INTERVAL DIGRAPHS

by

Elizabeth Honda Reiland

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

May, 2019

© Elizabeth Honda Reiland 2019

All rights reserved
Abstract

We say that a simple, undirected graph $G = (V, E)$ is an interval graph if there exists a function $f : V \rightarrow \mathcal{I}$ where $\mathcal{I}$ is the set of compact intervals in $[0, 1]$ (or $\mathbb{R}$) such that for all $u, v \in V$, $u$ is adjacent to $v$ if and only if $f(u) \cap f(v) \neq \emptyset$ (i.e. the corresponding intervals have a non-empty intersection). Undirected interval graphs were originally of interest for applications in biology, and the class of graphs has been thoroughly studied. In 1989, Das, Sen, Roy, and West introduced directed interval graphs or interval digraphs as a natural extension of the undirected interval graph. In this dissertation, we introduce a second type of interval digraphs and demonstrate that there exist separating examples for the two classes of graphs. We study properties of and present a recognition algorithm for the newly defined class, as well as comparing and contrasting the random properties of both models.

Primary Reader: Edward Scheinerman

Secondary Reader: Amitabh Basu
Acknowledgements

First and foremost, I would like to thank my advisor, Dr. Ed Scheinerman, for all his help over the course of my six years as a PhD student. Your insightful advice and guidance through the research process was invaluable, and your patience, support, and understanding meant so much when I was frustrated and struggling (which was more often than I would like to admit). This dissertation would not be anything close to what it is without your help.

I would also like to thank Dr. Amitabh Basu, my second reader, for being another pillar in my PhD experience. Your passion for mathematics is inspiring, and I feel very lucky I got to work with you as both a student and teaching assistant.

I am also incredibly grateful to Dr. Donniel Fishkind, Dr. John Wierman, and Dr. Edinah Gnang not only for serving on my committee, but for teaching me so much throughout my PhD studies.

Thank you to Dr. Art Benjamin, my undergraduate advisor, who has continued to consistently encourage and support me. When I teach, I always strive to make math as magical as you do.
To Zach and Joe, thank you for being not just officemates and peers, but amazing friends. We have shared good food, cool theorems, fun games, failure, success, and so much more over the years. You are both incredible, and I know you will succeed at whatever you put your mind to.

To Katia, thanks for putting up with all of us math people.

To Jenny, Sarah, Tim, Emily, and Kathy, thank you for your jokes, your commiseration, your words of wisdom, and most of all your friendship.

To Halorin, chesslooter, Dr. Toe, and DargoSun, thank you for cheering me on and for reminding me that there is more to life than just work. You are truly heroes of the storm.

To Carolyn, thank you for helping keep me sane.

To Mom, Dad, Mikeie, and Arcelia, your love and acceptance have been a rock for me. Thank you for always being there.

To Tristan, I cannot put into words everything you have done for me and everything you mean to me, but I hope you know.

Lastly, to Entropy, thank you for being the best little companion anyone could ask for.
## Contents

List of Figures viii

<table>
<thead>
<tr>
<th>1 Introduction</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Background</td>
<td>1</td>
</tr>
<tr>
<td>1.2 New Models</td>
<td>9</td>
</tr>
<tr>
<td>1.2.1 Type I Interval Digraphs</td>
<td>9</td>
</tr>
<tr>
<td>1.2.2 Terminology, Notation, and Assumptions</td>
<td>12</td>
</tr>
</tbody>
</table>

| 2 Differentiating Type I and Type II Interval Digraphs | 15 |

| 3 Recognizing Type I Interval Digraphs | 22 |
| 3.1 Special Subclasses of Type I Interval Digraphs | 23 |
| 3.2 Type Ib Interval Digraphs | 28 |
| 3.2.1 Recognition Algorithm | 32 |
| 3.2.2 Examples | 36 |
| 3.2.3 Proof of Correctness | 38 |
CONTENTS

3.2.3.1 Forbidden Subgraphs/Patterns . . . . . . . . . . . . 42
3.2.3.2 Conflicting Patterns . . . . . . . . . . . . . . . . . . . 43
3.2.4 Runtime Analysis . . . . . . . . . . . . . . . . . . . . . . . . . 49
3.2.5 Relation to Type I Interval Digraphs . . . . . . . . . . . . . . 50

4 Properties of Random Interval Digraphs 55
4.1 Edges in Random Interval Digraphs . . . . . . . . . . . . . . . . . . 55
4.1.1 Type I Interval Digraphs . . . . . . . . . . . . . . . . . . . . . . 56
4.1.2 Type II Interval Digraphs . . . . . . . . . . . . . . . . . . . . . 58
4.2 Degree in Random Interval Digraphs . . . . . . . . . . . . . . . . . . 60
4.2.1 Type I Interval Digraphs . . . . . . . . . . . . . . . . . . . . . . 60
4.2.2 Type II Interval Digraphs . . . . . . . . . . . . . . . . . . . . . 69
4.3 Diameter of Random Interval Digraphs . . . . . . . . . . . . . . . . 76
4.3.1 Type I Interval Digraphs . . . . . . . . . . . . . . . . . . . . . . 76
4.3.2 Type II Interval Digraph . . . . . . . . . . . . . . . . . . . . . . 90
4.3.3 Comparison with Erdős-Renyi random digraphs . . . . . . . 92
4.4 Cliques and Independent Sets in Random Interval Digraphs . . . . . 92
4.4.1 Type I Interval Digraphs . . . . . . . . . . . . . . . . . . . . . . 93
4.4.2 Type II Interval Digraphs . . . . . . . . . . . . . . . . . . . . . 98
4.4.3 Comparison with Erdős-Renyi random digraphs . . . . . . . . 103

5 Future Work 105
## CONTENTS

5.1 Type I and Type II Interval Digraphs ........................................... 105

5.2 Random Undirected Unit Interval Graphs ................................. 108

5.3 Random Random Interval Graphs ............................................ 110

6 Appendix: Proofs for Forbidden Subgraphs .................................. 114

6.1 Type I Interval Digraphs ...................................................... 114

6.2 Type Ib Interval Digraphs .................................................... 117

Bibliography .................................................................................. 120
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Four combinatorially distinct representations for an edge $uv$</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Interval overlaps and corresponding edges for a type I interval digraph</td>
<td>10</td>
</tr>
<tr>
<td>2.1</td>
<td>A type II interval digraph (with both its looped and non-looped interval representations) that is not a type I interval digraph</td>
<td>16</td>
</tr>
<tr>
<td>2.2</td>
<td>A type I interval digraph (with a corresponding interval representation) that is not a type II interval digraph</td>
<td>17</td>
</tr>
<tr>
<td>2.3</td>
<td>A type I interval digraph (with a corresponding interval representation) that is not a type II interval digraph</td>
<td>18</td>
</tr>
<tr>
<td>2.4</td>
<td>A fully looped orientation of $C_4$ along with its corresponding type II interval representation</td>
<td>21</td>
</tr>
<tr>
<td>3.1</td>
<td>Known minimal forbidden induced subdigraphs for type I interval digraphs</td>
<td>23</td>
</tr>
<tr>
<td>3.2</td>
<td>Interval configurations creating dashed and solid edges</td>
<td>28</td>
</tr>
<tr>
<td>3.3</td>
<td>A list of forbidden subgraphs for type I$\beta$ interval digraphs relevant to the algorithm</td>
<td>31</td>
</tr>
<tr>
<td>3.4</td>
<td>Interval configurations that must be rearranged in order to interact properly with $\ell$. Throughout our algorithm we use $o$ to indicate an interval that should overlap $I_\ell$, $c$ to indicate an interval that should be contained in $I_\ell$, and $x$ to indicate an interval that should be disjoint from $I_\ell$. We call these conflicts type (i), (ii), and (iii) respectively based on the order we address them in the algorithm</td>
<td>33</td>
</tr>
<tr>
<td>3.5</td>
<td>Interval configurations that result in incorrect edges formed relative to $\ell$. $o$ indicates an interval that should overlap $I_\ell$, $c$ indicates an interval that should be contained in $I_\ell$ and $x$ indicates an interval that should be disjoint from $I_\ell$.</td>
<td>40</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

3.6 Interval configurations that imply forbidden induced subgraphs in combination with \( \ell \). Here \( c \in N_{H}^{++}(\ell), \alpha, \alpha' \in N_{H}^{+}(\ell), \) and \( x \in N_{H}^{0}(\ell) \). ........................................... 43

4.1 Approximate picture of the intervals corresponding to the special vertices. .................................................... 88

4.2 Directed \((u, w)\)-paths in a random type I interval digraph. ............... 89
Chapter 1

Introduction

1.1 Background

A simple, undirected graph $G$ is an ordered pair $(V, E)$ consisting of a set of vertices $V$ and a set of edges $E$ which consist of 2-element subsets of $V$. Note that edges in a simple graph must be between two distinct vertices (i.e. there are no loop edges with both endpoints being the same vertex), and for vertices $u, v \in V$, the edge $uv$ is the same as the edge $vu$. There is at most one edge between any pair of vertices $u, v \in V$.

A directed graph or digraph $G$ is again ordered pair $(V, E)$ consisting of a set of vertices $V$ and a set of edges $E$, but in this case an edge is a 2-tuple (giving the edge a direction). This implies that the edge $uv$ is distinct from the edge $vu$, and neither, only $uv$, only $vu$, or both $uv$ and $vu$ may appear in a given digraph. We call $uv \in E$
A **single edge** if $vu \notin E$, and we refer to the antiparallel edges $uv, vu \in E$ as a **double edge**. Note that while there may multiple edges between a pair of vertices if those edges are antiparallel, we still do not allow multiple parallel edges between a pair of vertices $u, v \in V$ (i.e. there is at most one edge $uv$ and at most one edge $vu$).

Unlike simple, directed graphs, we allow loops in directed graphs.

We say that a simple, undirected graph $G = (V, E)$ is an **interval graph** if there exists a function $f : V \to I$ where $I$ is the set of compact intervals in $[0, 1]$ (or $\mathbb{R}$) such that for all $u, v \in V$, $u$ is adjacent to $v$ if and only if $f(u) \cap f(v) \neq \emptyset$ (i.e. the corresponding intervals have a non-empty intersection). We call such a function an **interval representation** for $G$. Observe that a graph $G$ may have combinatorially distinct interval representations; for example, the graph $K_2$ has the following possible interval representations (see Figure 1.1).

Interval graphs originally drew researchers’ interest in the 1960’s, motivated in part by the geneticist Benzer’s 1959 paper [1] studying the topology of the internal structure of genes. It was known that genes were arranged linearly within a chro-
mosome, so Benzer posited that this might be the case for sub-elements within a gene as well. In the course of his research, Benzer examines whether virus cells with various mutations could recombine to form a non-mutated version. He discovered that it was possible to model the mutants as intervals such that two mutations could successfully recombine if and only if the corresponding intervals did not intersect.

Through the 1960s, several different characterizations for interval graphs were proven. In particular, each of the following statements are equivalent to a graph being an interval graph:

- $G$ is triangulated and $G$ contains no astroidal triples (i.e. any three vertices of $G$ can be ordered such that every path from the first vertex to the third vertex must pass through a neighbor of the second vertex). (Lekkerkerker and Boland 1962 [2])

- $G$ is an interval graph if and only if $G$ contains no chordless 4-cycle and its complement $\bar{G}$ is a comparability graph (i.e. $G$ has a transitive orientation). (Gilmore and Hoffman 1964 [3])

- The maximal cliques of $G$ can be linearly ordered such that for all $v \in V$, the maximal cliques containing $v$ appear consecutively. (Gilmore and Hoffman 1964 [3])

- The clique matrix of $G$ (maximal cliques versus vertices) has the consecutive
CHAPTER 1. INTRODUCTION

1’s property for columns (i.e. the rows can be permuted such the 1’s in each column occur consecutively). (Fulkerson and Gross 1965 [4])

Lekkerkerker and Boland also include a forbidden subgraph characterization for interval graphs in [2]. Booth and Lueker later used Fulkerson and Gross’s consecutive 1’s characterization of the clique matrix to formulate an algorithm that recognizes interval graphs in linear time, presented in their 1976 paper [5]. Golumbic provides a comprehensive survey of many of these results on interval graphs and more in [6].

Two particularly noteworthy subclasses of interval graphs are unit interval graphs and proper interval graphs. We say a graph $G$ is a unit interval graph if there exists an interval representation for $G$ such that all intervals are the same length. We say a graph $G$ is a proper interval graph if there exists an interval representation for $G$ such that no interval properly contains any other interval. It is clear that the unit interval graphs must be a subset of the proper interval graphs since proper containment can only occur if a representation has intervals of varying length. In his 1969 paper [7], Roberts proved that in fact the two classes are the same, that is a graph is a unit interval graph if and only if it is a proper interval graph. He additionally showed that these two classes of graphs are exactly the interval graphs that contain no induced copy of $K_{1,3}$.

In 1979, Trotter and Harary introduced the concept of the interval number of a graph as a generalization of interval graphs (see [8]). The interval number of a
CHAPTER 1. INTRODUCTION

Graph $G$ is the minimum number $t$ for which there exists a function $f : V \rightarrow 2^I$ which assigns each vertex of $G$ a subset of $I$ that is the union of $t$ intervals (not necessarily disjoint) such that for all $u, v \in V$, we have $uv \in E$ if and only if $f(u) \cap f(v) \neq \emptyset$. Thus interval graphs are exactly those graphs with interval number 1. Additionally, for any graph $G$ on $n$ vertices, the interval number of $G$ is at most $n - 1$, so every graph has a finite interval number.

There is an extensive body of work looking at the interval numbers of graphs and recognizing graphs by their interval numbers. In their original paper introducing interval numbers, Trotter and Harary proved that the interval number of a tree is at most two and that the complete bipartite graph $K_{m,n}$ has interval number exactly $\lfloor (mn + 1) / (m + n) \rfloor$. In 1983, Scheinerman and West proved in their paper that the interval number of any planar graph is at most three. In 1984, West and Shmoys proved that while interval graphs can by recognized in linear time, for any value $t \geq 2$ determining whether the interval number of a graph $G$ is at most $t$ is NP-complete (see [10]). In a 1985 paper [11], Erdős and West showed that almost every graph on $n$ vertices has an interval number between $\frac{n}{4 \log n}$ and $\frac{n}{4}$.

Interest also began to surface in studying random sets of intervals and random interval graphs. In 1988, Scheinerman [12] introduced two equivalent models for random interval graphs. For the first model, we consider $2n$ random variables $X_1, Y_1, \ldots, X_n, Y_n$ which are independent and chosen uniformly from the unit interval $[0, 1]$. We then build an interval graph, taking $[X_i, Y_i]$ to be the interval asso-
CHAPTER 1. INTRODUCTION

ciated with the vertex \( v_i \). Note that we consider \([X_i, Y_i] = [Y_i, X_i]\) if \( X_i > Y_i \).

For the second model, we select a permutation \( x_1, y_1, \ldots, x_n, y_n \) of \([2n]\) uniformly at random. We take the resulting arrangement to be the order in which the endpoints of the intervals appear where \([x_i, y_i]\) is the interval associated with vertex \( v_i \).

Conveniently, these two models are equivalent, that is the probability of obtaining a particular graph under either model is equal. Observe that there is implicitly an assumption in the second model that the endpoints of the intervals are distinct, and in the first model the independently chosen values are distinct with probability 1. Additionally, each of the possible orderings of the endpoints of the intervals is equally likely in the first model, which is precisely the case in the second model (in fact, Scheinerman demonstrated that this is true for any continuous distribution, not just the uniform distribution, though this will not be relevant to our work here).

Note that while these models yield graphs that are uniformly distributed over all possible interval representations, they are not uniformly distributed over all possible interval graphs. For example, when \( n = 2 \), we see that the probability of a graph being \( K_2 \) is \( \frac{2}{3} \), whereas the probability of obtaining \( \overline{K_2} \) is \( \frac{1}{3} \). Note that this also implies that the probability of any particular edge existing in a random interval graph is \( \frac{2}{3} \).

Using these models, many interesting properties of random interval graphs
have been found. In [12], Scheinerman derived results concerning the number of edges and degrees in interval graphs as well as showing that with high probability a random interval graph is Hamiltonian. In [13], Justicz, Scheinerman, and Winkler gave a combinatorial proof of the result that given a collection of $n$ random intervals, the probability that some interval intersects all other intervals is exactly $\frac{2}{3}$ (in terms of random interval graphs, this implies the probability there is a vertex of degree $n-1$ is exactly $\frac{2}{3}$). This result is particularly interesting as it runs to counter to typical Erdős-Rényi random graph properties, which typically have probabilities tending to 0 or 1 (see Shelah and Spencer’s work in [14]).

Looking to expand on the idea of interval graphs, in 1989 Das, Sen, Roy, and West introduced the concept of directed interval graphs in [15]. Their generalization uses two intervals for each vertex in order to encode direction; essentially one is a send interval and the other is a receive interval. Formally, we say a directed graph $G = (V, E)$ is an interval digraph if and only if there exists an interval representation $f : V \to I, g : V \to I$ such that for all $u, v \in V$, there exists an edge from $u$ to $v$ in $G$ if and only if $f(u) \cap g(v) \neq \emptyset$.

In addition to introducing the idea of directed interval graphs, Das, Sen, Roy, and West gave a characterization of interval digraphs based on their adjacency matrices. Defining interval-point digraphs to be interval digraphs such that the receive intervals consist of single points, they prove that a digraph $D$ is an interval-point digraph if and only if its adjacency matrix has the consecutive 1’s property.
for rows. This parallels the result for clique matrices in the undirected case. They also prove a similar albeit more complicated result for interval digraphs in general: a digraph $D$ is an interval digraph if and only if the rows and columns of $D$’s adjacency matrix can be independently permuted in a way that each 0 can be replaced by one of $\{R, C\}$ such that every $R$ has only $R$’s to its right and every $C$ has only $C$’s below it.

In 1997 [16], Müller presented a dynamic programming algorithm for recognizing interval digraphs in polynomial time. His algorithm built on the idea of interpreting interval digraphs as “interval bigraphs,” bipartite graphs that encode the direction of edges by having one vertex in each partite set representing a particular vertex of the digraph.

While we have a matrix description of interval digraphs and an efficient recognition algorithm, there is not yet a complete forbidden subdigraph characterization. In 2007 [17], Brown, Busch, and Lundgren looked specifically at which tournaments are interval digraphs and obtained a characterization of this special case by forbidden subtournaments. Even more recently in 2016 [18], Das, Das, and Sen identified new forbidden substructures for interval digraphs. They also, however, exhibited an example that demonstrates the known list of forbidden substructures is still not complete.

Special classes of interval digraphs have also been studied in the form of indifference digraphs. In [19] Sen and Sanyal demonstrated that indifference digraphs
CHAPTER 1. INTRODUCTION

were the same class of digraphs as unit interval digraphs and proper interval digraphs, implying that as in the undirected case, unit and proper mean the same thing for interval digraphs.

In this dissertation, we build on this previous work to examine the properties of random interval digraphs. Additionally, we introduce an alternative definition of interval digraphs and contrast this interpretation with the version presented by Das, Sen, Roy, and West.

1.2 New Models

In addition to the model for interval digraphs of Das, Sen, Roy, and West, we introduce a second variety of interval digraphs.

1.2.1 Type I Interval Digraphs

We define type I interval digraphs as follows: a graph $G = (V, E)$ is a type I interval digraph if there exists an interval representation $f : V \rightarrow \mathcal{I}$ such that for all $u, v \in V$, there exists an edge from $u$ to $v$ in $G$ if and only if $f(u) \subseteq f(v)$ or $f(u) \cap f(v) \neq \emptyset$ and the left endpoint of $f(u)$ is less than the left endpoint of $f(v)$. In particular if $f(u) = [x_1, y_1]$ and $f(v) = [x_2, y_2]$ (with $x_1 < y_1$, $x_2 < y_2$), then there is an edge from $u$ to $v$ if the order of the endpoints is $x_1 < x_2 < y_2 < y_1$ or $x_2 < x_1 < y_1 < y_2$ or $x_1 < x_2 < y_1 < y_2$. A visual of this is given in Figure 1.2.
CHAPTER 1. INTRODUCTION

Figure 1.2: Interval overlaps and corresponding edges for a type I interval digraph

For the sake of clarity, we refer to Das, Sen, Roy, and West’s model as type II interval digraphs for the remainder of this dissertation. Although their definition of interval digraphs was presented first, we select this naming convention for convenience: type I interval digraphs have interval representations that associate one interval with each vertex, while type II interval digraphs have representations that associate two intervals with each vertex. It is interesting to note that these two notions of interval digraphs are not equivalent (see Chapter 2 for further details).

Note that we generally assume that type I interval digraphs are loopless. This decision is somewhat arbitrary; we could also choose to have a loop at each vertex. For type II interval digraphs, on the other hand, the presence and absence of loops matters for each vertex and is determined by the specific interval representation. In particular for a vertex $v$, there will be a loop at $v$ if and only if the send interval and receive interval for $v$ in the type II interval representation have a nonempty intersection.
CHAPTER 1. INTRODUCTION

Since Scheinerman’s models for creating random interval graphs center around selecting random intervals and building the associated graph, we use the same methods to generate random interval digraphs of either type. In the case of type II interval digraphs, we simply generate $2n$ intervals (a send and receive interval for each vertex) rather than $n$.

We also implement a third model that allows us to take advantage of both the independence of the first model and discreteness of the second. Consider $2n$ random variables $X_1, Y_1, \ldots, X_n, Y_n$ which are independent and chosen uniformly at random from the set \{0, $\frac{1}{n^3}$, $\frac{2}{n^3}$, $\ldots$, $\frac{n^3 - 1}{n^3}$, 1\}. We take $[X_i, Y_i]$ to be the interval associated with the vertex $v_i$ and build the corresponding interval graph from the resulting representation.

Note that for small values of $n$, this model is not necessarily equivalent to the previous two models, as there will be a non-trivial possibility of a particular value from \{0, $\frac{1}{n^3}$, $\ldots$, 1\} being chosen more than once as an endpoint. If two endpoints are assigned the same value, we cannot uniquely translate the result into a permutation of the endpoints. For large values of $n$, however, we argue that this model behaves the same as the others. We say a graph property holds \textit{with high probability} if the probability that the property occurs goes to 1 as the number of vertices $n$ goes to infinity. In this case, we show that with high probability no pair of interval endpoints chosen by the model have the same value:

$$\lim_{n \to \infty} P(\text{no two endpoints have the same value}) = \lim_{n \to \infty} \frac{(n^3)!}{(n^3-2n)! \cdot (n^3)^2n} \to 1.$$
CHAPTER 1. INTRODUCTION

Thus, asymptotically this model behaves identically to the other two.

In this dissertation, we study type I interval digraphs and their relationship to type II interval digraphs. In particular we give differentiating examples between type I and type II interval digraphs, and we introduce a variation on type I interval digraphs to obtain a recognition algorithm. We also investigate random properties of both varieties of interval digraph, including degree distribution, diameter, clique number, and independence number.

1.2.2 Terminology, Notation, and Assumptions

For convenience, we introduce some vocabulary and notation that allows us to more concisely and precisely refer to elements of interval digraphs and their corresponding interval representations.

Consider two intervals $I, J \in \mathcal{I}$. We say $I$ intersects $J$ if $I \cap J \neq \emptyset$. We say $I$ contains $J$ if $J \subseteq I$, and we say $I$ overlaps $J$ if $I \cap J \neq \emptyset$ but $I$ is not contained in $J$ and $J$ is not contained in $I$.

Given $G = (V, E)$, a type I interval digraph, $u, v \in V$, and an interval representation $f : V \to \mathcal{I}$:

- We denote the interval corresponding to $v$ by $I_v$ (i.e. $f(v) = I_v$).

- We denote the left endpoint of $I_v$ by $L(I_v)$ and the right endpoint of $I_v$ by $R(I_v)$.
CHAPTER 1. INTRODUCTION

- We write $I_u < I_v$ if $I_u$ overlaps $I_v$ and $L(I_u) < L(I_v)$. Note that this implies there is an edge from $u$ to $v$.

- We write $I_u \ll I_v$ if $R(I_u) < L(I_v)$ (i.e. $I_u$ is completely to the left of $I_v$). Note that this implies $u$ and $v$ are not adjacent.

- We denote the vertex corresponding to the interval in $f$ with the leftmost endpoint by $v_\ell$. Similarly, we denote the vertex corresponding to the interval in $f$ with the rightmost endpoint by $v_r$. Note that it is possible that $v_\ell = v_r$ if we have an interval that contains all other intervals.

As a final introductory note, consider the following lemma:

**Lemma 1.1.** A set of intervals $S = \{I_1, I_2, \ldots, I_n\}$ can be transformed into another set $S' = \{I'_1, I'_2, \ldots, I'_n\}$ satisfying the following properties:

1. All interval endpoints are distinct.

2. For all $1 \leq j, k \leq n$, $I_j \ll I_k$ if and only if $I'_j \ll I'_k$.

3. For all $1 \leq j, k \leq n$, $I_j \subseteq I_k$ if and only if $I'_j \subseteq I'_k$.

4. For all $1 \leq j, k \leq n$, $I_j < I_k$ if and only if $I'_j < I'_k$.

The proof of this lemma can be found in Golumbic and Trenk’s book [20]. Note that Golumbic and Trenk’s version of the lemma does not explicitly specify the final condition ($I_j < I_k$ if and only if $I'_j < I'_k$), but it is a direct implication of their
CHAPTER 1. INTRODUCTION

construction. Since any interval representation can be adjusted to one with distinct interval endpoints while preserving interactions between all pairs of intervals, any interval digraph (of either type) has a corresponding interval representation such that all interval endpoints are distinct. Thus for convenience, we assume throughout this dissertation that all endpoints in interval representations are distinct.
Chapter 2

Differentiating Type I and Type II Interval Digraphs

Before delving into the properties of interval digraphs of either type, we first establish that type I and type II interval digraphs are distinct classes of digraphs. Due to the assumption that type I interval digraphs either have loops at every vertex or are completely loopless, any type II interval digraph with loops at some but not all vertices will not be a type I interval digraph. However, even restricting to type II interval digraphs with loops at every vertex (or loops at none of the vertices) we exhibit type II interval digraphs that are not type I interval digraphs.

In Figure 2.1 we provide two interval representations for the pictured digraph: the top representation assumes loops at each vertex (not pictured); the bottom representation is for the pictured digraph as is (i.e. with no loops). Though this is a
CHAPTER 2. DIFFERENTIATING TYPE I AND TYPE II INTERVAL DIGRAPHS

Figure 2.1: A type II interval digraph (with both its looped and non-looped interval representations) that is not a type I interval digraph

type II interval digraph, we prove there exists no type I interval representation for this digraph.

Observe first that in order for $uv$ to exist without $vu$, we must have $I_u < I_v$. Similarly, since we have $uw$ but not $wu$, it must be the case that $I_u < I_w$. This implies that the right endpoint of $I_u$ must be contained in both $I_v$ and $I_w$. However, this implies $I_v \cap I_w \neq \emptyset$, so either $vw, wv$, or both would have to appear in the digraph. Thus the digraph pictured in Figure 2.1 has a type II representation but no type I representation.

Next, consider the type I interval digraph in Figure 2.2 with its corresponding interval representation (note that there are no loops). We show there is no type II interval representation for this digraph. Observe that each of the following must hold:

- $S_u \cap R_v \neq \emptyset$ and $S_u \cap R_w \neq \emptyset$, but $S_u \cap R_u = \emptyset$
CHAPTER 2. DIFFERENTIATING TYPE I AND TYPE II INTERVAL DIGRAPHS

Figure 2.2: A type I interval digraph (with a corresponding interval representation) that is not a type II interval digraph

- $S_v \cap R_u \neq \emptyset$ and $S_v \cap R_w \neq \emptyset$, but $S_v \cap R_v = \emptyset$
- $S_w \cap R_u \neq \emptyset$ and $S_w \cap R_v \neq \emptyset$, but $S_w \cap R_w = \emptyset$

Observe that no pair of the receive intervals can be nested. For example, if $R_u \subseteq R_v$, since $S_v \cap R_u \neq \emptyset$, this would imply $S_v \cap R_v \neq \emptyset$, a contradiction. Suppose then without loss of generality that the receive intervals are $R_u = [a_1, b_1], R_v = [a_2, b_2], R_w = [a_3, b_3]$ with $a_1 < a_2 < a_3$ and $b_1 < b_2 < b_3$.

There are two cases to consider: if $R_u \ll R_w$, then observe that either $b_1 \in R_v, a_3 \in R_v$, or $R_v \subseteq [b_1, a_3]$. Since we must have $b_1 \in S_v$ and $a_3 \in S_v$, this implies $S_v \cap R_v \neq \emptyset$, a contradiction. On the other hand, if $R_u < R_w$, then $a_2 < a_3 < b_1 < b_2$ and thus $R_u \cap R_w \subseteq R_v$. Since some point of $R_u \cap R_w$ must be contained in $S_v$, this again implies $S_v \cap R_v \neq \emptyset$, a contradiction.

Since there is no way to properly position the intervals to form a valid type II representation, the digraph in Figure 2.2 is a type I interval digraph but not a type II interval digraph. However, if we place a loop at each vertex, there exists a type
II interval representation for the resulting graph: simply construct intervals such that \( S_u \cap R_u \cap S_v \cap R_v \cap S_w \cap R_w \neq \emptyset \).

However, even if we assume the existence of loops at every vertex in a type I interval digraph, there exists an example that has no type II interval representation. In particular, the digraph in Figure 2.3 is not a type II interval digraph. In this case, we have

- \( S_a \cap R_a \neq \emptyset \) and \( S_a \cap R_b \neq \emptyset \), but \( S_a \cap R_c = \emptyset \)
- \( S_c \cap R_b \neq \emptyset \) and \( S_c \cap R_c \neq \emptyset \), but \( S_c \cap R_a = \emptyset \)
- \( S_d \cap R_a \neq \emptyset \) and \( S_d \cap R_c \neq \emptyset \), but \( S_d \cap R_b = \emptyset \)

By replacing \( S_u \) with \( S_a \), \( S_v \) with \( S_c \), \( S_w \) with \( S_d \), \( R_u \) with \( R_c \), \( R_v \) with \( R_a \), and \( R_w \) with \( R_b \) in the argument used for the loopless example in Figure 2.2, we prove that it is not possible to arrange these six intervals in such a way that satisfies the
conditions above. Since we are unable to properly configure these six intervals, there is no way to create an interval representation with all eight intervals needed to represent the graph. Thus we conclude this is not a type II interval digraph.

Though we now have examples of a loopless and a fully looped type I interval digraph without type II interval representations, we know type II interval digraphs could have loops at all, none, or just some of their vertices. In fact, there exist type II interval representations for the digraph given in Figure 2.2 if we add loops to every vertex (as noted above) and similarly for the digraph in Figure 2.3 if we remove all loops. This leads to the question: Given a type I interval digraph, if we are allowed to assign loops to any combination of vertices we choose, can we always find some set of loops such that the resulting digraph has a type II interval representation? We apply a Ramsey theory based argument to demonstrate that this is not necessarily the case.

**Theorem 2.1.** There exists a type I interval digraph $D$ such that for all $S \subseteq V(D)$, the digraph $D$ with loops at exactly the vertices in $S$ is not a type II interval digraph.

**Proof.** For $n \in \mathbb{N}$, $n \geq 2$ let $G_n$ be the type I interval digraph with

$$V(G_n) = \{[a, b] \mid a, b \in \mathbb{N}, 1 \leq a < b \leq n\}$$

and edge set defined in the normal way for a type I interval digraph. We show that, for $n$ sufficiently large, $G_n$ is the digraph whose existence is asserted in the Theorem.
From Figure 2.2 above, we have a loopless type I interval digraph that is not a type II interval digraph. Observe that this graph (call it $H_1$) is an induced subgraph of $G_6$ (with vertices $u = [1, 6] \supseteq v = [2, 5] \supseteq w = [3, 4]$).

Similarly, from Figure 2.3 we have a type I interval digraph on 4 vertices with loops at every vertex that is not a type II interval digraph. This graph (call it $H_2$) is an induced subgraph of $G_8$ (with vertices $a = [3, 4], b = [1, 7], c = [5, 6], d = [2, 8]$).

Suppose we color every interval of $G_n$ either red (indicating no loop at the corresponding vertex) or blue (indicating the presence of a loop at that vertex). We create a bijection between intervals with integer endpoints in $[n]$ and edges of $K_n$. In particular, map $[a, b]$ to the edge $ab$. Then the red/blue coloring of the intervals is equivalent to a red/blue coloring of the edges of $K_n$.

From Ramsey theory, we know that there exists some $N$ such that any red/blue edge coloring of $K_N$ will contain either a red $K_6$ or a blue $K_8$. Notice that a red $K_6$ represents an induced $G_6$ with no loops, which we know contains an induced $H_1$. Similarly a blue $K_8$ represents an induced $G_8$ with loops at every vertex, which we know contains an induced $H_2$. Thus no matter what subset of the vertices of $G_N$ we choose to place loops at, there will be an induced subgraph that has no type II interval representation, so $G_N$ is not a type II interval digraph regardless of how vertices are assigned loops.

Another interesting observation is that for any type I interval digraph, the underlying undirected graph must itself be an interval graph. On the other hand, it
is possible to assign directions (allowing double edges) to the edges of some undirected non-interval graphs such that the resulting digraph has a type II interval representation. For example, any loopless cycle oriented to form a directed cycle has a type II interval representation. Even if we have a fully looped graph, there exists an orientation of $C_4$ that has a type II interval representation (see Figure 2.4), though $C_4$ is of course not an interval graph.

Now that we have established that type I and type II interval digraphs are distinct classes of digraphs and that neither class is contained within the other, we further explore some special cases of type I interval digraphs in an attempt to better understand the class of digraphs as a whole.
Chapter 3

Recognizing Type I Interval Digraphs

As mentioned previously, type I interval graphs are restricted in that their underlying undirected graph structure must be an interval graph, implying that any digraph formed by assignment of directions (including double edges) to a forbidden induced subgraph of an interval graph is a forbidden induced subdigraph for a type I interval graph. A more interesting consideration is forbidden induced subdigraphs for type I interval graphs such that the underlying undirected graph is an interval graph. Figure 3.1 shows five forms determined to be forbidden for type I interval digraphs. The dotted line in the rightmost example indicates that the edges must exist, but they may be oriented in any way (including a double edge). Note that this list is not exhaustive.

For proofs that these examples have no type I interval representations, see the Appendix Section 6.1. We do not currently have a complete list of forbidden in-
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

Figure 3.1: Known minimal forbidden induced subdigraphs for type I interval digraphs

duced subdigraphs for type I interval digraphs, but there are some special subclasses of type I interval digraphs that we fully characterize.

3.1 Special Subclasses of Type I Interval Digraphs

We begin by examining trees. Suppose $T$ is a tree and $L \subseteq V(T)$ is the set of leaves of $T$. We say that $T$ is a caterpillar if the subgraph induced by the vertices $V(T) \setminus L$ is a path graph. In [21], Eckhoff showed that the connected triangle-free interval graphs are exactly the caterpillar trees. This implies that a tree is an interval graph if and only if it is a caterpillar. Extending this to type I interval digraphs, we have the following.

**Proposition 3.1.** Let $T$ be a caterpillar with at least two non-leaf vertices. Let $L$ be the set of leaves in $V(T)$, and let $u$ and $v$ be the distinct endpoints of the path graph induced by $V(T) \setminus L$. There exist exactly $2(d(u) + 1)(d(v) + 1)$ ways to assign directions to the edges of $T$ (allowing double edges) such that the resulting digraph is a type I interval digraph.
Proof. From Figure 3.1 observe that any induced path must be a directed path. We thus have 2 choices for the orientation of the path graph induced by all non-leaf vertices of $T$. Additionally, because of the first two forbidden induced subdigraphs, we know that all edges with one endpoint in $L$ and one endpoint in $V(T) \setminus (L \cup \{u,v\})$ must be double edges.

This leaves the orientation of the edges with one endpoint in $L$ and the other endpoint in $\{u,v\}$. We are again restricted by the first two forbidden induced subdigraphs. Without loss of generality suppose the “body” path is oriented from $u$ to $v$. Then we may have at most one edge with one endpoint $u$ and the other endpoint in $L$ oriented from the vertex in $L$ to $u$ and at most one edge with one endpoint $v$ and the other endpoint in $L$ oriented from $v$ to the vertex in $L$. All other such edges must be double edges. This gives us $d(u) + 1$ and $d(v) + 1$ options respectively (we either pick one edge to be oriented differently or opt to orient all as double edges).

Next, we look at digraphs $D$ that have the complete graph as their underlying structure. Let $D_1$ denote the subdigraph of $D$ whose edge set is exactly the single edges of $D$. From Figure 3.1 we see that if $D$ is a type I interval digraph, $D_1$ must be a transitive orientation; that is, if $uv$ and $vw$ are edges in $D_1$, $uw$ must also be an edge. This allows us to naturally associate a partially ordered set with $D_1$ by setting $u < v$ if and only if the edge $uv$ is in $D_1$.

Recall that a linear extension of a poset $P$ is a totally ordered set that preserves
the relation of \( P \). A realizer of \( P \) is a set of linear extensions whose intersection is \( P \), and the dimension of \( P \) is the minimum size of a realizer for that poset. We have the following result.

**Proposition 3.2.** Suppose \( D \) is a directed complete graph. \( D \) is a type I interval digraph if and only if the poset associated with \( D_1 \) has dimension at most 2.

**Proof.** Suppose \( D \) is a type I interval digraph with vertex set \( \{v_1, v_2, \ldots, v_n\} \) and consider a type I interval representation for \( D \). Since the underlying structure of \( D \) is a complete graph, we have that for all \( 1 \leq i, j \leq n, L(v_i) < R(v_j) \), that is all left endpoints must appear before any right endpoint in the representation. If not, then \( R(v_{i'}) < L(v_{j'}) \) for some \( v_{i'} \) and \( v_{j'} \), which implies \( I_{v_{i'}} \leq I_{v_{j'}} \) and thus neither \( v_{i'} v_{j'} \) nor \( v_{j'} v_{i'} \) is in the edge set, contradicting the completeness of the underlying structure.

Now consider the poset associated with \( D_1 \). Suppose \( v_i < v_j \) in the poset. Then we know \( v_i v_j \) is in \( D_1 \), and from the interval representation \( L(v_i) < L(v_j) \) and \( R(v_i) < R(v_j) \). On the other hand, if \( v_i \) and \( v_j \) are incomparable, neither \( v_i v_j \) nor \( v_j v_i \) are in \( D_1 \). Thus in the interval representation, \( L(v_i) < L(v_j) \) and \( R(v_i) > R(v_j) \) or vice versa.

This implies that the linear extensions defined by the left endpoints of the intervals (read left to right) and the right endpoints of the intervals (read left to right) form a realizer for the poset corresponding to \( D_1 \), so we conclude the dimension of the poset is at most 2.
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

Now suppose that the poset associated with $D_1$ has dimension at most 2. Let $<_L$ and $<_R$ be the two linear extensions in a minimum size realizer. Create an interval representation by ordering the left endpoints from left to right according to $<_L$ and to the right of that ordering the right endpoints from left to right according to $<_R$.

If $v_i < v_j$ then $v_i <_L v_j$ and $v_i <_R v_j$. This implies $L(v_i) < L(v_j)$ and $R(v_i) < R(v_j)$ so the edge $v_i v_j$ is in $D_1$ (and thus in $D$). If $v_i$ and $v_j$ are incomparable, either $v_i <_L v_j$ and $v_i >_R v_j$, or vice versa. In either case, this implies neither $v_i v_j$ nor $v_j v_i$ exist in $D_1$, but a double edge $v_i v_j$ exists in $D$. Since we have created an interval representation for $D$, we conclude $D$ is a type I interval digraph.

Another special case we characterize is the set of proper or unit type I interval digraphs. Generalizing from their descriptions from undirected interval graphs, we say a digraph is a proper type I interval digraph if there exists a type I interval representation for the graph such that no interval is properly contained in another interval. We say a digraph is a unit type I interval digraph if there exists a type I interval representation for the graph such that each interval is the same length.

In [22], Bogart and West presented a result for undirected interval graphs proving that an interval graph is unit if and only if it is proper, and furthermore that such graphs are characterized as the set of interval graphs that do not contain $K_{1,3}$ as an induced subgraph. We provide a similar result for type I interval digraphs.

**Theorem 3.3.** Suppose $G$ is a type I interval digraph. The following are equivalent:

(i) $G$ is proper.
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

(ii) $G$ is unit.

(iii) For all $u, v \in V(G)$, if $uv$ and $vu$ are both edges, then $u$ and $v$ are twin vertices (i.e. the set of in-neighbors of $u$ is identical to the set of in-neighbors of $v$, and the set of out-neighbors of $u$ is identical to the set of out-neighbors of $v$).

Proof. The proof that (i) is true if and only if (ii) is true is identical to the proof for undirected interval graphs (for details, see [22]).

Next, we show that (i) implies (iii). Suppose that $G$ is a proper type I interval digraph. In order for two vertices $u, v \in V(G)$ to have both $uv$ and $vu$ as edges, either $I_u \subseteq I_v$ or $I_v \subseteq I_u$. However since $G$ is proper, we cannot have proper containment of intervals, so it must be the case that $I_u = I_v$. Thus for each $w \in V \setminus u, v$, $I_u$ and $I_v$ must interact with $I_w$ in the same way, so $u$ and $v$ are twin vertices.

Finally we show (iii) implies (i). Suppose that $G$ is a type I interval digraph such that for all $u, v \in V(G)$ with $uv, vu \in E$, $u$ and $v$ are twin vertices. Pick an endpoint $w$ of a double edge in $G$ and consider the induced subdigraph of $G$ on $V \setminus w$. Repeat this process until we have a graph $G'$ with no double edges. Because $G'$ has no double edges, the type I interval representation for $G'$ has no proper containment of intervals. Thus $G'$ is proper. Adding back the vertices we removed, since each removed vertex $w$ is a twin vertex to some vertex $w'$ in $G'$, we simply set $I_w = I_{w'}$. The resulting interval representation is still proper, so we conclude $G$ is proper. \qed
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

We have fully characterized all type I interval digraphs that are proper/unit and those that have an underlying structure that is a complete graph or a tree. To facilitate characterization of general type I interval digraphs, we modify the graphs to encode more information about the potential interval representation in its structure.

3.2 Type Ib Interval Digraphs

Consider a variation on directed graphs with two types of edges; we differentiate edges of different types using solid and dashed lines and thus refer to this style of graph as a solid-dashed digraph. We say a solid-dashed digraph $G = (V, E)$ is a type Ib interval digraph if there exists a one-to-one interval representation $f : V \rightarrow \mathcal{I}$ such that for all $u, v \in V$, there exists a dashed edge from $v$ to $u$ if and only if $I_u \subseteq I_v$ and a solid edge from $v$ to $u$ if and only if $I_v < I_u$. We require the existence of a one-to-one interval representation to ensure the existence of an interval representation with distinct interval endpoints.

![Interval configurations creating dashed and solid edges](image)

**Figure 3.2:** Interval configurations creating dashed and solid edges
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

Given $G = (V, E)$, a type Ib interval digraph, $v \in V$, and an interval representation $f : V \rightarrow \mathcal{I}$, we use the following notation:

- $N_G^+(v)$ is the set of out-neighbors of $v$ in $G$ along solid edges and $d_G^+(v) = |N_G^+(v)|$.
- $N_G^{++}(v)$ is the set of out-neighbors of $v$ in $G$ along dashed edges and $d_G^{++}(v) = |N_G^{++}(v)|$.
- $N_G^0(v)$ is the set of vertices that are not out-neighbors of $v$ in $G$, that is $V \setminus (N^+(v) \cup N^{++}(v))$.

We make a key observation which serves as the foundation for the recognition of type Ib (and indirectly type I) interval digraphs.

**Proposition 3.4.** Let $G = (V, E)$ be a type Ib interval digraph. For each connected component of $G$, there exists exactly one vertex with in-degree 0 (for both solid and dashed edges).

**Proof.** Consider a connected component of $G$. First, we demonstrate that a vertex with in-degree 0 must exist. Consider an interval representation for $G$ such that all interval endpoints are distinct and let $v$ be the vertex whose corresponding interval has the leftmost left endpoint in $G$. Then the possible intersections $I_v$ can have with other intervals are as seen in Figure 3.2.

Notice that every edge incident on $v$ is directed out of $v$. Thus, we know that for each component of $G$ there exists at least one vertex with in-degree 0.
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

Now we show that \( v \) is unique in its component. Notice that in order for a vertex to have in-degree 0, \( L(I_v) \) cannot be contained in any other interval. This implies the existence of some \( \epsilon > 0 \) such that \( L(I_v) - \epsilon \) is not contained in any interval; in simpler (albeit less rigorous terms), there is a gap between \( L(I_v) \) and any intervals to the left of \( L(I_v) \). Observe that everything to the left of \( L(I_v) - \epsilon \) must be in a different connected component, so only the vertex whose interval has the leftmost left endpoint of any in its connected component can have in-degree 0. This implies \( v \) is unique as desired.

Taking advantage of this result along with our knowledge of type I interval digraphs, we have the following characterization of type Ib interval digraphs.

**Theorem 3.5.** A solid-dashed digraph \( G \) is a type Ib interval digraph if all of the following hold:

1. The underlying undirected graph is an interval graph.
2. No forbidden subgraph from Figure 3.3 is an induced subgraph of \( G \).
3. Every induced subgraph has exactly one vertex in each connected component with in-degree 0 (for both solid and dashed edges).

Since being a type Ib interval digraph is a hereditary property (i.e. if \( G \) is a type Ib interval digraph and \( H \) is an induced subdigraph of \( G \), then \( H \) is also a type Ib interval digraph), we can use this characterization to construct a recursive
Figure 3.3: A list of forbidden subgraphs for type I\textsubscript{b} interval digraphs relevant to the algorithm.
algorithm for recognizing type Ib interval digraphs. The proof of correctness for the algorithm also serves to prove the theorem.

3.2.1 Recognition Algorithm

Given as input a solid-dashed digraph $G = (V, E)$, the algorithm determines whether $G$ is a type Ib interval digraph. If it is, we return an interval representation for $G$ with unique interval endpoints in $[2n]$. The recursive algorithm proceeds as follows:

- Check whether the underlying structure of $G$ (i.e. the graph with edge directions ignored and solid/dashed edges treated the same) is an interval graph. If not, return FALSE.

- Check $G$ for forbidden induced subgraphs (see Figure 3.3). If such a subgraph is found, $G$ is not a type Ib interval digraph. Return FALSE.

- Create a list of components in $G$. We address each component individually, then combine the resulting interval representations to get an interval representation for $G$.

- Given a component $H$, find a vertex $\ell$ with in-degree 0.
  
  – If no such vertex exists or if multiple such vertices exist, by Proposition 3.4 we know $H$ is not a type Ib interval digraph. Return FALSE.
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

Figure 3.4: Interval configurations that must be rearranged in order to interact properly with $\ell$. Throughout our algorithm we use $o$ to indicate an interval that should overlap $I_\ell$, $c$ to indicate an interval that should be contained in $I_\ell$, and $x$ to indicate an interval that should be disjoint from $I_\ell$. We call these conflicts type (i), (ii), and (iii) respectively based on the order we address them in the algorithm.

- If $\ell$ is the only vertex of component $H$, set $I_\ell = [1, 2]$ and return.
- Otherwise, set $I_\ell = [1, d^+ \ell (\ell) + 2d^{++} \ell (\ell) + 2]$, remove $\ell$ from the graph, and call the algorithm on $H - \ell$ to obtain an interval representation for $H - \ell$.

- Search the interval representation for $H - \ell$ for conflicting patterns (see Figure 3.4).

Consider a pair of intervals involved in the conflicting pattern (say $I$ and $J$). We define the sub-component of $J$ with respect to $I$ to be the collection of intervals that are in the same connected component as $J$ when we remove any intervals that contain both $I$ and $J$.

Let $j$ be the minimum index assigned to an endpoint of an interval in the sub-component of $J$ with respect to $I$ (for convenience, denote this $H_{j,l}$). In
particular, this means \( j \) is the value assigned to the leftmost left endpoint of any interval corresponding to a vertex in \( H_{I,I} \). Similarly, let \( j' \) be the maximum index assigned to an endpoint of an interval in \( H_{J,I} \), that is the value assigned to the rightmost right endpoint of any interval corresponding to a vertex in \( H_{J,I} \). Additionally, let \( i \) be the minimum index over all intervals in the sub-component of \( I \) with respect to \( J \) (denoted \( H_{I,J} \)), and let \( i' \) be the maximum index over all intervals in \( H_{I,J} \).

Suppose \( I \ll J \), so \( i < i' < j < j' \). We shift the intervals corresponding to \( H_{J,I} \) to have minimum index \( i \) by adding \( i - j \) to each endpoint value from \( j \) to \( j' \) (the new indices are \( i \) to \( i - j + j' \)). We then shift any intervals between \( H_{I,J} \) and \( H_{J,I} \) by adding \( (j' - j) - (i' - i) \) to each endpoint value from \( i' + 1 \) to \( j - 1 \) (the new indices are \( i - j + j' + 1 \) to \( j' + i - i' - 1 \)). Finally, we shift \( H_{I,J} \) to have maximum index \( j' \) by adding \( j' - i' \) to each endpoint values from \( i \) to \( i' \) (the new indices are \( j' + i - i' \) to \( j' \)). This process changes the relative positioning of \( H_{J,I} \) and \( H_{I,J} \) without altering the relative positioning of any pair of intervals outside of these two sub-components.

Consider each interval \( I_o \) associated with a vertex \( o \in N_{H}^+(\ell) \), moving from left to right in the interval representation for \( H - \ell \). For each such \( I_o \), find the rightmost conflicting \( I_x \) such that \( x \in N_{H}^0(\ell) \) and swap the corresponding sub-components. If no such conflict exists, move to the next interval associated with a vertex in \( N_{H}^+(\ell) \). Repeat this process until
all type (i) conflicts are resolved.

- Look at each interval $I_c$ associated with a vertex $c \in N_H^{++}(\ell)$, moving from left to right in the interval representation for $H - \ell$. For each such $I_c$, find the rightmost conflicting $I_x$ such that $x \in N_H^0(\ell)$ and swap the corresponding sub-components. If no such conflict exists, move to the next interval associated with a vertex in $N_H^{++}(\ell)$ Repeat this process until all type (ii) conflicts are resolved.

- Look at each interval $I_o$ associated with a vertex $o \in N_H^{+}(\ell)$, moving from left to right in the interval representation for $H - \ell$. For each such $I_o$, find the leftmost conflicting $I_c$ such that $c \in N_H^{++}(\ell)$ and swap the corresponding sub-components. If no such conflict exists, move to the next interval associated with a vertex in $N_H^{+}(\ell)$ Repeat this process until all type (iii) conflicts are resolved.

- Shift the indices of $H - \ell$ to accommodate $\ell$ by adjusting each endpoint value in the interval representation of $H - \ell$. Add 1 to each endpoint value from 1 to $d_H^+(\ell) + 2d_H^{++}(\ell)$, and add 2 to each endpoint value that is $d_H^+(\ell) + 2d_H^{++}(\ell) + 1$ or greater.

- Return the resulting interval representation for $H$.

- If applicable, repeat this process for other connected components of $G$. Concatenate the interval representations of the components (shifting indices for
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

each component accordingly) to obtain and return an overall interval representation for $G$.

3.2.2 Examples

To give some intuition as to how the algorithm operates, consider the following examples for how it deals with some small cases.

First, we examine the case of adding back a vertex to the single vertex graph. At this point, the algorithm has an interval representation $I_v = [1, 2]$ for $H - \ell$, which consists only of the vertex $v$. We now reintroduce the vertex $\ell$.

- If $v \in N_H^+(\ell)$, the right endpoint of $\ell$ is assigned to $1 + 2 \cdot 0 + 2 = 3$. We then adjust the endpoints of the original interval, giving us $[1, 3]$ and $[2, 4]$ as the interval representation.

- If $v \in N_H^{++}(\ell)$, the right endpoint of $\ell$ is assigned to $0 + 2 \cdot 1 + 2 = 4$. We then adjust the endpoints of the original interval, giving us $[1, 4]$ and $[2, 3]$ as the interval representation.

- If $v \in N_H^0(\ell)$, the right endpoint of $\ell$ is assigned to $0 + 2 \cdot 0 + 2 = 2$. We then adjust the endpoints of the original interval, giving us $[1, 2]$ and $[3, 4]$ as the interval representation.

No further adjustments need to be made in any of these cases.
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

For a second example, consider the following situation: $H - \ell$ consists of three vertices $x, o, c$ with $I_x = [1, 2]$, $I_o = [3, 4]$, and $I_c = [5, 6]$. We have $x \in N_H^0(\ell)$, $o \in N_H^+(\ell)$, and $c \in N_H^{++}(\ell)$. We now reintroduce the vertex $\ell$.

- The algorithm first computes the right endpoint of $\ell$ as $1 + 2 \cdot 1 + 2 = 5$. Thus $I_\ell = [1, 5]$.

- We have a type (i) conflict involving $x$ and $o$. Shifting indices to swap the position of the sub-component of $I_x$ with respect to $I_o$ and the sub-component of $I_o$ with respect to $I_x$, we end up with $I_x = [3, 4]$, $I_o = [1, 2]$, and $I_c = [5, 6]$. At this point we have no further type (i) conflicts.

- We have a type (ii) conflict involving $x$ and $c$. Shifting indices to swap the position of the sub-component of $I_x$ with respect to $I_c$ and the sub-component of $I_c$ with respect to $I_x$, we end up with $I_x = [5, 6]$, $I_o = [1, 2]$, and $I_c = [3, 4]$. At this point we have no further type (ii) conflicts.

- We have a type (iii) conflict involving $o$ and $c$. Shifting indices to swap the position of the sub-component of $I_o$ with respect to $I_c$ and the sub-component of $I_c$ with respect to $I_o$, we end up with $I_x = [5, 6]$, $I_o = [3, 4]$, and $I_c = [1, 2]$. At this point we have no further type (iii) conflicts.

- Adjusting the endpoints of these intervals to accommodate $I_\ell$, we have the final interval representation with $I_\ell = [1, 5]$, $I_x = [7, 8]$, $I_o = [4, 6]$, and $I_c = [2, 3]$. 

37
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

3.2.3 Proof of Correctness

We show that the algorithm terminates and correctly identifies whether G is a type Ib interval digraph (finding an interval representation if it is).

First note that if the underlying undirected structure of a solid/dashed digraph G is not an interval graph, G is not a type Ib interval digraph. If G were type Ib interval, we would be able to use the resulting interval representation to encode the undirected graph, thus implying the underlying structure would be an interval graph. Additionally, observe that the forbidden subgraphs in Figure 3.3 are indeed impossible to encode with an interval representation (for complete proofs, see the Appendix Section 6.2).

With these special cases dealt with, we proceed by appealing to a loop invariant: the interval representation for the induced subgraph $H - \ell$ before adding back $\ell$ is a valid interval representation for $H - \ell$. The base cases are straightforward; trivially the empty interval representation is correct for the empty graph, and for a single vertex graph we have the single interval $[1, 2]$. Suppose that up to some vertex $\ell$, the recursive construction of the interval representation for G is correct. First we demonstrate that the assignment of the interval for $\ell$ yields a valid interval representation for $H$.

Lemma 3.6. If H is a connected type Ib interval digraph and $\ell$ is the unique vertex in H with in-degree 0, then in any interval representation of H with unique interval endpoints
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

in $[2n]$ we must have

$$I_{\ell} = [1, d_{H}^{+}(\ell) + 2d_{H}^{+}(\ell) + 2].$$

Proof. We know from the proof of Proposition 3.4 that $\ell$ must correspond to the interval with leftmost left endpoint, so $L(I_{\ell}) = 1$.

Observe that if an interval representation exists, it must be the case that

- For each vertex $o$ such that $o \in N_{H}^{+}(\ell)$, $L(I_{\ell}) < L(I_{o}) < R(I_{o}) < R(I_{\ell})$.
- For each vertex $c$ such that $c \in N_{H}^{0+}(\ell)$, $L(I_{\ell}) < L(I_{c}) < R(I_{c}) < R(I_{\ell})$.
- For each vertex $x$ such that $x \in N_{H}^{0}(\ell)$, $L(I_{\ell}) < R(I_{\ell}) < L(I_{x}) < R(I_{x})$.

Thus in any interval representation, we must have exactly $R(I_{\ell}) = d_{H}^{+}(\ell) + 2d_{H}^{+}(\ell) + 2$.

Next we show that if an interval from the representation for $H - \ell$ does not interact with $I_{\ell}$ correctly, there must be a forbidden pattern or conflicting pattern (as seen in Figures 3.6 and 3.4 respectively) appearing in the representation. Figure 3.5 shows the different possibilities for how intervals could be misplaced relative to the new interval $I_{\ell}$.

- Suppose we have $x \in N_{H}^{0}(\ell)$ such that $I_{x} \subseteq I_{\ell}$ or $I_{\ell} < I_{x}$. In either case, $L(I_{x}) < R(I_{\ell})$. Since $L(I_{\ell}) = 1$ and $R(I_{\ell}) = d_{H}^{+}(\ell) + 2d_{H}^{+}(\ell) + 2$, this leaves only $d_{H}^{+}(\ell) + 2d_{H}^{+}(\ell) - 1$ more values between 2 and $d_{H}^{+}(\ell) + 2d_{H}^{+}(\ell) + 1$ to assign to the endpoints of intervals corresponding to vertices in $N^{+}(\ell)$ and
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

\[ \begin{array}{c}
\text{x} & \text{c} & \text{o'} \\
\hline
\text{\ell} \\
\text{o} & \text{x'} & \text{c'} \\
\end{array} \]

**Figure 3.5:** Interval configurations that result in incorrect edges formed relative to \( \ell \). \( o \) indicates an interval that should overlap \( I_\ell \), \( c \) indicates an interval that should be contained in \( I_\ell \) and \( x \) indicates an interval that should be disjoint from \( I_\ell \).

\[ N^{++}(\ell). \] Thus, there exists some \( c \in N^{++}(\ell) \) such that \( R(I_c) > R(I_\ell) \) or some \( o \in N^+(\ell) \) such that \( L(I_o) > R(I_\ell) \).

- If \( R(I_c) > R(I_\ell) \), this implies \( L(I_x) < R(I_c) \), which in turn indicates an interaction between \( I_x \) and \( I_c \) that gives either a forbidden or conflicting pattern.

- If \( L(I_o) > R(I_\ell) \), this implies \( L(I_x) < L(I_o) \), which in turn indicates an interaction between \( I_x \) and \( I_o \) that gives either a forbidden or conflicting pattern.

- Suppose we have \( c \in N^{++}(\ell) \) such that \( I_\ell < I_c \) or \( I_\ell \ll I_c \). In either case, \( R(I_\ell) < R(I_c) \). Since \( L(I_\ell) = 1 \) and \( R(I_\ell) = d_H^+(\ell) + 2d_H^{++}(\ell) + 2 \), and at most one of the \( d_H^+(\ell) + 2d_H^{++}(\ell) \) values between 2 and \( d_H^+(\ell) + 2d_H^{++}(\ell) + 1 \) is assigned to an endpoint of \( I_c \), there exists some \( x \in N^0(\ell) \) such that \( L(I_x) < R(I_\ell) \) or some \( o \in N^+(\ell) \) such that \( R(I_o) < R(I_\ell) \).
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

- If $L(I_x) < R(I_c)$, this implies $L(I_x) < R(I_c)$, which in turn indicates an interaction between $I_x$ and $I_c$ that gives either a forbidden or conflicting pattern.

- If $R(I_o) < R(I_o)$, this implies $R(I_o) < R(I_c)$, which in turn indicates an interaction between $I_o$ and $I_c$ that gives either a forbidden or conflicting pattern.

- Suppose we have $o \in N^+(\ell)$ such that $I_o \subseteq I_{\ell}$. Since $L(I_{\ell}) = 1$ and $R(I_{\ell}) = d^+_H(\ell) + 2d^+_H(\ell) + 2$, this leaves $d^+_H(\ell) + 2d^+_H(\ell) - 1$ more values between 2 and $d^+_H(\ell) + 2d^+_H(\ell) + 1$ to assign to the endpoints of intervals corresponding to vertices in $N^+(\ell)$ and $N^{++}(\ell)$. Thus, there exists some $c \in N^{++}(\ell)$ such that $R(I_c) > R(I_{\ell})$ or some $o' \in N^+(\ell)$ such that $L(I_{o'}) > R(I_{\ell})$.

- If $R(I_c) > R(I_{\ell})$, this implies $R(I_o) < R(I_c)$, which in turn indicates an interaction between $I_o$ and $I_c$ that gives either a forbidden or conflicting pattern.

- If $L(I_{o'}) > R(I_{\ell})$, this implies $R(I_o) < L(I_{o'})$, which in turn indicates an interaction between $I_o$ and $I_{o'}$ that gives a forbidden pattern.

- Suppose we have $o' \in N^+(\ell)$ such that $I_{\ell} \ll I_{o'}$. Since $L(I_{\ell}) = 1$, $R(I_{\ell}) = d^+_H(\ell) + 2d^+_H(\ell) + 2$, and none of the $d^+_H(\ell) + 2d^+_H(\ell)$ values between 2 and $d^+_H(\ell) + 2d^+_H(\ell) + 1$ are assigned to an endpoint of $I_{o'}$, there exists some $x \in N^0(\ell)$ such that $L(I_x) < R(I_{\ell})$ or some $o \in N^+(\ell)$ such that $R(I_o) < R(I_{\ell})$. 41
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

– If \( L(I_x) < R(I_\ell) \), this implies \( L(I_x) < L(I_{o'}) \), which in turn indicates an interaction between \( I_x \) and \( I_{o'} \) that gives either a forbidden or conflicting pattern.

– If \( R(I_o) < R(I_\ell) \), this implies \( R(I_o) < L(I_{o'}) \), which in turn indicates an interaction between \( I_o \) and \( I_{o'} \) that gives a forbidden pattern.

Thus we see that the existence of an incorrect interval interaction between \( I_\ell \) and an interval generated for the representation of \( H - \ell \) implies the presence of a forbidden or conflicting pattern. It remains to show that forbidden patterns imply induced subgraphs that are forbidden for type Ib interval digraphs and that adjustments to conflicting patterns by the algorithm result in fewer conflicting patterns in the altered representation.

3.2.3.1 Forbidden Subgraphs/Patterns

Figure 3.3 shows the forbidden subgraphs for type Ib interval digraphs. While some of these are also forbidden subdigraphs for type I interval digraphs when considered as directed graphs, not all of them are.

Note that in the context of the algorithm, if we consider the interval representation for an intermediate subgraph \( H - \ell \), the interval arrangements seen in Figure 3.6 represent forbidden patterns that cannot appear (as in combination with \( \ell \), they imply a forbidden subgraph).
3.2.3.2 Conflicting Patterns

Suppose the algorithm has constructed an interval representation up through vertex $\ell$. This implies that the input graph contains none of the forbidden sub-graphs for a type Ib interval digraph. Note that the goal is to swap the relative positions of two sub-components; shifting these sub-components alters the interval representation, but does not change the structure of the corresponding graph. Thus we know the swaps do not introduce any forbidden patterns into the representation. There are three possible conflicting patterns we could encounter in $H - \ell$. Let $o \in N^+_H(\ell), c \in N_{H++}^+(\ell)$, and $x \in N_H^0(\ell).$
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

**Type (i) Conflict**

Consider a type (i) conflict: \( o \in N_H^+(\ell), x \in N_H^0(\ell) \) such that \( I_x \ll I_o \). In particular, for a given \( o \), let \( x \) be such that there exists no \( x' \in N_H^0(\ell) \) with \( I_{x'} \ll I_o \) but \( R(I_{x'}) > R(I_x) \). We claim that the sub-component of \( I_o \) with respect to \( I_x \) (call this \( H_o \)) and the sub-component of \( I_x \) with respect to \( I_o \) (call this \( H_x \)) are not the same.

Consider \( R(I_x) \). We know from the forbidden patterns (see Figure 3.6) that \( R(I_x) \notin I_c \) for any \( c \in N_H^{++}(\ell) \). Similarly, we cannot have \( R(I_x) \in I_{x'} \) for some \( x' \in N_H^0(\ell) \). This is because we know that \( I_0 \cap I_x \neq \emptyset \) by the selection of \( x \); however, we also know \( L(I_x) < L(I_o) \) and the forbidden patterns show that we cannot have \( I_{x'} < I_o \) nor can we have \( I_o \subseteq I_{x'} \). Thus it is not possible to have \( R(I_x) \in I_{x'} \).

Finally, we cannot have \( R(I_x) \in I_o \) for some \( o' \in N_H^{++}(\ell) \). We know we cannot have \( I_x < I_{o'} \) or \( I_{o'} \ll I_0 \). Thus \( I_x \subseteq I_{o'} \) but since we also know that \( I_{o'} \) is in \( H_x \), we cannot have \( I_o \subseteq I_{o'} \). The only remaining possibility is \( I_{o'} < I_o \), but this leads to the first three-interval forbidden pattern in Figure 3.6. Thus it is not possible to have \( R(I_x) \in I_{o'} \).

Since \( R(I_x) \) is not contained in any other interval in \( H_x \), we know that this sub-component does not contain \( I_o \), and thus \( H_x \) and \( H_o \) are disjoint. Thus we swap the order of \( H_x \) and \( H_o \) in the interval representation for \( H - \ell \). We claim that the resulting interval representation does not create any new type (i) conflicts.

Observe that since no pair of intervals outside \( H_x \) and \( H_o \) has moved relative to
their original positioning, some interval in either \( H_x \) or \( H_0 \) must be involved in any newly created conflicts. However we know that there is no \( o' \in N_{H}^+(\ell) \) such that \( I_{o'} \) was originally to the left of the intervals representing \( H_{o'} \), as this would imply \( I_{o'} \ll I_0 \) in the original representation, which is a forbidden configuration. Thus there are no new type (i) conflicts generated by this swap, so the overall number of type (i) conflicts decreases by at least one after this operation.

**TYPE (ii) CONFLICT**

Once we have resolved all type (i) conflicts in the interval representation for \( H - \ell \), we move on to address type (ii) conflicts. Consider a type (ii) conflict: \( c \in N_{H}^{++}(\ell), x \in N_{H}^0(\ell) \) such that \( I_x \ll I_c \) and \( I_c \) is the leftmost interval involved in a \( I_x \ll I_c \) conflict. For a given \( c \), let \( x \) be such that there exists no \( x' \in N_{H}^0(\ell) \) with \( I_{x'} \ll I_c \) but \( R(I_{x'}) > R(I_x) \). We claim that the sub-component of \( I_c \) with respect to \( I_x \) (call this \( H_c \)) and the sub-component of \( I_x \) with respect to \( I_c \) (call this \( H_x \)) are not the same.

Consider \( R(I_x) \). We know from the forbidden patterns (see Figure 3.6) that \( R(I_x) \notin I_{c'} \) for any \( c' \in N_{H}^{++}(\ell) \). Similarly, we cannot have \( R(I_x) \in I_{x'} \) for some \( x' \in N_{H}^0(\ell) \) since \( I_{x'} \cap I_c = \emptyset \) by the forbidden patterns but \( R(I_{x'}) > R(I_x) \) and thus we cannot have \( I_{x'} \ll I_c \). Observe that if \( R(I_x) \) is not contained in any other interval in \( H_x \), we know that this sub-component does not contain \( I_c \), and thus \( H_x \) and \( H_c \) are disjoint.
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

Finally, observe that if we have \( R(I_x) \in I_o \) for some \( o \in N^+_H(\ell) \), we know we cannot have \( I_x < I_o \) or \( I_o < I_c \). Thus \( I_x \subseteq I_o \) and \( I_o \ll I_c \). We show in the type (iii) case below that this implies \( H_x \) and \( H_c \) are disjoint sub-components.

Since \( H_x \) and \( H_c \) are disjoint, we swap their order in the interval representation for \( H - \ell \). We claim that the resulting interval representation does not add new type (i) conflicts, and the number of type (ii) conflicts has decreased by at least one. Observe that since no pair of intervals outside \( H_x \) and \( H_c \) has moved relative to their original positioning, some interval in either \( H_x \) or \( H_c \) must be involved in any newly created conflicts.

There cannot be a type (i) or type (ii) conflict involving an interval in \( H_x \) and some interval between \( H_x \) and \( H_c \), as this would imply the existence of some \( x' \) that originally satisfied \( I_{x'} \ll I_c \) and \( I_x \ll I_{x'} \). This contradicts \( x \) being the rightmost \( x \in N^0(\ell) \) with \( I_x \ll I_c \) in the starting representation.

Similarly, there cannot be a type (ii) conflict involving an interval in \( H_c \) and some interval between \( H_c \) and \( H_x \), as this would imply the existence of some \( c' \) that originally satisfied \( I_c \ll I_{c'} \) and \( I_{c'} \ll I_c \). This contradicts \( c \) being the leftmost \( c \in N^{++}(\ell) \) involved in a type (ii) conflict. There cannot be a new type (i) conflict involving \( H_c \) and some interval between \( H_c \) and \( H_x \), as this would imply the existence of some \( o \) that originally satisfied \( I_x \ll I_o \), which is a type (i) conflict. We have already addressed all type (i) conflicts by this stage of the algorithm, so this case is impossible.
Lastly, observe that if there is a new type (i) or type (ii) conflict involving some interval in $H_x$ and some interval in $H_c$, it must be the case that there exists some $x' \in N_H^0(\ell)$ with $x' \in H_c$. However, from Figure 3.6 we see that we must have $I_c \cap I_{x'} = \emptyset$, so there must exist $o \in N_H^+(\ell)$ with $o \in H_c$. By similar logic, we must have some $o' \in H_x$. However this implies $I_o \ll I_{o'}$, which is a forbidden pattern and thus a contradiction.

Putting all this together, we conclude that the swap does not create type (i) conflicts in the interval representation, and the number of type (ii) conflicts decreases by at least one.

**Type (iii) Conflict**

Once we have resolved all type (i) and type (ii) conflicts in the interval representation for $H - \ell$, we move on to address type (iii) conflicts. Consider a type (iii) conflict: $o \in N_H^+(\ell), c \in N_H^{++}(\ell)$ such that $I_o \ll I_c$. In particular, for a given $o$, let $c$ be such that there exists no $c' \in N_H^{++}(\ell)$ with $I_o \ll I_{c'}$ but $L(I_{c'}) < L(I_c)$. We claim that the sub-component of $I_o$ with respect to $I_c$ (call this $H_o$) and the sub-component of $I_c$ with respect to $I_o$ (call this $H_c$) are not the same.

Consider $L(I_c)$. We know from the forbidden patterns (see Figure 3.6) that $L(I_c) \notin I_x$ for any $x \in N_H^0(\ell)$. Similarly, we cannot have $L(I_c) \in I_{c'}$ for some $c' \in N_H^{++}(\ell)$ as this would imply $I_o \cap I_{c'} \neq \emptyset$ by the selection of $c$, which in turn implies either $I_o < I_{c'}$ or $I_o \subseteq I_{c'}$. Both of these are forbidden configurations. Fi-
Finally, we cannot have \( L(I_c) \in I_0' \) for some \( o' \in N_H^+(\ell) \). We know we cannot have \( I_0' < I_c \) or \( I_0 < I_0' \) from the forbidden patterns, so the only case left to check is \( I_c \subseteq I_0' \). However, since \( I_0' \) is in \( H_c \), we cannot have \( I_0 \subseteq I_0' \). The only remaining possibility is \( I_o < I_0' \), but this leads to the second three-interval forbidden pattern in Figure 3.6. Thus it is not possible to have \( L(I_c) \in I_0' \).

Since \( L(I_c) \) is not contained in any other interval in \( H_c \), we know that this subcomponent does not contain \( I_o \), and thus \( H_c \) and \( H_o \) are disjoint. As a result, we swap \( H_c \) and \( H_o \) in the interval representation for \( H - \ell \). We claim that the resulting interval representation does not create any type (i) or type (ii) conflicts, and the number of type (iii) conflicts decreases by at least one. Observe that since no pair of intervals outside \( H_o \) and \( H_c \) has moved relative to their original positioning, some interval in either \( H_o \) or \( H_c \) must be involved in any newly created conflicts.

First, we note that there cannot be a new type (iii) conflict caused by the swap, as this would imply \( o' \in N_H^+(\ell) \) such that \( I_0' \ll I_0 \) in the original representation, which is a forbidden configuration. The same also holds true for any type (i) or type (ii) conflict with some \( x \in N_H^0(\ell) \) such that \( I_x \in H_c \) due to the forbidden patterns forcing \( I_x \cap I_c = \emptyset \).

This leaves the case where we have a type (i) or type (ii) conflict with some \( x \in N_H^0(\ell) \) such that \( I_x \) is between \( H_o \) and \( H_c \). However, note that this implies that in the interval representation before the swap, it must have been the case that \( I_x \ll I_c \), which implies there was a type (ii) conflict that was not yet addressed. Since all
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

type (i) and type (ii) conflicts must be resolved by this point in this algorithm, this is impossible.

From all this, we see that the swap does not create any type (i) or type (ii) conflicts, and the number of type (iii) conflicts must have decreased by at least one.

We conclude that since each swap we make decreases the number of conflicts by at least one, the algorithm eventually eliminates all conflicts and terminate.

3.2.4 Runtime Analysis

We show that the type Ib interval digraph recognition algorithm we have described runs in polynomial time.

We first check whether the underlying structure of the input solid/dashed digraph \( G = (V, E) \) is an interval graph. This can be done in linear time using Booth and Leuker’s algorithm from [5].

We next check \( G \) for forbidden subgraphs. Consider each subset \( S \subseteq V \) with \( |S| = 3 \). For each forbidden subgraph \( H \) on 3 vertices in Figure 3.3 and each possible arrangement of \( S \), check whether the induced subgraph on \( S \) matches \( H \). Similarly, consider each subset \( S \subseteq V \) with \( |S| = 4 \) and check for the forbidden subgraphs on 4 vertices. This can be done in constant time for each subset, so overall this takes \( O(n^4) \) time.

In the loop, the indices for \( \ell \) can be set in constant time. The majority of time
in a loop iteration comes from checking for conflicting patterns. Since there are at least 2 vertices involved in any conflicting pattern, there are at most \( \binom{n}{2} \) conflicting patterns to correct in any iteration of the algorithm. For a particular conflicting pattern, the algorithm adjusts the indices of at most \( n \) intervals. Thus the process of correcting conflicting patterns for a single iteration takes \( O(n^3) \) time. Since there are \( n \) iterations, overall the loop takes \( O(n^4) \) time.

Looking at all steps together, the type Ib interval digraph recognition algorithm runs in \( O(n^4) \) time. Note that this analysis is using a brute force approach to checking for forbidden subgraphs, and we are almost certainly overestimating the actual amount of work that must be done in each iteration, so it may be possible to obtain a better asymptotic runtime with a more precise analysis or more efficient implementation.

### 3.2.5 Relation to Type I Interval Digraphs

We have an algorithm that successfully identifies whether a solid/dashed directed graph is a type Ib interval digraph. Our original goal, however, was to identify type I interval digraphs; we demonstrate how to use the type Ib recognition algorithm to achieve this goal.

Let \( G = (V, E) \) be a directed graph. Ideally, we would like to be able to convert \( G \) to a solid/dashed directed graph \( G' = (V, E') \) such that \( G' \) is a type Ib interval digraph if and only if \( G \) is a type I interval digraph. Intuitively, it makes sense
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

to create $G'$ with the same vertex set as $G$, all single edges of $G$ maintained, and double edges of $G$ converted to dashed edges. In the case where $G$ is a type I interval digraph, it also seems intuitive that we should map $G$ to the $G'$ with the same interval representation.

The primary challenge here is determining which way to orient the dashed edges that replace double edges in the original $G$. One option is to simply try all possible orientations of dashed edges. In this case, if there is some orientation that yields a type Ib interval representation, we know that $G$ is a type I interval digraph. Otherwise, no orientation of $G$’s double edges yields a type Ib interval representation and we conclude $G$ is not a type I interval digraph. The downside to this approach is that we may end up needing to run the type Ib recognition algorithm $2^d$ times where $d$ is the number of double edges in the graph. Ideally we would like to find some other method that yields the correct result without the need to run a potentially exponential number of tests.

Let $G = (V, E)$ be a directed graph and $u, v \in V$ be vertices such that $uv \in E$ and $vu \in E$ (i.e. there is a double edge between $u$ and $v$). We say a vertex $w \in V \setminus \{u, v\}$ is a distinguisher for the double edge $uv$ if $vw$ or $wv$ (or both) exist, but neither $uw$ nor $wu$ exist. If there exists a distinguisher for double edge $uv$, we say it is distinguished, otherwise we say it is undistinguished. Suppose we have a directed graph $G$ such that the double edge $uv$ is distinguished. If there exists a type I interval representation for $G$, it must be the case that $I_u \subseteq I_v$ since $w$ is adjacent
to \( v \) but not to \( u \). This implies the dashed edge between \( u \) and \( v \) in \( G' \) must be oriented from \( v \) to \( u \), which reduces the number of possible orientations we must check.

We have the following result:

**Proposition 3.7.** Let \( G = (V, E) \) be a random type I interval digraph with \( n \) vertices. With high probability, every double edge \( uv \) is distinguished.

**Proof.** Let \( I_v = [a, b] \) and \( I_u = [c, d] \) with \( a < c < d < b \). For \( w \in V \setminus \{u,v\} \), the probability that \( w \) is not a distinguisher for double edge \( uv \) given \( I_v = [a, b] \) and \( I_u = [c, d] \) is

\[
p = 1 - 2a(c - a) - 2(1 - b)(b - d) - (c - a)^2 - (b - d)^2
\]

\[
\leq 1 - a(c - a) - (1 - b)(b - d) - (c - a)^2 - (b - d)^2
\]

\[
\leq 1 - c(c - a) - (1 - d)(b - d)
\]

\[
\leq 1 - (c - a)^2 - (b - d)^2
\]

Observe that for two vertices \( w, w' \in V \), \( w \) being a distinguisher for \( uv \) is independent from \( w' \) being a distinguisher for \( uv \). Thus,

\[
P(uv \text{ is undistinguished} | I_v = [a, b], I_u = [c, d]) = \leq [1 - (c - a)^2 - (b - d)^2]^{n-2}.
\]

Integrating, we have

\[
P(uv \text{ is undistinguished}) \leq 4 \int_{b=0}^{1} \int_{d=0}^{b} \int_{c=0}^{d} \int_{a=0}^{c} [1 - (c - a)^2 - (b - d)^2]^{n-2} da \, dc \, dd \, db.
\]
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

We approximate this integral by breaking into cases based on the value of \( c - a \) and \( b - d \) relative to \( \frac{1}{\log n} \):

- If \( c - a < \frac{1}{\log n} \) and \( b - d < \frac{1}{\log n} \), 1 is an upper bound for the integrand.
- If \( c - a < \frac{1}{\log n} \) and \( b - d > \frac{1}{\log n} \), \( \left(1 - \frac{1}{\log^2 n}\right)^{n-2} \) is an upper bound for the integrand.
- If \( c - a > \frac{1}{\log n} \) and \( b - d < \frac{1}{\log n} \), \( \left(1 - \frac{2}{\log^2 n}\right)^{n-2} \) is an upper bound for the integrand.
- If \( c - a > \frac{1}{\log n} \) and \( b - d > \frac{1}{\log n} \), \( \left(1 - \frac{2}{\log^2 n}\right)^{n-2} \) is an upper bound for the integrand.

Using these bounds, we see the probability is bounded above by

\[
4 \left[ \left( \frac{1}{\log n} \right)^2 + 2 \left( \frac{1}{\log n} \right) \left(1 - \frac{1}{\log n}\right) \left(1 - \frac{1}{\log^2 n}\right)^{n-2} + \left(1 - \frac{1}{\log n}\right)^2 \left(1 - \frac{2}{\log^2 n}\right)^{n-2} \right]
\]

which goes to 0 as \( n \) goes to infinity. Thus, with high probability edge \( uv \) is distinguished.

Define \( X_{uv} \) to be the indicator random variable such that \( X_{uv} = 1 \) when \( uv \) is a double edge that is undistinguished. Define \( X = \sum_{u,v \in V} X_{uv} \). Then by Markov’s inequality,

\[
P(X > 0) \leq E[X] = \binom{n}{2} \cdot P(X_{uv} = 1) \to 0
\]

as \( n \) goes to infinity. Thus, with high probability every double edge \( uv \) in a random type I interval digraph is distinguished.
CHAPTER 3. RECOGNIZING TYPE I INTERVAL DIGRAPHS

One possible algorithm for type I interval digraph recognition is to take as input a directed graph, convert all distinguished double edges to the appropriate dashed edges, then check every possible orientation of the remaining double edges. As noted before, this algorithm still may not run in polynomial time since there could be exponentially many orientations to check. It is interesting to note, however, that for most type I interval digraphs this algorithm finishes after checking a single orientation.
Chapter 4

Properties of Random Interval Digraphs

Now that we have established a foundation about each type of interval digraph, we move on to looking at random interval digraphs of both types. Throughout this chapter, we use the notation \( f \sim g \) where \( f \) and \( g \) are functions of \( n \) to indicate \( f \) and \( g \) are asymptotically equivalent, that is

\[
\lim_{n \to \infty} \frac{f(n)}{g(n)} = 1.
\]

4.1 Edges in Random Interval Digraphs

One of the most basic random graph properties is the edge probability. In a random undirected interval graph with \( u, v \in V \), the probability of an edge \( uv \) existing
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

is $\frac{2}{3}$, determined by observing that the presence of an edge depends solely on the order of the endpoints of the intervals corresponding to $u$ and $v$. In [12], Scheinerman showed that with high probability a random undirected interval graph on $n$ vertices has $n^2/3 + o(n^2)$ edges, in contrast to Erdős-Rényi random graphs with edge probability $\frac{2}{3}$, which have $2n^2/3 + o(n^2)$ edges. We show similar results for both types of interval digraphs, appealing to the following result by Hoeffding (see [23] and [20]):

**Theorem 4.1.** Let $1 \leq k \leq n$ be integers and let $X_1, \ldots, X_n$ be independent, identically distributed random variables taking values in $\mathcal{X}$. Let $h : \mathcal{X}^k \to \mathbb{R}$ be a symmetric $k$-fold function and suppose $h$ is bounded, that is there exist $a, b \in \mathbb{R}$ such that for all $x_1, x_2, \ldots, x_k \in \mathcal{X}$, $a \leq h(x_1, x_2, \ldots, x_k) \leq b$. Let

$$U_n = \frac{1}{\binom{n}{k}} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} h(X_{i_1}, X_{i_2}, \ldots, X_{i_k}) \quad \text{and} \quad \mu = E[U_n].$$

Then for $t > 0$ we have

$$P[|U_n - \mu| \geq t] \leq 2 \exp \left\{ -\frac{2t^2}{b-a} \right\}.$$

We say a function $h(x_1, \ldots, x_k)$ is symmetric if for any permutation $\pi : [k] \to [k]$, we have $h(x_1, x_2, \ldots, x_k) = h(x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(k)})$.

### 4.1.1 Type I Interval Digraphs

In a random type I interval digraph, the edge probabilities are a bit more complicated than the undirected case.
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

**Proposition 4.2.** Let $G = (V, E)$ be a random type I interval digraph. For $u, v \in V$, the probability that $uv \in E$ is $\frac{1}{2}$.

To be precise, the probability that edge $uv$ exists (but not $vu$) is $\frac{1}{6}$, the probability that edge $vu$ exists (but not $uv$) is $\frac{1}{6}$, the probability that both edges $uv$ and $vu$ exist is $\frac{1}{3}$, and the probability that $u$ and $v$ are not adjacent is $\frac{1}{3}$. The expected number of edges in a random type I interval digraph is thus $\frac{1}{2}n(n - 1) \sim \frac{1}{2}n^2$.

**Proposition 4.3.** With high probability, the number of edges in a random type I interval digraph on $n$ vertices is $n^2/2 + o(n^2)$.

**Proof.** For real intervals $I$ and $J$, define the function

\[
h(I, J) = \begin{cases} 
0 & \text{if } I \cap J = \emptyset \\
1 & \text{if } I < J \text{ or } J < I \\
2 & \text{if } I \subseteq J \text{ or } J \subseteq I
\end{cases}
\]

Note that $h$ is a symmetric, bounded, 2-fold kernel. Consider a random type I interval digraph on $n$ vertices. Suppose the intervals in the interval representation are $I_1, I_2, \ldots, I_n$. Then

\[
U_n = \frac{1}{\binom{n}{2}} \sum_{i<j} h(I_i, I_j)
\]

is the number of edges in the graph divided by $\binom{n}{2}$. We compute $\mu = 1$ based on the edge probabilities for a type I interval digraph. By Theorem 4.1

\[
P[|U_n - 1| \geq t] \leq 2 \exp \left\{ -\frac{2|n/2|t^2}{4} \right\}.
\]
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

Let $Y$ denote the number of edges in a random type I interval digraph. The above inequality implies

$$P \left[ |Y - E(Y)| \geq t \left( \frac{n}{2} \right) \right] \leq 2 \exp \left\{ -\frac{|n/2| t^2}{2} \right\}.$$  

Let $\delta = t \left( \frac{n}{2} \right)$ (so $t \sim 2\delta/n^2$). Substituting in yields

$$P \left[ |Y - E(Y)| \geq \delta \right] \leq 2 \exp \left\{ -c\delta^2 \right\}.$$  

for a positive constant $c$. Setting $\delta = n^{7/4}$, we see that with high probability, the number of edges in a random type I interval digraph is $\binom{n}{2} + o(n^2)$.

Based on the derived edge probability, we compare and contrast the properties of random type I interval digraphs with Erdős-Rényi digraphs with edge probabilities $\frac{1}{2}$ throughout this chapter.

4.1.2 Type II Interval Digraphs

We make a similar argument for type II interval digraphs.

**Proposition 4.4.** Let $G = (V, E)$ be a random type II interval digraph. For $u, v \in V$, the probability that $uv \in E$ is $\frac{2}{3}$.

The logic is very similar to the undirected interval graph case; for type II interval digraphs we simply consider the send interval corresponding to $u$ and the receive interval corresponding to $v$. The expected number of edges in a random type II interval digraph on $n$ vertices is $\frac{2}{3}n(n-1) \sim \frac{2}{3}n^2$. 

58
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

**Proposition 4.5.** With high probability, the number of edges in a random type II interval digraph on \( n \) vertices is \( \frac{2}{3} n^2 + o(n^2) \).

**Proof.** Let \( I_s, I_r, J_s, J_r \) be real intervals and let \( I = (I_s, I_r), J = (J_s, J_r) \). Define the function

\[
h(I, J) = \begin{cases} 
0 & \text{if } I_s \cap J_r = \emptyset \text{ and } I_r \cap J_s = \emptyset \\
2 & \text{if } I_s \cap J_r \neq \emptyset \text{ and } I_r \cap J_s \neq \emptyset \\
1 & \text{otherwise}
\end{cases}
\]

Note that \( h \) is a symmetric, bounded, 2-fold kernel. Consider a random type II interval digraph on \( n \) vertices. For \( 1 \leq i \leq n \), define \( I_i = (R_i, S_i) \) where \( R_i \) is the receive interval associated with the \( i \)th vertex and \( S_i \) is the send interval associated with the \( i \)th vertex. Then

\[
U_n = \frac{1}{\binom{n}{2}} \sum_{i < j} h(I_i, I_j)
\]

is the number of edges in the graph divided by \( \binom{n}{2} \). We compute \( \mu = \frac{4}{3} \) based on the edge probabilities for a type II interval digraph. The remainder of the argument is identical to the type I interval digraph case. By Theorem 4.1

\[
P \left[ \left| U_n - \frac{4}{3} \right| \geq t \right] \leq 2 \exp \left\{ -\frac{2|n/2|^2}{4} \right\}.
\]

Let \( Y \) denote the number of edges in a random type II interval digraph and \( \delta = t \binom{n}{2} \) (so \( t \sim 2\delta / n^2 \)). Substituting in yields

\[
P \left[ |Y - E(Y)| \geq \delta \right] \leq 2 \exp \left\{ -\frac{c\delta^2}{n^3} \right\}.
\]
for a positive constant $c$. Setting $\delta = n^{7/4}$, we see that with high probability, the number of edges in a random type II interval digraph is $\frac{2}{3}n^2 + o(n^2)$. \hfill\square

Based on the derived edge probability, we compare and contrast the properties of random type I interval digraphs with Erdős-Rényi digraphs with edge probabilities $\frac{2}{3}$ throughout this chapter.

### 4.2 Degree in Random Interval Digraphs

Knowing the number of edges in a random interval digraph gives us a general idea of the structure of the graph. We look at the in-degree and out-degree of individual vertices to obtain a more nuanced view. We begin by examining the extreme cases of in-degree or out-degree 0 or $n - 1$, then look at the distribution of degrees.

#### 4.2.1 Type I Interval Digraphs

**Theorem 4.6.** The probability that a random type I interval digraph on $n$ vertices has a vertex with out-degree $n - 1$ is

$$
\sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} \frac{1}{2k+1}.
$$

**Proof.** Observe that when we are given a set of $n$ intervals, if any vertex in the corresponding type I interval digraph has out-degree $n - 1$, then $v_\ell$ must have
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

out-degree $n - 1$. If we have some vertex other than $v_\ell$ (say $u$) that has out-degree $n - 1$, its interval must be entirely contained in the interval corresponding to $v_\ell$, as that is the only way it could possibly have an edge to $v_\ell$.

Every interval contained in the interval corresponding to $u$ will also be contained in the interval corresponding to $v_\ell$ and thus there will be an edge from $v_\ell$ to the corresponding vertex. Any interval containing the interval of $u$ will have its left endpoint contained in $v_\ell$, and thus there will be an edge from $v_\ell$ to the corresponding vertex. Lastly, if there is an interval whose left endpoint is contained in the interval corresponding to $u$ (but whose right endpoint is not contained in $u$), this interval’s left endpoint will also be in the interval of $v_\ell$ and thus there will be an edge from $v_\ell$ to the corresponding vertex. Thus is $u$ has out-degree $n - 1$, $v_\ell$ must as well. As a result we focus on counting interval representations such that $v_\ell$ has out-degree $n - 1$.

Consider a permutation $a_1, b_1, \ldots, a_n, b_n$ selected uniformly at random from all permutations of $[2n]$. There are $(2n)!$ permutations total. Construct intervals $I_{v_1} = [a_1, b_1], I_{v_2} = [a_2, b_2], \ldots, I_{v_n} = [a_n, b_n]$.

We now count permutations such that at least $k$ intervals do not overlap with the interval $I_{v_\ell}$ (the interval with the leftmost endpoint). There are $n$ possibilities for which vertex $v_i$ from $\{v_1, \ldots, v_n\}$ is $v_\ell$, and there are 2 possibilities for which endpoint of $I_{v_i}$ is the leftmost endpoint (that is, either $a_i = 1$ or $b_i = 1$). There are then $\binom{n-1}{k}$ ways to pick $k$ of the other $n - 1$ intervals which are disjoint from $I_{v_\ell}$.
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

Once we have selected these \( k \) intervals, we select \( 2k + 1 \) positions in the permutation: one for the other endpoint of \( I_{v_i} \) and one for each of the two endpoints of the \( k \) intervals we selected to be disjoint from \( I_{v_i} \). There are \( 2n - 1 \) possible open positions in the permutation (since we have already assigned the leftmost endpoint). Thus there are \( \binom{2n-1}{2k+1} \) possible selections of \( 2k + 1 \) positions. In order for \( I_{v_i} \) to be disjoint from the selected \( k \) intervals, we know the leftmost of these \( 2k + 1 \) selected positions must be the second endpoint of \( I_{v_i} \). There are then \( (2k)! \) ways to distribute the endpoints of the \( k \) intervals among the remaining \( 2k \) selected positions. Finally, there are \( (2n - 2k - 2)! \) ways to arrange all the remaining endpoints to positions in the permutation. Overall, this gives us

\[
2n \cdot (2k)! \cdot (2n - 2k - 2)! \cdot \binom{n-1}{k} \binom{2n-1}{2k+1} = \binom{n-1}{k} \frac{(2n)!}{2k+1}
\]

permutations such that at least \( k \) intervals are disjoint from \( I_{v_i} \). Applying the Principle of Inclusion-Exclusion for possible values of \( k \) from 0 to \( n - 1 \), we conclude that

\[
P(\exists \text{ a vertex with out-degree } n-1) = \frac{\sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} \frac{(2n)!}{2k+1}}{(2n)!}
\]

\[
= \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} \frac{1}{2k+1}.
\]

\( \square \)

Letting \( n \) go to infinity, we have the following asymptotic result.
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

Corollary 4.7. The probability that a random type I interval digraph on $n$ vertices has a vertex of out-degree $n - 1$ is asymptotically equivalent to

$$\frac{1}{2} \sqrt{\frac{\pi e}{n}}.$$

Thus, with high probability a random type I interval digraph on $n$ vertices has no vertex of out-degree $n - 1$.

Proof. We find from Mathematica that the sum in the previous theorem evaluates to

$$\frac{\sqrt{\pi}(n - 1)!}{2(n - \frac{1}{2})!}.$$

By Stirling’s formula, we know that $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$, so we have

$$\frac{\sqrt{\pi}(n - 1)!}{2(n - \frac{1}{2})!} \sim \frac{\sqrt{\pi} \sqrt{2\pi(n - 1)(n - 1)^{n-1}e^{-(n-1)}}}{2\sqrt{2\pi(n - \frac{1}{2})(n - \frac{1}{2})^{n-\frac{1}{2}}e^{-(n-\frac{1}{2})}}}
\sim \frac{\sqrt{\pi e}}{2} \left(\frac{n - 1}{n - \frac{1}{2}}\right)^{n-\frac{1}{2}} (n - 1)^{-\frac{1}{2}}
\sim \frac{1}{2} \frac{\sqrt{\pi e}}{n} \left(1 - \frac{1}{n - \frac{1}{2}}\right)^{n-\frac{1}{2}}
\sim \frac{1}{2} \frac{\sqrt{\pi e}}{n},$$

which we see goes to zero as $n$ goes to infinity. Thus we conclude that with high probability, a type I interval digraph has no vertex of out-degree $n - 1$ as desired.

We also have the following corollary by simply altering the argument to focus on $v_r$ rather than $v_{\ell}$.
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

**Corollary 4.8.** The probability that a random type I interval digraph on \( n \) vertices has a vertex with in-degree \( n - 1 \) is

\[
\sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} \frac{1}{2k+1} \sim \frac{1}{2} \sqrt{\frac{\pi e}{n}},
\]

and with high probability, a random type I interval digraph on \( n \) vertices has no vertex of in-degree \( n - 1 \).

We use a similar argument to the proof of Theorem 4.6 to examine vertices with out-degree 0. In this case we only have an asymptotic result rather than an exact probability due to some additional assumptions needed in the proof.

**Theorem 4.9.** With high probability, a random type I interval digraph on \( n \) vertices has no vertex with out-degree 0.

*Proof.* Given a type I interval representation of a digraph, observe that in order for a vertex to have out-degree 0, the right endpoint of the corresponding interval in the interval representation must not be contained in any other interval. In the case that there are no breaks in the interval representation (i.e. every right endpoint of an interval is contained in some other interval save for the right endpoint of \( I_v \)), this means we only need to check whether \( v_r \) has out-degree 0.

Next, observe that we have no breaks in the interval representation if and only if the corresponding undirected interval graph is connected. We have

\[
P(\exists v \text{ with } d^+(v) = 0) \leq P(\text{connected and } \exists v \text{ with } d^+(v) = 0) + P(\text{disconnected})
\]

\[
\leq P(d^+(v_r) = 0) + P(\text{disconnected}).
\]
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

We next compute $P(d^+(v_r) = 0)$. The logic here is essentially identical to the previous proof, except now we look at the case that at least $k$ intervals are contained within $I_{v_r}$ (resulting in the out-degree of $v_r$ being nonzero).

As before, we consider a permutation $a_1, b_1, \ldots, a_n, b_n$ selected uniformly at random from all permutations of $[2n]$. We consider the intervals to be $[a_1, b_1], [a_2, b_2], \ldots, [a_n, b_n]$. There are $(2n)!$ such permutations total.

We count permutations such that at least $k$ intervals are contained within $I_{v_r}$. There are $2n$ possibilities for which point is the rightmost endpoint. There are then $\binom{n-1}{k}$ ways to pick $k$ of the other $n - 1$ intervals which are contained in $I_{v_r}$. Once we have selected these $k$ intervals, we select $2k + 1$ positions in the permutation to distribute these endpoints and the left endpoint of $I_{v_r}$. There are $2n - 1$ possible open positions in the permutation, so there are $\binom{2n-1}{2k+1}$ possible selections of $2k + 1$ positions. In order for $I_{v_r}$ to contain each of the selected $k$ intervals, we know the leftmost of these $2k + 1$ selected positions must be the second endpoint of $I_{v_r}$. There are then $(2k)!$ ways to distribute the endpoints of the $k$ intervals among the remaining $2k$ selected positions. Finally, there are $(2n - 2k - 2)!$ ways to arrange all the remaining endpoints to positions in the permutation. Overall, this gives

\[
2n \cdot (2k)! \cdot (2n - 2k - 2)! \cdot \binom{n-1}{k} \binom{2n-1}{2k+1} = \binom{n-1}{k} \frac{(2n)!}{2k+1}
\]

permutations such that at least $k$ intervals are contained in $I_{v_r}$. Applying the Principle of Inclusion-Exclusion for possible values of $k$ from 0 to $n - 1$, we conclude
that

\[ P(d^+(v_r) = 0) = \frac{\sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} \frac{(2n)!}{2k+1}}{(2n)!} = \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} \frac{1}{2k+1}. \]

We know from Corollary 4.7 that as \( n \) goes to infinity, this expression goes to zero. Additionally, we know from [12] that almost all interval graphs are Hamiltonian, implying that almost all interval graphs are connected. Thus as \( n \) goes to infinity, the probability of us encountering a disconnected interval representation goes to 0. We conclude that then it must be the case that the probability of having a vertex of out-degree 0 in an interval digraph of type I goes to 0 as \( n \) goes to infinity, and thus with high probability a type I interval digraph has no vertex with out-degree 0.

These results for vertices with out-degree/in-degree \( n - 1 \) or 0 agree with the results we typically expect from a random Erdős-Rényi digraph with edge probability \( \frac{1}{2} \); with high probability, an Erdős-Rényi digraph with edge probability \( \frac{1}{2} \) has no vertex with out-degree/in-degree \( n - 1 \) or 0.

We also look more generally at the distribution of out-degrees and in-degrees of vertices in a random type I interval graph. We use the following Chernoff bound:

**Proposition 4.10** (Chernoff Bound). Let \( n \) be a positive integer, \( X_1, \ldots, X_n \) be iid indicator random variables with \( P(X_i = 1) = p \), \( X = \sum_{i=1}^{n} X_i \), \( \mu = np \), and \( 0 < \delta < 1 \).
Then
\[ P[|X - \mu| \geq \delta \mu] \leq 2 \exp \left[-\frac{\mu \delta^2}{3}\right]. \]

We model the argument after the proof of Theorem 4.2 in [12]. The result is as follows:

**Theorem 4.11.** Let \( G = (V, E) \) be a type I interval digraph with \( v \in V \). For a fixed \( \alpha \in [0, 1] \) we have
\[
\lim_{n \to \infty} P(d^+(v) \leq \alpha n) = \frac{1}{4} \left(2 - 2 \sqrt{1 - \alpha} + 2 \sqrt{\alpha} + \sqrt{2}(1 - 2\alpha) \log \left[\frac{1 - \sqrt{2} - 2\alpha}{1 - \sqrt{2} + (-2 + \sqrt{2})\sqrt{\alpha}}\right]\right)
\]
if \( \alpha \leq \frac{1}{2} \) and
\[
\lim_{n \to \infty} P(d^+(v) \leq \alpha n) = 1 - \frac{1}{4} \left(-4 + 2 \left((1 - 2\sqrt{1 - \alpha}) \sqrt{2 - 2\sqrt{1 - \alpha}} - \alpha + \sqrt{\alpha}\right) + 8\sqrt{1 - \alpha} - 4\sqrt{\alpha} + 4\alpha + \sqrt{2}(-1 + 2\alpha) \log(1 - 2\sqrt{1 - \alpha} + \sqrt{2}\sqrt{2 - 2\sqrt{1 - \alpha}} - \alpha) - \sqrt{2}(-1 + 2\alpha) \log(1 + \sqrt{2}\alpha)\right).
\]
if \( \alpha > \frac{1}{2} \). Note that when \( \alpha = \frac{1}{2} \), both of these values equal \( \frac{1}{2} \).

**Proof.** Consider a vertex \( v \) in the graph. We define indicator random variables \( X_u \) for each \( u \in V \) with \( u \neq v \) such that \( X_u = 1 \) if \( uv \in E \). We define \( X = \sum_{u \in V \setminus \{v\}} X_u \); \( X \) is the degree of vertex \( v \).

Suppose \( I_v = [a, b] \). Observe that with the \( I_v \) fixed, the \( X_u \) are independent. We have
\[
p = P(X_u = 1) = 1 - a^2 - (1 - b)^2 - 2ab - a = a^2 + 2b - 2ab - b^2.
\]
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

Then $E[X] = (n - 1)p$. Applying the Chernoff bound in Proposition 4.10 yields

$$P \left( \left| X - (n - 1)p \right| \geq \sqrt{3(n - 1) \log n} \right) \leq 2 \exp \left( \frac{-3(n - 1) \log n}{3(n - 1)p} \right) = 2n^{-1/p} \ll \frac{1}{n}$$

since $0 < p < 1$. Thus with high probability we have $d^+(v) = np + o(n)$ for a fixed value of $p$ (say $p = r$). Observe that if we have some $\epsilon > 0$ such that $\alpha > r + \epsilon$, then with high probability $d^+(v) = rn + o(n) < (r + \epsilon)n + o(n) < an$. On the other hand, if we have some $\epsilon > 0$ such that $\alpha < r - \epsilon$, then with high probability $d^+(v) = rn + o(n) > (r - \epsilon)n + o(n) > an$. That is, for any $\epsilon > 0$,

$$P(d^+(v) \leq an | p = r) = \begin{cases} 1 - o(1) & \text{if } r < \alpha - \epsilon \\ o(1) & \text{if } r > \alpha + \epsilon \end{cases}$$

Now we see

$$P(d^+(v) \leq an) = \int_0^1 P(d^+(v) \leq an | p = r) dP(p \leq r)$$

$$= [1 - o(1)]P(p \leq \alpha) + o(1) \to P(p \leq \alpha).$$

We know $a$ and $b$ may take values between 0 and 1. When $\alpha \leq \frac{1}{2}$, we compute

$$P(a^2 + 2b - 2ab - b^2 \leq \alpha) = \int_{a=1-\sqrt{\alpha}}^{1} \left[ 1 - (\sqrt{\alpha + 2a^2 - 2a + 1} + a - 1) \right] da$$

$$+ \int_{a=0}^{\sqrt{\alpha}} \left[ -\sqrt{\alpha + 2a^2 - 2a + 1} - a + 1 \right] da$$

yielding

$$P(d^+(v) \leq an) = \frac{1}{4} \left( 2 - 2\sqrt{1-\alpha} + 2\sqrt{\alpha} + \sqrt{2}(1-2\alpha) \log \left[ \frac{1 - \sqrt{2 - 2\alpha}}{1 - \sqrt{2 + (-2 + \sqrt{2})\sqrt{\alpha}}} \right] \right).$$

68
When $\alpha > \frac{1}{2}$, we have

\[
P(a^2 + 2b - 2ab - b^2 \leq \alpha) = \int_{b=1-\sqrt{1-\alpha}}^1 \left[ b - \sqrt{\alpha + 2b^2 - 2b} \right] \, db \\
+ \int_{b=0}^{\sqrt{1-\alpha}} \left[ 1 - (b + \sqrt{\alpha + 2b^2 - 2b}) \right] \, db
\]

giving

\[
P(d^+(v) \leq \alpha n) = 1 - \frac{1}{4} \left( -4 + 2 \left( (1 - 2\sqrt{1-\alpha})\sqrt{2 - 2\sqrt{1-\alpha} - \alpha} + \sqrt{\alpha} \right) + 8\sqrt{1-\alpha} \\
- 4\sqrt{\alpha} + 4\alpha + \sqrt{2}(-1 + 2\alpha) \log(1 - 2\sqrt{1-\alpha} + \sqrt{2}\sqrt{2 - 2\sqrt{1-\alpha} - \alpha}) \\
- \sqrt{2}(-1 + 2\alpha) \log(1 + \sqrt{2\alpha}) \right).
\]

Finally, when $\alpha = \frac{1}{2}$, $P(p \leq \frac{1}{2}) = \frac{1}{2}$. \hfill \Box

We conclude the following:

**Corollary 4.12.** Let $G = (V, E)$ be a type I interval digraph. With high probability, the number of vertices $v \in V$ such that $d^+(v) \leq \alpha n$ is $P(d^+(v) \leq \alpha)n + o(n)$ (where $P(d^+(v) \leq \alpha)$ is given by the function in Theorem 4.11).

### 4.2.2 Type II Interval Digraphs

Similar to the analysis of type I interval digraphs, we examine probability of a vertex with out-degree $n$ in a type II interval digraph. In the case of a random undirected interval graph, [13] showed that the probability of having a vertex of degree $n - 1$ is exactly $\frac{2}{3}$ regardless of the value of $n$. This immediately implies
that the probability of a random type II interval digraph having a vertex with out-degree $n$ must be bounded away from zero, contrasting with the result for type I interval digraphs as well as with the usual Erdős-Rényi results.

**Proposition 4.13.** The probability a type II interval digraph has a vertex with out-degree $n$ is

\[
1 - 4n(n-1) \int_{y=0}^{1} \int_{x=0}^{1-y} xy(1-x^2-y^2)^{n-2}(1-2xy)^n dx dy \\
- 2n^2 \int_{y=0}^{1} \int_{x=1-y}^{1} (2(1-x)(1-y))^{n-1}((1-x)^2 + (1-y)^2)^n dx dy
\]

**Proof.** Consider a random type II interval digraph selected using the continuous method (i.e., by selecting endpoints for the $2n$ intervals independently at random from $[0,1]$). Let $x$ be the value associated with the leftmost right endpoint of any receive interval and $1-y$ be the value associated with the rightmost left endpoint of any receive interval. Observe that in order for a vertex to have out-degree $n$, it must have a left endpoint in $[0,x]$ and a right endpoint in $[1-y,1]$. We consider two cases, in each determining the probability that there is no vertex with out-degree $n$:

- If $x \leq 1-y$, it must be the case that $x$ and $1-y$ are associated with the endpoints of different intervals. There are $n$ choices for which vertex $x$ is associated with and $n-1$ choices for which vertex $1-y$ is associated with. There are then 2 choices in each case for which endpoint $x$ (respectively $1-y$) is associated with. Since $x$ is a right endpoint, the range of values for the
corresponding left endpoint is \([0, x]\); similarly since \(1 - y\) is a left endpoint, the range of values for the corresponding right endpoint is \([1 - y, 1]\) which has length \(y\).

The probability that no send interval intersects all receive intervals is then \((1 - 2xy)^n\), and the probability that the remaining \(n - 2\) receive intervals do not have right endpoint less than \(x\) or left endpoint greater than \(1 - y\) is \((1 - x^2 - y^2)^{n-2}\). Integrating over all possible values of \(x\) and \(y\) (keeping in mind \(x \leq 1 - y\)) yields

\[
4n(n - 1) \int_{y=0}^{1} \int_{x=0}^{1-y} xy(1 - x^2 - y^2)^{n-2}(1 - 2xy)^n dx dy.
\]

- If \(x > 1 - y\) and there exists some vertex \(v\) such that \(L(R_v) = 1 - y\) and \(R(R_v) = x\), there are \(n\) choices for which vertex \(v\) and 2 choices for which endpoint is which. In order to avoid intersecting all receive intervals, each send interval \(S\) must satisfy either \(S \ll R_v\) or \(R_v \ll S\). The probability of this happening is \(((1 - x)^2 + (1 - y)^2)^n\). The probability that the remaining \(n - 1\) receive intervals do not have right endpoint less than \(x\) or left endpoint greater than \(1 - y\) is \((2(1 - y)(1 - x))^{n-1}\). Integrating over all possible values of \(x\) and \(y\) (keeping in mind \(x > 1 - y\)) yields

\[
2n \int_{y=0}^{1} \int_{x=1-y}^{1} (2(1 - x)(1 - y))^{n-1}((1 - x)^2 + (1 - y)^2)^n dx dy.
\]

- If \(x > 1 - y\) and there is no vertex \(v\) satisfying \(L(R_v) = 1 - y\) and \(R(R_v) = x\), there are \(n\) choices for which vertex \(x\) is associated with and \(n - 1\) choices for
which vertex $1 - y$ is associated with. There are then 2 choices in each case for which endpoint $x$ (respectively $1 - y$) is associated with. The range of values for the left endpoint corresponding to $x$ is $[0, 1 - y]$ to ensure $1 - y$ is the rightmost left endpoint; similarly the range of values for the right endpoint corresponding to $1 - y$ is $[x, 1]$ which has length $1 - x$.

The probability that no send interval intersects all receive intervals is then $((1 - x)^2 + (1 - y)^2)^n$, and the probability that the remaining $n - 2$ receive intervals do not have right endpoint less than $x$ or left endpoint greater than $1 - y$ is $(2(1 - x)(1 - y))^{n-2}$. Integrating over all possible values of $x$ and $y$ (keeping in mind $x > 1 - y$) yields

$$4n(n-1) \int_{y=0}^{1} \int_{x=1-y}^{1} (1-x)(1-y)(2(1-x)(1-y))^{n-2}((1-x)^2(1-y)^2)^n dx dy.$$  

Putting all of these cases together (noting that they are disjoint) gives a final probability of a random type II interval digraph having a vertex with out-degree $n$ of

$$1 - 4n(n-1) \int_{y=0}^{1} \int_{x=0}^{1-y} xy(1 - x^2 - y^2)^{n-2}(1 - 2xy)^n dx dy$$  

$$- 2n^2 \int_{y=0}^{1} \int_{x=1-y}^{1} (2(1-x)(1-y))^{n-1}((1-x)^2 + (1-y)^2)^n dx dy$$  

Unfortunately this integral does not evaluate as nicely as the undirected case, but we are still able to analyze the probability asymptotically.
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

**Proposition 4.14.** As \( n \to \infty \), the probability that a random type II interval digraph has a vertex with out-degree \( n \) is at most \( \frac{2}{3} \).

**Proof.** Observe that the latter two cases from the proof of Proposition 4.13 only occur if all left endpoints of receive intervals appear to the left of all right endpoint of receive intervals. Using the discrete model, we see this happens with probability \( \frac{n!^2}{(2n)!} \), which is asymptotically zero. Thus for any asymptotic analysis, we only need to worry about the first term of the probability formula:

\[
1 - 4n(n - 1) \int_0^1 \int_0^{1 - y} xy(1 - x^2 - y^2)^{n-2}(1 - 2xy)^n dx \, dy.
\]

We observe that

\[
4n(n - 1) \int_0^1 \int_0^{1 - y} xy(1 - x^2 - y^2)^{n-2}(1 - 2xy)^n dx \, dy = \int_0^1 \int_0^{1 - y} xy \left( (1 - x^2 - y^2)(1 - 2xy) \right)^{n-2} (1 - 2xy)^2 dx \, dy
\]

\[
= \int_0^1 \int_0^{1 - y} xy \left( 1 - x^2 - y^2 - 2xy + 2x^3y + 2xy^3 \right)^{n-2} (1 - 2xy)^2 dx \, dy
\]

\[
\geq 4n(n - 1) \int_0^1 \int_0^{1 - y} xy(1 - x^2 - y^2 - 2xy)^{n-2}(1 - 2xy)^2 dx \, dy
\]

\[
\geq 4n(n - 1) \int_0^1 \int_0^{1 - y} xy(1 - x^2 - y^2 - 2xy)^{n} dx \, dy
\]

\[
\to \frac{1}{3}
\]

as \( n \to \infty \). Thus asymptotically \( \frac{2}{3} \) is an upper bound on the actual probability. \( \square \)

We do not yet have a satisfactory lower bound on this probability asymptotically. The trivial lower bound is \( \frac{1}{3} \) based on the knowledge that with probability
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

some interval intersects all others. This interval is equally likely to be a send or receive interval, so $\frac{1}{3}$ is a lower bound. It is, however, possible to have both send and receive intervals intersecting all others. Additionally even in the case where there is no interval that intersects all others, there may be a vertex with out-degree $n$ provided there is a send interval that intersects all receive intervals. This second consideration especially seems like it could result in far more than $\frac{1}{3}$ of type II interval digraphs having vertices with out-degree $n$. And indeed, based on simulations it appears the actual value is much closer to $\frac{2}{3}$.

We also look at the other extreme, analyzing the probability of a vertex with out-degree 0.

**Proposition 4.15.** With high probability, a random type II interval digraph has no vertex with out-degree 0.

**Proof.** We begin by considering only the receive intervals. Notice we have an upper bound on the desired property by taking advantage of the following:

$$P(\exists v \text{ with } d^+(v) = 0) \leq P(\text{receive intervals connected and } \exists v \text{ with } d^+(v) = 0)$$

$$+ P(\text{receive intervals disconnected and } \exists v \text{ with } d^+(v) = 0)$$

From [12], we know that with high probability there are no breaks among these $n$ intervals (i.e. every right endpoint of an interval is contained in some other interval save for the rightmost endpoint), the probability that the receive intervals are disconnected goes to zero as $n$ goes to infinity. Thus we just need to consider
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

the case where the receive intervals are connected.

Observe that the results from Lemma 4.18 and Corollary 4.19 can be applied to the set of receive intervals for the random type II interval digraph. In particular, we know that with high probability the leftmost endpoint among receive intervals is less than $\frac{\log n}{n}$ and with high probability the rightmost endpoint among receive intervals is greater than $1 - \frac{\log n}{n}$.

Since we are considering the case where the receive intervals are connected, this implies every value in the interval $\left[\frac{\log n}{n}, 1 - \frac{\log n}{n}\right]$ is contained in some receive interval. Since the only way for a vertex to have out-degree zero is if its send interval intersects no receive interval, this implies such a send interval would have to be contained in either $[0, \frac{\log n}{n}]$ or $[1 - \frac{\log n}{n}, 1]$. However from the proof of Lemma 4.18 and Corollary 4.19 applied this time to just the set of send intervals, we know with high probability no such send interval exists. Thus we conclude that with high probability, when the receive intervals are connected, there is no vertex with $d^+(v) = 0$.

Since the probability of a vertex with out-degree zero goes to zero both when the receive intervals are connected (due to the argument above) and when the receive intervals are disconnected (since with high probability this is not the case), overall we conclude that with high probability, a random type II interval digraph has no vertex with out-degree 0.

As always, due to symmetry these results are also applicable to in-degree $n$ and
0 as well. While the out-degree/in-degree 0 case aligns with the expectations for Erdős-Rényi random digraphs, the out-degree/in-degree \( n \) case does not. However this is not too surprising given that it follows the same pattern seen in random undirected interval graphs.

### 4.3 Diameter of Random Interval Digraphs

Recall that we know from [12] that almost all undirected interval graphs are Hamiltonian, implying that almost all interval graphs are connected. In the directed case, we show that with high probability both types of random interval digraphs are strongly connected and furthermore we are able to analyze the expected diameter of each type of digraph.

#### 4.3.1 Type I Interval Digraphs

We begin by examining the length of paths from \( v_r \) to \( v_\ell \) (i.e., the vertex corresponding to the interval with rightmost endpoint to the vertex corresponding to the interval with leftmost endpoint in a given interval representation). Intuitively, we expect \( d(v_r, v_\ell) \) to be among the largest distances between pairs of vertices in the digraph due to it being more difficult based on the edge definitions to move “right-to-left” relative to the interval representation. We begin with the following observation:
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

Proposition 4.16. Consider a random type I interval digraph. The shortest path from $v_r$ to $v_\ell$ (assuming such a path exists) has even length.

Proof. If $v_r = v_\ell$, the shortest path is length 0 and the result is trivial. In any other case, suppose $v_r, v_1, v_2, \ldots, v_k, v_\ell$ is the shortest path from $v_r$ to $v_\ell$. Clearly for any path we must have $I_{v_r} \supseteq I_{v_1}$ and $I_{v_\ell} \supseteq I_{v_k}$. We claim for a shortest path that in addition for $i \in [k]$,

$$
\begin{cases}
  I_{v_i} \subseteq I_{v_{i+1}} & \text{if } i \text{ odd} \\
  I_{v_i} \supseteq I_{v_{i+1}} & \text{if } i \text{ even}
\end{cases}
$$

Suppose we have a path from $v_r$ to $v_\ell$ such that this is not the case. We show that there exists some shorter path. Let $v_j, v_{j+1}$ be the first pair of consecutive vertices that diverges from the pattern described above. We look at the possible intersections of the corresponding intervals:

- If $I_{v_j} \subseteq I_{v_{j+1}}$, since this is a deviation from the prescribed pattern we know $I_{v_{j-1}} \subseteq I_{v_j}$. But then $I_{v_{j-1}} \subseteq I_{v_{j+1}}$, so we could find a shorter path by skipping vertex $v_j$ and taking the edge directly from $v_{j-1}$ to $v_{j+1}$.

- If $I_{v_j} \supseteq I_{v_{j+1}}$, since this is a deviation from the prescribed pattern we know $I_{v_{j-1}} \supseteq I_{v_j}$. But then $I_{v_{j-1}} \supseteq I_{v_{j+1}}$, so we could find a shorter path by skipping vertex $v_j$ and taking the edge directly from $v_{j-1}$ to $v_{j+1}$.

- The final possibility (since we know there exists an edge from $v_j$ to $v_{j+1}$) is $I_{v_j} < I_{v_{j+1}}$. We break this down further into subcases:
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

- If $I_{v_{j-1}} \supseteq I_{v_j}$, we know either $I_{v_{j-1}} \supseteq I_{v_{j+1}}$ or $I_{v_{j-1}} < I_{v_{j+1}}$ (since both $I_{v_{j-1}}$ and $I_{v_{j+1}}$ contain $R(I_{v_j})$). Thus we could find a shorter path by skipping vertex $v_j$ and taking the edge directly from $v_{j-1}$ to $v_{j+1}$.

- If $I_{v_{j-1}} \subseteq I_{v_j}$, either $I_{v_{j-1}} \subseteq I_{v_{j+1}}$, $I_{v_{j-1}} < I_{v_{j+1}}$, or $I_{v_{j-1}} \ll I_{v_{j+1}}$. In the former two scenarios, we could find a shorter path by skipping vertex $v_j$ and taking the edge directly from $v_{j-1}$ to $v_{j+1}$.

The final case is $I_{v_{j-1}} \ll I_{v_{j+1}}$. Since we know $I_{v_{j-1}} \subseteq I_{v_j}$, $v_{j-1} \neq v_r$, so there is a vertex $v_{j-2}$ with $I_{v_{j-2}} \supseteq I_{v_{j-1}}$. We know if $I_{v_{j-2}} \subseteq I_{v_j}$, we could have a shorter path by taking the edge directly from $v_{j-2}$ to $v_j$ and omitting $v_{j-1}$.

Otherwise, $R(I_{v_j}) \in I_{v_{j-2}}$ and $R(I_{v_{j-1}}) \in I_{v_{j-2}}$, so since $R(I_{v_{j-1}}) < L(I_{v_{j+1}}) < R(I_{v_j})$, $L(I_{v_{j+1}}) \in I_{v_{j-2}}$. Then $I_{v_{j-2}} < I_{v_{j+1}}$ or $I_{v_{j-2}} \supseteq I_{v_{j+1}}$, so we could use the edge from $v_{j-2}$ to $v_{j+1}$ to avoid $v_{j-1}$ and $v_j$ and obtain a shorter path.

Since in any case where the prescribed pattern is not followed there exists a shorter path, we conclude that the shortest path from $v_r$ to $v_{\ell}$ must follow this subset/superset type pattern. Since $I_{v_{\ell}}$ and $I_{v_r}$ cannot be subsets of any other intervals, we know a shortest $(v_r, v_{\ell})$-path must have even length to ensure the alternating order. $\square$

The strict containment structure of intervals found in the above proof implies
that all edges in a minimum length \((v_r, v_\ell)\)-path are double edges. We can thus leverage this \((v_r, v_\ell)\)-path to build paths between arbitrary pairs of vertices, allowing us to bound the diameter of the entire digraph in terms of \(d(v_r, v_\ell)\).

**Proposition 4.17.** Consider a random type I interval digraph \(G\). If a shortest \((v_r, v_\ell)\)-path has length \(2\alpha\), we have

\[
2\alpha \leq \text{diam}(G) \leq 2\alpha + 2.
\]

**Proof.** Let \(P = (v_0, v_1, \ldots, v_{2\alpha-1}, v_{2\alpha})\) be a shortest \((v_r, v_\ell)\)-path (i.e., \(v_r = v_0\) and \(v_\ell = v_{2\alpha-1}\)). Given any \(u, w \in V\), observe that it must be the case that \(I_u \subseteq I_{v_{2\alpha'}}\) or \(I_{v_{2\alpha'+1}} \subseteq I_u\) for some \(0 \leq k' \leq \alpha\). Similarly it must be the case that \(I_w \subseteq I_{v_{2j'}}\) or \(I_{v_{2j'+1}} \subseteq I_w\) for some \(0 \leq j' \leq \alpha\). In particular, this implies there exist \(0 \leq k, j \leq 2\alpha\) such that \(uv_k\) is an edge and \(v_jw\) is an edge.

Now we look at two different cases: if \(k \leq j\), observe that \(u, v_k, v_{k+1}, \ldots, v_{j-1}, v_j, w\) is a \((u, w)\)-path of length \(j - k + 2 \leq 2\alpha + 2\). If \(k > j\), we have \(u, v_k, v_{k-1}, \ldots, v_{j+1}, v_j, w\) is a \((u, w)\)-path of length \(k - j + 2 \leq 2\alpha + 2\). Thus for any two vertices \(u, w \in V\) we have \(d(u, w) \leq 2\alpha + 2\), giving the desired result. \(\square\)

We next narrow the range of likely diameters by determining the probability of different values of \(d(v_r, v_\ell)\). To accomplish this, we first establish some results involving the endpoints of \(I_{v_\ell}\) and \(I_{v_r}\). We also prove the existence of other intervals with high probability which gives us the intermediate vertices in a \((v_r, v_\ell)\)-path, limiting the distance from \(v_r\) to \(v_\ell\).
Lemma 4.18. Let $G$ be a random type I interval digraph with $n$ vertices such that interval endpoints are selected independently from $[0, 1]$. With high probability $L(I_v) < \frac{\log n}{n}$ and $R(I_v) > \frac{1}{\log^2 n}$.

Proof. We show that with high probability there exists an interval $I$ with $L(I) < \frac{\log n}{n}$ and with high probability there is no interval $I$ with $L(I) < \frac{\log n}{n}$ and $R(I) < \frac{1}{\log^2 n}$.

The probability that no interval $I$ in the interval representation has $L(I) < \frac{\log n}{n}$ and $R(I) < \frac{1}{\log^2 n}$ is

$$
\left(1 - 2 \cdot \frac{\log n}{n} \cdot \frac{1}{\log^2 n} + \left[\frac{\log n}{n}\right]^2\right)^n = \left(1 - 2 \cdot \frac{1}{n \log n} + \left[\frac{\log n}{n}\right]^2\right)^n
\sim \exp\left(-2 \cdot \frac{1}{\log n} + \frac{\log^2 n}{n}\right)
$$

Then we see that

$$
\lim_{n \to \infty} \exp\left(-2 \cdot \frac{1}{\log n} + \frac{\log^2 n}{n}\right) = \exp(0) = 1,
$$

that is with high probability no interval $I$ in the interval representation has $L(I) < \frac{\log n}{n}$ and $R(I) < \frac{1}{\log^2 n}$.

Now we show that with high probability there exists an interval with $L(I) < \frac{\log n}{n}$ and $R(I) > \frac{1}{\log^2 n}$ (and thus $I_v$ satisfies these criteria). The probability that no interval satisfies these conditions is

$$
\left(1 - 2 \cdot \frac{\log n}{n} \cdot \left[1 - \frac{1}{\log^2 n}\right]\right)^n = \left(1 - 2 \cdot \frac{\log n}{n} + \frac{2}{n \log n}\right)^n
\sim \exp\left(-2 \cdot \log n + \frac{2}{\log n}\right)
$$
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

Taking the limit, we get

$$\lim_{n \to \infty} \exp \left( -2 \cdot \log n + \frac{2}{\log n} \right) = \exp(-\infty + 0) = 0$$

so with high probability some interval satisfies the bounds and thus $I_{v_i}$ must with high probability have $L(I_{v_i}) < \frac{\log n}{n}$ and $R(I_{v_i}) > \frac{1}{\log n}$.

By symmetry, we make similar statements regarding the endpoints of $I_{v_r}$.

**Corollary 4.19.** Let $G$ be a random type I interval digraph with $n$ vertices such that interval endpoints are selected independently from $[0, 1]$. With high probability $L(I_{v_r}) < 1 - \frac{1}{\log^2 n}$ and $R(I_{v_r}) > 1 - \frac{\log n}{n}$. □

Next we address the intermediate vertices in a $(v_r, v_\ell)$-path.

**Lemma 4.20.** Let $G$ be a random type I interval digraph with $n$ vertices such that interval endpoints are selected independently from $[0, 1]$. With high probability there exist vertices that satisfy the following:

- A vertex $v_1$ with both endpoints of $I_{v_1}$ in $\left[\frac{1}{2\log^2 n}, \frac{1}{\log^2 n}\right]$.
- A vertex $v_2$ with $L(I_{v_2}) \in \left[\frac{\log n}{n}, \frac{1}{2\log^2 n}\right]$ and $R(I_{v_2}) \in \left[1 - \frac{1}{2\log^2 n}, 1 - \frac{\log n}{n}\right]$.
- A vertex $v_3$ with both endpoints of $I_{v_3}$ in $\left[1 - \frac{1}{\log^2 n}, 1 - \frac{1}{2\log^2 n}\right]$.

**Proof.** For both $v_1$ and $v_3$, the probability that no interval satisfies the given bounds is

$$\left(1 - \left[\frac{1}{2\log^2 n}\right]^2\right)^n = \left(1 - \frac{1}{4\log^4 n}\right)^n \sim \exp \left(-\frac{n}{4\log^4 n}\right)$$

81
Taking the limit, we have
\[
\lim_{n \to \infty} \exp \left( -\frac{n}{4 \log^4 n} \right) = \exp(-\infty) = 0
\]
so we know with high probability, vertices satisfying the conditions for \( v_1 \) and \( v_3 \) exist.

Next we show some vertex satisfies the bounds given for \( I_{v_2} \). The probability that no interval fits these bounds is
\[
\left( 1 - \left[ \frac{1}{\log^2 n} - \frac{\log n}{n} \right]^2 \right)^n = \left( 1 - \frac{1}{\log^4 n} + 2 \cdot \frac{1}{n \log n} - \frac{\log^2 n}{n^2} \right)^n
\]
\[
\sim \exp \left( -\frac{n}{\log^4 n} + 2 \cdot \frac{1}{\log n} - \frac{\log^2 n}{n} \right)
\]
Taking the limit, we have
\[
\lim_{n \to \infty} \exp \left( -\frac{n}{\log^4 n} + 2 \cdot \frac{1}{\log n} - \frac{\log^2 n}{n} \right) = \exp(-\infty + 0 - 0) = 0
\]
so we know with high probability, vertices satisfying the conditions for \( v_2 \) exist.

Now that we have established the existence of \( v_1, v_2, \) and \( v_3 \) with high probability, we bound \( d(v_r, v_\ell) \) precisely.

**Proposition 4.21.** For a type I interval representation chosen uniformly at random, \( P(d(v_r, v_\ell) = 2) = \frac{1}{2} \) and \( P(d(v_r, v_\ell) = 4) = \frac{1}{2} \) asymptotically as \( n \) goes to \( \infty \).

**Proof.** Note that, \( d(v_r, v_\ell) = 0 \) if and only if \( v_\ell = v_r \). The probability of this is \( \frac{1}{2\pi n-1} \), which goes to 0 as \( n \) goes to infinity. Based on this, we can safely ignore this case in the investigation.
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

Given \( v_\ell \neq v_r \), observe that either \( I_{v_\ell} \ll I_{v_r} \) or \( I_{v_\ell} < I_{v_r} \), since we know by definition of \( I_{v_\ell} \) and \( I_{v_r} \) that it is not the case that \( I_{v_\ell} \subseteq I_{v_r} \) or \( I_{v_r} \subseteq I_{v_\ell} \). Each of these occurs with limiting probability \( \frac{1}{2} \) based on the relative ordering of \( R(I_{v_\ell}) \) and \( L(I_{v_r}) \).

If \( I_{v_\ell} < I_{v_r} \), we know \( L(I_{v_\ell}) < L(I_{v_r}) < R(I_{v_\ell}) < R(I_{v_r}) \). Then letting \( L(I_{v_\ell}) = x_\ell, L(I_{v_r}) = x_r, R(I_{v_\ell}) = y_\ell, R(I_{v_r}) = y_r \), the limiting probability that there is no vertex \( u \) such that \( I_u \subseteq I_{v_\ell} \) and \( I_u \subseteq I_{v_r} \) is

\[
\lim_{n \to \infty} \left( \int_{y_r=0}^{1} \int_{y_\ell=0}^{y_r} \int_{x_r=0}^{1} \int_{x_\ell=0}^{x_r} ( (y_r - x_\ell)^2 - (y_\ell - x_r)^2 )^{n-2} dx_\ell \, dx_r \, dy_\ell \, dy_r \right).
\]

Since \( y_r - x_\ell > y_\ell - x_r \), observe that

\[
(y_r - x_\ell)^2 \geq (y_r - x_\ell)^2 - (y_\ell - x_r)^2.
\]

This is positive everywhere, so we bound the limiting probability above by replacing the integrand and expanding the bounds of integration. This gives an upper bound of

\[
\lim_{n \to \infty} \left( \int_{y_r=0}^{1} \int_{y_\ell=0}^{1} \int_{x_r=0}^{1} \int_{x_\ell=0}^{x_r} ( (y_r - x_\ell)^2 )^{n-2} dx_\ell \, dx_r \, dy_\ell \, dy_r \right)
\]

which evaluates to zero (according to Mathematica). Thus with high probability there is some vertex \( u \) such that \( I_u \subseteq I_{v_\ell} \) and \( I_u \subseteq I_{v_r} \). Then \( (v_r, u, v_\ell) \) gives a \((v_r, v_\ell)\)-path. Thus as \( n \) goes to infinity, we conclude \( P(d(v_r, v_\ell) = 2) = P(I_{v_\ell} < I_{v_r}) = \frac{1}{2} \).

From Lemma 4.18, Corollary 4.19 and Lemma 4.20 we know that with high probability there exist vertices \( v_1, v_2, v_3 \) such that \( I_{v_1} \subseteq I_{v_\ell}, I_{v_1} \subseteq I_{v_2}, I_{v_3} \subseteq I_{v_2} \), and
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

$I'_{v_3} \subseteq I_{v_r}$. This implies $(v_r, v_3, v_2, v_1, v_{\ell})$ is a path of length 4 from $v_r$ to $v_{\ell}$. Thus with high probability $d(v_r, v_{\ell}) \leq 4$.

Since $P(d(v_r, v_{\ell}) = 2) = \frac{1}{2}$, $P(d(v_r, v_{\ell}) = 0) = 0$, and from Proposition 4.16 we know $d(v_r, v_{\ell})$ is even, we conclude that as $n$ goes to infinity, $P(d(v_r, v_{\ell}) = 4) = \frac{1}{2}$. \hfill \Box

Now that we have bounded $d(v_r, v_{\ell})$, we show the existence of a set of vertices which in combination with $v_{\ell}, v_1, v_2, v_3, v_r$ allow us to create short paths between arbitrary pairs of vertices in the digraph. The proof uses the same method of examining asymptotics as the proofs of Lemma 4.18 and Lemma 4.20.

**Lemma 4.22.** Let $G$ be a random type I interval digraph with $n$ vertices such that interval endpoints are selected independently from $[0, 1]$. With high probability, there exist vertices that satisfy the following:

i. A vertex $a_1$ with $L(I_{a_1}) \in [0, \log_2 n/n]$ and $R(I_{a_1}) \in [\frac{1}{2} + n^{-1/4}, 1]$. By symmetry there is also a vertex $a_3$ with $L(I_{a_3}) \in [0, \frac{1}{2} - n^{-1/4}]$ and $R(I_{a_3}) \in [1 - \frac{\log_2 n}{n}, 1]$.

ii. A vertex $a_2$ with both endpoints in $[\frac{1}{2} - n^{-1/4}, \frac{1}{2} + n^{-1/4}]$.

iii. A vertex $b_1$ with $L(I_{b_1}) \in [0, n^{-2/3}]$ and $R(I_{b_1}) \in [1 - \frac{1}{2\log^2 n}, 1]$. By symmetry there is also a vertex $c_1$ with $R(I_{c_1}) \in [1 - n^{-2/3}, 1]$ and $L(I_{c_1}) \in [0, \frac{1}{2\log^2 n}]$.

iv. A vertex $b_2$ with both endpoints in $[n^{-2/3}, n^{-5/12}]$. By symmetry there is also a vertex $c_2$ with both endpoints in $[1 - n^{-5/12}, 1 - n^{-2/3}]$. 

84
v. Additionally, with high probability there does not exist any vertex \( v \) with \( L(I_v) \in [0, n^{-2/3}] \) and \( R(I_v) \in [0, n^{-5/12}] \). By symmetry, there also does not exist any vertex \( v \) with \( L(I_v) \in [1 - n^{-5/12}, 1] \) and \( R(I_v) \in [1 - n^{-2/3}, 1] \).

**Proof.** We prove each assertion in turn

i. The probability that no interval \( I \) in the interval representation has \( L(I) \in [0, \frac{\log n}{n}] \) and \( R(I) \in [\frac{1}{2} + n^{-1/4}, 1] \) is

\[
\left( 1 - 2 \cdot \frac{\log n}{n} \cdot \left( \frac{1}{2} - n^{-1/4} \right) \right)^n = \left( 1 - \frac{\log n}{n} + 2 \frac{\log n}{n^{5/4}} \right)^n
\]

\[ \sim \exp \left( -\log n + 2n^{-1/4} \log n \right) \]

Taking the limit, we have

\[
\lim_{n \to \infty} \exp \left( -\log n + 2n^{-1/4} \log n \right) = \exp(-\infty + 0) = 0
\]

so we know with high probability, there exists a vertex \( a_1 \) with \( L(I_{a_1}) \in [0, \frac{\log n}{n}] \) and \( R(I_{a_1}) \in [\frac{1}{2} + n^{-1/4}, 1] \).

ii. The probability that no interval \( I \) in the interval representation has both endpoints in \([\frac{1}{2} - n^{-1/4}, \frac{1}{2} + n^{-1/4}]\) is

\[
\left( 1 - 2(n^{-1/4})^2 \right)^n \sim \exp \left( -2n^{1/2} \right)
\]

Taking the limit, we have

\[
\lim_{n \to \infty} \exp \left( -2n^{1/2} \right) = \exp(-\infty) = 0
\]
so we know with high probability, there exists a vertex $a_2$ with both endpoints in $\left[ \frac{1}{2} - n^{-1/4}, \frac{1}{2} + n^{-1/4} \right]$.

iii. The probability that no interval $I$ in the interval representation has $L(I) \in [0, n^{-2/3}]$ and $R(I) \in [1 - \frac{1}{2\log^2 n}, 1]$ is

$$\left( 1 - 2n^{-2/3} \cdot \frac{1}{2\log^2 n} \right)^n \sim \exp \left( -2 \frac{n^{1/3}}{2\log^2 n} \right)$$

Taking the limit, we have

$$\lim_{n \to \infty} \exp \left( -2 \frac{n^{1/3}}{2\log^2 n} \right) = \exp(-\infty) = 0$$

so we know with high probability, there exists a vertex $b_1$ with $L(I_{b_1}) \in [0, n^{-2/3}]$ and $R(I_{b_1}) \in [1 - \frac{1}{2\log^2 n}, 1]$.

iv. The probability that no interval $I$ in the interval representation has both endpoints in $[n^{-2/3}, n^{-5/12}]$ is

$$\left( 1 - (n^{-5/12} - n^{-2/3})^2 \right)^n = \left( 1 - n^{-5/6} + 2n^{-13/12} - n^{-4/3} \right)^n \sim \exp \left( -n^{1/6} + 2n^{-1/12} - n^{-1/3} \right)$$

Taking the limit, we have

$$\lim_{n \to \infty} \exp \left( -n^{1/6} + 2n^{-1/12} - n^{-1/3} \right) = \exp(-\infty + 0 - 0) = 0$$

so we know with high probability, there exists a vertex $b_2$ with both $L(I_{b_2})$ and $R(I_{b_2})$ in $[n^{-2/3}, n^{-5/12}]$. 

86
v. The probability that no interval $I$ in the interval representation has $L(I) \in [0, n^{-2/3}]$ and $R(I) \in [0, n^{-5/12}]$ is 

\[
\left(1 - 2n^{-2/3} \cdot (n^{-5/12} - n^{-2/3}) - (n^{-2/3})^2\right)^n = \left(1 - 2n^{-13/12} + n^{-4/3}\right)^n
\]

\[
\sim \exp\left(-2n^{-1/12} + n^{-1/3}\right)
\]

so we have

\[
\lim_{n \to \infty} \exp\left(-2n^{-1/12} + n^{-1/3}\right) = \exp(0 + 0) = 1
\]

so with high probability there does not exist any vertex $v$ with $L(I_v) \in [0, n^{-2/3}]$ and $R(I_v) \in [0, n^{-5/12}]$.

To help give a more intuitive picture of what is going on with these vertices and their corresponding intervals, see Figure 4.1. Note that this arrangement is not exact, but more an example of a likely configuration of the intervals relative to one another. The intervals are color coded according to which “category” of path they belong. This is clarified in the proof of the next result, which uses these vertices as building blocks for paths between arbitrary pairs of vertices.

**Theorem 4.23.** With high probability, the diameter of a random type I interval digraph is 3 or 4.

**Proof.** We first obtain a lower bound by using the fact that with high probability, there exists some vertex $w$ such that $I_w \ll I_v$. This is derived from an argument
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

Figure 4.1: Approximate picture of the intervals corresponding to the special vertices.

similar to those used in proving Lemma 4.19. Note now that we immediately have $d(v_r, w) > 1$. The only way $d(v_r, w) = 2$ would be if there exists a vertex $u$ such that $I_w \subseteq I_u$ and $I_{v_r} \subseteq I_u$. However since by definition $R(v_r)$ is the rightmost right endpoint in the interval representation, no such $u$ exists. This implies $d(v_r, u) > 2$ and thus $\text{diam}(G) \geq 3$.

We next must prove an upper bound. Observe that for any interval $I \subseteq [0, 1]$, we know that at least one of the following must be true: $I \subseteq I_v$, $I_v \subseteq I$, $I \subseteq I_{v_1}$, $I_{v_3} \subseteq I$, or $I \subseteq I_{v_3}$.

Take any two vertices $u, w \in V(G)$. We consider each case to find a $(u, v)$-path of length at most 4. Figure 4.2 exhibits these paths for most cases. The three remaining cases require more careful analysis and use the vertices $a_1, a_2, a_3, b_1, b_2, c_1, c_2$ found in Lemma 4.22.

- Suppose $I_{v_3} \subseteq I_u$ and $I_w \subseteq I_{v_1}$ but it is not the case that $I_{v_1} \subseteq I_w$. We know from Lemma 4.22 that there is no vertex $v$ such that $I_v \ll I_{b_1}$ and if $I_v < I_{b_1}$,
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

<table>
<thead>
<tr>
<th>$I_u \subseteq I_{v_1}$</th>
<th>$I_{v_1} \subseteq I_w$</th>
<th>$I_w \subseteq I_{v_2}$</th>
<th>$I_{v_3} \subseteq I_w$</th>
<th>$I_w \subseteq I_{v_r}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(u, v_1, w)$</td>
<td>$(u, v_1, w)$</td>
<td>$(u, v_1, v_2, w)$</td>
<td>$(u, v_1, v_2, v_r, w)$</td>
<td>$(u, v_1, v_2, v_r, w)$</td>
</tr>
<tr>
<td>$(u, v_1, v_2, w)$</td>
<td>$(u, v_1, v_2, w)$</td>
<td>$(u, v_1, v_2, w)$</td>
<td>$(u, v_1, v_2, w)$</td>
<td>$(u, v_1, v_2, w)$</td>
</tr>
<tr>
<td>$(u, v_2, v_1, v_2, w)$</td>
<td>$(u, v_2, v_1, v_2, w)$</td>
<td>$(u, v_2, v_1, v_2, w)$</td>
<td>$(u, v_2, v_1, v_2, w)$</td>
<td>$(u, v_2, v_1, v_2, w)$</td>
</tr>
<tr>
<td>see proof</td>
<td>$(u, v_3, v_2, v_1, w)$</td>
<td>$(u, v_3, v_2, w)$</td>
<td>$(u, v_3, v_2, w)$</td>
<td>$(u, v_3, v_2, w)$</td>
</tr>
<tr>
<td>see proof</td>
<td>see proof</td>
<td>see proof</td>
<td>(u, v_r, v_3, v_2, w)</td>
<td>(u, v_r, v_3, v_2, w)</td>
</tr>
<tr>
<td>see proof</td>
<td>see proof</td>
<td>see proof</td>
<td>(u, v_r, v_3, v_2, w)</td>
<td>(u, v_r, v_3, v_2, w)</td>
</tr>
</tbody>
</table>

**Figure 4.2:** Directed $(u, w)$-paths in a random type I interval digraph.

it must be the case that $I_{b_2} \subseteq I_v$. We now consider the only two possible subcases:

- If $I_{b_2} \subseteq I_w$, then $u, v_3, b_1, b_2, w$ is a path of length 4.
- If $I_w \subseteq I_{b_1}$, then $u, v_3, b_1, w$ is a path of length 3.

Thus in either case we have a path of length at most 4.

• Suppose $I_u \subseteq I_{v_r}$ and $I_{v_1} \subseteq I_w$ but it is not the case that $I_{v_3} \subseteq I_u$. This case is symmetric to the previous one, using $c_1$ and $c_2$ rather than $b_1$ and $b_2$.

  - If $I_{c_2} \subseteq I_u$, then $u, c_2, c_1, v_1, w$ is a path of length 4.
  - If $I_u \subseteq I_{c_1}$, then $u, c_1, v_1, w$ is a path of length 3.

Thus in either case we have a path of length at most 4.

• If $I_u \subseteq I_{v_r}$ and $I_w \subseteq I_{v_1}$ but it is not the case that $I_{v_3} \subseteq I_u$ or $I_{v_1} \subseteq I_w$, observe
that $L(I_u) > 1 - \frac{1}{\log^2 n}$ and thus by with high probability $R(I_v) < 1 - \frac{\log n}{n}$ by Lemma 4.19. This implies $I_u \subseteq I_{a_3}$. By similar logic, we conclude $I_v \subseteq I_{a_1}$.

Thus $u, a_3, a_2, a_1, v$ is a path of length 4 in the graph.

We have now shown that for any pair of vertices $u, w \in V(G)$, it must be the case that $d(u, w) \leq 4$. Thus we have $	ext{diam}(G) \leq 4$, which in combination with the lower bound yields the desired result, namely that the diameter of a random type I interval digraph must be 3 or 4.

Notice that based on the result from Proposition 4.21, Theorem 4.23 implies that asymptotically the probability of a type I random interval digraph having diameter 4 is at least $\frac{1}{2}$ (whenever $d(v_r, v_{\ell}) = 4$). Based on simulations, we conjecture that in fact almost all type I interval digraphs have diameter 4, but this remains to be shown.

4.3.2 Type II Interval Digraph

Relative to its type I counterpart, the main result for type II interval digraphs has a fairly straightforward proof.

**Theorem 4.24.** With high probability, the diameter of a random type II interval digraph is 2 or 3.

**Proof.** Note first that in order to have diameter 1, it would have to be the case that every possible edge exists in $G$ (i.e. $G$ is the complete digraph with all double
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

edges). With high probability this is not the case, so \( \text{diam}(G) \geq 2 \).

Next, let \( G \) be a random type II interval digraph with \( n \) vertices such that interval endpoints are selected independently from \([0, 1]\). We claim that there exists some vertex \( v_1 \) such that \( L(S_{v_1}) \in [0, \frac{\log n}{n}] \), \( R(S_{v_1}) > \frac{1}{2} \), and \( \frac{1}{2} \in R_{v_1} \). The probability that \( \frac{1}{2} \in R_{v_1} \) is exactly \( \frac{1}{2} \), so overall the probability of such a \( v_1 \) existing is

\[
\left( 1 - 2 \frac{\log n}{n} \cdot \frac{1}{2} \cdot \frac{1}{2} \right)^n \sim \exp \left( -\frac{1}{2} \log n \right).
\]

Taking the limit, we have

\[
\lim_{n \to \infty} \exp \left( -\frac{1}{2} \log n \right) = \exp(-\infty) = 0
\]

so we know with high probability, such a \( v_1 \) exists.

By symmetry, we also conclude vertices \( v_2, v_3, v_4 \) satisfying the following conditions exist:

- \( R(S_{v_2}) \in [1 - \frac{\log n}{n}, 1] \), \( L(S_{v_2}) < \frac{1}{2} \), and \( \frac{1}{2} \in R_{v_2} \).
- \( L(R_{v_3}) \in [0, \frac{\log n}{n}] \), \( R(R_{v_3}) > \frac{1}{2} \), and \( \frac{1}{2} \in S_{v_3} \).
- \( R(R_{v_4}) \in [1 - \frac{\log n}{n}, 1] \), \( L(R_{v_4}) < \frac{1}{2} \), and \( \frac{1}{2} \in S_{v_4} \).

Observe now that for any vertex \( u \), either \( uv_3 \) or \( uv_4 \) is an edge in \( G \). Similarly, for any vertex \( w \), either \( v_1 w \) or \( v_2 w \) is an edge in \( G \). We additionally have the edges \( v_3 v_1, v_3 v_2, v_4 v_1, v_2 v_4 v_2 \). Using these edges, we construct a length 3 path from \( u \) to \( w \) for any \( u \) and \( w \) in \( V(G) \). Then since \( d(u, w) \leq 3 \) for all \( u, w \in V(G) \), we conclude \( \text{diam}(G) \leq 3 \).
Combining these two bounds yields \(2 \leq \text{diam}(G) \leq 3\) as desired. \(\square\)

Based on simulations, it appears both diameter 2 and diameter 3 occur with some nonzero probability as \(n\) goes to infinity. We do not yet have a proof of this though.

### 4.3.3 Comparison with Erdős-Renyi random digraphs

Recall that an undirected Erdős-Renyi random graph with any fixed edge probability has diameter 2 with high probability (see for example [24]). By nearly identical logic (simply using \(n(n - 1)\) as the number of possible edges instead of \((n^2)\)), we conclude that with high probability, an Erdős-Renyi random digraph with any fixed edge probability has diameter 2. Thus a type I interval digraph with high probability has greater diameter than the Erdős-Renyi random digraph with edge probability \(\frac{1}{2}\). Type II interval digraphs, on the other hand, may have the same or greater diameter.

### 4.4 Cliques and Independent Sets in Random Interval Digraphs

Another property of interest in random (di)graphs is the size of a maximum clique or maximum independent set. Recall that in a directed graph a clique must
have all possible edges (i.e. if \( u \) and \( v \) are two vertices in a clique, both \( uv \in E \) and \( vu \in E \)). We first introduce some useful variations on Chernoff bounds.

**Proposition 4.25** (Alternate Chernoff Bound). Let \( X_1, \ldots, X_m \) be independent random indicator variables such that \( \Pr(X_i = 1) = p \). Define \( X = \sum_i X_i \) and \( \mu = E[X] = mp \). For any \( 0 < \delta \leq \mu \),

\[
\Pr[X \geq \mu + \delta] \leq \exp \left[ -\frac{\delta^2}{4\mu} \right]
\]

**Proposition 4.26** (Another Chernoff Bound). Let \( X_1, \ldots, X_m \) be independent random variables such that \( X_i \) always lies in the interval \([0, 1]\). Define \( X = \sum_i X_i \) and \( \mu = E[X] \). Let \( p_i = E[X_i] \) and note that \( \mu = \sum_i p_i \). For any \( \delta \geq 1 \),

\[
\Pr[X \geq (1 + \delta)\mu] \leq \exp(-\mu \delta / 3)
\]

where

**Proposition 4.27** (One More Chernoff Bound). Let \( X_1, \ldots, X_m \) be independent random variables such that \( X_i \) always lies in the interval \([0, 1]\). Define \( X = \sum_i X_i \) and \( \mu = E[X] \). Let \( p_i = E[X_i] \) and note that \( \mu = \sum_i p_i \). For any \( \delta \in [0, 1] \),

\[
\Pr[X \leq (1 - \delta)\mu] \leq \exp(-\mu \delta^2 / 3).
\]

### 4.4.1 Type I Interval Digraphs

In the case of type I interval digraphs, analyzing the maximum independent set size is identical to the undirected interval graph case.
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

Proposition 4.28. With high probability, the maximum size of an independent set in a random type I interval digraph on \( n \) vertices is \( \frac{2}{\sqrt{\pi}} \sqrt{n} + o(\sqrt{n}) \).

This is derived directly from a result of Justicz, Scheinerman, and Winkler for undirected random interval graphs. For details on the proof, see [13]. For now, we turn our attention to the maximum clique size.

Proposition 4.29. With high probability, the maximum size of a clique in a random type I interval digraph is \( \sqrt{2n} + o(\sqrt{n}) \).

Proof. We know we must have double edges between any two vertices in the clique, which implies that in order for a set of vertices to form a clique in a type I interval digraph, their intervals must be nested in the corresponding interval representation.

Taking advantage of this observation, we find a maximum clique in a type I interval digraph (given an interval representation with endpoints chosen from \([0, 1]\)) by using the following procedure: first, pick a point \( x \in [0, 1] \). Let \( S \subseteq V \) be the set of vertices such that \( v \in S \) if and only if \( x \in I_v \). Next, label the intervals \( I_v \) in \( S \) with \( 1, 2, \ldots, |S| \) in increasing order based on the position of the left endpoints (read from left to right). Consider the permutation of \(|S|\) generated by reading the intervals by the order of their right endpoints (again from left to right). We then find the maximum set of nested intervals by finding the longest decreasing subsequence in the resulting permutation.
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

Note that because every left endpoint of $S$ is to the left of $x$ while every right endpoint is to the right of $x$, if we select a random type I interval digraph and perform this procedure, every permutation of $|S|$ is equally likely. Thus the problem of finding a maximum clique in a random type I interval digraph is now reduced to the problem of finding a longest decreasing subsequence in a random permutation (looking at all permutations generated by all possible points $x \in [0, 1]$).

Using this idea with $x = \frac{1}{2}$, we show that with high probability there exists a clique in a random type I interval digraph of size $\sqrt{2n} + o(\sqrt{n})$. Given a random type I interval digraph, for each $v \in V$ define $X_v$ to be the indicator random variable that is 1 when $\frac{1}{2} \in I_v$. We have $P(X_v = 1) = \frac{1}{2}$ so $\mu = \frac{1}{2} n$. Applying the Chernoff bound from Proposition 4.27 with $\delta = n^{-1/3}$, we have

$$P \left[ X \leq \left(1 - n^{-1/3}\right) \frac{1}{2} n \right] \leq \exp(-n^{1/3}/12) \to 0$$

as $n \to \infty$. This implies that with high probability, there are $\frac{1}{2} n + o(n)$ vertices $v \in V$ with $\frac{1}{2} \in I_v$.

From [25], we know that with high probability the length of a longest decreasing subsequence in a random permutation of $[k]$ is $2\sqrt{k}$. Thus, applying the procedure above to the set of $\frac{1}{2} n + o(n)$ vertices whose intervals contain $\frac{1}{2}$, the largest clique has size

$$2\sqrt{\frac{1}{2} n + o(n)} \leq 2 \left(\sqrt{\frac{1}{2} n} + \sqrt{o(n)}\right) = \sqrt{2n} + o(\sqrt{n}).$$

This shows that with high probability a random type I interval digraph has max
clique size at least $\sqrt{2n} + o(\sqrt{n})$.

Next we must show that with high probability there is no clique in a random type I interval digraph of size greater than $\sqrt{2n} + o(\sqrt{n})$. We use the third model for random type I interval digraphs here, choosing endpoints for the intervals uniformly at random from the set
\[
\left\{ \frac{0}{n^3}, \frac{1}{n^3}, \ldots, \frac{n^3 - 1}{n^3}, 1 \right\}.
\]

We may assume that no two endpoints are the same (since with high probability this is the case).

We only need to check the size of the cliques based on the possible common points $\left\{ \frac{1}{2n^3}, \frac{3}{2n^3}, \ldots, \frac{2n^3 - 1}{2n^3} \right\}$. Consider the point $\frac{2k - 1}{2n^3}$. For any vertex $v \in V$ let $X_{k,v}$ be the indicator random variable that is 1 when $\frac{2k - 1}{2n^3} \in I_v$.

In order to have $X_{k,v} = 1$, $I_v$ must have one endpoint selected from the set $\left\{ 0, \frac{1}{n^3}, \ldots, \frac{k - 1}{n^3} \right\}$ and one endpoint from the set $\left\{ \frac{k}{n^3}, \ldots, \frac{n^3 - 1}{n^3}, 1 \right\}$. These sets are size $k$ and $n^3 - k + 1$ respectively, and there are $n^3 + 1$ choices for each endpoint, so we have
\[
P(X_{k,v} = 1) = \left( 2 \cdot \frac{k}{n^3 + 1} \cdot \frac{n^3 - k + 1}{n^3 + 1} \right).
\]

This is maximized when $k = \frac{n^3 + 1}{2}$, yielding $P(X_{k,v} = 1) = \frac{1}{2}$.

Let $X_k = \sum_{v \in V} X_{k,v}$. Then $X_k$ gives the size of the clique made up of all vertices with corresponding interval containing $\frac{2k - 1}{n^3}$. Let $\mu_k = E[X_k]$. Note that $0 \leq \mu_k \leq \frac{n}{2}$.
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

We employ Chernoff bounds to show that with high probability the clique corresponding to a particular point \( \frac{2k-1}{n^3} \) is at most size \( \sqrt{2n} + o(\sqrt{n}) \). Let \( \delta_k = \left( \frac{1}{2} + n^{-1/3} \right)n - \mu_k \). If \( \mu_k \geq \frac{1}{2} \left( \frac{1}{2} + n^{-1/3} \right)n \), we apply the Chernoff bound from Proposition 4.25 to get

\[
P(X_k \geq \left( \frac{1}{2} + n^{-1/3} \right)n) \leq \exp \left[ - \frac{\left( \frac{1}{2} + n^{-1/3} \right)n - \mu_k \right]^{2}{4\mu_k}].
\]

This is maximized when \( \mu_k = \frac{n}{2} \). Similarly when \( \mu_k \geq \frac{1}{2} \left( \frac{1}{2} + n^{-1/3} \right)n \), observe that \( \left( \frac{1}{2} + n^{-1/3} \right)n > 2\mu_k \) so we apply Proposition 4.26 with \( \gamma = \frac{1}{\mu_k} \left( \frac{1}{2} + n^{-1/3} \right)n - 1 \) to obtain

\[
P(X_k \geq \left( \frac{1}{2} + n^{-1/3} \right)n) \leq \exp \left[ - \frac{\left( \frac{1}{2} + n^{-1/3} \right)n - \mu_k \right]^{3}{3}].
\]

This is maximized when \( \mu_k = \frac{1}{2} \left( \frac{1}{2} + n^{-1/3} \right)n \). Using the maximum possible probability and summing over all \( n^3 \) possible common/gap points, we derive an upper bound the probability of having a set of more than \( \left( \frac{1}{2} + n^{-1/3} \right)n \) vertices all containing one of the common points:

\[
n^3 \left( \exp \left[ - \frac{1}{4} n^{-7/3} \right] + \exp \left[ - \frac{\left( \frac{1}{2} + n^{-1/3} \right)n \right]^{6}{6} \right]
\]

which goes to zero as \( n \) goes to infinity. This implies that with high probability, over all the common points \( \left\{ 0, \frac{1}{n^3}, \frac{2}{n^3}, \ldots, \frac{n^3-1}{n^3}, 1 \right\} \), the largest set of vertices sharing any given point is of size \( \frac{1}{2}n + o(n) \).

Again applying [25], we find that the largest clique has size

\[
2\sqrt{\frac{1}{2}n + o(n)} \leq 2 \left( \sqrt{\frac{1}{2}n + \sqrt{o(n)}} \right) = \sqrt{2n} + o(\sqrt{n}).
\]
This shows that with high probability a random type I interval digraph has max clique size at most $\sqrt{2n} + o(\sqrt{n})$.

4.4.2 Type II Interval Digraphs

We now shift our attention to type II interval digraphs. When looking at maximum clique size, the situation is significantly different from the type I case.

Proposition 4.30. With high probability, the maximum size of a clique in a random type II interval digraph is $\frac{n}{4} + o(n)$.

Proof. We first show that with high probability, there exists a clique in a random type II interval digraph of size $\frac{n}{4} + o(n)$. Given a random type II interval digraph with interval endpoints chosen independently from $[0, 1]$, for each $v \in V$ define $X_v$ to be the indicator random variable that is 1 when $\frac{1}{2} \in S_v$ and $\frac{1}{2} \in R_v$. We have $P(X_v = 1) = \frac{1}{4}$ so $\mu = \frac{1}{4}n$. Applying the Chernoff bound from Proposition 4.27 with $\delta = n^{-1/3}$, we have

$$P \left[ X \leq \left( 1 - n^{-1/3} \right) \frac{1}{4} \right] \leq \exp \left( -n^{1/3}/12 \right) \to 0$$

as $n \to \infty$. This implies that with high probability, the clique formed by considering vertices with $\frac{1}{2} \in S_v$ and $\frac{1}{2} \in R_v$ is of size at least $\frac{1}{4}n + o(n)$. Thus the max clique in the graph must be of size at least $\frac{1}{4}n + o(n)$. 

98
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

Next we must show that with high probability there is no clique in a random type II interval digraph of size greater than $\frac{1}{4}n + o(n)$. We use the third model for random type II interval digraphs here, choosing endpoints for the intervals uniformly at random from the set

$$\left\{ 0, \frac{1}{n^3}, \frac{2}{n^3}, \ldots, \frac{n^3-1}{n^3}, 1 \right\}.$$ 

We assume that no two endpoints are the same (since with high probability this is the case).

In order for a set of vertices to form a clique, there are three possibilities for the interval representation. First, it is possible that all intervals (both send and receive) corresponding to the vertices of the clique share a common point. Another possibility allows for a gap among the receive intervals, but in this case each send intervals must contain the entirety of the gaps. The final symmetric case allows a gap in the send intervals, but requires all receive intervals contain the entirety of the gaps. We examine each of these cases individually.

In the case where all intervals share a common point, we only need to check the size of the cliques based on the possible common points $\left\{ \frac{1}{2n^3}, \frac{3}{2n^3}, \ldots, \frac{2n^3-1}{2n^3} \right\}$. Consider the point $\frac{2k-1}{2n^3}$. For any vertex $v \in V$ let $X_{k,v}$ be the indicator random variable that is 1 when $\frac{2k-1}{2n^3} \in R_v$ and $\frac{2k-1}{2n^3} \in S_v$.

In order to have $X_{k,v} = 1$, each of $S_v$ and $R_v$ must have one endpoint selected from the set $\left\{ 0, \frac{1}{n^3}, \ldots, \frac{k-1}{n^3} \right\}$ and one endpoint from the set $\left\{ \frac{k}{n^3}, \ldots, \frac{n^3-1}{n^3}, 1 \right\}$. These sets are size $k$ and $n^3 - k + 1$ respectively, and there are $n^3 + 1$ choices for each
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

endpoint, so we have

\[ P(X_{k,v} = 1) = \left( 2 \cdot \frac{k}{n^3 + 1} \cdot \frac{n^3 - k + 1}{n^3 + 1} \right)^2. \]

This is maximized when \( k = \frac{n^3 + 1}{2} \), yielding \( P(X_{k,v} = 1) = \frac{1}{4} \).

Let \( X_k = \sum_{v \in V} X_{k,v} \). Then \( X_k \) gives the size of the clique made up of all vertices with both send and receive intervals containing \( \frac{2k-1}{n^3} \). Let \( \mu_k = E[X_k] \). Note that \( 0 \leq \mu_k \leq \frac{n}{4} \).

We now employ Chernoff bounds to show that with high probability the clique corresponding to a particular point \( \frac{2k-1}{n^3} \) is at most size \( \frac{1}{4} + o(n) \). Let \( \delta_k = (\frac{1}{4} + n^{-1/3})n - \mu_k \). If \( \mu_k \geq \frac{1}{2}(\frac{1}{4} + n^{-1/3})n \), we apply the Chernoff bound from Proposition 4.25 to get

\[ P(X_k \geq (\frac{1}{4} + n^{-1/3})n) \leq \exp \left[ -\frac{\left( \left( \frac{1}{4} + n^{-1/3} \right)n - \mu_k \right)^2}{4\mu_k} \right]. \]

This is maximized when \( \mu_k = \frac{n}{4} \). Similarly when \( \mu_k \geq \frac{1}{2}(\frac{1}{4} + n^{-1/3})n \), observe that \( (\frac{1}{4} + n^{-1/3})n > 2\mu_k \) so we apply Proposition 4.26 with \( \gamma = \frac{1}{\mu_k}(\frac{1}{4} + n^{-1/3})n - 1 \) to obtain

\[ P(X_k \geq (\frac{1}{4} + n^{-1/3})n) \leq \exp \left[ -\frac{(\frac{1}{4} + n^{-1/3})n - \mu_k}{3} \right]. \]

This is maximized when \( \mu_k = \frac{1}{2} \left( \frac{1}{4} + n^{-1/3} \right) n \). Using the maximum possible probability and summing over all \( n^3 \) possible common/gap points, we find an upper bound on the probability of a clique of size at least \( (\frac{1}{4} + n^{-1/3})n \) such that all inter-
vals share a common point:

\[ n^3 \left( \exp \left( -n^{-7/3} \right) + \exp \left( -\frac{\left( \frac{1}{4} + \frac{1}{6} n^{-1/3} \right)}{n} \right) \right) \]

which goes to zero as \( n \) goes to infinity.

Now consider the case where there is a gap amongst the receive intervals. Notice that since every send interval must intersect every receive interval, the send intervals all must contain the entirety of the gap. Note that we only need to examine possible gaps at the points from before: \( \left\{ \frac{1}{2n^3}, \frac{3}{2n^3}, \ldots, \frac{2n^3-1}{2n^3} \right\} \). Consider the point \( \frac{2k-1}{2n^3} \). For any vertex \( v \in V \), define \( Y_{k,v} \) to be the indicator random variable that is 1 when \( \frac{2k-1}{2n^3} \in S_v \) but \( \frac{2k-1}{n^3} \notin R_v \).

In order to have \( Y_{k,v} = 1 \), one endpoint of \( S_v \) must be selected from the set \( \{0, \frac{1}{n^3}, \ldots, \frac{k-1}{n^3}\} \) and the other from \( \{\frac{k}{n^3}, \ldots, \frac{n^3-1}{n^3}, 1\} \). The two endpoints of \( R_v \), on the other hand, must both be selected from the same set. This gives

\[ P(Y_{k,v} = 1) = \left( 2 \cdot \frac{k}{n^3+1} \cdot \frac{n^3 - k + 1}{n^3 + 1} \right) \left( 1 - 2 \cdot \frac{k}{n^3+1} \cdot \frac{n^3 - k + 1}{n^3 + 1} \right). \]

Similar to the previous case, this is maximized when \( k = \frac{n^3+1}{2} \), yielding \( P(Y_{k,v} = 1) = \frac{1}{4} \).

Let \( Y_k = \sum_{v \in V} Y_{k,v} \). Notice that \( Y_k \) gives an upper bound on the size of any clique made up of only vertices with send intervals containing \( \frac{2k-1}{n^3} \) but a gap in the receive intervals at \( \frac{2k-1}{n^3} \). This is only an upper bound due to the fact that not all vertices obeying these constraints necessarily have send intervals intersecting all the receive intervals and vice versa.
CHAPTER 4. PROPERTIES OF RANDOM INTERVAL DIGRAPHS

We now use Chernoff bounds with the same logic as before to bound $Y_k$. The argument is identical, again yielding

$$n^3 \left( \exp \left[ -n^{-7/3} \right] + \exp \left[ -\frac{\left( \frac{1}{4} + n^{-1/3} \right)n}{6} \right] \right)$$

which goes to zero as $n$ goes to infinity.

The final case is symmetric to the second case, now having a gap between send intervals with a common point in all receive intervals. Summing the results from all three cases shows that with high probability the largest clique in a random type II interval digraph is $\frac{1}{4}n + o(n)$.

\[ \square \]

The independent set case also differs from the type I case. Here we must do a separate analysis from the undirected case, as we can allow intervals to intersect as long as they are of the same type (either send or receive) and still have an independent set.

**Proposition 4.31.** With high probability, the maximum size of an independent set in a random type II interval digraph is at least $\frac{n}{16} + o(n)$.

**Proof.** We show that with high probability, there exists an independent set in a random type II interval digraph of size $\frac{n}{16} + o(n)$. This implies the maximum size independent set must be at least this large. Given a random type II interval digraph with interval endpoints chosen independently from $[0, 1]$, for each $v \in V$ define $X_v$ to be the indicator random variable that is 1 when $S_v \subseteq [0, \frac{1}{2})$ and $R_v \subseteq (\frac{1}{2}, 1]$. 

102
Observe that $P(X_v = 1) = \frac{1}{16}$ so $\mu = \frac{1}{16} n$. Applying the Chernoff bound from Proposition 4.27 with $\delta = n^{-1/3}$, we have

$$P \left[ X \leq \left( 1 - n^{-1/3} \right) \frac{1}{16} n \right] \leq \exp \left( -n^{1/3} / 48 \right) \rightarrow 0$$

as $n \rightarrow \infty$. This implies that with high probability, the independent set formed by considering vertices with send intervals completely to the left of $\frac{1}{2}$ and receive intervals completely to the right of $\frac{1}{2}$ is of size $\frac{1}{16} n + o(n)$. Thus the max independent set in the graph must be of size at least $\frac{1}{16} n + o(n)$.\[
\square
\]

Based on simulations we conjecture that with high probability the maximum independent set size in a type II interval digraph is $\frac{n}{16} + o(n)$, but we do not yet have a proof for the upper bound needed for this result.

### 4.4.3 Comparison with Erdős-Renyi random digraphs

Recall that for a random undirected Erdős-Rényi graph with a constant edge probability $p$, with high probability we have a maximum independent set of size $2 \log \frac{1}{1-p} n$ and a maximum clique of size $2 \log \frac{1}{p} n$ (see for example [24]).

For a random Erdős-Rényi digraph with constant edge probability $p$, we then find the maximum clique size by mapping to the undirected case with edge probability $p^2$ (i.e. there is an edge $uv$ in the undirected case if and only if both $uv$ and $vu$ are edges in the directed case), giving maximum clique size $2 \log \frac{1}{p^2} n$. Similarly, we
find the maximum independent set size by mapping to the undirected case with
edge probability $1 - (1 - p)^2$, yielding maximum independent set size $2 \log \frac{1}{(1-p)^2} n$.

Looking at random Erdős-Rényi digraphs with edge probability $\frac{1}{2}$, the maximum clique size is asymptotically equivalent to $\frac{2}{\log 4} \log n$, asymptotically smaller than the maximum clique size for random type I interval digraphs $\left( \sqrt{2n} + o(\sqrt{n}) \right)$.

Examining instead random Erdős-Rényi digraphs with edge probability $\frac{2}{3}$, we find the maximum clique size is asymptotically equivalent to $\frac{2}{\log 3} \log n$, asymptotically smaller than the maximum clique size for random type II interval digraphs $\left( \frac{n}{4} + o(n) \right)$.

The maximum independent set size in a random Erdős-Rényi digraph with edge probability $\frac{2}{3}$ is asymptotically equivalent to $\frac{2}{\log 9} \log n$. While we cannot yet prove that the maximum size of an independent set in a random type II interval digraph is exactly $\frac{n}{16} + o(n)$, this lower bound is still asymptotically larger than the Erdős-Rényi case.
Chapter 5

Future Work

While we have investigated several aspects of interval digraphs in this thesis, there is still much left to discover.

5.1 Type I and Type II Interval Digraphs

The immediate subject matter of this dissertation is type I and type II interval digraphs, and there are still many interesting questions left to answer about them. We have introduced the concept of type I interval digraphs for the first time in this work, so nothing is known about them beyond what is included here.

Currently, we have an algorithm for recognizing type I interval digraphs, but it is based on type Ib interval digraphs. Since there is no bijective mapping from type I to type Ib, it would be ideal to have a more direct characterization of type I
CHAPTER 5. FUTURE WORK

interval digraphs that does not rely on type Ib interval digraphs at all. In particular, we would like to find a complete list of classes of forbidden subgraphs (see Figure 3.1 for a partial list), constraints on the adjacency or incidence matrices, or any other easier to state equivalent properties.

Open Problem 1. Find a complete characterization of type I interval digraphs (not reliant on type Ib interval digraphs).

The current recognition algorithm does not require checking every possible interval representation, but in the worst case it can theoretically run the type Ib recognition algorithm exponentially many times. Since there exist polynomial time recognition algorithms for both undirected interval graphs and type II interval digraphs, we conjecture that there exists a guaranteed polynomial time algorithm for type I interval digraphs as well. Relying on type Ib interval digraphs may not be the best approach in searching for an efficient algorithm, so finding a characterization as mentioned above becomes even more critical, as it may in turn lead to a polynomial time recognition algorithm (or vice versa).

Open Problem 2. Find a polynomial time recognition algorithm for type I interval digraphs.

There is also room to make some of the random results for both type I and type II interval digraphs more precise. In particular, there are open questions with respect to the diameter of either type of interval digraph. Based on simulations,
CHAPTER 5. FUTURE WORK

the number of random type I interval digraphs with diameter 3 is approximately
equal to the number of random type I interval digraphs with diameter 5. We have
shown that with high probability the diameter of a random type I interval digraph
is not 5, but we have thus far been unable to prove that the diameter is not 3 in the
case where $d(v_r, v_l) = 2$, which leads us to the following conjecture:

**Conjecture 1.** With high probability, a random type I interval digraph has diameter
4.

In the case of type II random interval digraphs, simulations appear to show that
diameters 2 and 3 both occur with some nonzero limiting probability. However,
we have not yet been able to identify what the limiting probabilities are for each of
these diameter or even prove that both occur with nonzero probability.

**Open Problem 3.** Find the limiting probabilities on diameter 2 and diameter 3 for
type II random interval digraphs.

There is also room for improvement on the bounds obtained for some other
graph properties for random type II interval digraphs. Recall that for a random
interval graph, the probability of a vertex with degree $n - 1$ is precisely $\frac{2}{3}$ inde-
pendent of $n$. For random type II interval digraphs, we demonstrated that $\frac{2}{3}$ is
an upper bound on the probability that we have a vertex with out-degree (or in-
degree) $n$. The probability in this case is not independent of $n$, however, and the
lower bounds we have do not seem tight based on simulations.
CHAPTER 5. FUTURE WORK

**Open Problem 4.** Find a more accurate lower bound describing the asymptotic behavior for a type II random interval digraph having a vertex with out-degree (or in-degree) $n$.

Lastly, we have been unable to prove an upper bound on the asymptotic size of a maximum independent set for random type II interval digraphs. Simulations, however, show that it appears this bound matches the lower bound we proved, giving the following conjecture:

**Conjecture 2.** With high probability, the size of a maximum independent set in a type II interval digraph is $\frac{n}{16} + o(n)$.

In addition to these open questions regarding (random) type I and type II interval digraphs, there are also related models we have not explored in this dissertation that could yield interesting results.

### 5.2 Random Undirected Unit Interval Graphs

While random undirected interval graphs have been studied quite a bit, random undirected unit interval graphs have been less explored. For a fixed interval length $t$, a random unit interval graph on $n$ vertices is chosen by selecting $n$ points $a_1, \ldots, a_n$ independently and uniformly at random from $[0, 1]$. The corresponding intervals are $[a_1, a_1 + t], \ldots, [a_n, a_n + t]$. We have the following basic results:
CHAPTER 5. FUTURE WORK

**Proposition 5.1.** In a random unit interval graph with fixed interval length $t$, the probability of an edge existing between any pair of vertices $u, v$ is $2t - t^2$.

**Proof.** Suppose the left endpoint of $I_u$ is $a$. If $0 \leq a \leq t$, the probability that $I_v$ intersects $I_u$ is $t + a$ (the left endpoint of $I_v$ could fall in $I_u$ or anywhere to the left of $a$). If $t \leq a \leq 1 - t$, the probability that $I_v$ intersects $I_u$ is $2t$ (the left endpoint of $I_v$ could fall in $I_u$ or within $t$ to the left of $a$). Finally if $1 - t \leq a \leq 1$, the probability that $I_v$ intersects $I_u$ is $1 - a + t$ (the left endpoint of $I_v$ could fall between $a - t$ and 1). Putting this together we have

$$
\int_{a=0}^{t} (a + t) \, dt + \int_{a=t}^{1} -t2t \, dt + \int_{a=1-t}^{1} (1 - a + t) \, dt = 2t - t^2.
$$

\qed

**Proposition 5.2.** The probability that a given set of $k$ vertices in a random unit interval graph form a clique is $kt^{k-1} - (k - 1)t^k$.

**Proof.** We condition on the leftmost left endpoint $a$ of any of the $k$ intervals. The left endpoints for the other $k - 1$ intervals must then fall in $[a, a + t]$ if $a + t \leq 1$ or $[a, 1]$ if $a + t > 1$. Putting this together gives,

$$
\int_{a=0}^{1-t} k^{k-1} \, da + \int_{a=1-t}^{1} k(1 - a)^{k-1} \, da = kt^{k-1} - (k - 1)t^k.
$$

\qed

**Proposition 5.3.** The probability that a given set of $k$ vertices in a random unit interval graph form an independent set is $(1 - (k - 1)t)^k$. 

109
CHAPTER 5. FUTURE WORK

Proof. To ensure no overlap in intervals, notice that the left endpoints of any pair of intervals must be at least $t$ apart. We create a bijection between $k$-tuples of points in $[0, 1 - (k-1)t]$ and $k$-tuples of points in $[0, 1]$ such that any pair of points is at least $t$ apart. In particular, we map

$$(a_1, a_2, \ldots, a_k) \mapsto (a_1, a_2 + t, a_3 + 2t, \ldots, a_k + (k-1)t).$$

The probability we select such a $k$-tuple is $(1 - (k-1)t)^k$, and thus this is the probability that $k$ vertices form an independent set.

All other properties of random unit interval graphs still remain to be explored. We expect some results may be easier to find and prove due to the increased restrictions meaning there are fewer variables involved. On the other hand, we are no longer able to take advantage of the discrete permutations model, which allows less flexibility in proof methods.

5.3 Random Random Interval Graphs

Another possible avenue for future exploration is “random random interval graphs.” Similar to normal random interval graphs, we create a random interval representation, associating each vertex with an interval by selecting endpoints independently and uniformly at random from the unit interval $[0, 1]$. We add an additional random step to determine whether or not $uv \in E$. In particular, the
CHAPTER 5. FUTURE WORK

Probability of edge $uv$ existing is based on the length of the intersection of the intervals corresponding to $u$ and $v$. We either use the length directly as a probability, that is $Pr(u \sim v) = \lambda(I_u \cap I_v)$, or we use the proportion of the two intervals’ total lengths that intersect, in which case $Pr(u \sim v) = \frac{\lambda(I_u \cap I_v)}{\lambda(I_u \cup I_v)}$.

Note that this still ensures that if $I_u \cap I_v = \emptyset$, $uv \notin E$, but in the case $I_u \cap I_v \neq \emptyset$, it is possible $uv \in E$ or $uv \notin E$. This also means each interval representation is associated with multiple graphs (in particular, the interval graph associated with that representation as well as all possible spanning subgraphs). This model is particularly interesting in that there are no longer any forbidden subgraphs; the case of $n$ nested intervals in particular could result in any possible graph on $n$ vertices.

We include here some basic preliminary results on random random interval graphs, but there is much more that could be done to explore this model.

**Proposition 5.4.** Suppose for fixed vertices $u, v$ we have $Pr(u \sim v) = \lambda(I_u \cap I_v)$. Then for any two vertices, $Pr(u \sim v) = \frac{2}{15}$.

**Proof.** We begin by computing $f(a, b)$, the probability that there is an edge between a vertex associated with a random interval and a vertex associated with an interval with endpoints $a$ and $b$. Recall that an interval $[x, y]$ on the real line can be associated with a point in the plane with coordinates $(x, y)$. Based on this, we divide the unit square into regions based on $a$ and $b$ as seen below (assuming $a < b$). The blue line is $y = x$, and the values in each region represent the probability that a vertex associated with an interval represented by a point in that region has an edge with
a vertex with corresponding interval \([a, b]\).

We integrate over each of the regions to obtain

\[
f(a, b) = 2 \int_{y=b}^{1} \int_{x=0}^{a} (b - a) \, dx \, dy + 2 \int_{y=b}^{1} \int_{x=0}^{a} (b - a) \, dx \, dy \\
+ 2 \int_{y=a}^{b} \int_{x=0}^{a} (y - a) \, dx \, dy + 2 \int_{y=a}^{b} \int_{x=0}^{y} (y - x) \, dx \, dy \\
= \frac{1}{3} (b - a)^3 + a(b - a)^2 + (1 - b)(b - a)^2 + 2a(1 - b)(b - a)
\]

where we scale each integral by 2 due to symmetry across the line \(y = x\). To obtain

\(Pr(u \sim v)\), we then compute

\[
Pr(u \sim v) = 2 \int_{b=0}^{1} \int_{a=0}^{b} f(a, b) = \frac{2}{15}
\]

where we scale by 2 to account for the symmetric case where \(b < a\). 

**Corollary 5.5.** Suppose for fixed vertices \(u, v\) we have \(Pr(u \sim v) = \lambda(I_u \cap I_v)\). Then for any three vertices \(u, v, w\), we have \(Pr(u \sim v \sim w) = \frac{17}{630}\).
CHAPTER 5. FUTURE WORK

Proof. We take advantage of the computation of $f(a, b)$ in the previous proof to aid us here. Note that if we condition on $I_v = [a, b]$, (or $I_v = [b, a]$), the probability that $u \sim v$ and the probability that $v \sim w$ are independent. Thus we simply compute

$$
Pr(u \sim v \sim w) = \int_{b=0}^{1} \int_{a=0}^{b} f(a, b)^2 \, da \, db = \frac{17}{630}.
$$

By similar logic, we also obtain the following result using the proportional variation on edge probability.

**Proposition 5.6.** Suppose $Pr(u \sim v) = \frac{\lambda(I_u \cap I_v)}{\lambda(I_u \cup I_v)}$. Then for any two vertices, $Pr(u \sim v) = \frac{2}{9}$.

For random random interval graphs we have touched on only the most basic of properties, leaving many avenues of exploration still open. It seems that computation gets quite complicated when looking at any properties more involved than what we have examined here, but there may be other perspectives that we did not try that could enable better investigation of these constructions.
Chapter 6

Appendix: Proofs for Forbidden Subgraphs

We include here the arguments demonstrating that the subgraphs shown in Figures 3.1 and 3.3 are in fact minimal forbidden subgraphs for Type I and Type Ib interval digraphs respectively.

6.1 Type I Interval Digraphs

We do not have a complete forbidden subgraph characterization of type I interval digraphs, but the following specific examples provably have no type I interval presentations.
CHAPTER 6. APPENDIX: PROOFS FOR FORBIDDEN SUBGRAPHS

In the first example, in order to have edge $uv$ but not $vu$, we know that $I_u < I_v$ and thus $L(I_v) \subseteq I_u$. By similar logic, in order to have edges $wv$ but not $vw$, it must be the case that $I_w < I_v$ and thus $L(I_v) \subseteq I_w$. However, this implies $I_u \cap I_w \neq \emptyset$, and thus it is impossible to avoid an edge between $u$ and $w$. The second example follows a similar logic. In this case, $I_v < I_u$ and $I_v < I_w$, implying $R(I_v) \subseteq I_u \cap I_w$, implying there must exist some edge between $u$ and $w$.

Suppose we attempt to create a type I interval representation for this example. Without loss of generality, let $I_u$ be the interval with the leftmost left endpoint. Since $wu$ is an edge but not $uw$, this implies $I_w < I_u$. However, since $I_u$ has the leftmost left endpoint, this is impossible. Thus there is no type I interval representation for this digraph.
If there exists a type I interval representation for this example, it must be the case that either $I_u \subseteq I_w$ or $I_w \subseteq I_u$ since $uw$ and $wu$ are both edges. If $I_u \subseteq I_w$, observe that since $uv$ is an edge but $vu$ is not an edge, $I_u < I_v$. This implies $L(I_v) \in I_w$, so either $I_w < I_v$ or $I_v \subseteq I_w$. Thus it is impossible to have $vw$ but not $wv$.

Similarly, if $I_w \subseteq I_u$, since $vw$ is an edge but $wv$ is not an edge, $I_v < I_w$. This implies $R(I_v) \in I_u$, and thus either $I_v < I_u$ or $I_v \subseteq I_u$. Since neither of these allow the edge $uv$ without $vu$, we conclude there is no type I interval representation for this digraph.

Consider a digraph such that the underlying structure is a path of length 3 with edges $ab, bc, cd$. Suppose also that both $bc$ and $cb$ are edges of the digraph. If there
exists a type I interval representation for such a digraph, it must be the case that
$I_b \subseteq I_c$ or $I_c \subseteq I_b$. If $I_b \subseteq I_c$, then in order to have an edge between $a$ and $b$ in
the digraph (regardless of whether that edge is $ab$, $ba$, or both), it must be the case
that $I_a \cap I_b \neq \emptyset$. However, this implies $I_a \cap I_c \neq \emptyset$, and thus either $ac, ca$, or both
would have to be edges of the digraph. Thus it cannot be the case that $I_b \subseteq I_c$. If $I_c \subseteq I_b$, however, a similar issue occurs. In order to have $cd, dc$, or both in the
digraph, it must be the case that $I_c \cap I_d \neq \emptyset$. However, this implies $I_b \cap I_d \neq \emptyset$, and thus there would have to be some edge between $b$ and $d$. Thus there is no type
I interval representation for this class of digraphs.

6.2 Type Ib Interval Digraphs

As a reminder, in addition to the specific examples included here, any solid-
dashed digraph that is not a type I interval digraph when its dashed edges are
replaced by double edges is not a type Ib interval digraph. Also, any solid-dashed
digraph digraph that does not have exactly one vertex with in-degree zero (for
both solid and dashed edges) in each component is not a type Ib interval digraph.

In order to have a type Ib interval representation for these examples, we must
have $I_v \subseteq I_u$ and $I_v \cap I_w \neq \emptyset$. However, if both of these are true, some point of $I_v$ is in $I_w$ and since every point of $I_v$ is in $I_u$, it must be the case that $I_u \cap I_w \neq \emptyset$ and thus there must be some edge between $u$ and $w$. Thus there is no type Ib interval representation in this case.

If there exists a type Ib interval representation for this example, it must be the case that $I_u \subseteq I_v$ and $I_v \subseteq I_w$. However, this implies $I_u \subseteq I_v \subseteq I_w$, which implies there must be a dashed edge from $u$ to $w$. Thus there is no type Ib interval representation for this digraph.

For this example, any interval representation must have $I_a < I_b$, $I_b < I_c$, and $I_a < I_c$. This implies the order of the endpoints of these intervals must be

$$L(I_a) < L(I_b) < L(I_c) < R(I_a) < R(I_b) < R(I_c).$$
The dashed edges from $a$ to $d$ and $c$ to $d$ imply that $I_d \subseteq I_a$ and $I_d \subseteq I_c$ so $I_d \subseteq I_a \cap I_c$. Notice, however, that $I_a \cap I_c \subseteq I_b$, so $I_d \subseteq I_b$ implying there must be a dashed edge from $b$ to $d$. Since this is not the case, we conclude there is no type Ib interval representation for this digraph.

By similar logic to the previous example, any interval representation for this digraph must satisfy

\[ L(I_a) < L(I_b) < L(I_c) < R(I_a) < R(I_b) < R(I_c). \]

Note that this means $I_b \subseteq I_a \cup I_c$. Having a dashed edge from $b$ to $d$ implies $I_d \subseteq I_b$. However, this means $I_d \subseteq I_a \cup I_c$, so there must be some edge between $a$ and $d$ or $c$ and $d$. Thus there is not type Ib interval representation for the given digraph.
Bibliography


BIBLIOGRAPHY

property, interval graphs, and graph planarity using pq-tree algorithms,”


BIBLIOGRAPHY


122
BIBLIOGRAPHY


BIBLIOGRAPHY


Curriculum Vitae

Elizabeth Honda Reiland was born in Yorba Linda, California on August 26, 1989. She attended 1st-12th grade at the University School of Nashville, graduating in May 2007. In Summer 2006, she attended the Governor’s School for the Sciences program with a focus in Mathematics. This was her first exposure to proofs, and it spurred her to consider math as a major for college.

She attended Harvey Mudd College for her undergraduate studies, where she studied under Dr. Arthur Benjamin, who inspired in her a love of both teaching and math. In 2009 she received the Courtney S. Coleman Prize, and in 2010 received the Giovanni Borrelli Mathematics Prize from the HMC Math department. In 2011 she received a Bachelor of Science degree in Mathematics.

She next enrolled in the Applied Mathematics and Statistics Ph.D. program at Johns Hopkins University in 2013, where she researched random graph theory under the guidance of Dr. Edward Scheinerman. In 2016, she won the Joel Dean Award for Excellence in Teaching, and in summer of 2017 she taught Discrete Mathematics for the JHU summer session.