

2020年度
修士論文

k -極大独立集合検証問題の 分散計算複雑性

名古屋工業大学大学院 情報工学専攻

所属: 片山・金研究室

2019年度入学 31414050 佐藤 僚祐

目次

第1章	はじめに	1
1.1	研究背景	1
1.2	関連研究	2
1.3	本研究の成果	3
1.4	論文の構成	3
第2章	諸定義	4
第3章	1-MIS 検証問題の1ラウンドアルゴリズム	7
第4章	k-MIS 検証問題の下限	8
4.1	2-MIS 検証問題の下限	8
4.2	3-MIS 検証問題の下限	11
4.3	k -MIS 検証問題の下限	13
第5章	まとめと今後の課題	14
5.1	まとめ	14
5.2	今後の課題	14
	謝辞	15
	参考文献	16

第1章

はじめに

1.1 研究背景

分散グラフアルゴリズムとは, 計算機を頂点, 辺を通信リンクとみなしてネットワークをモデル化したグラフ上において, そのネットワーク自身を入力として様々な問題を解く枠組みである. 分散アルゴリズムにおける代表的なモデルのひとつとして *CONGEST* モデルが存在する. *CONGEST* モデルにおいて, 各ノードは同期して同じアルゴリズムを実行して入力グラフ上の問題を解決する. 各ノードは各ラウンドで (i) b ビットのメッセージを隣接ノードに送信 (ii) 隣接ノードからメッセージを受信 (iii) 内部計算の3つの動作をする. 標準的には, $b = \log n$ を想定する. *CONGEST* モデルにおいて, ある1つのノードにグラフ全体のトポロジの情報を集め, そのノード上で逐次アルゴリズムを実行するという素朴なアプローチから自明に $O(n^2)$ ラウンドの上界を得ることができる. *CONGEST* モデルにおける下界の証明では, 下界をこの n^2 にどれだけ近づけることができるかに興味をもたれている.

ネットワーク上の最大独立集合を発見する最大独立集合問題に対する数多くの分散グラフアルゴリズムが研究されている. 大きなサイズの独立集合は経済学, 計算生物学, 符号理論, 実験計画法など様々な分野への応用に用いられる [1] が, そもそも最大独立集合問題は NP 完全であり, 頂点数 n に対して n の多項式時間で解くことは, その近似を含めて絶望的であるとされている. 内部計算に指数時間かかることを許した *CONGEST* モデルの下でラウンド数を n の多項式でおさえるような最大独立集合問題の複雑性の結果はいくつか知られている [1], [2] が, 内部計算に指数時間かかることを許容して分散グラフアルゴリズムの複雑性の議論を行うことの妥当性にはやや疑問が残る. そこで今

回,我々は最大独立集合の局所最適解である k -極大独立集合 (k -Maximal Independent Set, k -MIS) について考える. ネットワーク上の k -MIS を発見する k -MIS 問題は集中型アルゴリズムによって $O(n^{k+2})$ 時間で解くことができ, $k = O(1)$ のとき n の多項式時間になるため, 上記の疑問点を解決することができる.

今回は k -MIS の検証問題 (verification) に着目し, その複雑性について議論を行う. k -MIS 検証問題とはネットワーク上に独立集合が与えられ, それが k -MIS かどうかを判定する問題である. k -MIS 問題に対する素朴な局所探索アルゴリズムは, ある 1 つの独立集合からスタートし,

(I) 現在の状態が k -MIS であるか判定.

(II) yes であればそれを出力, no であれば解を更新

というフェーズを繰り返す. k -MIS 検証問題は上記の (I) に対応しており, k -MIS 問題と関連付いた問題ととらえることができる.

今回, 我々は極大独立集合検証問題に対するいくつかの複雑性を示した. 最初に, 1-MIS 検証問題が $O(1)$ ラウンドで解けることを証明した. 次に, 2-MIS 検証問題に対する $\tilde{\Omega}(\sqrt{n})$ ラウンドの下界と 3-MIS 検証問題に対する $\tilde{\Omega}(n)$ ラウンドの下界を証明した. 最後に, k -MIS 検証問題 ($k = 4l + 5, l \geq 1$) に対する $\tilde{\Omega}(n^{2-\frac{1}{l+1}})$ ラウンドの下界を証明した. 特に, 下界の証明のアイデアは 2 者間通信の枠組みにおける交叉判定問題からの帰着に基づいている.

1.2 関連研究

CONGEST モデルにおける最大独立集合問題の通信複雑性としては, 頂点の最大次数を Δ としたとき, 最大重み付き独立集合の $(1 + \epsilon) \cdot \Delta$ -近似を高確率で見つけるアルゴリズムに対する $(\frac{\text{poly}(\log \log n)}{\epsilon})$ ラウンドの上界 [1] や, 最大独立集合の $(\frac{1}{2} + \epsilon)$ -近似を見つけるアルゴリズムに対する $\Omega(\frac{n}{(\log n)^3})$ ラウンドの下界, $(\frac{3}{4} + \epsilon)$ -近似を見つけるアルゴリズムに対する $\Omega(\frac{n^2}{(\log n)^3})$ ラウンドの下界 [2] が知られている.

本研究の下界に関する結果は 2 者間通信の枠組みにおける交叉判定問題からの帰着に基づいているが, 交叉判定問題からの帰着によって下限を示すという証明方法は多くの問題に対して用いられている. 例えば, 最小カット発見 [3] や部分グラフ検出 [4], 近似最大クリーク検出 [5] といったさまざまな問題に対する下界の証明がされている.

1.3 本研究の成果

今回, 我々は極大独立集合検証問題に対するいくつかの複雑性を示した. 最初に, 1-MIS 検証問題が $O(1)$ ラウンドで解けることを証明した. 次に, 2-MIS 検証問題に対する $\tilde{\Omega}(\sqrt{n})$ ラウンドの下界と 3-MIS 検証問題に対する $\tilde{\Omega}(n)$ ラウンドの下界を証明した. 最後に, k -MIS 検証問題 ($k = 4l + 5, l \geq 1$) に対する $n^{2-\frac{1}{4l+1}}$ ラウンドの下界を証明した.

1.4 論文の構成

本論文は全 5 章で構成される. 第 2 章ではグラフの構造と用語の定義をしている. 第 3 章では 1-MIS 検証問題に対する $O(1)$ ラウンドアルゴリズムについて述べている. 第 4 章では k -MIS 検証問題 ($k = 2, 3, 4l + 5 (l \geq 1)$) に対する下界について述べている. 第 5 章ではまとめについて述べている.

第2章

諸定義

CONGEST モデル

本稿で考える *CONGEST* モデルは, 単純無向グラフ連結グラフ $G = (V, E)$ により表現される. ここで V はノードの集合で $|V| = n$ とし, E は通信リンクの集合である. *CONGEST* モデルでは計算機はラウンドに従って同期して動作するものとする. 1 ラウンド内で, 隣接頂点へのメッセージ送信, 隣接頂点からのメッセージ受信, 内部計算を行う. 各辺は単位ラウンドあたり b ビットを双方向に伝送可能であり, 各ノードは同一ラウンドに異なる接続辺に異なるメッセージを送信可能である. 標準的には, $b = \log n$ を想定する. また, 各ノードには $O(\log n)$ ビットの自然数値による ID が付与されており, 自身の隣接ノードすべての ID を既知であるとする. 各ノードはグラフのトポロジに関する事前知識を持たないものとする.

2 者間通信複雑性

2 者間通信複雑性の枠組みでは, アリスとボブの二人のプレイヤーがそれぞれ k ビットの $0/1$ のデータ列で構成されるプライベートな入力 x および y を持っているとする. プレイヤーの目標は, 結合関数 $f(x, y)$ を計算することであり, 複雑性の尺度として $f(x, y)$ を計算するためにアリスとボブが通信によって交換する必要があるビット数が用いられる.

この枠組みにおける重要な問題として, 交叉判定問題 (set-disjointness) がある. この問題では, アリスとボブはそれぞれ $x \in \{0, 1\}^k$ と $y \in \{0, 1\}^k$ を入力として持ち, 目的は $DISJ_k(x, y) := \bigvee_{i=1}^k x_i \wedge y_i$ を計算することである. k ビットの交叉判定問題を解くため

に, アリスとボブは通信によって $\Omega(k)$ ビット交換する必要があることが知られており [6], この事実を用いて最小カット発見 [3] や部分グラフ検出 [4], 近似最大クリーク検出 [5] といったさまざまな問題に対する下界の証明がされている.

CONGEST モデルにおいて, 入力グラフ上に特性 P があるかどうかの判定に対する下限の証明を 2 者間交叉判定問題から帰着するアプローチは以下のとおりである. 最初にアリスとボブは特殊なグラフ $G = (V, E)$ の構築と G を G_A と G_B に分割するカット辺 C の決定を行う. 次に, アリスとボブは入力文字列に基づいてそれぞれ G_A と G_B に辺 E_A と E_B を追加する. このとき, $DISJ_k(x, y) = 1$ のときのみ, 何らかの特性 P (例えば, P : 「グラフに与えられた MIS が 2-MIS でない」) を持つように辺を追加する. また, カット辺 C は入力文字列に依存しないようにする. グラフ G に辺を追加したグラフを $G' = (V', E')$ とすると $V' = V, E' = E \cap (E_A \cap E_B)$ 表すことができる. グラフ G' の構造の概要を図 2.1 に示す.

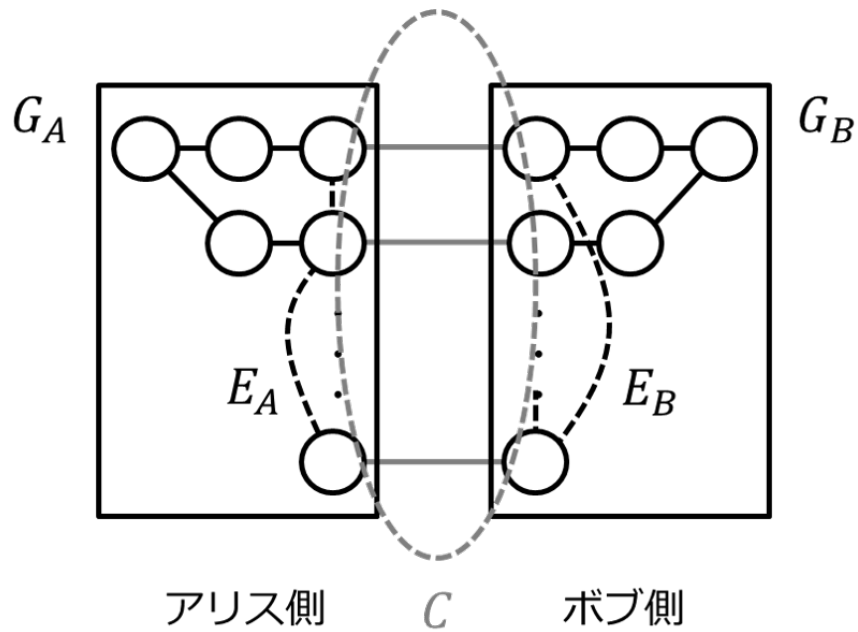


図 2.1: $G' = (V, E')$

アリスとボブは, 入力グラフ上に特性 P があるかを判定する分散アルゴリズムをシミュレートできる. 2 者間通信複雑性モデルでのシミュレートは, 次のように実行される. G_A 中の辺で送信されるメッセージ, あるいは G_B 中の辺で送信されるメッセージは, アリスとボブがそれぞれお互いと通信せずにシミュレートできる. カット辺 C を通じて送信さ

れるメッセージに対しては, お互い情報を交換する必要がある. *CONGEST* モデルにおいてグラフ上に特性 P があるかどうかを r ラウンドで判定するアルゴリズム A が存在したとすると, アリスとボブは特性 P の判定のために $O(r \cdot |C| \cdot b)$ ビット通信したことになる. これは, 各ラウンドで, アルゴリズムが各辺で $O(b)$ ビットの通信を行っているからである. このグラフにおいてアルゴリズム A を実行すると同時に 2 者間交叉判定問題も解けていることになる. 例えばアルゴリズムを実行した結果, 入力グラフに特性 P があると判定されれば $DISJ_k(x, y) = 1$ であるとわかるからである. 交叉判定問題の通信複雑性よりアリスとボブは少なくとも $\Omega(k)$ ビットは通信しているはずである. したがって, *CONGEST* モデルにおいて特性 P があるかどうかを判定する任意のアルゴリズムに対して $r = \Omega(k/|C| \cdot b)$ ラウンドの下界を得ることができる. カット辺の大きさが小さくなるほど下界が強くなる.

k -極大独立集合

定義 2.1. I が k -極大独立集合 \Leftrightarrow

$$\neg (\exists I' \subseteq I |I'| = k, \exists S \subseteq V |S| \geq k+1, (I \setminus I') \cap S \text{ が独立集合})$$

つまり, ある独立集合 I に対して, サイズ k の I の部分集合 I' を取り除いてサイズ $k+1$ 以上の V の部分集合 S をくっつけたものが新たな独立集合になり得ないとき, I を k -MIS と定義する.

第3章

1-MIS 検証問題の1ラウンドアルゴリズム

dddd

第4章

k -MIS 検証問題の下限

この章では, MIS 検証問題の下界についての議論を行う. 4.1 節では, 2-MIS 検証問題の下界についての定理とその証明を述べる. 4.2 節では, 3-MIS 検証問題の下界についての定理とその証明を述べる. 4.3 節では, k -MIS 検証問題 ($k = 4l + 5, l \geq 1$) の下界についての定理とその証明を述べる.

4.1 2-MIS 検証問題の下限

この節では 2-MIS 検証問題の下界についての議論を行う. 具体的には, 次の定理を証明する.

定理 4.1. *CONGEST* モデルにおいて, 2-MIS 検証問題を解く全てのアルゴリズムは $\tilde{\Omega}(\sqrt{n})$ の通信ラウンド数を必要とする.

証明. まず初めにアリスとボブが構築するグラフ $G' = (V', E')$ を図 4.1 に示す.

図中の頂点のうち, 灰色のものは独立集合に含まれる頂点とする. 図中で省略されている辺と $K_{W(y)}$ の構造は次のとおりである.

- $\forall i((a_i^1, b_i^1), (a_i^2, b_i^2)) \in E$
- $(a_i^1, a_j^2) \in E_A \Leftrightarrow x_{i,j} = 0$
- $W(y)$ は y 中に含まれる 1 の個数を表す. $K_{W(y)}$ の頂点 $c_{i,j}$ は $y_{i,j} = 1$ であるような (i, j) でインデックスづけされているとする. このとき, $(c_{i,j}, b_i^1) \in E_B$ かつ $(c_{i,j}, b_j^2) \in E_B$

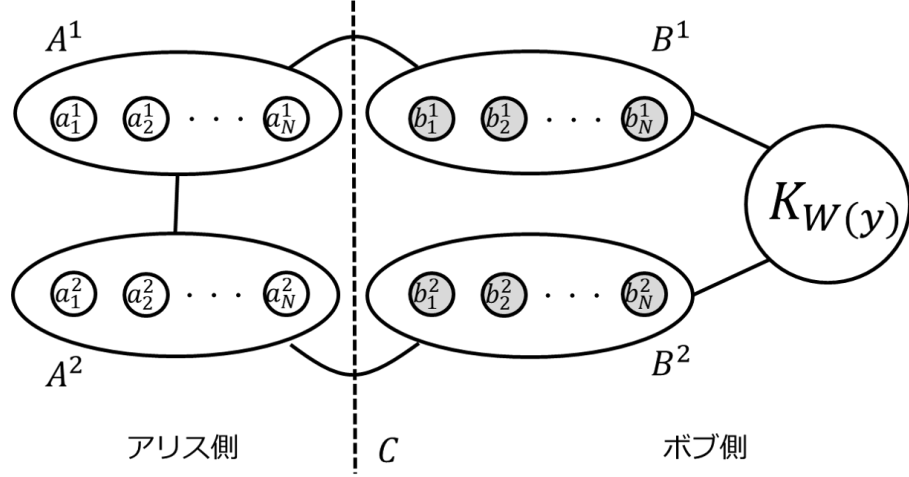


図 4.1: $G' = (V, E')$

このグラフは、「 G' 中に与えられている独立集合が, $DISJ_{N \times N}(x, y) = 1$ のときのみ 2-MIS でない」という特性を持つ. (なお, $DISJ_{N \times N}(x, y) := \bigvee_{i=1}^N \bigvee_{j=1}^N x_{i,j} \wedge y_{i,j}$ で定義される.) このグラフが上記の特性を満たしていることを示すために, 次の 2 点を確認する.

(i) $DISJ_{N \times N}(x, y) = 1$ のとき, グラフに与えられている独立集合が 2-MIS でない:

$x_{i,j} = y_{i,j} = 1$ とすると, s と c_i^1 と c_j^2 の 3 点を取り除いて a_i^1 , b_i^1 , a_j^2 , c_j^2 の 4 点を追加できることから確認できる.

(ii) $DISJ_{N \times N}(x, y) = 0$ のとき, グラフに与えられている独立集合が 2-MIS である:

グラフに与えられている独立集合が 3-MIS でないと仮定する. このとき, ある 3 点を取り除くことで独立集合に追加できる 4 点が存在するが, A^1, A^2, B^1, B^2 がそれぞれクリークであるため, 4 点を追加するためにはそれぞれから 1 点を選ぶ必要がある. c_i^1 を取り除いて a_i^1 と b_i^1 を追加し, c_j^2 を取り除いて a_j^2 と b_j^2 を独立集合に追加したとする. a_i^1 と a_i^2 が両方とも追加できるのは $x_{i,j} = 1$ のときのみであり, b_i^1 と b_i^2 が両方とも追加できるのは $y_{i,j} = 1$ のときのみであるが, これは $DISJ_{N \times N}(x, y) = 0$ に矛盾する. したがってグラフに与えられている独立集合は 3-MIS である.

今回, $N \times N$ ビットの交叉判定インスタンスをグラフに埋め込んでおり, カット辺のサイズ $|C| = 4N$ であることがわかる. したがって, $CONGEST$ モデルにおいてグラフ上

に与えられた独立集合が 3-MIS であるかどうかを r ラウンドで判定するアルゴリズム A が存在したとすると, アリスとボブは $O(r \cdot |C| \cdot b)$ ビット通信したことになる. このグラフにおいてアルゴリズム A を実行すると同時に 2 者間交叉判定問題も解けていることになるので, 交叉判定問題の通信複雑性よりアリスとボブは少なくとも $\Omega(N \times N)$ ビットは通信しているはずである. よって, $r = \Omega(N/4b) = \tilde{\Omega}(N)$ ラウンドの下界を得ることができる. 図??からわかる通り, $A^1, A^2, B^1, B^2, C^1, C^2$ は N 頂点で構成されており, グラフ全体の頂点数 n は $n = 6N + 1$ である. したがって $N = \Omega(n)$ になるため, $\tilde{\Omega}(n)$ ラウンドの下界を得ることができる.

4.2 3-MIS 検証問題の下限

この節では 3-MIS 検証問題の下界についての議論を行う. 具体的には, 次の定理を証明する.

定理 4.2. *CONGEST* モデルにおいて, 3-MIS 検証問題を解く全てのアルゴリズムは $\tilde{\Omega}(n)$ の通信ラウンド数を必要とする.

証明. まず初めにアリスとボブが構築するグラフ $G' = (V', E')$ を図 4.2 に示す.

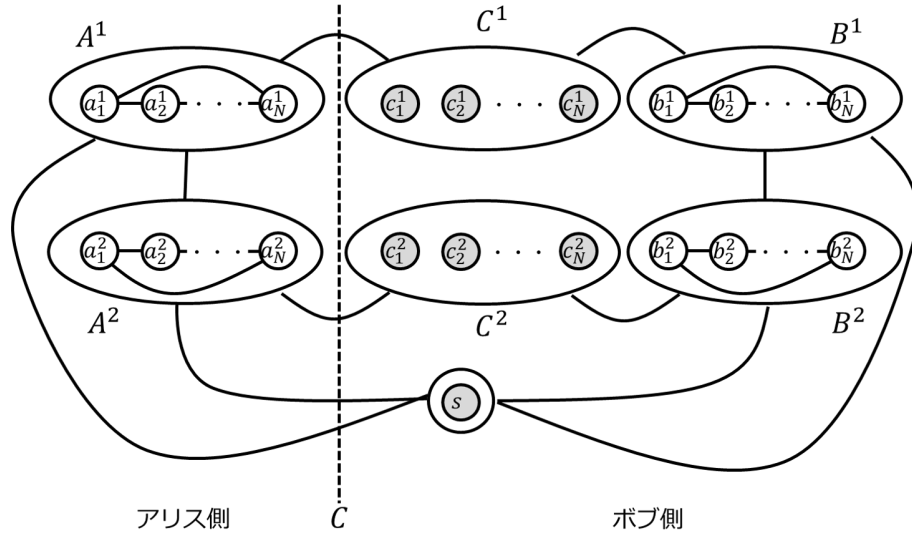


図 4.2: $G' = (V, E')$

図中の頂点のうち, 灰色のものは独立集合に含まれる頂点とする. 図中で省略されている辺は次のとおりである.

- A^1, A^2, B^1, B^2 は N 頂点のクリーク K_N
- $\forall i((a_i^1, c_i^1), (a_i^2, c_i^2), (b_i^1, c_i^1), (b_i^2, c_i^2)) \in E$
- $\forall i((a_i^1, s), (a_i^2, s), (b_i^1, s), (b_i^2, s)) \in E$
- $(a_i^1, a_j^2) \in E_A \Leftrightarrow x_{i,j} = 0, (b_i^1, b_j^2) \in E_B \Leftrightarrow y_{i,j} = 0$

このグラフは、「 G' 中に与えられている独立集合が、 $DISJ_{N \times N}(x, y) = 1$ のときのみ 3-MIS でない」という特性を持つ。このグラフが上記の特性を満たしていることを示すために、次の 2 点を確認する。

(i) $DISJ_{N \times N}(x, y) = 1$ のとき、グラフに与えられている独立集合が 3-MIS でない:

$x_{i,j} = y_{i,j} = 1$ とすると、 s と c_i^1 と c_j^2 の 3 点を取り除いて $a_i^1, b_i^1, a_j^2, c_j^2$ の 4 点を追加できることから確認できる。

(ii) $DISJ_{N \times N}(x, y) = 0$ のとき、グラフに与えられている独立集合が 3-MIS である:

グラフに与えられている独立集合が 3-MIS でないと仮定する。このとき、ある 3 点を取り除くことで独立集合に追加できる 4 点が存在するが、 A^1, A^2, B^1, B^2 がそれぞれクリークであるため、4 点を追加するためにはそれぞれから 1 点を選ぶ必要がある。 c_i^1 を取り除いて a_i^1 と b_i^1 を追加し、 c_j^2 を取り除いて a_j^2 と b_j^2 を独立集合に追加したとする。 a_i^1 と a_j^2 が両方とも追加できるのは $x_{i,j} = 1$ のときのみであり、 b_i^1 と b_j^2 が両方とも追加できるのは $y_{i,j} = 1$ のときのみであるが、これは $DISJ_{N \times N}(x, y) = 0$ に矛盾する。したがってグラフに与えられている独立集合は 3-MIS である。

今回、 $N \times N$ ビットの交叉判定インスタンスをグラフに埋め込んでおり、カット辺のサイズ $|C| = 4N$ であることがわかる。したがって、*CONGEST* モデルにおいてグラフ上に与えられた独立集合が 3-MIS であるかどうかを r ラウンドで判定するアルゴリズム \mathcal{A} が存在したとすると、アリスとボブは $O(r \cdot |C| \cdot b)$ ビット通信したことになる。このグラフにおいてアルゴリズム \mathcal{A} を実行すると同時に 2 者間交叉判定問題も解けていることになるので、交叉判定問題の通信複雑性よりアリスとボブは少なくとも $\Omega(N \times N)$ ビットは通信しているはずである。よって、 $r = \Omega(N/4b) = \tilde{\Omega}(N)$ ラウンドの下界を得ることができる。図??からわかる通り、 $A^1, A^2, B^1, B^2, C^1, C^2$ は N 頂点で構成されており、グラフ全体の頂点数 n は $n = 6N + 1$ である。したがって $N = \Omega(n)$ になるため、 $\tilde{\Omega}(n)$ ラウンドの下界を得ることができる。

4.3 k -MIS 検証問題の下限

このセクションでは 3-MIS 検証問題の下界についての議論を行う. 具体的には, 次の定理を証明する.

第5章

まとめと今後の課題

5.1 まとめ

5.2 今後の課題

4.3 節で一般の k に対する k -MIS 検証問題の下界を証明したが, $k = 4, \dots, 8$ については現在 3-MIS 検証問題と同じ下界しか得られていない. この下界をよりタイトにできるかが今後の課題である.

謝辞

本研究の機会を与え、数々の御指導を賜りました泉泰介准教授に深く感謝致します。
また、本研究を進めるにあたり多くの助言を頂き、様々な御協力を頂きました泉研究室
の学生みなさんに深く感謝致します。

参考文献

- [1] Ken-ichi Kawarabayashi, Seri Khoury, Aaron Schild, and Gregory Schwartzman. Improved distributed approximation to maximum independent set. *arXiv preprint arXiv:1906.11524*, 2019.
- [2] Yuval Efron, Ofer Grossman, and Seri Khoury. Beyond alice and bob: Improved inapproximability for maximum independent set in congest. *arXiv preprint arXiv:2003.07427*, 2020.
- [3] Mohsen Ghaffari and Fabian Kuhn. Distributed minimum cut approximation. In *International Symposium on Distributed Computing*, pages 1–15. Springer, 2013.
- [4] Orr Fischer, Tzlil Gonen, Fabian Kuhn, and Rotem Oshman. Possibilities and impossibilities for distributed subgraph detection. In *Proceedings of the 30th on Symposium on Parallelism in Algorithms and Architectures*, pages 153–162, 2018.
- [5] Artur Czumaj and Christian Konrad. Detecting cliques in congest networks. *Distributed Computing*, 33(6):533–543, 2020.
- [6] Bala Kalyanasundaram and Georg Schintger. The probabilistic communication complexity of set intersection. *SIAM Journal on Discrete Mathematics*, 5(4):545–557, 1992.