EXTENSIONS OF DOMINATION NUMBER AND THEIR DISTRIBUTION FOR RANDOM INTERVAL CATCH DIGRAPH FAMILIES

by

Enes Özel

A Thesis Submitted to the

Graduate School of Sciences and Engineering

in Partial Fulfillment of the Requirements for

the Degree of

Master of Science

in

Mathematics

Koç University

September 2011

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis

by

Enes Özel

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

examining committee have been made.

Thesis Committee Members:

Assoc. Prof. Elvan Ceyhan (Advisor) ....

Assoc. Prof. Mine Çağlar ....

Asst. Prof. Tevfik Metin Sezgin ....

Date: 13 September 2011

ABSTRACT

In this thesis we provide some extensions of the concept of domination in graph theory, namely, exact p-domination, weak p-domination and strong p-domination. We illustrate these new concepts on some random families of geometric graphs called class cover catch digraphs (CCCDs) and proximity catch digraphs (PCDs). PCDs and CCCDs are closely related to each other and have applications in pattern classification and spatial point pattern analysis. Furthermore, PCDs are parameterized by an expansion parameter and a centrality parameter. Previously, usual domination has been investigated thoroughly for these digraph families. We investigate the distribution of various extensions of domination number for these digraph families. In particular, we demonstrate that the asymptotic distribution of strong p-domination number is degenerate for CCCDs. We also derive the asymptotic distribution of the various domination number concepts for PCDs based on one dimensional uniform data. We also perform Monte Carlo simulation experiments, which support our theoretical findings. This study lays the foundation for the study of the various forms of domination on PCDs based on higher dimensional data.

ÖZET

Bu tezde çizge kuramsal baskınlık kavramının yeni versiyonlarını geliştiriyoruz. Bunlar, tam p-baskınlık, zayıf p-baskınlık ve güçlü p-baskınlık kavramlarıdır. Bu yeni kavramları, Küme Kapsayıcı Yakalama Yönlü Çizgeleri (KKYYÇler) ve Yakınlık Bölgesi Yakalama Yönlü Çizgeleri (YBYYÇler) olarak isimlendirdiğimiz bazı rassal geometrik çizge ailelerinde göstereceğiz. KKYYÇler ve YBYYÇlerin birbirleri ile yakın bir ilişkisi vardır ve bu rassal çizgelerin desen sınıflandırmaları ve uzaysal nokta desen analizinde uygulamaları vardır. Dahası, YBYYÇler genişleme ve merkez katsayıları ile parametrize edilmiştir. Literatürdeki baskınlık kavramı bu yönlü çizge aileleri üzerinde çalışılmıştır. Bu yönlü çizge aileleri üzerinde çeşitli baskınlık sayısı versiyonlarının dağılımlarını araştırdık. Özellikle, KKYYÇler için, güçlü p-baskınlık sayısının asimptotik dağılımının dejenere olduğu gösterilmiş ve tek boyutlu, düzgün dağılıma sahip verili YBYYÇler için çeşitli baskınlık sayılarının asimptotik dağılımını hesapladık. Teorik bulgular, Monte Carlo simülasyonları ile desteklenmiştir. Bu çalışma, çok boyutlu YBYYÇlerde çeşitli baskınlık formlarının analizine de temel oluşturacaktır.

To the memory of my grandfather, may he rest in peace.

ACKNOWLEDGEMENTS

I have an eternal gratitude to my advisor, Prof. Elvan Ceyhan, for his guidance, order and enthusiasm. At each step of my studies, I knew that he would be ready for help with accurate insight and without him, my efforts would carry little value. I am also indebted to all my professors in Mathematics Department of Koç University, through their knowledge and guidance, I started to figure out who I want to be, how can I become that person, but most importantly, I felt at home.

I thank Prof. Mine Çağlar and Prof. Tevfik Metin Sezgin for their effort, participating in my thesis committee.

Additionally I want to thank TÜBİTAK, for its financial support during my graduate studies.

My roommate, Adil Tolga Görgülü has been a great source of joint wisdom and happiness during my studies, I can never repay him. I am also grateful to my friends in the graduate office of mathematics students for the fun and healthy environment they created. I know all of them will be both successful and as entertaining as ever.

Last, but not the least, I want to thank my father, mother, two beautiful sisters and brother (and yes, the little one). I know that any speck of courage, dedication, responsibility and love I may have, is because of them. They are my warmth and laughter.

Contents

ABSTRACT **iii**

ÖZET **iv**

ACKNOWLEDGEMENTS **vi**

LIST OF SYMBOLS & ABBREVIATIONS **viii**

LIST OF FIGURES **ix**

1 PRELIMINARIES **1**

2 DOMINATION AND ITS EXTENSIONS **7**

2.1 Dominating Sets and Domination Number 7

2.2 Various Extensions of Domination 11

3 CLASS COVER CATCH DIGRAPHS, DOMINATION AND STRONG p-DOMINATION **16**

3.1 CCCDs and Previous Results on Domination 16

3.2 Strong p-domination on CCCDs 19

3.3 Asymptotic Distribution of $\gamma_{\geq p}$ on CCCDs 26

3.3.1 Method I 26

3.3.2 Method II 30

3.4 Monte Carlo Simulations 31

4 PROPORTIONAL EDGE PROXIMITY CATCH DIGRAPHS, DOMINATION AND STRONG p-DOMINATION **33**

4.1 PCDs and Previous Results on Domination 33

4.2 Exact and Weak p-Domination on PCDs 35

4.3 Strong p-domination on PCDs 38

4.4 Asymptotic Distribution of $\gamma_{\geq p}$ for PCDs 40

4.4.1 Proof of Theorem 4.12 44

4.4.2 Proof of Theorem 4.14 51

4.5 Monte Carlo Simulations 54

5 CONCLUSIONS **60**

REFERENCES **61**

LIST OF SYMBOLS & ABBREVIATIONS

G	Graph
D	Digraph
K_n	Complete graph with n vertices
$K_{n,m}$	Complete bipartite graph of size n and m
C_n	Cycle of length n
$V(G)$	Vertex set of the graph G
$E(G)$	Edge set of the graph G
$d(v)$	Degree of a vertex $v \in V(G)$, where G is a graph
$d^+(v), d^-(v)$	Outdegree and indegree of a vertex $v \in V(D)$; where D is a digraph
$\Delta(G)$	Maximum degree of the graph G
$\delta(G)$	Minimum degree of the graph G
$\gamma(G)$	The domination number of a graph G
$\gamma_{\geq p}(G)$	The strong p-domination number of a graph G, for $p \in [0,1]$
$\gamma_p^e(G)$	The exact p-domination number of a graph G, for $p \in [0,1]$
$\gamma_p^w(G)$	The weak p-domination number of a graph G, for $p \in [0,1]$
pdf	Probability Distribution Function
cdf	Cumulative Distribution Function
iid	independent and identically distributed
$X_{i:n}$	i^{th} order statistic of a random sample of size n
$CCCD$	Class Cover Catch Digraph
PCD	Proximity Catch Digraph
$PE-PCD$	Proportional-Edge Proximity Catch Digraph
$f(x) = O(g(x))$	$\lvert f(x)\rvert \leq Cg(x)$ for all $x < x_0$, where $C, x_0 \in \mathbb{R}$

List of Figures

1 An illustration of the pdf of $X_{i:n}$ given in Equation (2). . . . 3
2 An illustration of the joint pdf of $X_{i:n}, X_{j:n}$ given in Equation (3). 4
3 An illustration for the case of $\gamma(G) = n/2$. 8
4 An illustration of Remark 2.11. 11
5 A graph satisfying the conditions of Theorem 2.16. 13
6 An example regarding Theorem 2.18. 14
7 An example illustrating the construction of catch sets associated with the elements of a random sample. 16
8 The CCCD constructed from the points and sets in Figure 7. 17
9 An example of the interval, $\mathcal{I}_p$, in Theorem 3.10. 22
10 Another example of the interval, $\mathcal{I}_p$. 22
11 An example of proximity sets in $\mathcal{R}$. 34
12 The PCD constructed from the sets in Figure 11. 34

1 PRELIMINARIES

Here we briefly provide the basic probabilistic and graph theoretical concepts that we use throughout this study. All the definitions and statements on probability theory can be found in [1], [3], [4] and [5], on graph theory can be found in [8] and [2], with perhaps minor modifications undertaken for the sake of consistency.

Definition 1. A *probability space* is a measure space, (Ψ, Σ, P), where Ψ is a set, Σ is a specified σ-field, (i.e., the power set of Ψ) and P will be a nonnegative measure with the condition that $P(\Psi) = 1$.

Definition 2. A (real valued) *random variable* X is a measurable function with real values on a given probability space, i.e., $X : \Psi \to \mathbb{R}$.

Definition 3. Given a real valued random variable, a (cumulative) *distribution function* F is the function $F(x) = P(X \leq x)$. We use the abbreviation 'cdf' for cumulative distribution function.

Two instant properties of a distribution function are that $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$. Also notice that as P is nonnegative, F is monotone increasing.

Definition 4. Assuming that there exists a function f satisfying

$$F(x) = \int_{-\infty}^{x} f(t)\, dt,$$

we say that f is the *(probability) density function* of the random variable X. We use the abbreviation 'pdf' for probability density function.

Definition 5. Let $g : \mathbb{R} \to \mathbb{R}$ be any real valued function, then

$$\mathbf{E}(g(X)) = \int_{\Psi} g(X) dP = \int_{-\infty}^{\infty} g(x) dF(x)$$

is called the *expectation of g(X)*, we use the notation $\mu = \mathbf{E}(X)$. Also

$$\mathbf{Var}(X) = \sigma^2(X) = \mathbf{E}[(X-\mu)^2] = \mathbf{E}(X^2) - \mu^2$$

is called the *variance of X*.

Below we provide a slightly modified version of Theorem 2.4.6 from [1].

Theorem 1.1 (Markov's Inequality)**.** *For any continuous random variable X with pdf f, nonnegative real-valued function u and any constant $c > 0$, we have*

$$P\left(u(X) \geq c\right) \leq \frac{\mathbf{E}\left[u(X)\right]}{c}$$

Proof. Consider the set $S = \{x \in \mathbb{R} | u(x) \geq c\}$, then we have for X,

$$\mathbf{E}\left[u(X)\right] = \int_{\mathbb{R}} u(x).f(x)\,dx = \int_{S} u(x).f(x)\,dx + \int_{S^c} u(x).f(x)\,dx$$

$$\geq \int_S u(x).f(x)\,dx \geq \int_S c.f(x)\,dx = cP\left(X \in S\right) = c.P\left(u(X) \geq c\right).$$

$\square$

As a corollary to Theorem 1.1, we obtain Chebyshev's Inequality (Theorem 2.4.7 in [1]).

Corollary 1.2 (Chebyshev's Inequality)**.** *Let X be a given continuous random variable with $\mathbf{E}\left[X\right] = \mu$ and $\mathbf{Var}\left[X\right] = \sigma^2$, then for any $k > 0$ we have $P\left(|X - \mu| \geq k\sigma\right) \leq 1/k^2$.*

Proof. Consider the previous result with the function $u(X) = (X - \mu)^2$, $c = k^2\sigma^2$. Then

$$P\left(|X - \mu| \geq k\sigma\right) = P\left(|X - \mu| \geq k^2\sigma^2\right) \leq \frac{\mathbf{E}\left[(X - \mu)^2\right]}{k^2\sigma^2} = \frac{1}{k^2}.$$

$\square$

Remark 1.3. See that letting $\varepsilon = k\sigma$ we have the equivalent form

$$P\left(|X - \mu| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2}, \text{ whenever } \sigma^2 > 0. \quad \square \tag{1}$$

If we have more than one random variable to consider at the same time, we need to analyze their densities together as follows.

Definition 6. For the n-dimensional random variable, $X = (X_1, X_2, \ldots, X_n)$, the *joint probability density function* is

$$f(x_1, x_2, \ldots, x_n) = P\left(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\right),$$

for any possible value $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ of X.

Definition 7. The set of random variables $\{X_1, X_2, \ldots, X_n\}$ is called a *random sample of size n* from a random variable with the pdf f, if their joint pdf is of the form

$$f(x_1, x_2, \ldots, x_n) = f(x_1)f(x_2)\ldots f(x_n),$$

that is, $\{X_1, X_2, \ldots, X_n\}$ is a random sample of size n, if X_i are independent and identically distributed (iid), with pdf f.

Most of the basic concepts we need in our treatment of our main problem are provided, we will mostly work with random samples in the real line and use their natural ordering.

Definition 8. Given a random sample of size n from a real valued random variable, $\mathcal{X}_n$, if we order the observations from the smallest to largest, we call the random variables the *order statistics* of $\mathcal{X}_n$ and the i^{th} greatest observation in $\mathcal{X}_n$ the i^{th} *order statistics*. We denote i^{th} order statistics by $X_{i:n}$.

We need the pdf of the i^{th} order statistics, for any $i \in \{1, 2, \ldots, n\}$, as well as the joint pdf of two distinct order statistics, say the i^{th} and j^{th}, where $i, j \in \{1, 2, \ldots, n\}$ and $i \neq j$. Thus we give the following propositions (Theorem 6.5.2 in [1]).

Proposition 1.4. *Given $a, b \in \mathbb{R}^+$ a random sample $\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\}$ from a real valued random variable with the continuous pdf, f, where $f(x) > 0$ for $a < x < b$, and the cdf F, the pdf of the i^{th} order statistic $X_{i:n}$ is given by*

$$g_i(x_i) = \frac{n!}{(i-1)!(n-i)!} \left[F(x_i)\right]^{i-1} \left[1 - F(x_i)\right]^{n-i} f(x_i) \tag{2}$$

if $a < x_i < b$, and zero otherwise.

Figure 1 illustrates the rationale behind Proposition 1.4. Observe that x_i is fixed, we need $i-1$ points before x_i and $n-i$ points after x_i.

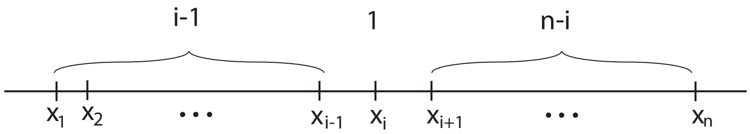

Figure 1: An illustration of the pdf of $X_{i:n}$ given in Equation (2).

Proposition 1.5. *Let $\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\}$ be as in Proposition 1.4, then the joint pdf of the i^{th} and the j^{th} order statistics, $X_{i:n}$ and $X_{j:n}$, is given by*

$$g_{i,j}(x_i, x_j) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \left[F(x_i)\right]^{i-1} f(x_i) \left[F(x_j) - F(x_i)\right]^{j-i-1} f(x_j) \left[1 - F(x_j)\right]^{n-j} \tag{3}$$

if $a < x_i < x_j < b$, and zero otherwise.

The rationale behind Proposition 1.5 is illustrated in Figure 2. Observe that x_i and x_j are fixed, we need $i-1$ points before x_i, $j-i-1$ points in between x_i and x_j as well as $n-j$ points after x_j.

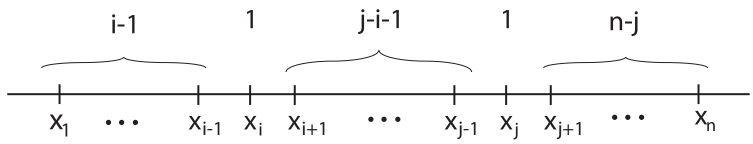

Figure 2: An illustration of the joint pdf of $X_{i:n}, X_{j:n}$ given in Equation (3).

More detail on order statistics and their analysis can be found in [5]. We also provide the concept of convergence of random variable sequences.

Definition 9. Let R_n be a sequence of real valued random variables, with distribution functions F_n. R_n is said to *converge in distribution* to R, with distribution function F, if for any real number $x \in \mathbb{R}$,

$$\lim_{n\to\infty} F_n(x) = F(x),$$

denoted $R_n \xrightarrow{D} R$.

Definition 10. Let R_n be a sequence of real valued random variables. R_n is said to *converge in probability* to R, if for any $\varepsilon > 0$

$$\lim_{n\to\infty} P\left(|R_n - R| < \varepsilon\right) = 1,$$

denoted $R_n \xrightarrow{P} R$.

Finally, we state the following theorem, it is stated as in [6].

Theorem 1.6 (Lebesgue's Dominated Convergence Theorem)**.** *Let $\{f_n\}$ be a sequence of integrable functions such that, (a) $f_n \to f$ a.e., and (b) there exists a nonnegative integrable function g such that $|f_n| < g$ a.e. for all $n \in \mathbb{N}$. Then f is integrable, and*

$$\int f = \lim_{n\to\infty} \int f_n.$$

In order to be able to study domination on graphs, we also need to introduce some graph theoretical tools. A *graph* G is a triple consisting of a vertex set $V(G)$, an edge set $E(G)$ and a relation that associates each edge with an unordered pair of vertices. The two vertices in the pair that is associated with an edge are called its *endpoints*. If two vertices are endpoints of an edge, then we call them *adjacent* to each other, and the edge is called *incident* to those vertices. A *digraph* (or a *directed graph*) is a triple consisting of a vertex set $V(G)$, an edge set $E(G)$ and a relation that associates each edge with an ordered pair of vertices. In digraphs, order is important, i.e., it is possible that a vertex is adjacent to another, but not vice versa. These directed edges of digraphs are usually called *arcs*.

A *loop* is an edge that has the same vertex as endpoints. If there are two vertices associated with more than one edge, we call those edges *multiple edges*. A graph is called *simple* if it has no loops or multiple edges. The number of vertices of G, $|V(G)|$ is called the *order* of G.

Note that we use the convention in which an edge of a graph, say e, is also denoted as $u\,v$, if u and v are its endpoints. Also in digraphs we denote an arc a as (u, v), if it starts from u and ends in v. We also denote a graph as $G = (\mathscr{V}, \mathscr{E})$ and a digraph as $D = (\mathscr{V}, \mathscr{A})$, for brevity of notation.

Definition 11. Given a graph G and a vertex $v \in V(G)$, the number of incident edges to v is called the *degree of* v and is denoted $d(v)$ (or $d_G(v)$). For a digraph D and a vertex $v \in V(D)$, the number of arcs going out from v is called its *outdegree*, denoted $d^+(v)$ and the number of arcs going into v is called its *indegree*, denoted $d^-(v)$.

Given a graph G, with vertices $V(G) = \{v_1, v_2, \ldots, v_n\}$, we call the set $\{d(v_1), d(v_2), \ldots, d(v_n)\}$ the *degree sequence* of G, and conventionally degree sequences are assumed to be ordered from the largest to the smallest degree. The maximum degree of vertices of G is denoted as $\Delta(G)$, whereas the minimum degree of its vertices is denoted as $\delta(G)$. In this sense, the degree sequence of G has the following form $\{\Delta(G), \ldots, \delta(G)\}$, where for some $i, j \in \{1, 2, \ldots, n\}$, $\Delta(G) = d(v_i)$ and $\delta(G) = d(v_j)$.

Definition 12. A *subgraph* of a graph $G = (V(G), E(G))$ is a graph G' with the vertex set $V'(G) \subset V(G)$ and with the edge set $E'(G) \subset E(G)$ under the condition that if an edge $e = u\,v$ of G is an edge of G', then $u, v \in V'(G)$. If G' is a subgraph of G, then we say $G' \subset G$. A *spanning subgraph* of G is a subgraph with the vertex set $V(G)$.

Definition 13. A *path* is a graph whose vertices can be ordered in a list so that two of its vertices are adjacent if and only if they are consecutive in the list. A *cycle* is a path which does not include any vertex twice, except for the first and the last one in the list. The number of edges in a path is called the *length* of the path.

Definition 14. The *distance* between two vertices u and v is the smallest length of any path between u and v. It is denoted as $d(u, v)$. If there exists no path between u and v, we say that $d(u, v) = \infty$. Given any vertex v, $d(v, v)$ is assumed to be 0. If, for all $u, v \in V(G)$, $d(u, v) < \infty$, then we say that G is *connected*. All graphs consist of connected parts that are called *components*.

Definition 15. A *tree* is a connected graph with no cycles. A *spanning tree* of G is a spanning subgraph of G, which is also a tree.

2 DOMINATION AND ITS EXTENSIONS

In this section we analyze the graph theoretical concept of domination and give previous results on domination. Then we introduce new extensions to this concept and provide several new results on them.

2.1 Dominating Sets and Domination Number

We define the concept of domination on graphs and digraphs, then we present results mostly on lower and upper bounds of the domination number. All of the results in this section can be found in [7], which includes a nice compilation of the results on domination. First we define the concept of domination in the usual sense.

Let $G = (\mathscr{V}, \mathscr{E})$ be a simple graph with vertex set $\mathscr{V}$ and (undirected) edge set $\mathscr{E}$. A vertex u of G is said to *dominate* itself and vertex v if there exists an edge $uv \in \mathscr{E}$. A subset $S_D \subseteq \mathscr{V}$ is said to be a *dominating set* of the graph G, if for any vertex $v \in \mathscr{V}$, either $v \in S_D$, or there exists a vertex $u \in S_D$ that dominates v. Clearly, $\mathscr{V}$ is a dominating set of G. A *minimum dominating set* S_D^* of G is a dominating set with the minimum cardinality. Also, a *minimal dominating set* S_D^* of G is a dominating set that does not have another dominating set as a proper subset. Notice any minimum dominating set is minimal, but not vice versa.

Theorem 2.1. *Every connected graph G, with the vertex set $\mathscr{V}$, of order $n \geq 2$ has a dominating set S_D, whose complement $\mathscr{V} - S_D$ is also a dominating set.*

Proof. Since G is connected, it has at least one spanning tree. Let T be any of these spanning trees of G, and let u be any vertex in $\mathscr{V}$. Then the vertices in T fall into two disjoint sets, S_D and S'_D consisting, respectively, of the vertices with an even and odd distance from u in T. This is due to the fact that there are no cycles in T, thus between any vertices u and v, there is a unique path. Clearly, both S_D and $S'_D = \mathscr{V} - S_D$ are dominating sets for G. □

Theorem 2.2. *If a graph G is without any isolated vertices, then the complement of $\mathscr{V} - S_D$ of every minimal dominating set S_D is a dominating set.*

Proof. Let S_D be any minimal dominating set of G. Assume that there exists a vertex $u \in S_D$ that is not dominated by any vertex in $\mathscr{V} - S_D$. Since G has no isolated vertices, u must be dominated by at least one vertex in $S_D - \{u\}$, that is, $S_D - \{u\}$ is a dominating set, contradicting the minimality of S_D. Thus every vertex in S_D is dominated by at least one vertex in $\mathscr{V} - S_D$, and also all elements of $\mathscr{V} - S_D$ dominate themselves, therefore $\mathscr{V} - S_D$ is a dominating set. □

For the digraphs, another term for domination was used in [7], namely *directed domination*. But as there is little difference in this term, we use the same term of domination both for graphs and digraphs. The definition for the directed graphs is very similar. Let $D = (\mathscr{V}, \mathscr{A})$ be a digraph with vertex set $\mathscr{V}$ and arc set $\mathscr{A}$. A vertex u of D is said to *dominate* vertex v if an arc $(u, v) \in \mathscr{A}$ exists. The definition of a dominating set and a minimum dominating set is the same as in the graphs.

Now given a graph G, consider the class of minimum dominating sets. Each set in this class has the same cardinality and this number is called the *domination number* of G and it is denoted $\gamma(G)$. Equivalently, $\gamma(G)$ is the minimum number of vertices required to dominate all vertices of G. As each vertex dominates itself, we directly have the upper bound, $\gamma(G) \leq n$. The following theorem gives a simple upper bound for the domination number of a wide class of graphs.

Theorem 2.3 (Ore). *If a graph G has no isolated vertices, then* $\gamma(G) \leq n/2$.

Proof. By the fact that no vertex of G is isolated, we see that any vertex dominates at least one other vertex, i.e., counting itself, it dominates at least two vertices. Therefore the result follows. $\square$

Let K_n be the simple graph on n vertices with each vertex being of degree $n-1$ (called the *complete graph on n vertices*). Note that the upper bound on the previous theorem is sharp, e.g., if the graph only consists of components that are K_2s, then the bound is obtained. See that in Figure 3 for each K_2 component we choose one vertex and obtain domination. As there are $n/2$ components, we have $\gamma(G) = n/2$. (Note the implicit assumption that n is even, and odd case is similar.)

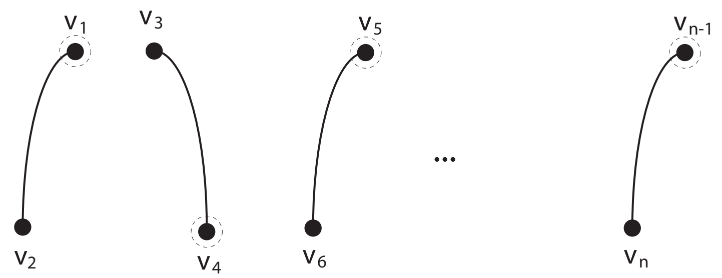

Figure 3: An illustration for the case of $\gamma(G) = n/2$.

Moreover, a more general result on the sharpness of the upper bound is given. We denote cycles of length k by C_k. By the *corona* $G \circ G'$ we mean that n copies of the graph G' are taken and each copy is made adjacent to a vertex of G. In particular $G \circ K_1$ is the graph where at each vertex of

G there is another edge and a vertex added. Note that K_2 is the corona, $K_1 \circ K_1$.

Theorem 2.4. *For a graph G with even order n and no isolated vertices, $\gamma(G) = n/2$ if and only if components of G are the cycle C_4 or the corona $H \circ K_1$ for any connected graph H.*

Increasing the minimum degree requirement, we further obtain tighter bounds.

Theorem 2.5 (Reed)**.** *If G is a connected graph with $\delta(G) \geq 3$, then $\gamma(G) \leq 3n/8$.*

For graphs with higher degree vertices, the following result is obtained.

Theorem 2.6. *For any graph G with $\delta(G) \geq 7$, we have*

$$\gamma(G) \leq n \left[1 - \delta \left(\frac{1}{\delta + 1} \right)^{1+1/\delta} \right].$$

We denote the *complete bipartite graph* defined on $n + m$ vertices with $K_{n,m}$ and it is the graph such that its vertices can be partitioned into two sets of sizes n and m, and that no two vertices that are in the same set are adjacent, whereas a vertex in a set is adjacent to all vertices in the other set. The graph $K_{1,3}$ is called a *claw.* Also the graph $K_3 \circ K_1$ is called a *net.* For a real number x, let $\lceil x \rceil$ be the smallest integer larger than or equal to x and $\lfloor x \rfloor$ be the largest integer smaller than or equal to x.

Theorem 2.7. *If a connected graph G is without any claws or nets, then $\gamma(G) \leq \lceil n/3 \rceil$.*

The following is a quite interesting result.

Theorem 2.8 (Weber)**.** *Let $k = \lfloor (\log(n) - 2\log(\log(n)) + \log(\log(e))) \rfloor$. Then for almost every graph G,*

$$k + 1 \leq \gamma(G) \leq k + 2.$$

The following is an intuitive result on bounds regarding the maximum degree of a graph.

Theorem 2.9. *For any graph G,*

$$\left\lceil \frac{n}{1 + \Delta(G)} \right\rceil \leq \gamma(G) \leq n - \Delta(G).$$

Proof. Let S_D be a minimum dominating set of G. See that each vertex can dominate at most itself and $\Delta(G)$ other vertices. So it follows that

$\gamma(G) \geq \lceil \frac{n}{1+\Delta(G)} \rceil$. Also see that there is a vertex with degree $\Delta(G)$, which also dominates itself. All the $n - \Delta(G) - 1$ vertices that cannot be dominated by the initial vertex can dominate themselves and adding the first vertex we obtain domination in $n - \Delta(G)$ vertices, which is not necessarily a minimum domination. Hence $\gamma(G) \leq n - \Delta(G)$. $\square$

The lower bound in this result can also be attained, i.e., $\gamma(G) = n/(1 + \Delta(G))$ if and only if there is a minimum dominating set S_D of G such that for any $u, v \in S_D$, u and v do not have any common vertex that is adjacent to both of them as well as if all vertices of S_D are of degree $\Delta(G)$. A complete bipartite graph, $K_{1,n-1}$, is called a *star* of size n. k copies of cycles of length 3, C_3's, and stars of size $\Delta(G)$, $K_{1,\Delta(G)-1}$ attain this bound, i.e., $\gamma = k = n/(1 + \Delta(G))$. Also note that the upper bound is attained for the graph $K_k \circ K_1$. The following result is for a graph with a given degree sequence.

Theorem 2.10. *If a graph G has the degree sequence $\{d_1, d_2, \ldots, d_n\}$, with $d_i \geq d_{i+1}$ for $i = 1, 2, \ldots, n-1$, then $\gamma(G) \geq \min(k : k+(d_1+d_2+\ldots+d_k) \geq n)$.*

Proof. We will construct a dominating set S_D in the following way. We add v_{d_1} to S_D, the vertex corresponding to the highest degree, d_1. We are done if $d_1 = n - 1$, otherwise we add v_{d_2} to S_D, the vertex corresponding to the degree d_2, and if v_{d_2} dominates all the $n - d_1 - 1 > 0$ vertices that v_{d_1} does not dominate, then we are done. Otherwise we continue adding vertices. This procedure is certain to stop as at the worst case all the points are isolated, i.e., $d_1 = d_2 = \ldots = d_n = 0$, therefore at $k = n$ we have the domination. $\square$

Remark 2.11. Note that the method in the proof of this theorem does not usually yield a minimum dominating set, it just gives a lower bound for the domination number. This is since for $i < j$, v_{d_i} may dominate vertices that were already dominated by v_{d_j} and instead of v_{d_i}, choice of another vertex could be more optimal. $\square$

An example for this remark is given in Figure 4, notice that the graph has the degree sequence $\{6, 4, 2, 2, 2, 2, 2, 1, 1, 1, 1\}$, and $n = 10$, thus by Theorem 2.10 we have $\gamma(G) \geq 2$. But the vertex set $\{v_{d_1}, v_{d_2}\}$ is not a dominating set, whereas if we choose v_{d_6}, we get the desired result, i.e., $\{v_{d_1}, v_{d_6}\}$ is a minimum dominating set.

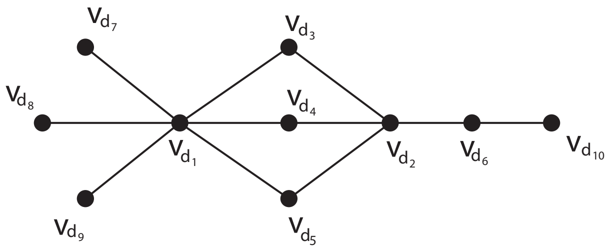

Figure 4: An illustration of Remark 2.11.

We state the following theorem, and its proof follows easily.

Theorem 2.12. *If a graph G has stars of size $d_1, d_2, \ldots, d_k$ as components, then the lower bound for $\gamma(G)$ in Theorem 2.10 is obtained.* □

Theorem 2.13. *For any graph G,*

$$\gamma(G) \leq \left(n + 1 - (\delta(G) - 1)\frac{\Delta(G)}{\delta(G)} \right) \Big/ 2.$$

And lastly, the following result is a quick corollary of the previous theorem.

Corollary 2.14. *If a graph has no isolated vertices, then we have*

$$\gamma(G) \leq \frac{n + 2 - \delta(G)}{2}.$$

2.2 Various Extensions of Domination

Here we define some extensions of the concept of domination, state and prove some related results. Note that there are quite a number of previous extensions of domination, namely, multiple domination, total domination, upper/lower domination, independent domination and distance domination. As they are not in the scope of this study, we do not define them. For more information see [7].

A new and really intuitive extension to the concept of domination is the exact p-domination.

Definition 16. Let G be any graph and $p \in [0, 1]$, we say that $S_D^p \subset V(G)$ is an *exact p-dominating set*, if S_D^p exists so that its vertices dominate **exactly** $\lceil np \rceil$ of all vertices of G. The cardinality of a minimum exact p-dominating set is called the *exact p-domination number*, and it is denoted $\gamma_p^e(G)$.

Notice that the case $p = 1$ gives the usual domination as we need the domination of all n vertices, whereas the case $p = 0$ is the null case, i.e., we need to dominate no vertices, so we assume γ_0^e to be 0. As intuitive as it may be, for a lot of different graphs, exact p-dominating sets, and therefore exact p-domination number may not exist. Consider the complete graph, K_n. For any value p that makes $\lceil np \rceil / n$ other than 0 or 1, there are no exact p-dominating sets and therefore $\gamma_p^e(K_n)$ is not defined. This is due to the fact that if you even choose one vertex, then you obtain the domination of all vertices. The graphs that give exact p-domination for any value of $p \in [0, 1]$ are not common.

We use the following elementary result.

Theorem 2.15 (Binary Expansion Theorem)**.** *For each natural number $n \in \mathbb{N}$, there is a unique expansion of the form*

$$n = \sum_{i=0}^{k} a_i 2^i,$$

where $k \in \mathbb{N}$ and a_i is either 0 or 1 for all $i \in \{0, 1, \ldots, k-1\}$, and $a_k = 1$.

Note that the condition $a_k = 1$ is not the case when $n = 0$. We have the following result.

Theorem 2.16. *For all $n \in \mathbb{N}$, of the form $n = 1 + 2 + 2^2 + \ldots + 2^k$, where $k \in \mathbb{N}$, there exists a graph of order n such that exact p-domination is well-defined, independent from the choice of p. Moreover, an example of such graphs is the graph with the components $K_1, K_2, \ldots, K_{2^k}$.*

Proof. Given a natural number $z \in \{0, 1, \ldots, n\}$, by Theorem 2.15, there is the binary expansion $z = \sum_{i=0}^{j} a_i 2^i$, where $j \in \mathbb{Z}^+$ and a_i are as in the theorem. Also $j \leq k$. We construct the exact p-dominating set as follows. For $i \in \{0, 1, \ldots, j\}$ such that $a_i = 1$, we choose any vertex in the component K_{2^i} and add it to the set. Then we have a total of z vertices dominated, and as this is true for any $z \in \{0, 1, \ldots, n\}$, the result follows. $\square$

We give an example in Figure 5. This is the graph with the components K_1, K_2, K_4 and K_8. For each value of p, $\lceil np \rceil$ changes from 0 to 15, taking any integer value between these two. By the construction of this graph, we can obtain, for all p, the exact p-domination number, and it is at most 4.

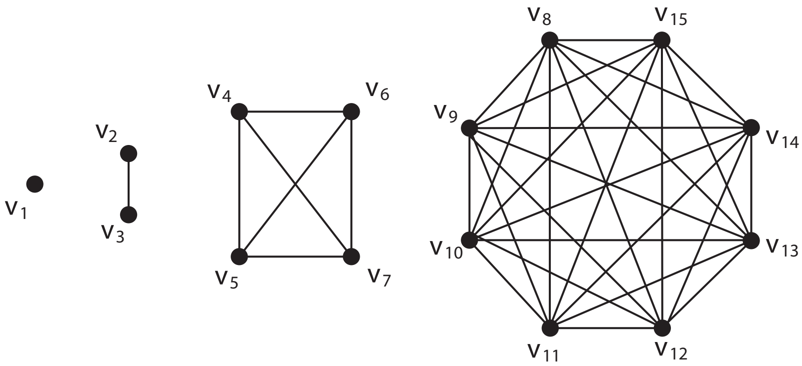

Figure 5: A graph satisfying the conditions of Theorem 2.16.

There is a more general result, which follows easily.

Theorem 2.17. *For any graph G of order n, exact p-domination is defined if and only if there is a set of k vertices of G, which together dominate $\lceil np \rceil - k$ vertices outside of this set.* $\square$

Now we define our main extension of domination, in which we loosen the 'exactness' condition in exact p-domination.

Definition 17. Let G be any graph and $p \in [0,1]$, we say that $S_D^p \subset V(G)$ is a *strong p-dominating set*, if vertices in S_D^p dominate **at least** $\lceil np \rceil$ of all vertices of G. The cardinality of a minimum strong p-dominating set is called the *strong p-domination number*, and it is denoted $\gamma_{\geq p}(G)$.

Notice, again, that the case $p = 1$ gives the usual domination as we need the domination of at least all n vertices, i.e., all vertices. Also notice that whenever exact p-domination makes sense, any exact p-dominating set is a strong p-dominating set, but not vice versa. Hence follows the result.

Theorem 2.18. *For any graph G and $p \in [0,1]$, whenever $\gamma_p^e(G)$ is well-defined, we have*

$$\gamma_{\geq p}(G) \leq \gamma_p^e(G). \quad \square$$

We give an example of this fact in Figure 6. See that in this graph $n = 10$ and for $p = 0.6$, $\gamma_{0.6}^e = 3$ (with the minimum exact p-dominating set $\{v_2, v_8, v_9\}$). On the other hand we can dominate 7 vertices with just 2, i.e., $\gamma_{\geq 0.6} = 2$ (with the minimum strong p-dominating set $\{v_2, v_5\}$).

Furthermore, observe that the strong p-domination number of a given graph G is increasing with respect to the parameter p, that is, the function $\varsigma_G : [0,1] \to \mathbb{R}$, $\varsigma_G(p) = \gamma_{\geq p}(G)$ is monotone increasing.

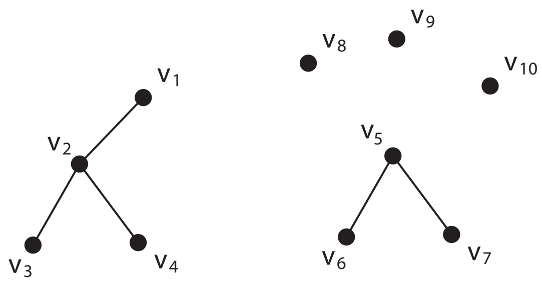

Figure 6: An example regarding Theorem 2.18.

Theorem 2.19. *For $p_1, p_2 \in [0,1]$ with $p_1 \leq p_2$, we have*

$$\gamma_{\geq p_1}(G) \leq \gamma_{\geq p_2}(G),$$

for any graph G.

Proof. See that any strong p_2-dominating set $S_D^{p_2}$, which dominates at least $\lceil \mathrm{n} \cdot p_2 \rceil$ vertices, certainly dominates at least $\lceil \mathrm{n} \cdot p_1 \rceil$ vertices, or, equivalently, any strong p_2-dominating set is a strong p_1-dominating set. This implies that $\gamma_{\geq p_1}(G) \leq \gamma_{\geq p_2}(G)$. □

Now we provide two results on the upper and lower bounds for $\gamma_{\geq p}$.

Theorem 2.20. *Let G be a graph with no isolated vertices and $p \in [0,1]$. Then*

$$\gamma_{\geq p}(G) \leq np/2.$$

Proof. Assuming G has no isolated vertices, each vertex is adjacent to at least one other vertex. We take p pairs of different such vertices, from each pair choose one vertex and obtain strong p-domination. Note the vertices we chose may not constitute a minimum strong p-dominating set. Hence follows the upper bound. □

Theorem 2.21. *Let G be a graph of size n, $p \in [0,1]$, with the degree sequence $\{d_1, d_2, \ldots, d_n\}$. Then we have the lower bound,*

$$\gamma_{\geq p} \geq \min\left(k : k + (d_1 + d_2 + \ldots + d_k) \geq np\right).$$

Proof. Proof is similar to the proof of Theorem 2.10, replacing n in the equation with np. □

Note that a result similar to Theorem 2.12 can also be given for strong p-domination. Finally, we have the following result, analogous to Theorem 2.9 and its proof is similar.

Theorem 2.22. *For any graph G of order n and $p \in [0,1]$, we have the following upper and lower bounds,*

$$\left\lceil \frac{\lceil np \rceil}{1+\Delta(G)} \right\rceil \leq \gamma_{\geq p}(G) \leq \lceil np \rceil - \Delta(G). \quad \square$$

Now we define an extension to exact p-domination that is well-defined in some cases, where exact p-domination is not well-defined.

Definition 18. Let $\mathscr{G} = (G_n)_{n\in\mathbb{N}}$ be a sequence of graphs with the corresponding orders, $n \in \mathbb{N}$. Also let $p \in [0,1]$ and $(p_n)_{n\in\mathbb{N}}$ a sequence of real numbers converging to p, with the property that for each $n \in \mathbb{N}$, there is a set $(S_D^{p_n})_{n\in\mathbb{N}} \subset V(G_n)$ that is a minimum exact p_n-dominating set of G_n. We call this sequence of vertex sets, $(S_D^{p_n})_{n\in\mathbb{N}}$ an asymptotically accurate *weak p-dominating class*. The limit of the sequence $(|S_D^{p_n}|)_{n\in\mathbb{N}}$ as n goes to infinity is called (asymptotically accurate) *weak p-domination number*, whenever it exists, and it is denoted $\gamma_p^w(\mathscr{G})$.

As complex as it may seem, (asymptotically accurate) weak p-domination number is of great use in the analysis of asymptotic behavior of random graphs that we investigate.

3 CLASS COVER CATCH DIGRAPHS, DOMINATION AND STRONG p-DOMINATION

In this section we introduce Class Cover Catch Digraphs (CCCDs), investigate the concept of domination, give previous results on the distribution of the domination number and present new results about the distribution of the strong p-domination number of CCCDs. CCCDs were first introduced in Priebe et al. [14] to construct a new method in statistical classification. They provided the distribution of the domination number of CCCDs constructed with two classes and with the uniform distribution over a bounded interval in $\mathbb{R}$. In [10], the distribution of the domination number was presented without the uniformness restriction on the two classes of points in $\mathbb{R}$.

3.1 CCCDs and Previous Results on Domination

Consider $\mathcal{Y}_m = \{y_1, y_2, \ldots, y_m\} \subset \mathbb{R}^q$, $m, q \in \mathbb{N}$. Given $\mathcal{Y}_m$ and $n \in \mathbb{N}$ assume $\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\}$ is a random sample of size n from a continuous distribution F in $\mathbb{R}^q$. For any $i \in \{1, 2, \ldots, n\}$ we associate the X_i with the ball (called *catch set*) centered around it, namely $B_i(X_i, r(X_i)) = \{z \in \mathbb{R}^q : d(z, X_i) < r(X_i)\}$, where $r(x) = d(x, \mathcal{Y}_m) = \min_{y \in \mathcal{Y}_m} d(x, y)$, i.e., the minimum distance of the a point x to any point in $\mathcal{Y}_m$. Here, $d : \mathbb{R}^q \times \mathbb{R}^q \to [0, \infty)$ is any distance function, such as Euclidian distance.

In Figure 7 we give an example in $\mathbb{R}^2$. There are two classes of points, namely $\mathcal{Y}_4 = \{y_1, y_2, y_3, y_4\}$ and $\mathcal{X}_3 = \{x_1, x_2, x_3\}$, and each ball, B_i associated with the element x_i of $\mathcal{X}_3$ is centered at x_i and has the minimum distance from that point to any point in $\mathcal{Y}_4$, $r(x_i)$, as its radius.

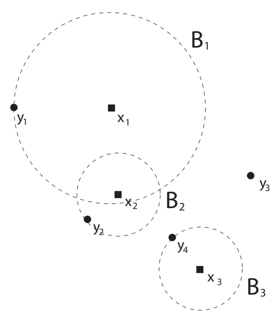

Figure 7: An example illustrating the construction of catch sets associated with the elements of a random sample.

Now we define the CCCDs.

Definition 19. *The class cover catch digraph* (CCCD) corresponding to the 'target set', $\mathcal{X}_n$ and the set $\mathcal{Y}_m$ is a simple digraph with n vertices, say $x_1, x_2, \ldots, x_n$, and such that there is an arc from x_i to x_j, as shown in Figure 8, if $x_j \in B_i(x_i, r(x_i))$, for any $1 \leq i, j \leq n$, where the ball B_i is defined as above.

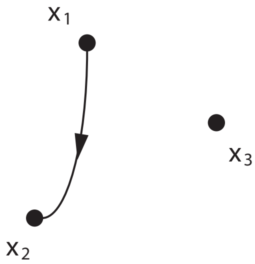

Figure 8: The CCCD constructed from the points and sets in Figure 7.

In [9], random graphs are classified with respect to the source of their randomness. There are *edge random graphs*, which are the first random graphs introduced by Erdös and Renyí [3], *vertex random graphs* and *vertex-edge random graphs*. In this sense, CCCDs are vertex random graphs, as their randomness lies in the vertices and their respective positions, i.e., once the vertices are given, edges are given by deterministic treatment of the vertices.

In this study we work on the case $q = 1$, i.e., the sets $\mathcal{X}_n$ and $\mathcal{Y}_m$ lie on the real line. In this case B_is are intervals centered at x_is, with the radii $d(x_i, \mathcal{Y}_m)$. For the case, $q = 1$, we have the order statistics of $\mathcal{X}_n$, $\{X_{1:n}, X_{2:n}, \ldots, X_{n:n}\}$. We follow the convention that random variables are denoted with upper case letters, while their realizations are with lower case letters.

Remark 3.1. Note that whenever a graph or a digraph is random, the various domination numbers as γ, $\gamma_{\geq p}$, γ_p^e and γ_p^w become random variables instead of just real numbers.

First, we present the previous results on CCCDs and their domination number. Priebe et al. proved the following results in [14].

Theorem 3.2. *Let D be a CCCD constructed with the sets $\mathcal{Y}_2 = \{y_1, y_2\}$, ($y_1, y_2 \in \mathbb{R}$, $y_1 < y_2$) and $\mathcal{X}_n$, a random sample of size n from the uniform* *distribution, $\mathcal{U}(y_1, y_2)$. Then the random variable $\gamma(D)$ has the distribution $1 + Bernoulli(1 - \vartheta(n))$, where*

$$\vartheta(n) = 5/9 + 4^{(1-n)} 4/9. \tag{4}$$

Here we let $1\{.\}$ be the indcator function, $Z_m = \{0, 1, \ldots, m-1\}$ and define

$$\Delta^S_{z,b} = \left\{ (z_1, z_2, \ldots, z_b) : \sum_{i=1}^{b} z_i = z; z_i \in S,\ \forall i \right\}.$$

Theorem 3.3. *Let D be the CCCD constructed with the sets $\mathcal{Y}_m$, an m-element subset of $\mathbb{R}$ and $\mathcal{X}_n$, a random sample of size n from the distribution $\mathcal{U}(y_1, y_m)$, where y_1 is the smallest and y_m is the largest element of $\mathcal{Y}_m$. Then the pdf of $\gamma(D)$ is given by*

$$f(d) = \frac{n!\, m!}{(n+m)!} \sum_{\vec{n} \in \Delta^{Z_{n+1}}_{n,m+1}} \sum_{\vec{d} \in \Delta^{Z_3}_{d,m+1}} \alpha(d_1, n_1)\, \alpha(d_{m+1}, n_{m+1}) \prod_{j=2}^{m} \beta(d_j, n_j),$$

where

$$\alpha(d, n) = \max(1\{n = d = 0\}, 1\{n \geq d = 1\})$$

and

$$\beta(d, n) = \max(1\{n = d = 0\}, 1\{n \geq d \geq 1\})\, \vartheta(n)^{1\{d=1\}}\, (1 - \vartheta(n))^{1\{d=2\}}.$$

Although this expression seems a bit complicated, it is intuitive and the theorem gives a comprehensive result. The following theorem gives the expectation of $\gamma(D)$.

Theorem 3.4. *Let D be a CCCD as above, then*

$$\mathbf{E}\left[\gamma(D)\right] = \frac{2n}{n+m} + \frac{n! m(m-1)}{(n+m)!} \sum_{i=1}^{n} \frac{(n+m-i-1)!}{(n-i)!} (2 - \vartheta(i)),$$

where $\vartheta(i)$ is defined as in Equation (4).

The next theorem gives the asymptotic distribution of $\gamma(D)$, depending on either n or m is held fixed.

Theorem 3.5. *Considering the CCCD, D, constructed with a set of cardinality m, $\mathcal{Y}_m \subset \mathbb{R}$ and the set $\mathcal{X}_n$ of random sample of size n from $\mathcal{U}(y_1, y_m)$ distribution, for fixed $n \in \mathbb{Z}^+$ we have*

$$\lim_{m \to \infty} \gamma(D) = n,$$

as well as for fixed $m \in \mathbb{Z}^+$, $\lim_{n\to\infty} \gamma(D)$ *has the distribution* $m + 1 + B$, *where* $B \sim Binomial(m-1, 4/9)$.

So far, the domination number was independent from the support of the distribution of the random sample, because $\mathcal{X}_n$ was a random sample from a uniform distribution. In [10] (Theorem(5.1)), the next result was proven and it does not require the distribution to be uniform, in fact, for $y_1, y_2 \in \mathbb{R}$, with $y_1 < y_2$, and for given $\varepsilon \in (0, (y_1 + y_2)/2)$, let

$$\mathcal{F}(y_1, y_2, \varepsilon) = \{F : (y_1, y_1+\varepsilon)\cup(y_2-\varepsilon, y_2)\cup((y_1+y_2)/2-\varepsilon, (y_1+y_2)/2+\varepsilon) \subset (y_1, y_2)\}.$$

For any distribution function with support as above, $F \in \mathcal{F}(y_1, y_2, \varepsilon)$, next result gives the distribution of the domination number, as $n \to \infty$.

Theorem 3.6. *Let* $\mathcal{Y}_2 = \{y_1, y_2\} \subset \mathbb{R}$ *with* $-\infty < y_1 < y_2 < \infty$. $\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\}$ *with* $X_i \overset{iid}{\sim} F \in \mathcal{F}(y_1, y_2, \varepsilon)$. *Let* D *be the CCCD constructed with* $\mathcal{Y}_2$ *and* $\mathcal{X}_n$. *Suppose* $k \geq 0$ *is the smallest integer for which* $F(\cdot)$ *has continuous right derivatives up to order (k+1) at* y_1, $(y_1 + y_2)/2$, $f^{(k)}(y_1^+)+2^{-(k+1)}f^{(k)}\left(\left(\frac{y_1+y_2}{2}\right)^+\right) \neq 0$ *and* $f^{(j)}(y_1^+) = 0$ *for all* $j = 0, 1, \ldots, k-1$*; and* $\ell \geq 0$ *is the smallest integer for which* $F(\cdot)$ *has continuous left derivatives up to order* $(\ell+1)$ *at* y_2, $(y_1+y_2)/2$, $f^{(\ell)}(y_2^-)+2^{-(\ell+1)}f^{(\ell)}\left(\left(\frac{y_1+y_2}{2}\right)^-\right) \neq 0$ *and* $f^{(j)}(y_2^-) = 0$ *for all* $j = 0, 1, \ldots, \ell-1$. *Then* $\gamma(D) \sim 1+Bernoulli(p_n(F))$, *where* $p_n(F) = P(\gamma(D) = 2)$, *and for bounded* $f^{(k)}(.)$ *and* $f^{(\ell)}(.)$ *we have* $\lim_{n\to\infty} p_n(F) =$

$$\frac{f^{(k)}(y_1^+)\, f^{(\ell)}(y_2^-)}{\left[f^{(k)}(y_1^+) + 2^{-(k+1)}f^{(k)}\left(\left(\frac{y_1+y_2}{2}\right)^+\right)\right]\left[f^{(\ell)}(y_2^-) + 2^{-(\ell+1)}f^{(\ell)}\left(\left(\frac{y_1+y_2}{2}\right)^-\right)\right]}.$$

If we choose F to be the uniform, i.e., $F = \mathcal{U}(y_1, y_2)$, then $k = \ell = 0$ and $f(y_1^+) = f(y_2^-) = f\left(\left(\frac{y_1+y_2}{2}\right)^+\right) = f\left(\left(\frac{y_1+y_2}{2}\right)^-\right) = 1/(y_1 - y_2)$. Then $\lim_{n\to\infty} p_n(F) = 4/9$, which agrees with Theorem 3.5.

3.2 Strong p-domination on CCCDs

In this section, we consider strong p-domination on CCCDs constructed with the set $\mathcal{Y}_2 = \{y_1, y_2\}$, and $\mathcal{X}_n$, a random sample from the uniform distribution between y_1 and y_2. We present the asymptotic distribution of $\gamma_{\geq p}(D)$, for such a CCCD, D, in the next section. Note that we choose $p \in [0, 1]$. For $p \in (0, 1/2]$, results are straightforward and we give them briefly. Unless stated otherwise, all p terms in the rest of the thesis are assumed to be from the interval $p \in (1/2, 1)$.

Uniformness enables $\gamma_{\geq p}$ to be independent from the support, i.e., y_1 and y_2.

Lemma 3.7 (Scale Invariance Property)**.** *Let* $-\infty < y_1 < y_2 < \infty$ *and* $-\infty < z_1 < z_2 < \infty$ *be two pairs of real numbers, and* $\mathcal{U}(a, b)$ *be the continuous uniform distribution on* (a, b)*. Then considering the linear bijection* $\omega : \mathbb{R} \to \mathbb{R}$, $\omega(x) = z_1.(y_2-x)/(y_2-y_1)+z_2.(x-y_1)/(y_2-y_1)$*, we have for the random variable* $X \sim \mathcal{U}(y_1, y_2)$*, the transformation* $Y = \omega(X) \sim \mathcal{U}(z_1, z_2)$*. Furthermore* $\omega(\cdot)$ *conserves the probability content on the nontrivial intervals.*

Proof. Clearly, ω takes the values from the interval $[y_1, y_2]$ to the interval $[z_1, z_2]$ and is a strictly increasing continuous bijection. Continuity and strict monotonity of ω implies its injectiveness. Hence we can apply Theorem (6.3.2) in [1] to determine the transformed random variable $Y = \omega(X)$. Let θ be the inverse transformation of ω and f_X be the pdf of X, then the pdf of Y, f_Y is, by the theorem, equal to $f_X(\theta).\left|\frac{\mathrm{d}}{\mathrm{dx}}\theta\right|$. $\theta(x) = y_1.\frac{z_2-x}{z_2-z_1} + y_2.\frac{x-z_1}{z_2-z_1}$, so we have its derivative $\frac{\mathrm{d}}{\mathrm{dx}}\theta(x) = \frac{y_2-y_1}{z_2-z_1}$ and as $f_X(x) = \frac{1}{y_1-y_2}$ is constant we obtain:

$$f_Y(x) = f_X(\theta(x)).\left|\frac{\mathrm{d}}{\mathrm{dx}}\theta(x)\right| = \frac{1}{y_1 - y_2}.\frac{y_2 - y_1}{z_2 - z_1} = \frac{1}{z_1 - z_2}$$

therefore we get $Y \sim \mathcal{U}(z_1, z_2)$.

Now, for any $c, d \in [y_1, y_2]$ and $X \sim \mathcal{U}(y_1, y_2)$ we have $P[X \in (c, d)] = \frac{d-c}{y_2-y_1}$ as well as $P(Y \in [\omega(c), \omega(d)]) = \frac{\omega(d)-\omega(c)}{z_2-z_1}$, for $Y \sim \mathcal{U}(z_1, z_2)$. Furthermore

$$P(Y \in [\omega(c), \omega(d)]) = \frac{\omega(d) - \omega(c)}{z_2 - z_1}$$

$$= \frac{z_1.\frac{y_2-d}{y_2-y_1} + z_2.\frac{d-y_1}{y_2-y_1} - z_1.\frac{y_2-c}{y_2-y_1} - z_2.\frac{c-y_1}{y_2-y_1}}{z_2 - z_1}$$

$$= \frac{z_1.y_2 - z_1.d + z_2.d - z_2.y_1 - z_1.y_2 + z_1.c - z_2.c + z_2.y_1}{(z_2 - z_1)(y_2 - y_1)}$$

$$= \frac{d(z_2 - z_1) - c(z_2 - z_1)}{(z_2 - z_1)(y_2 - y_1)} = \frac{d - c}{y_2 - y_1} = P(X \in [c, d])$$

□

Corollary 3.8. *For* $-\infty < y_1 < y_2 < \infty$*, let* $\mathcal{X}_n$ *be a random sample from the distribution* $\mathcal{U}(y_1, y_2)$*. Consider the CCCD, D, constructed with the sets* $\mathcal{X}_n$ *and* $\mathcal{Y}_2 = \{y_1, y_2\}$*, then the pdf of* $\gamma_{\geq p}(D)$ *is independent of* $\mathcal{Y}_2$*, i.e., it is equal to the pdf of* $\gamma_{\geq p}(D^*)$*, where* D^* *is the CCCD constructed with the sets* $\mathcal{Y}_2^* = \{y_1^*, y_2^*\}$*,* $y_1^*, y_2^* \in \mathbb{R}, y_1^* < y_2^*$*, and* $\mathcal{X}_n^*$ *a random sample from* $\mathcal{U}(y_1^*, y_2^*)$*.*

Proof. The result follows as the pdf of $\gamma_{\geq p}(D)$ is determined by the probability of uniform random variables being in the specified intervals. □

So we can choose, without loss of generality, $\mathcal{Y}_2 = \{0, 1\}$, and then generalize the results for any $\mathcal{Y}_m = \{y_1, y_2\}, y_1, y_2 \in \mathbb{R}, y_1 < y_2$. Hence $\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\}$ can be assumed to be a random sample from $\mathcal{U}(0, 1)$.

The domination number, $\gamma(D)$, of any nontrivial CCCD, D, with sets $\mathcal{X}_n$ and $\mathcal{Y}_2$ in $\mathbb{R}$, is either 1 or 2, and depends on the predicate whether the set $\mathcal{X}_n$ and the interval $\mathcal{I}_1 = (\frac{X_{n:n}}{2}, \frac{1+X_{1:n}}{2})$ are mutually exclusive or not [14]. Since the strong p-domination number $\gamma_{\geq p}(D)$ is the minimum cardinality of the vertex sets that cover at least $(p \cdot 100)\%$ of all n vertices, i.e., at least $\lceil n \cdot p \rceil$ of the vertices, we demonstrate that the distribution of $\gamma_{\geq p}(D)$ can be found in a similar fashion. We start with a simple observation.

Lemma 3.9. *Let D be the CCCD formed with sets $\mathcal{X}_n$ and $\mathcal{Y}_2$. Then we have $\gamma_{\geq p}(D) \leq 2$.*

Proof. Assuming that D is nontrivial, consider the cases when all elements of $\mathcal{X}_n$ are less than (resp. greater than) 1/2. Taking the greatest (resp. least) element, say x^* of $\mathcal{X}_n$ we directly obtain the coverage of all vertices with the ball $B(x^*, r(x^*))$, i.e., $\gamma_{[p]} = 1$. Otherwise, take the greatest element of $\mathcal{X}_n$, less than 1/2, say x_*, and the least element greater than 1/2, say x^*. Then we have the coverage of all vertices with the unions of the balls $B(x_*, r(x_*))$ and $B(x^*, r(x^*))$, i.e., $\gamma_{\geq p}(D) \leq 2$. (Note that this is an upper bound, i.e., desired domination may be achieved with just one of these vertices.) □

The following theorem is the main approach for finding the distribution of $\gamma_{\geq p}(D)$.

Theorem 3.10. *Let $k = n - \lfloor n \cdot p \rfloor$ and $\ell = \lceil n \cdot p \rceil$, consider the random variables $X_{k:n}$ and $X_{\ell:n}$ and define $\mathcal{I}_p = (X_{\ell:n}/2, (1+X_{k:n})/2)$. If $\mathcal{X}_n \cap \mathcal{I}_p \neq \emptyset$, then $\gamma_{\geq p}(D) = 1$, otherwise $\gamma_{\geq p}(D) = 2$.*

Proof. Observe that $X_{\ell:n} < 1$, so that $X_{\ell:n}/2 < 1/2$ as well as $1 + X_{k:n} > 1$, so that $X_{\ell:n}/2 < 1/2 < (1 + X_{k:n})/2$. Thus $1/2 \in \mathcal{I}_p$. Letting $\mathcal{I}_{p_1} = (X_{\ell:n}/2, 1/2]$ and $\mathcal{I}_{p_2} = (1/2, (1 + X_{k:n})/2)$, we have if $\mathcal{X}_n \cap \mathcal{I}_{p_1} \neq \emptyset$, then at least $X_{1:n}, X_{2:n}, \ldots, X_{\ell:n}$ is covered by the covering ball centered at any point in the nonempty set $\mathcal{X}_n \cap \mathcal{I}_{p_1}$. Similarly, if $\mathcal{X}_n \cap \mathcal{I}_{p_2} \neq \emptyset$, then at least $X_{k:n}, X_{k+1:n}, \ldots, X_{n:n}$ is covered by the covering ball centered at any point in the nonempty set $\mathcal{X}_n \cap \mathcal{I}_{p_2}$. As $\mathcal{X}_n \cap \mathcal{I}_p \neq \emptyset$ iff $\mathcal{X}_n \cap \mathcal{I}_{p_1} \neq \emptyset$ or $\mathcal{X}_n \cap \mathcal{I}_{p_2} \neq \emptyset$ and as at least either $n - k + 1 = n - (n - \lfloor n \cdot p \rfloor) + 1 = \lfloor n \cdot p \rfloor + 1 = \lceil n \cdot p \rceil$ or $\ell = \lceil n \cdot p \rceil$ vertices are dominated, we are done.

Now assume that $\mathcal{X}_n \cap \mathcal{I}_p = \emptyset$ and D is nontrivial. Then either $S_1 = [0, X_\ell/2] \cap \mathcal{X}_n \neq \emptyset$ or $S_2 = [\frac{1+X_k}{2}, 1] \cap \mathcal{X}_n \neq \emptyset$. Even if our aim is to dominate all the vertices, then we can do it with the set consisting of the greatest element of S_1 and the least element of S_2, as the former dominates all $X \in \mathcal{X}_n$ less than 1/2 and the latter dominates all $X \in \mathcal{X}_n$ greater than or equal to 1/2, so $\gamma_{\geq p}(D) = 2$ holds. □

We give two examples to this result. Observe that in Figure 9, B' is the interval associated with a possible element of random sample $\mathcal{X}_n$ that is exactly on $X_{\ell:n}/2$, and it covers all points up to $X_{\ell:n}$. Therefore any element of $\mathcal{X}_n$ that is in the highlighted interval will cover from $X_{1:n}$ up to at least $X_{\ell:n}$, thus give strong p-domination. Also in Figure 10, B'' is the interval associated with a possible element of random sample $\mathcal{X}_n$ that is exactly on $(1 + X_{k:n})/2$, and it covers all points down to $X_{k:n}$. Therefore any element of $\mathcal{X}_n$ that is in the highlighted interval will cover from $X_{n:n}$ at least down to $X_{k:n}$, thus give strong p-domination.

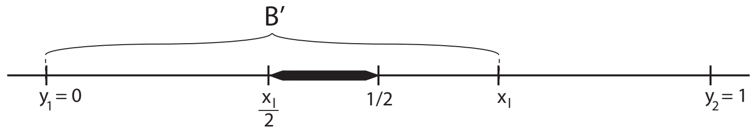

Figure 9: An example of the interval, $\mathcal{I}_p$, in Theorem 3.10.

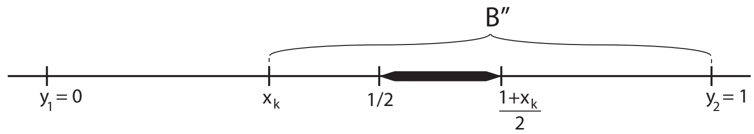

Figure 10: Another example of the interval, $\mathcal{I}_p$.

Theorem 3.10 directly implies the following corollary.

Corollary 3.11. *Let D be a CCCD constructed as above, then we have $\gamma_{\geq p}(D) \sim 1 + Bernoulli(\varsigma_n)$, where $\varsigma_n = P(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset)$.* $\square$

Remark 3.12. The closer there is a point $v \in \mathcal{X}_n \cap \mathcal{I}_{p_1}$ or $\mathcal{X}_n \cap \mathcal{I}_{p_2}$ to $1/2$, the more likely it is that the ball centered at v will cover all data points (hence at least $\lceil n \cdot p \rceil$ vertices), i.e., the more likely it is that $\gamma_{\geq p}(G) = \gamma(G) = 1$.

Now, instead of the probability of $\mathcal{X}_n \cap \mathcal{I}_p \neq \emptyset$, we determine the probability of $\mathcal{X}_n \cap \mathcal{I}_p = \emptyset$, for simplicity of calculations. First, we need the joint pdf of $X_{k:n}$ and $X_{\ell:n}$ for uniform data, $\mathcal{U}(0,1)$. For any pair (k,ℓ) with $1 \leq k < \ell \leq n$,

$$f_{k,\ell}(x_k, x_\ell) = \frac{n!}{(k-1)!(\ell-k-1)!(n-\ell)!} x_k^{k-1} \cdot (x_\ell - x_k)^{\ell-k-1} \cdot (1 - x_\ell)^{n-\ell}$$

for $0 < x_k < x_\ell < 1$. Conditioning on $X_{k:n} = x$ and $X_{\ell:n} = y$ we define

$$g(x,y) = P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset | X_{k:n} = x, X_{\ell:n} = y\right).$$

Observe that for $\mathcal{X}_n \cap \mathcal{I}_p$ to be empty, we must have

$$X_{k:n} < \frac{X_{\ell:n}}{2} \text{ and } \frac{1 + X_{k:n}}{2} < X_{\ell:n},$$

for otherwise, $X_{k:n}$ or $X_{\ell:n}$ would be contained in $\mathcal{I}_p$. Moreover, we need none of the $\ell - k - 1$ data points, between $X_{k:n}$ and $X_{\ell:n}$, to be in $\mathcal{I}_p$ and each such element has the probability

$$\frac{\frac{1+X_{k:n}}{2} - \frac{X_{\ell:n}}{2}}{X_{\ell:n} - X_{k:n}} = \frac{1 + X_{k:n} - X_{\ell:n}}{2(X_{\ell:n} - X_{k:n})}$$

to be contained in $\mathcal{I}_p$. Thus we have

$$g(x,y) = \left(1 - \frac{x - y + 1}{2(y - x)}\right)^{\ell-k-1}.$$

Considering the previous conditioning, we obtain the inequalities $X_{k:n} < \frac{X_{\ell:n}}{2} \Rightarrow X_{\ell:n} > 2X_{k:n}$ and $\frac{1+X_{k:n}}{2} < X_{\ell:n} \Rightarrow X_{\ell:n} > \frac{1+X_{k:n}}{2}$ and finally we obtain

$$P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset\right) = \int_0^{1/2} \int_{\max\left(\frac{x_k+1}{2}, 2x_k\right)}^1 f_{k,\ell}(x_k, x_\ell) g(x_k, x_\ell)\, dx_k\, dx_\ell,$$

where the bounds for x_ℓ are determined by the previous inequalities and the fact that $0 < x_\ell < 1$, and the bounds for x_k are determined by the fact that $0 < x_k < 1$ and if $x_k > 1/2$, then $1/2 < x_k < \frac{x_\ell}{2} < 1/2$ would lead to a contradiction. Following the simple facts, $x \in (0, 1/3) \Rightarrow \frac{1+x}{2} > 2x$ and $x \in (1/3, 1/2) \Rightarrow 2x > \frac{1+x}{2}$, we can divide this integral into two pieces and obtain

$$P(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset) = \left(\int_0^{1/3} \int_{\frac{x_k+1}{2}}^1 + \int_{1/3}^{1/2} \int_{2x_k}^1\right) f_{k,\ell}(x_k, x_\ell) g(x_k, x_\ell)\, dx_\ell\, dx_k$$

or, equivalently,

$$P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset\right) = \kappa \cdot \left(\int_0^{1/3} \int_{\frac{x_k+1}{2}}^{1} + \int_{1/3}^{1/2} \int_{2x_k}^{1} \right) x_k^{k-1} (x_\ell - x_k)^{\ell-k-1}$$
$$(1 - x_\ell)^{n-\ell} \left(1 - \frac{x_k - x_\ell + 1}{2(x_\ell - x_k)} \right)^{\ell-k-1} dx_\ell \, dx_k \quad (5)$$

where

$$\kappa = \frac{n!}{(k-1)!(\ell-k-1)!(n-\ell)!}.$$

For the CCCD with the target class $\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\}$ and $\mathcal{Y}_m = \{0, 1\}$, we can determine the distribution of $\gamma_{\geq p}(D)$, if we can calculate this probability. This is, again, due to the fact that (assuming $n \neq 0$)

$$\gamma_{\geq p}(D) = \begin{cases} 1, & \text{if } \mathcal{X}_n \cap \mathcal{I}_p \neq \emptyset, \\ 2, & \text{if } \mathcal{X}_n \cap \mathcal{I}_p = \emptyset. \end{cases}$$

Remark 3.13 (Coefficient κ of the Integral)**.** Consider the joint pdf, $f_{x_k,x_\ell}(x_k, x_\ell)$ of x_k and x_ℓ. There we have the constant

$$\kappa = \frac{n!}{(k-1)!(\ell-k-1)!(n-\ell)!}$$

and recall that $k = n - \lfloor n \cdot p \rfloor$, $\ell = \lceil n \cdot p \rceil$ and $p \in (1/2, 1)$. We would like to see if the factorial terms above are all well-defined.

First, $k - 1 = n - \lfloor n \cdot p \rfloor - 1$ may have a problematic case when we choose p too close to 1, but even then, as we take the floor integer, $\lfloor n \cdot p \rfloor$, we will have at the worst case $n - \lfloor n \cdot p \rfloor = 1$, and no less, so we will never have $k - 1 < 0$.

By the same reasoning, for $n - \ell = n - \lceil n \cdot p \rceil$ we may choose p too close to 1 so that $0 < n - n \cdot p < 1$, but even so $\lceil n \cdot p \rceil$ will be at most n, and the case $n - \ell < 0$ will not occur.

Now, considering $(\ell - k - 1) = \lceil n \cdot p \rceil + \lfloor n \cdot p \rfloor - n - 1$, if we choose p close enough to 1, we get $n + (n - 1) - n - 1 = n - 2$, which is greater than 0 as long as $n \geq 2$. This case is not problematic. On the other hand, if we take p too close to 1/2 (but still $p > 1/2$)

- for even $n \in \mathbb{N}$, $\ell - k - 1 = (\frac{n}{2} + 1) + \frac{n}{2} - n - 1 = 0$, which is clearly non-negative.
- for odd $n \in \mathbb{N}$, $\ell - k - 1 = \frac{n+1}{2} - (n - \frac{n-1}{2}) - 1 = -1$.

Therefore the only case that is problematic is when p is close enough to 1/2, i.e., when $\lceil n \cdot p \rceil = \frac{n+1}{2}$ and n is odd. Letting $I^* = \{p \in (1/2, 1) :$

$\lceil n \cdot p \rceil = \frac{n+1}{2}\}$, we have $p \in I^*$ iff $\frac{n}{2} < \lceil n \cdot p \rceil = \frac{n+1}{2}$ iff $\frac{n}{2} < n \cdot p \leq \frac{n+1}{2}$ iff $1/2 < p \leq \frac{n+1}{2n}$. We see that as n increases, measure of the interval I^* converges to 0. In other words, in the limit as $n \to \infty$ the problematic case occurs with probability 0. $\square$

Note that the integral in Equation (5) is not analytically tractable. So it is not possible to calculate $\varsigma_n = P(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset)$ and determine the distribution of $\gamma_{\geq p}(D)$ for finite n. Thus Corollary 3.11 is not very useful. From now on we tackle with the asymptotic distribution $\gamma_{\geq p}(D)$, i.e., distribution of the random variable $\gamma_{\geq p}(D)$, as $n \to \infty$.

3.3 Asymptotic Distribution of $\gamma_{\geq p}$ on CCCDs

Recall that we consider the CCCDs based on $\mathcal{Y}_2 = \{0,1\}$ and the random sample, $\mathcal{X}_n$, from uniform distribution on $(0,1)$. In this section we show that as $n \to \infty$, $\gamma_{\geq p}(D) \xrightarrow{P} 1$, i.e., asymptotic distribution of $\gamma_{\geq p}(D)$ is degenerate at 1, for any $p \in (1/2, 1)$. Thus in the limit it is possible to dominate at least $\lceil n \cdot p \rceil$ points of $\mathcal{X}_n$ with just one point, almost surely. We do this by two different methods.

(i) In the first method, we use the asymptotic behavior of the interval $\mathcal{I}_p = \left(\frac{X_{\ell:n}}{2}, \frac{1+X_{k:n}}{2}\right)$, as $n \to \infty$. In particular, we show that length (i.e., Lebesgue measure) of $\mathcal{I}_p$ converges to a positive quantity as $n \to \infty$. The interval $\mathcal{I}_p$ is determined by the random variables $\frac{X_{\ell:n}}{2}$ and $\frac{1+X_{k:n}}{2}$, and these converge to some degenerate distributions, with one is strictly greater than the other almost surely, so that it will be almost sure that there will be a point in $\mathcal{X}_n$ that is guaranteed to dominate at least $\lceil n \cdot p \rceil$ elements of $\mathcal{X}_n$. Hence $P\left(\gamma_{\geq p} = 1\right) \to 1$ as $n \to \infty$ will hold.

(ii) Consider the integral

$$P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset\right) = \int_0^{1/2} \int_{\max\left(\frac{x_k+1}{2}, 2x_k\right)}^{1} x_k^{k-1}(1-x_\ell)^{n-\ell} \left(\frac{3x_\ell - x_k - 1}{2}\right)^{\ell-k-1} dx_\ell \, dx_k. \quad (6)$$

Here we show that all the integrands have absolute value strictly smaller than 1 and have divergent powers dependent to n, and this will imply that the integral will converge to 0 as $n \to \infty$.

3.3.1 Method I

We start with a lemma that says in the limit, for any interval in $(0,1)$ there must exist a point of $\mathcal{X}_n$ in that interval.

Lemma 3.14. *For any $a, b \in \mathbb{R}$ with $0 < a < b < 1$ and $\mathcal{X}_n$ a random sample from $\mathcal{U}(0,1)$, as $n \to \infty$,*

$$P\left(\mathcal{X}_n \cap (a,b) \neq \emptyset\right) \to 1.$$

Proof. $P\left(\mathcal{X}_n \cap (a,b) = \emptyset\right) = \left(1 - \left(\frac{b-a}{1-0}\right)\right)^n = (a+1-b)^n$. Because $0 < 1 + a - b < 1$, as $n \to \infty$
$P\left(\mathcal{X}_n \cap (a,b) = \emptyset\right) \to 0$, i.e., $P\left((\mathcal{X}_n \cap (a,b) \neq \emptyset\right) \to 1$. ☐

We state a new proposition on the limiting behavior of sequences of random variables with specified properties.

Theorem 3.15. *Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of continuous random variables with $\mathbf{E}[X_n] = \mu_n \to \mu$, for some $\mu \in \mathbb{R}$, and $\mathbf{Var}[X_n] = \sigma_n^2 \to 0$ (with $\sigma_n^2 > 0$ for any $n \in \mathbb{N}$), as $n \to \infty$. Then we have the limiting random variable $X = \lim_{n\to\infty} X_n$ which has the degenerate distribution at μ.*

Proof. For each $n \in \mathbb{N}$, the event $\{X_n \neq \mu_n\}$ is equivalent to the union of events, $\cup_{i=1}^{\infty} \left\{|X_n - \mu_n| \geq \frac{1}{i}\right\}$. Also as $\sigma_n^2 > 0$ for any $n \in \mathbb{N}$ we have the inequality (1) and by the countable subadditivity of the probabilities of events we obtain

$$P(X_n \neq \mu_n) \leq \sum_{i=1}^{\infty} P\left(|X_n - \mu_n| \geq \frac{1}{i}\right) \leq \sum_{i=1}^{\infty} i^2\sigma_n^2. \tag{7}$$

For a fixed n, this implies the following. For any $k \in \mathbb{N}$ there exists a $n_k \in \mathbb{N}$ such that

$$\sum_{i=1}^{k} P\left(|X_n - \mu_n| \geq \frac{1}{i}\right) \leq \sum_{i=1}^{n_k} i^2\sigma_n^2. \tag{8}$$

Otherwise, there would be a natural number k s.t. for any natural number $n_k \in \mathbb{N}$

$$\sum_{i=1}^{k} P\left(|X_n - \mu_n| \geq \frac{1}{i}\right) > \sum_{i=1}^{n_k} i^2\sigma_n^2$$

and this would contradict with the inequality (7). Furthermore, we can always choose $\{n_k\}_{k\in\mathbb{N}}$ to be an increasing sequence, so that as $k \to \infty$, we have $n_k \to \infty$. Now let $X = \lim_{n\to\infty} X_n$. As (8) holds for any $n \in \mathbb{N}$, we can take the limit as $n \to \infty$ to obtain

$$\sum_{i=1}^{k} P\left(|X - \mu| \geq \frac{1}{i}\right) \leq \sum_{i=1}^{n_k} i^2 \lim_{n\to\infty} \sigma_n^2 = 0.$$

Therefore, for any $k \in \mathbb{N}$, we have $\sum_{i=1}^{k} P\left(|X - \mu| \geq \frac{1}{i}\right) = 0$, hence (as $k \to \infty$ implies $n_k \to \infty$)

$$P(X \neq \mu) \leq \sum_{i=1}^{\infty} P\left(|X - \mu| \geq \frac{1}{i}\right) = 0$$

Thence $X = \lim_{n\to\infty} X_n$ has the degenerate distribution at μ.

$\square$

Lemma 3.16. *Let $\mathcal{U}(0,1)$ be the uniform distribution on $(0,1)$, $X_{k:n}$ and $X_{\ell:n}$ the k^{th} and the ℓ^{th} order statistics of the random sample (with n elements) taken from $\mathcal{U}(0,1)$. Then $\mathbf{E}\left[\frac{X_{\ell:n}}{2}\right] = \frac{\ell}{2(n+1)}$ (Thus $\mathbf{E}\left[\frac{X_{\ell:n}}{2}\right] \to p/2$*

as $n \to \infty$*). Similarly* $\mathbf{E}\left[\frac{1+X_{k:n}}{2}\right] = \frac{n+k+1}{2(n+1)}$ *(Thus* $\mathbf{E}\left[\frac{1+X_{k:n}}{2}\right] \to 1 - p/2$ *as* $n \to \infty$*).*

Proof. The pdf of $X_{\ell:n}$ is

$$f_\ell(x_\ell) = \frac{n!}{(n-\ell)!(\ell-1)!}.x_\ell^{\ell-1}.(1-x_\ell)^{n-\ell}$$

so that the expectation

$$\mathbf{E}\left[\frac{X_{\ell:n}}{2}\right] = \frac{n!}{(n-\ell)!(\ell-1)!}\int_0^1 \frac{x_\ell}{2}.x_\ell^{\ell-1}.(1-x_\ell)^{n-\ell}\,dx_\ell$$

$$= \frac{n!}{2(n-\ell)!(\ell-1)!}.\frac{(n-\ell)!\ell!}{(n+1)!} = \frac{\ell}{2(n+1)} = \frac{\lceil n.p\rceil}{2(n+1)}$$

Now see that

$$\frac{n.p}{2(n+1)} < \mathbf{E}\left[\frac{X_{\ell:n}}{2}\right] < \frac{(n+1).p}{2(n+1)}$$

and both the upper and the lower bounds go to $\frac{p}{2}$ as $n \to \infty$. Therefore

$$\lim_{n\to\infty} \mathbf{E}\left[\frac{X_{\ell:n}}{2}\right] = \frac{p}{2}.$$

The result for $\mathbf{E}\left[\frac{1+X_{k:n}}{2}\right]$ follows similarly.

□

Lemma 3.17. $\mathbf{Var}\left[\frac{X_{\ell:n}}{2}\right] = \frac{\ell.(n-\ell+1)}{4(n+1)^2(n+2)}$ *(Thus* $\mathbf{Var}\left[\frac{X_{\ell:n}}{2}\right] \to 0$ *as* $n \to \infty$*). Moreover* $\mathbf{Var}\left[\frac{1+X_{k:n}}{2}\right] = \frac{k.(n-k+1)}{4(n+1^2)(n+2)}$ *(Thus* $\mathbf{Var}\left[\frac{1+X_{k:n}}{2}\right] \to 0$ *as* $n \to \infty$*).*

Proof.

$$\mathbf{Var}\left[\frac{X_{\ell:n}}{2}\right] = \mathbf{E}\left[\frac{X_{\ell:n}^2}{4}\right] - \mathbf{E}\left[\frac{X_{\ell:n}}{2}\right]^2 = \frac{n!}{4(n-\ell)!(\ell-1)!}.\frac{(n-\ell)!(\ell+1)!}{(n+2)!} - \frac{\ell^2}{4(n+1)^2}$$

$$= \frac{\ell}{4(n+1)}.\left(\frac{\ell+1}{n+2} - \frac{\ell}{n+1}\right) = \frac{\ell.(n-\ell+1)}{4(n+1)^2(n+2)}$$

Observe that

$$0 < \frac{\ell.(n-\ell+1)}{4(n+1)^2(n+2)} < \frac{(n+1)(n-\ell+1)}{4(n+1)^2(n+2)} < \frac{1}{4(n+2)}$$

and the upper bound goes to 0 as $n \to \infty$. Hence

$$\mathbf{Var}\left[\frac{X_{\ell:n}}{2}\right] \to 0, \text{ as } n \to \infty.$$

The result for $\mathbf{Var}\left[\frac{1+X_{k:n}}{2}\right]$ follows similarly.

$\square$

Now consider the two sequences of random variables, $\{A_n\}_{n\in\mathbb{N}}$ and $\{B_n\}_{n\in\mathbb{N}}$, where $A_n = \frac{X_{\ell:n}}{2}$ and $B_n = \frac{1+X_{k:n}}{2}$. By the properties of $\{A_n\}_{n\in\mathbb{N}}$ and $\{B_n\}_{n\in\mathbb{N}}$ stated in Lemma 3.16, Lemma 3.17, and applying Theorem 3.15, we obtain the following corollary.

Corollary 3.18. *The sequences of random variables $\{A_n\}_{n\in\mathbb{N}}$ and $\{B_n\}_{n\in\mathbb{N}}$ have degenerate limiting distributions, namely, $A = \lim_{n\to\infty} A_n$ has the degenerate distribution at $p/2$ and $B = \lim_{n\to\infty} B_n$ has the degenerate distribution at $1 - p/2$.* $\square$

As $p/2$ is strictly less than $1-p/2$ we have the interval $\mathcal{I}_p = (p/2, 1-p/2)$ with positive length. Furthermore, by Lemma 3.14 we have the probability, $P(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset) \to 0$, as $n \to \infty$. Thence by Corollary 3.18 the main result follows immediately.

Theorem 3.19 (Main Result I)**.** *Let $p \in (1/2, 1)$ and D be a CCCD based on $\mathcal{Y}_2 = \{0, 1\}$ and the random sample of size $n \in \mathbb{Z}^+$ $\mathcal{X}_n$, from the uniform distribution on $(0,1)$. Then as $n \to \infty$ we have $\gamma_{\geq p}(D) \xrightarrow{P} 1$, i.e., in the limit $\gamma_{\geq p}(D)$ has the degenerate distribution at 1.* $\square$

As a corollary, we state the distribution of the strong p-domination number for CCCDs with $\mathcal{Y}_m$ with more than two points.

Corollary 3.20. *Let $a, b \in \mathbb{R}$ with $a < b$, $n, m \in \mathbb{Z}^+$, D be a CCCD based on $\mathcal{Y}_m = \{y_1, y_2, \ldots, y_m\} \subset (a, b)$ and $\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\}$, where $X_i \overset{iid}{\sim} \mathcal{U}(a, b)$. Then we have $\gamma_{\geq p}(D) \xrightarrow{P} m+1$, i.e., in the limit, $\gamma_{\geq p}(D)$ has the degenerate distribution at $m+1$.*

Proof. For $i \in \{1, 2, \ldots, m-1\}$, let $\mathcal{J}_i = (y_i, y_{i+1}) \cap \mathcal{X}_n$, $\mathcal{J}_0 = (-\infty, y_1) \cap \mathcal{X}_n$ and $\mathcal{J}_m = (y_m, \infty) \cap \mathcal{X}_n$. See that there are $m+1$ disjoint sets, $\mathcal{J}_i, i \in \{0, 1, \ldots, m\}$. Observe that for a given random sample $\mathcal{X}_n$, the corresponding CCCD has at most $m+1$ components, and as $n \to \infty$, by Lemma 3.14, there is certainly some element of $\mathcal{X}_n$ in each $\mathcal{J}_i$, i.e., the CCCD will consist of exactly $m+1$ components.

Now consider the case $i = 0$, we can cover all vertices in $\mathcal{J}_0$ with the interval associated with the least element of $\mathcal{X}_n$, and similarly, we can cover all vertices in $\mathcal{J}_m$ with the interval associated with the largest element of $\mathcal{X}_n$. For $i \in \{1, 2, \ldots, m-1\}$ note that $|\mathcal{J}_i|$ will have infinitely many elements as well in the limit. This is implied by the fact that we can partition $\mathcal{J}_i$ into infinitely many disjoint intervals, and by Lemma 3.14, in each interval there will be at least one element of the random sample. Therefore each $m-1$ component is formed with $\{y_i, y_{i+1}\}$ and an associated random sample with number of elements going to infinity. So in the limit, they have the same distribution of strong p-domination number as the CCCD in Theorem 3.19, i.e., degenerate distribution at 1. Thus we can dominate each of these $m-1$ components with one vertex each, giving out the result that as $n \to \infty$, the CCCD, D, has strong p-domination number $m-1+1+1 = m+1$, i.e.,

$$\gamma_{\geq p}(D) \xrightarrow{P} m+1.$$

□

Finally, see that instead of $p \in (1/2, 1)$, taking $p \in (0, 1/2]$, we have the following corollary of the Main Result I.

Corollary 3.21. *Let D be a CCCD as in Theorem 3.19. For $p \in (0, 1/2]$, we have as $n \to \infty$, $\gamma_{\geq p}(D) \xrightarrow{P} 1$.*

Proof. Proof follows easily. Fix any $p \in (1/2, 1)$, by Theorem 3.19 $\gamma_{\geq p}(D)$ has the degenerate distribution at 1 in the limit. For any $p^* \in (0, 1/2]$, by Theorem 2.19, $\gamma_{\geq p^*}(D) \xrightarrow{P} 1$ too. □

3.3.2 Method II

We consider all the term in the integrands of the integral

$$I = \int_0^{1/2} \int_{\max\left(\frac{x_k+1}{2}, 2x_k\right)}^{1} x_k^{k-1}(1-x_\ell)^{n-\ell} \left(\frac{3x_\ell - x_k - 1}{2}\right)^{\ell-k-1} dx_\ell \, dx_k, \tag{9}$$

namely, x_k, $1-x_\ell$ and $\left(\frac{3x_\ell - x_k - 1}{2}\right)$, with their corresponding powers.

Remark 3.22. Observe that it naturally follows from Lemma 3.16 and Lemma 3.17 that in the limit the order statistics $X_{k:n}$ has the degenerate distribution at $1-p$ and $X_{\ell:n}$ has the degenerate distribution at p, i.e., $X_{k:n} \xrightarrow{P} 1-p$ and $X_{\ell:n} \xrightarrow{P} p$ This immediately leads to the following result.

Lemma 3.23. *Let $\varepsilon \in (0, 1-p)$, then $P(X_{k:n} > \varepsilon) \to 1$ and $P(X_{\ell:n} < 1-\varepsilon) \to 1$ as $n \to \infty$.* □

Observe that as the term x_k takes values in between $(0, 1/2)$, it is strictly less than 1 and it cannot get arbitrarily close to 1 to make the integral critical, i.e., not convergent to 0 as $n \to \infty$ (possibly divergent, or convergent to a positive value). Similarly, as $x_\ell \in \left(\max\{\frac{1+x_k}{2}, 2x_k\}, 1\right)$ it is at least 1/2, and thus the term $(1-x_\ell)$ cannot get arbitrarily close to 1. So we need only to consider the last term, $\left(\frac{3x_\ell - x_k - 1}{2}\right)$. For a specific $n \in \mathbb{N}$ this term can be arbitrarily close to 1 and make the integral in Equation (9) critical. But by Lemma 3.23 we know that for $n \in \mathbb{N}$ large enough and $\varepsilon \in (0, 1-p)$ we have the approximation

$$I \approx \int_{\varepsilon}^{1/2} \int_{\max\left(\frac{x_k+1}{2}, 2x_k\right)}^{1-\varepsilon} x_k^{k-1}(1-x_\ell)^{n-\ell} \left(\frac{3x_\ell - x_k - 1}{2}\right)^{\ell-k-1} dx_\ell \, dx_k.$$

In the limit $X_{\ell:n}$ can only get arbitrarily close to a real number $1-\varepsilon$, strictly less than one, and $X_{k:n}$ can only get arbitrarily close to the real number ε, strictly positive. Hence the third integrand $\left(\frac{3x_\ell - x_k - 1}{2}\right)$ can get arbitrarily close to at most $\frac{3(1-\varepsilon)-\varepsilon-1}{2} = 1 - 2\varepsilon < 1$.

Remark 3.24. Recall that $k = n - \lfloor n.p \rfloor$ and $\ell = \lceil n.p \rceil$. Hence

- $k - 1 \to \infty$ as $n \to \infty$
- $n - \ell \to \infty$ as $n \to \infty$
- $\ell - k - 1 \to \infty$ as $n \to \infty$ □

Now let $\varepsilon \in (0, 1-p)$, then we have the integral (9) dominated by the integral

$$\int_0^{1/2} \int_{\max\left(\frac{x_k+1}{2}, 2x_k\right)}^{1} (1-\varepsilon)^{p(n)} \, dx_k \, dx_\ell \tag{10}$$

where $p(n)$ is a function of n as $n \to \infty$, which can chosen to be $\min(k-1, n-\ell, \ell-k-1)$. The integral in Equation (10) converges to zero as $n \to \infty$, thus by Theorem 1.6 the integral in Equation (9) (the probability that $\mathcal{X}_n \cap \mathcal{I}_p = \emptyset$) converges to zero too. This implies

$$P\left(\mathcal{X}_n \cap \mathcal{I}_p \neq \emptyset\right) = P\left(\gamma_{\geq p}(D) = 1\right) \to 1, \text{ as } n \to \infty,$$

yielding the Main Result I, Theorem 3.19.

3.4 Monte Carlo Simulations

In this section we present Monte Carlo simulation results to verify our findings. We consider two different set of values, $p \in (1/2, 1)$, domination parameter, and $n \in \mathbb{N}$, number of vertices. We pick a random sample of size n, subject to uniform distribution between 0 and 1, and then observe the behavior of $\gamma_{\geq p}$, estimating the probability of the event $\gamma_{\geq p}(D) = 1$ for various combinations of n and p, out of 100000 trials. We use *R*, version 2.13.0 for simulations (*http://cran.r-project.org/*). The following is the R code snippet we used.

```
#This is the simulation for the rate that at least one element will be
#in Ip for p \in(1/2,1), considering a CCCD with ~U(0,1)
n= 100
p= 0.9
k= n - floor(n*p)
ell= ceiling(n*p)
count= 0
for (i in 1:100000){
data0 = runif(n)
data = sort(data0)
xk= data[k]
xl= data[ell]
z <- data[data > xl/2 & data < (1+xk)/2]
if (length(z) > 0) count = (count + 1)}
rate = count/i
rate
```

We present the simulation results.

$n\backslash p$	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	0.9999
100	1	1	1	1	1	1	1	0.99973	0.98878	0.55547
200	1	1	1	1	1	1	1	1	0.99974	0.55613
500	1	1	1	1	1	1	1	1	1	0.55455
1000	1	1	1	1	1	1	1	1	1	0.55443
2000	1	1	1	1	1	1	1	1	1	0.55569
5000	1	1	1	1	1	1	1	1	1	0.55499
10000	1	1	1	1	1	1	1	1	1	0.70435
20000	1	1	1	1	1	1	1	1	1	0.86881
50000	1	1	1	1	1	1	1	1	1	1
100000	1	1	1	1	1	1	1	1	1	1

Notice that even for small n convergence of $\gamma_{\geq p}$ to 1 is easily seen except for the case $p = 0.9999$. In this case the value of p is quite close to 1, and therefore for small n, $\gamma_{\geq p}(D)$ behaves like $\gamma(D)$. Thus we see that $P(\gamma_{\geq p}(D) = 1)$ is approximately equal to $P(\gamma(D) = 1) = 5/9$. Naturally after some point as we increase n, $P(\gamma_{\geq p}(D) = 1)$ converges to 1. The closer p is to 1, slower the convergence is observed.

4 PROPORTIONAL EDGE PROXIMITY CATCH DIGRAPHS, DOMINATION AND STRONG p-DOMINATION

In this section we introduce Proportional Edge-Proximity Catch Digraphs (PE-PCDs or PCDs) and study domination on them. We give several results previous from [11] on the distribution of domination number on PCDs. Then we present new results about strong p-domination and the strong p-domination number on the specified PCDs. PCDs were first introduced by Ceyhan and Priebe as an alternative to and a generalization of CC-CDs, because in higher dimensions finite and asymptotic distributions of the domination number on CCCDs are not analytically tractable. In [12] the asymptotic distribution of the domination number on PCDs in one dimension was presented, whereas in [11] finite and asymptotic distributions of the domination number on PCDs with both uniform and non-uniform data in one dimension are presented under various conditions and restrictions.

4.1 PCDs and Previous Results on Domination

Let $\mathcal{Y}_m = \{y_1, y_2, \ldots, y_m\} \subset \mathbb{R}$ (a subset of $\mathbb{R}$ of whose elements are ordered from smallest to largest), where $m \in \mathbb{N}^+$. For convenience of construction, let $y_0 = -\infty$ and $y_{m+1} = \infty$. Observe that $\mathcal{Y}_m$ divides the real line into $m+1$ pieces. Based on this observation, we define the intervals, $\mathcal{I}_i = (y_i, y_{i+1})$ for $i = 0, 1, \ldots, m$. Now, given $\mathcal{Y}_m$ and $n \in \mathbb{N}$ assume $\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\}$ is a random sample from a continuous distribution F in $\mathbb{R}$. For given $r \geq 1$ and $i \in \{1, 2, \ldots, n\}$, we define the set (called *proximity set*) that we will associate with any $x \in \mathcal{I}_i$, $N(x, r)$ as follows. For $i \in \{1, 2, \ldots, m-1\}$

$$N(x,r) = \begin{cases} (y_i, y_i + r\,(x - y_i)) \cap \mathcal{I}_i & \text{if } x \in (y_i, \frac{y_i + y_{i+1}}{2}), \\ (y_{i+1} - r\,(y_{i+1} - x)\,, y_{i+1}) \cap \mathcal{I}_i & \text{if } x \in (\frac{y_i + y_{i+1}}{2}, y_{i+1}), \end{cases}$$

and also for $i \in \{0, m+1\}$

$$N(x,r) = \begin{cases} (y_1 - r\,(y_1 - x)\,, y_1) \cap \mathcal{I}_i & \text{if } x < y_1, \\ (y_m, y_m + r(x - y_m)) \cap \mathcal{I}_i & \text{if } x > y_m. \end{cases}$$

Now that we have the proximity sets to associate with each of the data points of $\mathcal{X}_n$, we give the definition.

Definition 20. *The proportional-edge proximity catch digraph* (PCD) corresponding to the 'target' set $\mathcal{X}_n$, the set $\mathcal{Y}_m$ and the extension parameter $r \geq 1$ is the digraph with n vertices and for each $i, j \in \{1, 2, \ldots, n\}$ there is an arc from the i^{th} vertex to the j^{th} vertex if and only if $X_j \in N(X_i, r)$.

An example of proximity sets associated with the sets $\mathcal{X}_3 = \{x_1, x_2, x_3\}$ and $\mathcal{Y}_2 = \{0, 1\}$ and the resulting PCD for the parameter $r = 3/2$ is given in Figure **??** and Figure 12.

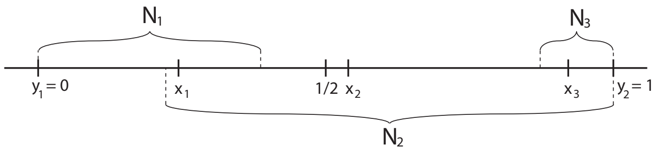

Figure 11: An example of proximity sets in $\mathcal{R}$.

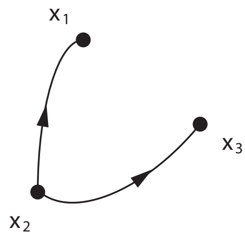

Figure 12: The PCD constructed from the sets in Figure 11.

Note that we consider the simpler case of $\mathcal{Y}_2 = \{y_1, y_2\}$, with $-\infty < y_1 < y_2 < \infty$. Moreover we will choose F to be the uniform distribution between y_1 and y_2, i.e., $\mathcal{X}_n$ is a random sample from $\mathcal{U}(y_1, y_2)$. By Lemma 3.7, we work on the case $(y_1, y_2) = (0, 1)$ and then generalize our results accordingly. From here on, the sets associated with the points in the random sample $\mathcal{X}_n$ are

$$N(x, r) = \begin{cases} (0, rx) \cap (0, 1) & \text{if } x \in (0, 1/2), \\ (1 - r(1 - x), 1) \cap (0, 1) & \text{if } x \in [1/2, 1). \end{cases} \tag{11}$$

Observe that the inclusion the number $1/2$ to the interval of the second case is discretionary, since for any $i \in \{1, 2, \dots, n\}$, we have $P(X_i = 1/2) = 0$. Also note that in [11] there is another parameter, called 'the centrality parameter' (and denoted by c), and it generalizes the choice of $1/2$ for the construction of $N(x, r)$. Here, we are only concerned with the case $c = 1/2$. Also see that the case $r = 2$ yields a CCCD, so any CCCD is a PCD with the condition that $r = 2$.

Now let D be a PCD constructed with the set $\mathcal{Y}_2 = \{0, 1\}$ and the random sample $\mathcal{X}_n$, from $\mathcal{U}(0, 1)$, for $r \geq 1$. The exact and asymptotic distribution of domination number of D is given in [11].

Theorem 4.1. *For D constructed as above, $\gamma(D) \sim 1 + Bernoulli(p_n)$, where*

$$p_n = \begin{cases} 1 - \frac{1+r^{2n-1}}{(2r)^{n-1}(r+1)} + \frac{(r-1)^n}{(r+1)^n}\left(1 - \left(\frac{r-1}{2r}\right)^{n-1}\right), & \text{for } 1 \le r < 2, \\ \frac{2r}{(r+1)^2}\left(\left(\frac{2}{r}\right)^{n-1} - \left(\frac{r-1}{r^2}\right)^{n-1}\right), & \text{for } r \ge 2. \end{cases}$$

Observe that for $r = 2$, $p_n = 4/9 - (16/9)4^{-n}$, i.e., identical to CCCDs.

Corollary 4.2. *In the limit, as $n \to \infty$, we have*

$$\gamma(D) \sim \begin{cases} 1, & \text{for } r > 2, \\ 1 + Bernoulli(4/9), & \text{for } r = 2, \\ 2, & \text{for } 1 \le r < 2. \end{cases}$$

4.2 Exact and Weak p-Domination on PCDs

In this section we study exact and weak p-domination on PCDs, and give the asymptotic distribution of weak p-domination number for PCDs constructed with the set $\mathcal{Y}_2 = \{0, 1\}$ and the random sample $\mathcal{X}_n$, from $\mathcal{U}(0, 1)$, for $r \ge 1$. As our concern is on asymptotic probability, we define and use the following concept.

Definition 21. For a sequence of random variables, $(X_n)_{n \in \mathbb{N}}$, a function f is called *asymptotically accurate probability density function*, if

$$\lim_{n \to \infty} \int_{-\infty}^{1} f(t)\, dt = 1.$$

Note that an asymptotically accurate probability density function may have n as argument, and this makes sense, since it is used to deal asymptotic situations.

Recalling the restrictions on the existence of exact p-dominating sets and exact p-domination number, we make the following conjecture and then focus on weak p-domination.

Conjecture 4.3. *Let $r \ge 1$, and D be a PCD constructed with the set $\mathcal{Y}_2 = \{0, 1\}$ and the random sample of size n, $\mathcal{X}_n$, from $\mathcal{U}(0, 1)$. Then $\gamma_p^e(D)$ has non-degenerate asymptotic distribution.*

We consider four random variables while dealing with weak p-domination. U_c is the largest element of $\mathcal{X}_n$ smaller than $1/2$, whereas U_d is the smallest element of $\mathcal{X}_n$ larger than $1/2$. Also we define Y_p and Y_{1-p}, closest points to p and $1 - p$. We provide the asymptotically accurate joint pdf of U_c and Y_p,

$$f_{c,p}(u_c, y_p) = n(n-1)\left(1 - ((1/2 - u_c) + |p - y_p|)\right)^{n-2}$$

$$= \begin{cases} n(n-1)\left(1/2 + u_c + y_p - p\right)^{n-2}, & \text{if } 1/2 < y_p < p \text{ and } 0 < x < 1/2, \\ n(n-1)\left(1/2 + u_c - y_p + p\right)^{n-2}, & \text{if } p < y_p < 1 \text{ and } 0 < x < 1/2. \end{cases}$$

We define the following events,

$$E_1 : \{rU_c \geq Y_p\}, \tag{12}$$

and

$$E_2 : \{1 - r(1 - U_d) \leq Y_{1-p}\}. \tag{13}$$

Observe that $\{\gamma_p^w = 1\}$ is asymptotically equivalent to $\{E_1 or E_2\}$, i.e.,

$$P(\gamma_p^e(D) = 1) = P(E_1) + P(E_2) - P(E_1 \cap E_2).$$

First, we show that there is symmetry between the events E_1 and E_2 in the following sense.

Lemma 4.4. *Considering the events, E_1 in Equation 12 and E_2 in Equation 13, we have $P(E_1) = P(E_2)$, as $n \to \infty$.*

Proof. The proof goes pretty much alike the proof of Proposition 4.21. We use the fact that as $n \to \infty$, U_c and U_d converge to 1/2, and this implies that for sufficiently large $n \in \mathbb{N}$, $U_c \stackrel{d}{=} 1 - U_d$, where $\stackrel{d}{=}$ stands for equal in distribution. Similarly, for sufficiently large $n \in \mathbb{N}$, $Y_p \stackrel{d}{=} 1 - Y_{1-p}$. Then it follows that

$$\{rU_c \geq Y_p\} \Leftrightarrow \{r(1 - U_d) \geq 1 - Y_{1-p}\} \Leftrightarrow \{Y_{1-p} \geq 1 - r(1 - U_d)\},$$

and the probabilities of these events are the same, i.e., $P(E_1) = P(E_2)$ in the limit. $\square$

We first calculate $\lim_{n\to\infty} P(E_1)$. See that $rU_c \geq Y_p$ implies $U_c \geq Y_p/r$, and that implies $Y_p < p$ and hence we obtain the integral

$$P(E_1) = \int_{1/2}^{p} \int_{y_p/r}^{1/2} n(n-1)\left(1/2 + u_c + y_p - p\right)^{n-2} \, du_c \, dy_p,$$

and we calculate it to be

$$P(E_1) = 1 - (3/2 - p)^n - \left(\frac{r}{r+1}\right)\left((p/r + 1/2)^n - (1 + 1/2r - p)^n\right).$$

Writing the Taylor expansion of $P(E_1)$ up to first degree, and substituting $r = 2p$, we get

$$P(E_1) = \frac{1}{1+2p} + O(n^{-1}),$$

and as we are concerned only with asymptotic case, sending $n \to \infty$ we get

$$\lim_{n\to\infty} P(E_1) = \frac{1}{1+2p}.$$

By Lemma 4.4, we have $\lim_{n\to\infty} P(E_2) = \frac{1}{1+2p}$ too. Now we calculate $P(E_1 \cap E_2)$. First, we provide the asymptotically accurate joint pdf U_c, U_d, Y_p and Y_{1-p}, it is

$$f_4(u_c, u_d, y_p, y_{1-p}) = n(n-1)(n-2)(n-3) \\ (1 - (|y_{1-p} - 1 + p| + (u_d - u_c) + |p - y_p|))^{n-4}.$$

Observe that $ru_c > y_p$ implies $y_p < p$ and $1 - r(1-u_d) < y_{1-p}$ implies $y_{1-p} > 1 - p$. Using Lemma 4.18, given $\varepsilon > 0$, close enough to zero, we obtain the integral

$$P(E_1 \cap E_2) \approx \int_{1-p}^{1-p+\varepsilon} \int_{p-\varepsilon}^{p} \int_{1/2}^{1-(1-y_{1-p})/r} \int_{y_p/r}^{1/2} n(n-1)(n-2)(n-3) \\ (2(1-p) + u_c + y_p - (u_d + y_{1-p}))^{n-4} \, du_c \, du_d \, dy_p \, dy_{1-p}.$$

Letting the change of variables, $u_1 = u_c$, $1 - u_2 = u_d$, $v_1 = y_p$ and $1 - v_2 = y_{1-p}$ we get the form, where we substitute $r = 2p$,

$$P(E_1 \cap E_2) \approx \int_{p}^{p-\varepsilon} \int_{p-\varepsilon}^{p} \int_{1/2}^{v_2/2p} \int_{v_1/2p}^{1/2} n(n-1)(n-2)(n-3) \\ (u_1 + v_1 + u_2 + v_2 - 2p)^{n-4} \, du_1 \, du_2 \, dv_1 \, dv_2.$$

Integrating and then calculating the Taylor expansion, we get in the limit, as $n \to \infty$, that

$$\lim_{n\to\infty} P(E_1 \cap E_2) = \frac{1}{(1+2p)^2} + O(\varepsilon),$$

and sending $\varepsilon \to 0$ we get

$$\lim_{n\to\infty} P(E_1 \cap E_2) = \frac{1}{(1+2p)^2}.$$

Now we have the probability, $P(\gamma_p^w = 1)$.

Theorem 4.5 (Main Result II)**.** *For a PCD, D, constructed with the set $\mathcal{Y}_2 = \{0,1\}$ and the random sample of size $n \in \mathbb{Z}^+$, $\mathcal{X}_n$, from $\mathcal{U}(0,1)$, with $r = 2p \geq 1$, we have as $n \to \infty$,*

$$P(\gamma_p^w = 1) = \frac{2}{1+2p} - \frac{1}{(1+2p)^2} = \frac{4p+1}{(2p+1)^2}. \quad \square$$

Remark 4.6. Observe that letting $p = 1$, we have $r = 2$ and $P(\gamma_p^w(D) = 1) = 5/9$, which agrees with the result in [14].

Remark 4.7. For $p = r/2 \in (1/2, 1)$, this probability also serves as a lower bound for the probability $P(\gamma_{\geq p}(D) = 1)$ in the limit, as $n \to \infty$.

4.3 Strong p-domination on PCDs

In this section we consider the concept of strong p-domination on PCDs constructed with $\mathcal{Y}_2 = \{0,1\}$ and $\mathcal{X}_n$, a random sample from $\mathcal{U}(0,1)$, for $r \geq 1$. The asymptotic distribution of $\gamma_{\geq p}$ is given in the next section. Again, we start with an easy lemma.

Lemma 4.8. *Let F be any probability distribution on $(0,1)$, D be the PCD formed with $\mathcal{Y}_2 = \{0,1\}$, $\mathcal{X}_n$, a random sample of size n subject to F and $r \geq 1$. Then we have $\gamma_{\geq p}(D) \leq 2$, for any $p \in [0,1]$.*

Proof. Identical with the proof of Lemma 3.9. $\square$

Now let $k = n - \lfloor n.p \rfloor$ and $\ell = \lceil n.p \rceil$ and define $\mathcal{I}_p = \left(\frac{X_{\ell:n}}{r}, 1/2\right) \cup \left[1/2, 1 + \frac{X_{k:n}-1}{r}\right)$, with the understanding that the interval (a,b) is empty whenever $a > b$. Note that $\mathcal{I}_p$ depends on the order statistics, $X_{k:n}$ and $X_{\ell:n}$. It has the following forms (written as cases(i)-(iv)),

- case(i): $\left(\frac{X_{\ell:n}}{r}, 1 + \frac{X_{k:n}-1}{r}\right)$,
- case(ii): $\left(\frac{X_{\ell:n}}{r}, \frac{1}{2}\right)$,
- case(iii): $\left[\frac{1}{2}, 1 + \frac{X_{k:n}-1}{r}\right)$,
- case(iv): $\emptyset$.

Observe that these four cases can be based on four conditions (which are 2-by-2 exhaustive) in the fashion that they determine the set $\mathcal{I}_p$, namely

(C1): $X_{\ell:n} \leq r/2$,

(C2): $X_{\ell:n} > r/2$,

(C3): $X_{k:n} \leq 1 - r/2$,

(C4): $X_{k:n} > 1 - r/2$.

The case(i) is implied by (C1) and (C4), case(ii) is implied by (C1) and (C3), case(iii) is implied by (C2) and (C4) and lastly, case(iv) is implied by (C2) and (C3). Thus we have the following result.

Lemma 4.9. *The set $\mathcal{I}_p$ is determined by the conditions (C1), (C2), (C3) and (C4). In particular,*

$$\mathcal{I}_p = \begin{cases} \left(\frac{X_{\ell:n}}{r}, 1 + \frac{X_{k:n}-1}{r}\right), & \textit{if case(i) holds,} \\ \left(\frac{X_{\ell:n}}{r}, \frac{1}{2}\right), & \textit{if case(ii) holds,} \\ \left[\frac{1}{2}, 1 + \frac{X_{k:n}-1}{r}\right), & \textit{if case(iii) holds,} \\ \emptyset, & \textit{if case(iv) holds.} \end{cases}$$

□

The following result forms the basis of the determination of the distribution of $\gamma_{\geq p}(D)$.

Theorem 4.10. *For the PCD, D, formed with $\mathcal{Y}_2 = \{0, 1\}$, F any probability distribution on $(0, 1)$ and $\mathcal{X}_n$ subject to F and $r \geq 1$, if $\mathcal{X}_n \cap \mathcal{I}_p \neq \emptyset$, then $\gamma_{\geq p}(D) = 1$, otherwise $\gamma_{\geq p}(D) = \gamma(D) = 2$.*

Proof. Consider any point $x_* \in \mathcal{X}_n$ with $x_* < 1/2$. The set $N(x_*, r)$ covers all elements of $\mathcal{X}_n$ that are smaller than x_* and if $x_* > x_{\ell:n}/r$ it will cover at least $\lceil n.p \rceil$ elements of $\mathcal{X}_n$, namely $X_{1:n}, X_{2:n}, \ldots, X_{\ell:n}$. This means that if there is an element $x_* \in \mathcal{X}_n$ with $x_* < 1/2$ and $x_* > X_{\ell:n}/r$, then x_* dominates at least $\lceil n.p \rceil$ points. Similarly for $x^* > 1/2$, $N(x^*, r)$ covers all elements of $\mathcal{X}_n$ that are greater than x^* and if $x^* < 1 + \frac{X_{k:n}-1}{r}$, it will cover at least $\lceil n.p \rceil$ elements of $\mathcal{X}_n$, namely $X_{k:n}, X_{k+1:n}, \ldots, X_{n:n}$. Again, that means if there is an element $x^* \in \mathcal{X}_n$ such that $x^* > 1/2$ and $x^* < 1 + \frac{X_{k:n}-1}{r}$, then x^* dominates at least $\lceil n.p \rceil$ points. Therefore $\gamma_{\geq p}(D) = 1$ if and only if there is an element of $\mathcal{X}_n$ either in $\left(\frac{X_{\ell:n}}{r}, \frac{1}{2}\right)$ or in $\left[\frac{1}{2}, 1 + \frac{X_{k:n}-1}{r}\right)$, i.e., $\mathcal{X}_n \cap \mathcal{I}_p \neq \emptyset$. Otherwise $\gamma_{\geq p}(D) = 2$ by Lemma 4.8. □

So the problem of finding the distribution of $\gamma_{\geq p}(D)$ boils down to the calculation of the probability $P(\mathcal{X}_n \cap \mathcal{I}_p \neq \emptyset)$ and thus we need to make calculations with respect to the conditions of the set $\mathcal{I}_p$ being empty, a single interval or in the limit, becoming a set consisting of one point.

4.4 Asymptotic Distribution of $\gamma_{\geq p}$ for PCDs

Recall that we proved $X_{k:n} \xrightarrow{P} 1-p$ and $X_{\ell:n} \xrightarrow{P} p$ as $n \to \infty$ in Corollary 3.18. From this result we see that as $n \to \infty$ the set $\mathcal{I}_p$ converges to the set $(p/r, 1/2) \cup [1/2, 1-p/r)$. Next result follows easily from this fact.

Lemma 4.11. *For a PCD constructed with $\mathcal{X}_n$, a random sample of size n from $\mathcal{U}(0,1)$, and $\mathcal{Y}_2 = \{0,1\}$, with the expansion parameter $r \in (1,2)$, we have $\gamma_{\geq p}(D)$ with degenerate distribution in the limit, as $n \to \infty$, whenever $r \neq 2p$. In particular, as $n \to \infty$,*

$$\gamma_{\geq p}(D) = \begin{cases} 1, & \text{if } r > 2p, \\ 2, & \text{if } r < 2p. \end{cases}$$

Proof. Notice that for $r > 2p$, $p/r < 1/2$ holds, i.e., in the limit the set $\mathcal{I}_p$ is a set with positive Lebesgue measure (i.e., positive length). Then by Lemma 3.14, there exists an element of the random sample $\mathcal{X}_n$, that is also in $\mathcal{I}_p$ with probability 1, i.e., $P(\gamma_{\geq p}(D) = 1)$ attains the value 1 as $n \to \infty$. Similarly, for $r > 2p$, $p/r > 1/2$ holds, i.e., in the limit the set $\mathcal{I}_p = \emptyset$, so fractional domination with just one point is not possible. Hence $\lim_{n\to\infty} P(\gamma_{\geq p}(D) = 2) = 1$. □

Next, we consider the case $r = 2p$. Proof of the following theorem is provided later.

Theorem 4.12. *(I) $P_{1,n} := P(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset, case(i)) \to 0$ as $n \to \infty$,*

(II) $P_{2,n} := P(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset, case(ii)) \to 0$ as $n \to \infty$,

(III) $P_{3,n} := P(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset, case(iii)) \to 0$ as $n \to \infty$.

By this theorem, we arrive to the fact that in the limit, the only case that $(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset)$ occurs is whenever case(iv) holds. Following corollary is trivial after Theorem 4.12.

Corollary 4.13. *As $n \to \infty$, $P\left(\gamma_{\geq p}(D) = 2\right) = P(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset) = P(case(iv))$.*

Proof. As the cases (i)-(ii)-(iii)-(iv) are disjoint and exhaustive, we have

$$P(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset) = P((\mathcal{X}_n \cap \mathcal{I}_p = \emptyset), case(i)) + P((\mathcal{X}_n \cap \mathcal{I}_p = \emptyset), case(ii)) + \\ P((\mathcal{X}_n \cap \mathcal{I}_p = \emptyset), case(iii)) + P((\mathcal{X}_n \cap \mathcal{I}_p = \emptyset), case(iv)).$$

Therefore, by Theorem 4.12 we have

$$\lim_{n\to\infty} P\left(\gamma_{\geq p}(D) = 2\right) = \lim_{n\to\infty} P(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset) = \lim_{n\to\infty} P(case(iv)).$$

□

Note that the following theorem, whose proof is presented later and is crucial in determining the distribution of $\gamma_{\geq p}(D)$.

Theorem 4.14. *Let $\theta_{n,p} = P\,(case(iv))$, then we have*

$$\lim_{n\to\infty} \theta_{n,p} = \frac{1}{4} - \frac{1}{2\pi} \arctan\left(\frac{1-p}{\sqrt{2p-1}}\right).$$

Let $\pi_p = \lim_{n\to\infty} \theta_{n,p} = \frac{1}{4} - \frac{1}{2\pi} \arctan\left(\frac{1-p}{\sqrt{2p-1}}\right)$. Assuming that Theorem 4.12 and Theorem 4.14 are proven, we have the following main result.

Theorem 4.15 (Main Result III)**.** *For a PCD, D, constructed with $\mathcal{X}_n$, a random sample of size n from $\mathcal{U}(0,1)$, and $\mathcal{Y}_2 = \{0,1\}$, with the expansion parameter $r \in (1,2)$ and $p \in (1/2,1)$ we have $\gamma_{\geq p}(D)$ with the following distribution in the limit, as $n \to \infty$,*

$$\gamma_{\geq p}(D) \sim \begin{cases} 1, & \text{if } r > 2p, \\ 2, & \text{if } r < 2p, \\ 1 + Bernoulli(\pi_p), & \text{if } r = 2p, \end{cases}$$

where $\pi_p = \frac{1}{4} - \frac{1}{2\pi} \arctan\left(\frac{1-p}{\sqrt{2p-1}}\right)$.

Proof. By Lemma 4.11, when $r \neq 2p$, we have degeneracy cases in the limit. Otherwise as $n \to \infty$, from Corollary 4.13, it follows that $\gamma_{\geq p}(D) = 2$ if and only if case(iv) holds. Thus follows the main result. $\square$

Letting a, b be real numbers with $a < b$, $n, m \in \mathbb{Z}^+$ we now consider the PCD, D, constructed with the set $\mathcal{Y}_m = \{y_1, y_2, \ldots, y_m\} \subset (a,b)$ and a random sample of size n, $\mathcal{X}_n$ from the $\mathcal{U}(a,b)$ distribution. Notice that $\mathcal{Y}_m$ divide the interval (a,b) into $m+1$ disjoint intervals, we denote them as follows. For $i \in \{2,3,\ldots,m\}$, we let $I_i = (y_{i-1}, y_i)$, then define $I_1 = (a, y_1)$ and $I_{m+1} = (y_m, b)$. Also define for any $i \in \{1,2,\ldots,m+1\}$, $J_i = \mathcal{X}_n \cap I_i$. Observe that J_i are disjoint and they yield disjoint components of D. So we need to consider domination in each component separately. We aim to find the distribution of $\gamma_{\geq p}(D)$ for such a PCD, D.

Now let $p_i \in \mathbb{R}$, where $\sum_{i=1}^{m+1} p_i \geq p$, with $\gamma_{\geq p}(D) = \sum_{i=1}^{m+1} \gamma_{\geq p_i}(D_i)$, where D_i is the sub digraph constructed with I_i and J_i. By Theorem 4.15 we know the distribution of each $\gamma_{\geq p_i}(D_i)$. Note that $\gamma_{\geq p_1}(D_1)$ and $\gamma_{\geq p_m}(D_m)$ are equal to one in the limit, as $n \to \infty$, as we can cover all vertices in J_0 with the smallest element of it as well as all vertices in J_m with the largest element of it. This implies

$$\gamma_{\geq p}(D) = 2 + \sum_{i=2}^{m} \gamma_{\geq p_i}(D_i),$$

and the result follows.

Theorem 4.16. *For any $a, b \in \mathbb{R}$ with $a < b$, $n, m \in \mathbb{Z}^+$, $\mathcal{Y}_m = \{y_1, y_2, \ldots, y_m\} \subset (a, b)$ and a random sample of size n, $\mathcal{X}_n$ from the $\mathcal{U}(a, b)$ distribution, consider the resulting PCD, D. For given $\{p_1, p_2, \ldots, p_{m+1}\}$ with $\sum_{i=1}^{m+1} p_i \geq p$, D has the strong p-domination number in the following form*

$$\gamma_{\geq p}(D) = 2 + \sum_{i=2}^{m} \left(I(r > 2p_i) + 2I(r < 2p_i) + (1 + \Gamma_i) I(r = 2p_i) \right),$$

where $\Gamma_i \sim Bernoulli(\pi_{p_i})$, with π_{p_i} as in Theorem 4.15. □

Additionally, note that whenever there is the case $\sum_{i=1}^{m+1} p_i = p$, for any $i \in \{2, 3, \ldots, m\}$ we have $r > 2p_i$ and that implies

$$\gamma_{\geq p}(D) = 2 + \sum_{i=2}^{m} 1 = m + 1.$$

Also, for the sake of completeness we present the following result.

Corollary 4.17. *Given $p \in (0, 1/2]$, the PCD, D, as in Theorem 4.15, has the strong p-domination number with the degenerate distribution at 1, as $n \to \infty$.* □

Proof. Proof easily follows from the fact that for $p \in (0, 1/2)$ the interval $\mathcal{I}_p$ has positive length. Therefore as $n \to \infty$ strong p-domination with just one vertex becomes almost sure. □

In the following two subsections we present the proofs of the two theorems, 4.12 and 4.14, which yield our main result. We need the following tools for these proofs. Next proposition follows directly from the fact that $X_{k:n}$ converges in probability to $1 - p$ and $X_{\ell:n}$ converges in probability to p, as $n \to \infty$.

Proposition 4.18. *For any $\varepsilon > 0$, as $n \to \infty$, we have*

$$P\left(X_{k:n} \in (1 - p - \varepsilon, 1 - p + \varepsilon), X_{\ell:n} \in (p - \varepsilon, p + \varepsilon)\right) \to 1. \quad \square$$

Proposition 4.19. *For $p \in (1/2, 1)$, let $(a_n)_{n \in \mathbb{N}}$, $(b_n)_{n \in \mathbb{N}}$ and $(c_n)_{n \in \mathbb{N}}$ be three sequences of real numbers defined as $a_n = n/(n - \ell)$, $b_n = n/(k - 1)$ and $c_n = n/(\ell - k - 1)$. Then we have*

$$a_n, b_n \to \frac{1}{1-p}, \text{ and}$$
$$c_n \to \frac{1}{2p-1}.$$

Proof. Recall that $k = n - \lfloor n.p \rfloor$ and $\ell = \lceil n.p \rceil$. The following inequality is immediate.

$$n - np - 1 < n - \ell < n - np$$

Hence follows

$$\underbrace{\frac{1}{1-p}}_{\alpha_n} = \frac{n}{n-np} < a_n = \frac{n}{n-\ell} < \frac{n}{n-np-1} = \underbrace{\frac{1}{1-p-1/n}}_{\beta_n}.$$

As $n \to \infty$, both α_n and β_n goes to $1/(1-p)$ and so does a_n. Similarly,

$$n - np - 1 < k - 1 < n - np$$

and

$$\underbrace{\frac{1}{1-p}}_{\alpha_n} = \frac{n}{n-np} < b_n = \frac{n}{k-1} < \frac{n}{n-np-1} = \underbrace{\frac{1}{1-p-1/n}}_{\beta_n}.$$

Again, as $n \to \infty$, both α_n and β_n goes to $1/(1-p)$ and so does b_n. Now see that

$$2np - n - 2 < \ell - k - 1 < 2np - n$$

and this implies that

$$\underbrace{\frac{1}{2p-1}}_{\alpha'_n} = \frac{n}{2np-n} < c_n = \frac{n}{\ell-k-1} < \frac{n}{2np-n-2} = \underbrace{\frac{1}{2p-1-2/n}}_{\beta'_n}.$$

As $n \to \infty$, both α'_n and β'_n goes to $1/(2p-1)$ and so does c_n. $\square$

Remark 4.20 (Stirling's Approximation)**.** For sufficiently large $n \in \mathbb{N}$ we have

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

in the sense that

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n} = 1. \quad \square$$

There is a short proof of this statement in [15]. Moreover, this is just a special case of the actual Stirling's approximation, which is possible since we deal with natural numbers. For the generalization concerning the gamma function and more information see [13].

4.4.1 Proof of Theorem 4.12

First, we will prove

$$(I) : P_{1,n} = P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset, case(i)\right) \to 0 \text{ as } n \to \infty.$$

Here, we act under the assumptions $X_{\ell:n} \leq r/2$ and $X_{k:n} > 1 - r/2$. As our calculations will depend on the order statistics $X_{k:n}$ and $X_{\ell:n}$, we need their joint probability distribution function (pdf), and it is

$$f_{k,\ell}(x_k, x_\ell) = \frac{n!}{(k-1)!(\ell-k-1)!(n-\ell)!} x_k^{k-1} \cdot (x_\ell - x_k)^{\ell-k-1} \cdot (1 - x_\ell)^{n-\ell},$$

for $0 < x_k < x_\ell < 1$ and the random sample $\mathcal{X}_n$ from the distribution $\mathcal{U}(0,1)$. Now we need to find the probability that $\mathcal{X}_n \cap \mathcal{I}_p = \emptyset$ given that $X_{k:n} = x$ and $X_{\ell:n} = y$. Recall that whenever case(i) holds, $\mathcal{I}_p = (X_{\ell:n}/r, 1 + (X_{k:n} - 1)/r)$. Let

$$g_1(x_k, x_\ell) = P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset | X_{k:n} = x_k, X_{\ell:n} = x_\ell\right)$$

Observe that for $\mathcal{X}_n \cap \mathcal{I}_p$ to be empty, we must have

$$X_{k:n} < X_{\ell:n}/r \text{ and } 1 + \frac{X_{k:n} - 1}{r} < X_{\ell:n},$$

for otherwise, $X_{k:n}$ or $X_{\ell:n}$ would be contained in $\mathcal{I}_p$. Moreover, we need none of the $\ell - k - 1$ sample points between $X_{k:n}$ and $X_{\ell:n}$ to be in $\mathcal{I}_p$. Hence

$$g_1(x_k, x_\ell) = \left(1 - \frac{1 + \frac{x_k - x_\ell - 1}{r}}{x_\ell - x_k}\right)^{\ell-k-1} = \left(\frac{r+1}{r} - \frac{r-1}{r}(x_\ell - x_k)^{-1}\right)^{\ell-k-1}.$$

Now that we have the probability g_1, conditioned on $X_{k:n} = x_k$ and $X_{\ell:n} = x_\ell$, we should multiply it with the joint pdf of $X_{k:n}$ and $X_{\ell:n}$ and integrate within their respective values. See that we want $X_{k:n} < X_{\ell:n}/r$ and $X_{\ell:n} > 1 + (X_{k:n} - 1)/r$, and hence both $X_{\ell:n} > rX_{k:n}$ and $X_{\ell:n} > (X_{k:n} + r - 1)/r$ must hold. The integral follows.

$$P_{1,n} = \int_0^{1/2} \int_{\max(rx_k, (x_k + r - 1)/r)}^1 f_{k,\ell}(x_k, x_\ell) g_1(x_k, x_\ell) \, dx_\ell \, dx_k \qquad (14)$$

Seeing that $rx_k > (x_k + r - 1)/r$ if and only if $x_k > 1/(r+1)$, we get the clearer form:

$$P_{1,n} = \left(\int_0^{1/(r+1)} \int_{(x_k+r-1)/r}^1 + \int_{1/(r+1)}^{1/2} \int_{rx_k}^1 \right) f_{k,\ell}(x_k, x_\ell) g_1(x_k, x_\ell)\, dx_\ell\, dx_k. \tag{15}$$

Equivalently we have

$$P_{1,n} = \kappa \left(\int_0^{1/(r+1)} \int_{(x_k+r-1)/r}^1 + \int_{1/(r+1)}^{1/2} \int_{rx_k}^1 \right) x_k^{k-1}(1-x_\ell)^{n-\ell} \left(\frac{r+1}{r}(x_\ell - x_k) - \frac{r-1}{r} \right)^{\ell-k-1} dx_\ell\, dx_k$$

where

$$\kappa = \frac{n!}{(k-1)!(\ell-k-1)!(n-\ell)!}.$$

By Lemma 4.11, for non-degeneracy in the limit, given $r > 2$, $p = r/2 > 1$ should hold, which is impossible. So the only case we investigate is $r \in (1,2)$, i.e., $p \in (1/2, 1)$. Let $m_p = \min\{1-p, 1/2-(1-p)\} = \min\{1-p, p-1/2\}$, and $\varepsilon \in (0, m_p/8) = (0, \min\{(1-p)/8, (p-1/2)/8\}) \subseteq (0, 1/16)$. By Proposition 4.18 and under the conditions (C1) and (C4) (i.e., $X_{\ell:n} \le p$ as well as $X_{k:n} > 1-p$) we see that for sufficiently large n

$$P_{1,n} \approx \kappa \int_{1-p}^{1-p+\varepsilon} \int_{p-\varepsilon}^{p} x_k^{k-1}(1-x_\ell)^{n-\ell} \left(\frac{r+1}{r}(x_\ell - x_k) - \frac{r-1}{r} \right)^{\ell-k-1} dx_\ell\, dx_k.$$

Now we will translate the variables to the origin, considering the change of variables $z_k := x_k - (1-p)$ and $z_\ell := p - x_\ell$, where $z_k, z_\ell \ge 0$, or, equivalently,

$$x_k = z_k + (1-p),$$
$$x_\ell = p - z_\ell,$$

and see that we have the following integral:

$$P_{1,n} \approx \kappa \int_0^{\varepsilon} \int_0^{\varepsilon} (z_k + 1 - p)^{k-1}(1 - p + z_\ell)^{n-\ell} \left(\frac{r+1}{r}(2p - 1 - z_\ell - z_k) - \frac{r-1}{r} \right)^{\ell-k-1} dz_\ell\, dz_k.$$

Using Stirling's Approximation we write κ in the following form.

$$\kappa \approx \frac{\sqrt{n}.n^n}{2\pi e^2(n-\ell)^{n-\ell}(\ell-k-1)^{\ell-k-1}(k-1)^{k-1}\sqrt{(n-\ell)(\ell-k-1)(k-1)}}$$

or, equivalently,

$$\kappa \approx \frac{\sqrt{n}.n^n}{2\pi e^2(n-\ell)^{(n-\ell+1/2)}(\ell-k-1)^{(\ell-k-1/2)}(k-1)^{(k-1/2)}}.$$

Then we reorder the terms in the following way so that calculations will be more convenient

$$\kappa \approx \frac{n}{2\pi e^2}\left(\frac{n}{n-\ell}\right)^{n-\ell+1/2}\left(\frac{n}{\ell-k-1}\right)^{\ell-k-1/2}\left(\frac{n}{k-1}\right)^{k-1/2},$$

and for large enough $n \in \mathbb{N}$ we have

$$\kappa \approx \frac{n}{2\pi e^2}\left(\frac{1}{1-p}\right)^{n(1-p)+1/2}\left(\frac{1}{2p-1}\right)^{n(2p-1)-1/2}\left(\frac{1}{1-p}\right)^{n(1-p)-1/2}. \tag{16}$$

Proposition 4.19 helps us to organize κ into terms such that they can easily be incorporated into the integrands at hand. For sufficiently large $n \in \mathbb{N}$ we have

$$P_{1,n} \approx \int_0^{\varepsilon}\int_0^{\varepsilon} \frac{n\sqrt{2p-1}}{2\pi e^2(1-p)}\left(1+\frac{z_k}{1-p}\right)^{n(1-p)}\left(1+\frac{z_\ell}{1-p}\right)^{n(1-p)} \left(1-\frac{2p+1}{2p(2p-1)}(z_\ell+z_k)\right)^{n(2p-1)} dz_\ell\, dz_k. \tag{17}$$

Observe that this integral is critical at $z_k = z_\ell = 0$, since we have all the integrands $(1)^{n(1-p)}$, $(1)^{n(1-p)}$ and $(1)^{n(2p-1)}$. So we need further to investigate its limit. As we want to make use of the fact that the integrands converge to exponential terms in the limit, we consider the following change of variables $u_k := nz_k$ and $u_\ell := nz_\ell$, or, equivalently,

$$z_k = u_k/n,$$
$$z_\ell = u_\ell/n,$$

then the integral (17) becomes

$$P_{1,n} \approx \int_0^{n\varepsilon}\int_0^{n\varepsilon} \frac{\sqrt{2p-1}}{2n\pi e^2(1-p)}\left(1+\frac{u_k}{n(1-p)}\right)^{n(1-p)}\left(1+\frac{u_\ell}{n(1-p)}\right)^{n(1-p)}$$
$$\left(1-\frac{2p+1}{n2p(2p-1)}(u_\ell+u_k)\right)^{n(2p-1)} du_\ell\, du_k.$$

As $n \to \infty$, in the limit,

$$P_{1,n} \approx O\left(n^{-1}\right)\int_0^{\infty}\int_0^{\infty} e^{(u_k+u_\ell)-(u_\ell+u_k)(1+1/2p)}\, du_\ell\, du_k.$$

since as $n \to \infty$, for any $f, g : \mathbb{R} \to \mathbb{R}$,

$$\left(1+\frac{f(x)}{n}\right)^{ng(x)} \to e^{g(x)f(x)}.$$

See that

$$P_{1,n} \approx O\left(n^{-1}\right)\int_0^{\infty}\int_0^{\infty} e^{(u_k+u_\ell)\frac{-1}{2p}}\, du_\ell\, du_k,$$

as $\alpha := -1/2p < 0$ we have

$$P_{1,n} \approx O\left(n^{-1}\right)\int_0^{\infty} e^{\alpha u_k}\left(\int_0^{\infty} e^{\alpha u_\ell}\, du_\ell\right) du_k,$$

Therefore

$$P_{1,n} \approx O\left(n^{-1}\right)\frac{1}{\alpha^2}.$$

Thus

$$\lim_{n\to\infty} P_{1,n} = \lim_{n\to\infty} P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset, case(i)\right) = 0.$$

Hence we proved (I). Now we will prove

$$(II) : P_{2,n} = P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset, case(ii)\right) \to 0 \text{ as } n \to \infty.$$

Now we have the assumptions that $X_{\ell:n} \leq r/2$ and $X_{k:n} \leq 1 - r/2$. Conditioning $X_{k:n} = x_k$ and $X_{\ell:n} = x_\ell$, we want the probability $\mathcal{X}_n \cap \mathcal{I}_p$ being empty. Recall that whenever it is the case(ii), $\mathcal{I}_p = (X_{\ell:n}, 1/2)$. Assuming the conditioning in question we define

$$g_2(x_k, x_\ell) = P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset | X_{k:n} = x_k \,\&\, X_{\ell:n} = x_\ell\right).$$

Notice that

$$X_{k:n} < \frac{X_{\ell:n}}{r} \qquad \text{and} \qquad X_{\ell:n} > 1/2$$

for $\mathcal{X}_n \cap \mathcal{I}_p$ to be empty as otherwise $X_{k:n}$ or $X_{\ell:n}$ would be in it. Also we want the $\ell - k - 1$ points of the random sample between $X_{k:n}$ and $X_{\ell:n}$ not to be in $\mathcal{I}_p$. Identical to what we did in case (I), we get

$$g_2(x_k, x_\ell) = \left(1 - \frac{1/2 - x_\ell/r}{x_\ell - x_k}\right)^{\ell-k-1} = \left(1 - \frac{r - 2x_\ell}{(x_\ell - x_k)2r}\right)^{\ell-k-1}.$$

Thus we obtain the integral

$$P_{2,n} = \int_{1/2}^{1} \int_{0}^{x_\ell/r} \kappa\, x_k^{k-1}(1 - x_\ell)^{n-\ell} \left((1 + 1/r)\, x_\ell - x_k - 1/2\right)^{\ell-k-1} \, dx_k \, dx_\ell.$$

Again, we calculate $\lim_{n\to\infty} P_{2,n}$. By Proposition 4.18 and under the conditions (C1) and (C3) (i.e., $X_{\ell:n} \leq p$ and $X_{k:n} \leq 1-p$) we write for sufficiently large $n \in \mathbb{N}$

$$P_{2,n} \approx \int_{p-\varepsilon}^{p} \int_{1-p-\varepsilon}^{1-p} \kappa\, x_k^{k-1}(1 - x_\ell)^{n-\ell} \left((1 + 1/r)x_\ell - x_k - 1/2\right)^{\ell-k-1} \, dx_k \, dx_\ell.$$

For the convenience we will translate the variables to origin, considering the change of variables $z_k := (1 - p) - x_k$ and $z_\ell := p - x_\ell$, where $z_k, z_\ell \geq 0$, or, equivalently,

$$\begin{aligned} x_k &= (1 - p) - z_k, \\ x_\ell &= p - z_\ell, \end{aligned}$$

to get that

$$\begin{aligned} P_{2,n} \approx \int_{0}^{\varepsilon} \int_{0}^{\varepsilon} \kappa\, &(1 - p - z_k)^{k-1}(1 - p + z_\ell)^{n-\ell} \\ &(2p - 1 + z_k - (1 + 1/2p)z_\ell)^{\ell-k-1} \, dz_k \, dz_\ell. \end{aligned}$$

We will again use the Stirling's Approximation for κ, the equation (16), to get that

$$P_{2,n} \approx \int_0^{\varepsilon}\int_0^{\varepsilon} \frac{n\sqrt{2p-1}}{2\pi e^2(1-p)} \left(1 - \frac{z_k}{1-p}\right)^{n(1-p)} \left(1 + \frac{z_\ell}{1-p}\right)^{n(p-1)} \left(1 + \frac{z_k - (1+1/2p)z_\ell}{2p-1}\right)^{n(2p-1)} dz_k\, dz_\ell. \quad (18)$$

Again, note that the integral is critical at $z_k = z_\ell = 0$. We will again use the fact that integrands converge to exponential terms with the help of the change of variables, $u_k := nz_k$ and $u_\ell := nz_\ell$, or, equivalently,

$$\begin{aligned} z_k &= u_k/n, \\ z_\ell &= u_\ell/n, \end{aligned}$$

the integral in Equation (18) becomes

$$P_{2,n} \approx \int_0^{n\varepsilon}\int_0^{n\varepsilon} \frac{\sqrt{2p-1}}{2n\pi e^2(1-p)} \left(1 - \frac{u_k}{n(1-p)}\right)^{n(1-p)} \left(1 + \frac{u_\ell}{n(1-p)}\right)^{n(1-p)} \left(1 + \frac{u_k - (1+1/2p)u_\ell}{n(2p-1)}\right)^{n(2p-1)} du_k\, du_\ell.$$

For sufficiently large $n \in \mathbb{N}$,

$$\begin{aligned} P_{2,n} &\approx \int_0^{n\varepsilon}\int_0^{n\varepsilon} \frac{\sqrt{2p-1}}{2n\pi e^2(1-p)} e^{-u_\ell/2p}\, du_k\, du_\ell \\ &= \int_0^{n\varepsilon} \frac{\varepsilon\sqrt{2p-1}}{2\pi e^2(1-p)} e^{-u_\ell/2p}\, du_\ell \end{aligned}$$

and as $n \to \infty$ we have

$$P_{2,n} \approx \int_0^{\infty} c_1 \varepsilon e^{-u_\ell/2p}\, du_\ell$$

where $c_1 \in \mathbb{R}$. Therefore

$$P_{2,n} \approx c_1 \alpha \varepsilon = c_2 \varepsilon$$

for some $c_2 \in \mathbb{R}$. Since this is true for any sufficiently small and arbitrary $\varepsilon > 0$, letting $\varepsilon \to 0$, we get $\lim_{n\to\infty} P_{2,n} = 0$. Thus follows the proof of claim (II). The last claim to prove is

$$(III) : P_{3,n} = P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset, case(iii)\right) \to 0 \text{ as } n \to \infty.$$

This is quite similar to the case (ii). We show that there is symmetry between cases (ii) and (iii) and then directly make use of it.

Proposition 4.21. *Considering the events of case(ii) and case(iii), there is symmetry in the following sense*

$$P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset, case(iii)\right) = P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset, case(ii)\right).$$

Proof. For case(iii) we have $\mathcal{I}_p = [1/2, 1 + (X_{k:n} - 1)/r)$ and for $\mathcal{X}_n \cap \mathcal{I}_p$ to be empty, $X_{k:n} < 1/2$, $X_{\ell:n} > 1 + (X_{k:n} - 1)/r$ and no element of the random sample between $X_{k:n}$ and $X_{\ell:n}$ should be in $\mathcal{I}_p$. Let $\xi_i := 1 - X_{(n+1-i):n}$ for $i = 1, 2, \ldots, n$. See that the set $\{\xi_1, \xi_2, \ldots, \xi_n\}$ is also a random sample from the distribution $\mathcal{U}(0,1)$. Recalling that $\ell = \lceil np \rceil$ and $k = n - \lfloor np \rfloor$, the equality $\ell = n + 1 - k$ follows from

$$n + 1 - k = n + 1 - n + \lfloor np \rfloor = \lfloor np \rfloor + 1 = \lceil np \rceil.$$

Thus we have

$$\begin{aligned}\xi_k &= 1 - X_{n+1-k:n} = 1 - X_{\ell:n},\\ \xi_\ell &= 1 - X_{n+1-\ell:n} = 1 - X_{k:n}.\end{aligned}$$

See that

$$X_{k:n} < 1/2 \Leftrightarrow 1 - \xi_\ell < 1/2 \Leftrightarrow \xi_\ell > 1/2$$

as well as

$$X_{\ell:n} > 1 + \frac{X_{k:n} - 1}{r} \Leftrightarrow 1 - \xi_k > 1 + \frac{1 - \xi_\ell - 1}{r} \Leftrightarrow \xi_k < \frac{\xi_\ell}{r}.$$

Also $X_{i:n} \in \mathcal{I}_p = [1/2, 1 + (X_{k:n} - 1)/r)$ if and only if

$$\begin{aligned}\frac{1}{2} \le X_{i:n} \le 1 + \frac{X_{k:n} - 1}{r} &\Leftrightarrow\\ \frac{-1}{2} \ge -X_{i:n} \ge -1 + \frac{1 - X_{k:n}}{r} &\Leftrightarrow\\ \frac{1}{2} \ge 1 - X_{i:n} \ge \frac{1 - X_{k:n}}{r} &\Leftrightarrow\\ \frac{1}{2} \ge \xi_i \ge \frac{\xi_\ell}{r} &\Leftrightarrow\\ \xi_i \in \left(\frac{\xi_\ell}{r}, \frac{1}{2}\right) &=: \mathcal{I}_p^*.\end{aligned}$$

Notice that the set $\mathcal{I}_p^*$ is the set $\mathcal{I}_p$ for case(ii). Hence cases(ii) and (iii) are symmetric in the sense that $\mathcal{X}_n \cap \mathcal{I}_p = \emptyset$ occurs, i.e., $P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset, case(iii)\right) = P\left(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset, case(ii)\right)$. ☐

Therefore, by Proposition 4.21 we see that $P(\mathcal{X}_n \cap \mathcal{I}_p = \emptyset, case(iii)) \to 0$ as $n \to \infty$. Hence we have the desired result, concluding the proof of Theorem 4.12.

4.4.2 Proof of Theorem 4.14

Recall that in this case we want $X_{k:n} < 1-p$ and $X_{\ell:n} > p$. So letting $P_{4,n} := P(case(iv))$ we have

$$P_{4,n} = \int_p^1 \int_0^{1-p} f_{k,\ell}(x_k, x_\ell)\, dx_k\, dx_\ell.$$

By Proposition 4.18 we have, for sufficiently large $n \in \mathbb{N}$,

$$P_{4,n} \approx \int_p^{p+\varepsilon} \int_{1-p-\varepsilon}^{1-p} \kappa x_k^{k-1}(1-x_\ell)^{n-\ell}(x_\ell - x_k)^{\ell-k-1}\, dx_k\, dx_\ell. \tag{19}$$

After we move the integral variables to the neighborhoods of zero with the change of variables $z_k := (1-p) - x_k$ and $z_\ell := x_\ell - p$, i.e.,

$$\begin{aligned} x_k &= (1-p) - z_k, \\ x_\ell &= z_\ell + p, \end{aligned}$$

the integral (19) becomes

$$P_{4,n} = \int_0^{\varepsilon} \int_0^{\varepsilon} \kappa(1-p-z_k)^{k-1}(1-p-z_\ell)^{n-\ell}(2p-1+z_\ell+z_k)^{\ell-k-1}\, dz_k\, dz_\ell,$$

or, equivalently,

$$P_{4,n} = \int_0^{\varepsilon} \int_0^{\varepsilon} \kappa \frac{1}{(1-p-z_k)(2p-1+z_k+z_\ell)}(1-p-z_k)^k(1-p-z_\ell)^{n-\ell} \\ (2p-1+z_\ell+z_k)^{\ell-k}\, dz_k\, dz_\ell.$$

Now we decompose κ so that we will get rid of the exponential powers in the Stirling's approximation, seeing that

$$\kappa = \frac{n!}{(k-1)!(\ell-k-1)!(n-\ell)!} = n(n-1)\frac{(n-2)!}{(k-1)!(\ell-k-1)!(n-\ell)!},$$

and applying the Stirling's approximation to the factorial terms, $(n-2)!$, $(k-1)!$, $(\ell-k-1)!$ and $(n-\ell)!$, we obtain for sufficiently large $n \in \mathbb{N}$,

$$\kappa \approx \frac{n}{2\pi}\frac{(n-1)}{\sqrt{k-1}\sqrt{n-\ell}}\sqrt{\frac{n-2}{\ell-k-1}}\left(\frac{n-2}{k-1}\right)^{k-1}\left(\frac{n-2}{n-\ell}\right)^{n-\ell}\left(\frac{n-2}{\ell-k-1}\right)^{\ell-k-1}.$$

By Lemma 4.19,

$$\kappa \approx \frac{n\sqrt{2p-1}}{2\pi}\left(\frac{1}{1-p}\right)^{n(1-p)}\left(\frac{1}{1-p}\right)^{n(1-p)}\left(\frac{1}{2p-1}\right)^{n(2p-1)},$$

and for sufficiently large $n \in \mathbb{N}$ we get,

$$\kappa \approx \frac{n\sqrt{2p-1}}{2\pi}\left(\frac{1}{1-p}\right)^{2n(1-p)}\left(\frac{1}{2p-1}\right)^{n(2p-1)}.$$

We organize the integral in the following way

$$P_{4,n} \approx \int_0^{\varepsilon}\int_0^{\varepsilon}\frac{n\sqrt{2p-1}}{2\pi}C_1^{-n}\frac{1}{(1-p-z_k)(2p-1+z_k+z_\ell)}g(z_k,z_\ell)\,dz_k\,dz_\ell,$$

where

$$C_1 = (1-p)^{2(1-p)}(2p-1)^{2p-1}$$

and

$$g(z_k,z_\ell) = (1-p-z_k)^{n(1-p)}(1-p-z_\ell)^{n(1-p)}(2p-1+z_k+z_\ell)^{n(2p-1)},$$

since, for large $n \in \mathbb{N}$, $k \approx n(1-p)$ and $\ell \approx np$. Now let

$$h(z_k,z_\ell) = (1-p-z_k)^{1-p}(1-p-z_\ell)^{1-p}(2p-1+z_k+z_\ell)^{2p-1},$$

so that

$$g(z_k,z_\ell) = (h(z_k,z_\ell))^n\,.$$

The integral is critical at $(z_k,z_\ell) = (0,0)$, since $h(0,0) = C_1$, and so

$$C_1^{-n}g(0,0) = 1.$$

We write the bivariate Taylor expansion of $\frac{1}{(1-p-z_k)(2p-1+z_k+z_\ell)}$ up to order one and $h(z_k,z_\ell)$ up to order two, around $(z_k,z_\ell) = (0,0)$. See

$$\frac{1}{(1-p-z_k)(2p-1+z_k+z_\ell)} = \frac{1}{(1-p)(2p-1)} + O(z_k) + O(z_\ell)$$

and

$$h(z_k, z_\ell) = \frac{4p^2 - 6p + 2 + p(z_k^2 + z_\ell^2) - 2z_k z_\ell(p-1)}{2(2p-1)(p-1)} + O(z_k^2) + O(z_\ell^2) + O(z_k z_\ell).$$

Organizing the terms

$$h(z_k, z_\ell) = \left(1 - \frac{p(z_k^2 + z_\ell^2) + 2z_k z_\ell(1-p)}{2(2p-1)(p-1)}\right) + O(z_k^2) + O(z_\ell^2) + O(z_k z_\ell),$$

letting

$$z_k = \frac{w_k}{\sqrt{n}} \text{ and } z_\ell = \frac{w_\ell}{\sqrt{n}},$$

we obtain

$$P_{4,n} \approx \int_0^{\sqrt{n}\varepsilon} \int_0^{\sqrt{n}\varepsilon} \frac{\sqrt{2p-1}}{2\pi} \left(\frac{1}{(1-p)(2p-1)} + O(n^{-1/2})\right)$$
$$\left(1 - 1/n \left(\frac{p(w_k^2 + w_\ell^2) + 2w_k w_\ell(1-p)}{2(2p-1)(p-1)}\right) + O(n^{-1})\right)^n dw_k \, dw_\ell,$$

and letting $n \to \infty$

$$\lim_{n\to\infty} P_{4,n} = \int_0^\infty \int_0^\infty \frac{1}{2\pi(1-p)\sqrt{2p-1}} \; e^{\left(-\frac{p(w_k^2+w_\ell^2)+2w_k w_\ell(1-p)}{2(2p-1)(1-p)}\right)} dw_k \, dw_\ell.$$

Now we switch to polar coordinates, so we make change of variables,

$$w_k = a\cos(t) \text{ and } w_\ell = a\sin(t),$$

$$\lim_{n\to\infty} P_{4,n} = \int_0^{\pi/2} \int_0^\infty \frac{1}{2\pi(1-p)\sqrt{2p-1}} \; ae^{\left(-\frac{pa^2+2(1-p)a^2\cos(t)\sin(t)}{2(2p-1)(1-p)}\right)} da \, dt,$$

$$
\begin{aligned}
\lim_{n\to\infty} P_{4,n} &= \int_0^{\pi/2} -\frac{\sqrt{2p-1}}{2\pi(\sin(2t)p-\sin(2t)-p)}\,dt, \\
&= \lim_{x\to\pi/2}\left[\frac{\arctan\left(\frac{\tan(t)p-p+1}{\sqrt{2p-1}}\right)}{2\pi}\right]_{t=0}^{t=x} \\
&= \frac{1}{4}-\frac{1}{2\pi}\arctan\left(\frac{1-p}{\sqrt{2p-1}}\right).
\end{aligned}
$$

Therefore,

$$
\pi_p = \lim_{n\to\infty} P_{4,n} = \lim_{n\to\infty} P\left(case(iv)\right) = \frac{1}{4}-\frac{1}{2\pi}\arctan\left(\frac{1-p}{\sqrt{2p-1}}\right), \tag{20}
$$

and the proof of Theorem 4.14 follows. □

4.5 Monte Carlo Simulations

In this section we present the Monte Carlo simulation results and compare them with the theoretical results we obtained in the previous sections. We will consider three different set of values, $p \in (1/2, 1)$, fractional domination parameter, $r \in (1, 2)$, extension parameter and $n \in \mathbb{N}$, number of vertices. We will get a random sample of size n, subject to uniform distribution, and then observe the behavior of $\gamma_{\geq p}$, estimating the probability of the event $\gamma_{\geq p}(D) = 1$ for various combinations of n, p and r. In particular we choose $r = 1.1$, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9; $p = 0.55$, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95 for $n = 100$, 200, 500, 1000, 2000, 5000, 10000, 20000, 50000, 100000. The following is an sample R code that was used in our simulations.

```
Nmc<-1000
p<-0.8
r<-2*p
n<-100

nlist<-(1:10)*100
set.seed(1)

pg1<-0
k<-n-floor(n*p)+1
ell<-ceiling(n*p)

for (i in 1:Nmc)
{
Xp<-runif(n)
```

```
SX<-sort(Xp)
indl<- (1:n)[Xp<=1/2]
XL<-Xp[indl]
XR<-Xp[-indl]
if (r*max(XL) >= SX[ell] | 1-r*(1-min(XR)) <= SX[k] )
{pg1<-pg1+1}
}

print(pg1/Nmc)
```

Here, we present the simulation results.

r=1.1

n\p	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
100	0.802	0.259	0.035	0	0	0	0	0	0
200	0.821	0.127	0.006	0	0	0	0	0	0
500	0.870	0.017	0	0	0	0	0	0	0
1000	0.879	0.001	0	0	0	0	0	0	0
2000	0.883	0	0	0	0	0	0	0	0
5000	0.895	0	0	0	0	0	0	0	0
10000	0.902	0	0	0	0	0	0	0	0
20000	0.908	0	0	0	0	0	0	0	0
50000	0.907	0	0	0	0	0	0	0	0
100000	0.905	0	0	0	0	0	0	0	0

The calculated value of $1-\pi_p$, from Equation (20), for $p = 0.55$ is 0.902.

r=1.2

n\p	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
100	.992	0.827	0.248	0.03	0	0	0	0	0
200	1	0.802	0.119	0.005	0	0	0	0	0
500	1	0.811	0.023	0	0	0	0	0	0
1000	1	0.845	0.001	0	0	0	0	0	0
2000	1	0.856	0	0	0	0	0	0	0
5000	1	0.865	0	0	0	0	0	0	0
10000	1	0.869	0	0	0	0	0	0	0
20000	1	0.872	0	0	0	0	0	0	0
50000	1	0.865	0	0	0	0	0	0	0
100000	1	0.860	0	0	0	0	0	0	0

The calculated value of $1 - \pi_p$ for $p = 0.6$ is 0.866.

r=1.3

n\p	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
100	1	0.990	0.792	0.030	0	0	0	0	0
200	1	1	0.795	0.005	0	0	0	0	0
500	1	1	0.792	0	0	0	0	0	0
1000	1	1	0.810	0	0	0	0	0	0
2000	1	1	0.819	0	0	0	0	0	0
5000	1	1	0.842	0	0	0	0	0	0
10000	1	1	0.853	0	0	0	0	0	0
20000	1	1	0.854	0	0	0	0	0	0
50000	1	1	0.838	0	0	0	0	0	0
100000	1	1	0.840	0	0	0	0	0	0

The calculated value of $1 - \pi_p$ for $p = 0.65$ is 0.840.

r=1.4

n\p	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
100	1	1	0.988	0.773	0	0	0	0	0
200	1	1	1	0.755	0	0	0	0	0
500	1	1	1	0.769	0	0	0	0	0
1000	1	1	1	0.793	0	0	0	0	0
2000	1	1	1	0.788	0	0	0	0	0
5000	1	1	1	0.821	0	0	0	0	0
10000	1	1	1	0.809	0	0	0	0	0
20000	1	1	1	0.831	0	0	0	0	0
50000	1	1	1	0.825	0	0	0	0	0
100000	1	1	1	0.833	0	0	0	0	0

The calculated value of $1 - \pi_p$ for $p = 0.7$ is 0.820.

r=1.5

n\p	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
100	1	1	1	0.986	0.732	0.173	0	0	0
200	1	1	1	0.998	0.728	0.076	0	0	0
500	1	1	1	1	0.745	0.004	0	0	0
1000	1	1	1	1	0.770	0	0	0	0
2000	1	1	1	1	0.785	0	0	0	0
5000	1	1	1	1	0.787	0	0	0	0
10000	1	1	1	1	0.805	0	0	0	0
20000	1	1	1	1	0.814	0	0	0	0
50000	1	1	1	1	0.800	0	0	0	0
100000	1	1	1	1	0.812	0	0	0	0

The calculated value of $1 - \pi_p$ for $p = 0.75$ is 0.804.

r=1.6

n\p	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
100	1	1	1	1	0.980	0.694	0.141	0.005	0
200	1	1	1	1	0.996	0.702	0.061	0	0
500	1	1	1	1	1	0.743	0.004	0	0
1000	1	1	1	1	1	0.765	0	0	0
2000	1	1	1	1	1	0.762	0	0	0
5000	1	1	1	1	1	0.776	0	0	0
10000	1	1	1	1	1	0.782	0	0	0
20000	1	1	1	1	1	0.783	0	0	0
50000	1	1	1	1	1	0.796	0	0	0
100000	1	1	1	1	1	0.795	0	0	0

The calculated value of $1 - \pi_p$ for $p = 0.8$ is 0.790.

r=1.7

n\p	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
100	1	1	1	1	0.999	0.977	0.657	0.100	0
200	1	1	1	1	1	0.997	0.686	0.034	0
500	1	1	1	1	1	1	0.749	0	0
1000	1	1	1	1	1	1	0.739	0	0
2000	1	1	1	1	1	1	0.753	0	0
5000	1	1	1	1	1	1	0.756	0	0
10000	1	1	1	1	1	1	0.761	0	0
20000	1	1	1	1	1	1	0.746	0	0
50000	1	1	1	1	1	1	0.748	0	0
100000	1	1	1	1	1	1	0.786	0	0

The calculated value of $1 - \pi_p$ for $p = 0.85$ is 0.778.

r=1.8

n\p	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
100	1	1	1	1	1	1	1	0.618	0.052
200	1	1	1	1	1	1	1	0.671	0.005
500	1	1	1	1	1	1	1	0.703	0
1000	1	1	1	1	1	1	1	0.725	0
2000	1	1	1	1	1	1	1	0.729	0
5000	1	1	1	1	1	1	1	0.738	0
10000	1	1	1	1	1	1	1	0.741	0
20000	1	1	1	1	1	1	1	0.751	0
50000	1	1	1	1	1	1	1	0.744	0
100000	1	1	1	1	1	1	1	0.760	0

The calculated value of $1 - \pi_p$ for $p = 0.9$ is 0.767.

r=1.9

n\p	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95
100	1	1	1	1	1	1	1	0.980	0.612
200	1	1	1	1	1	1	1	0.999	0.634
500	1	1	1	1	1	1	1	1	0.659
1000	1	1	1	1	1	1	1	1	0.692
2000	1	1	1	1	1	1	1	1	0.694
5000	1	1	1	1	1	1	1	1	0.719
10000	1	1	1	1	1	1	1	1	0.723
20000	1	1	1	1	1	1	1	1	0.758
50000	1	1	1	1	1	1	1	1	0.725
100000	1	1	1	1	1	1	1	1	0.744

The calculated value of $1 - \pi_p$ for $p = 0.95$ is 0.758.

The results of our simulations are in agreement with our main result in Theorem 4.15. Whenever $r > 2p$, the probability that $(\gamma_{\geq p} = 1)$ converges rapidly to 1, i.e., strong p-domination with just one element occurs almost surely. On the other hand, whenever $r < 2p$, the probability that $(\gamma_{\geq p} = 1)$ converges rapidly to 0, i.e., it is impossible to obtain strong p-domination with just one element, hence in the limit $\gamma_{\geq p}$ has the degenerate distribution at 2. For the case $r = 2p$, for which we showed the non-degeneracy, we added the calculated value for $1 - \pi_p$ (which is equivalent to the probability $(\gamma_{\geq p} = 1)$), where $p = r/2$ for specified values of r, and despite the fluctuations due to randomness caused by Monte Carlo simulations, we can see that values in simulations and calculated values are in considerable agreement.

5 CONCLUSIONS

In this study, we provide various graph theoretical extensions of domination, and their relation to the usual form of domination in literature. The intuition of these extensions is the possibility of the need for fractional analysis of graphs and domination. By eliminating the restriction of dominating all vertices, may be excluding most of the isolated ones, the 'effort' to acquire the proposed domination might decreases immensely. The extensions of domination we proposed are exact p-domination, weak p-domination and strong p-domination. We mainly focused on strong and weak p-domination on Class Cover Catch Digraphs (CCCDs) and Proximity Catch Digraphs (PCDs). For the exact p-domination, we conjectured non-degeneracy in the limit. We also showed that on CCCDs, whenever we decrease p from 1 to any real number in $(0, 1)$, we switch from non-degenerate to degenerate distribution of $\gamma_{\geq p}$ in the limit. That is, the pdf of $\gamma_{\geq p}$ for the case $p = 1$ is a threshold pdf, switching from non-degeneracy to degeneracy. The degeneracy and non-degeneracy of the asymptotic distribution of $\gamma_{\geq p}$ is due to the relation between the domination parameter p, and the extension parameter r of PCDs and our main result is based on this fact. In particular, we demonstrate that as $n \to \infty$, the PCD with $c = 1/2$ has degenerate distribution when $r \neq 2p$ and non-degenerate distribution only if $r = 2p$. This gives a rough description of the domination-related behavior of the CCCDs and PCDs for large sample populations.

Using Monte Carlo simulations, we showed that our estimated probabilities are in agreement with the theoretical probabilities. In quest for more generality, the 'centrality parameter' of PCDs may be added to the study of these extensions of domination on PCDs. This work also forms the foundation of future work about domination on CCCDs and PCDs based on higher dimensional data.

References

[1] L.J. Bain, M. Engelhardt, *Introduction to Probability and Mathematical Statistics*, Duxburry Press, 1992.

[2] B. Bollabas, *Modern Graph Theory*, Springer, 1998.

[3] B. Bollabas, *Random Graphs*, Cambridge University Press, 2001.

[4] K.L. Chung, *A Course in Probability Theory*, Academic Press, 1974.

[5] H.A. David, H.N. Nagaraja, *Order Statistics*, Wiley-Interscience, 2003.

[6] G.B. Folland, *Real Analysis: Modern Techniques and Their Applications*, Wiley-Interscience, 1999.

[7] T.W. Haynes, S.T. Hedetniemi, P.J. Slater, *Fundamentals of Domination in Graphs*, Marcel Dekker, 1998.

[8] D.B. West, *Introduction to Graph Theory*, Pearson, 2001.

[9] E. Beer, J.A. Fill, S. Janson, E. Scheinerman, On Vertex, Edge, and Vertex-Edge Random Graphs, arxiv.org/pdf/0812.1410, 2010.

[10] E. Ceyhan, The distribution of the domination number of class cover catch digraphs for non-uniform one-dimensional data, Discrete Mathematics - 308, 2008, p5376-5393.

[11] E. Ceyhan, The Distribution of the domination number of a family of random interval catch digraphs, 2010. (To appear)

[12] E. Ceyhan, C.E. Priebe, The Use of Domination Number of a Random Proximity Catch Digraph for Testing Spatial Patterns of Segregation and Association, Statistics & Probability Letters - 73, 2005, p37-50.

[13] Y.-C. Li, A Note on an identity of the gamma function and Stirling's formula, Real Analysis Exchange - 32, 2006/2007, p267-272.

[14] C.E. Priebe, J.G. DeVinney, D.J. Marchette, On the distribution of the domination number of random class catch cover digraphs, Statistics and Probability Letters - 55, 2001, p239-246.

[15] D. Romik, Stirling's approximation for n!: the ultimate short proof?, The American Mathematical Monthly - 107, 2000, p556-557.