2}} \right) = \frac{1}{e^{l_1-l_2}} \left(\frac{l_1}{l_2}\right)^{\frac{l_2+1/2}{l_1-l_2}} \cdot l_1$$

Findet man die optimalen Schranken, sind die optimalen Parameter damit auch bestimmt. Dabei sind nicht alle Kombinationen von l_1 und l_2 möglich. Aus den Optimalitätsbedingungen lassen sich mögliche Werte für λ_2 in Abhängigkeit von l_2 eingrenzen. Ist die untere Schranke bei l_2 , so gilt für λ_2 und λ_3 :

$$\mathbb{P}_{\lambda_3}(N = l_2) \geq \mathbb{P}_{\lambda_2}(N = l_2) \quad \text{und} \quad \mathbb{P}_{\lambda_3}(N = l_2 + 1) \leq \mathbb{P}_{\lambda_2}(N = l_2 + 1). \quad (1)$$

Aus der ersten Ungleichung erhält man eine untere Grenze für λ_2 :

$$\lambda_3^{l_2} \exp(-\lambda_3) \geq \lambda_2^{l_2} \exp(-\lambda_2) \quad \text{impliziert} \quad \lambda_2 - l_2 \log \lambda_2 \geq \lambda_3 - l_2 \log \lambda_3,$$

und da $\lambda_3 = l_2!^{1/l_2}$ ergibt dies

$$\lambda_2 - l_2 \log \lambda_2 \geq l_2!^{1/l_2} - \log(l_2!).$$

Eine obere Grenze folgt aus der zweiten Ungleichung in 1 und lässt sich analog bestimmen:

$$\lambda_2 - (l_2 + 1) \log \lambda_2 \leq l_2!^{1/l_2} - \frac{l_2 + 1}{l_2} \log l_2!.$$

Hält man die untere Schranke l_2 fest, ist das dazu optimale λ_3 bereits bestimmt und man befindet sich wieder im bekannten Milieu: Zu bestimmen bleibt nur, wie in Abhängigkeit von λ_1 die obere Schranke l_1 zu wählen ist, damit die Detektionswahrscheinlichkeit maximiert wird. Dies gleicht der Frage im vorherigen Abschnitt. Gesucht sind also die Werte von λ_1 , für die die obere Schranke gerade um 1 nach oben springt. Bezeichne

$$\lambda_2^{l_1, l_2} := \left(\frac{l_1!}{l_2!} \right)^{1/(l_1-l_2)},$$

dann ist die Lösung λ_1 dieser Gleichung für festes l_1 und l_2 gesucht:

Korollar 2.4. *Gegeben 3 Stimuli S_1, S_2, S_3 mit $\tau_1 = \tau_2 = \tau_3$ und $\lambda_3 < \lambda_2 < \lambda_1$. Für eine feste untere Schranke $l_2 < l_1$ lassen sich die Sprungstellen der optimalen oberen Schranke l_1 bestimmen über*

$$\mathbb{P}_{\lambda_2^{l_1, l_2}}(l_2 + 1 \leq N \leq l_1) + \mathbb{P}_{\lambda_1}(N = l_1 + 1) = \mathbb{P}_{\lambda_2^{l_1+1, l_2}}(l_2 + 1 \leq N \leq l_1 + 1).$$

Beweis. Analog zu Lemma 2.4 ist die Gleichung

$$\mathbb{P}_{\lambda_2^{l_1, l_2}}(l_2 + 1 \leq N \leq l_1) + \mathbb{P}_{\lambda_1}(N \geq l_1 + 1) = \mathbb{P}_{\lambda_2^{l_1+1, l_2}}(l_2 + 1 \leq N \leq l_1 + 1) + \mathbb{P}_{\lambda_1}(N \geq l_1 + 2)$$

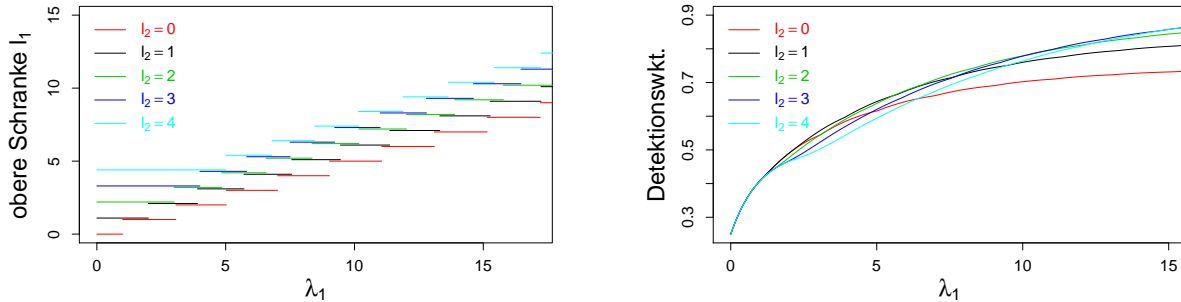
zu lösen. \square

In Abbildung 7 ist die optimale obere Schranke für verschiedene untere Schranken in Abhängigkeit von λ_1 dargestellt. Zum besseren Auseinanderhalten sind die Treppenfunktionen leicht gegeneinander verschoben, auftreten können aber nur ganzzahlige Schranken. Die Detektionswahrscheinlichkeit lässt sich für jede Schrankenkombination bestimmen über

$$p_D^{l_1, l_2} = \frac{1}{4} \left(1 + \mathbb{P}_{\lambda_3^{l_2}}(0 < N \leq l_2) + \mathbb{P}_{\lambda_2^{l_1, l_2}}(l_2 + 1 \leq N \leq l_1) + \mathbb{P}_{\lambda_1}(N \geq l_1 + 1) \right)$$

Für $\lambda_1 \leq 1$ besitzen alle Schrankenkombinationen dieselbe Detektionswahrscheinlichkeit, da sich stets nur die Gewichte $\mathbb{P}_{\lambda_1}(N > 0)$ ergeben. Man beachte, dass z.B für $l_2 = 4$ es nicht möglich ist, $\lambda_3 = 4!^{1/4}$ zu setzen, da $\lambda_3 \leq \lambda_1$ gefordert ist, und damit das bestmögliche $\lambda_3 = \lambda_1$ ist.

Weiter fällt auf, dass sich die Funktionen nach ihrem anfänglichen gemeinsamen Verlauf genau einmal schneiden. Zuerst liegt die Funktion mit der geringeren unteren Schranke oberhalb, ab einem bestimmten λ_1 dominiert aber die obere Schranke. Zuerst schneiden sich aber die Funktionen, die sich genau um 1 in der unteren Schranke unterscheiden. Dabei dominiert jede Funktion solange alle anderen, bis sie von ihrem Nachfolger, mit der um 1 erhöhten unteren Schranke, geschnitten wird. Demnach erwartet man folgenden Verlauf



(a) Optimale obere Schranke l_1 für verschiedene feste untere Schranken l_2 in Abhängigkeit von λ_1 .

(b) Maximale Detektionswahrscheinlichkeit für verschiedene festgehaltene untere Schranken l_2 in Abhängigkeit von λ_1 .

Abbildung 7: Optimale obere Schranke l_1 und maximale Detektionswahrscheinlichkeit für festgehaltene untere Schranke l_2 in Abhängigkeit von λ_1 . Oberhalb von l_1 entscheidet man sich für Stimulus 1, oberhalb von l_2 bis einschließlich l_1 für Stimulus 2 und bei einem bis maximal l_2 Spike für Stimulus 3.

für die optimalen Schranken: Die untere Schranke l_2 springt bei gewissen Werten von λ_1

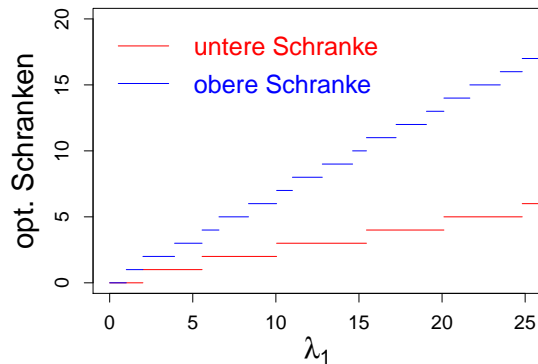
stets um 1 nach oben, die obere Schranke l_1 kann in der Zwischenzeit häufiger Sprünge einlegen.

Die Sprungstellen der unteren Schranke ließen sich demnach bestimmen, indem man folgende Gleichung nach λ_1 für die „richtigen“ oberen Schranken l_1 und l'_1 löst:

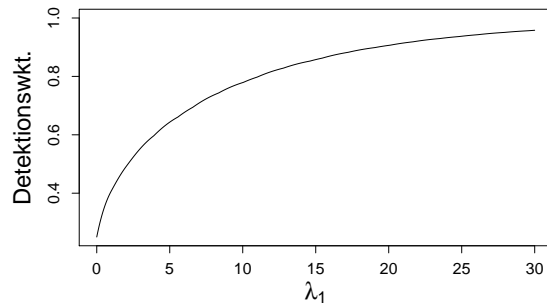
$$\begin{aligned} & \mathbb{P}_{\lambda_3^{l_2}}(0 < N \leq l_2) + \mathbb{P}_{\lambda_2^{l_1, l_2}}(l_2 < N \leq l_1) + \mathbb{P}_{\lambda_1}(N > l_1) \\ &= \mathbb{P}_{\lambda_3^{l_2+1}}(0 < N \leq l_2 + 1) + \mathbb{P}_{\lambda_2^{l'_1, l_2+1}}(l_2 + 1 < N \leq l'_1) + \mathbb{P}_{\lambda_1}(N > l'_1). \end{aligned}$$

Was sind aber die „richtigen“ oberen Schranken? Dazu muss man die Gleichung für alle l_1 und l'_1 lösen und hat erst dann die richtige Lösung gefunden, wenn die oberen Schranken l_1 und l'_1 auch für die Lösung λ_1 der Gleichung die optimalen oberen Schranken sind für die jeweils untere Schranke l_2 und $l_2 + 1$.

Oder man bestimmt stur über alle Schrankenkombinationen (untere Schranke und dazu die optimale obere Schranke) die maximale Detektionswahrscheinlichkeit. Auf diese Weise erhält man die in Abbildung 8 dargestellten optimalen Schranken. Die obere Schranke springt dabei öfters und unregelmäßiger als die untere Schranke. Außerdem fällt auf, dass



(a) Die optimalen oberen und unteren Schranken in Abhängigkeit von λ_1 .



(b) Die maximale Detektionswahrscheinlichkeit in Abhängigkeit von λ_1 .

Abbildung 8: Optimale obere und untere Schranke und die daraus resultierende Detektionswahrscheinlichkeit in Abhängigkeit von λ_1 für 3 Stimuli.

die obere Schranke auch immer dann springt, wenn die untere Schranke springt. Für kleine λ_1 sind dies gerade die Stellen, an denen die obere Schranke auch für die niedrigere untere Schranke gesprungen wäre. Dies bleibt aber nicht für größere λ_1 erhalten.

Um zu den optimalen Schranken im Fall von drei Stimuli zu gelangen, haben wir die untere Schranke festgehalten, dazu jeweils die in Abhängigkeit von λ_1 beste obere Schranke gesucht und dann die Detektionswahrscheinlichkeit über diese Möglichkeiten maximiert.

Für mehr als drei Stimuli ist dieses Vorgehen nicht praktikabel:

Schon im Fall von vier Stimuli müssten zwei Schranken festgehalten werden, und in Abhängigkeit davon die beste oberste Schranke bestimmt werden. Dies müsste für alle möglichen Kombinationen der beiden unteren Schranken durchgeführt werden und über alle Möglichkeiten müsste dann die maximale Detektionswahrscheinlichkeit gesucht werden.

Bemerkung 2.5. *In Abbildung 8 lässt sich aber eine elegantere Herangehensweise erkennen: Die obere Schranke l_1 springt immer auch dann, wenn die untere Schranke l_2 springt. Also ist jede obere Schranke genau einer unteren Schranke zugeordnet. Der Trick ist, die umgekehrte Herangehensweise einzuschlagen: In Abhängigkeit von der obersten Schranke l_1 sind die optimalen unteren Schranken gesucht. Diese Idee wird im nächsten Abschnitt verfolgt.*

2.3.4 Beliebige Anzahl Stimuli

Betrachtet werden nun $k \in \mathbb{N}$ Stimuli mit $\tau_1 = \dots, \tau_k$ und $0 < \lambda_k < \dots < \lambda_1 \leq \lambda_{\max}$. Gesucht ist die Parameterwahl in Abhängigkeit von λ_1 , sodass

$$p_D^{(0, l_{k-1}, l_{k-2}, \dots, l_1)} = \frac{1}{k} \left(1 + \sum_{i=2}^k \mathbb{P}_{\lambda_i}(l_i < N \leq l_{i-1}) + \mathbb{P}_{\lambda_1}(N > l_1) \right),$$

die Detektionswahrscheinlichkeit, für frei-wählbare Schranken $L = (0, l_{k-1}, l_{k-2}, \dots, l_1)$ maximiert wird. Mit den Ergebnissen aus den vorherigen Abschnitten lässt sich nun ein Algorithmus entwickeln, mit dem die optimalen Ratenparametern effizient berechnet werden können.

Aus Lemma 2.2 und Lemma 2.5 ist bekannt:

Satz 2.5. *Sei $\tau_1 = \dots = \tau_k$ und $0 < \lambda_k < \dots < \lambda_1 \leq \lambda_{\max}$, $\lambda_{\max} \in \mathbb{R}^+$ fest. Dann sind für feste Schranken $L = (0, l_{k-1}, \dots, l_1)$ die dazu optimalen Ratenparameter gegeben durch*

$$\lambda_i = \left(\frac{l_{i-1}!}{l_i!} \right)^{\frac{1}{l_{i-1} - l_i}}, \quad i = 2, \dots, k.$$

Die eigentliche Aufgabe ist damit, die optimalen Schranken L zu finden. Hält man die oberste Schranke l_1 fest, sind die optimalen unteren Schranken eindeutig festgelegt.

Definition 2.4. *Für k Stimuli bezeichne $L^k(l_1)$ die optimalen Schranken l_2, \dots, l_k in Abhängigkeit von l_1 , sodass also*

$$p_D^{L^k(l_1)} \geq p_D^{(l_k, l_{k-1}, \dots, l_1)} \quad \forall l_k < l_{k-1} < \dots < l_2 < l_1$$

gilt. $L_i^k(l_1)$ bezeichne die optimale i -größte Schranke l_i .

Die Idee dahinter ist, dass die unteren Schranken erst dann springen können, wenn die oberste l_1 springt. In der Zwischenzeit bleiben alle gleich.

Bemerkung 2.6. Für kleines λ_1 sind noch nicht alle Schranken $l_i > 0$ und damit liegen mehrere Stimuli mit Rate $\lambda_i = 0$ vor. Diese werden nicht detektiert, da in diesem Fall $l_i = l_{i-1}$ ist.

Beispiel k = 3 Stimuli: Gesucht sind die zwei optimalen Schranken l_1 und l_2 . Definiere wieder

$$\lambda^{l_{i+1}, l_i} := \left(\frac{l_i!}{l_{i+1}!} \right)^{1/(l_i - l_{i+1})},$$

der optimale Parameter für das Schrankenpaar (l_{i+1}, l_i) .

Für jedes $l_1 \in \mathbb{N}$ ist die dazu optimale untere Schranke $l_2 \leq l_1$ zu bestimmen. Für $l_1 = 0$ ist die einzig mögliche Wahl $l_2 = 0$. Für $l_1 > 0$ ist die minimale Schranke l_2 gesucht, sodass gilt:

$$\mathbb{P}_{\lambda^{0, l_2}}(0 < N \leq l_2) + \mathbb{P}_{\lambda^{l_2, l_1}}(l_2 < N \leq l_1) \geq \mathbb{P}_{\lambda^{0, l_2+1}}(0 < N \leq l_2 + 1) + \mathbb{P}_{\lambda^{l_2+1, l_1}}(l_2 + 1 < N \leq l_1).$$

In dieser Gleichung kommt nicht das λ_1 vor. Erst beim Bestimmen der Sprungstellen der Schranken wird die bestmögliche Wahl in Abhängigkeit von λ_1 berechnet. Zu lösen ist die Gleichung

$$\begin{aligned} & \mathbb{P}_{\lambda^{0, l_2}}(0 < N \leq l_2) + \mathbb{P}_{\lambda^{l_2, l_1}}(l_2 + 1 < N \leq l_1) + \mathbb{P}_{\lambda_1}(N > l_1) \\ &= \mathbb{P}_{\lambda^{0, l'_2}}(0 < N \leq l'_2) + \mathbb{P}_{\lambda^{l'_2, l_1+1}}(l'_2 + 1 < N \leq l_1 + 1) + \mathbb{P}_{\lambda_1}(N > l_1 + 1) \end{aligned}$$

nach λ_1 für $l_2 = L_2^3(l_1)$ und $l'_2 = L_2^3(l_1 + 1)$.

Beispiel k = 4 Stimuli: Dieses Vorgehen lässt sich leicht auf den Fall $k = 4$ übertragen. Hier sind die Schranken l_3 , l_2 und l_1 gesucht. Aus den optimalen Schrankenkombinationen für $k = 3$ lassen sich die optimalen Kombinationen im Fall $k = 4$ bestimmen. In Abhängigkeit von $l_1 \in \mathbb{N}$ muss folgende Gleichung nach dem minimalem $l_2 \leq l_1$ gelöst werden:

$$\begin{aligned} & \mathbb{P}_{\lambda^{0, l_3}}(0 < N \leq l_3) + \mathbb{P}_{\lambda^{l_3, l_2}}(l_3 < N \leq l_2) + \mathbb{P}_{\lambda^{l_2, l_1}}(l_2 < N \leq l_1) \\ & \geq \mathbb{P}_{\lambda^{0, l'_3}}(0 < N \leq l'_3) + \mathbb{P}_{\lambda^{l'_3, l_2+1}}(l'_3 < N \leq l_2 + 1) + \mathbb{P}_{\lambda^{l_2+1, l_1}}(l_2 + 1 < N \leq l_1), \end{aligned}$$

für $l_3 = L_2^3(l_2)$ und $l'_3 = L_2^3(l_2 + 1)$.

Die optimalen Schranken für den Fall k lassen sich also aus den optimalen Schranken im Fall $k - 1$ bestimmen, da durch die 2-te Schranke alle niedrigeren festgelegt sind und damit aus dem Fall $k - 1$ bekannt sind. Es muss also nur untersucht werden, welches die optimale zweite Schranke ist für vorgegebenes λ_1 , die niedrigeren Schranken sind dazu dann schon bekannt.

Die Sprungstellen findet man wieder, indem man für festes l_1

$$p_D^{(l_4, l_3, \dots, l_1)}(\lambda_1) = p_D^{(l'_4, l'_3, \dots, l_1+1)}(\lambda_1) \quad (2)$$

nach λ_1 löst. Dabei ist $l_i = L_i^4(l_1)$ und $l'_i = L_i^4(l_1 + 1)$.

Beachtet werden muss aber, dass nicht jede als optimal bestimmte Schrankenkombination auch tatsächlich vorkommt. Es kann passieren, dass die in Gleichung 2 bestimmten λ_1 sich für l_1 und $l_1 + 1$ überschneiden, die obere Schranke $l_1 + 2$ bereits besser ist als $l_1 + 1$, bevor die Schranke $l_1 + 1$ die Schranke l_1 dominiert. Dann kommt $l_1 + 1$ als oberste Schranke überhaupt nicht vor und folgende Gleichung ist zu lösen:

$$p_D^{(l_4, l_3, \dots, l_1)}(\lambda_1) = p_D^{(l'_4, l'_3, \dots, l_1 + 2)}(\lambda_1),$$

für $l_i = L_i^4(l_1)$ und $l'_i = L_i^4(l_1 + 2)$.

Zusammenfassend lässt sich im allgemeinen Fall folgendes Vorgehen festhalten:

Algorithmus. Induktiv: Verwende die optimalen Schranken im Fall von $k - 1$ Stimuli um die Schranken für k Stimuli zu berechnen.

1. Bestimme die optimale Schrankenkombinationen: Suche für l_1 fest das minimale l_2 , sodass gilt:

$$\begin{aligned} & \sum_{i=2}^{k-1} \mathbb{P}_{\lambda_i}(l_i < N \leq l_{i-1}) \\ & \geq \sum_{i=4}^{k-1} \mathbb{P}_{\lambda'_i}(l'_i < N \leq l'_{i-1}) + \mathbb{P}_{\lambda'_3}(l'_3 < N \leq l_2 + 1) + \mathbb{P}_{\lambda'_2}(l_2 + 1 < N \leq l_1), \end{aligned}$$

wobei $l_i = L_{i-1}^{k-1}(l_2)$ und $l'_i = L_{i-1}^{k-1}(l_2 + 1)$ für $i = 3, \dots, k - 1$.

2. Bestimme die Sprungstellen auf die nächste Schrankenkombination: Löse die Gleichung für festes l_1

$$p_D^{(l_{k-1}, l_{k-2}, \dots, l_1)}(\lambda_1) = p_D^{(l'_{k-1}, l'_{k-2}, \dots, l_1 + 1)}(\lambda_1)$$

nach λ_1 für $l_i = L_i^k(l_1)$ und $l'_i = L_i^k(l_1 + 1)$.

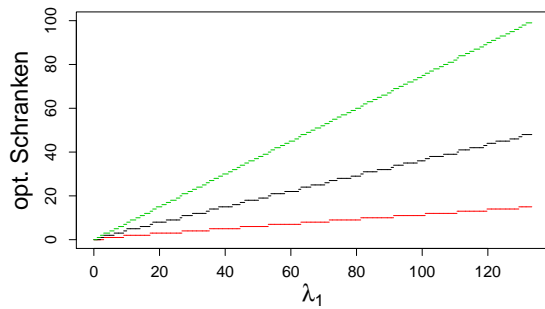
3. Teste, ob die Lösungen λ für l_1 und $l_1 + 1$ monoton wachsen. Ansonsten entferne Schrankenkombination mit oberer Schranke $l_1 + 1$ und löse die Gleichung

$$p_D^{(l_{k-1}, l_{k-2}, \dots, l_1)}(\lambda_1) = p_D^{(l'_{k-1}, l'_{k-2}, \dots, l_1 + 2)}(\lambda_1)$$

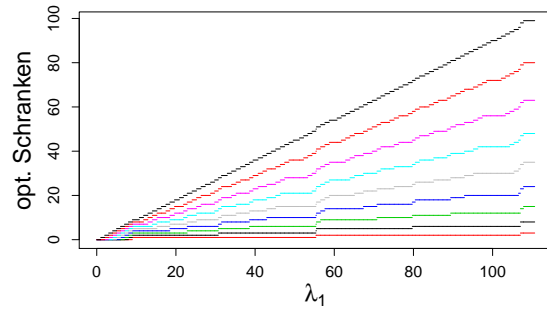
nach λ_1 für $l_i = L_i^k(l_1)$ und $l'_i = L_i^k(l_1 + 2)$.

In Abbildung 9 sind die optimalen Schranken für eine verschiedene Anzahl Stimuli berechnet.

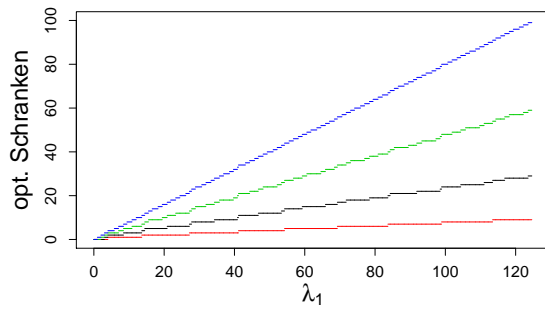
In Abbildung 10 sind deren Detektionswahrscheinlichkeiten gegenübergestellt. Wie zu erwarten, nimmt die Detektionswahrscheinlichkeit mit einer wachsenden Anzahl Stimuli sehr



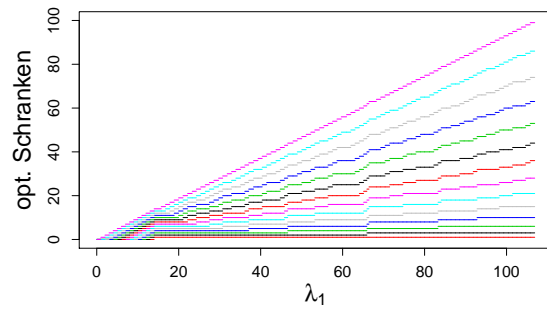
(a) Optimale Schranken für 4 Stimuli.



(c) Optimale Schranken für 10 Stimuli.



(b) Optimale Schranken für 5 Stimuli.



(d) Optimale Schranken für 15 Stimuli.

Abbildung 9: Optimale Schranken in Abhängigkeit von λ_1 für verschieden Anzahl Stimuli. In den eingegrenzten Gebieten entscheidet man sich für die jeweiligen der Größe nach geordneten Stimuli. Im Falle von Null Spikes entscheidet man sich für den Nullstimulus.

stark ab.

In Abbildung 11 sind die optimalen λ_2 in Abhängigkeit von λ_1 für unterschiedliche Anzahl Stimuli dargestellt.

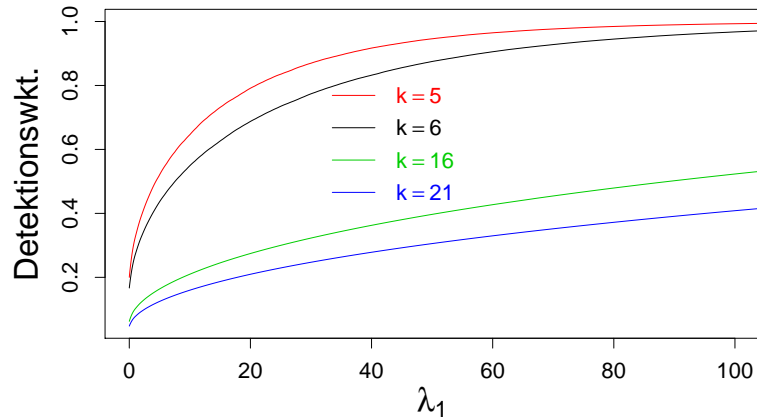


Abbildung 10: Maximale Detektionswahrscheinlichkeit in Abhängigkeit von λ_1 für 4, 5, 15 und 20 Stimuli.

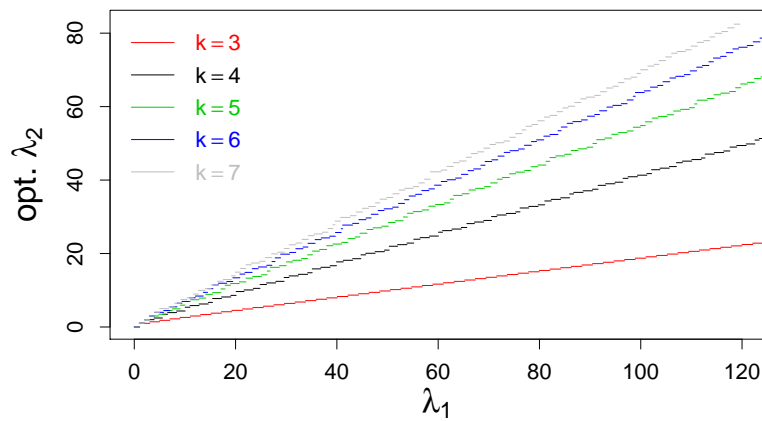


Abbildung 11: Für verschiedene Anzahl Stimuli sind die optimalen λ_2 in Abhängigkeit von λ_1 dargestellt. Die optimalen Rateparameter sind eindeutig durch die optimalen Schranken festgelegt, vergleiche Satz 2.5.

2.3.5 Asymptotisch beste Wahl

Mit dem vorgestellten Algorithmus lassen sich numerisch die optimalen λ_i , $i = 1, \dots, k$, berechnen. In den Grafiken lässt sich erkennen, dass abgesehen von den Sprungstellen, die auf den diskreten Werten der Poisson-Verteilung beruhen, ein in etwa linearer Zusammenhang zwischen λ_1 und den restlichen λ_i besteht. Für λ_1 groß genug, sollte man, ohne Rechnen zu müssen, wissen, welche Wahl der übrigen λ_i , $i = 1, \dots, k$, die Detektionswahrscheinlichkeit maximiert.

Die Poisson-Verteilung lässt sich für große λ auf Grund ihrer Reproduktivität nach dem Zentralen Grenzwertsatz durch die Normalverteilung approximieren. Reproduktivität bedeutet:

Lemma 2.6. *Für unabhängige $N_1 \sim Pois(\lambda_1)$ und $N_2 \sim Pois(\lambda_2)$ gilt:*

$$N_1 + N_2 \sim Pois(\lambda_1 + \lambda_2).$$

Beweis. Leicht lässt sich berechnen

$$\begin{aligned} \mathbb{P}(N_1 + N_2 = n) &= \sum_{k=0}^n \mathbb{P}(N_1 = k) \mathbb{P}(N_2 = n - k) \\ &= \sum_{k=0}^n \frac{\lambda_1^k}{k!} e^{-\lambda_1} \frac{\lambda_2^{n-k}}{(n-k)!} e^{-\lambda_2} \\ &= \frac{1}{n!} e^{-(\lambda_1 + \lambda_2)} \sum_{k=0}^n \binom{n}{k} \lambda_1^k \lambda_2^{n-k} \\ &= \frac{(\lambda_1 + \lambda_2)^n}{n!} e^{-(\lambda_1 + \lambda_2)}. \end{aligned}$$

□

Aus der Reproduktivität und dem Satz von Berry-Essen (für einen Beweis siehe z.B. [Senatov (1998)]) folgt mit einer kurzen Rechnung dieses bekannte Resultat über die Poisson-Verteilung:

Lemma 2.7. *Für $N \sim Pois(\lambda)$ gilt:*

$$\mathbb{P}(N \leq x) = \Phi\left(\frac{x - \lambda}{\sqrt{\lambda}}\right) + O(\lambda^{-1/2}),$$

wobei $\Phi(\cdot)$ für die Verteilungsfunktion der Standardnormalverteilung steht.

Beweis. Nach Lemma 2.6 ist für $N \sim Pois(\lambda)$

$$N \cong X_1 + X_2 + \dots + X_n,$$

wobei $X_i \sim \text{Pois}\left(\frac{\lambda}{n}\right)$ und unabhängig. Nach dem Zentralen Grenzwertsatz gilt für λ groß (n kann hier beliebig groß gewählt werden)

$$\frac{N - \lambda}{\sqrt{\lambda}} \stackrel{a}{\sim} \mathcal{N}(0, 1).$$

Bezeichne $F_\lambda(x)$ die Verteilungsfunktion von N und $\Phi(x)$ die Verteilungsfunktion der Standardnormalverteilung. Nach dem Satz von Berry-Essen lässt sich der Approximationsfehler abschätzen durch

$$|F_\lambda(x) - \Phi(x)| \leq \frac{C \mathbb{E}[X^3]}{\sigma_X^3 \sqrt{n}},$$

für eine Konstante C .

Das dritte Moment einer Poissonverteilung beträgt

$$\begin{aligned} \mathbb{E}[N^3] &= \sum_{k=0}^{\infty} k^3 \frac{\lambda^k}{k!} e^{-\lambda} \\ &= \sum_{k=0}^{\infty} (k(k-1)(k-2) + 3k(k-1) + k) \frac{\lambda^k}{k!} e^{-\lambda} \\ &= \sum_{k=3}^{\infty} \frac{\lambda^k}{(k-3)!} e^{-\lambda} + 3 \sum_{k=2}^{\infty} \frac{\lambda^k}{(k-2)!} e^{-\lambda} + \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\ &= \lambda^3 e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} + 3\lambda^2 e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} + \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\ &= \lambda^3 + 3\lambda^2 + \lambda. \end{aligned}$$

Damit ergibt sich für den Approximationsfehler

$$\begin{aligned} |F_\lambda(x) - \Phi(x)| &\leq \frac{C \left(\frac{\lambda^3}{n^3} + 3 \frac{\lambda^2}{n^2} + \frac{\lambda}{n} \right)}{\sqrt{\lambda/n}^3 \sqrt{n}} \\ &= \frac{C \left(\frac{\lambda^2}{n^2} + 3 \frac{\lambda}{n} + 1 \right)}{\sqrt{\lambda}}. \end{aligned}$$

Da n beliebig groß gewählt werden kann, folgt insgesamt

$$\mathbb{P}(N \leq x) = \Phi\left(\frac{x - \lambda}{\sqrt{\lambda}}\right) + O\left(\lambda^{-1/2}\right)$$

□

Aus Lemma 2.2 ist bekannt, dass stets der Nullstimulus ($\lambda_k = 0$) als optimaler Zustand vertreten ist. Im einfachsten Fall von zwei Stimuli war das ursprüngliche Ziel gewesen, zwei

Poisson-Verteilungen möglichst stark voneinander zu trennen und möglichst keine Masse auf der Null zu haben. Es sollte also für vorgegebenes λ_1

$$p_D(\lambda_2) = \mathbb{P}(0 < N_2 \leq l) + \mathbb{P}(N_1 > l), \quad \text{wobei } l = \frac{\lambda_1 - \lambda_2}{\log(\lambda_1/\lambda_2)},$$

nach $\lambda_2 \leq \lambda_1$ maximiert werden.

Für großes λ_1 lassen sich nach Satz 2.7 anstatt der beiden Poisson-Verteilungen zwei Normalverteilungen betrachten.

Lemma 2.8. Für $\lambda_2 \leq \lambda_1$ fest und $N_1 \sim \mathcal{N}(\lambda_1, \lambda_1)$ und $N_2 \sim \mathcal{N}(\lambda_2, \lambda_2)$ wird

$$p_D(\lambda_2) = \mathbb{P}(0 < N_2 \leq l) + \mathbb{P}(N_1 > l)$$

maximiert durch

$$l = \sqrt{\frac{\log(\lambda_1) - \log(\lambda_2) + \lambda_1 - \lambda_2}{1/\lambda_2 - 1/\lambda_1}}.$$

Beweis. Die optimale Grenze l , bis zu der man sich für die jeweilige Verteilung entscheidet, also der Schnittpunkt der beiden Dichten, erhält man wieder über den Maximum-Likelihood-Ansatz, vgl. Lemma 2.1:

$$\begin{aligned} & \varphi_{\lambda_1}(l) = \varphi_{\lambda_2}(l) \\ \Leftrightarrow & \frac{1}{\sqrt{2\pi\lambda_1}} e^{-\frac{(l-\lambda_1)^2}{2\lambda_1}} = \frac{1}{\sqrt{2\pi\lambda_2}} e^{-\frac{(l-\lambda_2)^2}{2\lambda_2}} \\ \Leftrightarrow & \frac{(l-\lambda_1)^2}{\lambda_1} + \log(\lambda_1) = \frac{(l-\lambda_2)^2}{\lambda_2} + \log(\lambda_2) \\ \Leftrightarrow & (\lambda_1 - \lambda_2)l^2 + \lambda_1\lambda_2^2 - \lambda_1^2\lambda_2 = \lambda_1\lambda_2(\log(\lambda_1) - \log(\lambda_2)) \\ \Rightarrow & l = \sqrt{\frac{\log(\lambda_1) - \log(\lambda_2) + \lambda_1 - \lambda_2}{1/\lambda_2 - 1/\lambda_1}}. \end{aligned}$$

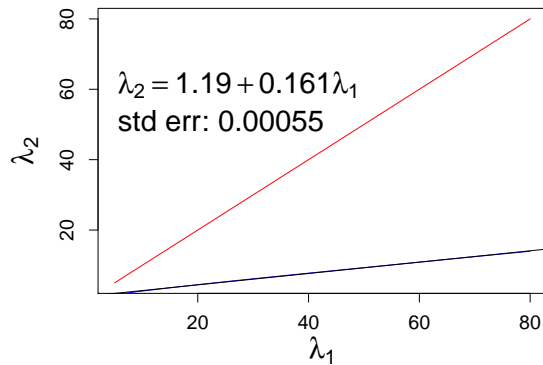
□

Gesucht ist jetzt also $\lambda_2 \leq \lambda_1$, sodass

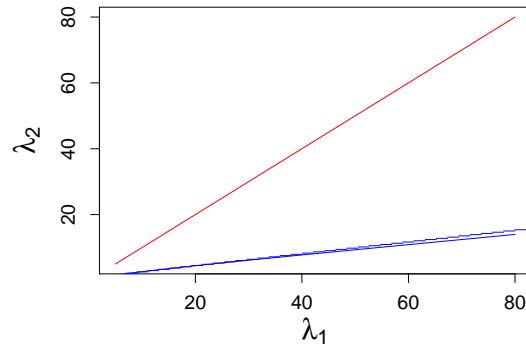
$$p_D(\lambda_2) := \Phi\left(\frac{l-\lambda_2}{\sqrt{\lambda_2}}\right) - \Phi\left(\frac{0-\lambda_2}{\sqrt{\lambda_2}}\right) + 1 - \Phi\left(\frac{l-\lambda_1}{\sqrt{\lambda_1}}\right)$$

für $\lambda_1 \rightarrow \infty$ maximiert wird.

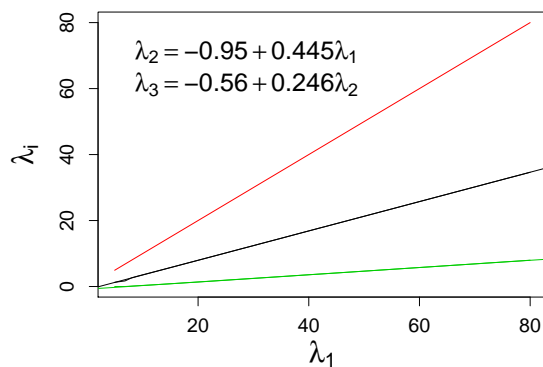
Maximiert man $p_D(\lambda_2)$ mit der Normalapproximation numerisch für vorgegebenes λ_1 , so lässt sich das approximierte λ_2 mit dem exakten λ_2 vergleichen. In Abbildung 12 erkennt man auf der rechten Seite, dass die Normalapproximation bereits für kleine λ_1 sehr ähnliche Ergebnisse zu der exakten Berechnung liefert.



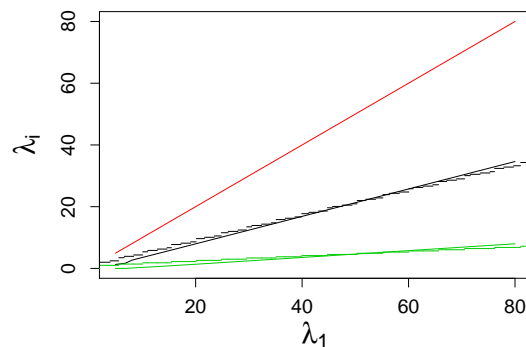
(a) Optimales λ_2 in Abhängigkeit von λ_1 gemäß Normalapproximation.



(c) Vergleich optimales λ_2 gemäß Normalapproximation mit numerisch exaktem optimalen λ_2 .



(b) Optimales λ_2 und λ_3 in Abhängigkeit von λ_1 gemäß Normalapproximation.



(d) Vergleich optimales λ_2 und λ_3 gemäß Normalapproximation mit numerisch exaktem optimalen λ_2 und λ_3 .

Abbildung 12: Vergleich exakte Lösung mit Normalapproximation für 2 und 3 Stimuli. Die Regressionsgeraden in (a) und (b) bestätigen bereits für kleine λ_1 einen annähernd linearen Zusammenhang zwischen λ_i und λ_1 .

Auf der linken Seite ist der lineare Zusammenhang bei der approximierten Lösung dargestellt, der schon für kleine λ_1 sehr stark vorhanden ist. Weiterhin fällt auf, dass bei 3 Stimuli für λ_3 in Abhängigkeit von λ_2 etwa derselbe Faktor auftritt wie bei 2 Stimuli für λ_2 in Abhängigkeit von λ_1 . Für $\lambda_1 \rightarrow \infty$ wird dieser identisch sein, da das kleinste λ_i nur von dem nächst höheren λ_{i-1} abhängt und dies gerade dem Zwei-Stimuli-Fall entspricht. Bei k -Stimuli werden dementsprechend dieselben Faktoren wie bei $(k-1)$ -Stimuli auftreten, bis auf das Verhältnis λ_2/λ_1 .

Der Faktor 0.25 bei 2-Stimuli deutet schon sehr stark auf eine asymptotisch optimale Lösung von $\lambda_2/\lambda_1 = 1/4$. $p_D(\cdot)$ ist jedoch sehr unhandlich, leichter geht es direkt über die Schranke l .

Lemma 2.9. Sei $\lambda_2 \leq \lambda_1$ und $N_1 \sim \mathcal{N}(\lambda_1, \lambda_1)$ und $N_2 \sim \mathcal{N}(\lambda_2, \lambda_2)$. Für $\lambda_1 \rightarrow \infty$ vereinfacht sich die Entscheidungsschranke aus Lemma 2.8

$$l = \sqrt{\frac{\log(\lambda_1) - \log(\lambda_2) + \lambda_1 - \lambda_2}{1/\lambda_2 - 1/\lambda_1}}$$

zu

$$l \approx \sqrt{c} \lambda_1,$$

für $c = \lambda_2/\lambda_1$.

Beweis. Substituiert man $\lambda_2 = c\lambda_1$, wobei $0 < c < 1$, erhält man für die Schranke l

$$\begin{aligned} l^2 &= \frac{\lambda_1 - c\lambda_1 + \log(\lambda_1) - \log(c\lambda_1)}{1/c\lambda_1 - 1/\lambda_1} \\ &= \frac{\lambda_1(1-c) - \log(c)}{1/\lambda_1(1/c - 1)} \\ &= c\lambda_1^2 - \frac{c}{1-c}\lambda_1 \log(c). \end{aligned}$$

Asymptotisch vereinfacht sich die Entscheidungsschranke l zu

$$l \approx \sqrt{c} \lambda_1.$$

□

In Abbildung 13 sind die Entscheidungsschranke und die Parameter der beiden Normalverteilungen dargestellt.

Lemma 2.10. Für $\lambda_2 = c \cdot \lambda_1$, ($0 < c < 1$), und $N_1 \sim \mathcal{N}(\lambda_1, \lambda_1)$ und $N_2 \sim \mathcal{N}(\lambda_2, \lambda_2)$ wird

$$p_D(\lambda_2) = \mathbb{P}(0 < N_2 \leq l) + \mathbb{P}(N_1 > l)$$

asymptotisch ($\lambda \rightarrow \infty$) maximiert durch $l = \sqrt{c} \lambda_1$ und

$$c = \frac{1}{4}.$$

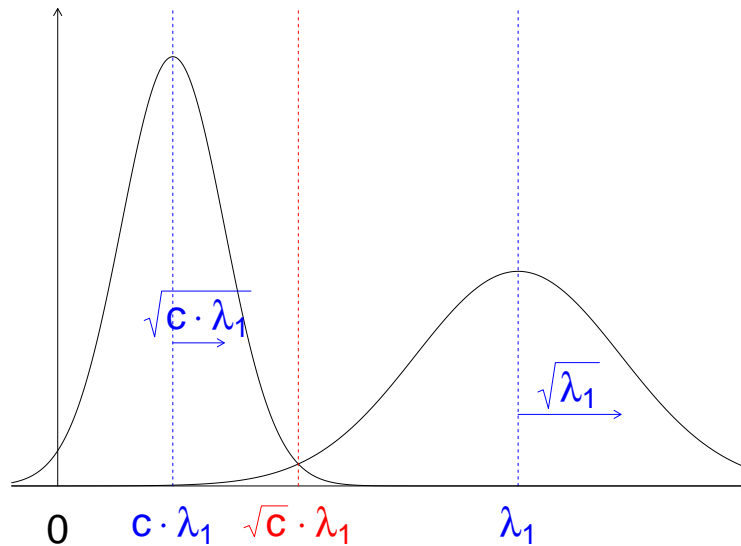


Abbildung 13: Trennung von zwei Normalverteilungen $N_1 \sim \mathcal{N}(\lambda_1, \lambda_1)$ und $N_2 \sim \mathcal{N}(c\lambda_1, c\lambda_1)$. Für $\lambda_1 \rightarrow \infty$ und $0 < c < 1$ liegt der Schnittpunkt der Dichten bei $\sqrt{c}\lambda_1$.

Beweis. Gegeben die Entscheidungsschranke l kann λ_2 nur dann die optimale Lösung sein, wenn λ_2 genau in der Mitte von 0 und l liegt. Sonst ließe sich bei festgehaltener Schranke l durch mittige Platzierung eine höhere Detektionswahrscheinlichkeit erreichen.

Nach Lemma 2.9 ist asymptotisch die optimale Schranke gegeben durch $l = \sqrt{c}\lambda_1$. Für das optimale $\lambda_2 = c\lambda_1$ muss damit gelten:

$$\begin{aligned} c\lambda_1 &= \frac{\sqrt{c}}{2}\lambda_1 \\ \implies 4c^2 &= c \\ \implies c(1 - 4c) &= 0. \end{aligned}$$

Da $c = 0$ offensichtlich die Detektionswahrscheinlichkeit minimiert, ist die eindeutige Lösung

$$c = \frac{1}{4}.$$

□

Beispiel k = 3: Analog lässt sich die asymptotisch beste Parameterwahl im Fall von drei Normalverteilungen ($\lambda_3 \leq \lambda_2 \leq \lambda_1$) bestimmen, dargestellt in Abbildung 14. Das optimale Verhältnis $c_3 := \lambda_3/\lambda_2$ entspricht gerade wieder der Lösung für 2 Verteilungen, also $c_3 =$

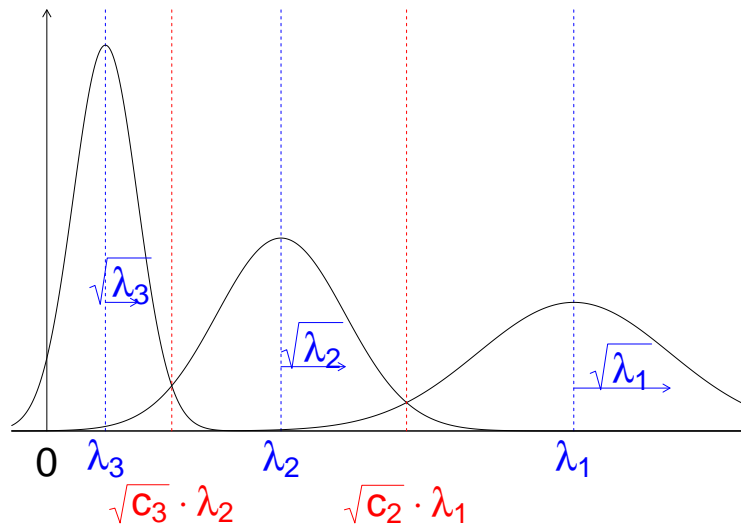


Abbildung 14: Trennung von drei Normalverteilungen $N_1 \sim \mathcal{N}(\lambda_1, \lambda_1)$, $N_2 \sim \mathcal{N}(c_2\lambda_1, c_2\lambda_1)$ und $N_3 \sim \mathcal{N}(c_3\lambda_2, c_3\lambda_2)$. Für $\lambda_1 \rightarrow \infty$ liegen die Schnittpunkte der benachbarten Dichten bei $\sqrt{c_3}\lambda_2$ und $\sqrt{c_2}\lambda_1$.

1/4. Das optimale λ_2 muss wieder in der Mitte der beiden Schranken liegen. Dies ergibt für $c_2 := \lambda_2/\lambda_1$

$$\begin{aligned} \frac{\sqrt{c_2} + \sqrt{c_3 c_2}}{2} \lambda_1 &= c_2 \lambda_1 \\ \xrightarrow{c_3=1/4} \quad \frac{3}{2} c_2 &= \sqrt{c_2} \\ \implies c_2 \left(1 - \frac{9}{4} c_2\right) &= 0. \end{aligned}$$

Für drei Normalverteilungen ist asymptotisch die beste Wahl gegeben durch

$$\frac{\lambda_3}{\lambda_2} = \frac{1}{4} = \left(\frac{1}{2}\right)^2 \quad \text{und} \quad \frac{\lambda_2}{\lambda_1} = \frac{4}{9} = \left(\frac{2}{3}\right)^2.$$

Die allgemeine Lösung lässt sich so schon erahnen und wir beweisen folgendes Resultat:

Satz 2.6. Für k Stimuli mit $\tau_k = \dots = \tau_1$ und $0 < \lambda_k < \dots < \lambda_1$ wird die Detektionswahrscheinlichkeit asymptotisch ($\lambda_1 \rightarrow \infty$) maximiert durch

$$\frac{\lambda_{i+1}}{\lambda_i} = \left(\frac{k-i}{k-i+1} \right)^2, \quad i = 1, \dots, k-2,$$

bzw.

$$\frac{\lambda_j}{\lambda_1} = \left(\frac{k-j+1}{k} \right)^2, \quad j = 2, \dots, k-1.$$

Beweis. Die Äquivalenz der beiden Darstellungen sieht man sofort über

$$\frac{\lambda_j}{\lambda_1} = \frac{\lambda_2}{\lambda_1} \cdot \frac{\lambda_3}{\lambda_2} \dots \frac{\lambda_j}{\lambda_{j-1}}.$$

Der Beweis erfolgt über Induktion. Der Induktionsanfang wurde bereits in Lemma 2.10 durchgeführt.

Induktionsschritt: $k \rightarrow k+1$

Zu zeigen ist

$$\frac{\lambda_2}{\lambda_1} = \left(\frac{k+1-2+1}{k+1} \right)^2 = \left(\frac{k}{k+1} \right)^2.$$

Die Verhältnisse der $\lambda_3, \dots, \lambda_{k+1}$ entsprechen gerade der Lösung für k -Normalverteilungen. Insbesondere also

$$c_3 := \frac{\lambda_3}{\lambda_2} = \left(\frac{k-1}{k} \right)^2.$$

Das das optimale $\lambda_2 = c_2 \lambda_1$ wieder in der Mitte der eingrenzenden Schranken liegen muss, erhält man

$$\begin{aligned} \frac{\sqrt{c_2} + \sqrt{c_3 c_2}}{2} \lambda_1 &= c_2 \lambda_1 \\ \frac{k+1}{k} c_2 &= \sqrt{c_2} \\ \implies c_2 \left(1 - \left(\frac{k+1}{k} \right)^2 c_2 \right) &= 0. \end{aligned}$$

Das optimale Verhältnis ist demnach

$$\frac{\lambda_2}{\lambda_1} = \left(\frac{k}{k+1} \right)^2.$$

□

Bemerkung 2.7. Die asymptotisch optimale Parameterwahl weist folgende Besonderheit auf: Die Verhältnisse der Abstände vom Erwartungswert λ_i zu der jeweiligen Grenze betragen bei k Normalverteilungen gerade

$$1 : 2 : 3 : \dots : k,$$

vergleiche Abbildung 15.

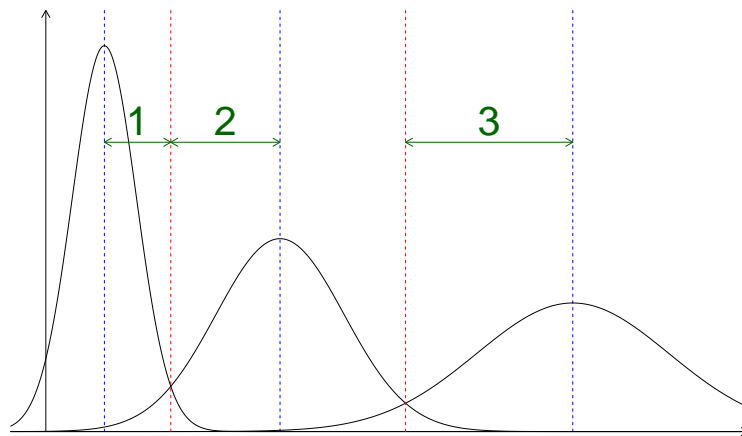


Abbildung 15: Asymptotische Verhältnis der Abstände von Mittelwert zum Schnittpunkt der benachbarten Normalverteilung. Für drei Normalverteilungen stehen die Abstände gerade im Verhältnis 1 : 2 : 3.

Für die Grenze l_i , die zwischen λ_i und λ_{i+1} entscheidet, gilt:

$$\begin{aligned} l_i &= \frac{k-i}{k-i+1} \lambda_i \\ &= \frac{k-i}{k-i+1} \left(\frac{k-i+1}{k} \right)^2 \lambda_1 = \frac{(k-i)(k-i+1)}{k^2} \lambda_1. \end{aligned}$$

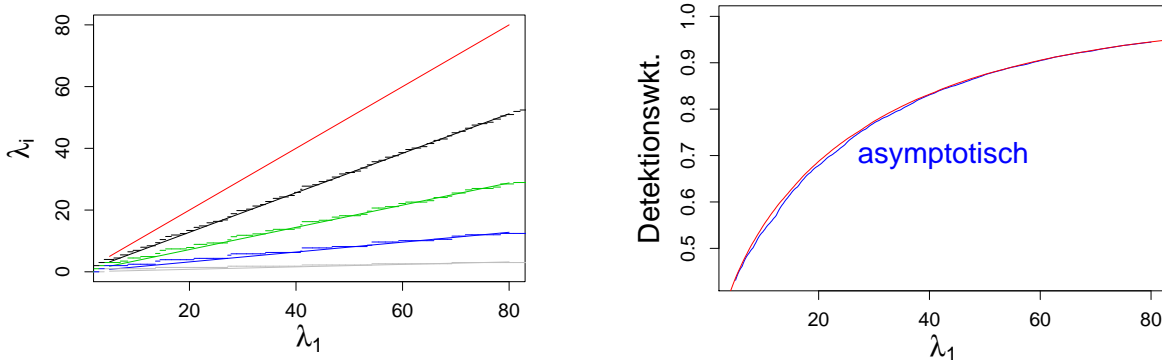
Der Abstand beträgt damit

$$\begin{aligned} l_i - \lambda_{i+1} &= \frac{(k-i)(k-i+1)}{k^2} \lambda_1 - \left(\frac{k-i}{k} \right)^2 \lambda_1 \\ &= \frac{k-i}{k^2} \lambda_1. \end{aligned}$$

Für das Verhältnis zweier benachbarter Abstände ergibt sich schließlich

$$\frac{l_i - \lambda_{i+1}}{l_{i-1} - \lambda_i} = \frac{k - 1}{k - i + 1}.$$

Wie in Abbildung 16 zu sehen ist, liefert die Asymptotik bereits für kleine Werte von λ_1 sehr ähnliche Ergebnisse im Vergleich zu den numerisch bestimmten Werten. Auch die Detektionswahrscheinlichkeit unterscheidet sich nur gering von dem numerischen Optimum.



(a) Asymptotische Lösung im Vergleich zur numerisch exakten Lösung für 5 Stimuli. Die asymptotische Lösung ist jeweils eine lineare Funktion durch den Ursprung.

(b) Vergleich der Detektionswkt. der asymptotischen Lösung und der numerisch exakten Lösung für 5 Stimuli.

Abbildung 16: Vergleich der numerisch exakten Berechnung mit der Asymptotik für 5 Stimuli. Die Asymptotik weicht nur geringfügig von den exakten λ_i ab und auch für die Detektionswahrscheinlichkeit werden bereits für kleines λ_1 annähernd gleich gute Werte erzielt.

2.4 Unterscheidung nur auf Basis von τ

Im vorangegangenen Abschnitt wurde dargelegt, wie die Parameter in der ersten Stufe zu wählen sind, um die Stimuli bestmöglich auseinanderhalten zu können. In diesem Abschnitt wird nur die zweite Stufe betrachtet: Es wird demnach angenommen, dass alle Stimuli dasselbe λ besitzen. O.B.d.A. kann wieder vorausgesetzt werden, dass die Stimuli geordnet sind durch $\tau_1 \geq \tau_2 \geq \dots \geq \tau_k$.

Die Frage ist nun, wie die Parameter τ_j , $j = 1 \dots k$, zu wählen sind, damit sich die Stimuli bestmöglich unterscheiden lassen, die Detektionswahrscheinlichkeit also maximiert wird. Auch hier müssen wieder zusätzliche Bedingungen eingeführt werden, um diese Frage sinnvoll beantworten zu können.

Betrachtet man etwa den Zwei-Stimuli-Fall: Für festes τ_1 würde das τ_2 beliebig klein gewählt werden und somit ließe sich für jedes τ_1 eine beliebig hohe Detektionswahrscheinlichkeit erreichen. Die Einführung eines künstlichen Stimuli mit Parameter $\tau = 0$ würde hier das Problem nicht lösen, da Stimuli 2 auch für einen sehr kleinen Parameter τ_2 fast sicher nicht Null liefert. Somit könnte man auch im Mehr-Stimuli-Fall alle Stimuli mit beliebig hoher Wahrscheinlichkeit auseinanderhalten.

Demnach muss das Verhältnis von τ_1/τ_2 , oder im Allgemeinen von τ_1/τ_k beschränkt werden. Wir werden sehen, dass die Detektionswahrscheinlichkeit im Zwei-Stimuli-Fall auch nur von dem Verhältnis τ_1/τ_2 abhängt.

Ziel des Kapitels ist eine analytische Lösung für die optimalen Phasenparameter zu finden.

Der Fragestellung wird auf folgender Weise nachgegangen: Zuerst wird die erste Stufe als deterministisch $N = 1$ angesetzt und die Optimierungsaufgabe erst für nur zwei, dann für drei Stimuli untersucht. Dann wird das Ergebnis auf mehrere Stimuli und beliebige deterministische erste Stufe $N = n$ übertragen und schließlich auf eine zufällige erste Stufe erweitert.

Für eine deterministische erste Stufe mit $N = 1$ sind k exponential-verteilte Zufallsvariablen

$$X_i \sim \exp(\tau_i), \quad i = 1 \dots, k \quad \text{mit } \tau_k \leq \tau_{k-1} \leq \dots \leq \tau_1,$$

zu unterscheiden. In Satz 2.3 vereinfachen sich die Entscheidungsschranken und es gilt:

Korollar 2.5. *Für eine deterministische erste Stufe $N = 1$ und $\tau_k \leq \tau_{k-1} \leq \dots \leq \tau_1$ lässt sich die Detektionswahrscheinlichkeit bestimmen über*

$$p_D(\tau_1, \dots, \tau_k) = \frac{1}{k} (\mathbb{P}(X_k \leq l_{k-1}) + \mathbb{P}(l_{k-1} < X_{k-1} \leq l_{k-2}) + \dots + \mathbb{P}(X_1 > l_1)),$$

wobei

$$l_j = \frac{\log(\tau_j/\tau_{j+1})}{1/\tau_{j+1} - 1/\tau_j}, \quad j = 1, \dots, k-1.$$

2.4.1 Zwei Stimuli, $N = 1$

In diesem Abschnitt wird die einfachste Ausgangssituation betrachtet: Gegeben zwei exponentialverteilte Zufallsvariablen

$$X_1 \sim \exp(\tau_1) \quad \text{und} \quad X_2 \sim \exp(\tau_2) \quad \text{mit} \quad \tau_2 < \tau_1,$$

entscheidet man sich bei einer Beobachtung x für den Parameter τ_1 , falls gilt:

$$\begin{aligned} & f_{\tau_1}(x) > f_{\tau_2}(x) \\ \Leftrightarrow & \frac{1}{\tau_1} \cdot e^{-x/\tau_1} > \frac{1}{\tau_2} \cdot e^{-x/\tau_2} \\ \Leftrightarrow & x > \frac{\log(\frac{\tau_1}{\tau_2})}{\frac{1}{\tau_2} - \frac{1}{\tau_1}}. \end{aligned}$$

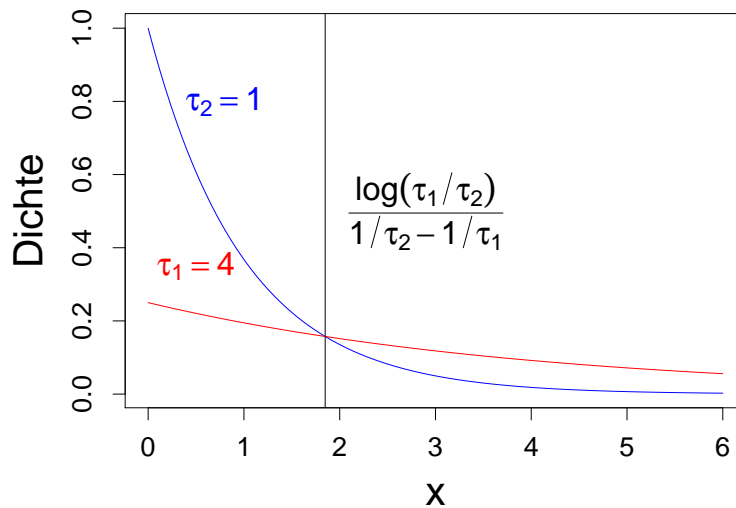


Abbildung 17: Entscheidungsschranke für eine deterministische erste Stufe $N = 1$ und $\tau_1 = 4$ und $\tau_2 = 1$.

In Abbildung 20 ist die Entscheidungsschranke, der Schnittpunkt der beiden Dichten, für den Fall $\tau_1 = 4$ und $\tau_2 = 1$ dargestellt.

Lemma 2.11. Für $N = 1$ und $0 < \tau_2 < \tau_1$ hängt die Detektionswahrscheinlichkeit allein vom Verhältnis

$$a := \frac{\tau_1}{\tau_2}$$

ab.

Beweis. Bezeichne $a := \tau_1/\tau_2 > 0$ das Verhältnis der beiden Parameter, dann lässt sich die Detektionswahrscheinlichkeit p_D schreiben als

$$\begin{aligned} p_D &= \frac{1}{2} \left(\mathbb{P} \left(X_1 > \frac{\log \frac{\tau_1}{\tau_2}}{1/\tau_2 - 1/\tau_1} \right) + \mathbb{P} \left(X_2 < \frac{\log \frac{\tau_1}{\tau_2}}{1/\tau_2 - 1/\tau_1} \right) \right) \\ &= \frac{1}{2} \left(\mathbb{P} \left(\frac{1}{\tau_1} X_1 > \frac{\log \frac{\tau_1}{\tau_2}}{\tau_1/\tau_2 - 1} \right) + \mathbb{P} \left(\frac{1}{\tau_2} X_2 < \frac{\log \frac{\tau_1}{\tau_2}}{1 - \tau_2/\tau_1} \right) \right) \\ &= \frac{1}{2} \left(\mathbb{P} \left(E > \frac{\log(a)}{a - 1} \right) + \mathbb{P} \left(E < \frac{\log(a)}{1 - 1/a} \right) \right), \end{aligned}$$

wobei E für eine standard-exponential-verteilte Zufallsvariable steht.

Wie man sieht, hängt die Detektionswahrscheinlichkeit nur von dem Quotienten τ_1/τ_2 ab. \square

Bemerkung 2.8. Weiterhin erkennt man, dass die untere Schranke $\log(a)/(a-1)$ gerade der Differenzenquotient der Logarithmusfunktion zwischen 1 und a ist und damit streng monoton fallend in Abhängigkeit von a ist. Damit ist der erste Ausdruck streng monoton wachsend für steigendes a . Analog ist die obere Schranke $\log(1/a)/(1/a-1)$ streng monoton wachsend für steigendes a , und damit auch der zweite Ausdruck.

Wie erwartet, wird das maximal mögliche Verhältnis von τ_1/τ_2 gewählt, um die Detektionswahrscheinlichkeit zu maximieren.

2.4.2 Drei Stimuli, $N = 1$

Der Zwei-Stimuli-Fall wird nun durch eine dritte exponential-verteilte Zufallsvariable erweitert. Gegeben sind die Zufallsvariablen

$$X_1 \sim \exp(\tau_1) \quad X_2 \sim \exp(\tau_2) \quad \text{und} \quad X_3 \sim \exp(\tau_3) \quad \text{mit} \quad \tau_3 < \tau_2 < \tau_1,$$

wobei das Verhältnis $a := \tau_1/\tau_3$ fest vorgegeben ist. Wir werden sehen, dass die Detektionswahrscheinlichkeit nur von dem Verhältnis abhängt, nicht aber von der genauen Wahl von τ_1 und τ_3 . Wie ist in Abhängigkeit davon nun das τ_2 zu wählen, damit die Detektionswahrscheinlichkeit maximiert wird?

Lemma 2.12. Für $N = 1$ und $0 < \tau_3 < \tau_2 < \tau_1$ ist die Detektionswahrscheinlichkeit gegeben durch

$$p_D(\tau_1, \tau_2, \tau_3) = \frac{1}{3} \left(\underbrace{1 - e^{-\frac{\log \frac{\tau_2}{\tau_3}}{1 - \tau_3/\tau_2}} - e^{-\frac{\log \frac{\tau_1}{\tau_2}}{1 - \tau_2/\tau_1}}}_{=: h_1(\tau_2)} + \underbrace{e^{-\frac{\log \frac{\tau_2}{\tau_3}}{\tau_2/\tau_3 - 1}} + e^{-\frac{\log \frac{\tau_1}{\tau_2}}{\tau_1/\tau_2 - 1}}}_{=: h_2(\tau_2)} \right).$$

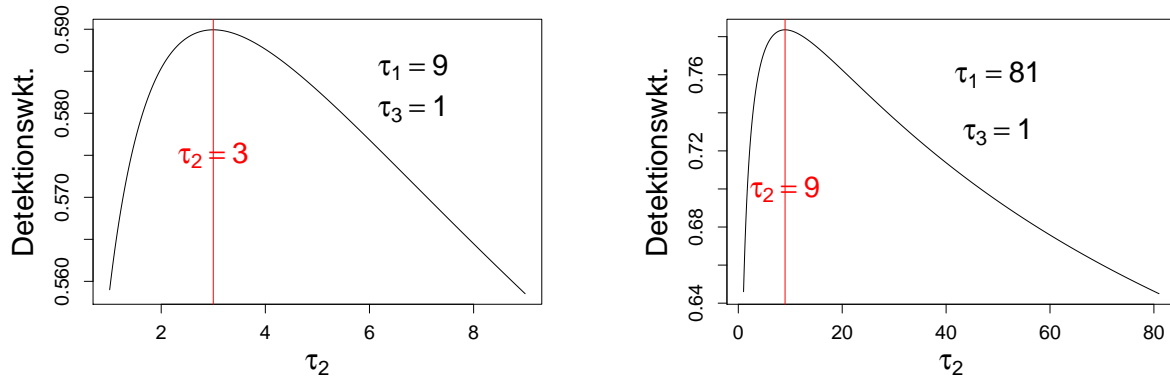
Beweis.

$$\begin{aligned}
 p_D(\tau_1, \tau_2, \tau_3) &= \frac{1}{3} \left(\mathbb{P} \left(X_3 < \frac{\log \frac{\tau_2}{\tau_3}}{1/\tau_3 - 1/\tau_2} \right) + \mathbb{P} \left(\frac{\log \frac{\tau_2}{\tau_3}}{1/\tau_3 - 1/\tau_2} < X_2 < \frac{\log \frac{\tau_1}{\tau_2}}{1/\tau_2 - 1/\tau_1} \right) \right. \\
 &\quad \left. + \mathbb{P} \left(X_1 > \frac{\log \frac{\tau_1}{\tau_2}}{1/\tau_2 - 1/\tau_1} \right) \right) \\
 &= \frac{1}{3} \left(\mathbb{P} \left(E < \frac{\log \frac{\tau_2}{\tau_3}}{1 - \tau_3/\tau_2} \right) + \mathbb{P} \left(\frac{\log \frac{\tau_2}{\tau_3}}{\tau_2/\tau_3 - 1} < E < \frac{\log \frac{\tau_1}{\tau_2}}{1 - \tau_2/\tau_1} \right) + \mathbb{P} \left(E > \frac{\log \frac{\tau_1}{\tau_2}}{\tau_1/\tau_2 - 1} \right) \right) \\
 &= \frac{1}{3} \left(1 - e^{-\frac{\log \frac{\tau_2}{\tau_3}}{1 - \tau_3/\tau_2}} - e^{-\frac{\log \frac{\tau_1}{\tau_2}}{1 - \tau_2/\tau_1}} + e^{-\frac{\log \frac{\tau_2}{\tau_3}}{\tau_2/\tau_3 - 1}} + e^{-\frac{\log \frac{\tau_1}{\tau_2}}{\tau_1/\tau_2 - 1}} \right).
 \end{aligned}$$

□

Dabei sind sich $h_1(\tau_2)$ und $h_2(\tau_2)$ sehr ähnlich, es ist nur τ_1 mit τ_3 vertauscht.

In Abbildung 18 ist die Detektionswahrscheinlichkeit in Abhängigkeit von τ_2 für $\tau_1 = 9$ und $\tau_1 = 81$ und jeweils $\tau_3 = 1$ dargestellt. Beide besitzen ein eindeutiges optimales τ_2 . Da die Exponentialverteilung nicht symmetrisch ist, erwartet man auch nicht, dass τ_2 das arithmetische Mittel von τ_1 und τ_3 ist, aber in Abbildung 18 erkennt man, dass τ_2 gerade das geometrische Mittel zu sein scheint.



- (a) Detektionswkt. in Abhängigkeit von τ_2 für $\tau_1 = 9$ und $\tau_3 = 1$. Das optimale τ_2 ist gerade $\tau_2 = \sqrt{9}$.
- (b) Detektionswkt. in Abhängigkeit von τ_2 für $\tau_1 = 81$ und $\tau_3 = 1$. Das optimale τ_2 ist gerade $\tau_2 = \sqrt{81}$.

Abbildung 18: Detektionswahrscheinlichkeit in Abhängigkeit von τ_2 für festes τ_1 und τ_3 und deterministische erste Stufe $N = 1$.

Wir wollen nun überprüfen, ob

$$\tau_2^{opt} = \sqrt{\tau_1 \cdot \tau_3}$$

ist. Dazu betrachten wir jeweils die Ableitungen von $h_1(\tau_2)$ und von $h_2(\tau_2)$. Für $h_1(\tau_2)$ ergibt sich

$$\begin{aligned} \frac{\partial h_1(\tau_2)}{\partial \tau_2} &= -\frac{\frac{1}{\tau_2}(\frac{\tau_3}{\tau_2} - 1) - \log\left(\frac{\tau_3}{\tau_2}\right) \frac{\tau_3}{\tau_2}}{\left(\frac{\tau_3}{\tau_2} - 1\right)^2} \cdot e^{-\frac{\log \frac{\tau_3}{\tau_2}}{\tau_3/\tau_2 - 1}} - \frac{-\frac{1}{\tau_2}(\frac{\tau_2}{\tau_1} - 1) + \log\left(\frac{\tau_2}{\tau_1}\right) \frac{1}{\tau_1}}{\left(\frac{\tau_2}{\tau_1} - 1\right)^2} \cdot e^{-\frac{\log \frac{\tau_2}{\tau_1}}{\tau_2/\tau_1 - 1}} \\ &\stackrel{!}{=} 0 \quad \text{für } \frac{\tau_2}{\tau_1} = \frac{\tau_3}{\tau_2} \text{ bzw. } \tau_2 = \sqrt{\tau_1 \cdot \tau_3}. \end{aligned}$$

Analog erhält man für $h_2(\tau_2)$

$$\begin{aligned} \frac{\partial h_2(\tau_2)}{\partial \tau_2} &= \frac{-\frac{1}{\tau_2}(\frac{\tau_2}{\tau_3} - 1) + \log\left(\frac{\tau_2}{\tau_3}\right) \frac{1}{\tau_3}}{\left(\frac{\tau_2}{\tau_3} - 1\right)^2} \cdot e^{-\frac{\log \frac{\tau_2}{\tau_3}}{\tau_2/\tau_3 - 1}} + \frac{\frac{1}{\tau_2}(\frac{\tau_1}{\tau_2} - 1) - \log\left(\frac{\tau_1}{\tau_2}\right) \frac{\tau_1}{\tau_2}}{\left(\frac{\tau_1}{\tau_2} - 1\right)^2} \cdot e^{-\frac{\log \frac{\tau_1}{\tau_2}}{\tau_1/\tau_2 - 1}} \\ &\stackrel{!}{=} 0 \quad \text{für } \frac{\tau_2}{\tau_1} = \frac{\tau_3}{\tau_2} \text{ bzw. } \tau_2 = \sqrt{\tau_1 \cdot \tau_3}. \end{aligned}$$

Damit ist $\tau_2 = \sqrt{\tau_1 \cdot \tau_3}$ eine Extremwert verdächtige Stelle. Es ist aber nicht klar, ob es sich für alle τ_1/τ_3 wirklich um ein Maximum handelt. Numerisch lässt sich erahnen, dass für größere τ_1/τ_3 beide Funktionen bei $\tau_2 = \sqrt{\tau_1 \cdot \tau_3}$ ein Maximum besitzen, vgl. (a) und (b) in Abbildung 19, aber für τ_1/τ_3 kleiner etwa 1.5^2 die Funktion $h_2(\tau_2)$ an dieser Stelle ein Minimum besitzt, vergleiche Abbildung 19 (d).

Wissen wollen wir aber eigentlich, ob $\tau_2 = \sqrt{\tau_1 \cdot \tau_3}$ wirklich die bestmögliche Wahl ist, an dieser Stelle also ein globales Maximum vorliegt. Auf jeden Fall müssen wir dazu $h_1(\tau_2)$ und $h_2(\tau_2)$ gemeinsam betrachten.

Lemma 2.13. *Für $N = 1$ und $0 < \tau_3 < \tau_2 < \tau_1$ mit $\tau_1/\tau_3 := a$ fest, wird die Detektionswahrscheinlichkeit durch $\tau_2 = \sqrt{\tau_1 \cdot \tau_3}$ maximiert. Es gilt also (vgl. Lemma 2.12)*

$$h_1(\tau_2) + h_2(\tau_2) \leq h_1(\sqrt{\tau_1 \cdot \tau_3}) + h_2(\sqrt{\tau_1 \cdot \tau_3}) \quad \text{für alle } \tau_3 \leq \tau_2 \leq \tau_1.$$

Beweis. Zu Vereinfachung setze

$$a^2 := \tau_3, \quad b^2 := \tau_2 \quad \text{und} \quad c^2 := \tau_1.$$

$h_1(b^2)$ lässt sich umschreiben zu

$$\begin{aligned} h_1(b^2) &= -e^{-\frac{\log(b^2/a^2)}{1-a^2/b^2}} - e^{-\frac{\log(c^2/b^2)}{1-b^2/c^2}} \\ &= -\left(e^{\log(a^2/b^2)}\right)^{\frac{1}{1-a^2/b^2}} - \left(e^{\log(b^2/c^2)}\right)^{\frac{1}{1-b^2/c^2}} \\ &= -\left(\frac{a^2}{b^2}\right)^{\frac{b^2}{b^2-a^2}} - \left(\frac{b^2}{c^2}\right)^{\frac{c^2}{c^2-b^2}}. \end{aligned}$$

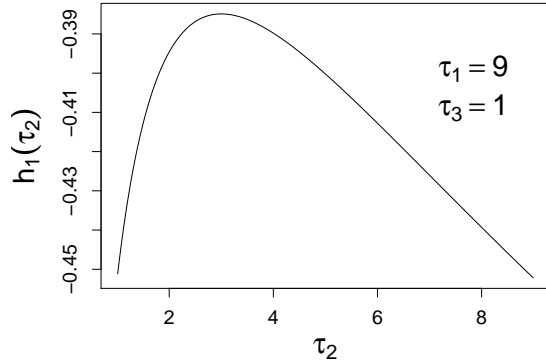
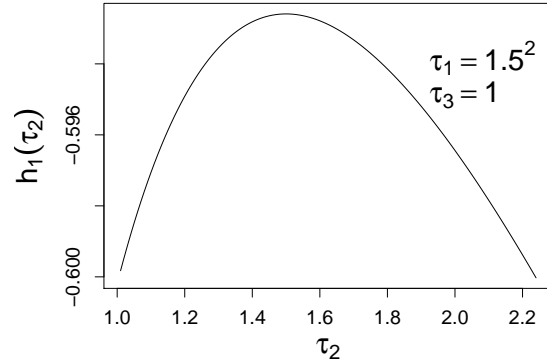
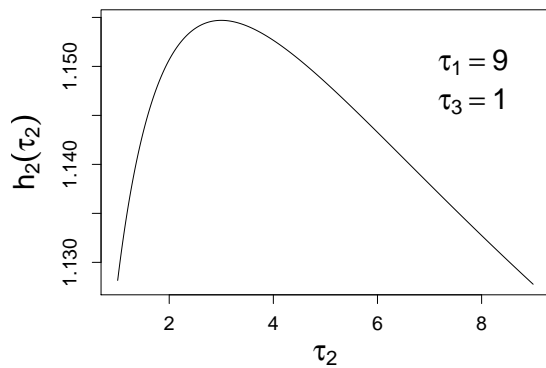
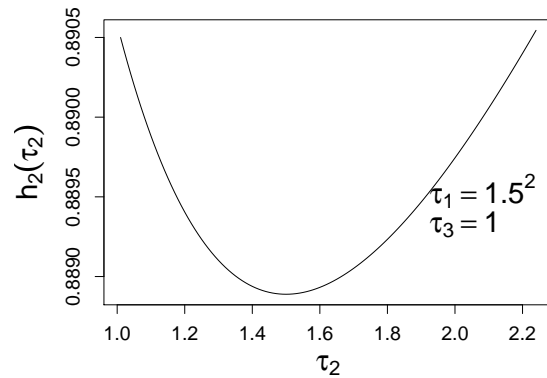
(a) $h_1(\tau_2)$ für $\tau_1 = 9$ und $\tau_3 = 1$.(c) $h_1(\tau_2)$ für $\tau_1 = 1.5^2$ und $\tau_3 = 1$.(b) $h_2(\tau_2)$ für $\tau_1 = 9$ und $\tau_3 = 1$.(d) $h_2(\tau_2)$ für $\tau_1 = 1.5^2$ und $\tau_3 = 1$.

Abbildung 19: $h_1(\tau_2)$ und $h_2(\tau_2)$ für verschiedene τ_1/τ_3 . Nicht für jedes τ_1/τ_3 besitzt $h_1(\tau_2)$ und $h_2(\tau_2)$ bei $\tau_2 = \sqrt{\tau_1/\tau_3}$ ein Maximum. Deshalb müssen beide Funktionen gemeinsam betrachtet werden.

Analog lässt sich $h_2(b^2)$ schreiben als

$$h_2(b^2) = \left(\frac{a^2}{b^2}\right)^{\frac{a^2}{b^2-a^2}} + \left(\frac{b^2}{c^2}\right)^{\frac{b^2}{c^2-b^2}}.$$

Zu zeigen ist die Ungleichung

$$\begin{aligned} & -\left(\frac{a^2}{b^2}\right)^{\frac{b^2}{b^2-a^2}} - \left(\frac{b^2}{c^2}\right)^{\frac{c^2}{c^2-b^2}} + \left(\frac{a^2}{b^2}\right)^{\frac{a^2}{b^2-a^2}} + \left(\frac{b^2}{c^2}\right)^{\frac{b^2}{c^2-b^2}} \leq -2\left(\frac{a}{c}\right)^{\frac{a}{c-a}} + 2\left(\frac{a}{c}\right)^{\frac{c}{c-a}} \\ \Leftrightarrow & \left(\frac{a^2}{b^2}\right)^{\frac{a^2}{b^2-a^2}} - \left(\frac{a^2}{b^2}\right)^{\frac{b^2}{b^2-a^2}} + \left(\frac{b^2}{c^2}\right)^{\frac{c^2}{c^2-b^2}} - \left(\frac{b^2}{c^2}\right)^{\frac{b^2}{c^2-b^2}} \geq 2\left(\frac{a}{c}\right)^{\frac{a}{c-a}} - 2\left(\frac{a}{c}\right)^{\frac{c}{c-a}}. \end{aligned}$$

Substituiert man

$$x := \frac{a^2}{b^2} \quad (0 < x < 1) \quad \text{und} \quad y := \frac{b^2}{c^2} \quad (0 < y < 1),$$

so vereinfacht sich die Ungleichung zu

$$x^{\frac{1}{1-x}} - x^{\frac{x}{1-x}} + y^{\frac{1}{1-y}} - y^{\frac{y}{1-y}} \geq 2\sqrt{xy}^{\frac{1}{1-\sqrt{xy}}} - 2\sqrt{xy}^{\frac{\sqrt{xy}}{1-\sqrt{xy}}}.$$

Setzt man nun

$$\begin{aligned} f(z) &:= z^{\frac{1}{1-z}} - z^{\frac{z}{1-z}} \\ &= z^{\frac{1}{1-z}} \cdot \left(1 - \frac{1}{z}\right), \quad (0 < z < 1), \end{aligned}$$

erhält man die angenehm scheinende Gleichung

$$\frac{1}{2}(f(x) + f(y)) \geq f(\sqrt{xy}) \quad \text{für } x, y \in (0, 1).$$

Für $f = id$ ist dies gerade die bekannte Ungleichung vom arithmetischen und geometrischen Mittel. Gilt diese Ungleichung auch für unsere Funktion $f(z)$? Schreibt man die Ungleichung allgemeiner in der Form

$$(1 - \lambda)f(x) + \lambda f(y) \geq f(x^{1-\lambda}y^\lambda), \quad \text{für } \lambda \in [0, 1] \text{ und } x, y \in I, \quad (3)$$

erkennt man deutlich die Ähnlichkeit zur Konvexität. Für $f : I \rightarrow \mathbb{R}$ zweimal stetig differenzierbar ist Ungleichung (3) äquivalent zu

$$x^2 \frac{\partial^2 f(x)}{\partial x^2} + x \frac{\partial f(x)}{\partial x} \geq 0. \quad (4)$$

Einsehen kann man dies, indem man die Funktion $g(\cdot)$ so wählt, dass $f(x) = g(\log(x))$ ist. Damit gilt:

$$\begin{aligned} &0 \leq x^2 f''(x) + x f'(x) \\ \iff &0 \leq x^2 \left(\frac{1}{x^2} g''(\log(x)) - \frac{1}{x^2} g'(\log(x)) \right) + x \left(\frac{1}{x} g'(\log(x)) \right) \\ \iff &0 \leq g''(\log(x)) \\ \iff &g \left(\log \left(x^{1-\lambda} y^\lambda \right) \right) \leq (1 - \lambda)g(\log(x)) + \lambda g(\log(y)). \end{aligned}$$

Im Folgenden wird gezeigt, dass unsere Funktion $f(\cdot)$ die Ungleichung (4) erfüllt. Dazu werden zuerst die ersten beiden Ableitungen bestimmt:

$$\begin{aligned} \frac{\partial f(x)}{\partial x} &= x^{(1-x)^{-1}} \left(\frac{\log(x)}{(1-x)^2} + \frac{1}{(1-x)x} \right) \left(1 - \frac{1}{x} \right) + \frac{x^{(1-x)^{-1}}}{x^2} \\ &= -\frac{\log(x)}{x(1-x)} \cdot x^{(1-x)^{-1}} \end{aligned}$$

$$\begin{aligned}\frac{\partial^2 f(x)}{\partial x^2} &= -\frac{x^{\frac{x}{1-x}} \log(x)}{(-1+x)^2} + x^{\frac{x}{1-x}} \left(\left(\frac{x}{(-1+x)^3} - 1 \right) \log(x) - 1 \right) \frac{\log(x)}{(1-x)^2} + \frac{x^{\frac{x}{1-x}}}{(-1+x) \cdot x} \\ &= -\frac{x^{(1-x)^{-1}}}{x^2(1-x)^3} \cdot \left(x \log(x) + 2x(1-x) \log(x) + (1-x)^2 \right)\end{aligned}$$

Ungleichung 4 lässt sich zusammenfassen zu

$$\begin{aligned}x^2 \frac{\partial^2 f(x)}{\partial x^2} + x \frac{\partial f(x)}{\partial x} &= \frac{x^{(1-x)^{-1}}}{(1-x)^3} \cdot \left(-\log(x) \left((1-x)^2 + 2x(1-x) \right) - x \log(x)^2 - (1-x)^2 \right) \\ &= \frac{x^{(1-x)^{-1}}}{(1-x)^3} \left(-\left(1-x^2\right) \log(x) - x \log(x)^2 - (1-x)^2 \right) \\ &= \frac{x^{(1-x)^{-1}}}{(1-x)^3} \underbrace{\left(x \log(x) + 1 - x \right)}_{:=g_1(x)} \underbrace{\left(-\log(x) - 1 + x \right)}_{:=g_2(x)} \\ &\stackrel{!}{\geq} 0\end{aligned}$$

Für $0 < x < 1$ ist der vordere Faktor offensichtlich positiv. Für $g_1(x)$ lässt sich dies auch leicht nachprüfen:

$$\begin{aligned}g_1(x) &\geq 0 \\ \iff \log(x) &\geq 1 - \frac{1}{x} \\ \iff 0 &\leq g''(\log(x)) \\ \stackrel{z:=1/x}{\iff} e^{z-1} &\leq z.\end{aligned}$$

Die letzte Ungleichung gilt offensichtlich. Auf dieselbe Ungleichung führt auch $g_2(x)$, womit unsere Funktion $f(x)$ die Ungleichung (3) erfüllt.

Damit ist gezeigt, dass $\tau_2 = \sqrt{\tau_1 \cdot \tau_3}$ die beste Wahl ist. \square

2.4.3 Beliebige Anzahl Stimuli, $N = 1$

In diesem Abschnitt wird eine beliebige Anzahl Stimuli betrachtet, aber weiterhin mit der Beschränkung $N = 1$, sodass nur exponential-verteilte Zufallsvariablen auseinandergelassen werden müssen. Gegeben sind also Zufallsvariablen

$$X_i \sim \exp(\tau_i), \quad i = 1 \dots, k \quad \text{mit } \tau_k \leq \tau_{k-1} \leq \dots \leq \tau_1,$$

wobei das Verhältnis von τ_1/τ_k wieder fest vorgegeben ist. Wie sind in Abhängigkeit von diesem Verhältnis die Parameter $\tau_2, \dots, \tau_{k-1}$ zu wählen, damit die Detektionswahrscheinlichkeit maximiert wird? In Abhängigkeit von τ_1 und τ_k gilt:

Lemma 2.14. Für eine deterministische erste Stufe $N = 1$ und $0 < \tau_k < \tau_{k-1} < \dots < \tau_1$ wird für festes τ_1/τ_k die Detektionswahrscheinlichkeit durch die eindeutige Lösung der Gleichungen

$$\frac{\tau_2}{\tau_1} = \frac{\tau_3}{\tau_2} = \frac{\tau_4}{\tau_3} = \dots = \frac{\tau_k}{\tau_{k-1}},$$

bzw. explizit geschrieben durch

$$\tau_i = \tau_1^{\frac{k-i}{k-1}} \cdot \tau_k^{\frac{i-1}{k-1}}, \quad i = 1, \dots, k,$$

maximiert.

Beweis. Aus Lemma 2.13 ist bereits bekannt, dass im Fall $k = 3$ die Aussage gilt. Der Rest folgt durch ein einfaches induktives Argument. Zur besseren Übersicht wird mit dem Fall $k = 4$ begonnen.

Angenommen, das optimale τ_2 und τ_3 wären bestimmt. Betrachtet man nur die Parameter

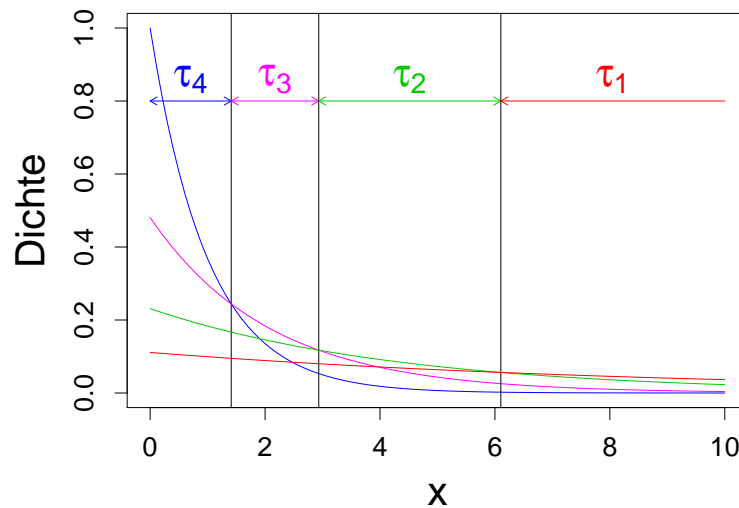


Abbildung 20: Entscheidungsschranken und optimale Phasenwahl für eine deterministische erste Stufe $N = 1$ für $k = 4$ Stimuli.

τ_4 , τ_3 und τ_2 , sieht man, dass τ_3 direkt nur von τ_4 und τ_2 abhängt. Würde das optimale τ_3 nicht die Gleichung $\tau_3/\tau_2 = \tau_4/\tau_3$ lösen, könnte man das τ_3 besser wählen und wir hätten keine optimale Lösung gehabt. Analog folgt, dass τ_2 die Gleichung $\tau_2/\tau_1 = \tau_3/\tau_2$ erfüllen muss. Beide Gleichungen kombiniert, ergibt die Aussage

$$\frac{\tau_2}{\tau_1} = \frac{\tau_3}{\tau_2} = \frac{\tau_4}{\tau_3}.$$

Dasselbe Vorgehen lässt sich auf k Stimuli übertragen. Gegeben sei wieder die optimale Lösung $\tau_2, \dots, \tau_{k-1}$. Betrachtet man nur die Parameter τ_2, \dots, τ_k erhält man aus der Induktionsannahme, dass gilt

$$\frac{\tau_3}{\tau_2} = \frac{\tau_4}{\tau_3} = \dots = \frac{\tau_k}{\tau_{k-1}},$$

andererseits erhält man für $\tau_1, \dots, \tau_{k-1}$, dass

$$\frac{\tau_2}{\tau_1} = \frac{\tau_3}{\tau_2} = \dots = \frac{\tau_{k-1}}{\tau_{k-2}}$$

ist, wodurch insgesamt der erste Teil der Aussage folgt.

Zu zeigen bleibt noch, dass sich τ_i , $i = 1, \dots, k$ darstellen lässt durch

$$\tau_i = \tau_1^{\frac{k-i}{k-1}} \cdot \tau_k^{\frac{i-1}{k-1}}.$$

Dies folgt wieder induktiv: Für $k = 3$ gilt diese Darstellung offensichtlich.

Induktionsschritt: $k - 1 \rightarrow k$

Für $\tau_1, \dots, \tau_{k-1}$ gilt

$$\tau_i = \tau_1^{\frac{k-1-i}{k-2}} \cdot \tau_{k-1}^{\frac{i-1}{k-2}}. \quad (5)$$

Aus dem ersten Teil und (5) folgt

$$\begin{aligned} \frac{\tau_{k-1}}{\tau_{k-2}} &= \frac{\tau_k}{\tau_{k-1}} \\ \stackrel{\text{I.A.}}{\iff} \tau_{k-1}^2 &= \tau_k \cdot \tau_1^{\frac{1}{k-2}} \cdot \tau_{k-1}^{\frac{k-3}{k-2}} \\ \iff \tau_{k-1}^{k-1} &= \tau_k^{k-2} \cdot \tau_1. \end{aligned}$$

Eingesetzt in (5) erhält man

$$\begin{aligned} \tau_i &= \tau_1^{\frac{k-1-i}{k-2}} \cdot \left(\tau_k \cdot \tau_1^{\frac{1}{k-2}} \right)^{\frac{i-1}{k-1}} \\ &= \tau_1^{\frac{k-1-i}{k-2} + \frac{i-1}{(k-2)(k-1)}} \cdot \tau_k^{\frac{i-1}{k-1}} \\ &= \tau_1^{\frac{(k-i)(k-2)}{(k-2)(k-1)}} \cdot \tau_k^{\frac{i-1}{k-1}} \\ &= \tau_1^{\frac{k-i}{k-1}} \cdot \tau_k^{\frac{i-1}{k-1}}. \end{aligned}$$

□

2.4.4 Zwei Stimuli, $N = n$

Nun wird der Ausgang der ersten Stufe etwas allgemeiner betrachtet: Für die Anzahl der Spikes N ist ein beliebiges, aber deterministisches $n \in \mathbb{N}$ möglich. Jetzt sind nicht mehr nur exponential-verteilte Zufallsvariablen, sondern zwei Zufallsvektoren

$$\begin{aligned} X_1 &= (X_{11}, X_{12}, \dots, X_{1n}) & \text{und} & & X_2 &= (X_{21}, X_{22}, \dots, X_{2n}) \\ X_{1i} &\sim \exp(\tau_1) & \text{und} & & X_{2i} &\sim \exp(\tau_2), \end{aligned}$$

mit $\tau_2 < \tau_1$, gegeben. Bei einer Beobachtung $\mathbf{x} = (x_1, \dots, x_n)$ entscheidet man sich für den Parameter τ_1 , falls gilt:

$$\begin{aligned} & f_{\tau_1}(\mathbf{x}) > f_{\tau_2}(\mathbf{x}) \\ \Leftrightarrow & \prod_{i=1}^n \frac{1}{\tau_1} \cdot e^{-x_i/\tau_1} > \prod_{i=1}^n \frac{1}{\tau_2} \cdot e^{-x_i/\tau_2} \\ \Leftrightarrow & e^{\left(\frac{1}{\tau_2} - \frac{1}{\tau_1}\right) \sum_{i=1}^n x_i} > \frac{\tau_1^n}{\tau_2^n} \\ \Leftrightarrow & \sum_{i=1}^n x_i > \frac{n \cdot \log\left(\frac{\tau_1}{\tau_2}\right)}{\frac{1}{\tau_2} - \frac{1}{\tau_1}}. \end{aligned}$$

Lemma 2.15. Die Summe von n unabhängigen exponential- (τ) -verteilten Zufallsvariablen entspricht einer Gammaverteilung mit Parameter n und $1/\tau$.

Beweis. Wie man durch Faltung leicht berechnet, besitzt die Summe von n unabhängigen exponential- (τ) -verteilten Zufallsvariablen die Dichte

$$f_n(x) = \begin{cases} \frac{x^{n-1}}{\tau^n \cdot (n-1)!} e^{-x/\tau} & \text{für } x \geq 0, \\ 0 & \text{für } x < 0, \end{cases}$$

und die Verteilungsfunktion

$$F_n(x) = \begin{cases} 1 - e^{-x/\tau} \sum_{i=0}^{n-1} \frac{(x/\tau)^i}{i!} & \text{für } x \geq 0, \\ 0 & \text{für } x < 0. \end{cases}$$

Dies entspricht einer Gammaverteilung mit Parameter n und $1/\tau$. Im speziellen ($n \in \mathbb{N}$) wird diese als Erlangverteilung bezeichnet. \square

Lemma 2.16. Für $N = n$ und $0 < \tau_2 < \tau_1$ hängt die Detektionswahrscheinlichkeit allein vom Verhältnis

$$\frac{\tau_1}{\tau_2}$$

ab.

Beweis. Analog zum Fall $N = 1$ lässt sich die Detektionswahrscheinlichkeit p_D schreiben als

$$\begin{aligned} p_D &= \frac{1}{2} \left(\mathbb{P} \left(\sum_{i=1}^n X_{1i} > \frac{n \cdot \log \frac{\tau_1}{\tau_2}}{1/\tau_2 - 1/\tau_1} \right) + \mathbb{P} \left(\sum_{i=1}^n X_{2i} < \frac{n \cdot \log \frac{\tau_1}{\tau_2}}{1/\tau_2 - 1/\tau_1} \right) \right) \\ &= \frac{1}{2} \left(\mathbb{P} \left(\sum_{i=1}^n \frac{1}{\tau_1} X_{1i} > \frac{n \cdot \log \frac{\tau_1}{\tau_2}}{\tau_1/\tau_2 - 1} \right) + \mathbb{P} \left(\sum_{i=1}^n \frac{1}{\tau_2} X_{2i} < \frac{n \cdot \log \frac{\tau_1}{\tau_2}}{1 - \tau_2/\tau_1} \right) \right) \\ &= \frac{1}{2} \left(\mathbb{P} \left(\sum_{i=1}^n E_i > \frac{n \cdot \log \frac{\tau_1}{\tau_2}}{\tau_1/\tau_2 - 1} \right) + \mathbb{P} \left(\sum_{i=1}^n E_i < \frac{n \cdot \log \frac{\tau_1}{\tau_2}}{1 - \tau_2/\tau_1} \right) \right), \end{aligned}$$

wobei E_i , $i = 1, \dots, n$, für unabhängige standard-exponential-verteilte Zufallsvariablen stehen. □

Bemerkung 2.9. *Wieder erkennt man mit demselben Argument aus Lemma 2.11, dass beide Wahrscheinlichkeiten für jedes n streng monoton steigen in Abhängigkeit von τ_1/τ_2 .*

2.4.5 Drei Stimuli, $N = n$

In diesem Abschnitt wird wieder ein weiterer Stimulus eingeführt. Gegeben die drei Zufallsvektoren

$$\begin{aligned} X_1 &= (X_{11}, X_{12}, \dots, X_{1n}) \quad \text{und} \quad X_2 = (X_{21}, X_{22}, \dots, X_{2n}) \quad \text{und} \quad X_3 = (X_{31}, X_{32}, \dots, X_{3n}) \\ X_{1i} &\sim \exp(\tau_1) \quad \text{und} \quad X_{2i} \sim \exp(\tau_2) \quad \text{und} \quad X_{3i} \sim \exp(\tau_3), \end{aligned}$$

mit $\tau_3 < \tau_2 < \tau_1$ ist wieder das $\tilde{\tau}_2$ gesucht, sodass für τ_1, τ_3 fest gilt:

$$p_D(\tau_1, \tilde{\tau}_2, \tau_3) \geq p_D(\tau_1, \tau_2, \tau_3) \quad \forall \tau_2 \text{ mit } \tau_3 < \tau_2 < \tau_1,$$

die Detektionswahrscheinlichkeit also maximiert wird.

Korollar 2.6. *Für $N = n$ und $0 < \tau_3 < \tau_2 < \tau_1$ ist die Detektionswahrscheinlichkeit gegeben durch*

$$p_D(\tau_1, \tau_2, \tau_3) = \frac{1}{3} \left(\mathbb{P} \left(\gamma_n < \frac{n \cdot \log \frac{\tau_2}{\tau_3}}{1 - \tau_3/\tau_2} \right) + \mathbb{P} \left(\frac{n \cdot \log \frac{\tau_2}{\tau_3}}{\tau_2/\tau_3 - 1} < \gamma_n < \frac{n \cdot \log \frac{\tau_1}{\tau_2}}{1 - \tau_2/\tau_1} \right) + \mathbb{P} \left(\gamma_n > \frac{n \cdot \log \frac{\tau_1}{\tau_2}}{\tau_1/\tau_2 - 1} \right) \right),$$

für eine gamma-verteilte Zufallsvariable γ_n mit Parameter n und 1.

Beweis. Nach Satz 2.3 wird die Detektionswahrscheinlichkeit bestimmt über

$$\begin{aligned} p_D(\tau_1, \tau_2, \tau_3) &= \frac{1}{3} \left(\mathbb{P} \left(\sum_{i=1}^n X_{3i} < \frac{n \cdot \log \frac{\tau_2}{\tau_3}}{1/\tau_3 - 1/\tau_2} \right) + \mathbb{P} \left(\frac{n \cdot \log \frac{\tau_2}{\tau_3}}{1/\tau_3 - 1/\tau_2} < \sum_{i=1}^n X_{2i} < \frac{n \cdot \log \frac{\tau_1}{\tau_2}}{1/\tau_2 - 1/\tau_1} \right) \right. \\ &\quad \left. + \mathbb{P} \left(\sum_{i=1}^n X_{1i} > \frac{n \cdot \log \frac{\tau_1}{\tau_2}}{1/\tau_2 - 1/\tau_1} \right) \right). \end{aligned}$$

Da $X_{ji} \sim \exp(\tau_j)$, $j = 1, 2, 3$, $i = 1, \dots, n$ unabhängig, ist nach Lemma 2.15 die Summe gamma-verteilt mit Parameter n und 1. □

Auch hier ergibt sich wieder:

Lemma 2.17. Für $N = n$ und $0 < \tau_3 < \tau_2 < \tau_1$ mit $\tau_1/\tau_3 := a$ fest, wird die Detektionswahrscheinlichkeit durch $\tau_2 = \sqrt{\tau_1 \cdot \tau_3}$ maximiert, bzw. durch die eindeutige Lösung der Gleichung

$$\frac{\tau_2}{\tau_1} = \frac{\tau_3}{\tau_2}.$$

Beweis. Der Beweis verläuft analog zu Lemma 2.13, nun mit etwas unhandlicheren Termen. Zur besseren Übersicht, vergleiche Beweisschritte in Lemma 2.13.

Bezeichne γ_n eine gamma-verteilte Zufallsvariable mit Parameter n und 1. Nach Korollar 2.6 gilt es

$$p_D(\tau_1, \tau_2, \tau_3) = \frac{1}{3} \left(\mathbb{P} \left(\gamma < \frac{n \cdot \log \frac{\tau_2}{\tau_3}}{1 - \tau_3/\tau_2} \right) + \mathbb{P} \left(\frac{n \cdot \log \frac{\tau_2}{\tau_3}}{\tau_2/\tau_3 - 1} < \gamma < \frac{n \cdot \log \frac{\tau_1}{\tau_2}}{1 - \tau_2/\tau_1} \right) + \mathbb{P} \left(\gamma > \frac{n \cdot \log \frac{\tau_1}{\tau_2}}{\tau_1/\tau_2 - 1} \right) \right)$$

nach τ_2 zu maximieren. Sei

$$g(x, y, n) := \frac{n \cdot \log \frac{y}{x}}{\frac{y}{x} - 1},$$

und

$$F(x, y, n) := 1 - e^{-g(x, y, n)} \cdot \sum_{i=0}^{n-1} \frac{g(x, y, n)^i}{i!},$$

dann lässt sich $p_D(\tau_1, \tau_2, n)$ schreiben als

$$p_D(\tau_1, \tau_2, \tau_3) = \frac{1}{3} \left(\underbrace{F(\tau_2, \tau_3, n) + F(\tau_1, \tau_2, n)}_{:= h_1(\tau_2)} - 1 + 2 \underbrace{-F(\tau_3, \tau_2, n) - F(\tau_2, \tau_1, n)}_{:= h_2(\tau_2)} \right)$$

Mit

$$\begin{aligned} \frac{\partial F(x, y, n)}{\partial x} &= \frac{\partial g(x, y, n)}{\partial x} \cdot e^{-g(x, y, n)} \cdot \sum_{i=0}^{n-1} \frac{g(x, y, n)^i}{i!} - e^{-g(x, y, n)} \cdot \sum_{i=1}^{n-1} \frac{g(x, y, n)^{i-1}}{(i-1)!} \\ &= \frac{\partial g(x, y, n)}{\partial x} \cdot \frac{g(x, y, n)^{n-1}}{(n-1)!} \cdot e^{-g(x, y, n)} \end{aligned}$$

lässt sich für die Ableitung von $h_1(\tau_2)$ einsehen, dass

$$\begin{aligned} \frac{\partial h_1(\tau_2)}{\partial \tau_2} &= - \frac{\frac{1}{\tau_2} \left(\frac{\tau_3}{\tau_2} - 1 \right) - \log \left(\frac{\tau_3}{\tau_2} \right) \frac{\tau_3}{\tau_2}}{\left(\frac{\tau_3}{\tau_2} - 1 \right)^2} \cdot \frac{g(\tau_2, \tau_3, n)}{(n-1)!} \cdot e^{-g(\tau_2, \tau_3, n)} \\ &\quad - \frac{-\frac{1}{\tau_2} \left(\frac{\tau_2}{\tau_1} - 1 \right) + \log \left(\frac{\tau_2}{\tau_1} \right) \frac{1}{\tau_1}}{\left(\frac{\tau_2}{\tau_1} - 1 \right)^2} \cdot \frac{g(\tau_1, \tau_2, n)}{(n-1)!} \cdot e^{-g(\tau_1, \tau_2, n)} \\ &\stackrel{!}{=} 0 \quad \text{für} \quad \frac{\tau_2}{\tau_1} = \frac{\tau_3}{\tau_2} \quad \text{bzw.} \quad \tau_2 = \sqrt{\tau_1 \cdot \tau_3}. \end{aligned}$$

Analog erhält man für $h_2(\tau_2)$

$$\begin{aligned} \frac{\partial h_2(\tau_2)}{\partial \tau_2} &= \frac{-\frac{1}{\tau_2} \left(\frac{\tau_2}{\tau_3} - 1 \right) + \log \left(\frac{\tau_2}{\tau_3} \right) \frac{1}{\tau_3}}{\left(\frac{\tau_2}{\tau_3} - 1 \right)^2} \cdot \frac{g(\tau_3, \tau_2, n)}{(n-1)!} \cdot e^{-g(\tau_3, \tau_2, n)} \\ &+ \frac{\frac{1}{\tau_2} \left(\frac{\tau_1}{\tau_2} - 1 \right) - \log \left(\frac{\tau_1}{\tau_2} \right) \frac{\tau_1}{\tau_2}}{\left(\frac{\tau_1}{\tau_2} - 1 \right)^2} \cdot \frac{g(\tau_2, \tau_1, n)}{(n-1)!} \cdot e^{-g(\tau_2, \tau_1, n)} \\ &\stackrel{!}{=} 0 \quad \text{für } \frac{\tau_2}{\tau_1} = \frac{\tau_3}{\tau_2} \text{ bzw. } \tau_2 = \sqrt{\tau_1 \cdot \tau_3}. \end{aligned}$$

Auch hier muss noch überprüft werden, ob ein Maximum vorliegt und ob dies auch ein globales ist.

Analog zum Fall $N = 1$ in Lemma 2.13 muss also überprüft werden, ob für alle $n \in \mathbb{N}$ und $\tau_2 \in [\tau_3, \tau_2]$ gilt

$$\begin{aligned} &\mathbb{P}(\gamma < g(\tau_2, \tau_3, n)) + \mathbb{P}(g(\tau_3, \tau_2, n) < \gamma < g(\tau_1, \tau_2, n)) + \mathbb{P}(\gamma > g(\tau_2, \tau_1, n)) \\ &\leq 2 \cdot \mathbb{P}(\gamma < g(\sqrt{\tau_1 \cdot \tau_3}, \tau_3, n)) + 2 \cdot \mathbb{P}(\gamma > g(\tau_1, \sqrt{\tau_1 \cdot \tau_3}, n)) \end{aligned}$$

Durch dieselben Substitutionen wie in Lemma 2.13 lässt sich diese Ungleichung in die Fragestellung überführen, ob die Funktion

$$f(x) := x^{\frac{n}{1-x}} \left(\sum_{i=0}^{n-1} \frac{1}{i!} \left(\frac{n \cdot \log x}{x-1} \right)^i - x^{-n} \sum_{i=0}^{n-1} \frac{1}{i!} \left(\frac{n \cdot x \cdot \log x}{x-1} \right)^i \right)$$

für $0 < x < 1$ Ungleichung (3) erfüllt.

Für die erste Ableitung erhält man

$$\begin{aligned} \frac{\partial f(x)}{\partial x} &= \frac{n(\log(x)x + 1 - x)}{(n-1)!x(-1+x)^2} \cdot x^{\frac{n}{1-x}} \left(\frac{n \log(x)}{-1+x} \right)^{n-1} - \frac{n(\log(x) + 1 - x)}{(n-1)!(-1+x)^2} \cdot x^{\frac{nx}{1-x}} \left(\frac{nx \log(x)}{-1+x} \right)^{n-1} \\ &= \frac{n^n}{(n-1)!} \frac{\log(x)^n}{x(x-1)^n} \cdot x^{\frac{n}{1-x}} \end{aligned}$$

und für die zweite Ableitung

$$\begin{aligned} \frac{\partial^2 f(x)}{\partial x^2} &= \\ &- \frac{n^{n+1}}{(n-1)!} \cdot \frac{\log(x)^{n-1}}{x^2(x-1)^{n+2}} \cdot x^{\frac{n}{1-x}} \left(-n \cdot x \log(x)^2 + \left((1+n)x^2 - 2x - n + 1 \right) \log x - n(1-x)^2 \right). \end{aligned}$$

Damit ergibt sich

$$\begin{aligned} x^2 \frac{\partial^2 f(x)}{\partial x^2} + x \frac{\partial f(x)}{\partial x} &= - \frac{n^{n+1}}{(n-1)!} \cdot \underbrace{\frac{\log(x)^{n-1}}{(-1+x)^{n+2}}}_{(*)} \cdot x^{\frac{n}{1-x}} (x \log x + 1 - x) (-\log x - 1 + x) \\ &\stackrel{!}{\geq} 0 \end{aligned}$$

Die beiden hinteren Faktoren sind bereits bekannt und positiv. Der vordere Teil ist für alle $n \in \mathbb{N}$ positiv, da in Abhängigkeit von n gerade oder ungerade in (*) genau ein Term negativ ist.

Damit ist auch für den allgemeinen Fall $N = n$ gezeigt, dass $\tau_2 = \sqrt{\tau_1 \cdot \tau_3}$ die beste Wahl ist. \square

2.4.6 Beliebige Anzahl Stimuli und zufällige erste Stufe

Analog zum Fall $N = 1$ lässt sich für $N = n$ (vgl. Lemma 2.14) der induktive Beweis führen, dass für eine beliebige Anzahl k an Stimuli die optimalen Parameter durch die Gleichung

$$\frac{\tau_2}{\tau_1} = \frac{\tau_3}{\tau_2} = \frac{\tau_4}{\tau_3} = \dots = \frac{\tau_k}{\tau_{k-1}}$$

gegeben sind. Bisher wurde die erste Stufe nur als deterministisch $N = n$ betrachtet. Da sich für jedes $n \in \mathbb{N}$ dieselbe optimale Parameterwahl ergeben hat, diese also unabhängig von n ist, gilt das Ergebnis auch für eine zufällige erste Stufe.

Somit halten wir noch einmal fest:

Satz 2.7. Für $k \in \mathbb{N}$ Stimuli mit $\lambda_1 = \lambda_2 = \dots = \lambda_k$ und $\tau_k \leq \tau_{k-1} \leq \dots \leq \tau_1$ mit τ_1/τ_k fest wird die Detektionswahrscheinlichkeit durch die eindeutige Lösung der Gleichungen

$$\frac{\tau_2}{\tau_1} = \frac{\tau_3}{\tau_2} = \frac{\tau_4}{\tau_3} = \dots = \frac{\tau_k}{\tau_{k-1}},$$

bzw. explizit geschrieben durch

$$\tau_i = \tau_1^{\frac{k-i}{k-1}} \cdot \tau_k^{\frac{i-1}{k-1}}, \quad i = 1, \dots, k,$$

maximiert.

Beweis. Der erste Teil folgt aus Lemma 2.17. Die zweite Teilaussage wurde bereits in Lemma 2.14 gezeigt. \square

3 Optimale Parameterwahl auf Basis von λ und τ

Für jede Stufe einzeln betrachtet haben wir die optimale Parameterwahl bereits analysiert. Auf Basis von λ lassen sich durch ein geschicktes Vorgehen die optimalen Parameter für eine beliebige Anzahl Stimuli sehr schnell numerisch berechnen. Asymptotisch existiert sogar eine analytische Lösung. Auf Basis von τ war es möglich eine explizite, analytische Darstellung der optimalen Parameter für eine beliebige Anzahl Stimuli zu beweisen.

Jetzt soll der Frage nachgegangen werden, wie die Parameter λ_i und τ_i , $i = 1, \dots, k$, zu wählen sind, damit die Stimuli mit möglichst hoher Wahrscheinlichkeit auseinandergehalten werden können. Auch hier bedarf es analog zur Unterscheidung auf nur einer Stufe einer zusätzlichen Beschränkung: Die Frage lässt sich nur sinnvoll beantworten, wenn man das maximale λ beschränkt und das Verhältnis τ_{\max}/τ_{\min} vorgibt.

O.B.d.A. kann davon ausgegangen werden, dass die λ 's der Größe nach geordnet sind: $\lambda_k \leq \lambda_{k-1} \leq \dots \leq \lambda_1$. Da bereits nur auf Basis von λ keine analytische Lösung möglich war (abgesehen von der Asymptotik), ist beim Zusammenfügen der beiden Stufen auch nur ein numerisches Vorgehen zu erwarten. Ziel ist es aber, das Zusammenspiel der beiden Stufen besser zu verstehen und exemplarisch für bestimmte Vorgaben die optimalen Parameter numerisch zu bestimmen. Insbesondere wird dabei der Einfluss der Phase herausgearbeitet und das deutlich bessere Detektieren durch Hinzunahme der Phase aufgezeigt. Begonnen wird wieder mit dem Zwei-Stimuli-Fall.

3.1 Nullstimulus

In diesem Abschnitt wird auch beim Zusammenspiel von λ und τ die Existenz des Nullstimulus motiviert, der unabhängig von den Randbedingungen stets als natürlich optimaler Stimulus vorhanden ist.

Gegeben zwei Stimuli S_1, S_2 mit $0 \leq \lambda_2 \leq \lambda_1 \leq \lambda_{\max}$ und $\tau_{\max}/\tau_{\min} \leq c$, ist das optimale $\tilde{\lambda}_2$ gesucht, sodass

$$p_D(\lambda_1, \tilde{\lambda}_2, \tau_1, \tau_2) \geq p_D(\lambda_1, \lambda_2, \tau_1, \tau_2) \quad \forall \lambda_2 \leq \lambda_1 \text{ und } \forall \tau_1, \tau_2 \text{ mit } \tau_{\max}/\tau_{\min} \leq c.$$

Gemäß Lemma 2.1 entscheidet man sich für Stimulus 1, falls gilt:

$$n \log \frac{\lambda_1 \tau_2}{\lambda_2 \tau_1} + \left(\frac{1}{\tau_2} - \frac{1}{\tau_1} \right) \sum_{i=1}^n x_i > \lambda_1 - \lambda_2.$$

Diese Entscheidungsregel ist in Abbildung 21 für simulierte Werte dargestellt. Um die Werte besser auseinanderhalten zu können, sind diese leicht gejittert.

Da λ_2 relativ klein ist, ergeben sich häufig Null Spikes. Für großes λ_1 erwartet man also wieder, dass das beste $\lambda_2 = 0$ ist, aber das ist für kleine λ_1 erstmal überhaupt nicht klar. Setzt man $\lambda_2 = 0$ verliert man komplett die zweite Stufe, egal wie viel Entscheidungsmacht man durch diese erhalten würde.

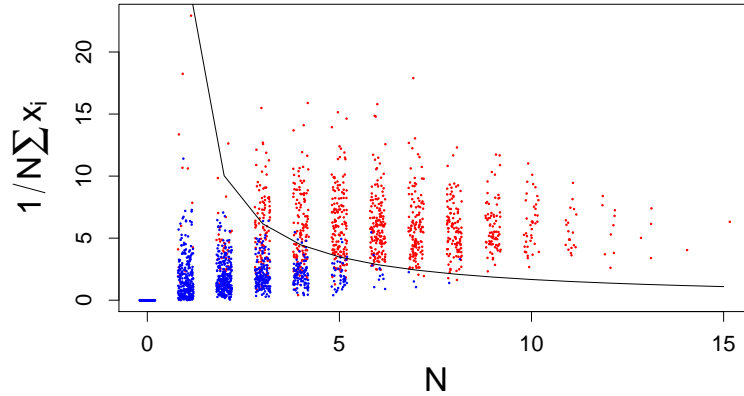


Abbildung 21: Entscheidungsranke für 2 Stimuli mit Parametern $\lambda_1 = 6$, $\tau_1 = 4$, $\lambda_2 = 2$ und $\tau_2 = 2$. Die roten Punkte entsprechen simulierten Werten gemäß λ_1 und τ_1 , die blauen Punkte gemäß λ_2 und τ_2 .

In Abbildung 22 ist die Detektionswahrscheinlichkeit in Abhängigkeit von λ_2 für festes λ_1 , τ_1 und τ_2 berechnet.

Auf den ersten Blick ergibt sich ein relativ erstaunliches Ergebnis: Auch für sehr kleines λ_1 und sehr großes τ_{\max}/τ_{\min} ist die beste Wahl immer $\lambda_2 = 0$. Auf den zweiten Blick ist dies aber gar nicht so erstaunlich: Ist das λ_1 klein, ergibt sich oft Null, und die Stimuli ließen sich bei $\lambda_2 = \lambda_1$ nicht unterscheiden. Wählt man $\lambda_2 = 0$ lassen sich diese unabhängig von der zweiten Stufe bei mindestens einem Spike stets sicher unterscheiden.

Im Allgemeinen gilt:

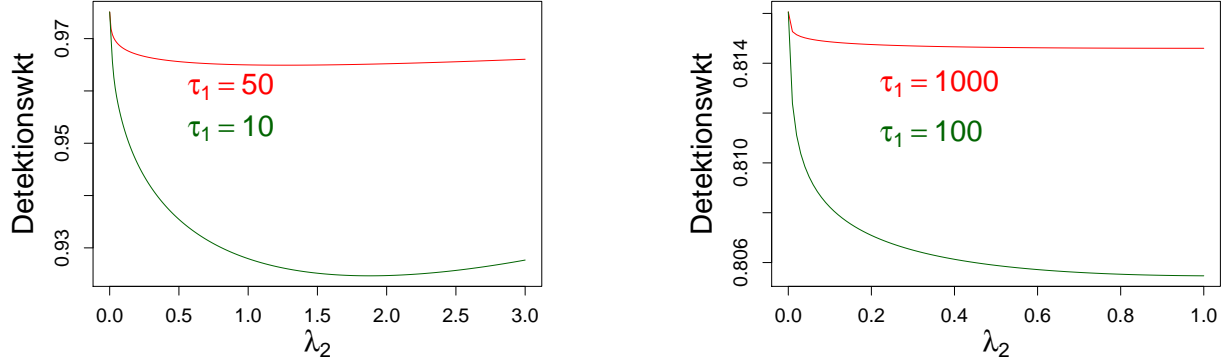
Satz 3.1. Für $k = 2$ Stimuli mit $\lambda_2 \leq \lambda_1 \leq c$, $c \in \mathbb{R}^+$ fest, und für beliebige τ_1, τ_2 wird die Detektionswahrscheinlichkeit stets durch

$$\lambda_2 = 0$$

maximiert.

Beweis. Es ist zu zeigen, dass für die Detektionswahrscheinlichkeit p_D gilt:

$$p_D(\lambda_1, 0, \tau_1, \tau_2) \geq p_D(\lambda_1, \lambda_2, \tau_1, \tau_2) \quad \forall \lambda_1, \tau_1, \tau_2 \in \mathbb{R} \text{ und } \lambda_2 \in [0, \lambda_1].$$



(a) Detektionswkt. in Abhängigkeit von λ_2 für $\lambda_1 = 3$, $\tau_2 = 1$ und $\tau_1 = 10$ bzw. $\tau_1 = 50$.

(b) Detektionswkt. in Abhängigkeit von λ_2 für $\lambda_1 = 1$, $\tau_2 = 1$ und $\tau_1 = 100$ bzw. $\tau_1 = 1000$.

Abbildung 22: Die Detektionswahrscheinlichkeit in Abhängigkeit von $\lambda_2 \leq \lambda_1$ für verschiedene Vorgaben λ_1 , τ_1 und τ_2 . Der optimale Wert ist stets $\lambda_2 = 0$.

Unabhängig von τ_1 und τ_2 ergibt sich für die linke Seite

$$\begin{aligned} p_D(\lambda_1, 0, \tau_1, \tau_2) &= \frac{1}{2} (\mathbb{P}(N_2 = 0) + \mathbb{P}(N_1 > 0)) \\ &= \frac{1}{2} (1 + 1 - e^{-\lambda_1}) \\ &= 1 - \frac{1}{2} e^{-\lambda_1}. \end{aligned}$$

Die rechte Seite ist monoton steigend in τ_{max}/τ_{min} und für eine beliebig mächtige zweite Stufe erhält man

$$\begin{aligned} p_D(\lambda_1, \lambda_2, \tau_1, \tau_2) &\leq \frac{1}{2} (\mathbb{P}(N_2 = 0) + \mathbb{P}(N_2 > 0) \cdot 1 + \mathbb{P}(N_1 > 0) \cdot 1) \\ &= \frac{1}{2} (e^{-\lambda_2} + (1 - e^{-\lambda_2}) \cdot 1 + (1 - e^{-\lambda_1}) \cdot 1) \\ &= 1 - \frac{1}{2} e^{-\lambda_1}. \end{aligned}$$

□

Dass das optimale $\lambda_2 = 0$ ist, lässt sich auch etwas intuitiver erkennen. Wir haben bereits gesehen, dass $\lambda_2 = 0$ die Detektionswahrscheinlichkeit in der ersten Stufe maximiert, auf eine gewisse Weise wird auch die zweite Stufe dadurch maximiert, da dort auch keine Fehlentscheidung mehr möglich ist.

Bemerkung 3.1. *Damit existiert unabhängig von λ_{max} und τ_{max}/τ_{min} stets der Nullstimulus, für den man sich im Fall von Null Spikes immer entscheidet. Dieser wird auch hier im Folgenden nicht mehr zu der Anzahl Stimuli dazu gezählt.*

3.2 Zwei Stimuli

Gegeben zwei Stimuli S_1, S_2 mit $0 < \lambda_2 \leq \lambda_1 \leq \lambda_{\max}$ und $\tau_{\max}/\tau_{\min} \leq c$, ist das optimale $\tilde{\lambda}_2$ gesucht, sodass

$$p_D(0, \lambda_1, \tilde{\lambda}_2, \tau_1, \tau_2) \geq p_D(0, \lambda_1, \lambda_2, \tau_1, \tau_2) \quad \forall \lambda_2 \leq \lambda_1 \text{ und } \forall \tau_1, \tau_2 \text{ mit } \tau_{\max}/\tau_{\min} \leq c.$$

Bestimmt man in Abhängigkeit von τ_{\max}/τ_{\min} die Detektionswahrscheinlichkeit für festes λ_1 und verschiedene λ_2 , lässt sich der grundlegende Verlauf des optimalen λ_2 bereits erkennen, vergleiche Abbildung 23.

Für $\tau_{\max}/\tau_{\min} = 1$ dominiert $\lambda_2 = 2$ unter den ausgewählten λ_2 . Das für diesen Fall optima-

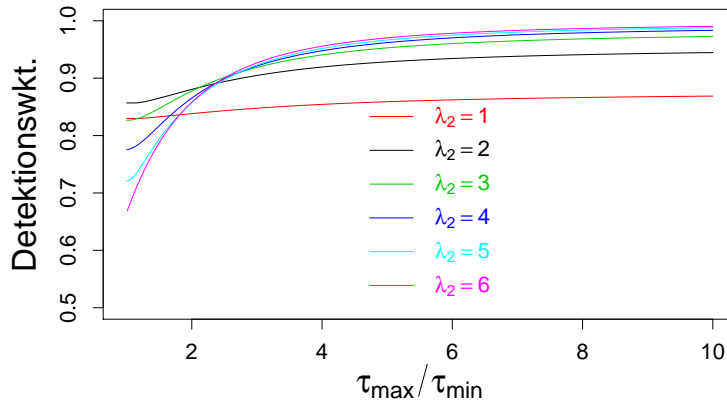


Abbildung 23: Detektionswahrscheinlichkeit in Abhängigkeit von τ_{\max}/τ_{\min} für $\lambda_1 = 6$ und ausgewählte λ_2 .

le λ_2 ist aber bereits bekannt, da dies der Unterscheidung nur auf Basis von λ entspricht. Bei diesem λ_2 beginnend, wächst es für steigendes τ_{\max}/τ_{\min} an, bis $\lambda_2 = \lambda_1$ erreicht wird und nur noch auf Basis von τ unterschieden wird.

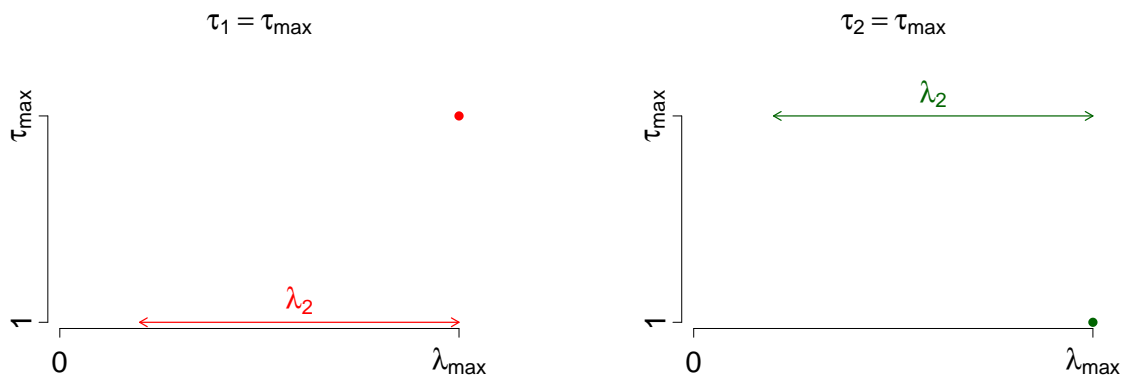
Im folgenden soll dieser Verlauf und die verschiedenen Kombinationsmöglichkeiten genauer untersucht werden.

Ob man sich für Stimulus 1 oder 2 entscheidet, hängt von der Höhe des mittleren Wertes $\bar{X} := 1/n \sum X_i$ ab.

Bemerkung 3.2. In Abhängigkeit von n ist für $\tau_1 \neq \tau_2$ die Entscheidungsschranke gegeben durch die Funktion

$$\begin{aligned} f(n) &= \frac{\frac{\lambda_1 - \lambda_2}{n} - \log \frac{\lambda_1 \tau_2}{\lambda_2 \tau_1}}{1/\tau_2 - 1/\tau_1} \\ &= \frac{1}{n} \cdot \frac{\lambda_1 - \lambda_2}{\tau_1 - \tau_2} \cdot \tau_1 \tau_2 - \frac{\log \frac{\lambda_1}{\lambda_2} \tau_1 \tau_2}{\tau_1 - \tau_2} + \frac{\log \frac{\tau_1}{\tau_2} \tau_1 \tau_2}{\tau_1 - \tau_2} \\ &=: \frac{1}{n} \cdot a - b + c \end{aligned}$$

Ist $\tau_1 > \tau_2$ entscheidet man sich für Stimulus 1, falls $1/n \sum X_i > f(n)$ ist, sonst umgekehrt. Bei der Parameterwahl gibt es zwei verschiedene Kombinationsmöglichkeiten: Entweder man fasst das größere τ mit dem größeren λ zusammen, vgl. Abbildung 24 (a), oder das kleinere τ zum größeren λ , siehe Abbildung 24 (b).



(a) Der maximalen Rate wird die maximale Phase zugeordnet.

(b) Der maximalen Rate wird die minimale Phase zugeordnet.

Abbildung 24: Zwei mögliche Kombinationen im Fall von zwei Stimuli. Der optimale Wert für λ muss berechnet werden.

Bemerkung 3.3. Für den Fall $\lambda_1 > \lambda_2$ und $\tau_1 > \tau_2$ (der rote Fall) gilt $a > 0$, $b > 0$ und $c > 0$. Für den Fall $\lambda_1 > \lambda_2$ und $\tau_1 < \tau_2$ (der grüne Fall) gilt $a' = -a$, $b' = -b$ und $c' = c$. Wie man auch in Abbildung 25 erkennen kann, sind die beiden Fälle achsensymmetrisch zu c .

Auf den ersten Blick könnte man zu dem Schluss kommen, dass der rote Fall stets eine höhere Detektionswahrscheinlichkeit besitzen sollte, da man sich im grünen Fall zum Beispiel bei diesen konkreten Werten bei $n = 1$ nie für die Kombination mit dem kleineren τ entscheiden würde, egal wie die zweite Stufe ausfällt und dazu auch noch das größere λ wählen würde. Aber wir werden sehen, dass in manchen Bereichen auch der grüne Fall eine höhere Detektionswahrscheinlichkeit aufweist. Meist dominiert aber der rote Fall, auch wenn die Unterschiede nur sehr gering sind.

In Abbildung 26 wird der rote Fall näher untersucht. Erwartungsgemäß basiert die Entscheidung für ähnliche τ fast nur noch auf Basis von λ , vgl. Abbildung 26 (a). Für ähnliche τ fließt für große n die zweite Stufe nicht mehr in die Entscheidung ein, ähnelt also dem grünen Fall in umgekehrter Weise, jetzt in der zweiten Stufe. Nun kann man sich schon besser vorstellen, dass in manchen Fällen auch die umgekehrte Parameterkombination vorteilhaft sein kann.

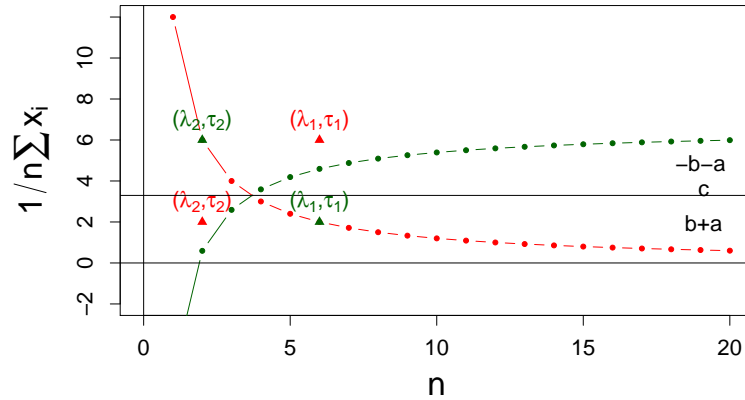


Abbildung 25: Beziehung zwischen den beiden Entscheidungsschranken für die Fälle $\lambda_1 > \lambda_2$, $\tau_1 > \tau_2$ (rot) und $\lambda_1 > \lambda_2$, $\tau_1 < \tau_2$ (grün). Die aus den beiden Kombinationsmöglichkeiten resultierenden Schranken sind symmetrisch zu

$$c = \frac{\log \frac{\tau_1 \tau_1 \tau_2}{\tau_1 - \tau_2}}{\tau_1 - \tau_2}$$

In Abbildung 27 ist der grüne Fall genauer dargestellt. Je dichter die τ aneinanderliegen, desto geringer fällt auch der Einfluss der zweiten Stufe aus. Aber auch für ein großes Verhältnis τ_{\max}/τ_{\min} hat für kleines n die zweite Stufe keinen Einfluss, vgl. Abbildung 27 (c). Also lässt sich erwarten, dass für größeres τ_{\max}/τ_{\min} und „größere“ λ_1 der rote Fall etwas höhere Detektionswahrscheinlichkeit besitzen wird. Das wird sich bei den weiteren Analysen größtenteils auch bestätigen wird.

Wie in Kapitel 2.2 beschrieben, lässt sich für vorgegebene Parameter λ_1 , λ_2 , τ_1 und τ_2 die Detektionswahrscheinlichkeit berechnen. Hält man nun λ_1 fest, kann man für den grünen und roten Fall in Abhängigkeit von τ_{\max}/τ_{\min} numerisch das λ_2 bestimmen, das die Detektionswahrscheinlichkeit maximiert.

Auf diese Weise wurde Abbildung 28 erzeugt. In (a) und (b) ist das optimale λ_2 in Abhängigkeit von τ_{\max}/τ_{\min} für $\lambda_1 = 4$ und $\lambda_1 = 6$ dargestellt. Im grünen Fall steigt das optimale λ_2 stets für wachsendes τ_{\max}/τ_{\min} an. Im roten Fall nimmt das optimale λ_2 zu Beginn sogar kurz ab, was etwas kontraintuitiv ist. Anschließend nimmt aber auch dieses kontinuierlich zu.

Ab einem gewissen τ_{\max}/τ_{\min} springt das optimale λ_2 auf λ_1 . Ab dieser Stelle ist es besser, die erste Stufe zu übergehen und nur auf Basis von τ zu unterscheiden. Für $\lambda_1 = 4$ springt das optimale λ_2 im roten Fall etwas früher als im grünen. Für $\lambda_1 = 6$ ist dieses allerdings umgekehrt.

Die Detektionswahrscheinlichkeit wird von keinem der beiden Fälle dominiert, doch besitzt

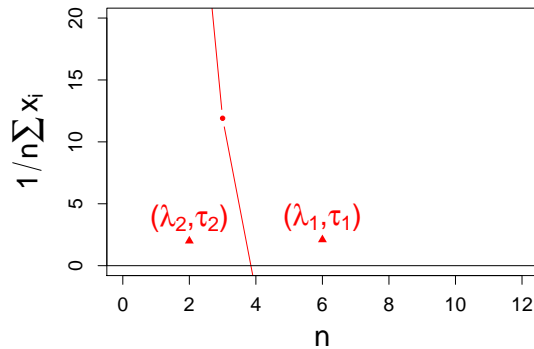
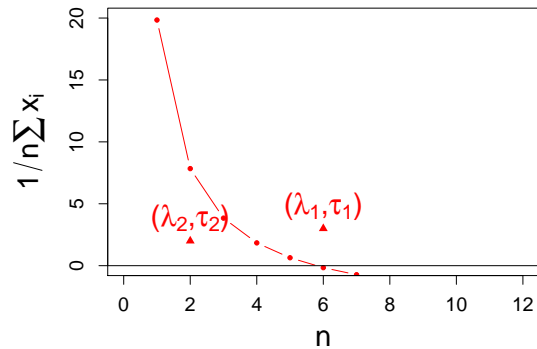
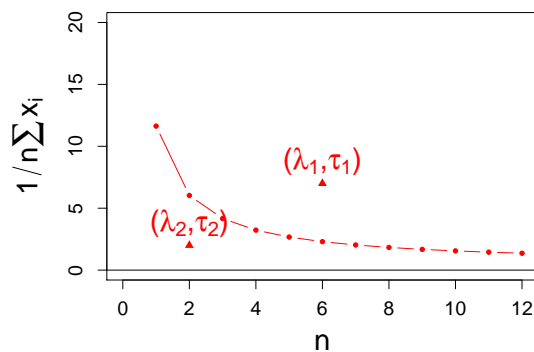
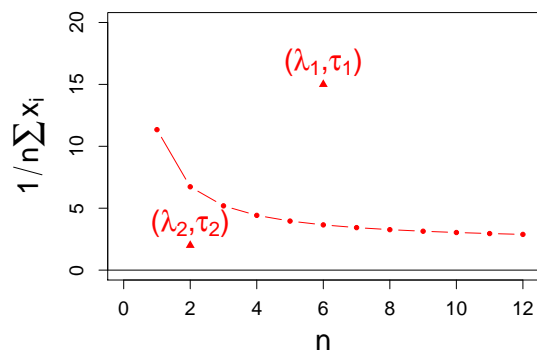
(a) $\tau_1 = 2.1$ und $\tau_2 = 2$.(c) $\tau_1 = 3$ und $\tau_2 = 2$.(b) $\tau_1 = 7$ und $\tau_2 = 2$.(d) $\tau_1 = 15$ und $\tau_2 = 2$.

Abbildung 26: Auswirkung auf die Entscheidungsschranken für den Fall $\lambda_1 > \lambda_2$, $\tau_1 > \tau_2$ durch Parameterveränderungen in kritischen Bereichen. Konstant bleibt dabei $\lambda_1 = 6$, $\lambda_2 = 2$ und $\tau_2 = 2$.

der rote Fall öfters eine höhere Detektionswahrscheinlichkeit. Der maximale Unterschied beträgt aber nur 0.004 und kommt in beiden Richtungen vor.

In Abbildung 28 (c) und (d) ist die optimale Detektionswahrscheinlichkeit in Abhängigkeit von τ_{\max}/τ_{\min} dargestellt. Dies entspricht der dickeren Linie. Die dünnere Linie entspricht der Detektionswahrscheinlichkeit, wenn man nur auf Basis von der ersten Stufe mit dem gewählten λ_2 entscheiden würde.

Für steigendes τ_{\max}/τ_{\min} nimmt die gemeinsame Detektionswahrscheinlichkeit wie erwartet kontinuierlich zu, auch wenn die Entscheidungskraft nur auf Basis der ersten Stufe kontinuierlich fällt. Dort wo das optimale λ_2 auf λ_1 springt, besitzt man hier einen Abwärtssprung auf $1/3 \cdot (2 - \exp(-\lambda_1))$. Die beiden Stimuli lassen sich auf Basis von λ nicht

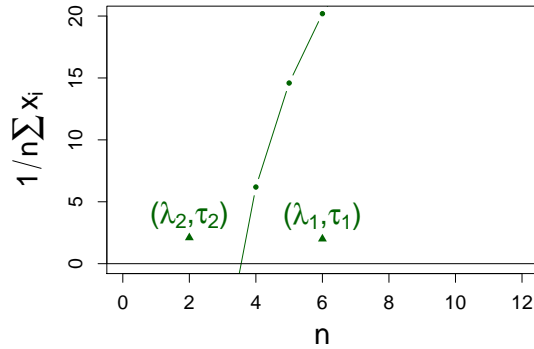
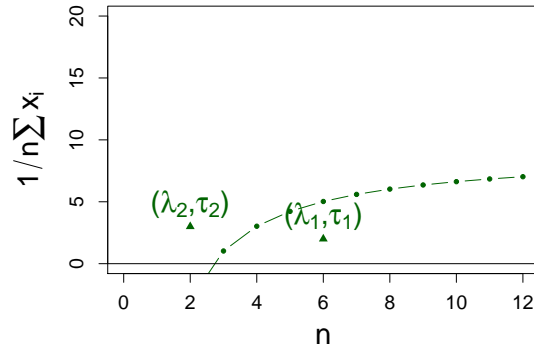
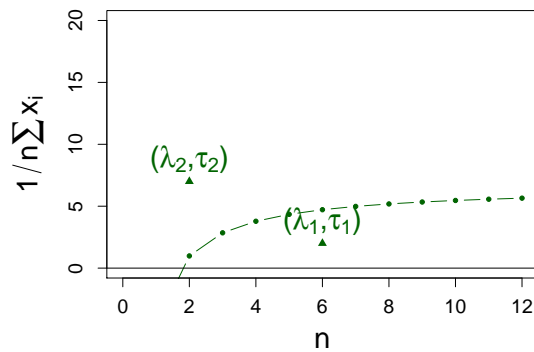
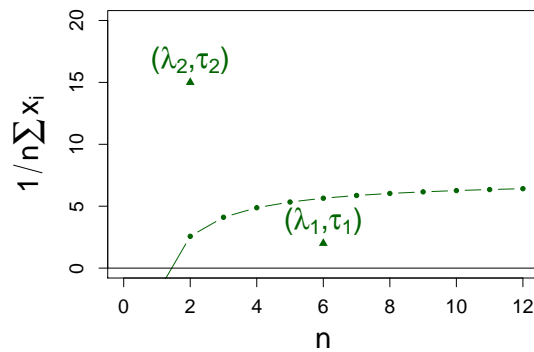
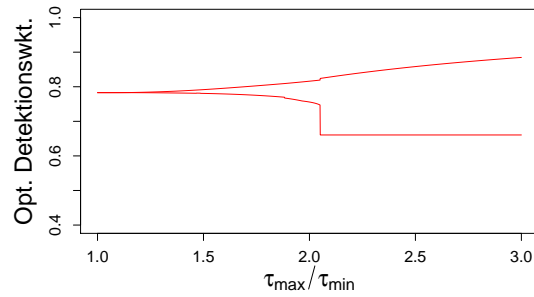
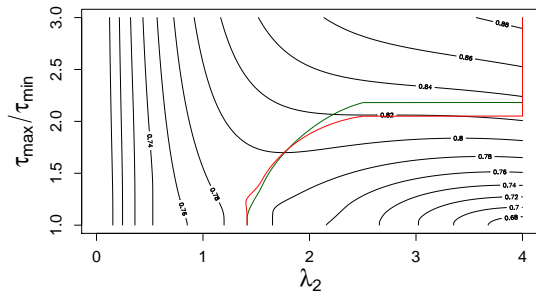
(a) $\tau_1 = 2$ und $\tau_2 = 2.1$.(c) $\tau_1 = 2$ und $\tau_2 = 3$.(b) $\tau_1 = 2$ und $\tau_2 = 7$.(d) $\tau_1 = 2$ und $\tau_2 = 15$.

Abbildung 27: Auswirkung auf die Entscheidungsschranken für den Fall $\lambda_1 > \lambda_2$, $\tau_1 < \tau_2$ durch Parameterveränderungen in kritischen Bereichen. Konstant bleibt dabei $\lambda_1 = 6$, $\lambda_2 = 2$ und $\tau_1 = 2$.

mehr auseinanderhalten, nur vom Nullstimulus lassen sich diese noch unterscheiden. Ab dem Zeitpunkt, zu dem nur noch auf Basis von τ unterschieden wird, besitzt die Detektionswahrscheinlichkeit anscheinend auch einen kleinen Aufwärtssprung.

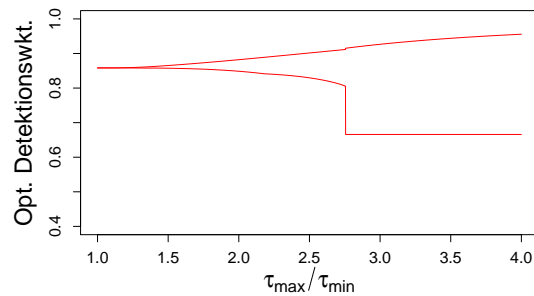
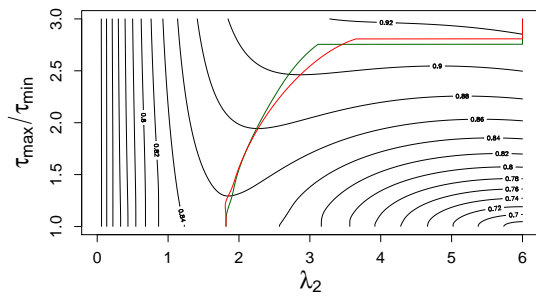
In Abbildung 29 ist genauer dargestellt, für welche τ_{\max}/τ_{\min} der rote oder der grüne Fall eine höhere Detektionswahrscheinlichkeit besitzt.

Für $\lambda_1 = 4$ dominiert zu Beginn, außer unmittelbar am Anfang, größtenteils der grüne Fall. Für $\lambda_1 = 4$ ist beim roten Fall λ_2 auch etwas früher auf λ_1 gesprungen und besitzt in diesem Bereich auch eine höhere Detektionswahrscheinlichkeit. Ab der Stelle, an der sowohl im grünen als auch im roten Fall das λ_2 auf λ_1 gesprungen ist, besitzen beide Fälle natürlich dieselbe Detektionswahrscheinlichkeit.



(a) Das optimale λ_2 für beide Kombinationsmöglichkeiten in Abhängigkeit von τ_{\max}/τ_{\min} für $\lambda_1 = 4$.

(c) Vergleich maximale Detektionswkt. für $\lambda_1 = 4$ auf Basis von λ und τ mit allein auf Basis des gewählten λ_2 .



(b) Das optimale λ_2 für beide Kombinationsmöglichkeiten in Abhängigkeit von τ_{\max}/τ_{\min} für $\lambda_1 = 6$.

(d) Vergleich maximale Detektionswkt. für $\lambda_1 = 6$ auf Basis von λ und τ mit allein auf Basis des gewählten λ_2 .

Abbildung 28: Optimales λ_2 und maximale Detektionswahrscheinlichkeit für den roten ($\lambda_1 > \lambda_2$ und $\tau_1 > \tau_2$) und grünen Fall ($\lambda_1 > \lambda_2$ und $\tau_1 < \tau_2$) für festes λ_1 in Abhängigkeit von τ_{\max}/τ_{\min} .

Für $\lambda_1 = 6$ ist der grüne Fall seltener vertreten, hauptsächlich nur zu Beginn. Aber für $\lambda_1 = 6$ ist der grüne Fall an der Sprungstelle kurzzeitig dominierend. In Abbildung 28 war zu erkennen, dass im grünen Fall das λ_2 knapp früher auf λ_1 springt als im roten Fall. Genau in diesem Bereich besitzt der grüne Fall eine höhere Detektionswahrscheinlichkeit als der rote Fall.

Zusammenfassend lässt sich festhalten, dass sich der grundsätzliche Verlauf für beide λ 's ähnelt, im Detail aber schon unterschiedliche Aufteilungen aufweist.

Das Zusammennehmen der beiden Stufen scheint die Frage nach den optimalen Parametern numerisch relativ aufwändig zu gestalten. Wird die Detektionswahrscheinlichkeit durch die Kombination der beiden Stufen überhaupt nennenswert beeinflusst. Bestimmt man das τ_{\max}/τ_{\min} , das benötigt wird, um die Detektionswahrscheinlichkeit gegenüber nur der Unterscheidung auf Basis von λ um 5% zu erhöhen, erhält man Abbildung 30.

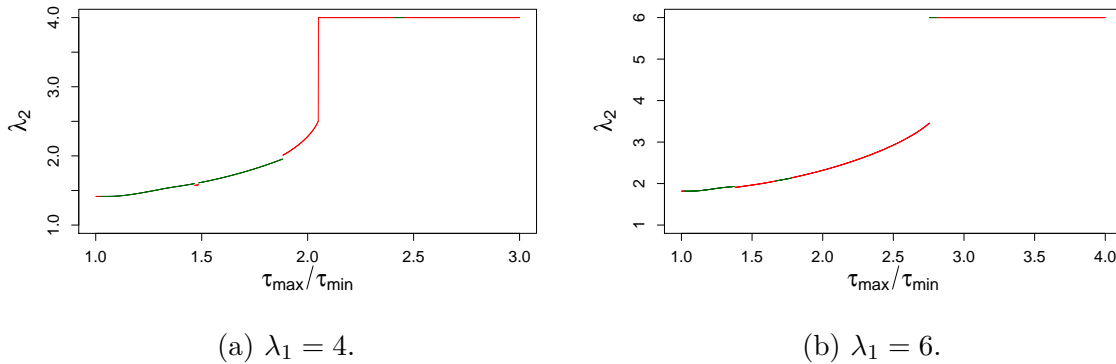


Abbildung 29: Für welche τ_{\max}/τ_{\min} besitzt der rote Fall ($\lambda_1 > \lambda_2$ und $\tau_1 > \tau_2$) oder der grüne Fall ($\lambda_1 > \lambda_2$ und $\tau_1 < \tau_2$) die höhere Detektionswahrscheinlichkeit.

Ab $\lambda_1 = 13$ ist ein Erhöhen der Detektionswahrscheinlichkeit um 5% gar nicht mehr mög-

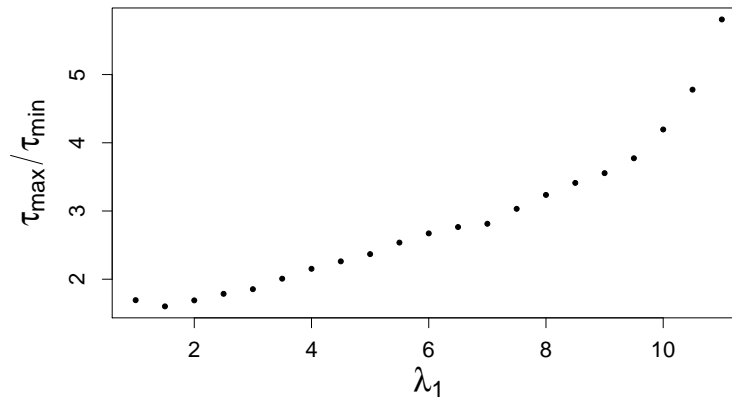
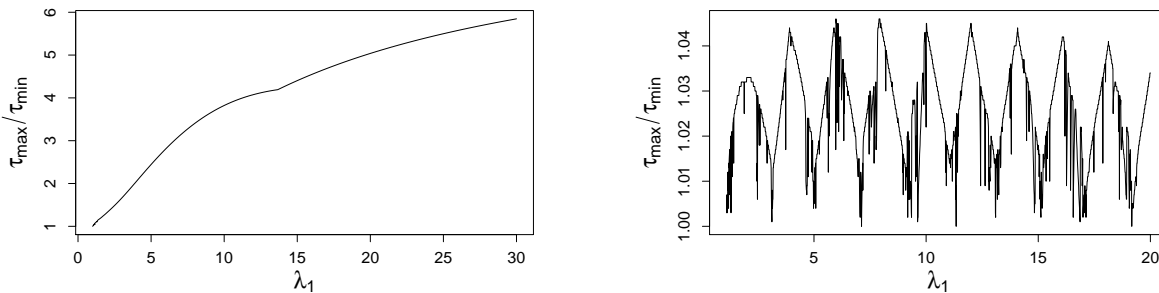


Abbildung 30: Welches τ_{\max}/τ_{\min} wird in Abhängigkeit von λ_1 benötigt, um die Detektionswahrscheinlichkeit gegenüber nur der Unterscheidung auf Basis von λ um 5% zu erhöhen.

lich. Dies liegt natürlich daran, dass bei zwei Stimuli für so hohes λ_1 ein Unterscheiden sehr sicher möglich ist. Bei mehreren Stimuli müsste dementsprechend das λ_1 immer größer gewählt werden und der Einfluss der zweiten Stufe dürfte stärker ausfallen.

Betrachtet man außerdem andersherum, ab welchem τ_{\max}/τ_{\min} man bereits auf die erste Stufe verzichtet, erkennt man, dass dieses im Verhältnis zu λ_1 relativ gering ausfällt. In Abbildung 31 (a) ist dies dargestellt. Selbst wenn man ein λ_1 von 30 zur Verfügung hat, würde man zur Unterscheidung nur auf Basis von τ bereits ab $\tau_{\max} = 6$ umstellen.



- (a) Ab welchem τ_{\max}/τ_{\min} ist es in Abhängigkeit von λ_1 optimal, nur noch auf Basis von τ zu unterscheiden.
- (b) Bis zu welchem τ_{\max}/τ_{\min} hat in Abhängigkeit von λ_1 die zweite Stufe keinen Einfluss auf λ_2 .

Abbildung 31: Die Bedeutung der Phase für 2 Stimuli.

Bemerkung 3.4. *Die zweite Stufe wird umso mächtiger, je größer die erste Stufe ausfällt, und damit, je größer das λ_1 ist. Die zweite Stufe besteht nämlich aus unabhängig, identisch verteilten Zufallsvariablen und zur Unterscheidung greift man nur auf deren Summe zurück. Für großes n unterliegt deren Mittelwert immer geringeren Schwankungen und folgt dem Gesetz der großen Zahlen.*

Für große λ_1 lassen sich mit der ersten Stufe die beiden Stimuli schon sehr sicher unterscheiden und die zweite Stufe kann die Detektionswahrscheinlichkeit nicht nennenswert steigern, doch benötigt man dort auch nur ein verhältnismäßig deutlich geringeres τ_{\max}/τ_{\min} um eine höhere Detektionswahrscheinlichkeit aufzuweisen. Interessant ist, wie sich dieser Effekt bei mehreren Stimuli auswirkt. Bei einer großen Anzahl Stimuli lassen sich diese auch für großes λ_1 in der ersten Stufe nur schwierig auseinanderhalten, doch greift dort in der zweiten Stufe das Gesetz der großen Zahlen. Ab wann dominiert die zweite Stufe, oder dominieren dort erst mal andere Kombinationsmöglichkeiten, die es im 2-Stimuli-Fall nicht gibt?

In Abbildung 31 (b) ist dargestellt, bis zu welchem τ_{\max}/τ_{\min} die zweite Stufe in Abhängigkeit von λ_1 das optimale λ_2 nicht beeinflusst. Wie man deutlich erkennen kann, wirkt sich selbst bei großem λ_1 ein sehr geringes τ_{\max}/τ_{\min} sofort auf die erste Stufe aus. Die Schwankungen vom minimalen τ_{\max}/τ_{\min} können dabei durchaus auf Rechenungenauigkeiten zurückgehen, eine sofortige Auswirkung der zweiten Stufe ist nicht auszuschließen.

3.3 Drei Stimuli

Diese Analysen haben im Zwei-Stimuli-Fall die Bedeutung der zweiten Stufe, auch für große λ_1 , aufgezeigt. Nun wird untersucht, wie sich für einen weiteren Stimulus die Aufteilung der optimalen Parameter verhält.

Gegeben drei Stimuli S_1, S_2, S_3 mit $0 < \lambda_3 \leq \lambda_2 \leq \lambda_1 \leq \lambda_{\max}$ und $\tau_{\max}/\tau_{\min} \leq c$, sind die optimalen $\tilde{\lambda}_2, \tilde{\lambda}_3, \tilde{\tau}_1, \tilde{\tau}_2, \tilde{\tau}_3$ gesucht, sodass

$$p_D(0, \lambda_1, \tilde{\lambda}_2, \tilde{\lambda}_3, \tilde{\tau}_1, \tilde{\tau}_2, \tilde{\tau}_3) \geq p_D(0, \lambda_1, \lambda_2, \lambda_3, \tau_1, \tau_2, \tau_3)$$

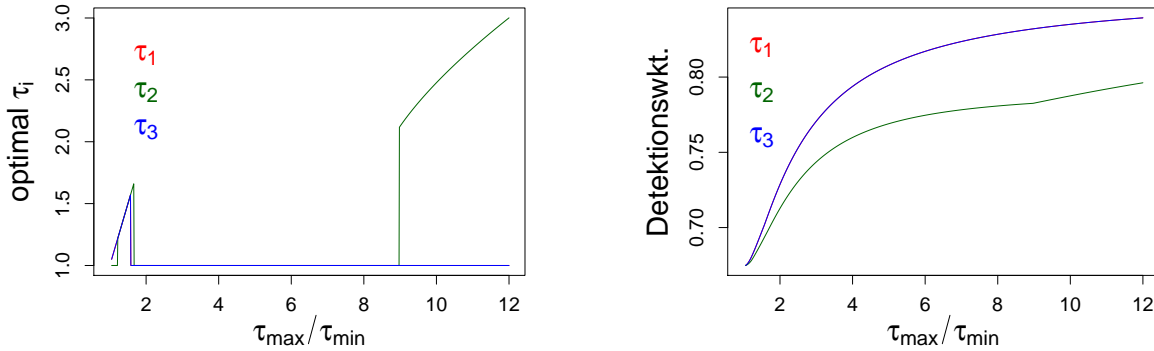
für alle $\lambda_3 \leq \lambda_2 \leq \lambda_1$ und für alle τ_1, τ_2, τ_3 mit $\tau_{\max}/\tau_{\min} \leq c$.

Für drei Stimuli ergibt sich eine Schwierigkeit: Bei Zweien konnte man die erste Stufe in Abhängigkeit von der zweiten Stufe betrachten. Dort wird nämlich offensichtlich das maximal mögliche Verhältnis τ_{\max}/τ_{\min} gewählt, da dies unabhängig von der ersten Stufe die Detektionswahrscheinlichkeit maximiert. Es ist nur nicht klar, ob τ_1 oder τ_2 maximal gewählt wird.

3.3.1 Zweite Stufe in Abhängigkeit von Stufe 1

Bei Dreien ist nur klar, dass ein τ minimal und eines maximal gewählt wird. Doch gibt es hier schon sechs mögliche Kombinationen. Offen bleibt die Frage, wie das mittlere τ gewählt werden soll. Deshalb wird hier das umgekehrte Vorgehen eingeschlagen: Vorgeben feste Werte für $\lambda_1 \geq \lambda_2 \geq \lambda_3$, wie lauten die optimalen Parameter τ_1, τ_2 und τ_3 ?

In Abbildung 32 ist dies für $\lambda_1 = 6, \lambda_2 = 3, \lambda_3 = 1$ dargestellt.



(a) Optimales τ_i in Abhängigkeit von τ_{\max}/τ_{\min} für festes $\lambda_1 = 6, \lambda_2 = 3, \lambda_3 = 1$.

(b) Maximale Detektionswkt. gemäß Wahl aus (a) für $\lambda_1 = 6, \lambda_2 = 3, \lambda_3 = 1$.

Abbildung 32: Optimale Detektionswahrscheinlichkeit und optimales τ_i für festes $\lambda_1 = 6, \lambda_2 = 3, \lambda_3 = 1$ in Abhängigkeit von τ_{\max}/τ_{\min} für die beiden anderen τ . Für die grüne Linie wurde das optimale τ_2 bestimmt, für die beiden Fälle $\tau_1 = \tau_{\max}$ und $\tau_3 = \tau_{\min}$ oder $\tau_1 = \tau_{\min}$ und $\tau_3 = \tau_{\max}$. Analog für die rote und blaue Linie.

Die Abbildung ist folgendermaßen entstanden: Es wurde jeweils das optimale τ_1 (rot), τ_2 (grün) und τ_3 (blau) berechnet, in Abhängigkeit von τ_{\max}/τ_{\min} für die beiden anderen τ . Dabei wurde gleichzeitig das Maximum über die beiden Möglichkeiten, welches der beiden

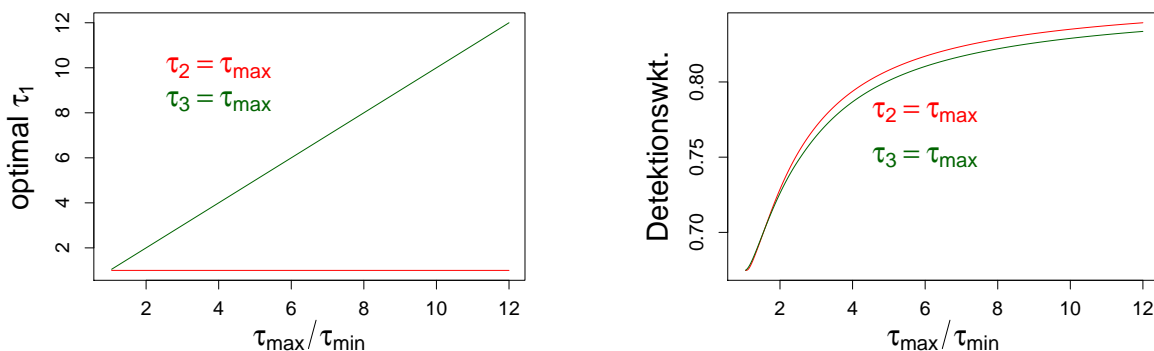
übrigen das maximale und das minimale τ annimmt, genommen.

Auf den ersten Blick fällt auf, dass in diesem Bereich das optimale τ_1 und τ_3 identisch sind und sich auch dieselbe Detektionswahrscheinlichkeit ergibt. Für kleines τ_{\max}/τ_{\min} ist es optimal, für τ_1 und τ_3 das maximale τ zu wählen. Kurz vor zwei springt die optimale Wahl aber direkt auf das minimale τ_1 , hier o.B.d.A $\tau_{\min} = 1$, und bleibt darauf in diesem Bereich konstant.

Das optimale τ_2 läuft etwas unregelmäßiger: Für sehr kleines τ_{\max}/τ_{\min} wird das optimale τ_2 gerade τ_{\min} gesetzt, dann springt es direkt auf das maximale τ und kurz darauf wieder auf τ_{\min} . Bei τ_{\max}/τ_{\min} etwa 9 springt es aber nochmal auf einen Wert zwischen τ_{\max} und τ_{\min} und steigt kontinuierlich.

Der grüne Fall besitzt aber im gesamten Bereich eine geringer Detektionswahrscheinlichkeit als die beiden anderen. Dies ist auch einleuchtend: Wählt man τ_1 und τ_3 als maximal und minimal, lassen sich gerade die Stimuli in der zweiten Stufe gut unterscheiden, die sich bereits in der ersten Stufe besser auseinandergelassen haben lassen. Der zweite Stimulus hingegen lässt sich in der ersten Stufe bereits schlechter unterscheiden und bei der zweiten Stufe bei dieser Kombination dann auch. Es sollte also besser sein, die τ 's „abwechselnd“ zu den λ 's zu wählen.

Um die Sprünge besser zu verstehen, werden für den roten Fall die beiden Möglichkeiten, ob τ_2 oder τ_3 als maximal gewählt werden, genauer analysiert. In Abbildung 33



(a) Optimales τ_1 in Abhängigkeit von τ_{\max}/τ_{\min} für die beiden Fälle $\tau_2 = \tau_{\max}$ oder $\tau_3 = \tau_{\max}$ für festes $\lambda_1 = 6$, $\lambda_2 = 3$, $\lambda_3 = 1$.

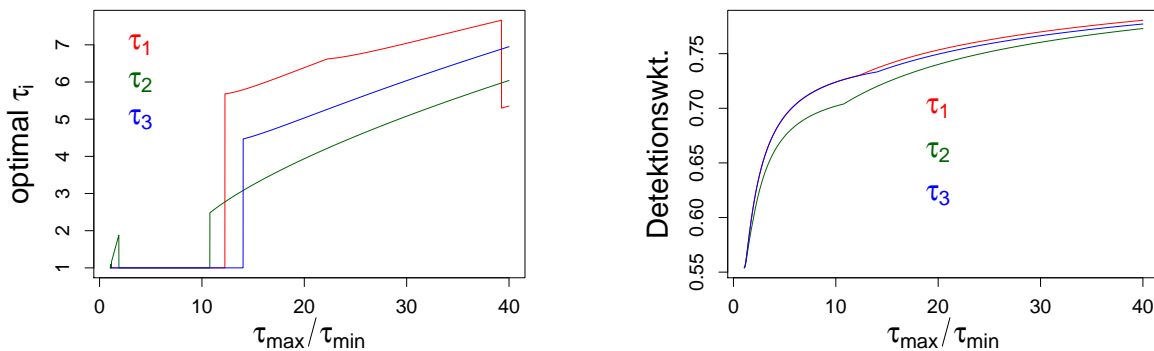
(b) Maximale Detektionswkt. gemäß Wahl aus (a) für $\lambda_1 = 6$, $\lambda_2 = 3$, $\lambda_3 = 1$.

Abbildung 33: Optimale Detektionswahrscheinlichkeit und optimales τ_1 für festes $\lambda_1 = 6$, $\lambda_2 = 3$, $\lambda_3 = 1$ in Abhängigkeit von τ_{\max}/τ_{\min} für die beiden Möglichkeiten $\tau_2 = \tau_{\max}$ oder $\tau_3 = \tau_{\max}$.

erkennt man, dass in diesem Bereich das für $\tau_3 = \tau_{\max}$ optimale τ_1 gerade das maximale τ ist und für $\tau_2 = \tau_{\max}$ gerade das minimale τ . τ_1 wird also in diesem Bereich gerade so gewählt, dass sich die benachbarten τ maximal unterscheiden. Wie zu erkennen, besitzt nur zu Beginn $\tau_3 = \tau_{\max}$ eine höhere Detektionswahrscheinlichkeit, wodurch der Sprung zustande kommt.

Für τ_3 erhält man dieselbe Abbildung. Ob man nach τ_1 oder τ_3 maximiert, ist also gerade in diesem Bereich symmetrisch, weshalb sich offensichtlich in beiden Fälle dieselbe Detektionswahrscheinlichkeit ergibt.

Intuitiv erwartet man aber für große τ_{\max}/τ_{\min} einen etwas anderen Verlauf. Bei den bisherigen Anordnungen lassen sich zwei Stimuli auf Basis von τ nicht unterscheiden, das sollte sich für im Verhältnis zu λ_1 sehr große τ_{\max}/τ_{\min} verändern. Deswegen ist in Abbildung 34 analog zur Abbildung 32 der Verlauf für $\lambda_1 = 3$, $\lambda_2 = 2$, $\lambda_3 = 1$ dargestellt. Dort treten



(a) Optimales τ_i in Abhängigkeit von τ_{\max}/τ_{\min} für festes $\lambda_1 = 3$, $\lambda_2 = 2$, $\lambda_3 = 1$.

(b) Maximale Detektionswkt. gemäß Wahl aus (a) für $\lambda_1 = 3$, $\lambda_2 = 2$, $\lambda_3 = 1$.

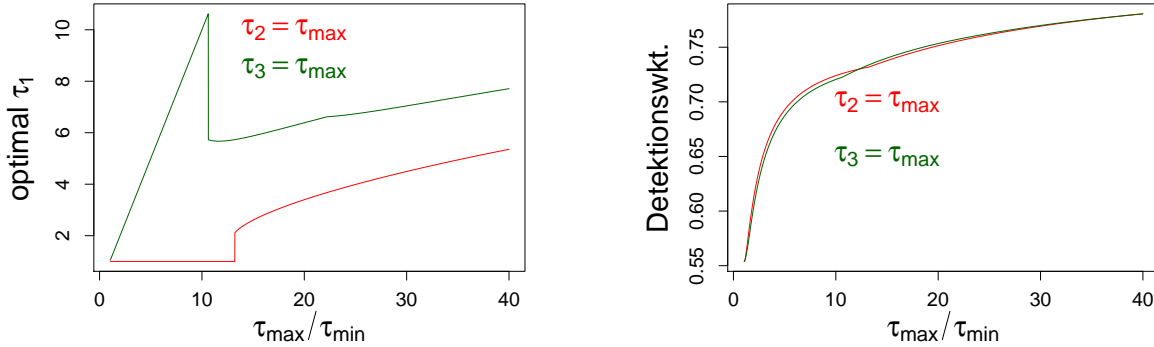
Abbildung 34: Optimale Detektionswahrscheinlichkeit und optimales τ_i für festes $\lambda_1 = 3$, $\lambda_2 = 2$, $\lambda_3 = 1$ in Abhängigkeit von τ_{\max}/τ_{\min} für die beiden anderen τ . Für die grüne Linie wurde das optimale τ_2 bestimmt, für die beiden Fälle $\tau_1 = \tau_{\max}$ und $\tau_3 = \tau_{\min}$ oder $\tau_1 = \tau_{\min}$ und $\tau_3 = \tau_{\max}$. Analog für die rote und blaue Linie.

die Sprünge früher auf und die numerische Berechnung verläuft schneller.

Nun lässt sich der vollständige Verlauf erkennen: Bis etwa $\tau_{\max}/\tau_{\min} = 11$ entspricht dies dem bisher diskutierten Verlauf. Doch springt ab hier auch das optimale τ_1 auf einen Wert zwischen τ_{\max} und τ_{\min} und nun hebt sich auch die rote von der blauen Kurve ab. Der rote Fall besitzt von da an auch eine höhere Detektionswahrscheinlichkeit.

Bis kurz vor $\tau_{\max}/\tau_{\min} = 40$ steigt das optimale τ_1 kontinuierlich an. Dann springt das τ_1 nochmals deutlich nach unten, von wo an es wieder weiter kontinuierlich wächst. Die blaue Kurve folgt der roten Kurve zeitverzögert.

In Abbildung 35 wird genauer betrachtet, wie das optimale λ_1 zustande kommt, in Abhängigkeit der beiden möglichen Kombinationen $\tau_2 = \tau_{\max}$ oder $\tau_3 = \tau_{\max}$.



(a) Optimales τ_1 in Abhängigkeit von τ_{\max}/τ_{\min} für die beiden Fälle $\tau_2 = \tau_{\max}$ oder $\tau_3 = \tau_{\max}$ für festes $\lambda_1 = 3$, $\lambda_2 = 2$, $\lambda_3 = 1$.

(b) Maximale Detektionswkt. gemäß Wahl aus (a) für $\lambda_1 = 3$, $\lambda_2 = 2$, $\lambda_3 = 1$.

Abbildung 35: Optimale Detektionswahrscheinlichkeit und optimales τ_1 für festes $\lambda_1 = 3$, $\lambda_2 = 2$, $\lambda_3 = 1$ in Abhängigkeit von τ_{\max}/τ_{\min} für die beiden Möglichkeiten $\tau_2 = \tau_{\max}$ oder $\tau_3 = \tau_{\max}$.

Der Anfang gleicht wieder dem bekannten Bild. Doch an der Stelle, an der das optimale τ_1 nach oben springt, springt hier im Falle $\tau_3 = \tau_{\max}$ das τ_1 von τ_{\max} nach unten. Ab hier übersteigt die Detektionswahrscheinlichkeit von $\tau_3 = \tau_{\max}$ nochmals den Fall $\tau_2 = \tau_{\max}$. Dies dauert aber nur solange, bis auch im Fall $\tau_2 = \tau_{\max}$ das τ_1 von τ_{\min} nach oben springt. Wie im rechten Teil der Abbildung zu erkennen, verliert man aber kaum an Detektionswahrscheinlichkeit, wenn man sich im gesamten Bereich immer für den Fall $\tau_2 = \tau_{\max}$ entscheiden würde.

Bemerkung 3.5. Für festgehaltene Raten lässt sich zusammenfassend hervorheben, dass größtenteils τ_2 und τ_3 als maximales und minimales τ gewählt werden und nach τ_1 maximiert wird. Genauer scheint folgender Verlauf vorzuliegen:

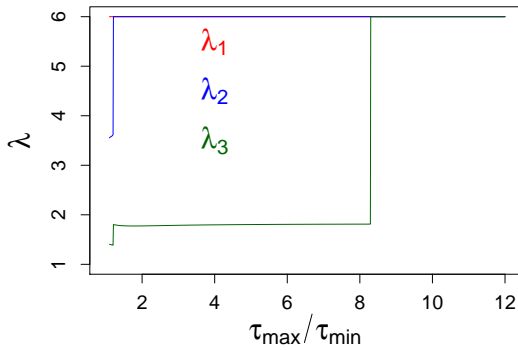
$$(\tau_{\max}, \tau_{\min}, \tau_{\max}) \rightarrow (\tau_{\min}, \tau_{\max}, \tau_{\min}) \rightarrow (\tau_1, \tau_{\min}, \tau_{\max}) \rightarrow (\tau_1, \tau_{\max}, \tau_{\min}) \quad \tau_{\min} < \tau_1 < \tau_{\max}.$$

Nun ist klarer, wie die τ am besten kombiniert werden. Es fehlt aber noch das Zusammenspiel zwischen λ und τ . Dieses wird im folgenden nun genauer betrachtet.

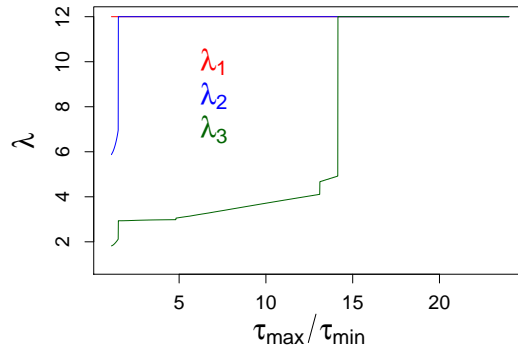
3.3.2 Global beste Wahl für λ und τ

Gesucht sind also die Wahlen für $\lambda_1 \geq \lambda_2 \geq \lambda_3$ und τ_1, τ_2, τ_3 , sodass die Detektionswahrscheinlichkeit maximiert wird. Wie bereits bekannt, ist diese Frage nur sinnvoll für festgehaltenes λ_1 und einem vorgegebenen Verhältnis τ_{\max}/τ_{\min} .

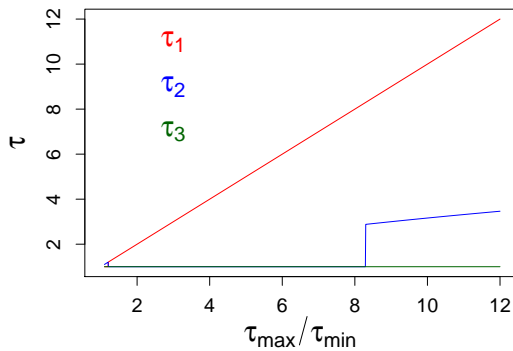
In Abbildung 36 wurde die optimale Parameterwahl für $\lambda_1 = 6$ und $\lambda_1 = 12$ in Abhängigkeit von τ_{\max}/τ_{\min} numerisch berechnet.



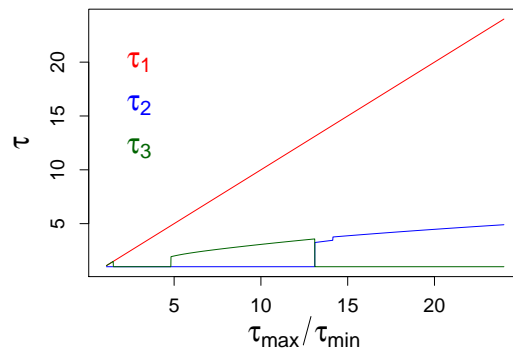
(a) Optimale Raten λ_2 und λ_3 für $\lambda_1 = 6$.



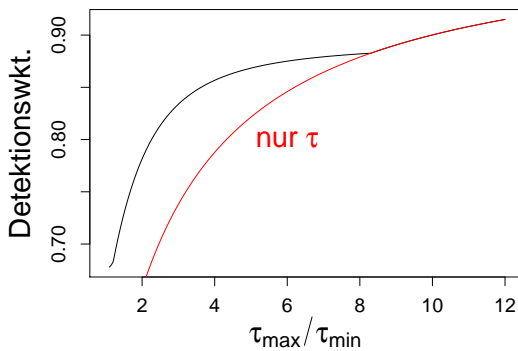
(d) Optimale Raten λ_2 und λ_3 für $\lambda_1 = 12$.



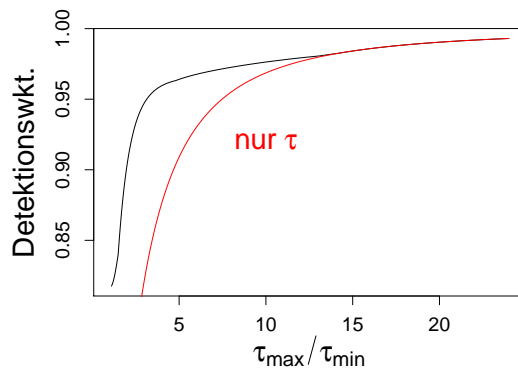
(b) Optimale Phasen τ_1 , τ_2 und τ_3 für $\lambda_1 = 6$.



(e) Optimale Phasen τ_1 , τ_2 und τ_3 für $\lambda_1 = 12$.



(c) Maximale Detektionswkt. für $\lambda_1 = 6$ im Vergleich zu einem reinen Phasen-Code.



(f) Maximale Detektionswkt. für $\lambda_1 = 12$ im Vergleich zu einem reinen Phasen-Code.

Abbildung 36: Optimale Parameterwahl und optimale Detektionswahrscheinlichkeit für 3 Stimuli für $\lambda_1 = 6$ und $\lambda_1 = 12$ in Abhängigkeit von τ_{\max}/τ_{\min} .

In Abbildung 37 sind zudem die vorkommenden Kombination dargestellt.

Für die optimalen λ gibt es drei grundsätzlich verschiedene Wahlen: Für sehr kleines τ_{\max}/τ_{\min} wird nur ein λ maximal gesetzt, es lassen sich also alle Stimuli auch nur auf Basis von λ unterscheiden. Für $\tau_{\max}/\tau_{\min} = 1$ wird nur noch auf Basis von λ unterschieden und man befindet sich in dem aus Kapitel 2.3 bekannten Fall.

Bereits im Verhältnis zu λ_1 kleinem τ_{\max}/τ_{\min} ist es vorteilhafter, zwei λ maximal zu wählen und damit das Unterscheiden zweier Stimuli nur noch auf Basis von τ zu ermöglichen. Erst für großes τ_{\max}/τ_{\min} , etwa um zwei bis drei größer als λ_1 , ist es optimal, alle λ maximal zu wählen und nur noch auf Basis von τ zu unterscheiden. In diesem Fall ist aus Kapitel 2.4 bereits die optimale Parameterkombination bekannt.

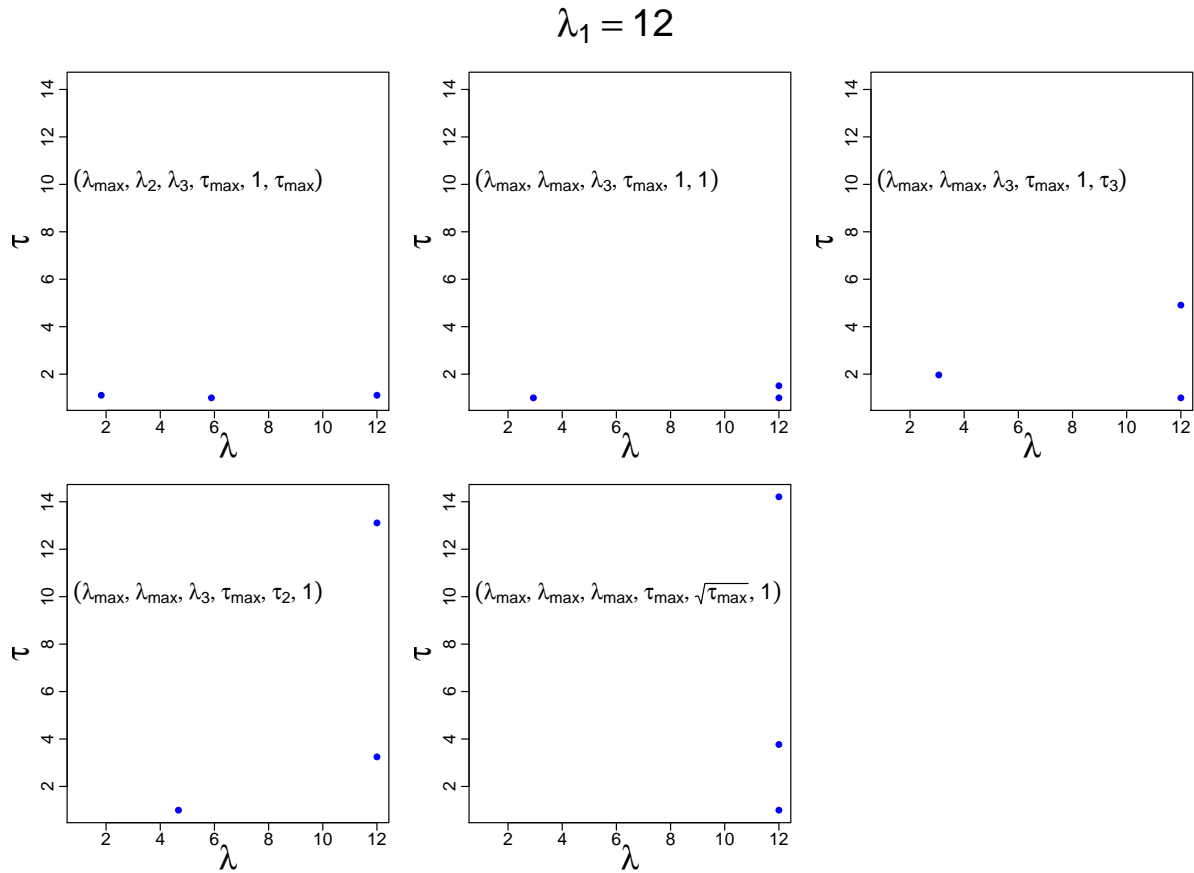


Abbildung 37: Auftretende Kombinationen für 3 Stimuli für $\lambda_1 = 12$. Die Koordinaten der blauen Punkte entsprechen den Werten von λ und τ . Das maximale Verhältnis τ_{\max}/τ_{\min} wird erhöht, bis man zu einem reinen Phasen-Code gelangt.

In Abbildung 36 ist der Unterschied zwischen der optimalen Detektionswahrscheinlichkeit und der Unterscheidung nur auf Basis von τ in Abhängigkeit von τ_{\max}/τ_{\min} zu erkennen.

Die Differenz nimmt erwartungsgemäß ab, bis die Kurven schließlich ineinander übergehen. Der andere Extremfall, nur auf Basis von λ zu unterscheiden, würde einer horizontalen Linie zu Beginn der optimalen Detektionswahrscheinlichkeit entsprechen.

Für die optimalen Wahlen von τ gibt es mehrere Sprungmöglichkeiten:

Grundsätzlich lässt sich auch hier festhalten: Für kleines τ_{\max}/τ_{\min} wird für jedes τ entweder τ_{\max} oder τ_{\min} , hier o.B.d.A $\tau_{\min} = 1$, gewählt. Für kleines τ_{\max}/τ_{\min} kommen dabei folgende Kombinationen vor

$$(\lambda_1, \lambda_2, \lambda_3, \tau_{\max}, 1, \tau_{\max}) \quad \text{oder} \quad (\lambda_1, \lambda_2, \lambda_3, 1, \tau_{\max}, 1),$$

wobei die erste Variante im Fall $\lambda_1 = 12$ und die zweite Variante für $\lambda_1 = 6$ auftritt. Anzumerken ist aber, dass sich diese beiden Möglichkeiten nur um maximale 0.01 unterscheiden. Diese Kombinationen erscheinen auch als sehr intuitiv, da so benachbarte Stimuli in der ersten Stufe, die also schwerer auseinanderzuhalten sind, in der zweiten Stufe gerade die stärkere Entscheidungsmacht erhalten. Es ist nicht auszuschließen, dass für anderes λ_1 beide Kombinationen als optimal auftreten.

Springt λ_2 auch auf das maximale λ ergibt sich zuerst folgende Kombination

$$(\lambda_1, \lambda_1, \lambda_3, \tau_{\max}, 1, 1).$$

Für wachsendes τ_{\max}/τ_{\min} treten für größere λ_1 , vergleiche $\lambda_1 = 12$, noch die beiden Fälle

$$(\lambda_1, \lambda_1, \lambda_3, \tau_{\max}, 1, \tau_3) \quad \text{und} \quad (\lambda_1, \lambda_1, \lambda_3, \tau_{\max}, \tau_2, 1)$$

auf. Für sehr großes τ_{\max}/τ_{\min} , etwas mehr als doppelt so groß wie λ_1 , ergibt sich dann die vertraute Parameterwahl

$$(\lambda_1, \lambda_1, \lambda_1, \tau_{\max}, \sqrt{\tau_{\max}}, 1).$$

Auffallend ist, dass größtenteils, außer für sehr kleine τ_{\max}/τ_{\min} bei $\lambda_1 = 6$, dem maximalen λ stets das maximale τ zugeordnet wird, vgl. Abbildung 36 (b). Für kleines λ_1 treten nur wenige der aufgeführten Kombinationsmöglichkeiten auf, da hier die Unterscheidung nur auf Basis von τ für relativ kleines τ_{\max}/τ_{\min} bereits bevorzugt wird. Dort ist ein zwischenzeitliches Unterscheiden (noch nicht alle λ gleich) aller Stimuli auf Basis von τ nicht optimal.

Wie verhält sich die bedingte Detektionswahrscheinlichkeit? Bedingt unter dem Ereignis, dass Stimulus i zum Zuge gekommen ist, mit welcher Wahrscheinlichkeit detektiert man auch diesen mit den gewählten Entscheidungsschranken. Dies ist in Abbildung 38 für $\lambda_1 = 6$ und $\lambda_1 = 12$ in Abhängigkeit von τ_{\max}/τ_{\min} dargestellt. Solange noch nicht alle λ maximal gewählt werden, ist kein einheitliches Muster erkennbar.

Wird nur noch auf Basis von τ unterschieden, lässt sich am besten der Stimulus mit

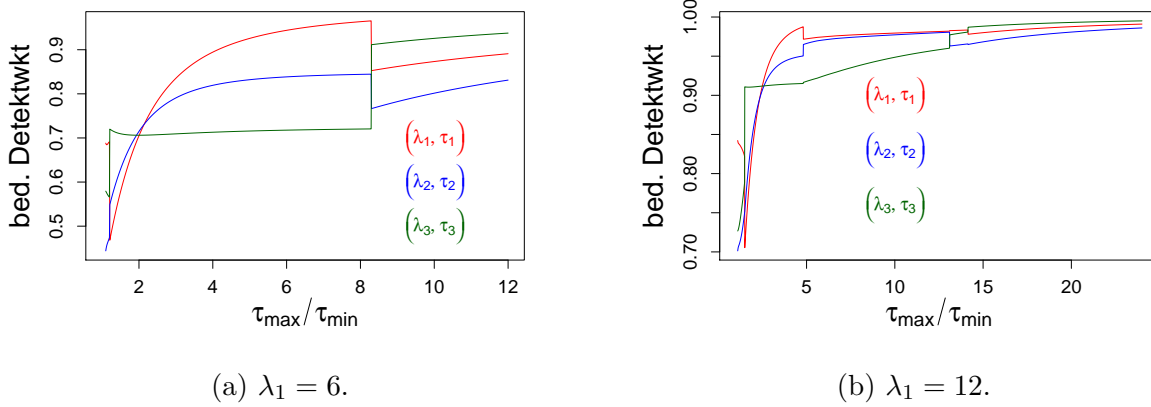


Abbildung 38: Bedingte Detektionswahrscheinlichkeit $\mathbb{P}_{\text{Stimulus } j}(\text{„Ent. für Stimulus } j\text{“})$ für $\lambda_1 = 6$ und $\lambda_1 = 12$ in Abhängigkeit von τ_{\max}/τ_{\min} . Für λ und τ wurden die als optimal berechneten Parameter gewählt.

dem minimalen τ unterscheiden, dann der mit dem maximalen τ und am schlechtesten der Stimulus mit dem mittleren τ . Dies war auch zu erwarten, da die Exponentialverteilung zu Beginn die höchste Dichte besitzt und die Entscheidungsschranken nur für das minimale τ den Anfang berücksichtigen (vgl. Abschnitt 2.4). Beim mittleren τ lässt sich in beiden Enden der Stimulus nicht detektieren und besitzt somit auch die geringste bedingte Detektionswahrscheinlichkeit.

3.4 Vier Stimuli

Für vier Stimuli erhält man sehr ähnliche Ergebnisse, nur ergeben sich hier zusätzliche Kombinationsmöglichkeiten. In Abbildung 39 ist die optimale Parameterwahl wieder für $\lambda_1 = 6$ und $\lambda_1 = 12$ in Abhängigkeit von τ_{\max}/τ_{\min} dargestellt. In Abbildung 40 sind zur besseren Übersicht die vorkommenden Kombinationen nochmals aufgeführt.

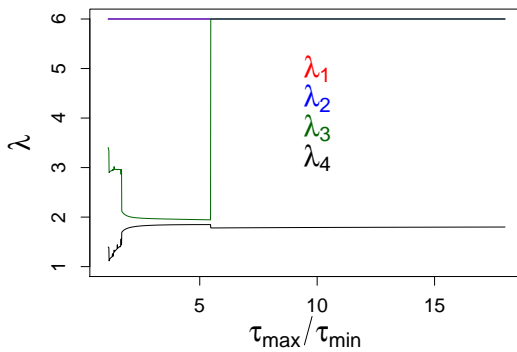
Für sehr kleines τ_{\max}/τ_{\min} werden alle λ unterschiedlich gewählt und die beiden Kombinationen

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \tau_{\max}, 1, 1, \tau_{\max}) \quad \text{und} \quad (\lambda_1, \lambda_2, \lambda_3, \lambda_4, 1, \tau_{\max}, \tau_{\max}, 1)$$

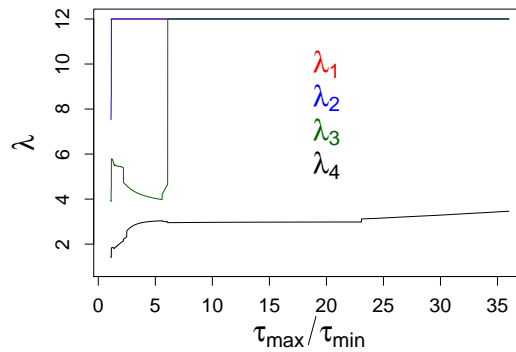
stehen im ständigen Wechsel. Steigt τ_{\max}/τ_{\min} nur leicht an, ist es von Vorteil, zwei Stimuli nur noch auf Basis von τ zu unterscheiden. Hierbei wechseln sich die Kombinationen

$$(\lambda_1, \lambda_1, \lambda_3, \lambda_4, \tau_{\max}, 1, 1, \tau_{\max}) \quad \text{und} \quad (\lambda_1, \lambda_1, \lambda_3, \lambda_4, \tau_{\max}, 1, \tau_{\max}, 1)$$

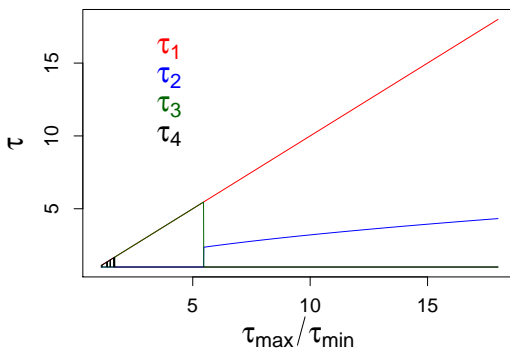
ab. Die Kombinationen auf Seiten von τ bleiben dabei unverändert. Die zweite Kombination tritt schließlich für größeres τ_{\max}/τ_{\min} (ungefähr ab 2) ausschließlich auf. Für $\lambda_1 = 6$ sind diese beiden Kombinationen neben der Unterscheidung nur auf Basis von τ die einzig vorkommenden Kombinationen.



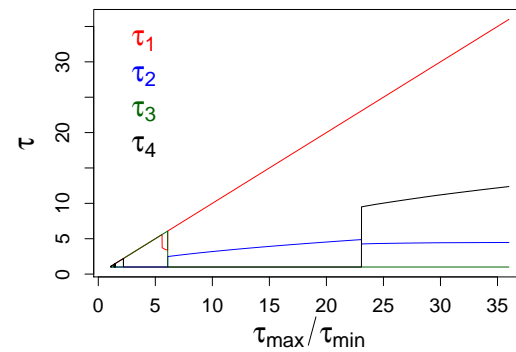
(a) Optimale Raten $\lambda_2, \lambda_3, \lambda_4$ für $\lambda_1 = 6$.



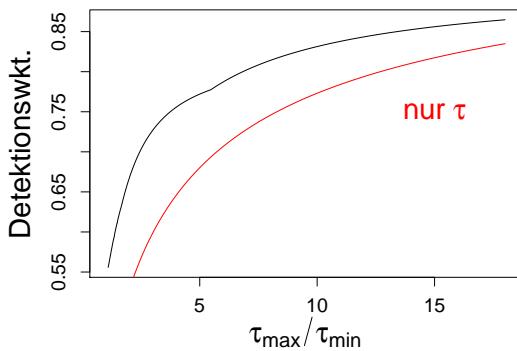
(d) Optimale Raten $\lambda_2, \lambda_3, \lambda_4$ für $\lambda_1 = 12$.



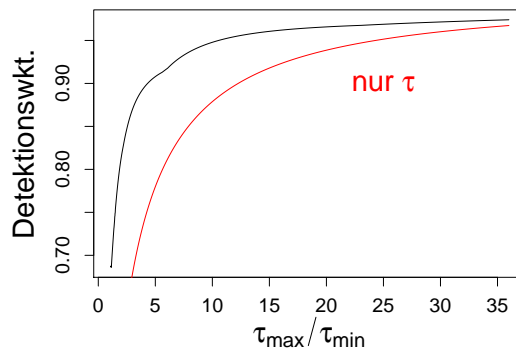
(b) Optimale Phasen $\tau_1, \tau_2, \tau_3, \tau_4$ für $\lambda_1 = 6$.



(e) Optimale Phasen $\tau_1, \tau_2, \tau_3, \tau_4$ für $\lambda_1 = 12$.



(c) Maximale Detektionswkt. für $\lambda_1 = 6$ im Vergleich zu einem reinen Phasen-Code.



(f) Maximale Detektionswkt. für $\lambda_1 = 12$ im Vergleich zu einem reinen Phasen-Code.

Abbildung 39: Optimale Parameterwahl und optimale Detektionswahrscheinlichkeit für 4 Stimuli für $\lambda_1 = 6$ und $\lambda_1 = 12$ in Abhängigkeit von τ_{\max}/τ_{\min} .

Steigt τ_{\max}/τ_{\min} an, bleiben weiterhin genau zwei λ maximal gewählt, doch ist es dann von

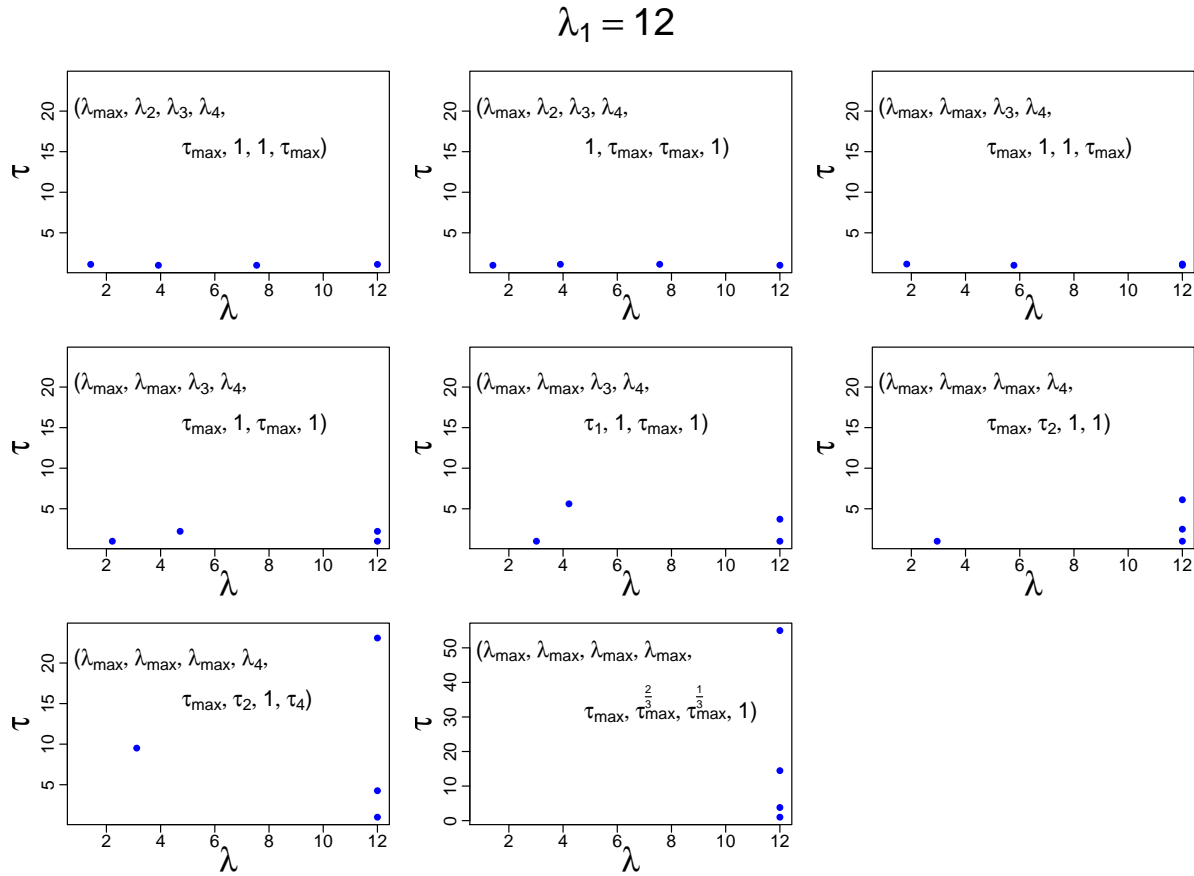


Abbildung 40: Auftretende Kombinationen für 4 Stimuli für $\lambda_1 = 12$. Die Koordinaten der blauen Punkte entsprechen den Werten von λ und τ . Das maximale Verhältnis τ_{\max}/τ_{\min} wird erhöht, bis man zu einem reinen Phasen-Code gelangt.

Vorteil, auf Basis von τ spezifischer zu unterscheiden. Optimal ist dabei die Anordnung

$$(\lambda_1, \lambda_1, \lambda_3, \lambda_4, \tau_1, 1, \tau_{\max}, 1).$$

Diese Anordnung ist untypisch, da hier das maximale λ nicht das maximale τ erhält. Sinnvoll erscheint die Wahl aber dennoch, da sich für $\lambda_1 = 12$ die ersten beiden Stimuli auf Basis von τ schon relativ gut unterscheiden lassen, und dort nicht so einen großen Phasenunterschied benötigen. λ_3 und λ_4 liegen ziemlich dicht beieinander und besitzen eine deutlich niedrigere Rate und benötigen somit einen größeren Unterschied auf Basis von τ , um sich gut auseinanderhalten zu lassen.

Steigt τ_{\max}/τ_{\min} weiter an, ist es schließlich von Vorteil drei Stimuli nur noch auf Basis von τ zu unterscheiden. Hierbei erhält das minimale λ auch das minimale τ .

$$(\lambda_1, \lambda_1, \lambda_1, \lambda_4, \tau_{\max}, \tau_2, 1, 1).$$

Diese Wahl dominiert über eine weite Spanne alle anderen Kombinationen. Erst für einen Wert von τ_{\max}/τ_{\min} , der mehr als viermal so groß wie λ_1 ist, wird nur noch auf Basis von τ unterschieden

$$\left(\lambda_1, \lambda_1, \lambda_1, \lambda_1, \tau_{\max}, \tau_{\max}^{2/3}, \tau_{\max}^{1/3}, 1\right).$$

3.5 Sieben Stimuli

Für sieben Stimuli sind die optimalen Parameterkombinationen für $\lambda_1 = 8$ und τ_{\max}/τ_{\min} zwischen 2 und 18 bestimmt worden. In Abbildung 41 sind die vorkommenden Kombinationen dargestellt und in Abbildung 42 die optimale Detektionswahrscheinlichkeit.

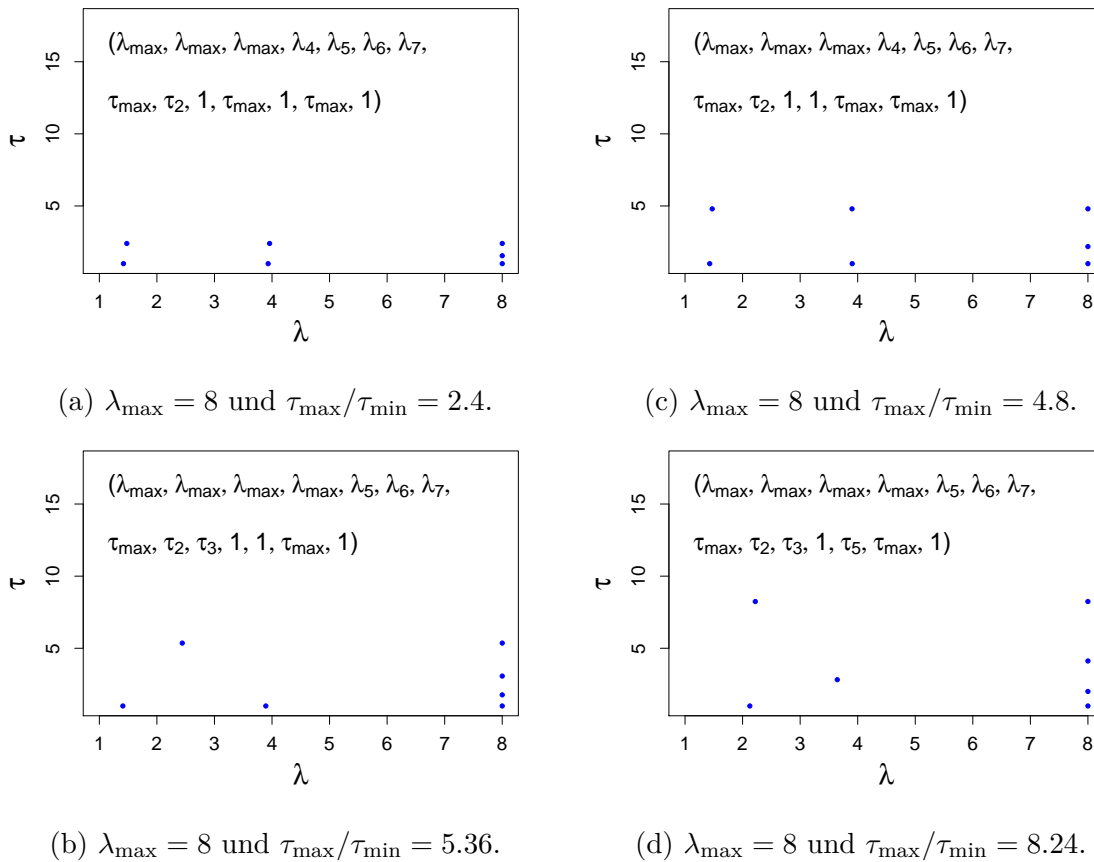


Abbildung 41: Optimale Kombinationen für 7 Stimuli für $\lambda_1 = 8$ und für τ_{\max}/τ_{\min} im Bereich zwischen 2 und 18. Für τ_{\max}/τ_{\min} wurde jeweils der Wert gewählt, ab dem die Kombination dominierend auftritt. Die Koordinaten der blauen Punkten entsprechen den Werten von λ und τ .

In dem untersuchten Parameterbereich beginnt die optimale Kombination bereits mit drei identischen Raten. Bei einer hohen Anzahl Stimuli ist also der Einfluss von τ schon für

niedriges τ_{\max}/τ_{\min} deutlich zu sehen. Zu Beginn, vgl. Abbildung 41 (a), dominiert die Kombination

$$(\lambda_1, \lambda_1, \lambda_1, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \tau_{\max}, \tau_2, 1, \tau_{\max}, 1, \tau_{\max}, 1).$$

Ab etwa $\tau_{\max}/\tau_{\min} = 3$ tritt zusätzlich die Kombination

$$(\lambda_1, \lambda_1, \lambda_1, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \tau_{\max}, \tau_2, 1, 1, \tau_{\max}, \tau_{\max}, 1)$$

auf, siehe Abbildung 41 (c). Zu Beginn dominiert jedoch noch die Startkombination, aber ab etwa $\tau_{\max}/\tau_{\min} = 4.8$ tritt die zweite Kombination überwiegend auf. Ist es von Vorteil, insgesamt vier Stimuli das maximale λ zuzuordnen, tritt ab etwa $\tau_{\max}/\tau_{\min} = 5.3$ die Kombination

$$(\lambda_1, \lambda_1, \lambda_1, \lambda_1, \lambda_5, \lambda_6, \lambda_7, \tau_{\max}, \tau_2, \tau_3, 1, 1, \tau_{\max}, 1)$$

auf, vgl. Abbildung 41 (b). Diese Kombination entspricht dabei den bisherigen Ergebnissen: Die Stimuli, die das maximale λ erhalten haben, werden auf Basis von τ mit der komplette Spanne τ_{\max}/τ_{\min} unterschieden. Für die drei übrigen Stimuli bleiben die Möglichkeiten bestehen, ob das maximale oder minimale τ gewählt wird. Ab etwa $\tau_{\max}/\tau_{\min} = 8.24$ überwiegt die Anordnung

$$(\lambda_1, \lambda_1, \lambda_1, \lambda_1, \lambda_5, \lambda_6, \lambda_7, \tau_{\max}, \tau_2, \tau_3, 1, \tau_5, \tau_{\max}, 1),$$

bei der nun ein spezifischeres Unterscheiden der drei übrigen Stimuli möglich ist. Dies ist die erste Kombination, bei dem einem mittleren λ auch ein mittleres τ zugeordnet wird, ein Stimulus also in die Mitte der anderen Stimuli gesetzt wird, siehe Abbildung 41 (d). Für eine größere Anzahl an Stimuli erwartet man mehr Stimuli im inneren Bereich. Dies wird im nächsten Abschnitt beim Einführen einer Nachbarschaftsbeziehung genauer untersucht. Ab etwa $\tau_{\max}/\tau_{\min} = 10$ tritt diese ausschließlich auf. Erst für τ_{\max}/τ_{\min} außerhalb des betrachteten Parameterbereichs werden schließlich fünf Stimuli nur noch auf Basis von τ unterschieden

$$(\lambda_1, \lambda_1, \lambda_1, \lambda_1, \lambda_1, \lambda_6, \lambda_7, \tau_{\max}, \tau_2, \tau_3, \tau_4, 1, \tau_{\max}, 1).$$

Diese Kombination besitzt aber auch für τ_{\max}/τ_{\min} am Ende des untersuchten Parameterbereichs nur eine um weniger als 0.5% reduzierte Detektionswahrscheinlichkeit.

Bemerkung 3.6. *In Abbildung 42 lässt sich - wie zu erwarten - erkennen, dass für eine größere Anzahl Stimuli das Detektieren durch geschicktes Kombinieren von λ und τ deutlich besser gegenüber eines reine Raten- oder Phasen-codes möglich ist.*

Für eine größere Anzahl Stimuli ist eine numerische Berechnung auf Grund des hochdimensionalen Problems nicht mehr realisierbar. Ein möglicher Ansatz ist, durch stochastische

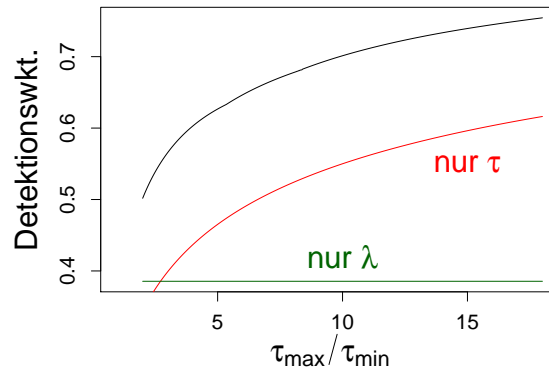


Abbildung 42: Maximale Detektionswahrscheinlichkeit für 7 Stimuli für $\lambda_1 = 8$ in Abhängigkeit von τ_{\max}/τ_{\min} . In rot (grün) ist die Detektionswahrscheinlichkeit dargestellt, die man durch einen reinen Phasen-Code (Raten-Code) erhält.

Optimierung besonders gute Anordnungen zu suchen. Bereits bei einer geringen Anzahl Stimuli ist zu beobachten, dass sich unterschiedliche Kombinationen über einen weiten τ_{\max}/τ_{\min} -Bereich nur geringfügig in der Höhe der Detektionswahrscheinlichkeit unterscheiden.

Die Frage nach der besten Lösung erscheint dabei auch etwas künstlich: Wie soll das Gehirn die optimale Wahl gefunden haben? Hier ist ein lernender Ansatz einleuchtender: Eine Ausgangsjustierung wird ständig durch Einflüsse aus der Umwelt weiter optimiert; der Grundgedanke, der auch der stochastischen Optimierung zugrunde liegt.

4 Einführen von Ähnlichkeitsbeziehungen

Innerhalb des ELO-Modells spielt die Anordnung der Stimuli keine Rolle und bisher war es das Ziel, die Detektionswahrscheinlichkeit global zu maximieren. Dabei werden jedoch keine strukturellen Beziehungen berücksichtigt.

Als Beispiel sei die Untersuchung mit Balken in unterschiedlichen Orientierungen genannt. Dort reagiert ein Neuron auf ähnliche Orientierungen auch mit einer ähnlichen Spikefolge. Um Ähnlichkeiten im ELO-Modell zu berücksichtigen, wird eine Fehlerfunktion eingeführt, die Fehlentscheidungen bei benachbarten Stimuli schwächer gewichtet.

Im Geist der Balkenorientierung bedeutet dies, dass Fehlentscheidungen bezüglich ähnlichen Orientierungen als nicht so gravierend angesehen werden, jedoch das Auswählen einer komplett anderen Orientierung stark bestraft wird. Ziel ist, den Gesamtfehler zu minimieren, dass man also mit großer Wahrscheinlichkeit in „etwa“ richtig liegt.

Das Einführen der Fehlerfunktion hebt den Rechenaufwand jedoch deutlich an: Nun reicht es nicht mehr aus, für jeden Stimulus nur die Wahrscheinlichkeit, sich für diesen korrekterweise zu entscheiden, zu bestimmen, sondern es muss jegliche bedingte Fehlentscheidung ermittelt werden. Damit bietet sich hier insbesondere stochastische Optimierung an.

Diese wird zuerst an die bekannten Ergebnisse aus der globalen Maximierung justiert. Anschließend wird diese auf die Fehlerfunktion angewendet und der Einfluss von unterschiedlichen Gewichtungen analysiert.

4.1 Konstruktion Fehlerfunktion

Gegeben sind k Stimuli S_1, \dots, S_k mit der Vorgabe $\lambda_i \leq \lambda_{\max}$ für alle $i \leq k$ und mit $\tau_i/\tau_j \leq c$ für alle $i, j \leq k$.

Nun sind die Stimuli nicht mehr gemäß λ geordnet, sondern ihre Anordnung untereinander ist relevant. Dabei werden die Stimuli gleichmäßig am Kreis angeordnet, sodass Stimulus 1 und k direkte Nachbarn sind, vergleiche Abbildung 43.

Definition 4.1. *Einer Parameterwahl mit*

$$(\boldsymbol{\lambda}, \boldsymbol{\tau}) = ((\lambda_1, \tau_1), (\lambda_2, \tau_2), \dots, (\lambda_k, \tau_k))$$

$\lambda_1, \dots, \lambda_k > 0$ wird der Fehler gemäß Entscheidungsregel aus Lemma 2.1

$$F((\boldsymbol{\lambda}, \boldsymbol{\tau})) := \sum_{i=1}^k \sum_{j=1}^k f_k(i, j) \cdot \mathbb{P}(\{E = j\} | \{S = i\}) + \sum_{i=1}^k \mathbb{P}(\{E = 0\} | \{S = i\})$$

zugeordnet, für eine Gewichtung $f_k(\cdot)$

$$f_k : (\mathbb{Z}, \mathbb{Z}) \mapsto [0, 1].$$

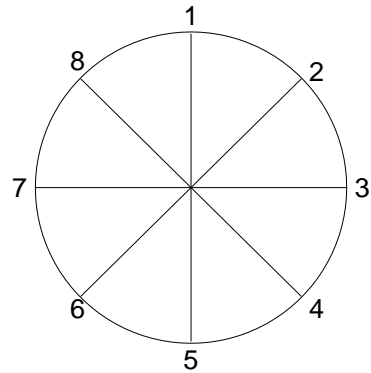


Abbildung 43: Anordnung der Stimuli im Fall $k = 8$. Je mehr Nachbarn zwischen zwei Stimuli liegen, desto schwächer fällt ihre Ähnlichkeitsbeziehung aus und desto höher ist die Gewichtung $f_k(i, j)$ der bedingten Fehlentscheidung von Stimulus i und j .

Im Fall einer linearen Gewichtung besitzt $f_k(\cdot)$ die Vorschrift

$$f_k(i, j) := \begin{cases} \frac{2 \cdot |i-j|}{k} & \text{für } |i-j| \leq \frac{k}{2}, \\ 2 - \frac{2 \cdot |i-j|}{k} & \text{sonst.} \end{cases}$$

Auf diese Weise wird eine Fehlentscheidung für zwei direkt benachbarte Stimuli im Fall von 8 Stimuli mit $1/4$ gewichtet, Stimuli mit dem Abstand 2 mit $1/2$, Stimuli mit dem Abstand 3 mit $3/4$ und gegenüberliegende Stimuli mit 1. Das fälschlicherweise Auswählen des Nullstimulus wird für alle Stimuli mit maximaler Gewichtung von 1 bestraft.

Mit Bezug auf Satz 2.3 lässt sich die bedingte Wahrscheinlichkeit folgendermaßen bestimmen.

Korollar 4.1. *Bezeichne $k \in \mathbb{N}$ die Zahl an Stimuli. Gegeben Ratenparameter $\lambda_1, \dots, \lambda_k > 0$ und Phasenparameter $\tau_1, \dots, \tau_k > 0$ ist $T_j := \{i \in \{1, \dots, k\} \setminus \{j\} \mid \tau_i = \tau_j \wedge \lambda_i \neq \lambda_j\}$ und $R_j := \{i \in \{1, \dots, k\} \setminus \{j\} \mid \tau_i = \tau_j \wedge \lambda_i = \lambda_j\}$. Die bedingte Fehlentscheidung lässt sich bestimmen über*

$$\mathbb{P}(\{E = j\} \mid \{S = i\}) = \frac{1}{|R_j|} \sum_{n=[l]}^{[u]} \frac{\lambda_i}{n!} e^{-\lambda_i} \cdot \mathbb{P}_i \left(\max_{\tau_r < \tau_j} \left(\frac{\lambda_j - \lambda_r - n \log \frac{\lambda_j \tau_r}{\lambda_r \tau_j}}{\frac{1}{\tau_r} - \frac{1}{\tau_j}} \right) < \sum_{k=1}^n X_k < \min_{\tau_r > \tau_j} \left(\frac{\lambda_j - \lambda_r - n \log \frac{\lambda_j \tau_r}{\lambda_r \tau_j}}{\frac{1}{\tau_r} - \frac{1}{\tau_j}} \right) \right).$$

für

$$l := \begin{cases} \max_{\lambda_r^{T_j} < \lambda_1} \left(\frac{\lambda_1 - \lambda_r^{T_j}}{\log \lambda_1 / \lambda_r^{T_j}} \right) & \text{falls } |\{\lambda_r^{T_j} < \lambda_1\}| > 0, \\ 1 & \text{sonst} \end{cases}$$

und

$$u := \begin{cases} \min_{\lambda_r^{T_j} > \lambda_1} \left(\frac{\lambda_1 - \lambda_r^{T_j}}{\log \lambda_1 / \lambda_r^{T_j}} \right) & \text{falls } |\{\lambda_r^{T_j} > \lambda_1\}| > 0, \\ \infty & \text{sonst.} \end{cases}$$

Für den Nullstimulus gilt

$$\mathbb{P}(\{E = j\} \mid \{S = 0\}) = 0 \quad \text{für } j > 0,$$

falls der Nullstimulus erscheint, wird dieser auch korrekterweise immer gewählt. Fälschlicherweise entscheidet man sich für den Nullstimulus mit Wahrscheinlichkeit

$$\mathbb{P}(\{E = 0\} \mid \{S = i\}) = \mathbb{P}(N_i = 0) = e^{-\lambda_i}.$$

Nun ist die Parameterwahl $(\boldsymbol{\lambda}, \boldsymbol{\tau})$ gesucht, sodass der Gesamtfehler $F((\boldsymbol{\lambda}, \boldsymbol{\tau}))$ minimiert wird. In anderen Worten wird die Wahl von k Punkten in der λ - τ -Ebene gesucht, sodass die Gesamtdistanz $F((\boldsymbol{\lambda}, \boldsymbol{\tau}))$ minimal ist.

4.1.1 Bedingte Fehlentscheidung

Um ein erstes Gefühl für $F((\boldsymbol{\lambda}, \boldsymbol{\tau}))$ zu bekommen, wird die bedingte Fehlentscheidung

$$d(i, j) := \mathbb{P}(\{E = j\} \mid \{S = i\}) + \mathbb{P}(\{E = i\} \mid \{S = j\})$$

zweier Stimuli betrachtet. In Abbildung 44 ist diese mit Hilfe von Isolinien dargestellt. Ist ein λ klein, nimmt die bedingte Fehlentscheidung besonders schnell durch ein noch kleineres λ ab, vgl. Abbildung 44 (a). Hierbei ist aber nicht die Fehlentscheidung zum Nullstimulus berücksichtigt. τ besitzt bei so geringem λ noch wenig Verbesserungsmöglichkeit.

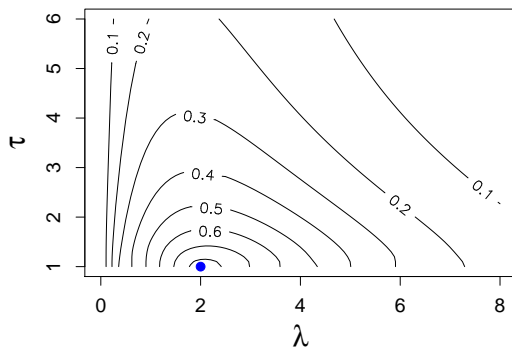
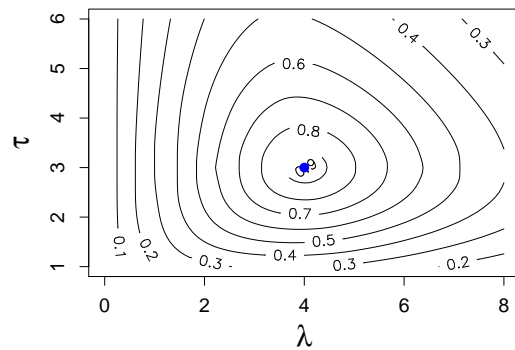
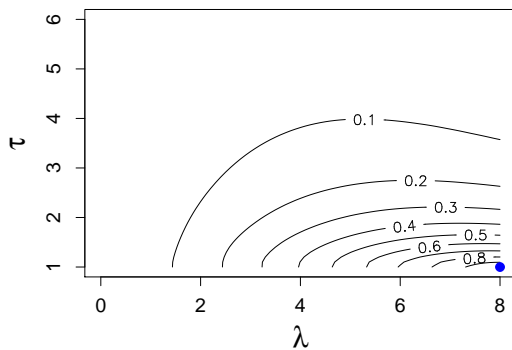
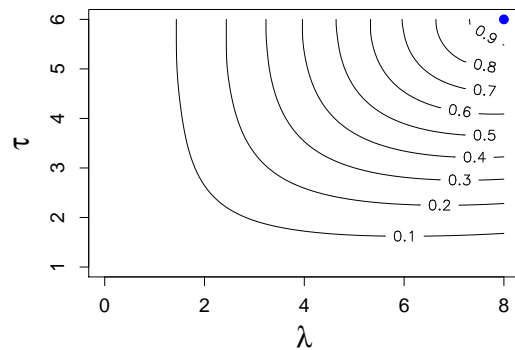
(a) $\lambda_{fix} = 2$ und $\tau_{fix} = 1$.(c) $\lambda_{fix} = 4$ und $\tau_{fix} = 3$.(b) $\lambda_{fix} = 8$ und $\tau_{fix} = 1$.(d) $\lambda_{fix} = 8$ und $\tau_{fix} = 6$.

Abbildung 44: Bedingte Fehlentscheidung zweier Stimuli im Konturdiagramm. Die Parameter eines Stimulus werden festgehalten (blauer Punkt). Die Parameter des anderen Stimulus werden im Bereich $\lambda = [0, 8]$ und $\tau = [1, 6]$ variiert und die bedingte Fehlentscheidung $d(i, j)$ durch Isolinien dargestellt.

Für ein mittleres λ ist das Unterscheiden generell schwieriger. Jedoch ist dort die Gewinnmöglichkeit durch die Phase deutlich zu sehen, siehe Abbildung 44 (c). Für großes λ funktioniert das Unterscheiden am besten nur noch auf Basis der Phase. Dabei verringert sich die Fehlentscheidung besonders schnell für eine kleine Phase, vgl. Abbildung 44 (b) und (d).

Bemerkung 4.1. *Will man viele Punkte in der λ - τ -Ebene platzieren, erwartet man, dass viele Punkte bei maximaler Rate liegen, insbesondere im unterem Phasenbereich. Eng aneinander werden die Punkte ebenso bei kleineren Raten liegen. Falls Punkte in der Mitte liegen, dann werden diese eher im unteren Phasenbereich platziert sein.*

Diese Ergebnisse leuchten auch ein, da für große Raten die mittlere Phase geringeren Schwankungen ausgesetzt ist. Weiterhin ist die Varianz sowohl für kleine Raten als auch Phasen geringer, weshalb sich in den unteren Bereichen mehr Punkte gut trennen lassen.

Da hierbei keine Gewichtung berücksichtigt wurde, decken sich die Ergebnisse mit dem Maximieren der Detektionswahrscheinlichkeit. Um die Auswirkung der Gewichtung zu analysieren, wird zuerst stochastische Optimierung an bereits bekannten Ergebnissen justiert und bewertet.

4.2 Stochastische Optimierung

Bezeichne k wieder die Anzahl Stimuli bzw. die Anzahl Punkte in der λ - τ -Ebene. Simuliert wird nun eine Irrfahrt, bei der nur Werte angenommen werden, die eine Verbesserung aufweisen. Das Vorgehen ist dabei wie folgt:

Wähle unabhängig k -Punkte uniform verteilt auf $[0, \lambda_{\max}] \times [\tau_{\min}, \tau_{\max}]$. Starte in

$$(\boldsymbol{\lambda}, \boldsymbol{\tau}) = ((\lambda_1, \tau_1), (\lambda_2, \tau_2), \dots, (\lambda_k, \tau_k)).$$

Ziehe nun unabhängig Kopien

$$\begin{aligned} N_1^\lambda, N_2^\lambda, \dots, N_k^\lambda &\sim \mathcal{N}(0, \sigma_\lambda^2) \\ \text{und} \quad N_1^\tau, N_2^\tau, \dots, N_k^\tau &\sim \mathcal{N}(0, \sigma_\tau^2) \end{aligned}$$

Die neue Punktconfiguration entsteht durch

$$(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\tau}}) = ((\lambda_1 + N_1^\lambda, \tau_1 + N_1^\tau), \dots, (\lambda_k + N_k^\lambda, \tau_k + N_k^\tau)),$$

wobei komponentenweise

$$\begin{aligned} \tilde{\lambda}_i &:= \min(\lambda_{\max}, \max(\tilde{\lambda}_i, 0)) \\ \text{und} \quad \tilde{\tau}_i &:= \min(\tau_{\max}, \max(\tilde{\tau}_i, \tau_{\min})) \end{aligned}$$

ersetzt wird. Setze $(\boldsymbol{\lambda}, \boldsymbol{\tau}) := (\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\tau}})$, falls

$$F((\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\tau}})) < F((\boldsymbol{\lambda}, \boldsymbol{\tau})).$$

Soll nicht die Fehlerfunktion minimiert werden, sondern die globale Detektionswahrscheinlichkeit maximiert werden, muss der letzte Schritt durch den Vergleich von p_D ersetzt werden.

Das zufällige Verändern der Punktconfiguration wird 1000-mal wiederholt.

Festzusetzen bleiben die Sprungvarianzen in λ - und τ -Richtung. Wendet man den Algorithmus auf die Detektionswahrscheinlichkeit an und simuliert für verschiedene Varianzen, lässt sich deren Auswirkung leicht erkennen. In Abbildung 45 wurde durch jeweils eine Simulation für 4 Stimuli die beste Anordnung gesucht, in (a) mit einer höheren Standardabweichung von 2, in (b) mit einer niedrigen von 0.1. Die optimale Lösung wurde numerisch bereits bestimmt und ist durch die rote Horizontale eingezeichnet.

Simuliert man mit einer hohen Varianz, verbessert sich die Detektionswahrscheinlichkeit

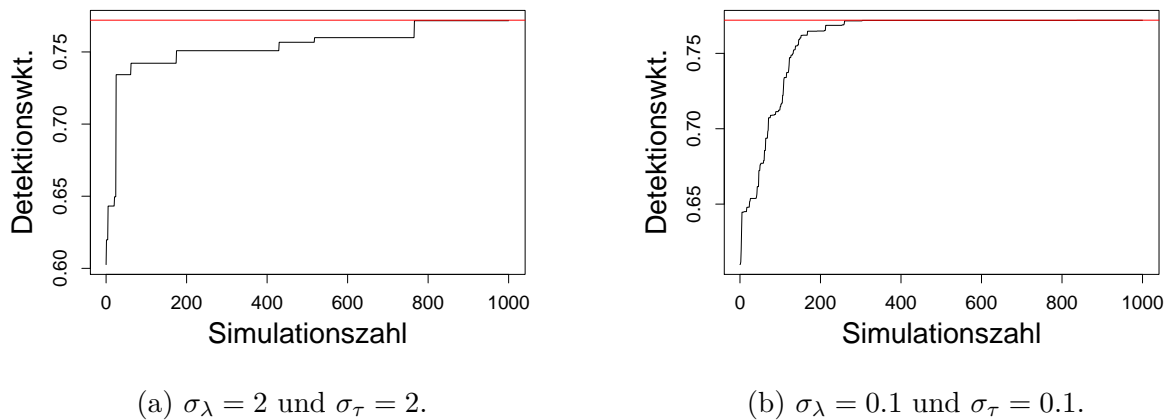


Abbildung 45: Verlauf der Detektionswahrscheinlichkeit einer Simulation für 4 Stimuli und eine feste Varianz. In rot ist die maximale Detektionswahrscheinlichkeit eingezeichnet

zu Beginn deutlich schneller. Man springt zwar in der Regel seltener, aber verbessert sich viel stärker. Befindet man sich aber erstmal in einem guten Zustand, dauert es immer länger, bis man in einen energetisch günstigeren Zustand springt. Dort kommt man mit einer geringeren Varianz schneller ans Ziel. Ein Nachteil geringer Varianz liegt im möglichen Verharren in einem lokalen Minimum, aus dem man nur mit größerer Varianz entkommen könnte.

Es liegt also nahe, mit einer hohen Varianz zu starten und diese im Laufe der Simulation immer weiter zu verringern. Somit nutzt man zu Beginn den positiven Effekt von starker Streuung, und gleichzeitig nähert man sich im fortgeschrittenen Stadium schneller der optimalen Lösung an.

Bemerkung 4.2. In Abbildung 46 wurden links für feste Varianz und rechts für variable Varianz jeweils 100-mal 1000 Simulationsläufe durchgeführt und es wurde jeweils die Abweichung von der bekannten optimalen Lösung bestimmt. Bei variabler Varianz wurde mit $\sigma = 2$ gestartet und jeweils nach 10 erfolglosen Verbesserungsversuchen um den Faktor 0.9 verringert.

Es ist deutlich zu sehen, dass man durch variable Varianz sowohl öfters sehr gute Lösungen findet als auch geringere Abweichungen zu verzeichnen hat.

Für 7 Stimuli funktioniert das stochastische Optimieren dementsprechend sehr überzeugend. Die Idee ist nun, diesen Algorithmus auf die Fehlerfunktion anzuwenden und bei mehrfachem Simulieren mit hoher Wahrscheinlichkeit eine Lösung zu finden, die dem Optimum sehr nahe kommt.

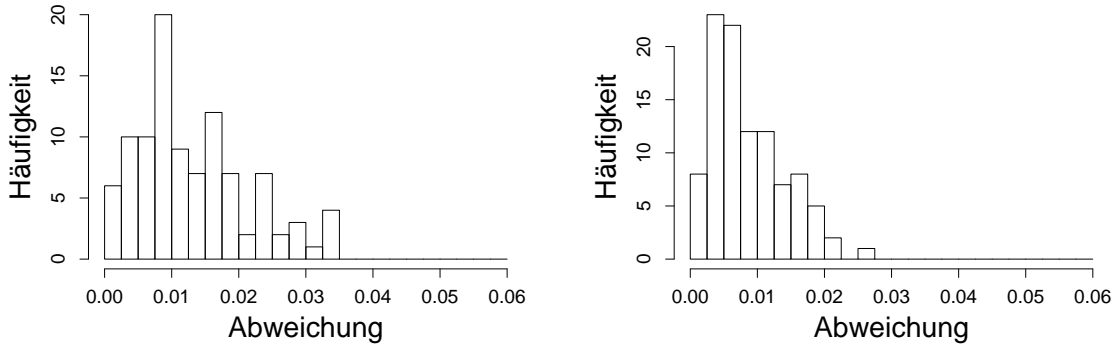
(a) Konstant $\sigma_\lambda = 0.1$ und $\sigma_\tau = 0.1$.(b) Variabel $\sigma_\lambda^{\text{start}} = 2$ und $\sigma_\tau^{\text{start}} = 2$.

Abbildung 46: Histogramm über die Abweichung der Detektionswahrscheinlichkeit vom Optimum für feste und variable Varianz für 7 Stimuli. Es wurden jeweils 100-mal 1000 Simulationsläufe durchgeführt. Bei variabler Varianz wurde bei 10-fachem Steckenbleiben des Algorithmus die Varianz um den Faktor 0.9 verringert.

4.3 Auswirkung Fehlerfunktion

Numerisch ließ sich das Optimum nur für bis zu 7 Stimuli bestimmen. Simuliert man mit dem beschriebenen Algorithmus 100-mal und wählt die Parameterkombination, die die maximale Detektionswahrscheinlichkeit besitzt, aus, lässt sich auch für 8 und 16 Stimuli eine Lösung bestimmen, die mit hoher Wahrscheinlichkeit sehr ähnlich dem wirklichen Optimum ist.

In Abbildung 47 sind die besten Lösungen dargestellt für $\lambda_{\max} = 8$ und $\tau_{\max}/\tau_{\min} = 6$. Die Lösung für 8 Stimuli ähnelt sehr dem bekannten Ergebnis für 7 Stimuli: Alle Stimuli

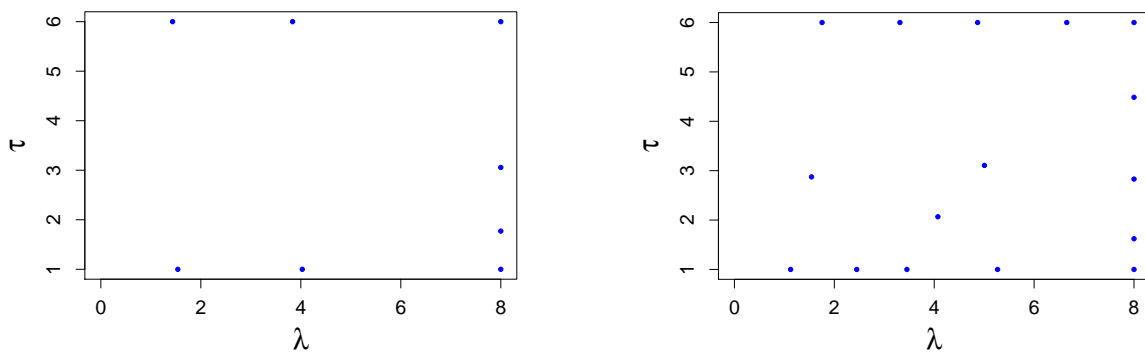
(a) Beste Simulation für 8 Stimuli mit $\lambda_{\max} = 8$ und $\tau_{\max}/\tau_{\min} = 6$.(b) Beste Simulation für 16 Stimuli mit $\lambda_{\max} = 8$ und $\tau_{\max}/\tau_{\min} = 6$.

Abbildung 47: Die Parameterkombination unter 100-Simulationen mit maximaler Detektionswahrscheinlichkeit für 8 und 16 Stimuli.

befinden sich auf dem äußeren Rand. Für 16 Stimuli ist jedoch auf dem Rand nicht mehr ausreichend Platz vorhanden und es wird vorteilhaft 3 Stimuli innerhalb des begrenzten „Rechtecks“ zu setzen.

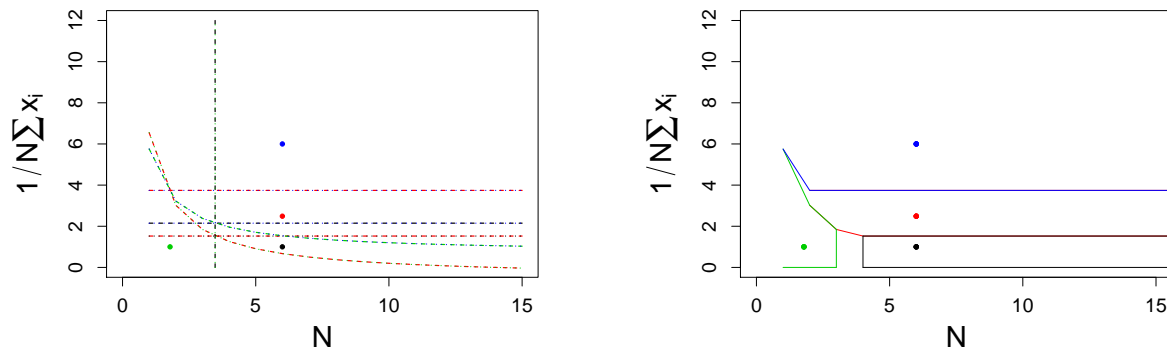
Die Stimuli mit einer mittleren Rate und Phase lassen sich von mehreren Stimuli nur schwer unterscheiden, da diese mehr Nachbarn mit ähnlichen Parametern besitzen. Dies wird beim Betrachten der Entscheidungsschranken klarer.

4.3.1 Analyse der Entscheidungsschranken

Um bei 2 Stimuli zu wissen, in welchem Bereich man sich für welchen Stimulus entscheidet, bestimmt man gemäß Bayesschen-Entscheidungsregel die Entscheidungsschranke, vergleiche Kapitel 2.2.1.

Will man bei mehreren Stimuli wissen, in welchem Bereich man sich für einen festen Stimulus entscheidet, müssen die Entscheidungsschranken zu all den anderen Stimuli bestimmt werden und das kleinste eingegrenzte Gebiet bestimmt werden.

In Abbildung 48 sind in (a) alle Entscheidungsschranken, farblich markiert mit den beteiligten Stimuli, eingezeichnet, in (b) die daraus resultierenden Gebiete. Zur besseren Übersicht



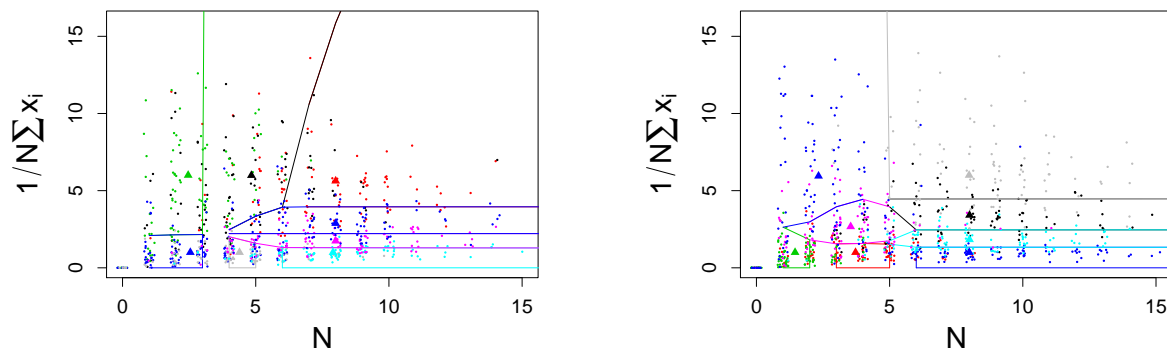
(a) Paarweise Entscheidungsschranken gemäß Lemma 2.1.

(b) Stimulus-spezifische Gebiete gemäß Satz 2.3.

Abbildung 48: Entscheidungsschranken und die daraus resultierenden Gebiete. Die Punkte entsprechen den gewählten Parametern in der λ - τ -Ebene. Die Entscheidungsschranken lassen sich farblich den beteiligten Stimuli zuordnen.

ist dies nur für 4 Stimuli dargestellt. Die Punkte repräsentieren die Wahl der Parameter der Stimuli.

Simuliert man nun Werte gemäß jedes Stimulus, lässt sich erkennen, in welchen Gebieten die simulierten Werte häufig korrekterweise fallen, und in welchen oft Fehlentscheidungen auftreten. In Abbildung 49 ist dies für 8 Stimuli links für eine Parameterkombination auf dem Rechteck und rechts für einen Stimulus in der Mitte umgesetzt.



(a) Alle Stimuli auf dem äußeren Rand.

(b) Ein Stimulus in der Mitte der anderen.

Abbildung 49: Vergleich Rechteckstruktur mit einer mittigen Platzierung. Die Dreiecke entsprechen den gewählten Parametern in der λ - τ -Ebene. Jeder Punkte entspricht einer Realisierung gemäß der Parameter der gewählten Farbe und ist zum besseren Unterscheiden leicht gejittert. Landet ein Punkt im Gebiet des gleichfarbigen Dreieckes, wird dieser korrekt detektiert.

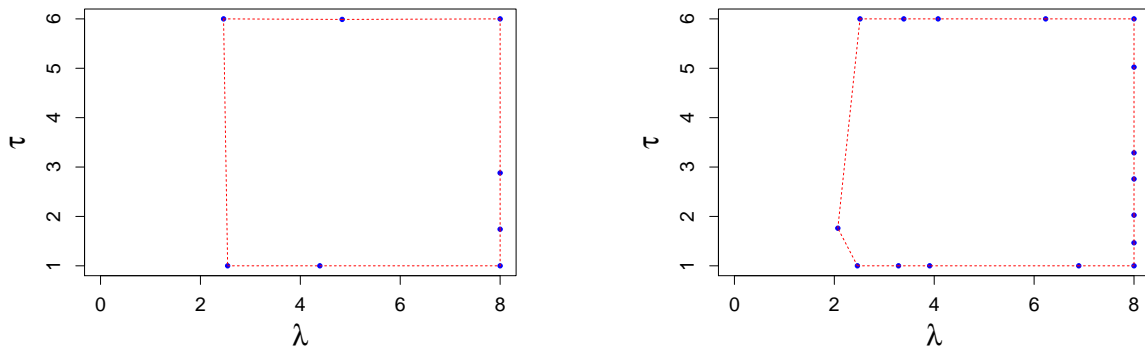
Schön lässt sich in dieser Grafik wieder erkennen, dass für niedriges λ und niedriges τ die Punkte nur sehr wenig streuen und sich damit die Stimuli gut unterscheiden lassen. Der mittlere Stimulus streut stark besonders in die Gebiete der linken 3 Stimuli und gleichzeitig streuen mehrere Stimuli Punkte in das mittlere Gebiet.

Für die globale Detektionswahrscheinlichkeit spielt das Streuen in viele verschiedene Gebiete keine Rolle, dort ist nur die Anzahl falscher Punkte entscheidend. Führt man jedoch Nachbarschaftsbeziehungen ein, wie z.B. durch eine lineare Fehlerfunktion, so werden die fehlerhaften Entscheidungen direkt benachbarter Stimuli nur sehr schwach gewichtet. Desto schwächer die Nachbarschaftsbeziehung ist, desto stärker fällt die Gewichtung aus. Will man den Effekt der Gewichtung gut ausnutzen, kommen Punkte in der Mitte nicht so in Frage, da sich dort auch oft für nicht direkt benachbarte Stimuli entschieden wird und sich somit ein höherer Gesamtfehler ergibt.

Bemerkung 4.3. Für die lineare Fehlerfunktion erwartet man also, dass auch für eine große Anzahl Stimuli die Punkte in der λ - τ -Ebene mehr auf dem äußeren Rechteck liegen und weniger oder keine Punkte in der Mitte Platz finden.

4.3.2 Lineare Fehlerfunktion

Nun wird der Algorithmus aus Kapitel 4.2 auf eine lineare Fehlerfunktion angewendet, bei der also die Fehlergewichtung für steigende Nachbarschaftsentfernung linear wächst. In Abbildung 50 ist die jeweils beste Simulation für 8 und 16 Stimuli in der λ - τ -Ebene dargestellt. Dabei wurde derselbe Parameterbereich wie bei der globalen Maximierung der Detektionswahrscheinlichkeit in Abbildung 47 gewählt.



(a) Beste Simulation für 8 Stimuli mit $\lambda_{max} = 8$ und $\tau_{max}/\tau_{min} = 6$.
 (b) Beste Simulation für 16 Stimuli mit $\lambda_{max} = 8$ und $\tau_{max}/\tau_{min} = 6$.

Abbildung 50: Die Parameterkombination unter 100-Simulationen mit minimalem linearem Fehler für 8 und 16 Stimuli. Die Punkte entsprechen den gewählten Parametern in der λ - τ -Ebene. Mit der rot gepunkteten Linie sind die benachbarten Stimuli verbunden.

Sowohl für 8 als auch 16 Stimuli fällt direkt auf, dass die Raten bei der Fehlerfunktion deutlich höher liegen. Grund dafür ist die starke Gewichtung des fehlerhaften Auswählens des Nullstimulus, der für jeden Stimulus maximal gewichtet wird.

Obwohl dadurch noch weniger Platz auf dem äußeren Rechteck ist, befinden sich auch für 16 Stimuli nun überhaupt keine Punkte mehr im inneren Parameterbereich. Auf diese Weise kann der Vorteil der Gewichtung maximal ausgenutzt werden. Die Punkte aus der Mitte werden insbesondere auf maximales λ gesetzt und es wird nur auf Basis von τ unterschieden, da dort im niedrigen τ -Bereich die Streuung sehr gering ausfällt und so weniger häufig für weiter entfernte Nachbarn entschieden wird.

Weiterhin fällt insbesondere bei 16 Stimuli auf, dass das Einführen von Nachbarschaftsbeziehungen Gruppchenbildung nach sich zieht. Paare oder dreier Gruppen liegen außerordentlich dicht zusammen, wobei der Abstand zu den nächsten Stimuli verhältnismäßig groß ist. Damit ver-

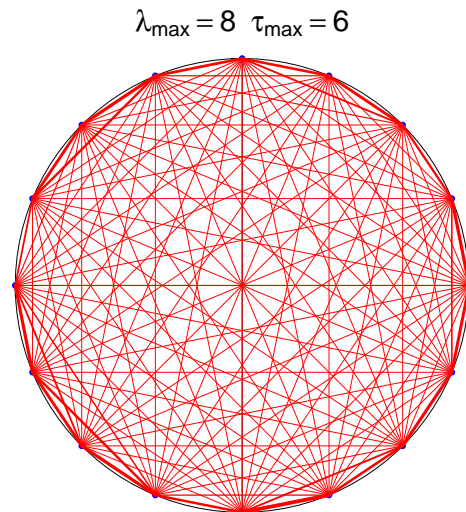


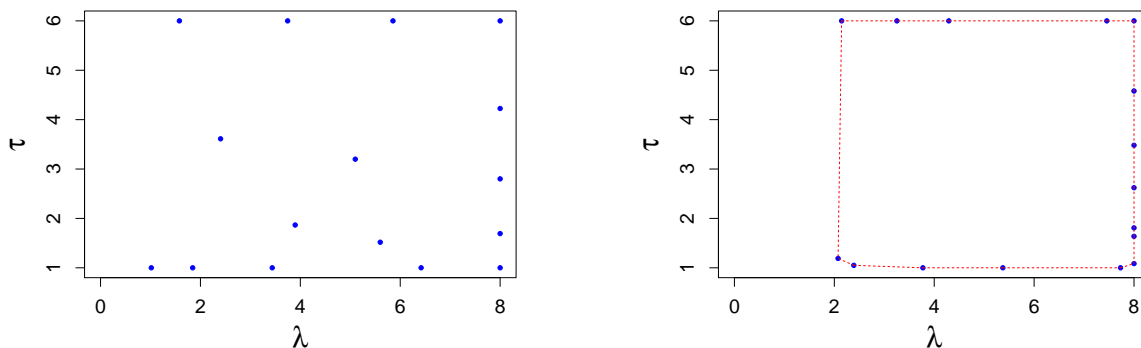
Abbildung 51: Bedingte Fehlentscheidung $d(i, j)$ für 16 Stimuli angedeutet anhand der Linienstärke.

ringert sich die Wahrscheinlichkeit, fälschlicherweise einen Stimulus zu wählen, der eine große Fehlergewichtung besitzt.

In Abbildung 51 sind 16 Stimuli am Kreis dargestellt, sodass die Nachbarschaftsentfernung deutlich zu erkennen ist. Alle Stimuli sind durch eine Linie miteinander verbunden, deren Stärke der Summe der jeweiligen Fehlentscheidungen entspricht. Die größten Fehlentscheidungen liegen auf dem Rand, gerade bei den direkten Nachbarn, bei denen die Fehler stark runter gewichtet werden. Vereinzelt treten noch dicke Linien zwischen Zweitnachbarn auf, welches auf die Dreier-Grüppchen zurückzuführen ist.

Das Ziel, die Fehlerfunktion zu minimieren, wirkt sich jedoch nur schwach auf die globale Detektionswahrscheinlichkeit aus: Beim globalen Maximieren der Detektionswahrscheinlichkeit hat sich für 16 Stimuli ein maximaler Wert von 34.1% (beste Simulation) ergeben. Würde man diese Parameterwahl so anordnen, dass der lineare Gesamtfehler minimal wird, so besitzt diese einen Fehler von etwa 4. Beim Minimieren der linearen Fehlerfunktion besitzt die beste Simulation eine Detektionswahrscheinlichkeit von 32.3% und einen Gesamtfehler von etwa 3.

Auch bei 16 Stimuli gibt es wieder viele mögliche Anordnungen, bei denen die Punkte sehr unterschiedlich platziert werden, die sich aber in Detektionswahrscheinlichkeit oder Gesamtfehler nur unwesentlich unterscheiden. In Abbildung 52 ist links eine Anordnung dargestellt, die dieselbe Detektionswahrscheinlichkeit bis in die 4-te Nachkommastelle auf-



(a) Anordnung mit vergleichbarer Detektionswkt. für 16 Stimuli mit $\lambda_{max} = 8$ und $\tau_{max}/\tau_{min} = 6$.

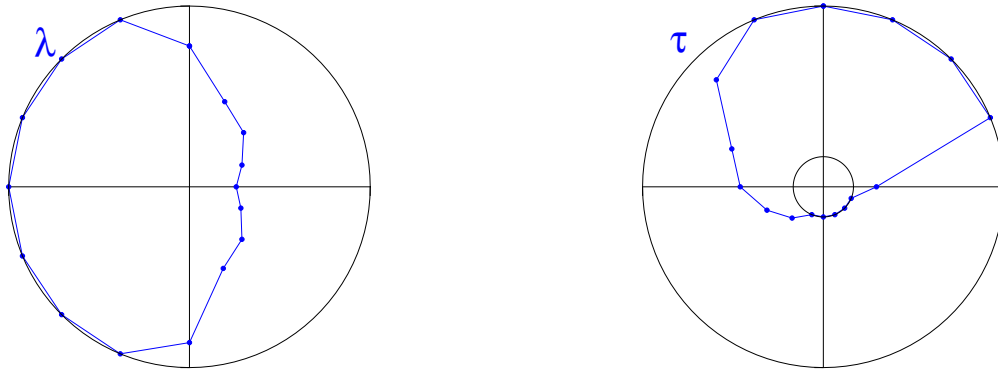
(b) Anordnung mit vergleichbarem linearem Gesamtfehler für 16 Stimuli mit $\lambda_{max} = 8$ und $\tau_{max}/\tau_{min} = 6$.

Abbildung 52: Anordnung mit vergleichbarer Detektionswahrscheinlichkeit (Abbildung 47 (b)) bzw. Gesamtfehler (Abbildung 50 (b)) zu dem simulierten Optimum.

weist, jedoch vier Punkte in der Mitte besitzt. Genauso ist rechts eine weitere Anordnung dargestellt, die denselben minimalen linearen Fehler besitzt.

In Abbildung 53 ist eine Lösung mit minimalen linearen Fehler als Tuning-Curve ein-

gezeichnet. Dabei entspricht der Abstand vom Ursprung der Größe des λ - bzw. τ -Wertes



(a) Anordnung der Raten am Einheitskreis. (b) Anordnung der Phasen am Einheitskreis.

Abbildung 53: Anordnung mit minimalen linearen Fehler als Tuning-Curve für 16 Stimuli und $\lambda_{\max} = 8$ und $\tau_{\max}/\tau_{\min} = 6$. Der Abstand vom Ursprung entspricht der Größe des λ - bzw. τ -Wertes im Verhältnis zum Maximalwert.

im Verhältnis zum Maximalwert. In der τ -Grafik ist zusätzlich ein innerer Kreis eingezeichnet, der τ_{\min} entspricht.

4.3.3 Sensibilität gegenüber der Fehlerfunktion

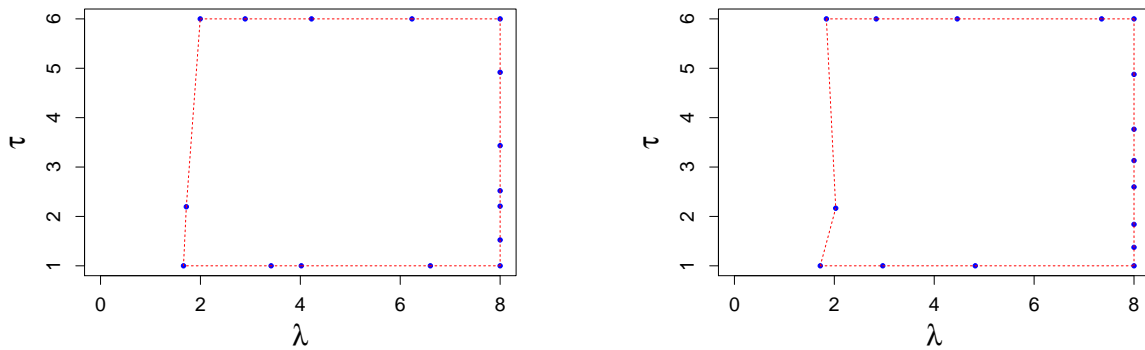
Die genaue Gewichtung der Fehlentscheidung scheint dabei strukturell keinen großen Einfluss auf die Parameterwahl zu haben. Von Bedeutung ist eher nur der Grundsatz, dass die Fehlergewichtung mit höherer Nachbarschaftordnung zunimmt.

Geht man von der linearen Gewichtung über in die 4-te Wurzel dieser Werte, wird der Unterschied der Gewichtung deutlich verringert.

In Abbildung 54 sind zwei Parameterwahlen eingezeichnet, die bei Simulationen mit etwa demselben minimalen Fehler gefunden wurden. Die allgemein stärkere Gewichtung hat zur Folge, dass die Punkte wieder näher zum Nullstimulus rutschen, da die Fehlentscheidung auf diesen nun relativ nicht mehr so hoch ist. Auch in der Höhe der Detektionswahrscheinlichkeit, die nun 33.2% beträgt, lässt sich erkennen, dass man sich durch die geringeren Unterschiede in der Gewichtung mehr zum globalen Fall bewegt hat.

Bemerkung 4.4. *Das Verschieben aller Stimuli auf den Rand bleibt jedoch auch für eine deutlich schwächere Gewichtung bestehen und ist somit ein Phänomen, das relativ stabil gegenüber der genauen Gewichtung ist.*

Offen bleibt die Frage, ob man dies auch für eine weitaus höhere Stimulizahl beobachten wird. Man erwartet, dass der Platz auf dem äußeren Rechteck irgendwann zu eng



(a) Beste Simulation für 16 Stimuli mit $\lambda_{max} = 8$ und $\tau_{max}/\tau_{min} = 6$.

(b) Annähernd gleich gute Simulation für 16 Stimuli mit $\lambda_{max} = 8$ und $\tau_{max}/\tau_{min} = 6$.

Abbildung 54: Die Parameterkombinationen unter 100-Simulationen, die unter der 4-ten Wurzel der linearen Fehlerfunktion einen minimalen Gesamtfehler für 16 Stimuli aufgewiesen haben. Die Unterschied beider Simulationen liegt unter 10^{-5} . Die Punkte entsprechen den gewählten Parametern in der λ - τ -Ebene. Mit der rot gepunkteten Linie sind die benachbarten Stimuli verbunden.

wird, aber gleichzeitig werden die Fehler benachbarter Stimuli immer schwächer gewichtet, sodass diese ohne große Auswirkungen auf den Gesamtfehler dichter zusammengesetzt werden können. Aber auch mit stochastischer Optimierung kommt man hier nicht weiter, da neben dem steigenden Aufwand der Auswertung der bedingten Fehlerwerte auch die Simulationszahl angehoben werden muss, um eine qualitativ hochwertige Lösung zu finden.

5 Schlussbetrachtung

5.1 Zusammenfassung

In dieser Arbeit wurde das ELO-Modell betrachtet, das sowohl die Dynamik von Einzelspiketrains als auch die Interaktion zwischen verschiedenen Spiketrains abbilden kann. Das ELO-Modell umfasst dabei die vier bekannten Beschreibungsebenen Feuerrate, Feuermuster (Phase), Oszillation und Synchronizität.

Kernfrage war, wie Rate und Phase zu wählen sind, damit die Detektion innerhalb eines einzelnen Oszillationszyklus am sichersten funktioniert. Jeder Stimulus erscheint mit derselben Wahrscheinlichkeit und es wird der Stimulus gewählt, der für den beobachteten Spiketrain am wahrscheinlichsten in Frage kommt.

Das Ziel war dabei, die globale Detektionswahrscheinlichkeit - die Wahrscheinlichkeit, den korrekten Stimulus zu wählen - zu maximieren.

Unter geeigneten Randbedingungen wurde zuerst die Informationsverarbeitung nur auf Basis von Rate untersucht. Hier ließ sich ein Algorithmus entwickeln, mit dem sich induktiv die optimale Wahl der Raten ohne großen Rechenaufwand bestimmen lässt.

Weiterhin ließ sich asymptotisch eine analytische Lösung beweisen, die bereits für kleine Raten eine außerordentlich gutes Ergebnis liefert.

Anschließend wurde die Stimuluserkennung allein auf Basis der Phase optimiert. Unabhängig von der globalen Rate konnte hier für eine beliebige Anzahl Stimuli eine analytische Lösung bewiesen werden

Ausgehend von diesen Ergebnissen wurden die optimalen Parameterkombinationen zwischen den beiden Extremfällen numerisch näher analysiert. Dabei wurde die Bedeutung der Phase zur effektiven Stimuluserkennung herausgearbeitet und das Zusammenspiel von Rate und Phase näher erforscht. Insbesondere für hohe Raten gewinnt die Phase stark an Bedeutung und übernimmt größtenteils die Unterscheidung der Stimuli.

Zum Abschluss wurde das ELO-Modell durch das Einführen einer Fehlerfunktion um Ähnlichkeiten erweitert. Ähnliche Stimuli sollten auch ähnliche Reaktionen im visuellen Kortex hervorbringen.

Da das Bestimmen von den bedingten Fehlentscheidungen einen deutlichen Mehraufwand bedeutet, wurde hierbei auf stochastische Optimierung zurückgegriffen, womit sich auch bei einer größeren Anzahl Stimuli fast optimale Parameterkombinationen finden lassen.

Insbesondere wurden dabei die Auswirkungen einer linearen Fehlerfunktion untersucht. Das niedrige Gewichten von Fehlentscheidungen bezüglich nahe verwandter Stimuli hat zur Folge, dass alle Parameter nun auf dem Rand liegen, d.h. entweder bei den Phasen nur minimale oder maximale Phase auftritt oder bei den Raten nur minimale (abhängig von

der Gewichtung zum Nullstimulus) oder maximale Rate vorliegt. Diese Effekte sind dabei ziemlich stabil gegenüber der genauen Gewichtung.

5.2 Ausblick

Die Ergebnisse dieser Arbeit beziehen sich auf ein Neuron innerhalb des ELO-Modells. Dies ist äquivalent zu dem Fall, dass man ein Netzwerk von Neuronen hat, die alle in Verteilung identisch reagieren. Dort zählt aufgrund der Unabhängigkeit der Prozesse allein die Gesamtanzahl der Spikes und deren mittleres zeitliches Auftreten.

Lässt man jedoch zu, dass Neuronen unterschiedliche Parameter wählen dürfen und sich somit spezifisch auf eine Gruppe Stimuli spezialisieren können, ergeben sich neue Konstruktionsmöglichkeiten, die das Detektieren optimieren können.

Im Fall von zwei Stimuli und zwei Neuronen würde sich jeweils ein Neuron auf genau einen Stimulus beschränken und dort mit maximaler Rate feuern. Feuert kein Neuron, ist höchstwahrscheinlich keiner der beiden Stimuli erschienen.

Interessant ist hierbei, ob sich für mehrere Neuronen bei Einschränkung auf Rate oder Phase die gezeigten Lösungen im Fall von einem Neuron wiederfinden.

Weiterhin ermöglichen mehrere spezifische Neuronen bzw. ein Netzwerk von Neuronen zusätzlich die Dimension des Lernens zu berücksichtigen.

Ein Netzwerk von Neuronen kann sich auf verschiedene Wiedererkennungswerte eines Stimulus, sogenannten Features, spezialisieren. Die einzelnen Neuronen geben weiter, welches Feature sie am ehesten erkannt haben und wie sicher sie sich sind. Aus den gesamten Informationen der Neuronen lässt sich identifizieren, welcher Stimulus am wahrscheinlichsten zu sehen war.

Dabei wird das neuronale Netzwerk, wenn es gut an die Umwelt angepasst ist, berücksichtigen, wie die gemeinsame Verteilung der Features in der Umwelt ist und diese Information ständig assimilieren.

Im Fall von einem Neuron entspricht dieses Lernen dem einfachen Anpassen der relativen Häufigkeit der Stimuli. Somit fließt - zusätzlich zu der raten- und phasenbasierten Entscheidung - die Information über die Statistik der Umwelt mit ein.

Eine große Herausforderung wird sein, den Rechenaufwand in einem kontrollierbaren Rahmen zu halten. Ein möglicher Ansatz ist, beim Maximieren der Detektionswahrscheinlichkeit nicht alle Entscheidungsschranken zu bestimmen, sondern nur auf die unmittelbaren benachbarten Stimuli zurückzugreifen, die das Gebiet maßgeblich vorgeben.

Abbildungsverzeichnis

1	Das ELO-Modell für zwei Neurone	4
2	Entscheidungsaufgabe im Falle von 2 Stimuli und ein Neuron	5
3	Entscheidungsschranke nur auf Basis von λ für zwei Stimuli	16
4	Das optimale λ_2 als Treppenfunktion von λ_1	18
5	Das optimale λ_2 und die maximale Detektionswahrscheinlichkeit in Abhängigkeit von λ_1 für 2 Stimuli.	19
6	Die Werte von λ_1 , für die die optimale Schranke um 1 springt.	20
7	Optimale obere Schranke für feste untere Schranke in Abhängigkeit von λ_1	23
8	Optimale obere und untere Schranke in Abhängigkeit von λ_1	24
9	Optimale Schranken in Abhängigkeit von λ_1 für verschieden Anzahl Stimuli	28
10	Maximale Detektionswahrscheinlichkeit in Abhängigkeit von λ_1	29
11	Optimales λ_2 in Abhängigkeit von λ_1	29
12	Vergleich exakte Lösung mit Normalapproximation für 2 und 3 Stimuli . .	33
13	Trennung von zwei Normalverteilungen	35
14	Trennung von drei Normalverteilungen	36
15	Asymptotische Verhältnis der Abstände vom Schnittpunkt für drei Normalverteilungen	38
16	Vergleich numerische Berechnung mit Asymptotik für 5 Stimuli	39
17	Entscheidungsschranke für $\tau_1 = 4$ und $\tau_2 = 1$	41
18	Detektionswahrscheinlichkeit in Abhängigkeit von τ_2	43
19	$h_1(\tau_2)$ und $h_2(\tau_2)$ für verschiedene τ_1/τ_3	45
20	Entscheidungsschranken für $k = 4$ Stimuli	48
21	Entscheidungsschranke für simulierte Werte für $\lambda_1 = 6$, $\tau_1 = 4$, $\lambda_2 = 2$ und $\tau_2 = 2$	56
22	Die Detektionswahrscheinlichkeit in Abhängigkeit von $\lambda_2 \leq \lambda_1$	57
23	Detektionswahrscheinlichkeit in Abhängigkeit von τ_{\max}/τ_{\min} für $\lambda_1 = 6$ und ausgewählte λ_2	58
24	Zwei mögliche Kombinationen im Fall von zwei Stimuli	59
25	Beziehung zwischen den beiden Fällen $\lambda_1 > \lambda_2$, $\tau_1 > \tau_2$ (rot) und $\lambda_1 > \lambda_2$, $\tau_1 < \tau_2$ (grün)	60
26	Auswirkung auf die Entscheidungsschranken für den Fall $\lambda_1 > \lambda_2$, $\tau_1 > \tau_2$ durch Parameterveränderungen in kritischen Bereichen	61
27	Auswirkung auf die Entscheidungsschranken für den Fall $\lambda_1 > \lambda_2$, $\tau_1 < \tau_2$ durch Parameterveränderungen in kritischen Bereichen	62
28	Optimales λ_2 und maximale Detektionswahrscheinlichkeit für den roten und grünen Fall für festes λ_1 in Abhängigkeit von τ_{\max}/τ_{\min}	63
29	Für welche τ_{\max}/τ_{\min} besitzt der rote oder der grüne Fall die höhere Detektionswahrscheinlichkeit	64
30	Welches τ_{\max}/τ_{\min} wird benötigt, um die Detektionswahrscheinlichkeit gegenüber nur der Unterscheidung auf Basis von λ um 5% zu erhöhen	64
31	Die Bedeutung der Phase für 2 Stimuli	65

32	Optimale Detektionswahrscheinlichkeit und optimales τ_i für festes $\lambda_1 = 6$, $\lambda_2 = 3$, $\lambda_3 = 1$ in Abhängigkeit von τ_{\max}/τ_{\min} für die beiden anderen τ . . .	66
33	Optimale Detektionswahrscheinlichkeit und optimales τ_1 für festes $\lambda_1 = 6$, $\lambda_2 = 3$, $\lambda_3 = 1$ in Abhängigkeit von τ_{\max}/τ_{\min} für die beiden Möglichkeiten $\tau_2 = \tau_{\max}$ oder $\tau_3 = \tau_{\max}$	67
34	Optimale Detektionswahrscheinlichkeit und optimales τ_i für festes $\lambda_1 = 3$, $\lambda_2 = 2$, $\lambda_3 = 1$ in Abhängigkeit von τ_{\max}/τ_{\min} für die beiden anderen τ . . .	68
35	Optimale Detektionswahrscheinlichkeit und optimales τ_1 für festes $\lambda_1 = 3$, $\lambda_2 = 2$, $\lambda_3 = 1$ in Abhängigkeit von τ_{\max}/τ_{\min} für die beiden Möglichkeiten $\tau_2 = \tau_{\max}$ oder $\tau_3 = \tau_{\max}$	69
36	Optimale Parameterwahl und optimale Detektionswahrscheinlichkeit für 3 Stimuli für $\lambda_1 = 6$ und $\lambda_1 = 12$ in Abhängigkeit von τ_{\max}/τ_{\min}	70
37	Auftretende Kombinationen für 3 Stimuli für $\lambda_1 = 12$	71
38	Bedingte Detektionswahrscheinlichkeit für $\lambda_1 = 6$ und $\lambda_1 = 12$ in Abhängigkeit von τ_{\max}/τ_{\min}	73
39	Optimale Parameterwahl und optimale Detektionswahrscheinlichkeit für 4 Stimuli für $\lambda_1 = 6$ und $\lambda_1 = 12$ in Abhängigkeit von τ_{\max}/τ_{\min}	74
40	Auftretende Kombinationen für 4 Stimuli für $\lambda_1 = 12$	75
41	Optimale Kombinationen für 7 Stimuli für $\lambda_1 = 8$ und für τ_{\max}/τ_{\min} im Bereich zwischen 2 und 18	76
42	Maximale Detektionswahrscheinlichkeit für 7 Stimuli für $\lambda_1 = 8$ in Abhängigkeit von τ_{\max}/τ_{\min} . In rot (grün) ist die Detektionswahrscheinlichkeit dargestellt, die man durch einen reinen Phasen-Code (Raten-Code) erhält.	78
43	Anordnung der Stimuli im Fall $k = 8$	80
44	Bedingte Fehlentscheidung zweier Stimuli im Konturdiagramm	82
45	Verlauf der Detektionswahrscheinlichkeit einer Simulation für feste Varianz	84
46	Histogramm über die Abweichung der Detektionswahrscheinlichkeit vom Optimum für feste und variable Varianz	85
47	Parameterkombination mit maximaler Detektionswahrscheinlichkeit simuliert für 8 und 16 Stimuli	85
48	Entscheidungsschranken und die daraus resultierenden Gebiete	86
49	Vergleich Rechteckstruktur mit einer mittigen Platzierung	87
50	Optimale Parameterkombination unter einer linearen Fehlerfunktion für 8 und 16 Stimuli	88
51	Bedingte Fehlentscheidungen angedeutet anhand der Linienstärke	88
52	Anordnung mit vergleichbarer Detektionswahrscheinlichkeit bzw. Gesamtfehler	89
53	Anordnung mit minimalen linearen Fehler als Tuning-Curve	90
54	Optimale Parameterkombination unter der 4-ten Wurzel der linearen Fehlerfunktion	91

Literatur

Adjamian u. a. 2004

ADJAMIAN, P. ; HILLIDAY, I. E. ; BARNES, G. R. ; HILLEBRAND, A. ; HADJIPAPAS, A. ; SINGH, K. D.: Induced visual illusions and gamma oscillations in human primary visual cortex. In: *European Journal of Neuroscience* 20(2) (2004), S. 587–592

Barlow 1972

BARLOW, H. B.: Single units and sensation: A neuron doctrine for perceptual psychology? In: *Perception* 1(4) (1972), S. 371–394

Bethge u. Pawelzik 2002

BETHGE, M. ; PAWELZIK, K.: Geheimsprache der Neuronen. In: *Gehirn und Geist* 02 (2002), S. 80–87

Bingmer u. a. 2011

BINGMER, M. ; SCHIEMANN, J. ; ROEPER, J. ; SCHNEIDER, G.: Measuring burstiness and regularity in oscillatory spike-trains. In: *Journal of Neuroscience Methods* 201 (2011), S. 426–437

Braitenberg u. Schüz 1991

BRAITENBERG, Valentino ; SCHÜZ, A.: *Anatomy of the cortex : statistics and geometry*. Springer-Verlag Berlin ; New York, 1991

Buzsáki u. Chrobak 1995

BUZSÁKI, G. ; CHROBAK, J.J.: Temporal structure in spatially organized neuronal ensembles: a role for interneuronal networks. In: *Current Opinion in Neurobiology* 5 (1995), S. 504–510

Camastra u. Vinciarelli 2007

CAMASTRA, F. ; VINCIARELLI, A.: *Machine Learning for Audio, Image and Video Analysis*. Springer Science & Business Media, 2007

Cox u. Lewis 1966

COX, D. R. ; LEWIS, P. A. W.: *The Statistical Analysis of Series of Events*. Chapman and Hall, 1966

Eckhorn u. a. 1988

ECKHORN, R. ; BAUER, R. ; JORDAN, W. ; KRUSE, W. ; MUNK, M. ; REITBOECK, H. J.: Coherent oscillations: A mechanism of feature linking in the visual cortex. In: *Biological Cybernetics* 60(2) (1988), S. 121–130

Engel u. a. 1992

ENGEL, A. K. ; KÖNIG, P. ; KREITER, A. K. ; SCHILLEN, T. B. ; SINGER, W.: Tecortex coding in the visual cortex: New vistas on integration in the nervous system. In: *Trends in Neurosciences* 15(6) (1992), S. 218–226

Frien u. a. 1994

FRIEN, A. ; ECKHORN, R. ; BAUER, R. ; WOELBERN, T. ; KEHR, H.: Stimulus-specific fast oscillations at zero phase between visual areas V1 and V2 of awake monkey. In: *Neuroreport* 5(17) (1994), S. 2273–2277

Fries 2009

FRIES, P.: Neuronal gamma-band Synchronization as a fundamental process in cortical computation. In: *Annual Review of Neuroscience* 32 (2009), S. 209–224

Fries u. a. 2007

FRIES, P. ; NIKOLIĆ, D. ; SINGER, W.: The gamma cycle. In: *Trends in Neurosciences* 30 (2007), S. 309–316

Gerstein u. Aertsen 1989

GERSTEIN, G. L. ; AERTSEN, A. M. H. J.: Neuronal assemblies. In: *IEEE Transactions on Biomedical Engineering* 36(1) (1989), S. 4–14

Gray u. a. 1992

GRAY, C. M. ; ENGEL, A. K. ; KÖNIG, P. ; SINGER, W.: Synchronization of oscillatory neuronal responses in cat striate cortex: Temporal properties. In: *Visual Neuroscience* 8(4) (1992), S. 337–347

Gray u. Singer 1989

GRAY, C. M. ; SINGER, W.: Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. In: *Proceedings of the National Academy of Sciences of the United States of America* 86(5) (1989), S. 1698–1720

Havenith u. a. 2011

HAVENITH, M. N. ; YU, S. ; BIEDERLACK, J. ; CHEN, N. ; SINGER, W. ; NIKOLIĆ, D.: Synchrony Makes Neurons Fire in Sequence, and Stimulus Properties Determine Who Is Ahead. In: *The Journal of Neuroscience* 31(23) (2011), S. 8570–8584

Hoogenboom u. a. 2006

HOOGENBOOM, N. ; SCHOFFELEN, J. M. ; OOSTENVELD, R. ; PARKES, L.M. ; FRIES, P.: Lattention human visual gamma-band activity in frequency, time and space. In: *Neuroimage* 29(3) (2006), S. 764–773

Hubel u. Wiesel 1962

HUBEL, D.H. ; WIESEL, T. N.: Receptive fields, binocular interactions and functional architecture in the cat's visual cortex. In: *Journal of Physiology* 160 (1962), S. 106–154

Kandel u. a. 1996

KANDEL, R. ; SCHWARTZ, J. H. ; JESSELL, T. M.: *Neurowissenschaften. Eine Einführung*. Spektrum akademischer Verlag, Heidelberg, 1996

König. u. a. 1995

KÖNIG., P. ; ENGEL, A.K. ; ROELFSEMA, P.R. ; SINGER, W.: How precise is neuronal synchronization? In: *Neural Computation* 7 (1995), S. 469–485

Larkman 1991

LARKMAN, A. U.: Dendritic morphology of pyramidal neurones of the visual cortex of the rat: III. Spine distributions. In: *The Journal of Comparative Neurology* 306(1) (1991), S. 332–343

Perrett u. a. 1982

PERRETT, D. I. ; ROLLS, E. T. ; CAAN, W.: Visual neurones responsive to faces in the monkey temporal cortex. In: *Experimental Brain Research* 47(3) (1982), S. 329–342

Roelfsema u. a. 1997

ROELFSEMA, P.R. ; ENGEL, A.K. ; KÖNIG, P. ; SINGER, W.: Visuomotor integration is associated with zero time-lag synchronization among cortical areas. In: *Nature* 385 (1997), S. 157–161

Schiemann u. a. 2012

SCHIEMANN, J. ; KLOSE, V. ; SCHLAUDRAFF, F. ; BINGMER, M. ; SEINO, S. ; MAGILL, P. J. ; SCHNEIDER, G. ; LISS, B. ; ROEPER, J.: K-ATP channels control in vivo burst firing of dopamine neuron in the medial substantia nigra and novelty-induced behavior. In: *Nature Neuroscience* 15 (2012), S. 1272–1280

Schneider 2008

SCHNEIDER, G.: Messages of oscillatory correlograms - a spike-train model. In: *Neural Computation* 20(5) (2008), S. 1211–1238

Schneider u. a. 2006

SCHNEIDER, G. ; HAVENITH, M. N. ; NIKOLIĆ, D.: Spatio-temporal structure in large neuronal networks detected from cross correlation. In: *Neural Computation* 18 (10) (2006), S. 2387–2413

Schneider u. Nikolić 2006

SCHNEIDER, G. ; NIKOLIĆ, D.: Detection and assessment of near-zero delays in neuronal spiking activity. In: *Journal of Neuroscience Methods* 152 (2006), S. 97–106

Senatov 1998

SENATOV, V. V.: *Normal Approximation: New Results, Methods And Problems*. VSP, 1998

Singer 1993

SINGER, W.: Synchronization of cortical activity and its putative role in information processing and learning. In: *Annual Review of Physiology* 55 (1993), S. 349–374

Taylor u. a. 2005

TAYLOR, K. ; MANDON, S. ; FREIWALD, W. A. ; KREITER, A. K.: Coherent oscillatory activity in monkey area V4 predicts successful allocation of attention. In: *Cereb. Cortex* 15(9) (2005), S. 1424–1437

Wyart u. Tallon-Baudry 2008

WYART, V. ; TALLON-BAUDRY, C.: Neural dissociation between visual awareness and spatial attention. In: *The Journal of Neuroscience* 28(10) (2008), S. 2667–2679

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass alle Stellen der Arbeit, die wörtlich oder sinngemäß aus anderen Quellen übernommen wurden, als solche kenntlich gemacht sind, und dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt wurde.

Frankfurt am Main, den 17. Juli 2014