

論文 / 著書情報
Article / Book Information

題目(和文)	人工ネットワークを用いた部分観測下におけるネットワーク構造の推定
Title(English)	
著者(和文)	小松三佐子
Author(English)	Misako Komatsu
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第9418号, 授与年月日:2014年3月26日, 学位の種別:課程博士, 審査員:中村 清彦,樺島 祥介,渡邊 澄夫,青西 亨,豊泉 太郎
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第9418号, Conferred date:2014/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

人工ネットワークを用いた部分観測下におけるネットワーク構造の推定

東京工業大学大学院 総合理工学研究科

知能システム科学専攻

小松 三佐子

目次

第1章 序論	1
1.1 目的	1
1.2 背景	1
1.2.1 脳内情報処理	1
1.2.2 部分観測問題	3
1.2.3 本論文の位置づけ	4
1.3 本論文の構成	5
第2章 ネットワークモデル	6
2.1 神経回路モデル	6
2.2 部分観測ネットワーク	7
2.3 学習手法	8
第3章 ネットワーク構造の推定	12
3.1 方法	12
3.1.1 皮質脳波	12
3.1.2 ネットワーク推定手法	14
3.2 皮質脳波間の機能結合	14
3.2.1 解析方法	15
3.2.2 結果	16
3.2.3 考察	23
3.3 皮質脳波から推定されたネットワーク構造	24
3.3.1 解析方法	24
3.3.2 結果	27
3.3.3 考察	29
第4章 結論	32
4.1 まとめ	32
4.2 本研究の意義	32
4.3 今後の展望 – 構造から機能へ	34
謝辞	38
参考文献	39

第 1 章 序論

1.1 目的

本研究の目的は、多点同時計測された脳信号の背景にある脳内ネットワークを推定することである。従来、多点計測された脳信号の解析に用いられている手法では、計測した信号間の関係性のみに焦点を当ててきたが、本来認知行動は全脳レベルあるいは外界からの影響も含めた情報処理により実現されていると考えられる。すなわち、観測していない因子からの動的な影響を考えることは、脳の情報処理を理解する上で重要である。本研究では、人工ネットワークモデルを用い、計測した信号の時系列学習を通して非観測因子をも含めたネットワーク構造を抽出する手法を提案した。また、提案した手法を実際の脳信号に適用し、推定結果の生理学的妥当性について検証を行った。

1.2 背景

1.2.1 脳内情報処理

神経科学の分野では近年の多点同時計測手法の発達により、脳内の情報処理の機構を解明する目的で、計測した信号間の関係性を推定する研究が注目を集めている。

同時計測された神経活動の関係を表す手法として、相互相関・コヒーレンス・位相同期といった、2変数間の統計的な独立性を表す指標が古くから用いられている (e.g., Griffith and Horn, 1963, Bressler et al., 1993, and Cobb et al., 1995, respectively)。これらの手法は、同時計測された 2 つの神経細胞あるいは脳領域間が機能的に結合している可能性を検証するには有効であったが、どちらが

情報の送り手でどちらが受け手であるかといった因果関係については分からなかった。したがって近年では、因果性を検証する手法として Granger causality (Granger, 1969) や transfer entropy (Schreiber, 2000) といった指標を用いる研究も増えている。Granger causality は最も広く使われている指標のひとつであり、もともとは社会経済学分野で用いられていた。遅延時間のある時系列データを用いた線形回帰分析を行うことで、ある変数の未来の状態について他の変数の過去の状態がどれだけ寄与しているのか定量化することを可能にしている (Granger, 1969) 。 Transfer entropy は、ある変数の状態が他の変数の状態遷移確率へおよぼす影響を考慮することで、2 変数の方向性を持った相互作用を検出する (Schreiber, 2000) 。しかしながら、これらの因果性の指標は因果性が直接的なものか間接的なものかを検出することができない。また、共通入力といった共通の外部からの影響を検出することができないことが知られている (Kaminski et al., 2001) 。

より最近の研究では、計測した信号間の関係性を推定する手法は多変数へと拡張され、神経回路内での因果関係の研究へと応用されている (e.g., Hesse et. al., 2003; Smith et. al., 2006) 。これらの研究は、多点計測された神経活動から、指

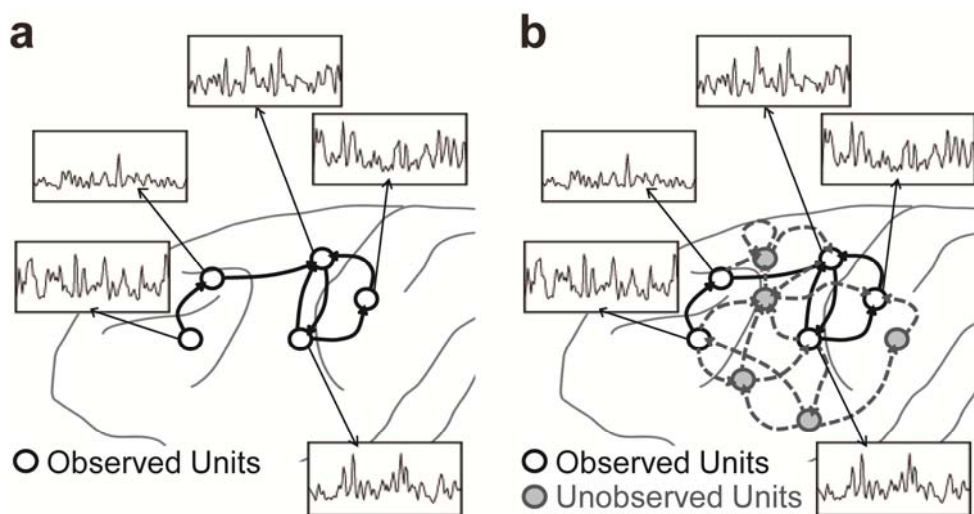


図 1.1 既存の研究 (a) および本研究 (b) の相違点

向性を持つ関係を任意の信号の組について決定することを目的としている。ただし、それらはいずれも計測した信号間の関係性のみに焦点を当てており、観測していない因子からの影響については考慮していない（図 1.1a）。しかしながら、認知行動は全脳レベルあるいは外界も含めたネットワーク内での情報処理により実現されていると考えられる（図 1.1b）。すなわち、非観測因子からの動的な影響を考えると、脳の情報処理を理解するために重要である。

1.2.2 部分観測問題

我々が観測可能なのはシステムの一部にすぎないという問題は、実世界の部分観測性として認識されている。部分的に観測したデータから、非観測部分に起因する不確実性を考慮に入れて特定のパラメータを決定する問題は、主に隠れマルコフモデルを用いて音声認識・ゲノミクス・自然言語処理などの分野で、また、部分観測マルコフ決定モデルを用いて適応的機械学習の分野で扱われてきた。これらの問題を解くアルゴリズムとしては Viterbi アルゴリズム (Viterbi, 1967) や Baum-Welch アルゴリズム (Baum et al., 1970) がある。しかしながら、これらのモデルは通常シンボルレベルの抽象化を行っており、高次元かつ時間的に連続な神経データを扱うには不十分である。さらに、計算論的神経科学の分野では、計算機実験により神経細胞の発火活動の部分観測問題を扱った例が数例ある (Pillow and Latham, 2008; Brea et al., 2011)。Pillow and Latham (2008) は感覚入力と神経発火活動とのマッピングの際に観測していない神経細胞からの影響を考慮する必要性を指摘している。Brea ら (2011) はスパイク列の学習の際に隠れユニットを導入することで、非マルコフ過程に従うデータや時間的にギャップのあるようなデータといった、それまでのモデルでは学習が困難であったスパイク列を学習できることを示した。ただし、これらの研究は計算機実験による検証を行ったのみで、実データへの

適用へは至っていない。また、人の脳内には大脳皮質だけでもおよそ 140 億個の神経細胞から構成されているため、スパイクニューロンモデルにより全脳レベルでのネットワーク構造を推定することは計算資源の面でも非現実的であり、何らかの抽象化が必要であると考えられる。本研究では、部分観測下において脳内ネットワーク構造を推定するための巨視的な神経回路のモデルとして、実数値リカレントネットワークを用いる。ネットワークの要素は互いに相互作用し、各要素の出力として実数値の時系列データを生成する。このようなネットワークは複雑な動的状態を獲得することができるが知られている (Funahashi and Nakamura, 1993; Beer, 1995)。したがって、各要素の出力が神経細胞集団の平均発火頻度を表現していると考えると、実数値要素を持つリカレントネットワークは巨視的な神経回路の複雑な振舞いを再現することのできる有望なモデルのひとつであると考えられる。実際、Yamashita らは同様のリカレントネットワークを用い、小鳥の歌文法の学習・生成にかかわる巨視的な神経回路の振舞いを説明することに成功している (Yamashita et al., 2008a; Yamashita et al., 2011)。また、これらのネットワークは逐次学習を通してその入出力からネットワークのパラメータを学習できるため、記録した脳信号を入出力として用いることで、脳信号における部分観測問題を取り扱うことが可能である。

1.2.3 本論文の位置づけ

本論文では、非観測因子からの動的な影響を陽に仮定し、多点同時計測した脳信号の背景にあるネットワーク構造 (図 1.1b) を抽出する手法を提案している。提案した手法では、リカレントニューラルネットワークを用い、計測した信号を教示データとした時系列学習を通して、信号の背景にあるネットワーク構造を推定する。リカレントニューラルネットワークを用いた時系列データの学習により教示データ

の動的性質をネットワーク上に再構築するという問題は、音声処理、経済学、知能ロボティクスなどの分野で用いられている手法であり、神経科学の分野でも神経活動と行動のマッピングのために用いられている例がある (Chapin et al., 1999; Wessberg et al., 2000) . しかしながら、脳信号そのものの再構築を試みた研究は本研究が初である.

我々はさらに、本手法を 2 個体のサルから計測した皮質脳波へ適用し、本研究で提案した手法が脳信号の背後にある脳内ネットワークを推定できる可能性について実データを用いた検証を行った.

1.3 本論文の構成

本論文は序論と結論を含めて 4 章で構成されている. 第 1 章の序論に続き、第 2 章では、本論文で提案するリカレントネットワークモデルの導入を行う. はじめに我々が神経回路をどのようなモデルとしてとらえているか説明し、次に、ネットワークの一部のみから得た観測データを用いてネットワーク全体のパラメータを推定するネットワークモデルとして、部分観測ネットワークを導入する. また、本研究で採用した学習手法について説明する. 次に、第 3 章で実際の脳信号に部分観測ネットワークモデルを適応した結果を示す. はじめに具体的なネットワーク推定手順について説明する. 続いて推定結果について、まず、観測した脳信号間の機能結合について、次に、非観測部分も含めたネットワーク全体の構造について解析を行った. 最後に第 4 章において本論文の総括を行い、今後の展望を示す.

第2章 ネットワークモデル

本章では、ネットワークモデルの導入を行う。本論文で提案する手法では、リカレントネットワークモデルを用い、計測した信号を教示データとした時系列学習を通して、信号の背景にあるネットワーク構造を推定する。はじめに我々が神経回路をどのようなモデルとしてとらえているか説明し、次に、ネットワークの一部のみから得た観測データを用いてネットワーク全体のパラメータを推定するリカレントネットワークモデルとして、部分観測ネットワークを導入する。最後に、観測データからネットワーク構造を推定する具体的な方法として、本研究で採用した時系列学習のアルゴリズムについて説明する。

2.1 神経回路モデル

本論文では、巨視的な神経回路は連続時間リカレントネットワークモデル（図 2.1a）で表されると仮定している。各ユニットは結合強度を通して相互作用し、各時刻の状態を更新する。ここで、 $c_t^{(i)}$ 、 $s_t^{(i)}$ をそれぞれ時刻 t におけるユニット i の内部状態および出力とする。また、ユニット j からユニット i への相互作用の強さを示す結合強度を W_{ij} で表すとすると、各ユニットの内部状態および出力の更新は以下の式に従う。

$$c_t^{(i)} = \left(1 - \frac{1}{\tau^{(i)}}\right)c_{t-1}^{(i)} + \frac{1}{\tau^{(i)}} \sum_{j=1}^N W_{ij} s_{t-1}^{(j)}, \quad (1)$$

$$s_t^{(i)} = \tanh\left(c_t^{(i)} + h^{(i)}\right), i = 1, 2, \dots, N, \quad (2)$$

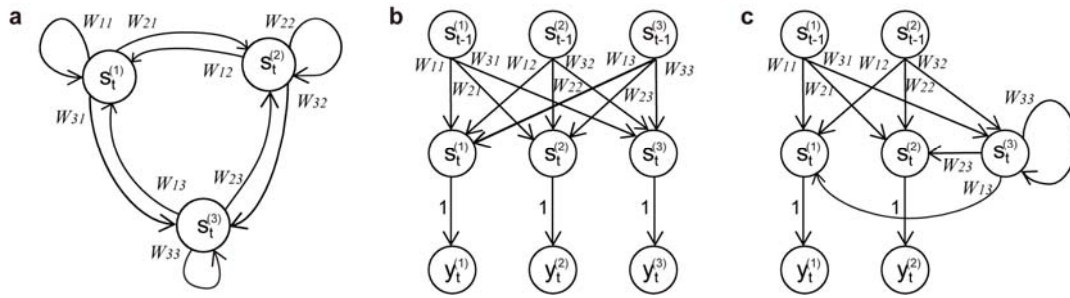


図 2.1 (a)3つのユニットで構成されるリカレントネットワークの例. これらのユニットは互いに結合強度 W_{ij} で相互作用し, それぞれの状態 s を出力する. (b)入力層は1ステップ前のユニットの出力, 出力層は現在の出力を表す. (c)部分観測ネットワーク. ネットワークの入出力は観測された信号のみからなる. 中間層は可観測ユニットおよび非観測ユニットで構成される.

ここで, N , $\tau^{(i)}$, $h^{(i)}$ はそれぞれネットワーク内のユニットの数, ユニット i の時定数, ユニット i のバイアスである. 式 (1) は時刻 t におけるユニット i の内部状態 $c_t^{(i)}$ がある結合強度で他のユニットから受け取る入力との総和と自身の過去の状態との内挿で与えられることを示している. 式 (2) はユニットの出力がある制限された値をとるとする仮定を示している.

これらの状態の更新は, フィードフォワード型の多入力・多出力ネットワークとして表することができる (図 2.1b). 1ステップ前のユニットの状態を入力層に持ち, 出力層は時刻 t の状態をとる. すなわち, $y_t^{(i)}$ を出力層のユニットの出力とすると,

$$y_t^{(i)} = s_t^{(i)}, i = 1, 2, \dots, N, \quad (3)$$

と表される.

2.2 部分観測ネットワーク

ここで, ネットワークの一部の状態しか観測できないと仮定する. すなわち, 図 2.1a の例では, ユニット 1, ユニット 2 の状態を観測可能, ユニット 3 の状態を非

観測とする。このとき、図 2.1b の多入力・多出力ネットワークは、観測した状態のみを入出力とするように変形して図 2.1c で表すことができる。このネットワークの入出力は観測した信号のみから構成され、中間層は可観測ユニットおよび非観測ユニットの両方で構成される。入力層から中間層への結合は全結合であり、中間層内では、非観測ユニットは観測ユニットおよび非観測ユニット両方へ結合する。中間層の可観測ユニットは出力層の対応するユニットのみに自身のコピーを送る。本論文では、この特別な制約を設けたリカレントネットワークを部分観測ネットワークと呼ぶことにする。同様の 3 層リカレントネットワークははじめに Elman (1990) によってコネクショニストモデルで時間を表現する目的で導入された。また、Jordan (1992) は運動制御の時系列学習のために Elman とは異なる制約の 3 層リカレントネットワークを提案した。これらのネットワークはコンテキストユニットと呼ばれる構造体を持っており、部分観測ネットワークではそれは非観測ユニットに対応している。Elman ネットワークおよび Jordan ネットワークでは、コンテキストユニットはそれぞれ中間層および出力層からの入力を受けている。それにより、コンテキストユニットはネットワークの前の状態を保持する役割を担っている。それらは複雑なダイナミクスを再現することができるため、時系列学習に利用されている (Elman, 1990; Jordan and Rumelhart, 1992) .

2.3 学習手法

次に、我々はネットワークのいくつかのユニットから記録された時系列データを持っており、ネットワークのパラメータは未知である場合を考える。ここでの目的は、記録された時系列データを用いてネットワークのパラメータを推定することである。記録されたデータは $z_t^{(k)}$, $k=1, 2, \dots, CH$, $t=1, 2, \dots, T$ で記述する。ここ

で、 CH は記録されたユニットの数、 T は全記録期間におけるデータ点の数に対応する。 W_{ij} , $\tau^{(i)}$, $h^{(i)}$, は推定したパラメータとすると、部分観測ネットワークの可観測および非観測ユニットの内部状態 $\hat{c}_t^{(i)}$ と出力 $\hat{s}_t^{(i)}$ は以下で更新される。

$$\hat{c}_t^{(i)} = \left(1 - \frac{1}{\tau^{(i)}}\right) \hat{c}_{t-1}^{(i)} + \frac{1}{\tau^{(i)}} \sum_{j=1}^{CH} W_{ij} z_{t-1}^{(j)} + \frac{1}{\tau^{(i)}} \sum_{j=CH+1}^N W_{ij} \hat{s}_{t-1}^{(j)}, \quad (4)$$

$$\hat{s}_t^{(i)} = \tanh(\hat{c}_t^{(i)} + h^{(i)}), i = 1, 2, \dots, N, \quad (5)$$

ここで、 N は中間層のユニットの数である。さらに、ネットワークの出力 $y_t^{(k)}$ は以下で更新される。

$$y_t^{(k)} = \hat{s}_t^{(k)}, k = 1, 2, \dots, CH. \quad (6)$$

ここで、 $y_t^{(k)} = z_t^{(k)}$ を満たす W_{ij} が見つけれられたら、それらの W_{ij} を持つネットワークは観測された信号の背後にあるネットワークの構造を表していると考えられる。このようなネットワークのパラメータは、ニューラルネットワークの時系列学習に用いられる一般的なアルゴリズムである勾配降下法を用いた最尤推定 (Namikawa and Tani, 2008) により推定することができる。学習のはじめに、ネットワークのパラメータの初期状態をランダムに設定し、それらのパラメータの値は記録された信号とネットワークの出力との誤差に基づいて変更される。この過程を多数回繰り返すことで、最終的に記録された信号とネットワークの出力の誤差は減少する、すなわち $y_t^{(k)} = z_t^{(k)}$ に近づく。観測された時系列データを用いた逐次学習を通し、そのデータの順モデルとしてネットワークのパラメータを得ることでデータの背後にあるネットワーク構造を推定する。

学習手法を実装するため、我々はネットワークに確率分布を導入した。ここで、ネットワークのパラメータを $\mathcal{G} = (\mathbf{W}, \boldsymbol{\tau}, \mathbf{h}, \hat{\mathbf{s}}_0)$ と記述する。 \mathbf{W} は結合強度 W_{ij} の $N \times N$ 行列、 $\boldsymbol{\tau}, \mathbf{h}, \hat{\mathbf{s}}_0$ はそれぞれ各ユニットの特徴量である時定数 $\tau^{(i)}$ 、バイアス $h^{(i)}$ 、初期状態 $\hat{\mathbf{s}}_0^{(i)}$ の長さ N のベクトルである。記録された信号は $\mathbf{Z} = (\mathbf{z}_t)_{t=1}^T$ で記述される。ここで \mathbf{z}_t は時刻 t における記録された信号 $z_t^{(k)}$ の長さ CH のベクトルである。任意の \mathbf{Z} および \mathcal{G} について、ネットワークの出力 \mathbf{y}_t の確率密度関数は以下で定義される。

$$p(\mathbf{y}_t | \mathbf{Z}, \mathcal{G}) = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^{CH} \exp\left(-\frac{\|\mathbf{z}_t - \mathbf{y}_t\|^2}{2\sigma^2} \right), \quad (7)$$

ここで、 \mathbf{y}_t は時刻 t におけるネットワークの出力 $y_t^{(k)}$ からなる長さ CH のベクトルである。全時系列 $\mathbf{Y} = (\mathbf{y}_t)_{t=1}^T$ の確率は以下で与えられる。

$$p(\mathbf{Y} | \mathbf{Z}, \mathcal{G}) = \prod_{t=1}^T p(\mathbf{y}_t | \mathbf{Z}, \mathcal{G}). \quad (8)$$

この式は観測された信号が時間領域に対してガウスノイズを含んでいるという仮定に基づいている。平均二乗誤差の最小化が正規分布で決定される尤度の最大化と等価であることはよく知られている。したがって、(7) 式はニューラルネットワークの学習への尤度最大化の自然な拡張を可能にしている (Jacobs et al., 1991)。 σ は正規分布の分散であり、scaling parameter と呼ばれている (tani and nolfi, 1999)。本研究では、ノイズについて特別な仮定をおいていないため、従来のモデルで用いられる定数値 $\sigma = 1$ を用いた。

教示データ \mathbf{Z} の下で尤度 $p(\mathbf{Y} | \mathbf{Z}, \mathcal{G})$ を最大化して最適なパラメータ \mathcal{G} を得るため、学習手順として慣性項を導入した勾配降下法を用いた。

ネットワークパラメータの更新ルールは以下で表される。

$$\Delta \mathcal{G}_n = \frac{\partial \ln p(\mathbf{Y}|\mathbf{Z}, \mathcal{G}_n)}{\partial \mathcal{G}_n} + \eta \Delta \mathcal{G}_{n-1}, \quad (9)$$

$$\mathcal{G}_{n+1} = \mathcal{G}_n + \alpha \Delta \mathcal{G}_n, \quad (10)$$

ここで \mathcal{G}_n は探索における n ステップ目のパラメータであり， α は学習率， η はパラメータの慣性項である． \mathcal{G} の偏微分は以下で与えられる．

$$\frac{\partial \ln p(\mathbf{Y}|\mathbf{Z}, \mathcal{G})}{\partial \mathcal{G}} = \sum_{t=1}^T \frac{-1}{2\sigma^2} \frac{\partial}{\partial \mathcal{G}} \|\mathbf{z}_t - \mathbf{y}_t\|^2. \quad (11)$$

ここで， $\frac{\partial}{\partial \mathcal{G}} \|\mathbf{z}_t - \mathbf{y}_t\|^2$ は通時的誤差逆伝搬法（Rumelhart, 1986）を用いて解くことができる．

学習の後，もしネットワークの出力が完全に教示信号と同じであるならば，ネットワークはその出力する状態を外部入力なしに更新することができる．さらに，もし得られたパラメータが唯一の解ならば，（4）式と（1）式は相同となる．ただし，多くの場合においてこれらのパラメータの推定は不良設定問題である．なぜならば，現実におけるほとんどの問題は単純自律系ではなく，単一解の存在は保証されないからである．しかしながら少なくとも，今回提案する手法を用いることで，過去の状態 \mathbf{z}_{t-1} と現在の状態 \mathbf{z}_t の矛盾を減少させるようなネットワークのパラメータが部分観測ネットワーク内に獲得されるはずである．

学習の初期状態として，結合強度および閾値は $[-1/N, 1/N]$ の範囲の一様分布乱数，初期状態 $\hat{\mathbf{s}}_0$ は $[-1, 1]$ の範囲の一様分布乱数とする．また， $p(\mathbf{Y}|\mathbf{Z}, \mathcal{G})$ の最大値は教示データの長さ T および出力ユニット数 CH に依存するため，学習率 α は $\alpha = \tilde{\alpha}/T \cdot CH$ でスケールされる．

第3章 ネットワーク構造の推定

本章では実際の脳信号に部分観測ネットワークモデルを適応した結果を示す。はじめに教示データとして用いた脳信号および具体的なネットワーク推定手順について説明する。続いて推定結果について、まず、観測した脳信号間の機能結合についてネットワークの学習成績を含めて解析を行った結果を示す。次に、非観測部分も含めたネットワーク全体の構造について解析を行った結果を示す。

3.1 方法

3.1.1 皮質脳波

我々は、前章で提案したネットワークモデルを実際の脳信号の解析に適用した。脳信号として、サルの硬膜下の皮質表面に埋め込まれたグリッド状の電極により記録された皮質脳波を用いた。脳波や脳磁図と比較して、皮質脳波は優れた空間解像度、高い信号対ノイズ比を持つなど、脳活動の時空間情報を持つ信号としていくつかの利点を持っていると考えられている。また、細胞外微小電極記録と比較すると、臨床的リスクが低く、長期にわたって安定的で信頼できる信号が記録できることが知られている (e.g., Leuthardt et al., 2004; Chao et al., 2010; Buzsaki et al., 2012)。我々は部分観測ネットワークを用いることで、皮質表面のネットワーク構造が推定できると考えている。本研究で用いた皮質脳波は Chao らにより記録され、インターネット上で公開されているデータを用いた (<http://neurotycho.org/data/20081127s1food-trackingazenaschao>; <http://neurotycho.org/data/20090527s1food-trackingk1zenaschao>) (Chao et al., 2010; Nagasaka et

al., 2011) . サル A では 32 チャンネルの電極を右半球の前頭前野から一次体性感覚野を覆う領域に, サル K では 64 チャンネルの電極を左半球の前頭前野から頭頂葉を覆う領域に埋め込んだ (図 3.1a) . 電極の位置はコンピューター断層撮影 (CT) および磁気共鳴画像 (MRI) により同定された. 皮質脳波は 1 kHz のサンプリングレートで記録された. 記録中, 2 頭のオスの日本サル (サル A および K) は反復エサ追従課題を行った (図 3.1b) . サルは実験者と対面してイスに座り, エサが実験者によって 3 次元空間内に提示された. その後, サルは電極が埋め込まれた半球と対側の手でエサを追従する運動を行う. 課題の詳細は (Chao et al., 2010) に記述されている.

推定に用いた教示データは以下の手順で作成された. 記録された信号は, すべての加算平均電位を基準電位とする平均基準誘導法 (common average reference) による波形を求め, その後, complex-Morlet ウェーブレット変換を用いて高ガンマ帯 (60-100 Hz) のパワー値を計算した. 高ガンマ帯を用いたのは, 皮質脳波における

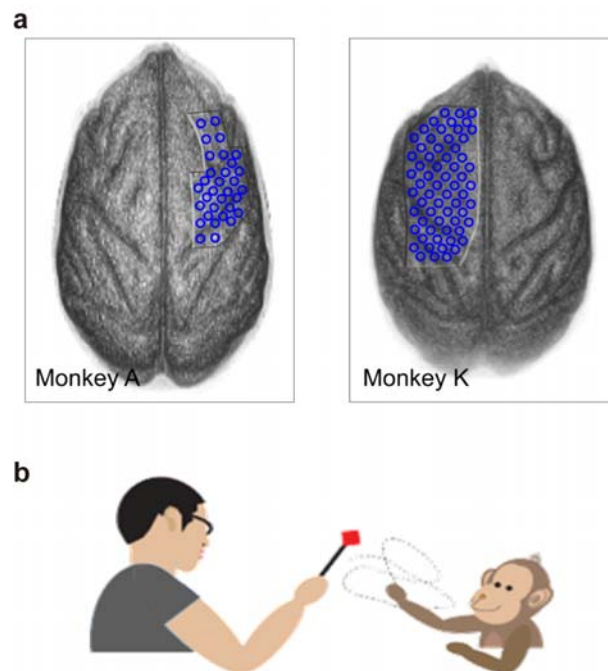


図 3.1 (a) 電極位置 (b) エサ追従課題. Chao ら (2010) Figure 1b から改変.

この周波数帯のパワー値が記録部位近傍に位置する神経集団の平均発火頻度と強く相関を持つことが知られているためである (Ray et al., 2008) . 信号はその後、咀嚼中のデータを除去し、サルAの1回の腕運動を含む1 トライアル毎の時系列データに分割された。最後に記録電極毎に z-score による標準化を行った。以下のネットワーク構造の推定では、各個体について初めの 32 トライアルのデータを教示データとして使用した。使用した教示データの長さの平均値 \pm SD はサルAが 9.219 ± 4.038 秒、サルKが 9.656 ± 4.397 秒であり、1 秒間に 1000 データ点が含まれている。

3.1.2 ネットワーク推定手法

2.3 節の学習手法を用い、皮質脳波を教示データとして部分観測ネットワークのパラメータの推定を行った。観測ユニット数 CH は皮質脳波の電極と同数、すなわち、サルAでは 32、サルKでは 64 とした。非観測ユニット数 $U (=N-CH)$ は 0, 1, 5, 10 としてそれぞれ推定を行った。すべての推定で共通のパラメータとして、 τ , $\tilde{\alpha}$, η はそれぞれ 0.2, 0.01, 0.9 とした。ほかのパラメータについては学習の初めにランダムに設定した。局所解を避けるため、5 つのランダムシードを用いて推定を行った。また、推定したネットワーク構造が特定の信号によるものではないことを確認するため、32 時系列を 8 つずつ 4 つのグループに分け、それぞれのデータグループについて独立に推定を行った。適切な結合強度を推定するため、50000 回の探索を行った。

3.2 皮質脳波間の機能結合

本節では、提案するネットワークの推定手法が生理学的に妥当な結果を得ることを確認するため、既存の生理学的知見と比較可能な皮質脳波間の機能結合、すなわ

ち可観測ユニット間の結合強度に焦点をあてて解析を行った。解析に先立って、部分観測ネットワークを用いて皮質脳波が学習できることを確認した。その後、推定した可観測ユニット間の結合強度の生理学的妥当性について検証を行った。

3.2.1 解析方法

推定したネットワーク構造の解析に先立って、部分観測ネットワークを用いて皮質脳波が学習できることを確認する必要がある。学習成績を評価する指標である学習誤差として、我々は各学習ステップにおいて、ネットワークの出力 \mathbf{Y} と教示データ \mathbf{Z} の平均二乗誤差 (MSE) を次式に従って計算した。

$$\text{MSE} = \frac{1}{2T \cdot CH} \sum_{t=1}^T \|\mathbf{z}_t - \mathbf{y}_t\|^2.$$

また、学習結果の汎化性能を評価するため、汎化誤差として推定したネットワークパラメータを用いて学習していないデータを入力としてネットワークの出力を計算し、同様に平均二乗誤差を計算した。

次に、異なる教示データから得た結合行列の平均値を計算し、異なるデータ間でも共通してみられる傾向について考察を行った。特に、ネットワークの構造的特徴としてクラスター構造に着目した。結合行列にみられるクラスター構造を定量的に比較するため、観測ユニット間の結合強度に対して Ward 法を用いた階層的クラスター分析を行った。結合行列毎にクラスターツリーを計算した。Ward 法では各クラスターの分散を計算し、その和が最小となるようにクラスターを作成する。ユニット i と j の間の非類似度 D_{ij} は以下で記述される。

$$D_{ij} = \sum_{k=1}^{CH} \|W_{ik} - W_{jk}\|^2.$$

ここで、 W_{ik} は状態 $\hat{s}_{i-1}^{(k)}$ の状態 $\hat{s}_i^{(i)}$ に対する影響を表しており、ユニット i とユニット j が他のユニットから受ける影響が同じほど非類似度 D_{ij} は小さな値を取り、同じクラスターに分類される。各クラスターツリーから任意のクラスター数に電極を分けることができる。

さらに、上述の手順で得たクラスター構造の生理学的妥当性について検証するため、同一クラスターに分類されたユニットの皮質上の電極位置の分布と皮質領野との対応について定量化を試みた。図 3.6b 上図を例にとると、四角で示されるクラスター内のユニットは、arcuate sulcus を境界として 6 つのユニットと 1 つのユニットに分けられる。ここで、我々はクラスター内の大多数のユニットが属する皮質領野をそのクラスターの「対応皮質領野」とした。そして、赤丸で囲まれた四角は対応皮質領野の外に位置している「はずれユニット」とした。同様に、十字で示されるクラスター、即ちもっとも大きな分散を持つクラスターを除く 5 つのクラスターについて調べると、はずれユニットの総数は図の例では 3 となる。もし、得られたクラスターの分布が皮質領野と一致しているならば、はずれユニットの数は小さくなることが期待される。皮質領野と一致しないクラスターの分布として、得られたクラスターのユニット数と同数のグループにランダムにグループ分けした皮質脳波の電極位置について、同様にはずれユニット数を調べ、推定した結果から得られたはずれユニット数がランダムなグループから得られる確率について検証した。

3.2.2 結果

各個体について、非観測ユニット数を 0, 1, 5, 10 と仮定した部分観測ネットワークを用い、5 つのランダムシードおよび 4 つのデータセットでそれぞれネットワ

ークパラメータの推定を行った。異なる非観測ユニット数を仮定した学習の進捗の例として、あるデータセットを用いた際の学習誤差の推移を図 3.2a 示す。非観測ユニット数によらず、すべてのランダムシードおよびデータセットについて、50000 回の学習ステップ後のネットワークの出力 \mathbf{Y} と教示データ \mathbf{Z} の平均二乗誤差は学習前の平均二乗誤差に比べて小さかった (Wilcoxon rank-sum test; $p < 0.01$)。すなわち学習誤差は学習の経過に従い減少した。次に、50000 回の学習ステップ後のパラメータを用いて学習していないデータセットの時系列を生成し、汎化誤差を計算した。その結果、サル K の非観測ユニット数 10 の場合 (Wilcoxon rank-sum test; $p = 0.0019$) を除いて学習誤差と汎化誤差には有意な差がみられなかった

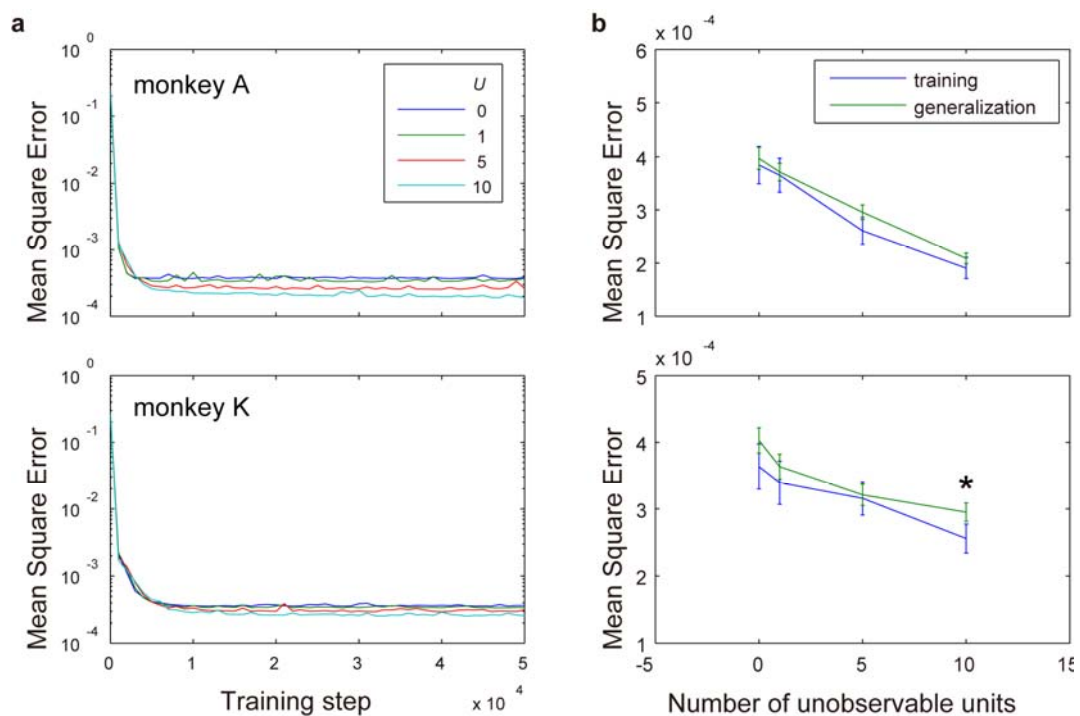


図 3.2 (a) サル A (上図), サル K (下図) における学習曲線。横軸は学習ステップ数, 縦軸は平均二乗誤差を表す。各線はそれぞれ異なる非観測ユニット数を持つ部分観測ネットワークを用いた際の 8 トライアルの平均値である。(b) 50000 ステップ後の学習誤差および推定したパラメータを学習していないデータに適用した際の汎化誤差。横軸は非観測ユニット数。平均二乗誤差は非観測ユニット数が増加するに伴い減少した (青線)。

	$U=0$	$U=1$	$U=5$	$U=10$
monkey A	35/1024	483/1024	207/1024	109/1024
monkey K	66/4096	271/4096	230/4096	304/4096

表 3.1 異なるデータセットから得た観測ユニット間の結合強度の平均値は、特定の皮質脳波のチャンネルにおいて統計的に有意にゼロではなかった ($p < 0.0001$) .

($p > 0.01$) . また、平均二乗誤差の値は非観測ユニット数 U の増加に伴い減少した (Spearman rank correlation test; $p < 1.0e-13$) (図 3.2b, 青線) . 推定したネットワークパラメータを用いて学習していないデータを生成させた場合 (図 3.2b, 緑線) にも平均二乗誤差は非観測ユニット数 U の増加に伴い減少したことから ($p < 1.0e-13$) , いずれの非観測ユニット数を用いて推定した結果についても過学習は起こっていないと考えられる. また、推定した結合強度が特定のデータに依存して得られたものではないことを確認するため、異なるデータセットから得た観測ユニット間の結合強度について比較を行った. 4つの異なるデータグループから得た結合強度の平均値は、特定の皮質脳波のチャンネルにおいて統計的に有意にゼロではなかった ($p < 0.0001$) (表 3.1) . つまり、異なるデータを用いても特定の電極の組にゼロではない結合強度 W_{ij} が得られた.

次に、我々は異なるデータグループから得た 4つの結合行列を平均し、ネットワークの構造について解析を行った. あるランダムシードを用いた際の結合行列を図 3.3 に示す. 上段にサル A, 下段にサル K の結果が示されており、各行列は左から、推定に用いた非観測ユニット数 0, 1, 5, 10 の平均結合行列および全教示データの相関係数行列である. 結合行列のパターンは非観測ユニット数 U に依存して異なっていた. $U=0$ の場合、すべての自己結合 W_{ii} が有意にゼロではなく、それ以外のほとんどの結合強度には統計的有意差はなかった. また、 U が増加するに従い、結

合行列のパターンは相関係数行列のパターンと似てくる傾向があった。さらに、少数の非観測ユニットを持つ場合 ($U=5$)，結合行列のパターンにクラスター構造がみられた。クラスターの例として、ユニット 11, 13, 14, 17, 18, 21 は相互に強い

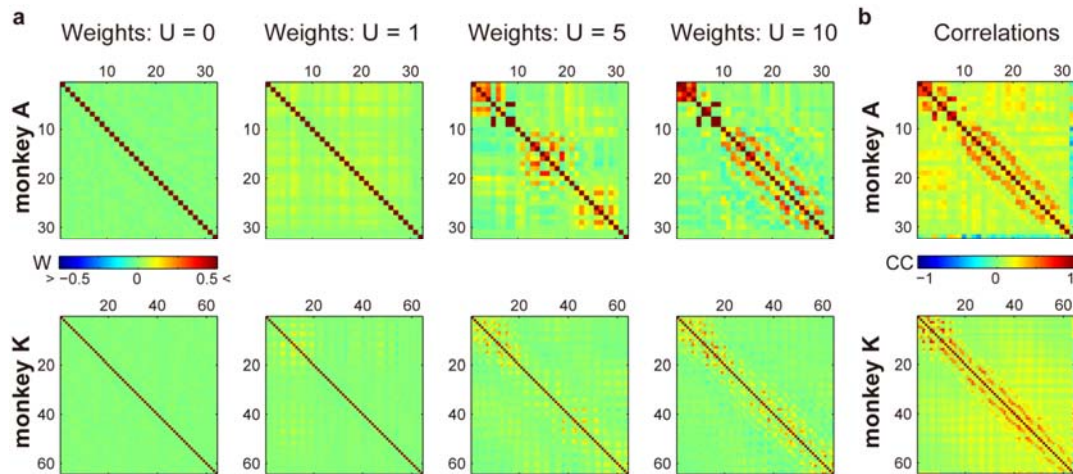


図 3.3 結合行列および相関係数行列，(a)観測ユニット間の結合行列．推定に用いた非観測ユニット数 U は左から 0, 1, 5, 10. (b)皮質脳波間の相関係数行列．x, y 軸のラベルはそれぞれ結合元および結合先の皮質脳波の電極番号に対応している．

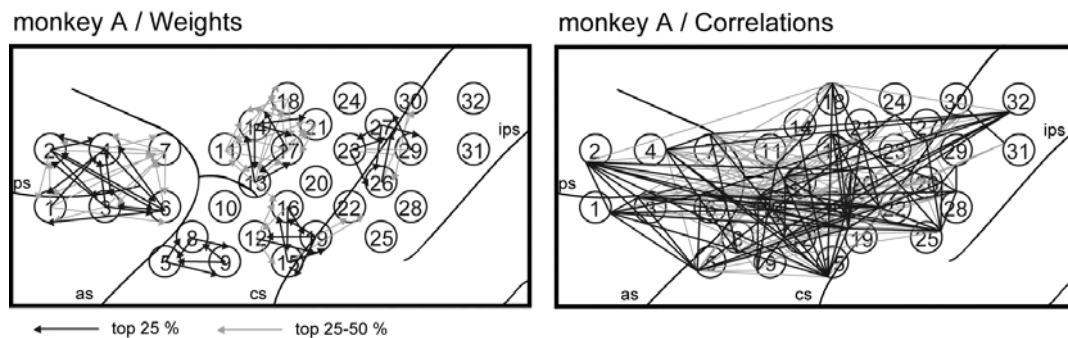


図 3.4 左右の図はそれぞれ統計的に有意であった結合強度と相関係数を皮質脳波の電極位置に表示したものである ($p < 0.0001$)．黒線は上位 25%，灰線は上位 25-50%の値に対応している．実線は皮質の主な溝 (ips, intraparietal sulcus; cs, central sulcus; as, arcuate sulcus; ps, principal sulcus)，点線はそれらの溝を元に決定した皮質領野 (PF, prefrontal area; PMd, dorsal premotor area; PMv, ventral premotor area; M1, primary motor area; S1, primary somatosensory area) の境界である (Matelli et al., 1985; Luppino and Rizzolatti, 2000)．

結合強度を持っているが、他のユニットとの結合強度は小さい。図 3.4 は統計的に有意であった結合強度と相関係数を皮質脳波の電極が留置されている皮質上の位置に表示した図である ($p < 0.0001$)。黒線は上位 25%、灰線は上位 25-50%の値に対応している。図 3.4 の左図を見ると、相互に強い結合強度を持つユニットは単一の皮質領野内に分布していることがわかる。例えば、上述の 11 から 21 の電極は背側運動前野に分布している。それに対し、相関係数は位置的に近いユニット同士で高い値を取り、記録領域全体に分布しており、クラスター構造は認められない(図 3.4 右図)。

結合行列にみられるクラスター構造を定量的に比較するため、観測ユニット間の結合強度に対して階層的クラスター分析を行った。各非観測ユニット数を用いて推定した結合行列に対してクラスターツリーを計算し、1-8 までのクラスター数の非類似度を図 3.5 に示した。図は、非類似度が非観測ユニット数およびクラスター数に依存し、クラスター数が大きな値のときには非観測ユニット数 5 の場合に最も小さな値を取ることを示している。このことは、非観測ユニット数 5 の時に、他の非観測ユニット数よりも小さな分散を持つクラスター構造が得られる、すなわち、各クラスター内のユニットは他の観測ユニットからより似たような影響を受けていることを意味している。したがって、クラスター構造は非観測ユニット数 5 の結合行列においてより顕著に現れている。我々は本研究において、以下では非観測ユニット数 5 の結合行列についてネットワーク構造の解析を行った。

次に、我々は得られたクラスター構造について生理学的妥当性を検証した。図 3.6b は上述のクラスターツリーにより得られたクラスターを皮質脳波電極位置に示したものである。図では観測ユニットを 6 つのクラスターに分割した結果を表示しており、同じシンボルで示されている電極が同じクラスターに分類されている。ここで、クラスター数は、クラスター内のユニット数が 1 となるクラスターが生じな

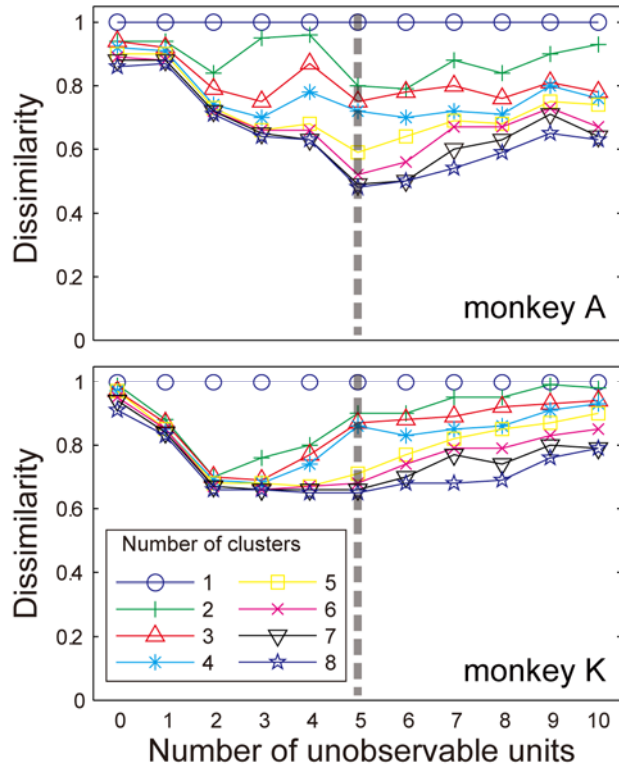


図 3.5 非観測ユニット数に依存した非類似度の変化. より分散の小さなクラスター構造では非類似度の値は小さくなる. 点線で示した非観測ユニット数が5のとき, クラスター数6以上の非類似度は最も小さい. したがって, 非観測ユニット数5でクラスター構造は顕著に現れている.

い最大の数とした (図 3.6a). 十字で表したクラスターは6つの中で最も大きな分散を持つクラスターである. 残りの5つのクラスターの分布は皮質上の溝から決定した皮質領野と概ね一致している. このことを統計的に検証するため, 我々は得られたクラスターの分布と, 電極をランダムにグループ分けした場合のグループの分布との比較を行った. 十字で示されるクラスターを除く5つのクラスターについてはずれユニット数を調べたところ, ユニットの合計数はサルAでは3, サルKでは8であった. さらに, 皮質脳波の電極を, それぞれのサルについて得られたクラスターのユニット数と同数のグループにランダムにグループ分けし, 同様にはずれユニット数を調べた. このランダムなグループ分けを100000回繰り返した際のはずれ

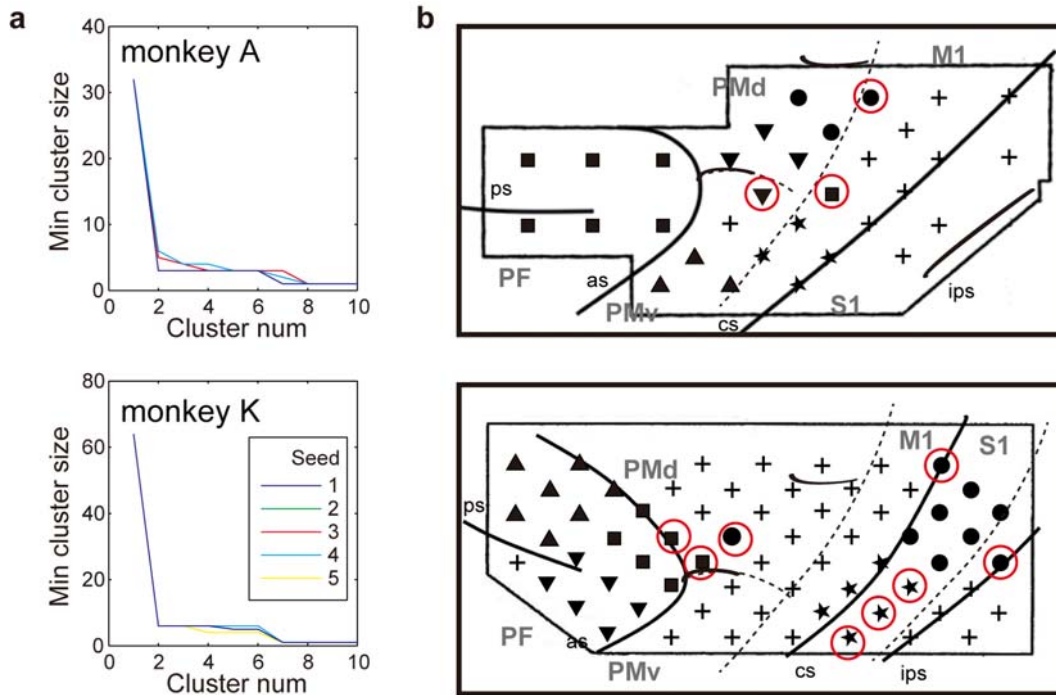


図 3.6 可観測ユニット間の結合行列に基づく皮質脳波電極の分類。(A) クラスタに含まれる最小のユニット数はクラスター数の増加に伴って減少する。サル A (上図), サル K (下図) のどちらの場合にも, クラスタ数 7 のときに単一のユニットしか含まないクラスターが出現する。(B) 6 つのクラスターに分類された皮質脳波電極。それぞれの図形は同じグループに分類された電極位置を示す。赤丸で囲まれた電極はクラスターの対応皮質領野外に位置するはずれユニットの電極である。

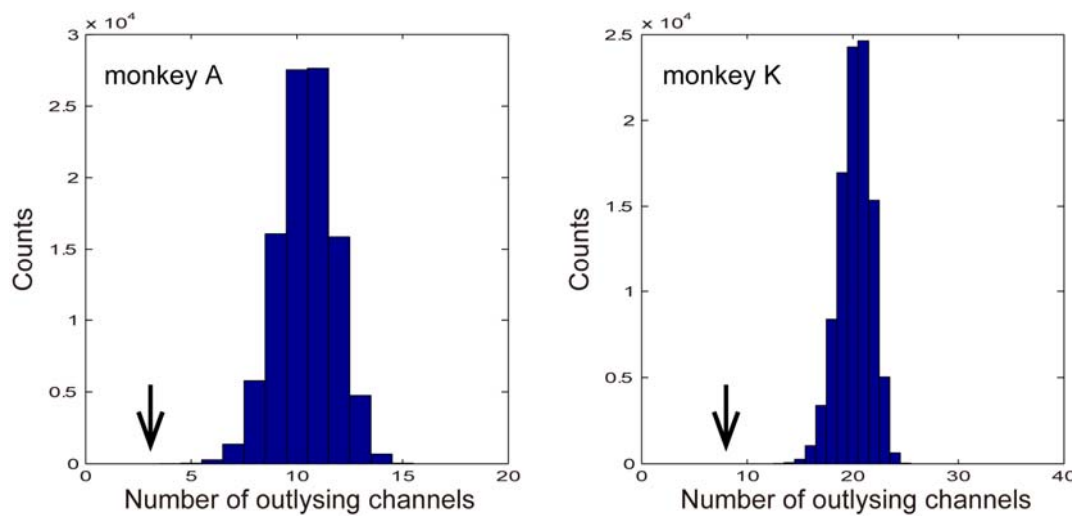


図 3.7 ランダムなグループ分けを 100000 回行って得られたはずれユニットの分布。矢印はそれぞれの個体の推定から得られたはずれユニットの数を示す。

ユニット数の平均値 \pm SD は、サル A では 10.44 ± 1.35 、サル K では 20.23 ± 1.56 であった。また、どちらの個体においても推定から得られたはずれユニット数よりも小さい値は観測されなかったため、推定から得られたはずれユニット数はランダムにグループ分けして得られるはずれユニット数よりも統計的に有意に小さいといえる ($p < 1.0e-5$) (図 3.7)。すなわち、結合行列にみられるクラスター構造はランダムなクラスター構造に比べて皮質領野の分布に一致しているといえる。

3.2.3 考察

ネットワークの出力 \mathbf{Y} と教示データ \mathbf{Z} の平均二乗誤差は学習の経過に従い減少したことから、部分観測ネットワークは皮質脳波を学習することができたと考えられる。また、汎化誤差が学習誤差と同程度に小さかったことから、推定された結果は過学習によるものではなく、皮質脳波の性質を反映したものであると考えられる。異なる学習データを用いても特定の電極の組にゼロではない結合強度が得られたことも、推定された結果が特定のデータに依存しない皮質脳波の何らかの構造を反映しているという結論を支持している。ここで、学習成績およびバリデーションの成績は非観測ユニット数の増加に従って減少していたことから、皮質脳波を再現するためにはより多くの非観測ユニットを用いて学習することが適切であると考えられる。ただし、本研究ではネットワークの構造に焦点を当てたため、皮質脳波の学習成績が劣っていても、クラスター構造が顕著にみられる非観測ユニット数の推定結果について詳細な検証を行った。

我々は異なるデータグループから得た 4 つの結合行列を平均し、データ間で共通してみられるネットワークの構造について解析を行った。結合行列のパターンは非観測ユニット数 U に依存して異なっていた。 $U=0$ の場合、すべての自己結合 W_{ii} が有意にゼロではなく、それ以外のほとんどの結合強度には統計的有意差はなかった。

これは得られたネットワークでは、入力として与えられた時刻 $t - 1$ の信号をそのまま出力していることを意味している。また、 U が増加するに従い、結合行列のパターンは相関係数行列のパターンと似てくる傾向があった。これはネットワークが多数の非観測ユニットを持つ場合、ユニット間の結合は時刻 t における信号間の相関に関する情報を保持できるようになるためと考えられる。さらに、少数の非観測ユニットを持つ場合 ($U=5$)、結合行列のパターンには相関係数行列にはみられなかったクラスター構造がみられた。

結合行列にみられるクラスター構造を定量的に比較するため、観測ユニット間の結合強度に対して階層的クラスター分析を行った。その結果、非観測ユニット数 5 の結合行列においてクラスター内の分散が小さくなっており、クラスター構造はより顕著に現れていることがわかった。さらに、そのクラスター構造は皮質領野に対応するような、神経科学の知見と一致する分布となっていた。このことは、我々の提案するネットワーク推定手法が神経科学的に妥当なネットワークを推定できている可能性を示唆している。

3.3 皮質脳波から推定されたネットワーク構造

本節では、5 つの非観測ユニット数で推定した際の、ネットワーク全体の構造に焦点をあてて解析を行った。まず、教示データやランダムシードに依存しない構造の安定性について評価を行った。もし、異なる教示データやランダムシードを用いても同様のネットワーク構造が推定されたならば、その構造は記録した信号の背後に存在する安定的な構造を反映していると考えられる。さらに、推定したネットワーク構造の機能的な意味について考察を行った。

3.3.1 解析方法

異なる推定から得られた結合行列間でネットワーク構造を比較するためには、非観測ユニットの同質性について考えなくてはならない。すなわち、前節で議論した可観測ユニットは皮質脳波という実体との対応があったため、異なるデータセットやランダムシードから得られた結合行列でも比較が可能であった。しかしながら、非観測ユニットはそのような実体との対応がないため、同じネットワーク構造を持つ結合行列であっても、推定の初期条件によって異なる行列が得られる可能性がある。したがって、異なる結合行列間でネットワーク構造を比較するためには、非観測ユニットの同質性を考慮する必要がある。

非観測ユニットには、2種類の間質性があると考えられる。ひとつは行列のインデックスである。可観測ユニットでは特定のインデックスに特定の皮質脳波の電極が対応しているが、非観測ユニットは学習によって獲得されるため、学習の初期状態などによって、同じ働きをする非観測ユニットが異なるインデックスで学習されることがあり得る。もうひとつは、可観測ユニットとの間の結合強度の符号である。

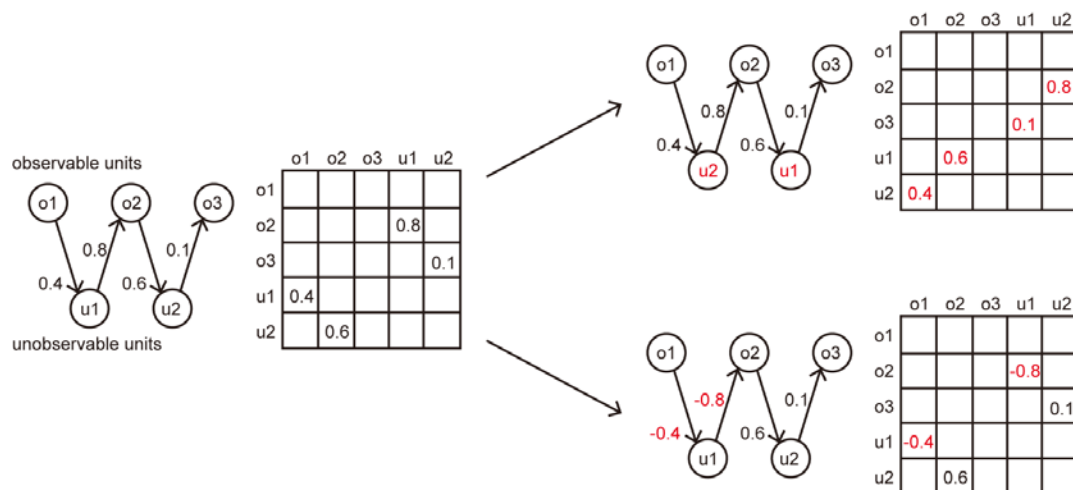


図 3.8 非観測ユニットの同質性. 左図のネットワークの非観測ユニットは、異なる行列のインデックスで学習されることがある (右上図). また、可観測ユニットとの入出力の符号が逆のユニット (左図 u1 および右下図 u1) は可観測ユニットに対して同じ影響を及ぼす (右下図). 右図赤字は左図のネットワークとの相違点.

ふたつの非観測ユニットについて、可観測ユニットとの入出力の符号がどちらも逆ならば、それらの非観測ユニットは可観測ユニットに対して機能的に同じ働きをする。具体的な例を図 3.8 に示す。

このような考察に基づき我々は、異なるデータセットおよびランダムシードから得た 20 個の結合行列に含まれる合計 100 個の非観測ユニットについて、これら 2 種類の同質性に基づいたラベル付けを行った。具体的な手順を、2 つの非観測ユニットを持つ 3 つの結合行列を例に、図 3.9 を用いて説明する。ここでは、合計 6 つの非観測ユニットがあり、それぞれ u_1 から u_6 とする（上段赤字）。まず、可観測ユニットから受ける入力（上段赤枠）の平均値が正となるように非観測ユニットの符

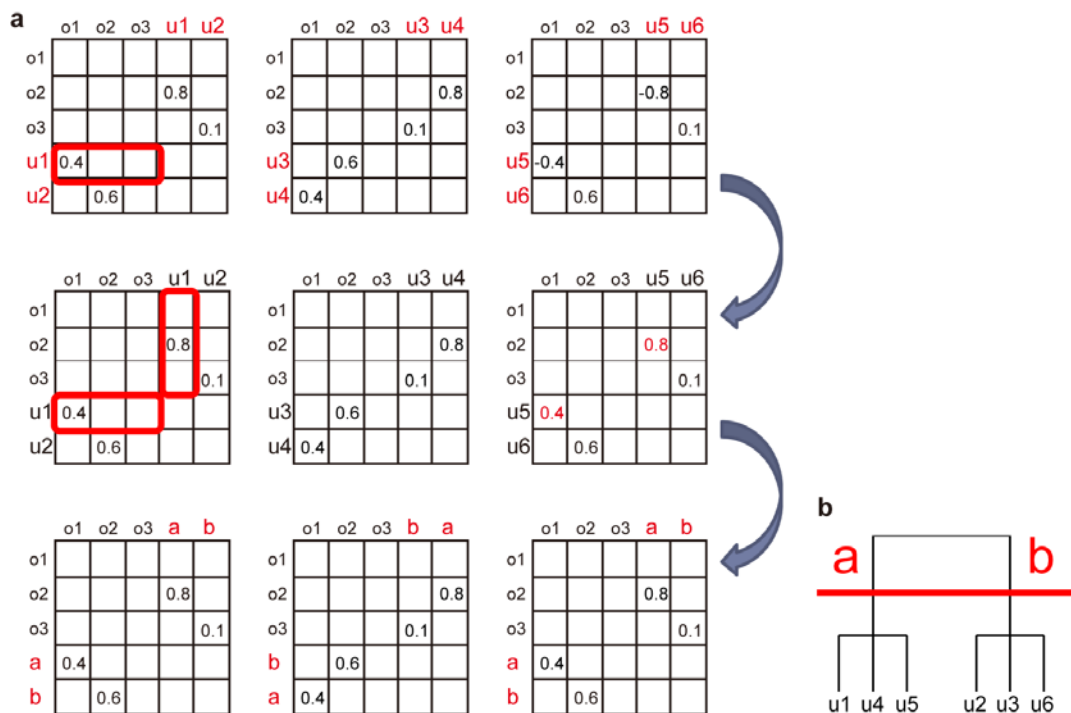


図 3.9 (a) 非観測ユニットのラベリング手順。可観測ユニットからの入力の平均値が正となるよう符号をそろえる（中段）。その後階層的クラスター分析により、可観測ユニットとの結合強度の入出力パターンに基づきラベル付けを行う（下段）。
(b) 階層的クラスター分析によるラベル付け。Ward 法を用いて、可観測ユニットとの結合強度の入出力パターンからクラスターツリーを計算し、非観測ユニット数と同数のクラスターに分け、ラベルを付与する。

号をそろえる．図 3.9a 上段の例では，u5 が可観測ユニットから受ける入力 の平均値は負であったため，符号を反転させている（中段赤字）．次に，Ward 法を用い，可観測ユニットとの結合強度の入出力パターン（中段赤枠）の違いを非類似度としてクラスタツリーを計算する（図 3.9b）．すなわち，ユニット i とユニット j の間の非類似度 D_{ij} は以下で定義される．

$$D_{ij} = \sum_{k=1}^{CH} \|W_{ik} - W_{jk}\|^2 + \sum_{k=1}^{CH} \|W_{ki} - W_{kj}\|^2 .$$

その後，計算したクラスタツリーから，推定に用いた非観測ユニット数と同数の 2 つのクラスターに分け，それぞれのクラスターに分類されたユニットに a, b とラベルをつける（下段赤字および図 3.9b）．以上の手順によって異なる結合行列間で非観測ユニットの比較が可能になり，上段の 3 つの結合行列は同じネットワーク構造を表していたことがわかる．

3.3.2 結果

5 つの非観測ユニット数で推定した，異なるデータセットおよびランダムシードから得た 20 個の結合行列について，非観測ユニットのラベル付けを行った．その後，各結合行列に含まれる非観測ユニットの組み合わせについて調べたところ，両個体において，推定したネットワークの大多数で同じ組み合わせの非観測ユニットを用いていることがわかった．サル A では 12/20，サル K では 18/20 の結合行列が非観測ユニット abcde の組み合わせとなっていた（図 3.10 左図）．

さらに，推定したネットワーク構造の機能的特徴を調べるため，各個体において最も多く得られたネットワーク構造の結合行列について解析を行った．いくつかの特徴が両個体で共通してみられた．そのうちのひとつは，可観測ユニット同士の結合について，統計的に有意であった結合強度の内約 80% が正の結合であった（two-

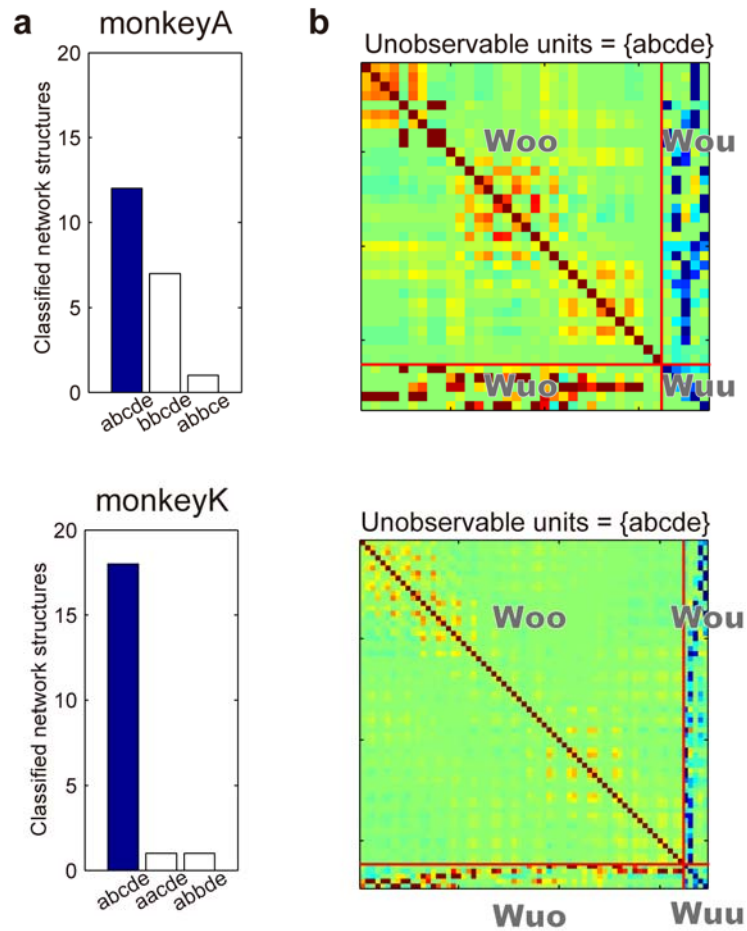


図 3.10 サル A (上) およびサル K (下) から推定された結合行列における非観測ユニットの構成. (a) 塗りつぶした棒線は最も多く得られた非観測ユニットの組み合わせを示す. サル A の例では, 最も頻繁にみられたネットワーク構造は a, b, c, d, e でラベル付けされた非観測ユニットで構成されており, 図中では “abcde” と表されている. 2 番目にみられたネットワークは 2 つの b と c, d, e で構成され, “bbcde” と表されている. (b) 最も頻繁にみられたネットワークの平均結合行列.

tailed t-test, $p < 0.01$) (図 3.11a). また, 非観測ユニット同士の結合強度について, 統計的に有意であった結合強度はすべて負であった (two-tailed t-test, $p < 0.01$) (図 3.11b). さらに, 可観測ユニットから非観測ユニットへの結合および非観測ユニットから可観測ユニットへの結合の間には負の相関があった (サル

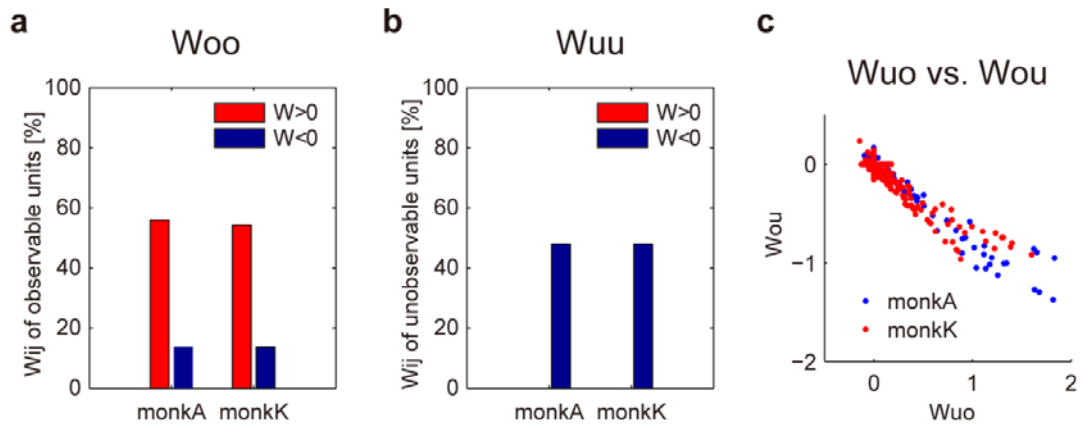


図 3.11 推定したネットワークの特徴. (a) 可観測ユニット間の結合強度 (W_{oo}) の符号. (b) 非観測ユニット間の結合強度 (W_{uu}) の符号. (c) 可観測ユニットから非観測ユニットへの結合強度 (W_{uo}) および非観測ユニットから可観測ユニットへの結合強度 (W_{ou}) は逆相関していた.

A およびサル K について相関係数はそれぞれ -0.97 および -0.94 であった. Pearson's linear correlation; $p < 0.01$) (図 3.11c).

3.3.3 考察

推定したネットワークの大多数で同じ組み合わせの非観測ユニットを用いていたことから、推定結果は学習データおよびランダムシードに対して頑健であるといえる.

両個体でもっとも多く得られたネットワークには 3 つの特徴が共通してみられた. それは、(1) 可観測ユニット間の結合の約 80% が正の値をとる、(2) 非観測ユニット間の結合が全て負の値をとる、(3) 可観測ユニットと非観測ユニット間の結合強度が反相関、というものである. この内、2 番目の結果は、可観測ユニットからの入力 of 平均値が正となるように結合強度の符号をそろえるという解析手法に依存して得られたものである. もし可観測ユニットからの入力 of 平均値が負となるよう

に符号をそろえたならば，非観測ユニット同士の結合強度は自己結合を除いて正の値となる．ただし，大脳皮質においてもっとも普遍的にみられる細胞は興奮性の出力を持つ錐体細胞であることから，可観測ユニットからの入力为正となるという仮定は生理学的に妥当なものであると考えられる．

これらの特徴は，興奮性および抑制性の神経細胞集団の逆相関を持った相互作用により皮質の動的状態が表されることを示唆している．このネットワークは，最近いくつかのグループが提唱している皮質の局所回路モデルとよく似ている（Bartos et al., 2007; Deco et al., 2009; Deco et al., 2011）．そのモデルは，皮質内のネットワークにおいて興奮性の細胞集団と抑制性の細胞集団が互いに相互作用し，その結果脳の周期活動が実現されるというものである（図 3.12）．推定したネットワークでは，興奮性細胞集団と細胞集団はそれぞれ可観測ユニットと非観測ユニットで表されていた．このことから，非観測ユニットは皮質脳波の高ガンマパワー値

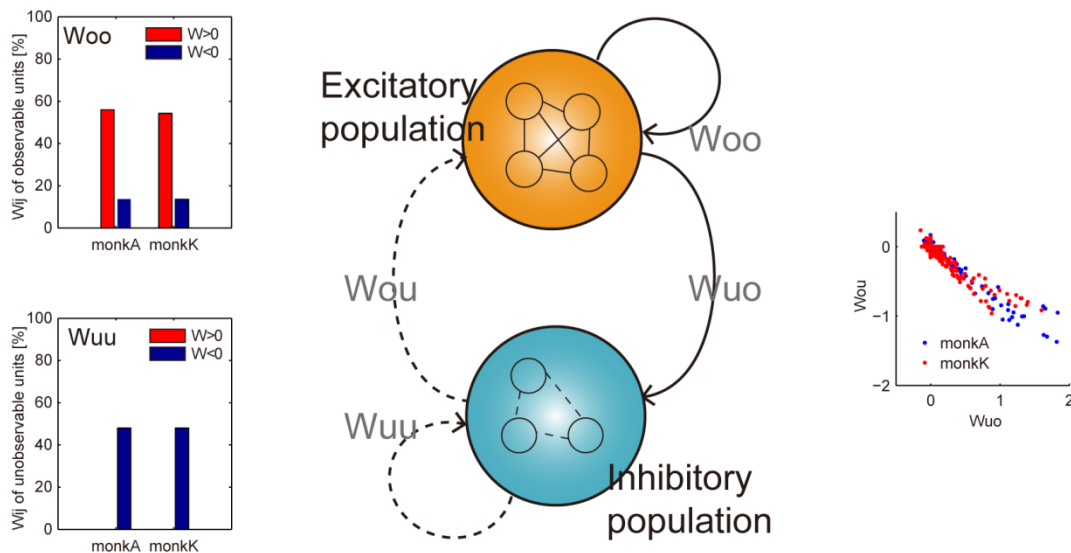


図 3.12 推定したネットワーク構造の機能的解釈．推定したネットワークの特徴の概略は図のように表される．このような構造のネットワークは周期的な信号の生成器として知られている．実線および点線はそれぞれ興奮性および抑制性の結合を表している．

における周期的な活動の再構成を助けている可能性がある。

また、我々が人工的に生成した周期的な時系列データを用いて行ったネットワークの推定の際には、非観測ユニットは必ずしも抑制性というわけではなかったので、可観測ユニットおよび非観測ユニットの興奮性および抑制性への分化はモデルから自動的に導かれる結果ではなく、皮質脳波の性質を反映していると考えられる。この結果のひとつの説明として、皮質脳波は主に皮質内の興奮性の活動を記録しているという報告があり (Buzsaki et al., 2012) , 非観測ユニットは皮質脳波では計測が困難な、皮質の介在ニューロンや皮質下の構造体の抑制性の活動を補完している可能性があると考えている。

本節の結果をまとめると、推定したネットワーク構造はデータセットやランダムシードに対して頑健であった。また、そのネットワーク構造は、皮質内のネットワークにより脳の周期活動を実現するという、既存の神経科学の知見に即したものになっていた。その知見と照らし合わせて考えると、非観測ユニットは観測していない脳活動を補完している可能性が示された。

第4章 結論

4.1 まとめ

本論文では、部分観測下におけるネットワーク推定手法の提案を行った。その後、提案した手法を多チャンネル同時記録された脳信号に適応し、以下の結果を得た。推定したネットワーク構造はデータセットやランダムシードに対して頑健であった。また、皮質脳波間の結合強度は、皮質内の脳領野の知見と一致していた。このことから、提案した手法が生理学的に妥当な脳信号間の関係性を推定している可能性が示された。さらに、得られたネットワーク構造は、皮質内のネットワークにより脳の周期活動を実現するという、既存の神経科学の知見に即したものになっていた。このことから、提案した手法は記録した信号間の関係だけでなく、信号の背後にあるネットワーク構造についても生理学的に妥当な解を得ている可能性が示された。

結論として、本研究は、部分観測ネットワークを用いた時系列学習を通してネットワークを推定するという手法が、脳信号の背後にある脳内ネットワークを推定できる可能性を示すことができたと考えている。

4.2 本研究の意義

本研究で提案する手法は、神経科学の分野に非観測動的因子の影響を含めたネットワーク構造の推定という視点を提示するだけでなく、行動と脳活動をつなぐ巨視的なモデル化という視点も提示している。

序論でも述べたように、観測していない因子を陽に仮定したネットワークの推定は、音声処理や遺伝学の分野では行われているが、神経科学ではまだない。しかし

ながら、脳全体のはたらきを理解するためには、非観測因子の影響を考慮することは絶対的な要請である。このような視点で脳信号を見直すことは、従来の解析では明らかにならなかった新たな知見を神経科学の分野にもたらすはずである。

また、現在の神経科学では、実際の神経活動と行動の間をつなぐ理論が存在しない。筆者は物理学における熱力学のように、行動と脳活動をつなぐ巨視的なモデルが必要であると考えている。そのような動機から、本研究ではリカレントニューラルネットワークを皮質脳波のモデルとして採用している。本研究で用いるリカレントニューラルネットワークはカオスなどの複雑なダイナミクスを実現できることが知られており (Funahashi and Nakamura, 1993) , 知能ロボティクスの分野ですでにロボットの認知行動を実現している (Paine and Tani, 2005; Yamashita and Tani, 2008b) 。一方で、皮質脳波は電極近傍の神経活動の足し合わせを計測していると考えられ (Buzsaki et al., 2012) , それはリカレントニューラルネットワークの各要素の状態として表現することができる。そして、皮質脳波は全皮質レベルでの多チャンネル同時計測が可能であり、腕位置のデコーディングの研究などから高精度の行動情報を持っていることが示されている (Chao et al., 2010) 。したがって、本研究によりリカレントニューラルネットワーク上に皮質脳波の動的性質が再構築できれば、行動と脳活動をつなぐ巨視的なモデルを提供することができると考えている。

本研究で提案したネットワーク推定手法は皮質脳波に限局したものではないため、異なるネットワークモデルを仮定することで信号処理や社会学などの分野にも応用可能である。また、獲得されたネットワークは学習した信号の時系列を自律的に生成可能なため、信号の再現精度を向上させることで数時刻未来の信号予測ができ、ブレインマシンインターフェイスなどのアプリケーションへの応用も可能であると考えられる。さらに、本研究で推定するネットワーク構造は知能ロボティクスで用

いられるリカレントニューラルネットワークとして使用可能であり，ロボットプラットフォームの向上に直接的に貢献できると考えられる．

4.3 今後の展望 ー構造から機能へ

本論文で紹介した研究によって，新たな疑問が生じた．例えば，なぜ非観測ユニットが 5 つの場合にクラスター構造が顕著に現れたのだろうか．また，本論文で用いた手法により推定された結合強度と従来脳信号間の機能結合の指標として用いられてきた granger causality や transfer entropy とはどのように異なるのであろうか．これらの問いに答えていくことで，全脳レベルにおける脳の情報処理機構の理解が進むと期待している．

我々は本研究において，部分観測ネットワークを到達運動課題遂行中のサルのエCoG に適用している．しかしながら使用した皮質脳波は二個体で電極数・留置部位が異なるため個体間での比較が困難であった．したがって今後の研究では，同じ部位に留置した電極から記録した皮質脳波を用いてネットワークの構造の推定を行い，個体間で共通してみられる特徴を明らかにしたい．個体間での比較を行うことは，認知行動を生成する脳回路に共通してみられる性質について理解を深めるために重要である．特に，結合強度および時定数の分布に着目し，皮質脳波記録部位での局所領域の機能表現・階層性について考察を深めたい．

知能ロボティクスの分野においては，認知行動における汎化および階層的コントロールという基本的な認知機能について，ネットワーク構造が重要な影響を及ぼすことが示されている．例えば，リカレントニューラルネットワークを用いたロボットシミュレーションにより，ネットワークのボトルネック構造が環境の汎化学習を促進するという報告がある (Paine and Tani, 2005) ．Paine らはボトルネック構造を持つネットワークおよび全結合のネットワークを用いてロボットに迷路を学習

させた (図 4.1a) . 迷路には 8 つのゴールがあり, ロボットは 12 回の試行の中でできるだけ多くのゴールへ到達することを学習する (図 4.1b) . 20 回の独立な学習の結果, 全結合ネットワークで学習したゴールの数は平均 2.3 であったのに対し, ボトルネック構造を持つネットワークでは平均 5.1 のゴールを学習した. さらに, ボトルネックユニットの活動の位相図からそれらの活動が右折・左折という, 実際のアクチュエータの動きよりも高度な表現となっていることが示唆された (図 4.1c) . 脳信号から部分観測ネットワークを用いて推定したネットワークにもボトルネック構造がみられるのだろうか. グラフ理論などで用いられる解析手法を推定した結合強度へ適用することで, この問いについて検証が可能であると考えている.

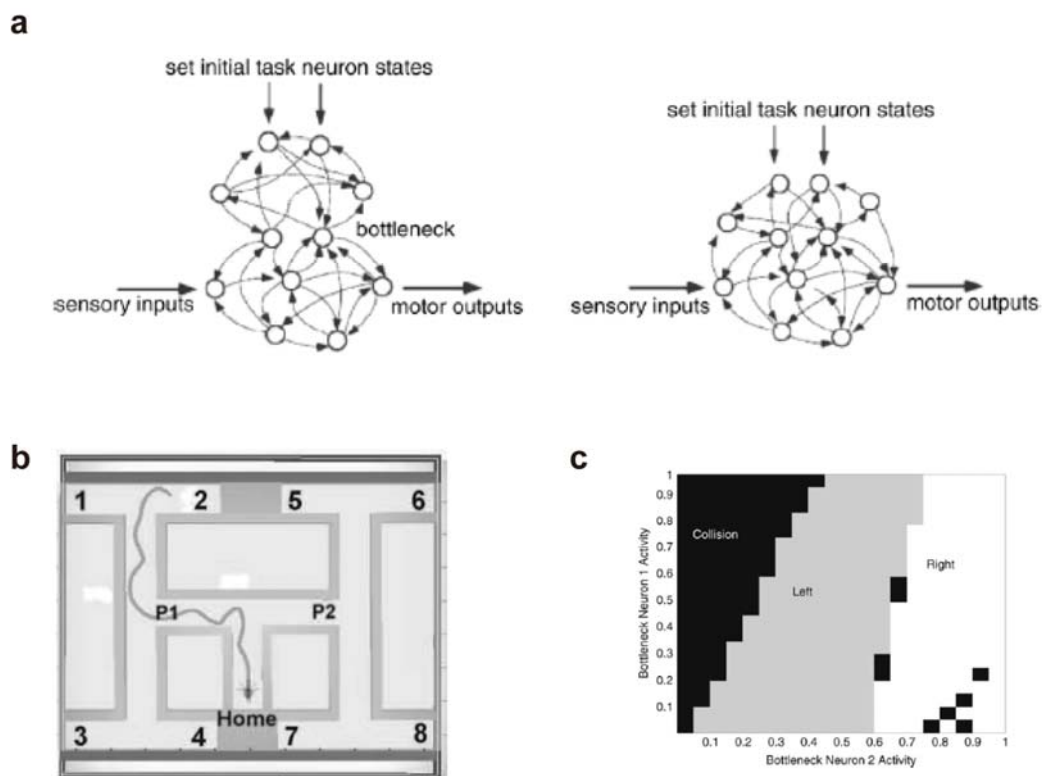


図 4.1 Paine ら (2005) の研究の概要. (a) ボトルネック構造を持つネットワーク (左) および全結合ネットワーク (右) . (b) ロボットは 12 試行の間にホームポジションからできるだけ多くのゴールへ移動することを学習する. (c) ボトルネックユニットの活動の位相図.

また、同じくりカレントニューラルネットワークを用いたロボットシミュレーションにおいて、Yamashita ら (2008b) は異なる時定数を持つ局所回路の存在によって行動の階層化が実現されることを示している (図 4.2) . Yamashita らの提案した Multiple Timescale Recurrent Neural Network では異なる時定数を持つ局所回路で構成されている. このようなネットワークを用いて, ロボットに 5 種類の運動要素を組み合わせた行動を学習させたところ, Fast ユニットに運動要素が学習され, Slow ユニットは Fast ユニットの制御して要素の組み合わせを制御するという, 機能の階層化が実現された. 本研究では結合強度のみに焦点を当てて解析を行ったが,

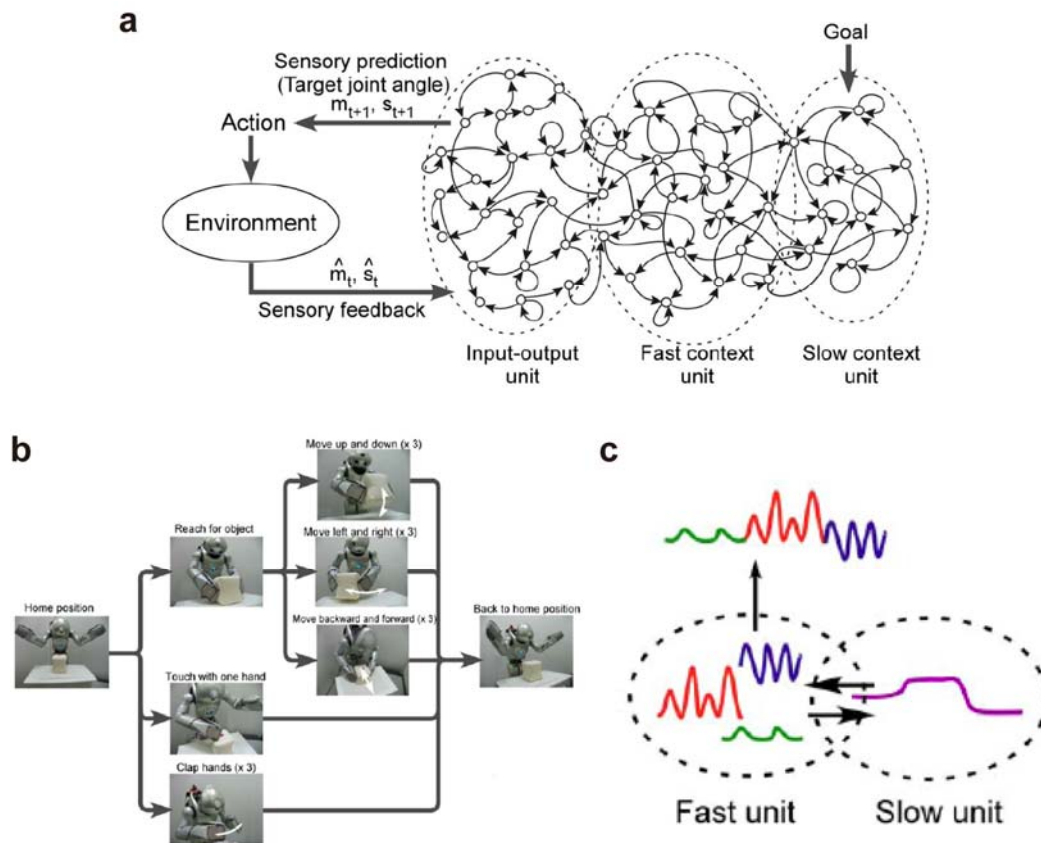


図 4.2 Yamashita ら (2008b) の研究の概要. (a) Multiple Timescale Recurrent Neural Network. (b) ロボットは 5 通りの運動要素の組み合わせを学習する. (c) Fast ユニットに運動要素が学習され, Slow ユニットは要素の組み合わせを学習する.

Yamashita らの提案したような機構が脳の情報処理にも働いているのかどうか，皮質脳波から得られたネットワーク内における時定数の分布を調べ，時定数の異なる局所ネットワークの存在を検証したい．このように，部分観測ネットワークにより推定した脳回路と，従来の人工ニューラルネットワークを用いた研究から得られた知見とを比較することで，認知行動を生成するネットワーク構造の普遍的な性質について理解が進むと期待している．

謝辞

この研究は多くの方々が支えてくださったおかげでここまでまとめることができました。ここに、心より感謝の意を表します。

東京工業大学大学院総合理工学研究科中村清彦教授には、筆者が同大学院修士課程に入学してから理化学研究所への就職を経て現在へと至るまでの紆余曲折の過程で、長年にわたりご指導とご教示を賜りました。また、樺島祥介教授、渡邊澄夫教授、青西亨准教授、豊泉太郎連携准教授には、博士論文の審査にあたり有益なご討論ご意見を頂きました。深く感謝いたします。

本研究は、筆者が理化学研究所脳科学総合研究センターにて行ったもので、研究の機会を与えていただき、日頃からご支援ご理解をいただいている同研究センター動的認知行動研究チーム谷淳チームリーダー、適応知性研究チーム藤井直敬チームリーダーに感謝いたします。とくに、動的認知行動研究チームの谷チームリーダーおよび並川研究員には、モデル構築・計算機実験の手法などについて研究の初期から多大なる尽力をいただきました。皮質脳波を提供してくれた適応知性研究チームのCiao研究員、長坂研究員は研究内容についてよき話し相手であり、本研究について貴重な俯瞰的意見を与えてくれました。どうもありがとうございました。

最後に、いつも精神的支えとなってくれたパートナー、えいさんの理解と支援に深く感謝いたします。どうもありがとう。

参考文献

- Bartos, M., Vida, I., Jonas, P., 2007. Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nat. Rev. Neurosci.* 8, 45-56.
- Baum, L.E., Petrie, T., Soules, G., Weiss, N., 1970. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Statist.* 41, 164-171.
- Beer, R.D., 1995. On the dynamics of small continuous-time recurrent neural networks. *Adapt. Behav.* 3, 469-509.
- Brea, J., Senn, W., Pfister, J.P., 2011. Sequence learning with hidden units in spiking neural networks. *Adv. Neural Inf. Process Syst.* 24, 1422-1430.
- Bressler, S.L., Coppola, R., Nakamura, R., 1993. Episodic multiregional cortical coherence at multiple frequencies during visual task performance. *Nature* 366, 153-156.
- Buzsáki, G., Anastassiou, C.A., Koch, C., 2012. The origin of extracellular fields and currents -EEG, ECoG, LFP and spikes. *Nat. Rev. Neurosci.* 13, 407-420.
- Chao, Z.C., Nagasaka, Y., Fujii, N., 2010. Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkeys. *Front. Neuroeng.* 30, 3.
- Chapin, J.K., Moxon, K.A., Markowitz, R.S., Nicolelis, M.A., 1999. Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *Nat. Neurosci.* 2, 664-70.
- Cobb, S.R., Buhl, E.H., Halasy, K., Paulsen, O., Somogyi, P., 1995. Synchronization of neuronal activity in hippocampus by individual GABAergic interneurons. *Nature* 378, 75-78.

- Deco, G., Jirsa, V., McIntosh, A.R., Sporns, O., Kotter, R., 2009. Key role of coupling, delay, and noise in resting brain fluctuations. *Proc. Natl. Acad. Sci. U. S. A.* 106, 10302-10307.
- Deco, G., Jirsa, V.K., McIntosh, A.R., 2011. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nat. Rev. Neurosci.* 12, 43-56.
- Elman, J., 1990. Finding structure in time. *Cogn. Sci.* 14, 179-211.
- Funahashi, K., Nakamura, Y., 1993. Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Netw.* 6, 801-806.
- Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424-438.
- Griffith, J.S., Horn, G., 1963. Functional coupling between cells in the visual cortex of the unrestrained cat. *Nature* 199, 876-895.
- Hesse, W., Moller, E., Arnold, M., Schack, B., 2003. The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. *J. Neurosci. Methods* 124, 27-44.
- Jacobs, R.A., Jordan, M.I., Nowlan, S.J., Hinton, G.E., 1991. Adaptive mixtures of local experts. *Neural Comput.* 3, 79-87.
- Jordan, M.I., Rumelhart, D.E., 1992. Forward models: Supervised learning with a distal teacher. *Cogn. Sci.* 16, 307-354.
- Kaminski, M., Ding, M., Truccolo, W.A., Bressler, S.L., 2001. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biol. Cybern.* 85, 145-157.
- Leuthardt, E.C., Schalk, G., Wolpaw, J.R., Ojemann, J.G., Moran, D.W., 2004. A brain-computer interface using electrocorticographic signals in humans. *J. Neural Eng.* 1, 63-71.
- Luppino, G., Rizzolatti, G., 2000. The organization of the frontal motor cortex. *News Physiol. Sci.* 15, 219-224.

- Nagasaka, Y., Shimoda, K., Fujii, N., 2011. Multidimensional Recording (MDR) and Data Sharing: An Ecological Open Research and Educational Platform for Neuroscience. *PLoS One* 6, e22561.
- Matelli, M., Luppino, G., Rizzolatti, G., 1985. Patterns of cytochrome oxidase activity in the frontal agranular cortex of macaque monkey. *Behav. Brain Res.* 18, 125-137.
- Namikawa, J., Tani, J., 2008. A model for learning to segment temporal sequences, utilizing a mixture of RNN experts together with adaptive variance. *Neural Netw.* 21, 1466-1475.
- Paine, R.W., Tani, J., 2005. How hierarchical control self-organizes in artificial adaptive systems. *Adapt. Behav.* 13, 211-225.
- Pillow, J., Latham, P., 2008. Neural characterization in partially observed populations of spiking neurons. *Adv. Neural Inf. Process Syst.* 20, 1161-1168.
- Ray, S., Crone, N.E., Niebur, E., Franaszczuk, P.J., Hsiao, S.S., 2008. Neural correlates of high-gamma oscillations (60-200 Hz) in macaque local field potentials and their potential implications in electrocorticography. *J. Neurosci.* 28, 11526-11536.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning internal representations by error propagation. In: Rumelhart, D.E., McClelland, J.L. (Eds.), *Parallel Distributed Processing*. MIT Press, Cambridge, MA, pp. 318-362.
- Schreiber, T., 2000. Measuring information transfer. *Phys. Rev. Lett.* 85, 461-464.
- Smith, V.A., Yu, J., Smulders, T.V., Hartemink, A.J., Jarvis, E.D., 2006. Computational inference of neural information flow networks. *PLoS Comput. Biol.* 2, e161.
- Tani, J., Nolfi, S., 1999. Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems. *Neural Netw.* 12, 1131-1141.

- Viterbi, A.J., 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans. Inf. Theory* 13, 260-269.
- Wessberg, J., Stambaugh, C.R., Kralik, J.D., Beck, P.D., Laubach, M., Chapin, J.K., Kim, J., Biggs, S.J., Srinivasan, M.A., Nicolelis, M.A., 2000. Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature* 408, 361-5.
- Yamashita, Y., Tani, J., 2008a. Emergence of functional hierarchy in a multiple timescale neural network model: a humanoid robot experiment. *PLoS Comput. Biol.* 4, e1000220.
- Yamashita, Y., Takahashi, M., Okumura, T., Ikebuchi, M., Yamada, H., Suzuki, M., Okanoya, K., Tani, J., 2008b. Developmental learning of complex syntactical song in the Bengalese finch: a neural network model. *Neural Netw.* 21, 1224-1231.
- Yamashita, Y., Okumura, T., Okanoya, K., Tani, J., 2011. Cooperation of deterministic dynamics and random noise in production of complex syntactical avian song sequences: a neural network model. *Front. Comput. Neurosci.* 5, 18.