TTMP vol 1(3), 2024
doi: 10.30511/TTMP.2024.2035668.1028

Transactions in Theoretical
and Mathematical Physics

Pair correlation function and its role in characterizing a system of randomly dispersed particles

Hamid Ghorbani[a]

Faculty of Mathematical Sciences, University of Kashan, Kashan, Iran

Received: 14 July 2024 / Accepted: 16 August 2024 / Published: 16 August 2024

Abstract The pair correlation function (or p.c.f.) is a statistical mechanical tool that quantifies the density variation as a function of distance from a reference particle in a system of randomly dispersed particles and is usually used to describe quantitatively, for example, the internal structure of fluids. In this paper, after recalling the p.c.f function to study the interaction behaviour between particles, i.e., attraction, repulsion, or independence, different particle systems are simulated in 2D, and the estimated p.c.fs are interpreted.

Keywords: pair correlation function, inhibition, spatial point process, completely spatial randomness, statistical mechanics, system of particles, intensity.

1 Introduction

In the terminology of spatial point process theory, we say that the centre of a system of randomly distributed particles in 3D is distributed as a spatial point process. Characterizing point processes by their second summary statistic allows us to find out, on average, how the particles interact with each other. Three main classes of interaction behaviour are completely random, inhibition, and attraction. Some authors consider the p.c.f. to be the most important second-order summary of point processes, and it is worth bringing to the attention of scientists willing to use satirical tools to interpret their data. The p.c.f. is not just important, but it is a cornerstone in many subfields of physics. It has been used, for example, in the study of molecular order, fluid dynamics, correlated light scattering, and liquid structures [1].

In the following sections, spatial point processes are first introduced very briefly. Then, the p.c.f. is defined as a second summary statistic for spatial point processes. Its role in characterizing the system of randomly dispersed particles, a crucial aspect of our research, is addressed. Finally, a simulation study improves the paper's readability for researchers working with experimental data.

1.1 A brief review of the spatial point processes theory

Diggle's text comprehensively presents the theory behind point pattern analysis [2]. Data in the form of a set of points irregularly distributed in a region of space arise in many different contexts; examples include the locations of trees in a forest, the locations of cell phone antennas in an urban environment, the centers of pores in a ceramic coating, or the nuclei in a microscopic section of tissue. We call such a data set a spatial point pattern and refer to the locations as events to distinguish them from arbitrary points in the region of interest.

A spatial point process is a stochastic mechanism that generates a countable set of events x_i in the $\mathbb{R}^d$ (in applications d = 2 or 3). A realization of a spatial point process is called to as spatial point pattern and consists of a countable set of points $\{x_1, x_2, \ldots, x_{N(A)}\}$ in d-dimensional space. We write N(A) for the number of events in a region $A \subset \mathbb{R}^d, N(A) = \#(x_i \in A)$. We usually deal with processes that are homogeneous (or stationary) and isotropic. A homogeneous process is the one whose distribution is invariant under translation; that is, for an integer k and regions $A_1, A_2, \ldots, A_k$, the joint distribution of $N(A_1), \ldots, N(A_k)$ is equal to the joint distribution of $N(A_1 + \boldsymbol{y}), \ldots, N(A_k + \boldsymbol{y})$ for an arbitrary vector $\boldsymbol{y}$. An isotropic point process is the one whose distribution is invariant under rotation. Fluids (liquids and gases) are spatially homogeneous and isotropic. Thus, the average properties of a fluid are invariant with respect to spatially uniform translations through any vector and with respect to arbitrary rotations about any axis [3]. An example of a homogeneous and isotropic point process is the spatial Poisson process, as a model for an ideal gas, referred to as complete spatial randomness (CSR), which is characterized by the following properties

---

[a] e-mail: hamidghorbani@kashanu.ac.ir

- The number of events in each region A,$N(A)$, follows a Poisson distribution with mean $\lambda.|A|$, where, |.| denotes the area or volume in 2D or 3D, respectively.
- Given $N(A) = n$, the locations of the n events in A are uniformly distributed, i.e. there is no interaction between the points.

The parameter λ, is the intensity of the process, which is the average number of points per unit area or volume, depending on the underlying mathematical space.

In the terminology of statistical mechanics, a CSR system is called a "total disorder" or "unrelated" system.

It is worth mentioning here that we have three main classes of spatial point processes.

The homogeneous Poisson process is completely random (uncorrelated), where a random point at any location does not affect the presence or absence of points at other locations. The inhibited (negatively correlated), where the distances between points are typically larger than expected for a CSR pattern of the same intensity, and the clustered (positively correlated), where the distances between points are typically smaller than expected for a CSR pattern of the same intensity.

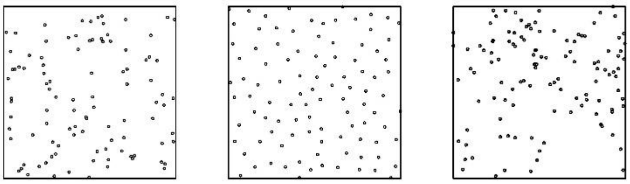

Figure 1 Simulated pattern showing three classes of point processes. From left to right: Poisson, inhibited and attractive patterns. All three patterns are in the unit square.

In other words, when examining the interaction energy of a point pattern at different scales.

The interaction can be attractive or repulsive, depending on geometrical features. The null interaction is associated with the Poisson point process as a reference or benchmark (see Fig. 1).

1.2 Basic Characteristics

Given a spatial point process, let dx and dy denote small regions containing the points x and y, respectively. The (first-order) intensity function is defined as

$$\lambda(x) = \lim_{|dx| \to 0} \frac{E[N(dx)]}{|dx|}. \tag{1}$$

The second-order intensity function is

$$\lambda_2(x, y) = \lim_{|dx| \to 0, |dy| \to 0} \frac{E[N(dx)N(dy)]}{|dx||dy|}. \tag{2}$$

Note that for a homogeneous Poisson process with intensity λ , $\lambda(x) = \lambda, \lambda_2(x, y) = \lambda^2$.

First-order characteristics, such as the intensity function, consider points as individuals, without interaction, while the second-order characteristics consider the interactions between points.

For a homogeneous and isotropic spatial point process, the intensity function is a constant equal to the expected number of events per unit area

$$\lambda(x) = \lambda = \frac{E[N(A)]}{|A|}, \tag{3}$$

and the second-order intensity is reduced to a function of distance:

$$\lambda_2(x, y) = \lambda_2(||x - y||) = \lambda_2(r). \tag{4}$$

The pair correlation function is the normalized second-order intensity function and is defined by:

$$g(x_1, x_2) = \frac{\lambda_2(x_1, x_2)}{\lambda(x_1)\lambda(x_2)}. \tag{5}$$

Again, for a homogeneous and isotropic spatial point process $g(x_1, x_2) = g(||x_1 - x_2||) = g(r)$.

The pair correlation function contains information about the short-range correlations. The pair correlation function is a measure that tells us how the density of particles varies as a function of distance from a reference particle. It gives us the probability of finding a particle at a distance from another particle relative to what we would expect if the particles were distributed according to the CSR process. For a CSR process, $g(r) = 1$. If $g(r) > 1$, it means that the particles with distance r are more frequent than in CSR process, which implies that the points tend to cluster relative to a Poisson process with the same intensity function. Similarly, if $g(r) < 1$, the particles with distance r are less frequent than in the CSR process, which implies that the points tend to epel relative to a Poisson process with the same intensity function. The former parenthèse from unity is a measure of the order that becomes less and less present as it increases

2 Importance of the pair correlation function

In a system of randomly dispersed particles, the p.c.f. quantifies how other particles surround the particle of interest. For example, as shown in Fig. 2, the p.c.f. has a constant value at all radial distances for a rarefied gas with no internal structure. In contrast, for a solid system in which molecules are almost uniformly distributed, the pair correlation function has sharply peaked values at the positions of particles. It is close to zero at the positions where particles are rarely found.

The p.c.f has almost the characteristics of the Dirac delta function for a solid system [5].

The p.c.f. describes the interaction between pairs of particles contained within a given volume. For a homogeneous and isotropically distributed system of particles, the probability of the interaction does not depend on the location or orientation of the particles. In summary, for a system of particles (atoms, molecules, colloids, etc.), the p.c.f. describes how the density varies as a function of distance from a reference particle. In simple terms, it measures the probability of finding a particle at a given distance of r away from a given reference pparticle relative to that for an ideal gas. Consider that the centers of a system of particles are modelled using a homogeneous isotropic point process with the intensity λ, which is the average number of points per unit volume. The p.c.f $g(r)$ is given by [6]:

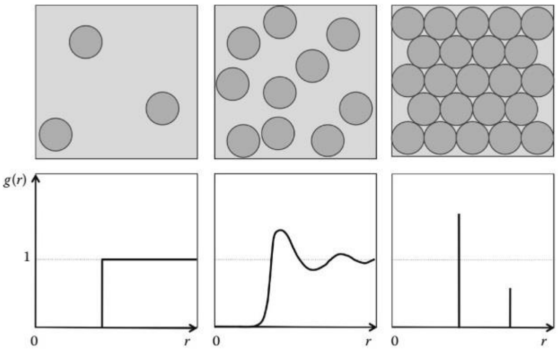

Figure 2 Schematic representation of the p.c.f for different materials. From left to right: liquid, gas and metal [4].

$$\begin{aligned}&\lambda^2 g(r) dV_1 dV_2\\&= P(\text{two points of the process are located in}\\&\quad \text{in the infinitesimal volume elements } \mathrm{dV}_1 \text{ and}\\&\quad \mathrm{dV}_2 \text{ centered in two points at the distance r}) \qquad (6)\end{aligned}$$

Figs 3, 4 and 5 show the p.c.f. of the point patterns of Fig. 1. The empirical functions (solid lines) deviate from the theoretically expected value in these plots, assuming the points are completely random (dashed lines).

To test whether the deviations in Figs 3,4 and 5 are statistically significant from what we expect under the CSR hypothesis, the standard approach is to use a Monte Carlo test based on envelopes of the p.c.f function obtained from simulated point patterns using the ‘spatstat’ package in the R statistical software [7]. Fig. 6 shows pointwise 95% confidence bands for the p.c.f. of the Poisson point

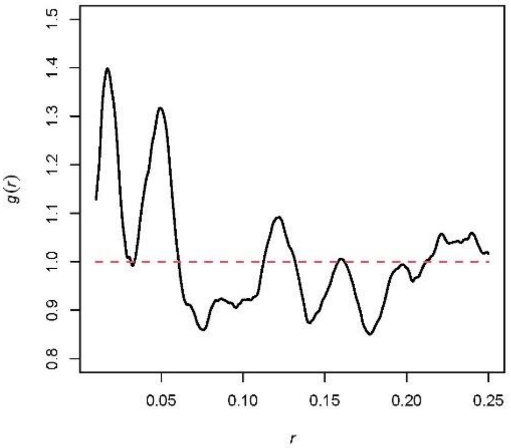

Figure 3: The empirical p.c.f. (solid line) of the Poisson point pattern of Fig. 1.

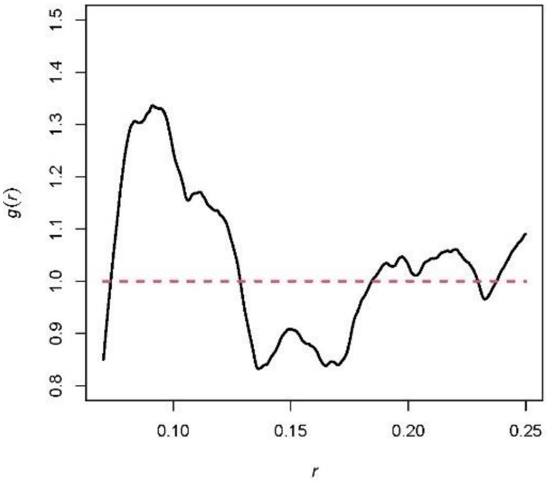

Figure 4: The empirical p.c.f. (solid line) of the repulsive point pattern of Fig. 1.

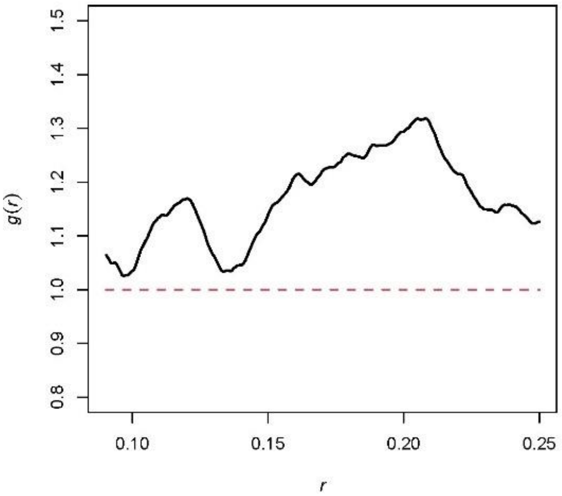

Figure 5: The empirical p.c.f. (solid line) of the attractive point pattern of Fig. 1.

the pattern in Fig.1. Solid lines represent the p.c.f. from Fig. 3. The shading indicates the envelope of values obtained from 199 simulations of complete spatial randomness (CSR). Dashed lines show the theoretical value of CSR. As we see from the figure, the empirical p.c.f. deviations from what we expect under CSR are not statistically significant at the 0.05 level. In Fig. 4, p.c.f. is below the Poisson value, $g(r) < 1$, for distances r up to 0.07 units, then it rises steeply to a value greater than the Poisson value, g(r) > 1 at about r = 0.09 units, before falling back to about 1 for long distances around 0.18. g (r) is zero for r ≤ 0.05, because the regular pattern in Fig. 1, has no interpoint distances shorter than 0.05 units.

Fig. 7 shows the pointwise 95% confidence bands for the p.c.f. of the regular point pattern in Fig. 1. It is clear that for short distances up to r=0.07 we have a negative correlation and for r>0.07, the deviation of the empirical p.c.fs from the Poisson theoretical value, is not statistically significant at the 0.05 level.

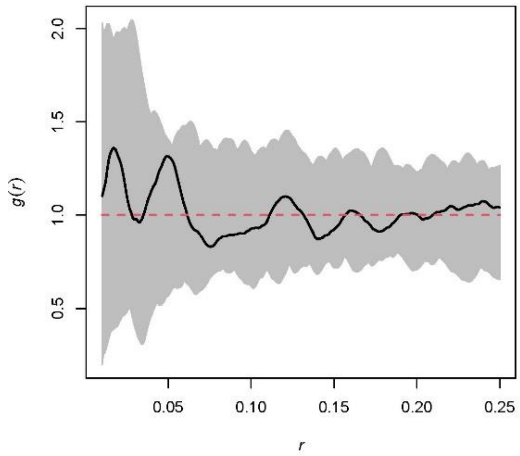

Figure 6: Pointwise 95% confidence bands for the p.c.f. in Fig. 3.

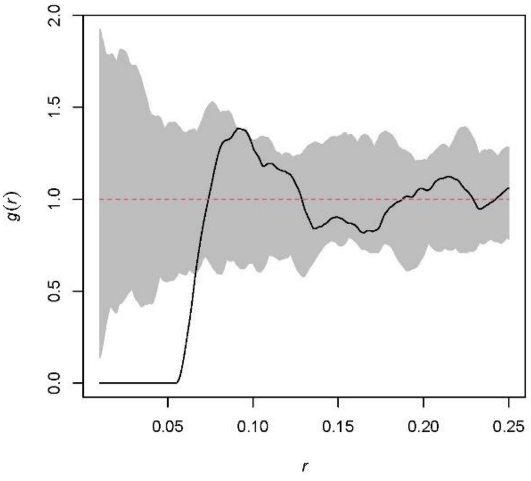

Figure 7: Pointwise 95% confidence bands for the p.c.f. in Fig.4.

Fig. 8 shows pointwise 95% confidence bands for the p.c.f. of the regular point pattern in Fig. 5. Contrary to our expectations, the envelopes here show that the deviations are not significant. By inference, we would say that by looking at the pattern in Fig. 1, the degree of clustering is not observable at any scale compared to what we observe under CSR.

3 Other characteristics defined based on p.c.f.

Two important characteristics, defined based on the p.c.f., are given in what follows.

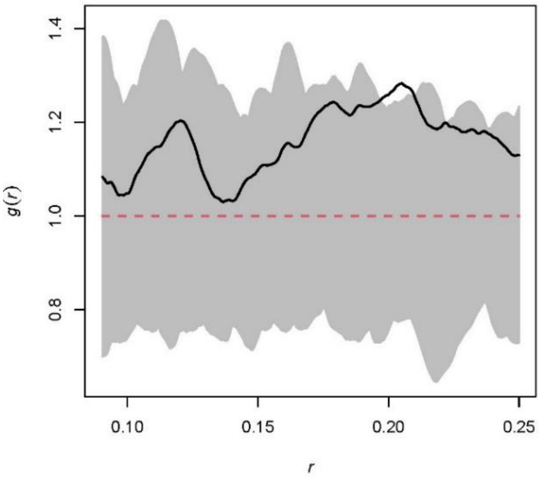

Figure 8: Pointwise 95% confidence bands for the p.c.f. in Fig. 5.

3.1 Radial distribution function

The position r_1 of the first maximum of $g(r)$ gives the average distance of the first nearest neighbors, and its width characterizes the dispersion of the nearest neighbor distances. By definition, the number $dN(r)$ of atoms contained between two concentric shells of radius r and $r + dr$ centered on an atom at the origin is

$$dN(r) = 4\pi\lambda g(r) r^2 dr \tag{7}$$

The radial distribution function (or r.d.f.) is defined as:

$$f(r) = 4\lambda r^2 g(r). \tag{8}$$

3.2 Coordination Number

The r.d.f. is a useful function for counting the number of neighbors by integration. The average number of neighbors around an atom, its coordination number $\bar{N}$, is obtained by integrating Eq.(7) from 0 to $R = r_{\mathrm{m}in}$, the first minimum of $g(r)$,

$$\bar{N} = 4\pi\lambda \int_0^{r_{min}} r^2 \,. \tag{9}$$

Finally, it is worth mentioning that the potential energy [8], [9], the Van Hove pair correlation function, and the Kirkwood-Buff integral [10] are among the other properties defined based on the p.c.f. The Fourier transform of the pair correlation function is called the static structure function S_k. The structure function contains information about the average relative positions of the atoms. Recently, [11] offered an approach based on the pair statistics of disordered particles, simplifying the inverse design of materials with desirable physical and chemical properties.

5. Conclusions

In this paper, we have reviewed the p.c.f., the well-known second summary tool in the theory of spatial point processes, to characterize the density variation as a function of distance from a reference particle in a system of randomly dispersed particles. Although it is defined for 2D and 3D systems of particles, we need a lot of programming skills to compute it in 3D. In the literature, some routines can be implemented for its handling in 3D. For the sake of simplicity, the 2D case has been considered in this paper.

References

1. K. Younge, C. Christenson, A. Bohara, J. Crnkovic, P. Saulnier, Am. J. Phys. **72**, 1247 (2004)
2. P. J. Diggle, *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns* (Chapman & Hall/CRC, Boca Raton, 2014)
3. P. M. Chaikin, T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, London, 1995)
4. N. Kanellopoulos, *Nanoporous Materials: Advanced Techniques for Characterization, Modeling, and Processing* (CRC Press, London, 2011)
5. Z. H Stachurski, *Fundamentals of Amorphous Solids Structure and Properties*. (Wiley-VCH Verlag, Weinheim, 2015)
6. S. N. Chiu, D. Stoyan, W. S. Kendall, J. Mecke, *Stochastic Geometry and Its Applications* (John Wiley and Sons Ltd, New York, 2013)
7. A. Baddeley, E. Rubak, R. Turner, *Spatial Point Patterns: Methodology and Applications with R* (CRC Press, London, 2015)
8. D. Frenkel, S. Berend, *Understanding Molecular Simulation from Algorithms to Applications* (Academic Press, San Diego, 2023)
9. J. D. Poll, S. Miller, J. Chem. Phys. **54** (2013)
10. K. Koga, B. Widom, J. Chem. Phys. **11** (2013)
11. H. Wang, S. Torquato, J. Chem. Phys. **160** (2024)