Zusammenfassung der Diplomarbeit anlässlich
der Verleihung des Hans-Pundt-Preises 2021 des VDE Dresden e.V.

Kanonische Korrelationen und die Berechnung von Informationsmaßen für unendlichdimensionale Verteilungen

Jonathan E. W. Huffmann
Betreuer: Dr.-Ing. Martin Mittelbach

Der folgenden Text gibt eine allgemeinverständliche Einführung in das Thema der Diplomarbeit, ohne dabei zu stark auf mathematischen Hintergründe und Feinheiten einzugehen. Der interessierte Leser sei auf die Diplomarbeit [2] verwiesen.

1 Einleitung und Problemstellung

Die abstrakte Aufgabe eines Kommunikationssystems ist die zuverlässige Übertragung von Nachrichten von einem Sender über einen Übertragungskanal an einen Empfänger.

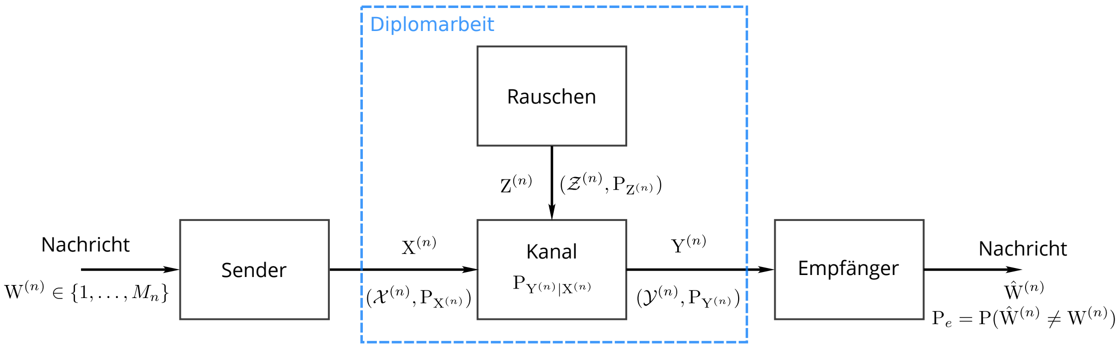

Abbildung 1: Abstraktes Übertragungsmodel nach Shannon.

Die Übertragung über den Kanal wird dabei im Allgemeinen durch Störungen, zum Beispiel in Form von Kanalrauschen, verfälscht. Die Realisierung sowohl der Nachrichten als auch des Kanalrauschens ist bei der Analyse und der Entwicklung eines Kommunikationssystems als unbekannt anzunehmen und wird deshalb jeweils über entsprechende Wahrscheinlichkeitsverteilungen mathematisch modelliert werden. Ein Kommunikationsmodell dieser Art ist in Abbildung 1 schematisch dargestellt, wobei die einzelnen Komponenten die folgenden Funktionen erfüllen.
Eine uncodierte Datenübertragung über einen Kanal ist aufgrund des Rauschens in der Regel mit einer hohen Fehlerwahrscheinlichkeit verbunden. Die Aufgabe des Senders ist es deshalb, jeweils eine von M_n Nachrichten, codiert als $\mathrm{W}^{(n)}$, in n Schritten an den Empfänger zu übertragen. Ein Übertragungsfehler P_e liegt genau dann vor, wenn die

Nachricht $\hat{\mathrm{W}}^{(n)}$ nach der Übertragung über den Kanal nicht mehr korrekt am Empfänger decodiert werden kann.
Die zentrale Fragestellung bei der Analyse eines Kommunikationssystems ist dabei, unter welchen Bedingungen an den Sender und den Empfänger, trotz Störungen bei der Übertragung, eine zuverlässige Datenübertragung möglich ist. Zusammengefasst ergibt sich daraus eine der zentralen Fragestellungen der Informationstheorie nach dem Zusammenhang zwischen der maximalen Anzahl an Nachrichten M_n, der maximaler Übertragungsrate R

$$R = \frac{1}{n} \log\left(M_n\right) \tag{1}$$

und der Decodierfehlerwahrscheinlichkeit P_e.
Ein erster Lösungsansatz für dieses klassische Problem konnte von Shannon [5] bereits 1948 mithilfe der von ihm eingeführten Transinformation $I(\mathrm{X};\mathrm{Y})$

$$I(\mathrm{X};\mathrm{Y}) = \int_{\mathcal{X}\times\mathcal{Y}} \log\left(\frac{dP_{\mathrm{X}\,\mathrm{Y}}}{dP_{\mathrm{X}} \otimes P_{\mathrm{Y}}}(x,y)\right) P_{\mathrm{X}\,\mathrm{Y}}(dx,dy) \tag{2}$$

bzw. der Kanalkapazität C

$$C = \sup_{\mathrm{P_X}} I(\mathrm{X};\mathrm{Y}) \tag{3}$$

gefunden werden. Dabei beschreibt die Transinformation die statistische Abhängigkeit zwischen dem Kanaleingang und dem Kanalausgang. Um eine möglichst hohe Datenrate zu erzielen, ist ein entsprechendes Codierverfahren/Decodierverfahren zu entwickeln, dass diese Abhängigkeit bezüglich der Verteilung der Kanaleingangssymbole maximiert.
Für die Fragestellung nach einer zuverlässigen Kommunikation über einen Kanal konnte von Shannon das folgende Kanalcodierungstheorem bewiesen werden: *Für alle Übertragungsraten R kleiner als die Kanalkapazität C kann für asymptotisch wachsende Codewortlängen n eine beliebig hohe Zuverlässigkeit, d.h. eine beliebig kleine Decodierfehlerwahrscheinlichkeit* P_e *erreicht werden.*
Die noch erstaunlichere Umkehrung dieses Resultates konnte einige Jahre später 1957 von Wolfowitz gezeigt werden. Dieser bewies (vgl. [6]), dass für Übertragungsraten R größer als die Kanalkapazität C die Decodierfehlerwahrscheinlichkeit für $n \to \infty$ gegen 1 konvergiert. Dies bedeutet eine zuverlässige Datenübertragung ist asymptotisch nur bis zu Datenraten R kleiner als die Kanalkapazität möglich.
Die gefundenen Resultate von Shannon und Wolfowitz sind dabei zwar von hoher theoretischer Relevanz, hängen jedoch asymptotisch von der Blocklänge n ab und benötigen eine entsprechend hohe bzw. asymptotisch unendlich lange Decodierzeit und damit hohe Latenzzeiten am Empfänger. Von entsprechend praktischer Bedeutung ist deshalb insbesondere die Analyse der Decodierfehlerwahrscheinlichkeit für endliche Blocklängen n.
Relevant ist dies insbesondere für moderne Kommunikationsanwendungen, bei denen geringe Latenzzeiten und eine hohe Zuverlässigkeit erforderlich sind wie, sie zum Beispiel in aktuellen 5G Mobilfunknetzen, in MIMO Systemen bzw. dem autonomen Fahren

vorkommen. Auch Anwendungen in denen Sensor- und Radardaten schnell und gemeinsam mit Kommunikationsdaten ausgewertet werden müssen, sogenannten joint-sensing-and-communication-Anwendungen, oder die Kommunikationskanäle hohen und vielen Schwankungen unterliegen, ist eine nicht-asymptotische, informationstheoretische Analyse von immanenter Bedeutung.

Erste Ergebnisse für Schranken der Decodierfehlerwahrscheinlichkeit in Verbindung mit der Blocklänge n ist schon in Arbeiten von Shannon, Strassen und Dobrushin zu finden. Weitere starke Verbesserungen in diesem Gebiet wurden insbesondere durch die Arbeit von Polyanskiy [4] kürzlich erzielt.

Für die genaue Analyse des Zusammenhanges der Fehlerwahrscheinlichkeit P_e und der Datenrate R bei endlicher Codewortlänge n sind insbesondere die Verteilung und die Momente der sogenannten Informationsdichte $i(\mathrm{X};\mathrm{Y})$,

$$i(\mathrm{X};\mathrm{Y}) := \log\left(\frac{dP_{\mathrm{X}\,\mathrm{Y}}}{dP_{\mathrm{X}} \otimes P_{\mathrm{Y}}}(\mathrm{X},\mathrm{Y})\right) \tag{4}$$

wobei die Transinformation $I(\mathrm{X};\mathrm{Y})$ dem Erwartungswert der Zufallsgröße $i(\mathrm{X};\mathrm{Y})$ entspricht, von Bedeutung.

Bei den zugrundeliegenden Kanalmodellierungen wird sich zumeist auf eine wertdiskrete und zeitdiskrete Modellierung mit weißem Rauschen, d.h. gedächtnislose Modelle beschränkt. Einfache verallgemeinerte Modelle beschäftigen sich zumeist mit wertkontinuierlichen aber immer noch zeitdiskreten Modellen ohne Gedächtnis. Für diesen Fall sind entsprechende Formeln für die Transinformation in der Literatur zu finden.

Weiterführend konnte in einer Arbeit des Autors zusammen mit Martin Mittelbach [3] gezeigt werden, dass für den praktisch wichtigen Spezialfall, gemeinsam normalverteilter Kanaleingangs-und-ausgangssymbole, die Verteilung und die Momente der Informationsdichte, analytisch berechnet werden können. Die Arbeit enthält dabei auch entsprechende numerische Berechnungen und die Ergebnisse sind insbesondere für die Analyse von Kanälen mit Gedächtnis anwendbar.

Auch eine zeitdiskrete Kanalmodellierung stellt für gewisse Szenarien/Anwendungen mitunter eine unzulässige Vereinfachung dar. Insgesamt können die oben genannten Kanalmodelle als Spezialfall wertkontinuierlicher und zeitkontinuierlicher Kanalmodelle, im Allgemeinen mit Gedächtnis z.B. farbiges Rauschen, verstanden werden, welche in der Diplomarbeit behandelt wurden.

2 Ergebnisse

Insbesondere wert- und zeitkontinuierliche Modelle können physikalischen Prozesse der Signalübertragung realitätsnäher abbilden, wodurch ein besseres Verständnis und eine weiterführende Optimierung von Sender- und Empfängerstrukturen ermöglicht wird.

Die Berechnung der Verteilung sowie der Momente der zentralen Größe Informationsdichte gestaltet sich dabei jedoch insbesondere für wertkontinuierliche Kanalmodelle mit kontinuierlicher Zeitstruktur als besonders schwierig und stellte die zentrale Problemstellung in der Diplomarbeit da. Dabei wurde sich, um die Berechnungen zu vereinfachen, auf den praktisch wichtigen Spezialfall normalverteilter Rausch- und Signalgrößen mit Gedächtnis beschränkt.

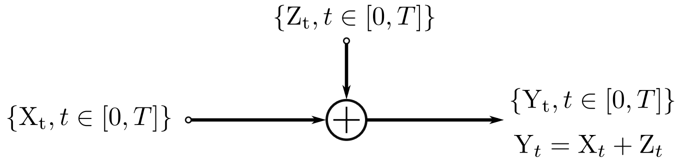

Abbildung 2: Abstraktes Kanalmodell

Für die abstrakte Modellierung des Rauschens sowie der Signale wurden Zufallsprozesse auf einem Hilbertraum gewählt. Anstelle von Kovarianzmatrizen in diskreten Modellen, treten damit Kovarianzoperatoren $\mathrm{K_{X\,X}}$, $\mathrm{K_{Y\,Y}}$ und $\mathrm{K_{X\,Y}}$ der beteiligten Zufallsgrößen auf.
Von besonderer Bedeutung ist in diesem Fall der additive Gaußkanal (siehe Abbildung 2), bei dem die Rausch- und Signalgrößen durch stationäre Zufallsprozesse mit rationaler Spektraldichte auf einem endlichen Zeitintervall $[0,T]$ gegeben sind.
Erste Ergebnisse für die Berechnung der Transinformation für diesen Fall sind in den Arbeiten von Gelfand [1] zu finden, wobei die Berechnung der Transinformation die Lösung eines verallgemeinertes Eigenwertproblem erfordert. Das beschriebene Verfahren in Gelfand ist jedoch nicht numerisch durchführbar und verlangt eine gezielte Analyse der Kovarianzoperatoren für jeden Spezialfall.
In der Diplomarbeit konnte gezeigt werden, dass aufbauend auf den Ideen von Gelfand auch die charakteristische Funktion $\varphi_{i(\mathrm{X;Y})-I(\mathrm{X;Y})}(t)$

$$\varphi_{i(\mathrm{X;Y})-I(\mathrm{X;Y})}(t) = \sqrt{\frac{\Phi(-1)}{\Phi(-(1+t^2))}}, \qquad t \in \mathbb{R}, \tag{5}$$

sowie die Momente $\mathbb{E}\left[(i(\mathrm{X;Y}) - I(\mathrm{X;Y}))^n\right]$ der Informationsdichte $i(\mathrm{X;Y})$ über ein verallgemeinertes Eigenwertproblem berechnet werden können.

$$\mathbb{E}\left[(i(\mathrm{X;Y}) - I(\mathrm{X;Y}))^n\right] = (-1)^n \left.\frac{d^n}{dt^n}\sqrt{\frac{\Phi(-1)}{\Phi(-(1+t^2))}}\right|_{t=0}, \quad \text{für gerade } n \tag{6}$$

Die Lösung dieses Problems insbesondere die verallgemeinerten Eigenwerte können dabei als Nullstellen einer speziellen Funktion Φ dargestellt werden, die sich gut für konkrete Berechnungen eignet.
Darüber hinaus wurde in der Diplomarbeit ein systematisches Verfahren zur Berechnung des oben genannten Eigenwertproblems entwickelt, welches es erlaubt, für den Fall rationaler Spektraldichten sowohl die Momente als auch die charakteristische Funktion der Informationsdichte analytisch zu berechnen. Das so entwickelte Verfahren eignet sich dabei insbesondere für computergestützte Berechnungen. Zusätzlich können beliebiege andere Spektraldichten durch diese rationalen Spektraldichten genau approximiert werden.

Dabei wird, in dem entwickelten Verfahren, eine simultane Diagonalisierung der Kovarianzoperatoren $\mathrm{K}_{\mathrm{X}\,\mathrm{X}}$ und $\mathrm{K}_{\mathrm{Y}\,\mathrm{Y}} = \mathrm{K}_{\mathrm{X}\,\mathrm{X}} + \mathrm{K}_{\mathrm{Z}\,\mathrm{Z}}$ ähnlich der zeitdiskreten kanonischen Korrelationsanalyse durchgeführt. Das Verfahren, welches es erlaubt, die Rauschprozesse simultan in zwei unendlichdimensionale diskrete Zufallsvektoren zu diagonalisieren, kann dabei auch auf andere Probleme wie zum Beispiel in der Rauschanalyse und der analogen Schaltungstechnik sowie in diversen physikalischen Analysen von Rauschprozessen Anwendung finden. Ein weiteres Anwendungsgebiet ist die Spektralanalyse in der Optik. Schlussendlich wurden Approximationen der Verteilungsfunktion der Informationsdichte sowie Fehlerabschätzungen für den zeitkontinuierlichen Fall, basierend auf der charakteristischen Funktion entwickelt, welche es erlauben insbesondere die Resultate auch aus dem zeitdiskreten Spezialfall anzuwenden.
Die Erweiterung und Anwendung der Ergebnisse insbesondere des entwickelten Diagonalisierungsverfahrens sind Gegenstand weiterer aktueller Forschungsarbeiten.

Literatur

[1] I. M. Gelfand und A. M. Yaglom. „Über die Berechnung der Menge an Information über eine zufällige Funktion, die in einer anderen zufälligen Funktion enthalten ist". In: *Arbeiten zur Informationstheorie*. Bd. II. Hrsg. von Heinrich Grell. Mathematische Forschungsberichte 6. Berlin: VEB Deutscher Verlag der Wissenschaften, 1958, S. 7–56.

[2] J.E.W. Huffmann. „Canonical Correlation and the Calculation of Information Measures for Infinite-Dimensional Distributions". Technische Universität Dresden, 2021. URL: `https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-742541`.

[3] J.E.W. Huffmann und M. Mittelbach. „On the Distribution of the Information Density of Gaussian Random Vectors: Explicit Formulas and Tight Approximations". In: *Entropy* 24.924 (Juli 2022). URL: `https://doi.org/10.3390/e24070924`.

[4] Yury Polyanskiy, H. Vincent Poor und Sergio Verdu. „Channel Coding Rate in the Finite Blocklength Regime". In: *IEEE Transactions on Information Theory* 56.5 (2010), S. 2307–2359.

[5] Claude E. Shannon und Warren Weaver. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press, 1964.

[6] Jacob Wolfowitz. *Coding Theorems of Information Theory*. Springer-Verlag Berlin Heidelberg, 1978.