

修士論文

レーザーネットワークを用いた
コミュニティ検出
Community Detection by Using
Laser Network

48-146419 坂口潤将

指導教員 安達淳 教授

2016年2月

東京大学大学院情報理工学系研究科電子情報学専攻

概要

ソーシャルネットワーク、脳機能ネットワーク、タンパク質ネットワークをはじめとする様々な現実世界のネットワークはコミュニティ構造を形成している。コミュニティとは、ネットワーク上のノードのグループのうち、グループ内では密に結合しており、グループ間では疎に結合しているものを指す。与えられたネットワークからこのようなコミュニティ構造を特定することは、ノードの分類、ネットワークの粗視化、ネットワークのもつ機能解明といった点で様々な分野において重要な問題である。本論文では、多数のレーザーを相互に注入してネットワークを形成することによって、それらの位相同期を利用した新しいコミュニティ検出手法を提案する。コミュニティ構造がネットワーク上に配置されたレーザーの挙動に影響を与えることを利用して、同一コミュニティ内でのみ位相の相関が大きくなるように発振させる。数値シミュレーションによって提案した手法を評価したところ、多くの既存アルゴリズムと比べても高い精度かつ短時間でコミュニティを検出できる可能性が示された。

目次

第 1 章	導入	1
1.1	複雑ネットワークとコミュニティ構造	1
1.2	コミュニティ検出のアプローチ	2
1.3	本研究の概要及び目的	3
第 2 章	レーザーのダイナミクスを記述する理論モデル	5
2.1	原子系のダイナミクス	5
2.2	レーザーのダイナミクスを表す量子ランジュバン方程式	6
2.3	複素ランジュバン方程式	10
2.4	相互注入によるネットワークのダイナミクス	11
第 3 章	コミュニティ検出	14
3.1	コミュニティ	14
3.2	コミュニティ検出問題	15
3.3	評価関数:モジュラリティ	16
3.4	コミュニティ検出アルゴリズム	16
3.5	アルゴリズムの性能評価:Normalized Mutual Information	19
第 4 章	レーザーネットワークを用いたコミュニティ検出	22
4.1	ネットワーク上の結合振動子の同期にコミュニティ構造が与える影響	22
4.2	レーザーネットワークの構成	24
4.3	コミュニティの特定	25
4.4	考えられる実験的な実装方法	27
4.5	数値シミュレーション	27
第 5 章	性能評価及びパラメータ依存性	32
5.1	LFR ベンチマークグラフ	32
5.2	同位相結合と逆位相結合の注入比	33
5.3	位相相関の閾値	37
5.4	rrNMI による既存アルゴリズムとの性能比較	41

iv 目次

5.5	計算時間	41
第 6 章	結論	48
	謝辭	49
	参考文献	50

第1章

導入

1.1 複雑ネットワークとコミュニティ構造

現実世界に存在する様々な系は、個々の要素及びそれらの相互作用により多様なダイナミクスや機能を生んでいる。個々の要素自体は単純な振る舞いであるにもかかわらず、それらが多数集まり互いに相互作用をなす系となった瞬間に系全体のダイナミクスは著しく複雑になり、多彩な現象を引き起こす。脳やソーシャルネットワークのように個々の要素そのものも非常に複雑なダイナミクスをもっており、それらが多数相互作用しあっているような系であれば、そのマクロな機能や振る舞いはなおさら複雑になる。更に系全体としての振るまいは、要素の繋がり方^{*1}にも大きく依存する。このようなネットワークとしての構造^{*2}を手がかりに、多数の要素が繋がっている系の多様な現象を理解する分野が複雑ネットワークである。

一般にネットワークは個々の要素を表す「ノード」とノード間の繋がりを表す「エッジ」から構成される「グラフ」によって表される。グラフ理論自体は古くから盛んに研究されており、組み合わせ最適化問題等のアルゴリズムや、データ構造など、情報科学の分野で幅広く応用されている。また、これまで多く研究されてきたグラフ構造は木構造、ランダムグラフ (Erdős-Rényi モデル)[1]、格子状等の単純な構造が主であった。現実に存在するネットワークは、単純にランダムに繋がっているだけではなく、より複雑な構造をもっていることが指摘されてきたが、ネットワーク構造の解析といった意味でのグラフの研究は特に進展がなかった。ブレークスルーが起きたのは1998年で、極めて単純なアルゴリズムによって、クラスター性^{*3}とスモールワールド性^{*4}を同時に兼ね備えたグラフ生成モデルである、Watts-Strogatz モデルが発表された [2]。更に翌年には Barabási-Albert モデルと呼ばれるスケールフリー性^{*5}をもったランダムグラフを生成するモデルも発表され [3]、これらは従来のものとは比べてより現実世界のネットワークの性質を反映したグラフモデルであるとして注目を浴びた。これ

*1 単純な格子に繋がっているのか、そのような構造をもたずにランダムに繋がっているのか、あるいは構造はもっているけれども複雑に繋がっているのか、など。

*2 ネットワークのトポロジーという言い方をする場合もある。

*3 あるノードの繋がっているノード同士が、また繋がっている割合が高い性質。

*4 ネットワーク上の任意のノード間は少数ノードしか介していないという性質。

*5 ノードの度数分布が冪分布となる性質。

らの研究を皮切りに、現実世界の様々な現象についてネットワークを通して理解しようとするという流れが急速に高まった。

Wattz-Strogatz モデル以来、約 15 年間で目覚ましい発展を遂げているネットワーク科学であるが、その中でもネットワークの構造として近年注目されているのがコミュニティ構造である。コミュニティ構造は密に結合しているノードの集合同士が互いに疎に結合しているようなグラフ構造であり、このような性質はソーシャルネットワーク [4]、脳機能ネットワーク [5, 6]、タンパク質ネットワーク [7]、代謝ネットワーク [8] 等、現実の様々なネットワークに現れる。一般的にコミュニティは似た性質を持つノードの集まりであることが多いと考えられている。例えばソーシャルネットワークであれば同じ会社に属していたり、同じ趣味をもった人同士の集まりであったりする。また、脳機能ネットワークやタンパク質ネットワークの場合は同じネットワークの中で同じ機能をもつ神経細胞集団であったり、タンパク質の集まりであったりする。

与えられたネットワークからこのようなコミュニティ構造を特定することは、以下の3つの点で重要である [9]。一つ目はノードの分類、いわゆるクラスタリングである。ネットワーク解析において、あるノードがどのようなクラスに属しているのかを知りたい状況は多々存在する。例えばマーケティングにおいて、顧客のネットワークの中で共通の趣向を持っているノードの集合を検出し、そのコミュニティの繋がり方を解析することは、自動推薦プログラムなどの開発に有効である。二つ目はネットワークの粗視化である。ノード数が大きくなっていくと繋がりが複雑で、ネットワーク全体を見たときにはその本質が見えない場合がある。このようなときに、コミュニティをノードとしてその繋がりを記述したネットワークを新たに構成することによって、本質的な構造が見える場合がある。3つ目はネットワークの機能解明^{*6}である。ネットワークの機能はコミュニティが重要な役割を担っていることが多い。例えば脳機能ネットワークでは脳の部位をノードとしてその活動の相関の有無でネットワークを構成しているが、異なるタスクや、また健常者と何らかの障害を持っている人でコミュニティ構造が異なる。脳機能と脳活動の関係は未だ未解明な部分が多く、その解明のためにネットワークによるアプローチは多数なされているが、コミュニティは一つのキーポイントと考えられている [5, 6]。これらの三つのコミュニティ検出の重要性は互いに重なっていることも多い。

1.2 コミュニティ検出のアプローチ

コミュニティ検出はノードの結合の仕方、すなわちネットワークの構造のみから密に結合しているノードの集合にネットワークを分割する問題である。3章で詳述するが、コミュニティの定義は目的によって様々であり複雑ネットワーク分野全体で広く受け入れられている定量的な定義は存在しない [10]。従ってコミュニティ検出アルゴリズムも多数開発されているが、基本的にはアルゴリズム独自の指標に基づくものであり、異なるアルゴリズムの定量的な評価は非常に難しい。また、アルゴリズムのアプローチも様々なものがあるが、基本的にはグラフ理

^{*6} ネットワークの構造に起因する、ネットワーク上で生じるあらゆる現象のことをここでは機能と呼んでいる。

論的に目的関数を定義してそれを最適化する種類のものが多い。多くのネットワークに対してこのようなアプローチは有効であり、実際に多く用いられている。しかしながら、多くの組み合わせ最適化問題とコミュニティ検出が異なる点は、その目的が目的関数の最適化そのものではないということである。すなわち「目的関数を最適化するコミュニティ分割」=「目的が満たされるのコミュニティ分割」という単純な関係が成り立っていない。そもそも、ネットワークの種類によって目的が異なるコミュニティを、ネットワークの背景にある情報を無視して、そのトポロジーのみから検出するという自体に限界があると考えられる。しかし、そのような限界があることは十分認識した上で、それでもその構造から意味のあるコミュニティを特定するというのが現在のコミュニティ検出の流れである。

また、最適化とは異なったコミュニティ検出のアプローチも存在する。ネットワーク上に、ある構造が存在するという事は、それによってネットワーク上の様々なダイナミクスは少なからず影響を受けるはずである。従って、当然コミュニティ構造がある場合と無い場合のネットワーク上で生じる大域的なダイナミクスや平衡状態は変化する。もちろんコミュニティ検出の目的にはそのようなダイナミクスや平衡状態を深く理解するといったことも含まれるが、逆にそれを利用してコミュニティ構造を検出するというアプローチも考えられる。このような立場からはスピン系 [11]、ランダムウォーク [12]、結合振動子 [13, 14, 15] などのダイナミクスを用いたアルゴリズムが提案されている。

1.3 本研究の概要及び目的

本研究では前節までの内容を踏まえて、ネットワーク構造が実際の物理系に与える影響を利用してコミュニティを検出する立場をとる。実際に物理系を構築して高速にコミュニティ検出を行なうことを視野に入れ、レーザーという周波数が高く素過程が非常に速いダイナミクスを利用する。4章で詳述するが、コミュニティ構造が結合振動子の位相ダイナミクスに与える影響は Arenas et al.[13] でなされている。これは結合振動子が同位相で結合されているときに、全振動子が同期するまでの過程でコミュニティ内でのみ位相相関をもつ領域が、ある時間スケールで安定して現れることを示したものである。レーザーもこのような非線形振動子の一種であるから同様の現象が生じると予想される。本研究では同位相结合だけでなく逆位相结合を導入することによって、コミュニティ内においてレーザーの位相相関が大きくなる状態が最終的に準平衡状態として現れることを利用したコミュニティ検出手法を提案する。本研究の結果は全て数値シミュレーションによるものであるが実際の物理系を実装することを視野にいれており、その具体的な方法についても述べる。

本論文の構成は以下の通りである。まず、第2章で単一レーザーのダイナミクスを記述するランジュバンを方程式を導出し、相互注入の効果を説明する。第3章ではコミュニティ及びコミュニティ検出についての一般論を説明する。また、現在用いられているコミュニティ検出アルゴリズム及びその評価方法を述べる。続いて第4章では、コミュニティ構造上の結合振動子のダイナミクスについての研究を紹介したあと、提案するレーザーネットワークを用いたコミュニティ検出方法を説明する。そして簡単な動作確認のベンチマーク結果を示す。第5章では数

4 第 1 章 導入

値シミュレーションによって、提案した手法を人工的なモデルによってコミュニティが埋め込まれたベンチマークグラフに適用し、評価した結果を報告する。最後に、第 6 章に結論を述べる。

第 2 章

レーザーのダイナミクスを記述する 理論モデル

一概にレーザーといっても気体、固体、半導体レーザーやファイバーレーザーなど様々な種類のレーザーが存在し、その発振方法も CW(Continuous Wave) やパルスなど様々である。しかし、共振器内にある、反転分布を形成した原子集団による誘導放出により、光が増幅されて発振に至るという点は全てのレーザーで同じである。この章では本研究に用いたモデルである、一般的な 4 準位系原子のレーザーのダイナミクスを記述する複素ランジュバン方程式を導出する。

2.1 原子系のダイナミクス

本研究で用いるモデルでは、レーザーは図 2.1 のような順位をもつ原子と共振器内電場から構成される。本説では共振器内電場との相互作用は考えずに、図 2.1 のような 4 準位原子にポンプを行なった際に生じる原子集団のみのダイナミクスを導出する。まず、各準位の演算子は、次式で定義される射影演算子によって表すことができる。

$$\hat{\sigma}_a^i = |a^i\rangle \langle a^i| \quad (2.1)$$

$$\hat{\sigma}_b^i = |b^i\rangle \langle b^i| \quad (2.2)$$

$$\hat{\sigma}_c^i = |c^i\rangle \langle c^i| \quad (2.3)$$

ただし $|a\rangle$ の上に存在する準位は素早く $|a\rangle$ に緩和すると仮定し、ここでは考えない。また、 i は各原子の番号を表し、直交性から次式が成り立つとする。

$$\langle x^i | y^j \rangle = \delta_{i,j} \delta_{x,y} \quad (2.4)$$

$\langle \hat{\sigma}_a^i \rangle$ は原子 i が状態 $|a\rangle$ にいる確率を表し、 $\langle \hat{\sigma}_b^i \rangle$ は状態 $|b\rangle$ にいる確率を表す。各準位に励起された原子は熱浴との相互作用により状態 $|c\rangle$ に γ_a 、 γ_b のレートで緩和する。また、 $|a\rangle$ への上の準位を介した $|c\rangle$ からのポンプは次式によって表される。

$$\hat{\Lambda}_a^i = \lambda_a \hat{\sigma}_c^i = \lambda_a |c\rangle \langle c| \quad (2.5)$$

ただし、ほとんどの原子が $|c\rangle$ に存在するため、励起による影響は受けず $\hat{\sigma}_c^i$ は時間依存しないと仮定する。従って $\hat{\Lambda}_a^i$ も時間依存しない。以上を考慮にいと、 $\hat{\sigma}_a^i$ 、 $\hat{\sigma}_b^i$ のダイナミクスは次式の量子ランジュバン方程式によって表すことができる [16]。

$$\frac{d}{dt}\hat{\sigma}_a^i(t) = \hat{\Lambda}_a^i - \gamma_a\hat{\sigma}_a^i - \gamma_{\text{sp}}\hat{\sigma}_a^i + \hat{F}_a^i(t) \quad (2.6)$$

$$\frac{d}{dt}\hat{\sigma}_b^i(t) = -\gamma_b\hat{\sigma}_b^i + \gamma_{\text{sp}}\hat{\sigma}_a^i + \hat{F}_b^i(t) \quad (2.7)$$

ここで、 γ_{sp} は共振器モードと結合しない自然放出レートであり、 $\hat{F}_{\{a,b\}}^i$ は雑音演算子である。また、各準位の演算子と同様に原子の分極は次式の遷移演算子によって表すことができる。

$$\hat{\sigma}^i = |b^i\rangle\langle a^i| \quad (2.8)$$

$\langle\hat{\sigma}^i\rangle$ は原子 i の双極子モーメントを表す。 $\hat{\sigma}^i$ のダイナミクスは、位相緩和レート γ を用いて次式のように表すことができる。

$$\frac{d}{dt}\hat{\sigma}^i(t) = -(\gamma + i\omega)\hat{\sigma}^i + \hat{f}_\sigma^i(t) \quad (2.9)$$

ただし ω は $|a\rangle$ 、 $|b\rangle$ 間の遷移に対応する周波数であり、 \hat{f}_σ^i は雑音演算子である。これら演算子に対して全ての原子の平均的な演算子を次式のように定義する。

$$\hat{\sigma}_{\{a,b\}} = \frac{1}{N} \sum_{i=1}^N \hat{\sigma}_{\{a,b\}}^i, \quad \hat{\sigma} = \frac{1}{N} \sum_{i=1}^N \hat{\sigma}^i, \quad \hat{\Lambda}_a = \frac{1}{N} \sum_{i=1}^N \hat{\Lambda}_a^i \quad (2.10)$$

雑音演算子も同様に平均化する。また、分極 $\hat{\sigma}$ とその雑音演算子については、次章で共振器との結合を扱う際の便宜上、共振器内電場の周波数 ν を用いて遅く変化する演算子 (slowly varying operator) $\hat{\Sigma} = e^{i\nu}\hat{\sigma}$ 、 $\hat{F}_\Sigma = e^{i\nu}\hat{f}_\sigma$ を導入する。これら平均化された演算子のダイナミクスは、式 (2.6)、(2.7)、(2.9) によって次式のように表すことができる。

$$\frac{d}{dt}\hat{\sigma}_a(t) = \hat{\Lambda}_a - \gamma_a\hat{\sigma}_a - \gamma_{\text{sp}}\hat{\sigma}_a + \hat{F}_a(t) \quad (2.11)$$

$$\frac{d}{dt}\hat{\sigma}_b(t) = -\gamma_b\hat{\sigma}_b + \gamma_{\text{sp}}\hat{\sigma}_a + \hat{F}_b(t) \quad (2.12)$$

$$\frac{d}{dt}\hat{\Sigma}(t) = -\{\gamma + i(\omega - \nu)\}\hat{\Sigma} + \hat{F}_\Sigma(t) \quad (2.13)$$

2.2 レーザーのダイナミクスを表す量子ランジュバン方程式

前節ではポンプ及び熱浴との相互作用による緩和の効果のみを取り入れた、共振器内における原子系のダイナミクスを導出した。本説では、共振器内電場と原子系の相互作用を考慮に入れ、レーザーのダイナミクスを導出する。まず、共振器内電場の生成消滅演算子を \hat{a}^\dagger 、 \hat{a} とすると、原子系との相互作用が無い場合その時間発展は次のように表すことができる。

$$\frac{d}{dt}\hat{a}(t) = -\left(\frac{1}{2}\gamma_c + i\Omega\right)\hat{a}(t) + \hat{f}(t) \quad (2.14)$$

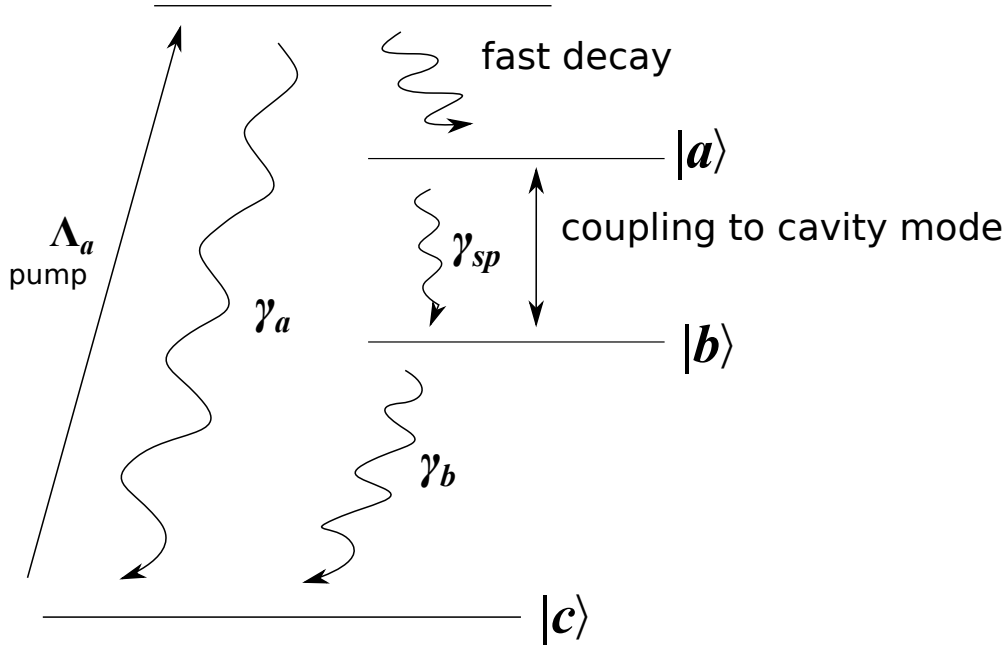


図 2.1. 4 準位系レーザーモデルの概念図

ここで $\gamma_c = \nu/Q$ は共振器の損失レートある。ただし Q は共振器の Q 値である。また Ω は共振器モードの周波数である。そして原子の分極演算子の場合と同様に遅く変化する演算子 $\hat{A}(t) = e^{i\nu t} \hat{a}(t)$ 、 $\hat{F}_A(t) = e^{i\nu t} \hat{f}(t)$ を導入し、式 (2.14) に代入すると

$$\frac{d}{dt} \hat{A}(t) = - \left\{ \frac{1}{2} \gamma_c + i(\Omega - \nu) \right\} \hat{A}(t) + \hat{F}_A(t) \quad (2.15)$$

となる。

次に原子系と共振器電場の相互作用を考える。回転波近似の元での相互作用ハミルトニアンは分極による摂動エネルギーとして次式のように表すことができる。

$$H_{\text{int}} = \hbar g N \Sigma \hat{A}^\dagger + \text{h.c.} \quad (2.16)$$

ただし g は相互作用の強さを表す定数である。これまで議論してきた任意の演算子 \hat{X} のダイナミクスに相互作用の効果を入れるためには $(1/i\hbar)[\hat{X}, H_{\text{int}}]$ による項を付け足せば良い。従って、相互作用の効果を含めた原子系及び共振器内電場のダイナミクスは次式のように書き直すことができる。

$$\frac{d}{dt} \hat{\sigma}_a = \hat{\Lambda}_a - \gamma_a \hat{\sigma}_a - \gamma_{\text{sp}} \hat{\sigma}_a + ig \left(\hat{A}^\dagger \hat{\Sigma} - \hat{\Sigma}^\dagger \hat{A} \right) + \hat{F}_a(t) \quad (2.17)$$

$$\frac{d}{dt} \hat{\sigma}_b = -\gamma_b \hat{\sigma}_b + \gamma_{\text{sp}} \hat{\sigma}_a - ig \left(\hat{A}^\dagger \hat{\Sigma} - \hat{\Sigma}^\dagger \hat{A} \right) + \hat{F}_b(t) \quad (2.18)$$

$$\frac{d}{dt} \hat{\Sigma} = -\{\gamma + i(\omega - \nu)\} \hat{\Sigma} + ig(\hat{\sigma}_a - \hat{\sigma}_b) \hat{A} + \hat{F}_\Sigma(t) \quad (2.19)$$

$$\frac{d}{dt} \hat{A} = - \left\{ \frac{1}{2} \gamma_c + i(\Omega - \nu) \right\} \hat{A} - ig N \hat{\Sigma} + \hat{F}_A(t) \quad (2.20)$$

ここで、式 (2.17)~(2.20) の任意の雑音演算子は以下のような性質をもつ。

$$\langle \hat{F}_X(t) \rangle = 0, \quad \langle \hat{F}_X(t) \hat{F}_Y(t') \rangle = \langle \hat{D}_{XY} \rangle \delta(t - t') \quad (2.21)$$

すなわち、雑音の平均は0であり異なる時間での相関がないマルコフ性を持つ。また、雑音の分散 $\langle \hat{D}_{XY} \rangle$ はアインシュタイン関係式より求めることができる [16]。

次に、原子系の緩和が共振器内電場よりも十分に速いという断熱近似を行なう。

$$\gamma, \gamma_a, \gamma_b \gg \gamma_c \quad (2.22)$$

このような近似のもとでは、原子系の緩和より速い雑音の周波数成分を無視することによって、原子系の時間変化についても無視することができる。すなわち、

$$\frac{d}{dt} \hat{\sigma}_a \simeq \frac{d}{dt} \hat{\sigma}_b \simeq \frac{d}{dt} \Sigma = 0 \quad (2.23)$$

を仮定する。従って、分極演算子は式 (2.19) より

$$\hat{\Sigma} \simeq D(\omega - \nu) \left\{ ig(\hat{\sigma}_a - \hat{\sigma}_b)A + \hat{F}_\Sigma \right\} \quad (2.24)$$

と表すことができる。ただし、 $D(\omega - \nu)$ は

$$D(\omega - \nu) = \frac{1}{\gamma + i(\omega - \nu)} \quad (2.25)$$

である。また式 (2.24) を式 (2.20) に代入することによって、

$$\frac{d}{dt} \hat{A}(t) = - \left\{ \frac{1}{2} \gamma_c + i(\Omega - \nu) \right\} \hat{A}(t) + g^2 N D(\omega - \nu) \hat{A}(t) (\hat{\sigma}_a - \hat{\sigma}_b) + \hat{G}(t) \quad (2.26)$$

が得られる。ここで、 $\hat{G}(t)$ は新たに定義された雑音演算子で、次式のように表される。

$$\hat{G}(t) = \hat{F}_A(t) - igND(\omega - \nu) \hat{F}_\Sigma(t) \quad (2.27)$$

もちろん、式 (2.21) より $\langle \hat{G}(t) \rangle = 0$ である。また、式 (2.17)、(2.18) に現れるコヒーレントな相互作用の項を

$$\hat{B}(t) = ig \left(\hat{A}^\dagger \hat{\Sigma} - \hat{\Sigma}^\dagger \hat{A} \right) \quad (2.28)$$

とおくと、式 (2.24) を代入して

$$\hat{B}(t) = -2g^2 \gamma^{-1} L(\omega - \nu) (\hat{\sigma}_a - \hat{\sigma}_b) \hat{A}^\dagger \hat{A} + \hat{F}'_B(t) \quad (2.29)$$

となる。B のノイズ部分を

$$\hat{F}'_B = igD(\omega - \nu) \hat{A}^\dagger \hat{F}_\Sigma - igD^*(\omega - \nu) \hat{F}_{\Sigma^\dagger} \hat{A} \quad (2.30)$$

と置き、

$$L(\omega - \nu) = \frac{\gamma^2}{\gamma^2 + (\omega - \nu)^2} \quad (2.31)$$

とした。また、断熱近似における時間スケールでは、以下のように $\hat{A}(t)$ と \hat{F}_Σ が同時刻相関をもつ。

$$\langle \hat{A}^\dagger(t) \hat{F}_\Sigma(t) \rangle = igND^*(\omega - \nu) \langle \hat{D}_{\Sigma+\Sigma} \rangle \quad (2.32)$$

$$\langle \hat{F}_{\Sigma^\dagger}(t) \hat{A}(t) \rangle = -igND(\omega - \nu) \langle \hat{D}_{\Sigma\Sigma^\dagger} \rangle \quad (2.33)$$

従って、 $\langle \hat{F}'_B(t) \rangle \neq 0$ であり、

$$\langle \hat{F}'_B(t) \rangle = -2g^2\gamma^{-1}L(\omega - \nu)N \langle \hat{D}_{\Sigma+\Sigma} \rangle \quad (2.34)$$

となる。また、アインシュタイン関係式より次式が導出できる。

$$2N \langle \hat{D}_{\Sigma+\Sigma} \rangle = \hat{\Lambda}_a + (2\gamma - \gamma_a - \gamma_{\text{sp}}) \langle \hat{\sigma}_a \rangle \quad (2.35)$$

$$2N \langle \hat{D}_{\Sigma\Sigma^\dagger} \rangle = (2\gamma - \gamma_b) \langle \hat{\sigma}_b \rangle + \gamma_{\text{sp}} \langle \hat{\sigma}_a \rangle \quad (2.36)$$

更に、断熱近似及び $\gamma \gg \Lambda_a$ を適用すると上式は次式のようなになる。

$$N \langle \hat{D}_{\Sigma+\Sigma} \rangle \simeq \gamma \langle \hat{\sigma}_a \rangle \quad (2.37)$$

$$N \langle \hat{D}_{\Sigma\Sigma^\dagger} \rangle \simeq \gamma \langle \hat{\sigma}_b \rangle \quad (2.38)$$

ここで、平均が 0 となるような新しい雑音演算子を次式のように定義する。

$$\hat{F}_B(t) = \hat{F}'_B(t) + 2g^2\gamma^{-1}L(\omega - \nu)\hat{\sigma}_a \quad (2.39)$$

これを用いて式 (2.29) を書き直すと、

$$\hat{B}(t) = -2g^2\gamma^{-1}L(\omega - \nu)(\hat{\sigma}_a - \hat{\sigma}_b)\hat{A}^\dagger\hat{A} - 2g^2\gamma^{-1}L(\omega - \nu)\hat{\sigma}_a + \hat{F}_B(t) \quad (2.40)$$

となる。

続いて以上の議論から $\hat{\sigma}_a - \hat{\sigma}_b$ を導出する。式 (2.26) において雑音に関する g^2 以上の項を無視すると、 $\hat{\sigma}_a - \hat{\sigma}_b$ の雑音項の影響は共振器内電場に寄与しない。従って演算子の平均 $\langle \hat{\sigma}_a \rangle$ 、 $\langle \hat{\sigma}_b \rangle$ のみを考えれば十分である。原子数 $N_a = N \langle \hat{\sigma}_a \rangle$ 、 $N_b = N \langle \hat{\sigma}_b \rangle$ とすると、式 (2.17)、(2.18)、(2.23)、(2.40) より、次式が得られる。

$$0 = N \langle \hat{\Lambda}_a \rangle - \gamma_a N_a - \gamma_{\text{sp}} N_a - \mathcal{A}(N_a - N_b) \langle \hat{A}^\dagger \hat{A} \rangle - \mathcal{A} N_a \quad (2.41)$$

$$0 = \gamma_b N_b + \gamma_{\text{sp}} N_a + \mathcal{A}(N_a - N_b) \langle \hat{A}^\dagger \hat{A} \rangle + \mathcal{A} N_a \quad (2.42)$$

ここで、

$$\mathcal{A} = \frac{2g^2}{\gamma} L(\omega - \nu) \quad (2.43)$$

である。これらより

$$R_0 = \frac{\langle \hat{\Lambda}_a \rangle}{\gamma_a} \quad (2.44)$$

$$R_{\text{opt}} = \left(\frac{1}{\gamma_a} + \frac{1}{\gamma_b} \right) \langle \hat{\Lambda}_a \rangle - \left(1 + \frac{\gamma_a}{\gamma_b} \right) \langle \hat{\sigma}_a \rangle \quad (2.45)$$

とおくと、次式のように分布数差を求めることができる。

$$N_a - N_b = N(R_0 - R_{\text{opt}}) \quad (2.46)$$

更に $\gamma_c \rightarrow 0$ 、 $\gamma_b \rightarrow \infty$ の極限を仮定すると、

$$N_a - N_b = \frac{N \langle \hat{\Lambda}_a \rangle}{\gamma_{\text{sp}} + \mathcal{A}(\langle \hat{A}^\dagger \hat{A} \rangle + 1)} \quad (2.47)$$

となる。これを用いて式 (2.26) の2項目、すなわち誘導放出による利得は次式のように表すことができる。

$$g(n) = \frac{D(\omega - \nu)\gamma p \beta}{1 + (n + 1)\beta} \quad (2.48)$$

ここで $p = N \langle \hat{\Lambda}_a \rangle$ 、 $\beta = \mathcal{A}/\gamma_{\text{sp}}$ 、 $n = \langle \hat{A}^\dagger \hat{A} \rangle$ とした。 β は飽和パラメータと呼ばれ、外部への自然放出に対する共振器モードへの自然放出の割合を表す。また、 n は共振器内光子数を表す。式 (2.26) に式 (2.48) を代入することにより、最終的に共振器内電場のダイナミクスは次式のようになる。

$$\frac{d}{dt} \hat{A}(t) = -\frac{1}{2} \left\{ \gamma_c + 2i(\Omega - \nu) - \frac{D(\omega - \nu)\gamma p \beta}{1 + (n + 1)\beta} \right\} \hat{A}(t) + \hat{G}(t) \quad (2.49)$$

2.3 複素ランジュバン方程式

前節では、レーザーの共振器内電場の演算子のダイナミクスを表す量子ランジュバン方程式を導出した。本説では共振器内電場はコヒーレント状態にあると仮定して平均値と量子ゆらぎによる半古典的な取り扱いに移行する。すなわち式 (2.49) において、

$$\hat{A}(t) \rightarrow A(t) = \sqrt{n} e^{i\theta} \quad (2.50)$$

として、演算子を複素数に変換する。更にここでは理想的な状況を考え $\omega = \nu = \Omega$ とすると、式 (2.49) は次式のように書き換えられる。

$$dA = -\frac{1}{2} \left\{ \gamma_c - \frac{p\beta}{1 + (n + 1)\beta} \right\} A dt + \sqrt{D} dW \quad (2.51)$$

ただし $n = |A|^2$ である。また、 dW は複素ウィーナー仮定であり、 D は次式で表される雑音の分散である。

$$D = \frac{\langle \hat{G}^\dagger \hat{G} \rangle + \langle \hat{G} \hat{G}^\dagger \rangle}{2} \quad (2.52)$$

次にレーザーの発振閾値を定常状態で光子数 n が1となるポンプレートで定義すると、

$$p_{\text{th}} = (1 + \beta^{-1})\gamma_c \quad (2.53)$$

となる。特に半導体レーザーのように $1 \gg \beta$ の場合には、

$$p_{\text{th}} = \beta^{-1}\gamma_c \quad (2.54)$$

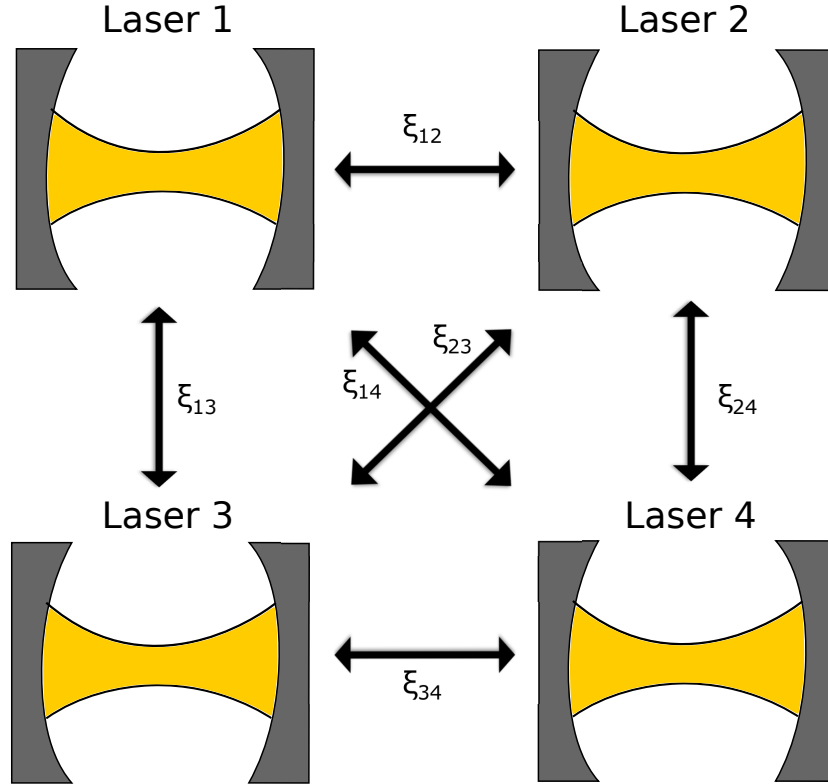


図 2.2. $N=4$ の場合のネットワークの概念図。

となる。更に閾値より上で動作すると仮定し $n \gg 1$ とすると式 (2.51) は最終的に次式のように書き表せる。

$$dA = -\frac{1}{2}\gamma_c \left(1 - \frac{p/p_{\text{th}}}{1 + \beta|A|^2}\right) A dt + \sqrt{D}dW \quad (2.55)$$

また、位相と光子数の式に分解すると

$$d\theta = \sin \theta \sqrt{D}dW \quad (2.56)$$

$$dn = -\gamma_c \left(1 - \frac{p/p_{\text{th}}}{1 + \beta n}\right) n + 2r \cos \theta \sqrt{D}dW \quad (2.57)$$

となる。すなわち単一レーザーの位相は雑音によって拡散し続け、完全にランダムとなる。

2.4 相互注入によるネットワークのダイナミクス

これまでの議論は単一のレーザーに対するものであったが、本説では図 2.2 のように、多数のレーザーを相互注入することによってネットワークを形成することを考える。図 2.2 では 4 つのレーザーを結合しているが、 N 個のレーザーを結合するとき、 j 番目のレーザーの複素振幅を A_j とするとそのダイナミクスは式 (2.55) に相互注入項を加えることによって次式のように表すことができる。

$$dA_j = -\frac{1}{2}\gamma_c \left(1 - \frac{p/p_{\text{th}}}{1 + \beta|A_j|^2}\right) A_j dt + \sum_{k=1, k \neq j}^N \xi_{jk} A_k + \sqrt{D} dW \quad (2.58)$$

ここで ξ_{jk} は j 番目のレーザーと k 番目のレーザーの結合の強さを表す結合係数である。簡単のため雑音の効果を無視すると、ネットワーク上の各レーザーの位相及び光子数の方程式は次式のようになる。

$$\frac{d\theta_j}{dt} = - \sum_{k=1, k \neq j}^N \xi_{jk} \sqrt{\frac{n_k}{n_j}} \sin(\theta_j - \theta_k) \quad (2.59)$$

$$\frac{dn_j}{dt} = -\gamma_c \left(1 - \frac{p/p_{\text{th}}}{1 + \beta n_j}\right) n_j + \sum_{k=1, k \neq j}^N \sqrt{n_j n_k} \cos(\theta_j - \theta_k) \quad (2.60)$$

全てのレーザーが同じ振幅値を持つとき、式 (2.59) は一般化された結合振動子系の位相のダイナミクスを表す蔵本モデルと等価となる。

また、レーザー間の結合係数については $\xi_{ij} > 0$ のとき同位相結合、 $\xi_{ij} < 0$ のときに逆位相結合となり、図 2.3 のように、対応するレーザーはそれぞれ同位相及び逆位相となるような効果を受ける。例えば二つのレーザーが ξ で結合している場合、雑音の影響を無視すると式 (2.59) から位相のダイナミクスは次式のように表すことが出来る。

$$\frac{d\theta_1}{dt} = -\xi \sqrt{\frac{n_2}{n_1}} \sin(\theta_1 - \theta_2) \quad (2.61)$$

$$\frac{d\theta_2}{dt} = -\xi \sqrt{\frac{n_1}{n_2}} \sin(\theta_2 - \theta_1) \quad (2.62)$$

$n_1 \simeq n_2$ のとき位相差 $\Delta\theta = \theta_1 - \theta_2$ のダイナミクス次式のようになる。

$$\Delta\theta = -\xi \sin(\Delta\theta) \quad (2.63)$$

従って $\xi > 0$ のとき安定点は $\Delta\theta = 2\pi l$ となり、 $\xi < 0$ のときに安定点は $\Delta\theta = \pi(2l + 1)$ となる。ただし l は任意の整数である。実際には雑音があるため、結合 ξ が雑音に対して非常に小さいときは位相は揃わない。

またネットワークの場合には、全てのレーザーは定常状態で同じ光子数 $n_i = n$ であると仮定すると熱力学極限 $N \rightarrow \infty$ で各レーザーの位相分布 $P(\theta)$ は次式のように、ボルツマン分布となる事が導かれる [17]。

$$P(\theta) \simeq N \exp\left(-\frac{2n_s}{D} H(\theta)\right) \quad (2.64)$$

ここで、 $n_s = (p - \beta^{-1}\gamma_c)/\gamma_c$ であり、エネルギー $H(\theta)$ は次式で表される。

$$H(\theta) = - \sum_{i < j} \xi_{ij} \cos(\theta_i - \theta_j) \quad (2.65)$$

また、この $H(\theta)$ はスピン系のモデルである xy モデルのハミルトニアンと等価である。

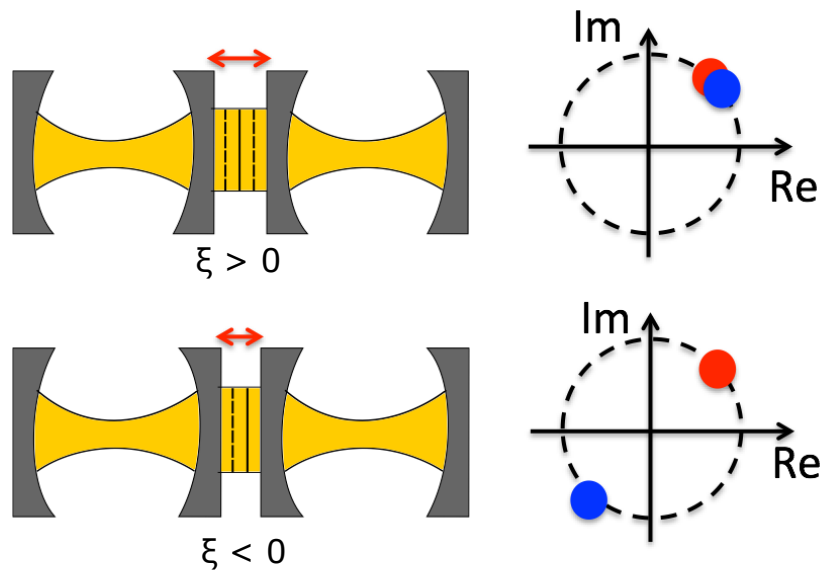


図 2.3. 結合の効果を表す概念図。

第3章

コミュニティ検出

導入でも述べたように、現実世界に存在する様々なネットワークはコミュニティ構造を形成している。ネットワークを解析する上でコミュニティ検出は、ノードの分類、ネットワークの粗視化、ネットワークの機能解明といった点で重要である [9]。ここでのネットワークの機能とはネットワークの構造に起因するあらゆる現象のことを指す。特に脳科学の分野では、様々なタスクにおいて fMRI(functional Magnetic Resonance Imaging) から得られたネットワークからコミュニティ検出を行い、脳機能と関連づける研究が近年多数行なわれている [6]。

しかしながら、コミュニティの定義もいくつか存在し、問題自体が組み合わせ最適化として一般化されていないため、アルゴリズムの評価が非常に難しい。それゆえ、コミュニティ検出のアルゴリズムは多数存在するが、その評価は目的によって様々であり、現在もその性能は活発に議論されている。本章ではまずコミュニティ検出について概観し、次に現在最も使用されている評価関数であるモジュラリティを紹介する。更に、本研究で主に用いたアルゴリズムの評価方法について説明する。

3.1 コミュニティ

複雑ネットワークにおいて、コミュニティとは密に結合しているノードの集合であり、それに加えて複数のコミュニティ間は疎に結合しているようなものを指す。このような曖昧な説明の他に、一般的に受け入れられているコミュニティの定量的な定義は存在せず、あるグラフ G の部分グラフ C がコミュニティであるための条件は目的やネットワークの種類に依存する。ただしコミュニティとは基本的には密に結合している部分グラフであるから、局所的なエッジ密度と関係づけて定義されることが多い。

コミュニティの定義について、一つ例を挙げる。 G のノード数を n 、 C のノード数を n_c とする。 C に属するあるノード $v \in C$ の次数のうち、 C 内の他のノードと結合しているエッジの本数を k_v^{int} 、 C に属さないノードと結合しているエッジの本数を k_v^{ext} と定義する。これらを用いて C の内部エッジ密度 (intra-cluster density) 及び外部エッジ密度 (inter-cluster

density) を次式で定義する [10]。

$$\delta_{int}(C) = \frac{\sum_{v \in C} k_v^{int} / 2}{n_c(n_c - 1) / 2} \quad (3.1)$$

$$\delta_{ext}(C) = \frac{\sum_{v \in C} k_v^{ext} / 2}{n_c(n - n_c)} \quad (3.2)$$

C がコミュニティであるためには $\delta_{int}(C) > \delta_{int}(G)$ でなければいけない。ここで $\delta_{int}(G)$ は明らかにグラフ全体のエッジ密度である。また、コミュニティ間の繋がりが疎であるということも考慮にいれると $\delta_{ext}(C) < \delta_{int}(G)$ という条件も必要であり、そのコミュニティの強さは $\delta_{int}(C) - \delta_{ext}(C)$ というように表すことができる。更に、 $\delta_{int}(C) - \delta_{ext}(C)$ がある閾値より大きいかがかでコミュニティを定義することも可能である。

このようなコミュニティの定義の他にも様々な定量的な定義が存在する。しかし、現在使用されているコミュニティ検出アルゴリズムは、陽に定義されたコミュニティを検出するというものではなく、「コミュニティ内の結合が密、コミュニティ間の結合は疎」というより一般的なコンセプトのもと、アルゴリズム独自の指標に基づいてグラフのクラスタリングを行なうというものが多い。

3.2 コミュニティ検出問題

コミュニティ検出問題とは具体的にはグラフ $G(V, E)$ を、その部分グラフであるコミュニティの集合 $\{C_i\}$ に直和分割する問題である*¹。一般のグラフ分割やデータクラスタリングと比較すると、ノードをグループに分類するといった点では同じであるが、その他の点においては幾分異なる。データのクラスタリングの場合は、基本的にデータ間の距離や類似度といった指標の「近さ」という概念に基づいてクラスタリングが行なわれる。それに対してコミュニティ検出の場合は前述の通りグラフの構造自体に起因するエッジの局所的な密度に基づいてノードを分類する。更に、グラフ分割やデータクラスタリングは、分割するクラスターの数は予め与えられていることが多い。しかしながらコミュニティ検出は、コミュニティの数はアルゴリズムが自ら見つけなくてはならない。

また、分割されたコミュニティの質は当然議論されるべき問題である。前述の通りこれは目的やネットワークの種類に依存するが、ネットワークの構造のみからコミュニティ検出を行なうというより一般的な立場では、ネットワークの背後に存在する情報に依存せずに評価されなければならない*²。分割されたコミュニティの良さを大域的に評価する指標は様々なものがあるが、その中でもモジュラリティが有名であり、これについては次節で説明する。また、アルゴリズムの性能を評価する場合には、あるモデルの元で人工的にコミュニティを埋め込んだネットワークに対してコミュニティ検出を行い、埋め込まれたコミュニティをどの程度再現で

*¹ コミュニティのオーバーラップを許す場合や、コミュニティが階層構造を成している場合もあるが、本研究では言及しない。

*² もちろん、特定分野のネットワークのコミュニティを限定された目的で検出するという立場からは、背後に存在する情報は考慮しなればいけない。

きたかの評価を行なうことが多い。これについては 3.5 節で説明する。

3.3 評価関数:モジュラリティ

モジュラリティは、あるグラフの分割がどの程度コミュニティ構造を上手く抽出しているかを与える評価関数であり、2004年に Newman らによって提案された [18]。この指標はランダムなグラフはコミュニティ構造をもたないという前提のもとで、元グラフと同じ分割をランダムなグラフで行った場合に、どの程度コミュニティ内のエッジに差があるかの期待値を比較したものである。このとき元のグラフと比較するランダムなグラフは、元グラフのなんらかの性質を保存したままコミュニティ構造を無くしたものであり、ヌルモデル (null model) と呼ばれる。モジュラリティは次式によって表される。

$$Q = \frac{1}{2m} \sum_{ij} (A_{ij} - P_{ij}) \delta(C_i, C_j) \quad (3.3)$$

ここで m はグラフの総エッジ数、 $\{A_{ij}\}$ はグラフの隣接行列、 $\{P_{ij}\}$ はヌルモデルにおいてノード間に存在するエッジの本数の期待値を表す行列である。また、 C_i はノード i が属するコミュニティ、 $\delta(x, y)$ はクロネッカーのデルタ関数である。

$$\delta(C_i, C_j) = \begin{cases} 1 & (C_i = C_j) \\ 0 & (C_i \neq C_j) \end{cases} \quad (3.4)$$

一般的に広く用いられているモジュラリティは、configuration model とよばれるモデル [19] をヌルモデルとしたものであり、次式で表される。

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(C_i, C_j) \quad (3.5)$$

configuration model はノードの次数のみを保存したまま、エッジのランダムな張り替えを行なうモデルであり括弧内第二項目はモデルにおいての (i, j) 間のエッジの本数の期待値を表す。組み合わせ最適化問題としてのモジュラリティ最大化は NP 困難であり、ノード数が大きくなると近似解を求めることしかできない。他にも解像度の限界やその統計的な正当性など様々な問題が議論されているが、その単純さから多数のアルゴリズムがモジュラリティ最大化を基礎としている。

3.4 コミュニティ検出アルゴリズム

コミュニティ検出のアルゴリズムは多数提案されている。最も良く用いられているのはモジュラリティを最大化するタイプのアルゴリズムで、その中でも貪欲法を基本とした Louvain 法は現在最も高速なコミュニティ検出アルゴリズムの一つである [20, 10]。Modularity 最大化以外にも様々なアルゴリズムが存在するが、Lancichinetti らによる人工グラフにおける評

価^{*3}において高精度な結果を得た Infomap[21] が有名である。本説では、これら二つのアルゴリズムについて紹介する。

Louvain 法

Louvain 法は貪欲法を基礎としたアルゴリズムであり、ノードのコミュニティへの所属を決定するステップ 1 と、コミュニティを新たなノードとみなすステップ 2 からなる。ステップ 1 では、まず最初に全てのノードを異なるコミュニティに所属させる。従って初期状態はノードの数だけコミュニティが存在する。次にあるノード i について、それと結合しているノード j と同じコミュニティに移動させた場合のモジュラリティの増分 ΔQ を計算する。 ΔQ の計算を結合しているノード全てについて行ない、ノード i を、 ΔQ が最大となるノードと同じコミュニティに移動させる。ただし、 ΔQ が正となるような移動がない場合はどのような操作も行なわない。また、ノード i がコミュニティ C に移動する場合のモジュラリティの増分 ΔQ は次式によって計算できる。

$$\Delta Q = \left[\frac{\sum_{in} + k_{i,in}}{2m} - \left(\frac{\sum_{in} + k_i}{2m} \right)^2 \right] - \left[\frac{\sum_{in}}{2m} - \left(\frac{\sum_{tot}}{2m} \right)^2 - \left(\frac{k_i}{2m} \right)^2 \right] \quad (3.6)$$

ここで、 \sum_{in} はコミュニティ C 内のエッジの重みの総和、 \sum_{tot} は、 C 内のノードと結合しているエッジの重み総和を表す。また、 k_i はノード i と結合しているエッジの重みの和、 $k_{i,in}$ はノード i と C 内のノードを結合するエッジの重みの和、 m はグラフ全体のエッジの重みの総和である。このような操作を全ノードに対して、 $\Delta Q > 0$ となるような移動がなくなるまで繰り返し行ない、ステップ 1 を終了する。このとき、同じノードを複数回操作する可能性もあることに注意する。

ステップ 1 でネットワークは複数のコミュニティに分割されたが、ステップ 2 ではコミュニティを新たなノードとして、新しいグラフを作成する。すなわち新しいグラフでは個々のノードがステップ 1 で分割されたコミュニティに対応する。このとき、各ノード間のエッジの重みは、ステップ 1 で分割されたコミュニティ間のエッジの重みの総和とする。またコミュニティ内のエッジの重みの総和をノードの自己結合のエッジの重みとする。このようなグラフに対して、再びステップ 1 に戻り、同じ操作を繰り返す。

著者らはステップ 1 とステップ 2 の一連の操作を「pass」と名付け、pass はモジュラリティの増分がなくなるまで繰り返される。図 3.1 にアルゴリズムの流れを示す。一般的にグラフのノード数は pass を経るごとに著しく減少するため、アルゴリズム全体の計算時間は主に最初の pass に費やされる。モジュラリティの増分 ΔQ はノードの次数オーダーで計算できるため、計算量はおおよそ $O(m)$ である^{*4}。

^{*3} この論文で用いられたグラフは LFR ベンチマークグラフとよばれ、アルゴリズムのベンチマークにしばしば用いられているがその詳細は 5 章に示す。また評価方法について次節で説明する。

^{*4} この計算量の評価は厳密ではなく経験的なものである。

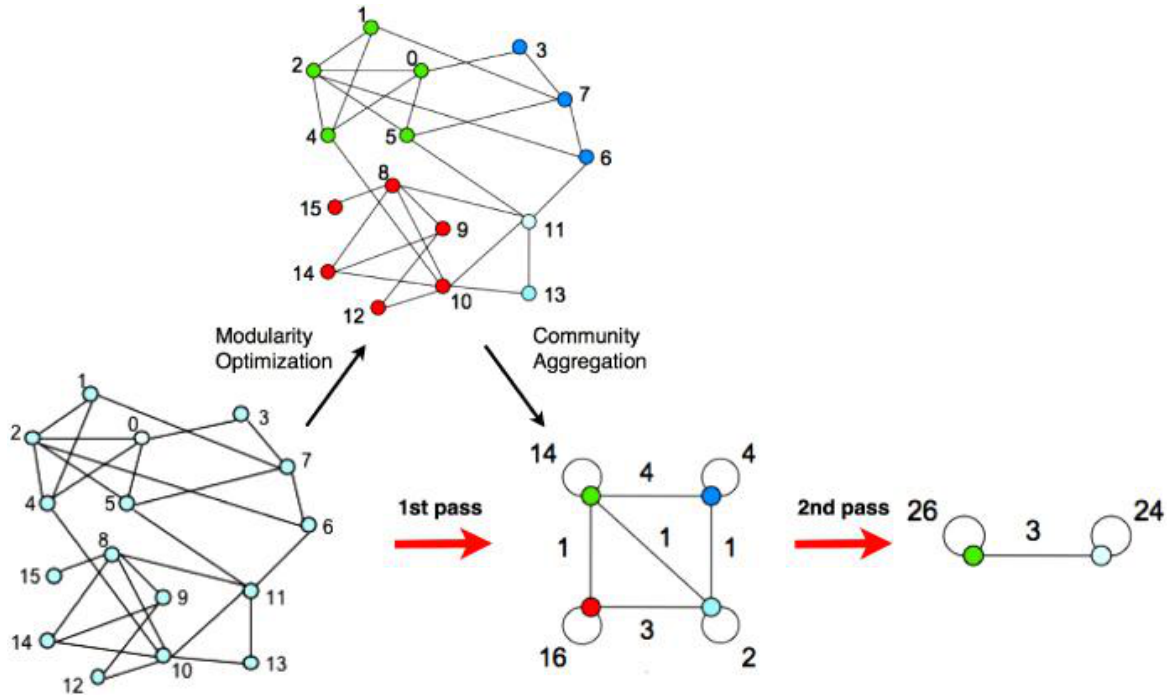


図 3.1. Louvain 法の流れ (Blondel et al. [20] より引用)

Infomap

Infomap は、グラフ構造における情報圧縮を最適化することによって、コミュニティ構造を特定するアルゴリズムである [21, 22]。グラフ上のランダムウォークの道 (Path)^{*5}を最も効率よく記述する符号化を特定することによってコミュニティを検出する。具体的には、次式で表されるランダムウォークにおける符号化の平均符号長 $L(M)$ をコスト関数として、それを最小化するグラフ分割を見つける。

$$L(M) = q_{\curvearrowright} H(Q) + \sum_{i=1}^m p_{\circlearrowleft}^i H(P^i) \quad (3.7)$$

ここで、 $M: \alpha \rightarrow i$ はノード $\alpha \in \{1, 2, \dots, n\}$ からコミュニティ $i \in \{1, 2, \dots, m\}$ へのグラフ分割を表す。(3.7) 式の第一項目はコミュニティ間の移動を表す符号帳の頻度で重み付けされた平均符号長を表す項で、 $q_{\curvearrowright} = \sum_{i=1}^m q_{i\curvearrowright}$ はランダムウォークが任意のステップで他のコミュニティに移動する確率を表す。ただし $q_{i\curvearrowright}$ はコミュニティ i を脱出する確率である。また $H(Q)$ はコミュニティ間の移動を表す平均符号長であり、次式によって表される。

$$H(Q) = - \sum_i \frac{q_{i\curvearrowright}}{q_{\curvearrowright}} \log \left(\frac{q_{i\curvearrowright}}{q_{\curvearrowright}} \right) \quad (3.8)$$

*5 ランダムウォークの道 (Path) とは、ネットワーク上のランダムウォーカーが通るノードの列のこと。

これはランダムウォークがコミュニティ間を移動する確率分布のエントロピーに他ならない。(3.7)式第二項目はコミュニティ内での移動を表す符号帳の頻度で重み付けされた平均符号長を表す項で、 $p_{\circ}^i = \sum_{\alpha \in i} p_{\alpha} + q_{i \circ}$ はランダムウォークがコミュニティ i 内に存在する割合と i から脱出する確率^{*6}の和である。ただし p_{α} はノード α を訪れる確率である。また $H(P^i)$ はコミュニティ内の移動及びコミュニティからの脱出を表す符号帳の平均符号長で、次式によって表される。

$$H(P^i) = -\frac{q_{i \circ}}{p_{\circ}^i} \log \left(\frac{q_{i \circ}}{p_{\circ}^i} \right) - \sum_{\alpha \in i} \frac{p_{\alpha}}{p_{\circ}^i} \log \left(\frac{p_{\alpha}}{p_{\circ}^i} \right) \quad (3.9)$$

具体的に (3.7) 式を最小化するための最適化には貪欲法と焼き鈍し法 (Simulated Annealing) が用いられる。貪欲法ではまず各ノードの滞在確率を計算し、初期状態では全てのノードが異なるコミュニティに割り当てられる。そして Louvain 法と同様に、最小の $L(M)$ が実現するようにまわりのコミュニティと統合していく。このような操作を $L(M)$ がどのコミュニティと統合しても小さくならない状態になるまで繰り返す。貪欲法が終了したあとに更に最適化を行なう場合は、得られた分割から焼き鈍し法を異なる温度で適用し、その中から最小の $L(M)$ を実現するものを選ぶ。

3.5 アルゴリズムの性能評価:Normalized Mutual Information

この節ではアルゴリズムの性能を評価する方法について説明する。3.2 節でも言及した通り、人工的にコミュニティを埋め込んだベンチマークグラフに対してアルゴリズムを適用し、図 3.2 のように特定されたコミュニティ構造と埋め込まれたコミュニティ構造がどの程度類似しているかに基づいて評価する。類似度を評価する指標は、Rand index、VI (Variation of Information)、NMI (Normalized Mutual Information) [23] など様々なものがあるが、本研究ではグラフ上のコミュニティ分割に特化した NMI を用いる。

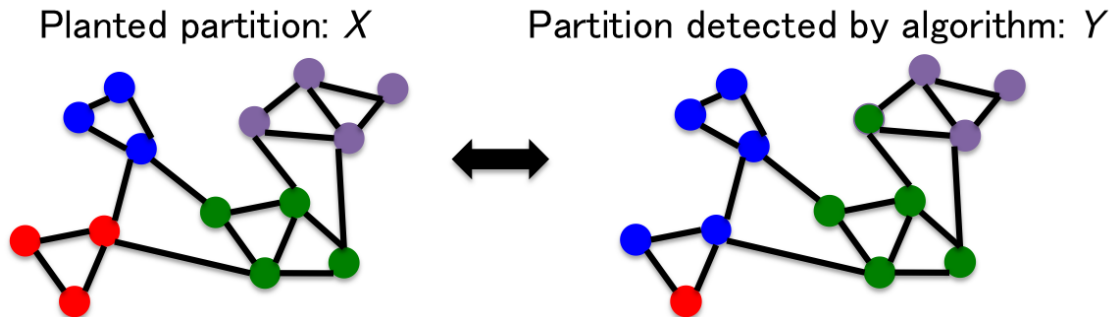


図 3.2. コミュニティ分割の比較例。

^{*6} ランダムウォークがコミュニティから脱出する際には、そのことを表す符号を用いるので $q_{i \circ}$ が必要。

NMI は同じグラフにおける二つの異なるコミュニティ分割の類似度を、相互情報量という形で定式化した指標である。情報理論において相互情報量とは、二つの確率変数 X, Y において、それらがどの程度相互に依存し合っているかを表す。まず $\{x_i\}, \{y_i\}$ を与えられたグラフに対するノードのコミュニティへの割り当てとする。ここで x_i, y_i はそれぞれ、分割 \mathcal{X}, \mathcal{Y} においてノード i が属するコミュニティのラベルを表す。また、ラベル x, y を二つの確率変数 X, Y の値とする。そしてランダムに選んだノードがコミュニティ x に属する確率を、 $P(X = x) = N_x/N$ と定義する。ただし N_x はコミュニティ x に属するノード数であり、 N は全ノード数である。同様に $P(Y = y) = N_y/N$ 、同時確率を $P(X = x, Y = y) = N_{xy}/N$ と定義する。また N_{xy} は、分割 \mathcal{X}, \mathcal{Y} における重複部分のノード数である。これらを用いて相互情報量は次式のように表すことができる。

$$I(\mathcal{X}, \mathcal{Y}) = \sum_{x,y} P(x,y) \log \frac{P(x,y)}{P(x)P(y)} \quad (3.10)$$

この相互情報量は \mathcal{Y} が \mathcal{X} を更に分割してできる全ての分割に対して同じ値をとる。このような状況を回避するために式 (3.10) をそれぞれの分割のエントロピー $H(X) = \sum_x P(x) \log P(x)$ 、 $H(Y) = \sum_y P(y) \log P(y)$ で規格化したものが NMI である [10, 23]。

$$\text{NMI}(\mathcal{X}, \mathcal{Y}) = \frac{2I(\mathcal{X}, \mathcal{Y})}{H(X) + H(Y)} \quad (3.11)$$

NMI は、 $[0,1]$ の値を取り、二つの分割 \mathcal{X} 及び \mathcal{Y} が完全に等しいときに “1” となる。

しかし、NMI はネットワークの有限サイズ効果によって、多い分割にバイアスされる問題点が指摘されている [24]。特に、一方の分割において全てのノードが別々のコミュニティに属している意味のない状況でも NMI は 0 でない比較的大きい値を取り得る。この問題の解決策として最近 NMI の統計的重要性に注目した rNMI(relative NMI) が提案された [24]。あるグラフに対して分割 \mathcal{X} と \mathcal{Y} を比較するとき、rNMI は次式のように表される。

$$\text{rNMI}(\mathcal{X}, \mathcal{Y}) = \text{NMI}(\mathcal{X}, \mathcal{Y}) - \langle \text{NMI}(\mathcal{X}, \mathcal{Z}) \rangle \quad (3.12)$$

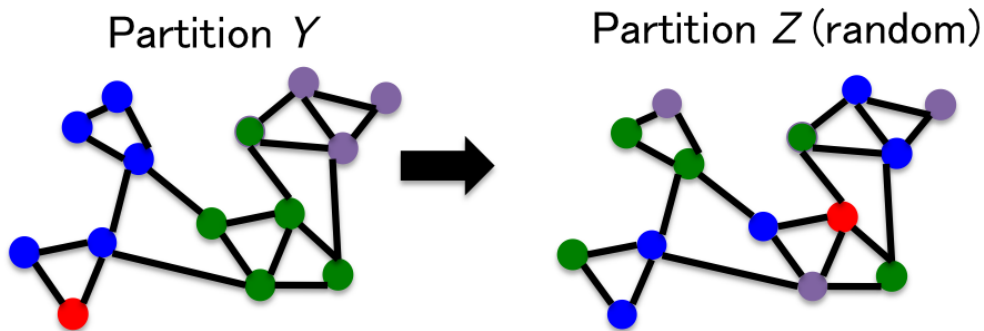


図 3.3. rNMI における \mathcal{Y} と \mathcal{Z} のコミュニティ分割の例。 \mathcal{Z} は \mathcal{Y} と完全に同じコミュニティサイズ分布を保ったランダム分割である。

ここで $\langle \cdot \rangle$ は統計平均を表し、 \mathcal{Z} は図 3.3 のように、 A と同じコミュニティサイズの分布を持ち、それ以外は完全にランダムな分割である。そして、 \mathcal{X} と \mathcal{Y} が全く同じ分割のときに 1 となるように規格化したものは rrNMI (ratio of relative NMI) と呼ばれている [25]。

$$\text{rrNMI}(\mathcal{X}, \mathcal{Y}) = \frac{\text{rNMI}(\mathcal{X}, \mathcal{Y})}{\text{rNMI}(\mathcal{X}, \mathcal{X})}. \quad (3.13)$$

rrNMI を計算するためには式 (3.12) の第 2 項目を計算するために、ランダムな分割 \mathcal{Z} に対して NMI の統計平均をとらなければならない。現状、解析的にこの統計的期待値を求めることはできないため、複数回のランダムな試行に対する平均で代用している。また、 rrNMI 、 NMI 共に二つのコミュニティ分割が完全に一致している場合は 1 となる。

第 4 章

レーザーネットワークを用いたコミュニティ検出

本章では、2章で述べたレーザーのダイナミクスを用いてコミュニティ検出を行なう方法を説明する。逆位相注入と同位相注入を利用して、ネットワーク上の繋がりが密なノード集団を近い値の位相で発振させることによりコミュニティを検出する。まず最初に、ネットワーク上の結合振動子の同期のダイナミクスとコミュニティ構造の関係について調べた先行研究を紹介する。そしてレーザーネットワークにおけるコミュニティ検出の方法を説明する。また、本研究では実際のレーザーを用いた評価は行なっていないが、そのような系を実験的に構築する方法についても述べる。最後に数値シミュレーションによって、提案した手法が簡単なネットワークにおいてコミュニティを検出できることを確認する。

4.1 ネットワーク上の結合振動子の同期にコミュニティ構造が与える影響

複雑ネットワーク上の非線形振動子の同期現象の振る舞いは盛んに研究されている。振動子のモデルとしては、一般の非線形振動子を位相のみの自由度に縮約した蔵本モデルが有名である [26]。 N 個の振動子が結合している蔵本モデルのダイナミクスは次式のように表される。

$$\frac{d\theta_i}{dt} = \omega_i - \sum_j^N \xi_{ij} \sin(\theta_i - \theta_j) \quad (4.1)$$

ここで θ_i は i 番目の振動子の位相であり、 ξ_{ij} は i 番目と j 番目の振動子の結合の強さを表す。また、 ω_i は各振動子に固有の周波数である。蔵本は、全ての振動子が他の全ての振動子と同じ強さで結合している平均場的な結合 $\xi_{ij} = \xi, \forall i, j$ において、熱力学極限 $N \rightarrow \infty$ で全ての振動子が同期する ξ の条件に相転移が存在することを示した。更に、複雑ネットワーク上の振動子に関しても、ネットワークの構造と同期の条件に関する研究は近年多数なされている。

このような背景の中、Arenas et al. [13] は蔵本振動子が最終的に同期状態に至るまでのダイナミクスに着目して、コミュニティ構造がある場合には同一コミュニティ内でのみ同期した

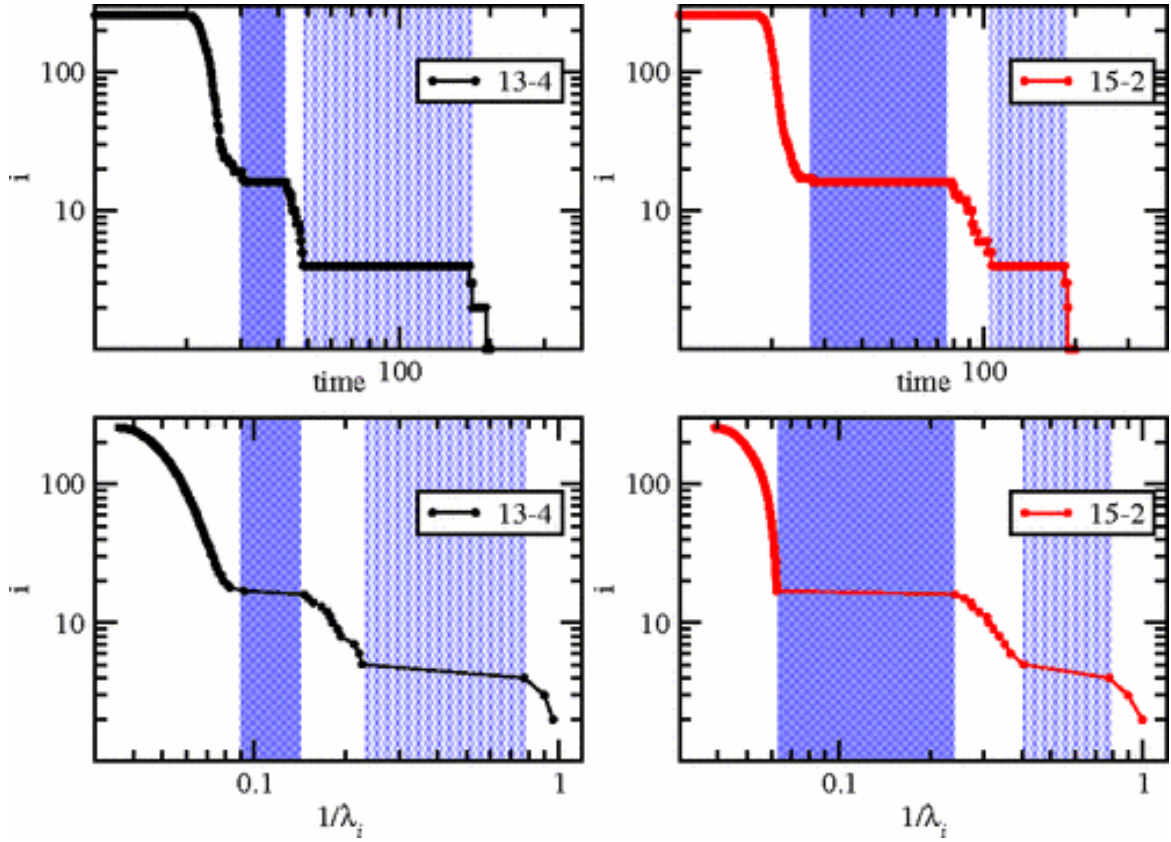


図 4.1. 蔵本モデルにおけるダイナミクス。(Arenas et al.[13] より引用。)

安定状態が出現することを示した。Arenas らは、初期値をランダムな位相状態として、次式で定義される複数回の位相相関のアンサンブル平均 ρ_{ij} のダイナミクスを解析した。

$$\rho_{ij}(t) = \langle \cos(\theta_i(t) - \theta_j(t)) \rangle \quad (4.2)$$

図 4.1 上段は $\rho_{ij}(t)$ を閾値 T でバイナリ化した行列 $\{D_{ij}(t)\}$

$$D_{ij}(t) = \begin{cases} 1 & \text{if } \rho_{ij}(t) > T \\ 0 & \text{otherwise} \end{cases} \quad (4.3)$$

の連結成分の個数の時間発展である。初期状態はランダムなので連結成分はノードの数だけ存在し、最終的には全振動子が同期するため連結成分は1つとなる、注目すべきはその中間に、同期した振動子集団が複数存在する状態がある時間スケールの間安定して出現することである。これは全振動子が同期する過程で、同一コミュニティ内の振動子のみがクラスターとなって同期した状態が存在することを示唆している。また、図 4.1 下段はネットワークのラプラシアン行列の固有値のスペクトルであり、固有値が連続している領域と固有値にギャップが存在する領域があり、このギャップの大きさと上段の安定状態が対応していることを示している。これはネットワーク構造に起因するグローバルな複数のモードがそれぞれ固有の同期状態への

緩和レートを持っており、その緩和レートのギャップの大きさによって安定状態の時間スケールが決定づけられていることを示唆している。

この研究はコミュニティ構造がネットワーク上で生じるダイナミクスに影響を与える好例であり、ネットワーク構造とクラスター状態の特徴的な時間スケールの対応を示したことは物理的にも興味深い。また、もっとも安定な時間スケールにおけるコミュニティ分割が、高いモジュラリティを持つことも示されている。

4.2 レーザーネットワークの構成

前節では、コミュニティ構造のあるネットワーク上において、蔵本振動子がコミュニティ内の大きな位相相関をもった特徴的な時間スケールが現れることを紹介した。蔵本モデルは非線形振動子の一般系であるから、様々な振動子にも同様の現象が生じると考えられる。実際2章でも示した通り、非線形振動子であるレーザーの位相ダイナミクスは、蔵本モデルと非常に似ている。この研究を受けて、本説では実際に物理系を構築することを視野にいれ、素過程の速いレーザーを用いたコミュニティ検出方法を提案する。ただし前節で紹介した内容は全のノードが同期するまでの中間状態においてコミュニティ内の位相相関が現れたことに対して、提案する手法は概ね位相のダイナミクスが収束した状態を用いる。従って位相相関を求めるためのアンサンブル平均の計算を時間にまたがって行なう必要がなく、最後の一度だけで良いことになる。

まず、与えられたグラフ $G(V, E)$ についてコミュニティを検出するために、レーザーネットワークを構成する。構成は非常にシンプルで、与えられたグラフ構造から、実際にエッジがあるノード間には同位相の相互注入を、エッジのないノード間には逆位相注入を行ないレーザーを発振させるというものである。即ち、与えられたグラフに対して図 4.2 のように、エッジがあるノード間には正の結合をいれ、エッジがないノード間には負の結合をいれたようなレーザーネットワークを構成する。これは、グラフ $G(V, E)$ が与えられたとき、次式

$$dA_j = -\frac{1}{2}\gamma_c \left(1 - \frac{p/p_{\text{th}}}{1 + \beta|A_j|^2}\right) A_j dt + \sum_{k=1, k \neq j}^N \xi_{jk} A_k + \sqrt{D} dW \quad (4.4)$$

において以下のような結合を入れる事に相当する。

$$\xi_{ij} = \begin{cases} \xi & \text{if } (i, j) \in E \\ -\gamma\xi & \text{otherwise} \end{cases} \quad (4.5)$$

ここで、 $\xi > 0$ であり $\gamma > 0$ は同位相结合に対する逆位相结合の強度比を表す。このような結合を入れる事により、図 4.3 のように、逆位相注入より同位相注入が支配的になるエッジが密なレーザー間でのみ同位相で発振する事が期待される。

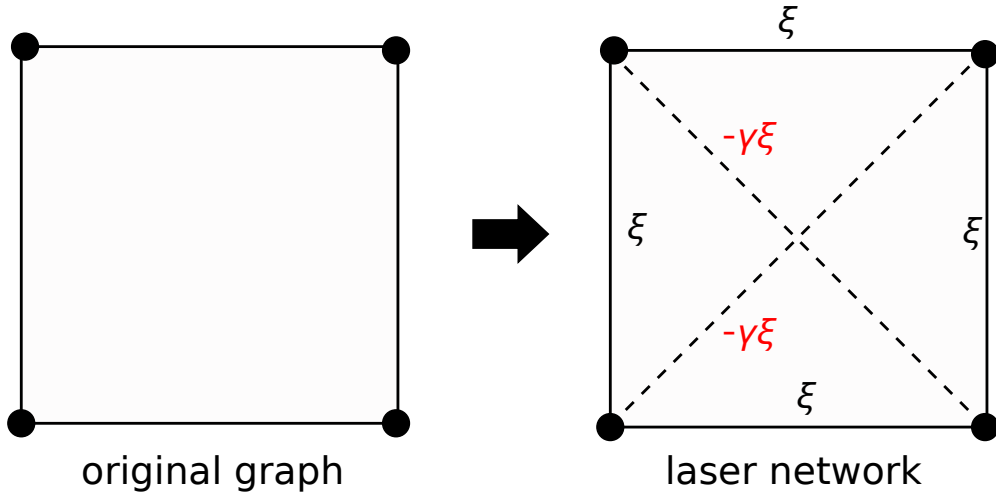


図 4.2. レーザーネットワーク構成の概念図。右図において実線が同位相結合、点線が逆位相結合を示す。また γ は同位相結合に対する逆位相結合の強度比である。

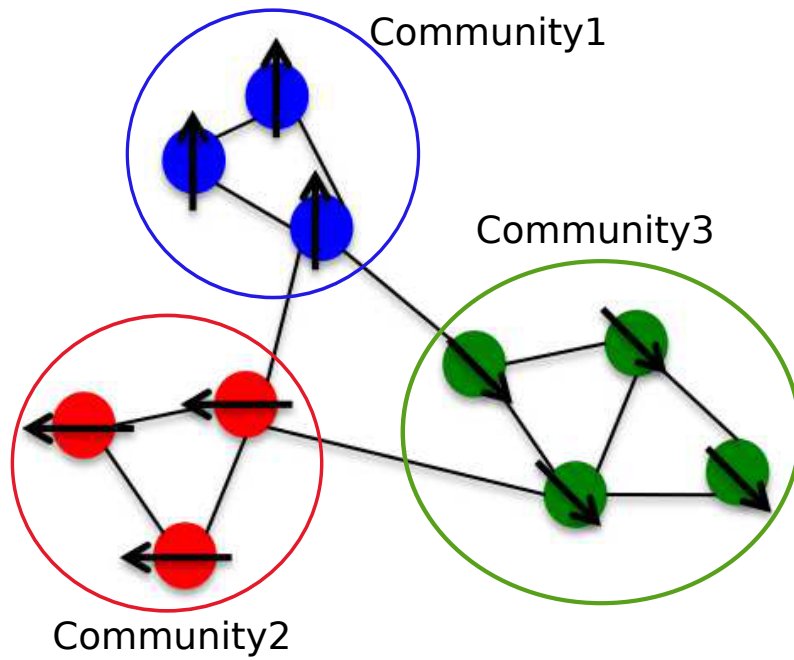


図 4.3. コミュニティ検出のイメージ。各ノードはレーザーを表し、矢印は位相を表す。

4.3 コミュニティの特定

前節で述べたレーザーネットワークの構成のもとで、式 (4.4) で表されるレーザーのダイナミクスが準平衡状態に達したあとに、各レーザーの位相状態からコミュニティを特定しなければならない。コミュニティの検出には振幅の自由度は使用せず位相情報のみを用いることにす

ると、360度の位相状態を使用できることになる。ただし負の結合が入っているため、同一コミュニティであっても位相空間上で幅を持って発振することになる。従って、一度の試行において位相状態のコミュニティへの対応を一気に決めるには工夫が必要である。一方、式(4.4)からも明らかな通り、時間発展にはランダムなノイズが含まれており、系は振幅0の状態からランダムなノイズによってドライブされることになる。すなわち、各試行において位相空間上の絶対位相は完全にランダムで意味を持たず、各ノード間の相対的な位相関係のみが意味を持つことになる。そこで、Arenasらと同様に複数回独立に試行した結果の各ノード間の位相差の余弦のアンサンブル平均(相関)を考える。

$$\rho_{ij} = \langle \cos(\theta_i - \theta_j) \rangle \quad (4.6)$$

ランダム性があるため、このようにして得られた ρ_{ij} は、エッジが密な構造(即ち同一コミュニティ内)のノード間のみで大きな値をとると期待される。更にある閾値 T で ρ_{ij} をバイナリ化した値を要素にもつ行列 $\{D_{ij}\}$ を考え、これを隣接行列としたグラフの連結成分を特定することによってコミュニティ検出を行なう。図4.4にコミュニティ検出の全体の流れを示す。

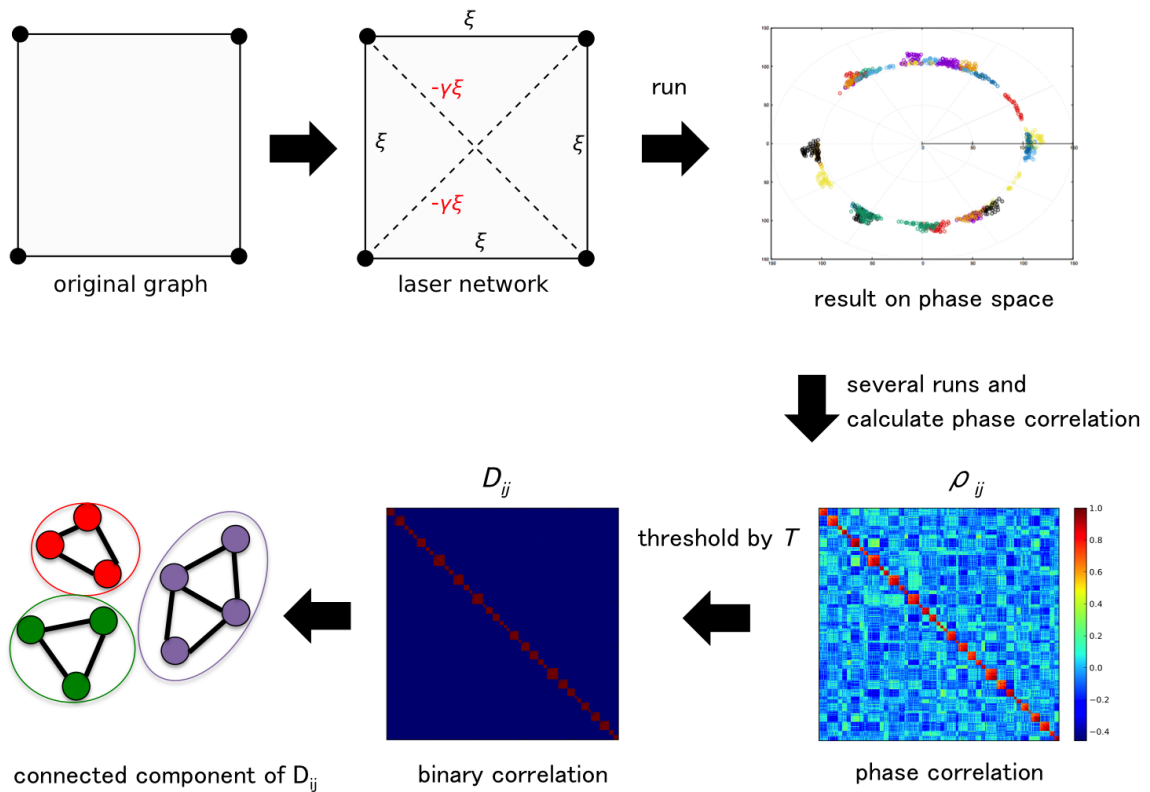


図 4.4. コミュニティ検出全体の流れ。

4.4 考えられる実験的な実装方法

本研究では実際にレーザーネットワークを構成した実験系を用いたベンチマークは行っていないが、現在考えられる実装方法を説明する。実験系を構築するにあたり最も重要であるのがスケラビリティである。2章で説明した理論モデルは、図 2.2 のように、単一モードレーザーを単純に複数結合したモデルであった。しかしながら実際に共振器をノード数 N 個分用意して更に $N(N-1)/2$ 個の結合路を実装することは現実的でない。

そこで以上のような問題を解決し、巨大なノード数及び任意のレーザー間での結合を実現するために、図 4.5 のように単一のファイバリング共振器の中に多数のレーザーパルスを生じさせるような時分割多重のネットワークを考える。このようなネットワークの実装方法は光パラメトリック発振器 (Optical Parametric Oscillator)^{*1}において既に提案されている [27]。この実装によって個々のパルスは互いに独立なレーザーと見なすことができ、遅延線を用いたパルス間の相互注入により結合を実現することによってネットワークとして動作させる。遅延線においては、周回ごとの光パルスの状態をホモダイン検波により読み出し、FPGA を用いた電気的なフィードバック回路により振幅・位相変調器を駆動させ信号パルスの隣を走る LO パルスを変調し、信号パルスに干渉させることによって個々のパルス間の結合を実装する。これにより、1本の遅延線で任意の結合を実装できる。

4.5 数値シミュレーション

本研究において、提案した手法の評価は全て数値シミュレーションによるものである。具体的には、オイラー・丸山法を用いてレーザーネットワークのダイナミクスを数値的に解くことによりベンチマークを行なった。計算の便宜上、数値計算に用いた式は共振器光子の寿命で規格化した時間 $\tau = t/\tau_c = \gamma_c t$ を用いて式 (4.4) を書き直した次式を用いる。

$$dA_j = -\frac{1}{2} \left(1 - \frac{p/p_{\text{th}}}{1 + \beta|A_j|^2} \right) A_j d\tau + \sum_{k=1, k \neq j}^N \tilde{\xi}_{jk} A_k + \sqrt{\tilde{D}} dW \quad (4.7)$$

ここで $\tilde{\xi} = \xi/\gamma_c$ 、 $\tilde{D} = D/\gamma$ である。本説では、コミュニティ構造が明確なベンチマークグラフを用いて簡単な動作確認を行なうことを目的とし、詳細な性能評価及び解析は次章に譲る。

動作確認に用いたネットワークは planted l -partition model [28] と呼ばれる生成モデルをもとに、Girvan と Newman によって提案されたモデル [29] を使用した。具体的には、グラフは図 4.6 のように 32 ノードで構成される 4 つの等価なコミュニティを持つ。すなわち、全ノード数 N は 128 である。また各ノードの次数 k は 16 であり、そのうち平均的に 0.9 の割

^{*1} 共振器の中に、レーザーにおけるゲイン媒質の代わりに非線形光学結晶を配置したもの。非線形光学効果によって、ポンプ光とは異なる波長の光で発振することが出来る。特にポンプ光の半分の波長で発振するものは、縮退光パラメトリック発振器 (Degenerate OPO) と呼ばれ、ポンプ光に対して位相は 0 か π の二つの値に限られる。

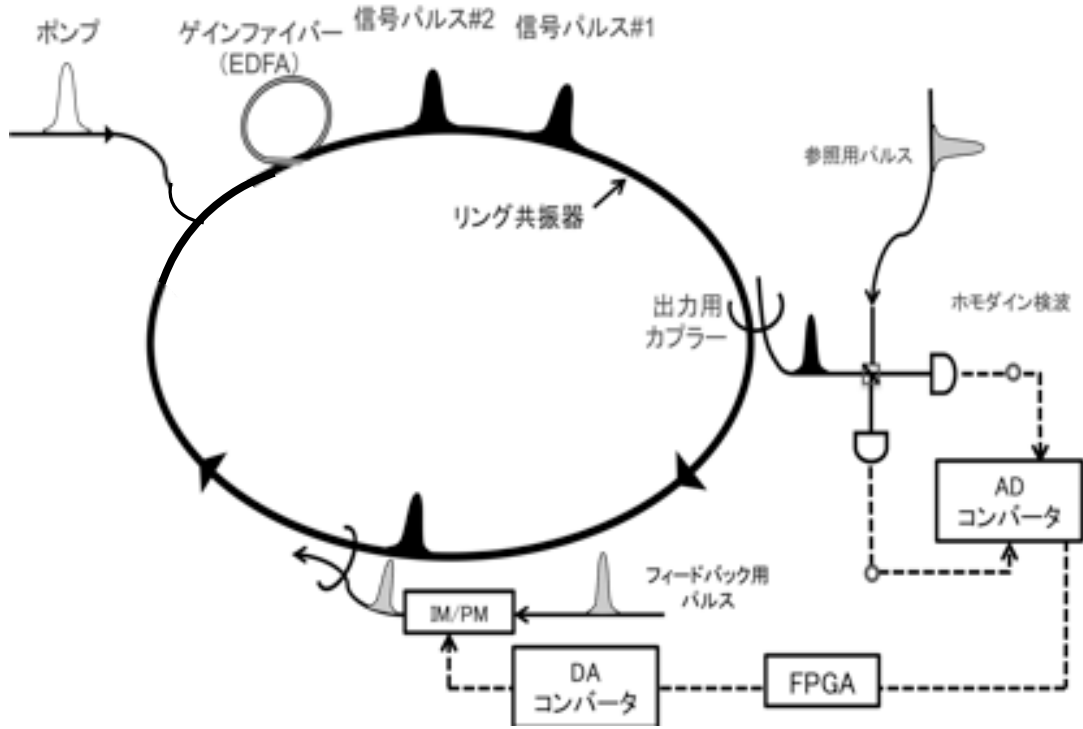


図 4.5. 時分割多重による実験的な実装方法。

合で同一コミュニティ内の他のノードと結合しており、0.1の割合で他のコミュニティのノード結合している。即ち全エッジ数のうち、90%がコミュニティ内で結合されているエッジであり、10%がコミュニティ間を繋ぐエッジである。数値計算に用いたパラメータは時間の刻み幅 $\Delta\tau = 0.05$ 、結合比 $\tilde{\gamma}_c = 0.1/\langle k \rangle$ 、自然放出雑音の分散 $\tilde{D} = 0.5$ 、ポンプレート $p/p_{th} = 2.0$ 、逆位結合と同位相結合の強度比 $\gamma = 1$ である。ステップ数は $n = 10000$ で、従って $\tau = 500$ まで計算を行なった。

図 4.7 にシミュレーションによって得られたレーザーネットワークの位相の時間発展を示す。位相空間上の点は各レーザーを表し、予め埋め込まれたコミュニティによって色づけされている。振幅が 0 の初期状態からランダムなノイズによって駆動され、 $\tau = 20$ で各レーザーの位相はほぼランダムで発振している。更に時間発展させると、結合の影響が効きはじめて $\tau = 100$ では埋め込まれたコミュニティに別れ始めている様子が確認できる。 $\tau = 300$ になると位相はコミュニティを表すクラスター状態となっており、準平衡状態に達していることがわかる。図 4.8 にこのようなダイナミクスを 100 回試行し、 $\tau = 500$ でのアンサンブル平均により計算した位相の相関 $\rho_{ij} = \langle \cos(\theta_i - \theta_j) \rangle$ を示す。行列の行 i 及び列 j は埋め込んだコミュニティによって並び変えている。図を見ると相関が非常に大きいブロックができており、これは埋め込んだコミュニティ内でのみ相関が大きいことを意味し、期待していた動作が実現していることが確認できる。

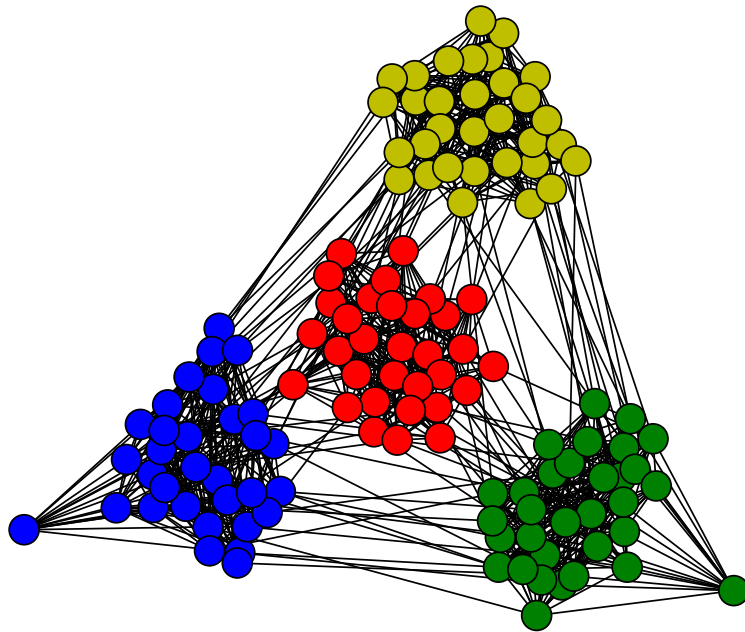


図 4.6. 動作確認に用いたネットワーク。

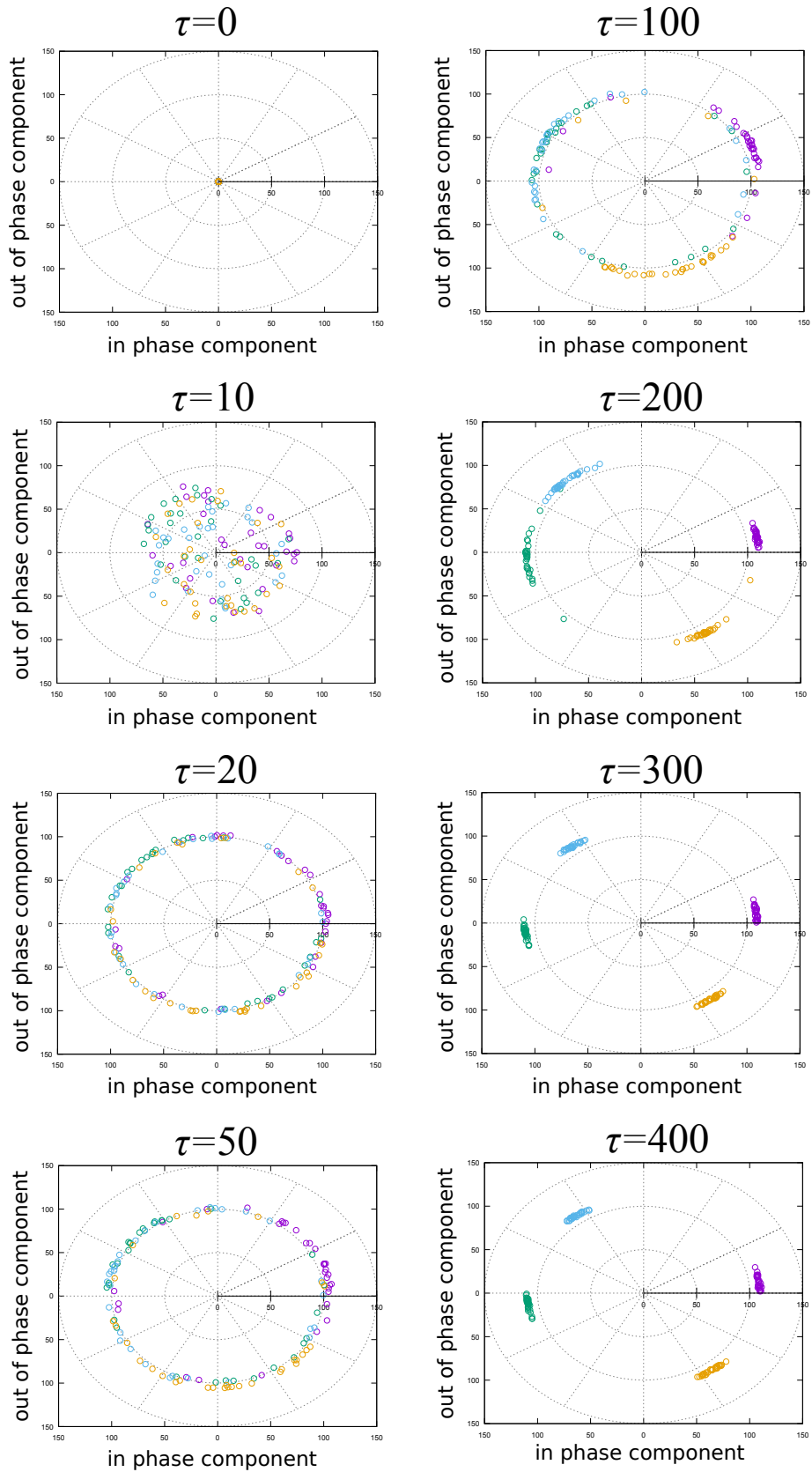


図 4.7. GN ベンチマークグラフにおける、レーザーネットワークの位相空間上のダイナミクス。色は予め埋め込まれたコミュニティを表す。

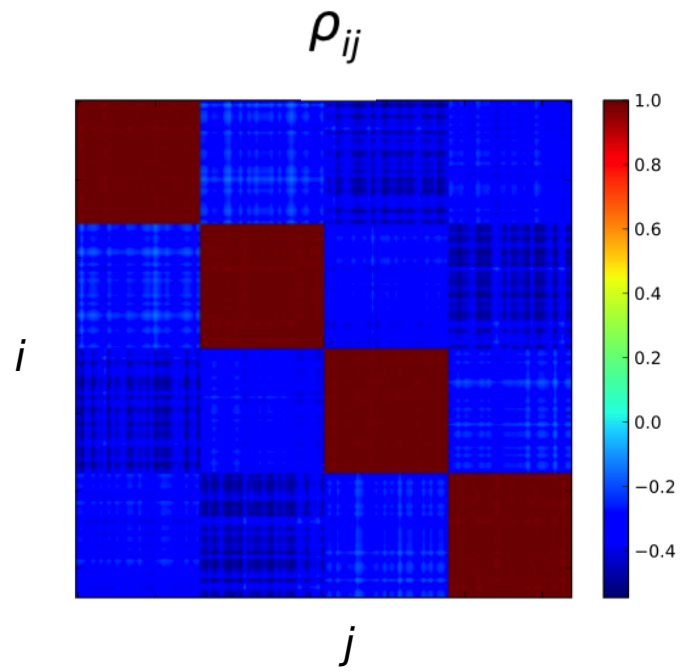


図 4.8. 100 回試行による位相の相関行列 $\{\rho_{ij}\}$

第 5 章

性能評価及びパラメータ依存性

本章では、前章で説明したレーザーネットワークを用いたコミュニティ検出手法について、数値シミュレーションを用いてより詳細な評価を行なった結果を報告する。手法の性能については、Lancichinetti らによって提案されたベンチマークグラフ (LFR ベンチマーク) を用いて rrNMI により評価を行なった。その結果、rrNMI による評価では既存手法と比較しても遜色のない性能を示す結果が得られた。更に系が準平衡状態に達するまでの時間、すなわちアルゴリズムの計算時間についての結果も示す。また、アルゴリズムが最適に動作するパラメータについても考察する。

5.1 LFR ベンチマークグラフ

4.5 節で動作確認に用いたグラフは、次数の分布が均一でコミュニティも全て同じサイズであるという非常に極端なモデルのもとに生成されたグラフであった。しかしながら、現実中存在するネットワークにおいてコミュニティのサイズ及び次数の分布が均一であるということは稀であり、特徴的な次数やコミュニティサイズを示さないスケールフリー性をもつことが多い。次数分布がスケールフリー性をもつランダムグラフとしては Barabási-Albert モデル [3] が有名であるが、これはコミュニティ構造をもたない。そこで Lancichinetti, Fortunato, Radicchi はコミュニティ構造を人工的に埋めこんだ planted l -partition model [30] を拡張して、次数及びコミュニティサイズの分布がべき分布になるようなモデルを提案し、それをを用いて様々な既存のコミュニティ検出アルゴリズムのベンチマークを行なった [31, 32]。この現実に近いコミュニティ構造をもつグラフ生成モデルは、提案した 3 人の頭文字をとって LFR ベンチマークグラフと呼ばれている。

LFR ベンチマークグラフは予め、次数分布の指数 τ_1 、コミュニティサイズの分布の指数 τ_2 及び各ノードの次数のうちコミュニティ間を結合するエッジの割合 μ を仮定する。このようなパラメータを設定することにより、ノード数 N 及び平均次数 $\langle k \rangle$ を持つグラフは以下の手順により生成される。

1. 各ノードに、仮定された指数 τ_1 をもつべき分布に従うように次数を割り当てる。この

とき最大次数 k_{max} 及び最小次数 k_{min} は平均次数が $\langle k \rangle$ となるように選ばれる。また、ノードは configuration model に従って次数分布を保ったまま他のノードと結合される。

2. 各ノードは $1 - \mu$ の割合で同じコミュニティに属するノードと結合し、 μ の割合で異なるコミュニティに属するノードと結合する。この μ は混合パラメータと呼ばれる。
3. コミュニティサイズの分布を、仮定された指数 τ_2 をもつべき分布に従うように決定する。このとき、全てのコミュニティサイズの合計が全ノード数 N となるようにする。また、最大コミュニティサイズ S_{max} 及び最小コミュニティサイズ S_{min} は、全てのノードがいずれかのコミュニティに属することができる条件 $S_{min} > k_{min}$ 及び $S_{max} > k_{max}$ を満たすように選ばれる。
4. 各コミュニティにノードを割り当てる。全てのノードがどのコミュニティにも属さない状態から出発し、ノードをランダムにコミュニティに割り当てていく。このとき、ランダムに選んだコミュニティのサイズが、 μ で定義されたノードの内部次数（次数と μ の積）を下回っていればそのノードはコミュニティに割り当てずそのままにしておく。これを全てのノードがコミュニティに割り当てられるまで繰り返す。
5. 各ノードの次数を保ったまま、 $1 - \mu$ の割合で同じコミュニティに属するノードと結合するようにエッジのランダムな繋ぎ替えを行なう。

総エッジ数を m とすると、このグラフ生成アルゴリズムは $O(m)$ 程度で収束することが示されている。図 5.1 に LFR ベンチマークグラフの作成例を示す。図 4.6 と比べて次数、コミュニティサイズが広く分布していることがわかる。

5.2 同位相結合と逆位相結合の注入比

4.5 節での動作確認では、同位相結合と逆位相結合の注入は同じ強度 ($\gamma=1$) であった。しかしながら、エッジの存在しないノード間には全て逆位相結合での注入を行なっているため、疎なグラフでは全体として逆位相結合が支配的になる。このような状況では、局所的にも同位相結合の効果よりも逆位相結合の効果が強くなり、上手くコミュニティを検出できない可能性がある。実際、現実のネットワークはほとんど疎でありそもそもグラフがコミュニティ構造をもつということはグラフ全体としてはある程度疎でなければならない。これはノード数が大きくなる程顕著である*1。

4.5 節でのシミュレーションでは μ が小さく、かつコミュニティサイズが均一というかなり理想的なグラフであったので期待した動作を示したが、LFR ベンチマークグラフなどの現実

*1 次数分布が同じままでグラフの全ノード数 N を増やした場合、エッジ密度 D は N に線形に増加する。

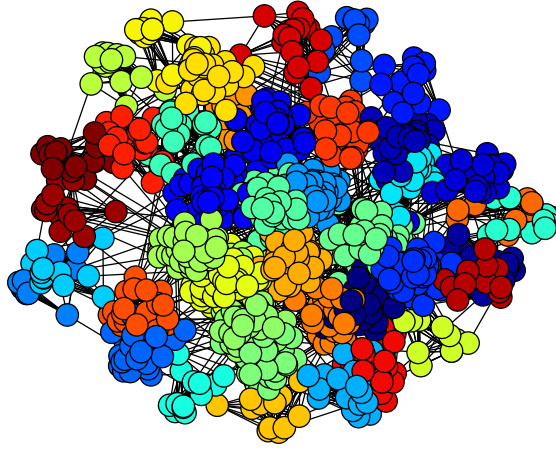


図 5.1. LFR ベンチマークグラフの作成例。ノード数 $N = 1000$ 、次数分布の指数 $\tau_1 = 2$ 、コミュニティサイズ分布の指数 $\tau_2 = 1$ 混合パラメータ $\mu = 0.1$ 。

に近いグラフで同じような動作をするかは自明ではない。一方で逆位相結合の効果を小さくするために γ を小さくしすぎると、同位相結合が効果が大きくなり、全体として同位相で発振してしまう。図 5.2 に 4.5 節での数値シミュレーションと同じ条件で γ のみを変化させた場合の位相状態を示す。 $\gamma = 0.02$ のときは、同位相結合が支配的になっているためにすべてのレーザーがほぼ同位相で発振している。 $\gamma = 0.04$ では逆位相結合の影響により各レーザーの位相はコミュニティのクラスターに別れはじめて、 $\gamma = 1$ では完全にコミュニティに別れている。更に逆位相結合を大きくした $\gamma = 2$ の場合、位相は 360° に散らばって発振する。この例ではたまたま $\gamma \sim 1$ 周辺が最適なパラメータであったが、これはグラフの総ノード数が 128 と少なく、エッジ密度 $D = 2m/N(N-1)$ が約 0.125 と比較的大きいことが原因であると考えられる。

コミュニティ検出においてこのような逆位相結合と同位相結合の強度バランスは非常に重要であり、また最適なパラメータはネットワークに依存する。基本的にはエッジ密度が小さくなる程全体として逆位相結合が支配的になるため、エッジ密度に依存して同位相結合と逆位相結合の強度比 γ を最適化しなければならない。そこでエッジ密度に依存して、強度比 γ が系全体に及ぼすマクロな影響の関係を明確にするために、ランダムなコミュニティ構造を持たないグラフにおいて同様のレーザーネットワークを構成し、次式で定義される系の秩序パラメータを評価した。

$$Re^{i\Phi} = \frac{1}{N} \sum_j e^{i\theta_j} \quad (5.1)$$

ここで R はレーザー集団のコヒーレンスを表し、全てのレーザーが同位相で発振していると 1 となり散らばって発振していると 0 となる。すなわち R によって、系全体としてどの程度同

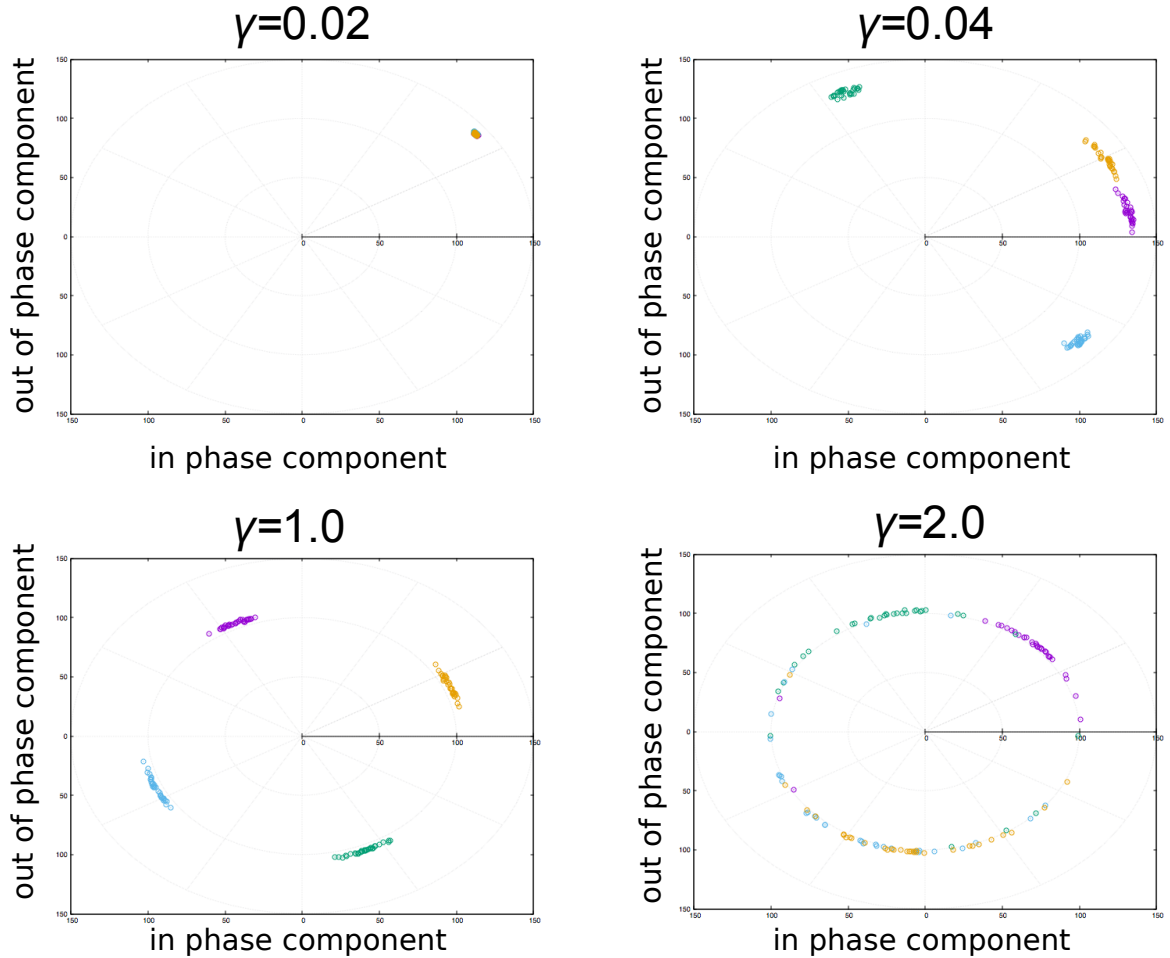


図 5.2. GN ベンチマークグラフにおける、位相状態の γ 依存性。

位相及び逆位相結合の影響を受けているのかを定量的に評価することができる。また Φ は系のマクロな位相であり、各レーザーの位相の平均値を表している。秩序パラメータの評価に用いたランダムグラフは Erdős-Rényi モデル [1] 及び Barabási-Albert モデル [3] よばれるグラフである。Erdős-Rényi モデルは通常のランダムグラフとして広く用いられており、各ノード間に独立に確率 p でエッジを張るというモデルである。従ってエッジ密度 D は p を変更することで調整できる。このグラフの次数分布は二項分布である。一方 Barabási-Albert モデルは次数分布がべき分布となるようなランダムグラフである。

図 5.3 に異なるエッジ密度 D に対する R の γ 依存性の数値シミュレーション結果を示す。Erdős-Rényi モデル及び Barabási-Albert モデル両者共に γ が小さい領域では R が大きく逆に γ を大きくすると R が小さくなっていることが分かる。またエッジ密度 D が大きいと、同位相結合に対する逆位相結合の影響が小さくなるため、同じ γ の場合は D が大きい方が R も大きい。ここで注目すべき点は、次数分布の違いによって Erdős-Rényi モデルと

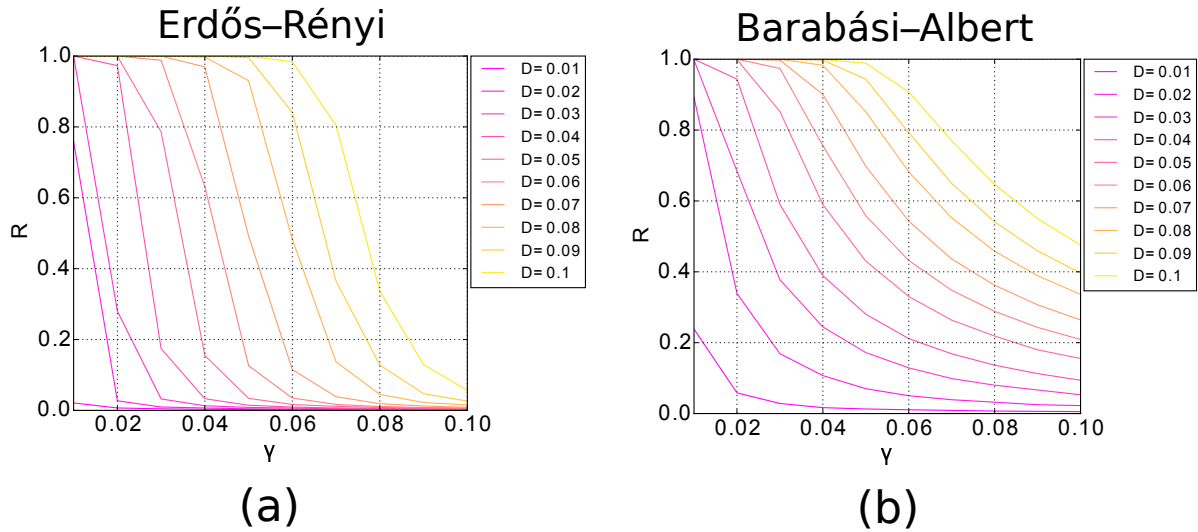


図 5.3. ランダムグラフにおける秩序パラメータ R の γ 依存性。全ての D について 100 回試行の平均をプロットしている。 $N = 1000$ で、実線の色は異なるエッジ密度 D を表している。(a)Erdős-Rényi モデル。(b)Barabási-Albert モデル。

Barabási-Albert モデルで R が 1 から減少し始める γ^{*2} が異なる点である。Barabási-Albert モデルでは、次数のスケールフリー性のために少数のハブノード^{*3}が存在する。ハブノードは同位相結合の効果が大きく、従ってグローバルな位相同期から外れにくい。一方でハブ以外のノードは逆位相結合が支配的になるため同位相の秩序からは外れやすい。その結果系全体としては次数が小さいノードが圧倒的に多いため、Erdős-Rényi モデルと比較してグローバルな位相同期から外れやすくなっていると考えられる。一方、グローバルな位相同期から外れた γ の領域では Barabási-Albert モデルの方が、 γ の増加に対する R の減少が遅いことが確認できる。これは、ハブノードを含む同期されたクラスターが同期から外れにくいためであると考えられる。このような現象は統計物理学におけるパーコレーション相転移と密接に関連していると考えられるが、本研究の主旨から逸れるためこれ以上の詳細な解析については本論分では触れない。

さて、我々の目的はコミュニティ検出を行なう際の γ の最適化である。我々のコミュニティ検出のアプローチは、ネットワークのコミュニティ構造に起因するレーザーの同期のダイナミクスを利用するものであった。すなわちランダムグラフと比較してコミュニティ構造が存在すると同じ D 及び γ でも局所的な同位相、逆位相結合の効果が変化し、その変化によってコミュニティを検出するというものである。従って同じエッジ密度 D において、ネットワークの構造の変化で秩序パラメータが変化しやすい γ の領域が望ましい。そのような観点から図 5.3 をみると、コミュニティ検出を行なうグラフと同じエッジ密度をもつランダムなグラフにおい

*2 全てのレーザーが同位相で発振しなくなる γ 。

*3 次数が他のノードよりも極端に大きいノード。

て、 $R \sim 0.5$ となる γ での動作が位相パターンが敏感に変化する領域であると考えられる。

以上の仮説を確かめるために、LFR ベンチマークグラフにおいて γ を変化させたコミュニティ検出の数値シミュレーションを行い、rrNMI により精度を評価した。図 5.4(a)~(c) は $N=1000$ 、(e)~(g) は $N=5000$ の LFR ベンチマークグラフで横軸 μ は混合パラメータである。 μ が大きくなる程コミュニティ間を結合するエッジの数が多くなるので、コミュニティ構造が弱いことを意味する。また、 T は位相の相関行列 $\{\rho_{ij}\}$ をバイナリ値に変換する際の閾値である。基本的に μ が大きくなる程、埋め込まれたコミュニティを検出するのは難しくなる。従って、どれだけ大きい μ までコミュニティを検出できるかというのは一つのアルゴリズムの性能を示す指標になっている。これらの図から、殆どの T において、 $N = 1000$ では $\gamma = 0.015$ 、 $N = 5000$ では $\gamma = 0.003$ の精度が良いことが確認できる。図 5.4(d)、(h) はそれぞれ $N = 1000$ 、 $N = 5000$ の Barabási-Albert モデルにおける R の γ 依存性で、各点は LFR ベンチマークグラフと同じエッジ密度かつ (a)~(c)、(e)~(g) と同じ γ における R を表している。LFR ベンチマークで精度の良い γ は、 $N = 1000$ 、 $N = 5000$ の両者共に同じエッジ密度のランダムグラフ (Barabási-Albert モデル) で $R \sim 0.5$ を実現するパラメータ領域にあることがわかる。

図 5.4 は barabási-Albert モデルにおいて、 $R = 0.5$ となる γ のエッジ密度依存性のグラフである。各点がシミュレーション値で実線が線形フィッティング、破線が次式で表される直線を示す。

$$\gamma = D - 5/N \quad (5.2)$$

$R = 0.5$ を満たす γ はエッジ密度に対してほぼ線形に増加しており、式 (5.2) によって良く近似できている。第 2 項は N が有限であることに起因する補正項であると考えられる。従ってコミュニティ検出の際の最適な γ は式 (5.2) により決定できる。

5.3 位相相関の閾値

前節では同位相結合と逆位相結合の強度比 γ の最適な値について考察したが、コミュニティ検出を行なうにあたってもう一つ重要なパラメータが、位相相関行列 ρ_{ij} をバイナリ値変換する際の閾値 T である。 T を小さくしすぎると、本来異なるコミュニティに所属するノード間の小さい相関も検知することになり、一方で T を大きくしすぎると同じコミュニティに属している程度相関を持っているにも関わらずそれを検知出来ない。コミュニティ構造が比較的明確な場合は位相相関のコントラストも明確になり T はあまり問題にならない。しかしながら LFR ベンチマークにおける μ が大きいグラフのように、コミュニティ構造が弱い場合には T によって検出されるコミュニティの差が顕著になる。実際、図 5.4 を見れば明らかな通りコミュニティ検出の精度は、大きい μ の領域において T にかかなり依存している。

図 5.6 に LFR ベンチマークにおいてレーザーネットワークのダイナミクスを 100 回試行した際の、同一コミュニティ内及び異なるコミュニティに属するノード間の位相相関 ρ_{ij} のヒストグラムを示す。数値計算に使用した γ は前節で考察した最適なパラメータを用いた。 $N = 1000$ 、 5000 のグラフ両者とも $\mu = 0.1$ では位相相関のコントラストがはっきりして

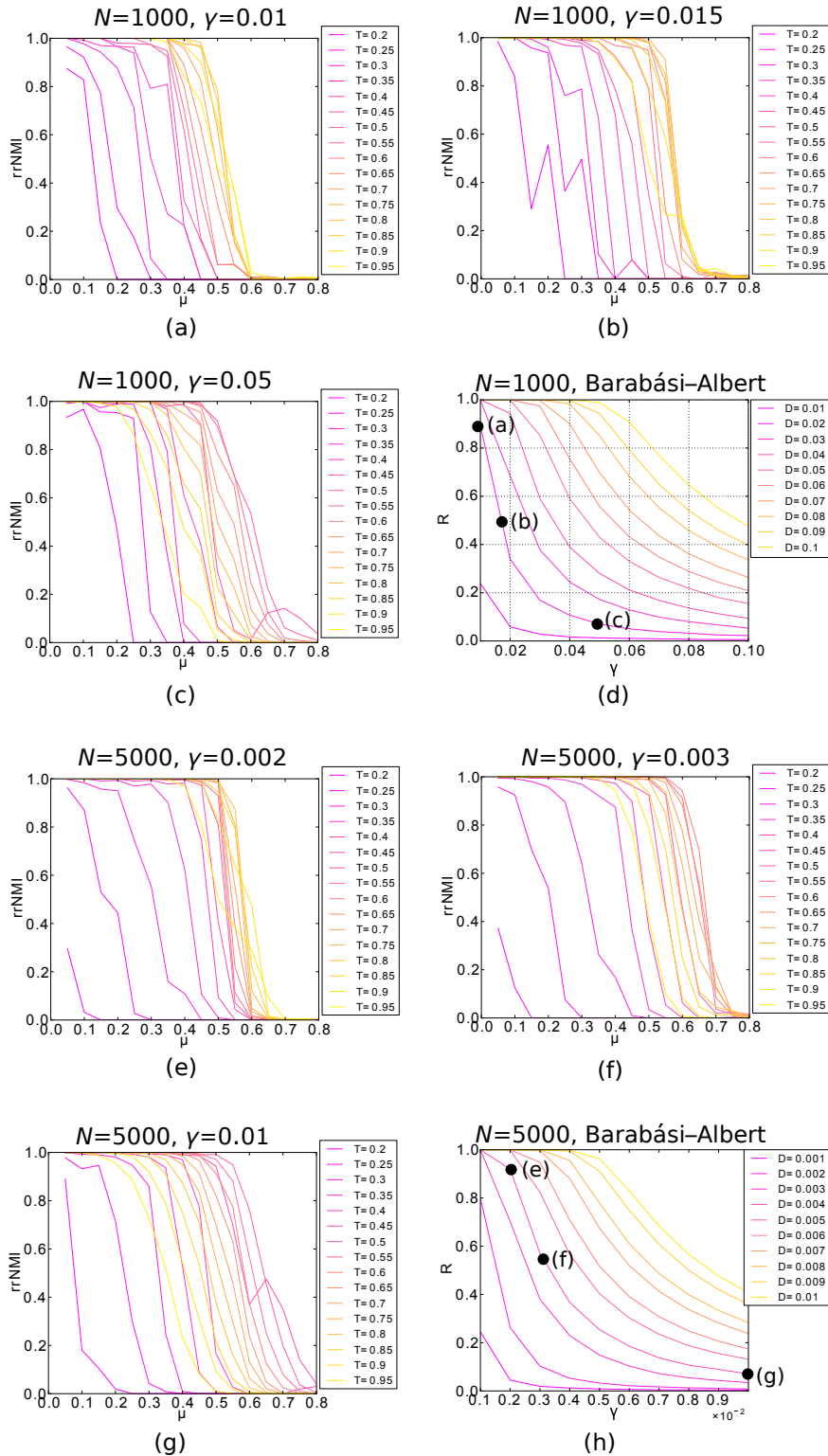


図 5.4. 異なる γ におけるコミュニティ検出の rrNMI による評価。(a)~(c) は $N = 1000$ 、(e)~(g) は $N = 5000$ の LFR ベンチマークグラフにおける結果で μ は混合パラメータを表す。また T は位相の相関をバイナリ値に変換する際の閾値である。(d)、(h) は $N = 1000$ 、 $N = 5000$ の Barabási-Albert モデルにおける R の γ 依存性のグラフで、点はそれぞれ (a)~(c)、(e)~(g) に対応する同じ D 、 γ における R を表す。

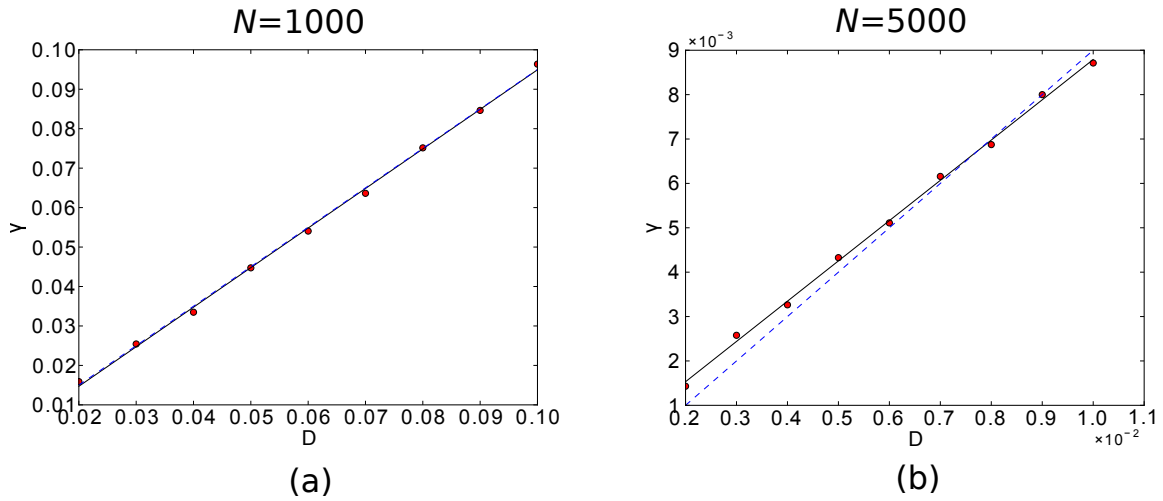


図 5.5. barabási-Albert モデルにおける $R = 0.5$ を実現する γ の D 依存性。各点が数値シミュレーションによる値で実線が線形フィッティング、破線は $\gamma = D - 5/N$ の直線を表す。

おり、コミュニティ内外で分布が明確に別れている。コミュニティ内とコミュニティ間の相関分布の別れている領域に閾値を設定すれば、正しいコミュニティを検出することができる。一方で μ が大きくなるに従ってコミュニティ間を繋ぐエッジの数が増え、二つの分布の距離が小さくなることを確認できる。 $N = 1000$ 、 $\mu = 0.5$ のグラフでは埋め込まれたコミュニティを完全に再現できる限界となっており、 $\mu = 0.8$ ではもはや閾値を何処に設定しても埋め込まれたコミュニティを検出するのは不可能である。このヒストグラムによって、少なくとも「LFR ベンチマークグラフに埋め込まれたコミュニティを検出する」という意味では、 N によらずに閾値 T を 0.6 あたりに設定するのが最適であると考えられる。

また $\mu = 0.8$ では $N = 1000$ 、5000 共のグラフともに、ヒストグラムのピークが 0 及び 1 に近い領域の二つに存在することが確認できる。これはレーザーネットワークを γ の値がランダムグラフで $R = 0.5$ となるパラメータ領域で動作させているので、 μ が大きいランダムに近いグラフではグローバルな秩序が現れ、高い相関が生じていることが原因であると考えられる。一方 μ が小さい領域では同じ γ のパラメータでも、コミュニティ間にまたがるグローバルな秩序は現れない。これは同じエッジ数の場合、コミュニティを形成している状態ではノード間の平均距離 (average path length)^{*4} がランダムグラフよりも大きくなるために、グローバルなレーザーの位相の情報伝達が困難になることが原因であると考えられる。言い換えると、このようなネットワーク構造の違いに起因するレーザーのダイナミクスの変化を利用してコミュニティ検出を行なっているのである。

*4 ノード間の距離 (path length) とは 2 ノード間の最短経路における辺数のことを言う。平均距離とは全てのノード間における距離の平均である。

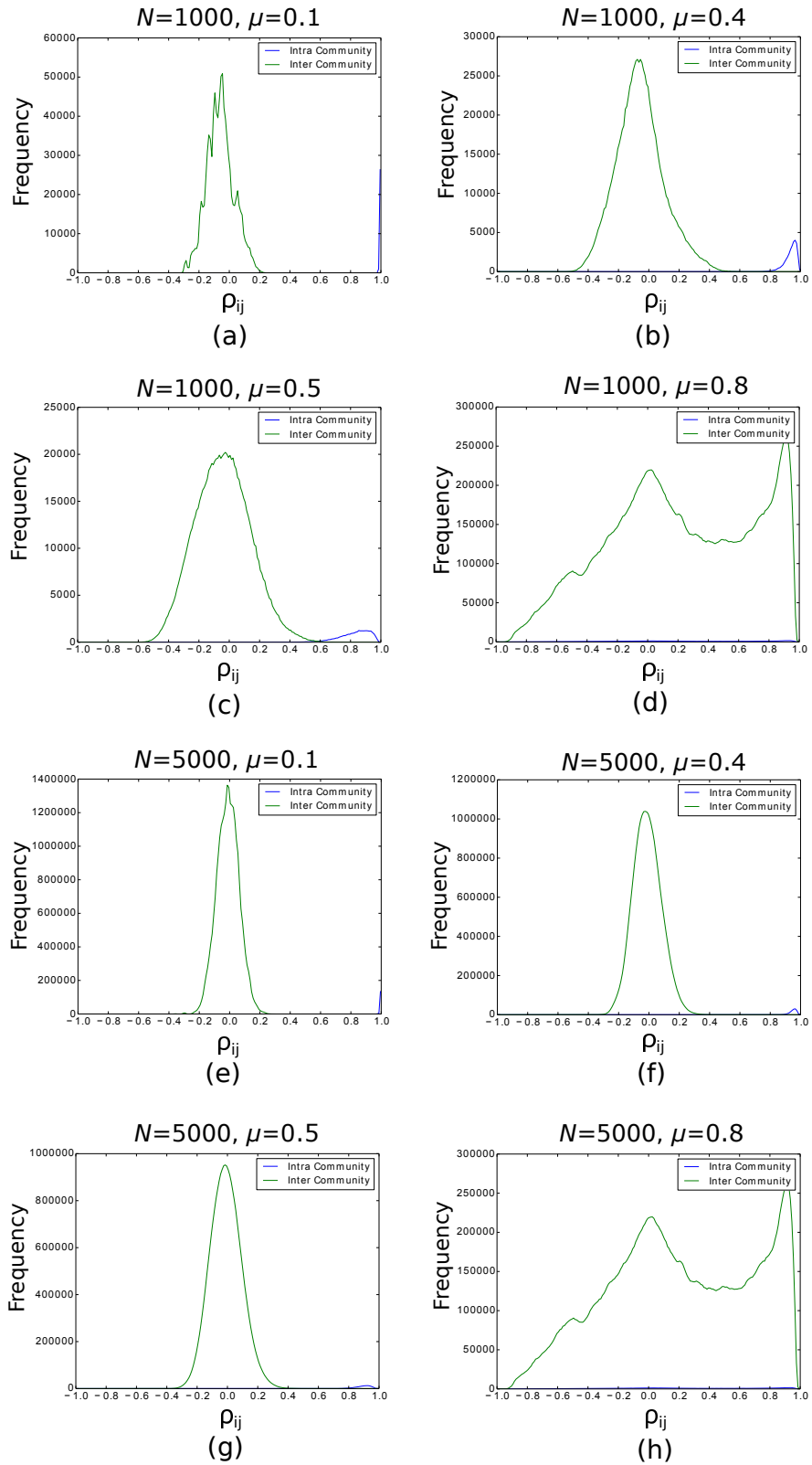


図 5.6. LFR ベンチマークにおける、コミュニティ内外における位相相関 ρ_{ij} のヒストグラム。緑がコミュニティ間、青がコミュニティ内の相関を表す。

5.4 rrNMI による既存アルゴリズムとの性能比較

3.4 節では代表的なコミュニティ検出アルゴリズムとして、Louvain 法と Infomap の二つを紹介した。本説ではこれらのアルゴリズムとレーザーネットワークについて、LFR ベンチマークグラフを用いた rrNMI による性能比較を行なう。

まず、多数のアルゴリズムを網羅的に評価し比較した例として有名な、Lancichinetti らによる LFR ベンチマーク [31] の結果を図 5.7 に示す。ここで、Blondel et al. が Louvain 法であることを注意されたい。この結果は $N = 1000$ 、 5000 の LFR ベンチマークグラフによるもので、rrNMI ではなく NMI による評価である。図中の S 及び B はコミュニティサイズが $S:10 \sim 50$ 、 $B:20 \sim 100$ であること表す。また、図 5.8 に全く同じパラメータを用いて作成した LFR ベンチマークグラフにおいて、レーザーネットワーク、Louvain 法そして Infomap のシミュレーション結果の rrNMI による評価を示す。ただしレーザーネットワークの結果については前節及び前々節で考察した最適なパラメータを用いた数値シミュレーションによるもので、位相相関は 100 回の試行における平均により求めた。比較対象としてこの 2 つのアルゴリズムを選んだのは Louvain 法は現在最も高速なアルゴリズムの一つであり、Infomap は図 5.7 より明らかな通り、Lancichinetti らによる LFR ベンチマークの精度が非常に良いことが理由である。また、Louvain 法についても Infomap と並んで LFR ベンチマークにおける NMI の評価は高いことが確認できる。

図 5.7 と図 5.8 の比較については、rrNMI は NMI の有限サイズ効果による問題を改善したものである。共に 1 である埋め込まれたコミュニティを完全に再現している領域のみ意味がある。従って、各アルゴリズムが埋め込まれたコミュニティを完全に再現できなくなる。すなわち rrNMI や NMI が 1 から落ち始める境界点が重要である。このような観点からレーザーネットワークを他のアルゴリズムと比較すると図 5.7 においては Infomap、RN を除くほとんどのアルゴリズムよりも精度が高いことが確認できる。また、rrNMI のみによる評価についても、レーザーネットワークは Louvain 法よりも精度が高いといえる。しかし、Infomap と比較すると精度は劣っている。今回行なった評価においては、レーザーネットワークは精度面において現行のほとんどのアルゴリズムと比較して遜色がないといえる。

5.5 計算時間

最後に、レーザーネットワークの計算時間について議論を行なう。レーザーネットワークがコミュニティの検出を行なうトータルの計算時間は、レーザーのダイナミクスが準平衡状態に達するまでの緩和時間と試行回数の積、及び位相相関によって連結成分を求める後処理にかかる時間の合計で見積もることができる。後者については、位相相関を元グラフにおいて結合しているノード間しか計算しないとすると、グラフの連結成分分解の計算量と同じなので $m \gg N$ を仮定すると $O(N + m) \approx O(m)$ である。前者については、次式で定義されるエッジの秩序

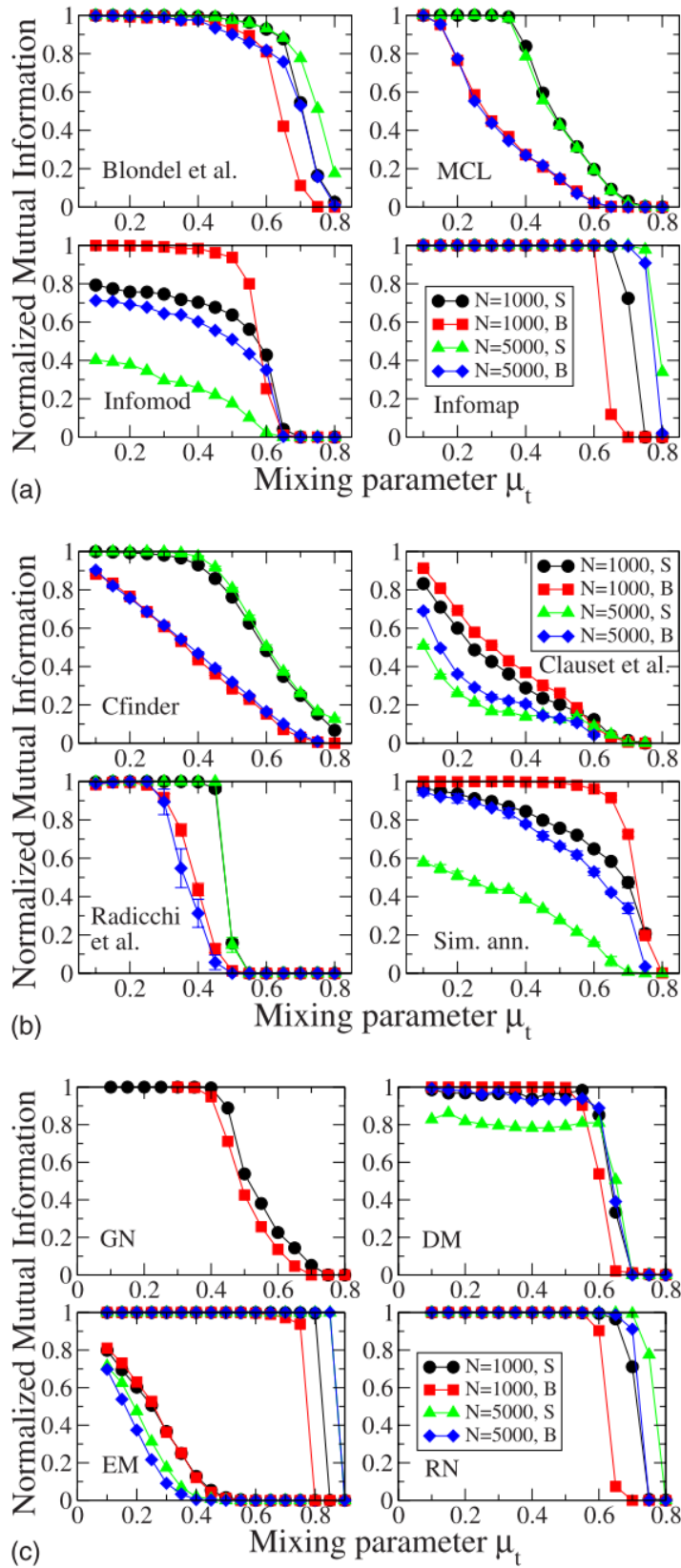


図 5.7. 様々なアルゴリズムにおける LFR ベンチマーク (Lancichinetti et al.[31] より引用)。

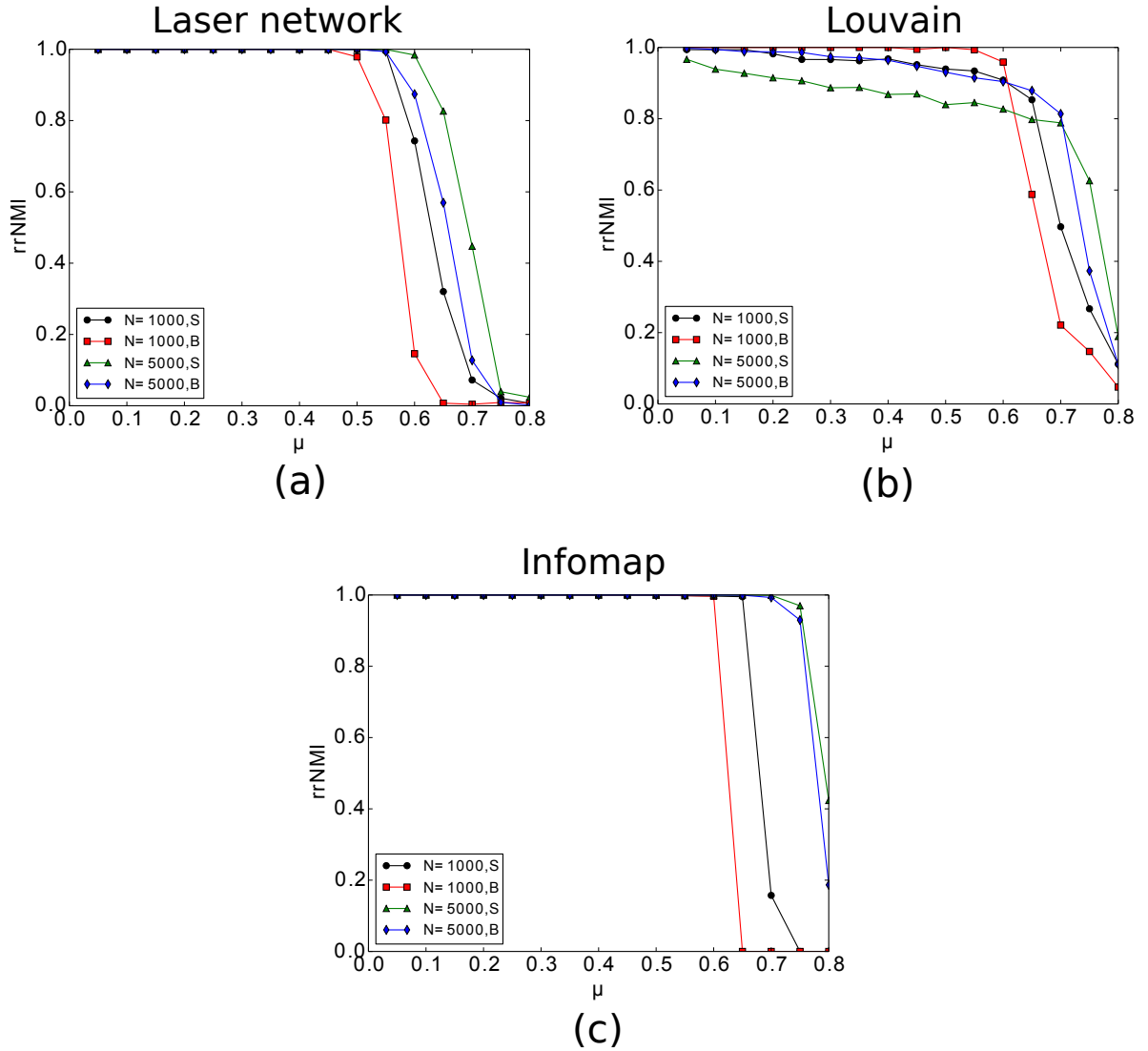


図 5.8. LFR ベンチマークによるコミュニティ検出アルゴリズムの性能評価。(a) レーザーネットワーク (b) Louvain 法 (c) Infomap。

パラメータ R_{link} を用いてダイナミクスの緩和時間を見積もった。

$$R_{\text{link}} = \frac{1}{m} \sum_{(i,j) \in E} \cos(\theta_i - \theta_j) \quad (5.3)$$

R_{link} は、同位相で結合しているノード間のグローバルな秩序を表し、それによってノードの秩序パラメータ R では表現できない同一位相のクラスターが複数散らばっているような状態も考慮に入れることが出来る。従って R は小さくても、元グラフにおいて結合が密なノードの位相が近い値であれば R_{link} は比較的大きい値となる。

図 5.9 に、前節で rrNMI を評価した数値シミュレーションにおける R_{link} の時間変化を示す。基本的にダイナミクスは μ に依存し、 μ が小さいほど緩和は速い傾向にある。通常の間

位相結合のみの場合の同期の場合は、ランダムグラフに近いほど平均距離が短いために緩和は早くなると予想されるが、今回の結果はその逆である。これは、元グラフにおいて結合がないノード間全てに逆位相結合が入っているために、同期のフラストレーション^{*5}が生じることに起因している。グラフがランダムに近い場合には局所的な構造がないためにこのようなフラストレーションはグラフ全体の全てのノードに広がり、結果的にダイナミクスの収束は遅くなると考えられる。一方でコミュニティ構造が強い場合には、同位相結合が密な領域がコミュニティ内で局在しているために、そのような領域においては逆位相の影響が少なくフラストレーションも小さい。結果的にフラストレーションはコミュニティ単位で生じることになり、ダイナミクスの収束はランダムなグラフと比べて小さくなっていると考えられる。しかし、緩和の速さではなく R_{link} 自体の大きさは μ が小さい領域では μ と R_{link} の大小関係は対応しているが、 $\mu > 0.6$ あたりの大きい領域では逆の大小関係になっている。これは、逆位相結合が複数ある領域では、必ずしも位相を離れさせようとするわけではないことを示唆している。例えば図 5.10 のように、あるノード間が逆位相結合を 2 つ介して結合している場合には、2 つのノードの位相は揃おうとする。実際のグラフは非常に複雑に結合しているため、高次の秩序^{*6}も存在し、このような単純な議論で説明できるわけではないが本研究の主旨から逸れるので、これ以上は立ち入らない。

さて、レーザーネットワークの緩和時間は R_{link} について $\gamma_c t = 5$ で 0.0001 以上の変化がなくなる時間で見積もった。図 5.11(a)、(b) に緩和時間のノード数依存性を示す。緩和時間はコミュニティサイズ S 、 B 及びノード数 N にはあまり依存しないが、前述の通りネットワーク構造にかなり依存している。同じ N においても、 μ が小さいコミュニティ構造が強いグラフでは μ が大きいランダムに近いグラフと比べて緩和時間は $1/3 \sim 1/5$ 程度になっている。コミュニティ検出アルゴリズムとしては、実際に適用するグラフはある程度コミュニティ構造が存在するグラフであることが多いので、ランダムに近いグラフで緩和が遅いことは問題にならないと考えられる^{*7}。また、緩和時間が N に依存しないのは非常に大きい。ただし実際の物理系を構築する場合には N が大きくなるにつれて共振器長が長くなり、共振器そのもので決まる光子の緩和レート γ_c が小さくなるために計算時間は N に線形に増加すると考えられる。ただし、一度共振器を作ってしまうと位相のダイナミクスそのものは問題サイズに依存しないといえる。実時間の目安としては、 $\gamma_c = 10^6 \sim 10^7 [1/\text{s}]$ 程度と仮定すると一回の試行につき、コミュニティ構造が強い場合には約 $0.5 \sim 5 [\text{ms}]$ 、100 回の試行では約 $50 \sim 500 [\text{ms}]$ となる。また、図 5.11(c)、(d) は最も高速なアルゴリズムである Louvain 法、(e)、(f) は Infomap の計算時間スケーリングである。これらのアルゴリズムの計算時間は N に対して線形に増加していることが確認できる^{*8}。Infomap に関しては 100 回の試行を考えてもレーザーネットワー

^{*5} 磁性体におけるスピンの自由度と結合の幾何学的な構造に起因する不安定性のことをフラストレーションと呼ぶ。レーザーネットワークの場合、逆位相結合がスピンの反強磁性結合に対応し、それによって原理的にレーザーの位相が完全に安定な配置をとることができなくなることをここではフラストレーションとよんでいる。

^{*6} 2 つではなく複数の結合を介した影響。

^{*7} 緩和が遅い場合はコミュニティ構造が存在しないとして、ある時間で切れば良い。

^{*8} 使用したグラフは N によらず同一の次数分布をもつため、エッジ数 m に対して線形に増加しているということもいえる。

クの緩和時間の方が速いが Louvain 法はそれよりも速い。ただ、レーザーネットワークは純粋な緩和時間のみを見積もっているため、この結果だけを見て単純に両者の比較はできない。

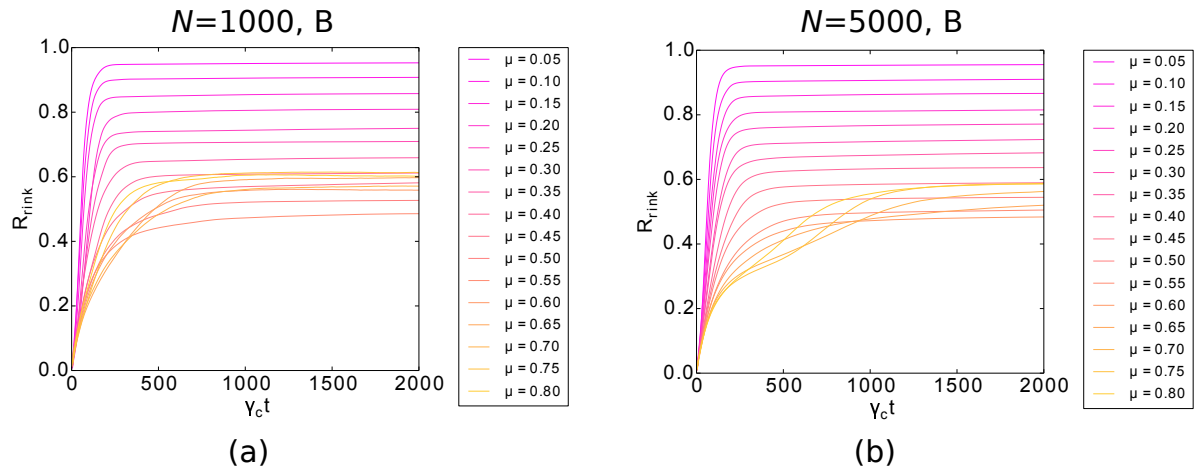


図 5.9. R_{link} の LFR ベンチマークグラフにおける時間依存性。

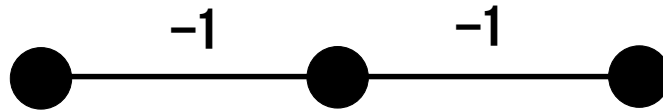


図 5.10. 二つの逆位相結合を介する結合例。

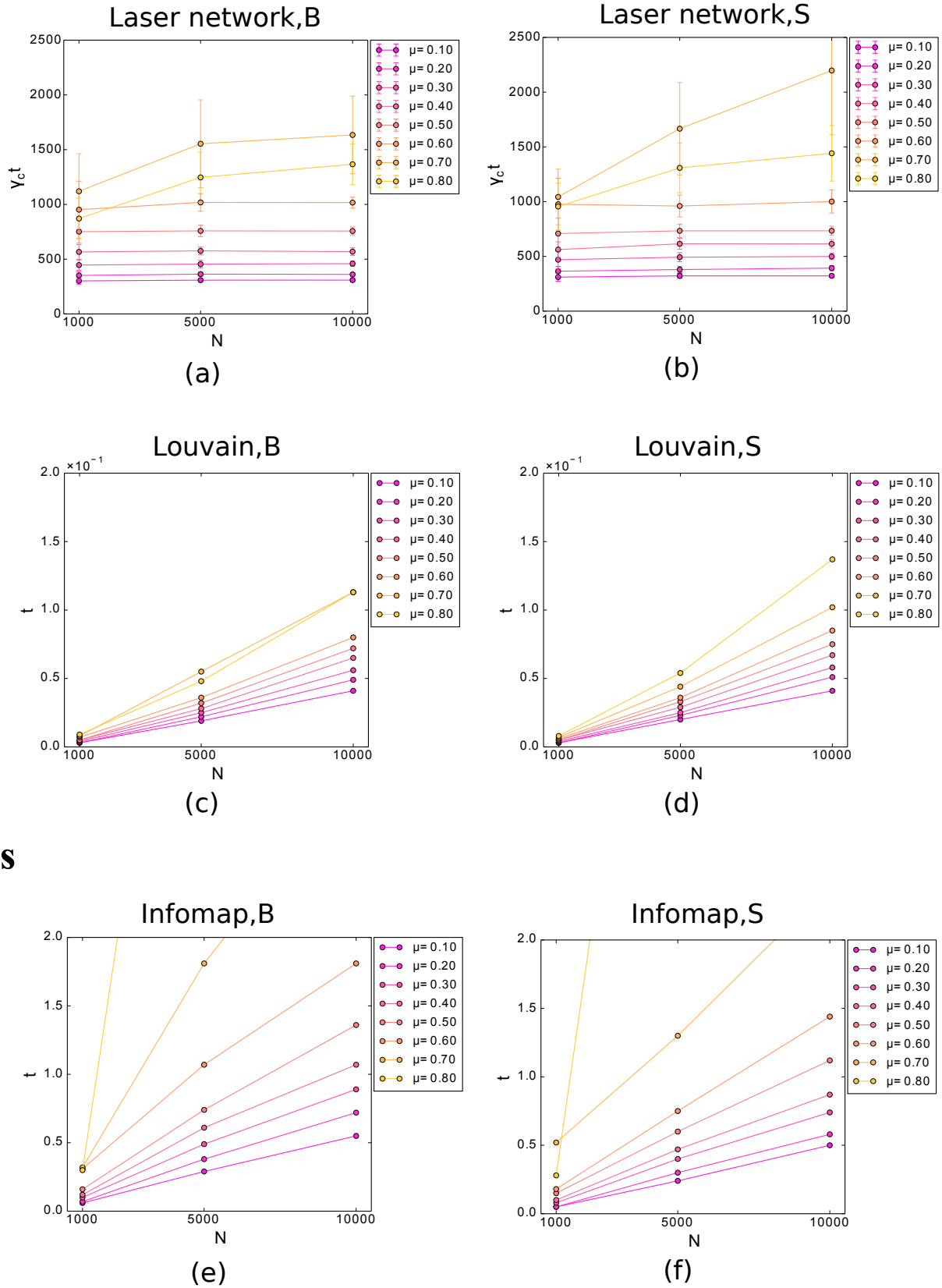


図 5.11. レーザーネットワーク及び Louvain 法の計算時間スケーリング。ただし (a)、(b) のレーザーネットワークは純粋な準平衡状態への緩和時間だけの結果であり、後処理は考慮に入れていない。

第6章

結論

レーザーネットワークの位相同期を用いたコミュニティ検出手法を提案した。元のネットワークにおいて結合していないノード間全てに適切な強さで逆位相注入を行なうことで、同一コミュニティ内の位相相関が継続することを数値シミュレーションによって確認した。精度については、LFR ベンチマークグラフにおいて予め埋め込まれたコミュニティ構造をどの程度再現できたかを $rrNMI$ によって評価した。その結果、既存の多くのアルゴリズムと比較して同程度以上の精度が得られた。また適切な逆位相結合の強さ γ はエッジ密度に依存し、コミュニティ検出を行なうネットワークと同じエッジ密度のランダムグラフにおいて秩序パラメータが 0.5 となる γ で動作させると、LFR ベンチマークにおける $rrNMI$ の意味での精度が良いことが示された。更に LFR ベンチマークグラフに埋め込まれたコミュニティ内外の位相相関のヒストグラムによって、多くの場合閾値 T を約 0.6 に設定すると精度が良くなることが示された。緩和時間についてはネットワーク構造に大きく依存し、コミュニティ構造がある場合には速く収束することを確認した。

今回は同位相注入及び逆位相注入はそれぞれ全て同じ大きさであったが、これを全て同じにせずモジュラリティ等の目的関数に対応する形で入れると精度はどう変化するのは興味深いので今後の課題としたい。また、現在の方法では位相相関を求めるために、位相差のアンサンブル平均を計算しなければならないため、一回のコミュニティ検出に複数の試行が必要である。従って、一度の試行でコミュニティ検出を行なうためにアンサンブル平均を時間平均で補えるかが重要になる。そのためには系にエルゴード性があるが鍵となるので、今後は系がそのような性質を持っているかについても調べていく予定である。

謝辞

まず最初に本研究を進めるにあたり、多大なるご指導を頂きました山本喜久教授に感謝致します。ご多忙の中、ミーティングに参加して頂いた際には常に広い視野で、かつ的確な視点で研究の方向性を示して頂きました。本当にありがとうございました。

本研究を進めるにあたり多くのご指導頂きました宇都宮聖子准教授に感謝致します。右も左も分からない状態で研究テーマも定まらない中、最初に研究の具体的な方向性について示して頂きました。

大学院に入学するにあたって、私たち学生を受け入れて頂いた安達淳教授に感謝致します。私が所属予定であった山本研究室が大学に存続しないことになり、急遽研究分野が異なるにもかかわらず快く受け入れて頂き、本当にお世話になりました。

研究を進めるにあたって多くのアドバイスを頂きました、博士研究員の玉手修平氏に感謝致します。席が隣であったこともあり、細かい議論にも付き合ってくださいました。

多くのアドバイスを頂いた、先輩の高田健太氏、松尾康弘氏、針原佳貴氏、神原太郎氏、同期の丸尾大貴君、後輩の竹田悠太可君にも感謝致します。また出張の際などあらゆる事務手続きを行なって頂き、研究をサポートして頂いた秘書の佐藤由紀子さん、相原理佐さん、能登屋智子さん、国崎みちるさん、久門敦子さんに感謝致します。最後に私を経済面及び精神面において支えて頂いた両親に感謝致します。

参考文献

- [1] Paul Erdős and Alfréd Rényi. On random graph. i. *Pub. Mat.*, 1959.
- [2] D. Watts and S. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.
- [3] Albert-László Barabási and Réka Albert. Emergence of Scaling in Random Networks. *Science (80-.)*, 286(5439):509–512, 1999.
- [4] Wayne W. Zachary. An information flow model for conflict and fission in small groups. *Journal of Ant. Res.*, 33(4):452–473, 1977.
- [5] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*, 10(3):186–198, 2009.
- [6] Olaf Sporns. Contributions and challenges for network models in cognitive neuroscience. *Nat. Neurosci.*, 17(5):652–60, 2014.
- [7] Pall F Jonsson, Tamara Cavanna, Daniel Zicha, and Paul a Bates. Cluster analysis of networks generated through homology: automatic identification of important protein communities involved in cancer metastasis. *BMC Bioinformatics*, 7:2, 2006.
- [8] Roger Guimerà and Luís A. Nunes Amaral. Functional cartography of complex metabolic networks. *Nature*, 433(7028):895–900, 2005.
- [9] 今野紀雄 増田直紀. 複雑ネットワーク 基礎から応用まで. 近代科学社, 2010.
- [10] Santo Fortunato. Community detection in graphs. *Phys. Rep.*, 486(3-5):75–174, 2010.
- [11] Jörg Reichardt and Stefan Bornholdt. Detecting fuzzy community structures in complex networks with a potts model. *Phys. Rev. Lett.*, 93(21):19–22, 2004.
- [12] Yanqing Hu, Menghui Li, Peng Zhang, Ying Fan, and Zengru Di. Community detection by signaling on complex networks. *Phys. Rev. E*, 78(1):1–8, 2008.
- [13] Alex Arenas, Albert Díaz-Guilera, and Conrad J. Pérez-Vicente. Synchronization reveals topological scales in complex networks. *Phys. Rev. Lett.*, 96(11):1–4, 2006.
- [14] Alex Arenas and A. Díaz-Guilera. Synchronization and modularity in complex networks. *Eur. Phys. J. Spec. Top.*, 143(1):19–25, 2007.
- [15] S. Boccaletti, M. Ivanchenko, V. Latora, A. Pluchino, and A. Rapisarda. Detecting complex network modularity by dynamical clustering. *Phys. Rev. E*, 75(4):1–4, 2007.
- [16] Wills E. Lamb Jr Muraay Sargent, Marlan O. Scully. Laser physics. *Perseus Books*

Group, 1978.

- [17] Hannes Risken and Till Frank. The fokker-planck equation: Methods of solution and applications. *Springer*, 1996.
- [18] M. E J Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69(2 2):1–15, 2004.
- [19] B. Reed M. Molloy. A critical point for random graphs with a given degree sequence. *Random Struct. Algor.*, 6(2-3):161–180, 1995.
- [20] Vincent D. Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *J. Stat. Mech.*, 10008:6, 2008.
- [21] Martin Rosvall and Carl T Bergstrom. Maps of random walks on complex networks reveal community structure. *PNAS*, 105(4):1118–1123, 2008.
- [22] Martin Rosvall, D. Axelsson, and C. T. Bergstrom. The map equation. *Eur. Phys. J. Spec. Top.*, 178(1):13–23, 2009.
- [23] Leon Danon, Jordi Duch, Albert Diaz-Guilera, and Alex Arenas. Comparing community structure identification. *J. Stat. Mech.*, 09008:10, 2005.
- [24] P Zhang. A revisit to evaluating accuracy of community detection using the normalized mutual information. *arXiv Prepr.*, 501.03844, 2015.
- [25] J. Hu J. Zhang, T. Chen. On the relationship between gaussian stochastic blockmodels and label propagation algorithms. *J. stat. mech.*, 2015.
- [26] Yoshiki Kuramoto. Chemical oscillations, waves, and turbulence. *Springer*, 1984.
- [27] Yoshitaka Haribara, Shoko Utsunomiya, and Yoshihisa Yamamoto. A coherent ising machine for MAX-CUT problems: Performance evaluation against semidefinite programming and simulated annealing. *arXiv Prepr.*, 1501.07030, 2016.
- [28] R.M. Karp A. Condon. Algorithms for graph partitioning on the planted partition model. *Random Struct. Algor.*, 18(2):116–140, 2001.
- [29] M. Girvan, M. Girvan, M. E. J. Newman, and M. E. J. Newman. Community structure in social and biological networks. *Proc. Natl. Acad. Sci.*, 99(12):7821–7826, 2002.
- [30] Anne Condon and Richard M Karp. Algorithms for graph partitioning on the planted partition model. *Random Struct. Algorithms*, 18(2):116–140, 2000.
- [31] Andrea Lancichinetti and Santo Fortunato. Community detection algorithms: A comparative analysis. *Phys. Rev. E*, 80(5):1–11, 2009.
- [32] Andrea Lancichinetti, Santo Fortunato, and Filippo Radicchi. Benchmark graphs for testing community detection algorithms. *Phys. Rev. E*, 78(4):1–6, 2008.