Model Assisted Statistics and Applications 1 (2005,2006) 231–255
IOS Press

On the distribution of the domination number of a new family of parametrized random digraphs[1]

Elvan Ceyhan[a,*] and Carey E. Priebe[b]
[a]*Department of Mathematics, Koç University, Sarıyer, 34450, Istanbul, Turkey*
[b]*Department of Applied Mathematics and Statistics, The Johns Hopkins University, Baltimore, MD, 21218, USA*

Abstract. We derive the asymptotic distribution of the domination number of a new family of random digraph called proximity catch digraph (PCD), which has application to statistical testing of spatial point patterns and to pattern recognition. The PCD we use is a parametrized digraph based on two sets of points on the plane, where sample size and locations of the elements of one is held fixed, while the sample size of the other whose elements are randomly distributed over a region of interest goes to infinity. PCDs are constructed based on the relative allocation of the random set of points with respect to the Delaunay triangulation of the other set whose size and locations are fixed. We introduce various auxiliary tools and concepts for the derivation of the asymptotic distribution. We investigate these concepts in one Delaunay triangle on the plane, and then extend them to the multiple triangle case. The methods are illustrated for planar data, but are applicable in higher dimensions also.

Keywords: Random graph, domination number, proximity map, delaunay triangulation, proximity catch digraph

1. Introduction

The *proximity catch digraphs* (PCDs) are a special type of *proximity graphs* which were introduced by [19]. A *digraph* is a directed graph with vertices V and arcs (directed edges) each of which is from one vertex to another based on a binary relation. Then the pair $(p, q) \in V \times V$ is an ordered pair which stands for an arc (directed edge) from vertex p to vertex q. For example, the *nearest neighbor* (*di*)*graph* of [15] is a proximity digraph. The nearest neighbor digraph has the vertex set V and (p, q) as an arc iff q is a nearest neighbor of p.

Our PCDs are based on the proximity maps which are defined in a fairly general setting. Let $(\Omega, \mathcal{M})$ be a measurable space. The *proximity map* $N(\cdot)$ is defined as $N : \Omega \to 2^{\Omega}$, where 2^{Ω} is the power set of Ω. The *proximity region* of $x \in \Omega$, denoted $N(x)$, is the image of $x \in \Omega$ under $N(\cdot)$. The points in $N(x)$ are thought of as being "closer" to $x \in \Omega$ than are the points in $\Omega \setminus N(x)$. Hence the term "proximity" in the name *proximity catch digraph*. Proximity maps are the building blocks of the *proximity graphs* of [19]; an extensive survey on proximity maps and graphs is available in [12].

The *proximity catch digraph* D has the vertex set $\mathcal{V} = \{p_1, \ldots, p_n\}$; and the arc set $\mathcal{A}$ is defined by $(p_i, p_j) \in \mathcal{A}$ iff $p_j \in N(p_i)$ for $i \neq j$. Notice that the proximity catch digraph D depends on the *proximity* map $N(\cdot)$ and if $p_j \in N(p_i)$, then we call $N(p_i)$ (and hence point p_i) *catches* p_j. Hence the term "catch" in the name *proximity catch digraph*. If arcs of the form (p_j, p_j) (i.e., loops) were allowed, D would have been called a *pseudodigraph* according to some authors (see, e.g. [7]).

[1]This research was supported by the Defense Advanced Research Projects Agency as administered by the Air Force Office of Scientific Research under contract DOD F49620-99-1-0213 and by Office of Naval Research Grant N00014-95-1-0777.
*Corresponding author. E-mail: elceyhan@ku.edu.tr.

ISSN 1574-1699/05/06/$17.00

In a digraph $D = (\mathcal{V}, \mathcal{A})$, a vertex $v \in \mathcal{V}$ *dominates* itself and all vertices of the form $\{u : (v, u) \in \mathcal{A}\}$. A *dominating set* S_D for the digraph D is a subset of $\mathcal{V}$ such that each vertex $v \in \mathcal{V}$ is dominated by a vertex in S_D. A *minimum dominating set* S_D^* is a dominating set of minimum cardinality and the *domination number* $\gamma(D)$ is defined as $\gamma(D) := |S_D^*|$ (see, e.g. [13]) where $|\cdot|$ denotes the set cardinality functional. See [7] and [20] for more on graphs and digraphs. If a minimum dominating set is of size one, we call it a *dominating vertex*.

Note that for $|\mathcal{V}| = n > 0$, $1 \leqslant \gamma(D) \leqslant n$, since $\mathcal{V}$ itself is always a dominating set.

In recent years, a new classification tool based on the relative allocation of points from various classes has been developed. Priebe et al. [16] introduced the *class cover catch digraphs* (CCCDs) and gave the exact and the asymptotic distribution of the domination number of the CCCD based on two sets, $\mathcal{X}_n$ and $\mathcal{Y}_m$, which are of size n and m, from classes, $\mathcal{X}$ and $\mathcal{Y}$, respectively, and are sets of iid random variables from uniform distribution on a compact interval in $\mathbb{R}$. Papers [9,10,14,17,18] applied the concept in higher dimensions and demonstrated relatively good performance of CCCD in classification. The methods employed involve *data reduction* (*condensing*) by using approximate minimum dominating sets as *prototype sets* (since finding the exact minimum dominating set is an NP-hard problem in general – e.g., for CCCD in multiple dimensions – (see [8]). DeVinney and Wierman [11] proved a SLLN result for the domination number of CCCDs for one-dimensional data. Although intuitively appealing and easy to extend to higher dimensions, exact and asymptotic distribution of the domination number of the CCCDs are not analytically tractable in $\mathbb{R}^2$ or higher dimensions. As alternatives to CCCD, two new families of PCDs are introduced in [2,4] and are applied in testing spatial point patterns (see [5,6]). These new families are both applicable to pattern classification also. They are designed to have better distributional and mathematical properties. For example, the distribution of the relative density (of arcs) is derived for one family in [5] and for the other family in [6]. In this article, we derive the asymptotic distribution of the domination number of the latter family called *r-factor proportional-edge PCD*. During the derivation process, we introduce auxiliary tools, such as, *proximity region* (which is the most crucial concept in defining the PCD), *Γ_1-region*, *superset region*, *closest edge extrema*, *asymptotically accurate distribution*, and so on. We utilize these special regions, extrema, and asymptotic expansion of the distribution of these extrema. The choice of the change of variables in the asymptotic expansion is also dependent on the type of the extrema used and crucial in finding the limits of the improper integrals we encounter. Our methodology is instructive in finding the distribution of the domination number of similar PCDs in $\mathbb{R}^2$ or higher dimensions.

In addition to the mathematical tractability and applicability to testing spatial patterns and classification, this new family of PCDs is more flexible as it allows choosing an optimal parameter for best performance in hypothesis testing or pattern classification.

The domination number of PCDs is first investigated for data in one Delaunay triangle (in $\mathbb{R}^2$) and the analysis is generalized to data in multiple Delaunay triangles. Some trivial proofs are omitted, shorter proofs are given in the main body of the article; while longer proofs are deferred to the Appendix.

2. Proximity maps and the associated PCDs

We construct the proximity regions using two data sets $\mathcal{X}_n$ and $\mathcal{Y}_m$ from two classes $\mathcal{X}$ and $\mathcal{Y}$, respectively. Given $\mathcal{Y}_m \subseteq \Omega$, the *proximity map* $N_{\mathcal{Y}}(\cdot) : \Omega \to 2^{\Omega}$ associates a *proximity region* $N_{\mathcal{Y}}(x) \subseteq \Omega$ with each point $x \in \Omega$. The region $N_{\mathcal{Y}}(x)$ is defined in terms of the distance between x and $\mathcal{Y}_m$. More specifically, our r-factor proximity maps will be based on the relative positions of points from $\mathcal{X}_n$ with respect to the Delaunay tessellation of $\mathcal{Y}_m$. In this article, a triangle refers to the closed region bounded by its edges. See Fig. 1 for an example with $n = 200$ $\mathcal{X}$ points iid $\mathcal{U}((0,1) \times (0,1))$, the uniform distribution on the unit square and the Delaunay triangulation is based on $m = 10$ $\mathcal{Y}$ which are points also iid $\mathcal{U}((0,1) \times (0,1))$.

If $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ is a set of Ω-valued random variables then $N_{\mathcal{Y}}(X_i)$ are random sets. If X_i are iid then so are the random sets $N_{\mathcal{Y}}(X_i)$. We define the data-random proximity catch digraph D – associated with $N_{\mathcal{Y}}(\cdot)$ – with vertex set $\mathcal{X}_n = \{X_1, \cdots, X_n\}$ and arc set $\mathcal{A}$ by

$$(X_i, X_j) \in \mathcal{A} \iff X_j \in N_{\mathcal{Y}}(X_i).$$

Since this relationship is not symmetric, a digraph is used rather than a graph. The random digraph D depends on the (joint) distribution of X_i and on the map $N_{\mathcal{Y}}(\cdot)$. For $\mathcal{X}_n = \{X_1, \cdots, X_n\}$, a set of iid random variables from

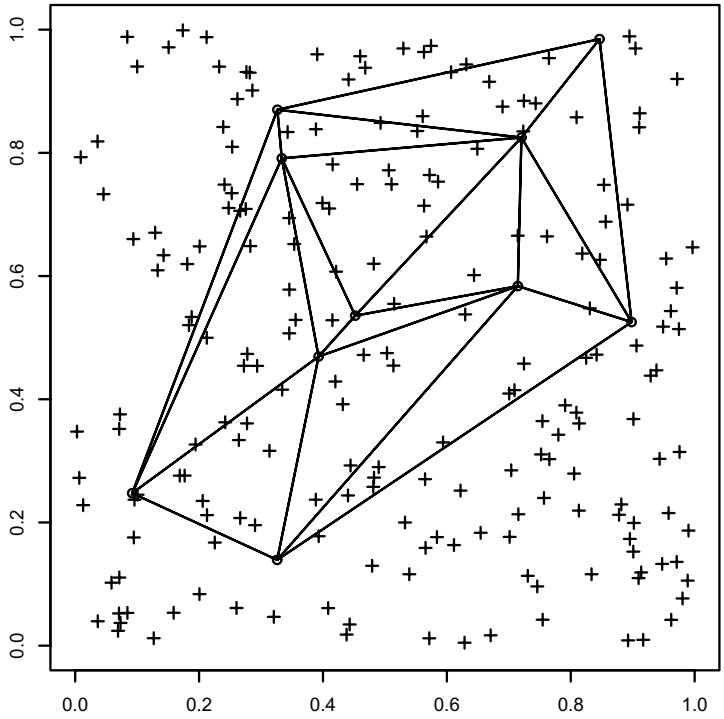

Fig. 1. A realization of 200 $\mathcal{X}$ points (crosses) and the Delaunay triangulation based on 10 $\mathcal{Y}$ points (circles).

F, the domination number of the associated data-random proximity catch digraph based on the proximity map $N(\cdot)$, denoted $\gamma(\mathcal{X}_n, N)$, is the minimum number of point(s) that dominate all points in $\mathcal{X}_n$.

The random variable $\gamma(\mathcal{X}_n, N)$ depends explicitly on $\mathcal{X}_n$ and $N(\cdot)$ and implicitly on F. Furthermore, in general, the distribution, hence the expectation $\mathbf{E}[\gamma(\mathcal{X}_n, N)]$, depends on n, F, and N; $1 \leqslant \mathbf{E}[\gamma(\mathcal{X}_n, N)] \leqslant n$. In general, the variance of $\gamma(\mathcal{X}_n, N)$ satisfies, $1 \leqslant \mathbf{Var}[\gamma(\mathcal{X}_n, N)] \leqslant n^2/4$.

For example, the CCCD of [16] can be viewed as an example of PCDs in $\mathbb{R}$ and is briefly discussed in the next section. We use many of the properties of the CCCD in $\mathbb{R}$ as guidelines in defining PCDs in higher dimensions.

2.1. Spherical proximity maps

Let $\mathcal{Y}_m = \{\mathsf{y}_1, \ldots, \mathsf{y}_m\} \subset \mathbb{R}$. Then the proximity map associated with CCCD is defined as the open ball $N_S(x) := B(x, r(x))$ for all $x \in \mathbb{R}$, where $r(x) := \min_{\mathsf{y} \in \mathcal{Y}_m} d(x, \mathsf{y})$ (see [16]) with $d(x, y)$ being the Euclidean distance between x and y. That is, there is an arc from X_i to X_j iff there exists an open ball centered at X_i which is "pure" (or contains no elements) of $\mathcal{Y}_m$ in its interior, and simultaneously contains (or "catches") point X_j. We consider the closed ball, $\overline{B}(x, r(x))$ for $N_S(x)$ in this article. Then for $x \in \mathcal{Y}_m$, we have $N_S(x) = \{x\}$. Notice that a ball is a sphere in higher dimensions, hence the notation N_S. Furthermore, dependence on $\mathcal{Y}_m$ is through $r(x)$. Note that in $\mathbb{R}$ this proximity map is based on the intervals $I_j = (\mathsf{y}_{(j-1):m}, \mathsf{y}_{j:m})$ for $j = 0, \ldots, m+1$ with $\mathsf{y}_{0:m} = -\infty$ and $\mathsf{y}_{(m+1):m} = \infty$, where $\mathsf{y}_{j:m}$ is the j^{th} order statistic in $\mathcal{Y}_m$. This interval partitioning can be viewed as the Delaunay tessellation of $\mathcal{Y}_m$ in $\mathbb{R}$. So *in higher dimensions, we use the Delaunay triangulation based on $\mathcal{Y}_m$ to partition the support.*

A natural extension of the proximity region $N_S(x)$ to $\mathbb{R}^d$ with $d > 1$ is obtained as $N_S(x) := B(x, r(x))$ where $r(x) := \min_{\mathsf{y} \in \mathcal{Y}_m} d(x, \mathsf{y})$ which is called the *spherical proximity map*. The spherical proximity map $N_S(x)$ is well-defined for all $x \in \mathbb{R}^d$ provided that $\mathcal{Y}_m \neq \emptyset$. Extensions to $\mathbb{R}^2$ and higher dimensions with the spherical proximity map – with applications in classification – are investigated by [9,10,14,17,18]. However, finding the minimum dominating set of CCCD (i.e., the PCD associated with $N_S(\cdot)$) is an NP-hard problem and the distribution of the domination number is not analytically tractable for $d > 1$. This drawback has motivated us to define new types

of proximity maps. Ceyhan and Priebe [4] introduced r-factor proportional-edge PCD, where the distribution of the domination number of r-factor PCD with $r = 3/2$ is used in testing spatial patterns of segregation or association. Ceyhan et al. [6] computed the asymptotic distribution of the relative density of the r-factor PCD and used it for the same purpose. Ceyhan and Priebe [2] introduced the central similarity proximity maps and the associated PCDs, and [5] computed the asymptotic distribution of the relative density of the parametrized version of the central similarity PCDs and applied the method to testing spatial patterns. An extensive treatment of the PCDs based on Delaunay tessellations is available in [1].

The following property (which is referred to as Property (1)) of CCCDs in $\mathbb{R}$ plays an important role in defining proximity maps in higher dimensions.

$$\textbf{Property}(\mathbf{1})\ \text{For } x \in I_j, N_S(x) \text{ is a proper subset of } I_j \text{ for almost all } x \in I_j. \tag{1}$$

In fact, Property (1) holds for all $x \in I_j \setminus \{(\mathsf{y}_{(j-1):m} + \mathsf{y}_{j:m})/2\}$ for CCCDs in $\mathbb{R}$. For $x \in I_j$, $N_S(x) = I_j$ iff $x = (\mathsf{y}_{(j-1):m} + \mathsf{y}_{j:m})/2$. We define an associated region for such points in the general context. The *superset region* for any proximity map $N(\cdot)$ in Ω is defined to be

$$\mathcal{R}_{S(N)} := \big\{x \in \Omega : N(x) = \Omega\big\}.$$

For example, for $\Omega = I_j \subsetneq \mathcal{R}$, $\mathcal{R}S(N_S) := \{x \in I_j : N_S(x) = I_j\} = \{(\mathsf{y}_{(j-1):m} + \mathsf{y}_{j:m})/2\}$ and for $\Omega = \mathcal{T}_j \subsetneq \mathbb{R}^d$, $\mathcal{R}S(N_S) := \{x \in \mathcal{T}_j : N_S(x) = \mathcal{T}_j\}$, where $\mathcal{T}_j$ is the j^{th} Delaunay cell in the Delaunay tessellation. Note that for $x \in I_j$, $\lambda(N_S(x)) \leqslant \lambda(I_j)$ and $\lambda(N_S(x)) = \lambda(I_j)$ iff $x \in \mathcal{R}_S(N_S)$ where $\lambda(\cdot)$ is the Lebesgue measure on $\mathbb{R}$. So the proximity region of a point in $\mathcal{R}_S(N_S)$ has the largest Lebesgue measure. Note also that given $\mathcal{Y}_m$, $\mathcal{R}_S(N_S)$ is not a random set, but $\mathbf{I}(X \in \mathcal{R}_S(N_S))$ is a random variable, where $\mathbf{I}(\cdot)$ stands for the indicator function. Property 1 also implies that $\mathcal{R}_S(N_S)$ has zero $\mathbb{R}$-Lebesgue measure.

Furthermore, given a set B of size n in $[\mathsf{y}_{1:m}, \mathsf{y}_{m:m}] \setminus \mathcal{Y}_m$, the number of disconnected components in the PCD based on $N_S(\cdot)$ is at least the cardinality of the set $\{j \in \{1, 2, \ldots, m\} : B \cap I_j \neq \emptyset\}$, which is the set of indices of the intervals that contain some point(s) from B.

Since the distribution of the domination number of spherical PCD (or CCCD) is tractable in $\mathbb{R}$, but not in $\mathbb{R}^d$ with $d > 1$, we try to mimic its properties in $\mathbb{R}$ while defining new PCDs in higher dimensions.

3. The r-factor proportional-edge proximity maps

First, we describe the construction of the r-factor proximity maps and regions, then state some of its basic properties and introduce some auxiliary tools.

3.1. Construction of the proximity map

Let $\mathcal{Y}_m = \{\mathsf{y}_1, \ldots, \mathsf{y}_m\}$ be m points in general position in $\mathbb{R}^d$ and $\mathcal{T}_j$ be the j^{th} Delaunay cell for $j = 1, \ldots, J_m$, where J_m is the number of Delaunay cells. Let $\mathcal{X}_n$ be a set of iid random variables from distribution F in $\mathbb{R}^d$ with support $\mathcal{S}(F) \subseteq \mathcal{C}_H(\mathcal{Y}_m)$.

In particular, for illustrative purposes, we focus on $\mathbb{R}^2$ where a Delaunay tessellation is a triangulation, provided that no more than three points in $\mathcal{Y}_m$ are cocircular (i.e., lie in the same circle). Furthermore, for simplicity, let $\mathcal{Y}_3 = \{\mathsf{y}_1, \mathsf{y}_2, \mathsf{y}_3\}$ be three non-collinear points in $\mathbb{R}^2$ and $T(\mathcal{Y}_3) = T(\mathsf{y}_1, \mathsf{y}_2, \mathsf{y}_3)$ be the triangle with vertices $\mathcal{Y}_3$. Let $\mathcal{X}_n$ be a set of iid random variables from F with support $\mathcal{S}(F) \subseteq T(\mathcal{Y}_3)$. If $F = \mathcal{U}(T(\mathcal{Y}_3))$, a composition of translation, rotation, reflections, and scaling will take any given triangle $T\mathcal{Y}_3$ to the basic triangle $T_b = T((0,0),(1,0),(c_1,c_2))$ with $0 < c_1 \leqslant 1/2$, $c_2 > 0$, and $(1-c_1)^2 + c_2^2 \leqslant 1$, preserving uniformity. That is, if $X \sim \mathcal{U}T(\mathcal{Y}_3)$ is transformed in the same manner to, say X', then we have $X' \sim \mathcal{U}(T_b)$.

For $r \in [1, \infty]$, define $N_{PE}^r(\cdot, M) := N(\cdot, M; r, \mathcal{Y}_3)$ to be the *r-factor proportional-edge proximity map* with M-vertex regions as follows (see also Fig. 2 with $M = M_C$ and $r = 2$). For $x \in T(\mathcal{Y}_3) \setminus \mathcal{Y}_3$, let $v(x) \in \mathcal{Y}_3$ be the vertex whose region contains x; i.e., $x \in R_M(v(x))$. In this article *M-vertex regions* are constructed by the lines joining any point $M \in \mathcal{R}^2 \setminus \mathcal{Y}_3$ to a point on each of the edges of $T(\mathcal{Y}_3)$. Preferably, M is selected to be in the interior of the triangle $T(\mathcal{Y}_3)^\circ$. For such an M, the corresponding vertex regions can be defined using the line

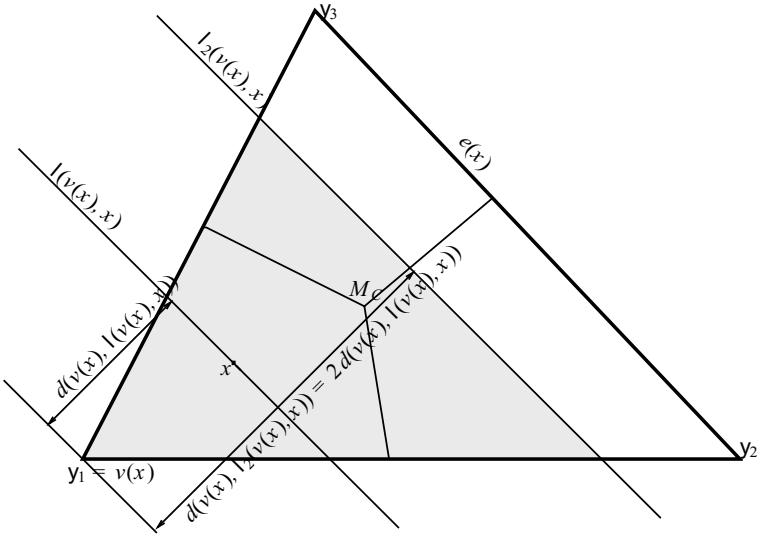

Fig. 2. Construction of r-factor proximity region, $N_{PE}^{r=2}(x)$ (shaded region).

segment joining M to e_j, which lies on the line joining y_j to M; e.g. see Fig. 3 (left) for vertex regions based on center of mass M_C, and (right) incenter M_I. With M_C, the lines joining M and $\mathcal{Y}_3$ are the *median lines*, that cross edges at M_j for $j = 1, 2, 3$. M-vertex regions, among many possibilities, can also be defined by the orthogonal projections from M to the edges. See [1] for a more general definition. The vertex regions in Fig. 3 are center of mass vertex regions or CM-vertex regions. If x falls on the boundary of two M-vertex regions, we assign $v(x)$ arbitrarily. Let $e(x)$ be the edge of $T(\mathcal{Y}_3)$ opposite of $v(x)$. Let $\ell(v(x), x)$ be the line parallel to $e(x)$ through x. Let $d(v(x), \ell(v(x), x))$ be the Euclidean (perpendicular) distance from $v(x)$ to $\ell(v(x), x)$. For $r \in [1, \infty)$, let $\ell_r(v(x), x)$ be the line parallel to $e(x)$ such that

$$d(v(x), \ell_r(v(x), x)) = r\, d(v(x), \ell(v(x), x)) \text{and } d(\ell(v(x), x), \ell_r(v(x), x)) < d(v(x), \ell_r(v(x), x)).$$

Let $T_r(x)$ be the triangle similar to and with the same orientation as $T(\mathcal{Y}_3)$ having $v(x)$ as a vertex and $\ell_r(v(x), x)$ as the opposite edge. Then the *r-factor proportional-edge proximity region* $N_{PE}^{r}(x, M)$ is defined to be $T_r(x) \cap T(\mathcal{Y}_3)$. Notice that $\ell(v(x), x)$ divides the edges of $T_r(x)$ (other than the one lies on $\ell_r(v(x), x)$) proportionally with the factor r. Hence the name *r-factor proportional edge proximity region*.

Notice that $r \geqslant 1$ implies $x \in N_{PE}^{r}(x, M)$ for all $x \in T(\mathcal{Y}_3)$. Furthermore, $\lim_{r\to\infty} N_{PE}^{r}(x, M) = T(\mathcal{Y}_3)$ for all $x \in T(\mathcal{Y}_3) \setminus \mathcal{Y}_3$, so we define $N_{PE}^{\infty}(x, M) = T(\mathcal{Y}_3)$ for all such x. For $x \in (\mathcal{Y}_3)$, we define $N_{PE}^{r}(x, M) = \{x\}$ for all $r \in [1, \infty]$.

Hence, r-factor proportional edge PCD has vertices $\mathcal{X}_n$ and arcs (x_i, x_j) iff $x_j \in N_{PE}^{r}(x_i, M)$. See Fig. 4 for a realization of $\mathcal{X}_n$ with $n = 7$ and $m = 3$. The number of arcs is 12 and $\gamma_n(r = 2, M_C) = 1$. By construction, note that as x gets closer to M (or equivalently further away from the vertices in vertex regions), $N_{PE}^{r}(x, M)$ increases in area, hence it is more likely for the outdegree of x to increase. So if more $\mathcal{X}$ points are around the center M, then it is more likely for γ_n to decrease. On the other hand, if more $\mathcal{X}$ points are around the vertices $\mathcal{Y}_3$, then the regions get smaller, hence it is more likely for the outdegree for such points to be smaller, thereby implying γ_n to increase. This probabilistic behaviour is utilized in [4] for testing spatial patterns.

Note also that, $N_{PE}^{r}(x, M)$ is a *homothetic transformation* (*enlargement*) with $r \geqslant 1$ applied on the region $N_{PE}^{r=1}(x, M)$. Furthermore, this transformation is also an *affine similarity transformation*.

3.2. Some basic properties and auxiliary concepts

First, notice that $X_i \overset{iid}{\sim} F$, with the additional assumption that the non-degenerate two-dimensional probability density function f exists with support $\mathcal{S}(F) \subseteq T(\mathcal{Y}_3)$, imply that the special case in the construction of N_{PE}^{r}–X

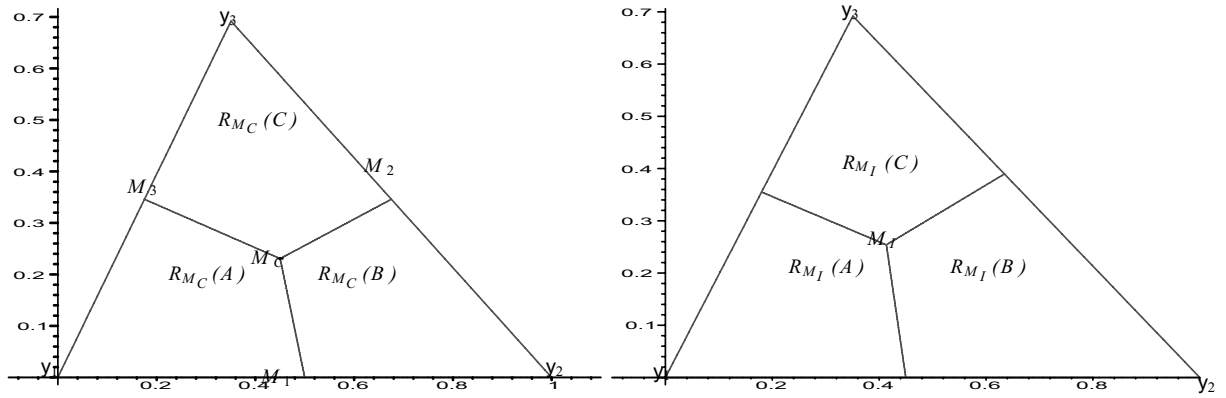

Fig. 3. The vertex regions constructed with center of mass M_C (left) and incenter M_I (right) using the line segments on the line joining M to $\mathcal{Y}_3$.

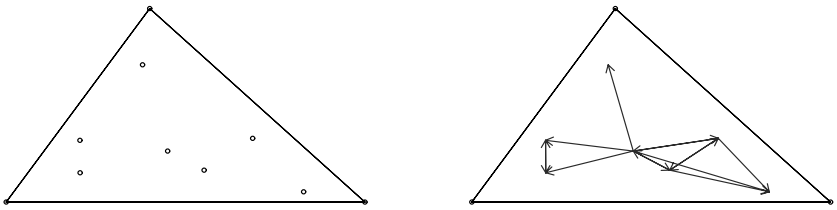

Fig. 4. A realization of 7 $\mathcal{X}$ points generated iid $\mathcal{U}T(\mathcal{Y}_3)$ (left) and the corresponding arcs of r-factor proportional edge PCD with $r = 3/2$ and $M = M_C$.

falls on the boundary of two vertex regions – occurs with probability zero. Note that for such an F, $N_{PE}^{r}(X)$ is a triangle a.s.

The similarity ratio of $N_{PE}^{r}(x, M)$ to $T(\mathcal{Y}_3)$ is given by $\frac{\min(d(v(x),e(x)),r\,d(v(x),\ell(v(x),x)))}{d(v(x),e(x))}$, that is, $N_{PE}^{r}(x, M)$ is similar to $T(\mathcal{Y}_3)$ with the above ratio. Property (1) holds depending on the pair M and r. That is, there exists an r_0 and a corresponding point $M(r_0) \in T(\mathcal{Y}_3)^\circ$ so that $N_{PE}^{r_0}(x, M)$ satisfies Property (1) for all $r \leqslant r_0$, but fails to satisfy it otherwise. Property (1) fails for all M when $r = \infty$. With CM-vertex regions, for all $r \in [1, \infty]$, the area $A\left(N_{PE}^{r}(x, M_C)\right)$ is a continuous function of $d(\ell_r(v(x), x), v(x))$ which is a continuous function of $d(\ell(v(x), x), v(x))$ which is a continuous function of x.

Note that if x is close enough to M, it is also possible to have $N_{PE}^{r}(x, M) = T(\mathcal{Y}_3)$ for $r = \sqrt{2}$.

In $T(\mathcal{Y}_3)$, drawing the lines $q_j(r, x)$ such that $d(\mathsf{y}_j, e_j) = r\,d(\mathsf{y}_j, q_j(r, x))$ for $j \in \{1, 2, 3\}$ yields a triangle, denoted $\mathcal{T}_r$, for $r < 3/2$. See Fig. 5 for $\mathcal{T}_r$ with $r = \sqrt{2}$.

The functional form of $\mathcal{T}_r$ in the basic triangle T_b is given by

$$\begin{aligned}\mathcal{T}_r = T(t_1(r), t_2(r), t_3(r)) &= \left\{(x, y) \in T_b : y \geqslant \frac{c_2\,(r-1)}{r};\ y \leqslant \frac{c_2\,(1 - r\,x)}{r\,(1-c_1)};\ y \leqslant \frac{c_2\,(r\,(x-1)+1)}{r\,c_1}\right\} \\ &= T\left(\left(\frac{(r-1)\,(1+c_1)}{r}, \frac{c_2\,(r-1)}{r}\right), \left(\frac{2-r+c_1\,(r-1)}{r}, \frac{c_2\,(r-1)}{r}\right),\right. \\ &\quad \left.\left(\frac{c_1\,(2-r)+r-1}{r}, \frac{c_2\,(r-2)}{r}\right)\right) \qquad (2)\end{aligned}$$

There is a crucial difference between the triangles $\mathcal{T}_r$ and $T(M_1, M_2, M_3)$. More specifically $T(M_1, M_2, M_3) \subseteq \mathcal{R}_S(r, M)$ for all M and $r \geqslant 2$, but $(\mathcal{T}_r)^\circ$ and $\mathcal{R}_S(r, M)$ are disjoint for all M and r. So if $M \in (\mathcal{T}_r)^\circ$, then $\mathcal{R}_S(r, M) = \emptyset$; if $M \in \partial(\mathcal{T}_r)$, then $\mathcal{R}_S(r, M) = \{M\}$; and if $M \notin \mathcal{T}_r$, then $\mathcal{R}_S(r, M)$ has positive area. Thus $N_{PE}^{r}(\cdot, M)$ fails to satisfy Property (1) if $M \notin \mathcal{T}_r$. See Fig. 6 for two examples of superset regions with M that

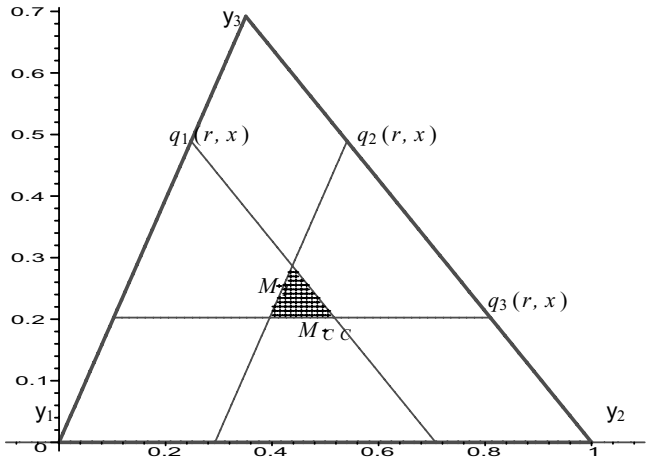

Fig. 5. The triangle $\mathcal{T}_r$ with $r = \sqrt{2}$ (the hatched region).

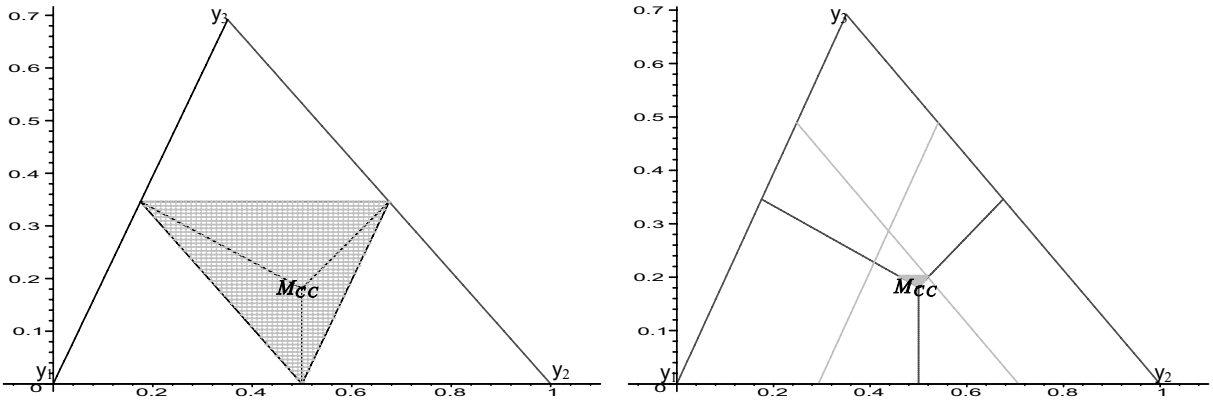

Fig. 6. The superset regions (the shaded regions) constructed with circumcenter M_{CC} with $r = \sqrt{2}$ (left) and $r = 2$ (right) with vertex regions constructed with orthogonal projections to the edges.

corresponds to circumcenter M_{CC} in this triangle and the vertex regions are constructed using orthogonal projections. For $r = 2$, note that $\mathcal{T}_r = \emptyset$ and the superset region is $T(M_1, M_2, M_3)$ (see Fig. 6 (left)), while for $r = \sqrt{2}$, $\mathcal{T}_r^\circ$ and $\mathcal{R}_S(r = \sqrt{2}, M)^\circ$ are disjoint (see Fig. 6 (right)).

The triangle $\mathcal{T}_r$ given in Eq. (2) and the superset region $\mathcal{R}_S(r, M)$ play a crucial role in computing the distribution of the domination number of the r-factor PCD.

3.3. Main result

Next, we present the main result of this article. Let $\gamma_n(r, M) := \gamma(\mathcal{X}_n, N_{PE}^r, M)$ be the domination number of the PCD based on N_{PE}^r with $\mathcal{X}_n$, a set of iid random variables from $\mathcal{U}(T(\mathcal{Y}_3))$, with M-vertex regions.

The domination number $\gamma_n(r, M)$ of the PCD has the following asymptotic distribution. As $n \to \infty$,

$$\gamma_n(r, M) \sim \begin{cases} 2 + \mathrm{BER}(1 - p_r), & \text{for } r \in [1, 3/2] \text{ and } M \in \{t_1(r), t_2(r), t_3(r)\}, \\ 1, & \text{for } r > 3/2, \\ 3, & \text{for } r \in [1, 3/2) \text{ and } M \in \mathcal{T}_r \setminus \{t_1(r), t_2(r), t_3(r)\}, \end{cases} \tag{3}$$

where $\mathrm{BER}(p)$ stands for Bernoulli distribution with probability of success p, $\mathcal{T}_r$ and $t_j(r)$ are defined in Eq. (2), and for $r \in [1, 3/2)$ and $M \in \{t_1(r), t_2(r), t_3(r)\}$,

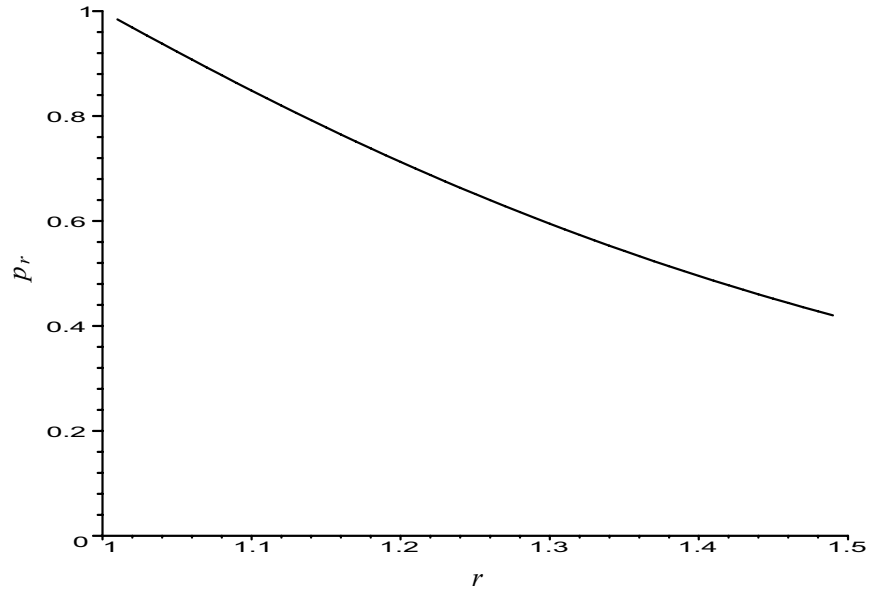

Fig. 7. Plotted is the probability $p_r = \lim_{n\to\infty} P(\gamma_n(r, M) = 2)$ given in Eq. (4) as a function of r for $r \in [1, 3/2)$ and $M \in \{t_1(r), t_2(r), t_3(r)\}$.

$$p_r = \int_0^\infty \int_0^\infty \frac{64\, r^2}{9\,(r-1)^2}\, w_1\, w_3 \exp\left(\frac{4\, r}{3\,(r-1)}\ (w_1^2 + w_3^2 + 2\, r\,(r-1)\, w_1\, w_3)\right) dw_3 w_1; \tag{4}$$

for example for $r = 3/2$ and $M = M_C$, $p_r \approx 0.7413$. See Fig. 7 for the plot of p_r for $r \in [1, 3/2)$.

In Eq. (3), the first line is referred as the non-degenerate case, the second and third lines are referred as degenerate cases with a.s. limits 1 and 3, respectively.

In the following sections, we define a region associated with $\gamma = 1$ case in general. Then we give finite sample and asymptotic upper bounds for $\gamma_n(r, M)$. Then we derive the asymptotic distribution of $\gamma_n(r, M)$.

4. The Γ_1-regions for N_{PE}^r

First, we define Γ_1-regions in general, and describe the construction of Γ_1-region of N_{PE}^r for one point and multiple point data sets, and provide some results concerning Γ_1-regions.

4.1. Definition of Γ_1-regions

Let $(\Omega, \mathcal{M})$ be a measurable space and consider the proximity map $N : \Omega \to 2^\Omega$. For any set $B \subseteq \Omega$, the Γ_1-*region* of B associated with $N(\cdot)$, is defined to be the region $\Gamma_1^N(B) := \{z \in \Omega : B \subseteq N(z)\}$. For $x \in \Omega$, we denote $\Gamma_1^N(\{x\})$ as $\Gamma_1^N(x)$.

If $\mathcal{X}_n = \{X_1, X_2, \cdots, X_n\}$ is a set of Ω-valued random variables, then $\Gamma_1^N(X_i)$, $i = 1, \cdots, n$, and $\Gamma_1^N(\mathcal{X}_n)$ are random sets. If the X_i are iid, then so are the random sets $\Gamma_1^N(X_i)$.

Note that $\gamma(\mathcal{X}_n, N) = 1$ iff $\mathcal{X}_n \cap \Gamma_1^N(\mathcal{X}_n) \neq \emptyset$. Hence the name Γ_1-*region*.

It is trivial to see the following.

Proposition 1. For any proximity map N and set $B \subseteq \Omega$, $\mathcal{R}_S(N) \subseteq \Gamma_1^N(B)$.

Lemma 1. For any proximity map N and $B \subseteq \Omega$, $\Gamma_1^N(B) = \cap_{x\in B}\Gamma_1^N(x)$.

Proof: Given a particular type of proximity map N and subset $B \subseteq \Omega$, $y \in \Gamma_1^N(B)$ iff $B \subseteq N(y)$ iff $x \in N(y)$ for all $x \in B$ iff $y \in \Gamma_1^N(x)$ for all $x \in B$ iff $y \in \cap_{x\in B}\Gamma_1^N(x)$. Hence the result follows. □

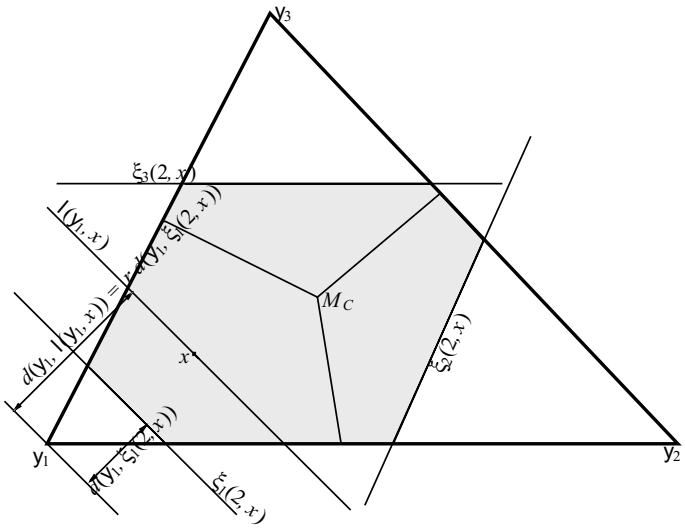

Fig. 8. Construction of the Γ_1-region, $\Gamma_1^{r=2}(x, M_C)$ (shaded region).

A problem of interest is finding, if possible, a (proper) subset of B, say $G \subsetneq B$, such that $\Gamma_1^N(B) = \cap_{x\in G}\Gamma_1^N(x)$. This implies that only the points in G will be *active* in determining $\Gamma_1^N(B)$. For example, in $\mathbb{R}$ with $\mathcal{Y}_2 = \{0, 1\}$, and $\mathcal{X}_n$ a set of iid random variables of size $n > 1$ from F in $(0, 1)$, $\Gamma_1^{N_S}(\mathcal{X}_n) = (X_{n:n}/2, (1 + X_{1:n})/2)$. So the extrema (minimum and maximum) of the set $\mathcal{X}_n$ are sufficient to determine the Γ_1-region; i.e., $G = \{X_{1:n}, X_{n:n}\}$ for $\mathcal{X}_n$ a set of iid random variables from a continuous distribution on $(0, 1)$. Unfortunately, in the multi-dimensional case, there is no natural ordering that yields natural extrema such as minimum or maximum.

4.2. Construction of Γ_1-region of a point for N_{PE}^r

For $N_{PE}^r(\cdot, M)$, the Γ_1-region, denoted as $\Gamma_1^r(\cdot, M) := \Gamma_1^{N_{PE}^r}(\cdot, M)$, is constructed as follows; see also Fig. 8. Let $\xi_j(r, x)$ be the line parallel to e_j such that $\xi_j(r, x) \cap T(\mathcal{Y}_3) \neq \emptyset$ and $r\, d(\mathsf{y}_j, \xi_j(r, x)) = d(\mathsf{y}_j, \ell(\mathsf{y}_j, x))$ for $j \in \{1, 2, 3\}$. Then

$$\Gamma_1^r(x, M) = \cup_{j=1}^3 \left[\Gamma_1^r(x, M) \cap R_M(\mathsf{y}_j)\right]$$

where $\Gamma_1^r(x, M) \cap R_M(\mathsf{y}_j) = \{z \in R_M(\mathsf{y}_j) : d(\mathsf{y}_j, \ell(\mathsf{y}_j, z)) \geqslant d(\mathsf{y}_j, \xi_j(r, x)\}$ for $j \in \{1, 2, 3\}$.

Notice that $r \geqslant 1$ implies that $x \in \Gamma_1^r(x, M)$. Furthermore, $\lim_{r\to\infty} \Gamma_1^r(x, M) = T(\mathcal{Y}_3)$ for all $x \in T(\mathcal{Y}_3) \setminus \mathcal{Y}_3$ and so we define $\Gamma_1^{r=\infty}(x, M) = T(\mathcal{Y}_3)$ for all such x. For $x \in \mathcal{Y}_3$, $\Gamma_1^r(x, M) = \{x\}$ for all $r \in [1, \infty]$.

Notice that $\Gamma_1^r(x, M_C)$ is a convex hexagon for all $r \geqslant 2$ and $x \in T(\mathcal{Y}_3) \setminus \mathcal{Y}_3$, (since for such an x, $\Gamma_1^r(x, M_C)$ is bounded by $\xi_j(r, x)$ and e_j for all $j \in \{1, 2, 3\}$, see also Fig. 8) else it is either a *convex hexagon* or a *non-convex* but *star-shaped polygon* depending on the location of x and the value of r.

4.3. The Γ_1-region of a multiple point data set for N_{PE}^r

So far, we have described the Γ_1-region for a point in $x \in T(\mathcal{Y}_3)$. For a set $\mathcal{X}_n$ of size n in $T(\mathcal{Y}_3)$, the region $\Gamma_1^r(\mathcal{X}_n, M)$ can be specified by the edge extrema only. The (closest) *edge extrema* of a set B in $T(\mathcal{Y}_3)$ are the points closest to the edges of $T(\mathcal{Y}_3)$, denoted x_{e_j} for $j \in \{1, 2, 3\}$; that is, $x_{e_j} \in \operatorname{arginf}_{x\in B} d(x, e_j)$. Note that if $B = \mathcal{X}_n$ is a set of iid random variables of size n from F then the edge extrema, denoted $X_{e_j}(n)$, are random variables. Below, we show that the edge extrema are the active points in defining $\Gamma_1^r(\mathcal{X}_n, M)$.

Proposition 2. Let B be any set of n distinct points in $T(\mathcal{Y}_3)$. For r-factor proportional-edge proximity maps with M-vertex regions, $\Gamma_1^r(B, M) = \cap_{k=1}^3 \Gamma_1^r(x_{e_k}, M)$.

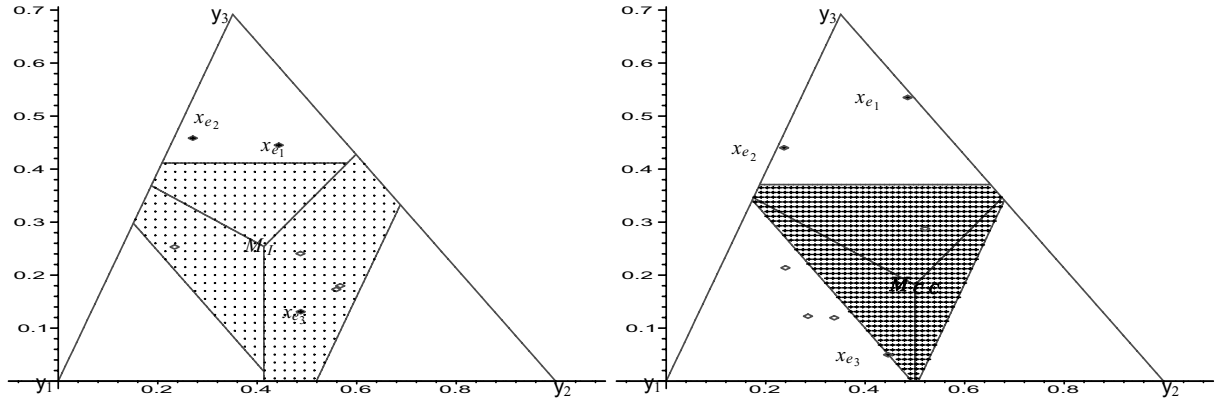

Fig. 9. The Γ_1-regions (the hatched regions) for $r = 2$ with seven $\mathcal{X}$ points iid $\mathcal{U}(T(\mathcal{Y}_3))$ where vertex regions constructed with incenter M_I (left) and circumcenter M_{CC} (right) with orthogonal projection.

Proof: Given $B = \{x_1, \ldots, x_n\}$ in $T(\mathcal{Y}_3)$. Note that

$$\Gamma_1^r(B, M) \cap R_M(\mathsf{y}_j) = \left[\cap_{i=1}^n \Gamma_1^r(x_i, M)\right] \cap R_M(\mathsf{y}_j),$$

but by definition $x_{e_j} \in \operatorname{argmax}_{x \in B} d(\mathsf{y}_j, \xi_j(r, x))$, so

$$\Gamma_1^r(B, M) \cap R_M(\mathsf{y}_j) = \Gamma_1^r(x_{e_j}, M) \cap R_M(\mathsf{y}_j) \text{for } j \in \{1, 2, 3\}. \tag{5}$$

Furthermore, $\Gamma_1^r(B, M) = \cup_{j=1}^3 \left[\Gamma_1^r(x_{e_j}, M) \cap R_M(\mathsf{y}_j)\right]$, and

$$\Gamma_1^r(x_{e_j}, M) \cap R_M(\mathsf{y}_j) = \cap_{k=1}^3 \left[\Gamma_1^r(x_{e_k}, M) \cap\ R_M(\mathsf{y}_j)\right] \text{ for } j \in \{1, 2, 3\}. \tag{6}$$

Combining these two results in Eqs (5) and (6), we obtain $\Gamma_1^r(B, M) = \cap_{k=1}^3 \Gamma_1^r(x_{e_k}, M)$. □

From the above proposition, we see that the Γ_1-region for B as in proposition can also be written as the union of three regions of the form

$$\Gamma_1^r(B, M) \cap R_M(\mathsf{y}_j) = \{z \in R_M(\mathsf{y}_j) : d(\mathsf{y}_j, \ell(\mathsf{y}_j, z)) \geqslant d(\mathsf{y}_j, \xi_j(r, x_{e_j}))\} \text{ for } j \in \{1, 2, 3\}.$$

See Fig. 9 for Γ_1-region for $r = 2$ with seven $\mathcal{X}$ points iid $\mathcal{U}(T(\mathcal{Y}_3))$. In the left figure, vertex regions are based on incenter, while in the right figure, on circumcenter with orthogonal projections to the edges. In either case $\mathcal{X}_n \cap \Gamma_1^{r=2}(\mathcal{X}_n, M)$ is nonempty, hence $\gamma_n(2, M) = 1$.

Below, we demonstrate that edge extrema are distinct with probability 1 as $n \to \infty$. Hence in the limit three distinct points suffice to determine the Γ_1-region.

Theorem 1. Let $\mathcal{X}_n$ be a set of iid random variables from $\mathcal{U}(T(\mathcal{Y}_3))$ and let $E_{c,3}(n)$ be the event that (closest) edge extrema are distinct. Then $P(E_{c,3}(n)) \to 1$ as $n \to \infty$.

We can also define the regions associated with $\gamma(\mathcal{X}_n, N) = k$ for $k \leqslant n$ called Γ_k-*region* for proximity map $N_{\mathcal{Y}_3}(\cdot)$ and set $B \subseteq \Omega$ for $k = 1, \ldots, n$ (see [1]).

5. The asymptotic distribution of $\gamma_n(r, M)$

In this section, we first present a finite sample upper bound for $\gamma_n(r, M)$, then present the degenerate cases, and the nondegenerate case of the asymptotic distribution of $\gamma_n(r, M)$ given in Eq. (3).

5.1. An upper bound for $\gamma_n(r, M)$

Recall that by definition, $\gamma(\mathcal{X}_n, N) \leqslant n$. We will seek an a.s. least upper bound for $\gamma(\mathcal{X}_n, N)$. Let $\mathcal{X}_n$ be a set of iid random variables from F on $T(\mathcal{Y}_3)$ and let $\gamma(\mathcal{X}_n, N)$ be the domination number for the PCD based on a proximity map N. Denote the general a.s. least upper bound for $\gamma(\mathcal{X}_n, N)$ that works for all $n \geqslant 1$ and is independent of n (which is called *κ-value* in [1]) as $\kappa(N) := \min\{k : \gamma(\mathcal{X}_n, N) \leqslant k \text{ a.s. for all } n \geqslant 1\}$.

In $\mathbb{R}$ with $\mathcal{Y}_2 = \{0, 1\}$, for $\mathcal{X}_n$ a set of iid random variables from $\mathcal{U}(0, 1)$, $\gamma(\mathcal{X}_n, N_S) \leqslant 2$ with equality holding with positive probability. Hence $\kappa(N_S) = 2$.

Theorem 2. Let $\mathcal{X}_n$ be a set of iid random variables from $\mathcal{U}(T\mathcal{Y}_3))$ and $M \in \mathbb{R}^2 \setminus \mathcal{Y}_3$. Then $\kappa\left(N^r_{PE}\right) = 3$ for $N^r_{PE}(\cdot, M)$.

Proof: For $N^r_{PE}(\cdot, M)$, pick the point closest to edge e_j in vertex region $R_M(\mathsf{y}_j)$; that is, pick $U_j \in \operatorname{argmin}_{X \in \mathcal{X}_n \cap R_M(\mathsf{y}_j)} d(X, e_j) = \operatorname{argmax}_{X \in \mathcal{X}_n \cap R_M(\mathsf{y}_j)} d(\ell(\mathsf{y}, X), \mathsf{y}_j)$ in the vertex region for which $X_n \cap R_M(\mathsf{y}_j) \neq \emptyset$ for $j \in \{1, 2, 3\}$ (note that as $n \to \infty$, U_j is unique a.s. for each j, since X is from $\mathcal{U}(T(\mathcal{Y}_3))$. Then $\mathcal{X}_n \cap R_M(\mathsf{y}_j) \subset N^r_{PE}(U_j, M)$. Hence $\mathcal{X}_n \subset \cup_{j=1}^3 N^r_{PE}(U_j, M)$. So $\gamma_n(r, M_C) \leqslant 3$ with equality holding with positive probability. Thus $\kappa\left(N^r_{PE}\right) = 3$. □

Below is a general result for the limiting distribution of $\gamma(\mathcal{X}_n, N)$ for $\mathcal{X}_n$ from a very broad family of distributions and for general $N(\cdot)$.

Lemma 2. Let $\mathcal{R}_S(N)$ be the superset region for the proximity map $N(\cdot)$ and $\mathcal{X}_n$ be a set of iid random variables from F with $P_F(X \in \mathcal{R}_S(N)) > 0$. Then $\lim_{n\to\infty} P_F(\gamma(\mathcal{X}_n, N) = 1) = 1$.

Proof: Suppose $P_F(X \in \mathcal{R}_S(N)) > 0$. Recall that for any $x \in \mathcal{R}_S(N)$, we have $N(x) = \Omega$, so $\mathcal{X}_n \subseteq N(x)$, hence if $\mathcal{X}_n \cap \mathcal{R}_S(N) \neq \emptyset$ then $\gamma(\mathcal{X}_n, N) = 1$. Then $P(\mathcal{X}_n \cap \mathcal{R}_S(N) \neq \emptyset) \leqslant P(\gamma(\mathcal{X}_n, N) = 1)$. But $P\big(\mathcal{X}_n \cap \mathcal{R}_S(N) \neq \emptyset\big) = 1 - P\big(\mathcal{X}_n \cap \mathcal{R}_S(N) = \emptyset\big) = 1 - \big[1 - P_F\big(X \in \mathcal{R}_S(N)\big)\big]^n \to 1$ as $n \to \infty$, since $P_F\big(X \in \mathcal{R}_S(N)\big) > 0$. Hence $\lim_{n\to\infty} P(\gamma(\mathcal{X}_n, N) = 1) = 1$. □

Remark 1. In particular, for $F = \mathcal{U}(T(\mathcal{Y}_3))$, the inequality $P_F(X \in \mathcal{R}_S(N)) > 0$ holds iff $A(\mathcal{R}_S(N)) > 0$, then $P(\mathcal{X}_n \cap \mathcal{R}_S(N) \neq \emptyset) \to 1$. □

For $\mathcal{Y}_2 = \{0, 1\} \subset \mathcal{R}$, $\mathcal{R}_S(N_S) = \{1/2\}$, so Lemma 2 does not apply to N_S in $\mathbb{R}$.

Recall that $\kappa\left(N^r_{PE}\right) = 3$, then

$$1 \leqslant \mathbf{E}\left[\gamma_n(r, M)\right] \leqslant 3 \text{ and } 0 \leqslant \mathbf{V}ar\left[\gamma_n(r, M)\right] \leqslant 9/4.$$

Furthermore, there is a stochastic ordering for $\gamma_n(r, M)$.

Theorem 3. Suppose $\mathcal{X}_n$ is a set of iid random variables from a continuous distribution F on $T(\mathcal{Y}_3)$. Then for $r_1 < r_2$, we have $\gamma_n(r_2, M) \leq^{ST} \gamma\left(\mathcal{X}_n, N^{r_1}_{PE}, M\right)$.

Proof: Suppose $r_1 < r_2$. Then $P\left(\gamma_n(r_2, M) \leqslant 1\right) > P\left(\gamma_n(r_1, M) \leqslant 1\right)$ since $\Gamma^{r_1}_1(\mathcal{X}_n, M) \subsetneq \Gamma^{r_2}_1(\mathcal{X}_n, M)$ for any realization of $\mathcal{X}_n$ and by a similar argument $P\big(\gamma_n(r_2, M) \leqslant 2\big) > P\left(\gamma_n(r_1, M) \leqslant 2\right)$ so $P\left(\gamma_n(r_2, M) \leqslant 3\right) = P\left(\gamma_n(r_1, M) \leqslant 3\right)$. Hence the desired result follows. □

5.2. Geometry invariance

We present a "geometry invariance" result for $N^r_{PE}(\cdot, M)$ where M-vertex regions are constructed using the lines joining $\mathcal{Y}_3$ to M, rather than the orthogonal projections from M to the edges. This invariance property will simplify the notation in our subsequent analysis by allowing us to consider the special case of the equilateral triangle.

Theorem 4. (Geometry Invariance Property) Suppose $\mathcal{X}_n$ is a set of iid random variables from $\mathcal{U}(T(\mathcal{Y}_3))$. Then for any $r \in [1, \infty]$ the distribution of $\gamma_n(r, M)$ is independent of $\mathcal{Y}_3$ and hence the geometry of $T(\mathcal{Y}_3)$.

Proof: Suppose $X \sim \mathcal{U}(T(\mathcal{Y}_3))$. A composition of translation, rotation, reflections, and scaling will take any given triangle $T(\mathcal{Y}_3) = T(\mathsf{y}_1, \mathsf{y}_2, \mathsf{y}_3)$ to the basic triangle $T_b = T((0, 0), (1, 0), (c_1, c_2))$ with $0 < c_1 \leqslant 1/2$, $c_2 > 0$, and $(1 - c_1)^2 + c_2^2 \leqslant 1$. Furthermore, when X is also transformed in the same manner, say to X', then X'

Table 1
The number of $\gamma_n(r, M) = k$ out of $N = 1000$ Monte Carlo replicates with $M = M_C$ and $r = 2$ (left) and $r = 5/4$ (right)

$k\backslash n$	10	20	30	50	100
1	961	1000	1000	1000	1000
2	34	0	0	0	0
3	5	0	0	0	0

$k\backslash n$	10	20	30	50	100
1	9	0	0	0	0
2	293	110	30	8	0
3	698	890	970	992	1000

is uniform on T_b, i.e., $X' \sim \mathcal{U}(T_b)$. The transformation $\phi_e : \mathbb{R}^2 \to \mathbb{R}^2$ given by $\phi_e(u, v) = \left(u + \frac{1-2\,c_1}{\sqrt{3}}\,v, \frac{\sqrt{3}}{2\,c_2}\,v\right)$ takes T_b to the equilateral triangle $T_e = \big((0,0),(1,0),(1/2,\sqrt{3}/2)\big)$. Investigation of the Jacobian shows that ϕ_e also preserves uniformity. That is, $\phi_e(X') \sim \mathcal{U}(T_e)$. Furthermore, the composition of ϕ_e, with the scaling and rigid body transformations, maps the boundary of the original triangle, T_o, to the boundary of the equilateral triangle, T_e, the lines joining M to y_j in T_b to the lines joining $\phi_e(M)$ to $\phi_e(\mathsf{y}_j)$ in T_e, and lines parallel to the edges of T_o to lines parallel to the edges of T_e. Since the distribution of $\gamma_n(r, M)$ involves only probability content of unions and intersections of regions bounded by precisely such lines and the probability content of such regions is preserved since uniformity is preserved; the desired result follows. □

Note that geometry invariance of $\gamma\left(\mathcal{X}_n, N_{PE}^{r=\infty}, M\right)$ also follows trivially, since for $r = \infty$, we have $\gamma_n(r = \infty, M) = 1$ a.s. for all $\mathcal{X}_n$ from any F with support in $T(\mathcal{Y}_3) \setminus \mathcal{Y}_3$.

Based on Theorem 4 we may assume that $T(\mathcal{Y}_3)$ is a standard equilateral triangle with $\mathcal{Y}_3 = \{(0,0),(1,0),(1/2,\sqrt{3}/\,2)\}$ for $N_{PE}^r(\cdot, M)$ with M-vertex regions.

Notice that, we proved the geometry invariance property for $N_{PE}^{\,r}$ where M-vertex regions are defined with the lines joining $\mathcal{Y}_3$ to M. On the other hand, if we use the orthogonal projections from M to the edges, the vertex regions, hence N_{PE}^r will depend on the geometry of the triangle. That is, the orthogonol projections from M to the edges will not be mapped to the orthogonal projections in the standard equilateral triangle. Hence with the choice of the former type of M-vertex regions, it suffices to work on the standard equilateral triangle. On the other hand, with the orthogonal projections, the exact and asymptotic distribution of γ_n will depend on c_1, c_2, so one needs to do the calculations for each possible combination of c_1, c_2.

5.3. *The degenerate case with* $\gamma_n(r, M) \xrightarrow{p} 1$

Below, we prove that $\gamma_n(r, M)$ is degenerate in the limit for $r > 3/2$.

Theorem 5. Suppose $\mathcal{X}_n$ is a set of iid random variables from a continuous distribution F on $T(\mathcal{Y}_3)$. If $M \notin \mathcal{T}_r$ (see Fig. 5 and Eq. (2) for $\mathcal{T}_r$), then $\lim_{n\to\infty} P\left(\gamma_n(r, M) = 1\right) = 1$ for all $M \in \mathbb{R}^2 \setminus \mathcal{Y}_3$.

Proof: Suppose $M \notin \mathcal{T}_r$. Then $\mathcal{R}_S\left(N_{PE}^r, M\right)$ is nonempty with positive area. Hence the result follows by Lemma 2. □

Corollary 1. Suppose $\mathcal{X}_n$ is a set of iid random variables from a continuous distribution F on $T(\mathcal{Y}_3)$. Then for $r > 3/2$, $\lim_{n\to\infty} P\left(\gamma_n(r, M) = 1\right) = 1$ for all $M \in R^2 \setminus \mathcal{Y}_3$.

Proof: For $r > 3/2$, $\mathcal{T}_r = \emptyset$, so $M \notin \mathcal{T}_r$. Hence the result follows by Theorem 6. □

We estimate the distribution of $\gamma_n(r, M)$ with $r = 2$ and $M = M_C$ for various n empirically. In Table 1 (left), we present the empirical estimates of $\gamma_n(r, M)$ with $n = 10,\ 20,\ 30,\ 50,\ 100$ based on 1000 Monte Carlo replicates in T_e. Observe that the empirical estimates are in agreement with the asymptotic distribution given in Corollary 1.

The asymptotic distribution of $\gamma_n(r, M)$ for $r < 3/2$ depends on the relative position of M with respect to the triangle $\mathcal{T}_r$.

5.4. *The degenerate case with* $\gamma_n(r, M) \xrightarrow{p} 3$

Theorem 6. Suppose $\mathcal{X}_n$ is a set of iid random variables from a continuous distribution F on $T(\mathcal{Y}_3)$. If $M \in (\mathcal{T}_r)^\circ$, then $P\left(\gamma_n(r, M) = 3\right) \to 1$ as $n \to \infty$.

Table 2
The number of $\gamma_n(r, M) = k$ out of $N = 1000$ Monte Carlo replicates with $r = 5/4$ and $M = \left(3/5, \sqrt{3}/10\right)$

$k\backslash n$	10	20	30	50	100	500	1000	2000
1	118	60	51	39	15	1	2	1
2	462	409	361	299	258	100	57	29
3	420	531	588	662	727	899	941	970

Table 3
The number of $\gamma_n(r, M) = k$ out of $N = 1000$ Monte Carlo replicates with $r = 5/4$ and $M = \left(7/10, \sqrt{3}/10\right)$

$k\backslash n$	10	20	30	50	100	500	1000	2000
1	174	118	82	61	22	5	1	1
2	532	526	548	561	611	617	633	649
3	294	356	370	378	367	378	366	350

We estimate the distribution of $\gamma_n(r, M)$ with $r = 5/4$ and $M = M_C$ for various n values empirically. In Table 1 (right), we present the empirical estimates of $\gamma_n(r, M)$ with $n = 10, 20, 30, 50, 100$ based on 1000 Monte Carlo replicates in T_e. Observe that the empirical estimates are in agreement with our result in Theorem 7.

Theorem 7. Suppose $\mathcal{X}_n$ is a set of iid random variables from $\mathcal{U}(T(\mathcal{Y}_3))$. If $M \in \partial(\mathcal{T}_r)$, then $P\left(\gamma_n(r, M) > 1\right) \to 1$ as $n \to \infty$.

For $M \in \partial(\mathcal{T}_r)$, there are two separate cases:

(i) $M \in \partial(\mathcal{T}_r) \setminus \{t_1(r), t_2(r), t_3(r)\}$ where $t_j(r)$ with $j \in \{1, 2, 3\}$ are the vertices of $\mathcal{T}_r$ whose explicit forms are given in Eq. (2).
(ii) $M \in \{t_1(r), t_2(r), t_3(r)\}$.

Theorem 8. Suppose $\mathcal{X}_n$ is a set of iid random variables from $\mathcal{U}(T(\mathcal{Y}_3))$. If $M \in \partial(\mathcal{T}_r) \setminus \{t_1(r), t_2(r), t_3(r)\}$, then $P\left(\gamma_n(r, M) = 3\right) \to 1$ as $n \to \infty$.

We estimate the distribution of $\gamma_n(r, M)$ with $r = 5/4$ and $M = \left(3/5, \sqrt{3}/10\right) \in \partial(\mathcal{T}_r) \setminus \{t_1(r), t_2(r), t_3(r)\}$ for various n empirically. In Table 2 we present empirical estimates of $\gamma_n(r, M)$ with $n = 10, 20, 30, 50, 100, 500, 1000, 2000$ based on 1000 Monte Carlo replicates in T_e. Observe that the empirical estimates are in agreement with our result in Theorem 9.

5.5. The nondegenerate case

Theorem 9. Suppose $\mathcal{X}_n$ is a set of iid random variables from $\mathcal{U}(T(\mathcal{T}_3))$. If $M \in \{t_1(r), t_2(r), t_3(r)\}$, then $P\left(\gamma_n(r, M) = 2\right) \to p_r$ as $n \to \infty$ where $p_r \in (0, 1)$ is provided in Eq. (4) but only numerically computable.

For example, $p_{r=5/4} \approx 0.6514$ and $p_{r=\sqrt{2}} \approx 0.4826$.

So the asymptotic distribution of $\gamma_n(r, M)$ with $r \in [1, 3/2)$ and $M \in \{t_1(r), t_2(r), t_3(r)\}$ is given by

$$\gamma_n(r, M) \sim 2 + \mathrm{BER}(1 - p_r). \tag{7}$$

We estimate the distribution of $\gamma_n(r, M)$ with $r = 5/4$ and $M = \left(7/10, \sqrt{3}/10\right)$ for various n empirically. In Table 3, we present the empirical estimates of $\gamma_n(r, M)$ with $n = 10, 20, 30, 50, 100, 500, 1000, 2000$ based on 1000 Monte Carlo replicates in T_e. Observe that the empirical estimates are in agreement with our result $p_{r=5/4} \approx 0.6514$.

Remark 2. For $r = 3/2$, as $n \to \infty$, $P\left(\gamma_n(r, M_C) > 1\right) \to 1$ at rate $O\left(n^{-1}\right)$. □

Theorem 10. Suppose $\mathcal{X}_n$ is a set of iid random variables from $\mathcal{U}(T(\mathcal{Y}_3))$. Then for $r = 3/2$, as $n \to \infty$,

$$\gamma_n(3/2, M_C) \sim 2 + \mathrm{BER}(p \approx 0.2487) \tag{8}$$

Table 4
The number of $\gamma_n(3/2, M_C) = k$ out of $N = 1000$ Monte Carlo replicates

$k\backslash n$	10	20	30	50	100	500	1000	2000
1	151	82	61	50	27	2	3	1
2	602	636	688	693	718	753	729	749
3	247	282	251	257	255	245	268	250

For the proof of Theorem 10, see [3,4].

Using Theorem 10,

$$\lim_{n\to\infty} \mathbf{E}\left[\gamma_n(3/2, M_C)\right] = 3 - p_{3/2} \approx 2.2587 \tag{9}$$

and

$$\lim_{n\to\infty} \mathbf{V}_{ar}\left[\gamma_n(3/2, M_C)\right] = 6 + p_{3/2} - p_{3/2}^2 \approx 0.1917. \tag{10}$$

Indeed, the finite sample distribution of $\gamma_n(3/2, M_C)$ hence the finite sample mean and variance can also be obtained by numerical methods.

We also estimate the distribution of $\gamma_n(3/2, M_C)$ for various n empirically. The empirical estimates for $n = 10, 20, 30, 50, 100, 500, 1000, 2000$ based on 1000 Monte Carlo replicates are given in Table 4 observe that the estimates are in agreement with our result $p_{r=3/2} \approx 0.7413$.

5.6. Distribution of the $\gamma_n(r, M)$ in multiple triangles

So far we have worked with data in one Delaunay triangle, i.e., $m = 3$ or $J_3 = 1$. In this section, we present the asymptotic distribution of the domination number of r-factor PCDs in multiple Delaunay triangles. Suppose $\mathcal{Y}_m = \{\mathsf{y}_1, \mathsf{y}_2, \ldots, \mathsf{y}_m\} \subset \mathbb{R}^2$ be a set of m points in general position with $m > 3$ and no more than 3 points are cocircular. Then there are $J_m > 1$ Delaunay triangles each of which is denoted as $\mathcal{T}_j$. Let M^j be the point in $\mathcal{T}_j$ that corresponds to M in T_e, $\mathcal{T}_r^j$ be the triangle that corresponds to $\mathcal{T}_r$ in T_e, and $t_i^j(r)$ be the vertices of $\mathcal{T}_r^j$ that correspond to $t_i(r)$ in T_e for $i \in \{1, 2, 3\}$. Moreover, let $n_j := |\mathcal{X}_n \cap \mathcal{T}_j|$, the number of X points in Delaunay triangle $\mathcal{T}_j$. For $\mathcal{X}_n \subset \mathcal{C}_H(\mathcal{Y}_m)$, let $\gamma_{n_j}(r, M^j)$ be the domination number of the digraph induced by vertices of $\mathcal{T}_j$ and $\mathcal{X}_n \cap \mathcal{T}_j$. Then the domination number of the r-factor PCD in J_m triangles is

$$\gamma_n(r, M, J_m) = \sum_{j=1}^{J_m} \gamma_{n_j}(r, M^j).$$

See Fig. 10 (left) for the 77 $\mathcal{X}$ points that are in $\mathcal{C}_H(\mathcal{Y}_m)$ out of the 200 $\mathcal{X}$ points plotted in Fig. 1. Observe that 10 $\mathcal{Y}$ points yield $J_{10} = 13$ Delaunay triangles. In Fig. 10 (right) are the corresponding arcs for $M = M_C$ and $r = 3/2$. The corresponding $\gamma_n = 22$. Suppose $\mathcal{X}_n$ is a set of iid random variables from $\mathcal{U}(\mathcal{C}_H(\mathcal{Y}_m))$, the uniform distribution on convex hull of $\mathcal{Y}_m$ and we construct the r-factor PCDs using the points M^j that correspond to M in T_e. Then for fixed m (or fixed J_m), as $n \to \infty$, so does each n_j. Furthermore, as $n \to \infty$, each component $\gamma_{n_j}(r, M^j)$ become independent. Therefore using Eq. (3), we can obtain the asymptotic distribution of $\gamma_n(r, M, J_m)$. As $n \to \infty$, for fixed J_m,

$$\gamma_n(r, M, J_m) \sim \begin{cases} 2\,J_m + \mathrm{BIN}(J_m, 1 - p_r), & \text{for } M^j \in \{t_1^j(r), t_2^j(r), t_3^j(r)\} \text{ and } r \in [1, 3/2], \\ J_m, & \text{for } r > 3/2, \\ 3\,J_m, & \text{for } M \in \mathcal{T}_r^j \setminus \{t_1^j(r), t_2^j(r), t_3^j(r)\} \text{ and } r \in [1, 3/2), \end{cases} \tag{11}$$

where $\mathrm{BIN}(n, p)$ stands for binomial distribution with n trials and probability of success p, for $r \in [1, 3/2)$ and $M \in \{t_1(r), t_2(r), t_3(r)\}$, p_r is given in Eq. (3) and for $r = 3/2$ and $M = M_C$, $p_r \approx 0.7413$ (see Eq. (8)).

Table 5
The number of $\gamma_n(4/3, M_C) = k$ out of $N = 1000$ Monte Carlo replicates

$k\backslash n$	10	20	30	40	50	100	200	500	1000	2000
1	52	18	5	5	4	0	0	0	0	0
2	385	308	263	221	219	155	88	41	31	19
3	348	455	557	609	621	725	773	831	845	862
4	215	219	175	165	156	120	139	128	124	119

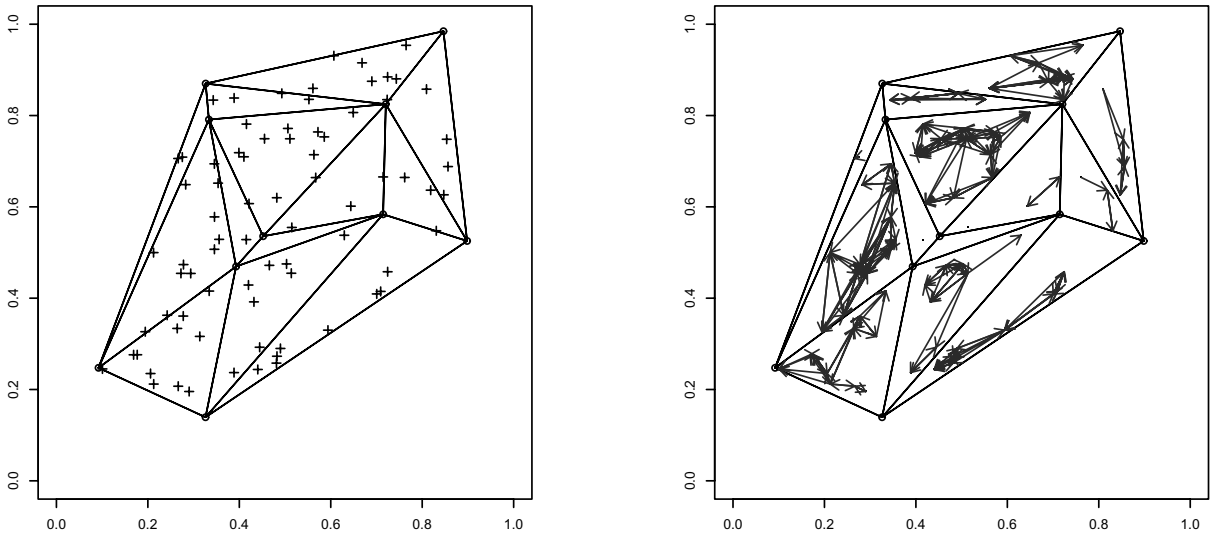

Fig. 10. The 77 $\mathcal{X}$ points (crosses) in the convex hull of $\mathcal{Y}$ points (circles) given in Fig. 1 (left) and the corresponding arcs (right) of r-factor proportional edge PCD with $r = 3/2$ and $M = M_C$.

5.7. Extension of N_{PE}^r to higher dimensions

The extension to $\mathbb{R}^d$ for $d > 2$ with $M = M_C$ is provided in [4], but the extension for general M is similar.

Let $\gamma_n(r, M, d) := \gamma(\mathcal{X}_n, N_{PE}^r, M, d)$ be the domination number of the PCD based on the extension of $N_{PE}^r(\cdot, M)$ to $\mathbb{R}^d$. Then it is easy to see that $\gamma_n(r, M, 3)$ is nondegenerate as $n \to \infty$ for $r = 4/3$. In $\mathbb{R}^d$, it can be seen that $\gamma_n(r, M, d)$ is nondegenerate in the limit only when $r = (d+1)/d$. Furthermore, for large d, asymptotic distribution of $\gamma_n(r, M, d)$ is nondegenerate at values of r closer to 1. Moreover, it can be shown that $\lim_{n\to\infty} P\big(2 \leqslant \gamma_n(r = (d+1)/d, M, d) \leqslant d+1\big) = 1$ and we conjecture the following.

Conjecture 1. Suppose $\mathcal{X}_n$ is a set of iid random variables from the uniform distribution on a simplex in $\mathbb{R}^d$. Then the domination number $\gamma_n(r, M)$ in the simplex satisfies

$$\lim_{n\to\infty} P\left(d \leqslant \gamma_n((d+1)/d, M, d) \leqslant d+1\right) = 1.$$

For instance, with $d = 3$ we estimate the empirical distribution of $\gamma(\mathcal{X}_n, 4/3)$ for various n. The empirical estimates for $n = 10, 20, 30, 40, 50, 100, 200, 500, 1000, 2000$ based on 1000 Monte Carlo replicates for each n are given in Table 5.

6. Discussion

The r-factor proportional-edge proximity catch digraphs (PCDs), when compared to class cover catch digraphs (CCCDs), have some advantages. The asymptotic distribution of the domination number $\gamma_n(r, M)$ of the r-factor PCDs, unlike that of CCCDs, is mathematically tractable (but computable by numerical integration). A minimum

dominating set can be found in polynomial time for r-factor PCDs in $\mathbb{R}^d$ for all $d \geqslant 1$, but finding a minimum dominating set is an NP-hard problem for CCCDs (except for $\mathbb{R}$). These nice properties of r-factor PCDs are due to the geometry invariance of distribution of $\gamma_n(r, M)$ for uniform data in triangles.

On the other hand, CCCDs are easily extendable to higher dimensions and are defined for all $\mathcal{X}_n \subset \mathbb{R}^d$, while r-factor PCDs are only defined for $\mathcal{X}_n \subset \mathcal{C}_H(\mathcal{Y}_m)$. Furthermore, the CCCDs based on balls use proximity regions that are defined by the obvious metric, while the PCDs in general do not suggest a metric. In particular, our r-factor PCDs are based on some sort of dissimilarity measure that has no underlying metric.

The finite sample distribution of $\gamma_n(r, M)$, although computationally tedious, can be found by numerical methods, while that of CCCDs can only be empirically estimated by Monte Carlo simulations. Moreover, we had to introduce many auxiliary tools to compute the distribution of $\gamma_n(r, M)$ in $\mathbb{R}^2$. Same tools will work in higher dimensions, perhaps with more complicated geometry.

The r-factor PCDs have applications in classification and testing spatial patterns of segregation or association. The former can be performed building discriminant regions for classification in a manner analogous to the procedure proposed in [17]; and the latter can be performed by using the asymptotic distribution of $\gamma_n(r, M)$ similar to the procedure used in [4].

Acknowledgements

This work was partially supported by the Defense Advanced Research Projects Agency as administered by the Air Force Office of Scientific Research under contract DOD F49620-99-1-0213 and by Office of Naval Research Grant N00014-95-1-0777. We also thank anonymous referees, whose constructive comments and suggestions greatly improved the presentation and flow of this article.

Appendix

First, we begin with a remark that introduces some terminology which we will use for asymptotics throughout this appendix.

Remark 3. Suppose $\mathcal{X}_n$ is a set of iid random variables from F with support $\mathcal{S}(F) \subseteq \Omega$. If over a sequence $\Omega_n \subseteq \Omega$, $n = 1, 2, 3, \ldots$, X restricted to Ω_n, $X|_{\Omega_n}$, has distribution F_n with $F_n(x) = F(x)/P_F(X \in \Omega_n)$ and $P_F(X \in \Omega_n) \to 1$ as $n \to \infty$, then we call F_n the *asymptotically accurate distribution* of X and Ω_n the *asymptotically accurate support* of F. If F has density f, then $f_n = f(x)/P_F(X \in \Omega_n)$ is called the *asymptotically accurate pdf* of X. In both cases, if we are concerned with asymptotic results, for simplicity we will, respectively, use F and f for asymptotically accurate distribution and pdf. Conditioning will be implied by stating that $X \in \Omega_n$ with probability 1, as $n \to \infty$ or for sufficiently large n. □

Proof of Theorem 1

Without loss of generality, assume $T(\mathcal{Y}_3) = T_b = T((0,0),(1,0),(c_1,c_2))$ Note that the probability of edge extrema all being equal to each other is $P(X_{e_1}(n) = X_{e_2}(n) = X_{e_3}(n)) = \mathbf{I}(n = 1)$. Let $E_{c,2}(n)$ be the event that there are only two distinct (closest) edge extrema. Then for $n > 1$,

$$P(E_{c,2}(n)) = P(X_{e_1}(n) = X_{e_2}(n)) + P(X_{e_1}(n) = X_{e_3}(n)) + P(X_{e_2}(n) = X_{e_3}(n))$$

since the intersection of the events $\{X_{e_i}(n) = X_{e_j}(n)\}$ and $\{X_{e_i}(n) = X_{e_k}(n)\}$ for distinct i, j, k is equivalent to the event $\{X_{e_1}(n) = X_{e_2}(n) = X_{e_3}(n)\}$. Notice also that $P(E_{c,2}(n = 2)) = 1$. So, for $n > 2$, there are two or three distinct edge extrema with probability 1. Hence $P(E_{c,3}(n)) + P(E_{c,2}(n)) = 1$ for $n > 2$.

By simple integral calculus, we can show that $P(E_{c,2}(n)) \to 0$ as $n \to \infty$, which will imply the desired result. □

Proof of Theorem 6

??

Note that $(\mathcal{T}_r)^\circ \neq \emptyset$ iff $r < 3/2$. Suppose $M \in (\mathcal{T}_r)^\circ$. Then for any point u in $R_M(\mathsf{y}_j)$, $N_{PE}^r(u, M) \subsetneq T(\mathcal{Y}_3)$, because there is a tiny strip adjacent to edge e_j not covered by $N_{PE}^r(u, M)$, for each $j \in \{1, 2, 3\}$. Then, $N_{PE}^r(u, M) \cup N_{PE}^r(v, M) \subsetneq T(\mathcal{Y}_3)$ for all $(u, v) \in R_M(\mathsf{y}_1) \times R_M(\mathsf{y}_2)$. Pick $\sup_{(u,v)\in R_M(\mathsf{y}_1)\times R_M(\mathsf{y}_2)} N_{PE}^r(u, M) \cup N_{PE}^r(v, M) \subsetneq T(\mathcal{Y}_3)$. Then $T(\mathcal{Y}_3) \setminus \left[\sup_{(u,v)\in R_M(\mathsf{y}_1)\times R_M(\mathsf{y}_2)} N_{PE}^r(u, M) \cup N_{PE}^r(v, M)\right]$ has positive area. So

$$\mathcal{X}_n \cap \left[T(\mathcal{Y}_3) \setminus \left[\sup_{(u,v)\in R_M(\mathsf{y}_1)\times R_M(\mathsf{y}_2)} N_{PE}^r(u, M) \cup N_{PE}^r(v, M)\right]\right] \neq \emptyset$$

with probability 1 for sufficiently large n. (The supremum of a set functional $A(x)$ over a range B is defined as the set $S := \sup_{x\in B} A(x)$ such that S is the smallest set satisfying $A(x) \subseteq S$ for all $x \in B$.) Then at least three points–one for each vertex region – are required to dominate $\mathcal{X}_n$. Hence for sufficiently large n, $\gamma_n(r, M) \geqslant 3$ with probability 1, but $\kappa\left(N_{PE}^r\right) = 3$ by Theorem 2. Then $\lim_{n\to\infty} P\left(\gamma_n(r, M) = 3\right) = 1$ for $r < 3/2$. □

Proof of Theorem 7

Let $M = (m_1, m_2) \in \partial(\mathcal{T}_r)$, say $M \in q_3(r, x)$ (recall that $q_j(r, x)$ are defined such that $d(\mathsf{y}_j, e_j) = r \cdot d(q_j(r, x), \mathsf{y}_j)$ for $j \in \{1, 2, 3\}$), then $m_2 = \frac{\sqrt{3}\,(2-r)}{2\,r}$ and $m_1 \in \left[\frac{3\,(r-1)}{2\,r}, \frac{3-r}{2\,r}\right]$. Let $X_{e_j}(n)$ be one of the closest point(s) to the edge e_j; i.e., $X_{e_j}(n) \in \operatorname{argmin}_{X\in\mathcal{X}_n} d(X, e_j)$ for $j \in \{1, 2, 3\}$. Note that $X_{e_j}(n)$ is unique a.s. for each j.

Notice that for all $j \in \{1, 2, 3\}$, $X_{e_j}(n) \notin N_{PE}^r(X)$ for all $X \in \mathcal{X}_n \cap R_M(\mathsf{y}_j)$ implies that $\gamma_n(r, M) > 1$ with probability 1. For sufficiently large n, $X_{e_j}(n) \notin N_{PE}^r(X)$ for all $X \in \mathcal{X}_n \cap R_M(\mathsf{y}_j)$ with probability 1, for $j \in \{1, 2\}$, by the choice of M. Hence we consider only $X_{e_3}(n)$. The asymptotically accurate pdf of $X_{e_3}(n)$ is

$$f_{e_3}(x, y) = n\left(\frac{A(S_U(x, y))}{A(T(\mathcal{Y}_3))}\right)^{n-1} \frac{1}{A(T(\mathcal{Y}_3))},$$

where $S_U(x, y)$ is the unshaded region in Fig. 11 (left) (for a given $X_{e_3}(n) = x_{e_3} = (x, y)$) whose area is $A(S_U(x, y)) = \sqrt{3}\left(2\,y - \sqrt{3}\right)^2/12$. Note that $X_{e_3}(n) \notin N_{PE}^r(X)$ for all $X \in \mathcal{X}_n \cap R_M(\mathsf{y}_3)$ iff $\mathcal{X}_n \cap [\Gamma_1^r(\mathcal{X}_n, M) \cap R_M(\mathsf{y}_3)] = \emptyset$. Then given $X_{e_3}(n) = (x, y)$,

$$P\left(\mathcal{X}_n \cap \left[\Gamma_1^r(\mathcal{X}_n, M) \cap R_M(\mathsf{y}_3)\right] = \emptyset\right) = \left(\frac{A(S_U(x, y)) - A\left(\Gamma_1^r(\mathcal{X}_n, M) \cap R_M(\mathsf{y}_3)\right)}{A(S_U(x, y))}\right)^{n-1},$$

where $A\left(\Gamma_1^r(\mathcal{X}_n, M) \cap R_M(\mathsf{y}_3)\right) = \frac{\sqrt{3}\,y^2}{3\,(r-1)\,r}$ (see Fig. 12 (right) where the shaded region is $\Gamma_1^r(\mathcal{X}_n, M) \cap R_M(\mathsf{y}_3)$ for a given $X_{e_3}(n) = (x, y)$), then for sufficiently large n

$$P\left(\mathcal{X}_n \cap [\Gamma_1^r(\mathcal{X}_n, M) \cap R_M(\mathsf{y}_3)] = \emptyset\right) \approx \int \left(\frac{A(S_U(x, y)) - A\left(\Gamma_1^r(\mathcal{X}_n, M) \cap R_M(\mathsf{y}_3)\right)}{A(S_U(x, y))}\right)^{n-1}$$

$$f_{e_3}(x, y)\,dy\,dx = \int \frac{n}{AT(\mathcal{Y}_3)} \left(\frac{A(S_U(x, y)) - A\left(\Gamma_1^r(\mathcal{X}_n, M) \cap R_M(\mathsf{y}_3)\right)}{AT(\mathcal{Y}_3)}\right)^{n-1} dy\,dx.$$

Let

$$G(x, y) = \frac{A(S_U(x, y)) - A\left(\Gamma_1^r(\mathcal{X}_n, M) \cap R_M(\mathsf{y}_3)\right)}{A(T(\mathcal{Y}_3))} = \frac{4}{\sqrt{3}}\left(\frac{\sqrt{3}\left(2\,y - \sqrt{3}\right)^2}{12} - \frac{\sqrt{3}\,y^2}{3\,(r-1)\,r}\right),$$

which is independent on x, so we denote it as $G(y)$.

Let $\varepsilon > 0$ be sufficiently small, then for sufficiently large n,

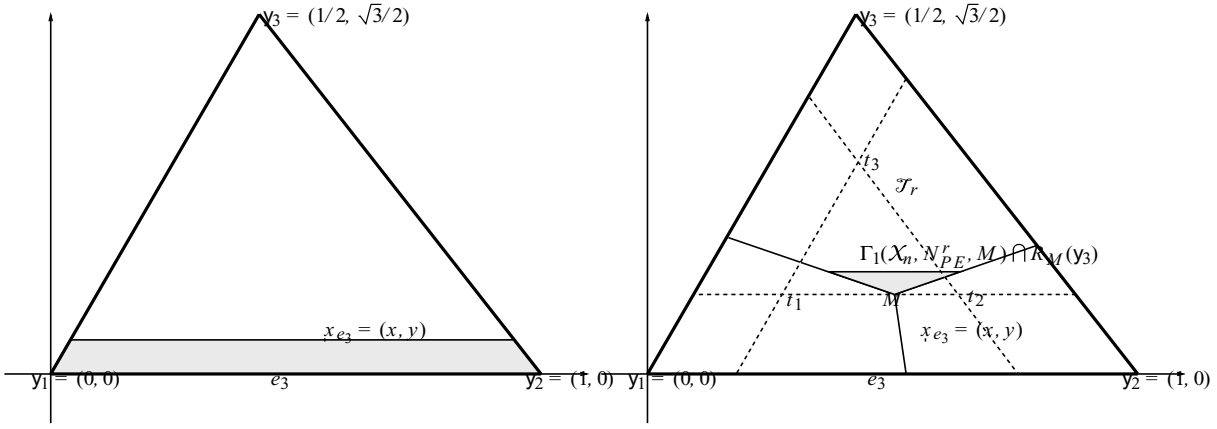

Fig. 11. A figure for the description of the pdf of $X_{e_3}(n)$ (left) and $\Gamma_1^r(\mathcal{X}_n, M)$ (right) given $X_{e_3}(n) = x_{e_3} = (x, y)$.

$$P\left(\mathcal{X}_n \cap [\Gamma_1^r(\mathcal{X}_n, M) \cap R_M(\mathsf{y}_3)] = \emptyset\right) \approx \int_0^{\varepsilon} \int_{y/\sqrt{3}}^{1-y/\sqrt{3}} n\, G(y)^{n-1}\, 4/\sqrt{3}\, dy\, dx$$

$$= \left(1 - 2\,y/\sqrt{3}\right) \int_0^{\varepsilon} n\, G(y)^{n-1}\, 4/\sqrt{3}\, dy.$$

The integrand is critical at $y = 0$, since $G(0) = 1$ (i.e., when $x_{e_3} \in e_3$). Furthermore, $G(y) = 1 - 4\,y/\sqrt{3} + O\left(y^2\right)$ around $y = 0$. Then letting $y = w/n$, we get

$$P\left(\mathcal{X}_n \cap [\Gamma_1^r(\mathcal{X}_n, M) \cap R_M(\mathsf{y}_3)] = \emptyset\right) \approx \left(1 - \frac{2\,w}{\sqrt{3}\,n}\right) \frac{4}{\sqrt{3}} \int_0^{n\varepsilon} \left(1 - \frac{4\,w}{\sqrt{3}\,n} + O\left(n^{-2}\right)\right)^{n-1} dw.$$

$$\text{letting } n \to \infty,\ \approx 4/\sqrt{3} \int_0^{\infty} \exp\left(-4\,w/\sqrt{3}\right)\, dw = 1.$$

Hence $\lim_{n\to\infty} P\left(\gamma_n(r, M) > 1\right) = 1$. For $M \in q_j(r, x) \cap \mathcal{T}_r$ with $j \in \{1, 2\}$ the result follows similarly. □

Proof of Theorem 8

Let $M = (m_1, m_2) \in \partial(\mathcal{T}_r) \setminus \{t_1(r),\, t_2(r),\, t_3(r)\}$, say $M \in q_3(r, x)$. Then $m_2 = \frac{\sqrt{3}\,(r-1)}{2\,r}$. Without loss of generality, assume $\frac{1}{2} \leqslant m_1 < \frac{3-r}{2\,r}$. See also Fig. 12.

Whenever $\mathcal{X}_n \cap R_M(\mathsf{y}_j) \neq \emptyset$, let

$$\widehat{Q}_j(n) \in \operatorname{argmin}_{X \in \mathcal{X}_n \cap R_M(\mathsf{y}_j)} d\left(X, e_j\right) = \operatorname{argmax}_{X \in \mathcal{X}_n \cap R_M(\mathsf{y}_j)} d(\ell(\mathsf{y}_j, X), \mathsf{y}_j) \text{ for } j \in \{1, 2, 3\}.$$

Note that at least one of the $\widehat{Q}_j(n)$ uniquely exists w.p. 1 for finite n and as $n \to \infty$, $\widehat{Q}_j(n)$ are unique w.p. 1. Then

$$\gamma_n(r, M) \leqslant 2 \text{ iff } \mathcal{X}_n \subset \left[N_{PE}^r\left(\widehat{Q}_1(n), M\right) \cup N_{PE}^r\left(\widehat{Q}_2(n), M\right)\right] \text{ or}$$

$$\mathcal{X}_n \subset \left[N_{PE}^r\left(\widehat{Q}_2(n), M\right) \cup N_{PE}^r\left(\widehat{Q}_3(n), M\right)\right] \text{ or } \mathcal{X}_n \subset \left[N_{PE}^r\left(\widehat{Q}_1(n), M\right) \cup N_{PE}^r\left(\widehat{Q}_3(n), M\right)\right].$$

Let $E_n^{i,j}$ be the event that $\mathcal{X}_n \subset N_{PE}^r\left(\widehat{Q}_i, M\right) \cup \left[{N_{PE}}^r\left(\widehat{Q}_j(n), M\right)\right]$ for $(i, j) \in \left\{(1,2), (1,3), (2,3)\right\}$. Then

$$P\left(\gamma_n(r, M) \leqslant 2\right) = P\left(E_n^{1,2}\right) + P\left(E_n^{2,3}\right) + P\left(E_n^{1,3}\right) - P\left(E_n^{1,2} \cap E_n^{2,3}\right) - P\left(E_n^{1,2} \cap E_n^{1,3}\right)$$
$$- P\left(E_n^{1,3} \cap E_n^{2,3}\right) + P\left(E_n^{1,2} \cap E_n^{2,3} \cap E_n^{1,3}\right).$$

But note that $P\left(E_n^{1,2}\right) \to 0$ as $n \to \infty$ by the choice of M since

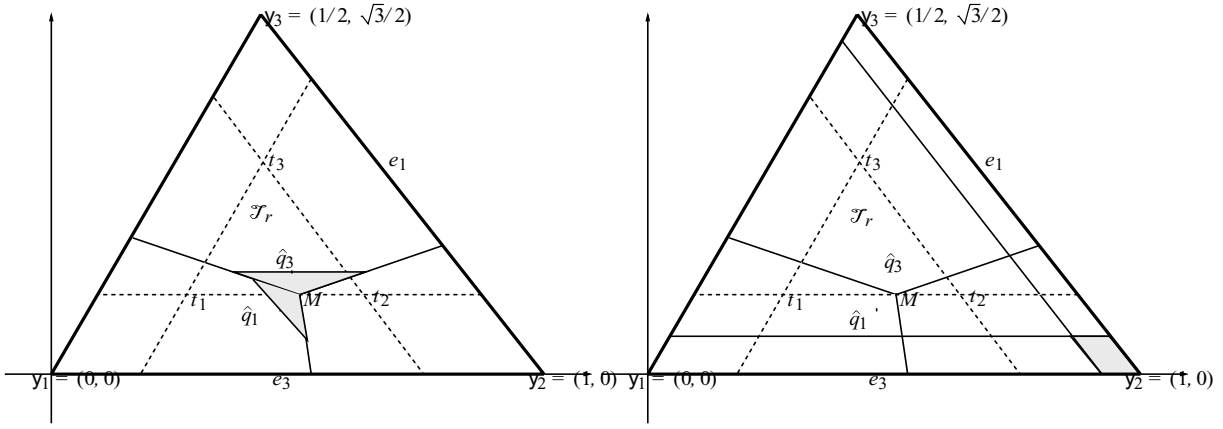

Fig. 12. A figure for the description of the pdf of $\hat{Q}_1(n)$ and $\hat{Q}_3(n)$ (left) and the unshaded region is $N_{PE}^r(\hat{q}_1, M) \cup N_{PE}^r(\hat{q}_3, M)$ (right).

$$\sup_{\substack{u \in R_M(\mathsf{y}_1) \\ v \in R_M(\mathsf{y}_2)}} N_{PE}^r(u, M) \cup N_{PE}^r(v, M) \subsetneq T(\mathcal{Y}_3),$$

and

$$P\left(\mathcal{X}_n \cap T(\mathcal{Y}_3) \setminus \left[\sup_{\substack{u \in R_M(\mathsf{y}_1) \\ v \in R_M(\mathsf{y}_2)}} N_{PE}^r(u, M) \cup N_{PE}^r(v, M)\right] \neq \emptyset\right) \to 1 \text{ as } n \to \infty.$$

Then,

$$P\left(E_n^{1,2}\right) - P\left(E_n^{1,2} \cap E_n^{2,3}\right) - P\left(E_n^{1,2} \cap E_n^{1,3}\right) + P\left(E_n^{1,2} \cap E_n^{2,3} \cap E_n^{1,3}\right) \leqslant 4\, P\left(E_n^{1,2}\right) \to 0 \text{ as } n \to \infty.$$

Therefore,

$$\lim_{n\to\infty} P\left(\gamma_n(r, M) \leqslant 2\right) = \lim_{n\to\infty} \left(P\left(E_n^{2,3}\right) + P\left(E_n^{1,3}\right)\right).$$

Furthermore, observe that $P\left(E_n^{1,3}\right) \geqslant P\left(E_n^{2,3}\right)$ by the choice of M. Then we first find $\lim_{n\to\infty} P\left(E_n^{1,3}\right)$. Given a realization of $\mathcal{X}_n$ with $\widehat{Q}_1(n) = \widehat{q}_1 = (x_1, y_1)$ and $\widehat{Q}_3(n) = \widehat{q}_3 = (x_3, y_3)$, the remaining $n-2$ points should fall, for example, in the undshaded region in Fig. 12 (left). Then the asymptotically accurate joint pdf of $\widehat{Q}_1(n), \widehat{Q}_3(n)$ is

$$f_{13}\left(\vec{\zeta}\right) = \frac{n\,(n-1)}{AT(\mathcal{Y}_3)^2} \left(\frac{AT(\mathcal{Y}_3) - A(S_R\left(\vec{\zeta}\right)\left(\vec{\zeta}\right))}{AT(\mathcal{Y}_3)}\right)^{n-2}$$

where $\vec{\zeta} = (x_1, y_1, x_3, y_3)$, $S_R\left(\vec{\zeta}\right)$ is the shaded region in Fig. 12 (left) whose area is $A(S_R\left(\vec{\zeta}\right)) = \frac{\sqrt{3}\left(2\,r\,y_3 - \sqrt{3}\,(r-1)\right)^2}{12\,r\,(r-1)} + \frac{\sqrt{3}\left[2\,\sqrt{3}\,r\,y_1 - 3\,(r-1) + 6\,r\,(x_1 - m_1)\right]^2}{72\,r\,(1-r\,(2\,m_1-1))}$.

Given $\widehat{Q}_j(n) = \widehat{q}_j = (x_j, y_j)$ for $j \in \{1, 3\}$,

$$P\left(E_n^{1,3}\right) = \left(\frac{A\left(N_{PE}^r\left(\widehat{q}_1, M\right) \cup N_{PE}^r\left(\widehat{q}_3, M\right)\right) - A(S_R\left(\vec{\zeta}\right))}{AT(\mathcal{Y}_3) - A(S_R\left(\vec{\zeta}\right))}\right)^{n-2}$$

then for sufficiently large n

$$P\left(E_n^{1,3}\right) \approx \int \left(\frac{A\left(N_{PE}^r\left(\widehat{q}_1, M\right) \cup N_{PE}^r\left(\widehat{q}_3, M\right)\right) - A(S_R\left(\vec{\zeta}\right))}{AT(\mathcal{Y}_3) - A(S_R\left(\vec{\zeta}\right))}\right)^{n-2} f_{13}\left(\vec{\zeta}\right) d\vec{\zeta},$$

$$= \int \frac{n\,(n-1)}{A(T(\mathcal{Y}_3))^2} \left(\frac{A\left(N_{PE}^r\left(\widehat{q}_1, M\right) \cup N_{PE}^r\left(\widehat{q}_3, M\right)\right) - A(S_R(\vec{\zeta}))}{AT(\mathcal{Y}_3)} \right)^{n-2} d\vec{\zeta}$$

where

$$A\left(N_{PE}^r\left(\widehat{q}_1, M\right) \cup N_{PE}^r\left(\widehat{q}_3, M\right)\right) = \frac{\sqrt{3}}{4} - \left(\frac{\left(\sqrt{3}\,r\,y_1 + 3\,r\,x_1 - 3\right)\,\left(\sqrt{3}\,(r-1) - 2\,r\,y_3\right)}{6} \right).$$

See Fig. 12 (right) for $N_{PE}^r(\widehat{q}_1, M) \cup N_{PE}^r(\widehat{q}_3, M)$. Let

$$G(\vec{\zeta}) = \frac{A\left(N_{PE}^r\left(\widehat{q}_1, M\right) \cup N_{PE}^r\left(\widehat{q}_3, M\right)\right) - A(S_R(\vec{\zeta}))}{AT(\mathcal{Y}_3)}.$$

Note that the integral is critical at $x_1 = x_3 = m_1$ and $y_1 = y_3 = m_2$, since $G(\vec{\zeta}) = 1$. Since $N_{PE}^r(x, M_C)$ depends on the distance $d(x, e_j)$ for $x \in R_M(\mathsf{y}_j)$, we make the change of variables $(x_1, y_1) \to (d(M, e_1) + z_1, y_1)$ where $d(M, e_1) = \frac{\sqrt{3}\,(r+1-2\,r\,m_1)}{4\,r}$ and $(x_3, y_3) \to (x_3, m_2 + z_3)$ then $G(\vec{\zeta})$ depends only on z_1, z_3, we denote it $G(z_1, z_3)$ which is

$$G(z_1, z_3) = 1 - \frac{8\,r\,z_1^2}{3\,(1 + r\,(1 - 2\,m_1))} - \frac{4\,r\,z_3^2}{3\,(r-1)} - \frac{2\,r\,z_3\,\left(\sqrt{3}\,(3-r)\right) + r\,\left(4\,z_1 - 2\,\sqrt{3}\,m_1\right)}{3}.$$

The new integrand is $\frac{n\,(n-1)}{AT(\mathcal{Y}_3)^2} G(z_1, z_3)^{n-2}$. Integrating with respect to x_3 and y_1 yields $\frac{2\sqrt{3}\,z_3\,r}{3\,(r-1)}$ and $\frac{4\,\sqrt{3}\,r\,z_1}{3\,(2\,r\,m_1 - r - 1)}$, respectively. Hence for sufficiently large n

$$P\left(E_n^{1,3}\right) \approx \int_0^{\varepsilon} \int_0^{\varepsilon} \frac{n\,(n-1)}{AT(\mathcal{Y}_3)^2} \left(\frac{2\,\sqrt{3}\,z_3\,r}{3\,(r-1)} \right) \left(\frac{4\,\sqrt{3}\,r\,z_1}{3\,(2\,r\,m_1 - r - 1)} \right) G(z_1, z_3)^{n-2} dz_1\,dz_3.$$

Note that the new integral is critical when $z_1 = z_3 = 0$, so we make the change of variables $z_1 = w_1/\sqrt{n}$ and $z_3 = w_3/n$ then $G(z_1, z_3)$ becomes

$$G(w_1, w_3) = 1 + \frac{1}{n} \left(\frac{2\,\sqrt{3}\,r\,(r - 3 + 2\,r\,m_1)}{3} w_3 + \frac{8\,r}{3\,(r + 1 - 2\,r\,m_1)}\,w_1^2 \right) + O\left(n^{-3/2}\right),$$

so for sufficiently large n

$$\begin{aligned} P\left(E_n^{1,3}\right) &\approx \int_0^{\sqrt{n}\,\varepsilon} \int_0^{n\,\varepsilon} \frac{(n-1)}{n^3}\,\frac{16}{3} \left(\frac{2\,\sqrt{3}\,r}{3\,(r-1)} \right) \left(\frac{4\,\sqrt{3}\,r}{3\,(2\,r\,m_1 - r - 1)} \right) (-4\,m_1 + 2 + \sqrt{2}) w_1\,w_3 \\ &\qquad \left[1 - \frac{1}{n} \left(\frac{2\,\sqrt{3}\,r\,(r - 3 + 2\,r\,m_1)}{3} w_3 + \frac{8\,r}{3\,(r + 1 - 2\,r\,m_1)}\,w_1^2 \right) + O\left(n^{-3/2}\right) \right]^{n-2} dw_3 w_1, \\ &\approx O\left(n^{-1}\right) \int_0^{\infty} \int_0^{\infty} w_1\,w_3\,\exp\left(-\frac{2\,\sqrt{3}\,r\,(r - 3 + 2\,r\,m_1) w_3}{3} - \frac{8\,r\,w_1^2}{3\,(r + 1 - 2\,r\,m_1)} \right) dw_3 w_1 \\ &= O\left(n^{-1}\right) \end{aligned}$$

since $\int_0^{\infty} \int_0^{\infty} w_1\,w_3\,\exp\left(-\frac{2\,\sqrt{3}\,r\,(r-3+2\,r\,m_1)}{3} w_3 - \frac{8\,r}{3\,(r+1-2\,r\,m_1)}\,w_1^2 \right) dw_3 w_1 = \frac{3}{8\,r\,(3 - r\,(2\,m_1 + 1))}$, which is a finite constant. Then $P\left(E_n^{1,3}\right) \to 0$ as $n \to \infty$, which also implies $P\left(E_n^{2,3}\right) \to 0$ as $n \to \infty$. Then $P\left(\gamma_n(r, M) \leqslant 2\right) \to 0$. Hence the desired result follows. □

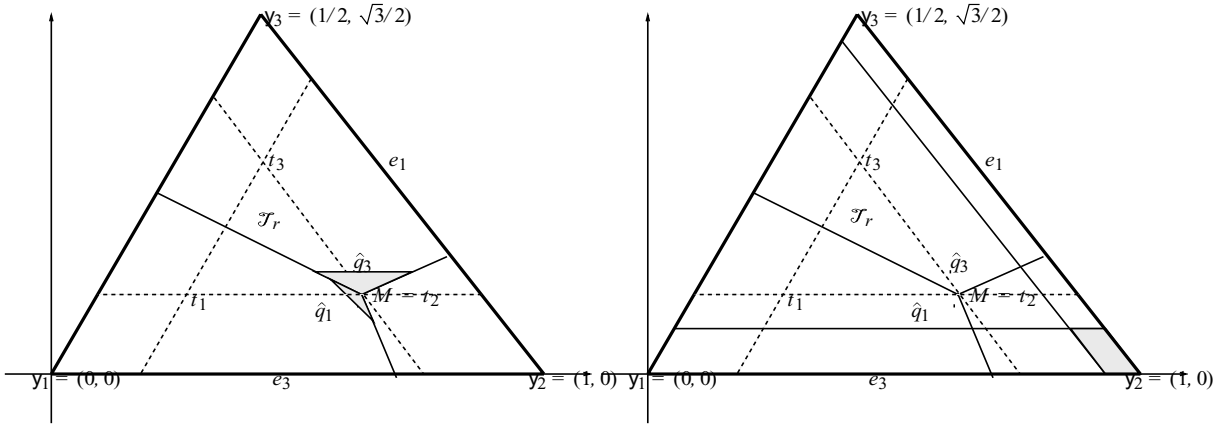

Fig. 13. A figure for the description of the pdf of $\hat{Q}_1(n)$ and $\hat{Q}_3(n)$ (left) and the unshaded region is $N_{PE}^r(\hat{q}_1, M) \cup N_{PE}^r(\hat{q}_3, M)$ (right) given $\hat{Q}_j(n) = \hat{q}_j$ for $j \in \{1, 3\}$.

Proof of Theorem 9

Let $M = (m_1, m_2) \in \{t_1(r), t_2(r), t_3(r)\}$. Without loss of generality, assume $M = t_2(r)$ then $m_1 = \frac{2-r+c_1\,(r-1)}{r}$ and $m_2 = \frac{c_2\,(r-1)}{r}$. See Fig. 13.

Let $\widehat{Q}_j(n)$ and the events $E_n^{i,j}$ be defined as in the proof of Theorem 8 for $(i, j) \in \left\{(1,2), (1,3), (2,3)\right\}$. Then as in the proof of Theorem 8,

$$P\left(\gamma_n(r, M) \leqslant 2\right) = P\left(E_n^{1,2}\right) + P\left(E_n^{2,3}\right) + P\left(E_n^{1,3}\right) - P\left(E_n^{1,2} \cap E_n^{2,3}\right) - P\left(E_n^{1,2} \cap E_n^{1,3}\right) - P\left(E_n^{1,3} \cap E_n^{2,3}\right) + P\left(E_n^{1,2} \cap E_n^{2,3} \cap E_n^{1,3}\right).$$

Observe that the choice of M implies that $P\left(E_n^{1,3}\right) \geqslant P\left(E_n^{2,3}\right)$ and by symmetry (in T_e) $P\left(E_n^{1,2}\right) = P\left(E_n^{2,3}\right)$. So first we find $P\left(E_n^{1,3}\right)$. As in the proof of Theorem 8 asymptotically accurate joint pdf of $\widehat{Q}_1(n), \widehat{Q}_3(n)$ is

$$f_{13}\left(\vec{\zeta}\right) = \frac{n\,(n-1)}{AT(\mathcal{Y}_3)^2}\left(\frac{AT(\mathcal{Y}_3) - A(S_R(\vec{\zeta}))}{AT(\mathcal{Y}_3)}\right)^{n-2}$$

where $\vec{\zeta} = (x_1, \mathsf{y}_1, x_3, \mathsf{y}_3)$ and $S_R\left(\vec{\zeta}\right)$ is the shaded region in Fig. 13 (left) whose area is

$$A(S_R(\vec{\zeta})) = \frac{\sqrt{3}\,\left(2\,r\,y_3 - \sqrt{3}\,(r-1)^2\right)}{12\,(r-1)\,r} + \frac{\sqrt{3}\,\left(\sqrt{3}\,r\,y_1 + 3\,x_1\,r - 3\right)^2}{36\,(r-1)\,r}.$$

Given $\widehat{Q}_j(n) = \widehat{q}_j = (x_j, y_j)$ for $j \in \{1, 3\}$,

$$P\left(E_n^{1,3}\right) = \left(\frac{A\left(N_{PE}^r\left(\widehat{q}_1, M\right) \cup N_{PE}^r\left(\widehat{q}_3, M\right)\right) - A(S_R(\vec{\zeta}))}{AT(\mathcal{Y}_3) - A(S_R(\vec{\zeta}))}\right)^{n-2},$$

then for sufficiently large n

$$P\left(E_n^{1,3}\right) \approx \int \left(\frac{A\left(N_{PE}^r\left(\widehat{Q} = \widehat{q}_1, M\right) \cup N_{PE}^r\left(\widehat{q}_3, M\right)\right) - A(S_R(\vec{\zeta}))}{AT(\mathcal{Y}_3) - A(S_R(\vec{\zeta}))}\right)^{n-2} f_{13}\left(\vec{\zeta}\right) d\vec{\zeta},$$

$$= \int \frac{n\,(n-1)}{AT(\mathcal{Y}_3)^2}\left(\frac{A\left(N_{PE}^r\left(\widehat{q}_1, M\right) \cup N_{PE}^r\left(\widehat{q}_3, M\right)\right) - A(S_R(\vec{\zeta}))}{AT(\mathcal{Y}_3)}\right)^{n-2} d\vec{\zeta}$$

where

$$A\left(N_{PE}^{r}\left(\widehat{q}_1, M\right) \cup N_{PE}^{r}\left(\widehat{q}_3, M\right)\right) = \frac{\sqrt{3}}{4} - \frac{\left(2\,r\,y_3 - \sqrt{3}\,(r-1)\right)\,\left(3 - \sqrt{3}\,r\,y_1 - 3\,r\,x_1\right)}{6}.$$

See Fig. 13 (right) for $N_{PE}^{r}(\widehat{q}_1, M) \cup N_{PE}^{r}(\widehat{q}_3, M)$. Let

$$G(\vec{\zeta}) = \frac{A\left(N_{PE}^{r}\left(\widehat{q}_1, M\right) \cup N_{PE}^{r}\left(\widehat{q}_3, M\right)\right) - A(S_R(\vec{\zeta}))}{AT(\mathcal{Y}_3)}.$$

Note that the integral is critical when $x_1 = x_3 = m_1$ and $y_1 = y_3 = m_2$, since $G(\vec{\zeta}) = 1$.

As in the proof of Theorem 8, we make the change of variables $(x\,_1, y_1) \to (d(M, e_1) + z_1, y_1)$ where $d(M, e_1) = \frac{\sqrt{3}\,(r-1)}{2\,r}$ and $(x_3, y_3) \to (x_3, m_2 + z_3)$. Then $G(\vec{\zeta})$ becomes

$$G(z_1, z_3) = 1 - \frac{4\,r}{3\,(r-1)}\,z_1^2 - \frac{4\,r}{3\,(r-1)}\,z_3^2 - \frac{8\,r^2}{3}\;z_1\,z_3.$$

The new integral is

$$\int \frac{n\,(n-1)}{AT(\mathcal{Y}_3)^2} G(z_1, z_3)^{n-2} dx_3 dy_1 dz_3 dz_1.$$

Note that $G(z_1, z_3)$ is independent of y_1, x_3, so integrating with respect to x_3 and y_1 yields $\frac{2\sqrt{3}\,r\,z_1}{3\,(r-1)}$ and $\frac{2\sqrt{3}\,r\,z_3}{3\,(r-1)}$, respectively. The new integral is critical at $z_1 = z_3 = 0$. Hence, for sufficiently large n and sufficiently small $\varepsilon > 0$, the integral becomes,

$$P\left(E_n^{1,3}\right) \approx \int_0^{\varepsilon}\int_0^{\varepsilon} \frac{n\,(n-1)}{AT(\mathcal{Y}_3)^2}\left(\frac{12\,r^2}{9\,(r-1)^2}\right)\,z_1\,z_3\,G(z_1, z_3)^{n-2}\;dz_1 dz_3.$$

Since the new integral is critical when $z_1 = z_2 = 0$, we make the change of variables $z_j = w_j/\sqrt{n}$ for $j \in \{1, 3\}$; then $G(z_1, z_3)$ becomes

$$G(w_1, w_3) = 1 - \frac{4\,r}{3\,n\,(r-1)}\left(w_1^2 + w_3^2 + 2\,r\,(r-1)\,w_1\,w_3)\right),$$

so

$$\begin{aligned} p_r := P\left(E_n^{1,3}\right) &\approx \int_0^{\sqrt{n}\,\varepsilon}\int_0^{\sqrt{n}\,\varepsilon} \frac{(n-1)}{n}\,\frac{16}{3}\left(2\left(\frac{12\,r^2}{9\,(r-1)^2}\right)w_1\,w_3\right) \\ &\quad \left[1 - \frac{4\,r}{3\,n\,(r-1)}\left(w_1^2 + w_3^2 + 2\,r\,(r-1)\,w_1\,w_3)\right)\right]^{n-2} dw_3 w_1, \;\; \text{letting } n \to \infty, \\ &\approx \int_0^{\infty}\int_0^{\infty} \frac{64}{9}\left(\frac{r}{r-1}\right)^2 w_1\,w_3\;\exp\left(\frac{4\,r}{3\,(r-1)}\left(w_1^2 + w_3^2 + 2\,r\,(r-1)\,w_1\,w_3\right)\right) dw_3 w_1 \end{aligned}$$

which is not analytically integrable, but p_r can be obtained by numerical integration, e.g., $p_{r=\sqrt{2}} \approx 0.4826$ and $p_{r=5/4} \approx 0.6514$.

Next, we find $\lim_{n\to\infty} P\left(E_n^{2,3}\right)$. The asymptotically accurate joint pdf of $\widehat{Q}_2(n), \widehat{Q}_3(n)$ is

$$f_{23}\left(\vec{\zeta}\right) = \frac{n\,(n-1)}{AT(\mathcal{Y}_3)^2}\left(\frac{AT(\mathcal{Y}_3) - A(S_R^2(\vec{\zeta}))}{AT(\mathcal{Y}_3)}\right)^{n-2}$$

where $\vec{\zeta} = (x_2, y_2, x_3, y_3)$ and $S_R^2(\vec{\zeta})$ is the shaded region in Fig. 14 (left) whose area is

$$A\left(S_R^2(\vec{\zeta})\right) = \frac{\sqrt{3}\,\left(2\,r\,\mathsf{y}_3 + \sqrt{3}\,(1-r)\right)}{12\,r\,(r-1)} + \frac{\sqrt{3}\,\left(\sqrt{3}\,r\,\mathsf{y}_2 - 3\,r\,x_2 - 3\,r + 6\right)}{36\,(2-r)\,r}.$$

As before,

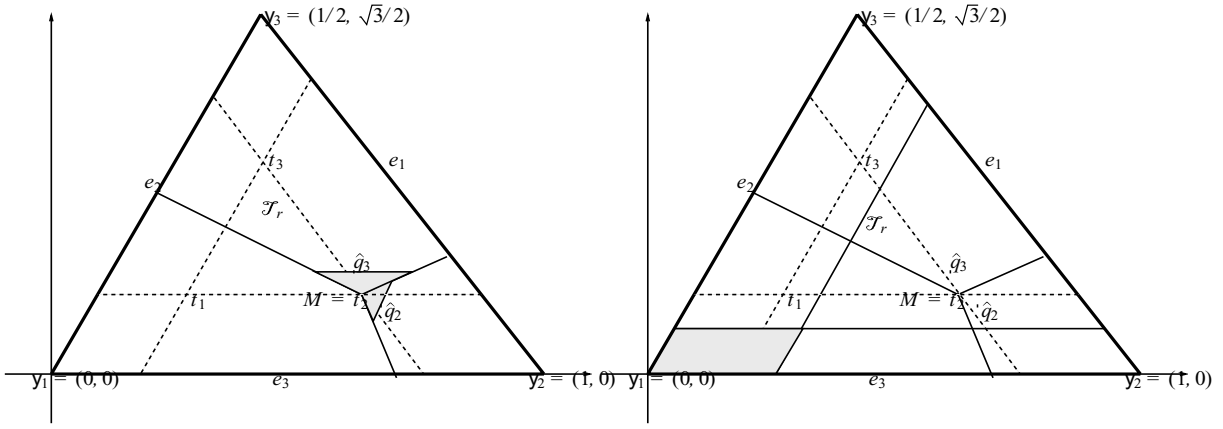

Fig. 14. A figure for the description of the pdf of $\hat{Q}_2(n)$ and $\hat{Q}_3(n)$ (left) and unshaded region is $N^r_{PE}(\hat{q}_2) \cup N^r_{PE}(\hat{q}_3)$ (right) given $\hat{Q}_j(n) = \hat{q}_j$ for $j \in \{2, 3\}$.

$$P\left(E_n^{2,3}\right) = \int \left(\frac{A\left(N^r_{PE}\left(\widehat{q}_2, M\right) \cup N^r_{PE}\left(\widehat{q}_3, M\right)\right) - A(S_R(\vec{\zeta}))}{AT(\mathcal{Y}_3) - A(S^2_R(\vec{\zeta}))} \right)^{n-2} f_{23}\left(\vec{\zeta}\right) d\vec{\zeta}$$

$$= \int \frac{n\,(n-1)}{AT(\mathcal{Y}_3)^2} \left(\frac{A\left(N^r_{PE}\left(\widehat{q}_2, M\right) \cup N^r_{PE}\left(\widehat{q}_3, M\right)\right) - A(S_R(\vec{\zeta}))}{AT(\mathcal{Y}_3)} \right)^{n-2} d\vec{\zeta},$$

where $A\left(N^r_{PE}\left(\widehat{q}_2, M\right) \cup N^r_{PE}\left(\widehat{q}_3, M\right)\right) = \frac{\sqrt{3}}{4} - \frac{\left(2\,r\,y_3 - \sqrt{3}\,(r-1)\right)\left(3 - \sqrt{3}\,r\,y_2 + 3\,r\,x_2 - 3\,r\right)}{6}$.

See Fig. 14 (right) for $N^r_{PE}(\widehat{q}_2) \cup N^r_{PE}(\widehat{q}_3, M)$. Let

$$G(\vec{\zeta}) = \frac{A\left(N^r_{PE}\left(\widehat{q}_2, M\right) \cup N^r_{PE}\left(\widehat{q}_3, M\right)\right) - A(S_R(\vec{\zeta}))}{AT(\mathcal{Y}_3)}.$$

Note that the integral is critical when $x_2 = x_3 = m_1$ and $y_2 = y_3 = m_2$, since $G(\vec{\zeta}) = 1$.

We make the change of variables $(x_3, y_3) \to (x_3, m_2 + z_3)$ and $(x_2, y_2) \to (d(M, e_2) + z_2, y_2)$ where $d(M, e_2) = \frac{\sqrt{3}\,(2-r)}{2\,r}$. Then $G(\vec{\zeta})$ becomes

$$G(z_2, z_3) = 1 - \frac{4\,r\,z_2^2}{3\,(2-r)} - \frac{4\,r\,z_2^2}{3\,(r-2)} - \frac{4\,\sqrt{3}\,r\,z_3\,(3-2\,r)}{3} - \frac{8\,r^2\,z_2\,z_3}{3}.$$

The new integral is

$$\int \frac{n\,(n-1)}{AT(\mathcal{Y}_3)^2} G(z_2, z_3)^{n-2} dx_3 dy_2 dz_3 dz_2.$$

The integrand is independent of x_3 and y_2, so integrating with respect to x_3 and y_2 yields $\frac{2\sqrt{3}\,r\,z_3}{3\,(r-1)}$ and $\frac{2\sqrt{3}\,r\,z_2}{3\,(2-r)}$, respectively. Hence, for sufficiently large n

$$P\left(E_n^{2,3}\right) \approx \int_0^{\varepsilon} \int_0^{\varepsilon} \frac{n\,(n-1)}{AT(\mathcal{Y}_3)^2} \left(\frac{4\,r^2}{3\,(r-1)\,(2-r)} \right) z_3\,z_2\,G(z_2, z_3)^{n-2} dz_2 dz_3.$$

Note that the new integral is critical when $z_2 = z_3 = 0$, so we make the change of variables $z_2 = w_2/\sqrt{n}$ and $z_3 = w_3/n$ then $G(z_2, z_3)$ becomes

$$G(w_2, w_3) = 1 - \frac{1}{n}\left[\frac{4\,r\,w_2^2}{3\,(2-r)} - \frac{4\,\sqrt{3}\,r\,w_3\,(3-2\,r)}{3} \right] + O\left(n^{-\frac{3}{2}}\right),$$

so for sufficiently large n

$$P\left(E_n^{2,3}\right) \approx \int_0^{\sqrt{n}\,\varepsilon} \int_0^{n\,\varepsilon} \frac{(n-1)}{n^2} \frac{64\,r^2}{9\,(r-1)\,(2-r)}\, w_2\, w_3 \left[1-\frac{1}{n}\left(\frac{4\,r\,w_2^2}{3\,(2-r)}-\frac{4\,\sqrt{3}\,r\,w_3\,(3-2\,r)}{3}\right)+O\left(n^{-\frac{3}{2}}\right)\right]^{n-2} dw_3 w_2,$$

$$\approx O\left(n^{-1}\right) \int_0^{\infty} \int_0^{\infty} w_2\, w_3\ \exp\left(-\frac{4\,r\,w_2^2}{3\,(2-r)}-\frac{4\,\sqrt{3}\,r\,u_3\,(3-2\,r)}{3}\right) dw_3 w_2 = O\left(n^{-1}\right)$$

since

$$\int_0^{\infty} \int_0^{\infty} w_2\, w_3\ \exp\left(-\frac{4\,r\,w_2^2}{3\,(2-r)}-\frac{4\,\sqrt{3}\,r\,u_3\,(3-2\,r)}{3}\right) dw_3 w_2 = \frac{27\,(2-r)}{384\,r^3\,(3-2\,r)^2}$$

which is a finite constant.

Thus we have shown that $P\left(E_n^{2,3}\right) \to 0$ as $n \to \infty$, which implies that as $n \to \infty$,

$$P\left(E_n^{2,3}\right) + P\left(E_n^{1,2}\right) - P\left(E_n^{1,2} \cap E_n^{2,3}\right) - P\left(E_n^{1,2} \cap E_n^{1,3}\right)$$
$$-P\left(E_n^{1,3} \cap E_n^{2,3}\right) + P\left(E_n^{1,2} \cap E_n^{2,3} \cap E_n^{1,3}\right) \leqslant 5\,P\left(E_n^{2,3}\right) \to 0.$$

Hence $\lim_{n\to\infty} P\left(\gamma_n(r,M) \leqslant 2\right) = \lim_{n\to\infty} P\left(E_n^{1,3}\right)$ and $\lim_{n\to\infty} P(\gamma_n(r,M) > 1) = 1$ together imply that $\lim_{n\to\infty} P(\gamma_n(r,M) = 2) = p_r$. □

References

[1] E. Ceyhan, *An Investigation of Proximity Catch Digraphs in Delaunay Tessellations*, PhD thesis, The Johns Hopkins University, Baltimore, MD, 2004, 21218.

[2] E. Ceyhan and C. Priebe, *Central similarity proximity maps in Delaunay tessellations*, In Proceedings of the Joint Statistical Meeting, Statistical Computing Section, American Statistical Association, 2003.

[3] E. Ceyhan and C. Priebe, *On the distribution of the domination number of random r-factor proportional-edge proximity catch digraphs. Technical Report 651*, Department of Applied Mathematics and Statistics, The Johns Hopkins University, Baltimore, 2004, MD 21218.

[4] E. Ceyhan and C.E. Priebe, The use of domination number of a random proximity catch digraph for testing spatial patterns of segregation and association, *Statistics and Probability Letters* **73** (2005), 37–50.

[5] E. Ceyhan, C.E. Priebe and D.J. Marchette, A new family of random graphs for testing spatial segregation, *Accepted for publication.* (*Available as Technical Report No. 644, with title Relative density of random τ-factor proximity catch digraph for testing spatial patterns of segregation and association, Department of Applied Mathematics and Statistics, The Johns Hopkins University, Baltimore, 2005, MD 21218.*).

[6] E. Ceyhan, C.E. Priebe and J.C. Wierman, Relative density of the random r-factor proximity catch digraphs for testing spatial patterns of segregation and association, *Computational Statistics & Data Analysis* **50**(8) (2006), 1925–1964.

[7] G. Chartrand and L. Lesniak, *Graphs & Digraphs*, Chapman & Hill, 1996.

[8] J. DeVinney, *The Class Cover Problem and its Applications in Pattern Recognition*, PhD thesis, The Johns Hopkins University, Baltimore, MD, 2003, 21218.

[9] J. DeVinney and C.E. Priebe, A new family of proximity graphs: Class cover catch digraphs, *Discrete Applied Mathematics* **154**(14) (2006), 1975–1982.

[10] J. DeVinney, C.E. Priebe, D.J. Marchette and D. Socolinsky, Random walks and catch digraphs in classification. http://www.galaxy.gmu.edu/interface/I02/I2002Proceedings/DeVinneyJason/DeVinneyJason.paper.pdf. Proceedings of the 34th Symposium on the Interface: Computing Science and Statistics, Vol. 34, 2002.

[11] J. DeVinney and J.C. Wierman, A SLLN for a one-dimensional class cover problem, *Statistics and Probability Letters* **59**(4) (2003), 425–435.

[12] J.W. Jaromczyk and G.T. Toussaint, Relative neighborhood graphs and their relatives, *Proceedings of IEEE* **80** (1992), 1502–1517.

[13] C. Lee, Domination in digraphs. *Journal of Korean Mathematical Society* **4** (1998), 843–853.

[14] D.J. Marchette and C.E. Priebe, Characterizing the scale dimension of a high dimensional classification problem, *Pattern Recognition* **36**(1) (2003), 45–60.

[15] M.S. Paterson and F.F. Yao, On nearest neighbor graphs, in: *Proceedings of 19^{th} Int. Coll. Automata, Languages and Programming, Springer LNCS*, (Vol. 623), 1992, pp. 416–426.

[16] C.E. Priebe, J.G. DeVinney and D.J. Marchette, On the distribution of the domination number of random class catch cover digraphs, *Statistics and Probability Letters* **55** (2001), 239–246.

[17] C.E. Priebe, D.J. Marchette, J. DeVinney and D. Socolinsky, Classification using class cover catch digraphs, *Journal of Classification* **20**(1) (2003), 3–23.
[18] C.E. Priebe, J.L. Solka, D.J. Marchette and B.T. Clark, Class cover catch digraphs for latent class discovery in gene expression monitoring by DNA microarrays, *Computational Statistics and Data Analysis on Visualization* **43-4** (2003), 621–632.
[19] G.T. Toussaint, The relative neighborhood graph of a finite planar set, *Pattern Recognition* (1980), 12.
[20] D.B. West, *Introduction to Graph Theory, 2nd ed*, Prentice Hall, NJ, 2001.