2018年8月2日

量子情報・物性の新潮流（東京大学 物性研究所）

量子情報スペクトル理論の発展と応用

電気通信大学 大学院情報理工学研究科

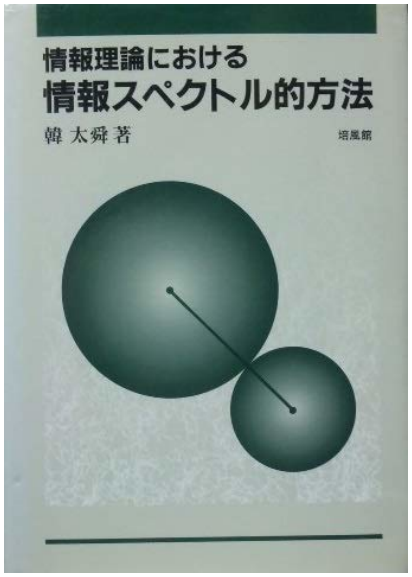
小川朋宏

目次

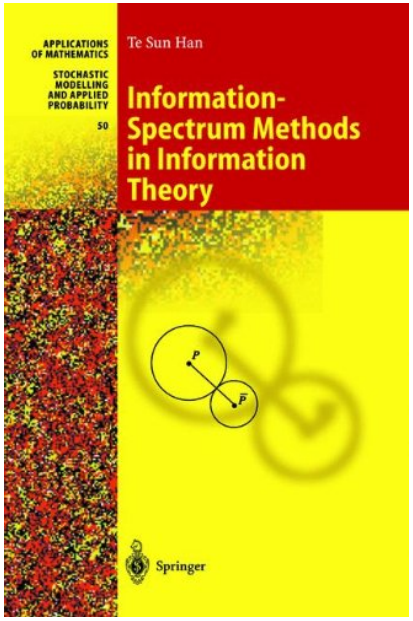
1. **エントロピーから情報スペクトルへ**
2. **ダイバージェンス・スペクトル**
3. **古典・量子通信路と相互情報量スペクトル**
4. cq wiretap channel coding, resolvability
5. **まとめと展望**

主要な参考文献（元祖）

- T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inform. Theory*, vol. 39, pp. 752–772, 1993.
- S. Verdú and T.S. Han, "A general formula for channel capacity," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1147–1157, 1994.
- **韓太舜**, **情報理論における情報スペクトル的方法**, **培風館**, 1998.
- T. S. Han, *Information-Spectrum Methods in Information Theory*, Springer, 2002.



$\Longrightarrow$ **英訳**



主要な参考文献（量子系）

- H. Nagaoka, "On asymptotic theory of quantum hypothesis testing," in Proc. *Symp. Statistical Inference Theory and its Information Theoretical Aspect*, pp. 49–52, 1998 (in Japanese).
- H. Nagaoka and M. Hayashi, "An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses," *IEEE Trans. Inform. Theory*, vol. 53, pp. 534–549, 2007.
- M. Hayashi and H. Nagaoka, "General formulas for capacity of classical-quantum channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1753–1768, 2003.
- M. Hayashi, *Quantum Information Theory: An Introduction*, Springer, 2006 (Japanese edition: Science-Press, 2004).
- G. Bowen and N. Datta, "Beyond i.i.d in quantum information theory," in Proc. *2006 IEEE ISIT*, Seattle, USA, July, pp. 451–455, arXiv:quant-ph/0604013.
- N. Datta and R. Renner, "Smooth Renyi entropies and the quantum information spectrum," *IEEE Trans. Inform. Theory*, vol. 55, pp. 2807–2815, 2009.

参考文献（wiretap channel, resolvabilityなど）

- N. Datta, "Min- and Max- relative entropies and a new entanglement monotone," *IEEE Trans. Inform. Theory*, vol. 55, pp. 2816–2826, 2009.
- A. D. Wyner, "The Wire-tap channel," The Bell System Technical Journal, vol. 54, pp. 1355–1387, 1975.
- I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, pp. 339–348, 1978.
- R. Ahlswede and A. Winter, "Strong converse for identification via quantum channels," *IEEE Trans. Inform. Theory*, vol. 48, No. 3, 569–579, 2002.
- I. Devetak, "The private classical capacity and quantum capacity of a quantum channel," *IEEE Trans. Inform. Theory*, vol. 51, pp. 44–55, 2005.
- M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1562–1575, 2006.
- F. Hiai and D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," *Commun. Math. Phys.*, vol. 143, pp. 99–114, 1991.
- T. Ogawa and H. Nagaoka, "Strong converse and Stein's lemma in quantum hypothesys testing," *IEEE Trans. Inform. Theory*, vol. 56, pp. 2428–2433, 2000.

1　エントロピーから情報スペクトルへ

- Shannon**エントロピー：**
  確率分布$P(x)$**に従う確率変数**X**に対して**

$$H(X) = H(P) = -\sum_x P(x) \log P(x) = E_P[-\log P(X)]$$

- Shannon**の情報源符号化定理をレビュー（**i.i.d. **情報源において）**

最適なデータ圧縮のレート（圧縮率）$= H(X)$

- **情報スペクトル理論のアイデアを概説**

i.i.d. (independently and identically distributed)

- **確率変数 $X_1 X_2 \ldots X_n$ が独立に同一の確率分布 $P(x)$ に従っている**

$$X^n := X_1 X_2 \ldots X_n \overset{\text{i.i.d.}}{\sim} P(x)$$

- **各データ列 $x^n := x_1 x_2 \ldots x_n$ の確率は**

$$P_{X^n}(x^n) := P(x_1) P(x_2) \cdots P(x_n)$$

- **固定長符号化をレビュー**

$x^n := x_1 x_2 \ldots x_n$

$\downarrow f_n$: **符号器**(encoder)

$k \in \{ 1, 2, \ldots, M_n \}$

$\downarrow g_n$: **復号器**(decoder)

$g_n(f_n(x^n))$

- **誤り確率**

$$\Pr\{ X^n \neq g_n(f_n(X^n)) \}$$

- **符号化レート（圧縮率）**

$$\frac{\text{符号語のビット長}}{\text{データのビット長}} = \frac{\log M_n}{n}$$

誤りゼロだと圧縮にならないことに注意

データ圧縮のアイデア（最適な符号化）

- **可変長符号化：平均符号長を短くする，誤りなし**
  - 出現確率の大きいデータは短く
  - 出現確率の小さいデータは長く

$$\text{データ列} x^n \text{の最適符号長：} \quad -\log P_{X^n}(x^n)$$

これにより平均的な圧縮率はエントロピー$H(X)$になる

- **固定長符号化：誤り確率と圧縮率のトレードオフ**
  - 出現確率の大きいデータx^nをまじめに符号化：

$$-\frac{1}{n}\log P_{X^n}(x^n) \le a \quad (\Leftrightarrow P_{X^n}(x^n) \ge e^{-na})$$

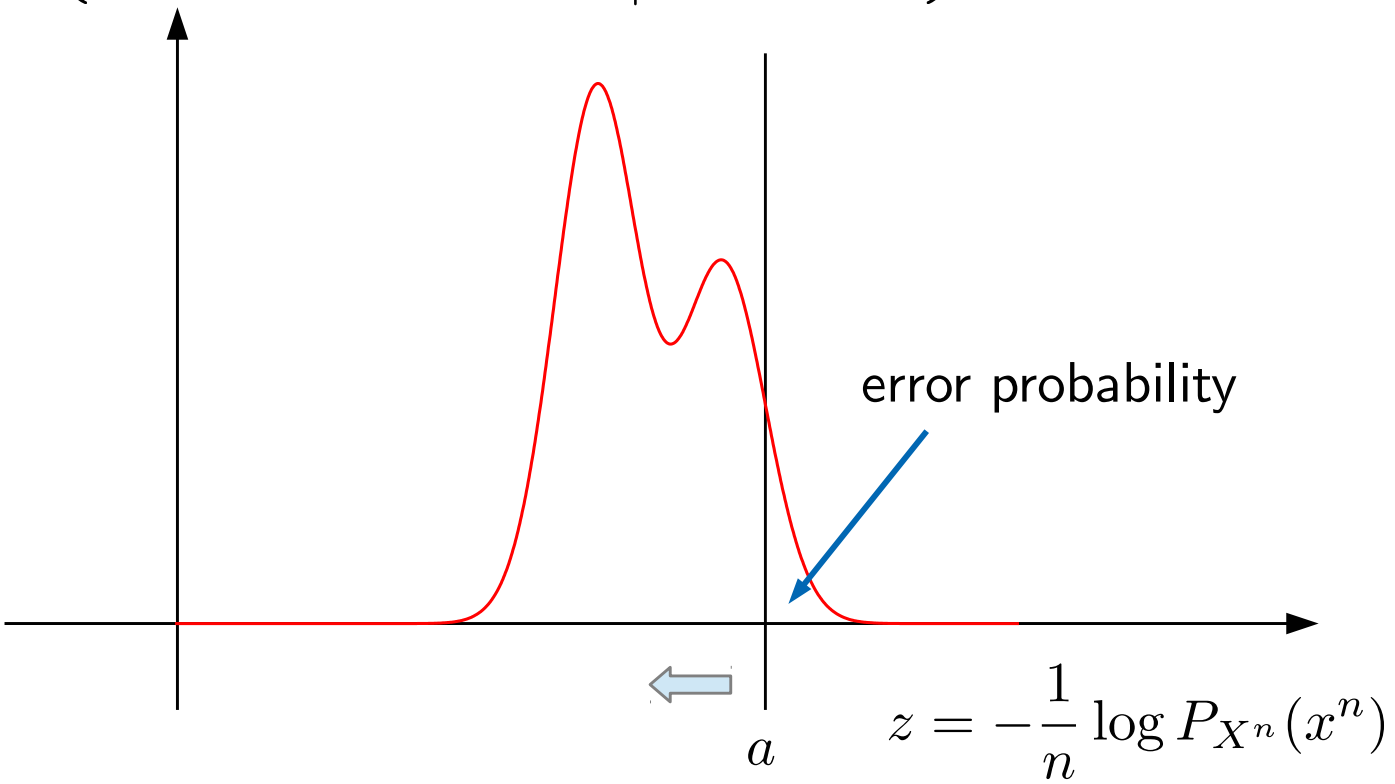
  - 出現確率の小さいデータはあきらめる（誤り確率になる）

$$-\frac{1}{n}\log P_{X^n}(x^n) > a \quad (\Leftrightarrow P_{X^n}(x^n) < e^{-na})$$

$a \in \mathbb{R}$はトレードオフパラメータ（$\simeq$ 圧縮率）

最適な符号化における誤り確率

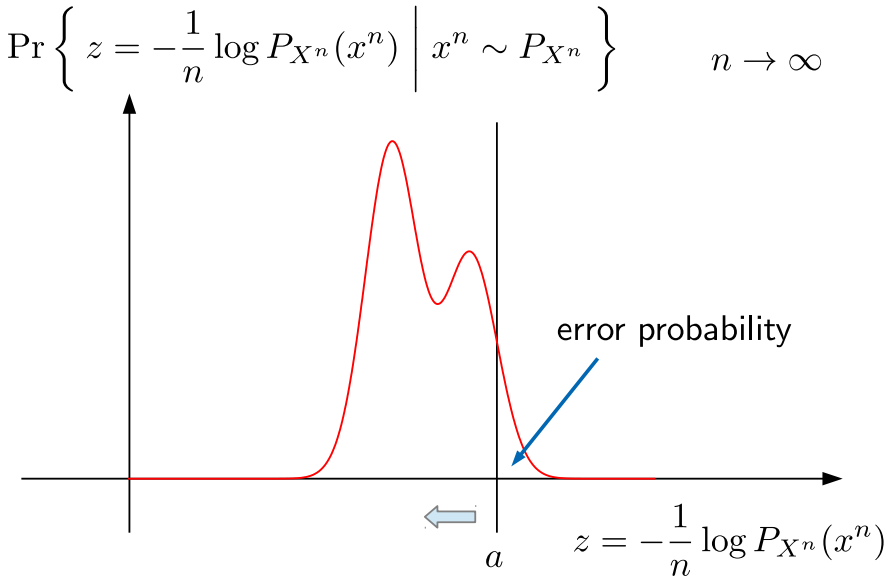
$$\Pr\left\{ z = -\frac{1}{n}\log P_{X^n}(x^n) \,\middle|\, x^n \sim P_{X^n} \right\} \qquad n \to \infty$$



復号成功確率： $\Pr\left\{ -\frac{1}{n}\log P_{X^n}(x^n) \le a \,\middle|\, x^n \sim P_{X^n} \right\}$

復号誤り確率： $\Pr\left\{ -\frac{1}{n}\log P_{X^n}(x^n) > a \,\middle|\, x^n \sim P_{X^n} \right\}$

最適な符号化における符号化レート（圧縮率）



正しく復号される要素数M_nは？

$$M_n := \#\left\{ x^n \,\middle|\, -\frac{1}{n}\log P_{X^n}(x^n) \le a \right\} \le e^{na}$$

$$\because \quad 1 \ge \sum_{x^n:\, P_{X^n}(x^n) \ge e^{-na}} P_{X^n}(x^n) \ge \sum_{x^n:\, P_{X^n}(x^n) \ge e^{-na}} e^{-na} = M_n \cdot e^{-na}$$

よって圧縮率はa以下である

$$\frac{\text{符号語のビット長}}{\text{データのビット長}} = \frac{\log M_n}{n} \le a$$

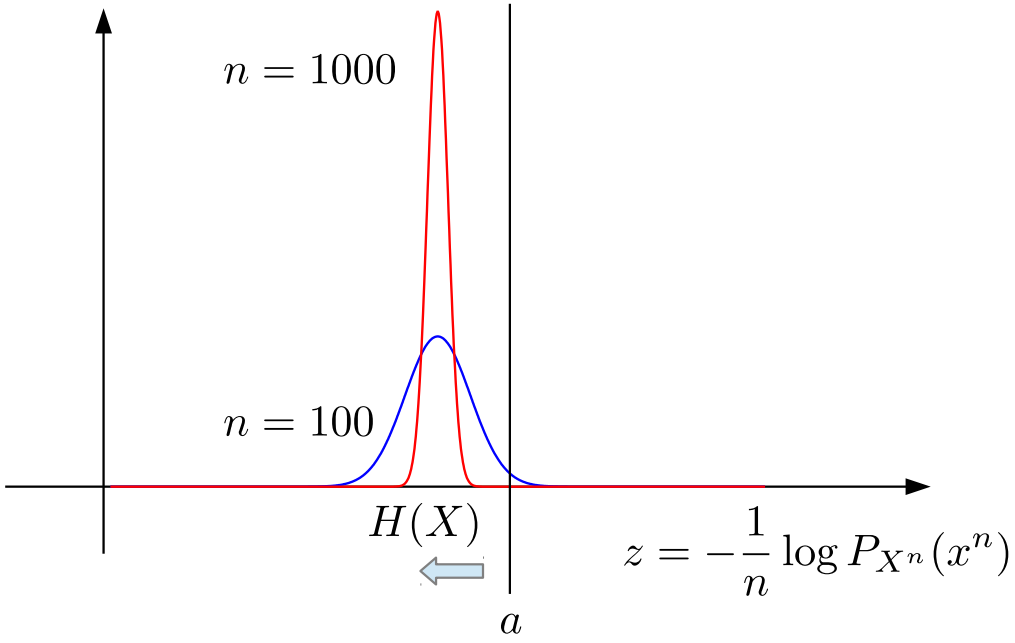
大数の法則とエントロピー

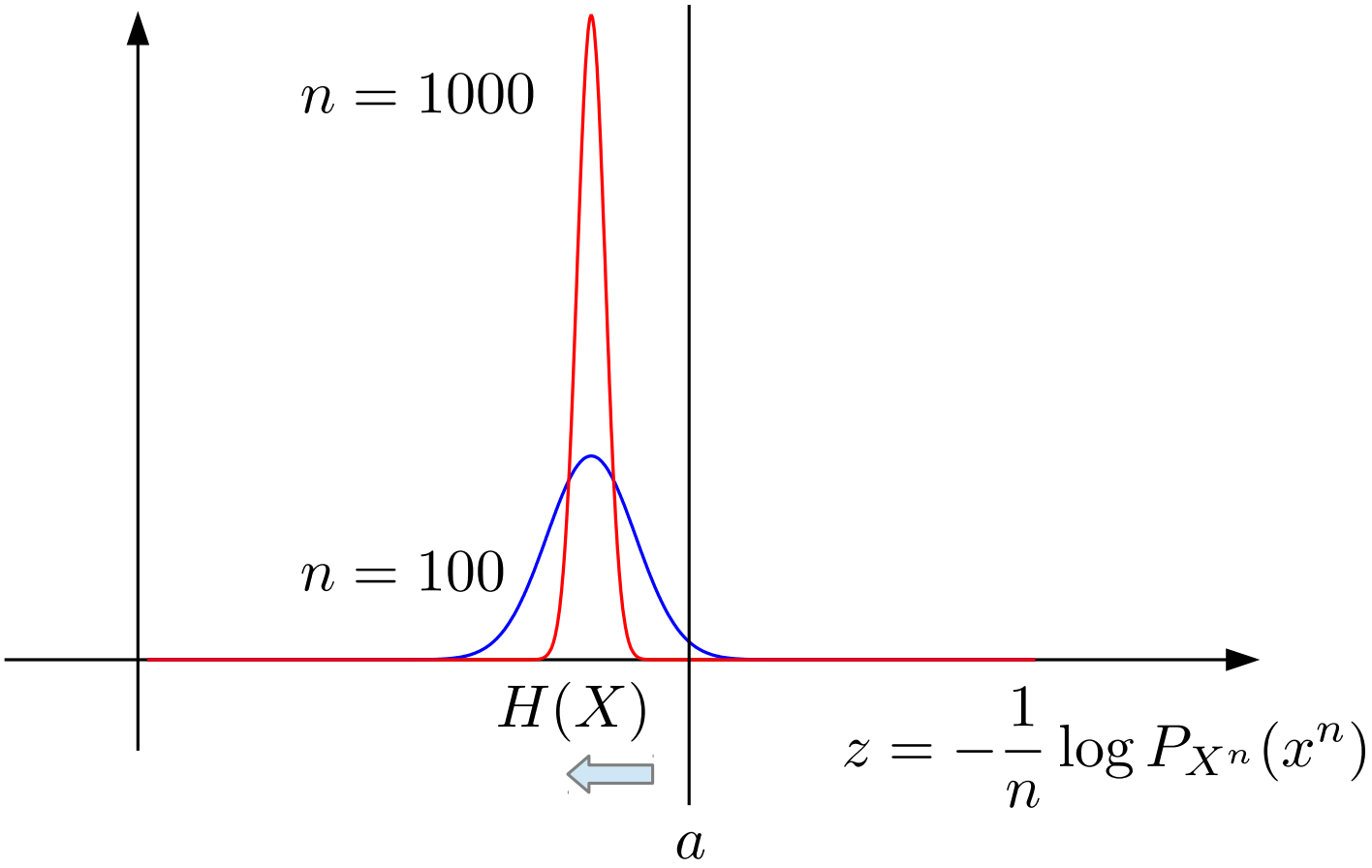
i.i.d. **条件** $P_{X^n}(x^n) = P(x_1)P(x_2)\cdots P(x_n)$ **では大数の法則より**

$$-\frac{1}{n}\log P_{X^n}(x^n) = -\frac{1}{n}\log P(x_1)P(x_2)\cdots P(x_n)$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\log P(x_i) \quad \text{（サンプル平均）}$$

$$\xrightarrow{\text{確率収束}\ (n\to\infty)} \quad E_P[-\log P(X)] = H(X) \quad \text{（アンサンブル平均）}$$





このとき圧縮率aにおける漸近的な復号成功確率は

$$\lim_{n\to\infty} \Pr\left\{ -\frac{1}{n} \log P_{X^n}(x^n) \le a \;\middle|\; x^n \sim P_{X^n} \right\} = \begin{cases} 1 & a > H(X) \\ 0 & a < H(X) \end{cases}$$

漸近的に誤りなしで圧縮できる圧縮率の限界値はエントロピー$H(X)$

情報スペクトル理論の設定

確率変数の列 $\widehat{X} = \{X^{(n)}\}_{n=1}^{\infty}$ **を考える（オリジナルの記法：**$\mathbb{X} = \{X^{(n)}\}_{n=1}^{\infty}$**）**

$$
\begin{aligned}
X^{(1)} &= X_1^{(1)} \\
X^{(2)} &= X_1^{(2)} X_2^{(2)} \\
&\vdots \\
X^{(n)} &= X_1^{(n)} X_2^{(n)} \cdots X_n^{(n)}
\end{aligned}
$$

- **定常性，エルゴード性など，確率構造を一切仮定しない**
- **整合性条件**(compatibility)**：実際上重要だが，理論上不要である**

$$
P_{X^{(n)}}(x_1, x_2, \ldots, x_n) = \sum_{x_{n+1}} P_{X^{(n+1)}}(x_1, x_2, \ldots, x_n, x_{n+1})
$$

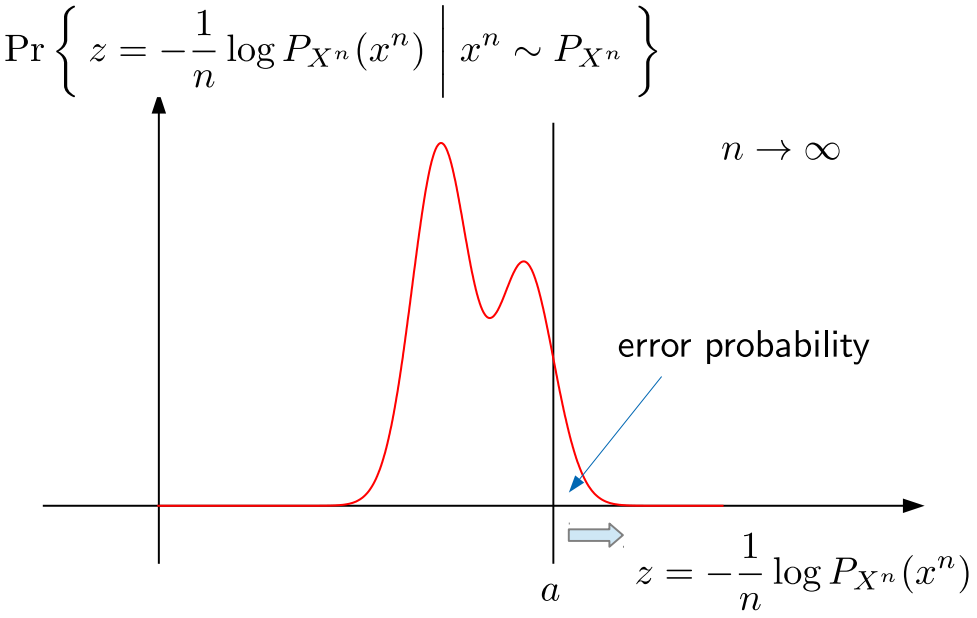
- $\widehat{X} = \{X^{(n)}\}_{n=1}^{\infty}$ **を一般情報源**(general source)**と呼ぶ**
- **数学的に簡単にするため離散確率変数の列を扱う**

エントロピースペクトルレート

一般情報源 $\widehat{X} = \{X^{(n)}\}_{n=1}^{\infty}$ **と** $\varepsilon \in [0,1]$ **に対して，**

定義（ε-エントロピースペクトル上限）

$$\overline{H}(\varepsilon|\widehat{X}) := \inf\left\{ a \,\middle|\, \limsup_{n\to\infty} \Pr\left\{ -\frac{1}{n}\log P_{X^{(n)}}(X^{(n)}) > a \right\} \le \varepsilon \right\}$$



- **解釈：漸近的な最悪の誤りがεでおさまっている圧縮レートaの最適値**
- **符号化問題の最適レートを定義に押し込めただけ？ →** YES **（尤度を使っているのがポイント）**

データ圧縮（固定長符号化）における一般公式

一般情報源 $\widehat{X} = \{X^{(n)}\}_{n=1}^{\infty}$ と $\varepsilon \in [0,1]$ に対して，

定義（ε-最適符号化レート）

$$R(\varepsilon|\widehat{X}) := \inf\Bigg\{ R \mid \exists \{(f_n, g_n)\}_{n=1}^{\infty},\ \text{s.t.}$$
$$\limsup_{n\to\infty} \Pr\left\{ X^{(n)} \neq f_n \circ g_n(X^{(n)}) \right\} \leq \varepsilon,$$
$$\limsup_{n\to\infty} \frac{1}{n} \log M_n \leq R \Bigg\}$$

上手い符号器・複合器の “列” があって，漸近的な誤り確率が ε 以下で，
圧縮率が R 以下となる R の限界値

定理 (Han-Verdú 93, Han 98 **の一般公式**)

$$R(\varepsilon|\widehat{X}) = \overline{H}(\varepsilon|\widehat{X})$$

尤度比を使うことの最適性から導かれる

情報スペクトル的方法の考え方

(韓&長岡1990年代後半)

- i.i.d. の場合はエントロピー $H(X)$ が最適レートであった

$$R(\varepsilon|\widehat{X}) \overset{(1)}{=} \overline{H}(\varepsilon|\widehat{X}) \overset{(2)}{=} H(X) \quad (0 \leq \forall \varepsilon < 1)$$

(skeleton)

- (1) は**一般公式** (general formula) と呼ばれる
  - ☺確率構造に依存しない
  - ☺シンプルな議論から導かれ，ある種の様式美を備える
  - ☹一般公式だけでは実際の問題に適用することは出来ない

個別に確率構造を付加すると，実際に適用可能な符号化定理が得られる

$$R(\varepsilon|\widehat{X}) \overset{(1)}{=} \overline{H}(\varepsilon|\widehat{X}) \overset{(2)}{=} H(X) \quad (0 \leq \forall \varepsilon < 1)$$

- **例：** i.i.d. **情報源**, **マルコフ情報源**, **定常エルゴード情報源**, **混合情報源**
- **量子系：** i.i.d. **状態**, Finitely Correlated State（AKLT **状態**），Gibbs **状態**
- ☹(2) **の証明も簡単とは限らない**
- ☺ **ただし，**(2) **は純粋に確率論，量子確率論の問題である**

情報スペクトル的方法によって，問題をクリアーに分離することが可能☺

(1) **符号化・操作の最適性に関する議論**　(2) **確率論の議論**

エントロピースペクトル上限と下限

定義（ε-エントロピースペクトル上限）

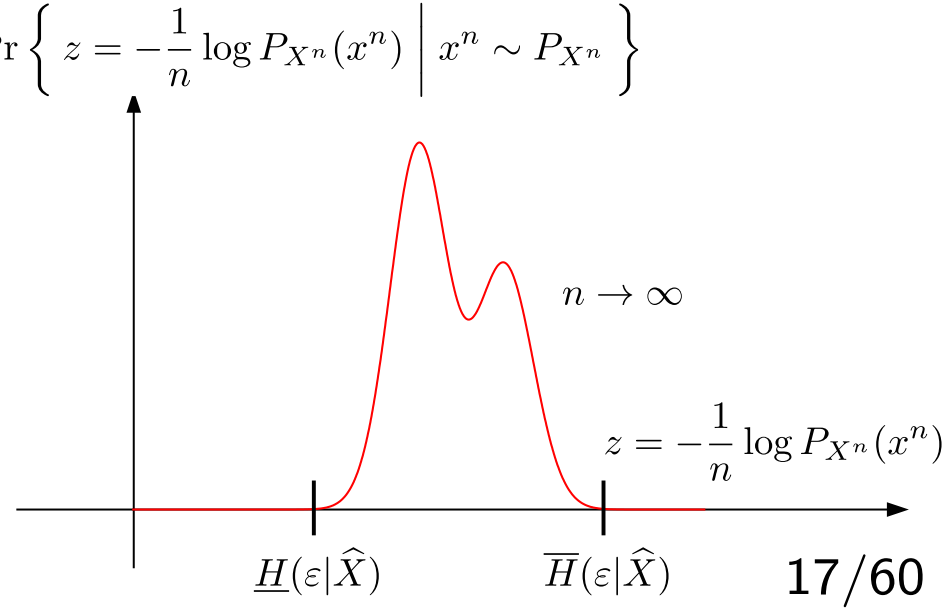
$$\overline{H}(\varepsilon|\widehat{X}) := \inf\left\{ a \,\middle|\, \limsup_{n\to\infty} \Pr\left\{ -\frac{1}{n}\log P_{X^{(n)}}(X^{(n)}) > a \right\} \le \varepsilon \right\}$$

定義（ε-エントロピースペクトル下限）

$$\underline{H}(\varepsilon|\widehat{X}) := \sup\left\{ a \,\middle|\, \liminf_{n\to\infty} \Pr\left\{ -\frac{1}{n}\log P_{X^{(n)}}(X^{(n)}) > a \right\} \ge \varepsilon \right\}$$

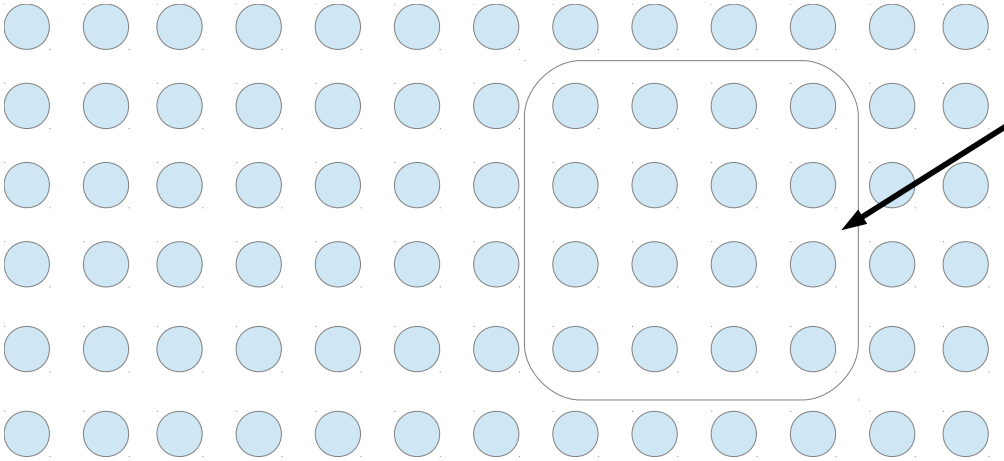
特に

$$\overline{H}(\widehat{X}) := \overline{H}(0|\widehat{X})$$
$$\underline{H}(\widehat{X}) := \underline{H}(1|\widehat{X})$$



情報スペクトル理論の応用

- 情報理論，量子情報理論（nは離散時間を表すことが多い）
  - データ圧縮，乱数生成（確率分布の変換），状態識別，通信路符号化，resolvability符号化，盗聴通信路符号化，レート歪み理論，同定符号
- d-次元スピンチェインの解析
  - 粒子数が無限になる極限（熱力学的極限）において，
    有限窓を増大させて漸近的な極限量を解析する手法を提供
  - 有限窓$A \subseteq \mathbb{Z}^d$が集合の包含関係で有向族（net，数列の一般化）
  - Finitely Correlated State（AKLT状態），Gibbs状態, Fermion Quasi-Free State (07 Hiai-Mosonyi-O, 08 HMO-Fannes)



$$\widehat{X} = \{X^{(A)}\}_{A \subseteq \mathbb{Z}^d}$$

量子系における情報スペクトル的方法

密度行列の列$\widehat{\rho} = \{\rho^{(n)}\}_{n=1}^{\infty}$を任意に考える

- **典型例：** i.i.d. **状態**

$$\begin{aligned}
\rho^{(1)} &= \rho \\
\rho^{(2)} &= \rho \otimes \rho \\
&\vdots \\
\rho^{(n)} &= \rho \otimes \rho \otimes \cdots \otimes \rho
\end{aligned}$$

- **その他の例：**Finitely Correlated State (AKLT**状態**), i.i.d. mixture, . . .

$$\text{(i.i.d. mixture)} \quad \rho^{(n)} = \sum_{i=1}^{m} \pi(i) \rho_i^{\otimes n}$$

量子系におけるエントロピースペクトル

- (コンパクト)エルミート作用素 $A = \sum_i a_i \, |a_i\rangle\langle a_i|$ (固有値分解)

$$A_+ := \sum_{i:a_i>0} a_i \, |a_i\rangle\langle a_i| \quad \text{(正部分)},$$

$$\{A > 0\} := \sum_{i:a_i>0} |a_i\rangle\langle a_i| \quad \text{(正部分への射影)}$$

- 密度行列の列 $\widehat{\rho} = \{\rho^{(n)}\}_{n=1}^{\infty}$ について固有値分解

$$\rho^{(n)} = \sum_{x^n} p^{(n)}(x^n) \, |x^n\rangle\langle x^n|$$

- 固有値の情報スペクトルを考える

$$\mathrm{Tr}\, \rho^{(n)} \{\rho^{(n)} - e^{-na} I^{(n)} > 0\} = \sum_{x^n : \, p^{(n)}(x^n) > e^{-na}} p^{(n)}(x^n)$$

$$= \Pr\left\{ -\frac{1}{n} \log p^{(n)}(x^{(n)}) > a \;\middle|\; x^{(n)} \sim p^{(n)} \right\}$$

$\widehat{\rho} = \{\rho^{(n)}\}_{n=1}^{\infty}$ **と** $\varepsilon \in [0,1]$ **に対して，**

定義（ε**-エントロピースペクトル上限）**

$$\overline{H}(\varepsilon|\widehat{\rho}) := \inf\left\{ a \,\middle|\, \limsup_{n\to\infty} \mathrm{Tr}\, \rho^{(n)}\{\rho^{(n)} - e^{-na} I^{(n)} > 0\} \le \varepsilon \right\}$$

定義（ε**-エントロピースペクトル下限）**

$$\underline{H}(\varepsilon|\widehat{\rho}) := \sup\left\{ a \,\middle|\, \liminf_{n\to\infty} \mathrm{Tr}\, \rho^{(n)}\{\rho^{(n)} - e^{-na} I^{(n)} > 0\} \ge \varepsilon \right\}$$

特に

$$\overline{H}(\widehat{\rho}) := \overline{H}(0|\widehat{\rho}), \quad \underline{H}(\widehat{\rho}) := \underline{H}(1|\widehat{\rho})$$

- i.i.d. **状態** $\rho^{(n)} = \rho^{\otimes n}$ **のとき，**

$$\overline{H}(\widehat{\rho}) = \underline{H}(\widehat{\rho}) = H(\rho) \quad \text{von Neumann}\,\textbf{エントロピー}$$

2　応用：純粋状態LOCC変換可能性

- **エンタングルメント蒸留，希釈の情報スペクトル理論：**
  Hayashi (2006), Bowen-Datta (2008)
- **二体系の純粋状態列** $\widehat{\psi}^{AB}=\{\psi_n^{AB}\}_{n=1}^{\infty}$, $\widehat{\varphi}^{AB}=\{\varphi_n^{AB}\}_{n=1}^{\infty}$ **について**
- LOCC**変換の列** $\widehat{\mathcal{L}}=\{\mathcal{L}_n\}_{n=1}^{\infty}$ **が存在して**

$$\lim_{n\to\infty}\|\mathcal{L}_n(\psi_n^{AB})-\varphi_n^{AB}\|_1=0$$

となるとき "漸近的にLOCC変換可能" であるという

定理 (Jiao-Wakakuwa-O, 2018)　**蒸留，希釈の理論を一般化**

- $\underline{H}(\widehat{\psi}^A)>\overline{H}(\widehat{\varphi}^A)$ **ならば "漸近的にLOCC変換可能"**
- **"漸近的にLOCC変換可能" ならば**

$$\overline{H}(\varepsilon|\widehat{\psi}^A)\geq\overline{H}(\varepsilon|\widehat{\varphi}^A),\ \underline{H}(\varepsilon|\widehat{\psi}^A)\geq\underline{H}(\varepsilon|\widehat{\varphi}^A)\quad(0\leq\forall\varepsilon\leq 1)$$

ただし， $\psi_n^A=\mathrm{Tr}_B\,\psi_n^{AB}$, $\varphi_n^A=\mathrm{Tr}_B\,\varphi_n^{AB}$

3　ダイバージェンス・スペクトル

ダイバージェンス（相対エントロピー, Kullback-Leibler**情報量）**

$$D(p||q) = \sum_x p(x)\{\log p(x) - \log q(x)\} = \sum_x p(x) \log \frac{p(x)}{q(x)}$$

エントロピーや相互情報量の元になる基本的な情報量

- $q(x) = 1(x) = 1$ **(定数関数) とおくとエントロピー**

$$D(p||1) = \sum_x p(x) \log p(x) = -H(p), \text{ i.e., } H(P) = -D(P||1)$$

- **相互情報量：**
  同時分布 $P_{XY}(x, y)$ **と周辺分布の積** $P_X(x)P_Y(y)$ **のダイバージェンス**

$$I(X;Y) = \sum_x \sum_y P_{XY}(x, y) \log \frac{P_{XY}(x, y)}{P_X(x)P_Y(y)}$$

情報スペクトルでもダイバージェンススペクトルを基本に考えるのが良い

ダイバージェンスと大数の法則

- **ダイバージェンス**

$$D(p||q) := \sum_x p(x) \log \frac{p(x)}{q(x)}$$

- p**と**q**の**i.i.d. **拡張**

$$p^n(x^n) = p(x_1)p(x_2)\dots p(x_n), \quad q^n(x^n) = q(x_1)q(x_2)\dots q(x_n)$$

where $x^n = (x_1, x_2, \dots, x_n)$

- $x^n = (x_1, x_2, \dots, x_n)$**が**$p^n(x^n)$**に従うとき，大数の法則より**

$$\frac{1}{n} \log \frac{p^n(x^n)}{q^n(x^n)} = \frac{1}{n} \sum_{i=1}^n \log \frac{p(x_i)}{q(x_i)} \xrightarrow{n\to\infty} E_p \left[\log \frac{p(x)}{q(x)} \right] = D(p||q)$$

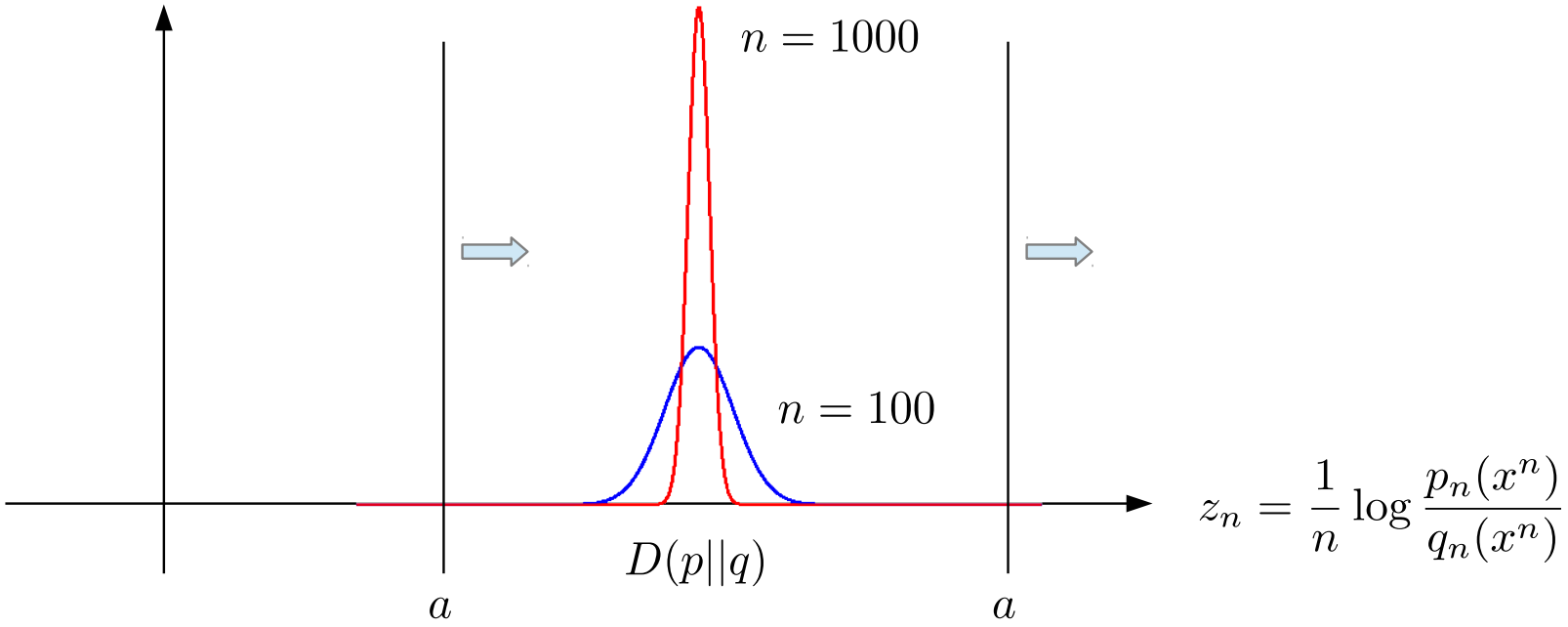
ダイバージェンスに確率収束する

i.i.d. における情報量スペクトル的な見方

大数の法則より

$$\Pr\left\{ \frac{1}{n}\log\frac{p^n(x^n)}{q^n(x^n)} > a \right\} \xrightarrow{n\to\infty} \begin{cases} 1 & a < D(p||q) \\ 0 & a > D(p||q) \end{cases}$$

$$\Pr\left\{ z_n = \tfrac{1}{n}\log\tfrac{p_n(x^n)}{q_n(x^n)} \;\middle|\; x^n \sim p^n(x^n) \right\}$$



エントロピーのときと，不等式の向きが逆

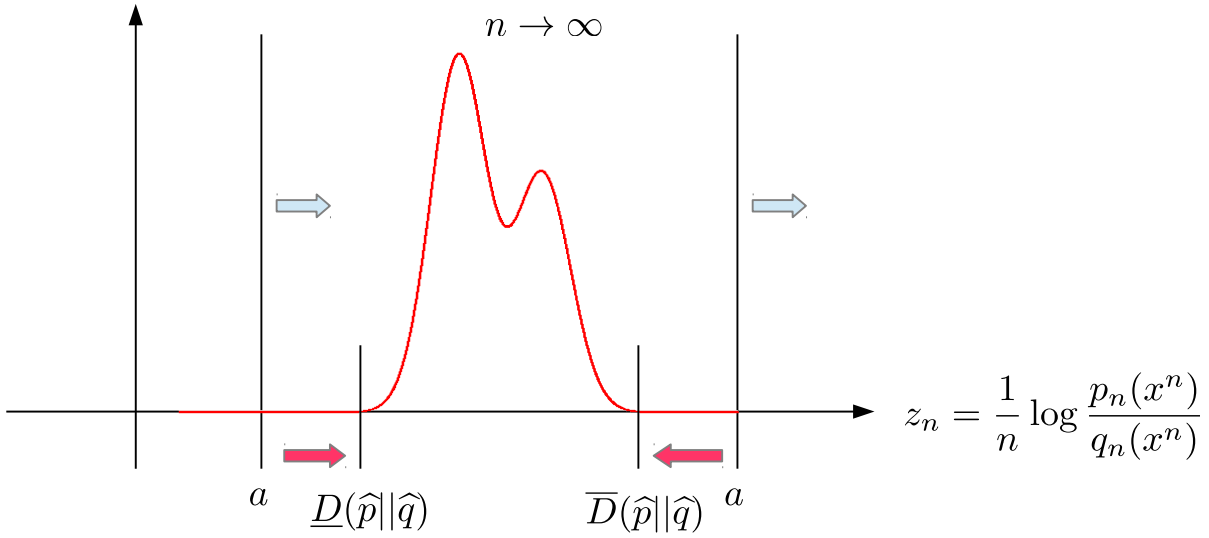
ダイバージェンス・スペクトル・レート下限と上限

確率分布の列$\widehat{p} = \{p_n\}_{n=1}^{\infty}$, $\widehat{q} = \{q_n\}_{n=1}^{\infty}$ **について**[*1]

$$\underline{D}(\widehat{p}||\widehat{q}) := \sup\left\{ a \;\middle|\; \lim_{n\to\infty} \Pr\left\{ \frac{1}{n}\log\frac{p_n(x^n)}{q_n(x^n)} > a \;\middle|\; x^n \sim p_n \right\} = 1 \right\}$$

$$\overline{D}(\widehat{p}||\widehat{q}) := \inf\left\{ a \;\middle|\; \lim_{n\to\infty} \Pr\left\{ \frac{1}{n}\log\frac{p_n(x^n)}{q_n(x^n)} > a \;\middle|\; x^n \sim p_n \right\} = 0 \right\}$$

$\Pr\left\{ z_n = \frac{1}{n}\log\frac{p_n(x^n)}{q_n(x^n)} \;\middle|\; x^n \sim p^n(x^n) \right\}$



[*1] $p^{(n)}$, $q^{(n)}$ **と書くのが面倒になってくるので**p_n, q_n **と書く**

量子系への準備（非可換性により $\log$ はダメ）

$$\frac{1}{n}\log\frac{p_n(x^n)}{q_n(x^n)} > a \Longleftrightarrow p_n(x^n) - e^{na}q_n(x^n) > 0$$

- $x^n \sim p_n(x^n)$ **のとき，**

$$\Pr\left\{\frac{1}{n}\log\frac{p_n(x^n)}{q_n(x^n)} > a\right\} = \Pr\{\, p_n(x^n) - e^{na}q_n(x^n) > 0\,\}$$

$$= \sum_{x^n} p_n(x^n)\mathbf{1}\{\, p_n(x^n) - e^{na}q_n(x^n) > 0\,\}$$

ただし， $\mathbf{1}\{\dots\}$ **は定義関数：**

$$\mathbf{1}\{\, p_n(x^n) - e^{na}q_n(x^n) > 0\,\} = \begin{cases} 1 & \text{if} \quad p_n(x^n) - e^{na}q_n(x^n) > 0 \\ 0 & \text{otherwise} \end{cases}$$

量子ダイバージェンス・スペクトル・レート

- （復習）エルミート作用素 $A = \sum_i a_i \, |a_i\rangle\langle a_i|$ (固有値分解) に対して

$$A_+ := \sum_{i:a_i>0} a_i \, |a_i\rangle\langle a_i| \quad \text{(正部分)}$$

$$\{A > 0\} := \sum_{i:a_i>0} |a_i\rangle\langle a_i| \quad \text{(正部分への射影子)}$$

定義 密度行列の列 $\widehat{\rho} = \{\rho_n\}_{n=1}^{\infty}$, $\widehat{\sigma} = \{\sigma_n\}_{n=1}^{\infty}$ に対して，

$$\underline{D}(\widehat{\rho}||\widehat{\sigma}) := \sup\left\{ a \,\middle|\, \lim_{n\to\infty} \mathrm{Tr}\, \rho_n \{ \rho_n - e^{na}\sigma_n > 0 \} = 1 \right\}$$

$$\overline{D}(\widehat{\rho}||\widehat{\sigma}) := \inf\left\{ a \,\middle|\, \lim_{n\to\infty} \mathrm{Tr}\, \rho_n \{ \rho_n - e^{na}\sigma_n > 0 \} = 0 \right\}$$

- $\mathrm{Tr}\, \rho_n \{ \rho_n - e^{na}\sigma_n > 0 \}$ は $\Pr\left\{ \frac{1}{n} \log \frac{p_n(x^n)}{q_n(x^n)} > a \right\}$ に相当

  量子Neyman-Pearson testの第一種成功確率

状態識別タスク（量子仮説検定）における一般公式

- POVM $\{T_n, I_n - T_n\}$ $(0 \le T_n \le I_n)$**による状態の識別**

$$\alpha_n(T_n) = \mathrm{Tr}\, \rho_n (I_n - T_n) \qquad \text{(1st kind error)}$$
$$\beta_n(T_n) = \mathrm{Tr}\, \rho_n T_n \qquad \text{(2nd kind error)}$$

- **最適指数** $R(\widehat{\rho}||\widehat{\sigma})$ **(大きい方が良い)**, $R^*(\widehat{\rho}||\widehat{\sigma})$ **(小さい方が良い)**

$$\beta_n(T_n) \simeq e^{-nR(\widehat{\rho}||\widehat{\sigma})} \quad \text{s.t.} \quad \lim_{n\to\infty} \alpha_n(T_n) = 0$$

$$\beta_n(T_n) \simeq e^{-nR^*(\widehat{\rho}||\widehat{\sigma})} \quad \text{s.t.} \quad \lim_{n\to\infty} \alpha_n(T_n) = 1$$

定理 (Nagaoka 98, Nagaoka-Hayashi 07 **の一般公式**)

$$R(\widehat{\rho}||\widehat{\sigma}) = \underline{D}(\widehat{\rho}||\widehat{\sigma}), \quad R^*(\widehat{\rho}||\widehat{\sigma}) = \overline{D}(\widehat{\rho}||\widehat{\sigma})$$

量子 i.i.d. 状態（$\rho_n = \rho^{\otimes n}$, $\sigma_n = \sigma^{\otimes n}$）

- **ダイバージェンス・スペクトルの定義より**

$$a < \underline{D}(\widehat{\rho}||\widehat{\sigma}) \Longleftrightarrow \lim_{n\to\infty} \mathrm{Tr}\,\rho_n \{\, \rho_n - e^{na}\sigma_n > 0 \,\} = 1$$

$$a > \overline{D}(\widehat{\rho}||\widehat{\sigma}) \Longleftrightarrow \lim_{n\to\infty} \mathrm{Tr}\,\rho_n \{\, \rho_n - e^{na}\sigma_n > 0 \,\} = 0$$

- i.i.d. **状態においてはダイバージェンス・スペクトル上限・下限は**

$$\underline{D}(\widehat{\rho}||\widehat{\sigma}) = \overline{D}(\widehat{\rho}||\widehat{\sigma}) = D(\rho||\sigma) := \mathrm{Tr}\,\rho(\log\rho - \log\sigma)$$

量子相対エントロピーと一致する．

$$\lim_{n\to\infty} \mathrm{Tr}\,\rho_n \{\, \rho_n - e^{na}\sigma_n > 0 \,\} = \begin{cases} 1 & \text{if} \quad a < D(\rho||\sigma) \text{ (Hiai-Petz 91)} \\ 0 & \text{if} \quad a > D(\rho||\sigma) \text{ (O-Nagaoka 00)} \end{cases}$$

ダイバージェンス・スペクトルの別表現

Bowen-Datta 06 の別表現

$$\underline{C}(\widehat{\rho}||\widehat{\sigma}) := \sup\left\{ a \,\middle|\, \lim_{n\to\infty} \mathrm{Tr}(\rho_n - e^{na}\sigma_n)_+ = 1 \right\}$$

$$\overline{C}(\widehat{\rho}||\widehat{\sigma}) := \inf\left\{ a \,\middle|\, \lim_{n\to\infty} \mathrm{Tr}(\rho_n - e^{na}\sigma_n)_+ = 0 \right\}$$

- Nagaoka 98, Nagaoka-Hayashi 07 **の定義と一致することが示される**

$$\underline{C}(\widehat{\rho}||\widehat{\sigma}) = \underline{D}(\widehat{\rho}||\widehat{\sigma}), \quad \overline{C}(\widehat{\rho}||\widehat{\sigma}) = \overline{D}(\widehat{\rho}||\widehat{\sigma})$$

- **意味：重み付き** Bayes **エラー（**Bayes **成功確率）の最適値**

$$\min_{0\le T_n\le I_n} \{\mathrm{Tr}\,\rho_n(I_n - T_n) + e^{na}\,\mathrm{Tr}\,\rho_n T_n\} = 1 - \mathrm{Tr}(\rho_n - e^{na}\sigma_n)_+$$

これは以下より導かれる

$$\max_{0\le T_n\le I_n} \mathrm{Tr}(\rho_n - e^{na}\sigma_n)T_n = \mathrm{Tr}(\rho_n - e^{na}\sigma_n)_+$$

Bowen-Datta表現のメリット

- **挙動が素直（$\mathrm{Tr}\, A_+$ の性質），トレードオフを一気に扱える☺**
  - **単調性：** $\mathrm{Tr}\, A_+ \geq \mathrm{Tr}\, \mathcal{F}(A)_+$ (for trace preserving positive maps)
  - order preserving: if $A \leq B$ then $\mathrm{Tr}\, A_+ \leq \mathrm{Tr}\, B_+$
  - subadditivity: $\mathrm{Tr}(A+B)_+ \leq \mathrm{Tr}\, A_+ + \mathrm{Tr}\, B_+$
- i.i.d. $(\rho_n = \rho^{\otimes n},\, \sigma_n = \sigma^{\otimes n})$ **の場合**

$$\lim_{n\to\infty} \mathrm{Tr}(\rho_n - e^{na}\sigma_n)_+ = \begin{cases} 1 & \text{if} \quad a < D(\rho||\sigma) \\ 0 & \text{if} \quad a > D(\rho||\sigma) \end{cases}$$

これは，一般に以下が成り立つことによる

$$\underline{C}(\widehat{\rho}||\widehat{\sigma}) = \underline{D}(\widehat{\rho}||\widehat{\sigma}) = D(\rho||\sigma) = \overline{D}(\widehat{\rho}||\widehat{\sigma}) = \overline{C}(\widehat{\rho}||\widehat{\sigma})$$

- **両方の定義を都合良く使い分ければよい**

4　古典・量子通信路と相互情報量スペクトル

- **古典・量子通信路**(classical-quantum channel, cq channel)

$$W : x \in \mathcal{X} \text{ (**有限集合，古典シグナルの集合**)} \longmapsto W_x \text{ (**密度行列**)}$$

- Holevo**相互情報量**

$$I(P;W) := \sum_x P(x) D(W_x || W_P) \quad \text{where} \quad W_P := \sum_x P(x) W_x$$

- Holevo**相互情報量は拡大状態のダイバージェンス**

$$R = \begin{pmatrix} P(1)\,W_1 & & 0 \\ & \ddots & \\ 0 & & P(N)\,W_N \end{pmatrix}, \quad S = \begin{pmatrix} P(1)\,W_P & & 0 \\ & \ddots & \\ 0 & & P(N)\,W_P \end{pmatrix}$$

$$I(P;W) = D(R||S)$$

cq channelのi.i.d. 拡張（記法の準備）

- cq channel W のi.i.d. 拡大

$$\begin{array}{ccc} x_1 & \to \boxed{W} \to & W_{x_1} \\ x_2 & \to \boxed{W} \to & W_{x_2} \\ \vdots & & \vdots \\ x_n & \to \boxed{W} \to & W_{x_n} \end{array}$$

- **記法：上を次のように書く**

$$x^n := (x_1, x_2, \ldots, x_n) \to \boxed{W^n} \to W_{x^n} := W_{x_1} \otimes W_{x_2} \otimes \cdots \otimes W_{x_n}$$

- **確率** $P(x)$ **の**i.i.d. **拡張**

$$P^n(x^n) = P(x_1)P(x_2)\ldots P(x_n)$$

一般化cq channelと相互情報量スペクトル

- cq channelの列 $\widehat{W} = \{W^n\}_{n=1}^{\infty}$ を考える
- 典型例はi.i.d. 拡張

定義：相互情報量スペクトル　拡大状態列 $\widehat{R} = \{R_n\}_{n=1}^{\infty}$,
$\widehat{S} = \{S_n\}_{n=1}^{\infty}$ のダイバージェンススペクトル として定義される

$$R_n = \bigoplus_{x^n} P^n(x^n) W_{x^n}^n,\ S_n = \bigoplus_{x^n} P^n(x^n) W_{P^n}^n \left(W_{P^n}^n := \sum_{x^n} P^n(x^n) W_{x^n}^n \right)$$

$$\underline{I}(\widehat{P}, \widehat{W}) := \underline{D}(\widehat{R}, \widehat{S}) = \sup \left\{ a \,\middle|\, \lim_{n\to\infty} \operatorname{Tr} R_n \{ R_n - e^{na} S_n > 0 \} = 1 \right\}$$

$$\overline{I}(\widehat{P}, \widehat{W}) := \overline{D}(\widehat{R}, \widehat{S}) = \inf \left\{ a \,\middle|\, \lim_{n\to\infty} \operatorname{Tr} R_n \{ R_n - e^{na} S_n > 0 \} = 0 \right\}$$

拡大状態はブロックごとに計算できるので，

$$\operatorname{Tr} R_n \{ R_n - e^{na} S_n > 0 \} = \sum_{x^n} P^n(x^n) \operatorname{Tr} W_{x^n}^n \{ W_{x^n}^n - e^{na} W_{P^n}^n > 0 \}$$

- **相互情報量スペクトルの定義より**

$$a < \underline{I}(\widehat{P}, \widehat{W}) \Longleftrightarrow \lim_{n\to\infty} \mathrm{Tr}\, R_n \{ R_n - e^{na} S_n > 0 \} = 1$$

$$a > \overline{I}(\widehat{P}, \widehat{W}) \Longleftrightarrow \lim_{n\to\infty} \mathrm{Tr}\, R_n \{ R_n - e^{na} S_n > 0 \} = 0$$

- i.i.d. **拡張通信路の場合は**Holevo**相互情報量に一致する**

$$\underline{I}(\widehat{P}, \widehat{W}) = \overline{I}(\widehat{P}, \widehat{W}) = D(R||S) = I(P, W)$$

これは，以下の事実による

$$R_n = R^{\otimes n} = (\oplus_x P(x) W_x)^{\otimes n}, \quad S_n = S^{\otimes n} = (\oplus_x P(x) W_P)^{\otimes n}$$

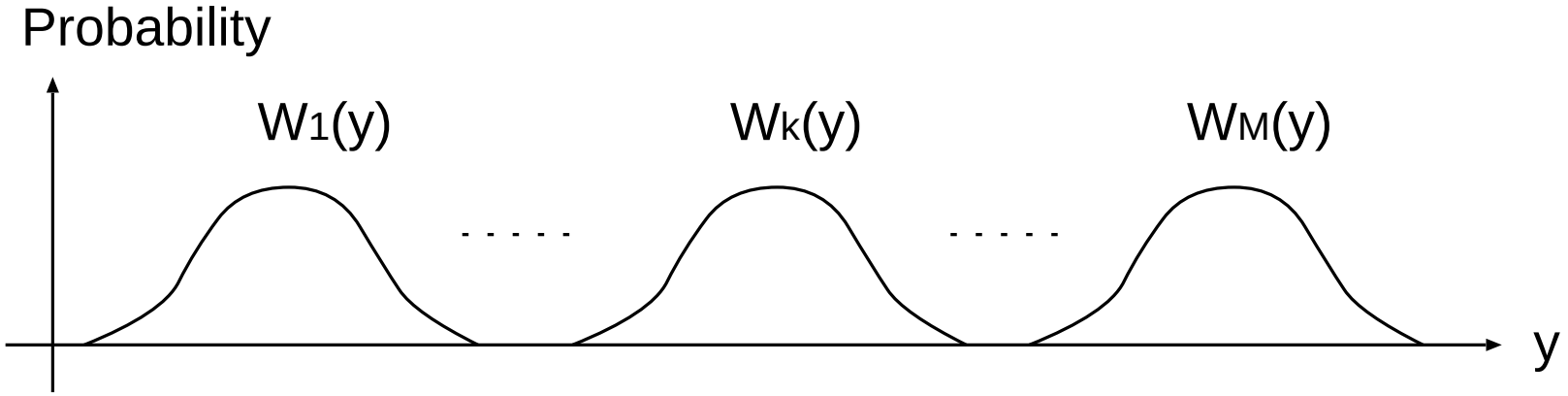
通信路符号化（メッセージ伝送）：古典系

古典通信路$W_1(y), W_2(y), \ldots, W_M(y)$の出力$y$から入力$k$を当てる問題

$$k \in \{1, 2, \ldots, M\} \longrightarrow \boxed{W(y|k) = W_k(y)} \longrightarrow y$$

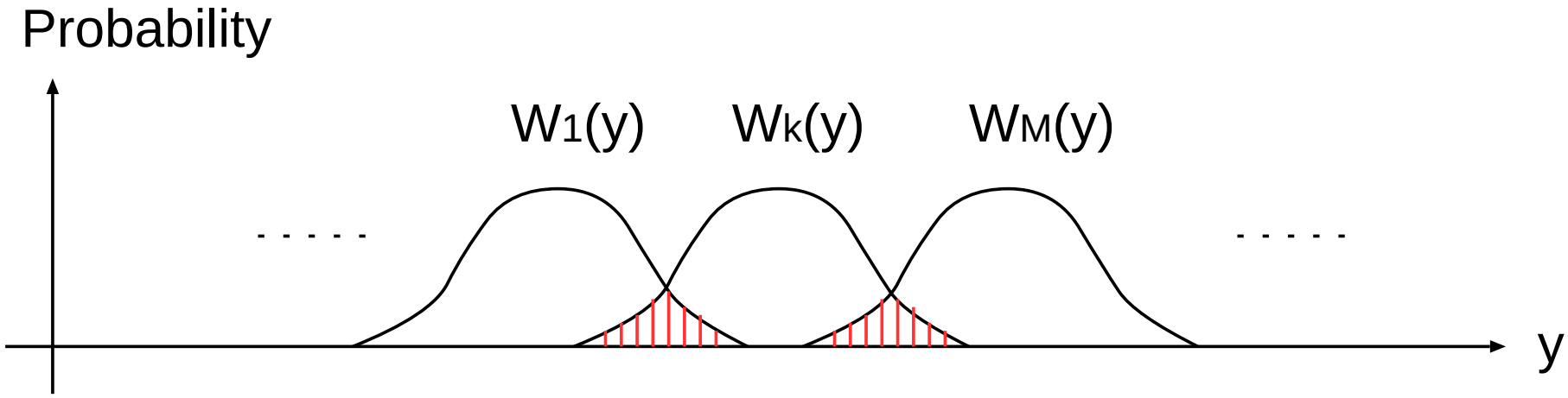
候補が3つ以上の仮説検定問題（候補が2つだけ：単純仮説検定）

- trivial zero error case



  $W_k(y)$とその他$W_l(y)$ $(l \neq k)$のオーバラップがない
- **yを観測することでkをゼロエラーで識別できる**

An Overlap Measure：単純仮説検定への帰着

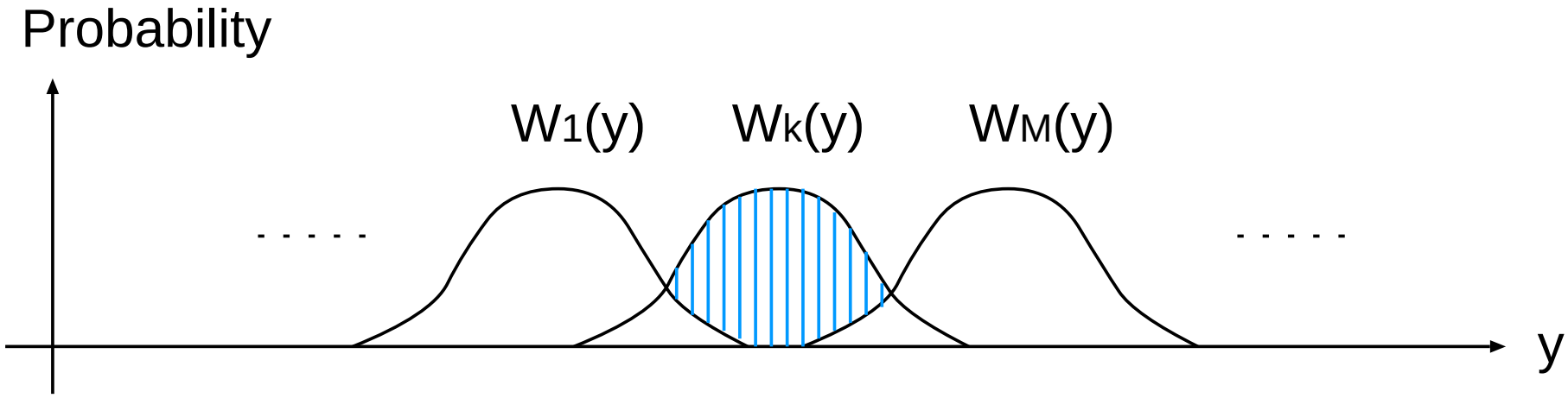


- W_k **とそれ以外** $W_l\ (l \neq k)$ **の** Bayes **重みエラーを考える**
- $T \subset \mathcal{Y}$ **を** k **の受容域** (acceptance region) **とすると**

$$P_e(k) := \min_{T \subset \mathcal{Y}} \left\{ \sum_{y \in T^c} W_k(y) + \sum_{l \neq k} \sum_{y \in T} W_l(y) \right\} \quad \text{1st kind} + \text{2nd kind}$$

$$= \min_{T \subset \mathcal{Y}} \left\{ 1 - \sum_{y \in T} \left(W_k(y) - \sum_{l \neq k} W_l(y) \right) \right\}$$

Bayes成功確率



- $P_s(k) := 1 - P_e(k)$： W_k **とそれ以外を識別するときの成功確率**

$$P_s(k) = 1 - P_e(k) = \max_{T \subset \mathcal{Y}} \sum_{y \in T} \left(W_k(y) - \sum_{l \neq k} W_l(y) \right)$$

$$= \sum_{y} \left(W_k(y) - \sum_{l \neq k} W_l(y) \right)_+$$

ただし， $(F(x))_+ = \max\{0, F(x)\}$

単純仮説検定への帰着（量子系）

古典・量子通信路$W_1, W_2, \ldots, W_M$（密度行列）が与えられているとき，出力量子状態をPOVM測定することでkを当てる問題

$$k \in \{1, 2, \ldots, M\} \longrightarrow \boxed{W_k} \longrightarrow \boxed{\text{POVM}} \longrightarrow \widehat{k}$$

古典系と同様にoverlap measure**を考える**

- $P_s(k) = 1 - P_e(k)$**：**W_k**とそれ以外を識別する**Bayes**成功確率を考える**

$$P_s(k) = \max_{0 \le T \le I} \mathrm{Tr}\left(W_k - \sum_{l \neq k} W_l\right) T = \mathrm{Tr}\left(W_k - \sum_{l \neq k} W_l\right)_+$$

漸近論：overlapの情報量スペクトル的解析

問題設定 cq channel列$\{W_{x^n}^n\}_{x^n\in\mathcal{X}^n}$に対して，$M_n = e^{nR}$個の入力を選んでoverlap $\max_k P_e(k) = \max_k\{1 - P_s(k)\}$がゼロになるようにする

$$\varphi_n : k \in \{1, 2, \ldots, M_n\} \longmapsto x_k^n = \varphi_n(k) \longmapsto W_{\varphi_n(k)}^n$$

定理 相互情報量スペクトルレートはsharp limitである

(1) (achievability) Choosing $\varphi_n(k)$ $(k = 1, 2, \ldots, M_n)$ randomly subject to $P^n(x^n)$, we have

$$\lim_{n\to\infty} E\left[\mathrm{Tr}\left(W_{\varphi_n(k)}^n - \sum_{l\neq k} W_{\varphi_n(l)}^n\right)_+\right] = 1 \quad \text{if } R < \underline{I}(\widehat{P}, \widehat{W})$$

(2) (strong converse) For any $\{\varphi_n\}_{n=1}^\infty$, we have

$$\lim_{n\to\infty} \max_{k\in[1,M_n]} \mathrm{Tr}\left(W_{\varphi_n(k)}^n - \sum_{l\neq k} W_{\varphi_n(l)}^n\right)_+ = 0 \quad \text{if } R > \overline{I}(\widehat{P}, \widehat{W}),$$

proof of (1) achievability

Given a cq channel $x \in \mathcal{X} \mapsto W_x$ and a map $k \in \{1, 2, \ldots, M\} \mapsto \varphi(k) \in \mathcal{X}$, for any $P(x)$, $a \in \mathbb{R}$, and $k \in \{1, 2, \ldots, M\}$, it obviously holds that

$$\mathrm{Tr}\left(W_{\varphi(k)} - \sum_{l \neq k} W_{\varphi(l)}\right)_+ \geq \mathrm{Tr}\left(W_{\varphi(k)} - \sum_{l \neq k} W_{\varphi(l)}\right)\left\{W_{\varphi(k)} - e^a W_P > 0\right\}.$$

Taking expectation, we have

$$E\left[\mathrm{Tr}\left(W^n_{\varphi_n(k)} - \sum_{l \neq k} W^n_{\varphi_n(l)}\right)_+\right]$$

$$\geq E\left[\mathrm{Tr}\, W_{\varphi_n(k)}\left\{W_{\varphi_n(k)} - e^a W_P > 0\right\}\right] - M_n \cdot E\left[\mathrm{Tr}\, W_P\left\{W_{\varphi_n(k)} - e^a W_P > 0\right\}\right]$$

$$\geq (1 - e^{nR} e^{-na}) E_{P^n}\left[\mathrm{Tr}\, W^n_{x^n}\left\{W^n_{x^n} - e^a W^n_{P^n} > 0\right\}\right].$$

If $R < \underline{I}(\widehat{P}, \widehat{W})$, then there exists $R < a < \underline{I}(\widehat{P}, \widehat{W})$, and the definition of $\underline{I}(\widehat{P}, \widehat{W})$ assures RHS $\to 1$ $(n \to \infty)$

通信路符号化の誤り確率との関係

cq channel $W : x \mapsto W_x$ **と符号器**$\varphi : k \mapsto \varphi(k) \in \mathcal{X}$

$$k \in \{1, 2, \ldots, M\} \to \varphi(k) \to \boxed{W_{\varphi(k)}} \to \boxed{\text{POVM } Y = \{Y_l\}_{l=0}^{M}} \to \widehat{k} = l,$$

復号器Y (POVM)**が与えられると，誤り確率**$\mathrm{Pe}(k; \varphi, Y)$**が自然に定まる**

定理 For any map $\varphi : \{1, 2, \ldots, M\} \to \mathcal{X}$
there exists POVM $Y = \{Y_k\}_{k=0}^{M}$ such that for any k and T $(0 \le T \le I)$

$$\mathrm{Pe}(k; \varphi, Y) \le \mathrm{Tr}\, W_{\varphi(k)}(I - T) + \mathrm{Tr}\left(\sum_{l \ne k} W_{\varphi(l)}\right) T$$

holds. Taking the minimum w.r.t. $0 \le T \le I$, we have

$$Pe(k; \varphi, Y) \le 1 - \mathrm{Tr}\left(W_k - \sum_{l \ne k} W_l\right)_+$$

sketch proof

- POVM $Y = \{Y_k\}_{k=0}^{M}$ is constructed by Beigi-Gohari (2014) [*2]

$$Y_k = \begin{cases} T_\varphi^{-1/2} W_{\varphi(k)} T_\varphi^{-1/2} & (k = 1, 2, \dots, M) \\ I - \mathrm{suppot}(T_\varphi) & (k = 0) \end{cases}, \quad T_\varphi = \sum_{l=1}^{M} W_{\varphi(l)}$$

  T_φ^{-1} is the generalized inverse satisfying $T_\varphi^{-1/2} T_\varphi T_\varphi^{-1/2} = s(T_\varphi)$

- monotonicity of the Sandwiched Renyi divergence for $\alpha = 2$
  (Collision relative entropy)

$$D_\alpha^*(A||B) = \frac{1}{\alpha - 1} \log Q_\alpha^*(A||B) - \frac{1}{\alpha - 1} \log \mathrm{Tr}\, A$$

$$Q_\alpha^*(A||B) = \mathrm{Tr} \left(B^{\frac{1-\alpha}{2\alpha}} A B^{\frac{1-\alpha}{2\alpha}} \right)^\alpha$$

$$Q_2^*(A||B) = \mathrm{Tr} \left(B^{-1/4} A B^{-1/4} \right)^2 = \mathrm{Tr} \left(A \cdot B^{-1/2} A B^{-1/2} \right)$$

[*2] $*$ S. Beigi and A. Gohari, "Quantum achievability proof via collision relative entropy," *IEEE Trans. Inform. Theory*, vol. 60, pp. 7980–7986, 2014.

Let

$$W_k = W_{\varphi(k)}, \quad V_k = \sum_{l \neq k} W_{\varphi(l)}$$

For simplicity, we neglect the support treatment. Then we have

$$Y_k = (W_k + V_k)^{-1/2} W_k (W_k + V_k)^{-1/2}$$

and

$$\begin{aligned}
Pe(k; \varphi, Y) &= 1 - \operatorname{Tr} W_k (W_k + V_k)^{-1/2} W_k (W_k + V_k)^{-1/2} \\
&= \operatorname{Tr} W_k (W_k + V_k)^{-1/2} V_k (W_k + V_k)^{-1/2} \\
&= \operatorname{Tr} W_k - Q_2^*(W_k || W_k + V_k) \\
&\leq \frac{\operatorname{Tr} W_k T \times \operatorname{Tr} V_k T}{\operatorname{Tr} W_k T + \operatorname{Tr} V_k T} + \frac{\operatorname{Tr} W_k (I - T) \times \operatorname{Tr} V_k (I - T)}{\operatorname{Tr} W_k (I - T) + \operatorname{Tr} V_k (I - T)} \quad \text{monotonicity} \\
&\leq \operatorname{Tr} W_k T + \operatorname{Tr} V_k (I - T)
\end{aligned}$$

Hayashi-Nagaokaの不等式

Hayashi-Nagaoka (2003)　**量子情報理論に飛躍的進歩をもたらした**

Given c-q channel $W : x \mapsto W_x$, for any $P(x)$ there exists (φ, Y) such that

$$Pe(\varphi, Y) \le 2 \sum_{x \in \mathcal{X}} P(x) \operatorname{Tr} W_x \{ W_x - e^a W_P \le 0 \} + 4(M-1) \operatorname{Tr} W_P \{ W_x - e^a W_P > 0 \}$$

for any $a \in \mathbb{R}$, where $W_P = \sum_{x \in \mathcal{X}} P(x) W_x$.

- **別証明と係数の改善が得られた**

$$Pe(\varphi, Y) \le \sum_{x \in \mathcal{X}} P(x) \operatorname{Tr} W_x \{ W_x - e^a W_P \le 0 \} + (M-1) \operatorname{Tr} W_P \{ W_x - e^a W_P > 0 \}$$

5 cq wiretap channel

- Given a quantum channel $\mathcal{E} : \mathcal{S}(\mathcal{H}_A) \to \mathcal{S}(\mathcal{H}_B)$, there exists an environment system $\mathcal{H}_E$ and an isometry $U : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$, such that

$$\mathcal{E}(\rho) = \mathrm{Tr}_E[U\rho U] \quad \text{(Stinespring dilation)}$$

- complementary channel $\mathcal{F} : \mathcal{S}(\mathcal{H}_A) \to \mathcal{S}(\mathcal{H}_E)$ is defined by

$$\mathcal{F}(\rho) = \mathrm{Tr}_B[U\rho U]$$

- U can be extended to unitary U_{ABE} on $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E$

$$\text{Alice } \rho \to \boxed{U_{ABE}} \to \mathcal{E}(\rho) \in \mathcal{S}(\mathcal{H}_B) \text{ Bob}$$

$$\searrow \mathcal{F}(\rho) \in \mathcal{S}(\mathcal{H}_E) \text{ Eve}$$

- c-q wiretap channel is a pair of c-q channels (W, V) such that

$$\text{Alice } x \to \rho_x \to \boxed{W} \to W_x := \mathcal{E}(\rho_x) \in \mathcal{S}(\mathcal{H}_B) \text{ Bob}$$

$$\searrow \boxed{V} \to V_x := \mathcal{F}(\rho_x) \in \mathcal{S}(\mathcal{H}_E) \text{ Eve}$$

that is a model of eavesdropper.

secure transmission of message over wiretap channel

- wiretap channel coding over the i.i.d. extension

  Alice (message) $k \to \boxed{Q_{n,k}(x^n)} \to x^n \to \boxed{W^n} \to \sum_{x^n} Q_{n,k}(x^n) W_{x^n}$

  Bob

  $k \in \{1, 2, \ldots, M_n\}$ $\qquad\qquad$ $\searrow \boxed{V^n} \to \sum_{x^n} Q_{n,k}(x^n) V_{x^n}$ Eve

- $Q_n = \{Q_{n,k}\}_{k=1}^{M_n}$ is a stochastic encoder:
  Alice will input signal $x^n = (x_1, \ldots, x_n)$ randomly depending on $Q_{n,k}(x^n)$
- density operator at Bob and Eve are, respectively,

$$Q_{n,k}W := \sum_{x^n} Q_{n,k}(x^n) W_{x^n}, \quad Q_{n,k}V := \sum_{x^n} Q_{n,k}(x^n) V_{x^n}$$

- Bob's decoder is a POVM $Y_n = \{Y_{n,k}\}_{k=1}^{M_n}$ on $\mathcal{S}(\mathcal{H}_B^{\otimes n})$
- pair of an encoder and a decoder (Q_n, Y_n) is called a code
- error probability

$$\mathrm{Pe}(Q_n, Y_n) := \frac{1}{M_n} \sum_{k=1}^{M_n} \{1 - \mathrm{Tr}(Q_{n,k}W) Y_{n,k}\}$$

the aim of wiretap channel coding

- error probability should go to zero asymptotically

$$\mathrm{Pe}(Q_n, Y_n) \xrightarrow{n\to\infty} 0$$

- Eve should not obtain any information of the message

$$d_e(Q_n) := \frac{1}{M_n(M_n-1)} \sum_{k,l:\, k\neq l} \|Q_{n,k}V_n - Q_{n,l}V_n\|_1 \xrightarrow{n\to\infty} 0$$

  where $\|\cdot\|_1$ is the trace norm. (Eve cannot distinguish messages)

- encoding rate $\dfrac{\log M_n}{n}$ should be large as $n \to \infty$
- (remark: another important measure)

$$I_e(Q_n) := \sum_{k=1}^{M_n} \frac{1}{M_n} D(Q_{n,k}V_n \| \overline{Q}_n V_n), \quad \text{where} \quad \overline{Q}_n(x^n) = \frac{1}{M_n}\sum_{k=1}^{M_n} Q_{n,k}(x^n)$$

  that is, the Holevo information between the message and Eve's state

private capacity

private capacity, secrecy capacity

$C_S(W,V) :=$ supremum of the rate $\lim_{n\to\infty} \frac{\log M_n}{n}$ such that

$$\text{Bob's error } \mathrm{Pe}(Q_n, Y_n) \xrightarrow{(n\to\infty)} 0, \quad \text{Eve's information } d_e(Q_n) \xrightarrow{(n\to\infty)} 0$$

Mathematically speaking,

$$C_S(W,V) := \sup\Big\{ R \,\Big|\, \exists\{(Q_n, Y_n)\}_{n=1}^{\infty} \quad \text{such that} \quad \liminf_{n\to\infty} \frac{\log M_n}{n} \ge R, \lim_{n\to\infty} \mathrm{Pe}(Q_n, Y_n) = 0, \lim_{n\to\infty} d_e(Q_n, Y_n) = 0 \Big\}$$

c-q wiretap channel coding theorem

- relative entropy: For density operators ρ and σ,

$$D(\rho||\sigma) := \mathrm{Tr}\,\rho(\log\rho - \log\sigma)$$

- Holevo information: For c-q channel $W : x \in \mathcal{X} \to W_x \in \mathcal{S}(\mathcal{H})$, and probability $P(x)$,

$$I(P;W) := \sum_x P(x) D(W_x||W_P) \quad \text{where} \quad W_P := \sum_x P(x) W_x$$

Devetak 2005, Hayashi 2006 textbook

$$C_S(W,V) = \lim_{n\to\infty} \frac{1}{n} \sup_{P_n, T_n} \left\{ I(P_n; T_n W^n) - I(P_n; T_n V^n) \right\},$$

where T_n is taken over the set of conditional probabilities from a set $\mathcal{T}^n$ to $\mathcal{X}^n$ and P_n is taken over the set of probability functions on $\mathcal{T}_n$.

- classical wiretap channel coding is developed by Wyner (1975) and Csiszár-Körner (1978)

What is channel resolvability?

- In wiretap channel coding,
  Alice should control output statistics of wiretap channel V^n

$$\Downarrow$$

- channel resolvability coding is a method to control the output statistics by stochastic encoder at the input side
- channel resolvability for a channel V is the required amount of randomness to make it a completely noisy channel asymptotically
- that is, channel with high resolvability requires much amount of randomness

Historical Remark

- resolvability coding is proposed by Han-Verdú (1993) in the classical case, concerning another problem (so called identification code)
- As resolvability coding has been used implicitly in wiretap channel coding both in the classical and the quantum case, Hayashi (textbook, 2006) applied it to wiretap channel coding explicitly both in the classical and the quantum case

c-q channel resolvability coding

- Consider the size L_n of uniform random number to approximate the output statistics

$$x^n \sim P^n(x^n) \to \boxed{V^n} \to V_{P^n} := \sum_{x^n} P^n(x^n) V_{x^n}$$

$$x^n \underset{\sim}{\text{uniform}} \{x_1^n, x_2^n, \ldots, x_{L_n}^n\} \to \boxed{V^n} \to V_{\Phi_n} := \frac{1}{L_n} \sum_{l=1}^{L_n} V_{x_l^n}$$

- $\Phi_n := \{x_1^n, x_2^n, \ldots, x_{L_n}^n\} \subset \mathcal{X}^n$ is called a code
- approximation measure (trace distance)

$$d(V_{\Phi_n}, V_{P^n}) = \left\| \frac{1}{L_n} \sum_{l=1}^{L_n} V_{x_l^n} - V_{P^n} \right\|_1 \overset{n\to\infty}{\longrightarrow} 0$$

- encoding rate $\frac{1}{n} \log L_n$ should be as small as possible

c-q channel resolvability theorem

channel resolvability

$$C_R(P,V) := \inf\Big\{ R \;\Big|\; \exists \{\Phi_n\}_{n=1}^{\infty},\ \lim_{n\to\infty} d(V_{\Phi_n}, V_{P^n}) = 0,\ \limsup_{n\to\infty} \frac{1}{n} \log L_n \le R \Big\}$$

Hayashi 2005 [implicitly: Winter 2002, Devetak 2005]

$$C_R(P,V) \le I(P,V) \quad \text{(Holevo information)}$$

- It means, for $R > I(P;V)$, there exists a code $\Phi_n = \{x_1^n, x_2^n, \ldots, x_{L_n}^n\}$ with size $L_n = e^{nR}$ such that

$$\left\| \frac{1}{L_n} \sum_{l=1}^{L_n} V_{x_l^n} - V_{P^n} \right\|_1 \overset{n\to\infty}{\longrightarrow} 0$$

c-q channel resolvability coding for general channels

- channel resolvability is defined in the same way as the i.i.d. case
- the size L_n of uniform random number to approximate the output statistics

$$x^n \sim P^n(x^n) \to \boxed{V^n} \to V_{P^n} := \sum_{x^n} P^n(x^n) V_{x^n}$$

$$x^n \underset{\sim}{\text{uniform}} \{x_1^n, x_2^n, \ldots, x_{L_n}^n\} \to \boxed{V^n} \to V_{\Phi_n} := \frac{1}{L_n} \sum_{l=1}^{L_n} V_{x_l^n}$$

- approximation measure (trace distance)

$$d(V_{\Phi_n}, V_{P^n}) = \left\| \frac{1}{L_n} \sum_{l=1}^{L_n} V_{x_l^n} - V_{P^n} \right\|_1 \overset{n\to\infty}{\longrightarrow} 0$$

channel resolvability for general channels Given probability $P^n(x^n)$ on $\mathcal{X}^n$ (denoted by $\widehat{P} = \{P^n\}_{n=1}^{\infty}$)

$$C_R(\widehat{P}, \widehat{V}) := \inf\left\{ R \;\middle|\; \exists \{\Phi_n\}_{n=1}^{\infty}, \lim_{n\to\infty} d(V_{\Phi_n}, V_{P^n}) = 0, \limsup_{n\to\infty} \frac{1}{n} \log L_n \le R \right\}$$

main theorem

- Let $\delta(a|P_n, V_n) := \sum_{x^n} P^n(x^n) \operatorname{Tr} V_{x^n}^n \{ V_{x^n}^n - e^{na} V_{P^n}^n > 0 \}$
- remind that $a > \overline{I}(\widehat{P}, \widehat{V}) \iff \lim_{n\to\infty} \delta(a|P_n, V_n) = 0$

Given $\widehat{V} = \{V^n\}_{n=1}^\infty$ and $\widehat{P} = \{P^n\}_{n=1}^\infty$, for any nonnegative operator B, there exists a sequence of codes $\{\Phi_n\}_{n=1}^\infty$ such that

$$\left\| \frac{1}{L_n} \sum_{l=1}^{L_n} V_{x_l^n} - V_{P^n} \right\|_1 \le 4\sqrt{2} \sqrt{\sum_{x^n} P^n(x^n) \operatorname{Tr}(V_{x^n} - B)_+} + \sqrt{\frac{\operatorname{Tr} B}{L_n}}$$

- Letting $B = e^{na} V_{P^n}$, we have

$$d(V_{\Phi_n}, V_{P^n}) = \left\| \frac{1}{L_n} \sum_{l=1}^{L_n} V_{x_l^n} - V_{P^n} \right\|_1 \le 4\sqrt{2} \sqrt{\delta(a|P_n, V_n)} + \sqrt{\frac{e^{na}}{L_n}}$$

- Thus $L_n = e^{nR}$ and $R > a > \overline{I}(\widehat{P}, \widehat{V}) \Longrightarrow \lim_{n\to\infty} d(V_{\Phi_n}, V_{P^n}) = 0$

main theorem

- We have shown that $R > \overline{I}(\widehat{P}, \widehat{V})$ $(L_n = e^{nR})$ is achievable.
- There is possibility that R can be smaller.

$$C_R(\widehat{P}, \widehat{V}) \leq \overline{I}(\widehat{P}, \widehat{V})$$

- These arguments have been expected in Hayashi's textbook.
- difference: no pinching $\kappa_{V_{P_n}}$, operator B can be arbitrary

$$\left\| \frac{1}{L_n} \sum_{l=1}^{L_n} V_{x_l^n} - V_{P^n} \right\|_1 \leq 4\sqrt{2} \sqrt{\sum_{x^n} P^n(x^n) \operatorname{Tr}(V_{x^n} - B)_+} + \sqrt{\frac{\operatorname{Tr} B}{L_n}}$$

$$\left\| \frac{1}{L_n} \sum_{l=1}^{L_n} V_{x_l^n} - V_{P^n} \right\|_1 \leq 3 \sqrt{\sum_{x^n} P^n(x^n) \operatorname{Tr} V_{P_n} \{\kappa_{V_{P_n}}(V_{x^n}) - e^{na} V_{P^n} > 0\}}$$

$$+ \sqrt{\frac{\nu_n e^{na}}{L_n}} \quad \text{(Hayashi)} \quad (\nu_n = \#\{\text{eigen-value of } V_{P^n}\})$$

resolvability and wiretap channel coding

encoding by Alice $\quad l \sim \{1, 2, \ldots, L\}$ (uniform random number)

$$
\begin{array}{cccccccc}
 & & & \downarrow & & & & \\
 & x^n_{1,1} & \cdots & x^n_{1,l} & \cdots & x^n_{1,L} & \to \boxed{V^n} \to \frac{1}{L}\sum_{l=1}^{L} V_{x^n_{1,l}} \approx V_{P^n} \\
 & \vdots & & \vdots & & \vdots & \\
\text{(message)} & & & & & & \\
k \to & x^n_{k,1} & \cdots & x^n_{k,l} & \cdots & x^n_{k,L} & \to \boxed{V^n} \to \frac{1}{L}\sum_{l=1}^{L} V_{x^n_{k,l}} \approx V_{P^n} \\
 & \vdots & & \vdots & & \vdots & \\
 & x^n_{M,1} & \cdots & x^n_{M,l} & \cdots & x^n_{M,L} & \to \boxed{V^n} \to \frac{1}{L}\sum_{l=1}^{L} V_{x^n_{M,l}} \approx V_{P^n}
\end{array}
$$

$\to \boxed{W^n} \to$ Bob can distinguish each $x^n_{k,l}$ Eve cannot distinguish k

- By c-q channel coding theorem $M_n L_n \approx e^{n\{I(P,W)-\varepsilon\}}$ messages can be sent from Alice to Bob with vanishing error
- By c-q resolvability coding theorem $L_n \approx e^{n\{I(P,V)+\varepsilon\}}$ random number is enough to introduce confusion to Eve
- Thus $M_n \approx e^{n\{I(P,V)-I(P,V)-2\varepsilon\}}$ private messages can be sent

a general formula for c-q wiretap channel coding

- Using the general formula (Hayashi-Nagaoka, 2002) for message transmission from Alice to Bob,

$$C(\widehat{W}) \text{ (message transmission capacity)} = \sup_{\widehat{P}} \underline{I}(\widehat{P}; \widehat{W})$$

we have the following theorem

$$C_S(\widehat{W}, \widehat{V}) = \sup_{\widehat{P}, \widehat{T}} \left\{ \underline{I}(\widehat{P}; \widehat{T}\widehat{W}) - \overline{I}(\widehat{P}; \widehat{T}\widehat{V}) \right\} = \sup_{\widehat{P}, \widehat{T}: \, \overline{I}(\widehat{P}; \widehat{T}\widehat{V}) = 0} \underline{I}(\widehat{P}; \widehat{T}\widehat{W}),$$

where $\widehat{T} = \{T_n\}_{n=1}^{\infty}$ is a sequence of Markov maps, $\widehat{T}\widehat{V} = \{T_n V^n\}_{n=1}^{\infty}$, and $\widehat{P} = \{P_n\}_{n=1}^{\infty}$ is a sequence of probability functions. The supremum is taken over such possible choice of sequences.

6　まとめと展望

情報スペクトル的方法　$\widehat{\rho} = \{\rho_n\}_{n=1}^{\infty}, \widehat{\sigma} = \{\sigma_n\}_{n=1}^{\infty}$

$$R(\widehat{\rho}||\widehat{\sigma}) \overset{(1)}{=} \underline{D}(\widehat{\rho}||\widehat{\sigma}) \overset{(2):\text{i.i.d.}}{=} D(\rho||\sigma)$$


情報スペクトル量と，関連する基本タスク

- **エントロピー・スペクトル・レート** (Data Compression)
- **ダイバージェンス・スペクトル・レート**
  - $\underline{D}(\widehat{\rho}||\widehat{\sigma}), \overline{D}(\widehat{\rho}||\widehat{\sigma})$: hypothesis testing (strong converse)
- **相互情報量スペクトル・レート**
  - **下限，**$\underline{I}(\widehat{P};\widehat{W})$: cq channel coding
  - **上限，**$\overline{I}(\widehat{P};\widehat{W})$: cq channel resolvability

展望と応用：エンタングルメント変換, spin chain**の解析**, **量子統計力学**