Discrete Manhattan and Chebyshev pair correlation functions in k dimensions

Alexander Lai De Oliveira and Benjamin J. Binder *
School of Mathematical Sciences, University of Adelaide, Adelaide 5005, Australia

(Received 5 January 2020; accepted 24 June 2020; published 13 July 2020)

Pair correlation functions provide a summary statistic which quantifies the amount of spatial correlation between objects in a spatial domain. While pair correlation functions are commonly used to quantify continuous-space point processes, the on-lattice discrete case is less studied. Recent work has brought attention to the discrete case, wherein on-lattice pair correlation functions are formed by normalizing empirical pair distances against the probability distribution of random pair distances in a lattice with Manhattan and Chebyshev metrics. These distance distributions are typically derived on an *ad hoc* basis as required for specific applications. Here we present a generalized approach to deriving the probability distributions of pair distances in a lattice with discrete Manhattan and Chebyshev metrics, extending the Manhattan and Chebyshev pair correlation functions to lattices in k dimensions. We also quantify the variability of the Manhattan and Chebyshev pair correlation functions, which is important to understanding the reliability and confidence of the statistic.

DOI: 10.1103/PhysRevE.102.012130

I. INTRODUCTION

Many biological and physical processes exhibit spatial patterning through the aggregation and segregation of objects or agents in a spatial domain. For example, considering cells as agents, spatial patterning via cell clustering can commonly be found in cell biology, where cultures of cells are grown *in vitro* for cancer research, developmental biology, and tissue engineering [1–7]. More generally, spatial patterning naturally manifests in pigmentation of animal skins [8], resource competition in ecology [9–14], and in particle physics and fluid mixing [15–20]. Quantifying spatial patterning can help with distinguishing different mechanisms and processes in physical systems. Therefore, the development and study of spatial statistics are of great relevance and importance to our understanding of mechanisms and processes in physical systems.

Many spatial statistics have been developed to quantify spatial correlations [9,21–24]. In this work, we restrict our attention to pair correlation functions [11,25–45]. A pair correlation function is a summary statistic which measures the amount of clustering in a spatial domain. More precisely, for a fixed distance in a spatial domain, a pair correlation function returns unity if the proportion of distances over all pairs of agents in a sample is equal to the proportion of distances which would occur in a system of agents spatially distributed uniformly at random. Values greater than unity indicate aggregation of agents and values less than unity indicate segregation of agents. One may think of a pair correlation function simply as a ratio between an empirical count of a fixed distance between all pairs of agents in a sample to the expected value of that distance count if all agents were distributed uniformly at random. We refer to this expected value as the normalization factor of the pair correlation function.

While pair correlation functions are studied extensively in their applications to spatially continuous point processes [34,38–40,43–45], their application to spatially discrete exclusion processes is relatively new [27–31]. Discrete pair correlation functions can be constructed by dividing the empirical count of distances between all pairs of agents in some on-lattice domain by an appropriate probability distribution of frequencies of pair distances in a lattice. Such a probability distribution arises from a choice of metric on the lattice. In the continuous setting, the most natural metric is the standard Euclidean metric induced by the ℓ_2 norm. However, carrying this over to the discrete setting presents some difficulties. As Markham *et al.* [26] and Gavagnin *et al.* [30] both point out, the distances between lattice sites increase irregularly when using the standard Euclidean metric. Markham *et al.* [26] overcome this issue by considering a partial differential equation representation of a system of ordinary differential equations governing nearest neighbor correlations, extending previous work by Baker and Simpson [25]. Gavagnin *et al.* [30] shift their focus away from the standard Euclidean metric since the irregular distance spacing causes the normalization factors in their approach to be inaccurate, making results hard to interpret. Gavagnin *et al.* [30] instead introduce new metrics whose distances on a lattice are always given in terms of integers, alleviating the issues of irregular distance spacing. The two main metrics introduced are the Manhattan (square taxicab) metric and the Chebyshev (square uniform) metric, induced by the ℓ_1 norm and ℓ_∞ norm, respectively. Our work focuses primarily on these two metrics, which we illustrate schematically in Fig. 1.

Binder and Simpson [27] first studied the case of one-dimensional pair correlation functions on a square lattice, and its application to several different types (e.g., scratch assay, multicellular aggregation, and evenly distributed single cells) of processed experimental images. Further developments of the approach by Binder and Simpson [27] included extension to periodic boundary conditions [28], as well as using a dis-

*benjamin.binder@adelaide.edu.au

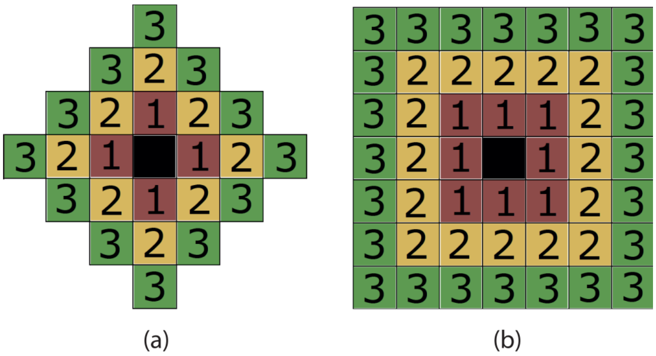

FIG. 1. Schematic of the Manhattan and Chebyshev distance metric in two dimensions: (a) Manhattan; (b) Chebyshev.

crete spectral analysis of the one-dimensional pair correlation function to objectively determine the size of the bin width, or bandwidth, in the analysis of experimental images of highly proliferative 231 breast cancer cells and highly motile murine 3T3 fibroblast cells [29].

Gavagnin *et al.* [30] further extended Binder and Simpson's [27] one-dimensional discrete pair correlation functions to Manhattan and Chebyshev distances in two and three dimensions by determining the frequencies of pair distances in a lattice. They compared the Manhattan and Chebyshev pair correlation functions with the one-dimensional pair correlation function for three highly regular spatial patterns (diagonal stripes, checkerboard, and concentric ring patterns). Their analysis showed that the one-dimensional pair correlation function was unable to detect the spatial correlation in the patterns, whereas both the Manhattan and Chebyshev pair correlation functions successfully detected the nonuniform spatial correlations. However, while the derivation of their method is elegant, it does not cover the full range of distances possible under these metrics. In particular, their analytic formulas for a $X \times Y$ square lattice are valid only for distances $s \leqslant \min\{X, Y\}$ under nonperiodic boundary conditions and $s \leqslant \min\{\lfloor X/2 \rfloor, \lfloor Y/2 \rfloor\}$ under periodic boundary conditions. In the best case scenario where $X = Y$, this is only half the domain of valid distances under the Manhattan metric. Therefore, while the work of Gavagnin *et al.* [30] can analyze distances at short to moderate length scales, it cannot be applied to long length scales. Recently, the remaining part of the two-dimensional nonperiodic Manhattan distance distribution has been obtained by Johnston and Crampin [31]. However, the remaining part of the periodic cases for both Manhattan and Chebyshev distances, as well as the nonperiodic Chebyshev case, have not yet been derived.

We extend the work of Gavagnin *et al.* [30] and Johnston and Crampin [31] by presenting an alternative method of obtaining the discrete Manhattan and Chebyshev distance distributions. In particular, we provide analytic formulas for the remaining portions of the two-dimensional distance distribution for the periodic Manhattan distribution, and both periodic and nonperiodic Chebyshev distributions. Moreover, in this work we derive explicit formulas for computing discrete Manhattan and Chebyshev distance distributions (periodic and nonperiodic), valid over the whole domain of any square lattice and also generalizable to k-dimensional lattices. We also examine the variability of the pair correlation functions over their domain of distances, which is important to establishing the reliability and confidence of the statistic. Hence variability analysis, which has thus far been neglected, is critical when using pair correlation functions in physical applications.

II. PAIR CORRELATION FUNCTIONS

In this section we consider an approach to obtaining the normalization factors used in the Manhattan and Chebyshev pair correlation functions, obtaining explicit formulas in two dimensions. For the case of the two-dimensional square lattice, Gavagnin *et al.* [30] and Johnston and Crampin [31] both used geometric approaches to determine the total number of pair distances between unoccupied sites in the lattice, providing the normalization factor for the pair correlation function. We will instead use a probabilistic approach which can be used for straightforward numerical calculations of the Manhattan and Chebyshev normalization factors in k dimensions, or as a means for deriving explicit closed-form formulas of those normalization factors if desired.

We first define the k-dimensional discrete pair correlation function, which generalizes the two-dimensional pair correlation function used in the previous works of Binder and Simpson [27], Gavagnin *et al.* [30], and Johnston and Crampin [31]. Consider a system of agents in a k-dimensional $v_1 \times \cdots \times v_k$ integer lattice of points $\mathbf{x} = (x_1, \ldots, x_k)$ with unit spacing and $1 \leqslant x_i \leqslant v_i$. We refer to such a lattice as a $\mathbf{v}_k$ lattice, denoted $\mathbb{L}_{\mathbf{v}_k}$, where $\mathbf{v}_k = (v_1, \ldots, v_k)$ is the vector of lattice dimensions. The lattice is equipped with the exclusion property such that a site may be occupied by at most one agent at any given time. Locations of agents in the system can be denoted by the k-dimensional occupancy array $\lambda = (\lambda_{\mathbf{x}})_{\mathbf{x} \in \mathbb{L}_{\mathbf{v}_k}}$ with entries

$$\lambda_{\mathbf{x}} = \begin{cases} 1, & \text{if } \mathbf{x} \text{ is occupied,} \\ 0, & \text{if } \mathbf{x} \text{ is vacant.} \end{cases}$$

The number of agents in the system is

$$N = \sum_{x_k=1}^{v_k} \cdots \sum_{x_1=1}^{v_1} \lambda_{\mathbf{x}} \leqslant \prod_{i=1}^{k} v_i.$$

Our main interest is the frequency of pair distances between distinct points $\mathbf{a} = (a_1, \ldots, a_k)$ and $\mathbf{b} = (b_1, \ldots, b_k)$ in the $\mathbf{v}_k$ lattice $\mathbb{L}_{\mathbf{v}_k}$. For some given distance function $\mathcal{D}(\mathbf{a}, \mathbf{b})$, we can consider the set

$$\mathcal{F}_\lambda(s) = \{(\mathbf{a}, \mathbf{b}) \in \mathbb{L}_{\mathbf{v}_k} \times \mathbb{L}_{\mathbf{v}_k} \mid \mathcal{D}(\mathbf{a}, \mathbf{b}) = s, \lambda_{\mathbf{a}} = \lambda_{\mathbf{b}} = 1, \mathbf{a} \neq \mathbf{b}\}$$

of pairs of distinct points which are separated by distance s in the $\mathbf{v}_k$ lattice $\mathbb{L}_{\mathbf{v}_k}$. The frequency of pair distances s is then given by the cardinality

$$f_\lambda(s) = |\mathcal{F}_\lambda(s)|. \tag{1}$$

To construct a pair correlation function, we require the function $d(s)$ which gives the count of pair distances in a completely filled lattice. More precisely, if $\lambda \equiv 1$, by which we mean $\lambda_{\mathbf{x}} = 1$ for all $\mathbf{x} \in \mathbb{L}_{\mathbf{v}_k}$, then we define

$$d(s) = f_{\lambda \equiv 1}(s). \tag{2}$$

Alternatively, one may regard $d(s)$ as the number of sites in the lattice $\mathbb{L}_{\mathbf{v}_k}$ separated by distance s, without consideration of whether the sites are occupied or unoccupied. In a completely filled lattice, we have $f_\lambda(s)/d(s) = 1$. For a $\mathbf{v}_k$ lattice that is not fully occupied, we will not have $f_\lambda(s)/d(s) = 1$, even when agents are uniformly distributed in the lattice. To adjust for this, if the lattice contains agents distributed uniformly at random, then the expected number of agents in the lattice separated by distance s should be the number of sites $d(s)$ separated by distance s scaled by the probability that any choice of two such sites are actually occupied. Hence we scale $d(s)$ by

$$\rho_2 = \frac{N}{\prod_{i=1}^{k} v_k} \frac{N-1}{\left(\prod_{i=1}^{k} v_k\right) - 1}, \tag{3}$$

the probability of selecting two agents at random from the lattice without replacement [27]. We then define the pair correlation function by

$$P(s) = \frac{f_\lambda(s)}{\rho_2 d(s)}. \tag{4}$$

In practice, the value $f_\lambda(s)$ is obtained from the data and ρ_2 is dependent on the lattice. The only unknown quantity in (4) is $d(s)$. In the proceeding subsections, we derive formulas for $d(s)$ for the discrete Manhattan and Chebyshev distance metrics in k dimensions, considering both nonperiodic and periodic boundary conditions.

A. Nonperiodic k-dimensional Manhattan and Chebyshev distance distributions

We begin by considering the nonperiodic case. In the one-dimensional setting, i.e., $k = 1$, both Manhattan and Chebyshev nonperiodic distance distributions coincide with what has thus far been referred to as the *rectilinear* distance distribution in the literature [30,31]. In our formulation, it makes sense to call this one-dimensional distribution a (one-dimensional) *component* distribution, since it turns out that the k-dimensional Manhattan and Chebyshev distributions are built from these one-dimensional components.

To illustrate this idea, let $\mathbf{a} = (a_1, \dots, a_k)$ and $\mathbf{b} = (b_1, \dots, b_k)$ be two points in a k-dimensional lattice with dimension vector $\mathbf{v}_k = (v_1, \dots, v_k)$. For each $i = 1, \dots, k$, we define the component distance function

$$\mathcal{C}_{v_i}(a_i, b_i) = |a_i - b_i|. \tag{5}$$

The nonperiodic Manhattan distance function $\mathcal{D}^M_{\mathbf{v}_k}$ and the nonperiodic Chebyshev distance function $\mathcal{D}^C_{\mathbf{v}_k}$ can then be written

$$\mathcal{D}^M_{\mathbf{v}_k}(\mathbf{a}, \mathbf{b}) = \mathcal{C}_{v_1}(a_1, b_1) + \cdots + \mathcal{C}_{v_k}(a_k, b_k), \tag{6}$$

$$\mathcal{D}^C_{\mathbf{v}_k}(\mathbf{a}, \mathbf{b}) = \max\{\mathcal{C}_{v_1}(a_1, b_1), \dots, \mathcal{C}_{v_k}(a_k, b_k)\}. \tag{7}$$

We now wish to find the distance between two random points in a k-dimensional lattice. Since the Manhattan and Chebyshev distances are constructed from the one-dimensional component distances, we first consider $k = 1$. Let $X_{i1}, X_{i2} \sim U(1, v_i)$ be independent discrete uniform random variables. Define the random variable $C_i = |X_{i1} - X_{i2}|$; this will be our nonperiodic v_i component. Since we are dealing with a finite lattice, all the probabilities that appear throughout the paper are of the form $d(s)/\prod_{i=1}^{k} v_i^2$. The denominator of this fraction is just a normalizing factor to ensure the distance counts $d(s)$ normalize to unity when summed over all valid s. In particular, for the one-dimensional v_i lattice, we have $d(s) = v_i^2 \mathbb{P}(C_i = s)$. Our primary goal is to obtain a calculable expression for $d(s)$, so we will work with $v_i^2 \mathbb{P}(C_i = s)$ rather than just $\mathbb{P}(C_i = s)$, and similarly in the k-dimensional setting. With this in mind, it can be shown via a geometric argument or direct computation that

$$v_i^2 \mathbb{P}(C_i = s) = \begin{cases} v_i & \text{if } s = 0 \\ 2(v_i - s) & \text{if } 1 \leqslant s \leqslant v_i - 1. \end{cases} \tag{8}$$

The cumulative distribution function, which will be required for the Chebyshev distribution, is therefore

$$v_i^2 \mathbb{P}(C_i \leqslant s) = v_i + 2\sum_{j=1}^{s}(v_i - j) \quad \text{if } 0 \leqslant s \leqslant v_i - 1. \tag{9}$$

We now turn our attention to the general setting of k dimensions. If $\mathbf{v}_k = (v_1, \dots, v_k)$ is a vector containing the dimensions of a k-dimensional square lattice, then, in terms of the independent components C_i, the random Manhattan distance, denoted D^M, and the random Chebyshev distance, denoted D^C, between two random points $\mathbf{X}_1 = (X_{11}, \dots, X_{k1})$ and $\mathbf{X}_2 = (X_{12}, \dots, X_{k2})$ within the k-dimensional lattice is

$$D^M = C_1 + \cdots + C_k,$$

$$D^C = \max\{C_1, \dots, C_k\}.$$

The distribution of D^M is given by the convolution of the components $C_1, \dots, C_k$. Convolving discrete random variables is the same as multiplying their probability generating functions, and the same holds for unnormalized distance counts $d(s)$. So if $G^M_{\mathbf{v}_k}(z) = \sum_{s=0}^{(\sum_{i=1}^{k} v_i)-k} d^M_{\mathbf{v}_k}(s) z^s$ is the generating function for the number of sites $d^M_{\mathbf{v}_k}(s)$ separated by Manhattan distance s in a $\mathbf{v}_k$ lattice, and $G_{v_i}(z) = \sum_{s=0}^{v_i-1} d_{v_i}(s) z^s$ is the generating function for the number of sites $d_{v_i}(s)$ separated by distance s in a one-dimensional lattice of length v_i, then we can recover the probability distribution of D^M from the generating function equation

$$G^M_{\mathbf{v}_k}(z) = \prod_{i=1}^{k} G_{v_i}(z). \tag{10}$$

Our goal is obtaining the coefficients of the generating function $G^M_{\mathbf{v}_k}(z)$; the coefficients for each generating function $G_{v_i}(z)$ are each given by the very simple formula described in (8). Writing (10) explicitly in terms of polynomials in the variable z gives

$$\sum_{s=0}^{(\sum_{i=1}^{k} v_i)-k} \left(\prod_{i=1}^{k} v_i^2\right) \mathbb{P}(D^M = s) z^s = \prod_{i=1}^{k} \sum_{s=0}^{v_i-1} v_i^2 \mathbb{P}(C_i = s) z^s. \tag{11}$$

By equating coefficients, we can determine $d^M_{\mathbf{v}_k}(s) = (\prod_{i=1}^{k} v_i^2)\mathbb{P}(D^M = s)$ from the component distances $d_{v_i}(s) = v_i^2 \mathbb{P}(C_i = s)$. Hence the unnormalized distance counts $d^M_{\mathbf{v}_k}(s) = (\prod_{i=1}^{k} v_i^2)\mathbb{P}(D^M = s)$ in (11) can be calculated numerically by iterative usage of the `conv` function on each nonperiodic component distance in MATLAB with no computational issues.

For the Chebyshev distribution, we first calculate the Chebyshev cumulative distribution function from the component cumulative distribution function (9). Recalling that we multiply each probability mass $\mathbb{P}(D^C = s)$ by the normalizing factor $\prod_{i=1}^k v_i^2$ to obtain the unnormalized distance count, we have

$$\left(\prod_{i=1}^k v_i^2\right)\mathbb{P}(D^C \leqslant s) = \left(\prod_{i=1}^k v_i^2\right)\mathbb{P}(\max\{C_1, \dots, C_k\} \leqslant s) = \left(\prod_{i=1}^k v_i^2\right)\mathbb{P}(C_1 \leqslant s, \dots, C_k \leqslant s).$$

By independence of each C_i, the joint cumulative probability is the product of the cumulative probabilities. Hence

$$\left(\prod_{i=1}^k v_i^2\right)\mathbb{P}(D^C \leqslant s) = \prod_{i=1}^k v_i^2 \mathbb{P}(C_i \leqslant s).$$

Consider the case when $s \geqslant 0$; i.e., when $\mathbb{P}(D^C \leqslant s)$ is nonzero. From (9), we have $\mathbb{P}(C_i \leqslant s) = 1$ if $s > v_i - 1$. Therefore, consider the subset $I_s \subseteq \{1, \dots, k\}$ of natural numbers satisfying $s \leqslant v_i - 1$. Our expression simplifies to

$$\left(\prod_{i=1}^k v_i^2\right)\mathbb{P}(D^C \leqslant s) = \left(\prod_{i\in\{1,\dots,k\}\setminus I_s} v_i^2\right)\prod_{i\in I_s} \times [(2s+1)v_i - s(s+1)].$$

This last expression is easily codable in MATLAB. Moreover, one recovers the Chebyshev distance counts via

$$\left(\prod_{i=1}^k v_i^2\right)\mathbb{P}(D^C = s) = \left(\prod_{i=1}^k v_i^2\right)\mathbb{P}(D^C \leqslant s) - \left(\prod_{i=1}^k v_i^2\right)\mathbb{P}(D^C \leqslant s-1). \quad (12)$$

Hence, in (11) and (12), respectively, we have obtained the nonperiodic Manhattan and Chebyshev distance counts in the k-dimensional setting.

B. The periodic case

The periodic case is entirely analogous to the nonperiodic case. We begin by defining the periodic component distance. Let $\mathbf{a} = (a_1, \dots, a_k)$ and $\mathbf{b} = (b_1, \dots, b_k)$ be two points in a k-dimensional lattice with dimension vector $\mathbf{v}_k$. Define the component periodic distance

$$\mathcal{C}_{v_i}^{\mathcal{P}}(a_i, b_i) := \begin{cases} |a_i - b_i| & \text{if } |a_i - b_i| \leqslant \lfloor v_i/2 \rfloor, \\ v_i - |a_i - b_i| & \text{otherwise.} \end{cases} \quad (13)$$

Here $\mathcal{P}$ indicates that we are operating with periodic boundary conditions. We then define the periodic Manhattan distance function $\mathcal{D}_{\mathbf{v}_k}^{M,\mathcal{P}}$ and periodic Chebyshev distance function $\mathcal{D}_{\mathbf{v}_k}^{C,\mathcal{P}}$ by

$$\mathcal{D}_{\mathbf{v}_k}^{M,\mathcal{P}}(\mathbf{a}, \mathbf{b}) := \mathcal{C}_{v_1}^{\mathcal{P}}(a_1, b_1) + \cdots + \mathcal{C}_{v_k}^{\mathcal{P}}(a_k, b_k), \quad (14)$$

$$\mathcal{D}_{\mathbf{v}_k}^{C,\mathcal{P}}(\mathbf{a}, \mathbf{b}) := \max\left\{\mathcal{C}_{v_1}^{\mathcal{P}}(a_1, b_1), \dots, \mathcal{C}_{v_k}^{\mathcal{P}}(a_k, b_k)\right\}. \quad (15)$$

Let $\mathbf{v}_k = (v_1, \dots, v_k)$ be the lattice dimension vector. Denoting the random periodic component distance by $C_i^{\mathcal{P}}$, the random periodic Manhattan distance by $D^{M,\mathcal{P}}$, and the random periodic Chebyshev distance by $D^{C,\mathcal{P}}$, we have

$$D^{M,\mathcal{P}} = C_1^{\mathcal{P}} + \cdots + C_k^{\mathcal{P}},$$
$$D^{C,\mathcal{P}} = \max\left\{C_1^{\mathcal{P}}, \dots, C_k^{\mathcal{P}}\right\}.$$

So we need only calculate the periodic component distance distributions. Then we may argue analogously to the nonperiodic case to transition from the periodic component distances to the periodic Manhattan and Chebyshev distances. Proceeding with this calculation, the unnormalized periodic distance counts for the periodic component distances are given by

$$v_i^2\mathbb{P}\left(C_i^{\mathcal{P}} = s\right) = \begin{cases} v_i & \text{if } s = 0, \\ 2v_i & \text{if } 1 \leqslant s \leqslant \lfloor v_i/2 \rfloor - 1, \\ 2v_i & \text{if } s = \lfloor v_i/2 \rfloor \text{ and } v_i \text{ odd}, \\ v_i & \text{if } s = \lfloor v_i/2 \rfloor \text{ and } v_i \text{ even}. \end{cases}$$

This gives the periodic cumulative distribution function

$$v_i^2\mathbb{P}\left(C_i^{\mathcal{P}} \leqslant s\right) = \begin{cases} 0 & \text{if } s < 0, \\ (2s+1)v_i & \text{if } 0 \leqslant s \leqslant \lfloor v_i/2 \rfloor - 1, \\ v_i^2 & \text{if } s \geqslant \lfloor v_i/2 \rfloor. \end{cases} \quad (16)$$

The periodic Manhattan distance is found similarly by convolution [see (10) and (11)]. So we have a generating function equation $G_{\mathbf{v}_k}^{M,\mathcal{P}}(z) = \prod_{i=1}^k G_{v_i}^{\mathcal{P}}(z)$, which gives rise to

$$\sum_{s=0}^{\sum_{i=1}^k \lfloor v_i/2 \rfloor}\left(\prod_{i=1}^k v_i^2\right)\mathbb{P}(D^{M,\mathcal{P}} = s)z^s = \prod_{i=1}^k \sum_{s=0}^{\lfloor v_i/2 \rfloor} v_i^2\mathbb{P}\left(C_i^{\mathcal{P}} = s\right)z^s. \quad (17)$$

For $s \geqslant 0$, the periodic Chebyshev cumulative distribution function is

$$\left(\prod_{i=1}^k v_i^2\right)\mathbb{P}(D^{C,\mathcal{P}} \leqslant s) = \left(\prod_{i\in\{1,\dots,k\}\setminus I_s} v_i^2\right)\prod_{i\in I_s}(2s+1)v_i = (2s+1)^{|I_s|}\prod_{i=1}^k v_i \prod_{i\in\{1,\dots,k\}\setminus I_s} v_i,$$

where $I_s \subseteq \{1, \dots, k\}$ is the set of natural numbers i such that $s \leqslant \lfloor v_i/2 \rfloor - 1$. Then the unnormalized distance counts can be obtained by

$$\left(\prod_{i=1}^{k} v_i^2\right)\mathbb{P}(D^{C,\mathcal{P}} = s) = \left(\prod_{i=1}^{k} v_i^2\right)\mathbb{P}(D^{C,\mathcal{P}} \leqslant s) - \left(\prod_{i=1}^{k} v_i^2\right)\mathbb{P}(D^{C,\mathcal{P}} \leqslant s-1). \tag{18}$$

We have now derived the formulas for the k-dimensional Manhattan and Chebyshev distributions: the nonperiodic Manhattan and Chebyshev distributions are given in (11) and (12), while the periodic Manhattan and Chebyshev distributions are given in (16) and (17). These formulas allow us to compute the normalizing factors in both Manhattan and Chebyshev distance metrics for periodic and nonperiodic distances in k dimensions.

The formulas also allow us to derive closed-form formulas. We consider the two-dimensional setting, which is often applicable to biological and ecological systems [27–31].

C. Closed-form formulas in two dimensions

We now show that the probabilistic formulation can also be used to derived closed-form formulas. To demonstrate the approach, we carefully derive the closed-form formula for the two-dimensional nonperiodic Manhattan distance distribution. We omit the details for the derivation of the periodic Manhattan, nonperiodic Chebyshev, and periodic Chebyshev.

Consider a $\mathbf{v}_2 = (v_1, v_2)$ lattice. Denote the number of sites separated by nonperiodic Manhattan distance s by $d^M_{\mathbf{v}_2}(s)$. Recall that the random Manhattan distance $D^M = C_1 + C_2$ can be given by the convolution of the two components C_1 and C_2. The convolution formula can be written as

$$d^M_{\mathbf{v}_2}(s) = v_1^2 v_2^2 \mathbb{P}(C_1 + C_2 = s) = \sum_{n=\max\{0, s-v_2+1\}}^{\min\{s, v_1-1\}} v_1^2\mathbb{P}(C_1 = n) v_2^2 \mathbb{P}(C_2 = s-n).$$

There are four situations that require consideration:

Situation	Lower index	Upper index	Condition
1	0	s	$s \leqslant v_1 - 1, s \leqslant v_2 - 1,$
2	0	$v_1 - 1$	$v_1 - 1 \leqslant s \leqslant v_2 - 1,$
3	$s - v_2 + 1$	s	$v_2 - 1 \leqslant s \leqslant v_1 - 1,$
4	$s - v_2 + 1$	$v_1 - 1$	$s \geqslant v_1 - 1, s \geqslant v_2 - 1.$

Notice that Situation 2 and Situation 3 cannot be simultaneously true. By imposing the condition that $v_1 \leqslant v_2$, we make Situation 3 impossible. We may impose this condition without loss of generality since any lattice with $v_1 > v_2$ can be rotated a quarter turn to make $v_1 \leqslant v_2$, and the pair distances remain unchanged under this rotation. The distance counts $d^M_{\mathbf{v}_2}(s)$ are then obtained by restriction to the natural numbers of the function

$$d^M_{\mathbf{v}_2}(s) = \begin{cases} v_1 v_2 & \text{if } s = 0, \\ 2s[2v_1v_2 - (v_1+v_2)s] + \frac{2}{3}s(s^2-1) & \text{if } 0 < s \leqslant v_1, \\ 2v_1^2(v_2 - s) + \frac{2}{3}v_1\left(v_1^2 - 1\right) & \text{if } v_1 \leqslant s \leqslant v_2, \\ \frac{2}{3}(v_1+v_2-s-1)(v_1+v_2-s)(v_1+v_2-s+1) & \text{if } v_2 \leqslant s \leqslant v_1+v_2-2, \end{cases} \tag{19}$$

which is actually continuous as a real-valued function, except at $s = 0$. This formula is consistent with the formulas presented in Gavagnin *et al.* [30] and Johnston and Crampin [31], up to an extra factor of 2 which arises due to our formulation considering symmetric distances as distinct.

We also find that the distance counts for the Chebyshev nonperiodic $d^C_{\mathbf{v}_2}$, Manhattan periodic $d^{M,\mathcal{P}}_{\mathbf{v}_2}$, and Chebyshev periodic $d^{C,\mathcal{P}}_{\mathbf{v}_2}$ are given by the formulas

$$d^C_{\mathbf{v}_2}(s) = \begin{cases} v_1 v_2, & s = 0, \\ 2s[4v_1v_2 - 3(v_1+v_2)s + 2s^2], & 0 < s \leqslant v_1, \\ 2(v_2 - s)v_1^2, & v_1 \leqslant s < v_2, \end{cases} \tag{20}$$

$$d^{M,\mathcal{P}}_{\mathbf{v}_2}(s) = \begin{cases} v_1 v_2, & s = 0, \\ 4v_1v_2 s, & 0 < s < \lfloor v_1/2 \rfloor, \\ 2v_1^2 v_2, & \lfloor v_1/2 \rfloor < s < \lfloor v_2/2 \rfloor, \\ 4v_1v_2\left(\frac{v_1+v_2}{2} - s\right), & \lfloor v_2/2 \rfloor < s < \lfloor v_1/2 \rfloor + \lfloor v_2/2 \rfloor, \end{cases} \tag{21}$$

$$d^{C,\mathcal{P}}_{\mathbf{v}_2}(s) = \begin{cases} v_1 v_2, & s = 0, \\ 8v_1v_2 s, & 0 < s < \lfloor v_1/2 \rfloor, \\ 2v_1^2 v_2, & \lfloor v_1/2 \rfloor < s < \lfloor v_2/2 \rfloor. \end{cases} \tag{22}$$

The periodic formulas have degeneracies at boundary points depending on whether v_1 or v_2 are even or odd. We have thus ignored these points in our closed-form formulas. There are no issues regarding these points when calculating the distributions using the numerical approach.

III. RESULTS

In this section, we illustrate three points of consideration in the study of discrete pair correlation functions. We extend the pair correlation function to k dimensions, explicitly considering the two- and three-dimensional case in our examples. We consider variability of the statistic by examining the behavior of the region between the 2.5th and 97.5th percentiles as we vary the pair distance s. Finally we examine and discuss the applicability of the periodic pair correlation functions.

We first verify that the normalization factors derived in Sec. II are correct by applying the pair correlation function to spatial domains with no correlation. For the two-dimensional case displayed in Fig. 2, we consider the count of pair distances over 1000 realizations of a 60×30 matrix populated with agents spatially distributed uniformly at random, with density $\rho = 0.5$. We extract the average frequency along with the 2.5th and 97.5th percentile for every pair distance possible under each of the four combinations of nonperiodic or periodic, and Manhattan or Chebyshev metrics. We then plot the raw average frequencies alongside the normalized pair correlation function. The three-dimensional case displayed in Fig. 3 follows the same methodology, except using a $60 \times 30 \times 40$ array with density $\rho = 0.01$. The reduced density in the three-dimensional case is for computational purposes; the amount of agents per realization is similar to the two-dimensional case.

In all simulations (Figs. 2 and 3), the average frequencies of pair distances generally converge to their expected values: the unnormalized case shows that these distances converge to the shape of the relevant distance distribution (left column, Figs. 2 and 3), and the normalized case shows convergence to unity (right column, Figs. 2 and 3). There is much to be gained by examining the percentiles of each pair correlation function. For example, in the two-dimensional case, the confidence interval in the nonperiodic cases (right column, top row and bottom-middle row in Fig. 2) is considerably larger than that of the periodic counterparts. In particular, we see that the nonperiodic pair correlation functions are not as reliable as the periodic pair correlation functions at long distances due to the division of small numbers in the statistic; i.e., at distances where $d(s)$, the count of all pair distances s in a completely filled lattice, is small. Moreover, the variability in the three-dimensional case (Fig. 3) is greater than the two-dimensional case (Fig. 2) at each pair distance s among all distance metrics. Notably, the variability increases at small s (right column, Fig. 3); this was not exhibited in the two-dimensional case (right column, Fig. 2). This shows that the three-dimensional pair correlation function is best behaved at medium length scales. It is particularly important that we consider the variability in the statistic when applying and interpreting the signal in practical situations.

There is a striking difference in the behavior of the unnormalized pair distances under the Manhattan and Chebyshev metrics in Fig. 2. This is interesting to note because the Manhattan and Chebyshev distances behave similarly in two dimensions. For example, in the continuous setting, a circle with radius r in the Manhattan metric is a square with side length $\sqrt{2}r$, rotated $\pi/4$ compared to the coordinate axes, while a circle with radius r in the Chebyshev metric is a square with side length $2r$, parallel to the coordinate axes. Rotation should not affect random correlation, which means that only the difference in side length should affect the pair distance frequencies. Despite this, the unnormalized pair distances in Fig. 2 are quite different. The main reason for this is the rectangular 60×30 spatial domain we have chosen; showcasing this was one of our motivations in choosing an irregular domain. We first notice that the Manhattan and Chebyshev distances are quite similarly behaved at short distances. There is some observable difference attributable to the differences in circle side length as previously discussed. The more noticeable differences occur past distance 30 in the nonperiodic case, and past distance 15 in the periodic case. Notice that 30 is the length of the shorter dimension of the 60×30 spatial domain. Since the Chebyshev distance is given by the maximum of the component distances, we know that any pair distance greater than the size of the shorter dimension must have arisen due to the remaining longer dimension. For example, for pair distances exceeding 30 in Fig. 2, we effectively have only one dimension affecting the pair distance frequency, whereas for distances shorter than 30, both dimensions affect the pair distance frequency.

We now examine the applicability of the simultaneous evaluation of the periodic and nonperiodic Manhattan and Chebyshev pair correlation functions. This notion has been previously exploited with a one-dimensional pair correlation function in identifying the necrotic zone boundary in experimental images of tumor spheroids [44]. To focus on the distinction between nonperiodic and periodic pair correlation functions in the present work, we will restrict attention to the Manhattan distance metric. Similar results were found in the case of the Chebyshev metric (not shown).

We examine a pattern consisting of diagonal bars, similar to the approach of Gavagnin *et al.* [30], except we now modify the pattern by thickening one of the diagonal bars (top row, Fig. 4). The choice of a stripy pattern is also partially motivated by its ubiquity in biological settings, for example, in two-dimensional images of the skin of a zebrafish [45]. Evaluating the nonperiodic pair correlation function for the two patterns with a different location of the thicker bar gives two distinct signals (middle row, Fig. 4), indicating that the spatial patterning is different in each example. However, in the periodic case we see that the signals are almost identical (bottom row, Fig. 4). To put it simply, the periodic evaluations cannot distinguish the patterns with a different location of the thicker bar.

On one hand, this can be advantageous in practical settings where a small window is sampled from a larger spatial domain (e.g., a Petri dish in a biological experiment). In this situation, periodic boundary conditions are often assumed [28,45,46], and with this assumption it is desirable that both sample patterns have a similar pair correlation function evaluation. On the other hand, if the samples are representative of the entire spatial domain, such as in the analysis of scratch wound assays [27] or in situations where the domain contains obstructions [31], the nonperiodic evaluations show that these two patterns are distinct. Our results therefore provide the means to examine patterns simultaneously with both periodic and nonperiodic Manhattan and Chebyshev pair correlation functions, and this was not considered in the previous works Gavagnin *et al.* [30] and Johnston and Crampin [31].

IV. DISCUSSION

The study and development of discrete pair correlation functions that quantify spatial patterning within discrete square lattices are applicable to furthering our physical understanding of biological processes, since these functions can be applied directly to images of cell biology experiments

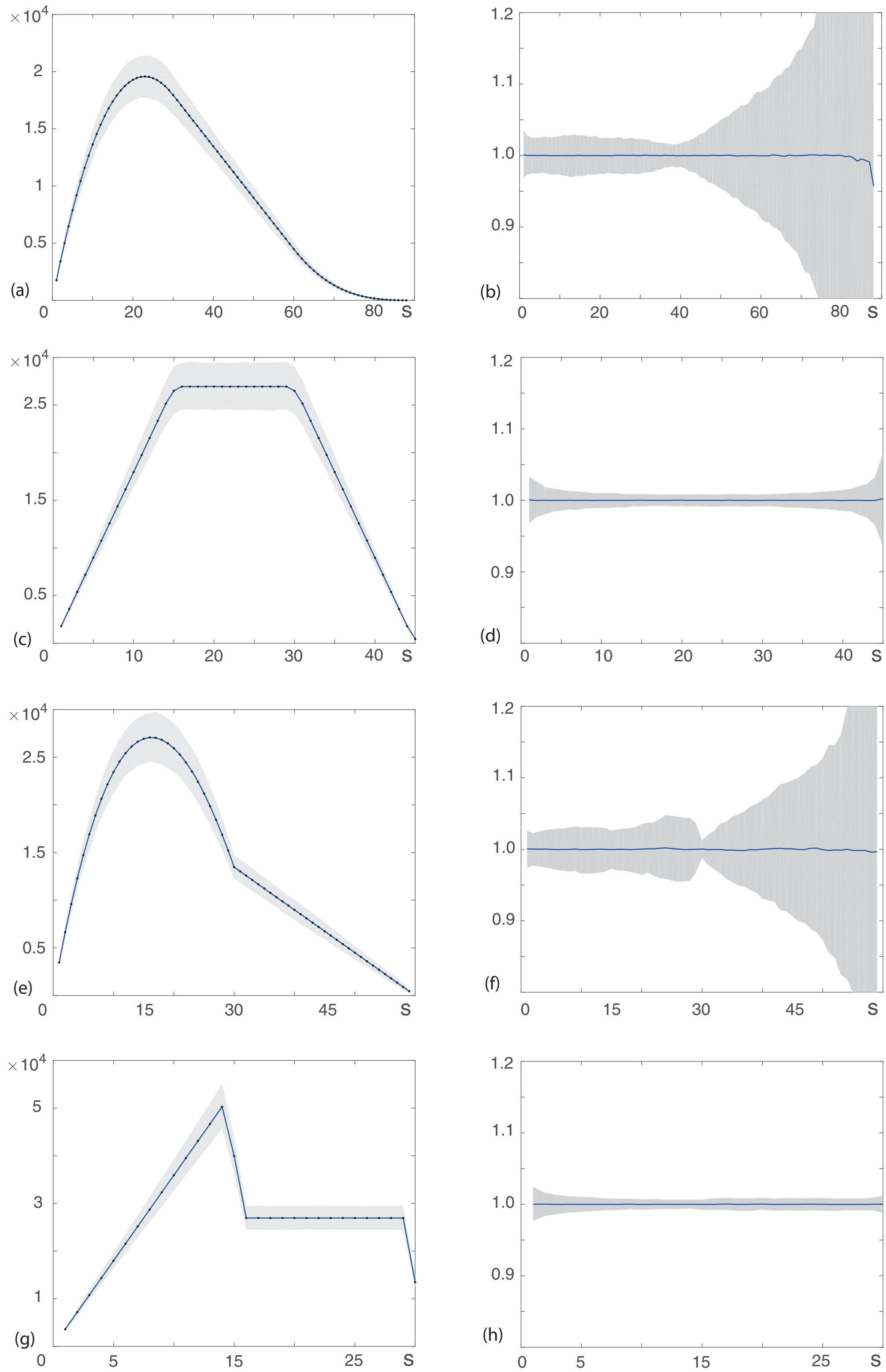

FIG. 2. Frequency and pair correlation function versus pair distance, s, for a two-dimensional 60×30 spatial domain. The gray shading indicates the region between the 2.5th and 97.5th percentiles. Realizations $R = 1000$, density $\rho = 0.5$ ($\approx$900 points). (a), (c), (e), (g) Left column: Pair distance frequency. Sample mean (solid curves) and distance distribution scaled by ρ_2 (dots). (b), (d), (f), (h) Right column: Pair correlation function. (a), (b) Top row: Nonperiodic Manhattan. (c), (d) Top-middle row: Periodic Manhattan. (e), (f) Bottom-middle row: Nonperiodic Chebyshev. (g), (h) Bottom row: Periodic Chebyshev.

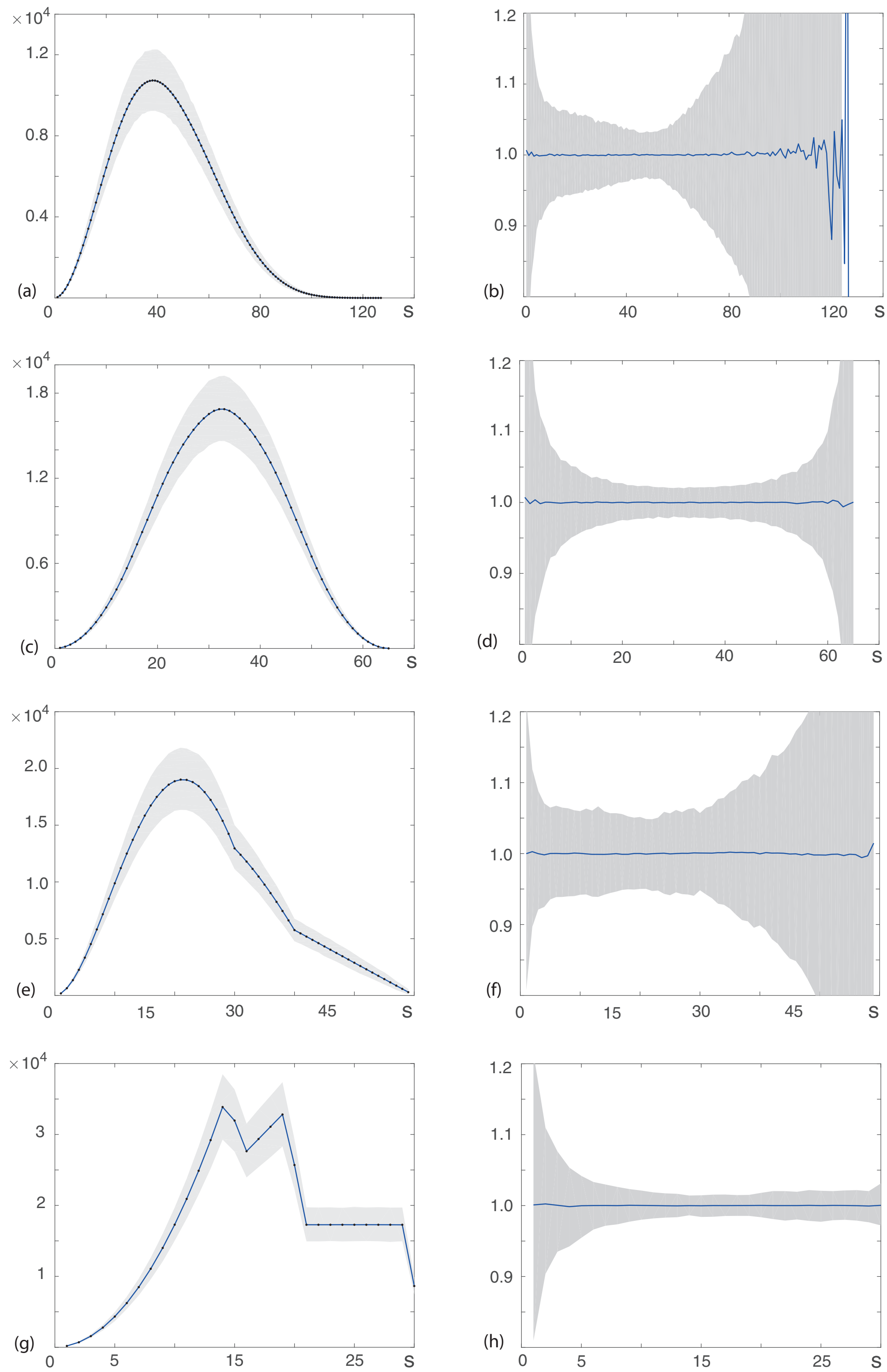

FIG. 3. Frequency and pair correlation function versus pair distance, s, for a three-dimensional $60 \times 30 \times 40$ spatial domain. The gray shading indicates the region between the 2.5th and 97.5th percentiles. Realizations $R = 1000$, density $\rho = 0.01$ ($\approx$720 points). (a), (c), (e), (g) Left column: Pair distance frequency. Sample mean (solid curves) and distance distribution scaled by ρ_2 (dots). (b), (d), (f), (h) Right column: Pair correlation function. (a), (b) Top row: Nonperiodic Manhattan. (c), (d) Top-middle row: Periodic Manhattan. (e), (f) Bottom-middle row: Nonperiodic Chebyshev. (g), (h) Bottom row: Periodic Chebyshev.

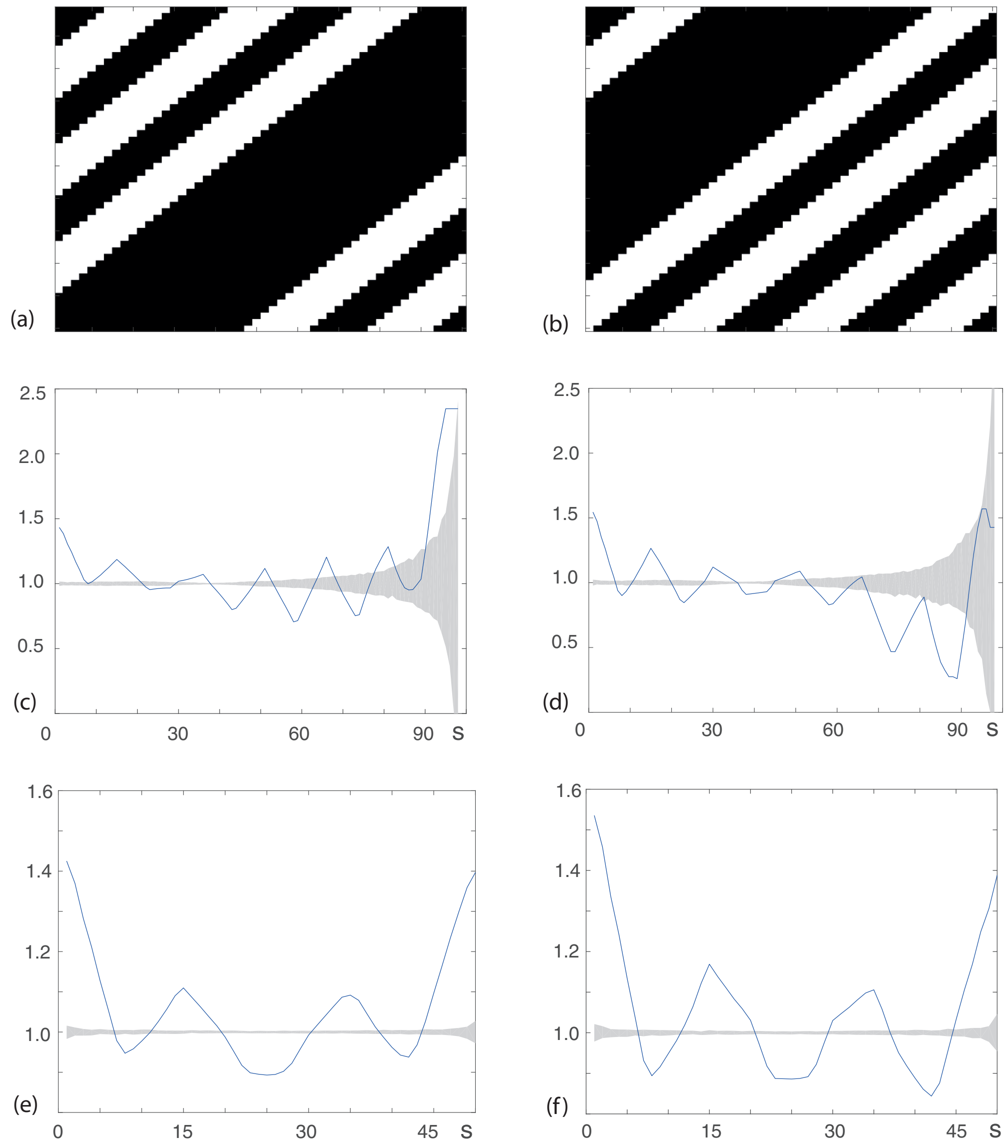

FIG. 4. Analysis of two-dimensional 50×50 diagonal bar patterns with Manhattan pair correlation function (solid curves). The gray shading indicates the region between the 2.5th and 97.5th percentiles of the domains populated uniformly at random with the same density as in the diagonal patterns, realizations $R = 1000$. (a), (c), (e) Left column: Thickened central bar pattern analysis. (b), (d), (f) Right column: Thickened off-center bar pattern analysis. (a), (b) Top row: Diagonal bar patterns. (c), (d) Middle row: Nonperiodic Manhattan pair correlation function. (e), (f) Bottom row: Periodic Manhattan pair correlation function.

and cellular automata simulations. Discrete pair correlation functions on square lattices have recently received attention through the works of Binder and Simpson [27], Gavagnin *et al.* [30], and Johnston and Crampin [31]. The purpose of this study is to build upon these works to provide a unified probabilistic framework to derive discrete pair correlation functions on a square lattice. We also examine the variability of pair correlation functions while using both nonperiodic and periodic metrics simultaneously to assess spatial patterning.

Our main contribution is the extension of the Manhattan and Chebyshev discrete pair correlation functions to k dimensions, via derivation of the corresponding k-dimensional distance distributions which serve as the normalizing factors of the pair correlation functions. Binder and Simpson [27], Gavagnin *et al.* [30], and Johnston and Crampin [31] each used a geometric combinatorial approach to determine the necessary normalization factors in one, two, and three dimensions. Specifically, their approach consisted of counting the total number of possible pair distances between agents in a completely filled lattice, which was effective when dealing with just one, two or three dimensions. The probabilistic formulation we have used, which instead consists of deriving the distance distribution of two random points in a k-dimensional lattice, provides a general framework for all discrete pair correlation functions.

In the case of two or three dimensions, normalization factors were derived partially for short to medium length scales for each of the Manhattan and Chebyshev distance metrics under both periodic and nonperiodic boundary conditions by Gavagnin *et al.* [30], while the two-dimensional nonperiodic Manhattan was derived completely by Johnston and Crampin

[31]. Therefore, our work, while generalizing to higher dimensions, also provides the derivation for the normalization term (at all distances) for periodic Manhattan and Chebyshev, and nonperiodic Chebyshev metrics.

In the work of Gavagnin *et al.* [30], the periodic distance metrics were used as a means to derive the nonperiodic distance count formulas. However, in their work they did not use the periodic distance metric to formulate or analyze the corresponding periodic pair correlation function. The periodic pair correlation functions were not applicable in the recent work of Johnston and Crampin [31] who considered the problem of modeling obstructions within the spatial domain. Therefore, the analysis of periodic Manhattan and Chebyshev pair correlation functions has not been considered.

The development of the higher dimensional periodic pair correlation functions in this study has immediate applications in extending the previous work of Agnew *et al.* [28] and Dini *et al.* [45]. Both works utilized periodic pair correlation functions in their analyses, but were limited to analyzing vertical and horizontal component distances independently using the one-dimensional rectilinear pair correlation functions since the higher dimensional pair correlation functions were not available. In a similar vein, Zhang *et al.* [46] recently applied the two-dimensional periodic Manhattan formulas of Gavagnin *et al.* [30] to an $L_x \times L_y$ lattice on the restricted domain $\min\{\lfloor L_x/2\rfloor, \lfloor L_y/2\rfloor\}$ since formulas for the periodic Manhattan pair correlation function were not available on the full domain $\{\lfloor L_x/2\rfloor + \lfloor L_y/2\rfloor\}$ prior to this work. We expect that our extension of the periodic pair correlation functions to higher dimensions, in particular dimension two, can allow for more thorough analysis in these biological applications.

We anticipate that the discrete k-dimensional pair correlation functions derived in this work will find application in the statistical analysis and classification of data in high-dimensional databases. For example, nonbinary image processing results in high-dimensional data to represent space, time, date, colors, intensity, contrast and so on. In particular, the Manhattan metric is commonly used as a measure of distance in high-dimensional nearest neighbor search problems of geometric and multimedia databases [47] and in data mining problems such as fraud detection and information retrieval [47,48]. Studies have also utilized both Manhattan and Chebyshev metrics in the formulation of the objective function in clustering algorithms [48,49].

Furthermore, it is interesting to note that in the continuous setting, the k-dimensional Manhattan distance distribution has been derived to assess the distribution of data relative to a random distribution (e.g., the uniform distribution with a continuous interval of unit support) [50]. However, to the best of our knowledge, there is currently no such measure for the discrete counterpart as is considered in this work. Therefore, the discrete distribution, which we have studied in this work, could potentially be used to obtain an approximation to the continuous distribution by considering the limiting process of the lattice spacing vanishing. This notion suggests that the k-dimensional discrete distribution can be used as an approximation to the k-dimensional continuous distribution, and this is the subject of our ongoing research.

Finally, we mention that a limitation of our method is that it relies on the regularity of the lattice spacing between agents because the distances can be realized as relatively well-understood random variables. Even in other low-dimensional generalizations where the connectivity is not constant (e.g., general tessellations [30] and punctured domains [31]), calculating the pair distances is difficult because it requires some fairly specific path-finding algorithms. While the problem is both interesting and challenging, it is beyond the scope of the present work, and this too is left to future research.

[1] M. S. Alber, Y. Jiang, and M. A. Kiskowski, Physica D **191**, 343 (2004).

[2] B. J. Binder, K. A. Landman, D. F. Newgreen, J. E. Simkin, Y. Takahashi, and D. Zhang, Bull. Math. Biol. **74**, 474 (2012).

[3] J. E. F. Green, S. L. Waters, J. P. Whiteley, L. Edelstein-Keshet, K. M. Shakesheff, and H. M. Byrne, J. Theor. Biol. **267**, 106 (2010).

[4] T. Höfer, J. A. Sherratt, and P. K. Maini, Physica D **85**, 425 (1995).

[5] N. J. Savill and J. A. Sherratt, Dev. Biol. **258**, 141 (2003).

[6] R. J. Thomas, A. Bennett, B. Thomson, and K. M. Shakesheff, Eur. Cell. Mater. **11**, 16 (2006).

[7] B. Vasiev and C. J. Weijer, J. Theor. Biol. **223**, 347 (2003).

[8] R. N. Kelsh, K. C. Sosa, J. P. Owen, and C. A. Yates, Bioessays **39**, 1600234 (2016).

[9] U. Dieckmann, R. Law, and J. A. J. Metz, *The Geometry of Ecological Interactions: Simplifying Spatial Complexity* (Cambridge University Press, Cambridge, 2000).

[10] A. F. G. Bourke and N. R. Franks, *Social Evolution in Ants* (Princeton University Press, Princeton, NJ, 1995).

[11] R. Law, J. Illian, D. F. R. P. Burslem, G. Gratzer, C. V. S. Gunatilleke, and I. A. U. N. Gunatilleke, J. Ecol **97**, 616 (2009).

[12] A. J. Flügge, S. C. Olhede, and D. J. Murrell, Ecology **93(7)**, 1540 (2012).

[13] T. Rajala, S. C. Olhede, and D. J. Murrell, J. Ecol. **107**, 711 (2019).

[14] J. B. Illian and D. F. R. P. Burslem, AStA Adv. Stat. Anal. **101**, 495 (2017).

[15] J. H. Phelps and C. L. Tucker III, Chem. Eng. Sci. **61**, 6826 (2006).

[16] M. D. Finn, S. M. Cox, and H. M. Byrne, J. Eng. Math. **48**, 129 (2004).

[17] B. J. Binder and S. M. Cox, Fluid Dyn. Res. **40**, 34 (2008).

[18] B. J. Binder, Phys. Lett. A **374**, 3483 (2010).

[19] B. J. Binder and K. A. Landman, Phys. Rev. E **83**, 041914 (2011).

[20] S. C. Yang, Powder Technol. **164**, 65 (2006).

[21] P. J. Diggle, J. Besag, and J. T. Gleaves, Biometrics **32**, 659 (1976).

[22] P. J. Diggle, *Statistical Analysis of Spatial Point Patterns* (Academic Press, London, 1983).

[23] B. D. Ripley, *Spatial Statistics* (Wiley, Hoboken, 1981).

[24] J. N. Perry, A. M. Liebhold, M. S. Rosenberg, J. Dungan, M. Miriti, A. Jakomulska, and S. Citron-Pousty, Ecography **25**, 578 (2002).

[25] R. E. Baker and M. J. Simpson, Phys. Rev. E **82**, 041905 (2010).

[26] D. C. Markham, M. J. Simpson, and R. E. Baker, Phys. Rev. E **87**, 062702 (2013).

[27] B. J. Binder and M. J. Simpson, Phys. Rev. E **88**, 022705 (2013).

[28] D. J. G. Agnew, J. E. F. Green, T. M. Brown, M. J. Simpson, and B. J. Binder, J. Theor. Biol. **352**, 16 (2014).

[29] B. J. Binder and M. J. Simpson, R. Soc. Open Sci. **2**, 140494 (2015).

[30] E. Gavagnin, J. P. Owen, and C. A. Yates, Phys. Rev. E **97**, 062104 (2018).

[31] S. T. Johnston and E. J. Crampin, Phys. Rev. E **99**, 032124 (2019).

[32] N. A. Bahcall and R. M. Soneira, Astrophys. J. **270**, 20 (1983).

[33] M. Holzmann and Y. Castin, Eur. Phys. J. D **7**, 425 (1999).

[34] W. R. Young, A. J. Roberts, and G. Stuhne, Nature (London) **412**, 328 (2001).

[35] R. Law, D. J. Murrell, and U. Dieckmann, Ecology **84**, 252 (2003).

[36] A. Donev, S. Torquato, and F. H. Stillinger, Phys. Rev. E **71**, 011105 (2005).

[37] J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan, *Statistical Analysis and Modelling of Spatial Point Patterns* (Wiley, Hoboken, 2008).

[38] E. Hinde, F. Cardarelli, M. A. Digman, and E. Gratton, Proc. Natl. Acad. Sci. USA **107**, 16560 (2010).

[39] M. Raghib, N. A. Hill, and U. Dieckmann, J. Math. Biol. **62**, 605 (2011).

[40] R. N. Binny, A. James, and M. J. Plank, Bull. Math. Biol. **78**, 2277 (2016).

[41] S. T. Johnston, M. J. Simpson, D. L. S. McElwain, B. J. Binder, and J. V. Ross, Open Biol. **4**, 140097 (2014).

[42] S. T. Johnston, J. V. Ross, B. J. Binder, D. L. S. McElwain, P. Haridas, and M. J. Simpson, J. Theor. Biol. **400**, 19 (2016).

[43] B. J. Binder, J. F. Sundstrom, J. M. Gardner, V. Jiranek, and S. G. Oliver, PLoS Comput. Biol. **11**, e1004070 (2015).

[44] S. Dini, B. J. Binder, S. C. Fischer, C. Mattheyer, A. Schmitz, E. H. K. Stelzer, N. G. Bean, and J. E. F. Green, J. R. Soc. Int. **13**, 20160649 (2016).

[45] S. Dini, B. J. Binder, and J. E. F. Green, J. Theor. Biol. **439**, 50 (2018).

[46] S. Zhang, A. Chong, and B. D. Hughes, Phys. Rev. E **100**, 042415 (2019).

[47] A. Hinneburg, C. C. Aggarwal, and D. A. Keim, in *VLDB 2000, Proceedings of 26th International Conference on Very Large Data Bases, September 10–14, 2000, Cairo, Egypt*, edited by A. El Abbadi, M. L. Brodie, S. Chakravarthy, U. Dayal, N. Kamel, G. Schlageter, and K.-Y. Whang (Morgan Kaufmann Publishers Inc., San Francisco, California, 2000), pp. 506–515.

[48] C. C. Aggarwal, A. Hinneburg, and D. A. Keim, in *Database Theory—ICDT 2001*, edited by J. Van den Bussche and V. Vianu (Springer, Berlin, 2001), pp. 420–434.

[49] A. Mousa and Y. Yusof, in *Innovation and Analytics Conference and Exhibition (IACE 2015): Proceedings of the 2nd Innovation and Analytics Conference & Exhibition*, edited by A. Ibrahim, J. Zulkepli, N. Aziz, N. Ahmad, and S. Abdul-Rahman, AIP Conf. Proc. No. 1691 (AIP, New York, 2015), p. 040020.

[50] O. Sugano, J. Jpn. Statist. Soc. **12**, 113 (1982).