

## BOUNDS FOR THE PROBABILITY GENERATING FUNCTIONAL OF A GIBBS POINT PROCESS

KASPAR STUCKI<sup>\*\*\*</sup> AND

DOMINIC SCHUHMACHER,<sup>\*\*\*</sup> *University of Bern*

### Abstract

We derive explicit lower and upper bounds for the probability generating functional of a stationary locally stable Gibbs point process, which can be applied to summary statistics such as the  $F$  function. For pairwise interaction processes we obtain further estimates for the  $G$  and  $K$  functions, the intensity, and higher-order correlation functions. The proof of the main result is based on Stein's method for Poisson point process approximation.

*Keywords:* Gibbs process; probability generating functional; intensity; correlation function;  $F$  function;  $G$  function;  $K$  function; Poisson saddlepoint approximation; Stein's method

2010 Mathematics Subject Classification: Primary 60G55

Secondary 62M30

### 1. Introduction

Gibbs processes are very popular point process models which are extensively used both in spatial statistics and in statistical physics; see e.g. [10] and [13]. In particular, the pairwise interaction processes allow a simple yet flexible modelling of point interactions. However, a major drawback of Gibbs processes is that, in general, there are no analytic formulas available for their intensities or higher-order correlation functions. Recently, Baddeley and Nair [2], [3] proposed an approximation method that is fast to compute and accurate as verified by Monte Carlo methods. There are, however, no theoretical results in this respect and hence no guarantees for accuracy in most concrete models nor quantifications of the approximation error.

The aim of the present paper is to derive rigorous lower and upper bounds for correlation functions and related quantities. These allow us to narrow down the true values quite precisely if the Gibbs process is not too far away from a Poisson process. Figure 1 shows our bounds on the intensity for a two-dimensional Strauss process, which is dependent on its interaction parameter  $\gamma$ . The plus symbols are estimates of the true intensity obtained as averages over the numbers of points in  $[0, 1]^2$  of 10 000 Strauss processes simulated by dominated coupling from the past. The point processes were simulated on a larger window  $([-0.5, 1.5]^2)$  in order to avoid noticeable edge effects. All simulations and numerical computations in this paper were performed in R (see [www.r-project.org](http://www.r-project.org)) using the contributed package `spatstat` [4].

Our main result, Theorem 1, below, more generally gives bounds on the probability generating functional of a Gibbs process. Let  $\Xi$  be an arbitrary point process on  $\mathbb{R}^d$ . The probability

---

Received 16 October 2012; revision received 27 March 2013.

\* Current address: Institute for Mathematical Stochastics, University of Göttingen, Goldschmidtstrasse 7, 37077 Göttingen, Germany

\*\* Email address: Email address: [kaspar.stucki@mathematik.uni-goettingen.de](mailto:kaspar.stucki@mathematik.uni-goettingen.de)

\*\*\* Email address: Email address: [dominic.schuhmacher@mathematik.uni-goettingen.de](mailto:dominic.schuhmacher@mathematik.uni-goettingen.de)

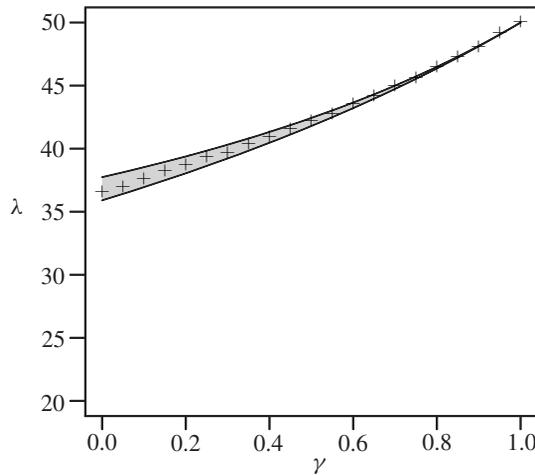


FIGURE 1: Bounds on the intensities of two-dimensional Strauss processes with  $\beta = 50$ ,  $r = 0.05$ , and values of  $\gamma$  ranging from 0 to 1. The plus symbols are estimates of the intensities based on 10 000 simulations each.

generating functional  $\Psi_{\Xi}$  is defined as

$$\Psi_{\Xi}(g) = \mathbb{E}\left(\prod_{y \in \Xi} g(y)\right), \tag{1}$$

for any measurable function  $g: \mathbb{R}^d \rightarrow [0, 1]$  for which  $1 - g$  has bounded support; see e.g. [7, p. 59] for details.

Many statistics of point processes, such as the *empty space function* ( $F$  function), contain expectations as in (1). For pairwise interaction processes the situation is even better. By the Georgii–Nguyen–Zessin equation (3), below, the *nearest neighbour function* ( $G$  function), Ripley’s  $K$  function, and the correlation functions of all orders can be rewritten using the probability generating functional.

The idea for proving Theorem 1, below, is to replace the Gibbs process  $\Xi$  in (1) by a suitable Poisson process and bound the error using Stein’s method.

The rest of the paper is organised as follows. In Section 2 we introduce some notation and state the main result. In Section 3 we provide bounds on the intensity, and in Section 4 bounds on other summary statistics are derived. Section 5 contains the proof of the main result.

## 2. Preliminaries and main result

Let  $(\mathfrak{N}, \mathcal{N})$  denote the space of locally finite point measures on  $\mathbb{R}^d$  equipped with the  $\sigma$ -algebra generated by the evaluation maps  $[\mathfrak{N} \ni \xi \mapsto \xi(A) \in \mathbb{Z}_+]$  for bounded Borel sets  $A \in \mathbb{R}^d$ . A point process is just a  $\mathfrak{N}$ -valued random element. We assume the point processes to be *simple*, i.e. do not allow multi-points. Thus, we can use set notation, i.e.  $x \in \xi$  means that the point  $x$  lies in the support of the measure  $\xi$ .

In spatial statistics, point processes are usually defined on a bounded window  $\mathcal{W} \subset \mathbb{R}^d$ . Let  $\mathfrak{N}|_{\mathcal{W}}$  denote the restriction of the  $\mathfrak{N}$  to  $\mathcal{W}$ . A point process  $\Xi$  on  $\mathcal{W}$  is called a *Gibbs process* if it has a hereditary density  $u$  with respect to the distribution of the Poisson process

with unit intensity. Heredity means that  $u(\xi) > 0$  implies  $u(\eta) > 0$  for all subconfigurations  $\eta \subset \xi$ . Invoking the Hammersley–Clifford–Ripley–Kelly theorem in [10, Theorem 6.1], such a process may be characterised by requiring a density of the form  $u(\xi) = \exp(-\sum_{n=1}^{\infty} \sum_{\{x_1, \dots, x_n\} \subset \xi} v(x_1, \dots, x_n))$ , where  $v: (\mathbb{R}^d)^n \rightarrow (-\infty, \infty]$  are symmetric functions, which yields the definition of a finite Gibbs process from statistical physics. By heredity we can define the *conditional intensity* as

$$\lambda(x \mid \xi) = \frac{u(\xi \cup \{x\})}{u(\xi)}, \tag{2}$$

where  $0/0 = 0$ . Roughly speaking, the conditional intensity is the infinitesimal probability that  $\Xi$  has a point at  $x$ , given that  $\Xi$  coincides with the configuration  $\xi$  everywhere else. Furthermore,  $\lambda(\cdot \mid \cdot)$  uniquely characterises the distribution of  $\Xi$ , since by (2) we can recursively recover an unnormalised density. It is well known that the conditional intensity is the  $dx \otimes \mathcal{L}(\Xi)$ -almost everywhere unique product measurable function that satisfies the Georgii–Nguyen–Zessin equation [10, p. 95],

$$\mathbb{E}\left(\int_{\mathcal{W}} h(x, \Xi \setminus \{x\}) \Xi(dx)\right) = \int_{\mathcal{W}} \mathbb{E}(h(x, \Xi)\lambda(x \mid \Xi)) dx, \tag{3}$$

for every measurable  $h: \mathcal{W} \times \mathfrak{N}|_{\mathcal{W}} \rightarrow \mathbb{R}_+$ .

So far,  $\lambda(\cdot \mid \cdot)$  is only a function on  $\mathcal{W} \times \mathfrak{N}|_{\mathcal{W}}$ , but in many cases there exists a natural extension to the whole space, which we shall also denote by  $\lambda(\cdot \mid \cdot)$ . One way to generalise Gibbs processes to the whole space  $\mathbb{R}^d$  is then by the so-called *integral characterisation*. A point process  $\Xi$  on  $\mathbb{R}^d$  is a Gibbs process corresponding to the conditional intensity  $\lambda(\cdot \mid \cdot)$  if it satisfies (3) with  $\mathcal{W}$  replaced by  $\mathbb{R}^d$  for all measurable  $h: \mathbb{R}^d \times \mathfrak{N} \rightarrow \mathbb{R}_+$ ; see [10, p. 95] or [12] for a more rigorous presentation. Unlike in the case of a bounded domain,  $\Xi$  may not be uniquely determined by (3). For the rest of this paper we will only deal with the conditional intensity, i.e. if we say that a result holds for a Gibbs process with conditional intensity  $\lambda(\cdot \mid \cdot)$ , we mean that it holds for *all* processes corresponding to this conditional intensity.

A Gibbs process  $\Xi$  is said to be a *pairwise interaction process* if its conditional intensity is of the form

$$\lambda(x \mid \xi) = \beta \prod_{y \in \xi} \varphi(x, y),$$

for a constant  $\beta > 0$  and a symmetric *interaction function*  $\varphi$ . We denote the distribution of  $\Xi$  by  $\text{PIP}(\beta, \varphi)$ . The process  $\Xi$  is called *inhibitory* if  $\varphi \leq 1$  and it is said to have a *finite interaction range* if  $1 - \varphi$  is compactly supported. The point process  $\Xi$  is *stationary* if its distribution is invariant under translations, and *isotropic* if it is also invariant under rotations. If  $\Xi$  is stationary we tacitly assume that  $\varphi(x, y)$  depends only on the difference  $x - y$ ; we then write  $\varphi(x, y) = \varphi(x - y)$ . For conditions on  $\varphi$  ensuring the existence of  $\Xi$  we refer the reader to [13].

If  $\Xi$  is a general point process, its *expectation measure* or *first order moment measure*  $\mathbb{E}\Xi$  on  $\mathbb{R}^d$  is simply given by  $(\mathbb{E}\Xi)(A) = \mathbb{E}(\Xi(A))$  for every Borel set  $A \subset \mathbb{R}^d$ . For  $k \geq 1$  the *kth order factorial moment measure* of  $\Xi$  is the expectation measure of the factorial product measure

$$\Xi^{[k]} = \sum_{\substack{X_1, \dots, X_k \in \Xi \\ \text{pairwise different}}} \delta_{(X_1, \dots, X_k)}$$

on  $(\mathbb{R}^d)^k$ . Any moment measure is said to *exist* if it is locally finite.

The *intensity (function)*  $\lambda(x)$  of a Gibbs process  $\Xi$  is the density of the first moment measure of  $\Xi$  with respect to the Lebesgue measure, provided that the first moment measure exists. For a bounded  $A \subset \mathbb{R}^d$ , (3) yields

$$\mathbb{E}\Xi(A) = \int_A \mathbb{E}(\lambda(x \mid \Xi)) \, dx;$$

hence, the existence of the intensity and the form  $\lambda(x) = \mathbb{E}(\lambda(x \mid \Xi))$ . For stationary processes the intensity is obviously constant and we just write  $\lambda$ . For a stationary pairwise interaction process, we get

$$\lambda = \mathbb{E}(\lambda(0 \mid \Xi)) = \beta \mathbb{E}\left(\prod_{y \in \Xi} \varphi(y)\right) = \beta \Psi_{\Xi}(\varphi). \tag{4}$$

In a similar manner, it is possible to obtain the densities of the higher-order factorial moment measures, the so-called *correlation functions*; see [9] and [11]. For a stationary process  $\Xi \sim \text{PIP}(\beta, \varphi)$ , the  $k$ th correlation function is given by

$$\begin{aligned} \lambda_k(x_1, \dots, x_k) &= \beta^k \left( \prod_{1 \leq i < j \leq k} \varphi(x_i - x_j) \right) \mathbb{E}\left( \prod_{y \in \Xi} \varphi(y - x_1) \cdots \varphi(y - x_k) \right) \\ &= \beta^k \left( \prod_{1 \leq i < j \leq k} \varphi(x_i - x_j) \right) \Psi_{\Xi}(\varphi(\cdot - x_1) \cdots \varphi(\cdot - x_k)). \end{aligned} \tag{5}$$

A frequently used function in spatial statistics is the *pair correlation function* which is defined as

$$\rho(x, y) = \frac{\lambda_2(x, y)}{\lambda(x)\lambda(y)}. \tag{6}$$

In the stationary isotropic case this simplifies to  $\rho(s) = \lambda_2(x, y)/\lambda^2$ , where  $s = \|x - y\|$ .

For our results we need a stability condition for the Gibbs processes. A Gibbs process  $\Xi$  is called *locally stable* if there exists a nonnegative function  $c^*$  such that  $\int_{\mathcal{W}} c^*(x) \, dx < \infty$  for all bounded domains  $\mathcal{W} \subset \mathbb{R}^d$  and the conditional intensity satisfies

$$\lambda(x \mid \xi) \leq c^*(x),$$

for all  $\xi \in \mathfrak{N}$ . For stationary Gibbs processes we require  $c^*$  to be constant. Local stability is a frequently used condition in spatial statistics; see [10, p. 84]. However, for  $\Xi \sim \text{PIP}(\beta, \varphi)$  it implies that either  $\Xi$  is inhibitory or it has a *hard core*, i.e. there exists an  $r_{\min}$  such that  $\varphi(x, y) = 0$  for all  $|x - y| \leq r_{\min}$ . This excludes, for example, the *Lennard–Jones process*, which is very popular in statistical physics.

The following result is the key theorem for obtaining the results in Sections 3 and 4. Its proof is the subject of Section 5.

**Theorem 1.** *Let  $\Xi$  be a stationary locally stable Gibbs process with intensity  $\lambda$  and local stability constant  $c^*$ , and let  $g: \mathbb{R}^d \rightarrow [0, 1]$  be a function for which  $1 - g$  has bounded support. Then*

$$1 - \lambda G \leq \mathbb{E}\left(\prod_{y \in \Xi} g(y)\right) \leq 1 - \frac{\lambda}{c^*} (1 - e^{-c^*G}),$$

where  $G = \int_{\mathbb{R}^d} 1 - g(x) \, dx$ .

**Remark 1.** With a slight adaptation of the proof of Theorem 1 it is possible to obtain bounds on the probability generating functionals of possibly nonstationary processes. Let  $\Xi$  be a locally stable Gibbs process with intensity function  $\lambda(x)$  and local stability function  $c^*(x)$ . Let  $g: \mathbb{R}^d \rightarrow [0, 1]$  be a function for which  $1 - g$  has bounded support. Then we obtain

$$\begin{aligned}
 & 1 - \int_{\mathbb{R}^d} (1 - g(x))\lambda(x) \, dx \\
 & \leq \mathbb{E} \left( \prod_{y \in \Xi} g(y) \right) \\
 & \leq 1 - \frac{\int_{\mathbb{R}^d} (1 - g(x))\lambda(x) \, dx}{\int_{\mathbb{R}^d} (1 - g(x))c^*(x) \, dx} \left( 1 - \exp \left( - \int_{\mathbb{R}^d} (1 - g(x))c^*(x) \, dx \right) \right). \tag{7}
 \end{aligned}$$

As the intensity appears under the integral sign, we cannot generalise the arguments of Section 3 based on (7).

For the rest of the paper we tacitly assume all point processes to be stationary.

### 3. Bounds on the intensity

For the intensity of a inhibitory pairwise interaction process we immediately obtain from Theorem 1 the following result.

**Theorem 2.** Let  $\Xi \sim \text{PIP}(\beta, \varphi)$  be inhibitory with finite interaction range. Then

$$\frac{\beta}{1 + \beta G} \leq \lambda \leq \frac{\beta}{2 - e^{-\beta G}}, \tag{8}$$

where  $G = \int_{\mathbb{R}^d} 1 - \varphi(x) \, dx$ .

*Proof.* Recall from (4) that  $\lambda = \beta \mathbb{E} \prod_{y \in \Xi} \varphi(y)$  and use  $c^* = \beta$ . Theorem 1 then yields

$$1 - \lambda G \leq \frac{\lambda}{\beta} \leq 1 - \frac{\lambda}{\beta} (1 - e^{-\beta G}),$$

which can be rearranged as (8).

**Remark 2.** The lower bound of (8) can also be found in [13, p. 96] with the restriction  $\beta G < e^{-1}$ , whereas our inequality holds for all values of  $\beta$  and  $G$ .

**Example 1.** Let  $\Xi$  be a Strauss process, i.e.

$$\varphi(x) = \begin{cases} \gamma & \text{if } \|x\| \leq r, \\ 1 & \text{if } \|x\| > r, \end{cases}$$

for some parameters  $r > 0$  and  $0 \leq \gamma \leq 1$ . Then  $G = (1 - \gamma)\alpha_d r^d$ , where  $\alpha_d$  denotes the volume of the unit ball. Figure 1 shows that, for a reasonable choice of the parameters  $(\beta, r, \gamma)$ , the bounds on  $\lambda$  are quite good. The maximal relative error between the bounds and the simulated values is about 3.5%.

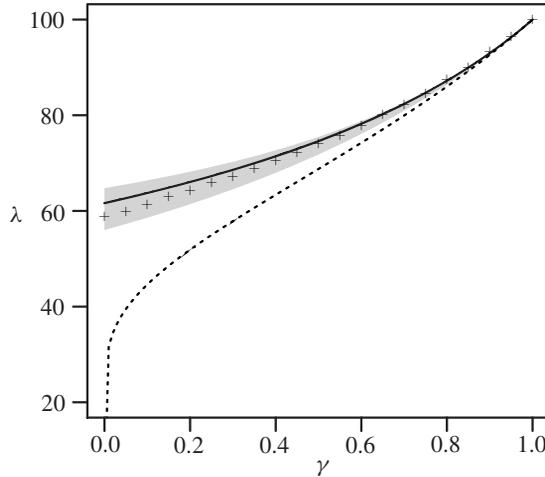


FIGURE 2: Intensities of two-dimensional Strauss processes with  $\beta = 100$ ,  $r = 0.05$ , and values of  $\gamma$  ranging from 0 to 1. The solid line is  $\lambda_{PS}$ , the dashed line is  $\lambda_{MF}$ , and the shaded area corresponds to the bounds in (8). The plus symbols are estimates of the intensities based on 10 000 simulations each.

**Remark 3.** For the special case of a Strauss hard core process on  $\mathbb{R}^2$ , i.e. setting  $\gamma = 0$  in Example 1, intensity estimates are derived in [16, p. 181]. The procedure used there can easily be refined to yield the bounds

$$\frac{\beta}{1 + \beta\pi r^2} \leq \lambda \leq \frac{\beta}{1 + \beta\pi r^2/4};$$

see [15]. The lower bound is the same as implied by Theorem 2. The upper bound is worse for small and moderate choices of  $\beta$ , but better if  $\beta$  is very large (e.g. greater than or equal to 500 if  $r = 0.05$ ).

For a comparison of our bounds on the intensity to known approximations from the literature we concentrate on two methods.

The first one is the *Poisson-saddlepoint* approximation proposed in [3]. The authors replaced the Gibbs process  $\Xi \sim \text{PIP}(\beta, \varphi)$  in (4) by a Poisson process  $H_{\lambda_{PS}}$  with intensity  $\lambda_{PS}$  such that the following equality holds:

$$\lambda_{PS} = \mathbb{E}\lambda(0 \mid H_{\lambda_{PS}}) = \beta \mathbb{E}\left(\prod_{y \in H_{\lambda_{PS}}} \varphi(y)\right).$$

Solving this equation yields

$$\lambda_{PS} = \frac{W(\beta G)}{G},$$

where  $W$  is *Lambert’s W function*, the inverse of  $x \mapsto xe^x$ , and  $G = \int_{\mathbb{R}^d} 1 - \varphi(x) \, dx$  as above.

The second method is the *mean-field* approximation that was also described in [3] and is given by

$$\lambda_{MF} = \frac{W(\beta \Gamma)}{\Gamma},$$

where  $\Gamma = - \int_{\mathbb{R}^d} \log(\varphi(x)) \, dx$ . Figure 2 shows the two approximations and our bounds from (8) for two-dimensional Strauss processes.

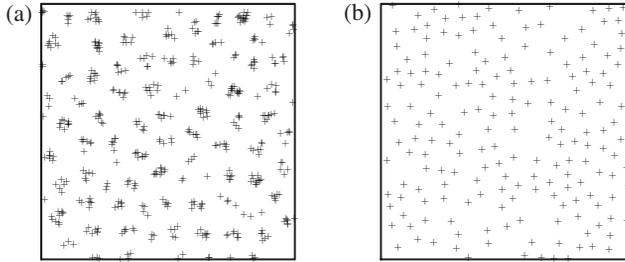


FIGURE 3: The two processes in Example 2 restricted to the unit square. Part (a) shows a realisation of the hard annulus process having 489 points, whereas  $\lambda_{PS} = 295.2$ . Part (b) shows a hard core process realisation having 188 points.

In [3] it was shown that, under the conditions of Theorem 2, we have  $\lambda \geq \lambda_{MF}$ . The authors also conjectured, based on simulations for Strauss processes, that  $\lambda_{PS}$  is an upper bound for  $\lambda$ . However, the next example indicates that this is not generally true.

**Example 2.** Consider the process  $\Xi \sim \text{PIP}(\beta, \varphi)$  with the interaction function

$$\varphi(x) = \begin{cases} 1 & \text{if } \|x\| \leq r, \\ 0 & \text{if } r < \|x\| \leq R, \\ 1 & \text{if } \|x\| > R, \end{cases}$$

for constants  $0 \leq r \leq R$ . We refer to this as a *hard annulus process*. It is a special case of a so-called *multiscale Strauss process*; see [10, Example 6.2]. Let  $d = 2$ ,  $\beta = 3000$ ,  $r = 0.05$ , and  $R = \sqrt{2}r$ . Then we have

$$\lambda_{MF} = 0, \quad \frac{\beta}{1 + \beta G} = 122.1, \quad \lambda_{PS} = 295.2, \quad \frac{\beta}{2 - e^{-\beta G}} = 1500.$$

An estimate of the intensity based on 300 simulations gives  $\hat{\lambda} = 493.8 > \lambda_{PS}$ . For comparison, we also estimated the intensity of a Strauss hard core process with the same  $\beta$  and  $G$  and obtained  $\hat{\lambda} = 193.3$ . Figure 3 shows that although the two processes have the same  $\beta$  and  $G$ , their realisations look quite different. All simulations were performed by long runs ( $10^7$  steps) of Markov Chain Monte Carlo.

We were not able to prove that  $\lambda > \lambda_{PS}$  in this case, but we do have the following heuristic argument for the observed phenomenon. The simulations show that, for large  $\beta$ , the points tend to cluster on ‘islands’ of radius less than or equal to  $r/2$  which are separated by a distance greater than or equal to  $R$ . Since the points within each island do not interact, we expect the intensity to grow linearly in  $\beta$  for large  $\beta$ . However,  $\lambda_{PS}$  only grows logarithmically for large  $\beta$ , so that at some point the intensity will overtake  $\lambda_{PS}$ .

Even if  $\lambda_{PS}$  may not serve as a bound on  $\lambda$  it remains useful as an approximation. Empirically its values stay relatively close to the simulated values, whereas the difference of our upper and lower bounds in (8) increases for large  $\beta G$ .

The following result in connection with Theorem 2 gives an upper bound on the error in Poisson saddlepoint approximation.

**Proposition 1.** *Under the conditions of Theorem 2 we have*

$$\frac{\beta}{1 + \beta G} \leq \lambda_{\text{PS}} \leq \frac{\beta}{2 - e^{-\beta G}}.$$

*Proof.* Since  $\lambda_{\text{PS}} = W(\beta G)/G$ , it suffices to show the following two inequalities:

$$\frac{x}{1 + x} \leq W(x) \quad \text{and} \quad W(x) \leq \frac{x}{2 - e^{-x}}, \tag{9}$$

for all  $x \geq 0$ . The first one follows from  $x/(1 + x) \leq \log(1 + x)$ , see [1, Equation 4.1.33], by transforming it to

$$\frac{x}{1 + x} \exp\left(\frac{x}{1 + x}\right) \leq x$$

and applying the increasing function  $W$  on both sides. For the second inequality, note that

$$\log(2 - e^{-x}) \leq \frac{x}{2 - e^{-x}}.$$

This holds because we have equality for  $x = 0$  and it is straightforward to see that the derivative of the left-hand side is less than or equal to the derivative of the right-hand side for all  $x \geq 0$ . A similar transformation as above and applying  $W$  on both sides again gives the second inequality in (9).

#### 4. Summary statistics

For a stationary point process  $\Xi$  the *empty space function* or *F function* is defined as the cumulative distribution function of the distance from the origin to the nearest point in  $\Xi$ , i.e.

$$\begin{aligned} F(t) &= \mathbb{P}(\text{there exists } y \in \Xi: \|y\| \leq t) \\ &= 1 - \mathbb{P}(\Xi(\mathbb{B}(0, t)) = 0) \\ &= 1 - \mathbb{E}\left(\prod_{y \in \Xi} \mathbf{1}_{\{y \notin \mathbb{B}(0, t)\}}\right) \\ &= 1 - \Psi_{\Xi}(\mathbf{1}_{\{\cdot \notin \mathbb{B}(0, t)\}}), \end{aligned}$$

where  $\mathbb{B}(x, t)$  denotes the closed ball centred at  $x \in \mathbb{R}^d$  with radius  $t \geq 0$  and  $\mathbf{1}$  denotes the indicator function. Thus, for a locally stable Gibbs process  $\Xi$  with constant  $c^*$ , from Theorem 1 we obtain

$$\frac{\lambda}{c^*} (1 - \exp(-c^* \alpha_d t^d)) \leq F(t) \leq \lambda \alpha_d t^d. \tag{10}$$

Note that, for a Poisson process with intensity  $\lambda$ , we may choose  $c^* = \lambda$ , in which case the lower bound in (10) is exact. A minor drawback of the bounds in (10) is that the intensity is in general not known and has to be estimated as well, e.g. by the methods of Section 3.

The *nearest neighbour function* or *G function* is defined as the cumulative distribution function of the distance from a typical point of  $\Xi$  (in the sense of the Palm distribution) to its nearest neighbour. For pairwise interaction processes  $\Xi \sim \text{PIP}(\beta, \varphi)$  the  $G$  function is computed in [9, Section 5] as

$$G(t) = 1 - \frac{\beta}{\lambda} \mathbb{E}\left(\prod_{y \in \Xi} \mathbf{1}_{\{y \notin \mathbb{B}(0, t)\}} \varphi(y)\right) = 1 - \frac{\beta}{\lambda} \Psi_{\Xi}(\mathbf{1}_{\{\cdot \notin \mathbb{B}(0, t)\}} \varphi(\cdot)).$$

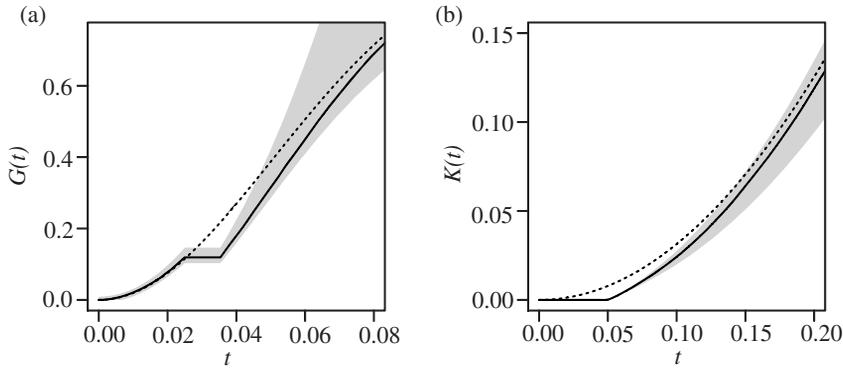


FIGURE 4: Part (a) shows the  $G$  function of the hard annulus process from Example 2 with parameters  $\beta = 70$ ,  $r = 0.025$ , and  $R = 0.035$ . Part (b) shows the  $K$  function of a Strauss hard core process with parameters  $\beta = 40$  and  $r = 0.05$ . In both panels the solid line is an estimate of the true function based on 1 000 simulations, the dashed line is the true function for a Poisson process (in the case of the  $G$  function with the same intensity as the one obtained by simulation for the hard annulus process), and the shaded area corresponds to the bounds computed in the text.

Thus, if  $\Xi$  is inhibitory and has finite interaction range, setting  $c^* = \beta$  in Theorem 1 yields

$$2 - \frac{\beta}{\lambda} - \exp(-\beta \tilde{G}_t) \leq G(t) \leq 1 - \frac{\beta}{\lambda} + \beta \tilde{G}_t,$$

where  $\tilde{G}_t = \int_{\mathbb{R}^d} [1 - \varphi(x) \mathbf{1}\{\|x\| > t\}] dx$ . Figure 4(a) shows these bounds for the hard annulus process from Example 2 with parameters  $\beta = 70$ ,  $r = 0.025$ , and  $R = 0.035$ .

Let us furthermore assume that  $\Xi$  is isotropic. Then the  $K$  function is defined as

$$K(t) = \alpha_d d \int_0^t s^{d-1} \rho(s) ds,$$

where  $\rho$  is the pair correlation function. By (5), (6), and Theorem 1, we obtain bounds on  $\rho$  as

$$\varphi(x) \left( \frac{\beta^2}{\lambda^2} - \frac{\beta^2 \tilde{G}_x}{\lambda} \right) \leq \rho(\|x\|) \leq \varphi(x) \left( \frac{\beta^2}{\lambda^2} - \frac{\beta}{\lambda} (1 - \exp(-\beta \tilde{G}_x)) \right), \tag{11}$$

where  $\tilde{G}_x = \int_{\mathbb{R}^d} [1 - \varphi(y)\varphi(y-x)] dy$ , and (in most cases numeric) integration of (11) yields bounds on the  $K$  function.

**Example 3.** Let  $\Xi$  be a Strauss process in two dimensions. Then

$$\tilde{G}_x = 2\pi r^2(1 - \gamma) - 2r^2(1 - \gamma)^2 \left( \arccos\left(\frac{\|x\|}{2r}\right) - \frac{\|x\|}{2r} \sqrt{1 - \left(\frac{\|x\|}{2r}\right)^2} \right);$$

see also [2]. Since we do not know the true intensity of the Strauss process, we substitute the bounds of (8) into (11) to obtain bounds on the  $K$  function. This procedure twice causes an error and therefore the estimates on  $K$  are good only for smaller values of  $\beta G$ . Figure 4(b) shows these estimates for  $\beta = 40$ ,  $r = 0.05$ , and  $\gamma = 0$ .

### 5. Proof of Theorem 1

The main strategy of the proof is to replace in (1) the process  $\Xi$  by a Poisson process  $H$ , and then use Stein’s method to bound the error

$$\mathbb{E}\left(\prod_{y \in \Xi} g(y)\right) - \mathbb{E}\left(\prod_{y \in H} g(y)\right).$$

In the context of Poisson process approximation, Stein’s method for bounding expressions of the form  $|\mathbb{E}f(\Xi) - \mathbb{E}f(H)|$  uniformly in  $f$  from a class of functions  $\mathcal{F}$  was first introduced in [6]. A nice exposition of the main body of the method with various ramifications and detailed proofs can be found in [17]. More recent developments include the generalisation to Gibbs process approximation [14].

The central idea of Stein’s method is to set up an equation of the form

$$f(\xi) - \mathbb{E}f(H) = \mathcal{A}h_f(\xi), \quad \text{for all } \xi \in \mathfrak{N}, \tag{12}$$

where  $\mathcal{A}$  is an operator on a space of functions  $h: \mathfrak{N} \rightarrow \mathbb{R}$  that characterises the distribution  $Q$  of  $H$  in the sense that  $\tilde{H}$  is  $Q$ -distributed if and only if  $\mathbb{E}\mathcal{A}h(\tilde{H}) = 0$  for all functions  $h$ . The hope is that a suitable choice of  $\mathcal{A}$  results in a right-hand side whose expectation is much easier to bound. A particularly successful approach is the generator method by Barbour [5], which suggested to choose for  $\mathcal{A}$  the infinitesimal generator of a Markov process whose stationary distribution is  $Q$ .

In the context of an approximation based on a Poisson process  $H_\nu$  with intensity  $\nu > 0$  on a compact set  $A$ , the default choice is a spatial immigration–death process on  $A$  with immigration rate  $\nu$  and unit per-capita death rate. This is a pure-jump Markov process on  $\mathfrak{N}|_A$  that holds any state  $\eta \in \mathfrak{N}|_A$  for an exponentially distributed time with mean  $1/(\nu|A| + \eta(A))$ , where  $|A|$  denotes the Lebesgue measure of  $A$ ; then a uniformly distributed point in  $A$  is added with probability  $\nu|A|/(\nu|A| + \eta(A))$ , or a uniformly distributed point in  $\eta$  is deleted with probability  $\eta(A)/(\nu|A| + \eta(A))$ . Let  $Z_\xi$  be such a process started at configuration  $\xi$ . The operator  $\mathcal{A}$  then takes the form

$$\mathcal{A}h(\xi) = \int_A [h(\xi + \delta_x) - h(\xi)]\nu \, dx + \int_A [h(\xi - \delta_x) - h(\xi)]\xi(dx).$$

Immigration–death processes have many nice properties. In particular, it is known that  $Z_\emptyset(t)$  is a Poisson process with intensity  $\nu(1 - e^{-t})$  for every  $t \geq 0$ . Let  $\{E_x\}_{x \in \xi}$  be independent and identically distributed and exponentially distributed random variables with mean 1, and introduce the death process  $D_\xi(t) = \sum_{x \in \xi} \mathbf{1}\{E_x > t\}\delta_x$ . Constructing  $Z_\emptyset$  and  $D_\xi$  independently on the same probability space,  $Z_\xi$  can be represented as  $Z_\xi(t) \stackrel{D}{=} Z_\emptyset(t) + D_\xi(t)$  for every  $t \geq 0$ ; see [17, Theorem 3.5].

We can then solve the Stein equation (12). Taking  $g: \mathbb{R}^d \rightarrow [0, 1]$  such that  $A = \text{supp}(1 - g)$  is compact, consider the function  $f: \mathfrak{N} \rightarrow [0, 1]$ ,

$$f(\xi) = f(\xi|_A) = \prod_{y \in \xi} g(y).$$

By [17, Theorem 5.2] the function  $h_f: \mathfrak{N} \rightarrow \mathbb{R}$ ,

$$h_f(\xi) = h_f(\xi|_A) = - \int_0^\infty [\mathbb{E}(f(Z_{\xi|_A}(t))) - \mathbb{E}(f(H_\nu))] \, dt,$$

is well-defined and satisfies (12). Thus,

$$\begin{aligned} \mathbb{E}f(\Xi) - \mathbb{E}f(H_\nu) &= \mathbb{E}\mathcal{A}h_f(\Xi) \\ &= \mathbb{E} \int_A [h_f(\Xi + \delta_x) - h_f(\Xi)]\nu \, dx + \mathbb{E} \int_A [h_f(\Xi - \delta_x) - h_f(\Xi)]\Xi \, dx \\ &= \mathbb{E} \int_A [h_f(\Xi + \delta_x) - h_f(\Xi)](\nu - \lambda(x \mid \Xi)) \, dx \\ &= \mathbb{E} \int_{\mathbb{R}^d} [h_f(\Xi + \delta_x) - h_f(\Xi)](\nu - \lambda(x \mid \Xi)) \, dx, \end{aligned} \tag{13}$$

where we applied the Georgii–Nguyen–Zessin equation on  $\mathbb{R}^d$  to the function  $[(x, \xi) \mapsto \mathbf{1}_A(x)(h_f(\xi) - h_f(\xi + \delta_x))]$  for obtaining the second equality.

Equation (13) is our starting point for further considerations.

**Proposition 2.** *Let  $\Xi$  be a stationary Gibbs process with intensity  $\lambda$  and conditional intensity  $\lambda(\cdot \mid \cdot)$ . Let  $g: \mathbb{R}^d \rightarrow [0, 1]$  be a function such that  $1 - g$  has bounded support. Then, for all  $\nu > 0$ ,*

$$\mathbb{E}\left(\prod_{y \in \Xi} g(y)\right) = 1 - \frac{\lambda}{\nu}(1 - e^{-\nu G}) + I_\nu(g),$$

where

$$I_\nu(g) = e^{-\nu G} \mathbb{E}\left(\int_0^1 e^{\nu Gs} \left(1 - \prod_{y \in \Xi} (1 - s(1 - g(y)))\right) ds \int_{\mathbb{R}^d} (1 - g(x))(\lambda(x \mid \Xi) - \nu) \, dx\right).$$

*Proof.* It is well known that, for the Poisson process  $H_\nu$ , we have

$$\mathbb{E}\left(\prod_{y \in H_\nu} g(y)\right) = \exp\left(-\nu \int_{\mathbb{R}^d} 1 - g(x) \, dx\right) = e^{-\nu G}; \tag{14}$$

see e.g. [7, Equation 9.4.17]. We then follow the main proof strategy laid out above in order to re-express

$$\mathbb{E}\left(\prod_{y \in \Xi} g(y)\right) - e^{-\nu G}.$$

Starting from (13), we may use the decomposition  $Z_{\xi+\delta_x} \stackrel{D}{=} Z_\xi + D_{\delta_x}$ , with independent  $Z_\xi$  and  $D_{\delta_x}$ , to see that, for any  $\xi \in \mathfrak{N}|_A$ ,

$$\begin{aligned} h_f(\xi + \delta_x) - h_f(\xi) &= \int_0^\infty \mathbb{E}f(Z_\xi(t)) - \mathbb{E}f(Z_{\xi+\delta_x}(t)) \, dt \\ &= \int_0^\infty \mathbb{E}f(Z_\xi(t)) - \mathbb{E}f(Z_\xi(t) + D_{\delta_x}(t)) \, dt \\ &= (1 - g(x)) \int_0^\infty \mathbb{E}f(Z_\xi(t))\mathbb{P}(D_{\delta_x}(t) \neq \emptyset) \, dt \\ &= (1 - g(x)) \int_0^\infty \mathbb{E}f(Z_\xi(t))e^{-t} \, dt \\ &= (1 - g(x)) \int_0^\infty \mathbb{E}(f(Z_\emptyset(t)) + (f(Z_\xi(t)) - f(Z_\emptyset(t))))e^{-t} \, dt. \end{aligned}$$

By further decomposing  $Z_\xi \stackrel{D}{=} Z_\emptyset + D_\xi$  with independent  $Z_\emptyset$  and  $D_\xi$ , we obtain

$$\begin{aligned} \mathbb{E}(f(Z_\xi(t)) - f(Z_\emptyset(t))) &= \mathbb{E}\left(\prod_{y \in Z_\emptyset(t)} g(y)\right) \mathbb{E}\left(\prod_{y \in D_\xi(t)} g(y) - 1\right) \\ &= \exp(-\nu(1 - e^{-t})G) \left(\prod_{y \in \xi} (1 - e^{-t}(1 - g(y))) - 1\right), \end{aligned}$$

where for the first expectation we used (14) and that  $Z_\emptyset(t)$  is a Poisson process with intensity  $\nu(1 - e^{-t})$ ; for the second expectation note that each point of  $\xi$  survives independently up to time  $t$  with probability  $e^{-t}$ . Thus, in total,

$$\begin{aligned} h_f(\xi + \delta_x) - h_f(\xi) &= (1 - g(x))e^{-\nu G} \int_0^\infty \left[ e^{\nu G e^{-t}} + e^{\nu G e^{-t}} \left(\prod_{y \in \xi} (1 - e^{-t}(1 - g(y))) - 1\right) \right] e^{-t} dt \\ &= (1 - g(x)) \frac{1 - e^{-\nu G}}{\nu G} + (1 - g(x))e^{-\nu G} \int_0^1 e^{\nu G s} \left(\prod_{y \in \xi} (1 - s(1 - g(y))) - 1\right) ds \end{aligned}$$

by the substitution  $s = e^{-t}$ .

Substituting this into (13) and using  $\mathbb{E}\lambda(x \mid \Xi) = \lambda$  finally yields

$$\mathbb{E}\left(\prod_{y \in \Xi} g(y)\right) - e^{-\nu G} = \frac{\nu - \lambda}{\nu} (1 - e^{-\nu G}) + I_\nu(g).$$

**Proposition 3.** *Let  $\Xi$  be a stationary locally stable Gibbs process with constant  $c^*$ . Then, for all  $0 < \nu < c^*$ ,*

$$\underline{I}_\nu(g) \leq I_\nu(g) \leq \overline{I}_\nu(g),$$

where

$$\underline{I}_\nu(g) = -\frac{1}{c^* - \nu} (c^*(1 - e^{-\nu G}) - \nu(1 - e^{-c^*G})) \leq 0, \tag{15}$$

$$\overline{I}_\nu(g) = \frac{1}{\nu} (c^*(1 - e^{-\nu G}) - \nu(1 - e^{-c^*G})) \geq 0. \tag{16}$$

Furthermore,  $I_\nu(g) \leq 0$  for all  $\nu \geq c^*$ .

*Proof.* Since  $-\nu \leq \lambda(x \mid \Xi) - \nu \leq c^* - \nu$  almost surely, we have, for  $\nu < c^*$ , the upper bound

$$I_\nu(g) \leq (c^* - \nu)e^{-\nu G} G \int_0^1 e^{\nu G s} \left(1 - \mathbb{E} \prod_{y \in \Xi} (1 - s(1 - g(y)))\right) ds, \tag{17}$$

and a similar lower bound, where  $c^* - \nu$  is replaced by  $-\nu$ . Because  $A = \text{supp}(1 - g)$  is bounded,  $\Xi$  can be replaced by  $\Xi|_A$  in (17). It is a known fact that every locally stable Gibbs process on a bounded domain can be obtained as a dependent random thinning of a Poisson process; see [8, Remark 3.4]. In particular, there exists a Poisson process  $H_{c^*}$  such that  $\Xi|_A \subset H_{c^*}$  almost surely. Since  $1 - s(1 - g(y)) \leq 1$  for all  $s \in [0, 1]$  and for all  $y \in \mathbb{R}^d$ , we obtain

$$1 - \mathbb{E} \prod_{y \in \Xi} (1 - s(1 - g(y))) \leq 1 - \mathbb{E} \prod_{y \in H_{c^*}} (1 - s(1 - g(y))) = 1 - e^{-sc^*G},$$

where the last equality follows by (14). Integrating and rearranging the terms yields (15) and (16). If  $\nu \geq c^*$ ,  $I_\nu(g)$  is obviously nonpositive.

We now continue with the proof of Theorem 1. Propositions 2 and 3 yield the upper bounds

$$\begin{aligned} \mathbb{E}\left(\prod_{y \in \Xi} g(y)\right) &\leq 1 - \frac{\lambda}{\nu}(1 - e^{-\nu G}) + \frac{c^*}{\nu}(1 - e^{-\nu G}) - (1 - e^{-c^*G}) \\ &= (c^* - \lambda)G \frac{1 - e^{-\nu G}}{\nu G} + e^{-c^*G}, \end{aligned}$$

for  $0 < \nu < c^*$ , and

$$\mathbb{E}\left(\prod_{y \in \Xi} g(y)\right) \leq 1 - \lambda G \frac{1 - e^{-\nu G}}{\nu G},$$

for  $\nu \geq c^*$ . Since the function  $[x \mapsto (1 - \exp(-x))/x]$  is monotonically decreasing for  $x \geq 0$ , we obtain the minimal upper bound for  $\nu = c^*$ , as

$$\mathbb{E}\left(\prod_{y \in \Xi} g(y)\right) \leq 1 - \frac{\lambda}{c^*}(1 - e^{-c^*G}).$$

For the lower bound recall the Weierstrass product inequality, which states

$$\prod_{i=1}^n (1 - a_i) \geq 1 - \sum_{i=1}^n a_i,$$

for  $0 \leq a_1, \dots, a_n \leq 1$ . Then, noting that the products below contain only finitely many factors not equal to 1 by the boundedness of  $\text{supp}(1 - g)$ , we have

$$\begin{aligned} \mathbb{E}\left(\prod_{y \in \Xi} g(y)\right) &= \mathbb{E}\left(\prod_{y \in \Xi} (1 - (1 - g(y)))\right) \\ &\geq 1 - \mathbb{E} \sum_{y \in \Xi} (1 - g(y)) \\ &= 1 - \mathbb{E} \int_{\mathbb{R}^d} 1 - g(x) \Xi(dx) \\ &= 1 - \lambda \int_{\mathbb{R}^d} 1 - g(x) dx \\ &= 1 - \lambda G \end{aligned}$$

by Campell’s formula; see [7, Section 9.5].

**Remark 4.** An alternative proof for the lower bound can be obtained by using Propositions 2 and 3 in the analogous way as for the upper bound.

### Acknowledgement

This work was supported by Swiss National Science Foundation Grant 200021-137527.

## References

- [1] ABRAMOWITZ, M. AND STEGUN, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Nat. Bureau Standards Appl. Math. Ser. 55). US Government Printing Office, Washington, DC.
- [2] BADDELEY, A. AND NAIR, G. (2012). Approximating the moments of a spatial point process. *Stat* **1**, 18–30.
- [3] BADDELEY, A. AND NAIR, G. (2012). Fast approximation of the intensity of Gibbs point processes. *Electron. J. Statist.* **6**, 1155–1169.
- [4] BADDELEY, A. AND TURNER, R. (2005). Spatstat: an R package for analyzing spatial point patterns. *J. Statist. Software* **12**, 42pp.
- [5] BARBOUR, A. D. (1988). Stein’s method and Poisson process convergence. In *A Celebration of Applied Probability* (J. Appl. Prob. Spec. Vol. **25A**), Applied Probability Trust, Sheffield, pp. 175–184.
- [6] BARBOUR, A. D. AND BROWN, T. C. (1992). Stein’s method and point process approximation. *Stoch. Process. Appl.* **43**, 9–31.
- [7] DALEY, D. J. AND VERE-JONES, D. (2008). *An Introduction to the Theory of Point Processes*, Vol. II, *General Theory and Structure*, 2nd edn. Springer, New York.
- [8] KENDALL, W. S. AND MØLLER, J. (2000). Perfect simulation using dominating processes on ordered spaces, with application to locally stable point processes. *Adv. Appl. Prob.* **32**, 844–865.
- [9] MASE, S. (1990). Mean characteristics of Gibbsian point processes. *Ann. Inst. Statist. Math.* **42**, 203–220.
- [10] MØLLER, J. AND WAAGEPETERSEN, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes* (Monogr. Statist. Appl. Prob. **100**). Chapman & Hall/CRC, Boca Raton, FL.
- [11] MØLLER, J. AND WAAGEPETERSEN, R. P. (2007). Modern statistics for spatial point processes. *Scand. J. Statist.* **34**, 643–684.
- [12] NGUYEN, X.-X. AND ZESSIN, H. (1979). Integral and differential characterizations of the Gibbs process. *Math. Nachr.* **88**, 105–115.
- [13] RUELLE, D. (1969). *Statistical Mechanics: Rigorous Results*. W. A. Benjamin, New York.
- [14] SCHUHMACHER, D. AND STUCKI, K. (2013). Gibbs point process approximation: total variation bound using Stein’s method. To appear in *Ann. Prob.* Preprint available at <http://uk.arxiv.org/abs/1207.3096>.
- [15] STOYAN, D. (2012). Personal communication.
- [16] STOYAN, D., KENDALL, W. S. AND MECKE, J. (1995). *Stochastic Geometry and Its Applications*, 2nd edn. John Wiley, Chichester.
- [17] XIA, A. (2005). Stein’s method and Poisson process approximation. In *An Introduction to Stein’s Method*, Singapore University Press, pp. 115–181.