

# ネットワーク上の情報熱力学とその生体情報処理への応用

伊藤 創祐

北海道大学 電子科学研究所

2017年9月27日

## 概要

ゆらぎの熱力学 (stochastic thermodynamics) の入門的な解説を行い、ゆらぎの熱力学において情報理論がどのように取り入れられるかをみる。さらに、多成分が相互に情報伝達を行っている場合を取り扱えるよう、ベイジアンネットワークを導入してゆらぎの熱力学を拡張する。また応用例として、生体の情報処理の一種である大腸菌の走化性シグナル伝達についての議論を行う。

## 1. 導入

古典的な熱力学はマクロな系を現象論的に取り扱うことで成功を取めた学問である [1, 2]。古典的な熱力学における特徴の一つとして、エントロピーがマクロな状態量として自然に導入されることが挙げられる。マクロな状態量としてのエントロピーは相加性<sup>\*1</sup>を示し、熱力学第二法則を満たすという性質を持っている。

一方近年注目を集めているのは、ゆらぎの効果が支配的になるメゾスコピックな領域<sup>\*2</sup>での熱力学である。これはゆらぎの熱力学 (stochastic thermodynamics)<sup>\*3</sup>と呼ばれ、コロイド粒子や細胞内部の化学反応など確率過程で記述される物理現象に適用することができる [3, 4]。ゆらぎの熱力学が成立する理由は、様々な系で熱力学第二法則を数学的に導出できるゆらぎの定理 (Fluctuation theorem) と呼ばれる理論が背景にある<sup>\*4</sup>。これにより、ゆらぎの熱力学で扱えるメゾスコピックな系では、古典的な熱力学に比べてより精密に非平衡系の熱力学を考察できるようになった。

古典的な熱力学とゆらぎの熱力学の違いの一つに、エントロピーの相加性を要請しないという点が挙げられる。このエントロピーの非相加性は、古典的な熱力学においては系のマクロ性から無視されていた部分であったが、メゾスコピック系においては無視できない。しかしこの事実は厄介なものではなく、むしろゆらぎの熱力学自体に大きな深みを与えてくれる。なぜなら、この非相加性こそが情報理論 [5, 6] で語られていた「情報」、すなわち相互情報量の定義<sup>\*5</sup>そのものになるからである。

情報と熱力学の関係は、従来「マクスウェルのデーモン」のパラドックスを解決する方法として長らく議論されていたものであった [7]。近年のゆらぎの熱力学と情報理論の融合研究の進展、すなわち「情報熱力学」の研究の進展も、マクスウェルのデーモンの問題を考察することによって発展してきた [8]。しかし、今や情報熱力学はもはやマクスウェルのデーモンの為の理論<sup>\*6</sup>というだけでなく、ゆらぎの熱力学において相互作用する系を扱うときの基本的な考え方になりつつある。本稿では情報熱力学の考え方について、平易な解説を行う

<sup>\*1</sup> 系  $A$  と系  $B$  を足した合成系  $\{A, B\}$  のエントロピー  $S_{\text{th}}(A, B)$  が、それぞれの系のエントロピーの和でかける  $S_{\text{th}}(A, B) = S_{\text{th}}(A) + S_{\text{th}}(B)$  ことを意味する。

<sup>\*2</sup> コロイド粒子の Brown 運動などに代表される、確率的な記述が有効なマクロとマイクロの中間領域 (nm ~  $\mu\text{m}$  オーダーの領域) である。

<sup>\*3</sup> もしくは、ゆらぎのエネルギー論 (stochastic energetics)、確率的熱力学とも。

<sup>\*4</sup> ゆらぎの定理はエントロピーの変化量のゆらぎについての定理である。この定理を用いると、「期待値の意味での」エントロピー増大則を導ける。本稿ではこの「期待値の意味での」エントロピー増大則を熱力学第二法則と呼称する。

<sup>\*5</sup> 二つの確率変数  $A, B$  の間の相互情報量は、Shannon エントロピー  $H$  を用いて  $I(A; B) = H(A) + H(B) - H(A, B)$  で表され、相加性の破れを意味する。Shannon エントロピーの定義や相互情報量については、3 節で詳しく述べる。

<sup>\*6</sup> マクスウェルのデーモンと情報熱力学の関係は、既に講義ノートが存在しているため [9]、本稿では敢えて触れない。情報熱力学自体はゆらぎの熱力学における一般的な考え方であり、本来「マクスウェルのデーモン」とは独立に語るができるものである。

ことを目標とする。

また、情報熱力学の応用例として筆者が面白いと思っているものの一つに、細胞内のシグナル伝達系がある。これは多成分の化学反応が相互作用することで情報を伝達する生体内の仕組みであり、メゾスコピック系のために「ゆらぎ」の影響が無視できない。そのため、熱力学量と「情報」の密接な関係を考える上で適した例題である。特に生体系では、正確に情報を伝達できるか否かが淘汰の条件となるため、情報通信が“熱力学的な意味で”効率化されている可能性がある。本稿では筆者のシグナル伝達系の研究を紹介することで、情報熱力学がいかに応用されるかを見ていきたい。

本稿の目的は、これからゆらぎの熱力学や情報熱力学を学びたい人への指針となるものを提供することである。まず2節で、ゆらぎの熱力学の考え方を紹介する。3節では情報理論での基本的な量を導入し、3節で定義された量から熱力学第二法則を導くことを4節で解説する。5節では情報熱力学の考え方を紹介し、3節で導入された相互情報量がどのように熱力学第二法則に現れるかをみる。6節で、情報熱力学を一般化する理論であるネットワーク上の情報熱力学の考え方を紹介する。7節では、相互作用する二体系に着目して、熱力学第二法則と情報熱力学の結果の違いを説明する。8節では、細胞内のシグナル系の一種である適応現象について説明する。9節では、適応のシグナル伝達系に情報熱力学を応用する。最後に10節で、近年のさまざまな情報熱力学の研究について紹介を行う。

## 2. ゆらぎの熱力学入門

まず、ゆらぎメゾスコピック系を記述する Langevin 方程式に注目して、ゆらぎの熱力学の基本的な考え方を紹介する。ゆらぎの熱力学に関するより詳しい内容は [3] の教科書を参考にしてほしい。

Langevin 方程式の一例として、次のようなノイズ項  $\sqrt{2\gamma_{\mathcal{X}}T_{\mathcal{X}}}\xi(t)$  を含んだ運動方程式が挙げられる。これは液体中のコロイド粒子の Brown 運動の記述などに使われる。

$$m_{\mathcal{X}}\ddot{x}(t) = -\gamma_{\mathcal{X}}\dot{x}(t) - \partial_x U_{\mathcal{X}}(x(t), \lambda_{\mathcal{X}}(t)) + \sqrt{2\gamma_{\mathcal{X}}T_{\mathcal{X}}}\xi_{\mathcal{X}}(t). \quad (1)$$

ここで、 $m_{\mathcal{X}}$  はコロイド粒子の質量、 $\ddot{x}(t)$  は加速度、 $\gamma_{\mathcal{X}}$  は摩擦係数、 $\dot{x}(t)$  は粒子の速度、 $\partial_x$  は  $x$  の偏微分を表す演算子、 $U_{\mathcal{X}}(x(t), \lambda(t))$  はポテンシャルエネルギー、 $x(t)$  は粒子の位置、 $\lambda_{\mathcal{X}}(t)$  は外部のコントロールパラメータ、 $T_{\mathcal{X}}$  は熱浴の温度<sup>\*7</sup>、 $\xi_{\mathcal{X}}(t)$  は平均0分散1の白色 Gauss ノイズである。また、この Langevin 方程式で記述される系を  $\mathcal{X}$  と呼ぼう。

以後の議論の簡単のために、 $m_{\mathcal{X}}/\gamma_{\mathcal{X}} \rightarrow 0$  の極限 (overdamped 極限) である、次の Langevin 方程式を考えよう (ただし、以後の議論の簡単のために  $\gamma_{\mathcal{X}} = 1$  とした)。

$$\dot{x}(t) = -\partial_x U_{\mathcal{X}}(x(t), \lambda_{\mathcal{X}}(t)) + \sqrt{2T_{\mathcal{X}}}\xi_{\mathcal{X}}(t). \quad (2)$$

この Langevin 方程式に、熱力学第一法則を用いて熱を導入するのが本節の目的である。

まず系  $\mathcal{X}$  のポテンシャルエネルギー変化量  $dU_{\mathcal{X}}$  は  $U_{\mathcal{X}}$  に関する Chain rule から

$$dU_{\mathcal{X}} = \dot{x}dt \circ \partial_x U_{\mathcal{X}}(x, \lambda_{\mathcal{X}}) + \dot{\lambda}_{\mathcal{X}}dt \circ \partial_{\lambda_{\mathcal{X}}} U_{\mathcal{X}}(x, \lambda_{\mathcal{X}}), \quad (3)$$

のように与えられる。ただしここで  $\circ$  は Stratonovich 積分<sup>\*8</sup>であり、 $\dot{\lambda}_{\mathcal{X}}$  は外部のコントロールパラメータの変化率、 $dt$  は無限小時間を意味する。

このとき、外部のパラメータに依存したポテンシャルエネルギー変化の項を系  $\mathcal{X}$  にした仕事  $dW_{\mathcal{X}} := \dot{\lambda}_{\mathcal{X}}dt \circ \partial_{\lambda_{\mathcal{X}}} U_{\mathcal{X}}(x, \lambda_{\mathcal{X}})$  とみなすと、熱力学第一法則から系  $\mathcal{X}$  に流れ込んだ熱  $dQ_{\mathcal{X}}$  は

$$\begin{aligned} dQ_{\mathcal{X}} &:= dU_{\mathcal{X}} - dW_{\mathcal{X}} \\ &= \dot{x}dt \circ \partial_x U_{\mathcal{X}}(x, \lambda_{\mathcal{X}}) \end{aligned} \quad (4)$$

<sup>\*7</sup> ただし Boltzmann 定数  $k_B$  を 1 に規格化した温度である

<sup>\*8</sup> 特異関数であるノイズ項  $\xi_{\mathcal{X}}(t)dt$  を含む量との積は、積の取るときの離散化の仕方によって値が変わりうる。特に chain rule が成り立つ離散化の方法として Stratonovich 積分がある [3, 10]。任意の関数  $A(t)$  とノイズ  $\xi_{\mathcal{X}}(t)dt := dB_{\mathcal{X}t}$  との Stratonovich 積分  $\circ$  は、 $A(t) \circ \xi_{\mathcal{X}}(t)dt := \frac{A(t+dt)+A(t)}{2}dB_{\mathcal{X}t}$  のように定義される。

のように定義される。よって単位時間あたりに熱浴<sup>\*9</sup>から系 $\mathcal{X}$ に流れ込んだ熱流 $j_{\mathcal{X}} := dQ_{\mathcal{X}}/dt$ は

$$\begin{aligned} j_{\mathcal{X}}(t) &= \dot{x}(t) \circ \partial_x U_{\mathcal{X}}(x(t), \lambda_{\mathcal{X}}(t)) \\ &= \dot{x}(t) \circ (\sqrt{2T_{\mathcal{X}}}\xi_{\mathcal{X}}(t) - \dot{x}(t)), \end{aligned} \quad (5)$$

となる。

このように Langevin 方程式で定義された熱流 $j_{\mathcal{X}}$ は、熱力学第一法則を満たすだけの数学的な定義でしかない。よって、熱流が持つべき性質の一つである熱力学第二法則が成り立つのかは非自明である。実際、熱流 $j_{\mathcal{X}}$ は確率的には様々な値、すなわち正の値も負の値も取り得る。

しかしながら、驚くべきことにアンサンブル平均 $\langle \dots \rangle$ の意味では次のような不等式が成り立つのである。

$$dS_{\mathcal{X}}(t) \geq \frac{\langle j_{\mathcal{X}}(t) \rangle}{T_{\mathcal{X}}} dt, \quad (6)$$

ここで $dS_{\mathcal{X}}(t)$ は時刻 $t$ での系 $\mathcal{X}$ の時間 $dt$ あたりのエントロピー変化量<sup>\*10</sup>であり、位置 $x(t) =: x_t$ に関する Shannon エントロピー

$$H(X_t) := - \int dx_t p_{X_t}(x_t) \ln p_{X_t}(x_t), \quad (7)$$

を用いて、 $dS_{\mathcal{X}}(t) := H(X_{t+dt}) - H(X_t)$ のように定義される。ここで $p_{X_t}(x_t)$ は時刻 $t$ における位置 $x_t$ の確率分布を意味する。特に定常状態においては確率分布が時間変化しないため、 $dS_{\mathcal{X}} = 0$ となり、

$$0 \geq \langle j_{\mathcal{X}} \rangle, \quad (8)$$

が成り立つ。よって、定常状態では系 $\mathcal{X}$ への熱の流入は(期待値の意味で)起こらないという結果が導かれる。

また、熱浴のエントロピー変化を $dS_{\text{bath}}(t) := -\langle j_{\mathcal{X}}(t) \rangle dt / T_{\mathcal{X}}$ と定義してしまえば、式(6)は系と熱浴のエントロピー変化量の和の増大則

$$dS_{\mathcal{X}}(t) + dS_{\text{bath}}(t) \geq 0, \quad (9)$$

だとみなせる。

このように、ゆらぐメゾスコピック系において熱力学第二法則を考える際には、自然と Shannon エントロピーのような情報理論における概念が現れてくる。また、熱力学第二法則[式(6)]が成り立つという事実は数学的に示せるものであり、この証明自体も情報理論の考え方をを用いて理解することができる。よって、式(6)を証明する前に、一旦情報理論の紹介を経由することにしよう。

### 3. 情報理論入門

この節では情報理論における各量についての数学的な定義と性質について述べる。情報理論に関する詳しいことをより知りたい場合は、教科書[6]を参考にしてほしい。

まず、前章で述べたように、確率変数 $X$ に関する Shannon エントロピー $H(X)$ は

$$H(X) := - \int dx p_X(x) \ln p_X(x), \quad (10)$$

のように定義される。ここで $p_X(x)$ は確率変数 $X$ が $x$ の値をとる確率であり、 $\int dx p_X(x) = 1$ を満たす。この Shannon エントロピーは分布 $p_X(x)$ のランダムさの尺度であり、分布が一様に近いほど大きな値を持つ。

<sup>\*9</sup> 熱浴はこのときノイズ項 $\xi_{\mathcal{X}}(t)$ の発生源に相当する。例えば液体中のコロイド粒子の運動を Langevin 方程式が記述している時には、液体が熱浴に相当する。

<sup>\*10</sup> この Shannon エントロピーによる定義が、系 $\mathcal{X}$ のエントロピーの定義として妥当かどうかは本稿では問わない。

同様に、確率変数  $X_1, \dots, X_n$  の集合に関する Shannon エントロピーは、確率変数  $X_1, \dots, X_n$  がそれぞれ  $x_1, \dots, x_n$  を値を持つ同時確率分布  $p_{X_1, \dots, X_n}(x_1, \dots, x_n)$  を用いて、

$$H(X_1, \dots, X_n) := - \int dx_1 \cdots dx_n p_{X_1, \dots, X_n}(x_1, \dots, x_n) \ln p_{X_1, \dots, X_n}(x_1, \dots, x_n), \quad (11)$$

のように定義される。<sup>\*11</sup> また、確率変数  $Y$  を知っているという条件の下での確率変数  $X$  の条件付き Shannon エントロピー  $H(X|Y)$  は、

$$H(X|Y) := H(X, Y) - H(Y), \quad (12)$$

のように定義される。

さらに、二つの確率変数  $X, Y$  間の相関を表す量として、相互情報量  $I(X; Y)$  を

$$I(X; Y) := H(X) - H(X|Y) = H(X) + H(Y) - H(X, Y), \quad (13)$$

のように定義する。この量は、 $X$  が  $Y$  に（もしくは  $Y$  が  $X$  に）関する「情報」をどれだけ持っているか、を意味する量である。この量は常に非負  $I(X; Y) \geq 0$  であり、等号成立条件は確率変数  $X$  と  $Y$  が独立、すなわち任意の  $x$  と  $y$  で  $p(x, y) = p(x)p(y)$  を満たす。

同様に  $Z$  という条件の元での確率変数  $X, Y$  の条件付き相互情報量は

$$I(X; Y|Z) := H(X|Z) - H(X|Y, Z), \quad (14)$$

のように定義される。

さて、この相互情報量の非負性を証明するために、相対エントロピーと呼ばれる量を導入しよう。二つの確率分布  $p_X(x)$  と  $q_X(x)$  の間の相対エントロピーは、

$$D(p_X||q_X) := \int dx p_X(x) \ln \frac{p_X(x)}{q_X(x)}, \quad (15)$$

で与えられる。この量は常に  $D(p_X||q_X) \geq 0$  を満たし、任意の  $x$  で  $p_X(x) = q_X(x)$  のときのみ  $D(p_X||q_X) = 0$  となる。この非負性は  $\ln$  の凸性より示すことが可能である。なぜならば、凸関数に対する Jensen の不等式<sup>\*12</sup>より

$$\begin{aligned} -D(p_X||q_X) &= \langle \ln[q_X(x)/p_X(x)] \rangle \\ &\leq \ln[\langle q_X(x)/p_X(x) \rangle] = \ln \left[ \int dx q_X(x) \right] = 0, \end{aligned} \quad (16)$$

となり示せる。この相対エントロピーの非負性を用いて、相互情報量の非負性も示せる。なぜなら、相互情報量  $I(X; Y)$  は

$$I(X; Y) = D(p_{X,Y}||p_X p_Y), \quad (17)$$

で与えられるからである。同様に、次のように定義される条件付き相対エントロピー

$$D(p_{X|Y}||q_{X|Y}) := \int dx dy p_{X,Y}(x, y) \ln \frac{p_{X|Y}(x|y)}{q_{X|Y}(x|y)}, \quad (18)$$

も非負であることが示せ<sup>\*13</sup>、そこから条件付き相互情報量の非負性<sup>\*14</sup>も示せる。

<sup>\*11</sup> 同様に、以下で出てくる条件付き Shannon エントロピー、相互情報量、条件付き相互情報量、相対エントロピーも確率変数の集合に対しても定義される。

<sup>\*12</sup> Jensen の不等式とは、上に凸な関数  $f$ 、 $x$  を変数を持つ関数  $g$  に対して、 $\langle f(g(x)) \rangle \leq f(\langle g(x) \rangle)$  が成り立つことをいう。ただし  $\langle \cdots \rangle$  は確率分布  $p_X(x)$  によるアンサンブル平均  $\langle \cdots \rangle := \int dx p_X(x) \cdots$  を意味する。

<sup>\*13</sup> 条件付き確率  $p_{X|Y}(x|y)$  は同時確率分布  $p_{X,Y}(x, y)$  を用いて、 $p_{X|Y}(x|y) = p_{X,Y}(x, y) / \int dx p_{X,Y}(x, y)$  のように定義される。

<sup>\*14</sup> すなわち  $I(X; Y|Z) = D(p_{X,Y|Z}||p_{X|Z} p_{Y|Z}) \geq 0$

#### 4. ゆらぎの定理と熱力学第二法則

この節ではゆらぎの定理として知られている考え方を經由して、前節で取り上げた情報理論を用いて熱力学第二法則 (6) を導出することを目標とする。<sup>\*15</sup>

まず、式 (2) を微小時間  $dt$  で離散化した式を考えよう。

$$x_{t+dt} - x_t = -\partial_x U_{\mathcal{X}}(x_t, \lambda_{\mathcal{X}t})dt + \sqrt{2T_{\mathcal{X}}}dB_{\mathcal{X}t}. \quad (19)$$

ここで、 $x_t := x(t)$ ,  $\lambda_{\mathcal{X}t} := \lambda_{\mathcal{X}}(t)$  であり、 $dB_{\mathcal{X}t} := \xi_{\mathcal{X}}(t)dt$  は次のような Gauss 分布に従う

$$p(dB_{\mathcal{X}t}) = \frac{1}{\sqrt{2\pi dt}} \exp \left[ -\frac{(dB_{\mathcal{X}t})^2}{2dt} \right]. \quad (20)$$

よってここから、 $x_t$  から  $x_{t+dt}$  への遷移確率  $\mathcal{T}$  は

$$\mathcal{T}(x_{t+dt}; x_t) := p_{X_{t+dt}|X_t}(x_{t+dt}|x_t) = \frac{1}{\sqrt{2\pi dt}\sqrt{2T_{\mathcal{X}}}} \exp \left[ -\frac{\left( \frac{x_{t+dt}-x_t}{dt} + \partial_x U_{\mathcal{X}}(x_t, \lambda_{\mathcal{X}t}) \right)^2 dt}{4T_{\mathcal{X}}} \right], \quad (21)$$

のように与えられる。ここで  $1/\sqrt{2T_{\mathcal{X}}}$  は  $\partial[dB_{\mathcal{X}t}]/\partial x_{t+dt}$  の Jacobian である。

ゆらぎの定理の基本的な考え方は、この遷移確率と、時間逆方向に定義した遷移確率

$$p_{B_{X_t}|X_{t+dt}}(x_t|x_{t+dt}) = \mathcal{T}(x_t; x_{t+dt}), \quad (22)$$

を比較することにある。いま、 $p_{X_{t+dt}|X_t}$  と  $p_{B_{X_t}|X_{t+dt}}$  の二つの遷移確率の比をとってみると、

$$\begin{aligned} \frac{p_{X_{t+dt}|X_t}(x_{t+dt}|x_t)}{p_{B_{X_t}|X_{t+dt}}(x_t|x_{t+dt})} &= \exp \left[ -(x_{t+dt} - x_t) \frac{\partial_x U_{\mathcal{X}}(x_t, \lambda_{\mathcal{X}t}) + \partial_x U_{\mathcal{X}}(x_{t+dt}, \lambda_{\mathcal{X}t+dt})}{2T_{\mathcal{X}}} \right] \\ &= \exp \left[ -\frac{1}{T_{\mathcal{X}}} \dot{x}(t)dt \circ \partial_x U_{\mathcal{X}}(x(t), \lambda_{\mathcal{X}}(t)) \right] \\ &= \exp \left[ -\frac{j_{\mathcal{X}}(t)}{T_{\mathcal{X}}} dt \right], \end{aligned} \quad (23)$$

のように、確率的な熱浴のエントロピー変化に相当する項  $ds_{\text{bath}}(t) := -j_{\mathcal{X}}(t)dt/T_{\mathcal{X}}$  が出てくる。この事実は一般に (詳細) ゆらぎの定理とよばれる <sup>\*16</sup>。

さて、このゆらぎの定理から熱力学第二法則 (6) を導くことにしよう。まず、次のような相対エントロピーの非負性を考える。

$$D(p_{X_{t+dt}|X_t} p_{X_t} || p_{B_{X_t}|X_{t+dt}} p_{X_{t+dt}}) \geq 0, \quad (24)$$

ここに、ゆらぎの定理 (23) を用いることで、

$$D(p_{X_{t+dt}|X_t} p_{X_t} || p_{B_{X_t}|X_{t+dt}} p_{X_{t+dt}}) = -\frac{\langle j_{\mathcal{X}}(t) \rangle}{T_{\mathcal{X}}} dt + H(X_{t+dt}) - H(X_t), \quad (25)$$

が導ける <sup>\*17</sup>。よって、式 (24)、式 (25) から熱力学第二法則 (6) が証明された。

<sup>\*15</sup> この導出は、主にゆらぎの定理 [11] や Jarzynski 等式 [12] と呼ばれている関係式から熱力学第二法則を導出する方法と等価である。本節での議論の流れに近い論文としては、例えば [13] がある。

<sup>\*16</sup> 二つの遷移確率の比  $p/p_B$  を取った時に熱浴のエントロピー変化が出現するのは、Langevin 方程式に限った話ではない。Master 方程式で記述される多くの妥当な熱力学的なモデルで、熱浴のエントロピー変化が出現するのである。よって、この式は非常に普遍的な関係式としてゆらぎの定理という名前が付いており、様々な設定で導出がなされている (例えば Refs. [11, 12, 14, 15, 16])。

<sup>\*17</sup> ただし、ここでアンサンブル平均  $\langle \dots \rangle$  は  $\langle \dots \rangle := \int dx_t dx_{t+dt} p(x_{t+dt}|x_t) p(x_t) \dots$  で定義される。

## 5. 情報熱力学入門

ここから情報理論における「情報」, すなわち相互情報量のような量が, ゆらぎの熱力学においていかに出現していくかを考えていこう<sup>\*18</sup>. もっとも単純な場合として, 次のように外部のコントロールパラメータ  $\lambda_{\mathcal{X}}$  が, 系  $\mathcal{M}$  の自由度  $m$  に依存するケースを考えていこう.

$$\dot{x}(t) = -\partial_x U_{\mathcal{X}}(x(t), \lambda_{\mathcal{X}}(m, t)) + \sqrt{2T_{\mathcal{X}}} \xi_{\mathcal{X}}(t). \quad (26)$$

この時, 熱力学第二法則が式 (6) と同じ形で成り立つか, を考えてみたい.

そのために前節の同様の導出を再度考え直してみよう. 詳細ゆらぎの定理を考え直すと, 以下のように  $m$  を条件付けた遷移確率の比が, 熱浴のエントロピー変化量と結びつくことがわかる.

$$\begin{aligned} \frac{p_{X_t|X_{t+dt}, M}(x_{t+dt}|x_t, m)}{p_{B_{X_{t+dt}|X_t, M}}(x_t|x_{t+dt}, m)} &= \exp \left[ -\frac{1}{T_{\mathcal{X}}} \dot{x}(t) dt \circ \partial_x U_{\mathcal{X}}(x(t), \lambda_{\mathcal{X}}(m, t)) \right] \\ &= \exp [ds_{\text{bath}}(t)]. \end{aligned} \quad (27)$$

ただし,  $p_{X_{t+dt}|X_t, M}(x_{t+dt}|x_t, m) = \mathcal{T}_m(x_{t+dt}; x_t)$  としたとき, 時間逆方向に定義した遷移確率は  $p_{B_{X_t|X_{t+dt}, M}}(x_t|x_{t+dt}, m) = \mathcal{T}_m(x_t; x_{t+dt})$  で定義されるとする. この量を用いて, 前節のように相対エントロピーを考えるとどうなるだろうか.

例えば次のような相対エントロピーの非負性を考えることができる.

$$D(p_{X_{t+dt}|X_t, M} p_{X_t, M} || p_{B_{X_t|X_{t+dt}, M}} p_{X_{t+dt}, M}) \geq 0, \quad (28)$$

ここに, ゆらぎの定理 (27) を用いることで,

$$\begin{aligned} D(p_{X_{t+dt}|X_t, M} p_{X_t, M} || p_{B_{X_t|X_{t+dt}, M}} p_{X_{t+dt}, M}) \\ = \langle ds_{\text{bath}}(t) \rangle + H(X_{t+dt}, M) - H(X_t, M), \end{aligned} \quad (29)$$

となる. すなわち,

$$dS_{\text{bath}}(t) + H(X_{t+dt}, M) - H(X_t, M) \geq 0, \quad (30)$$

がこの場合に成り立つべき熱力学第二法則である. これを式 (9) からのずれを明示した形で書くと, 次のような表記を得る.

$$dS_{\text{bath}}(t) + dS_{\mathcal{X}}(t) \geq I(X_{t+dt}; M) - I(X_t; M), \quad (31)$$

これがいわゆる「情報熱力学第二法則」と呼ばれるものである. 系  $\mathcal{M}$  が存在することによる熱力学的な補正項が, 相互情報量の形で右辺にでてきている. この項  $dI_{\mathcal{X}; M}(t) := I(X_{t+dt}; M) - I(X_t; M)$  は, 一般には負の値を取り得るため, 系  $\mathcal{M}$  の存在によって式 (9) の制限を破りうることを意味する<sup>\*19</sup>.

しかし, 重要なのは, この式 (31) を持って熱力学第二法則自体が破れていることを意味しない. というより, 系  $\mathcal{M}$  が存在するときの正しい熱力学第二法則はこの式 (31) で与えられると考えたほうが自然である. 補正項  $dI_{\mathcal{X}; M}(t)$  は系  $\mathcal{M}$  と系  $\mathcal{X}$  の間の相互情報量の変化で表されるため, もし系  $\mathcal{M}$  と系  $\mathcal{X}$  が独立であったり, もしくは系  $\mathcal{M}$  と系  $\mathcal{X}$  が十分に独立だと見なせるような状況<sup>\*20</sup>では無視できる項であり, 式 (9) と同じ形の式に

<sup>\*18</sup> 本節での議論の流れに相当する論文は例えば [17, 18, 19] である. 例えば, [17] の論文では  $m$  はメモリ自由度としたときのフィードバック制御を考えている.

<sup>\*19</sup> この  $dI_{\mathcal{X}; M}(t)$  が負の値を取り得ることから, マクスウェルのデーモンにおける「一見して熱力学第二法則を破るような効果」を理解することが可能である.

<sup>\*20</sup> マクロな系の熱力学は, この状況に対応すると考えてもよい. しかしながら, 現象論であるマクロな系の熱力学と, 確率過程論であるゆらぎの熱力学を正しく比較することは一般には難しい.

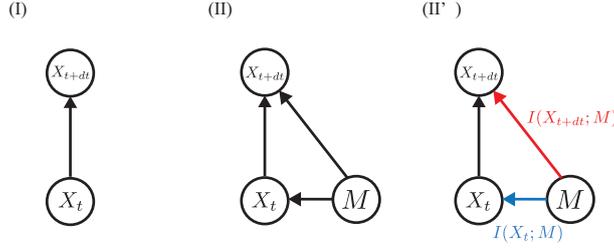


図1 ベイジアンネットワークによる図示. (I) 式 (32) を表現するベイジアンネットワーク. (II) 式 (33) を表現するベイジアンネットワーク. (II') ベイジアンネットワークと情報項の対応関係の概念図.

なる. しかし, ゆらぎが支配的になるメゾスコピックな領域では, この補正項  $dI_{X;M}(t)$  が  $dS_{\text{bath}}(t) + dS_{\mathcal{X}}(t)$  と同等の寄与を与える.

このようにゆらぎの熱力学においては, 情報理論の量である相互情報量項  $dI_{X;M}(t)$  と, 熱力学量として導入された  $dS_{\text{bath}}(t)$  が, 式 (31) を介してトレードオフ関係のある同種の物理量として扱うことができる. よって, メゾスコピック系においては情報通信の熱力学的性質を考察する「情報熱力学」という分野が成立するのである.

## 6. ネットワーク上の情報熱力学入門

これから情報熱力学第二法則を, 複雑な状況でも成立するように一般化していくことにしよう\*<sup>21</sup>. まず, 熱力学第二法則 (6) と情報熱力学第二法則 (31) の違いの起源は, 系  $\mathcal{M}$  が存在するかどうかである. (I) 熱力学第二法則 (6) が成り立つ設定と, (II) 情報熱力学第二法則 (31) が成り立つ設定を端的に記述すると次のようになる.

(I) 確率分布  $p_{X_t}(x_t)$  で分布している系  $\mathcal{X}$  の状態  $x_t$  から, 遷移確率  $p_{X_{t+dt}|X_t}(x_{t+dt}|x_t) = \mathcal{T}(x_{t+dt}; x_t)$  で  $x_{t+dt}$  へ遷移する.

(II) 系  $\mathcal{M}$  の状態  $m$  が確率分布  $p_M(m)$  で分布しているとする. 系  $\mathcal{M}$  の状態  $m$  の下での確率分布  $p_{X_t|M}(x_t|m)$  で分布している系  $\mathcal{X}$  の状態  $x_t$  から, 遷移確率  $p_{X_{t+dt}|X_t, M}(x_{t+dt}|x_t, m) = \mathcal{T}_m(x_{t+dt}; x_t)$  で  $x_{t+dt}$  へ遷移する.

(I) と (II) について数式を用いて端的に表現すると,

$$p_{X_t, X_{t+dt}}(x_t, x_{t+dt}) = p_{X_t}(x_t)p_{X_{t+dt}|X_t}(x_{t+dt}|x_t), \quad (32)$$

$$p_{M, X_t, X_{t+dt}}(m, x_t, x_{t+dt}) = p_M(m)p_{X_t|M}(x_t|m)p_{X_{t+dt}|X_t, M}(x_{t+dt}|x_t, m), \quad (33)$$

という同時確率分布の分解の仕方 \*<sup>22</sup>に行き着く. この分解の仕方を直感的に理解するために, 図1のような図示を導入しよう. これはベイジアンネットワーク \*<sup>23</sup>と呼ばれるものであり, 同時確率分布がいかに分解されるかを表現したものになる.

以下, ベイジアンネットワークの表記を説明しよう. 例えば, 図1(I)に注目してみよう. 丸で記述されている各ノードには, それぞれ確率変数が対応しており, この場合は  $X_t$  と  $X_{t+dt}$  の二つがある. この間の有向なエッジ (矢印) が二つの確率変数の関係を表現しており, 矢印の向きが確率変数間の条件付けを表現する. この場合,  $X_t$  に伸びる矢印は存在しないので,  $X_t$  に関する分布は条件付けなしの  $p_{X_t}(x_t)$  で記述され,  $X_{t+dt}$  に伸びる矢印は  $X_t$  があるため,  $X_{t+dt}$  に関する分布は  $X_t$  の条件の下での分布  $p_{X_{t+dt}|X_t}(x_{t+dt}|x_t)$  で記述される. よって, 図1(I) は同時確率分布の分解の仕方  $p_{X_t}(x_t)p_{X_{t+dt}|X_t}(x_{t+dt}|x_t)$  を表す. 同様に, 図1(II) は  $p_M(m)p_{X_t|M}(x_t|m)p_{X_{t+dt}|X_t, M}(x_{t+dt}|x_t, m)$  を表す.

\*<sup>21</sup> 本節に相当する内容の論文は [20] になる. また, より詳しい内容が本の一章分として出版されている [21].

\*<sup>22</sup> いわゆる確率分布の chain rule. 式 (3) の微分に関する chain rule とは異なる意味で使っているので注意.

\*<sup>23</sup> ベイジアンネットワークについては例えば次の教科書 [22, 23] を参照. ベイジアンネットワークは有向非循環グラフで記述され, 各ノード間の順序を導入できる (トポロジー順序). この順序を時間方向に取るベイジアンネットワークを作ることが可能であり, 以降の議論ではそのようなベイジアンネットワークを考察する.

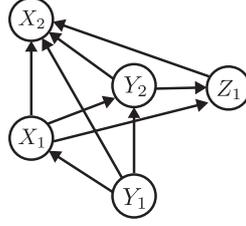


図2 複雑な相互作用を表現するベイジアンネットワークの一例.

さて、ここで (I) で成り立つ熱力学第二法則 (6) と (II) で成り立つ情報熱力学第二法則 (31) の違いを考えてみよう. この二つの不等式の違いは、情報の項  $I(X_{t+dt}; M) - I(X_t; M)$  である. 一方、二つのベイジアンネットワーク (I) と (II) の違いはノード  $M$  の存在と、 $M$  から  $X_t$  と  $X_{t+dt}$  に伸びる二つのエッジである. よって、情報熱力学第二法則とベイジアンネットワークによる設定の記述になんらかの対応関係があるのではないか、という素朴な期待を持てるだろう (図 1(II')).

また、式 (6) と式 (31) の導出を思い出すと、導出に用いた道具は相対エントロピーの非負性とゆらぎの定理だけである. 詳細ゆらぎの定理は同時確率分布における遷移確率  $\mathcal{T}(\mathcal{T}_m)$  の部分から導入されており、また相対エントロピーの定義そのものも同時確率分布によって与えられる. よって、導出の差は同時確率分布に関する分解の差、すなわちベイジアンネットワークの差に対応しているといえる.

これらの事実は、いかに複雑な相互作用のある確率過程であっても、一般に情報熱力学第二法則を与えるベイジアンネットワーク上の理論が構成可能である、ということを示唆する. 実際、ベイジアンネットワークのトポロジーと情報熱力学第二法則における情報の項が一對一に対応する理論を構築することができる [20]. この理論を我々はネットワーク上の情報熱力学とよんでいる.

本節では一般のネットワークで情報熱力学第二法則を導出することはせず、複雑な設定の具体例を用いて情報熱力学第二法則を導出することで、ネットワーク上の情報熱力学の考え方を紹介することにしよう. 例えば図 2 のようなベイジアンネットワークを考えてみたい. このベイジアンネットワークは系  $\mathcal{X}$ ,  $\mathcal{Y}$ ,  $\mathcal{Z}$  がお互いに複雑に相互作用する時間発展を記述しており、その時間発展を記述する同時確率分解の仕方は以下のようになる.

$$p_{X_1, X_2, Y_1, Y_2, Z_1}(x_1, x_2, y_1, y_2, z_1) = p_{Y_1}(y_1)p_{X_1|Y_1}(x_1|y_1)p_{Y_2|X_1, Y_1}(y_2|x_1, y_1)p_{Z_1|X_1, Y_2}(z_1|x_1, y_2)p_{X_2|X_1, Y_1, Y_2, Z_1}(x_2|x_1, y_1, y_2, z_1), \quad (34)$$

ここで  $x_i$  は系  $\mathcal{X}$  の時間順序的に  $i$  番目の状態を意味する ( $x_1$  は初期状態). ただし、異なる系の変数間の添え字の関係は時間順序を考慮していない. すなわち  $x_i$  と  $y_j$  を対して、 $i < j$  であっても  $x_i$  が  $y_j$  よりも前の出来事であることは意味しない. 例えば、この例だと  $z_1$  は  $y_2$  よりも後の出来事に相当する.

ここで、遷移確率  $p_{X_2|X_1, Y_1, Y_2, Z_1}(x_2|x_1, y_1, y_2, z_1) =: \mathcal{T}_{y_1, y_2, z_1}(x_2; x_1)$  を用いて、時間逆方向に定義した遷移確率を  $p_{B X_1|X_2, Y_1, Y_2, Z_1}(x_1|x_2, y_1, y_2, z_1) := \mathcal{T}_{y_1, y_2, z_1}(x_1; x_2)$  とすると、ゆらぎの定理を仮定することで確率的な熱浴のエントロピーの変化量は  $\exp(ds_{\text{bath}}) := \mathcal{T}_{y_1, y_2, z_1}(x_2; x_1)/\mathcal{T}_{y_1, y_2, z_1}(x_1; x_2)$  で与えられる. そこで、次のような相対エントロピーを考える.

$$D(p_{X_1, X_2, Y_1, Y_2, Z_1} \| p_{B X_1|X_2, Y_1, Y_2, Z_1} p_{X_2, Y_1, Y_2, Z_1}) \geq 0. \quad (35)$$

この相対エントロピーを式 (34) を用いて書き直すと、次のような式が得られる.

$$D(p_{X_1, X_2, Y_1, Y_2, Z_1} \| p_{B X_1|X_2, Y_1, Y_2, Z_1} p_{X_2, Y_1, Y_2, Z_1}) = \langle ds_{\text{bath}} \rangle + H(X_2) - H(X_1) + I(X_1; Y_1) + I(X_1; Y_2|Y_1) + I(X_1; Z_1|Y_1, Y_2) - I(X_2; \{Z_2, Y_2, Y_1\}). \quad (36)$$

ただし、ここで式 (34) から得られる関係  $p_{Z_1|X_1, Y_2}(z_1|x_1, y_2) = p_{Z_1|X_1, Y_1, Y_2}(z_1|x_1, y_1, y_2)$  を用いた.

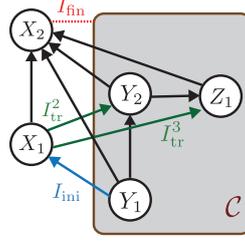


図3 ベイジアンネットワークと情報項の対応関係.  $I_{\text{ini}}$  が  $x_1$  に伸びるエッジに,  $I_{\text{tr}}^l$  が  $X_1$  から  $C$  に伸びるエッジに,  $I_{\text{fin}}$  が  $X_2$  にエッジが伸びている  $C$  の集合に対応している.

式 (36) において得られた各項の意味を考えていこう. まず, これまでと同様に  $dS_{\text{bath}} = \langle ds_{\text{bath}} \rangle$  は熱浴のエントロピー変化量,  $dS_{\mathcal{X}} := H(X_2) - H(X_1)$  は系  $\mathcal{X}$  のエントロピー変化量になる. では, 次に相互情報量の項を見ていこう. まず  $I_{\text{ini}} := I(X_1; Y_1)$  は, 系  $\mathcal{X}$  が初期状態で他の系  $\mathcal{Y}$  と持っていた相関を意味する項になる.  $I_{\text{fin}} := I(X_2; \{Z_2, Y_2, Y_1\})$  は系  $\mathcal{X}$  が終状態で他の系  $\mathcal{Y}$ ,  $\mathcal{Z}$  と持っていた相関を意味する項になっている.

では, 残りの条件付き相互情報量の項を考えていこう. まず移動エントロピー (transfer entropy) と呼ばれる概念 [24] を導入する. 系  $\mathcal{X}$  のある状態  $x_k$  と, 別の系  $\mathcal{C}$  の時系列  $c_1, \dots, c_n$  があったときに, 系  $\mathcal{X}$  から系  $\mathcal{C}$  への移動エントロピーは

$$I_{\text{tr}}^l = I(X_k; C_l | C_{l-1}, \dots, C_1), \quad (37)$$

で定義される量である. この量は  $I_{\text{tr}}^l = I(X_k; \{C_l, \dots, C_1\}) - I(X_k; \{C_{l-1}, \dots, C_1\})$  であるため, 確率変数  $X_k$  と系  $\mathcal{X}$  の時系列の確率変数の組の間の相互情報量が, 時刻  $l-1$  から  $l$  になったときにどれだけ増加したかを表している. すなわち, 系  $\mathcal{C}$  の時刻  $l-1$  から  $l$  の時間発展の間に確率変数  $X_k$  からどれだけ情報が流れたかを定量化する情報の流れの量になっている.

この移動エントロピーの定義を用いて, 残りの条件付き相互情報量の項を考えると次のようになる. いま, 他の系  $\mathcal{Y}$ ,  $\mathcal{Z}$  を纏めた系  $\mathcal{C}$  の時系列  $\{y_1, y_2, z_1\}$  を考慮すると, 条件付き相互情報量の項は  $I_{\text{tr}}^2 := I(X_1; Y_2 | Y_1)$ ,  $I_{\text{tr}}^3 := I(X_1; Z_2 | Y_2, Y_1)$  のように移動エントロピーで与えられる系  $\mathcal{X}$  から他の系  $\mathcal{C}$  に流れた情報の流れだともみなせる.

よって, 式 (36) の非負性は次のように書き直すことができる.

$$dS_{\text{bath}} + dS_{\mathcal{X}} \geq I_{\text{fin}} - I_{\text{ini}} - (I_{\text{tr}}^2 + I_{\text{tr}}^3), \quad (38)$$

この式が, ネットワーク上の情報熱力学における情報熱力学第二法則になる. この式の情報項は, やはりベイジアンネットワークにおけるトポロジーと対応関係を持っている (図3). この関係式は, 今回具体的な図2のベイジアンネットワークに対して成り立つ関係式であったが,

$$dS_{\text{bath}} + dS_{\mathcal{X}} \geq I_{\text{fin}} - I_{\text{ini}} - \sum_l I_{\text{tr}}^l, \quad (39)$$

という関係そのものは, 一般の \*24 ベイジアンネットワーク上で常に成立する [20]. ただし一般の場合, 各情報項  $I_{\text{fin}}$ ,  $I_{\text{ini}}$ ,  $I_{\text{tr}}^l$  はグラフのトポロジーから計算する量になっている.

\*24 この議論はゆらぎの定理を仮定として用いているため, この仮定が成り立つ設定を記述するベイジアンネットワークに限られる. 系  $\mathcal{X}$  の時間発展が他の系  $\mathcal{C}$  を介さない形で非マルコフである ( $x_1 \rightarrow x_2 \rightarrow x_3$  かつ  $x_1 \rightarrow x_3$  のようなエッジが存在する) ときは, ゆらぎの定理を単純適用できないため成り立たない.

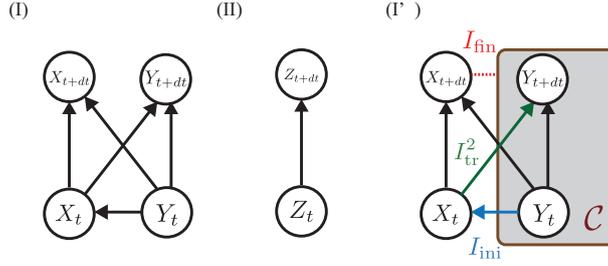


図4 二次元 Langevin 系における二つのベイジアンネットワーク. (I) 式 (42) に対応したベイジアンネットワーク. (II) 式 (43) に対応したベイジアンネットワーク. (I') (I) のベイジアンネットワークから得られる情報熱力学第二法則の, 情報項とベイジアンネットワークの対応関係.

## 7. 情報熱力学第二法則と熱力学第二法則

ネットワーク上の情報熱力学が与える結果は, 通常の熱力学第二法則とはどのように違うのだろうか<sup>\*25</sup>. この問いに答えるため, 次のような二次元の Langevin 方程式を例に考えてみよう.

$$\begin{aligned}\dot{x}(t) &= -\partial_x U_{\mathcal{X}}(x(t), y(t)) + \sqrt{2T_{\mathcal{X}}}\xi_{\mathcal{X}}(t), \\ \dot{y}(t) &= -\partial_y U_{\mathcal{Y}}(x(t), y(t)) + \sqrt{2T_{\mathcal{Y}}}\xi_{\mathcal{Y}}(t).\end{aligned}\quad (40)$$

ここで  $\xi_{\mathcal{X}}(t)$  と  $\xi_{\mathcal{Y}}(t)$  は独立な分散 1 平均 0 の白色 Gauss ノイズ,  $T_{\mathcal{Y}}$  と  $T_{\mathcal{X}}$  は系  $\mathcal{X}$  と系  $\mathcal{Y}$  の熱浴の温度,  $U_{\mathcal{X}}$  と  $U_{\mathcal{Y}}$  はそれぞれ系  $\mathcal{X}$  と系  $\mathcal{Y}$  のポテンシャルエネルギーで, 必ずしも  $U_{\mathcal{X}} = U_{\mathcal{Y}}$  を要請しない.  $\dot{x}(t)$ ,  $\dot{y}(t)$  は粒子の速度を,  $x(t)$ ,  $y(t)$  は粒子の位置を意味する. これを微小時間  $dt$  で離散化した式は次のようになる.

$$\begin{aligned}x_{t+dt} - x_t &= -\partial_x U_{\mathcal{X}}(x_t, y_t)dt + \sqrt{2T_{\mathcal{X}}}dB_{\mathcal{X}t}, \\ y_{t+dt} - y_t &= -\partial_y U_{\mathcal{Y}}(x_t, y_t)dt + \sqrt{2T_{\mathcal{Y}}}dB_{\mathcal{Y}t}.\end{aligned}\quad (41)$$

ここで,  $x_t := x(t)$ ,  $y_t := y(t)$  であり,  $dB_{\mathcal{X}t}$  と  $dB_{\mathcal{Y}t}$  は互いに独立な同分布に従い, その分布は平均 0 分散  $dt$  の Gauss 分布である.

それぞれの遷移確率を  $p_{X_{t+dt}|X_t, Y_t}(x_{t+dt}|x_t, y_t) = \mathcal{T}_{y_t}^{\mathcal{X}}(x_{t+dt}; x_t)$ ,  $p_{Y_{t+dt}|X_t, Y_t}(y_{t+dt}|x_t, y_t) = \mathcal{T}_{x_t}^{\mathcal{Y}}(y_{t+dt}; y_t)$  とした時に, ゆらぎの定理はそれぞれ  $\exp(-j_{\mathcal{X}}(t)dt/T_{\mathcal{X}}) = \mathcal{T}_{y_t}^{\mathcal{X}}(x_{t+dt}; x_t)/\mathcal{T}_{x_t}^{\mathcal{X}}(x_t; x_{t+dt})$ ,  $\exp(-j_{\mathcal{Y}}(t)dt/T_{\mathcal{Y}}) = \mathcal{T}_{x_t}^{\mathcal{Y}}(y_{t+dt}; y_t)/\mathcal{T}_{y_t}^{\mathcal{Y}}(y_t; y_{t+dt})$  で与えられる.<sup>\*26</sup>

この系をベイジアンネットワークで表記するとき, 同時確率分布の分解の仕方

$$\begin{aligned}p_{X_t, Y_t, X_{t+dt}, Y_{t+dt}}(x_t, y_t, x_{t+dt}, y_{t+dt}) \\ = p_{Y_t}(y_t)p_{X_t|Y_t}(x_t|y_t)p_{X_{t+dt}|X_t, Y_t}(x_{t+dt}|x_t, y_t)p_{Y_{t+dt}|X_t, Y_t}(y_{t+dt}|x_t, y_t),\end{aligned}\quad (42)$$

に対応した表記 (I) と, 系  $\mathcal{X}$  と系  $\mathcal{Y}$  の合成系  $\mathcal{Z}$  を導入し,  $z_t := \{x_t, y_t\}$  としたときの同時確率分布の分解の仕方

$$p_{Z_t, Z_{t+dt}}(z_t, z_{t+dt}) = p_{Z_t}(z_t)p_{Z_{t+dt}|Z_t}(z_{t+dt}|z_t),\quad (43)$$

に対応した表記 (II) の二種類の表現の仕方がある (図 4). ここで表記 (II) では, 通常の熱力学第二法則が導ける.

$$dS_{\text{bath}} + dS_{\mathcal{Z}} \geq 0.\quad (44)$$

<sup>\*25</sup> 本節での議論と同等の議論が論文 [25] のサブリメントにある. この熱力学第二法則と情報熱力学第二法則の違いに関しては, 論文 [26, 27, 28, 29, 30] など様々な研究がなされている. Langevin 系に特化した先行的な研究としては論文 [31] が挙げられる.

<sup>\*26</sup> ただし, ここで熱流は  $j_{\mathcal{X}}(t) := \dot{x}(t) \circ (\sqrt{2T_{\mathcal{X}}}\xi_{\mathcal{X}}(t) - \dot{x}(t))$ ,  $j_{\mathcal{Y}}(t) := \dot{y}(t) \circ (\sqrt{2T_{\mathcal{Y}}}\xi_{\mathcal{Y}}(t) - \dot{y}(t))$  と定義される.

ここで  $dS_Z = H(X_{t+dt}, Y_{t+dt}) - H(X_t, Y_t)$  である。また遷移確率は  $p(z_{t+dt}|z_t) := \mathcal{T}(z_{t+dt}; z_t) = T_{y_t}^{\mathcal{X}}(x_{t+dt}; x_t)T_{x_t}^{\mathcal{Y}}(y_{t+dt}; y_t)$  とかけることより、ゆらぎの定理は  $\exp(-j_{\mathcal{X}}(t)dt/T_{\mathcal{X}} - j_{\mathcal{Y}}(t)dt/T_{\mathcal{Y}}) = \mathcal{T}(z_{t+dt}; z_t)/\mathcal{T}(z_t; z_{t+dt})$  で与えられる。よって  $dS_{\text{bath}} = -\langle j_{\mathcal{X}}(t) \rangle dt/T_{\mathcal{X}} - \langle j_{\mathcal{Y}}(t) \rangle dt/T_{\mathcal{Y}}$  となる。以上より、通常の熱力学第二法則は

$$\frac{\langle j_{\mathcal{X}}(t) \rangle}{T_{\mathcal{X}}} dt \leq -\frac{\langle j_{\mathcal{Y}}(t) \rangle}{T_{\mathcal{Y}}} dt + H(X_{t+dt}, Y_{t+dt}) - H(X_t, Y_t), \quad (45)$$

で与えられる。一方表記 (I) でのベイジアンネットワークにおける情報熱力学第二法則は、

$$dS_{\text{bath}} + dS_{\mathcal{X}} \geq I_{\text{fin}} - I_{\text{ini}} - I_{\text{tr}}^2, \quad (46)$$

であり、それぞれ  $dS_{\text{bath}} = -\langle j_{\mathcal{X}}(t) \rangle dt/T_{\mathcal{X}}$ ,  $dS_{\mathcal{X}} := H(X_{t+dt}) - H(X_t)$ ,  $I_{\text{fin}} := I(X_{t+dt}; \{Y_t, Y_{t+dt}\})$ ,  $I_{\text{ini}} := I(X_t; Y_t)$ ,  $I_{\text{tr}}^2 := I(X_t; Y_{t+dt}|Y_t)$  で定義される (図 4 (I) も参照)。よって、

$$\begin{aligned} \frac{\langle j_{\mathcal{X}}(t) \rangle}{T_{\mathcal{X}}} dt &\leq H(X_{t+dt}) - H(X_t) + I(X_t; Y_t) + I(X_t; Y_{t+dt}|Y_t) - I(X_{t+dt}; \{Y_t, Y_{t+dt}\}) \\ &= H(X_{t+dt}, Y_t, Y_{t+dt}) - H(X_t, Y_t, Y_{t+dt}), \end{aligned} \quad (47)$$

となる。すなわちこれらの事実から、熱力学第二法則と情報熱力学第二法則は同時に成り立つ異なるものであるということがわかる。

さらに式 (45) と式 (47) をみると、系  $\mathcal{X}$  の熱流の項  $\langle j_{\mathcal{X}}(t) \rangle dt/T_{\mathcal{X}}$  に対してそれぞれ異なる上限を与えることがわかる。通常の熱力学第二法則 (45) は系  $\mathcal{Y}$  側の熱流の項  $-\langle j_{\mathcal{Y}}(t) \rangle dt/T_{\mathcal{Y}}$  を考慮しなければいけないのに対し、情報熱力学第二法則 (47) では系  $\mathcal{Y}$  側の熱流の項を考える必要がなくなっている。また、実際に情報熱力学第二法則 (47) の方が熱力学第二法則 (45) に比べ、熱流の項  $\langle j_{\mathcal{X}}(t) \rangle dt/T_{\mathcal{X}}$  に対してタイトな制限を与えられることがわかっている。

以上の事実は、複雑に絡み合った系の熱力学を考察する上で情報熱力学の利点になりうる。もしある特定の箇所の熱散逸に対して熱力学第二法則から制限を出そうとするならば、複雑に絡み合った全ての系の熱散逸を数え上げる必要がある。一方で、情報熱力学第二法則は特定の箇所と相互作用しているところの情報の項さえわかれば<sup>\*27</sup>、絡み合った系すべてを考える必要がないということを主張している。このような情報熱力学の利点は、複雑に絡み合った生体系で有用たりうる。

## 8. 生体センサーと適応現象

この節ではネットワーク上の情報熱力学を、生体シグナル伝達系に応用した研究 [25] を紹介しよう。今回考えるのは、大腸菌の走化性のシグナル伝達である。

単細胞生物である大腸菌 (*Escherichia coli* もしくは *E. coli*) には、入力信号となる餌 (リガンド<sup>\*28</sup>) の濃度勾配を受容体が検知して、濃度の濃い方に進もうとする走化性という現象がある。リガンドの濃度変化の入力情報は受容体を通して細胞表面から内部に伝わり、出力として内部のキナーゼ<sup>\*29</sup>の活性を変化させる。その結果として、鞭毛モーターが推進モード<sup>\*30</sup>に切り替わり、大腸菌は媒質の中を泳ぐことができるようになる (図 5(I)) [32].

<sup>\*27</sup> 例えば、系  $\mathcal{X}$  と系  $\mathcal{Y}$  が相互作用し、系  $\mathcal{Y}$  と系  $\mathcal{Z}$  が相互作用する系を考えると、熱力学第二法則では系  $\mathcal{Z}$  のダイナミクス (熱散逸) を考える必要がある。一方、情報熱力学は系  $\mathcal{X}$  と系  $\mathcal{Y}$  の確率分布さえ知っていれば系  $\mathcal{Z}$  は陽には考える必要はない。一般には系  $\mathcal{X}$  と系  $\mathcal{Y}$  の確率分布を得るために系  $\mathcal{Z}$  を考える必要が出てくるが、例えば実験などで測定から確率分布を得る場合は系  $\mathcal{Z}$  の測定を必要はなくなる。

<sup>\*28</sup> 特定の受容体に特異的に結合する物質の総称である。大腸菌の場合、例えばアスパラギン酸受容体 (Tar) があるため、アスパラギン酸がリガンドに相当する。

<sup>\*29</sup> リン酸化酵素の総称である。

<sup>\*30</sup> 鞭毛モーターが反時計回りに回ると推進モードに、時計回りに回るとランダムな方向転換モード (タンブル) になる。この二つの状態の切り替わる比率を図 5(I) におけるタンパク質 CheY のリン酸化が決めている。

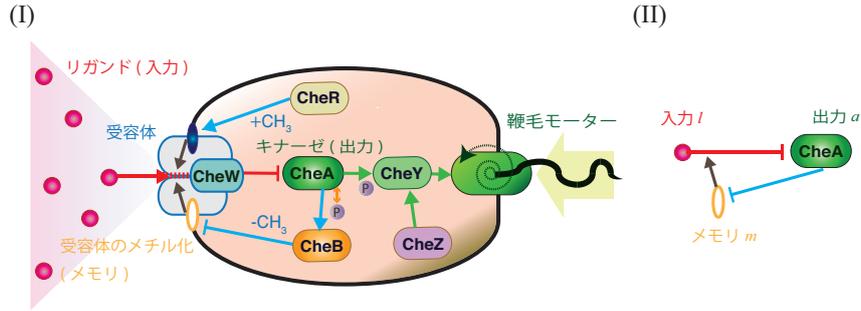


図5 (I) 大腸菌走化性のシグナル伝達ネットワーク. CheA ~ CheZ は大腸菌の走化性に関わるタンパク質である. キナーゼはここで CheA に相当し, これが活性化することで CheY にリン酸基に転移させる役割をもっている. CheY がリン酸化すると, 鞭毛モーターの回転を推進モードに変換することができる. (II) 適応現象における重要な三つの要素. キナーゼ活性の出力と受容体のメチル化のメモリの間にはフィードバックループが存在しており, このフィードバックループにより適応現象を達成することができる.

この大腸菌の走化性の特徴の一つに適応とよばれる現象がある. これは生体センサーである受容体がきちんと濃度の「変化」を感知するために, 現在の濃度に慣れるという現象である. この適応を達成するためには, どこかに現在の濃度の状態を記録するメモリの存在が必要であり, 大腸菌の走化性の場合には受容体のメチル化がメモリに相当する. この受容体のメチル化のおかげで, 濃度が変化しないときはキナーゼ活性を変化させず, 濃度が変化したときのみ一過的にキナーゼ活性を変化させることができるのである (図5(I)).

この適応現象は生体センサーにおいて普遍的にみられる構造であり, 様々な適応現象は入力, 出力, メモリの三つの要素によって構成されることが知られている (図5(II)). この入力と出力, メモリの役割を担っているタンパク質やその状態は系によって様々であるが, この三つの要素の構造自体は適応現象において普遍的なものである. よって単細胞生物<sup>\*31</sup>である大腸菌の走化性を考え, 入力, 出力, メモリの役割に着目するのは生体シグナル伝達を理解する上で自然な発想であろう.

特に (大腸菌走化性の) 適応現象は, 次のような Langevin 方程式で記述されることが知られている [32, 33, 34],

$$\begin{aligned} \dot{a}(t) &= -\frac{1}{\tau_A} [a(t) - \bar{a}(l(t), m(t))] + \sqrt{2T_A} \xi_A(t), \\ \dot{m}(t) &= -\frac{1}{\tau_M} a(t) + \sqrt{2T_M} \xi_M(t), \end{aligned} \quad (48)$$

ここで,  $a(t)$  はキナーゼの活性レベル,  $m(t)$  は受容体のメチル化レベル,  $l(t)$  はリガンドの入力,  $\tau_A, \tau_M$  はそれぞれキナーゼ活性の出力系  $\mathcal{A}$  と受容体メチル化のメモリ系  $\mathcal{M}$  の時定数で  $\tau_M \gg \tau_A$  を満たす,  $T_A, T_M$  はそれぞれ系  $\mathcal{A}$  と  $\mathcal{M}$  のノイズの強度<sup>\*32</sup>,  $\xi_A(t), \xi_M(t)$  はそれぞれ独立な分散 1, 平均 0 の白色 Gauss ノイズである.  $\bar{a}(l(t), m(t))$  は各  $l(t), m(t)$  の値に応じたキナーゼ活性レベル  $a(t)$  の定常的な値であり, たとえば大腸菌の走化性では  $\bar{a}(l(t), m(t)) = \alpha m(t) - \beta l(t)$  のように線形領域でかける領域が比較的広いことが知られている. ここで  $\alpha, \beta$  は正の係数である.

この Langevin 方程式の典型的な振る舞いを見てみよう. 例えばリガンドの入力として階段関数的なもの考えると図6のような振る舞いになる. このように, 瞬間的な入力  $l(t)$  の変化に対して, 出力  $a(t)$  は一過的に変化しあと元のレベルに戻る. 最終的に現在のリガンドの状態はメモリ  $m(t)$  の変化として蓄えられる.

このように二次元 Langevin 方程式によって, 普遍的な生体シグナル伝達の一つである適応現象を記述することがわかった. では, この適応現象のモデルに情報熱力学を応用することを考えてみよう.

<sup>\*31</sup> 単細胞生物である大腸菌は, 生物の構成要素である細胞の個数が最小の生物である. 構成要素が最小の系に着目するのは物理学の手法としては自然だと思われる.

<sup>\*32</sup> 今, 考えている Langevin 方程式は Brown 粒子の位置の運動ではなく化学的な状態のレベルのため, この  $T_A, T_M$  は必ずしも媒質の温度を意味しない量である.

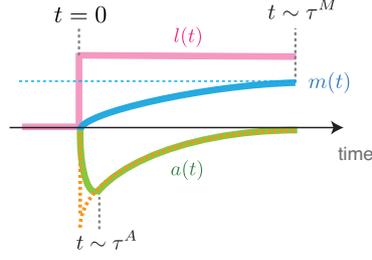


図6 Langevin 方程式 (48) が記述する典型的な適応の振る舞い.

## 9. 生体シグナル伝達の情報熱力学

情報熱力学第二法則 (47) を, 適応現象を記述する Langevin 方程式 (48) に適用することを考えてみよう. このとき情報熱力学第二法則は

$$H(X_{t+dt}) - H(X_t) + I(X_t; Y_t) + I(X_t; Y_{t+dt}|Y_t) - I(X_{t+dt}; \{Y_t, Y_{t+dt}\}) \geq \frac{\langle j_{\mathcal{A}}(t) \rangle}{T_{\mathcal{A}}} dt \quad (49)$$

で与えられる. 特に定常状態のときは  $p_{X_t, Y_t}(x_t, y_t)$  と  $p_{X_{t+dt}, Y_{t+dt}}(x_{t+dt}, y_{t+dt})$  は同分布のため,  $H(X_{t+dt}) = H(X_t)$ ,  $I(X_t; Y_t) = I(X_{t+dt}; Y_{t+dt})$  となる. よって, 定常状態のとき式 (49) は

$$\frac{\langle j_{\mathcal{A}}(t) \rangle}{T_{\mathcal{A}}} dt \leq I(X_t; Y_{t+dt}|Y_t) - I(X_{t+dt}; Y_t|Y_{t+dt}) \quad (50)$$

$$\leq I(X_t; Y_{t+dt}|Y_t), \quad (51)$$

が得られる. ここで相互情報量の非負性  $I(X_{t+dt}; Y_t|Y_{t+dt}) \geq 0$  を使った<sup>\*33</sup>. このように移動エントロピー  $I_{tr}^2 = I(X_t; Y_{t+dt}|Y_t)$  がある量  $\langle j_{\mathcal{A}}(t) \rangle dt / T_{\mathcal{A}}$  の定常状態における上限になっていることがわかる.

では, この  $\langle j_{\mathcal{A}}(t) \rangle dt / T_{\mathcal{A}}$  という量は, 適応現象において具体的に何を意味しているのだろうか. 実際に計算してみると

$$\begin{aligned} \frac{\langle j_{\mathcal{A}}(t) \rangle}{T_{\mathcal{A}}} dt &= \frac{dt}{T_{\mathcal{A}}} \left\langle \dot{a}(t) \circ \frac{1}{\tau_{\mathcal{A}}} [a(t) - \bar{a}(l(t), m(t))] \right\rangle \\ &= \frac{dt}{T_{\mathcal{A}}} \left\langle \left[ -\frac{1}{\tau_{\mathcal{A}}} [a(t) - \bar{a}(l(t), m(t))] + \sqrt{2T_{\mathcal{A}}} \xi_{\mathcal{A}}(t) \right] \circ \frac{1}{\tau_{\mathcal{A}}} [a(t) - \bar{a}(l(t), m(t))] \right\rangle \\ &= \frac{dt}{(\tau_{\mathcal{A}})^2 T_{\mathcal{A}}} \left[ \tau_{\mathcal{A}} T_{\mathcal{A}} - \langle [a(t) - \bar{a}(l(t), m(t))]^2 \rangle \right], \end{aligned} \quad (52)$$

となる. ここで任意の関数  $g(a(t))$  に対し  $\langle \xi_{\mathcal{A}}(t) \circ g(a(t)) \rangle = \sqrt{T_{\mathcal{A}}/2} \langle \partial_a g(a(t)) \rangle$  であることを用いた<sup>\*34</sup>.

ここで  $\tau_{\mathcal{A}} T_{\mathcal{A}}$  は環境のノイズの強度に相当する項である. また  $\bar{a}(l(t), m(t))$  を与えられたリガンド入力  $l(t)$  に対する適切な出力  $a(t)$  の値だと考えると  $\langle [a(t) - \bar{a}(l(t), m(t))]^2 \rangle$  は出力の分散に相当する. よって,  $\frac{\langle j_{\mathcal{A}}(t) \rangle}{T_{\mathcal{A}}} dt$  の項は, 「ノイズのある環境で, いかに正しくリガンド入力に対してキナーゼ活性を出力するか (適応するか)」という適応のロバストさ (頑健さ) の尺度とみなせる.<sup>\*35</sup>

<sup>\*33</sup> この量は我々が backward transfer entropy とよんでいる量である [35]. 移動エントロピー (transfer entropy)  $I(X_t; Y_{t+dt}|Y_t)$  だけの上限 (51) よりも, この量を考えた上限 (50) の方がタイトな上限になっている.

<sup>\*34</sup> この導出には Stratonovich 積分の性質と  $dB_{\mathcal{A}t} := \xi_{\mathcal{A}}(t) dt$  に関する性質を使う [3].  $\langle \xi_{\mathcal{A}}(t) \circ g(a(t)) \rangle = \langle \xi_{\mathcal{A}}(t) [g(a(t)) + g(a(t+dt))] / 2 \rangle = \langle \xi_{\mathcal{A}}(t) [g(a(t)) + \dot{a}(t) dt \partial_a g(a(t))] / 2 \rangle = \sqrt{T_{\mathcal{A}}/2} \langle \xi_{\mathcal{A}}(t) \xi_{\mathcal{A}}(t) dt \partial_a g(a(t)) \rangle$  であり, ここで  $\xi_{\mathcal{A}}(t) \xi_{\mathcal{A}}(t) dt = (dB_{\mathcal{A}t})^2 / dt$ ,  $(dB_{\mathcal{A}t})^2 = dt$ ,  $\langle \xi_{\mathcal{A}}(t) g(a(t)) \rangle = 0$  である性質を使うと,  $\langle \xi_{\mathcal{A}}(t) \circ g(a(t)) \rangle = \sqrt{T_{\mathcal{A}}/2} \langle \partial_a g(a(t)) \rangle$  が示せる.

<sup>\*35</sup> もしコロイド粒子系において同様の表現を得ると, この表現は揺動散逸関係の一般化に相当する [3]. 一般に  $\langle j_{\mathcal{A}}(t) \rangle \neq 0$  の非平衡の場合は,  $\tau_{\mathcal{A}} T_{\mathcal{A}}$  の散逸項と  $\langle [a(t) - \bar{a}(l(t), m(t))]^2 \rangle$  の揺動項が釣り合わない. しかしその差が (驚くべきことに) 熱流  $\langle j_{\mathcal{A}}(t) \rangle$  に対応し,  $\langle j_{\mathcal{A}}(t) \rangle = 0$  の平衡の場合には通常の揺動散逸関係を回復する.

この式 (52) の表現を用いて情報熱力学第二法則 (50), (51) を考察し直そう。不等式の左辺はロバストさであるため、ロバストさの上限が情報の項 (式 (51) の場合は移動エントロピー) によって抑えられていることになる。特に式 (51) の右辺は正であるということは、一般に出力の分散の項  $\langle [a(t) - \bar{a}(l(t), m(t))]^2 \rangle$  が環境のノイズの強度  $\tau_A T_A$  に比べて小さくなりうることを意味する。よって、情報の効果の分だけ環境のノイズの影響を減らせる。<sup>\*36</sup>

このような情報熱力学に基づいて、生体シグナル伝達を議論することでどのようなご利益があるだろうか。一つの答えは、情報熱力学の観点から情報伝達効率を定量化できることにある。すなわち、情報熱力学第二法則 (50), (51) の等号達成にどれだけ近いかをもって、どれだけ実際のロバストさに対して情報を効率よく利用していたかを定量的に議論できる。また同様に通常の熱力学第二法則 (45) の等号達成にどれだけ近いのかと、情報熱力学第二法則 (50), (51) の等号達成にどれだけ近いかを比較することによって、系全体の熱力学的な効率と情報伝達に関わる部分の効率のどちらを重視しているのかも定量化できる。実際我々は、大腸菌の走化性のモデルで現実的な実験パラメータを用いてシミュレーションを行い、通常の熱力学効率に比べて情報熱力学的な効率が高いことが明らかにした<sup>\*37</sup>。このように情報熱力学を生体シグナル伝達に応用することで、生体情報処理の設計原理を理解する上での有用な示唆を与える。

## 10. さらに情報熱力学を学びたい人のために

最後に、情報熱力学に関連する近年の様々な進展を紹介する<sup>\*38</sup>。今後研究したい人のための指針になれば幸いである。

検証実験については、マックスウェルのデーモンの検証という文脈でコロイド粒子系 [37, 38] や単電子箱 [39] で実験的に検証されている。

また、Markov jump 過程に特化した情報熱力学の研究は、様々な進展をみせている。例えば、マックスウェルのデーモンを模した Markov jump 過程を考察する「自律的なマックスウェルのデーモン」という研究が行われた [40, 41, 42]。これらの研究で提唱されたモデルを用いて、Markov jump 過程で情報浴という概念 [40, 43] や情報熱力学第二法則が考察されることになった [26, 27, 28, 29, 30]。またこの Markov jump 過程では、情報熱力学第二法則に対応するオンサーガ相反定理が成り立つことが示された [44]。

情報熱力学の生体系への応用は、主にセンサーにおける情報処理に関する理論的な研究が行われている [45, 46, 47, 48]。また最近では酵素反応や複製過程などの生体系で応用しようとする動きも見られる [49, 50]。

最近では、機械学習方向への応用も考えられ始めてきた [51]。特に移動エントロピー [24] やベイジアンネットワークのような概念は、学習理論やニューラルネットワークといった機械学習分野の理論と相性が良いため [52, 53]、今後もさらなる発展が期待できるだろう。

## 参考文献

- [1] 田崎晴明, 『熱力学—現代的な視点から』 (培風館, 2000).
- [2] 清水明, 『熱力学の基礎』 (東京大学出版会, 2007).
- [3] K. Sekimoto, “Stochastic Energetics” (Springer, 2010); 関本謙, 『ゆらぎのエネルギ論』 (岩波書店, 2004).
- [4] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).
- [5] C. Shannon, Bell System Technical Journal 27, 379-423, 623-656 (1948); 和訳:クロード・シャノン, 『通信の数学的理論』 (ちくま学芸文庫, 2009).

<sup>\*36</sup> この事実は、Shannon の情報理論における通信路符号化定理と類似しており、弱い不等式 (51) についてはアナロジーが成り立つ。

<sup>\*37</sup> この結果の本質は、 $\tau_M \gg \tau_A$  の時定数の差に起因している。この時定数の差により、合成系  $\mathcal{A}, \mathcal{M}$  全体としては平衡から遠く離れているが、部分系  $A$  にとっては局所的には可逆になっているといえる。この結果より、大腸菌の走化性のシグナル伝達は、適応のロバストさを高めるために受容体のメチル化の部分で熱散逸している系と考えることもできる。

<sup>\*38</sup> 情報熱力学に関するレビュー論文 [8] も参照のこと。また量子系のレビューとしては論文 [36] が挙げられる。

- [6] T. M. Cover and J. A. Thomas, “Elements of Information Theory” (Wiley-Interscience, 2nd edition, 2006). 和訳: 『情報理論—基礎と広がり—』 (共立出版, 2012).
- [7] H. S. Leff and A. F. Rex (eds.), “Maxwell’s demon 2: Entropy, Classical and Quantum Information, Computing” (Princeton University Press, 2003).
- [8] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, *Nature physics* **11**, 131 (2015).
- [9] 沙川貴大, 『< 講義ノート > 情報処理の熱力学 (第 59 回物性若手夏の学校: 集中ゼミ)』, *物性研究・電子版* **4**, 1 (2015).
- [10] N. G. van Kampen, “Stochastic processes in physics and chemistry” (North Holland, 3rd edition, 1992).
- [11] Crooks, G. E. *Phys. Rev. E*, **60**, 2721 (1999).
- [12] C. Jarzynski, *J. Stat. Phys.* **98**, 77 (2000).
- [13] R. Kawai, J. M. R. Parrondo and C. Van den Broeck, *Phys. Rev. Lett.* **98**, 080602 (2007).
- [14] D. J. Evans, E. G. D. Cohen and G. P. Morriss, *Phys. Rev. Lett.* **71**, 2401 (1993).
- [15] G. Gallavotti and E. G. D. Cohen, *Phys. Rev. Lett.* **74**, 2694 (1995).
- [16] J. L. Lebowitz and H. Spohn, *J. Stat. Phys.* **95**, 333 (1999).
- [17] T. Sagawa and M. Ueda, *Phys. Rev. Lett.* **104**, 090602 (2010).
- [18] T. Sagawa and M. Ueda, *Phys. Rev. Lett.* **109**, 180602 (2012).
- [19] S. Still, D. A. Sivak, A. J. Bell and G. E. Crooks, *Phys. Rev. Lett.* **109**, 120604 (2012).
- [20] S. Ito and T. Sagawa, *Phys. Rev. Lett.* **111**, 180603 (2013).
- [21] S. Ito and T. Sagawa, arXiv:1506.08519 (2015); これは次の本に収録されている. M. Dehmer, F. Emmert-Streib, Z. Chen, X. Li, and Y. Shi (Eds.). “Mathematical foundations and applications of graph entropy” (Wiley-VCH, 2016 )
- [22] T. D. Nielsen and F. V. Jensen, “Bayesian networks and decision graphs” (Springer, 2009).
- [23] J. Pearl, “Causality”, (Cambridge University Press, 2nd edition, 2009).
- [24] T. Schreiber, *Phys. Rev. Lett.* **85**, 461 (2000).
- [25] S. Ito and T. Sagawa, *Nat. Commun.* **6**, 7498 (2015).
- [26] D. Hartich, A. C. Barato and U. Seifert, *J. Stat. Mech. Theor. Exp.* (2014). P02016.
- [27] J. M. Horowitz and M. Esposito, *Phys. Rev. X* **4**, 031015 (2014).
- [28] J. M. Horowitz and H. Sandberg, *New J. Phys.* **16**, 125007 (2014).
- [29] N. Shiraishi and T. Sagawa, *Phys. Rev. E* **91**, 012130 (2015).
- [30] N. Shiraishi, S. Ito, K. Kawaguchi and T. Sagawa, *New J. Phys.* **17**, 045012 (2015).
- [31] A. E. Allahverdyan, D. Janzing and G. Mahler, *J. Stat. Mech. Theor. Exp.* (2009). P09011.
- [32] Y. Tu, T. S. Shimizu and H. C. Berg, *Proc. Natl. Acad. Sci.* **105**, 14855 (2008).
- [33] F. Tostevin and P. R. Ten Wolde, *Phys. Rev. Lett.* **102**, 218101 (2009).
- [34] G. Lan, P. Sartori, S. Neumann, V. Sourjik and Y. Tu, *Nat. Phys.* **8**, 422 (2012).
- [35] S. Ito, *Sci. Rep.* **6**, 36831 (2016).
- [36] J. Goold, M. Huber, A. Riera, L. del Rio and P. Skrzypczyk, *J. Phys. A: Math. Theor.* **49**, 143001 (2016).
- [37] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki and M. Sano, *Nat. Phys.* **6**, 988-992 (2010).
- [38] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider and E. Lutz, *Nature* **483**, 187 (2012).
- [39] J. V. Koski, V. F. Maisi, T. Sagawa and J. P. Pekola, *Phys. Rev. Lett.* **113**, 030601 (2014).
- [40] D. Mandal and C. Jarzynski, *Proc. Natl. Acad. Sci.* **109**, 11641 (2012).
- [41] P. Strasberg, G. Schaller, T. Brandes and M. Esposito, *Phys. Rev. Lett.* **110**, 040601 (2013).
- [42] A. C. Barato, D. Hartich and U. Seifert, *Phys. Rev. E*, **87**, 042104 (2013).
- [43] A. C. Barato and U. Seifert, *Phys. Rev. Lett.* **112**, 090601 (2014).
- [44] S. Yamamoto, S. Ito, N. Shiraishi and T. Sagawa, *Phys. Rev. E*, **94**, 052121 (2016).
- [45] A. C. Barato, D. Hartich and U. Seifert, *New J. Phys.* **16**, 103024 (2014).
- [46] P. Sartori, L. Granger, C. F. Lee and J. M. Horowitz, *PLoS Comput. Biol.* **10**, e1003974 (2014).
- [47] S. Bo, M. Del Giudice and A. Celani, *J. Stat. Mech. Theor. Exp.* (2015). P01014.
- [48] D. Hartich, A. C. Barato and U. Seifert, *Phys. Rev. E*, **93**, 022116 (2016).
- [49] T. McGrath, N. S. Jones, P. R. ten Wolde and T. E. Ouldridge, *Phys. Rev. Lett.* **118**, 028101 (2017).
- [50] T. E. Ouldridge, C. C. Govern and P. R. ten Wolde, *Phys. Rev. X*, **7**, 021004 (2017).
- [51] S. Goldt and U. Seifert, *Phys. Rev. Lett.* **118**, 010601 (2017).
- [52] N. Ay and D. Polani, *Advances in complex systems*, **11**, 17 (2008).
- [53] M. Oizumi, N. Tsuchiya and S. I. Amari, *Proc. Natl. Acad. Sci.* **113**, 14817 (2016).