THE LIMIT SHAPE AND EMERGENCE OF THE DISCRETE GAUSSIAN LEVEL LINES

JOSEPH CHEN AND EYAL LUBETZKY

ABSTRACT. Consider the (2+1)D Discrete Gaussian model ($\mathbb{Z}$GFF, integer-valued Gaussian free field) on an $L \times L$ box with a hard floor at height zero and zero boundary conditions, at low temperature. The second author, Martinelli and Sly (2016) showed that the surface has a plateau, filling nearly the full square, at height either H or $H+1$ for an explicit function $H(L)$. In a companion paper, we studied the local laws of the top level lines near the four sides of the box, and showed that after rescaling each by $(L^{2/3-o(1)}, L^{1/3-o(1)})$, they converge to a product of Ferrari–Spohn diffusions. Two key features of the top level lines remained unaddressed: their global limit shape, and the critical window marking the transition from a top plateau at height H to one at height $H+1$. These features are intrinsically linked: deriving the global limit of the top level line is needed for determining whether it is preferable to be at height H or $H+1$ near criticality.

This work completes this picture as follows. First, we obtain the global limit of the top level lines: for every fixed n, the n-th from-the-top level line converges in Hausdorff distance to a deterministic shape $\mathscr{L}_n$ that features the Wulff shape at scale $N_n = L^{1-o(1)}$ near the four corners of the box. Second, we identify, for every h, the point of emergence of a macroscopic h level line: the probability of this event is monotone increasing in L (up to a $o(1)$ error), and undergoes a sharp transition from near 0 to near 1 in a critical window of width $\leq L^{1/2+o(1)}$ around a side length $L = L_c^{(h)}$. This transition is discontinuous in that, once a macroscopic level h emerges, it immediately occupies nearly all the box, and the above global and local scaling limits (Wulff, Ferrari–Spohn) hold for it. The new results extend to the $(2+1)$D $|\nabla\phi|^p$-models ($\mathbb{Z}$GFF is the case $p=2$) for every fixed $p>1$.

1. Introduction

The $(2+1)$-dimensional Discrete Gaussian model ($\mathbb{Z}$GFF), also known as the integer-valued Gaussian free field, is a random surface model extensively studied in the context of the roughening transition in crystals (see the work of Chui and Weeks [15] in 1976, and the related models in [6] dating back to the 1950's). It is dual to the Villain XY model [39], and as such undergoes a Berezinskii–Kosterlitz–Thouless (BKT) phase transition [3,25]—as established, along with this result for the Solid-On-Solid (SOS) model, by Fröhlich and Spencer in their celebrated papers [18, 19].

At inverse temperature $\beta > 0$, with zero boundary conditions and a hard floor at height zero, the model is the probability measure over height functions $\phi : \Lambda \to \mathbb{Z}_+$ on $\Lambda = [\![1, L]\!]^2$ given by

$$\pi^0_\Lambda(\phi) \propto \exp\Big(-\beta \sum_{x\sim y} |\phi_x - \phi_y|^2 \Big)\,, \tag{1.1}$$

where $x \sim y$ denotes nearest neighbors in $\mathbb{Z}^2$, and the boundary conditions are $\phi_x = 0$ for $x \notin \Lambda$. Let $\widehat{\pi}^0_\Lambda$ be the floor-free analogue ($\phi : \Lambda \to \mathbb{Z}$). The $\mathbb{Z}$GFF phase transition for $\phi \sim \widehat{\pi}_\Lambda$ occurs at a critical $\beta_{\rm R} > 0$ (empirically, $\beta_{\rm R} \approx 0.665$) as follows:

- for $\beta \leq \beta_{\rm R}$ (delocalized regime) the surface is rough, in that $\lim_{L\to\infty} \mathrm{Var}(\phi_o) = \infty$;
- for $\beta > \beta_{\rm R}$ (localized regime) the surface is rigid, with $\mathrm{Var}(\phi_o) = O(1)$ (moreover, $|\phi_o|$ has a finite exponential moment).

The localization result for $\beta \gg 1$ was established in [4] via an elementary Peierls argument. The delocalization for $\beta \ll 1$ was shown in the aforementioned works of Fröhlich and Spencer [18, 19] (see also [26, 27, 31, 38]). The regimes extend to $(0, \beta_{\rm R})$, $(\beta_{\rm R}, \infty)$ via the monotonicity of $\beta \mapsto \mathrm{Var}(\phi_0)$ due to [1] (this includes the exponential tail in the localized regime by log concavity of the marginals, see [35, §8.2]). Delocalization at the critical $\beta = \beta_{\rm R}$ was thereafter proved by Lammers [26].

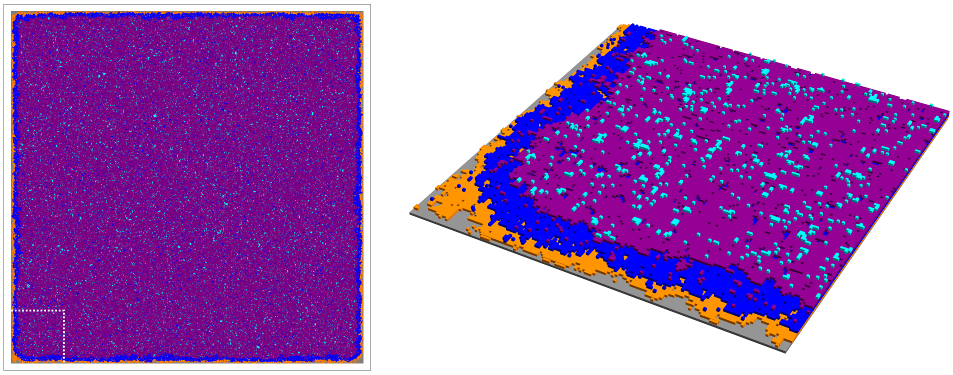

FIGURE 1. Simulation of the low temperature $(2+1)$D $\mathbb{Z}$GFF on a box Λ of side length $L = 1000$, zooming in on the corner of the box, where the macroscopic limit shape is visible.

Throughout this paper, we focus on a fixed large enough β, where the surface is localized, and we let $\widehat{\pi}_\infty$ denote the infinite-volume weak limit of $\widehat{\pi}_\Lambda$ as $L \to \infty$ (well-known to exist for such β). The presence of a hard floor creates a nontrivial competition: the boundary conditions favor a flat surface at height zero in the localized regime, but the floor limits the entropy of surfaces near it, prohibiting some downward fluctuations. This induces *entropic repulsion*: the surface lifts itself to gain entropy from downward spikes, producing a macroscopic plateau far above the floor (Fig. 1). In a recent companion paper [14], we studied the *local geometry* of this plateau—the mesoscopic law of the top level lines near the sides of the box. Our goal here is to study the *global geometry* of the plateau—its macroscopic shape and the sharp transition at which each new layer appears.

1.1. **Previous work on the level lines above a hard floor.** Following is an account of the works most relevant to the current paper; see [14] for a more detailed review of the related literature.

1.1.1. *Entropic repulsion and the surface plateau.* Bricmont, El-Mellouki, and Fröhlich, in their breakthrough paper [5], were the first to rigorously analyze the effect of a hard floor on two closely related $\mathbb{Z}$-valued height functions above a hard floor: the $\mathbb{Z}$GFF from Eq. (1.1), and the (absolute-value) SOS model, where the gradient cost $|\phi_x - \phi_y|^2$ in Eq. (1.1) is replaced by $|\phi_x - \phi_y|$. They established that, in both models at low enough temperature, the floor pushes the average surface height to polylog(L), with the heuristic being that a surface at height h gains entropy from the extra downward fluctuations (of order h) that become accessible. Specifically, they showed that $\phi \sim \pi_\Lambda^0$ has $\frac{1}{L^2}\sum_x \mathbb{E}[\phi_x] \geq \frac{c}{\beta}\sqrt{\log L}$, and that the analogous SOS model has $\frac{1}{L^2}\sum_x \mathbb{E}[\phi_x] \geq \frac{c}{\beta}\log L$. (This is the correct surface height for SOS, while for $\mathbb{Z}$GFF it is off by a $\sqrt{\log\log L}$ factor; see Eq. (1.2).)

For SOS, Caputo et al. [11–13] made significant progress in understanding the geometry of the surface above a floor at low temperature. They showed that the surface typically forms a plateau: all but an ε-fraction of the sites are all at the same height, which is either $\lfloor \frac{1}{4\beta}\log L\rfloor - 1$ or $\lfloor \frac{1}{4\beta}\log L\rfloor$. They further obtained a shape theorem for the SOS surface. The h *level lines* are the loops formed by dual-edges separating sites with $\phi < h$ from ones with $\phi \geq h$ (see Definition 2.1 for more details). It was shown in [13] that there exists $c_*(\beta)$ tending to 4 as $\beta \to \infty$, so that for every fixed $\varepsilon > 0$, if

$$(1+\varepsilon)c_*\beta e^{4\beta h} < L < (1-\varepsilon)c_*\beta e^{4\beta(h+1)}$$

then, with high probability (w.h.p.) as $L \to \infty$ (equivalently, as $h \to \infty$), one has that:

(i) At least $\frac{9}{10}$ of the sites have height exactly h ([13, Thm. 1]).

(ii) For each $j = 0, \ldots, h$ there is a unique j level-line loop of length $\geq (\log L)^2$ (none for $j > h$). These loops are nested, and after rescaling $[\![1, L]\!]^2$ to $[0,1]^2$ they converge in Hausdorff distance to a deterministic limit obtained by translates of Wulff shapes ([13, Thm. 2]).

This delineated for the $(2+1)$D SOS the sequence of critical side-lengths $L_c^{(h)} = c_*\beta e^{4\beta h}$ that mark the emergence of each new layer in the surface, where the plateau increases from height $h-1$ to h. (The SOS behavior at the critical side-lengths $L = (1+o(1))L_c^{(h)}$ was not addressed by [13].)

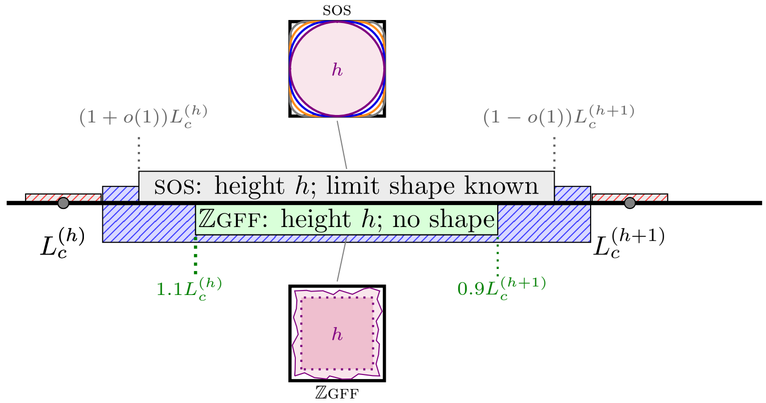

FIGURE 2. Previous results for SOS (gray region) in [13] and $\mathbb{Z}$GFF (green region) in [29]. In SOS, the height and limit shape of the top level was identified outside a $1+o(1)$ window around $\lambda_*\beta/\widehat{\pi}_\infty(\phi_o=h)$, the natural candidate for $L_c^{(h)}$. The $\mathbb{Z}$GFF results excluded a larger $1+\delta/\beta$ window and missed the limit shape. Theorems 1.2 and 1.4 extend the range where the top height is identified (to the blue region), excluding now a $1+(L_c^{(h)})^{-1/2+o(1)}$ window. Theorem 1.1 gives the $\mathbb{Z}$GFF limit shape for all L, regardless of whether the top height can be identified. Theorem 1.4 provides this latter extension also for SOS.

For $\mathbb{Z}$GFF, the second author, Martinelli, and Sly [29] proved an analogue of the SOS Item (i). Let

$$H(L) := \max\left\{h : \widehat{\pi}_\infty(\phi_o = h) \geq \frac{5\beta}{L}\right\} \asymp \sqrt{(1/\beta)\log L \log\log L}\,. \tag{1.2}$$

Call a loop γ `large` if its length is at least $\log L$ (one could even use a threshold of $(C/\beta)\log L$). The following result of [29] shows that the surface typically forms a sequence of nested `large` loops, one per level $h=0,\ldots,H$ and none at height $H+2$ or above:

Theorem A ([29, Thm. 2][1])**.** *Fix β large and consider the $\mathbb{Z}$GFF $\phi\sim\pi_\Lambda^0$ per Eq. (1.1) and $H(L)$ per Eq. (1.2). Then, w.h.p, there is a unique* `large` *h level-line loop for each $h=0,\ldots,H$, and there are no* `large` *h level-line loops for any $h\geq H+2$. This sequence of loops is nested; the loops for $h\leq H-1$ have area $(1-o(1))L^2$, and the loop for $h=H$ has area at least $(1-\varepsilon_\beta)L^2$.*

The proof argument in [29] (see §4.4 there) further showed that if $\widehat{\pi}_\infty(\phi_o=H+1)>4.1\beta/L$ then there is an $(H+1)$ level-line loop with area at least $(1-\varepsilon_\beta)L^2$, and that if $\widehat{\pi}_\infty(\phi_o=H+1)<3.9\beta/L$ then there are no `large` $(H+1)$ level-line loops. (The constants 3.9 and 4.1 can be taken to be $4-\delta$ and $4+\delta$ if β is large enough as a function of δ; namely, the aforementioned argument holds when replacing them by $4-5/\beta$ and $4+2/\beta$, respectively.) Together, these results of [29] show that the critical side-length $L_c^{(h)}$ marking the onset of level h behaves as $\approx(4\pm\varepsilon_\beta)\beta/\widehat{\pi}_\infty(\phi_o=h)$. (See Fig. 2 comparing these results with the SOS picture summarized in Items (i) and (ii) above.)

For all but an exceptional set $\mathscr{B}$ of side lengths of zero logarithmic density (the values near which the plateau transitions from height h to $h+1$; see Eq. (1.5)), the top loop at height H also has area $(1-o(1))L^2$ and is the unique macroscopic level line at that height.

1.1.2. *Local limit at the flat boundary.* If the low temperature $\mathbb{Z}$GFF model is to take after its SOS counterpart, then the macroscopic limit shapes of its top level lines should press against all four sides of the box, producing *flat portions* coinciding with each side. Along these flat portions, the level line can be viewed as a 1D interface pinned near the side wall, and the relevant question is the scale and law of its transversal fluctuations. Near the four corners, one wishes to both characterize the curved limit, and obtain the law of the random fluctuations around it.

[1]The cutoff for `large` loops in [29] (called *macroscopic* loops there) was $\log^2 L$, though the proofs hold also for a threshold of $\log L$ provided β is large enough; see the footnote below the definition of macroscopic loops in [29, §1.2].

In the SOS model, the top level line was shown in [13] to have random fluctuations of at most $L^{1/3+o(1)}$ from the flat portions of its limit. A key ingredient was showing that the law of the top level line in the relevant region resembles that of a random walk, conditioned to be nonnegative, and penalized exponentially in the area below it. Such area-tilted walks are known [24] to have $L^{1/3}$ fluctuations and their scaling limit is a Ferrari–Spohn diffusion [17]. It was recently established [7] that the order of the SOS top level line fluctuations is at least $L^{1/3}$, and one expects that to be the correct order. Notably, the conjectured SOS scaling limit for an individual level line in the presence of other level lines is *not* Ferrari–Spohn but a variant that accounts for the interaction between the different level lines. Rather, the conjectured SOS limit should be that of an ensemble of non-crossing random walks with geometric area tilts (see, e.g., [2,8–10,16,21,34] for studies of this line ensemble in the discrete and continuous settings, which for a single curve reverts to a Ferrari–Spohn limit).

For the $\mathbb{Z}$GFF, the picture turns out to be different. It was shown in the companion paper [14] that if one excludes a set $\mathscr{B}$ of near-critical side lengths (akin to the set excluded in [29]; see Eq. (1.5)), then the following separation of scales occurs for the top level lines. The n-th from the top level line $\mathfrak{L}_n$ lies typically at distance $N_n^{1/3}$ from the flat portion of the boundary for

$$N_n := \frac{1}{\widehat{\pi}_\infty(\phi_o = H+1-n)} \quad (= L^{1-o(1)}) \qquad (n = 0,1,\ldots)\,, \tag{1.3}$$

which further satisfies $N_n = N_{n-1}e^{-\Theta(\sqrt{\beta \log L/\log\log L})}$ (so $\mathfrak{L}_n$ is supported in "a scale of its own"). This separation of scales allowed the authors to decouple the different level lines, and show that the top m level lines, each after an $(N_n^{2/3}, N_n^{1/3})$ space-time rescaling of its distance from the side boundary, converge to *independent* Ferrari–Spohn diffusions. Specifically, the Ferrari–Spohn diffusion FS_σ on $(0,\infty)$ is the diffusion with Dirichlet boundary condition at 0 and generator

$$\mathsf{L}_\sigma = \tfrac{\sigma^2}{2}\tfrac{\mathrm{d}^2}{\mathrm{d}x^2} + \sigma^2 \tfrac{\varphi_\sigma'(x)}{\varphi_\sigma(x)}\tfrac{\mathrm{d}}{\mathrm{d}x}\,, \tag{1.4}$$

where $\varphi_\sigma(x) = \mathsf{Ai}((2/\sigma^2)^{1/3}x - \omega_1)$ for Ai the Airy function and $\omega_1 = \min\{x > 0 : \mathsf{Ai}(-x) = 0\}$. This process is ergodic and reversible against the probability density $\frac{(2/\sigma^2)^{1/3}}{\mathrm{Ai}'(-\omega_1)^2}\varphi_\sigma(x)^2\mathbb{1}_{\{x>0\}}$. Let

$$\mathscr{B} = \bigcup_{h\geq 1} [\![\tfrac34 \overline{L}^{(h)}, \overline{L}^{(h)}]\!] \qquad \text{where} \qquad \overline{L}^{(h)} = \lceil 5\beta/\widehat{\pi}_\infty(\phi_o = h)\rceil \qquad (h = 1,2,\ldots) \tag{1.5}$$

be the set of excluded side length. Recalling that the transition from a plateau at height $h-1$ to one at height h was known to occur at $(4\pm\varepsilon_\beta)\beta/\widehat{\pi}_\infty(\phi_o = h)$, by excluding $\mathscr{B}$ one avoids this delicate window: for $L \notin \mathscr{B}$, the plateau is w.h.p. at height H from Eq. (1.2) (rather than $H+1$). (The previous work [29] excluded $[\![3.9\beta/\widehat{\pi}_\infty(\phi_o = h), 4.1\beta/\widehat{\pi}_\infty(\phi_o = h)]\!]$ with the exact same effect.)

Theorem B ([14, Thm. 1.1])**.** *Fix $\beta > 0$ large enough and $m \geq 1$. Let $L \notin \mathscr{B}$ for $\mathscr{B}$ as in Eq. (1.5), and let $\mathfrak{L}_n$ $(n = 1,\ldots,m)$ be the* `large` *$(H+1-n)$ level line[2] of $\phi \sim \pi_\Lambda^0$ on $\Lambda = [\![1,L]\!]^2$. Set N_n as in Eq. (1.3), and let I_n be the interval of length $2N_n^{2/3}$ co-centered on the bottom side of Λ. Denote by $\psi_n(x) = \min\{y \geq 0 : (\frac{L}{2}+x, y) \in \mathfrak{L}_n\}$ the vertical distance of $\mathfrak{L}_n$ from I_n. Then the joint law of the rescaled distances $Y_n(t) := N_n^{-1/3}\psi_n(tN_n^{2/3})$ of the top m level lines $(n = 1,\ldots,m)$ converges weakly to m i.i.d. stationary Ferrari–Spohn diffusions FS_σ on $[-1,1]$ for an explicit fixed $\sigma > 0$.*

The above theorem showed, in particular, that the top level line fluctuations at the center of the side are of order $N_1^{1/3} = L^{1/3-o(1)}$, as previously conjectured in [29]. As stressed in [14, Rem. 1.2], the $L^{o(1)}$ correction in the estimate for the scale N_n is necessary: even though $N_n \asymp L$ for infinitely many values of L, for infinitely many others one has $N_n \approx L\exp(-c\sqrt{\beta \log L \log\log L})$.

While Theorem B addressed only the local law in the flat portion of the limit shape (where it coincides with the side boundaries), the $N_n^{1/3}$ fluctuations are also connected to the global limit

[2]Excluding the set $\mathscr{B}$ meant that w.h.p. there is no `large` $H+1$ level line, so $\mathfrak{L}_n$ is the n-th highest `large` level line.

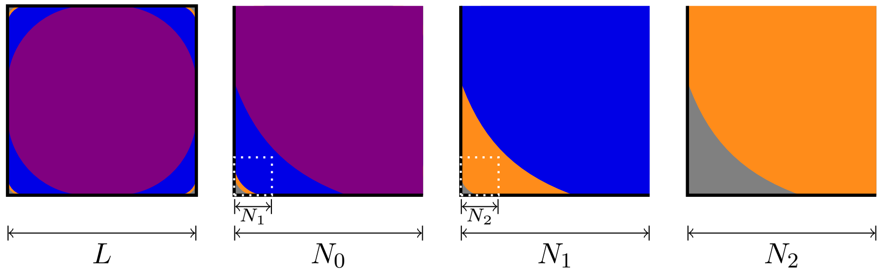

FIGURE 3. Schematic of the macroscopic limit shape of the top level lines $n = 0, 1, \dots$ as established in Theorem 1.1, with the associated scales N_n near the corner of the box Λ.

shape near the corners of the box. As in SOS, the droplet delimited by the $\mathbb{Z}$GFF level line $\mathfrak{L}_n$ behaves in an $N_n^{2/3} \times N_n^{1/3}$ rectangle as an area-tilted random walk, induced to either advance or retreat as per the behavior of the corresponding Wulff shape in the rectangle. This was leveraged in [14] (following [11] for SOS) via a "growth gadget" ([14, Thm. 4.4]; see Theorem 2.14 below) with two ramifications: (i) $\mathfrak{L}_n$ is pressed against the flat boundary to (nearly) the correct scale; and (ii) $\mathfrak{L}_n$ contains a Wulff shape near the corners of the box. In the first—the focus of Theorem B—there was no need for a "retreat gadget" as $\mathfrak{L}_n$ is trivially bounded by the boundary of the box in its flat portion. However, in the second, without a retreat gadget, [14] gave only an "inner bound" on the global limit shape. A matching "outer bound" seemed highly nontrivial: the growth gadget for $\mathbb{Z}$GFF already required significant new ideas compared to the SOS setting, and several of these relied crucially on monotonicity arguments that go in the wrong direction for a retreat gadget.

1.2. **Main results.** Two central open problems left in [29] (see §1.5 there) were to determine the *global limit shape* of the $\mathbb{Z}$GFF level lines—in particular, whether the Wulff construction that governs the SOS limit shape also governs the $\mathbb{Z}$GFF—and to prove that the top level line has fluctuations of order $L^{1/3+o(1)}$ along the flat portions of its limit. We resolved the latter in the companion work [14] (see Theorem B above), except at side lengths $L \in \mathscr{B}$ that are near the critical points $\{L_c^{(h)} : h \geq 1\}$. In this paper we address the former, as well as extend the analysis of [14] to cover every L.

The first theorem provides the global limit shape of the top finitely many level lines, namely, the $H + 1 - n$ level line for every fixed $n \geq 0$ (as per [29], w.h.p. there are no `large` $H + 2$ level lines, but that there may or may not be such a loop at height $H + 1$; the prequel [14] looked only at the latter L's, whence the top height was H). To construct the limit shape for $\mathfrak{L}_n$, we begin with the Wulff shape $\mathcal{W}_1$ which is the convex body with unit area whose boundary minimizes the surface tension integral of the model. Let $\mathcal{W}$ denote $\mathcal{W}_1$ scaled by $\mathsf{w}_1/2$, where w_1 is the value of said surface-tension integral along $\partial\mathcal{W}_1$. (See Section 2.2 for the definitions related to the Wulff shape. The proofs will use a slightly more accurate rescaling of $\mathcal{W}$, relevant if one wishes to improve the quantitative convergence estimates; see Eq. (2.15).) The Wulff shape $\mathcal{W}$ will appear in the limit shape of each $\mathfrak{L}_n$ after a rescaling by N_n; see Fig. 3 for a depiction of this self-similar limit.

Theorem 1.1 (Limit Shape)**.** *Fix $\beta > 0$ large enough and consider $\phi \sim \pi_\Lambda^0$, the $(2+1)$D $\mathbb{Z}$GFF on $\Lambda = [\![1, L]\!]^2$ with a floor and zero boundary conditions. Set $H(L)$ and N_n per Eqs. (1.2) and (1.3). Fix $m \geq 1$, and let $\mathscr{L}_n$ $(n = 0, \dots, m)$ be obtained by placing the 4 quadrants of the Wulff shape $\mathcal{W}$ defined above, rescaled by N_n, in the 4 corners of Λ, and joining them via straight lines along $\partial\Lambda$. Let $\mathfrak{L}_n$ be the* `large` *$(H + 1 - n)$ level line loop(s) of ϕ. Then, with probability $1 - O(L^{-10})$:*

(1) For each $1 \leq n \leq m$, $\mathfrak{L}_n$ is a single loop that satisfies $d_{\mathcal{H}}(\mathfrak{L}_n, \mathscr{L}_n) < N_n e^{-\sqrt{\log L}}$.
(2) For $n = 0$, either $\mathfrak{L}_0$ is empty (no `large` *$(H+1)$ level-line loops), or it is a single loop that satisfies $d_{\mathcal{H}}(\mathfrak{L}_0, \mathscr{L}_0) < N_0 e^{-\sqrt{\log L}}$.*

An important aspect of the above theorem is that it holds for all L, rather than avoiding an exceptional set of L's around the emergence of the $H + 1$ level line. Consequentially, for any L, w.h.p. there cannot be a loop $\mathfrak{L}_0$ which only occupies, say, half of Λ. That is, there are no values

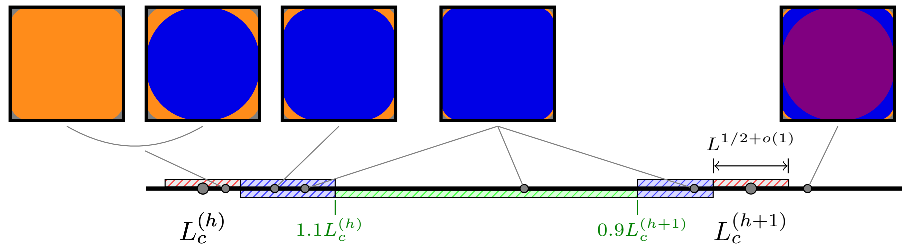

FIGURE 4. Evolution of the surface, as a new layer (plateau at height h) emerges. In the red interval around $L_c^{(h)}$, the top level may be h (dark blue) or $h-1$ (orange), possibly both with constant probability. In the blue and green intervals, the top level is always h, with a larger and larger limit shape as L increases until it occupies a $1-o(1)$ fraction of the box. This was established in the green interval (without identifying the limit shape) in [29].

of L where, with constant probability, the surface features a plateau at height $H+1$ coexisting with a plateau at height H. This byproduct of Theorem 1.1 is reminiscent of a discontinuous phase transition, an interpretation that can be made precise as per the following remark.

Remark. By FKG, one may couple all $\mathbb{Z}$GFF models $\phi^{(L)} \sim \pi^0_{\Lambda_L}$ on $\Lambda_L = [\![1, L]\!]^2$ for $L = 1, 2, \ldots$ such that if $L_1 < L_2$ then $\phi^{(L_1)} \leq \phi^{(L_2)}\restriction_{\Lambda_{L_1}}$ pointwise. As we increase L starting from 1 and sample within this coupling, for every height $h \in \mathbb{Z}_+$ there will be a random side length $L(h)$ which is the minimum L such that $\phi^{(L)}$ contains a `large` h level line. Theorem 1.1 shows that, eventually a.s. for h, every $L > L(h)$ will also feature a `large` h level line (so the phase transition is well-defined), and, immediately upon its emergence in $\phi^{(L(h))}$, the h level line will have the same "super-critical" features (asymptotically, after an appropriate rescaling) of the h level line at $\phi^{(L(h+1)-1)}$.

Although the $H+1$ and H phases cannot coexist, there may still be a critical window of L where both phases have a constant probability to exist; e.g., for every h we can set $L_1(h)$ as the first L where the probability of a `large` h is at least $\frac{1}{10}$, and $L_2(h)$ as the last L where it is at most $\frac{9}{10}$. Our second result states that the width of this critical window is at most $L^{1/2+o(1)}$, and the probability for ϕ to feature a `large` h level line undergoes a sharp transition from $o(1)$ to $1-o(1)$ (see Fig. 5). Fixing h, let $\underline{L}^{(h)} := \frac{3}{4}\overline{L}^{(h)}$ from Eq. (1.5), and define $\lambda_* := \frac{1}{\beta}(2\tau(0) + \frac{\mathsf{w}_1}{2}) \approx 4(1 \pm \varepsilon_\beta)$ where τ and w_1 are the surface tension at angle 0 and surface-tension integral along the unit area Wulff shape.

Theorem 1.2 (Critical window). *Fix any $h \in \mathbb{Z}_+$, and define with respect to $\underline{L}^{(h)}$ the probability*

$$p_h(L) := \pi^0_\Lambda(\textit{there exists an } h \textit{ level-line loop with interior area} \geq .8\underline{L}^{(h)})\,.$$

Then the function $L \mapsto p_h(L)$ is increasing, and the probability that there exists a `large` h level-line loop is $p_h(L) + o(1)$, where the $o(1)$-term goes to 0 as $L \to \infty$. Furthermore, if we define

$$L_c^{(h)} = \min\{L : p_h(L) \geq 1/2\}\,,$$

then we have

$$L_c^{(h)} = (1 + o(1))\frac{\lambda_* \beta}{\widehat{\pi}_\infty(\phi_o = h)}$$

and the following holds for every h:

(i) if $L < L_c^{(h)} - (L_c^{(h)})^{1/2+o(1)}$, then $p_h(L) = o(1)$;
(ii) if $L > L_c^{(h)} + (L_c^{(h)})^{1/2+o(1)}$, then $p_h(L) = 1 - o(1)$.

Remark. We further prove that the $(1 + o(1))$ correction is necessary in the expression for $L_c^{(h)}$:

$$L_c^{(h)} > \frac{\lambda_* \beta}{\widehat{\pi}_\infty(\phi_o = h)} + L^{1-o(1)}\,; \tag{1.6}$$

in particular, $L = \lfloor \frac{\lambda_* \beta}{\widehat{\pi}_\infty(\phi_o = h)} \rfloor$ lies in the "sub-critical" regime where $p_h(L) = o(1)$, as Fig. 5 depicts.

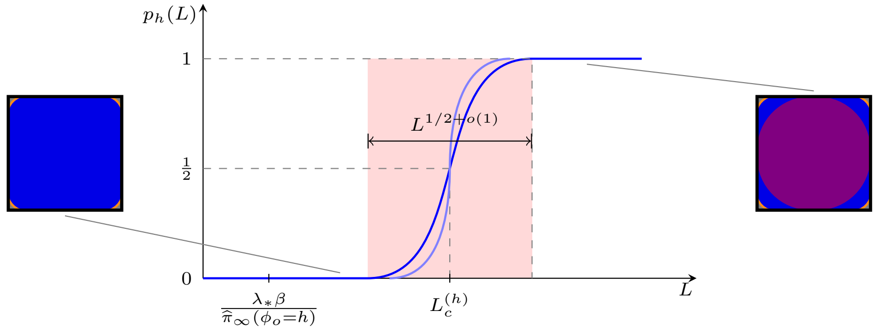

FIGURE 5. The sharp transition of $p_h(L)$, a proxy for the probability that there exists a `large` h level line. $p_h(L)$ is monotone, and increases from near $o(1)$ to near $1-o(1)$ within a window of size at most $(L_c^{(h)})^{1/2+o(1)}$. It is possible that the window is even smaller (depicted in light blue), and identifying the real window size is still open (see Section 1.5). The prediction $\frac{\lambda_*\beta}{\widehat{\pi}_\infty(\phi_o=h)}$ for $L_c^{(h)}$ from previous work on SOS lies outside the critical window.

To give context to the last remark, let us sketch why $\frac{\lambda_*\beta}{\widehat{\pi}_\infty(\phi_o=h)}$ was the natural prediction for $L_c^{(h)}$, appearing also in the previous work [13] on the SOS model. As argued in the pioneering work of Bricmont, El-Mellouki, and Fröhlich [5], an h level line $\mathfrak{L}$, with length $|\mathfrak{L}|$ and interior $\mathrm{Int}(\mathfrak{L})$, costs $\beta|\mathfrak{L}|$ in energy, and gains an entropy of $\log(1+\mathsf{p}_h)\,\mathrm{Int}(\mathfrak{L}) \approx \mathsf{p}_h\,\mathrm{Int}(\mathfrak{L})$ where p_h is the probability of a spike of depth h, thanks to an independent choice of placing such a spike in every $x \in \mathrm{Int}(x)$. If $\mathfrak{L}$ were an $L \times L$ square, then the energetic cost would be $4\beta L$ and, for SOS, the entropic gain would be $e^{-4\beta h}L^2$, leading to the prediction of $L_c^{(h)} = \frac{4\beta}{e^{-4\beta h}}$ as the value of L balancing these. This heuristic was refined in [13] via replacing (a) the ansatz on the shape of $\mathfrak{L}$ from a box to its Wulff-type limit shape, changing 4 into λ_*; and (b) the probability of a downward spike p_h by the large deviation shape under $\widehat{\pi}_\infty$ (which becomes a spike after $O(1)$ levels), changing $e^{-4\beta h}$ into $\widehat{\pi}_\infty(\phi_o = h) \sim c_0 e^{-4\beta h}$. The gap between $L_c^{(h)}$ and $\frac{\lambda_*\beta}{\widehat{\pi}_\infty(\phi_o=h)}$ for the $\mathbb{Z}$GFF highlights a more subtle aspect of this phenomenon. The entropy term $\widehat{\pi}_\infty(\phi_o = h)|\,\mathrm{Int}(\mathfrak{L})|$ treated every x in the interior of $\mathfrak{L}$ as having an independent choice for a downward deviation, ignoring local correlations... Since the large deviations of the $\mathbb{Z}$GFF are shown in [29] to resemble harmonic functions, a downward deviation drags with it nearby points, and this effect is significant enough to change the location of $L_c^{(h)}$. See Remark 5.3 for the technical details.

The next theorem extends Theorem B to all L, removing the exceptional set. As we discussed, for L within the critical window, both events $\{\mathfrak{L}_0$ exists$\}$ and $\{\mathfrak{L}_0$ does not exist$\}$ may have constant probabilities. Nevertheless, when we condition on them, the local limit law is still Ferrari–Spohn. That is, in the same flavor of the discussion following Theorem 1.1, there are no L for which the $H+1$ level line exists yet exhibits a different local behavior along the flat parts of the limit shape.

Theorem 1.3. *Fix $\beta > 0$ large enough and consider $\phi \sim \pi_\Lambda^0$, the (2+1)D $\mathbb{Z}$GFF model on $\Lambda = [\![1, L]\!]^2$ with a floor and zero boundary conditions. Set $H(L)$ and N_n per Eqs. (1.2) and (1.3). For fixed m, let $\mathfrak{L}_n$ $(n = 0, \dots, m)$ be the* `large` *$(H + 1 - n)$ level lines, where possibly $\mathfrak{L}_0$ does not exist. Let $I_n = [\![\frac{L}{2} - N_n^{2/3}, \frac{L}{2} + N_n^{2/3}]\!]$, let $\rho_n(x) := \min\{y \geq 0 : (\frac{L}{2} + x, y) \in \mathfrak{L}_n\}$ be the vertical distance of $\mathfrak{L}_n$ from I_n, and $Y_n(t) := N_n^{-1/3}\rho_n(tN_n^{2/3})$. Then there exists a constant $\sigma > 0$ such that:*

1. *The law of $(Y_n(t))_{n=1}^m$ under π_Λ^0 converges weakly to the law of m i.i.d. copies of the stationary Ferrari–Spohn diffusion FS_σ on $[-1, 1]$.*
2. *For L with $\pi_\Lambda^0(\mathfrak{L}_0$ exists$) > L^{-10}$, the law of $(Y_n(t))_{n=0}^m$ under $\pi_\Lambda^0(\cdot \mid \mathfrak{L}_0$ exists$)$ converges weakly to that of $m+1$ i.i.d. copies of FS_σ.*
3. *For L with $\pi_\Lambda^0(\mathfrak{L}_0$ does not exist$) > L^{-10}$, the law of $(Y_n(t))_{n=1}^m$ under $\pi_\Lambda^0(\cdot \mid \mathfrak{L}_0$ does not exist$)$ converges weakly to that of m i.i.d. copies of FS_σ.*

The same statements hold for $\bar{Y}_n(t)$ corresponding to $\bar{\rho}_n(x) = \max\{y \leq \frac{L}{2} : (\frac{L}{2} + x, y) \in \mathfrak{L}_n\}$.

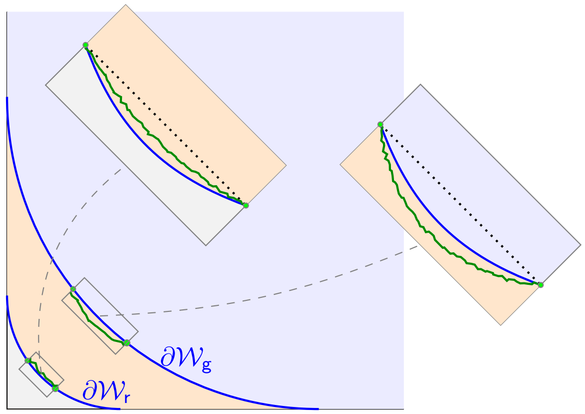

FIGURE 6. Identifying the size ℓ of the Wulff shape featured in $\mathfrak{L}_n$ near the corners of box: A smaller ℓ leads to a bigger area of the limit shape. A retreat mechanism compares the Wulff boundary $\partial\mathcal{W}_{\mathsf{r}}$ (blue) to the location of an area-tilted random walk (green). When the random walk does not drop past the Wulff boundary, neither will $\mathfrak{L}_n$. Similarly, a growth mechanism says when the random walk stays below the Wulff boundary $\partial\mathcal{W}_{\mathsf{g}}$, so will $\mathfrak{L}_n$. As a result, we can sandwich the corner of $\mathfrak{L}_n$ between two arcs of Wulff shapes.

Finally, our proofs hold more generally for the family of $|\nabla\phi|^p$ models for $p \geq 1$, of which the $\mathbb{Z}$GFF ($p=2$) and SOS ($p=1$) are special cases, given by

$$\pi_\Lambda^{(p),0}(\phi) \propto \exp\Big(-\beta\sum_{x\sim y}|\phi_x-\phi_y|^p\Big)\,. \tag{1.7}$$

As before, the infinite-volume weak limit $\pi_\infty^{(p)}$ is well-defined, and leads to analogous definitions of $H(L)$ and N_n per Eqs. (1.2) and (1.3). Similarly, the surface tension is well-defined, leading to analogous definitions of the Wulff shape $\mathcal{W}$ from Theorem 1.1 and λ_* from Theorem 1.2.

Theorem 1.4. *There exist absolute constants $\delta_p, \delta, c > 0$ such that the following hold. The statements of Theorems 1.1 to 1.3 extend to the $|\nabla\phi|^p$ model for any fixed $p > 1$: Theorems 1.2 and 1.3 extend verbatim and Theorem 1.1 extends after modifying the Hausdorff distance bound to*

$$d_{\mathcal{H}}(\mathfrak{L}_n,\mathscr{L}_n) < \begin{cases} N_n^{1-\delta_p} & \text{if } 1<p<2\,,\\ N_n e^{-c\sqrt{\beta\log L}} & \text{if } p>2\,.\end{cases}$$

Finally, the statements of Theorems 1.1 and 1.2 also extend to $p=1$ (SOS): Theorem 1.2 extends verbatim and Theorem 1.1 extends after modifying the bound on the Hausdorff distance to $N_n^{1-\delta}$.

For SOS, the previous work [13] only considered $L \notin \bigcup_h (1\pm\varepsilon)L_c^{(h)}$, avoiding the near-critical ("exceptional") side lengths. The extension of the new results to SOS separates the question of establishing a limit shape from that of determining whether the top level is at H or $H+1$. For the former, the extension of Theorem 1.1 establishes the SOS limit shape for all L (no exceptional side lengths); for the latter, the extension of Theorem 1.2 places the transition point in a window of width $L^{1/2+o(1)}$ as opposed to just $o(L)$. Note that Theorem 1.3 does not extend to $p=1$, and as discussed above, it is not expected to (the conjectured local limit law for SOS is *not* Ferrari–Spohn).

1.3. **Proof ideas.** In what follows, we outline the approach used to prove the new results.

The limit shape. The level lines of the $\mathbb{Z}$GFF above a floor feature a tradeoff between an energy cost for the size of the level line and an entropic reward for its interior area. Variational problems of this type have been well-studied, and the optimal shape in a square domain Λ is precisely the shape $\mathscr{L}_n$ from Theorem 1.1, featuring flat sides and four quadrants of a Wulff shape (defined via the surface tension associated to the model) at the corners (see, e.g., [33]). To identify the limit shape, one must determine the size ℓ of the Wulff shape featured at the four corners (see Fig. 6).

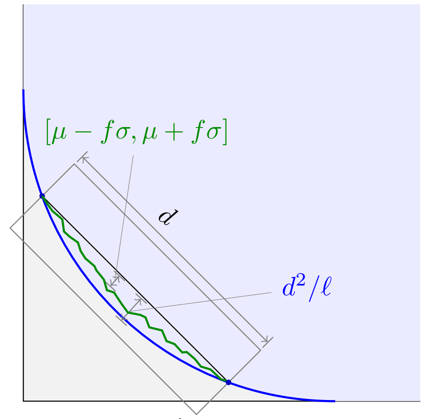

FIGURE 7. Take a chord of length $d = N_n^{2/3} f$ for some $f = f(L) \to \infty$ as $L \to \infty$. The Wulff shape of length ℓ dips $\asymp d^2/\ell$ below the midpoint of the chord, while the area-tilted random walk dips to mean $\mu \asymp N_n^{1/3} f^2$ with fluctuations $\sigma = N_n^{1/3}\sqrt{f} = o(\mu)$. For $\mathcal{W}_{\mathsf{g}}, \mathcal{W}_{\mathsf{r}}$ depicted in Fig. 6, the optimal choices of their sizes $\ell_{\mathsf{g}}, \ell_{\mathsf{r}}$ satisfy $d^2/\ell_{\mathsf{g}} = \mu - f\sigma$ and $d^2/\ell_{\mathsf{r}} = \mu + f\sigma$.

The following heuristic was at the heart of [13], both answering the question above on the scale of the Wulff shape, and yielding the $L^{1/3+o(1)}$ upper bound on the fluctuations at the flat portions of the SOS level-line limit. Suppose we know that locally, between two points A, B distance d apart, the $(H-n)$ level line $\mathfrak{L}_n$ can be coupled to a random walk with area tilt of strength $1/N_n$. Without the area tilt, the normalized area above $\mathfrak{L}_n$ is $O(d^{3/2})$. Hence when $d^{3/2} \gg N_n$, the area tilt will push $\mathfrak{L}_n$ downwards, so that the mean distance μ of $\mathfrak{L}_n$ from $\overline{AB}$ is of a larger order than its fluctuations σ. Explicitly, if $d = N_n^{2/3} f$ for $f = f(L) \to \infty$ as $L \to \infty$, then $\mu = O(N_n^{1/3} f^2)$ and $\sigma = O(N_n^{1/3} f^{1/2})$. Hence, the larger d is, the smaller σ/μ will be.

Now place A, B somewhere along the bottom left boundary of a Wulff shape $\mathcal{W}$ of length ℓ, as in Fig. 7. The distance between $\partial\mathcal{W}$ and $\overline{AB}$ in the middle is $\asymp d^2/\ell$. Hence, if $d^2/\ell > \mu + f\sigma$, then even if at A, B the level line $\mathfrak{L}_n$ reaches all the way up to $\partial\mathcal{W}$, it still will stay above $\partial\mathcal{W}$ in the middle. Hence, such a Wulff shape overestimates how close $\mathfrak{L}_n$ is to the corner of Λ, and we can retreat the Wulff shape away from the corners by taking a larger and larger ℓ until we reach $\mathcal{W}_{\mathsf{r}}$ with radius $\ell_{\mathsf{r}} := d^2/(\mu + f\sigma)$. Similarly, we can grow any Wulff shape of length ℓ such that $d^2/\ell < \mu - f\sigma$ towards the corner, until we reach $\mathcal{W}_{\mathsf{g}}$ with radius $\ell_{\mathsf{g}} := d^2/(\mu - f\sigma)$.

The growth procedure was carried out rigorously in [14] for the $\mathbb{Z}$GFF, giving "half the limit." However, a serious obstacle stood in the way of deriving a matching retreat gadget. In the next subsection we explain this obstacle in more detail, as well as our new approach to circumvent it.

Provided a matching retreat gadget can be established, one would find, as depicted in Fig. 6, that $\mathfrak{L}_n$ is sandwiched between two arcs of $\partial\mathcal{W}_{\mathsf{g}}$ and $\partial\mathcal{W}_{\mathsf{r}}$, which gives the limit shape if the Hausdorff distance X between the two arcs is $o(L)$. To compute the distance X, as per Fig. 8 we see that $X = O(\ell_{\mathsf{g}} - \ell_{\mathsf{r}})$. Since $\sigma = o(\mu)$, up to first order, we have $\ell_{\mathsf{g}} = O(d^2/\mu) = O(N_n)$, so that $X = O((1 - \frac{\ell_{\mathsf{r}}}{\ell_{\mathsf{g}}})N_n)$. Then, the expressions for d^2/ℓ_i give $\frac{\ell_{\mathsf{r}}}{\ell_{\mathsf{g}}} = \frac{\mu - f\sigma}{\mu + f\sigma} = 1 - O(\frac{f\sigma}{\mu})$. Combining these, $X = O(\frac{f\sigma}{\mu} N_n) = o(N_n)$, and showing $\frac{\sigma}{\mu} < e^{-\sqrt{\log L}}$ leads to the bounds of Theorem 1.1.

Area-tilted random walk representation on larger domains. As mentioned above, it is crucial to show that locally (i.e., in a mesoscopic box), $\mathfrak{L}_n$ behaves like a random walk with an area tilt of strength $1/N_n$. The main challenge of this paper is to extend this picture to larger scales, even up to the macroscopic scale. This directly depends on estimating the probability that $\phi_x \geq 0$ for all x inside a domain D with boundary conditions $H + 1 - n$. In the low temperature setting, decorrelation estimates imply that this floor event is realized nearly independently for each x, so

$$\widehat{\pi}_D^{H+1-n}(\phi_x \geq 0, \forall x \in D) \approx \prod_{x \in D} \widehat{\pi}_D^{H+1-n}(\phi_x \geq 0)\,. \tag{1.8}$$

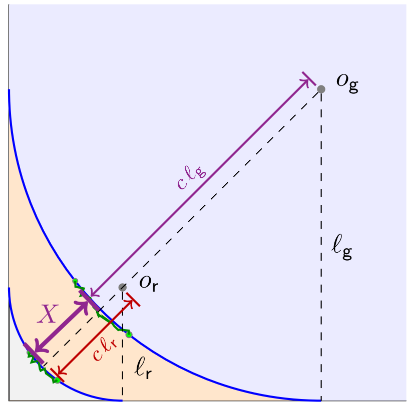

FIGURE 8. The Hausdorff distance X between the Wulff shapes $\mathcal{W}_{\mathsf{g}}$ and $\mathcal{W}_{\mathsf{r}}$ driven by the growth and retreat mechanisms, respectively. If $\mathcal{W}_{\mathsf{g}}, \mathcal{W}_{\mathsf{r}}$ have sizes $\ell_{\mathsf{g}}, \ell_{\mathsf{r}}$, then for some constant $c = c(\beta)$ with $1 < c < \sqrt{2}$ we have $X = (\ell_{\mathsf{g}} - \ell_{\mathsf{r}})(\sqrt{2} - c)$. For an optimal choice of $\ell_{\mathsf{g}}, \ell_{\mathsf{r}}$ via the mechanism depicted in Fig. 7, one gets $X \asymp N_n \frac{f\sigma}{\mu} = N_n f^{-1/2} = o(N_n)$.

However, the error in this approximation grows with the size of the domain $|D|$, and the error produced by the standard inclusion-exclusion analysis becomes too large once $|D| \geq L^{1+o(1)}$.

From the above discussion on the limit shape, we need to exhibit this random walk coupling in a box R of width $d \gg N_n^{2/3}$, and to avoid pinning issues we require the height of R to be at least the same order (stemming from the fact that the pinning effect can be $\exp(cd)$, which, if the height is $o(d)$, would overwhelm the large deviation probability of the random walk). This leads to a minimum area of $|D| \gg N_n^{4/3}$, far exceeding the threshold $L^{1+o(1)}$ allowed by prior arguments. For comparison:

(1) the same requirement of $|D| \gg N_n^{4/3}$ would have been needed in [14] were it not for fortuitous monotonicity arguments: by FKG, there we could relax the floor constraints on some of the area (reducing $|D|$ down to $L^{1+o(1)}$) in the context of a growth gadget, but this would go in the wrong direction in the context of a retreat gadget.
(2) In the SOS model, the error in Eq. (1.8) is much smaller, leading to a less restrictive domain requirement of $|D| \geq L^{3/2}$, easily allowing for D with $|D| \approx L^{4/3}$ in [7,13].

In fact, not only do we need to allow $|D| \approx L^{4/3}$, but to study the critical window about $L_c^{(h)}$ where the h level line first appears, we will need to treat $|D|$ all the way up to $O(L^2)$.

To this end, we need a finer analysis of the probability $\widehat{\pi}_D^{H-n}(\phi_x \geq 0, \forall x \in D)$. Rather than attempt to reduce the error in Eq. (1.8), we take a different approach and say that the floor event occurs nearly independently across (but not within) mesoscopic $\ell \times \ell$ boxes. (A choice of $\ell = 1$ corresponds to the old approach described above.) That is, we express the above probability via

$$\xi_{\ell,h} := -\frac{1}{\ell^2} \log \widehat{\pi}_\infty \left(\phi_x \geq -h,\, \forall x \in Q_\ell := [\![1,\ell]\!]^2\right) ,$$

for an appropriately chosen $\ell = \ell_* \asymp \sqrt{L}$ (see Theorem 3.2 for the precise statement). Handling mesoscopic vs. macroscopic domains D calls for different choices of ℓ to control the associated errors. We do so by first obtaining a generic bound applicable to a range of ℓ's (Proposition 3.9), and thereafter showing that $\xi_{\ell,h}$ does not change much, so we can use the specific $\xi_{\ell_*,h}$ across these various domains D. From there, we conclude that the area tilt felt by $\mathfrak{L}_n$ on larger scales is given by $\mathsf{\rho}_n/N_n$ for a new parameter $\mathsf{\rho}_n = 1 - o(1)$, vs. the old approach of $\ell = 1$ that had a tilt of $1/N_n$. (See below for a further discussion on $\mathsf{\rho}_n$ and its effect on the location of $L_c^{(h)}$.)

Discontinuity of the transition. Once the key estimate of Theorem 3.2 is obtained, the proof of the limit shape of the level lines for L away from $L_c^{(h)}$, as well as a preliminary estimate on $L_c^{(h)}$, follows by combining the methods used for the SOS model in [13] with the disagreement polymer machinery established in [14] to treat the $\mathbb{Z}$GFF (and the $|\nabla\phi|^p$ for $p > 1$). We push these results further in two directions (new even for SOS as far as the critical window and global limit are concerned):

(i) We tighten the estimate on the location of $L_c^{(h)}$ to a near square root window. This involves a quantitative estimate on a functional $\mathcal{F}$ on loops, which captures a tradeoff between the loop length and interior area (see Lemma 2.12).

(ii) We extend both the global limit shape and local scaling limit to all L, no longer excluding a window around $L_c^{(h)}$. For L close to $L_c^{(h)}$, we do not know if the h level line exists, so we must work on the measures $\pi_L(\cdot \mid h \text{ level line exists})$ and $\pi_L(\cdot \mid h \text{ level line does not exist})$. These conditional measures no longer have FKG, so the previous proofs need to be modified so as to be less reliant on monotonicity.

Our analysis in Item (ii) implies that the moment the h level line appears, it attains the limit shape $\mathscr{L}_n$ occupying nearly the whole box Λ, and features Ferrari–Spohn fluctuations along the flat sides of $\mathscr{L}_n$. In particular, there is no critical phase around $L_c^{(h)}$ where the h level line only occupies, say, a third of the space. Roughly speaking, this is because if the area inside an h level line $\mathfrak{L}$ is smaller than a threshold $C_1 L^2$, it cannot exist as the gain from the interior of $\mathfrak{L}$ is too small compared to the gradient cost. At the same time, if the interior area of $\mathfrak{L}$ is larger than another threshold $C_2 L^2$, then a growth procedure shows that $\mathfrak{L}$ must contain the entire aforementioned limit shape. (More precisely, these thresholds relate not just to the interior area but also to the shape contained by $\mathfrak{L}$. See Lemma 5.4 and the definition of $\mathcal{E}_{x,\ell}$ in Claim 6.2 for details.) The dichotomy on the h level line is due to the fact that these thresholds satisfy $C_1 > C_2$. (In fact, $C_2 \leq \frac{1}{4} + \varepsilon_\beta$, pertaining to the level line $\mathfrak{L}_0$ encompassing a Wulff shape of diameter $\approx \frac{1}{2}\mathsf{w}_1 N_0/L$, whereas $C_1 \geq 1 - \varepsilon_\beta$.)

Analysis of the area tilt ρ_n. As we mentioned above, the area tilt ρ_n/N_n as opposed to $1/N_n$ was a result of the new approach of controlling Eq. (1.8) via the tail for the minimum in a mesoscopic square of side length $\sqrt{L}$, as opposed to the singleton rate $\widehat{\pi}_\infty(\phi_o \geq -h)$. Initially, the authors viewed this strategy as merely a technical workaround for the large error associated with Eq. (1.8). However, as it turns out, $(1-\rho_n)L \geq L^{1-o(1)}$. In view of the upper bound $L^{1/2+o(1)}$ on the critical window, this positions the original prediction for $L_c^{(h)}$ (using an area tilt of $1/N_0$) in the subcritical regime, where the probability of a `large` h level line is $o(1)$. We further extended this separation of the a-priori prediction for $L_c^{(h)}$ from the critical window to the $|\nabla\phi|^p$ model for all $1 < p \leq 2$.

Along the way, we prove two results which are essential for the estimates on ρ_n, and may be of independent interest. Firstly, for the $\mathbb{Z}$GFF ($p = 2$), we show that $\widehat{\pi}_\infty(\cdot \mid \phi_x = h, \forall x \in B_r(o))$ has a weak form of rigidity about the discrete harmonic function ϕ^* which is h on $B_r(o)$ and 0 on $\partial B_R(o)$ for large R. In the $\mathbb{R}$-valued setting—the discrete GFF—one can look at $\tilde{\phi} = \phi - \phi^*$, which is nothing but a discrete GFF with 0 boundary conditions. If the same were true in the $\mathbb{Z}$GFF, one would be able to infer rigidity of ϕ about ϕ^* for large β. Unfortunately, in the $\mathbb{Z}$-valued setting, said $\tilde{\phi}$ becomes a field taking values in $\{a_i + \mathbb{Z}\}_{i\in\Lambda}$ for some shifts $a_i \in [0, 1)$. Since the fractional parts a_i are complicated (the complement of those of ϕ^*), they could destroy the full rigidity of $\tilde{\phi}$ (see, e.g., [20], where such fields are studied in detail). Nevertheless, we are able to show rigidity beyond height $h/\log h$ (see Proposition 3.10) by controlling the corresponding integer rounding errors. Secondly, for $1 < p < 2$, we show (see Lemma 8.11) that for any finite set A, the large deviation rate function for the event that $\phi\restriction_A = h$ is given by the p-capacity of A.

1.4. **Organization/Reader's guide.** The paper is organized as follows. In Section 2 we list the required preliminaries, mostly focusing on disagreement polymers and the Wulff shape. Section 3 proves estimates on the probability of the floor event, leading to the area prefactor of ρ_n/N_n. Section 4 uses these area estimates and cluster expansion to obtain the law of the disagreement polymers. Section 5 proves Theorem 1.2, showing the critical window has width $\leq L^{1/2+o(1)}$.

Section 6 proves Theorem 1.1, showing the limit shape of the level lines is a translation of Wulff shapes. Section 7 proves Theorem 1.3, extending the Ferrari–Spohn limit law to all L. Section 8.1 proves Theorem 1.4, extending the main results to the $|\nabla\phi|^p$ model for $p \geq 1$.

Several of the main results are of fairly different flavors—targeting different phenomena in the $\mathbb{Z}$GFF and featuring proofs that build on different methods. We include the following guide for the benefit of a reader interested in one of these results in particular.

- For the proof of Theorems 1.1 and 1.3 and the resulting discontinuous transition, one should read Section 4 for the estimates on the polymer law, and then read Sections 6 and 7.
- For the proof of Theorem 1.2 that the critical window has size at most $L^{1/2+o(1)}$, one should read particularly Lemmas 2.7 and 2.12 from the preliminaries, and then Sections 4 and 5.
- For the discussion on the new area tilt of $\mathsf{\rho}_n/N_n$ vs. $1/N_n$ and its effect on the value of $L_c^{(h)}$, one should read Section 3 and then the start of Section 5 (in particular Remark 5.3). Thereafter, one should read Section 8.1 to see how this varies for different p.
- For the extension of the area-tilted random walk representation to larger domains, one should read Section 3, and may skip Section 3.4 (the effect of a $\mathsf{\rho}_n/N_n$ vs. $1/N_n$ area term).
- For a proof that the $\mathbb{Z}$GFF conditioned on $\phi_o = h$—or more generally, conditioned on $\phi_x = h$ for all $x \in B_r(o)$, a ball around the origin—is rigid beyond scale $h/\log h$ about the harmonic function solving the corresponding $\mathbb{R}$-valued Dirichlet problem, see Proposition 3.10.
- For the extension to the $|\nabla\phi|^p$ models for $1 \leq p < \infty$, see Section 8.1.

1.5. **Open problems.** The results in Theorems 1.1 and 1.3 yield, for all L, the global limit shape of the $\mathbb{Z}$GFF measure π^0_Λ and the Ferrari–Spohn law of the local fluctuations around said limit in its flat portion. It thus remains to establish the local limit law of the level lines about the corner of the limit shape. The long-standing conjecture (for $\mathbb{Z}$GFF as well as other related models) is that the fluctuations $O(L^{1/2})$, and moreover the scaling limit should be Brownian motion. For versions of this conjecture see, e.g., [32, Eq. (3.5) and the paragraph below it] in the 2D Ising model with an external field, [13, §1.2] in SOS and [29, §1.5] in $\mathbb{Z}$GFF.

In light of Theorem 1.2, it would also be interesting to establish the correct order of the critical window for the $\mathbb{Z}$GFF, or more generally for the $|\nabla\phi|^p$ models at any $p \geq 1$. Improving our upper bound of $L^{1/2+o(1)}$ on the width to $O(\sqrt{L})$ for $\mathbb{Z}$GFF or SOS would seem to require a new idea to overcome the $L^{1/2+o(1)}$ error in the law of the level lines in a macroscopic scale (see Corollary 4.4). Any lower bound on the width of critical window would be interesting.

2. Preliminaries

In this section, we will define level lines and disagreement polymers for height functions, review prior results needed for this paper, and set up general notation. We will consider various subgraphs of $\mathbb{Z}^2$, as well as dual-edges (referred to as bonds) of $(\mathbb{Z}^2)^*$. If Λ is a subgraph of $\mathbb{Z}^2$ and γ is a collection of bonds, then $\Lambda \setminus \gamma$ refers to the result of removing from Λ every edge for which its dual-edge is in γ. For any point $u \in \mathbb{R}^2$, we will denote its x, y coordinates by u_1, u_2 respectively. For a set of vertices $U \subset \mathbb{Z}^2$, let ∂U be the boundary bonds of U, i.e., the set of bonds dual to uv with $u \in U$ and $v \notin U$, and let $\partial_{\mathsf{v}} U$ be the external vertex boundary of U, i.e., every vertex $v \notin U$ adjacent to some $u \in U$. Occasionally we will refer to two vertices as $*$-adjacent, which will mean that their $L^\infty(\mathbb{R}^2)$ distance is 1. Finally, any object, such as a level line, disagreement polymer or a connected component, will be called `large` if its size is at least $\log L$.

2.1. **Level lines and disagreement polymers.** We begin by formally defining the level lines.

Definition 2.1 (level lines)**.** Given a height function ϕ on a domain $V \subset \mathbb{Z}^2$, its h level lines are defined as the collection of loops obtained from bonds that are dual to nearest-neighbors $x \sim y$ such that $\phi_x < h$ and $\phi_y \geq h$. As standard in the literature, we separate these into a collection of self-avoiding paths and loops via a NORTHEAST splitting rule (in case 4 dual-edges share an

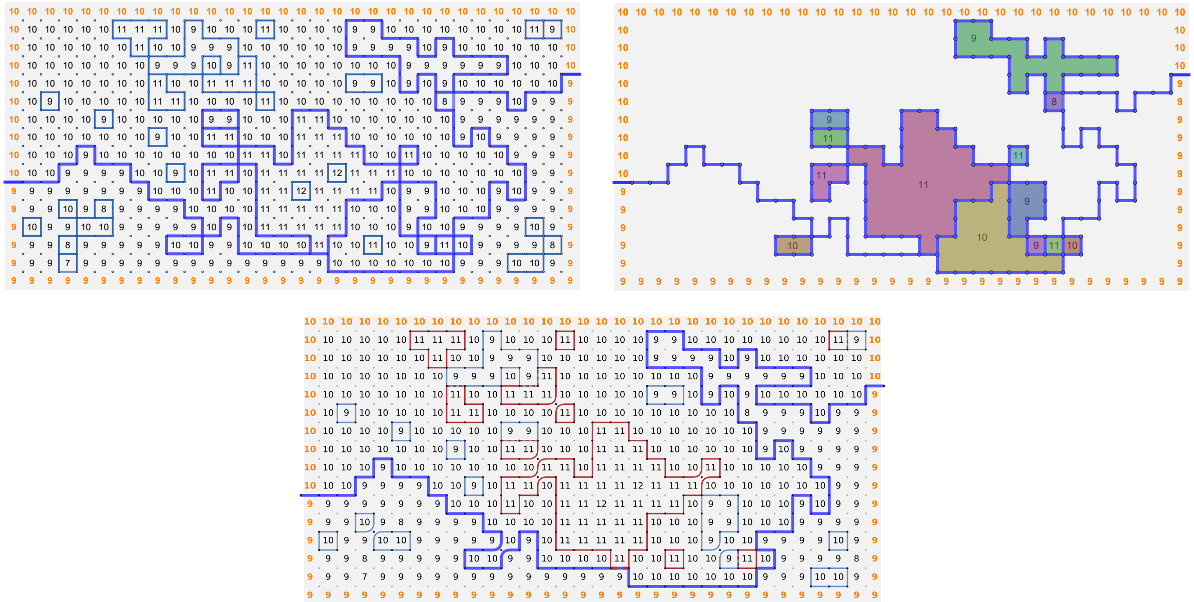

FIGURE 9. Comparison of the disagreement polymer (top left: in thick blue, the bonds that γ consists of; top right: the regions D_i are marked along with the corresponding h_i) and the level lines (bottom: the 10 level lines in blue, the 11 level lines in red) in the same ϕ.

endpoint, one splits them apart along the NORTHEAST diagonal). For an h level-line loop $\mathfrak{L}$, denote its length (number of bonds) by $|\mathfrak{L}|$, its interior (the finite connected region in $\mathbb{R}^2 \setminus \mathfrak{L}$) by $\mathrm{Int}(\mathfrak{L})$, and the area of its interior by $A(\mathfrak{L}) = |\mathrm{Int}(\mathfrak{L})|$. We call $\mathfrak{L}$ an *up-loop* if $\phi_x \geq h$ for every $x \in \mathrm{Int}(\mathfrak{L})$ adjacent to an edge dual to some $b \in \mathfrak{L}$, and otherwise ($\phi_x < h$ for all such x) we call it a *down-loop*.

The key object which we will study are not the level lines themselves but rather "disagreement polymers" as introduced in [14], which are connected components of dual edges corresponding to height differences. The primary issue with studying an h level line $\mathfrak{L}$ directly is that it induces boundary conditions of $\geq h$ and $\leq h-1$ above and below $\mathfrak{L}$, as opposed to $= h$ and $= h-1$ boundaries. We would like to use a cluster expansion analysis which studies the law of $\mathfrak{L}$ by paying a cost along the length of $\mathfrak{L}$, and then separately analyzing the cost of height configurations on the domains above and below $\mathfrak{L}$. In the case of the SOS model ($p = 1$), the linear penalty on gradients allows one to split up a gradient cost of $|x-y|$ for $x \geq h$ and $y \leq h-1$ into $|x-h| + |h-(h-1)| + |(h-1)-y|$. This fits the described framework – every edge of $\mathfrak{L}$ contributes $|h-(h-1)| = 1$, and then the terms $|x-h|$ and $|(h-1)-y|$ can be analyzed separately. However, this decomposition clearly fails when the cost of the gradient is nonlinear, highlighting why boundaries of $\geq h$ and $\leq h-1$ are difficult to work with for the $|\nabla\phi|^p$ models with $p > 1$.

We next summarize the definition and key properties of disagreement polymers as proven in [14].

Definition 2.2 (Disagreement polymer)**.** Let ϕ be any $\mathbb{Z}$-valued height function on a domain $V \subset \mathbb{Z}^2$. Associate to each bond $e \in (\mathbb{Z}^2)^*$, dual to some edge $(x,y) \in \mathbb{Z}^2$, the gradient $(\nabla\phi)_e := \phi_x - \phi_y$, where x is taken to be the NORTH vertex if (x,y) is vertical and the WEST vertex if (x,y) is horizontal. Call e a *disagreement bond* of ϕ if $(\nabla\phi)_e \neq 0$. A *disagreement polymer* γ is a (maximal) connected component of the disagreement bonds of ϕ. For such a polymer, let D_i be the connected components of $V \setminus \gamma$. By construction, within each D_i, all the vertices that are $*$-adjacent to $V \setminus D_i$ must have the same height h_i in ϕ. Hence, we can view each γ as a triple $(\gamma, \{D_i\}, \{h_i\})$, and it makes sense to talk about the components D_i of γ and their corresponding heights h_i.

Under Dobrushin-type boundary conditions $h+1$ and h, the h level lines $\mathfrak{L}$ from Definition 2.1 are a subset of the corresponding disagreement polymer γ from Definition 2.2; see Fig. 9.

It will be convenient to reference the sets $\Delta^*_\gamma := \{u \in V \,:\, \mathrm{dist}(u,\gamma) \le 1/\sqrt{2}\}$, and $D_i^\circ := D_i \backslash \Delta^*_\gamma$.

Definition 2.3 (Energy, length and decorations of a disagreement polymer)**.** Let γ be a disagreement polymer. Its *length* $\mathscr{N}(\gamma)$ and *energy* $\mathscr{E}_\beta(\gamma)$ are defined as

$$\mathscr{N}(\gamma) = \sum_{e\in\gamma} |(\nabla\phi)_e| \quad , \quad \mathscr{E}_\beta(\gamma) := \beta \sum_{e\in\gamma} |(\nabla\phi)_e|^2 \,.$$

To describe the law of the disagreement polymer, we will use a family of decoration functions Φ on connected subsets $\mathsf{W} \subset \mathbb{Z}^2$ satisfying the following properties.

(i) The function Φ is invariant under translations, rotations by $\pi/2$, and reflection across the x and y axes.

(ii) There exists a constant $C > 0$ such that for every W, we have $|\Phi(\mathsf{W})| \le \exp(-(\beta - C)\mathbf{d}(\mathsf{W}))$, where $\mathbf{d}(\mathsf{W})$ is the size of the smallest connected set of bonds in $(\mathbb{Z}^2)^*$ containing all the boundary bonds of W.

When the boundary conditions induce a unique disagreement polymer γ, the law of γ in the $\mathbb{Z}$GFF model with no floor can be written in terms of $\mathscr{E}_\beta$ and Φ. The effect of Φ will be in the form of a sum over $\Phi(\mathsf{W})$ for $\mathsf{W} \subset V$ that intersects Δ^*_γ. That is, for a general set of dual bonds γ, let

$$\mathfrak{I}_U(\gamma) := \sum_{\substack{\mathsf{W}\subset U\\ \mathsf{W}\cap\Delta^*_\gamma\neq\emptyset}} \Phi(\mathsf{W})\,. \tag{2.1}$$

We have the following representation for $\widehat{\pi}^\eta_V(\gamma)$ under 0/1 Dobrushin-type boundary conditions η.

Proposition 2.4 ([14, Prop. 2.12])**.** *Let $V \subset \mathbb{Z}^2$ be a connected domain, and consider the $\mathbb{Z}$GFF model $\widehat{\pi}^\xi_V$ with boundary conditions ξ that are 1 on a $*$-connected path in $\partial_{\mathrm{v}} V$ and 0 elsewhere so that they induce a unique disagreement polymer $(\gamma, \{D_i\}, \{h_i\})$ in $V \cup \partial_{\mathrm{v}} V$ that contains boundary disagreements. Then for $\beta \ge \beta_0$, the law of this unique disagreement polymer is given by*

$$\widehat{\pi}^\eta_V(\gamma) = \frac{1}{\widehat{Z}^\eta_V} \exp\Big(-\mathscr{E}_\beta(\gamma) + \mathfrak{I}_V(\gamma)\Big) \tag{2.2}$$

for $\widehat{Z}^\eta_V = \widehat{Z}^\eta_V(\beta)$ and $\mathfrak{I}_V$ from Eq. (2.1) featuring a decoration function Φ as per Definition 2.3.

The case of π^η_V, the measure with the floor, is harder and requires conditions on the domain and on γ to obtain a similar form for the law. In fact, precisely studying how these conditions change as the domain size varies from mesoscopic to macroscopic is the primary technical contribution of this paper, and will be the goal of Sections 3 and 4. For reference, we first provide here the preliminary result used in [14].

Proposition 2.5 ([14, Prop. 2.3])**.** *Fix $n \ge 1$. Let $V \subset \mathbb{Z}^2$ be a connected domain, and consider the $\mathbb{Z}$GFF model $\pi^\xi_{V;F}$ with a floor at 0 imposed only on a subset $F \subset V$, and boundary conditions ξ that are $H+1-n$ on a $*$-connected path in $\partial_{\mathrm{v}} V$ and $H-n$ elsewhere so that they induce a unique disagreement polymer $(\gamma, \{D_i\}, \{h_i\})$ in $V \cup \partial_{\mathrm{v}} V$ that contains boundary disagreements. Then for $\beta \ge \beta_0$, the law of this unique disagreement polymer γ is given by*

$$\pi^\eta_{V;F}(\gamma) = \frac{1}{Z^\eta_{V;F}} \exp\Big(-\mathscr{E}_\beta(\gamma) + \mathfrak{I}_V(\gamma)\Big) \prod_{i\ge 0} \widehat{\pi}^{h_i}_{D_i^\circ}\big(\phi_x \ge 0, \forall x \in D_i^\circ \cap F\big)\,, \tag{2.3}$$

for $Z^\eta_{V;F} = Z^\eta_{V;F}(\beta, n)$ and $\mathfrak{I}_V$ from Eq. (2.1) featuring a decoration function Φ as per Definition 2.3.

Moreover, if we further have $|F| \le Le^{\kappa\sqrt{\log L}}$ and $|\partial F| \le L^{1-\delta}$ for fixed $\kappa > 0$ and $0 < \delta < \frac{1}{3}$, denote by D_0 and D_1 the regions of γ containing the boundary vertices of V at heights $H+1-n$ and $H-n$, respectively, and let

$$E := \Big\{|\gamma| \le L^{1-\delta}\,,\ \ |F| - |D_0 \cap F| - |D_1 \cap F| \le L^{1-\delta}\Big\}\,,$$

then the following holds for all $\beta \geq \beta_0$. The probability distribution given by

$$\mathfrak{p}^{\eta}_{V;F}(\gamma) := \frac{1}{\widetilde{Z}^{\eta}_{V;F}} \exp\Big(- \mathscr{E}_\beta(\gamma) + \frac{|D_0 \cap F|}{N_n} + \mathfrak{I}_V(\gamma)\Big) \prod_{i\geq 2} \widehat{\pi}^{h_i}_{D_i^\circ}(\phi_x \geq 0,\, \forall x \in D_i^\circ \cap F) \tag{2.4}$$

for $\gamma \in E$, with N_n from Eq. (1.3) and $\widetilde{Z}^{\eta}_{V;F}$ a normalizer, satisfies that for every $\gamma \in E$,

$$\pi^{\eta}_{V;F}(\gamma \mid E) = (1 + o(1))\mathfrak{p}^{\eta}_{V;F}(\gamma)\,. \tag{2.5}$$

It will be convenient, for brevity, to fold the product in Eq. (2.4) into the energy of γ, defining[3]

$$\mathscr{E}^*_\beta(\gamma) := \mathscr{E}_\beta(\gamma) - \sum_{i\geq 2} \log \widehat{\pi}^{h_i}_{D_i^\circ}(\phi_x \geq 0,\, \forall x \in D_i^\circ)\,, \tag{2.6}$$

so $\mathfrak{p}^{\eta}_{V;F}(\gamma) \propto \exp\big(- \mathscr{E}^*_\beta(\gamma) + \frac{|D_0\cap F|}{N_n} + \mathfrak{I}_V(\gamma)\big)$, an area-tilted law of a low-temperature polymer.

In Section 4, we will also be interested in studying the probability of a disagreement polymer existing when the domain has uniform boundary conditions. We begin here with a preliminary bound on the probability of a level line.

Definition 2.6. Let $\mathcal{C}_{\mathfrak{L},h}$ denote the event that $\mathfrak{L}$ is an h level line. Similarly, let $\mathcal{C}_{\gamma,h}$ denote the event that γ is a disagreement polymer such that an h level line can be formed using a subset of the bonds of γ. (To differentiate between the two, $\mathfrak{L}$ is reserved for level lines and γ is reserved for disagreement polymers.)

Lemma 2.7 ([29, Prop 4.1][4])**.** *For sufficiently large β, and h such that $\widehat{\pi}_\infty(\phi_o \geq h) \leq (1/\log L)^2$, we have for any constant boundary conditions $j \geq 0$ that*

$$\pi^{j}_{V;F}(\mathcal{C}_{\mathfrak{L},h+1}) \leq \exp\big(- (\beta - o(1))|\mathfrak{L}| + \widehat{\pi}_\infty(\phi_o > h)|\operatorname{Int}(\mathfrak{L}) \cap F|\big)\,. \tag{2.7}$$

Corollary 2.8. *For sufficiently large β, and h such that $\widehat{\pi}_\infty(\phi_o \geq h) \leq (1/\log L)^2$, suppose $F \subset V \subset \mathbb{Z}^2$ and*

$$|F| \leq \Big(\frac{3\beta}{\widehat{\pi}_\infty(\phi_o > h)}\Big)^2\,. \tag{2.8}$$

Then, under the measure $\pi^h_{V;F}$, there are no `large` *disagreement polymers in V except with probability $O(L^{-10})$.*

Proof. It suffices to show that w.h.p. there are no `large` $h+1$ level lines. (By definition, this also rules out higher level lines, and level line contours which have lower heights on the interior as opposed to the exterior can always be ruled out by a standard Peierls argument.) This then follows from Lemma 2.7 because the area term cannot make up for the cost of $|\mathfrak{L}|$. Indeed, we know by the isoperimetric inequality that $|\operatorname{Int}(\mathfrak{L}) \cap F| \leq \sqrt{|F|}\frac{|\mathfrak{L}|}{4}$. For an $h+1$ level line, the area term is then bounded above by $\widehat{\pi}_\infty(\phi_o > h)|\operatorname{Int}(\mathfrak{L}) \cap F| \leq \frac{3}{4}\beta|\mathfrak{L}|$, so that Lemma 2.7 gives $\pi^h_{V;F}(\mathcal{C}_{\mathfrak{L},h+1}) \leq \exp(-(\frac{\beta}{4} - o(1))|\mathfrak{L}|)$, whence the standard Peierls argument enumerating over $\mathfrak{L}$ concludes. ∎

As seen in the above results, the law of level lines and disagreement polymers is intimately related to the probability of large height deviations. The following theorem summarizes a few important results concerning such large deviations. These results were first proven in [29, Thm. 3.1] for the case of $V = \mathbb{Z}^2$, and extended to more general V in [14]. (The latter work also sharpened the bound on the conditional probability in Eq. (2.11).)

[3]In [14, Eq. 3.5], there is an extra term $3c(\beta)|\gamma|$ which was absorbed into the definition of the Φ functions to make the latter non-negative. This was to set up for a random walk coupling, which is not used here, hence we do not write it. Of course, the law of γ is not changed – we just add and subtract $3c(\beta)|\gamma|$.

[4]The original result was for $F = V$, but the proof extends to $F \subseteq V$.

Theorem 2.9 ([14, Thm. 2.5]). *There exist constants $\beta_0 > 0$ and $c > c' > 0$ so that the following holds for every $\beta \geq \beta_0$ and integer $h \geq 2$. Let $V \subset \mathbb{Z}^2$ be a region containing $\mathcal{B}_r(o)$, the ball of radius $r = \lceil 2ch/\log h \rceil$ centered at the origin o, as well as $\mathcal{B}_{r+1}(z)$ for a site $z \in V$. Then*

$$\exp\Big(-c\beta\frac{h}{\log h}\Big) \leq \frac{\widehat{\pi}_V^0(\phi_o = h)}{\widehat{\pi}_V^0(\phi_o = h-1)} \leq \exp\Big(-c'\beta\frac{h}{\log h}\Big)\,, \tag{2.9}$$

$$\exp\Big(-2\pi\beta\frac{h^2}{\log h} - c\beta\frac{h^2}{\log^2 h}\Big) \leq \quad \widehat{\pi}_V^0(\phi_o = h) \quad \leq \exp\Big(-2\pi\beta\frac{h^2}{\log h} + c\beta\frac{h^2}{\log^2 h}\Big)\,, \tag{2.10}$$

$$\widehat{\pi}_V^0(\phi_z = h \mid \phi_o = h) \leq \exp\Big(-c\beta\frac{h^2}{\log^2 h}\Big)\,. \tag{2.11}$$

2.2. **Polymer model, surface tension, Wulff shape.** The form of the law of γ in Propositions 2.4 and 2.5 motivate the general study of polymer models outside the context of the measures $\widehat{\pi}_V^\eta$ and π_V^η. We begin by defining the set of admissible polymers.

Definition 2.10 (Polymers from A to B). Let $V \subset \mathbb{Z}^2$ be finite and simply connected, with two marked points A, B on ∂V. Assuming without loss of generality that A is left of B, let ξ be boundary conditions of h along the upper[5] arc of ∂V from A to B and $h-1$ along the lower arc. For every $\mathbb{Z}$-height function on V with boundary condition ξ, there is a unique disagreement polymer γ which contains the boundary disagreements. Define $\mathcal{P}_V(A,B)$ as the set of all such possible disagreement polymers in this setting.[6] If V is an infinite volume domain, then define $\mathcal{P}_V(A,B) = \bigcup_{V' \subset V} \mathcal{P}_{V'}(A,B)$, where the union is over all simply connected $V' \subset V$ which have finite volume.

Consider the polymer model with weights given by

$$\widetilde{q}_U(\gamma) = \exp\big(-\mathscr{E}_\beta(\gamma) + \mathfrak{I}_U(\gamma)\big)\,,$$

and partition function

$$\widetilde{Z}_{V,U}(A,B) := \sum_{\gamma \in \mathcal{P}_V(A,B)} \widetilde{q}_U(\gamma)\,.$$

(Just as in the previous work [14], this will be different from the weights appearing in Section 4, and the two will be compared at the level of the surface tensions via [14, Prop. 4.14].)

Now fix a unit vector $\vec{\mathsf{n}}$ with angle θ. Let N be such that the point $N\vec{\mathsf{n}}$ lies on the lattice. In [14, §3.3] it was shown that the surface tension for this polymer model is well-defined:

$$\tau_\beta(\theta) := \tau_\beta(\vec{\mathsf{n}}) := -\lim_{N\to\infty} \frac{1}{\|N\vec{\mathsf{n}}\|_1} \log \widetilde{Z}_{\mathbb{Z}^2,\mathbb{Z}^2}(\mathsf{o}^*, N\vec{\mathsf{n}})\,. \tag{2.12}$$

We can extend τ_β to a function over all of $\mathbb{R}^2$ by homogeneity, and in [14, Prop. 3.11] it was shown that τ_β is analytic and satisfies and strict convexity: for any u, v not on the same line,

$$\tau_\beta(\mathsf{u}) + \tau_\beta(\mathsf{v}) > \tau_\beta(\mathsf{u} + \mathsf{v})\,.$$

Moreover, τ_β is symmetric under rotations by $\pi/4$, reflections across the x and y axis and the diagonals $y = \pm x$.

We now move away from our specific τ_β and recall some general facts concerning the Wulff shape, see, e.g., [33, §2 and §4] for a more comprehensive overview. We can define the Wulff shape

$$\mathcal{W} = \mathcal{W}(\tau) := \bigcap_{\mathsf{y} \in \mathbb{R}^2} \{\mathsf{h} \in \mathbb{R}^2 : \mathsf{h} \cdot \mathsf{y} \leq \tau(\mathsf{y})\}\,. \tag{2.13}$$

[5]The choice of which arc is upper can be made precise via a conformal map sending ∂V to the unit circle, A to $(0,-1)$, and B to $(0,1)$, and selecting the arc which maps to the upper half plane.

[6]Note that the choice of h is irrelevant here, all that matters here is that the boundary heights differ by 1.

Define also the Wulff functional on closed rectifiable curves in $\mathbb{R}^2$ as $W(\gamma) := \int_\gamma \tau(\theta_s)ds$. If $A(\gamma)$ denotes the area interior to γ, then $\partial\mathcal{W}$ is the minimizer of the Wulff functional over all curves such that $A(\gamma) = A(\partial\mathcal{W})$. Let $\mathcal{W}_1$ denote the Wulff shape with unit area, and $\mathsf{w}_1 = \mathsf{w}_1(\tau) := W(\partial\mathcal{W}_1)$. Let ℓ_τ denote the length of the smallest square which contains $\mathcal{W}_1$. One can compute that $\ell_\tau = 4\tau(0)/\mathsf{w}_1$. Observe that $\ell_\tau \geq 1$, whence this immediately implies that $4\tau(0) \geq \mathsf{w}_1$.

Now suppose we wish to minimize the Wulff functional, but we require that γ lies in the unit square and that the area $A(\gamma)$ is equal to some fixed constant $\alpha \in (0,1)$. When $\alpha \leq \frac{1}{\ell_\tau^2}$, the above discussion immediately gives an answer of $\sqrt{\alpha}\mathcal{W}_1$. When $\alpha > \frac{1}{\ell_\tau^2}$, the answer is given by a translation of Wulff shapes denoted $\mathscr{L}(\lambda)$ defined below (for λ such that $A(\mathscr{L}(\lambda)) = \alpha$).

Definition 2.11. Given τ, let $\mathscr{L}(\lambda)$ be the set obtained by first taking the union of all translates of $\frac{\mathsf{w}_1}{2\beta\lambda}\mathcal{W}_1$ inside the unit square.[7]

Next, fix $\beta, \lambda > 0$, and consider the functional

$$\mathcal{F}_\lambda(\gamma) = \mathcal{F}_\lambda(\gamma, \tau) := -\int_\gamma \tau(\theta_s)ds + \lambda\beta A(\gamma)\,.$$

This functional captures quantitatively a tradeoff between the length of a curve and its interior area, which are precisely the two main terms governing the law of a level line. Hence, in order to identify the point at which a new top level line forms, we need to know when $\mathcal{F}_\lambda(\gamma) = 0$ and provide bounds on how sensitive $\mathcal{F}_\lambda(\gamma)$ is to changes in λ. This is the content of the following lemma, which will be very useful in Section 5 (a non-quantitative version was already known in [33]).

Lemma 2.12. *Fix a constant $\alpha \in (0,1)$. Set $\lambda_* = \frac{1}{\beta}(2\tau(0) + \frac{\mathsf{w}_1}{2})$. For any loop γ in the unit square with an interior area of $|A(\gamma)| \geq \alpha$, we have that*

$$\mathcal{F}_\lambda(\gamma) \leq \begin{cases} \beta(\lambda - \lambda_*), & \text{if } \lambda \geq \lambda_* \\ -\beta(\lambda_* - \lambda)\alpha, & \text{if } \lambda < \lambda_* \text{ and } 1 - (4\tau(0)^2 - \frac{\mathsf{w}_1^2}{4})\frac{1}{\beta^2\lambda^2} \geq \alpha. \end{cases}$$

Moreover, if $\lambda \geq \lambda_*$, then $\sup_\gamma \mathcal{F}_\lambda(\gamma) = \mathcal{F}_\lambda(\mathscr{L}(\lambda))$ and $\mathcal{F}_\lambda(\mathscr{L}(\lambda)) \geq (\beta - (4\tau(0)^2 - \frac{\mathsf{w}_1^2}{4})\frac{1}{\beta\lambda^2})(\lambda - \lambda_*)$.

Proof. Consider all γ with a fixed interior area $A(\gamma) = \alpha$. As discussed above, when $\alpha \leq (\frac{\mathsf{w}_1}{4\tau(0)})^2$, $\mathcal{F}_\lambda$ is maximized over such γ when $\mathcal{F}_\lambda(\sqrt{\alpha}\mathcal{W}_1)$, and when $\alpha > (\frac{\mathsf{w}_1}{4\tau(0)})^2$ it is maximized at $\mathcal{F}_\lambda(\mathscr{L}(\mathfrak{u}))$ where $\mathfrak{u} = \mathfrak{u}(\alpha)$ is such that $\mathscr{L}(\mathfrak{u})$ has area α. Hence, to study $\sup_\gamma \mathcal{F}_\lambda(\gamma)$, it suffices to study these two terms for various α.

For the case of $\sqrt{\alpha} > \frac{\mathsf{w}_1}{4\tau(0)}$, we have

$$\int_{\partial\mathscr{L}(\mathfrak{u})} \tau(\theta_s)ds = \frac{\mathsf{w}_1^2}{2\beta\mathfrak{u}} + 4\tau(0)(1 - \frac{\mathsf{w}_1}{2\beta\mathfrak{u}}\frac{4\tau(0)}{\mathsf{w}_1}) = \frac{\mathsf{w}_1^2}{2\beta\mathfrak{u}} + 4\tau(0)(1 - \frac{2\tau(0)}{\beta\mathfrak{u}})\,,$$

$$\alpha = 1 + \frac{\mathsf{w}_1^2}{4\beta^2\mathfrak{u}^2} - \frac{4\tau(0)^2}{\beta^2\mathfrak{u}^2}\,.$$

Hence, we have

$$\mathcal{F}_\lambda(\partial\mathscr{L}(\mathfrak{u})) = -4\tau(0) + \lambda\beta + (4\tau(0)^2 - \frac{\mathsf{w}_1^2}{4})\frac{2\mathfrak{u} - \lambda}{\beta\mathfrak{u}^2}\,. \tag{2.14}$$

Since $\mathsf{w}_1 \leq 4\tau(0)$ implies that $4\tau(0)^2 - \frac{\mathsf{w}_1^2}{4} \geq 0$, we can compute the first and second derivatives of $\frac{2\mathfrak{u}-\lambda}{\mathfrak{u}^2}$ to get that $\mathcal{F}_\lambda(\partial\mathscr{L}(\mathfrak{u}))$ is maximized at $\mathfrak{u} = \lambda$. Hence it suffices to upper bound the function

$$F(\lambda) := -4\tau(0) + \lambda\beta + (4\tau(0)^2 - \frac{\mathsf{w}_1^2}{4})\frac{1}{\beta\lambda}\,.$$

[7]The factor of β is for our application with $\tau = \tau_\beta$, as we chose the convention of not normalizing the surface tension by β. For applications where τ has already been normalized, take $\beta = 1$.

We can compute that $F(\lambda) = 0$ when $\lambda = \lambda_*$. Moreover, $F'(\lambda) = \beta - (4\tau(0)^2 - \frac{\mathsf{w}_1^2}{4})\frac{1}{\beta\lambda^2} \leq \beta$. Hence, for $\lambda > \lambda_*$ we have $F(\lambda) \leq \beta(\lambda - \lambda_*)$, and for $\lambda < \lambda_*$ we have $F(\lambda) \leq (\beta - (4\tau(0)^2 - \frac{\mathsf{w}_1^2}{4})\frac{1}{\beta\lambda^2})(\lambda - \lambda_*)$.

For the case of $\sqrt{\alpha} \leq \frac{\mathsf{w}_1}{4\tau(0)}$, we have

$$\mathcal{F}_\lambda(\sqrt{\alpha}\mathcal{W}_1) = -\sqrt{\alpha}\mathsf{w}_1 + \beta\lambda\alpha\,.$$

Let $\lambda_0 := \mathsf{w}_1/(\beta\sqrt{\alpha})$, the value where $\mathcal{F}_\lambda(\sqrt{\alpha}\mathcal{W}_1) = 0$. Then, we can write $\mathcal{F}_\lambda(\sqrt{\alpha}\mathcal{W}_1) = \beta(\lambda - \lambda_0)\alpha$. By a direct computation, we obtain that $\lambda_0 > \lambda_*$ when $\sqrt{\alpha} \leq \frac{\mathsf{w}_1}{4\tau(0)}$. Thus, $\mathcal{F}_\lambda(\sqrt{\alpha}\mathcal{W}_1) \leq \beta(\lambda - \lambda_*)\alpha$. To summarize, we have proven that when $\lambda \geq \lambda_*$, we have $\sup_\gamma \mathcal{F}_\lambda(\gamma) \leq \beta(\lambda - \lambda_*)$, and when $\lambda < \lambda_*$, we have $\sup_\gamma \mathcal{F}_\lambda(\gamma) \leq (\lambda - \lambda_*)\min(\beta - (4\tau(0)^2 - \frac{\mathsf{w}_1^2}{4})\frac{1}{\beta\lambda^2}, \beta|A(\gamma)|)$.

Finally we prove the last statement. We wish to show that $F(\lambda) \geq \mathcal{F}_\lambda(\sqrt{\alpha}\mathcal{W}_1)$ for all $\sqrt{\alpha} \leq \frac{\mathsf{w}_1}{4\tau(0)}$. Fix such an α, which defines a λ_0 as above. When $\lambda \in [\lambda_*, \lambda_0]$, we have $F(\lambda) \geq 0$ while $\mathcal{F}_\lambda(\sqrt{\alpha}\mathcal{W}_1) \leq 0$. When $\lambda > \lambda_0$, we lower bound $F(\lambda) \geq \tilde{F}(\lambda) := -4\tau(0) + \lambda\beta$. The assumption on α implies that $\tilde{F}(\lambda_0) = -4\tau(0) + \frac{\mathsf{w}_1}{\sqrt{\alpha}} \geq 0$. We now conclude that $\tilde{F}(\lambda) \geq \mathcal{F}_\lambda(\sqrt{\alpha}\mathcal{W}_1)$ in this regime since both are linear functions of λ, and $\tilde{F}(\lambda)$ has a larger slope and value at λ_0. The lower bound on $F(\lambda)$ follows from the formula for the derivative computed above. ■

Remark 2.13. In our applications in Sections 5 and 6, it will often be simpler to let the argument of $\mathscr{L}$ be the scaling factor of $\mathcal{W}_1(\tau)$ directly, so that $\mathscr{L}(\ell)$ is the union of translates of $\ell\mathcal{W}_1(\tau)$. By abuse of notation, we will use both, and it will be clear from context whether the argument refers to the area tilt or the scaling factor of the Wulff shape.

When the functional $\mathcal{F}_\lambda(\gamma)$ is positive, it is maximized at $\mathcal{F}_\lambda(\mathscr{L}(\lambda))$, suggesting that the limit shape of the level lines is given by $\mathscr{L}(\lambda)$ for some λ. We recall here the partial result proven in [14] that the level lines contain such a shape, in the notation of the above remark.

Theorem 2.14 ([14, Thm. 4.4 revised as per Rem. 4.7])**.** *Consider the $\mathbb{Z}$GFF on an $L \times L$ box, for any L. Fix $n \geq 1$. Then, for any constant $C > 0$, w.h.p., the $H + 1 - n$ level line contains*

$$(1 - \frac{N_n^{1/3}e^{3C\sqrt{\log L}}}{L})L\mathscr{L}(\ell_n(1 + e^{-\frac{C}{3}\sqrt{\log L}}))\,,$$

where $\ell_n = \frac{\mathsf{w}_1(\tau_\beta)N_n}{2L}$.

In fact, in this work we will use a slightly different side length, defining

$$\ell_n^* = \frac{\mathsf{w}_1(\tau_\beta)N_n}{2L\rho_n} \tag{2.15}$$

for ρ_n defined immediately below in Eq. (3.2). Note that after we prove Proposition 3.1, the above theorem also holds with ℓ_n^* instead of ℓ_n, because the difference between ρ_n and 1 can be absorbed into the $1 + e^{-\frac{C}{3}\sqrt{\log L}}$ term.

Finally, we conclude this section with a few definitions related to disagreement polymers which will be generally useful throughout this paper.

Definition 2.15. For any disagreement polymer γ, define its outer envelope $\mathsf{OE}(\gamma)$ as the outermost loop that consists only of bonds in γ.

Definition 2.16. For a bond $b \in \gamma$, call $\mathcal{D}_b = \mathcal{D}_b(\gamma)$ the connected component of regions in $\Lambda' \setminus (\gamma \cup D_0 \cup D_1)$ containing a region that has b as part of its boundary. If there is no such region adjacent to b, then $\mathcal{D}_b = \emptyset$. Denote by $\mathfrak{D} = \mathfrak{D}(\gamma)$ the collection of all nonempty sets $\mathcal{D}_b$.

Note that $\bigcup_{\mathcal{D}\in\mathfrak{D}} \mathcal{D} = \bigcup_{b\in\gamma} \mathcal{D}_b = \bigcup_{i\geq 2} D_i$, and that each $\mathcal{D} \in \mathfrak{D}$ can be written as a union of sets $\bigcup_{i\in I} D_i$ for some finite set $I \subset \{i \geq 2\}$.

Definition 2.17. We call v a cut-point of γ if it is the end-vertex of a bond $b \in \gamma$ with $\mathcal{D}_b = \emptyset$.

3. Minimum height in a region without a floor

In this section we will study the large deviation rates for the minimum height of the set of sites in a square of side length $\ell_* \in [\sqrt{L}, 2\sqrt{L})$ under the infinite-volume measure $\widehat{\pi}_\infty$:

$$\mathsf{\xi}_n := -\frac{1}{\ell_*^2} \log \widehat{\pi}_\infty \left(\phi_x \geq -(H+1-n),\, \forall x \in Q_{\ell_*} := [\![1, \ell_*]\!]^2\right) \quad \text{for} \quad \ell_* := 2^{\lceil \frac{1}{2}\log_2 L \rceil} \tag{3.1}$$

$$\mathsf{\rho}_n := (\mathsf{\xi}_{n+1} - \mathsf{\xi}_n)\, N_n\,. \tag{3.2}$$

One of the main results of this section will establish the following estimates for $\mathsf{\xi}_n$ and $\mathsf{\rho}_n$.

Proposition 3.1. *Fix $n \geq 0$. There are absolute constants $c, C > 0$ such that*

$$\frac{\mathsf{\xi}_n}{\widehat{\pi}_\infty(\phi_o > H+1-n)} \in [1 - e^{-C\frac{\log L}{\log\log L}}, 1 - e^{-c\frac{\log L}{\log\log L}}]\,, \tag{3.3}$$

$$\mathsf{\rho}_n \in [1 - e^{-C\frac{\log L}{\log\log L}}, 1 - e^{-c\frac{\log L}{\log\log L}}]\,. \tag{3.4}$$

We will see in Section 4 that $\mathsf{\rho}_n$ is the correct pre-factor of the area tilt in the cluster expansion for the disagreement polymers. Towards that, we will show in this section that $\mathsf{\xi}_n$ is the correct rate for the probability of a region to lie above height $-h$ for $h = H+1-n$, as per the next result. In what follows, a connected subset F of the vertices of $\mathbb{Z}^2$ is said to have $g \geq 0$ *holes* if we can decompose its vertex boundary $\partial_{\mathrm{v}} F$ into $g+1$ disjoint $*$-connected closed paths.

Theorem 3.2. *Fix $g \geq 0$, $n \geq 0$, and set $h = H+1-n$. The following hold for every subsets $F \subseteq V \subset \mathbb{Z}^2$ where F is connected with g holes.*

(1) [Mesoscopic range] If $|\partial F| \vee |\partial V| \leq O(L^{2/3}\exp(\sqrt{\log L}))$ then

$$\widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in F) = (1+o(1))\exp\left(-\mathsf{\xi}_n |F|\right)\,. \tag{3.5}$$

(2) [Macroscopic range] If $|\partial F| \vee |\partial V| \leq O(L\exp(\sqrt{\log L}))$ then

$$\widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in F) \geq \exp\left(-\mathsf{\xi}_n |F| + O\big(\sqrt{L} e^{\frac{\log L}{\log\log L}}\big)\right)\,, \tag{3.6}$$

and if $\mathfrak{S}$ is the event that there are no disagreement polymers γ in ϕ with $|\gamma| \geq \log L$, then

$$\widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in F \mid \mathfrak{S}) \leq \exp\left(-\mathsf{\xi}_n |F| + O\big(\sqrt{L} e^{\frac{\log L}{\log\log L}}\big)\right)\,. \tag{3.7}$$

If $|F| \leq \left(\frac{3\beta}{\widehat{\pi}_\infty(\phi_o > h)}\right)^2$, then the last upper bound also applies to $\widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in F)$.

The above statements also hold when replacing $\widehat{\pi}^0_V$ by $\widehat{\pi}_\infty$.

3.1. Rates for the minimum height in a box.

This section is devoted to proving lower and upper bounds on the following rates: for every $\ell \geq 1$ and $h \geq 0$, define

$$\xi_{\ell,h} := -\frac{1}{\ell^2} \log \widehat{\pi}_\infty \left(\phi_x \geq -h,\, \forall x \in Q_\ell := [\![1, \ell]\!]^2\right)\,, \tag{3.8}$$

$$\bar{\xi}_{\ell,h} := -\frac{1}{\ell^2} \log \widehat{\pi}_\infty \left(\phi_x > -h,\, \forall x \in Q_\ell \;\middle|\; \phi_x \geq -h,\, \forall x \in Q_\ell\right)\,, \tag{3.9}$$

noting that

$$\bar{\xi}_{\ell,h} = \xi_{\ell,h-1} - \xi_{\ell,h}\,, \tag{3.10}$$

and that $\mathsf{\xi}_n$ and $\mathsf{\rho}_n$ from Eqs. (3.1) and (3.2) are simply $\xi_{\ell_*,H+1-n}$ and $\bar{\xi}_{\ell_*,H+1-n} N_n$, respectively.

3.1.1. *Estimating the rate* $\xi_{\ell,h}$. By FKG, if $\ell = km$ then the probability in the right-hand of Eq. (3.8) is bounded from below by $\widehat{\pi}_\infty(\phi_x \geq -h,\, \forall x \in Q_m)^k$, leading to the following observation.

Observation 3.3. *For all* $k, m \geq 1$, $\xi_{km,h} \leq \xi_{m,h}$. *In particular,* $\xi_{\ell,h} \leq \xi_{1,h} = -\log \widehat{\pi}_\infty(\phi_o \geq -h)$.

The next lemma readily infers an upper bound on $\xi_{\ell,h}$ from the above observation for the ℓ, h relevant to our application in Section 3.2, and complements it with a sharper lower bound.

Lemma 3.4. *There exist absolute constants* $\beta_0, c_0 > 0$ *such that the following holds for all* $\beta > \beta_0$. *Let* $h = H + 1 - n$ *for fixed* $n \geq 0$. *Then for all* $1 \leq \ell \leq \sqrt{L}\exp(-\frac{1}{2}\frac{\log L}{\log\log L})$ *we have*

$$1 - e^{-c_0 \frac{\log L}{\log\log L}} \leq \frac{\xi_{\ell,h}}{\widehat{\pi}_\infty(\phi_o < -h)} \leq 1 + L^{-1+o(1)}\,. \tag{3.11}$$

Moreover, this holds when replacing $\xi_{\ell,h}$ *by* $-\frac{1}{|S|}\log\widehat{\pi}_\infty(\phi_x \geq -h,\, \forall x \in S)$ *for any set* $S \subset \mathbb{Z}^2$, *not necessarily connected, of size at most* ℓ^2. *In addition, the upper bound holds for all* $\ell \geq 1$.

Proof. We begin with the upper bound. Recall from Observation 3.3 that $\xi_{\ell,h} \leq \xi_{1,h}$. Writing $\xi_{1,h} = -\log(1-t)$ for $t = \widehat{\pi}_\infty(\phi_o < -h)$, we use the fact that $1 - t \geq e^{-t-t^2}$ for $0 < t < \frac{1}{2}$ (here $t < \varepsilon_\beta < \frac{1}{2}$ for every $h \geq 0$ by the standard Peierls estimate; cf. [4]), and find that for every ℓ, h,

$$\frac{\xi_{\ell,h}}{\widehat{\pi}_\infty(\phi_o < -h)} \leq 1 + \widehat{\pi}_\infty(\phi_o < -h)\,.$$

When $h = H + 1 - n$ for fixed $n \geq 0$, as in our hypothesis, we have $\widehat{\pi}_\infty(\phi_o < -h) = L^{-1+o(1)}$ (see, e.g., [14, Eq. (2.3)] and the definition of H), establishing the upper bound in Eq. (3.11).

We proceed to the lower bound. The Bonferroni inequalities imply that

$$\widehat{\pi}_\infty(\phi_x \geq -h,\, \forall x \in Q_\ell) \leq 1 - |Q_\ell|\widehat{\pi}_\infty(\phi_o < -h) + \sum_{x,y \in Q_\ell} \widehat{\pi}_\infty(\phi_x < -h\,,\, \phi_y < -h)\,.$$

If $\mathrm{dist}(x,y) > \log L$ then $\widehat{\pi}_\infty(\phi_y < -h \mid \phi_x < -h) < \widehat{\pi}_\infty(\phi_y < -h) + L^{-10}$ by the well-known decorrelation estimates $\widehat{\pi}_\infty$ at low temperature (stemming from a routine Peierls argument; see also Eq. (3.24) below), and for all other distinct x, y we have that $\widehat{\pi}_\infty(\phi_y < -h \mid \phi_x < -h) < e^{-c\beta h^2/\log^2 h}$ by [14, Thm. 2.5, Eq. (2.8)]. Hence,

$$\begin{aligned}\frac{\sum_{x,y\in Q_\ell}\widehat{\pi}_\infty(\phi_x < -h\,,\, \phi_y < -h)}{|Q_\ell|\widehat{\pi}_\infty(\phi_o < -h)} &\leq |Q_\ell|\left(\widehat{\pi}_\infty(\phi_o < -h) + L^{-10}\right) + O((\log L)^2)e^{-c\beta h^2/\log^2 h}\\ &\leq e^{-c'\frac{\log L}{\log\log L}}\,,\end{aligned} \tag{3.12}$$

using $h^2/\log^2 h \asymp \frac{1}{\beta}\frac{\log L}{\log\log L}$, so the last summand in the first line is at most $\exp(-c\frac{\log L}{\log\log L})$, as is the first summand since $|Q_\ell| = \ell^2 \leq L\exp(-\frac{\log L}{\log\log L})$ and $\widehat{\pi}_\infty(\phi_o < -h) \leq L^{-1}\exp(o(\sqrt{\log L}))$. Overall,

$$\begin{aligned}\widehat{\pi}_\infty(\phi_x \geq -h,\, \forall x \in Q_\ell)^{1/\ell^2} &\leq \left(1 - \left(1 - e^{-c'\frac{\log L}{\log\log L}}\right)\ell^2\widehat{\pi}_\infty(\phi_o < -h)\right)^{1/\ell^2}\\ &\leq \exp\left(-\left(1 - e^{-c'\frac{\log L}{\log\log L}}\right)\widehat{\pi}_\infty(\phi_o < -h)\right)\,,\end{aligned}$$

establishing the lower bound in Eq. (3.11). ∎

Note the asymmetry between the lower and upper bounds on $\xi_{\ell,h}/\widehat{\pi}_\infty(\phi_o < -h)$ in Eq. (3.11). The following lemma refines the upper bound to show that $\xi_{\ell,h}/\widehat{\pi}_\infty(\phi_o < -h)$ is in fact bounded from above by $1 - e^{c_1\frac{\log L}{\log\log L}}$.

Lemma 3.5. *There exist absolute constants $\beta_0, c_1 > 0$ such that the following holds for all $\beta > \beta_0$. Fix $n \geq 0$ and $\delta > 0$, and let $h = H + 1 - n$. Then for all $\ell \geq L^\delta$ we have*

$$\frac{\xi_{\ell,h}}{\widehat{\pi}_\infty(\phi_o < -h)} \leq 1 - e^{-c_1 \frac{\log L}{\log\log L}}\,.$$

We need the following result, showing that the upper bound $\widehat{\pi}^0_V(\phi_x = h \mid \phi_o = h) \leq \exp(-c\beta \frac{h^2}{\log^2 h})$ for any $x \neq o$ ([14, Thm. 2.5, Eq. (2.8)]) is tight when x is a neighbor of the origin.

Claim 3.6. *Let $h \geq 1$, and let $V \subseteq \mathbb{Z}^2$ be a connected region with $V \supset B_R(o)$, where $R = \lfloor h/\log h \rfloor$. There is an absolute constant $c > 0$ such that, for large β, if o' is a neighbor of the origin o then*

$$\widehat{\pi}^0_V\left(\phi_{o'} \geq h \mid \phi_o \geq h\right) \geq e^{-c\beta \frac{h^2}{\log^2 h}}\,.$$

Proof. It was shown in [29, Eq. (3.6)] that, for a large enough absolute constant $C_1 > 0$,

$$\widehat{\pi}^0_V\Big(\phi_{o'} \geq h - C_1 \frac{h}{\log h} \mid \phi_o = h\Big) > \frac{1}{2}\,.$$

Consequently,

$$\widehat{\pi}^0_V\Big(\phi_{o'} \geq h - C_1\frac{h}{\log h}\;\Big|\;\phi_o \geq h\Big) \geq \widehat{\pi}^0_V\Big(\phi_{o'} \geq h - C_1\frac{h}{\log h}\;\Big|\;\phi_o = h\Big)\frac{\widehat{\pi}^0_V(\phi_o = h)}{\widehat{\pi}^0_V(\phi_o \geq h)} > \frac{1}{3}\,,$$

where the last inequality used that $\widehat{\pi}^0_V(\phi_o = h) = (1 - o(1))\widehat{\pi}^0_V(\phi_o \geq h)$ with the $o(1)$-term going to 0 as $h \to \infty$ (see for instance [14, Thm. 2.5, Eq. (2.6)]). Therefore,

$$\begin{aligned}\widehat{\pi}^0_V(\phi_{o'} \geq h \mid \phi_o \geq h) &\geq \frac{1}{3}\widehat{\pi}^0_V\Big(\phi_{o'} \geq h\;\Big|\;\phi_{o'} \geq h - C_1\frac{h}{\log h}\,,\,\phi_o \geq h\Big)\\ &\geq \frac{1}{3}\widehat{\pi}^0_V\Big(\phi_{o'} \geq h\;\Big|\;\phi_{o'} \geq h - C_1\frac{h}{\log h}\Big) \geq e^{-C_2\beta\frac{h^2}{\log^2 h}}\,,\end{aligned}$$

where the inequality between the lines follows from FKG, and the last inequality used the fact that $\widehat{\pi}^0_V(\phi_o = h')/\widehat{\pi}^0_V(\phi_o = h'-1) \geq e^{-c'\beta\frac{h'}{\log h'}} \geq e^{-c'\beta\frac{h}{\log h}}$ for an absolute constant c' and all $h' \in [\![h - C_1\frac{h}{\log h}, h]\!]$ (see, e.g., [14, Thm. 2.5, Eq. (2.6)]), yielding the final constant $C_2 = c'C_1$. ∎

Proof of Lemma 3.5. By dividing Q_ℓ into 1×2 and 2×1 boxes ("dominoes") and applying FKG, we see that, for a neighbor of the origin $o' \sim o$,

$$\begin{aligned}\widehat{\pi}_\infty(\phi_x \geq -h,\, \forall x \in Q_\ell) &\geq \widehat{\pi}_\infty(\phi_o \geq -h,\, \phi_{o'} \geq -h)^{\lceil \ell^2/2\rceil}\\ &= \Big(1 - 2\widehat{\pi}_\infty(\phi_o < -h) + \widehat{\pi}_\infty(\phi_o < -h,\, \phi_{o'} < -h)\Big)^{\lceil \ell^2/2\rceil}\,.\end{aligned}$$

Again using that $1 - t \geq e^{-t-t^2}$ for $0 < t < \frac{1}{2}$, if we denote

$$q = \widehat{\pi}_\infty(\phi_o < -h)\Big(1 - \frac{1}{2}\widehat{\pi}_\infty(\phi_{o'} < -h \mid \phi_o < -h)\Big) = L^{-1+o(1)}$$

then we see that

$$-\log\widehat{\pi}_\infty(\phi_x \geq -h,\, \forall x \in Q_\ell) \leq \lceil \ell^2/2\rceil 2q(1+2q) \leq (\ell^2+1)q(1+2q)\,.$$

Thus,

$$\begin{aligned}\frac{\xi_{\ell,h}}{\widehat{\pi}_\infty(\phi_o < -h)} &\leq \Big(1 + \frac{1}{\ell^2}\Big)(1+2q)\frac{q}{\widehat{\pi}_\infty(\phi_o < -h)}\\ &\leq \Big(1 + L^{-2\delta}\Big)\Big(1 + 2L^{-1+o(1)}\Big)\Big(1 - \frac{1}{2}\widehat{\pi}_\infty(\phi_{o'} < -h \mid \phi_o < -h)\Big)\,,\end{aligned}$$

and the desired result now follows from Claim 3.6. ∎

3.1.2. *Estimating the rate $\bar\xi_{\ell,h}$.* Throughout this subsection, denote

$$\mathbf{P}_{\ell,h} := \widehat{\pi}_\infty(\cdot \mid \mathscr{F}_{\ell,h}) \qquad \text{where} \qquad \mathscr{F}_{\ell,h} := \{\phi_x \ge -h,\, \forall x \in Q_\ell\}$$

for brevity, so that $\bar\xi_{\ell,h} = -\frac{1}{\ell^2}\log \mathbf{P}_{\ell,h}(\phi_x > -h,\, \forall x \in Q_\ell)$. Note that $\mathbf{P}_{\ell,h}$ is simply a $\mathbb{Z}$GFF model with a floor at $-h$ on Q_ℓ, and thus enjoys FKG (see, e.g., [14, Claim A.1]). In particular, the same reasoning that led to Observation 3.3 for $\xi_{\ell,h}$ gives its following analogue for $\bar\xi_{\ell,h}$.

Observation 3.7. *For all $k,m \ge 1$, $\bar\xi_{km,h} \le \bar\xi_{m,h}$. In particular, $\bar\xi_{\ell,h} \le \bar\xi_{1,h}$.*

As $\xi_{\ell,h} \sim \widehat{\pi}_\infty(\phi_o < -h)$, it would be natural guess when looking at Eq. (3.10) that one would have $\bar\xi_{\ell,h} \sim \widehat{\pi}_\infty(\phi_o = -h)$. This is indeed the case, as shown in the following analogue of Lemma 3.4.

Lemma 3.8. *In the setting of Lemma 3.4, we have*

$$1 - e^{-c_0\frac{\log L}{\log\log L}} \le \frac{\bar\xi_{\ell,h}}{\widehat{\pi}_\infty(\phi_o = -h)} \le 1 + L^{-1+o(1)}\,. \tag{3.13}$$

Proof. The proof will be a simple modification of the argument used to prove Lemma 3.4.

The upper bound follows directly from Observation 3.7 and the fact that $\bar\xi_{1,h} = -\log(1-t)$ for $t = \widehat{\pi}_\infty(\phi_o = -h \mid \phi_o \ge -h) = \widehat{\pi}_\infty(\phi_o = -h)/\widehat{\pi}_\infty(\phi_o \ge -h)$. As $t \le (1 + L^{-1+o(1)})\widehat{\pi}_\infty(\phi_o = -h)$ for our h, arguing as for the upper bound in Lemma 3.4 yields $\bar\xi_{1,h} \le (1+L^{-1+o(1)})\widehat{\pi}_\infty(\phi_o = -h)$.

For the lower bound, we start, as in said lemma, with

$$\mathbf{P}_{\ell,h}(\phi_x > -h,\, \forall x \in Q_\ell) \le 1 - |Q_\ell|\mathbf{P}_{\ell,h}(\phi_o = -h) + \sum_{x,y\in Q_\ell}\mathbf{P}_{\ell,h}(\phi_x = -h\,,\, \phi_y = -h)\,. \tag{3.14}$$

Recalling $\mathbf{P}_{\ell,h}(\phi_x = -h\,,\,\phi_y = -h) = \widehat{\pi}_\infty(\phi_x \le -h\,,\,\phi_y \le -h \mid \mathscr{F}_{\ell,h})$, we deduce from FKG, and the fact that $\widehat{\pi}_\infty(\phi_o = h) = (1-o(1))\widehat{\pi}_\infty(\phi_o \ge h)$ (see [29, Thm. 3.1]) that

$$\frac{\sum_{x,y\in Q_\ell}\mathbf{P}_{\ell,h}(\phi_x = -h\,,\,\phi_y = -h)}{|Q_\ell|\widehat{\pi}_\infty(\phi_o = -h)} \le \frac{\sum_{x,y\in Q_\ell}\widehat{\pi}_\infty(\phi_x \le -h\,,\,\phi_y \le -h)}{(1-o(1))|Q_\ell|\widehat{\pi}_\infty(\phi_o \le -h)} \le e^{-c\frac{\log L}{\log\log L}}\,, \tag{3.15}$$

where the last transition is by Eq. (3.12). Further note that

$$\begin{aligned}\mathbf{P}_{\ell,h}(\phi_o = -h) &\ge \widehat{\pi}_\infty(\phi_o = -h) - \widehat{\pi}_\infty(\phi_o = -h,\, \mathscr{F}^c_{\ell,h})\\ &= (1 - \widehat{\pi}_\infty(\mathscr{F}^c_{\ell,h} \mid \phi_o = -h))\widehat{\pi}_\infty(\phi_o = -h)\,,\end{aligned}$$

and as usual, we can control $\widehat{\pi}_\infty(\mathscr{F}^c_{\ell,h} \mid \phi_o = -h)$ via a union bound, splitting the treatment of $x \in Q_\ell$ into x at distance larger than $\log L$ from the origin and those within said distance. As argued above Eq. (3.12), in the former case,

$$|Q_\ell|\widehat{\pi}_\infty(\phi_x < -h \mid \phi_o = -h) < \ell^2\left(\widehat{\pi}_\infty(\phi_o < -h) + L^{-10}\right) < e^{-(1-o(1))\frac{\log L}{\log\log L}}\,,$$

by the assumption on ℓ and the fact $\widehat{\pi}_\infty(\phi_o < -h) \le L^{-1}e^{c\beta\frac{h}{\log h}} \le L^{-1}e^{o(\sqrt{\log L})}$; in the latter case,

$$O(\log^2 L)\widehat{\pi}_\infty(\phi_x < -h \mid \phi_o = -h) < O(\log^2 L)\exp\Big(-c\beta\frac{h^2}{\log^2 h}\Big) \le \exp\Big(-c'\frac{\log L}{\log\log L}\Big)\,.$$

Combined, we find that

$$\widehat{\pi}_\infty(\mathscr{F}^c_{\ell,h} \mid \phi_o = -h) \le e^{-c'\frac{\log L}{\log\log L}}\,, \tag{3.16}$$

and infer that

$$\mathbf{P}_{\ell,h}(\phi_o = -h) \ge \Big(1 - e^{-c\frac{\log L}{\log\log L}}\Big)\,\widehat{\pi}_\infty(\phi_o = -h)\,. \tag{3.17}$$

Plugging Eqs. (3.15) and (3.17), into Eq. (3.14) yields

$$\begin{aligned}\mathbf{P}_{\ell,h}(\phi_x > -h,\, \forall x \in Q_\ell) &\le 1 - \Big(1 - e^{-c\frac{\log L}{\log\log L}}\Big)|Q_\ell|\widehat{\pi}_\infty(\phi_o = -h)\\ &\le \exp\bigg(-\Big(1 - e^{-c\frac{\log L}{\log\log L}}\Big)|Q_\ell|\widehat{\pi}_\infty(\phi_o = -h)\bigg)\,,\end{aligned}$$

yielding the required lower bound on $\bar{\xi}_{\ell,h}$. ■

3.2. **Minimum height under the no-floor measure.** The following proposition is our main tool for expressing the probability under $\widehat{\pi}^0_V$ that $\phi_x \ge -h$ on a subset F to the rate $\xi_{\ell_0,h}$ for a smaller scale ℓ_0. It will be used to derive Theorem 3.2 using appropriate choices of ℓ_0 (namely, $\ell_0 = L^{-1/3+o(1)}$ for the mesoscopic scale and $\ell_0 = L^{-1/2-o(1)}$ for the macroscopic scale), as well as to relate the corresponding $\xi_{\ell_0,h}$ to the rate $\xi_{\ell_*,h}$, for ℓ_* from Eq. (3.1), appearing in Theorem 3.2.

Proposition 3.9. *Fix $g \ge 0$, $n \ge 0$, and set $h = H + 1 - n$. There are absolute constants $c, C > 0$ such that the following holds for every $\ell_0 \in [\![\log^2 L, \sqrt{L}e^{-\frac{1}{2}\frac{\log L}{\log\log L}}]\!]$ and subsets $F \subseteq V \subset \mathbb{Z}^2$ where F has g holes and $|\partial F| \le L^{3/2}$. Let $\Xi = \Xi_1 + \Xi_2 + \Xi_3$, where*

$$\Xi_1 = \frac{|\partial V|}{L} e^{C\frac{\log L}{\log\log L}}\,, \qquad \Xi_2 = \frac{|F|}{\ell_0 L} e^{\sqrt{\log L}}\,, \qquad \Xi_3 = \frac{|\partial F|\ell_0}{L} e^{-c\frac{\log L}{\log\log L}}\,. \tag{3.18}$$

(i) If $|F| \le \ell_0 L e^{-\sqrt{\log L}}$, then

$$\widehat{\pi}^0_V(\phi_x \ge -h,\, \forall x \in F) = \exp\left(-\xi_{\ell_0,h}|F| + O(\Xi) + o(L^{-5})\right)\,. \tag{3.19}$$

(ii) Otherwise, if $\mathfrak{S}$ is the event that there are no disagreement polymers γ with $|\gamma| \ge \log L$, then

$$\widehat{\pi}^0_V(\phi_x \ge -h,\, \forall x \in F) \ge \exp\left(-\xi_{\ell_0,h}|F| + O(\Xi)\right)\,, \tag{3.20}$$

$$\widehat{\pi}^0_V(\phi_x \ge -h,\, \forall x \in F \mid \mathfrak{S}) \le \exp\left(-\xi_{\ell_0,h}|F| + O(\Xi) + o(L^{-5})\right)\,. \tag{3.21}$$

If $|F| \le \big(\frac{3\beta}{\widehat{\pi}_\infty(\phi_o > h)}\big)^2$, then the last upper bound also applies to $\widehat{\pi}^0_V(\phi_x \ge -h,\, \forall x \in F)$.

The same holds under $\widehat{\pi}_\infty$, in which case the error term Ξ_1 can be omitted from Ξ. Separately, in the special case where $F = [\![1,\ell]\!]^2$ with $\ell_0 + 10\lceil\log L\rceil \mid \ell$, the error term Ξ_3 can be omitted from Ξ.

Proof. Let us first establish Eqs. (3.19) to (3.21), then argue why the proof extends to $\widehat{\pi}_\infty$ and how to omit Ξ_1 in that case, as well as Ξ_3 when F is a square of side length ℓ divisible by $\ell_0 + 10\lceil\log L\rceil$.

Partition $\mathbb{Z}^2$ into squares R_i of side length $\ell_0 + 10\lceil\log L\rceil$, and let $Q_i \subset R_i$ be the concentric squares of side length ℓ_0. With this tiling of $\mathbb{Z}^2$ in hand, partition F into the sets

$$\begin{aligned}
F_1 &= \{x \in F \,:\, \operatorname{dist}(x, \partial V) < 5\log L\} && \text{(near boundary)}\,,\\
F_2 &= \bigcup_{i:\, R_i \subset F} R_i \setminus (Q_i \cup F_1) && \text{(square annuli)}\,,\\
F_3 &= \bigcup_{i:\, R_i \cap F^c \neq \emptyset} Q_i \cap (F \setminus F_1) && \text{(cut off squares)}\,,\\
F_4 &= F \setminus (F_1 \cup F_2 \cup F_3) && \text{(full squares)}\,.
\end{aligned}$$

We will first argue that

$$\widehat{\pi}^0_V(\phi_x \ge -h,\, \forall x \in F) = \widehat{\pi}^0_V\left(\phi_x \ge -h,\, \forall x \in F_3 \cup F_4\right) e^{O(\Xi_1) + O(\Xi_2)}\,, \tag{3.22}$$

$$\xi_{\ell_0,h}(|F_1| + |F_2|) = O(\Xi_1) + O(\Xi_2)\,, \tag{3.23}$$

for Ξ_1, Ξ_2 from Eq. (3.18), both assumed to be $o(1)$ or the bound is trivial. To see this, note first that the right-hand of Eq. (3.22) is clearly an upper bound on the left-hand (without any error term) by event inclusion. For a lower bound, let us denote

$$p_i := \max_{x\in F_i} \widehat{\pi}^0_V(\phi_x < -h) \qquad \text{for } i=1,2\,,$$

and apply FKG to see that

$$\begin{aligned}\widehat{\pi}^0_V(\phi_x \geq -h\,, \forall x\in F) &\geq \widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in F_3\cup F_4)\prod_{i=1,2}(1-p_i)^{|F_i|}\\ &\geq \widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in F_3\cup F_4)\prod_{i=1,2}(1-p_i|F_i|)\,,\end{aligned}$$

and it remains to show that $p_i|F_i| \leq \Xi_i$ and $\xi_{\ell_0,h}|F_i|\leq \Xi_i$ for $i=1,2$ and the Ξ_i's as per Eq. (3.18).

(1) Consider $i=1$. It was shown in [14, §2.1] (specifically, combining Lemma 2.8 with Eq. (2.7) there) that there exists some absolute constant $c>0$ such that, for every V, $h\geq 2$ and large enough β,

$$\widehat{\pi}^0_V(\phi_x = h) \leq \widehat{\pi}_\infty(\phi_o = h)e^{c\beta\frac{h^2}{\log^2 h}}\,.$$

As $\widehat{\pi}_\infty(\phi_o = h)/\widehat{\pi}_\infty(\phi_o = h-1)\leq \exp(-c\beta h/\log h)$ for all $h\geq 2$ ([29, Eq. (3.1)]), in our setting of $h=H+1-n$ we have $\widehat{\pi}_\infty(\phi_o<-h) = L^{-1}\exp(O(\sqrt{\frac{\log L}{\log\log L}}))$ and so, for some other absolute constant $c>0$,

$$p_1 \leq L^{-1}e^{c\frac{\log L}{\log\log L}}\,.$$

The fact that $|F_1| = O(|\partial V|\log^2 L)$ now shows $p_1F_1\leq \Xi_1$ if the constant C is large enough. The bound $\xi_{\ell,h}\leq (1+o(1))\widehat{\pi}_\infty(\phi_o<-h)$ as per Lemma 3.4 shows that $\xi_{\ell_0,h}|F_1|\leq\Xi_1$ as well (with room to spare, as $\widehat{\pi}_\infty(\phi_o = h)$ is better controlled than $\widehat{\pi}^0_V(\phi_x=h)$).

(2) Consider $i=2$. The upper bound we gave on p_1 also holds for p_2, however it is imperative that in this case we improve the term $\exp(c\frac{\log L}{\log\log L})$ given there. To this end, we use the standard decay of correlations in the low temperature $\mathbb{Z}$GFF model (e.g., [4]), whereby

$$\|\widehat{\pi}^0_W(\phi\restriction_U\in\cdot) - \widehat{\pi}_\infty(\phi\restriction_U\in\cdot)\|_{\mathrm{TV}} \leq |\partial U|e^{-c\beta r} \qquad \text{if } U\subset W\subset \mathbb{Z}^2 \text{ has } \mathrm{dist}(U,\partial W)>r \tag{3.24}$$

(see, e.g., [14, Eq. (2.20)] where this was stated for $U=\mathcal{B}_r(o)$, a ball of radius r about the origin; the exact same reasoning holds when U is a general domain, as it hinges on a Peierls argument to encapsulate U in a loop of height 0 in $W\setminus U$). By definition, we excluded F_1 from the rectangular annuli $R_i\setminus Q_i$ within F_2, and thus, for large enough β,

$$p_2 \leq \widehat{\pi}_\infty(\phi_o<-h) + e^{-c\beta\log L} \leq (1+o(1))\widehat{\pi}_\infty(\phi_o<-h)\,.$$

Using $\widehat{\pi}_\infty(\phi_o = h)/\widehat{\pi}_\infty(\phi_o=h-1)\leq\exp(-c\beta h/\log h)$, we get

$$p_2 \leq (1+o(1))\widehat{\pi}_\infty(\phi_o>h) \leq \frac{1}{L}e^{c(n+1)\sqrt{\beta\frac{\log L}{\log\log L}}} = \frac{1}{L}e^{o(\sqrt{\log L})}\,.$$

Combining this with the fact that

$$|F_2| \leq 2\frac{10\lceil\log L\rceil}{\ell_0+10\lceil \log L\rceil}|F| = O\Big(\frac{|F|\log L}{\ell_0}\Big)$$

shows $p_2|F_2| = o(\Xi_2)$ if L is large enough. Since $\xi_{\ell_0,h}\leq(1+o(1))\widehat{\pi}_\infty(\phi_o<-h)$ as per Lemma 3.4, we obtain the exact same bound for $\xi_{\ell_0,h}|F_2|$.

We have thus established Eqs. (3.22) and (3.23), reducing the proof of Eq. (3.20) to proving that

$$\widehat{\pi}^0_V(\phi_x\geq -h,\,\forall x\in F_3\cup F_4) \geq \exp\Big(-\xi_{\ell_0,h}\left(|F_3|+|F_4|\right)+o(\Xi_3)\Big)\,, \tag{3.25}$$

with Ξ_3 as per Eq. (3.18) with the following specific constant: letting $c_0 > 0$ denote the absolute constant from Lemma 3.4, we set

$$\Xi_3 := \frac{|\partial F|\ell_0}{L} e^{-(c_0/2)\frac{\log L}{\log\log L}}\,. \tag{3.26}$$

Similarly, once Eq. (3.25) is established, the proof of Eq. (3.19) will be reduced to showing that, for $|F| \leq \ell_0 L e^{-\sqrt{\log L}}$,

$$\widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in F_3 \cup F_4) \leq \exp\Big(-\xi_{\ell_0,h}\left(|F_3| + |F_4|\right) + o(\Xi_3) + o(L^{-5})\Big)\,, \tag{3.27}$$

and the proof of Eq. (3.21) will be reduced to showing that, without this restriction on $|F|$,

$$\widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in F_3 \cup F_4 \mid \mathfrak{S}) \leq \exp\Big(-\xi_{\ell_0,h}\left(|F_3| + |F_4|\right) + o(\Xi_3) + o(L^{-5})\Big)\,. \tag{3.28}$$

We begin with the lower bound in Eq. (3.25), which will be a direct consequence of FKG:

$$\widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in F_3 \cup F_4) \geq \Big(\min_{x\in F_3} \widehat{\pi}^0_V(\phi_x \geq -h)\Big)^{|F_3|} \prod_{Q_i \subset F_4} \widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in Q_i)\,.$$

Appealing once more to the coupling in Eq. (3.24) (bearing in mind that every $x \in F_3 \cup F_4$ is at distance at least $\log L$ from ∂F by construction), it follows that, for every $x \in F_3$,

$$\widehat{\pi}^0_V(\phi_x \geq -h) \geq 1 - \widehat{\pi}_\infty(\phi_o < -h) - L^{-10} \geq \exp\Big(-\widehat{\pi}_\infty(\phi_o < -h) - L^{-2}e^{o(\sqrt{\log L})}\Big)\,,$$

provided that β is large enough (using $\widehat{\pi}_\infty(\phi_o < -h) \leq L^{-1}\exp(o(\sqrt{\log L}))$ in the last inequality). Similarly, for every Q_i accounted for in F_4,

$$\widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in Q_i) \geq \widehat{\pi}^0_\infty(\phi_x \geq -h,\, \forall x \in Q_i) - L^{-10} \geq e^{-\xi_{\ell_0,h}|Q_i| - L^{-9-o(1)}}\,.$$

Therefore, summing the $L^{-9-o(1)}$ error over a crude bound on $|F_4| \leq |F| \leq |\partial F|^2 \leq L^3$,

$$\widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in F_3 \cup F_4) \geq \exp\Big(-\Big(\widehat{\pi}_\infty(\phi_o < -h) + L^{-2}e^{o(\sqrt{\log L})}\Big)|F_3| - \xi_{\ell_0,h}|F_4| - o(L^{-5})\Big)\,.$$

To infer Eq. (3.25) from the last display, it remains to show that

$$\Big|\widehat{\pi}_\infty(\phi_o < -h) + L^{-2}e^{o(\sqrt{\log L})} - \xi_{\ell_0,h}\Big|\,|F_3| \leq o(\Xi_3)\,. \tag{3.29}$$

By Lemma 3.4, $\xi_{\ell_0,h} = (1 + O(e^{-c_0\frac{\log L}{\log\log L}}))\widehat{\pi}_\infty(\phi_o < -h)$ for the same $c_0 > 0$ from above; thus,

$$\Big|\widehat{\pi}_\infty(\phi_o < -h) + L^{-2}e^{o(\sqrt{\log L})} - \xi_{\ell_0,h}\Big|\,|F_3| \leq e^{-(c_0 - o(1))\frac{\log L}{\log\log L}}\frac{|F_3|}{L}\,,$$

where we again plugged in the fact that $\widehat{\pi}_\infty(\phi_o < -h) = L^{-1}\exp(o(\sqrt{\log L}))$ for $h = H + 1 - n$ with fixed $n \geq 0$. Revisiting the definition of Ξ_3 in Eq. (3.26), it is thus left to show that

$$|F_3| = O(|\partial F|\ell_0)\,. \tag{3.30}$$

(This is in lieu of the trivial bound $|F_3| \leq O(|\partial F|\ell_0^2)$, which would not suffice.) We will argue that

$$|F_3| \leq 4(|\partial F|\ell_0 + (g+1)\ell_0^2) = O(|\partial F|\ell_0)\,.$$

To see this, view the squares R_i partitioning the plane $\mathbb{Z}^2$ as the squares of a chessboard, and let R_{i_j} be their subset corresponding to selecting only the light squares that lie in even files. By symmetry, the claim will follow once we show $|F_3 \cap \bigcup R_{i_j}| \leq |\partial F|\ell_0 + (g+1)\ell_0^2$. Recall our hypothesis that F has at most g holes (this is the only point in the proof that requires that assumption); that is, $\partial_{\mathsf{v}} F$ can be decomposed into at most $g+1$ disjoint $*$-connected closed paths $P_0, \ldots, P_g$. Each P_i visits at most $1 + |P_i|/\ell_0$ distinct squares R_{i_j} (the shortest path from $R_{i_{j_1}}$ to the next $R_{i_{j_2}}$ has length at least $\ell_0 + 10\lceil \log L\rceil > \ell_0$), and each such square contributes at most ℓ_0^2 sites to F_3 via its corresponding Q_{i_j}. Thus, $|F_3 \cap \bigcup R_{i_j}| \leq (g + 1 + |\partial F|/\ell_0)\ell_0^2$, as claimed. Having obtained

Eq. (3.30), we arrive at Eq. (3.29) and thus conclude the proof of the lower bound as per Eq. (3.25). Consequently, Eq. (3.20) is established.

We now turn to the proofs of the upper bounds as per Eqs. (3.27) and (3.28). To that end, we wish to expose the disagreement polymers along a grid of sites that will separate the Q_i's in $F_3 \cup F_4$, thereby inducing a product measure on Q_i's that are each surrounded by a loop of sites at height 0. Recalling that every Q_i that is part of $F_3 \cup F_4$ has $\mathrm{dist}(Q_i, F_3 \cup F_4 \setminus Q_i) \geq 5 \log L$, and moreover, distinct such Q_i, Q_j are at distance at least $10 \log L$ from each other, we define

$$U := \{x \in F \ : 2\log L \leq \mathrm{dist}(x, F_3 \cup F_4) \leq 3 \log L\} \ ,$$

and let

$$\Gamma_U := \{\gamma \in \phi \text{ incident to some } x \in U\} \ , \quad \Gamma_U^\dagger := \{\gamma \in \Gamma_U \ : \ |\gamma| \geq \log L\} \ .$$

If we further define the events

$$A = \left\{ \exists \gamma \in \Gamma_U^\dagger \text{ with } |\gamma| \geq \tfrac{1}{10}\ell_0 \right\} \ , \quad B_k = \left\{ |\Gamma_U^\dagger| = k \right\}$$

then the standard Peierls estimate yields that

$$\widehat{\pi}_V^0(A) \leq e^{-(\beta - C)\ell_0} \ ,$$

and, for every $k \geq 1$,

$$\widehat{\pi}_V^0(B_k) \leq |F|^k e^{-(\beta-C)k \cdot \frac{1}{10}\log L} \leq L^{-(\beta - C')k/10} \, .$$

For Eq. (3.27), recall $\widehat{\pi}_\infty(\phi_o < -h) \leq L^{-1} e^{o(\sqrt{\log L})}$, and note that if $|F| \leq \ell_0 L e^{-\sqrt{\log L}}$ then

$$|F| \widehat{\pi}_\infty(\phi_o < -h) \leq e^{-(1-o(1))\sqrt{\log L}} \ell_0 = o(\ell_0) \, , \tag{3.31}$$

in which case we can absorb the additive $\widehat{\pi}_V^0(A)$ term as a multiplicative $1 + O(e^{-(\beta - C - o(1))\ell_0})$ factor in front of our main term of $e^{-\xi_{\ell_0,h}|F|}$; that is, this adds a negligible term of $O(e^{-c\ell_0}) = o(L^{-5})$ to our exponent (using here that $\ell_0 \geq \log^2 L$). Therefore, it will suffice to prove Eq. (3.27) conditional on A^c. (NB. that the same holds for $\widehat{\pi}_V^0(\bigcup\{B_k \ : \ k \geq \ell_0/\log L\})$, so in principle we could have restricted also to $\bigcup_{k \leq \ell_0} B_k$, but the bound on $\widehat{\pi}_V^0(B_k)$ is good enough to allow to sum over all k's without needing this threshold.)

Consider some Q_i intersecting $F_3 \cup F_4$. We say that Q_i is `good` if one can find in U, for each connected component S of $Q_i \cap (F_3 \cup F_4)$, a $*$-adjacent loop of sites $\mathcal{C}$ at distance at most $3 \log L$ from S such that $\phi\restriction_{\mathcal{C}} = 0$. (NB. If Q_i is not `good` then there must be some $\gamma \in \Gamma_U^\dagger$ intersecting R_i.)

By Eq. (3.24) and the definition of $\xi_{\ell_0,h}$ (as well as $e^{-\xi_{\ell_0,h}|Q_i|} \geq 1 - \ell_0^2 L^{-1} e^{o(\sqrt{\log L})} \geq 1 - o(1)$, so we can write the additive $O(L^{-10})$ coupling error as a $1 + O(L^{-10})$ factor), for every $Q_i \subset F_4$,

$$\widehat{\pi}_V^0(\phi_x \geq -h, \, \forall x \in Q_i \mid \Gamma_U, \, Q_i \text{ is } \texttt{good}) = \exp\left(-\xi_{\ell_0,h}|Q_i| + O(L^{-10})\right) \, .$$

Similarly, if $Q_i \cap F_3 \neq \emptyset$ then by Lemma 3.4 (noting that indeed $\ell_0 \leq \sqrt{L}\exp(-\frac{\log L}{\log\log L})$ and appealing to the remark in that lemma that allows its application to $Q_i \cap F_3$, a region that is not necessarily connected), for the constant $c_0 > 0$ from that lemma we have

$$\widehat{\pi}_V^0(\phi_x \geq -h, \, \forall x \in Q_i \cap F_3 \mid \Gamma_U, \, Q_i \text{ is } \texttt{good}) \leq \exp\left(-\widehat{\pi}_\infty(\phi_o < -h)\left(1 + O(e^{-c_0 \frac{\log L}{\log\log L}}\right)|Q_i \cap F_3|\right) \, .$$

All the `good` Q_i's are conditionally independent given Γ_U (stemming from the fact that their pairwise distances are all at least $10 \log L$). Finally, for each $k = 0, \ldots, \ell_0/\log L$ and every realization of Γ_U consistent with $A^c \cap B_k$, all the Q_i's accounted in $F_3 \cup F_4$ are `good` except for at most $4k$.

Note that the two terms that we have encountered in the last two displays (when conditioning on Q_i being `good`), namely $\xi_{\ell_0,h}|Q_i|$ or $\widehat{\pi}_\infty(\phi_o < -h)|Q_i \cap F_3|$, are each at most

$$(1 + o(1))\widehat{\pi}_\infty(\phi_o < -h)\ell_0^2 \leq e^{-(1-o(1))\frac{\log L}{\log\log L}} < e^{-\frac{3}{4}\frac{\log L}{\log\log L}}$$

for any sufficiently large L, using here our upper bound on ℓ_0. We can therefore account for the (at most $4k$) missing Q_i's that are not `good` from each of the last two displays—increasing $\sum |Q_i|$

to $|F_4|$ in the first and $\sum |Q_i \cap F_3|$ to $|F_3|$ in the second—by adding a term $4ke^{-\frac34 \frac{\log L}{\log\log L}}$ in the exponent. That is, summing the $O(L^{-10})$ error over $|F_4| \le |F| \le L^3$ we can infer that

$$\widehat\pi_V^0(\phi_x \ge -h,\, \forall x \in F_3 \cup F_4 \mid \Gamma_U, A^c \cap B_k) \le \exp\bigg(- \xi_{\ell_0,h}|F_4| + O(L^{-7}) \\ - \widehat\pi_\infty(\phi_o < -h)\big(1 + O(e^{-c_0 \frac{\log L}{\log\log L}})\big)|F_3| \\ + 4ke^{-\frac34 \frac{\log L}{\log\log L}}\bigg)\,.$$

Turning our attention to the term $4ke^{-\frac34 \frac{\log L}{\log\log L}}$, we have

$$\sum_{k\ge 0} \widehat\pi_V^0(B_k) e^{4ke^{-\frac34 \frac{\log L}{\log\log L}}} \le \sum_{k\ge0} \exp\bigg(- \Big[(\beta - C)\log L - 4e^{-\frac34\frac{\log L}{\log\log L}}\Big]k\bigg) \le 1 + O(L^{-10})\,,$$

provided that β is large enough. Combining the last two displays, we see that

$$\widehat\pi_V^0\big(\phi_x \ge -h, \forall x\in F_3\cup F_4 \mid \Gamma_U, A^c\big) \\ \le \exp\Big(- \widehat\pi_\infty(\phi_o < -h)\big(1 + O(e^{-c_0 \frac{\log L}{\log\log L}})\big)|F_3| - \xi_{\ell_0,h}|F_4| + o(L^{-5})\Big)\,. \tag{3.32}$$

We have already seen in Eq. (3.29) that we may replace $\widehat\pi_\infty(\phi_o < -h)|F_3|$ by $\xi_{\ell_0,h}|F_3| + o(\Xi_3)$. The same reasoning allows us to neglect the term $\exp(-c_0 \frac{\log L}{\log\log L})\widehat\pi_\infty(\phi_o < -h)|F_3|$, as Eq. (3.30) gives

$$e^{-c_0 \frac{\log L}{\log\log L}} \widehat\pi_\infty(\phi_o < -h)|F_3| \le e^{-(c_0 - o(1))\frac{\log L}{\log\log L}} \frac{|\partial F|\ell_0}{L} = o(\Xi_3)\,.$$

again by the definition of Ξ_3 in Eq. (3.26) and the choice of the constant in the exponent there. We have therefore established Eq. (3.27), which, when combined with Eq. (3.25), completes Eq. (3.19).

To obtain Eq. (3.28), the situation is simpler, since A is precluded by the conditioning on $\mathfrak{S}$, which in addition implies that B_0 holds (that is, $\Gamma_U^\dagger = \emptyset$). Each Q_i intersecting $F_3 \cup F_4$ is thus guaranteed to be `good` conditionally on $\Gamma_U, \mathfrak{S}$ (and these Q_i's are conditionally independent). Thus, the only modification we need to make to the preceding argument is to control the effect that the conditioning on $\mathfrak{S}$ has on the probability that $\phi_x \ge -h$ in Q_i. To this end, we may bound

$$\widehat\pi_V^0(\phi_x \ge -h,\, \forall x\in Q_i \mid \Gamma_U,\, \mathfrak{S}) \le \frac{\widehat\pi_V^0(\phi_x \ge -h,\, \forall x \in Q_i \mid \Gamma_U,\, Q_i \text{ is } \texttt{good})}{\widehat\pi_V^0(\mathfrak{S})} \\ \le \big(1 + O(\widehat\pi_V^0(\mathfrak{S}^c)\big)\,\widehat\pi_V^0(\phi_x \ge -h,\, \forall x\in Q_i \mid \Gamma_U,\, Q_i \text{ is } \texttt{good})\,,$$

and the same holds for $\widehat\pi_V^0(\phi_x \ge -h,\, \forall x \in Q_i \cap F_3 \mid \Gamma_U,\, \mathfrak{S})$. Since $\widehat\pi_V^0(\mathfrak{S}^c) < L^{-10}$, accumulating this error over all the Q_i's results in an additive term of $O(L^{-8})$ in the exponent, hence the analog of Eq. (3.32) holds true for $\widehat\pi_V^0(\cdot \mid \Gamma_U,\, \mathfrak{S})$. This establishes Eq. (3.28), and hence also Eq. (3.21).

In the special case where $|F| \le \big(\frac{3\beta}{\widehat\pi_\infty(\phi_o > h)}\big)^2$, we obtain the unconditional version of Eq. (3.21) as a consequence of Corollary 2.8. Let us apply the elementary bound $\mathbb{P}(A) \le \frac{\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A|B)}{\mathbb{P}(B|A)}$, valid whenever $\mathbb{P}(A\cap B) > 0$, to the events $A = \{\phi_x \ge -h,\, \forall x \in F\}$ and $B = \mathfrak{S}$ under the measure $\widehat\pi_V^0$. In that notation, we wish to extend our bound on $\mathbb{P}(A \mid B)$ to $\mathbb{P}(A)$; it thus suffices to show that

$$\widehat\pi_V^0(\mathfrak{S} \mid \phi_x \ge -h,\, \forall x \in F) \ge 1 - o(L^{-5})\,.$$

Indeed, the quantity on the left is nothing but $\pi_{V;F}^h(\mathfrak{S})$, which is $1 - O(L^{-10})$ by Corollary 2.8.

It remains to address the remark in the proposition pertaining to the improved error terms. When working under $\widehat\pi_\infty$, as opposed to $\widehat\pi_V^0$, we may apply the same proof only with $F_1 = \emptyset$, as there is no longer a need to exclude the sites near ∂V in order to support a coupling with $\widehat\pi_\infty$. Hence, we avoid the error term Ξ_1 formerly associated with F_1. Similarly, when F is a square of side length ℓ such that $\ell_0 + 10\lceil \log L\rceil \mid \ell$, we can partition F into squares R_i's without any residue, so $F_3 = \emptyset$, forgoing the error term Ξ_3 that was formerly associated with F_3, as claimed. ∎

3.3. **Proofs of the bound on ξ_n in Proposition 3.1, and Theorem 3.2.** Recall that ξ_n is $\xi_{\ell_*,h}$ for ℓ_* as per Eq. (3.1). In Section 3.1 we gave bounds on $\xi_{\ell,h}$ for $\ell < \sqrt{L}\exp(-C\frac{\log L}{\log\log L})$, and with the help of Proposition 3.9, we can now extend those to $\ell_* \asymp \sqrt{L}$ to derive Eq. (3.3) from Proposition 3.1. Thereafter, a more delicate application of that proposition will yield Theorem 3.2.

Proof of Proposition 3.1, Eq. (3.3). We will relate $\xi_{\ell_*,h}$ to $\xi_{\ell_0,h}$ for

$$\ell_0 = \lfloor L^{1/4}\rfloor\,.$$

Applying Proposition 3.9 for $\widehat\pi_\infty$ and $F = [\![1,\ell_*]\!]^2$, we get from Eq. (3.19) (noting that $|F| = \ell_*^2 \asymp L$, which is indeed smaller than $\ell_0 L e^{-\sqrt{\log L}} = L^{5/4-o(1)}$, fulfilling the hypothesis[8]) that

$$\widehat\pi_\infty(\phi_x \geq -h,\,\forall x\in F) = \exp\bigg(-\xi_{\ell_0,h}\ell_*^2 + O\Big(\frac{\ell_*^2}{\ell_0}L^{-1+o(1)} + \ell_*\ell_0 L^{-1-o(1)}\Big) + o(L^{-5})\bigg)\,.$$

The dominant error term is $(\ell_*^2/\ell_0)L^{-1+o(1)}$ (originating from Ξ_2 in the above proposition), and so

$$\xi_{\ell_*,h} = -\frac{1}{\ell_*^2}\log\widehat\pi_\infty(\phi_x \geq -h,\,\forall x\in F) = \xi_{\ell_0,h} + L^{-5/4+o(1)}\,. \tag{3.33}$$

In particular, since $\xi_{\ell_0,h} = (1+o(1))\widehat\pi_\infty(\phi_o < -h) = L^{-1+o(1)}$, we infer that

$$\frac{\xi_n}{\widehat\pi_\infty(\phi_o < -h)} = \frac{\xi_{\ell_*,h}}{\widehat\pi_\infty(\phi_o < -h)} = \Big(1 + L^{-1/4+o(1)}\Big)\frac{\xi_{\ell_0,h}}{\widehat\pi_\infty(\phi_o < -h)}\,;$$

thus, Eq. (3.3) follows from Lemmas 3.4 and 3.5 for $\xi_{\ell_0,h}$.

Having established Eq. (3.3), we note that at this point we can also infer the following weaker version of the bound in Eq. (3.4) on ρ_n:

$$\rho_n \in [1 - e^{-C\frac{\log L}{\log\log L}}, 1 + L^{-1+o(1)}]\,. \tag{3.34}$$

Indeed, Eq. (3.33) implies that $\bar\xi_{\ell_*,h} = \xi_{\ell_*,h-1} - \xi_{\ell_*,h}$ satisfies

$$\bar\xi_{\ell_*,h} = \bar\xi_{\ell_0,h} + L^{-5/4+o(1)}\,,$$

and the fact that $\bar\xi_{\ell_0,h} = L^{-1+o(1)}$ shows that

$$\rho_n = \frac{\bar\xi_{\ell_*,h}}{\widehat\pi_\infty(\phi_o = -h)} = \Big(1 + L^{-1/4+o(1)}\Big)\frac{\bar\xi_{\ell_0,h}}{\widehat\pi_\infty(\phi_o = -h)}\,; \tag{3.35}$$

hence, Eq. (3.34) follows from Lemma 3.8 applied to $\bar\xi_{\ell_0,h}$. ∎

Proof of Theorem 3.2. We aim to apply Proposition 3.9 for $\ell_0 = L^{1/3+o(1)}$ to obtain Eq. (3.5) (mesoscopic range), and for $\ell_0 = L^{1/2+o(1)}$ to establish Eqs. (3.6) and (3.7) (macroscopic range). In order to relate the corresponding $\xi_{\ell_0,h}$ rates to ξ_n, unlike the proof of Proposition 3.1, it will be imperative to fulfill the divisibility condition of Proposition 3.9 and avoid the error term Ξ_3 there.

Let $\ell_1 \in [\![L^{1/3}, \sqrt{L}e^{-\frac12\frac{\log L}{\log\log L}}]\!]$ be an integer satisfying $\ell_1 \mid \ell_*$, and let $\ell_0 = \ell_1 - 10\lceil\log L\rceil$. Applying Proposition 3.9 for $\widehat\pi_\infty$ and $F = [\![1,\ell_*]\!]^2$ (noting that $\ell_0 L e^{-\sqrt{\log L}} \geq L^{4/3-o(1)}$, whereas $\ell_*^2 \asymp L$, qualifying an application of Eq. (3.19)), we find that

$$\widehat\pi_\infty(\phi_x \geq -h,\,\forall x\in F) = \exp\bigg(-\xi_{\ell_0,h}\ell_*^2 + O\Big(\frac{\ell_*^2}{\ell_0 L}e^{\sqrt{\log L}}\Big) + o(L^{-5})\bigg)\,,$$

[8]This application of Proposition 3.9 would be valid for any $\ell_* \leq L^{5/8-o(1)}$. More generally, if we set $\ell_0 = L^{1/2-o(1)}$, we could have extended the result up to $\ell_* \leq L^{3/4-o(1)}$.

and therefore

$$|\xi_{\ell_*,h} - \xi_{\ell_0,h}| = O\Big(\frac{1}{\ell_0 L} e^{\sqrt{\log L}}\Big)\,. \tag{3.36}$$

For Eq. (3.5), let $\ell_0 = \ell_1 - \lceil 10 \log L \rceil$ for

$$\ell_1 = 2^{\lceil \frac13 \log L + (4\log 2)\sqrt{\log L}\rceil} \asymp L^{1/3} e^{4\sqrt{\log L}}\,.$$

Our assumption that $|\partial F| \vee |\partial V| \leq O(L^{2/3} e^{\sqrt{\log L}})$ implies $|\partial F| \leq 10L$ and $|F| \leq O(L^{4/3} e^{2\sqrt{\log L}})$, whereas $\ell_0 L e^{-\sqrt{\log L}} \geq L^{4/3} e^{3\sqrt{\log L}}$. Hence, we may appeal to Eq. (3.19) of Proposition 3.9, so

$$\widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in F) = \exp\Big(- \xi_{\ell_0,h}|F| + O(\Xi) + o(L^{-5})\Big)\,,$$

with $\Xi = \Xi_1 + \Xi_2 + \Xi_3$ for

$$\begin{aligned}
\Xi_1 &= \frac{|\partial V|}{L} e^{C\frac{\log L}{\log\log L}} \leq L^{-1/3+o(1)}\,,\\
\Xi_2 &\leq \frac{|\partial F|^2}{\ell_0 L} e^{\sqrt{\log L}} \leq e^{-\sqrt{\log L}}\,,\\
\Xi_3 &= \frac{|\partial F|\ell_0}{L} e^{-c\frac{\log L}{\log\log L}} \leq e^{-(c-o(1))\frac{\log L}{\log\log L}}\,,
\end{aligned}$$

whence

$$\Xi = o(1)\,.$$

Appealing to Eq. (3.36), we may replace $\xi_{\ell_0,h}$ by $\xi_n = \xi_{\ell_*,h}$ and incur an error of at most

$$\frac{|F|}{\ell_0 L} e^{\sqrt{\log L}} \leq e^{-\sqrt{\log L}}\,,$$

thus establishing Eq. (3.5).

For Eqs. (3.6) and (3.7), let $\ell_0 = \ell_1 - \lceil 10 \log L\rceil$ for

$$\ell_1 = 2^{\lceil \frac12 \log L - (\frac34 \log 2)\frac{\log L}{\log\log L}\rceil} \asymp \sqrt{L} e^{-\frac34 \frac{\log L}{\log\log L}}\,.$$

By assumption, $|\partial F| \vee |\partial V| \leq O(L e^{\sqrt{\log L}})$, and Eqs. (3.20) and (3.21) yield

$$\begin{aligned}
\widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in F) &\geq \exp\Big(- \xi_{\ell_0,h}|F| + O(\Xi) + o(L^{-5})\Big)\,,\\
\widehat{\pi}^0_V(\phi_x \geq -h,\, \forall x \in F \mid \mathfrak{S}) &\leq \exp\Big(- \xi_{\ell_0,h}|F| + O(\Xi) + o(L^{-5})\Big)\,,
\end{aligned}$$

with $\Xi = \Xi_1 + \Xi_2 + \Xi_3$ for

$$\begin{aligned}
\Xi_1 &= \frac{|\partial V|}{L} e^{C\frac{\log L}{\log\log L}} \leq e^{(C+o(1))\frac{\log L}{\log\log L}}\,,\\
\Xi_2 &\leq \frac{|\partial F|^2}{\ell_0 L} e^{\sqrt{\log L}} \leq \sqrt{L} e^{(\frac34 - o(1))\frac{\log L}{\log\log L}}\,,\\
\Xi_3 &= \frac{|\partial F|\ell_0}{L} e^{-c\frac{\log L}{\log\log L}} \leq \sqrt{L} e^{-(c+\frac34-o(1))\frac{\log L}{\log\log L}}\,,
\end{aligned}$$

so that, for any sufficiently large L,

$$\Xi \leq \sqrt{L} e^{\frac{\log L}{\log\log L}}\,.$$

(The remark regarding the validity of the upper bound without conditioning on $\mathfrak{S}$ in the special case where $|F| \leq \big(\frac{3\beta}{\widehat{\pi}_\infty(\phi_o > h)}\big)^2$ follows from the statement below Eq. (3.21) in Proposition 3.9.) Revisiting Eq. (3.36), we replace $\xi_{\ell_0,h}$ by $\xi_n = \ell_{\ell_*,h}$ at a cost of

$$\frac{|F|}{\ell_0 L} e^{\sqrt{\log L}} \leq \sqrt{L} e^{(\frac34 - o(1))\frac{\log L}{\log\log L}}\,,$$

obtaining Eqs. (3.6) and (3.7) and completing the proof (NB. Had we chosen $\ell_1 \nmid \ell_*$, Eq. (3.36) would have instead resulted in an untenable cost of $(\ell_0/\ell_*)|F|L^{-1+o(1)} = L^{1-o(1)}$.) ∎

3.4. **Sharper upper bound on ρ_n.** We conclude this section with the proof of Eq. (3.4), featuring a sharper upper bound compared to the one in Eq. (3.34), which already established the sought lower bound. Although this sharper bound is not needed for the proof of Theorem 1.2, it is necessary for the remark following it, that the critical window from Theorem 1.2 excludes the natural prediction $\lambda_*\beta/\widehat{\pi}_\infty(\phi_o = h)$; see Remark 5.3 for more details.

We will need the following result (which may be of independent interest) establishing that the $\mathbb{Z}$GFF, conditional on $\phi\restriction_{B_r(o)} = h$, is rigid, beyond scale $h/\log h$, about the corresponding real-valued harmonic solution ϕ^*.

Proposition 3.10. *Fix $r \geq 0$. Let $h \geq 1$, $R = \lceil 20h/\log h\rceil$, and let ϕ^* be the $\mathbb{R}$-valued harmonic function in $A_{r,R} := B_R(o) \setminus B_r(o)$ with boundary conditions h on $\partial B_r(o)$ and 0 on $\partial B_R(o)$. There exist an absolute constant $C > 0$ such that, for every sufficiently large β and every $z \notin B_r(o)$,*

$$\widehat{\pi}_\infty\left(\phi_z \notin \left[\phi^*_z - C\frac{h}{\log h}, \phi^*_z + C\frac{h}{\log h}\right] \;\Big|\; \phi_x = h,\, \forall x \in B_r(o)\right) \leq e^{-\frac{1}{5}(\beta - C)\frac{h}{\log h}}\,.$$

Proof. From the explicit solution to the Dirichlet problem stated in the proposition in terms of hitting times of simple random walk S_t in $\mathbb{Z}^2$ (see, e.g., [28, §1.4]), one has that

$$\phi^*_x = h\mathbb{P}_x(\tau_{\partial B_r(o)} < \tau_{\partial B_R(o)}) \qquad \text{for all } x \in A_{r,R}\,, \tag{3.37}$$

where $\mathbb{P}_x$ denote the law of the random walk S_t started at x and τ_A is the first time it hits a set A. It is well-known that the potential kernel $a(x)$ satisfies that $a(S_t)$ is a martingale, and from known estimates on $a(x)$ (see, e.g., [28, §1.6], and in particular Prop. 1.6.7 and Example 1.6.8 there),

$$\mathbb{P}_x(\tau_{\partial B_r(o)} > \tau_{\partial B_R(o)}) = \frac{\log|x| - \log r + O(1/r)}{\log R - \log r}\,, \tag{3.38}$$

where $|x|$ denotes the Euclidean distance of x from the origin. We will use the following fact on ϕ^*, which, just like the last display, follows from the Optional Stopping Theorem and known estimates for $a(x)$ (e.g., those in [37, p. 125–127], combined with a Taylor approximation to account for the lattice error $|S_{\tau_{B_r(o)}}| \in [r, r+1]$, whence $\log|S_{\tau_{B_r(o)}}| \in [\log r, \log r + \frac{1}{r}]$, and similarly for $B_R(o)$):

$$\phi^*_x \leq \frac{1}{64} \qquad \text{for all } x \in A_{r,R} \text{ adjacent to some } y \in \partial B_R(o)\,. \tag{3.39}$$

Several steps will be needed to deal with boundary effects and different large deviation scenarios.

Step 1. (*Subtracting the real-valued solution ϕ^*.*) Let

$$\sigma := \phi - \phi^*\,.$$

Denoting the discrete Laplacian by $(\Delta\phi)_x = \frac{1}{4}\sum_{y\sim x}(\nabla\phi)_{xy}$ for $(\nabla\phi)_{xy} = \phi_y - \phi_x$, we see that

$$\sum_{x\sim y}|(\nabla\phi)_{xy}|^2 = \sum_{x\sim y}|(\nabla\sigma)_{xy}|^2 + \sum_{x\sim y}|(\nabla\phi^*)_{xy}|^2 - 8\sum_x \sigma_x(\Delta\phi^*)_x\,.$$

Since $\sum_{x\sim y}|(\nabla\phi^*)_{xy}|^2$ is but a constant, $(\Delta\phi^*)_x = 0$ for all $x \notin \partial_{\mathrm{v}} A_{r,R}$ (the external vertex boundary of $A_{r,R}$), and $\sigma_x = 0$ for $x \in B_r(o)$, we see that the law of σ under $\widehat{\pi}_\infty$ can be written as

$$\widehat{\mu}_\infty(\sigma) = \frac{1}{Z_{\widehat{\mu}}}\exp\bigg(-\beta\sum_{x\sim y}|(\nabla\sigma)_{xy}|^2 + 8\beta\sum_{x\in\partial_{\mathrm{v}} B_R(o)}(\Delta\phi^*)_x\sigma_x\bigg)\,, \tag{3.40}$$

where $Z_{\widehat{\mu}}$ sums the exponent on the right-hand over all configurations $\sigma : \mathbb{Z}^2 \to \mathbb{R}$ with
(i) $\sigma_x = 0$ for all $x \in B_r(o)$;
(ii) $\sigma_x \in -\phi^*_x + \mathbb{Z}$ for all x.

Using Eq. (3.39) (as well as that $\phi^* \geq 0$ and $\phi^*\restriction_{B_R(o)^c} = 0$) we further have

$$0 \leq (\Delta\phi^*)_x \leq \frac{1}{64} \quad \text{for all } x \in \partial_{\mathsf{v}} B_R(o)\,, \tag{3.41}$$

and see that the distribution $\widehat{\mu}_\infty$ is nothing but a $\mathbb{Z}$GFF distribution with a small positive external field of $\Delta\phi^*$ on $\partial_{\mathsf{v}} B_R(o)$ (and values that are the integers translated by the fractional parts of $-\phi^*$).

Step 2. (*Ruling out many* `steep` *level lines in* $\lfloor\sigma\rfloor$.) The fact that σ is $\mathbb{R}$-valued hinders the Peierls approach for establishing rigidity: for instance, if one were to look at $\lfloor\sigma\rfloor$ (point-wise floor of σ), erasing a level-line loop via an appropriate vertical shift of its interior may actually increase the energetic cost if (consider, for instance, the situation when the gradients along it have fractional parts $0.1, 0.9$). A similar phenomenon occurs when rounding σ to the nearest integers (e.g., consider fractional parts $0.4, 0.6$ in that setting). To deal with this, we introduce the following notion of a `steep` level line.

Definition 3.11. We say $\mathscr{A}$ is a j level-line *annulus* in $\lfloor\sigma\rfloor$, for some $j \in \mathbb{Z}$, if it consists of two j level-line loops $\mathscr{A}^{\mathsf{in}}, \mathscr{A}^{\mathsf{out}}$ in $\lfloor\sigma\rfloor$ as per Definition 2.1 with reversed types—$\mathscr{A}^{\mathsf{out}}$ is an up-loop and $\mathscr{A}^{\mathsf{in}}$ is a down-loop (whence $\mathscr{A}$ is an up-annulus) or vice versa (a down-annulus)—so that either
(a) (trivial annulus, one external loop) $\mathscr{A}^{\mathsf{in}} = \emptyset$ and $B_r(o) \cap \mathrm{Int}(\mathscr{A}^{\mathsf{out}}) = \emptyset$; or
(b) (nontrivial annulus) $B_r(o) \subset \mathrm{Int}(\mathscr{A}^{\mathsf{in}}) \subset \mathrm{Int}(\mathscr{A}^{\mathsf{out}})$.
Write $|\mathscr{A}| := |\mathscr{A}^{\mathsf{in}}| + |\mathscr{A}^{\mathsf{out}}|$ and $\mathrm{Int}(\mathscr{A}) := \mathrm{Int}(\mathscr{A}^{\mathsf{out}}) \setminus \mathrm{Int}(\mathscr{A}^{\mathsf{in}})$.

Note that for any level line $\mathfrak{L}$ of $\lfloor\sigma\rfloor$, either $B_r(o) \subset \mathrm{Int}(\mathfrak{L})$ or $B_r(o) \subset \mathrm{Int}(\mathfrak{L})^c$ since $\sigma\restriction_{B_r(o)} = 0$. The above definition requires that if $B_r(o) \subset \mathrm{Int}(\mathscr{A}^{\mathsf{out}})$ then there must be an internal loop, of a reversed type, separating $B_r(o)$ from $\mathscr{A}^{\mathsf{out}}$. Further note that, since $\mathscr{A}^{\mathsf{in}}, \mathscr{A}^{\mathsf{out}}$ are two j level-line loops that are nested ($\mathrm{Int}(\mathscr{A}^{\mathsf{in}}) \subset \mathrm{Int}(\mathscr{A}^{\mathsf{out}})$) and of reversed types, then they must be disjoint.

Observe that if $\lfloor\sigma_z\rfloor = k > 0$, then there must exist level line up-annuli $\mathscr{A}_1, \ldots, \mathscr{A}_k$ in $\lfloor\sigma\rfloor$ corresponding to heights $k, k-1, \ldots, 1$ such that $z \in \mathrm{Int}(\mathscr{A}_1) \subset \ldots \subset \mathrm{Int}(\mathscr{A}_k)$. (This would be easy to see if in lieu of $\widehat{\pi}_\infty$ we considered $\widehat{\pi}^0_{B_M(o)}$ for $M \gg R$ (and the same for $\widehat{\mu}_\infty$), and one can couple the two measures to agree on $B_{M/2}(o)$ up to a probability $e^{-c\beta M}$, vanishing as $M \to \infty$.) Similarly, if $\lfloor\sigma_z\rfloor = -k$ for $k > 0$ we can find such down-annuli $\mathscr{A}_1, \ldots, \mathscr{A}_k$ at heights $-k+1, -k+2, \ldots, 0$.

Definition 3.12. We say a j level-line loop $\mathfrak{L}$ in $\lfloor\sigma\rfloor$ is `steep` if at least $\frac{2}{3}|\mathfrak{L}|$ of its bonds b satisfy $|(\nabla\lfloor\sigma\rfloor)_{xy}| \geq 2$, where xy is the edge of $\mathbb{Z}^2$ dual to b. Similarly, we say that a j level-line annulus $\mathscr{A}$ in $\lfloor\sigma\rfloor$ is `steep` if $\frac{2}{3}|\mathscr{A}|$ of the bonds b in $\mathscr{A}^{\mathsf{in}} \cup \mathscr{A}^{\mathsf{out}}$ satisfy $|(\nabla\lfloor\sigma\rfloor)_{xy}| \geq 2$ for the xy dual to b.

The following claim will control the probability of the occurrence of k `steep` level lines annuli.

Claim 3.13. *Let $\mathscr{S}^{\uparrow}_{z,k}$ be the event that there is a sequence $\{\mathscr{A}_i\}_{i=1}^k$ of* `steep` *j_i level line up-annuli in $\lfloor\sigma\rfloor$ for some $j_1 > \ldots > j_k > 0$, such that $z \in \mathrm{Int}(\mathscr{A}_1) \subset \mathrm{Int}(\mathscr{A}_2) \subset \ldots \subset \mathrm{Int}(\mathscr{A}_k)$. There exists an absolute constant $C > 0$ such that, for every large enough $\beta > 0$ and every $k \geq 1$,*

$$\widehat{\mu}_\infty(\mathscr{S}^{\uparrow}_{z,k}) \leq e^{-\frac{1}{5}(\beta - C)k}\,. \tag{3.42}$$

The same holds for the event $\mathscr{S}^{\downarrow}_{z,k}$ corresponding to j_i level line down-annuli for $j_1 < \ldots < j_k < 0$.

Proof. Fix a sequence $\{\mathscr{A}_i\}_{i=1}^k$ of level-line annuli with $z \in \mathrm{Int}(\mathscr{A}_1) \subset \ldots \subset \mathrm{Int}(\mathscr{A}_k)$, and let $\mathscr{S}^{\uparrow}_{z,\{\mathscr{A}_i\}}$ denote the event that these occur as `steep` j_i level line up-annuli in $\lfloor\sigma\rfloor$ for some $j_1 > \ldots > j_k > 0$. The sought Eq. (3.42) will follow from showing that

$$\widehat{\mu}_\infty(\mathscr{S}^{\uparrow}_{z,\{\mathscr{A}_i\}}) \leq e^{-\frac{1}{5}\beta\sum_{i=1}^k |\mathscr{A}_i|}\,. \tag{3.43}$$

This is a consequence of a routine enumeration over the $\mathscr{A}_i$'s: for an upper bound on the number of such sequences, we may ignore the consistency constraints between the $2k$ loops $\{\mathscr{A}_i^{\mathtt{in}}, \mathscr{A}_i^{\mathtt{out}}\}_{i=1}^k$; every external loop $\mathscr{A}_i^{\mathsf{out}}$ must cross the horizontal line through y at distance at most $|\mathscr{A}_i^{\mathsf{out}}|/2$

from y, and every internal loop $\mathscr{A}_i^{\mathtt{in}}$ must cross the horizontal line through o at distance at most $|\mathscr{A}_i^{\mathtt{in}}|/2$ from o. We see that the enumeration cost is at most $\exp(C\sum_i |\mathscr{A}_i|)$ for some absolute constant $C>0$, and hence Eq. (3.42) will follow from Eq. (3.43) by a union bound.

To obtain Eq. (3.43), consider the Peierls map T that subtracts the height of x by the number of level-line annuli that contain x in their interior:

$$(T\sigma)_x = \sigma_x - \sum_{i=1}^k \mathbb{1}_{\{x\in \mathrm{Int}(\mathscr{A}_i)\}}\,. \tag{3.44}$$

Observe that, by definition, T does not modify σ_x for $x\in B_r(o)$, and elsewhere, it only modifies σ via an integer shift (specifically, a shift by 1), and therefore $T\sigma$ remains a legal configuration; clearly, as we have fixed $\{\mathscr{A}_i\}_{i=1}^k$ ahead of time, it is an injection, and thus Eq. (3.43) will follow once we establish that, for every $\sigma\in\mathscr{S}^{\uparrow}_{z,\{\mathscr{A}_i\}}$,

$$\widehat{\mu}_\infty(\sigma)\le e^{-\frac14\beta\sum_{i=1}^k|\mathscr{A}_i|}\widehat{\mu}_\infty(T\sigma)\,. \tag{3.45}$$

In view of the expression for $\widehat{\mu}$ in Eq. (3.40),

$$\frac{\widehat{\mu}_\infty(\sigma)}{\widehat{\mu}_\infty(T\sigma)} = e^{-\beta(\Upsilon_1-\Upsilon_2)}\,,$$

where

$$\Upsilon_1 = \sum_{xy\text{ dual to } b\in\bigcup\mathscr{A}_i} |(\nabla\sigma)_{xy}|^2 - |(\nabla(T\sigma))_{xy}|^2 \qquad,\qquad \Upsilon_2 = 8\sum_{x\in\partial_{\mathsf{v}}B_R(o)}(\Delta\phi^*)_x(\sigma_x-(T\sigma)_x)\,. \tag{3.46}$$

The handling of the gradient term Υ_1 shows the rationale behind the definition of a `steep` level line. Within this argument, consider xy dual to some $b\in\bigcup_i\mathscr{A}_i$. W.l.o.g., $(\nabla\lfloor\sigma\rceil)_{xy}\ge 1$ (otherwise reverse the roles of x,y; the gradient is nonzero as xy is dual to a bond in one of the $\mathscr{A}_i$'s). Since

$$(\nabla\sigma)_{xy} > (\nabla\lfloor\sigma\rceil)_{xy}-1\,,$$

we see that every bond $b\in\bigcup_i\mathscr{A}_i$ dual to xy with $(\nabla\lfloor\sigma\rceil)_{xy}\ge 2$ contributes at least 1 to Υ_1 for each $\mathscr{A}_i$ it belongs to (and more than that if it belongs to multiple $\mathscr{A}_i$'s, due to the quadratic effect). On the other hand, if $b\in\bigcup_i\mathscr{A}_i$ is dual to xy with $(\nabla\lfloor\sigma\rceil)_{xy}=1$ then it decreases Υ_1 by at most 1 (and by definition belongs to just a single $\mathscr{A}_i$). Thus, the fact that every $\mathscr{A}_i$ is `steep` implies that

$$\Upsilon_1 \ge \frac13\sum_i|\mathscr{A}_i|\,.$$

To handle the external field term Υ_2, we use Eq. (3.41) to write

$$\Upsilon_2\le\frac18\sum_i|\,\mathrm{Int}(\mathscr{A}_i)\cap\partial_{\mathsf{v}}B_R(o)|\,,$$

and proceed to make the elementary yet important observation that, for every i,

$$|\mathscr{A}_i|\ge|\,\mathrm{Int}(\mathscr{A}_i)\cap\partial_{\mathsf{v}}B_R(o)|\,.$$

(One way to see this would send disjoint paths P_x from every $x\in\partial B_R(o)$ to ∞—e.g., assigning the 4 sectors $(\frac{\pi}{4},\frac{3\pi}{4})$, $(\frac{3\pi}{4},\frac{5\pi}{4})$, $(\frac{5\pi}{4},\frac{7\pi}{4})$, $(\frac{7\pi}{4},\frac{\pi}{4})$ paths going north, west, south, east, respectively. This would map each $x\in\mathrm{Int}(\mathscr{A}_i)\cap\partial_{\mathsf{v}}B_R(o)$ to a distinct bond of $\mathscr{A}_i$, dual to some $e\in P_x$.)

Combining the last three displays shows that

$$\Upsilon_1-\Upsilon_2\ge\frac15\sum_i|\mathscr{A}_i|\,,$$

establishing Eq. (3.45) and thus completing the proof of Eq. (3.42). The analogous bound for $\mathscr{S}^{\downarrow}_{z,k}$ follows via the Peierls map $(T\sigma)_x=\sigma_x+\sum_{i=1}^k\mathbb{1}_{\{x\in\mathrm{Int}(\mathscr{A}_i)\}}$, where the exact same argument holds for analyzing the gradient change Υ_1 (and one can use the simple bound $\Upsilon_2 \leq 0$, as the external field makes it only more preferable to increase the heights of the configuration). ∎

Step 3. (*Bounding the energy of non-`steep` level lines.*) To treat a large deviation stemming from many non-`steep` level lines, we will show establish the following lower bound on the number of gradient-1 bonds in a collection of non-`steep` level lines surrounding a site y.

Claim 3.14. *Every sequence $\mathscr{A}_1, \ldots, \mathscr{A}_k$ of level-line annuli with $z \in \operatorname{Int}(\mathscr{A}_1) \subset \ldots \subset \operatorname{Int}(\mathscr{A}_k)$ that are non-`steep` in $\lfloor\sigma\rfloor$ contains at least $\frac{2}{9}k^2$ bonds b whose dual xy satisfies $|(\nabla\lfloor\sigma\rfloor)_{xy}| = 1$.*

Proof. We will prove by induction that, $|\operatorname{Int}(\mathscr{A}_j)| \geq j^2/9$ for all $j \geq 1$. The base case is trivial (one always has $|\operatorname{Int}(\mathscr{A}_j)| \geq 1$ since the loops $\mathscr{A}_j^{\texttt{in}}, \mathscr{A}_j^{\texttt{out}}$ are disjoint). For the inductive step, recall Definition 3.11, and observe that the fact that $\operatorname{Int}(\mathscr{A}_j) \subset \operatorname{Int}(\mathscr{A}_{j+1})$ implies that

$$\operatorname{Int}(\mathscr{A}_j^{\texttt{out}}) \subset \operatorname{Int}(\mathscr{A}_{j+1}^{\texttt{out}}) \qquad \text{and} \qquad \operatorname{Int}(\mathscr{A}_{j+1}^{\texttt{in}}) \subset \operatorname{Int}(\mathscr{A}_j^{\texttt{in}}).$$

By the isoperimetric inequality in $\mathbb{Z}^2$ and the induction hypothesis, $|\mathscr{A}_j| \geq 4\sqrt{|\operatorname{Int}(\mathscr{A}_j)|} \geq 4j/3$, where we recall that $|\mathscr{A}_j| = |\mathscr{A}_j^{\texttt{in}}| + |\mathscr{A}_j^{\texttt{out}}|$. Recalling Definition 3.12, the fact that $\mathscr{A}_j$ is non-`steep` (and that it is a level line of $\lfloor\sigma\rfloor$, whence each of its bonds b has a dual xy with $(\nabla\lfloor\sigma\rfloor)_{xy} \neq 0$) implies that it has at least $|\mathscr{A}_j|/3 \geq 4j/9$ bonds b whose dual xy has $(\nabla\lfloor\sigma\rfloor)_{xy} = 1$. Each such bond $b \in \mathscr{A}_j^{\texttt{out}}$ corresponds to a distinct unit square in $\operatorname{Int}(\mathscr{A}_{j+1}^{\texttt{out}}) \setminus \operatorname{Int}(\mathscr{A}_j^{\texttt{out}})$, whereas each such $b \in \mathscr{A}_j^{\texttt{in}}$ corresponds to a distinct unit square in $\operatorname{Int}(\mathscr{A}_j^{\texttt{in}}) \setminus \operatorname{Int}(\mathscr{A}_{j+1}^{\texttt{in}})$. Hence,

$$|\operatorname{Int}(\mathscr{A}_{j+1})| = \left|\operatorname{Int}(\mathscr{A}_{j+1}^{\texttt{out}}) \setminus \operatorname{Int}(\mathscr{A}_{j+1}^{\texttt{in}})\right| \geq |\operatorname{Int}(\mathscr{A}_j)| + \frac{4j}{9} \geq \frac{j(j+4)}{9} \geq \frac{(j+1)^2}{9},$$

completing the proof of the induction. At the same time, since the gradient-1 bonds counted above in the collection $\{\mathscr{A}_j\}_{j=1}^k$ are all distinct by definition, we find that this collection of loops contains at least $\frac{4}{9}\sum_{j=1}^k j \geq \frac{2}{9}k^2$ such bonds, as required. ∎

Step 4. (*Ruling out many non-`steep` level lines in $\lfloor\sigma\rfloor$.*) Let $\widehat{\nu}_\infty$ be the $\mathbb{Z}$GFF on configurations $\bar\sigma : \mathbb{Z}^2 \to \mathbb{Z}$ satisfying $\bar\sigma_x = 0$ for all $x \in B_r(o)$, with the external field from Eq. (3.40), and a modified interaction of $\beta/2$ (as opposed to β) for edges within $B_R(o)$ or intersecting it:

$$\widehat{\nu}_\infty(\bar\sigma) = \frac{1}{Z_{\widehat{\nu}}} \exp\Bigg(-\frac{\beta}{2} \sum_{\substack{x\sim y\\ \{x,y\}\cap B_R(o)\neq\emptyset}} |(\nabla\bar\sigma)_{xy}|^2 - \beta \sum_{\substack{x\sim y\\ \{x,y\}\cap B_R(o)=\emptyset}} |(\nabla\bar\sigma)_{xy}|^2 + 8\beta \sum_{x\in\partial_{\mathsf{v}} B_R(o)} (\Delta\phi^*)_x \bar\sigma_x \Bigg). \tag{3.47}$$

The analysis of the Peierls map in T from the proof of Claim 3.13 will readily show that $\widehat{\nu}_\infty$ is rigid:

Claim 3.15. *The following holds for β large enough. Let $\{\mathscr{A}_i\}_{i=1}^k$ be a sequence of level-line annuli with $z \in \operatorname{Int}(\mathscr{A}_1) \subset \ldots \subset \operatorname{Int}(\mathscr{A}_k)$, and let $\mathscr{S}^{\uparrow}_{z,\{\mathscr{A}_i\}}$ denote the event that these occur as j_i level line up-annuli for some $j_1 > \ldots > j_k > 0$. There exists an absolute constant $C > 0$ such that*

$$\widehat{\nu}_\infty(\mathscr{S}^{\uparrow}_{z,\{\mathscr{A}_i\}}) \leq e^{-\frac{1}{4}\beta\sum_1^k |\mathscr{A}_i|}.$$

The same holds for $\mathscr{S}^{\downarrow}_{z,\{\mathscr{A}_i\}}$ corresponding to down-annuli for some $j_1 < \ldots < j_k < 0$.

Proof. Let T be the map from Eq. (3.44), decrementing the height of each of the regions $\operatorname{Int}(\mathscr{A}_i)$, and define Υ_1, Υ_2 as in Eq. (3.46). Unlike the situation in the proof of Claim 3.13, where the $\mathbb{R}$-valued nature of the configuration $\sigma \sim \widehat{\mu}_\infty$ called for the notion of `steep` level lines to handle the energy gain in the gradients of $T\sigma$ compared σ, here $\bar\sigma \sim \widehat{\nu}_\infty$ is $\mathbb{Z}$-valued; thus, there is no loss in the gradient term Υ_1, and

$$\Upsilon_1 \geq \sum_i |\mathscr{A}_i|.$$

The analysis of the external field term Υ_2 holds verbatim as in said proof, giving

$$\Upsilon_2 \le \frac{1}{8}\sum_i |\mathscr{A}_i|\,.$$

Combining these gives the sought bound on $\mathscr{S}^{\uparrow}_{z,\{\mathscr{A}_i\}}$, as here we have

$$\frac{\widehat{\nu}_\infty(\bar\sigma)}{\widehat{\nu}_\infty(T\bar\sigma)} \le e^{-\beta(\frac12\Upsilon_1-\Upsilon_2)}$$

due to the modified interaction of $\beta/2$ within $B_R(o)$. As in the proof of Claim 3.13, the same applies to $\mathscr{S}^{\downarrow}_{z,\{\mathscr{A}_i\}}$ via the counterpart of T that increments the height of each region $\mathrm{Int}(\mathscr{A}_i)$. ■

Let $\overline{\mathscr{S}}^{\uparrow}_{z,k}$ be the event that there is a sequence $\{\mathscr{A}_i\}_{i=1}^k$ of non-`steep` level line up-annuli with $z\in\mathrm{Int}(\mathscr{A}_1)\subset\ldots\subset\mathrm{Int}(\mathscr{A}_k)$ (the analogue of $\mathscr{S}^{\uparrow}_{z,k}$ from Claim 3.13 for non-`steep` annuli). By combining Claims 3.14 and 3.15 with the usual enumeration over the annuli $\{\mathscr{A}_i\}_{i=1}^k$, we find that

$$\widehat{\nu}_\infty(\overline{\mathscr{S}}^{\uparrow}_{z,k}) \le e^{-\frac{1}{18}(\beta-C)k^2}\,. \tag{3.48}$$

The same applies for the analogously defined $\overline{\mathscr{S}}^{\downarrow}_{z,k}$ pertaining down-annuli.

The next standard result (see, e.g., [29, Claim 3.5], an analogue when there is no external field on $\partial_{\mathsf{v}}B_R(o)$ as we encounter here) controls the Radon–Nikodym derivative between $\widehat{\mu}_\infty$ and $\widehat{\nu}_\infty$.

Claim 3.16. *There exists an absolute constant $c_0>0$ such that*

$$\sup_\sigma \frac{\widehat{\mu}_\infty(\sigma)}{\widehat{\nu}_\infty(\lfloor\sigma\rceil)} \le e^{c_0\beta R^2}\,. \tag{3.49}$$

In particular, $\widehat{\mu}_\infty(\overline{\mathscr{S}}^{\uparrow}_{z,k}) \le e^{-\frac{1}{25}(\beta-C)k^2}$ for $k\ge 10\sqrt{c_0}R$, and similarly for the analogous $\overline{\mathscr{S}}^{\downarrow}_{z,k}$.

Proof. Let $\phi^*_{\mathsf{f}} := \phi^*-\lfloor\phi^*\rceil$ denote the fractional part of ϕ^*, and recall that by definition $\sigma=\lfloor\sigma\rceil-\phi^*_{\mathsf{f}}$. For every $x\sim y$ we have

$$|(\nabla\sigma)_{xy}|^2 = |(\nabla\lfloor\sigma\rceil)_{xy}|^2 + |(\nabla\phi^*_{\mathsf{f}})_{xy}|^2 - 2(\nabla\lfloor\sigma\rceil)_{xy}(\nabla\phi^*_{\mathsf{f}})_{xy}\,.$$

As ϕ^*_{f} is fixed, we may rewrite $\widehat{\mu}_\infty$ from Eq. (3.40) via another partition function $\tilde Z_{\widehat{\mu}}$ so that

$$\widehat{\mu}_\infty(\sigma) = \frac{1}{\tilde Z_{\widehat{\mu}}}\exp\Bigg(-\beta\sum_{x\sim y}|(\nabla\lfloor\sigma\rceil)_{xy}|^2 + 2\beta\sum_{\substack{x\sim y\\ \{x,y\}\cap B_R(o)\neq\emptyset}}(\nabla\lfloor\sigma\rceil)_{xy}(\nabla\phi^*_{\mathsf{f}})_{xy} + 8\beta\sum_{x\in\partial_{\mathsf{v}}B_R(o)}(\Delta\phi^*)_x\lfloor\sigma\rceil_x\Bigg)\,,$$

where we were allowed to sum the cross-term $(\nabla\lfloor\sigma\rceil)_{xy}(\nabla\phi^*_{\mathsf{f}})_{xy}$ only on x,y with at least one of the sites falling within $B_R(o)$ as otherwise $\phi^*_x=\phi^*_y=0$ (and so $(\nabla\phi^*_{\mathsf{f}})_{xy}=0$) by construction. Comparing this with Eq. (3.47) for $\bar\sigma:=\lfloor\sigma\rceil$, we get

$$\begin{aligned}\frac{\widehat{\mu}_\infty(\sigma)}{\widehat{\nu}_\infty(\lfloor\sigma\rceil)} &= \frac{Z_{\widehat{\nu}}}{\tilde Z_{\widehat{\mu}}}\exp\Bigg(-\frac{\beta}{2}\sum_{\substack{x\sim y\\ \{x,y\}\cap B_R(o)\neq\emptyset}}\Big(|(\nabla\lfloor\sigma\rceil)_{xy}|^2 - 4(\nabla\lfloor\sigma\rceil)_{xy}(\nabla\phi^*_{\mathsf{f}})_{xy}\Big)\Bigg)\\ &\le \frac{Z_{\widehat{\nu}}}{\tilde Z_{\widehat{\mu}}}\exp\Bigg(2\beta\sum_{\substack{x\sim y\\ \{x,y\}\cap B_R(o)\neq\emptyset}}|(\nabla\phi^*_{\mathsf{f}})_{xy}|^2\Bigg)\\ &\le \frac{Z_{\widehat{\nu}}}{\tilde Z_{\widehat{\mu}}}\exp\left(c\beta R^2\right)\end{aligned}$$

for an absolute constant $c>0$, using that $4ab\le a^2+4b^2$ for any real a,b for the inequality in the second line, and that $|\nabla\phi^*_{\mathsf{f}}|\in[0,1)$ for the inequality in the third line.

To bound $Z_{\widehat{\nu}}/\tilde{Z}_{\widehat{\mu}}$ from below, we infer from the identity in the first line above that, for $\bar{\sigma} \sim \widehat{\nu}_\infty$ and $\mathbb{E}_{\widehat{\nu}}$ denoting expectation w.r.t. $\widehat{\nu}_\infty$,

$$\begin{aligned}\frac{\tilde{Z}_{\widehat{\mu}}}{Z_{\widehat{\nu}}} &= \mathbb{E}_{\widehat{\nu}}\bigg[\exp\bigg(-\frac{\beta}{2}\sum_{\substack{x\sim y\\ \{x,y\}\cap B_R(o)\neq\emptyset}}\Big(|(\nabla\bar{\sigma})_{xy}|^2 - 4(\nabla\bar{\sigma})_{xy}(\nabla\phi^*_{\mathsf{f}})_{xy}\Big)\bigg)\bigg]\\ &\geq \exp\bigg(-\frac{\beta}{2}\sum_{\substack{x\sim y\\ \{x,y\}\cap B_R(o)\neq\emptyset}}\Big(\mathbb{E}_{\widehat{\nu}}\left[|(\nabla\bar{\sigma})_{xy}|^2\right] - 4(\nabla\phi^*_{\mathsf{f}})_{xy}\mathbb{E}_{\widehat{\nu}}\left[(\nabla\bar{\sigma})_{xy}\right]\Big)\bigg),\end{aligned}$$

using here Jensen's inequality. The proof will be concluded once we show that for every $x \sim y$ we have $\mathbb{E}_{\widehat{\nu}}[|(\nabla\bar{\sigma})_{xy}|^2] \leq C$ for $C > 0$ (thus also $\mathbb{E}_{\widehat{\nu}}[|(\nabla\bar{\sigma})_{xy}|] \leq \sqrt{C}$) of the form $C = \varepsilon_\beta$. This is a consequence of Claim 3.15: if $\bar{\sigma}_y = \bar{\sigma}_x + k$ for some $k > 0$ then we there must be k level-line annuli going through the bond b dual to xy—either up-annuli with y in their interior, or down-annuli with x in their interior (or a combination of these). Enumerating over their types and geometry is a factor of at most $\exp(C\sum_{i=1}^k |\mathscr{A}_i|)$, and hence the probability of this event, in light of Claim 3.15, is at most $\exp(-\frac{1}{4}(\beta - C)k)$. This yields that $\mathbb{E}[\exp(\frac{1}{5}(|\nabla\bar{\sigma})_{xy}] < \varepsilon_\beta$, thus establishing Eq. (3.49). The proof of the final assertion of the claim follows from combining Eqs. (3.48) and (3.49). ∎

We now conclude the proof of the proposition by combining the result on `steep` annuli from Step 2 with the one on non-`steep` annuli from Step 4. Recall that our aim is to rule out $\{|\sigma_z| > C\frac{h}{\log h}\}$ for $\sigma \sim \widehat{\mu}_\infty$. Let $C = 250\sqrt{c_0}$ for the absolute constant $c_0 > 0$ from Claim 3.16, and $k := \lceil (C/2)\frac{h}{\log h}\rceil$. Recalling $R = \lceil 20h/\log h\rceil$, for this choice we have $k > 10\sqrt{c_0}R$.

By definition, $\{\sigma_z > C\frac{h}{\log h}\}$ implies $\mathscr{S}^{\uparrow}_{z,k} \cup \overline{\mathscr{S}}^{\uparrow}_{z,k}$. To treat $\mathscr{S}^{\uparrow}_{z,k}$, we appeal to Claim 3.13 (the dominant term in the final estimate) and to treat $\overline{\mathscr{S}}^{\uparrow}_{z,k}$ we appeal to Claim 3.16. The same holds for $\{\sigma_z < -C\frac{h}{\log h}\}$ via the analogous argument for $\mathscr{S}^{\downarrow}_{z,k} \cup \overline{\mathscr{S}}^{\downarrow}_{z,k}$. This completes the proof of the proposition. ∎

We will be interested in bounding the probability that, given that ϕ exceeds height h at the origin and its neighbor, we have $\phi_y \geq h$ at some other site y. The proposition above will provide such a bound when y is such that $\phi^*_y < h - C\frac{h}{\log h}$. The next claim will explain this is the case for all y with $|y| > r_1$ for some absolute constant $r_1 > 0$, and then extend that bound also to $|y| \leq r_1$.

Claim 3.17. *Let o be the origin and o' be a neighbor of o. There exists an absolute constant $c > 0$ such that, for large enough β, every $h \geq 1$ and every $y \neq o, o'$,*

$$\widehat{\pi}_\infty(\phi_y \geq h \mid \phi_o \geq h,\, \phi_{o'} \geq h) \leq e^{-c\beta\frac{h}{\log h}}\,.$$

Proof. Put $o' = (-1, 0)$. Let $C > 0$ be the absolute constant from Proposition 3.10 and ϕ^* be the harmonic function defined there. We first tweak the conditioning to the one in said proposition:

$$\begin{aligned}\widehat{\pi}_\infty(\phi_y \geq h \mid \phi_o \geq h,\, \phi_{o'} \geq h) &\leq \widehat{\pi}_\infty(\phi_y \geq h \mid \phi_x \geq h, \forall x \in B_1(o))\\ &\leq (1 + \varepsilon_\beta)\widehat{\pi}_\infty(\phi_y \geq h - 1 \mid \phi_x = h, \forall x \in B_1(o))\,,\end{aligned}$$

where the first transition is by FKG, and the second one is by a Peierls argument via the map that decreases ϕ_z by 1 for each z in a connected component of $\{v : \phi_v > h\}$ that intersects some $z \in B_1(o)$ (this may also decrease ϕ_y by 1, hence the weaker inequality $\phi_y \geq h-1$). The asymptotic behavior of ϕ^*_x in terms of $|x|$ given by Eqs. (3.37) and (3.38) shows that there some absolute constant $r_1 > 0$ (behaving as e^C for the above $C > 0$) such that $\phi^*_y < h - (C+1)\frac{h}{\log h}$ for every y with $|y| \geq r_1$. Thus, applying Proposition 3.10 to the last display shows that

$$\sup_{y:\,|y|\geq r_1} \widehat{\pi}_\infty\Big(\phi_y \geq h - \frac{h}{\log h}\,\Big|\, \phi_o \geq h,\, \phi_{o'} \geq h\Big) \leq \exp\Big(-\frac{1}{5}(\beta - C)\frac{h}{\log h}\Big)\,. \tag{3.50}$$

Next, fix $y \neq o, o'$ with $|y| < r_1$. We aim to show that

$$\widehat{\pi}^0_\infty(\phi_y \geq h \mid \phi_o, \phi_{o'} \geq h) \leq \exp\Big(-\frac{1}{10\lceil r_1 \rceil}(\beta - C)\frac{h}{\log h}\Big)\,. \tag{3.51}$$

In the special case where $y \sim o$ (or $y \sim o'$, symmetrically) this is simple, and will demonstrate the basic principle of the argument. Recalling $o' = (-1, 0)$, suppose $y = (1, 0)$. Letting $y_k = (k, 0)$ for $k \in \mathbb{Z}$ (so that o', o, y are y_{-1}, y_0, y_1, respectively), and setting $K = \lceil r_1 \rceil$, one has

$$\begin{aligned}\widehat{\pi}_\infty(\phi_{y_K} \geq h \mid \phi_o \geq h,\, \phi_{o'} \geq h) &\geq \widehat{\pi}_\infty\Big(\bigcap_{k=1}^K \phi_{y_k} \geq h \;\Big|\; \phi_o \geq h,\, \phi_{o'} \geq h\Big) \\ &\geq \prod_{k=1}^K \widehat{\pi}_\infty\left(\phi_{y_k} \geq h \mid \phi_{y_{k-1}} \geq h,\, \phi_{y_{k-2}} \geq h\right) \\ &= \widehat{\pi}_\infty\left(\phi_y \geq h \mid \phi_o \geq h,\, \phi_{o'} \geq h\right)^K\,, \end{aligned} \tag{3.52}$$

where the second inequality is by FKG (replacing the conditioning on $\bigcap_{i=-1}^{k-1}\{\phi_{y_i} \geq h\}$ by the one on $\{\phi_{y_{k-1}} \geq h\} \cap \{\phi_{y_{k-2}} \geq h\}$). As $|y_K| \geq r_1$, it is covered by Eq. (3.50), whose upper bound thus applies to the left-hand of Eq. (3.52), thereby yielding Eq. (3.51) (moreover, with the slightly better constant $1/(5\lceil r_1 \rceil)$ in the exponent). The case $y = (0, 1)$ is very similar: letting $y_k = (\lfloor k/2 \rfloor, \lceil k/2 \rceil)$ (so again o', o, y are y_{-1}, y_0, y_1, respectively), and (say) $K = 2\lceil r_1 \rceil$ the above display holds unchanged, and as $|y_K| \geq r_1$ (with room to spare), we again get Eq. (3.51).

It remains to show Eq. (3.51) for $|y| < r_1$ that is not incident to o, o'. The strategy would be to first show that, for any (arbitrarily chosen) neighbor $y' \sim y$,

$$\widehat{\pi}_\infty\Big(\phi_y \wedge \phi_{y'} \geq h - \frac{h}{\log h} \;\Big|\; \phi_o \wedge \phi_{o'} \geq h\Big) \leq \exp\Big(-\frac{1}{5\lceil r_1 \rceil}(\beta - C)\frac{h}{\log h}\Big)\,, \tag{3.53}$$

and thereafter use this for the sought bound on y. Towards establishing Eq. (3.53), we will set aside the fact that y was chosen ahead of y', and treat these as two interchangeable adjacent sites.

If yy' is a horizontal edge—denoting $y = (a, b)$ and $y' = (a - 1, b)$—then we let

$$y_k := (ka, kb)\,, \qquad y'_k := y_k - (1, 0)\,.$$

If it is a vertical edge—denoting $y = (a, b)$ and $y' = (a, b - 1)$, and here we assume w.l.o.g. that if $a > 0$ then $b \geq 0$ and if $a < 0$ then $b \leq 0$ (otherwise, reflect ϕ about the x-axis)—then we let

$$y_k := \begin{cases} (\frac{k}{2}(a + b), \frac{k}{2}(b - a)) & k \text{ is even} \\ (\frac{k-1}{2}(a + b) + a, \frac{k-1}{2}(b - a) + b) & k \text{ is odd}\end{cases}\,, \qquad y'_k := \begin{cases} y_k - (1, 0) & k \text{ is even} \\ y_k - (0, 1) & k \text{ is odd}\end{cases}\,.$$

By $\mathbb{Z}^2$ symmetry (translation, $\frac{\pi}{2}$-rotation), in both cases the law of $\phi\restriction_{\{y_k, y'_k\}}$ given $\phi\restriction_{\{y_{k-1}, y'_{k-1}\}}$ is equal to that of $\phi\restriction_{\{y, y'\}}$ given $\phi\restriction_{\{o, o'\}}$. Moreover, $|y_k| \geq k$ (in the second case, $|a + b| \geq 2$ by the choice of sign(b) and since $y \not\sim o, o'$). In particular, setting $K = \lceil r_1 \rceil$ we guarantee that $|y_K| \geq r_1$.

Revisiting Eq. (3.52), we replace the events $\{\phi_{y_k} \geq h\}$ there by $\{\phi_{y_k} \wedge \phi_{y'_k} \geq h - \frac{kh}{K \log h}\}$, to find that by the exact same reasoning,

$$\begin{aligned}&\widehat{\pi}_\infty\Big(\phi_{y_K} \wedge \phi_{y'_K} \geq h - \frac{h}{\log h} \;\Big|\; \phi_o \wedge \phi_{o'} \geq h\Big) \\ &\qquad\qquad \geq \prod_{k=1}^K \widehat{\pi}_\infty\Big(\phi_{y_k} \wedge \phi_{y'_k} \geq h - \frac{kh}{K\log h} \;\Big|\; \phi_{y_{k-1}} \wedge \phi_{y'_{k-1}} \geq h - \frac{(k-1)h}{K\log h}\Big)\,.\end{aligned}$$

Notice that for every $k \geq 1$ and M, chosen sufficiently large such that $y_{k-1}, y'_{k-1}, y_k, y'_k \in B_M(o)$, we have by shift invariance and monotonicity that

$$\begin{aligned}\widehat{\pi}^0_{B_M(o)}\Big(\phi_{y_k} \wedge \phi_{y'_k} \geq h - \frac{kh}{K\log h} \;\Big|\; \phi_{y_{k-1}} \wedge \phi_{y'_{k-1}} \geq h - \frac{(k-1)h}{K\log h}\Big) \\ = \widehat{\pi}^{\frac{(k-1)h}{K\log h}}_{B_M(o)}\Big(\phi_{y_k} \wedge \phi_{y'_k} \geq h - \frac{h}{K\log h} \;\Big|\; \phi_{y_{k-1}} \wedge \phi_{y'_{k-1}} \geq h\Big) \\ \geq \widehat{\pi}^0_{B_M(o)}\Big(\phi_{y_k} \wedge \phi_{y'_k} \geq h - \frac{h}{K\log h} \;\Big|\; \phi_{y_{k-1}} \wedge \phi_{y'_{k-1}} \geq h\Big)\,.\end{aligned}$$

Taking $M \to \infty$ yields this inequality under $\widehat{\pi}_\infty$, and combined with the previous display we get

$$\widehat{\pi}_\infty\Big(\phi_{y_K} \wedge \phi_{y'_K} \geq h - \frac{h}{\log h} \;\Big|\; \phi_o \wedge \phi_{o'} \geq h\Big) \geq \widehat{\pi}_\infty\Big(\phi_y \wedge \phi_{y'} \geq h - \frac{h}{\log h} \;\Big|\; \phi_o \wedge \phi_{o'} \geq h\Big)^K\,.$$

The left-hand is at most $\widehat{\pi}_\infty\big(\phi_{y_K} \geq h - \frac{h}{\log h} \mid \phi_o \wedge \phi_{o'} \geq h\big)$, which we may bound from above via Eq. (3.50) since $|y_K| \geq r_1$, thus establishing Eq. (3.53).

To derive the required bound on ϕ_y, first note that one can infer from Eq. (3.53) that either y satisfies Eq. (3.51) or y' does, after using FKG to bound the arithmetic mean of these probabilities. In case y satisfies Eq. (3.51) for any choice of y' as one of its neighbors, we are done. Otherwise, every $y' \sim y$ has $\widehat{\pi}_\infty(\phi_{y'} \geq h - \frac{h}{\log h} \mid \phi_o \wedge \phi_{o'} \geq h) \leq \exp(-\frac{1}{10\lceil r_1\rceil}(\beta - C)\frac{h}{\log h})$, whence

$$\begin{aligned}&\widehat{\pi}_\infty(\phi_y \geq h \mid \phi_o \wedge \phi_{o'} \geq h) \\ &\quad\leq \widehat{\pi}_\infty\Big(\max_{y'\sim y}\phi_{y'} \geq h - \frac{h}{\log h} \;\Big|\; \phi_o \wedge \phi_{o'} \geq h\Big) + \widehat{\pi}_\infty\Big(\phi_y \geq h \;\Big|\; \max_{y'\sim y}\phi_{y'} \leq h - \frac{h}{\log h}\Big) \\ &\quad\leq 4e^{-\frac{1}{10\lceil r_1\rceil}(\beta-C)\frac{h}{\log h}} + e^{-4\beta(\frac{h}{\log h})^2}\end{aligned}$$

(using for the last line a Peierls map that decreases ϕ_y by $\lceil\frac{h}{\log h}\rceil$), giving the required Eq. (3.51). ∎

As a consequence of Claim 3.17, we can now infer an analogue of Claim 3.6 for $\mathbf{P}_{\ell,h}$, which will be the key to improving the upper bound in Lemma 3.8.

Corollary 3.18. *Let $h = H + 1 - n$ for n fixed, and $\ell \leq \sqrt{L}\exp(-\frac{1}{2}\frac{\log L}{\log\log L})$. If o' is a neighbor of the origin o, then there exists an absolute constant $c > 0$ such that*

$$\mathbf{P}_{\ell,h}\left(\phi_{o'} = -h \mid \phi_o = -h\right) \geq e^{-c\frac{\log L}{\log\log L}}\,.$$

Proof. Recalling the definition $\mathbf{P}_{\ell,h} = \widehat{\pi}_\infty(\cdot \mid \mathscr{F}_{\ell,h})$, we have

$$\begin{aligned}\mathbf{P}_{\ell,h}(\phi_{o'} = -h \mid \phi_o = -h) &\geq \widehat{\pi}_\infty\left(\phi_{o'} = -h,\, \mathscr{F}_{\ell,h} \mid \phi_o = -h\right) \\ &= \widehat{\pi}_\infty\left(\mathscr{F}_{\ell,h} \mid \phi_{o'} = \phi_o = -h\right)\widehat{\pi}_\infty\left(\phi_{o'} = -h \mid \phi_o = -h\right)\,.\end{aligned}$$

Since $\widehat{\pi}_\infty(\phi_{o'} = -h \mid \phi_o = -h) \geq e^{-c\beta\frac{h^2}{\log^2 h}}$ by Claim 3.6, which for the value of h here is $e^{-c\frac{\log L}{\log\log L}}$, it will suffice to show that

$$\widehat{\pi}_\infty\left(\mathscr{F}_{\ell,h} \mid \phi_o = \phi_{o'} = -h\right) = 1 - o(1)\,. \tag{3.54}$$

To this end, recall that in the proof of Lemma 3.8, we bounded $\widehat{\pi}_\infty(\mathscr{F}^c_{\ell,h} \mid \phi_o = -h)$ by splitting the treatment of $x \in Q_\ell$ into those at distance larger than $\log L$ from the origin and those within said distance, leading to Eq. (3.16). The same reasoning applies here, showing

$$\widehat{\pi}_\infty(\mathscr{F}^c_{\ell,h} \mid \phi_o = \phi_{o'} = -h) \leq e^{-(1-o(1))\frac{\log L}{\log\log L}} + O(\log^2 L)\max_{y\in B_{\log L}(o)}\widehat{\pi}_\infty(\phi_y < -h \mid \phi_o = \phi_{o'} = -h)\,.$$

(Note that here we used the assumption on ℓ to control $y \in Q_\ell$ at distance larger than $\log L$ from the origin via a union bound over all $|Q_\ell| \leq L\exp(-\frac{\log L}{\log\log L})$ such sites).

It remains to bound $\widehat{\pi}_\infty(\phi_y > h \mid \phi_o = \phi_{o'} = h)$. To this end, can use a Peierls argument via the map that decreases by 1 the heights of all sites in the connected component of $\{x : \phi_x > h\}$ intersecting $\{o, o'\}$, to find that $\widehat{\pi}_\infty(\phi_o = \phi_{o'} = h) \geq (1 - o(1))\widehat{\pi}_\infty(\phi_o \geq h,\, \phi_{o'} \geq h)$. Hence,

$$\widehat{\pi}_\infty(\phi_y > h \mid \phi_o = \phi_{o'} = h) \leq (1 + o(1))\widehat{\pi}_\infty(\phi_y > h \mid \phi_o \geq h,\, \phi_{o'} \geq h)$$
$$\leq e^{-c\beta \frac{h}{\log h}}\,,$$

where the last inequality is by Claim 8.8. For h as in our hypothesis, this at most $\exp(-c\sqrt{\beta \frac{\log L}{\log\log L}})$, which outweighs the $O(\log^2 L)$ factor, yielding Eq. (3.54) and thus completing the proof. ∎

Equipped with the above result, we can now establish the desired upper bound on $\bar{\xi}_{\ell,h}$ via the following analog of Lemma 3.5.

Lemma 3.19. *There exist absolute constants $\beta_0, c_1 > 0$ such that the following holds for all $\beta > \beta_0$. For $h = H + 1 - n$ with fixed $n \geq 0$ and ℓ with $\ell \geq L^\delta$ for fixed $\delta > 0$,*

$$\frac{\bar{\xi}_{\ell,h}}{\widehat{\pi}_\infty(\phi_o = -h)} \leq 1 - e^{-c_1 \frac{\log L}{\log\log L}}\,.$$

Proof. The proof of Lemma 3.5 extends verbatim to our setting of $\mathbf{P}_{\ell,h}$: by FKG,

$$\mathbf{P}_{\ell,h}(\phi_x > -h,\, \forall x \in Q_\ell) \geq \Big(1 - 2\mathbf{P}_{\ell,h}(\phi_o = -h) + \mathbf{P}_{\ell,h}(\phi_o = \phi_{o'} = -h)\Big)^{\lceil \ell/2 \rceil}\,,$$

and following the same argument in that proof, after replacing in it $\widehat{\pi}_\infty(\phi_o < -h)$ by $\mathbf{P}_{\ell,h}(\phi_o = -h)$ and the application of Claim 3.6 by that of Corollary 3.18, we find that there exists some absolute constant $c > 0$ (given by said corollary) such that

$$\frac{\bar{\xi}_{\ell,h}}{\mathbf{P}_{\ell,h}(\phi_o = -h)} \leq 1 - e^{-c \frac{\log L}{\log\log L}}\,.$$

Now, $\mathbf{P}_{\ell,h}(\phi_o = -h) = \mathbf{P}_{\ell,h}(\phi_o \leq -h)$, which, by FKG and our assumption on h, satisfies

$$\mathbf{P}_{\ell,h}(\phi_o \leq -h) \leq \widehat{\pi}_\infty(\phi_o \leq -h \mid \phi_o \geq -h) \leq \frac{\widehat{\pi}_\infty(\phi_o = -h)}{1 - \widehat{\pi}_\infty(\phi_o < -h)} \leq (1 + L^{-1+o(1)})\widehat{\pi}_\infty(\phi_o = -h)\,.$$

Combining the last two displays concludes the proof. ∎

Proof of Proposition 3.1, Eq. (3.4). As mentioned above, we already established the lower bound as part of Eq. (3.34). For the upper bound, set $\ell_0 = \lfloor L^{1/4} \rfloor$ as in the proof of Proposition 3.1. The sought bound now follows from substituting the bound on $\bar{\xi}_{\ell_0,h}$ from Lemma 3.19 in Eq. (3.35). ∎

4. Refined estimates on the law of the disagreement polymer

This section is devoted to establishing several refinements of Proposition 2.5. As mentioned in the beginning of Section 3, these will involve the rate p_n from Eq. (3.2).

The following result establishes the asymptotic law of the disagreement polymer in a box of area $L^{4/3} e^{O(\sqrt{\log L})}$ as opposed to $L(\log L)^{O(1)}$. Throughout the rest of this paper, the letter $\mathcal{G}$ will be used to denote the conditions on γ needed to apply the relevant polymer law proposition, which will change depending on the context.

Proposition 4.1 (Law of a mesoscopic disagreement polymer)**.** *Fix $n \geq 0$ and $g \geq 0$. There exists β_0 such that the following holds for all $\beta \geq \beta_0$. Let $V \subset \mathbb{Z}^2$ be a connected domain with g holes, and consider the* $\mathbb{Z}$GFF *model $\pi^\eta_{V;F}$ with a floor at 0 imposed only on a subset $F \subset V$, and boundary conditions η that are $H + 1 - n$ on a $*$-connected path in $\partial_{\mathrm{v}} V$ and $H - n$ elsewhere so that they induce a unique disagreement polymer $(\gamma, \{D_i\}, \{h_i\})$ in $V \cup \partial_{\mathrm{v}} V$ that contains boundary* *disagreements. Further suppose* $|\partial_{\mathsf{v}} F| \vee |\partial_{\mathsf{v}} V| \leq O(L^{2/3}e^{\sqrt{\log L}})$*, denote by* D_0 *and* D_1 *the regions of* γ *containing the boundary vertices of* V *at heights* $H+1-n$ *and* $H-n$*, respectively, and let*

$$\mathcal{G} := \left\{|\gamma| \leq L^{2/3}e^{\sqrt{\log L}}\right\},$$
$$\mathcal{G}^\dagger := \mathcal{G} \cap \left\{\max_{i\geq 2} |\partial D_i| < \log L\right\}.$$

Then

$$\pi^\eta_{V;F}(\mathcal{G}^\dagger \mid \mathcal{G}) = 1 - o(1)\,, \tag{4.1}$$
$$\pi^\eta_{V;F}(\gamma \mid \mathcal{G}^\dagger) = (1+o(1))\mathfrak{q}^\eta_{V;F}(\gamma) \quad \textit{uniformly over all} \quad \gamma \in \mathcal{G}^\dagger\,, \tag{4.2}$$

where $\mathfrak{q}^\eta_{V;F}$ *is the probability distribution on* $\gamma \in \mathcal{G}^\dagger$ *given by*

$$\mathfrak{q}^\eta_{V;F}(\gamma) := \frac{1}{Z^\eta_{V;F}} \exp\Big(- \mathscr{E}^*_\beta(\gamma) + \frac{\rho_n}{N_n}|D_0 \cap F| + \mathfrak{I}_V(\gamma)\Big)$$

with N_n *from Eq.* (1.3)*,* $\mathfrak{I}_V$ *from Eq.* (2.1)*,* ρ_n *from Eq.* (3.2)*, and a normalizer* $Z^\eta_{V;F}$*.*

Proof. Let $(\gamma, \{D_i\}, \{h_i\})$ be a disagreement polymer. We start with the expression for $\pi^\eta_{V;F}(\gamma)$ given in Eq. (2.3) from Proposition 2.5, which holds without any restrictions on the D_i's or V, F, and recall the notation $\mathscr{E}^*_\beta(\gamma)$ from Eq. (2.6) that absorbed the $\widehat{\pi}^{h_i}_{D_i^\circ}$ terms for $i > 1$, so that

$$\pi^\eta_{V;F}(\gamma) \propto \exp\Big(- \mathscr{E}^*_\beta(\gamma) + \mathfrak{I}_V(\gamma)\Big) \prod_{i=0,1} \widehat{\pi}^{H+1-n-i}_{D_i^\circ}\big(\phi_x \geq 0,\, \forall x \in D_i^\circ \cap F\big)\,.$$

In view of this representation for $\pi^\eta_{V;F}$, we now further impose $\gamma \in \mathcal{G}$, and turn our attention to

$$\widehat{\pi}^{h_i}_{D_i^\circ}\big(\phi_x \geq 0,\, \forall x \in D_i^\circ \cap F\big) \qquad \text{for } i = 0,1\,.$$

The event $\mathcal{G}$ will enable us to appeal to the mesoscopic range of Theorem 3.2 for $\widehat{\pi}^0_{D_i^\circ}$ $(i = 0,1)$. Let us verify the required hypotheses:

1. The condition on the domain: $|\partial D_i^\circ| \leq O(|\gamma| + |\partial V|) = O(L^{2/3}e^{\sqrt{\log L}})$ as per $\mathcal{G}$.
2. The condition on the subset where we wish to control the minimum height: again we have that $|\partial(D_i^\circ \cap F)| \leq O(|\gamma| + |\partial V| + |\partial F|) = O(L^{2/3}e^{\sqrt{\log L}})$ as per $\mathcal{G}$.

Having qualified for an application of Eq. (3.5) from Theorem 3.2, we deduce that, for $i = 0,1$,

$$\begin{aligned}\widehat{\pi}^{H+1-n-i}_{D_i^\circ}\big(\phi_x \geq 0,\, \forall x \in D_i^\circ \cap F\big) &= \widehat{\pi}^0_{D_i^\circ}\big(\phi_x \geq (H+1-n-i),\, \forall x \in D_i^\circ \cap F\big)\\ &= (1+o(1))\exp\big(-\xi_{n+i}|D_i \cap F|\big)\,,\end{aligned}$$

using here that $\xi_{n+i} = (1+o(1))\widehat{\pi}_\infty(\phi_o > H+1-n-i) \leq L^{-1+o(1)}$, and so

$$\xi_{n+i}|(D_i \setminus D_i^\circ) \cap F| \leq L^{-1+o(1)}O(|\gamma|) = L^{-1/3+o(1)} = o(1)\,.$$

We now move to multiply $\prod_{i\geq 0} \widehat{\pi}^{h_i}_{D_i^\circ}(\phi_x \geq 0,\, \forall x \in D_i^\circ \cap F)$ by $\exp(\xi_{n+1}|F|)$, which can be absorbed into the partition function for $\widehat{\pi}^\eta_{V;F}$, being independent of γ. This cancels the term $-\xi_{n+1}|D_1 \cap F|$ in the exponent, whereas $(\xi_{n+1} - \xi_n)|D_0 \cap F| = (\rho_n/N_n)|D_0 \cap F|$ by the definition of ρ_n in Eq. (3.2); thus,

$$\pi^\eta_{V;F}(\gamma \mid \mathcal{G}) = (1+o(1))\mathfrak{p}^\eta_{V;F}(\gamma)$$

uniformly over $\gamma \in \mathcal{G}$, where $\mathfrak{p}^\eta_{V;F}$ is the probability distribution on $\gamma \in \mathcal{G}$ given by

$$\mathfrak{p}^\eta_{V;F}(\gamma) := \frac{1}{\tilde{Z}^\eta_{V;F}} \exp\Big(- \mathscr{E}^*_\beta(\gamma) + \frac{\rho_n}{N_n}|D_0 \cap F| + \mathfrak{I}_V(\gamma) + \xi_{n+1}\sum_{i\geq 2}|D_i \cap F|\Big)\,, \tag{4.3}$$

in which $\tilde{Z}^{\eta}_{V;F}$ is a normalizer. We now move to show Eq. (4.1) using this intermediary form of the law of γ, which throughout this proof we denote by $\mathfrak{p}(\gamma)$ in lieu of $\mathfrak{p}^{\eta}_{V;F}(\gamma)$ for brevity. Recall Definition 2.16, according to which

$$\mathcal{G} \setminus \mathcal{G}^{\dagger} = \left\{ \gamma \in \mathcal{G} \,:\, \max_{\mathcal{D} \in \mathfrak{D}(\gamma)} |\partial \mathcal{D}| > \log L \right\}.$$

We will use a Peierls argument to show that, for some constant $c > 0$,

$$\sum_{\gamma \in \mathcal{G} \setminus \mathcal{G}^{\dagger}} \mathfrak{p}(\gamma) \leq L^{-c\beta} \sum_{\gamma \in \mathcal{G}} \mathfrak{p}(\gamma). \tag{4.4}$$

There are only L^2 locations for a bond b, so fix b and consider the map T_b on γ defined as follows. If $b \notin \gamma$ or if $\mathcal{D}_b = \emptyset$, then T_b is the identity map. Otherwise, let u, v be the two cut-points delineating $\mathcal{D}_b$. Let $T_b(\gamma)$ be the disagreement polymer which replaces $\mathcal{D}_b$ by a (arbitrarily chosen, say the outermost such) minimal length path from u to v that remains within $\mathcal{D}_b$. Then,

$$\mathscr{E}_{\beta}(T_b(\gamma)) \leq \mathscr{E}_{\beta}(\gamma) - \tfrac{\beta}{2} |\partial \mathcal{D}_b|.$$

Moreover, we only removed components D_i for $i \geq 2$, and each such component contributes a nonnegative term (namely $-\log \widehat{\pi}^{h_i}_{D_i^{\circ}}(\phi_x \geq 0,\, \forall x \in D_i^{\circ} \cap F)$) to $\mathscr{E}^*_{\beta}$, hence we also have

$$\mathscr{E}^*_{\beta}(T_b(\gamma)) \leq \mathscr{E}^*_{\beta}(\gamma) - \tfrac{\beta}{2} |\partial \mathcal{D}_b|.$$

We also could only have increased $|D_0|$ in $T_b(\gamma)$. However, $T_b(\gamma)$ did in fact lose the contribution of $\mathcal{D}_b$ to $\xi_{n+1} \sum_{i \geq 2} |D_i \cap F|$. By the decay properties of Φ, we therefore have for some C that

$$\mathfrak{p}(\gamma) \leq \mathfrak{p}(T_b(\gamma)) e^{-\frac{\beta - C}{2} |\partial \mathcal{D}_b| + \xi_{n+1} |\mathcal{D}_b|}.$$

Note that we necessarily have the crude bound that $\max_{\mathcal{D} \in \mathfrak{D}(\gamma)} |\partial \mathcal{D}| \leq L^{2/3} e^{\sqrt{\log L}}$. By isoperimetry and this crude bound, we have $|\mathcal{D}_b| \leq |\partial \mathcal{D}_b|^2/16 \leq |\partial \mathcal{D}_b| L^{2/3} e^{\sqrt{\log L}}/16$ which after multiplying by ξ_{n+1} is $o(1)$. Hence, the above display implies that for some $c > 0$,

$$\mathfrak{p}(\gamma) \leq \mathfrak{p}(T_b(\gamma)) e^{-c\beta |\partial \mathcal{D}_b|}.$$

Moreover, the number of preimages γ under T_b for a given disagreement polymer γ' such that $|\partial \mathcal{D}_b(\gamma)| = k$ is bounded above by the number of connected components of bonds of size k times the number of possible points u. There are C^k choices for the connected component, and k^2 choices for the location of the cut-point u at which to attach the component.

Putting the above all together proves Eq. (4.4):

$$\begin{aligned} \sum_{\gamma \in \mathcal{G} \setminus \mathcal{G}^{\dagger}} \mathfrak{p}(\gamma) &\leq \sum_{\gamma' \in \mathcal{G}} \sum_{\substack{k \geq \log L \\ b \in \Lambda}} \sum_{\substack{\gamma \in \mathcal{G} \cap T_b^{-1}(\gamma') \\ |\partial \mathcal{D}_b(\gamma)| = k}} \mathfrak{p}(\gamma) \\ &\leq \sum_{\gamma' \in \mathcal{G}} \sum_{k \geq \log L} L^2 k^2 C^k \mathfrak{p}(\gamma') e^{-c\beta k} \\ &\leq e^{-c\beta \log L} \sum_{\gamma' \in \mathcal{G}} \mathfrak{p}(\gamma'). \end{aligned} \tag{4.5}$$

This proves Eq. (4.1).

Finally, for all $\gamma \in \mathcal{G}^{\dagger}$, the total area of $F \cap (\bigcup_{i \geq 2} D_i)$ is at most $L^{2/3} e^{\sqrt{\log L}} \log L$, hence we find that $\xi_{n+1} \sum_{i \geq 2} |D_i \cap F| = L^{-1/3 + o(1)} = o(1)$. ∎

Proposition 4.2 (Law of a macroscopic disagreement polymer)**.** *Fix $n \geq 0$ and $g \geq 0$. In the setting of Proposition 4.1, under the relaxed condition $|\partial_{\mathsf{v}} F| \vee |\partial_{\mathsf{v}} V| \leq O(Le^{\sqrt{\log L}})$ and the modified definition of the event $\mathcal{G}$ as*

$$\mathcal{G} := \Big\{|\gamma| \leq Le^{\sqrt{\log L}}\Big\} \cap \bigcap_{i=0,1} \Big\{|D_i \cap F| \leq \Big(\frac{3\beta}{\widehat{\pi}_\infty(\phi_o > H+1-n-i)}\Big)^2\Big\}\,,$$

the following holds. The probability distribution $\mathfrak{p}^\eta_{V;F}$ on $\gamma \in \mathcal{G}$ given by

$$\mathfrak{p}^\eta_{V;F}(\gamma) := \frac{1}{Z^\eta_{V;F}} \exp\Big(- \mathscr{E}^*_\beta(\gamma) + \frac{\rho_n}{N_n}|D_0 \cap F| + \mathfrak{I}_V(\gamma) + \xi_{n+1} \sum_{i\geq 2} |D_i \cap F|\Big)$$

with N_n from Eq. (1.3), ρ_n from Eq. (3.2), and a normalizer $Z^\eta_{V;F}$, satisfies that

$$\pi^\eta_{V;F}(\gamma \mid \mathcal{G}) = \exp\Big(O\Big(\sqrt{L}e^{\frac{\log L}{\log\log L}}\Big)\Big)\, \mathfrak{p}^\eta_{V;F}(\gamma) \tag{4.6}$$

uniformly over all $\gamma \in \mathcal{G}$.

Proof. The proof will follow via exactly the same argument that produced Eq. (4.3) in the proof of Proposition 4.1. As argued there, we have

$$\pi^\eta_{V;F}(\gamma) \propto \exp\Big(- \mathscr{E}^*_\beta(\gamma) + \mathfrak{I}_V(\gamma)\Big) \prod_{i=0,1} \widehat{\pi}^{H+1-n-i}_{D_i^\circ}\big(\phi_x \geq 0,\, \forall x \in D_i^\circ \cap F\big)\,,$$

and next wish to qualify for an application of Theorem 3.2, albeit this time in the macroscopic range. Let us indeed verify the required hypotheses for applying said theorem for $\widehat{\pi}^0_{D_i^\circ}$ $(i = 0, 1)$.

(1) The condition on the domain boundary: $|\partial D_i^\circ| \leq O(|\gamma| + |\partial V|) = O(Le^{\sqrt{\log L}})$ as per $\mathcal{G}$.
(2) The stronger requirement needed to forgo the conditioning on $\mathfrak{S}$: we wish to satisfy

$$|D_i^\circ \cap F| \leq \Big(\frac{3\beta}{\widehat{\pi}_\infty(\phi_o > H+1-n-i)}\Big)^2\,, \tag{4.7}$$

which is guaranteed (moreover for $D_i \cap F$) as per $\mathcal{G}$.

Therefore, Eqs. (3.6) and (3.7) are applicable, where in the latter we have the unconditional version secured by that theorem via the additional requirement Eq. (4.7), and we infer that for $i = 0, 1$,

$$\widehat{\pi}^{H+1-n-i}_{D_i^\circ}\big(\phi_x \geq 0,\, \forall x \in D_i^\circ \cap F\big) = \exp\Big(-\xi_{n+i}|D_i \cap F| + O\big(\sqrt{L}e^{\frac{\log L}{\log\log L}}\big)\Big)\,.$$

Note that, in the above display, we further replaced $\xi_{n+i}|D_i^\circ \cap F|$ by $\xi_{n+i}|D_i \cap F|$, absorbing the difference between the two into the $O(L^{1/2+o(1)})$ error term since

$$\xi_{n+i}|D_i \setminus D_i^\circ| \leq L^{-1+o(1)} O(|\gamma|) \leq L^{o(1)}\,.$$

The $1 + o(1)$ factor that was associated with (3.5) from that theorem in the mesoscopic range, in the proof of Proposition 4.1, will thereby be replaced by the $\exp(L^{1/2+o(1)})$ from Eqs. (3.6) and (3.7) in the macroscopic range. Proceeding as in the proof of Eq. (4.3) by introducing a factor of $\exp(\xi_{n+1}|F|)$, that is shifted into the partition function, we arrive at Eq. (4.6), as required. ∎

Proposition 4.3 (probability of a given disagreement polymer with uniform boundary conditions)**.** *Fix $g \geq 0$, $n \geq 0$ and set $h = H+1-n$. Let $F \subseteq V \subset \mathbb{Z}^2$ where V is connected, contains the origin and has g holes. Suppose $|\partial F| \vee |\partial V| \leq O(Le^{\sqrt{\log L}})$. Let γ be a disagreement polymer where D_0 is the region containing the origin with $h_0 = h$, and D_1 is the region containing the boundary vertices ∂V with $h_1 = h - 1$. Moreover, assume that $|D_1 \cap F| \leq \big(\frac{3\beta}{\widehat{\pi}_\infty(\phi_o > h-1)}\big)^2$ and $|\gamma| \leq Le^{\sqrt{\log L}}$. Then, letting*

$$\mathfrak{p}(\gamma) = \exp\Big(- \mathscr{E}^*_\beta(\gamma) + \frac{\rho_n}{N_n}|D_0 \cap F| + \mathfrak{I}_V(\gamma) + \xi_{n+1} \sum_{i\geq 2} |D_i \cap F|\Big)\,,$$

we have

$$\pi^{h-1}_{V;F}(\mathfrak{S})e^{O(L^{1/2+o(1)})} \leq \frac{\pi^{h-1}_{V;F}(\mathcal{C}_{\gamma,h})}{\mathfrak{p}(\gamma)} \leq \frac{1}{\pi^{h}_{D_0^\circ\cap F}(\mathfrak{S})}e^{O(L^{1/2+o(1)})}\,. \tag{4.8}$$

Proof. By the cluster expansion for disagreement polymers (see [14, Eq. (2.28)]), we know that if $w(\phi) := \prod_{\gamma'\in\phi}\exp(-\mathscr{E}_\beta(\gamma'))$ and $\widehat{Z}^h_V = \sum_\phi w(\phi)$, then

$$\frac{\prod_i \widehat{Z}^{h_i}_{D_i^\circ}}{\widehat{Z}^0_V} = \exp\Bigg(\sum_{\substack{\mathsf{W}\subset V\\ \mathsf{W}\cap\Delta^*_\gamma\neq\emptyset}}\Phi(\mathsf{W})\Bigg) = \exp\left(\mathfrak{I}_V(\gamma)\right)\,. \tag{4.9}$$

Letting $Z^{h-1}_{V;F} = \sum_\phi w(\phi)\mathbb{1}_{\{\phi_x\geq 0,\,\forall x\in F\}}$ for the $\mathbb{Z}$GFF with boundary conditions $h-1$ and a zero floor in F (as well as defining $Z^{h_i}_{D_i^\circ;F}$ analogously), we have

$$\pi^{h-1}_{V;F}(\mathcal{C}_{\gamma,h}) = \frac{1}{Z^{h-1}_{V;F}}\exp\left(-\mathscr{E}_\beta(\gamma)\right)\prod_{i\geq 0}Z^{h_i}_{D_i^\circ;F}\,,$$

and since $Z^{h-1}_{V;F} = \widehat{Z}^h_{V;F}\widehat{\pi}^{h-1}_V(\phi_x\geq 0,\,\forall x\in F)$ (and analogously for $Z^{h_i}_{D_i^\circ;F}$), we infer that

$$\begin{aligned}\pi^{h-1}_{V;F}(\mathcal{C}_{\gamma,h}) &= \frac{1}{\widehat{Z}^{h-1}_V\widehat{\pi}^{h-1}_V(\phi_x\geq 0,\,\forall x\in F)}\exp\left(-\mathscr{E}_\beta(\gamma)\right)\prod_{i\geq 0}\widehat{Z}^{h_i}_{D_i^\circ}\widehat{\pi}^{h_i}_{D_i^\circ}(\phi_x\geq 0,\,\forall x\in D_i^\circ\cap F)\\ &= \exp\Big(-\mathscr{E}^*_\beta(\gamma)+\mathfrak{I}_V(\gamma)\Big)\frac{\prod_{i=0,1}\widehat{\pi}^{h-i}_{D_i^\circ}(\phi_x\geq 0,\,\forall x\in F\cap D_i^\circ)}{\widehat{\pi}^{h-1}_V(\phi_x\geq 0,\,\forall x\in F)}\,,\end{aligned} \tag{4.10}$$

where the last transition used Eq. (4.9) and the definition of $\mathscr{E}^*_\beta$ from Eq. (2.6).

For the upper bound on Eq. (4.8), we can lower bound the denominator in the right of Eq. (4.10) using Eq. (3.6) of Theorem 3.2, showing that

$$\begin{aligned}\pi^{h-1}_{V;F}(\mathcal{C}_{\gamma,h}) \leq &\exp\Bigg(-\mathscr{E}^*_\beta(\gamma)+\mathfrak{I}_V(\gamma)+\xi_{n+1}\sum_{i\geq 0}|D_i\cap F|+O(L^{1/2+o(1)})\Bigg)\\ &\cdot\prod_{i=0,1}\widehat{\pi}^{h-i}_{D_i^\circ}\left(\phi_x\geq 0,\,\forall x\in D_i^\circ\cap F\right)\,,\end{aligned} \tag{4.11}$$

and we move our attention to the two terms in the product over $i=0,1$.

For $i=1$, the assumption on $|D_1\cap F|$ allows us to infer from Eq. (3.7) the unconditional bound

$$\widehat{\pi}^{h-1}_{D_1^\circ}\left(\phi_x\geq 0,\,\forall x\in D_1^\circ\cap F\right)\leq\exp\Big(-\xi_{n+1}|D_1^\circ\cap F|+O(L^{1/2+o(1)})\Big)\,.$$

Since $\xi_{n+1}|D_1\setminus D_1^\circ|\leq L^{-1+o(1)}O(|\gamma|)=L^{o(1)}$, we can replace $\exp(-\xi_{n+1}|D_1^\circ\cap F|)$ in the above by $\exp(-\xi_{n+1}|D_1\cap F|)$, which cancels the corresponding term in the right hand of Eq. (4.11).

For $i=0$, by Eq. (3.7) (in the conditional version, as we did not assume that $|D_0^\circ\cap F|$ is small):

$$\widehat{\pi}^{h}_{D_0^\circ}\left(\phi_x\geq 0,\,\forall x\in D_0^\circ\cap F\mid\mathfrak{S}\right)\leq\exp\Big(-\xi_n|D_0^\circ\cap F|+O(L^{1/2+o(1)})\Big)\,.$$

As in the proof of Proposition 3.9, the routine bound $\mathbb{P}(A)\leq\frac{\mathbb{P}(A|B)}{\mathbb{P}(B|A)}$ for $A=\{\phi_x\geq -h,\,\forall x\in F\}$ and $B=\mathfrak{S}$ shows that

$$\widehat{\pi}^{h}_{D_0^\circ}\left(\phi_x\geq 0,\,\forall x\in D_0^\circ\cap F\right)\leq\frac{\exp\Big(-\xi_n|D_0\cap F|+O(L^{1/2+o(1)})\Big)}{\widehat{\pi}^h_{D_0^\circ;F}(\mathfrak{S})}\,,$$

where again we absorbed the move from $\xi_n|D_0^\circ \cap F|$ to $\xi_n|D_0 \cap F|$ into the error term. Together with the $\xi_{n+1}|D_0 \cap F|$ from Eq. (4.11), we see that the prefactor of $|D_0 \cap F|$ is $\xi_{n+1} - \xi_n = \rho_n/N_n$ as per Eq. (3.2), thus arriving at the upper bound of Eq. (4.8).

For the lower bound in Eq. (4.8), since $\pi^{h-1}_{V;F}(\mathfrak{S})$ is nothing but $\widehat{\pi}^{h-1}_V(\mathfrak{S} \mid \phi_x \geq 0,\, \forall x \in F)$, we see from Eq. (4.10) that

$$\begin{aligned}\frac{\pi^{h-1}_{V;F}(\mathcal{C}_{\gamma,h})}{\pi^{h-1}_{V;F}(\mathfrak{S})} &= \exp\Big(-\mathscr{E}^*_\beta(\gamma) + \mathfrak{I}_V(\gamma)\Big)\frac{\prod_{i=0,1}\widehat{\pi}^{h-i}_{D_i^\circ}(\phi_x \geq 0,\, \forall x \in F \cap D_i^\circ)}{\widehat{\pi}^{h-1}_V(\{\phi_x \geq 0,\, \forall x \in F\} \cap \mathfrak{S})} \\ &\geq \exp\Big(-\mathscr{E}^*_\beta(\gamma) + \mathfrak{I}_V(\gamma)\Big)\frac{\prod_{i=0,1}\widehat{\pi}^{h-i}_{D_i^\circ}(\phi_x \geq 0,\, \forall x \in F \cap D_i^\circ)}{\widehat{\pi}^{h-1}_V(\phi_x \geq 0,\, \forall x \in F \mid \mathfrak{S})}\,. \end{aligned} \tag{4.12}$$

By Eqs. (3.6) and (3.7) of Theorem 3.2,

$$\begin{aligned}\widehat{\pi}^{h-1}_V(\phi_x \geq 0,\, \forall x \in F \mid \mathfrak{S}) &\leq \exp\Big(-\xi_{n+1}|F| + O(L^{1/2+o(1)})\Big)\,, \\ \widehat{\pi}^{h-1}_V(\phi_x \geq 0,\, \forall x \in D_i^\circ \cap F) &\geq \exp\Big(-\xi_{n+1}|D_i \cap F| + O(L^{1/2+o(1)})\Big) \quad \text{for } i = 0,1\,,\end{aligned}$$

where, as before, we moved from $|D_i \cap F|$ from $|D_i^\circ \cap F|$ via an additive (negligible) error of $O(L^{o(1)})$. Substituting this in Eq. (4.12) establishes the lower bound in Eq. (4.8), completing the proof. ∎

Corollary 4.4. *Consider the setting of Proposition 4.3 for $n = 0$, where the assumption on $|D_1 \cap F|$ is implied by $|D_1 \cap F| \leq L^2/3$. If we further assume that $|F| \leq L^2$ then we have*

$$\pi^H_{V;F}(\mathfrak{S})\exp\Big(O(L^{1/2+o(1)})\Big) \leq \frac{\pi^H_{V;F}(\mathcal{C}_{\gamma,H+1})}{\mathfrak{p}(\gamma)} \leq \exp\Big(O(L^{1/2+o(1)})\Big)\,. \tag{4.13}$$

Proof. By the definition of H, we have $\widehat{\pi}_\infty(\phi_o = H) \geq \frac{5\beta}{L} > \widehat{\pi}_\infty(\phi_o = H+1)$, and we immediately observe that the condition $|D_1 \cap F|^{1/2} \leq 3\beta/\widehat{\pi}_\infty(\phi_o > H)$ would be satisfied if $|D_1 \cap F|^{1/2} \leq \frac{3}{5}L$. The lower bound in Eq. (4.13) is verbatim the one given in Eq. (4.8), and it remains to show that $\pi^{H+1}_{D_0^\circ \cap F}(\mathfrak{S}) = O(L^{-1})$, say, so that the factor $1/\pi^{H+1}_{D_0^\circ \cap F}(\mathfrak{S})$ in the upper bound in Eq. (4.8) would be overtaken by the $O(L^{1/2+o(1)})$ in its exponent. To see this, recall as usual that ruling out `large` $h-$level lines for $h < H+1$ is immediate from the standard Peierls argument, which is unperturbed by the presence of a floor (see, e.g., [29, Eq. (4.2) of Prop. 4.1]). Consider an $(H+2)$ level line $\mathfrak{L}$. Recalling that $\widehat{\pi}_\infty(\phi_o > h) \leq e^{-c\beta\frac{h}{\log h}}\widehat{\pi}_\infty(\phi_o = h) = o(\widehat{\pi}_\infty(\phi_o = h))$ by Eq. (2.9), we see that $\widehat{\pi}_\infty(\phi_o = H+2) = o(1/L)$. Therefore, by Lemma 2.7, combined with the isoperimetric inequality $|\operatorname{Int}(\mathfrak{L}) \cap F| \leq |\mathfrak{L}|\sqrt{|F|}/4 \leq O(|\mathfrak{L}|L)$ (using here our assumption on $|F|$), it follows that

$$\pi^{H+1}_{D_0^\circ \cap F}(\mathcal{C}_{\mathfrak{L},H+2}) \leq \exp\Big(-\Big(\beta - \widehat{\pi}_\infty(\phi_o > H+1)O(L) - o(1)\Big)|\mathfrak{L}|\Big) \leq \exp\left(-\left(\beta - o(1)\right)|\mathfrak{L}|\right)\,.$$

At this point, a union bound over all `large` level-line loops $\mathfrak{L}$ shows that, say, $\pi^{H+1}_{D_0^\circ \cap F}(\mathfrak{S}) \geq 1 - L^{-10}$, and the proof is complete. ∎

5. Asymptotically small critical window

In this section we prove Theorem 1.2, showing that except for an asymptotically small critical window about $L_c^{(h)}$, we can determine w.h.p. whether the top level line is at height H or height $H+1$. We broadly follow the proof strategy of [13, Sections 4.2, 4.3]. However, we will require some extra quantitative estimates utilizing Lemma 2.12, pushing towards a critical window of width at most $(L_*^{(h)})^{1/2+o(1)}$. We will also need the machinery on disagreement polymers developed in [14], and most importantly the law of disagreement polymers in macroscopic domains obtained in Section 4.

Define $L_*^{(h)} = \lceil \frac{\lambda_* \beta}{\rho_0 \widehat{\pi}_\infty(\phi_o = h)} \rceil$, recalling that $\lambda_* \approx 4(1 \pm \varepsilon_\beta)$. Throughout this section, we will fix $h \geq h_0$ for some h_0 sufficiently large depending on β, and define $\lambda = \lambda(L)$ to be such that $L = \frac{\lambda \beta}{\widehat{\pi}_\infty(\phi_o = h)}$, or equivalently $\widehat{\pi}_\infty(\phi_o = h) = \frac{\lambda \beta}{L}$. We split Theorem 1.2 into two propositions, showing separately that below the window $[L_*^{(h)} - (L_*^{(h)})^{1/2+o(1)}, L_*^{(h)} + (L_*^{(h)})^{1/2+o(1)}]$ the probability of having a `large` h level line is $o(1)$, and above the window this probability is $1 - o(1)$.

Proposition 5.1. *There exists $h_0 > 0$ such that for all $h \geq h_0$, for all $L \leq L_*^{(h)} - (L_*^{(h)})^{1/2+o(1)}$, the $\mathbb{Z}$GFF model on w.h.p. has no* `large` *level line at height h.*

Proposition 5.2. *There exists $h_0 > 0$ such that for all $h \geq h_0$, for all $L \geq L_*^{(h)} + (L_*^{(h)})^{1/2+o(1)}$, the $\mathbb{Z}$GFF model on Λ w.h.p. has a* `large` *level line at height h.*

The exponents of $1/2 + o(1)$ come from the error term in Proposition 4.3, which ultimately comes from Proposition 3.9.

Remark 5.3. The term $L_*^{(h)}$ predicts where $L_c^{(h)}$ is by solving an optimization problem using this prefactor of ρ_0 on the area tilt. One could define analogously a $\widetilde{L}_*^{(h)} := \lfloor \frac{\lambda_* \beta}{\widehat{\pi}_\infty(\phi_o = h)} \rfloor$, which predicts the location of $L_c^{(h)}$ using the same analysis but using an area tilt of $\frac{1}{N_0}$, as done in previous works (e.g. [13,14]). Indeed, [13] proves for SOS that $L_c^{(h)} \in \widetilde{L}_*^{(h)} \pm (\widetilde{L}_*^{(h)})^{o(1)}$, and at first glance it appears that perhaps $L_c^{(h)} = \widetilde{L}_*^{(h)}$ since the errors leading to this estimate seem purely technical.

However, at least for the $\mathbb{Z}$GFF, this is not the case. The difference between $L_*^{(h)}$ and $\widetilde{L}_*^{(h)}$ is of order $L_*^{(h)}(1 - \rho_n)$. At the same time, in this section we prove that the window around $L_*^{(h)}$ where we can guarantee $L_c^{(h)}$ is has size $(L_*^{(h)})^{1/2+o(1)}$ (primarily coming from the error term in the macroscopic range results of Theorem 3.2). Hence, if $1 - \rho_n \leq L^{-1/2+o(1)}$, then the gap between $L_*^{(h)}$ and $\widetilde{L}_*^{(h)}$ is larger than the window around $L_*^{(h)}$. This implies not only that $L_c^{(h)} \neq \widetilde{L}_*^{(h)}$, but moreover that the gap between the two is wide enough such that at $\widetilde{L}_*^{(h)}$ there is w.h.p. no `large` h level line. In other words, using an area tilt of $\frac{1}{N_0}$ gives an erroneous prediction for the real transition point for the h level line.

To connect to the notation of $H(L)$, note that for every L in the interval $[\frac{5\beta}{\widehat{\pi}_\infty(\phi_o = h-1)}, \frac{5\beta}{\widehat{\pi}_\infty(\phi_o = h)})$ we have $H(L) = h - 1$. Hence, the above propositions are equivalent to saying that for any $h \geq h_0$,

$$L \in \left[\tfrac{5\beta}{\widehat{\pi}_\infty(\phi_o = h-1)}, L_*^{(h)} - (L_*^{(h)})^{1/2+o(1)}\right] \leftrightarrow \text{ top level line at height } H(L)\,,$$
$$L \in \left(L_*^{(h)} - (L_*^{(h)})^{1/2+o(1)}, L_*^{(h)} + (L_*^{(h)})^{1/2+o(1)}\right) \leftrightarrow \text{ top l.l. at height } H(L) \text{ or } H(L)+1\,,$$
$$L \in \left[L_*^{(h)} + (L_*^{(h)})^{1/2+o(1)}, \tfrac{5\beta}{\widehat{\pi}_\infty(\phi_o = h)}\right) \leftrightarrow \text{ top l.l. at height } H(L)+1\,.$$

We will henceforth write $H + 1$ instead of h to be consistent with the rest of the paper.

5.1. **Preliminary lemmas.** We begin with preliminary bounds on the level lines. When L is too small, we can rule out the existence of $\mathfrak{L}_0$ using the crude estimate of Lemma 2.7. When L too big, we can guarantee the existence of $\mathfrak{L}_0$ by the stitching procedure of [29], and ensure $\mathfrak{L}_0$ contains a large square. For L in between, we cannot a-priori say anything about the existence of $\mathfrak{L}_0$ (indeed refining this is the whole goal of this section), but we can still ensure that if $\mathfrak{L}_0$ exists it must still contain a large square.

Lemma 5.4. *Recall that λ satisfies $L = \frac{\lambda \beta}{\widehat{\pi}_\infty(\phi_o = H+1)}$. For β large, there exists $\varepsilon_\beta^\square \downarrow 0$ such that the following holds w.h.p.:*

(1) If $\lambda \leq 4(1 - \frac{3}{\beta})$, there are no `large` *$H + 1$ level lines.*
(2) If $\lambda \geq 4(1 + \frac{3}{4\beta})$, there is a `large` *$H + 1$ level line containing a square of length $L(1 - \varepsilon_\beta^\square)$.*

(3) For $\lambda \in [4(1-\frac{3}{\beta}), 4(1+\frac{3}{4\beta})]$, if a `large` *$H+1$ level line exists, then it must contain a square of length $L(1-\varepsilon_\beta^\square)$.*

Proof. For any loop $\mathfrak{L}$, clearly $|\operatorname{Int}(\mathfrak{L})| \le L^2$. So, we have that if $|\mathfrak{L}| \ge \lambda L(1-\frac{3}{\beta})^{-1}$, then Lemma 2.7 gives

$$\pi_V^0(\mathcal{C}_{\mathfrak{L},H+1}) \le \exp(-(\beta+o(1))|\mathfrak{L}| + \lambda\beta L) \le \exp(-(3+o(1))|\mathfrak{L}|)\,. \tag{5.1}$$

Additionally, since $|\operatorname{Int}(\mathfrak{L})| \le |\mathfrak{L}|^2/16$ by the isoperimetric inequality on $\mathbb{Z}^2$, we have that if $|\mathfrak{L}| \le \frac{16L}{\lambda}(1-\frac{3}{\beta})$, then

$$\pi_V^0(\mathcal{C}_{\mathfrak{L},H+1}) \le \exp(-(\beta+o(1))|\mathfrak{L}| + \beta(1-\tfrac{3}{\beta})|\mathfrak{L}|) \le \exp(-(3+o(1))|\mathfrak{L}|)\,.$$

Hence we can rule out the existence of any $\mathfrak{L}$ such that $|\mathfrak{L}| \ge \lambda L(1-\frac{3}{\beta})^{-1}$ or $\log L \le |\mathfrak{L}| \le \frac{16L}{\lambda}(1-\frac{3}{\beta})$ by the standard Peierls argument. Hence, if $\lambda \le 4(1-\frac{3}{\beta})$, there is not gap between the two ranges on $|\mathfrak{L}|$, so there are w.h.p. no `large` $H+1$ level lines.

Next, Item 2 was essentially already proved via the stitching argument of [29, §4.4] (this was already summarized in [14, Clm. 4.19], so we will not provide the details again here). Indeed, for such L, we have $\frac{4\beta+4}{\widehat{\pi}_\infty(\phi_o \ge H)} \le L^{1-o(1)}$, so either using the stitching argument of [29] or just Theorem 2.14 gives that in π_Λ^0 there is w.h.p. an H level line distance at most $L^{1-o(1)}$ from $\partial\Lambda$. Then, the stitching argument says that as long as there exists $\frac{4\beta+2}{\widehat{\pi}_\infty(\phi_o \ge H+1)} \le \ell \le \frac{4\beta+4}{\widehat{\pi}_\infty(\phi_o \ge H+1)}$ such that a box of length $\ell + (\log \ell)^2$ fits inside the H level line, then w.h.p. there exists a `large` $H+1$ level line surrounding $\Lambda_{L(1-\varepsilon_\beta^\square)}$. So we are done as long as $L - L^{1-o(1)} > \frac{4\beta+2}{\widehat{\pi}_\infty(\phi_o \ge H+1)} + (\log \frac{4\beta+2}{\widehat{\pi}_\infty(\phi_o \ge H+1)})^2$. Plugging in the relation between λ and L, this is equivalent to $1 - L^{-o(1)} > \frac{4\beta+2}{\lambda\beta}$, which is satisfied by $\lambda \ge 4(1+\frac{3}{4\beta})$.

Finally to show Item 3, recall from above that we only need to consider $\mathfrak{L}$ such that $|\mathfrak{L}| \ge \frac{16L}{\lambda}(1-\frac{3}{\beta})$. Plugging this and the bound $\lambda \le 4(1+\frac{3}{4\beta})$ into Lemma 2.7 gives that for some choice of ε_β, we can assume that $|\operatorname{Int}(\mathfrak{L})| \ge L^2(1-\varepsilon_\beta)$. The fact that $\mathfrak{L}$ now contains an appropriately sized square follows deterministically by [13, Lem. 2.6]. ■

5.2. **No large $H+1$ level lines below the window.** In this subsection we prove Proposition 5.1. By Lemma 5.4, we can assume additionally that $\lambda \ge 4(1-\varepsilon_\beta)$, and that we can already rule out level lines $\mathfrak{L}$ which do not contain a macroscopic square. This simplified geometry will allows us to study the law of disagreement polymers containing $\mathfrak{L}$ using cluster expansion. Also important is that a level line containing a square is an increasing property. Hence, leveraging monotonicity, we can use a domain enlargement trick to handle interactions between $\mathfrak{L}$ and the domain boundary.

Let $\mathcal{E}_\square$ be the event that there is an $H+1$ level line $\mathfrak{L}$ containing a square $\square$ of side length $L(1-\varepsilon_\beta^\square)$. Let Λ' be the square of side length $5L$, concentric with Λ. By FKG (and since 0 is the minimum value of the $\mathbb{Z}$GFF model with a floor at 0), we have $\pi_\Lambda^0 \preceq \pi_{\Lambda';\Lambda}^H$ where the latter is the $\mathbb{Z}$GFF model on Λ' with boundary conditions H with a floor at 0 only on sites in Λ. In particular since $\mathcal{E}_\square$ is increasing, we have $\pi_\Lambda^0(\mathcal{E}_\square) \le \pi_{\Lambda';\Lambda}^H(\mathcal{E}_\square)$. Abusing notation, we will say that $\mathfrak{L} \in \mathcal{E}_\square$ if $\mathfrak{L}$ is an $H+1$ level line containing a square of side length $L(1-\varepsilon_\beta^\square)$. We will also write $\gamma \in \mathcal{E}_\square$ if γ is a disagreement polymer such that there exists $\mathfrak{L} \in \mathcal{E}_\square$ and $\mathfrak{L}$ is a subset of the bonds of γ. Now fix a $\gamma \in \mathcal{E}_\square$. In order to show that $\pi_{\Lambda';\Lambda}^H(\mathcal{E}_\square) = o(1)$, we need a precise control over the law of γ. We begin with a claim which rules out a bad set of γ not conducive to an application of Corollary 4.4.

Claim 5.5. *Let $\mathcal{G}$ denote the set of γ such that $\mathscr{E}_\beta(\gamma) \le 4.1L$. Then, under $\pi_{\Lambda';\Lambda}^H$, the event $\gamma \in \mathcal{G} \cap \mathcal{E}_\square$ implies that $|\gamma| \le 4.1L$, $h_0 = H+1$, and $h_1 = H$. Moreover, $\pi_{\Lambda';\Lambda}^H(\gamma \in \mathcal{E}_\square \setminus \mathcal{G}) = o(1)$.*

Proof. By FKG and the decorrelation estimate of Eq. (3.24), a standard computation gives a lower bound on the floor event:

$$\widehat{\pi}^H_{\Lambda';\Lambda}(\phi_x \geq 0,\, \forall x \in \Lambda) \geq \prod_{x\in\Lambda} \widehat{\pi}^H_{\Lambda';\Lambda}(\phi_x \geq 0) \geq \tfrac{1}{2}\exp\big(-|\Lambda|\widehat{\pi}^H_\infty(\phi_x < 0)\big) \geq \tfrac{1}{2}e^{-4(1+\varepsilon_\beta)\beta L}\,,$$

where the last line uses the fact that $\widehat{\pi}^H_\infty(\phi_x<0) = \widehat{\pi}^0_\infty(\phi_x \geq h) = (1+o(1))\frac{\lambda\beta}{L}$ and the bound on λ. Hence, it suffices to show that in the no-floor measure, we have $\widehat{\pi}^H_{\Lambda'}(\gamma \in \mathcal{E}_\square \setminus \mathcal{G}) = o(e^{-4(1+\varepsilon_\beta)\beta L})$. But here we can use Proposition 2.4. In particular, since $\mathfrak{I}_{\Lambda'}(\gamma) \leq \varepsilon_\beta|\gamma|$, we can conclude by a standard Peierls argument mapping γ to the disagreement polymer with $4L$ bonds forming the boundary of a square, that $\widehat{\pi}^H_{\Lambda'}(\exists\gamma,\, \mathscr{E}_\beta(\gamma) \geq 4.1L) \leq e^{-4.1\beta L}$.

Finally, it is immediate that $|\gamma| > 4.1L$ implies $\mathscr{E}_\beta(\gamma) > 4.1L$. Moreover, if $h_0 \neq H+1$ or if $h_1 \neq H$, the condition that $\gamma \in \mathcal{E}_\square$ implies that either there is a height gradient of ≥ 2 along $4(1-\varepsilon_\beta)L$ of the bonds in γ, or there are at least $8(1-\varepsilon_\beta)L$ bonds in γ, both of which imply that $\mathscr{E}_\beta(\gamma) > 4.1L$. ∎

For $\gamma \in \mathcal{E}_\square$, using the bound on λ and the large deviation ratios in Theorem 2.9, we also have $|D_1 \cap \Lambda| \leq 4\varepsilon^\square_\beta L^2$. Hence, for all $\gamma \in \mathcal{G} \cap \mathcal{E}_\square$, the conditions of Corollary 4.4 are satisfied, and we have

$$\begin{aligned}\pi^H_{\Lambda';\Lambda}(\mathcal{C}_{\gamma,H+1}) &\leq \exp\Big(-\mathscr{E}^*_\beta(\gamma) + \mathfrak{I}_{\Lambda'}(\gamma) + \tfrac{\rho_0\lambda\beta}{L}|\Lambda \cap D_0| + \xi_1\sum_{i\geq 2}|D_i \cap F| + O(L^{1/2+o(1)})\Big)\\ &=: \mathfrak{p}_{\Lambda';\Lambda}(\gamma)e^{O(L^{1/2+o(1)})}\,. \end{aligned} \tag{5.2}$$

With Lemma 5.4 and Claim 5.5, our goal now is to show that

$$\sum_{\gamma\in\mathcal{G}\cap\mathcal{E}_\square} \mathfrak{p}_{\Lambda';\Lambda}(\gamma)e^{CL^{1/2+o(1)}} = o(1), \qquad \text{or equivalently} \qquad \sum_{\gamma\in\mathcal{G}\cap\mathcal{E}_\square} \mathfrak{p}_{\Lambda';\Lambda}(\gamma) = o(e^{-CL^{1/2+o(1)}})\,.$$

We begin by showing that the term $\xi_1\sum_{i\geq 2}|D_i \cap F|$ is negligible. In fact, the proof is essentially the same as in Proposition 4.1, only this time we are in a macroscopic setting and need to ensure that we do not violate the additional condition $\mathcal{E}_\square$. For completeness we provide the details this time, though when similar estimates are needed in future lemmas we will simply refer back to the proof here or in Proposition 4.1 for details.

To start, it is easy to show a crude bound on the maximum size of a component $\mathcal{D} \in \mathfrak{D}(\gamma)$, say by $|\partial\mathcal{D}| \leq L/5$. Let u, v be the two cut-points delineating $\mathcal{D}$. Then, $|\partial\mathcal{D}| \geq 2\|u-v\|_1$ since $\|u-v\|_1$ is the length of the shortest path from u to v, and $\partial\mathcal{D}$ consists of two such paths. Observe that u, v must also lie on the level line $\mathfrak{L}$ which γ contains. Since $\mathfrak{L} \in \mathcal{E}_\square$, the minimal length path between all of its cut-points is necessarily $\geq 4L(1-\varepsilon^\square_\beta)$. Hence, we have $|\gamma| \geq 4L(1-\varepsilon^\square_\beta) + \sum_{\mathcal{D}\in\mathfrak{D}(\gamma)}|\partial\mathcal{D}|$, whence the assumption $|\gamma| \leq 4.1L$ immediately implies the bound

$$\max_{\mathcal{D}\in\mathfrak{D}(\gamma)} |\partial\mathcal{D}| \leq L/5\,. \tag{5.3}$$

We next use a Peierls argument to show that for some constant $c > 0$,

$$\sum_{\substack{\gamma\in\mathcal{G}\cap\mathcal{E}_\square\\ \max_{\mathcal{D}\in\mathfrak{D}(\gamma)}|\partial\mathcal{D}| > \log L}} \mathfrak{p}_{\Lambda';\Lambda}(\gamma) \leq e^{-c\beta\log L} \sum_{\gamma\in\mathcal{G}\cap\mathcal{E}_\square} \mathfrak{p}_{\Lambda';\Lambda}\,. \tag{5.4}$$

Again we can fix b and consider the map T_b on γ which replaces $\mathcal{D}_b$ by a (arbitrarily chosen, say the outermost such) minimal length path from u to v that remains within $\mathcal{D}_b$. As before, we have $\mathscr{E}^*_\beta(T_b(\gamma)) \leq \mathscr{E}^*_\beta(\gamma) - \frac{\beta}{2}|\partial\mathcal{D}_b|$. Hence, by the decay properties of Φ and the bound on λ, we therefore have after accounting for the area terms that

$$\mathfrak{p}_{\Lambda';\Lambda}(\gamma) \leq \mathfrak{p}_{\Lambda';\Lambda}(T_b(\gamma))e^{\frac{\beta-C}{2}|\partial\mathcal{D}_b| + \frac{4(1+\varepsilon_\beta)\beta}{L}|\mathcal{D}_b|}\,.$$

By isoperimetry and the crude bound in Eq. (5.3), we have $|\mathcal{D}_b| \le |\partial \mathcal{D}_b|^2/16 \le |\partial \mathcal{D}_b| L/80$, so that the above display implies that for some $c > 0$,

$$\mathfrak{p}_{\Lambda';\Lambda}(\gamma) \le \mathfrak{p}_{\Lambda';\Lambda}(T_b(\gamma)) e^{-c\beta|\partial \mathcal{D}_b|}\,.$$

Moreover, we claim that the resulting disagreement polymer satisfies $T_b(\gamma) \in \gamma \in \mathcal{G} \cap \mathcal{E}_\square$. The bound on $\mathscr{E}_\beta(T_b(\gamma))$ is easily satisfied since the map T_b only reduces the length of γ. Showing that $T_b(\gamma) \in \mathcal{E}_\square$ amounts to showing that we can select a minimal length path from u to v staying within $\mathcal{D}_b$ that stays outside of the square, $\square$, which $\mathfrak{L}$ contains. Suppose for contradiction that there is no such path. Then, take a shortest path P, and let w, z be two points of P on $\partial\square$ which mark an excursion into $\square$ in the sense that in between w, z, P lies inside $\square$. Recalling the outer envelope in Definition 2.15, observe that the region sandwiched between P and the arc of $\mathsf{OE}(\gamma)$ between u, v is contained in $\mathcal{D}_b$. Since $\mathfrak{L}$ lies outside of $\square$, so does $\mathsf{OE}(\gamma)$. Since the arc of P from w to z is inside $\square$, the line segment $\overline{wz}$ (or the L shaped path from w to z, if w, z are on two different sides of the square $\square$) is contained in this sandwiched region, and is therefore in $\mathcal{D}_b$. Hence, the path P' which replaces the arc of P between w and z by $\overline{wz}$ is in $\mathcal{D}_b$. Since $|P'| < |P|$, this is a contradiction.

Finally, as before, the number of preimages γ under T_b for a given disagreement polymer γ' is at most $k^2 C^k$. The exact same computation in Eq. (4.5) now concludes the proof of Eq. (5.4).

With Eq. (5.4) proven, we can now restrict our attention to γ which additionally satisfies $\max_{\mathcal{D}\in\mathfrak{D}(\gamma)} |\partial\mathcal{D}| \le \log L$. Together with the isoperimetric inequality and the restriction that $|\gamma| \le 4.1L$, this in particular implies that $\mathfrak{x}_1 \sum_{i\ge 2} |D_i \cap F| \le O(\log L)$. Hence, our updated goal is now to show that

$$\sum_{\substack{\gamma\in\mathcal{G}\cap\mathcal{E}_\square \\ \max_{\mathcal{D}\in\mathfrak{D}(\gamma)}|\partial\mathcal{D}|\le \log L}} e^{-\mathscr{E}^*_\beta(\gamma) + \mathfrak{J}_{\Lambda'}(\gamma) + \frac{\rho_0 \lambda\beta}{L}|\Lambda\cap D_0|} =: \sum_{\substack{\gamma\in\mathcal{G}\cap\mathcal{E}_\square \\ \max_{\mathcal{D}\in\mathfrak{D}(\gamma)}|\partial\mathcal{D}|\le \log L}} \mathfrak{q}_{\Lambda';\Lambda}(\gamma) \le o(e^{-CL^{1/2+o(1)}})\,.$$

We next continue to rule out a set of unlikely γ by Peierls type arguments.

Definition 5.6. For any two cut-points $v, v' \in \gamma$, let $d_\gamma(v, v')$ denote the number of bonds in between v and v'. More precisely, removing v and v' splits γ into two connected components, and $d_\gamma(v, v')$ is the number of bonds in the smaller of the two components.

Definition 5.7. We say that γ has a button hole if there are two cut-points $v, v' \in \gamma$, not on opposite sides of the square in the event $\mathcal{E}_\square$, such that $d_\gamma(v, v') \ge \log L$ while $\|v - v'\|_1 \le \frac{1}{2} d_\gamma(v, v')$.

Lemma 5.8. *Write $\gamma \in \mathcal{B}$ to mean that γ has a button hole. Then,*

$$\sum_{\substack{\gamma\in\mathcal{G}\cap\mathcal{E}_\square \\ \max_{\mathcal{D}\in\mathfrak{D}(\gamma)}|\partial\mathcal{D}|\le \log L \\ \gamma\in\mathcal{B}}} \mathfrak{q}_{\Lambda';\Lambda}(\gamma) \le e^{-c\beta\log L} \sum_{\substack{\gamma\in\mathcal{G}\cap\mathcal{E}_\square \\ \max_{\mathcal{D}\in\mathfrak{D}(\gamma)}|\partial\mathcal{D}|\le \log L}} \mathfrak{q}_{\Lambda';\Lambda}(\gamma) \tag{5.5}$$

Proof. Let γ be a disagreement polymer in the left sum, and let v, v' be two cut-points resulting in a button hole (if there are multiple candidates for v, v', choose arbitrarily, say the minimal pair under lexicographic ordering). Let $T(\gamma)$ be the result of replacing the portion between v, v' by a shortest length path from v to v' with an edge disagreement of one along the path. This shortest length path can always be chosen such that $T(\gamma)$ still contains a square of side length $L(1-\varepsilon'_\beta)$ (we can assume that v, v' are not on opposite sides of the square since this would violate the condition $|\gamma| \le 4.1L$). By the button hole criterion, we have $\mathscr{E}_\beta(T(\gamma)) \le \mathscr{E}_\beta(\gamma) - \frac{\beta}{2} d_\gamma(v, v')$. Since we are only possibly removing some finite clusters D_i, the same inequality holds with $\mathscr{E}^*_\beta$ instead of $\mathscr{E}_\beta$. The change in the interaction term is also at most $\varepsilon_\beta d_\gamma(v, v')$ by the decay properties of Φ. Moreover, $\big||\Lambda \cap D_0(\gamma)| - |\Lambda \cap D_0(T(\gamma))|\big| \le \min(d_\gamma(v, v')^2, \varepsilon'_\beta L^2)$, which is always negligible compared to $\frac{\beta}{2} d_\gamma(v, v')$. At the same time, the number of preimages γ under the map T is at most $L^2 C^{d_\gamma(v,v')}$ for a constant C, given by the L^2 choices for the two cut-points v, v' and the $C^{d_\gamma(v,v')}$ choices of how to reconstruct the portion of γ between v and v'. Hence, we can conclude by a Peierls map argument completely analogous to the computation in Eq. (4.5). That is, enumerate over the possible preimages γ to a given image γ' such that $d_\gamma(v, v')$ is equal to a fixed value $k \geq \log L$, and sum over k. ∎

Given Eq. (5.5), it now suffices to show that

$$\sum_{\substack{\gamma \in \mathcal{G} \cap \mathcal{E}_\square \\ \max_{\mathcal{D} \in \mathfrak{D}(\gamma)} |\partial \mathcal{D}| \leq \log L \\ \gamma \notin \mathcal{B}}} \mathfrak{q}_{\Lambda';\Lambda}(\gamma) \leq o(e^{-CL^{1/2+o(1)}})\,. \tag{5.6}$$

We finally note that the bound $|\gamma| \leq 4.1L$ and the domain enlargement trick to move $\partial\Lambda'$ far away implies that we can replace $\mathfrak{I}_{\Lambda'}(\gamma)$ by $\mathfrak{I}_{\mathbb{Z}^2}(\gamma)$ at the cost of $o(1)$ in the exponent. We denote the weights with this modification by $\mathfrak{q}_{\mathbb{Z}^2;\Lambda}(\gamma)$ (This step could have been done earlier, but is not needed until now.)

The benefit of ruling out button-holes is that we can now control the interaction between different parts of γ in the following sense. Suppose $\{v_1, \dots v_s\}$ is a set of ordered cut-points on γ. Then we can naturally decompose $\gamma = \gamma_1 \circ \gamma_2 \dots \circ \gamma_s$ such that γ_i is the portion of γ between v_i and v_{i+1}, with the convention that $v_{s+1} = v_1$.

Lemma 5.9. *There exists a constant $C > 0$ such that for any γ with no button holes satisfying $\max_{\mathcal{D} \in \mathfrak{D}(\gamma)} |\partial \mathcal{D}| \leq \log L$, and any choice of ordered cut-points $\{v_1, \dots v_s\}$, we have*

$$|\mathfrak{I}_{\mathbb{Z}^2}(\gamma) - \sum_{i=1}^s \mathfrak{I}_{\mathbb{Z}^2}(\gamma_i)| \leq sC \log L\,.$$

Proof. By definition of the $\mathfrak{I}_{\mathbb{Z}^2}$ and the decay properties of Φ, we have

$$|\mathfrak{I}_{\mathbb{Z}^2}(\gamma) - \sum_{i=1}^s \mathfrak{I}_{\mathbb{Z}^2}(\gamma_i)| \leq \sum_{i \neq j} \sum_{\substack{\mathsf{W} \cap \Delta_{\gamma_i} \neq \emptyset \\ \mathsf{W} \cap \Delta_{\gamma_j} \neq \emptyset}} \Phi(\mathsf{W}) \leq \sum_{i=1}^s \sum_{\substack{x \in \gamma_i \\ y \in \gamma \setminus \gamma_i}} e^{-c\|x-y\|_1}\,.$$

Now fix i. Let W be the set of points $x \in \gamma_i$ such that there exists $z \in \gamma \setminus \gamma_i$ with $\|z - x\|_1 \leq \log L$. By the condition that $\max_{\mathcal{D} \in \mathfrak{D}(\gamma)} |\partial \mathcal{D}| \leq \log L$, we can find cut-points $z^{\mathsf{cut}}, x^{\mathsf{cut}}$ such that $\|z^{\mathsf{cut}} - x^{\mathsf{cut}}\|_1 \leq 3\log L$. By the button-hole criterion, this implies that $d_\gamma(z^{\mathsf{cut}}, x^{\mathsf{cut}}) \leq 6 \log L$. Now assume that the v_i are ordered clockwise. Starting from v_i and going clockwise, let x_i^{cut} be the first cut-point such that $d_\gamma(v_i, x_i^{\mathsf{cut}}) \geq 6\log L$. Let x_{i+1}^{cut} be defined analogously by starting from v_{i+1} and going counter-clockwise. By the above argument, every $x \in W$ must either be in between v_i and x_i^{cut}, or be in between x_{i+1}^{cut} and v_{i+1}. In particular, $|W| \leq 12 \log L$. We can bound

$$\sum_{\substack{x \in W \\ y \in \gamma \setminus \gamma_i}} e^{-c\|x-y\|_1} \leq |W| \sum_{y \in \mathbb{Z}^2} e^{-c\|x-y\|_1} \leq C \log L\,,$$

where the point x in the second sum is arbitrary. At the same time, we have

$$\sum_{\substack{x \in \gamma_i \setminus W \\ y \in \gamma \setminus \gamma_i}} e^{-c\|x-y\|_1} \leq |\gamma_i| \sum_{\substack{y \in \mathbb{Z}^2 \\ d(x,y) \geq \log L}} e^{-c\|x-y\|_1} \leq e^{-c \log L}\,.$$

Combining the above two displays and summing over i completes the proof. ∎

For brevity, to prove Eq. (5.6), let $\widetilde{\mathcal{E}}_\square$ be the set of all $\gamma \in \mathcal{G} \cap \mathcal{E}_\square$ such that $\max_{\mathcal{D} \in \mathfrak{D}(\gamma)} |\partial \mathcal{D}| \leq \log L$, and γ has no button-holes. We will enumerate over $\gamma \in \widetilde{\mathcal{E}}_\square$ by marking a set of special anchor points $\underline{v}$. For a sequence of points $\underline{v} = \{v_1, \dots v_s\}$ in the dual graph, define the set $\mathcal{P}(\underline{v})$ as disagreement polymers γ such that $\gamma \in \widetilde{\mathcal{E}}_\square$, $\mathsf{OE}(\gamma)$ contains all the points v_i in order from 1 to s, and each v_{i+1} is the next cut-point in $\mathsf{OE}(\gamma)$ after v_i which has $\|\cdot\|_1$ distance at least $L^{1/2}$ from v_i. In particular, we assume from the condition that the maximum size of any $|\partial\mathcal{D}| \le \log L$ that $L^{1/2} \le \|v_{i+1}-v_i\|_1 \le L^{1/2}+\log L$. We may further assume that $\|v_s-v_1\|_1 \le L^{1/2}$. Since $|\gamma| \le 4.1L$, this implies that $s \le 4.1L^{1/2}$ in order for $\mathcal{P}(\underline{v})$ not to be empty.

Now fix $\underline{v}$. We will bound the sum Eq. (5.6) first over all $\gamma \in \mathcal{P}(\underline{v})$. For convenience of notation, let $v_{s+1} := v_1$. Define also $L_{\underline{v}}$ as the linear interpolation of the points $\underline{v}$. Recall the definitions of τ_β and $\mathcal{P}_{\mathbb{Z}^2}(v_i, v_{i+1})$ from Section 2.2. Using Lemma 5.9 and the relationship between τ_β and partition function given by [14, Prop. 3.14 (ii) and Prop. 4.14], we have a bound on the sum of the weights without the area term:

$$\begin{aligned}\sum_{\gamma\in\mathcal{P}(\underline{v})} \exp(-\mathscr{E}^*_\beta(\gamma)+\mathfrak{I}_{\mathbb{Z}^2}(\gamma)) &\le e^{sC\log L}\prod_{i=1}^{s}\sum_{\gamma_i\in\mathcal{P}_{\mathbb{Z}^2}(v_i,v_{i+1})}\exp(-\mathscr{E}^*_\beta(\gamma_i)+\mathfrak{I}_{\mathbb{Z}^2}(\gamma_i))\\ &\le e^{sC\log L}\prod_{i=1}^{s}\exp(-\tau_\beta(v_{i+1}-v_i))\\ &= \exp(-\int_{L_{\underline{v}}}\tau_\beta(\theta_s)ds + sC\log L)\,.\end{aligned}$$

Define $K_{\underline{v}}$ as the convex hull of $\underline{v}$. By convexity of the surface tension, we have

$$\int_{L_{\underline{v}}}\tau_\beta(\theta_s)ds \ge \int_{\partial K_{\underline{v}}}\tau_\beta(\theta_s)ds \ge \int_{\partial(K_{\underline{v}}\cap\Lambda)}\tau_\beta(\theta_s)ds\,.$$

To handle the area term, we observe that $|\Lambda\cap D_0| \le |K_{\underline{v}}\cap\Lambda| + sL \le |K_{\underline{v}}\cap\Lambda| + 4.1L^{3/2}$. Hence, we have

$$\begin{aligned}&\sum_{\gamma\in\mathcal{P}(\underline{v})}\exp(-\mathscr{E}^*_\beta(\gamma)+\mathfrak{I}_{\mathbb{Z}^2}(\gamma)+\tfrac{\rho_0\lambda\beta}{L}|\Lambda\cap D_0|)\\ &\qquad\le \exp(-\int_{\partial(K_{\underline{v}}\cap\Lambda)}\tau_\beta(\theta_s)ds + \tfrac{\rho_0\lambda\beta}{L}|K_{\underline{v}}\cap\Lambda| + 4.1\rho_0\lambda\beta L^{1/2} + 4.1CL^{1/2}\log L)\\ &\qquad\le \exp(-\int_{\partial(K_{\underline{v}}\cap\Lambda)}\tau_\beta(\theta_s)ds + \tfrac{\rho_0\lambda\beta}{L}|K_{\underline{v}}\cap\Lambda| + O(L^{1/2}\log L))\,.\end{aligned}$$

By rescaling to the unit square, we can use Lemma 2.12 to obtain that

$$-\int_{\partial(K_{\underline{v}}\cap\Lambda)}\tau_\beta(\theta_s)ds + \tfrac{\rho_0\lambda\beta}{L}|K_{\underline{v}}\cap\Lambda| \le -.9\beta(\lambda_*-\rho_0\lambda)L\,,$$

whence we have a combined upper bound of $\exp(-.9\beta(\lambda_*-\rho_0\lambda)L + O(L^{1/2}\log L))$ for the total weight of all the $\gamma\in\mathcal{P}(\underline{v})$ contributing to the sum in Eq. (5.6), for a fixed $\underline{v}$. Since the number of possible $\underline{v}$ is bounded by $|\Lambda'|^s \le O(L^{8.2L^{1/2}}) = O(\exp(8.2L^{1/2}\log L))$, we conclude by summing over all choices of $\underline{v}$ that Eq. (5.6) will hold as long as $(\lambda_*-\rho_0\lambda)L \gg L^{1/2+o(1)}$. Finally, we conclude by checking that for $L \le L_*^{(H+1)} - (L_*^{(H+1)})^{1/2+o(1)}$, we have

$$(\tfrac{\lambda_*}{\rho_0}-\lambda)L \ge \frac{\widehat{\pi}_\infty(\phi_o = H+1)}{\beta}(L_*^{(H+1)} - 1 - (L_*^{(H+1)} - (L_*^{(H+1)})^{1/2+o(1)}))L = O(L^{1/2+o(1)})\,,$$

concluding the proof of Proposition 5.1.

5.3. **Existence of a large $H+1$ level line above the window.** The goal of this subsection is to prove Proposition 5.2. By Lemma 5.4 we can additionally assume that $\lambda < 4(1+\frac{3}{4\beta})$, and we know that w.h.p. only `large` $H+1$ level lines containing a square $\Lambda_{L(1-\varepsilon_\beta^\square)}$ can exist. We follow the strategy of [13, Section 4.3].

We now proceed via a proof by contradiction. Let $\mathfrak{S}_0$ be the event that there are no `large` $H+1$ level lines with area larger than $(1-\varepsilon_\beta^\square)^2L^2$. We wish to show that $\pi_\Lambda^0(\mathfrak{S}_0)=o(1)$. Recall the shapes defined in Definition 2.11 and the notational comment of Remark 2.13. Recall also that $\ell_n^* = \ell_n^*(L) := \frac{\mathsf{w}_1(\tau_\beta)N_n}{2L\rho_n}$. By Theorem 2.14 and the bounds on ρ_n in Proposition 3.1, w.h.p. there exists an H level line $\mathfrak{L}_1$ containing $L(1-L^{-2/3-o(1)})\mathscr{L}(\ell_0^*)$. We can reveal it and by monotonicity it suffices to bound $\pi_{\mathrm{Int}(\mathfrak{L}_1)}^H(\mathfrak{S}_0)$, as $\mathfrak{S}_0$ is decreasing.

Now define $\mathfrak{S}$ as having no `large` disagreement polymers. The goal is to use Corollary 4.4 to show that $\pi_{\mathrm{Int}(\mathfrak{L}_1)}^H(\mathfrak{S})=o(1)$ by summing over a set of nice γ. Since `large` level lines where the interior is lower than the exterior can always be rule out via the standard Peierls argument, this will imply that $\pi_{\mathrm{Int}(\mathfrak{L}_1)}^H(\mathfrak{S}_0)=o(1)$. Define the shape $\mathcal{K}=\partial((L(1-L^{-1/2+o(1)}))\mathscr{L}(\ell_0^*))$. Define also "cigar shapes" as follows: for any two points $u=(u_1,u_2), v=(v_1,v_2)$, let θ be the angle of $\overline{uv}$ with the x-axis, and $M_{u,v}=v_1-u_1$. Assume for simplicity that $\theta\in[0,\pi/4]$. Define $\mathcal{C}_{u,v}^{\pm}$ by

$$\mathcal{C}_{u,v}^{\pm}(t)=\tan(\theta_{u,v})(t-u_1)+\left(\frac{(t-u_1)(M_{u,v}-t+u_1)}{M_{u,v}}\right)^{1/2}(\log L)^2\,. \tag{5.7}$$

Define the cigar shape $\mathscr{C}=\mathscr{C}(u,v)$ as the region in between the curves $\mathcal{C}_{u,v}^+$ and $\mathcal{C}_{u,v}^-$. For u,v with an angle not in $[0,\pi/4]$, we can define $\mathscr{C}(u,v)$ via reflection across x and y axes.

Let $\{v_i\}_{i=1}^s$ be a (arbitrarily chosen) set of points on $\mathcal{K}$ such that the distance between any two points is between $3L/s$ and $5L/s$. For simplicity of notation, define $v_{s+1}=v_1$. Our nice set of γ, denoted $\mathcal{G}$, will be all γ contained in $\bigcup_{i=1}^s\mathscr{C}(v_i,v_{i+1})$ with length $|\gamma|\le 10L$ such that if D_0 is the region containing the origin, then $h_0=H$, if D_1 as the region containing the boundary vertices $\partial\,\mathrm{Int}(\mathfrak{L}_1)$, we have $h_1=H-1$, and $\max_{\mathcal{D}\in\mathfrak{D}(\gamma)}|\partial\mathcal{D}|\le\log L$. By construction, this implies that $|D_1^\circ\cap F|\le\varepsilon_\beta L^2$. All $\gamma\in\mathcal{G}$ thus satisfy the conditions of Corollary 4.4, and together with the bound on $|\partial\mathcal{D}|$ we have

$$\pi_{\mathrm{Int}(\mathfrak{L}_1)}^H(\mathcal{C}_{\gamma,H+1})/\pi_{\mathrm{Int}(\mathfrak{L}_1)}^H(\mathfrak{S})\ge\exp(-\mathscr{E}_\beta^*(\gamma)+\mathfrak{I}_{\mathrm{Int}(\mathfrak{L}_1)}(\gamma)+\tfrac{\rho_0\lambda\beta}{L}|D_0|+L^{1/2+o(1)})\,.$$

Since $|D_0|\ge|\mathrm{Int}(\mathcal{K})|-s(5L/s)^2=|\mathrm{Int}(\mathcal{K})|-5L^2/s$, so we can replace $|D_0|$ with $|\mathrm{Int}(\mathcal{K})|$ at a cost of L/s as error. Moreover, γ stays distance $L^{1/2+o(1)}$ away from $\mathfrak{L}_1$, so we can replace $\mathfrak{I}_{\mathrm{Int}(\mathfrak{L}_1)}(\gamma)$ with $\mathfrak{I}_{\mathbb{Z}^2}(\gamma)$. Finally, letting γ_i denote the portion of γ between v_i and v_{i+1}, we can write $\mathscr{E}_\beta^*(\gamma)=\sum_i\mathscr{E}_\beta^*(\gamma_i)$, and by construction of the cigar shapes we have $|\mathfrak{I}_{\mathbb{Z}^2}(\gamma)-\sum_{i=1}^s\mathfrak{I}_{\mathbb{Z}^2}(\gamma_i)|=O(s(\log L)^2)$. Hence, we have

$$\pi_{\mathrm{Int}(\mathfrak{L}_1)}^H(\mathcal{C}_{\gamma,H+1})/\pi_{\mathrm{Int}(\mathfrak{L}_1)}^H(\mathfrak{S})\ge\exp(\tfrac{\rho_0\lambda\beta}{L}|\mathrm{Int}(\mathcal{K})|+L^{1/2+o(1)}+O(L/s)+O(s(\log L)^2))\prod_i e^{-\mathscr{E}_\beta^*(\gamma_i)+\mathfrak{I}_{\mathbb{Z}^2}(\gamma_i)}\,.$$

We study the product first, which captures the weight of the polymer γ_i with no area tilt terms. We claim that

$$\sum_{\substack{\gamma_i:v_i\mapsto v_{i+1}\\ \gamma_i\in\mathscr{C}(v_i,v_{i+1})\\ |\gamma_i|\le 10L/s\\ \max_{\mathcal{D}\in\mathfrak{D}(\gamma_i)}|\partial\mathcal{D}|\le\log L}} e^{-\mathscr{E}_\beta^*(\gamma_i)+\mathfrak{I}_{\mathbb{Z}^2}(\gamma_i)}\ge e^{-\tau_\beta(v_{i+1}-v_i)+O(1)}\,.$$

Indeed, if the sum did not have any restrictions besides $\gamma_i:v_i\mapsto v_{i+1}$, this holds by [14, Prop. 3.14 (ii) and Prop. 4.14]. We then note that each of the additional restrictions only affect the sum by a multiplicative factor of $1-o(1)$. Indeed, we can control the effect of restricting to the cigar shape by [14, Lem. 3.15], control the length by a simple Peierls argument mapping γ_i to a minimal length path from v_i to v_{i+1}, and control the size of $\mathcal{D}\in\mathfrak{D}(\gamma_i)$ by the same argument used in the proof of Proposition 4.1 (only simpler since there are no area terms here).

Now we can take $s = L^{1/2}$ to absorb all the error terms together, define $\overline{\mathcal{K}}$ as the linear interpolation of the points $\{v_i\}$, and combine the above two displays to obtain

$$\sum_{\gamma\in\mathcal{G}} \frac{\pi^H_{\mathrm{Int}(\mathfrak{L}_1)}(\mathcal{C}_{\gamma,H+1})}{\pi^H_{\mathrm{Int}(\mathfrak{L}_1)}(\mathfrak{S})} \geq \exp\Big(- \int_{\overline{\mathcal{K}}} \tau_\beta(\theta_s)ds + \tfrac{\rho_0\lambda\beta}{L}|\operatorname{Int}(\mathcal{K})| + L^{1/2+o(1)}\Big)$$
$$\geq \exp\Big(- \int_{\mathcal{K}} \tau_\beta(\theta_s)ds + \tfrac{\rho_0\lambda\beta}{L}|\operatorname{Int}(\mathcal{K})| + L^{1/2+o(1)}\Big)$$
$$= \exp\big(L\mathcal{F}^0_{\rho_0\lambda}((1-L^{-1/2+o(1)})\mathscr{L}(\ell_0^*)) + L^{1/2+o(1)}\big)\,, \tag{5.8}$$

using the convexity of the surface tension for the second inequality.

Since $\sum_{\gamma\in\mathcal{G}} \pi^H_{\mathrm{Int}(\mathfrak{L}_1)} \leq 1$, then if we show that the right side of Eq. (5.8) goes to ∞ as $L\to\infty$, this implies that $\pi^H_{\mathrm{Int}(\mathfrak{L}_1)} = o(1)$. To reference Lemma 2.12, it will be easier to use the notational exchange $\mathscr{L}(\ell_0^*) \equiv \mathscr{L}(\rho_0\lambda)$ via the identification in Remark 2.13. For any $C\in[0,1]$ and any λ, we have $\mathcal{F}_\lambda(C\partial\mathscr{L}(\lambda')) = C\mathcal{F}_\lambda(\partial\mathscr{L}(\lambda)) - C(1-C)\lambda\beta|A(\mathscr{L}(\lambda))|$ by the definition of $\mathcal{F}_\lambda$, so that when $C = 1-L^{-1/2+o(1)}$, we have

$$\mathcal{F}_{\rho_0\lambda}((1-L^{-1/2+o(1)})\partial\mathscr{L}(\rho_0\lambda)) \geq (1-L^{-1/2+o(1)})\mathcal{F}_{\rho_0\lambda}(\partial\mathscr{L}(\rho_0\lambda)) - \beta\rho_0\lambda L^{-1/2+o(1)}\,.$$

From Lemma 2.12, we have that $\mathcal{F}_{\rho_0\lambda}(\partial\mathscr{L}(\rho_0\lambda)) \geq (\rho_0\lambda-\lambda_*)(\beta-1)$. Put altogether, we have

$$L\mathcal{F}_{\rho_0\lambda}((1-L^{-1/2+o(1)})\partial\mathscr{L}(\rho_0\lambda)) \geq L(\rho_0\lambda-\lambda_*)(\beta-1)(1-L^{-1/2+o(1)}) - \beta\rho_0\lambda L^{1/2+o(1)}\,,$$

and we want to ensure that $L\mathcal{F}_{\rho_0\lambda}((1-L^{-1/2+o(1)})\partial\mathscr{L}(\rho_0\lambda)) \gg L^{1/2+o(1)}$. For this, it suffices to take λ such that $\rho_0\lambda-\lambda_* \gg L^{-1/2+o(1)}$. We can check that for $L \geq L_*^{(H+1)} + (L_*^{(H+1)})^{1/2+o(1)}$, we indeed have

$$\rho_0\lambda-\lambda_* \geq \rho_0\frac{\widehat{\pi}_\infty(\phi_o = H+1)}{\beta}(L-L_*^{(H+1)}) \geq O((L_*^{(H+1)})^{1/2+o(1)}/L) = O(L^{-1/2+o(1)})\,.$$

Hence, for such L, we can conclude from Eq. (5.8) that $\pi^H_{\mathrm{Int}(\mathfrak{L}_1)}(\mathfrak{S}) = o(1)$, and this concludes the proof of Proposition 5.2.

6. Limit shape of the top level lines

In this section we prove Theorem 1.1, establishing the global limit shape of the top level lines (even in the critical window). In Section 6.1, we extend Theorem 2.14 to the case where the $H+1$ level line exists, showing one side of the shape theorem, that the level line contains a shape. In Section 6.2, we prove the other side, that the level line is contained in a shape.

Throughout this section, there will be several occasions where we state that a bad event only occurs with probability $o(1)$, and then take a union bound over polynomially many such bad events. This is not an issue, as in fact every $o(1)$ is at most $e^{-c\beta\log L}$ for some constant c, the dominant error arising from Peierls maps which delete level lines of size $\log L$.

6.1. Top level line contains limit shape in the critical window.

The results of Theorem 2.14 show that for any L, w.h.p. in π^0_Λ the $H, H-1,\dots H-m$ level lines contain their respective limit shapes, up to a distance of $o(N_n)$ for the $H+1-n$ level line for $n\geq 1$. As we have shown, for certain L, there is also a `large` $H+1$ level line $\mathfrak{L}_0$, possibly with probability bounded away from both 0 and 1. An immediate corollary of Theorem 2.14 is that for L such that $\pi^0_\Lambda(\mathfrak{L}_0=\emptyset) > \delta_L$, the theorem still holds under the measure $\pi^0_\Lambda(\cdot\mid\mathfrak{L}_0=\emptyset)$.

Here we prove that if there is an $H+1$ level line, then it must also contain its respective limit shape. Note that this easily implies the uniqueness of the `large` $H+1$ level line when it exists, as we can reveal the innermost such level line and rule out any others via Lemma 2.7.

Recall again that $\ell_n^* = \ell_n^*(L) = \frac{\mathsf{w}_1(\tau_\beta)N_n}{2L\rho_n}$. Fix any constant $C > 0$, and define the following shapes

$$\mathfrak{R}_1 := (1 - \frac{N_1^{1/3}e^{3C\sqrt{\log L}}}{L})L\mathscr{L}(\ell_1^*(1 + e^{-\frac{C}{3}\sqrt{\log L}}))$$
$$\mathfrak{R}_0 := (1 - 2\frac{N_0^{1/3}e^{3C\sqrt{\log L}}}{L})L\mathscr{L}(\ell_0^*(1 + e^{-\frac{C}{3}\sqrt{\log L}}))\,.$$

We will show the following proposition.

Proposition 6.1. *Let $\mathcal{E}$ be the event that there exists a* `large` *$H+1$ level line which does not contain $\mathfrak{R}_0$ in its interior. Then, $\pi_\Lambda^0(\mathcal{E}) = o(1)$. Consequently, for every L such that $\pi_\Lambda^0(\mathfrak{L}_0 \neq \emptyset) > \delta_L$, we have $\pi_\Lambda^0(\mathcal{E} \mid \mathfrak{L}_0 \neq \emptyset) = o(1)$.*

The proof of Theorem 2.14 does not immediately extend to studying the $H+1$ level line since it assumes the level line exists w.h.p., and needs to be modified to remove this assumption. The difference is the following: if L is such that w.h.p. there is a unique `large` $H+1$ level line, then we can try to show that w.h.p., this level line exists and contains $\mathfrak{R}_0$, and this is a monotone increasing event. If we don't know whether the $H+1$ level line exists, we can only aim to show that the event $\mathcal{E}$ occurs with probability $o(1)$. However, $\mathcal{E}$ is no longer an monotone event! Hence, we need to be more careful when we make monotonicity arguments. A reader familiar with [14, §4.2] will see that this is a technical modification, and the overall proof strategy remains the same.

Recall that $\mathcal{W}_1(\tau_\beta)$ is the Wulff shape of area 1. For $x \in \Lambda$, define $\mathcal{W}(x,\ell)$ as the rescaled Wulff shape $L\ell\mathcal{W}_1(\tau_\beta)$ centered at x. Let ℓ_x be the largest value of ℓ before $\mathcal{W}(x,\ell)$ reaches within distance $N_0^{1/3}e^{3C\sqrt{\log L}}$ from $\mathfrak{R}_1$. For any x,ℓ, we define the event $\mathcal{E}_{x,\ell}$ to be that there is an $H+1$ level line which contains $\mathcal{W}(x,\ell)$ in its interior, but not $\mathcal{W}'(x,\ell) := (1+L^{-3/4})\mathcal{W}(x,\ell)$. The main claim is the following:

Claim 6.2. *For any domain $\Lambda \supset \mathfrak{R}_1$ and boundary conditions $\eta \geq H$, for all x,ℓ such that $\ell_0^*(1 + e^{-\frac{C}{3}\sqrt{\log L}}) \leq \ell \leq \ell_x$, we have $\pi_\Lambda^\eta(\mathcal{E}_{x,\ell}) = o(1)$.*

Proof. Fix x,ℓ such that $(1 + e^{-\frac{C}{3}\sqrt{\log L}})\ell_0^* \leq \ell < \ell_x$. Fix an angle $\theta \in [0,\pi/4]$ and let A,B on the bottom right quarter of $\mathcal{W}(x,\ell)$ be such that $B_1 - A_1 = N_0^{2/3}e^{C\sqrt{\log L}}$ and the angle of $\overline{AB}$ is θ. (By symmetry, this is equivalent to looking at points on the top half of $\mathcal{W}(x,\ell)$, or fixing instead the vertical distance $B_2 - A_2$ and looking at the right/left halves.) Let R be the rectangle of width $N_0^{2/3}e^{C\sqrt{\log L}}$ and height $2N_0^{2/3}e^{C\sqrt{\log L}}$ with A,B on its sides, such that the distance from A to the bottom of R is $N_0^{1/3}e^{3C\sqrt{\log L}}$.

Let $\partial^1 R, \partial^2 R$ be the arcs of ∂R above and below A,B, respectively. For every vertex $v \in \partial^2 R$, reveal the connected component of sites with height $\leq H-1$ containing it. This reveals an external boundary of sites with height $\geq H$. Note that the standard Peierls argument implies that with probability $1-o(1)$, each connected component revealed has size $\leq \log L$. Similarly, for every vertex $v \in \partial^1 R$ that has not already been revealed, reveal the connected component of sites with height $\leq H$ containing it. Except with probability $o(1)$, the event that there is an $H+1$ level line containing $\mathcal{W}(x,\ell)$ implies that each of these connected components have size $\leq \log L$. This reveals an external boundary of sites with height $\geq H+1$. Piecing these external boundaries together obtains a circuit of sites $\mathcal{C}_*$ distanced at most $\log L$ from ∂R with two marked points A',B' with $d(A,A'), d(B,B') \leq \log L$, such that the heights on $\mathcal{C}$ are $\geq H+1$ on the arc above A',B' and $\geq H$ on the arc below.

Hence, the event $\mathcal{E}_{x,\ell}$ implies that at every angle θ, except with probability $o(1)$, we can find such a circuit $\mathcal{C}_*$. However, it also implies that there is some point $z \in \partial\mathcal{W}'(x,\ell)$ which is not contained by the $H+1$ level line containing $\mathcal{W}(x,\ell)$. That is, on $\mathcal{E}_{x,\ell}$, there is some angle θ such that for any such circuit $\mathcal{C}_*$ as above, the connected component of sites with height $\geq H+1$ containing the arc of $\mathcal{C}_*$ above A', B' does not contain z. Note that for any fixed circuit, this latter event is decreasing. We will next argue that it occurs with probability $o(1)$.

Indeed, given θ, we can reveal the outermost such circuit $\mathcal{C}_*$, which in particular does not reveal anything about its interior. Since we want to upper bound the probability of a decreasing event, we can now use monotonicity to lower the heights along the top and bottom arcs of $\mathcal{C}_*$ to be exactly $H+1, H$ respectively. By Domain Markov, the law of the heights interior to $\mathcal{C}_*$ is now a Discrete Gaussian with $H+1, H$ boundary conditions. To satisfy the inputs of [14, Thm. 4.9], we can then further lower some of the $H+1$ heights to H, as done in step 5 of the proof of [14, Lem. 4.8], to obtain the regularity needed in $\mathcal{C}_*$ around points where boundary conditions change. The desired bound on the size of the $\geq H+1$ component containing the $H+1$ boundary conditions now follows from [14, Thm. 4.9][9] together with the computation preceding [14, Eq. 4.34]. Finally, taking a union bound over all possible angles θ concludes the proof of the claim. ∎

Proof of Proposition 6.1. By Theorem 2.14 and Lemma 5.4, the following hold w.h.p. under π_Λ^0:

(1) There is an H level line containing $\mathfrak{R}_1$ in its interior, and
(2) If a `large` $H+1$ level line exists, it also contains $\mathcal{W}(o, \ell_0^*(1+e^{-\frac{C}{3}\sqrt{\log L}}))$, where o is the center of Λ.

We want to bound the probability that a `large` $H+1$ level line exists, contains $\mathcal{W}(o, \ell_0^*(1+e^{-\frac{C}{3}\sqrt{\log L}}))$, yet does not contain $\mathfrak{R}_0$. Reveal the outermost H level line containing $\mathfrak{R}_1$. This brings us into the setting of Claim 6.2. Now define $\widetilde{\mathfrak{R}}_0$ as

$$\widetilde{\mathfrak{R}}_0 := \bigcup_{x:\, \ell_0^*(1+e^{-\frac{C}{3}\sqrt{\log L}}) \leq \ell_x} \mathcal{W}(x, \ell_0^*(1+e^{-\frac{C}{3}\sqrt{\log L}}))\,.$$

With Claim 6.2 in mind, we first argue that for any domain $V \subset \Lambda$ which contains $\mathcal{W}(o, \ell_0^*(1+e^{-\frac{C}{3}\sqrt{\log L}}))$ yet does not contain $\widetilde{\mathfrak{R}}_0$, there must exist x such that $\mathcal{W}(x, \ell_0^*(1+e^{-\frac{C}{3}\sqrt{\log L}})) \subset V$ and $\ell_0^*(1+e^{-\frac{C}{3}\sqrt{\log L}}) \leq \ell_x$, but $V \not\subset \mathcal{W}'(x, \ell_0^*(1+e^{-\frac{C}{3}\sqrt{\log L}}))$. Indeed, one can simply start with $\mathcal{W}(o, \ell_0^*(1+e^{-\frac{C}{3}\sqrt{\log L}}))$ and keep shifting the center by one up, down, left, or to the right within $\widetilde{\mathfrak{R}}_0$ until it intersects a point in $\widetilde{\mathfrak{R}}_0 \setminus V$. The last shift right before will then be in $V \cap \widetilde{\mathfrak{R}}_0$, which produces the desired shift $\mathcal{W}(x, \ell_0^*(1+e^{-\frac{C}{3}\sqrt{\log L}}))$.

Thus, by Claim 6.2 and a union bound over all x such that $\ell_0^*(1+e^{-\frac{C}{3}\sqrt{\log L}}) \leq \ell_x$, we have that the probability that a `large` $H+1$ level line exists and does not contain $\widetilde{\mathfrak{R}}_0$ is $o(1)$. Since $\mathfrak{R}_0 \subset \widetilde{\mathfrak{R}}_0$, this concludes the proof. ∎

6.2. **Top level lines are contained within limit shapes.** In this section, we prove that the top level line is contained in a translation of Wulff shapes. Together with [14, Thm. 4.4] (restated here as Theorem 2.14) and Proposition 6.1, this proves Theorem 1.1.

Theorem 6.3. *Fix β sufficiently large and $n \geq 0$. Then, the event that $\mathfrak{L}_n$ exists but is not contained in the shape $L\mathscr{L}(\ell_n^*(1-e^{-\frac{C}{3}\sqrt{\log L}}))$ occurs with probability $o(1)$ in π_Λ^0. Consequently, if L is such that $\pi_\Lambda^0(\mathfrak{L}_0 \neq \emptyset) > \delta_L$, then the theorem also holds under $\pi_\Lambda^0(\cdot \mid \mathfrak{L}_0 \neq \emptyset)$, and if $\pi_\Lambda^0(\mathfrak{L}_0 = \emptyset) > \delta_L$, then the theorem also holds under $\pi_\Lambda^0(\cdot \mid \mathfrak{L}_0 = \emptyset)$.*

We begin with Proposition 6.4 below, a counterpart to [14, Thm. 4.9] which shows that the level line in a smaller domain will not deviate too far from its initial height. To prove this, it is crucial that we have an expression for the law of γ with an error of at most a multiplicative $(1+o(1))$ in the larger domain of size $\approx L^{2/3} \times L^{2/3}$, as in Proposition 4.1. The previous work [14] also looked

[9] In particular, a version of [14, Thm. 4.9] where the rectangle has dimensions as described above. As discussed in [14, Rem. 4.7], this poses no issues. One may also compare the equations leading to Eq. (6.4) with those leading to [14, Eq. 4.34] for details.

at the restriction of the level line on an $L^{2/3} \times L^{2/3}$ rectangle, yet the nonnegative conditioning was only enforced on a much smaller $\approx L^{2/3} \times L^{1/3}$ rectangle, reducing the effective size of the domain. This monotonicity trick no longer works here, being in the wrong direction here. Hence, the main obstacle lies in obtaining Proposition 4.1.

Fix $n \geq 0$. We define a slight modification of the "near rectangular" domains used in [14, Thm 4.9]. Let R be a $N_n^{2/3} e^{C\sqrt{\log L}} \times 3N_n^{2/3} e^{C\sqrt{\log L}}$ rectangle, with two marked points A, B on the left and right sides of ∂R, respectively. Assume that if θ is the angle of $\overline{AB}$, then $|\theta| \leq \pi/4$, and that both A, B are distance at least $N_n^{2/3} e^{C\sqrt{\log L}}$ from the top and bottom sides of R. We will consider domains Q satisfying the following:

1. Q is simply connected,
2. $\mathrm{dist}(\partial Q, \partial R) \leq \log L$,
3. There exists $A', B' \in \partial Q$ such that $\mathscr{C}(\overline{A'B'}) \subset Q$ (see the definition of $\mathscr{C}$ below Eq. (5.7)), and $\max\{d(A, A'), \mathrm{dist}(B, B')\} \leq 2(\log L)^5$.[10]
4. The boundary conditions ξ assigns height $H + 1 - n$ on the top arc from A' to B', and $H - n$ on the bottom arc.

Proposition 6.4. *The following holds uniformly over all possible Q, ξ as above. Fix $n \geq 0$. Let $\mathfrak{L}_n$ be the $(H + 1 - n)$ level line induced by ξ. Let θ denote the angle of $\overline{AB}$ with the x-axis. Then, with π_Q^ξ-probability $1 - o(1)$, $\mathfrak{L}_n$ lies above the point $X = (0, Y - \sigma e^{C\sqrt{\log L}} - \log L)$, where*

$$Y = -\frac{\rho_n N_n^{1/3} e^{2C\sqrt{\log L}}}{8(\tau_\beta(\theta) + \tau_\beta''(\theta)) \cos(\theta)^3}, \tag{6.1}$$

$$\sigma^2 = \frac{N_n^{2/3} e^{C\sqrt{\log L}}}{4(\tau_\beta(\theta) + \tau_\beta''(\theta)) \cos(\theta)^3}.$$

Proof. The proof is essentially the same as that of [14, Thm. 4.9], which we recall showed instead that $\mathfrak{L}_n$ lies below $(0, Y + \sigma e^{C\sqrt{\log L}} + \log L)$. We only comment on the few necessary modifications. First, note that we are imposing the additional assumption that A, B are distance at least $N_n^{2/3} e^{C\sqrt{\log L}}$ from both the top and bottom sides of R. We will shortly bound the length of γ, the level line containing $\mathfrak{L}_n$, such that w.h.p. it will not reach the top or bottom of Q, preventing any potential pinning issues. (The setting of [14, Thm. 4.9] did not allow for this and relied on a conditioning argument instead, which we no longer need here.)

To set up for an application of Proposition 4.1, let $\mathcal{G}$ denote the set of γ with $\mathscr{E}_\beta(\gamma) \leq 1.1|A-B|_1$. We can show that $\widehat{\pi}_Q^\eta(\mathcal{G}^c) \leq e^{-1.1(\beta - C)|A-B|_1}$ by using the form of the law given by Proposition 2.4 and applying a Peierls map to a minimal length disagreement polymer. At the same time, we can lower bound $\widehat{\pi}_Q^\eta(\phi_x \geq 0, \forall x \in Q) \geq e^{-cL^{1/3-o(1)}}$ by the standard FKG computation explained in Claim 5.5. Together this implies that $\pi_Q^\eta(\mathcal{G}^c) = e^{-1.1(\beta - C')|A-B|_1}$, and we can assume we are only considering $\gamma \in \mathcal{G}$.

The law of such γ is given by Proposition 4.1. Compared to Proposition 2.5, the area tilt now has an additional prefactor of ρ_n, but the preliminary lemmas needed for [14, Thm. 4.9] (i.e., bounding the size of components D_i in [14, Lem. 4.12,4.15] and bounding the partition function in [14, Lem. 4.17]) allowed for any $\mu > 0$ prefactor, and therefore still hold. We can thus proceed exactly as in the proof of [14, Thm. 4.9], observing that the proof actually showed the distribution of the height of γ at the line $x = 0$ is comparable to a Gaussian centered at Y with variance σ^2,

[10] In [14], there was an additional requirement that linear cones emanating from A and B did not intersect their respective sides of ∂Q. This follows also from our construction of Q in the proof of Claim 6.9. However, it is not used in the proof of [14, Thm. 4.9], and hence not needed for Proposition 6.4, so we do not include it.

so controlling the probability that the height exceeds $Y + \sigma e^{C\sqrt{\log L}} + \log L$ is no different than controlling the probability that it is at most $Y - \sigma e^{C\sqrt{\log L}} - \log L$.[11] ■

We will prove Theorem 6.3 first for $n = 0$, using an induction over the radius ℓ of the Wulff shape. The base case will be proven in Lemma 6.6 and the induction step in Lemma 6.8. Then we will use an induction over n to prove Theorem 6.3 for lower level lines. This involves combining the framework of the disagreement polymer developed in [14] with the "retreat of droplets" proof method found in [13, Section 6.2]. Throughout the rest of the section, we will refer to the error quantity $\varepsilon_n := N_n^{-1/3}$.

Definition 6.5. Define $\mathcal{E}_n(\ell)$ as the (decreasing) event that there is no `large` chain of sites with height $\geq H + 1 - n$ intersecting the exterior of the shape $L(1+\varepsilon_n)\mathscr{L}(\ell)$.

Observe that to prove Theorem 6.3, it suffices to show that w.h.p. we have $\mathcal{E}_n(\ell_n^*(1-\frac{1}{2}e^{-\frac{C}{3}\sqrt{\log L}}))$, as the factor of $1/2$ makes up for the scaling factor of $(1+\varepsilon_n)$. More precisely, the additional scaling only comes into effect at the corners of the shape since we are still restricted to Λ regardless. It is then easy to check by starting at the corners of the shapes and looking at the intersection points with $\partial\Lambda$ that $(1+\varepsilon_n)L\mathscr{L}(\ell_n^*(1-\frac{1}{2}e^{-\frac{C}{3}\sqrt{\log L}})) \cap \Lambda \subset L\mathscr{L}(\ell_n^*(1-e^{-\frac{C}{3}\sqrt{\log L}}))$.

Lemma 6.6 (Base case for $n = 0$)**.** *There exists a constant $s > 0$ such that for $\ell_0 := sN_0/L$, we have that $\mathcal{E}_0(\ell_0)$ holds w.h.p.*

Proof. We will write the proof in terms of n so it is clear how to generalize it later on, but here $n = 0$. It suffices by symmetry to just look at one corner of the shape $L\mathscr{L}(\ell_n)$, say the south-west corner, and show that there is no `large` chain of height $\geq H + 1 - n$ sites which exits $L\mathscr{L}(\ell_n)$ in this corner except with probability $o(1)$. As we want to upper bound the probability of an increasing event, we can first make monotone increasing adjustments to the measure. Call A, B the two points where $L\mathscr{L}(2\ell_n)$ lifts off the west, south boundaries of $\partial\Lambda$, respectively (they are at distance $\leq (1+\varepsilon_\beta)sN_n$ from the bottom left corner of Λ). Let R be the rectangle with width $B_1 - A_1$ and height sN_n, such that its top right corner is at B. Let $\widetilde{\Lambda} := \Lambda \cup R$. Finally, let ξ denote boundary conditions on $\partial\widetilde{\Lambda}$ taking value $H - n$ on the smaller arc of $\partial\widetilde{\Lambda}$ between A and B, and $H+1-n$ otherwise. By monotonicity, we can move to the measure $\pi^{\eta}_{\widetilde{\Lambda}}$. Let $\mathfrak{L}$ be the $H+1-n$ level line including the boundary disagreements of ξ, and let γ be the disagreement polymer containing $\mathfrak{L}$. Let D_1 denote the region below γ (which has $H-n$ b.c.). It suffices to show that γ does not exit $L\mathscr{L}(\ell_n)$ except with probability $o(1)$, as then a `large` chain of $\geq H + 1 - n$ sites can be excluded by a standard Peierls argument on D_1.

Once again, we begin by setting up for Proposition 4.2 with the following claim:

Claim 6.7. *Let $\mathcal{G}$ denote the set of γ such that $\mathscr{E}_\beta(\gamma) \leq Le^{\sqrt{\log L}}$, $|D_1| \leq (\frac{3\beta}{\widehat{\pi}_\infty(\phi_o > H-n)})^2$. Then, $\pi^{\eta}_{\widetilde{\Lambda}}(\mathcal{G}) = 1 - o(1)$.*

Proof. We first bound the probability that $\mathcal{E}_\beta(\gamma) > Le^{\sqrt{\log L}}$. Similarly to Claim 5.5, by FKG, the decorrelation estimate of Eq. (3.24), and the bound on the large deviation ratios in Theorem 2.9, we have the lower bound

$$\widehat{\pi}^{\eta}_{\widetilde{\Lambda}}(\phi_x \geq 0,\ \forall x \in \widetilde{\Lambda}) \geq \prod_{x\in\Lambda} \widehat{\pi}^{H-n}_{\widetilde{\Lambda}}(\phi_x \geq 0) \geq \tfrac{1}{2}\exp\big(-|\widetilde{\Lambda}|\widehat{\pi}^{H-n}_{\infty}(\phi_x < 0)\big) \geq \tfrac{1}{2}e^{-Le^{c\sqrt{\beta\log L/\log\log L}}}.$$

(The $\sqrt{\log L/\log\log L}$ is not needed for $n = 0$, but is there to clarify how to generalize for larger n.) Hence, it suffices to show in the no floor measure that $\widehat{\pi}^{\eta}_{\widetilde{\Lambda}}(\mathcal{E}_\beta(\gamma) > Le^{\sqrt{\log L}}) = o(e^{-Le^{c\sqrt{\beta\log L/\log\log L}}})$.

[11] In [14], we arbitrarily decided to bound the size of components using Peierls maps by $(\log L)^2$ instead of $\log L$. This makes no difference.

Using the form of the law in Proposition 2.4, we can conclude by a Peierls map to a minimal length path from A to B that $\widehat{\pi}^{\eta}_{\widetilde{\Lambda}}(\mathscr{E}_\beta(\gamma) \geq Le^{\sqrt{\log L}}) \leq e^{-c\beta Le^{\sqrt{\log L}}}$.

We turn now to bound $|D_1|$. We will in fact show that for some c, c', we have

$$\pi^{\eta}_{\widetilde{\Lambda}}(|D_1| > cs^2N_n^2) = e^{-c'sN_n}\,.$$

Let D_1' be the connected component of sites with height $\leq H-n$ containing the $H-n$ boundary condition of ξ. Then $D_1 \subset D_1'$, and the event $|D_1'| > cs^2N_n^2$ is decreasing. Hence by monotonicity it suffices to drop the floor and bound $\widehat{\pi}^{\eta}_{\widetilde{\Lambda}}(|D_1'| > cs^2N_n^2)$. Without the floor, we can again appeal to Proposition 2.4 and apply a Peierls map argument. More specifically, by the isoperimetric inequality and the fact that $|\partial D_1' \setminus \partial\widetilde{\Lambda}| \leq |\gamma|$, we have that $|D_1'| > cs^2N_n^2$ implies that $|\gamma| \geq (\frac{\sqrt{c}}{4} - 4(1+\varepsilon_\beta))sN_n$. Taking c large enough, we can rule this out by considering the Peierls map which sends γ to a minimal length path from A to B. ■

With the claim proven, we can restrict our attention to $\gamma \in \mathcal{G}$. Moreover, specific for $n = 0$, we have $|D_0| \leq |\Lambda| \leq (\frac{3\beta}{\widehat{\pi}_\infty(\phi_o > H+1-n)})^2$. Hence by Proposition 4.2 we have

$$\pi^{\eta}_{\Lambda}(\gamma) \propto \exp\Big(-\mathscr{E}^*_\beta(\gamma) + \mathfrak{I}_\Lambda(\gamma) + \tfrac{\mathsf{p}_n}{N_n}|D_0| + \mathsf{p}_n\widehat{\pi}_\infty(\phi_x \geq H+1-n)|(\bigcup_{i\geq 2} D_i)| + O(L^{1/2+o(1)})\Big) \tag{6.2}$$

$$=: e^{O(L^{1/2+o(1)})}\mathfrak{p}^{\eta}_V(\gamma) \tag{6.3}$$

Suppose that we did not have the extra $e^{O(L^{1/2+o(1)})}$ error term and the weight was just $\mathfrak{p}^{\eta}_V(\gamma)$. Then, using the same argument as leading to Eq. (5.4), we can further restrict to γ such that $\max_{\mathcal{D}\in\mathcal{D}} |\partial\mathcal{D}| \leq \log L$, so that the term $\mathsf{p}_n\widehat{\pi}_\infty(\phi_x \geq H)|(\bigcup_{i\geq 2} D_i)| = o(1)$. In more detail, with Eq. (6.2), we can rule out the case that $|\gamma| > 5sN_n$ by taking the Peierls map sending γ to the disagreement contour which coincides with the small arc of the boundary from A to B on $\partial\widetilde{\Lambda}$. Since the minimal length path from A to B has length $2\ell(1+\varepsilon_\beta) \leq 2sN_n(1+\varepsilon_\beta)$, this implies that we can assume $\max_{\mathcal{D}\in\mathfrak{D}(\gamma)} |\partial\mathcal{D}| \leq 5sN_n$ as $\frac{1}{2}|\partial\mathcal{D}|$ is excess length. With this input, the steps leading up to Eq. (5.4) can be followed exactly, considering the map which replaces a component $\mathcal{D}$ with a minimal length path.

Hence, we will focus on the polymer model of disagreement polymers from A to B, restricted to lie in $\widetilde{\Lambda}$, with weight given by $\exp(-\mathscr{E}^*_\beta(\gamma) + \mathfrak{I}_\Lambda(\gamma) + \frac{\mathsf{p}_n}{N_n}|D_0|)$, and show that under this model, γ exits $L\mathscr{L}(\ell)$ with probability $o(e^{-L^{1/2+o(1)}})$. Call $\mathcal{B}$ the set of bad γ which exits $L\mathscr{L}(\ell)$. Denote the probability and expectation with respect to this model by $\mathbf{P}$ and $\mathbf{E}$. We can also consider the same model without the area term $\frac{\mathsf{p}_n}{N_n}|D_0|$, denoted by $\hat{\mathbf{P}}$ and $\hat{\mathbf{E}}$. Finally, let $\mathcal{A}(D_0)$ denote the area of $|D_0|$ normalized to be a signed area with respect to the line segment $\overline{AB}$ (where the area above $\overline{AB}$ is signed positive). This normalization doesn't change the measure, so we can replace $|D_0|$ with $\mathcal{A}(D_0)$ in the definitions of $\mathbf{P}$ and $\mathbf{E}$.

Automatically, we have that for some absolute constant c, $\mathcal{A}(D_0) \leq cs^2N_n^2$. Then, by Jensen's inequality we can write

$$\mathbf{P}(\mathcal{B}) = \frac{\hat{\mathbf{E}}[e^{\mathsf{p}_n\mathcal{A}(D_0)/N_n}\mathbb{1}_{\mathcal{B}}]}{\hat{\mathbf{E}}[e^{\mathsf{p}_n\mathcal{A}(D_0)/N_n}]} \leq \frac{e^{cs^2N_n}\hat{\mathbf{E}}[\mathbb{1}_{\mathcal{B}}]}{\hat{\mathbf{E}}[e^{\mathsf{p}_n\mathcal{A}(D_0)/N_n}]} \leq \frac{e^{cs^2N_n}\hat{\mathbf{E}}[\mathbb{1}_{\mathcal{B}}]}{e^{\hat{\mathbf{E}}[\mathsf{p}_n\mathcal{A}(D_0)/N_n]}}\,.$$

From [14, Prop. 3.14, Lem. 3.16], for some constants $c', c'' > 0$, the expectation in the numerator can be upper bounded by $e^{-c'sN_n}$ and the expectation in the denominator can be lower bounded by $-c''(sN_n)^{1/2}$. For s sufficiently small depending only on the constants c, c', the dominant term is $e^{-c'sN_n}$, which proves the desired upper bound on $\mathbf{P}(\mathcal{B})$. ■

Given the local control over the level lines in Proposition 6.4, the induction step below follows as in [13], only with certain calculations involving N_n instead of L. For completeness, the main ideas of the proof are presented below, and we refer to [13] for more details. Note that unlike the base case, the proof for the induction step works for any fixed $n \geq 0$.

Lemma 6.8 ([13, Lem. 6.12], Induction step for $n \geq 0$)**.** *Fix $n \geq 0$. Let $\ell_n = sN_n/L$ for s from Lemma 6.6. If $\ell_n \leq \ell \leq (1 - \frac{1}{2}e^{-\frac{C}{3}\sqrt{\log L}})\ell_n^*$ and $\mathcal{E}_n(\ell_n)$ holds w.h.p., then so does $\mathcal{E}_n(\ell + L^{-3/4})$.*

Proof. By symmetry it suffices to consider one of the four corners of Λ, say the bottom left one. Take the convention that the origin is at the center of Λ, so in particular the bottom side of $\partial\Lambda$ is on the line $y = -\frac{L}{2}$. Let $f_\ell(x)$ be the function on $(-\frac{1+\varepsilon_n}{2}L, 0]$ whose graph is the bottom left section of $\partial L(1+\varepsilon_n)\mathscr{L}(\ell)$. Let $\hat{x}(\ell)$ be the solution to $f_\ell(x) = x$, and $x_{\mathsf{L}}(\ell)$, $x_{\mathsf{R}}(\ell)$ the first points where $f_\ell(x)$ is smaller than $-\frac{L}{2}$, $-\frac{L}{2}(1+\varepsilon_n)$, respectively.

Now let $\ell' = \ell + L^{-3/4}$, and fix $x \in [\hat{x}(\ell), x_{\mathsf{L}}(\ell')]$. Let $x^{\pm} := x + \frac{1}{2}N_n^{2/3}e^{C\sqrt{\log L}}$. Let $\theta \in [0, \pi/4]$ satisfy $\tan(\theta) = |f_\ell'(x)|$. By construction, $x_{\mathsf{R}}(\ell) - x_{\mathsf{L}}(\ell') \geq c\sqrt{\varepsilon_n}N_n = cN_n^{5/6}$ for some constant $c > 0$, so that in particular $x^+ < x_{\mathsf{R}}(\ell)$. Thus, we can use the deterministic bound on Wulff shapes in [13, Lem. 3.9] with $d = N_n^{2/3}e^{C\sqrt{\log L}}/L$ and then scale by $L(1+\varepsilon_n)$ to obtain that

$$f_\ell(x) = \tfrac{1}{2}(f_\ell(x^-) + f_\ell(x^+)) - \frac{\mathsf{w}_1(\tau_\beta)N_n^{4/3}e^{2C\sqrt{\log L}}}{L\ell(1+\varepsilon_n)16(\tau_\beta(\theta) + \tau_\beta''(\theta))\cos^3(\theta)} + o(1)\,.$$

Since $\ell' = \ell + L^{-3/4}$, the scaling of the Wulff shape implies that $f_{\ell'}(x) \leq f_\ell(x) + CL^{1/4}$ for some constant C. We aim to show that w.h.p., there is no `large` chain of sites with height $\geq H + 1 - n$ which drops below the point $(x, f_\ell(x'))$. Motivated by Proposition 6.4, define

$$Z_\ell(x) := \tfrac{1}{2}(f_\ell(x_-) + f_\ell(x_+)) - \frac{\rho_n N_n^{1/3}e^{2C\sqrt{\log L}}}{8(\tau_\beta(\theta) + \tau_\beta''(\theta))\cos^3(\theta)} - \sigma e^{C\sqrt{\log L}}\,.$$

Suppose we can show the following claim:

Claim 6.9. *For any choice of $x \in [\hat{x}(\ell), x_{\mathsf{L}}(\ell')]$, with probability $1 - o(1)$, there is no `large` chain of sites with height $\geq H + 1 - n$ which drops below the point $(x, Z_\ell(x))$.*

This is sufficient as long as $Z_\ell(x) \geq f_\ell(x) + CL^{1/4}$, and it is a straightforward computation to check that this is satisfied as long as ℓ satisfies

$$\frac{\mathsf{w}_1(\tau_\beta)N_n}{2L\ell(1+\varepsilon_n)\rho_n} - 1 \geq e^{-\frac{C}{2}\sqrt{\log L}}c(\beta,\theta) \tag{6.4}$$

for some constant $c(\beta, \theta)$ (similar to [14, Eq. 4.34]), or equivalently,

$$\ell \leq \frac{w_1(\tau_\beta)N_n}{2L(1+\varepsilon_n)(1 + e^{-\frac{C}{2}\sqrt{\log L}}c(\beta,\theta))\rho_n}\,.$$

This in turn is satisfied by $\ell \leq (1 - \frac{1}{2}e^{-\frac{C}{3}\sqrt{\log L}})\ell_n^*$. We can then conclude the proof by iterating over all $x \in [\hat{x}(\ell), x_{\mathsf{L}}(\ell')]$ and then using symmetry across the line $y = x$.

We now prove Claim 6.9. This will be a monotonicity argument to justify the application of Proposition 6.4. Observe that the event in the claim is decreasing, so we are allowed to increase heights en-route to showing it occurs w.h.p. Let R be the rectangle of width $N_n^{2/3}e^{C\sqrt{\log L}}$ and height $3N_n^{2/3}e^{C\sqrt{\log L}}$, centered at x. Call A, B the intersection of $\partial L(1+\varepsilon_n)\mathscr{L}(\ell)$ with the left, right sides of ∂R respectively. (If R does not fit inside Λ, we can again enlarge the domain, taking $\widetilde{\Lambda} = \Lambda \cup R$ with 0 boundary conditions.) Call the portion of the boundary ∂R above A, B by $\partial^1 R$, and let $\partial^2 R := \partial R \setminus \partial^1 R$. If we reveal the components of sites with height $\geq H + 2 - n$ along $\partial^1 R$, the external boundaries of such revealed components stitch together to form a chain of sites with height $\leq H+1-n$. By a Peierls argument, w.h.p. these components will have size at most $\log L$. Similarly, if we reveal the components of sites with height $\geq H+1-n$ along $\partial^2 R$, then the external boundaries can be stitched together to form a chain of sites with height $\leq H-n$, and by definition on the event $\mathcal{E}_n(\ell)$ these components will have size at most $\log L$. Hence, w.h.p, there exists a circuit of sites passing through A, B which has height $\leq H-n$ in the arc below A, B and height $\leq H+1-n$ in the arc above A, B. Revealing the outermost such circuit and raising the heights to be equal to $H-n-1$ and $H-n$ respectively, the result is a domain Q with boundary conditions ξ satisfying Items 1, 2 and 4 with respect to R (and notably, no information about ϕ in the interior of Q was revealed). Finally, the technical condition Item 3 can be satisfied by a procedure completely analogous to steps 5, 6 in the proof of [14, Lem. 4.8]. Hence, we are in the position to apply Proposition 6.4, which immediately implies Claim 6.9. ∎

This proves Theorem 6.3 for $n=0$, and now it remains to extend this to any fixed $n \geq 0$.

Lemma 6.10. *For $n \geq 1$, if $\mathcal{E}_{n-1}(\ell^*_{n-1}(1-\frac{1}{2}e^{-\frac{C}{3}\sqrt{\log L}}))$ holds w.h.p., then $\mathcal{E}_n(\ell^*_n(1-\frac{1}{2}e^{-\frac{C}{3}\sqrt{\log L}}))$ holds w.h.p. as well.*

Proof. Since we already have Lemma 6.8, the only thing missing is an analog of Lemma 6.6 for general n. That is, with s as in Lemma 6.6 and $\ell_n := sN_n/L$, we want to show that $\mathcal{E}_n(\ell_n)$ holds with probability $1-o(1)$ for all $n \geq 1$. The difference from the $n=0$ case is that we now need to work to bound $|D_0| \leq (\frac{3\beta}{\widehat{\pi}_\infty(\phi_o > H+1-n)})^2$ for the application of Proposition 4.2, while for $n=0$ this was satisfied from the trivial bound $|D_0| \leq |\Lambda|$.

For each vertex v in the southwest corner of $L(1+\delta_{n-1})\mathscr{L}(\ell^*_{n-1}(1-\frac{1}{2}e^{-\frac{C}{3}\sqrt{\log L}}))$, reveal the connected component of sites with height $\geq H+2-n$. This reveals a boundary of $\leq H+1-n$ sites. Since we are assuming that $\mathcal{E}_{n-1}(\ell^*_{n-1}(1-\frac{1}{2}e^{-\frac{C}{3}\sqrt{\log L}})$ holds w.h.p., then w.h.p. all of these components have size at most $\log L$. We can stitch these together to form a boundary Γ of $\leq H+1-n$ sites, and then by monotonicity we can raise the boundary to height $H+1-n$. Now, we can follow the proof of Lemma 6.6 starting with the region below Γ, as opposed to all of Λ. We now have $|D_0| \leq CN^2_{n-1}$ for a constant C depending only on s (and hence independent of β), the only missing input needed for the application of Proposition 4.2. Moreover, Γ is sufficiently far away from the end-points of γ so that it does not interfere with any of the random walk estimates used. (More precisely, one would need Γ to stay distance CN_n away for a constant C depending on s. We have Γ is distance $O(N_{n-1})$ away, which is much further.) Hence, the proof of Lemma 6.6 shows that with probability $1-o(1)$ we have $\mathcal{E}_n(\ell_n)$, and then Lemma 6.8 concludes. ∎

Remark 6.11. The proof of Lemma 6.8 gives details behind the heuristic described in the proof ideas in Section 1.3. Here we have an area-tilted random walk at scale $d = N_n^{2/3} f(L)$ for $f(L) = e^{C\sqrt{\log L}}$. The quantity $f_\ell(x)$ is where the boundary of the Wulff shape of size ℓL is, and $Z_\ell(x)$ is where the level line $\mathfrak{L}_n$ is w.h.p. above. Let us ignore the centering by $\frac{1}{2}(f_\ell(x^-)+f_\ell(x^+))$ present in both terms. Let us also ignore the factor of $1+\varepsilon_n$, which was necessary only for technical reasons, since ε_n is negligibly small. We are left with

$$\bar{f}_\ell(x) := \frac{\mathsf{w}_1(\tau_\beta) N_n^{4/3} e^{2C\sqrt{\log L}}}{L\ell 16(\tau_\beta(\theta)+\tau''_\beta(\theta))\cos^3(\theta)},$$

$$\bar{Z}(x) := \frac{\rho_n N_n^{1/3} e^{2C\sqrt{\log L}}}{8(\tau_\beta(\theta)+\tau''_\beta(\theta))\cos^3(\theta)} + \sigma e^{C\sqrt{\log L}} =: \mu + f\sigma\,.$$

$\bar{Z}(x)$ has a mean term μ, representing the mean of the level line's location above x, and a fluctuation term $f\sigma$ representing the most the level line will deviate from μ except with probability $o(1)$. Then, first observe that ℓ^*_n is indeed the value such that $\bar{f}_{\ell^*_n}(x) = \mu$, giving a first order estimate for the location of $\mathfrak{L}_n$. Secondly, observe that Eq. (6.4) is equivalent to the condition $f_\ell(x) \geq \mu + f\sigma$.

Since $f_\ell(x) = C(\beta,\theta)d^2/\ell$ depends inversely on ℓ, we have altogether that the induction works up to $\ell/\ell_* \le 1 - O(f\sigma/\mu)$. On the other side, we have a matching bound in Proposition 6.1 which is based off an induction holding up to $\ell/\ell_* \ge 1 + O(f\sigma/\mu)$ (see [14, Eq. 4.34]).

Moreover, $f\sigma/\mu$ is a power of f^{-1}, and this is the form of the Hausdorff distance bound in Theorem 1.1. Hence, we obtain a better estimate when f is larger. The size we can take for f comes from the bound on $|\partial F| \vee |\partial V|$ in Proposition 4.1, which in turn depends on the respective bound in Theorem 3.2, measuring how well we can estimate the probability of staying above a floor. In particular, this area bound will change for different p, resulting in the different Hausdorff bounds in Theorem 1.4.

7. Ferrari–Spohn limit law for $\mathfrak{L}_0$

In this section we prove Theorem 1.3, which extends the limit law theorem of [14, Thm. 1.1] to handle the $\mathfrak{L}_0$ level line, and to cover all L (whereas the previous result avoided an exceptional set of L close to $L_c^{(h)}$). As in Section 6.1, this is a matter of modifying the original proof to circumvent an initial lack of monotonicity. The main point is that even though a measure such as $\pi^0_\Lambda(\cdot \mid \mathfrak{L}_0 \neq \emptyset)$ does not have FKG, by zooming in to a mesoscopic box, we can still compare level lines in this measure to level lines in a $\mathbb{Z}$GFF measure which does have FKG. We split the proof into four statements in the proposition below, being the analog of [14, Thms. 5.1 and 6.1] in the settings where we condition on $\mathfrak{L}_0 = \emptyset$ and $\mathfrak{L}_0 \neq \emptyset$.

Fix $n \ge 0$ and $K > 0$. Let $\rho_n(x)$ be the maximum vertical distance of $\mathfrak{L}_n$ above $x + (\frac{L}{2}, 0)$ for $-N_n^{2/3} \le x \le N_n^{2/3}$, and let σ_n be the same constant as in Theorem B. Define the process $Y_n(t) := N_n^{-1/3}\rho_n(tN_n^{2/3})$ $(t \in [-K, K])$.

Proposition 7.1. *Fix $m \ge 0$. For every sequence of $L \to \infty$ such that each L satisfies $\pi^0_\Lambda(\mathfrak{L}_0 \neq \emptyset) > \delta_L$, then sampling $\mathfrak{L}_n$ under the measures $\pi^0_\Lambda(\cdot \mid \mathfrak{L}_0 \neq \emptyset)$, every weak limit point $(\mathbf{Y}_n(t))_{0\le n\le m}$ of the processes $(Y_n(t))_{0\le n\le m}$ satisfies*

$$(\mathbf{Y}_n)_{0\le n\le m} \preceq \bigotimes_{0\le n\le m} \mathsf{FS}_{\sigma_n}\,, \tag{7.1}$$

$$(\mathbf{Y}_n)_{0\le n\le m} \succeq \bigotimes_{0\le n\le m} \mathsf{FS}_{\sigma_n}\,. \tag{7.2}$$

If the sequence $L \to \infty$ is such that each L satisfies $\pi^0_\Lambda(\mathfrak{L}_0 = \emptyset) > \delta$, then sampling $\mathfrak{L}_n$ under the measures $\pi^0_\Lambda(\cdot \mid \mathfrak{L}_0 = \emptyset)$ we have

$$(\mathbf{Y}_n)_{1\le n\le m} \preceq \bigotimes_{1\le n\le m} \mathsf{FS}_{\sigma_n}\,, \tag{7.3}$$

$$(\mathbf{Y}_n)_{1\le n\le m} \succeq \bigotimes_{1\le n\le m} \mathsf{FS}_{\sigma_n}\,. \tag{7.4}$$

Observation 7.2. *For any L, under π^0_Λ, by a Peierls argument, w.h.p. there are no* `large` *level-line down-loops (as per Definition 2.1). Moreover by Theorem 1.1, w.h.p. there is at most one* `large` *level line for heights $1,\dots,H+1$, and no* `large` *$H+2$ level line[12]. Hence, for an appropriate choice of δ_L, these events occur w.h.p. also under the measure $\pi^0_\Lambda(\cdot \mid \mathfrak{L}_0 \neq \emptyset)$ when $\pi^0_\Lambda(\mathfrak{L}_0 \neq \emptyset) > \delta_L$, and under $\pi^0_\Lambda(\cdot \mid \mathfrak{L}_0 = \emptyset)$ when $\pi^0_\Lambda(\mathfrak{L}_0 = \emptyset) > \delta_L$.*

Proof. As a technical step, let us redefine $\mathfrak{L}_0$ to be the `large` $H+1$ level line surrounding the origin; by Theorem 1.1 this w.h.p. does not change $\mathfrak{L}_0$. We start with Eq. (7.1). Let L be such that $\pi^0_\Lambda(\mathfrak{L}_0 = \emptyset) > \delta_L$. Fixing $n \ge 0$, Proposition 6.1 and the discussion at the beginning of Section 6 show that w.h.p. under $\pi^0_\Lambda(\cdot \mid \mathfrak{L}_0 \neq \emptyset)$, if we look above the bulk of the bottom of $\partial\Lambda$ (say, at least

[12] This characterization was already shown for all heights except $H, H+1$ in [29, Thm. 2]

distance $9L/10$ from the corners), then $\mathfrak{L}_n$ is at most distance $N_n^{1/3}(\log L)^{16}$ from the bottom of Λ and $\mathfrak{L}_{n+1}$ is at most distance $N_n^{1/3}\exp(-c\sqrt{\beta\frac{\log L}{\log\log L}})$ for some constant $c>0$. (The control on $\mathfrak{L}_{n+1}$ follows from the bound on Eq. (2.9).)

Now let R_n be the rectangle of size $N_n^{2/3}(\log L)^{25}\times 2N_n^{2/3}(\log L)^{25}$ positioned so that its bottom side distance $2N_n^{1/3}\exp(-c\sqrt{\beta\log L/\log\log L})$ above the bottom side of Λ, and its center is at $x=0$. Let A,B be the two points on the left and right sides of ∂R_n such that the distance from A,B to the bottom of R_n is $N_n^{1/3}(\log L)^{16}$. As discussed, w.h.p. under $\pi_\Lambda^0(\cdot\mid\mathfrak{L}_0\neq\emptyset)$ the points A,B lie in the interior of $\mathfrak{L}_n$ and R_n lies in the interior of $\mathfrak{L}_{n+1}$.

The main claim to argue is that the stochastic domination lemma of [14, Lem. 4.8] still holds in the conditional measure $\pi_\Lambda^0(\cdot\mid\mathfrak{L}_0\neq\emptyset)$. That is, we will show there is a $\pi_\Lambda^0(\cdot\mid\mathfrak{L}_0\neq\emptyset)$ measurable distribution on simply connected regions $Q\subset R$ and $H+1-n, H-n$ Dobrushin boundary conditions ξ such that w.h.p. (in the sampling of Q and ξ), the restriction to Q of $\mathfrak{L}_n$ under $\pi_\Lambda^0(\cdot\mid\mathfrak{L}_0\neq\emptyset)$ is stochastically dominated by (i.e., lies below) the $H+1-n$ level line in π_Q^ξ. Moreover, the distribution on Q,ξ is supported on pairs such that $d(\partial Q,\partial R_n)<\log L$, and the boundary change in ξ occurs in a sufficiently regular part of ∂Q (see [14, Lem. 4.8 (3)] for details). Consequently, to prove Eq. (7.1), it suffices to prove the bound in π_Q^ξ for every such Q,ξ satisfying the aforementioned properties. Crucially, for every fixed Q,ξ, the measure π_Q^ξ does have FKG.

To prove this claim, call $\partial^1 R_n,\partial^2 R_n$ the top and bottom arcs of ∂R_n going from A to B. Consider the following revealing procedure: first reveal all of $\phi\restriction_{\Lambda\setminus R_n}$. Then, for every $x\in\partial^1 R_n$, reveal its (possibly empty) connected component of sites with height $\leq H-n$, which in turn reveals a boundary of $\geq H+1-n$ sites. Similarly, for every $x\in\partial^2 R_n$, reveal its (possibly empty) connected component of sites with height $\leq H-n-1$, which in turn reveals a boundary of $\geq H-n$ sites. Observe that the set of unrevealed sites, call it Q, has induced boundary conditions of part $\geq H+1-n$ and part $\geq H-n$. A consequence of Observation 7.2 (specifically concerning level-line down-loops), A,B being in the interior of $\mathfrak{L}_n$, and R_n being in the interior of $\mathfrak{L}_{n+1}$, is that throughout this revealing process, w.h.p. every revealed component has size at most $\log L$. Hence, Q is a simply connected component with $d(\partial Q,\partial R_n)<\log L$, and the above procedure gives a $\pi_\Lambda^0(\cdot\mid\mathfrak{L}_0\neq\emptyset)$ measurable distribution on sets $Q\subset R$ and (not yet Dobrushin) b.c. $\tilde\xi$ such that the restriction to Q of $\mathfrak{L}_n$ in $\pi_\Lambda^0(\cdot\mid\mathfrak{L}_0\neq\emptyset)$ is equal to the $H+1-n$ level line in $\pi_Q^{\tilde\xi}(\cdot\mid\mathfrak{L}_0\neq\emptyset)$.[13] Moreover, the event $\{\mathfrak{L}_0\neq\emptyset\}$ is measurable w.r.t. $\phi\restriction_{\Lambda\setminus R_n}$, so that in fact $\pi_Q^{\tilde\xi}(\cdot\mid\mathfrak{L}_0\neq\emptyset)=\pi_Q^{\tilde\xi}$. (Here, we use that $\mathfrak{L}_0$ has been redefined as the `large` $H+1$ level line containing the origin, so its existence is determined by $\phi\restriction_{\Lambda\setminus R_n}$.) The latter is simply a $\mathbb{Z}$GFF measure above a floor, and has FKG. Hence, we can now lower the $\geq H+1-n$ heights to be $H+1-n$, and the $\geq H-n$ heights to be $H-n$, in the manner outlined in steps (5), (6) of the proof of [14, Lem. 4.8] to ensure that the boundary conditions change from $H-n$ to $H+1-n$ at a sufficiently regular part of ∂Q. This concludes the proof of the claim.

Having reduced to studying π_Q^ξ, the proof of Eq. (7.1) now follows exactly as in [14, Thm. 5.1]. First the Ferrari–Spohn domination is established for a single level line $\mathfrak{L}_m$. Then we observe that this domination holds for $\mathfrak{L}_{m-1}$ conditionally on $\mathfrak{L}_m$, since w.h.p. $\mathfrak{L}_m$ (along with all the lower level lines) was anyways revealed during the exposing of $\phi\restriction_{\Lambda\setminus R_{m-1}}$ in the revealing process above. Continuing in this way until $\mathfrak{L}_0$, this suffices to show the stochastic domination of the joint law as written in Eq. (7.1).

[13]At this point the $H+1-n$ level line may not be unique in $\pi_Q^{\tilde\xi}$. Hence, it is more accurate to say that if $\mathcal{A}_1$ is the intersection of the interior of $\mathfrak{L}_n$ with Q under $\pi_\Lambda^0(\cdot\mid\mathfrak{L}_0\neq\emptyset)$ and $\mathcal{A}_2$ is the connected component of the $\geq H+1-n$ sites containing the top side of ∂Q, then $\mathcal{A}_1=\mathcal{A}_2$. This is resolved immediately after when the b.c. $\tilde\xi$ are changed to Dobrushin b.c. ξ.

The proof of Eq. (7.3) is exactly the same, only we consider $n \geq 1$ since there is no longer a $\mathfrak{L}_0$. The proof of Eq. (7.2) is very similar so we will be more brief. This time, we must begin with $n = 0$, but we will leave things in terms of n for clarity on how to handle $n \geq 1$ later. Again we zoom in to a mesoscopic rectangle R_n, this time having dimensions $3TN_n^{2/3} \times N_n^{1/3}(\log L)$ for some $T > K$, and placed so the bottom of ∂R_n coincides with the bottom of $\partial\Lambda$ and R_n is centered at $x = 0$. Let $A = (-TN_n^{2/3}, 0)$ and $B = (TN_n^{2/3}, 0)$. The main claim is that there is a $\pi_\Lambda^0(\cdot \mid \mathfrak{L}_0 \neq \emptyset)$-measurable distribution on simply connected $Q \subset R$ and $H + 1 - n, H - n$ Dobrushin b.c. ξ such that w.h.p., the restriction to Q of $\mathfrak{L}_n$ under $\pi_\Lambda^0(\cdot \mid \mathfrak{L}_0 \neq \emptyset)$ stochastically dominates (i.e., lies above) the $H + 1 - n$ level line in π_Q^ξ. This distribution is supported on Q such that $d(\partial Q, \partial R_n) < \log L$, and the boundary conditions change exactly at A, B.

To prove the claim, under $\pi_\Lambda^0(\cdot \mid \mathfrak{L}_0 \neq \emptyset)$, reveal all of $\phi\restriction_{\Lambda\setminus R_n}$. Then along the top and sides of ∂R_n, reveal each connected component of sites with height $\geq H + 2 - n$, which in turn reveals a boundary of $\leq H + 1 - n$ sites. W.h.p. every component revealed has size at most $\log L$ because we are starting with $n = 0$ and by Observation 7.2 there is no `large` $H + 2 - n$ level line. Hence we can reduce to the measure $\pi_Q^{\tilde{\xi}}$ for some simply connected Q and b.c. $\tilde{\xi}$ which is $\leq H + 1 - n$ on the sides and top of ∂Q, and 0 on the bottom flat side of ∂Q. Then by monotonicity we can raise heights to obtain Dobrushin boundary conditions ξ satisfying the desired properties.

Having reduced to the measure π_Q^ξ, the proof of Eq. (7.2) in the case of a single level line ($m = 0$) now follows exactly as in [14, Thm. 6.1]. In particular, the top level line $\mathfrak{L}_0$ lives at scale $N_0^{1/3}$ and w.h.p. does not intersect the rectangle R_1. Thus, we can now run the proof again for $n = 1$, and this time the revealed components of sites with height $\geq H + 2 - 1$ along the top and sides of ∂R_1 are at most size $\log L$ because by Observation 7.2 there is at most one `large` $H + 2 - 1$ level line and it does not intersect R_1. This establishes the domination of Ferrari–Spohn for $\mathfrak{L}_1$, which holds even conditionally on $\mathfrak{L}_0$ due to the revealing procedure as before. This then shows that w.h.p. $\mathfrak{L}_1$ does not intersect R_2, and the process can be repeated inductively to obtain Eq. (7.2).

Finally, the proof of Eq. (7.4) is the same, only we start with $n = 1$ instead and the needed input that there is no `large` $H + 2 - 1$ level line intersecting R_1 now follows by the conditioning on $\{\mathfrak{L}_0 = \emptyset\}$ and Proposition 6.1. ∎

8. Extension to $|\nabla\phi|^p$ models for $1 \leq p < \infty$

Here we will prove Theorem 1.4, extending results to the $|\nabla\phi|^p$ model from Eq. (1.7). To aid this, the following is a summary of the previous sections and how they relate to each other:

- Section 2 provides Lemma 2.12, which is not model dependent. Otherwise, the preliminary definitions and cluster expansion results there have already been established for general $p > 1$, see in particular [14, Rem. 3.4, Prop. 7.12].
- Section 3 provides estimates on the probability of staying above a floor in a given region. The results of this section will require some new arguments to extend to $p > 1$.
- Section 4 uses the area estimates of Section 3 to study the law of the disagreement polymer in larger domains than covered in Section 2. The proof of these results given Proposition 3.1 and Theorem 3.2 generalize verbatim for $p > 1$.
- Sections 5 to 7 then prove Theorems 1.1 to 1.3 by moving to a polymer model using the results of Section 4, or otherwise using routine FKG and Peierls arguments which hold for all $p \geq 1$.

Hence, it remains to prove analogs of Proposition 3.1 and Theorem 3.2, modifying only the log factors which arise from large deviation probabilities under $\widehat{\pi}_\infty^{(p)}$, and then check that modifying these log factors do not cause any problems. We will do this in Section 8.1 modulo some missing large deviation estimates, and then fill in these missing estimates in Section 8.2. Finally in Section 8.3 we discuss extensions to the SOS model ($p = 1$).

8.1. **Extending Proposition 3.1 and Theorem 3.2 to $p > 1$.** Here we will show how the results of Section 3 extend for general $p > 1$. We define the natural generalizations

$$\xi_n^{(p)} := -\frac{1}{\ell_*^2} \log \widehat{\pi}_\infty^{(p)} \left(\phi_x \geq -(H+1-n), \forall x \in Q_{\ell_*} := [\![1, \ell_*]\!]^2 \right) \quad \text{for} \quad \ell_* := 2^{\lceil \frac{1}{2} \log_2 L \rceil} \tag{8.1}$$

$$\rho_n^{(p)} := \left(\xi_{n+1}^{(p)} - \xi_n^{(p)} \right) N_n \,. \tag{8.2}$$

Then we have the following analogs of Proposition 3.1 and Theorem 3.2. Amidst the many $\log L$ and L^{δ_p} factors, the key properties to note are

(1) The error term in the macroscopic range of Theorem 8.2 is always of the form $L^{1/2+o(1)}$,
(2) The upper bounds on $\frac{\xi_n^{(p)}}{\widehat{\pi}_\infty^{(p)}(\phi_o > H+1-n)}$ give a margin of at least $L^{-1/2+\delta_p}$ from 1,
(3) This upper bound is transferred to $\rho_n^{(p)}$ for $1 < p < 2$, but not for $p > 2$.

Hence, Remark 5.3 applies for $1 < p < 2$ as well. The remark should also hold for $p > 2$, as suggested by the upper bound on $\frac{\xi_n^{(p)}}{\widehat{\pi}_\infty^{(p)}(\phi_o > H+1-n)}$, but we are unable to prove a matching upper bound for $\rho_n^{(p)}$ in this setting. To make progress on this, the first step would be to obtain the rate constant for the large deviation event $\widehat{\pi}_\infty^{(p)}(\phi_o \geq h)$, which was proven for $p \leq 2$ in [29] but still missing for $p > 2$.

Proposition 8.1 (Cf. Proposition 3.1)**.** *Fix $1 < p < 2$, and $n \geq 0$. There exists a constant $\delta_p > 0$ such that*

$$\frac{\xi_n^{(p)}}{\widehat{\pi}_\infty^{(p)}(\phi_o > H+1-n)} \in [1 - L^{-\delta_p + o(1)}, 1 - L^{-1/2+\delta_p+o(1)}] \,, \tag{8.3}$$

$$\rho_n^{(p)} \in [1 - L^{-\delta_p + o(1)}, 1 - L^{-1/2+\delta_p+o(1)}] \,. \tag{8.4}$$

For $p > 2$, there exists $c_0(\beta), c_1(\beta) > 0$ such that

$$\frac{\xi_n^{(p)}}{\widehat{\pi}_\infty^{(p)}(\phi_o > H+1-n)} \in [1 - e^{-c_0(\beta)\sqrt{\log L}^{\frac{p}{p-1}}}, 1 - e^{-c_1(\beta)\sqrt{\log L}^{\frac{p}{p-1}}}] \,, \tag{8.5}$$

$$\rho_n^{(p)} \in [1 - e^{-c_0(\beta)\sqrt{\log L}^{\frac{p}{p-1}}}, 1 + L^{-1+o(1)}] \,. \tag{8.6}$$

Theorem 8.2 (Cf. Theorem 3.2)**.** *Fix $g \geq 0$, $n \geq 0$, and set $h = H + 1 - n$. The following hold for every subsets $F \subseteq V \subset \mathbb{Z}^2$ where F is connected with g holes. There exists constants $\delta_p, c > 0$ such that*

*(1) [*Mesoscopic range*] If $1 < p < 2$ and $|\partial F| \vee |\partial V| \leq O(L^{\frac{2}{3}+\frac{\delta_p}{5}})$, or if $p > 2$ and $|\partial F| \vee |\partial V| \leq O(L^{2/3} e^{c\sqrt{\beta \log L}})$, then*

$$\widehat{\pi}_V^{(p),0}(\phi_x \geq -h, \forall x \in F) = (1 + o(1)) \exp\left(-\xi_n^{(p)} |F| \right) \,. \tag{8.7}$$

*(2) [*Macroscopic range*] If $1 < p < 2$ and $|\partial F| \vee |\partial V| \leq O(Le^{\sqrt{\log L}})$, then*

$$\widehat{\pi}_V^{(p),0}(\phi_x \geq -h, \forall x \in F) \geq \exp\left(-\xi_n^{(p)} |F| + O\big(\sqrt{L} e^{3\sqrt{\log L}}\big) \right) \,, \tag{8.8}$$

and if $\mathfrak{S}$ is the event that there are no disagreement polymers γ in ϕ with $|\gamma| \geq \log L$, then

$$\widehat{\pi}_V^{(p),0}(\phi_x \geq -h, \forall x \in F \mid \mathfrak{S}) \leq \exp\left(-\xi_n^{(p)} |F| + O\big(\sqrt{L} e^{3\sqrt{\log L}}\big) \right) \,. \tag{8.9}$$

If $|F| \leq \big(\frac{3\beta}{\widehat{\pi}_\infty^{(p)}(\phi_o > h)}\big)^2$, then the last upper bound also applies to $\widehat{\pi}_V^0(\phi_x \geq -h, \forall x \in F)$. The same holds if $p > 2$ and $|\partial F| \vee |\partial V| \leq O(Le^{c\sqrt{\beta \log L}}))$, replacing the error of $O\big(\sqrt{L} e^{3\sqrt{\log L}}\big)$ by $O\big(\sqrt{L} e^{5c\sqrt{\beta \log L}}\big)$.

The above statements also hold when replacing $\widehat{\pi}_V^{(p),0}$ *by* $\widehat{\pi}_\infty^{(p)}$.

Before we begin the proofs, we explain why the modified log factors do not cause any problems in the application of Theorem 8.2. There are two changes, firstly the change in the requirement on $|\partial F| \vee |\partial V|$, and secondly the error terms in Eqs. (8.8) and (8.9). These changes directly transfer over to generalizations of Section 4 for general $p > 1$.

The second change is never an issue, since these error terms anyways appear in the form $O(L^{1/2+o(1)})$ in all applications in Sections 5 and 7.

The first change is relevant in Sections 5 to 7 when we initially check that we can rule out a-priori any $|\gamma|$ of abnormally large length, so that the regions above and below γ (or inside and outside γ) have boundary lengths satisfying the above requirement. In particular, this is done in Claim 5.5 and Claim 6.7. In Claim 5.5 the log factors do not appear so there is no difference, while the proof of Claim 6.7 requires only that the ratio of large deviation rates $\frac{\widehat{\pi}_\infty^{(p)}(\phi_o=h-1)}{\widehat{\pi}_\infty^{(p)}(\phi_o=h)}$ (given by [14, Thm. 7.1, Eq. (7.4)]) is smaller than the log term in the upper bound on $|\partial F| \vee |\partial V|$ in the macroscopic regime. For $1 < p < 2$ this is true as $e^{c\beta(\log L)^{\frac{p-1}{p}}} \ll e^{\sqrt{\log L}}$. For $p > 2$, both are of the form $e^{c\sqrt{\beta \log L}}$, but the constant in Theorem 8.2 can be taken to be arbitrarily large (it comes from Proposition 8.7, where a larger c gives a weaker result), and in particular larger than c_0 from [14, Thm. 7.1].

Moreover, the new upper bound on $|\partial F| \vee |\partial V|$ in the mesoscopic range determines the size of the rectangles R we can choose in Section 6 (in particular, in the "retreat" mechanism of Proposition 6.4 and its "growth" analog [14, Thm. 4.9]). As mentioned in Remark 6.11, this results in the change in the Hausdorff distance bounds in Theorem 1.4.[14]

We now begin towards the proof of Proposition 8.1. Recalling the proof of Proposition 3.1, we will first need bounds on $\xi_{\ell,h}^{(p)}/\widehat{\pi}_\infty^{(p)}(\phi_o < -h)$.

Lemma 8.3 (Cf. Lemma 3.4)**.** *Fix* $1 < p < 2$. *The following holds for all* β *sufficiently large. Let* $h = H + 1 - n$ *for fixed* $n \geq 0$. *Then for any* $\delta > 0$ *and all* $1 \leq \ell \leq L^{\frac{1}{2}-\delta}$, *we have*

$$1 - L^{-2\delta+o(1)} \leq \frac{\xi_{\ell,h}^{(p)}}{\widehat{\pi}_\infty^{(p)}(\phi_o < -h)} \leq 1 + L^{-1+o(1)}\,. \tag{8.10}$$

If $p > 2$ *instead, there exists* $c(\beta), c_0(\beta)$ *such that for* $1 \leq \ell \leq \sqrt{L} e^{-c(\beta)\sqrt{\log L}^{\frac{p}{p-1}}}$, *we have*

$$1 - e^{-c_0(\beta)\sqrt{\log L}^{\frac{p}{p-1}}} \leq \frac{\xi_{\ell,h}^{(p)}}{\widehat{\pi}_\infty^{(p)}(\phi_o < -h)} \leq 1 + L^{-1+o(1)}\,. \tag{8.11}$$

Moreover, this holds when replacing $\xi_{\ell,h}$ *by* $-\frac{1}{|S|}\log \widehat{\pi}_\infty^{(p)}(\phi_x \geq -h,\, \forall x \in S)$ *for any set* $S \subset \mathbb{Z}^2$, *not necessarily connected, of size at most* ℓ^2. *In addition, the upper bound holds for all* $\ell \geq 1$.

Proof. The proof of the upper bound is unchanged. The proof of the lower bound for $p > 2$ remains unchanged after plugging in the upper bound on $\widehat{\pi}_\infty^{(p)}(\phi_x < -h \mid \phi_y < -h)$ from [14, Thm. 7.1, Eq. (7.6)]. It is noteworthy that $1 < \frac{p}{p-1} < 2$ in this range, so that $e^{-c(\beta)\sqrt{\log L}^{\frac{p}{p-1}}} = L^{-o(1)}$. In contrast, for $1 < p < 2$, the bound on $\widehat{\pi}_\infty^{(p)}(\phi_x < -h \mid \phi_y < -h)$ is $e^{-c\beta h^p}$ which is a polynomial in L. Hence, repeating the proof for $1 < p < 2$ without modification would restrict the allowed ℓ to be $\leq L^{1/2-c}$ for some c, and this will turn out to be too restrictive for later.

[14] At this point, a keen reader will notice that in fact the requirement on $|\partial F| \vee |\partial V|$ allows for larger domains in $p > 2$ than for $p = 2$, contrary to what should be expected. This is because the bound for $p = 2$ was not sharp; e.g. in the mesoscopic range we could have allowed for all the way up to $O(L^{2/3}e^{c\log L/\log\log L})$ for some constant c depending on previous constants, but the choice of $O(L^{2/3}e^{\sqrt{\log L}})$ suffices and is more reader-friendly.

Thus, we will modify the proof to allow for a larger range of ℓ, at the cost of a suboptimal lower bound in Eq. (8.10). Consider the key equation, Eq. (3.12). Plugging in the bound $e^{-c\beta h^p}$ (which replaces the $e^{-c\beta h^2/\log^2 h}$ factor), we obtain

$$\ell^2(\widehat{\pi}_\infty^{(p)}(\phi_o < -h) + L^{-10}) + O((\log L)^2)e^{-c\beta h^p}\,.$$

Rather than choosing ℓ which causes the above two terms to be of the same order, we choose ℓ that simply ensures the above is $o(1)$. In particular, since $\widehat{\pi}_\infty^{(p)}(\phi_o < -h) \leq \frac{1}{L}e^{c\beta(\log L)^{\frac{p-1}{p}}}$, the conditions on ℓ in the lemma ensure that the display is bounded above by $L^{-2\delta+o(1)}$, which as before directly translates to the desired lower bound. ∎

For a refined upper bound on $\xi_{\ell,h}^{(p)}/\widehat{\pi}_\infty^{(p)}(\phi_o < -h)$, we will first need the following conditional large deviation lower bound. As the proof is very different from the $p=2$ case[15], we relegate this to Section 8.2.

Claim 8.4 (Cf. Claim 3.6)**.** *Fix $1 < p < 2$. There exists $\delta_p, \bar{\delta}_p > 0$ such that for large β, if o' is a neighbor of the origin o then*

$$\widehat{\pi}_\infty^{(p)}(\phi_{o'} \geq h \mid \phi_o \geq h) \geq e^{-\beta(\frac{1}{2}\mathscr{E}^{(p)}(\phi_o^*) - \bar{\delta}_p + o(1))h^p} = L^{-1/2+\delta_p+o(1)}\,.$$

If $p > 2$ instead, then

$$\widehat{\pi}_\infty^{(p)}(\phi_{o'} \geq h \mid \phi_o \geq h) \geq e^{-c\beta h^{p/(p-1)}} = e^{-c(\beta)\sqrt{\log L}^{\frac{p}{p-1}}}\,.$$

Lemma 8.5 (Cf. Lemma 3.5)**.** *Fix $1 < p < 2$, and let δ_p be from Claim 8.4. For $h = H + 1 - n$ with fixed $n \geq 0$, β sufficiently large, and $\ell \geq L^{\frac{1}{4}}$, we have*

$$\frac{\xi_{\ell,h}^{(p)}}{\widehat{\pi}_\infty^{(p)}(\phi_o < -h)} \leq 1 - L^{-1/2+\delta_p+o(1)}\,.$$

If instead $p > 2$, then for any constant $\delta > 0$ we can allow $\ell \geq L^\delta$, and we have

$$\frac{\xi_{\ell,h}^{(p)}}{\widehat{\pi}_\infty^{(p)}(\phi_o < -h)} \leq 1 - e^{-c(\beta)\sqrt{\log L}^{\frac{p}{p-1}}}\,.$$

Proof. The proof remains unchanged from the $p = 2$ case given the new bounds on $\widehat{\pi}_\infty^{(p),0}(\phi_{o'} \geq h \mid \phi_o \geq h)$ and ℓ. ∎

We also insert here the bound on $\bar{\xi}_{\ell,h}^{(p)}/\widehat{\pi}_\infty^{(p)}(\phi_o = -h)$.

Lemma 8.6 (Cf. Lemma 3.8)**.** *In the setting of Lemma 8.3, we have for $1 < p < 2$ that*

$$1 - L^{-2\delta+o(1)} \leq \frac{\bar{\xi}_{\ell,h}^{(p)}}{\widehat{\pi}_\infty^{(p)}(\phi_o = -h)} \leq 1 + L^{-1+o(1)}\,, \tag{8.12}$$

and for $p > 2$ that

$$1 - e^{-c_0(\beta)\sqrt{\log L}^{\frac{p}{p-1}}} \leq \frac{\xi_{\ell,h}^{(p)}}{\widehat{\pi}_\infty^{(p)}(\phi_o < -h)} \leq 1 + L^{-1+o(1)}\,. \tag{8.13}$$

Proof. The proof remains unchanged from the $p = 2$ case, besides the modifications described for Lemma 8.3. ∎

[15]There the claim was stronger and allowed for finite volume settings, but only $V = \mathbb{Z}^2$ was used in application.

This leads to the main proposition controlling the probability of the floor event. (Note that, just as for $p = 2$, we do not need Claim 8.4 and Lemmas 8.5 and 8.6 for this. In particular, we can take ℓ_0 as small as $\log^2 L$, rather than the restriction $\ell \geq L^{1/4}$.)

Proposition 8.7 (Cf. Proposition 3.9). *Fix $g \geq 0$, $n \geq 0$, and set $h = H + 1 - n$. Assume β is sufficiently large. Let δ_p be the minimum of the constant from Claim 8.4 and $\frac{1}{6}$. There exists $c > 0$ such that the following holds for every $\log^2 L \leq \ell_0 \leq L^{\frac{1}{2}-\frac{\delta_p}{2}}$, and subsets $F \subseteq V \subset \mathbb{Z}^2$ where F has g holes and $|\partial F| \leq L^{3/2}$. Let $\Xi = \Xi_1 + \Xi_2 + \Xi_3$, where for $1 < p < 2$ we define*

$$\Xi_1 = \frac{|\partial V|}{L} e^{c\beta(\log L)^{\frac{2p-2}{p}}}\,, \qquad \Xi_2 = \frac{|F|}{\ell_0 L} e^{c\beta(\log L)^{\frac{p-1}{p}}}\,, \qquad \Xi_3 = \frac{|\partial F|\ell_0}{L} L^{-\delta_p + o(1)}\,, \tag{8.14}$$

and for $p > 2$ we define

$$\Xi_1 = \frac{|\partial V|}{L} L^{\varepsilon_\beta}\,, \qquad \Xi_2 = \frac{|F|}{\ell_0 L} e^{c\sqrt{\beta \log L}}\,, \qquad \Xi_3 = \frac{|\partial F|\ell_0}{L} e^{-c(\beta)\sqrt{\log L}^{\frac{p}{p-1}}}\,. \tag{8.15}$$

(i) If $1 < p < 2$ and $|F| \leq \ell_0 L e^{-c\beta(\log L)^{\frac{p-1}{p}}}$, or $p > 2$ and $|F| \leq \ell_0 L e^{-c\sqrt{\beta \log L}}$, then

$$\widehat{\pi}_V^{(p),0}(\phi_x \geq -h,\ \forall x \in F) = \exp\Big(-\xi^{(p)}_{\ell_0,h}|F| + O(\Xi) + o(L^{-5})\Big)\,. \tag{8.16}$$

(ii) Otherwise, if $\mathfrak{S}$ is the event that there are no disagreement polymers γ with $|\gamma| \geq \log L$, then

$$\widehat{\pi}_V^{(p),0}(\phi_x \geq -h,\ \forall x \in F) \geq \exp\Big(-\xi^{(p)}_{\ell_0,h}|F| + O(\Xi)\Big)\,, \tag{8.17}$$

$$\widehat{\pi}_V^{(p),0}(\phi_x \geq -h,\ \forall x \in F \mid \mathfrak{S}) \leq \exp\Big(-\xi^{(p)}_{\ell_0,h}|F| + O(\Xi) + o(L^{-5})\Big)\,. \tag{8.18}$$

If $|F| \leq \big(\frac{3\beta}{\widehat{\pi}^{(p)}_\infty(\phi_o > h)}\big)^2$, then the last upper bound also applies to $\widehat{\pi}_V^{(p),0}(\phi_x \geq -h,\ \forall x \in F)$.

The same holds under $\widehat{\pi}^{(p)}_\infty$, in which case the error term Ξ_1 can be omitted from Ξ. Separately, in the special case where $F = [\![1,\ell]\!]^2$ with $\ell_0 + 10\lceil \log L\rceil \mid \ell$, the error term Ξ_3 can be omitted from Ξ.

Proof. The proof essentially remains the same after plugging in the new large deviation estimates, so we just comment on the differences. The first error term Ξ_1 accounts for points near ∂V, which depends on the control of $\widehat{\pi}_V^{(p),0}(\phi_x < -h)$ for x near ∂V. For this we plug in the bound from [14, Lem. 7.10] (for $1 < p < 2$, the proof of this lemma actually gives the sharper bound $\widehat{\pi}_V^{(p)}(\phi_x \geq h) \leq \widehat{\pi}^{(p)}_\infty(\phi_x \geq h)e^{\varepsilon_\beta C h^{2p-2}}$, whence the fact that $h = O((\log L)^p)$ gives the term $\frac{1}{L}e^{c\beta(\log L)^{\frac{2p-2}{p}}}$).

The second Ξ_2 accounts for points in the annuli of squares (recall the proof involves decomposing V into squares), and only involves points far away from ∂V where we can couple to $\widehat{\pi}^{(p)}_\infty$. The changed terms arise from the upper bound for $\widehat{\pi}^{(p)}_\infty(\phi_o < -h)$, which is obtained by the definition of H and the bound on $\widehat{\pi}^{(p)}_\infty(\phi_o = h)/\widehat{\pi}^{(p)}_\infty(\phi_o = h-1)$ in [14, Thm. 7.1]. Finally, the third term Ξ_3 accounts for squares partially cut off by the boundary. This error depends on the comparison between $\widehat{\pi}^{(p)}_\infty(\phi_o < -h)$ and $\xi_{\ell_0,h}$, in particular, the lower bound of Lemma 8.3 applied at $\delta = \delta_p/2$.

The bound on $|F|$ needed for Eq. (8.16) is to satisfy $|F|\widehat{\pi}^{(p)}_\infty(\phi_o < -h) = o(\ell_0)$ (see Eq. (3.31)). ∎

Proof of Proposition 8.1, except the upper bound of Eq. (8.4). Recall that $\ell_* := 2^{\lceil \frac{1}{2}\log_2 L\rceil} \asymp \sqrt{L}$. As before, the proof is to relate $\xi^{(p)}_{\ell_*,h}, \bar{\xi}^{(p)}_{\ell_*,h}$ to $\xi^{(p)}_{\ell_0,h}, \bar{\xi}^{(p)}_{\ell_0,h}$ for a specific choice of ℓ_0, and then use Lemmas 8.3 and 8.5 to conclude for $\xi^{(p)}$ and Lemma 8.6 for $\rho^{(p)}$. The proof for $p > 2$ remains unchanged.

For $1 < p < 2$ however, we must be more careful with our choice of ℓ_0. (If one were to follow the proof of $p = 2$ as is, the factor of $1 + L^{-1/4}$ would overwhelm the upper bound in Lemma 8.5.) We will take ℓ_0 to be the largest integer less than $L^{\frac{1}{2}-\frac{\delta_p}{2}}$ such that $\ell_0 + 10\lceil \log L\rceil$ divides ℓ_*.

In what follows, the constant c may change from line to line. Applying Proposition 8.7 in the special case for $\widehat{\pi}^{(p)}_\infty$ and $F = [\![1, \ell_*]\!]^2$, we get from Eq. (8.16) that

$$\widehat{\pi}^{(p)}_\infty(\phi_x \geq -h,\, \forall x \in F) = \exp\Big(- \xi^{(p)}_{\ell_0,h}\ell_*^2 + O\Big(\frac{\ell_*^2}{\ell_0 L}e^{c\beta(\log L)^{\frac{p-1}{p}}}\Big) + o(L^{-5})\Big)\,.$$

(In the $p = 2$ proof, we chose not to use the divisibility condition as it was unnecessary. It is needed here.) Plugging in ℓ_* and ℓ_0, we get

$$\xi^{(p)}_{\ell_*,h} = -\frac{1}{\ell_*^2}\log \widehat{\pi}^{(p)}_\infty(\phi_x \geq -h,\, \forall x \in F) = \xi^{(p)}_{\ell_0,h} + O\Big(\frac{1}{\ell_0 L}e^{c\beta(\log L)^{\frac{p-1}{p}}}\Big)\,. \tag{8.19}$$

Since $\xi^{(p)}_{\ell_0,h} = (1 + o(1))\widehat{\pi}^{(p)}_\infty(\phi_o < -h)$ from Lemma 8.3, we get

$$\frac{\xi_n}{\widehat{\pi}^{(p)}_\infty(\phi_o < -h)} = \frac{\xi^{(p)}_{\ell_*,h}}{\widehat{\pi}^{(p)}_\infty(\phi_o < -h)} = (1 + L^{-\frac{1}{2}+\frac{\delta_p+o(1)}{2}})\frac{\xi^{(p)}_{\ell_0,h}}{\widehat{\pi}^{(p)}_\infty(\phi_o < -h)}\,;$$

whence Eq. (8.3) follows from the bounds on $\xi^{(p)}_{\ell_0,h}$ in Lemmas 8.3 and 8.5.

Similarly, Eq. (8.19) implies that $\bar{\xi}^{(p)}_{\ell_*,h}$ also satisfies $\bar{\xi}^{(p)}_{\ell_*,h} = \bar{\xi}^{(p)}_{\ell_0,h} + O(\frac{1}{\ell_0 L}e^{c\beta(\log L)^{\frac{p-1}{p}}})$, so that the bound $\bar{\xi}^{(p)}_{\ell_0,h} = (1 + o(1))\widehat{\pi}^{(p)}_\infty(\phi_o = -h)$ from Lemma 8.6 implies that

$$\rho_n = \frac{\bar{\xi}^{(p)}_{\ell_*,h}}{\widehat{\pi}^{(p)}_\infty(\phi_o = -h)} = (1 + L^{-\frac{1}{2}+\frac{\delta_p+o(1)}{2}})\frac{\bar{\xi}^{(p)}_{\ell_0,h}}{\widehat{\pi}^{(p)}_\infty(\phi_o = -h)}\,; \tag{8.20}$$

hence, Lemma 8.6 gives the following weak version of Eq. (8.4):

$$\rho^{(p)}_n \in [1 - e^{-\sqrt{\log L}}, 1 + L^{-1+o(1)}]\,. \qquad \blacksquare$$

We emphasize that the above calculations show the $1 < p < 2$ case is more delicate than the $p \geq 2$ case. Here the upper bound estimates on $\xi^{(p)}_n$ (and eventually $\rho^{(p)}_n$) feature an error term of $1 + L^{-\frac{1}{2}+\frac{\delta_p+o(1)}{2}}$ multiplying a bound of $1 - L^{-1/2+\delta_p+o(1)}$. The only reason the final upper bound stays strictly < 1 is because $L^{\delta_p/2} \ll L^{\delta_p}$. Moreover, recall we already had to sacrifice with a suboptimal lower bound in Lemma 8.3 to even have an error term this small to begin with (as otherwise the choice of ℓ_0 in the above proof would not be allowed). In comparison, for $p \geq 2$ we have an error term of $1 + L^{-1/4}$ multiplying a bound of $1 - L^{-o(1)}$, so the choice of $\ell_0 = L^{1/4}$ there is fairly arbitrary.

Proof of Theorem 8.2. The heavy lifting is done in obtaining Proposition 8.7; proving the theorem amounts to verifying that particular choices of ℓ_0 yield the desired statements. As in Proposition 8.7, we choose δ_p to be the minimum of the constant from Claim 8.4 and $\frac{1}{6}$.

Begin with $1 < p < 2$ and Eq. (8.7). Let $\ell_0 \approx L^{\frac{1}{3}+\frac{\delta_p}{2}}$ be the largest integer less than $L^{\frac{1}{3}+\frac{\delta_p}{2}}$ such that $\ell_0 + 10\lceil \log L\rceil$ divides ℓ_*. (The condition $\delta_p \leq \frac{1}{6}$ now ensures that $L^{\frac{1}{3}+\frac{\delta_p}{2}} \leq L^{\frac{1}{2}-\frac{\delta_p}{2}}$, so we stay in the regime of Proposition 8.7.) The assumption that $|\partial F| \vee |\partial V| \leq O(L^{\frac{2}{3}+\frac{\delta_p}{5}})$ ensures that $|F| \leq \ell_0 L e^{-c\beta(\log L)^{\frac{p-1}{p}}}$. Hence, we may appeal to Eq. (8.16) of Proposition 8.7 and check that ℓ_0 was carefully chosen so that $\Xi = o(1)$. The cost of moving from $\xi^{(p)}_{\ell_0,h}$ to $\xi^{(p)}_{\ell_*,h}$ with a $|F|$ prefactor in the exponent of Eq. (8.7) is also $\frac{|F|}{\ell_0 L}e^{c\beta(\log L)^{\frac{p-1}{p}}} = o(1)$.

For Eqs. (8.8) and (8.9) (still for $1 < p < 2$), take $\ell_0 \approx L^{1/2}e^{-\sqrt{\log L}}$ to be the largest integer less than $\ell_0 \approx L^{1/2}e^{-\sqrt{\log L}}$ such that $\ell_0 + 10\lceil \log L\rceil$ divides ℓ_*. We apply Proposition 8.7, and compute $\Xi \le \sqrt{L}e^{3\sqrt{\log L}}$. Finally, the cost of moving from $\xi^{(p)}_{\ell_0,h}$ to $\xi^{(p)}_{\ell_*,h}$ with a $|F|$ prefactor in the exponent is $\frac{|F|}{\ell_0 L}e^{c\beta(\log L)^{\frac{p-1}{p}}} \le \sqrt{L}e^{3\sqrt{\log L}}$.

The $p > 2$ case is essentially the same. We can take $\ell_0 \approx L^{1/3}e^{3c\sqrt{\beta\log L}}$ in the mesoscopic case and $\ell_0 \approx L^{1/2}e^{-\sqrt{\log L}}$ in the macroscopic case, where $\approx$ has the same meaning as above and the constant c is from Proposition 8.7. Then as above, one can verify that the assumptions of Proposition 8.7 are met, that Ξ satisfies the desired bounds (either $o(1)$ or $\sqrt{L}e^{(5+o(1))c\sqrt{\beta\log L}}$), and that the cost of moving from $\xi^{(p)}_{\ell_0,h}$ to $\xi^{(p)}_{\ell_*,h}$ is the same order as Ξ. ■

It remains to prove the upper bound of Eq. (8.4). Recall this only applies for $1 < p < 2$. We will need the following conditional large deviation estimate, the proof of which we relegate to Section 8.2 as the proof is very different from the $p = 2$ case. One should think of the following as a replacement for Proposition 3.10 that is weaker, but sufficient for our purposes.

Claim 8.8. *Let $1 < p < 2$. Fix a neighbor of the origin o'. There exists $\delta'_p > 0$ and h_0 such that for all $h \ge h_0$ and any $y \notin \{o, o'\}$,*

$$\widehat{\pi}^{(p)}_\infty(\phi_y = h \mid \phi_o = \phi_{o'} = h) \le e^{-\beta\delta'_p h^p}\,.$$

Then, define $\mathbf{P}^{(p)}_{\ell,h} := \widehat{\pi}^{(p)}_\infty(\cdot \mid \mathscr{F}_{\ell,h})$, where $\mathscr{F}_{\ell,h} := \{\phi_x \ge -h,\ \forall x \in Q_\ell\}$. This is motivated by the fact that we can then write $\bar\xi_{\ell,h} = -\frac{1}{\ell^2}\log \mathbf{P}_{\ell,h}(\phi_x > -h,\ \forall x \in Q_\ell)$.

Corollary 8.9 (Cf. Corollary 3.18)**.** *Fix $1 < p < 2$. Let $h = H + 1 - n$ for n fixed, and $\ell \le \sqrt{L}e^{-\sqrt{\log L}}$. If o' is a neighbor of the origin o, then for the constant δ_p from Claim 8.4, we have*

$$\mathbf{P}_{\ell,h}\left(\phi_{o'} = -h \mid \phi_o = -h\right) \ge 1 - L^{-1/2+\delta_p+o(1)}\,.$$

Proof. The proof remains unchanged from the $p = 2$ case, using Claim 8.4 and Claim 8.8 as inputs instead of Claim 3.6 and Proposition 3.10. ■

Lemma 8.10 (Cf. Lemma 3.19)**.** *Fix $1 < p < 2$, and let δ_p be from Claim 3.6. For $h = H + 1 - n$ with fixed $n \ge 0$, β sufficiently large, and $\ell \ge L^{\frac14}$, we have*

$$\frac{\bar\xi_{\ell,h}}{\widehat{\pi}_\infty(\phi_o = -h)} \le 1 - L^{-1/2+\delta_p+o(1)}\,.$$

Proof. The proof remains unchanged from the $p = 2$ case. ■

Proof of Proposition 8.1, upper bound of Eq. (8.4)*.* Plug in Lemma 8.10 into Eq. (8.20). ■

8.2. **Missing large deviation estimates for $p > 1$.** Finally, we conclude with the proofs of Claim 8.4 and Claim 8.8. We will start with Claim 8.4 for $p > 2$, which seeks to prove a lower bound on $\widehat{\pi}^{(p)}_\infty(\phi_{o'} \ge h \mid \phi_o \ge h)$. We recall from [14, Thm. 7.1, Eq. (7.4)] that, for some absolute constant $c_0 > 0$ and every $h \ge 1$,

$$\frac{\widehat{\pi}(\phi_o = h)}{\widehat{\pi}(\phi_o = h-1)} > \exp(-c_0\beta h)\,, \tag{8.21}$$

and from [14, Thm. 7.1, Eq. (7.4)], for another constant $c_1 > 0$ and every $h \ge 1$,

$$\widehat{\pi}^{(p)}_\infty(\phi_y = h \mid \phi_x = h) \le \exp(-c_1\beta h^{p/(p-1)})\,. \tag{8.22}$$

Proof of Claim 8.4, $p>2$. Let $M := (c_0/2)^{1/(p-1)}$ for the constant c_0 from Eq. (8.21), and define

$$q_M := \widehat{\pi}_\infty^{(p)}\big(\phi_{o'} \geq h - Mh^{1/(p-1)} \mid \phi_o \geq h\big)\,.$$

We will argue that

$$q_M > e^{-2c_0\beta h}\,. \tag{8.23}$$

To see this, assume the opposite, and let $\mathcal{S} = \big\{\phi \,:\, \max_{z\sim o}\phi_z \leq h - M^{1/(p-1)}\big\}$. By a union bound, we have $\widehat{\pi}_\infty^{(p)}(\mathcal{S}^c \mid \phi_o \geq h) \leq 4q_M$. Now consider the bijection that decreases ϕ_o by 1. For every $\phi \in \mathcal{S}\cap\{\phi_o \geq h+1\}$,

$$\sum_{z\sim o}(|\phi_o-\phi_z|^p - |\phi_o - 1 - \phi_z|^p) \geq \sum_{z\sim o} p|\phi_o - 1 - \phi_z|^{p-1} \geq 4pM^{p-1}h = 2c_0 h\,,$$

using the definition of M in the last step. Thus, $\widehat{\pi}_\infty^{(p)}(\mathcal{S},\, \phi_o \geq h+1 \mid \phi_o \geq h) \leq e^{-2c_0\beta h}$, and we can conclude that

$$\begin{aligned}\widehat{\pi}_\infty^{(p)}(\phi_o \geq h+1 \mid \phi_o \geq h) &\leq \widehat{\pi}_\infty^{(p)}(\mathcal{S},\, \phi_o \geq h+1 \mid \phi_o \geq h) + \widehat{\pi}_\infty^{(p)}(\mathcal{S}^c \mid \phi_o \geq h)\\ &\leq e^{-2c_0\beta h} + 4q_M\\ &\leq 6e^{-2c_0\beta h}\,.\end{aligned}$$

For large h, this would contradict the fact that $\widehat{\pi}_\infty^{(p)}(\phi_o \geq h+1)/\widehat{\pi}_\infty^{(p)}(\phi_o \geq h) \geq (1-\varepsilon_\beta)e^{-c_0\beta h}$ via Eq. (8.21) (as well as that $\widehat{\pi}_\infty(\phi_o = h) \geq (1-\varepsilon_\beta)\widehat{\pi}_\infty(\phi_o \geq h)$ by a routine Peierls map). This establishes Eq. (8.23).

To derive the sought bound from Eq. (8.23), we write

$$\begin{aligned}\widehat{\pi}_\infty^{(p)}(\phi_{o'} \geq h \mid \phi_o \geq h) &= q_M \prod_{k=h-Mh^{1/(p-1)}}^{h-1} \widehat{\pi}_\infty^{(p)}(\phi_{o'} \geq k+1 \mid \phi_o \geq h,\, \phi_{o'} \geq k)\\ &\geq q_M \exp\Big(-c_0\beta h\cdot Mh^{1/(p-1)}\Big)\\ &\geq \exp\Big(-(c_0M\beta - o(1))h^{p/(p-1)}\Big)\,,\end{aligned}$$

where the inequality in the second line used FKG to drop the conditioning on $\phi_o \geq h$, followed by an application of Eq. (8.21) to each of the terms in the product (where for each of these $Mh^{1/(p-1)}$ terms we further increase k to h). Note that, of the two terms in the second line, $q_M \geq e^{-2c_0\beta h}$ is negligible compared to the term $\exp(-c\beta h^{p/(p-1)})$. ∎

Next, to set up for the proofs of Claim 8.4 for $1<p<2$ and Claim 8.8, we begin by recalling some preliminaries about p-harmonic functions from nonlinear potential analysis. Define the p-Laplacian as $(\Delta_p\phi)_x = \frac14\sum_{y\sim x}|(\nabla\phi)_{xy}|^{p-1}(\nabla\phi)_{xy}$ for $(\nabla\phi)_{xy} = \phi_y - \phi_x$. Define the energy $\mathscr{E}^{(p)}(f) := \sum_e |\nabla f|^p$, and the p-capacity for $A\subset\mathbb{Z}^2$ by

$$\mathrm{Cap}_p(A) := \inf\{\mathscr{E}^{(p)}(f) :\ f\to 0 \text{ at } \infty,\ f\geq 1 \text{ on } A\}\,.$$

Now fix a finite set A. For $1<p<2$, the graph $\mathbb{Z}^2$ is p-hyperbolic (see, e.g. [36, pp. 176–178]), meaning $\mathrm{Cap}_p(\{A\})>0$. As a consequence, there is a unique p-harmonic function with boundary conditions 1 on A and 0 at infinity, call this ϕ_A^*. Moreover, we have $\mathrm{Cap}_p(A) = \mathscr{E}^{(p)}(\phi_A^*)$, and for any $\varepsilon>0$, there exists a radius r and an approximation φ_A^* of ϕ_A^* such that φ_A^* is supported on $U = \bigcup_{x\in A}B_r(x)$ and $\mathscr{E}(\varphi_A^*) \leq \mathscr{E}(\phi_A^*) + \frac{\varepsilon}{2}$.

By looking at the maximum of the p-harmonic functions, an elementary fact is that

$$\mathrm{Cap}_p(A\cup B) \leq \mathrm{Cap}_p(A) + \mathrm{Cap}_p(B)\,. \tag{8.24}$$

Moreover, the p-capacity for $1<p<2$ is asymptotically additive, i.e.,

$$\mathrm{Cap}_p(A\cup(B+x)) = \mathrm{Cap}_p(A) + \mathrm{Cap}_p(B) + o(1) \quad\text{as}\quad |x|\to\infty \tag{8.25}$$

(this follows from quasi-additivity of the p-capacity combined with decay estimates for p-Green functions; see, e.g., [22,30] for the general framework, and [23] for the graph setting).

It is known ([29, Thm. 5.1, Eq. (5.1)]) that for every $1 < p < 2$ there exists some $c_p^* > 0$ such that

$$\widehat{\pi}_\infty^{(p)}(\phi_o \geq h) = e^{-(c_p^*\beta + o(1))h^p}\,, \tag{8.26}$$

where the $o(1)$-term goes to 0 as $h \to \infty$. In fact, the proof shows that c_p^* is precisely $\mathscr{E}^{(p)}(\phi_o^*)$. The next lemma extends this one-point large deviation result to a general finite set.

Lemma 8.11. *Let $1 < p < 2$. For any constants $k, \varepsilon > 0$, there exists constants h_0 such that for all $h \geq h_0$, the following holds. Let $A \subset \mathbb{Z}^2$ have k vertices. Then,*

$$\exp(-\beta(\mathscr{E}^{(p)}(\phi_A^*) + \varepsilon)h^p) \leq \widehat{\pi}_\infty^{(p)}(\phi\restriction_A = h) \leq \exp(-\beta(\mathscr{E}^{(p)}(\phi_A^*) - \varepsilon)h^p)\,.$$

The same statement holds for $\widehat{\pi}_\infty^{(p)}(\phi\restriction_A \geq h)$.

Proof. As usual, an easy Peierls argument shows that $\widehat{\pi}_\infty^{(p)}(\phi\restriction_A = h \mid \phi\restriction_A \geq h) \geq 1 - \varepsilon_\beta$, so it suffices to prove the statement for $\widehat{\pi}_\infty^{(p)}(\phi\restriction_A = h)$. Call the k points of A by $\{x_i\}_{i=1}^k$. Let $R = Ch^{p-1}$ for a constant C to be chosen later. Let $V = \bigcup_i B_R(x_i)$. Let $\mathcal{E}$ be the event that there is no connected component of sites with height ≥ 1 with diameter (in $\mathbb{Z}^2$) larger than $R/2$ that intersects V. First we prove that

$$\widehat{\pi}_\infty^{(p)}(\phi\restriction_A = h, \mathcal{E}) \leq \exp(-\beta(\mathscr{E}^{(p)}(\phi_A^*) - o(1))h^p)\,. \tag{8.27}$$

For every $v \in \partial V$, reveal its connected component of sites with height ≥ 1. On the event $\mathcal{E}$, none of these component reach $\{x_i\}$. Thus, this reveals circuits of ≤ 0 sites around the $\{x_i\}$, where different circuits have disjoint interiors. Group together the points sharing the same circuit into $j \leq k$ groups A_j. Note that no information about heights inside the circuit has been revealed. Hence after the revealing, we can forget the event $\mathcal{E}$, increase the event $\phi\restriction_A = h$ to $\phi\restriction_A \geq h$, use monotonicity to raise the ≤ 0 sites to height 0, and use Domain Markov to obtain

$$\widehat{\pi}_\infty^{(p)}(\phi\restriction_A \geq h, \mathcal{E}) \leq \max_{\{V_j\}} \prod_j \widehat{\pi}_{V_j}^{(p)}(\phi\restriction_{A_j} \geq h)\,,$$

where the maximum is over all collections of disjoint subsets $V_j \subset V$ such that each $A_j \subset V_j$. To upper bound the terms $\widehat{\pi}_{V_j}^{(p)}(\phi\restriction_{A_j} \geq h)$, observe that each configuration attaining $\phi\restriction_{A_j} \geq h$ has probability at most $e^{-\beta\mathscr{E}^{(p)}(\phi_{A_j}^*)h^p}$. Hence we can write

$$\widehat{\pi}_{V_j}^{(p)}(\phi\restriction_{A_j} \geq h) \leq \widehat{\pi}_{V_j}^{(p)}(\max_{x \in A_j} |\phi_x| \geq h^2) + e^{-\beta\mathscr{E}^{(p)}(\phi_{A_j}^*)h^p}(2h^2+1)^{|V_j|} \leq e^{\beta(\mathscr{E}^{(p)}(\phi_{A_j}^*) - o(1))h^p}\,.$$

The bound Eq. (8.27) now follows by Eq. (8.24).

Now let Γ_i be the outermost 1 level line containing x_i. To conclude the desired upper bound of the lemma, it suffices to show that for some $C' > 0$, we have

$$\widehat{\pi}_\infty^{(p)}(\mathcal{E}^c \mid \phi\restriction_A = h) \leq \widehat{\pi}_\infty^{(p)}(\mathcal{E}^c \mid \max_i |\Gamma_i| \leq R,\, \phi\restriction_A = h) + \widehat{\pi}_\infty^{(p)}(\max_i |\Gamma_i| > R \mid \phi\restriction_A = h) \leq e^{-\beta C' h^{p-1}}\,. \tag{8.28}$$

The first inequality is trivial. An upper bound on $\widehat{\pi}_\infty^{(p)}(\max_i |\Gamma_i| > R \mid \phi\restriction_A = h)$ follows via essentially the same proof of [29, Lem. 5.3]. Firstly, for any neighbor y of A, we can bound $\widehat{\pi}_\infty^{(p)}(\phi_y \leq -h \mid \phi\restriction_A = h) \leq e^{-c\beta h^p}$ by a Peierls map on the down-loops surrounding x. We can afford this as an additive error. Otherwise, if all the neighbors of A have height $\geq -h$, then consider the Peierls map T which lowers the height of all vertices in $\bigcup_i \mathsf{Int}(\Gamma_i)$ by 1, and then raises the height of $\phi\restriction_A$ back up to h. The condition on the neighbors of A implies that we have an energy difference of

$$\widehat{\pi}_\infty^{(p)}(T(\phi)) \geq \widehat{\pi}_\infty^{(p)}(\phi) e^{\beta(|\bigcup_i |\Gamma_i| - |\partial A|\beta p(2h)^{p-1})}\,.$$

Hence, the standard Peierls argument enumerating over Γ_i shows that for a sufficiently large constant $C > 0$ depending on $|\partial A|$, we have $\widehat{\pi}_\infty^{(p)}(\max_i |\Gamma_i| \geq Ch^{p-1} \mid \phi\restriction_A = h) \leq e^{C'\beta h^{p-1}}$ for some other constant C'. Finally, an upper bound of the same form on $\widehat{\pi}_\infty^{(p)}(\mathcal{E}^c \mid \max_i |\Gamma_i| \leq R, \phi\restriction_A = h)$ follows by observing that the large ≥ 1 component realizing $\mathcal{E}^c$ must be in the exterior of all the Γ_i, so revealing the Γ_i reveals ≤ 0 boundary conditions and $\mathcal{E}^c$ can be ruled out by a Peierls argument.

We turn now to showing the lower bound of the lemma. For the choice of ε in the lemma, there exists a radius r and an approximation φ_A^* of ϕ_A^* such that φ_A^* is supported on $U = \bigcup_{x\in A} B_r(x)$ and $\mathscr{E}(\varphi_A^*) \leq \mathscr{E}(\phi_A^*) + \frac{\varepsilon}{2}$. The maximum value attained by $h\varphi_A^*$ is h, so we have an integer rounding cost of $\mathscr{E}(\lfloor h\varphi_A^*\rfloor) \leq \mathscr{E}(h\varphi_A^*]) + h^{p-1}r^2$. At the same time, forcing all heights in U to be 0 under $\widehat{\pi}_\infty^{(p)}$ has a probability of at least $(1-\varepsilon_\beta)^{|U|}$ by absolute value FKG (valid for $1 \leq p \leq 2$). The weight of a specific configuration ϕ in U, relative to the all-0 configuration, is simply $\exp(-\beta E(\phi))$. Hence, in total we have that $\widehat{\pi}_\infty^{(p)}(\phi\restriction_A = h) \geq \exp(-\beta(\mathscr{E}(\phi_A^*) + \frac{\varepsilon}{2})h^p - \beta h^{p-1}|U| - \varepsilon_\beta |U|)$, where the extra error terms can be absorbed into another factor of $\beta\frac{\varepsilon}{2}h^p$ for sufficiently large h since $|U| \leq kr^2$. ■

Proof of Claim 8.4, $1 < p < 2$. Consider the p-harmonic function ϕ_o^* giving the one point large deviation. The energy $\mathscr{E}^{(p)}(\phi_o^*)$ is the sum of the energy over horizontal bonds $\mathscr{E}_{\mathsf{h}}^{(p)}(\phi_o^*)$ and vertical bonds $\mathscr{E}_{\mathsf{v}}^{(p)}(\phi_o^*)$. W.l.o.g, assume that $\mathscr{E}_{\mathsf{v}}^{(p)}(\phi_o^*) \leq \mathscr{E}_{\mathsf{h}}^{(p)}(\phi_o^*)$. Define φ via block-copying ϕ_o^* on $2\mathbb{Z}\times\mathbb{Z}$, that is,

$$\varphi_{2k,l} = \varphi_{2k+1,l} = (\phi_o^*)_{k,l} \qquad \text{for all } k,l \in \mathbb{Z}\,.$$

Then

$$\mathscr{E}^{(p)}(\varphi) = \mathscr{E}_{\mathsf{h}}^{(p)}(\phi_o^*) + 2\mathscr{E}_{\mathsf{v}}^{(p)}(\phi_o^*) \leq \tfrac{3}{2}\mathscr{E}^{(p)}(\phi_o^*)\,.$$

Since φ takes values 1 on o, o' and 0 at infinity, then $\mathscr{E}^{(p)}(\phi_{o,o'}^*) \leq \mathscr{E}^{(p)}(\varphi)$. We next verify that $\varphi \neq \phi_{o,o'}^*$, so that there exists some $\delta_p > 0$ such that

$$\mathscr{E}^{(p)}(\phi_{o,o'}^*) \leq \mathscr{E}^{(p)}(\varphi) - \delta_p \leq \tfrac{3}{2}\mathscr{E}^{(p)}(\phi_o^*) - \delta_p\,. \tag{8.29}$$

We need to verify that the function φ is not p-harmonic. One elementary way to show this is to take $(k,l) \neq o$, let $x^- = (k-1,l)$, $x = (k,l)$, $x^+ = (k+1,l)$, and observe that if $(\Delta_p\varphi)_{2k,l} = 0$ then $|(\nabla\phi_o^*)_{xx^-}|^{p-2}(\nabla\phi_o^*)_{xx^-} = -\xi$, where ξ is the contribution of the vertical bonds to $(\Delta_p\phi_o^*)_x$, and similarly, if $(\Delta_p\varphi)_{2k+1,l} = 0$ then $|(\nabla\phi_o^*)_{xx^+}|^{p-2}(\nabla\phi_o^*)_{xx^+} = -\xi$. Thus, $(\Delta_p\phi_o^*)_x = -\xi$, and since ϕ_o^* is p-harmonic, we get $(\nabla\phi_o^*)_{x^-x} = (\nabla\phi_o^*)_{xx^+} = 0$. Iterating this argument (as well as in the vertical axes), while recalling that ϕ_o^* is 0 at infinity, we find $\phi_o^* = \delta_o$, a contradiction.

The proof now concludes by combining Eq. (8.29) and Lemma 8.11. ■

Proof of Claim 8.8. The statement is immediately true by Lemma 8.11 if δ_p can depend on y, in particular with $\delta_p(y) := \mathscr{E}^{(p)}(\phi_{o,o',y}^*) - \mathscr{E}^{(p)}(\phi_{o,o'}^*) - o(1)$. So it remains to show that $\mathscr{E}^{(p)}(\phi_{o,o',y}^*) - \mathscr{E}^{(p)}(\phi_{o,o'}^*)$ is uniformly bounded away from 0, whence we can take h large enough so that the $o(1)$ is negligible. This in turn follows from Eq. (8.25) combined with the observation that for any $y \notin \{o,o'\}$, we always have $\mathscr{E}^{(p)}(\phi_{o,o',y}^*) > \mathscr{E}^{(p)}(\phi_{o,o'}^*)$. ■

8.3. **Extension to $p = 1$.** Finally, we discuss the SOS model. Consider the summary at the beginning of this section. The crucial input needed for the proofs of the main theorems is an analog of Theorem 3.2. But for $p = 1$ this was already proven, with the mesoscopic range in [13, Prop. A.1] and the macroscopic range in [13, Prop. 2.12]. (For $p = 1$, there is no need to define ξ or ρ; one may work directly in terms of $\widehat{\pi}_\infty^{(1)}(\phi_o < -h)$.) Hence, the results of Theorems 1.1 and 1.2 also follow. However, Theorem 1.3 does not follow because the quantities $\widehat{\pi}_\infty^{(p)}(\phi_o = h)$ and $\widehat{\pi}_\infty^{(p)}(\phi_o = h-1)$ are now comparable up to a factor of $C(\beta)$, instead of being on different scales, the latter of which is a crucial part of the proof of [14] which Section 7 is modifying. (In fact the limit law for $p = 1$ is not expected to be Ferrari–Spohn, see, e.g., [14, §1] for more details).

Acknowledgments. This research was supported by the NSF grant DMS-2451083.

References

[1] M. Aizenman, M. Harel, R. Peled, and J. Shapiro. Depinning in integer-restricted Gaussian fields and BKT phases of two-component spin models, 2021. Preprint, arXiv:2110.09498.
[2] M. Basu Roy Chowdhury, P. Caputo, and S. Ganguly. Characterizing Gibbs states for area-tilted Brownian lines. *Ann. Probab.* To appear.
[3] V. L. Berezinskiĭ. Destruction of long-range order in one-dimensional and two-dimensional systems having a continuous symmetry group. I. Classical systems. *Sov. Phys. JETP*, (32):493–500, 1971.
[4] R. Brandenberger and C. E. Wayne. Decay of correlations in surface models. *J. Statist. Phys.*, 27(3):425–440, 1982.
[5] J. Bricmont, A. El Mellouki, and J. Fröhlich. Random surfaces in statistical mechanics: roughening, rounding, wetting,. . . . *J. Statist. Phys.*, 42(5-6):743–798, 1986.
[6] W. K. Burton, N. Cabrera, and F. C. Frank. The growth of crystals and the equilibrium structure of their surfaces. *Philos. Trans. Roy. Soc. London Ser. A*, 243:299–358, 1951.
[7] P. Caddeo, Y. H. Kim, and E. Lubetzky. On level line fluctuations of SOS surfaces above a wall. *Forum Math. Sigma*, 12:Paper No. e91, 59, 2024.
[8] P. Caputo and S. Ganguly. Uniqueness, mixing, and optimal tails for Brownian line ensembles with geometric area tilt. *Probab. Math. Phys.*, 6(1):195–239, 2025.
[9] P. Caputo, D. Ioffe, and V. Wachtel. Confinement of Brownian polymers under geometric area tilts. *Electron. J. Probab.*, 24:Paper No. 37, 21, 2019.
[10] P. Caputo, D. Ioffe, and V. Wachtel. Tightness and line ensembles for Brownian polymers under geometric area tilts. In *Statistical mechanics of classical and disordered systems*, volume 293 of *Springer Proc. Math. Stat.*, pages 241–266. Springer, Cham, 2019.
[11] P. Caputo, E. Lubetzky, F. Martinelli, A. Sly, and F. L. Toninelli. The shape of the $(2+1)D$ SOS surface above a wall. *C. R. Math. Acad. Sci. Paris*, 350(13-14):703–706, 2012.
[12] P. Caputo, E. Lubetzky, F. Martinelli, A. Sly, and F. L. Toninelli. Dynamics of $(2+1)$-dimensional SOS surfaces above a wall: Slow mixing induced by entropic repulsion. *Ann. Probab.*, 42(4):1516–1589, 2014.
[13] P. Caputo, E. Lubetzky, F. Martinelli, A. Sly, and F. L. Toninelli. Scaling limit and cube-root fluctuations in SOS surfaces above a wall. *J. Eur. Math. Soc. (JEMS)*, 18(5):931–995, 2016.
[14] J. Chen and E. Lubetzky. The limiting law of the Discrete Gaussian level lines, 2025. Preprint, arXiv:2509.04333.
[15] S. T. Chui and J. D. Weeks. Phase transition in the two-dimensional coulomb gas, and the interfacial roughening transition. *Phys. Rev. B*, 14:4978–4982, Dec 1976.
[16] A. Dembo, E. Lubetzky, and O. Zeitouni. On the limiting law of line ensembles of Brownian polymers with geometric area tilts. *Ann. Inst. Henri Poincaré Probab. Stat.*, 60(1):113–125, 2024.
[17] P. L. Ferrari and H. Spohn. Constrained Brownian motion: fluctuations away from circular and parabolic barriers. *Ann. Probab.*, 33(4):1302–1325, 2005.
[18] J. Fröhlich and T. Spencer. Kosterlitz-Thouless transition in the two-dimensional plane rotator and Coulomb gas. *Phys. Rev. Lett.*, 46(15):1006–1009, 1981.
[19] J. Fröhlich and T. Spencer. The Kosterlitz-Thouless transition in two-dimensional abelian spin systems and the Coulomb gas. *Comm. Math. Phys.*, 81(4):527–602, 1981.
[20] C. Garban and A. Sepúlveda. Statistical reconstruction of the GFF and KT transition. *Journal of the European Mathematical Society*, 26(2):639–694, 2023.
[21] M. Hegde, Y. H. Kim, and C. Serio. Scaling limit and tail bounds for a random walk model of SOS level lines, 2025. Preprint, arXiv:2502.10384.
[22] J. Heinonen, T. Kilpeläinen, and O. Martio. *Nonlinear Potential Theory of Degenerate Elliptic Equations.* Oxford Mathematical Monographs. Oxford University Press, Clarendon Press, Oxford, 1993.
[23] I. Holopainen and P. M. Soardi. p-harmonic functions on graphs and manifolds. *Manuscripta Mathematica*, 94(1):95–110, 1997.
[24] D. Ioffe, S. Shlosman, and Y. Velenik. An invariance principle to Ferrari-Spohn diffusions. *Comm. Math. Phys.*, 336(2):905–932, 2015.
[25] J. M. Kosterlitz and D. J. Thouless. Ordering, metastability and phase transitions in two-dimensional systems. *Journal of Physics C: Solid State Physics*, 6(7):1181–1203, 1973.
[26] P. Lammers. Height function delocalisation on cubic planar graphs. *Probab. Theory Related Fields*, 182(1-2):531–550, 2022.
[27] P. Lammers and S. Ott. Delocalisation and absolute-value-FKG in the Solid-On-Solid model. *Probab. Theory Related Fields*, 188(1-2):63–87, 2024.

[28] G. F. Lawler. *Intersections of random walks*. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2013. Reprint of the 1996 edition.

[29] E. Lubetzky, F. Martinelli, and A. Sly. Harmonic pinnacles in the discrete Gaussian model. *Comm. Math. Phys.*, 344(3):673–717, 2016.

[30] V. Maz'ya. *Sobolev Spaces: With Applications to Elliptic Partial Differential Equations*, volume 342 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin/Heidelberg, 2 edition, 2011.

[31] S. Ott and F. Schweiger. Quantitative delocalization for Solid-On-Solid models at high temperature and arbitrary tilt, 2025. Preprint, `arxiv:2505.16804`.

[32] R. H. Schonmann and S. B. Shlosman. Complete analyticity for 2D Ising completed. *Comm. Math. Phys.*, 170(2):453–482, 1995.

[33] R. H. Schonmann and S. B. Shlosman. Constrained variational problem with applications to the Ising model. *Journal of statistical physics*, 83(5):867–905, 1996.

[34] C. Serio. Scaling limit for line ensembles of random walks with geometric area tilts. *Electron. J. Probab.*, 28:Paper No. 124, 14, 2023.

[35] S. Sheffield. Random surfaces. *Astérisque*, (304):vi+175, 2005.

[36] P. M. Soardi. *Potential theory on infinite networks*. Springer, 2006.

[37] F. Spitzer. *Principles of random walk*, volume Vol. 34 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Heidelberg, second edition, 1976.

[38] D. van Engelenburg and M. Lis. An elementary proof of phase transition in the planar XY model. *Comm. Math. Phys.*, 399(1):85–104, 2023.

[39] J. Villain. Theory of one- and two-dimensional magnets with an easy magnetization plane. II. The planar, classical, two-dimensional magnet. *Journal de Physique*, 36(6):581–590, 1975.

J. Chen

LPSM, Sorbonne Université, 4 Place Jussieu 75252, Paris, France.

Email address: `jchen@lpsm.paris`

E. Lubetzky

Courant Institute, New York University, 251 Mercer Street, New York, NY 10012, USA.

Email address: `eyal@courant.nyu.edu`