

**Kapazität von vollständig und partiell
bekannten MIMO-Kanälen**

Jan Krohn

Matrikel-Nr. 214147

Diplomarbeit in Mathematik

vorgelegt der

**FAKULTÄT FÜR MATHEMATIK, INFORMATIK
UND NATURWISSENSCHAFTEN
der RWTH Aachen**

im Februar 2005

angefertigt am

Lehrstuhl für Theoretische Informationstechnik

Danksagung

Ich danke Herrn Univ.-Prof. Dr. Rudolf Mathar und Herrn Dipl.-Phys. Daniel Catrein für die Bereitstellung des Themas und für die Geduld bei der Betreuung der Diplomarbeit. Ich danke Herrn Rolf Wildenburg, Herrn Oliver Vieten, Herrn Roland Schmidt und meinen Eltern für das Korrekturlesen. Insbesondere danke ich meinen Eltern für die Unterstützung während meines gesamten Studiums.

Inhaltsverzeichnis

Kapitel 1. Einleitung	5
1.1. Motivation	5
1.2. Übersicht	5
Kapitel 2. Mathematische Grundlagen	7
2.1. Grundlagen aus der Stochastik	7
2.2. Informationstheoretische Grundlagen	9
2.3. Die komplexe Normalverteilung	14
2.4. MMSE-Schätzer	15
Kapitel 3. Das MIMO-Modell	17
Kapitel 4. Komplett bekannte Kanäle	21
4.1. Herleitung der Kapazität bei fester Kanalmatrix	21
4.2. Herleitung der Kapazität bei zufälliger Kanalmatrix	28
4.3. Weitere Modellmöglichkeiten	32
Kapitel 5. Empfängerseitig partiell bekannte Kanäle	33
5.1. Modellierung von Synchronised Detection	33
5.2. Abhängigkeitsstrukturen der Synentropie	38
5.3. Berechnung einiger Kapazitätsschranken	42
5.4. Weitere Optimierungsansätze	47
Kapitel 6. Senderseitig partiell bekannte Kanäle	49
6.1. Modellierung von senderseitigem P-CSI	49
6.2. Bestimmung einiger Kapazitätsschranken	50
Kapitel 7. Simulationen	65
Kapitel 8. Zusammenfassung und Ausblick	69
Anhang A. Quellcode	71
Literaturverzeichnis	75

KAPITEL 1

Einleitung

1.1. Motivation

Es ist anzunehmen, dass in der Zukunft datenintensive Anwendungen in der Mobilkommunikation eine immer größere Rolle spielen werden, beispielsweise bei Übertragung von Videodaten statt Audiodaten [10]. Eine nahe liegende aber kostenintensive Lösung wäre eine Erhöhung der Link- oder Kanalbandweite. Eine ökonomischere Lösung ist die Verteilung des Datenvolumens auf mehrere Sende- und Empfangsantennen. MIMO steht für “Multiple In Multiple Out”, und bedeutet, dass an Mobilgeräten und Basisstationen jeweils mehrere Antennen vorhanden sind. Es werden also keine einzelnen Symbole mehr gesendet und empfangen, sondern Vektoren von Symbolen. Das MIMO-Modell ist eine Verallgemeinerung des SISO (Single In Single Out) Gauß-Kanals. In der Literatur finden sich zahlreiche verschiedene Möglichkeiten, solch ein System zu modellieren. Dabei wird im Wesentlichen nach der Kanalmatrix unterschieden, ob sie vollständig, partiell oder nicht bekannt ist. Für einfache Modelle kann man die Kanalkapazität explizit angeben [6] [20], in der Regel ist die Maximierung der Synentropie jedoch ein noch offenes Problem, und es können nur obere und untere Schranken angegeben werden [3] [8] [10] [13].

1.2. Übersicht

Im folgenden Kapitel 2 werden die grundlegenden Definitionen und Sätze aus der Informationstheorie, sowie einige Ergebnisse aus anderen Bereichen der Stochastik zusammengefasst, die im Laufe der Arbeit benötigt werden. Kapitel 3 führt das MIMO-Modell ein. Es werden technische Hintergründe dargestellt sowie grundlegende Bezeichnungen und Definitionen angegeben.

Kapitel 4 beschäftigt sich mit vollständig bekannten MIMO-Kanälen. Das bedeutet, dass die Kanalmatrix beim Empfänger als bekannt vorausgesetzt wird. Man kann unterscheiden zwischen den Fällen, dass die Kanalmatrix bereits beim Sender bekannt ist, oder erst nach der Übertragung beim Empfänger. Für beide Fälle werden die Kapazität angegeben, sowie einige Spezialfälle und Grenzfälle genauer betrachtet.

Kapitel 5 behandelt “Synchronised Detection”, ein Verfahren zur Schätzung von empfängerseitig unbekanntem Kanalmatrixen durch vorangehende Testsymbole [3]. Es wird gezeigt, dass die Synentropie dieses Kanals nicht von den gesendeten und empfangenen Testsymbolen abhängt, sondern allein von den durch MMSE

geschätzten Kanalmatrizen. Die Berechnung der Kapazität ist hier ein noch offenes Problem, es können aber obere und untere Schranken angegeben werden.

In Kapitel 6 wird der Fall behandelt, dass beim Sender bereits partielles Wissen über den Kanal bekannt ist. Dieser Fall wird so modelliert, dass die Kanalmatrix in eine Summe von einem festen und einen zufälligen Teil zerfällt. Auch hier ist die Bestimmung der Kanalkapazität ein offenes Problem, mit den gleichen Methoden wie in Kapitel 5 lassen sich aber auch hier obere und untere Schranken bestimmen.

In Kapitel 7 werden einige Simulationsergebnisse vorgestellt, und die errechneten Kapazitäten und Kapazitätsschranken der vorigen Kapitel grafisch miteinander verglichen.

Kapitel 8 enthält eine Zusammenfassung der Ergebnisse sowie einen Ausblick.

Mathematische Grundlagen

Die Arbeit gehört in den Bereich der Informationstheorie. Zum Verständnis sind Grundkenntnisse aus der Informationstheorie notwendig, wie sie etwa in [11], [12] oder [7] nachgelesen werden können. Außerdem werden Kenntnisse in Linearer Algebra, Analysis und Stochastik vorausgesetzt.

Im dieser Arbeit werden Spaltenvektoren jeweils mit einem Unterstrich bezeichnet. Zufallsvektoren werden mit Großbuchstaben ausgedrückt ($\underline{X}, \underline{Y}, \underline{Z}, \dots$), feste Vektoren mit Kleinbuchstaben ($\underline{x}, \underline{y}, \underline{z}, \dots$). Zufallsmatrizen werden mit Skriptbuchstaben bezeichnet ($\mathcal{H}, \mathcal{X}, \dots$), feste Matrizen mit lateinischen oder griechischen Großbuchstaben (H, X, Λ, \dots). Univariate Zufallsvariablen werden ebenfalls mit Großbuchstaben bezeichnet. Von dieser Notation wird in Einzelfällen abgewichen, etwa wenn in einer Matrix sowohl feste als auch zufällige Komponenten enthalten sind.

Definition 2.1. *Folgende Bezeichnungen werden in der Arbeit verwendet:*

- (1) Sei $z := x + iy \in \mathbb{C}$ eine komplexe Zahl. Dann wird mit $\bar{z} := x - iy$ die **komplex konjugierte** Zahl bezeichnet.
- (2) Sei $Z \in \mathbb{C}^{m \times n}$ eine komplexe Matrix. Dann wird mit Z' die **Transponierte** und mit $Z^* := \bar{Z}'$ die **konjugiert Transponierte** oder **hermitesch Transponierte** bezeichnet.
- (3) Für eine quadratische Matrix $Z \in \mathbb{C}^{n \times n}$ bezeichne

$$\text{tr } Z := \sum_{i=1}^n z_{ii}$$

die **Spur** der Matrix.

- (4) Sei $x \in \mathbb{R}$ eine reelle Zahl. Dann wird mit

$$x^+ := \begin{cases} x, & \text{falls } x \geq 0 \\ 0, & \text{falls } x < 0 \end{cases}$$

der **Positivanteil** von x bezeichnet.

2.1. Grundlagen aus der Stochastik

Grundlegende Ergebnisse aus der Stochastik, etwa aus [18] oder [15], [16] werden in der Regel vorausgesetzt. Da in dieser Arbeit im wesentlichen mit komplexwertigen Zufallsvariablen und Zufallsvektoren gearbeitet wird, wird dieser Abschnitt dazu verwendet, einige Definitionen aus dem Reellen ins Komplexe zu erweitern, sowie einige Sätze, die benötigt werden, entsprechend anzupassen.

Für Zufallsvariablen wird jeweils eine vereinfachende bezeichnende Schreibweise verwendet. Die Bezeichnung $X \in \Omega'$ bedeutet dabei, dass X als Abbildung $X : \Omega \rightarrow \Omega'$ in den Raum Ω' abbildet, wobei in der Regel $\Omega' = \mathbb{R}$ oder $\Omega' = \mathbb{C}$ ist.

Definition 2.2. Gegeben sei eine Zufallsvariable $X \in \mathbb{C}$. Die **Varianz** ist definiert als

$$\text{Var}(X) := \mathbb{E}(X - \mathbb{E}X) \overline{(X - \mathbb{E}X)}.$$

Nach [16] gelten für die komplexe Varianz dieselben grundlegenden Rechenregeln wie für die reelle Varianz. Direkt aus der Definition kann man erkennen, dass die Varianz stets nichtnegativ ist. Insbesondere ist die Varianz einer komplexen Zufallsvariable hiermit stets reell. Hat man einen mehrdimensionalen Zufallsvektor gegeben, kann man die Kovarianzmatrix folgendermaßen definieren:

Definition 2.3. Gegeben sei ein Zufallsvektor $\underline{X} \in \mathbb{C}^k$. Die **Kovarianzmatrix** ist definiert als

$$\text{Cov}(\underline{X}) := \mathbb{E}(\underline{X} - \mathbb{E}\underline{X})(\underline{X} - \mathbb{E}\underline{X})^*.$$

Ist zusätzlich ein weiterer Zufallsvektor $\underline{Y} \in \mathbb{C}^l$ gegeben, so definiert man

$$\text{Cov}(\underline{X}, \underline{Y}) := \mathbb{E}(\underline{X} - \mathbb{E}\underline{X})(\underline{Y} - \mathbb{E}\underline{Y})^*.$$

Falls $k := l := 1$ gilt, vereinfacht sich diese Definition zu

$$\text{Cov}(X) := \mathbb{E}(X - \mathbb{E}X)(X - \mathbb{E}X)^* = \mathbb{E}(X - \mathbb{E}X) \overline{(X - \mathbb{E}X)}$$

bzw. zu

$$\text{Cov}(X, Y) := \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y)^* = \mathbb{E}(X - \mathbb{E}X) \overline{(Y - \mathbb{E}Y)}.$$

Man sieht, dass für jede Komponente aus der Kovarianzmatrix gilt

$$\text{Cov}(\underline{X}, \underline{Y})_{ij} = \text{Cov}(X_i, Y_j) := \mathbb{E}(X_i - \mathbb{E}X_i)(Y_j - \mathbb{E}Y_j)$$

für $1 \leq i \leq k$ und $1 \leq j \leq l$. Kovarianzmatrizen sind in der Regel nicht komponentenweise positiv. Man kann jedoch erkennen, dass Kovarianzmatrizen der Form $\text{Cov}(\underline{X})$ stets nichtnegativ definit sind. Für solche Matrizen gilt dann das folgende

Lemma 2.4. Gegeben sei eine nichtnegativ definite Matrix $A \in \mathbb{C}^{m \times m}$. Es ist

$$\det(A) \geq 0.$$

Beweis. Da A nichtnegativ definit ist, sind auch alle Eigenwerte nichtnegativ, und damit auch die Determinante als Produkt aller Eigenwerte. \square

Aus diesem Lemma folgt sofort das folgende

Korollar 2.5. Gegeben sei ein Zufallsvektor $\underline{X} \in \mathbb{C}^k$. Es ist

$$\det(\text{Cov}(\underline{X})) \geq 0.$$

Weiterhin sieht man, dass die Kovarianz im Komplexen nicht symmetrisch ist, sondern, dass das folgende Korollar gilt:

Korollar 2.6. Gegeben seien Zufallsvektoren $X, Y \in \mathbb{C}$. Es ist

$$\text{Cov}(X, Y) = \overline{\text{Cov}(Y, X)}.$$

Während man bei reellwertigen Zufallsvariablen X und Y den Korrelationskoeffizienten

$$\text{Corr}(X, Y) := \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var} X \text{Var} Y}}$$

definieren kann, ist diese Definition im Komplexen nicht sinnvoll, da dieser Ausdruck im Allgemeinen nicht reellwertig ist. Man möchte jedoch eine analoge Aussage zu der reellen Ungleichung

$$\text{Corr}(X, Y) \leq 1$$

erhalten. Diese findet sich im folgenden

Satz 2.7. Gegeben seien komplexwertige Zufallsvariablen X und Y . Dann ist

$$\begin{aligned} |\text{Cov}(X, Y)| &\leq \sqrt{\text{Var}(\Re X) \text{Var}(\Re Y)} + \sqrt{\text{Var}(\Re X) \text{Var}(\Im Y)} \\ &\quad + \sqrt{\text{Var}(\Im X) \text{Var}(\Re Y)} + \sqrt{\text{Var}(\Im X) \text{Var}(\Im Y)}. \end{aligned}$$

Beweis. Es gilt mit Zuhilfenahme der reellen Korrelationsungleichung

$$\begin{aligned} |\text{Cov}(X, Y)| &= |\text{Cov}(\Re X + i\Im X, \Re Y + i\Im Y)| \\ &\leq |\text{Cov}(\Re X, \Re Y)| + |\text{Cov}(\Re X, \Im Y)| \\ &\quad + |\text{Cov}(\Im X, \Re Y)| + |\text{Cov}(\Im X, \Im Y)| \\ &\leq \sqrt{\text{Var}(\Re X) \text{Var}(\Re Y)} + \sqrt{\text{Var}(\Re X) \text{Var}(\Im Y)} \\ &\quad + \sqrt{\text{Var}(\Im X) \text{Var}(\Re Y)} + \sqrt{\text{Var}(\Im X) \text{Var}(\Im Y)}. \end{aligned}$$

□

2.2. Informationstheoretische Grundlagen

Unter einem Kanal ist eine Vorrichtung zum Übertragen von Daten über eine räumliche Distanz zu verstehen. Mathematisch ist der Kanal eindeutig durch Übergangswahrscheinlichkeiten zwischen den Input- und Outputsymbolen festgelegt. Ebenso wie die weiter unten folgenden Definitionen der Entropie, der Transinformation und der Kanalkapazität, geht auch diese auf Claude Elwood Shannon zurück.

Definition 2.8. Bezeichne \mathcal{B}^n die Menge der n -dimensionalen Borel-Mengen. Gegeben sei eine Menge $X \in \mathcal{B}^k$ von **Inputsymbolen** und eine Menge $Y \in \mathcal{B}^l$ von **Outputsymbolen**. Ein **Kanal** auf diesen beiden Mengen wird definiert durch eine Übergangswahrscheinlichkeitsverteilung $f(\underline{x}, \underline{y}) \geq 0$ für alle $\underline{x} \in X$ und $\underline{y} \in Y$, so dass

$$\int_Y f(\underline{x}, \underline{y}) d\underline{y} = 1$$

für alle $\underline{x} \in X$ gilt.

Im Verlauf dieser Arbeit werden Kanäle niemals durch Übergangswahrscheinlichkeiten charakterisiert, sondern stets durch Funktionen. Dabei schreibt man

$$\underline{Y} := F(\underline{X}),$$

wobei \underline{X} ein gemäß einer gewissen Inputverteilung ausgewürfeltes Inputsymbol ist. In der Regel sind Komponenten der Funktion F ebenfalls zufallsabhängig.

Beispiel 2.9. Seien $X := [0, 1]$ und $Y := [-1, 2]$. Man definiere

$$f(x, y) := \begin{cases} \frac{1}{2}, & \text{falls } |x - y| \leq 2 \\ 0, & \text{falls } |x - y| > 2 \end{cases}$$

für $x \in X$ und $y \in Y$. Dann ist

$$\int_Y f(x, y) dy = \int_{-1}^2 f(x, y) dy = \int_{x-1}^{x+1} \frac{1}{2} dy = 1$$

für alle $x \in X$. Diesen Kanal kann man auch ausdrücken in der Schreibweise

$$Y := X + Z,$$

wobei X Rechteckverteilt ist mit $X \sim U[-1, 1]$.

Im Verlauf der Arbeit werden nur absolut-stetige Kanäle betrachtet. Das heißt, im Gegensatz zu diskreten Kanälen sind die Inputverteilung und Übergangswahrscheinlichkeitsverteilungen absolut-stetig. Diskrete Kanäle werden z.B. in [11] ausführlich behandelt. Dieser Abschnitt soll dazu dienen, sich die grundlegenden Definitionen aus der Informationstheorie ins Gedächtnis zu rufen, die wichtigsten Sätze der Vorlesung "Informationstheorie II" [12] zusammenzufassen und zu verallgemeinern (z.B. auf komplexwertige Zufallsvektoren). Ebenso beinhaltet dieser Abschnitt einige Ergebnisse, die nicht in [12] enthalten sind, aber eng in diesen Zusammenhang gehören.

Definition 2.10 (Entropie). Gegeben sei eine absolut-stetige (komplexwertige) Zufallsvariable X mit Dichte $f(x)$. Falls $f(x) \log f(x)$ integrierbar ist, definiert man die (**differentielle**) **Entropie** von X als

$$H(X) := - \int_{\mathbb{C}} f(x) \log f(x) dx,$$

wobei $0 \log(0) := 0$ gesetzt wird.

Sei $\underline{X} = (X_1, \dots, X_n)$ ein Zufallsvektor mit λ^n -Dichte $f(\underline{x}) = f(x_1, \dots, x_n)$, so heißt

$$H(\underline{X}) := H(X_1, \dots, X_n) := - \int_{\mathbb{C}^n} f(\underline{x}) \log f(\underline{x}) d\underline{x}$$

die **gemeinsame (differentielle) Entropie** von \underline{X} .

Sei $(\underline{X}, \underline{Y})$ ein Zufallsvektor mit λ^{m+n} -Dichte $f(\underline{x}, \underline{y})$ sowie bedingter Dichte

$$f(\underline{x} | \underline{y}) := \begin{cases} \frac{f(\underline{x}, \underline{y})}{f(\underline{y})}, & \text{falls } f(\underline{y}) > 0 \\ 0, & \text{falls } f(\underline{y}) = 0, \end{cases}$$

wobei $f(\underline{y})$ die Dichte von \underline{Y} sei. Definiere die **bedingte (differentielle) Entropie** als

$$H(\underline{X} | \underline{Y}) := - \int_{\mathbb{C}^m} \int_{\mathbb{C}^n} f(\underline{x}, \underline{y}) \log f(\underline{x} | \underline{y}) d\underline{y} d\underline{x}.$$

Definition 2.11 (Synentropie). Gegeben sei ein absolut-stetiger Zufallsvektor $(\underline{X}, \underline{Y})$ mit λ^{m+n} -Dichte $f(\underline{x}, \underline{y})$ sowie Randdichten $f(\underline{x})$ und $f(\underline{y})$. Dann heißt

$$I(\underline{X}; \underline{Y}) := \int_{\mathbb{C}^n} \int_{\mathbb{C}^m} f(\underline{x}, \underline{y}) \log \frac{f(\underline{x}, \underline{y})}{f(\underline{x}) f(\underline{y})} d\underline{x} d\underline{y}$$

die **Synentropie** oder **Transinformation** zwischen \underline{X} und \underline{Y} .

Sei $(\underline{X}, \underline{Y}, \underline{Z})$ ein Zufallsvektor mit λ^{l+m+n} -Dichte $f(\underline{x}, \underline{y}, \underline{z})$ sowie bedingten Dichten

$$f(\underline{x}, \underline{y} | \underline{z}) := \begin{cases} \frac{f(\underline{x}, \underline{y}, \underline{z})}{f(\underline{z})}, & \text{falls } f(\underline{z}) > 0 \\ 0, & \text{falls } f(\underline{z}) = 0, \end{cases}$$

$$f(\underline{x} | \underline{z}) := \begin{cases} \frac{f(\underline{x}, \underline{z})}{f(\underline{z})}, & \text{falls } f(\underline{z}) > 0 \\ 0, & \text{falls } f(\underline{z}) = 0 \end{cases}$$

und

$$f(\underline{y} | \underline{z}) := \begin{cases} \frac{f(\underline{y}, \underline{z})}{f(\underline{z})}, & \text{falls } f(\underline{z}) > 0 \\ 0, & \text{falls } f(\underline{z}) = 0, \end{cases}$$

wobei $f(\underline{z})$ die Dichte von \underline{Z} sei, $f(\underline{x}, \underline{z})$ die gemeinsame Dichte von $(\underline{X}, \underline{Z})$ und $f(\underline{y}, \underline{z})$ die von $(\underline{Y}, \underline{Z})$. Definiere die **bedingte Synentropie** oder **bedingte Transinformation** als

$$I(\underline{X}; \underline{Y} | \underline{Z}) := \int_{\mathbb{C}^l} \int_{\mathbb{C}^m} \int_{\mathbb{C}^n} f(\underline{x}, \underline{y}, \underline{z}) \log \frac{f(\underline{x}, \underline{y} | \underline{z})}{f(\underline{x} | \underline{z}) f(\underline{y} | \underline{z})} d\underline{z} d\underline{y} d\underline{x}.$$

Für die differentielle Entropie und Synentropie gelten verschiedene Gleichungen und Ungleichungen, die zusammengefasst werden im folgenden

Lemma 2.12. Seien \underline{X} und \underline{Y} absolut-stetige Zufallsvektoren. Dann gilt

(1)

$$I(\underline{X}; \underline{Y}) = H(\underline{X}) - H(\underline{X} | \underline{Y}) = H(\underline{Y}) - H(\underline{Y} | \underline{X}),$$

(2)

$$I(\underline{X}; \underline{Y} | \underline{Z}) = H(\underline{X} | \underline{Z}) - H(\underline{X} | \underline{Y}, \underline{Z}) = H(\underline{Y} | \underline{Z}) - H(\underline{Y} | \underline{X}, \underline{Z}),$$

(3)

$$I(\underline{X}; \underline{Y}) \geq 0$$

mit Gleichheit genau dann, wenn \underline{X} und \underline{Y} stochastisch unabhängig sind,

(4)

$$H(\underline{X} | \underline{Y}) \leq H(\underline{X})$$

mit Gleichheit genau dann, wenn \underline{X} und \underline{Y} stochastisch unabhängig sind (Shannon-Ungleichung),

(5)

$$H(X_1, \dots, X_n) = \sum_{i=1}^n H(X_i | X_1, \dots, X_{i-1}),$$

(6)

$$H(X_1, \dots, X_n) \leq \sum_{i=1}^n H(X_i),$$

Beweis. Siehe [12] (1), (3) – (6) bzw. [17] (2). □

Der folgende Satz ist eine Verallgemeinerung eines Satzes aus [12].

Satz 2.13. *Gegeben seien zwei absolut-stetige Zufallsvektoren $\underline{X} \in \mathbb{C}^m$ und $\underline{Y} \in \mathbb{C}^n$, eine invertierbare Matrix $S \in \mathbb{C}^{m \times m}$, sowie ein Vektor $\underline{\mu} \in \mathbb{C}^m$. Dann gilt*

$$H(S\underline{X} + \underline{\mu}) = H(\underline{X}) + \log |\det S|.$$

Analog gilt für die bedingte Entropie

$$H(S\underline{X} + \underline{\mu} | T\underline{Y}) = H(\underline{X} | \underline{Y}) + \log |\det S|$$

für eine invertierbare Matrix $T \in \mathbb{C}^{n \times n}$.

Beweis. Zunächst ist

$$\begin{aligned} H(\underline{X} + \underline{\mu}) &= - \int_{\mathbb{C}^m} f_{\underline{x}}(\underline{x} + \underline{\mu}) \log f_{\underline{x}}(\underline{x} + \underline{\mu}) d\underline{x} \\ &= - \int_{\mathbb{C}^m} f_{\underline{x}}(\underline{x}) \log f_{\underline{x}}(\underline{x}) d\underline{x} = H(\underline{X}). \end{aligned}$$

Zum Beweis der zweiten Aussage seien $\underline{X}' := S\underline{X}$ und $\underline{Y}' := T\underline{Y}$. Dann gilt für $B \subseteq \mathbb{C}^m$ und $C \subseteq \mathbb{C}^n$ beliebig,

$$P(\underline{X}' \in B, \underline{Y}' \in C) = P(S\underline{X} \in B, T\underline{Y} \in C) = P(\underline{X} \in S^{-1}B, \underline{Y} \in T^{-1}C).$$

Damit ergibt sich

$$\begin{aligned} \int_{S^{-1}B} \int_{T^{-1}C} f_{\underline{x}, \underline{y}}(\underline{x}', \underline{y}') d\underline{y}' d\underline{x}' &= \\ \int_B \int_C f_{\underline{x}, \underline{y}}(S^{-1}\underline{x}', T^{-1}\underline{y}') |\det S^{-1}| |\det T^{-1}| d\underline{y}' d\underline{x}' &. \end{aligned}$$

Also ist

$$f_{\underline{x}', \underline{y}'}(\underline{x}', \underline{y}') := \frac{1}{|\det S| |\det T|} f_{\underline{x}, \underline{y}}(S^{-1}\underline{x}', T^{-1}\underline{y}')$$

eine Dichte des Zufallsvektors $(\underline{X}', \underline{Y}')$. Für die bedingte Entropie ergibt sich damit

$$\begin{aligned}
H(S\underline{X} | T\underline{Y}) &= H(\underline{X}' | \underline{Y}') = - \int_{\mathbb{C}^n} \int_{\mathbb{C}^m} f_{\underline{x}', \underline{y}'}(\underline{x}', \underline{y}') \log f_{\underline{x}' | \underline{y}'}(\underline{x}' | \underline{y}') d\underline{x}' d\underline{y}' \\
&= - \int_{\mathbb{C}^n} \int_{\mathbb{C}^m} \frac{1}{|\det S| |\det T|} f_{\underline{x}, \underline{y}}(S^{-1}\underline{x}', T^{-1}\underline{y}') \log \frac{\frac{1}{|\det S| |\det T|} f_{\underline{x}', \underline{y}'}(S^{-1}\underline{x}', T^{-1}\underline{y}')}{\frac{1}{|\det T|} f_{\underline{y}}(T^{-1}\underline{y}')} d\underline{x}' d\underline{y}' \\
&= - \int_{\mathbb{C}^n} \int_{\mathbb{C}^m} \frac{1}{|\det S| |\det T|} f_{\underline{x}, \underline{y}}(S^{-1}\underline{x}', T^{-1}\underline{y}') \log \frac{f_{\underline{x} | \underline{y}}(S^{-1}\underline{x}' | T^{-1}\underline{y}')}{|\det S|} d\underline{x}' d\underline{y}' \\
&= - \int_{\mathbb{C}^n} \int_{\mathbb{C}^m} \frac{1}{|\det S| |\det T|} f_{\underline{x}, \underline{y}}(S^{-1}\underline{x}', T^{-1}\underline{y}') \log f_{\underline{x} | \underline{y}}(S^{-1}\underline{x}' | T^{-1}\underline{y}') d\underline{x}' d\underline{y}' \\
&\quad + \int_{\mathbb{C}^n} \int_{\mathbb{C}^m} \frac{1}{|\det S| |\det T|} f_{\underline{x}, \underline{y}}(S^{-1}\underline{x}', T^{-1}\underline{y}') \log |\det S| d\underline{x}' d\underline{y}' \\
&= - \int_{\mathbb{C}^n} \int_{\mathbb{C}^m} f_{\underline{x}, \underline{y}}(\underline{x}, \underline{y}) \log f_{\underline{x} | \underline{y}}(\underline{x} | \underline{y}) d\underline{x} d\underline{y} + \log |\det S| \\
&= H(\underline{X} | \underline{Y}) + \log |\det S|.
\end{aligned}$$

Die erste Aussage folgt aus der zweiten mit Hilfe der Shannon-Ungleichung. Denn man benötigt im Beweis die Darstellung von \underline{Y} nicht, und kann daher \underline{X} und \underline{Y} stochastisch unabhängig wählen. Damit folgt mit der Shannon-Ungleichung

$$H(S\underline{X}) = H(S\underline{X} | T\underline{Y}) = H(\underline{X} | \underline{Y}) + \log |\det S| = H(\underline{X}) + \log |\det S|.$$

□

Der folgende Satz stammt aus [13]. Dieses Ergebnis lässt sich ohne Probleme auch auf andere Fälle (z.B. diskrete Verteilungen) übertragen.

Satz 2.14. *Seien \underline{X} und \underline{Z} stochastisch unabhängige Zufallsvektoren, sowie \underline{Y} ein weiterer Zufallsvektor, im Allgemeinen abhängig von den vorigen beiden. Dann gilt*

$$I(\underline{X}; \underline{Y}) \leq I(\underline{X}; \underline{Y} | \underline{Z}).$$

Beweis. Es gilt

$$I(\underline{X}; \underline{Y}) = H(\underline{X}) - H(\underline{X} | \underline{Y})$$

sowie

$$I(\underline{X}; \underline{Y} | \underline{Z}) = H(\underline{X} | \underline{Z}) - H(\underline{X} | \underline{Y}, \underline{Z}).$$

Da \underline{X} und \underline{Z} stochastisch unabhängig sind, gilt

$$H(\underline{X} | \underline{Z}) = H(\underline{X}).$$

Damit folgt

$$I(\underline{X}; \underline{Y} | \underline{Z}) - I(\underline{X}; \underline{Y}) = H(\underline{X} | \underline{Y}) - H(\underline{X} | \underline{Y}, \underline{Z}) = I(\underline{X}; \underline{Z} | \underline{Y}) \geq 0.$$

□

2.3. Die komplexe Normalverteilung

In diesem Abschnitt werden die komplexe Normalverteilung definiert, sowie die Ergebnisse aus [12] und [20] zusammengefasst, die im Laufe der Arbeit benötigt werden.

Definition 2.15. (1) Der Zufallsvektor $\underline{X} = \underline{U} + i\underline{V} \in \mathbb{C}^n$ heißt **komplex normalverteilt**, falls $\begin{pmatrix} \underline{U} \\ \underline{V} \end{pmatrix} = \begin{pmatrix} \Re \underline{X} \\ \Im \underline{X} \end{pmatrix}$ $2n$ -dimensional normalverteilt ist.

(2) Der Zufallsvektor \underline{X} heißt **zirkulär symmetrisch** oder **zirkulär symmetrisch komplex normalverteilt**, falls \underline{X} symmetrisch komplex normalverteilt und

$$\text{Cov} \begin{pmatrix} \underline{U} \\ \underline{V} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \Re Q & -\Im Q \\ \Im Q & \Re Q \end{pmatrix}$$

für eine hermitesche nichtnegativ definite Matrix $Q \in \mathbb{C}^{n \times n}$ ist. Sei $\underline{\mu} := \mathbb{E} \underline{X}$. Dann schreibt man

$$\underline{X} \sim \text{SCN}(\underline{\mu}, Q).$$

Lemma 2.16. Sei $\underline{X} \sim \text{SCN}(\underline{\mu}, Q)$ mit Q invertierbar. Dann besitzt \underline{X} die Dichte

$$f(\underline{x}) = \frac{1}{\det(\pi Q)} \exp(-(\underline{x} - \underline{\mu})^* Q^{-1} (\underline{x} - \underline{\mu})), \underline{x} \in \mathbb{C}^n.$$

Beweis. Siehe [20]. □

Satz 2.17. Sei \underline{X} eine absolut-stetige komplexe Zufallsvariable mit $\mathbb{E} \underline{X} = \underline{0}$ und $\text{Cov} \underline{X} = Q$. Dann gilt

$$H(\underline{X}) \leq \log \det(\pi e Q)$$

mit Gleichheit genau dann, wenn $\underline{X} \sim \text{SCN}(\underline{0}, Q)$.

Beweis. Siehe [20]. □

Lemma 2.18. (1) Seien $\underline{X} \sim \text{SCN}(\underline{\mu}, Q)$ und $A \in \mathbb{C}^{m \times n}$. Dann gilt

$$A\underline{X} \sim \text{SCN}(A\underline{\mu}, AQA^*).$$

(2) Seien $\underline{X} \sim \text{SCN}(\underline{\mu}_1, Q_1)$ und $\underline{Y} \sim \text{SCN}(\underline{\mu}_2, Q_2)$ stochastisch unabhängig. Dann gilt

$$\underline{X} + \underline{Y} \sim \left(\underline{\mu}_1 + \underline{\mu}_2, Q_1 + Q_2 \right).$$

Insbesondere sind Linearkombinationen von symmetrisch komplex normalverteilten Zufallsvariablen wieder symmetrisch komplex normalverteilt.

Beweis. Siehe [20]. □

Die SCN-Verteilung hat die Eigenschaft, dass man die bedingte Erwartung von gemeinsam SCN-verteilten Vektoren explizit angeben kann. Die Rechnung ist pure Matrix-Algebra, deshalb wird hier nur das Ergebnis angegeben.

Satz 2.19. *Seien \underline{X} und \underline{Y} symmetrisch komplex normalverteilt mit gewissen Parametern. Dann ist*

$$E(\underline{Y} | \underline{X}) = E\underline{Y} + \text{Cov}(\underline{Y}, \underline{X}) \text{Cov}(\underline{X})^{-1} (\underline{X} - E\underline{X}).$$

Beweis. Siehe [7], Kapitel 4.8. □

2.4. MMSE-Schätzer

Im Folgenden werden einige Ergebnisse über MMSE (minimum mean square error) Schätzer zusammengefasst, die in Kapitel 5 benötigt werden. Univariate analoge Aussagen können in [16] nachgeschlagen werden, die hier zitierten multivariaten Aussagen werden in [1] bzw. [7] bewiesen, oder können in anderer Literatur über Schätztheorie nachgeschlagen werden.

Definition 2.20. *Gegeben sei ein statistisches Experiment $\mathcal{E} = (\mathcal{X}, \mathcal{C}, \mathcal{F})$, $\mathcal{F} = \{P_{\underline{\theta}} | \underline{\theta} \in \Theta\}$, $\Theta \subseteq \mathbb{C}^k$ mit einer Schätzfunktion $\underline{\delta}: (\mathcal{X}, \mathcal{C}) \rightarrow (\Theta, \sigma(\Theta))$ für $\underline{\theta}$.*

- (1) Die **mittlere quadratische Abweichung** des Schätzers ist definiert als

$$\text{MSE}_{\underline{\theta}}(\underline{\delta}) := E((\underline{\delta}(\underline{X}) - \underline{\theta})(\underline{\delta}(\underline{X}) - \underline{\theta})^*).$$

- (2) Falls für alle Schätzer $\underline{\delta}'$ gilt $\text{MSE}_{\underline{\theta}}(\underline{\delta}') \geq \text{MSE}_{\underline{\theta}}(\underline{\delta})$, so ist $\underline{\delta}$ ein **MMSE-Schätzer** für $\underline{\theta}$.

Satz 2.21. *Gegeben sei ein statistisches Experiment wie oben in Definition 2.20. Dann gilt:*

- (1) Der MMSE-Schätzer für $\underline{\theta}$ ist gegeben durch

$$\underline{\delta}(\underline{X}) = E(\underline{\theta} | \underline{X}).$$

- (2) Dieser Schätzer ist erwartungstreu.
 (3) Der Fehler eines MMSE-Schätzers ist orthogonal zum zugehörigen Datenvektor (Orthogonalitätsprinzip), d.h.

$$E((\underline{\theta} - \underline{\delta})(\underline{\theta} - \underline{\delta})^*) = 0.$$

Diese Aussage gilt sogar für jede Funktion abhängig von den Daten, insbesondere also für den Schätzer selbst.

Beweis. Siehe [7], Kapitel 4.11. □

KAPITEL 3

Das MIMO-Modell

Das Modell baut auf dem Modell auf, wie es auch von Emre Telatar [20] und in der Vorlesung “Informationstheorie II” von Rudolf Mathar [12] verwendet wird.

Anschaulich kann man sich ein MIMO-Netz so vorstellen, dass es eine gewisse Anzahl von Sende- und Empfangsantennen gibt, und von jeder Sendeantenne zu jeder Empfangsantenne Daten gesendet werden (Abbildung 3.1).

Im Folgenden bezeichne \underline{X} den Inputvektor, \mathcal{H} eine multiplikative Fading-Matrix, \underline{Z} eine additive Störung, und \underline{Y} den Outputvektor. Solch ein Übertragungssystem wird in Abbildung 3.2 gezeigt. Bei komponentenweiser Betrachtung wird das Diagramm wie in Abbildung 3.3 erweitert.

Für eine Definition des MIMO-Modells werden vorerst noch keine weiteren Annahmen über die Kanalmatrix \mathcal{H} getroffen. Sie kann entweder zufällig, bekannt oder partiell bekannt sein. Dies wird in den folgenden Kapiteln nach Bedarf angepasst.

Modell 3.1. Es seien $t, r \in \mathbb{N}$, die Anzahl der Sende (transmit)- bzw. Empfangs (receive)-Antennen. Sei $\mathcal{H} \in \mathbb{C}^{r \times t}$ eine in der Regel zufällige Fading-Matrix, $\underline{X} \in \mathbb{C}^t$ ein zufälliges Inputsymbol, $\underline{Z} \in \mathbb{C}^r, \underline{Z} \sim \text{SCN}(\underline{0}, Q)$ die additive Störung und $P > 0$ die Leistungsbeschränkung. Es gelte $\mathbb{E}(\underline{X}^* \underline{X}) = \text{tr}(\mathbb{E}(\underline{X} \underline{X}^*)) \leq P$. Die Zufallsvektoren $\underline{X} \in \mathbb{C}^t$ und $\underline{Z} \in \mathbb{C}^r$ und die Zufallsmatrix \mathcal{H} seien gemeinsam stochastisch unabhängig.

ABBILDUNG 3.1. Darstellung eines MIMO-Kanals

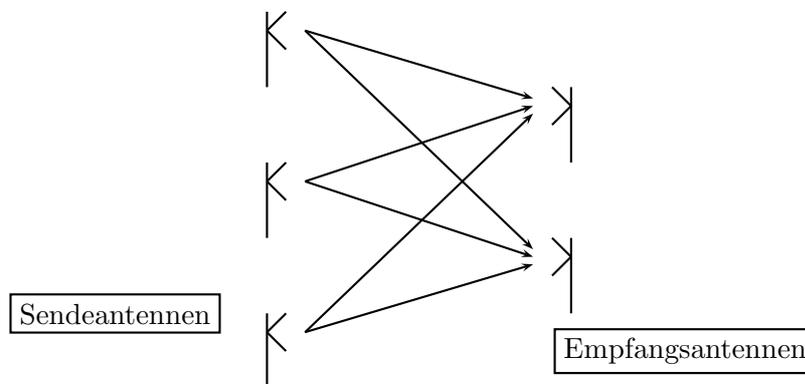


ABBILDUNG 3.2. Übertragungsmodell eines MIMO-Kanals

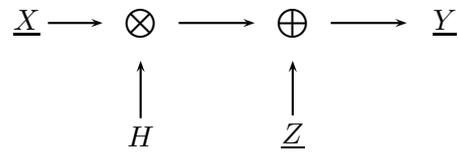
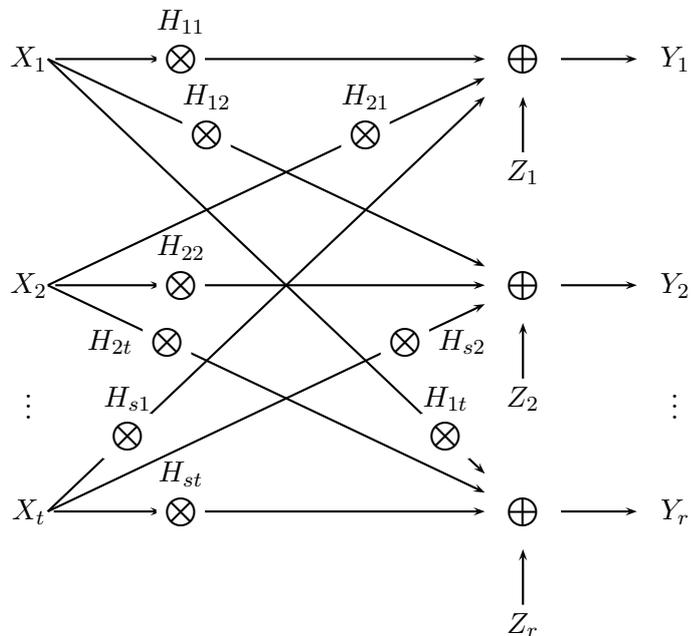


ABBILDUNG 3.3. Komponentenweises Übertragungsmodell eines MIMO-Kanals



Der Outputvektor $\underline{Y} \in \mathbb{C}^r$ ist definiert durch

$$\underline{Y} := \mathcal{H}\underline{X} + \underline{Z}.$$

Eine Interpretation aller verwendeten Symbole ist in Tabelle 3.1 wiedergegeben.

Die zweite wichtige Definition in dieser Arbeit ist die der Kanalkapazität, die für verschiedene MIMO-Modelle ausgerechnet bzw. falls das nicht möglich ist, abgeschätzt werden soll. Auch diese Definition muss in einigen der folgenden Kapitel verändert und leicht an die Gegebenheiten des jeweiligen Modells angepasst werden.

TABELLE 3.1. Interpretation der Symbole im MIMO-Modell

Mathematische Definition	Interpretation
$r \in \mathbb{N}$	Empfangsantennen (receive)
$t \in \mathbb{N}$	Sendeantennen (transmit)
$H \in \mathbb{C}^{r \times t}$	Kanalmatrix
$P > 0$	Leistungsbeschränkung
$\underline{X} \in \mathbb{C}^t$	Inputvektor
$\underline{Y} \in \mathbb{C}^r$	Outputvektor
$\underline{Z} \in \mathbb{C}^r$	additive Störung

Definition 3.2 (Kapazität). Die **Kapazität eines Kanals** ist definiert als

$$C = \max_{\mathbb{E}(\underline{X}^* \underline{X}) \leq P} I(\underline{X}; \underline{Y}),$$

die maximal erreichbare Transinformation unter allen Inputverteilungen unter der Leistungsbeschränkung P .

Es gibt verschiedene Arten von MIMO-Kanälen, die danach unterschieden werden, wie viel über die Kanalmatrix \mathcal{H} beim Empfänger bekannt ist. Falls \mathcal{H} beim Empfänger bekannt ist, schreibt man für \mathcal{H} auch H , und der Kanal wird mit C-CSI (“completely known channel state information”) bezeichnet (Kapitel 4). Falls \mathcal{H} nur partiell bekannt ist, etwa durch Schätzung, wird der Kanal mit P-CSI (“partially known channel state information”) bezeichnet (Kapitel 5). Wenn weder beim Sender noch beim Empfänger die Kanalmatrix bekannt ist, bezeichnet man den Kanal mit N-CSI (“no channel state information”). Solche Modelle werden in dieser Arbeit nicht behandelt. In [9] wird die Synentropie dieses Kanals angegeben. Die Maximierung ist ein offenes Problem.

Man unterscheidet ebenfalls zwischen dem Kanalwissen an Sender- und Empfängerseite.

Komplett bekannte Kanäle

Dieses Kapitel beschäftigt sich mit komplett bekannten Kanälen (C-CSI, “completely known channel state information”). Im Abschnitt 4.1 wird dabei zunächst auf die Kanalkapazität bei deterministischen Kanalmatrizen eingegangen, in Abschnitt 4.2 bei zufälligen. Die Ergebnisse aus [20] und [12] sind durch weitere Ergebnisse fortgesetzt und vor allem durch eine Verallgemeinerung der additiven Störung erweitert worden, sowie durch einige Beispiele und Grenzwertbetrachtungen veranschaulicht.

4.1. Herleitung der Kapazität bei fester Kanalmatrix

In diesem Abschnitt wird der Fall betrachtet, dass H deterministisch ist, d.h. man setzt überall $H \in \mathbb{C}^{r \times t}$ fest vorgegeben voraus. Dieses Szenario ist gleichbedeutend damit, dass die Realisation der zufälligen Kanalmatrix H bereits beim Sender bekannt ist.

Modell 4.1. *Es seien die Voraussetzungen wie in Modell 3.1. Dabei sei $H \in \mathbb{C}^{r \times t}$ fest.*

Der Outputvektor $\underline{Y} \in \mathbb{C}^r$ ist definiert durch

$$\underline{Y} := H\underline{X} + \underline{Z}.$$

Definition 4.2 (Erweiterte Diagonalmatrix). *Seien $\xi_1, \dots, \xi_{\min(r,t)} \in \mathbb{C}$. Dann wird die **erweiterte Diagonalmatrix** $\Xi \in \mathbb{C}^{r \times t}$ definiert mit den Einträgen*

$$\xi_{ij} := \begin{cases} \xi_i, & \text{falls } i = j \\ 0, & \text{falls } i \neq j \end{cases}, 1 \leq i \leq r, 1 \leq j \leq t.$$

Schreibweise: $\Xi =: (\xi_{ij})_{1 \leq i \leq r, 1 \leq j \leq t} =: \text{diag}_{r \times t}(\xi_1, \dots, \xi_{\min(r,t)})$.

Beispiel 4.3. *Es ist*

$$\text{diag}_{3 \times 7}(1, 2, 3) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Die Kapazität wird zunächst für den Fall berechnet, dass H eine erweiterte Diagonalmatrix ist. Der allgemeine Fall wird auf diesen Spezialfall zurückgeführt. Die Lösung erfolgt über “Water-Filling”, so ähnlich wie bei parallelen Gauß-Kanälen (siehe z.B. [12]). Die hier gewählte Beweismethode ist zwar etwas aufwändiger als die in [12], macht aber die Verbindung des MIMO-Modells zu

den parallelen Gauß-Kanälen deutlicher. Weiterhin ist die Aussage der Sätze verallgemeinert worden. In [12] und [20] war für die additive Störung nur die Einheitsmatrix als Kovarianzmatrix zugelassen, also weißes Rauschen. Neu ist, dass hier auch additive Störungen mit beliebiger Kovarianzmatrix möglich sind.

Lemma 4.4. *Seien $n \in \mathbb{N}$ und $\lambda_1 \dots \lambda_n > 0$. Dann wird $\prod_{i=1}^n (x_i + \lambda_i)$ unter der Bedingung $\sum_{i=1}^n x_i \leq P$ maximal, wenn $x_i = (\mu - \lambda_i)^+$, wobei μ die eindeutige Lösung von $\sum_{i=1}^n (\mu - \lambda_i)^+ = P$ ist.*

Beweis. Siehe [12]. □

Lemma 4.5 (Hadamard-Ungleichung). *Im Folgenden finden sich zwei Spezialfälle der Hadamard-Ungleichung.*

(1) *Für jede symmetrische nichtnegative reelle Matrix A gilt*

$$\det A \leq \prod_{i=1}^n a_{ii}.$$

(2) *Für jede hermitesche Matrix A gilt*

$$|\det A| \leq \prod_{i=1}^n |a_{ii}|.$$

Beweis. Siehe [12]. □

Satz 4.6 (“Water-Filling”). *Gegeben sei ein Kanal wie in Modell 4.1 mit*

$$\underline{Z} \sim \text{SCN}(\underline{0}, Q)$$

für $Q \in \mathbb{C}^{r \times r}$ hermitesch und nichtnegativ definit, sowie der Einschränkung

$$H := \Lambda := \text{diag}_{r \times t} \left(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_{\min\{t,r\}}} \right)$$

für $\lambda_i \geq 0, 1 \leq i \leq \min\{t,r\}$. Es ist somit

$$\underline{Y} := \Lambda \underline{X} + \underline{Z}.$$

Dieser Kanal hat die Kapazität

$$C = \max_{\mathbb{E}(\underline{X}^* \underline{X}) \leq P} \mathbb{I}(\underline{X}; \underline{Y}) = \sum_{i=1}^{\min\{t,r\}} (\log(\mu \lambda_i))^+ - \log \det Q,$$

wobei μ so gewählt, dass

$$P = \sum_{i=1}^{\min\{t,r\}} \left(\mu - \frac{q_{ii}}{\lambda_i} \right)^+.$$

Beweis. Ohne Einschränkung seien die λ_i sogar echt positiv, da sonst die zugehörigen X_i keine Auswirkung auf den Kanal haben. Zunächst gilt für die Synentropie

$$\begin{aligned} I(\underline{X}; \underline{Y}) &= H(\underline{Y}) - H(\underline{Y} | \underline{X}) \\ &= H(\underline{Y}) - H(\Lambda \underline{X} + \underline{Z} | \underline{X}) \\ &= H(\underline{Y}) - H(\underline{Z} | \underline{X}) = H(\underline{Y}) - H(\underline{Z}). \end{aligned}$$

Es ist

$$\begin{aligned} \text{Cov}(\underline{Y}) &= \text{Cov}(\Lambda \underline{X} + \underline{Z}) \\ &= \Lambda \text{Cov}(\underline{X}) \Lambda' + \text{Cov}(\underline{Z}) \\ &= \Lambda \text{Cov}(\underline{X}) \Lambda' + Q =: \Lambda \Sigma_X \Lambda' + Q. \end{aligned}$$

Dann gilt $H(\underline{Y}) \leq \log \det(\pi e (\Lambda \Sigma_X \Lambda' + Q))$ mit Gleichheit, genau dann wenn $\underline{Y} \sim \text{SCN}(\underline{0}, \Lambda \Sigma_X \Lambda' + Q)$, also $\underline{X} \sim \text{SCN}(\underline{0}, \Sigma_X)$. Diese obere Schranke bleibt zu maximieren, d.h.

$$\max_{\text{E}(\underline{X}^* \underline{X}) \leq P} \det(\Lambda \Sigma_X \Lambda' + Q) = \max_{\text{tr}(\Sigma_X) \leq P} \det(\Lambda \Sigma_X \Lambda' + Q).$$

Nach der Hadamard-Ungleichung 4.5 ist $\det(B) \leq \prod_{i=1}^n b_{ii}$, falls $B \geq 0$. In diesem Fall hat man also

$$\det(\Lambda \Sigma_X \Lambda' + Q) \leq \prod_{i=1}^{\min\{t,r\}} (\lambda_i \text{Var}(X_i) + q_{ii})$$

mit Gleichheit, falls $\Lambda \Sigma_X \Lambda' + Q$ Diagonalform hat. Es bleibt somit zu lösen $\max \prod_{i=1}^{\min\{t,r\}} (\lambda_i \text{Var}(X_i) + q_{ii})$ unter der Bedingung $\sum_{i=1}^{\min\{t,r\}} \text{Var}(X_i) \leq P$. Dazu betrachte man zunächst den Fall, dass $Q = I_r$ ist. Man benutze Lemma 4.4 sowie $\lambda_i \text{Var}(X_i) + 1 = \lambda_i (\text{Var}(X_i) + \lambda_i^{-1})$ und erhalte die Lösung $\text{Var}(X_i) = (\mu - \lambda_i^{-1})^+$. Für beliebige Kovarianzmatrizen Q folgt unter Zuhilfenahme dieses Ergebnisses aus der Gleichung

$$\begin{aligned} \prod_{i=1}^{\min\{t,r\}} (\lambda_i \text{Var}(X_i) + q_{ii}) &= \prod_{i=1}^{\min\{t,r\}} q_{ii} \left(\frac{\lambda_i}{q_{ii}} \text{Var}(X_i) + 1 \right) \\ &= \left(\prod_{i=1}^{\min\{t,r\}} q_{ii} \right) \left(\prod_{i=1}^{\min\{t,r\}} \frac{\lambda_i}{q_{ii}} \text{Var}(X_i) + 1 \right) \end{aligned}$$

die Lösung

$$\text{Var}(X_i) = \left(\mu - \frac{q_{ii}}{\lambda_i} \right)^+.$$

Die Kapazität errechnet sich als

$$\begin{aligned}
C &= \max_{\text{tr}(\Sigma_X) \leq P} \mathbb{H}(\underline{Y}) - \mathbb{H}(\underline{Z}) \\
&= \log \left(\pi e \left(\prod_{i=1}^{\min\{t,r\}} \lambda_i \right) \mu^{\min\{t,r\}} \right) - \log(\pi e \det Q) \\
&= \sum_{i=1}^{\min\{t,r\}} (\log(\mu \lambda_i)) - \log \det Q.
\end{aligned}$$

Falls einige $\lambda_i = 0$ sind, folgt die Lösung mit der folgenden Rechnung:

$$\begin{aligned}
\prod_{i=1}^{\min\{t,r\}} (\lambda_i \text{Var}(X_i) + q_{ii}) &= \prod_{\substack{i=1 \\ \lambda_i \neq 0}}^{\min\{t,r\}} (\lambda_i \text{Var}(X_i) + q_{ii}) \\
&= \prod_{\substack{i=1 \\ \lambda_i \neq 0}}^{\min\{t,r\}} (\lambda_i (\mu - \lambda_i^{-1})^+ + q_{ii}) \\
&= \prod_{\substack{i=1 \\ \lambda_i \neq 0}}^{\min\{t,r\}} ((\lambda_i \mu - 1)^+ + q_{ii}) \\
&= \prod_{\substack{i=1 \\ \lambda_i \neq 0}}^{\min\{t,r\}} (\lambda_i \mu)^+.
\end{aligned}$$

Mit der Schreibweise $(\log(0))^+ = -\infty^+ = 0$ hat man damit auch

$$\begin{aligned}
\log \left(\prod_{i=1}^{\min\{t,r\}} (\lambda_i \text{Var}(X_i) + q_{ii}) \right) &= \log \left(\prod_{\substack{i=1 \\ \lambda_i \neq 0}}^{\min\{t,r\}} (\lambda_i \mu)^+ \right) \\
&= \sum_{\substack{i=1 \\ \lambda_i \neq 0}}^{\min\{t,r\}} (\log(\lambda_i \mu))^+ = \sum_{i=1}^{\min\{t,r\}} (\log(\lambda_i \mu))^+
\end{aligned}$$

Damit ist insgesamt

$$C = \sum_{i=1}^{\min\{t,r\}} (\log(\mu \lambda_i))^+ - \log \det Q.$$

Abschließend gilt für die maximierende Inputverteilung, dass Real- und Imaginärteile jeder einzelnen Komponente unabhängig identisch verteilt sind. Daraus ergibt sich

$$\mathbb{E}(\Re(X_i)^2) = \mathbb{E}(\Im(X_i)^2) = \frac{1}{2} (\mu - \lambda_i^{-1})^+.$$

□

Den allgemeinen Fall führt man auf diesen zurück, indem man die Kanalmatrix durch eine Singulärwertzerlegung ausdrückt, und sich neue Input- und Outputvektoren für die neue Kanalmatrix zusammensetzt. Dann kann man zeigen, dass die beiden Kanäle die gleiche Kapazität besitzen.

Satz 4.7 (Singulärwertzerlegung, SVD). *Sei $A \in \mathbb{C}^{r \times t}$ beliebig. Dann existieren unitäre Matrizen $U \in \mathbb{C}^{r \times r}$, $V \in \mathbb{C}^{t \times t}$ und reelle Zahlen $\lambda_1, \dots, \lambda_{\min(r,t)} \geq 0$, so dass $A = U\Lambda V^*$, wo $\Lambda = \text{diag}_{r \times t}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_{\min(r,t)}})$. Dabei sind die Spalten von U die Eigenvektoren von AA^* , die Spalten von V die Eigenvektoren von A^*A und die λ_i die Eigenwerte von AA^* .*

Beweis. Siehe [5], Kapitel 5.4. □

Satz 4.8 (Kanalkapazität). *Der Kanal aus Modell 4.1 mit deterministischer Kanalmatrix H und additivem Fehler $\underline{Z} \sim \text{SCN}(\underline{0}, Q)$ für beliebiges $Q \in \mathbb{C}^{r \times r}$ hermitesch und nichtnegativ definit besitzt die Kapazität*

$$C = \sum_{i=1}^{\min(r,t)} (\log(\mu \lambda_i))^+ - \log \det Q,$$

wo $\mu \in \mathbb{R}$ die eindeutige reelle Zahl ist mit

$$P = \sum_{i=1}^{\min(r,t)} \left(\mu - \frac{q_{ii}}{\lambda_i} \right)^+.$$

Die λ_i sind dabei die Eigenwerte von HH^* .

Beweis. Nach Satz 4.7 existiert eine Singulärwertzerlegung $H = U\Lambda V^*$ mit $U \in \mathbb{C}^{r \times r}$, $V \in \mathbb{C}^{t \times t}$ unitär, $\Lambda = \text{diag}_{r \times t}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_{\min(r,t)}})$. Damit ergibt sich die Darstellung

$$\underline{Y} = H\underline{X} + \underline{Z} = U\Lambda V^*\underline{X} + \underline{Z}. \quad (4.1)$$

Seien $\tilde{\underline{Y}} := U^*\underline{Y}$, $\tilde{\underline{X}} := V^*\underline{X}$ und $\tilde{\underline{Z}} := U^*\underline{Z}$. Damit ist $\tilde{\underline{Z}}$ zirkulär symmetrisch verteilt als Linearkombination von zirkulär symmetrisch verteilten Zufallsvariablen.

Aus (4.1) folgt

$$\tilde{\underline{Y}} = U^*\underline{Y} = U^*U\Lambda V^*\underline{X} + U^*\underline{Z} = \Lambda\tilde{\underline{X}} + \tilde{\underline{Z}}, \quad (4.2)$$

wobei

$$\tilde{\underline{Z}} \sim \text{SCN}(U^*\underline{0}, UQU^*) = \text{SCN}(\underline{0}, Q).$$

Komponentenweises Betrachten von (4.2) liefert

$$\tilde{Y}_i = \sqrt{\lambda_i}\tilde{X}_i + \tilde{Z}_i, \quad 1 \leq i \leq \min\{r, t\}.$$

Die Matrizen U und V sind invertierbar, und es gilt

$$\mathbb{E}(\tilde{\underline{X}}^*\tilde{\underline{X}}) = \mathbb{E}((V\underline{X})^*V\underline{X}) = \mathbb{E}(\underline{X}^*V^*V\underline{X}) = \mathbb{E}(\underline{X}^*\underline{X}).$$

Die Gleichheit der Kapazitäten ergibt sich mit Hilfe von Satz 2.13 schließlich aus

$$\begin{aligned} \max_{\mathbb{E}(\tilde{\mathbf{X}}^* \tilde{\mathbf{X}}) \leq P} \mathbf{I}(\tilde{\mathbf{X}}; \tilde{\mathbf{Y}}) &= \max_{\mathbb{E}(\tilde{\mathbf{X}}^* \tilde{\mathbf{X}}) \leq P} \left(\mathbf{H}(\tilde{\mathbf{X}}) - \mathbf{H}(\tilde{\mathbf{X}} | \tilde{\mathbf{Y}}) \right) \\ &= \max_{\mathbb{E}(\tilde{\mathbf{X}}^* \tilde{\mathbf{X}}) \leq P} \left(\mathbf{H}(\tilde{\mathbf{X}}) + \log |\det U| - \mathbf{H}(\tilde{\mathbf{X}} | \tilde{\mathbf{Y}}) - \log |\det U| \right) \\ &= \max_{\mathbb{E}(\tilde{\mathbf{X}}^* \tilde{\mathbf{X}}) \leq P} \left(\mathbf{H}(\tilde{\mathbf{X}}) - \mathbf{H}(\tilde{\mathbf{X}} | \tilde{\mathbf{Y}}) \right) = \max_{\mathbb{E}(\tilde{\mathbf{X}}^* \tilde{\mathbf{X}}) \leq P} \mathbf{I}(\tilde{\mathbf{X}}; \tilde{\mathbf{Y}}) \end{aligned}$$

Aus der komponentenweisen Darstellung des neuen Kanals sieht man, dass für $i > \min\{t, r\}$ gilt $\tilde{Y}_i = \tilde{Z}_i$, also ist hier \tilde{Y}_i unabhängig von \tilde{X}_i .

Nach Satz 4.6 kann man die Kapazität des neuen Kanals mit “Water-Filling” angeben:

$$C = \max_{\mathbb{E}(\tilde{\mathbf{X}}^* \tilde{\mathbf{X}}) \leq P} \mathbf{I}(\tilde{\mathbf{X}}; \tilde{\mathbf{Y}}) = \max_{\mathbb{E}(\tilde{\mathbf{X}}^* \tilde{\mathbf{X}}) \leq P} \mathbf{I}(\tilde{\mathbf{X}}; \tilde{\mathbf{Y}}) = \sum_{i=1}^{\min\{t, r\}} (\log(\mu \lambda_i))^+ - \log \det Q$$

mit

$$P = \sum_{i=1}^{\min\{t, r\}} \left(\mu - \frac{q_{ii}}{\lambda_i} \right)^+.$$

□

Aus den beiden Sätzen 4.6 und 4.8 folgen sofort die Ergebnisse aus [12] und [20], die im Folgenden wiedergegeben sind.

Korollar 4.9. *Gegeben sei ein Kanal wie in Modell 4.1 mit*

$$\underline{\mathbf{Z}} \sim \text{SCN}(\underline{\mathbf{0}}, I_r)$$

und der Einschränkung

$$H := \Lambda := \text{diag}_{r \times t} \left(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_{\min\{t, r\}}} \right)$$

für $\lambda_i \geq 0, 1 \leq i \leq \min\{t, r\}$. Es ist somit

$$\underline{\mathbf{Y}} := \Lambda \underline{\mathbf{X}} + \underline{\mathbf{Z}}.$$

Dieser Kanal hat die Kapazität

$$C = \max_{\mathbb{E}(\tilde{\mathbf{X}}^* \tilde{\mathbf{X}}) \leq P} \mathbf{I}(\tilde{\mathbf{X}}; \tilde{\mathbf{Y}}) = \sum_{i=1}^{\min\{t, r\}} (\log(\mu \lambda_i))^+,$$

wobei μ so gewählt wird, dass

$$P = \sum_{i=1}^{\min\{t, r\}} (\mu - \lambda_i^{-1})^+.$$

Beweis. Es ist

$$\log \det I_r = \log 1 = 0.$$

Die Behauptung folgt aus Satz 4.6. □

Mit demselben Argument erhält man das folgende

Korollar 4.10. *Der Kanal aus Modell 4.1 mit deterministischer Kanalmatrix H und additivem Fehler $\underline{Z} \sim \text{SCN}(\underline{0}, I_r)$ besitzt die Kapazität*

$$C = \sum_{i=1}^{\min(r,t)} (\log(\mu\lambda_i))^+,$$

wo $\mu \in \mathbb{R}$ die eindeutige reelle Zahl ist mit

$$P = \sum_{i=1}^{\min(r,t)} (\mu - \lambda_i^{-1})^+.$$

Die λ_i sind dabei die Eigenwerte von HH^* .

Korollar 4.11. *Die Kapazität ändert sich nicht, wenn die Kanalmatrix H^* anstatt H lautet.*

Beweis. Die Matrizen H^*H und HH^* besitzen die gleichen nichtnegativen Eigenwerte. \square

Beispiel 4.12. *Sei $H = \mathbf{1}_{r \times t}$, d.h. $H_{ij} = 1, 1 \leq i \leq r, 1 \leq j \leq t$. Die SVD sieht folgendermaßen aus:*

$$H = \begin{pmatrix} \sqrt{1/r} & \star & \dots & \star \\ \vdots & \vdots & & \vdots \\ \sqrt{1/r} & \star & \dots & \star \end{pmatrix}_{r \times r} \begin{pmatrix} \sqrt{rt} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}_{r \times t} \begin{pmatrix} \sqrt{1/t} & \dots & \sqrt{1/t} \\ \star & \dots & \star \\ \vdots & & \vdots \\ \star & \dots & \star \end{pmatrix}_{t \times t},$$

wobei die Einträge \star uninteressant sind, da nur ein Eigenwert $\neq 0$ ist.

Nach Satz 4.8 folgt also

$$P = \sum_{i=1}^{\min\{t,r\}} (\mu - \lambda_1^{-1})^+ = \mu - \lambda_1^{-1} = \mu - \frac{1}{rt}$$

und damit die Kapazität

$$C = \sum_{i=1}^{\min\{t,r\}} (\log(\mu\lambda_i))^+ = \log(\mu\lambda_1) = \log(\mu rt) = \log(1 + rtP).$$

Für das \underline{X} , das die Synentropie maximiert, ergibt sich damit

$$\mathbb{E}(X_i \bar{X}_j) = \frac{P}{t},$$

denn

$$\begin{aligned} \text{Cov}(\underline{X}) &= \text{Cov}(V\tilde{X}) = V \text{Cov}(\tilde{X}) V^* \\ &= \begin{pmatrix} \sqrt{1/r} & \star & \dots & \star \\ \vdots & \vdots & & \vdots \\ \sqrt{1/r} & \star & \dots & \star \end{pmatrix} \begin{pmatrix} P & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} \sqrt{1/t} & \dots & \sqrt{1/t} \\ \star & \dots & \star \\ \vdots & & \vdots \\ \star & \dots & \star \end{pmatrix} = \frac{P}{t} 1_{t \times t} \end{aligned}$$

Beispiel 4.13. Sei $r = t =: n$ und $H := I_n$. Die SVD ist trivial: $H = I_n I_n I_n$. Damit ergibt sich sofort

$$P = \sum_{i=1}^n (\mu - \lambda_i^{-1})^+ = n(\mu - 1)^+ = n(\mu - 1)$$

und

$$C = \sum_{i=1}^n (\log \mu \lambda_i)^+ = \sum_{i=1}^n (\log \mu)^+ = n(\log \mu)^+ = n \log \left(\frac{P}{n} + 1 \right).$$

Da $V^* = I_n$, folgt sofort $\underline{X} = \tilde{X}$, und für das maximierende \underline{X} gilt damit

$$\begin{aligned} \mathbb{E}(X_i \bar{X}_j) &= \mathbb{E}(\tilde{X}_i \tilde{X}_j) = \delta_{ij} \left(\mathbb{E} \left(\Re(\tilde{X}_i)^2 \right) + \mathbb{E} \left(\Im(\tilde{X}_i)^2 \right) \right) \\ &= \delta_{ij} (\mu - \lambda_i^{-1}) = \frac{\delta_{ij} P}{n}. \end{aligned}$$

4.2. Herleitung der Kapazität bei zufälliger Kanalmatrix

Dieser Abschnitt behandelt den Fall, dass $H := \mathcal{H} := (H_{ij}) \in \mathbb{C}^{r \times t}$ stochastisch ist. Weiterhin sei angenommen, dass die Realisationen H von \mathcal{H} beim Empfänger bekannt sind, nicht aber beim Sender. Es werden die Hauptergebnisse aus [20] und [12] zusammengefasst, sowie einige Spezialfälle genauer betrachtet und Grenzwertbetrachtungen gemacht.

Modell 4.14. Es seien die Voraussetzungen aus Modell 3.1 gegeben. Es sei $H := \mathcal{H} := (H_{ij}) \in \mathbb{C}^{r \times t}$ eine Zufallsmatrix, die stochastisch unabhängig von \underline{X} und \underline{Z} ist. Der Kanal ist gegeben durch den Input \underline{X} und den Output

$$(\underline{Y}, \mathcal{H}) = (\mathcal{H}\underline{X} + \underline{Z}, \mathcal{H}).$$

Die Einträge H_{ij} von \mathcal{H} seien untereinander stochastisch unabhängig, und haben stochastisch unabhängige Real- und Imaginärteile mit jeweils $\Re(H_{ij}) \sim N(0, 1/2)$ und $\Im(H_{ij}) \sim N(0, 1/2)$.

Lemma 4.15. Gegeben seien \mathcal{H} wie oben, $U \in \mathbb{C}^{r \times r}$ und $V \in \mathbb{C}^{t \times t}$ unitär. Dann besitzt $U\mathcal{H}V^*$ dieselbe Verteilung wie \mathcal{H} .

Satz 4.16 (Kanalkapazität). Die Kapazität des obigen Kanals ist gegeben durch

$$C = \max_{\mathbb{E}(\underline{X}^* \underline{X}) \leq P} \mathbb{I}(\underline{X}; \underline{Y}, \mathcal{H}) = \mathbb{E} \left(\log \det \left(I_r + \frac{P}{t} \mathcal{H} \mathcal{H}^* \right) \right).$$

Beweis. Für die Synentropie gilt

$$\begin{aligned}
I(\underline{X}; \underline{Y}, \mathcal{H}) &= I(\underline{X}; \mathcal{H}) + I(\underline{X}; \underline{Y} \mid \mathcal{H}) = I(\underline{X}; \underline{Y} \mid \mathcal{H}) \\
&= \mathbb{E}_{\mathcal{H}} (I(\underline{X}; \underline{Y} \mid \mathcal{H} = H)) \\
&= \mathbb{E}_{\mathcal{H}} (H(\underline{Y}) - H(\mathcal{H}\underline{X} + \underline{Z} \mid \underline{X})) \\
&= \mathbb{E}_{\mathcal{H}} (H(\underline{Y}) - H(\underline{Z})) \\
&\leq \mathbb{E}_{\mathcal{H}} (\log \det (I_r + \mathcal{H}Q\mathcal{H}^*))
\end{aligned}$$

mit $Q = \mathbb{E}(\underline{X}\underline{X}^*) = \text{Cov}(\underline{X})$ mit Gleichheit, wenn $\underline{X} \sim \text{SCN}(\underline{0}, Q)$. Dieser Ausdruck ist zu maximieren unter den Bedingungen $\text{tr} Q \leq P$ und $Q \geq 0$.

Dazu sei \mathcal{P}_t der Raum aller $t \times t$ Permutationsmatrizen. Nach [12] gilt, dass $\mathcal{H}\Pi$ und \mathcal{H} dieselbe Verteilung besitzen für jede Permutationsmatrix $\Pi \in \mathcal{P}_t$. Damit folgt mit der Bezeichnung $Q^\Pi := \Pi Q \Pi'$ für den obigen Ausdruck

$$\begin{aligned}
\mathbb{E}(\log \det (I_r + \mathcal{H}Q\mathcal{H}^*)) &= \mathbb{E}(\log \det (I_r + \mathcal{H}\Pi Q \Pi' \mathcal{H}^*)) \\
&= \mathbb{E}(\log \det (I_r + \mathcal{H}Q^\Pi \mathcal{H}^*)) \\
&= \frac{1}{t!} \sum_{\Pi \in \mathcal{P}_t} \mathbb{E}(\log \det (I_r + \mathcal{H}Q^\Pi \mathcal{H}^*)) \\
&= \mathbb{E} \left(\frac{1}{t!} \sum_{\Pi \in \mathcal{P}_t} \log \det (I_r + \mathcal{H}Q^\Pi \mathcal{H}^*) \right) \\
&\leq \mathbb{E} \left(\log \det \left(I_r + \mathcal{H} \left(\frac{1}{t!} \sum_{\Pi} Q^\Pi \right) \mathcal{H}^* \right) \right),
\end{aligned}$$

da $\Phi(Q) := \log \det (I_r + \mathcal{H}Q\mathcal{H}^*)$ nach [12] konkav auf der Menge der nichtnegativ definiten Matrizen ist.

Die Matrix $\frac{1}{t!} \sum_{\Pi \in \mathcal{P}_t} Q^\Pi = \alpha I_t$ ist ein Vielfaches der Einheitsmatrix. Insgesamt gilt

$$\begin{aligned}
\max_{\text{tr} Q \leq P, Q \geq 0} \mathbb{E}(\log \det (I_r + \mathcal{H}Q\mathcal{H}^*)) &\leq \max_{t\alpha \leq P} \mathbb{E}(\log \det (I_r + \mathcal{H}\alpha I_t \mathcal{H}^*)) \\
&= \mathbb{E} \left(\log \det \left(I_r + \frac{P}{t} \mathcal{H}\mathcal{H}^* \right) \right).
\end{aligned}$$

Somit hat man auch die Gleichheit. Das Maximum wird angenommen, falls $\underline{X} \sim \text{SCN}(\underline{0}, \frac{P}{t} I_t)$. \square

An dieser Stelle endet die Zusammenfassung von [12] und [20], und es folgen einige Spezialfälle und Grenzwertbetrachtungen. Die Ergebnisse aus Satz 4.16 lassen sich für $t = 1$ bzw. $r = 1$ relativ leicht ausrechnen. Das Ergebnis ist das folgende

Beispiel 4.17. Gegeben sei der obige Kanal.

(1) Für den Fall $r = 1$ ist die Kapazität gegeben durch

$$C = \int_0^\infty \log \left(1 + \frac{Pm}{4t} \right) \frac{1}{2^t \Gamma(t)} m^{t-1} e^{-\frac{m}{2}} dm.$$

(2) Für den Fall $t = 1$ ist die Kapazität gegeben durch

$$C = \int_0^\infty \log \left(1 + \frac{Pm}{4} \right) \frac{1}{2^r \Gamma(r)} m^{r-1} e^{-\frac{m}{2}} dm.$$

Beweis. (1) Sei $r = 1$. Die Kapazität errechnet sich als

$$\begin{aligned} C &= \mathbb{E} \left(\log \det \left(I_r + \frac{P}{t} \mathcal{H} \mathcal{H}^* \right) \right) \\ &= \mathbb{E} \left(\log \left(1 + \frac{P}{t} \sum_{i=1}^t H_{1i} \bar{H}_{1i} \right) \right) \\ &= \mathbb{E} \left(\log \left(1 + \frac{P}{t} \sum_{i=1}^t \Re(H_{1i})^2 + \Im(H_{1i})^2 \right) \right) \\ &= \mathbb{E} \left(\log \left(1 + \frac{P}{4t} \sum_{i=1}^{2t} N_i^2 \right) \right) \quad \text{mit } N_i \sim N(0, 1) \text{ iid} \\ &= \mathbb{E} \left(\log \left(1 + \frac{P}{4t} M \right) \right) \quad \text{mit } M \sim \chi^2(2t) \\ &= \int_0^\infty \log \left(1 + \frac{Pm}{4t} \right) \frac{1}{2^t \Gamma(t)} m^{t-1} e^{-\frac{m}{2}} dm \end{aligned}$$

(2) Sei $t = 1$. Die Kapazität errechnet sich als

$$\begin{aligned} C &= \mathbb{E} \left(\log \det \left(I_r + \frac{P}{t} \mathcal{H} \mathcal{H}^* \right) \right) \\ &= \mathbb{E} (\log \det (I_t + P \mathcal{H}^* \mathcal{H})) \\ &= \mathbb{E} \left(\log \left(1 + P \sum_{i=1}^r \bar{H}_{i1} H_{i1} \right) \right) \\ &= \mathbb{E} \left(\log \left(1 + P \sum_{i=1}^r \Re(H_{i1})^2 + \Im(H_{i1})^2 \right) \right) \\ &= \mathbb{E} \left(\log \left(1 + \frac{P}{4} \sum_{i=1}^{2r} N_i^2 \right) \right) \quad \text{mit } N_i \sim N(0, 1) \text{ iid} \\ &= \mathbb{E} \left(\log \left(1 + \frac{P}{4} M \right) \right) \quad \text{mit } M \sim \chi^2(2r) \\ &= \int_0^\infty \log \left(1 + \frac{Pm}{4} \right) \frac{1}{2^r \Gamma(r)} m^{r-1} e^{-\frac{m}{2}} dm \end{aligned}$$

□

Man betrachte in (1) den Grenzwert $t \rightarrow \infty$, und in (2) den Grenzwert $r \rightarrow \infty$. Dann ergibt sich bei (1) mit dem starken Gesetz großer Zahlen, und bei (2) mit der Bezeichnung $M(2r) \sim \chi^2(2r)$

(1)

$$\begin{aligned}
\lim_{t \rightarrow \infty} \mathbb{E} \left(\log \det \left(1 + \frac{P}{t} \mathcal{H} \mathcal{H}^* \right) \right) &= \lim_{t \rightarrow \infty} \mathbb{E} \left(\log \left(1 + \frac{P}{4t} \sum_{i=1}^{2t} N_i^2 \right) \right) \\
&= \mathbb{E} \left(\log \left(1 + \lim_{t \rightarrow \infty} \frac{P}{4t} \sum_{i=1}^{2t} N_i^2 \right) \right) \\
&= \mathbb{E} \left(\log \left(1 + \frac{P}{2} \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{i=1}^t N_i^2 \right) \right) \\
&= \log \left(1 + \frac{P}{2} \right),
\end{aligned}$$

(2)

$$\begin{aligned}
\lim_{r \rightarrow \infty} \mathbb{E} (\log \det (1 + P \mathcal{H} \mathcal{H}^*)) &= \lim_{r \rightarrow \infty} \mathbb{E} \left(\log \left(1 + \frac{P}{4} \sum_{i=1}^{2r} N_i^2 \right) \right) \\
&\geq \lim_{r \rightarrow \infty} \mathbb{E} \left(\log \left(\frac{P}{4} \sum_{i=1}^{2r} N_i^2 \right) \right) \\
&= \lim_{r \rightarrow \infty} \mathbb{E} (\log M(2r)) + \log \left(\frac{P}{4} \right) \\
&\geq \lim_{r \rightarrow \infty} \log 2r + \log \left(\frac{P}{4} \right) = \infty,
\end{aligned}$$

und damit die Gleichheit. D.h. bei wachsender Anzahl von Empfangsantennen r und einer Sendeantenne $t = 1$ steigt die Kapazität des Kanals ungefähr wie

$$C \approx \log \left(\frac{2rP}{4} \right).$$

In der Praxis ist es eher nicht anzunehmen, dass bei wachsender Anzahl von Sendeantennen die Leistungsbeschränkung konstant bleibt. Man setze in (1) $P := P(t) := tc$ für ein konstantes $c > 0$ und berechne den Grenzwert neu. Dabei ergibt sich mit den bekannten Bezeichnungen

$$\begin{aligned}
\lim_{t \rightarrow \infty} \mathbb{E} \left(\log \det \left(1 + \frac{P(t)}{t} \mathcal{H} \mathcal{H}^* \right) \right) &= \lim_{t \rightarrow \infty} \mathbb{E} (\log \det (1 + c \mathcal{H} \mathcal{H}^*)) \\
&= \lim_{t \rightarrow \infty} \mathbb{E} \left(\log \left(1 + \frac{c}{4} \sum_{i=1}^{2t} N_i^2 \right) \right) \\
&\geq \lim_{t \rightarrow \infty} \mathbb{E} \left(\log \left(\frac{c}{4} \sum_{i=1}^{2t} N_i^2 \right) \right) \\
&= \lim_{t \rightarrow \infty} \mathbb{E} (\log M(2t)) + \log \left(\frac{c}{4} \right) \\
&\geq \lim_{t \rightarrow \infty} \log 2t + \log \left(\frac{c}{4} \right) = \infty.
\end{aligned}$$

Man sieht, dass in diesem Fall die Kapazität bei sehr großen t ungefähr steigt wie

$$C \approx \log \left(\frac{2tc}{4} \right).$$

Man sieht, dass die Kapazität bei einer Sendeantenne und wachsender Anzahl Empfangsantennen ungefähr genau so stark wächst wie bei einer Empfangsantenne und wachsender Anzahl von Sendeantennen. Denn für $t = 1$ und $r \rightarrow \infty$ gilt mit den neuen Bezeichnungen

$$C \approx \log \left(\frac{2rP(t)}{4} \right) = \log \left(\frac{2rtc}{4} \right) = \log \left(\frac{2rc}{4} \right).$$

4.3. Weitere Modellmöglichkeiten

In der Literatur werden viele weitere Modelle vorgeschlagen, so etwa in [10]. Dort wird grundsätzlich von Modell 4.14 ausgegangen, jedoch werden über die gleiche Kanalmatrix $N \in \mathbb{N}$ Datenvektoren übertragen. Dieses Modell ist realistisch, wenn die Kanalmatrix in der Praxis über N Zeiteinheiten annähernd konstant bleibt. Die N Inputvektoren werden zu einer Matrix $\mathcal{X} \in \mathbb{C}^{t \times N}$ zusammengefasst, ebenso die additiven Störungsvektoren zu $\mathcal{Z} \in \mathbb{C}^{r \times N}$, sowie die Outputvektoren zu $\mathcal{Y} \in \mathbb{C}^{r \times N}$. Das Übertragungsmodell ist damit

$$\mathcal{Y} = \mathcal{X} \mathcal{H} + \mathcal{Z}.$$

Das wichtigste Ergebnis des Papers ist, dass die Kapazität für $t > N$ gleich der Kapazität für $t = N$ ist, also durch Aufstocken der Sendeantennen auf mehr als N -Stück kein Kapazitätsgewinn mehr erzielt wird.

Eine weitere interessante Möglichkeit wird in [4] vorgestellt. Es wird ebenfalls das Modell 4.14 als Grundlage genommen, jedoch sind die Einträge der Kanalmatrix hier untereinander korreliert. Die Anzahl der Sendeantennen sei gleich der Anzahl der Empfangsantennen, $n := t = r$. Die Hauptaussage des Papers ist, dass die Kapazität dann linear in n steigt. In der Praxis wird dieses Modell relevant, wenn die Antennen im Sendearray bzw. im Empfangsarray nahe beieinander sind. Dadurch sind die einzelnen Datenströme untereinander korreliert.

Empfängerseitig partiell bekannte Kanäle

5.1. Modellierung von Synchronised Detection

Die Erweiterung des letzten Modells durch Synchronised Detection beinhaltet die Schätzung der zufälligen Kanalmatrix durch vorangehende Testsymbole. Der wesentliche Unterschied zum vorhergehenden Modell ist dabei, dass die Kanalmatrix \mathcal{H} beim Empfänger nicht bekannt ist, sondern geschätzt wird. Dieses Modell baut auf [2] und [3] auf. Diese beiden Papers liefern kein mathematisches Modell für Synchronised Detection. Dieser Umstand wird behoben, und ein passendes Modell hier nachgeliefert. Auf weitere Fehler in den Papern wird an jeweils passender Stelle hingewiesen, und die Unklarheiten beseitigt. Außerdem werden die Ergebnisse aus [3] derart erweitert, dass auch eine obere Schranke für die Kanalkapazität folgert werden kann.

Modell 5.1. *Es seien $r, t, N_D, N_T \in \mathbb{N}$ und*

$$N := N_D + N_T.$$

Sei $\mathcal{H}_i^D \in \mathbb{C}^{r \times t}, i = 1, \dots, N_D$ eine iid-Folge von Zufallsmatrizen, wobei für jeden Eintrag

$$H_{ijk}^D \sim \text{SCN}(0, \sigma_h^2)$$

für ein $\sigma_h^2 > 0$ gelte. Die einzelnen Einträge sind dabei stochastisch unabhängig. Sei $\mathcal{H}_j^T \in \mathbb{C}^{r \times t}, j = 1, \dots, N_T$ eine weitere Folge mit den gleichen Eigenschaften. Im Allgemeinen ist die Folge \mathcal{H}_j^T dabei stochastisch abhängig von \mathcal{H}_i^D . Es sei $N_0 > 0, \underline{Z}_i^D$ iid mit

$$\underline{Z}_i^D \sim \text{SCN}(\underline{0}, N_0 I_r), i = 1, \dots, N_D$$

unabhängig von den oben definierten \mathcal{H}_i^D und \mathcal{H}_j^T . Es sei $\underline{Z}_j^T \in \mathbb{C}^r$ eine davon unabhängige Folge mit den gleichen Eigenschaften. Sei $\underline{X}_i^D \in \mathbb{C}^t, i = 1, \dots, N_D$ iid, unabhängig von \mathcal{H}_i^D und \mathcal{H}_j^T sowie \underline{Z}_i^D und \underline{Z}_j^T . Seien $\underline{X}_j^T \in \mathbb{C}^t, j = 1, \dots, N_T$ fest vorgegeben. Sei $P > 0$, und es gelte

$$\mathbb{E}(\underline{X}_1^{D*} \underline{X}_1^D) \leq P$$

sowie

$$\underline{X}_j^{T*} \underline{X}_j^T \leq P, j = 1, \dots, N_T.$$

*Definiere den **Datenkanal** durch*

$$\underline{Y}_i^D := \mathcal{H}_i^D \underline{X}_i^D + \underline{Z}_i^D, i = 1, \dots, N_D, \quad (5.1)$$

*sowie den **Testkanal** durch*

$$\underline{Y}_j^T := \mathcal{H}_j^T \underline{X}_j^T + \underline{Z}_j^T, j = 1, \dots, N_T. \quad (5.2)$$

Weiterhin seien

$$\begin{aligned}\hat{\mathcal{H}}_i^D &:= \hat{\mathcal{H}}_i^D(\underline{X}_1^T, \dots, \underline{X}_{N_T}^T, \underline{Y}_1^T, \dots, \underline{Y}_{N_T}^T) \\ &:= \mathbb{E}(\mathcal{H}_i^D \mid \underline{X}_1^T, \dots, \underline{X}_{N_T}^T, \underline{Y}_1^T, \dots, \underline{Y}_{N_T}^T), i = 1, \dots, N_D\end{aligned}$$

MMSE-Schätzer für \mathcal{H}_i^D .

Alle in diesem Modell verwendeten Symbole sind in Tabelle 5.1 mit ihrer Interpretation zusammengefasst.

Technische Realisierung:

In der Realität sind die Kanalmatrizen im Allgemeinen korreliert in der Zeit, d.h. es liegt ein Prozess \mathcal{H}_k , $k = 1, \dots, N$ mit Autokorrelationsfunktion

$$\alpha_n^* := \mathbb{E}(\mathcal{H}_{k,i,j} \bar{\mathcal{H}}_{k+n,i,j}), n = 1, \dots, N,$$

sonst aber gleichen Eigenschaften, zugrunde. Dieses Problem wird durch so genanntes Interleaving und De-Interleaving umgangen. Dazu werden die Daten- und Testsymbole eines Blockes der Länge N in einen Interleaver \mathcal{I} gegeben, der sie auf Komponentenebene so permutiert, dass die Komponenten eines Inputvektors so weit auseinander gezogen werden, dass sie über Kanalmatrizen mit Autokorrelation $\alpha_n^* < \varepsilon$ übertragen werden, dabei aber auch über jede Kanalmatrix genügend Testsymbole zum Schätzen geschickt werden. Im Modell kann man dann wegen der geringen Korrelation sogar stochastische Unabhängigkeit voraussetzen. Auf der Empfängerseite werden die übertragenen Symbole durch den De-Interleaver \mathcal{D} wieder nach Daten- und Testsymbolen sortiert. Dieser Vorgang ist in Abbildung 5.1 dargestellt. Auf eine Modellierung des Interleaving wird in dieser Arbeit verzichtet. Das Kanalmodell

$$\underline{Y}_k := \mathcal{H}_k \underline{X}_k + \underline{Z}_k$$

wird auch als innerer Kanal bezeichnet, und der Datenkanal

$$\underline{Y}_i^D := \mathcal{H}_i^D \underline{X}_i^D + \underline{Z}_i^D$$

als äußerer Kanal. In dieser Arbeit wird auf die gesamte Modellierung des inneren Kanals verzichtet, und die Kanalmatrizen von vornherein als stochastisch unabhängig vorausgesetzt.

Im Folgenden ist es rechentechnisch oft einfacher, die Werte der Inputverteilung in Matrixform und die Fading-Koeffizienten in Vektorform vorliegen zu haben. Zu einer einfachen Darstellung der dazu nötigen Transformationen wird das Kronecker-Produkt zweier Matrizen eingeführt.

Definition 5.2. Gegeben seien die Matrizen $A \in \mathbb{C}^{m \times n}$ und $B \in \mathbb{C}^{p \times q}$. Das **Kronecker-Produkt** $A \otimes B \in \mathbb{C}^{mp \times nq}$ ist definiert als

$$A \otimes B := \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}.$$

ABBILDUNG 5.1. Übertragungsmodell eines MIMO-Kanals mit Synchronised Detection

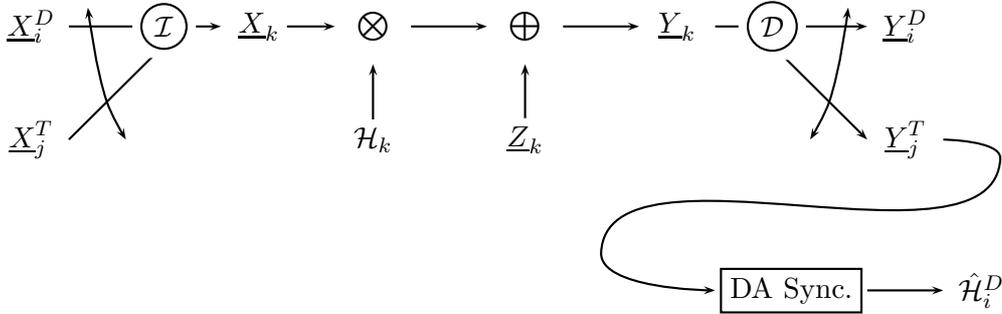


TABELLE 5.1. Interpretation der Symbole im MIMO-Modell mit Synchronised Detection

Mathematische Definition	Interpretation
$r \in \mathbb{N}$	Empfangsantennen (receive)
$t \in \mathbb{N}$	Sendeantennen (transmit)
$N_D \in \mathbb{N}$	Anzahl Datensymbole pro Block
$N_T \in \mathbb{N}$	Anzahl Testsymbole pro Block
$N \in \mathbb{N}$	Blocklänge
$P > 0$	Leistungsbeschränkung
$\underline{X}_i^D \in \mathbb{C}^t$	i -tes gesendetes Datensymbol (Datenkanal)
$\underline{Y}_i^D \in \mathbb{C}^r$	i -tes empfangenes Datensymbol (Datenkanal)
$\mathcal{H}_i^D \in \mathbb{C}^{r \times t}$	Kanalmatrix zur Zeit i (Datenkanal)
$\underline{Z}_i^D \in \mathbb{C}^r$	additive Störung zur Zeit i (Datenkanal)
$\underline{X}_j^T \in \mathbb{C}^t$	j -tes gesendetes Testsymbol (Testkanal)
$\underline{Y}_j^T \in \mathbb{C}^r$	j -tes empfangenes Testsymbol (Testkanal)
$\mathcal{H}_j^T \in \mathbb{C}^{r \times t}$	Kanalmatrix zur Zeit j (Testkanal)
$\underline{Z}_j^T \in \mathbb{C}^r$	additive Störung zur Zeit j (Testkanal)
$\hat{\mathcal{H}}_i^D \in \mathbb{C}^{r \times t}$	Schätzer für Kanalmatrix zur Zeit i
$\sigma_h^2 > 0$	Fading-Varianz
$N_0 > 0$	Varianz der additiven Störung

Sei also

$$\mathcal{X}_i^D := \underline{X}_i^{D'} \otimes I_r := (X_{i1}^D I_r, \dots, X_{it}^D I_r) \in \mathbb{C}^{r \times rt}, i = 1, \dots, N_D$$

die **Signalmatrix** und

$$\underline{H}_i^D := (H_{i11}^D, H_{i21}^D, \dots, H_{irt}^D)' \in \mathbb{C}^{rt}, i = 1, \dots, N_D$$

der **Kanalvektor** mit allen Spalten von \mathcal{H}_i^D übereinander.

Damit lässt sich der Datenkanal umschreiben zu

$$\underline{Y}_i^D = \mathcal{X}_i^D \underline{H}_i^D + \underline{Z}_i^D, i = 1, \dots, N_D. \quad (5.3)$$

Seien

$$\begin{aligned} \underline{Y}^D &:= (\underline{Y}_1^D, \dots, \underline{Y}_{N_D}^D)' \in \mathbb{C}^{rN_D}, \\ \underline{H}^D &:= (\underline{H}_1^D, \dots, \underline{H}_{N_D}^D)' \in \mathbb{C}^{rtN_D}, \\ \underline{Z}^D &:= (\underline{Z}_1^D, \dots, \underline{Z}_{N_D}^D)' \in \mathbb{C}^{rN_D} \end{aligned}$$

und

$$\mathcal{X}^D := \begin{pmatrix} \mathcal{X}_1^D & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \mathcal{X}_{N_D}^D \end{pmatrix} \in \mathbb{C}^{rN_D \times rtN_D}.$$

Damit ergibt sich das Modell für einen Block von N_D Datensymbolen

$$\underline{Y}^D = \mathcal{X}^D \underline{H}^D + \underline{Z}^D. \quad (5.4)$$

Analog definiere für den Testkanal

$$\mathcal{X}_j^T := \underline{X}_j^{T'} \otimes I_r := (X_{j1}^T I_r, \dots, X_{jt}^T I_r) \in \mathbb{C}^{r \times rt}, j = 1, \dots, N_T$$

und

$$\underline{H}_j^T := (H_{j11}^T, H_{j21}^T, \dots, H_{jrt}^T)' \in \mathbb{C}^{rt}, j = 1, \dots, N_T,$$

so dass

$$\underline{Y}_j^T = \mathcal{X}_j^T \underline{H}_j^T + \underline{Z}_j^T, j = 1, \dots, N_T. \quad (5.5)$$

Sei

$$\underline{Y}^T := (\underline{Y}_1^T, \dots, \underline{Y}_{N_T}^T)' \in \mathbb{C}^{rN_T}$$

der Vektor mit allen Outputvektoren $\underline{Y}_j^T, j = 1, \dots, N_T$ übereinander. Analog definiere die zugehörigen

$$\begin{aligned} \underline{H}^T &:= (\underline{H}_1^T, \dots, \underline{H}_{N_T}^T)' \in \mathbb{C}^{rtN_T}, \\ \underline{Z}^T &:= (\underline{Z}_1^T, \dots, \underline{Z}_{N_T}^T)' \in \mathbb{C}^{rN_T} \end{aligned}$$

und

$$\mathcal{X}^T := \begin{pmatrix} \mathcal{X}_1^T & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \mathcal{X}_{N_T}^T \end{pmatrix} \in \mathbb{C}^{rN_T \times rtN_T},$$

so dass

$$\underline{Y}^T = \mathcal{X}^T \underline{H}^T + \underline{Z}^T. \quad (5.6)$$

Außerdem definiere man für die geschätzten Kanalmatrizen analog

$$\hat{\underline{H}}_i^D := (\hat{H}_{i11}^D, \hat{H}_{i21}^D, \dots, \hat{H}_{irt}^D)' \in \mathbb{C}^{rt}, i = 1, \dots, N_D$$

und

$$\hat{\underline{H}}^D := (\hat{\underline{H}}_1^D, \dots, \hat{\underline{H}}_{N_D}^D)' \in \mathbb{C}^{rtN_D}.$$

TABELLE 5.2. Zusätzliche Symbole im MIMO-Modell mit Synchronised Detection

Mathematische Definition	Interpretation
$\mathcal{X}_i^D \in \mathbb{C}^{r \times rt}$	i -te gesendete Signalmatrix (Datenkanal)
$\underline{H}_i^D \in \mathbb{C}^{rt}$	Kanalvektor zur Zeit i (Datenkanal)
$\mathcal{X}_j^T \in \mathbb{C}^{r \times rt}$	j -te gesendete Signalmatrix (Testkanal)
$\underline{H}_j^T \in \mathbb{C}^{rt}$	Kanalvektor zur Zeit j (Testkanal)
$\hat{\underline{H}}_i^D \in \mathbb{C}^{rt}$	Schätzer für Kanalvektor zur Zeit i
$\mathcal{X}^D \in \mathbb{C}^{rN_T \times rtN_D}$	Signalmatrizen nebeneinander (Datenkanal)
$\underline{Y}^D \in \mathbb{C}^{rN_D}$	Outputvektoren übereinander (Datenkanal)
$\underline{H}^D \in \mathbb{C}^{rtN_D}$	Kanalvektoren übereinander (Datenkanal)
$\underline{Z}^D \in \mathbb{C}^{rN_D}$	Additive Störungen übereinander (Datenkanal)
$\mathcal{X}^T \in \mathbb{C}^{rN_T \times rtN_T}$	Signalmatrizen nebeneinander (Testkanal)
$\underline{Y}^T \in \mathbb{C}^{rN_T}$	Outputvektoren übereinander (Testkanal)
$\underline{H}^T \in \mathbb{C}^{rtN_T}$	Kanalvektoren übereinander (Testkanal)
$\underline{Z}^T \in \mathbb{C}^{rN_T}$	Additive Störungen übereinander (Testkanal)
$\hat{\underline{H}}^D \in \mathbb{C}^{rtN_D}$	Schätzer für die Kanalvektoren übereinander
$C > 0$	Kanalkapazität

Alle zusätzlichen Symbole, die nicht in Tabelle 5.1 vorkommen, sind noch einmal in der Tabelle 5.2 zusammengefasst.

Bei der Definition der Kapazität dieses Kanals ist zu beachten, dass von den N Symbolen nur N_D wirkliche Daten enthalten. Deshalb wird die Kapazität folgendermaßen definiert:

Definition 5.3. Gegeben sei der obige Kanal. Die Kapazität ist definiert als

$$C := \frac{1}{N} \max_{\mathbb{E}(\underline{X}^D \star \underline{X}^D) \leq N_D P} \mathbb{I}(\underline{Y}^D, \underline{Y}^T; \underline{X}^D | \underline{X}^T).$$

Das gegebene \underline{X}^T ist hierbei eine Schreibweise. Es handelt sich nicht um die bedingte Synentropie, da \underline{X}^T eine bekannte Größe ist, und nicht zufallsabhängig. In den Rechnungen kann \underline{X}^T auch als einpunktverteilte Zufallsvariable betrachtet werden. Da \mathcal{X}^D die gleichen Informationen enthält wie \underline{X}^D bzw. \mathcal{X}^T wie \underline{X}^T , kann man die Kapazität umformen zu

$$C = \frac{1}{N} \max_{\mathbb{E}(\underline{X}^D \star \underline{X}^D) \leq N_D P} \mathbb{I}(\underline{Y}^D, \underline{Y}^T; \mathcal{X}^D | \mathcal{X}^T).$$

Durch die Schätzung der \mathcal{H}_i^D ist der Kanal nur partiell bekannt. Es handelt sich daher um einen partiell bekannten Kanal (P-CSI, “partially known channel state information”).

5.2. Abhängigkeitsstrukturen der Symentropie

Die Kanalmatrix wird in diesem Modell durch MMSE (minimum mean square error) geschätzt. Bei der Berechnung der Abhängigkeitsstrukturen wird auf einige wichtige Eigenschaften von MMSE-Schätzern zurückgegriffen, die in Abschnitt 2.4 zusammengefasst sind.

Satz 5.4. *Die Symentropie des oben definierten Kanals beträgt*

$$I(\underline{Y}^D, \underline{Y}^T; \underline{X}^D | \underline{X}^T) = I(\underline{Y}^D; \mathcal{X}^D | \hat{\underline{H}}^D) = \mathbb{E} \left(J(\underline{Y}^D, \mathcal{X}^D, \hat{\underline{H}}^D) \right),$$

wobei

$$J(\underline{y}^D, X^D, \hat{\underline{h}}^D) = \log \frac{f(\underline{y}^D | X^D, \hat{\underline{h}}^D)}{\int f(\underline{y}^D | X^{D'}, \hat{\underline{h}}^D) f(X^{D'}) dX^{D'}}$$

und $f(\underline{y}^D | X^D, \hat{\underline{h}}^D)$ die Dichte der Zufallsvariablen $\underline{Y}^D | \mathcal{X}^D = X^D, \hat{\underline{H}}^D = \hat{\underline{h}}^D$ und $f(X^D)$ die Dichte von \mathcal{X}^D seien.

Beweis. Für die Symentropie gilt

$$\begin{aligned} I(\underline{Y}^D, \underline{Y}^T; \mathcal{X}^D | \mathcal{X}^T) &= \underbrace{I(\underline{Y}^T; \mathcal{X}^D | \mathcal{X}^T)}_{=0, \text{ da } \underline{Y}^T \text{ u. } \mathcal{X}^D \text{ stoch. unabh.}} + I(\underline{Y}^D; \mathcal{X}^D | \mathcal{X}^T, \underline{Y}^T) \\ &= \int_{\mathbb{C}^{rN_D}} \int_{\mathbb{C}^{tN_D}} \int_{\mathbb{C}^{rN_T}} f(\underline{y}^D, X^D, \underline{y}^T | \mathcal{X}^T) \log \frac{f(\underline{y}^D | X^D, \underline{y}^T, \mathcal{X}^T)}{f(\underline{y}^D | \underline{y}^T, \mathcal{X}^T)} d\underline{y}^D dX^D d\underline{y}^T, \end{aligned}$$

wobei f jeweils die zu den Argumenten zugehörige Dichte bezeichne. Sei

$$J'(\underline{y}^D, X^D, \underline{y}^T) := \log \frac{f(\underline{y}^D | X^D, \underline{y}^T, \mathcal{X}^T)}{f(\underline{y}^D | \underline{y}^T, \mathcal{X}^T)}.$$

Dann ist

$$I(\underline{Y}^D, \underline{Y}^T; \mathcal{X}^D | \mathcal{X}^T) = \mathbb{E} (J'(\underline{Y}^D, \mathcal{X}^D, \underline{Y}^T)).$$

Es gilt

$$\log \frac{f(\underline{y}^D | X^D, \underline{y}^T, \mathcal{X}^T)}{f(\underline{y}^D | \underline{y}^T, \mathcal{X}^T)} = \log \frac{f(\underline{y}^D | X^D, \underline{y}^T, \mathcal{X}^T)}{\int f(\underline{y}^D | X^{D'}, \underline{y}^T, \mathcal{X}^T) f(X^{D'}) dX^{D'}}.$$

Damit erhält man die Darstellung

$$J'(\underline{y}^D, X^D, \underline{y}^T) = \log \frac{f(\underline{y}^D | X^D, \underline{y}^T, \mathcal{X}^T)}{\int f(\underline{y}^D | X^{D'}, \underline{y}^T, \mathcal{X}^T) f(X^{D'}) dX^{D'}}. \quad (5.7)$$

Da der Datenkanal gedächtnislos ist, sind die einzelnen Komponenten des Vektors \underline{Y}^D stochastisch unabhängig. Für die Randdichten gilt daher

$$f(\underline{y}^D | X^D, \underline{y}^T, \mathcal{X}^T) = \prod_{i=1}^{N_D} f(y_i^D | X_i^D, \underline{y}^T, \mathcal{X}^T). \quad (5.8)$$

Jede einzelne Randdichte $f(\underline{y}_i^D | \mathcal{X}_i^D, \underline{y}^T, \mathcal{X}^T)$ ist symmetrisch komplex normalverteilt als Linearkombination von symmetrisch komplex verteilten Zufallsvariablen. Die Verteilung einer einzelnen Randdichte ist also eindeutig bestimmt durch den Erwartungswertvektor und die Kovarianzmatrix.

Man definiere einen MMSE-Schätzer

$$\hat{\underline{Y}}_i := \mathbb{E}(\underline{Y}_i^D | \mathcal{X}_i^D, \underline{Y}^T, \mathcal{X}^T)$$

für den Erwartungswertvektor. Der Schätzer errechnet sich als

$$\begin{aligned} \hat{\underline{Y}}_i &= \mathbb{E}(\underline{Y}_i^D | \mathcal{X}_i^D, \underline{Y}^T, \mathcal{X}^T) = \mathbb{E}(\mathcal{X}_i^D \underline{H}_i^D | \mathcal{X}_i^D, \underline{Y}^T, \mathcal{X}^T) \\ &= \mathcal{X}_i^D \mathbb{E}(\underline{H}_i^D | \mathcal{X}_i^D, \underline{Y}^T, \mathcal{X}^T) = \mathcal{X}_i^D \mathbb{E}(\underline{H}_i^D | \mathcal{X}^T, \underline{Y}^T) = \mathcal{X}_i^D \hat{\underline{H}}_i^D. \end{aligned}$$

In dieser Rechnung bleibt fraglich, ob \mathcal{X}_i^D vor den Erwartungswert gezogen werden kann. Dies ist möglich, falls es $\mathfrak{A}(\mathcal{X}_i^D, \underline{Y}^T, \mathcal{X}^T)$ -messbar ist. Dies trifft in der Tat zu, denn es ist nun \mathcal{X}_i^D nach Definition $\mathfrak{A}(\mathcal{X}_i^D)$ -messbar, und, weil $\mathfrak{A}(\mathcal{X}_i^D, \underline{Y}^T, \mathcal{X}^T) \supseteq \mathfrak{A}(\mathcal{X}_i^D)$ ist, somit auch $\mathfrak{A}(\mathcal{X}_i^D, \underline{Y}^T, \mathcal{X}^T)$ -messbar.

Der entsprechende Schätzer für die Kovarianz-Matrix lautet

$$\begin{aligned} \hat{Q}_i &= \mathbb{E}\left(\left(\underline{Y}_i^D - \hat{\underline{Y}}_i\right)\left(\underline{Y}_i^D - \hat{\underline{Y}}_i\right)^* \mid \mathcal{X}_i^D, \mathcal{X}^T, \underline{Y}^T\right) \\ &= \mathbb{E}\left(\left(\mathcal{X}_i^D \underline{H}_i^D + \underline{Z}_i^D - \mathcal{X}_i^D \hat{\underline{H}}_i^D\right)\left(\mathcal{X}_i^D \underline{H}_i^D + \underline{Z}_i^D - \mathcal{X}_i^D \hat{\underline{H}}_i^D\right)^* \mid \mathcal{X}_i^D, \mathcal{X}^T, \underline{Y}^T\right) \\ &= \mathbb{E}\left(\left(\mathcal{X}_i^D \left(\underline{H}_i^D - \hat{\underline{H}}_i^D\right) + \underline{Z}_i^D\right)\left(\mathcal{X}_i^D \left(\underline{H}_i^D - \hat{\underline{H}}_i^D\right) + \underline{Z}_i^D\right)^* \mid \mathcal{X}_i^D, \mathcal{X}^T, \underline{Y}^T\right) \\ &= \mathcal{X}_i^D \mathbb{E}\left(\left(\underline{H}_i^D - \hat{\underline{H}}_i^D\right)\left(\underline{H}_i^D - \hat{\underline{H}}_i^D\right)^* \mid \mathcal{X}^T, \underline{Y}^T\right) \mathcal{X}_i^{D*} + \mathbb{E}\left(\underline{Z}_i^D \underline{Z}_i^{D*}\right) \\ &\quad + \underbrace{\mathcal{X}_i^D \mathbb{E}\left(\left(\underline{H}_i^D - \hat{\underline{H}}_i^D\right) \underline{Z}_i^{D*} \mid \mathcal{X}^T, \underline{Y}^T\right)}_{=0, \text{ denn } \underline{Z}_i^D \text{ und } \underline{H}_i^D \text{ sind unabhängig und } \mathbb{E} \underline{Z}_i^D = 0} \\ &\quad + \underbrace{\mathbb{E}\left(\underline{Z}_i^D \left(\underline{H}_i^D - \hat{\underline{H}}_i^D\right)^* \mid \mathcal{X}^T, \underline{Y}^T\right) \mathcal{X}_i^{D*}}_{=0, \text{ denn } \underline{Z}_i^D \text{ und } \hat{\underline{H}}_i^D \text{ sind unabhängig und } \mathbb{E} \underline{Z}_i^D = 0} \\ &= \mathcal{X}_i^D \text{MSE}_{\underline{H}_i^D} \left(\hat{\underline{H}}_i^D\right) \mathcal{X}_i^{D*} + N_0 I_r. \end{aligned}$$

Mit der Bezeichnung

$$C_i^\varepsilon := \text{MSE}_{\underline{H}_i^D} \left(\hat{\underline{H}}_i^D\right)$$

ergibt sich die Darstellung

$$\hat{Q}_i = \mathcal{X}_i^D C_i^\varepsilon \mathcal{X}_i^{D*} + N_0 I_r.$$

Man sieht, dass die Verteilungen nur noch von den Schätzern des Kanalvektors $\hat{\underline{H}}_i^D$ und den Datensymbolen \mathcal{X}^D abhängen, nicht aber von den Outputsymbolen \underline{Y}^T . Wenn man die MMSE-Schätzer insbesondere als bedingte Erwartungen betrachtet, folgt, dass die Erwartung von \underline{Y}_i^D bzw. von $\left(\underline{Y}_i^D - \hat{\underline{Y}}_i\right)\left(\underline{Y}_i^D - \hat{\underline{Y}}_i\right)^*$ mit gegebener a-priori Information $\mathcal{X}_i^D, \mathcal{X}^T, \underline{Y}^T$ ebenfalls nur von $\hat{\underline{H}}_i^D$ und

\mathcal{X}^D abhängen [7]. Die gemeinsame bedingte Dichte des Zufallsvektors $\underline{Y}^D \mid \mathcal{X}^D, \underline{Y}^T, \mathcal{X}^T$ ist nach (5.8) daher nur abhängig von $\hat{\underline{H}}^D$ und \mathcal{X}^D . Damit erhält man die Darstellung

$$\begin{aligned}
J'(\underline{y}^D, X^D, \underline{y}^T) &= \log \frac{f(\underline{y}^D \mid X^D, \underline{y}^T, \mathcal{X}^T)}{\int f(\underline{y}^D \mid X^{D'}, \underline{y}^T, \mathcal{X}^T) f(X^{D'}) dX^{D'}} \\
&= \log \frac{f(\underline{y}^D \mid X^D, \hat{\underline{h}}^D)}{\int f(\underline{y}^D \mid X^{D'}, \hat{\underline{h}}^D) f(X^{D'}) dX^{D'}} \\
&= \log \frac{f(\underline{y}^D \mid X^D, \hat{\underline{h}}^D)}{f(\underline{y}^D \mid \hat{\underline{h}}^D)} \\
&= J(\underline{y}^D, X^D, \hat{\underline{h}}^D \mid \mathcal{X}^T, \underline{y}^T) = J(\underline{y}^D, X^D, \hat{\underline{h}}^D). \quad (5.9)
\end{aligned}$$

Daraus ergibt sich

$$I(\underline{Y}^D, \underline{Y}^T; \mathcal{X}^D \mid \mathcal{X}^T) = I(\underline{Y}^D; \mathcal{X}^D \mid \mathcal{X}^T, \underline{Y}^T) = I(\underline{Y}^D; \mathcal{X}^D \mid \hat{\underline{H}}^D)$$

bzw. auf komponentenweiser Ebene

$$I(\underline{Y}_i^D, \underline{Y}_i^T; \mathcal{X}_i^D \mid \mathcal{X}^T) = I(\underline{Y}_i^D; \mathcal{X}_i^D \mid \hat{\underline{H}}_i^D).$$

□

Korollar 5.5. *Die Symentropie kann man als Summe der komponentenweisen Symentropien ausdrücken,*

$$I(\underline{Y}^D, \underline{Y}^T; \mathcal{X}^D \mid \mathcal{X}^T) = \sum_{i=1}^{N_D} I(\underline{Y}_i^D, \underline{Y}_i^T; \mathcal{X}_i^D \mid \mathcal{X}^T).$$

Beweis. Das Ergebnis (5.8) lässt sich äquivalent als Gleichung zwischen den Synentropien ausdrücken,

$$\begin{aligned}
I(\underline{Y}^D, \underline{Y}^T; \mathcal{X}^D | \mathcal{X}^T) &= I(\underline{Y}^D; \mathcal{X}^D | \mathcal{X}^T, \underline{Y}^T) = I(\underline{Y}_1^D, \dots, \underline{Y}_{N_D}^D; \mathcal{X}^D | \mathcal{X}^T, \underline{Y}^T) \\
&= H(\underline{Y}_1^D, \dots, \underline{Y}_{N_D}^D | \mathcal{X}^T, \underline{Y}^T) - H(\underline{Y}_1^D, \dots, \underline{Y}_{N_D}^D | \mathcal{X}^D, \mathcal{X}^T, \underline{Y}^T) \\
&= \sum_{i=1}^{N_D} H(\underline{Y}_i^D | \mathcal{X}^T, \underline{Y}^T) - \sum_{i=1}^{N_D} H(\underline{Y}_i^D | \mathcal{X}^D, \mathcal{X}^T, \underline{Y}^T) \\
&= \sum_{i=1}^{N_D} I(\underline{Y}_i^D; \mathcal{X}^D | \mathcal{X}^T, \underline{Y}^T) = \sum_{i=1}^{N_D} I(\underline{Y}_i^D; \mathcal{X}_1^D, \dots, \mathcal{X}_{N_D}^D | \mathcal{X}^T, \underline{Y}^T) \\
&= \sum_{i=1}^{N_D} H(\mathcal{X}_1^D, \dots, \mathcal{X}_{N_D}^D | \mathcal{X}^T, \underline{Y}^T) - \sum_{i=1}^{N_D} H(\mathcal{X}_1^D, \dots, \mathcal{X}_{N_D}^D | \mathcal{X}^T, \underline{Y}^T, \underline{Y}_i^D) \\
&= \sum_{i=1}^{N_D} \sum_{i'=1}^{N_D} H(\mathcal{X}_{i'}^D | \mathcal{X}^T, \underline{Y}^T) - \sum_{i=1}^{N_D} \sum_{i'=1}^{N_D} H(\mathcal{X}_{i'}^D | \mathcal{X}^T, \underline{Y}^T, \underline{Y}_i^D) \\
&= \sum_{i=1}^{N_D} \sum_{i'=1}^{N_D} I(\underline{Y}_i^D; \mathcal{X}_{i'}^D | \mathcal{X}^T, \underline{Y}^T) = \sum_{i=1}^{N_D} I(\underline{Y}_i^D; \mathcal{X}_i^D | \mathcal{X}^T, \underline{Y}^T) \\
&= \sum_{i=1}^{N_D} I(\underline{Y}_i^D, \underline{Y}^T; \mathcal{X}_i^D | \mathcal{X}_T).
\end{aligned}$$

Die vorletzte Gleichung gilt, da \underline{Y}_i^D und $\mathcal{X}_{i'}^D$ nur abhängig sind, wenn $i = i'$. \square

Satz 5.6. Für die Schätzer $\hat{\underline{H}}_i^D$ gilt

$$\hat{\underline{H}}_i^D \sim \text{SCN}(\underline{0}, C_i),$$

wobei

$$C_i = C_i^\varepsilon + \sigma_h^2 I_{rt} = \text{MSE}_{\underline{H}_i^D}(\hat{\underline{H}}_i^D) + \sigma_h^2 I_{rt}.$$

Dieses Ergebnis weicht leicht von dem entsprechenden Ergebnis in [3] ab, wo die Kovarianzmatrix mit

$$C_i = \sigma_h^2 I_{rt} = \text{MSE}_{\underline{H}_i^D}(\hat{\underline{H}}_i^D) + \sigma_h^2 I_{rt} - C_i^\varepsilon$$

angegeben wird.

Beweis. Der Schätzer $\hat{\underline{H}}_i^D$ ist eine Linearkombination symmetrisch komplex normalverteilter Zufallsvariablen, und damit selbst auch SCN-verteilt und durch Erwartungswertvektor und Kovarianzmatrix eindeutig bestimmt. Aus der Erwartungstreue des MMSE-Schätzers ergibt sich sofort

$$\text{E}(\hat{\underline{H}}_i^D) = \text{E}(\underline{H}_i^D) - \text{E}(\underline{H}_i^D - \hat{\underline{H}}_i^D) = \underline{0} - \underline{0} = \underline{0}.$$

Nach dem Orthogonalitätsprinzip für MMSE-Schätzer gilt

$$\text{E}((\underline{H}_i^D - \hat{\underline{H}}_i^D) \underline{H}_i^{D*}) = 0.$$

Für die Kovarianzmatrix folgt damit

$$\begin{aligned}
C_i &:= \mathbb{E} \left(\hat{\underline{H}}_i^D \hat{\underline{H}}_i^{D*} \right) = \mathbb{E} \left(\hat{\underline{H}}_i^D \hat{\underline{H}}_i^{D*} \right) + \mathbb{E} \left(\left(\underline{H}_i^D - \hat{\underline{H}}_i^D \right) \underline{H}_i^{D*} \right) \\
&= \mathbb{E} \left(\hat{\underline{H}}_i^D \hat{\underline{H}}_i^{D*} \right) + \mathbb{E} \left(\underline{H}_i^D \underline{H}_i^{D*} \right) - \mathbb{E} \left(\hat{\underline{H}}_i^D \underline{H}_i^{D*} \right) \\
&= \mathbb{E} \left(\hat{\underline{H}}_i^D \hat{\underline{H}}_i^{D*} \right) + \mathbb{E} \left(\underline{H}_i^D \underline{H}_i^{D*} \right) - \mathbb{E} \left(\hat{\underline{H}}_i^D \underline{H}_i^{D*} \right) \\
&\quad + \mathbb{E} \left(\underline{H}_i^D \hat{\underline{H}}_i^{D*} \right) - \mathbb{E} \left(\underline{H}_i^D \hat{\underline{H}}_i^{D*} \right) \\
&= \mathbb{E} \left(\left(\underline{H}_i^D - \hat{\underline{H}}_i^D \right) \left(\underline{H}_i^D - \hat{\underline{H}}_i^D \right)^* \right) + \mathbb{E} \left(\underline{H}_i^D \hat{\underline{H}}_i^{D*} \right) \\
&= \mathbb{E} \left(\left(\underline{H}_i^D - \hat{\underline{H}}_i^D \right) \left(\underline{H}_i^D - \hat{\underline{H}}_i^D \right)^* \right) + \mathbb{E} \left(\underline{H}_i^D \underline{H}_i^{D*} \right) = \text{MSE}_{\underline{H}_i^D} \left(\hat{\underline{H}}_i^D \right) + \sigma_h^2 I_{rt}
\end{aligned}$$

□

Da alle Verteilungen der Gleichung (5.9) jetzt bekannt sind, lässt sich die Syn-entropie $I(\underline{Y}^D, \underline{Y}^T; \underline{X}^D | \underline{X}^T)$ numerisch berechnen.

5.3. Berechnung einiger Kapazitätsschranken

Die Maximierung der Synentropie ist ein noch offenes Problem. Die Kanalkapazität kann man daher nicht als geschlossenen Ausdruck angeben. Es lassen sich jedoch obere wie untere Schranken angeben. Die untere Schranke stammt ebenfalls aus [3], die obere Schranke ist eine Folgerung der Ergebnisse aus [3]. Für die Berechnung der Schranken benötigt man zunächst das folgende

Lemma 5.7 (Matrix Inversion Lemma). *Gegeben sei eine positiv definite Matrix $A \in \mathbb{C}^{m \times m}$ der Form*

$$A = B^{-1} + CD^{-1}C^*$$

wobei $B \in \mathbb{C}^{m \times m}$ positiv definit, $C \in \mathbb{C}^{m \times n}$ und $D \in \mathbb{C}^{n \times n}$ positiv definit. Die Inverse der Matrix A ist dann gegeben durch

$$A^{-1} = B - BC(D + C^*BC)^{-1}C^*B.$$

Beweis. Dieser Beweis stammt aus [19]. Man beachte zunächst, dass sich die Einheitsmatrix schreiben lässt als

$$I_m = A^{-1}A = A^{-1}B^{-1} + A^{-1}CD^{-1}C^*.$$

Man multipliziere diesen Ausdruck von links mit B und erhalte

$$B = A^{-1} + A^{-1}CD^{-1}C^*B. \quad (5.10)$$

Daraus folgt

$$B - A^{-1} = A^{-1}CD^{-1}C^*B. \quad (5.11)$$

Multiplizieren von (5.10) von rechts mit CD^{-1} gibt

$$BCD^{-1} = A^{-1}CD^{-1} + A^{-1}CD^{-1}C^*BCD^{-1} = A^{-1}CD^{-1} (I_n + C^*BCD^{-1}).$$

Daraus ergibt sich

$$A^{-1}CD^{-1} = BCD^{-1} (I_n + C^*BCD^{-1})^{-1}.$$

Weiteres Multiplizieren mit C^*B von rechts ergibt

$$A^{-1}CD^{-1}C^*B = BCD^{-1}(I_n + C^*BCD^{-1})^{-1}C^*B.$$

Mit (5.11) folgt

$$B - A^{-1} = BCD^{-1}(I_n + C^*BCD^{-1})^{-1}C^*B,$$

und somit

$$A^{-1} = B - BCD^{-1}(I_n + C^*BCD^{-1})^{-1}C^*B = B - BC(D + C^*BC)^{-1}C^*B.$$

□

Satz 5.8. *Für den obigen P-CSI Kanal gilt*

$$C \geq \frac{1}{N} \sum_{i=1}^{N_D} \mathbb{E} \left(\log \det \left(I_r + \frac{P}{t(P\sigma_h^2 + N_0)} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} \right) \right).$$

Beweis. Man lege die Inputverteilung fest als

$$\underline{X}_i^D \sim \text{SCN} \left(\underline{0}, \frac{P}{t} I_t \right).$$

Man schreibe jetzt die Synentropie um zu

$$I(\underline{Y}_i^D; \underline{X}_i^D | \hat{\mathcal{H}}_i^D) = H(\underline{X}_i^D | \hat{\mathcal{H}}_i^D) - H(\underline{X}_i^D | \underline{Y}_i^D, \hat{\mathcal{H}}_i^D).$$

Für den ersten Summanden gilt nach der festgelegten Inputverteilung

$$H(\underline{X}_i^D | \hat{\mathcal{H}}_i^D) = H(\underline{X}_i^D) = \log \det (\pi e \text{Cov}(\underline{X}_i^D)) = \log \det \left(\pi e \frac{P}{t} I_t \right).$$

Für den zweiten Summanden benutze man die Methoden aus [13] wie folgt. Es gilt zunächst

$$\begin{aligned} H(\underline{X}_i^D | \underline{Y}_i^D, \hat{\mathcal{H}}_i^D) &= H(\underline{X}_i^D - A\underline{Y}_i^D | \underline{Y}_i^D, \hat{\mathcal{H}}_i^D) \\ &\leq H(\underline{X}_i^D - A\underline{Y}_i^D) \leq \log \det (\pi e \text{Cov}(\underline{X}_i^D - A\underline{Y}_i^D)) \end{aligned}$$

für beliebige Matrix $A \in \mathbb{C}^{t \times r}$. Die Kovarianzmatrix in dieser Formel errechnet sich als

$$\text{Cov}(\underline{X}_i^D - A\underline{Y}_i^D) = \text{Cov}(\underline{X}_i^D) - A \text{Cov}(\underline{Y}_i^D) A^* = \frac{P}{t} I_t - A \text{Cov}(\underline{Y}_i^D) A^*.$$

Die Kovarianzmatrix von \underline{Y}_i^D wiederum errechnet sich als

$$\text{Cov}(\underline{Y}_i^D) = \text{Cov}(\mathcal{H}_i^D \underline{X}_i^D + \underline{Z}_i^D) = \text{Cov}(\mathcal{H}_i^D \underline{X}_i^D) + N_0 I_r.$$

Die Kovarianzmatrix aus dieser Formel errechnet sich koordinatenweise

$$\text{Cov}(\mathcal{H}_i^D \underline{X}_i^D)_{jk} = \text{Cov} \left(\sum_{l=1}^t H_{ijl}^D X_{il}^D, \sum_{m=1}^t H_{ikm}^D X_{im}^D \right) = \delta_{jk} \sum_{l=1}^t \text{Var}(H_{ijl}^D X_{il}^D),$$

wobei gilt

$$\begin{aligned} \text{Var} (H_{ijl}^D X_{il}^D) &= \text{E} \left(H_{ijl}^D X_{il}^D \overline{H_{ijl}^D X_{il}^D} \right) - \underbrace{\text{E} (H_{ijl}^D X_{il}^D) \text{E} \left(\overline{H_{ijl}^D X_{il}^D} \right)}_{=0} \\ &= \text{E} \left(|H_{ijl}^D X_{il}^D|^2 \right) = \text{E} \left(|H_{ijl}^D|^2 \right) \text{E} \left(|X_{il}^D|^2 \right) = \frac{P\sigma_h^2}{t}. \end{aligned}$$

Jeweiliges Einsetzen in die Gleichungen liefert dann insgesamt

$$\begin{aligned} \text{H} \left(\underline{X}_i^D \mid \underline{Y}_i^D, \hat{\mathcal{H}}_i^D \right) &\leq \log \det \left(\pi e \text{Cov} \left(\underline{X}_i^D - A \underline{Y}_i^D \right) \right) \\ &= \log \det \left(\pi e \left(\frac{P}{t} I_t - A \text{Cov} \left(\underline{Y}_i^D \right) A^* \right) \right) \\ &= \log \det \left(\pi e \left(\frac{P}{t} I_t - A \left(\text{Cov} \left(\mathcal{H}_i \underline{X}_i^D \right) + N_0 I_r \right) A^* \right) \right) \\ &= \log \det \left(\pi e \left(\frac{P}{t} I_t - A \left(P\sigma_h^2 I_r + N_0 I_r \right) A^* \right) \right) \\ &= \log \det \left(\pi e \left(\frac{P}{t} I_t - \left(P\sigma_h^2 + N_0 \right) A A^* \right) \right). \end{aligned}$$

Man wähle die Matrix A jetzt speziell

$$A := \sqrt{\frac{P^2/t^2}{P\sigma_h^2 + N_0}} \hat{\mathcal{H}}_i^{D*} \text{Cov} \left(\hat{\mathcal{H}}_i^D \underline{X}_i^D + \underline{Y}_i^D \right)^{-\frac{1}{2}}.$$

Damit folgt

$$\begin{aligned} \text{H} \left(\underline{X}_i^D \mid \underline{Y}_i^D, \hat{\mathcal{H}}_i^D \right) &\leq \log \det \left(\pi e \left(\frac{P}{t} I_t - \frac{P^2}{t^2} \hat{\mathcal{H}}_i^{D*} \text{Cov} \left(\hat{\mathcal{H}}_i^D \underline{X}_i^D + \underline{Y}_i^D \right)^{-1} \hat{\mathcal{H}}_i^D \right) \right), \quad (5.12) \end{aligned}$$

und somit auch

$$\begin{aligned} \text{H} \left(\underline{X}_i^D \mid \underline{Y}_i^D, \hat{\mathcal{H}}_i^D \right) &\leq \text{E} \left(\log \det \left(\pi e \left(\frac{P}{t} I_t - \frac{P^2}{t^2} \hat{\mathcal{H}}_i^{D*} \text{Cov} \left(\hat{\mathcal{H}}_i^D \underline{X}_i^D + \underline{Y}_i^D \right)^{-1} \hat{\mathcal{H}}_i^D \right) \mid \hat{\mathcal{H}}_i^D \right) \right) \\ &= \text{E} \left(\log \det \left(\pi e Q \right) \mid \hat{\mathcal{H}}_i^D \right) \end{aligned}$$

mit

$$Q := \frac{P}{t} I_t - \frac{P^2}{t^2} \hat{\mathcal{H}}_i^{D*} \left(\frac{P}{t} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} + P\sigma_h^2 I_r + N_0 I_r \right)^{-1} \hat{\mathcal{H}}_i^D.$$

Nochmaliges Anwenden des Erwartungswertes ergibt

$$\text{H} \left(\underline{X}_i^D \mid \underline{Y}_i^D, \hat{\mathcal{H}}_i^D \right) \leq \text{E} \left(\log \det \left(\pi e Q \right) \right).$$

Nach dem ‘‘Matrix Inversion Lemma’’ 5.7 ist

$$\begin{aligned} Q^{-1} &= \frac{t}{P}I_t - \hat{\mathcal{H}}_i^{D*} \left(-\frac{P}{t}\hat{\mathcal{H}}_i^D\hat{\mathcal{H}}_i^{D*} - P\sigma_h^2I_r - N_0I_r + \frac{P}{t}\hat{\mathcal{H}}_i^D\hat{\mathcal{H}}_i^{D*} \right)^{-1} \hat{\mathcal{H}}_i^D \\ &= \frac{t}{P}I_t + \frac{1}{P\sigma_h^2 + N_0}\hat{\mathcal{H}}_i^{D*}\hat{\mathcal{H}}_i^D. \end{aligned}$$

Somit ergibt sich

$$\begin{aligned} \mathbb{I}(\underline{Y}_i^D; \underline{X}_i^D \mid \hat{\mathcal{H}}_i^D) &\leq \mathbb{E} \left(\log \det \left(\pi e \frac{P}{t} I_t \right) - \log \det (\pi e Q) \right) \\ &= \mathbb{E} \left(\log \det \left(\pi e \frac{P}{t} I_t \right) + \log \det \left(\frac{1}{\pi e} Q^{-1} \right) \right) \\ &= \mathbb{E} \left(\log \det \left(\frac{P}{t} Q^{-1} \right) \right) \\ &= \mathbb{E} \left(\log \det \left(I_t + \frac{P}{t(P\sigma_h^2 + N_0)} \hat{\mathcal{H}}_i^{D*} \hat{\mathcal{H}}_i^D \right) \right) \\ &= \mathbb{E} \left(\log \det \left(I_r + \frac{P}{t(P\sigma_h^2 + N_0)} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} \right) \right), \end{aligned}$$

und für die Kapazität damit nach Korollar 5.5

$$\begin{aligned} C &\geq \frac{1}{N} \sum_{i=1}^{N_D} \mathbb{I}(\underline{Y}_i^D; \underline{X}_i^D \mid \hat{\mathcal{H}}_i^D) \\ &\geq \frac{1}{N} \sum_{i=1}^{N_D} \mathbb{E} \left(\log \det \left(I_r + \frac{P}{t(P\sigma_h^2 + N_0)} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} \right) \right). \end{aligned}$$

□

Auch hier weicht das Ergebnis leicht vom Ergebnis in [3] ab, wo die Summe zusammengefasst wird. Diese Zusammenfassung der Summe ist jedoch nur möglich, falls $\hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*}$ unabhängig von i ist, was im Allgemeinen jedoch nicht der Fall ist.

Bemerkung 5.9. Zur Wahl der Matrix A ist zu bemerken, dass nach Bemerkung 6.3

$$A = \text{Cov}(\underline{X}_i^D, \underline{Y}_i^D) \text{Cov}(\underline{Y}_i^D)^{-1}$$

möglicherweise eine gute Wahl wäre. Dass dies hier nicht der Fall ist, wird ersichtlich, wenn man beachtet, dass (vergleiche Beweis zu Satz 6.2, insbesondere (6.2)) $\text{Cov}(\underline{X}_i^D, \underline{Y}_i^D) = 0$ gilt, woraus lediglich die triviale Schranke $C \geq 0$ folgt. Die Wahl von A erfolgt lediglich aus rechentechnischen Gründen. Es ist nicht auszuschließen, dass es eine bessere Wahl für A gibt, die ein besseres Ergebnis liefert.

Satz 5.10. Für den obigen P-CSI Kanal gilt

$$\begin{aligned} & \mathbb{I}(\underline{Y}_i^D; \underline{X}_i^D \mid \hat{\mathcal{H}}_i^D) \\ & \leq r \mathbb{E} \left(\log \left(\frac{P\sigma_h^2}{\sigma_h^2 \underline{X}_i^{D*} \underline{X}_i^D + tN_0} \right) \right) + \mathbb{E} \left(\log \det \left(I_r + \frac{1}{\sigma_h^2} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} \right) \right). \end{aligned}$$

Beweis. Die Syntropie lässt sich umschreiben zu

$$\mathbb{I}(\underline{Y}_i^D; \underline{X}_i^D \mid \hat{\mathcal{H}}_i^D) = \mathbb{H}(\underline{Y}_i^D \mid \hat{\mathcal{H}}_i^D) - \mathbb{H}(\underline{Y}_i^D \mid \underline{X}_i^D, \hat{\mathcal{H}}_i^D).$$

Für den ersten Summanden folgt

$$\mathbb{H}(\underline{Y}_i^D \mid \hat{\mathcal{H}}_i^D) \leq \log \det \left(\pi e \text{Cov}(\underline{Y}_i^D \mid \hat{\mathcal{H}}_i^D) \right),$$

wobei

$$\text{Cov}(\underline{Y}_i^D \mid \hat{\mathcal{H}}_i^D) = \frac{P}{t} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} + \frac{P\sigma_h^2}{t} I_r + N_0 I_r.$$

Für den zweiten Summanden ist zunächst zu beachten, dass die zugehörige bedingte Verteilung nach dem Beweis von Satz 5.4 SCN ist mit

$$\underline{Y}_i^D \mid \underline{X}_i^D, \hat{\mathcal{H}}_i^D \sim \text{SCN} \left(\mathcal{A}_i^D \hat{\underline{H}}_i^D, \mathcal{A}_i^D C_i^\varepsilon \mathcal{A}_i^{D*} + N_0 I_r \right).$$

Damit ist

$$\begin{aligned} & \text{Cov}(\underline{Y}_i^D \mid \underline{X}_i^D, \hat{\mathcal{H}}_i^D) \\ & = \mathcal{A}_i^D \mathbb{E} \left(\left(\underline{H}_i^D - \hat{\underline{H}}_i^D \right) \left(\underline{H}_i^D - \hat{\underline{H}}_i^D \right)^* \mid \hat{\mathcal{H}}_i^D, \underline{X}_i^D \right) \mathcal{A}_i^{D*} + N_0 I_r \\ & = \mathbb{E} \left(\left(\left(\mathcal{H}_i^D - \hat{\mathcal{H}}_i^D \right) \underline{X}_i^D \right) \left(\left(\mathcal{H}_i^D - \hat{\mathcal{H}}_i^D \right) \underline{X}_i^D \right)^* + \underline{Z}_i^D \underline{Z}_i^{D*} \mid \hat{\mathcal{H}}_i^D, \underline{X}_i^D \right) \\ & = \mathbb{E} \left(\left(\underline{Y}_i^D - \hat{\mathcal{H}}_i^D \underline{X}_i^D \right) \left(\underline{Y}_i^D - \hat{\mathcal{H}}_i^D \underline{X}_i^D \right)^* \mid \hat{\mathcal{H}}_i^D, \underline{X}_i^D \right) \\ & = \frac{\sigma_h^2}{t} \underline{X}_i^{D*} \underline{X}_i^D I_r + N_0 I_r. \end{aligned}$$

Damit folgt für die Syntropie

$$\begin{aligned} & \mathbb{I}(\underline{Y}_i^D; \underline{X}_i^D \mid \hat{\mathcal{H}}_i^D) \leq \mathbb{E} \left(\log \frac{\det \left(\pi e \left(\frac{P}{t} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} + \frac{P\sigma_h^2}{t} I_r + N_0 I_r \right) \right)}{\det \left(\pi e \left(\frac{\sigma_h^2}{t} \underline{X}_i^{D*} \underline{X}_i^D I_r + N_0 I_r \right) \right)} \right) \\ & = \mathbb{E} \left(\log \det \left(\frac{P}{t} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} + \left(\frac{P\sigma_h^2}{t} + N_0 \right) I_r \right) \left(\frac{t}{\sigma_h^2 \underline{X}_i^{D*} \underline{X}_i^D + tN_0} I_r \right) \right) \\ & = \mathbb{E} \left(\log \det \left(\frac{P}{\sigma_h^2 \underline{X}_i^{D*} \underline{X}_i^D + tN_0} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} + \frac{P\sigma_h^2}{\sigma_h^2 \underline{X}_i^{D*} \underline{X}_i^D + tN_0} I_r \right) \right) \\ & = \mathbb{E} \left(\log \left(\frac{P\sigma_h^2}{\sigma_h^2 \underline{X}_i^{D*} \underline{X}_i^D + tN_0} \right)^r \det \left(\frac{1}{\sigma_h^2} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} + I_r \right) \right) \\ & = r \mathbb{E} \left(\log \left(\frac{P\sigma_h^2}{\sigma_h^2 \underline{X}_i^{D*} \underline{X}_i^D + tN_0} \right) \right) + \mathbb{E} \left(\log \det \left(\frac{1}{\sigma_h^2} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} + I_r \right) \right). \end{aligned}$$

□

Das folgende Korollar ist nicht in [3] enthalten.

Korollar 5.11. *Für den obigen P-CSI Kanal gilt*

$$C \leq \frac{1}{N} \sum_{i=1}^{N_D} \left(r \log \left(\frac{P\sigma_h^2}{tN_0} \right) + \mathbb{E} \left(\log \det \left(I_r + \frac{1}{\sigma_h^2} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} \right) \right) \right).$$

Beweis. Es ist $\underline{X}_i^{D*} \underline{X}_i^D \geq 0$, damit folgt

$$\mathbb{E} \left(\log \left(\frac{P\sigma_h^2}{\sigma_h^2 \underline{X}_i^{D*} \underline{X}_i^D + tN_0} \right) \right) \leq \log \left(\frac{P\sigma_h^2}{tN_0} \right).$$

Mit Satz 5.10 folgt die Behauptung. □

Die Ergebnisse über die Kanalkapazität werden noch einmal zusammengefasst in dem folgenden

Korollar 5.12. *Für den obigen P-CSI Kanal gilt*

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^{N_D} \mathbb{E} \left(\log \det \left(I_r + \frac{P}{t(P\sigma_h^2 + N_0)} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} \right) \right) &\leq C \\ &\leq \frac{1}{N} \sum_{i=1}^{N_D} \left(r \log \left(\frac{P\sigma_h^2}{tN_0} \right) + \mathbb{E} \left(\log \det \left(I_r + \frac{1}{\sigma_h^2} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} \right) \right) \right). \end{aligned}$$

5.4. Weitere Optimierungsansätze

Um die Kanalkapazität zu verbessern, ist es möglich, das Verhältnis von Test- zu Datensymbolen zu verändern. Im bisherigen Modell aus [3] erfordert dies die äußerst aufwändige Berechnung der $\hat{\mathcal{H}}_i^D = \mathbb{E}(\mathcal{H}_i^D | \underline{X}_1^T, \dots, \underline{X}_{N_T}^T, \underline{Y}_1^T, \dots, \underline{Y}_{N_T}^T)$, $i = 1, \dots, N_D$, die durch die Anzahl der Testsymbole indirekt beeinflusst werden. In [8] wird ein MIMO-Kanal mit Synchronised Detection daher anders als in dieser Arbeit modelliert, um dieses Problem abzuschwächen. Es wird davon ausgegangen, dass die Kanalmatrix \mathcal{H} zufällig ausgewürfelt wird, und dann über N Zeiteinheiten konstant bleibt. Über diesen konstanten Kanal werden hintereinander die Test- und Datensymbole mit unterschiedlichen Leistungszuweisungen übertragen.

Dieses System wird ebenfalls durch einen Test- und einen Datenkanal modelliert, wobei der Testkanal durch

$$\mathcal{Y}^T := \sqrt{\frac{\rho_T}{t}} \mathcal{H} \mathcal{X}^T + \mathcal{Z}^T$$

mit $\mathcal{X}^T \in \mathbb{C}^{t \times N_T}$ den Testsymbolen nebeneinander, $\mathcal{Y}^T \in \mathbb{C}^{r \times N_T}$ den Outputsymbolen nebeneinander und $\mathcal{Z}^T \in \mathbb{C}^{r \times N_T}$ den additiven Störungen nebeneinander, jeweils in einer Matrix zusammengefasst, definiert ist. Der Wert ρ_T bezeichne die SNR (Signal to Noise Ratio), die hier angegeben wird, weil

während der Trainingsphase und der Datenübertragungsphase mit unterschiedlicher Leistung gesendet wird. Analog modelliert man den Datenkanal durch

$$\mathcal{Y}^D := \sqrt{\frac{\rho_D}{t}} \mathcal{H} \mathcal{X}^D + \mathcal{Z}^D$$

Im Paper wird für das Modell eine untere Schranke ähnlich wie in diesem Kapitel ermittelt. Diese wird in Abhängigkeit von den gewählten Testsymbolen, der Leistungszuweisung sowie dem Verhältnis von Test- zu Datensymbolen maximiert. Im Synchronised Detection Modell, wie es in den vorigen Abschnitten eingeführt wurde, sind die ersten beiden Optimierungstechniken nicht sinnvoll, da in der Realität die Test- und Datensymbole bei der Übertragung durch Interleaving derart durcheinander gemischt sind, dass sie weder mit unterschiedlichen Leistungszuweisungen übertragen werden können, noch die Struktur der Testsymbole bei der Übertragung erhalten bleibt und analysiert werden könnte.

Es bleibt die Optimierung des Verhältnisses von Test- zu Datensymbolen. Der Ansatz aus [8] ist problematisch. In dieser Arbeit wird eine untere Schranke der Kapazität der Form

$$C \geq \frac{N_D}{N} \mathbb{E} \log(1 + \rho \lambda)$$

gefolgert, wobei $\rho := \rho(r, t, N_D)$ die optimale Leistungszuweisung ist, und λ ein Eigenwert der Matrix $(1/\sigma_h^2 t) \hat{H}^D \hat{H}^{D*}$. Dieser Ausdruck wird nach N_D abgeleitet. Dabei hängt die optimale Leistungszuweisung ρ von N_D ab, aber insbesondere auch die geschätzte Kanalmatrix, und damit auch ihre Eigenwerte, die bei wachsendem N_D eine bessere Qualität haben. Die Autoren beachten dies in ihrem Paper nicht, wodurch ihnen die innere Ableitung des λ fehlt. Es ist unbekannt, ob sich \hat{H}^D oder λ überhaupt geschlossen in Abhängigkeit von N_D darstellen lassen.

Senderseitig partiell bekannte Kanäle

6.1. Modellierung von senderseitigem P-CSI

Die bisherigen Kapitel, ausgenommen Abschnitt 4.1, differenzierten nur zwischen der Kenntnis der Kanalinformation an der Empfängerseite. In diesem Kapitel wird der Fall betrachtet, dass schon senderseitig die Kanalinformation partiell bekannt ist, und empfängerseitig kein zusätzliches Wissen über den Kanal erlangt wird. Das lässt sich dadurch modellieren, dass die Kanalmatrix \mathcal{H} in einen konstanten Teil $\bar{\mathcal{H}}$, der z.B. durch Kanalschätzung bekannt ist, sowie einen zufälligen Fehler $\tilde{\mathcal{H}}$ zerfällt. Bei der Bestimmung von Kapazitätsschranken eines Kanals mit senderseitigem P-CSI tauchen viele Parallelen zur Synchronised Detection auf. Insbesondere der direkte Vergleich der Schranken wird sich als interessant erweisen. In diesem Kapitel werden diese beiden Kanalmodelle an zwei Beispielen konkret verglichen. Eine allgemeinere Gegenüberstellung findet sich im nachfolgenden Kapitel 7. Das Modell in diesem Kapitel basiert auf einem Kanalmodell aus [13]. Die Ergebnisse werden erweitert, Unklarheiten aus dem Paper beseitigt, Grenzwertbetrachtungen gemacht, sowie aus den Abschätzungen der Synentropie auch eine obere Schranke für die Kanalkapazität gefolgert. Anhand einiger Beispiele werden die obere und die untere Schranke in geschlossener Darstellung ausgerechnet.

Modell 6.1. Es seien $r, t \in \mathbb{N}$ und $\mathcal{H} \in \mathbb{C}^{r \times t}$ mit

$$\mathcal{H} = \bar{\mathcal{H}} + \tilde{\mathcal{H}},$$

wobei $\bar{\mathcal{H}} \in \mathbb{C}^{r \times t}$ fest und $\tilde{\mathcal{H}} \in \mathbb{C}^{r \times t}$ eine Zufallsmatrix, wobei für jeden Eintrag

$$\tilde{H}_{ij} \sim \text{SCN}(0, \tilde{\sigma}_h^2)$$

gelte. Die einzelnen Einträge seien dabei stochastisch unabhängig. Es sei $\underline{Z} \in \mathbb{C}^r$ mit

$$\underline{Z} \sim \text{SCN}(\underline{0}, N_0 I_r)$$

für ein $N_0 > 0$ unabhängig von dem oben definierten $\tilde{\mathcal{H}}$. Seien $P > 0$ und $\underline{X} \in \mathbb{C}^t$ unabhängig von $\tilde{\mathcal{H}}$ und \underline{Z} . Weiterhin sei

$$\mathbb{E}(\underline{X}^* \underline{X}) \leq P.$$

Der Outputvektor ist definiert als

$$\underline{Y} := \mathcal{H} \underline{X} + \underline{Z} = \bar{\mathcal{H}} \underline{X} + \tilde{\mathcal{H}} \underline{X} + \underline{Z}.$$

Die im Modell verwendeten Symbole sind in Tabelle 6.1 zusammengefasst.

TABELLE 6.1. Interpretation der Symbole im senderseitigen P-CSI MIMO-Modell

Mathematische Definition	Definition
$r \in \mathbb{N}$	Empfangsantennen (receive)
$t \in \mathbb{N}$	Sendeantennen (transmit)
$P > 0$	Leistungsbeschränkung
$\underline{X} \in \mathbb{C}^t$	Inputvektor
$\underline{Y} \in \mathbb{C}^r$	Outputvektor
$\underline{Z} \in \mathbb{C}^r$	additive Störung
$\mathcal{H} \in \mathbb{C}^{r \times t}$	Kanalmatrix
$\bar{\mathcal{H}} \in \mathbb{C}^{r \times t}$	Konstanter Teil der Kanalmatrix
$\tilde{\mathcal{H}} \in \mathbb{C}^{r \times t}$	Zufälliger Teil der Kanalmatrix
$\tilde{\sigma}_h^2 > 0$	Fading-Varianz
$N_0 > 0$	Varianz der additiven Störung

Technische Realisierung:

Man legt diesem Modell ein Modell mit Synchronised Detection zugrunde. Den Schätzer $\hat{\mathcal{H}}^D$ für die Kanalmatrix \mathcal{H} sendet der Empfänger über einen Feedback-Kanal zurück zum Sender. Durch weitere Störungen erhält der Sender die Matrix $\bar{\mathcal{H}}$. Dabei geht man davon aus, dass der Fehler $\tilde{\mathcal{H}} = \mathcal{H} - \bar{\mathcal{H}}$ ebenfalls wieder SCN-verteilt ist. Auf eine Modellierung mit Synchronised Detection im selben Modell wird in dieser Arbeit verzichtet, ebenso wie auf die Modellierung des Feedback-Kanals. Eine Veranschaulichung des Modells findet sich in Abbildung 6.1. Dabei ist noch zu beachten, dass die zusätzlichen Störungen beim Feedback im hier verwendeten Modell keine Rolle spielen, sondern vernachlässigt werden. Für die Varianz $\tilde{\sigma}_h^2$ ist zu beachten, dass

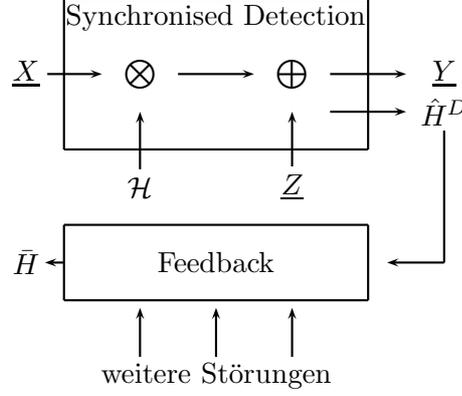
$$\tilde{\sigma}_h^2 = \sigma_h^2 - \hat{\sigma}_h^2$$

gilt, wobei $\hat{\sigma}_h^2$ die Varianz der Komponenten der geschätzten Matrix $\hat{\mathcal{H}}^D$ ist. Insbesondere gilt in der Realität daher in der Regel $\tilde{\sigma}_h^2 < \sigma_h^2$. Weiterhin ist zu beachten, dass in der Realität die Qualität der Schätzmatrix $\hat{\mathcal{H}}^D$ unter anderem von der Stärke der additiven Störung abhängt, und damit die Restvarianz $\tilde{\sigma}_h^2$ nicht unabhängig von der Varianz N_0 des Rauschens ist. Im Modell wird diese Abhängigkeitsstruktur vernachlässigt, und die beiden Zufallselemente $\tilde{\mathcal{H}}$ und \underline{Z} als unabhängig vorausgesetzt.

6.2. Bestimmung einiger Kapazitätsschranken

Man geht davon aus, dass $\bar{\mathcal{H}}$ sowohl beim Sender als auch beim Empfänger bekannt ist, $\tilde{\mathcal{H}}$ jedoch weder beim Sender noch beim Empfänger. Die Ausgabe des Kanals ist daher lediglich das \underline{Y} . Die Synentropie des Kanals ist daher gegeben durch $I(\underline{X}; \underline{Y})$. Die Maximierung dieses Ausdrucks ist ein offenes Problem. Es lassen sich jedoch sowohl obere als auch untere Schranken angeben. Eine untere Schranke für die Kapazität wird im folgenden Satz angegeben:

ABBILDUNG 6.1. Übertragungsmodell eines senderseitigen P-CSI MIMO-Kanals



Satz 6.2. Für den obigen senderseitigen P-CSI-Kanal aus Modell 6.1 gilt

$$C \geq \log \det \left(I_r + \frac{P}{t(P\tilde{\sigma}_h^2 + N_0)} \bar{\mathcal{H}}\bar{\mathcal{H}}^* \right).$$

Beweis. Da die Kapazität die Maximierung über alle Inputverteilungen bedeutet, kann eine untere Schranke durch Festhalten einer bestimmten Inputverteilung gefunden werden. Sei also

$$\underline{X} \sim \text{SCN} \left(\underline{0}, \frac{P}{t} I_t \right).$$

Dann weiß man schon, dass

$$\mathbb{H}(\underline{X}) = \log \det \left(\frac{\pi e P}{t} I_t \right) = t \log \left(\frac{\pi e P}{t} \right) \quad (6.1)$$

ist. Außerdem gilt analog zum Beweis von Satz 5.8

$$\mathbb{H}(\underline{X} | \underline{Y}) = \mathbb{H}(\underline{X} - A\underline{Y} | \underline{Y}) \leq \mathbb{H}(\underline{X} - A\underline{Y}) \leq \log \det (\pi e \text{Cov}(\underline{X} - A\underline{Y}))$$

für beliebige Matrizen $A \in \mathbb{C}^{t \times r}$. Man wähle nun speziell

$$A := \text{Cov}(\underline{X}, \underline{Y}) \text{Cov}(\underline{Y})^{-1}$$

und erhalte die folgende Darstellung

$$\begin{aligned} A &= \text{Cov}(\underline{X}, \underline{Y}) \text{Cov}(\underline{Y})^{-1} = \text{Cov} \left(\underline{X}, \bar{\mathcal{H}}\underline{X} + \tilde{\mathcal{H}}\underline{X} + \underline{Z} \right) \text{Cov}(\underline{Y})^{-1} \\ &= \left(\text{Cov}(\underline{X}, \bar{\mathcal{H}}\underline{X}) + \text{Cov}(\underline{X}, \tilde{\mathcal{H}}\underline{X}) \right) \text{Cov}(\underline{Y})^{-1}. \end{aligned}$$

Komponentenweise gilt

$$\begin{aligned}
\text{Cov} \left(\underline{X}, \tilde{\mathcal{H}} \underline{X} \right)_{ij} &= \text{Cov} \left(X_i, \left(\tilde{\mathcal{H}} \underline{X} \right)_j \right) = \text{Cov} \left(X_i, \sum_{k=1}^t \tilde{H}_{jk} X_k \right) \\
&= \sum_{k=1}^t \text{Cov} \left(X_i, \tilde{H}_{jk} X_k \right) = \text{Cov} \left(X_i, \tilde{H}_{ji} X_i \right) \\
&= \text{E} \left((X_i - \text{E} X_i) \left(\tilde{H}_{ji} X_i - \text{E} \tilde{H}_{ji} \text{E} X_i \right) \right) \\
&= \text{E} \left(X_i^2 \tilde{H}_{ji} \right) = \text{E} X_i^2 \text{E} \tilde{H}_{ji} = 0,
\end{aligned} \tag{6.2}$$

damit folgt

$$A = \text{Cov} \left(\underline{X}, \bar{\mathcal{H}} \underline{X} \right) \left(\text{Cov} \left(\bar{\mathcal{H}} \underline{X} \right) + \text{Cov} \left(\tilde{\mathcal{H}} \underline{X} \right) + N_0 I_r \right)^{-1}.$$

Unter Zuhilfenahme der beiden Formeln

$$\text{Cov} \left(\bar{\mathcal{H}} \underline{X} \right) = \bar{\mathcal{H}} \text{Cov} \left(\underline{X} \right) \bar{\mathcal{H}}^* = \frac{P}{t} \bar{\mathcal{H}} \bar{\mathcal{H}}^*$$

und

$$\text{Cov} \left(\underline{X}, \bar{\mathcal{H}} \underline{X} \right) = \text{Cov} \left(\underline{X} \right) \bar{\mathcal{H}}^* = \frac{P}{t} \bar{\mathcal{H}}^*$$

gilt

$$A = \frac{P}{t} \bar{\mathcal{H}}^* \left(\text{Cov} \left(\tilde{\mathcal{H}} \underline{X} \right) + \frac{P}{t} \bar{\mathcal{H}} \bar{\mathcal{H}}^* + N_0 I_r \right)^{-1},$$

und damit

$$\begin{aligned}
\text{Cov} \left(\underline{X} - A \underline{Y} \right) &= \text{Cov} \underline{X} - A \text{Cov} \left(\underline{Y} \right) A^* = \\
&= \frac{P}{t} I_t - \frac{P^2}{t^2} \bar{\mathcal{H}}^* \left(\text{Cov} \left(\tilde{\mathcal{H}} \underline{X} \right) + \frac{P}{t} \bar{\mathcal{H}} \bar{\mathcal{H}}^* + N_0 I_r \right)^{-1} \bar{\mathcal{H}}.
\end{aligned}$$

Nach Anwendung des Matrix Inversion Lemma 5.7 in umgekehrter Richtung folgt

$$\text{Cov} \left(\underline{X} - A \underline{Y} \right)^{-1} = \frac{t}{P} I_t + \bar{\mathcal{H}}^* \left(\text{Cov} \left(\tilde{\mathcal{H}} \underline{X} \right) + N_0 I_r \right)^{-1} \bar{\mathcal{H}}.$$

Damit gilt für die bedingte Entropie

$$\begin{aligned}
\text{H} \left(\underline{X} \mid \underline{Y} \right) &\leq \text{H} \left(\underline{X} - A \underline{Y} \right) \leq \log \det \left(\pi e \text{Cov} \left(\underline{X} - A \underline{Y} \right) \right) \\
&\leq -\log \det \left(\frac{1}{\pi e} \left(\frac{t}{P} I_t + \bar{\mathcal{H}}^* \left(\text{Cov} \left(\tilde{\mathcal{H}} \underline{X} \right) + N_0 I_r \right)^{-1} \bar{\mathcal{H}} \right) \right). \tag{6.3}
\end{aligned}$$

Damit folgt für die Synentropie insgesamt mit (6.1) und (6.3)

$$\begin{aligned}
I(\underline{X}; \underline{Y}) &= H(\underline{X}) - H(\underline{X} | \underline{Y}) \\
&\geq \log \det \left(\frac{\pi e P}{t} I_t \right) \\
&\quad + \log \det \left(\frac{1}{\pi e} \left(\frac{t}{P} I_t + \bar{\mathcal{H}}^* \left(\text{Cov}(\tilde{\mathcal{H}}\underline{X}) + N_0 I_r \right)^{-1} \bar{\mathcal{H}} \right) \right) \\
&= \log \det \left(I_t + \frac{P}{t} \bar{\mathcal{H}}^* \left(\text{Cov}(\tilde{H}\underline{X}) + N_0 I_r \right)^{-1} \bar{\mathcal{H}} \right) \\
&= \log \det \left(I_r + \frac{P}{t} \left(\text{Cov}(\tilde{H}\underline{X}) + N_0 I_r \right)^{-1} \bar{\mathcal{H}} \bar{\mathcal{H}}^* \right).
\end{aligned}$$

Die Kovarianzmatrix errechnet sich komponentenweise

$$\text{Cov}(\tilde{\mathcal{H}}\underline{X})_{ij} = \text{Cov} \left(\sum_{k=1}^t \tilde{H}_{ik} X_k, \sum_{l=1}^t \tilde{H}_{jl} X_l \right) = \delta_{ij} \sum_{k=1}^t \text{Var}(\tilde{H}_{ik} X_k),$$

wobei gilt

$$\begin{aligned}
\text{Var}(\tilde{H}_{ik} X_k) &= \text{E} \left(\tilde{H}_{ik} X_k \overline{\tilde{H}_{ik} X_k} \right) - \underbrace{\text{E}(\tilde{H}_{ik} X_k) \text{E}(\overline{\tilde{H}_{ik} X_k})}_{=0} \\
&= \text{E} \left(|\tilde{H}_{ik} X_k|^2 \right) = \text{E} \left(|\tilde{H}_{ik}|^2 \right) \text{E} \left(|X_k|^2 \right) = \frac{P \tilde{\sigma}_h^2}{t}.
\end{aligned}$$

Damit gilt

$$\begin{aligned}
I(\underline{X}; \underline{Y}) &\geq \log \det \left(I_r + \frac{P}{t} \left(\text{Cov}(\tilde{H}\underline{X}) + N_0 I_r \right)^{-1} \bar{\mathcal{H}} \bar{\mathcal{H}}^* \right) \\
&= \log \det \left(I_r + \frac{P}{t(P \tilde{\sigma}_h^2 + N_0)} \bar{\mathcal{H}} \bar{\mathcal{H}}^* \right)
\end{aligned}$$

und somit auch

$$C \geq \log \det \left(I_r + \frac{P}{t(P \tilde{\sigma}_h^2 + N_0)} \bar{\mathcal{H}} \bar{\mathcal{H}}^* \right).$$

□

Bemerkung 6.3. *Im Beweis wird die Matrix A gewählt als*

$$A = \text{Cov}(\underline{X}, \underline{Y}) \text{Cov}(\underline{Y})^{-1}.$$

Es stellt sich hier die Frage, wie gut diese Wahl des A ist, da keinerlei Einschränkung für die Wahl gemacht wurde. Für die Abschätzung nach oben möchte man den Vektor $\underline{X} - A\underline{Y}$ möglichst klein halten. Da $\text{E}\underline{X} = \text{E}\underline{Y} = \underline{0}$ gilt, folgt für die bedingte Erwartung nach Satz 2.19, dass

$$A\underline{Y} = \text{Cov}(\underline{X}, \underline{Y}) \underline{Y} = \text{E}(\underline{X} | \underline{Y})$$

gilt. Somit sieht man, dass die Wahl von A in der Tat gut ist. Dass diese Strategie zur Wahl von A nicht die beste sein muss, kann man am Beweis von Satz 5.8 und Bemerkung 5.9 sehen.

Vergleicht man dieses Ergebnis direkt mit dem Ergebnis aus Satz 5.8 für eine untere Schranke der Kapazität

$$\frac{1}{N} \sum_{i=1}^{N_D} \mathbb{E} \left(\log \det \left(I_r + \frac{P}{t(P\sigma_h^2 + N_0)} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} \right) \right)$$

im Synchronised Detection Modell, sieht man, dass die Summe bei von i unabhängigen $\hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*}$ zusammengefasst werden kann zu

$$\frac{N_D}{N} \mathbb{E} \left(\log \det \left(I_r + \frac{P}{t(P\sigma_h^2 + N_0)} \hat{\mathcal{H}}_i^D \hat{\mathcal{H}}_i^{D*} \right) \right),$$

und die untere Schranke bis auf den Faktor N_D/N gleich ist. Weiterhin ist zu beachten, dass im senderseitigen P-CSI-Modell in der Realität ebenfalls Testsymbole vorangeschickt werden, um den konstanten Teil der Kanalmatrix zu bestimmen, die die Kapazität dieses Kanals um den Faktor N_D/N verringern. Im Modell in diesem Kapitel ist vorausgesetzt, dass die Kanalschätzung und das Feedback der geschätzten Kanalmatrix zum Zeitpunkt der Übertragung bereits geschehen sind und keinen Einfluss auf die Kapazität ausüben. Des Weiteren wurden zur Bestimmung der unteren Schranke sowohl im senderseitigen P-CSI-Modell als auch im Synchronised Detection Modell jeweils die Inputvektoren auf der gleichen Inputverteilung festgehalten, nämlich symmetrisch komplex normalverteilt mit Parametern $(\underline{0}, \frac{P}{t} I_r)$. Das heißt, dass der entscheidende Faktor hier das σ_h^2 ist, das in beiden Modellen unterschiedlich ist. Während im Synchronised Detection Modell diese Varianz die der gesamten Kanalmatrix ist, ist beim senderseitigen P-CSI $\tilde{\sigma}_h^2$ die Varianz der Fehlermatrix, die in der Regel deutlich kleiner ist, je nach Güte der Kanalschätzung. Man sieht, dass die untere Kapazitätsschranke im senderseitigen P-CSI dadurch deutlich größer werden kann als bei Synchronised Detection.

Beispiel 6.4. Gegeben sei der obige Kanal mit der unteren Kapazitätsschranke aus Satz 6.2. Sei $r := 1$ und $t \gg 0$. Man nehme an, dass die Summe

$$\sum_{i=1}^t \left(\Re(\bar{H}_{1i})^2 + \Im(\bar{H}_{1i})^2 \right) \approx qt$$

ungefähr linear in t sei. Dann vereinfacht sich die Kapazitätsschranke zu

$$\begin{aligned} C &\geq \log \det \left(1 + \frac{P}{t(P\tilde{\sigma}_h^2 + N_0)} \bar{\mathcal{H}} \bar{\mathcal{H}}^* \right) \\ &= \log \left(1 + \frac{P}{t(P\tilde{\sigma}_h^2 + N_0)} \sum_{i=1}^t \bar{H}_{1i} \bar{H}_{1i} \right) \\ &= \log \left(1 + \frac{P}{t(P\tilde{\sigma}_h^2 + N_0)} \sum_{i=1}^t \left(\Re(\bar{H}_{1i})^2 + \Im(\bar{H}_{1i})^2 \right) \right) \\ &\approx \log \left(1 + \frac{Pq}{(P\tilde{\sigma}_h^2 + N_0)} \right). \end{aligned}$$

Man erkennt, dass die rechte Seite unabhängig von t ist, und damit auch insbesondere

$$\lim_{t \rightarrow \infty} \log \det \left(1 + \frac{P}{t(P\tilde{\sigma}_h^2 + N_0)} \bar{\mathcal{H}} \bar{\mathcal{H}}^* \right) = \log \left(1 + \frac{Pq}{(P\tilde{\sigma}_h^2 + N_0)} \right)$$

gilt.

Falls man in diesem Beispiel annimmt, dass die Leistungsbeschränkung P abhängig von der Anzahl t der Sendeantennen ist, etwa $P := P(t) := tc$, ergibt sich die Rechnung

$$\begin{aligned} C &\geq \log \left(1 + \frac{tc}{t(tc\tilde{\sigma}_h^2 + N_0)} \sum_{i=1}^t \left(\Re(\bar{H}_{1i})^2 + \Im(\bar{H}_{1i})^2 \right) \right) \\ &\approx \log \left(1 + \frac{cqt}{tc\tilde{\sigma}_h^2 + N_0} \right). \end{aligned}$$

Damit ergibt sich für den Grenzwert

$$\lim_{t \rightarrow \infty} \log \det \left(1 + \frac{P}{t(P\tilde{\sigma}_h^2 + N_0)} \bar{\mathcal{H}} \bar{\mathcal{H}}^* \right) = \log \left(1 + \frac{q}{\tilde{\sigma}_h^2} \right)$$

Beispiel 6.5. Sei nun $t := 1$ und $r \gg 0$, und man nehme an, dass die Summe

$$\sum_{i=1}^r \left(\Re(\bar{H}_{1i})^2 + \Im(\bar{H}_{1i})^2 \right) \approx qr$$

ungefähr linear in r sei. Damit vereinfacht sich die Schranke zu

$$\begin{aligned} C &\geq \log \det \left(I_r + \frac{P}{P\tilde{\sigma}_h^2 + N_0} \bar{\mathcal{H}} \bar{\mathcal{H}}^* \right) \\ &= \log \det \left(I_t + \frac{P}{P\tilde{\sigma}_h^2 + N_0} \bar{\mathcal{H}}^* \bar{\mathcal{H}} \right) \\ &= \log \det \left(1 + \frac{P}{P\tilde{\sigma}_h^2 + N_0} \sum_{i=1}^r \bar{H}_{1i} \bar{H}_{1i} \right) \\ &= \log \det \left(1 + \frac{P}{P\tilde{\sigma}_h^2 + N_0} \sum_{i=1}^t \left(\Re(\bar{H}_{1i})^2 + \Im(\bar{H}_{1i})^2 \right) \right) \\ &\approx \log \left(1 + \frac{Pqr}{P\tilde{\sigma}_h^2 + N_0} \right). \end{aligned}$$

Für den Grenzwert in r erhält man somit

$$\lim_{r \rightarrow \infty} \log \det \left(I_r + \frac{P}{P\tilde{\sigma}_h^2 + N_0} \bar{\mathcal{H}} \bar{\mathcal{H}}^* \right) = \infty$$

Um eine obere Schranke der Kapazität zu finden, betrachtet man zunächst die Syntropie und schätzt diese nach oben ab.

Satz 6.6. Für den obigen senderseitigen P-CSI-Kanal aus Modell 6.1 gilt

$$I(\underline{X}; \underline{Y}) \leq \log \det \left(\frac{1}{N_0} \bar{\mathcal{H}} \text{Cov}(\underline{X}) \bar{\mathcal{H}}^* + \left(\frac{\tilde{\sigma}_h^2}{N_0} \sum_{k=1}^t \text{Var}(X_k) + 1 \right) I_r \right).$$

Beweis. Mit Anwendung von Satz 2.14 und analogen Schlussfolgerungen wie im Beweis von Satz 4.16 erhält man

$$\begin{aligned} I(\underline{X}; \underline{Y}) &\leq I(\underline{X}; \underline{Y} | \mathcal{H}) = \mathbb{E}_{\mathcal{H}} (\text{H}(\underline{Y}) - \text{H}(\underline{Z})) \\ &= \mathbb{E}_{\mathcal{H}} (\text{H}(\underline{Y}) - \log \det(\pi e N_0 I_r)). \end{aligned}$$

Dieser Ausdruck bleibt zu maximieren. Die Entropie $\text{H}(\underline{Y})$ wird maximal, wenn $\underline{Y} \sim \text{SCN}(\underline{0}, \text{Cov}(\underline{Y}))$. In diesem Fall ist

$$\text{H}(\underline{Y}) = \log \det(\pi e \text{Cov}(\underline{Y})).$$

Dann ist wegen der Konvexität der Logarithmusfunktion und der Linearität der Determinante

$$\begin{aligned} \mathbb{E}_{\mathcal{H}} (\text{H}(\underline{Y})) &= \mathbb{E}_{\mathcal{H}} (\log \det(\pi e \text{Cov}(\underline{Y}))) \\ &\leq \log (\mathbb{E}_{\mathcal{H}} (\det(\pi e \text{Cov}(\mathcal{H}\underline{X} + \underline{Z})))) \\ &= \log (\mathbb{E}_{\mathcal{H}} (\det(\pi e (\mathcal{H} \text{Cov}(\underline{X}) \mathcal{H}^* + N_0 I_r)))) \\ &= \log \left(\mathbb{E}_{\mathcal{H}} \left(\det \left(\pi e \left((\bar{\mathcal{H}} + \tilde{\mathcal{H}}) \text{Cov}(\underline{X}) (\bar{\mathcal{H}} + \tilde{\mathcal{H}})^* + N_0 I_r \right) \right) \right) \right) \\ &= \log \det \left(\pi e \left(\bar{\mathcal{H}} \text{Cov}(\underline{X}) \bar{\mathcal{H}}^* + \mathbb{E} \left(\tilde{\mathcal{H}} \text{Cov}(\underline{X}) \tilde{\mathcal{H}}^* \right) + N_0 I_r \right) \right). \end{aligned}$$

Komponentenweise gilt

$$\begin{aligned} \left(\mathbb{E} \left(\tilde{\mathcal{H}} \text{Cov}(\underline{X}) \tilde{\mathcal{H}}^* \right) \right)_{ij} &= \mathbb{E} \left(\sum_{k=1}^t \sum_{l=1}^t \tilde{H}_{il} \text{Cov}(X_l, X_k) \tilde{H}_{jk} \right) \\ &= \sum_{k=1}^t \sum_{l=1}^t \text{Cov}(X_l, X_k) \mathbb{E} \left(\tilde{H}_{il} \tilde{H}_{jk} \right) \\ &= \delta_{ij} \sum_{k=1}^t \tilde{\sigma}_h^2 \text{Var}(X_k) \\ &= \delta_{ij} \tilde{\sigma}_h^2 \sum_{k=1}^t \text{Var}(X_k), \end{aligned}$$

und damit folgt insgesamt

$$\text{H}(\underline{Y}) \leq \log \det \left(\pi e \left(\bar{\mathcal{H}} \text{Cov}(\underline{X}) \bar{\mathcal{H}}^* + \left(\tilde{\sigma}_h^2 \sum_{k=1}^t \text{Var}(X_k) + N_0 \right) I_r \right) \right).$$

Damit folgt für die Synentropie insgesamt

$$I(\underline{X}; \underline{Y}) \leq \log \det \left(\frac{1}{N_0} \bar{\mathcal{H}} \text{Cov}(\underline{X}) \bar{\mathcal{H}}^* + \left(\frac{\tilde{\sigma}_h^2}{N_0} \sum_{k=1}^t \text{Var}(X_k) + 1 \right) I_r \right).$$

□

Dieses Ergebnis aus [13] ist nicht unabhängig von der Inputverteilung, daher liefert es an sich noch keine Abschätzung für die Kanalkapazität. An dieser Stelle ist es interessant zu ergründen, ob es möglich ist, durch geschicktes Abschätzen die Inputverteilung zu eliminieren und gleichzeitig die obere Schranke nicht zu groß zu machen. Ein Versuch dazu findet sich im folgenden

Satz 6.7. *Für den obigen senderseitigen P-CSI-Kanal aus Modell 6.1 gilt*

$$C \leq \sum_{i=1}^r \log \left(\frac{P \left(4H_i + H_{i,max}^2 + \tilde{\sigma}_h^2 \right)}{N_0} + 1 \right),$$

wobei

$$H_{i,max}^2 := \max \left\{ |\bar{H}_{il}|^2 \mid l = 1, \dots, t \right\}$$

und

$$H_i := \sqrt{\sum_{l=1}^t |\bar{H}_{il}|^2 \sum_{k=l+1}^t |\bar{H}_{ik}|^2}.$$

Beweis. Nach der Hadamard-Ungleichung 4.5 gilt

$$\begin{aligned} & \det \left(\frac{1}{N_0} \bar{\mathcal{H}} \text{Cov}(\underline{X}) \bar{\mathcal{H}}^* + \left(\frac{\tilde{\sigma}_h^2}{N_0} \sum_{k=1}^t \text{Var}(X_k) + 1 \right) I_r \right) \\ & \leq \prod_{i=1}^r \left| \frac{1}{N_0} \sum_{l=1}^t \sum_{k=1}^t \bar{H}_{il} \text{Cov}(X_l, X_k) \bar{H}_{ik} + \frac{\tilde{\sigma}_h^2}{N_0} \sum_{k=1}^t \text{Var}(X_k) + 1 \right| \\ & \leq \prod_{i=1}^r \left(\frac{1}{N_0} \sum_{l=1}^t \sum_{k=1}^t |\bar{H}_{il}| |\text{Cov}(X_l, X_k)| |\bar{H}_{ik}| + \frac{\tilde{\sigma}_h^2}{N_0} \sum_{k=1}^t \text{Var}(X_k) + 1 \right) \\ & \leq \prod_{i=1}^r \left(\frac{1}{N_0} \sum_{l=1}^t \sum_{k=1}^t |\bar{H}_{il}| |\text{Cov}(X_l, X_k)| |\bar{H}_{ik}| + \frac{\tilde{\sigma}_h^2 P}{N_0} + 1 \right). \end{aligned}$$

Zur Abschätzung der $|\text{Cov}(X_l, X_k)|$ ist zu beachten, dass für die Korrelationskoeffizienten nach Satz 2.7 gilt

$$\begin{aligned} |\text{Cov}(X_l, X_k)| & \leq \sqrt{\text{Var}(\Re X_l) \text{Var}(\Re X_k)} + \sqrt{\text{Var}(\Re X_l) \text{Var}(\Im X_k)} \\ & \quad + \sqrt{\text{Var}(\Im X_l) \text{Var}(\Re X_k)} + \sqrt{\text{Var}(\Im X_l) \text{Var}(\Im X_k)}. \end{aligned}$$

Mit dieser Abschätzung folgt

$$\begin{aligned}
& \sum_{l=1}^t \sum_{k=1}^t |\bar{H}_{il}| |\text{Cov}(X_l, X_k)| \left| \bar{\bar{H}}_{ik} \right| \\
& \leq 2 \sum_{l=1}^t \sum_{k=l+1}^t \sqrt{\text{Var}(\Re X_l) \text{Var}(\Re X_k)} \left| \bar{H}_{il} \bar{\bar{H}}_{ik} \right| \\
& \quad + 2 \sum_{l=1}^t \sum_{k=l+1}^t \sqrt{\text{Var}(\Re X_l) \text{Var}(\Im X_k)} \left| \bar{H}_{il} \bar{\bar{H}}_{ik} \right| \\
& \quad + 2 \sum_{l=1}^t \sum_{k=l+1}^t \sqrt{\text{Var}(\Im X_l) \text{Var}(\Re X_k)} \left| \bar{H}_{il} \bar{\bar{H}}_{ik} \right| \\
& \quad + 2 \sum_{l=1}^t \sum_{k=l+1}^t \sqrt{\text{Var}(\Im X_l) \text{Var}(\Im X_k)} \left| \bar{H}_{il} \bar{\bar{H}}_{ik} \right| \\
& \quad + \sum_{l=1}^t \text{Var}(X_l) \left| \bar{H}_{il} \right|^2,
\end{aligned}$$

und damit auch

$$\begin{aligned}
& \sum_{l=1}^t \sum_{k=1}^t |\bar{H}_{il}| |\text{Cov}(X_l, X_k)| \left| \bar{\bar{H}}_{ik} \right| \\
& \leq 2 \sum_{l=1}^t \sum_{k=l+1}^t \sqrt{\text{Var}(\Re X_l) \text{Var}(\Re X_k)} \left| \bar{H}_{il} \bar{\bar{H}}_{ik} \right| \\
& \quad + 2 \sum_{l=1}^t \sum_{k=l+1}^t \sqrt{\text{Var}(\Re X_l) \text{Var}(\Re X_k)} \left| \bar{H}_{il} \bar{\bar{H}}_{ik} \right| \\
& \quad + 2 \sum_{l=1}^t \sum_{k=l+1}^t \sqrt{\text{Var}(\Re X_l) \text{Var}(\Im X_k)} \left| \bar{H}_{il} \bar{\bar{H}}_{ik} \right| \\
& \quad + 2 \sum_{l=1}^t \sum_{k=l+1}^t \sqrt{\text{Var}(\Re X_l) \text{Var}(\Im X_k)} \left| \bar{H}_{il} \bar{\bar{H}}_{ik} \right| + PH_{i,max}^2.
\end{aligned}$$

Mit zweimaliger Anwendung der Ungleichung von Cauchy-Schwarz folgt die Abschätzung

$$\begin{aligned}
& \sum_{l=1}^t \sum_{k=l+1}^t \sqrt{\text{Var}(\Re X_l) \text{Var}(\Re X_k)} |\bar{H}_{il} \bar{H}_{ik}| & (6.4) \\
&= \sum_{l=1}^t \sqrt{\text{Var}(\Re X_l)} |\bar{H}_{il}| \sum_{k=l+1}^t \sqrt{\text{Var}(\Re X_k)} |\bar{H}_{ik}| \\
&\leq \sum_{l=1}^t \sqrt{\text{Var}(\Re X_l)} |\bar{H}_{il}| \sqrt{\sum_{k=l+1}^t \text{Var}(\Re X_k) \sum_{k=l+1}^t |\bar{H}_{ik}|} \\
&\leq \sqrt{\sum_{l=1}^t \text{Var}(\Re X_l) \sum_{l=1}^t |\bar{H}_{il}|^2 \sum_{k=l+1}^t \text{Var}(\Re X_k) \sum_{k=l+1}^t |\bar{H}_{ik}|} \\
&\leq \sum_{l=1}^t \text{Var}(\Re X_l) \sqrt{\sum_{l=1}^t |\bar{H}_{il}|^2 \sum_{k=l+1}^t |\bar{H}_{ik}|^2}.
\end{aligned}$$

Durch analoge Rechnungen ergeben sich

$$\begin{aligned}
& \sum_{l=1}^t \sum_{k=l+1}^t \sqrt{\text{Var}(\Re X_l) \text{Var}(\Im X_k)} |\bar{H}_{il} \bar{H}_{ik}| & (6.5) \\
&\leq \sum_{l=1}^t \sqrt{\text{Var}(\Re X_l) \text{Var}(\Im X_l)} \sqrt{\sum_{l=1}^t |\bar{H}_{il}|^2 \sum_{k=l+1}^t |\bar{H}_{ik}|^2},
\end{aligned}$$

$$\begin{aligned}
& \sum_{l=1}^t \sum_{k=l+1}^t \sqrt{\text{Var}(\Im X_l) \text{Var}(\Re X_k)} |\bar{H}_{il} \bar{H}_{ik}| & (6.6) \\
&\leq \sum_{l=1}^t \sqrt{\text{Var}(\Re X_l) \text{Var}(\Im X_l)} \sqrt{\sum_{l=1}^t |\bar{H}_{il}|^2 \sum_{k=l+1}^t |\bar{H}_{ik}|^2}
\end{aligned}$$

und

$$\begin{aligned}
& \sum_{l=1}^t \sum_{k=l+1}^t \sqrt{\text{Var}(\Im X_l) \text{Var}(\Im X_k)} |\bar{H}_{il} \bar{H}_{ik}| & (6.7) \\
&\leq \sum_{l=1}^t \text{Var}(\Im X_l) \sqrt{\sum_{l=1}^t |\bar{H}_{il}|^2 \sum_{k=l+1}^t |\bar{H}_{ik}|^2}.
\end{aligned}$$

Man erkennt, dass die beiden mittleren Summanden (6.5) und (6.6) die gleiche obere Schranke haben, sowie dass für die Summe des ersten (6.4) und letzten

Summanden (6.7)

$$\begin{aligned} & \sum_{l=1}^t \sum_{k=l+1}^t \left(\sqrt{\text{Var}(\Re X_l) \text{Var}(\Re X_k)} + \sqrt{\text{Var}(\Im X_l) \text{Var}(\Im X_k)} \right) |\bar{H}_{il} \bar{H}_{ik}| \\ & \leq \left(\sum_{l=1}^t \text{Var}(\Re X_l) + \sum_{l=1}^t \text{Var}(\Im X_l) \right) H_i = \left(\sum_{l=1}^t \text{Var}(X_l) \right) H_i \leq P H_i \end{aligned}$$

gilt.

Um die beiden mittleren Summanden (6.5) und (6.6) abzuschätzen, betrachtet man zunächst die binomische Formel

$$\begin{aligned} 0 & \leq \left(\sqrt{\text{Var}(\Re X_l)} - \sqrt{\text{Var}(\Im X_l)} \right)^2 \\ & = \text{Var}(\Re X_l) - 2\sqrt{\text{Var}(\Re X_l)}\sqrt{\text{Var}(\Im X_l)} + \text{Var}(\Im X_l) \end{aligned}$$

für alle $l = 1, \dots, t$.

Damit erhält man

$$2\sqrt{\text{Var}(\Re X_l) \text{Var}(\Im X_l)} \leq \text{Var}(\Re X_l) + \text{Var}(\Im X_l),$$

und somit für die Summe der mittleren Summanden (6.5) und (6.6)

$$\begin{aligned} & \sum_{l=1}^t \sum_{k=l+1}^t \left(\sqrt{\text{Var}(\Re X_l) \text{Var}(\Im X_k)} + \sqrt{\text{Var}(\Im X_l) \text{Var}(\Re X_k)} \right) |\bar{H}_{il} \bar{H}_{ik}| \\ & \leq \sum_{l=1}^t 2\sqrt{\text{Var}(\Re X_l) \text{Var}(\Im X_l)} H_i \\ & \leq \sum_{l=1}^t (\text{Var}(\Re X_l) + \text{Var}(\Im X_l)) H_i \leq P H_i \end{aligned}$$

Damit ergibt sich die Abschätzung der Determinante

$$\begin{aligned} \det \left(\frac{1}{N_0} \bar{\mathcal{H}} \text{Cov}(\underline{X}) \bar{\mathcal{H}}^* + \left(\frac{\tilde{\sigma}_h^2}{N_0} \sum_{k=1}^t \text{Var}(X_k) + 1 \right) I_r \right) \\ \leq \prod_{i=1}^r \left(\frac{P(4H_i + H_{i,max}^2 + \tilde{\sigma}_h^2)}{N_0} + 1 \right). \end{aligned}$$

Somit hat man für die Synentropie

$$\begin{aligned}
I(\underline{X}; \underline{Y}) &\leq \log \det \left(\frac{1}{N_0} \bar{\mathcal{H}} \text{Cov}(\underline{X}) \bar{\mathcal{H}}^* + \left(\frac{\tilde{\sigma}_h^2}{N_0} \sum_{k=1}^t \text{Var}(X_k) + 1 \right) I_r \right) \\
&\leq \log \left(\prod_{i=1}^r \left(\frac{P(4H_i + H_{i,max}^2 + \tilde{\sigma}_h^2)}{N_0} + 1 \right) \right) \\
&= \sum_{i=1}^r \log \left(\frac{P(4H_i + H_{i,max}^2 + \tilde{\sigma}_h^2)}{N_0} + 1 \right).
\end{aligned}$$

□

Im Gegensatz zu den unteren Kapazitätsschranken im Synchronised Detection Modell 5.1 und diesem, die ähnlich aussehen, weicht die obere Schranke sehr von der aus Korollar 5.11 ab. Aus den gleichen Überlegungen wie vorher kann man die obere Schranke des Synchronised Detection Modells vereinfachen zu

$$r \log \left(\frac{P\sigma_h^2}{tN_0} \right) + \log \det \left(I_r + \frac{1}{\sigma_h^2} \bar{\mathcal{H}} \bar{\mathcal{H}}^* \right).$$

Zunächst kann man die beiden Schranken direkt an einigen einfachen Beispielen vergleichen.

Beispiel 6.8. Seien $r = t =: n$ und $\bar{\mathcal{H}} := I_n$. Für die Schranke im Synchronised Detection Modell folgt dann

$$\begin{aligned}
C &\leq n \log \left(\frac{P\sigma_h^2}{nN_0} \right) + \log \det \left(I_n + \frac{1}{\sigma_h^2} I_n \right) \\
&= n \log \left(\frac{P\sigma_h^2}{nN_0} \right) + \log \left(\left(1 + \frac{1}{\sigma_h^2} \right)^n \right) \\
&= n \log \left(\frac{P\sigma_h^2}{nN_0} \right) + n \log \left(1 + \frac{1}{\sigma_h^2} \right) \\
&= n \log \left(\frac{P(1 + \sigma_h^2)}{nN_0} \right).
\end{aligned}$$

Im senderseitig partiell bekannten Modell hingegen gilt

$$\begin{aligned}
C &\leq \sum_{i=1}^n \log \left(\frac{P(4H_i + H_{i,max}^2 + \tilde{\sigma}_h^2)}{N_0} + 1 \right) \\
&= n \log \left(\frac{P(1 + \tilde{\sigma}_h^2)}{N_0} + 1 \right).
\end{aligned}$$

Man kann erkennen, dass die zweite Schranke in der Regel etwas größer ist. Dieses Ergebnis war anzunehmen, da durch das Feedback eine Vergrößerung der Kapazität angestrebt wurde. Ob dies tatsächlich erreicht wurde, ist durch diese Beispiele jedoch nicht beweisbar. Zu beachten ist außerdem noch, dass im

Fall einer einzelnen Sende- und Empfangsantenne (SISO) $n = 1$ gilt, und bei sehr guter Kanalschätzung die Varianz des Fehlers $\tilde{\sigma}_h^2$ im senderseitigen P-CSI Modell sehr klein wird. So ist es durchaus möglich, dass die obere Schranke in diesem Modell sogar etwas kleiner wird als im Synchronised Detection Modell. Dies ist genau der Fall, wenn $\tilde{\sigma}_h^2 < \sigma_h^2 - N_0/P$ gilt, sofern die rechte Seite positiv bleibt.

Beispiel 6.9. Sei $\bar{\mathcal{H}} := 1_{r \times t}$. Zur Berechnung der Schranken benötigt man noch die Determinante der Matrix $I_r + (1/\sigma_h^2)1_{r \times t}1_{r \times t}^* = I_r + (t/\sigma^2)1_{r \times r}$. Durch Induktion folgt

$$\begin{aligned}
\det \left(I_r + \frac{t}{\sigma_h^2} 1_{r \times r} \right) &= \det \begin{pmatrix} 1 + \frac{t}{\sigma_h^2} & \frac{t}{\sigma_h^2} & \cdots & \frac{t}{\sigma_h^2} \\ \frac{t}{\sigma_h^2} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{t}{\sigma_h^2} \\ \frac{t}{\sigma_h^2} & \cdots & \frac{t}{\sigma_h^2} & 1 + \frac{t}{\sigma_h^2} \end{pmatrix}_{r \times r} \\
&= \det \begin{pmatrix} 1 + \frac{t}{\sigma_h^2} & \frac{t}{\sigma_h^2} & \cdots & \cdots & \frac{t}{\sigma_h^2} \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}_{r \times r} \\
&= 1 + \frac{t}{\sigma_h^2} + \det \begin{pmatrix} \frac{t}{\sigma_h^2} & \cdots & \cdots & \cdots & \frac{t}{\sigma_h^2} \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}_{(r-1) \times (r-1)} \\
&= 1 + \frac{t}{\sigma_h^2} + \frac{t}{\sigma_h^2} + \det \begin{pmatrix} \frac{t}{\sigma_h^2} & \cdots & \cdots & \cdots & \frac{t}{\sigma_h^2} \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}_{(r-2) \times (r-2)} \\
&= \dots = 1 + \frac{rt}{\sigma_h^2}.
\end{aligned}$$

Damit erhält man für die obere Schranke im Fall von Synchronised Detection

$$C \leq r \log \left(\frac{P\sigma_h^2}{tN_0} \right) + \log \det \left(I_r + \frac{t}{\sigma_h^2} 1_{r \times r} \right) = r \log \left(\frac{P\sigma_h^2}{tN_0} \right) + \log \left(1 + \frac{rt}{\sigma_h^2} \right).$$

Im Fall des senderseitig partiell bekannten Kanals hat man, wenn man beachtet, dass $H_i \equiv \sqrt{t(t-1)}/2$ und $H_{i,max}^2 \equiv 1$ gilt,

$$C \leq r \log \left(\frac{P \left(2\sqrt{2t(t-1)} + 1 + \tilde{\sigma}_h^2 \right)}{N_0} + 1 \right).$$

Dann gilt

$$r \log \left(\frac{P\sigma_h^2}{tN_0} \right) + \log \left(1 + \frac{rt}{\sigma_h^2} \right) \leq r \log \left(\frac{P \left(2\sqrt{2t(t-1)} + 1 + \tilde{\sigma}_h^2 \right)}{N_0} + 1 \right)$$

genau dann, wenn

$$\frac{\sigma_h^2}{t} \sqrt{1 + \frac{rt}{\sigma_h^2}} \leq 2\sqrt{2t(t-1)} + 1 + \tilde{\sigma}_h^2 + \frac{N_0}{P}.$$

Durch den letzten Summanden kann die rechte Seite beliebig groß werden. Da er auf jeden Fall größer als Null ist, kann man ihn vernachlässigen. In der Realität hätte eine Erhöhung von N_0 eine Verschlechterung der Kanalschätzung, und damit eine gleichzeitige Erhöhung von $\tilde{\sigma}_h^2$ zur Folge, da dies aber ebenfalls die rechte Seite vergrößert, und $\tilde{\sigma}_h^2$ auch nicht größer werden kann als σ_h^2 , wird dieser Effekt hier nicht weiter beachtet.

Im Fall, dass $r = t =: n = 1$ und $\tilde{\sigma}_h^2 = \sigma_h^2$ gilt, unterscheiden sich die beiden Seiten nur durch den Summanden N_0/P . Bei wachsenden n wird die linke Seite kleiner, sowie die rechte Seite größer. Wenn allerdings $\tilde{\sigma}_h^2$ kleiner wird, wird auch die rechte Seite kleiner. Damit ist auch in diesem Beispiel eine Möglichkeit gefunden, wo die obere Schranke im senderseitigen P-CSI kleiner ist als bei Synchronised Detection.

KAPITEL 7

Simulationen

In den vorigen Kapiteln wurden insgesamt vier Modelle behandelt:

- Die Kanalmatrix ist fest (d.h. senderseitig vollständig bekannt),
- zufällig (d.h. empfängerseitig vollständig bekannt),
- Synchronised Detection (d.h. die Kanalmatrix ist empfängerseitig partiell bekannt),
- und schließlich ein senderseitig partiell bekannter Kanal.

Für diese Modelle wurden jeweils Ausdrücke für die Kapazität hergeleitet bzw. eine obere und eine untere Schranke gefunden.

Um diese Ergebnisse grafisch darzustellen und vergleichen zu können, wurden einige Simulationen durchgeführt. Der zugehörige Quellcode findet sich im Anhang A. Das Szenario wird mit konstanter Anzahl Empfangsantennen $r := 3$ gewählt, da hier die Determinantenberechnung noch nicht zu aufwändig wird. Die Leistungsbeschränkung wird gewählt als $P := 10$, und die Varianzen als $\sigma_h^2 := 1$ und $N_0 := 1$. Die Varianz $\tilde{\sigma}_h^2$ der Fehlermatrix \tilde{H} sei 1/10 der Varianz der Kanalmatrix. Die Anzahl der Empfangsantennen t läuft von 1 bis 15. Für jedes t werden 10000 unabhängige Kanalmatrizen erzeugt, die für jedes Modell einheitlich verwendet werden, entweder als bekannte Kanalmatrix, als geschätzte Kanalmatrix oder als unbekannte Kanalmatrix innerhalb des Erwartungswertes. Als Basis der Logarithmen wird jeweils e gewählt. Alle Ergebnisse finden sich im Diagramm in Abbildung 7.1

Die blaue Linie bezeichnet die Kapazität bei fester Kanalmatrix aus Kapitel 4.1, die grüne Linie die Kapazität bei zufälliger Kanalmatrix aus Abschnitt 4.2 Die schwarzen Linien sind die Kapazitätsschranken von Synchronised Detection aus Kapitel 5 und die roten Linien die Schranken von senderseitigem P-CSI aus Kapitel 6.

Man sieht sofort, dass die Kapazität bei fester Kanalmatrix wie erwartet wesentlich schneller wächst als bei zufälliger Kanalmatrix. Bei geringer Anzahl von Sendeantennen ist jedoch kein großer Unterschied feststellbar.

Weiterhin fällt auf, dass die obere Schranke des Synchronised Detection Modells für große t sich der Kapazität bei zufälliger Kanalmatrix annähert. D.h., dass in diesem Fall die Kapazität bei senderseitig partiell bekannter Kanalmatrix nicht besser wird als bei vollständig bekannter Kanalmatrix.

Verlauf der Kapazität bei wachsendem t und konst. P

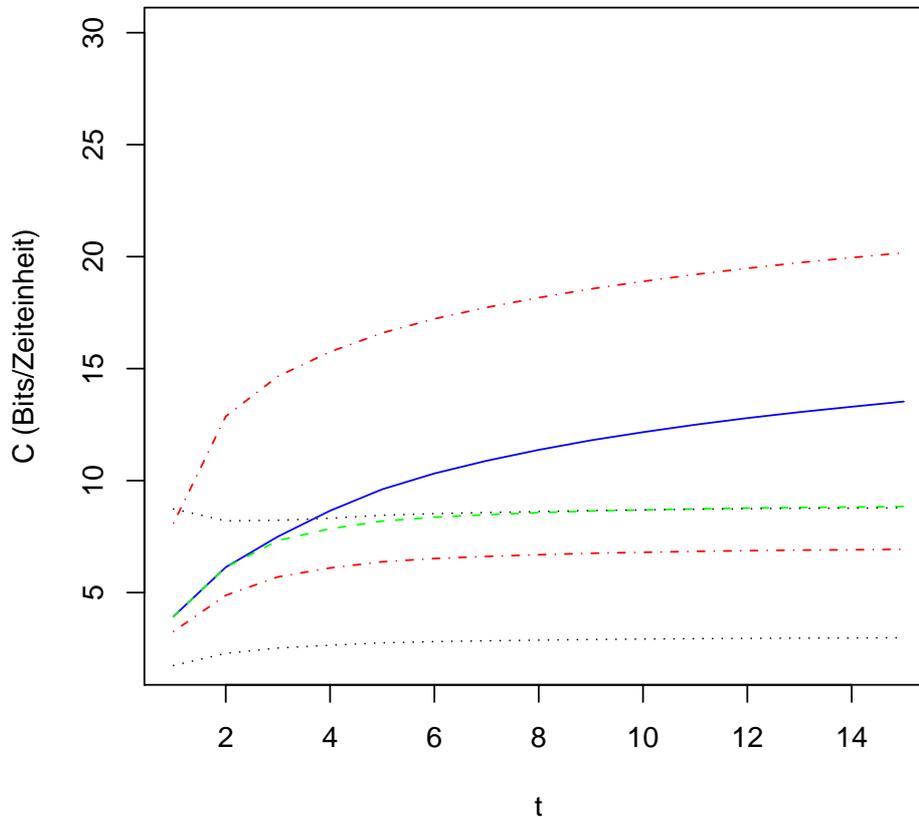


ABBILDUNG 7.1. Simulationsergebnisse für konstantes P

Bisher wurde davon ausgegangen, dass die Leistungsbeschränkung P unabhängig von der Anzahl der Sendeantennen t konstant ist. Es ist davon auszugehen, dass in der Realität jede einzelne Sendeantenne die gleiche Leistung besitzt, und die Leistungsbeschränkung daher linear in t ist. Wenn man in der Simulation alle P durch Pt ersetzt, erhält man das Diagramm in Abbildung 7.2, wo jede Sendeantenne mit Leistung $P = 10$ sendet.

Insgesamt erkennt man, dass eine Erhöhung der Anzahl t der Sendeantennen eine wesentlich größere Auswirkung auf die Kapazität hat, wenn t klein ist, als wenn t groß ist, und zwar in jedem Modell.

Man kann erkennen, dass die Kapazität bei fester Kanalmatrix im Vergleich zur zufälligen Kanalmatrix zwar bei wachsendem t noch schneller wächst, aber nicht mehr so stark wie im vorigen Diagramm.

Die obere Schranke von Synchronised Detection wächst jetzt weitaus schneller als alle anderen Werte.

Verlauf der Kapazität bei wachsendem t und konst. P/t

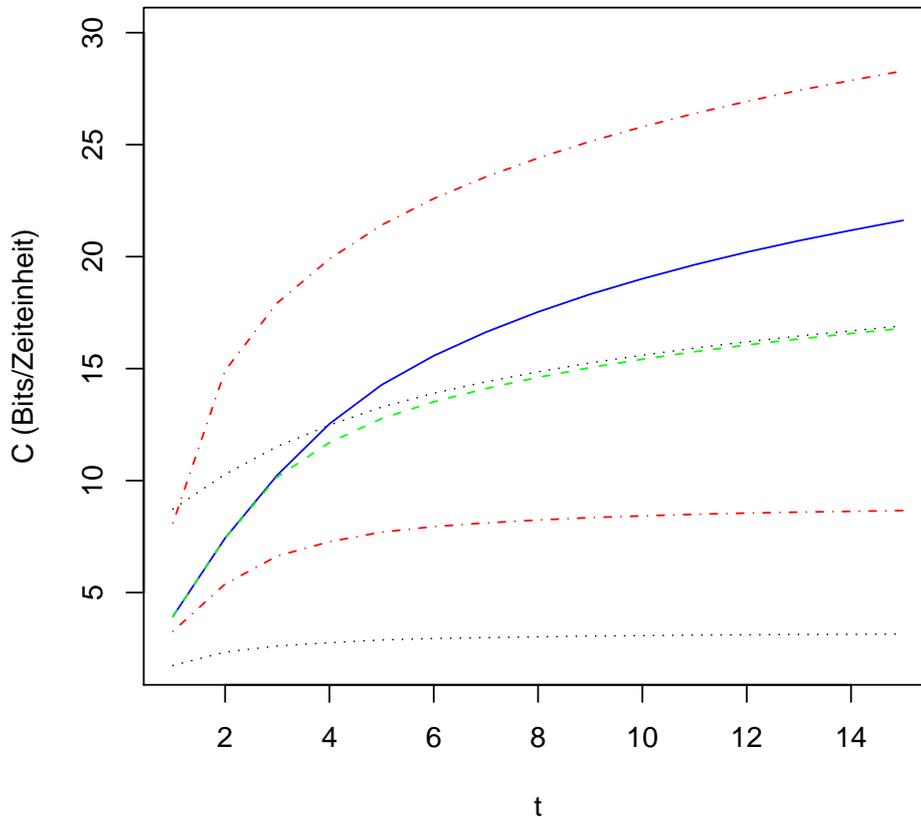


ABBILDUNG 7.2. Simulationsergebnisse für konstantes P/t

Intuitiv geht man davon aus, dass die Kapazität von senderseitig partiell bekannten Kanälen nach oben durch die Kapazität von senderseitig vollständig bekannten Kanälen beschränkt ist, da der Sender in diesem Fall weniger Kanalinformation besitzt. Es ist unbekannt, ob sich diese Vermutung beweisen lässt. Die in dieser Arbeit angegebene obere Schranke liegt jedenfalls in diesem Simulationsszenario weit über der Kapazität des senderseitig vollständig bekannten Kanals.

Analog kann man vermuten, dass die Kapazität des empfängerseitig vollständig bekannten Kanals eine obere Schranke für die Kapazität des empfängerseitig partiell bekannten Kanals ist. Ob dies zutrifft, ist ebenfalls unbekannt. Die hier angegebene obere Schranke ist bei konstantem P/t ebenfalls immer größer als die Kapazität von empfängerseitig vollständig bekannten Kanälen, nähert sich bei wachsendem t aber an. Bei konstantem P werden die beiden Werte bei $t = 10$ annähernd gleich, bei $t > 10$ bleibt dann die obere Schranke für die

Legende

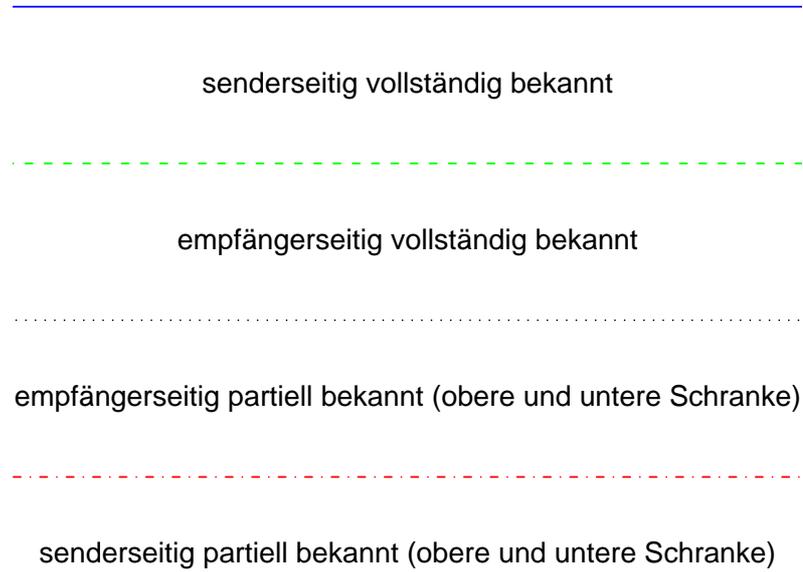


ABBILDUNG 7.3. Legende

Kapazität von empfängerseitig partiell bekannten Kanälen ein wenig kleiner als die Kapazität von empfängerseitig vollständig bekannten Kanälen.

Die Simulation wurde mit dem Statistikprogramm R programmiert [14].

Zusammenfassung und Ausblick

In dieser Arbeit wird zunächst die Kapazität von senderseitig vollständig bekannten MIMO-Kanälen bei fester sowie bei stochastischer Kanalmatrix angegeben. Im Falle der festen Kanalmatrix geschieht dies mit Hilfe von “Water-Filling”, im Falle der zufälligen Kanalmatrix durch explizites Ausrechnen der Synentropie, die durch einen Erwartungswert ausgedrückt wird. An einigen Beispielen werden Grenzwertfälle untersucht, wie die Kapazität wächst, wenn die Anzahl der Sende- bzw. Empfangsantennen gegen Unendlich läuft.

Danach wird das Synchronised Detection Modell behandelt. Zunächst wird dieses Modell mathematisch modelliert; anschließend gezeigt, dass die Synentropie dieses Kanals nur noch von den geschätzten Kanalmatrizen und den Datensymbolen abhängt. Daraus folgt, dass die Synentropie als Summe der komponentenweisen Synentropien ausgedrückt werden kann. Durch komponentenweises Ausrechnen können also sowohl eine untere als auch eine obere Schranke der Kanalkapazität gefunden werden. Die untere Schranke erhält man dabei, indem die Inputverteilung als symmetrisch komplex normalverteilt festgehalten wird und die Synentropie explizit ausgerechnet; die obere, indem zunächst eine Schranke für die Transinformation gefunden wird, aus der die Verteilung der Inputsymbole noch eliminiert wird. Auch hier werden Grenzwertbetrachtungen für wachsende Anzahl von Sende- bzw. Empfangsantennen gemacht. Außerdem wird die Verteilung der Kanalschätzer explizit angegeben.

Durch Feedback kann der Schätzer der Kanalmatrix zurück zum Sender übertragen werden, der dann ebenfalls partielles Wissen über den Kanal besitzt. Dieses System wird modelliert, sowie Schranken für die Kapazität angegeben. Wie bei Synchronised Detection findet man die untere Schranke durch explizites Errechnen der Synentropie bei symmetrisch komplex normalverteilten Inputsymbolen, und die obere Schranke durch Errechnen einer Schranke der Synentropie und folgender Eliminierung der Inputverteilung. Anschließend werden die Schranken von Synchronised Detection und senderseitigem P-CSI an einigen Beispielen explizit verglichen, und gezeigt, dass je nach Rahmenbedingungen die obere Schranke von Synchronised Detection sogar besser sein kann.

Die Arbeit schließt ab mit Simulationen aller behandelten Modelle, entweder durch explizite Berechnungen der eigentlichen Kapazitäten oder der Schranken.

Im Paper [6] werden die wichtigsten offenen Probleme im Bereich der Kapazitätsbestimmung vom MIMO-Kanälen zusammengefasst, die hier auszugsweise erläutert werden. Zum einen sei hier das Kovarianz-Feedback Modell genannt,

bei dem vorausgesetzt wird, dass die Kanalmatrix mittelwertfrei ist, und in eine senderseitige und empfängerseitige Kovarianzmatrix zerfällt. Dieses Modell ist durch Feldversuche als hinreichend akkurat bestätigt worden, die Fadingkorrelationen in Mobilnetzen zu repräsentieren, jedoch ist für viele Szenarien die Kapazität noch unbekannt. Als zweites seien die Fälle genannt, in denen senderseitig nichts über die Kanalmatrix bekannt sei, und empfängerseitig nur die Verteilung der Kanalmatrix, nicht aber einzelne Komponenten. In solchen Szenarien ist in fast allen Fällen die Bestimmung der Kapazität ungelöst.

Weiterhin seien Mehrbenutzer MIMO Kanäle genannt, wo es wegen der höheren Komplexität der Modelle weitaus mehr offene Probleme gibt als bei Einbenutzer MIMO Kanälen. Auch hier sei das Paper [6] genannt, in dem der aktuelle Stand der Forschung zusammengefasst wird, und ungelöste Probleme gesammelt werden.

ANHANG A

Quellcode

Hier finden sich ein paar Auszüge des Quelltextes zu den Simulationsversuchen mit einigen Erläuterungen.

Die Variablennamen entsprechen weitestgehend den Bezeichnungen aus der Arbeit. Die Anzahl der Sendeantennen t variiert von 1 bis $tmax$.

```
# Empfangsantennen (nicht aendern wegen Determinantenberechnung!)
r <- 3
# Anzahl der Sendeantennen variabel bis tmax
tmax <- 15
# Varianz der Kanalmatrix
sigma <- 1
# Varianz des Rauschens
NO <- 1
# Varianz der Fehlermatrix
fsigma <- sigma/10
# Leistungsbeschaenkung
P <- 10
# Einheitsmatrix
I <- diag(r)
```

Die Kanalmatrix H wird jeweils mit stochastisch unabhängigen normalverteilten Einträgen gefüllt. Real- und Imaginärteil sind dabei jeweils wieder stochastisch unabhängig.

```
H <- matrix(nrow = r, ncol = t)
for (i in 1:r)
{
  for (j in 1:t)
  {
    H[i,j] <- complex(real = rnorm(1, mean=0, sd=sigma),
                      imag = rnorm(1, mean=0, sd=sigma))
  }
}
HHstern <- H %*% Conj(t(H))
```

Die Simulation von senderseitig vollständig bekannten Kanälen unterscheidet sich von den weiteren Simulationen stark. Hier muss das μ anhand von "Water-Filling" bestimmt werden. Der Einfachheit halber geht das Programm davon aus, dass spätestens nach dem zweiten Eigenwert das Water-Filling erreicht

ist. Sollte es bis dahin noch nicht erreicht sein, gibt das Programm eine Fehlermeldung aus. Bei etwa 500000 zufällig gewählten Kanalmatrizen trat dieses Problem genau ein mal auf. Somit kann man davon ausgehen, dass dieser Effekt vernachlässigbar ist und für die Simulationsergebnisse keine Rolle spielt.

```

# Senderseitig vollstaendig bekannt
lambda <- eigen(HHstern)$values
my <- P/min(r,t)
for (i in 1:min(r,t))
{
  my <- my + lambda[i]^1/min(r,t)
}
summe <- 0
for (i in 1:min(r,t))
{
  if (my>lambda[i]^1)
    summe <- summe + (my - lambda[i]^1)
}
if (abs(P - summe) > fehler)
{
  my <- P/(min(r,t)-1)
  for (i in 1:(min(r,t)-1))
  {
    my <- my + lambda[i]^1/(min(r,t)-1)
  }
  summe <- 0
  for (i in 1:(min(r,t)-1))
  {
    if (my>lambda[i]^1)
      summe <- summe + (my - lambda[i]^1)
  }
  if (abs(P - summe) > fehler)
  {
    # Dieser Fall tritt so selten ein,
    # dass er nicht beruecksichtigt wird.
    cat("Fehler! Bitte starten
        Sie die Simulation neu. \n")
  }
}
summe <- 0
for (i in 1:min(r,t))
{
  if (my>0)
  if (log(my*lambda[i], base=exp(1))>0)
    summe <- summe + log(my*lambda[i], base=exp(1))
}
sv[t] <- sv[t] + summe

```

Die Programmierung der anderen Modelle geschieht vergleichsweise einfach. Bei der Anzahl der Empfangsantennen $r = 3$ ergeben sich bloß Determinanten der Dimension 3 zum berechnen, so dass die Programmierung keine Schwierigkeiten darstellt. R besitzt zwar eine Funktion zur Berechnung von Determinanten, da diese jedoch nur für reelle Matrizen funktioniert, lässt sie sich für dieses Programm nicht verwenden. Aus diesem Grund werden die Determinanten im Programm mit der Sarrus-Regel berechnet. Als Beispiel findet sich hier die Simulation des empfängerseitig vollständig bekannten Kanals.

```
# Empfaengerseitig vollstaendig bekannt
evmat <- I + P/t * HHstern
evdet <- evmat[1,1]*evmat[2,2]*evmat[3,3]
      + evmat[2,1]*evmat[3,2]*evmat[1,3]
      + evmat[1,2]*evmat[2,3]*evmat[3,1]
      - evmat[1,3]*evmat[2,2]*evmat[3,1]
      - evmat[2,3]*evmat[3,2]*evmat[1,1]
      - evmat[1,2]*evmat[2,1]*evmat[3,3]
# Imaginaerteil = 0. Rundungsfehler
# muessen ausgeschlossen werden.
evdet <- Re(evdet)
ev[t] <- ev[t] + log(evdet, base=exp(1))
```

Auch die obere Schranke des senderseitig partiell bekannten Kanals ist ein Spezialfall, bei dem sich aber keine weiteren Probleme ergeben. Es werden zunächst die Arrays $Hmax$ und Hi berechnet, und in die Formel für die obere Schranke eingesetzt.

```
# Senderseitig partiell bekannt (obere Schranke)
for (i in 1:r)
{
  Hmax[i] <- max(abs(H[i,1:t])^2)
  summeR <- 0
  if (t>1)
  for (l in 1:(t-1))
  {
    summeR <- summeR
      + abs(H[i,l])^2*sum(abs(H[i,(l+1):t])^2)
  }
  Hi[i] <- sqrt(summeR)
  spo[t] <- spo[t]
  + log(P*(4*Hi[i]+Hmax[i]+fsigma)/N0+1, base=exp(1))
}
```


Literaturverzeichnis

- [1] Brian D. O. Anderson, John B. Moore, *Optimal Filtering* (Prentice-Hall, 1979)
- [2] Jens Baltersee, Gunnar Fock, Heinrich Meyr, *An Information Theoretic Foundation of Synchronized Detection* (IEEE Transactions on Communications, Vol. 49, No. 12, S. 2115-2123, Dezember 2001)
- [3] Jens Baltersee, Gunnar Fock, Heinrich Meyr, *Achievable Rate of MIMO Channels with Data-Aided Channel Estimation and Perfect Interleaving* (IEEE Journal on Selected Areas in Communications, Vol. 19, No. 12, S. 2358-2368, Dezember 2001)
- [4] Chen-Nee Chuah, David N. C. Tse, Joseph M. Kahn, Reinaldo A. Valenzuela, *Capacity Scaling in MIMO Wireless Systems Under Correlated Fading* (IEEE Transactions on Information Theory, Vol. 48, No. 3, S. 637-650, März 2002)
- [5] Peter Deuffhard, Andreas Hohmann, *Numerische Mathematik I* (de Gruyter, 1993)
- [6] Andrea Goldsmith, Syed Ali Jafar, Nihar Jindal, Sriram Vishwanath, *Capacity Limits of MIMO channels* (IEEE Journal on Selected Areas in Communications, Vol. 21, No. 5, Juni 2003)
- [7] Robert M. Gray, Lee D. Davison, *An Introduction to Statistical Signal Processing* (Cambridge University Press, 2004)
- [8] Babak Hassibi, Bertrand M. Hochwald *How Much Training is Needed in Multiple-Antenna Wireless Links?* (IEEE Transactions on Information Theory, Vol. 49, No. 4, S. 951-963, April 2003)
- [9] Babak Hassibi, Thomas L. Marzetta, *Multiple-Antennas and Isotropically Random Unitary Inputs: The Received Signal Density in Closed Form* (IEEE Transactions on Information Theory, Vol. 48, No. 6, S. 1473-1484, Juni 2002)
- [10] Thomas L. Marzetta, Bertrand M. Hochwald, *Capacity of a Mobile Multiple-Antenna Communication Link in Rayleigh Flat Fading* (IEEE Transactions on Information Theory, Vol. 45, No. 1, S. 139-157, Januar 1999)
- [11] Rudolf Mathar, *Informationstheorie* (Teubner, 1996)
- [12] Rudolf Mathar, *Informationstheorie II* (Vorlesung RWTH Aachen, WS 2002/03)
- [13] Muriel Médard, *The Effect upon Channel Capacity in Wireless Communications of Perfect and Imperfect Knowledge of the Channel* (IEEE Transactions on Information Theory, Vol. 46, No. 3, S. 933-946, Mai 2000)
- [14] R Development Core Team, *R: A language and environment for statistical computing* (R Foundation for Statistical Computing, <http://www.R-project.org/>)
- [15] Burkhard Rauhut, *Stochastik I* (Vorlesung RWTH Aachen, WS 1998/99)
- [16] Burkhard Rauhut, *Stochastik II* (Vorlesung RWTH Aachen, SS 1999)
- [17] Roman Rosipal, *Relation Analysis in Stochastic Systems* (České vysoké učení technické v Praze, 1993)
- [18] Norbert Schmitz, *Vorlesungen über Wahrscheinlichkeitstheorie* (Teubner, 1996)
- [19] Bruno J. Schnekenburger, *A Modified Extended Kalman Filter as a Parameter Estimator for Linear Discrete-Time System* (New Jersey Institute of Technology, Juni 1988)
- [20] Emre Telatar, *Capacity of Multi-antenna Gaussian Channels* (European Transactions on Telecommunications, Vol. 10, No. 6, S. 585-595, November/Dezember 1999)

Ich versichere, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe

Aachen, den 3. Februar 2005