Environ Ecol Stat (2017) 24:69–108
DOI 10.1007/s10651-016-0361-z

Nearest neighbor methods for testing reflexivity

Elvan Ceyhan[1] · **Selim Bahadır**[2]

Received: 15 July 2015 / Revised: 19 October 2016 / Published online: 5 November 2016

Abstract Nearest neighbor (NN) methods are widely employed for drawing inferences about spatial point patterns of two or more classes. We introduce a method for testing reflexivity in the NN structure (i.e., NN reflexivity) based on a contingency table which will be called reflexivity contingency table (RCT) henceforth. The RCT is based on the NN relationships among the data points and was used for testing niche specificity in literature, but we demonstrate that it is actually more appropriate for testing the NN reflexivity pattern. We derive the asymptotic distribution of the entries of the RCT under random labeling and introduce tests of reflexivity based on these entries. We also consider Pielou's approach on RCT and show that it is not appropriate for completely mapped spatial data. We determine the appropriate null hypotheses and the underlying conditions/assumptions required for all tests considered. We investigate the finite sample performance of the tests in terms of empirical size and power by extensive Monte Carlo simulations and illustrate the methods on two real-life ecological data sets.

Keywords Association · Completely mapped data · Complete spatial randomness · Habitat/niche specificity · Independence · Random labeling · Segregation · Sparse sampling

Handling Editor: Pierre Dutilleul.

✉ Elvan Ceyhan
elvanceyhan@gmail.com

1 Department of Statistics, University of Pittsburgh, Pittsburgh, PA 15260, USA

2 Department of Mathematics, Koç University, 34450 Sarıyer, Istanbul, Turkey

1 Introduction

The spatial point patterns in natural populations of one or multiple classes have received considerable attention in statistical literature. In this article, the term “class” refers to species or any other characteristic of the subjects such as sex, age group, health condition, etc. Two frequently studied spatial patterns between multiple classes or species are segregation and association (Dixon 2002a). In our setting, the data consists of the locations of events and the mark (e.g., species) at each location. Usually, these data are completely mapped and are from a marked spatial point pattern, although other sampling schemes are possible. Previously, the patterns on this type of data were analyzed by using a contingency table called nearest neighbor contingency table (NNCT) based on the class labels (i.e., marks) of the NN pairs (see, e.g., Dixon 1994, 2002a; Ceyhan 2010, 2008). Here we construct another contingency table, called *reflexivity contingency table* (RCT), by classifying NN pairs according to whether they are reflexive or not and whether they are of the same class or not. That is, a NNCT is obtained by cross-classifying NN pairs according to their class labels, whereas RCT is constructed by cross-classifying NN pairs according to reflexivity and type of NN pairs (as self or mixed). Indeed, NN reflexivity pattern and tests are inspired by one of Pielou’s tests which was introduced for testing niche specificity (Pielou 1961), but the specifics of Pielou’s analysis of an RCT (e.g., the use of the usual Pearson’s test of independence on it) are not correct for completely mapped data. In this manuscript, we develop the correct sampling distribution of the entries in the RCT based on completely mapped data and demonstrate that this contingency table intended for testing niche specificity is more appropriate for testing reflexivity in the NN structure, hence the name *reflexivity contingency table*.

A pair of points (p_1, p_2) is called a *base-NN pair*, if p_2 is a NN of p_1 where p_1 is the *base point* and p_2 is the *NN point*. In the base-NN pair (p_1, p_2), we also write $NN(p_1) = p_2$. *NN reflexivity* occurs if for a pair of points (p_1, p_2), p_1 is a NN of p_2 and p_2 is a NN of p_1. NN reflexivity is a type of spatial interaction (i.e., interdependence) of points from the same or different classes. A base-NN pair is called a *reflexive pair*, if the elements of the pair are NN to each other; a *non-reflexive pair*, if the elements of the pair are not NN to each other; a *self pair*, if the elements of the pair are from the same class; a *mixed pair*, if the elements of the pair are from different classes. Thus, NN reflexivity patterns for multiple classes are of four types: self-reflexivity (p_1 and p_2 are NNs to each other and are from the same class), mixed-reflexivity (p_1 and p_2 are NNs to each other but are from different classes), self-nonreflexivity and mixed-nonreflexivity which are defined similarly. The reflexivity in the NN structure is also referred to as “NN reflexivity ” or for brevity “reflexivity” henceforth. In NN reflexivity, the members in the NN pair exhibit a spatially symmetric interdependence (i.e., one subject occurs closest to the other and vice versa in a particular environment). In the self-reflexivity pattern, conspecifics or members of the same class exhibit such interdependence, while in the mixed-reflexivity pattern, members from different classes exhibit such interdependence (as, e.g., in mutualistic symbiosis). Unless stated otherwise, NN relationships are based on the usual Euclidean distance throughout the article.

NN reflexivity is closely related to spatial interaction patterns like niche specificity, segregation and mutualism. The pattern of niche specificity is much broader than self-NN reflexivity. *Niche/habitat specificity* is the collection of biotic and abiotic conditions favoring the development and hence the existence and abundance of a species on a spatial scale (Ranker and Haufler 2008). That is, niche specificity is the dependence of an organism on an environment (i.e., niche or habitat). Niche specificity can be determined by tolerance to various factors such as climate, exposure to light, soil and nutrient properties and so on (Lindenmayer and Burgman 2005). NN self-reflexivity can be viewed as a special form of spatial niche specificity (i.e., Pielou's approach was not entirely missing the point). Self-reflexivity might be a result of niche specificity or segregation. On the other hand, mixed-reflexivity might be a result of association or interspecific interactions like mutualism or beneficial symbiosis.

There are many methods available for testing various types of spatial patterns in literature. These methods include tests of segregation (Pielou 1961), nearest neighbor (NN) methods (Dixon 2002b), K-function (Ripley 2004), J-function (Lieshout and Baddeley 1999) and so on. An extensive survey for the tests of spatial point patterns is provided by Kulldorff (2006) who categorized and compared more than 100 such tests. These tests are for testing spatial clustering in a one-class setting or testing segregation of points in a multi-class setting. Most of the tests for multiple classes deal with presence or lack of spatial interaction usually in the form of spatial segregation or association between the classes. The second order methods such as K-function, J-function, pair-correlation function, and mark-connection function $p_{12}(r)$ (Ripley 2004; Stoyan and Stoyan 1994; Illian et al. 2008, respectively) are used for testing bivariate spatial interaction at various scales, but K and J functions are unreliable at large scales and pair- and mark-correlation functions are unreliable at small scales. We use the NN relationships for testing spatial pattern of reflexivity. The scale of NN distances is data-dependent; that is, NN distances are of smaller (resp. larger) scale for a given data set when the class sizes are large (resp. small). Hence NN methods provide valuable information about the pattern in addition to the information provided by the above second-order methods. Furthermore, NN relations, including reflexivity, are studied by many authors (see, e.g., Henze 1987; Cox 1981; Schilling 1986). But, these references almost exclusively pertain to the NN patterns in a one class setting. The number of reflexive pairs is also of interest in the NNCT tests as variances and covariances of the cell counts in the NNCT depend on it (Dixon 1994). The approach in this article explicitly requires a multi-class setting; moreover, to the best of authors' knowledge, the proposed test of reflexivity in the NN structure is the only method available in literature for assessing such a pattern.

We investigate the underlying assumptions for the less known—hence less applied compared to tests of segregation—tests of NN reflexivity and Pielou's approach on RCT (i.e., the use of Pearson's test of independence on RCT). We show why Pielou's approach on RCT is not appropriate for completely mapped spatial data. We derive the asymptotic normality for the entries of the RCT for completely mapped data under RL, hence propose Z-tests for the diagonal entries in the RCT and then show joint normality of the diagonal entries of the RCT thereby introducing an overall χ^2 test of NN reflexivity combining the Z-tests. Finite sample empirical size and power

Table 1 A list of abbreviations used in the article

CSR	Complete spatial randomness
NN	Nearest neighbor
NNCT	Nearest neighbor contingency table
RCT	Reflexivity contingency table
RL	Random labeling

comparisons are performed by Monte Carlo simulations. A list of abbreviations used in the article is provided in Table 1.

We describe and discuss the patterns of reflexivity and related patterns together with the appropriate null patterns for them in Sect. 2. We discuss the tests for RCT (namely, the newly introduced tests of reflexivity and Pielou's approach), provide the asymptotic distribution for the cell counts (i.e., entries) of the RCT in Sect. 3. We prove consistency of the tests in Sect. 4, and provide an extensive empirical size and power analysis by Monte Carlo simulations in Sect. 5. We illustrate the methodology on two ecological data sets in Sect. 6 and provide discussion and guidelines for using the tests in Sect. 7.

2 NN reflexivity and related patterns

NN reflexivity, segregation/association and niche specificity are related but different patterns. In this article, we only consider multivariate spatial patterns which are basically concerned with the (spatial) interaction between two or more classes of points. For multivariate spatial data analysis, the benchmark pattern is usually complete spatial randomness (CSR) independence or random labeling (RL) (Diggle 2003) depending on the context. Under CSR independence, the points from each class are independently uniformly distributed in the region of interest conditioned on the class sizes. That is, the points from each class are independent realizations of Homogeneous Poisson Process (HPP) with fixed class sizes (i.e., from the uniform binomial process). However we call independence with uniform binomial processes as (restricted) CSR independence in this article. On the other hand, under RL, class labels are independently and randomly assigned to a set of given locations which could be a realization from any pattern such as HPP or some clustered or regular pattern. There are two major forms of deviation from these benchmark patterns in the multivariate spatial pattern analysis. These interaction patterns are segregation and association. Under segregation, the members of a class or species enjoy the company of the conspecifics, hence form one class clumps or clusters, while under association they tend to coexist with members of other class(es) and form mixed clumps or clusters (see, e.g., Ceyhan 2008 for more detail). Hence, under segregation, it is more conceivable to have the entries in the self column (i.e., the self-reflexive and self-nonreflexive counts) to be larger than their expected values, as segregation does not impose a restriction on reflexivity, but implies an abundance of self NN pairs. Similarly, under association, the entries in the mixed column (i.e., mixed-reflexive and mixed-nonreflexive counts) are usually larger than their expected values. NN reflexivity pertains to the pair types as self or mixed among the reflexive base-NN pairs. In particular, self-reflexivity in NN structure can account for segregation and so can niche specificity. But there is an important distinction between the two patterns. Niche specificity is a biological cause of where a species occurs (i.e., the locations) while reflexivity is more of a consequence of locations, not a biological cause of the locations. However, niche specificity and self-reflexivity in the NN structure are not mutually exclusive, in the sense that they can coexist under a segregation pattern. In the presence of segregation, if the niches of the classes are about the same, then self-reflexivity in NN structure accounts more for segregation. If the niches of the classes are considerably different, niche specificity accounts for segregation, but still self-reflexivity in NN structure might partially account for segregation. On the other hand, mixed-reflexivity in the NN structure might imply association in the form of, e.g., mutualistic symbiosis between the species. Symbiosis is an interaction between species in which there is a close physical contact during most of lives of both participants in the form of physiological connection or integration. Note that the definition makes no statement about direction of interaction, which may be mutualistic, parasitic, or commensalistic. See, e.g., Freeman (2002). Each type of these symbiosis patterns can be viewed as a factor causing association between the species. Both CSR independence and RL patterns imply that self-reflexivity or mixed-nonreflexivity in the NN structure exist at the expected levels.

As the reflexivity, segregation/association and niche specificity patterns are different, the corresponding null hypotheses for them are also different. For segregation/association alternatives, the null case is that there is some sort of randomness in the spatial pattern as in random labeling (RL) or independence. However, for independence, we will only consider a restricted type of complete spatial randomness (CSR) independence with fixed class sizes. In particular, this null hypothesis follows provided that there is randomness in the NN structure in such a way that the probability of a NN of a point being from a class is proportional to the relative frequency of that class. This assumption holds, e.g., under RL or CSR independence of the points from each class. Both CSR independence and RL patterns imply that self-reflexivity or mixed-nonreflexivity in the NN structure exists at its expected levels. In fact, it is conceivable that other independence patterns (in which all classes are independently generated from the same process or distribution) can yield the same null hypotheses, but we restrict our attention to the (restricted) CSR independence or RL as they are commonly considered to be the benchmark patterns in spatial data analysis. The null case for the niche specificity is that there is no relation between the spatial distribution of a class/species and its niche or habitat, and the null case for NN reflexivity is that values of self-reflexive and mixed-nonreflexive pairs are as expected under CSR independence or RL (these expected values will be explicitly provided in Sect. 3.1 below).

For $m = 2$ classes, we label the classes as X and Y (or interchangeably 1 and 2, respectively). Let $\mathcal{X}_{n_1}$ be a data set of size n_1 from class X and $\mathcal{Y}_{n_2}$ be a data set of size n_2 from class Y. Then under CSR independence, we have $\mathcal{X}_{n_1} = \{X_1, X_2, \ldots, X_{n_1}\}$ and $\mathcal{Y}_{n_2} = \{Y_1, Y_2, \ldots, Y_{n_2}\}$ which are independent random samples from $\mathcal{U}(S)$, the uniform distribution on the common support $S \subset \mathbb{R}^d$ for classes X and Y. Unless stated otherwise, for simplicity and practical purposes, we take $d = 2$ throughout the article. We combine $\mathcal{X}_{n_1}$ and $\mathcal{Y}_{n_2}$ into one data set together with the labels of the points and obtain $\mathcal{Z}_n = \{(Z_1, L_1), (Z_2, L_2), \ldots, (Z_n, L_n)\}$ where $n = n_1 + n_2$ and L_i is in $\{0, 1\}$ or $\{X, Y\}$ which are the class labels. Notice that under CSR independence, the randomness is in the locations of the points for each class and the class label is a fixed characteristic of the subject that occupies the point. Under the RL pattern, the class labels or marks are assigned randomly to points whose locations are given. The spatial pattern generating these point locations for RL is referred to as the *background pattern* henceforth. n_1 (resp. n_2) of these background points are assigned as class X (resp. Y) randomly; i.e., the labels L_i are 1 or X with probability approximately n_1/n (resp. 2 or Y with probability approximately n_2/n) independently for $i = 1, 2, \ldots, n$.

The distinction between CSR independence and RL could be very important in practice. Under CSR independence the (locations of the) points from two classes are *a priori* the result of different processes (for instance, individuals of different species or age cohorts). On the other hand, under RL, some processes affect the individuals of a single population *a posteriori* (for instance, diseased versus non-diseased individuals of a single plant species) (Goreaud and Pélissier 2003).

3 Reflexivity contingency table and the associated tests

Recall that RCT is based on the cross-tabulation of the points with respect to NN reflexivity and pair type. The resultant categories are (*self,reflexive*), (*mixed,reflexive*), (*self,non-reflexive*), and (*mixed,non-reflexive*) pairs. Hence the NN reflexivity patterns are essentially of two types: self-reflexivity or mixed-nonreflexivity. We ignore the patterns of mixed-reflexivity and self-nonreflexivity under RL, since these patterns are just the opposites of self-reflexivity and mixed-nonreflexivity, respectively. That is, under RL, self-reflexivity occurs if and only if there is lack of mixed-reflexivity and mixed-nonreflexivity occurs if and only if there is lack of self-nonreflexivity.

The patterns of self-reflexivity and mixed-nonreflexivity in the NN structure can be tested by using the RCT. Let $N_{s,r}$ be the observed number of self-reflexive pairs, $N_{s,nr}$ be the observed number of self-nonreflexive pairs, $N_{m,r}$ be the observed number of mixed-reflexive pairs, and $N_{m,nr}$ be the observed number of mixed-nonreflexive pairs. With the partitioning of base-NN pairs according to NN reflexivity and pair type as self or mixed, we obtain a 2×2 RCT. See also Table 2 where the column sum C_s is the number of self pairs, and C_m is the number of mixed pairs, while the row sum N_r is the number of reflexive pairs, and N_{nr} is the number of nonreflexive pairs.

Table 2 The contingency table for self-reflexivity or mixed-nonreflexivity in the NN structure, i.e., the RCT

	Pair type		Total
	Self	Mixed	
NN reflexivity			
Reflexive	$N_{s,r}$	$N_{m,r}$	N_r
Non-reflexive	$N_{s,nr}$	$N_{m,nr}$	N_{nr}
Total	C_s	C_m	n

We write the entries in the RCT as linear combinations of indicator random variables. Under RL, let

$$r_{ij} = \begin{cases} 1 & \text{if } z_i \text{ and } z_j \text{ are reflexive NNs} \\ 0 & \text{otherwise ;} \end{cases}$$

and for every $1 \leq u \leq m$, let

$$z_{ij}^u = \begin{cases} 1 & \text{if } L_i = L_j = u, \\ 0 & \text{otherwise .} \end{cases}$$

Then, letting $N_r^u = \sum_{i=1}^n \sum_{j=1}^n r_{ij} z_{ij}^u$, it is easy to see that

$$N_{s,r} = \sum_{u=1}^m N_r^u.$$

For every $1 \leq a \neq b \leq m$, let

$$z_{ij}^{ab} = \begin{cases} 1 & \text{if } L_i = a \text{ and } L_j = b, \\ 0 & \text{otherwise ,} \end{cases}$$

then $N_r^{ab} = \sum_{i=1}^n \sum_{j=1}^n r_{ij} z_{ij}^{ab}$. So

$$N_{m,r} = \sum_{1 \leq a \neq b \leq m} N_r^{ab}.$$

Let

$$y_{ij} = \begin{cases} 1 & \text{if } NN(z_i) = z_j \text{ and } NN(z_j) \neq z_i, \\ 0 & \text{otherwise .} \end{cases}$$

and $N_{nr}^u = \sum_{i=1}^n \sum_{j=1}^n (y_{ij} + y_{ji}) z_{ij}^u$, then it follows that

$$N_{s,nr} = \sum_{u=1}^m N_{nr}^u.$$

Finally, letting $N_{nr}^{ab} = \sum_{i=1}^n \sum_{j=1}^n y_{ij} z_{ij}^{ab}$, we get

$$N_{m,nr} = \sum_{1 \leq a \neq b \leq m} N_{nr}^{ab}.$$

Notice also that $\sum_{1 \leq i,j \leq n} y_{ij} = n - R$. Moreover, we have $N_r = R = \sum_{i=1}^n \sum_{j=1}^n r_{ij}$ which is the number of ordered reflexive NNs (Dixon 1994), and so $N_{nr} = n - R$.

Remark 3.1 (Ties in the NN structure) If $\mathcal{X}_n$ is a random sample of size n from a continuous distribution in its support (e.g., under CSR independence), then each point has one NN a.s. However, under RL, it is possible to have ties in the NN relations; i.e., a point might have multiple NNs. In this case, the ties should be broken and corresponding counts should be modified so as to take the actual NN relations into account and thus avoid information loss and to keep $N_r + N_{nr} = n$. Along this line, we let N_i^{nn} be the number of NNs of point Z_i, then any quantity (implicitly) involving the quantity $\mathbf{I}(Z_j \text{ is a NN of } Z_i)$ should be scaled by $1/N_i^{nn}$ where $\mathbf{I}()$ stands for the indicator function. For example, without ties in the NN structure, $r_{ij} = \mathbf{I}(z_j \text{ is a NN of } z_i)\mathbf{I}(z_i \text{ is a NN of } z_j)$; but in the presence of ties, it is adjusted as $r_{ij} = (1/N_i^{nn})\mathbf{I}(z_j \text{ is a NN of } z_i)(1/N_j^{nn})\mathbf{I}(z_i \text{ is a NN of } z_j)$. Similarly, y_{ij} can be adjusted as $y_{ij} = (1/N_i^{nn})\mathbf{I}(z_j \text{ is a NN of } z_i)\mathbf{I}(z_i \text{ is not a NN of } z_j)$.

3.1 Sampling distribution of the entries of the RCT under RL

Let z_i, z_j, z_k, z_l be four distinct points from the sample (i.e., distinct events) and u, v, w, t be not necessarily distinct class labels. Define the probabilities

$$\begin{aligned} p_{uv} &= P(L_i = u, L_j = v), \\ p_{uvw} &= P(L_i = u, L_j = v, L_k = w), \\ p_{uvwt} &= P(L_i = u, L_j = v, L_k = w, L_l = t). \end{aligned}$$

In particular, for pairwise distinct class labels a, b, c, d, we have

$$\begin{aligned} p_{aa} &= \frac{n_a(n_a-1)}{n(n-1)} \text{ and } p_{ab} = \frac{n_a n_b}{n(n-1)}; \\ p_{aaa} &= \frac{n_a(n_a-1)(n_a-2)}{n(n-1)(n-2)},\ p_{aab} = \frac{n_a(n_a-1)n_b}{n(n-1)(n-2)}, \text{ and} \\ p_{abc} &= \frac{n_a n_b n_c}{n(n-1)(n-2)}; \\ p_{aaaa} &= \frac{n_a(n_a-1)(n_a-2)(n_a-3)}{n(n-1)(n-2)(n-3)},\ p_{aaab} = \frac{n_a(n_a-1)(n_a-2)n_b}{n(n-1)(n-2)(n-3)}, \\ p_{aabb} &= \frac{n_a(n_a-1)n_b(n_b-1)}{n(n-1)(n-2)(n-3)}, \\ p_{aabc} &= \frac{n_a(n_a-1)n_b n_c}{n(n-1)(n-2)(n-3)}, \text{ and } p_{abcd} = \frac{n_a n_b n_c n_d}{n(n-1)(n-2)(n-3)}. \end{aligned}$$

Notice that the probabilities above are independent of i, j, k, l, and invariant under any permutation of the indices; i.e., $p_{ab} = p_{ba}$, $p_{aab} = p_{aba} = p_{baa}$ and so on. Also, let

$$P_{aa} = \sum_{a=1}^{m} p_{aa},\ P_{ab} = \sum_{1 \le a \ne b \le m} p_{ab},\ P_{aab} = \sum_{1 \le a \ne b \le m} p_{aab},\ P_{abc} = \sum_{1 \le a \ne b \ne c \le m} p_{abc},$$

$$P_{aaaa} = \sum_{a=1}^{m} p_{aaaa},\ P_{aaab} = \sum_{1 \le a \ne b \le m} p_{aaab},\ P_{aabb} = \sum_{1 \le a \ne b \le m} p_{aabb},$$

$$P_{aabc} = \sum_{1 \le a \ne b \ne c \le m} p_{aabc},\ P_{abcd} = \sum_{1 \le a \ne b \ne c \ne d \le m} p_{abcd},$$

where $a \neq b \neq c$ means that a, b, c are pairwise distinct, and $a \neq b \neq c \neq d$ means that a, b, c, d are pairwise distinct. Notice that we have $P_{aa} = \frac{(\sum_{u=1}^{m} n_u^2) - n}{n(n-1)}$ and $P_{ab} = \frac{\sum_{u \neq v} n_u n_v}{n(n-1)}$ and P_{aa} is the probability of a reflexive pair being a self pair and P_{ab} is the probability of a non-reflexive pair being a mixed pair. Notice that $P_{aa} + P_{ab} = 1$; and for $m = 2$, we have $P_{aa} = \frac{n_a(n_a-1)}{n(n-1)} + \frac{n_b(n_b-1)}{n(n-1)} = \frac{n_a^2+n_b^2-n}{n(n-1)}$ and $P_{ab} = \frac{2n_a n_b}{n(n-1)}$.

Let Q be the number of shared NNs (i.e., number of triplets (z_i, z_j, z_k) with $NN(z_i) = NN(z_j) = z_k$) and T be the number of triplets (z_i, z_j, z_k) with "$NN(z_i) = NN(z_j) = z_k$ and $NN(z_k) = z_j$". One can construct a NN digraph for a data set $\mathcal{X}_n$ with vertices $\mathcal{X}_n$ and there is an arc from x_i to x_j if $NN(x_i) = x_j$ for $x_i, x_j \in \mathcal{X}_n$. Note that in the NN digraph, $T + R$ is the sum of the indegrees of the points in the reflexive pairs.

For $m \geq 4$ classes, we show that

$$\mathbf{E}[N_{s,r}] = R P_{aa},\ \mathbf{E}[N_{m,nr}] = (n-R) P_{ab}, \tag{1}$$

$$\mathbf{Var}[N_{s,r}] = R^2 \left(P_{aaaa} + P_{aabb} - (P_{aa})^2\right) + 2R\left(P_{aa} - P_{aaaa} - P_{aabb}\right), \tag{2}$$

$$\begin{aligned}\mathbf{Var}[N_{m,nr}] = {} & (n-R)^2(2P_{aabb} + 4P_{aabc} + P_{abcd} - (P_{ab})^2) + (n-R)P_{ab} \\ & + (2n - 2R + Q - 4T)(P_{aab} + P_{abc}) \\ & + (-3n + 3R - Q + 4T)(2P_{aabb} + 4P_{aabc} + P_{abcd}),\end{aligned} \tag{3}$$

$$\begin{aligned}\mathbf{Cov}[N_{s,r}, N_{m,nr}] = {} & R(n-R)(2P_{aaab} + P_{aabc} - P_{aa}P_{ab}) \\ & + 2T(P_{aab} - 2P_{aaab} - P_{aabc}).\end{aligned} \tag{4}$$

See Appendix for the derivation of these quantities under RL. Notice that for the case $m \leq 3$, one can consider a null or void class (i.e. a class with no elements) or classes as if they exist no elements to make the number of classes equal to 4; hence in the equations above, the sum P_{abcd} vanishes for $m = 3$, and the sums P_{abcd}, P_{abc} and P_{aabc} vanish for $m = 2$.

In the self-reflexivity (resp. mixed-nonreflexivity) pattern, self-reflexive (resp. mixed-nonreflexive) pairs are more frequent than expected under RL. The other patterns of mixed-reflexivity and self-nonreflexivity are defined similarly. The null case for NN reflexivity is that values of self-reflexive and mixed-nonreflexive pairs are as expected under CSR independence or RL (these expected values are explicitly provided below). Under RL, the total number of reflexive and non-reflexive base-NN pairs are fixed quantities.

Then, for the NN reflexivity tests, we have H_o : "The entries of the RCT are equal to their expected values" or equivalently,

$$H_o: \mathbf{E}[N_{s,r}] = RP_{aa} \quad \text{and} \quad \mathbf{E}[N_{m,nr}] = (n-R)P_{ab} \tag{5}$$

as our null hypothesis. The other entries are omitted, since the row sums are fixed under RL. That is, $N_{m,r} = n_r - N_{s,r}$ and $N_{s,nr} = n_{nr} - N_{m,nr}$, and so if (5) holds, the other entries are also equal to their expected values. The alternative hypotheses for self-reflexivity and mixed-nonreflexivity in the NN structure are

$$H_a: \mathbf{E}[N_{s,r}] > RP_{aa} \text{ and } H_a: \mathbf{E}[N_{m,nr}] > (n-R)P_{ab},$$

respectively. We also show that the joint distribution of all $N_{s,r}$ and $N_{m,nr}$ is asymptotically bivariate normal under RL where Q, R and T are fixed quantities. In our analysis, we assume that $R/n \to r$, $Q/n \to q$, $T/n \to t$ and $n_i/n \to \lambda_i$ for all $1 \le i \le m$, as $n \to \infty$ (see Appendix A4 for the derivations).

Then, the test statistics for self-reflexivity and mixed-nonreflexivity,

$$Z_{s,r} = \frac{N_{s,r} - \mathbf{E}[N_{s,r}]}{\sqrt{\mathbf{Var}[N_{s,r}]}} \quad \text{and} \quad Z_{m,nr} = \frac{N_{m,nr} - \mathbf{E}[N_{m,nr}]}{\sqrt{\mathbf{Var}[N_{m,nr}]}}, \tag{6}$$

approximately have $N(0,1)$ distribution for large large n_r and n_{nr}, respectively. Notice that under RL, we have $Z_{s,r} = -Z_{m,r}$ since $N_{m,r} = N_r - N_{s,r}$ where N_r is fixed and $Z_{m,nr} = -Z_{s,nr}$ since $N_{s,nr} = N_{nr} - N_{m,nr}$ where N_{nr} is fixed. We can combine the test statistics $Z_{s,r}$ and $Z_{m,nr}$ into an overall test of reflexivity as follows. Let $\mathbf{V_{ref}} = (N_{s,r}, N_{m,nr})^t$ and $\mathbf{E}[\mathbf{V_{ref}}] = (\mathbf{E}[N_{s,r}], \mathbf{E}[N_{m,nr}])^t$, and Σ_{ref} be the variance-covariance matrix of $\mathbf{V_{ref}}$, then $\mathbf{V_{ref}} \xrightarrow{\mathcal{L}} N(\mathbf{E}[\mathbf{V_{ref}}], \Sigma_{ref})$ as n_r and n_{nr} tend to infinity. So, it follows that the quadratic form

$$\mathcal{X}_R^2 = (\mathbf{V_{ref}} - \mathbf{E}[\mathbf{V_{ref}}])^t \Sigma_{ref}^{-1} (\mathbf{V_{ref}} - \mathbf{E}[\mathbf{V_{ref}}]) \tag{7}$$

has asymptotically χ_2^2 distribution as both R and $n - R$ are tending to infinity (Moser 1996).

Here alternatively, for the overall test, we could also combine the entries of the RCT and obtain an overall test of reflexivity as follows. Let

$$\widetilde{\mathbf{V}}_{\mathbf{ref}} = (N_{s,r}, N_{m,r}, N_{m,r}, N_{m,nr})^t$$

and

$$\mathbf{E}\left[\widetilde{\mathbf{V}}_{\mathbf{ref}}\right] = \left(\mathbf{E}[N_{s,r}], \mathbf{E}[N_{m,r}], \mathbf{E}[N_{m,r}], \mathbf{E}[N_{m,nr}]\right)^t,$$

and $\widetilde{\Sigma}_{ref}$ be the variance-covariance matrix of $\widetilde{\mathbf{V}}_{\mathbf{ref}}$, then one can get the following quadratic form

$$\widetilde{\mathcal{X}}_R^2 = \left(\widetilde{\mathbf{V}}_{\mathbf{ref}} - \mathbf{E}\left[\widetilde{\mathbf{V}}_{\mathbf{ref}}\right]\right)^t \widetilde{\Sigma}_{ref}^- \left(\widetilde{\mathbf{V}}_{\mathbf{ref}} - \mathbf{E}\left[\widetilde{\mathbf{V}}_{\mathbf{ref}}\right]\right) \quad (8)$$

where $\widetilde{\Sigma}_{ref}^-$ is the generalized inverse of $\widetilde{\Sigma}_{ref}$ (Searle 2006). However, under RL, $N_{m,r} = R - N_{s,r}$ and $N_{m,r} = n - R - N_{m,nr}$, where R is a fixed quantity. Hence all the information in the RCT is contained in $\mathbf{V}_{\mathbf{ref}}$, so $\widetilde{\Sigma}_{ref}$ is of rank 2 (hence the use of generalized inverse in the quadratic form above). Therefore, this version of the overall reflexivity test, $\widetilde{\mathcal{X}}_R^2$, also has asymptotically χ_2^2 distribution as both R and $n - R$ are tending to infinity (Moser 1996). Indeed, in practice one would observe that $\widetilde{\mathcal{X}}_R^2 = \mathcal{X}_R^2$, so for convenience, we only use the overall test provided in Eq. (7) in the rest of the article.

When $\mathcal{X}_R^2$ is significant, it would only imply a significant deviation from the NN reflexivity structure under H_o in Eq. (5). But the two types of alternatives of self-reflexivity or mixed-nonreflexivity in the NN structure are not distinguishable by the usual χ^2 test. To determine the direction of this deviation, one can use both of the Z-tests, $Z_{s,r}$ and $Z_{m,nr}$, for the left- and right-sided alternatives.

Remark 3.2 (Status of Q, R, and T under RL and CSR independence) The quantities Q, R and T are fixed under RL, while they are random under CSR independence. Hence the tests in Eqs. (6) and (7) are conditional on the observed values of Q, R and T under CSR independence while no such conditioning is required under RL. The quantities given in Eqs. (1)–(4), and all the quantities depending on them also depend on Q, R, and T. Hence these expressions are appropriate under the RL pattern. Under the CSR independence pattern, they are conditional on Q, R, and T, as these are random quantities. But under CSR independence, it is conceivable that they converge in probability to their expected values and hence, by an appropriate application of Slutsky's Theorem, one can replace Q, R, and T with their expectations and can still use asymptotic normal approximation for the test statistics. However, given the difficulty of calculating the expectations of Q, R, and T for spatial data under CSR independence, it is convenient and reasonable to use the observed values of Q, R, and T as plug-in estimators even when assessing their behavior under CSR independence. Alternatively, one can estimate the expected values of Q, R, and T empirically and substitute these estimates in the expressions. For example, for the uniform binomial process on the unit square, we have $\mathbf{E}[Q/n] \approx .6324$, $\mathbf{E}[R/n] \approx 0.6219$, and $\mathbf{E}[T/n] \approx 0.2753$. (estimated empirically based on 10000 Monte Carlo simulations for various values of n). When Q, R, and T are replaced by $0.63\,n$, $0.62\,n$, and $0.275\,n$, respectively, we obtain the so-called *QRT-adjusted* reflexivity tests.

3.2 Pielou's test for the reflexivity contingency table

We provide a detailed treatment of Pielou's approach on RCT to determine when it is appropriate to use and why it fails for completely mapped spatial data. Pielou (1961) uses the usual χ^2 test of independence on the RCT in order to detect presence of niche specificity, but RCT is by construction more appropriate for testing the spatial pattern of NN reflexivity. Moreover, Pielou's test on the RCT is used to detect independence between the row and column categories (i.e., NN reflexivity of the pairs and pair type as self or mixed). Such independence would imply

$$\mathbf{E}\left[N_{s,r}\right] = C_s N_r/n, \quad \mathbf{E}\left[N_{m,r}\right] = C_m N_r/n,$$
$$\mathbf{E}\left[N_{s,nr}\right] = C_s N_{nr}/n, \text{ and } \mathbf{E}\left[N_{m,nr}\right] = C_m N_{nr}/n. \tag{9}$$

Hence an excess of $N_{s,r}$ from its expected value above would imply a positive dependence; on the other hand, if $N_{s,r}$ is less than expected, it would imply a negative dependence between a pair being reflexive and self pair. The deviations of other entries from their expected values have similar interpretations. Therefore, in Pielou's approach, the actual null hypothesis is

$\mathbf{H_o}$ There is independence between row and column categories (i.e., independence of "NN reflexivity" and "pair type as self or mixed"). (10)

For the RCT, recall that Pielou (1961) suggests the use of Pearson's usual χ^2 test with 1 df,

$$\mathcal{X}_P^2 = \frac{(N_{s,r} - C_s N_r/n)^2}{C_s N_r/n} + \frac{(N_{m,r} - C_m N_r/n)^2}{C_m N_r/n} + \frac{(N_{s,nr} - C_s N_{nr}/n)^2}{C_s N_{nr}/n} + \frac{(N_{m,nr} - C_m N_{nr}/n)^2}{C_m N_{nr}/n} \tag{11}$$

and H_o is rejected when the p-value based on this χ^2 test is significant. Equivalently, the independence between the NN reflexivity and pair type can also be tested using the directional test statistic

$$Z_{dir} = \left(\frac{N_{s,r}}{N_r} - \frac{N_{s,nr}}{N_{nr}}\right)\sqrt{\frac{N_r\, N_{nr}\, n}{C_s\, C_m}}. \tag{12}$$

Notice that the tests in Eqs. (11) and (12) are used to test the same null hypothesis with the same underlying assumptions, but the former is for the two-sided alternative only, while the latter can be used for both two- and one-sided alternatives.

Under positive (resp. negative) dependence, we expect $Z_{dir} > 0$ (resp. $Z_{dir} < 0$). Under the usual row-wise multinomial framework, for large n, Z_{dir} approximately has a $N(0, 1)$ distribution. Thus for the negative dependence alternative, H_o is rejected when $Z_{dir} < z_\alpha$ where z_α is the $100\alpha^{th}$ percentile of the standard normal distribution; and for the positive dependence alternative, H_o is rejected when $Z_{dir} > z_{1-\alpha}$. The two types of deviations from independence are not distinguishable by the usual χ^2 test. Hence one can resort to the directed test statistic, Z_{dir}, for this purpose.

3.3 Problems with Pielou's approach

The *niche* (or *habitat*) of a species might have an impact on or account for the existence of segregation. If (spatial) niche specificity is operating, among the reflexive pairs, self pairs will be more frequent than mixed pairs (Pielou 1961). But this does not necessarily imply that the entries in the RCT would be significantly different from their expected frequencies under RL. Pielou described a test based on the RCT and suggested the use of Pearson's usual χ^2 test of independence (hence, implicitly the corresponding one-sided directed tests). However, a class can be restricted to a niche (i.e., can have niche specificity), but still the self-reflexive pairs can be similar to their expected frequencies under RL. Additionally, independence of row and column categories in a RCT does not necessarily imply that RCT has the cell counts at the expected levels of self-reflexivity or mixed-nonreflexivity. Therefore RCT is more useful to test the existence of self-reflexivity or mixed-nonreflexivity in the NN structure, rather than niche specificity.

In general, a contingency table may result from four sampling models: (i) Poisson sampling model: cell counts are independent Poisson random variables, and hence all margins and the overall sum are all random quantities; (ii) (Overall) Multinomial sampling model: the total sample size n is fixed but row and column sums are random, so, the cell counts are from a multinomial distribution; (iii) Independent multinomial sampling or product multinomial sampling or row-wise multinomial sampling: row totals are assumed fixed and each row (independent of other rows) is from the same multinomial distribution; and (iv) Hypergeometric sampling: both row and column sums are assumed to be fixed. In the 2×2 contingency table, the rows will have two entries, so the row-wise multinomial distribution reduces to a binomial distribution. If the data set for each class independently comes from a HPP, then Poisson framework would be the most appropriate. But we assume the overall sample size is fixed which renders the Poisson framework inappropriate in our setting. The hypergeometric framework does not seem to fit to a RCT under any reasonable setting as well (see Remark 3.3). In the *overall multinomial framework*, all of the cell counts (viewed as a vector) are assumed to arise from independent multinomial trials. Conditional on the row sums, the overall multinomial framework reduces to the row-wise multinomial framework. In the RCT, row sums N_r and N_{nr} are fixed under RL; hence, under RL, row-wise multinomial framework is more appropriate compared to the overall multinomial framework. Under CSR independence we consider, row sums would be random quantities, so the overall multinomial framework would be more appropriate compared to the row-wise multinomial framework. However, a RCT based on spatial data is unlikely to result from either row-wise or overall multinomial frameworks. In a RCT, a trial is the "categorization of a base-NN pair with respect to NN reflexivity and pair type as self or mixed". In general, in a 2×2 contingency table under the row-wise framework, the entries (N_{11}, N_{12}) and (N_{21}, N_{22}) are assumed to be independent and so are the individual trials. This assumption is invalid when the RCT is based on completely mapped spatial data, because independence between rows is violated. A similar result holds under CSR independence with overall multinomial framework, since again independence between trials is violated. *Thus Pielou's test is influenced* *by deviations not only from the null case in Eq.* (10) *but also from dependence of trials.*

If the RCT is constructed using a random sample of labels of base-NN pairs in terms of NN reflexivity and pair type as self or mixed, then the usual contingency table assumptions under the row-wise multinomial framework would hold. The spatial dependence can not merely be avoided by random sub-sampling but can be circumvent by an appropriate sparse sampling (Diggle 1979). When the data were properly sparsely sampled, we will assume that the RCT satisfies the usual independence assumptions in the row-wise multinomial framework henceforth. In this framework, the explicit form of the null hypothesis is as in Eq. (10). Under CSR independence with sparse sampling, the overall multinomial framework is able to model the RCT, but again this would only be an approximate modeling, because of the inherent correlation between the components or entries of a multinomially distributed random variable. The assessment of various sparse sampling schemes for these tests is a topic of prospective research.

Our suggestion for Pielou's test on the RCT is that if the data is properly sparsely sampled, then it is safe to use. But if the data is completely mapped, to remove the influence of spatial dependence on Pielou's test on RCT, we suggest the usual Monte Carlo randomization where class labels are randomly assigned to the given points a large number of times and test statistics are computed. The corresponding p-value of the test is based on the rank (divided by the number of Monte Carlo replications) of the test statistic of the original data in the sample of test statistics obtained from the Monte Carlo randomization procedure. This suggestion is similar to the one by Meagher and Burdick (1980) except for the fact that in our suggestion, the class sizes are fixed at each Monte Carlo replication, while in Meagher and Burdick (1980), in a two-class setting each point is labeled as a class with probability being equal to proportion of that class size to the total sample size (in the original data) at each Monte Carlo replication, hence in their simulated data class sizes are random as well. Alternatively, if the supports of the classes in an independence setting satisfy the superposition setting (i.e., if the supports are rectangular and the processes generating the data sets are stationary), the tests could be employed with toroidal-shifts as well (Dixon 2002c).

Remark 3.3 (Fisher's exact test for the RCT) For the RCT, one could also consider Fisher's exact test which is frequently used for contingency tables with small cell counts and marginal sums (see Agresti 1992). However, the counts in the reflexivity contingency table are not independent and hence there are two sufficient statistics for the 2×2 RCT, while the sampling distribution for Fisher's exact test (i.e., hypergeometric sampling distribution) has only one. Yet, our simulations suggested that this dependence between the counts is negligible. Hence in the technical report Ceyhan (2014), we investigate the use of Fisher's exact test with the four variants for the one-sided alternative as in Ceyhan (2010) and demonstrate that one of the variants (called the table-inclusive version) has the appropriate significance level while others are liberal in rejecting the null hypothesis. But this variant of the exact test has a relatively low power performance under various alternatives (Ceyhan 2014), so we do not pursue nor present this line of research in this article.

3.4 How to perform the tests for $m > 2$ classes

The extension of the RCT to multi-class case with $m > 2$ is straightforward, since the base-NN pairs in any multi-class data set can be categorized into the four groups as in Table 2 based on the relation between reflexiveness and pair type (as self or mixed). Hence RCT is of dimension 2×2 regardless of the number of classes, m. Although the dimension of the RCT is same for any number of classes, the distribution of the column sums (C_s, C_m) depends on m. In particular, if m gets larger, the likelihood of reflexive NN pairs being mixed increases and hence C_m tends to increase with increasing m. But this will not confound the expected cell counts in the contingency table, since the expected values of the cell counts take the class sizes into account in our approach (in Eq. (5)) and the row and column sums in Pielou's approach (in Eq. (9)). Thus a test of deviation from the expected cell counts in the RCT would not be (substantially) affected by the number of classes in the multi-class case.

In the multi-class case with $m > 2$, we recommend the following strategy for the practical implementation of these tests: First perform an overall omnibus test (as in ANOVA F-test for multi-group comparisons) and then if the omnibus test is significant, then perform post-hoc (or follow-up) tests to determine the specifics of the differences. These post-hoc tests could be pairwise tests (as in the pairwise t-tests) or one-vs-rest tests, where one class is compared with respect to all other classes combined. More specifically, with $m > 2$ classes, in the pairwise comparison, we only restrict our attention to two classes, i and j with $i \neq j$, at a time, and treat the classes as in the two-class case. In the one-vs-rest type of test for class i, we pool the remaining classes and treat them as the "other class" in a two-class setting, hence the name *one-vs-rest test*. In a multi-class setting with m classes, there are m one-vs-rest type tests and $m(m-1)/2$ pairwise tests. As m increases, the first version is computationally less intensive and potentially easier to interpret. For any number of classes and each type of post-hoc test, the RCT is of dimension 2×2, and self-reflexivity and mixed-nonreflexivity relations are defined with different class types.

In a RCT, if all column entries are significant in the same direction, then self or mixed abundance might be operating. In such a case, one can follow up by cell-specific NNCT tests (Dixon 2002a). For example, if the self column entries are significantly larger than their expected values, this implies an abundance of self NN pairs which is indicating segregation of some classes from the others. In order to determine which classes exhibit segregation, one can perform the cell-specific tests for the diagonal entries in the NNCT. On the other hand, if the mixed column entries are significantly larger than their expected values, then mixed pairs abound indicating association of some classes with others. To determine which classes are associated, one can perform cell-specific tests in the off-diagonal entries in the NNCT. The more interesting reflexivity patterns are when entries in the same column are significant in opposite direction or only one of them is significant, which is suggestive of a pattern more intricate than segregation/association.

In all the above cases, the post-hoc tests can give different and seemingly conflicting results; e.g., one class can exhibit self-reflexivity in the NN structure with respect to some other class, while mixed-reflexivity in the NN structure with respect to another class. Thus extra care should be taken about which post-hoc/follow-up test is used and how it is interpreted.

Remark 3.4 (Further partitioning of the entries of RCT) Notice that it is possible to refine the RCT in various ways. In the current form discussed in this article, RCT tests provide information on whether self-reflexivity or mixed-nonreflexivity is present or not. However, each entry can be partitioned into class- or pair-specific forms. For example, we can write $N_{s,r}$ as the sum $N^1_{s,r} + N^2_{s,r} + \cdots + N^m_{s,r}$ where $N^i_{s,r}$ is the number of self-reflexive pairs belonging to class i and test each $N^i_{s,r}$ for deviations from its expected value under the null case or test $(N^1_{s,r}, N^2_{s,r}, \ldots, N^m_{s,r})$ jointly. Similarly, we can write $N_{m,nr} = N^{12}_{m,nr} + N^{13}_{m,nr} + \cdots + N^{(m-1)m}_{m,nr}$ where $N^{ij}_{m,nr}$ is the number of nonreflexive pairs with one member from class i and the other from class j with $i \neq j$ and test each $N^{ij}_{m,nr}$ for deviations from its expected value under the null case. Such a partition might provide further ecological information about the pattern/data in question. However, this approach does not necessarily provide a post-hoc analysis after a significant overall reflexivity test on the RCT. In particular, one might have $N^i_{s,r}$ values significantly deviating from their expected values, but the overall reflexivity test can be insignificant, and vice versa. Hence the study of these class- or pair-specific reflexivity patterns/tests is a topic of ongoing research and deferred to a prospective article.

4 Consistency of tests

The null hypotheses are different for our tests of reflexivity in the NN structure and Pielou's test in Sect. 3.2 (see also Eq. (10)) and so are the alternatives. Hence the comparison of the tests is inappropriate even for large samples; but a reasonable test should have more power as the sample size increases. So, we prove the consistency of the tests in question under appropriate hypotheses. Let $\chi^2_{\nu,\alpha}$ be the $100 \times \alpha$th percentile of χ^2 distribution with ν degrees of freedom.

We first prove the consistency of the self-reflexivity and mixed-nonreflexivity tests and the overall reflexivity tests and then that of Pielou's test (under the corresponding appropriate settings).

Theorem 4.1 *Under RL, we have:*

(i) The self-reflexivity test, $Z_{s,r}$ in Eq. (6), rejecting $H_o : \mathbf{E}[N_{s,r}] = RP_{aa}$ against the two-sided (and one-sided alternatives) for $|Z_{s,r}| > z_{1-\alpha/2}$ (and $Z_{s,r} > z_{1-\alpha}$ or $Z_{s,r} < z_\alpha$) is consistent.

(ii) The mixed-nonreflexivity test, $Z_{m,nr}$ in Eq. (6) rejecting $H_o : \mathbf{E}[N_{m,nr}] = RP_{ab}$ against the two-sided (and one-sided alternatives) for $|Z_{m,nr}| > z_{1-\alpha/2}$ (and $Z_{m,nr} > z_{1-\alpha}$ or $Z_{m,nr} < z_\alpha$) is consistent.

(iii) The overall reflexivity test, $\mathcal{X}^2_R$, in Eq. (7) rejecting $H_o : \mathbf{E}[N_{s,r}] = RP_{aa}$ and $\mathbf{E}[N_{m,nr}] = (n-R)P_{ab}$ against the alternative $H_a : \mathbf{E}[N_{s,r}] \neq RP_{aa}$ or $\mathbf{E}[N_{m,nr}] \neq (n-R)P_{ab}$ for $\mathcal{X}^2_R > \chi^2_{2,1-\alpha}$ is consistent.

Proof (i) Consider the one-sided alternative with H_a : $\mathbf{E}[N_{s,r}] > RP_{aa}$. Under RL, $\mathbf{E}[Z_{s,r}] = 0$ and $\mathbf{Var}[N_{s,r}]$ is given in Eq. (2). Consider the parametrization of the alternative as H_a : $\mathbf{E}[N_{s,r}] > R(P_{aa} + \varepsilon)$ for $\varepsilon \in (0, 1 - P_{aa})$. Then under H_a, $\mathbf{E}[N_{s,r}|H_a] = R\varepsilon > 0$ which implies $\mathbf{E}[Z_{s,r}] > \varepsilon'$ for some $\varepsilon' > 0$. Under both null and alternative hypothesis, $Z_{s,r}$ has asymptotic normal distribution. Then by the standard arguments for the consistency of Z-tests, the test statistic, $Z_{s,r}$, is consistent for the right-sided alternative. The consistency for the other types of alternatives follow similarly.

(ii) The consistency of $Z_{m,nr}$ for the one- and two-sided alternatives follows as in case (i) above.

(iii) Under RL, $\mathbf{Y} := \mathbf{V_{ref}} - \mathbf{E}[\mathbf{V_{ref}}]$ is approximately distributed as $\mathcal{N}(\mathbf{0}, \Sigma)$ for large n where Σ is the variance-covariance matrix of $\mathbf{V_{ref}}$ whose rank is 2. Then by Theorem 3.1.2 of Moser (1996), under H_o, $\mathbf{Y}'\Sigma^{-1}\mathbf{Y} \sim \chi^2_2(\lambda = 0)$. Consider any deviation from H_o. Then under H_a, $\mathbf{E}[\mathbf{Y}|H_a] = \boldsymbol{\mu}_a$ and $\mathbf{Y}$ have multivariate normal distribution with mean $\boldsymbol{\mu}_a$. Then by Theorem 3.1.2 of Moser (1996), under $H_a, \mathbf{Y}'\Sigma^{-1}\mathbf{Y} \sim \chi^2_2(\lambda = \boldsymbol{\mu}_a'\Sigma^{-1}\boldsymbol{\mu}_a)$. Since Σ^{-1} is positive definite and $\boldsymbol{\mu}_a$ is nonzero, the mean of the quadratic form $\mathcal{X}^2_R$ is $2 + \lambda$ with $\lambda > 0$. Then for large n, the null and alternative hypotheses are equivalent to $H_o : \lambda = 0$ and $H_a : \lambda > 0$, respectively. Then by standard arguments for consistency of the χ^2 tests, the desired result follows. □

Theorem 4.2 *Let the RCT be constructed by a random sample of labels of base-NN pairs in terms of NN reflexivity and pair type as self or mixed (or data is obtained by an appropriate sparse sampling) under a row-wise multinomial framework with $m \geq 2$ classes. Then, Pielou's test, $\mathcal{X}^2_P$, for the RCT given in Eq.* (11) *(i.e., the test rejecting independence in the RCT for $\mathcal{X}^2_P > \chi^2_{1,1-\alpha}$) is consistent. The same holds under the overall multinomial framework. The one-sided tests (hence the two-sided test) using Z_{dir} given in Eq.* (12) *are also consistent.*

Proof Under the null hypothesis of independence with the row-wise multinomial framework, as $n \to \infty$, we have $Z_{dir} \sim N(0, 1)$ and Z_{dir} also has a normal distribution under the alternative hypothesis. So under H_o, $\mathbf{E}[Z_{dir}] = 0$ and under H_a : *positive dependence*, we have $\mathbf{E}[Z_{dir}|H_a] = \varepsilon > 0$ and under H_a : *negative dependence*, $\mathbf{E}[Z_{dir}|H_a] = \varepsilon < 0$ where ε is a parameterization of the alternative. Then by the standard arguments for the consistency of Z-tests, the test using Z_{dir} is consistent for the one-sided (hence the two-sided) alternatives. Furthermore, we have $Z^2_{dir} = \mathcal{X}^2_P$. The α-level test based on $\mathcal{X}^2_P$ is equivalent to the α-level two-sided test based on Z_{dir}. Hence the consistency of $\mathcal{X}^2_P$ follows as well. The cases for the overall multinomial framework are similar. □

5 Empirical size and power analysis

In this section we investigate the finite sample behavior of the tests under their appropriate null hypotheses and under various alternatives via Monte Carlo simulations.

5.1 Empirical size analysis

To determine empirical size performance of the tests, we use CSR independence and RL as our null hypotheses. Under these patterns, self-reflexivity or mixed-nonreflexivity in the NN structure would not deviate significantly from their expected behavior. That is, under these cases, NN reflexivity would occur at expected levels: H_o : $\mathbf{E}[N_{s,r}] = RP_{aa}$ and $\mathbf{E}[N_{m,nr}] = (n - R)P_{ab}$ in Eq. (5) would hold for reflexivity in the NN structure.

We estimate the empirical sizes (i.e. significance levels) based on the asymptotic critical values. For example, let T be a test with a χ^2_{df} distribution asymptotically, and let T_i be the value of test statistic for the sample generated at i^{th} Monte Carlo replication for $i = 1, 2, \ldots, N_{mc}$. Then the empirical size of T at level $\alpha = 0.05$, denoted $\widehat{\alpha}_T$ is computed as $\widehat{\alpha}_T = \frac{1}{N_{mc}} \sum_{i=1}^{N_{mc}} \mathbf{I}\left(T_i \geq \chi^2_{df,0.95}\right)$. Furthermore, let Z be a test with a $N(0, 1)$ asymptotic distribution, and let Z_i be the value of test statistic for i^{th} sample generated. Then the empirical size of Z for the left-sided alternative at level $\alpha = 0.05$, denoted $\widehat{\alpha}_Z$ is computed as $\widehat{\alpha}_Z = \frac{1}{N_{mc}} \sum_{i=1}^{N_{mc}} \mathbf{I}(Z_i \leq z_{0.05} = -1.645)$. The empirical size for the right-sided alternative is computed as $\widehat{\alpha}_Z = \frac{1}{N_{mc}} \sum_{i=1}^{N_{mc}} \mathbf{I}(Z_i \geq z_{0.95} = 1.645)$.

5.1.1 Empirical size analysis under CSR independence

We consider the two-class case, with classes X and Y (also referred to as classes 1 and 2) of sizes n_1 and n_2, respectively. Under H_o, at each of $N_{mc} = 10000$ replicates, we generate n_1 X points $\mathcal{X}_{n_1} = \{X_1, \ldots, X_{n_1}\}$ and n_2 Y points $\mathcal{Y}_{n_2} = \{Y_1, \ldots, Y_{n_2}\}$ independently of each other and iid from $\mathcal{U}((0, 1) \times (0, 1))$ and combine X and Y points as $\mathcal{Z}_n = \mathcal{X}_{n_1} \cup \mathcal{Y}_{n_2} = \{Z_1, Z_2, \ldots, Z_n\}$. We consider two cases for CSR independence:

Case 1 $n_1 = n_2 = n = 10, 20, 30, 40, 50$, **Case 2** $n_1 = 20$ and $n_2 = 20, 30, \ldots, 60$.

In case 1, the sample sizes are equal and increasing, in order to determine the influence of the increasing balanced sample sizes on the empirical levels of the tests. On the other hand, case 2 is designed to determine the influence of differences in the sample sizes (i.e., differences in relative abundances of classes) on the empirical levels of the tests.

The empirical significance levels for the tests under CSR independence cases 1 and 2 are presented in Table 3, where $\widehat{\alpha}_P$ is the empirical significance level for $\mathcal{X}^2_P$, Pearson's χ^2 test of independence with 1 df for the RCT (as suggested by Pielou), $\widehat{\alpha}^{>}_{dir}$ (resp. $\widehat{\alpha}^{<}_{dir}$) is for the right(resp. left)-sided alternative, i.e., positive (resp. negative) dependence between NN reflexivity and self pairs, for the directional test, Z_{dir}, in Eq. (12); $\widehat{\alpha}_R$ is for the χ^2 test statistic, $\mathcal{X}^2_R$, for self-reflexivity or mixed-nonreflexivity in the NN structure in Eq. (11); $\widehat{\alpha}^Z_{s,r}$ is for the self-reflexivity in the NN structure test statistic, $Z_{s,r}$; $\widehat{\alpha}^Z_{m,nr}$ is for the mixed-nonreflexivity test statistic, $Z_{m,nr}$ (see Eq. (6)). For $N_{mc} = 10000$ replications, an empirical size estimate is deemed conservative, if smaller than .046 while it is deemed liberal, if larger than .054 at .05 level (based on

Table 3 The empirical significance levels of the tests under CSR independence cases 1 and 2 with $N_{mc} = 10000$ at $\alpha = .05$

Case 1: $n_1 = n_2 = n = 10, 20, \ldots, 50$						
n	$\widehat{\alpha}_P$	$\widehat{\alpha}_{dir}^{>}$	$\widehat{\alpha}_{dir}^{<}$	$\widehat{\alpha}_R$	$\widehat{\alpha}_{s,r}^{Z}$	$\widehat{\alpha}_{m,nr}^{Z}$
10	.044	.103	.058	.047	.054	.056
20	.056	.082	.051	.045	.051	.051
30	.056	.082	.046	.046	.048	.048
40	.065	.082	.047	.048	.048	.048
50	.068	.087	.052	.046	.046	.052
Case 2: $n_1 = 20, n_2 = 20, 30, \ldots, 60$						
n_2	$\widehat{\alpha}_P$	$\widehat{\alpha}_{dir}^{>}$	$\widehat{\alpha}_{dir}^{<}$	$\widehat{\alpha}_R$	$\widehat{\alpha}_{s,r}^{Z}$	$\widehat{\alpha}_{m,nr}^{Z}$
20	.053	.083	.048	.045	.052	.048
30	.057	.084	.050	.052	.045	.052
40	.055	.081	.040	.047	.046	.045
50	.056	.084	.040	.046	.046	.047
60	.051	.076	.034	.047	.051	.046

$\widehat{\alpha}_P$ is the empirical significance level for the χ^2 test of independence with 1 df for the RCT, $\widehat{\alpha}_{dir}^{>}$ (resp. $\widehat{\alpha}_{dir}^{<}$) for the right (resp. left)-sided alternative for the directional test, Z_{dir}, in Eq. (12); $\widehat{\alpha}_R$ for the χ^2 test statistic $\mathcal{X}_R^2$ for self-reflexivity or mixed-nonreflexivity in the NN structure; $\widehat{\alpha}_{s,r}^{Z}$ for the self-reflexivity in the NN structure test statistic, $Z_{s,r}$; $\widehat{\alpha}_{m,nr}^{Z}$ for the mixed-nonreflexivity test statistic, $Z_{m,nr}$. Size estimates larger than .054 (smaller than .046) are liberal (conservative)

binomial critical values with $n = 10000$ trials and probability of success 0.05). Under CSR independence case 1, notice that $\widehat{\alpha}_{dir}^{>}$ is significantly larger than 0.05 (i.e., Z_{dir} is significantly liberal) for all sample size combinations and the χ^2 test of independence for the RCT is liberal for large samples (i.e., for $n \geq 40$). The reflexivity tests (namely, $\mathcal{X}_R^2$, $Z_{s,r}$ and $Z_{m,nr}$) seem to be of the desired level for each sample size considered under case 1. Under case 2, observe that Z_{dir} is liberal at .05 level (although less liberal compared to case 1), and contrary to case 1, χ^2 test of independence for the RCT, $\mathcal{X}_P^2$ is at about the desired level for each sample size combination. Furthermore, $\widehat{\alpha}_{dir}^{<}$ seems to be significantly less than .05 (i.e., the corresponding tests are conservative) when the relative abundance ratio gets larger than two (i.e., when $n_2/n_1 \geq 2$). The reflexivity tests show size performance similar to that in case 1 (i.e., they are at about the nominal level of .05). In both cases, $\mathcal{X}_P^2$ has larger size estimates compared to $\mathcal{X}_R^2$ and the right-sided directional test Z_{dir} has larger size estimates compared to Z-test for self-reflexivity in the NN structure, $Z_{s,r}$. Furthermore, for balanced sample sizes, we recommend the use of the Monte Carlo randomized versions or the use of Monte Carlo critical values for $\mathcal{X}_P^2$ and the right-sided alternative for Z_{dir}. Also, for unbalanced sample sizes, we recommend the use of the Monte Carlo randomized versions or the use of Monte Carlo critical values for the directional test Z_{dir}. A Monte Carlo critical value is determined as the appropriately ranked value of the test statistic in a certain number of generated data sets from the distribution under the null hypothesis. The class sizes are said to be *balanced*, if the relative abundances of the classes are close to one, and they are called *unbalanced*, if the relative abundances deviate substantially from one.

5.1.2 Empirical size analysis under RL

Under the RL pattern, the class labels or marks are assigned randomly to points whose locations are given. Recall that $\mathcal{Z}_n = \{z_1, z_2, \ldots, z_n\}$ is the given set of locations for n points from the background pattern. At each background realization, n_1 of the points are labeled as class 1 or X and the remaining $n_2 = n - n_1$ points are labeled as class 2 or Y.

Types of the background patterns:

Case 1 The background points are a realization of $Z_i \overset{iid}{\sim} \mathcal{U}((0, 1) \times (0, 1))$ for $i = 1, 2, \ldots, n$. That is, the background points, $\mathcal{Z}_n$, are generated iid uniform in the unit square $(0, 1) \times (0, 1)$. We consider $n_1 = n_2 = 10, 20, \ldots, 50$.
Case 2 The background points, $\mathcal{Z}_n$, are generated as in case 1 above with $n_1 = 20$ and $n_2 = 20, 30, \ldots, 60$.
Case 3 The background points, $\mathcal{Z}_n$, are generated from a Matérn cluster process, MatClust(κ, r, μ) (Baddeley and Turner 2005). In this process, first “parent” points are generated from a Poisson process with intensity κ. Then each parent point is replaced by $N \sim$ Poisson(μ) new points which are generated iid inside the circle of radius r centered at the parent point. Each background realization is a realization of $\mathcal{Z}_n$ and is generated from MatClust(κ, r, μ). Let n be the number of points in a particular realization. Then $n_1 = \lfloor n/2 \rfloor$ of these points are labeled as class 1 where $\lfloor x \rfloor$ stands for the floor of x, and $n_2 = n - n_1$ as class 2. In our simulations, we use $\kappa = 2, 4, \ldots, 10$, $\mu = \lfloor 100/\kappa \rfloor$, and $r = 0.1$. That is, we take $(\kappa, \mu) \in \{(2, 50), (4, 25) \ldots, (10, 10)\}$ so as to have about 100 background points on the average with about half being from class 1 and the other half being from class 2.

To reduce the influence of a particular background realization on the size performance of the tests, we generate 100 different realizations of each background pattern. For each case, the RL scheme described is repeated 1000 times for each (n_1, n_2) combination at each of 100 different background realizations. So we have $N_{mc} = 100000$. In RL with background cases 1 and 2, the points are from HPP in the unit square with fixed n_1 and n_2 (i.e., from binomial process), where case 1 is for assessing the effect of equal but increasing sample sizes on the tests, while case 2 is for assessing the effect of increasing differences in sample sizes of the classes (with one class size being fixed, while the other is increasing). On the other hand, in the background realizations of case 3, centers and numbers of clusters are random. On the average, with increasing κ, the cluster sizes tend to decrease and the number of clusters tend to increase (so as to have fixed class sizes on the average). Hence in case 3, we investigate the influence of increasing number of clusters with randomly determined centers on the size performance of the tests.

The empirical size estimates of the tests under RL with background cases 1-3 are presented in Table 4. For $N_{mc} = 100000$ replications, an empirical size estimate is

Table 4 The empirical significance levels of the tests under RL with background cases 1–3 with $N_{mc} = 100000$ (1000 replications for each of 100 background realizations) at $\alpha = .05$

Case 1: $\mathcal{U}((0,1)\times(0,1))$ with $n_1 = n_2 = n$						
n	$\widehat{\alpha}_P$	$\widehat{\alpha}_{dir}^{>}$	$\widehat{\alpha}_{dir}^{<}$	$\widehat{\alpha}_R$	$\widehat{\alpha}_{s,r}^{Z}$	$\widehat{\alpha}_{m,nr}^{Z}$
10	.045	.101	.052	.047	.054	.051
20	.056	.083	.051	.046	.054	.051
30	.060	.084	.045	.048	.047	.048
40	.067	.084	.048	.049	.050	.048
50	.070	.085	.050	.048	.050	.050
Case 2: $\mathcal{U}((0,1)\times(0,1))$ with $n_1 = 20$ and n_2						
n_2	$\widehat{\alpha}_P$	$\widehat{\alpha}_{dir}^{>}$	$\widehat{\alpha}_{dir}^{<}$	$\widehat{\alpha}_R$	$\widehat{\alpha}_{s,r}^{Z}$	$\widehat{\alpha}_{m,nr}^{Z}$
20	.056	.082	.052	.047	.054	.052
30	.058	.086	.046	.049	.044	.048
40	.053	.080	.042	.047	.042	.046
50	.055	.080	.038	.047	.050	.045
60	.050	.074	.033	.046	.051	.046
Case 3: MatClust($\kappa, r = 0.1, \mu = \lfloor 100/\kappa \rfloor$)						
κ	$\widehat{\alpha}_P$	$\widehat{\alpha}_{dir}^{>}$	$\widehat{\alpha}_{dir}^{<}$	$\widehat{\alpha}_R$	$\widehat{\alpha}_{s,r}^{Z}$	$\widehat{\alpha}_{m,nr}^{Z}$
2	.067	.085	.049	.048	.051	.048
4	.067	.084	.051	.048	.049	.050
6	.067	.085	.051	.049	.051	.051
8	.068	.084	.048	.047	.048	.048
10	.067	.083	.051	.049	.049	.050

The empirical size labeling for the tests is as in Table 3. Size estimates larger than .051 (smaller than .049) are liberal (conservative). MatClust(κ, r, μ) stands for the Matérn cluster process

deemed conservative, if smaller than .049 while it is deemed liberal, if larger than .051 at .05 level (based on binomial critical values with $n = 100000$ trials and probability of success 0.05). The size performance under cases 1 and 2 are similar to that under CSR independence cases 1 and 2, respectively. However, under RL with background case 3, $\mathcal{X}_P^2$ is liberal for each κ value, which would be expected, since for each κ value $n_1 \approx n_2 \approx 50$ (and this test was liberal for $n_1 = n_2 = 50$ under RL background case 1). Notice also that the size estimates of the tests are not influenced by the number of clusters, κ, when the class sizes are fixed.

Based on the empirical size performance of the tests under CSR independence and RL, we conclude that the newly proposed reflexivity tests, i.e., $\mathcal{X}_R^2$, $Z_{s,r}$ and $Z_{m,nr}$, are at about the desired nominal level, hence they are robust to the imbalance in the class sizes and their asymptotic approximation can be employed for both balanced and unbalanced class sizes. Furthermore, the directional test, Z_{dir}, is liberal for the right-sided alternative for small to large samples and Pielou's χ^2 test of independence on the RCT, $\mathcal{X}_P^2$, is liberal for large samples. So we recommend Monte Carlo randomization for these tests under these situations. The left-sided directional Z-test is conservative when the relative abundances of the classes are very different. That is, this test is severely confounded by the differences in relative abundances of the classes. Therefore,

Table 5 The power estimates under the cases I, II, and III alternatives in Eqs. (13), (14), (15), respectively, with $N_{mc} = 10000$, $n_1 = n_2 = 40$ at $\alpha = .05$. $\widehat{\beta}_R$ is power estimate for the χ^2 test statistic, $\mathcal{X}^2_R$ in Eq. (10); $\widehat{\beta}^Z_{sr}$ for the self-reflexivity in the NN structure test statistic, $Z_{s,r}$; $\widehat{\beta}^Z_{mn}$ for the mixed-nonreflexivity test statistic, $Z_{m,nr}$. The ">" (resp. "<") sign in the superscript implies the power is estimated for the right-sided (resp. left-sided) alternative

	Case I alternatives				Case II alternatives				Case III alternatives		
	$\widehat{\beta}_R$	$\widehat{\beta}^{Z,>}_{s,r}$	$\widehat{\beta}^{Z,<}_{m,nr}$		$\widehat{\beta}_R$	$\widehat{\beta}^{Z,<}_{s,r}$	$\widehat{\beta}^{Z,<}_{m,nr}$		$\widehat{\beta}_R$	$\widehat{\beta}^{Z,>}_{s,r}$	$\widehat{\beta}^{Z,<}_{m,nr}$
H_I^1	.24	.27	.26	H_{II}^1	.23	.48	.03	H_{III}^1	.46	.45	.42
H_I^2	.995	.98	.97	H_{II}^2	.63	.88	.01	H_{III}^2	.94	.87	.87
H_I^3	.994	.97	.96	H_{II}^3	.79	.97	$<.01$	H_{III}^3	.9999	.998	.995

we recommend the use of these tests when the sample sizes are balanced, otherwise we recommend Monte Carlo randomization for these tests.

5.2 Empirical power analysis

To compare the empirical power performance of the tests, we consider various alternative cases for self-reflexivity or mixed-nonreflexivity in the NN structure. The empirical power estimates are computed at $\alpha = .05$ as in the size estimates in Sect. 5.1.

Case I For this class of alternatives, we generate $X_i \stackrel{iid}{\sim} \mathcal{U}((0,1) \times (0,1))$ for $i = 1, \ldots, n_1$ and $Y_j \stackrel{iid}{\sim} \text{BVN}(1/2, 1/2, \sigma_1, \sigma_2, \rho)$ for $j = 1, \ldots, n_2$, where $\text{BVN}(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ is the bivariate normal distribution with mean (μ_1, μ_2) and covariance $\begin{bmatrix} \sigma_1 & \rho \\ \rho & \sigma_2 \end{bmatrix}$. In our simulations, we set $\sigma_1 = \sigma_2 = \sigma$ and $\rho = 0$. We consider the following three alternatives:

$$H_I^1 : \sigma = 1/5, \quad H_I^2 : \sigma = 2/15, \text{ and } H_I^3 : \sigma = 1/10. \tag{13}$$

The classes 1 and 2 (i.e., X and Y) have different distributions with different local intensities. In particular, X points are a realization of uniform distribution in the unit square, while Y points are clustered around the center of the unit square $(1/2, 1/2)$ with the level of clustering increasing as σ decreases. This suggests a high level of segregation of Y points from X points.

The empirical power estimates under the alternatives, $H_I^1 - H_I^3$, with $n_1 = n_2 = 40$ are presented in Table 5, where $\widehat{\beta}_R$ is power estimate for the χ^2 test statistic, $\mathcal{X}^2_R$, for self-reflexivity or mixed-nonreflexivity in the NN structure; $\widehat{\beta}^Z_{sr}$ is for the self-reflexivity test statistic, $Z_{s,r}$; $\widehat{\beta}^Z_{mn}$ is for the mixed-nonreflexivity test statistic, $Z_{m,nr}$. *We omit the power estimates for the χ^2 test of independence and one-sided directional tests on the RCT, since they are undefined when an entire column of the RCT is zero, which happens with substantial probability under case I alternatives.* Under the case I alternatives, the power estimates increase as σ decreases. In particular, χ^2 test of NN reflexivity, and right-sided test of self-reflexivity in the NN structure, $Z_{s,r}$ and left-sided mixed-nonreflexivity test, $Z_{m,nr}$ have high power estimates, which implies significant self-reflexivity in the NN structure, which is caused by the substantial level of clustering of Y points around the center of the unit square.

Case II For this type of alternatives, first, we generate $X_i \stackrel{iid}{\sim} \mathcal{U}((0,1)\times(0,1))$ for $i = 1, 2, \ldots, n_1$ and for each $j = 1, 2, \ldots, n_2$, we generate Y_j around a randomly picked X_i with probability p in such a way that $Y_j = X_i + R_j\,(\cos T_j, \sin T_j)^t$ where v^t stands for transpose of the vector v, $R_j \sim \mathcal{U}(0, \min_{i\neq j} d(X_i, X_j))$ and $T_j \sim \mathcal{U}(0, 2\pi)$ or generate Y_j uniformly in the unit square with probability $1-p$. In the pattern generated, Y_j are more associated with X_i. The three values of p constitute the following alternatives:

$$H_{II}^1 : p = .25, \quad H_{II}^2 : p = .50, \text{ and } H_{II}^3 : p = .75. \tag{14}$$

In this case, X points constitute a realization of the uniform distribution in the unit square, while Y points are clustered around the X points, and the level of clustering increases as p increases. The empirical power estimates under the alternatives, $H_{II}^1 - H_{II}^3$, with $n_1 = n_2 = 40$ are presented in Table 5. Notice that χ^2 test of NN reflexivity has high power which increases as p increases, and $Z_{s,r}$ has higher power for the left-sided alternative. On the other hand, $Z_{m,nr}$ has very low power for both left- and right-sided alternatives (even significantly lower than the nominal level of .05), but slightly higher for the left-sided alternative (hence only power estimates for the left-sided alternative are presented). Therefore, we can conclude that there is significant mixed-reflexivity but there is lack of mixed-nonreflexivity (in fact, mixed-nonreflexivity in the NN structure deviates from the expected level less frequently than it would under the null case).

Case III For this class of alternatives, we consider $X_i \stackrel{iid}{\sim} \mathcal{U}((0,1-s)\times(0,1-s))$ for $i = 1, \ldots, n_1$, and $Y_j \stackrel{iid}{\sim} \mathcal{U}((s,1)\times(s,1))$ for $j = 1, \ldots, n_2$. The three values of s constitute the following alternatives;

$$H_{III}^1 : s = 1/6, \quad H_{III}^2 : s = 1/4, \text{ and } H_{III}^3 : s = 1/3. \tag{15}$$

Notice that these alternatives are the segregation alternatives considered for Monte Carlo analysis in Ceyhan (2010). The empirical power estimates under the segregation alternatives $H_{III}^1 - H_{III}^3$ are presented in Table 5. The NN reflexivity tests have high power which increases as s increases. Furthermore, $Z_{s,r}$ has high power for the right-sided alternative and $Z_{m,nr}$ has high power for the left-sided alternative, which indicates significant self-reflexivity in the NN structure.

Case IV We generate $X_i \stackrel{iid}{\sim} S_1$ for $i = 1, \ldots, \lfloor n_1/2 \rfloor$ and $Y_j \stackrel{iid}{\sim} S_2$ for $j = 1, \ldots, \lfloor n_2/2 \rfloor$. Then for $k = \lfloor n_1/2 \rfloor + 1, \ldots, n_1$, we generate $X_k = X_{k-\lfloor n_1/2 \rfloor} + r\,(\cos T_j, \sin T_j)^t$ and for $l = \lfloor n_2/2 \rfloor + 1, \ldots, n_2$, we generate $Y_l = Y_{l-\lfloor n_1/2 \rfloor} + r\,(\cos T_j, \sin T_j)^t$ where $r \in (0,1)$ and $T_j \sim \mathcal{U}(0, 2\pi)$. Appropriate small choices of r will yield an abundance of self-reflexive pairs. The three values of r we consider constitute the self-reflexivity alternatives at each support pair (S_1, S_2). Then the nine alternative combinations we consider are given by

Table 6 The power estimates under the case IV alternatives in Eq. (16) with $N_{mc} = 10000$, $n_1 = n_2 = 40$ at $\alpha = .05$. The empirical power labeling and superscripting for "<" and ">" are as in Table 5

r	$\widehat{\beta}_R$	$\widehat{\beta}^{Z,>}_{s,r}$	$\widehat{\beta}^{Z,>}_{m,nr}$
	H^1_{IV}		
1/7	.06	.08	.08
1/8	.08	.13	.08
1/9	.10	.20	.09
	H^2_{IV}		
1/7	.06	.08	.07
1/8	.07	.11	.07
1/9	.08	.16	.06
	H^3_{IV}		
1/7	.06	.07	.06
1/8	.06	.11	.06
1/9	.08	.15	.05

$$
\begin{aligned}
&(i) H^1_{IV} : S_1 = S_2 = (0,1) \times (0,1), (a)\, r = 1/7, (b)\, r = 1/8, (c)\, r = 1/9,\\
&(ii) H^2_{IV} : S_1 = (0, 5/6) \times (0, 5/6) \text{ and } S_2 = (1/6, 1) \times (1/6, 1), (a)\, r = 1/7, (b)\\
&\quad r = 1/8, (c)\, r = 1/9,\\
&(iii) H^3_{IV} : S_1 = (0, 3/4) \times (0, 3/4) \text{ and } S_2 = (1/4, 1) \times (1/4, 1)(a)\\
&\quad r = 1/7, (b)\, r = 1/8, (c)\, r = 1/9. \qquad (16)
\end{aligned}
$$

In this case, under H^2_{IV} and H^3_{IV}, by construction, there is segregation of the classes due to the choices of the supports. Additionally, with decreasing r, the self-reflexive pairs will be more abundant. The empirical power estimates under the alternatives $H^1_{IV} - H^3_{IV}$ are presented in Table 6. Notice that the NN reflexivity tests have relatively low power estimates. Furthermore, $Z_{s,r}$ and $Z_{m,nr}$ have higher power for the right-sided alternative, which indicates significant presence of self-reflexivity and mixed non-reflexivity in the NN structure (but each at a mild level). The power estimates for these tests decrease from H^1_{IV} to H^3_{IV} but they increase as r decreases from (a) to (c) at each (S_1, S_2) combination.

Case V In this case, first, we generate $X_i \overset{iid}{\sim} \mathcal{U}((0,1) \times (0,1))$ and then generate Y_j as $Y_j = X_i + r\,(\cos T_j, \sin T_j)^t$ where $r \in (0,1)$ and $T_j \sim \mathcal{U}(0, 2\pi)$. In the pattern generated, appropriate choices of r will cause Y_j and X_i more associated, that is, a Y point will be more likely to be the NN of an X point, and vice versa. The three values of r we consider constitute the three association alternatives;

$$H^1_V : r = 1/4, \quad H^2_V : r = 1/7, \text{ and } H^3_V : r = 1/10. \qquad (17)$$

These are also the association alternatives considered for Monte Carlo analysis in Ceyhan (2010).

Table 7 The power estimates under the case V alternatives in Eq. (17) with $N_{mc} = 10000$, $n_1 = n_2 = 40$ at $\alpha = .05$. The empirical power labeling and superscripting for "<" and ">" are as in Table 5

	$\widehat{\beta}_R$	$\widehat{\beta}^{Z,<}_{s,r}$	$\widehat{\beta}^{Z,>}_{m,nr}$
H_V^1	.16	.30	.12
H_V^2	.45	.63	.20
H_V^3	.70	.84	.24

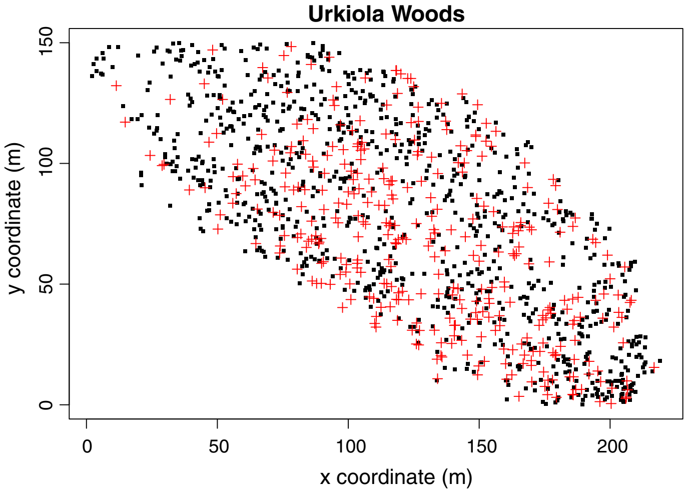

Fig. 1 The scatter plot of the locations of birch trees (*solid squares*), and oak trees (*plus symbols*) in the Urkiola Natural Park, Basque region, northern Spain

The empirical power estimates under $H_V^1 - H_V^3$ are presented in Table 7. Notice that the χ^2 test of NN reflexivity has high power (which increases as r decreases). But $Z_{s,r}$ has high power for the left-sided alternative and $Z_{m,nr}$ has high power for the right-sided alternative only, which indicates significant mixed-reflexivity and presence of moderate mixed-nonreflexivity in the NN structure. The power estimates for these tests increase as r decreases.

6 Example data sets

To illustrate the methodology, we use two example data sets: the Urkiola Woods data (Laskurain 2008) which is available in the spatstat package in R (Baddeley and Turner 2005) and the swamp tree data of Good and Whipple (1982).

6.1 Urkiola woods data

The Urkiola Woods data contains locations of trees (in meters) in a secondary wood in Urkiola Natural Park, Basque region, northern Spain (Laskurain 2008). The data set contains 886 birch trees (Betula celtiberica) and 359 oak trees (Quercus robur).

Table 8 The RCT for Urkiola Woods data

	Pair type		Total
	Self	Mixed	
Reflexivity			
Reflexive	474 (431.34)	258 (300.66)	732
Non-reflexive	323 (302.29)	190 (210.71)	513
Total	797	448	1245

The expected values of the cell counts under RL are provided in parentheses

Table 9 The test statistics and the p values for Urkiola Woods data

	$\mathcal{X}_P^2$	$Z_{dir}^{>}$	$Z_{dir}^{<}$	$T_F^{>}$	$T_F^{<}$	$\mathcal{X}_R^2$	$Z_{sr}^{>}$	$Z_{mn}^{<}$
TS	.35	.65		1.08		11.37	2.50	−2.00
p_{asy}	.5564	.2584	.7415	.2780*	.7607*	.0034	.0062	.0230
p_{rand}	.5648	.3079	.6921	.3079	.6928	.0030	.0070	.0209
p_{tor}	.4275	.4275	.5763	.4275	.5763	.4948	.4293	.5523

Z_{sr}, Z_{mn}, $\mathcal{X}_P^2$, and $\mathcal{X}_R^2$ are as defined in the text; $Z_{dir}^{>}$ and $Z_{dir}^{<}$ are for the right- and left-sided alternatives with Z_{dir}; $T_F^{>}$ and $T_F^{<}$ are one-sided Fisher's exact test (for the right and left-sided tests on the RCT, respectively). "TS" stands for the test statistic, p_{asy} for the p values based on asymptotic critical values (except for the exact tests), p_{rand} for the p values based on Monte Carlo randomization, and p_{tor} for the p values based on random toroidal shifts. *The p values for the exact tests are computed using the table inclusive version as described in Ceyhan (2014)

This data set is actually a part of a more extensive data set collected and analyzed by Laskurain (2008). The scatter plot of the tree locations are presented in Fig. 1.

The RCT for this data set is presented in Table 8 where the expected values of the entries of the RCT are provided in parentheses. We estimate the "self-enrichment" by the ratio of number of observed self pairs to the number of expected self pairs for each row (reflexive or non-reflexive), i.e., $N_{s,r}/\mathbf{E}[N_{s,r}]$ and $N_{s,nr}/\mathbf{E}[N_{s,nr}]$. In this data set, these ratios are $474/431.34 \approx 1.10$ for reflexive and $323/302.29 \approx 1.07$ for non-reflexive self-pairs. The similarity of the row-specific ratios in the self column suggests that the pattern in this data set is the overall enrichment of self-pairs.

We compute $Q = 812$, $R = 732$, and $T = 360$, and the corresponding empirical ratios are $Q/n \approx .65$, $R/n \approx .59$ and $T/n \approx .29$. We present the test statistics and the associated p values in Table 9, where Z_{sr}, Z_{mn}, $\mathcal{X}_P^2$, and $\mathcal{X}_R^2$ are as defined in the text, and $Z_{dir}^{>}$ and $Z_{dir}^{<}$ are for the right-sided and left-sided versions of the directional test Z_{dir}, respectively. Furthermore, $T_F^{>}$ and $T_F^{<}$ are one-sided Fisher's exact test (for the right and left-sided tests on the RCT, respectively) where the test statistic is the odds ratio. In this table p_{asy} stands for the p-value based on the asymptotic approximation (i.e., asymptotic critical value), except for the exact tests; p_{rand} is based on Monte Carlo randomization of the labels on the given locations of the trees 10000 times; and p_{tor} is based on 100000 random toroidal shifts of the data set. For the exact tests, the p-value written for the p_{asy} row is computed using the table inclusive version as described in Ceyhan (2014). Notice that p_{asy} and p_{rand} are similar for all the tests. However, p_{tor} values are very different from p_{asy} and p_{rand} values, especially for the reflexivity tests. The reason for this discrepancy is that the supports of the classes in Urkiola Woods data are not rectangular, hence the use of toroidal shifts is not appropriate (Dixon 2002c). Notice that Pearson's χ^2 test of independence and Fisher's exact test on the RCT suggest no significant deviation from independence. However, these tests do not have the correct sampling distribution and also the exact tests are valid for small sample sizes (less than about 50). Hence the asymptotic approximation for Pielou's test and the exact tests would not be reliable. However, the more reliable Monte Carlo randomized p values, p_{rand}, are also not significant for Pearson's χ^2 test of independence and Fisher's exact test suggesting independence between NN reflexivity and pair type as self or mixed. On the other hand, the Z-test for self-reflexivity in the NN structure is significant for the right-sided alternative and the Z-test for mixed-nonreflexivity is significant for the left-sided alternative and χ^2 test for NN reflexivity, $\mathcal{X}^2_R$, is significant. Thus, the overall reflexivity test implies there is significant deviation of diagonal cells from their expected values. In particular, there is significant self-reflexivity and significant lack of self-nonreflexivity in the NN structure. In other words, we show that reflexive pairs are more likely to be self-self and non-reflexive pairs are less likely to be mixed (and hence non-reflexive pairs are more likely to be self-self also). Hence the total number of self NN pairs significantly exceeding its expected value leads towards an analysis of species-specific contributions to these totals. Regardless of reflexivity, this analysis of the species-specific contributions can be performed by an analysis of the diagonals of the usual NNCT (Dixon 1994). The cell-specific test (i.e., the test for the diagonal cell) for birch trees in the NNCT is $Z = 2.91$ with $p = .0018$ and for oaks is $Z = 2.71$ with $p = .0034$ which implies significant segregation of both species from each other. This significant segregation of the species supports self enrichment in NN pairs, and this enrichment is at about the same level for reflexive and non-reflexive pairs.

6.2 Swamp tree data

From the swamp tree data of Good and Whipple (1982), Dixon (2002b) used a single 50 m × 200 m rectangular plot (denoted as the (0, 200) × (0, 50) rectangle) to illustrate his nearest neighbor contingency table (NNCT) methods. All live or dead trees with 4.5 cm or more dbh (diameter at breast height) were recorded together with their species labels. The plot contains 13 different tree species, but for illustrative purposes (and for brevity in presentation) we only consider the three groups, namely, Carolina ashes, bald cypresses, and miscellaneous other trees (i.e., trees not belonging to one of the first five most frequent species). In practice, such an omission of classes or species is not recommended (see Remark 6.1). The scatter plot of these tree locations are presented in Fig. 2 where for convenience in presentation x and y coordinates are flapped which do not influence the methods based on NN distances.

The RCT for this data together with the expected values of the entries is presented in Table 10. In this data set, self-enrichment ratios are 1.73 for reflexive and 1.07 for non-reflexive self-pairs. The substantial difference between the row-specific ratios in the self column suggests that the pattern in this data set is the abundance of self reflexive pairs.

We compute $Q = 160$, $R = 210$, and $T = 75$ and the corresponding empirical proportions are .51, .67 and .24. We present the test statistics and the associated p val-

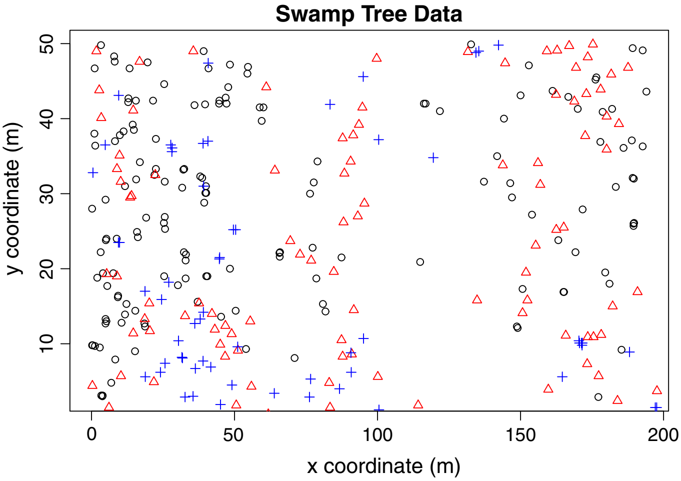

Fig. 2 The scatter plot of the locations of Carolina ashes (*circles*), bald cypresses (*triangles*), and miscellaneous trees (*plus symbols*) in the Swamp Tree data

Table 10 The RCT for the Swamp Tree data

	Pair type		Total
	Self	Mixed	
Reflexivity			
Reflexive	138 (79.54)	72 (130.46)	210
Non-reflexive	42 (39.39)	62 (64.61)	104
Total	180	134	314

The expected values of the cell counts under RL are provided in parentheses

Table 11 The test statistics and the p values for Swamp Tree data

	$\mathcal{X}^2_P$	$Z^{>}_{dir}$	$Z^{<}_{dir}$	$T^{>}_F$	$T^{<}_F$	$\mathcal{X}^2_R$	$Z^{>}_{sr}$	$Z^{<}_{mn}$
TS	17.22	4.27		2.81		38.40	6.14	−.54
p_{asy}	<.0001	<.0001	≈1.00	<.0001	≈1.00	<.0001	<.0001	.2950
p_{rand}	.0001	.0001	.9999	.0001	.9999	<.0001	<.0001	.2987
p_{tor}	.2170	.2170	.7886	.2170	.7886	.0731	.6233	.9310

The column and row labeling are as in Table 9

ues in Table 11, where p_{asy} and p_{rand} are similar for all the tests, but are very different from p_{tor} values since the processes generating the points seem to be non-stationary. Furthermore, toroidal shift might produce biased results when there are clusters around the edges. Pearson's χ^2 test of independence and Fisher's exact test on the RCT suggests significant deviation from independence based on the more reliable Monte Carlo randomized p values. This result supports dependence (in the form of positive association) between NN reflexivity and pair type as self or mixed. On the other hand, the overall reflexivity test is also highly significant, which implies a significant devi-

Table 12 The test statistics and the p values for Swamp Tree data in the pairwise setting (top) and one-vs-rest setting (bottom)

Pairwise analysis for Swamp Tree data

	Carolina ashes versus bald cypresses			Carolina ashes versus others			Bald cypresses versus others		
	$\mathcal{X}^2_R$	$Z^>_{sr}$	$Z^<_{mn}$	$\mathcal{X}^2_R$	$Z^>_{sr}$	$Z^<_{mn}$	$\mathcal{X}^2_R$	$Z^>_{sr}$	$Z^<_{mn}$
TS	10.14	3.11	−.59	52.27	6.46	−2.53	17.99	4.02	−1.21
p_{asy}	.0063	.0009	.2777	<.0001	<.0001	.0058	.0001	<.0001	.1139
p_{rand}	.0054	.0014	.2814	<.0001	<.0001	.0047	.0001	<.0001	.1419
p_{tor}	.7534	.7894	.3877	.6366	.8789	.2081	.7839	.2284	.8096

One-vs-rest analysis for Swamp Tree data

	Carolina ashes versus rest			Bald cypresses versus rest			Others versus rest		
	$\mathcal{X}^2_R$	$Z^>_{sr}$	$Z^<_{mn}$	$\mathcal{X}^2_R$	$Z^>_{sr}$	$Z^<_{mn}$	$\mathcal{X}^2_R$	$Z^>_{sr}$	$Z^<_{mn}$
TS	24.59	4.81	−1.21	9.21	2.85	1.28	35.56	5.47	−1.00
p_{asy}	<.0001	<.0001	.1134	.0100	.0022	.0999	<.0001	<.0001	.1590
p_{rand}	<.0001	<.0001	.1135	.0095	.0021	.1405	<.0001	<.0001	.1537
p_{tor}	.4244	.8192	.2246	.5666	.6933	.7700	.8422	.8854	.6578

The column and row labeling are as in Table 9 ation of one or both of the diagonal cells from their expected values. In particular, the Z-test for self-reflexivity is significant for the right-sided alternative, but the Z-test for mixed-nonreflexivity is not significant for either of the one-sided alternatives. Hence there is strong self-reflexivity in the NN structure and no significant deviation from mixed-nonreflexivity in the NN structure (i.e., no significant deviation from self-nonreflexivity). In other words, reflexive pairs are more likely to be self NN pairs but non-reflexive pairs are at about their expected levels. To figure out which species contribute more to the abundance of self-reflexive NN pairs, we conduct cell-specific tests on the diagonals of the NNCT. The cell-specific test for Carolina ashes in the NNCT is $Z = 4.86$ with $p < .0001$, for bald cypresses is $Z = 1.07$ with $p = .1414$ and for others is $Z = 6.76$ with $p < .0001$ which implies significant segregation of Carolina ashes and other trees, which partly accounts for abundance of self-reflexive NN pairs.

Post hoc tests as pairwise or one-vs-tests are performed as described in Sect. 3.4 each in a two-class setting as in Sect. 6.1 to determine which (pairs) of species exhibit reflexivity in the NN structure. The corresponding test statistics and the p values are presented in Table 12. The p_{asy} and p_{rand} values are very close, but very different from p_{tor} values, due to inappropriateness of the toroidal shift for this data set. In the pairwise test setting, we observe a significant presence of self-reflexivity and self-nonreflexivity when Carolina ashes and others are compared. But when bald cypresses and others are compared there is significant self-reflexivity and no deviation from mixed-nonreflexivity; the same conclusion is reached when Carolina ashes and bald cypresses are compared. In the one-vs-rest setting, we observe a significant presence of self-reflexivity and no deviation from mixed-nonreflexivity when each group is compared to the rest.

Remark 6.1 (Caveat on exclusion of species) The exclusion of classes or species in a data set is not recommended in practice, as it will very likely yield spurious results. In the Swamp Tree data set, we omit some of the tree species in order to simplify our presentation, and this is solely done for the purpose of illustration, in the sense that, the conclusions of the analysis of this data set would hold only if there were only three groups, whose locations were as in Fig. 2. However, regarding ecological relevance, removing species changes the neighborhood relationships. In fact, this can be observed in the pairwise analysis provided in Table 12. Hence, in practice all (relevant) groups/species should be included. For example, if only the five most prevalent species were of interest (as in Dixon 2002a), they should be included in the analysis, or else all 13 species should be included.

7 Discussion and conclusions

In this article, we discuss various tests of reflexivity in the NN structure. In particular, we investigate Pielou's test proposed for niche specificity (Pielou 1961) and introduce new tests of NN reflexivity using a contingency table based on the NN relations between classes or species. We consider Pielou's test, determine its appropriate null hypothesis and the underlying assumptions and demonstrate that Pielou's contingency table intended for niche specificity is actually more appropriate for NN reflexivity (hence called reflexivity contingency table (RCT) in this article). We demonstrate that

Pielou's approach is not appropriate for completely mapped spatial data. We provide the (correct) asymptotic distribution for the entries of the RCT and thus propose new tests of NN reflexivity. Pearson's χ^2 test of independence (whose use on RCT is suggested by Pielou 1961) and the one-sided versions on the RCT are slightly liberal with the asymptotic approximation, but our new NN reflexivity tests are at about the desired level. We also provide extensions of the methodology to multi-class case with more than two classes and describe the possible post hoc tests in such a multi-class case when the overall test of NN reflexivity yields a significant result.

Pielou's approach and the new approach on the RCT are testing different null and alternative hypotheses (hence they would have different rejection and acceptance regions). In particular, Pielou's approach is based on the usual χ^2 test for the independence of NN reflexivity and pair type, while the new NN reflexivity tests are based on the normal approximation of the entries with their expected values under CSR independence or RL with completely mapped data. Hence Pielou's test is appropriate when, under H_o, we have a random sample of labels of base-NN pairs in terms of NN reflexivity and pair type as self or mixed, and such a random sample can be obtained by an appropriate sparse sampling from the data. On the other hand, the tests in Eqs. (6) and (7) (i.e., the reflexivity tests) are appropriate for completely mapped spatial data under RL (and under CSR independence conditional on various quantities such as the number of reflexive NN pairs).

Throughout this article, we assume that the total sample size and class sizes are all fixed. If it is desired to have the sample size be a random variable, we may consider a spatial Poisson point process on the region of interest instead of the uniform binomial process. In fact, this case is also a realistic situation for a data collection scheme in the plant ecology. That is, in the region of interest, one can examine each subject, determine its species and that of its NN. In this framework, all margins in the RCT would be random. The effects of such randomness on the behavior (e.g., distribution), size and power performance of the tests is a topic of ongoing research. For the cases where CSR independence is the appropriate benchmark (see Sect. 2), this framework might be more realistic, but for the cases where RL is the appropriate benchmark, then our approach in this article might be more realistic.

In the literature usually NN relationships are based on the distance metrics. For example, in this article, Euclidean distance in $\mathbb{R}^2$ is the only metric used. The NN relations based on dissimilarity measures is an extension of NN relations based on distance metrics. In such an extension, NN of an object, x, refers to the object with the minimum dissimilarity to x. We assume that the objects (events) lie in a finite or infinite dimensional space satisfying the lack of any inter-dependence which implies self-reflexivity or mixed-nonreflexivity in the NN structure. Under RL, the objects' locations are fixed yielding fixed interpoint dissimilarity measures, but the labels are assigned randomly to the objects. The extensions of Pielou's test of independence and our newly proposed tests on the RCT are straightforward.

Acknowledgements We would like to thank an anonymous associate editor and two referees, whose constructive comments and suggestions greatly improved the presentation and flow of the paper. EC was supported by the research agency TUBITAK via Project # 111T767 and by the European Commission under the Marie Curie International Outgoing Fellowship Programme via Project # 329370 titled PRinHDD.

Appendix: Sampling distribution of the diagonal cell counts in the RCT under RL

$N_{s,r}$: Number of self-reflexive NNs

Recall that $N_{s,r} = \sum_{u=1}^{m} N_r^u$ where $N_r^u = \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} z_{ij}^u$ (see Sect. 3).

Since z_{ij}^u is an indicator random variable, we have $\mathbf{E}[z_{ij}^u] = P(L_i = L_j = u) = p_{uu}$, and hence by the linearity of the expectation we have

$$\mathbf{E}[N_r^u] = \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} \mathbf{E}[z_{ij}^u] = p_{uu} \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} = R p_{uu}.$$

Therefore, we obtain

$$\mathbf{E}[N_{s,r}] = R \sum_{a=1}^{m} p_{aa} = R P_{aa}.$$

To compute the variance of $N_{s,r}$, we need to evaluate $\mathbf{E}[N_r^u N_r^v]$ for every u, v.

$$\mathbf{E}[N_r^u N_r^v] = \mathbf{E}\left[\sum_{i,j} r_{ij} z_{ij}^u \sum_{k,l} r_{kl} z_{kl}^v\right] = \sum_{i,j,k,l} r_{ij} r_{kl} \mathbf{E}[z_{ij}^u z_{kl}^v].$$

First note that $r_{ij} = 0$ whenever $i = j$. We also have $r_{ij} r_{kl} = 0$ if $\{i, j\} \cap \{k, l\}$ has exactly one element, since each point has a unique NN. Moreover, observe that $r_{ij}^2 = r_{ij} r_{ji} = r_{ij}$. Therefore, we obtain

$$\mathbf{E}[N_r^u N_r^v] = \sum_{i,j} r_{ij} \mathbf{E}[z_{ij}^u z_{ij}^v] + \sum_{i,j} r_{ij} \mathbf{E}[z_{ij}^u z_{ji}^v] + \sum_{i \neq j \neq k \neq l} r_{ij} r_{k,l} \mathbf{E}[z_{ij}^u z_{kl}^v].$$

Notice that

$$\sum_{i \neq j \neq k \neq l} r_{ij} r_{k,l} = R^2 - 2R \text{ and } \mathbf{E}[z_{ij}^u z_{kl}^v] = p_{uuvv} \text{ if } i \neq j \neq k \neq l,$$

and for distinct class labels a and b we have

$$\mathbf{E}[z_{ij}^a z_{ij}^b] = \mathbf{E}[z_{ij}^a z_{ji}^b] = 0.$$

Thus, we obtain

$$\mathbf{E}[(N_r^a)^2] = 2R p_{aa} + (R^2 - 2R) p_{aaaa} \text{ and } \mathbf{E}[N_r^a N_r^b] = (R^2 - 2R) p_{aabb},$$

and

$$\mathbf{E}\,[N_{s,r}^2] = 2R\sum_{a=1}^{m} p_{aa} + (R^2 - 2R)\sum_{a=1}^{m} p_{aaaa} + (R^2 - 2R)\sum_{1\le a\neq b\le m} p_{aabb}.$$

Hence, we get

$$\begin{aligned}\mathbf{Var}\,[N_{s,r}] = R^2 &\left(\sum_{a=1}^{m} p_{aaaa} + \sum_{1\le a\neq b\le m} p_{aabb} - \left(\sum_{a=1}^{m} p_{aa}\right)^2\right)\\ &+ 2R\left(\sum_{a=1}^{m} p_{aa} - \sum_{a=1}^{m} p_{aaaa} - \sum_{1\le a\neq b\le m} p_{aabb}\right).\end{aligned}$$

$N_{m,nr}$: Number of mixed-nonreflexive NNs

Recall that $N_{m,nr} = \sum_{1\le a\neq b\le m} N_{nr}^{ab}$ (see Sect. 3). Since

$$\mathbf{E}\,[N_{nr}^{ab}] = \sum_{i=1}^{n}\sum_{j=1}^{n} y_{ij}\mathbf{E}\,[z_{ij}^{ab}] = (n-R)\,p_{ab},$$

we get

$$\mathbf{E}\,[N_{m,nr}] = (n-R)\sum_{1\le a\neq b\le m} p_{ab} = (n-R)\,P_{ab}.$$

We now compute $\mathbf{E}\,[N_{nr}^{ab} N_{nr}^{cd}]$ where $a \neq b$ and $c \neq d$.

$$\mathbf{E}\,[N_{nr}^{ab} N_{nr}^{cd}] = \mathbf{E}\left[\sum_{i\neq j} y_{ij} z_{ij}^{ab} \sum_{k\neq l} y_{kl} z_{kl}^{cd}\right] = \sum_{1\le i,j,k,l\le n} y_{ij}\, y_{kl}\mathbf{E}\,[z_{ij}^{ab}\, z_{kl}^{cd}].$$

First note that $y_{ij} = 0$ whenever $i = j$. Also, observe that the sum in the right hand side of equation above can be rewritten as sum of seven summations which are grouped with respect to the indices of $y_{ij}\, y_{kl}$ as follows:

$$\begin{aligned}&\sum_{i\neq j} y_{ij}^2\mathbf{E}\,[z_{ij}^{ab} z_{ij}^{cd}] + \sum_{i\neq j} y_{ij}\, y_{ji}\mathbf{E}\,[z_{ij}^{ab} z_{ji}^{cd}] + \sum_{i\neq j\neq k} y_{ij}\, y_{ik}\mathbf{E}\,[z_{ij}^{ab} z_{ik}^{cd}] + \sum_{i\neq j\neq k} y_{ij}\, y_{ki}\mathbf{E}\,[z_{ij}^{ab} z_{ki}^{cd}]\\ &+ \sum_{i\neq j\neq k} y_{ij}\, y_{jk}\mathbf{E}\,[z_{ij}^{ab} z_{jk}^{cd}] + \sum_{i\neq j\neq k} y_{ij}\, y_{kj}\mathbf{E}\,[z_{ij}^{ab} z_{kj}^{cd}] + \sum_{i\neq j\neq k\neq l} y_{ij}\, y_{kl}\mathbf{E}\,[z_{ij}^{ab} z_{kl}^{cd}].\end{aligned}$$

Clearly we have $y_{ij}^2 = y_{ij}$ and $y_{ij}\, y_{ji} = y_{ij}\, y_{ik} = 0$, so we may ignore the second and the third summands above. Then we have

$$\sum_{i\neq j} y_{ij}^2 = n - R.$$

Let w_{ij} be the indicator of the event $NN(z_i) = z_j$. Then observe that $y_{ij} = w_{ij}(1 - w_{ji})$ and $w_{ij}w_{ik} = 0$ whenever $j \neq k$. So, $y_{ij}y_{ki} = w_{ij}w_{ki} - w_{ij}w_{ji}w_{ki}$, and summing over $i \neq j \neq k$ yields

$$\sum_{i\neq j\neq k} y_{ij}y_{ki} = (n - R) - T = n - R - T.$$

Similarly we get

$$\sum_{i\neq j\neq k} y_{ij}y_{jk} = n - R - T.$$

Also, $y_{ij}y_{kj} = w_{ij}w_{kj} - w_{ij}w_{ji}w_{kj} - w_{ij}w_{kj}w_{jk}$ and summing up over $i \neq j \neq k$ gives

$$\sum_{i\neq j\neq k} y_{ij}y_{kj} = Q - 2T.$$

Moreover, we have

$$\begin{aligned}\sum_{i\neq j\neq k\neq l} y_{ij}y_{kl} &= (n - R)^2 - ((n - R) + 2(n - R - T) + (Q - 2T)) \\ &= n^2 - n(2R + 3) + R^2 + 3R + 4T - Q.\end{aligned}$$

There are again seven cases for the pairs a, b and c, d. Since

$$\mathbf{E}[z_{ij}^{ab}z_{ij}^{ab}] = p_{ab}, \mathbf{E}[z_{ij}^{ab}z_{ki}^{ab}] = 0, \mathbf{E}[z_{ij}^{ab}z_{jk}^{ab}] = 0, \mathbf{E}[z_{ij}^{ab}z_{kj}^{ab}] = p_{aab},$$
$$\mathbf{E}[z_{ij}^{ab}z_{kl}^{ab}] = p_{aabb},$$

we have

$$\begin{aligned}\mathbf{E}[N_{nr}^{ab}N_{nr}^{ab}] &= (n - R)p_{ab} + (Q - 2T)p_{aab} \\ &\quad + (n^2 - n(2R + 3) + R^2 + 3R + 4T - Q)p_{aabb}. \qquad (18)\end{aligned}$$

Since

$$\mathbf{E}[z_{ij}^{ab}z_{ij}^{ba}] = 0, \mathbf{E}[z_{ij}^{ab}z_{ki}^{ba}] = p_{abb}, \mathbf{E}[z_{ij}^{ab}z_{jk}^{ba}] = p_{aab}, \mathbf{E}[z_{ij}^{ab}z_{kj}^{ba}] = 0,$$
$$\mathbf{E}[z_{ij}^{ab}z_{kl}^{ba}] = p_{aabb},$$

we have

$$\mathbf{E}\,[N_{nr}^{ab} N_{nr}^{ba}] = (n - R - T)(p_{abb} + p_{aab}) + (n^2 - n(2R + 3) + R^2 + 3R + 4T - Q)\,p_{aabb}. \tag{19}$$

Since

$$\mathbf{E}\,[z_{ij}^{ab} z_{ij}^{ac}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{ki}^{ac}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{jk}^{ac}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{kj}^{ac}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{kl}^{ac}] = p_{aabc},$$

we get

$$\mathbf{E}\,[N_{nr}^{ab} N_{nr}^{ac}] = (n^2 - n(2R + 3) + R^2 + 3R + 4T - Q)\,p_{aabc}. \tag{20}$$

Since

$$\mathbf{E}\,[z_{ij}^{ab} z_{ij}^{ca}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{ki}^{ca}] = p_{abc}, \mathbf{E}\,[z_{ij}^{ab} z_{jk}^{ca}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{kj}^{ca}] = 0,$$
$$\mathbf{E}\,[z_{ij}^{ab} z_{kl}^{ca}] = p_{aabc},$$

we obtain

$$\mathbf{E}\,[N_{nr}^{ab} N_{nr}^{ca}] = (n - R - T)\,p_{abc} + (n^2 - n(2R + 3) + R^2 + 3R + 4T - Q)\,p_{aabc}. \tag{21}$$

As

$$\mathbf{E}\,[z_{ij}^{ab} z_{ij}^{bc}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{ki}^{bc}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{jk}^{bc}] = p_{abc}, \mathbf{E}\,[z_{ij}^{ab} z_{kj}^{bc}] = 0,$$
$$\mathbf{E}\,[z_{ij}^{ab} z_{kl}^{bc}] = p_{abbc},$$

we obtain

$$\mathbf{E}\,[N_{nr}^{ab} N_{nr}^{bc}] = (n - R - T)\,p_{abc} + (n^2 - n(2R + 3) + R^2 + 3R + 4T - Q)\,p_{abbc}. \tag{22}$$

As

$$\mathbf{E}\,[z_{ij}^{ab} z_{ij}^{cb}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{ki}^{cb}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{jk}^{cb}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{kj}^{cb}] = p_{abc},$$
$$\mathbf{E}\,[z_{ij}^{ab} z_{kl}^{cb}] = p_{abbc},$$

we have

$$\mathbf{E}\,[N_{nr}^{ab} N_{nr}^{cb}] = (Q - 2T)\,p_{abc} + (n^2 - n(2R + 3) + R^2 + 3R + 4T - Q)\,p_{abbc}. \tag{23}$$

Finally, since

$$\mathbf{E}\,[z_{ij}^{ab} z_{ij}^{cd}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{ki}^{cd}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{jk}^{cd}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{kj}^{cd}] = 0, \mathbf{E}\,[z_{ij}^{ab} z_{kl}^{cd}] = p_{abcd},$$

we have

$$\mathbf{E}\,[N_{nr}^{ab} N_{nr}^{cd}] = (n^2 - n(2R+3) + R^2 + 3R + 4T - Q)\,p_{abcd}. \tag{24}$$

Adding up the Eqs. (18)–(24) over pairwise distinct a, b, c, d, we get

$$\begin{aligned}\mathbf{E}\,[N_{m,nr}^2] &= (n-R)P_{ab} + 2(n-R-T)(P_{aab}+P_{abc}) + (Q-2T)(P_{aab}+P_{abc})\\ &\quad + (n^2 - n(2R+3) + R^2 + 3R + 4T - Q)(2P_{aabb} + 4P_{aabc} + P_{abcd}),\end{aligned}$$

and hence

$$\begin{aligned}\mathbf{Var}\,[N_{m,nr}] = (n-R)^2&(2P_{aabb} + 4P_{aabc} + P_{abcd} - P_{ab}^2)\\ &+ (n-R)P_{ab} + (2n - 2R + Q - 4T)(P_{aab} + P_{abc})\\ &+ (-3n + 3R - Q + 4T)(2P_{aabb} + 4P_{aabc} + P_{abcd}).\end{aligned}$$

Covariance of $N_{s,r}$ and $N_{m,nr}$

Since $N_{s,r} = \sum_{a=1}^{m} N_r^a$ and $N_{m,nr} = \sum_{1\le u\ne v\le m} N_{nr}^{uv}$, we first compute $\mathbf{E}\,[N_r^a N_{nr}^{uv}]$.

$$\mathbf{E}\,[N_r^a N_{nr}^{uv}] = \mathbf{E}\left[\sum_{i,j} r_{ij} z_{ij}^a \sum_{k,l} y_{kl} z_{kl}^{uv}\right] = \sum_{i,j,k,l} r_{ij} y_{kl} \mathbf{E}\,[z_{ij} z_{kl}^{uv}].$$

Observe that $r_{ij} y_{ij} = r_{ij} y_{ji} = r_{ij} y_{ik} = r_{ij} y_{jk} = 0$, because each point has only one NN. Therefore the sum above can be written as sum of three summations as follows;

$$\sum_{i\ne j\ne k} r_{ij} y_{ki} \mathbf{E}\,[z_{ij}^a z_{ki}^{uv}] + \sum_{i\ne j\ne k} r_{ij} y_{kj} \mathbf{E}\,[z_{ij}^a z_{kj}^{uv}] + \sum_{i\ne j\ne k\ne l} r_{ij} y_{kl} \mathbf{E}\,[z_{ij}^a z_{kl}^{uv}].$$

Note that $r_{ij} = w_{ij} w_{ji}$, and recall $y_{kl} = w_{kl}(1 - w_{lk})$ and $w_{ij} w_{ik} = 0$. Thus, $r_{ij} y_{ki} = w_{ij} w_{ji} w_{ki}$ and we get $\sum_{i\ne j\ne k} r_{ij} y_{ki} = T$. Similarly we obtain $\sum_{i\ne j\ne k} r_{ij} y_{kj} = T$, and so

$$\sum_{i\ne j\ne k\ne l} r_{ij} y_{kl} = \sum_{i\ne j} r_{ij} \sum_{k\ne l} y_{kl} - 2T = R(n-R) - 2T.$$

Since $\mathbf{E}\,[z_{ij}^a z_{ki}^{ab}] = 0$, $\mathbf{E}\,[z_{ij}^a z_{kj}^{ab}] = 0$ and $\mathbf{E}\,[z_{ij}^a z_{kl}^{ab}] = p_{aaab}$ for $a \ne b$, we have

$$\mathbf{E}\,[N_r^a N_{nr}^{ab}] = (R(n-R) - 2T)\,p_{aaab}. \tag{25}$$

Since $\mathbf{E}[z^a_{ij}z^{ba}_{ki}] = p_{aab}$, $\mathbf{E}[z^a_{ij}z^{ba}_{kj}] = p_{aab}$ and $\mathbf{E}[z^a_{ij}z^{ba}_{kl}] = p_{aaab}$ for $a \neq b$, we obtain

$$\mathbf{E}[N^a_r N^{ba}_{nr}] = 2Tp_{aab} + (R(n-R) - 2T)p_{aaab}. \tag{26}$$

As $\mathbf{E}[z^a_{ij}z^{bc}_{ki}] = 0$, $\mathbf{E}[z^a_{ij}z^{bc}_{kj}] = 0$ and $\mathbf{E}[z^a_{ij}z^{bc}_{kl}] = p_{aabc}$ for $a \neq b \neq c$, we get

$$\mathbf{E}[N^a_r N^{bc}_{nr}] = (R(n-R) - 2T)p_{aabc}, \tag{27}$$

and hence summing up the equations in (25)–(27) over all $a \neq b \neq c$ gives

$$\mathbf{E}[N_{s,r}N_{m,nr}] = (R(n-R) - 2T)(2P_{aaab} + P_{aabc}) + 2TP_{aab}.$$

Therefore, we obtain

$$\begin{aligned}\mathbf{Cov}[N_{s,r}, N_{m,nr}] &= (R(n-R) - 2T)(2P_{aaab} + P_{aabc}) \\ &\quad + 2TP_{aab} - R(n-R)P_{aa}P_{ab} \\ &= R(n-R)(2P_{aaab} + P_{aabc} - P_{aa}P_{ab}) \\ &\quad + 2T(P_{aab} - 2P_{aaab} - P_{aabc}).\end{aligned}$$

Asymptotic (joint) distribution of $N_{s,r}$ and $N_{m,nr}$

We show that the asymptotic joint distribution of $N_{s,r}$ and $N_{m,nr}$ is bivariate normal under RL where Q, R and T are fixed quantities. This is required for the test $\mathcal{X}^2_R$ in Eq. (7) to have a χ^2 distribution in the limit. In what follows, we assume that $R/n \to r$, $Q/n \to q$, $T/n \to t$ and $n_i/n \to \lambda_i$ for all $1 \leq i \leq m$, as $n \to \infty$.

At first glance, each of **Var** $(N_{s,r})$, **Var** $(N_{m,nr})$ and **Cov** $(N_{s,r}, N_{m,nr})$ looks like of order n^2 (i.e., $O(n^2)$). However, each one of them is of order n. To show that, it suffices to prove that the coefficients of the terms of order n^2 goes to 0 as $n \to \infty$. Note that as $n \to \infty$ we have $P_{aa} \to \sum_{1\leq a\leq m} \lambda^2_a$, $P_{ab} \to \sum_{1\leq a\neq b\leq m} \lambda_a\lambda_b$, $P_{aab} \to \sum_{1\leq a\neq b\leq m} \lambda^2_a\lambda_b$ and so on. Then, as $n \to \infty$ we get

$$P_{aaaa} + P_{aabb} - (P_{aa})^2 \to \sum_{1\leq a\leq m} \lambda^4_a + \sum_{1\leq a\neq b\leq m} \lambda^2_a\lambda^2_b - \left(\sum_{1\leq a\leq m} \lambda^2_a\right)^2 = 0,$$

$$\begin{aligned}2P_{aabb} + 4P_{aabc} + P_{abcd} - (P_{ab})^2 &\to 2\sum_{1\leq a\neq b\leq m} \lambda^2_a\lambda^2_b + 4\sum_{1\leq a\neq b\neq c\leq m} \lambda^2_a\lambda_b\lambda_c \\ &\quad + \sum_{1\leq a\neq b\neq c\neq d\leq m} \lambda_a\lambda_b\lambda_c\lambda_d - \left(\sum_{1\leq a\neq b\leq m} \lambda_a\lambda_b\right)^2 = 0,\end{aligned}$$

$$2P_{aaab} + P_{aabc} - P_{aa}P_{ab} \to 2 \sum_{1\le a\ne b\le m} \lambda_a^3\lambda_b + \sum_{1\le a\ne b\ne c\le m} \lambda_a^2\lambda_b\lambda_c - \left(\sum_{1\le a\le m}\lambda_a^2\right)\left(\sum_{1\le a\ne b\le m}\lambda_a\lambda_b\right) = 0.$$

Therefore, $\mathbf{Var}\,(N_{s,r})/n \to \sigma_1^2$, $\mathbf{Var}\,(N_{m,nr})/n \to \sigma_2^2$ and $\mathbf{Cov}\,(N_{s,r}, N_{m,nr})/n \to \rho$ for some constants σ_1^2, σ_2^2 and ρ.

We next prove the asymptotic joint normality of self reflexive and mixed non-reflexive counts. By Cramér-Wold device, it suffices to prove the asymptotic normality of any linear combination of $N_{s,r}$ and $N_{m,nr}$, $c_1N_{s,r} + c_2N_{m,nr}$, where c_1 and c_2 are arbitrary real numbers. That is,

$$\frac{c_1N_{s,r} + c_2N_{m,nr} - c_1\mathbf{E}\,[N_{s,r}] - c_2\mathbf{E}\,[N_{m,nr}]}{\sigma\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1),$$

where $\sigma^2 = c_1^2\sigma_1^2 + 2c_1c_2\rho + c_2^2\sigma_2^2$, $\mathcal{N}(0, 1)$ is the standard normal distribution and $\xrightarrow{\mathcal{L}}$ stands for convergence in law. Notice that we may assume that $\sigma^2 > 0$ since otherwise $(c_1N_{s,r} + c_2N_{m,nr} - c_1\mathbf{E}\,[N_{s,r}] - c_2\mathbf{E}\,[N_{m,nr}])/\sqrt{n}$ converges in law to the constant random variable 0 (i.e., degenerate normal distribution with mean and variance 0).

Let $a_{ij} = c_1r_{ij} - c_2(y_{ij} + y_{ji})/2$ for every $1 \le i, j \le n$. Fix a labeling of $z_1, \ldots, z_n$ and let $b_{ij} = \mathbf{1}_{\{L_i = L_j\}}$ for all $1 \le i, j \le n$. Then, observe that

$$U := U(\pi) := \sum_{1\le i\ne j\le n} a_{ij}b_{\pi(i)\pi(j)} \stackrel{d}{=} c_1N_{s,r} - c_2N_{s,nr}, \tag{28}$$

where π is uniformly distributed over all permutations of $\{1, 2, \ldots, n\}$ and $\stackrel{d}{=}$ denotes equality in distribution. Moreover, note that (a_{ij}) and (b_{ij}) are symmetric $n \times n$ matrices. Let $s^2 = \mathbf{Var}\,(U)$ and note that by (28) we have $s^2/n \to \sigma^2 > 0$ as $n \to \infty$. Also, recall that $N_{s,nr} + N_{m,nr} = n - R$ which is non-random. Therefore, by (28) we see that U is asymptotically normal if and only if

$$c_1N_{s,r} - c_2N_{s,nr} + c_2(N_{s,nr} + N_{m,nr}) = c_1N_{s,r} + c_2N_{m,nr}$$

is asymptotically normal. To prove that U is asymptotically normally distributed we use Corollary 2.1 in Barbour and Eagleson (1986).

Let $(n)_k$ denote $n(n-1)\cdots(n-k+1)$ and

$$a_1 = \frac{1}{(n)_2}\sum_{1\le i\ne j\le n} |a_{ij}|, \qquad a_2 = \frac{1}{(n)_3}\sum_{1\le i\ne j\ne k\le n} |a_{ij}a_{ik}|,$$

$$a_3 = \frac{1}{(n)_4}\sum_{1\le i\ne j\ne k\ne l\le n} |a_{ij}a_{ik}a_{il}|, \qquad a_4 = \frac{1}{(n)_4}\sum_{1\le i\ne j\ne k\ne l\le n} |a_{ij}a_{ik}a_{jl}|,$$

$$a_5 = \frac{1}{(n)_3} \sum_{1 \le i \ne j \ne k \le n} |a_{ij}^2 a_{ik}|, \qquad a_6 = \frac{1}{(n)_2} \sum_{1 \le i \ne j \le n} |a_{ij}|^3,$$
$$a_7 = \frac{1}{(n)_3} \sum_{1 \le i \ne j \ne k \le n} |a_{ij} a_{ik} a_{jk}|, \qquad a_8 = \frac{1}{(n)_2} \sum_{1 \le i \ne j \le n} a_{ij}^2.$$

Similarly, define $b_1, \ldots, b_8$ for b_{ij}'s and let

$$\epsilon = s^{-3}(n^4(a_1^3 + a_1a_2 + a_3 + a_4)(b_1^3 + b_1b_2 + b_3 + b_4) \\ + n^3(a_5 + a_1a_8)(b_5 + b_1b_8) + n^2a_6b_6).$$

Barbour and Eagleson (1986) show that if $\epsilon \to 0$ as $n \to \infty$, then U is asymptotically normal. (Notice that a_7 and b_7 do not appear in ϵ, but these quantities are used in the proof of Theorem 2.1 in Barbour and Eagleson (1986) and hence, to be consistent with their notation we define both quantities.)

Let $C = \max\{|c_1|, |c_2|\}$. It is easy to see that a point in the sample on a plane can be the NN of at most 6 points. Then we have $a_1 \le 7C/(n-1) \le K_1/n$ where $K_1 = 14C$. Similarly, one can obtain $a_2, a_5 \le K_2/n^2$, $a_3, a_4 \le K_3/n^3$ and $a_6, a_8 \le K_4/n$ for some constants K_2, K_3 and K_4. On the other hand, $|b_{ij}| \le 1$ for each i and j and hence, we get $b_i \le 1$ for every $1 \le i \le 8$. Therefore, we have that ϵ is $O(1/\sqrt{n})$ since s is of order $\sqrt{n}$, and thus the desired result follows.

References

Agresti A (1992) A survey of exact inference for contingency tables. Stat Sci 7(1):131–153

Baddeley AJ, Turner R (2005) spatstat: an R package for analyzing spatial point patterns. J Stat Softw 12(6):1–42

Barbour AD, Eagleson GK (1986) Random association of symmetric arrays. Stoch Anal Appl 4(3):239–281

Ceyhan E (2008) Overall and pairwise segregation tests based on nearest neighbor contingency tables. Comput Stat Data Anal 53(8):2786–2808

Ceyhan E (2010a) Exact inference for testing spatial patterns by nearest neighbor contingency tables. J Probab Stat Sci 8(1):45–68

Ceyhan E (2010b) On the use of nearest neighbor contingency tables for testing spatial segregation. Environ Ecol Stat 17(3):247–282

Ceyhan E (2014) Nearest neighbor methods for testing reflexivity and species-correspondence. arXiv:1405.3689 [stat.ME]. Technical report # KU-EC-14-1, Koç University, Istanbul

Cox TF (1981) Reflexive nearest neighbours. Biometrics 37(2):367–369

Diggle PJ (1979) On parameter-estimation and goodness-of-fit testing for spatial point patterns. Biometrics 35(1):87–101

Diggle PJ (2003) Statistical analysis of spatial point patterns, 2nd edn. Hodder Arnold, London

Dixon PM (1994) Testing spatial segregation using a nearest-neighbor contingency table. Ecology 75(7):1940–1948

Dixon PM (2002a) Nearest-neighbor contingency table analysis of spatial segregation for several species. Ecoscience 9(2):142–151

Dixon PM (2002) Nearest neighbor methods. In: El-Shaarawi Abdel H, Piegorsch WW (eds) Encyclopedia of environmetrics, vol 3. Wiley, New York, pp 1370–1383

Dixon PM (2002) Ripley's K function. In: El-Shaarawi Abdel H, Piegorsch WW (eds) Encyclopedia of environmetrics, vol 3. Wiley, New York, pp 1796–1803

Freeman S (2002) Biological science. Prentice Hall, Upper Saddle River

Good BJ, Whipple SA (1982) Tree spatial patterns: South Carolina bottomland and swamp forests. Bull Torrey Bot Club 109(4):529–536
Goreaud F, Pélissier R (2003) Avoiding misinterpretation of biotic interactions with the intertype K_{12}-function: population independence vs. random labelling hypotheses. J Veg Sci 14(5):681–692
Henze N (1987) On the fraction of random points with specified nearest-neighbour interactions and degree of attraction. Adv Appl Probab 19(4):873–895
Illian J, Penttinen A, Stoyan H, Stoyan D (2008) Statistical analysis and modelling of spatial point patterns. Wiley, Chichester
Kulldorff M (2006) Tests for spatial randomness adjusted for an inhomogeneity: a general framework. J Am Stat Assoc 101(475):1289–1305
Laskurain NA (2008) Dinámica espacio-temporal de un bosque secundario en el Parque Natural de Urkiola (Bizkaia). PhD thesis, Universidad del País Vasco/Euskal Herriko Unibertsitatea
Lindenmayer D, Burgman M (2005) Practical conservation biology. CSIRO Publishing, Collingwood
Meagher TR, Burdick DS (1980) The use of nearest neighbor frequency analysis in studies of association. Ecology 61(5):1253–1255
Moser BK (1996) Linear models: a mean model approach. Academic Press, Cambridge
Pielou EC (1961) Segregation and symmetry in two-species populations as studied by nearest-neighbor relationships. J Ecol 49(2):255–269
Ranker TA, Haufler CHE (2008) Biology and evolution of ferns and lycophytes. Cambridge University Press, Cambridge
Ripley BD (2004) Spatial statistics, 2nd edn. Wiley-Interscience, New York
Schilling MF (1986) Mutual and shared neighbor probabilities: finite- and infinite-dimensional results. Adv Appl Probab 18(2):388–405
Searle SR (2006) Matrix algebra useful for statistics. Wiley-Interscience, New York
Stoyan D, Stoyan H (1994) Fractals, random shapes and point fields: methods of geometrical statistics. Wiley, New York
van Lieshout MNM, Baddeley AJ (1999) Indices of dependence between types in multivariate point patterns. Scand J Stat 26(4):511–532

Elvan Ceyhan is visiting associate professor at the Department of Statistics, University of Pittsburgh, Pittsburgh, PA (elvanceyhan@gmail.com). His research interests include spatial data and pattern analysis (by nearest neighbor and graph theoretic methods) and their applications; random geometric graphs called proximity catch digraphs (PCDs), their construction and characterization; domination, edge and arc density, esp. for PCDs; pattern recognition and classification; and statistical methods for medical data and image analysis.

Selim Bahadır has graduated in July 2016 from the doctoral program of Department of Mathematics, Koç University, Istanbul, Turkey (selimbahadir@ku.edu.tr). His research interests include random k nearest neighbor digraphs and their applications in spatial data analysis, random directed graphs and their classifications.