# Quantenkanalkapazität

## Seminar: Quanteninformation - Entwicklung, Protokolle, Technologien

Marco Möller

24.05.2007

# Inhalt


1. Information
   - Definition
   - Bedingte -/ Transinformation
2. Kodierung ohne Rauschen
   - Problemstellung
   - typische Unterräume
   - Satz von Schumacher
   - Beispiel: Schumacher-Kompression
3. Kodierung unter Rauschen
   - klassisch
   - QM
4. Ende
   - Literatur
   - Fragen

# Gliederung


1. Information
2. Kodierung ohne Rauschen
3. Kodierung unter Rauschen
4. Ende

# Information

## Definitionen

- *Shannon Entropie*

$$H(X) \equiv -\sum_x p_x \log p_x$$

- *Von Neumann Entropie*

$$S(X) \equiv tr(\rho \log \rho)$$

## Korrolar (Shannon Entropie ist Spezialfall)

*falls* $\rho = \sum_x p_x |x\rangle\langle x|$ *und* $\langle x|y\rangle = \delta_{xy}$

$$S(X) = H(X)$$

- Einheit ist *bit* bzw. *qubit*
- log ist zur Basis 2

# Information

## Definitionen

- *Shannon Entropie*

$$H(X) \equiv -\sum_x p_x \log p_x$$

- *Von Neumann Entropie*

$$S(X) \equiv tr(\rho \log \rho)$$

## Korrolar (Shannon Entropie ist Spezialfall)

*falls* $\rho = \sum_x p_x |x\rangle\langle x|$ *und* $\langle x|y\rangle = \delta_{xy}$

$$S(X) = H(X)$$

- Einheit ist *bit* bzw. *qubit*
- log ist zur Basis 2

# Information

## Definitionen

- *Shannon Entropie*

$$H(X) \equiv -\sum_x p_x \log p_x$$

- *Von Neumann Entropie*

$$S(X) \equiv tr(\rho \log \rho)$$

## Korrolar (Shannon Entropie ist Spezialfall)

*falls* $\rho = \sum_x p_x |x\rangle\langle x|$ *und* $\langle x|y\rangle = \delta_{xy}$

$$S(X) = H(X)$$

- Einheit ist *bit* bzw. *qubit*
- log ist zur Basis 2

# Information

## Definitionen

- *Shannon Entropie*

$$H(X) \equiv -\sum_x p_x \log p_x$$

- *Von Neumann Entropie*

$$S(X) \equiv tr(\rho \log \rho)$$

## Korrolar (Shannon Entropie ist Spezialfall)

*falls* $\rho = \sum_x p_x |x\rangle\langle x|$ *und* $\langle x|y\rangle = \delta_{xy}$

$$S(X) = H(X)$$

- Einheit ist *bit* bzw. *qubit*
- log ist zur Basis 2

# weitere Informationen

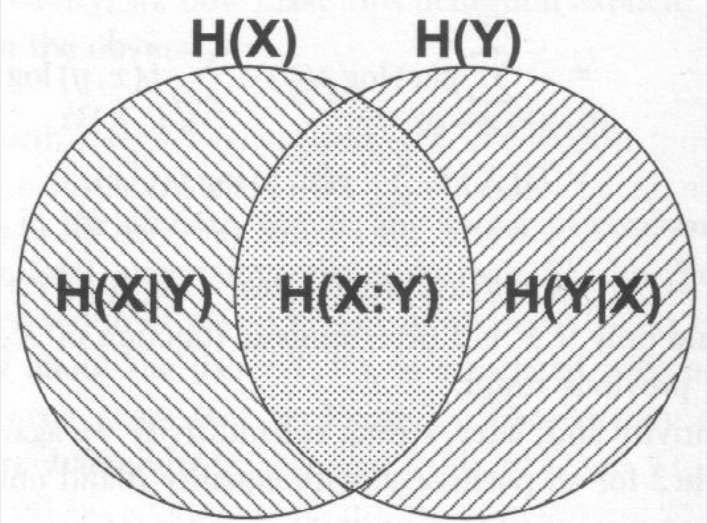
## Definitionen

- *bedingte Entropie*

$$H(X|Y) = H(X,Y) - H(Y)$$

- *Transinformation (mutual information)*

$$\begin{aligned} H(X:Y) &= H(X) + H(Y) - H(X,Y) \\ &= H(X) - H(X|Y) \end{aligned}$$




Analog für $S(X)$

# weitere Informationen

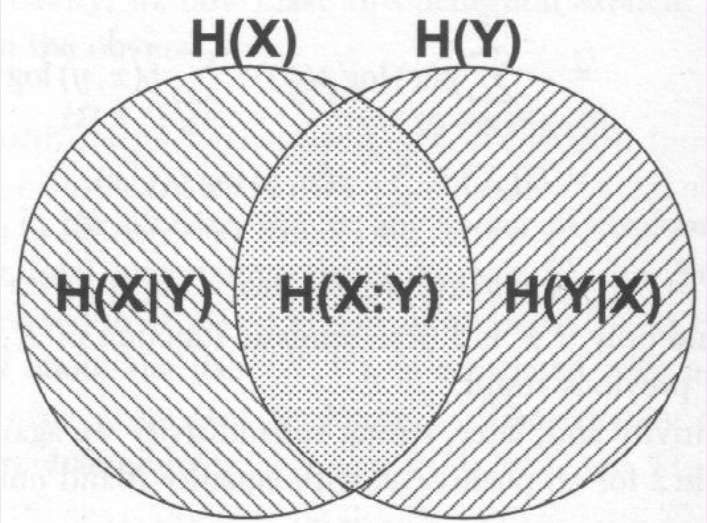
## Definitionen

- *bedingte Entropie*

$$H(X|Y) = H(X,Y) - H(Y)$$

- *Transinformation (mutual information)*

$$\begin{aligned} H(X:Y) &= H(X) + H(Y) - H(X,Y) \\ &= H(X) - H(X|Y) \end{aligned}$$




Analog für $S(X)$

# weitere Informationen

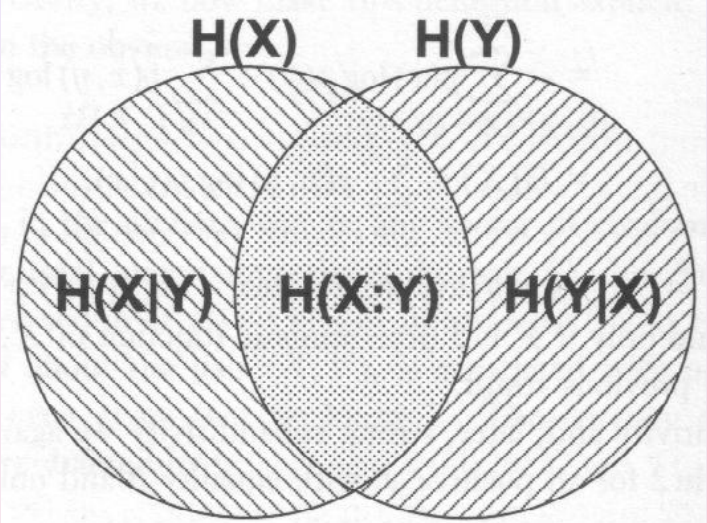
## Definitionen

- *bedingte Entropie*

$$H(X|Y) = H(X,Y) - H(Y)$$

- *Transinformation (mutual information)*

$$\begin{aligned} H(X:Y) &= H(X) + H(Y) - H(X,Y) \\ &= H(X) - H(X|Y) \end{aligned}$$




Analog für $S(X)$

# Maße für Operatoren

- Wie gleich sind sich zwei Quantenzustände? → benötigen besseres Kriterium als Identität, um Grenzwerte bilden zu können
- Viele Möglichkeiten, so etwas zu definieren

**Definition (Fidelity)**

- *Statische Fidelity:*

$$F(\rho,\sigma) = tr\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}}$$

- *Dynamische Fidelity ($\rho$ verglichen mit $\varepsilon(\rho)$): Sei $\varepsilon$ eine spurerhaltene Quantenoperation mit den Kraus-Operatoren $E_i$*

$$F(\rho,\epsilon) = \sum_i |tr(\rho E_i)|^2$$

- $F(\rho, \mathrm{id}) = 1$

# Maße für Operatoren

- Wie gleich sind sich zwei Quantenzustände? $\rightarrow$ benötigen besseres Kriterium als Identität, um Grenzwerte bilden zu können
- Viele Möglichkeiten, so etwas zu definieren

**Definition (Fidelity)**

- *Statische Fidelity:*

$$F(\rho, \sigma) = tr\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}}$$

- *Dynamische Fidelity ($\rho$ verglichen mit $\varepsilon(\rho)$): Sei $\varepsilon$ eine spurerhaltene Quantenoperation mit den Kraus-Operatoren $E_i$*

$$F(\rho, \epsilon) = \sum_i |tr(\rho E_i)|^2$$

- $F(\rho, \mathrm{id}) = 1$

# Maße für Operatoren

- Wie gleich sind sich zwei Quantenzustände? $\rightarrow$ benötigen besseres Kriterium als Identität, um Grenzwerte bilden zu können
- Viele Möglichkeiten, so etwas zu definieren

## Definition (Fidelity)

- *Statische Fidelity:*

$$F(\rho, \sigma) = tr\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}}$$

- *Dynamische Fidelity ($\rho$ verglichen mit $\varepsilon(\rho)$): Sei $\varepsilon$ eine spurerhaltene Quantenoperation mit den Kraus-Operatoren $E_i$*

$$F(\rho, \epsilon) = \sum_i |tr(\rho E_i)|^2$$

- $F(\rho, \mathrm{id}) = 1$

# Maße für Operatoren

- Wie gleich sind sich zwei Quantenzustände? $\rightarrow$ benötigen besseres Kriterium als Identität, um Grenzwerte bilden zu können
- Viele Möglichkeiten, so etwas zu definieren

## Definition (Fidelity)

- *Statische Fidelity:*

$$F(\rho, \sigma) = tr\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}}$$

- *Dynamische Fidelity ($\rho$ verglichen mit $\varepsilon(\rho)$): Sei $\varepsilon$ eine spurerhaltene Quantenoperation mit den Kraus-Operatoren $E_i$*

$$F(\rho, \epsilon) = \sum_i |tr(\rho E_i)|^2$$

- $F(\rho, \mathrm{id}) = 1$

# Maße für Operatoren

- Wie gleich sind sich zwei Quantenzustände? $\rightarrow$ benötigen besseres Kriterium als Identität, um Grenzwerte bilden zu können
- Viele Möglichkeiten, so etwas zu definieren

## Definition (Fidelity)

- *Statische Fidelity:*

$$F(\rho, \sigma) = tr\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}}$$

- *Dynamische Fidelity ($\rho$ verglichen mit $\varepsilon(\rho)$): Sei $\varepsilon$ eine spurerhaltene Quantenoperation mit den Kraus-Operatoren $E_i$*

$$F(\rho, \epsilon) = \sum_i |tr(\rho E_i)|^2$$

- $F(\rho, \mathrm{id}) = 1$

# Maße für Operatoren

- Wie gleich sind sich zwei Quantenzustände? $\rightarrow$ benötigen besseres Kriterium als Identität, um Grenzwerte bilden zu können
- Viele Möglichkeiten, so etwas zu definieren

## Definition (Fidelity)

- *Statische Fidelity:*

$$F(\rho, \sigma) = tr\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}}$$

- *Dynamische Fidelity ($\rho$ verglichen mit $\varepsilon(\rho)$): Sei $\varepsilon$ eine spurerhaltene Quantenoperation mit den Kraus-Operatoren $E_i$*

$$F(\rho, \epsilon) = \sum_i |tr(\rho E_i)|^2$$

- $F(\rho, \mathrm{id}) = 1$

# Gliederung


1. Information
2. Kodierung ohne Rauschen
3. Kodierung unter Rauschen
4. Ende

# Problemstellung klassisch

**Quelle** entspricht einer Folge von Zufallsvariablen $X_1, X_2, \ldots$ über einem endlichen Alphabet

- die Werte entsprechen einer möglichen Ausgabe der Quelle
- die $X_i$ sind identisch verteilt und unabhängig zueinander (indipendent, identical distributed = i.i.d.)

**Frage:** wie oft müsste man einen störungsfreien binären Kanal mindestens im Mittel pro generierten Symbol benutzen ($\hat{=}$ Rate).

- Kann auch als *Kompression* aufgefasst werden.

**Satz (Shannon: rauschfreie Kanalkodierung)**

*Sei $\{X_i\}$ eine i.i.d Informationsquelle mit einer Entropierate $H(X)$. Sei $R > H(X)$. Dann existiert ein zuverlässiges Kompressionsschema der Rate $R$ für diese Quelle. Umgekehrt ist wäre jedes Kompressionsschema für $R < H(X)$ nicht zuverlässig.*

# Problemstellung klassisch

**Quelle** entspricht einer Folge von Zufallsvariablen $X_1, X_2, \ldots$ über einem endlichen Alphabet

- die Werte entsprechen einer möglichen Ausgabe der Quelle
- die $X_i$ sind identisch verteilt und unabhängig zueinander (indipendent, identical distributed = i.i.d.)

**Frage:** wie oft müsste man einen störungsfreien binären Kanal mindestens im Mittel pro generierten Symbol benutzen ($\hat{=}$ Rate).

- Kann auch als *Kompression* aufgefasst werden.

**Satz (Shannon: rauschfreie Kanalkodierung)**

*Sei $\{X_i\}$ eine i.i.d Informationsquelle mit einer Entropierate $H(X)$. Sei $R > H(X)$. Dann existiert ein zuverlässiges Kompressionsschema der Rate $R$ für diese Quelle. Umgekehrt ist wäre jedes Kompressionsschema für $R < H(X)$ nicht zuverlässig.*

# Problemstellung klassisch

**Quelle** entspricht einer Folge von Zufallsvariablen $X_1, X_2, \ldots$ über einem endlichen Alphabet

- die Werte entsprechen einer möglichen Ausgabe der Quelle
- die $X_i$ sind identisch verteilt und unabhängig zueinander (indipendent, identical distributed = i.i.d.)

**Frage:** wie oft müsste man einen störungsfreien binären Kanal mindestens im Mittel pro generierten Symbol benutzen ($\hat{=}$ Rate).

- Kann auch als *Kompression* aufgefasst werden.

**Satz (Shannon: rauschfreie Kanalkodierung)**

*Sei $\{X_i\}$ eine i.i.d Informationsquelle mit einer Entropierate $H(X)$. Sei $R > H(X)$. Dann existiert ein zuverlässiges Kompressionsschema der Rate $R$ für diese Quelle. Umgekehrt ist wäre jedes Kompressionsschema für $R < H(X)$ nicht zuverlässig.*

# Problemstellung klassisch

**Quelle** entspricht einer Folge von Zufallsvariablen $X_1, X_2, \ldots$ über einem endlichen Alphabet

- die Werte entsprechen einer möglichen Ausgabe der Quelle
- die $X_i$ sind identisch verteilt und unabhängig zueinander (**i**ndipendent, **i**dentical **d**istributed = i.i.d.)

**Frage:** wie oft müsste man einen störungsfreien binären Kanal mindestens im Mittel pro generierten Symbol benutzen ($\hat{=}$ Rate).

- Kann auch als *Kompression* aufgefasst werden.

**Satz (Shannon: rauschfreie Kanalkodierung)**

*Sei $\{X_i\}$ eine i.i.d Informationsquelle mit einer Entropierate $H(X)$. Sei $R > H(X)$. Dann existiert ein zuverlässiges Kompressionsschema der Rate $R$ für diese Quelle. Umgekehrt ist wäre jedes Kompressionsschema für $R < H(X)$ nicht zuverlässig.*

# Problemstellung klassisch

**Quelle** entspricht einer Folge von Zufallsvariablen $X_1, X_2, \ldots$ über einem endlichen Alphabet

- die Werte entsprechen einer möglichen Ausgabe der Quelle
- die $X_i$ sind identisch verteilt und unabhängig zueinander (**i**ndipendent, **i**dentical **d**istributed = i.i.d.)

**Frage:** wie oft müsste man einen störungsfreien binären Kanal mindestens im Mittel pro generierten Symbol benutzen ($\hat{=}$ Rate).

- Kann auch als *Kompression* aufgefasst werden.

**Satz (Shannon: rauschfreie Kanalkodierung)**

*Sei $\{X_i\}$ eine i.i.d Informationsquelle mit einer Entropierate $H(X)$. Sei $R > H(X)$. Dann existiert ein zuverlässiges Kompressionsschema der Rate $R$ für diese Quelle. Umgekehrt ist wäre jedes Kompressionsschema für $R < H(X)$ nicht zuverlässig.*

# Problemstellung klassisch

**Quelle** entspricht einer Folge von Zufallsvariablen $X_1, X_2, \ldots$ über einem endlichen Alphabet

- die Werte entsprechen einer möglichen Ausgabe der Quelle
- die $X_i$ sind identisch verteilt und unabhängig zueinander (**i**ndipendent, **i**dentical **d**istributed = i.i.d.)

**Frage:** wie oft müsste man einen störungsfreien binären Kanal mindestens im Mittel pro generierten Symbol benutzen ($\hat{=}$ Rate).

- Kann auch als *Kompression* aufgefasst werden.

**Satz (Shannon: rauschfreie Kanalkodierung)**

*Sei $\{X_i\}$ eine i.i.d Informationsquelle mit einer Entropierate $H(X)$. Sei $R > H(X)$. Dann existiert ein zuverlässiges Kompressionsschema der Rate $R$ für diese Quelle. Umgekehrt ist wäre jedes Kompressionsschema für $R < H(X)$ nicht zuverlässig.*

# Problemstellung klassisch

**Quelle** entspricht einer Folge von Zufallsvariablen $X_1, X_2, \ldots$ über einem endlichen Alphabet

- die Werte entsprechen einer möglichen Ausgabe der Quelle
- die $X_i$ sind identisch verteilt und unabhängig zueinander (**i**ndipendent, **i**dentical **d**istributed = i.i.d.)

**Frage:** wie oft müsste man einen störungsfreien binären Kanal mindestens im Mittel pro generierten Symbol benutzen ($\hat{=}$ Rate).

- Kann auch als *Kompression* aufgefasst werden.

**Satz (Shannon: rauschfreie Kanalkodierung)**

*Sei $\{X_i\}$ eine i.i.d Informationsquelle mit einer Entropierate $H(X)$. Sei $R > H(X)$. Dann existiert ein zuverlässiges Kompressionsschema der Rate $R$ für diese Quelle. Umgekehrt ist wäre jedes Kompressionsschema für $R < H(X)$ nicht zuverlässig.*

# Motivation: typische Sequenzen

- Im binären Fall:
  $p(x_1, \ldots, x_n) = p(x_1)\, p(x_2) \cdots p(x_n) \approx p^{np} (1-p)^{n(1-p)}$
- $\log p(x_1, \ldots, x_n) \approx np \log p + (1-p)\, n \log(1-p) = -nH(X)$
- $p(x_1, \ldots, x_n) \approx 2^{-nH(X)}$
- Es kann also nur ca. $2^{nH(X)}$ Sequenzen geben, die häufig vorkommen.
- $\Rightarrow$ ein um den Faktor $H(X)$ kürzere umkodierten Sequenz reicht aus, um die typischen Sequenzen zu kodieren.
- Kodierung, die das ausnutzt: Huffmann-Kodierung[2]

# Motivation: typische Sequenzen

- Im binären Fall:
  $p(x_1, \ldots, x_n) = p(x_1)\, p(x_2) \cdots p(x_n) \approx p^{np} (1-p)^{n(1-p)}$
- $\log p(x_1, \ldots, x_n) \approx np \log p + (1-p)\, n \log(1-p) = -nH(X)$
- $p(x_1, \ldots, x_n) \approx 2^{-nH(X)}$
- Es kann also nur ca. $2^{nH(X)}$ Sequenzen geben, die häufig vorkommen.
- $\Rightarrow$ ein um den Faktor $H(X)$ kürzere umkodierten Sequenz reicht aus, um die typischen Sequenzen zu kodieren.
- Kodierung, die das ausnutzt: Huffmann-Kodierung[2]

# Motivation: typische Sequenzen

- Im binären Fall:
  $p(x_1, \ldots, x_n) = p(x_1)\, p(x_2) \cdots p(x_n) \approx p^{np} (1-p)^{n(1-p)}$
- $\log p(x_1, \ldots, x_n) \approx np \log p + (1-p)\, n \log(1-p) = -nH(X)$
- $p(x_1, \ldots, x_n) \approx 2^{-nH(X)}$
- Es kann also nur ca. $2^{nH(X)}$ Sequenzen geben, die häufig vorkommen.
- $\Rightarrow$ ein um den Faktor $H(X)$ kürzere umkodierten Sequenz reicht aus, um die typischen Sequenzen zu kodieren.
- Kodierung, die das ausnutzt: Huffmann-Kodierung[2]

# Motivation: typische Sequenzen

- Im binären Fall:
  $p(x_1, \ldots, x_n) = p(x_1)\, p(x_2) \cdots p(x_n) \approx p^{np} (1-p)^{n(1-p)}$
- $\log p(x_1, \ldots, x_n) \approx np \log p + (1-p)\, n \log(1-p) = -nH(X)$
- $p(x_1, \ldots, x_n) \approx 2^{-nH(X)}$
- Es kann also nur ca. $2^{nH(X)}$ Sequenzen geben, die häufig vorkommen.
- $\Rightarrow$ ein um den Faktor $H(X)$ kürzere umkodierten Sequenz reicht aus, um die typischen Sequenzen zu kodieren.
- Kodierung, die das ausnutzt: Huffmann-Kodierung[2]

# Motivation: typische Sequenzen

- Im binären Fall:
  $p(x_1, \ldots, x_n) = p(x_1)\, p(x_2) \cdots p(x_n) \approx p^{np} (1-p)^{n(1-p)}$
- $\log p(x_1, \ldots, x_n) \approx np \log p + (1-p)\, n \log(1-p) = -nH(X)$
- $p(x_1, \ldots, x_n) \approx 2^{-nH(X)}$
- Es kann also nur ca. $2^{nH(X)}$ Sequenzen geben, die häufig vorkommen.
- $\Rightarrow$ ein um den Faktor $H(X)$ kürzere umkodierten Sequenz reicht aus, um die typischen Sequenzen zu kodieren.
- Kodierung, die das ausnutzt: Huffmann-Kodierung[2]

# Motivation: typische Sequenzen

- Im binären Fall:
  $p(x_1, \ldots, x_n) = p(x_1)\, p(x_2) \cdots p(x_n) \approx p^{np} (1-p)^{n(1-p)}$
- $\log p(x_1, \ldots, x_n) \approx np \log p + (1-p)\, n \log(1-p) = -nH(X)$
- $p(x_1, \ldots, x_n) \approx 2^{-nH(X)}$
- Es kann also nur ca. $2^{nH(X)}$ Sequenzen geben, die häufig vorkommen.
- $\Rightarrow$ ein um den Faktor $H(X)$ kürzere umkodierten Sequenz reicht aus, um die typischen Sequenzen zu kodieren.
- Kodierung, die das ausnutzt: Huffmann-Kodierung[2]

# Problemstellung quantenmechanisch

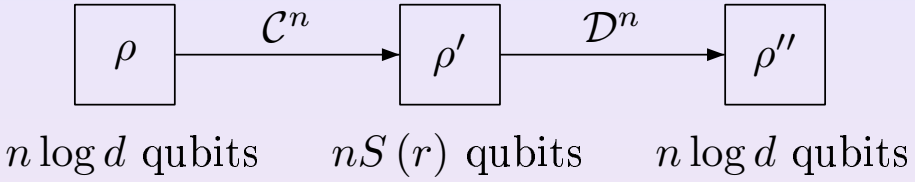



Abbildung: (De)Kompression einer Quanteninformationsquelle [1]

**Quelle** (i.i.d.) wird durch einen Hilbertraum $H$ und einem Dichteoperator $\rho$ darauf beschrieben

**Kompressionsschema** der Rate $R$ besteht aus zwei Familien von Quantenoperationen $\mathcal{C}^n$ und $\mathcal{D}^n$

$$\text{Kompression } \mathcal{C}^n : H^{\otimes n} \to 2^{Rn}$$
$$\text{Dekompression } \mathcal{D}^n : 2^{Rn} \to H^{\otimes n}$$

**Zuverlässig** ist ein Kompressionsschema, wenn die Fehlerrate $\mathcal{D}^n \circ \mathcal{C}^n$ auf $\rho^{\otimes n}$ für große $n$ gegen 0 geht, bzw. $F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) \to 1$.

**Verallgemeinerung** von klassischem Fall

# Problemstellung quantenmechanisch

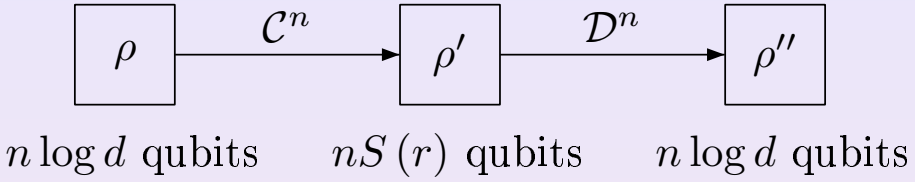



Abbildung: (De)Kompression einer Quanteninformationsquelle [1]

**Quelle** (i.i.d.) wird durch einen Hilbertraum $H$ und einem Dichteoperator $\rho$ darauf beschrieben

**Kompressionsschema** der Rate $R$ besteht aus zwei Familien von Quantenoperationen $\mathcal{C}^n$ und $\mathcal{D}^n$

$$\text{Kompression } \mathcal{C}^n : H^{\otimes n} \to 2^{Rn}$$
$$\text{Dekompression } \mathcal{D}^n : 2^{Rn} \to H^{\otimes n}$$

**Zuverlässig** ist ein Kompressionsschema, wenn die Fehlerrate $\mathcal{D}^n \circ \mathcal{C}^n$ auf $\rho^{\otimes n}$ für große $n$ gegen 0 geht, bzw. $F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) \to 1$.

**Verallgemeinerung** von klassischem Fall

# Problemstellung quantenmechanisch

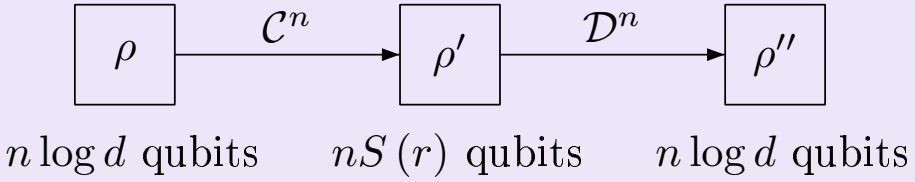



Abbildung: (De)Kompression einer Quanteninformationsquelle [1]

**Quelle** (i.i.d.) wird durch einen Hilbertraum $H$ und einem Dichteoperator $\rho$ darauf beschrieben

**Kompressionsschema** der Rate $R$ besteht aus zwei Familien von Quantenoperationen $\mathcal{C}^n$ und $\mathcal{D}^n$

Kompression $\mathcal{C}^n : H^{\otimes n} \to 2^{Rn}$
Dekompression $\mathcal{D}^n : 2^{Rn} \to H^{\otimes n}$

**Zuverlässig** ist ein Kompressionsschema, wenn die Fehlerrate $\mathcal{D}^n \circ \mathcal{C}^n$ auf $\rho^{\otimes n}$ für große $n$ gegen 0 geht, bzw. $F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) \to 1$.

**Verallgemeinerung** von klassischem Fall

# Problemstellung quantenmechanisch

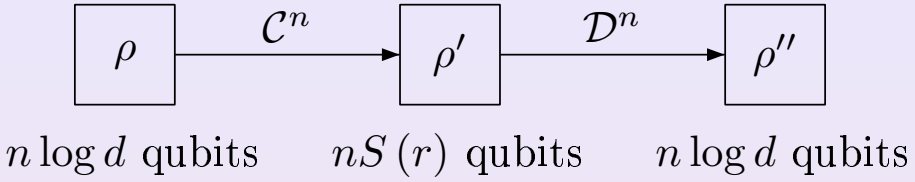



Abbildung: (De)Kompression einer Quanteninformationsquelle [1]

**Quelle** (i.i.d.) wird durch einen Hilbertraum $H$ und einem Dichteoperator $\rho$ darauf beschrieben

**Kompressionsschema** der Rate $R$ besteht aus zwei Familien von Quantenoperationen $\mathcal{C}^n$ und $\mathcal{D}^n$

**Kompression** $\mathcal{C}^n : H^{\otimes n} \to 2^{Rn}$

**Dekompression** $\mathcal{D}^n : 2^{Rn} \to H^{\otimes n}$

**Zuverlässig** ist ein Kompressionsschema, wenn die Fehlerrate $\mathcal{D}^n \circ \mathcal{C}^n$ auf $\rho^{\otimes n}$ für große $n$ gegen 0 geht, bzw. $F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) \to 1$.

**Verallgemeinerung** von klassischem Fall

# Problemstellung quantenmechanisch

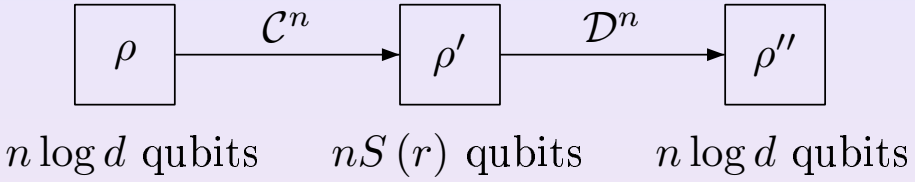



Abbildung: (De)Kompression einer Quanteninformationsquelle [1]

**Quelle** (i.i.d.) wird durch einen Hilbertraum $H$ und einem Dichteoperator $\rho$ darauf beschrieben

**Kompressionsschema** der Rate $R$ besteht aus zwei Familien von Quantenoperationen $\mathcal{C}^n$ und $\mathcal{D}^n$

**Kompression** $\mathcal{C}^n : H^{\otimes n} \to 2^{Rn}$
**Dekompression** $\mathcal{D}^n : 2^{Rn} \to H^{\otimes n}$

**Zuverlässig** ist ein Kompressionsschema, wenn die Fehlerrate $\mathcal{D}^n \circ \mathcal{C}^n$ auf $\rho^{\otimes n}$ für große $n$ gegen 0 geht, bzw. $F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) \to 1$.

**Verallgemeinerung** von klassischem Fall

# Problemstellung quantenmechanisch

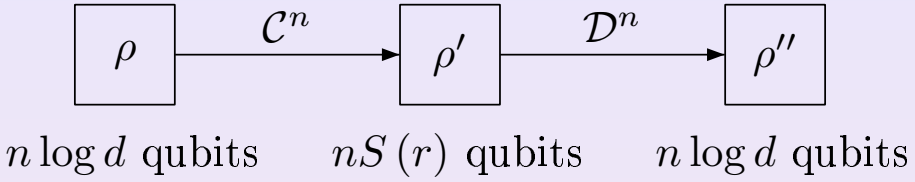



Abbildung: (De)Kompression einer Quanteninformationsquelle [1]

**Quelle** (i.i.d.) wird durch einen Hilbertraum $H$ und einem Dichteoperator $\rho$ darauf beschrieben

**Kompressionsschema** der Rate $R$ besteht aus zwei Familien von Quantenoperationen $\mathcal{C}^n$ und $\mathcal{D}^n$

**Kompression** $\mathcal{C}^n : H^{\otimes n} \to 2^{Rn}$
**Dekompression** $\mathcal{D}^n : 2^{Rn} \to H^{\otimes n}$

**Zuverlässig** ist ein Kompressionsschema, wenn die Fehlerrate $\mathcal{D}^n \circ \mathcal{C}^n$ auf $\rho^{\otimes n}$ für große $n$ gegen 0 geht, bzw. $F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) \to 1$.

**Verallgemeinerung** von klassischem Fall

# Problemstellung quantenmechanisch

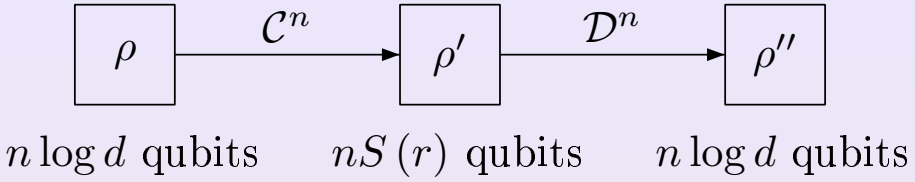



Abbildung: (De)Kompression einer Quanteninformationsquelle [1]

**Quelle** (i.i.d.) wird durch einen Hilbertraum $H$ und einem Dichteoperator $\rho$ darauf beschrieben

**Kompressionsschema** der Rate $R$ besteht aus zwei Familien von Quantenoperationen $\mathcal{C}^n$ und $\mathcal{D}^n$

**Kompression** $\mathcal{C}^n : H^{\otimes n} \to 2^{Rn}$
**Dekompression** $\mathcal{D}^n : 2^{Rn} \to H^{\otimes n}$

**Zuverlässig** ist ein Kompressionsschema, wenn die Fehlerrate $\mathcal{D}^n \circ \mathcal{C}^n$ auf $\rho^{\otimes n}$ für große $n$ gegen 0 geht, bzw. $F\left(\rho^{\otimes n}, \mathcal{D}^n \circ C^n\right) \to 1$.

**Verallgemeinerung** von klassischem Fall

# Gesetz der großen Zahlen

**Satz (Gesetz der großen Zahlen)**

*Seien $X_1, X_2, \ldots$ i.i.d. mit endlichem ersten und zweiten Moment. Dann ist für jedes $\epsilon > 0$*

$$\lim_{n\to\infty} p\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - E(X)\right| \leq \epsilon\right) = 1$$

# $\epsilon$-typisch

## Definition ($\epsilon$-typisch)

*Sei* $\rho = \sum_x p(x) |x\rangle \langle x|$ *orthonormale Dekomposition. Eine Sequenz* $x_1, \ldots, x_n$ *wird* $\epsilon$*-typisch genannt, falls*

$$\left| \frac{1}{n} \log \left( \frac{1}{p(x_1)\, p(x_2) \cdots p(x_n)} \right) - S(\rho) \right| \leq \epsilon$$

*gilt. Entsprechend wird ein zugehöriger Zustand* $|x_1\rangle |x_2\rangle \cdots |x_n\rangle$ $\epsilon$*-typisch genannt.*
*Der Unterraum aller* $\epsilon$*-typischen Zustände wird* $T(n, \epsilon)$ *genannt. Der entsprechende Projektor auf diesen Unterraum ist*

$$P(n, \epsilon) = \sum_{x\ \epsilon\text{-typisch}} |x_1\rangle \langle x_1| \otimes \cdots \otimes |x_n\rangle \langle x_n|$$

# Satz über typische Unterräume

## Satz (über typische Unterräume)

1. *Sei $\epsilon > 0$. Für jedes $\delta > 0$ und genügend große $n$ gilt:*

$$tr\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) \geq 1-\delta$$

2. *Für jedes $\epsilon > 0$ und $\delta > 0$ und genügend große $n$ erfüllt die Dimension $|T\left(n,\epsilon\right)| = tr\left(P\left(n,\epsilon\right)\right)$ von $T\left(n,\epsilon\right)$*

$$\left(1-\delta\right)2^{n\left(S(\rho)-\epsilon\right)} \leq |T\left(n,\epsilon\right)| \leq 2^{n\left(S(\rho)+\epsilon\right)}$$

3. *Sei $S\left(n\right)$ ein Projektor auf einen beliebigen Unterraum von $H^{\otimes n}$ mit Dimension kleiner $2^{nR}$. Sei $R < S\left(\rho\right)$. Dann gilt für alle $\delta > 0$ und genügend große $n$*

$$tr\left(S\left(n\right)\rho^{\otimes n}\right) \leq \delta$$

# Satz über typische Unterräume

## Satz (über typische Unterräume)

1. *Sei $\epsilon > 0$. Für jedes $\delta > 0$ und genügend große $n$ gilt:*

$$tr\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) \geq 1-\delta$$

2. *Für jedes $\epsilon > 0$ und $\delta > 0$ und genügend große $n$ erfüllt die Dimension $|T\left(n,\epsilon\right)| = tr\left(P\left(n,\epsilon\right)\right)$ von $T\left(n,\epsilon\right)$*

$$\left(1-\delta\right)2^{n\left(S(\rho)-\epsilon\right)} \leq |T\left(n,\epsilon\right)| \leq 2^{n\left(S(\rho)+\epsilon\right)}$$

3. *Sei $S\left(n\right)$ ein Projektor auf einen beliebigen Unterraum von $H^{\otimes n}$ mit Dimension kleiner $2^{nR}$. Sei $R < S\left(\rho\right)$. Dann gilt für alle $\delta > 0$ und genügend große $n$*

$$tr\left(S\left(n\right)\rho^{\otimes n}\right) \leq \delta$$

# Satz über typische Unterräume

## Satz (über typische Unterräume)

1. *Sei $\epsilon > 0$. Für jedes $\delta > 0$ und genügend große $n$ gilt:*

$$tr\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) \geq 1-\delta$$

2. *Für jedes $\epsilon > 0$ und $\delta > 0$ und genügend große $n$ erfüllt die Dimension $|T\left(n,\epsilon\right)| = tr\left(P\left(n,\epsilon\right)\right)$ von $T\left(n,\epsilon\right)$*

$$\left(1-\delta\right)2^{n\left(S(\rho)-\epsilon\right)} \leq |T\left(n,\epsilon\right)| \leq 2^{n\left(S(\rho)+\epsilon\right)}$$

3. *Sei $S\left(n\right)$ ein Projektor auf einen beliebigen Unterraum von $H^{\otimes n}$ mit Dimension kleiner $2^{nR}$. Sei $R < S\left(\rho\right)$. Dann gilt für alle $\delta > 0$ und genügend große $n$*

$$tr\left(S\left(n\right)\rho^{\otimes n}\right) \leq \delta$$

# Satz über typische Unterräume

## Satz (über typische Unterräume)

1. *Sei $\epsilon > 0$. Für jedes $\delta > 0$ und genügend große $n$ gilt:*

$$tr\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) \geq 1-\delta$$

2. *Für jedes $\epsilon > 0$ und $\delta > 0$ und genügend große $n$ erfüllt die Dimension $|T\left(n,\epsilon\right)| = tr\left(P\left(n,\epsilon\right)\right)$ von $T\left(n,\epsilon\right)$*

$$\left(1-\delta\right)2^{n\left(S(\rho)-\epsilon\right)} \leq |T\left(n,\epsilon\right)| \leq 2^{n\left(S(\rho)+\epsilon\right)}$$

3. *Sei $S\left(n\right)$ ein Projektor auf einen beliebigen Unterraum von $H^{\otimes n}$ mit Dimension kleiner $2^{nR}$. Sei $R < S\left(\rho\right)$. Dann gilt für alle $\delta > 0$ und genügend große $n$*

$$tr\left(S\left(n\right)\rho^{\otimes n}\right) \leq \delta$$

# Beweis: Satz typische Unterräume - Teil 1

- Zu zeigen: $\operatorname{tr}\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) \geq 1-\delta$
- $\operatorname{tr}\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) = \sum_{x\in T(n,\epsilon)} p\left(x_1\right)p\left(x_2\right)\cdots p\left(x_n\right) =$W-keit, dass eine Sequenz $\epsilon$-typisch ist
- Entspricht

$$p\left(\left|\frac{1}{n}\log\left(\frac{1}{p\left(x_1\right)p\left(x_2\right)\cdots p\left(x_n\right)}\right)-S\left(\rho\right)\right|\leq\epsilon\right)\geq 1-\delta$$

- Sei $X$ eine Zufallsvariable mit der Verteilung der $x_i$.

$$p\left(\left|\sum_{i=1}^{n}\frac{-\log p\left(x_i\right)}{n}-E\left(-\log X\right)\right|\leq\epsilon\right)\geq 1-\delta$$

- Gesetz der großen Zahlen

# Beweis: Satz typische Unterräume - Teil 1

- Zu zeigen: $\operatorname{tr}\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) \geq 1-\delta$
- $\operatorname{tr}\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) = \sum_{x\in T(n,\epsilon)} p\left(x_1\right)p\left(x_2\right)\cdots p\left(x_n\right) =$W-keit, dass eine Sequenz $\epsilon$-typisch ist
- Entspricht

$$p\left(\left|\frac{1}{n}\log\left(\frac{1}{p\left(x_1\right)p\left(x_2\right)\cdots p\left(x_n\right)}\right)-S\left(\rho\right)\right|\leq\epsilon\right)\geq 1-\delta$$

- Sei $X$ eine Zufallsvariable mit der Verteilung der $x_i$.

$$p\left(\left|\sum_{i=1}^{n}\frac{-\log p\left(x_i\right)}{n}-E\left(-\log X\right)\right|\leq\epsilon\right)\geq 1-\delta$$

- Gesetz der großen Zahlen

# Beweis: Satz typische Unterräume - Teil 1

- Zu zeigen: $\operatorname{tr}\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) \geq 1-\delta$
- $\operatorname{tr}\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) = \sum_{x\in T(n,\epsilon)} p\left(x_1\right)p\left(x_2\right)\cdots p\left(x_n\right) =$W-keit, dass eine Sequenz $\epsilon$-typisch ist
- Entspricht

$$p\left(\left|\frac{1}{n}\log\left(\frac{1}{p\left(x_1\right)p\left(x_2\right)\cdots p\left(x_n\right)}\right)-S\left(\rho\right)\right|\leq\epsilon\right)\geq 1-\delta$$

- Sei $X$ eine Zufallsvariable mit der Verteilung der $x_i$.

$$p\left(\left|\sum_{i=1}^{n}\frac{-\log p\left(x_i\right)}{n}-E\left(-\log X\right)\right|\leq\epsilon\right)\geq 1-\delta$$

- Gesetz der großen Zahlen

# Beweis: Satz typische Unterräume - Teil 1

- Zu zeigen: $\operatorname{tr}\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) \geq 1-\delta$
- $\operatorname{tr}\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) = \sum_{x\in T(n,\epsilon)} p\left(x_1\right)p\left(x_2\right)\cdots p\left(x_n\right) =$W-keit, dass eine Sequenz $\epsilon$-typisch ist
- Entspricht

$$p\left(\left|\frac{1}{n}\log\left(\frac{1}{p\left(x_1\right)p\left(x_2\right)\cdots p\left(x_n\right)}\right)-S\left(\rho\right)\right|\leq\epsilon\right)\geq 1-\delta$$

- Sei $X$ eine Zufallsvariable mit der Verteilung der $x_i$.

$$p\left(\left|\sum_{i=1}^{n}\frac{-\log p\left(x_i\right)}{n}-E\left(-\log X\right)\right|\leq\epsilon\right)\geq 1-\delta$$

- Gesetz der großen Zahlen

# Beweis: Satz typische Unterräume - Teil 1

- Zu zeigen: $\operatorname{tr}\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) \geq 1-\delta$
- $\operatorname{tr}\left(P\left(n,\epsilon\right)\rho^{\otimes n}\right) = \sum_{x\in T(n,\epsilon)} p\left(x_1\right)p\left(x_2\right)\cdots p\left(x_n\right) =$W-keit, dass eine Sequenz $\epsilon$-typisch ist
- Entspricht

$$p\left(\left|\frac{1}{n}\log\left(\frac{1}{p\left(x_1\right)p\left(x_2\right)\cdots p\left(x_n\right)}\right)-S\left(\rho\right)\right|\leq\epsilon\right)\geq 1-\delta$$

- Sei $X$ eine Zufallsvariable mit der Verteilung der $x_i$.

$$p\left(\left|\sum_{i=1}^{n}\frac{-\log p\left(x_i\right)}{n}-E\left(-\log X\right)\right|\leq\epsilon\right)\geq 1-\delta$$

- Gesetz der großen Zahlen

# Beweis: Satz typische Unterräume - Teil 2

- Zu zeigen: $(1-\delta)\,2^{n(S(\rho)-\epsilon)} \leq |T(n,\epsilon)| \leq 2^{n(S(\rho)+\epsilon)}$
- Aus Def. lässt sich Abschätzung gewinnen:

$$\left|\frac{1}{n}\log\left(\frac{1}{p(x_1)\,p(x_2)\cdots p(x_n)}\right) - S(\rho)\right| \leq \epsilon$$

$$S(\rho)-\epsilon \leq \frac{1}{n}\log\left(\frac{1}{p(x)}\right) \leq S(\rho)+\epsilon$$

$$-n\left(S(\rho)-\epsilon\right) \geq \log\left(p(x)\right) \geq -n\left(S(\rho)+\epsilon\right)$$

$$2^{-n(S(\rho)-\epsilon)} \geq p(x) \geq 2^{-n(S(\rho)+\epsilon)}$$

- $1 \geq \sum_{x\in T(n,\epsilon)} p(x) \geq \sum_{x\in T(n,\epsilon)} 2^{-n(S(\rho)+\epsilon)} = |T(n,\epsilon)|\,2^{-n(S(\rho)+\epsilon)}$
- $|T(n,\epsilon)| \leq 2^{n(S(\rho)+\epsilon)}$
- $1-\delta \leq \sum_{x\in T(n,\epsilon)} p(x) \leq \sum_{x\in T(n,\epsilon)} 2^{-n(S(\rho)-\epsilon)} = |T(n,\epsilon)|\,2^{-n(S(\rho)-\epsilon)}$
- $|T(n,\epsilon)| \geq (1-\delta)\,2^{n(S(\rho)-\epsilon)}$

# Beweis: Satz typische Unterräume - Teil 2

- Zu zeigen: $(1-\delta)\,2^{n(S(\rho)-\epsilon)} \leq |T(n,\epsilon)| \leq 2^{n(S(\rho)+\epsilon)}$
- Aus Def. lässt sich Abschätzung gewinnen:

$$\left|\frac{1}{n}\log\left(\frac{1}{p(x_1)\,p(x_2)\cdots p(x_n)}\right) - S(\rho)\right| \leq \epsilon$$

$$S(\rho)-\epsilon \leq \frac{1}{n}\log\left(\frac{1}{p(x)}\right) \leq S(\rho)+\epsilon$$

$$-n\left(S(\rho)-\epsilon\right) \geq \log\left(p(x)\right) \geq -n\left(S(\rho)+\epsilon\right)$$

$$2^{-n(S(\rho)-\epsilon)} \geq p(x) \geq 2^{-n(S(\rho)+\epsilon)}$$

- $1 \geq \sum_{x\in T(n,\epsilon)} p(x) \geq \sum_{x\in T(n,\epsilon)} 2^{-n(S(\rho)+\epsilon)} = |T(n,\epsilon)|\,2^{-n(S(\rho)+\epsilon)}$
- $|T(n,\epsilon)| \leq 2^{n(S(\rho)+\epsilon)}$
- $1-\delta \leq \sum_{x\in T(n,\epsilon)} p(x) \leq \sum_{x\in T(n,\epsilon)} 2^{-n(S(\rho)-\epsilon)} = |T(n,\epsilon)|\,2^{-n(S(\rho)-\epsilon)}$
- $|T(n,\epsilon)| \geq (1-\delta)\,2^{n(S(\rho)-\epsilon)}$

# Beweis: Satz typische Unterräume - Teil 2

- Zu zeigen: $(1-\delta)\,2^{n(S(\rho)-\epsilon)} \leq |T(n,\epsilon)| \leq 2^{n(S(\rho)+\epsilon)}$
- Aus Def. lässt sich Abschätzung gewinnen:

$$\left|\frac{1}{n}\log\left(\frac{1}{p(x_1)\,p(x_2)\cdots p(x_n)}\right) - S(\rho)\right| \leq \epsilon$$

$$S(\rho)-\epsilon \leq \frac{1}{n}\log\left(\frac{1}{p(x)}\right) \leq S(\rho)+\epsilon$$

$$-n\left(S(\rho)-\epsilon\right) \geq \log\left(p(x)\right) \geq -n\left(S(\rho)+\epsilon\right)$$

$$2^{-n(S(\rho)-\epsilon)} \geq p(x) \geq 2^{-n(S(\rho)+\epsilon)}$$

- $1 \geq \sum_{x\in T(n,\epsilon)} p(x) \geq \sum_{x\in T(n,\epsilon)} 2^{-n(S(\rho)+\epsilon)} = |T(n,\epsilon)|\,2^{-n(S(\rho)+\epsilon)}$
- $|T(n,\epsilon)| \leq 2^{n(S(\rho)+\epsilon)}$
- $1-\delta \leq \sum_{x\in T(n,\epsilon)} p(x) \leq \sum_{x\in T(n,\epsilon)} 2^{-n(S(\rho)-\epsilon)} = |T(n,\epsilon)|\,2^{-n(S(\rho)-\epsilon)}$
- $|T(n,\epsilon)| \geq (1-\delta)\,2^{n(S(\rho)-\epsilon)}$

# Beweis: Satz typische Unterräume - Teil 2

- Zu zeigen: $(1-\delta)\,2^{n(S(\rho)-\epsilon)} \leq |T(n,\epsilon)| \leq 2^{n(S(\rho)+\epsilon)}$
- Aus Def. lässt sich Abschätzung gewinnen:

$$\left|\frac{1}{n}\log\left(\frac{1}{p(x_1)\,p(x_2)\cdots p(x_n)}\right) - S(\rho)\right| \leq \epsilon$$

$$S(\rho)-\epsilon \leq \frac{1}{n}\log\left(\frac{1}{p(x)}\right) \leq S(\rho)+\epsilon$$

$$-n\left(S(\rho)-\epsilon\right) \geq \log\left(p(x)\right) \geq -n\left(S(\rho)+\epsilon\right)$$

$$2^{-n(S(\rho)-\epsilon)} \geq p(x) \geq 2^{-n(S(\rho)+\epsilon)}$$

- $1 \geq \sum_{x\in T(n,\epsilon)} p(x) \geq \sum_{x\in T(n,\epsilon)} 2^{-n(S(\rho)+\epsilon)} = |T(n,\epsilon)|\,2^{-n(S(\rho)+\epsilon)}$
- $|T(n,\epsilon)| \leq 2^{n(S(\rho)+\epsilon)}$
- $1-\delta \leq \sum_{x\in T(n,\epsilon)} p(x) \leq \sum_{x\in T(n,\epsilon)} 2^{-n(S(\rho)-\epsilon)} = |T(n,\epsilon)|\,2^{-n(S(\rho)-\epsilon)}$
- $|T(n,\epsilon)| \geq (1-\delta)\,2^{n(S(\rho)-\epsilon)}$

# Beweis: Satz typische Unterräume - Teil 2

- Zu zeigen: $(1-\delta)\,2^{n(S(\rho)-\epsilon)} \leq |T(n,\epsilon)| \leq 2^{n(S(\rho)+\epsilon)}$
- Aus Def. lässt sich Abschätzung gewinnen:

$$\left|\frac{1}{n}\log\left(\frac{1}{p(x_1)\,p(x_2)\cdots p(x_n)}\right) - S(\rho)\right| \leq \epsilon$$

$$S(\rho)-\epsilon \leq \frac{1}{n}\log\left(\frac{1}{p(x)}\right) \leq S(\rho)+\epsilon$$

$$-n\left(S(\rho)-\epsilon\right) \geq \log\left(p(x)\right) \geq -n\left(S(\rho)+\epsilon\right)$$

$$2^{-n(S(\rho)-\epsilon)} \geq p(x) \geq 2^{-n(S(\rho)+\epsilon)}$$

- $1 \geq \sum_{x\in T(n,\epsilon)} p(x) \geq \sum_{x\in T(n,\epsilon)} 2^{-n(S(\rho)+\epsilon)} = |T(n,\epsilon)|\,2^{-n(S(\rho)+\epsilon)}$
- $|T(n,\epsilon)| \leq 2^{n(S(\rho)+\epsilon)}$
- $1-\delta \leq \sum_{x\in T(n,\epsilon)} p(x) \leq \sum_{x\in T(n,\epsilon)} 2^{-n(S(\rho)-\epsilon)} = |T(n,\epsilon)|\,2^{-n(S(\rho)-\epsilon)}$
- $|T(n,\epsilon)| \geq (1-\delta)\,2^{n(S(\rho)-\epsilon)}$

# Beweis: Satz typische Unterräume - Teil 2

- Zu zeigen: $(1-\delta)\,2^{n(S(\rho)-\epsilon)} \leq |T(n,\epsilon)| \leq 2^{n(S(\rho)+\epsilon)}$
- Aus Def. lässt sich Abschätzung gewinnen:

$$\left|\frac{1}{n}\log\left(\frac{1}{p(x_1)\,p(x_2)\cdots p(x_n)}\right) - S(\rho)\right| \leq \epsilon$$

$$S(\rho)-\epsilon \leq \frac{1}{n}\log\left(\frac{1}{p(x)}\right) \leq S(\rho)+\epsilon$$

$$-n\left(S(\rho)-\epsilon\right) \geq \log\left(p(x)\right) \geq -n\left(S(\rho)+\epsilon\right)$$

$$2^{-n(S(\rho)-\epsilon)} \geq p(x) \geq 2^{-n(S(\rho)+\epsilon)}$$

- $1 \geq \sum_{x\in T(n,\epsilon)} p(x) \geq \sum_{x\in T(n,\epsilon)} 2^{-n(S(\rho)+\epsilon)} = |T(n,\epsilon)|\,2^{-n(S(\rho)+\epsilon)}$
- $|T(n,\epsilon)| \leq 2^{n(S(\rho)+\epsilon)}$
- $1-\delta \leq \sum_{x\in T(n,\epsilon)} p(x) \leq \sum_{x\in T(n,\epsilon)} 2^{-n(S(\rho)-\epsilon)} = |T(n,\epsilon)|\,2^{-n(S(\rho)-\epsilon)}$
- $|T(n,\epsilon)| \geq (1-\delta)\,2^{n(S(\rho)-\epsilon)}$

# Beweis: Satz typische Unterräume - Teil 3

- Zu zeigen: $|S(n)| < 2^{nR} < 2^{nS(\rho)} \Rightarrow \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\right) \leq \delta$
- Aufteilen der Spur in den typischen Unterraum und den Rest

$$\operatorname{tr}\left(S(n)\,\rho^{\otimes n}\right) = \operatorname{tr}\left(S(n)\,\rho^{\otimes n} P(n,\epsilon)\right) + \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right)$$

- Da der Projektor $P(n,\epsilon)$ mit $\rho^{\otimes n}$ kommutiert, gilt:

$$\rho^{\otimes n} P(n,\epsilon) = P(n,\epsilon)\,\rho^{\otimes n} P(n,\epsilon)$$

- Da die Eigenwerte beschränkt sind, gilt:

$$\operatorname{tr}\left(S(n)\,P(n,\epsilon)\,\rho^{\otimes n} P(n,\epsilon)\right) \leq 2^{nR} 2^{-n(S(\rho)+\epsilon)}$$

$$= 2^{-n(S(\rho)-R+\epsilon)} \to 0$$

- Es gilt $S(n) \leq I$. Da $S(n)$ und $\rho^{\otimes n}\left(I - P(n,\epsilon)\right)$ positive Operatoren sind, gilt:

$$0 \leq \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right) \leq \operatorname{tr}\left(\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right) \to 0$$

□

# Beweis: Satz typische Unterräume - Teil 3

- Zu zeigen: $|S(n)| < 2^{nR} < 2^{nS(\rho)} \Rightarrow \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\right) \leq \delta$
- Aufteilen der Spur in den typischen Unterraum und den Rest

$$\operatorname{tr}\left(S(n)\,\rho^{\otimes n}\right) = \operatorname{tr}\left(S(n)\,\rho^{\otimes n} P(n,\epsilon)\right) + \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right)$$

- Da der Projektor $P(n,\epsilon)$ mit $\rho^{\otimes n}$ kommutiert, gilt:

$$\rho^{\otimes n} P(n,\epsilon) = P(n,\epsilon)\,\rho^{\otimes n} P(n,\epsilon)$$

- Da die Eigenwerte beschränkt sind, gilt:

$$\operatorname{tr}\left(S(n)\,P(n,\epsilon)\,\rho^{\otimes n} P(n,\epsilon)\right) \leq 2^{nR} 2^{-n(S(\rho)+\epsilon)}$$

$$= 2^{-n(S(\rho)-R+\epsilon)} \to 0$$

- Es gilt $S(n) \leq I$. Da $S(n)$ und $\rho^{\otimes n}\left(I - P(n,\epsilon)\right)$ positive Operatoren sind, gilt:

$$0 \leq \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right) \leq \operatorname{tr}\left(\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right) \to 0$$

□

# Beweis: Satz typische Unterräume - Teil 3

- Zu zeigen: $|S(n)| < 2^{nR} < 2^{nS(\rho)} \Rightarrow \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\right) \leq \delta$
- Aufteilen der Spur in den typischen Unterraum und den Rest

$$\operatorname{tr}\left(S(n)\,\rho^{\otimes n}\right) = \operatorname{tr}\left(S(n)\,\rho^{\otimes n} P(n,\epsilon)\right) + \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right)$$

- Da der Projektor $P(n,\epsilon)$ mit $\rho^{\otimes n}$ kommutiert, gilt:

$$\rho^{\otimes n} P(n,\epsilon) = P(n,\epsilon)\,\rho^{\otimes n} P(n,\epsilon)$$

- Da die Eigenwerte beschränkt sind, gilt:

$$\operatorname{tr}\left(S(n)\,P(n,\epsilon)\,\rho^{\otimes n} P(n,\epsilon)\right) \leq 2^{nR} 2^{-n(S(\rho)+\epsilon)}$$

$$= 2^{-n(S(\rho)-R+\epsilon)} \to 0$$

- Es gilt $S(n) \leq I$. Da $S(n)$ und $\rho^{\otimes n}\left(I - P(n,\epsilon)\right)$ positive Operatoren sind, gilt:

$$0 \leq \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right) \leq \operatorname{tr}\left(\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right) \to 0$$

□

# Beweis: Satz typische Unterräume - Teil 3

- Zu zeigen: $|S(n)| < 2^{nR} < 2^{nS(\rho)} \Rightarrow \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\right) \leq \delta$
- Aufteilen der Spur in den typischen Unterraum und den Rest

$$\operatorname{tr}\left(S(n)\,\rho^{\otimes n}\right) = \operatorname{tr}\left(S(n)\,\rho^{\otimes n} P(n,\epsilon)\right) + \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right)$$

- Da der Projektor $P(n,\epsilon)$ mit $\rho^{\otimes n}$ kommutiert, gilt:

$$\rho^{\otimes n} P(n,\epsilon) = P(n,\epsilon)\,\rho^{\otimes n} P(n,\epsilon)$$

- Da die Eigenwerte beschränkt sind, gilt:

$$\operatorname{tr}\left(S(n)\,P(n,\epsilon)\,\rho^{\otimes n} P(n,\epsilon)\right) \leq 2^{nR} 2^{-n(S(\rho)+\epsilon)}$$

$$= 2^{-n(S(\rho)-R+\epsilon)} \to 0$$

- Es gilt $S(n) \leq I$. Da $S(n)$ und $\rho^{\otimes n}\left(I - P(n,\epsilon)\right)$ positive Operatoren sind, gilt:

$$0 \leq \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right) \leq \operatorname{tr}\left(\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right) \to 0$$

□

# Beweis: Satz typische Unterräume - Teil 3

- Zu zeigen: $|S(n)| < 2^{nR} < 2^{nS(\rho)} \Rightarrow \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\right) \leq \delta$
- Aufteilen der Spur in den typischen Unterraum und den Rest

$$\operatorname{tr}\left(S(n)\,\rho^{\otimes n}\right) = \operatorname{tr}\left(S(n)\,\rho^{\otimes n} P(n,\epsilon)\right) + \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right)$$

- Da der Projektor $P(n,\epsilon)$ mit $\rho^{\otimes n}$ kommutiert, gilt:

$$\rho^{\otimes n} P(n,\epsilon) = P(n,\epsilon)\,\rho^{\otimes n} P(n,\epsilon)$$

- Da die Eigenwerte beschränkt sind, gilt:

$$\operatorname{tr}\left(S(n)\,P(n,\epsilon)\,\rho^{\otimes n} P(n,\epsilon)\right) \leq 2^{nR} 2^{-n(S(\rho)+\epsilon)}$$

$$= 2^{-n(S(\rho)-R+\epsilon)} \to 0$$

- Es gilt $S(n) \leq I$. Da $S(n)$ und $\rho^{\otimes n}\left(I - P(n,\epsilon)\right)$ positive Operatoren sind, gilt:

$$0 \leq \operatorname{tr}\left(S(n)\,\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right) \leq \operatorname{tr}\left(\rho^{\otimes n}\left(I - P(n,\epsilon)\right)\right) \to 0$$

$\square$

# Satz von Schumacher

**Satz (Schumacher: quantenmechanische rauschfreie Kanalkodierung)**

*Sei $\{H, \rho\}$ eine i.i.d. Quanteninformationsquelle. Für $R > S(\rho)$ exisitiert ein zuverlässiges Kompressionsschema der Rate $R$ für diese Quelle. Umgekehrt ist wäre jedes Kompressionsschema für $R < S(\rho)$ nicht zuverlässig.*

- Zuverlässig entspricht $F(\rho, \cdot) \to 1$ für $n \to \infty$

# Satz von Schumacher

**Satz (Schumacher: quantenmechanische rauschfreie Kanalkodierung)**

*Sei $\{H, \rho\}$ eine i.i.d. Quanteninformationsquelle. Für $R > S(\rho)$ exisitiert ein zuverlässiges Kompressionsschema der Rate $R$ für diese Quelle. Umgekehrt ist wäre jedes Kompressionsschema für $R < S(\rho)$ nicht zuverlässig.*

- Zuverlässig entspricht $F(\rho, \cdot) \to 1$ für $n \to \infty$

# Beweis: Satz von Schumacher - Existenz - Teil 1

- Zu zeigen: Wenn $R > S(\rho)$, dann $F(\rho, \cdot) \to 1$
- wähle $\epsilon > 0$ so, dass $S(\rho) + \epsilon \leq R$.
- Aus Theorem über typische Unterräume: Für jedes $\delta > 0$ und genügend große $n$ gilt:

$$\begin{aligned} \operatorname{tr}\left(\rho^{\otimes n} P(n, \epsilon)\right) &\geq 1 - \delta \\ |T(n, \epsilon)| &\leq 2^{nR} \end{aligned}$$

- Sei $H_c^n$ ein $2^{nR}$-dimensionaler Hilbertraum, der $T(n, \epsilon)$ enthält.

# Beweis: Satz von Schumacher - Existenz - Teil 1

- Zu zeigen: Wenn $R > S(\rho)$, dann $F(\rho, \cdot) \to 1$
- wähle $\epsilon > 0$ so, dass $S(\rho) + \epsilon \leq R$.
- Aus Theorem über typische Unterräume: Für jedes $\delta > 0$ und genügend große $n$ gilt:

$$\begin{aligned} \operatorname{tr}\left(\rho^{\otimes n} P(n, \epsilon)\right) &\geq 1 - \delta \\ |T(n, \epsilon)| &\leq 2^{nR} \end{aligned}$$

- Sei $H_c^n$ ein $2^{nR}$-dimensionaler Hilbertraum, der $T(n, \epsilon)$ enthält.

# Beweis: Satz von Schumacher - Existenz - Teil 1

- Zu zeigen: Wenn $R > S(\rho)$, dann $F(\rho, \cdot) \to 1$
- wähle $\epsilon > 0$ so, dass $S(\rho) + \epsilon \leq R$.
- Aus Theorem über typische Unterräume: Für jedes $\delta > 0$ und genügend große $n$ gilt:

$$\begin{aligned} \operatorname{tr}\left(\rho^{\otimes n} P(n, \epsilon)\right) &\geq 1 - \delta \\ |T(n, \epsilon)| &\leq 2^{nR} \end{aligned}$$

- Sei $H_c^n$ ein $2^{nR}$-dimensionaler Hilbertraum, der $T(n, \epsilon)$ enthält.

# Beweis: Satz von Schumacher - Existenz - Teil 1

- Zu zeigen: Wenn $R > S(\rho)$, dann $F(\rho, \cdot) \to 1$
- wähle $\epsilon > 0$ so, dass $S(\rho) + \epsilon \leq R$.
- Aus Theorem über typische Unterräume: Für jedes $\delta > 0$ und genügend große $n$ gilt:

$$\begin{aligned} \operatorname{tr}\left(\rho^{\otimes n} P(n, \epsilon)\right) &\geq 1 - \delta \\ |T(n, \epsilon)| &\leq 2^{nR} \end{aligned}$$

- Sei $H_c^n$ ein $2^{nR}$-dimensionaler Hilbertraum, der $T(n, \epsilon)$ enthält.

# Beweis: Satz von Schumacher - Existenz - Teil 2

- Komprimiere mit $A_i \equiv |0\rangle \langle i|$, wobei
  $|i\rangle$ orthonormale Basis, die orthogonal zum typischen Unterraum ist;
  $|0\rangle$ ein beliebiger Vektor aus dem typischen Unterraum.

$$\mathcal{C}^n(\sigma) = P(n,\epsilon)\,\sigma P(n,\epsilon) + \sum_i A_i \sigma A_i^\dagger$$

- Dekodieren $\mathcal{D}^n$ ist als Identität auf $H_c^n$ definiert.

$$\begin{aligned} F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) &= \left|\mathrm{tr}\left(\rho^{\otimes n} P(n,\epsilon)\right)\right|^2 + \sum_i \left|\mathrm{tr}\left(\rho^{\otimes n} A_i\right)\right|^2 \\ &\geq \left|\mathrm{tr}\left(\rho^{\otimes n} P(n,\epsilon)\right)\right|^2 \\ &\geq |1-\delta|^2 \geq 1 - 2\delta \end{aligned}$$

# Beweis: Satz von Schumacher - Existenz - Teil 2

- Komprimiere mit $A_i \equiv |0\rangle \langle i|$, wobei
  $|i\rangle$ orthonormale Basis, die orthogonal zum typischen Unterraum ist;
  $|0\rangle$ ein beliebiger Vektor aus dem typischen Unterraum.

$$\mathcal{C}^n(\sigma) = P(n,\epsilon)\,\sigma P(n,\epsilon) + \sum_i A_i \sigma A_i^\dagger$$

- Dekodieren $\mathcal{D}^n$ ist als Identität auf $H_c^n$ definiert.

$$\begin{aligned} F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) &= \left|\mathrm{tr}\left(\rho^{\otimes n} P(n,\epsilon)\right)\right|^2 + \sum_i \left|\mathrm{tr}\left(\rho^{\otimes n} A_i\right)\right|^2 \\ &\geq \left|\mathrm{tr}\left(\rho^{\otimes n} P(n,\epsilon)\right)\right|^2 \\ &\geq |1-\delta|^2 \geq 1 - 2\delta \end{aligned}$$

# Beweis: Satz von Schumacher - Optimal - Teil 1

- Zu zeigen: Wenn $R < S(\rho)$ dann $F(\rho, \cdot) \to 0$
- oBdA $\mathcal{C}^n$ bildet in einen $2^{nR}$-dim Unterraum von $H^{\otimes n}$ ab (zugehöriger Projektor $S(n)$)
- Seien $C_j$ die Operatorelemente von $\mathcal{C}^n$ und $D_k$ die von $\mathcal{D}^n$.

$$F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) = \sum_{jk} \left|\operatorname{tr}\left(D_k C_j \rho^{\otimes n}\right)\right|^2$$

- Jeder der Operatoren $C_j$ bildet in den Unterraum von $S(n)$ ab: $C_j = S(n) C_j$
- Sei $S^k(n)$ der Projektor auf den Unterraum, auf den $D_k\ S(n)$ abbildet:
  $S^k(n) D_k S(n) = D_k S(n)$

$$\begin{aligned} D_k C_j &= D_k S(n) C_j \\ &= S^k(n) D_k S(n) C_j \\ &= S^k(n) D_k C_j \end{aligned}$$

# Beweis: Satz von Schumacher - Optimal - Teil 1

- Zu zeigen: Wenn $R < S(\rho)$ dann $F(\rho, \cdot) \to 0$
- oBdA $\mathcal{C}^n$ bildet in einen $2^{nR}$-dim Unterraum von $H^{\otimes n}$ ab (zugehöriger Projektor $S(n)$)
- Seien $C_j$ die Operatorelemente von $\mathcal{C}^n$ und $D_k$ die von $\mathcal{D}^n$.

$$F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) = \sum_{jk} \left|\mathrm{tr}\left(D_k C_j \rho^{\otimes n}\right)\right|^2$$

- Jeder der Operatoren $C_j$ bildet in den Unterraum von $S(n)$ ab: $C_j = S(n) C_j$
- Sei $S^k(n)$ der Projektor auf den Unterraum, auf den $D_k\ S(n)$ abbildet: $S^k(n) D_k S(n) = D_k S(n)$

$$\begin{aligned} D_k C_j &= D_k S(n) C_j \\ &= S^k(n) D_k S(n) C_j \\ &= S^k(n) D_k C_j \end{aligned}$$

# Beweis: Satz von Schumacher - Optimal - Teil 1

- Zu zeigen: Wenn $R < S(\rho)$ dann $F(\rho, \cdot) \to 0$
- oBdA $\mathcal{C}^n$ bildet in einen $2^{nR}$-dim Unterraum von $H^{\otimes n}$ ab (zugehöriger Projektor $S(n)$)
- Seien $C_j$ die Operatorelemente von $\mathcal{C}^n$ und $D_k$ die von $\mathcal{D}^n$.

$$F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) = \sum_{jk} \left|\mathrm{tr}\left(D_k C_j \rho^{\otimes n}\right)\right|^2$$

- Jeder der Operatoren $C_j$ bildet in den Unterraum von $S(n)$ ab: $C_j = S(n) C_j$
- Sei $S^k(n)$ der Projektor auf den Unterraum, auf den $D_k\ S(n)$ abbildet: $S^k(n) D_k S(n) = D_k S(n)$

$$\begin{aligned} D_k C_j &= D_k S(n) C_j \\ &= S^k(n) D_k S(n) C_j \\ &= S^k(n) D_k C_j \end{aligned}$$

# Beweis: Satz von Schumacher - Optimal - Teil 1

- Zu zeigen: Wenn $R < S(\rho)$ dann $F(\rho, \cdot) \to 0$
- oBdA $\mathcal{C}^n$ bildet in einen $2^{nR}$-dim Unterraum von $H^{\otimes n}$ ab (zugehöriger Projektor $S(n)$)
- Seien $C_j$ die Operatorelemente von $\mathcal{C}^n$ und $D_k$ die von $\mathcal{D}^n$.

$$F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) = \sum_{jk} \left|\mathrm{tr}\left(D_k C_j \rho^{\otimes n}\right)\right|^2$$

- Jeder der Operatoren $C_j$ bildet in den Unterraum von $S(n)$ ab: $C_j = S(n) C_j$
- Sei $S^k(n)$ der Projektor auf den Unterraum, auf den $D_k\ S(n)$ abbildet:
  $S^k(n) D_k S(n) = D_k S(n)$

$$\begin{aligned} D_k C_j &= D_k S(n) C_j \\ &= S^k(n) D_k S(n) C_j \\ &= S^k(n) D_k C_j \end{aligned}$$

# Beweis: Satz von Schumacher - Optimal - Teil 1

- Zu zeigen: Wenn $R < S(\rho)$ dann $F(\rho, \cdot) \to 0$
- oBdA $\mathcal{C}^n$ bildet in einen $2^{nR}$-dim Unterraum von $H^{\otimes n}$ ab (zugehöriger Projektor $S(n)$)
- Seien $C_j$ die Operatorelemente von $\mathcal{C}^n$ und $D_k$ die von $\mathcal{D}^n$.

$$F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) = \sum_{jk} \left|\mathrm{tr}\left(D_k C_j \rho^{\otimes n}\right)\right|^2$$

- Jeder der Operatoren $C_j$ bildet in den Unterraum von $S(n)$ ab: $C_j = S(n) C_j$
- Sei $S^k(n)$ der Projektor auf den Unterraum, auf den $D_k\ S(n)$ abbildet:
  $S^k(n) D_k S(n) = D_k S(n)$

$$\begin{aligned} D_k C_j &= D_k S(n) C_j \\ &= S^k(n) D_k S(n) C_j \\ &= S^k(n) D_k C_j \end{aligned}$$

# Beweis: Satz von Schumacher - Optimal - Teil 2

- Damit ergibt sich (mit Cauchy-Schwarz-Ungleichung)

$$\begin{aligned} F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) &= \sum_{jk} \left|\mathrm{tr}\left(D_k C_j \rho^{\otimes n} S^k(n)\right)\right|^2 \\ &\leq \sum_{jk} \mathrm{tr}\left(D_k C_j \rho^{\otimes n} C_j^\dagger D_k^\dagger\right) \mathrm{tr}\left(S^k(n) \rho^{\otimes n} S^k(n)\right) \end{aligned}$$

- Nach dem Satz über typische Unterräume gilt mit der Spurerhaltung von $C_j, D_k$:

$$\begin{aligned} F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) &\leq \delta \sum_{jk} \mathrm{tr}\left(D_k C_j \rho^{\otimes n} C_j^\dagger D_k^\dagger\right) \\ &= \delta \end{aligned}$$

□

# Beweis: Satz von Schumacher - Optimal - Teil 2

- Damit ergibt sich (mit Cauchy-Schwarz-Ungleichung)

$$\begin{aligned} F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) &= \sum_{jk} \left|\operatorname{tr}\left(D_k C_j \rho^{\otimes n} S^k(n)\right)\right|^2 \\ &\leq \sum_{jk} \operatorname{tr}\left(D_k C_j \rho^{\otimes n} C_j^\dagger D_k^\dagger\right) \operatorname{tr}\left(S^k(n) \rho^{\otimes n} S^k(n)\right) \end{aligned}$$

- Nach dem Satz über typische Unterräume gilt mit der Spurerhaltung von $C_j, D_k$:

$$\begin{aligned} F\left(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n\right) &\leq \delta \sum_{jk} \operatorname{tr}\left(D_k C_j \rho^{\otimes n} C_j^\dagger D_k^\dagger\right) \\ &= \delta \end{aligned}$$

$\square$

# Beispiel: Schumacher-Kompression 1

- Sei ein i.i.d. Quantenquelle, die

$$\begin{aligned} |\psi_0\rangle &= |0\rangle \\ |\psi_1\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \end{aligned}$$

  jeweils mit Wahrscheinlichkeit $\frac{1}{2}$ emittiert
- dies entspricht der Dichtematrix

$$\begin{aligned} \rho &= \frac{1}{2}|\psi_0\rangle\langle\psi_0| + \frac{1}{2}|\psi_1\rangle\langle\psi_1| \\ &= \frac{1}{4}\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} \end{aligned}$$

# Beispiel: Schumacher-Kompression 1

- Sei ein i.i.d. Quantenquelle, die

$$\begin{aligned} |\psi_0\rangle &= |0\rangle \\ |\psi_1\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \end{aligned}$$

  jeweils mit Wahrscheinlichkeit $\frac{1}{2}$ emittiert
- dies entspricht der Dichtematrix

$$\begin{aligned} \rho &= \frac{1}{2}|\psi_0\rangle\langle\psi_0| + \frac{1}{2}|\psi_1\rangle\langle\psi_1| \\ &= \frac{1}{4}\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} \end{aligned}$$

# Beispiel: Schumacher-Kompression 2

- mit orthonormaler Dekomposition:

$$\begin{aligned}
\rho &= p\left|\overline{0}\right\rangle\left\langle\overline{0}\right| + (1-p)\left|\overline{1}\right\rangle\left\langle\overline{1}\right| \\
\left|\overline{0}\right\rangle &= \cos\frac{\pi}{8}\left|0\right\rangle + \sin\frac{\pi}{8}\left|1\right\rangle \\
\left|\overline{1}\right\rangle &= -\sin\frac{\pi}{8}\left|0\right\rangle + \cos\frac{\pi}{8}\left|1\right\rangle \\
p &= \frac{1}{4}\left(3+\tan\frac{\pi}{8}\right) \approx 0.85
\end{aligned}$$

- Bzgl. dieser Basis lässt sich ein Block aus $n$ qubits schreiben als

$$\sum_{X=\left\{\overline{00}\ldots\overline{0},\overline{0}\ldots\overline{01},\ldots,\overline{11}\ldots\overline{1}\right\}} C_X\left|X\right\rangle$$

- Nur typische Sequenzen $\left|X\right\rangle$ müssen übertragen werden. Bei diesen ist das Hamminggewicht (Anz. $\overline{1}$en) ca. $n(1-p)$
- Da $p \gg (1-p)$ müssen nur die $\left|X\right\rangle$ berücksichtigt werden, die viele $\overline{0}$en enthalten

# Beispiel: Schumacher-Kompression 2

- mit orthonormaler Dekomposition:

$$\begin{aligned}
\rho &= p\left|\overline{0}\right\rangle\left\langle\overline{0}\right| + (1-p)\left|\overline{1}\right\rangle\left\langle\overline{1}\right| \\
\left|\overline{0}\right\rangle &= \cos\frac{\pi}{8}\left|0\right\rangle + \sin\frac{\pi}{8}\left|1\right\rangle \\
\left|\overline{1}\right\rangle &= -\sin\frac{\pi}{8}\left|0\right\rangle + \cos\frac{\pi}{8}\left|1\right\rangle \\
p &= \frac{1}{4}\left(3+\tan\frac{\pi}{8}\right) \approx 0.85
\end{aligned}$$

- Bzgl. dieser Basis lässt sich ein Block aus $n$ qubits schreiben als

$$\sum_{X=\left\{\overline{00}\ldots\overline{0},\overline{0}\ldots\overline{01},\ldots,\overline{11}\ldots\overline{1}\right\}} C_X\left|X\right\rangle$$

- Nur typische Sequenzen $\left|X\right\rangle$ müssen übertragen werden. Bei diesen ist das Hamminggewicht (Anz. $\overline{1}$en) ca. $n(1-p)$
- Da $p \gg (1-p)$ müssen nur die $\left|X\right\rangle$ berücksichtigt werden, die viele $\overline{0}$en enthalten

# Beispiel: Schumacher-Kompression 2

- mit orthonormaler Dekomposition:

$$\begin{aligned}
\rho &= p\left|\overline{0}\right\rangle\left\langle\overline{0}\right| + (1-p)\left|\overline{1}\right\rangle\left\langle\overline{1}\right| \\
\left|\overline{0}\right\rangle &= \cos\frac{\pi}{8}\left|0\right\rangle + \sin\frac{\pi}{8}\left|1\right\rangle \\
\left|\overline{1}\right\rangle &= -\sin\frac{\pi}{8}\left|0\right\rangle + \cos\frac{\pi}{8}\left|1\right\rangle \\
p &= \frac{1}{4}\left(3+\tan\frac{\pi}{8}\right) \approx 0.85
\end{aligned}$$

- Bzgl. dieser Basis lässt sich ein Block aus $n$ qubits schreiben als

$$\sum_{X=\left\{\overline{00}\ldots\overline{0},\overline{0}\ldots\overline{01},\ldots,\overline{11}\ldots\overline{1}\right\}} C_X\left|X\right\rangle$$

- Nur typische Sequenzen $\left|X\right\rangle$ müssen übertragen werden. Bei diesen ist das Hamminggewicht (Anz. $\overline{1}$en) ca. $n(1-p)$
- Da $p \gg (1-p)$ müssen nur die $\left|X\right\rangle$ berücksichtigt werden, die viele $\overline{0}$en enthalten

# Beispiel: Schumacher-Kompression 2

- mit orthonormaler Dekomposition:

$$\begin{aligned}
\rho &= p\left|\overline{0}\right\rangle\left\langle\overline{0}\right| + (1-p)\left|\overline{1}\right\rangle\left\langle\overline{1}\right| \\
\left|\overline{0}\right\rangle &= \cos\frac{\pi}{8}\left|0\right\rangle + \sin\frac{\pi}{8}\left|1\right\rangle \\
\left|\overline{1}\right\rangle &= -\sin\frac{\pi}{8}\left|0\right\rangle + \cos\frac{\pi}{8}\left|1\right\rangle \\
p &= \frac{1}{4}\left(3+\tan\frac{\pi}{8}\right) \approx 0.85
\end{aligned}$$

- Bzgl. dieser Basis lässt sich ein Block aus $n$ qubits schreiben als

$$\sum_{X=\left\{\overline{00}\ldots\overline{0},\overline{0}\ldots\overline{01},\ldots,\overline{11}\ldots\overline{1}\right\}} C_X\left|X\right\rangle$$

- Nur typische Sequenzen $\left|X\right\rangle$ müssen übertragen werden. Bei diesen ist das Hamminggewicht (Anz. $\overline{1}$en) ca. $n(1-p)$
- Da $p \gg (1-p)$ müssen nur die $\left|X\right\rangle$ berücksichtigt werden, die viele $\overline{0}$en enthalten

# Beispiel: Schumacher-Kompression 3

- Bsp. $n = 3$. Sortiere alle Möglichkeiten aufsteigend nach Hamminggewicht und ordne ihnen eine Kodierung zu

$$\begin{array}{ll} 000 \longrightarrow 000 & 011 \longrightarrow 100 \\ 001 \longrightarrow 001 & 101 \longrightarrow 101 \\ 010 \longrightarrow 010 & 110 \longrightarrow 110 \\ 100 \longrightarrow 011 & 111 \longrightarrow 111 \end{array}$$

- Erwartetes Hamminggewicht liegt bei $(1-p)\,3 \approx 0.44$.
- Da $3S(\rho) = 3(-p\log p - (1-p)\log(1-p)) \approx 1.8 < 2$, reicht folgende Zuordnung

$$\begin{array}{ll} 000 \longrightarrow 00 & 011 \longrightarrow 11 \\ 001 \longrightarrow 01 & 101 \longrightarrow 11 \\ 010 \longrightarrow 10 & 110 \longrightarrow 11 \\ 100 \longrightarrow 11 & 111 \longrightarrow 11 \end{array}$$

- Dieser Code ist "relativ" zuverlässig. Im Grenzwert $n \longrightarrow \infty$ geht die Fehlerwahrscheinlichkeit gegen 0.
- Dies ist nicht der optimale Algorithmus $\longrightarrow$ quantum arithmetic coding

# Beispiel: Schumacher-Kompression 3

- Bsp. $n = 3$. Sortiere alle Möglichkeiten aufsteigend nach Hamminggewicht und ordne ihnen eine Kodierung zu

$$\begin{array}{ll} 000 \to 000 & 011 \to 100 \\ 001 \to 001 & 101 \to 101 \\ 010 \to 010 & 110 \to 110 \\ 100 \to 011 & 111 \to 111 \end{array}$$

- Erwartetes Hamminggewicht liegt bei $(1-p)\,3 \approx 0.44$.
- Da $3S(\rho) = 3(-p\log p - (1-p)\log(1-p)) \approx 1.8 < 2$, reicht folgende Zuordnung

$$\begin{array}{ll} 000 \to 00 & 011 \to 11 \\ 001 \to 01 & 101 \to 11 \\ 010 \to 10 & 110 \to 11 \\ 100 \to 11 & 111 \to 11 \end{array}$$

- Dieser Code ist "relativ" zuverlässig. Im Grenzwert $n \to \infty$ geht die Fehlerwahrscheinlichkeit gegen 0.
- Dies ist nicht der optimale Algorithmus $\to$ quantum arithmetic coding

# Beispiel: Schumacher-Kompression 3

- Bsp. $n = 3$. Sortiere alle Möglichkeiten aufsteigend nach Hamminggewicht und ordne ihnen eine Kodierung zu

$$\begin{array}{ll} 000 \to 000 & 011 \to 100 \\ 001 \to 001 & 101 \to 101 \\ 010 \to 010 & 110 \to 110 \\ 100 \to 011 & 111 \to 111 \end{array}$$

- Erwartetes Hamminggewicht liegt bei $(1-p)\,3 \approx 0.44$.
- Da $3S(\rho) = 3(-p\log p - (1-p)\log(1-p)) \approx 1.8 < 2$, reicht folgende Zuordnung

$$\begin{array}{ll} 000 \to 00 & 011 \to 11 \\ 001 \to 01 & 101 \to 11 \\ 010 \to 10 & 110 \to 11 \\ 100 \to 11 & 111 \to 11 \end{array}$$

- Dieser Code ist "relativ" zuverlässig. Im Grenzwert $n \to \infty$ geht die Fehlerwahrscheinlichkeit gegen 0.
- Dies ist nicht der optimale Algorithmus $\to$ quantum arithmetic coding

# Beispiel: Schumacher-Kompression 3

- Bsp. $n = 3$. Sortiere alle Möglichkeiten aufsteigend nach Hamminggewicht und ordne ihnen eine Kodierung zu

$$\begin{array}{ll} 000 \to 000 & 011 \to 100 \\ 001 \to 001 & 101 \to 101 \\ 010 \to 010 & 110 \to 110 \\ 100 \to 011 & 111 \to 111 \end{array}$$

- Erwartetes Hamminggewicht liegt bei $(1-p)\,3 \approx 0.44$.
- Da $3S(\rho) = 3(-p\log p - (1-p)\log(1-p)) \approx 1.8 < 2$, reicht folgende Zuordnung

$$\begin{array}{ll} 000 \to 00 & 011 \to 11 \\ 001 \to 01 & 101 \to 11 \\ 010 \to 10 & 110 \to 11 \\ 100 \to 11 & 111 \to 11 \end{array}$$

- Dieser Code ist "relativ" zuverlässig. Im Grenzwert $n \to \infty$ geht die Fehlerwahrscheinlichkeit gegen 0.
- Dies ist nicht der optimale Algorithmus $\to$ quantum arithmetic coding

# Beispiel: Schumacher-Kompression 3

- Bsp. $n = 3$. Sortiere alle Möglichkeiten aufsteigend nach Hamminggewicht und ordne ihnen eine Kodierung zu

$$\begin{array}{ll} 000 \to 000 & 011 \to 100 \\ 001 \to 001 & 101 \to 101 \\ 010 \to 010 & 110 \to 110 \\ 100 \to 011 & 111 \to 111 \end{array}$$

- Erwartetes Hamminggewicht liegt bei $(1-p)\,3 \approx 0.44$.
- Da $3S(\rho) = 3(-p\log p - (1-p)\log(1-p)) \approx 1.8 < 2$, reicht folgende Zuordnung

$$\begin{array}{ll} 000 \to 00 & 011 \to 11 \\ 001 \to 01 & 101 \to 11 \\ 010 \to 10 & 110 \to 11 \\ 100 \to 11 & 111 \to 11 \end{array}$$

- Dieser Code ist "relativ" zuverlässig. Im Grenzwert $n \to \infty$ geht die Fehlerwahrscheinlichkeit gegen 0.
- Dies ist nicht der optimale Algorithmus $\to$ quantum arithmetic coding

# Gliederung


1. Information
2. Kodierung ohne Rauschen
3. Kodierung unter Rauschen
4. Ende

# Klassisch: Verrauschter Kanal

**Satz (Shannon: verrauschte Kanalkodierung)**

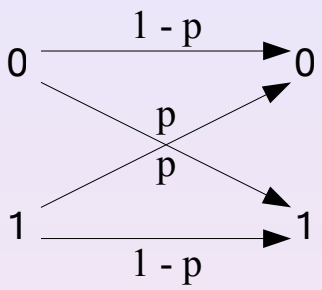
*Für einen verrauschten Kanal $\mathcal{N}$ ist die Kapazität gegeben durch*

$$\mathcal{C}(\mathcal{X}) = \max_{p(x)} H(X : Y)$$

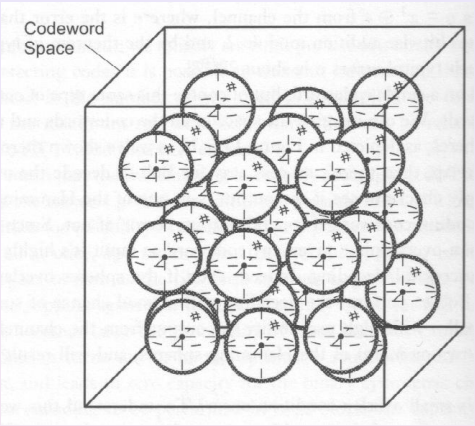
*wobei das Maximum über alle Eingangsverteilungen $p(x)$ für $X$ gebildet wird und $Y$ die zugehörige induzierte Zufallsvariable am Ausgang des Kanals ist.*

- Kanal wird durch einen Satz von bedingten Wahrscheinlichkeiten $p(x|y) \geq 0$ beschrieben.

# Klassischer binärer verrauschter Kanal



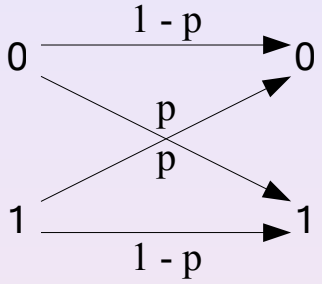

klassischer binärer Kanal mit Fehlerwahrscheinlichkeit $p$



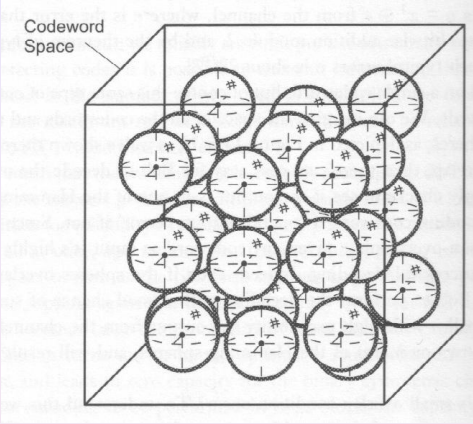

## Optimal ist eine Gleichverteilung am Eingang

$$
\begin{aligned}
\mathcal{C}(\mathcal{X}) &= \max_{p(x)} H(X:Y) \\
&= \max_{p(x)} \left(H(Y) - H(Y|X)\right) \\
&= \max_{p(x)} \left(H(Y) - \sum_x p(x) H(Y|X=x)\right) \\
&= 1 - H_{bin}(p)
\end{aligned}
$$

# Klassischer binärer verrauschter Kanal




klassischer binärer Kanal mit Fehlerwahrscheinlichkeit $p$




## Optimal ist eine Gleichverteilung am Eingang

$$
\begin{aligned}
\mathcal{C}(\mathcal{X}) &= \max_{p(x)} H(X:Y) \\
&= \max_{p(x)} \left(H(Y) - H(Y|X)\right) \\
&= \max_{p(x)} \left(H(Y) - \sum_x p(x)\, H(Y|X=x)\right) \\
&= 1 - H_{bin}(p)
\end{aligned}
$$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 1

- Zu zeigen: für Raten $R < C$ geht die Fehlerwahrscheinlichkeit gegen 0
- Sei $p(x)$ Verteilung, mit der wir ein Wörterbuch generieren.
- Wörterbuch (ist Sender und Empfänger bekannt)

$$\mathcal{C} = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ \vdots & \vdots & \ddots & \vdots \\ x_1\left(2^{nR}\right) & x_2\left(2^{nR}\right) & \dots & x_n\left(2^{nR}\right) \end{bmatrix}$$

- Wahrscheinlichkeit, ein spezielles Wörterbuch zu generieren

$$\Pr(\mathcal{C}) = \prod_{\omega=1}^{2^{nR}} \prod_{i=1}^{n} p(x_i(\omega))$$

- Nachrichten werden gleichverteilt übertragen:

$$\Pr(W = \omega) = 2^{-nR}, \quad \omega = 1, 2, \dots, 2^{nR}$$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 1

- Zu zeigen: für Raten $R < C$ geht die Fehlerwahrscheinlichkeit gegen 0
- Sei $p(x)$ Verteilung, mit der wir ein Wörterbuch generieren.
- Wörterbuch (ist Sender und Empfänger bekannt)

$$\mathcal{C} = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ \vdots & \vdots & \ddots & \vdots \\ x_1\left(2^{nR}\right) & x_2\left(2^{nR}\right) & \dots & x_n\left(2^{nR}\right) \end{bmatrix}$$

- Wahrscheinlichkeit, ein spezielles Wörterbuch zu generieren

$$\Pr(\mathcal{C}) = \prod_{\omega=1}^{2^{nR}} \prod_{i=1}^{n} p(x_i(\omega))$$

- Nachrichten werden gleichverteilt übertragen:

$$\Pr(W = \omega) = 2^{-nR}, \quad \omega = 1, 2, \dots, 2^{nR}$$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 1

- Zu zeigen: für Raten $R < C$ geht die Fehlerwahrscheinlichkeit gegen 0
- Sei $p(x)$ Verteilung, mit der wir ein Wörterbuch generieren.
- Wörterbuch (ist Sender und Empfänger bekannt)

$$\mathcal{C} = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ \vdots & \vdots & \ddots & \vdots \\ x_1\left(2^{nR}\right) & x_2\left(2^{nR}\right) & \dots & x_n\left(2^{nR}\right) \end{bmatrix}$$

- Wahrscheinlichkeit, ein spezielles Wörterbuch zu generieren

$$\Pr(\mathcal{C}) = \prod_{\omega=1}^{2^{nR}} \prod_{i=1}^{n} p(x_i(\omega))$$

- Nachrichten werden gleichverteilt übertragen:

$$\Pr(W = \omega) = 2^{-nR}, \quad \omega = 1, 2, \dots, 2^{nR}$$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 1

- Zu zeigen: für Raten $R < C$ geht die Fehlerwahrscheinlichkeit gegen 0
- Sei $p(x)$ Verteilung, mit der wir ein Wörterbuch generieren.
- Wörterbuch (ist Sender und Empfänger bekannt)

$$\mathcal{C} = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ \vdots & \vdots & \ddots & \vdots \\ x_1\left(2^{nR}\right) & x_2\left(2^{nR}\right) & \dots & x_n\left(2^{nR}\right) \end{bmatrix}$$

- Wahrscheinlichkeit, ein spezielles Wörterbuch zu generieren

$$\Pr(\mathcal{C}) = \prod_{\omega=1}^{2^{nR}} \prod_{i=1}^{n} p(x_i(\omega))$$

- Nachrichten werden gleichverteilt übertragen:

$$\Pr(W = \omega) = 2^{-nR}, \quad \omega = 1, 2, \dots, 2^{nR}$$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 1

- Zu zeigen: für Raten $R < C$ geht die Fehlerwahrscheinlichkeit gegen 0
- Sei $p(x)$ Verteilung, mit der wir ein Wörterbuch generieren.
- Wörterbuch (ist Sender und Empfänger bekannt)

$$\mathcal{C} = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ \vdots & \vdots & \ddots & \vdots \\ x_1\left(2^{nR}\right) & x_2\left(2^{nR}\right) & \dots & x_n\left(2^{nR}\right) \end{bmatrix}$$

- Wahrscheinlichkeit, ein spezielles Wörterbuch zu generieren

$$\Pr(\mathcal{C}) = \prod_{\omega=1}^{2^{nR}} \prod_{i=1}^{n} p(x_i(\omega))$$

- Nachrichten werden gleichverteilt übertragen:

$$\Pr(W = \omega) = 2^{-nR}, \quad \omega = 1, 2, \dots, 2^{nR}$$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 2

- Empfänger empfängt $y^n$ mit Wahrscheinlichkeit

$$P\left(y^n | x^n\left(\omega\right)\right) = \prod_{i=1}^{n} p\left(y_i | x_i\left(\omega\right)\right)$$

- Dekodieren: Es wurde Nachricht $\hat{W}$ geschickt, falls
  - $\left(X^n\left(\hat{W}\right), Y^n\right)$ gleichzeitig typisch sind
  - es keinen anderen Index $k$ gibt, so dass $\left(X^n\left(k\right), Y^n\right)$ gleichzeitig typisch sind
- Falls es kein solches $\hat{W}$ gibt, wird ein Standardeintrag ausgegeben.
- Dekodierfehler bei $W \neq \hat{W}$: entspricht dem Ereignis $\mathcal{E}$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 2

- Empfänger empfängt $y^n$ mit Wahrscheinlichkeit

$$P\left(y^n | x^n\left(\omega\right)\right) = \prod_{i=1}^{n} p\left(y_i | x_i\left(\omega\right)\right)$$

- Dekodieren: Es wurde Nachricht $\hat{W}$ geschickt, falls
  - $\left(X^n\left(\hat{W}\right), Y^n\right)$ gleichzeitig typisch sind
  - es keinen anderen Index $k$ gibt, so dass $\left(X^n\left(k\right), Y^n\right)$ gleichzeitig typisch sind
- Falls es kein solches $\hat{W}$ gibt, wird ein Standardeintrag ausgegeben.
- Dekodierfehler bei $W \neq \hat{W}$: entspricht dem Ereignis $\mathcal{E}$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 2

- Empfänger empfängt $y^n$ mit Wahrscheinlichkeit

$$P\left(y^n | x^n\left(\omega\right)\right) = \prod_{i=1}^{n} p\left(y_i | x_i\left(\omega\right)\right)$$

- Dekodieren: Es wurde Nachricht $\hat{W}$ geschickt, falls
  - $\left(X^n\left(\hat{W}\right), Y^n\right)$ gleichzeitig typisch sind
  - es keinen anderen Index $k$ gibt, so dass $\left(X^n\left(k\right), Y^n\right)$ gleichzeitig typisch sind
- Falls es kein solches $\hat{W}$ gibt, wird ein Standardeintrag ausgegeben.
- Dekodierfehler bei $W \neq \hat{W}$: entspricht dem Ereignis $\mathcal{E}$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 2

- Empfänger empfängt $y^n$ mit Wahrscheinlichkeit

$$P\left(y^n | x^n\left(\omega\right)\right) = \prod_{i=1}^{n} p\left(y_i | x_i\left(\omega\right)\right)$$

- Dekodieren: Es wurde Nachricht $\hat{W}$ geschickt, falls
  - $\left(X^n\left(\hat{W}\right), Y^n\right)$ gleichzeitig typisch sind
  - es keinen anderen Index $k$ gibt, so dass $\left(X^n\left(k\right), Y^n\right)$ gleichzeitig typisch sind
- Falls es kein solches $\hat{W}$ gibt, wird ein Standardeintrag ausgegeben.
- Dekodierfehler bei $W \neq \hat{W}$: entspricht dem Ereignis $\mathcal{E}$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 2

- Empfänger empfängt $y^n$ mit Wahrscheinlichkeit

$$P\left(y^n | x^n\left(\omega\right)\right) = \prod_{i=1}^{n} p\left(y_i | x_i\left(\omega\right)\right)$$

- Dekodieren: Es wurde Nachricht $\hat{W}$ geschickt, falls
  - $\left(X^n\left(\hat{W}\right), Y^n\right)$ gleichzeitig typisch sind
  - es keinen anderen Index $k$ gibt, so dass $\left(X^n\left(k\right), Y^n\right)$ gleichzeitig typisch sind
- Falls es kein solches $\hat{W}$ gibt, wird ein Standardeintrag ausgegeben.
- Dekodierfehler bei $W \neq \hat{W}$: entspricht dem Ereignis $\mathcal{E}$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 2

- Empfänger empfängt $y^n$ mit Wahrscheinlichkeit

$$P\left(y^n | x^n\left(\omega\right)\right) = \prod_{i=1}^{n} p\left(y_i | x_i\left(\omega\right)\right)$$

- Dekodieren: Es wurde Nachricht $\hat{W}$ geschickt, falls
  - $\left(X^n\left(\hat{W}\right), Y^n\right)$ gleichzeitig typisch sind
  - es keinen anderen Index $k$ gibt, so dass $\left(X^n\left(k\right), Y^n\right)$ gleichzeitig typisch sind
- Falls es kein solches $\hat{W}$ gibt, wird ein Standardeintrag ausgegeben.
- Dekodierfehler bei $W \neq \hat{W}$: entspricht dem Ereignis $\mathcal{E}$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 3

- Fehlerwahrscheinlichkeit hängt nicht vom Sendewort ab:

$$\begin{aligned}\Pr(\mathcal{E}) &= \sum_{\mathcal{C}} P(\mathcal{C}) P_e^{(n)}(\mathcal{C}) \\ &= \sum_{\mathcal{C}} P(\mathcal{C}) \lambda_1(\mathcal{C}) \\ &= \Pr(\mathcal{E}|W=1)\end{aligned}$$

- Definiere Ereignis $E_i = \{(X^n(i), Y^n)$ sind gleichzeitig typisch$\}$. Für genügend große $n$ und $R < H(X:Y) - 3\epsilon$ gilt:

$$\begin{aligned}\Pr(\mathcal{E}|W=1) &= P\left(E_1^c \cup E_2 \cup \ldots \cup E_{2^{nR}}\right) \\ &\leq P(E_1^c) + \sum_{i=2}^{2^{nR}} P(E_i) \\ &\leq \epsilon + \sum_{i=2}^{2^{nR}} 2^{-n(H(X:Y)-3\epsilon)} \\ &\leq 2\epsilon\end{aligned}$$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 3

- Fehlerwahrscheinlichkeit hängt nicht vom Sendewort ab:

$$\begin{aligned}\Pr(\mathcal{E}) &= \sum_{\mathcal{C}} P(\mathcal{C}) P_e^{(n)}(\mathcal{C}) \\ &= \sum_{\mathcal{C}} P(\mathcal{C}) \lambda_1(\mathcal{C}) \\ &= \Pr(\mathcal{E}|W=1)\end{aligned}$$

- Definiere Ereignis $E_i = \{(X^n(i), Y^n)$ sind gleichzeitig typisch$\}$. Für genügend große $n$ und $R < H(X:Y) - 3\epsilon$ gilt:

$$\begin{aligned}\Pr(\mathcal{E}|W=1) &= P\left(E_1^c \cup E_2 \cup \ldots \cup E_{2^{nR}}\right) \\ &\leq P(E_1^c) + \sum_{i=2}^{2^{nR}} P(E_i) \\ &\leq \epsilon + \sum_{i=2}^{2^{nR}} 2^{-n(H(X:Y)-3\epsilon)} \\ &\leq 2\epsilon\end{aligned}$$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 4

- Da die durchschnittliche Fehlerwahrscheinlichkeit über alle Wörterbücher $\leq 2\varepsilon$ ist, gibt es mindestens ein $\mathcal{C}^*$ mit $P_e^n(\mathcal{C}^*) \leq 2\varepsilon$.
- Verwerfe im Wörterbuch die schlechteste Hälfte der Zuordnungen. Aus

$$P_e^n(\mathcal{C}^*) = \frac{1}{2^{nR}} \sum_i \lambda_i(\mathcal{C}^*) \leq 2\epsilon$$

  folgt, das für alle übrigen Zuordnungen gilt: $\lambda_i(\mathcal{C}^*) \leq 4\epsilon$.
- Somit haben wir nur noch $2^{nR-1}$ Wörterbucheinträge, womit die Rate von $R$ auf $R - \frac{1}{n}$ gesunken ist.
- Insgesamt haben wir einen Code konstruiert mit $R' = R - \frac{1}{n}$ und Fehlerwahrscheinlichkeit $\lambda^{(n)} \leq 4\epsilon$.

$\square$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 4

- Da die durchschnittliche Fehlerwahrscheinlichkeit über alle Wörterbücher $\leq 2\varepsilon$ ist, gibt es mindestens ein $\mathcal{C}^*$ mit $P_e^n(\mathcal{C}^*) \leq 2\varepsilon$.
- Verwerfe im Wörterbuch die schlechteste Hälfte der Zuordnungen. Aus
$$P_e^n(\mathcal{C}^*) = \frac{1}{2^{nR}} \sum_i \lambda_i(\mathcal{C}^*) \leq 2\epsilon$$
folgt, das für alle übrigen Zuordnungen gilt: $\lambda_i(\mathcal{C}^*) \leq 4\epsilon$.
- Somit haben wir nur noch $2^{nR-1}$ Wörterbucheinträge, womit die Rate von $R$ auf $R - \frac{1}{n}$ gesunken ist.
- Insgesamt haben wir einen Code konstruiert mit $R' = R - \frac{1}{n}$ und Fehlerwahrscheinlichkeit $\lambda^{(n)} \leq 4\epsilon$.

$\square$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 4

- Da die durchschnittliche Fehlerwahrscheinlichkeit über alle Wörterbücher $\leq 2\varepsilon$ ist, gibt es mindestens ein $\mathcal{C}^*$ mit $P_e^n(\mathcal{C}^*) \leq 2\varepsilon$.
- Verwerfe im Wörterbuch die schlechteste Hälfte der Zuordnungen. Aus
$$P_e^n(\mathcal{C}^*) = \frac{1}{2^{nR}} \sum_i \lambda_i(\mathcal{C}^*) \leq 2\epsilon$$
folgt, das für alle übrigen Zuordnungen gilt: $\lambda_i(\mathcal{C}^*) \leq 4\epsilon$.
- Somit haben wir nur noch $2^{nR-1}$ Wörterbucheinträge, womit die Rate von $R$ auf $R - \frac{1}{n}$ gesunken ist.
- Insgesamt haben wir einen Code konstruiert mit $R' = R - \frac{1}{n}$ und Fehlerwahrscheinlichkeit $\lambda^{(n)} \leq 4\epsilon$.

$\square$

# Beweis: Shannon - verrauschte Kanalkodierung - Teil 4

- Da die durchschnittliche Fehlerwahrscheinlichkeit über alle Wörterbücher $\leq 2\varepsilon$ ist, gibt es mindestens ein $\mathcal{C}^*$ mit $P_e^n(\mathcal{C}^*) \leq 2\varepsilon$.
- Verwerfe im Wörterbuch die schlechteste Hälfte der Zuordnungen. Aus
$$P_e^n(\mathcal{C}^*) = \frac{1}{2^{nR}} \sum_i \lambda_i(\mathcal{C}^*) \leq 2\epsilon$$
folgt, das für alle übrigen Zuordnungen gilt: $\lambda_i(\mathcal{C}^*) \leq 4\epsilon$.
- Somit haben wir nur noch $2^{nR-1}$ Wörterbucheinträge, womit die Rate von $R$ auf $R - \frac{1}{n}$ gesunken ist.
- Insgesamt haben wir einen Code konstruiert mit $R' = R - \frac{1}{n}$ und Fehlerwahrscheinlichkeit $\lambda^{(n)} \leq 4\epsilon$.
$\square$

# Beweis: Shannon - verrauschte Kanalkodierung - Umkehrung

- Umgekehrt gilt:
- $nR = H(W) = H(W|Y^n) + H(W:Y^n)$
- $\leq H(W|Y^n) + I(X^n(W)|Y^n)$
- Fanos Ungleichung
  $\leq 1 + P_e^{(n)} nR + I(X^n(W)|Y^n)$
- $\leq 1 + P_e^{(n)} nR + nC$
- $P_e^{(n)} \geq 1 - \frac{1}{nR} - \frac{C}{R} \geq 0 \Rightarrow R > C$
  $\square$

# Beweis: Shannon - verrauschte Kanalkodierung - Umkehrung

- Umgekehrt gilt:
- $nR = H(W) = H(W|Y^n) + H(W : Y^n)$
- $\leq H(W|Y^n) + I(X^n(W)|Y^n)$
- Fanos Ungleichung
  $\leq 1 + P_e^{(n)} nR + I(X^n(W)|Y^n)$
- $\leq 1 + P_e^{(n)} nR + nC$
- $P_e^{(n)} \geq 1 - \frac{1}{nR} - \frac{C}{R} \geq 0 \Rightarrow R > C$
  $\square$

# Beweis: Shannon - verrauschte Kanalkodierung - Umkehrung

- Umgekehrt gilt:
- $nR = H(W) = H(W|Y^n) + H(W : Y^n)$
- $\leq H(W|Y^n) + I(X^n(W)|Y^n)$
- Fanos Ungleichung
  $\leq 1 + P_e^{(n)} nR + I(X^n(W)|Y^n)$
- $\leq 1 + P_e^{(n)} nR + nC$
- $P_e^{(n)} \geq 1 - \frac{1}{nR} - \frac{C}{R} \geq 0 \Rightarrow R > C$
  $\square$

# Beweis: Shannon - verrauschte Kanalkodierung - Umkehrung

- Umgekehrt gilt:
- $nR = H(W) = H(W|Y^n) + H(W : Y^n)$
- $\leq H(W|Y^n) + I(X^n(W)|Y^n)$
- Fanos Ungleichung
  $\leq 1 + P_e^{(n)} nR + I(X^n(W)|Y^n)$
- $\leq 1 + P_e^{(n)} nR + nC$
- $P_e^{(n)} \geq 1 - \frac{1}{nR} - \frac{C}{R} \geq 0 \Rightarrow R > C$
  $\square$

# Beweis: Shannon - verrauschte Kanalkodierung - Umkehrung

- Umgekehrt gilt:
- $nR = H(W) = H(W|Y^n) + H(W : Y^n)$
- $\leq H(W|Y^n) + I(X^n(W)|Y^n)$
- Fanos Ungleichung
  $\leq 1 + P_e^{(n)} nR + I(X^n(W)|Y^n)$
- $\leq 1 + P_e^{(n)} nR + nC$
- $P_e^{(n)} \geq 1 - \frac{1}{nR} - \frac{C}{R} \geq 0 \Rightarrow R > C$
  $\square$

# Beweis: Shannon - verrauschte Kanalkodierung - Umkehrung

- Umgekehrt gilt:
- $nR = H(W) = H(W|Y^n) + H(W : Y^n)$
- $\leq H(W|Y^n) + I(X^n(W)|Y^n)$
- Fanos Ungleichung
  $\leq 1 + P_e^{(n)} nR + I(X^n(W)|Y^n)$
- $\leq 1 + P_e^{(n)} nR + nC$
- $P_e^{(n)} \geq 1 - \frac{1}{nR} - \frac{C}{R} \geq 0 \Rightarrow R > C$
  $\square$

# HSW-Theorem

**Satz (Holevo-Schumacher-Westmoreland (HSW))**

*Sei $\varepsilon$ eine spurerhaltende Quantenoperation. Die Produktzustandskapazität ist*

$$C^{(1)}(\varepsilon) = \max_{\{p_j, \rho_j\}} \left[ S\left(\varepsilon\left(\sum_j p_j \rho_j\right)\right) - \sum_j p_j S\left(\varepsilon\left(\rho_j\right)\right) \right]$$

*wobei das Maximum über alle Ensemble $\{p_j, \rho_j\}$ von möglichen Eingangszuständen $\rho_j$ des Kanals gebildet wird.*

- Semiklassisch: Alice hat eine Menge von Symbolen $\{i\}$ in Form von Dichteoperatoren $\rho_i$ kodiert.
- Eine Sequenz der Länge $n$ bildet Alice per Produktzustand

$$\rho_M = \rho_{M_1} \otimes \rho_{M_2} \otimes \ldots \otimes \rho_{M_n}$$

- Bob kann beliebige Operationen auf $\varepsilon^{\otimes n}(\rho_M)$ durchführen.

# HSW-Theorem

**Satz (Holevo-Schumacher-Westmoreland (HSW))**

*Sei $\varepsilon$ eine spurerhaltende Quantenoperation. Die Produktzustandskapazität ist*

$$C^{(1)}(\varepsilon) = \max_{\{p_j, \rho_j\}} \left[ S\left(\varepsilon\left(\sum_j p_j \rho_j\right)\right) - \sum_j p_j S\left(\varepsilon\left(\rho_j\right)\right) \right]$$

*wobei das Maximum über alle Ensemble $\{p_j, \rho_j\}$ von möglichen Eingangszuständen $\rho_j$ des Kanals gebildet wird.*

- Semiklassisch: Alice hat eine Menge von Symbolen $\{i\}$ in Form von Dichteoperatoren $\rho_i$ kodiert.
- Eine Sequenz der Länge $n$ bildet Alice per Produktzustand

$$\rho_M = \rho_{M_1} \otimes \rho_{M_2} \otimes \ldots \otimes \rho_{M_n}$$

- Bob kann beliebige Operationen auf $\varepsilon^{\otimes n}(\rho_M)$ durchführen.

# HSW-Theorem

**Satz (Holevo-Schumacher-Westmoreland (HSW))**

*Sei $\varepsilon$ eine spurerhaltende Quantenoperation. Die Produktzustandskapazität ist*

$$C^{(1)}(\varepsilon) = \max_{\{p_j, \rho_j\}} \left[ S\left(\varepsilon\left(\sum_j p_j \rho_j\right)\right) - \sum_j p_j S\left(\varepsilon\left(\rho_j\right)\right) \right]$$

*wobei das Maximum über alle Ensemble $\{p_j, \rho_j\}$ von möglichen Eingangszuständen $\rho_j$ des Kanals gebildet wird.*

- Semiklassisch: Alice hat eine Menge von Symbolen $\{i\}$ in Form von Dichteoperatoren $\rho_i$ kodiert.
- Eine Sequenz der Länge $n$ bildet Alice per Produktzustand

$$\rho_M = \rho_{M_1} \otimes \rho_{M_2} \otimes \ldots \otimes \rho_{M_n}$$

- Bob kann beliebige Operationen auf $\varepsilon^{\otimes n}(\rho_M)$ durchführen.

# HSW-Theorem

## Satz (Holevo-Schumacher-Westmoreland (HSW))

*Sei $\varepsilon$ eine spurerhaltende Quantenoperation. Die Produktzustandskapazität ist*

$$C^{(1)}(\varepsilon) = \max_{\{p_j, \rho_j\}} \left[ S\left(\varepsilon\left(\sum_j p_j \rho_j\right)\right) - \sum_j p_j S\left(\varepsilon\left(\rho_j\right)\right) \right]$$

*wobei das Maximum über alle Ensemble $\{p_j, \rho_j\}$ von möglichen Eingangszuständen $\rho_j$ des Kanals gebildet wird.*

- Semiklassisch: Alice hat eine Menge von Symbolen $\{i\}$ in Form von Dichteoperatoren $\rho_i$ kodiert.
- Eine Sequenz der Länge $n$ bildet Alice per Produktzustand

$$\rho_M = \rho_{M_1} \otimes \rho_{M_2} \otimes \ldots \otimes \rho_{M_n}$$

- Bob kann beliebige Operationen auf $\varepsilon^{\otimes n}(\rho_M)$ durchführen.

# Gliederung


1. Information
2. Kodierung ohne Rauschen
3. Kodierung unter Rauschen
4. Ende

# Literatur

Michael A. Nielsen and Isaac L. Chuang.
*Quantum Computation and Quantum Information*.
Cambridge University Press, October 2000.

Wikipedia.
Shannon-Fano-Kodierung — Wikipedia, Die freie Enzyklopädie, 2007.
[Online; Stand 18. Mai 2007].

# Fragen?

Kuhle Sache:
Mit KUHbits
gibt's eimerweise
tolle Prämien ...
KUHbit
MILCHLAND
NIEDERSACHSEN
www.milchwirtschaft.de