

京都大学数理解析研究所共同利用研究会「情報物理学の数学的構造」2006 年 6 月 28 日-30 日

センシングと符号化の情報物理学¹

日本電信電話株式会社・NTT コミュニケーション科学基礎研究所
村山 立人 (Tatsuto Murayama)

NTT Communication Science Laboratories,
NIPPON TELEGRAPH AND TELEPHONE CORPORATION

1 はじめに

センシングには符号化が有効である。より詳しく述べると、多数のセンサー端末を用いて特定の情報源を調査する大自由度システムは、センサー情報を独立に不可逆圧縮することで情報利得が発生するメカニズムを内在する。本稿では、最近になって発見されたこの非自明な現象を説明する過程においても、情報理論の数学的ツールと統計力学の処方箋が整合的な結論を与えることを示す。

今、データセンターがあるデータ系列 $\{X(t)\}_{t=1}^{\infty}$ に興味を持っているが、これを直接計測できないものとする。そこで、データセンターは L 個のセンサーを周囲に配置したとしよう。各センサーはノイズのある環境で計測した系列 $\{Y_i(t)\}_{t=1}^{\infty}$ をそれぞれ独立に符号化する。つまり、各センサーは互いに通信することができず、したがって、事前にデータセンターに伝送する内容についていっさい協調できないものとする。データセンターは、 L 個の符号語を通信回線を利用して回収し、元の系列 $\{X(t)\}_{t=1}^{\infty}$ をできるだけ復元したいと考えている。しかし、このデータ系列だけがデータセンターにとっての重要な事柄ではないので、各センサーが利用できるデータ伝送率（通信速度）の合計 R は厳密に制限されている。つまり、データセンターは一定の回線速度でしか符号語を回収できない。このような推定操作を伴った分散型通信のモデルは Berger-Zhang-Viswanathan によって定式化され [1]、情報理論の立場からセンサーネットの理論的枠組みを提供していると解釈されている。

彼らの仕事によって、大規模観測系におけるいくつかの興味深い性質が明らかにされた。もしセンサーが互いに通信できるなら、センサー数 L が無限の極限において、独立に発生している計測ノイズを完全に除去することが可能となる。したがって、 $D(\cdot)$ を $\{X(t)\}$ の歪み・レート関数（レート・歪み関数の逆関数）として、データセンターは任意の忠実度で「歪み」 $D(R)$ を達成することが保障されている。逆に、センサーが互いに通信できないなら、有限の合計伝送率 R で歪み D を無限に小さくすることはできない。たとえ、無限個のセンサーが利用可能であったとしても、それは実現できないと証明できるのである [1]。

本稿では、有限の合計伝送率 R での分散化の極限 $L \rightarrow \infty$ における限界効率を解析的に議論するために、レート・歪み関数 $R(D)$ を利用する。より詳細にいうと、各センサーが符号化のメリットを最大限享受できると仮定し、データセンターは L 個の復号系列に

¹共同研究者：Peter Davis（日本電信電話株式会社・NTT コミュニケーション科学基礎研究所）

ビットごとの「多数決」を行うことによってベイズ最適な推定操作を実現するものとする [2]. このとき, 合計伝送率 R を既与として, どの程度のセンサー数 L が最適であるかを議論する不可逆性分散観測問題を本稿では提案する. 本稿の議論によって, 全システムの効率を $L \rightarrow \infty$ の極限で評価することが可能であるが, これは個々のセンサーが送信に利用できる伝送率がゼロに収束することを意味する. ここで, レート・歪み関数の連続性を直接的に利用した情報理論的アプローチと, 最適性を有する可解モデルを構成的に調べる統計物理学的アプローチが整合的な結果を与える事実は興味深い. また, 後者のアプローチにおいて, レプリカ法とスケールング的処方を組み合わせることにより, 不定形の極限を巧みに排除した理論構成を行ったのも本稿の技術的特色である.

2 システムモデル

データ系列 $\{X(t)\} \in \mathcal{X}$ に共通の確率分布を $P(x)$ とする. また, $\mathcal{Y}$ を計測系列 $\{Y_i(t)\}$ に共通のアルファベットとし, $\mathcal{X} \times \mathcal{Y}$ 上で定義される確率行列を $W(y|x)$ とする ($i = 1, \dots, L$, $t \geq 1$). まず, 無記憶情報源 $\{X(t)\}_{t=1}^{\infty}$ に対し, 同時確率分布を次のように仮定する.

$$\Pr[x, y_1, \dots, y_L] = P(x) \prod_{i=1}^L W(y_i|x).$$

ここで, 確率変数 $Y_i(t)$ は $X(t)$ に対して独立であり, 条件つき確率 $W[y_i(t)|x(t)]$ の値はすべての i と t に関して同一である. さらに本稿では, この問題をもっとも単純な 2 値系列で議論していく. つまり, データ系列 $\{X(t)\}$ と, それを一定のノイズレベルで計測した系列 $\{Y_i(t)\}$ はすべて 2 値系列であると仮定される. したがって, 確率行列は次のようにパラメータ化できる.

$$W(y|x) = \begin{cases} 1-p & (y=x) \\ p & (y \neq x) \end{cases}.$$

ここで, $p \in [0, 1]$ は計測におけるノイズのレベルを意味し, アルファベットは $\mathcal{X} = \mathcal{Y}$ と選択されている. さらに, 簡単のため, $P(x) = 1/2$ がいつも成立すると仮定しよう. これは, 全く冗長性のないランダムな情報源を計測していることに対応する.

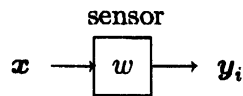


Figure 1: 計測のモデル: 各センサーは, 情報源を独立に計測する.

符号化の段階では, センサー i が計測系列 $\{y_i(t)\}_{t=1}^{\infty}$ から長さ n のブロック $\mathbf{y}_i = [y_i(1), \dots, y_i(n)]^T$ を切り取り, $\mathcal{Z}$ 上で定義された長さ m のブロック $\mathbf{z}_i = [z_i(1), \dots, z_i(m)]^T$ にブロックごと符号化する. 以後, ブール代数の表記にならい, $\mathcal{X} = \{0, 1\}$, したがって $\mathcal{Y} = \mathcal{Z} = \{0, 1\}$ とする. いま, $\hat{\mathbf{y}}_i$ をこのブロックの復号系列で, 圧縮系列の長さ $m (< n)$ が既与だとする. 本稿では, ハミング距離 $d_H(\cdot, \cdot)$ を歪み測度として採用し, 各エージェントは独立にレート・歪み関数を達成できると仮定する.

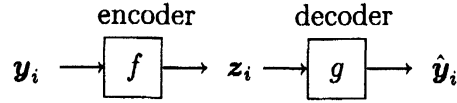


Figure 2: 符号化と復号: 各センサーの計測値は独立に符号化され, 符号語は受信したデータセンターによって復号される.

復号・推定の段階では, データセンターは L 個の符号語の系列 $z_1, \dots, z_L$ を回収することになる. 符号語の長さはすべて m なので, 合計の伝送率は $R = L \times m/n$ となる. そのため, この枠組みでは, データセンターは同一程度の歪みを持つ復号系列 $\hat{y}_1, \dots, \hat{y}_L$ を提供する交換可能なセンサーを配置していることになる. 最後に, 推定系列 $\hat{x} = [\hat{x}(1), \dots, \hat{x}(n)]^T$ の第 t 番目のビットは復号系列の対応する L 個のビットの多数決によって計算される [2]:

$$\hat{x}(t) = \begin{cases} 0 & (\hat{y}_1(t) + \dots + \hat{y}_L(t) \leq L/2) \\ 1 & (\hat{y}_1(t) + \dots + \hat{y}_L(t) > L/2) \end{cases} \quad (1)$$

よって, システム全体の性能は多数決 (1) によるビット誤り率の期待値 $P_e = \Pr[x \neq \hat{x}]$ によって定義するのが自然である. 本稿では, 分散化のレベルをシステムが持つ自由度と解釈して, 次の2つの「戦略」を考える.

1. 無限個のセンサーで系列を無限に圧縮する: $L \rightarrow \infty$
2. R 個のセンサーで系列を圧縮しない: $L = R$

前者の戦略では各センサーに配分される伝送率はゼロに収束し, 後者の戦略では符号化を行わないで通信をすることになる. しかし, 一般には, どちらの戦略がある特定の系列の推定に適しているのかを決定するのは難しい. つまり, どちらの戦略がより小さいビット誤り率の期待値 P_e を与えるのかが自明ではないのである. 実際, レート・歪み符号を用いることによって, データセンターはより多くの数のセンサーを利用することが可能になる. しかし, 同時に各センサーが提供する復号系列の歪みはより大きくなるだろう. 最適な分散化レベルの選択は, 計測におけるノイズレベル p と通信における合計伝送率 R に依存して決定されるはずである.

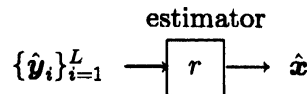


Figure 3: 推定: 計測された系列は, 多数の復号系列を利用して推定される.

3 大システム化戦略の利得構造

さて, 前章での議論を継承して計測におけるノイズレベルを p , そして通信における合計伝送率が有限の実数 R だとする. このとき $L \rightarrow \infty$ の極限では, データセンターの推定

におけるビット誤り率の期待値は

$$P_e(p, R) = \int_{-\infty}^{-(1-2p)c_g\sqrt{R}} \frac{dr}{\sqrt{2\pi}} \exp\left(-\frac{r^2}{2}\right) \quad (2)$$

となる。ただし、理想的なレート・歪み限界では $c_g = \sqrt{2\ln 2}$ と計算される（次章以降を参照）。これは、強制的に実行される独立復号過程を含むモデルの最高性能である。では、有限の合計伝送率 R が与えられたとき、推定系列の相対的品質がノイズの大きさ p にどのように依存するのかを議論しよう。Fig. 4 には、デシベル (dB) 単位で測られた $P_e(p, R)$ の典型的な挙動を示している。ただし、ここでは R を整数に制限し、参照レベルは次のように設定した。

$$P_e^{(0)}(p, R) = \begin{cases} \sum_{l=0}^{(R-1)/2} \binom{R}{l} (1-p)^l p^{R-l} & (R \text{ is odd}) \\ \sum_{l=0}^{R/2-1} \binom{R}{l} (1-p)^l p^{R-l} + \frac{1}{2} \binom{R}{R/2} (1-p)^{R/2} p^{R/2} & (R \text{ is even}) \end{cases} \quad (3)$$

参照レベル (3) はセンサー数 L を合計伝送率 R に一致させたときの P_e であり、これはセンサーが系列を全く圧縮しないシナリオに相当している。このとき、デシベル単位でのビット誤り率の期待値は

$$P_e^{(\text{dB})}(p, R) = 10 \log \frac{P_e(p, R)}{P_e^{(0)}(p, R)}, \quad (4)$$

と定義する。ただし、対数 $\log$ の底は 10 にとった。この単位でビット誤り率を測ることにすると、 $P_e(p, R)$ が参照レベル $P_e^{(0)}(p, R)$ と同じになるときゼロになる。定義 (4) より、デシベルで測った量は負の値をとる可能性がある。そのようなときは、測定している分散化レベルでのビット誤り率の期待値が、参照レベルのものより小さくなっていることを意味している。数値解析によると、合計伝送率 R が小さいとき（ナローバンド）は、整数 R の偶奇性に強く依存したビット誤り率の挙動が観測できる [Fig. 4 (a)]。ここで、本来なら実数でも定義されている合計圧縮率を整数に限定しているのは、比較している参照レベルを自然に導入したいからである。特に、 $R = 2$ のケースでは、最も小さな閾値の値を持ち、 $p_c = 0.082$ となっている。この閾値 p_c は、ここを超えると分散化の極限 $L \rightarrow \infty$ が参照レベル $L = R$ より大きな情報利得をもたらすことを意味する。しかし、 $R = 1$ のケースは特別で、このような閾値 p_c が存在していないのは興味深い。これとは対照的に、合計伝送率 R が大きいとき（ブロードバンド）は、デシベル単位で測ったビット誤り率の期待値の差 $P_e^{(\text{dB})}(p, R)$ が、定性的には良く似た挙動を示す [Fig. 4 (b)]。この図を見ると、 R が大きくなるにしたがって前述の閾値 p_c が 0.146 という値に漸近している様子がうかがえるが、この値の理論的導出には成功していない。

本章の結果より、次のことが主張できる。つまり、レート・歪み符号による不可逆圧縮の自由度を各センサー i に与えるとき、それを利用した計測精度の向上が見込めるノイズ領域が存在する。それは、特徴的な閾値 p_c を超えた高ノイズ領域であり、この区間 $[p_c, 1/2]$ では「数」の効果が「質」の効果を凌駕していると解釈できる。逆に、低ノイズ領域 $[0, p_c]$ では、「質」の効果が「数」の効果を凌駕しているので符号化による分散化の利得は得にくい。この結果は、計測と通信をうまく干渉させることで、システム全体として相乗効果が享受できる事実を理論的に示している。

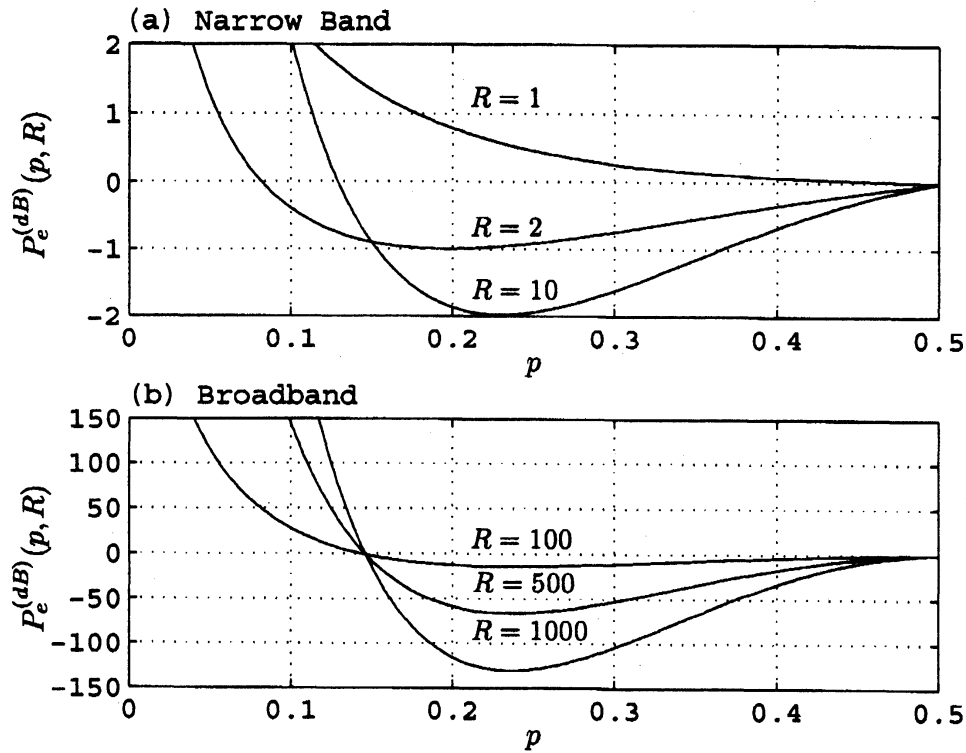


Figure 4: レート・歪み関数による分散符号化と独立復号過程による推定操作の性能評価. デシベル (dB) 単位で測った参照レベル $P_e^{(0)}(p, R)$ に対する $P_e(p, R)$ の相対的大きを測った. (a) 合計伝送率 R の小さいナローバンド回線. $R = 1, 2, 10$. (b) 合計伝送率 R の大きいブロードバンド回線. $R = 100, 500, 1000$.

4 情報理論

本章では、レート・歪み理論における Shannon の定理を最初に紹介する [3]. この定理は情報源符号化定理 (Shannon の第一定理), 通信路符号化定理 (Shannon の第二定理) に次ぐ第三の Shannon の定理であり, それは不可逆圧縮における限界記述長を与える. いま, ビット誤り率 D を許してデータを圧縮するとしよう. このとき, データの圧縮率が $r(D)$ より大きい限り, ビット誤り率が D より大きくならない符号化の方法が存在する. この限界の圧縮率 $r(D)$ を歪み D の関数とみなし, レート・歪み関数と呼ぶ. 特に, 単純な情報源のクラスでは, 簡単に $r(D)$ が構成できることが知られている. 仮に, n ビットのデータ $y_i = [y_i(1), \dots, y_i(n)]^T$ が m ビットの符号語 $z_i = [z_i(1), \dots, z_i(m)]^T$ に圧縮され, それが復号されて系列 $\hat{y}_i = [\hat{y}_i(1), \dots, \hat{y}_i(n)]^T$ を得たとしよう. 簡単のため, 元のデータ系列が全くのランダム系列であり, 冗長性を利用した圧縮が不可能な場合を考えることにする (一般化は容易である). すると, 圧縮率を $r = m/n$ で定義して, 復元におけるビット誤り率をハミング距離で測ることにすると, 上述のレート・歪み関数は

$$r(D) = \begin{cases} 1 - h(D) & (0 \leq D \leq 1/2) \\ 0 & (\text{otherwise}) \end{cases}, \quad (5)$$

と求まる. ただし, $h(\cdot)$ は 2 値エントロピー関数であり, 次のように定義された.

$$h(D) = -D \log_2 D - (1 - D) \log_2 (1 - D).$$

このレート・歪み関数 (5) は本稿で扱うシステム・モデルの分析にも便利な数学的道具である.

以後, センサー数が限りなく大きくなる極限を想定し, レート・歪み関数 $r(D)$ の $(D, r) = (1/2, 0)$ 近傍について議論していく. 関数 (5) の連続性により, $D \in [0, 1/2]$ においてテイラー展開すると

$$\begin{aligned} r(D) &= 1 - h(D) \\ &= \frac{2}{\ln 2} \left(\frac{1}{2} - D \right)^2 + \mathcal{O} \left(\left(\frac{1}{2} - D \right)^3 \right) \end{aligned}$$

となる. ただし, ランダウの記号 $\mathcal{O}(\cdot)$ はその引数より高次の無限小を意味する [4]. ここで, 関係式 $R/L = m/n$ を考慮すれば, センサー数 L と歪み D の間には漸近的に

$$\frac{R}{L} \approx \frac{2}{\ln 2} \left(\frac{1}{2} - D \right)^2. \quad (6)$$

が成立する.

一方, 各センサーが独立に計測系列の符号化を行うと仮定すると, 歪みに由来するビット誤りは Bernoulli 試行でモデル化できる [5]. そのため, 復号系列 $\hat{y}_i$ のビット誤り率は次のように求まる.

$$e = \Pr[x(t) \neq \hat{y}_i(t)] = p(1 - D) + (1 - p)D.$$

よって、推定系列 $\hat{x}$ のビット誤り率の期待値は累積二項分布 [6]

$$P_{\text{BER}}(e, L) = \Pr[x(t) \neq \hat{x}(t)]$$

$$= \begin{cases} B(\frac{L-1}{2} : e, L) & (L \text{ is odd}) \\ B(\frac{L}{2} - 1 : e, L) + \frac{1}{2}b(\frac{L}{2} : e, L) & (L \text{ is even}) \end{cases}$$

で記述できる。ただし、簡単のため

$$B(L' : e, L) = \sum_{l=0}^{L'} b(l : e, L),$$

$$b(l : L, q) = \binom{L}{l} (1-e)^l e^{L-l}$$

と略記した。ここで、整数 l は $\hat{y}(t)$ における反転していない要素の合計を表し、特に項 $(1/2)b(L/2 : e, L)$ は (ちょうど) $l = L/2$ となったときのランダムな推量を意味する。もちろん、記号 $\binom{L}{l}$ は L 個から l 個を選ぶ組み合わせの総数である。

では、センサーの数 L が十分に大きいとしよう。すると、累積二項分布は

$$P_{\text{BER}}(e, L) \approx B\left(\frac{L}{2} : e, L\right)$$

$$= \sum_{l=0}^{L/2} \binom{L}{l} (1-e)^l e^{L-l},$$

と近似できる。さらに、統計学における基本的定理によると、二項分布と正規分布には

$$P_e(p, R) = \lim_{L \rightarrow \infty} P_{\text{BER}}(e, L)$$

$$= \int_0^{L/2} du \, N(L(1-e), Le(1-e)) \quad (7)$$

という関係式が成立する。ただし、 $N(X, Y)$ は平均 X で分散 Y の正規分布を表している [6]。積分 (7) を標準正規分布の形式にするには、測度を $r = (u - L(1-e))/\sqrt{Le(1-e)}$, so $dr = du/\sqrt{Le(1-e)}$ と置き換えればよいことが知られている。その結果、ビット誤り率は

$$P_e(p, R) = \lim_{L \rightarrow \infty} \int_{-\sqrt{L}}^{-r_c} dr \, N(0, 1)$$

と求まる。ただし r_c は次式を満たす。

$$r_c = \frac{L(\frac{1}{2} - e)}{\sqrt{Le(1-e)}} \approx 2\sqrt{L}(1-2p) \left(\frac{1}{2} - D \right). \quad (8)$$

この関係式は与えられた L , p , D の値に対しては常に成立し、符号化の個性は D の値に集約される。結局、分散化利得の理論限界は (6) と (8) を連立して

$$P_e(p, R) = \int_{-\infty}^{(1-2p)\sqrt{2\ln 2R}} dr \, N(0, 1) \quad (9)$$

と求まる。これは前章の (2) に他ならない。

5 統計力学

一般に統計力学の方法では、構成的に定義されたシステムの挙動を解析的に調べることができる。特にシステムが本質的に持つ対称性が高く、平均場概念による近似の正当化がしやすい「可解モデル」に対して、その威力を発揮することが知られている。実際、物性科学の分野において、気体・磁性体・スピングラスなどの物質系を抽象化した数理モデルにこの方法はうまく適用されてきた。しかし近年は、グラフ上で定義された符号の性能評価問題が全く同じ数学的構造をしていることが発見され、情報理論の分野においても統計力学の方法が重要視されつつある。本稿ではまず、誤り訂正符号として提唱された低密度生成行列 (LDGM) 符号を逆問題的に用いた不可逆圧縮の方法を提案し、レート・歪み関数を理論上達成できる可解モデルを構成する。そして、このモデルにスピングラスの平均場理論である「レプリカ法」の処方を用い、大システム化戦略の利得構造を直接的に評価する。この議論では、レート・歪み関数の具体的形式を仮定せず、それをあるパラメータ極限で達成する符号族を分析対象にしているのが特徴である。このアプローチは、同様に定義された計算的複雑さの低い符号族の性能評価問題にも適用可能であり、その汎用性は広いと考えられる。

このシナリオでは、 $n \times m$ 型行列の 2 値行列 A_i ($i = 1, \dots, L$) をエージェント分だけ準備し、 m ビットの系列 $\mathbf{z}_i = [z_i(1), \dots, z_i(m)]^T$ が線形復号条件

$$\hat{\mathbf{y}}_i = A_i \mathbf{z}_i \pmod{2}, \quad (10)$$

と忠実度規範

$$D = \frac{1}{n} d_H(\mathbf{y}_i, \hat{\mathbf{y}}_i)$$

を満足するときに符号語 (のひとつ) として定義する [7]。ここで、ハミング距離 $d_H(\cdot, \cdot)$ が歪み測度として採用されている。また、式 (10) では 2 を法とした加法を用いていることに注意。今、行列 A_i の各行にそれぞれ K 個、各列に C 個だけ非ゼロ要素の 1 が存在するように作成したとする。このとき、有限でしかも通常は小さい値を持つ K と C によって、LDGM 符号の符号族が指定されることになる。ここで、パラメータ K の値が非常に大きくなると LDGM 符号はレート・歪み関数を達成することが知られている。そのため、LDGM 符号化における $K \rightarrow \infty$ の極限を構成的に議論できるのなら、レート・歪み関数の存在を仮定した情報理論の分析と整合的な結論を与えるはずである。

最近、統計力学の方法を利用して LDGM 符号の性能を分析する方法が確立された [7]。本稿ではこの方法を用いて LDGM 符号の理論限界を導出し、分散符号化による効果をスケール理論的な処方箋にしたがって計算してみる [8]。まず第一に、アルファベット $\mathcal{Z} = \{0, 1\}$ を統計力学で多用するアルファベット $\mathcal{S} = \{+1, -1\}$ に翻訳する。すると、表現の整合性を維持するため、 $\mathcal{Z} = \{0, 1\}$ 上で定義される「加法」も、 $\mathcal{S} = \{+1, -1\}$ 上で定義される「乗法」に翻訳する必要がある。例えば、 $z_i(s) + z_i(s') \pmod{2}$ という数式は、 $\sigma_i(s) \times \sigma_i(s') \in \mathcal{S}$ と書き換わる。同様にして、 $y_i(t)$ を $J_i(t)$ に書き換えることができる。ただ簡単のため、 L 個のセンサーを区別するための指標 i は省略することにする。以後、Sourlas の処方箋 [9] に従い、Gibbs-Boltzmann 分布

$$\Pr[\boldsymbol{\sigma}] = \frac{\exp[-\beta H(\boldsymbol{\sigma}|\mathbf{J})]}{Z(\mathbf{J})} \quad (11)$$

を計算する。ただし、分配関数（規格化定数）

$$Z(J) = \sum_{\sigma} e^{-\beta H(\sigma|J)}$$

とハミルトニアン（エネルギー関数）

$$H(\sigma|J) = - \sum_{s_1 < \dots < s_K} \mathcal{A}_{s_1 \dots s_K} J[t(s_1, \dots, s_K)] \sigma(s_1) \dots \sigma(s_K) \quad (12)$$

を適切に定義して用いた。ここで、時系列の指標 $t(s_1, \dots, s_K)$ は、符号語の指標 $s_1, \dots, s_K$ の集合に対応した t の値を指定し、パリティ検査条件 (10) を満足させている。さらに、相互作用の希釈性を表現している対称テンソル $\mathcal{A}_{s_1 \dots s_K}$ の各要素は、指標集合 $(s_1, \dots, s_K)$ の組み合わせに依存して 0 か 1 の値をとる。この符号化では指標 s に対して C 個の 1 がランダムに選択されるので、 $\sum_{s_2, \dots, s_K} \mathcal{A}_{s s_2 \dots s_K} = C$ が成立している。このとき、復号系列はひとつの指標 s に対して C 個のビットを持つが、これは K 個のビットを符号語から抽出していることになる。よって、符号化のレートは $R/L = K/C$ となっている。また、ハミルトニアン (12) が復号したときのエラー $[1 - J[t(s_1, \dots, s_K)] \cdot \sigma(s_1) \dots \sigma(s_K)]/2$ を記録していることも容易に理解できる。

さらに統計力学によると、客観的な観測にかかる測定量（オブザーバブル）は自由エネルギーを利用することで解析的に計算が可能になっている。ここで、自由エネルギーとは、次式のように定義される関数である。

$$f = -\frac{1}{\beta} \langle \ln Z(J) \rangle_{A,J} \quad (13)$$

ここで、 β は Gibbs-Boltzmann 分布 (11) のパラメータで「逆温度」と呼ばれる。記号 $\langle \cdot \rangle_{A,J}$ は配位平均を意味する。このため、自由エネルギーを計算するためには、分配関数 $Z(J)$ の対数に関する配位平均 $\langle \cdot \rangle_{A,J}$ を実行する必要がある。しかし、これは数学的に困難な課題なので、いわゆるレプリカ法が利用される [10]。そして、自由エネルギー (13) が解析的に計算できると、LDGM 符号化における平均歪み D も次の関係式より直ちに求まることが知られている [7]。

$$D = \frac{1+f}{2}.$$

よって、自由エネルギー (13) の $L \rightarrow \infty$ での振舞いを分析できれば、関係式 (8) より P_e の評価が可能となる。以下、この処方箋にしたがった結果を要約し、統計力学による方法の汎用性と情報理論の結果との整合性を確認しよう。

線形復号条件 (10) による単一情報源の符号化は、自由エネルギー

$$\begin{aligned} f = & -\frac{1}{\beta n} \left[\ln \cosh \beta - K \langle \ln [1 + \tanh(\beta x) \tanh(\beta \hat{x})] \rangle_{\pi(x), \hat{\pi}(\hat{x})} \right. \\ & + \frac{1}{2} \left\langle \sum_{J=\pm 1} \ln \left[1 + \tanh(\beta J) \prod_{l=1}^K \tanh(\beta x_l) \right] \right\rangle_{\pi(x)} \\ & \left. + \frac{C}{K} \left\langle \ln \sum_{\sigma=\pm 1} \prod_{l=1}^C [1 + \sigma \tanh(\beta \hat{x}_l)] \right\rangle_{\hat{\pi}(\hat{x})} \right], \quad (14) \end{aligned}$$

で記述される。ただし、 $\langle \cdot \rangle_{\pi(x)}$ は $p(x_l)$ に関する平均操作などを意味する。ここで、秩序変数 $\pi(x)$ と $\hat{\pi}(\hat{x})$ に関する自由エネルギー (14) の変分問題を考えれば、鞍点方程式として

$$\pi(x) = \left\langle \delta \left[x - \sum_{l=1}^{C-1} \hat{x}_l \right] \right\rangle_{\hat{\pi}(\hat{x})},$$

$$\hat{\pi}(\hat{x}) = \left\langle \frac{1}{2} \sum_{J=\pm 1} \delta [\hat{x} - \mu(x_1, \dots, x_{K-1}; J)] \right\rangle_{\pi(x)}$$

を得る。ただし、簡単のため

$$\mu(x_1, \dots, x_{K-1}; J) = \frac{1}{\beta} \tanh^{-1} \left[\tanh(\beta J) \prod_{l=1}^{K-1} \tanh(\beta x_l) \right]$$

と略記した。

一般に、自由エネルギー (14) によって記述される符号化が実現されるとき、システム全体のビット誤り率の評価は数値的に実施されるべきである。ところが、 $K \rightarrow \infty$ の漸近論では、積分に有意な項だけを抽出する操作が解析的に可能である。実際、秩序変数 $\hat{\pi}(\hat{x})$ に対する主要項は $\sqrt{C\sigma^2}$ 近傍の x のみからであり、

$$\hat{\pi}(\hat{x}) \approx \left\langle \delta \left[\hat{x} - y\beta^K (C\sigma^2)^{\frac{K}{2}} \right] \right\rangle$$

と評価できる。すると、レプリカ法の枠組では、自由エネルギー (14) は方程式

$$\sqrt{L}f = -\frac{\sqrt{\alpha_c}}{2} - \frac{R}{\sqrt{\alpha_c}} \ln 2,$$

$$0 = -\frac{1}{2} + \frac{R}{\alpha_c} \ln 2$$

を満足することになる。ただし、 $\alpha_c = \beta^2 L$ とした。この結果は (2) という関数形と $c_g = \sqrt{2 \ln 2}$ を与えるので、前章で紹介したレート・歪み関数を前提とした議論と一致している。

6 おわりに

本稿で扱った大自由度センシングのモデル化では、センサー情報の通信コストをセンサー自体の設置コストより大きく見積もっているのがその最大の特徴である。実は、このようにエネルギー効率を重要視する省資源化の技術動向は、あらゆる分野に波及しつつある。例えば、半導体分野におけるプロセッサのマルチコア戦略も、全く同じ設計思想を持つと考えられる [11]。つまり、本稿で示した新しい最適化の概念は、より汎用性の高いシステムの設計原理として再解釈される可能性もある。そして、このような革新技術を支える理論的基盤として、古典的なレート・歪み理論と最新のレプリカ法が再発見される事実も興味深い。

Acknowledgments

本稿を作成するにあたり、議論に参加してくれた村松純氏と上田修功氏に感謝する。

References

- [1] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Trans. Inform. Theory*, vol. 42, pp. 887–902, May 1996.
- [2] D. J. C. MacKay, *Information Theory, Inference and Learning Algorithms*. Cambridge, UK: Cambridge University Press, 2003.
- [3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 1991.
- [4] S. Lang, *A First Course in Calculus*. Springer-Verlag, 1986.
- [5] K. L. Chung, *A Course in Probability Theory*. Academic Press, 2000.
- [6] W. Hays, *Statistics (5th Edition)*. Belmont, CA: Wadsworth Publishing, 1994.
- [7] T. Murayama and M. Okada, "Rate distortion function in the spin glass state: a toy model," in *Advances in Neural Information Processing Systems 15 (NIPS'02)*, Vancouver, Canada, Dec. 2002, pp. 423–430.
- [8] T. Murayama and P. Davis, "Rate distortion codes in sensor networks: A system-level analysis," in *Advances in Neural Information Processing Systems 18 (NIPS'05)*, Vancouver, Canada, Dec. (in press).
- [9] N. Sourlas, "Spin-glass models as error-correcting codes," *Nature*, vol. 339, pp. 693–695, June 1989.
- [10] V. Dotsenko, *Introduction to the Replica Theory of Disordered Statistical Systems*. Cambridge, UK: Cambridge University Press, 2001.
- [11] (2006) Intel@Multi-Core. [Online]. Available: <http://www.intel.com/multi-core/>