Probab. Theory Relat. Fields 97, 35–63 (1993)

Probability Theory and Related Fields

Finite clusters in high-density continuous percolation: compression and sphericality

Kenneth S. Alexander*

Department of Mathematics DRB 155, University of Southern California, Los Angeles, CA 90089-1113, USA

Received August 23, 1991; in revised form November 2, 1992

Summary. A percolation process in $\mathbb{R}^d$ is considered in which the sites are a Poisson process with intensity ρ and the bond between each pair of sites is open if and only if the sites are within a fixed distance r of each other. The distribution of the number of sites in the cluster C of the origin is examined, and related to the geometry of C. It is shown that when ρ and k are large, there is a characteristic radius λ such that conditionally on $|C| = k$, the convex hull of C closely approximates a ball of radius λ, with high probability. When the normal volume k/ρ that k points would occupy is small, the cluster is compressed, in that the number of points per unit volume in this λ-ball is much greater than the ambient density ρ. For larger normal volumes there is less compression. This can be compared to Bernoulli bond percolation on the square lattice in two dimensions, where an analog of this compression is known not to occur.

Mathematics Subject Classifications (1991): 60K35, 82B43

I Introduction

Traditionally the study of percolation has focused on lattice models. But continuous models, in which the set of sites is a random point process, are more natural for many applications, particularly those arising in statistics and in the study of impurities in materials, as examined by Hall [Ha2], Men'shikov et al. [MMS], Given and Stell [GS1, GS2], Stell and Xu [SX] and the references therein. Other recent works on continuous models are by Penrose [Pe] on the cluster size distribution, and by Roy [Ro] and Zuev and Sidorenko [ZS] on equality of critical points.

Most interesting, perhaps, are results for continuous models which are not direct extensions of known results for lattice models. In [ACC] it was shown that for Bernoulli bond percolation on the square lattice in two dimensions, for each

* Research supported by NSF grant number DMS-9006395

supercritical probability there is a characteristic shape W with the following property: for large k, conditionally on the cluster of the origin containing k sites, with high probability the shape of the cluster approximates a multiple $ck^{1/2}W$ of this characteristic shape. Further, the constant c is such that the density of the k-cluster in the region $ck^{1/2}W$ is the same as the ambient density of the infinite cluster in all of $\mathbb{Z}^2$. That is, the k-cluster looks like a broken-off piece of the infinite cluster – it is not any more tightly compressed. In contrast, we will see for our continuous model that an analog of such compression does sometimes occur.

It should be mentioned that the correspondence underlying this analogy is not an exact one – a different analog for lattice percolation, closer than that considered in [ACC], to the questions we consider here is possible, and might well exhibit answers similar to what we will obtain here for the continuous model. But we will not consider that alternate analog, outside of a brief description following Theorem 2.4 below, because it is not as natural a question for lattice models.

Two main types of continuous models have been considered. In the first, the *lily pad* model, independent regions of random size and/or shape are placed with center at each site of a (usually Poisson) point process; the clusters are then the connected components of the subset of $\mathbb{R}^d$ that is covered by these regions. This model is considered in [Ha1, Ha2, MMS, Ro, ZS], and [Gr, Sect. 10.5]. In the second type of model, there is a nonincreasing function $f: (0, \infty) \to [0, 1]$ with $f(s) \to 0$ as $s \to \infty$, and there is a bond between every pair of sites of the point process; these bonds are independently open with probability given by the function f evaluated at the distance between the two endpoints. We will call this the *random connection model*; it is considered in [GS1, GS2, SX] and [Pe].

Here we will consider what is perhaps the simplest continuous percolation model, the *Poisson blob* model, which lies in the intersection of the above two types. Our point process will be a Poisson process X in $\mathbb{R}^d$ ($d \geqq 2$) with intensity ρ, with a point added at 0; thus $X \cup \{0\}$ is a Poisson process "as viewed from one of its sites." (We use X to denote both the random counting measure and the corresponding set of sites.) For a fixed $r > 0$, each pair of sites is then connected by an open bond if and only if the sites are separated by distance $\leqq r$. This is of course equivalent to the lily pad model with nonrandom discs of radius $r/2$, or to the random connection model with $f = 1_{(0,r]}$. The cluster C of the origin is defined to be the set of sites which are connected to 0 by a path of open bonds. We wish to examine the distribution of the cardinality $|C|$ when ρ is large, and determine what a typical configuration looks like given $|C| = k$.

In a forthcoming paper we will examine similar questions for the random connection model with general smooth f. It will be shown that this smoothness makes the qualitative features of finite clusters quite different from the present case with discontinuous $f = 1_{(0,r]}$.

We will use P_ρ to denote probability when the intensity is ρ. In both the lily pad and random connection models, it is known [ZS, Pe] that under mild hypotheses satisfied for the Poisson blob model, there is a critical intensity $0 < \rho_0 < \infty$ such that $P_\rho(|C| = \infty)$ is positive for $\rho > \rho_0$ and zero for $\rho < \rho_0$.

II Statement of results

Penrose [Pe] showed that for the random connection model with finite range,

$$\lim_{\rho \to \infty} P_\rho(|C| = 1)/P_\rho(|C| < \infty) = 1 .$$

Since $\lim_{\rho\to\infty} P_\rho(|C| = \infty) = 1$, this is a statement about rare events. What we will obtain in our special case is much more detailed information about the probabilities of the rare events $[|C| = k + 1]$. (The " +1" here takes account of the origin.) Given a subset A of $\mathbb{R}^d$ and $s > 0$, let

$$A^s := \{x \in \mathbb{R}^d : d(x, A) \leqq s\}$$
$$A_s := \{x \in \mathbb{R}^d : d(x, A^C) \geqq s\},$$

where $d(\cdot, \cdot)$ denotes Euclidean distance. For j-dimensional sets A in $\mathbb{R}^d$, $j \leqq d$, let $|A|$ denote the j-dimensional volume; recall that $|A|$ denotes cardinality for countable sets A. When confusion is possible we use $\mathrm{vol}(\cdot)$ to denote d-dimensional volume. Let U denote the unit ball of $\mathbb{R}^d$, and $\pi_d := |U|$. What makes the event $[|C| = k + 1]$ rare for large ρ is that the region C^r must be empty except for the k sites of X in C; stated another way, C must be surrounded by a skin of thickness at least r which is completely empty of sites. One way for this to occur is that, for some $\lambda > 0$, there be k sites of X in the ball λU, and 0 sites of X in the annular region $(\lambda + r)U \backslash \lambda U$. Thus define

$$\begin{aligned} q(\rho, k) &:= \sup_{\lambda > 0} P_\rho[X(\lambda U) = k] P_\rho[X((\lambda + r)U \backslash \lambda U) = 0] \\ &= \sup_{\lambda > 0} \exp(-\rho\pi_d(\lambda + r)^d)(\rho\pi_d\lambda^d)^k/k!\,. \end{aligned}$$

Let λ_α denote the positive solution of the equation

$$\lambda(\lambda + r)^{d-1} = \alpha/\pi_d; \tag{2.1}$$

it is easily checked that the above sup occurs at $\lambda = \lambda_{k/\rho}$.

If we ignored the possibility that the k sites of X in λU are not all connected, which is unlikely for large ρ and not-too-large λ, we would obtain a seemingly crude lower bound:

$$P_\rho(|C| = k + 1) \geqq q(\rho, k)\,. \tag{2.2}$$

(This would be exactly true if we restricted the sup defining q to $0 < \lambda < r$.) But in fact, the lower bound (2.2) is not so crude at all, as our main result will show. When $|C| = k + 1$, C will be shown to approximate the ball $\lambda_{k/\rho} U$. This is analogous to phenomena observed in the study of "density of states," where in systems at low temperature most of the probability is concentrated on the small subcollection of states which have some particular geometry; see [LGP].

Let C_H denote the convex hull of C. There are two natural smoothings of C to use in comparing C to a ball or other region of $\mathbb{R}^d$: C_H and the set $(C^r)_r$. Note $C \subset (C^r)_r \subset C_H$. For C small relative to r, $(C^r)_r$ can be thought of as C_H with dents in it. It is easily checked that any convergence to a convex set which we prove for the shape of $(C^r)_r$ is also valid for C_H, provided C stays bounded. Define

$$d_m(A, B) := \inf_x \mathrm{vol}((A + x) \triangle B), \quad A, B \subset \mathbb{R}^d\,.$$

Theorem 2.1 *For the Poisson blob model,*

$$\log P_\rho(|C| = k)/\log q(\rho, k) \to 1 \quad \text{as } \rho \to \infty, \text{ uniformly in } k \geqq 0\,. \tag{2.3}$$

Further, conditionally on $|C| = k + 1$, C is approximately a ball of radius $\lambda_{k/\rho}$, in the sense that

$$d_m((\lambda_{k/\rho}^{-1}(C^r)_r), U) \to 0 \text{ in probability as } \rho \to \infty \text{ and } k \to \infty\,. \tag{2.4}$$

Theorem 2.1 is most easily understood by decomposing it into three different cases: $k/\rho \to 0$, k/ρ bounded away from 0 and ∞, and $k/\rho \to \infty$. The corresponding optimal radius $\lambda_{k/\rho}$ then also approaches 0, stays bounded, or approaches ∞, respectively. The first case has two subcases: k fixed and $k \to \infty$. In the cases $k/\rho \to 0$ and k/ρ bounded we will be able to strengthen (2.4), as follows. Let $H(A, B)$ denote the Hausdorff distance between subsets A and B of $\mathbb{R}^d$, and let d_H be the translated Hausdorff distance given by

$$d_H(A, B) := \inf_{x \in \mathbb{R}^d} H(A, x + B)$$

where $x + B$ is the translation of B by x. The statement

$$d_H(\lambda_{k/\rho}^{-1}(C^r)_r, U) \to 0 \quad \text{in probability} \tag{2.5}$$

is not in general equivalent to (2.4). This is because (2.4) allows for the possibility for example that $(C^r)_r$ is shaped like a ball with a long thin spike attached, while (2.5) allows $(C^r)_r$ to be a "spherical sponge", i.e. ball-shaped with many small holes inside. When $\lambda_{k/\rho}$ stays bounded, however, (2.4) and (2.5) are equivalent. This fact involves the special nature of $(C^r)_r$; specifically, no two components of $(C^r)_r$ can be separated by distance greater than r, and any point not in $(C^r)_r$ is part of an r-ball which doesn't meet $(C^r)_r$.

We define the *relative density* θ of the cluster C to be $|C|/\rho \operatorname{vol}(C_H)$, which should be thought of as the ratio of the density $|C|/\operatorname{vol}(C_H)$ of the cluster C to the ambient density ρ.

Theorem 2.2 *Let $k \geqq 0$ be fixed and $\rho \to \infty$. Then*

$$\text{the optimal radius } \lambda_{k/\rho} \sim (\pi_d r^{d-1})^{-1} k/\rho \to 0\,; \tag{2.6}$$

$$P_\rho(|C| = k + 1) = \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + O(1)])\,; \tag{2.7}$$

and for $k \geqq 1$, conditionally on $[|C| = k + 1]$,

$$\lambda_{k/\rho}^{-1} \operatorname{diam}(C) \text{ is bounded away from 0 and } \infty \text{ in probability;} \tag{2.8}$$

$$\text{the relative density } \theta \to \infty \text{ in probability}\,. \tag{2.9}$$

We call the phenomenon in (2.9) *compression.*

For $k = 0$ the factor $k \log(\rho/k)$ in (2.7) should be interpreted as 0, and the $O(1)$ is unnecessary.

Theorem 2.3 *Suppose $k \to \infty$ and $\rho \to \infty$ with $k/\rho \to 0$. Then*

$$\text{the optimal radius } \lambda_{k/\rho} \sim (\pi_d r^{d-1})^{-1} k/\rho \to 0\,; \tag{2.10}$$

$$\begin{aligned} &P_\rho(|C| = k + 1) \\ &= \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + (d-1)k\log(e\pi_d r^d) + o(k)])\,; \end{aligned} \tag{2.11}$$

and conditionally on $[|C| = k + 1]$,

$$d_H(\lambda_{k/\rho}^{-1}(C^r)_r, U) \to 0 \quad \text{and } d_m(\lambda_{k/\rho}^{-1}(C^r)_r, U) \to 0 \quad \text{in probability}\,; \tag{2.12}$$

$$\text{the relative density } \theta \to \infty \text{ in probability}\,. \tag{2.13}$$

Note the decay in (2.11) is locally roughly exponential in k, with the local rate shrinking from a very large value as k grows.

The ratio k/ρ represents the amount of volume which typically contains about k points when the intensity of the Poisson process is ρ. We therefore call k/ρ the *normal volume*. (2.9) and (2.13) show that clusters with small normal volume are typically highly compressed.

Theorem 2.4 *Let $0 < m < M < \infty$ and $\alpha = k/\rho$. Then as $\rho \to \infty$, uniformly in those k for which $m \leqq \alpha \leqq M$,*

$$\text{the optimal radius } \lambda_\alpha \text{ is bounded away from } 0 \text{ and } \infty; \tag{2.14}$$

$$P_\rho(|C| = k + 1) = \exp(-[r/\lambda_\alpha + (d-1)\log(1 + r/\lambda_\alpha)]k(1 + o(1))) ; \tag{2.15}$$

and conditionally on $[|C| = k + 1]$,

$$d_H(\lambda_\alpha^{-1}(C^r)_r, U) \to 0 \quad \text{and } d_m(\lambda_{k/\rho}^{-1}(C^r)_r, U) \to 0 \text{ in probability} ; \tag{2.16}$$

$$\text{the relative density satisfies } |\theta - (1 + r/\lambda_\alpha)^{d-1}| \to 0 \text{ in probability} . \tag{2.17}$$

We call the phenomenon in (2.17) *partial compression*; the cluster is denser than the ambient density ρ but only by a finite factor. Thus clusters with moderate normal volume are typically partially compressed.

Note that, as in (2.11), the decay in (2.15) is roughly exponential in k.

Taking the low-temperature limit $\rho \to \infty$ is of course equivalent to keeping ρ fixed and letting $r \to \infty$. For lattice percolation, the analog of $\rho \to \infty$ would be to have the bond density $p \to 1$, whereas $r \to \infty$ corresponds to the completely different phenomenon of extending the range, having a possibly open bond between every pair of lattice sites separated by distance at most r. The results of [ACC] discussed above in the introduction considered p fixed, $r = 1$, and $k \to \infty$, where it was found that no compression occurred. A better analog, for lattice percolation, to our Theorems 2.3 and 2.4 might be to consider p fixed and $k, r \to \infty$; k/ρ corresponds to $k/\pi_d r^d$ wherever limits of this ratio are taken. We do not know if any compression would then occur. Our closest analog to the results in [ACC] is Theorem 2.5 below, in which no compression occurs.

In [ACC] only dimension $d = 2$ was considered, but the results covered all supercritical probabilities. Here the dimension $d \geqq 2$ is arbitrary but only very high intensities (low temperatures) are considered; this allows the use of Peierls-type arguments in which one sums over all possible contours which could bound a lattice approximation to the cluster.

Theorem 2.5 *Let $k \to \infty$ and $\rho \to \infty$ with $k/\rho \to \infty$. Then*

$$\text{the optimal radius } \lambda_{k/\rho} \sim (k/\rho\pi_d)^{1/d} \to \infty ; \tag{2.18}$$

$$P_\rho(|C| = k + 1) = \exp(-d\pi_d^{1/d} rk(\rho/k)^{1/d}(1 + o(1))) ; \tag{2.19}$$

and conditionally on $[|C| = k + 1]$,

$$d_m((\lambda_{k/\rho}^{-1}(C^r)_r), U) \to 0 \text{ in probability} ; \tag{2.20}$$

$$\text{the relative density } \theta \to 1 \text{ in probability} . \tag{2.21}$$

Thus clusters with large normal volume are typically not compressed. The decay in (2.19), in contrast to (2.11) and (2.15), is subexponential in k. The power $k^{(d-1)/d}$ in the exponent corresponds to that found by Kesten and Zhang [KZ] for Bernoulli bond percolation on the lattice.

In dimension $d = 2$ one can substitute d_H for d_m in (2.20); see the remarks after (2.5) and see Remark 4.8 below.

By rescaling one can obtain results in which r varies, in addition to k and ρ. Compression, partial compression, and no compression correspond to $k/\rho r^d$ approaching 0, staying bounded away from 0 and ∞, and approaching ∞ respectively. Thus for example if $k \to \infty$, $\rho \to \infty$, and $r \to 0$ with $k/\rho \to \alpha \in (0, \infty)$, then conditionally on $|C| = k + 1$, with high probability C will approximate a ball of radius $(\alpha/\pi_d)^{1/d}$ surrounded by a thin shell of empty space; there is no compression. Allowing $r \to 0$ in this way makes for a more natural continuum limit.

When $k \to \infty$ and k/ρ stays bounded, one can say more than just that C approximates a ball; the next theorem shows that the $k + 1$ points approximate a uniform distribution over a randomly translated ball.

Theorem 2.6 *Let $k \to \infty$ and $\rho \to \infty$ with k/ρ bounded. On the event $[|C| = k + 1]$, let Y_k denote the centroid of* Co(C), *let ν denote the uniform distribution on U and define the empirical measure*

$$\mu_k := (k + 1)^{-1} \sum_{x \in C} \delta_{\lambda_{k/\rho}^{-1} x} .$$

Then

$$\mathscr{L}(\mu_k(\cdot + Y_k) \,|\, |C| = k + 1) \Rightarrow \nu \textit{ in probability} . \tag{2.22}$$

The weak convergence in probability in (2.22) means that the sup-norm distance between the corresponding d.f.'s approaches 0 in probability.

Theorem 2.1 is a routine consequence of Theorems 2.2–2.5, so we will only prove the latter results and Theorem 2.6.

III Proofs when the normal volume is small

In this section we will prove Theorems 2.2 and 2.3. Throughout, $c_1, c_2, \ldots$ represent positive constants which do not depend on k or ρ. The case $k = 0$ is trivial, so we henceforth assume $k \geqq 1$.

We begin with an easy lower bound. As $k/\rho \to 0$ we have $\lambda_{k/\rho} \sim c_1 k/\rho$ where $c_1 := (\pi_d r^{d-1})^{-1}$. When $c_1 k/\rho < r$ all sites in $(c_1 k/\rho)U$ are connected; therefore using Stirling's formula,

$$\begin{aligned} P_\rho[|C| = k + 1] &\geqq P_\rho[X((c_1 k/\rho)U) = k] P_\rho[X((c_1 k/\rho + r)U \backslash (c_1 k/\rho)U) = 0] \\ &\geqq (2\pi k)^{-1/2} (e\pi_d c_1^d (k/\rho)^{d-1})^k \exp(-[\rho\pi_d r^d + d\pi_d r^{d-1} c_1 k + c_2 k^2/\rho]) \\ &\geqq \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + (d-1)k\log(e\pi_d r^d) + c_2 k^2/\rho + c_3 \log k]) . \end{aligned} \tag{3.1}$$

Note that the right side of (3.1) includes both a term $(d-1)k\log(\rho/k)$ arising from the need to have k points in a very small ball, and a term $\rho\pi_d r^d$ arising from the need to have a shell of empty space surrounding this ball. The optimal radius $\lambda_{k/\rho}$ represents an optimal tradeoff between the probabilities of these two events; at the optimal radius, both events are rare. This contrasts with the situation for large k/ρ, examined in Sect. IV.

The idea of the proof of (2.8) and (2.12) is to show that all possible ways for the event $[|C| = k + 1]$ to occur, other than those in (2.8) and (2.12), together have probability which is much less than the right side of (3.1). Our first four lemmas deal with the possibility of spatially very large clusters.

Lemma 3.1 *Let* $\mu > 1$ *and define*

$$\psi_\mu(y) := \pi_{d-1} r^{d-1} y/4 - \log(e\pi_d \mu^d y^d), \quad y > 0 .$$

There exists $c_4 = c_4(\mu, r, d) > 0$ *such that if* $yk/\rho < c_4$ *then*

$$P_\rho(|C| = k+1, yk/\rho < \operatorname{diam}(C) \leqq \mu yk/\rho)$$
$$\leqq \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + k\psi_\mu(y)]) .$$

Note that $\psi_\mu(y) \to \infty$ as $y \to \infty$ and as $y \to 0$.

Proof. Let C_* be the set of sites of $X \cup \{0\}$ in the ball $(\mu yk/\rho)U$. Define events

$$A := [X((\mu yk/\rho)U) = k]$$
$$B := [yk/\rho < \operatorname{diam}(C_*) \leqq \mu yk/\rho] .$$

Then, provided $c_4 < r/\mu$ so that $\mu yk/\rho < r$, we have

$$P_\rho(|C| = k+1, yk/\rho < \operatorname{diam}(C) \leqq \mu yk/\rho) \tag{3.2}$$
$$= E(P_\rho(|C| = k+1, yk/\rho < \operatorname{diam}(C) \leqq \mu yk/\rho | C_*))$$
$$= E(P_\rho(X(C_*^r \backslash (\mu yk/\rho)U) = 0 | C_*) 1_A 1_B) .$$

Let x and z be the endpoints of a diameter of C_*, let v be any vector perpendicular to $x - z$, and let w be the point of C_* which maximizes the inner product with v. Let H_x and H_z be the hyperplanes through x and z perpendicular to $x - z$, and let S be the slab between H_x and H_z. Let H_w be the hyperplane through w perpendicular to v; there then exist halfspaces H_x^+, H_z^+, and H_w^+ bounded by H_x, H_z, and H_w which have no points of C_* in their interiors. Then on the event B,

$$|C_*^r| \geqq |H_x^+ \cap (x + rU)| + |H_z^+ \cap (z + rU)| + |S \cap H_w^+ \cap (w + rU)| \tag{3.3}$$
$$\geqq \pi_d r^d + \pi_{d-1} r^{d-1} yk/4\rho$$

provided c_4 is small enough. Thus, using Stirling's formula, since X inside and outside this ball are independent, the right side of (3.2) is bounded above by

$$\exp(-\rho[\pi_d r^d + \pi_{d-1} r^{d-1} yk/4\rho - |(\mu yk/\rho)U|]) P_\rho(A)$$
$$= \exp(-\rho[\pi_d r^d + \pi_{d-1} r^{d-1} yk/4\rho])(\rho|(\mu yk/\rho)U|)^k/k!$$
$$\leqq \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + k\psi_\mu(y)]). \quad \square$$

We need to consider some lattice approximations. Define

$$\mathscr{G}(\delta) := \{x + [-\delta/2, \delta/2]^d : x \in \delta\mathbb{Z}^d\} .$$

For $A \subset \mathbb{R}^d$ let

$$Q^\delta(A) := \cup\{G \in \mathscr{G}(\delta) : G \cap A \neq \phi\}$$
$$Q_\delta(A) := \cup\{G \in \mathscr{G}(\delta) : G \subset A\}$$

and define the *outer boundary* of A to be

$$\partial_0 A := \{x \in \partial A : x \text{ connected to } \infty \text{ in } A^c\} .$$

A union H of cubes from some $\mathscr{G}(\delta)$ is *strongly connected* if every $x, y \in H$ can be connected by a path in H which does not pass through any intersection of dimension $d-2$ or less of two cubes in $\mathscr{G}(\delta)$. Define

$$\mathscr{H}(\delta) := \{H \subset \mathbb{R}^d : H \text{ is a strongly connected finite union of cubes in } \mathscr{G}(\delta)\} .$$

A *δ-plaquette* is a face of any cube in $\mathscr{G}(\delta)$ with sides parallel to the axes. A *δ-contour* is a set of δ-plaquettes which is of the form $\partial_0 H$ for some $H \in \mathscr{H}(\delta)$; let

$$\mathscr{C}_n(\delta) := \{S : S \text{ is a } \delta\text{-contour of } n \text{ plaquettes enclosing } 0\} .$$

It is well-known that there exists a constant a_d such that

$$|\mathscr{C}_n(\delta)| \leqq a_d^n . \tag{3.4}$$

A union S of δ-plaquettes is *strongly connected* if every $x, y \in S$ can be connected by a path which does not pass through any intersection of dimension $d-3$ or less of two δ-plaquettes. Given a δ-contour S, let $I(S)$ denote the closed region enclosed by S and let

$$D(S) := \cup \{G \in \mathscr{G}(\delta) : G \subset I(S), G \text{ has a face in } S\} .$$

Then

$$S \in \mathscr{C}_n(\delta) \text{ implies } |D(S)| \geqq \delta^d n/2d . \tag{3.5}$$

If $\delta <$ some c_6 then the inner lattice approximation $Q_\delta(C^r)$ is always strongly connected, and its outer layer of cubes satisfies

$$X(D(\partial_0 Q_\delta(C^r))) = 0 \text{ when } \delta < c_6 . \tag{3.6}$$

Lemma 3.2 *For every $\beta < \infty$ there exist $R < \infty$ and $\rho_1 < \infty$ such that for $\rho \geqq \rho_1$,*

$$P_\rho[|C| < \infty, \operatorname{diam}(C) \geqq R] \leqq e^{-\rho\beta} .$$

Proof. Fix $\delta < c_6$. The contour $\partial_0 Q_\delta(C^r)$ consists of at least R/δ plaquettes, so by (3.4), (3.5), and (3.6),

$$\begin{aligned} P_\rho[|C| < \infty, \operatorname{diam}(C) \geqq R] &\leqq \sum_{n \geqq R/\delta} \sum_{S \in \mathscr{C}_n(\delta)} P_\rho(X(D(S)) = 0) \\ &\leqq \sum_{n \geqq R/\delta} a_d^n \exp(-\rho\delta^d n/2d) \end{aligned}$$

and the lemma follows easily. □

Lemma 3.3 *For every $0 < \zeta < r$, $R < \infty$ there exist $c_7 > 0$ and $\rho_2 < \infty$ such that for $\rho \geqq \rho_2$,*

$$P_\rho[\zeta \leqq \operatorname{diam}(C) \leqq R] \leqq \exp(-\rho[\pi_d r^d + c_7]) .$$

Proof. As in the proof of Lemma 3.1 (cf. (3.3) and the preceding definitions), when $\zeta \leqq \operatorname{diam}(C) \leqq R$ there exists a region $A \subset [-(R+r), R+r]^d$ such that (i) A consists of two half-balls of radius r and the intersection of a third r-ball with a slab of thickness ζ or more, and (ii) $X(A) = 0$. Therefore, for some c_8, c_9 and c_{10}, we have $|A| \geqq \pi_d r^d + c_8\zeta$, and for $\delta < c_9$ we have $|Q_\delta(A)| \geqq \pi_d r^d + c_{10}\zeta$. The lemma now follows readily from the fact there are only finitely many possible sets $Q_\delta(A)$. □

Lemma 3.4 *For every $\beta < \infty$ there exist $c_{11} < \infty$, $\zeta > 0$ and $\rho_3 < \infty$ such that when $\rho \geqq \rho_3$,*

$$P_\rho[|C| = k+1, c_{11}k/\rho \leqq \operatorname{diam}(C) \leqq \zeta] \leqq \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + \beta k]) .$$

Proof. Fix $\mu > 1$ and $\zeta \leqq c_4$ (of Lemma 3.1). Fix c_{11} such that $\psi_\mu(c_{11}) > \beta + \log 2$. We may assume that $c_{11}k/\rho \leqq \zeta$. Let N be the largest integer such that $c_{11}\mu^N k/\rho \leqq \zeta$. Then provided c_{11} is sufficiently (depending on μ) large, by Lemma 3.1,

$$\begin{aligned} P_\rho[|C| = k+1, c_{11}k/\rho \leqq \operatorname{diam}(C) \leqq \zeta] &\leqq \sum_{j=0}^{N} P_\rho[|C| = k+1, c_{11}\mu^j k/\rho \leqq \operatorname{diam}(C) \leqq c_{11}\mu^{j+1}k/\rho] \\ &\leqq \sum_{j=0}^{N} \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + k\psi_\mu(c_{11}\mu^j)]) \\ &\leqq 2\exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + k\psi_\mu(c_{11})]) , \end{aligned}$$

and the lemma follows. □

The next lemma covering very small clusters follows from essentially the same argument as in Lemma 3.4.

Lemma 3.5 *For every $\beta < \infty$ there exist $c_{12} > 0$, $\zeta > 0$ and $\rho_4 < \infty$ such that when $\rho \geqq \rho_4$ and $c_{12}k/\rho \leqq \zeta$,*

$$P_\rho[|C| = k+1, \operatorname{diam}(C) \leqq c_{12}k/\rho] \leqq \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + \beta k]) .$$

Proof of Theorem 2.2 From Lemma 3.1 with $\mu = c_{11}/c_{12}$ we have

$$P_\rho[|C| = k+1, c_{12}k/\rho \leqq \operatorname{diam}(C) \leqq c_{11}k/\rho] \leqq \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + k\psi_\mu(c_{12})]) .$$

With (3.1) this and Lemmas 3.2–3.5, with β fixed but arbitrarily large, prove (2.7) and (2.8). (2.6) is trivial, and (2.9) is an easy consequence of (2.6) and (2.8). □

To prove Theorem 2.3 we will need more precise results than Lemma 3.1 to handle $c_{12}k/\rho \leqq \operatorname{diam}(C) \leqq c_{11}k/\rho$. The following result is due to Minkowski; see e.g. [Bu, Chap. 2].

Lemma 3.6 *Let A and B be nonempty convex sets in $\mathbb{R}^d$. There exist positive "mixed volumes" $V_i(A, B)$, $0 \leqq i \leqq d$, such that*

$$|xA + yB| = \sum_{i=0}^{d} \binom{d}{i} V_i(A, B)x^i y^{d-i} \quad \textit{for all } x, y \geqq 0 .$$

Further, $V_0(A, B) = |B|$, $V_d(A, B) = |A|$, and $dV_{d-1}(A, U) = |\partial A|$. For fixed $1 \leqq i \leqq d-1$ and $a > 0$, *as A varies with $|A| = a|U|$ held fixed, $V_i(A, U)$ has a minimum, unique up to translations, at $A = a^{1/d}U$.*

Remark 3.7 The convex subsets of a fixed bounded set in $\mathbb{R}^d$ form a compact set with respect to the topology of the metric d_H; therefore the infimum of

$\chi(A) = V_i(A, U)$, as A varies over any bounded collection which has $|A| = a|U|$ and is d_H-bounded away from $a^{1/d}U$, is strictly greater than $\chi(a^{1/d}U)$.

The next lemma will cover the possibility that the cluster C occupies a volume significantly different from the optimal volume.

Lemma 3.8 *For every* $\varepsilon > 0$ *and* $0 < t < T$ *there is a* $\zeta_0 > 0$ *such that for every* $0 < \zeta < \zeta_0$ *there exist* $k_1 \geqq 1$ *and* $c_i > 0$ $(i = 13, 14)$ *such that if* $k \geqq k_1$ *and* $k/\rho \leqq c_{13}$ *then for* $\delta = \zeta k/\rho$,

$$P_\rho[|C| = k + 1, tk/\rho \leqq \operatorname{diam}(C) \leqq Tk/\rho, |(|Q^\delta(C)|/\pi_d(c_1 k/\rho)^d)^{1/d} - 1| > \varepsilon]$$
$$\leqq \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + (d-1)k\log(e\pi_d r^d) + c_{14}k]) .$$

Proof. Leaving aside the specification of ζ_0 for the moment, let $D \ni 0$ be an element of $\mathscr{H}(\delta)$ with $tk/\rho \leqq \operatorname{diam}(D) \leqq (T + \zeta d^{1/2})k/\rho$ and $|(|D|/\pi_d(c_1 k/\rho)^d)^{1/d} - 1| \geqq \varepsilon$. Note the number of such D is bounded uniformly in k and ρ, and $\operatorname{diam}(C) \leqq Tk/\rho$ ensures $Q^\delta(C)$ is such a set D. Also $(T + \zeta d^{1/2})k/\rho < r$ provided c_{13} is small enough, so that all sites of X in D are connected. As in the proof of Lemma 3.1 it follows from the independence of X inside and outside of D that

$$\begin{aligned} &P_\rho[|C| = k+1, Q^\delta(C) = D] \\ &\leqq P_\rho[X(D) = k] \sup_{A\in\mathscr{U}} P_\rho[X(A^r\backslash D) = 0] \\ &= (k!)^{-1}(\rho|D|)^k \exp(-\rho \inf_{A\in\mathscr{U}} |A^r|) \end{aligned} \tag{3.7}$$

where $\mathscr{U}$ denotes the collection of all finite sets $A \subset D$ with $|A| = k + 1$, $0 \in A$, $\operatorname{diam}(A) \leqq Tk/\rho$ and with every δ-cube comprising D containing at least one point of A. Fix $A \in \mathscr{U}$; we need a lower bound for $|A^r|$. Note we cannot apply Lemma 3.6 to $|A^r| = |A + rU|$ because A is not convex.

Suppose $x \in A_H\backslash(A^r)_r$. Recalling that $(A^r)_r$ can be thought of as A_H with dents in it, we wish to show these dents are not very deep, i.e. x is close to ∂A_H. There exists $y \in \partial A^r$ with $d(x, y) < r$. Let a be a point of A which minimizes $(x - y)\cdot a$. There exists z on the line from x to y such that $z\cdot a = (x - y)\cdot a$; note y and A_H are on opposite sides of the hyperplane through a and z perpendicular to $x - y$. Let φ be the angle between $a - y$ and $x - y$; then

$$d(x, \partial A_H) \leqq d(x, z) \leqq r - d(y, z) = r - d(y, a)\cos\varphi \leqq r(1 - \cos\varphi) .$$

But for some $c_{16}(r, T)$, $\operatorname{diam}(A) \leqq Tk/\rho$ ensures that

$$r(1 - \cos\varphi) \leqq \alpha := c_{16}(k/\rho)^2$$

and it follows that

$$G := (A_H)_\alpha \subset (A^r)_r ,$$

and then that

$$A^r \supset G^r . \tag{3.8}$$

Since G is convex, we will be able to apply Lemma 3.6 to $|G + rU|$ once we know $|G|$.

We need to show that the G is not much smaller than D. We have

$$|D\backslash G| \leqq |D_H\backslash A_H| + |A_H\backslash(A_H)_\alpha| . \tag{3.9}$$

Now

$$A_H \subset D_H \subset (A_H)^{\delta d^{1/2}},$$

so for some c_{17} and c_{18},

$$|D_H \backslash A_H| \leqq c_{17}\delta d^{1/2}(\mathrm{diam}(A_H))^{d-1} \leqq c_{18}\zeta(k/\rho)^d \tag{3.10}$$

while

$$|A_H \backslash (A_H)_\alpha| \leqq c_{17}\alpha(\mathrm{diam}(A_H))^{d-1} \leqq c_{18}\zeta(k/\rho)^d . \tag{3.11}$$

Define τ by

$$|D| = \tau^d \pi_d (c_1 k/\rho)^d ,$$

so that the radii of the optimal ball and the ball of volume $|D|$ differ by a factor of τ, and

$$|\tau - 1| \geqq \varepsilon .$$

Suppose first that D is large enough so $\tau \geqq 1/4$. Given $\eta > 0$ to be specified later, from (3.9)–(3.11), if ζ_0 is small enough then

$$|D \backslash G| \leqq (1 - (1-\eta)^d)|D| \quad \text{so} \quad |G| \geqq (1-\eta)^d |D| = |((1-\eta)\tau c_1 k/\rho)U| .$$

Using (3.8) and Lemma 3.6, then,

$$|A^r| \geqq |G + rU| \geqq |((1-\eta)\tau c_1 k/\rho)U + rU| \geqq \pi_d r^d + d\pi_d r^{d-1}(1-\eta)\tau c_1 k/\rho .$$

Using (3.7), the definition of c_1, and Stirling's formula, we then obtain

$$\begin{aligned} &P_\rho[|C| = k+1, Q^\delta(C) = D] \\ &\leqq (e\rho|D|/k)^k \exp(-[\rho\pi_d r^d + d(1-\eta)\tau k]) \\ &= (e^d \tau^d e^{-d(1-\eta)\tau})^k \exp(-[\rho\pi_d r^d \\ &\quad + (d-1)k\log(\rho/k) + (d-1)k\log(e\pi_d r^d)]) . \end{aligned} \tag{3.12}$$

Note that the factor $(e^d\tau^d e^{-d(1-\eta)\tau})^k$ comes from the nonoptimal size $(\tau \neq 1)$ of D. The function $f(\tau) := e^d \tau^d e^{-d\tau}$ achieves a unique maximum at $f(1) = 1$, so there exists $c_{19}(\varepsilon, d) > 0$ small enough so

$$|\tau - 1| \geqq \varepsilon \quad \text{implies} \quad f(\tau) \leqq e^{-2c_{19}} .$$

We can then choose $\eta(d)$ so that

$$|\tau - 1| \geqq \varepsilon \quad \text{implies} \quad e^d \tau^d e^{-d(1-\eta)\tau} \leqq e^{-c_{19}} . \tag{3.13}$$

Under the alternate possibility that $\tau < 1/4$, the extra term $d\pi_d r^{d-1}(1-\eta)\tau c_1 k/\rho$ in the lower bound for $|A^r|$ is not needed, and the above inqualities remain valid with $\eta = 1$, provided we take c_{19} less than $\log(4/e)$. The lemma now follows from (3.12) and (3.13), since, as discussed above, the number of possible D is bounded. □

It remains to cover the possibility that C has near-optimal volume but a non-spherical shape.

Lemma 3.9 *For every $T, \gamma > 0$ there is a $\zeta_1 > 0$ such that for every $0 < \zeta < \zeta_1$ there exist $\varepsilon > 0$, $k_2 \geqq 1$ and $c_i > 0$ $(i = 20, 21)$ such that if $k \geqq k_2$ and $k/\rho \leqq c_{20}$*

then for $\delta = \zeta k/\rho$,

$$P_\rho[|C| = k+1, \operatorname{diam}(C) \leqq Tk/\rho, (1-\varepsilon)^d \leqq |Q^\delta(C)|/\pi_d(c_1 k/\rho)^d \leqq (1+\varepsilon)^d, \quad d_H(\lambda_{k/\rho}^{-1} C_H, U) > \gamma]$$
$$\leqq \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + (d-1)k\log(e\pi_d r^d) + c_{21}k]).$$

Proof. The proof is similar to that of Lemma 3.8, so we will continue with the notation of that proof. It is enough to prove the lemma with $\lambda_{k/\rho}^{-1}$ replaced by $\rho/c_1 k$, as the ratio can be made arbitrarily close to 1 by taking c_{20} small. Again we fix D and A, with the additional restriction that

$$d_H((\rho/c_1 k)A_H, U) > \gamma. \tag{3.14}$$

Let $G^* := (|U|/|G|)^{1/d}G$, so $|G^*| = |U|$. As in Lemma 3.8, and using Lemma 3.6, we have

$$|A^r| \geqq |G + rU| \geqq \pi_d r^d + dV_1(G^*, U)(|G|/|U|)^{1/d} r^{d-1}.$$

For some $\varepsilon > 0$ to be specified later, let us first consider the case of $|G| \geqq |((1+\varepsilon)\tau c_1 k/\rho)U|$. Here analogously to (3.12),

$$\begin{aligned} P_\rho[|C| = k+1, Q^\delta(C) = D] &\leqq (e^d\tau^d e^{-d(1+\varepsilon)\tau})^k \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) \\ &\quad + (d-1)k\log(e\pi_d r^d)]) \qquad (3.15)\\ &\leqq (e^{-d\varepsilon\tau})^k \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + (d-1)k\log(e\pi_d r^d)]). \end{aligned}$$

The other possibility is

$$|G| < |((1+\varepsilon)\tau c_1 k/\rho)U|. \tag{3.16}$$

If ζ_1 is small enough then as in Lemma 3.8,

$$|G| \geqq (1-\varepsilon)^d |D| \geqq |((1-\varepsilon)\tau c_1 k/\rho)U|. \tag{3.17}$$

Also,

$$\begin{aligned} &d_H((\rho/c_1 k)A_H, G^*) \qquad (3.18)\\ &\leqq d_H((\rho/c_1 k)A_H, (\rho/c_1 k)(A_H)_\alpha) + d_H((\rho/c_1 k)G, (|U|/|G|)^{1/d}G). \end{aligned}$$

Since $|D| \geqq (1-\varepsilon)\pi_d(c_1 k/\rho)^d$ and $\operatorname{diam}(D) \leqq (T + \zeta d^{1/2})k/\rho$, D and also A_H cannot be too flat, and therefore provided ζ_1 is small, for some c_{23},

$$d_H((\rho/c_1 k)A_H, (\rho/c_1 k)(A_H)_\alpha) \leqq c_{23}(\rho/c_1 k)\alpha \leqq \gamma/4. \tag{3.19}$$

Meanwhile,

$$d_H((\rho/c_1 k)G, (|U|/|G|)^{1/d}G) \leqq |(\rho/c_1 k) - (|U|/|G|)^{1/d}| \operatorname{diam}(G). \tag{3.20}$$

Since by (3.16) and (3.17), using $|\tau - 1| < \varepsilon$,

$$(1+\varepsilon)^2(\rho/c_1 k) \geqq (|U|/|G|)^{1/d} \geqq (1-\varepsilon)^2(\rho/c_1 k),$$

we have from (3.20)

$$d_H((\rho/c_1 k)G, (|U|/|G|)^{1/d}G) \leqq (2\varepsilon + \varepsilon^2)(\rho/c_1 k)c_{11}k/\rho \leqq \gamma/4$$

if ε is sufficiently small. This with (3.14), (3.18) and (3.19) gives

$$d_H(G^*, U) > \gamma/2 . \tag{3.21}$$

By Remark 3.7 there is a constant $\sigma(\gamma, d) > 0$ such that (3.21) implies

$$V_1(G^*, U) \geqq V_1(U, U) + \sigma = \pi_d + \sigma .$$

Therefore analogously to (3.12),

$$P_\rho[|C| = k + 1, Q^\delta(C) = D] \tag{3.22}$$
$$\leqq (e^d \tau^d e^{-d(1-\varepsilon)\tau(1+\sigma/\pi_d)})^k \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + (d-1)k\log(e\pi_d r^d)]) .$$

If ε is small enough then uniformly in τ,

$$e^d \tau^d e^{-d(1-\varepsilon)\tau(1+\sigma/\pi_d)} \leqq e^{-c_{24}}$$

for some $c_{24} > 0$. (Note this uses the positiveness of the nonsphericality term σ.) The lemma now follows from (3.15) and (3.22), since the number of possible D is bounded. □

To show that the lower bound in (3.1) really gives the right rate, we will use an upper bound for C of near-optimal volume, without restriction on the shape.

Lemma 3.10 *For every $T, v > 0$ there is a $\zeta_2 > 0$ such that for every $0 < \zeta < \zeta_2$ there exist $\varepsilon > 0$, $k_3 \geqq 1$ and $c_{25} > 0$ such that if $k \geqq k_3$ and $k/\rho \leqq c_{25}$ then for $\delta = \zeta k/\rho$,*

$$P_\rho[|C| = k + 1, \operatorname{diam}(C) \leqq Tk/\rho, (1-\varepsilon)^d \leqq |Q^\delta(C)|/\pi_d(c_1 k/\rho)^d \leqq (1+\varepsilon)^d]$$
$$\leqq \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + (d-1)k\log(e\pi_d r^d) - vk]) .$$

Proof. The proof is again similar to that of Lemma 3.8. Fixing D as in that proof, if ζ_2 is small enough we have

$$|G| \geqq |((1-\varepsilon)\tau c_1 k/\rho)U|$$

and analogously to (3.12),

$$P_\rho[|C| = k + 1, Q^\delta(C) = D]$$
$$\leqq (e^d \tau^d e^{-d(1-\varepsilon)\tau})^k \exp(-[\rho\pi_d r^d + (d-1)k\log(\rho/k) + (d-1)k\log(e\pi_d r^d)]) .$$

If ε is small enough then since $|\tau - 1| < \varepsilon$,

$$e^d \tau^d e^{-d(1-\varepsilon)\tau} < e^{v/2}$$

and the lemma follows as in Lemma 3.8. □

Proof of Theorem 2.3 Fix $v, \gamma > 0$ and $\beta > \pi_d r^d$. Let $T = c_{11}$ and $t = c_{12}$ be as in Lemmas 3.4 and 3.5, R as in Lemma 3.2, and ζ as in Lemmas 3.4 and 3.5. Then let ε be as in Lemmas 3.9 and 3.10. (2.11) follows from (3.1) and Lemmas 3.2–3.5, 3.8, and 3.10. Let $\kappa := \min(c_{14}, c_{21})/2$; then from (3.1) and Lemmas 3.2–3.5, 3.8, and 3.9, if k and ρ are large and k/ρ is small,

$$P_\rho[d_{\mathrm{H}}(\lambda_{k/\rho}^{-1} C_H, U) > \gamma \,|\, |C| = k + 1]$$
$$\leqq 6e^{-2\kappa k}/\exp(-[c_2 k^2/\rho + c_3 \log k]) \leqq e^{-\kappa k}$$

and (2.12) for d_H, with $(C^r)_r$ replaced by C_H, follows. (2.12) for d_H as written then follows from the fact that, for $\alpha := c_{16}(k/\rho)^2$ as in the proof of Lemma 3.8, we have

$$(C_H)_\alpha \subset (C^r)_r \subset C_H$$

except possibly on the event $[\operatorname{diam}(C) \geqq c_{11} k/\rho]$, which has conditional probability decreasing exponentially in k by Lemma 3.4. Equivalence of convergence for d_H and d_m follows from the remarks after (2.5). (2.10) is trivial, and (2.13) follows easily from (2.10) and (2.12). □

IV Proofs when the normal volume is large

In this section we will prove Theorem 2.5. When k/ρ is large and $|C| = k+1$, the cluster C tends to be spread over a large region, resulting in two new features not present in Sect. III where k/ρ was smaller. First, when k/ρ was small, we approximated C^r by something close to $(C_H)^r$ (specifically by G^r) which enabled us to use Lemma 2.6 from convex geometry. In the present situation, C can be very irregularly shaped, and C^r is just C with a thin skin around it so C^r need not be anywhere close to convex. Second, a spread-out cluster C, or more precisely $(C^r)_r$, is likely to surround a number of separate small clusters which sit inside holes in C. For large λ the conditions $X(\lambda U) = k$, $X((\lambda + r)U \backslash \lambda U) = 0$ imply only that $|C| \leqq k+1$, not $|C| = k+1$. But for large ρ these conditions do imply that with high probability $|C|$ is a large fraction of k. Therefore instead of directly estimating $P_\rho[|C| = k+1]$ we will make some of our estimates at first for $P_\rho[(1-2\varepsilon)k \leqq |C| < \infty]$, with $\varepsilon > 0$ small.

In place of Lemma 3.6 we will use the following lemma, which is well-known when restricted to convex sets (see e.g. [Bo].) The second statement of the lemma is false in dimension $d \geqq 3$ for d_H in place of d_m, because of the possibility for example that B is shaped like a ball with a long arbitrarily thin spike attached.

Lemma 4.1 (See [Ta1, Ta2]) *For $a > 0$ the region $B = a^{1/d} U$ uniquely minimizes $|\partial B|$ among all (not necessarily connected) regions $B \subset \mathbb{R}^d$ with piecewise C^1 boundary and $|B| = |a^{1/d} U|$. Given $\varepsilon > 0$, there exists $\eta > 0$ such that for every polyhedron B with $|B| = |a^{1/d} U|$ and $d_m(B, a^{1/d} U) > \varepsilon$, we have $|\partial B| \geqq (1+\eta)|\partial(a^{1/d} U)|$.*

If A is finite or polyhedral and $t > 0$, then one can find a polyhedron B with $|B|$ and $|\partial B|$ arbitrarily close to $|A^t|$ and $|\partial A^t|$. Thus the second conclusion in Lemma 4.1 is also valid when B is such a set A^t. The lemma can then be used to estimate volumes $|A^t \backslash A|$ using the relationship

$$|A^t \backslash A| = \int_0^t |\partial A^s| \, ds \,. \tag{4.1}$$

In particular we have the following.

Lemma 4.2 *Given $\gamma > 0$, there exist $\eta > 0$ and $t_0 > 0$ such that for A with $|A| = |U|$, $d_m(A, U) > \gamma$, and ∂A^s piecewise C^1 for all $0 \leqq s \leqq t_0$, for $0 < t \leqq t_0$,*

$$|A^t \backslash A| \geqq (1+\eta)|U^t \backslash U| \,.$$

Proof. The difficulty here in applying (4.1) and Lemma 4.1 is that A^s could be spherical for arbitrarily small s, though A is not.

For $s > 0$ let b_s denote the radius of the multiple of U with volume $|A^s|$. Let $\alpha := \inf\{s > 0: |A^s| < |U^s|\}$. Now $|A^0| = |U^0|$ and by Lemma 4.1 for $0 < s < \alpha$ we have

$$\frac{d}{ds}|A^s| \geqq \frac{d}{ds}|U^s| .$$

It follows that $\alpha = \infty$ and thus

$$|A^s| = |b_s U| \geqq |U^s| \quad \text{for all } s . \tag{4.2}$$

If for some s we have $b_s > (1 + s)(1 + \varepsilon/8)$ then

$$|A^s| = |b_s U| > |(1 + \varepsilon/8)U^s|$$

and therefore

$$|\partial A^s| > |\partial U^s|(1 + \varepsilon/8) . \tag{4.3}$$

On the other hand, presuming $\varepsilon < 1$ and $s < t_0 := \varepsilon/16$, if $b_s \leqq (1 + s)(1 + \varepsilon/8)$ then $d_m(U, b_s U) < \varepsilon/4$. But our assumptions $d_m(A, U) > \varepsilon$ and $|A| = |U|$ ensure that for all x,

$$\varepsilon/2 \leqq |A\backslash(U + x)| \leqq |A^s\backslash(U + x)|$$

so $d_m(A^s, U) \geqq \varepsilon/2$ and therefore $d_m(A^s, b_s U) > \varepsilon/4$. By Lemma 4.1, for some $\eta > 0$ not depending on s or A, this implies

$$|\partial(b_s^{-1} A^s)| > (1 + \eta)|\partial U|$$

and therefore by (4.2)

$$|\partial A^s| > (1 + \eta)|\partial(b_s U)| \geqq (1 + \eta)|\partial U^s| . \tag{4.4}$$

Taking $\eta \leqq \varepsilon/8$ the theorem then follows from (4.1), (4.3) and (4.4). □

The lower bound is more complicated than (3.1) here, because not all sites of X in the optimal ball are necessarily interconnected. The first inequality in the following lemma reflects the fact that, with no compression, having about k points in the ball of optimal radius is not a particularly rare event. In contrast to Sect. III, the rarity of clusters of size near k here is almost exclusively due to the need to have a shell of empty space surrounding a ball of optimal radius. See the remarks following (3.1).

Lemma 4.3 *Given $\varepsilon, \eta > 0$ there exist $\rho_5 < \infty$ and $c_{27} < \infty$ such that if $\rho \geqq \rho_5$ and $k/\rho \geqq c_{27}$, then for $B := (k/\pi_d \rho)^{1/d} U$,*

$$\begin{aligned} P_\rho[(1 - 2\varepsilon)k \leqq |C| < \infty] &\geqq P_\rho[X(B^r\backslash B) = 0]/4 \\ &\geqq \exp(-(1 + \eta) d\pi_d^{1/d} rk(\rho/k)^{1/d}) . \end{aligned}$$

Proof. Let C_* denote the cluster of 0 for the restriction X_B of X to B. Then X outside B is independent of C_*, so

$$\begin{aligned} &P_\rho[(1 - 2\varepsilon)k \leqq |C| < \infty] \qquad (4.5)\\ &\geqq P_\rho[X(B^r\backslash B) = 0](1 - P_\rho[X(B) \leqq (1 - \varepsilon)k] - P_\rho[X(B\backslash C_*^r) \geqq \varepsilon k]) . \end{aligned}$$

Now $|B| = k/\rho$ so $E_\rho X(B) = k$; hence for large k,

$$P_\rho[X(B) \leqq (1-\varepsilon)k] \leqq 1/4 .$$

We need to show next that $B \backslash C_*^r$ is likely to be small. For some c_{28} to be specified, let

$$M := \max((c_{28}/\rho)\log(k/\rho), r) .$$

Let $\delta := \min(c_6, r/2d^{1/2})$ (cf. (3.6).) Given $x \in B \backslash C_*^r$ one of the following four possibilities must occur:

$$d(x, \partial B) \leqq M \tag{4.6}$$

$$d(x, \partial B) > M \text{ and there exists a } \delta\text{-contour } S \subset B \text{ enclosing } x \text{ with } X(D(S)) = 0 \tag{4.7}$$

$$\text{there exists a } \delta\text{-contour } S \subset B \text{ enclosing } 0 \text{ with } X(D(S)) = 0 \tag{4.8}$$

$$d(x, \partial B) > M \text{ and there exists a } \delta\text{-contour } S\text{, enclosing } 0 \text{ but excluding } x$$
$$\text{and a point } y \in \partial B_{\delta d^{1/2}}\text{, with } X(D(S) \cap B) = 0 . \tag{4.9}$$

Presuming ρ and k/ρ are large, we have

$$\operatorname{vol}(\{x \in B\colon (4.6) \text{ holds}\}) \leqq (\varepsilon/8)|B| = (\varepsilon/8)k/\rho \tag{4.10}$$

and for some c_{29}

$$\begin{aligned} P_\rho[(4.8) \text{ occurs}] &\leqq \sum_{n \geqq 2d} \sum_{S \in \mathscr{C}_n(\delta)} P_\rho[X(D(S)) = 0] \\ &\leqq \sum_{n \geqq 2d} a_d^n e^{-\rho n \delta^d/2d} \\ &\leqq \exp(-c_{29}\rho) \\ &\leqq 1/12 . \end{aligned} \tag{4.11}$$

If (4.9) occurs then for some n there exists a strongly connected set S of n δ-plaquettes in B, each intersecting $B_{\delta d^{1/2}}$, such that (i) S separates $B_{\delta d^{1/2}}$ into a region containing both x and y and a region H containing 0; (ii) the union D of all δ-cubes in H adjacent to S has $X(D) = 0$; and (iii) $n \geqq [(M - \delta d^{1/2})/\delta] \geqq M/4\delta$. Here $[\cdot]$ denotes the integer part. One can take S, for example, to be one of the strongly-connected components of the set of those plaquettes in $\partial_0 Q_\delta(C_*^r)$ which intersect $B_{\delta d^{1/2}}$. Note D consists of at least $n/2d$ δ-cubes. Similarly to (3.3), since there are at most $2d|B|/\delta^d$ δ-plaquettes in B, the number of such sets S that can occur is at most $(2d|B|/\delta^d)\tilde{a}_d^n$, where $\tilde{a}_d$ is some constant depending only on the dimension d. Therefore provided ρ is large, and c_{28} is chosen large enough, for some c_{30},

$$\begin{aligned} P_\rho[(4.9) \text{ occurs for some } x \in B] &\leqq \sum_{n \geqq M/4\delta} (2d|B|/\delta^d)\tilde{a}_d^n \exp(-\rho n \delta^d/2d) \\ &\leqq |B|\exp(-c_{30}\rho M) \\ &\leqq 1/12 . \end{aligned} \tag{4.12}$$

The calculations in (4.11) give

$$E_\rho(\operatorname{vol}(\{x \in B\colon (4.7) \text{ holds for } x\})) \leqq |B|\exp(-c_{29}\rho)$$

so that if ρ is large,

$$P_\rho[\mathrm{vol}(\{x \in B: (4.7) \text{ holds for } x\}) \geqq (\varepsilon/8)k/\rho] \leqq 1/12 . \tag{4.13}$$

Combining (4.10)–(4.13) gives

$$P_\rho[|B \backslash C^r_*| \geqq (\varepsilon/4)k/\rho] \leqq 1/4 . \tag{4.14}$$

It is easy to see that the region C^r_* can be constructed without any knowledge of the configuration X in $B\backslash C^r_*$. Therefore $X(B\backslash C^r_*)$ depends on C^r_* only through $|B\backslash C^r_*|$, and we have by (4.14)

$$P_\rho[X(B\backslash C^r_*) \geqq \varepsilon k] \leqq 1/4 + P[Z \geqq \varepsilon k] \leqq 1/2 ,$$

where Z is Poisson $(\varepsilon k/4)$.

The first inequality in the lemma now follows from (4.5); the second inequality is trivial. □

As in Sect. III, (2.20) will be proved by showing that all possible ways for the event $[|C| = k + 1]$ to occur other than that in (2.20) have probability which is much less than the lower bound in Lemma 4.3. Given a δ-contour S, let $N(S)$ denote the number of δ-plaquettes comprising S. The next lemma covers the possibility of a very large cluster; the proof is similar to Lemma 3.2 and is omitted.

Lemma 4.4 *Given $0 < \delta \leqq \min(c_6, r/2d^{1/2})$ there exists $\rho_6 < \infty$ depending on δ, r, d such that if $\rho \geqq \rho_6$ then*

$$P_\rho[|C| = k + 1, N(\partial_0 Q_\delta(C^r)) \geqq (k/\rho)^{(d-1)/d}] \leqq \exp(-2d\pi_d^{1/d} rk(\rho/k)^{1/d}) .$$

We next consider clusters occupying regions of somewhat greater than optimal volume.

Lemma 4.5 *Given $\varepsilon > 0$ there is a $\delta_0 > 0$ such that for every $0 < \delta \leqq \delta_0$ there exist $\rho_7 < \infty$ and $c_{31}, c_{32} < \infty$ depending on $\varepsilon, \delta, r, d$ such that if $\rho \geqq \rho_7$ and $k/\rho \geqq c_{31}$ then for $s := r - \delta d^{1/2}$,*

$$\begin{aligned} P_\rho[|C| = k + 1, N(\partial_0 Q_\delta(C^r)) < (k/\rho)^{(d-1)/d}, |(I(\partial_0 Q_\delta(C^r)))_s| \geqq (1 + \varepsilon)^d k/\rho] \\ \leqq \exp(-(1 + c_{32}) d\pi_d^{1/d} rk(\rho/k)^{1/d}) . \end{aligned}$$

Proof. Let D be the region enclosed by some δ-contour, with $|D_s| \geqq (1 + \varepsilon)^d k/\rho$. If $I(\partial_0 Q_\delta(C^r)) = D$, then $X((D_s)^s \backslash D_s) = 0$, while for b_s given by

$$|b_s U| = |D_s|$$

we have by Lemma 4.2

$$\begin{aligned} |(D_s)^s \backslash D_s| &\geqq |(b_s + s)U \backslash b_s U| \\ &\geqq (1 + \varepsilon)^{d-1} d\pi_d^{1/d} s(k/\rho)^{(d-1)/d} \\ &\geqq (1 + \varepsilon/2) d\pi_d^{1/d} r(k/\rho)^{(d-1)/d} \end{aligned}$$

provided δ_0 is small enough. Hence using (3.3), for $m := (k/\rho)^{(d-1)/d}$,

$$\begin{aligned} P_\rho[|C| = k + 1, N(\partial_0 Q_\delta(C^r)) < (k/\rho)^{(d-1)/d}, |(I(\partial_0 Q_\delta(C^r)))_s| \geqq (1 + \varepsilon)^d k/\rho] \\ \leqq \sum_{2d \leqq n < m} a_d^n \exp(-(1 + \varepsilon/2) d\pi_d^{1/d} rk(\rho/k)^{1/d}) \end{aligned}$$

from which the lemma follows easily. □

We continue now with the case of clusters occupying less than optimal volume.

Lemma 4.6 *Given* $\varepsilon > 0$ *there exist* $\rho_8 < \infty$ *and* $c_{33}, c_{34} < \infty$ *depending on* ε, r, d *such that if* $\rho \geqq \rho_8$ *and* $k/\rho \geqq c_{34}$ *then for* $0 < \delta \leqq \min(c_6, r/2d^{1/2})$ *and* $s := r - \delta d^{1/2}$,

$$P_\rho[|C| = k+1, N(\partial_0 Q_\delta(C^r)) < (k/\rho)^{(d-1)/d}, |(I(\partial_0 Q_\delta(C^r)))_s| \leqq (1-\varepsilon)^d k/\rho]$$
$$\leqq \exp(-c_{33} k) .$$

Proof. Let D be the region enclosed by some δ-contour, with $|D_s| \leqq (1-\varepsilon)^d k/\rho$. If $I(\partial_0 Q_\delta(C^r)) = D$, then $X(D_s) \geqq k$, while $EX(D_s) \leqq (1-\varepsilon)^d k$. Therefore for some c_{35},

$$P_\rho[|C| = k+1, I(\partial_0 Q_\delta(C^r)) = D] \leqq P[Z \geqq k] \leqq \exp(-c_{35} k) ,$$

where Z is Poisson $((1-\varepsilon)^d k)$. The result now follows by summing over n as in Lemma 4.5. □

Next we consider clusters of near-optimal volume but nonspherical shape.

Lemma 4.7 *Given* $\gamma > 0$ *one can find* ε, $\delta_0 > 0$ *such that for every* $0 < \delta \leqq \delta_0$ *there exist* $\rho_9 < \infty$ *and* $c_{36}, c_{37} < \infty$ *depending on* ε, γ, δ, r, d *such that if* $\rho \geqq \rho_9$ *and* $k/\rho \geqq c_{36}$ *then for* $s := r - \delta d^{1/2}$,

$$P_\rho[|C| = k+1, N(\partial_0 Q_\delta(C^r)) < (k/\rho)^{(d-1)/d},$$
$$(1-\varepsilon)^d k/\rho \leqq |(I(\partial_0 Q_\delta(C^r)))_s| \leqq (1+\varepsilon)^d k/\rho, d_m(\lambda_{k/\rho}^{-1}(C^r)_r, U) > \gamma]$$
$$\leqslant \exp(-(1+c_{37}) d\pi_d^{1/d} r k(\rho/k)^{1/d}) .$$

Proof. There are two ways for the event of nonsphericality

$$d_m(\lambda_{k/\rho}^{-1}(C^r)_r, U) > \gamma \tag{4.15}$$

to occur, under the condition of near-optimal volume

$$(1-\varepsilon)^d k/\rho \leqq |(I(\partial_0 Q_\delta(C^r))_s| \leqq (1+\varepsilon)^d k/\rho . \tag{4.16}$$

Roughly, either the outer boundary $\partial_0(C^r)_r$ is nonspherical, or the outer boundary is essentially spherical but $(C^r)_r$ has holes in it, i.e. a substantial fraction of the interior of $\partial_0(C^r)_r$ is not in $(C^r)_r$.

The first possibility can be dealt with by contour-counting, as in the last two lemmas. With ε, δ_0 to be specified later and $\delta < \delta_0$, let D be the region enclosed by some δ-contour, with

$$N(\partial_0 D) < (k/\rho)^{(d-1)/d} \tag{4.17}$$

$$(1-\varepsilon)^d k/\rho \leqq |D_s| \leqq (1+\varepsilon)^d k/\rho \tag{4.18}$$

and

$$d_m((|U|/|D_s|)^{1/d} D_s, U) > \gamma/4 . \tag{4.19}$$

The idea is that D_s and $(D_s)^s$ are possible values of approximations for $(C^r)_r$ and for C^r, each "with holes filled in." Let $\sigma := (|D_s|/|U|)^{1/d}$, so

$$|\sigma^{-1} D_s| = |U| \tag{4.20}$$

and

$$\sigma \geqq (1-\varepsilon)(k/\pi_d \rho)^{1/d} .$$

By Lemma 4.2 there is a constant $\eta(\gamma, d) \in (0, 1)$ such that

$$\begin{aligned} |D \backslash D_s| &\geqq |(D_s)^s \backslash D_s| \\ &\geqq (1+\eta)\sigma^d |U^{s/\sigma} \backslash U| \\ &\geqq (1+\eta)\sigma^{d-1} d\pi_d s \end{aligned}$$

which gives

$$|D \backslash D_s| \geqq (1+\eta/2) d\pi_d^{1/d} r(k/\rho)^{(d-1)/d}, \tag{4.21}$$

the last inequality being valid provided ε, δ_0 are sufficiently small. If $I(\partial_0 Q_\delta(C^r)) = D$, then $X(D \backslash D_s) = 0$. Therefore by (3.4), for some c_{38}, provided ρ_9 is large,

$$\begin{aligned} P_\rho[|C| = k+1, I(\partial_0 Q_\delta(C^r)) = D \text{ for some } D \text{ satisfying (4.17) and (4.21)}] \\ \leqq \exp(c_{38}(k/\rho)^{(d-1)/d}) \exp(-(1+\eta/2) d\pi_d^{1/d} rk(\rho/k)^{1/d}) \qquad (4.22) \\ \leqq \exp(-(1+\eta/4) d\pi_d^{1/d} rk(\rho/k)^{1/d}). \end{aligned}$$

The other possibility, when $|C| = k+1$, under (4.15) and (4.16) is that $I(\partial_0 Q_\delta(C^r)) = D$ for some D satisfying (4.17) and (4.18) but not (4.21), hence also not (4.19). Fix such a D. We claim that

$$|D_s \backslash (C^r)_r| > (\gamma/9) k/\rho. \tag{4.23}$$

Note that

$$(C^r)_r \subset D_s \subset D \subset I(\partial_0 C^r)$$

and that, provided c_{36} is large,

$$(1-\varepsilon)(k/\pi_d \rho)^{1/d} < \sigma, \lambda_{k/\rho} < (1+\varepsilon)(k/\pi_d \rho)^{1/d}. \tag{4.24}$$

Since (4.19) does not hold, for some x and $\tilde{U} := U + x$,

$$|\sigma^{-1} D_s \triangle \tilde{U}| \leqq \gamma/4. \tag{4.25}$$

Now provided ε is small, using (4.20), (4.24) and (4.25),

$$|\lambda_{k/\rho}^{-1}(C^r)_r \backslash \tilde{U}| \leqq |\lambda_{k/\rho}^{-1} D_s \backslash \tilde{U}| \leqq (\sigma/\lambda_{k/\rho})^d (|\sigma^{-1} D_s \backslash \tilde{U}| + |U \backslash (\lambda_{k/\rho}/\sigma) U|) \leqq \gamma/6.$$

Therefore by (4.15),

$$|\tilde{U} \backslash \lambda_{k/\rho}^{-1}(C^r)_r| > 5\gamma/6. \tag{4.26}$$

Also

$$|\tilde{U} \backslash \lambda_{k/\rho}^{-1} D_s| \leqq (\sigma/\lambda_{k/\rho})^d (|(\lambda_{k/\rho}/\sigma) U \backslash U| + |\tilde{U} \backslash \sigma^{-1} D_s|) \leqq \gamma/6$$

which with (4.26) gives

$$|\lambda_{k/\rho}^{-1} D_s \backslash \lambda_{k/\rho}^{-1}(C^r)_r| > 2\gamma/3$$

which with (4.24) yields the claim (4.23), since $\pi_d \leqq 6$ for all d.

By (4.18) and failure of (4.19), taking ε small we may assume

$$|D| \leqq (1+\gamma/40) k/\rho. \tag{4.27}$$

Under (4.23) there are two possibilities: either

$$|D_s \cap (C^r \backslash (C^r)_r)| > (\gamma/18)k/\rho \tag{4.28}$$

or

$$|D_s \backslash C^r| > (\gamma/18)k/\rho \, . \tag{4.29}$$

Under (4.28) there is an excessive amount of empty space in D, as follows. On the event $I(\partial_0 Q_\delta(C^r)) = D$ we have

$$\cup \{G \in \mathscr{G}(\delta) : G \subset D, X(G) = 0\} \supset Q_\delta(C^r \backslash (C^r)_r) \, .$$

If δ_0 is small enough relative to r, and $\gamma < 1$, (4.27) and (4.28) imply

$$|Q_\delta(C^r \backslash (C^r)_r)| \geqq |C^r \backslash (C^r)_r|/2 \geqq (\gamma/36)k/\rho \geqq (\gamma/40)|D|$$

so that

$$M := \operatorname{card}(\{G \in \mathscr{G}(\delta) : G \subset D, X(G) = 0\})$$

and

$$m := \operatorname{card}(\{G \in \mathscr{G}(\delta) : G \subset D\})$$

satisfy

$$M \geqq (\gamma/40)m \, .$$

But $P[X(G) = 0] = \exp(-\rho\delta^d)$, $m = |D|/\delta^d \geqq k/2\rho\delta^d$ by (4.18), and the random variables $X(G)$ are iid, so by Bennett's inequality ([Be], or see [Ho]) still assuming (4.28) and ρ_9 large, for some c_{39},

$$\begin{aligned} P_\rho[|C| = k+1, I(\partial_0 Q_\delta(C^r)) = D] & \\ \leqq P[M \geqq (\gamma/40)m] & \\ \leqq \exp(-\gamma m \rho \delta^d/160) & \\ \leqq \exp(-c_{39}k) \, . & \end{aligned} \tag{4.30}$$

Therefore by (3.4) and (4.30), provided ρ_9 is large, for some c_{40},

$$\begin{aligned} P_\rho[|C| = k+1, I(\partial_0 Q_\delta(C^r)) = D & \\ \text{for some } D \text{ such that (4.17), (4.18), (4.27) hold and (4.19) fails}] & \\ \leqq \exp(-c_{40}k) \, . & \end{aligned} \tag{4.31}$$

Under the alternate possibility (4.29) there are likely to be too many sites of X in D, as follows. Let C_* denote the cluster of 0 for the restriction of X to D_s. As in the proof of Lemma 4.3, C_*^r can be constructed without any knowledge of the configuration X in $D_s \backslash C_*^r$. On the event $I(\partial_0 Q_\delta(C^r)) = D$ we have $C = C_*$. It follows under (4.29) that given C_*, $X(D_s \backslash C_*^r)$ is stochastically larger than Poisson $((\gamma/18)k)$. Therefore using (4.27), for Z a Poisson $((\gamma/18)k)$ r.v., still assuming (4.29), for some c_{41},

$$\begin{aligned} P_\rho[|C| = k+1, I(\partial_0 Q_\delta(C^r)) = D, \text{(4.28) holds}] \; P[Z > (\gamma/36)k] & \\ \leqq P[X(D) > (1 + \gamma/36)k] & \\ \leqq \exp(-c_{41}k) \, . & \end{aligned} \tag{4.32}$$

Therefore by (3.4) and (4.32), provided ρ_9 is large, for some c_{42},

$$P_\rho[|C| = k + 1, I(\partial_0 Q_\delta(C^r)) = D \quad \text{for some } D \text{ such that (4.17), (4.18), (4.29) hold and (4.19) fails}] \leqq \exp(-c_{42}k) . \tag{4.33}$$

With (4.31) and (4.22) this proves the lemma. □

Remark 4.8 In dimension $d = 2$, d_m can be replaced by d_H in (2.20) of Theorem 2.5. This is because (4.21) is valid when we have d_H in place of d_m in (4.19), provided that, for σ and D as in the last proof, σ is sufficiently large and D is a possible value of $I(\partial_0 Q_\delta(C^r))$. Indeed, roughly, suppose $d_H(\sigma^{-1}D_s, U)$ is large but $d_m(\sigma^{-1}D_s, U)$ is not, so $|\sigma^{-1}D_s \triangle (U + x)|$ is small for some x, which we take for simplicity to be 0. Then $\sigma^{-1}D_s$ must contain a point y a large distance from U; from the nature of C^r, $\sigma^{-1}D_s$ must actually contain a string of points, with adjacent ones separated by distance at most r/σ, connecting y to a point z near ∂U, so $(D_s)^s$ is connected. But $|(D_s)^s \backslash D_s|$ is at least roughly s times the length of $\partial(D_s)^s$. Because $d = 2$ and $(D_s)^s$ is connected, the length of $\partial(D_s)^s$ is at least approximately as great as the length of the convex hull boundary $\partial(\sigma(U \cup \{y\})_H)$, which is significantly more than the length of $\partial(\sigma U)$. This in turn gives (4.21). □

Corresponding to Lemma 3.10, to ensure that the lower bound in Lemma 4.3 really gives the right rate, we need an upper bound for C of near-optimal volume, without restriction on the shape.

Lemma 4.9 *Given $\eta > 0$ one can find $\varepsilon, \delta_0 > 0$ such that for every $0 < \delta \leqq \delta_0$ there exist $\rho_{10} < \infty$ and $c_{43} < \infty$ such that if $\rho \geqq \rho_{10}$ and $k/\rho \geqq c_{43}$ then for $s := r - \delta d^{1/2}$,*

$$P_\rho[|C| = k + 1, N(\partial_0 Q_\delta(C^r)) < (k/\rho)^{(d-1)/d}, |(I(\partial_0 Q_\delta(C^r)))_s| \geqq (1 - \varepsilon)k/\rho] \leqq \exp(-(1 - \eta) d\pi_d^{1/d} rk(\rho/k)^{1/d}) .$$

Proof. This is essentially the same as the derivation of (4.22) in the last lemma, with $1 + \eta$, $1 + \eta/2$, $1 + \eta/4$ replaced in (4.21) and (4.22) by constants slightly less than 1. □

Proposition 4.10 *Given $\alpha > 0$ there exist $\rho_{11} < \infty$ and $c_{44} < \infty$ such that if $\rho \geqq \rho_{11}$ and $k/\rho \geqq c_{44}$ then*

$$\exp(-(1 + \alpha) d\pi_d^{1/d} rk(\rho/k)^{1/d}) \leqq P_\rho[k + 1 \leqq |C| < \infty] \leqq \exp(-(1 - \alpha) d\pi_d^{1/d} rk(\rho/k)^{1/d}) .$$

Proof. Assume $\alpha < 1$. The lower bound is a slight reformulation of Lemma 4.3. For the upper bound, let ε, δ_0 be as in Lemma 4.9 with $\eta = \alpha/2$, and let δ be as in Lemmas 4.4, 4.6, and 4.9. The latter three lemmas show, for ρ_{11} and c_{43} large enough, that

$$P_\rho[|C| = k + 1] \leqq 3 \exp(-(1 - \alpha/2) d\pi_d^{1/d} rk(\rho/k)^{1/d}) .$$

The proposition now follows by summing the series. □

In order to obtain the desired results about $P_\rho[|C| = k + 1]$ from Proposition 4.10, we will use the following, which is similar to a result of Kunz and Souillard

[KS] for site percolation on a lattice, though our use of a continuous model introduces additional complications. The terms "+1" in Lemma 4.11 do not appear in the result for lattices, and are an artifact of the existence in our model of a fixed site at 0.

Lemma 4.11 *For every* $\rho > 0$ *and* $j, \ell \geqq 0$,

$$\frac{P_\rho[|C| = j + \ell + 1]}{j + \ell + 1} \geqq \frac{P_\rho[|C| = j+1]}{j+1} \frac{P_\rho[|C| = \ell + 1]}{\ell + 1}. \tag{4.34}$$

Proof. Let us use "left" and "right" in this proof to signify relative values of the dth coordinate, i.e. x is left of y means x has a smaller dth coordinate.

Let C_L and C_R denote the cluster of 0 for the restriction of X to the half spaces H_L and H_R to the left and right of 0, respectively. By translation invariance we have

$$P_\rho[|C| = j+1]/(j+1) = P_\rho[|C| = j + 1 \text{ and } 0 \text{ is the rightmost point of } C] . \tag{4.35}$$

Define events

$$A := [|C_L| = j + 1], B := [|C_R| = \ell + 1] .$$

Now A is not the same as the event on the right side of (4.35), but we do have

$$\begin{aligned} &P_\rho[|C| = j + 1 \text{ and } 0 \text{ is the rightmost point of } C] \\ &= P_\rho(A \cap [X(C_L^r \cap H_R) = 0]) \\ &= \int_A \exp(-\rho|C_L^r \cap H_R|)dP . \end{aligned} \tag{4.36}$$

Similarly, reversing left and right,

$$P_\rho[|C| = \ell + 1]/(\ell + 1) = \int_B \exp(-\rho|C_R^r \cap H_L|)\, dP . \tag{4.37}$$

Observing that knowledge of C_L does not affect the configuration X outside $C_L^r \cap H_L$, and similarly for C_R, we obtain

$$\begin{aligned} &P_\rho[|C| = j + \ell + 1]/(j + \ell + 1) \\ &\quad = P_\rho[|C| = j + \ell + 1 \text{ and } 0 \text{ is the } (j+1)\text{st leftmost point of } C] \\ &\quad \geqq P_\rho(A \cap B \cap [X(C_L^r \cap H_R \backslash C_R^r) = 0] \cap [X(C_R^r \cap H_L \backslash C_L^r) = 0]) \\ &\quad = \int_{A \cap B} \exp(-\rho|C_L^r \cap H_R \backslash C_R^r|)\exp(-\rho|C_R^r \cap H_L \backslash C_L^r|)dP \\ &\quad \geqq \int_{A \cap B} \exp(-\rho|C_L^r \cap H_R|)\exp(-\rho|C_R^r \cap H_L|)dP . \end{aligned} \tag{4.38}$$

Since C_L and C_R are independent, (4.35)–(4.37) show that the right side of (4.38) is equal to the right side of (4.34). □

Somewhat as in [KS], Lemma 4.11 will help us to reduce the problem of lower bounds for $P_\rho[|C| = k + 1]$ to the same problem for a much smaller value of k, where the following crude bound will suffice.

Lemma 4.12 *For every $M < \infty$ there exist c_{45} and ρ_{12} such that if $\rho \geqq \rho_{12}$ and $k < M\rho$ then*

$$P_\rho[|C| = k+1] \geqq \exp(-c_{45}\rho)\,.$$

Proof. Let $B := (k/\rho\pi_d)^{1/d}U$, so $|B| = k/\rho < M$, and let $\delta := r/2d^{1/2}$. Then

$$P_\rho[|C| = k+1] \tag{4.39}$$

$$\geqq P_\rho[X(B^r \backslash B) = 0, X(B) = k, X(G) \geqq 1 \text{ for all } G \in \mathscr{G}(\delta) \text{ with } G \subset B]\,.$$

Now

$$\begin{aligned} &P_\rho[X(G) = 0 \text{ for some } G \in \mathscr{G}(\delta) \text{ with } G \subset B | X(B) = k] \\ &\quad \leqq (|B|/\delta^d)(1 - \delta^d/|B|)^k \\ &\quad \leqq (M/\delta^d)\exp(-\rho\delta^d) \\ &\quad \leqq 1/2 \end{aligned}$$

provided ρ_{12} is large. Therefore by (4.39) and Stirling's formula, for some c_{46},

$$\begin{aligned} P_\rho[|C| = k+1] &\geqq P_\rho[X(B^r \backslash B) = 0, X(B) = k]/2 \\ &\geqq \exp(-c_{46}\rho)/4(2\pi k)^{1/2} \end{aligned}$$

and the lemma follows easily. □

Proof of Thorem 2.5 Fix $0 < \eta < 1/8$. By Proposition 4.10, for some ρ_{13} and c_{47}, if $\rho \geqq \rho_{13}$ and $k/\rho \geqq c_{47}$,

$$P_\rho[(1 - 5\eta)k \leqq |C| \leqq k] \geqq \exp(-(1-\eta)\pi_d^{1/d} r\rho^{1/d}k^{(d-1)/d})\,.$$

Therefore there exists $1 \leqq j_1 \leqq 5\eta k + 1$ such that

$$\begin{aligned} &P_\rho[|C| = k - j_1 + 1]/(k - j_1 + 1) \\ &\quad \geqq (5\eta k + 1)^{-1}\exp(-(1-\eta)\pi_d^{1/d} r\rho^{1/d}k^{(d-1)/d})/(k - j_1 + 1) \\ &\quad \geqq \exp(-\pi_d^{1/d} r\rho^{1/d}k^{(d-1)/d})\,. \end{aligned}$$

If $j_1/\rho \geqq c_{47}$ we then similarly obtain $j_2 \leqq 5\eta j_1$ such that

$$P_\rho[|C| = j_1 - j_2 + 1]/(j_1 - j_2 + 1) \geqq \exp(-\pi_d^{1/d} r\rho^{1/d} j_1^{(d-1)/d})\,.$$

We can continue in this manner, using $j_0 := k$, until we reach the largest index n for which $j_n/\rho \geqq c_{47}$. By Lemma 4.12 applied with $M := c_{47}$, for some c_{48},

$$P_\rho[|C| = j_{n+1} + 1]/(j_{n+1} + 1) \geqq \exp(-c_{45}\rho)/C_{47}\rho \geqq \exp(-c_{48}\rho)\,.$$

Successive applications of Lemma 4.11 give the lower bound

$$\begin{aligned} P_\rho[|C| = k+1] &\geqq \left[\prod_{i=0}^{n} P_\rho[|C| = j_i - j_{i+1} + 1]/(j_i - j_{i+1} + 1)\right] \\ &\quad \cdot P_\rho[|C| = j_{n+1} + 1]/(j_{n+1} + 1) \\ &\geqq \exp(-\pi_d^{1/d} r\rho^{1/d}(\sum_{i=0}^{n} j_i^{(d-1)/d}) - c_{48}\rho) \\ &\geqq \exp(-(1 - (6\eta)^{(d-1)/d})^{-1}\pi_d^{1/d} r\rho^{1/d}k^{(d-1)/d})\,, \end{aligned}$$

the last inequality being valid provided k/ρ is sufficiently large. Proposition 4.10 provides a similar upper bound for $P_\rho[|C| = k+1]$, and, since $\eta > 0$ is arbitrary, the theorem then follows from these two bounds and Lemmas 4.4–4.7, analogously to the proof of Theorem 2.3. Equivalence of convergence for d_H and d_m follows from the discussion after (2.5). □

V Proofs when the normal volume is moderate

In this section we will prove Theorem 2.4. Some of the ideas are relatively similar to the proofs of Theorems 2.3 and 2.5, so we will be a little sketchy at times. From the proof of Lemma 4.12 we have for $B = \lambda_\alpha U$

$$P_\rho[|C| = k+1] \geqq P_\rho[X(B^r \backslash B) = 0, X(B) = k]/2 .$$

Fix $\zeta > 0$. Using Stirling's formula and these restatements of (2.1):

$$\pi_d(\lambda_\alpha + r)^d/\alpha = 1 + r/\lambda_\alpha; \qquad \alpha/\pi_d \lambda_\alpha^d = (1 + r/\lambda_\alpha)^{d-1} , \tag{5.1}$$

this gives, provided k is large,

$$\begin{aligned} P_\rho[|C| = k+1] &\geqq \exp(-[r/\lambda_\alpha + (d-1)\log(1 + r/\lambda_\alpha)]k)/3(2\pi k)^{1/2} \\ &\geqq \exp(-(1+\zeta)[r/\lambda_\alpha + (d-1)\log(1 + r/\lambda_\alpha)]k) . \end{aligned} \tag{5.2}$$

By Lemma 3.2 we may restrict our attention to $\mathrm{diam}(C) \leqq R$ for some large R. Let $\delta > 0$, $s := r - \delta d^{1/2}$, and let $D \in \mathscr{H}(\delta)$ be a possible value for $Q_\delta(C^r)$. Define τ, τ^* by

$$|D_s| = |(1+\tau)\lambda_\alpha U|, \quad |D| = |(s + (1+\tau^*)\lambda_\alpha)U| .$$

Then $\tau \leqq \tau^*$ by (4.1) and Lemma 4.1. Note τ is near 0 when D_s has near-optimal volume, and τ^* is near 0 if D_s is also approximately spherical. We have

$$\begin{aligned} P_\rho[|C| = k+1, Q_\delta(C^r) = D] &\leqq P_\rho[X(D_s) = k, X(D \backslash D_s) = 0] \\ &= \exp(-\rho|D|)(\rho|D_s|)^k/k! . \end{aligned} \tag{5.3}$$

Using (4.1) and Lemma 4.1 again, along with (5.1) and the definition (2.1) of λ_α, we obtain

$$\begin{aligned} |D| &= \pi_d(s + (1+\tau^*)\lambda_\alpha)^d \\ &\geqq \pi_d(\lambda_\alpha + r)^d + d\alpha(\tau^* - \delta d^{1/2}/\lambda_\alpha) \\ &= \alpha(1 + r/\lambda_\alpha + d(\tau^* - \delta d^{1/2}/\lambda_\alpha)) . \end{aligned}$$

With (5.1) and (5.3) this yields

$$\begin{aligned} &P_\rho[|C| = k+1, Q_\delta(C^r) = D] \\ &\leqq \exp(-[r/\lambda_\alpha + (d-1)\log(1 + r/\lambda_\alpha) \\ &\quad + d(\tau^* - \log(1+\tau) - \delta d^{1/2}/\lambda_\alpha)]k) . \end{aligned} \tag{5.4}$$

Fix $0 < \varepsilon < 1/2$. Note λ_α is bounded between λ_m and λ_M. If δ is sufficiently small (depending on ε), then there exists $c_{48} > 0$ such that

$$|\tau| > \varepsilon \quad \text{implies } d(\tau - \log(1+\tau) - \delta d^{1/2}/\lambda_\alpha) > c_{48}$$

and

$$\tau^* > \tau + \varepsilon \quad \text{implies } d(\tau^* - \log(1+\tau) - \delta d^{1/2}/\lambda_\alpha) > c_{48} .$$

Under either of these two conditions we get from (5.4):

$$\begin{aligned} &P_\rho[|C| = k+1, \quad Q_\delta(C^r) = D] \\ &\leqq \exp(-[r/\lambda_\alpha + (d-1)\log(1 + r/\lambda_\alpha) + c_{48}]k) . \end{aligned} \tag{5.5}$$

Thus let us henceforth assume D is such that $|\tau| \leqq \varepsilon$ and $\tau^* \leqq \tau + \varepsilon$. Fix $\gamma > 0$, and fix $a > 0$ to be specified later. Let us show that $Q_\delta(C^r) = D$ implies

$$d_m(\lambda_\alpha^{-1}(C^r)_r, U) \leqq \gamma . \tag{5.6}$$

Our proof will be somewhat like that of Lemma 4.7, but simplified by the fact that here D approximates not only the outer boundary of C^r but also the holes, if any. We claim that if ε is small enough (depending on μ, γ and λ_M) then $\tau^* \leqq \tau + \varepsilon$ ensures that

$$d_m(((1+\tau)\lambda_\alpha)^{-1} D_s, U) \leqq \min((a\gamma)^{(d+1)/2}, \gamma/8) . \tag{5.7}$$

For suppose that (5.7) fails. Let $\sigma := (1+\tau)\lambda_\alpha$, so $|\sigma^{-1} D_s| = |U|$. By Lemma 4.2 there exists $0 < t_0 < s/(1+\tau)\lambda_\alpha$, $c_{49} > 0$ and $\eta > 0$ depending on a, γ such that

$$\begin{aligned} |D \backslash D_s| &\geqq |(D_s)^s \backslash D_s| \\ &= \sigma^d(|(\sigma^{-1} D_s)^{t_0} \backslash (\sigma^{-1} D_s)| + |(\sigma^{-1} D_s)^{s/\sigma} \backslash (\sigma^{-1} D_s)^{t_0}|) \\ &\geqq \sigma^d(|U^{s/\sigma} \backslash U| + \eta |U^{t_0} \backslash U|) \\ &\geqq |(\sigma U)^s \backslash \sigma U| + c_{49} . \end{aligned}$$

Therefore

$$|D| \geqq |(s + (1+\tau)\lambda_\alpha)U| + c_{49} ,$$

which if ε is small enough (depending on a, γ and λ_M) implies $\tau^* > \tau + \varepsilon$. This proves the claim (5.7).

By (5.7) we may assume that

$$|((1+\tau)\lambda_\alpha)^{-1} D_s \triangle U| \leqq \min((a\gamma)^{(d+1)/2}, \gamma/8); \tag{5.8}$$

if not then replace U throughout by a fixed translation of U for which this inequality holds.

If (5.6) fails then either

$$|\lambda_\alpha^{-1}(C^r)_r \backslash U| > \gamma/2 \tag{5.9}$$

or

$$|U \backslash \lambda_\alpha^{-1}(C^r)_r| > \gamma/2 . \tag{5.10}$$

But if $Q_\delta(C^r) = D$, (5.9) and (5.8) would imply

$$\lambda_\alpha^d |(1+\tau)U \backslash U| = |\sigma U \backslash \lambda_\alpha U| \geqq |D_s \backslash \lambda_\alpha U| - |D_s \backslash \sigma U| \geqq (\lambda_\alpha^d/2 - \sigma^d/8)\gamma \geqq \lambda_\alpha^d \gamma/4$$

implying $|\tau| > \varepsilon$, provided ε is small enough, depending on γ and λ_m. This rules out (5.9).

To rule out (5.10) when $Q_\delta(C^r) = D$, let $y \in U \backslash \lambda_\alpha^{-1}(C^r)_r$ maximize $d(y, \partial U)$ and define b by $d(y, \partial U) = b\gamma$. Note that

$$|U \backslash \lambda_\alpha^{-1}(C^r)_r| \leqq b\gamma |\partial U| , \tag{5.11}$$

so we wish to bound b. There exists $z \notin \lambda_\alpha^{-1} C^r$ for which $d(z, y) \leqq \lambda_\alpha^{-1} r$; since $z \notin \lambda_\alpha^{-1} D$, the ball $V := z + s\lambda_\alpha^{-1} U$ does not meet $\lambda_\alpha^{-1} D_s$, hence satisfies

$$U \cap V \subset U \backslash \lambda_\alpha^{-1} D_s .$$

Let $x \in U \cap V$ maximize $d(x, \partial U)$; then $d(x, \partial U) \geqq b\gamma - \delta d^{1/2}$. Since the radius of V is at least $r/2\lambda_M$, there exists a constant c_{50} depending only on $r/2\lambda_M$ and d such that

$$|U \backslash \lambda_\alpha^{-1} D_s| \geqq |U \cap V| \geqq c_{50}(b\gamma - \delta d^{1/2})^{(d+1)/2} . \tag{5.12}$$

But from (5.8), if ε is small enough, depending on a and γ,

$$\begin{aligned} |U \backslash \lambda_\alpha^{-1} D_s| &\leqq |U \backslash (1+\tau) U| + (1+\tau)^d |U \backslash \sigma^{-1} D_s| \\ &\leqq |\partial U| \varepsilon + (a\gamma)^{(d+1)/2} \\ &\leqq 2(a\gamma)^{(d+1)/2} . \end{aligned} \tag{5.13}$$

From (5.12) and (5.13), either

$$b\gamma/2 \leqq \delta d^{1/2}$$

or

$$c_{50}(b\gamma/2)^{(d+1)/2} \leqq 2(a\gamma)^{(d+1)/2} .$$

Thus for some c_{51} depending only on $r/2\lambda_M$ and d,

$$b \leqq \max(c_{51} a, 2\delta d^{1/2}/\gamma) .$$

Thus if we choose a small enough (depending on $r/2\lambda_M$ and d) and δ small enough (depending on γ and d), it follows that (5.10) does not hold.

We have shown that if $Q_\delta(C^r) = D$ for some D such that $|\tau| \leqq \varepsilon$ and $\tau^* \leqq \tau + \varepsilon$, then (5.6) holds. Recalling that we have restricted to $\operatorname{diam}(C) \leqq R$, so that the number of possible D is finite, with (5.4) and (5.5) this proves Theorem 2.4, similarly to the proofs of Theorems 2.3 and 2.5. □

VI The uniform distribution within the cluster

This section contains the proof of Theorem 2.6. Recall that Y_k denotes the centroid of $\operatorname{Co}(C)$. It is sufficient to show that for every $\tau > 0$, if $\varepsilon > 0$ is sufficiently small (not depending on k or ρ) then there exist events $A_k \subset [|C| = k+1]$ satisfying

$$P_\rho(A_k \,|\, |C| = k+1) \to 1 \quad \text{as } k, \rho \to \infty \tag{6.1}$$

such that for $N := X(Y_k + (1-3\varepsilon)\lambda_{k/\rho} U)$

$$A_k \subset [N \geqq (1-\tau)k] \tag{6.2}$$

$$Y_k + (1-\varepsilon)\lambda_{k/\rho} U \subset \operatorname{Co}(C) \subset Y_k + (1+\varepsilon)\lambda_{k/\rho} U \tag{6.3}$$

and

$$\text{given } A_k \cap [N = n], \text{ the } n \text{ points } \{x - Y_k : x \in X \cap (Y_k + (1 - 3\varepsilon)\lambda_{k/\rho} U)\} \quad (6.4)$$
$$\text{are independent and uniformly distributed in } (1 - 3\varepsilon)\lambda_{k/\rho} U \,.$$

Let $\tau > 0$, let $\varepsilon > 0$, let $\zeta > 0$ be a constant to be specified, and let $\delta = \zeta k/\rho$. We consider the case of $k/\rho \leqq c_{51}$, with $c_{51}(\zeta, \varepsilon)$ to be specified. The proof for $k/\rho > c_{51}$ but bounded is similar, so we omit it. For a bounded subset D of $\mathbb{R}^d$ let z_D denote the centroid of $\mathrm{Co}(D)$ and

$$G_1^D := z_D + (1 - 2\varepsilon)\lambda_{k/\rho} U, \quad G_2^D := z_D + (1 + 2\varepsilon)\lambda_{k/\rho} U \,.$$

Let A_k be the event that

$$|C| = k + 1 \tag{6.5}$$

$$Y_k + (1 - \varepsilon)\lambda_{k/\rho} U \subset \mathrm{Co}(C) \subset Y_k + (1 + \varepsilon)\lambda_{k/\rho} U \tag{6.6}$$

and

$$X(G_2^D \backslash G_1^D) < \tau k \quad \text{for } D = Q^\delta(\mathrm{Co}(C)) \,. \tag{6.7}$$

Let D be a possible value of $Q^\delta(\mathrm{Co}(C))$. Assume now that A_k occurs and $Q^\delta(\mathrm{Co}(C)) = D$. For some $c_{52}(d)$, provided ζ and c_{51} are small,

$$\| Y_k - z_D \| \leqq c_{52}\delta < \varepsilon\lambda_{k/\rho} \,. \tag{6.8}$$

Therefore

$$\mathrm{Co}(C) = \mathrm{Co}(C \cap (G_2^D \backslash G_1^D)) \,. \tag{6.9}$$

This means that Y_k is a function of the restriction of X to $G_2^D \backslash G_1^D$. It also means that the occurrence of $A_k \cap [Q^\delta(\mathrm{Co}(C)) = D]$ is unaffected by the positions of the sites of X in G_1^D. Therefore, given $A_k \cap [Q^\delta(\mathrm{Co}(C)) = D]$, the sites of X in G_1^D are iid uniform in G_1^D. Since D is arbitrary, and since by (6.8) $Y_k + (1 - 3\varepsilon)\lambda_{k/\rho} U \subset G_1^D$ on $A_k \cap [Q^\delta(\mathrm{Co}(C)) = D]$, (6.4) follows.

Now (6.2) follows from (6.6) and (6.7). From Theorems 2.3 and 2.4,

$$P_\rho[(6.6) \mid |C| = k + 1] \to 1 \quad \text{as } k, \rho \to \infty \,,$$

so to prove (6.1) it is enough to show

$$P_\rho[(6.6) \text{ holds, } (6.7) \text{ fails} \mid |C| = k + 1] \to 0 \quad \text{as } k, \rho \to \infty \,. \tag{6.10}$$

Let D be a possible value of $Q^\delta(\mathrm{Co}(C))$ under (6.6). Then as in (3.7)

$$\begin{aligned} &P_\rho[|C| = k + 1, Q^\delta(\mathrm{Co}(C)) = D, X(D \cap G_2^D \backslash G_1^D) \geqq \tau k] \\ &\leqq P_\rho[X(D) = k, X(D \cap G_2^D \backslash G_1^D) \geqq \tau k] \sup_{A \in \mathscr{U}} P_\rho[X(A^r \backslash D) = 0] \qquad (6.11) \\ &= (k!)^{-1}(\rho|D|)^k \exp(-\rho \inf_{A \in \mathscr{U}} |A^r|) P_\rho[X(D \cap G_2^D \backslash G_1^D) \geqq \tau k | X(D) = k] \,, \end{aligned}$$

where now $\mathscr{U}$ denotes the collection of all finite sets $A \subset D$ with $|A| = k + 1, 0 \in A$, and $Q^\delta(\mathrm{Co}(A)) = D$. Provided ε is small, we have

$$|G_2^D \backslash G_1^D| \leqq \tau |\lambda_{k/\rho} U|/4 \,,$$

while

$$|D| \geqq (1 - \varepsilon)^d |\lambda_{k/\rho} U|$$

so that, provided ε is small,

$$|D \cap G_2^D \backslash G_1^D|/|D| \leqq \tau/2 .$$

This implies that the probability on the right side of (6.11) satisfies

$$P_\rho[X(D \cap G_2^D \backslash G_1^D) \geqq \tau k | X(D) = k] \leqq e^{-c_{53}k} \tag{6.12}$$

for some $c_{53}(\tau)$.

Provided c_{51} is taken small enough (depending on ε), we have

$$(1-2\varepsilon)\lambda_{k/\rho} \geqq (1-4\varepsilon)c_1 k/\rho \quad \text{and} \quad (1+2\varepsilon)\lambda_{k/\rho} \leqq (1+4\varepsilon)c_1 k/\rho . \tag{6.13}$$

As in the proof of (3.8) within the proof of Lemma 3.8, it is easily checked that, provided ζ is small, for $A \in \mathscr{U}$,

$$A^r \supset (Y_k + (1-2\varepsilon)\lambda_{k/\rho} U)^r$$

so that using (6.13)

$$|A^r| \geqq \pi_d r^d + d\pi_d r^{d-1}(1-4\varepsilon)c_1 k/\rho = \pi_d r^d + d(1-4\varepsilon)k/\rho$$

while, provided ζ is small,

$$|D| \leqq \pi_d((1+2\varepsilon)\lambda_{k/\rho})^d \leqq \pi_d((1+4\varepsilon)c_1 k/\rho)^d .$$

Therefore from (6.11) and (6.12), using Stirling's formula,

$$\begin{aligned} &P_\rho[|C| = k+1, Q^\delta(\mathrm{Co}(C)) = D, X(D \cap G_2^D \backslash G_1^D) \geqq \tau k] \\ &\leqq (\rho|D|e/k)^k \exp(-\rho\pi_d r^d - (1-4\varepsilon)dk) \\ &\leqq \exp(-[\rho\pi_d r^d + (d-1)k\log\rho/k + (d-1)k\log(e\pi_d r^d) \\ &\qquad - (4\varepsilon + \log(1+4\varepsilon)kd + c_{53}k)]) . \end{aligned}$$

Since the number of possible D remains bounded as $k, \rho \to \infty$, this and (3.1) prove (6.10), provided c_{51} and ε are small. □

References

[ACC] Alexander, K.S., Chayes, J.T., Chayes, L.: The Wulff construction and asymptotics of the finite cluster distribution for two-dimensional Bernoulli percolation. Commun. Math. Phys. **131**, 1–50 (1990)

[Be] Bennett, G.: Probability inequalities for the sum of independent random variables. J. Am. Stat. Assoc. **57**, 33–45 (1962)

[Bo] Bonessen, T.: Les problèmes des isopérimètres et des isépiphanes. Paris: Gauthier-Villars 1929

[Bu] Busemann, H.: Convex surfaces. New York: Interscience 1958

[GS1] Given, J.A., Stell, G.: The continuum Potts model and continuum percolation. Physica A **161**, 152–180 (1989)

[GS2] Given, J.A., Stell, G.: Scaled-particle theory and the short distance behavior of continuum percolation. J. Chem. Phys. **92**, 4433–4446 (1990)

[Gr] Grimmett, G.: Percolation. Berlin Heidelberg New York: Springer 1989

[Ha1] Hall, P.: On continuum percolation. Ann. Probab. **13**, 1250–1266 (1985)

[Ha2] Hall, P. Introduction to the theory of coverage processes. New York: Wiley 1988

[Ho] Hoeffding, W.: Probability inequalities for sums of bounded random variables. J. Am. Stat. Assoc. **58**, 13–30 (1963)

[KZ] Kesten, H., Zhang, Y.: The probability of a large finite cluster in supercritical Bernoulli percolation. Ann. Probab. **18**, 537–555 (1990)

[KS] Kunz, H., Souillard, B.: Essential singularity in percolation problems and asymptotic behavior of the cluster size distribution. J. Stat. Phys. **19**, 77–106 (1978)

[LGP] Lifshits, I.M., Gredeskul, S.A., Pastur, L.A.: Introduction to the theory of disordered systems. New York: Wiley 1988

[MMS] Men'shikov, M.V., Molchanov, S.A., Sidorenko, A.F.: Percolation theory and some applications (in Russian). Itogi Nauki Tekh. Ser. Teor. Veroyatn., Mat. Stat., Teor. Kibern., **24**, 53–110 (1986) J. Sov. Math. **42**, 1766–1810 (1988)

[Pe] Penrose, M.: On a continuum percolation model. Adv. Appl. Probab. **23**, 536–556 (1991)

[Ro] Roy, R.: The Russo-Seymour-Welsh theorem and the equality of critical densities and the "dual" critical densities for continuum percolation on $\mathbb{R}^2$. Ann. Probab. **18**, 1563–1575 (1990)

[SX] Stell, G., Xu, J.: An analytic treatment of percolation in simple fluids. J. Chem. Phys. **89**, 1101–1111 (1988)

[Ta1] Taylor, J.E.: Existence and structure of solutions to a class of nonelliptic variational problems. Symp. Math. **14**, 499–508 (19xx)

[Ta2] Taylor, J.E.: Unique structure of solutions to a class of nonelliptic variational problems. Proc. Symp. Pure Math. **27**, 419–427 (1975)

[ZS] Zuev, S.A., Sidorenko, A.F.: Continuous models of percolation theory I, II. Theor. Math. Phys. **62**, 72–86, 253–262 (1985)