Mathematische Codierungstheorie

Das Sicherstellen von Klarheit in einer verrauschten Welt

Ein Skript zur Lehrerfortbildung
basierend auf einem Vortrag von Prof. Joachim Rosenthal (Universität Zürich)

Passau, 12. Dezember 2025

Zusammenfassung

Dieses Skript bietet eine Einführung in die mathematischen Grundlagen der Codierungstheorie. Ausgehend von den fundamentalen Arbeiten Claude Shannons werden die Konzepte der Entropie und der Quellencodierung erläutert. Der Schwerpunkt liegt auf der Kanalcodierung: Es werden lineare Blockcodes, die Kugelpackungsschranke, perfekte Codes (am Beispiel des Hamming-Codes) sowie MDS-Codes (Reed-Solomon) detailliert behandelt. Ergänzend werden praktische Anwendungen wie ISBN-10 und ISBN-13 analysiert und moderne LDPC-Codes vorgestellt.

Inhaltsverzeichnis

1 Einleitung: Shannons Vermächtnis

Claude Shannon (1916–2001) gilt als der Begründer der mathematischen Informationstheorie. In seinen bahnbrechenden Arbeiten von 1948 und 1949 [Sha48, Sha49] legte er den Grundstein für drei fundamentale Fragen der modernen Kommunikation:

1. **Redundanz (Quellencodierung):** Wie kann Information mit einer minimalen Anzahl von Bits beschrieben werden?
2. **Zuverlässigkeit (Kanalcodierung):** Wie überträgt man Information fehlerfrei über einen rauschbehafteten Kanal?
3. **Sicherheit (Kryptographie):** Wie überträgt man Information privat zwischen zwei Parteien?

Shannon formulierte das **Separationsprinzip**, welches besagt, dass diese Aufgaben getrennt voneinander optimiert werden können (Quellencodierung → Verschlüsselung → Kanalcodierung), ohne die theoretische Leistungsfähigkeit des Gesamtsystems zu beeinträchtigen.

$$\boxed{\text{Nachrichtenquelle}} \longrightarrow \boxed{\text{Quellencodierer}} \longrightarrow \boxed{\text{Verschlüsseler}} \rightarrow$$

$$\longrightarrow \boxed{\text{Kanalcodierer}} \longrightarrow \boxed{\text{Modulator}} \rightarrow$$

$$\rightsquigarrow \text{ Rauschen } \rightsquigarrow \text{ Übertragung } \rightsquigarrow \text{ Rauschen } \rightsquigarrow$$

$$\longrightarrow \boxed{\text{Demodulator}} \longrightarrow \boxed{\text{Kanaldecodierer}} \longrightarrow$$

$$\rightarrow \boxed{\text{Entschlüsseler}} \longrightarrow \boxed{\text{Quellendecodierer}} \longrightarrow \boxed{\text{Nachricht}}$$

Eine gute Einführung in die Quellencodierung und die Kanalcodierung findet man im Buch von Dirk Hoffmann [Hof23].

2 Quellencodierung: Eliminierung von Redundanz

Das Ziel der Quellencodierung ist die Datenkompression, also das Entfernen von Redundanz aus einer Nachricht.

2.1 Entropie als Maß für Information

Die theoretische Untergrenze für die Kompression wird durch die Entropie bestimmt.

Definition 2.1 (Entropie nach Shannon [Sha48]). *Sei X eine diskrete Zufallsvariable mit den Werten $x_1, \ldots, x_n$ und der Wahrscheinlichkeitsfunktion $f_X(x_i) = p_i$. Die Entropie von X ist definiert als:*

$$H(X) := -\sum_{i=1}^{n} p_i \log_2 p_i.$$

Es gilt dabei stets $0 \leq H(X) \leq \log_2 n$.

Beispiel 2.2 (DNA-Sequenz). *Betrachten wir einen Text, der nur aus den Buchstaben A, C, G, T besteht. Sind alle Buchstaben gleichverteilt ($p_i = 1/4$), benötigen wir $\log_2 4 = 2$ Bits pro Buchstabe. Angenommen jedoch, die Wahrscheinlichkeiten sind ungleich verteilt:*

- $P(A) = 1/2$ *(Codierung z.B. 0)*
- $P(C) = 1/4$ *(Codierung z.B. 10)*
- $P(G) = 1/8$ *(Codierung z.B. 110)*
- $P(T) = 1/8$ *(Codierung z.B. 111)*

Die Entropie berechnet sich zu:

$$H(X) = \frac{1}{2}\log_2(2) + \frac{1}{4}\log_2(4) + 2 \cdot \frac{1}{8}\log_2(8) = 0.5 + 0.5 + 0.75 = 1.75 \text{ Bits.}$$

2.2 Satz über die rauschfreie Codierung

Satz 2.3 (Satz über die rauschfreie Codierung [Sha48]). *Seien $a_1, a_2, \ldots$ eine Zufallsstichprobe einer diskreten Zufallsvariablen X, die einen Text repräsentiert. Dann beträgt die durchschnittliche Anzahl von Bits pro Symbol in jedem binären Codierungsschema mindestens $H(X)$. Außerdem gibt es Codierungsschemata, die im Durchschnitt höchstens ϵ Bits pro Symbol mehr als $H(X)$ benötigen, für jedes beliebige $\epsilon > 0$.*

Praktische Verfahren, die diese Grenzen annähernd erreichen, sind die Huffman-Codierung [Huf52] (benötigt Kenntnis der Verteilung) und die universellen Lempel-Ziv-Algorithmen [ZL77, ZL78] (Basis von ZIP, PNG etc.).

Die Huffman-Codierung kann auf gymnasialer Stufe gut behandelt werden.

3 Kanalcodierung: Strukturierte Redundanz

Vor Shannon galt das Prinzip: *Repetitio est mater communicationis.* Um Fehler zu vermeiden, wiederholte man Signale. Shannon zeigte jedoch, dass dies ineffizient ist.

Satz 3.1 (Kanalcodierungssatz [Sha48]). *Jeder Übertragungskanal hat eine Kapazität C. Solange die Übertragungsrate $R < C$ ist, existiert ein Codierungsschema, das eine Symbolfehlerrate von höchstens ϵ ermöglicht. Ist $R > C$, ist eine zuverlässige Übertragung unmöglich.*

Der Beweis war existentiell und probabilistisch. Die Aufgabe der Codierungstheorie ist es, konstruktive Codes zu finden, die effizient decodierbar sind.

3.1 Lineare Blockcodes

Wir betrachten Codes über einem endlichen Körper $\mathbb{F}_q$ (oft $\mathbb{F}_2 = \{0, 1\}$).

Definition 3.2. *Ein **linearer Blockcode** $\mathcal{C}$ der Länge n und Dimension k ist ein k-dimensionaler Unterraum von $\mathbb{F}_q^n$. Man notiert dies als $[n, k]$-Code oder $[n, k, d]$-Code, wenn die Minimaldistanz d spezifiziert wird.*

Ein solcher Code kann durch zwei Matrizen beschrieben werden:

1. **Generatormatrix** G $(k \times n)$: Bildet Nachrichten $m \in \mathbb{F}_q^k$ auf Codewörter $c \in \mathcal{C}$ ab: $c = m \cdot G$. Die Zeilen von G bilden eine Basis von $\mathcal{C}$.
2. **Paritätsprüfmatrix** H $((n-k) \times n)$: Beschreibt den Code als Kern einer linearen Abbildung. Es gilt:
$$\mathcal{C} = \{c \in \mathbb{F}_q^n \mid Hc^T = 0\}.$$

Es gilt stets $HG^T = 0$.

3.2 Hamming-Distanz und Fehlerkorrektur

Die Hamming-Distanz $d(x, y)$ ist die Anzahl der Stellen, an denen sich zwei Vektoren unterscheiden. Die Minimaldistanz des Codes ist $d(\mathcal{C}) = \min_{x \neq y \in \mathcal{C}} d(x, y)$. Ein Code kann t Fehler korrigieren, wenn:
$$t = \left\lfloor \frac{d(\mathcal{C}) - 1}{2} \right\rfloor.$$

Für lineare Codes entspricht die Minimaldistanz dem minimalen Gewicht (Anzahl Nicht-Null-Einträge) der Codewörter ungleich Null. Bezüglich der Matrix H ist $d(\mathcal{C})$ die kleinste Anzahl linear abhängiger Spalten von H.

4 Perfekte Codes und die Hamming-Schranke

Eine zentrale Frage der Codierungstheorie ist: Wie viele Codewörter kann man in einem Raum $\mathbb{F}_q^n$ unterbringen, sodass sie alle einen Mindestabstand $d = 2t + 1$ haben?

4.1 Die Kugelpackungsschranke (Hamming-Bound)

Jedes Codewort $c \in \mathcal{C}$ ist von einer "Kugel" aus Vektoren umgeben, die sich in höchstens t Positionen von c unterscheiden. Damit eine eindeutige Decodierung möglich ist, dürfen sich diese Kugeln nicht überschneiden.

Das Volumen einer solchen Kugel mit Radius t im Raum $\mathbb{F}_q^n$ ist:

$$V(n, t) = \sum_{i=0}^{t} \binom{n}{i} (q-1)^i.$$

Da es $|\mathcal{C}|$ Codewörter gibt und der gesamte Raum q^n Vektoren enthält, gilt die Schranke:

Satz 4.1 (Hamming-Schranke). *Für einen t-fehlerkorrigierenden Code $\mathcal{C} \subseteq \mathbb{F}_q^n$ gilt:*

$$|\mathcal{C}| \cdot \sum_{i=0}^{t} \binom{n}{i} (q-1)^i \leq q^n.$$

Definition 4.2 (Perfekter Code). *Ein Code heißt **perfekt**, wenn er die Hamming-Schranke mit Gleichheit erfüllt. Anschaulich bedeutet dies, dass die Fehlerkugeln den Raum lückenlos ausfüllen. Jeder Vektor im Raum lässt sich eindeutig einem Codewort zuordnen.*

4.2 Der binäre Hamming-Code [7, 4]

Hamming-Codes sind eine Klasse von perfekten Codes mit Distanz $d = 3$ ($t = 1$). Wir konstruieren den binären Hamming-Code der Länge $n = 2^r - 1$ mit $r = 3$. Die Paritätsprüfmatrix H enthält alle Vektoren aus $\mathbb{F}_2^r \setminus \{0\}$ als Spalten.

Für $r = 3$ ist $n = 7$. Die Matrix H ist eine 3×7 Matrix:

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

(Hier sind die Spalten lexikographisch als Binärzahlen 1 bis 7 sortiert).

Satz 4.3. *Der binäre Hamming-Code mit den Parametern $[7, 4, 3]$ ist perfekt.*

Beweis 1. *Die Parameter sind: $n = 7$, $k = 4$ (da $n - k = 3$ Zeilen in H), $d = 3$ (minimale Anzahl linear abhängiger Spalten ist 3, z.B. $e_1 + e_2 + e_3$ ist nicht 0, aber Spalte 1 + Spalte 2 = Spalte 3). Er korrigiert $t = 1$ Fehler. Einsetzen in die Hamming-Schranke ($q = 2$):*

$$\textit{Linke Seite: } |\mathcal{C}| \cdot \sum_{i=0}^{1} \binom{7}{i} = 2^4 \cdot \left(\binom{7}{0} + \binom{7}{1} \right) = 16 \cdot (1 + 7) = 16 \cdot 8 = 128.$$

$$\textit{Rechte Seite: } q^n = 2^7 = 128.$$

Da $128 = 128$, ist der Code perfekt.

Decodierung: Empfängt man y, berechnet man das Syndrom $S = Hy^T$. Ist $S = 0$, war kein Fehler. Ist $S \neq 0$, entspricht S (bei obiger Sortierung der Spalten) der Binärdarstellung der Fehlerposition.

5 MDS-Codes und die Singleton-Schranke

Eine weitere fundamentale Schranke verbindet die Länge n, die Dimension k und die Distanz d.

Satz 5.1 (Singleton-Schranke). *Für jeden linearen $[n, k, d]$-Code gilt:*

$$d \leq n - k + 1.$$

Dies folgt sofort von der Tatsache, dass irgendwelche $n - k + 1$ Spalten der $(n - k) \times n$ Paritätsprüfmatrix H notwendigerweise linear abhängig sind.

Definition 5.2 (MDS-Code)**.** *Ein Code, der die Singleton-Schranke mit Gleichheit erfüllt ($d = n - k + 1$), heißt **MDS-Code** (Maximum Distance Separable).*

MDS-Codes sind insofern optimal, als sie für eine gegebene Redundanz $n-k$ die maximal mögliche Distanz bieten. Eine Paritätsprüfmatrix H repräsentiert einen MDS code genau dann, wenn jede maximale quadratische Untermatrix der Grösse $n - k$ invertierbar ist.

5.1 Beispiele für MDS-Codes

1. **Der Paritäts-Code:** Ein $[n, n-1, 2]$-Code über $\mathbb{F}_2$. $d = 2$ und $n - (n-1) + 1 = 2$. Er ist MDS.
2. **Der Wiederholungscode:** Ein $[n, 1, n]$-Code (z.B. $0 \to 000, 1 \to 111$). $d = n$ und $n - 1 + 1 = n$. Er ist MDS.
3. **Reed-Solomon-Codes (RS-Codes):** Dies ist die wichtigste Klasse von MDS-Codes in der Praxis (QR-Codes, Satellitenkommunikation). Sie basieren auf Polynomen über einem endlichen Körper $\mathbb{F}_q$. Ein Codewort entsteht durch Auswertung eines Polynoms $f(x)$ vom Grad $< k$ an n verschiedenen Stellen $\alpha_1, \ldots, \alpha_n$:

   $$c = (f(\alpha_1), f(\alpha_2), \ldots, f(\alpha_n)).$$

   Da ein Polynom vom Grad $k-1$ höchstens $k-1$ Nullstellen hat, haben zwei verschiedene Polynome höchstens $k-1$ Übereinstimmungen. Sie unterscheiden sich also an mindestens $n-(k-1) = n-k+1$ Stellen. Damit ist die MDS-Eigenschaft bewiesen. Für Details siehe [LG88].

6 Anwendung: ISBN-Codes

Ein klassisches Beispiel für Fehlererkennung im Alltag sind die International Standard Book Numbers (ISBN).

6.1 ISBN-10

ISBN-10 verwendet den Körper $\mathbb{Z}_{11}$ (Symbole 0-9 und X für 10). Eine Nummer $x = (x_1, \ldots, x_{10})$ ist gültig, wenn:

$$\sum_{i=1}^{10} i \cdot x_i \equiv 0 \pmod{11}.$$

Satz 6.1. *ISBN-10 erkennt jeden Einzelfehler und jede Vertauschung (Transposition) zweier Ziffern.*

Beweis 2 (Vertauschung)**.** *Sei x gültig. Durch Vertauschung an Positionen j und k ($j < k$) entsteht y. Differenz der Prüfsummen:*

$$S_y - S_x = (jx_k + kx_j) - (jx_j + kx_k) = (j-k)(x_k - x_j).$$

Da 11 eine Primzahl ist, ist das Produkt modulo 11 nur dann 0, wenn ein Faktor durch 11 teilbar ist. $|j-k| < 10$ (ungleich 0 mod 11) und $|x_k - x_j| < 10$. Der Fehler wird also erkannt, es sei denn $x_k = x_j$ (keine Änderung).

6.2 ISBN-13 (EAN-13)

Seit 2007 wird ISBN-13 verwendet, um den Adressraum zu vergrößern und Kompatibilität zum EAN-Barcode herzustellen. Es wird im Ring $\mathbb{Z}_{10}$ gerechnet. Die Gewichte alternieren zwischen 1 und 3. Prüfgleichung:

$$x_1 + 3x_2 + x_3 + 3x_4 + \cdots + x_{13} \equiv 0 \pmod{10}.$$

Problem der Transposition: Da 10 keine Primzahl ist ($10 = 2 \cdot 5$), werden nicht alle Vertauschungen erkannt. Betrachten wir eine Vertauschung benachbarter Ziffern a und b. Die Gewichte sind 1 und 3. Änderung der Summe: $\Delta = (1 \cdot b + 3 \cdot a) - (1 \cdot a + 3 \cdot b) = 2a - 2b = 2(a - b)$. Ist $2(a - b)$ durch 10 teilbar, wird der Fehler nicht erkannt. Dies passiert, wenn $a - b = 5$ (z.B. Vertauschung von 0 und 5 oder 2 und 7). ISBN-13 ist mathematisch schwächer als ISBN-10, aber logistisch kompatibler.

7 LDPC-Codes: An der Grenze des Möglichen

Die "Nearest-Neighbor"-Decodierung allgemeiner Codes ist NP-schwer [BMvT78]. Daher benötigt man Codes mit spezieller Struktur für effiziente Decodierung. Low-Density Parity-Check (LDPC) Codes wurden von Gallager 1962 eingeführt [Gal62].

Definition 7.1. *Ein LDPC-Code ist ein linearer Blockcode, der durch eine* ***dünnbesetzte*** *(sparse) Paritätsprüfmatrix H definiert ist. D.h., die Anzahl der Einsen in H wächst nur linear mit der Blocklänge n, während die Matrixgröße quadratisch wächst.*

7.1 Tanner-Graphen

LDPC-Codes werden oft durch bipartite Graphen (Tanner-Graphen) visualisiert:

- **Variable Nodes (VN):** Repräsentieren die Bits des Codeworts (Spalten von H).
- **Check Nodes (CN):** Repräsentieren die Paritätsgleichungen (Zeilen von H).

Eine Kante existiert zwischen VN j und CN i, wenn $H_{ij} = 1$.

Die Decodierung erfolgt durch "Message Passing": Informationen über Wahrscheinlichkeiten werden iterativ zwischen Variablen- und Prüfknoten ausgetauscht. Dieser Algorithmus ist sehr effizient und erlaubt es LDPC-Codes (ebenso wie Turbo-Codes), Übertragungsraten extrem nahe an der theoretischen Shannon-Grenze zu erreichen. Sie sind heute Standard in WLAN (802.11n/ac/ax) und 5G Mobilfunk.

8 Übungsaufgaben

1. **Entropie und Codierung:** Eine Wetterstation sendet jeden Tag einen Status: “Sonne” (50%), “Wolken” (25%), “Regen” (12.5%), “Schnee” (12.5%).

   (a) Berechnen Sie die Entropie $H(X)$ dieser Quelle.

   (b) Entwerfen Sie einen binären Code, dessen mittlere Wortlänge genau der Entropie entspricht.

2. **Arithmetik in endlichen Körpern:** Es sei

   $$\mathbb{F}_{101} := \{0, 1, 2, \ldots, 100\}$$

   der endliche Körper mit $q = 101$ Elementen. Berechnen Sie von Hand:

   (a) $25 \times 36 - 49$.

   (b) 64^{-1}.

   (c) 81^{100}.

3. **Hamming-Code [7,4] in Aktion:** Gegeben sei die Paritätsprüfmatrix H (Spalten binär aufsteigend sortiert):

   $$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

   Sie empfangen das Wort $y = (1, 0, 1, 1, 0, 0, 0)$.

   (a) Berechnen Sie das Syndrom $S = Hy^T$.

   (b) An welcher Position liegt der Fehler? Wie lautet das korrigierte Codewort?

4. **ISBN-13 Analyse** Gegeben sei eine fehlerhafte ISBN-13, bei der zwei Ziffern vertauscht wurden: 978-3-03-02**50**42-5 (fiktives Beispiel). Die korrekte Sequenz an dieser Stelle wäre $\ldots$**05**$\ldots$. Zeigen Sie rechnerisch, dass das Prüfsummenverfahren (Gewichte 1, 3 modulo 10) diesen spezifischen Fehler (Vertauschung von 0 und 5 an benachbarter Position) *nicht* erkennen kann.

5. **MDS-Eigenschaft:** Betrachten Sie einen Code der Länge $n = 4$ über $\mathbb{F}_2$ mit nur zwei Codewörtern $\mathcal{C} = \{(0, 0, 0, 0), (1, 1, 1, 1)\}$.

   (a) Bestimmen Sie Dimension k und Minimaldistanz d.

   (b) Prüfen Sie mittels der Singleton-Schranke, ob dieser Code ein MDS-Code ist.

6. **Tanner-Graph zeichnen:** Zeichnen Sie den Tanner-Graphen für folgende Paritätsprüfmatrix:

   $$H = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \end{pmatrix}$$

   Markieren Sie die Variablenknoten $x_1 \ldots x_5$ und die Prüfknoten $p_1 \ldots p_3$.

7. **Zwerghut Problem 1:**

   Szenario: Es gibt 3 Zwerge. Jeder bekommt einen Hut aufgesetzt mit zufälliger Farbe, entweder Rot oder Blau.

   Die Regeln: Jeder Zwerg sieht die Hüte der anderen, aber nicht seinen eigenen. Sie dürfen nicht kommunizieren. Alle müssen **gleichzeitig** eine Entscheidung treffen: "Rot", "Blau" oder "Ich passe".

   Das Ziel: Die Gruppe gewinnt als Ganzes, wenn mindestens einer die Farbe korrekt rät und kein einziger falsch rät. (Wer passt, zählt nicht als falsch).

   Entwickle eine Strategie die garantiert, dass die Gruppe mit Wahrscheinlichkeit $\frac{3}{4}$ gewinnt.

8. **Zwerghut Problem 2:**

   Szenario: Es gibt 7 Zwerge. Jeder bekommt einen Hut aufgesetzt mit zufälliger Farbe, entweder Rot oder Blau.

   Die Regeln: Jeder Zwerg sieht die Hüte der anderen, aber nicht seinen eigenen. Sie dürfen nicht kommunizieren. Die Zwerge sind erlaubt **sequentiell** zu antworten und müssen die eigene Farbe erraten. Jeder Zwerg muss eine Antwort geben, entweder "Rot" oder"Blau".

   Das Ziel: Die Gruppe gewinnt als Ganzes, wenn jeder seine Farbe korrekt rät.

   Entwickle eine Strategie die garantiert, dass die Gruppe mit Wahrscheinlichkeit $\frac{7}{8}$ gewinnt.

9. **Zwerghut Problem 3:**

   Szenario: Es gibt n Zwerge. Jeder bekommt einen Hut aufgesetzt mit zufälliger Farbe, von k verschiedenen Farbmöglichkeiten.

   Die Regeln: Jeder Zwerg sieht die Hüte der anderen, aber nicht seinen eigenen. Sie dürfen nicht kommunizieren. Die Zwerge sind erlaubt **sequentiell** zu antworten und müssen die eigene Farbe erraten. Jeder Zwerg muss eine Antwort geben.

   Das Ziel: Die Gruppe gewinnt als Ganzes, wenn bis auf einen alle ihre Farbe korrekt erraten.

   Entwickle eine Strategie die garantiert, dass die Gruppe gewinnt.

Literatur

[BMvT78] E. R. Berlekamp, R. J. McEliece, and H. C. A. van Tilborg. On the inherent intractability of certain coding problems. *IEEE Trans. Information Theory*, IT-24(3):384–386, 1978.

[Gal62] R.G. Gallager. Low-density parity-check codes. *IRE Trans. on Info. Theory*, IT-8:21–28, 1962.

[Gop88] V. D. Goppa. *Geometry and Codes.* Kluwer Academic Publisher, 1988.

[Hof23] D. Hoffmann. *Einführung in die Informations- und Codierungstheorie.* Springer Verlag, 2023.

[Huf52] D. A. Huffman. A method for the construction of minimum redundancy codes. *Proc. IRE*, 40:1098–1101, 1952.

[LG88] J. H. van Lint and G. van der Geer. *Introduction to Coding Theory and Algebraic Geometry.* Birkhäuser Verlag, 1988.

[Sha48] C.E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423 and 623–656, 1948.

[Sha49] C. E. Shannon. Communication theory of secrecy systems. *Bell System Tech. J.*, 28:656–715, 1949.

[ZL77] J. Ziv and A. Lempel. A universal algorithm for sequential data compression. *IEEE Trans. Information Theory*, IT-23(3):337–343, 1977.

[ZL78] J. Ziv and A. Lempel. Compression of individual sequences via variable-rate coding. *IEEE Trans. Inform. Theory*, 24(5):530–536, 1978.