GIANT COMPONENT FOR THE SUPERCRITICAL LEVEL-SET PERCOLATION OF THE GAUSSIAN FREE FIELD ON REGULAR EXPANDER GRAPHS

JIŘÍ ČERNÝ

ABSTRACT. We consider the zero-average Gaussian free field on a certain class of finite d-regular graphs for fixed $d \geq 3$. This class includes d-regular expanders of large girth and typical realisations of random d-regular graphs. We show that the level set of the zero-average Gaussian free field above level h has a giant component in the whole supercritical phase, that is for all $h < h_\star$, with probability tending to one as the size of the graphs tends to infinity. In addition, we show that this component is unique. This significantly improves the result of [4], where it was shown that a linear fraction of vertices is in mesoscopic components if $h < h_\star$, and together with the description of the subcritical phase from [4] establishes a fully-fledged percolation phase transition for the model.

1. INTRODUCTION

Level-set percolation of the Gaussian free field is a significant representative of percolation models with long range correlations. It has attracted attention for a long time, dating back to [24, 21, 7]. In the last decade, it has been subject to intensive research after a non-trivial percolation phase transition has been identified for this model on $\mathbb{Z}^d$ in [27], see for instance [2, 9, 12, 29]. Only very recently, in the remarkable paper [15], it has been shown that this phase transition is sharp, and, rather amazingly, the critical exponents have been identified for a related model of the level-set percolation on the cable system of $\mathbb{Z}^d$ in [13].

In coherence with a long line of past percolation results, it is natural to consider the level-set percolation of the Gaussian free field on finite graphs as well. In this context, [1] introduced a suitable version of the Gaussian free field which can be defined on finite graphs, the *zero-average* Gaussian free field, and studied its properties on discrete tori of growing side length in dimension $d \geq 3$. In [4] (with preparatory steps conducted in [3]), A. Abächerli and the author initiated the investigations of the zero-average Gaussian free field on a certain class of finite locally tree-like d-regular graphs. The present paper continues these investigations.

In [4] it has been shown that the level-set percolation of the zero-average Gaussian field on this class of graphs exhibits a percolation phase transition at a critical level $h_\star$ in the following sense: With probability tending to one as the size N of the graphs tends to infinity, whenever $h > h_\star$, the level set of the zero-average Gaussian free field above level h does not contain any connected component of size larger than $C_h \log N$, and, on the contrary, whenever $h < h_\star$, a linear fraction of the vertices is contained in 'mesoscopic' connected components of the level set above level h, that is in components having a size of at least a small fractional power of N. The critical level $h_\star$ agrees with the percolation threshold of the level set percolation of the usual Gaussian free field on a d-regular tree which was identified in [28].

Date: September 30, 2021.

Key words and phrases. Gaussian free field; level-set percolation; giant component; expander graphs.

In the subcritical phase, $h > h_\star$, this description of the behaviour of the level set is satisfactory. On the other hand, in the supercritical phase, $h < h_\star$, it leaves open the question whether the mesoscopic components form a giant component, that is a component of size of order N, cf. [4, Remark 5.7].

This is a natural question since for other probabilistic models on essentially the same class of graphs the emergence of the unique giant component has been shown in the corresponding supercritical phases. The first example is the Bernoulli bond percolation on d-regular expanders of large girth considered in [5] (for more recent results, see [18]). A second example is the percolation of the vacant set left by the simple random walk on the same class of graphs as considered here and in [4], see [8]. In particular the latter result gives a strong indication that a giant component should emerge also in the supercritical phase of the level-set percolation, as the two models share many common features, like similar decay of correlations.

We answer this question affirmatively. To state our results we first recall the setting of [4]. We fix $d \geq 3$ and assume that $(\mathcal{G}_n)_{n\geq 1}$ is a sequence of graphs satisfying the following conditions.

Assumption 1.1. There exist $\alpha \in (0,1)$, $\beta > 0$, and an increasing sequence of positive integers $(N_n)_{n\geq 1}$ with $\lim_{n\to\infty} N_n = \infty$ such that for all $n \geq 1$:

(a) $\mathcal{G}_n = (\mathcal{V}_n, \mathcal{E}_n)$ is a simple connected graph with N_n vertices which is d-regular (that is all its vertices have degree d).
(b) For all $x \in \mathcal{V}_n$ there is at most one cycle in the ball of radius $\lfloor \alpha \log_{d-1}(N_n) \rfloor$ around x.
(c) The spectral gap of $\mathcal{G}_n$, denoted by $\lambda_{\mathcal{G}_n}$, satisfies $\lambda_{\mathcal{G}_n} \geq \beta$.

We refer to [4] for a more detailed discussion of these assumptions, but recall that they are satisfied for two important classes of graphs: (a) random d-regular graphs, (b) d-regular expanders of large girth. We also remark that assumptions very similar to ours were used in recent studies of quantum ergodicity on graphs, and in related studies of percolation of the level sets of the adjacency eigenvectors (see for instance [16, 6]).

On $\mathcal{G}_n$ we consider the zero-average Gaussian free field $\Psi_{\mathcal{G}_n} = (\Psi_{\mathcal{G}_n}(x) : x \in \mathcal{V}_n)$ which is a centred Gaussian process on $\mathcal{V}_n$ whose law is determined by its covariance function

$$E[\Psi_{\mathcal{G}_n}(x)\Psi_{\mathcal{G}_n}(y)] = G_{\mathcal{G}_n}(x,y) \qquad \text{for all } x, y \in \mathcal{V}_n, \tag{1.1}$$

where $G_{\mathcal{G}_n}(\cdot,\cdot)$ is the zero-average Green function on $\mathcal{G}_n$ (see (2.3), (2.5) for its definition).

The zero-average Gaussian free field is a natural version of the Gaussian free field for finite graphs. However, due to the zero-average property, namely

$$\sum_{x\in\mathcal{V}_n} \Psi_{\mathcal{G}_n}(x) = 0, \qquad \text{a.s.}, \tag{1.2}$$

it comes with some peculiarities like the lack of an FKG-inequality and of the domain Markov property which are instrumental when studying Gaussian free field on infinite graphs, cf. [4, Section 2.2].

We analyse the properties of the level sets of $\Psi_{\mathcal{G}_n}$ above level $h \in \mathbb{R}$, that is of

$$E^{\geq h}(\Psi_{\mathcal{G}_n}) := \{x \in \mathcal{V}_n : \Psi_{\mathcal{G}_n}(x) \geq h\}. \tag{1.3}$$

In particular, we are interested in the sizes of its largest and second largest connected components $\mathcal{C}^{\mathcal{G}_n,h}_{\max}$ and $\mathcal{C}^{\mathcal{G}_n,h}_{\mathrm{sec}}$.

For our investigations it is important that the field $\Psi_{\mathcal{G}_n}$ is locally well approximated by the Gaussian free field $\varphi_{\mathbb{T}_d} = (\varphi_{\mathbb{T}_d}(x) : x \in \mathbb{T}_d)$ on the infinite rooted d-regular tree

$\mathbb{T}_d$ (see the paragraph containing (2.6) for the definition). For now, we only define its percolation function

$$\eta(h) \coloneqq P\big(|\mathcal{C}_{\mathrm{o}}^h| = \infty\big), \tag{1.4}$$

where $\mathcal{C}_{\mathrm{o}}^h$ is the connected component of the set $E^{\geq h}(\varphi_{\mathbb{T}_d}) \coloneqq \{x \in \mathbb{T}_d : \varphi_{\mathbb{T}_d}(x) \geq h\}$ containing the root $\mathrm{o} \in \mathbb{T}_d$, and we set

$$h_\star \coloneqq \inf\big\{h \in \mathbb{R} : \eta(h) = 0\big\}, \tag{1.5}$$

to be its critical value. From [28] it is known that $h_\star$ is positive and finite.

We can now state our main result.

Theorem 1.2. *If $h < h_\star$, then for every sequence of graphs $(\mathcal{G}_n)_{n\geq 1}$ satisfying Assumption 1.1 and every $\delta > 0$*

$$\lim_{n\to\infty} P\Big(\frac{|\mathcal{C}_{\max}^{\mathcal{G}_n,h}|}{N_n} \in (\eta(h) - \delta, \eta(h) + \delta) \text{ and } \big|\mathcal{C}_{\sec}^{\mathcal{G}_n,h}\big| \leq \delta N_n\Big) = 1. \tag{1.6}$$

Theorem 1.2 confirms the emergence of the giant component in the supercritical phase of the model, gives its typical size, and provides its uniqueness. Together with the description of the subcritical behaviour from Theorem 4.1 of [4] (which states that $\lim_{n\to\infty} P(|\mathcal{C}_{\max}^{\mathcal{G}_n,h}| \leq C_h \log N_n) = 1$ for $h > h_\star$) it establishes a fully-fledged percolation phase transition for the level-set percolation of the zero-average Gaussian free field on $\mathcal{G}_n$.

Assumption 1.1 of Theorem 1.2 can be weakened slightly, as explained in Remark 9.1 at the end of the paper. However, for these weakened assumptions we do not have the corresponding subcritical description. We thus prefer to work in the same setting as in [4].

Similarly as in [5, 8, 18], we use a sprinkling technique to show that the mesoscopic components (that we know to exist due to [4, Theorem 5.1]) indeed form a giant component. Making the sprinkling work in the settings of dependent percolation is however rather challenging, as was already observed in [8], in the context of the vacant set left by a random walk. In the context of Gaussian free field, sprinkling was previously used in [14], to show the existence of an infinite connected component of the supercritical level set on $\mathbb{Z}^d$ when $d \to \infty$. The diverging dimension is important for the arguments therein, since the correlations of the field decay with the dimension (as d^{-1} for the neighbouring vertices). Several sprinkling steps are also used in the recent paper [15], which proves the sharpness of the phase transition for the level set of the Gaussian free field on $\mathbb{Z}^d$, $d \geq 3$. Note also that the results of [15] can be combined with [1] to show the existence of the giant component for the supercritical level set of the zero-average Gaussian free field on a large discrete torus (cf. [15, Section 1.2]).

Very recently, a result similar to Theorem 1.2 was proved by G. Conchon-Kerjan in [10]. Namely, assuming that $\mathcal{G}_n$ is a *uniformly random* d-regular graph with n vertices, he shows that under the *annealed* probability measure P_{ann} (that is taking into account the randomness of the graph and the field),

$$\lim_{n\to\infty} P_{\mathrm{ann}}\Big(\frac{|\mathcal{C}_{\max}^{\mathcal{G}_n,h}|}{n} \in (\eta(h) - \delta, \eta(h) + \delta) \text{ and } \frac{|\mathcal{C}_{\sec}^{\mathcal{G}_n,h}|}{\log n} \in (c, c')\Big) = 1, \tag{1.7}$$

for every $\delta > 0$ and some $0 < c < c' < \infty$, and further gives a rather detailed description of the geometry of $\mathcal{C}_{\max}^{\mathcal{G}_n,h}$. The arguments of [10] are completely different from the ones used in this paper and rely strongly on the assumption that $\mathcal{G}_n$ is a random regular graph, and thus can be revealed progressively using the usual pairing construction. As discussed above, this assumption is stronger than our Assumption 1.1.

The estimate on $|\mathcal{C}_{\rm sec}^{\mathcal{G}_n,h}|$ in (1.7) is essentially optimal and better than our estimate on the same quantity in (1.6). Incidentally, this resembles the previously known results for the vacant set of random walk: On the same class of graphs as here [8, Theorem 1.3], only shows that the second largest connected component of the vacant set is $o(N_n)$, while on random regular graphs it can be proved that it is $O(\log N_n)$, by combining the techniques of [11, 25]. Note also that [18] shows that for any $\omega < 1$ there are regular expanders with an arbitrarily large girth such that the second connected component of the Bernoulli bond percolation has size growing at least as N_n^{ω}, which indicates that the exact asymptotic behaviour of $|\mathcal{C}_{\rm sec}^{\mathcal{G}_n,h}|$ might be a delicate issue.

Let us now comment on the proof of Theorem 1.2. To explain its main ideas, it is useful to discuss the sprinkling construction for the Bernoulli percolation from [5] first. This construction relies on the fact that a percolation configuration $(\omega^p(x))_{x\in\mathcal{V}_n} \in \{0,1\}^{\mathcal{V}_n}$ at level p can be obtained as the maximum of two independent Bernoulli configurations ω^{p_1} and ω^{p_2} with the levels p_1, p_2 satisfying $1-p=(1-p_1)(1-p_2)$. For the techniques of [5] to work, it is very important that (a) ω^{p_2} is independent of ω^{p_1}, (b) ω^{p_2} is a Bernoulli percolation, that is the random variables $(\omega^{p_2}(x))_{x\in\mathcal{V}_n}$ are independent, and (c) that the maximum function is monotonous, in particular $\{x : \omega^p(x) = 1\} \supset \{x : \omega^{p_1}(x) = 1\}$. While (c) is important for the sprinkling not to destroy the mesoscopic components of ω^{p_1}, (a) and (b) play a key role in estimating the probability of a certain bad event which needs to be much smaller than e^{-cN_n/m_n}, with m_n denoting the minimal size of mesoscopic components (cf. proof of Proposition 3.1 in [5]). In [5], the proof of this estimate is just a simple large deviation argument for i.i.d. Bernoulli random variables. Unfortunately, a corresponding estimate is mostly simply not true in the setting of correlated Gaussian fields.

Before describing our approach, let us very quickly mention two natural ideas how to adapt the sprinkling construction of [5] to the zero-average Gaussian free field which, unfortunately, cannot easily be converted into a rigorous proof, mostly due to the lack of independence. The first one is to use the existence of many mesoscopic components at a level $h' \in (h, h_\star)$ and prove that by lowering the level from h' to h those components merge. This preserves the monotonicity, that is the point (c) from the last paragraph, but completely destroys the independence (a) and (b), making the above mentioned estimate on the bad event essentially impossible to prove. The second one is to write $\Psi^{\mathcal{G}_n}$ as a linear combination $\sqrt{1-t^2}\Psi'_{\mathcal{G}_n} + t\Psi''_{\mathcal{G}_n}$ (with a small t) of its independent copies $\Psi'_{\mathcal{G}_n}, \Psi''_{\mathcal{G}_n}$. Here, the monotonicity (c) is lost (but probably could be salvaged by some technical work), (a) is preserved, but the correlations of $\Psi''_{\mathcal{G}_n}$ make the estimate on the bad event fail again. Remark also that the zero average property (1.2) excludes writing $\Psi_{\mathcal{G}_n}$ as a sum $X+Y$ of two non-trivial independent fields X, Y such that $Y \geq 0$, or as $\max(X,Y)$ for X, Y independent; both of these decompositions would be desirable for the monotonicity (c).

In this paper we thus develop a new decomposition of the zero-average Gaussian free field which provides enough independence to be useful in a sprinkling argument and which is of independent interest, see Section 3. It is inspired by a similar decomposition of the (usual) Gaussian free field on $\mathbb{Z}^d$ from [15]. Using this decomposition we will write $\Psi_{\mathcal{G}_n}$ as a sum of two independent components $\Psi_{\mathcal{G}_n} = \Psi^1_{\mathcal{G}_n} + \Psi^2_{\mathcal{G}_n}$, where

$$\Psi^2_{\mathcal{G}_n}(x) \coloneqq t_n\Big(Z_0(x) - N^{-1}\sum_{y\in\mathcal{V}_n} Z_0(y)\Big), \qquad x \in \mathcal{V}_n, \tag{1.8}$$

for some family $(Z_0(x))_{x\in\mathcal{V}_n}$ of i.i.d. Gaussian random variables. Since the field $\Psi^2_{\mathcal{G}_n}$ is essentially an i.i.d. field, up to the zero-average property, this writing preserves (a) and (b) from the above discussion, but gives up on the monotonicity (c). We will deal with the non-monotonicity issue by taking t_n small and by restricting the connected components of the level set to certain subgraphs of $\mathcal{G}_n$ where Z_0 is not too small. These arguments are relatively straightforward and are given in Sections 8, 9.

The decomposition, however, introduces a new problem: the field $\Psi^1_{\mathcal{G}_n}$ is not longer a zero-average Gaussian free field and we thus do not know that it has many mesoscopic components in the whole supercritical phase $h < h_\star$. To show this we will use a perturbative argument. More precisely, we use the fact that $\Psi^1_{\mathcal{G}_n} = \Psi_{\mathcal{G}_n} - \Psi^2_{\mathcal{G}_n}$ (with $\Psi_{\mathcal{G}_n}$ and $\Psi^2_{\mathcal{G}_n}$ dependent!) and that $\Psi_{\mathcal{G}_n}$ has many mesoscopic components at any level $h' \in (h, h_\star)$, by [4, Theorem 5.1]. We then show that, typically, these components are robust to certain perturbations and are thus not destroyed by subtracting $\Psi^2_{\mathcal{G}_n}$. The proof of the existence of the robust components is based on multi-type branching process arguments developed in [4]. Its details are given in Sections 5–7. On the way, in Section 4, we use the decomposition of $\Psi_{\mathcal{G}_n}$ from Section 3 to construct a new coupling of $\Psi_{\mathcal{G}_n}$ and $\varphi_{\mathbb{T}_d}$.

2. Preliminaries

In this section we introduce the notation and recall few useful facts that we use throughout the paper. For an arbitrary locally-finite, simple, non-oriented graph G we denote by $V(G)$ and $E(G)$ the sets of its vertices and edges. For $x, y \in V(G)$, we write $x \sim y$ when $(x, y) \in E(G)$, $d_G(\cdot,\cdot)$ denotes their graph distance, and $\deg_G(x)$ the degree of x in G. For any $U \subset V(G)$, $|U|$ stands for its cardinality, and $\partial_G U := \{y \in V(G) \setminus U : \exists x \in U \text{ s.t. } x \sim y\}$ denotes its (outer vertex) boundary. For any $r \ge 0$ and $x \in V(G)$ we define the ball and sphere of radius r around x to be $B_G(x, r) := \{y \in V(G) : d_G(x, y) \le r\}$ and $S_G(x, r) := \{y \in V(G) : d_G(x, y) = r\}$.

We write $\bar{P}^G_x$ for the canonical law on $V(G)^{\mathbb{N}}$ of the *lazy* simple random walk $X = (X_k)_{k\ge0}$ on G starting at $x \in V(G)$, and $\bar{E}^G_x$ for the corresponding expectation. Under $\bar{P}^G_x$, the transition probabilities of X are given by

$$\bar{P}^G_x(X_{k+1} = z \mid X_k = y) = \begin{cases} \frac{1}{2}, & \text{if } z = y, \\ \frac{1}{2\deg_G(x)}, & \text{if } (z, y) \in E(G). \end{cases} \tag{2.1}$$

If G is a finite connected graph, we denote the unique invariant distribution of X by

$$\pi_G(x) := \frac{\deg_G(x)}{2|E(G)|}. \tag{2.2}$$

The zero-average Green function $\bar{G}_G$ of X and its density are given by

$$\begin{aligned} \bar{G}_G(x, y) &:= \sum_{k\ge0} (\bar{P}^G_x(X_k = y) - \pi_G(y)), \qquad x, y \in V(G), \\ \bar{g}_G(x, y) &:= (\deg_G(y))^{-1} \bar{G}_G(x, y). \end{aligned} \tag{2.3}$$

It is easy to check from the reversibility of the random walk that $\bar{g}_G(x, y)$ is a symmetric function. Zero-average Gaussian free field on G is a centred Gaussian process $(\Psi_G(x))_{x\in V(G)}$ whose law is determined by its covariance function

$$E(\Psi_G(x)\Psi_G(y)) = C_0 \bar{g}_G(x, y) \qquad \text{for all } x, y \in V(G). \tag{2.4}$$

The constant C_0 only influences the scaling of the field and is introduced for convenience. If G is d-regular, as in Assumption 1.1(a), it customary to take $C_0 = d/2$. With this choice,

$$\bar{g}_G(x, y) = \tfrac{1}{2}\bar{G}_G(x, y) = G_G(x, y), \tag{2.5}$$

where $G_G(\cdot,\cdot)$ is the zero-average Green function of the usual continuous-time random walk on G (the factor $\frac{1}{2}$ disappears due to the laziness), and thus the covariance from (2.4) agrees with the one used in [4], cf. (2.17) therein.

For any field f on G we denote by $E^{\geq h}(f) \coloneqq \{x \in V(G) : f(x) \geq h\}$ its level set above the level $h \in \mathbb{R}$.

We use $\mathbb{T}_d$ to denote the d-regular infinite tree with root o. For every $x \in V(\mathbb{T}_d)$ we denote by $\mathrm{desc}(x)$ the set of its direct descendants, and for $x \in V(\mathbb{T}_d) \setminus \{\mathrm{o}\}$ we use $\mathrm{anc}(x)$ to denote the direct ancestor of x in $\mathbb{T}_d$. The Gaussian free field on $\mathbb{T}_d$ is a centred Gaussian process $(\varphi_{\mathbb{T}_d}(x))_{x \in V(\mathbb{T}_d)}$ whose distribution is determined by

$$E(\varphi_{\mathbb{T}_d}(x)\varphi_{\mathbb{T}_d}(y)) = g_{\mathbb{T}_d}(x, y) \qquad \text{for all } x, y \in V(\mathbb{T}_d), \tag{2.6}$$

where $g_{\mathbb{T}_d}$ is the Green function of the (usual discrete-time) simple random walk on $\mathbb{T}_d$.

As mentioned earlier, we consider for fixed $d \geq 3$ the d-regular graphs $(\mathcal{G}_n)_{n\geq 1}$ satisfying Assumption 1.1, and we abbreviate $\mathcal{V}_n = V(\mathcal{G}_n)$, $\mathcal{E}_n = E(\mathcal{G}_n)$. For $r \geq 0$, we say that a vertex $x \in \mathcal{V}_n$ is r-treelike, if there is no cycle in $B_{\mathcal{G}_n}(x, r)$. If x is r-treelike, then we fix a graph isomorphism $\rho_{x,r} : B_{\mathcal{G}_n}(x, r) \to B_{\mathbb{T}_d}(\mathrm{o}, r)$ such that $\rho_{x,r}(x) = \mathrm{o}$.

We recall from [4, Proposition 2.2] that there is $\varepsilon \in (0, 1)$ such that for every $n \geq 1$ and $x, y \in \mathcal{V}_n$,

$$\bar{g}_{\mathcal{G}_n}(x, y) \leq C(d-1)^{-d_{\mathcal{G}_n}(x,y)} + N_n^{-\varepsilon}. \tag{2.7}$$

Finally, Assumption 1.1(a,c) imply (by Cheeger's inequality, for the argument see e.g. [8, (2.11)]) the uniform isoperimetric inequality for the sequence $\mathcal{G}_n$:

$$\begin{aligned}&\text{There is } \beta' > 0 \text{ such that } \tfrac{|\partial_{\mathcal{G}_n} A|}{|A|} \geq \beta' \text{ for}\\ &\text{all } n \geq 1 \text{ and } A \subset \mathcal{V}_n \text{ with } |A| \leq |\mathcal{V}_n|/2.\end{aligned} \tag{2.8}$$

We use c, c', C, ... to denote positive constants with values changing from place to place and which only depend on the degree d and the constants α and β from Assumption 1.1.

3. Decomposition of the field

The goal of this section is to construct a decomposition of the zero-average Gaussian free field into independent components. We believe that this decomposition is of independent interest. It is the main ingredient of our sprinkling construction, as described in the introduction, but also will be used in Section 4 to construct a new coupling of $\Psi_{\mathcal{G}_n}$ and $\varphi_{\mathbb{T}_d}$. The construction of this decomposition is inspired by a similar decomposition for the usual Gaussian free field on $\mathbb{Z}^d$ from [15], see Lemma 3.1 therein. However, the zero-average property introduces certain complications making the decomposition less straightforward.

For sake of generality, we consider an arbitrary finite, simple, non-oriented, connected graph $G = (V(G), E(G))$ in this section. That is we do not require that Assumption 1.1 holds.

Recall the definition of Ψ_G from (2.4). To introduce its decomposition we need more notation. We write $\tilde{G}$ for the graph obtained from G by adding an additional vertex to the middle of every edge of G, formally $\tilde{G} = (V(\tilde{G}), E(\tilde{G}))$ with

$$V(\tilde{G}) \coloneqq V(G) \cup E(G), \tag{3.1}$$

$$E(\tilde{G}) \coloneqq \{(x, e) : x \in V(G), e \in E(G), e \ni x\}. \tag{3.2}$$

Observe that $\tilde{G}$ is a bipartite graph. For $\tilde{x} \in V(\tilde{G})$, let

$$\tilde{\pi}_G(\tilde{x}) \coloneqq \deg_{\tilde{G}}(\tilde{x}) = \begin{cases} \deg_G(\tilde{x}), & \text{if } \tilde{x} \in V(G),\\ 2, & \text{if } \tilde{x} \in E(G).\end{cases} \tag{3.3}$$

Let

$$\tilde{Q}_G(\tilde{x}, \tilde{y}) = \mathbf{1}_{(\tilde{x},\tilde{y}) \in E(\tilde{G})}/\tilde{\pi}_G(\tilde{x}), \qquad \tilde{x}, \tilde{y} \in V(\tilde{G}), \tag{3.4}$$

be the transition matrix of the *usual* simple random walk on $\tilde{G}$. $\tilde{Q}_G$ acts on the space $\ell^2(\tilde{\pi}_G)$ by $\tilde{Q}_G f(\tilde{x}) = \sum_{\tilde{y}\in V(\tilde{G})} \tilde{Q}_G(\tilde{x},\tilde{y}) f(\tilde{y})$. Due to the reversibility, $\tilde{Q}_G$ is a self-adjoint operator on $\ell^2(\tilde{\pi}_G)$. Since $\tilde{G}$ is connected and bipartite, $\tilde{Q}_G$ has simple eigenvalues 1 and -1 with respective eigenfunctions $\mathbf{1}$ and w, where $w(\tilde{x}) = 1$ if $\tilde{x} \in V(G)$ and $w(\tilde{x}) = -1$ if $\tilde{x} \in E(G)$. Denoting by $\|\cdot\|_{\tilde{\pi}_G}$ the $\ell^2(\tilde{\pi}_G)$-norm and by $\langle \cdot,\cdot\rangle_{\tilde{\pi}_G}$ the corresponding scalar product, we have

$$\|\mathbf{1}\|^2_{\tilde{\pi}_G} = \|w\|^2_{\tilde{\pi}_G} = \sum_{\tilde{x}\in V(\tilde{G})} \tilde{\pi}_G(\tilde{x}) = 4|E(G)|, \qquad \langle \mathbf{1}, w\rangle_{\tilde{\pi}_G} = 0. \tag{3.5}$$

Let Π_G be the orthogonal projection (in $\ell^2(\tilde{\pi}_G)$) on $\operatorname{span}(\mathbf{1}, w)$. Then $\Pi_G^2 = \Pi_G$ and Π_G is self-adjoint in $\ell^2(\tilde{\pi}_G)$. Moreover, since $\mathbf{1}$ and w are eigenvectors of $\tilde{Q}_G$, the operators Π_G and $\tilde{Q}_G$ commute, and thus, for every $k \in \mathbb{N}$, the operators $\tilde{Q}_G^k \Pi_G = \Pi_G \tilde{Q}_G^k$ and $(\mathrm{Id} - \Pi_G)\tilde{Q}_G^k$ are self-adjoint as well (here Id stands for the identity operator). For later use we observe that for $f \in \ell^2(\tilde{\pi}_G)$ and $y \in V(G)$ (so that $w(y) = 1$),

$$\begin{aligned}(\Pi_G f)(y) &= \|\mathbf{1}\|^{-2}_{\tilde{\pi}_G} \langle f, \mathbf{1}\rangle_{\tilde{\pi}_G} \mathbf{1}(y) + \|w\|^{-2}_{\tilde{\pi}_G}\langle f, w\rangle_{\tilde{\pi}_G} w(y) \\ &\overset{(3.5)}{=} \frac{1}{4|E(G)|} \sum_{\tilde{x}\in V(\tilde{G})} f(\tilde{x})(1 + w(\tilde{x}))\tilde{\pi}_G(\tilde{x}) \\ &\overset{(3.3)}{=} \frac{1}{2|E(G)|} \sum_{x\in V(G)} f(x)\deg_G(x) \overset{(2.2)}{=} \sum_{x\in V(G)} f(x)\pi_G(x).\end{aligned} \tag{3.6}$$

With C_0 as in (2.4), let $(Z_k(\tilde{x}))_{k\in\mathbb{N},\tilde{x}\in V(\tilde{G})}$ be independent centred Gaussian random variables with

$$\operatorname{Var} Z_k(\tilde{x}) = C_0/\tilde{\pi}_G(\tilde{x}), \tag{3.7}$$

defined on a probability space $(\Omega, \mathcal{A}, P)$. Finally, for $x \in V(G)$, set

$$\xi^k_G(x) := \sum_{\tilde{y}\in V(\tilde{G})} ((\mathrm{Id} - \Pi_G)\tilde{Q}^k_G)(x,\tilde{y}) Z_k(\tilde{y}) \overset{\text{(not.)}}{=} \big((\mathrm{Id} - \Pi_G)\tilde{Q}^k_G Z_k\big)(x), \tag{3.8}$$

$$\tilde{\Psi}_G(x) := \sum_{k\in\mathbb{N}} \xi^k_G(x). \tag{3.9}$$

We now show that (3.9) provides the desired decomposition of Ψ_G.

Proposition 3.1. *The series on the right-hand side of* (3.9) *converges in $L^2(P)$ and P-a.s., and the law of $\tilde{\Psi}_G$ agrees with the law of Ψ_G, that is $\tilde{\Psi}_G$ is a zero-average Gaussian free field on G.*

Proof. We start by computing the covariances of the fields ξ^k_G. Using the independence of $Z_k(\tilde{x})$'s, the self-adjointness of $(\mathrm{Id} - \Pi_G)\tilde{Q}^k_G$ and the fact that Π_G and $\tilde{Q}_G$ commute, for every $x, y \in V(G)$,

$$\begin{aligned}\operatorname{Cov}(\xi^k_G(x), \xi^k_G(y)) &= \sum_{\tilde{z}\in V(\tilde{G})} ((\mathrm{Id} - \Pi_G)\tilde{Q}^k_G)(x,\tilde{z})((\mathrm{Id} - \Pi_G)\tilde{Q}^k_G)(y,\tilde{z}) \frac{C_0}{\tilde{\pi}_G(\tilde{z})} \\ &= \frac{C_0}{\tilde{\pi}_G(y)} \sum_{\tilde{z}\in V(\tilde{G})} ((\mathrm{Id} - \Pi_G)\tilde{Q}^k_G)(x,\tilde{z})((\mathrm{Id} - \Pi_G)\tilde{Q}^k_G)(\tilde{z}, y) \\ &= \frac{C_0}{\tilde{\pi}_G(y)} ((\mathrm{Id} - \Pi_G)\tilde{Q}^{2k}_G)(x,y).\end{aligned} \tag{3.10}$$

To compute the terms involving Π_G, let $(v_i)_{i=1,\dots,|V(\tilde{G})|}$ be an orthonormal basis of $\ell^2(\tilde{\pi}_G)$ composed by the eigenvectors of $\tilde{Q}_G$ such that $v_1 = \mathbf{1}/\|\mathbf{1}\|_{\tilde{\pi}_G}$ and $v_2 = w/\|w\|_{\tilde{\pi}_G}$, and let

$(\lambda_i)_{i=1,\dots,|V(\tilde G)|}$ be the corresponding eigenvalues. Observe $\Pi_G v_i = 0$ for $i \geq 3$ and that $\tilde Q_G^{2k}(f) = \sum_{i=1}^{|V(\tilde G)|} \langle v_i, f\rangle_{\tilde\pi_G} \lambda_i^{2k} v_i$. Hence, for $x, y \in V(G)$,

$$
\begin{aligned}
(\Pi_G \tilde Q_G^{2k})(x,y) &= \frac{1}{\tilde\pi_G(x)} \langle \mathbf{1}_x, (\Pi_G \tilde Q_G^{2k})\mathbf{1}_y\rangle_{\tilde\pi_G} \\
&= \frac{1}{\tilde\pi_G(x)} \sum_{i=1}^{|V(\tilde G)|} \big\langle \mathbf{1}_x, \langle v_i, \mathbf{1}_y\rangle_{\tilde\pi_G} \lambda_i^{2k} \Pi_G v_i \big\rangle_{\tilde\pi_G} \\
&= \langle \mathbf{1}_y, v_1\rangle_{\tilde\pi_G} v_1(x) + \langle \mathbf{1}_y, v_2\rangle_{\tilde\pi_G} v_2(x) \\
&= (\Pi_G \mathbf{1}_y)(x) \overset{(3.6)}{=} \pi_G(y).
\end{aligned}
\tag{3.11}
$$

Hence, by (3.10), (3.11) and (3.3),

$$
\mathrm{Cov}(\xi_G^k(x), \xi_G^k(y)) = \frac{C_0}{\deg_G(y)} (\tilde Q_G^{2k}(x,y) - \pi_G(y)), \qquad x, y \in V(G), k \in \mathbb{N}. \tag{3.12}
$$

The matrix $\tilde Q_G^{2k}$ restricted to $V(G)$ agrees with the k-step transition matrix of the lazy random walk on G, that is $Q_G^{2k}(x,y) = \bar P_x^G(X_k = y)$. In particular, due to standard convergence results for Markov chains, $|Q_G^{2k}(x,y) - \pi_G(y)| \leq Ce^{-ck}$ for all $x, y \in V(G)$, and thus also $|\mathrm{Cov}(\xi_G^k(x), \xi_G^k(y))| \leq Ce^{-ck}$. This implies that the series in (3.9) converges in $L^2(P)$. The a.s. convergence is then standard, e.g. using Kolmogorov's maximal inequality. Finally, (3.12) implies that

$$
\begin{aligned}
\mathrm{Cov}\left(\tilde\Psi_G(x), \tilde\Psi_G(y)\right) &= \sum_{k\in\mathbb{N}} \mathrm{Cov}\left(\xi_G^k(x), \xi_G^k(y)\right) \\
&= \frac{C_0}{\deg_G(y)} \sum_{k\in\mathbb{N}} \left(\bar P_x^G(X_k = y) - \pi_G(y)\right),
\end{aligned}
\tag{3.13}
$$

which agrees with the covariance of Ψ_G from (2.4). Since $\tilde\Psi_G$ is obviously a centred Gaussian field, this completes the proof. □

4. Coupling with a tree

We now come back to our original setting of Assumption 1.1 and construct, in Proposition 4.1 below, a coupling between the zero-average Gaussian free field $\Psi_{\mathcal{G}_n}$ and the Gaussian free field $\varphi_{\mathbb{T}_d}$. A similar coupling is provided by Theorem 3.1 of [4]. However, our Proposition 4.1 has several advantages: First, it has a much simpler proof which is based on the decomposition from Section 3. Second, in its proof we also couple the underlying Z-fields (cf. Remark 4.2 below) which will be important later. And third, in contrast to [4], we use two independent fields $\varphi_{\mathbb{T}_d}$, $\varphi'_{\mathbb{T}_d}$ in its statement; this will simplify the application of the coupling in the second moment computation in the proof of Proposition 7.1 below.

For the statement recall from Section 2 that $\rho_{x,r}$ denotes a fixed isomorphism of $B_{\mathcal{G}_n}(x,r)$ and $B_{\mathbb{T}_d}(\mathrm{o}, r)$, if $x \in \mathcal{V}_n$ is r-treelike.

Proposition 4.1. *There are $c, C \in (0,\infty)$ such that for all $n, r \in \mathbb{N}$, and for all $x, x' \in \mathcal{V}_n$ which are $2r$-treelike and satisfy $B_{\mathcal{G}_n}(x, 2r) \cap B_{\mathcal{G}_n}(x', 2r) = \emptyset$ there exists a coupling $\mathbb{Q}_n^{x,x'}$ of $\Psi_{\mathcal{G}_n}$ and two independent Gaussian free fields $\varphi_{\mathbb{T}_d}$, $\varphi'_{\mathbb{T}_d}$ such that for all $\varepsilon > 0$*

$$
\begin{aligned}
&\mathbb{Q}_n^{x,x'}\left[\max\{D(x,r), D(x',r)\} > \varepsilon\right] \\
&\qquad \leq Cd(d-1)^r \Big(\exp\Big(-\frac{\varepsilon^2 e^{cr}}{18}\Big) + \exp\Big(-\frac{\varepsilon^2 N_n}{9(r+1)}\Big)\Big),
\end{aligned}
\tag{4.1}
$$

where

$$D(x,r) := \max_{y\in B_{\mathcal{G}_n}(x,r)} \big|\Psi_{\mathcal{G}_n}(y) - \varphi_{\mathbb{T}_d}(\rho_{x,2r}(y))\big|. \tag{4.2}$$

Proof. We use the decomposition of $\Psi_{\mathcal{G}_n}$ from Section 3 and a corresponding decomposition of $\varphi_{\mathbb{T}_d}$. Using the notation of Section 3, let $\tilde{\mathcal{V}}_n := V(\tilde{\mathcal{G}}_n)$, and let $Z = (Z_k(\tilde{x}))_{k\in\mathbb{N},x\in\tilde{\mathcal{V}}_n}$ be a collection of independent Gaussian random variables on some probability space $(\Omega,\mathcal{A},\mathbb{Q}_n^{x,x'})$ with variances (cf. (3.7), (3.3), we take $C_0 = d/2$ as explained below (2.4))

$$\operatorname{Var} Z_k(\tilde{x}) = \begin{cases} \frac{1}{2}, & \text{if } \tilde{x}\in\mathcal{V}_n,\\ \frac{d}{4}, & \text{if } \tilde{x}\in\mathcal{E}_n. \end{cases} \tag{4.3}$$

Set $\xi^k_{\mathcal{G}_n}(x) := \big((\mathrm{Id}-\Pi_{\mathcal{G}_n})\tilde{Q}^k_{\mathcal{G}_n} Z_k\big)(x)$ and

$$\Psi_{\mathcal{G}_n}(x) := \sum_{k\ge 0} \xi^k_{\mathcal{G}_n}(x) = \sum_{k\ge 0} \big((\mathrm{Id}-\Pi_{\mathcal{G}_n})\tilde{Q}^k_{\mathcal{G}_n} Z_k\big)(x). \tag{4.4}$$

By Proposition 3.1, $\Psi_{\mathcal{G}_n}$ has the law of zero-average Gaussian free field.

We now introduce an analogous decomposition for the field $\varphi_{\mathbb{T}_d}$, similarly to [15, Lemma 3.1]. Let $\tilde{\mathbb{T}}_d$ be a graph obtained from $\mathbb{T}_d$ by adding a vertex to the middle of every edge, and let $\mathtt{Z} = (\mathtt{Z}_k(\tilde{x}))_{k\in\mathbb{N},\tilde{x}\in V(\tilde{\mathbb{T}}_d)}$ be a collection of independent Gaussian random variables on the same probability space $(\Omega,\mathcal{A},\mathbb{Q}_n^{x,x'})$ such that (cf. (4.3))

$$\operatorname{Var} \mathtt{Z}_k(\tilde{x}) = \begin{cases} \frac{1}{2}, & \text{if } \tilde{x}\in V(\mathbb{T}_d),\\ \frac{d}{4}, & \text{if } \tilde{x}\in V(\tilde{\mathbb{T}}_d)\setminus V(\mathbb{T}_d). \end{cases} \tag{4.5}$$

Denoting $\tilde{\mathtt{Q}}$ the transition matrix of the usual simple random walk on $\tilde{\mathbb{T}}_d$, we set $\zeta^k(x) := (\tilde{\mathtt{Q}}^k \mathtt{Z}_k)(x)$ and

$$\varphi_{\mathbb{T}_d}(x) := \sum_{k\ge 0}\zeta^k(x) = \sum_{k\ge 0}(\tilde{\mathtt{Q}}^k\mathtt{Z}_k)(x). \tag{4.6}$$

Then $\varphi_{\mathbb{T}_d}$ is a Gaussian free field on $\mathbb{T}_d$. This can be shown by a straightforward adaptation of the proof for the Gaussian free field on $\mathbb{Z}^d$ from [15] (or by adapting the proof of Proposition 3.1, leaving out all terms involving the projection Π_G). By introducing an independent copy $\mathtt{Z}' = (\mathtt{Z}'_k(x))_{k\in\mathbb{N},x\in V(\tilde{\mathbb{T}}_d)}$ of $\mathtt{Z}$, we further define the field $\varphi'_{\mathbb{T}_d}$ by a formula analogous to (4.6), with $\mathtt{Z}'$ instead of $\mathtt{Z}$.

Let $\tilde{\rho}_{x,2r} : B_{\tilde{\mathcal{G}}_n}(x,4r)\to B_{\tilde{\mathbb{T}}_d}(\mathrm{o},4r)$ be the natural extension of the isomorphism $\rho_{x,2r}$ to the balls in graphs $\tilde{\mathcal{G}}_n$ and $\tilde{\mathbb{T}}_d$. (Note that the ball $B_{\tilde{\mathcal{G}}_n}(x,4r)$ is related to $B_{\mathcal{G}_n}(x,2r)$, since in $\tilde{\mathcal{G}}_n$ there are additional vertices in the middle of every edge of $\mathcal{G}_n$.) We now require that under $\mathbb{Q}_n^{x,x'}$ the underlying fields Z, $\mathtt{Z}$, and $\mathtt{Z}'$ satisfy the following equalities while otherwise being independent:

$$\begin{aligned} \mathtt{Z}_k(\tilde{\rho}_{x,2r}(\tilde{y})) &= Z_k(\tilde{y}) && \text{for every } k\le 2r, \tilde{y}\in B_{\tilde{\mathcal{G}}_n}(x,4r),\\ \mathtt{Z}'_k(\tilde{\rho}_{x',2r}(\tilde{y})) &= Z_k(\tilde{y}) && \text{for every } k\le 2r, \tilde{y}\in B_{\tilde{\mathcal{G}}_n}(x',4r). \end{aligned} \tag{4.7}$$

Observe that this can be done without changing the distribution of Z, $\mathtt{Z}$ and $\mathtt{Z}'$, in particular the assumption $B_{\mathcal{G}_n}(x,2r)\cap B_{\mathcal{G}_n}(x',2r)=\emptyset$ is necessary for the independence of $\mathtt{Z}$ and $\mathtt{Z}'$. The assumption that x is $2r$-treelike implies that the law of the image by $\tilde{\rho}_{x,2r}$ of the random walk on $\tilde{\mathcal{G}}_n$ started in $\tilde{y}\in B_{\tilde{\mathcal{G}}_n}(x,2r)$ and stopped on exiting $B_{\tilde{\mathcal{G}}_n}(x,4r)$ agrees with the law of the random walk on $\tilde{\mathbb{T}}_d$ started in $\tilde{\rho}_{x,2r}(\tilde{y})$ and stopped on exiting $B_{\tilde{\mathbb{T}}_d}(\mathrm{o},4r)$, and that this random walk makes at least $2r$ steps before being stopped. As consequence the corresponding transition probabilities agree in the sense of

$$\begin{aligned}&\tilde{Q}^k_{\mathcal{G}_n}(\tilde{y},\tilde{y}') = \tilde{\mathsf{Q}}^k(\tilde{\rho}_{x,2r}(\tilde{y}),\tilde{\rho}_{x,2r}(\tilde{y}'))\\ &\text{for } \tilde{y}\in B_{\tilde{\mathcal{G}}_n}(x,2r), \tilde{y}'\in B_{\tilde{\mathcal{G}}_n}(x,4r), k\le 2r.\end{aligned} \tag{4.8}$$

From (4.4) and (4.6)–(4.8) it follows that for every $y \in B_{\mathcal{G}_n}(x,r)$

$$\begin{aligned}&\Psi_{\mathcal{G}_n}(y) - \varphi_{\mathbb{T}_d}(\rho_{x,2r}(y))\\ &\quad= \sum_{k>2r}\xi^k_{\mathcal{G}_n}(y) - \sum_{0\le k\le 2r}(\Pi_{\mathcal{G}_n}\tilde{Q}^k_{\mathcal{G}_n}Z_k)(y) - \sum_{k>2r}\zeta^k(\rho_{x,2r}(y)),\end{aligned} \tag{4.9}$$

and a similar equality holds when x is replaced by x' and ζ^k by $\zeta'^k := \tilde{\mathsf{Q}}^k\mathsf{Z}'_k$.

We now estimate the three sums on the right-hand side of (4.9). For the last one, we claim that there is a constant $c>0$ such that for every $\varepsilon>0$, $k_0\ge 1$, and $y\in V(\mathbb{T}_d)$,

$$\mathbb{Q}^{x,x'}_n\Big(\Big|\sum_{k>k_0}\zeta^k(y)\Big| \ge \frac{\varepsilon}{3}\Big) \le 2\exp\Big(-\frac{\varepsilon^2 e^{ck_0}}{18}\Big). \tag{4.10}$$

Indeed, observe that $\zeta^k(y)$, $k\ge 0$, are independent Gaussian random variables with $\operatorname{Var}\zeta^k(y) = \tilde{\mathsf{Q}}^{2k}(y,y)/2$ (which can be proved by a similar computation as in (3.10), recalling $C_0 = d/2$). Therefore,

$$\operatorname{Var}\Big(\sum_{k>k_0}\zeta^k(y)\Big) = \frac{1}{2}\sum_{k>k_0}\tilde{\mathsf{Q}}^{2k}(y,y) \le e^{-ck_0}, \tag{4.11}$$

where we used the fact that the lazy random walk $(X_k)_{k\ge 0}$ on $\mathbb{T}_d$ satisfies $\tilde{\mathsf{Q}}^{2k}(y,y)\le e^{-ck}$, which can easily be proved by observing that $d_{\mathbb{T}_d}(\mathrm{o},X_n)$ is a random walk on $\mathbb{N}$ with a drift pointing away from 0. Claim (4.10) then follows by the usual Gaussian tail estimates.

We proceed similarly for the first sum in (4.9). Using (3.12) and the standard estimate on the convergence to stationarity for finite Markov chains (see e.g. [22, (12.11), p.155]),

$$\operatorname{Var}\xi^k_{\mathcal{G}_n}(x) \overset{(3.12)}{=} \frac{1}{2}\Big(\tilde{Q}^{2k}_{\mathcal{G}_n}(x,x) - \frac{1}{N_n}\Big) \le e^{-\lambda_{\mathcal{G}_n}k}, \tag{4.12}$$

where $\lambda_{\mathcal{G}_n}\ge\beta$ is the spectral gap appearing in Assumption 1.1(c) (due to the laziness, there is no 2 in the exponent). Therefore, for every $\varepsilon>0$, $k_0\ge 1$, and $y\in\mathcal{V}_n$,

$$\mathbb{Q}^{x,x'}_n\Big(\Big|\sum_{k>k_0}\xi^k_{\mathcal{G}_n}(y)\Big| \ge \frac{\varepsilon}{3}\Big) \le 2\exp\Big(-\frac{\varepsilon^2 e^{\beta k_0}}{18}\Big). \tag{4.13}$$

Finally, for the second sum in (4.9), we claim that for every $\varepsilon>0$, $k_0\ge 1$ and $y\in\mathcal{V}_n$,

$$\mathbb{Q}^{x,x'}_n\Big(\Big|\sum_{0\le k\le k_0}\sum_{\tilde{z}\in\tilde{\mathcal{V}}_n}(\Pi_{\mathcal{G}_n}\tilde{Q}^k_{\mathcal{G}_n})(y,\tilde{z})Z_k(\tilde{z})\Big| \ge \frac{\varepsilon}{3}\Big) \le 2\exp\Big(-\frac{\varepsilon^2 N_n}{9(k_0+1)}\Big). \tag{4.14}$$

Indeed, by the same computation as in (3.10)–(3.11),

$$\operatorname{Var}\Big(\sum_{0\le k\le k_0}\sum_{\tilde{z}\in\tilde{\mathcal{V}}_n}(\Pi_{\mathcal{G}_n}\tilde{Q}^k_{\mathcal{G}_n})(y,\tilde{z})Z_k(\tilde{z})\Big) = \tfrac{1}{2}(k_0+1)\pi_{\mathcal{G}_n}(y) = \tfrac{k_0+1}{2N_n}, \tag{4.15}$$

this follows by the same reasoning as above.

Claim (4.1) then follows from (4.9), (4.10), (4.13), and (4.14) using the triangle inequality, a union bound, and the fact that $|B_{\mathbb{T}_d}(\mathrm{o},r)| = |B_{\mathcal{G}_n}(x,r)| = \frac{d(d-1)^r-2}{d-2} \le d(d-1)^r$, if x is $2r$-treelike. $\square$

Remark 4.2. Later, it will play the key role that the coupling $\mathbb{Q}_n^{x,x'}$ also couples the underlying Z-fields. In particular, we will use that $\mathbb{Q}_n^{x,x'}$-a.s.

$$\begin{aligned} Z_0(y) &= \mathtt{Z}_0(\rho_{x,2r}(y)), \qquad y \in B_{\mathcal{G}_n}(x, 2r), \\ Z_0(y) &= \mathtt{Z}'_0(\rho_{x',2r}(y)), \qquad y \in B_{\mathcal{G}_n}(x', 2r). \end{aligned} \tag{4.16}$$

which follows directly from (4.7).

5. Robust components of the GFF on the tree

In what follows, we assume that $\Psi_{\mathcal{G}_n}$, $\varphi_{\mathbb{T}_d}$, and the underlying fields Z, $\mathtt{Z}$ are constructed on some probability space $(\Omega, \mathcal{A}, P)$ and (4.4), (4.6) hold. As explained in the introduction (cf. (1.8)), in the sprinkling construction we will write $\Psi_{\mathcal{G}_n}$ as a sum of two independent fields $\Psi^1_{\mathcal{G}_n}$ and $\Psi^2_{\mathcal{G}_n}$. To this end, let $t \in (0, 1)$ be a parameter which will later depend on n, and write $Z_0 = \sqrt{1-t^2} Z_0^1 + t Z_0^2$, where $Z_0^i = (Z_0^i(x) : x \in \tilde{\mathcal{V}}_n)$, $i \in \{1, 2\}$, are two independent copies of Z_0. Similarly as above (4.4), we define

$$\xi^{0,i}_{\mathcal{G}_n}(x) := (\mathrm{Id} - \Pi_{\mathcal{G}_n}) Z_0^i(x) = Z_0^i(x) - \frac{1}{N_n} \sum_{y \in \mathcal{V}_n} Z_0^i(y), \quad x \in \mathcal{V}_n, i \in \{1, 2\}, \tag{5.1}$$

(where in the second equality we used (3.6) and $\pi_{\mathcal{G}_n}(x) = 1/N_n$), and set

$$\Psi^1_{\mathcal{G}_n}(x) \coloneqq \sqrt{1-t^2}\,\xi^{0,1}_{\mathcal{G}_n}(x) + \sum_{k\ge1} \xi^k_{\mathcal{G}_n}(x) \qquad \text{and} \qquad \Psi^2_{\mathcal{G}_n}(x) \coloneqq t\xi^{0,2}_{\mathcal{G}_n}(x). \tag{5.2}$$

Then $\Psi_{\mathcal{G}_n} = \Psi^1_{\mathcal{G}_n} + \Psi^2_{\mathcal{G}_n}$, and $\Psi^1_{\mathcal{G}_n}$, $\Psi^2_{\mathcal{G}_n}$ are independent.

Next, we introduce two independent copies $\mathtt{Z}_0^1$, $\mathtt{Z}_0^2$ of $\mathtt{Z}_0$ so that $\mathtt{Z}_0 = \sqrt{1-t^2}\mathtt{Z}_0^1 + t\mathtt{Z}_0^2$, and define (cf. (4.6))

$$\varphi^1_{\mathbb{T}_d}(x) \coloneqq \sqrt{1-t^2}\,\mathtt{Z}_0^1(x) + \sum_{k\ge1} \zeta^k(x) \qquad \text{and} \qquad \varphi^2_{\mathbb{T}_d}(x) \coloneqq t\mathtt{Z}_0^2(x). \tag{5.3}$$

Then $\varphi_{\mathbb{T}_d} = \varphi^1_{\mathbb{T}_d} + \varphi^2_{\mathbb{T}_d}$ and the summands are independent.

The goal of the next three sections is to show that the supercritical level sets of $\varphi^1_{\mathbb{T}_d}$, and as consequence also of $\Psi^1_{\mathcal{G}_n}$, have large connected components. Unfortunately, we cannot apply the results of [3, 4] directly, because $\Psi^1_{\mathcal{G}_n}$ and $\varphi^1_{\mathbb{T}_d}$ are no longer Gaussian free fields. In this section, we thus show that for any $h < h_\star$ the level set $E^{\ge h}(\varphi_{\mathbb{T}_d})$ of the unmodified field $\varphi_{\mathbb{T}_d}$ has infinite components which are robust to certain perturbations. In the next section, we use this result to show that the level set $E^{\ge h}(\varphi^1_{\mathbb{T}_d})$ percolates if $h < h_\star$ and t is small enough. Finally, in Section 7, we transfer these result to the field $\Psi^1_{\mathcal{G}_n}$, using the coupling from Section 4.

For the sprinkling construction of Section 9, we need to consider two types of perturbations of $E^{\ge h}(\varphi_{\mathbb{T}_d})$. The first one comes from the field $\varphi^2_{\mathbb{T}_d}$, as already explained, and the second one from an independent Bernoulli percolation. For the latter, let $\iota = (\iota(x))_{x \in V(\mathbb{T}_d)}$ be i.i.d. Bernoulli random variables with $P(\iota(x) = 1) = p$ which are independent of $\mathtt{Z}$ and thus of $\varphi_{\mathbb{T}_d}$. The robustness against the perturbation by $\varphi^2_{\mathbb{T}_d}$ involves certain averaging property for $\varphi_{\mathbb{T}_d}$ and is driven by a parameter $\gamma \in [-\infty, 0]$. Formally, for $x \in V(\mathbb{T}_d)$ (recalling that $\mathrm{desc}(x)$ is the set of direct descendants of x in $\mathbb{T}_d$), let $\mathcal{K}(h, p, \gamma)$ be the set of *robust* vertices in $E^{\ge h}(\varphi_{\mathbb{T}_d})$ defined by

$$\mathcal{K}(h, p, \gamma) := \Big\{ x \in \mathcal{V}_n : \varphi_{\mathbb{T}_d}(x) \ge h, \iota(x) = 1, \text{ and } \sum_{y \in \mathrm{desc}(x)} \varphi_{\mathbb{T}_d}(y) \ge \gamma \Big\}, \tag{5.4}$$

and let $\mathcal{C}_{\rm o}^{h,p,\gamma}$ be the connected component of $\mathcal{K}(h,p,\gamma)$ containing the root o. Note that if $p=1$ and $\gamma=-\infty$, then $\mathcal{C}_{\rm o}^{h,p,\gamma}$ agrees with the connected component $\mathcal{C}_{\rm o}^{h}$ of the level set $E^{\geq h}(\varphi_{\mathbb{T}_d})$ containing the root o. We set

$$\eta(h,p,\gamma) \coloneqq P(|\mathcal{C}_{\rm o}^{h,p,\gamma}|=\infty), \tag{5.5}$$

$$\mathcal{S} \coloneqq \{(h,p,\gamma)\in\mathbb{R}\times[0,1]\times[-\infty,0] : \eta(h,p,\gamma)>0\}. \tag{5.6}$$

The main result of this section is the following proposition which shows that, in the supercritical regime, $\mathcal{C}_{\rm o}^{h,p,\gamma}$ has similar properties as $\mathcal{C}_{\rm o}^{h}$, cf. [3, Theorems 5.1, 5.3] or [4, (2.14), (2.16)].

Proposition 5.1. *(a) If $(h,p,\gamma)\in\mathcal{S}$ and $h'<h$, $p'>p$, $\gamma'<\gamma$, then also $(h',p',\gamma')\in\mathcal{S}$. Moreover, for every $h<h_\star$ there is $p<1$ and $\gamma>-\infty$ such that $(h,p,\gamma)\in\mathcal{S}$.*

(b) For every (h,p,γ) in the interior $\mathcal{S}^0$ of $\mathcal{S}$ there is $\lambda_h^{p,\gamma}>1$ such that

$$\lim_{k\to\infty} P\big(|\mathcal{C}_{\rm o}^{h,p,\gamma}\cap S_{\mathbb{T}_d}({\rm o},k)|\geq(\lambda_h^{p,\gamma})^k/k^2\big)=\eta(h,p,\gamma)>0. \tag{5.7}$$

(c) The functions $(h,p,\gamma)\mapsto\lambda_h^{p,\gamma}$ and $(h,p,\gamma)\mapsto\eta(h,p,\gamma)$ are continuous on $\mathcal{S}^0$ (this includes the continuity at points $(h,p,-\infty)\in\mathcal{S}^0$).

Remark 5.2. We expect that $\mathcal{S}$ is open, that is $\mathcal{S}^0=\mathcal{S}$. Proving this would require to study the critical behaviour of $\mathcal{C}_{\rm o}^{h,p,\gamma}$ which goes beyond the scope of this paper.

Proof of Proposition 5.1. The proof uses multi-type branching process techniques and is a modification of the arguments given in Section 3 of [28] and in Sections 3–5 of [3]. Here, we only explain how these arguments should be adapted to our setting and leave out the parts that are relatively standard in the context of the multi-type branching processes.

We first recall the recursive construction of $\varphi_{\mathbb{T}_d}$ from [3, Section 2.1]. Define random variables

$$\begin{aligned} Y_{\rm o} &\coloneqq \varphi_{\mathbb{T}_d}({\rm o}) \\ Y_x &\coloneqq \varphi_{\mathbb{T}_d}(x)-\frac{1}{d-1}\varphi_{\mathbb{T}_d}({\rm anc}(x)), \qquad \text{for } x\in V(\mathbb{T}_d)\setminus\{{\rm o}\}. \end{aligned} \tag{5.8}$$

Then, by the domain Markov property of $\varphi_{\mathbb{T}_d}$, see [3, (2.6),(2.7)], $(Y_x)_{x\in V(\mathbb{T}_d)}$ are independent random variables such that $Y_{\rm o}\sim\mathcal{N}(0,\frac{d-1}{d-2})$ and $Y_x\sim\mathcal{N}(0,\frac{d}{d-1})$ for $x\neq{\rm o}$.

The definition (5.8) can be written as $\varphi_{\mathbb{T}_d}({\rm o})=Y_{\rm o}$ and

$$\varphi_{\mathbb{T}_d}(x)=\frac{1}{d-1}\varphi_{\mathbb{T}_d}({\rm anc}(x))+Y_x \quad \text{for } x\in V(\mathbb{T}_d)\setminus\{{\rm o}\}. \tag{5.9}$$

The field $\varphi_{\mathbb{T}_d}$ is thus determined by Y_x's, by applying (5.9) recursively. This also implies that $\varphi_{\mathbb{T}_d}$ can be viewed as a multi-type branching process. Indeed, we can view every $x\in S_{\mathbb{T}_d}({\rm o},k)$ as an individual in the k-th generation of the branching process with an attached type $\varphi_{\mathbb{T}_d}(x)\in\mathbb{R}$. (5.9) can be then rephrased as: every individual x has $d-1$ children (d children if $x={\rm o}$) whose types, conditionally on $\varphi_{\mathbb{T}_d}(x)$, are chosen independently according to the normal distribution $\mathcal{N}(\frac{1}{d-1}\varphi_{\mathbb{T}_d}(x),\frac{d}{d-1})$. This point of view can easily be adapted to $\mathcal{C}_{\rm o}^{h}$, namely by considering the same multi-type branching process but instantly killing all individuals whose type does not exceed h. Relying on this point of view, [3] investigates the properties of $\mathcal{C}_{\rm o}^{h}$ using branching process techniques.

We now modify this construction to apply to $\mathcal{C}_{\rm o}^{h,p,\gamma}$. In addition to instantly killing the individuals whose type does not exceed h, we also kill individuals x for which $\iota(x)=0$, and we kill all direct descendants of x if $\sum_{y\in{\rm desc}(x)}\varphi_{\mathbb{T}_d}(y)<\gamma$. Then the surviving individuals form a component $\bar{\mathcal{C}}_{\rm o}^{h,p,\gamma}$ which is slightly larger than $\mathcal{C}_{\rm o}^{h,p,\gamma}$. More precisely, since we only kill the direct descendants of non-robust vertices, and not those vertices themselves,

$$\bar{\mathcal{C}}_{\mathrm{o}}^{h,p,\gamma} = \mathcal{C}_{\mathrm{o}}^{h,p,\gamma} \cup \Big\{ x \in \partial_{\mathbb{T}_d} \mathcal{C}_{\mathrm{o}}^{h,p,\gamma} : \varphi_{\mathbb{T}_d}(x) \geq h, \iota(x) = 1, \sum_{y \in \mathrm{desc}(x)} \varphi_{\mathbb{T}_d}(y) \overset{(!)}{<} \gamma \Big\}. \tag{5.10}$$

As consequence, $|\bar{\mathcal{C}}_{\mathrm{o}}^{h,p,\gamma}| = \infty$ iff $|\mathcal{C}_{\mathrm{o}}^{h,p,\gamma}| = \infty$, and $|\bar{\mathcal{C}}_{\mathrm{o}}^{h,p,\gamma} \cap S_{\mathbb{T}_d}(\mathrm{o}, k)| \geq a$ implies that $|\mathcal{C}_{\mathrm{o}}^{h,p,\gamma} \cap S_{\mathbb{T}_d}(\mathrm{o}, k-1)| \geq a/(d-1)$. Hence, it is sufficient to show claims (b,c) for $\bar{\mathcal{C}}_{\mathrm{o}}^{h,p,\gamma}$ instead of $\mathcal{C}_{\mathrm{o}}^{h,p,\gamma}$ (with an additional constant $(d-1)$). The advantage of the former is that it can be interpreted as a multi-type branching process.

The key role in the investigations of [3] plays certain operator introduced in [28] in order to give a spectral characterisation of the critical value $h_\star$. This operator is defined as follows, cf. [3, Section 2.2]: Let ν be the centred Gaussian measure of variance $\frac{d-1}{d-2}$. For $h \in \mathbb{R}$, $f \in L^2(\nu)$ and $a \in \mathbb{R}$, set

$$(L_h f)(a) \coloneqq (d-1) \mathbf{1}_{[h,\infty)}(a)\, E^Y \big[(f \mathbf{1}_{[h,\infty)})(\tfrac{a}{d-1} + Y) \big], \tag{5.11}$$

where $Y \sim \mathcal{N}(0, \frac{d}{d-1})$ and E^Y is the expectation with respect to Y. The operator L_h is the 'mean value' operator corresponding to $\mathcal{C}_{\mathrm{o}}^h$ when it is viewed as a multi-type branching process, more precisely, for any $x \neq \mathrm{o}$ and $a \geq h$,

$$(L_h f)(a) = E\Big[\sum_{y \in \mathcal{C}_{\mathrm{o}}^h \cap \mathrm{desc}(x)} f(\varphi_{\mathbb{T}_d}(y)) \,\Big|\, \varphi_{\mathbb{T}_d}(x) = a, x \in \mathcal{C}_{\mathrm{o}}^h \Big]. \tag{5.12}$$

Denoting λ_h the largest eigenvalue of L_h, the critical point $h_\star$ is given as the unique solution of the equation $\lambda_h = 1$, see [28, Proposition 3.3]

For $\bar{\mathcal{C}}_{\mathrm{o}}^{h,p,\gamma}$, the corresponding operator has a similar, slightly more complicated, form: For $f \in L^2(\nu)$ and $a \in \mathbb{R}$ (and for $x \neq \mathrm{o}$, $a \geq h$ in the formula on the right-hand side of the first line, which is only included to motivate the definition),

$$\begin{aligned}
(L_h^{p,\gamma} f)(a) &= E\Bigg[\sum_{x \in \bar{\mathcal{C}}_{\mathrm{o}}^{h,p,\gamma} \cap \mathrm{desc}(x)} f(\varphi_{\mathbb{T}_d}(y)) \,\Bigg|\, \varphi_{\mathbb{T}_d}(x) = a, x \in \bar{\mathcal{C}}_{\mathrm{o}}^{h,p,\gamma} \Bigg] \\
&\coloneqq p \mathbf{1}_{[h,\infty)}(a) E^Y \Bigg[\mathbf{1}_{[\gamma,\infty)} \Big(\sum_{i=1}^{d-1} (\tfrac{a}{d-1} + Y_i) \Big) \sum_{i=1}^{d-1} (f \mathbf{1}_{[h,\infty)})(\tfrac{a}{d-1} + Y_i) \Bigg], \\
&= p(d-1) \mathbf{1}_{[h,\infty)}(a) E^Y \Bigg[\mathbf{1}_{[\gamma,\infty)} \Big(\sum_{i=1}^{d-1} (\tfrac{a}{d-1} + Y_i) \Big) (f \mathbf{1}_{[h,\infty)})(\tfrac{a}{d-1} + Y_1) \Bigg],
\end{aligned} \tag{5.13}$$

where $(Y_i)_{i=1,\dots,d-1}$ are i.i.d. $\mathcal{N}(0, \frac{d}{d-1})$ and E^Y is the corresponding expectation. Note that $L_h = L_h^{1,-\infty}$.

Contrary to L_h, the operator $L_h^{p,\gamma}$ is not self-adjoint in $L^2(\nu)$. We thus need an additional argument to show that (cf. [28, Proposition 3.1]):

The value $\lambda_h^{p,\gamma} \coloneqq \|L_h^{p,\gamma}\|_{L^2(\nu)} = \sup\{\langle g, L_h^{p,\gamma} g\rangle_{L^2(\nu)} : \|g\|_{L^2(\nu)} = 1\}$ is a simple eigenvalue of $L_h^{p,\gamma}$. Moreover, there is a unique, non-negative eigenfunction $\chi_h^{p,\gamma} \in L^2(\nu)$ of $L_h^{p,\gamma}$ corresponding to $\lambda_h^{p,\gamma}$ with $\|\chi_h^{p,\gamma}\|_{L^2(\nu)} = 1$. (5.14)

To show this, observe first that from (5.11), (5.13) it follows that there exist functions $K_h, K_h^{p,\gamma} : [h,\infty)^2 \to (0,\infty)$ such that, for $a \geq h$ (which is the relevant range since

$L_h f(a) = L_h^{p,\gamma} f(a) = 0$ for $a < h$),

$$\begin{aligned}(L_h f)(a) &= \int_{[h,\infty)} K_h(a,y) f(y) \nu(\mathrm{d}y),\\ (L_h^{p,\gamma} f)(a) &= \int_{[h,\infty)} K_h^{p,\gamma}(a,y) f(y) \nu(\mathrm{d}y).\end{aligned} \tag{5.15}$$

Moreover, $K_h^{p,\gamma} \le K_h$ for all admissible values of h, p, and γ. Since L_h is a Hilbert-Schmidt operator on $L^2(\nu)$ (see [28, (3.16)]), it follows that $L_h^{p,\gamma}$ is a Hilbert-Schmidt and thus compact operator on $L^2(\nu)$ as well. By Riesz-Schauder theorem (see e.g. [26, Theorem 6.15]), every $\lambda \neq 0$ in the spectrum $\sigma(L_h^{p,\gamma})$ of $L_h^{p,\gamma}$ is an eigenvalue of $L_h^{p,\gamma}$ and 0 is the only possible limit point of $\sigma(L_h^{p,\gamma})$. Since $\lambda_h^{p,\gamma} = \|L_h^{p,\gamma}\|_{L^2(\nu)} = \sup\{|\lambda| : \lambda \in \sigma(L_h^{p,\gamma})\}$, it follows that there is $\lambda \in \mathbb{C}$ with $|\lambda| = \lambda_h^{p,\gamma}$ such that $L_h^{p,\gamma}\chi = \lambda\chi$ for some (possibly complex valued) $\chi \in L^2(\nu)$ with $\|\chi\|_{L^2(\nu)} = 1$.

Next, we verify that $\lambda > 0$ and $\chi \ge 0$. If χ is not of the form $\chi = \beta g$ for some $\beta \in \mathbb{C}$ and a real-valued non-negative function g, then $\||\chi|\|_{L^2(\nu)} = 1$ and $\langle |\chi|, L_h^{p,\gamma}|\chi|\rangle_{L^2(\nu)} > \lambda_h^{p,\gamma}$ which leads to contradiction with the definition of $\lambda_h^{p,\gamma}$ in (5.14). Hence, $\chi = \beta g$. Since the multiplication by scalars preserves eigenfunctions, we can assume that $\beta = 1$, that is $\chi = g$ is non-negative as required. The equality $L_h^{p,\gamma}\chi = \lambda\chi$ then implies that $\lambda > 0$ as well, and thus $\lambda = \lambda_h^{p,\gamma}$.

To finish the proof of (5.14), it remains to show that $\lambda_h^{p,\gamma}$ is a simple eingenvalue. We proceed similarly to [28]: If $f \in L^2(\nu)$ is an eigenfunction of $L_h^{p,\gamma}$ attached to $\lambda_h^{p,\gamma}$, then we can assume that it is non-negative, as explained in the last paragraph. Thus $\langle f - \alpha\chi, \mathbf{1}\rangle_{L^2(\nu)} = 0$ for some $\alpha \ge 0$. The function $f - \alpha\chi$ is also an eigenfunction attached to $\lambda_h^{p,\gamma}$, so it is a multiple of a non-negative function. It follows that $f - \alpha\chi = 0$ in $L^2(\nu)$. That is $\lambda_h^{p,\gamma}$ is a simple eigenvalue corresponding to $\chi = \chi_h^{p,\gamma}$, completing the proof of (5.14).

Using (5.15), one easily shows that $\chi_h^{p,\gamma}(a) > 0$ for $a \in [h,\infty)$. From (5.13), (5.14) it follows that $\lambda_h^{p,\gamma}$ is decreasing in h and γ, and stricly increasing in p. Strict monotonicity in h can be proved as in [28, (3.23)]. Since $\chi_h^{p,\gamma} > 0$ on $[h,\infty)$, for $\gamma > \gamma'$ we have

$$\begin{aligned}\lambda_h^{p,\gamma} = \langle \chi_h^{p,\gamma}, L_h^{p,\gamma}\chi_h^{p,\gamma}\rangle_{L^2(\nu)} &\overset{(5.13)}{<} \langle \chi_h^{p,\gamma}, L_h^{p,\gamma'}\chi_h^{p,\gamma}\rangle_{L^2(\nu)}\\ &\overset{(5.14)}{\le} \langle \chi_h^{p,\gamma'}, L_h^{p,\gamma'}\chi_h^{p,\gamma'}\rangle_{L^2(\nu)} = \lambda_h^{p,\gamma'},\end{aligned} \tag{5.16}$$

yielding the strict monotonicity in γ. The continuity of $\lambda_h^{p,h}$ can be shown using the same arguments as in the proof of [28, (3.20)]. In particular, the continuity at $\gamma = -\infty$ follows from the lower-semicontinuity of $\lambda_h^{p,\gamma}$ (cf. (5.14)) and its monotonicity.

The rest of the proof of Proposition 5.1 follows the lines of [28, 3] with mostly obvious modifications, frequently relying on the fact that $L_h^{p,\gamma}$ is "smaller" than L_h (in the sense explained under (5.15)): First, as in [28, Proposition 3.3], it can be shown that the value of $\lambda_h^{p,\gamma}$ dictates whether the process is sub- or supercritical,

$$\{(h,p,\gamma) : \lambda_h^{p,\gamma} \ge 1\} \supset \mathcal{S} \supset \mathcal{S}^0 = \{(h,p,\gamma) : \lambda_h^{p,\gamma} > 1\}, \tag{5.17}$$

where $\mathcal{S}$ is as in (5.6), and the equality in (5.17) follows from the strict monotonicities of $\lambda_h^{p,\gamma}$ discussed in the last paragraph. Second, the same argument as in the proof of [3, Proposition 3.1(i)] provides a control on the growth of $\chi_h^{p,\gamma}$, which is necessary for the further steps. Third, Section 4 of [3] (studying an functional equation for the non-percolation probability) needs to be adapted: besides changing the defininition of the non-linear operator R_h from [3, (4.3),(4.4)] accordingly, only relatively straightforward changes are required there.

After these preparatory steps, the continuity of η in claim (c) of the proposition can be proved in the same way as Theorem 5.1, and claim (b) in the same way as Theorem 5.3 in [3]. The first part of claim (a) follows by monotonicity. Finally, using the continuity of $\lambda_h^{p,\gamma}$ from (c)

$$\lim_{\gamma\to-\infty} \lambda_h^{p,\gamma} \overset{(5.13)}{=} p \lim_{\gamma\to-\infty} \lambda_h^{1,\gamma} = p\lambda_h. \tag{5.18}$$

Hence if $h < h_\star$ and thus $\lambda_h > 1$, then there exist $p \in (0,1)$ and $\gamma \in \mathbb{R}$ with $\lambda_h^{p,\gamma} > 1$, proving the second part of (a). $\square$

6. Percolation for the pruned field on the tree

We now consider the *pruned* field $\varphi^1_{\mathbb{T}_d}$ defined in (5.3) (recall that $\varphi^1_{\mathbb{T}_d}$ implicitly depends on the sprinkling strength t) and show that for $h < h_\star$ and t small enough its level set $E^{\geq h}(\varphi^1_{\mathbb{T}_d})$ percolates.

Let $\mathcal{C}_{\rm o}^{h,p}(t)$ be the connected component of $\{x \in V(\mathbb{T}_d) : \varphi_{\mathbb{T}_d}(x) \geq h,\ \varphi^1_{\mathbb{T}_d}(x) \geq h,\ \iota(x) = 1\}$ containing o, and abbreviate $\eta(h,p) \coloneqq \eta(h,p,-\infty)$.

Proposition 6.1. *For every $\delta \in (0,1)$, $h < h_\star$, and $p \in [0,1]$ such that $(h,p,-\infty) \in \mathcal{S}^0$,*

$$\lim_{\substack{k\to\infty\\ t\to 0}} P\Big(|\mathcal{C}_{\rm o}^{h,p}(t) \cap S_{\mathbb{T}_d}({\rm o},k)| \geq \big(p(1-\delta)\lambda_h\big)^k\Big) = \eta(h,p). \tag{6.1}$$

(In the limit we allow $k \to \infty$ and then $t \to 0$, or $t \to 0, k \to \infty$ together.)

Proof. Since $\mathcal{C}_{\rm o}^{h,p}(t) \subset \mathcal{C}_{\rm o}^{h,p,-\infty}$, the left-hand side of (6.1) is bounded from above by $\lim_{k\to\infty} P(|\mathcal{C}_{\rm o}^{h,p,-\infty} \cap S_{\mathbb{T}_d}({\rm o},k)| \geq 1) = P(|\mathcal{C}_{\rm o}^{h,p,-\infty}| = \infty) = \eta(h,p)$, by (5.5), yielding the upper bound in (6.1).

To prove the lower bound, we will use Proposition 5.1(b) with $h' \in (h, h_\star)$ and $\gamma > -\infty$, and show that when $t > 0$ is small enough, then subtracting $\varphi^2_{\mathbb{T}_d}$ "does not destroy $\mathcal{C}_{\rm o}^{h',p,\gamma}$ too much". To this end, recall from (5.3) that $\varphi^2_{\mathbb{T}_d}$ is an i.i.d. field. However, it is not independent of $\varphi_{\mathbb{T}_d}$, so we need to compute its conditional distribution given $\varphi_{\mathbb{T}_d}$.

Lemma 6.2. *Conditionally on $\varphi_{\mathbb{T}_d}$, $\varphi^2_{\mathbb{T}_d}$ is a Gaussian field determined by*

$$E\big(\varphi^2_{\mathbb{T}_d}(x) \mid \sigma(\varphi_{\mathbb{T}_d})\big) = \tfrac{t^2}{2}\Big(\varphi_{\mathbb{T}_d}(x) - \tfrac{1}{d}\sum_{z\sim x}\varphi_{\mathbb{T}_d}(z)\Big), \tag{6.2}$$

$$E\big(\varphi^2_{\mathbb{T}_d}(x)\varphi^2_{\mathbb{T}_d}(y) \mid \sigma(\varphi_{\mathbb{T}_d})\big) = \tfrac{t^2}{2}\delta_{x,y} - \tfrac{t^4}{4}\big(\delta_{x,y} - \tfrac{1}{d}\mathbf{1}_{x\sim y}\big). \tag{6.3}$$

Proof. By (2.6), (4.5), and (5.3), the fields $\varphi^2_{\mathbb{T}_d}$ and $\varphi_{\mathbb{T}_d}$ are centred jointly Gaussian fields satisfying

$$\begin{aligned} &E(\varphi^2_{\mathbb{T}_d}(x)\varphi^2_{\mathbb{T}_d}(y)) = E(\varphi^2_{\mathbb{T}_d}(x)\varphi_{\mathbb{T}_d}(y)) = t^2\delta_{x,y}/2,\\ &E(\varphi_{\mathbb{T}_d}(x)\varphi_{\mathbb{T}_d}(y)) = g_{\mathbb{T}_d}(x,y) \end{aligned} \tag{6.4}$$

for every $x, y \in V(\mathbb{T}_d)$. Denoting Q the transition matrix of the usual random walk on $\mathbb{T}_d$, we observe that that for every $x, y \in V(\mathbb{T}_d)$

$$\begin{aligned} &\operatorname{Cov}\big(\varphi^2_{\mathbb{T}_d}(x) - \tfrac{t^2}{2}((\mathrm{Id}-Q)\varphi_{\mathbb{T}_d})(x), \varphi_{\mathbb{T}_d}(y)\big)\\ &\quad= \tfrac{t^2}{2}\Big(\delta_{x,y} - \sum_{z\in V(\mathbb{T}_d)} (\mathrm{Id}-Q)(x,z)\operatorname{Cov}\big(\varphi_{\mathbb{T}_d}(z), \varphi_{\mathbb{T}_d}(y)\big)\Big)\\ &\quad\overset{(2.6)}{=} \tfrac{t^2}{2}\Big(\delta_{x,y} - \sum_{z\in V(\mathbb{T}_d)} (\mathrm{Id}-Q)(x,z) g_{\mathbb{T}_d}(z,y)\Big) = 0, \end{aligned} \tag{6.5}$$

where in the last equality we used the well-known identity $(\mathrm{Id}-Q)g_{\mathbb{T}_d} = \mathrm{Id}$ for the Green function. It follows that the field $\psi := \varphi^2_{\mathbb{T}_d} - \frac{t^2}{2}(\mathrm{Id}-Q)\varphi_{\mathbb{T}_d}$ is independent of $\sigma(\varphi_{\mathbb{T}_d})$. Hence,

$$E\big(\varphi^2_{\mathbb{T}_d} \mid \sigma(\varphi_{\mathbb{T}_d})\big) = E\big(\psi + \tfrac{t^2}{2}(\mathrm{Id}-Q)\varphi_{\mathbb{T}_d} \mid \sigma(\varphi_{\mathbb{T}_d})\big) = \tfrac{t^2}{2}(\mathrm{Id}-Q)\varphi_{\mathbb{T}_d}, \tag{6.6}$$

from which (6.2) follows.

The conditional covariance of $\varphi^2_{\mathbb{T}_d}$ agrees with the covariance of ψ (see e.g. [19, Corollary 1.10]), which is

$$\begin{aligned} E(\psi(x)\psi(y)) = {} & E(\varphi^2_{\mathbb{T}_d}(x)\varphi^2_{\mathbb{T}_d}(y)) \\ & - \tfrac{t^2}{2}\sum_{z\in V(\mathbb{T}_d)} (\mathrm{Id}-Q)(y,z)E(\varphi^2_{\mathbb{T}_d}(x)\varphi_{\mathbb{T}_d}(z)) \\ & - \tfrac{t^2}{2}\sum_{z\in V(\mathbb{T}_d)} (\mathrm{Id}-Q)(x,z)E(\varphi^2_{\mathbb{T}_d}(y)\varphi_{\mathbb{T}_d}(z)) \\ & + \tfrac{t^4}{4}\sum_{z,z'\in V(\mathbb{T}_d)} (\mathrm{Id}-Q)(x,z)(\mathrm{Id}-Q)(y,z')E(\varphi_{\mathbb{T}_d}(z)\varphi_{\mathbb{T}_d}(z')). \end{aligned} \tag{6.7}$$

Statement (6.3) then follows by inserting the values of the expectations from (6.4), and by applying once more the above identity for the Green function. □

We continue with the proof of Proposition 6.1. Consider an arbitrary $h' > h$. If $x \in \mathcal{C}_{\mathrm{o}}^{h',p,\gamma} \setminus \{\mathrm{o}\}$, then $\varphi_{\mathbb{T}_d}(x) \ge h'$ and $\varphi_{\mathbb{T}_d}(\mathrm{anc}(x)) \ge h'$. Therefore, by the robustness condition (5.4), for $x \in \mathcal{C}_{\mathrm{o}}^{h',p,\gamma} \setminus \{\mathrm{o}\}$,

$$\begin{aligned} E\big(\varphi^1_{\mathbb{T}_d}(x) \mid \sigma(\varphi_{\mathbb{T}_d})\big) &= E\big(\varphi_{\mathbb{T}_d}(x) - \varphi^2_{\mathbb{T}_d}(x) \mid \sigma(\varphi_{\mathbb{T}_d})\big) \\ &\overset{(6.2)}{=} \varphi_{\mathbb{T}_d}(x) - \tfrac{t^2}{2}\Big(\varphi_{\mathbb{T}_d}(x) - \tfrac{1}{d}\sum_{z\sim x}\varphi_{\mathbb{T}_d}(z)\Big) \\ &= (1-\tfrac{t^2}{2})\varphi_{\mathbb{T}_d}(x) + \tfrac{t^2}{2d}\varphi_{\mathbb{T}_d}(\mathrm{anc}(x)) + \tfrac{t^2}{2d}\sum_{z\in\mathrm{desc}(x)}\varphi_{\mathbb{T}_d}(z) \\ &\overset{(5.4)}{\ge} h' + t^2(\tfrac{\gamma}{2d} - \tfrac12 + \tfrac{h'}{2d}). \end{aligned} \tag{6.8}$$

By (6.3), $\mathrm{Var}\big(\varphi^1_{\mathbb{T}_d}(x) \mid \sigma(\varphi_{\mathbb{T}_d})\big) \le ct^2$. For t small, $h' + t^2(\frac{\gamma}{2d} - \frac12 - \frac{h'}{2d}) > h$ and thus

$$\lim_{t\downarrow 0} P\big(\varphi^1_{\mathbb{T}_d}(x) \ge h \mid \sigma(\varphi_{\mathbb{T}_d})\big) = 1, \qquad \text{uniformly for } x \in \mathcal{C}_{\mathrm{o}}^{h',p,\gamma}\setminus\{\mathrm{o}\}. \tag{6.9}$$

A similar computation implies that (6.9) holds for $x = \mathrm{o}$ as well. In addition, by (6.3), conditionally on $\varphi_{\mathbb{T}_d}$, the random variables $\varphi^1_{\mathbb{T}_d}(x)$, $\varphi^1_{\mathbb{T}_d}(y)$ are independent whenever $d_{\mathbb{T}_d}(x,y) \ge 2$. By the domination argument of [23], the family $(\mathbf{1}_{[h,\infty)}(\varphi^1_{\mathbb{T}_d}(x)) : x \in \mathcal{C}_{\mathrm{o}}^{h',p,\gamma})$ dominates (conditionally on $\varphi_{\mathbb{T}_d}$) an independent Bernoulli percolation on $\mathcal{C}_{\mathrm{o}}^{h',p,\gamma}$ with parameter $g(t)$ and $g(t) \uparrow 1$ as $t \downarrow 0$. As consequence:

$$\begin{aligned} &\text{For every } \varepsilon > 0,\ h' > h \text{ and } \gamma \in \mathbb{R} \text{ there is } t_0 = t_0(h,h',\gamma,\varepsilon) \\ &\text{such that } \mathcal{C}_0^{h,p}(t) \text{ dominates } \mathcal{C}_0^{h',p(1-\varepsilon),\gamma} \text{ for all } t < t_0. \end{aligned} \tag{6.10}$$

We now fix $\delta > 0$ and $\varepsilon < \delta/4$. By the continuity of $\lambda_h^{p,\gamma}$ proved in Proposition 5.1(c), there is a neighbourhood $\mathcal{U}_\delta \subset \mathcal{S}^0$ of $(h,p,-\infty)$ such that ,

$$\lambda_{h'}^{p(1-\varepsilon),\gamma} = p(1-\varepsilon)\lambda_{h'}^{1,\gamma} \ge p\big(1-\tfrac{\delta}{2}\big)\lambda_h \qquad \text{for every } (h',p,\gamma) \in \mathcal{U}_\delta. \tag{6.11}$$

In particular, $(p(1-\delta)\lambda_h)^k \le (\lambda_{h'}^{p(1-\varepsilon),\gamma})^k/k^2$ for $k \ge k_0(\delta)$. Hence, by (6.10), for such k, for every $h' > h$ and γ such that $(h', p, \gamma) \in \mathcal{U}_\delta$, and for every $t < t_0(h, h', \gamma, \varepsilon)$, the probability in (6.1) satisfies

$$\begin{aligned} &P\big(|\mathcal{C}_{\mathrm{o}}^{h,p}(t) \cap S_{\mathbb{T}_d}(\mathrm{o}, k)| \ge \big(p(1-\delta)\lambda_h\big)^k\big) \\ &\quad \ge P\big(|\mathcal{C}_{\mathrm{o}}^{h',p(1-\varepsilon),\gamma} \cap S_{\mathbb{T}_d}(\mathrm{o}, k)| \ge (\lambda_{h'}^{p(1-\varepsilon),\gamma})^k/k^2\big). \end{aligned} \tag{6.12}$$

Observe that the probability on the right-hand side is independent of t. Therefore, by Proposition 5.1(b), for every $\varepsilon < \delta/4$, $h' \in (h, h_\star)$ and γ such that $(h', p, \gamma) \in \mathcal{U}_\delta$,

$$\liminf_{\substack{k\to\infty \\ t\to 0}} P\big(|\mathcal{C}_{\mathrm{o}}^{h,p}(t) \cap S_{\mathbb{T}_d}(\mathrm{o}, k)| \ge (p(1-\delta)\lambda_h)^k\big) \ge \eta(h', p(1-\varepsilon), \gamma), \tag{6.13}$$

where we can take $k \to \infty$ and then $t \to 0$, or $k \to \infty$, $t \to 0$ simultaneously. Since $\varepsilon > 0$ can be taken arbitrarily close to 0 and (h', γ) close to $(h, -\infty)$, the lower bound for (6.1) follows using the continuity of η from Proposition 5.1(c). □

7. Many mesoscopic components for the pruned field on finite graphs

As a corollary of Proposition 6.1 and the coupling stated in Proposition 4.1, we now prove the existence of many mesoscopic components for the level set of the field $\Psi^1_{\mathcal{G}_n}$.

To state this result precisely, we need to introduce an additional notation. Let $\bar{Z}^2_0 = (\bar{Z}^2_0(x), x \in \mathcal{V}_n)$ be a copy of Z^2_0 which is independent of Z, Z^1_0 and Z^2_0, and set (cf. (5.2))

$$\bar{\Psi}^2_{\mathcal{G}_n}(x) := t\Big(\bar{Z}^2_0(x) - \frac{1}{N_n}\sum_{y\in\mathcal{G}_n} \bar{Z}^2_0(y)\Big). \tag{7.1}$$

The field $\bar{\Psi}^2_{\mathcal{G}_n}$ has the same law as $\Psi^2_{\mathcal{G}_n}$ and thus $\bar{\Psi}_{\mathcal{G}_n} := \Psi^1_{\mathcal{G}_n} + \bar{\Psi}^2_{\mathcal{G}_n}$ has the same law as $\Psi_{\mathcal{G}_n}$, that is it is a zero-average Gaussian free field on $\mathcal{G}_n$. For $p > 1/2$ we define $L = L(p) < 0$ by

$$P(\bar{Z}^2_0(x) \ge L) = p. \tag{7.2}$$

We set

$$\bar{\mathcal{V}}_n := \{x \in \mathcal{V}_n : \bar{Z}^2_0(x) \ge L\}, \tag{7.3}$$

and use $\bar{\mathcal{G}}_n$ to denote the subgraph of $\mathcal{G}_n$ induced by $\bar{\mathcal{V}}_n$. Finally, for $x \in \bar{\mathcal{V}}_n$, let $\mathcal{C}^h_x(t)$ be the connected component component of the set $E^{\ge h}(\Psi^1_{\mathcal{G}_n}) \cap E^{\ge h}(\Psi_{\mathcal{G}_n})$ in $\bar{\mathcal{G}}_n$.

To understand the reason for this notation, note that eventually, in Section 9, we will use $\bar{\Psi}_{\mathcal{G}_n}$, and not $\Psi_{\mathcal{G}_n}$, to show that the supercritical level set has a giant component. In particular, we will use the field $\bar{\Psi}^2_{\mathcal{G}_n}$ for the sprinkling. At the sites where this field is very small, it can potentially destroy the connected components of the level set. To avoid this, we will restrict to $\bar{\mathcal{G}}_n$ in our sprinkling construction. It is also useful to compare the definition of $\mathcal{C}^h_x(t)$ with the definition of $\mathcal{C}^{h,p}_{\mathrm{o}}(t)$ in Section 6, in particular note that the role of the percolation ι is taken by the subgraph $\bar{\mathcal{G}}_n$.

Proposition 7.1. *Let $h < h_\star$ and let p be such that $(h, p, -\infty) \in \mathcal{S}^0$. Then there exists $c_h \in (0,1)$ such that for any $\delta > 0$ and any sequence $t_n \downarrow 0$,*

$$\lim_{n\to\infty} P\Big(\sum_{x\in\mathcal{V}_n} \mathbf{1}_{\{|\mathcal{C}^h_x(t_n)| \ge N_n^{c_h}\}} \ge (1-\delta)\eta(h,p)N_n\Big) = 1. \tag{7.4}$$

Proof. The proof follows the steps of Section 5 of [4] and is an application of the second moment method. Some simplifications, compared to [4], are due to the fact that our Proposition 4.1 uses two independent copies of $\varphi_{\mathbb{T}_d}$, so we do not need to use the decoupling inequalities for $\varphi_{\mathbb{T}_d}$ as in [4].

Let $r_n = c_1 \log N_n$ with $c_1 > 0$, and set

$$\begin{aligned} W_n &:= \{x \in \mathcal{V}_n : x \text{ is } 2r_n\text{-treelike}\}, \\ \widetilde{W}_n &:= \big\{(x,x') \in W_n \times W_n : B_{\mathcal{G}_n}(x,2r_n) \cap B_{\mathcal{G}_n}(x',2r_n) = \emptyset\big\}. \end{aligned} \tag{7.5}$$

By [4, (5.6), (5.7)], it is possible to fix c_1 small, such that, for some some $c > 0$ and for all n large enough,

$$|W_n| \geq N_n(1 - N_n^{-c}) \quad \text{and} \quad |\widetilde{W}_n| \geq N_n^2(1 - N_n^{-c}). \tag{7.6}$$

We will prove (7.4) with

$$c_h := c_1 \log\big(p\lambda_h(1-2\delta')\big), \tag{7.7}$$

where $\delta' > 0$ is small enough so that $p\lambda_h(1-2\delta') > 1$, which is possible since $(h,p,-\infty) \in \mathcal{S}^0$ implies $1 < \lambda_h^{p,-\infty} = p\lambda_h$.

Let $\tilde{\mathcal{C}}_x^h(t_n) \subset \mathcal{C}_x^h(t_n)$ be the connected component of $\mathcal{C}_x^h(t_n) \cap B(x,2r_n)$ containing x, and define events

$$\begin{aligned} A_x^{\mathcal{G}_n,h} &:= \big\{|\tilde{\mathcal{C}}_x^h(t_n) \cap S_{\mathcal{G}_n}(x,r_n)| \geq N_n^{c_h}\big\}, \qquad \text{for } x \in \mathcal{V}_n, \\ A_{\mathrm{o}}^{\mathbb{T}_d,h} &:= \big\{|\mathcal{C}_{\mathrm{o}}^{h,p}(t_n) \cap S_{\mathbb{T}_d}(\mathrm{o},r_n)| \geq N_n^{c_h}\big\}. \end{aligned} \tag{7.8}$$

We now show

$$\lim_{n\to\infty} P\Big(\sum_{x \in W_n} \mathbf{1}_{A_x^{\mathcal{G}_n,h}} \geq (1-\delta)\eta(h,p)N_n\Big) = 1, \tag{7.9}$$

from which (7.4) directly follows.

To show (7.9), for every pair $x, x' \in \widetilde{W}_n$, we use the coupling $\mathbb{Q}_n^{x,x'}$ from Proposition 4.1 (with $r = r_n$) to couple $\Psi_{\mathcal{G}_n}$ with two independent copies of $\varphi_{\mathbb{T}_d}$, $\varphi'_{\mathbb{T}_d}$ of the Gaussian free field on $\mathbb{T}_d$. By Remark 4.2, this coupling also couples the underlying fields Z_0, Z_0 and Z_0' as in (4.16). In addition, we write $\mathsf{Z}_0 = \sqrt{1-t_n^2}\mathsf{Z}_0^1 + t_n \mathsf{Z}_0^2$ and assume that $Z_0^2(y) = \mathsf{Z}_0^2(\rho_{x,2r_n}(y))$ for every $y \in B_{\mathcal{G}_n}(x,2r_n)$, where Z_0^1, Z_0^2 are independent copies of Z_0. We also use analogous statements for Z', $\mathsf{Z}_0'^2$ in the ball $B_{\mathcal{G}_n}(x',2r_n)$. We also couple the site percolation ι (introduced in the paragraph above (5.4)) and its independent copy ι' with the field $\bar{Z}_0^2$ so that

$$\begin{aligned} \iota(\rho_{x,2r_n}(y)) &= \mathbf{1}_{[L,\infty)}(\bar{Z}_0^2(y)), \qquad y \in B_{\mathcal{G}_n}(x,2r_n), \\ \iota'(\rho_{x',2r_n}(y)) &= \mathbf{1}_{[L,\infty)}(\bar{Z}_0^2(y)), \qquad y \in B_{\mathcal{G}_n}(x',2r_n), \end{aligned} \tag{7.10}$$

which is always possible due to the choice (7.2) of L and since $B_{\mathcal{G}_n}(x,2r_n)$ and $B_{\mathcal{G}_n}(x',2r_n)$ are disjoint. Note that (7.10) implies

$$\rho_{x,2r_n}(\bar{\mathcal{V}}_n \cap B(x,2r_n)) = \{y \in B_{\mathbb{T}_d}(\mathrm{o},2r_n) : \iota(y) = 1\}. \tag{7.11}$$

We now fix $\varepsilon > 0$ arbitrary but small enough so that

$$\lambda_{h+\varepsilon} > (1-\delta')\lambda_h \qquad \text{and} \qquad (h+\varepsilon,p,-\infty) \in \mathcal{S}_0. \tag{7.12}$$

where δ' was fixed in (7.7). When all coupling equalities from the last paragraph hold and when the coupling $\mathbb{Q}_n^{x,x'}$ succeeds, that is the complement of the event on the left-hand side of (4.1) occurs, then it follows from the definitions of the components $\tilde{\mathcal{C}}_x^h(t_n)$ and $\mathcal{C}_{\mathrm{o}}^{h,p}(t_n)$ that $A_{\mathrm{o}}^{\mathbb{T}_d,h-\varepsilon} \supset A_x^{\mathcal{G}_n,h} \supset A_{\mathrm{o}}^{\mathbb{T}_d,h+\varepsilon}$, and similarly for x', replacing $A_{\mathrm{o}}^{\mathbb{T}_d,h\pm\varepsilon}$ by their independent copies defined in terms of the field $\varphi'_{\mathbb{T}_d}$. Hence, for $x \in W_n$,

$$P(A_x^{\mathcal{G}_n,h}) \geq P(A_{\mathrm{o}}^{\mathbb{T}_d,h+\varepsilon}) - e(n,\varepsilon), \tag{7.13}$$

where $e(n,\varepsilon)$ is the probability that the coupling fails. By Proposition 4.1, $e(n,\varepsilon)$ is bounded by right-hand side of (4.1) with $r = r_n$, in particular $0 \leq e(n,\varepsilon) \leq c(\varepsilon)N_n^{-k}$ for any $k \in \mathbb{N}$.

By Proposition 6.1 (applied with $h+\varepsilon$ instead of h, and δ' instead of δ), using also (7.12), we obtain that

$$\liminf_{n\to\infty} P(A_{\mathrm{o}}^{\mathbb{T}_d,h+\varepsilon}) \geq \eta(h+\varepsilon,p). \tag{7.14}$$

Hence, by (7.13), also $\liminf_{n\to\infty} P(A_x^{\mathcal{G}_n,h}) \geq \eta(h+\varepsilon,p)$ for every $x \in W_n$. As consequence, since $\varepsilon > 0$ is arbitrary, using (7.6) and the continuity of η from Proposition 5.1(c),

$$\liminf_{n\to\infty} \frac{1}{N_n} E\Big(\sum_{x\in W_n} \mathbf{1}_{A_x^{\mathcal{G}_n,h}}\Big) \geq \eta(h,p). \tag{7.15}$$

We now compute the variance of the sum in the last display. Expanding it, and then using the coupling $\mathbb{Q}_n^{x,x'}$ again,

$$\begin{aligned}\operatorname{Var}\Big(\sum_{x\in W_n} \mathbf{1}_{A_x^{\mathcal{G}_n,h}}\Big) &= \sum_{x,x'\in W_n} \Big(P\big(A_x^{\mathcal{G}_n,h} \cap A_{x'}^{\mathcal{G}_n,h}\big) - P\big(A_x^{\mathcal{G}_n,h}\big)P\big(A_{x'}^{\mathcal{G}_n,h}\big)\Big).\\ &\leq |(W_n \times W_n)\setminus \widetilde{W}_n| + \sum_{(x,x')\in \widetilde{W}_n} \Big(P(A_{\mathrm{o}}^{\mathbb{T}_d,h-\varepsilon})^2 - P(A_{\mathrm{o}}^{\mathbb{T}_d,h+\varepsilon})^2\Big) + e(n,\varepsilon).\end{aligned} \tag{7.16}$$

By definition of $A_{\mathrm{o}}^{\mathbb{T}_d,h}$, using also that $\mathcal{C}_{\mathrm{o}}^{h-\varepsilon,p}(t_n) \subset \mathcal{C}_{\mathrm{o}}^{h-\varepsilon,p,-\infty}$,

$$\begin{aligned}\limsup_{n\to\infty} P(A_{\mathrm{o}}^{\mathbb{T}_d,h-\varepsilon}) &= \limsup_{n\to\infty} P\big(|\mathcal{C}_{\mathrm{o}}^{h-\varepsilon,p}(t_n) \cap S_{\mathbb{T}_d}(\mathrm{o},r_n)| \geq N_n^{c_h}\big)\\ &\leq \limsup_{n\to\infty} P\big(|\mathcal{C}_{\mathrm{o}}^{h-\varepsilon,p,-\infty} \cap S_{\mathbb{T}_d}(\mathrm{o},r_n)| \geq 1\big)\\ &= P(|\mathcal{C}_{\mathrm{o}}^{h-\varepsilon,p,-\infty}| = \infty) = \eta(h-\varepsilon,p).\end{aligned} \tag{7.17}$$

Inequalities (7.6), (7.14), (7.16) and (7.17) together imply that

$$\limsup_{n\to\infty} \frac{1}{N_n^2} \operatorname{Var}\Big(\sum_{x\in W_n} \mathbf{1}_{A_x^{\mathcal{G}_n,h}}\Big) \leq \eta(h-\varepsilon,p)^2 - \eta(h+\varepsilon,p)^2. \tag{7.18}$$

The right-hand side of this inequality can be made arbitrary small by taking $\varepsilon \downarrow 0$, using the continuity of η. Statement (7.9) then follows from (7.15) and (7.18) by applying Chebyshev inequality. ☐

We finish this section by a simple lemma which gives a lower bound on the number of vertices that are contained in small components of the (non-pruned) field $\Psi_{\mathcal{G}_n}$. This lower bound will be used to show the upper bound on $|\mathcal{C}_{\max}^{\mathcal{G}_n,h}|$ in the proof of Theorem 1.2. In its statement we use $\mathcal{C}_x^{\mathcal{G}_n,h}$ to denote connected component of $E^{\geq h}(\Psi_{\mathcal{G}_n})$ containing $x \in \mathcal{V}_n$.

Lemma 7.2. *Let $\mathcal{H}_n := \{x \in \mathcal{V}_n : \mathcal{C}_x^{\mathcal{G}_n,h} \subset B_{\mathcal{G}_n}(x, r_n/2)\}$ with $r_n = c_1 \log N_n$ as in the last proof. Then for every $h < h_\star$ and $\delta > 0$*

$$\lim_{n\to\infty} P(|\mathcal{H}_n| > (1-\eta(h)-\delta)N_n) = 1. \tag{7.19}$$

Proof. The proof is very similar to the previous one. Due to (7.6) it is sufficient to prove the claim for $|\mathcal{H}_n \cap W_n|$ instead of $|\mathcal{H}_n|$. For $x \in W_n$ define the events $A_x^{\mathcal{G}_n,h} := \{\mathcal{C}_x^{\mathcal{G}_n,h} \subset B_{\mathcal{G}_n}(x,r_n/2)\}$, $A_{\mathrm{o}}^{\mathbb{T}_d,h} := \{\mathcal{C}_{\mathrm{o}}^h \subset B_{\mathbb{T}_d}(\mathrm{o},r_n/2)\}$. Using Proposition 4.1, we can couple those events so that $A_{\mathrm{o}}^{\mathbb{T}_d,h-\varepsilon} \subset A_x^{\mathcal{G}_n,h} \subset A_{\mathrm{o}}^{\mathbb{T}_d,h+\varepsilon}$. By (1.4), $\lim_{n\to\infty} P(A_x^{\mathbb{T}_d,h}) = 1-\eta(h)$. Using the same first and second moment arguments as in the previous proof, the lemma easily follows. ☐

8. Expansion properties of reduced graphs

Before going to the final sprinkling step, we need a little lemma that show that particular subgraphs of $\mathcal{G}_n$ still have good expansion properties. To this end recall from (7.3) the definition of the subgraph $\bar{\mathcal{G}}_n$. For $K \in \mathbb{R}$, let

$$\hat{\mathcal{V}}_n = \{x \in \mathcal{V}_n : \bar{Z}_0^2(x) \geq L, \Psi^1_{\mathcal{G}_n}(x) \geq K\} \subset \bar{\mathcal{V}}_n, \tag{8.1}$$

and let $\hat{\mathcal{G}}_n$ be the subgraph of $\mathcal{G}_n$ induced by $\hat{\mathcal{V}}_n$. The additional condition $\Psi^1_{\mathcal{G}_n}(x) \geq K$ will later ensure that, in the sprinkling step, the sites in $\hat{\mathcal{V}}_n$ have a reasonable chance to be in the level set $E^{\geq h}(\bar{\Psi}_{\mathcal{G}_n})$ of the field $\bar{\Psi}_{\mathcal{G}_n}$ (defined under (7.1)). We will always assume that $K \leq h$, so that the mesoscopic connected components $\mathcal{C}^h_x(t_n)$ (in $\bar{\mathcal{G}}_n$ as considered in Proposition 7.1) are also connected components in $\hat{\mathcal{G}}_n$.

We now show that $\hat{\mathcal{G}}_n$ has good expansion properties, at least when we only consider its large subsets. Recall from (2.8) that β' is the lower bound on the isoperimetric constants of $\mathcal{G}_n$.

Lemma 8.1. *For every $\delta > 0$, there exist $K_0 = K_0(\delta)$ and $L_0 = L_0(\delta)$ such that for every $K < K_0$ and $L < L_0$*

$$P\Big(\inf_{A\subset\mathcal{V}(\hat{\mathcal{G}}_n):\delta N_n\leq|A|\leq N_n/2} \frac{|\partial_{\hat{\mathcal{G}}_n} A|}{|A|} \geq \frac{\beta'}{2}\Big) \geq 1 - N^{-\varepsilon} \tag{8.2}$$

with $\varepsilon > 0$ independent of δ.

Proof. We show that for K, L sufficiently negative, $B_1(n) = \{x \in \mathcal{V}_n : \bar{Z}_0^2(x) < L\}$ and $B_2(n) = \{x \in \mathcal{V}_n : \Psi^1_{\mathcal{G}_n}(x) < K\}$ satisfy

$$P\big(|B_1(n)| + |B_2(n)| \leq \beta'\delta N_n/2\big) \geq 1 - N^{-\varepsilon}. \tag{8.3}$$

The claim of the lemma then follows from (2.8). Indeed, on the event in (8.3), for A as in (8.2),

$$|\partial_{\hat{\mathcal{G}}_n} A| \geq |\partial_{\mathcal{G}_n} A| - (|B_1(n)| + |B_2(n)|) \geq \beta'|A| - \beta'\delta N_n/2 \geq \beta'|A|/2. \tag{8.4}$$

To prove (8.3), observe first that $|B_1(n)|$ is a binomial random variable with parameters N_n and $p = P(\bar{Z}_0^2(x) < L)$. Hence, by taking L_0 depending on δ sufficiently small, we obtain by the standard large deviation estimates that $P(|B_1(n)| \geq \beta'\delta N_n/4) \leq e^{-cN_n}$ for all $L < L_0$.

For $B_2(n)$, we use the second moment method again. Observe first that by (5.2),

$$\begin{aligned}\operatorname{Cov}(\Psi^1_{\mathcal{G}_n}(x), \Psi^1_{\mathcal{G}_n}(y)) &= \operatorname{Cov}(\Psi_{\mathcal{G}_n}(x), \Psi_{\mathcal{G}_n}(y)) - \operatorname{Cov}(\Psi^2_{\mathcal{G}_n}(x), \Psi^2_{\mathcal{G}_n}(y)) \\ &\overset{(2.4)}{=} G_{\mathcal{G}_n}(x,y) - t_n^2 \operatorname{Cov}(\xi_0^2(x), \xi_0^2(y)) \overset{(3.12)}{=} G_{\mathcal{G}_n}(x,y) - \tfrac{t_n^2}{2}(\delta_{x,y} + \tfrac{1}{N_n}).\end{aligned} \tag{8.5}$$

In particular, using the estimate (2.7) on $G_{\mathcal{G}_n}$, since $t_n \to 0$, $\sigma_x^2 := \operatorname{Var}(\Psi^1_{\mathcal{G}_n}(x)) = G_{\mathcal{G}_n}(x,x) + O(t_n^2) \in (c, c')$ for some $0 < c < c' < \infty$, and if $x \neq y$, for some $\varepsilon \in (0,1)$,

$$\operatorname{Cov}(\Psi^1_{\mathcal{G}_n}(x), \Psi^1_{\mathcal{G}_n}(y)) \leq C(d-1)^{-d_{\mathcal{G}_n}(x,y)} + N_n^{-\varepsilon}. \tag{8.6}$$

As consequence, we can fix K_0 small enough so that

$$E(|B_2(n)|) = \sum_{x\in\mathcal{V}_n} P(\Psi^1_{\mathcal{G}_n}(x) < K) \leq \sum_{x\in\mathcal{V}_n} e^{-K^2/(2c)} \leq \beta'\delta N_n/8 \tag{8.7}$$

for every $K \leq K_0$. By the normal comparison lemma, see e.g. [20, Theorem 4.2.1], for $x \neq y \in \mathcal{V}_n$, we then obtain

$$\begin{aligned}&P(\Psi^1_{\mathcal{G}_n}(x) \leq K, \Psi^1_{\mathcal{G}_n}(y) \leq K) - P(\Psi^1_{\mathcal{G}_n}(x) \leq K)P(\Psi^1_{\mathcal{G}_n}(y) \leq K) \\ &\quad\leq C(\operatorname{Cov}(\Psi^1_{\mathcal{G}_n}(x)/\sigma_x, \Psi^1_{\mathcal{G}_n}(y)/\sigma_y) \vee 0) \leq C(\operatorname{Cov}(\Psi^1_{\mathcal{G}_n}(x), \Psi^1_{\mathcal{G}_n}(y)) \vee 0).\end{aligned} \tag{8.8}$$

As consequence, using also the fact that diameter of $\mathcal{G}_n$ is smaller than $C\log N_n$ (see e.g. [17, Proposition 3.1.5]) and that $|S_{\mathcal{G}_n}(x,r)| \le d(d-1)^{r-1}$, we obtain

$$\begin{aligned}
&\mathrm{Var}(|B_2(n)|)\\
&\quad= \sum_{x,y\in\mathcal{V}_n} P(\Psi^1_{\mathcal{G}_n}(x)\le K, \Psi^1_{\mathcal{G}_n}(y)\le K) - P(\Psi^1_{\mathcal{G}_n}(x)\le K)P(\Psi^1_{\mathcal{G}_n}(y)\le K)\\
&\quad\le N_n + C\sum_{x\in\mathcal{V}_n}\sum_{r=1}^{C\log N_n}\sum_{y\in\mathcal{V}_n: d_{\mathcal{G}_n}(x,y)=r} (\mathrm{Cov}(\Psi^1_{\mathcal{G}_n}(x),\Psi^1_{\mathcal{G}_n}(y))\vee 0) \qquad (8.9)\\
&\quad\overset{(8.6)}{\le} N_n + C\sum_{x\in\mathcal{V}_n}\sum_{r=1}^{C\log N_n}\sum_{y\in\mathcal{V}_n: d_{\mathcal{G}_n}(x,y)=r} ((d-1)^{-r} + N_n^{-\varepsilon}) \le N_n^{2-\varepsilon}.
\end{aligned}$$

By Chebyshev inequality, using (8.7), (8.9), $P(|B_2(n)|\ge \beta'\delta N_n/4)\le N_n^{-\varepsilon}$.

Combining the conclusions of the last two paragraphs then implies (8.3) and completes the proof. □

9. Sprinkling / Proof of Theorem 1.2

With all preparations of the previous sections, the sprinkling construction is relatively straightforward and follows the steps of [5].

We start by showing that the field $\bar\Psi_{\mathcal{G}_n}$ defined under (7.1) contains a giant component of size at least $\eta(h)(1-\delta)N_n$, with probability tending to one as $n\to\infty$. Since $\bar\Psi_{\mathcal{G}_n}$ is a zero-average Gaussian free field on $\mathcal{G}_n$, this will imply the lower bound on $|\mathcal{C}^{\mathcal{G}_n,h}_{\max}|$ for Theorem 1.2.

For $h<h_\star$ and $\delta\in(0,1/8)$ as in the statement of Theorem 1.2, we fix an arbitrary $h'\in(h,h_\star)$ and set $\varepsilon := h'-h$. We further fix K, L small and p close to 1 so that L, p are linked by (7.2) and

$$\begin{aligned}
&K = h\wedge K_0(\delta\eta(h')/2), && L < L_0(\delta\eta(h')/2),\\
&(h',p,-\infty)\in\mathcal{S}^0, && \eta(h',p) > \eta(h')/2,
\end{aligned} \qquad (9.1)$$

where $K_0(\delta\eta(h')/2)$, $L_0(\delta\eta(h')/2)$ are as in Lemma 8.1, and the last inequality in (9.1) can be satisfied by Proposition 5.1(c). We let $t_n\to 0$ slowly so that

$$P(\bar Z_0^2(x)\ge t_n^{-1}(h+1-K)) \ge N_n^{-c_{h'}\beta'\delta\eta(h',p)/8}, \qquad (9.2)$$

where $c_{h'}$ is as in Proposition 7.1, and β' as in Lemma 8.1.

Due to Lemma 8.1, using also (9.1), we know that:

$$\text{For } \mathcal{A}^1_n := \Big\{ \inf_{\substack{A\subset V(\hat{\mathcal{G}}_n):\\ \delta\eta(h')\frac{N_n}{2}<|A|<\frac{N_n}{2}}} \frac{|\partial_{\hat{\mathcal{G}}_n}A|}{|A|} \ge \frac{\beta'}{2}\Big\} \text{ we have } P(\mathcal{A}^1_n)\ge 1-N_n^{-c}. \qquad (9.3)$$

By Gaussian tail estimates, the zero-averaging term in the definition (7.1) of $\bar\Psi^2_{\mathcal{G}_n}$ is negligible with high probability:

$$\text{For } \mathcal{A}^2_n := \Big\{ \Big| N_n^{-1}\sum_{y\in\mathcal{V}_n} \bar Z_0^2(y)\Big| \le \varepsilon\Big\} \text{ we have } P(\mathcal{A}^2_n)\ge 1-e^{-cN_n}. \qquad (9.4)$$

Introducing $m_n := N_n^{c_{h'}}$ to denote the minimal size of mesoscopic components and writing $a_k = (1-k\delta)\eta(h',p)$ for $k\in\{1,2\}$, Proposition 7.1 implies that:

$$\text{For } \mathcal{A}^3_n := \Big\{ \sum_{x\in\mathcal{V}_n} \mathbf{1}_{\{|\mathcal{C}^{h'}_x(t_n)|\ge m_n\}} \ge a_1 N_n\Big\} \text{ we have } \lim_{n\to\infty} P(\mathcal{A}^3_n) = 1, \qquad (9.5)$$

that is $E^{\geq h'}(\Psi^1_{\mathcal{G}_n}) \cap \hat{\mathcal{V}}_n$ has many mesoscopic components, with high probability. Finally, since $\Psi^1_{\mathcal{G}_n}$ is independent of $\bar{Z}^2_0$ and the graph $\hat{\mathcal{G}}_n$ depends on $\bar{Z}^2_0$ only via $\mathbf{1}_{[L,\infty)}(\bar{Z}^2_0(x))$, it follows that:

$$\text{Conditionally on } \sigma(\Psi^1_{\mathcal{G}_n}, \hat{\mathcal{G}}_n), \text{ the random variables } (\bar{Z}^2_0(x))_{x\in\hat{\mathcal{V}}_n} \text{ are i.i.d. distributed as } \mathcal{N}(0,1/2) \text{ random variable conditioned on being larger than } L. \tag{9.6}$$

In particular, since $L < 0$,

$$\begin{aligned} p_n &\coloneqq P(\bar{Z}^2_0(x) \geq t_n^{-1}(h + t_n\varepsilon - K) \mid \sigma(\Psi^1_{\mathcal{G}_n}, \hat{\mathcal{G}}_n), \{x \in \hat{\mathcal{V}}_n\}) \\ &\overset{(9.2)}{\geq} N_n^{-c_{h'}\beta'\delta\eta(h',p)/8}. \end{aligned} \tag{9.7}$$

Assume now that $\mathcal{A}_n \coloneqq \mathcal{A}^1_n \cap \mathcal{A}^2_n \cap \mathcal{A}^3_n$ occurs. On $\mathcal{A}^3_n$, we can fix a set of at most $a_1 N_n/m_n$ mesoscopic components of $E^{\geq h'}(\Psi^1_{\mathcal{G}_n}) \cap \hat{\mathcal{V}}_n$ that together contain at least $a_1 N_n$ vertices. Any x in those components satisfies $\bar{Z}^2_0(x) \geq L$ (by definition of $\hat{\mathcal{G}}_n$) and thus (on $\mathcal{A}^2_n$),

$$\bar{\Psi}_{\mathcal{G}_n}(x) = \Psi^1_{\mathcal{G}_n}(x) + t_n\Big(\bar{Z}^2_0(x) - N_n^{-1}\sum_{y\in\mathcal{V}_n}\bar{Z}^2_0(y)\Big) \geq h' + t_n(L-\varepsilon) \geq h \tag{9.8}$$

for all n large enough. It follows that these fixed mesoscopic components are contained in $E^{\geq h}(\bar{\Psi}_{\mathcal{G}_n})$. If $E^{\geq h}(\bar{\Psi}_{\mathcal{G}_n})$ has no component of size at least $a_2 N_n$, then one can split these fixed components into two groups A, B, each having at least $\delta\eta(h',p)N_n$ vertices, which are not connected within $E^{\geq h}(\bar{\Psi}_{\mathcal{G}_n})$. There are at most $2^{a_1 N_n/m_n}$ ways to split the fixed mesoscopic components into two groups. By (9.1), $\delta\eta(h',p)N_n > \delta\eta(h')N_n/2$. Therefore, on $\mathcal{A}^1_n$, we can use Menger's theorem to show that there are at least $\beta'\delta N_n\eta(h',p)/2$ pairwise vertex-disjoint paths from A to B in $\hat{\mathcal{G}}_n$. Since $\hat{\mathcal{G}}_n$ has at most N_n vertices, at last half of those paths are of length at most $4/\beta'\delta\eta(h',p)$ each. For every $x \in \hat{\mathcal{V}}_n$, $\Psi^1_{\mathcal{G}_n}(x) \geq K$. Therefore, if $\bar{Z}^2_0(x) \geq t_n^{-1}(h + t_n\varepsilon - K)$ and $\mathcal{A}^2_n$ occurs, then

$$\bar{\Psi}_{\mathcal{G}_n}(x) = \Psi^1_{\mathcal{G}_n}(x) + t_n\Big(\bar{Z}^2_0(x) - \frac{1}{N_n}\sum_{x\in\mathcal{V}_n}\bar{Z}^2_0(x)\Big) \geq h. \tag{9.9}$$

Hence for the groups A and B being disconnected in $\hat{\mathcal{G}}_n \cap E^{\geq h}(\bar{\Psi}_{\mathcal{G}_n})$, there must be at least one vertex with $\bar{Z}^2_0(x) < t_n^{-1}(h + t_n\varepsilon - K)$ on every of these paths. Due to (9.6) and (9.7), this has probability at most

$$(1 - p_n^{4/\beta'\delta\eta(h',p)})^{\beta'\delta N_n\eta(h',p)/4} \leq \exp(-c(\delta,h',p)N_n^{1-c_{h'}/2}). \tag{9.10}$$

It follows that the probability that $\mathcal{A}_n$ occurs and there is no connected component of $E^{\geq h}(\bar{\Psi}_{\mathcal{G}_n})$ of size at least $a_2 N_n$ (that is there is some partition of the fixed mesoscopic components into groups A and B as above that are disconnected form each other in $E^{\geq h}(\bar{\Psi}_{\mathcal{G}_n})$) is at most

$$2^{a_1 N_n/m_n}\exp(-c(\delta,h',p)N_n^{1-c_{h'}/2}) \leq \exp\{-c' N_n^{1-c_{h'}/2}\}, \tag{9.11}$$

which converges to 0 as $n \to \infty$.

Together with (9.3)–(9.5), this implies that with probability tending to one with n, $E^{\geq h}(\bar{\Psi}_{\mathcal{G}_n})$ has a connected component of size at least $a_2 N_n = (1-2\delta)\eta(h',p)N_n$. Taking h' close to h, p close to 1, using the continuity of $\eta(h,p)$ from Proposition 5.1(c), and recalling that $\bar{\Psi}_{\mathcal{G}_n}$ has the same distribution as $\Psi_{\mathcal{G}_n}$ then proves the lower bound on $|\mathcal{C}^{\mathcal{G}_n,h}_{\max}|$ for our main result (1.6) of Theorem 1.2.

The upper bounds on $|\mathcal{C}^{\mathcal{G}_n,h}_{\max}|$ and $|\mathcal{C}^{\mathcal{G}_n,h}_{\sec}|$ in (1.6) then follow from Lemma 7.2 and the lower bound on $|\mathcal{C}^{\mathcal{G}_n,h}_{\max}|$. This completes the proof of Theorem 1.2.

Remark 9.1. We conclude this paper with a short discussion of the assumptions of Theorem 1.2. Assumption 1.1(a) is clearly necessary in all our considerations (besides Section 3).

Assumption 1.1(c), that is the assumption on the spectral gap, is only used to imply the uniform isoperimetric inequality (2.8), and also in (4.12). For our results to be true, we only need (2.8) to hold for macroscopic sets (cf. proof of Lemma 8.1). Also, the argument around (4.12) can be easily adapted if $\lambda_{\mathcal{G}_n} \to 0$ sufficiently slowly.

Assumption 1.1(b) is only used very implicitly in this paper, namely to ensure that a majority of vertices of $\mathcal{G}_n$ are r_n-treelike with $r_n = c_1 \log N_n$, cf. (7.6) which is proved in [4, (5.6)] using [8, Lemma 6.1]. In Sections 7–9 of this paper, we even do not need that r_n grows so quickly. $r_n = C \log\log N_n$ for C sufficiently large would be sufficient for our purposes. Hence Assumption 1.1(b) can be replaced by: For some C sufficiently large,

$$|\{x \in \mathcal{V}_n : x \text{ is } (C \log\log N_n)\text{-treelike}\}| \geq N_n(1 - o(1)). \tag{9.12}$$

For the existence of the giant component (not necessary of size $(1-\delta)\eta(h)N_n$), the factor $(1 - o(1))$ in the last inequality could even be replaced by a $c \in (0, 1)$.

Acknowledgment. The author wish to thank P.-F. Rodriguez for useful discussions.

References

[1] Abächerli, A. Local picture and level-set percolation of the Gaussian free field on a large discrete torus. *Stochastic Process. Appl.* **129** (2019), no. 9, 3527–3546. Available at: https://doi.org/10.1016/j.spa.2018.09.017

[2] Abächerli, A.; Sznitman, A.-S. Level-set percolation for the Gaussian free field on a transient tree. *Ann. Inst. H. Poincaré Probab. Statist.* **54** (2018), no. 1, 173–201. Available at: https://doi.org/10.1214/16-AIHP799

[3] Abächerli, A.; Černý, J. Level-set percolation of the Gaussian free field on regular graphs I: Regular trees. *Electron. J. Probab.* **25** (2020), Paper No. 65, 24. Available at: https://doi.org/10.1214/20-ejp468

[4] Abächerli, A.; Černý, J. Level-set percolation of the Gaussian free field on regular graphs II: finite expanders. *Electron. J. Probab.* **25** (2020), Paper No. 130, 39. Available at: https://doi.org/10.1214/20-ejp532

[5] Alon, N.; Benjamini, I.; Stacey, A. Percolation on finite graphs and isoperimetric inequalities. *Ann. Probab.* **32** (2004), no. 3A, 1727–1745. Available at: https://doi.org/10.1214/009117904000000414

[6] Anantharaman, N.; Le Masson, E. Quantum ergodicity on large regular graphs. *Duke Math. J.* **164** (2015), no. 4, 723–765. Available at: https://doi.org/10.1215/00127094-2881592

[7] Bricmont, J.; Lebowitz, J. L.; Maes, C. Percolation in strongly correlated systems: the massless Gaussian field. *J. Statist. Phys.* **48** (1987), no. 5-6, 1249–1268. Available at: https://doi.org/10.1007/BF01009544

[8] Černý, J.; Teixeira, A.; Windisch, D. Giant vacant component left by a random walk in a random d-regular graph. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** (2011), no. 4, 929–968. Available at: http://dx.doi.org/10.1214/10-AIHP407

[9] Chiarini, A.; Nitzschner, M. Entropic repulsion for the Gaussian free field conditioned on disconnection by level-sets. *Probab. Theory Related Fields* **177** (2020), no. 1-2, 525–575. Available at: https://doi.org/10.1007/s00440-019-00957-7

[10] Conchon-Kerjan, G.: Anatomy of a Gaussian giant: supercritical level-sets of the free field on random regular graphs, Preprint, available at: *arXiv:2102.10975*. 2021.

[11] Cooper, C.; Frieze, A. Component structure of the vacant set induced by a random walk on a random graph. *Random Structures Algorithms* **42** (2013), no. 2, 135–158. Available at: https://doi.org/10.1002/rsa.20402

[12] Drewitz, A.; Prévost, A.; Rodriguez, P.-F. The sign clusters of the massless Gaussian free field percolate on $\mathbb{Z}^d$, $d \geq 3$ (and more). *Comm. Math. Phys.* **362** (2018), no. 2, 513–546. Available at: https://doi.org/10.1007/s00220-018-3209-6

[13] Drewitz, A.; Prévost, A.; Rodriguez, P.-F.: Critical exponents for a percolation model on transient graphs, Preprint, available at: *arXiv:2101.05801*. 2021.

[14] Drewitz, A.; Rodriguez, P.-F. High-dimensional asymptotics for percolation of Gaussian free field level sets. *Electron. J. Probab.* **20** (2015), 1–39. Available at: https://doi.org/10.1214/EJP.v20-3416

[15] Duminil-Copin, H.; Goswami, S.; Rodriguez, P.-F.; Severo, F.: Equality of critical parameters for percolation of Gaussian free field level-sets, Preprint, available at: *arXiv:2002.07735*. 2020.

[16] Elon, Y.; Smilansky, U. Percolating level sets of the adjacency eigenvectors of d-regular graphs. *Journal of Physics A: Mathematical and Theoretical* **43** (2010), no. 45, 455 209. Available at: https://doi.org/10.1088/1751-8113/43/45/455209

[17] Kowalski, E. *An introduction to expander graphs*, *Cours Spécialisés [Specialized Courses]*, vol. 26, Société Mathématique de France, Paris, 2019.

[18] Krivelevich, M.; Lubetzky, E.; Sudakov, B. Asymptotics in percolation on high-girth expanders. *Random Structures Algorithms* **56** (2020), no. 4, 927–947. Available at: https://doi.org/10.1002/rsa.20903

[19] Le Gall, J.-F. *Brownian motion, martingales, and stochastic calculus*, *Graduate Texts in Mathematics*, vol. 274, Springer [Cham], 2016. Available at: https://doi.org/10.1007/978-3-319-31089-3

[20] Leadbetter, M. R.; Lindgren, G.; Rootzén, H. *Extremes and related properties of random sequences and processes*, Springer Series in Statistics, Springer-Verlag, New York, 1983.

[21] Lebowitz, J. L.; Saleur, H. Percolation in strongly correlated systems. *Phys. A* **138** (1986), no. 1-2, 194–205. Available at: https://doi.org/10.1016/0378-4371(86)90180-9

[22] Levin, D. A.; Peres, Y.; Wilmer, E. L. *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.

[23] Liggett, T. M.; Schonmann, R. H.; Stacey, A. M. Domination by product measures. *Ann. Probab.* **25** (1997), no. 1, 71–95. Available at: http://doi.org/10.1214/aop/1024404279

[24] Molchanov, S. A.; Stepanov, A. K. Percolation in random fields. I. *Teoret. Mat. Fiz.* **55** (1983), no. 2, 246–256.

[25] Molloy, M.; Reed, B. The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.* **7** (1998), no. 3, 295–305. Available at: https://doi.org/10.1017/S0963548398003526

[26] Reed, M.; Simon, B. *Methods of modern mathematical physics. I*, Academic Press, Inc., New York, 1980, 2nd ed.

[27] Rodriguez, P.-F.; Sznitman, A.-S. Phase transition and level-set percolation for the Gaussian free field. *Comm. Math. Phys.* **320** (2013), no. 2, 571–601. Available at: https://doi.org/10.1007/s00220-012-1649-y

[28] Sznitman, A.-S. Coupling and an application to level-set percolation of the Gaussian free field. *Electron. J. Probab.* **21** (2016), 1–26. Available at: https://doi.org/10.1214/16-EJP4563

[29] Sznitman, A.-S. On macroscopic holes in some supercritical strongly dependent percolation models. *Ann. Probab.* **47** (2019), no. 4, 2459–2493. Available at: https://doi.org/10.1214/18-AOP1312

Jiří Černý, Department of Mathematics and Computer Science, University of Basel, Spiegelgasse 1, 4051 Basel, Switzerland

Email address: jiri.cerny@unibas.ch