

**Multiple Input Multiple
Output (MIMO) - Drahtlose
Nachrichtenübertragung
hoher Bitrate und Qualität
mit Mehrfachantennen**

von

Prof. Dr.-Ing. Joachim Speidel

Universität Stuttgart

Institut für Nachrichtenübertragung

Inhaltsangabe

Drahtlose Übertragungssysteme mit mehreren Antennen auf der Sende- und Empfangsseite haben in den letzten Jahren großes Interesse in Forschung und Entwicklung erzielt. Solche „Multiple-Input-Multiple-Output (MIMO)“-Systeme erlauben bei gleicher Bandbreite eine höhere Bitrate und zeigen eine geringere Anfälligkeit gegenüber Störungen auf dem Übertragungsweg. Daher sind die Verfahren für künftige Erweiterungen von UMTS und WLAN, besonders aber für neue, breitbandige Mobilkommunikationssysteme der 4. Generation vorgesehen. Dieser Beitrag gibt eine Übersicht über die wesentlichen Prinzipien und zeigt deren Leistungsfähigkeit.

Stichwörter:

MIMO, UMTS, WLAN, drahtlose Kommunikation, Mobilkommunikation, Mehrfachantennen.

Inhaltsverzeichnis

1. Einleitung
2. Das MIMO-Übertragungssystem
 - 2.1 Das Prinzip
 - 2.2 Der äquivalente MIMO-Tiefpasskanal
 - 2.3 Einige MIMO-Kanalmodelle
 - 2.3.1 Zeitvariantes Transversalfilter
 - 2.3.2 Statistische MIMO-Kanalmodelle
 - 2.3.2.1 IID-Kanalmodell
 - 2.3.2.2 MIMO-Kanal mit räumlicher Korrelation
 - 2.3.2.3 Kanalmodell mit exponentieller Korrelation
 - 2.3.3 Rauschen bei MIMO-Kanälen
3. Empfänger für MIMO-Systeme
 - 3.1 Empfänger nach dem Prinzip „Zero Forcing“
 - 3.1.1 Gleiche Antennenzahl $M = N$
 - 3.1.2 Ungleiche Antennenzahl $M < N$
 - 3.1.3 Ungleiche Antennenzahl $M > N$
 - 3.2 Empfänger nach dem Prinzip des „minimalen mittleren quadratischen Fehlers“
 - 3.3 Empfänger nach dem Prinzip des „Maximum Likelihood“
4. Kanalkapazität von MIMO-Systemen
 - 4.1 Einführung
 - 4.2 Instantane MIMO-Kapazität
 - 4.2.1 Voraussetzungen
 - 4.2.2 Zerlegung des MIMO-Systems in Eigenmoden
 - 4.2.3 Kanalkapazität
 - 4.3 Maximierung der instantanen MIMO-Kanalkapazität
 - 4.3.1 Verteilung der Sendeleistung bei bekanntem Kanal („Water Filling“)
 - 4.3.2 Ermittlung des optimalen Sendefilters
 - 4.4 Kapazität des statistischen MIMO-Kanals
 - 4.4.1 Ergodische Kanalkapazität
 - 4.4.2 Ausfallkapazität
5. Raum-Zeit-Codierung
 - 5.1. Das Prinzip
 - 5.2. Raum-Zeit-Codierung nach Alamouti
 - 5.3. Orthogonale Raum-Zeit-Block-Codes
 - 5.4. Raum-Zeit-Trellis-Codierung
6. Räumlicher Multiplex und Demultiplex
 - 6.1 Räumlicher Multiplex ohne zeitliche Codierung
 - 6.2 Vertikale Codierung (V-BLAST)
 - 6.3 Horizontale Codierung (H-BLAST)

-
7. Berechnung der Symbolfehlerwahrscheinlichkeit
 8. Zero-Forcing-Empfänger mit schrittweiser Interferenzauslöschung
 9. Anhang
 - 9.1 Wichtige Beziehungen der linearen Algebra
 - 9.2 Ermittlung der ZF-Empfängermatrix
 - 9.3 Herleitung der MMSE-Empfängermatrix \mathbf{W}
 - 9.4 Singulärwertzerlegung einer Matrix
 - 9.5 Berechnung der MIMO-Kanalkapazität
 - 9.5.1 Herleitung von $\mathcal{C} = \text{ld} \left| \mathbf{I}_N + \mathbf{H}\mathbf{H}^H \frac{E_S}{\sigma_n^2} \right|$
 - 9.5.2 Herleitung von $\mathcal{C} = \text{ld} \left| \mathbf{I}_M + \mathbf{H}^H \mathbf{H} \frac{E_S}{\sigma_n^2} \right|$
 10. Literaturverzeichnis
 11. Verfasserportrait

1. Einleitung

Mit der Einführung von GSM (Global System for Mobile Communications) im Jahr 1992 begann für die Mobilkommunikation eine neue Zeit. Die Nachfrage nach mobilen Telefonanschlüssen hat auch sehr optimistische Erwartungen weit übertroffen. Schon früh wurde erkannt, dass die angebotenen Bitraten für die Sprachenübertragung mehr als ausreichend sind, für die beginnende Nachfrage nach breitbandigeren Multimedia-Anwendungen aber eine Grenze darstellen. Mit EDGE (Enhanced Data for GSM Evolution) und GPRS (Generalized Packet Radio System) wurde daher die Bitrate von GSM erweitert. Parallel dazu begann die Entwicklung des breitbandigen Mobilkommunikationssystems UMTS (Universal Mobile Telecommunications System), auch als IMT 2000 (International Mobile Telecommunications) bekannt, das um 2003 mit Japan beginnend in den Weltmarkt eingeführt wurde. Neben diesen zellularen Systemen wurde ab dem Jahr 2000 mit WLAN (Wireless Local Area Network) nach den Standards IEEE 802.11 b bis g ein weiteres, deutlich breitbandigeres System eingeführt, das Bitraten bis etwa 50 Mbit/s erlaubt. Sein Einsatzgebiet ist hauptsächlich innerhalb von privaten und öffentlichen Gebäuden mit Reichweiten von etwa 30 bis 100 Metern.

Da die verfügbare Bandbreite in Funkkanälen begrenzt und teuer ist, wird die Erhöhung der Bitrate hauptsächlich durch Codierungs- und Modulationsverfahren erreicht, deren spektrale Effizienz (bit/s pro Hz Bandbreite) groß ist. Hierzu gehören die Quadratur-Amplitudenmodulation (QAM) in Verbindung mit leistungsfähigen (Turbo-) Codes. Mehr und mehr setzt sich in der Mobilkommunikation auch das Mehrträgerverfahren OFDM (Orthogonal Frequency Division Multiplex) durch, das beim digitalen Ton- und Fernseh Rundfunk (Digital Multimedia Broadcast DMB, Digital Video Broadcast-Terrestrial DVB-T) schon erfolgreich verwendet wird. OFDM erhöht zwar die spektrale Effizienz nicht, bietet aber vergleichsweise einfache Möglichkeiten der Entzerrung oder Schätzung von breitbandigen Funkkanälen beim Empfänger.

Will man die Bitrate bei gleicher Bandbreite weiter deutlich vergrößern, dann stößt eine Erhöhung der Modulationsstufen der QAM auf Grenzen, da die Sendeleistung aus Gründen der elektromagnetischen Verträglichkeit der Umwelt nicht wie erforderlich gesteigert werden kann.

Anfang der 90er Jahre tauchten die ersten Vorschläge auf, mehrere Sende- und Empfangsantennen zur Erhöhung der Kanalkapazität (in bit pro Kanalnutzung) einzusetzen [1]-[4]. Durch weitere Forschungsarbeiten in den Bell Labs [5]-[8] wurde ein kreativer Sturm in Forschung und Entwicklung ausgelöst, der bis heute andauert.

Es leuchtet zunächst unmittelbar ein, dass durch eine Hinzunahme von weiteren Sende- und Empfangsantennen die Bitrate gesteigert werden kann, weil jedes Paar aus Sender und Empfänger einen zusätzlichen Kanal bildet. Da die zusätzlichen Signale jedoch im selben Frequenz- und Zeitbereich übertragen werden, beeinflussen sich die Kanäle gegenseitig. Es wird sich zeigen, dass diese Korrelation i. A. die MIMO-Kanalkapazität reduziert. Eine spannende Frage ist auch, welchen Einfluss die Zahl M der Sendeantennen und die Zahl N der Empfangsantennen auf die Leistungsfähigkeit des Systems besitzen. Wie bereits erwähnt, ist die maximale Sendeleistung begrenzt und die Leistung pro Sendeantenne geht bei Mehrfachantennen somit um den Faktor M zurück. Ferner muss das Problem gelöst werden, wie man die einzelnen Sendedaten, die sich auf dem

MIMO-Kanal überlagern, durch einen geeigneten Empfänger wieder trennen und trotz Kanalstörungen durch Fading und Rauschen möglichst fehlerfrei detektieren kann.

Mehrfachantennen beim Empfänger wurden schon recht früh in der Funktechnik eingesetzt. Zielsetzung war allerdings nicht die Erhöhung der Bitrate, sondern eine Verbesserung des Empfangs durch Raum-Diversität. Werden z. B. zwei Antennen mit entsprechenden Empfängern eingesetzt, so ist die Wahrscheinlichkeit, dass beide Empfangssignale gleichzeitig stark gestört sind, sehr gering, und eines der beiden Empfangssignale ist häufig brauchbar.

Erweitert man ein drahtloses Funksystem zu einem MIMO-System, dann kommt zur Frequenz- und Zeit- noch eine weitere, höchst interessante Koordinate, die Raumkoordinate hinzu, die neue Möglichkeiten des Systementwurfs bietet.

Dieser Beitrag ist wie folgt gegliedert: In Kapitel 2 werden der MIMO-Übertragungskanal und verschiedene Kanalmodelle, die eine geeignete mathematische Beschreibung erlauben, vorgestellt. Kapitel 3 behandelt MIMO-Empfänger. In Kapitel 4 werden die MIMO-Kanalkapazität und Möglichkeiten zu deren Maximierung erörtert. Kapitel 5 behandelt die Codierung in Raum und Zeit, Kapitel 6 stellt bevorzugte räumliche Multiplexverfahren wie z. B. V-BLAST vor. Kapitel 7 befasst sich mit der quantitativen Beschreibung der Güte der digitalen Übertragung über MIMO-Kanäle, und die Symbolfehlerwahrscheinlichkeit wird berechnet. Kapitel 8 stellt abschließend einige verbesserte MIMO-Empfänger aus der jüngsten Forschung vor. Da für das tiefere Verständnis von MIMO-Systemen solide Kenntnisse der linearen Algebra erforderlich sind, werden einige mathematische Grundlagen in den Anhängen des Kapitels 9 zusammengestellt.

Mein Dank gilt Frau Dorothee Kritzer und Frau Agnes Schön-Abiry für das sorgfältige Erstellen des Schriftsatzes und der Zeichnungen. Herrn Nabil Muhammad danke ich für das sorgfältige Korrekturlesen des Manuskripts.

2. Das MIMO-Übertragungssystem

2.1 Das Prinzip

Bild 2.1 zeigt das Blockschaltbild eines MIMO-Übertragungssystems. Die codierte Bitfolge $b(n)$ wird von einem Seriell-Parallel-Wandler (S/P-Wandler) auf M parallele Zweige des Senders aufgeteilt, die ähnlich aufgebaut sind. Greifen wir Zweig v ($v = 1, \dots, M$) heraus. Gruppen von m_v aufeinander folgender Bits aus dem S/P-Wandler werden durch einen Symbolcodierer (im Weiteren als Mapper bezeichnet) einem komplexen QAM-Symbol $s_v(k)$ zugeordnet, das $L_v = 2^{m_v}$ verschiedene Werte in der komplexen Ebene annehmen kann. k ist die diskrete Zeit. Die Symbolfolge $s_v(k)$ wird durch den Impulsformer auf die Grenzfrequenz f_g bandbegrenzt und in der nachfolgenden Modulatorstufe mit der komplexen Trägerschwingung $e^{j\omega_0 t}$ multipliziert. Nach Bildung des Realteils wird das Signal der v -ten Sendeantenne zugeführt und als $u_v(t)$ abgestrahlt. Die Symbole $s_v(k)$ folgen im Abstand T aufeinander. Die Symbolrate ist also $\nu_S = 1/T$.

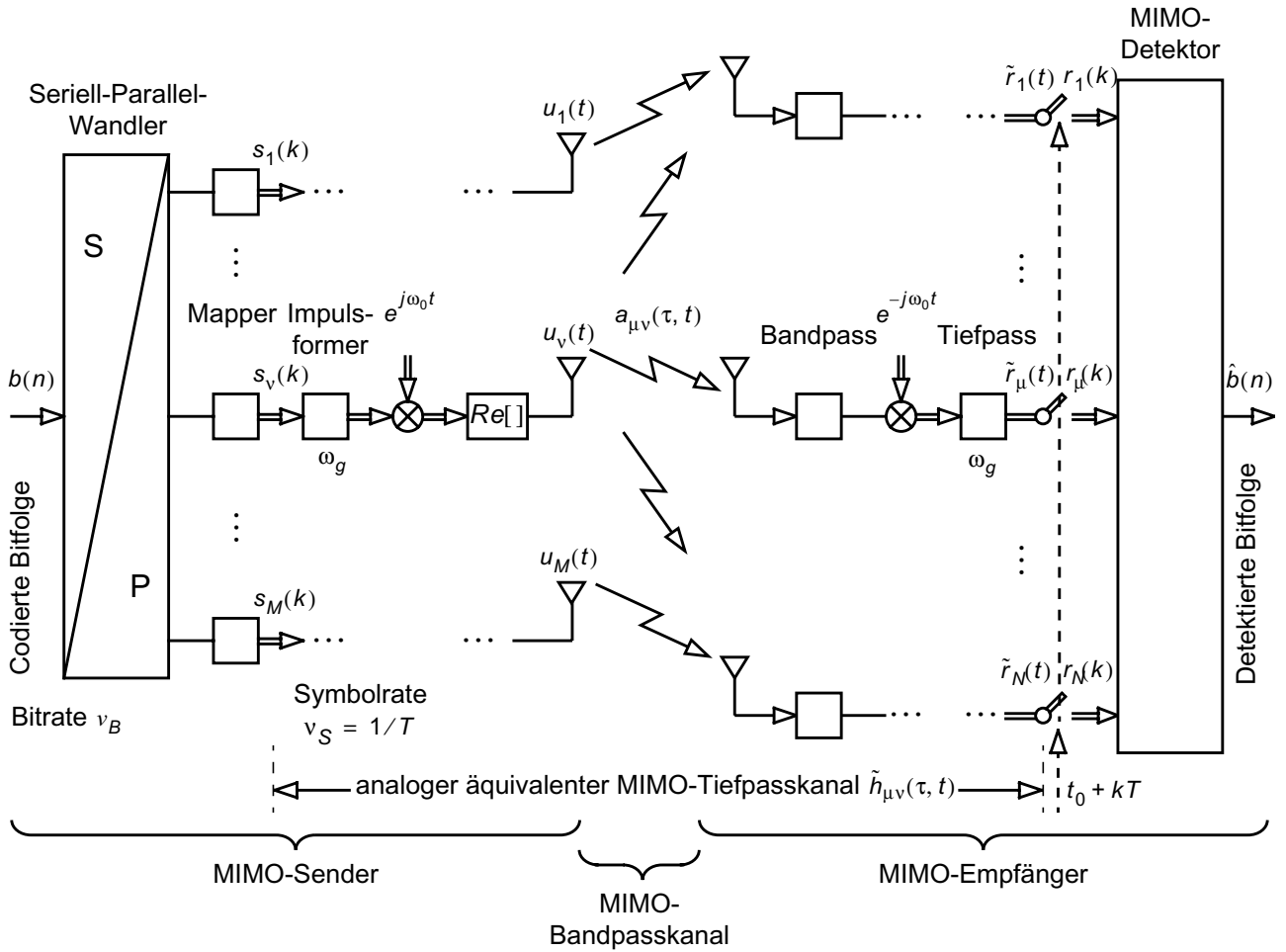


Bild 2.1 Blockschaltbild eines MIMO-Übertragungssystems mit M Sende- und N Empfangsantennen

Diese Funktionen werden in allen M Zweigen parallel ausgeführt. Ist die Trägerfrequenz $f_0 = \frac{\omega_0}{2\pi}$ in allen Zweigen gleich wie in Bild 1 dargestellt liegt ein MIMO-Sender mit Einträgermodulation vor. Bei einer Mehrträgermodulation (z. B. OFDM) sind die Trägerfrequenzen verschieden. Die Kombination von MIMO- und OFDM-Verfahren ist für künftige Systeme ebenfalls sehr interessant, soll aber in diesem Artikel nicht vertieft werden, um den Umfang zu begrenzen.

Über den MIMO-Kanal mit der zeitvarianten Impulsantwort $a_{\mu\nu}(\tau, t)$ wirkt das Sendesignal $u_\nu(t)$ auf alle N Empfangsantennen ($\mu = 1, \dots, N$). Der Empfänger besteht im Prinzip aus N parallelen Zweigen. Das Empfangssignal im Zweig μ durchläuft zunächst einen Bandpass mit der Bandbreite $2f_g$, der im Wesentlichen eine Filterung des Rauschanteils im Empfangssignal vornehmen soll. Durch Multiplikation mit dem komplexen Träger $e^{-j\omega_0 t}$ wird das Signal kohärent demoduliert und anschließend im Tiefpass mit Grenzfrequenz f_g gefiltert. Beim Empfänger ist man nicht am vollständigen Kurvenverlauf des Tiefpass-Ausgangssignals $\tilde{r}_\mu(t)$ interessiert, sondern nur an den Abtastwerten

$$r_\mu(k) = \tilde{r}_\mu(t_0 + kT) \quad , \quad k \in \mathbb{Z} \quad (2.1)$$

im Sende-Zeitraster kT , in dem die gesendeten Symbole $s_\nu(k)$ ($\nu = 1, \dots, M$) erwartet werden. Die

Laufzeit t_0 und das Abtastraster kT müssen durch eine in Bild 2.1 nicht dargestellte Taktrückgewinnung aus dem Empfangssignal gewonnen werden. Ebenso muss eine nicht näher gekennzeichnete Trägerrückgewinnung die exakte Trägerfrequenz f_0 und die geeignete Trägerphase ermitteln. Alle N Empfangszweige sind ähnlich aufgebaut. Aus den zeitdiskreten Signalen $r_\mu(k)$ ($\mu = 1, \dots, N$) ermittelt ein Detektor die Bitfolge $\hat{b}(n)$. Sie enthält lediglich Schätzwerte der gesendeten Folge $b(n)$, da durch Störungen auf dem Übertragungskanal Verfälschungen aufgetreten sein können. Oft ermittelt man auch zunächst Schätzwerte $\hat{s}_v(k)$ ($v = 1, \dots, M$) aus den Abtastwerten $r_\mu(k)$ und in einem zweiten Schritt durch einen Symboldecodierer (Demapper) die Bitfolge $\hat{b}(n)$.

Bild 2.2 (a) zeigt die häufig verwendete Antennenkonfiguration, bei der M Antennenelemente längs einer Linie im gleichen Abstand d angeordnet sind. Man spricht daher auch von einer linearen, äquidistanten Antennenanordnung (linear uniform antenna array). Aus Untersuchungen des abgestrahlten elektromagnetischen Feldes weiß man, dass die emittierten Signale der einzelnen Elemente sich gegenseitig beeinflussen, d. h. korreliert sind, wenn näherungsweise $d \leq \lambda_0/2$ ist. Dabei ist $\lambda_0 = c/f_0$, wobei $f_0 = \frac{\omega_0}{2\pi}$ die Trägerfrequenz in Bild 2.1 und c die Lichtgeschwindigkeit bezeichnen. Unter dieser Bedingung kann man z. B. das abgestrahlte Feld in eine Richtung lenken (Strahlformung, beam forming), in die gesendet werden soll. Bringt man die Antennenelemente weiter auseinander, dann nimmt die Korrelation ab. Für $d > 3\lambda_0$ sind die einzelnen abgestrahlten Signale nahezu nicht mehr korreliert. Diesen Fall kann man für das Senden derselben Daten über mehrere Elemente einsetzen und damit durch Sende-Diversität die Qualität der Funkübertragung erhöhen. Bild 2.2 (b) zeigt das Beispiel einer linearen, nicht äquidistanten Antennenanordnung, die aus zwei Gruppen besteht. Wählt man $d_1 \leq \lambda_0/2$ und $d_2 > 3\lambda_0$, dann kann man Strahlformung und Diversität kombinieren, ein Verfahren, das auch in der Standardisierung [9] für die UMTS-Weiterentwicklung diskutiert wird. Die Antennen in Bild 2.2 können auch als Empfänger betrieben werden, da Sende- und Empfangsmodus weitgehend symmetrisch sind. Wenn nicht anders vermerkt, soll im Weiteren eine lineare, äquidistante Antennenanordnung nach Bild 2.2 (a) mit M Elementen beim Sender und N Elementen beim Empfänger zu Grunde gelegt werden.

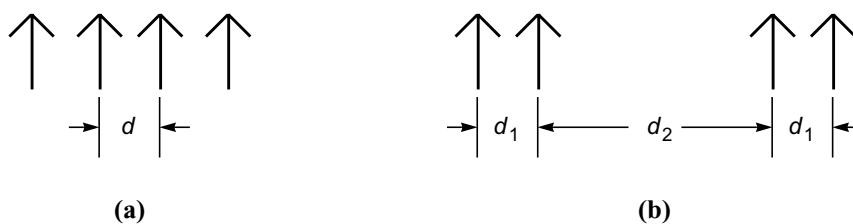


Bild 2.2 Lineare Antennenanordnungen. (a) äquidistante Elemente, (b) Elemente mit unterschiedlichem Abstand

2.2 Der äquivalente MIMO-Tiefpasskanal

Um die weitere Beschreibung des MIMO-Übertragungssystems zu vereinfachen, wird in Bild 2.1 zwischen dem Eingang des Impulsformers im Zweig v des Senders und dem Ausgang des Tiefpasses im Zweig μ des Empfängers eine Ersatzschaltung, der sog. analoge äquivalente Tiefpasskanal mit der Impulsantwort $\tilde{h}_{\mu v}(\tau, t)$, gebildet. $\tilde{h}_{\mu v}(\tau, t)$ enthält die Funktion der Modulation und Demodulation sowie die zeitvariante Impulsantwort $a_{\mu v}(\tau, t)$ des Funkkanals. Während $a_{\mu v}(\tau, t)$ immer reell ist, liefert der Einfluss der Modulation und Demodulation ein i. A. komplexes $\tilde{h}_{\mu v}(\tau, t)$. Die äquivalenten Tiefpasskanäle, die für alle Kombinationen $v = 1, \dots, M$ und $\mu = 1, \dots, N$ auftreten, bilden den analogen äquivalenten MIMO-Tiefpasskanal. Er ist in Bild 2.3 gestrichelt eingetragen. Die gekennzeichneten Antennen haben nur symbolische Bedeutung, denn sie enthalten noch Filterung und Modulation bzw. Demodulation.

Die Signale $\tilde{r}_\mu(t)$ im Empfänger werden wie bereits erwähnt zu den Zeiten $t = t_0 + kT$ abgetastet, und man erhält die zeitdiskreten Signale $r_\mu(k)$; $\mu = 1, \dots, N$; $k \in \mathbb{Z}$. Durch eine weitere Abstraktion kann man nun den in Bild 2.3 gekennzeichneten zeitdiskreten äquivalenten MIMO-Tiefpasskanal bilden, der Grundlage für die weiteren Betrachtungen sein wird. Er wird durch die zeitdiskreten und zeitvarianten Impulsantworten $h_{\mu v}(m, k)$; $\mu = 1, \dots, N$; $v = 1, \dots, M$; $m, k \in \mathbb{Z}$ festgelegt. Dabei kennzeichnet k die „Zeitvarianz“ und m die „Laufzeit“.

Nehmen wir an, dass nur der Eingang v des Senders in Bild 2.3 aktiv ist, und alle übrigen Eingangssignale $s_j(k) = 0$ ($j \neq v$; $j = 1, \dots, M$) sind, dann erhält man im Zweig μ des Empfängers

$$r_\mu(k) = s_v(k) * h_{\mu v}(m, k) + n_\mu(k) \quad \mu = 1, \dots, N \quad (2.2)$$

Dabei ist die zeitvariante Faltung $*$ definiert als

$$s_v(k) * h_{\mu v}(m, k) = \sum_{m=-\infty}^{\infty} h_{\mu v}(m, k) s_v(k-m) = \sum_{m=-\infty}^{\infty} s_v(m) h_{\mu v}(k-m, k) \quad (2.3)$$

Obigen Übertragungsmodus kann man als Single Input Multiple Output (SIMO) bezeichnen.

$n_\mu(k)$ ist dabei das zeitdiskrete Rauschen im Empfangszweig μ , das im Wesentlichen von der Antenne aufgenommen wird und dem sich weitere Anteile der elektronischen Empfängerschaltung überlagert haben. Die statistische Beschreibung geschieht später.

Sind nun die Sender aller Zweige $v = 1, \dots, M$ aktiv, dann erhält man in Matrixschreibweise

$$\underbrace{\begin{bmatrix} r_1(k) \\ r_2(k) \\ \vdots \\ r_N(k) \end{bmatrix}}_{\mathbf{r}(k)} = \underbrace{\begin{bmatrix} h_{11}(m, k) & h_{12}(m, k) & \dots & h_{1M}(m, k) \\ h_{21}(m, k) & h_{22}(m, k) & \dots & h_{2M}(m, k) \\ \dots & \dots & \dots & \dots \\ h_{N1}(m, k) & h_{N2}(m, k) & \dots & h_{NM}(m, k) \end{bmatrix}}_{\mathbf{H}(m, k)} * \underbrace{\begin{bmatrix} s_1(k) \\ s_2(k) \\ \vdots \\ s_M(k) \end{bmatrix}}_{\mathbf{s}(k)} + \underbrace{\begin{bmatrix} n_1(k) \\ n_2(k) \\ \vdots \\ n_N(k) \end{bmatrix}}_{\mathbf{n}(k)} \quad (2.4)$$

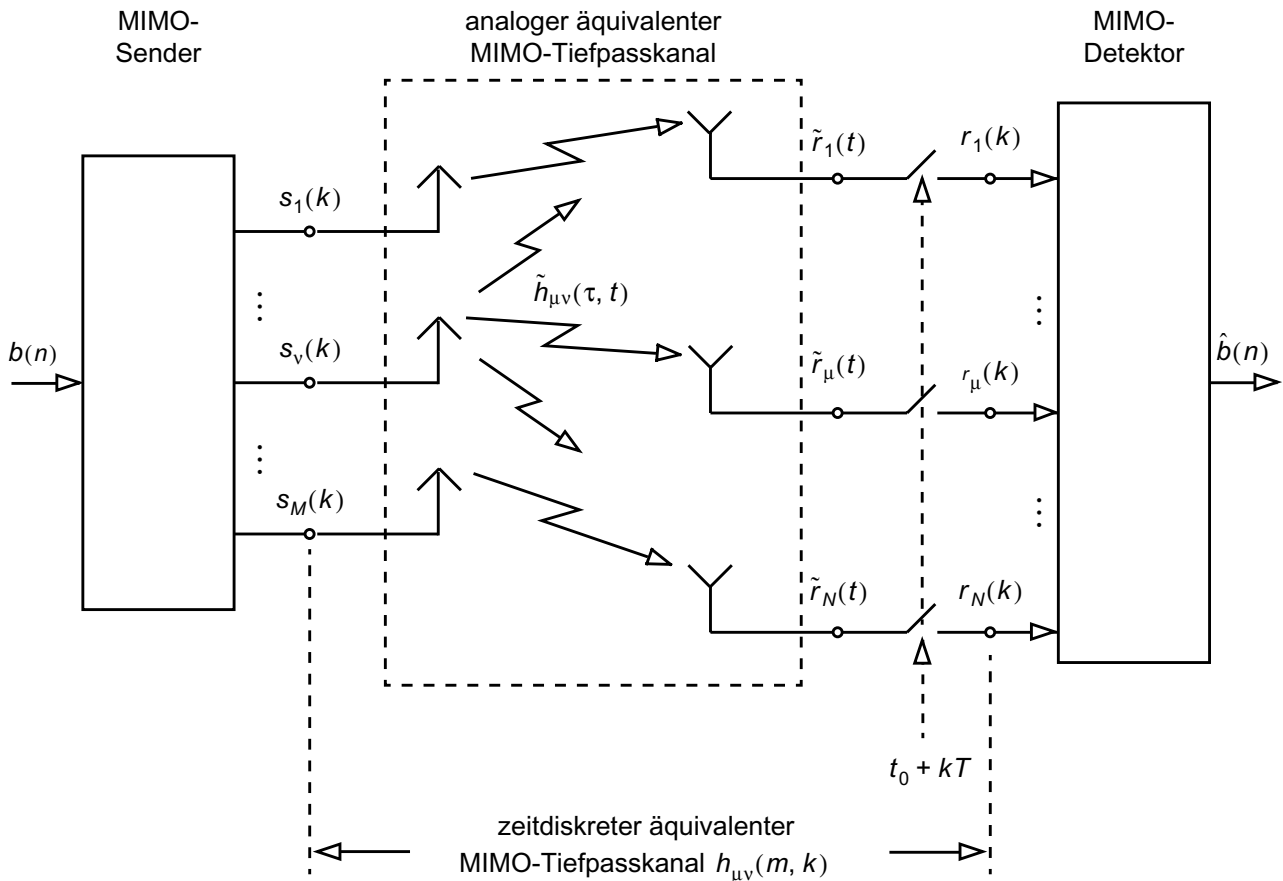


Bild 2.3 Äquivalentes Tiefpassmodell eines MIMO-Systems

$$\mathbf{r}(k) = \mathbf{H}(m, k) * \mathbf{s}(k) + \mathbf{n}(k) \quad (2.5)$$

Dabei ist $\mathbf{r}(k)$ der Empfangsvektor, $\mathbf{H}(m, k)$ die $N \times M$ Matrix des zeitdiskreten, zeitvarianten äquivalenten MIMO-Tiefpasskanals mit den Elementen $h_{\mu\nu}(m, k)$, $\mathbf{s}(k)$ der Vektor der Sendesignale und $\mathbf{n}(k)$ der Rauschvektor. $*$ ist nach (2.3) definiert.

2.3 Einige MIMO-Kanalmodelle

2.3.1 Zeitvariantes Transversalfilter

Realistische Kanäle sind kausal. Durch die Filter im Zweig ν und μ ist die zeitvariante Impulsantwort $h_{\mu\nu}(m, k)$ bandbegrenzt, und ihre Beträge nehmen mit m rasch ab. Daher gilt die Näherung

$$|h_{\mu\nu}(m, k)| = 0 \quad \text{für } m < 0 \text{ und } m > K_{\mu\nu} \quad (2.6)$$

Daraus lässt sich für die Verbindung von Zweig ν nach μ das Ersatzschaltbild in Form eines zeitvarianten Transversalfilters (FIR-Filter, Finite Impulse Response) in Bild 2.4 ableiten. Die Verzögerungselemente arbeiten im Symbolraster T . Eine Verfeinerung des Modells erhält man durch

Überabtastung mit dem Raster $T' < T$. Dies führt auf mehr Filterkoeffizienten, ändert die prinzipiellen Zusammenhänge aber nicht.

Durch Fourier-Transformation von $h_{\mu\nu}(m, k)$ bezüglich m erhält man die zeitvariante Übertragungsfunktion $\tilde{H}_{\mu\nu}(e^{j\omega T}, k)$ zwischen ν und μ . Ist der Verlauf von $|\tilde{H}_{\mu\nu}|$ als Funktion vom ω innerhalb der primitiven Periode $|\omega| \leq \pi/T$ (bzw. π/T') zerklüftet, dann spricht man von einem frequenzselektiven Kanal. Er tritt z. B. auf, wenn zwischen Sendeantenne ν und Empfangsantenne μ eine Mehrwegeausbreitung mit wenigen Pfaden unterschiedlicher Laufzeiten auftritt. Ist dagegen $|\tilde{H}_{\mu\nu}|$ näherungsweise konstant (flach) bezüglich ω , dann liegt ein nicht-frequenzselektiver Kanal vor (frequency flat). Das Kanalmodell in Bild 2.4 enthält dann lediglich den Koeffizienten $h_{\mu\nu}(0, k)$. Alle übrigen sind Null.

$$h_{\mu\nu}(m, k) = h_{\mu\nu}(k) \cdot \delta(m) = \begin{cases} h_{\mu\nu}(k) \neq 0 & , m = 0 \\ 0 & , m = \pm 1, \pm 2, \dots \end{cases} \quad (2.7)$$

mit dem Einheitsimpuls (Kroneckersymbol)

$$\delta(m) = \begin{cases} 1 & , m = 0 \\ 0 & , m = \pm 1, \pm 2, \dots \end{cases} \quad (2.8)$$

In der Informationstheorie nennt man einen solchen Kanal auch gedächtnislos oder speicherfrei. $h_{\mu\nu}(k)$ ist zeitvariant, aber nicht-frequenzselektiv. Gelten die Eigenschaften frequenzselektiv bzw. nicht-frequenzselektiv für alle $h_{\mu\nu}(m, k)$ ($\nu = 1, \dots, M$; $\mu = 1, \dots, N$; alle k), dann nennt man den MIMO-Kanal (im strengen Sinne) frequenzselektiv bzw. nicht-frequenzselektiv.

Aus (2.2) und (2.3) folgt mit (2.7) für den nicht-frequenzselektiven SIMO-Kanal

$$r_{\mu}(k) = s_{\nu}(k) \cdot h_{\mu\nu}(k) + n_{\mu}(k) \quad \mu = 1, \dots, N \quad (2.9)$$

da die Faltung in die Multiplikation übergeht.

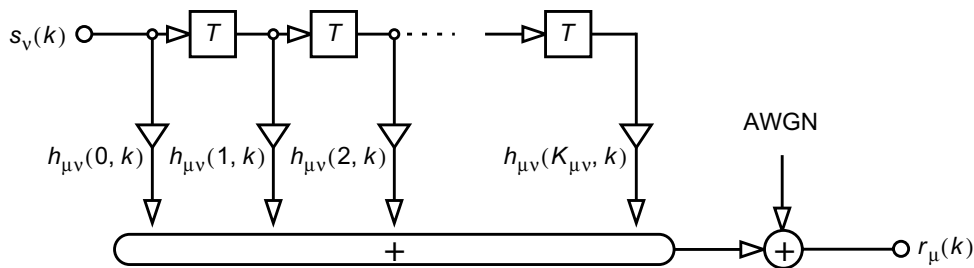


Bild 2.4 Zeitdiskretes Modell des zeitvarianten äquivalenten MIMO-Tiefpasskanals zwischen dem Zweig ν des Senders und μ des Empfängers

Gilt (2.7) für alle Impulsantworten, dann erhält man aus (2.4) und (2.5) für den MIMO-Kanal

$$\mathbf{H}(m, k) = \mathbf{H}(k) = (h_{\mu\nu}(k)) \quad (2.10)$$

$$\mathbf{r}(k) = \mathbf{H}(k) \cdot \mathbf{s}(k) + \mathbf{n}(k) \quad (2.11)$$

Das durch (2.11) beschriebene MIMO-System ist nicht-frequenzselektiv, jedoch nach wie vor zeitvariant. Dies wird leicht durch folgendes Gedankenexperiment deutlich. Nehmen wir an, ein Gleichsignal $s(k) = c_1 = \text{const.}$ werde gesendet, und Rauschen sei nicht vorhanden, d. h. $\mathbf{n}(k) = 0$, dann erwartet man bei einem zeitinvarianten System ebenfalls ein konstantes Empfangssignal. Der zeitvariante MIMO-Kanal $\mathbf{H}(k)$ lässt $\mathbf{r}(k)$ allerdings zeitlich schwanken. Man spricht von Schwund oder Fading.

Bildet man die Fouriertransformierte von $h_{\mu\nu}(m, k)$ bezüglich k , so erhält man die dopplervariante Impulsantwort $\tilde{H}(m, e^{j\omega_D T})$. Dabei ist ω_D die entsprechende Frequenzvariable. Ändert sich $h_{\mu\nu}(m, k)$ schnell mit k , dann ist die Grenzfrequenz von \tilde{H} groß, und man spricht von schnellem Schwund (fast fading), andernfalls von langsamem Schwund (slow fading).

2.3.2 Statistische MIMO-Kanalmodelle

Hier wie auch im Weiteren sollen die Betrachtungen auf nicht-frequenzselektive Kanäle beschränkt werden, d. h. im Weiteren ist stets

$$h_{\mu\nu}(m, k) = h_{\mu\nu}(k) \quad \mu = 1, \dots, N ; \nu = 1, \dots, M \quad (2.12)$$

Ähnlich wie bei Übertragungssystemen mit jeweils nur einer Antenne (Single Input Single Output, SISO), treten auch bei MIMO-Kanälen Reflexionen, Beugungen und Mehrwegeausbreitungen auf, die so vielfältig sind, dass eine statistische Beschreibung der deterministischen vorzuziehen ist. Tritt bei der Mehrwegeausbreitung eine Vielzahl von statistisch unabhängigen Pfaden auf, und ist kein dominanter Pfad (z. B. Sichtverbindung, Line Of Sight, LOS) oder Reflektor vorhanden, dann kann man mit dem zentralen Grenzwertsatz

$$h_{\mu\nu}(k) = \text{Re} [h_{\mu\nu}(k)] + j \text{Im} [h_{\mu\nu}(k)] \quad (2.13)$$

als komplexwertigen Gaußprozess modellieren. Real- und Imaginärteil sind mittelwertfrei, statistisch unabhängig und besitzen dieselbe Varianz. Alle $|h_{\mu\nu}(k)|$ besitzen somit eine Rayleigh-Wahrscheinlichkeitsdichtefunktion, weshalb man häufig von Rayleigh-Fading spricht.

2.3.2.1 IID-Kanalmodell

Dieses Modell geht davon aus, dass alle $h_{\mu\nu}(k)$ statistisch unabhängig sind und identische Wahrscheinlichkeitsdichtefunktionen besitzen (independent, identical distributed, IID). Es gilt für alle $\mu = 1, \dots, N ; \nu = 1, \dots, M$:

$$h_{\mu\nu} \text{ komplex gaußisch} \quad (2.14)$$

$$E[h_{\mu\nu}] = 0 \quad (2.15)$$

$$E[h_{\mu\nu} \cdot h_{\lambda\rho}^*] = \begin{cases} \sigma^2 = 1 & \text{für } \mu = \lambda, \nu = \rho \\ 0 & \text{für } \mu \neq \lambda, \nu \neq \rho \end{cases} \quad (2.16)$$

(2.16) besagt auch, dass die zeitvarianten Impulsantworten $h_{\mu\nu}(k)$ räumlich unkorreliert, d. h. „weiß“ sind. Die Matrix mit Eigenschaften (2.14)-(2.16) soll als

$$\mathbf{H}(k) = \mathbf{H}_W(k) \quad (2.17)$$

bezeichnet werden.

Befindet sich zwischen den Sende- und Empfangsantennen ein dominanter Pfad, z. B. eine Sichtverbindung, dann kann man den $|h_{\mu\nu}(k)|$ näherungsweise eine Rice-Wahrscheinlichkeitsdichtefunktion zuordnen mit $E[h_{\mu\nu}] \neq 0$.

Zur Ermittlung der Korrelationsmatrix \mathbf{R}_{HH} kann man die Kanalmatrix zunächst durch ihre Spaltenvektoren $\mathbf{h}_l(k)$ ($l = 1, \dots, M$) darstellen.

$$\mathbf{H} = (\mathbf{h}_1 \ \mathbf{h}_2 \ \dots \ \mathbf{h}_M) \quad (2.18)$$

und dann einen Vektor der Spaltenvektoren

$$\text{vec } \mathbf{H} = \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_M \end{bmatrix} \quad (2.19)$$

bilden. Die Autokorrelationsmatrix \mathbf{R}_{HH} der Matrix \mathbf{H} lässt sich dann definieren

$$\mathbf{R}_{HH} = E[(\text{vec } \mathbf{H})(\text{vec } \mathbf{H})^H] \quad (2.20)$$

Dabei ist

$$\mathbf{H}^H = (\mathbf{H}^T)^* = (\mathbf{H}^*)^T \quad (2.21)$$

die zu \mathbf{H} Hermitesche Matrix (konjugiert-transponierte Matrix). \mathbf{R}_{HH} nach (2.20) wird oft auch als Kovarianzmatrix bezeichnet, was zutrifft, wenn alle Kanalimpulsantworten $h_{\mu\nu}$ mittelwertfrei sind.

Beispiel: $M = N = 2$

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}$$

$$\mathbf{R}_{HH} = E[(\text{vec } \mathbf{H})(\text{vec } \mathbf{H})^H] = E \left[\begin{pmatrix} h_{11} \\ h_{21} \\ h_{12} \\ h_{22} \end{pmatrix} \cdot (h_{11}^* \ h_{21}^* \ h_{12}^* \ h_{22}^*) \right]$$

$$\mathbf{R}_{HH} = \begin{pmatrix} E[|h_{11}|^2] & E[h_{11} h_{21}^*] & E[h_{11} h_{12}^*] & E[h_{11} h_{22}^*] \\ E[h_{21} h_{11}^*] & E[|h_{21}|^2] & E[h_{21} h_{12}^*] & E[h_{21} h_{22}^*] \\ E[h_{12} h_{11}^*] & E[h_{12} h_{21}^*] & E[|h_{12}|^2] & E[h_{12} h_{22}^*] \\ E[h_{22} h_{11}^*] & E[h_{22} h_{21}^*] & E[h_{22} h_{12}^*] & E[|h_{22}|^2] \end{pmatrix} \quad (2.22)$$

Man kann leicht zeigen, dass \mathbf{R}_{HH} eine positiv-semidefinite Hermitesche Matrix der Dimension $M \cdot N$ ist. Daraus folgt, dass ihre Eigenwerte nicht negativ sind. Für einen räumlich unkorrelierten MIMO-Kanal verschwinden die Elemente außerhalb der Hauptdiagonale in \mathbf{R}_{HH} . Mit (2.16) erhält man

$$\mathbf{R}_{HH} = \mathbf{I}_{MN} \quad (2.23)$$

wobei \mathbf{I}_{MN} die $MN \times MN$ Einheitsmatrix ist. (2.23) gilt insbesondere für $\mathbf{H} = \mathbf{H}_W$ in (2.17).

2.3.2.2 MIMO-Kanal mit räumlicher Korrelation

Die Korrelation der individuellen Kanalimpulsantworten $h_{\mu\nu}(k)$ hängt stark von den Umgebungsbedingungen auf dem Weg zwischen der Sende- und der Empfangsantenne ab. Wie in Bild 2.5 dargestellt, treten besonders im Stadtbereich in der Nähe der Antennen die einflussreichen Hindernisse wie Bäume, Hauswände u. dgl. auf, die Reflexionen und Beugungen der Wellen verursachen und damit zur bereits früher erwähnten Mehrwegeausbreitung führen.

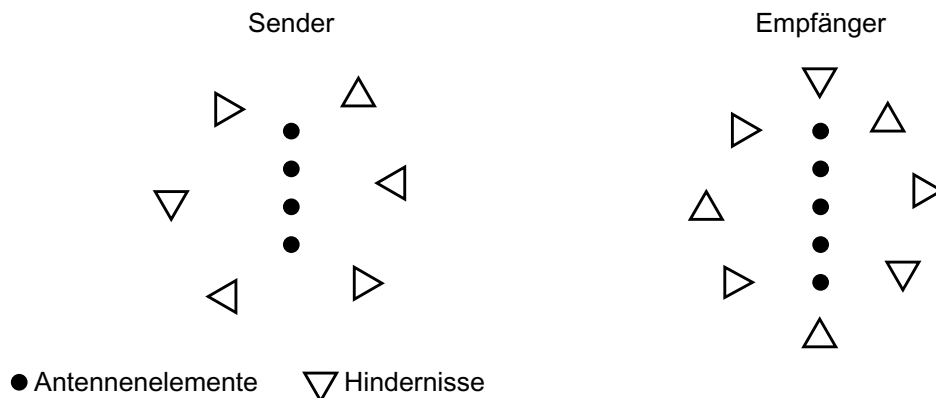


Bild 2.5 MIMO-Verbindung mit Reflexionen in der lokalen Umgebung von Sender und Empfänger

Beeinflussen sich die Wellen gegenseitig, dann gilt die in Abschnitt 2.3.2.1 genannte statistische Unabhängigkeit der Pfade nicht mehr. Um dem Einfluss der Korrelation dicht beim Sender und Empfänger Rechnung zu tragen, ist ein Modell zweckmäßig, das eine senderseitige (\mathbf{R}_{tx}) und eine empfangsseitige Autokorrelationsmatrix (\mathbf{R}_{rx}) vorsieht. Man kann dann das Modell mit $\mathbf{H}_W(k)$ nach (2.17) wie folgt erweitern

$$\mathbf{H}(k) = \mathbf{A}^H \mathbf{H}_W(k) \mathbf{B} \quad (2.24)$$

wobei $\mathbf{R}_{tx} = \mathbf{B}^H \cdot \mathbf{B} \quad (M \times M) \quad (2.25)$

und $\mathbf{R}_{rx} = \mathbf{A}^H \mathbf{A} \quad (N \times N) \quad (2.26)$

Die Elemente von $\mathbf{H}(k)$ sind nun räumlich korreliert. Aus (2.25) und (2.26) erkennt man, dass \mathbf{R}_{tx} und \mathbf{R}_{rx} zeitinvariante Hermitesche Matrizen sind. Die Matrizen \mathbf{A} und \mathbf{B} erhält man also aus einer Zerlegung von \mathbf{R}_{tx} und \mathbf{R}_{rx} . Daher verwendet man oft die Schreibweise $\mathbf{B} = \mathbf{R}_{tx}^{1/2}$ und $\mathbf{A} = \mathbf{R}_{rx}^{1/2}$. Man beachte, dass nur dann, wenn \mathbf{R}_{tx} und \mathbf{R}_{rx} Diagonalmatrizen sind, die Elemente von \mathbf{B} bzw. \mathbf{A} durch die einfache Operation $(\)^{1/2} = \sqrt{(\)}$ aus den Elementen von \mathbf{R}_{tx} bzw. \mathbf{R}_{rx} gebildet werden können.

Man kann zeigen, dass für die Autokorrelationsmatrix \mathbf{R}_{HH} von $\mathbf{H}(k)$ in (2.24) gilt

$$\mathbf{R}_{HH} = \mathbf{R}_{tx}^* \otimes \mathbf{R}_{rx} \quad (2.27)$$

Dabei ist \otimes das Kronecker-Produkt. Das Übertragungssystem mit korreliertem MIMO-Kanal lässt sich dann durch Bild 2.6 darstellen. Wegen (2.27) nennt man diese Anordnung auch das Kronecker-Korrelationsmodell. Im Fall des unkorrelierten MIMO-Kanals ist $\mathbf{R}_{tx} = \mathbf{I}_M$ und $\mathbf{R}_{rx} = \mathbf{I}_N$.

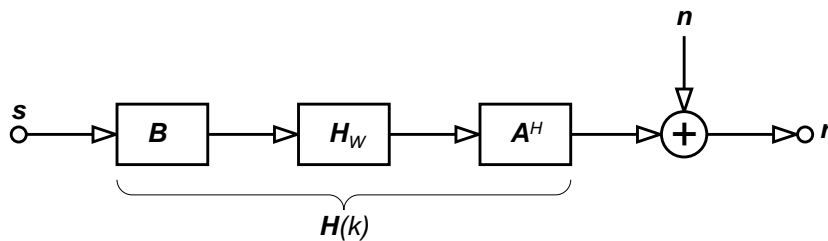


Bild 2.6 Kronecker-Korrelationsmodell des MIMO-Kanals

2.3.2.3 Kanalmodell mit exponentieller Korrelation

Messtechnisch kann man näherungsweise folgendes feststellen: Die Korrelation von zwei Signalpfaden beim Sender ist umso kleiner, je weiter die Antennenelemente des Senders voneinander entfernt sind, in denen die Signalpfade beginnen. Nimmt man näherungsweise eine Exponentialfunktion als Modell, dann erhält man im Umfeld des Senders

$$(\mathbf{R}_{tx})_{p,q} = \rho_{tx}^{|p-q|} \quad p, q = 1, \dots, M \quad (2.28)$$

mit $|\rho_{tx}| < 1 \quad (2.29)$

Entsprechend für das Umfeld des Empfängers

$$(\mathbf{R}_{rx})_{p,q} = \rho_{rx}^{|p-q|} \quad p, q = 1, \dots, N \quad (2.30)$$

$$|\rho_{rx}| < 1 \quad (2.31)$$

2.3.3 Rauschen bei MIMO-Kanälen

Man kann mit guter Näherung annehmen, dass jedem Empfangssignal $r_\mu(k)$ in (2.4) Gaußsches Rauschen $n_\mu(k)$ überlagert ist (Additive White Gaussian Noise, AWGN). Das Rauschen wird einerseits von der Antenne aus dem Funkkanal aufgenommen. Andererseits enthält der Empfänger elektronische Schaltungen, die einen zusätzlichen Rauschbeitrag liefern. Das gesamte Rauschen an der Empfangsantenne μ wird mit $n_\mu(k)$ bezeichnet. Die Wahrscheinlichkeitsdichtefunktion (WDF) $p_n(\mathbf{n})$ des Rauschvektors \mathbf{n} in (2.5) wird nachfolgend schrittweise ermittelt.

- Sei x_μ reelles, mittelwertfreies Gaußsches Rauschen mit der Varianz ρ^2 . Dann gilt für die WDF

$$p(x_\mu) = \frac{1}{\sqrt{2\pi} \rho} e^{-\frac{1}{2} \frac{x_\mu^2}{\rho^2}} \quad (2.32)$$

- Sei $n_\mu = x_\mu + jy_\mu$ komplexes, mittelwertfreies Gaußsches Rauschen, wobei x_μ und y_μ die WDF in (2.32) und gleiche Varianz ρ^2 besitzen und statistisch unabhängig sind. Dann gilt für die WDF von n_μ , d. h. die Verbunddichte aus x_μ und y_μ

$$p_{n_\mu}(n_\mu) = p(x_\mu) p(y_\mu) = \left(\frac{1}{\sqrt{2\pi} \rho} \right)^2 e^{-\frac{1}{2} \frac{x_\mu^2 + y_\mu^2}{\rho^2}} = \frac{1}{\pi 2\rho^2} e^{-\frac{1}{2} \frac{n_\mu^* n_\mu}{\rho^2}} = \frac{1}{\pi \sigma_n^2} e^{-\frac{n_\mu^* n_\mu}{\sigma_n^2}} \quad (2.33)$$

wobei $\sigma_n^2 = 2\rho^2$ die Varianz von n_μ ist.

- Sei $\mathbf{n} = (n_1 \dots n_N)^T$ ein Rauschvektor mit Komponenten n_μ nach (2.33). Die Komponenten n_1, \dots, n_N seien statistisch unabhängig, insbesondere gilt also

$$E[\mathbf{n} \mathbf{n}^H] = \sigma_n^2 \mathbf{I}_N \quad (2.34)$$

Dann kann man die WDF von \mathbf{n} als Verbund-Dichtefunktion aller Komponenten darstellen

$$p_n(\mathbf{n}) = \prod_{\mu=1}^N p_{n_\mu}(n_\mu) = \left(\frac{1}{\pi \sigma_n^2} \right)^N e^{-\frac{\mathbf{n}^H \cdot \mathbf{n}}{\sigma_n^2}} = \left(\frac{1}{\pi \sigma_n^2} \right)^N e^{-\frac{\|\mathbf{n}\|^2}{\sigma_n^2}} \quad (2.35)$$

wobei

$$\mathbf{n}^H \cdot \mathbf{n} = \|\mathbf{n}\|^2$$

3. Empfänger für MIMO-Systeme

Wir gehen vom Blockschaltbild in Bild 2.7 aus. Der Empfangsvektor $\mathbf{r}(k)$ ist in (2.11) berechnet worden. Der MIMO-Kanal sei nicht frequenzselektiv, jedoch zeitvariant durch Fading-Effekte. Daher sind die zeitvarianten Impulsantworten $h_{\mu\nu}(m, k) = h_{\mu\nu}(k)$ und die MIMO-Kanalmatrix $\mathbf{H}(m, k) = \mathbf{H}(k)$ die kennzeichnenden Größen. Wie man aus (2.11) erkennt, wird das Empfangssignal $r_\mu(k)$ der Empfangsantenne μ durch Überlagerung der Sendesignale $s_1(k) \dots s_M(k)$ aller M Sendeantennen, gewichtet mit den Kanalimpulsantworten $h_{\mu\nu}(k)$, gebildet.

Hinzu kommt noch das Empfängerrauschen $n_\mu(k)$. Der Empfänger mit Matrix \mathbf{W} in Bild 2.7 hat die Aufgabe, die Sendesignale $s_1(k)\dots s_M(k)$ mit möglichst geringen Verzerrungen aus den Empfangssignalen $r_1(k)\dots r_N(k)$ zu gewinnen. Hierzu werden in diesem Abschnitt mehrere Empfänger vorgeschlagen und untersucht. Dabei wird angenommen, dass die Kanalmatrix $\mathbf{H}(k)$ beim Empfänger bekannt ist. Dies ist in der Realität zunächst nicht der Fall. Mit einer Kanalschätzeinrichtung, die in Bild 2.7 nicht dargestellt ist, müssen die Impulsantworten $h_{\mu\nu}(k)$ näherungsweise ermittelt und laufend aktualisiert werden. Ein alternatives, sehr leistungsfähiges Verfahren ist auch, das Filter im Sender anzuordnen [10]-[12]. Zur Vereinfachung der Schreibweise werde im Weiteren k weggelassen. Mit (2.11) folgt für das Signal \mathbf{y} am Ausgang des Empfängers in Bild 2.7

$$\mathbf{y} = \mathbf{W} \cdot \mathbf{r} = \mathbf{W}\mathbf{H} \mathbf{s} + \mathbf{W} \cdot \mathbf{n} \quad (3.1)$$

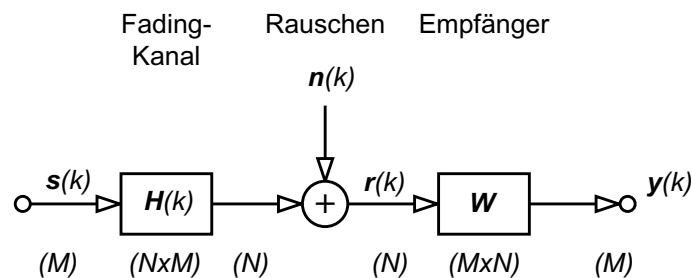


Bild 2.7 MIMO-System mit nicht-frequenzselektivem Fading-Kanal, Gaußischem Rauschen und Empfängermatrix \mathbf{W}

3.1 Empfänger nach dem Prinzip „Zero Forcing“

Um das Prinzip zu zeigen, nehmen wir für einen Augenblick an, das Rauschen sei vernachlässigbar klein. Wie bereits erwähnt, hängt das Signal r_μ im Empfangszweig μ in Bild 2.3 von allen Sendesignalen $s_1 \dots s_M$ ab. Das gilt zunächst auch für die Komponente y_μ des Signalvektors \mathbf{y} in (3.1). y_μ lässt sich also wie folgt darstellen

$$y_\mu = A_\mu \cdot s_\mu + I(s_1, \dots, s_{\mu-1}, s_{\mu+1}, \dots, s_M) \quad (3.2)$$

Dabei ist A_μ eine Konstante. $I(\)$ kann man als Interkanalinterferenz auffassen, die den Nutzanteil $A_\mu s_\mu$ verzerrt. ¹⁾ Das Ziel des Zero-Forcing (ZF)-Empfängers ist, $I(\)$ zu beseitigen.

3.1.1 Gleiche Antennenzahl $M = N$

Wir nehmen ferner an, dass \mathbf{H} vollen Rang besitze, d. h.

$$\text{rang } \mathbf{H} = M = N \quad (3.3)$$

1) Normalerweise tritt in Datenübertragungssystemen immer auch Intersymbolinterferenz auf. Diese ist hier nicht vorhanden, da ein nicht-frequenzselektiver Kanal $\mathbf{H}(m, k) = \mathbf{H}(k)$ vorausgesetzt wurde.

Aus (3.1) folgt bei vernachlässigbarem Rauschen

$$\mathbf{y} = \mathbf{W}\mathbf{H}\mathbf{s} \quad (3.4)$$

Soll die Kanalinterferenz verschwinden, muss gefordert werden

$$\mathbf{y} = \mathbf{s}$$

Aus (3.4) erhält man dann die Matrix \mathbf{W} des ZF-Empfängers

$$\mathbf{W} = \mathbf{H}^{-1} \quad (3.5)$$

Da (3.3) vorausgesetzt wurde, existiert \mathbf{H}^{-1} .

Berücksichtigt man nun das Rauschen, dann folgt mit (3.5) aus (3.1)

$$\mathbf{y} = \mathbf{s} + \mathbf{H}^{-1} \cdot \mathbf{n} \quad (3.6)$$

Das Empfängerausgangssignal \mathbf{y} ist damit nur durch das gefilterte Rauschen $\mathbf{H}^{-1}\mathbf{n}$ gestört, dessen Varianz durch \mathbf{H}^{-1} jedoch vergrößert sein kann. Diesen Nachteil muss man hinnehmen, da bei der Festlegung von \mathbf{W} das Rauschen nicht berücksichtigt wurde.

3.1.2 Ungleiche Antennenzahl $M < N$

Nun werde der Fall betrachtet, dass die Zahl M der Sendeantennen kleiner als die Zahl N der Empfangsantennen ist. Bei Vernachlässigung des Rauschens erhält man hierfür aus (2.11)

$$\mathbf{r} = \mathbf{H} \cdot \mathbf{s} \quad (3.7)$$

Eine realistische Anordnung könnte durch $M = 2$ und $N = 3$ gegeben sein.

$$\begin{aligned} r_1 &= h_{11} s_1 + h_{12} s_2 \\ r_2 &= h_{21} s_1 + h_{22} s_2 \\ r_3 &= h_{31} s_1 + h_{32} s_2 \end{aligned} \quad (3.8)$$

Offensichtlich ist das Gleichungssystem $(N - M)$ -fach überbestimmt, da N Gleichungen für $M < N$ Unbekannte $s_1 \dots s_M$ vorhanden sind. Das Gleichungssystem ist eindeutig lösbar, wenn \mathbf{H} den Rangabfall $N - M$ besitzt. Dies ist gleichbedeutend damit, dass $N - M$ Gleichungen linear abhängig sein müssen und daher gestrichen werden können. Man überlegt leicht, dass sich bei der vorliegenden physikalischen Problemstellung im rauschfreien Fall eine solche lineare Abhängigkeit von selbst einstellt. Kommt auf der rechten Seite von (3.7) und (3.8) noch Rauschen hinzu oder bei einem rein mathematischen Problem muss das nicht der Fall sein. Bis zu $N - M$ Gleichungen könnten zu den übrigen im Widerspruch stehen. Man kann (3.7) dann im Sinne des minimalen quadratischen Fehlers lösen. Dazu definiert man einen Fehler aus linker und rechter Seite des Gleichungssystems

$$\Delta = \mathbf{r} - \mathbf{H} \cdot \mathbf{s} \quad (3.9)$$

Der quadratische Fehler J wird nun bezüglich s minimiert

$$J = \|\Delta\|^2 = \Delta^H \cdot \Delta = (\mathbf{r} - \mathbf{H} \mathbf{s})^H (\mathbf{r} - \mathbf{H} \mathbf{s}) = \min_{\{\mathbf{s} = \hat{\mathbf{s}}\}} \quad (3.10)$$

Aus der notwendigen Bedingung

$$\frac{\partial J}{\partial s_v} = 0 \quad v = 1, \dots, M \quad (3.11)$$

erhält man bereits die Lösung, da J eine quadratische und konvexe Funktion der $s_1 \dots s_M$ ist.

$$\hat{\mathbf{s}} = \mathbf{H}_p \cdot \mathbf{r} \quad (3.12)$$

mit
$$\mathbf{H}_p = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \quad (M \times N) \quad (3.13)$$

Der Beweis ist im Anhang 9.2 geführt.

\mathbf{H}_p heißt Pseudoinverse von \mathbf{H} , in der amerikanischen Literatur auch Moore-Penrose-Inverse.

Damit ist die Matrix \mathbf{W} des Empfängers gefunden:

$$\mathbf{W} = \mathbf{H}_p = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \quad (3.14)$$

Man zeigt leicht, dass

$$\mathbf{H}_p \cdot \mathbf{H} = \mathbf{I}_M \quad (3.15)$$

ist, und das Empfängerausgangssignal folgt aus (3.1) mit (3.14) und (3.15)

$$\hat{\mathbf{s}} = \mathbf{H}_p \cdot \mathbf{H} \mathbf{s} + \mathbf{H}_p \mathbf{n} = \mathbf{s} + \mathbf{H}_p \mathbf{n} \quad (3.16)$$

Auch hier ist festzustellen, dass die Varianz des Rauschanteils $\mathbf{H}_p \cdot \mathbf{n}$ größer sein kann als die Varianz von \mathbf{n} .

Was passiert nun, wenn $M = N$ ist? Dann existiert \mathbf{H}^{-1} wegen der Voraussetzung (3.3). Aus (3.13) folgt

$$\mathbf{I}_p = \mathbf{H}^{-1} (\mathbf{H}^H)^{-1} \mathbf{H}^H = \mathbf{H}^{-1} \mathbf{I}_M = \mathbf{H}^{-1} \quad (3.17)$$

und die Pseudoinverse wird zur inversen Matrix. Damit kann man den Anwendungsbereich von \mathbf{H}_p nach (3.13) auf $M \leq N$ ausdehnen.

3.1.3 Ungleiche Antennenzahl $M > N$

Vernachlässigt man das Rauschen, dann folgt auch hier aus (2.11)

$$\mathbf{r} = \mathbf{H} \mathbf{s} \quad (3.18)$$

und für den realitätsnahen Fall $M = 3$ und $N = 2$

$$\begin{aligned} r_1 &= h_{11} s_1 + h_{12} s_2 + h_{13} s_3 \\ r_2 &= h_{21} s_1 + h_{22} s_2 + h_{23} s_3 \end{aligned} \quad (3.19)$$

(3.18) ist $(M - N)$ -fach unterbestimmt, weil $M - N$ Gleichungen fehlen. Da \mathbf{H} nicht quadratisch ist, existiert \mathbf{H}^{-1} auch nicht. Aus der Theorie der linearen Algebra ist bekannt, dass man im Fall (3.19) eine Variable, z. B. s_3 , beliebig wählen darf (einfach-unendliche Lösungsmannigfaltigkeit). Zunächst erscheint dies als eine Verschwendung von Sendeantennen. Doch bei zeitvarianten Übertragungskanälen kann man daraus Vorteile durch Raum-Zeit-Codierung ziehen, wie in Kapitel 5 dargestellt wird. Dann kann man durch Sender-Diversität die Übertragungsqualität verbessern, da z. B. durch Fading augenblicklich $h_{12}(k) = h_{22}(k) = 0$, jedoch $h_{23}(k) \neq 0$ sein können.

An diesem Beispiel können die Vorzüge eines MIMO-Systems besonders deutlich gemacht werden. Nehmen wir den einfachen Fall $M = N = 2$

$$\begin{aligned} r_1 &= h_{11} s_1 + h_{12} s_2 \\ r_2 &= h_{21} s_1 + h_{22} s_2 \end{aligned} \quad (3.20)$$

Offensichtlich strahlt der Sender zum Zeitpunkt k zwei Symbole $s_1(k)$ und $s_2(k)$ ab, d. h. die Symbolrate ist doppelt so groß wie im Fall $M = 1$. Störungen durch Fading können in einem gewissen Maß beim Empfänger behoben werden. Ist z. B. augenblicklich $h_{11}(k) = 0$, so kann s_1 dennoch ermittelt werden, nämlich aus der 2. Gleichung in (3.20). Dies wäre bei zwei parallel betriebenen, konventionellen Einantennensystemen (Single Input Single Output, SISO) nicht möglich, da hierfür die nicht gekoppelten Übertragungsgleichungen gelten

$$r_1 = h_{11} s_1 \quad \text{und} \quad r_2 = h_{22} s_2$$

Bei $h_{11}(k) = 0$ kann s_1 aus r_1 und r_2 nicht gewonnen werden. Offensichtlich tragen die Impulsantworten h_{12} und h_{21} zur Diversität und damit zur Verbesserung der Empfangsqualität bei.

3.2 Empfänger nach dem Prinzip des „minimalen mittleren quadratischen Fehlers“

Wir betrachten wieder das MIMO-Übertragungssystem in Bild 2.7 und wollen die Empfängermatrix \mathbf{W} nunmehr so wählen, dass der mittlere quadratische Fehler zwischen dem Sendesignal \mathbf{s} und dem Empfänger Ausgangssignal \mathbf{y} minimal wird (Minimum Mean Squared Error, MMSE).

Wir betrachten also

$$J = E [\|\mathbf{s} - \mathbf{y}\|^2] \quad (3.21)$$

Im Gegensatz zum ZF-Empfänger schließt dieser Ansatz damit das Rauschen ein.

Sei

$$\mathbf{R}_{ss} = E [\mathbf{s} \mathbf{s}^H] \quad (M \times M) \quad (3.22)$$

die Autokorrelationsmatrix des Sendesignals und

$$\mathbf{R}_{nn} = E [\mathbf{n} \mathbf{n}^H] = \sigma_n^2 \mathbf{I}_N \quad (N \times N) \quad (3.23)$$

die Autokorrelationsmatrix des räumlich unkorrelierten Rauschens, dann wird im Anhang 9.3 gezeigt, dass die optimale MMSE-Empfängermatrix gegeben ist durch

$$\mathbf{V} = \mathbf{R}_{ss} \mathbf{H}^H (\mathbf{H} \mathbf{R}_{ss} \mathbf{H}^H + \mathbf{R}_{nn})^{-1} \quad (3.24)$$

Wie bereits zu Beginn des Kapitels 3 vorausgesetzt, ist auch für (3.24) die perfekte Kenntnis des Kanals $\mathbf{H}(k)$ erforderlich, um $\mathbf{W} = \mathbf{W}(k)$ zu jedem Zeitpunkt k zu ermitteln.

Häufig kann man davon ausgehen, dass die Sendesignale $s_v(k)$ räumlich nicht korreliert sind. Ist die mittlere Sendeleistung in jedem Zweig v des Senders gleich, d. h.

$$E [|s_v|^2] = E_S \quad , \quad v = 1, \dots, M \quad (3.25)$$

dann wird

$$\mathbf{R}_{ss} = E_S \mathbf{I}_M \quad (3.26)$$

Mit (3.23) und (3.26) folgt dann aus (3.24)

$$\begin{aligned} &= E_S \mathbf{I}_M \mathbf{H}^H (\mathbf{H} E_S \mathbf{I}_M \mathbf{H}^H + \sigma_n^2 \mathbf{I}_N)^{-1} = E_S \mathbf{H}^H \left(E_S \left[\mathbf{H} \mathbf{H}^H + \frac{\sigma_n^2}{E_S} \mathbf{I}_N \right] \right)^{-1} \\ \mathbf{V} &= \mathbf{H}^H (\mathbf{H} \mathbf{H}^H + \alpha \mathbf{I}_N)^{-1} \quad ; \quad \alpha = \sigma_n^2 / E_S \end{aligned} \quad (3.27)$$

Mit dem Matrixinversions-Lemma der linearen Algebra

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{C}^{-1} + \mathbf{DA}^{-1} \mathbf{B})^{-1} \mathbf{DA}^{-1}$$

folgt dann aus (3.27) alternativ

$$\mathbf{V} = (\mathbf{H}^H \cdot \mathbf{H} + \alpha \mathbf{I}_M)^{-1} \mathbf{H}^H \quad (3.28)$$

$\alpha = \sigma_n^2 / E_S$ ist offensichtlich das Verhältnis der mittleren Leistung σ_n^2 des komplexwertigen Gaußschen Rauschens in einem Empfängerzweig zur mittleren Signalleistung E_S in einem Senderzweig. Es ist interessant mit (3.28) festzustellen, dass $\alpha = 0$ den MMSE- in den ZF-Empfänger nach (3.14) überführt. Beim Entwurf des letzteren blieb ja das Rauschen unberücksichtigt.

Setzt man (3.28) in (3.1) ein, dann erhält man das Empfängerausgangssignal

$$\mathbf{r} = \underbrace{(\mathbf{H}^H \mathbf{H} + \alpha \mathbf{I}_M)^{-1} \mathbf{H}^H \mathbf{H}}_{\mathbf{Q}} \mathbf{s} + \mathbf{W} \quad (3.29)$$

Die Matrix \mathbf{Q} bestimmt den noch verbleibenden Anteil an der Interkanalinterferenz. Beim ZF-Empfänger ist $\alpha = 0$ und damit verschwindet mit $\mathbf{Q} = \mathbf{I}_M$ die Interkanalinterferenz vollständig.

3.3 Empfänger nach dem Prinzip des „Maximum Likelihood“

Wir betrachten wieder die Anordnung in Bild 2.7. Im Gegensatz zu den beiden vorangehenden Lösungen soll das Empfängerausgangssignal $y(k)$ nicht durch eine Matrix \mathbf{W} , sondern durch eine statistische Schätzung aus $\mathbf{r}(k)$ für festes k bestimmt werden. Hierzu definieren wir die bedingte WDF

$$p_L(\mathbf{r} | \mathbf{H} \mathbf{s}) \quad (3.30)$$

die auch als Likelihood-WDF bezeichnet wird. Dabei ist \mathbf{r} der beobachtete und verrauschte Empfangsvektor, und $\mathbf{H} \mathbf{s}$ ist der Sendevektor, bewertet mit der (bekannten) Kanalmatrix. $\mathbf{H} \mathbf{s}$ ist jedoch beim Empfänger nicht bekannt. (3.30) beschreibt nicht nur eine WDF, sondern für alle möglichen Sendevektoren \mathbf{s} eine Menge von WDF. Der Maximum-Likelihood (ML)-Empfänger wählt aus allen Möglichkeiten $\mathbf{H} \mathbf{s}$ diejenige aus, die \mathbf{r} „am ähnlichsten“ ist. Das liefert den Schätzwert $\mathbf{y}(k)$ des Sendevektors \mathbf{s} zum Zeitpunkt k . Der Empfänger sucht also aus der Menge der WDF $p_L(\mathbf{r} | \mathbf{H} \mathbf{s})$ die größte aus, d. h. die ML-Entscheidungsregel ist

$$p_L(\mathbf{r} | \mathbf{H} \mathbf{s}) = \max_{\{\text{alle } \mathbf{s}\}} \quad (3.31)$$

Daraus folgt der Schätzwert \mathbf{y} für \mathbf{s}

$$\mathbf{y} = \arg \max p_L(\mathbf{r} | \mathbf{H} \mathbf{s}) \quad (3.32)$$

Treten alle Sendesymbole s_v ($v = 1, \dots, M$) gleich häufig auf, dann lässt sich zeigen, dass durch (3.31) auch die A-Posteriori-WDF $p_{AP}(\mathbf{H} \mathbf{s} | \mathbf{r})$ maximiert und die Symbolfehlerwahrscheinlichkeit minimiert werden.

Mit (2.35) erhält man

$$p_L(\mathbf{r} | \mathbf{H} \mathbf{s}) = p_n(\mathbf{r} - \mathbf{H} \cdot \mathbf{s}) = \left(\frac{1}{\pi \sigma_n^2} \right)^N e^{-\frac{\|\mathbf{r} - \mathbf{H} \cdot \mathbf{s}\|^2}{\sigma_n^2}} \quad (3.33)$$

Mit (3.31) folgt dann, dass der Exponent in (3.33) minimiert werden muss

$$\|\mathbf{r} - \mathbf{H} \cdot \mathbf{s}\|^2 = \min_{\{\text{alle } \mathbf{s}\}} \quad (3.34)$$

und man erhält

$$\mathbf{y} = \arg \min \|\mathbf{r} - \mathbf{H} \cdot \mathbf{s}\|^2 \quad (3.35)$$

Offensichtlich führt nach (3.34) die ML-Schätzung auf die Minimierung des quadratischen Fehlers zwischen \mathbf{r} und $\mathbf{H} \mathbf{s}$ oder zur Minimierung der quadratischen Distanz zweier Vektoren \mathbf{r} und $\mathbf{H} \mathbf{s}$. Die statistische Schätzaufgabe wird demnach eine geometrische Problemstellung.

Der ML-Empfänger mit dem Algorithmus nach (3.34) betrachtet also alle möglichen Sendevektoren,

deren Menge natürlich vorab bekannt sein muss, und wählt den Vektor $\mathbf{y} = \mathbf{s}$ als besten Schätzwert aus, der den quadratischen Fehler in (3.34) minimiert.

Beispiel: M Sendeantennen

Jedes Symbol s_v kann L verschiedene Werte annehmen ($v = 1, \dots, M$).

Dann gibt es L^M verschiedene Vektoren \mathbf{s} .

Der Rechenaufwand steigt also exponentiell mit der Zahl der Sendeantennen.

4. Kanalkapazität von MIMO-Systemen

4.1 Einführung

Wir betrachten zunächst ein SISO-System, d. h. je eine Antenne beim Sender und Empfänger. Wie in Bild 4.1 dargestellt, sei am Eingang das komplexe Signal $s_v(k)$ wirksam. Real- und Imaginärteil von s_v seien statistisch unabhängig. s_v sei mittelwertfrei und besitze die mittlere Leistung $E_S = E[|s_v|^2]$. Es wird ein AWGN-Kanal betrachtet. Er überlagert dem Eingangssignal komplexwertiges weißes Gaußsches Rauschen $n_v(k)$, das mittelwertfrei ist und die mittlere Rauschleistung $\sigma_n^2 = E[|n_v|^2]$ besitzt. Real- und Imaginärteil sind statistisch unabhängig.

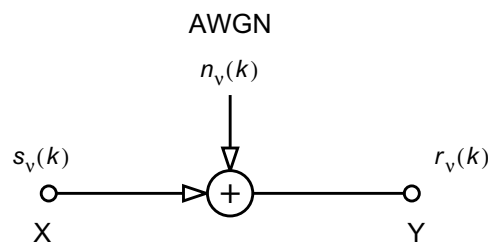


Bild 4.1 Zeitdiskreter AWGN-Kanal

In der Informationstheorie stellt man nun die wichtige Frage nach der Kanalkapazität C , d. h. wieviel bit/Symbol der Kanal übertragen kann. Für den SISO-Kanal ist das Ergebnis bekannt und wurde von Shannon [13] angegeben. Das Eingangssymbol s_v wird als Zufallsvariable X und das Ausgangssymbol r_v als Zufallsvariable Y aufgefasst. Die Kanalkapazität C ist dann die maximale wechselseitige Information (Transinformation) $I(X, Y)$. $I(X, Y)$ ist dabei die Information, die wir aus der Beobachtung von Y über X gewinnen.

$$C = \max_{p_x(X)} I(X, Y) \quad (4.1)$$

Die Maximierung muss über alle freien Parameter erfolgen. Das ist hier die WDF $p_x(X)$ von X , wobei die Nebenbedingung der begrenzten Sendeleistung E_S berücksichtigt werden muss. Mit dem Logarithmus zur Basis 2 (ld) erhält man nach [13]

$$C = \text{ld} \left(1 + \frac{E[|s_v|^2]}{E[|n_v|^2]} \right) = \text{ld} \left(1 + \frac{E_S}{\sigma_n^2} \right) \quad (4.2)$$

Man misst C in bit/Symbol. Da die Signaleinspeisung einem Kanalzugriff gleichkommt, ist als Einheit auch bit/Kanalbenutzung üblich. Man kann zeigen, dass C in (4.2) erreicht wird, wenn s_v eine kontinuierliche Zufallsvariable mit Gaußscher WDF ist. Dann besitzt r_v ebenfalls eine Gaußsche WDF. Für reelle Signale tritt in (4.2) noch ein Faktor $1/2$ hinzu. Wir nehmen an, dass die Symbole $s_v(k)$ im zeitlichen Abstand T aufeinander folgen und über einen Bandpasskanal mit der Bandbreite $B = 2 f_g$ mit QAM übertragen werden sollen. Sind die $s_v(k)$ Abtastwerte eines zeitkontinuierlichen tiefpassförmigen Nutzsymbols $\tilde{s}_v(t)$, dann kann dessen Grenzfrequenz f_g nach dem Abtasttheorem höchstens

$$f_g = \frac{1}{2T} = \frac{B}{2} \quad (4.3)$$

betragen. Das ist auch die kleinste Grenzfrequenz eines Nyquist-Rechtecktieffasses zur Datenübertragung ohne Intersymbolinterferenz. Damit lässt sich eine zweite Form angeben. Mit (4.3) und (4.2) erhält man

$$C' = C/T = B C = B \text{ld} \left(1 + \frac{E_S}{\sigma_n^2} \right) \quad (4.4)$$

und misst C' in $\frac{\text{Symbol}}{s} \cdot \frac{\text{bit}}{\text{Symbol}} = \frac{\text{bit}}{s}$. C' wird oft als der maximale Informationsfluss bezeichnet. Auch wird die auf B bezogene Größe angegeben

$$C'' = C'/B = \text{ld} \left(1 + \frac{E_S}{\sigma_n^2} \right) \text{ in } \frac{\text{bit/s}}{\text{Hz}} \quad (4.5)$$

Wie man aus (4.4) erkennt, ist die Zunahme von C' linear mit der Bandbreite B und lediglich logarithmisch mit dem Signal-Rauschverhältnis (SNR) E_S/σ_n^2 .

4.2 Instantane MIMO-Kapazität

Die Kapazität eines Systems mit mehreren Eingängen und Ausgängen wird vom Shannonschen Theorem [13] zunächst nicht abgedeckt. Telatar [5], [14] hat diese Berechnung für den MIMO-Kanal mit unkorreliertem Rayleigh-Fading und Gaußschen Rauschen durchgeführt. Er konnte dadurch sehr große Kapazitätsgewinne vorhersagen und hat damit intensive Forschungsarbeiten ausgelöst. In [15]-[17] werden obere Schranken der Kapazität für den IID-MIMO-Kanal und korreliertem Rayleigh-Fading aufgezeigt. [18]-[22] behandeln den allgemeinen Fall sender- und empfängerseitig korrelierter MIMO-Kanäle mit beliebigen Antennenzahlen. Über eine mathematisch geschlossene Form der momentengenerierenden Funktion gelingt dabei auch eine einheitliche Darstellung der bisher in der Literatur vorhandenen Teillösungen, besonders für die ergodische MIMO-Kapazität. Die Darstellung in diesem Kapitel gibt eine Einführung in die wichtigsten Zusammenhänge.

4.2.1 Voraussetzungen

Wir gehen wieder von der Anordnung in Bild 2.7 aus. Dabei ist

$$\mathbf{s} \in \mathcal{C}^{M \times 1}, \quad \mathbf{n}, \quad \mathbf{r} \in \mathcal{C}^{N \times 1}, \quad \mathbf{H} \in \mathcal{C}^{N \times M} \quad (4.6)$$

Sendevektor und Rauschvektor seien räumlich unkorreliert, d. h.

$$\mathbf{R}_{ss} = E_S \mathbf{I}_M, \quad \mathbf{R}_{nn} = \sigma_n^2 \mathbf{I}_N \quad (4.7)$$

Die gesamte mittlere Sendeleistung ist dann

$$E [\|\mathbf{s}\|^2] = M E_S \quad (4.8)$$

und die gesamte mittlere Rauschleistung wird

$$E [\|\mathbf{n}\|^2] = N \sigma_n^2 \quad (4.9)$$

4.2.2 Zerlegung des MIMO-Systems in Eigenmoden

Um die Kanalkapazität des MIMO-Systems zu berechnen, wird hier eine Zerlegung in parallele, unabhängige SISO-Zweige durchgeführt, auf die dann die Ergebnisse in Abschnitt 4.1 angewandt werden sollen. Wir führen hierzu eine Singulärwertzerlegung (Singular Value Decomposition, SVD) der MIMO-Kanalmatrix $\mathbf{H}(k)$ durch. Auch im Weiteren vereinfachen wir die Schreibweise und lassen die diskrete Zeit k weg. Man erhält:

$$\mathbf{H} = \mathbf{U} \cdot \mathbf{D} \cdot \mathbf{V}^H \quad (4.10)$$

wobei

$$\mathbf{D} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \vdots & 0 \\ 0 & \sqrt{\lambda_P} & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & 0 \end{bmatrix} \quad (N \times M) \quad (4.11)$$

und

$$\mathbf{U} \in \mathbb{C}^{N \times N}, \quad \mathbf{U}^{-1} = \mathbf{U}^H \quad (4.12)$$

$$\mathbf{V} \in \mathbb{C}^{M \times M}, \quad \mathbf{V}^{-1} = \mathbf{V}^H \quad (4.13)$$

\mathbf{U} ist die unitäre Matrix der Eigenvektoren zu den Eigenwerten $\lambda_1, \dots, \lambda_N$ der Matrix $\mathbf{H}\mathbf{H}^H$. \mathbf{V} ist eine unitäre Matrix, die aus \mathbf{U} und \mathbf{H} ermittelt werden kann. Da $\mathbf{H}\mathbf{H}^H$ eine Hermitesche Matrix ist, sind ihre Eigenwerte λ_v reell und nicht negativ, wie im Anhang 9.1 gezeigt wird.

$\sqrt{\lambda_1} \dots \sqrt{\lambda_P}$ heißen Singulärwerte von \mathbf{H} und

$$P = \text{rang}(\mathbf{H}\mathbf{H}^H), \quad P \leq \min\{M, N\} \quad (4.14)$$

$$\lambda_1 \geq \lambda_2 \geq \dots \lambda_P > 0; \lambda_{P+1} = \dots = \lambda_N = 0 \quad (4.15)$$

sind die Eigenwerte von $\mathbf{H}\mathbf{H}^H$. Der Beweis wird in Anhang 9.4 geführt.

Setzen wir (4.10) in (2.11) ein, erhält man

$$\mathbf{r} = \mathbf{U} \cdot \mathbf{D} \cdot \mathbf{V}^H \cdot \mathbf{s} + \mathbf{n} \quad (4.16)$$

Multipliziert man (4.16) von links mit \mathbf{U}^H , dann folgt

$$\underbrace{\mathbf{U}^H \mathbf{r}}_{\bar{\mathbf{r}}} = \mathbf{D} \cdot \underbrace{\mathbf{V}^H \cdot \mathbf{s}}_{\bar{\mathbf{s}}} + \underbrace{\mathbf{U}^H \cdot \mathbf{n}}_{\bar{\mathbf{n}}} \quad (4.17)$$

$$\bar{\mathbf{r}} = \mathbf{D} \cdot \bar{\mathbf{s}} + \bar{\mathbf{n}} \quad (4.18)$$

und mit den Abkürzungen (4.18) wird aus (4.17)

$$\bar{\mathbf{r}} = \mathbf{D} \cdot \bar{\mathbf{s}} + \bar{\mathbf{n}} \quad (4.19)$$

(4.19) beschreibt in Matrixform eine alternative Darstellung des MIMO-Systems. Die neue Kanalmatrix \mathbf{D} nach (4.11) ist nur in der Diagonalen besetzt, d. h. die einzelnen Zweige sind entkoppelt. Interkanalinterferenz tritt also in diesem neuen Modell nicht mehr auf. Dies erkennt man auch leicht, wenn man die μ -te Gleichung aus (4.19) schreibt:

$$r_\mu = \sqrt{\lambda_\mu} \bar{s}_\mu + \bar{n}_\mu \quad \mu = 1, \dots, N \quad (4.20)$$

(4.20) lässt sich in das Modell in Bild 4.2 (a) umsetzen. Die Zweige $P+1$ bis N tragen nur zum Empfängerrauschen bei, da der Signalanteil wegen der verschwindenden Singulärwerte Null ist. Diese Zweige werden daher oft nicht mehr berücksichtigt. Der Index μ in (4.20) läuft dann von $\mu = 1, \dots, P$. Die Matrixnotation (4.17) bis (4.19) lässt sich anschaulich in Bild 4.2 (b) umsetzen.

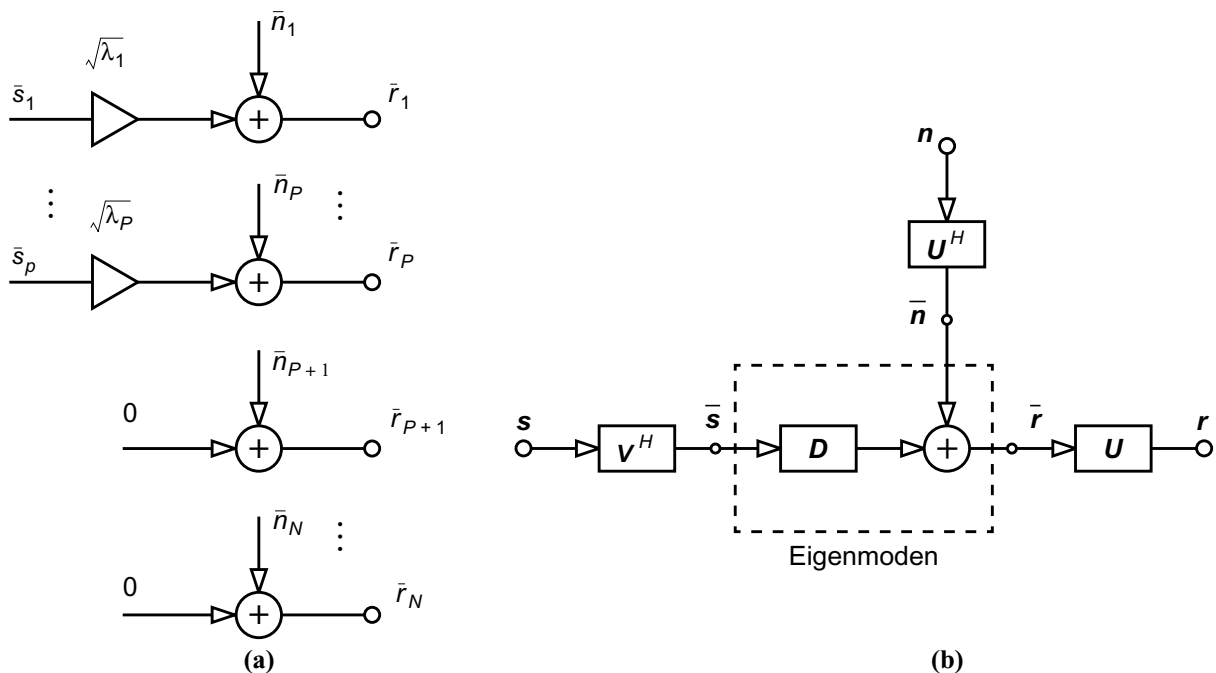


Bild 4.2 (a) Zerlegung des MIMO-Systems in parallele, unabhängige Eigenmoden, (b) Blockschaltbild der Eigenmoden in Matrixform

Aus (4.18) erhalten wir die Autokorrelationsmatrix von \bar{s}

$$\mathbf{R}_{\bar{s}\bar{s}} = E [\bar{s} \bar{s}^H] = E [V^H \cdot s \cdot s^H V] = V^H E [s s^H] V = V^H \mathbf{R}_{ss} V \quad (4.21)$$

Speziell für \mathbf{R}_{ss} nach (4.7) folgt mit (4.13)

$$\mathbf{R}_{\bar{s}\bar{s}} = E_S \cdot \mathbf{I}_M \quad (4.22)$$

und aus (4.18) ermittelt man die Autokorrelationsmatrix von \bar{n}

$$\begin{aligned} \mathbf{R}_{\bar{n}\bar{n}} &= E [\bar{n} \cdot \bar{n}^H] = E[U^H n \cdot n^H U] \\ &= U^H E [n \cdot n^H] U = U^H \mathbf{R}_{nn} U \end{aligned} \quad (4.23)$$

Speziell für \mathbf{R}_{nn} nach (4.7) folgt mit (4.12)

$$\mathbf{R}_{\bar{n}\bar{n}} = \sigma_n^2 \mathbf{I}_N \quad (4.24)$$

Da \bar{n} aus n durch eine lineare Transformation entsteht, bleibt \bar{n} gaußisch. Mit $E[n] = 0$ folgt auch

$$E[\bar{n}] = U^H \cdot E[n] = 0 \quad (4.25)$$

d. h. \bar{n} ist ebenfalls mittelwertfreies Gaußsches Rauschen (AWGN).

4.2.3 Kanalkapazität

Mit den Voraussetzungen im Kapitel 4.2.1 und der Zerlegung in Eigenmoden nach 4.2.2 kann nun die instantane Kanalkapazität des MIMO-Systems ermittelt werden. Dazu werden noch die mittleren Leistungen des Signal- und Rauschanteils in \bar{r}_μ nach (4.20) benötigt: Mit (4.22) erhält man

$$E [|\sqrt{\lambda_\mu} \bar{s}_\mu|^2] = \lambda_\mu E [|\bar{s}_\mu|^2] = \lambda_\mu \cdot E_S \quad (4.26)$$

Aus (4.24) folgt

$$E [|\bar{n}_\mu|^2] = \sigma_n^2 \quad (4.27)$$

Nun können (4.26) und (4.27) auf (4.2) angewandt werden, und man erhält die Kanalkapazität für Zweig μ in Bild 4.2 (a)

$$C_\mu = \text{Id} \left(1 + \frac{\lambda_\mu \cdot E_S}{\sigma_n^2} \right) \quad (4.28)$$

Für das gesamte MIMO-System folgt damit die instantane Kanalkapazität als Summe aus den unabhängigen Zweigen:

$$C = \sum_{\mu=1}^P C_\mu = \sum_{\mu=1}^P \text{Id} \left(1 + \frac{\lambda_\mu E_S}{\sigma_n^2} \right) \quad (4.29)$$

C wird ebenfalls in bit/Kanalbenutzung gemessen.

(4.29) lässt sich noch in anderer Form schreiben. Bei genauer Betrachtung kann man das Argument des Logarithmuses als μ -tes Element der folgenden Determinante einer Diagonalmatrix auffassen

$$\left| \mathbf{I}_P + \Lambda_P \frac{E_S}{\sigma_n^2} \right| = \prod_{\mu=1}^P \left(1 + \lambda_{\mu} \frac{E_S}{\sigma_n^2} \right) \quad (4.30)$$

Dabei ist Λ_P die $P \times P$ Diagonalmatrix

$$\Lambda_P = \text{diag} (\lambda_1, \dots, \lambda_P) \quad (4.31)$$

Wendet man auf beide Seiten von (4.30) den Logarithmus an, folgt

$$\text{ld} \left| \mathbf{I}_P + \Lambda_P \frac{E_S}{\sigma_n^2} \right| = \sum_{\mu=1}^P \text{ld} \left(1 + \lambda_{\mu} \frac{E_S}{\sigma_n^2} \right) \quad (4.32)$$

Mit (4.29) erhält man schließlich die alternative Form der Kanalkapazität

$$C = \text{ld} \left| \mathbf{I}_P + \Lambda_P \frac{E_S}{\sigma_n^2} \right|$$

Ergänzt man Λ_P durch die Eigenwerte mit dem Wert Null und berücksichtigt, dass $\text{ld} 1 = 0$, dann kann man die weiteren beiden Beziehungen angeben

$$C = \text{ld} \left| \mathbf{I}_N + \Lambda_N \frac{E_S}{\sigma_n^2} \right| = \text{ld} \left| \mathbf{I}_M + \Lambda_M \frac{E_S}{\sigma_n^2} \right| \quad (4.33)$$

Im Anhang 9.5 wird ferner gezeigt, dass aus (4.33) auch die folgenden Beziehungen für die MIMO-Kanalkapazität hergeleitet werden können

$$C = \text{ld} \left| \mathbf{I}_N + \mathbf{H}\mathbf{H}^H \frac{E_S}{\sigma_n^2} \right| = \text{ld} \left| \mathbf{I}_M + \mathbf{H}^H \mathbf{H} \frac{E_S}{\sigma_n^2} \right| \quad (4.34)$$

Aus (4.29) wird der Kapazitätsgewinn durch mehrere Antennen unmittelbar deutlich. Alle Beiträge C_{μ} ($\mu = 1, \dots, P$) sind positiv und tragen somit zur Gesamtkapazität C bei, die demnach im Allgemeinen größer ist als bei einem System mit nur einer Sende- und einer Empfangsantenne (SISO). Allerdings geht nicht die Zahl M der Sende- und N der Empfangsantennen direkt in die Gesamtkapazität ein, sondern nur die Zahl $P \leq \min\{M, N\}$. Sie gibt die Anzahl der unabhängigen Moden an und ist höchstens gleich der minimalen Antennenzahl, die entweder durch den Sender oder den Empfänger festgelegt wird. Da ferner $P = \text{rang}(\mathbf{H}\mathbf{H}^H)$ ist, und der Rang einer Matrix die maximale Zahl ihrer linear unabhängigen Zeilen oder Spalten angibt, spielt die Unabhängigkeit der Pfade im MIMO-System eine entscheidende Rolle. Abhängigkeiten (im weiteren Sinne Korrelationen) reduzieren also die maximal erreichbare Steigerung der Kanalkapazität. Z. B. kann es Zeitpunkte k geben, bei denen nur $P = 2$ Moden aktiv sind, obwohl $M = 4$ Sende- und $N = 4$ Empfangsantennen vorhanden sind. Das heißt andererseits nicht, dass man je 2 Antennen „abschalten“ kann, denn nach Bild 4.2 (b) liegen zwischen den realen Antennen die $M \times M$ Matrix \mathbf{V}^H und die $N \times N$ Matrix \mathbf{U} .

Die λ_μ können im praktischen Anwendungsfall stark unterschiedliche Werte annehmen. Die Kanalkapazität nach (4.29) wird dann hauptsächlich von den Moden mit den größten Eigenwerten bestimmt.

Bild 4.3 aus [19] zeigt die ergodische Kanalkapazität als Funktion des sendeseitigen Korrelationskoeffizienten ρ_{tx} nach (2.28) und (2.29) für das Kronecker-Kanalmodell mit exponentieller Korrelation und ein 4x4-MIMO-System. Empfangskorrelation ist nicht vorhanden. Man erkennt, dass eine starke sendeseitige Korrelation die Kanalkapazität verringert, besonders für hohes SNR. Dabei wird definiert

$$\text{SNR} = E_b/N_0$$

wobei E_b die gesamte Energie pro empfangenem Informationsbit und $N_0 = \sigma_n^2$ ist.

In Bild 4.4 aus [19] ist die ergodische Kanalkapazität (EC) als Funktion des SNR für ein 4x4-MIMO-System für verschiedene sendeseitige Korrelationskoeffizienten von 0.3 bis 0.97 und festem Empfangskorrelationskoeffizienten 0.5 aufgetragen. Man erkennt, dass die Kanalkapazität mit steigender Korrelation abnimmt. In Bild 4.4 sind theoretisch ermittelte obere Schranken und Ergebnisse einer Simulation eingetragen, die recht gut übereinstimmen.

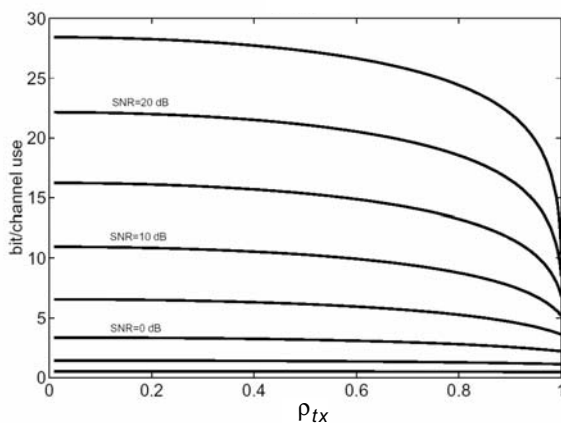


Bild 4.3 Ergodische Kanalkapazität eines 4x4-MIMO-Systems als Funktion des sendeseitigen Korrelationskoeffizienten. Kanalmodell mit exponentieller Korrelation ohne Empfangskorrelation

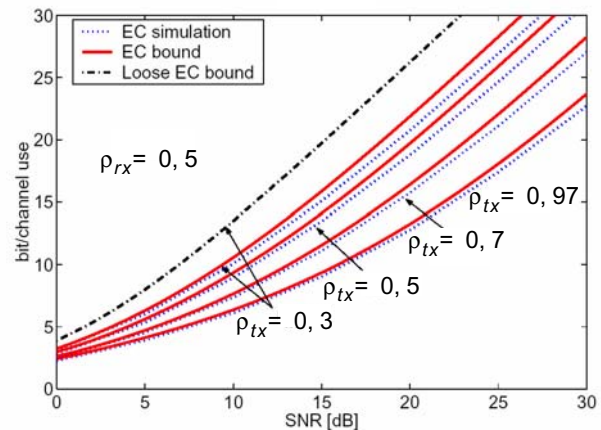


Bild 4.4 Ergodische Kanalkapazität eines 4x4-MIMO-Systems als Funktion des Signal-Rauschverhältnisses SNR für verschiedene Sendekorrelationskoeffizienten. Kanalmodell mit exponentieller Korrelation und festem Empfangskorrelationskoeffizienten 0.5

4.3 Maximierung der instantanen MIMO-Kanalkapazität

4.3.1 Verteilung der Sendeleistung bei bekanntem Kanal („Water Filling“)

In diesem Kapitel soll die Frage beantwortet werden, wie die Kanalkapazität des MIMO-Systems bei gegebener Antennenzahl maximiert werden kann. Denn dann wird es möglich, einen gegebenen Frequenzbereich bestmöglich auszunutzen, d. h. die spektrale Effizienz in bit/s pro Hz Bandbreite zu maximieren. Wir nehmen dazu an, dass die Kanalmatrix $\mathbf{H}(k)$ zu jedem Zeitpunkt k beim Empfänger durch geeignete Kanalschätzverfahren ermittelt wurde, und die Matrixelemente von

$\mathbf{H}(k)$ über einen Rückkanal auch dem Sender zur Verfügung stehen. Welche Freiheitsgrade gibt es denn? Betrachten wir (4.29), so sind das zunächst nur die mittlere Sendeleistung E_S pro Antenne und die mittlere Rauschleistung σ_n^2 in jedem Empfängerzweig. Das Rauschen liegt durch die Wahl des Kanals fest, ebenso die λ_μ und die Modenzahl P . Einzig verbleibender Parameter ist also die mittlere Sendeleistung, die nun - anders als in (4.29) - für jede Sendeantenne individuell so gewählt werden könnte, dass C maximal wird. Dies erfordert eine geringfügige Verallgemeinerung von (4.29). Dazu gehen wir zurück auf die Modenzerlegung des MIMO-Systems in Bild 4.2 (a) und die Autokorrelationsmatrix $\mathbf{R}_{\bar{s}\bar{s}}$ des Eingangssignalvektors \bar{s} . Durch die Annahme, dass das ursprüngliche Sendesignal \mathbf{s} die Autokorrelationsmatrix $\mathbf{R}_{ss} = E_S \mathbf{I}_M$ in (4.7) besitzt, zeigt auch $\mathbf{R}_{\bar{s}\bar{s}} = E_S \cdot \mathbf{I}_M$ in (4.22) Diagonalfom mit gleichen Diagonalelementen. Die mittleren Sendeleistungen $E[|\bar{s}_\mu|^2]$ sind aber gerade die Elemente in der Diagonalen und sollen nun individuell für jeden Modenzweig angepasst werden

$$E[|\bar{s}_\mu|^2] = a_\mu E_S \quad \mu = 1, \dots, M \quad (4.35)$$

$$\text{mit} \quad a_\mu \geq 0 \quad \mu = 1, \dots, M \quad (4.36)$$

Eine wichtige Nebenbedingung bei der Maximierung der Kanalkapazität ist eine konstante Sendeleistung, da sonst der Optimierungsalgorithmus die Trivillösung einer unendlichen Sendeleistung liefern würde. Soll die gesamte mittlere Sendeleistung von \bar{s}

$$E[\bar{s}^H \bar{s}] = M \cdot E_S \quad (4.37)$$

sein, dann muss die Nebenbedingung lauten

$$\sum_{\mu=1}^M a_\mu = M \quad (4.38)$$

Da wir für die Maximierung von C wieder auf das Modenmodell zurückgehen, müssen die Parallelzweige voneinander unabhängig sein. Das gilt nach wie vor auch für die Eingangssignale \bar{s}_μ , d. h. $\mathbf{R}_{\bar{s}\bar{s}}$ muss weiterhin eine Diagonalmatrix sein, allerdings nun im Vergleich zu (4.22) mit ungleichen Elementen in der Diagonalen. Man erhält also

$$\mathbf{R}_{\bar{s}\bar{s}} = E_S \text{diag}(a_1, \dots, a_M) \quad (4.39)$$

Man beachte, dass $\mathbf{R}_{\bar{s}\bar{s}}$ wie bisher nach (4.21) eine $M \times M$ Matrix ist. Nach Bild 4.2 (a) tragen allerdings \bar{s}_μ für $\mu > P$ nichts zur Übertragung bei, da die entsprechenden $\lambda_\mu = 0$ sind.

Aus (4.21) folgt durch einfache Umformung

$$\mathbf{R}_{ss} = \mathbf{V} \mathbf{R}_{\bar{s}\bar{s}} \mathbf{V}^H \quad (4.40)$$

Wegen $\mathbf{R}_{\bar{s}\bar{s}}$ nach (4.39), ist \mathbf{R}_{ss} keine Diagonalmatrix mehr wie in (4.7), sondern eine im Allgemeinen voll besetzte Matrix. Man arbeitet (4.35) leicht in (4.29) ein, denn man muss dort nur E_S durch $a_\mu E_S$ ersetzen. Damit erhält man die Optimierungsaufgabe:

$$C = \sum_{\mu=1}^P \text{ld} \left(1 + \frac{\lambda_\mu a_\mu E_S}{\sigma_n^2} \right) \Big|_{\{a_1, \dots, a_P\}} = \max \quad (4.41)$$

mit der Nebenbedingung

$$\sum_{\mu=1}^P a_{\mu} = M \quad \Leftrightarrow \quad g = \sum_{\mu=1}^P a_{\mu} - M = 0 \quad (4.42)$$

Da (4.41) nur a_1, \dots, a_P enthält, muss (4.38) angepasst und die obere Grenze der Summe in (4.42) P sein. Das bedeutet physikalisch, dass die gesamte mittlere Sendeleistung $M E_S$ auf a_1, \dots, a_P verteilt wird und $a_{P+1} = \dots = a_M = 0$. Die Lösung findet man mit der Methode nach Lagrange. Hierzu definieren wir die Zielfunktion

$$J = C + L \cdot g = \max_{\{a_1, \dots, a_P; L\}} \quad (4.43)$$

L ist eine neue Variable, der Lagrangemultiplikator. $J(a_1, \dots, a_P; L)$ ist eine konkave Funktion. Daher gibt es ein einziges Maximum. Es genügt bereits, die 1. Ableitung auszuwerten. Aus (4.43) folgt

$$\frac{\partial J}{\partial a_v} = \frac{\partial C}{\partial a_v} + L \cdot \frac{\partial g}{\partial a_v} = 0 \quad (4.44)$$

Mit (4.41), (4.42) und $\ln x = A \cdot \ln x$ ($A = 1/\ln 2$) sowie der Abkürzung

$$b_{\mu} = \lambda_{\mu} E_S / \sigma_n^2 \quad (4.45)$$

folgt aus (4.44)

$$\frac{\partial}{\partial a_v} [A \ln(1 + b_v a_v)] + L \frac{\partial g}{\partial a_v} = 0 \quad (4.46)$$

und man erhält

$$A b_v + L(1 + b_v a_v) = 0 \quad v = 1, \dots, P \quad (4.47)$$

Die Ableitung von J nach L liefert

$$\frac{\partial J}{\partial L} = g = 0$$

oder

$$\sum_{\mu=1}^P a_{\mu} = M \quad (4.48)$$

Aus (4.47) folgt mit (4.45) und $-A/L = K$

$$a_v = \left(K - \frac{\sigma_n^2}{\lambda_v E_S} \right)_+ \quad v = 1, \dots, P \quad (4.49)$$

(4.48) und (4.49) sind $P+1$ Gleichungen für die $P+1$ Unbekannten a_1, \dots, a_P und K . Man beachte, dass $\lambda_v > 0$; $v = 1, \dots, P$. Der Operator in (4.49) ist definiert durch

$$(x)_+ = \begin{cases} x & , x \geq 0 \\ 0 & , x < 0 \end{cases} \quad (4.50)$$

Er ist erforderlich, da Bedingung (4.36) eingehalten werden muss.

Bild 4.5 verdeutlicht das Berechnungsverfahren, das in der Informationstheorie als „Water-Filling-Algorithmus“ bezeichnet wird. Man sortiert zunächst die Eigenwerte $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_P > 0$, und trägt die Werte $\sigma_n^2 / (\lambda_v E_S)$ aus (4.49) ins Diagramm ein. Sie bilden gleichsam den treppenförmigen Boden eines Behälters. Die gesamte „Wassermenge“ ist durch die Summe der a_μ nach (4.48) fest vorgegeben. K ist der Wasserstand, der sich einstellt, wenn das Wasser eingefüllt ist. Wegen (4.50) kann man K nur in Sonderfällen in einem Schritt ermitteln. Meist sind mehrere Iterationen nötig. Das nachfolgende Beispiel wird dies deutlich machen. Aus Bild 4.5 erkennt man, dass ein übertragungstechnisch schlechter Modus, der ja kleines λ_μ besitzt, einen kleinen Koeffizienten a_μ und damit wenig Sendeleistung erhält. Dominante Pfade wie z. B. durch λ_1 gekennzeichnet transportieren die größte Sendeleistung $a_1 E_S$.

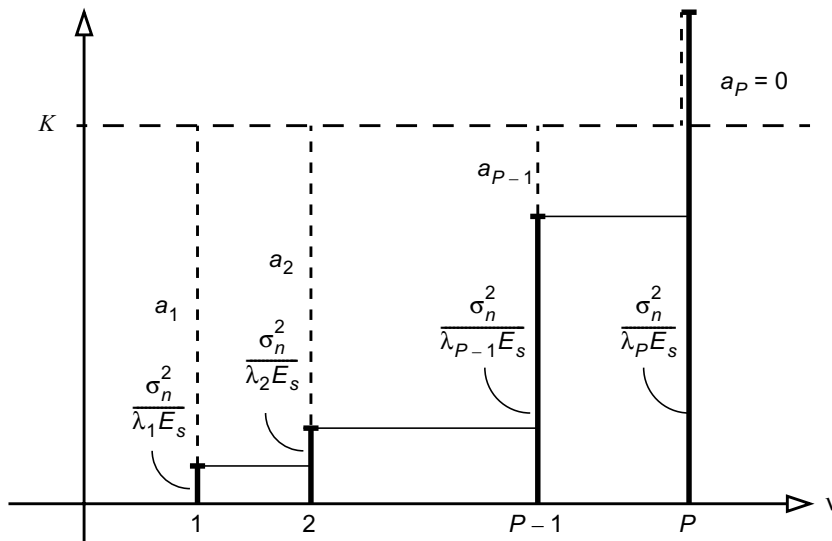


Bild 4.5 Ermittlung der optimalen Leistung $a_v E_S$ je Sendeantenne durch „Water Filling“

Nun zurück zum Algorithmus. Zur Vereinfachung der Schreibweise sei $c_v = \sigma_n^2 / (\lambda_v E_S)$. Dann wird aus (4.49)

$$a_v = (K - c_v)_+ \quad v = 1, \dots, P \quad (4.51)$$

Beispiel:

$$c_1 = 1, \quad c_2 = 2, \quad c_3 = 10, \quad c_4 = 19, \quad M = 8, \quad P = 4$$

1. Schritt

$$a_4 = K - c_4$$

$$a_3 = K - c_3$$

$$a_2 = K - c_2$$

$$a_1 = K - c_1$$

$$M = 4K - (c_1 + \dots + c_4)$$

$$\Rightarrow K = 10$$

$$a_4 = (10 - 19)_+ = 0$$

2. Schritt

$$a_3 = K - c_3$$

$$a_2 = K - c_2$$

$$a_1 = K - c_1$$

$$M = 3K - (c_1 + \dots + c_3)$$

$$\Rightarrow K = 7$$

$$a_3 = (7 - 10)_+ = 0$$

3. Schritt

$$a_2 = K - c_2$$

$$a_1 = K - c_1$$

$$M = 2K - (c_1 + c_2)$$

$$\Rightarrow K = 5.5$$

$$a_2 = (K - c_2)_+ = 3.5$$

$$a_1 = (K - c_1)_+ = 4.5$$

Lösung: $a_1 = 4.5$, $a_2 = 3.5$, $a_3 = a_4 = 0$; alle übrigen Null.

Dieser Algorithmus kann leicht von einem Signalprozessor im Endgerät ausgeführt werden.

4.3.2 Ermittlung des optimalen Sendefilters

Im vorigen Abschnitt wurde ein Verfahren angegeben, wie die Elemente a_v der Matrix $\mathbf{R}_{\bar{s}\bar{s}}$ in (4.39) ermittelt werden können, um die Sendeleistung so auf die Einzelmoden zu verteilen, dass die MIMO-Kanalkapazität maximal wird. Nach (4.40) ergab sich damit, dass der Sendevektor \mathbf{s} in Bild 4.2 (b) nicht länger räumlich unkorreliert ist, sondern die Autokorrelationsmatrix \mathbf{R}_{ss} nach (4.40) besitzen muss. Nun stellt sich die Frage, wie man ein solches Sendesignal \mathbf{s} erzeugen kann. Eine recht einfache Lösung ist ein Sendefilter nach Bild 4.6 mit der Filtermatrix \mathbf{F} .

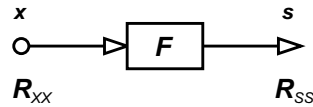


Bild 4.6 Sendefilter mit Matrix \mathbf{F} zur Maximierung der MIMO-Kanalkapazität

Wir gehen von einem Signalvektor $\mathbf{x} \in \mathcal{C}^{M \times 1}$ beim Sender aus, der räumlich unkorreliert ist und dessen Elemente x_v ($v = 1, \dots, M$) gleiche mittlere Leistung E_S besitzen. x_v bezeichnen dann die zu sendenden Symbole. Die Autokorrelationsmatrix von \mathbf{x} ist demnach

$$\mathbf{R}_{xx} = E [\mathbf{x} \mathbf{x}^H] = E_S \mathbf{I}_M \quad (4.52)$$

Nach Bild 4.6 ist

$$\mathbf{s} = \mathbf{F} \mathbf{x} \quad (4.53)$$

und

$$\mathbf{R}_{ss} = E [\mathbf{s} \mathbf{s}^H] = E [\mathbf{F} \mathbf{x} \mathbf{x}^H \mathbf{F}^H] = \mathbf{F} E [\mathbf{x} \mathbf{x}^H] \mathbf{F}^H$$

Mit (4.52) folgt

$$\mathbf{R}_{ss} = \mathbf{F} \mathbf{R}_{xx} \mathbf{F}^H = E_S \cdot \mathbf{F} \mathbf{F}^H \quad (4.54)$$

Mit (4.39) und (4.40) folgt

$$\mathbf{F} \mathbf{F}^H = \mathbf{V} \mathbf{A}^{1/2} \mathbf{A}^{1/2} \mathbf{V}^H \quad (4.55)$$

wobei

$$\mathbf{A}^{1/2} = \text{diag} (\sqrt{a_1}, \dots, \sqrt{a_M}) \quad (4.56)$$

Da $a_v \geq 0$ ($v = 1, \dots, M$), sind die Elemente von $\mathbf{A}^{1/2}$ reell.

Damit kann man die rechte Seite von (4.55) in das Produkt zweier Matrizen \mathbf{F} und \mathbf{F}^H zerlegen und erhält für die Matrix \mathbf{F} des Sendefilters

$$\mathbf{F} = \mathbf{V} \text{diag} (\sqrt{a_1}, \dots, \sqrt{a_M}) \quad (4.57)$$

Bild 4.7 zeigt den relativen Gewinn an ergodischer Kanalkapazität (EC) durch "Water Filling" als Funktion des SNR für verschiedene sendeseitige Korrelationskoeffizienten nach (2.28) und (2.29) zwischen 0.5 und 0.97 für ein Kanalmodell mit exponentieller Korrelation. Der Empfangskorrelationskoeffizient ist fest und beträgt 0.7. Wie man erkennt, schmilzt der Gewinn mit zunehmendem SNR. Exakte Berechnung und Näherung stimmen gut überein, [19].

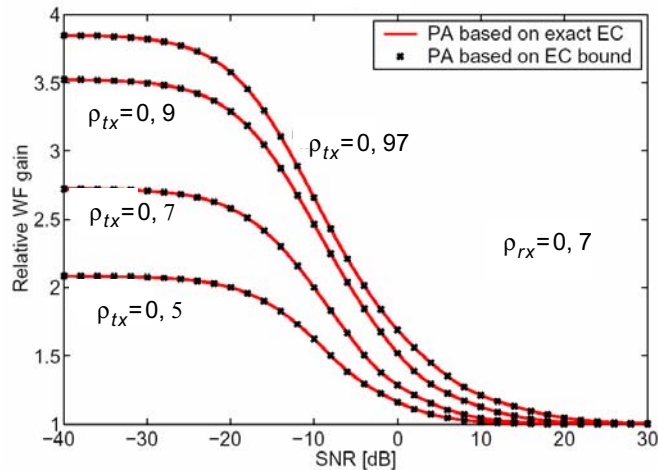


Bild 4.7 Relativer Kanalkapazitätsgewinn durch optimale Verteilung der Sendeleistung auf die Sendeantennen durch „Water Filling“ als Funktion des Signal-Rauschverhältnisses SNR für verschiedene Sendekorrelationskoeffizienten. 4x4-MIMO-System, Kanalmodell mit exponentieller Korrelation und festem Empfangskorrelationskoeffizienten 0.7

4.4 Kapazität des statistischen MIMO-Kanals

4.4.1 Ergodische Kanalkapazität

Bisher wurde die instantane Kapazität C betrachtet, die von der Kanalmatrix $\mathbf{H}(k)$ zum Zeitpunkt k ausgeht. $\mathbf{H}(k)$ wurde dabei als zeitvariant, aber nicht als zufällig aufgefasst. Daher ist C eigentlich eine Funktion der Zeit k , wie man z. B. aus (4.34) direkt erkennen kann. Um eine konstante Aussage über einen längeren Zeitabschnitt zu erhalten, kann man den zeitlichen Mittelwert über $C(k)$ nehmen. Unter der Voraussetzung einer ergodischen Statistik, kann man auch den Ensemble-Mittelwert nehmen, den man dann als ergodische Kapazität C_{erg} bezeichnet

$$C_{\text{erg}} = E[C] = \int_0^{\infty} C p_C(C) dC \quad (4.58)$$

Dabei ist $p_C(C)$ die WDF von C . Das Problem bei der Auswertung von (4.58) ist allerdings die Kenntnis von $p_C(C)$. Hierzu gibt es in der Literatur verschiedene Ansätze, z. B. [19], die hier nicht näher vertieft werden sollen.

4.4.2 Ausfallkapazität

Der Sender speist im Allgemeinen eine vorgegebene Bitrate in das MIMO-System ein, die nach der Informationstheorie an die Kanalkapazität C angepasst sein muss. Wie bereits erwähnt, ändert sich C aufgrund der zeitlichen Schwankungen (Fading) der Kanaleigenschaften. Fällt C unter eine vorgegebene Grenze C_{Aus} , dann kann das MIMO-System den eingespeisten Informationsfluss nicht mehr tragen, und es kommt kurzzeitig zu starken Bitfehlern beim Empfänger, die einem momentanen Betriebsausfall des Systems gleichkommen. C_{Aus} heißt daher Ausfallkapazität (outage capacity). Bei bekannter WDF $p_C(C)$ muss entsprechend Bild 4.8 die Ausfallwahrscheinlichkeit (outage probability) P_{Aus} ermittelt werden. Man erhält

$$P_{\text{Aus}} = P [C \leq C_{\text{Aus}}] = \int_0^{C_{\text{Aus}}} p_C(C) dC$$

Mit C_{Aus} und P_{Aus} kann man die Dienstqualität einstellen.

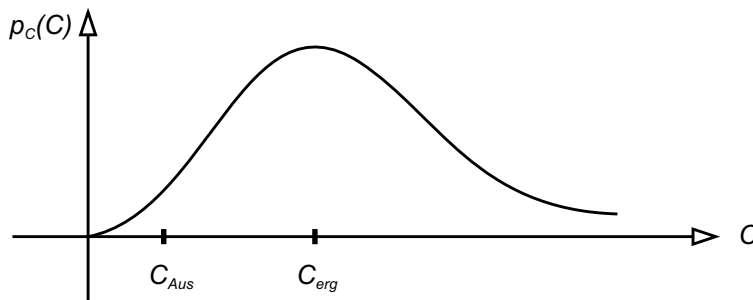


Bild 4.8 Prinzipieller Verlauf der Wahrscheinlichkeitsdichtefunktion $p_C(C)$ der MIMO-Kanalkapazität mit Ausfallkapazität C_{Aus}

5. Raum-Zeit-Codierung

5.1 Das Prinzip

Bild 5.1 zeigt das Prinzip eines MIMO-Senders mit Raum-Zeit-Codierung. Die Bitfolge $b(n)$ der Datenquelle durchläuft einen zeitlichen Codierer (einschließlich Interleaver), z. B. einen Faltungscodierer, der $q N_1$ Eingangsbits Redundanzbits hinzufügt, so dass im gleichen Zeitintervall der Länge LT an seinem Ausgang $q N_2$ codierte Bits auftreten. Die zeitliche Coderate ist also

$$r_t = N_1/N_2 \quad (5.1)$$

q aufeinanderfolgende Bits $b'(n')$ werden durch den QAM-Mapper einem Symbol $c(k)$ zugeordnet. Am Ausgang des Mappers treten also im Betrachtungszeitraum N_2 Symbole auf. Der Raum-Zeit-Codierer verteilt die einlaufenden Symbole nach einer noch zu besprechenden Codiervorschrift auf die Antennen. Im Beobachtungszeitraum LT laufen N_2 Symbole $c(k)$ ein, und L Vektorsymbole $s(k), \dots, s(k+L-1)$ werden von den Antennen ausgesendet. Die räumliche (spatiale) Coderate ist also

$$r_s = N_2/L = M \quad (5.2)$$

Man beachte, dass als Ausgangsgrößen der Antennenanordnung der Vektor \mathbf{s} gezählt wird und nicht die einzelnen Komponenten (Symbole). Die Coderate vom Eingang zum Ausgang der gesamten Anordnung in Bild 5.1 beträgt somit

$$r_{\text{total}} = \frac{q N_1}{L} = q r_t r_s \quad (5.3)$$

Die Besonderheit von Mehrantennensystemen besteht nun darin, dass man sowohl „konventionell“ zeitlich als auch zusätzlich räumlich-zeitlich codieren kann. Beim SISO-System ist nur die zeitliche Codierung vorhanden.¹⁾ Die Bezeichnung ist leider nicht eindeutig. Man spricht von Raum-Zeit-Codierung, wenn beide Codierer in Bild 5.1 oder nur der Raum-Zeit-Codierer vorhanden sind.

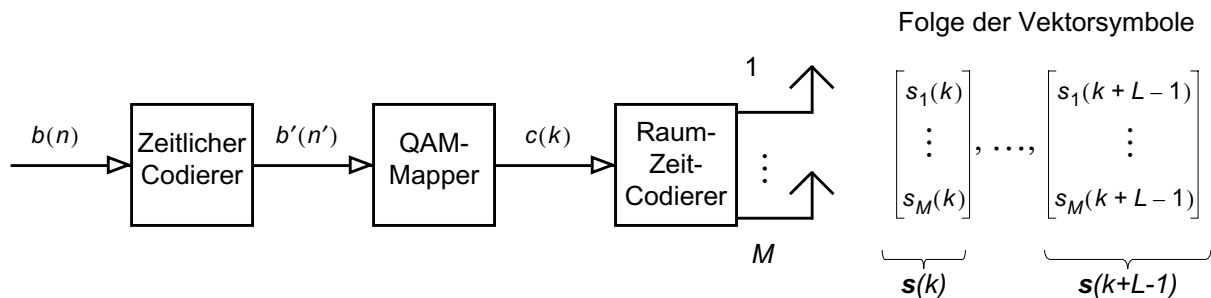


Bild 5.1 Prinzip eines MIMO-Senders mit zeitlicher und räumlicher Codierung

Am Ausgang des MIMO-Senders in Bild 5.1 kann man folgende $M \times L$ Matrix angeben

$$\mathbf{S} = \begin{matrix} & \text{Zeit } k & & \\ & \rightarrow & & \\ \begin{pmatrix} s_1(k) & \dots & s_1(k+L-1) \\ \vdots & & \vdots \\ s_M(k) & \dots & s_M(k+L-1) \end{pmatrix} & & \downarrow \text{Ort } \mathbf{v} & \end{matrix} \quad (5.4)$$

Für \mathbf{S} haben sich mehrere Bezeichnungen gebildet: Raum-Zeit-Codiermatrix, Raum-Zeit-Codewortmatrix oder Raum-Zeit-Codeblock. \mathbf{S} verknüpft die Symbole $c(k), c(k+1), \dots, c^*(k), c^*(k+1), \dots$ mit den Sendesymbolen $s_1(k), \dots, s_M(k+L-1)$ in geeigneter Weise, um die Symbolfehlerwahrscheinlichkeit beim Empfänger zu verringern. Hierbei gibt es viele Möglichkeiten, die man in zwei Gruppen einteilen kann:

- Raum-Zeit-Block-Codes: Die Zuordnung geschieht blockweise, und die Verknüpfung ist (meist) linear. Beispiele sind der Alamouti-Code [23] und die Verallgemeinerung nach Tarokh [8].
- Raum-Zeit-Trellis-Codes: Die Verknüpfung geschieht über einen Gittergraphen (Trellis).

1) Mehrträgersysteme erlauben eine Zeit-Frequenz-Codierung, die hier nicht betrachtet wird.

5.2 Raum-Zeit-Codierung nach Alamouti

Ein einfaches und dennoch wirksames Verfahren zur Verbesserung der Übertragungsqualität ohne Erhöhung der Bitrate wurde von Alamouti [23] angegeben. Es wird auch für Erweiterungen von UMTS in 3GPP vorgesehen [9]. Der Sender arbeitet mit $M = 2$, der Empfänger mit $N = 1$. Eine reine zeitliche Codierung wie in Bild 5.1 (links) dargestellt ist nicht vorhanden. Die Codiermatrix \mathbf{S} wird wie folgt festgelegt

$$\mathbf{S} = \begin{pmatrix} s_1(k) & s_1(k+1) \\ s_2(k) & s_2(k+1) \end{pmatrix} = \begin{pmatrix} c(k+1) & -c(k)^* \\ c(k) & c(k+1)^* \end{pmatrix} \quad \downarrow \text{Ort } v \quad (5.5)$$

$$\mathbf{S} = \begin{pmatrix} s_1(k) & -s_2^*(k) \\ s_2(k) & s_1^*(k) \end{pmatrix} \quad (5.6)$$

Da offensichtlich zwei Eingangssymbole $c(k)$, $c(k+1)$ zwei Vektorsymbolen am Ausgang zugeordnet werden ($L = 2$), beträgt die spatiale Coderate $r_s = 1$.

Man zeigt leicht, dass $\tilde{\mathbf{S}} = \mathbf{S} / \sqrt{|s_1|^2 + |s_2|^2}$ mit \mathbf{S} nach (5.6) eine unitäre Matrix ist, d. h.

$$\tilde{\mathbf{S}} \cdot \tilde{\mathbf{S}}^H = \mathbf{I}_2 \quad (5.7)$$

$$\text{Damit gilt auch } \tilde{\mathbf{S}}^H \cdot \tilde{\mathbf{S}} = \mathbf{I}_2 \quad (5.8)$$

(5.7) kann man auch als Orthogonalitätsbedingung bezüglich der Zeit (horizontale Richtung) und (5.8) bezüglich des Orts (vertikale Richtung) auffassen. Zur Berechnung des Empfangssignals wird angenommen, dass sich der MIMO-Kanal nur langsam ändert (slow fading), d. h.

$$h_{\mu\nu}(k) \approx h_{\mu\nu}(k+1)$$

gilt. Man erhält mit $\mathbf{r}(k) = r(k)$ und $\mathbf{n}(k) = n(k)$

$$\mathbf{r}(k) = \mathbf{H}(k) \mathbf{s}(k) + \mathbf{n}(k) \quad (5.9)$$

und daraus mit (5.6)

$$\mathbf{r}(k) = h_{11}(k) s_1(k) + h_{12}(k) s_2(k) + n(k) \quad (5.10)$$

$$\mathbf{r}(k+1) = -h_{11}(k) s_2^*(k) + h_{12}(k) s_1^*(k) + n(k+1) \quad (5.11)$$

oder in Matrixdarstellung

$$\begin{pmatrix} r(k) \\ r^*(k+1) \end{pmatrix} = \begin{pmatrix} h_{11}(k) & h_{12}(k) \\ h_{12}^*(k) & -h_{11}^*(k) \end{pmatrix} \begin{pmatrix} s_1(k) \\ s_2(k) \end{pmatrix} + \begin{pmatrix} n(k) \\ n^*(k+1) \end{pmatrix} \quad (5.12)$$

$$= \mathbf{U}(k) \quad (5.13)$$

U besitzt die Eigenschaft

$$U^H \cdot U = \beta \cdot I_2 \quad (5.14)$$

mit
$$\beta = |h_{11}(k)|^2 + |h_{12}(k)|^2 \quad (5.15)$$

d. h. $\frac{1}{\sqrt{\beta}} U$ ist unitär, und die inverse Matrix $U^{-1} = U^H$ kann einfach verwirklicht werden. Wählt man also wie in Bild 5.2 dargestellt $U^H(k)$ als Empfängermatrix, dann erhält man mit (5.12)

$$\underbrace{\begin{pmatrix} y_1(k) \\ y_2(k) \end{pmatrix}}_{\mathbf{y}(k)} = \underbrace{\beta(k)}_{\text{Signalanteil}} \underbrace{\begin{pmatrix} s_1(k) \\ s_2(k) \end{pmatrix}}_{\text{Signalanteil}} + \underbrace{U^H(k)}_{\text{gefiltertes Rauschen}} \underbrace{\begin{pmatrix} n(k) \\ n^*(k+1) \end{pmatrix}}_{\text{gefiltertes Rauschen}} \quad (5.16)$$

Man erkennt, dass diese Empfängermatrix den Empfangsvektor auch entkoppelt, d. h. $y_1(k)$ hängt nicht von $s_2(k)$ und $y_2(k)$ hängt nicht von $s_1(k)$ ab. Damit lässt sich die nachfolgende Maximum-Likelihood-Detektion einfacher gestalten.

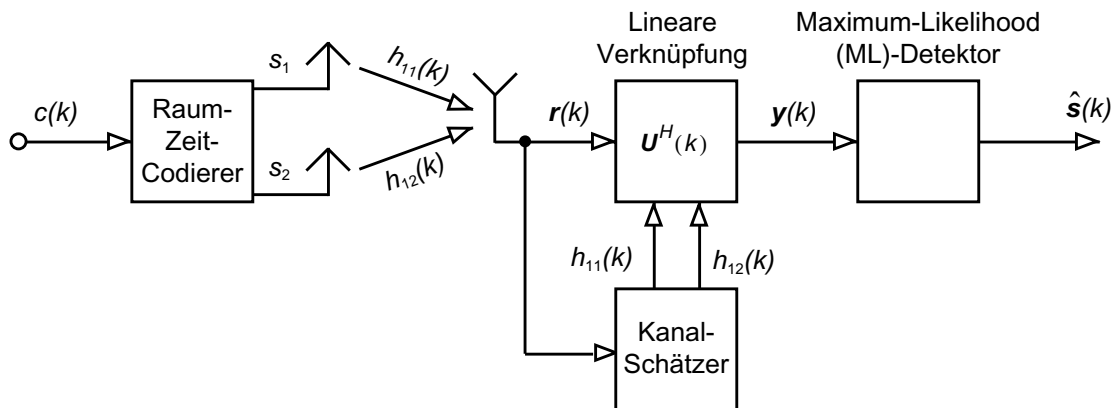


Bild 5.2 Raum-Zeit-Codierung nach Alamouti

Um $U^H(k)$ zu gewinnen, muss beim Empfänger ein Kanalschätzer vorhanden sein, der Näherungswerte von $h_{11}(k)$ und $h_{12}(k)$ bereitstellt. Schätzfehler beeinflussen natürlich die Leistungsfähigkeit des Systems, was hier nicht erörtert werden soll. Aus (5.15) erkennt man, dass sich $\mathbf{y}(k)$ aus dem mit Faktor $\beta(k)$ gewichteten Sendesignal und additivem, gefiltertem Gaußischem Rauschen zusammensetzt. In $\beta(k)$ verbirgt sich die Sendediversität durch die beiden Sendeantennen. h_{11} und h_{12} sind zwei unabhängige Signalwege. Wenn durch Fading $\beta(k) = 0$ wird, verschwindet der Signalanteil in $\mathbf{y}(k)$, und ein Übertragungsfehler ist aufgetreten. Dazu müssen $h_{11}(k)$ und $h_{12}(k)$ zum selben Zeitpunkt k Null werden. Diese Wahrscheinlichkeit ist aber klein. Mit der räumlich-zeitlichen Verteilung der Symbole $s_1(k)$ und $s_2(k)$ auf beide Sendeantennen wird also eine Verringerung der Symbolfehlerwahrscheinlichkeit erreicht.

5.3 Orthogonale Raum-Zeit-Block-Codes

In Anlehnung an das einfache Codierverfahren nach Alamouti kann man einen allgemeineren Ansatz für ein $M \times N$ MIMO-System machen [8], [24]-[27], und dabei $M \times L$ Codiermatrizen \mathbf{S} mit der Eigenschaft

$$\mathbf{S}\mathbf{S}^H = \beta \mathbf{I}_M, \quad \beta = K \sum_{v=1}^M |s_v|^2, \quad K = \text{const.} \quad (5.17)$$

einsetzen. Ähnlich wie (5.16) kann eine Empfängermatrix \mathbf{S}^H wieder eine Entkopplung der Signale liefern. Die Elemente von \mathbf{S} können aus s_v , s_v^* , $\text{Re } s_v$, $\text{Im } s_v$ und Verknüpfungen hieraus gebildet werden. \mathbf{S} muss bevorzugt so gewählt werden, dass sich die maximale Diversitätsordnung $M \cdot N$ (Sendediversität M , Empfangsdiversität N) ergibt. Die Synthese solcher Codematrizen ist mathematisch sehr anspruchsvoll. Ein Beispiel für eine Codiermatrix \mathbf{S} für $M = 3$ Sendeantennen und einer räumlichen Coderate $r_s = 1/2$ ist

$$\mathbf{S} = \begin{bmatrix} s_1 & -s_2 & -s_3 & -s_4 & s_1^* & -s_2^* & -s_3^* & -s_4^* \\ s_2 & s_1 & s_4 & -s_3 & s_2^* & s_1^* & s_4^* & -s_3^* \\ s_3 & -s_4 & s_1 & s_2 & s_3^* & -s_4^* & s_1^* & s_2^* \end{bmatrix} \quad (5.18)$$

Man überprüft leicht selbst, dass \mathbf{S} die Bedingung (5.17) erfüllt.

5.4 Raum-Zeit-Trellis-Codierung

Diese Codierungsart soll am Beispiel eines MIMO-Senders mit 4-PSK (Phase Shift Keying) in Bild 5.3 (a) dargestellt werden. Die PSK-Symbole $c(k)$ können nach Bild 5.3 (b) vier verschiedene komplexe Werte c_1, \dots, c_4 annehmen. Da das Signal $s_2(k)$ der Antenne 2 das verzögerte Signal $s_1(k-1)$ von Antenne 1 darstellt, spricht man im vorliegenden Fall auch von Verzögerungs-Diversität (delay diversity). Die Zustandsvariable $c(k-1)$ des diskreten Automaten kann nach Bild 5.3 (b) offensichtlich 4 Zustände einnehmen. Die Funktion des Senders stellt man in Form eines Zustandsübergangsdiagramms dar, das zum Trellis-Diagramm in Bild 5.3 (c) wird, fügt man eine Zeitachse k hinzu. Die Übergänge sind durch Pfeile sowie mit dem Eingangs- und Ausgangssignal des Verzögerungselements T gekennzeichnet. Befindet sich die Schaltung zur Zeit $k-1$ im Zustand 1, d. h. $c(k-1) = c_1$, dann muss am Eingang $c(k) = c_4$ wirksam werden, damit die Schaltung in den Zustand 4 gelangt.

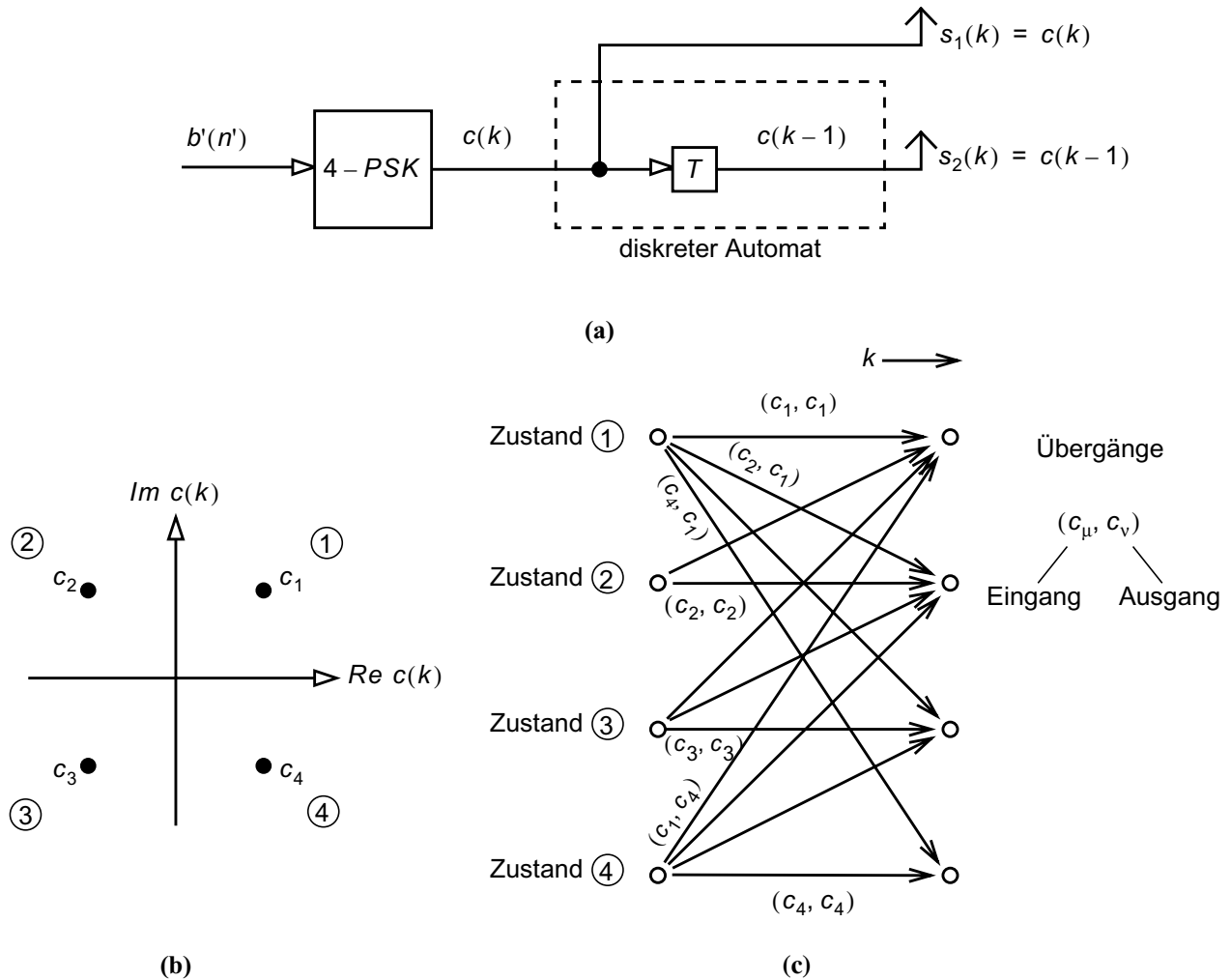


Bild 5.3 Raum-Zeit-Trellis-Codierer. (a) Blockdiagramm des Senders mit $M=2$ Antennen, (b) Konstellationsdiagramm für 4-PSK, (c) Trellis-Diagramm mit einigen (nicht allen) Zustandsübergängen

Die Decodierung beim Empfänger geschieht durch einen Maximum-Likelihood-Sequenz-Schätzer, der wie der Codierer mit einem Trellis arbeitet. Da seine Komplexität exponentiell mit der Zahl der Zustände wächst, ist der Empfänger aufwendiger als bei einer Block-Codierung, jedoch kann man einen zusätzlichen Codiergewinn erzielen.

6. Räumlicher Multiplex und Demultiplex

In diesem Kapitel werden die prinzipiellen Möglichkeiten einer Zuordnung der zeitlich aufeinander folgenden Symbole auf die Sendeantennen vorgestellt.

6.1 Räumlicher Multiplex ohne zeitliche Codierung

Bild 6.1 zeigt die grundsätzliche Anordnung zur Aufteilung der Symbole $\dots, c(0), c(1), \dots$ der Symbolfolge $c(k)$ auf die M Sendeantennen. Dabei wird angenommen, dass weder die Bitfolge noch die Symbolfolge in zeitlicher Richtung codiert wurde. Der Fall mit Codierung wird in den folgenden

Kapiteln behandelt. Die Symbole laufen im zeitlichen Abstand T , d. h. mit der Symbolrate $1/T$ in einen Seriell-Parallel-Wandler (S/P-Wandler, räumlicher Demultiplexer) ein. Dieser schreibt die Symbole nacheinander in Speicher ein. Nach der Zeit MT sind alle M Speicher mit neuen Symbolen $c(0), \dots, c(M-1)$ gefüllt. Dann erfolgt mit dem Vektorsymbol-Takt der Frequenz $1/(MT)$ die Ausgabe an die Antennen $1, \dots, M$ und die Bildung des Symbolvektors $s(k/M)$. Die spatiale Coderate ist

$$r_s = M \quad (6.1)$$

da M Symbole in der Zeit MT einlaufen und 1 Symbolvektor ausgegeben wird. Räumliches Demultiplexen liefert also die Besonderheit einer Coderate größer als Eins.

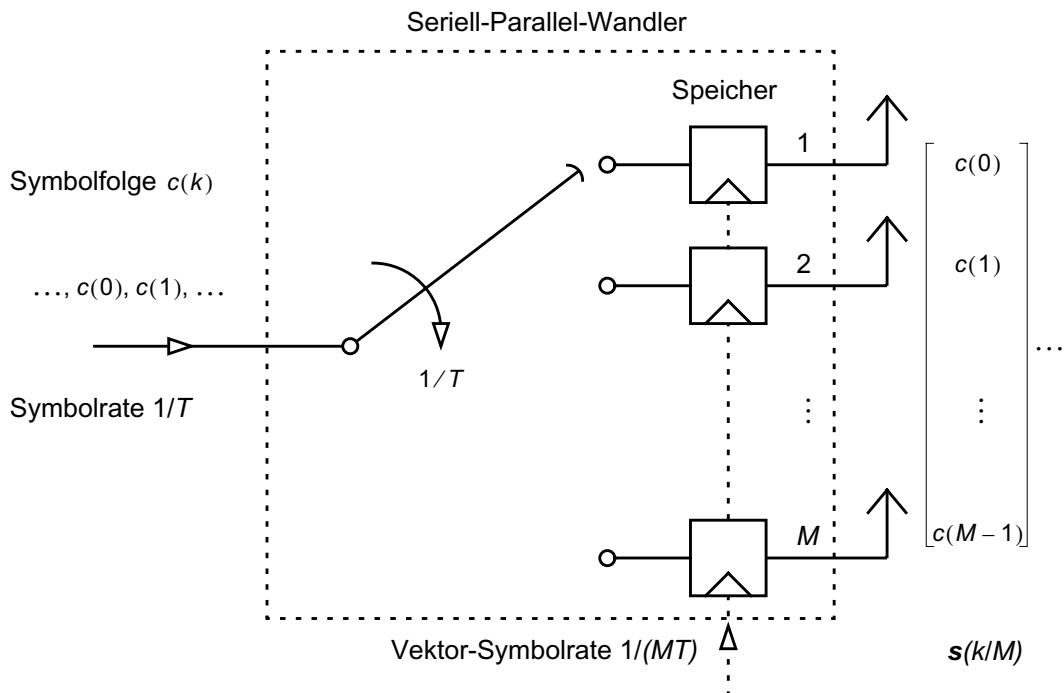


Bild 6.1 Seriell-Parallel-Wandler für räumliches Multiplexen

6.2 Vertikale Codierung (V-BLAST)

Wir nehmen nun zum räumlichen Demultiplexen noch eine zeitliche Codierung hinzu. Das räumliche Demultiplexen wird bereits schon als Raum-Zeit-Codierung bezeichnet, auch ohne zusätzliche zeitliche Codierung. Diese Anordnung wurde zuerst in den Bell Labs untersucht und als Vertical Bell Labs Layered Architecture for Space-Time Coding (V-BLAST) bezeichnet [6]. In der Literatur hat sich auch die kürzere Bezeichnung Vertical Encoding (VE) oder Vertical Layered Space-Time (VLST) verbreitet. Bild 6.2 zeigt die Anordnung. Die von der Datenquelle kommende Bitfolge $b(n)$ wird (zeitlich) codiert, z. B. mit einem Faltungscodierer der Coderate r_t . Im Interleaver (Verschränker) wird die Reihenfolge der Bits verändert, um die Bitfehlerwahrscheinlichkeit bei Büschelfehlern (Bursts) zu reduzieren. Die Coderate des Interleavers ist $r_t = 1$. Der QAM-Mapper ordnet q aufeinanderfolgende Bits einem Symbol c zu. Der Demultiplexer (S/P-Wandler) verteilt die

Symbole $c(k)$ auf die M Antennen. Im Zeitintervall MT tritt ein Symbolvektor am Ausgang des S/P-Wandlers auf. In dieser Zeit laufen M Symbole in den S/P-Wandler, $q M$ Bits in den Mapper und $q M r_t$ Bits in den zeitlichen Codierer ein. Die Coderate des gesamten Senders ist also

$$r_{\text{total}} = q M r_t / 1 = r_t q M \text{ mit } r_s = M \quad (6.2)$$

Besitzt der Empfänger N Antennen, dann liegt die Diversitätsordnung von V-BLAST zwischen N und MN . Wie man aus Bild 6.2 erkennt, verteilt sich die codierte Bitfolge $b'(n')$ in Form der Symbolfolge $c(k)$ über die M Sendeantennen, zeichnerisch in vertikaler Richtung, daher auch der Name vertikale Codierung. Für 2-PSK ($q = 1$) ohne zeitliche Codierung ($r_t = 1$) wird bei V-BLAST $r_{\text{total}} = M$, wie beim einfachen spatialen Demultiplexer in Bild 6.1.

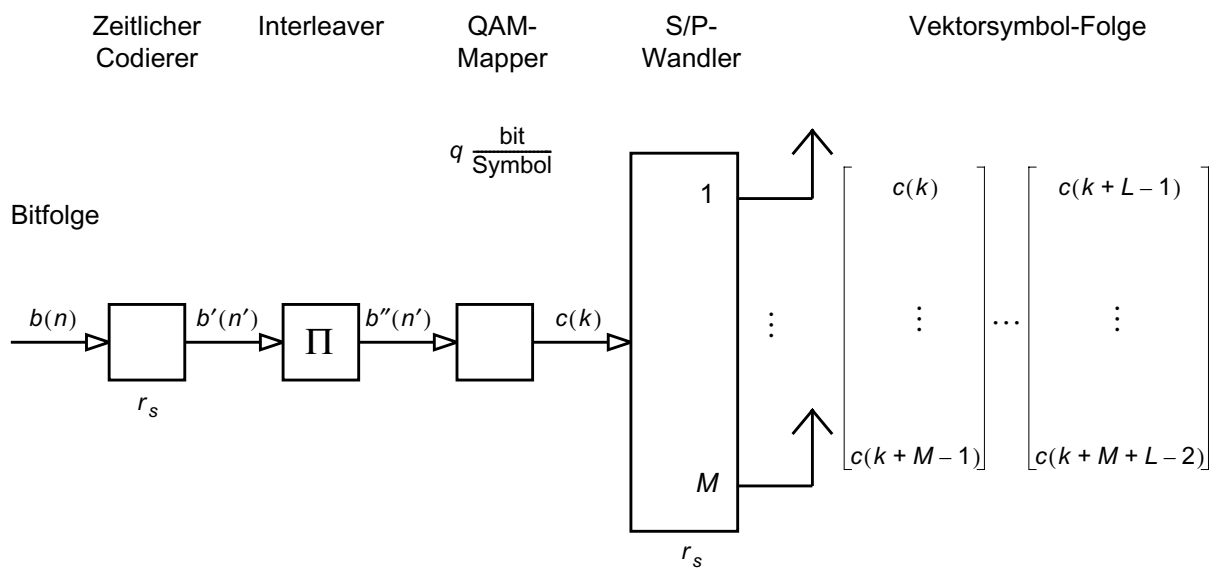


Bild 6.2 MIMO-Sender mit vertikaler Codierung (Vertical Bell Labs Layered Architecture for Space-Time Coding, V-BLAST)

Bild 6.3 zeigt das Prinzip eines Empfängers für V-BLAST. Die erste Stufe ist ein ZF- oder MMSE-Empfänger nach Abschnitt 3. Die M Komponenten $y_1(k), \dots, y_M(k)$ des Ausgangssignalvektors laufen in einem Parallel-Seriell-Wandler (P/S-Wandler) ein und bilden Schätzwerte $\hat{c}(k)$ der Sendesymbole $c(k)$. Demapper und Entscheider liefern Schätzwerte $\hat{b}''(n')$ der Sendefolge $b''(n')$. Der Deinterleaver führt die inverse Operation des Interleavers aus und liefert die Folge der Schätzwerte $\hat{b}'(n')$. Der zeitliche Decodierer, der im Fall einer Faltungscodierung meist als Maximum-Likelihood-Sequenz-Schätzer mit Viterbi-Algorithmus ausgeführt wird, liefert schließlich Schätzwerte $\hat{b}(n)$ der Sendebitfolge $b(n)$.

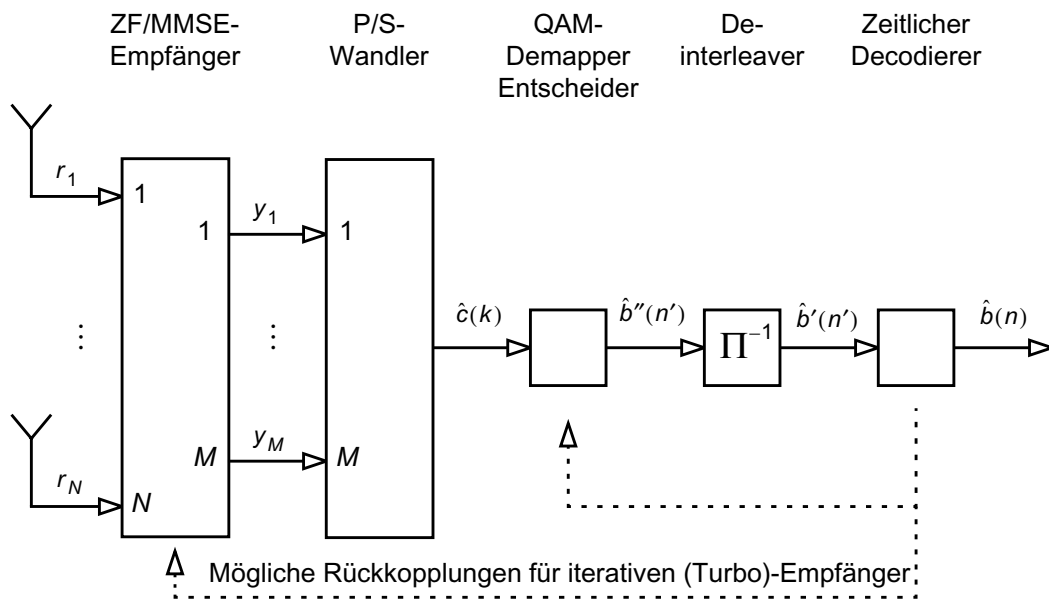


Bild 6.3 MIMO-Empfänger für Sender mit vertikaler Codierung (V-BLAST)

In Bild 6.4 ist als Ergebnis einer Computersimulation die Bitfehlerhäufigkeit (Bit Error Ratio, BER) als Funktion des Signal-Rauschverhältnisses SNR dargestellt [28]. Auch für die folgenden Bilder wird vorausgesetzt, dass alle M Antennen mit gleicher Leistung senden. Der MIMO-Kanal wird durch Rayleigh-Fading und AWGN gestört. Ein Zero-Forcing (ZF)-Empfänger wird verwendet. Das Modulationsverfahren ist 4-PSK mit Gray-Mapping. Verschiedene $M \times N$ -MIMO-Systeme mit und ohne zeitliche Faltungscodierung (Coderate $r_t = 1/2$) werden verglichen. Man erkennt in allen Fällen, dass die zeitliche Codierung eine wesentliche Verringerung des SNR liefert. Zunächst etwas überraschend zeigt sich auch, dass eine Verringerung der Zahl M der Sendeantennen bei gleicher Empfangsantennenzahl N die Ergebnisse verbessert. Der Grund liegt in der gesamten Sendeleistung, die in allen Fällen, auch in den folgenden Bildern immer konstant ist. Wird M von 4 auf 2 verringert, verdoppelt sich die Sendeleistung pro Antenne, was offensichtlich die BER reduziert.

Bild 6.5 zeigt die Bitfehlerhäufigkeit BER als Funktion des SNR für ein MIMO-System mit $M = 6$ Sende- und $N = 6$ Empfangsantennen, 16-QAM mit Gray-Mapping mit und ohne zeitliche Faltungscodierung. Verglichen wird die Leistungsfähigkeit von ZF- und MMSE-Empfänger. Wie man sieht, schneidet der MMSE-Empfänger besser ab, da, wie in Kapitel 3 erwähnt, bei seinem Entwurf das Kanalrauschen berücksichtigt wird. Es zeigt sich auch sehr deutlich, dass die zeitliche Faltungscodierung erhebliche Verbesserungen bringt. Daher wird man MIMO-Systeme nur mit zusätzlicher zeitlicher Kanalcodierung betreiben.

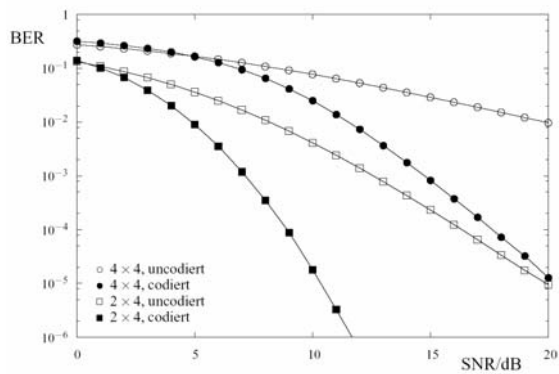


Bild 6.4 Bitfehlerhäufigkeit (BER) als Funktion des Signal-Rauschverhältnisses SNR für V-BLAST mit Zero-Forcing (ZF)-Empfänger und unterschiedlicher Zahl von Sende- und Empfangsantennen $M \times N$, 4-PSK mit Gray-Mapping. Mit und ohne zeitlicher Faltungscodierung, Rayleigh-Fading

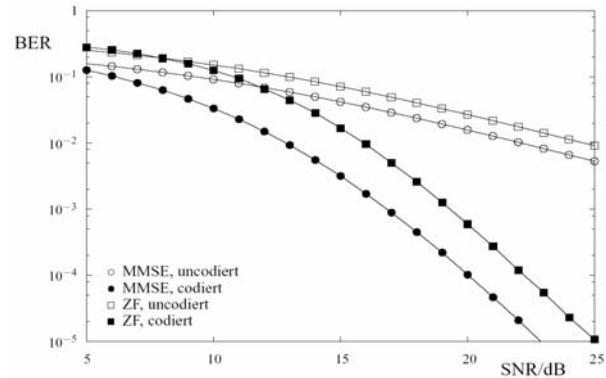


Bild 6.5 Bitfehlerhäufigkeit (BER) als Funktion des Signal-Rauschverhältnisses SNR für 6x6 V-BLAST mit 16-QAM und Gray-Mapping. Vergleich ZF- und MMSE-Empfänger mit und ohne zeitlicher Faltungscodierung, Rayleigh-Fading

In Bild 6.6 ist die Bitfehlerhäufigkeit BER als Funktion des SNR für einen iterativen Empfänger dargestellt [28]. Als Modulationsverfahren wird 16-QAM und Gray-Mapping verwendet. Das MIMO-System arbeitet mit $M = 6$ Sende- und $N = 6$ Empfangsantennen. Die Kurve in Bild 6.6 mit der Kennzeichnung 0. Iteration zeigt das Ergebnis für den nicht-iterativen Empfänger in Bild 6.3. Wie zu erwarten, nimmt BER mit zunehmendem SNR ab, allerdings relativ schwach. Einen steileren Abfall kann man durch einen iterativen Empfänger, der nach dem Turbo-Prinzip [29]-[35] arbeitet, erreichen. U. a. in [28] und [36]-[41] wurde dieses Prinzip auf MIMO-Empfänger in Bild 6.3 angewandt. Dabei wird in einer Rückkopplungsschleife, die in Bild 6.3 gestrichelt angedeutet ist, extrinsische Information, die bei einem zeitlichen Soft-In-Soft-Out-Decoder anfällt, als A-Priori-Information an die vorangehenden Stufen wie QAM-Demapper und ZF/MMSE-Empfänger zurückgegeben. Zur Decodierung eines Bits kann diese Schleife mehrmals durchlaufen werden (Iteration). Bild 6.6 zeigt die beeindruckende Verbesserung nach 5 und nach 9 Iterationen. Die Kurven fallen wie gewünscht bei verringertem SNR steil ab (Turbo Cliff).

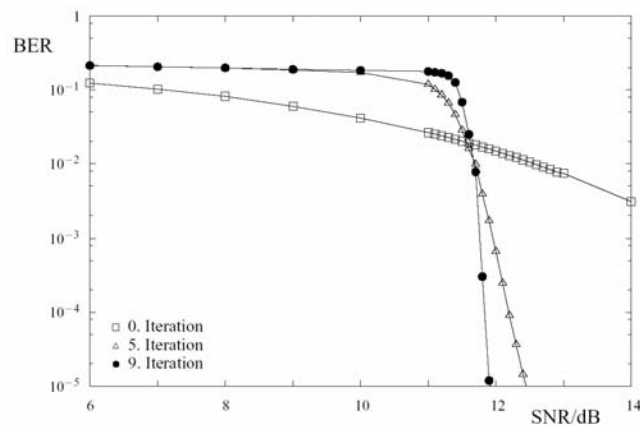


Bild 6.6 Bitfehlerhäufigkeit (BER) als Funktion des Signal-Rauschverhältnisses SNR für V-BLAST mit $M=6$ Sende- und $N=6$ Empfangsantennen, 16-QAM Gray-Mapping, Rayleigh-Fading. Empfänger mit Iteration über den Demapper und die Symbol-Auslöschung (s. Kapitel 8)

Das Ausführungsbeispiel eines iterativen V-BLAST-Empfängers zeigt Bild 6.7 aus [28]. Die Iteration wird dabei über den Demapper geführt, der nunmehr ein iterativer oder Soft-Demapper ist [34], [35] und nicht mehr nur eine einfache Zuordnungstabelle verwirklicht.

In Bild 6.6 hat es den Anschein, als ob die Kurven bei Vergrößerung des SNR weiter stark abfallen. Dies ist bei iterativen Empfängern, auch im Fall SISO ($M = N = 1$) meist nicht der Fall. Vielmehr biegen die Kurven nach rechts ab und verlaufen dann wesentlich flacher. Es bildet sich also eine Bitfehlerschulter (bit error floor) aus. Sehr kleine BER, z. B. 10^{-8} werden erst bei großen SNR erreicht. Dies ist ein Nachteil von Turbo-Codes und iterativen Decodierungsverfahren. Es gibt verschiedene Entwurfsverfahren, z. B. [42], [43], um diesen Effekt zu mildern. Im besonderen ist ein weiterer zeitlicher Code als „äußerer Code“ sehr wirksam, der allerdings die gesamte Coderate r_{total} verkleinert und damit die spektrale Effizienz (bit/s in 1 Hz Bandbreite) etwas abschwächt.

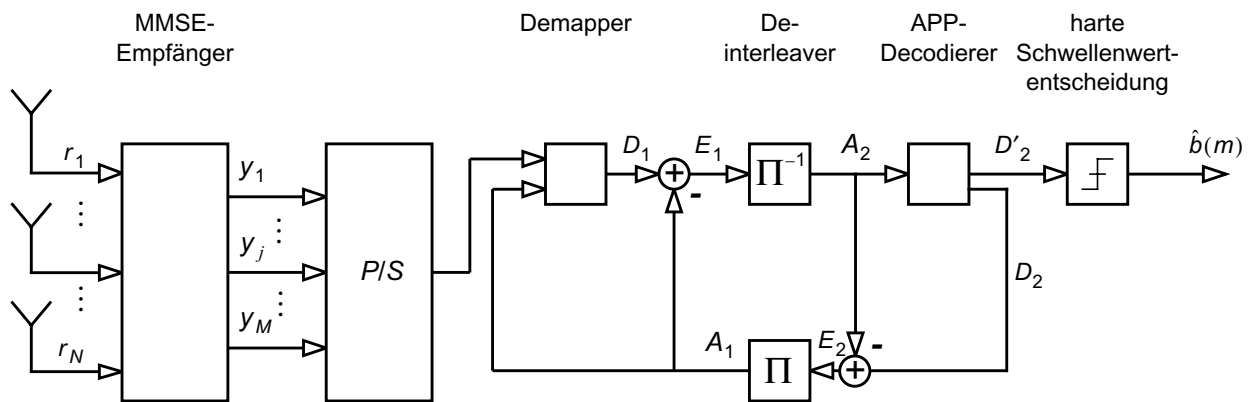


Bild 6.7 V-BLAST-Empfänger mit Iteration über den Demapper

6.3 Horizontale Codierung (H-BLAST)

Bild 6.8 zeigt den MIMO-Sender für das Verfahren „Horizontale Codierung“, die auch als Horizontal Bell Labs Layered Architecture for Space-Time Coding (H-BLAST) bezeichnet wird. Die Bitfolge $b(n)$ der Datenquelle wird im S/P-Wandler in M parallele Zweige, entsprechend der Zahl M der Sendeantennen, aufgespalten. Die Bitrate reduziert sich um Faktor M . In jedem Zweig werden eine zeitliche Codierung (meist eine Faltungscodierung) und ein Interleaving durchgeführt. Der QAM-Mapper ordnet je q aufeinanderfolgenden Bits $b''_v(n)$ ein komplexes Symbol $c_v(k)$ zu, das in Antenne v eingespeist wird ($v = 1, \dots, M$). Die Raum-Zeit-Codierung ist durch die S/P-Wandlung bereits erledigt worden. Die Antennen geben die Symbolvektoren mit der Taktfrequenz $1/(TM)$ ab, wobei $1/T$ die Rate der Symbole $c_v(k)$ ist. Man überlegt sich leicht, dass die Coderate des gesamten Senders

$$r_{\text{total}} = r_t q M \quad (6.3)$$

ist, also mit der des V-BLAST-Systems in (6.2) übereinstimmt. Aus Bild 6.8 erkennt man deutlich, dass die Codierung in den parallelen „horizontalen“ M Zweigen unabhängig voneinander ist, weshalb der Name horizontale Codierung gewählt wurde.

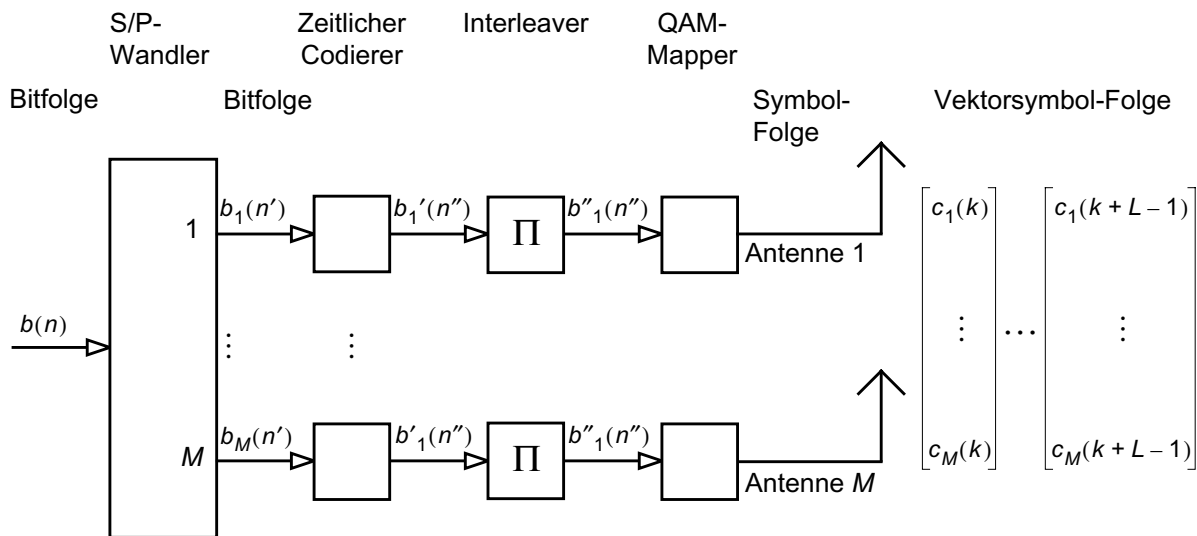


Bild 6.8 MIMO-Sender mit horizontaler Codierung (Horizontal Bell Labs Layered Architecture for Space-Time Coding, H-BLAST)

Bild 6.9 zeigt eine Möglichkeit eines V-BLAST-Empfängers. Ein ZF- oder MMSE-Empfänger ermittelt aus dem Empfangsvektor $\mathbf{r}(k)$ die Signalkomponenten $y_1(k), \dots, y_M(k)$. In M parallelen Zweigen werden dann die in Bild 6.9 dargestellten, uns bereits bekannten Operationen ausgeführt. Der Parallel-Seriell-Wandler fügt die Teilfolgen zu einer Bitfolge $\hat{b}(n)$ zusammen, einem Schätzwert zur Sendefolge $b(n)$.

Man erkennt, dass die im V- und H-BLAST-System auftretenden Operationen sehr ähnlich sind. H-BLAST besitzt parallele Zweige mit reduzierter Taktfrequenz, V-BLAST nur einen Zweig mit entsprechend höherer Taktfrequenz. Da in beiden Systemen die Struktur der räumlich-zeitlichen Codierung und Decodierung unterschiedlich ist, zeigen beide auch nicht die gleichen Eigenschaften, insbesondere BER als Funktion des SNR. Zur Vertiefung wird auf die Literatur, z. B. [44]-[46] verwiesen. Da die von den M Antennen abgestrahlten Signale unabhängig voneinander codiert wurden, liefert der Sender keine Sendediversität. Im MIMO-System verbleibt also nur noch die Ordnung der Empfangsdiversität, die höchstens N sein kann. Damit besitzt H-BLAST eine kleinere Diversitätsordnung als V-BLAST.

Neben der Ordnung der Diversität ist bei einem MIMO-System noch der Codiergewinn (coding gain) wichtig. Er hängt von der Leistungsfähigkeit des zeitlichen Codierverfahrens ab. In der Literatur wird noch eine weitere MIMO-Architektur, die „Diagonale Codierung“ (D-BLAST), aufgezeigt. Hierzu wird z. B. auf [44] verwiesen.

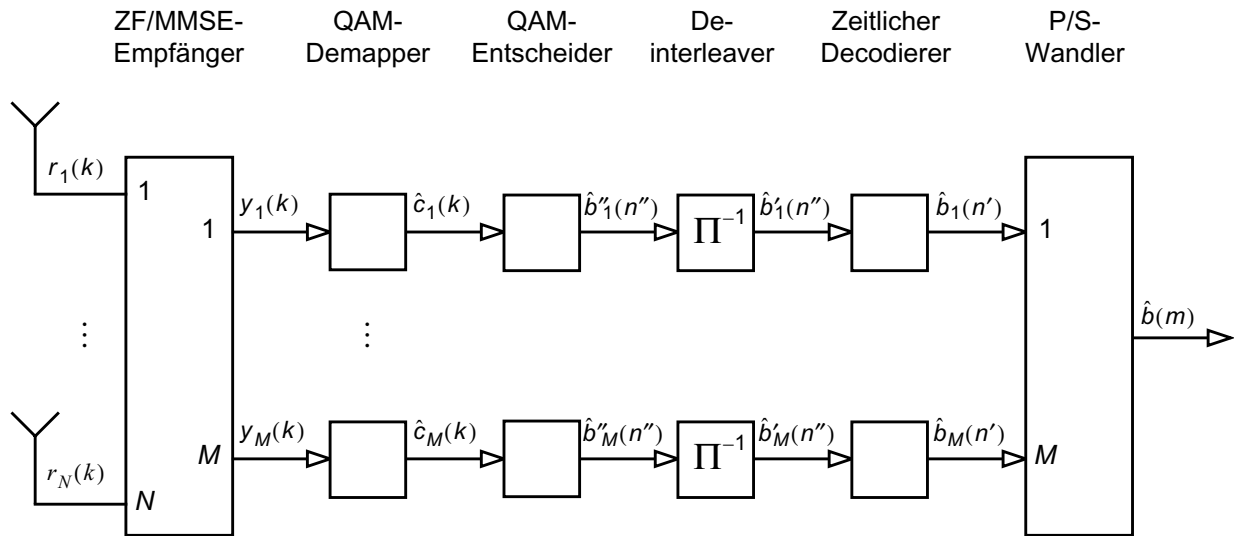


Bild 6.9 MIMO-Empfänger für Sender mit horizontaler Codierung (H-BLAST)

7. Berechnung der Symbolfehlerwahrscheinlichkeit

In diesem Kapitel soll die Berechnung der Symbolfehlerwahrscheinlichkeit P_S in groben Zügen dargestellt werden. Enthält das MIMO-System einen zeitlichen Codierer, dann sind mathematisch-analytische Lösungen kaum mehr möglich. Vielmehr ist man auf die Computersimulation und asymptotische Abschätzungen angewiesen. Um eine prinzipielle, analytische Einsicht in die Zusammenhänge zu erhalten, konzentrieren wir uns hier auf die Symbolfehlerwahrscheinlichkeit und nehmen an, dass eine zeitliche Codierung nicht vorhanden ist.

Für das Empfangssignal erhalten wir nach (2.11)

$$\mathbf{r} = \mathbf{H} \cdot \mathbf{s} + \mathbf{n} \quad (7.1)$$

Alle QAM-Signalpunkte sollen gleiche Auftretswahrscheinlichkeiten besitzen. Zunächst wird noch der allgemeine Fall $M \geq 1$ betrachtet. Für das Rauschen \mathbf{n} gelten die früheren Voraussetzungen in Abschnitt 2.3.2, insbesondere

$$\mathbf{R}_{nn} = \sigma_n^2 \mathbf{I}_N$$

Der Signalanteil in (7.1) ist

$$\mathbf{z} = \mathbf{H} \mathbf{s} \quad (7.2)$$

mit der Autokorrelationsmatrix

$$\mathbf{R}_{zz} = E[\mathbf{z} \mathbf{z}^H] = \mathbf{H} \mathbf{R}_{ss} \mathbf{H}^H \quad (7.3)$$

und mit $\mathbf{R}_{ss} = E_S \cdot \mathbf{I}_M$ folgt

$$\mathbf{R}_{zz} = E_S \mathbf{H} \mathbf{H}^H \quad (7.4)$$

Für die mittlere Leistung des Signals z erhält man mit (7.4)

$$E[z^H \cdot z] = \text{spur}(\mathbf{R}_{zz}) = E_S \sum_{v=1}^M \sum_{\mu=1}^N |h_{\mu v}|^2 \quad (7.5)$$

Die mittlere Leistung von räumlich unkorreliertem Rauschen \mathbf{n} ist

$$E[\mathbf{n}^H \mathbf{n}] = N \sigma_n^2 \quad (7.6)$$

Aus (7.5) und (7.6) erhält man damit das Signal-Rauschverhältnis beim Empfänger

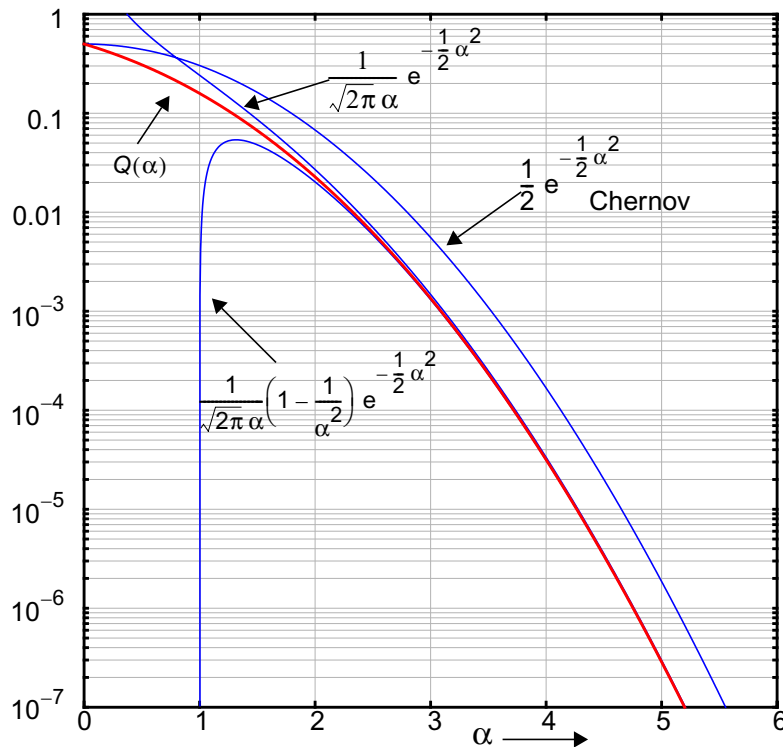
$$\gamma = E[z^H z] / E[\mathbf{n}^H \mathbf{n}] = \frac{E_S}{N \sigma_n^2} \sum_{v=1}^M \sum_{\mu=1}^N |h_{\mu v}|^2 \quad (7.7)$$

und speziell für den Fall $M = 1$, der zunächst weiter verfolgt werden soll

$$\gamma = \frac{E_S}{N \sigma_n^2} \sum_{\mu=1}^N |h_{\mu 1}|^2 \quad (7.8)$$

Bei Maximum-Likelihood-Detektion und QAM gilt bekanntlich für die Symbolfehlerwahrscheinlichkeit P_S näherungsweise [47]

$$P_S \approx K_1 Q(\sqrt{K_2} \gamma) \quad (7.9)$$



$$Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\infty} e^{-\frac{1}{2}u^2} du$$

Bild 7.1 Q-Funktion und Näherungen

Dabei ist $Q(\cdot)$ die bekannte Q-Funktion nach Bild 7.1. K_1 und K_2 hängen von der QAM-Konstellation ab. Insbesondere ist K_2 proportional zum dichtesten Abstand zweier QAM-Signalpunkte. Im Beispiel eines SISO-Systems mit $h_{11} = 1$ und 4-QAM mit den 4 Signalpunkten $(\pm b \pm jb)/\sqrt{2}$ ist bekanntlich $P_S = 2Q(b/\sigma_n) - Q^2(b/\sigma_n) \approx 2Q(b/\sigma_n)$ für $b/\sigma_n \gg 1$.

Mit (7.8) folgt dann aus (7.9) näherungsweise die Symbolfehlerwahrscheinlichkeit

$$P_S \approx K_1 Q\left(\left[\frac{K_2 E_S}{N \sigma_n^2} \sum_{\mu=1}^N |h_{\mu 1}|^2\right]^{1/2}\right) \quad (7.10)$$

Ist der MIMO-Kanal zeitinvariant, dann sind die $h_{\mu 1}$ und damit P_S fest. Bei dem meist vorhandenen Fading-Kanal schwankt P_S , da $h_{\mu 1} = h_{\mu 1}(k)$ ($\mu = 1, \dots, N$) ist. Dann kann man den Erwartungswert der Symbolfehlerwahrscheinlichkeit $\bar{P}_S = E[P_S]$ als Maß wählen. Ist die Wahrscheinlichkeitsdichtefunktion $p_x(x)$ für die Zufallsvariable

$$x = \sum_{\mu=1}^N \sum_{v=1}^M |h_{\mu v}|^2 \quad (\geq 0) \quad (7.11)$$

bekannt, dann erhält man aus (7.10)

$$\bar{P}_S = E[P_S] = K_1 \int_0^{\infty} Q\left(\left[\frac{K_2 E_S}{N \sigma_n^2} x\right]^{1/2}\right) p_x(x) dx \quad (7.12)$$

mit $M = 1$ für das SIMO-System.

Für das Kanalmodell mit $\mathbf{H}_{\mathcal{Y}}$ nach (2.17), bei dem die $h_{\mu v}$ unabhängig und identisch komplex gaußverteilt sind mit $E[|h_{\mu v}|^2] = 1$ und $E[h_{\mu v}] = 0$, kann man zeigen, dass x eine Chi-Quadrat-Dichte mit $2 MN$ Freiheitsgraden besitzt

$$p_x(x) = \begin{cases} \frac{x^{MN-1}}{(MN-1)!} e^{-x} & x \geq 0 \\ 0 & \text{sonst} \end{cases} \quad (7.13)$$

Damit lässt sich dann das Integral in (7.12) ermitteln. Daraus kann man folgende asymptotische Abschätzung für großes Signal-Rauschverhältnis gewinnen [44]:

$$\bar{P}_S \leq K_3 \frac{1}{\left(\frac{E_S}{N \sigma_n^2}\right)^N} \quad \text{für} \quad \frac{E_S}{\sigma_n^2} \gg 1 \quad (7.14)$$

wobei K_3 eine feste Konstante ist. Bildet man den Logarithmus, dann folgt

$$\log \bar{P}_S \leq \log K_3 - N \log \frac{E_S}{\sigma_n^2} + N \log N \quad (7.15)$$

und erhält die Geraden in Bild 7.2, die umso steiler verlaufen, je größer N ist. Die Diversitätsordnung drückt sich also in einer steilen Abnahme von \bar{P}_S aus.

Eine Verallgemeinerung auf MIMO-Systeme mit $M > 1$ ist möglich [44]. Die Steigung der Kurven in Bild 7.2 ist dann näherungsweise durch $-MN$ gegeben, d. h. die Sende- und Empfangsdiversität. In [19] und [48]-[50] konnten mit mathematisch anspruchsvollen Verfahren erstmalig analytische Beziehungen für die Symbolfehlerwahrscheinlichkeit, besonders auch asymptotische Näherungen, angegeben werden.

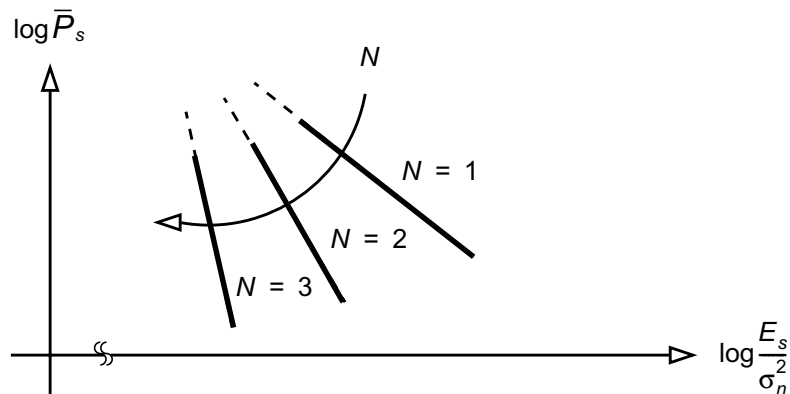


Bild 7.2 Asymptotischer Verlauf der mittleren Symbolfehlerwahrscheinlichkeit für große Signal-Rauschverhältnisse und unterschiedlicher Zahl N von Empfangsantennen

8. Zero-Forcing-Empfänger mit schrittweiser Interferenzauslöschung

Nachdem in den vorangegangenen Kapiteln verschiedene Empfangs- und Entscheidungsprinzipien behandelt wurden, soll in diesem letzten Kapitel abschließend noch ein Verfahren vorgestellt werden, bei dem Zero-Forcing-Empfänger und Entscheider nicht nacheinander, sondern ineinander verwoben arbeiten. Es wurde in [51] und [52] mit V-BLAST verwirklicht und abgekürzt als OSIC (Ordered Successive Interference Cancellation). Es ist vom Grundsatz her nichts anderes als das Gaußsche Eliminationsverfahren, bei dem man eine ermittelte Unbekannte in das Gleichungssystem einsetzt, und so schrittweise die Zahl der Unbekannten reduziert.

Wir betrachten das MIMO-Übertragungssystem in Bild 2.7. \mathbf{W} ist die Matrix des Empfängers, d. h. die Pseudoinverse nach (3.13) von \mathbf{H} oder \mathbf{W} nach (3.28). Daran schließt sich bekanntlich eine symbolweise Entscheidung an, d. h. die Operation

$$\hat{s}_v = q(y_v) \quad v = 1, \dots, M \quad (8.1)$$

Dabei bezeichnet $q(\cdot)$ die Entscheidungsregel. Sind z. B. die Quellensymbole s_v binär, d. h. $s_v \in \{-1, 1\}$, $v = 1, \dots, M$, dann wird

$$q(y_v) = \text{sign } y_v = \begin{cases} 1 & , y_v \geq 0 \\ -1 & , y_v < 0 \end{cases} \quad v = 1, \dots, M \quad (8.2)$$

gewählt. \hat{s}_v ist also ein Schätzwert des Sendesymbols s_v , der aus dem verzerrten und verrauschten

Wert y_v nach (8.1) gewonnen wird. Das OSIC-Verfahren stellt man am besten anhand eines Beispiels dar. Wir wählen $M = N = 3$. Dann erhält man aus Bild 2.7 das Gleichungssystem

$$\mathbf{H} \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} + \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \quad (8.3)$$

$\underbrace{\hspace{1.5cm}}_{\mathbf{h}_1} \quad \underbrace{\hspace{1.5cm}}_{\mathbf{h}_2} \quad \underbrace{\hspace{1.5cm}}_{\mathbf{h}_3} \quad \underbrace{\hspace{1.5cm}}_{\mathbf{s}} \quad \underbrace{\hspace{1.5cm}}_{\mathbf{n}} \quad \underbrace{\hspace{1.5cm}}_{\mathbf{r}}$

Ferner $\mathbf{y} = \mathbf{W} \mathbf{r}$ (8.4)

Wir nehmen an, dass \mathbf{H} und \mathbf{r} bekannt sind. s_1, \dots, s_3 sind gesucht. Bevor der OSIC-Algorithmus vorgestellt wird, noch eine wichtige Umformung von (8.3), die für das Verständnis wichtig ist. Angenommen, s_1 sei bekannt, dann kann man (8.3) auf die Form bringen

$$(\mathbf{h}_2 \ \mathbf{h}_3) \begin{pmatrix} s_2 \\ s_3 \end{pmatrix} + \mathbf{n} = \mathbf{r} - \mathbf{h}_1 s_1 \quad (8.5)$$

Im Weiteren ist $\mathbf{W}^{(i)}$ die ZF-Matrix im Verfahrensschritt $i = 1, 2, \dots$. Ferner werde angenommen, dass nach (8.1) immer richtig entschieden werde, d. h.

$$\hat{s}_v = s_v \quad (8.6)$$

1. Schritt

Sei $\mathbf{H}^{(1)} = \mathbf{H}$, $\mathbf{r}^{(1)} = \mathbf{r}$

Ermittle $\mathbf{W}^{(1)}$ aus $\mathbf{H}^{(1)}$; sei $\mathbf{w}_1^{(1)}$ erster Zeilenvektor von $\mathbf{W}^{(1)}$

Aus (8.4): $y_1 = \mathbf{w}_1^{(1)} \cdot \mathbf{r}^{(1)}$

Aus (8.1): $\hat{s}_1 = q(y_1)$

Aus (8.3) bzw. (8.5): $(\mathbf{h}_2 \ \mathbf{h}_3) \begin{pmatrix} s_2 \\ s_3 \end{pmatrix} + \mathbf{n} = \underbrace{\mathbf{r}^{(1)} - \mathbf{h}_1 \hat{s}_1}_{\mathbf{r}^{(2)}} \quad (8.7)$

2. Schritt

Sei $\mathbf{H}^{(2)} = (\mathbf{h}_2 \ \mathbf{h}_3)$ in (8.7)

Ermittle $\mathbf{W}^{(2)}$ aus $\mathbf{H}^{(2)}$; sei $\mathbf{w}_1^{(2)}$ erster Zeilenvektor von $\mathbf{W}^{(2)}$

Aus (8.4): $y_2 = \mathbf{w}_1^{(2)} \cdot \mathbf{r}^{(2)}$

Aus (8.1): $\hat{s}_2 = q(y_2)$

Aus (8.7): $\mathbf{h}_3 s_3 + \mathbf{n} = \underbrace{\mathbf{r}^{(2)} - \mathbf{h}_2 \hat{s}_2}_{\mathbf{r}^{(3)}} \quad (8.8)$

3. Schritt

Sei $\mathbf{H}^{(3)} = \mathbf{h}_3$ in (8.8)

Ermittle $\mathbf{W}^{(3)}$ aus $\mathbf{H}^{(3)}$; sei $\mathbf{w}_1^{(3)}$ erster Zeilenvektor von $\mathbf{W}^{(3)}$

Aus (8.4): $y_3 = \mathbf{w}_1^{(3)} \cdot \mathbf{r}^{(3)}$

Aus (8.1): $\hat{s}_3 = q(y_3)$

In jedem Schritt wird ein Symbol \hat{s}_v berechnet und durch Anwendung von (8.5) das Gleichungssystem um eine Unbekannte verringert. Man sagt auch, jeweils ein Symbol werde „gelöscht“ und bezeichnet das Verfahren daher oft als schrittweise Symbol-Auslöschung (SAL). Bild 8.1 zeigt das Ergebnis einer Computersimulation für ein V-BLAST-System mit $M = 4$ Sende- und $N = 4$ Empfangsantennen, Rayleigh-Fading und 4-PSK mit Gray-Mapping. Parallele Detektion bezeichnet die Anwendung der ZF-Matrix \mathbf{W} ohne SAL. Man erkennt, dass SAL im Fall ohne zeitliche Codierung („uncodiert“ in Bild 8.1) eine deutliche Verbesserung liefern kann. Die Hinzunahme von SAL zur zeitlichen Codierung kann das Ergebnis verschlechtern, wenn die Entscheidungsmechanismen auf Symbolebene (SAL) und auf Bitebene (zeitliche Codierung) nicht sorgfältig aufeinander abgestimmt werden. Dieser Fragestellung wird in [28] nachgegangen. Darüber hinaus wird dort auch der Einfluss der Fehlerfortpflanzung von Schritt zu Schritt beim OSIC-Verfahren analysiert. Sie entsteht, wenn die Annahme (8.6) verletzt wird, ein in der Realität oft auftretender Fall. Darauf aufbauend werden in [28], [38]-[42] wesentlich verbesserte Verfahren vorgeschlagen, die prüfen, mit welcher Zuverlässigkeit in Form von Likelihood-Werten die Bedingung (8.6) erfüllt ist, und dann eine SAL durchführen oder nicht. Darüber hinaus kann man bei Einsatz einer zeitlichen Codierung eine iterative Detektion einsetzen, wie in Abschnitt 6.2 dargestellt. Als besonders leistungsfähig bei vertretbarem Realisierungsaufwand hat sich die Kombination von OSIC und iterativer Detektion (Iteration über MMSE-Empfänger und Soft-Demapper) erwiesen [28]. Wesentliche Ergebnisse wurden im Abschnitt 6.2 (Bild 6.6) bereits vorgestellt.

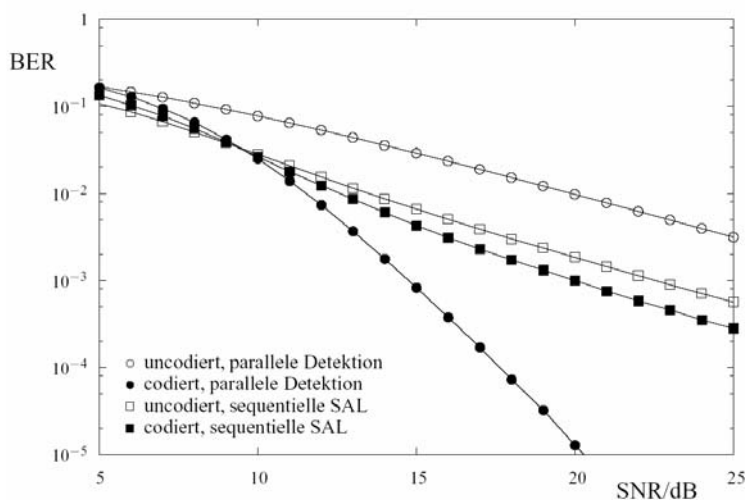


Bild 8.1 Bitfehlerhäufigkeit (BER) als Funktion des Signal-Rauschverhältnisses für ein V-BLAST-ZF-System mit 4x4 Antennen, 4-PSK, Gray Mapping und Rayleigh-Fading, mit und ohne zeitlicher Faltungscodierung und Symbol-Auslöschung (SAL)

9. Anhang

9.1 Wichtige Beziehungen der linearen Algebra

Sei \mathbf{Q} eine Hermitesche $N \times N$ Matrix mit Eigenwerten λ_ν ($\nu = 1, \dots, N$), d. h.

$$\mathbf{Q} = \mathbf{Q}^H \quad (9.1.1)$$

A) $\mathbf{z}^H \mathbf{Q} \mathbf{z}$ ist eine quadratische Form mit

$$\mathbf{z}^H \mathbf{Q} \mathbf{z} \geq 0, \text{ alle } \mathbf{z} \neq \mathbf{0} \quad (9.1.2)$$

$\mathbf{Q} = \mathbf{a}\mathbf{a}^H$ mit $\mathbf{a} \in \mathbb{C}^{N \times 1}$ ein Spaltenvektor.

Beweis: $\mathbf{z}^H \mathbf{Q} \mathbf{z} = \mathbf{z}^H \mathbf{a}\mathbf{a}^H \mathbf{z} = \mathbf{z}^H \mathbf{a} (\mathbf{z}^H \cdot \mathbf{a})^H = \|\mathbf{z}^H \mathbf{a}\|^2 \geq 0$ q. e. d.

B) Eine Matrix \mathbf{Q} mit Eigenschaft (9.1.2) heißt positiv-semidefinit. Folgende Äquivalenz gilt:

$$\mathbf{Q} \text{ mit (9.1.2)} \Leftrightarrow \lambda_\nu \geq 0 \quad (\nu = 1, \dots, N) \quad (9.1.3)$$

d. h. insbesondere, alle λ_ν sind reell.

Beweis: Sei \mathbf{u}_ν der Eigenvektor zu λ_ν , d. h.

$$\mathbf{Q} \mathbf{u}_\nu = \lambda_\nu \mathbf{u}_\nu \quad (9.1.4)$$

Wir berechnen

$$\begin{aligned} \mathbf{u}_\nu^H \mathbf{Q} \mathbf{u}_\nu &= \mathbf{u}_\nu^H \lambda_\nu \mathbf{u}_\nu = \lambda_\nu \mathbf{u}_\nu^H \mathbf{u}_\nu = \lambda_\nu \|\mathbf{u}_\nu\|^2 \geq 0 \\ \Rightarrow \lambda_\nu &\geq 0 \quad \text{q. e. d.} \end{aligned}$$

C) Seien \mathbf{u}_1 und \mathbf{u}_2 Eigenvektoren zu den Eigenwerten λ_1 und λ_2 von \mathbf{Q} und $\lambda_1 \neq \lambda_2 \neq 0$.

$$\mathbf{u}_1 \text{ und } \mathbf{u}_2 \text{ sind orthogonal, d. h. } \mathbf{u}_1^H \cdot \mathbf{u}_2 = 0 \quad (9.1.5)$$

Beweis: Wir benutzen die Definition der Eigenvektoren

$$\mathbf{Q} \mathbf{u}_1 = \lambda_1 \mathbf{u}_1 \quad (9.1.6)$$

$$\mathbf{Q} \mathbf{u}_2 = \lambda_2 \mathbf{u}_2 \quad (9.1.7)$$

$$(\lambda_1 \mathbf{u}_1)^H \mathbf{u}_2 = (\mathbf{Q} \mathbf{u}_1)^H \mathbf{u}_2 = \mathbf{u}_1^H \mathbf{Q}^H \mathbf{u}_2 = \mathbf{u}_1^H \mathbf{Q} \mathbf{u}_2 = \mathbf{u}_1^H \cdot \mathbf{u}_2 \cdot \lambda_2$$

$$\lambda_1 \mathbf{u}_1^H \mathbf{u}_2 = \lambda_2 \mathbf{u}_1^H \cdot \mathbf{u}_2$$

da $\lambda_1 \neq 0$ und $\lambda_2 \neq 0$, folgt $\mathbf{u}_1^H \mathbf{u}_2 = 0$ q.e.d.

Man beachte, dass (9.1.5) auch für nicht normalisierte Vektoren gilt, d. h.

$$(\alpha_1 \cdot \mathbf{u}_1^H)(\alpha_2 \mathbf{u}_2) = 0 \quad \alpha_1, \alpha_2 \neq 0 \quad (9.1.8)$$

Sind die Eigenvektoren normalisiert, d. h. $\|\mathbf{u}_v\|^2 = 1$, dann heißen sie orthonormal.

Sei \mathbf{U} die Matrix orthonormaler Eigenvektoren

$$\mathbf{U} = (\mathbf{u}_1 \dots \mathbf{u}_N)$$

Dann gilt

$$\mathbf{U}^{-1} = \mathbf{U}^H \quad (9.1.9)$$

\mathbf{U} heißt dann unitäre Matrix.

Eigenwertzerlegung von \mathbf{Q}

$$\mathbf{U}^H \mathbf{Q} \mathbf{U} = \Lambda \quad \text{mit} \quad \Lambda = \text{diag}(\lambda_1 \dots \lambda_N) \quad (9.1.10)$$

Aus (9.1.10) folgt mit (9.1.9)

$$\begin{aligned} \mathbf{U}^H \mathbf{Q} &= \Lambda \mathbf{U}^{-1} \\ \mathbf{Q} &= (\mathbf{U}^H)^{-1} \Lambda \mathbf{U}^{-1} = \mathbf{U} \Lambda \mathbf{U}^H \end{aligned} \quad (9.1.11)$$

9.2 Ermittlung der ZF-Empfängermatrix

Wir gehen vom quadratischen Fehler J in (3.10) aus. Bei der Differentiation muss man berücksichtigen, dass

$$\mathbf{s} = \mathbf{a} + j \mathbf{b} \quad \mathbf{a}, \mathbf{b} \in \mathbb{R}^{M \times 1} \quad (9.2.1)$$

ein Vektor mit komplexen Komponenten ist. Wir bilden nun die Gradientenvektoren, setzen sie zu Null und schreiben symbolisch

$$\frac{\partial J}{\partial \mathbf{a}} = 0 \quad , \quad \frac{\partial J}{\partial \mathbf{b}} = 0 \quad (9.2.2)$$

Aus (3.10) folgt mit (9.2.1)

$$J = (\mathbf{r} - \mathbf{H}(\mathbf{a} + j\mathbf{b}))^H (\mathbf{r} - \mathbf{H}(\mathbf{a} + j\mathbf{b})) \quad (9.2.3)$$

Bei Anwendung der Produktregel erhält man

$$\begin{aligned} \frac{\partial J}{\partial \mathbf{a}} &= -\mathbf{1}^H (\mathbf{r} - \mathbf{H} \mathbf{s}) + [(\mathbf{r}^H - \mathbf{s}^H \mathbf{H}^H)(-\mathbf{H})]^T \\ &= -\mathbf{1}^H \mathbf{r} + \mathbf{H}^H \mathbf{H} \mathbf{s} - \underbrace{(\mathbf{r}^H \mathbf{H})^T}_{(\mathbf{H}^H \mathbf{r})^*} + \underbrace{(\mathbf{s}^H \mathbf{H}^H \mathbf{H})^T}_{(\mathbf{H}^H \mathbf{H} \mathbf{s})^*} = 0 \end{aligned} \quad (9.2.4)$$

Damit

$$\begin{aligned} -2 \operatorname{Re}(\mathbf{H}^H \mathbf{r}) + 2 \operatorname{Re}(\mathbf{H}^H \mathbf{H} \mathbf{s}) &= 0 \\ 2 \operatorname{Re}[-\mathbf{H}^H \mathbf{r} + \mathbf{H}^H \mathbf{H} \mathbf{s}] &= 0 \end{aligned} \quad (9.2.5)$$

(9.2.5) ist im Besonderen erfüllt, wenn $[] = 0$ ist. Daraus folgt

$$\mathbf{s} = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{r} \quad (9.2.6)$$

also die Behauptung (3.12) und (3.13).

Die Auswertung von $\frac{\partial J}{\partial \mathbf{b}} = 0$ verläuft entsprechend und führt ebenfalls auf (9.2.6), es gibt keinen Widerspruch.

9.3 Herleitung der MMSE-Empfängermatrix \mathbf{W}

Gegeben: \mathbf{H} ($N \times M$)

Sendesignal \mathbf{s} mit Autokorrelationsmatrix $\mathbf{R}_{ss} = E[\mathbf{s}\mathbf{s}^H]$

Rauschen \mathbf{n} mit Autokorrelationsmatrix $\mathbf{R}_{nn} = E[\mathbf{n}\mathbf{n}^H]$

\mathbf{s} und \mathbf{n} sind statistisch unabhängig

Fehler $\mathbf{e} = \mathbf{s} - \mathbf{y} = \mathbf{s} - \mathbf{W} \mathbf{r}$; quadratischer Fehler $\|\mathbf{e}\|^2 = \mathbf{e}^H \mathbf{e}$ (9.3.1)

MMSE $J = E[\mathbf{e}^H \cdot \mathbf{e}] = \min \Rightarrow \mathbf{W} = \arg \min J$

Ergebnis $\mathbf{W} = \mathbf{R}_{ss} \mathbf{H}^H (\mathbf{H} \cdot \mathbf{R}_{ss} \mathbf{H}^H + \mathbf{R}_{nn})^{-1}$ (9.3.2)

Beweis von (9.3.2):

Autokorrelationsmatrix von \mathbf{e} $\mathbf{R}_{ee} = E[\mathbf{e}\mathbf{e}^H]$

Dann ist $J = \text{spur} \mathbf{R}_{ee} = \text{spur} E[\mathbf{e}\mathbf{e}^H]$.

Die Spur einer (quadratischen) Matrix ist die Summe der Hauptdiagonal-Elemente. Mit (9.3.1)

$$\begin{aligned} J &= \text{spur} E[(\mathbf{s} - \mathbf{W}\mathbf{r})(\mathbf{s}^H - \mathbf{r}^H \mathbf{W}^H)] = \text{spur} E[\mathbf{s}\mathbf{s}^H - \mathbf{s}\mathbf{r}^H \mathbf{W}^H - \mathbf{W}\mathbf{r}\mathbf{s}^H + \mathbf{W}\mathbf{r}\mathbf{r}^H \mathbf{W}^H] \\ &= \text{spur} \left(\underbrace{E[\mathbf{s}\mathbf{s}^H]}_{\mathbf{R}_{ss}} - \underbrace{E[\mathbf{s}\mathbf{r}^H]}_{\mathbf{R}_{sr}} \mathbf{W}^H - \mathbf{W} \underbrace{E[\mathbf{r}\mathbf{s}^H]}_{\mathbf{R}_{rs}} + \mathbf{W} \underbrace{E[\mathbf{r}\mathbf{r}^H]}_{\mathbf{R}_{rr}} \mathbf{W}^H \right) \end{aligned} \quad (9.3.3)$$

Man überzeugt sich leicht, dass \mathbf{R}_{rr} eine Hermitesche Matrix ist. Dann sind nach (9.1.3) alle Eigenwerte $\lambda_v \geq 0$ ($v = 1, \dots, N$).

Allgemein gilt

$$\text{spur}(\mathbf{C} + \mathbf{D}) = \text{spur} \mathbf{C} + \text{spur} \mathbf{D} \quad (9.3.4)$$

Eigenwertzerlegung von

$$\mathbf{R}_{rr} = \mathbf{U}^H \mathbf{\Lambda} \mathbf{U} = \mathbf{U}^H \mathbf{\Lambda}^{1/2} \mathbf{\Lambda}^{1/2} \mathbf{U} = (\mathbf{\Lambda}^{1/2} \mathbf{U})^H \cdot (\mathbf{\Lambda}^{1/2} \mathbf{U}) = \mathbf{A}^H \cdot \mathbf{A} \quad (9.3.5)$$

mit

$$\mathbf{\Lambda} = \text{diag}(\lambda_1 \dots \lambda_N)$$

Setzt man

$$\begin{aligned}
\mathbf{R}_{rr} &= \mathbf{A}^H \cdot \mathbf{A} \text{ und } \mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I}_M \text{ in (9.3.3) ein, folgt} \\
J &= \text{spur} (\mathbf{R}_{ss} - \mathbf{R}_{sr} \mathbf{A}^{-1} \cdot \mathbf{A} \cdot \mathbf{W}^H - \mathbf{W} \mathbf{A}^H (\mathbf{A}^H)^{-1} \mathbf{R}_{rs} + \mathbf{W} \mathbf{A}^H \mathbf{A} \mathbf{W}^H) \\
&= \text{spur} (\mathbf{R}_{ss} - \mathbf{R}_{sr} \mathbf{A}^{-1} (\mathbf{A}^H)^{-1} \mathbf{R}_{rs} + \mathbf{R}_{sr} \mathbf{A}^{-1} (\mathbf{A}^H)^{-1} \mathbf{R}_{rs} \\
&\quad \mathbf{R}_{sr} \mathbf{A}^{-1} \mathbf{A} \mathbf{W}^H - \mathbf{W} \mathbf{A}^H (\mathbf{A}^H)^{-1} \mathbf{R}_{rs} + \mathbf{W} \mathbf{A}^H \mathbf{A} \mathbf{W}^H)
\end{aligned} \tag{9.3.6}$$

In (9.3.6) wurden die Terme Nummer 2 und 3 künstlich eingefügt und heben sich weg.

Mit $\mathbf{R}_{sr}^H = \mathbf{R}_{rs}$ folgt

$$= \text{spur} (\mathbf{R}_{ss} - \mathbf{R}_{sr} \mathbf{A}^{-1} (\mathbf{A}^H)^{-1} \mathbf{R}_{rs} + [\mathbf{W} \mathbf{A}^H - \mathbf{R}_{sr} \mathbf{A}^{-1}] [\mathbf{W} \mathbf{A}^H - \mathbf{R}_{sr} \mathbf{A}^{-1}]^H) \tag{9.3.7}$$

Da $J \geq 0$ immer gilt, d. h. für alle \mathbf{W} , kann man aus (9.3.7) mit (9.3.4) schließen, dass die Spuren aus den Termen links und rechts des Pluszeichens in (9.3.7) jeweils ≥ 0 sein müssen. Folglich ist die Bedingung für das Minimum von J

$$\begin{aligned}
\text{spur} (\mathbf{W} \mathbf{A}^H - \mathbf{R}_{sr} \mathbf{A}^{-1}) &= 0 \\
\mathbf{W} \mathbf{A}^H &= \mathbf{R}_{sr} \mathbf{A}^{-1} \\
\mathbf{W} &= \mathbf{R}_{sr} \mathbf{A}^{-1} \cdot (\mathbf{A}^H)^{-1} = \mathbf{R}_{sr} (\mathbf{A}^H \mathbf{A})^{-1} = \mathbf{R}_{sr} \mathbf{R}_{rr}^{-1}
\end{aligned} \tag{9.3.8}$$

Für \mathbf{R}_{rr} erhält man mit (2.11)

$$\begin{aligned}
\mathbf{R}_{rr} &= \\
[\mathbf{r} \mathbf{r}^H] &= E[(\mathbf{H} \cdot \mathbf{s} + \mathbf{n})(\mathbf{H} \cdot \mathbf{s} + \mathbf{n})^H] = \mathbf{H} E[\mathbf{s} \mathbf{s}^H] \mathbf{H}^H + \mathbf{H} E[\mathbf{s} \cdot \mathbf{n}^H] + E[\mathbf{n} \mathbf{s}^H] \mathbf{H}^H + E[\mathbf{n} \mathbf{n}^H] \\
&= \mathbf{H} \mathbf{R}_{ss} \mathbf{H}^H + \mathbf{H} \mathbf{R}_{sn} + \mathbf{R}_{ns} \mathbf{H}^H + \mathbf{R}_{nn}
\end{aligned} \tag{9.3.9}$$

mit

$$\mathbf{R}_{sn} = E[\mathbf{s} \mathbf{n}^H], \quad \mathbf{R}_{ns} = E[\mathbf{n} \mathbf{s}^H] \tag{9.3.10}$$

$$\mathbf{R}_{sr} = E[\mathbf{s} \mathbf{r}^H] = E[\mathbf{s} (\mathbf{H} \cdot \mathbf{s} + \mathbf{n})^H] = E[\mathbf{s} \mathbf{s}^H] \mathbf{H}^H + E[\mathbf{s} \mathbf{n}^H] = \mathbf{R}_{ss} \cdot \mathbf{H}^H + \mathbf{R}_{sn} \tag{9.3.11}$$

Da \mathbf{s} und \mathbf{n} statistisch unabhängig sind, sind sie auch unkorreliert. Ferner ist \mathbf{n} nach Voraussetzung mittelwertfrei.

$$\text{Daher} \quad \mathbf{R}_{sn} = 0, \quad \mathbf{R}_{ns} = 0 \tag{9.3.12}$$

$$(9.3.12) \text{ in } (9.3.9): \quad \mathbf{R}_{rr} = \mathbf{H} \mathbf{R}_{ss} \mathbf{H}^H + \mathbf{R}_{nn} \tag{9.3.13}$$

$$(9.3.12) \text{ in } (9.3.11): \quad \mathbf{R}_{sr} = \mathbf{R}_{ss} \cdot \mathbf{H}^H \tag{9.3.14}$$

(9.3.13) und (9.3.14) in (9.3.8) schließt den Beweis ab

$$\mathbf{W} = \mathbf{R}_{ss} \mathbf{H}^H (\mathbf{H} \mathbf{R}_{ss} \mathbf{H}^H + \mathbf{R}_{nn})^{-1} \quad \text{q.e.d.}$$

Wichtiger Anwendungsfall: $\mathbf{R}_{ss} = E_s \mathbf{I}_M$, $\mathbf{R}_{nn} = \sigma_n^2 \mathbf{I}_N$

Aus (9.3.2) folgt nach wenigen Umformungen

$$\mathbf{V} = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H + \alpha \mathbf{I}_N)^{-1} ; \quad \alpha = \sigma_n^2 / E_s \quad (9.3.15)$$

Mit dem Inversionsatz der linearen Algebra

$$(\mathbf{A} + \mathbf{B} \mathbf{C} \mathbf{D})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{C}^{-1} + \mathbf{D} \mathbf{A}^{-1} \mathbf{B})^{-1} \mathbf{D} \mathbf{A}^{-1}$$

folgt aus (9.3.15)

$$\mathbf{V} = (\mathbf{H}^H \mathbf{H} + \alpha \mathbf{I}_M)^{-1} \cdot \mathbf{H}^H \quad (9.3.16)$$

Beweis:

$$\mathbf{H}^H (\mathbf{H}\mathbf{H}^H + \alpha \mathbf{I}_N)^{-1} = \mathbf{H}^H \underbrace{(\alpha \mathbf{I}_N + \mathbf{H} \mathbf{I}_M \mathbf{H}^H)^{-1}}_{\mathbf{A} \quad \mathbf{B} \quad \mathbf{C} \quad \mathbf{D}} =$$

$$\mathbf{H}^H (\mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{H} (\mathbf{I}_M + \mathbf{H}^H \mathbf{A}^{-1} \mathbf{H})^{-1} \mathbf{H}^H \mathbf{A}^{-1}) = \mathbf{H}^H \mathbf{A}^{-1} - \mathbf{H}^H \mathbf{A}^{-1} \mathbf{H} (\mathbf{I}_M + \mathbf{H}^H \mathbf{A}^{-1} \mathbf{H})^{-1} \mathbf{H}^H \mathbf{A}^{-1}$$

$$(\underbrace{\mathbf{I}_M + \mathbf{H}^H \mathbf{A}^{-1} \mathbf{H}}_{\mathbf{I}_M}) (\mathbf{I}_M + \mathbf{H}^H \mathbf{A}^{-1} \mathbf{H})^{-1} \mathbf{H}^H \cdot \mathbf{A}^{-1} - \mathbf{H}^H \mathbf{A}^{-1} \mathbf{H} (\mathbf{I}_M + \mathbf{H}^H \mathbf{A}^{-1} \mathbf{H})^{-1} \mathbf{H}^H \mathbf{A}^{-1}$$

$$\underbrace{[\mathbf{I}_M + \mathbf{H}^H \mathbf{A}^{-1} \mathbf{H} - \mathbf{H}^H \mathbf{A}^{-1} \mathbf{H}]}_{= \mathbf{I}_M} (\mathbf{I}_M + \mathbf{H}^H \mathbf{A}^{-1} \mathbf{H})^{-1} \mathbf{H}^H \mathbf{A}^{-1} = \left(\mathbf{I}_M + \mathbf{H}^H \frac{1}{\alpha} \mathbf{I}_N \cdot \mathbf{H} \right)^{-1} \mathbf{H}^H \frac{1}{\alpha} \mathbf{I}_N$$

$$\alpha (\mathbf{H}^H \mathbf{H} + \alpha \mathbf{I}_M)^{-1} \cdot \mathbf{H}^H \cdot \frac{1}{\alpha} = (\mathbf{H}^H \mathbf{H} + \alpha \mathbf{I}_M)^{-1} \mathbf{H}^H \quad \text{q. e. d.}$$

9.4 Singulärwertzerlegung einer Matrix

Sei $\mathbf{H} \in \mathbb{C}^{N \times M}$. Eine Eigenwertzerlegung von \mathbf{H} ist bekanntlich nicht immer möglich, besonders wenn $M \neq N$. Dann kann man die Singulärwertzerlegung wählen

$$\mathbf{H} = \mathbf{U} \cdot \mathbf{D} \cdot \mathbf{V}^H \quad \text{mit} \quad \mathbf{D} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots \\ & \ddots & \\ 0 & \sqrt{\lambda_P} & \cdots \\ \cdots & 0 & \ddots & 0 \end{bmatrix} \in \mathbb{R}^{N \times M} \quad (9.4.1)$$

wobei $\mathbf{U} \in \mathbb{C}^{N \times N}$, $\mathbf{V} \in \mathbb{C}^{M \times M}$ unitäre Matrizen sind, für die gilt

$$\mathbf{U}^{-1} = \mathbf{U}^H \quad , \quad \mathbf{V}^{-1} = \mathbf{V}^H \quad (9.4.2)$$

$\sqrt{\lambda_v}$ ($v = 1, \dots, P$) heißen Singulärwerte von \mathbf{H} .

λ_v ($v = 1, \dots, P$) sind die positiven, von Null verschiedenen Eigenwerte der Hermiteschen Matrix

$$\mathbf{Q}_N = \mathbf{H}\mathbf{H}^H \in \mathbb{C}^{N \times N} \quad (9.4.3)$$

wobei $\text{rang } \mathbf{H} = \text{rang } \mathbf{Q}_N = P$ und $P \leq \min\{M, N\}$ (9.4.4)

Die übrigen Eigenwerte λ_v ($v = P + 1, \dots, N$) von \mathbf{Q}_N sind Null. Für \mathbf{Q}_N existiert immer eine Eigenwertzerlegung

$$\mathbf{U}^H \mathbf{Q}_N \mathbf{U} = \Lambda_N \quad \text{mit} \quad \Lambda_N = \begin{bmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \lambda_P & \\ 0 & & & \ddots & \\ & & & & 0 \end{bmatrix} \in \mathbb{R}^{N \times N} \quad (9.4.5)$$

λ_v ($v = 1, \dots, N$) sind die Eigenwerte von \mathbf{Q}_N . \mathbf{U} ist die Matrix der Eigenvektoren zu den Eigenwerten $\lambda_1, \dots, \lambda_N$.

Beweis von (9.4.1)-(9.4.5):

Da $\mathbf{Q}_N = \mathbf{Q}_N^H$, ist die Zerlegung (9.4.5) immer möglich.

Mit $\mathbf{V}\mathbf{V}^H = \mathbf{I}_M$ und (9.4.3) folgt aus (9.4.5)

$$\underbrace{\mathbf{U}^H \mathbf{H} \mathbf{V}}_{\mathbf{A}} \underbrace{\mathbf{V}^H \mathbf{H}^H \mathbf{U}}_{\mathbf{A}^H} = \Lambda_N \quad ; \quad \mathbf{A} \in \mathbb{C}^{N \times M} \quad (9.4.6)$$

Nun kann man Λ_N zerlegen

$$\Lambda_N = \underbrace{\begin{bmatrix} \sqrt{\lambda_1} & & & 0 \\ & \ddots & & \\ & & \sqrt{\lambda_P} & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix}}_{\mathbf{D} \ (N \times M)} \cdot \underbrace{\begin{bmatrix} \sqrt{\lambda_1} & & & 0 \\ & \ddots & & \\ & & \sqrt{\lambda_P} & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix}}_{\mathbf{D}^H \ (M \times N)} = \mathbf{D}\mathbf{D}^H \quad (9.4.7)$$

Aus dem Vergleich von (9.4.6) und (9.4.7) folgt

$$\mathbf{U}^H \mathbf{H} \mathbf{V} = \mathbf{D} \quad (9.4.8)$$

und daraus durch einfache Umformung (9.4.1).

Bisher musste $\mathbf{V} \in \mathbb{C}^{M \times M}$ nur die Bedingung $\mathbf{V}^{-1} = \mathbf{V}^H$ erfüllen. Damit ist \mathbf{V} aber noch nicht vollständig festgelegt. Eine Möglichkeit ist, sie aus (9.4.8) wie folgt zu bestimmen: Dazu multipliziert man von links mit \mathbf{U} , dann mit \mathbf{H}^H und schließlich mit $(\mathbf{H}^H \mathbf{H})^{-1}$ und erhält

$$\mathbf{V} = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{U} \mathbf{E} \quad (9.4.9)$$

Dabei muss man voraussetzen, dass $\mathbf{H}^H \mathbf{H}$ vollen Rang besitzt, da sonst die Inverse nicht existiert. Ist diese Voraussetzung nicht erfüllt, können die Eigenvektoren einzeln wie folgt berechnet werden. Multipliziert man (9.4.1) von links mit \mathbf{U}^H , und wendet danach auf beide Seiten die Hermitesche Operation an, so wird

$$\mathbf{V} \mathbf{D}^H = \mathbf{H}^H \mathbf{U} \quad (9.4.10)$$

Mit $\mathbf{V} = (\mathbf{v}_1 \dots \mathbf{v}_M)$ und $\mathbf{U} = (\mathbf{u}_1 \dots \mathbf{u}_N)$ sowie \mathbf{D}^H nach (9.4.7) folgt

$$(\mathbf{v}_1 \dots \mathbf{v}_M) \begin{bmatrix} \sqrt{\lambda_1} & & & \vdots & 0 \\ & \ddots & & & \vdots \\ & & \sqrt{\lambda_P} & & \vdots \\ & & & \ddots & \vdots \\ 0 & & & & \vdots & 0 \end{bmatrix} = \mathbf{H}^H (\mathbf{u}_1 \dots \mathbf{u}_N) \quad (9.4.11)$$

und daraus $\mathbf{v}_i = \mathbf{H}^H \mathbf{u}_i / \sqrt{\lambda_i} ; i = 1, \dots, P$ (9.4.12)

$\mathbf{v}_{P+1}, \dots, \mathbf{v}_M$ lassen sich damit nicht gewinnen, da $\lambda_i = 0$ für $i > P$. Diese Eigenvektoren muss man orthonormal zu den übrigen, z. B. mit dem Gram-Schmidtschen Orthogonalisierungsverfahren ermitteln.

9.5 Berechnung der MIMO-Kanalkapazität

9.5.1 Herleitung von $\mathcal{C} = \text{ld} \left| \mathbf{I}_N + \mathbf{H} \mathbf{H}^H \frac{E_S}{\sigma_n^2} \right|$

Gegeben $\mathcal{C} = \text{ld} \left| \mathbf{I}_N + \Lambda_N \frac{E_S}{\sigma_n^2} \right|$ (9.5.1)

Die Eigenwertzerlegung von $\mathbf{H} \mathbf{H}^H$ liefert

$$\mathbf{J}^H \mathbf{H} \mathbf{H}^H \mathbf{U} = \Lambda_N = \begin{bmatrix} \Lambda_P & \vdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \vdots & 0 \end{bmatrix} \quad (N \times N \quad ; \quad \Lambda_P = \text{diag}(\lambda_1, \dots, \lambda_P) \quad (9.5.2)$$

Damit $\mathcal{C} = \text{ld} \left| \mathbf{I}_N + \mathbf{U}^H \mathbf{H} \mathbf{H}^H \mathbf{U} \frac{E_S}{\sigma_n^2} \right|$ (9.5.3)

Für Determinanten gilt allgemein

$$|\mathbf{I}_M + \mathbf{A} \mathbf{B}| = |\mathbf{I}_N + \mathbf{B} \mathbf{A}|, \quad \mathbf{A} \quad (M \times N), \quad \mathbf{B} \quad (N \times M, \text{ auch } M = N) \quad (9.5.4)$$

Somit

$$\mathcal{C} = \text{ld} \left| \mathbf{I}_N + \underbrace{\mathbf{H} \mathbf{H}^H \mathbf{U} \mathbf{U}^H}_{\mathbf{I}_N} \frac{E_S}{\sigma_n^2} \right| = \text{ld} \left| \mathbf{I}_N + \mathbf{H} \mathbf{H}^H \frac{E_S}{\sigma_n^2} \right| \quad \text{q.e.d.}$$

9.5.2 Herleitung von $C = \text{ld} \left| \mathbf{I}_M + \mathbf{H}^H \mathbf{H} \frac{E_S}{\sigma_n^2} \right|$

Gegeben $C = \text{ld} \left| \mathbf{I}_M + \Lambda_M \frac{E_S}{\sigma_n^2} \right|$ (9.5.5)

Eigenwertzerlegung von $\mathbf{H}^H \mathbf{H}$ liefert

$$\mathbf{V}^H \mathbf{H}^H \mathbf{H} \mathbf{V} = \Lambda_M = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & 0 \end{bmatrix} \quad (M \times M) \quad (9.5.6)$$

zusammen mit (9.5.4) folgt

$$C = \text{ld} \left| \mathbf{I}_M + \mathbf{H}^H \mathbf{H} \frac{E_S}{\sigma_n^2} \right| \quad \text{q. e. d.}$$

10. Literaturverzeichnis

- [1] Roy, R.; Ottersten, B.: Spatial division multiple access wireless communication systems. US Patent, 5 515 378, May 1996.
- [2] Paulraj, A.; Kailath, T.: Increasing capacity in wireless broadcast systems using distributed transmission/directional reception. US Patent, 5 345 599, 1994.
- [3] Wittneben, A.: A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation. *IEEE International Conf. on Communications (ICC 93)*, May 1993, 1630- 1634, 1993.
- [4] Seshadri, N.; Winters, J.: Two signaling schemes for improving the error performance of FDD transmission systems using transmitter antenna diversity. *Proc. 1993 IEEE Vehicular Technology Conf. (VTC 43rd)*, 508-511, May 1993.
- [5] Telatar, I.: Capacity of multi-antenna Gaussian channels. Technical Report #BL0112170-950615-07TM, AT&T Bell Laboratories, 1995.
- [6] Foschini, G.: Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Tech. J.*, 41-59, 1996.
- [7] Foschini, G.; Gans, M.: On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Pers. Comm.*, 6(3), 311-335, March 1998.
- [8] Tarokh, V.; Seshadri, N.; Calderbank, A.: Space-time codes for high data rate wireless communication: Performance criterion and code construction. *IEEE Trans. Inf. Theory*, 44(2), 744-765, March 1998.
- [9] 3rd Generation Partnership Project. Technical specification Group Radio Access Network. *MIMO in UTRA*, 2005.
- [10] Fischer, R.; Windpassinger, C.; Lampe, A.; Huber, J.: Space-Time Transmission using Tomlinson-Harashima Precoding. *Proceedings of 4. ITG Conference on Source and Channel Coding*, 139-147, Berlin, January 2002.
- [11] Fischer, R.; Windpassinger, C.; Lampe, A.; Huber, J.: MIMO Precoding for Decentralized Receivers. *Proceedings of International Symposium on Information Theory (ISIT '02)*, Lausanne, Switzerland, June/July 2002.
- [12] Baier, P.; Meurer, M.; Weber, T.; Troeger, H.: Joint Transmission (JT), an alternative rationale for the downlink of Time Division CDMA using multi-element transmit antennas. *Proc. IEEE Intern. Symp. on Spread Spectrum Techniques and Applications (ISSSTA)*, 1-5, Parsippany, NJ, Sept. 2000.
- [13] Shannon, C.: A mathematical theory of communication. *Bell Labs Tech. J.*, 27, 379-423, 623-656, July and October 1948.
- [14] Telatar, I.: Capacity of multi-antenna Gaussian channels. *European Trans. Tel.*, 10(6), 585-595, November/December 1999.
- [15] Grant, A.: Rayleigh fading multiple-antenna channels. *EURASIP J. Appl. Signal Processing*, 2002(3), 316-329, March 2002.
- [16] Ivrlac, M.T.; Kurpjuhn, T.P.; Brunner, C.; Utschick, W.: Efficient use of fading correlations in MIMO systems. *IEEE Vehicular Technology Conference, VTC*, October 2001.

-
- [17] Shiu, D.; Foschini, J.; Gans, M.; Kahn, J.: Fading correlation and its effect on the capacity of multielement antenna systems. *IEEE Transactions on Communications*, 48, 502-513, March 2000.
- [18] Ivrlac, M.; Kurpjuhn, T.; Brunner, C.; Nossek, J.: On Capacity of MIMO-Systems in Correlated Channels. *ITG-Diskussionssitzung "Systeme mit intelligenten Antennen"*, Ilmenau, March 2001.
- [19] Kießling, M.: Statistical analysis and transmit prefiltering for MIMO wireless systems in correlated fading environments. Dissertation Universität Stuttgart, 2004. Shaker Verlag, ISBN-Nr. 3-8322-3444-6. www.inue.uni-stuttgart.de/german/forschung/disserta/index.html
- [20] Kießling, M.: Unifying Analysis of Ergodic MIMO Capacity in Correlated Rayleigh Fading Environments. *European Transactions on Telecommunications*, 16(1), 17-35, Jan. 2005.
- [21] Kießling, M.; Speidel, J.: Mutual Information of MIMO Channels in Correlated Rayleigh Fading Environments - a General Solution. *IEEE International Conference on Communications (ICC)*, Paris, France, June 2004.
- [22] Kießling, M.; Speidel, J.; Reinhardt, M.: Unifying analysis of ergodic MIMO capacity in correlated Rayleigh fading environments. *European Wireless Conference*, February 2004.
- [23] Alamouti, S.: A simple transmit diversity technique for wireless communications. *IEEE J. Sel. Areas Comm.*, 16(8), 1451-1458, October 1998.
- [24] Tarokh, V.; Jafarkhani, H.; Calderbank, A.: Space-time block coding for wireless communications: Performance results. *IEEE J. Sel. Areas Comm.*, 17(3), 451-460, March 1999.
- [25] Tarokh, V.; Jafarkhani, H.; Calderbank, A.: Space-time block codes from orthogonal designs. *IEEE Trans. Inf. Theory*, 45(5), 1456-1467, July 1999.
- [26] Tarokh, V.; Naguib, A.; Seshadri, N.; Calderbank, A.: Combined array processing and space-time coding. *IEEE Trans. Inf. Theory*, 45(4), 1121-1128, May 1999.
- [27] Tarokh, V.; Naguib, A.; Seshadri, N.; Calderbank, A.: Space-time codes for high data rate wireless communication: Performance criteria in the presence of channel estimation errors, mobility, and multiple paths. *IEEE Trans. Comm.*, 47(2), 199-207, February 1999.
- [28] Boronka, A.: Verfahren mit adaptiver Symbolauslöschung zur iterativen Detektion codierter MIMO-Signale. Dissertation Universität Stuttgart, 2004. Shaker Verlag 2004, ISBN-Nr. 3-8322-3759-3. www.inue.uni-stuttgart.de/german/forschung/disserta/index.html
- [29] Berrou, C.; Glavieux, A.; Thitimajshima, P.: Near Shannon limit error-correcting coding and decoding: Turbo-codes. *IEEE International Conference on Communications (ICC) 1993*, 1064-1070, Genua, Italien, 1993.
- [30] Hagenauer, J.: The turbo principle: Tutorial introduction and state of the art. *Proc. 1st Internat. Symposium on Turbo Codes (1997)*, 1-12, 1997.
- [31] ten Brink, S.: Designing iterative decoding schemes with the Extrinsic Information Transfer Chart. *AEÜ International Journal of Electronics and Communications*, 54(6), 389-398, November 2000.
- [32] ten Brink, S.: Convergence behaviour of iteratively decoded parallel concatenated codes. *IEEE Transactions on Communications*, 49(10), October 2001.

-
- [33] ten Brink, S.: Design of concatenated coding schemes based on iterative decoding convergence. Dissertation Universität Stuttgart, 2001. Shaker-Verlag, ISBN-Nr. 3-8322-0684-1. www.inue.uni-stuttgart.de/german/forschung/disserta/index.html
- [34] ten Brink, S.; Speidel, J.; Yan, R.-H.: Iterative demapping for QPSK modulation. *IEEE Electronics Letters*, 34(15), 1459-1460, July 1998.
- [35] ten Brink, S.; Speidel, J.; Yan, R.-H.: Iterative demapping and decoding for multilevel modulation. *IEEE International Conference on Global Communications (GLOBECOM)*, 579-584, Sydney, Australia, November 1998.
- [36] Båro, S.; Hagenauer, J.; Witzke, M.: Iterative detection of MIMO transmission using a list-sequential (LISS) detector. *IEEE International Conference on Communications (ICC) 2003*, 2653-2657, Anchorage, USA, 2003.
- [37] Hochwald, B.; ten Brink, S.: Achieving near-capacity on a multiple-antenna channel. *IEEE Transactions on Communications*, 51(3), 389-399, March 2003.
- [38] Boronka, A.; Speidel, J.: A low complexity MIMO system based on BLAST and iterative anti-Gray-demapping. *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Beijing, China, September 2003.
- [39] Boronka, A.; Efinger, D.; Speidel, J.: Improving MIMO detection by L-value analysis and adaptive threshold-based cancellation. *IEEE International Conference on Global Communications (GLOBECOM)*, San Francisco, USA, December 2003.
- [40] Boronka, A.; Rankl, T.; Speidel, J.: Iterative nonlinear detection of MIMO signals using an MMSE-OSIC detector and adaptive cancellation. *ITG Conference on Source and Channel Coding (SCC)*, 17-324, Erlangen, Germany, January 2004.
- [41] Boronka, A.: Analysis and results on optimized MIMO detection with threshold-based adaptive cancellation. *IEEE Vehicular Technology Conference (VTC)*, Milan, Italy, May 2004.
- [42] Boronka, A.; Muhammad, N.; Speidel, J.: Removing error floor for bit interleaved coded modulation MIMO transmission with iterative detection. *International Conference on Communications (ICC)*, Seoul, Korea, May 2005.
- [43] Muhammad, N.; Speidel, J.: Joint optimization of signal constellation and bit labeling for bit-interleaved coded modulation with iterative decoding. *IEEE Communications Letters*, 9(9), 775-777, September 2005.
- [44] Paulraj, A.; Nabar, R.; Gore, D.: Introduction to Space-Time Wireless Communications. Cambridge University Press, 2003.
- [45] Larsson, E.; Stoica, P.: Space-Time Block Coding for Wireless Communications. Cambridge University Press, 2003.
- [46] Vucetic, B.; Yuan, J.: Space-Time Coding. John Wiley Publisher, 2003.
- [47] Proakis, J.: Digital Communications. Mc Graw-Hill Book Company, 3rd edition, 1995.
- [48] Kießling, M.; Speidel, J.: Unifying Performance Analysis of Linear MIMO Receivers in Correlated Rayleigh Fading Environments. *IEEE International Symposium on Spread Spectrum Techniques and Applications (ISSSTA)*, Sydney, Australia, August 2004.
- [49] Kießling, M.; Speidel, J.: Asymptotically Tight Bound on SER of MIMO Zero-Forcing Receivers with Correlated Fading. *IEEE International Symposium on Information Theory (ISIT)*, Chicago, USA, June 2004.

- [50] Kießling, M.; Speidel, J.; Geng, N.; Reinhardt, M.: Performance analysis of MIMO maximum likelihood receivers with channel correlation, colored Gaussian noise, and linear prefiltering. *IEEE International Conference on Communications (ICC)*, Anchorage, USA, May 2003.
- [51] Wolniansky, P. W.; Foschini, G. J.; Golden, G. D.; Valenzuela, R. A.: V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel. *URSI International Symposium on Signals, Systems, and Electronics (ISSSE)*, 295-300, Pisa, Italy, September-October 1998.
- [52] Golden, G. D.; Foschini, G. J.; Valenzuela, R. A.; Wolniansky, P. W.: Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture. *Electronic letters*, 35(1), 14-15, January 1999.

11. Verfasserportrait

Joachim Speidel studierte Elektrotechnik (Nachrichtentechnik) an der Universität Stuttgart und schloss 1975 mit dem Dipl.-Ing. und 1980 mit dem Dr.-Ing. ab.

Seit 1992 ist er Professor in der Fakultät Informatik, Elektrotechnik und Informationstechnik an der Universität Stuttgart und Leiter des Instituts für Nachrichtenübertragung. Seine Gruppe arbeitet auf dem breiten Gebiet der Nachrichtenübertragung in drahtlosen, festen, elektrischen und optischen Netzen mit den Schwerpunkten Codierung, Modulation und Detektion. Anwendungsgebiete sind Digitaler Rundfunk, Mobilfunk, MIMO sowie die optische Nachrichtenübertragung (www.inue.uni-stuttgart.de).

Von 1980 bis 1992 war er bei der Philips Kommunikations Industrie AG (der heutigen Lucent Technologies Bell Labs Innovations) tätig. Als Entwicklungsingenieur hat er 1985 den ersten Videocodec für 2 Mbit/s gebaut und die Standardisierung neuer Codierverfahren mit seinem Team vorangetrieben. Seine Tätigkeiten in Entwicklung und Forschung umspannten ein breites Feld wie Video- und Teilnehmerzugangssysteme, einschließlich der optischen Zugangstechnik. Er war in unterschiedlichen Positionen tätig, als Entwicklungsingenieur, Leiter von Labors, Prokurist sowie als Philips-Direktor verantwortlich für einen Produktbereich.

Mit mehreren Patenten konnte er zum Erfolg von Produkten und internationalen Standards beitragen. Professor Speidel ist Mitglied verschiedener nationaler und internationaler Organisationen sowie Beratungs- und Gutachtergremien.
