社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

# ばらつきを含む多層ニューラルネットワークモデルにおける確率共鳴の理論解析

佐橋　透†　宇田川　玲†　浅井　哲也†　雨宮　好仁†

† 北海道大学大学院情報科学研究科　〒060–0814 札幌市北区北 14 西 9 丁目
E-mail: †sahashi@lalsie.ist.hokudai.ac.jp

**あらまし**　我々は [1] において、素子ばらつきを持つ多層ニューラルネットワークにおける確率共鳴の数値シミュレーションを行った。ネッワークの入出力間の相関値を網羅的に計算した結果、この相関値を最大にする「しきい素子を駆動する動的な雑音の最適な強度」及び「しきい素子の最適な受容野の大きさ」が存在することを見いだした。この仕組みを解明することが本論文の目的である。提案する解析手法は、局所結合およびばらつきを持つセンサアレイ（イメージセンサ等）に適用可能であり、将来的には、極低消費電力で動作する "素子ばらつきを利用する" イメージセンサの研究と結びつく可能性がある。そこで、可能な限り理論的なアプローチで上記の仕組みを明らかにする。

**キーワード**　確率共鳴，はらつき，多層ニューラルネットワーク

# Theoretical Analysis of Stochastic Resonance with Population Heterogeneity in a Multi-Layer Neural Network Model

Toru SAHASHI†, Akira UTAGAWA†, Tetsuya ASAI†, and Yoshihito AMEMIYA†

† Graduate School of Information Science and Technology, Hokkaido University
Kita 14, Nishi 9, Kita–ku, Sapporo, 060–0814 Japan
E-mail: †sahashi@lalsie.ist.hokudai.ac.jp

**Abstract**　In [1] we examined stochastic resonance (SR) behaviors in a multilayer neural network with population heterogeneity. The results showed that correlation values between the input and output of the network had a peak for not only strength of temporal noises applied to the threshold elements in the network but also receptive-field sizes of the elements. Our purpose in this paper is to clarify the reason why the network exhibited such a SR behavior for the receptive-field size. Since the proposed scheme could be applied to sensor arrays having random offsets and local couplings, we here try to clarify the SR-like characteristics above as theoretically as possible.

**Key words**　stochastic resonance, noises and fluctuation, multi-layer neural network

## 1. まえがき

近年、電子工学の分野で確率共鳴（Stochastic Resonance : SR）の応用に関する研究が進んでおり、特に、確率共鳴を利用した微弱信号や暗画像の検出に関する研究が盛んに行われている [2]。暗画像を撮像（微弱な光信号を検出）するためには、通常、高精度—高価な撮像素子を用意する必要があるが、確率共鳴を利用して（生物的な方法で）暗画像の検出ができれば、低コストな暗画像撮像デバイス（微弱光検出回路）を実現できそうである。ただし、確率共鳴により一つの画素（フォトセンサ）における微弱光の検出ができても、フォトセンサをアレイ状に並べた場合（通常のイメージセンサの構造をとった場合）は、センサ間の特性ばらつきも確率共鳴により検出されてしまう [1]。したがって、確率共鳴を暗画像の撮像に応用することは、巷で言われているほど容易ではなさそうだ。

近年、猫の一次視覚野において、視覚経路にある神経細胞群が信号を検出するために確率共鳴を有効利用していることが示された（雑音により誘起された神経活動を抑えつつ、SN 比を高く保つ）[3]。このメカニズムはまだ解明されていないが、その背景には Collins らにより提唱された確率共鳴のモデル [4] があるものと考えられる。また、視覚野の神経細胞の多くは比較的広い受容野を持つことから、個々の神経素子が独立して確率共鳴を起こしているとは考えにくく、神経素子は（例えば近傍の）細胞と協調して確率共鳴を誘因している可能性がある。個々の神経素子が大きなばらつきを持っているにも関わらず、我々が暗画像を映像として認知できる理由の一つがここにある

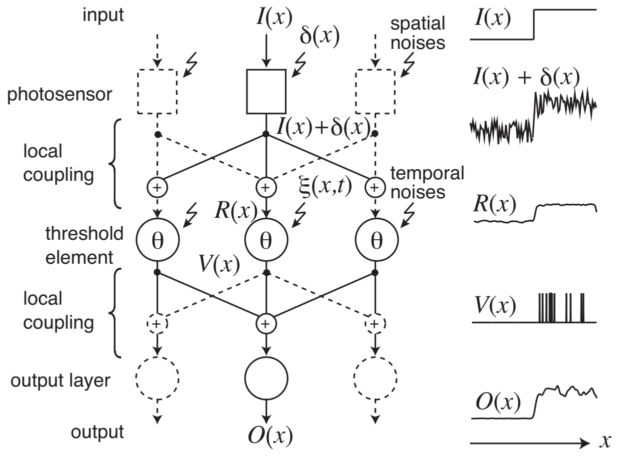


図 1　提案した多層ニューラルネットワークモデル [1]。空間ばらつき $\delta(x)$ を持つフォトセンサは入力 $I(x)$ を受け、その出力 $I(x)+\delta(x)$ は近傍結合を経てしきい素子に与えられる。しきい素子は雑音 $\xi(x,t)$ も同時に受ける。しきい素子の出力 $V(x)$ は近傍結合を経て出力層へ伝搬される。

のかもしれない。

上記の仮説を確かめるために、我々は [1] において、多層ニューラルネットワークのモデルを構築し、シミュレーションによりその性質を調べた。このモデルは、Collins らにより提唱された確率共鳴のモデルの単安定ダイナミクスを単純なしきい素子に置き換えたものであり、個々のしきい素子が隣接するしきい素子と重なり合うような構造を持つ（個々のしきい素子が受容野を持つ）。ネッワークの入出力間の相関値を網羅的に計算した結果、相関値を最大にする「しきい素子を駆動する動的な雑音の最適な強度」及び「しきい素子の最適な受容野の大きさ」が存在することを見いだした。この仕組みを理論的に解明することが本論文の目的である。

以下、2 章にて提案モデルおよびその数値シミュレーションの結果を概観する。3 章にて、提案モデルの理論解析を行う。評価関数として、相関値の代わりに入出力間の誤差関数を定義し、誤差の主要因となるパラメータを明らかにする。4 章にて、数値シミュレーションの結果と理論解析結果の比較検討を行う。

## 2. 提案モデル

図 1 に提案した多層ニューラルネットワークモデルを示す。構造を極力単純化するために、ばらつきを持つフォトセンサ, 近傍結合ネットワーク, しきい素子のみによりモデルを構成した。入力光 $I(x)$, フォトセンサのオフセットばらつき $\delta(x)$ を用いて、フォトセンサの出力を $I(x)+\delta(x)$ で定義する。$\delta(x)$ は $N(0,m^2)$ で与えられるランダム値である（平均 0, 分散 $m^2$ に従うガウス分布）。しきい素子への入力 $R(x)$ は

$$R(x) = \int \bigl(I(y)+\delta(y)\bigr)\cdot g(x-y)\,dy$$
$$\equiv \bigl(I(x)+\delta(x)\bigr) * g(x), \qquad (1)$$
$$g(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{x^2}{2\sigma^2}\right]$$

で定義する。ここで、$g(x)$ はフォトセンサ-しきい素子間の近傍結合ネットワークのインパルス応答である（しきい素子の受容野の大きさは $3\sigma$ 程度となる）。また、しきい素子の出力 $V(x,t)$ を $V(x,t)=H\bigl(R(x)+\xi(x,t)-\theta\bigr)$ で定義する。ここで、$H(\cdot)$ はステップ関数, $\xi(x,t)$ は振幅 $-A$〜$A$ の一様分布に従う乱数, $\theta$ はしきい値である。また、ネットワークの出力 $O(x,t)$ を $V(x,t)*g(x)$ で定義する。

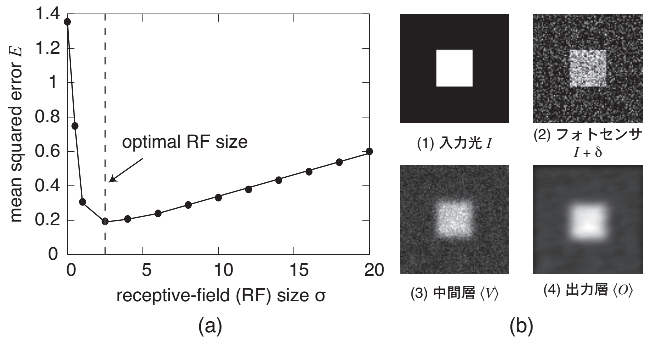


図 2　提案した多層ニューラルネットワークモデルのシミュレーション結果。(a) 入出力間の二乗誤差 $E$ の受容野サイズ $\sigma$ に対する変化。$\sigma=3$ 付近で誤差 $E$ が最小となる。(b) 二次元モデルにおけるシミュレーション結果。(1)〜(4) の画像はそれぞれ図 1 の $I(x)$, $I(x)+\delta(x)$, $V(x)$, $O(x)$ に対応する。

このモデルのシミュレーション結果を図 2(a) に示す。入力光 $I(x)$ を $S\cdot H(x)$（$S$ は入力光の明部強度）, $A=0.5$, $S=0.3$, $\theta=0.5$, 時間平均を計算するための試行回数を 1000 回としてシミュレーションを行った。しきい素子に与える雑音の振幅（$A$）を変化させる代わりに、受容野の大きさ（$\sigma$）を変化させて、$I(x)$ と $O(x)$ の間の二乗誤差 $E$ の変化を調べた。受容野が重ならない場合（$\sigma=0$）は $E$ は大きな値を持ったが（$E\approx 1.4$）、受容野が重なるにつれて $E$ は減少し、$\sigma=3$ 付近で最小となった（$E\approx 0.2$）。さらに受容野を広げると再び $E$ が上昇するという現象を確認した。

入出力間のエラーが最小となる $\sigma$（$=3$）を用いて、二次元モデルにおけるシミュレーションも併せて行った。その結果を図 2(b) に示す。出力層の画像 (4) は入力光 (1) の分布のコントラストが下がった画像として認識できるが、適度なレベル調整を行うことで、$g(x)$ の効果（受容野の重なり）による中間層の出力 (3) よりも明らかに入力画像に近い出力を得られることがわかった。以上より、入力にオフセットばらつきがある場合は、雑音強度のみならず受容野の大きさにも $E$ を最小にする最適値が存在することがわかった。次章にてこのメカニズムを理論的に明らかにする。

## 3. 理 論 解 析

### 3.1 誤差式の導出

通常、確率共鳴の評価には入出力間の相関係数を用いるが、相関係数を計算するためには、$I(x)$ および $O(x,t)$ の共分散と標準偏差を求める必要があり、その導出には少々骨が折れそうだ。そこで、2 章のシミュレーションと同様に、入力光 $I(x)$ と

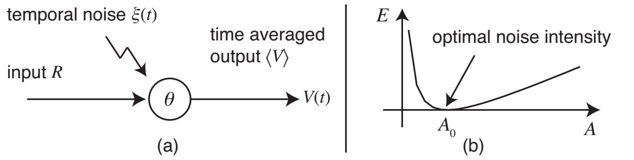


図 3　しきい素子単体による確率共鳴モデル。(a) しきい素子の入出力。入力 $R$ と振幅 $A$ の一様分布に従う雑音 $\xi(t)$ を加え、その出力 $V(t)$ の時間平均 $\langle V\rangle$ を計算する。(b) しきい素子単体の確率共鳴曲線。雑音振幅をパラメータとし、入出力間の誤差 $E$ をプロットしたもの。雑音強度 $A_0$ にて誤差が最小となり、このとき、$R=\langle V\rangle$ となる。

十分に時間平均をとった $\langle O(x)\rangle$ との間の二乗誤差 $E$ を評価関数として用いる。この二乗誤差を

$$E=\frac{1}{X}\int_{-X/2}^{X/2}\big(\langle O(x)\rangle-I(x)\big)^2dx, \tag{2}$$

で定義する。ここで、$X$ は $I(x)$ および $\langle O(x)\rangle$ を計算する空間範囲を表す。上式中の $\langle O(x)\rangle$ を導出するために、まず図 3(a) に示されるような、雑音を受けるしきい素子単体のモデルを考える。入力 $R$（$<\theta$）と雑音 $\xi(t)$ の和が $\theta$ より大きければ（または小さければ）、しきい素子の出力 $V(t)$ は 1（または 0）となる。この出力は、$\xi(t)$ に強く依存する 1 ビットの乱数列であるが、十分な時間をかけて平均化すると、$\langle V\rangle$ は定常値となり、後の理論計算が楽になる。この $\langle V\rangle$ は雑音強度 $A$ にも強く依存し、$A$ が大きければ $\langle V\rangle$ の振幅も大きくなる。図 3(b) に示すように、$\langle V\rangle=R$ となるような最適な強度 $A_0$ を持つ $\xi(t)$ をしきい素子に与える事により、誤差 $E$ は 0 になると考えられる。

上記の関係を用いて、図 1 における $I(x)$ と $\langle O(x)\rangle$ の関係を導く。2 章の定義より、時間平均されたネットワークの出力は、$\langle O(x)\rangle=\langle V(x)\rangle*g(x)$ となる。上述のとおり、最適な雑音強度下では $\langle V(x)\rangle=R(x)$ とおけるので、$\langle O(x)\rangle=R(x)*g(x)$ となる。また、2 章にて $R(x)=\big(I(x)+\delta(x)\big)*g(x)$ と定義したため、$\langle O(x)\rangle$ は

$$\langle O(x)\rangle=\Big[\big(I(x)+\delta(x)\big)*g(x)\Big]*g(x),$$

で表される。これを展開すると、時間平均された出力は

$$\langle O(x)\rangle=\big[I(x)*g(x)*g(x)\big]+\big[\delta(x)*g(x)*g(x)\big], \tag{3}$$

となる。式 3 を式 2 に代入し、ガウス分布同士の畳み込み $g(x)*g(x)$ を $g'(x)$ とおけば、二乗誤差 $E$ は

$$E=\frac{1}{X}\int_{-X/2}^{X/2}\big[I(x)*g'(x)+\delta(x)*g'(x)-I(x)\big]^2dx$$

$$\equiv E_1+2E_0+E_2, \tag{4}$$

で表される。ここで、$E_0$, $E_1$, $E_2$ はそれぞれ

$$E_0 = \frac{1}{X}\int_{-X/2}^{X/2}\big(I(x)*g'(x)-I(x)\big)\big(\delta(x)*g'(x)\big)dx,$$

$$E_1 = \frac{1}{X}\int_{-X/2}^{X/2}\Big[\big(I(x)*g'(x)\big)-I(x)\Big]^2dx, \tag{5}$$

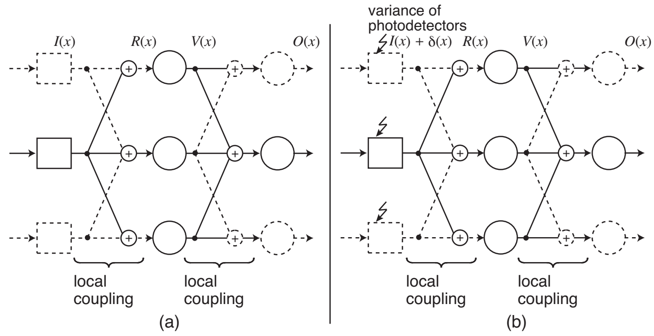


図 4　解析を行うモデルの概略図。(a) 入力ばらつきのないネットワーク, (b) 入力バラツキのあるネットワーク。

$$E_2 = \frac{1}{X}\int_{-X/2}^{X/2}\big(\delta(x)*g'(x)\big)^2dx, \tag{6}$$

で与えられる。

$E_0$ は入力 $I(x)$ とばらつき $\delta(x)$ の相関を表すため、$I(x)$ と $\delta(x)$ が互いに無相関である場合は、$E_0\approx 0$ としてよい。したがって、$E$ は互いに独立した誤差 $E_1$ と $E_2$ の和のみで表される。$E_1$ は $g'(x)$ による $I(x)$ の歪みにより生じる誤差、$E_2$ は $\delta(x)$ による誤差を表している。

ここでまず最も簡単な例として、ばらつきがなく（$m=0$）近傍結合を持たないモデル（受容野サイズ $\sigma=0$）における誤差を考える。$\sigma=0$ なので $g(x)$ はデルタ関数となり、$I(x)*g(x)=I(x)$ となるので $E_1=0$ となる。一方、$m=0$ なので $\delta(x)=0$ となり、その積分 $E_2$ も 0 となる。したがって、式 4 より、ばらつきがなく近傍結合を持たない場合の誤差 $E$ は 0 となる。

次に、ばらつきがなく近傍結合を持つ（受容野の広がりのある）モデル（図 4(a)）における誤差を考える。このモデルの場合、上述のとおり $E_0$ は 0 であり、また $\delta(x)=0$ より $E_2$ も 0 となる。したがって、誤差 $E$ は $E_1$（式 5）のみにより表される。$g'(x)$ は二つのガウス分布 $g(x)$ の畳み込みであり、ガウス分布同士の畳み込みもまたガウス分布となる。よって、

$$g'(x)=\frac{1}{\sqrt{2\pi}\sigma'}\exp\left[-\frac{x^2}{2\sigma'^2}\right], \tag{7}$$

を得る。畳み込んだガウス分布の標準偏差 $\sigma'$ は元の分布の標準偏差 $\sigma$ の $\sqrt{2}$ 倍となる ($\sigma'=\sqrt{2}\sigma$)。この $g'(x)$ を用いて式 5 の $I(x)*g'(x)$ を導出する。2 章にて入力 $I(x)$ を $S\cdot H(x)$ で定義したので、入力とガウス分布の畳み込み $I(x)*g'(x)$ は

$$I(x)*g'(x)=S\int_0^{X/2}g'(x-y)dy,$$

で表される。ガウス分布の積分はエラー関数で表すことができるため、

$$I(x)*g'(x)=\frac{S}{2}\left\{\mathrm{erf}\left(\frac{x}{\sqrt{2}\sigma'}\right)-\mathrm{erf}\left(\frac{x-X/2}{\sqrt{2}\sigma'}\right)\right\},$$

を得る。ここで、$x\ll X/2$ かつ $X\ll\sigma'$ であるとき、$\mathrm{erf}\big((x-X/2)/\sqrt{2}\sigma'\big)\approx -1$ と近似できるため、

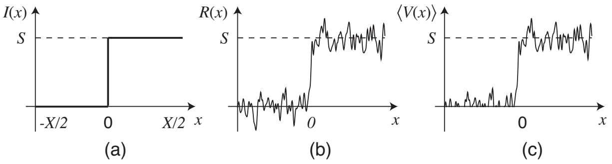


図 5　ステップ入力に対するしきい素子への入力およびしきい素子の出力。(a) ステップ入力 $I(x)$（$\equiv S \cdot H(x)$）, (b) しきい素子への入力 $R(x)$（$= \big(I(x)+\delta(x)\big) * g(x)$）。$R(x)$ は $x<0$ のときに負となる場合がある。(c) しきい素子の時間平均出力 $\langle V(x)\rangle$。この $\langle V(x)\rangle$ は必ず正となる。

$$I(x) * g'(x) \approx \frac{S}{2}\left(\mathrm{erf}\left(\frac{x}{\sqrt{2}\sigma'}\right)+1\right),$$

を得る。これを式 5 に代入すると、

$$E_1 = \frac{1}{X}\int_{-X/2}^{X/2}\left(\frac{S}{2}\left(\mathrm{erf}\left(\frac{x}{\sqrt{2}\sigma'}\right)+1\right)-I(x)\right)^2 dx,$$

を得る。$I(x)$ の定義より被積分関数は奇関数となり、その二乗は偶関数になる。また、$\sigma' = \sqrt{2}\sigma$ なので、誤差は

$$E_1 = \frac{S^2}{2X}\int_{0}^{X/2}\left(1-\mathrm{erf}\left(\frac{x}{2\sigma}\right)\right)^2 dx, \tag{8}$$

で表される。$E_0$ と $E_2$ は 0 であるため、この $E_1$ が「ばらつきがなく受容野の広がりのあるモデル」（図 4(a)）における誤差である。以上が、ばらつきのないネットワーク（近傍結合あり／なし）における誤差の導出法である。

次いで、入力ばらつきのあるネットワークにおける誤差の導出を行う。まず、近傍結合（受容野の重なり）がない場合の誤差を求める。この場合、時間平均された出力は

$$\langle O(x)\rangle = \langle V(x)\rangle = R(x) = I(x) + \delta(x),$$

で表されるため、誤差 $E$ は

$$\begin{aligned} E &= \frac{1}{X}\int_{-X/2}^{X/2}\big(\langle O(x)\rangle - I(x)\big)^2 dx \\ &= \frac{1}{X}\int_{-X/2}^{X/2}\delta^2(x)dx, \end{aligned} \tag{9}$$

で表される。つまり、ばらつきがあり受容野の重なりがない場合の誤差は、ばらつきの大きさ $m$ のみに依存する。

最後に、入力ばらつきがあり、近傍結合（受容野の重なり）があるネットワークにおける誤差を求める（図 4 (b)）。二乗誤差 $E$ は式 4 で表されるが、ここでも入力とばらつきは無相関であるとする（$E_0 = 0$）。また、式 5 で表される $E_1$ に $\delta(x)$ が含まれないため、このモデルの $E_1$ は式 8 で与えられる。よって、あとは式 6 で表される $E_2$ を求めればよい。ただし、式 6 は $R(x) = \langle V(x)\rangle$ という条件の下で成立する。$R(x)$ は $\big(I(x)+\delta(x)\big) * g(x)$ で与えられ、$\delta(x)$ は平均 0 のガウス分布に従う確率変数なので、$R(x)$ は負になり得る。一方、$V(x,t) = H\big(R(x)+\xi(x,t)-\theta\big)$ であるので、$V(x,t)$ の値は 0 または 1 のどちらかである。そのため、$V(x,t)$ の時間平均値 $\langle V(x)\rangle$ は必ず 0 以上となることから、$R(x)<0$ となる $x$ では上記の $R(x) = \langle V(x)\rangle$ という条件は成立しない。そこで、図 5 を用いて $R(x)<0$ となる領域の $E_2$ への影響を考える。$0<x<X/2$ の領域では $S>3m$ と定義しているので $R(x)<0$ となる確率は限りなく 0 に近い（図 5(b)）。よってこの領域では $R(x) = \langle V(x)\rangle$ が成り立つため、式 6 を用いて $E_2$ を求められる。一方、$-X/2<x<0$ の領域では $I(x)=0$ なので、$R(x) = \delta(x) * g(x)$（$\delta(x)$ は平均 0 のガウス分布に従うランダム値）より、$R(x)<0$ となることがある。$R(x)<0$ かつ $\xi(x,t)-\theta=0$ であれば $V(x,t)$ は必ず 0 になり、その時間平均 $\langle V(x)\rangle$ も 0 となる。$\langle O(x)\rangle = \langle V(x)\rangle * g$ より、$\langle V(x)\rangle = 0$ であれば $\langle O(x)\rangle = 0$ となり、この $x$ におけるばらつきによる誤差は 0 となる。したがって、この $x$ の集団が $-X/2<x<0$ の領域を占める割合に比例してばらつきによる誤差が小さくなる。この割合を知るために $\langle V(x)\rangle = 0$（つまり $R(x)<0$)）となる確率を求める。

$x<0$ では $R(x) = \delta(x) * g(x)$ となり、$\delta(x)$ は平均 0 のガウス分布に従うランダム値であるため、1/2 の確率で $\delta(x)<0$ となる。したがって、$R(x)<0$ となる確率は 1/2 となるので、$-X/2<x<0$ でのばらつきによる誤差は式 6 を用いて計算した誤差の 1/2 になる。$E_2$ は $-X/2<x<0$ のときのばらつきによる誤差と $0<x<X/2$ のそれとの和で表されるので、式 6 で計算した $E_2$ の 3/4 倍となる。よって、この条件下でのばらつきによる誤差 $E_2'$ は

$$E_2' = \frac{3}{4}E_2 = \frac{3}{4X}\int_{-X/2}^{X/2}\big(\delta(x) * g'(x)\big)^2 dx, \tag{10}$$

で表される。

次に、式 10 中の $\delta(x) * g'(x)$ を求める必要がある。しかし、$\delta(x)$ は確率変数であるため、$\delta(x) * g'(x)$ を直接的に求めることは困難である。しかし、$X$ が十分に大きいときは確率変数 $\delta(x)$ を直接扱う必要はなく、$\delta(x)$ の統計的な特性のみがわかればよい。簡便のため、昇順に並び替えた $\delta(x) * g'(x)(\equiv z(x))$ を図 6(b) に示す。式 10 より、$E_2$ は $z^2(x)$ の積分により得られるので、

$$E_2 = \frac{1}{X}\int_{-X/2}^{X/2} z^2(x)dx, \tag{11}$$

を得る。ここで $z(x)$ を求めるために、図 6(b) に示される $z(x_1) = \Delta$ について考える。図 6(c) は、$z$ の確率分布 $P(z)$(ガウス分布) を示している。$x_1$ は $z = 0$〜$\Delta$ までの累積確率密度の $X$ 倍に対応するので、

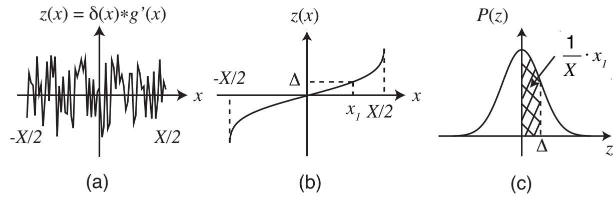


図 6　(a) 式 10 における $\delta(x) * g'(x)$（$\equiv z(x)$）の分布, (b) $z(x)$ を昇順に並び替えた分布, (c) 確率変数 $z(x)$ が従う標準偏差 $\sigma_o$ のガウス分布。

$$x_1 \;=\; X\int_0^{\Delta} P(z)dz \;=\; \frac{X}{2}\mathrm{erf}\left(\frac{\Delta}{\sqrt{2}\sigma_{\mathrm{o}}}\right),$$

を得る。ここで、$\sigma_{\mathrm{o}}$ は $\delta(x)*g'(x)$ の標準偏差である。これより、$\Delta = \sqrt{2}\sigma_{\mathrm{o}}\mathrm{erf}^{-1}(2x_1/X) = z(x_1)$ を得る。$x_1$ はすべての $x$ について成り立つので、$x_1 = x$ として式 11 に代入すると

$$\begin{aligned} E_2 \;&=\; \frac{1}{X}\int_{-X/2}^{X/2}\left(\sqrt{2}\sigma_{\mathrm{o}}\mathrm{erf}^{-1}\left(\frac{2x}{X}\right)\right)^2 dx \\ &=\; \frac{4\sigma_{\mathrm{o}}^2}{X}\int_0^{X/2}\left(\mathrm{erf}^{-1}\left(\frac{2x}{X}\right)\right)^2 dx, \end{aligned} \tag{12}$$

を得る。$E_2$ は $R(x) = \langle V(x)\rangle$ という条件の下で成り立ち、$x<0$ の領域で $R(x)<0$ になる場合のばらつきによる誤差 $E_2'$ は式 10 で与えられるので、式 12 を代入すると

$$E_2' = \frac{3\sigma_{\mathrm{o}}^2}{X}\int_0^{X/2}\left(\mathrm{erf}^{-1}\left(\frac{2x}{X}\right)\right)^2 dx,$$

が求まる、ここで $\delta(x)*g'(x)$ の分散は $\sigma_{\mathrm{o}}^2 = m^2/(2\sqrt{\pi}\sigma^2)$ で与えられる。したがって、式 12 は

$$E_2' = \frac{3m^2}{2X\sqrt{\pi}\sigma^2}\int_0^{X/2}\left(\mathrm{erf}^{-1}\left(\frac{2}{X}x\right)\right)^2 dx, \tag{13}$$

となる。以上より、近傍結合（受容野の重なり）があるモデルの入出力間の誤差 $E$ は式 8 の $E_1$ と式 13 の $E_2'$ の和で表される。

### 3.2 入力ばらつきのあるネットワークにおける最適な受容野サイズの導出

前節にて図 4(b) のモデルの誤差 $E$ を理論的に導出することができた。ここではその誤差が最小となる最適な受容野のサイズ（$\sigma_{\mathrm{m}}$）を求める。$\sigma_{\mathrm{m}}$ は $E$ を $\sigma$ で微分し、その微分係数が 0 となる $\sigma$ である。$E$ は受容野の重なりによる誤差 $E_1$ とばらつきによる誤差 $E_2'$ の和で表され、かつそれらは互いに独立であるので、$E$ の微分係数は

$$\frac{dE}{d\sigma} = \frac{dE_1}{d\sigma} + \frac{dE_2'}{d\sigma}, \tag{14}$$

となる。まず $E_2'$ の微分を求める。式 13 を $\sigma$ で微分すると、

$$\frac{dE_2'}{d\sigma} = -\frac{3m^2}{X\sqrt{\pi}\sigma^3}\int_0^{X/2}\left(\mathrm{erf}^{-1}\left(\frac{2}{X}x\right)\right)^2 dx, \tag{15}$$

が得られる。次に、$E_1$ の微分を求める。式 8 で表される $E_1$ は被積分関数に $\sigma$ が含まれるため、積分関数の微分を行う必要がある。したがって、$E_1$ の微分は

$$\begin{aligned} \frac{dE_1}{d\sigma} &= \frac{S^2}{2X}\int_0^{X/2}\frac{d}{d\sigma}\left(1-\mathrm{erf}\left(\frac{x}{2\sigma}\right)\right)^2 dx \\ &= -\frac{S^2}{X}\int_0^{X/2}\left(1-\mathrm{erf}\left(\frac{x}{2\sigma}\right)\right)\frac{d}{d\sigma}\mathrm{erf}\left(\frac{x}{2\sigma}\right)dx, \end{aligned} \tag{16}$$

で表される。次いで、被積分関数内の $\mathrm{erf}\big(x/2\sigma\big)$ の微分を行う。エラー関数は

$$\mathrm{erf}\left(\frac{x}{2\sigma}\right) = \frac{2}{\sqrt{\pi}}\int_0^{x/2\sigma}\exp\left(-t^2\right)dt, \tag{17}$$

で定義されるガウス分布の積分である。これを $\sigma$ で微分すると、積分の基本定理より

$$\frac{d}{d\sigma}\mathrm{erf}\left(\frac{x}{2\sigma}\right) = -\frac{x}{\sigma^2\sqrt{\pi}}\exp\left(\frac{x^2}{4\sigma^2}\right), \tag{18}$$

を得る。これを式 16 に代入し、$t = x/2\sigma$ とおくと

$$\frac{dE_1}{d\sigma} = \frac{4S^2}{X\sqrt{\pi}}\int_0^{X/4\sigma} t\cdot\exp\left(-t^2\right)\left(1-\mathrm{erf}(t)\right)dt, \tag{19}$$

が得られる。この積分を行うために、$1-\mathrm{erf}(t) \equiv f(t)$ と $t\cdot\exp(-t^2) \equiv g'(t)$ に分けて部分積分を行う。部分積分の定義より、式 19 は

$$\frac{dE_1}{d\sigma} = \frac{4S^2}{X\sqrt{\pi}}\left(\Big[f(t)\cdot g(t)\Big]_0^{X/4\sigma} - \int_0^{X/4\sigma} f'(t)\cdot g(t)dt\right),$$

で表される。ここで $g(t)$ は $\int g'(t)dt = -\exp(-t^2)/2$ となるので

$$\Big[f(t)\cdot g(t)\Big]_0^{X/4\sigma} = \left[-\frac{1}{2}\exp(-t^2)\cdot(1-\mathrm{erf}(t))\right]_0^{X/4\sigma}, \tag{20}$$

を得る。$I(x)$ および $O(x)$ の領域 $X$ は、受容野サイズ $\sigma$ より十分に大きいと仮定している ($X \gg \sigma$) ため、$\exp(-(X/4\sigma)^2) \approx 0$ および $\mathrm{erf}(X/4\sigma) \approx 1$ と近似できる。よって、式 20 は

$$\Big[f(t)\cdot g'(t)\Big]_0^{X/4\sigma} \approx \frac{1}{2}, \tag{21}$$

と近似できる。一方、$f'(t)$ はエラー関数の微分なのでガウス分布で表される。式 17 を用いると、$f'(t) = -2\exp(-t^2)/\sqrt{\pi}$ を得る。したがって、$f'(t)\cdot g(t) = \exp(-2t^2)/\sqrt{\pi}$ となる。この式と式 21 を用いると、式 19 は

$$\begin{aligned} \frac{dE_1}{d\sigma} \;&=\; \frac{2S^2}{X\sqrt{\pi}} - \frac{4S^2}{X\pi}\int_0^{X/4\sigma}\exp\left(-2t^2\right)dt \\ &=\; \frac{2S^2}{X\sqrt{\pi}} - \frac{\sqrt{2}S^2}{X\sqrt{\pi}}\mathrm{erf}\left(\frac{X}{2\sqrt{2}\sigma}\right), \end{aligned} \tag{22}$$

となる。$X \ll \sigma$ と仮定しているので、$\mathrm{erf}(X/2\sqrt{2}\sigma) \approx 1$ と近似できる。したがって、拡散結合による誤差 $E_1$ の微分は

$$\frac{dE_1}{d\sigma} \approx \frac{S^2}{X}\cdot\frac{2-\sqrt{2}}{\sqrt{\pi}} \equiv \frac{S^2}{X}\cdot\alpha, \tag{23}$$

となる（$\alpha \equiv (2-\sqrt{2})/\sqrt{\pi}$）。式 15 および式 23 を式 14 に代入すると、二乗誤差 $E$ の微分は

$$\frac{dE}{d\sigma} = \frac{\alpha S^2}{X} - \frac{3m^2}{X\sqrt{\pi}\sigma^3}\int_0^{X/2}\left(\mathrm{erf}^{-1}\left(\frac{2x}{X}\right)\right)^2 dx, \tag{24}$$

で表される。最適な拡散結合 $\sigma_{\mathrm{m}}$ は二乗誤差の微分が 0 となる $\sigma$ なので、

$$\sigma_{\mathrm{m}} = \sqrt[3]{\frac{3Xm^2}{2\alpha S^2\sqrt{\pi}}\int_0^1\left(\mathrm{erf}^{-1}(x)\right)^2 dx}, \tag{25}$$

を得る。エラー関数の二乗の積分を定数とみなすと、$\sigma_{\mathrm{m}}$ は $I(x)$ および $O(x)$ の領域 $X$, 空間ばらつきの標準偏差 $m$, 入力の振幅 $S$ の三つのパラメータにより表される。式 25 より、$\sigma_{\mathrm{m}}$ は $\sqrt[3]{X}$ および $m^{2/3}$ に比例し、$S^{2/3}$ に反比例する。このことより、$m$ あるいは $X$ が大きい場合は、誤差を最小にするために受容野を広げる必要があること, および入力の振幅が大きい場合は、誤差を最小にする受容野サイズは小さくて済むことがわかった。

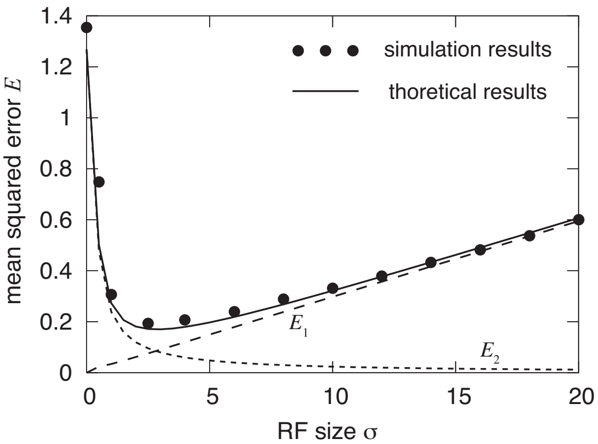


図 7　誤差の理論値 (実線) とシミュレーション値 (点) の比較

## 4. 理論とシミュレーションの比較検討

3 章で得られた誤差 $E$ と数値シミュレーションの比較検討を行う。最適受容野サイズ $\sigma_{\mathrm{m}}$ は、$E$ を $\sigma$ で微分して得られたものであるため、まず式 4, 8, 13 とシミュレーション結果が合致するか確認する。最適雑音強度 ($A = 0.5$) をしきい素子に与え、$S = 0.3$, $\theta = 0.5$, $m = 0.06$, $X = 500$, 時間平均のための試行回数を 1000 回としてシミュレーションを行った。図 7 に理論およびシミュレーションによって得られた $\sigma$ に対する $E$ の変化を示す。実線の曲線は理論値を, 点はシミュレーション値を表し、二つの破線の曲線はそれぞれ $E_1$ と $E_2$ の理論値を表す。この図より、理論値がシミュレーション値と一致することが確認できた。$E_1$ は $\sigma$ が大きいときに支配的になり、$\sigma$ の増加に対して単調増加する。一方、$E_2$ は $\sigma$ が小さいときに支配的になり、$\sigma$ の増加に対して単調減少する。$E$ は式 4 より $E_1$ と $E_2$ の和で与えられるので、図 7 からも $E$ を最小とする $\sigma$ の存在が確認できる。以上より、入出力信号の誤差を最小にする受容野サイズの最適値の存在が確認でき、かつ理論から得られた $E$ の式がシミュレーションと合致することが確認できた。

次に、式 24 で表される $dE/d\sigma$ と数値シミュレーション結果の比較を行った。モデルのパラメータは上記のシミュレーションと同じものを用いた。図 8(a) に理論およびシミュレーションによって得られた $\sigma$ に対する $dE/d\sigma$ の変化を示す。実線の曲線が理論値, 点がシミュレーション値を示し、破線はそれぞれ $dE_1/d\sigma$ および $dE_2/d\sigma$ を表す。この図より、$dE/d\sigma$ の理論値とシミュレーション値がほぼ一致することがわかる。

最適受容野サイズは、$dE/d\sigma = 0$ となる $\sigma$（$= \sigma_{\mathrm{m}}$）である。式 25 で計算した $\sigma_{\mathrm{m}}$ とシミュレーションにより得られた $\sigma_{\mathrm{m}}$ を比較するために、図 8(a) の $\sigma = \sigma_{\mathrm{m}}$ の付近を拡大したものを図 8(b) に示す。式 25 に用いたシミュレーションのパラメータを代入して計算した結果、$\sigma_{\mathrm{m}} = 2.82$ となった。一方、図 8(b) より $dE/d\sigma = 0$ となる $\sigma(\sigma_{\mathrm{m}})$ は約 2.82 であった。したがって、式 25 はシミュレーション結果と合致するといえる。以上より、入出力信号の誤差を最小にする受容野サイズの最適値を理論的に示すことができた。

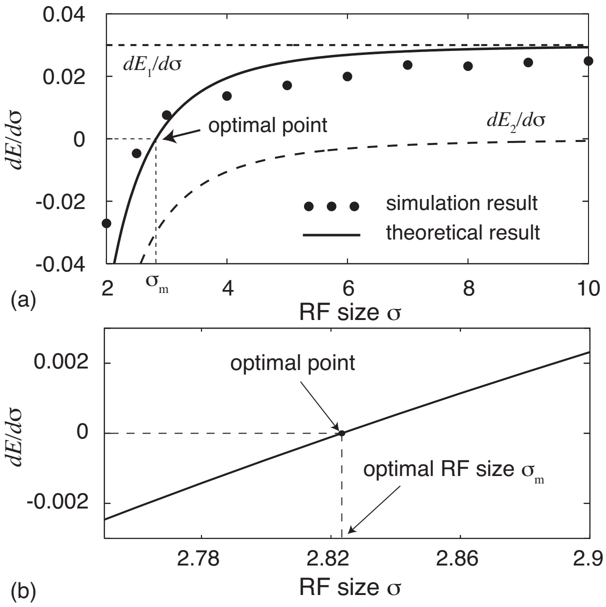


図 8　誤差の受容野サイズ変化量（$dE/d\sigma$）の受容野サイズ（$\sigma$）依存性。(a) $dE/d\sigma$ の理論値（実線）とシミュレーション値（点）。$dE_1/d\sigma$ および $dE_2/d\sigma$ の理論値を破線で示す。$dE/d\sigma = 0$ となる $\sigma$ が受容野サイズの最適値（$\sigma_{\mathrm{m}}$）である。(b) $\sigma = \sigma_{\mathrm{m}}$ 近傍の拡大図。

## 5. ま　と　め

素子ばらつきを持つ多層ニューラルネットワークにおける確率共鳴 [1] の理論解析を行った。まず、新たな評価関数として、ネッワークの入出力間の二乗誤差を定義し、[1] のモデルの再シミュレーションを行った。その結果、この誤差を最小にする「しきい素子を駆動する動的な雑音の最適な強度」及び「しきい素子の最適な受容野の大きさ」が存在することが確認できた。次いで、提案モデルの理論解析を行い、誤差を最小にする受容野のサイズが $\sqrt[3]{X}$（$X$ は空間幅）および $m^{2/3}$（$m$ は素子バラツキの標準偏差）に比例し、$S^{2/3}$（$S$ は信号強度）に反比例することを明らかにした。また、この結果がシミュレーション結果とよく合致することを示した。

## 文　　献

[1] A. Utagawa, T. Asai, T. Sahashi and Y. Amemiya, “Stochastic resonance in an array of locally-coupled McCulloch-Pitts neurons with population heterogeneity,” IEICE Trans. Fundamentals., vol. E92-A, no. 10, pp. 2508-2513, 2009.

[2] E. Simonotto, M. Riani, C. Seife, M Roberts, J. Twitty and F. Moss, “Visual Perception of Stochastic Resonance,” Phys. Rev. Lett. vol. 78, no. 6, pp. 1186 - 1189, 1997

[3] K. Funke, N. J. Kerscher and F. W ¨ org ¨ otter, “Noiseimproved signal detection in cat primary visual cortex via a well-balanced stochastic resonance like procedure,” European Journal of Neuroscience, vol. 5, pp. 1 - 35, 2007.

[4] P. C. Gailey, A. Neiman, J. J. Collins, and F. Moss, “Stochastic Resonance in Ensembles of Nondynamical Elements: The Role of Internal Noise,” Phys. Rev. Lett, vol. 79, no. 23, pp. 4701 - 4704, 1997.