

Codierungsgewinn und höherstufige Modulation

Bernd Friedrichs, 7.12.2018

In den CCSDS-Standards sind BER-Kurven über E_b/N_0 für einen AWGN gegeben, aber jeweils nur mit binärer Modulation. Die Umrechnung auf andere höherstufige Modulationsverfahren wird hier erklärt.

Mit einer kleinen Näherung erweist sich der Zusammenhang zwischen verschiedenen Modulationsverfahren als sehr einfach, neue Rechnungen oder Simulationen sind nicht erforderlich.

Diese Notiz ist eine Kurzfassung eines Teils von Chapter 2: Digital Passband Modulation over AWGN Channels aus meinem Buch.

1. Grundlagen

Für die Umrechnung von BER-Kurven für binär modulierten AWGN auf höherstufige Modulationsverfahren mit einem 2^M -stufigen Modulationsalphabet ist zu beachten, dass es zwei gegenläufige Effekte gibt:

- die Modulationspunkte in der komplexen Ebene liegen enger zusammen, bei gleichem Signal-Rausch Abstand wird dadurch die Fehlerwahrscheinlichkeit mit wachsendem M ansteigen.
- Ausgehend von einer Energie pro Bit ist die Energie pro Modulationssymbol bzw. die Signalenergie höher und damit auch der Signal-Rausch Abstand. (Natürlich kann man das auch genau gegenteilig sehen: wenn man die Energie pro Modulationssymbol aus Ausgangspunkt nimmt, bedeutet wachsendes M weniger Energie pro Codebit.)

Der erste Effekt dominiert jedoch und deshalb erfordert höherstufige Modulation immer sowohl eine höhere Energie pro Bit wie auch eine höhere Energie pro Symbol.

Expertenanmerkung: Die Codeperformance (Fehlerwahrscheinlichkeit eines Kanal-codierungs-Verfahrens) wird immer unter der Annahme berechnet, dass die Fehler bei der Übertragung (also vor dem Decoder) statistisch unabhängig auftreten, das steckt im Buchstaben $W=White$ des Modells AWGN. Leider folgt aus der statistischen Unabhängigkeit der übertragenen Modulationssymbole nicht die eigentlich notwendige statistische Unabhängigkeit der Codebits (die Umkehrung würde zwar gelten, ist aber leider nicht relevant), denn die Verfälschung eines Modulationssymbols bei der Übertragung kann zu einem oder mehreren Codebitfehlern führen. Wenn man allerdings eine Gray-Codierung für das Mapping der Codebits auf die Modulationssymbole unterstellt, bleibt die statistische Unabhängigkeit näherungsweise erhalten, zumindest bei guten Kanälen.

Noch verzwickter wird es bei komplizierteren Relationen zwischen hochstufigen Modulationssymbolen und hochstufigen Codesymbolen: Bei 8-PSK entspricht 1 Modulationssymbol genau 3 Codebits und bei einem RS(255,k) Code entspricht 1 RS-

Codesymbol einem Byte und damit genau 8 Codebits. Es kann also keine simple Ausrichtung der beiden Bitgruppen geben, lediglich eine Gruppe von 3 RS-Symbolen kann einer Gruppe von 8 8-PSK Symbolen zuordnen.

Alle diese Details werden müssen ignoriert werden und Modulationssymbole, Codesymbole und Codebits werden als statistisch unabhängige unterstellt. Bei guten Kanälen ist das näherungsweise erfüllt bzw. asymptotisch exakt, schlechte Kanäle mit hoher Fehlerrate können notfalls simuliert werden falls man in diesem Bereich wirklich exakte Ergebnisse benötigen sollte.

Ein Sonderfall stellt 8-PSK 4D-TCM dar, bei dem die spezielle Abbildung der Bits auf die Symbole ein essentieller Teil des Codierungsprinzips ist.

Grundlegende Bezeichnungen und Zusammenhänge für die Codierung:

E_b = Energie pro Infobit
 $E_c = R \cdot E_b$ = Energie pro Codebit, $R = \text{Coderate} (< 1)$
 $E_s = R \cdot M \cdot E_b$ = Energie pro Modulationssymbol bei 2^M -stufiger Modulation

Ohne Codierung gilt vereinfacht $E_c = E_b$ und $E_s = M \cdot E_b = M \cdot E_c$.

Grundlegende Charakterisierung des Modulationsverfahrens:

Δ_0 = minimaler Abstand zwischen den Modulationssymbolen in der komplexen Ebene bei Normierung auf $E_s = 1$.

Im allgemeinen Fall werden Modulationsverfahren hier in diesem Papier und auch bei ReSAT ausschließlich über diesen einzigen Parameter Δ_0 charakterisiert.

M	2^M	ASK	PSK	QAM
1	2	2.0000	2.0000	
2	4	0.8944	1.4142	$\sqrt{2} = 1.4142$
3	8	0.4364	0.7654	
4	16	0.2169	0.3902	$\sqrt{2/5} = 0.6325$
5	32	0.1083	0.1960	$\sqrt{2/10} = 0.4472$
6	64	0.0541	0.0981	$\sqrt{2/21} = 0.3086$
7	128	0.0271	0.0491	$\sqrt{2/41} = 0.2209$
8	256	0.0135	0.0245	$\sqrt{2/85} = 0.1534$

Tabelle 1: Minimale Euklidische Distanz Δ_0 für ASK, PSK, QAM

M	2^M	Modulation	$\gamma_1 = R_2/R_1, \gamma_2 = R_3/R_1, \gamma_3 = R_4/R_1,$	Δ_0
4	16	16 APSK (4+12)	$\gamma_1 = 2.57$	0.583
5	32	32 APSK (4+8+16)	$\gamma_1 = 2.53, \gamma_2 = 4.30$	0.382
6	64	64 APSK (4+12+20+28)	$\gamma_1 = 2.73, \gamma_2 = 4.52, \gamma_3 = 6.31$	0.281

Tabelle 2: Minimale Euklidische Distanz Δ_0 für APSK

Expertenanmerkung: bei APSK ist die Spezifikation der Ringradien abhängig von der Coderate, vom Codeverfahren, von der Rahmenlänge und auch unterschiedlich bei ETSI und CCSDS.

2. BER-Berechnung bei uncodierter Übertragung

Für BER-Kurven kann man entweder die Entscheidungsbereiche der Modulationsverfahren korrekt berechnen (z.B. Tortenstücke bei PSK) oder als general upper bound über Kugeln in der komplexen Ebene vom Radius $(\Delta_0/2) \cdot \sqrt{E_s}$ wie folgt abschätzen:

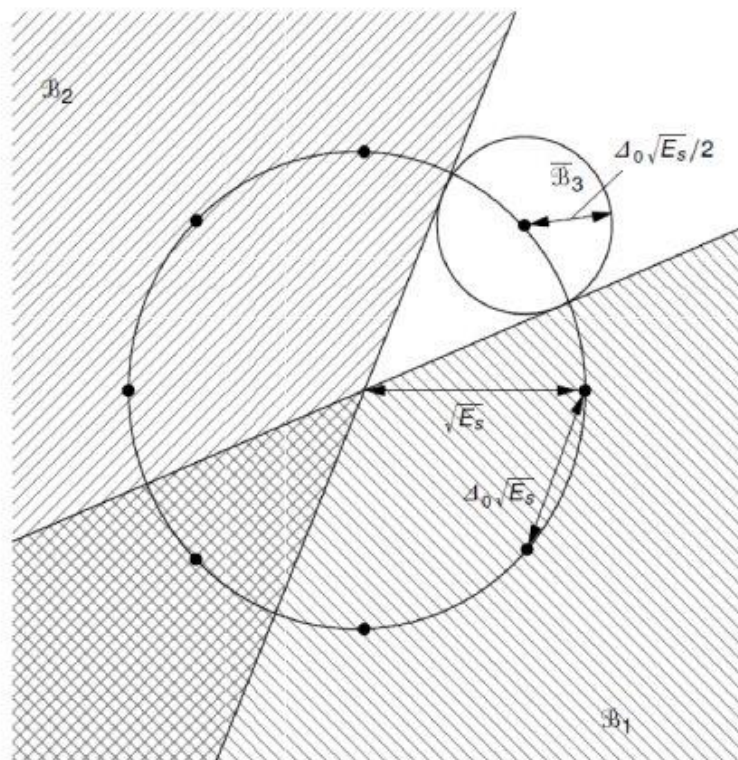


Abb. 1: Entscheidungsbereiche am Beispiel 8-PSK

Jedes Empfangssymbol außerhalb der Kugel wird als fehlerhaft angesehen, das ist also eine (geringfügige) Verschlechterung des ML-Decoders. Diese Methode wird bei ReSAT angewendet da die exakten Entscheidungsbereiche zumindest bei APSK geometrisch viel zu kompliziert sind. Asymptotisch für kleines BERs ist die Kugel-Methode exakt, lediglich bei großem BER bzw. schlechten Kanälen wird BER nach oben mit der Kugelmethode überschätzt.

Expertenanmerkung: Das Kugelprinzip im weiteren Sinne kommt auch bei der Kanalcodierung zur Anwendung in Form von Decodierkugeln (Entscheidungsbereiche), dann sind die Kugeln allerdings im n -dimensionalen Raum (n =Blocklänge) zu verstehen und der Kugelrand bzw. Kugeloberfläche ist über den Hamming-Abstand anstelle des Euklidischen Abstands definiert. In der Kanalcodierung wird das dann als BMD-Decoder bezeichnet, der nur eine geringe Verschlechterung des optimalen ML-Decoders bedeutet und asymptotisch sogar gleichwertig ist, aber wesentlich einfacher im Decoder implementierbar ist.

Der Unterschied zwischen Bit- und Symbol-Fehlerwahrscheinlichkeit kann auch weitgehend ignoriert werden: Der Unterschied beträgt im ungünstigsten Fall den Faktor M , angesichts der

typischen Steilheit der Fehlerwahrscheinlichkeitskurven bei guten Kanälen oder in Kombination mit der Kanalcodierung bedeutet der Unterschied nur Bruchteile von einem dB.

Nach Formel (2.4.3) für die Kugelmethode gilt für jedes Modulationsverfahren folgendes:

$$\text{BER} = \text{Funktion} \left(\Delta_0^2 \cdot \frac{ME_c}{N_0} \right) = \text{Funktion} \left(\Delta_0^2 \cdot \frac{E_s}{N_0} \right)$$

Beispiele: Für die niederstufigen PSK-Verfahren gelten folgende Zusammenhänge zwischen E_s und E_b sowie für Δ_0

Modulation	Relation E_s zu E_c		Δ_0
2-PSK	$E_s = E_c$		$\Delta_0 = 2$
4-PSK	$E_s = 2 \cdot E_c$	$E_{s,\text{dB}} = E_{c,\text{dB}} + 3.01$	$\Delta_0 = 1.41$
8-PSK	$E_s = 3 \cdot E_c$	$E_{s,\text{dB}} = E_{c,\text{dB}} + 4.77$	$\Delta_0 = 0.77$

Also gilt für den wesentlichen Term in der BER-Formel:

Modulation	Term für BER über E_c/N_0	Term für BER über E_s/N_0
2-PSK	$\Delta_0^2 \cdot \frac{ME_c}{N_0} = 4 \cdot 1 \cdot \frac{E_c}{N_0}$	$\Delta_0^2 \cdot \frac{E_s}{N_0} = 4 \cdot \frac{E_s}{N_0}$
4-PSK	$\Delta_0^2 \cdot \frac{ME_c}{N_0} = 2 \cdot 2 \cdot \frac{E_c}{N_0}$	$\Delta_0^2 \cdot \frac{E_s}{N_0} = 2 \cdot \frac{E_s}{N_0}$
8-PSK	$\Delta_0^2 \cdot \frac{ME_c}{N_0} = 0.5858 \cdot 3 \cdot \frac{E_c}{N_0}$	$\Delta_0^2 \cdot \frac{E_s}{N_0} = 0.5858 \cdot \frac{E_s}{N_0}$

Daraus ergeben sich folgende Folgerungen:

- Bezüglich E_b/N_0
 - 2-PSK und 4-PSK sind identisch
 - 8-PSK erfordert $10 \cdot \log_{10}(4/(0.5858 \cdot 3)) = 3.57$ dB mehr
- Bezüglich E_s/N_0
 - 4-PSK erfordert $10 \cdot \log_{10}(4/2) = 3.01$ dB mehr gegenüber 2-PSK
 - 8-PSK erfordert $10 \cdot \log_{10}(4/0.5858) = 8.34$ dB mehr gegenüber 2-PSK

Diese Differenzen bzw. Abstände aus dem Beispiel findet man genau so auch wieder in den beiden folgenden Bildern mit Fehlerraten über E_s/N_0 bzw. über $E_c/N_0 = E_b/N_0$:

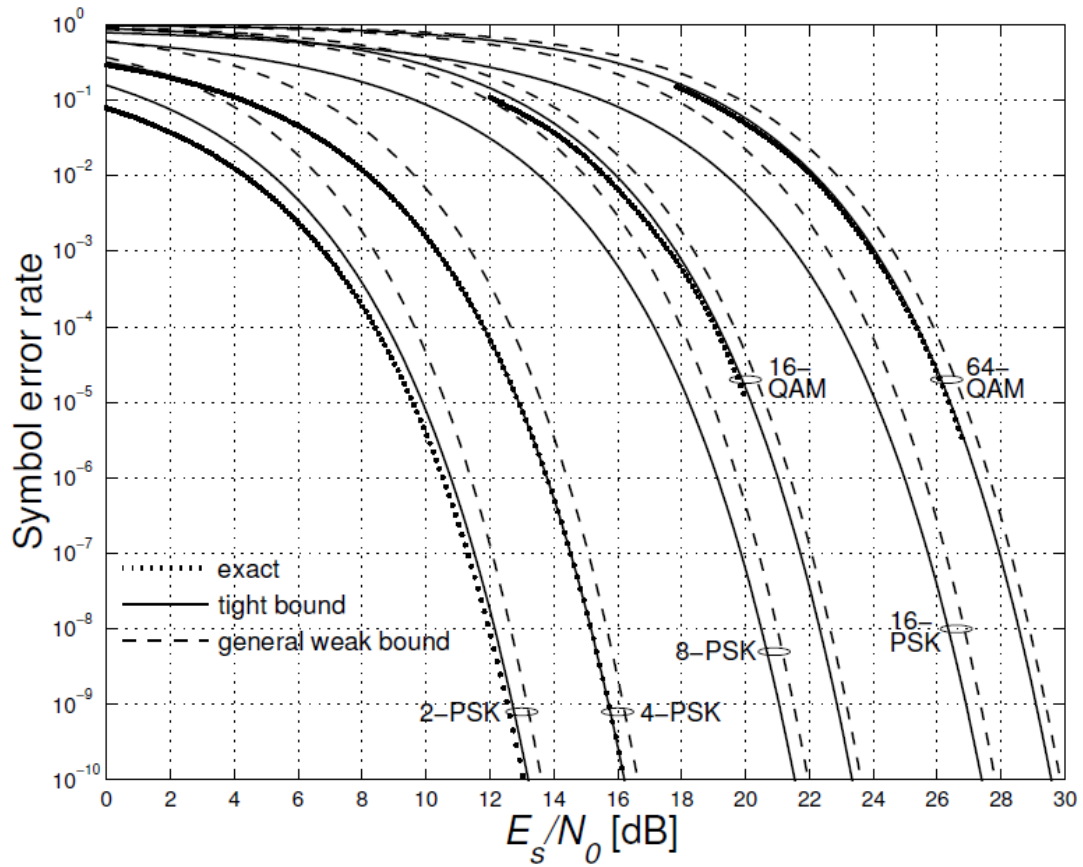


Abb. 2: Fehlerwahrscheinlichkeit für uncodierte PSK und QAM über E_s/N_0

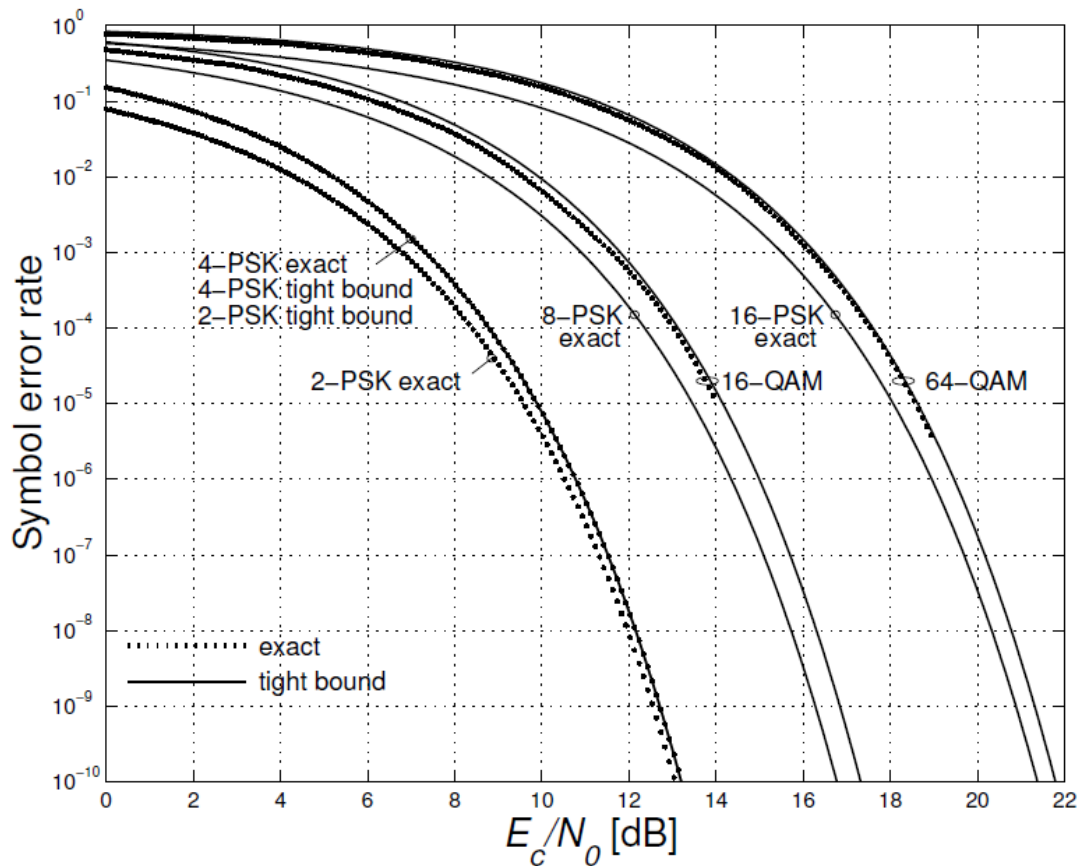


Abb. 3: Fehlerwahrscheinlichkeit für uncodierte PSK und QAM über E_c/N_0

3. BER-Berechnung bei codierter Übertragung

3.1 Allgemeines

Der zusätzlich notwendige Betrag an E_c/N_0 wurde vorangehend schon für die Modulationsverfahren 2-PSK, 4-PSK und 8-PSK ausgerechnet. In den beiden nachfolgenden Bildern wird der zusätzlich notwendige Betrag für weitere Modulationsverfahren aufgelistet.

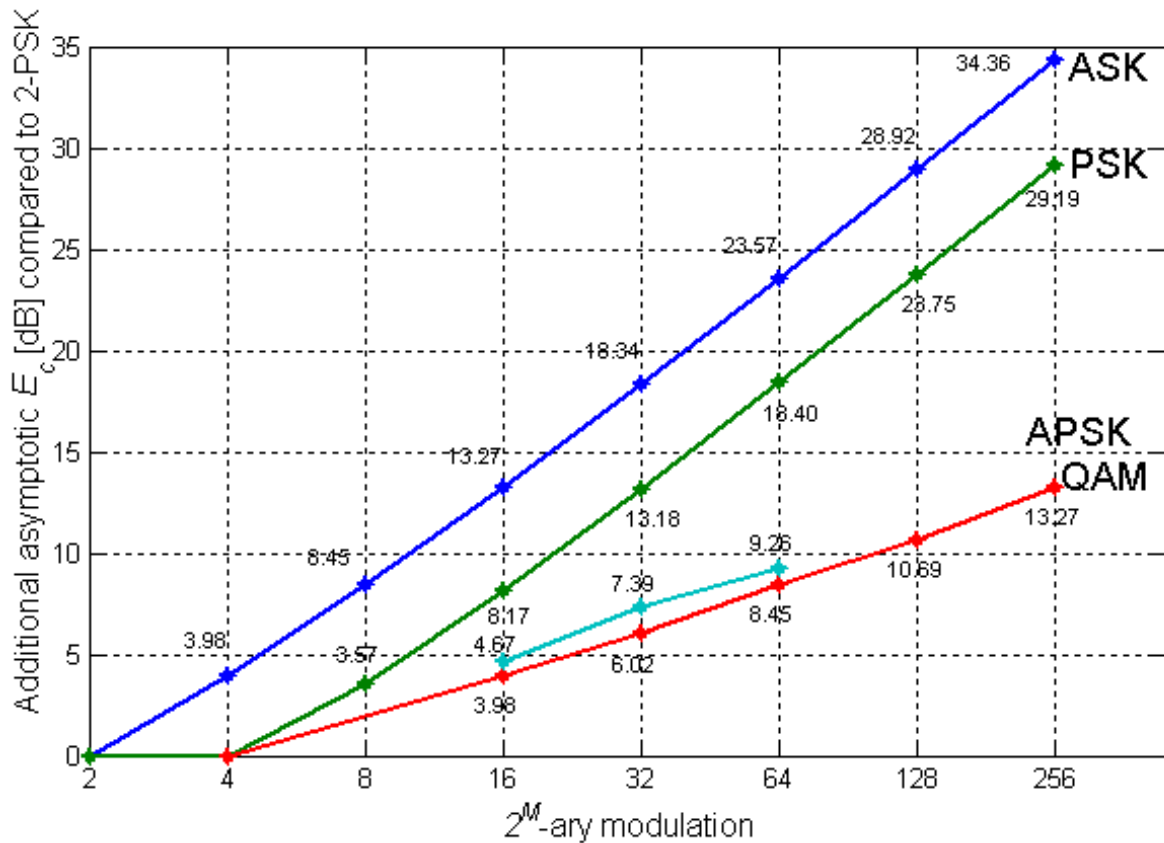


Abb.4 : Notwendiges E_c/N_0 [dB] für 2^M -stufige Modulation im Vergleich zu BPSK (asymptotisch für $E_c/N_0 \rightarrow \infty$)

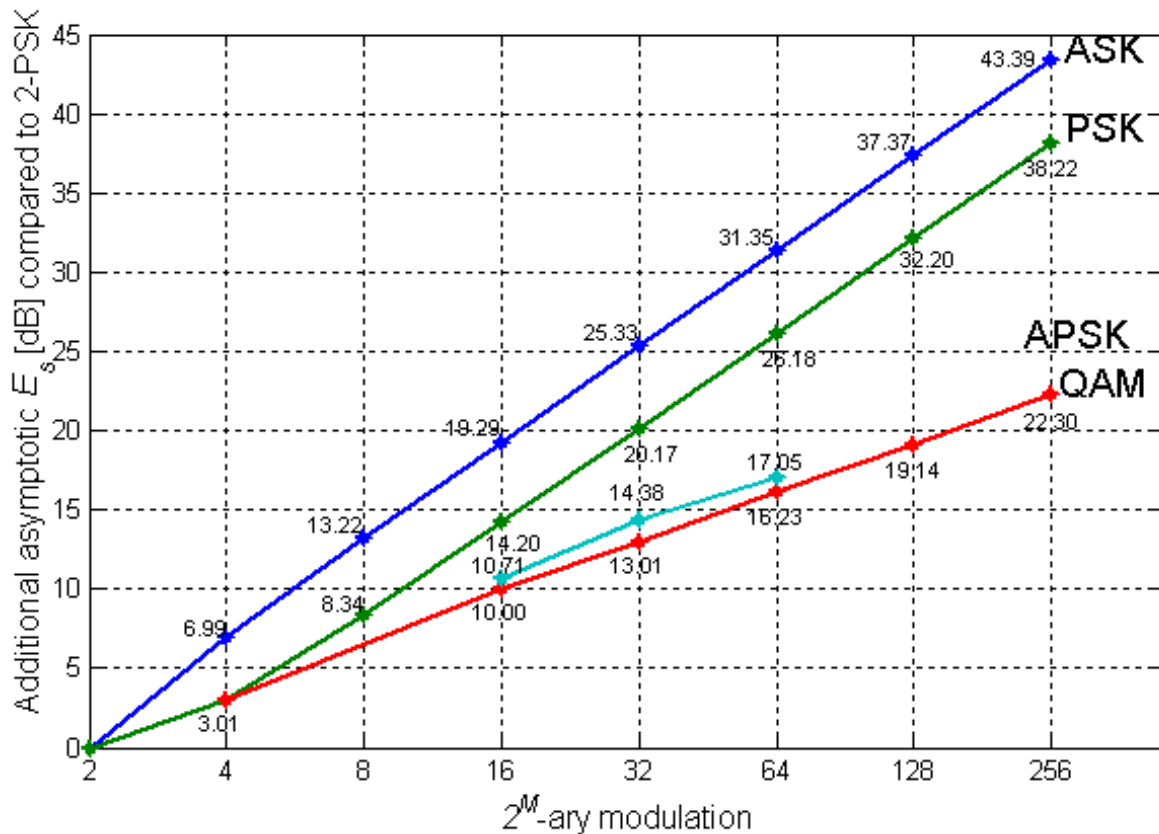


Abb. 5: Notwendiges E_s/N_0 [dB] für 2^M -stufige Modulation im Vergleich zu BPSK (asymptotisch für $E_s/N_0 \rightarrow \infty$)

Beispielsweise gilt $M=4$ für 16-stufige und $M=6$ für 64-stufige Modulation. Entsprechend gilt für das zusätzlich notwendige E_c/N_0 wieder folgender Zusammenhang

$$\frac{E_c}{N_0} = \frac{1}{M} \cdot \frac{E_s}{N_0}$$

Mit dieser Umrechnung kann man Codierungskurven für BPSK einfach auf andere Modulationsverfahren übertragen.

3.2 RS-Code

Die Wort-Fehlerwahrscheinlichkeit P_w für RS-Codes mit der Blocklänge $n=255$ und $t=E$ korrigierbare Fehler ist im nachfolgenden Bild angegeben, und zwar sowohl über E_b/N_0 wie über der Input-Fehlerrate. In beiden Fällen ist ein AWGN mit 2-PSK (binärer Input) und Hard-Decision (binärer Output) unterstellt.

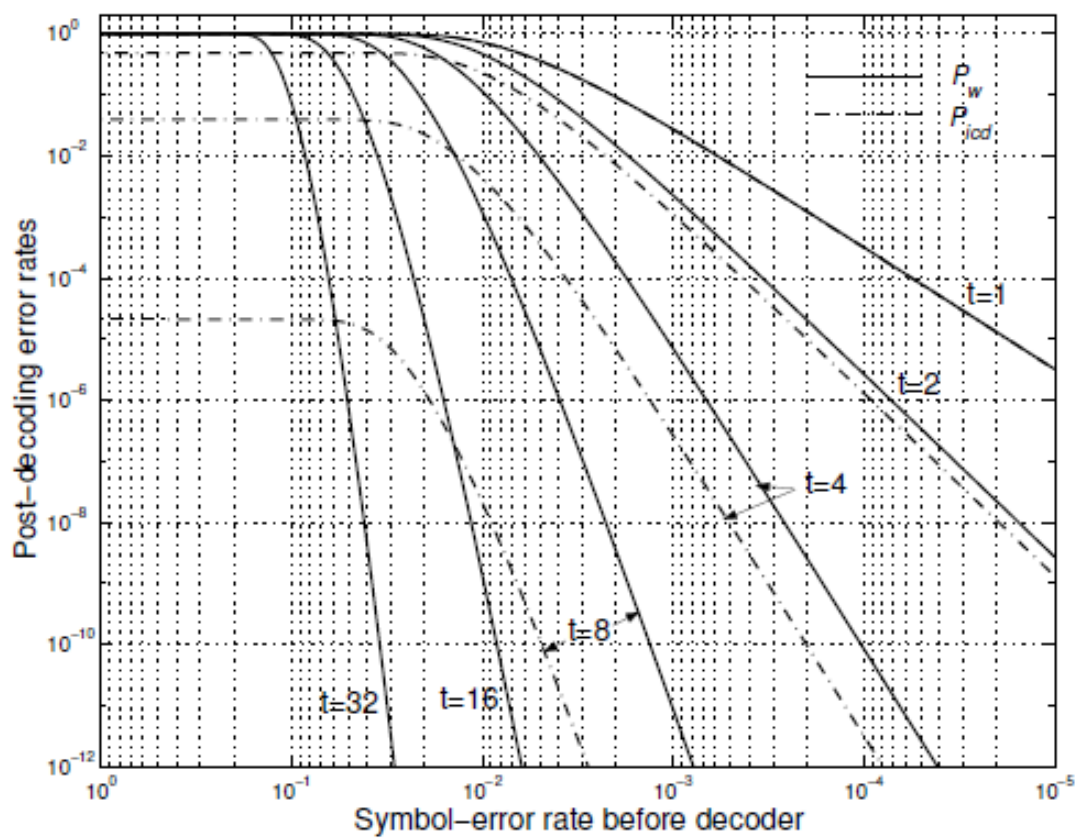
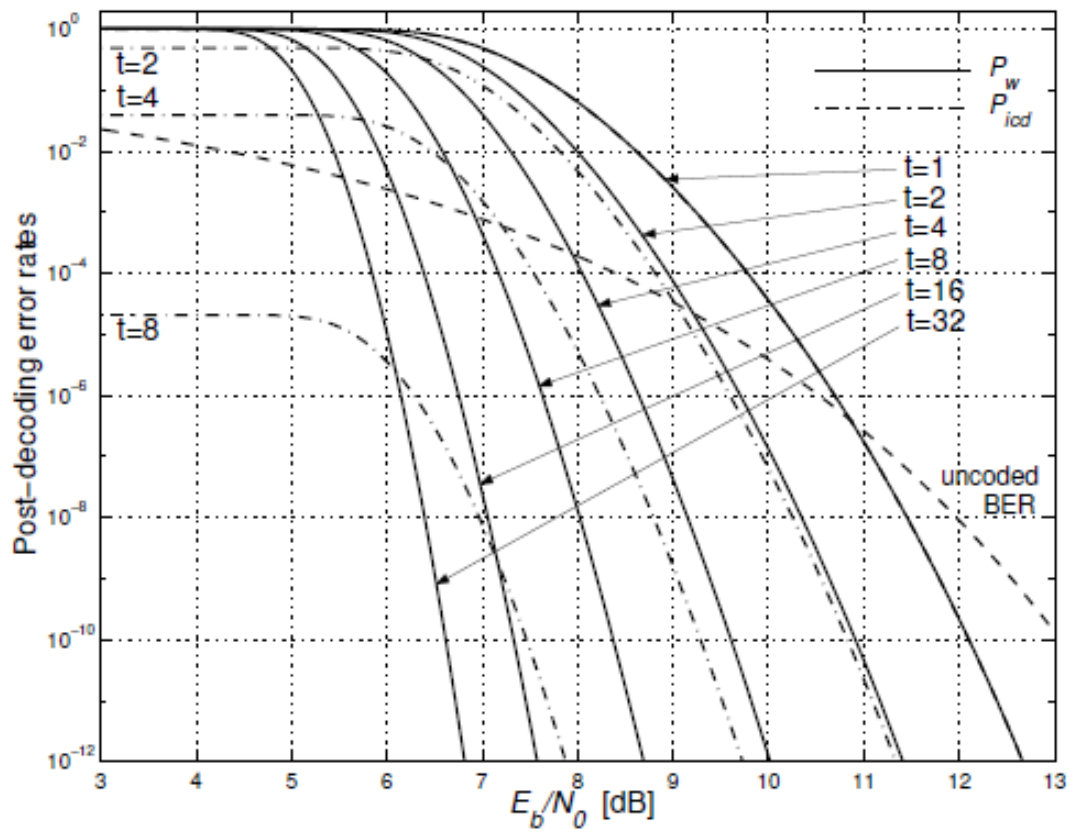


Abb. 6: Wort-Fehlerwahrscheinlichkeit P_w für $RS(255, 255-2t, 2t+1)_{256}$ -Code für verschiedene t über E_b/N_0 bzw. über der Input-BER

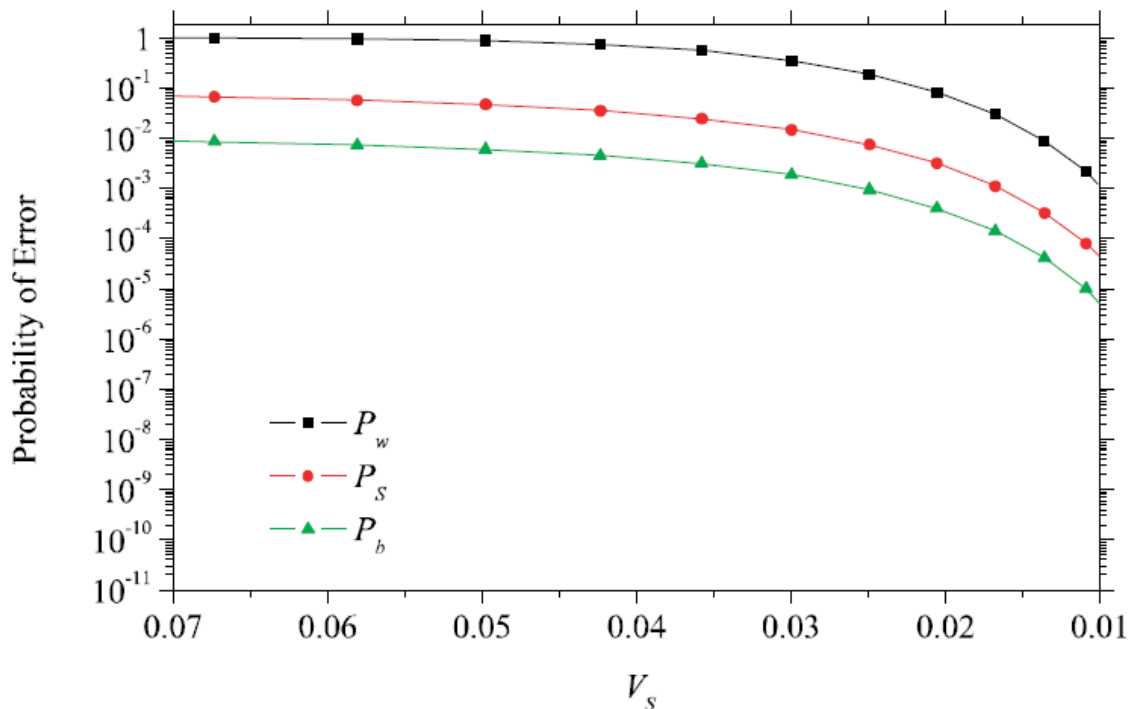


Abb. 7: Wort-/Symbol-/Bit-Fehlerwahrscheinlichkeiten für RS(255, 239, 17)₂₅₆-Code (t=8) über der Input-BER [Quelle: CCSDS 130.1-G-1]

Für das Beispiel t=E=8 und eine Input-Fehlerrate von 10^{-2} beträgt $P_w=10^{-3}$ übereinstimmend in beiden Bildern.

Die Symbol- und Bit-Fehlerwahrscheinlichkeiten sind etwas geringer als die Wort-Fehlerwahrscheinlichkeit, aber wegen des zunehmend steileren Kurvenverlaufs werden diese Differenzen zunehmend unerheblicher.

Beispiel: RS(255,239) mit t=E=8 bei einer Fehlerrate von 10^{-8} :

- Beim AWGN mit binärer Modulation
 - werden etwa $E_b/N_0 = 8$ dB für 10^{-8} benötigt (unter der Annahme Hard-Decision, also auch binärer Output des Demodulators), ...
 - das entspricht $E_s/N_0 = E_c/N_0 = R \cdot E_b/N_0 = 10 \log_{10}(239/255) + 8.00 = -0.28 + 8.00 = 7.72$ dB
- Für 8-PSK
 - erhöht sich das notwendige E_s/N_0 um 8.34 dB nach der vorangehenden Tabelle. Also erfordert 8-PSK nun $E_s/N_0 = 7.72 + 8.34 = 16.06$ dB, ...
 - dies entspricht $E_c/N_0 = 1/M \cdot E_s/N_0 = 10 \log_{10}(1/3) + 16.06 = -4.77 + 16.06 = 11.29$ dB, bzw. ...
 - dies entspricht $E_b/N_0 = 1/(RM) \cdot E_s/N_0 = 10 \log_{10}(255/239 \cdot 1/3) + 16.06 = -4.49 + 16.06 = 11.57$ dB.
- Ein alternativer Rechenweg für 8-PSK liefert das gleiche Ergebnis:
 - E_b/N_0 und E_c/N_0 erhöhen beim Übergang von 2-PSK auf 8-PSK jeweils um 3.57 dB, also auf $E_b/N_0 = 8.00 + 3.57 = 11.57$ und $E_c/N_0 = 7.72 + 3.57 = 11.29$ dB.

Dieses Beispiel ist auch im nachfolgenden Bild illustriert: Von E_b/N_0 für 2-PSK kommt man allein mit der Coderate und der vorangehenden Tabelle für zusätzliches E_s/N_0 zum tatsächlich notwendigen E_s/N_0 .

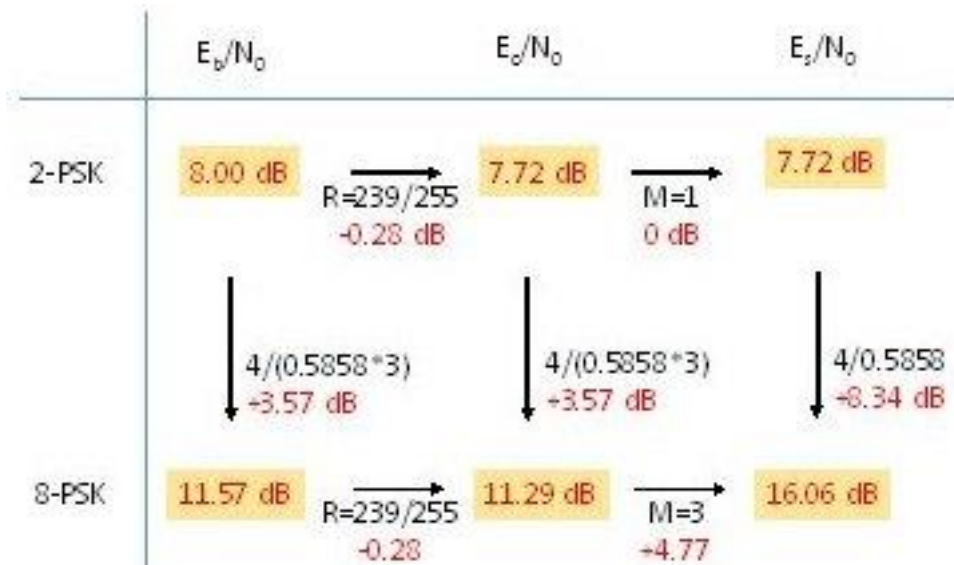


Abb. 8: Illustration zur Umrechnung von 2-PSK auf 8-PSK, ausgehend von 8 dB bei Coderate $R=239/255$

3.3 RS*CC-Code

Siehe CCSDS 130.1-G-1 ab Seite 6-6 für die Konkatenierung von äußerem RS(255,k) mit $t=16$ oder $t=8$ und innerem Faltungscodex (CC) mit Rate R_{CC} und Interleavertiefe I :

Fig 6-5/6-6 zeigt BER/WER für $t=16$, abhängig von Interleavertiefe $I=1 \dots 16$, $R_{CC}=1/2$ fix

Fig 6-7/6-8 zeigt BER/WER für $t=8$, abhängig von Interleavertiefe $I=1 \dots 16$, $R_{CC}=1/2$ fix

Fig 6-9/6-10 zeigt BER/WER für $t=16$, abhängig von CC-Rate $R_{CC}=1/2 \dots 7/8$, $I=5$ fix

Fig 6-11/6-12 zeigt BER/WER für $t=8$, abhängig von CC-Rate $R_{CC}=1/2 \dots 7/8$, $I=5$ fix