PERCOLATION FOR LEVEL-SETS OF GAUSSIAN FREE FIELDS ON METRIC GRAPHS

BY JIAN DING[1],[*] AND MATEO WIRTH[1],[**]

[1]*Department of Statistics, University of Pennsylvania,* [*]*dingjian@wharton.upenn.edu;* [**]*mwirth@wharton.upenn.edu*

We study level-set percolation for Gaussian free fields on metric graphs. In two dimensions, we give an upper bound on the chemical distance between the two boundaries of a macroscopic annulus. Our bound holds with high probability conditioned on connectivity and is sharp up to a poly-logarithmic factor with an exponent of one-quarter. This substantially improves a previous result by Li and the first author. In three dimensions and higher, we provide rather precise estimates of percolation probabilities in different regimes which altogether describe a sharp phase transition.

1. Introduction.

1.1. *Gaussian free fields on metric graphs.* In this paper, we study Gaussian free fields on metric graphs of integer lattices, which are closely related to (discrete) Gaussian free fields on integer lattices. We begin with some basic definitions before stating our main results. Let $\{S_t : t \geq 0\}$ be a continuous-time simple random walk on $\mathbb{Z}^d$ with transition rates $\frac{1}{2d}$. For $d \geq 3$, the (discrete) Gaussian free field on $\mathbb{Z}^d$, $\{\phi_v : v \in \mathbb{Z}^d\}$, is defined as a mean-zero Gaussian process with covariance $\mathbb{E}(\phi_u \phi_v)$ given by (denoting below by $\mathbb{1}_A$ the indicator function of the event A)

$$(1) \qquad G(u, v) = \mathbb{E}_u \left[\int_0^\infty \mathbb{1}_{S_t = v} \, dt \right], \quad u, v \in \mathbb{Z}^d.$$

It is clear that the preceding definition cannot extend to $d = 2$ because simple random walk is recurrent in the two-dimensional lattice. For this reason (as usual), for $d = 2$ we define the Gaussian free field on a finite set $V \subset \mathbb{Z}^2$ with Dirichlet boundary conditions, denoted by $\{\phi_v : v \in V\}$, to be a mean zero Gaussian process with covariance $\mathbb{E}(\phi_u \phi_v)$ given by

$$(2) \qquad G(u, v) = \mathbb{E}_u \left[\int_0^\zeta \mathbb{1}_{S_t = v} \, dt \right], \quad u, v \in V,$$

where $\zeta = \inf\{t \geq 0 : S_t \in \partial V\}$ is the hitting time of the internal boundary $\partial V = \{v \in V : \exists u \in V^c, |u - v| = 1\}$.

Let $\mathcal{G} = \mathcal{G}(V, E)$ be the subgraph of $\mathbb{Z}^d$ on V, where we usually let V be a finite box for $d = 2$ and we take $V = \mathbb{Z}^d$ for $d \geq 3$. To each $e \in E$, we associate a different compact interval I_e of length d and identify the endpoints of this interval with the two vertices adjacent to e. The metric graph $\tilde{\mathcal{G}}$ associated to V is then defined to be $\tilde{\mathcal{G}} = \bigcup_{e \in E} I_e$. With this definition, it was shown in [11] that the Gaussian free field on $\tilde{\mathcal{G}}$, denoted by $\{\tilde{\phi}_v : v \in \tilde{\mathcal{G}}\}$, can be constructed in two equivalent ways. The first is by extending ϕ to $\tilde{\mathcal{G}}$ in the following manner: for adjacent vertices u, v, the value of $\tilde{\phi}$ on the edge $e(u, v)$, conditioned on ϕ_u and ϕ_v, is given by an independent bridge of length d of a Brownian motion with variance 2 at time 1. We note in passing that we have chosen the convention that each edge of $\mathcal{G}$ has conductance $\frac{1}{2d}$ in order to be consistent with [10].

Alternatively, one can construct $\tilde{\phi}$ by first defining a Brownian motion $\{\tilde{B}_t : t \geq 0\}$ on $\tilde{\mathcal{G}}$ as in [11], Section 2. $\tilde{B}$ behaves like a standard Brownian motion in the interior of the edges, while on the vertices (i.e., lattice points) it chooses to do excursions on each incoming edge uniformly at random (see [11] for further details). We let $\tilde{\zeta} = \inf\{t \geq 0 : \tilde{B}_t \in \partial V\}$ for $d = 2$, $\tilde{\zeta} = \infty$ for $d \geq 3$, and by an abuse of notation let $\{G(u, v) : u, v \in \tilde{\mathcal{G}}\}$ be the density of the 0-potential of $\{\tilde{B}_t : 0 \leq t < \tilde{\zeta}\}$ (with respect to the Lebesgue measure on $\tilde{\mathcal{G}}$), where u and v are now arbitrary points in $\tilde{\mathcal{G}}$ (not necessarily vertices). It was shown in [11] that the trace of $\tilde{B}$ on V (when parametrized by its local time at the vertices) is exactly the continuous-time simple random walk on V (killed at ∂V for $d = 2$), and that therefore the two definitions of G coincide for $u, v \in V$, justifying the abuse of notation. The Gaussian free field $\{\tilde{\phi}_v : v \in \tilde{\mathcal{G}}\}$ is then the continuous, mean-zero Gaussian field on $\tilde{\mathcal{G}}$ with covariance given by $\mathbb{E}[\tilde{\phi}_u \tilde{\phi}_v] = G(u, v)$. It was also shown in [11] that the value of G on the edges of $\tilde{\mathcal{G}}$ can be obtained by interpolation from the value on the vertices. For two pairs of adjacent vertices (u_1, v_1) and (u_2, v_2) in V, and two points $w_1 \in e(u_1, v_1)$ and $w_2 \in e(u_2, v_2)$ on the corresponding edges, taking the convention that either the edges are distinct or $(u_1, v_1) = (u_2, v_2)$ and letting $r_1 = |w_1 - u_1|$ and $r_2 = |w_2 - u_2|$ (here we are measuring the standard Euclidean distance), we have (cf. [11], equation (2.1))

$$
G(w_1, w_2)
$$
(3)
$$
= (1 - r_1)(1 - r_2)G(u_1, u_2) + r_1 r_2 G(v_1, v_2) + (1 - r_1)r_2 G(u_1, v_2)
$$
$$
+ r_1(1 - r_2)G(v_1, u_2) + 2d(r_1 \wedge r_2 - r_1 r_2)\mathbb{1}_{(u_1, v_1) = (u_2, v_2)}.
$$

1.2. *Main results.* The main goal of the present paper is to study level-set percolation for Gaussian free fields on metric graphs. For $r \geq 1$, we let

(4)
$$
V_r = [-r, r]^d \cap \mathbb{Z}^d
$$

be the points in the lattice contained in the box of side-length $2r$ centered at the origin (we choose this convention so that all boxes can be centered at the origin). For $d = 2$, we take a sequence $\tilde{\phi}_N$ of fields defined on the metric graphs $\tilde{\mathcal{G}}_N$ associated to V_N (with Dirichlet boundary conditions). For $h \in \mathbb{R}$, we let $\tilde{E}_N^{\geq h} = \{v \in \tilde{\mathcal{G}}_N : \tilde{\phi}_{N,v} \geq h\}$ be the level set, or excursion set, of $\tilde{\phi}_N$ above h—note that our choice of level set is different from that of [6] by a flipping symmetry, in order to be consistent with the majority of the literature. Further, for $u, v \in \tilde{E}_N^{\geq h}$, we let the chemical distance $D_{N,h}(u, v)$ be the graph distance between u and v in $\tilde{E}_N^{\geq h}$, with $D_{N,h}(u, v) = \infty$ if u and v are disconnected in $\tilde{E}_N^{\geq h}$. For two subsets $A, B \subset \tilde{\mathcal{G}}_N$, we let $D_{N,h}(A, B) = \inf\{D_{N,h}(u, v) : u \in A, v \in B\}$. The following result is an upper bound on the chemical distance between two boundaries of a macroscopic annulus, conditioned on percolation.

THEOREM 1. *For any fixed $h \in \mathbb{R}$, and $0 < \alpha < \beta < \gamma < 1$, let $\mathbb{P}_{N,h}^{\alpha,\gamma}$ be the law of $\tilde{\phi}_N$ conditioned on the event $\{D_{N,h}(V_{\alpha N}, \partial V_{\gamma N}) < \infty\}$. Then for any $\epsilon > 0$, there exists a constant C such that*

(5)
$$
\limsup_{N \to \infty} \mathbb{P}_{N,h}^{\alpha,\gamma}\big(D_{N,h}(V_{\alpha N}, \partial V_{\beta N}) > CN(\log N)^{\frac{1}{4}}\big) \leq \epsilon.
$$

REMARK 2. Note that $\mathbb{P}(D_{N,h}(V_{\alpha N}, \partial V_{\gamma N}) < \infty)$ stays above 0 uniformly in N (See Lemma 18). In a work in preparation by Aru–Lupu–Sepúlveda, it is expected that the following may be deduced as a consequence of their main results: for any $\epsilon > 0$, there exist $\delta, N_0 >$

0 such that for all $N > N_0$ and $\gamma < \beta + \delta$, we have $\mathbb{P}_{N,h}^{\alpha,\beta}(D_{N,h}(V_{\alpha N}, \partial V_{\gamma N}) < \infty) \geq 1 - \epsilon$. Provided with this continuity of the percolation probability, one would then be able to derive from Theorem 1 that

$$
(6) \qquad \limsup_{N \to \infty} \mathbb{P}_{N,h}^{\alpha,\beta}(D_{N,h}(V_{\alpha N}, \partial V_{\beta N}) > CN(\log N)^{\frac{1}{4}}) \leq \epsilon.
$$

For $d \geq 3$, we let $V = \mathbb{Z}^d$, let $\mathbf{0}$ be the origin in $\mathbb{Z}^d$ and let $\tilde{\mathcal{G}}$ be the metric graph associated to V. In the present paper, we will focus on the behavior of $p_{N,h} = \mathbb{P}(\mathbf{0} \xleftrightarrow{\geq h} \partial V_N)$ as $N \to \infty$, where $\{\mathbf{0} \xleftrightarrow{\geq h} \partial V_N\}$ denotes the event that $\mathbf{0}$ is connected to ∂V_N in $\tilde{E}^{\geq h} = \{v \in \tilde{\mathcal{G}} : \tilde{\phi}_v \geq h\}$. We obtain the following results for supercritical, subcritical and critical percolation, respectively.

The first result is an explicit characterization of the probability that $\mathbf{0}$ is in an infinite connected component of $\tilde{E}^{\geq -h}$ (for the rest of this section, we take h to be positive). Below and throughout the paper, $\text{sign}(x) = 1$ for $x \geq 0$ and $\text{sign}(x) = -1$ for $x < 0$.

THEOREM 3. *Let $\sigma_d^2 = \text{Var}[\tilde{\phi}_{\mathbf{0}}]$. Then for any $h > 0$,*

$$
(7) \qquad \lim_{N \to \infty} p_{N,-h} = \mathbb{E}\big[(1 - e^{-2h(\tilde{\phi}_{\mathbf{0}}+h)/\sigma_d^2})\mathbb{1}_{\tilde{\phi}_{\mathbf{0}} > -h}\big] = \mathbb{E}[\text{sign}(\tilde{\phi}_{\mathbf{0}} + h)].
$$

We note that the second equality follows from an elementary calculation. The next result establishes the exponential decay of $p_{N,h}$ as $N \to \infty$ for $d > 3$, with an extra log factor for $d = 3$.

THEOREM 4. *For any $h > 0$ and $d \geq 3$, there exists a constant c such that for $N > 1$,*

$$
(8) \qquad p_{N,h} \leq \exp(-ch^2 N / \log N), \quad d = 3,
$$

$$
(9) \qquad p_{N,h} \leq \exp(-ch^2 N), \quad d > 3.
$$

The third result establishes the polynomial decay of $p_{0,N}$ as $N \to \infty$ and provides bounds on the exponent.

THEOREM 5. *For any $d \geq 3$, there exist constants $C, c > 0$ such that for $N > 1$,*

$$
(10) \qquad \frac{c}{\sqrt{N}} \leq p_{N,0} \leq C\sqrt{\frac{\log N}{N}}, \quad d = 3,
$$

$$
(11) \qquad \frac{c}{N^{d/2-1}} \leq p_{N,0} \leq \frac{C}{\sqrt{N}}, \quad d > 3.
$$

Finally, for $d = 3$ we provide some bounds on the critical window for h. Below we take $h_N > 0$ to be a sequence of levels that converges to 0, and write p_N^+ for p_{N,h_N} and p_N^- for $p_{N,-h_N}$ to simplify notation.

THEOREM 6. *Let $d = 3$. Suppose that there exists a constant $C > 0$ such that $h_N \leq CN^{-1/2}$ for $N \geq 1$. Then there exists a constant $c > 0$ such that for $N \geq 1$,*

$$
(12) \qquad p_N^+ \geq \frac{c}{\sqrt{N}}.
$$

Conversely, there exists a constant $C > 0$ such that if

$$
\liminf_{N \to \infty} h_N \sqrt{\frac{N}{\log N \log \log N}} \geq C,
$$

then

$$\lim_{N\to\infty} \frac{p_N^+}{p_{N,0}} = 0. \tag{13}$$

Furthermore, if there exists a constant $C > 0$ such that $h_N \leq C(\log N/N)^{1/2}$ for $N \geq 1$, then there exists a constant $c > 0$ such that for $N > 1$

$$p_N^- \leq c\sqrt{\frac{\log N}{N}}. \tag{14}$$

Conversely, if $\lim_{N\to\infty} h_N \sqrt{\frac{N}{\log N}} = \infty$, then

$$\lim_{N\to\infty} \frac{p_N^-}{p_{N,0}} = \infty. \tag{15}$$

1.3. *Related work.* The chemical distance on level sets for $d = 2$ has been previously studied in [6] (we refer the reader to [6] for another extensive discussion of related work), where it was proved that with positive probability the chemical distance between two boundaries of a macroscopic annulus is at most $Ne^{(\log N)^\alpha}$ for any fixed $\alpha > 1/2$. Our Theorem 1 improves on [6] in the following two ways:

- Instead of proving a positive probability bound as in [6], Theorem 1 states that the upper bound on the chemical distance holds with high probability given connectivity. At the moment, we can only show the with high probability result as in (5); as noted in Remark 2, it is possible that with some expected future input, one would be able to derive a stronger version as in (6).
- The upper bound is sharpened from $Ne^{(\log N)^\alpha}$ to $N(\log N)^{1/4}$, which is somewhat surprising. In fact, the authors of the present article as well as a few people we talked to believed that the chemical distance should be at least $N \log N$.

A major difference between our proof of Theorem 1 and the proof of the corresponding result in [6] is that our proof does not rely on Makarov's theorem (on the dimension of the support of planar harmonic measures) which was a fundamental ingredient in [6]. Instead of applying Makarov's theorem, we study the *intrinsic structure* of the "exploration martingale" introduced in Section 2.

Additionally, we remark that the result proved in [6] applies to level-set percolation for Gaussian free fields on the integer lattice (as well as on the metric graph). Since percolation on the metric graph is dominated by the percolation on the integer lattice, our Theorem 1 implies that with nonvanishing probability the chemical distance in the level-set cluster on the integer lattice between the boundary of the annulus is $O(N(\log N)^{1/4})$. We feel it is possible that the methods we employ in proving Theorem 1 together with some technical work might be sufficient to show that the chemical distance on the integer lattice is $O(N(\log N)^{1/4})$ with high probability given connectivity. However, we prefer not to consider this problem here to avoid further complications. Furthermore, we note that percolation clusters for level sets on the metric graph in two dimensions are of fundamental importance since their first passage sets converge to those of (continuous) Gaussian free fields [1, 2] (thus, we believe Theorem 1 is of substantial interest on its own). Finally, [6] also established the chemical distance for critical random walk loop soup clusters. We chose not to consider proving an analogue of Theorem 1 for random walk clusters in the present paper.

For $d \geq 3$, level-set percolation for Gaussian free fields (on metric graphs) has been studied in [11], which in particular computed the connectivity probability between any two points

and showed that the critical threshold is at $h = 0$. Our methods allow us to improve on those results by deriving more quantitative information on the phase transition, especially when $d = 3$. In this case, we can compute the connectivity exponent at criticality (we remark that the real contribution of the present paper is on its upper bound, since the lower bound can be deduced easily from [11]), prove an almost exponential decay at subcriticality and provide an explicit description of the percolation probability in the supercritical regime (which can be rarely achieved in percolation models). Our results seem to describe the phase transition of the percolation model for the metric graph Gaussian free field in three dimensions in rather precise detail. This is somewhat interesting since percolation models in three dimensions are in general rather difficult.

That being said, we would like to mention that level-set percolation for Gaussian free fields on integer lattices for $d \geq 3$ has already been extensively studied (see [5, 7, 8, 13, 14, 16–18, 20–22]). Contrary to the case of two dimensions, percolation is substantially different on metric graphs and on integer lattices for $d \geq 3$, (roughly speaking) for the reason that there is a phase transition in higher dimensions but in two dimensions the percolation has the same qualitative behavior for any fixed h (see Theorem 1). We remark that percolation on integer lattices is considerably more challenging than on metric graphs. In fact, despite intensive research, it remains an open question what the exact critical threshold is for $d \geq 3$ on integer lattices (but it was proved in [7] that the critical threshold is strictly positive), as well as whether a sharp phase transition exists. It would be difficult to apply methods in the present paper to prove something on the integer lattices for $d \geq 3$, for the reason that we do not have a precise control over the "exploration martingale" (as introduced in Section 2) in the case of integer lattices.

1.4. *Discussions on future directions.*

Our work suggests a number of interesting directions for further research, which we list below:

- The factor of $N(\log N)^{1/4}$ in Theorem 1 reinstates the (now even more intriguing) question of whether the chemical distance is linear or not. Our bound of $N(\log N)^{1/4}$ strongly suggests that this is a highly delicate problem.
- Our method can give some nontrivial bounds on the exponent for chemical distances for $d \geq 3$ at criticality, but it seems challenging to compute the exact exponent.
- The difference of a factor of $\sqrt{\log N}$ between the upper and lower bounds of (10) hides important information about the geometry of critical clusters for $d = 3$. For example, whether the capacity of the critical cluster containing $\mathbf{0}$ is of order N, conditioned on the cluster intersecting with ∂V_N. It would be interesting to prove an up-to-constants bound for the connectivity probability at criticality.
- It would be very interesting to construct an incipient infinite cluster measure for critical percolation in three dimensions, as has been done for Bernoulli percolation in two dimensions in [9] (see also [3] for a nicely streamlined presentation).

1.5. *Notation conventions and organization.*

For a real vector $\mathbf{x}$ (in any dimension), we denote by $|\mathbf{x}|$ the Euclidean norm of $\mathbf{x}$, by $|\mathbf{x}|_{\ell_1}$ its ℓ_1-norm and by $|\mathbf{x}|_{\infty}$ its ℓ_{∞}-norm. We will also use $|A|$ to denote the cardinality of a finite set A. The meaning will be clear from context. We will denote by $\bar{A}$ and A^o the closure and interior of a subset $A \subset \tilde{\mathcal{G}}$, respectively. We use A^c to denote the complement of the set (or event) A.

Throughout, we will use φ and Φ to denote the density function and distribution function of the standard normal distribution, and $\bar{\Phi}$ to denote its survival function. That is, $\bar{\Phi}(x) = 1 - \Phi(x)$.

To simplify certain statements, we use the following notation to describe the asymptotic behavior of functions of N as N tends to infinity. For two positive functions f and g, we

say $f(N) = O(g(N))$ as $N \to \infty$ if there exist constants $c > 0$ and $N_0 > 0$ (possibly depending on d, h or other parameters) such that for all $N \geq N_0$, $f(N) \leq cg(N)$. We say $f(N) = o(g(N))$ if for every constant $c > 0$, there exists $N_0 > 0$ such that for all $N \geq N_0$, $f(N) \leq cg(N)$. Similarly, we say $f(N) = \Omega(g(N))$ if $g(N) = O(f(N))$ and $f(N) = \omega(g(N))$ if $g(N) = o(f(N))$. Finally, we say $f(N) = \Theta(g(N))$ if $f(N) = O(g(N))$ and $f(N) = \Omega(g(N))$.

The rest of the paper is organized as follows. In Section 2, we introduce a family of martingales which is the key to proving all the results in the present paper. In Section 3, we prove the results concerning $d \geq 3$ (as the proof is substantially simpler than that for $d = 2$), and in Section 4 we prove Theorem 1 concerning $d = 2$ (we remark that the proof of Theorem 1 encapsulates all the technical ideas of the present article).

2. Exploration martingale.

In this section, we introduce the "exploration martingale" and demonstrate some of its basic properties. We note that the approach of applying martingales in the study of percolation for Gaussian free fields has appeared before (cf. [2, 12]).

The discussion in this section applies to all dimensions, and we will denote by $\{\tilde{\phi}_v : v \in \tilde{\mathcal{G}}\}$ the Gaussian free field under consideration, without further specifying $\tilde{\mathcal{G}}$. For a finite subset $A \subset V$ (as usual, $\tilde{\mathcal{G}}$ is the metric graph associated to the vertex set V), we define the "observable" X_A to be the average of $\tilde{\phi}$ on A,

$$X_A = \frac{1}{|A|} \sum_{v \in A} \tilde{\phi}_v.$$

Let $\mathcal{I}_0$ be a deterministic, closed, bounded, connected subset of $\tilde{\mathcal{G}}$ and let $\mathcal{I}_t = \{v : D_h(\mathcal{I}_0, v) \leq t\}$ be the closed ball of radius $t > 0$ around $\mathcal{I}_0$ with respect to D_h, the graph distance on $\tilde{E}^{\geq h}$ (here we use the following convention: if u and v are distinct and $u \notin \tilde{E}^{\geq h}$ we let $D_h(u, v) = D_h(v, u) = \infty$, but for any $u \in \tilde{\mathcal{G}}$ we set $D_h(u, u) = 0$ even if $u \notin \tilde{E}^{\geq h}$). For $U \subset \tilde{\mathcal{G}}$, let $\mathcal{F}_U$ be the σ-field generated by $\{\tilde{\phi}_v : v \in U\}$, and

$$\mathcal{F}_{\mathcal{I}_t} = \{\mathcal{E} \in \mathcal{F}_{\tilde{\mathcal{G}}} : \mathcal{E} \cap \{\mathcal{I}_t \subset U\} \in \mathcal{F}_U \text{ for all open } U \supset \mathcal{I}_0\}.$$

We then define the continuous-time martingale M_A by

$$(16) \qquad\qquad M_{A,t} = \mathbb{E}[X_A \mid \mathcal{F}_{\mathcal{I}_t}].$$

We will call M_A the exploration martingale with source $\mathcal{I}_0$ and target A. Before proceeding further, we show the following measurability property of $\mathcal{I}_t$.

PROPOSITION 7. *For any open subset U of $\tilde{\mathcal{G}}$ containing $\mathcal{I}_0$ and any $t \geq 0$, we have* $\{\mathcal{I}_t \subset U\} \in \mathcal{F}_U$.

PROOF. Since $\mathcal{I}_0$ is deterministic, we assume without loss of generality that $t > 0$. Write $U_h = \bar{U} \cap \tilde{E}^{\geq h}$, let D_{U_h} be the graph distance on U_h, and $\mathcal{I}_{U,t} = \{u \in \bar{U} : D_{U_h}(\mathcal{I}_0, v) \leq t\}$ be the closed ball of radius t around $\mathcal{I}_0$ with respect to D_{U_h}. We will show that $\{\mathcal{I}_t \subset U\} = \{\mathcal{I}_{U,t} \subset U\} \in \mathcal{F}_U$. First, it is clear that for any $u, v \in \tilde{\mathcal{G}}$, $D_h(u, v) \leq D_{U_h}(u, v)$ so we have $\mathcal{I}_{U,t} \subseteq \mathcal{I}_t$ and $\{\mathcal{I}_t \subset U\} \subseteq \{\mathcal{I}_{U,t} \subset U\}$. Now, assume $\mathcal{I}_{U,t} \subset U$ and that there exists $v \in \mathcal{I}_t$ with $v \notin U$. Then there exists $0 < s \leq t$ and a path $\Gamma : [0, s] \to \tilde{\mathcal{G}}$ parametrized by D_h with $\Gamma(0) = u_0 \in \partial \mathcal{I}_0$, $\Gamma(s) = v$, and $\Gamma([0, s]) \subseteq \tilde{E}^{\geq h}$. Setting $s' = \sup\{x \in [0, s] : \Gamma(x) \in \mathcal{I}_{U,t}\}$ we have $\Gamma(s') \in \mathcal{I}_{U,t} \subset U$ and thus $s' < s$. Hence, there exists an open neighborhood (a, b) of s' such that $\Gamma((a, b)) \subset U$. By assumption, $\Gamma((a, b)) \subset \tilde{E}^{\geq h}$, so we get $\Gamma((a, b)) \subset U_h$, and thus $\Gamma((a, b)) \subset \mathcal{I}_{U,t}$, which contradicts the maximality of s'. This concludes the proof that $\{\mathcal{I}_t \subset U\} = \{\mathcal{I}_{U,t} \subset U\}$, and thus the proposition follows. □

Proposition 7, together with the following strong Markov property of the Gaussian free field (see [19], Theorem 4 in Chapter 2, Section 2.4, for a proof) will provide useful formulas for M_A and its quadratic variation $\langle M_A \rangle$.

THEOREM 8. *Let $\mathcal{K}$ be a random compact connected subset of $\tilde{\mathcal{G}}$ such that for every deterministic open subset U of $\tilde{\mathcal{G}}$, the event $\{\mathcal{K} \subseteq U\} \in \mathcal{F}_U$. Then conditioned on $\mathcal{K}$ and $\mathcal{F}_{\mathcal{K}}$, $\{\tilde{\phi}_v : v \in \tilde{\mathcal{G}} \setminus \mathcal{K}\}$ is equal in distribution to $\{\mathbb{E}[\tilde{\phi}_v \mid \mathcal{F}_{\mathcal{K}}] + \tilde{\psi}_v : v \in \tilde{\mathcal{G}} \setminus \mathcal{K}\}$ where $\tilde{\psi}$ is a Gaussian free field (with Dirichlet boundary condition) on $\tilde{\mathcal{G}} \setminus \mathcal{K}$. Additionally, for $v \in \tilde{\mathcal{G}} \setminus \mathcal{K}$ and $T = \inf\{t \geq 0 : \tilde{B}_t \in \mathcal{K}\}$,*

$$(17) \qquad \mathbb{E}[\tilde{\phi}_v \mid \mathcal{F}_{\mathcal{K}}] = \mathbb{E}_v[\tilde{\phi}_{\tilde{B}_T} \mathbb{1}_{T < \zeta} \mid \mathcal{F}_{\mathcal{K}}] = \sum_{u \in \partial \mathcal{K}} \mathrm{Hm}(v, u; \mathcal{K}) \tilde{\phi}_u.$$

Here, the harmonic measure Hm is given by $\mathrm{Hm}(v, u; \mathcal{K}) = \mathbb{P}_v(T < \zeta, \tilde{B}_T = u)$.

Note that if $v \in \mathcal{K}$, the harmonic measure is a point mass of mass 1 at v. To avoid having to account for this case separately, we will sometimes let the sum in (17) range over all $u \in \mathcal{K}$. The summation notation is justified since in either case the harmonic measure is supported on a finite number of points. Writing $\mathrm{Hm}_t(v, u)$ for $\mathrm{Hm}(v, u; \mathcal{I}_t)$ and

$$\mathrm{Hm}_t(A, u) = \frac{1}{|A|} \sum_{v \in A} \mathrm{Hm}_t(v, u),$$

we get from (17)

$$(18) \qquad M_{A,t} = \sum_{u \in \mathcal{I}_t} \mathrm{Hm}_t(A, u) \tilde{\phi}_u.$$

Writing G_t for the Green's function on $\tilde{\mathcal{G}} \setminus \mathcal{I}_t$. We get from Theorem 8 that

$$(19) \qquad \mathrm{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_t}] = \frac{1}{|A|^2} \sum_{v, v' \in A} G_t(v, v').$$

The following lemma establishes the continuity of M_A and will allow us to compute its quadratic variation. We defer the proof to the end of this section.

LEMMA 9. *$M_{A,t}$ and $\mathrm{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_t}]$ are almost-surely continuous as functions of t.*

Recall that for a continuous martingale M its quadratic variation $\langle M \rangle$ is the unique increasing continuous process vanishing at zero such that $M^2 - \langle M \rangle$ is a martingale (see [15], Theorem 1.3 in Chapter IV). From this, we deduce the following.

COROLLARY 10. *The quadratic variation of M_A is given by*

$$\langle M_A \rangle_t = \mathrm{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_0}] - \mathrm{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_t}].$$

PROOF. By definition the process $\mathrm{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_0}] - \mathrm{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_t}]$ is adapted to $\mathcal{F}_{\mathcal{I}_t}$ and vanishes at zero. It is increasing because $\mathcal{I}_t$ is increasing (so $\mathrm{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_t}]$ is decreasing), and it is continuous by Lemma 9. Therefore, by the characterization of $\langle M_A \rangle$ we only need to show that the process $\{Y_t : t \geq 0\}$ is a martingale, where

$$Y_t = M_{A,t}^2 - \mathrm{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_0}] + \mathrm{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_t}].$$

To this end, note that for any times $0 \le s < t$,

$$\mathbb{E}[M_{A,t}^2 \mid \mathcal{F}_{\mathcal{I}_s}] = M_{A,s}^2 + \mathbb{E}[(M_{A,t} - M_{A,s})^2 \mid \mathcal{F}_{\mathcal{I}_s}],$$

$$\mathbb{E}[\operatorname{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_t}] \mid \mathcal{F}_{\mathcal{I}_s}] = \mathbb{E}[(X_A - M_{A,t})^2 \mid \mathcal{F}_{\mathcal{I}_s}]$$

$$= \operatorname{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_s}] - \mathbb{E}[(M_{A,t} - M_{A,s})^2 \mid \mathcal{F}_{\mathcal{I}_s}].$$

We can then calculate $\mathbb{E}[Y_t \mid \mathcal{F}_{\mathcal{I}_s}]$ as follows:

$$\mathbb{E}[Y_t \mid \mathcal{F}_{\mathcal{I}_s}] = \mathbb{E}[M_{A,t}^2 \mid \mathcal{F}_{\mathcal{I}_s}] + \mathbb{E}[\operatorname{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_t}] \mid \mathcal{F}_{\mathcal{I}_s}] - \operatorname{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_0}]$$

$$= M_{A,s}^2 + \operatorname{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_s}] - \operatorname{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_0}]$$

$$= Y_s,$$

completing the verification that $\{Y_t : t \ge 0\}$ is a martingale. $\quad\square$

Next, we recall that the quadratic variation relates M to Brownian motion. In particular [15], Theorem 1.7 in Chapter V, stated below, gives the appropriate extension of the Dubins–Schwarz theorem for martingales of bounded quadratic variation.

THEOREM 11.   *Let be M a continuous martingale*, $T_t = \inf\{s : \langle M \rangle_s > t\}$, *and W be the following process*:

$$W_t = \begin{cases} M_{T_t} - M_0 & t < \langle M \rangle_\infty, \\ M_\infty - M_0 & t \ge \langle M \rangle_\infty. \end{cases}$$

Then W is a Brownian motion stopped at $\langle M \rangle_\infty$.

When applying this theorem, we will generally denote by B a Brownian motion which satisfies $B_t = M_{T_t} - M_0$ for $t < \langle M \rangle_\infty$ but is not stopped at $\langle M \rangle_\infty$, so that $W_t = B_{t \wedge \langle M \rangle_\infty}$. We note that by Theorem 8, $\{M_{A,t} - M_{A,0} : t \ge 0\}$ is independent of $\mathcal{F}_{\mathcal{I}_0}$, so we will take B to be independent of $\mathcal{F}_{\mathcal{I}_0}$ as well. Next, we present some formulas which will be used in the later sections to compute the quadratic variation of exploration martingales. First, we note that G_t may be written in terms of G and Hm_t as follows:

$$(20) \qquad G_t(v, v') = G(v, v') - \sum_{w \in \mathcal{I}_t} \operatorname{Hm}_t(v, w) G(w, v').$$

This can be seen, for example, by noting that for any $v' \in \tilde{\mathcal{G}}$, $G(v, v') - G_t(v, v')$ is harmonic in v for $v \notin \mathcal{I}_t$ and converges to $G(v, v')$ as $v \to \partial \mathcal{I}_t$ and that the same holds for $\sum_{w \in \mathcal{I}_t} \operatorname{Hm}_t(v, w) G(w, v')$. By similar reasoning, for $0 \le s \le t$ and $v, v' \in A$ we have the following two expressions for $G_s(v, v') - G_t(v, v')$:

$$G_s(v, v') - G_t(v, v') = \sum_{w \in \mathcal{I}_t} \operatorname{Hm}_t(v, w) G_s(w, v')$$

$$= \sum_{w \in \mathcal{I}_t} \operatorname{Hm}_t(v, w) G(w, v') - \sum_{w' \in \mathcal{I}_s} \operatorname{Hm}_s(v, w') G(w', v').$$

In later sections, it will be useful to consider the total mass of the harmonic measure. For a set $B \subset \tilde{\mathcal{G}}$, we let $\operatorname{Hm}(v, B) = \sum_{w \in B} \operatorname{Hm}(v, w; B)$. For a finite subset $A \subset \tilde{\mathcal{G}}$, we let $\operatorname{Hm}(A, B) = |A|^{-1} \sum_{v \in A} \operatorname{Hm}(v, B)$. To simplify notation, we will write $\pi_{A,t}$ for $\operatorname{Hm}(A, \mathcal{I}_t)$.

Finally, we prove that $M_{A,t}$ and $\operatorname{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_t}]$ are indeed continuous.

PROOF OF LEMMA 9. It suffices to show that for any $v, v' \in V$, $\mathbb{E}[\tilde{\phi}_v \mid \mathcal{F}_{\mathcal{I}_t}]$ and $\mathrm{Cov}[\tilde{\phi}_v \tilde{\phi}'_v \mid \mathcal{F}_{\mathcal{I}_t}]$ are continuous (then it is clear that $M_{A,t}$ and $\mathrm{Var}[X_A \mid \mathcal{F}_{\mathcal{I}_t}]$ are averages of a finite number of continuous functions and are thus continuous). Since both functions are constant for $t \geq D_h(\mathcal{I}_0, v)$, we let $0 \leq t \leq D_h(\mathcal{I}_0, v)$. By (17) and (20), it suffices to show that for any continuous function f on $\tilde{\mathcal{G}}$, the following function is continuous:

$$F(t) = \sum_{u \in \partial \mathcal{I}_t} \mathrm{Hm}_t(v, u) f(u).$$

Let D_{ℓ_1} denote the graph distance on $\tilde{\mathcal{G}}$ (i.e., ℓ_1 distance on $\mathbb{R}^d$). Since $K = |\partial \mathcal{I}_t| < \infty$,

$$\delta_1 = \min\{D_{\ell_1}(u, (\partial \mathcal{I}_t \cup V) \setminus \{u\}) : u \in \partial \mathcal{I}_t\} > 0.$$

For s such that $|t - s| < \delta_1/2$ and $u \in \partial \mathcal{I}_t$, let $\psi_s(u) = \partial \mathcal{I}_s \cap B_{\ell_1}(u, \delta_1/2)$ (here $B_{\ell_1}(u, \delta_1/2)$ is the open ball of radius $\delta_1/2$ around u with respect to D_{ℓ_1}). The sets $\psi_s(u)$ are nonempty and disjoint. If $s > t$, $\partial \mathcal{I}_s = \bigcup_{u \in \partial \mathcal{I}_t} \psi_s(u)$; if $s < t$, we let $\mathcal{R}_s = \partial \mathcal{I}_s \setminus (\bigcup_{u \in \partial \mathcal{I}_t} \psi_s(u))$. We have

$$|F(t) - F(s)| \leq \sum_{u \in \partial \mathcal{I}_t} \left| \mathrm{Hm}_t(v, u) f(u) - \sum_{u' \in \psi_s(u)} \mathrm{Hm}_s(v, u') f(u') \right|$$

$$+ \mathbb{1}_{s < t} \left| \sum_{u' \in \mathcal{R}_s} \mathrm{Hm}_s(v, u') f(u') \right|.$$

Since $\mathcal{I}_t$ is compact, $M = \max\{f(u) : u \in \mathcal{I}_t\} < \infty$. Since $\partial \mathcal{I}_t$ is finite, for any $\epsilon > 0$ there exists $0 < \delta_2 \leq \delta_1/2$ such that $|f(u) - f(u')| < \epsilon/2$ for $u \in \partial \mathcal{I}_t$ and $u' \in B_{\ell_1}(u, \delta_2)$. Thus, for $|t - s| < \delta_2$,

$$|F(t) - F(s)| \leq M \sum_{u \in \partial \mathcal{I}_t} |\mathrm{Hm}_t(v, u) - \mathrm{Hm}_s(v, \psi_s(u))| + \mathbb{1}_{s < t} M \mathrm{Hm}_s(v, \mathcal{R}_s) + \frac{\epsilon}{2}.$$

Finally, it follows from the construction of $\tilde{B}_t$ (by considering the excursions of a standard Brownian motion) that for any $u \in \tilde{\mathcal{G}}$, $b \leq D_{\ell_1}(u, V \setminus \{u\})$, and $u' \in \tilde{\mathcal{G}}$ such that $a = |u - u'|_{\ell_1} \leq b$,

$$\mathrm{Hm}(u, u'; \{u'\} \cup \partial B_{\ell_1}(u, b)) \geq \frac{b}{b + (2d - 1)a}.$$

Combining this with the previous bound, it follows from a straightforward calculation that there exists $\delta_3 \leq \delta_2$ such that if $|t - s| < \delta_3$,

$$|F(t) - F(s)| \leq \epsilon. \qquad \square$$

3. Percolation in three and higher dimensions.

In the case $d \geq 3$, we let the vertex set $V = \mathbb{Z}^d$ be the whole lattice and study the behavior of $p_{N,h} = \mathbb{P}(0 \xleftrightarrow{\geq h} \partial V_N)$ as $N \to \infty$. The theorems characterizing the behavior of $p_{N,h}$ will follow from Proposition 12, stated below. The idea is to consider a sequence of exploration martingales indexed by K, with source set $\mathcal{I}_0 = \{0\}$ and target set $A_K = \{x_K\}$, where $x_K \in \mathbb{Z}^d$ satisfies $|x_K| = K$, and let $K \to \infty$. We note that the calculations below are valid for any sequence $\{x_K\}_{K \geq 1}$ satisfying this condition so we choose one such sequence arbitrarily and consider it fixed for the rest of the section. For ease of notation, we will write M_K for the exploration martingale instead of M_{x_K}. Recall $\sigma_d^2 = G(0, 0)$ and note that by translation invariance $\sigma_d^2 = G(u, u)$ for all $u \in \mathbb{Z}^d$. Recall also the process $\pi_{K,t} = \sum_{v \in \mathcal{I}_t} \mathrm{Hm}_t(x_K, v)$.

PROPOSITION 12.   *Let B be a standard one-dimensional Brownian motion that is independent of $\tilde{\phi}_0$ and τ_h be the following stopping time*:

$$\tau_h = \inf\left\{t \geq 0 : B_t \leq ht - \frac{\tilde{\phi}_0 - h}{\sigma_d^2}\right\}.$$

We have

$$\lim_{K \to \infty} \mathbb{P}\left(\frac{\pi_{K,\infty} - \pi_{K,0}}{G(0, x_K)} \leq s\right) = \mathbb{P}(\tau_h \leq s) \quad \forall s \geq 0.$$

That is, $(\pi_{K,\infty} - \pi_{K,0})/G(0, x_K)$ converges in law to τ_h as $K \to \infty$.

Next, we state a few results which will be needed to prove Proposition 12. The proofs are deferred to the end of the section. First, we state a lemma that will allow us to relate the quadratic variation of M_K to the harmonic measure of $\mathcal{I}_t$.

LEMMA 13.   *Let $N \geq 1$ be given. Recall that for any subset $\mathcal{I} \subset \tilde{\mathcal{G}}$, $\mathrm{Hm}(x_K, \mathcal{I})$ denotes the probability that a metric graph Brownian motion started at x_K hits $\mathcal{I}$. The following holds uniformly over all compact, connected subsets $\mathcal{I} \subset \tilde{\mathcal{G}}$ satisfying $0 \in \mathcal{I} \subset \tilde{\mathcal{G}} \cap [-N, N]^d$*:

$$\frac{\mathrm{Var}[\tilde{\phi}_{x_K} \mid \mathcal{F}_0] - \mathrm{Var}[\tilde{\phi}_{x_K} \mid \mathcal{F}_\mathcal{I}]}{G(0, x_K)(\mathrm{Hm}(x_K, \mathcal{I}) - \mathrm{Hm}(x_K, 0))} = 1 + O\left(\frac{1}{K}\right) \quad \text{as } K \to \infty.$$

The second lemma gives upper and lower bounds for $(\pi_{K,t} - \pi_{K,0})/G(0, x_K)$ at the time $\mathcal{I}_t$ hits ∂V_N.

LEMMA 14.   *Let f_1 and f_2 be the following functions*:

$$f_1(N) = \limsup_{K \to \infty} \frac{\mathrm{Hm}(x_K, V_N) - \mathrm{Hm}(x_K, 0)}{G(0, x_K)},$$

$$f_2(N) = \liminf_{K \to \infty} \inf_{\mathcal{I}} \frac{\mathrm{Hm}(x_K, \mathcal{I}) - \mathrm{Hm}(x_K, 0)}{G(0, x_K)},$$

where the infimum in the definition of f_2 is taken over all compact, connected subsets $\mathcal{I} \subset \tilde{\mathcal{G}}$ such that $0 \in \mathcal{I}$ and $\mathcal{I} \cap \partial V_N \neq \varnothing$. We have

$$f_1(N) = O(N^{d-2}),$$

$$f_2(N) = \Omega\left(\frac{N}{\log(N)^{\mathbb{1}_{d=3}}}\right).$$

Finally, the following proposition gives the distribution of τ_h, and will allow us to obtain quantitative estimates from Proposition 12.

PROPOSITION 15 ([4], equation (2.0.2) in Part II).   *For $m \in \mathbb{R}$ and $b > 0$, let $\tau = \inf\{t > 0 : B_t \leq mt - b\}$. Then for $T > 0$,*

$$\mathbb{P}(\tau \leq T) = \bar{\Phi}\left(\frac{b}{\sqrt{T}} - m\sqrt{T}\right) + e^{2bm}\bar{\Phi}\left(\frac{b}{\sqrt{T}} + m\sqrt{T}\right).$$

We note that it follows from this proposition that the function $\mathbb{P}(\tau \leq T)$ is continuous in (m, b, T). We now turn to proving Proposition 12 using these three results.

PROOF OF PROPOSITION 12. First, we note that the case $s = 0$ is trivial, as $\{\pi_{K,\infty} = \pi_{K,0}\} = \{\tilde{\phi}_0 \leq h\} = \{\tau_h = 0\}$. Therefore, we will take $s > 0$ and assume that $\tilde{\phi}_0 > 0$. We let f_2 be as in Lemma 14, and N be such that $f_2(N) > 3s$. By Lemma 13, for any $0 < \epsilon < 1/2$ there exists K_0 such that for all $K \geq K_0$ the following holds almost surely. For all $0 \leq t \leq D_h(\mathbf{0}, \partial V_N)$,

$$(21) \qquad \frac{\pi_{K,t} - \pi_{K,0}}{G(\mathbf{0}, x_K)}(1 - \epsilon) \leq \frac{\langle M_K \rangle_t}{G(\mathbf{0}, x_K)^2} \leq \frac{\pi_{K,t} - \pi_{K,0}}{G(\mathbf{0}, x_K)}(1 + \epsilon).$$

Additionally, by our choice of N we can take K_0 large enough that for any $K \geq K_0$ and any compact connected set $\mathcal{I}$ connecting $\mathbf{0}$ to ∂V_N we have

$$\frac{\mathrm{Hm}(x_K, \mathcal{I}) - \mathrm{Hm}(x_K, \mathbf{0})}{G(\mathbf{0}, x_K)} > 2s.$$

By the definition of M_K (and the assumption $\tilde{\phi}_0 > h$), we have $M_K \geq h\pi_{K,t}$ for all $t \geq 0$, with equality if and only if $\pi_{K,t} = \pi_{K,\infty}$. That is, if and only if the explored set $\mathcal{I}_t$ has stopped growing (when viewed from infinity). Letting $\eta_h = \inf\{t \geq 0 : M_{K,t} = h\pi_{K,t}\} \wedge D_h(\mathbf{0}, \partial V_N)$ be the time the exploration stops or reaches ∂V_N, we have by our assumptions on K,

$$\left\{\frac{\pi_{K,\infty} - \pi_{K,0}}{G(\mathbf{0}, x_K)} \leq s\right\} = \left\{\frac{\pi_{K,\eta_h} - \pi_{K,0}}{G(\mathbf{0}, x_K)} \leq s\right\}$$

since $\pi_{K,\infty} - \pi_{K,0} \leq sG(\mathbf{0}, x_K)$ implies $D_h(\mathbf{0}, \partial V_N) = \infty$, and consequently $\pi_{K,\eta_h} = \pi_{K,\infty}$. Now, we let $T_s = \inf\{t \geq 0 : \langle M_K \rangle_t / G(\mathbf{0}, x_K)^2 > s\}$ and B_K be a standard Brownian motion satisfying $B_{K,s} = (M_{K,T_s} - M_{K,0})/G(\mathbf{0}, x_K)$ for $0 \leq s \leq \langle M_K \rangle_\infty / G(\mathbf{0}, x_K)^2$. As usual, we take B_K to be independent of $\tilde{\phi}_0$. We note that $\pi_{K,0}\sigma_d^2 = \mathrm{Hm}(x_K, \mathbf{0})G(\mathbf{0}, \mathbf{0}) = G(\mathbf{0}, x_k)$ which gives

$$\eta_h = \inf\left\{t \geq 0 : \frac{M_{K,t} - M_{K,0}}{G(\mathbf{0}, x_K)} \leq h\frac{\pi_{K,t} - \pi_{K,0}}{G(\mathbf{0}, x_K)} - \frac{\tilde{\phi}_0 - h}{\sigma_d^2}\right\} \wedge D_h(\mathbf{0}, \partial V_N).$$

We let $\hat{\tau}_{K,h} = \langle M_K \rangle_{\eta_h} / G(\mathbf{0}, x_K)^2$ and note that by (21)

$$\frac{\hat{\tau}_{K,h}}{1 + \epsilon} \leq \frac{\pi_{K,\eta_h} - \pi_{K,0}}{G(\mathbf{0}, x_K)} \leq \frac{\hat{\tau}_{K,h}}{1 - \epsilon}.$$

For $m \in \mathbb{R}$, we let $\tau_{K,m} = \inf\{t \geq 0 : B_{K,t} \leq mt - (\tilde{\phi}_0 - h)/\sigma_d^2\}$ and note that $\tau_{K,h}$ is equal in distribution to τ_h. We claim that

$$\tau_{K,h/(1+\epsilon)} \wedge s \leq \hat{\tau}_{K,h} \leq \tau_{K,h/(1-\epsilon)}, \quad h < 0,$$

$$\tau_{K,h/(1-\epsilon)} \wedge s \leq \hat{\tau}_{K,h} \leq \tau_{K,h/(1+\epsilon)}, \quad h \geq 0.$$

To see this, assume that $h \geq 0$ (the case $h < 0$ is similar) and let $S_t = \langle M_K \rangle_t / G(\mathbf{0}, x_k)^2$ be the inverse of T_s. Almost surely, S_t is continuous on the interval $[0, \eta_h]$. For $t < \eta_h$, we have by the definition of η_h,

$$\frac{M_{K,t} - M_{K,0}}{G(\mathbf{0}, x_K)} > h\frac{\pi_{K,t} - \pi_{K,0}}{G(\mathbf{0}, x_K)} - \frac{\tilde{\phi}_0 - h}{\sigma_d^2}.$$

By (21), this is equivalent to

$$B_{K,S_t} > \frac{h}{1 + \epsilon}S_t - \frac{\tilde{\phi}_0 - h}{\sigma_d^2} \quad \forall 0 \leq t < S_{\eta_h}.$$

By continuity of S, $\hat{\tau}_{K,h} = S_{\eta_h} \leq \tau_{K,h/(1+\epsilon)}$. For the lower bound, note that if $\eta_h = D_h(\mathbf{0}, \partial V_N)$ then by (21) and our assumptions on N and K,

$$\hat{\tau}_{K,h} \geq \frac{\pi_{K,D_h(\mathbf{0},\partial V_N)} - \pi_{K,0}}{G(\mathbf{0}, x_K)^2}(1-\epsilon) > s.$$

If $\eta_h < D_h(\mathbf{0}, \partial V_N)$, then arguing as above

$$B_{K,\hat{\tau}_{K,h}} = h\frac{\pi_{K,\eta_h} - \pi_{K,0}}{G(\mathbf{0}, x_K)^2} - \frac{\tilde{\phi}_{\mathbf{0}} - h}{\sigma_d^2} \leq \frac{h}{1-\epsilon}\hat{\tau}_{K,h} - \frac{\tilde{\phi}_{\mathbf{0}} - h}{\sigma_d^2}.$$

It follows that $\hat{\tau}_{K,h} \geq \tau_{K,h/(1-\epsilon)} \wedge s$ as desired.

Therefore, we conclude that for all $K \geq K_0$,

$$\mathbb{P}\left(\frac{\tau_{K,h/(1-\epsilon)}}{1-\epsilon} \leq s\right) \leq \mathbb{P}\left(\frac{\pi_{K,\eta_h} - \pi_{K,0}}{G(\mathbf{0}, x_K)} \leq s\right) \leq \mathbb{P}\left(\frac{\tau_{K,h/(1+\epsilon)}}{1+\epsilon} \leq s\right), \quad h < 0,$$

$$\mathbb{P}\left(\frac{\tau_{K,h/(1+\epsilon)}}{1-\epsilon} \leq s\right) \leq \mathbb{P}\left(\frac{\pi_{K,\eta_h} - \pi_{K,0}}{G(\mathbf{0}, x_K)} \leq s\right) \leq \mathbb{P}\left(\frac{\tau_{K,h/(1-\epsilon)}}{1+\epsilon} \leq s\right), \quad h \geq 0.$$

Letting $\epsilon \downarrow 0$ and noting that $\mathbb{P}(\tau_{K,m} \leq x)$ is continuous in (m, x) by Proposition 15 concludes the proof. $\square$

3.1. *Proof of main theorems.* In this subsection, we prove Theorems 3, 4 and 5. The following corollary of Proposition 12 will be used in each case.

COROLLARY 16. *Let f_1 and f_2 be as in Lemma 14, and τ_h be as in Proposition 12. We have*

$$\mathbb{P}(\tau_h \geq f_1(N)) \leq p_{N,h} \leq \mathbb{P}(\tau_h \geq f_2(N)).$$

PROOF. Let $f_{1,K}$ and $f_{2,K}$ be defined as follows:

$$f_{1,K} = \frac{\mathrm{Hm}(x_K, V_N) - \mathrm{Hm}(x_K, \mathbf{0})}{G(\mathbf{0}, x_K)}, \qquad f_{2,K} = \inf_{\mathcal{I}} \frac{\mathrm{Hm}(x_K, \mathcal{I}) - \mathrm{Hm}(x_K, \mathbf{0})}{G(\mathbf{0}, x_K)},$$

where as in Lemma 14, the infimum is over all compact connected subsets of $\tilde{\mathcal{G}}$ connecting $\mathbf{0}$ to ∂V_N. We have for all K,

$$\mathbb{P}\left(\frac{\pi_{K,\infty} - \pi_{K,0}}{G(\mathbf{0}, x_K)} \geq f_{1,K}(N)\right) \leq p_{h,N} \leq \mathbb{P}\left(\frac{\pi_{K,\infty} - \pi_{K,0}}{G(\mathbf{0}, x_K)} \geq f_{2,K}(N)\right).$$

Since $\mathbb{P}(\tau_h \leq s)$ is continuous in s by Proposition 15, taking the limit $K \to \infty$ gives the desired result. $\square$

PROOF OF THEOREM 3. For $h > 0$ and $b > 0$, we obtain by letting $T = \infty$ in Proposition 15 that

$$\mathbb{P}(B_t > -ht - b, \forall t \geq 0) = 1 - \exp(-2bh).$$

Since $f_2(N), f_1(N) \to \infty$ as $N \to \infty$, Corollary 16 gives

$$\lim_{N \to \infty} p_{N,-h} = \mathbb{P}(\tau_{-h} = \infty) = \mathbb{E}\left[(1 - e^{-2h(\tilde{\phi}_{\mathbf{0}}+h)/\sigma_d^2})\mathbb{1}_{\tilde{\phi}_{\mathbf{0}}>-h}\right]. \qquad \square$$

PROOF OF THEOREM 4. In this case, we use the following simple bound, which is a direct consequence of Proposition 15 and the fact that $\bar{\Phi}(x) \leq \exp(-x^2/2)$ for all $x \geq 0$,

$$\mathbb{P}(\tau_h \geq s \mid \mathcal{F}_{\mathbf{0}}) \leq \bar{\Phi}\left(h\sqrt{s} - \frac{\tilde{\phi}_{\mathbf{0}} - h}{\sigma_d^2\sqrt{s}}\right) \leq \exp\left(-\frac{h^2 s}{2} + \frac{h(\tilde{\phi}_{\mathbf{0}} - h)}{\sigma_d^2}\right).$$

This gives

$$p_{N,h} \le \mathbb{P}(\tau_h \ge f_2(N)) \le \exp\left(-\frac{h^2 f_2(N)}{2}\right) \mathbb{E}\left[\exp\left(\frac{h(\tilde{\phi}_0 - h)}{\sigma_d^2}\right)\right]$$

$$= \exp\left[-\Omega\left(\frac{h^2 N}{(\log N)^{\mathbb{1}_{d=3}}}\right)\right],$$

completing the proof of the theorem. $\square$

PROOF OF THEOREM 5.   As in the previous cases, we apply Proposition 15, Corollary 16 and Lemma 14 to obtain

$$p_{N,0} \ge \mathbb{P}(\tau_0 \ge f_1(N))$$

$$= \mathbb{E}\left[\left(\Phi\left(\frac{\tilde{\phi}_0}{\sigma_d^2 \sqrt{f_1(N)}}\right) - \Phi\left(-\frac{\tilde{\phi}_0}{\sigma_d^2 \sqrt{f_1(N)}}\right)\right)\mathbb{1}_{\tilde{\phi}_0 > 0}\right]$$

$$= \Omega\left(\frac{1}{N^{d/2-1}}\right).$$

The upper bound follows by the same reasoning

$$p_{N,0} \le \mathbb{P}(\tau_0 \ge f_2(N))$$

$$= \mathbb{E}\left[\left(\Phi\left(\frac{\tilde{\phi}_0}{\sigma_d^2 \sqrt{f_2(N)}}\right) - \Phi\left(-\frac{\tilde{\phi}_0}{\sigma_d^2 \sqrt{f_2(N)}}\right)\right)\mathbb{1}_{\tilde{\phi}_0 > 0}\right]$$

$$= O\left(\frac{\sqrt{\log N}^{\mathbb{1}_{d=3}}}{\sqrt{N}}\right),$$

completing the proof of the theorem. $\square$

3.2. *Critical window in three dimensions.*   In this section, we prove Theorem 6. That is, we take $d = 3$ and give rates of decay for h (now considered as a function of N) such that $p_{N,\pm h}$ is of the same order as $p_{N,0}$. Throughout, we let $h_N > 0$ be a sequence such that $h_N \to 0$. To simplify notation, we will write σ^2 for σ_3^2, $p_N^\pm$ for $p_{N,\pm h_N}$, and similarly with other quantities. Additionally, we write $(a)_+$ for $\max\{a, 0\}$ and will use $(\tilde{\phi}_0 \mp h_N)_+$ instead of $(\tilde{\phi}_0 \mp h_N)$ when applying Proposition 12 to avoid writing $\mathbb{1}_{\tilde{\phi}_0 > \pm h_N}$ when taking expectations.

We first prove (12). Letting $b = (\tilde{\phi}_0 - h_N)_+/\sigma^2$, we have from Proposition 16 and Proposition 15,

(22)
$$p_N^+ \ge \mathbb{E}\left[\bar{\Phi}\left(h_N \sqrt{f_1(N)} - \frac{b}{\sqrt{f_1(N)}}\right)\right.$$
$$\left. - e^{2h_N b}\bar{\Phi}\left(h_N \sqrt{f_1(N)} + \frac{b}{\sqrt{f_1(N)}}\right)\right].$$

To bound the right-hand side of the preceding inequality, we use the following lemma.

LEMMA 17.   *Let $x \in \mathbb{R}$ and $y \ge 0$. Define f and g by*

$$f(x, y) = \bar{\Phi}(x - y) - e^{2xy}\bar{\Phi}(x + y) \quad and \quad g(x, y) = 1 - \frac{x\bar{\Phi}(x + y)}{\varphi(x + y)}.$$

We have $g(x, y) > 0$ and

(23)
$$2g(x, 0)\varphi(x)(y \wedge x) \le f(x, y) \le 2g(x, y)\varphi((x - y)_+)y, \quad x \ge 0,$$
$$2g(x, y)\varphi(x - y)y \le f(x, y) \le 2g(x, 0)\varphi(x)y, \quad x \le 0.$$

PROOF.   It is clear that $g(x, y) > 0$ for $x \leq 0$. For $x > 0$, the fact that $g(x, 0) > 0$ is equivalent to the well-known (and straightforward to check) bound $\bar{\Phi}(x) < \varphi(x)/x$ for all $x > 0$ and it directly implies $g(x, y) \geq g(x + y, 0) > 0$.

To prove (23), note that

$$\frac{\partial f}{\partial y}(x, y) = 2\varphi(x - y) - 2xe^{2xy}\bar{\Phi}(x + y) = 2g(x, y)\varphi(x - y).$$

Using the fact that $\bar{\Phi}(x)/\varphi(x)$ is decreasing in x (for all values of x), we conclude that $g(x, y)$ is decreasing in y for $x < 0$ and increasing in y for $x > 0$. The desired bounds follow by integrating $\partial f/\partial y$. For instance, for $x > 0$ we have

$$f(x, y) = \int_0^y \frac{\partial f}{\partial y}(x, s) \, ds \leq 2g(x, y) \int_0^y \varphi(x - s) \, ds$$

$$= 2g(x, y)[\Phi(x) - \Phi(x - y)]$$

$$\leq 2g(x, y)\varphi((x - y)_+)y$$

The other three bounds follow by similar arguments.   □

Note now that $h_N = O(N^{-1/2})$ implies $h_N\sqrt{f_1(N)} = O(1)$ and, therefore, $g(h_N\sqrt{f_1(N)}, 0) = \Omega(1)$. Recalling (22) and applying Lemma 17 with $x = h_N\sqrt{f_1(N)}$ and $y = b/\sqrt{f_1(N)}$ give

$$p_N^+ \geq 2g(h_N\sqrt{f_1(N)}, 0)\varphi(h_N\sqrt{f_1(N)})\mathbb{E}\left[\frac{b}{\sqrt{f_1(N)}} \wedge h_N\sqrt{f_1(N)}\right]$$

$$= \Omega\left(\frac{1}{\sqrt{N}}\right),$$

which proves (12).

We turn next to (13). We let b be as above, and note the trivial bound $g(x, y) \leq 1$ for $x \geq 0$. Combining this with Lemma 17, we obtain

$$p_N^+ \leq \mathbb{P}(\tau_N^+ \geq f_2(N))$$

$$\leq \mathbb{E}\left[\frac{2b}{\sqrt{f_2(N)}}\varphi\left(\left(h_N\sqrt{f_2(N)} - \frac{b}{\sqrt{f_2(N)}}\right)_+\right)\right]$$

$$\leq \frac{\exp(-h_N^2 f_2(N)/2)}{\sqrt{f_2(N)}}\mathbb{E}[be^{h_N b}].$$

Recalling $f_2(N) = \Omega(N/\log N)$, we have under the assumption

$$\liminf_{N\to\infty} \frac{h_N\sqrt{N}}{\sqrt{\log N \log\log N}} \geq C$$

for a large enough constant C, that $e^{-h_N^2 f_2(N)/2} = o(\sqrt{\log N})$. This gives $p_N^+ = o(\frac{1}{\sqrt{N}})$ as required for (13).

The bounds on p_N^- are obtained by a similar argument. Let $b = (\tilde{\phi}_0 + h_N)_+/\sigma^2$, and note that for $x \geq 0$ we have $g(-x, 0) \leq 1 + x/\varphi(x)$. Applying Lemma 17 with $x = h_N\sqrt{f_2(N)}$ and $y = b/\sqrt{f_2(N)}$, this gives

$$p_N^- \leq \mathbb{P}(\tau_N^- \geq f_2(N))$$

$$\leq 2\frac{\varphi(h_N\sqrt{f_2(N)}) + h_N\sqrt{f_2(N)}}{\sqrt{f_2(N)}}\mathbb{E}[b]$$

$$= O(h_N + f_2(N)^{-1/2}).$$

Recalling $f_2(N) = \Omega(N/\log(N))$, we see that if $h_N = O(\sqrt{\log N/N})$, then the above implies $p_N^- = O(\sqrt{\frac{\log N}{N}})$, as required for (14).

Conversely, we have for $x, y \geq 0$ that $g(-x, y) \geq x\bar{\Phi}(y)/\varphi(-x+y)$ so Lemma 17 gives

$$\bar{\Phi}(-x-y) - e^{-2xy}\bar{\Phi}(-x+y) \geq 2xye^{-2xy}\bar{\Phi}(y).$$

Letting $x = h_N\sqrt{f_1(N)}$ and $y = b/\sqrt{f_1(N)}$, we obtain

$$p_N^- \geq \mathbb{P}(\tau_N^- \geq f_1(N))$$

$$\geq \mathbb{E}\left[2h_N b e^{-2h_N b}\bar{\Phi}\left(\frac{b}{\sqrt{f_1(N)}}\right)\right].$$

It follows that if $h_N = o(1)$, $p_N^- = \Omega(h_N)$. In particular, $h_N = \omega(\sqrt{\log N/N})$ implies $p_N^- = \omega(\sqrt{\log N/N})$, as required for (15).

3.3. *Proof of technical lemmas.*

PROOF OF LEMMA 13. By [10], Theorem 4.3.1, there exists a constant c_d such that the following holds:

(24) $$G(\mathbf{0}, x) = \frac{c_d}{|x|^{d-2}} + O\left(\frac{1}{|x|^d}\right).$$

Since $G(x, y) = G(\mathbf{0}, y - x)$ for $x, y \in \mathbb{Z}^d$, we can deduce that for any point $v \in V_N$,

$$|G(v, x_K) - G(\mathbf{0}, x_K)| = G(\mathbf{0}, x_K)O\left(\frac{|v|}{K}\right) \quad \text{as } K \to \infty.$$

By this, we mean that the suppressed constant does not depend on v or K. We note that by (3) this bound extends to $v \in \tilde{\mathcal{G}} \cap [-N, N]^d$, and in particular to points v on edges incident to $\mathbf{0}$ (i.e. such that $|v| < 1$). From this, we obtain

$$\text{Var}[\tilde{\phi}_{x_K} \mid \mathcal{F}_\mathcal{I}] = G(x_K, x_K) - \sum_{v \in \partial\mathcal{I}} \text{Hm}(x_K, v; \mathcal{I})G(v, x_K)$$

$$= G(x_K, x_K) - G(\mathbf{0}, x_K)\text{Hm}(x_K, \mathcal{I})$$

$$+ G(\mathbf{0}, x_K)\sum_{v \in \partial\mathcal{I}} \text{Hm}(x_K, v; \mathcal{I})O\left(\frac{|v|}{K}\right).$$

Similarly, we have $\text{Var}[\tilde{\phi}_{x_K} \mid \mathcal{F}_\mathbf{0}] = G(x_K, x_K) - G(\mathbf{0}, x_K)\text{Hm}(x_K, \mathbf{0})$. Therefore,

$$\frac{\text{Var}[\tilde{\phi}_{x_K} \mid \mathcal{F}_\mathbf{0}] - \text{Var}[\tilde{\phi}_{x_K} \mid \mathcal{F}_\mathcal{I}]}{G(\mathbf{0}, x_K)} = \text{Hm}(x_K, \mathcal{I}) - \text{Hm}(x_K, \mathbf{0})$$

$$+ \sum_{v \in \partial\mathcal{I}} \text{Hm}(x_K, v; \mathcal{I})O\left(\frac{|v|}{K}\right).$$

To conclude, we bound the last term above by the difference $\text{Hm}(x_K, \mathcal{I}) - \text{Hm}(x_K, \mathbf{0})$. Let $\tau_\mathbf{0} = \inf\{t \geq 0 : \tilde{B}_t = \mathbf{0}\}$ be the hitting time of $\mathbf{0}$ by a metric graph Brownian motion. We have

$$\text{Hm}(x_K, \mathcal{I}) - \text{Hm}(x_K, \mathbf{0}) = \sum_{v \in \partial\mathcal{I}} \text{Hm}(x_K, v; \mathcal{I})\mathbb{P}_v(\tau_\mathbf{0} = \infty)$$

$$= \sum_{v \in \partial\mathcal{I}} \text{Hm}(x_K, v; \mathcal{I})\Omega(|v| \wedge 1),$$

where we have used the fact that there exists a constant $p_d > 0$ such that $\mathbb{P}_v(\tau_{\mathbf{0}} = \infty) \geq p_d$ for all $v \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$. It follows that

$$\frac{\text{Var}[\tilde{\phi}_{x_K} \mid \mathcal{F}_{\mathbf{0}}] - \text{Var}[\tilde{\phi}_{x_K} \mid \mathcal{F}_{\mathcal{I}}]}{G(\mathbf{0}, x_K)(\text{Hm}(x_K, \mathcal{I}) - \text{Hm}(x_K, \mathbf{0}))} = 1 + O\left(\frac{1}{K}\right). \qquad \square$$

PROOF OF LEMMA 14.   Both bounds are proved by similar arguments so we only provide the details for the bound on $f_2(N)$. First, note that $\text{Hm}(x_K, \mathbf{0})/G(\mathbf{0}, x_K) = 1/\sigma_d^2$ so it suffices to show $\text{Hm}(x_K, \mathcal{I})/G(\mathbf{0}, x_K) = \Omega(N/(\log N)^{\mathbb{1}_{d=3}})$ uniformly over compact connected sets $\mathcal{I}$ containing $\mathbf{0}$ and intersecting ∂V_N and K large enough (say $K \geq 100N$). Begin by noting that any such $\mathcal{I}$ contains a set $\mathcal{U} = \{u_j\}_{j=0}^N$ where $u_j \in \mathbb{Z}^d$ and $|u_j|_\infty = j$ so it suffices to lower bound $\text{Hm}(x_K, \mathcal{U})$. To this end, let Y be the number of visits to $\mathcal{U}$ by a random walk started at x_K. We have

$$\text{Hm}(x_K, \mathcal{U}) = \mathbb{P}(Y > 0) = \frac{\mathbb{E}[Y]}{\mathbb{E}[Y \mid Y > 0]}.$$

By (24), we have

$$\mathbb{E}[Y] = \sum_{u \in \mathcal{U}} G(u, x_K) = \Omega(N)G(\mathbf{0}, x_K)$$

On the other hand, if we let S be a simple random walk on $\mathbb{Z}^d$ and $\tau_{\mathcal{U}} = \inf\{n \geq 1 : S_n \in \mathcal{U}\}$ be the hitting time of $\mathcal{U}$, the following holds uniformly over $u \in \mathcal{U}$,

$$\mathbb{E}[Y \mid S_{\tau_{\mathcal{U}}} = u] = \sum_{u' \in \mathcal{U}} G(u, u') = O\left(\sum_{j=1}^N \frac{1}{j^{d-2}}\right) = O(\log(N)^{\mathbb{1}_{d=3}}).$$

The result follows by noting that

$$\mathbb{E}[Y \mid Y > 0] = \frac{1}{\text{Hm}(x_K, \mathcal{U})} \sum_{u \in \mathcal{U}} \mathbb{E}[Y \mid S_{\tau_{\mathcal{U}}} = u] \text{Hm}(x_K, u). \qquad \square$$

4. Chemical distance in two dimensions.   This section is devoted to the proof of Theorem 1. Recall that $\tilde{\phi}_N$ is the Gaussian free field on the metric graph of V_N with Dirichlet boundary conditions, and that G_N is Green's function on V_N as in (2). Our goal is prove that for any constants $0 < \alpha < \beta < \gamma < 1$, the chemical distance $D_{N,h}(V_{\alpha N}, \partial V_{\beta,N})$ is of order at most $N(\log N)^{1/4}$, conditioned on the event that $V_{\alpha N}$ is connected to $\partial V_{\gamma N}$.

The proof employs the same type of exploration martingale as in the case of $d \geq 3$. Below we prove Theorem 1 while postponing proofs of a few lemmas to later subsections.

PROOF OF THEOREM 1.   For $h \in \mathbb{R}$ and $0 < \alpha < \gamma < 1$, define $\mathcal{E}_{N,1} = \{D_{N,h}(V_{\alpha N}, \partial V_{\gamma N}) < \infty\}$. That is, $\mathcal{E}_{N,1}$ is the event that $V_{\alpha N}$ is connected to $\partial V_{\gamma,N}$ in $\tilde{E}_N^{\geq h}$.

LEMMA 18.   *We have*

$$c_1 = \inf\{\mathbb{P}(\mathcal{E}_{N,1}) : N \geq 1\} > 0,$$

where c_1 depends on h, α and γ.

REMARK 19.   Despite the fact that the statement of Lemma 18 is formally slightly stronger than [6], Proposition 4, (since percolation on metric graph is a sub-event of percolation on discrete lattice), the proof of [6], Proposition 4, adapts with essentially no change. Thus, we omit further details of the proof.

Now, let $\mu = (1 + \gamma)/2$ and $M_{\mu N}$ be the exploration martingale with target set $\partial V_{\mu N}$ and source set $\mathcal{I}_0 = V_{\alpha N}$, as defined in (16). That is to say,

$$X_{\mu N} = \frac{1}{|\partial V_{\mu N}|} \sum_{v \in \partial V_{\mu N}} \tilde{\phi}_{N,v} \quad \text{and} \quad M_{\mu N,t} = \mathbb{E}[X_{\mu N} \mid \mathcal{F}_{\mathcal{I}_t}],$$

where $\mathcal{I}_t = \{v \in V_N : D_{N,h}(V_{\alpha N}, v) \leq t\}$. From now on, we take N large enough that the boxes $V_{\alpha N}, V_{\beta N}, V_{\gamma N}, V_{\mu N}$ are distinct. For $t \geq 0$, we let $\partial \mathcal{I}_t^+ = \partial \mathcal{I}_t \cap \tilde{E}_N^{>h}$ be the points on $\partial \mathcal{I}_t$ where $\tilde{\phi}$ is strictly above h, which we will refer to as the active points at time t (by active here we mean that these are the points from which the metric ball exploration can proceed further), and let $\partial \mathcal{I}_t^- = \partial \mathcal{I}_t \setminus \tilde{E}_N^{\geq h}$ be the points on $\partial \mathcal{I}_t$ where $\tilde{\phi}$ is strictly below h. We then define the "positive" and "negative" parts of $M_{\mu N}$ (which we denote by $M_{\mu N}^{\pm}$) as

$$M_{\mu N,t}^{\pm} = \sum_{u \in \partial \mathcal{I}_t^{\pm}} \mathrm{Hm}_{N,t}(\partial V_{\mu N}, u)(\tilde{\phi}_u - h),$$

where $\mathrm{Hm}_{N,t}(v, u) = \mathrm{Hm}(v, u; \mathcal{I}_t \cup \partial V_N)$. We note that $\partial \mathcal{I}_t^- = \partial \mathcal{I}_0^-$ for all t, which, combined with the fact $\mathcal{I}_t$ is increasing, implies $M_{\mu N}^-$ is increasing. For $c \in \mathbb{R}$, define

$$\mathcal{E}_{N,2}(c) = \{M_{\mu N,t}^+ \geq c \text{ for all } 0 \leq t \leq D_{N,h}(V_{\alpha N}, \partial V_{\beta N})\}.$$

LEMMA 20. *There exists a constant $c_2 = c_2(h, \beta, \gamma) > 0$ such that*

$$\mathbb{P}(\mathcal{E}_{N,2}(c_2 \epsilon) \mid \mathcal{E}_{N,1}) \geq 1 - \epsilon \quad \text{for all } \epsilon > 0.$$

For the rest of the section, we let $\mathcal{E}_{N,2} = \mathcal{E}_{N,2}(c_2 \epsilon / 2)$ for convenience. The core idea in proving Theorem 1 is to bound from below the rate at which the quadratic variation increases as a function of $M_{\mu N,t}^+$. Combined with an upper bound on $\langle M_{\mu N} \rangle_{D_{N,h}(V_{\alpha N}, \partial V_{\beta N})}$, this then yields an upper bound on $D_{N,h}(V_{\alpha N}, \partial V_{\beta N})$. In order to carry out the proof, we first give the upper bound on the quadratic variation of $M_{\mu N}$ (which is easier than the lower bound).

LEMMA 21 ([6], Lemma 2). *For $0 < \mu < 1$, there exist constants $c, c' > 0$ such that*

$$\sum_{v \in \partial V_{\mu N}} G_N(u, v) \leq cN \quad \forall u \in \partial V_{\mu N}; \tag{25}$$

$$G_N(u, v) \geq c' \quad \forall u, v \in V_{\mu N}. \tag{26}$$

By (25) and Corollary 10, we get that for some constant $c_3 = c_3(\mu) > 0$,

$$\langle M_{\mu N} \rangle_\infty \leq \frac{1}{|\partial V_{\mu N}|^2} \sum_{v, v' \in \partial V_{\mu N}} G(v, v') \leq c_3. \tag{27}$$

The remaining main task for proving Theorem 1 is to show that on some event $\mathcal{E}_N^*$ with $\mathbb{P}(\mathcal{E}_N^* \mid \mathcal{E}_{N,1}) \geq 1 - \epsilon$, we have

$$\langle M_{\mu N} \rangle_{D_{N,h}(V_{\alpha N}, \partial V_{\beta N})} \geq \kappa \frac{D_{N,h}(V_{\alpha N}, \partial V_{\beta N})^2}{N^2 \sqrt{\log N}}, \tag{28}$$

for some constant $\kappa = \kappa(\epsilon, \alpha, \beta, \gamma, h) > 0$. Indeed, assuming (28), we can then combine it with (27), and conclude that on $\mathcal{E}_N^*$,

$$D_{N,h}(V_{\alpha N}, \partial V_{\beta N}) \leq c_3 \kappa^{-1/2} N (\log N)^{\frac{1}{4}},$$

completing the proof of Theorem 1.

It remains to show (28). To this end, we first bound the quadratic variation from below in terms of the ℓ_2-norm of the harmonic measure on the active points, as in the next lemma.

LEMMA 22. *There exists a constant $c_4 > 0$ such that the following holds almost surely for all integers $K \geq 1$:*

$$\langle M_{\mu N}\rangle_K \geq c_4 \sum_{k=1}^{K} \sum_{u \in \partial \mathcal{I}_k^+} \mathrm{Hm}_{N,k}(\partial V_{\mu N}, u)^2.$$

Note that the chemical distance $D_{N,h}(V_{\alpha N}, \partial V_{\beta N})$ is an integer whenever it is finite since (by our convention as in (4)) $V_{\alpha N}, \partial V_{\beta N} \subset \mathbb{Z}^2$. We let $\mathcal{I}^+ = \bigcup_{k=1}^{D_{N,h}(V_{\alpha N}, \partial V_{\beta N})} \partial \mathcal{I}_k^+$ and since the sets $\{\partial \mathcal{I}_k^+ : k \geq 1\}$ are disjoint, for $u \in \mathcal{I}^+$ we can define

$$W(u) = \mathrm{Hm}_{N,k}(\partial V_{\mu N}, u),$$

where k is the unique positive integer such that $u \in \partial \mathcal{I}_k^+$. We rewrite the conclusion of Lemma 22 as

(29) $$\langle M_{\mu N}\rangle_{D_N(V_{\alpha N}, \partial V_{\beta N})} \geq c_4 \sum_{u \in \mathcal{I}^+} W(u)^2.$$

In order to bound the right-hand side of (29) from below by $M_{\mu N}^+$, we need some control on the empirical profile of $\{\tilde{\phi}_{N,v} : v \in \mathcal{I}^+\}$. To this end, we define

$$\mathcal{B}_0 = \{u \in \mathcal{I}^+ : \tilde{\phi}_{N,u} - h \leq \sqrt{\log N}\},$$

$$\mathcal{B}_j = \{u \in \mathcal{I}^+ : 2^{j-1}\sqrt{\log N} < \tilde{\phi}_{N,u} - h \leq 2^j \sqrt{\log N}\}, \quad j \geq 1.$$

Here, the scale $\sqrt{\log N}$ is chosen to match the order of $\sqrt{\mathbb{E}[\tilde{\phi}_u^2]}$ for $u \in V_{\beta N}$. Letting $\mathcal{W}_j = \sum_{u \in \mathcal{B}_j} W(u)$, we get from the Cauchy–Schwartz inequality that

(30) $$\sum_{u \in \mathcal{I}^+} W(u)^2 \geq \sum_{j=0}^{\infty} \frac{\mathcal{W}_j^2}{|\mathcal{B}_j|},$$

where we use the convention $0/0 = 0$. The appearance of $|\mathcal{B}_j|$ in the denominator in the preceding inequality calls for an upper bound on $|\mathcal{B}_j|$, as incorporated in the next lemma (the reason for the specific form of the bound will be made clear below).

LEMMA 23. *Let f be a function on the positive integers such that $f(j) = e^{O(j)}$. Then there exists $c_5 = c_5(\alpha, \beta, h, f) > 0$ such that*

$$\sup\{f(j)\mathbb{E}[|\mathcal{B}_j|] : j \geq 0\} \leq c_5 \frac{N^2}{\sqrt{\log N}}.$$

We are now ready to give a lower bound on $\langle M_{\mu N}\rangle_{D_{N,h}(V_{\alpha N}, \partial V_{\beta N})}$. By the definition of $\mathcal{W}_j$, we have

$$\mathcal{W}_j \geq 2^{-j}(\log N)^{-\frac{1}{2}} \sum_{u \in \mathcal{B}_j} W(u)(\tilde{\phi}_{N,u} - h).$$

In addition, on the event $\mathcal{E}_{N,1} \cap \mathcal{E}_{N,2}$ we have

$$\sum_{j=1}^{\infty} \sum_{u \in \mathcal{B}_j} W(u)(\tilde{\phi}_{N,u} - h) = \sum_{k=1}^{D_{N,h}(V_{\alpha N}, \partial V_{\beta N})} M_{\mu N,k}^+$$

$$\geq \frac{c_2 \epsilon D_{N,h}(V_{\alpha N}, \partial V_{\beta N})}{2}.$$

Letting $c_6 = 6/\pi^2$ so that $c_6 \sum_{j=0}^{\infty}(j+1)^{-2} = 1$, we see that

$$(31) \qquad \mathcal{E}_{N,1} \cap \mathcal{E}_{N,2} \subseteq \bigcup_{j=0}^{\infty} \left\{ \mathcal{W}_j \geq \frac{c_6 c_2 \epsilon D_{N,h}(V_{\alpha N}, \partial V_{\beta N})}{2^{j+1}(j+1)^2 \sqrt{\log N}} \right\}.$$

Letting $c_7 = 20/(c_1 \epsilon)$, we define $\mathcal{E}_{N,3} = \bigcap_{j=0}^{\infty} \mathcal{E}_{N,3,j}$, where

$$\mathcal{E}_{N,3,j} = \{ |\mathcal{B}_j| \leq c_7 \mathbb{E}[|\mathcal{B}_j|](1.1)^{j+1} \}, \quad j \geq 0.$$

By Markov's inequality, we get that

$$(32) \qquad \mathbb{P}(\mathcal{E}_{N,3}^c \mid \mathcal{E}_{N,1}) \leq \frac{\mathbb{P}(\mathcal{E}_{N,3}^c)}{c_1} \leq \frac{\epsilon}{2}.$$

Let $\mathcal{E}_N^* = \mathcal{E}_{N,1} \cap \mathcal{E}_{N,2} \cap \mathcal{E}_{N,3}$. By Lemma 20 and (32), we get that

$$(33) \qquad \mathbb{P}(\mathcal{E}_N^* \mid \mathcal{E}_{N,1}) \geq 1 - \epsilon.$$

We deduce from (29), (30) and (31) that on $\mathcal{E}_N^*$,

$$\langle M_{\mu N} \rangle_{D_{N,h}(V_{\alpha N}, \partial V_{\beta N})} \geq \frac{c_6^2 c_4 c_2^2 c_1 \epsilon^3}{20} \inf_{j \geq 0} \frac{D_{N,h}(V_{\alpha N}, \partial V_{\beta N})^2}{(4.4)^{j+1}(j+1)^4 \mathbb{E}[|\mathcal{B}_j|] \log N}.$$

Combined with Lemma 23, this gives that on $\mathcal{E}_N^*$,

$$\langle M_{\mu N} \rangle_{D_{N,h}(V_{\alpha N}, \partial V_{\beta N})} \geq \frac{c_6^2 c_4 c_2^2 c_1 \epsilon^3}{20 c_5} \frac{D_{N,h}(V_{\alpha N}, \partial V_{\beta N})^2}{N^2 \sqrt{\log N}}.$$

Combining with (33), we complete the verification of (28) as promised. $\quad \square$

4.1. *Proof of Lemma* 20. We first give the main intuition behind the proof of Lemma 20 in the case when $h = 0$. On the event $\mathcal{E}_{N,1} \cap \mathcal{E}_{N,2}(\epsilon)^c$, we have $M_{\mu N,s} \leq \epsilon + M_{\mu N,s}^-$ for some $s \leq D_{N,h}(V_{\alpha N}, \partial V_{\beta N})$. However, we also have $M_{\mu N,t} \geq M_{\mu N,s}^-$ for all $t \geq s$. Since $D_{N,h}(V_{\alpha N}, \partial V_{\gamma N}) < \infty$ on $\mathcal{E}_{N,1}$, the martingale must stay above $M_{\mu N,s}^-$ after time s and yet accumulate an order 1 amount of quadratic variation—this happens with small probability. The case for general h is similar but a bit more complicated. We carry out a detailed proof below.

In this subsection and the ones that follow, we let $c, c', c'' > 0$ be arbitrary constants whose values may change each time they appear, and may depend on h, α, β, γ but not on N. Since $M_{\mu N,t}^-$ is increasing in t (recalling $\partial \mathcal{I}_t^- = \partial \mathcal{I}_0^-$ for all t), we have that for $0 \leq s < t < D_{N,h}(V_{\alpha N}, \partial V_{\mu N})$,

$$M_{\mu N,t} - M_{\mu N,s} \geq h[\pi_{\mu N,t} - \pi_{\mu N,s}] - M_{\mu N,s}^+,$$

with equality if and only if $M_{\mu N,t}^+ = M_{\mu N,s}^+ = 0$, where $\pi_{\mu N,t} = \mathrm{Hm}(\partial V_{\mu N}, \mathcal{I}_t)$. Next, we claim that there exist constants c, c' such that for any $0 \leq s < t < D_{N,h}(V_{\alpha N}, \partial V_{\gamma N})$,

$$c'[\pi_{\mu N,t} - \pi_{\mu N,s}] \leq \langle M_{\mu N} \rangle_t - \langle M_{\mu N} \rangle_s,$$

$$\langle M_{\mu N} \rangle_t - \langle M_{\mu N} \rangle_s \leq c[\pi_{\mu N,t} - \pi_{\mu N,s}].$$

To see this, note that we have

$$\langle M_{\mu N} \rangle_t - \langle M_{\mu N} \rangle_s = \frac{1}{|\partial V_{\mu N}|} \sum_{w \in \partial \mathcal{I}_t} \mathrm{Hm}_{N,t}(\partial V_{\mu N}, w) G_{N,s}(w, \partial V_{\mu N}),$$

$$\pi_{\mu N,t} - \pi_{\mu N,s} = \sum_{w \in \partial \mathcal{I}_t} \mathrm{Hm}_{N,t}(\partial V_{\mu N}, w) \mathbb{P}_w(\tau_{\partial V_N} < \tau_{\mathcal{I}_s}),$$

where $G_{N,s}$ is the Green's function on $\tilde{\mathcal{G}}_N \setminus \mathcal{I}_s$,

$$G_{N,s}(w, \partial V_{\mu N}) = \sum_{v \in \partial V_{\mu N}} G(w, v),$$

and τ_A is as usual the hitting time of A by a metric graph Brownian motion. By Lemma 21 and the assumption that $\mathcal{I}_s \subset V_{\gamma N}$,

$$c' \mathbb{P}_w(\tau_{\partial V_{\mu N}} < \tau_{\mathcal{I}_s}) \leq \frac{G_{N,s}(w, \partial V_{\mu N})}{|\partial V_{\mu N}|} \leq c \mathbb{P}_w(\tau_{\partial V_{\mu N}} < \tau_{\mathcal{I}_s}).$$

Similarly, we have by straightforward random walk considerations (namely the invariance principle)

$$c' \mathbb{P}_w(\tau_{\partial V_{\mu N}} < \tau_{\mathcal{I}_s}) \leq \mathbb{P}_w(\tau_{\partial V_N} < \tau_{\mathcal{I}_s}) \leq \mathbb{P}_w(\tau_{\partial V_{\mu N}} < \tau_{\mathcal{I}_s}).$$

Thus, the upper and lower bounds on the quadratic variation follow. Altogether, this implies

$$(34) \qquad M_{\mu N, t} - M_{\mu N, s} > ch[\langle M_{\mu N} \rangle_t - \langle M_{\mu N} \rangle_s] - M_{\mu N, s}^+,$$

for all $0 \leq s < t \leq D_{N,h}(V_{\alpha N}, \partial V_{\gamma N})$ such that $M_{\mu N, s}^+ > 0$. For any $x > 0$, let $\eta_x = \inf\{t : M_{\mu N, t}^+ \leq x\}$ and define the martingale $\tilde{M}_{\mu N}^x$ (with respect to $\mathcal{G}_t = \mathcal{F}_{\mathcal{I}_{\eta_x + t}}$) by

$$\tilde{M}_{\mu N, t}^x = \begin{cases} M_{\mu N, \eta_x + t} - M_{\mu N, \eta_x} & \eta_x < \infty, \\ 0 & \eta_x = \infty. \end{cases}$$

Let $\Delta = D_{N,h}(V_{\alpha N}, \partial V_{\gamma N}) - \eta_x$ and note that on $\mathcal{E}_{N,1} \cap \mathcal{E}_{N,2}(x)^c$ we get from (34) that

$$\tilde{M}_{\mu N, t}^x > ch \langle \tilde{M}_{\mu N}^x \rangle_t - x, \quad 0 \leq t < \Delta.$$

Using the lower bound on the quadratic variation by the Harmonic measure that was proven above, we see that the following bound holds almost surely on $\mathcal{E}_{N,1} \cap \mathcal{E}_{N,2}(x)^c$:

$$\langle \tilde{M}_{\mu N}^x \rangle_\Delta \geq \langle M_{\mu N} \rangle_{D_{N,h}(V_{\alpha N}, \partial V_{\gamma N})} - \langle M_{\mu N} \rangle_{D_{N,h}(V_{\alpha N}, \partial V_{\beta N})}$$

$$\geq c'[\pi_{\mu N, D_{N,h}(V_{\alpha N}, \partial V_{\gamma N})} - \pi_{\mu N, D_{N,h}(V_{\alpha N}, \partial V_{\beta N})}] \geq c',$$

where the last inequality follows from a simple adaptation of the proof of [6], Proposition 4.

Write $T_t = \inf\{s : \langle \tilde{M}_{\mu N}^x \rangle_s > t\}$ and let B be a standard Brownian motion that satisfies $B_t = \tilde{M}_{\mu N, T_t}^x - \tilde{M}_{\mu N, 0}^x$ for $t < \langle \tilde{M}_{\mu N}^x \rangle_\infty$ and is independent of $\mathcal{F}_{\mathcal{I}_{\eta_x}}$. Letting $\tau_{h,x} = \inf\{t : B_t \leq cht - x\}$, it follows from Proposition 15 and Lemma 17 that for some c''

$$\mathbb{P}(\mathcal{E}_{N,1} \cap \mathcal{E}_{N,2}(x)^c) \leq \mathbb{P}(\tau_{h,x} \geq c') \leq c'' x.$$

Since $\mathbb{P}(\mathcal{E}_{N,1})$ is bounded away from 0 by Lemma 18, the conclusion follows.

4.2. *Proof of Lemma 22.* Define $d_k = \langle M_{\mu N} \rangle_k - \langle M_{\mu N} \rangle_{k-1}$ (throughout this section k and K are positive integers), and $\mathcal{A}_k = \partial \mathcal{I}_k \setminus \mathcal{I}_{k-1}$. Let $\mathcal{I}^{+,K} = \bigcup_{k=1}^K \partial \mathcal{I}_k^+$, $\mathcal{A} = \bigcup_{k=1}^K \mathcal{A}_k$, and note that for all $k \geq 0$, $\partial \mathcal{I}_k^+ \subset \mathcal{A}_k \cap V$. By Corollary 10, we have

$$d_k = \frac{1}{|\partial V_{\mu N}|} \sum_{v \in \partial V_{\mu N}} \sum_{u \in \mathcal{A}_k} \mathrm{Hm}_{N,k}(\partial V_{\mu N}, u) G_{N,k-1}(u, v)$$

$$\geq \sum_{u \in \mathcal{A}_k} (\mathrm{Hm}_{N,k}(\partial V_{\mu N}, u))^2 G_{N,k-1}(u, u),$$

where the inequality follows from $G_{N,k-1}(u,v) \geq \text{Hm}_{N,k}(v,u)G_{N,k-1}(u,u)$. Consequently,

$$(35) \qquad \langle M_{\mu N} \rangle_K \geq \sum_{k=1}^{K} \sum_{u \in \mathcal{A}_k} \left(\text{Hm}_{N,k}(\partial V_{\mu N}, u) \right)^2 G_{N,k-1}(u,u).$$

Comparing (35) to the desired inequality in Lemma 22, we see two differences: (1) the summation in (35) is over $\mathcal{A}_k$ as opposed to $\partial \mathcal{I}_k^+$; (2) there is a term $G_{N,k-1}(u,u)$ in (35) which we need to bound from below. To address this, we will define a function $\psi : \mathcal{I}^{+,K} \to \mathcal{A}$ which, roughly speaking, allows us to bound $(\text{Hm}_{N,k}(\partial V_{\mu N}, u))^2 G_{N,k-1}(u,u)$ from below by $(\text{Hm}_{N,\tau}(\partial V_{\mu N}, \psi(u)))^2$ (where $\psi(u) \in \mathcal{A}_\tau$). We next carry out the details.

To specify ψ, let D_{ℓ_1} be ℓ_1-distance on $\mathbb{R}^2$ (up to scaling, this is the graph distance on $\tilde{\mathcal{G}}$), and $u \in \partial \mathcal{I}_k^+ \subseteq \mathcal{A}_k$ be an active point at time k. If $D_{\ell_1}(u, \mathcal{I}_{k-1}) \geq 1/2$, then $G_{N,k-1}(u,u) \geq 1/2$ (see [11]) and we let $\psi(u) = u$. If, on the other hand, $D_{\ell_1}(u, \mathcal{I}_{k-1}) < 1/2$, there exist at most four points on $\partial \mathcal{I}_{k-1} \cap B_{\ell_1}(u, 1/2)$ (here $B_{\ell_1}(u,r)$ denotes the open ball of radius r centered at u with respect to D_{ℓ_1}). For every $w \in \partial \mathcal{I}_{k-1} \cap B_{\ell_1}(u, 1/2)$, there is a unique (random) integer $\tau_w \leq k-1$ such that $w \in \mathcal{A}_{\tau_w}$. We let $\psi(u) = w$ be the point in $\partial \mathcal{I}_{k-1} \cap B_{\ell_1}(u, 1/2)$ that minimizes τ_w (i.e., the "oldest" w), breaking ties by distance to u (choosing the w closest to u). With this choice, $\partial \mathcal{I}_{\tau_w} \cap B_{\ell_1}(u, |u-w|_{\ell_1}) = \varnothing$ so $\text{Hm}_{N,\tau_w}(u,w) \geq 1/4$, and hence

$$(36) \qquad \text{Hm}_{N,\tau_w}(\partial V_{\mu N}, w) \geq \frac{1}{4} \text{Hm}_{N,k}(\partial V_{\mu N}, u).$$

Also, $D_{\ell_1}(u, \mathcal{I}_{\tau_w-1}) \geq 1/2$ so by (3) we have for $\delta = D_{\ell_1}(w,u) < 1/2$,

$$(37) \qquad G_{N,\tau_w-1}(w,w) = 4\delta(1-\delta) + (1-\delta)^2 G_{N,\tau_w-1}(u,u) \geq \frac{1}{2}.$$

Finally, for distinct $u, u' \in \mathcal{I}^{+,K}$, $B_{\ell_1}(u, 1/2) \cap B_{\ell_1}(u', 1/2) = \varnothing$ so ψ is injective. Recalling (35), we get that

$$\langle M_{\mu N} \rangle_K \geq \sum_{k=1}^{K} \sum_{u \in \partial \mathcal{I}_k^+} \left(\text{Hm}_{N,\tau_{\psi(u)}}(\partial V_{\mu N}, \psi(u)) \right)^2 G_{N,\tau_{\psi(u)}-1}(\psi(u), \psi(u))$$

$$\geq \frac{1}{32} \sum_{k=1}^{K} \sum_{u \in \partial \mathcal{I}_k^+} \text{Hm}_{N,k}(\partial V_{\mu N}, u)^2,$$

where the factor of $\frac{1}{32}$ comes from $(\frac{1}{4})^2$ (which accounts for the ratio on the square of harmonic measures; see (36)) and $\frac{1}{2}$ (which accounts for the Green function term; see (37)).

4.3. *Proof of Lemma* 23. Let $A_N = V_{\beta N} \setminus V_{\alpha N}$. We will bound $\mathbb{E}[|\mathcal{B}_j|]$ by bounding the probability that each vertex $v \in A_N$ belongs to $\mathcal{B}_j$. Note that

$$\mathcal{I}^+ \subseteq \left\{ v \in A_N, D_{N,h}(V_{\alpha N}, v) < \infty \right\}.$$

By Theorem 8, $\mathbb{P}(D_{N,h,\beta}(V_{\alpha N}, v) < \infty \mid \tilde{\phi}_{N,v})$ is increasing in $\tilde{\phi}_{N,v}$. Therefore, letting $a_j = h + 2^{j-1}\sqrt{\log N} \mathbb{1}_{j>0}$ and $b_j = h + 2^j \sqrt{\log N}$ we have for $j \geq 0$ and $v \in A_N$,

$$(38) \qquad \mathbb{P}(v \in \mathcal{B}_j) \leq \mathbb{P}(\tilde{\phi}_{N,v} > a_j)\mathbb{P}(D_{N,h}(V_{\alpha N}, v) < \infty \mid \tilde{\phi}_{N,v} = b_j).$$

Since (see, e.g., [10], Theorem 4.4.4, Proposition 4.6.2)

$$(39) \qquad \text{Var}[\tilde{\phi}_{N,v}] = G_N(v,v) = \frac{2}{\pi}\log N + O(1),$$

we see that there exists a constant $c = c(h, \beta) > 0$ such that for all $j \geq 0$ and $v \in A_N$,

$$(40) \qquad\qquad\qquad\qquad \mathbb{P}(\tilde{\phi}_{N,v} > a_j) \leq e^{-c4^j}.$$

We will bound the second term of (38) in terms of k for $v \in \partial V_{\alpha N + k}$. We state the result here and defer the proof to the end of this section.

LEMMA 24.   *Let $k^* = (\beta - \alpha)N/\sqrt{\log N}$. There exists a positive constant $c = c(h, \beta, \alpha)$ such that for all $j \geq 0$, $k \geq k^*$ and $v \in \partial V_{\alpha N + k}$,*

$$\mathbb{P}\big(D_{N,h}(V_{\alpha N}, v) < \infty \mid \tilde{\phi}_{N,v} = b_j\big) \leq \frac{c2^j}{\sqrt{\log N}} \sqrt{\log N - \log k}.$$

Using Lemma 24 and (40) (and the fact that $\mathbb{E}[|\mathcal{B}_j|] = \sum_{v \in A_N} \mathbb{P}(v \in \mathcal{B}_j)$), we have

$$\mathbb{E}[|\mathcal{B}_j|] \leq c2^j e^{-c'4^j} \left[\frac{N^2}{\sqrt{\log N}} + \frac{N}{\sqrt{\log N}} \sum_{k=k^*}^{(\beta-\alpha)N} \sqrt{\log N - \log k} \right],$$

where we used the fact that there are $O(N^2/\sqrt{\log N})$ lattice points in $V_{\alpha N + k^*} \setminus V_{\alpha N}$. Finally, we note

$$\sum_{k=k^*}^{(\beta-\alpha)N} \sqrt{\log N - \log k} = O(1) \int_{k^*}^{(\beta-\alpha)N} \sqrt{\log N - \log x}\, dx$$

$$= O(N) \int_{(\beta-\alpha)^{-1}}^{(\beta-\alpha)^{-1}\log N} \frac{\sqrt{\log u}}{u^2}\, du = O(N).$$

This completes the proof of Lemma 23.

PROOF OF LEMMA 24.   As usual, the proof consists of analyzing an exploration martingale. To specify the martingale, we let $A_{N,v} = \partial(V_{(1-\gamma)N} + v)$ be the boundary of a box of radius $(1-\gamma)N$ around v. Note that $A_{N,v} \subset V_N$ for all $v \in V_{\beta N}$. We then take $M_{N,v}$ to be the exploration martingale with source $\mathcal{I}_0 = \{v\}$ and observable $X_{A_{N,v}}$ on $\tilde{E}_N^{\geq h}$. We also let $r_{N,v} = k \wedge ((1-\gamma)N/2)$ and $W_{N,v} = \partial(V_{r_{N,v}} + v)$, and note that

$$\{D_{N,h}(v, V_{\alpha N}) < \infty\} \subset \{D_{N,h}(v, W_{N,v}) < \infty\}.$$

As before, we let $\mathrm{Hm}_{N,t}(u, w) = \mathrm{Hm}(u, w; \mathcal{I}_t \cup \partial V_N)$, $\mathrm{Hm}_N(u, \mathcal{I}_t) = \mathbb{P}_u(\tau_{\mathcal{I}_t} < \tau_{\partial V_N})$, and $\pi_{N,v,t} = \mathrm{Hm}(A_{N,v}, \mathcal{I}_t)$. We have

$$M_{N,v,t} - M_{N,v,0} \geq h[\pi_{N,v,t} - \pi_{N,v,0}] - 2^j \sqrt{\log N}\, \pi_{N,v,0},$$

with equality if and only if $\pi_{N,v,t} = \pi_{N,v,\infty}$ (recall we assume $\tilde{\phi}_v = 2^j \sqrt{\log N} + h$). In particular, on $\{D_{N,h}(v, W_{N,v}) < \infty\}$ we have strict inequality for $0 \leq t < D_{N,h}(v, W_{N,v})$. Arguing as in the proof of Lemma 20, we can show that there exist constants $c_-, c_+ > 0$ (independent of v and k) such that the following holds for all $0 \leq s \leq t \leq D_{N,h}(v, W_{N,v})$:

$$c_-[\pi_{N,v,t} - \pi_{N,v,s}] \leq \langle M_{N,v} \rangle_t - \langle M_{N,v} \rangle_s \leq c_+[\pi_{N,v,t} - \pi_{N,v,s}].$$

Additionally, it follows from Lemma 21 and (39) that there exists a constant $c_1 > 0$ (also independent of v and k) such that $\pi_{N,v,0} \leq c_1/\log N$. All together, this shows that the following holds almost surely on $\{D_{N,h}(v, W_{N,v}) < \infty\}$:

$$M_{N,v,t} - M_{N,v,0} \geq c_2 h \langle M_{N,v} \rangle_t - \frac{c_1 2^j}{\sqrt{\log N}}, \qquad 0 \leq t < D_{N,h}(v, W_{N,v}).$$

To simplify notation, we let $m = c_2 h$ and $b = c_1 2^j / \sqrt{\log N}$. As usual, we write $T_t = \inf\{s : \langle M_{N,v} \rangle_s > t\}$ and let B be a standard Brownian motion independent of $\tilde{\phi}_v$ such that $B_t = M_{N,v,T_t} - M_{N,v,0}$ for $t < \langle M_{N,v} \rangle_\infty$. Then we can define

$$\tau_{h,N} = \inf\{t : B_t \leq mt - b\},$$

$$\pi_{N,v}^- = \inf_{\mathcal{I}}\{\mathrm{Hm}_N(A_{N,v}, \mathcal{I})\} - \mathrm{Hm}_N(A_{N,v}, v),$$

where the infimum is taken over all closed, connected subsets $\mathcal{I} \subset \tilde{\mathcal{G}}_N$ containing v and intersecting $W_{N,v}$. For notational convenience, we write $T = c_- \pi_{N,v}^-$. We have that

$$\mathbb{P}(D_{N,h}(v, W_{N,v}) < \infty) \leq \mathbb{P}(\tau_{h,N} \geq T).$$

Note that the right-hand side of the inequality is decreasing in h, so we assume $h \leq 0$. Applying Proposition 15 and Lemma 17, we get that

$$\mathbb{P}(\tau_{h,N} \geq T) \leq 2g(m\sqrt{T}, 0)\varphi(m\sqrt{T})\frac{b}{\sqrt{T}}$$

$$\leq 2[\varphi(m\sqrt{T}) - m\sqrt{T}]\frac{b}{\sqrt{T}}.$$

To conclude the proof, we need the following bound on $\pi_{N,v}^-$:

(41) $$\pi_{N,v}^- \geq c(\log N - \log k)^{-1}.$$

Provided with (41) and noting the trivial bound $T \leq c_-$, we have $-m\sqrt{T} = O(1)$ and $b = O(2^j / \sqrt{\log N})$ which implies

$$\mathbb{P}(\tau_{h,N} \geq T) = O\left(\frac{2^j \sqrt{\log N - \log k}}{\sqrt{\log N}}\right),$$

as claimed.

It remains to prove (41). In fact, we will prove

$$\pi_{N,v}^- \geq c(\log N - \log r_{N,v})^{-1}.$$

To this end, let u be a point on $\mathcal{I}$ such that $r_{N,v}/2 \leq |u - v|_{\ell_\infty} < 1 + r_{N,v}/2$, and let B_u, B_u', B_u'' be boxes centered at u of side length $r_{N,v}/4$, $r_{N,v}/8$, $r_{N,v}/16$ respectively. By [10], Proposition 6.4.1, we get that there exists $c' = c'(\alpha, \beta, \gamma) > 0$ such that for any $x \in A_{N,v}$,

$$\mathrm{Hm}_N(x, B_u') \geq c'(\log N - \log r_{N,v})^{-1}.$$

It is also obvious that once the random walk arrives at $\partial B_u'$, there is a probability bounded uniformly from below that the random walk range before exiting B_u will contain a contour in $B_u \setminus B_u''$. In this case, the random walk will hit at least one point in $\mathcal{I} \cap B_u$. Therefore, we get that

$$\mathrm{Hm}_N(x, \mathcal{I} \cap B_u) \geq c''(\log N - \log k)^{-1},$$

where $c'' > 0$ depends on c'. In addition, for any $w \in \mathcal{I} \cap B_u$, we have

$$\mathrm{Hm}(w, \partial V_N; \partial V_N \cup \{v\}) \geq 1/2$$

(see, e.g., [10], Theorem 4.4.4., Proposition 4.6.2). Altogether, this means that for any $x \in A_{N,v}$, we have

$$\Delta\pi(x, \mathcal{I}) := \mathbb{P}_x(\text{random walk hits } \mathcal{I} \text{ but not } v \text{ before it hits } \partial V_N)$$

$$\geq c(\log N - \log r_{N,v})^{-1}$$

for a constant $c > 0$. Noting that

$$\mathrm{Hm}_N(A_{N,v}, \mathcal{I}) - \mathrm{Hm}_N(A_{N,v}, v) = \frac{1}{|A_{N,v}|} \sum_{x \in A_{N,v}} \Delta \pi(x, \mathcal{I}),$$

this completes the verification of (41). $\square$

REFERENCES

[1] ARU, J., LUPU, T. and SEPÚLVEDA, A. The first passage sets of the 2D Gaussian free field: Convergence and isomorphisms. Preprint, arXiv:1805.09204.

[2] ARU, J., LUPU, T., and SEPÚLVEDA, A. First passage sets of the 2D continuum Gaussian free field. Preprint, arXiv:1706.07737.

[3] BASU, D. and SAPOZHNIKOV, A. (2017). Kesten's incipient infinite cluster and quasi-multiplicativity of crossing probabilities. *Electron. Commun. Probab.* **22** Paper No. 26, 12. MR3652039 https://doi.org/10.1214/17-ECP56

[4] BORODIN, A. N. and SALMINEN, P. (2002). *Handbook of Brownian Motion—Facts and Formulae*, 2nd ed. *Probability and Its Applications*. Birkhäuser, Basel. MR1912205 https://doi.org/10.1007/978-3-0348-8163-0

[5] BRICMONT, J., LEBOWITZ, J. L. and MAES, C. (1987). Percolation in strongly correlated systems: The massless Gaussian field. *J. Stat. Phys.* **48** 1249–1268. MR0914444 https://doi.org/10.1007/BF01009544

[6] DING, J. and LI, L. (2018). Chemical distances for percolation of planar Gaussian free fields and critical random walk loop soups. *Comm. Math. Phys.* **360** 523–553. MR3800790 https://doi.org/10.1007/s00220-018-3140-x

[7] DREWITZ, A., PRÉVOST, A. and RODRIGUEZ, P.-F. (2018). The sign clusters of the massless Gaussian free field percolate on $\mathbb{Z}^d, d \geq 3$ (and more). *Comm. Math. Phys.* **362** 513–546. MR3843421 https://doi.org/10.1007/s00220-018-3209-6

[8] DREWITZ, A. and RODRIGUEZ, P.-F. (2015). High-dimensional asymptotics for percolation of Gaussian free field level sets. *Electron. J. Probab.* **20** no. 47, 39. MR3339867 https://doi.org/10.1214/EJP.v20-3416

[9] KESTEN, H. (1986). The incipient infinite cluster in two-dimensional percolation. *Probab. Theory Related Fields* **73** 369–394. MR0859839 https://doi.org/10.1007/BF00776239

[10] LAWLER, G. F. and LIMIC, V. (2010). *Random Walk: A Modern Introduction. Cambridge Studies in Advanced Mathematics* **123**. Cambridge Univ. Press, Cambridge. MR2677157 https://doi.org/10.1017/CBO9780511750854

[11] LUPU, T. (2016). From loop clusters and random interlacements to the free field. *Ann. Probab.* **44** 2117–2146. MR3502602 https://doi.org/10.1214/15-AOP1019

[12] LUPU, T. and WERNER, W. (2018). The random pseudo-metric on a graph defined via the zero-set of the Gaussian free field on its metric graph. *Probab. Theory Related Fields* **171** 775–818. MR3827222 https://doi.org/10.1007/s00440-017-0792-y

[13] MOLCHANOV, S. A. and STEPANOV, A. K. (1983). Percolation in random fields. I. *Teoret. Mat. Fiz.* **55** 246–256. MR0734878

[14] POPOV, S. and RÁTH, B. (2015). On decoupling inequalities and percolation of excursion sets of the Gaussian free field. *J. Stat. Phys.* **159** 312–320. MR3325312 https://doi.org/10.1007/s10955-015-1187-z

[15] REVUZ, D. and YOR, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **293**. Springer, Berlin. MR1725357 https://doi.org/10.1007/978-3-662-06400-9

[16] RODRIGUEZ, P.-F. (2014). Level set percolation for random interlacements and the Gaussian free field. *Stochastic Process. Appl.* **124** 1469–1502. MR3163210 https://doi.org/10.1016/j.spa.2013.12.009

[17] RODRIGUEZ, P.-F. (2017). A 0–1 law for the massive Gaussian free field. *Probab. Theory Related Fields* **169** 901–930. MR3719059 https://doi.org/10.1007/s00440-016-0743-z

[18] RODRIGUEZ, P.-F. and SZNITMAN, A.-S. (2013). Phase transition and level-set percolation for the Gaussian free field. *Comm. Math. Phys.* **320** 571–601. MR3053773 https://doi.org/10.1007/s00220-012-1649-y

[19] ROZANOV, YU. A. (1982). *Markov Random Fields*. *Applications of Mathematics*. Springer, New York–Berlin. Translated from the Russian by Constance M. Elson. MR0676644

[20] SZNITMAN, A.-S. (2015). Disconnection and level-set percolation for the Gaussian free field. *J. Math. Soc. Japan* **67** 1801–1843. MR3417515 https://doi.org/10.2969/jmsj/06741801

[21] SZNITMAN, A.-S. (2016). Coupling and an application to level-set percolation of the Gaussian free field. *Electron. J. Probab.* **21** Paper No. 35, 26. MR3492939 https://doi.org/10.1214/16-EJP4563

[22] SZNITMAN, A.-S. (2019). On macroscopic holes in some supercritical strongly dependent percolation models. *Ann. Probab.* **47** 2459–2493. MR3980925 https://doi.org/10.1214/18-AOP1312