

# Error bounds in stochastic-geometric normal approximation

Mathew Penrose, Tom Rosoman

► **To cite this version:**

Mathew Penrose, Tom Rosoman. Error bounds in stochastic-geometric normal approximation. Roesler, Uwe. Fifth Colloquium on Mathematics and Computer Science, 2008, Kiel, Germany. Discrete Mathematics and Theoretical Computer Science, DMTCS Proceedings vol. AI, Fifth Colloquium on Mathematics and Computer Science, pp.71-94, 2008, DMTCS Proceedings. <hal-01194682>

**HAL Id: hal-01194682**

**<https://hal.inria.fr/hal-01194682>**

Submitted on 7 Sep 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Error bounds in stochastic-geometric normal approximation

Mathew D. Penrose<sup>1†</sup> and Tom Rosoman<sup>1‡</sup>

<sup>1</sup> *Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK*

---

We provide normal approximation error bounds for sums of the form  $\sum_x \xi_x$ , indexed by the points  $x$  of a Poisson process (not necessarily homogeneous) in the unit  $d$ -cube, with each term  $\xi_x$  determined by the configuration of Poisson points near to  $x$  in some sense. We consider geometric graphs and coverage processes as examples of our general results.

**Keywords:** Central limit theorem, Berry Esséen bound, stochastic geometry, coverage process

---

## 1 Introduction

On first acquaintance with the central limit theorem, it is not uncommon for undergraduates to ask how accurate the normal approximation is for a sum of (say) twenty random variables. One answer is provided by the Berry-Esséen theorem. This says that there exists a constant  $C$  such that if  $T = \sum_i \xi_i$  is a finite sum of independent identically distributed random variables with finite third moments, then

$$d_K \left( \frac{T - ET}{\sqrt{\text{Var}(T)}}, Z \right) \leq C (\text{Var} T)^{-3/2} \sum_i E[|\xi_i - E\xi_i|^3], \quad (1.1)$$

where  $Z$  denotes a normal random variable with mean zero and variance 1, and for any two random variables  $X, Y$  we set

$$d_K(X, Y) := \sup_{t \in \mathbb{R}} |P[X \leq t] - P[Y \leq t]|,$$

the so-called Kolmogorov distance between the distribution of  $X$  and that of  $Y$ . A possible value for the constant  $C$  is given as  $22/5$  in (6) (page 525). Note that without a numerical value for  $C$ , the result, while providing information about the rate of convergence in the central limit theorem, does not provide an answer to the undergraduate question mentioned above.

In the present paper we are concerned with sums arising in stochastic geometry with spatial dependence between terms. The summands  $\xi_i$  are now indexed by the points  $x_i$  of a Poisson point process in a region

---

<sup>†</sup>Partially supported by the Alexander von Humboldt Foundation through a Friedrich Wilhelm Bessel Research Award

<sup>‡</sup>Supported by an EPSRC studentship

of  $d$ -space (so the number of terms in the sum is random), and are assumed to be determined by the local configuration of points in some sense. Such spatial random sums have been much studied recently, motivated by amongst other things, spatial statistics and theoretical computer science; see (8; 14; 21; 22). Depending on the precise nature of the local spatial dependence, numerous central limit theorems are available for such sums (see e.g. (1; 7; 8; 15; 19)).

Again in the geometric setting, Berry-Esséen type error bounds are of interest. Barbour and Xia (2), improving slightly on similar results of Penrose and Yukich (18), provide such a bound in a general geometric setting, but with a non-explicit constant  $C$ . For some geometrical examples, Chatterjee (4) provides error bounds which allow for (among other things) a non-Poisson number of points, though only for the weaker Kantorovich-Wasserstein distance, rather than the Kolmogorov distance (see the discussion at the start of Section 2 of (4)), and also with a non-explicit  $C$ .

In the present paper we provide a general Berry-Esséen error bound similar to that of (2), but with a more explicit constant (Theorem 2.3 below). As an application, we consider normal approximation for sums (over the Poisson points) of functions of each point's  $k$  nearest neighbour configuration, and compare this result with the similar one in (4) (again we have a more explicit constant). In these examples, as in (1.1) above, the error bound depends on the variance of the sum in question; finding a lower bound for this variance can be viewed as a separate problem to be considered case by case, although for the purpose of asymptotics, general methodologies for finding lower bounds for variances have been proposed in (17; 19).

The notion of local dependence is formalized by the notions of *radius of stabilization* or *radius of localization*, as used in (2; 18; 20) and elsewhere. Considerable simplifications are possible where this radius has bounded range, and for this case we give general results with simpler bounds. We also discuss some concrete examples where the range is bounded; namely (i) the count of the number of Poisson points with  $k$ th nearest neighbour less than some given distance away, and (ii) the covered volume in a union of balls centred at the Poisson points. We compare our results for these examples with those in (4) and (7).

As in (18), the general idea here is to express the variable of interest, or some close approximation to it, as a sum of variables forming a  $k$ -dependent random field, and use results of Chen and Shao (5) who provide normal approximation for such sums (using Stein's method). A refinement of this method is also used in (2), while Stein's method is used a different way in (4).

As this suggests, there is nothing very new about the methodology used here, although it has apparently not been well publicised (e.g. in (4), the ability of previously known methods to provide error bounds is somewhat understated). Alongside our principal aim of providing more explicit error bounds, we also aim in the present paper to draw further attention to these methods.

We restrict attention here to Poisson processes in the unit cube  $[0, 1]^d$ , but we believe this approach should also be applicable in the case of Poisson processes in a nice  $d$ -dimensional manifold embedded in a higher-dimensional Euclidean space, which provides much of the motivation for considering more general metric spaces in section 3.4 of (4). In the special case where the Poisson process on  $[0, 1]^d$  is homogeneous of intensity  $\lambda$ , it can be viewed, by scaling, as equivalent to a Poisson processes of unit intensity over  $[0, \lambda^{1/d}]^d$ . There is an extensive literature on central limit theorems (without concrete error bounds) in this setting (e.g. (7; 8; 9; 10; 12; 13; 19)).

## 2 General notation and general results

For  $x \in \mathbb{R}^d$  and  $r > 0$ , let  $B_r(x)$  denote the Euclidean ball of radius  $r$  around  $x$ , and let  $\theta_d$  denote the volume of the unit ball  $B_1(x)$ . For  $\mu > 0$ , write  $\text{Po}(\mu)$  for a Poisson variable with mean  $\mu$ .

Let  $(f(x), x \in [0, 1]^d)$  be a probability density function on the unit cube  $[0, 1]^d$ , and assume its supremum  $f_{\max}$  is finite; let  $f_{\min}$  denote its infimum. For  $\lambda > 0$  let  $\mathcal{P}_\lambda$  be a Poisson point process with intensity function  $\lambda f(\cdot)$  on the unit cube  $[0, 1]^d$ . List as  $x_1, x_2, \dots, x_N$  the points of  $\mathcal{P}_\lambda$ .

Suppose  $\xi(x, \Psi)$  is a measurable real-valued functional defined on all pairs of the form  $(x, \Psi)$  where  $\Psi$  is a finite point set in  $[0, 1]^d$  and  $x$  is a point in this point set. For  $x \in [0, 1]^d \setminus \Psi$  we write  $\xi(x, \Psi)$  for  $\xi(x, \Psi \cup \{x\})$ . Let

$$T := \sum_{i=1}^N \xi(x_i, \mathcal{P}_\lambda); \quad \sigma := \sqrt{\text{Var}[T]}. \quad (2.1)$$

We are interested in normal approximation for  $(T - ET)/\sigma$ .

Given  $a > 0$ , we can divide  $[0, 1]^d$  up into disjoint small cubes (boxes)  $C_i$  of side  $a\lambda^{-1/d}$  (if  $a^{-1}\lambda^{1/d}$  is not an integer, these will not form an exact fit). Define

$$Y_j = \sum_{i: x_i \in C_j} \xi(x_i, \mathcal{P}_\lambda); \quad X_j = Y_j - E[Y_j]. \quad (2.2)$$

Note that  $T = \sum Y_j$ . For  $p > 0$ , set

$$\beta_p := \beta_p(a) := \max_i E|X_i|^p; \quad \tilde{\beta}_p := \tilde{\beta}_p(a) := \max_i E|Y_i|^p.$$

For  $t > 0$  we say that the *range of interaction* of  $\xi$  is at most  $t$ , if  $\xi(x, \Psi) = \xi(x, \Psi \cap B_t(x))$  for all  $x$  and all  $\Psi$ . Our first result concerns the case when  $\xi$  has a range of interaction that is not too large compared to  $\lambda^{-1/d}$ . As we shall see, it follows rather easily from known results, but nevertheless has its uses.

**Theorem 2.1** *Suppose for some  $a > 0$  that the range of interaction of  $\xi$  is at most  $\lambda^{-1/d}a$ , and  $\beta_3 < \infty$ . Then with  $T$  defined as above,*

$$d_K\left(\frac{T - ET}{\sigma}, Z\right) \leq 75 * 21^{2d} \left(\lceil \frac{\lambda^{1/d}}{a} \rceil\right)^d \beta_3(a) \sigma^{-3}.$$

In Theorem 2.1, we have assumed for simplicity that a third moment bound holds, i.e.  $\beta_3 < \infty$ . One can also obtain weaker bounds under the weaker assumption that  $\beta_q$  is finite for some  $q$  with  $2 < q < 3$ , but we do not do this here.

It is often more natural to start from having a bound on the  $p$ th moment of the contribution from a single point inserted into  $\mathcal{P}$ , rather than on the contribution from all Poisson points in a cube of side  $a\lambda^{-1/d}$ . Set

$$\gamma_p := \sup_{x \in \mathbb{R}^d} E[|\xi(x, \mathcal{P}_\lambda)|^p]. \quad (2.3)$$

For  $x > 0$  we define the notation  $h_d(x) := (1 + x^{-1/d})^d$ . This is useful because for any  $x > 0$  we have

$$h_d(x) = \left(\frac{x^{1/d} + 1}{x^{1/d}}\right)^d \geq \frac{\lceil x^{1/d} \rceil^d}{x} \quad (2.4)$$

Note that  $h_d(1) = 2^d$  and  $h_d$  is decreasing with  $h_d(x) \rightarrow 1$  as  $x \rightarrow \infty$ .

By combining Theorem 2.1 and Lemma 5.2 below, we obtain a similar result to Theorem 2.1 but with the error bound given in terms of  $\gamma_p$ , some  $p > 3$ , rather than  $\beta_3$ . For simplicity we give this only for the case with  $p = 4$ .

**Theorem 2.2** *Suppose for some  $a \geq f_{\max}^{-1/d}$  that the range of interaction of  $\xi$  is at most  $a\lambda^{-1/d}$ , and that  $\gamma_4 < \infty$ . Then*

$$d_K \left( \frac{T - ET}{\sigma}, Z \right) \leq 21^{2d} (174600) h_d(\lambda a^{-d}) \gamma_4^{3/4} f_{\max}^3 a^{2d} \lambda \sigma^{-3}. \quad (2.5)$$

This result is similar to the case  $Q(\rho) = 0$  of Theorem 3.2 of (2), but provides an explicit constant as the coefficient of  $\lambda \sigma^{-3}$  in the right hand side.

We give a further general result where we assume only that the range of interaction of  $\xi$  is  $a\lambda^{-1/d}$  with high probability, rather than almost surely as in Theorems 2.1 and 2.2. For this we require the notions of *stabilization* and *localization*, also used in (2; 15; 18) and elsewhere (for a survey see (20)). Consider measurable  $\mathbb{R}^+$ -valued functionals of the form  $R(x, \Psi)$  where  $\Psi$  is a finite point set in  $[0, 1]^d$  and  $x$  is a point in this point set. For  $x \in [0, 1]^d \setminus \Psi$  we write  $R(x, \Psi)$  for  $R(x, \Psi \cup \{x\})$ . We say that  $R$  is a *radius of stabilization* for  $\xi$  if  $\xi(x, \Psi') = \xi(x, \Psi)$ , for all  $x \in [0, 1]^d$  and all finite point sets  $\Psi, \Psi'$  in  $[0, 1]^d$  such that  $\Psi' \cap B_{R(x, \Psi)}(x) = \Psi \cap B_{R(x, \Psi)}(x)$ .

We shall say that  $R$  is a *stopping radius of localization* for  $\xi$  if for all  $r > 0$ , the value of the function  $\mathbf{1}\{R(x, \Psi) \leq r\}$  is determined by the pair  $(x, \Psi \cap B_r(x))$ , and also for all  $(x, \Psi)$  we have

$$\xi(x, \Psi) = \xi(x, \Psi \cap B_{R(x, \Psi)}(x)).$$

The notion of ‘stopping radius of localization’ is used (without being assigned any particular terminology) in (2). It is similar to, though not the same as, the notion of ‘radius of localization’ described in (20). The notion of ‘radius of stabilization’ is the same as in e.g. Definition 2.1 of (18), up to a multiplicative factor of  $\lambda^{1/d}$  (where  $\lambda$  is the Poisson process intensity); for the non-asymptotic error bounds given here it is more natural not to include  $\lambda$  in the definition of radius of stabilization.

**Theorem 2.3** *Suppose  $R$  is either a radius of stabilization or a stopping radius of localization for  $\xi$ ,  $\lambda \geq 1$ , and  $a \geq f_{\max}^{-1/d}$  is a constant such that*

$$\sup_{x \in [0, 1]^d} P[R(x, \mathcal{P}_\lambda) > a\lambda^{-1/d}] \leq \lambda^{-9}. \quad (2.6)$$

Then

$$d_K \left( \frac{T - ET}{\sigma}, Z \right) \leq (3.012a^{2d} + 2) 174600 (21)^{2d} h_d(\lambda a^{-d}) \gamma_4^{3/4} f_{\max}^3 \lambda \sigma^{-3}. \quad (2.7)$$

Theorem 2.3 is similar to Theorem 3.2 of (2). By taking the  $\rho$  of (2) to be the same as our  $a$ , and setting the  $q$  and  $p$  of (2) to be 3 and 4 respectively, one sees that the middle term in the bound provided by (2) is similar to the bound in our Theorem 2.3, except that we provide an explicit constant, while (2) has a factor of  $f_{\max}^2$  where we have  $f_{\max}^3$ .

In many cases, a condition known as ‘exponential stabilization’ holds, i.e. for each  $\lambda$  there exists a radius of stabilization  $R(\lambda)$  such that

$$\limsup_{t \rightarrow \infty} t^{-1} \sup_{\lambda \geq 1} \log P[\lambda^{1/d} R(\lambda) > t] < 0.$$

When this condition holds, we can take the  $a$  in (2.6) to be a constant times  $\log \lambda$ , and hence the bound in (2.7) is at most a constant times  $(\log \lambda)^{2d} \lambda \sigma^{-3}$ , which is a slight improvement on Theorem 2.1 of (18), and would be the same as in Corollary 3.1 of (2), except that the latter result has a typo: the factor of  $\vartheta^{-p_3/2}$  in (2) should be  $\vartheta^{-p_3}$ .

To conclude the discussion of Theorem 2.3, we note that it is possible to conceive of a radius of stabilization which is *not* a stopping radius of localization. For example, suppose  $\xi(x, \Psi)$  is the distance from  $x$  to its nearest neighbour in  $\Psi \setminus \{x\}$  (or 0 if  $\Psi \setminus \{x\}$  is empty). Suppose for  $i = 1, 2$  we set  $R_i(x, \Psi)$  to be the distance from  $x$  to its  $i$ th nearest neighbour in  $\Psi \setminus \{x\}$  (or  $\sqrt{d}$  if  $\Psi \setminus \{x\}$  has fewer than  $i$  elements). Set  $R(x, \Psi)$  to be  $R_1(x, \Psi)$  if the first non-zero digit in the decimal expansion of  $R_2(x, \Psi) - R_1(x, \Psi)$  is even and set  $R(x, \Psi)$  to be  $R_2(x, \Psi)$  otherwise.

Then  $R$  is a radius of stabilization for  $\xi$ , since  $R \geq R_1$  and  $R_1$  is a radius of stabilization. However, in this example, given  $r > 0$ , in the event that  $R_1(x, \Psi) \leq r < R_2(x, \Psi)$ , we cannot tell from looking at  $\Psi \cap B_r(x)$  whether  $R = R_1$  or  $R = R_2$ , so that the value of the indicator  $\mathbf{1}\{R(x, \Psi) \leq r\}$  is *not* determined by the configuration of  $\Psi \cap B_r(x)$ . Therefore  $R$  is not a stopping radius of localization (it fails the ‘stopping time’ condition).

Of course, in this example a more natural choice of radius of stabilization, instead of  $R$ , would be  $R_1$  which is in fact a stopping radius of localization. Lemma 6.5 (in Section 6) makes rigorous the idea that given any radius of stabilization we can find a stopping radius of localization that is smaller.

Essentially because of examples such as the one above, there is in fact a minor error in the proof (though not the statements) of the main results of (18), which can be fixed by using Lemma 6.5. We shall discuss this issue at the start of Section 7.

For our last general result, instead of making an assumption on the range of influence of  $\xi(x, \Psi)$  we assume for some  $k \in \mathbb{N}$  that  $\xi(x, \Psi)$  is determined by  $x$  and the  $k$  nearest neighbours of  $x$  in the point set  $\Psi$ . Throughout, our logarithms are to base  $e$ .

**Theorem 2.4** *Suppose  $f_{\min} > 0$ . Let  $k \in \mathbb{N}$ , and suppose  $\lambda \geq 1$ . Suppose that  $T$  is given by (2.1), and that  $\xi(x, \Psi)$  is determined by  $x$  and the  $k$  nearest neighbours of  $x$  in the point set  $\Psi$ . Suppose also that*

$$2^{2d+1} \theta_d^{-1} \max \left( \frac{9 \log \lambda}{1 - \log 2}, k \right) \leq \lambda f_{\min}. \quad (2.8)$$

Then

$$d_K \left( \frac{T - ET}{\sigma}, Z \right) \leq \frac{(1764)^d (2452781) \gamma_4^{3/4} f_{\max}^3 \max(861(\log \lambda)^2, k^2) \lambda}{(f_{\min} \theta_d)^2 \sigma^3} \times h_d \left( \frac{\lambda f_{\min} \theta_d}{2^{d+1} \max(9 \log \lambda / (1 - \log 2), k)} \right). \quad (2.9)$$

The condition (2.8) is a mild lower bound saying that  $\lambda$  is reasonably large, both in absolute terms and relative to  $k$ . The final factor of  $h_d(\cdot)$  in the bound (2.9) is at most  $(3/2)^d$  by condition (2.8), and will be close to 1 if  $\lambda$  is large enough both in absolute terms and relative to  $k$ .

Theorem 2.4 is similar to Theorem 3.4 of (4), with a Poisson number of points in  $[0, 1]^d$ , rather than with a fixed number of points in a general metric space as in (4), and with an error bound on the Kolmogorov distance rather than the Kantorovich-Wasserstein distance as in (4). Setting aside the differences in scope, we may compare the error bound in Theorem 2.4 with that in (4). Our  $\lambda$  corresponds to the  $n$  of (4), and  $\tilde{\gamma}_p := \lambda^{p/2} \gamma_p$  corresponds to the quantity denoted  $\gamma_p$  in (4); our factor of  $\gamma_4^{3/4} \lambda$  could be written as  $\tilde{\gamma}_4^{3/4} \lambda^{-1/2}$ . Our bound is then competitive with that in (4). Indeed, if we consider only the second term in the bound in (4), we see that our bound has a more explicit constant factor, and has a factor of  $\max(861(\log \lambda)^2, k^2)$  where (4) has a factor of  $n^{3/p} k^3$  ((4) requires a  $p$ th moment condition for some  $p \geq 8$ ).

Much of the motivation for consideration of nearest neighbours in (4) comes from applications to tests of dimensionality such as those in (11). As mentioned earlier, these are often concerned with data in a submanifold of  $\mathbb{R}^d$ , for which our methods may well be adaptable but are not directly applicable. There are many other applications of nearest-neighbour statistics for which Theorem 2.4 should be directly applicable; see e.g. the discussion and references in (19), and see (3) for a recent discussion of one particular application.

It should be possible to prove a similar result to Theorem 2.4 (with a different constant) for summands which depend, not on the  $k$  nearest neighbours, but on the the Voronoi cells of the Poisson points; see the arguments in Section 6.3 of (15).

All of the results stated in this section carry over to the case where the points carry independent identically distributed marks, as in (18). Here we have not given the results with this degree of generality for the sake of clarity of exposition.

If the functional  $\xi(x, \Psi)$  takes only nonnegative values, then each of the bounds (2.5), (2.7), and (2.9) still hold with the right hand side divided by a factor of 4. This is because in that case, in the proof of (5.1) in Lemma 5.2 below, one can use the bound (5.5) instead of (5.4) and this improvement carries through to the subsequent proofs of Theorems 2.2, 2.3 and 2.4.

As they stand, the constants in our bounds might be too large to be of much practical use except for very large samples (i.e., large values of  $\lambda$ ). There may be scope to reduce the bound in the constant in Theorem 2.4. We discuss this at the conclusion of Section 7.

We prove Theorems 2.1 and 2.2 in Section 5. We prove Theorems 2.3 and 2.4 in Section 7 using estimates from Section 6. We discuss two concrete examples in Sections 3 and 4.

### 3 Degrees in geometric graphs

Theorem 2.1 reduces the problem of finding normal approximations to finding bounds on the variance and third moments. In this section we apply it to the count of vertices of given degree in the geometric graph. Given a finite point set  $\Psi \subset \mathbb{R}^d$ , and  $r > 0$ , the geometric graph  $G(\Psi, r)$  is the graph with vertex set  $\Psi$  and with each pair of points distance at most  $r$  from each other connected by an edge (see (14) for much more about random geometric graphs).

**Theorem 3.1** *Let  $k \in \mathbb{N}$ . Suppose  $T$  denotes the number of points  $x_i$  in the Poisson process  $\mathcal{P}_\lambda$  with degree at least  $k$  in  $G(\mathcal{P}, a\lambda^{-1/d})$ . Suppose  $f_{\min} > 0$ . Then, setting  $J := P[\text{Po}(2^{-d}\theta_d a^d f_{\min}) \geq k]$ , and  $\mu := a^d f_{\max}$ , we have:*

$$d_K \left( \frac{T - ET}{\sigma}, Z \right) \leq 75 * 21^{2d} \left( \lceil \frac{\lambda^{1/d}}{a} \rceil \right)^d (2\mu^3 + 3\mu^2 + \mu)(\lambda J)^{-3/2}, \quad (3.1)$$

and also, setting  $\sigma = (\text{Var}T)^{1/2}$ , we have

$$d_K \left( \frac{T - ET}{\sigma}, Z \right) \leq 75 * 21^{2d} h_d(a^{-d} \lambda) f_{\max}^3 (2a^{2d} + 3a^d + 1) \lambda \sigma^{-3}. \quad (3.2)$$

The bound (3.1) is more explicit but we have also included the bound (3.2) for comparison with (4); this example is discussed as application (i) of Theorem 3.4 of (4). Correcting an inconsistency between the definition of  $\sigma$  there and in Theorem 3.4 of (4) itself, the displayed bound in (4) should be

$$C \left( k^4 n^{1/2} \sigma^{-2} + k^3 n \sigma^{-3} \right)$$

which is rather like our bound in (3.2) except that (4) has powers of  $k$  in the bound, while we have powers of  $a$  in the bound. If  $a$  is fixed but  $k$  grows as  $\lambda$  becomes large, then our bound will grow more slowly.

*Proof of Theorem 3.1.* Our specific choice of  $T$  here is given by the general formula (2.1), when we set  $\xi(x, \Psi) = I[\#\{\Psi \cap B_{a\lambda^{-1/d}}(x)\} > k]$ . By Theorem 2.1,

$$d_K \left( \frac{T - ET}{\sigma}, Z \right) \leq 75 * 21^{2d} \left( \lceil \frac{\lambda^{1/d}}{a} \rceil \right)^d \beta_3(a) \sigma^{-3}. \quad (3.3)$$

We need to find an upper bound for  $\beta_3 := \beta_3(a)$ . Set  $\mu := a^d f_{\max}$  and  $V = \text{Po}(\mu)$ . Then

$$\beta_3 = E|X_i|^3 = E|Y_i - E[Y_i]|^3 \leq E[V^3] + (EV)^3 = 2\mu^3 + 3\mu^2 + \mu.$$

Using this in (3.3), together with (2.4), we obtain (3.2).

For (3.1), we need to find a lower bound for  $\sigma$ . For  $x, y$  in  $\mathbb{R}^d$ , let  $E_x$  (respectively  $E_x^{(y)}$ ) denote the event that  $x$  has degree at least  $k$  in  $G(x; \mathcal{P}_\lambda \cup \{x\})$  (respectively, in  $G(x; \mathcal{P}_\lambda \cup x, y)$ ). Using Palm theory for Poisson processes (see e.g. (14)) we get:

$$E[T] = \lambda \int P[E_x] f(x) dx$$

and

$$\begin{aligned} E[T(T-1)] &= \lambda^2 \int \int P[E_x^{(y)} \cap E_y^{(x)}] f(x) f(y) dx dy \\ &\geq \lambda^2 \int \int P[E_x] P[E_y] f(x) f(y) dx dy = (E[T])^2. \end{aligned}$$

Therefore

$$\text{Var}[T] = E[T(T-1)] + E[T] - (E[T])^2 \geq E[T].$$

Then  $P[E_x] \geq J$  as the worst case scenario is  $x$  being at a corner of  $[0, 1]^d$  and having  $f(y) = f_{\min}$  for all  $y$  near  $x$ . Therefore

$$E[T] = \lambda \int P[E_x] f(x) dx \geq \lambda \int J f(x) dx = \lambda J.$$

Plugging this into (3.3) gives (3.1).  $\square$



## 4 A coverage process

In this section, we let the random variable  $T$  be the total area covered in the unit cube by the union of balls of radius  $a/\lambda^{1/d}$  centred at points of  $\mathcal{P}_\lambda$ . We assume throughout the section that  $a\lambda^{-1/d} < 1/2$ .

**Theorem 4.1** *Suppose  $f_{\min} > 0$ . Then with  $T$  as just described, and  $\sigma = \sqrt{\text{Var}(T)}$ ,*

$$d_K\left(\frac{T - ET}{\sigma}, Z\right) \leq 75(144)^d (\theta_d a^d \lambda^{-1})^2 \sigma^{-3} \quad (4.1)$$

and

$$d_K\left(\frac{T - ET}{\sigma}, Z\right) \leq \frac{75(144)^d (a^d \theta_d)^{1/2} \exp(3f_{\max} a^d \theta_d)}{(1 - 2a\lambda^{-1/d})^{3d/2} (\exp(2^{-d} f_{\min} a^d \theta_d) - 1)^{3/2} \lambda^{1/2}}. \quad (4.2)$$

As an example, if we take  $d = 2$ ,  $f$  uniform and  $\theta_d a^d = 1$  then we get

$$\begin{aligned} d_K\left(\frac{T - ET}{\sigma}, Z\right) &\leq \frac{75 * 144^2 * e^3}{(\exp(1/4) - 1)^{3/2} \lambda^{1/2}} \left(1 - \frac{2}{\sqrt{\pi\lambda}}\right)^3 \\ &\approx \frac{2.06 * 10^8}{\lambda^{1/2}} \left(1 - \frac{2}{\sqrt{\pi\lambda}}\right)^3. \end{aligned}$$

A reduction in this bound is possible if working on the *torus* rather than the unit cube. See the remarks at the end of this section.

Again we have included a bound (4.1) which includes  $\sigma$  to aid comparison with (4). The quantity  $\theta_d a^d \lambda^{-1}$  corresponds to the quantity denoted  $M_\epsilon$  in Proposition 3.3 of (4). The second bound (4.2) agrees with the bound described in (4) for the case  $a = 1$ , but provides an explicit constant; a related bound (without an explicit constant) is also implicit in earlier work of Götze *et al.* (7).

To prove bounds of the form of those in Theorem 2.2, taking  $\xi(x, \Psi)$  to be the area of the region formed by the intersection of  $B_{a\lambda^{-1/d}}(x)$ ,  $[0, 1]^d$  and the Voronoi cell around  $x$ . However, we can obtain better error bounds by splitting the unit cube into a lattice of little cubes  $Z_i$ , each of side  $\epsilon$  with  $\epsilon^{-1} \in \mathbb{N}$ , and thinking of  $T$  as  $\sum Y_i$  where  $Y_i$  is the area covered in cube  $i$ . Let  $X_i = Y_i - E[Y_i]$ . Then the centred variable  $\bar{T} := T - E[T]$  is equal to  $\sum X_i$ . We also do better to use the following result of Chen and Shao (5) (not the same one used to prove our general results):

Let  $\mathcal{J}$  be a finite index set. Given a set of subsets  $\{C_i \subset \mathcal{J}; i \in \mathcal{J}\}$  define  $N(C_i)$  as  $\{j \in \mathcal{J} : C_i \cap C_j \neq \emptyset\}$ .

**Lemma 4.1** (5) *Suppose  $(X_i, i \in \mathcal{J})$  is a family of variables with mean zero and finite variances. Assume for each  $i \in \mathcal{J}$  there exists  $A_i \subset B_i \subset C_i \subset \mathcal{J}$  such that  $X_i$  is independent of  $X_{A_i^c}$ ,  $X_{A_i}$  is independent of  $X_{B_i^c}$  and  $X_{B_i}$  is independent of  $X_{C_i^c}$ . Also assume  $\max(|N(C_i)|, |\{j : i \in C_j\}|) \leq \kappa$ . Let  $2 < p \leq 3$ . Let  $W := \sum_i X_i$  and  $\sigma := \text{Var}(W)^{1/2}$ . Then:*

$$d_K\left(\frac{W}{\sigma}, Z\right) \leq 75\kappa^{p-1} \sum E|X_i|^p \sigma^{-p}$$

*Proof of Theorem 4.1.* We split the unit cube into very small cubes of side  $\epsilon$ . Given any  $\delta > 0$  we can find an  $\epsilon > 0$  such that given the cube of side  $\epsilon$  centred at  $i$  we can let  $A_i, B_i$  and  $C_i$  be made up of cubes

of side  $\epsilon$  inside the balls of radius  $2a\lambda^{-1/d} + \delta$ ,  $4a\lambda^{-1/d} + 2\delta$  and  $6a\lambda^{-1/d} + 3\delta$  respectively around  $i$ . This ensures that  $N(C_i) \leq \epsilon^{-d}\theta_d(12a\lambda^{-1/d} + 6\delta)^d$ . Also,  $i \in C_j$  iff  $j \in C_i$  so  $|\{j : i \in C_j\}| \leq |N(C_i)|$ . Therefore we can take  $\kappa$  to be  $\epsilon^{-d}\theta_d(12a\lambda^{-1/d} + 6\delta)^d$ .

Now we need an upper bound for the third moment. We can simply say  $E|X_i|^3 = E|Y_i - E[Y_i]|^3 \leq \epsilon^{3d}$ . Therefore  $\sum E|X_i|^3 \leq \epsilon^{-d}\epsilon^{3d} = \epsilon^{2d}$ .

So using Lemma 4.1 with  $p = 3$  gives

$$d_K\left(\frac{\bar{T}}{\sigma}, Z\right) \leq 75[\epsilon^{-d}\theta_d(12a\lambda^{-1/d} + 6\delta)^d]^2[\epsilon^{2d}]\sigma^{-3}$$

for any  $\delta > 0$ . Taking the limit as  $\delta$  tends to zero gives

$$d_K\left(\frac{\bar{T}}{\sigma}, Z\right) \leq 75\theta_d^2(12a\lambda^{-1/d})^{2d}\sigma^{-3} \quad (4.3)$$

which implies (4.1).

For the variance we have

$$\begin{aligned} \text{Var}[T] &= E[T^2] - (E[T])^2 = \int_{[0,1]^d} \int_{[0,1]^d} (P[E_x \cap E_y] - P[E_x]P[E_y]) dx dy \\ &\geq \int_{[a\lambda^{-1/d}, 1-a\lambda^{-1/d}]^d} \int_{D_y \cap [0,1]^d} (P[E_x \cap E_y] - P[E_x]P[E_y]) dx dy \end{aligned}$$

where  $D_y$  is the ball of radius  $a\lambda^{-1/d}$  around  $y$  and  $E_y$  is the event that  $y$  is covered (note that events  $E_x$  and  $E_y$  are either independent or positively correlated for all  $x, y$ ). Now given  $x$  and  $y$ , let the expected number of points of  $\mathcal{P}_\lambda$  in  $D_x$ ,  $D_y$  and  $D_x \cap D_y$  be denoted  $v_1, v_2$  and  $u$  respectively. Then

$$\begin{aligned} &P[E_x \cap E_y] - P[E_x]P[E_y] \\ &= (1 - e^{-u}) + e^{-u}(1 - e^{-v_1})(1 - e^{-v_2}) - (1 - e^{-v_1})(1 - e^{-v_2}) \\ &= e^{-(v_1+v_2)}(e^u - 1). \end{aligned}$$

Now with  $x$  in  $D_y$ , the ball of radius  $(a/2)\lambda^{-1/d}$  around the midpoint of  $xy$  is contained in  $D_x \cap D_y$ , and for  $y \in [a\lambda^{-1/d}, 1 - a\lambda^{-1/d}]^d$  (i.e.,  $y$  away from the boundary of the unit cube), this ball lies inside the unit cube, so we can bound  $u$  by

$$u \geq \lambda f_{\min} \left(\frac{a}{2\lambda^{1/d}}\right)^d \theta_d = 2^{-d} f_{\min} a^d \theta_d$$

and we can bound  $v_1$  by

$$v_1 \leq \lambda f_{\max} \left(\frac{a}{\lambda^{1/d}}\right)^d \theta_d = f_{\max} a^d \theta_d.$$

The same bound also holds for  $v_2$ . The volume of  $D_y \cap [0, 1]^d$  is at least  $\frac{a^d \theta_d}{\lambda}$  so putting all of this together gives

$$\text{Var}[T] \geq \frac{a^d \theta_d}{\lambda} \exp(-2f_{\max} a^d \theta_d) (\exp(2^{-d} f_{\min} a^d \theta_d) - 1) (1 - 2a\lambda^{-1/d})^d.$$

Since  $\sigma^2 = \text{Var}[\bar{T}] = \text{Var}[T]$ , we can now substitute this into (4.3) to get

$$d_K \left( \frac{\bar{T}}{\sigma}, Z \right) \leq 75\theta_d^2 (12a\lambda^{-1/d})^{2d} \\ \times \left( \frac{a^d \theta_d}{\lambda} \exp(-2f_{\max} a^d \theta_d) (\exp(2^{-d} f_{\min} a^d \theta_d) - 1) (1 - 2a\lambda^{-1/d})^d \right)^{-3/2}$$

and (4.2) follows.  $\square$

With  $d = 2$  and the coverage process on the unit torus we can also reduce the variance integral to

$$\text{Var}[T] = \frac{4\pi a^2 \exp(-2\pi a^2)}{\lambda} \int_0^{\pi/2} \cos \theta [\exp(3a^2(2\theta - \sin 2\theta)) - \exp(a^2(2\theta - \sin 2\theta))] d\theta$$

With  $a = 1/\sqrt{\pi}$  this gives  $\text{Var}[T] \approx 0.45/\lambda$  which gives the bound

$$d_K \left( \frac{T - ET}{\sigma}, Z \right) \leq 75(144)^d (\theta_d a^d \lambda^{-1})^2 \sigma^{-3} \approx 5.1 * 10^6 \lambda^{-1/2}.$$

## 5 Proof of Theorems 2.1 and 2.2

Our proof of Theorem 2.1 is based on a result of Chen and Shao (5), which we now describe.

Let  $Z^d$  denote the  $d$ -dimensional integer lattice. The distance between two points  $i = (i_1, \dots, i_d)$  and  $j = (j_1, \dots, j_d)$  in  $Z^d$  is defined by  $|i - j| = \max_{1 \leq l \leq d} |i_l - j_l|$  and the distance between two sets  $A$  and  $B$  is defined by  $\rho(A, B) = \inf\{|i - j| : i \in A, j \in B\}$ . For a given subset  $\mathcal{J}$  of  $Z^d$ , a set of random variables  $\{X_i, i \in \mathcal{J}\}$  is said to be an  $m$ -dependent random field if  $\{X_i, i \in A\}$  and  $\{X_j, j \in B\}$  are independent whenever  $\rho(A, B) > m$  for any subsets  $A$  and  $B$  of  $\mathcal{J}$ .

**Lemma 5.1** (5) *Let  $\{X_i, i \in \mathcal{J}\}$  be an  $m$ -dependent random field with zero means and finite  $E|X_i|^p$  for  $2 < p \leq 3$ . Let  $W = \sum X_i$ , let  $\sigma^2 = \text{Var}[T]$ , then*

$$d_K \left( \frac{W}{\sigma}, Z \right) \leq 75(10m + 1)^{(p-1)d} \sum E|X_i|^p \sigma^{-p}.$$

*Proof of Theorem 2.1.* The random variables  $X_i$  are now a 2-dependent random field so we can apply Lemma 5.1 with  $m = 2$  and  $p = 3$  to get the result.  $\square$

**Lemma 5.2** *Suppose  $\gamma_p < \infty$  for some  $p > 3$ , with  $\alpha := p/(p-3) \in \mathbb{N}$ . Then for any  $a > 0$ ,*

$$\beta_3(a) \leq 8\gamma_p^{3/p} a^d f_{\max} (1 + 6\alpha + 16\alpha^2 + 6a^d f_{\max} + 4a^{2d} f_{\max}^2). \quad (5.1)$$

and

$$\tilde{\beta}_3(a) \leq \gamma_p^{3/p} a^d f_{\max} (1 + 6\alpha + 16\alpha^2 + 6a^d f_{\max} + 4a^{2d} f_{\max}^2), \quad (5.2)$$

while if  $\gamma_4 < \infty$  then

$$\beta_2(a) \leq 4\gamma_4^{1/2} a^d f_{\max} (5 + 2a^d f_{\max}). \quad (5.3)$$

Note that in Lemma 5.2 we do *not* make any assumption on range of interaction of the functional  $\xi$ . Lemma 5.2 is similar to Lemma 4.3 of (18), but with a more explicit constant and a different proof than there.

*Proof of Lemma 5.2.* We wish to provide a bound on  $\beta_3$  in terms of  $\gamma_p$ , which we assume to be finite for some  $p > 3$ . Note first that by the Hölder or Minkowski inequality, with  $X_m$  and  $Y_m$  defined by (2.2),

$$E[|X_m|^3] = E[|Y_m - EY_m|^3] \leq E[(|Y_m| + |EY_m|)^3] \leq 8E[|Y_m|^3]. \quad (5.4)$$

In the case where  $\xi$  takes only nonnegative values, we have the better bound

$$\begin{aligned} E[|X_m|^3] &= E[|Y_m - EY_m|^3] \leq E[\max(Y_m, EY_m)^3] \leq E[Y_m^3 + (EY_m)^3] \\ &\leq 2E[Y_m^3]. \end{aligned} \quad (5.5)$$

Setting  $N_m$  to be the number of points of  $\mathcal{P}_\lambda$  in  $C_m$ , and listing the points of  $\mathcal{P}_\lambda \cap C_m$  as  $y_1, \dots, y_{N_m}$ , we have that

$$\begin{aligned} E[|Y_m|^3] &\leq E\left[\left(\sum_{i=1}^{N_m} |\xi(y_i, \mathcal{P}_\lambda)|\right)^3\right] \\ &= E\sum_i |\xi(y_i, \mathcal{P}_\lambda)|^3 + 3E\sum_{(i,j):i \neq j} |\xi(y_i, \mathcal{P}_\lambda)|^2 |\xi(y_j, \mathcal{P}_\lambda)| \\ &\quad + E\sum_{(i,j,k):i \neq j \neq k \neq i} |\xi(y_i, \mathcal{P}_\lambda)| |\xi(y_j, \mathcal{P}_\lambda)| |\xi(y_k, \mathcal{P}_\lambda)| \end{aligned}$$

so that by Palm theory for the Poisson process (see e.g. Theorem 1.6 of (14)), writing  $\mathcal{P}_\lambda^x$  for  $\mathcal{P}_\lambda \cup \{x\}$  and  $\mathcal{P}_\lambda^{x,y}$  for  $\mathcal{P}_\lambda \cup \{x, y\}$ , we have that

$$E[|Y_m|^3] \leq A_1 + 3A_2 + A_3 \quad (5.6)$$

where we set

$$\begin{aligned} A_1 &:= \lambda \int_{C_m} E|\xi(x, \mathcal{P}_\lambda)|^3 f(x) dx; \\ A_2 &:= \lambda^2 \int_{C_m} f(x) dx \int_{C_m} f(y) dy E[|\xi(x, \mathcal{P}_\lambda^y)|^2 |\xi(y, \mathcal{P}_\lambda^x)|] \\ A_3 &:= \lambda^3 \int_{C_m} f(x) dx \int_{C_m} f(y) dy \int_{C_m} f(z) dz E[|\xi(x, \mathcal{P}_\lambda^{y,z})| |\xi(y, \mathcal{P}_\lambda^{x,z})| |\xi(z, \mathcal{P}_\lambda^{x,y})|]. \end{aligned}$$

Also, similar arguments to (5.4) and (5.6) yield

$$E[|X_m|^2] \leq 4E[|Y_m|^2] \leq 4A_4 + 4A_5, \quad (5.7)$$

where we set

$$A_4 := \lambda \int_{C_m} E[|\xi(x, \mathcal{P}_\lambda)|^2] f(x) dx; \quad (5.8)$$

$$A_5 := \lambda^2 \int_{C_m} f(x) dx \int_{C_m} f(y) dy E[|\xi(x, \mathcal{P}_\lambda^y)| |\xi(y, \mathcal{P}_\lambda^x)|] \quad (5.9)$$

Consider in turn the terms in the right hand side of (5.6). First, since  $EN_m = \lambda \int_{C_m} f(x)dx$ , Hölder's inequality implies for any  $p > 3$  that

$$A_1 \leq \lambda \int_{C_m} \gamma_3 f(x)dx \leq \gamma_p^{3/p} (EN_m). \quad (5.10)$$

Turning to  $A_2$ , let  $U, V, W$  be independent continuous random  $d$ -vectors taking values in  $C_m$  with common density function  $f(\cdot)/\int_{C_m} f(x)dx$ , and independent of  $\mathcal{P}_\lambda$ . Then by Hölder's inequality,

$$\begin{aligned} A_2 &= (EN_m)^2 \cdot E[|\xi(U, \mathcal{P}_\lambda^V)^2 \xi(V, \mathcal{P}_\lambda^U)|] \\ &\leq (EN_m)^2 (E[|\xi(U, \mathcal{P}_\lambda^V)|^3])^{2/3} (E[|\xi(V, \mathcal{P}_\lambda^U)|^3])^{1/3}. \end{aligned} \quad (5.11)$$

Now let  $\mathcal{P}'_\lambda$  be a Poisson process on  $(0, 1)^d \setminus C_m$  with intensity  $\lambda f(\cdot)$ , independent of  $(U, V, W)$ . For  $j \in \mathbb{N}$  let  $\mathcal{X}_j$  denote a binomial point process consisting of  $j$  independent identically distributed random  $d$ -vectors taking values in  $C_m$  with common density function  $f(\cdot)/\int_{C_m} f(x)dx$ , independent of  $(U, V, W, \mathcal{P}'_\lambda)$ . Then

$$\begin{aligned} E[|\xi(U, \mathcal{P}'_\lambda)|^3] &= \sum_{j=0}^{\infty} E[|\xi(U, \mathcal{X}_{j+1} \cup \mathcal{P}'_\lambda)|^3] P[N_m = j] \\ &= \sum_{j=0}^{\infty} E[|\xi(U, \mathcal{X}_{j+1} \cup \mathcal{P}'_\lambda)|^3] P[N_m = j+1] \left(\frac{j+1}{EN_m}\right) \\ &\leq \sum_{i=0}^{\infty} E[|\xi(U, \mathcal{X}_i \cup \mathcal{P}'_\lambda)|^3] P[N_m = i] \left(\frac{i}{EN_m}\right) \\ &= E \left[ |\xi(U, \mathcal{P}_\lambda)|^3 \cdot \left(\frac{N_m}{EN_m}\right) \right] \\ &\leq \gamma_p^{3/p} \left( E \left[ \left(\frac{N_m}{EN_m}\right)^{p/(p-3)} \right] \right)^{(p-3)/p} \end{aligned} \quad (5.12)$$

where the last line comes from Hölder's inequality.

Assuming  $p$  is chosen so that  $p > 3$  and  $\alpha := p/(p-3) \in \mathbb{N}$ , and writing  $(n)_\alpha$  for the descending factorial  $n(n-1)\cdots(n-\alpha+1)$ , we have for  $n \geq 2\alpha$  that  $n^\alpha \leq 2^\alpha (n)_\alpha$  so that since  $N_m$  is Poisson distributed,

$$\begin{aligned} E[N_m^\alpha] &= E[N_m^\alpha \mathbf{1}\{N_m > 2\alpha\}] + E[N_m^\alpha \mathbf{1}\{N_m \leq 2\alpha\}] \\ &\leq 2^\alpha E[(N_m)_\alpha] + (2\alpha)^\alpha = 2^\alpha (EN_m)^\alpha + 2^\alpha \alpha^\alpha \end{aligned}$$

so

$$E \left[ \left(\frac{N_m}{EN_m}\right)^\alpha \right]^{1/\alpha} \leq 2 \left( 1 + \left(\frac{\alpha}{EN_m}\right)^\alpha \right)^{1/\alpha} \leq 2 \left( 1 + \frac{\alpha}{EN_m} \right) \quad (5.13)$$

so that by (5.12),

$$E[|\xi(U, \mathcal{P}'_\lambda)|^3] \leq 2\gamma_p^{3/p} (1 + \alpha/EN_m)$$

and the same bound naturally holds for  $E[\xi(V, \mathcal{P}_\lambda^U)^3]$ . Hence by (5.11),

$$A_2 \leq 2(EN_m)\gamma_p^{3/p}(EN_m + \alpha) \quad (5.14)$$

Next we argue similarly for  $A_3$ . We have that

$$A_3 = (EN_m)^3 E[|\xi(X, \mathcal{P}_\lambda^{Y,Z})\xi(Y, \mathcal{P}_\lambda^{X,Z})\xi(Z, \mathcal{P}_\lambda^{X,Y})|]$$

so by Hölder's inequality and the equality in distribution of  $\xi(X, \mathcal{P}_\lambda^{Y,Z})$ ,  $\xi(Y, \mathcal{P}_\lambda^{X,Z})$ , and  $\xi(Z, \mathcal{P}_\lambda^{X,Y})$ ,

$$A_3 \leq (EN_m)^3 E[|\xi(X, \mathcal{P}_\lambda^{Y,Z})|^3]. \quad (5.15)$$

Also, with the same notation as before,

$$\begin{aligned} E[|\xi(X, \mathcal{P}_\lambda^{Y,Z})|^3] &= \sum_{j=0}^{\infty} E[|\xi(X, \mathcal{X}_{j+2} \cup \mathcal{P}'_\lambda)|^3] P[N_m = j] \\ &= \sum_{j=0}^{\infty} E[|\xi(X, \mathcal{X}_{j+2} \cup \mathcal{P}'_\lambda)|^3] P[N_m = j+2] \left( \frac{(j+1)(j+2)}{(EN_m)^2} \right) \\ &\leq \sum_{i=0}^{\infty} E[|\xi(X, \mathcal{X}_i \cup \mathcal{P}'_\lambda)|^3] P[N_m = i] \left( \frac{i(i-1)}{(EN_m)^2} \right) \\ &= E \left[ |\xi(X, \mathcal{P}_\lambda)|^3 \left( \frac{N_m(N_m-1)}{(EN_m)^2} \right) \right] \\ &\leq \gamma_p^{3/p} E \left[ \left( \frac{N_m(N_m-1)}{(EN_m)^2} \right)^{p/(p-3)} \right]^{(p-3)/p}. \end{aligned} \quad (5.16)$$

Assuming  $\alpha := p/(p-3) \in \mathbb{N}$  as before, we have that

$$\begin{aligned} E[(N_m(N_m-1))^\alpha] &= E[(N_m(N_m-1))^\alpha \mathbf{1}\{N_m > 4\alpha\}] \\ &\quad + E[(N_m(N_m-1))^\alpha \mathbf{1}\{N_m \leq 4\alpha\}] \\ &\leq 2^{2\alpha} E[(N_m)_{2\alpha}] + (4\alpha)^{2\alpha} \\ &= 4^\alpha (EN_m)^{2\alpha} + (4\alpha)^{2\alpha} \end{aligned}$$

so that

$$E \left[ \left( \frac{N_m(N_m-1)}{(EN_m)^2} \right)^\alpha \right]^{1/\alpha} \leq \left( 4^\alpha + \frac{16^\alpha \alpha^{2\alpha}}{(EN_m)^{2\alpha}} \right)^{1/\alpha} \leq 4 + \frac{16\alpha^2}{(EN_m)^2}$$

so that by (5.15) and (5.16),

$$\begin{aligned} A_3 &\leq (EN_m)^3 \gamma_p^{3/p} \left( 4 + \frac{16\alpha^2}{(EN_m)^2} \right) \\ &= 4\gamma_p^{3/p} (EN_m) ((EN_m)^2 + 4\alpha^2). \end{aligned} \quad (5.17)$$

Combining (5.6), (5.10) (5.14) and (5.17), we have that

$$E[|Y_m|^3] \leq \gamma_p^{3/p}(EN_m) (1 + 6(EN_m + \alpha) + 4((EN_m)^2 + 4\alpha^2)). \quad (5.18)$$

By the choice of box side as  $a/\lambda^{1/d}$ , we have the bound  $EN_m \leq \lambda f_{\max}(a/\lambda^{1/d})^d = a^d f_{\max}$ , for all  $m$ . Hence, by (5.18), (5.2) holds. Also, using (5.4), we obtain (5.1).

It remains to prove (5.3), which we do by bounding  $A_4$  and  $A_5$  given by (5.8) and (5.9). First, note that by (5.8) and an analogous argument to (5.10),

$$A_4 \leq \lambda \int_{C_m} \gamma_2 f(x) dx \leq \gamma_4^{1/2}(EN_m). \quad (5.19)$$

Also, by (5.9), and a similar argument to (5.11),

$$A_5 \leq (EN_m)^2 (E[|\xi(U, \mathcal{P}_\lambda^V)|^2])^{1/2} (E[|\xi(V, \mathcal{P}_\lambda^U)|^2])^{1/2} = (EN_m)^2 E[|\xi(U, \mathcal{P}_\lambda^V)|^2]$$

and by a similar argument to (5.12),

$$\begin{aligned} E[|\xi(U, \mathcal{P}_\lambda^V)|^2] &\leq E \left[ |\xi(U, \mathcal{P}_\lambda)|^2 \cdot \left( \frac{N_m}{EN_m} \right) \right] \\ &\leq \gamma_4^{1/2} \left( E \left[ \left( \frac{N_m}{EN_m} \right)^2 \right] \right)^{1/2} \end{aligned}$$

so by the case  $\alpha = 2$  of (5.13),

$$A_5 \leq 2(EN_m)\gamma_4^{1/2}(EN_m + 2). \quad (5.20)$$

Since  $EN_m \leq a^d f_{\max}$ , combining (5.20) with (5.7) and (5.19) gives us the bound

$$E[|X_m|^2] \leq 4\gamma_4^{1/2}(a^d f_{\max})(2a^d f_{\max} + 5)$$

which implies (5.3).  $\square$

*Proof of Theorem 2.2.* By Theorem 2.1 and Lemma 5.2, with  $\alpha = p/(p-3)$  we have that

$$\begin{aligned} d_K \left( \frac{T - ET}{\sigma}, Z \right) &\leq 600 * 21^{2d} \left( \lceil \frac{\lambda^{1/d}}{a} \rceil \right)^d \sigma^{-3} \gamma_p^{3/p} a^d f_{\max} \\ &\quad \times (1 + 6\alpha + 16\alpha^2 + 6a^d f_{\max} + 4a^{2d} f_{\max}^2). \end{aligned} \quad (5.21)$$

Putting  $x = a^{-d}\lambda$  in (2.4), we obtain the bound  $(\lceil a^{-1}\lambda^{1/d} \rceil)^d \leq a^{-d}\lambda h_d(a^{-d}\lambda)$ .

Taking  $p = 4$ , so that  $\alpha = 4$ , in (5.21) yields

$$\begin{aligned} d_K \left( \frac{T - ET}{\sigma}, Z \right) &\leq 600 h_d(a^{-d}\lambda) 21^{2d} \lambda \sigma^{-3} \gamma_4^{3/4} f_{\max} \\ &\quad \times (1 + 24 + 256 + 6a^d f_{\max} + 4a^{2d} f_{\max}^2). \end{aligned}$$

If also  $a^d f_{\max} \geq 1$  then

$$d_K \left( \frac{T - ET}{\sigma}, Z \right) \leq 600 h_d(a^{-d} \lambda) 21^{2d} \lambda \sigma^{-3} \gamma_4^{3/4} f_{\max} (291 a^{2d} f_{\max}^2)$$

and (2.5) follows.

## 6 Preliminaries to the proof of Theorem 2.3

In this section we provide a series of estimates, which we shall ultimately use in the Section 7 to prove Theorem 2.3. First we give a simple estimate for the tail of the standard normal distribution.

**Lemma 6.1** *For all  $x > 0$ , it is the case that*

$$P[|Z| > x] \leq x^{-3}. \quad (6.1)$$

*Proof.* A standard integration by parts argument shows that  $P[Z > x]$  is bounded above by  $x^{-1}(2\pi)^{-1/2}e^{-x^2/2}$ , so that

$$x^3 P[|Z| > x] \leq (2/\pi)^{1/2} x^2 e^{-x^2/2}$$

and this bound takes a maximum value (at  $x = \sqrt{2}$ ) of  $(2/\pi)^{1/2}(2/e)$  which is less than 1.  $\square$

Let  $a$  and  $R$  be as in the statement of Theorem 2.3, and set

$$\begin{aligned} \xi'(x, \Psi) &:= \xi(x, \Psi) \mathbf{1}\{R(x, \Psi) \leq a\lambda^{-1/d}\}; \\ \tilde{T} &= \sum_{i=1}^N \xi'(x_i, \mathcal{P}_\lambda); \quad \tilde{\sigma} := \text{Var}(\tilde{T}). \end{aligned} \quad (6.2)$$

**Lemma 6.2** *Suppose the assumptions of Theorem 2.3 hold. Then*

$$P[T \neq \tilde{T}] \leq \lambda^{-8}, \quad (6.3)$$

$$E[|T - \tilde{T}|] \leq \gamma_4^{1/4} \lambda^{-23/4}. \quad (6.4)$$

and

$$E[|T|] \leq \gamma_4^{1/4} \lambda; \quad E[|\tilde{T}|] \leq \gamma_4^{1/4} \lambda. \quad (6.5)$$

*Proof.* By Palm theory for the Poisson process (see e.g. Theorem 1.6 of (14)),

$$\begin{aligned} P[T \neq \tilde{T}] &\leq E \sum_{i=1}^N \mathbf{1}\{R_k(x_i, \mathcal{P}_\lambda) > a\lambda^{-1/d}\} \\ &= \lambda \int_{(0,1)^d} P[R(x, \mathcal{P}_\lambda) > a\lambda^{-1/d}] f(x) dx \end{aligned}$$



and (6.3) follows from (2.6); similarly,

$$\begin{aligned} E[|T - \tilde{T}|] &= E \left[ \left| \sum_{i=1}^N \xi(x_i, \mathcal{P}_\lambda) \mathbf{1}\{R_k(x_i; \mathcal{P}_\lambda) > a\lambda^{-1/d}\} \right| \right] \\ &\leq E \left[ \sum_{i=1}^N \left| \xi(x_i, \mathcal{P}_\lambda) \mathbf{1}\{R(x_i; \mathcal{P}_\lambda) > a\lambda^{-1/d}\} \right| \right] \\ &= \lambda \int_{(0,1)^d} E[|\xi(x, \mathcal{P}_\lambda)| \mathbf{1}\{R(x; \mathcal{P}_\lambda) > a\lambda^{-1/d}\}] f(x) dx \end{aligned}$$

so that by Hölder's inequality,

$$E[|T - \tilde{T}|] \leq \lambda \int_{(0,1)^d} (E[|\xi(x, \mathcal{P}_\lambda)|^4])^{1/4} (P[R(x; \mathcal{P}_\lambda) > a\lambda^{-1/d}])^{3/4} f(x) dx$$

and (6.4) follows by (2.6). Also, similar (simpler) arguments yield (6.5).  $\square$

Given any random variable  $X$ , and given  $p \geq 1$ , we use the standard notation  $\|X\|_p$  for  $(E[|X|^p])^{1/p}$ . We obtain the following bounds using Lemma 5.2 and Minkowski's inequality.

**Lemma 6.3** *Under the assumptions of Theorem 2.3, it is the case that*

$$\|T\|_3 \leq (291)^{1/3} h_d(\lambda) \gamma_4^{1/4} f_{\max} \lambda, \quad (6.6)$$

$$\|\tilde{T}\|_3 \leq (291)^{1/3} h_d(\lambda) \gamma_4^{1/4} f_{\max} \lambda, \quad (6.7)$$

and

$$\sigma \leq \sqrt{28} h_d(\lambda) \gamma_4^{1/4} f_{\max} \lambda. \quad (6.8)$$

*Proof.* We can write  $T = \sum_i \xi(x_i, \mathcal{P}_\lambda)$  as a sum of  $\lceil \lambda^{1/d} \rceil^d$  contributions from boxes of side  $\lambda^{-1/d}$ . By Minkowski's inequality and (5.2) with  $\alpha = 4$ , and the fact that  $f_{\max} \geq 1$ , we obtain the bound

$$\|T\|_3 \leq \lceil \lambda^{1/d} \rceil^d (\tilde{\beta}_3(1))^{1/3} \leq \lceil \lambda^{1/d} \rceil^d (291 \gamma_4^{3/4} f_{\max}^3)^{1/3}$$

and (6.6) follows by (2.4). Since  $\gamma_4$  (given by (2.3)) is also a bound for  $E[|\xi'(x, \mathcal{P}_\lambda)|^4]$ , a similar argument yields (6.7). Also, a further application of Minkowski's inequality, with  $X_m$  denoting the centred contribution from the  $m$ th box of side  $\lambda^{-1/d}$ , yields

$$\sigma = \sqrt{E \left[ \left( \sum_m X_m \right)^2 \right]} = \left\| \sum_m X_m \right\|_2 \leq \lceil \lambda^{1/d} \rceil^d \beta_2(1)^{1/2},$$

so by (2.4) and the case  $a = 1$  of (5.3), we obtain (6.8).  $\square$

The final estimate in this section is a bound, in terms of  $\sigma$ , on the relative difference between  $\sigma$  and  $\tilde{\sigma}$ .

**Lemma 6.4** *Provided  $\lambda \geq 1$ , it is the case that*

$$\left| \frac{\tilde{\sigma} - \sigma}{\sigma} \right| \leq 90h_d(\lambda)^2 \gamma_4^{1/2} f_{\max}^2 \lambda^{-2/3} \sigma^{-2} \quad (6.9)$$

*Proof.* By the Hölder and Minkowski inequalities,

$$\begin{aligned} |E[T^2] - E[\tilde{T}^2]| &= |E[(T^2 - \tilde{T}^2)\mathbf{1}\{T \neq \tilde{T}\}]| \\ &\leq \|T^2 - \tilde{T}^2\|_{3/2} (P[T \neq \tilde{T}])^{1/3} \\ &\leq (\|T^2\|_{3/2} + \|\tilde{T}^2\|_{3/2}) (P[T \neq \tilde{T}])^{1/3} \\ &= (\|T\|_3^2 + \|\tilde{T}\|_3^2) (P[T \neq \tilde{T}])^{1/3}, \end{aligned}$$

so by (6.3), (6.6), and (6.7),

$$|E[T^2] - E[\tilde{T}^2]| \leq 2(291)^{2/3} h_d(\lambda)^2 \gamma_4^{1/2} f_{\max}^2 \lambda^{-2/3}. \quad (6.10)$$

Also, using (6.4) and (6.5), we have that

$$|(ET)^2 - (E\tilde{T})^2| = |(ET + E\tilde{T})(E[T - \tilde{T}])| \leq 2\gamma_4^{1/2} \lambda^{-19/4}.$$

Combined with (6.10), using the fact that  $(291)^{2/3} < 44$ , this gives

$$\begin{aligned} |\sigma^2 - \tilde{\sigma}^2| &= |ET^2 - (ET)^2 - (E\tilde{T}^2 - (E\tilde{T})^2)| \\ &\leq 88h_d(\lambda)^2 \gamma_4^{1/2} f_{\max}^2 \lambda^{-2/3} + 2\gamma_4^{1/2} \lambda^{-19/4} \\ &\leq 90h_d(\lambda)^2 \gamma_4^{1/2} f_{\max}^2 \lambda^{-2/3} \end{aligned}$$

We can deduce (6.9) from this, together with the estimate

$$\left| \frac{\tilde{\sigma} - \sigma}{\sigma} \right| = \frac{|\tilde{\sigma}^2 - \sigma^2|}{\sigma(\sigma + \tilde{\sigma})} \leq \frac{|\tilde{\sigma}^2 - \sigma^2|}{\sigma^2}.$$

□

We now deal with the fact that in the statement of Theorem 2.3,  $R$  may be either a radius of stabilization or a stopping radius of localization.

**Lemma 6.5** *Suppose  $R$  is a radius of stabilization for  $\xi$ , and let  $\lambda \geq 1$ , and  $c > 0$ . Then there exists a stopping radius of localization for  $\xi$ , denoted  $\tilde{R}$ , such that for all  $x \in [0, 1]^d$*

$$P[\tilde{R}(x, \mathcal{P}_\lambda) > c] \leq P[R(x, \mathcal{P}_\lambda) > c]. \quad (6.11)$$

*Proof.* Let  $\mathcal{F}$  be the collection of finite point sets in  $[0, 1]^d$ . For  $x \in [0, 1]^d$ , and  $\Psi \in \mathcal{F}$ , set

$$\begin{aligned} S(x, \Psi) &:= \sup\{\xi(x, \Psi') : \Psi' \in \mathcal{F}, \Psi' \cap B_c(x) = \Psi \cap B_c(x)\}; \\ I(x, \Psi) &:= \inf\{\xi(x, \Psi') : \Psi' \in \mathcal{F}, \Psi' \cap B_c(x) = \Psi \cap B_c(x)\} \end{aligned}$$

and set

$$\tilde{R}(x, \Psi) = \begin{cases} c & \text{if } S(x, \Psi) = I(x, \Psi) \\ 2\sqrt{d} & \text{otherwise} \end{cases}$$

Then  $S(x, \Psi)$  is a measurable function of  $(x, \Psi)$  (see (16), page 1128), as is  $I(x, \Psi)$ . Hence  $\tilde{R}(x, \Psi)$  is a measurable function of  $(x, \Psi)$ . Also, by definition the value of  $\mathbf{1}\{\tilde{R}(x, \Psi) = c\}$  is determined by the pair  $(x, \Psi \cap B_c(x))$ , and for all  $x \in [0, 1]^d$  and  $\Psi \in \mathcal{F}$  we have  $\xi(x, \Psi) = \xi(x, \Psi \cap B_{\tilde{R}(x, \Psi)}(x))$  (this is easily checked for both the possible values of  $\tilde{R}(x, \Psi)$ ). In other words,  $\tilde{R}$  is indeed a stopping radius of localization. Moreover, it is easy to see from the definition that if  $R(x, \Psi) \leq c$  then  $\tilde{R}(x, \Psi) \leq c$ . Hence, (6.11) holds.  $\square$

## 7 Proof of Theorems 2.3 and 2.4

Before proceeding with the proof of Theorem 2.3, we note first that there is a small error in the proof (but not the statement) of the main results of (18). That proof proceeds similarly to the one here, via a variable  $\tilde{T}_\lambda$  defined similarly to our  $\tilde{T}$  given by (6.2); see Section 4.3 of (18). Shortly before (4.12) of that paper, a dependency graph argument is set out based on the assertion that the contribution to  $\tilde{T}_\lambda$  from a given box is determined by the restriction of the Poisson process to the union of neighbouring boxes, i.e. boxes within distance  $a\lambda^{-1/d}$  (in the notation of our (6.2)) from that box. This assertion is valid when  $R$  (in the definition at (6.2)) is a stopping radius of localization, but could fail if  $R$  is a radius of stabilization (as in the statement of results in (18)). Indeed, if  $R$  is as in the example discussed shortly after the statement of our Theorem 2.3, then there could be Poisson points  $x$  in a given box, for which the value of the indicator  $\mathbf{1}\{R(x, \mathcal{P}_\lambda) \leq a\lambda^{-1/d}\}$  is not determined by the restriction of  $\mathcal{P}$  to the union of neighbouring boxes, and hence the value of  $\tilde{T}$  is also not determined by the restriction to these boxes.

However, it is possible to fix the proof in (18) by using our Lemma 6.5 to replace the original radius of stabilization by a stopping radius of localization satisfying the same tail probability bound, as we do in the present paper.

*Proof of Theorem 2.3.* Define the ‘‘constant’’ (strictly speaking, a bounded function of  $\lambda$ ,  $\lambda \geq 1$ )

$$B := B(d, \lambda) := 21^{2d} 174600 h_d(\lambda).$$

Assume the conditions of Theorem 2.3 hold; by Lemma 6.5 we can (and do) assume without loss of generality that  $R$  is a stopping radius of localization. Let  $\lambda \geq 1$  and let  $a \geq f_{\max}^{-1/d}$  satisfying (2.6). Then  $\tilde{T}$  given by (6.2) is a sum of terms with range of interaction  $a\lambda^{-1/d}$ , and so by Theorem 2.2, we have

$$d_K \left( \frac{\tilde{T} - E\tilde{T}}{\tilde{\sigma}}, Z \right) \leq B\gamma_4^{3/4} f_{\max}^3 a^{2d} \lambda \tilde{\sigma}^{-3}. \quad (7.1)$$

By the proof of (4.19) of (18), for any two random variables  $Y$  and  $Y'$  on the same probability space, and any  $\varepsilon > 0$ , we have the estimate

$$d_K(Y, Z) \leq (2\pi)^{-1/2} \varepsilon + P[|Y - Y'| > \varepsilon] + d_K(Y', Z) \quad (7.2)$$

which we shall apply in the case where we take  $Y = (T - ET)/\sigma$ , and  $Y' = (\tilde{T} - E\tilde{T})/\tilde{\sigma}$ , and (motivated by (7.1)) take

$$\varepsilon := B\gamma_4^{3/4} f_{\max}^3 \lambda \sigma^{-3}. \quad (7.3)$$

We need to estimate the middle term in the right hand side of (7.2). Note first that for any  $\lambda \geq 1$ , since  $h_d(\lambda) \leq 2^d$  it is the case that

$$\frac{2(28)^{3/2}h_d(\lambda)^3}{B} \leq \frac{2 \times 149}{174600} \left( \frac{4}{21^2} \right)^d \leq 1 \quad (7.4)$$

and therefore by (6.8),

$$\sigma^3 \leq (B/2)\gamma_4^{3/4} f_{\max}^3 \lambda^3 \leq (B/2)\gamma_4^{3/4} f_{\max}^3 \lambda^9$$

so that by (6.3) and (7.3),

$$P[T \neq \tilde{T}] \leq \lambda^{-8} \leq (B/2)\gamma_4^{3/4} f_{\max}^3 \lambda \sigma^{-3} = \varepsilon/2. \quad (7.5)$$

Also, (7.4) implies that  $56h_d(\lambda)^2 \leq Bf_{\max}$ , so that by (6.8),

$$2\sigma^2 \leq B\gamma_4^{1/2} f_{\max}^3 \lambda^2 \leq B\gamma_4^{1/2} f_{\max}^3 \lambda^{27/4}$$

and therefore by (6.4),

$$|ET - E\tilde{T}| \leq \gamma_4^{1/4} \lambda^{-23/4} \leq (B/2)\gamma_4^{3/4} f_{\max}^3 \lambda \sigma^{-2} = \varepsilon\sigma/2. \quad (7.6)$$

Now we can estimate the middle term in the right hand side of (7.2), by

$$\begin{aligned} & P \left[ \left| \frac{T - ET}{\sigma} - \frac{\tilde{T} - E\tilde{T}}{\tilde{\sigma}} \right| > \varepsilon \right] \\ & \leq P \left[ \left| \frac{T - ET}{\sigma} - \frac{\tilde{T} - E\tilde{T}}{\sigma} \right| > \frac{\varepsilon}{2} \right] + P \left[ \left| \frac{\tilde{T} - E\tilde{T}}{\sigma} - \frac{\tilde{T} - E\tilde{T}}{\tilde{\sigma}} \right| \geq \frac{\varepsilon}{2} \right] \\ & \leq P[T \neq \tilde{T}] + P \left[ \left| \frac{\tilde{T} - E\tilde{T}}{\tilde{\sigma}} \left( \frac{\tilde{\sigma} - \sigma}{\sigma} \right) \right| \geq \frac{\varepsilon}{2} \right], \end{aligned} \quad (7.7)$$

where for the last line we have used (7.6). By (6.9),

$$\begin{aligned} P \left[ \left| \frac{\tilde{T} - E\tilde{T}}{\tilde{\sigma}} \left( \frac{\tilde{\sigma} - \sigma}{\sigma} \right) \right| \geq \frac{\varepsilon}{2} \right] & \leq P \left[ |Z| \geq \frac{\varepsilon \lambda^{2/3} \sigma^2}{180 h_d(\lambda)^2 \gamma_4^{1/2} f_{\max}^2} \right] \\ & \quad + 2d_K \left( \frac{\tilde{T} - E\tilde{T}}{\tilde{\sigma}}, Z \right), \end{aligned}$$

so by (7.7) and (7.5), we have that

$$\begin{aligned} P \left[ \left| \frac{T - ET}{\sigma} - \frac{\tilde{T} - E\tilde{T}}{\tilde{\sigma}} \right| > \varepsilon \right] & \leq \frac{\varepsilon}{2} + P \left[ |Z| \geq \frac{\varepsilon \lambda^{2/3} \sigma^2}{180 h_d(\lambda)^2 \gamma_4^{1/2} f_{\max}^2} \right] \\ & \quad + 2d_K \left( \frac{\tilde{T} - E\tilde{T}}{\tilde{\sigma}}, Z \right). \end{aligned}$$

Using (7.2), and the fact that  $(2\pi)^{-1/2} \leq 1/2$ , then yields

$$d_K \left( \frac{T - ET}{\sigma}, Z \right) \leq \varepsilon + P \left[ |Z| \geq \frac{\varepsilon \lambda^{2/3} \sigma^2}{180 h_d(\lambda)^2 \gamma_4^{1/2} f_{\max}^2} \right] + 3d_K \left( \frac{\tilde{T} - E\tilde{T}}{\tilde{\sigma}}, Z \right). \quad (7.8)$$

To estimate the last term in the right hand side of (7.8), we would like to replace the  $\tilde{\sigma}^{-3}$  in the right hand side of (7.1) by  $\sigma^{-3}$  times a suitable constant. Before doing this, assuming for now that  $\varepsilon \leq 1$ , we make the estimate

$$\begin{aligned} \frac{90 h_d(\lambda)^2 \varepsilon^{2/3} \lambda^{-4/3}}{B^{2/3}} &\leq \frac{90 h_d(\lambda)^{4/3}}{(21)^{4d/3} (174600)^{2/3}} \leq \left( \frac{2^{4/3}}{21^{4/3}} \right)^d \frac{90}{(174600)^{2/3}} \\ &\leq \left( \frac{2}{21} \right)^{4/3} \frac{90}{(174600)^{2/3}} \leq \frac{1}{799}. \end{aligned}$$

Then use (6.9) and (7.3) to make the estimate

$$\begin{aligned} \frac{\tilde{\sigma}^{-3}}{\sigma^{-3}} &= \left( 1 + \frac{\tilde{\sigma} - \sigma}{\sigma} \right)^{-3} \leq \left( 1 - 90 h_d(\lambda)^2 \gamma_4^{1/2} f_{\max}^2 \lambda^{-2/3} \sigma^{-2} \right)^{-3} \\ &= \left( 1 - \frac{90 h_d(\lambda)^2 \varepsilon^{2/3} \lambda^{-4/3}}{B^{2/3}} \right)^{-3} \leq \left( \frac{799}{798} \right)^3 \leq 1.004 \end{aligned}$$

so that (7.1) implies

$$d_K \left( \frac{\tilde{T} - E\tilde{T}}{\tilde{\sigma}}, Z \right) \leq (1.004) B \gamma_4^{3/4} f_{\max}^3 a^{2d} \lambda \sigma^{-3} = (1.004) \varepsilon a^{2d}.$$

Now we bound the middle term in the right side of (7.8). It is easily checked for  $\lambda \geq 1$  that  $5040 h_d(\lambda)^4 \leq B^{4/3}$ . Then by (6.8),

$$\begin{aligned} \lambda^2 &\geq 180 h_d(\lambda)^2 B^{-4/3} \gamma_4^{-1/2} f_{\max}^{-2} (28 h_d(\lambda)^2 \gamma_4^{1/2} f_{\max}^2 \lambda^2) \\ &\geq 180 h_d(\lambda)^2 B^{-4/3} \gamma_4^{-1/2} f_{\max}^{-2} \sigma^2 \end{aligned}$$

so that with  $\varepsilon$  given by (7.3),

$$\frac{B \gamma_4^{1/4} f_{\max} \lambda^{5/3}}{180 h_d(\lambda)^2 \sigma} \geq \frac{\sigma}{B^{1/3} \lambda^{1/3} f_{\max} \gamma_4^{1/4}} = \varepsilon^{-1/3}.$$

Then (by (7.3) again) the middle term of the right side of (7.8) satisfies

$$\begin{aligned} P \left[ |Z| \geq \frac{\varepsilon \lambda^{2/3} \sigma^2}{180 h_d(\lambda)^2 \gamma_4^{1/2} f_{\max}^2} \right] &= P \left[ |Z| > \frac{B \gamma_4^{1/4} f_{\max} \lambda^{5/3}}{180 h_d(\lambda)^2 \sigma} \right] \\ &\leq P \left[ |Z| \geq \varepsilon^{-1/3} \right] \\ &\leq \varepsilon \end{aligned}$$

where the last line comes from (6.1).

Putting together the preceding estimates for the terms in the right side of (7.8), gives

$$d_K \left( \frac{T - ET}{\sigma}, Z \right) \leq (3.012a^{2d} + 2) \varepsilon \quad (7.9)$$

which yields (2.7). Finally, we assumed earlier in the proof that  $\varepsilon \leq 1$ , but if this fails then clearly (7.9) holds anyway, and hence so does (2.7).  $\square$

*Proof of Theorem 2.4.* Assume that  $T$  is given by (2.1), with  $\xi(x, \Psi)$  is determined by  $x$  and the  $k$  nearest neighbours of  $x$  in the point set  $\Psi$ . Let  $R_k(x, \Psi)$  denote the distance from  $x$  to its  $k$ th nearest neighbour in the point set  $\Psi \setminus \{x\}$ , with  $R_k(x, \Psi) := d^{1/2}$  if  $\Psi \setminus \{x\}$  has fewer than  $k$  points. Then  $R_k$  is a stopping radius of localization for  $\Psi$ .

Set  $b := 18/(1 - \log 2)$ . Let  $\lambda \geq 1$  satisfying (2.8), and then set

$$a := \left( \frac{2}{(f_{\min} \theta_d)^{1/d}} \right) \max(b \log \lambda, 2k)^{1/d}. \quad (7.10)$$

Since the Euclidean unit ball is contained in the  $\ell_\infty$  unit ball in  $\mathbb{R}^d$ , for all  $d$  we have the bound

$$\theta_d \leq 2^d. \quad (7.11)$$

Therefore, since we assume  $k \geq 1$ , and since  $f_{\min} \leq 1$ , the value of  $a$  given by (7.10) satisfies  $a \geq 1$ , and hence  $a \geq f_{\max}^{-1/d}$ .

For  $x > 0$  set  $H(x) = 1 - x + x \log x$ , and recall the Chernoff-type bound for the Poisson distribution (see e.g. (1.11) of (14)):

$$P[\text{Po}(\mu) < \mu'] \leq \exp(-\mu H(\mu'/\mu)), \quad 0 < \mu' < \mu.$$

By the choice of  $b$ ,  $H(\frac{1}{2})b = \frac{1}{2}(1 - \log 2)b = 9$ . Also, the condition (2.8) ensures that  $a\lambda^{-1/d} \leq 1/2$ , so that for  $x \in [0, 1]^d$ ,

$$\begin{aligned} P[R_k(x, \mathcal{P}_\lambda) > a\lambda^{-1/d}] &\leq P[\text{Po}(2^{-d}\lambda f_{\min} \theta_d a^d / \lambda) < k] \\ &= P[\text{Po}(\max(b \log \lambda, 2k)) < k] \\ &\leq \exp(-H(\frac{1}{2})b \log \lambda) = \lambda^{-9}, \end{aligned}$$

and therefore if we make this choice of  $R = R_k$ , and this choice (7.10) of  $a$ , all the conditions of Theorem 2.3 hold. Hence that result, along with the definition (7.10) of  $a$ , yields

$$d_K \left( \frac{T - ET}{\sigma}, Z \right) \leq \left( 3.012 \left( \frac{4^d}{(f_{\min} \theta_d)^2} \right) \max(b \log \lambda, 2k)^2 + 2 \right) \varepsilon$$

with

$$\varepsilon = 174600(21)^{2d} h_d(\lambda a^{-d}) \gamma_4^{3/4} f_{\max}^3 \lambda \sigma^{-3}.$$

and since  $\frac{b^2}{4} \leq 861$ ,

$$d_K \left( \frac{T - ET}{\sigma}, Z \right) \leq \left( \frac{4^d(12.048) \max(861(\log \lambda)^2, k^2)}{(f_{\min} \theta_d)^2} + 2 \right) \varepsilon.$$

To simplify this further, note that by (7.11) we have  $2 \leq 2(4^d \theta_d^{-2})$ , and since also  $f_{\min} \leq 1$ , and we assume  $k \geq 1$ , we obtain

$$\begin{aligned} d_K \left( \frac{T - ET}{\sigma}, Z \right) &\leq \left( \frac{4^d(14.048) \max(861(\log \lambda)^2, k^2)}{(f_{\min} \theta_d)^2} \right) \\ &\quad \times 21^{2d} 174600 h_d(\lambda a^{-d}) \gamma_4^{3/4} f_{\max}^3 \lambda \sigma^{-3} \\ &\leq \frac{(1764)^d (2452781) h_d(\lambda a^{-d}) \gamma_4^{3/4} f_{\max}^3 \max(861(\log \lambda)^2, k^2) \lambda}{(f_{\min} \theta_d)^2 \sigma^3}. \end{aligned}$$

This gives Theorem 2.4 as required.  $\square$

**Remarks.** In proving Theorem 2.3, we have followed the scheme of (18). This argument involves a rather coarse discretization of the unit cube so as to use a result of (5) on normal approximation for  $k$ -dependent random fields. Barbour and Xia (2) use a different result of (5), and avoid this coarse discretization. It is possible that by following the proof of Theorem 3.4 of (2), and keeping track of the constants there, one could reduce the constant in the bound of Theorem 2.3, and hence also in Theorem 2.4.

Even staying within the scheme of the proof given here, it might be possible to reduce the bounding constant somewhat by adopting a less crude bound on  $\sigma^2$  than the one we have used at (6.8). As was done in (18) and (2), one can by using the local dependence of the terms contributing to  $\tilde{\sigma}^2$ , bound  $\tilde{\sigma}$  by a similar bound to that in (6.8) but with the factor of  $\lambda$  changed to  $(\lambda a^d)^{1/2}$  (and the constant factor changed), and then using the proof of Lemma 6.4 deduce bound of similar form for  $\sigma$ . Following the effect of this in our proofs it appears to be that we might be able to take  $\lambda^{-6}$  instead of  $\lambda^{-9}$  in Theorem 2.3, and hence reduce the constant of 861 in Theorem 2.4, although some of the other constants might also change. We have not checked how this would work out in detail.

## Acknowledgements

We thank Joseph Yukich for several helpful comments on earlier versions of this paper which led to improved presentation.

## References

- [1] Avram, F. and Bertsimas, D. (1993). On central limit theorems in geometrical probability. *Ann. Appl. Probab.* **3**, 1033–1046.
- [2] Barbour, A. D. and Xia, A. (2006). Normal approximation for random sums. *Adv. in Appl. Probab.* **38**, 693–728.

- [3] Baryshnikov, Yu., Penrose, M. D. and Yukich J. E. (2008). Gaussian Limits for Generalized Spacings. *Ann. Appl. Probab.*, to appear.
- [4] Chatterjee, S. (2008). A new method of normal approximation. *Ann. Probab.* **36**, 1584–1610.
- [5] Chen, L. H. Y. and Shao, Q.-M. (2004). Normal approximation under local dependence. *Ann. Probab.* **32**, 1985–2028.
- [6] Feller, W. (1966). *An Introduction to Probability Theory and its Applications, Volume II*. Wiley, New York.
- [7] F. Götze, L. Heinrich and C. Hipp (1995).  $m$ -dependent random fields with analytic cumulant generating function. *Scand. J. Statist* **22**, 183–195.
- [8] P. Hall (1988). *Introduction to the Theory of Coverage Processes*. Wiley, New York.
- [9] L. Heinrich (1993). Asymptotic properties of minimum contrast estimators for parameters of Boolean models. *Metrika* **40**, 67–94.
- [10] L. Heinrich and I. Molchanov (1999). Central limit theorem for a class of random measures associated with germ-grain models. *Adv. Appl. Probab.* **31**, 283-314.
- [11] Levina, E. and Bickel, P. J. (2005). Maximum likelihood estimation of intrinsic dimension. In *Advances in NIPS 17*, eds. L.K. Saul Y. Weiss, and L. Bottou.
- [12] S. Mase (1982). Asymptotic properties of stereological estimators of volume fraction for stationary random sets. *J. Appl. Probab.* **19** 111–126.
- [13] I. Molchanov and D. Stoyan (1994). Asymptotic properties of estimators for parameters of the Boolean model. *Adv. Appl. Probab.* **26**, 301–323.
- [14] Penrose, M. (2003). *Random Geometric Graphs*. Oxford University Press, Oxford.
- [15] Penrose, M. D. (2007). Gaussian limits for random geometric measures. *Electron. J. Probab.* **12**, 989–1035
- [16] Penrose, M. D. (2007). Laws of large numbers in stochastic geometry with statistical applications. *Bernoulli* **13**, 1124-1150.
- [17] Penrose, M. D. and Wade, A.R. (2008). Multivariate normal approximation in geometric probability. *J. Stat. Theory Pract.* **2**, 293-326.
- [18] Penrose, M. D. and Yukich, J. E. (2005). Normal approximation in geometric probability. *Stein's Method and Applications*, 37–58, Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., **5**, Singapore Univ. Press, Singapore.
- [19] M.D. Penrose and J.E. Yukich (2001). Central limit theorems for some graphs in computational geometry. *Ann. Appl. Probab.* **11**, 1005-1041.



- [20] Schreiber, T. (2008). Limit theorems in stochastic geometry. *New Perspectives in Stochastic Geometry*, eds. W.S. Kendall and I. Molchanov, Oxford University Press, to appear.
- [21] Steele, J. M. (1997). *Probability theory and combinatorial optimization*. Society for Industrial and Applied Mathematics, Philadelphia.
- [22] Yukich, J. E. (1998). *Probability theory of classical Euclidean optimization problems*. Springer, Berlin.