# LOGARITHMIC DELOCALIZATION OF LOW TEMPERATURE 3D ISING AND POTTS INTERFACES ABOVE A HARD FLOOR

JOSEPH CHEN, REZA GHEISSARI, AND EYAL LUBETZKY

ABSTRACT. We study the entropic repulsion of the low temperature 3D Ising and Potts interface in an $n \times n \times n$ box with blue boundary conditions on its bottom face (the hard floor), and red boundary conditions on its other five faces. For Ising, Fröhlich and Pfister proved in 1987 that the typical interface height above the origin diverges (non-quantitatively), via correlation inequalities special to the Ising model; no such result was known for Potts. We show for both the Ising and Potts models that the entropic repulsion fully overcomes the potentially attractive interaction with the floor, and obtain a logarithmically diverging lower bound on the typical interface height. This is complemented by a conjecturally sharp upper bound of $\lfloor \xi^{-1} \log n \rfloor$ where $\xi$ is the rate function for a point-to-plane non-red connection under the infinite volume red measure. The proof goes through a coupled random-cluster interface to overcome the potentially attractive interaction with the boundary, and a coupled fuzzy Potts model to reduce the upper bound to a simpler setting where the repulsion is attained by conditioning a no-floor interface to lie in the upper half-space.

## 1. INTRODUCTION

The Potts model on a finite graph $G = (V, E)$ is the following probability distribution over $\sigma \in \{1, \ldots, q\}^V$ (coloring each site in $V$ by one of $q$ colors),

$$\mu_{G,\beta,q}(\sigma) \propto \exp\left(-\beta \sum_{(u,v) \in E} \mathbf{1}_{\{\sigma_v \neq \sigma_w\}}\right), \tag{1.1}$$

with $\beta > 0$ the inverse-temperature (in this paper, a large constant). The case $q = 2$ is the Ising model. A boundary condition for the model is an assignment $\eta$ of specific values to a subset of the sites $S \subset V$, whence one considers Eq. (1.1) restricted to configurations agreeing with $\eta$ on $S$ (equivalently, one looks at $\mu_{G,\beta,q}$ conditional on $\sigma$ agreeing with $\eta$ on $S$). We will be interested in the $q$-state Potts model on the $n \times n \times n$ cube $\Lambda_n$ with boundary conditions, denoted by fl, that are blue boundary conditions on its bottom face (the *hard floor*), and red boundary conditions on the other five faces. We denote this distribution by $\mu_{\Lambda_n}^{\mathsf{fl}}$.

These boundary conditions induce a blue-to-non-blue interface, denoted here by $\mathcal{I}_{\mathsf{blue}}$, consisting of the set of plaquettes separating vertices in the blue phase (either in the blue component of the boundary, or in non-blue bubbles encapsulated by that component), from the rest (see Definition 2.1 for the formal definition). (The plaquettes of the interface are unit squares bounded by vertices of $\mathbb{Z}^3$, and two plaquettes are considered adjacent if they share an edge.) Regard the height of each plaquette $f \in \mathcal{I}_{\mathsf{blue}}$ as the difference of the $\mathfrak{e}_3$-coordinate of its midpoint and the $\mathfrak{e}_3$-coordinate of the floor, which we view as having height zero, and denote it by $\mathrm{ht}(f)$. In this paper, we study the typical height of plaquettes in the bulk of the interface $\mathcal{I}_{\mathsf{blue}}$, as it balances competing forces of rigidity (which would hold at large $\beta$ if instead the boundary conditions changed color in the middle of the box), entropic repulsion away from the hard floor, and a complicated interaction between the interface and the floor through the potential finite bubbles of non-blue below the interface that the hard floor precludes.

The analysis of low-temperature 3D Ising interfaces has a long history, dating fifty years to the classical work of Dobrushin [6] who established rigidity of the interface when there is no nearby floor; namely, one imposes on the $n \times n \times n$ cube $\Lambda_n$ *Dobrushin boundary conditions*, denoted here by dob, that are blue on its lower half and red on its top half. This spurred several other works on rigidity (localization) of integer-valued surfaces at low-temperatures, including the order-order interface of

the super-critical random-cluster model and interfaces of low-temperature Potts model [14] under $\mu_{\Lambda_n}^{\mathsf{dob}}$. The behavior of localized interfaces in the presence of a nearby floor (sometimes called a wall) introduces a competition between the forces preferring rigidity and *entropic repulsion*, a desire to delocalize from the floor to make room for thermal fluctuations that would otherwise be precluded. The work of Fröhlich and Pfister [7, 8] studied this and related surface transitions like wetting and layering rigorously for the Ising model. Using correlation inequalities special to the Ising model, they established non-quantitative delocalization of the Ising interface in $\mu_{\Lambda_n}^{\mathsf{fl}}$; i.e., the statement that the typical height of the interface above the origin diverges with $n$ (at an unknown rate). To our knowledge, this remains the only rigorous bound on the Ising interface height above a hard floor (see Section 1.2 for logarithmic bounds in simpler settings), and not even a non-quantitative delocalization was known for the Potts interface.

Rigorous quantitative results on entropic repulsion have generally been restricted to models only on interfaces (as opposed to entire spin configurations) and that are height functions (no overhangs). The model closest to the Ising/Potts interfaces is the—well-studied in its own right— Solid-On-Solid (SOS) model. There, the picture of entropic repulsion at low-temperatures is now precisely understood; the SOS interface in $\Lambda_n$ with 0 boundary conditions and nonnegative heights is such that all but $\varepsilon_\beta n^2$ many plaquettes of the interface are at height $\lfloor \frac{1}{4\beta} \log n \rfloor + O(1)$ [2] with more precise quantities like the shape and fluctuations of the level curve for this highest height themselves being the subject of much rich investigation [1, 3]. Entropic repulsion, and the ensuing shape of the interface above a hard floor, have also been studied extensively for related height function models like [21]; see the survey [20] for more rich interface phenomena in SOS and related height function models. One element of all such analyses is a good understanding of the exponential rates for typical and maximal height deviations of the rigid surface before placing it near a floor.

In recent years, the precise law of height deviations, and maximum height, in the 3D Ising interface itself were derived in [10, 11], and for the random-cluster and Potts interfaces in [5]. In the Ising case, using this, a phenomenon of delocalization due to entropic repulsion was sharply studied when the repulsion is generated by what we will refer to as a *soft floor*, where the model has no boundary condition floor, but the interface is conditioned to stay above height zero [13]. This may feel qualitatively the same, and in the case of the height function models like SOS it is indeed the same. However, the boundary conditions at height zero are an $\exp(cn^2)$ tilt of the measure without them, and the interactions of the interface to the boundary conditions on the floor are non-explicit and quite complicated. Moreover, since the interface is uniformly within height $O(\log n)$ of the floor, these interactions are on the same order as the entropy-induced repulsion effects and compete. In particular, these interactions could pin the surface to the floor and prevent it rising to logarithmic heights. We describe this and how we overcome it in more detail in Section 1.2.

Our main results concern the low-temperature Ising and Potts interfaces in $\mu_{\Lambda_n}^{\mathsf{fl}}$, i.e., with a "hard floor" formed by boundary conditions at height zero. We show a logarithmically growing lower bound on the typical height of $\mathcal{I}$, together with a conjecturally sharp upper bound. Moreover, as $\beta \to \infty$, the ratio of the lower and upper bounds goes to 1. Our analyses require moving back and forth between Ising and Potts, a coupled random-cluster model, and the fuzzy Potts model [23], the latter two giving certain advantages in dealing with the possible (attractive and repulsive, respectively) interactions of the interface with the hard floor at height zero.

1.1. **Main result.** Under the infinite-volume measure $\mu_{\mathbb{Z}^3}^{\mathsf{red}}$, at large enough $\beta$, there is almost surely an infinite connected component of red vertices, call it $\mathcal{V}_{\mathsf{red}}$, and all connected components of $\mathcal{V}_{\mathsf{red}}^c = \mathbb{Z}^3 \setminus \mathcal{V}_{\mathsf{red}}$ are finite. See, e.g., [15, Chapter 7] for more on the $\beta$ large regime of the Potts model in $d \geq 3$. Define the finite point-to-plane connectivity rate by

$$\xi_h = \xi_h(q, \beta) := -\log \mu_{\mathbb{Z}^3}^{\mathsf{red}}\big((0, 0, 0) \longleftrightarrow \mathbb{R}^2 \times \{h\} \text{ in } \mathcal{V}_{\mathsf{red}}^c\big). \tag{1.2}$$

As we will see in Section 4, a sub-additivity argument using Fekete's Lemma can show that

$$\lim_{h \to \infty} \frac{\xi_h}{h} =: \xi = \xi(q, \beta) \tag{1.3}$$

exists (in fact, we show in Lemma 4.6 that $\xi_h$ is approximately additive, so that $\xi_h = \xi h + O(1)$ for all $h$). A simple Peierls argument (and the fact that such a non-red connection implies a finite random-cluster component of size $h$)—see, e.g., [15, Thm 7.3.2]—implies that $\xi \in [4\beta - C, 4\beta + C]$ for all $h \geq 1$, for an absolute constant $C > 0$. With the rate quantity $\xi$ in hand, we define

$$h_n^* = h_n^*(q, \beta) := \lfloor \xi^{-1} \log n \rfloor. \tag{1.4}$$

Our main result is the following.

**Theorem 1.1.** *Fix $q \geq 2$ and $\beta$ large enough. Consider $\mu_{\Lambda_n}^{\mathsf{fl}}$, the $q$-state Potts model on the $n \times n \times n$ box $\Lambda_n$ with boundary conditions* blue *on its bottom face and* red *elsewhere, and let $\mathcal{I}_{\mathsf{blue}}$ be the interface separating the* blue *and non-*blue *phases as defined above (formally defined in Definition 2.1). There exists a sequence $\varepsilon_\beta \downarrow 0$ as $\beta \uparrow \infty$ such that with probability $1 - o(1)$,*

$$|\{f \in \mathcal{I}_{\mathsf{blue}} : \mathrm{ht}(f) \notin [(1 - \varepsilon_\beta)h_n^*, h_n^*]\}| \leq \varepsilon_\beta n^2.$$

*In addition, the same bound holds for the interface $\mathcal{I}_{\mathsf{red}}$ separating the* red *and non-*red *phases.*

As mentioned, in terms of lower bounds on the interface height with a hard floor, to the best of our knowledge in the $q = 2$ case there was only the non-quantitative bound of [7], and in Potts for $q \geq 3$, there was no proof of delocalization due to the entropic repulsion, even non-quantitatively.

The upper bound of Theorem 1.1 is conjecturally sharp. Indeed, for general $q \geq 2$, the height $h_n^*$ is exactly half of the maximum height oscillation $M_n = \max\{|\mathrm{ht}(f)| : f \in \mathcal{I}_{\mathsf{blue}}\}$ under the no-floor Dobrushin boundary condition measure $\mu_{\Lambda_n}^{\mathsf{dob}}$ as identified in [5,10]. This matches the scaling relation between the repelled typical height, and no-floor maximum height in the SOS approximation [2]. To be precise, [5, Theorem 1.3] identified $M_n$ less explicitly in terms of a different rate, which they called the pillar rate. Part of our work in this paper (Theorem 4.1) identifies that pillar rate with $\xi_h$. This implies that under $\mu_{\Lambda_n}^{\mathsf{dob}}$, one has $M_n = 2h_n^* + O(1)$ with probability $1 - o(1)$.

**Remark 1.2.** Our proofs go through an analysis of the random-cluster model, in order to handle the interaction with the hard floor—this is even needed in the special case $q = 2$ (the Ising model) where the model has FKG. Specifically, the interaction between the random-cluster interface $\mathcal{I}_{\mathsf{top}}$ (the boundary of the connected random-cluster component of the red boundary sites; see Definition 2.1) and the hard floor is amenable to rigorous analysis in a way Ising and Potts interactions are not.

As a consequence of Lemma 2.27, with high probability the typical height of $\mathcal{I}_{\mathsf{top}}$ (and the analogously defined random-cluster interface $\mathcal{I}_{\mathsf{bot}}$), and the typical height of the Potts interface $\mathcal{I}_{\mathsf{blue}}$ are the same. As a byproduct, Theorem 1.1 also holds (with the same $h_n^*$) for the interfaces $\mathcal{I}_{\mathsf{top}}, \mathcal{I}_{\mathsf{bot}}$. At the same time, the intuition behind the entropic repulsion phenomenon would suggest that the interface $\mathcal{I}_{\mathsf{bot}}$ would rise to height $\xi_{\mathsf{bot}}^{-1} \log n$, where $\xi_{\mathsf{bot}}$ is the relevant random-cluster connection rate. Perhaps surprisingly, the upper bound of $h_n^*$ proves that the typical height of $\mathcal{I}_{\mathsf{bot}}$ is $\Omega(\log n)$ smaller than this prediction (as $\xi_{\mathsf{bot}}^{-1}$ is known to be strictly smaller than $\xi$ by [5]).

**Remark 1.3.** The rate $\xi$ governing Eq. (1.4) was the large deviation of a point-to-plane connection probability; for the Ising model, it can be shown (albeit for a slightly different notion of adjacency) that it is equivalent to a point-to-point rate, e.g., using the CLT in [11, Cor. 3]. We expect the same to hold for the notion of adjacency used here, as well as for the Potts model for any $q \geq 3$.

**Remark 1.4.** Our lower bound of $(1 - \varepsilon_\beta)h_n^*$ in Theorem 1.1 is of the form $(1 - C/\beta)h_n^*$. See Remark 2.26 for a discussion on a sketch of how this could be improved to the point-to-plane random-cluster connection rate in the complement of the (a.s. unique) infinite component. This would asymptotically be $(1 - (C/\beta)e^{-\beta})h_n^*$, but that would be the limit of our approach.

1.2. **Further prior work and proof ideas.** Let us begin with a history of quantitative lower bounds for entropic repulsion of Ising-style interfaces that simplify some aspect to not have the delicate potentially attractive interaction to the boundary of the hard-floor case. In 1991, Meličerčík [24] studied a distribution over Ising-type interfaces (allowing overhangs), but with the distribution only being over the interface, not having any finite bubbles above and below the interface, and established that the interface rises to height order $\log n$ when conditioned to be non-negative. The presence of the full spin configurations (including the possibility of finite bubbles below the interface) causes a well-known difficulty for studying Ising interfaces in the presence of nearby boundary conditions due to complicated, possibly attractive, interactions between the interface and the boundary floor. (In $d = 2$ where the phenomenology is very different, the works [18, 19] required significant ideas to deal with the potential for these interactions to attract the interface to the boundary floor. The latter of these works showed first that the interface is typically at least distance $n^\varepsilon$ away from the floor, whence the interactions become negligible. However, in the $d = 3$ case where the repulsion is only of order $\log n$, there is no hope for such an approach to controlling the interactions.) Holický and Zahradník [17] sketched in 1993 how some of the steps in [24] can be extended to incorporate the interactions in the true 3D Ising measure, postponing the details to a future version (but to our knowledge there was still no rigorous proof of this even in this Ising case). The recent work [13] proved an entropic repulsion transition for the interface in the true 3D Ising measure (including the bubbles) at height $h_n^*$ when the floor is soft, i.e., where the interface is conditioned to be above height zero. Effectively, this incorporated self-interactions of the interface through the bubbles, but still did not have to handle interactions with a nearby hard floor.

Dealing with the potentially attractive interactions between the interface and the hard floor is a significant part of establishing our lower bound of $(1 - \varepsilon_\beta)h_n^*$ and is achieved by instead working with a certain "top" subset $\mathcal{I}_{\mathsf{top}}$ of a coupled random-cluster model's order-order interface (see Definition 2.1). For this specific marginal of the random-cluster interface, the interaction with the boundary is approximately in the right direction, in that it is less costly to apply maps that lift the interface in the presence of the hard floor, than in its soft-floor analogue. The reason this lower bound does not match $h_n^*$ is because the rates for downwards oscillations of $\mathcal{I}_{\mathsf{top}}$ are the source of our lower bound, but are distinct from those of $\mathcal{I}_{\mathsf{blue}}$ that we expect to govern the repulsion in $\mu_{\Lambda_n}^{\mathsf{fl}}$. This lower bound (and a logarithmic order upper bound) are proven in Section 2. Notably, these proofs of logarithmic order delocalization are essentially self-contained and do not use any of the more combinatorial walls, ceilings, and pillars machinery core to recent work in the subject.

The proof of the conjecturally sharp upper bound of $h_n^*$ goes via a different comparison to the soft-floor model, directly in the spin space. This comparison shows that the complicated interactions through finite bubbles between the interface $\mathcal{I}_{\mathsf{blue}}$ and the hard floor do not induce any additional repulsion beyond that visible in the soft-floor model. In the Ising case, this follows from a sequence of delicate monotonicity relations between the soft-floor and hard-floor interfaces. In the Potts case, we first build on [5] to get sharp large deviation bounds on maximal oscillations of a rigid Potts interface, and get the $h_n^*$ upper bound in the soft-floor model. Notably, we then replace the Ising model's monotonicity step with a monotonicity step in the coupled *fuzzy Potts model* where all non-blue vertices are relabeled as fuzzy. Since the interface $\mathcal{I}_{\mathsf{blue}}$ is measurable with respect to the coupled fuzzy Potts model and monotone in the set of blue vertices, this allows for an application of the FKG inequalities developed for the fuzzy Potts model in [4]. The comparison between the soft and hard floor measures is in Section 2.2 and the proof of the sharp upper bound is in Section 3.

The above arguments would establish Theorem 1.1 with $h_n^*$ defined less explicitly in terms of a rate $\alpha_h$ (3.2) for a certain *pillar* of the Dobrushin interface to exceed some height, in place of $\xi_h$. Section 4 establishes equivalence between $\alpha_h$ and the finite point-to-plane connectivity rate of $\xi_h$, and furthermore shows that the latter is additive so that $h_n^*$ has the nice characterization of (1.4) with the limiting $\xi$. This goes via the random-cluster representation of the interface and certain swapping operations to show the equivalences of rates on the appropriate high probability events

for the connections. As mentioned, this equivalence has the added benefit of refining the previous tightness results of the no-floor maximum oscillation $M_n$ of $\mathcal{I}_{\mathsf{blue}}$ from [5] to become tightness around the explicit value $2\lfloor \xi^{-1} \log n \rfloor = 2h_n^*$.

## 2. Logarithmic upper and lower bounds

The main goal of this section is to establish the lower bound of $(1 - \varepsilon_\beta)h_n^*$ for the typical height of the plaquettes of $\mathcal{I}_{\mathsf{blue}}$. As mentioned in Section 1.2, we approach this via a coupled random-cluster model where the bottom boundary face of $\Lambda_n$ is conditioned not to be connected to the five other boundary faces, inducing an interface $\mathcal{I}_{\mathsf{full}}$ of plaquettes dual to closed edges separating the two components. (By a simple energy-entropy tradeoff, when $\beta$ is large, the number of plaquettes in interfaces will not exceed $(1 + \varepsilon_\beta)n^2$, and so most of the plaquettes of $\mathcal{I}_{\mathsf{blue}}$ coincide with the plaquettes of $\mathcal{I}_{\mathsf{top}}$ (which bounds the open connected component of the five red boundary faces). Thus, lower and upper bounds on the typical height of one of these interfaces also bound the other.)

Since this section contains many of the new ideas in the paper to control the interactions of interfaces with the hard floor in the two possible directions (attractive and repulsive), we sketch the logic of the argument below.

(1) *Control of attractive interaction with hard floor via $\mathcal{I}_{\mathsf{top}}$.* When seeking a lower bound on the interface height, we are concerned about potentially attractive interactions between the interface and the hard floor. In the random-cluster world, these are expressed in terms of differences in probabilities, under a soft-floor interface, of the bottom face boundary edges all being open (to force a hard floor): see Lemma 2.18. If we *only* reveal the $\mathcal{I}_{\mathsf{top}}$ portion of the random-cluster interface, it induces free boundary conditions on the vertices immediately below it. Pushing these free boundary conditions up, as would happen under a map that lifts $\mathcal{I}_{\mathsf{top}}$, heuristically moves the free part of the boundary away from height zero, increasing the marginals at height zero, and hence interactions with the boundary only favor the higher realization of $\mathcal{I}_{\mathsf{top}}$, at which point the new entropy we can inject via downwards $\mathcal{I}_{\mathsf{top}}$ spikes beats the energy cost paid for lifting the interface up to $(1 - \varepsilon)h_n^*$.

(2) *Necessity for a crude upper bound.* The above argument, however, has problems due to the potential overhangs in $\mathcal{I}_{\mathsf{top}}$ resulting in its vertical shift not being entirely above it. Our solution is to lift $\mathcal{I}_{\mathsf{top}}$ by an exaggerated amount, more than its maximum height oscillation, and as long as this is $n^{o(1)}$, the extra cost this induces only changes the $\varepsilon_\beta$ in the $(1 - \varepsilon_\beta)h_n^*$ we can ensure the interface reaches up to.

(3) *Control on repulsive interactions with hard floor via fuzzy Potts.* Step (2) necessitates some non-trivial upper bound on the maximum height of the interface in the hard-floor setting. In Proposition 2.9, we use a delicate revealing and FKG argument with the fuzzy Potts representation of the model to show a monotonicity in the opposite direction of step (1) for the $\mathcal{I}_{\mathsf{blue}}$ interface: the interactions of $\mathcal{I}_{\mathsf{blue}}$ with the hard floor are only more attractive than in the soft-floor case—actually with boundary conditions flipping color at height $\mathsf{h} = C \log n$ instead of 0. In this latter setting, an $O(\log n)$ upper bound is easy to deliver. This same monotonicity is used in Section 3 to give the sharp $h_n^*$ upper bound on the height of $\mathcal{I}_{\mathsf{blue}}$.

These results are presented in a different order in the section than the above logic, starting with preliminaries introducing the coupled random-cluster, Potts, and fuzzy Potts interfaces in Section 2.1, then the monotonicity relation of item (3) is established in Section 2.2, and finally items (1)–(2) to give the logarithmic lower bound using maps on $\mathcal{I}_{\mathsf{top}}$ are developed in Sections 2.3–2.4.

### 2.1. Random-cluster, FK–Potts and fuzzy Potts models.
Let $G = (V, E)$ be a finite graph. The *random-cluster* model on $G$, also known as the *Fortuin–Kasteleyn* (FK) model, is a probability

distribution over subsets $\omega$ of the edge set $E$, with parameters $0 < p < 1$ and $q > 0$, given by

$$\pi_G(\omega) \propto \left(\frac{p}{1-p}\right)^{|\omega|} q^{\kappa(\omega)}, \tag{2.1}$$

where $\kappa(\omega)$ is the number of connected components of the graph $(V, \omega)$. One views $\omega$ as a Boolean function on $E$, assigning $e \in E$ the value $\omega_e = 1$ if it is present ("*open*") and $\omega_e = 0$ if it is missing ("*closed*"); we refer to the $\kappa(\omega)$ connected components of the graph of open edges as *open clusters*. Every edge is dual to the plaquette it is normal to and bisected by, and we say that plaquette is dual-to-open or dual-to-closed if said edge is open or closed, respectively.

Consider the $q$-state Potts model on a finite graph $G$ at inverse-temperature $\beta$, and the FK model on the same graph with $p = 1 - e^{-\beta}$ and the same (integer) $q$. These two models can be coupled via the FK–Potts coupling, given by the following joint measure on edge subsets $\omega \subset E$ and vertex colorings $\sigma : V \to \{1, \dots, q\}$:

$$\mathbf{P}_G(\omega, \sigma) \propto \left(\frac{p}{1-p}\right)^{|\omega|} \prod_{(u,v) \in \omega} \mathbf{1}_{\{\sigma_u = \sigma_v\}}.$$

We will refer to the joint distribution $\mathbf{P}_G$ as the FK–Potts law. In order to sample $(\omega, \sigma) \sim \mathbf{P}_G$ conditional on a given edge configuration $\omega \sim \pi_G$, one independently and uniformly at random draws a random color $\sigma_{\mathcal{C}}$ for each open cluster $\mathcal{C}$ of $\omega$, and assigns every $v \in \mathcal{C}$ the color $\sigma_{\mathcal{C}}$. For the converse, to sample $(\omega, \sigma) \sim \mathbf{P}_G$ conditional on a given coloring $\sigma \sim \mu_G$, one closes all edges, then independently samples every monochromatic edge as a Bernoulli($p$) random variable. We will often use blue and red for the colors 1 and 2, respectively, as well as nonB (resp., nonR) to denote all colors other than blue (resp., red).

The $q$-state *fuzzy Potts* model[1] is the pushforward measure

$$\phi_G = \mu_G \circ f_{\mathsf{bf}}^{-1}$$

for $f_{\mathsf{bf}} : \{1, \dots, q\}^V \to \{\mathsf{blue}, \mathsf{fuzzy}\}^V$ that replaces every color other than blue by the same fuzzy color. Equivalently, one can sample $\sigma \sim \phi_G$ by first sampling $\omega \sim \pi_G$, then coloring the vertices of each open cluster of $\omega$ independently, in blue with probability $1/q$ or in the color fuzzy otherwise. It was shown in [4] (see also [16]) that this model satisfies the FKG lattice condition. In particular, for any arbitrary fixed boundary condition $\eta$ (taking values in blue, fuzzy), the fuzzy Potts model $\phi_G^\eta$ (that is, $\phi_G$ conditional on the boundary vertices of $G$ agreeing with the assignment $\eta$) has for any pair of events $A, B$ that are increasing in the set of blue vertices, that

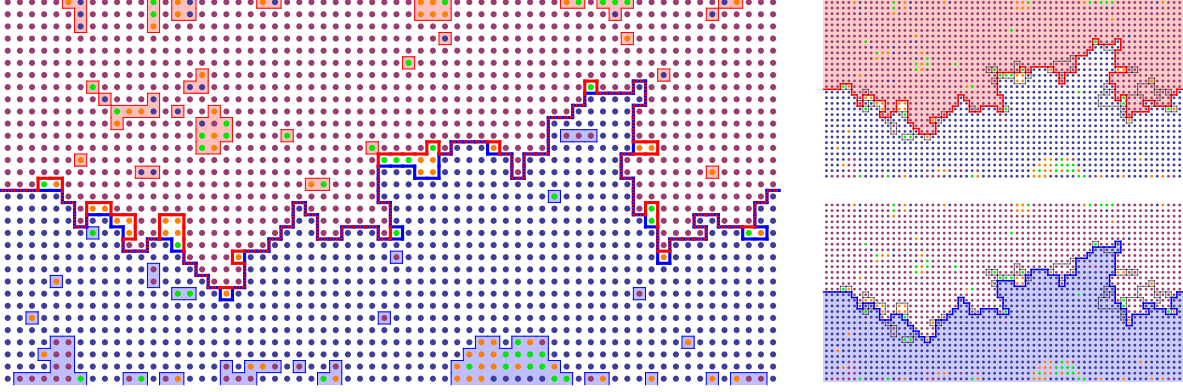$$\phi_G^\eta(A \cap B) \geq \phi_G^\eta(A)\phi_G^\eta(B). \tag{2.2}$$

In the context of Theorem 1.1, it will be useful to view the $n \times n \times m$ box $\Lambda_{n,m}$ as a subgraph of $(\mathbb{Z} + \frac{1}{2})^3$, so that the Potts model assigns colors to the midpoints of unit cubes in $\mathbb{Z}^3$, whence the interface between, say, blue and nonB vertices, is a set of plaquettes whose corners lie in $\mathbb{Z}^3$:

$$\Lambda_{n,m} = \left([-\tfrac{n}{2}, \tfrac{n}{2}]^2 \times [0, m]\right) \cap (\mathbb{Z} + \tfrac{1}{2})^3. \tag{2.3}$$

To impose a boundary condition $\eta$ on $\Lambda_{n,m}$, one adds to the graph its outer vertex boundary in $(\mathbb{Z} + \frac{1}{2})^3$ to get the outer vertex boundary set $\partial \Lambda_{n,m} = \{v \in (\mathbb{Z} + \frac{1}{2})^3 \setminus \Lambda_{n,m} : v \sim \Lambda_{n,m}\}$, and forces those vertices to take certain colors $\eta$, e.g., red on boundary vertices in the upper half-space and blue on the lower half-space to obtain $\mu_{\Lambda_{n,m}}^{\mathsf{fl}}$. We write $\mu_n^{\mathsf{fl}} = \mu_{\Lambda_{n,n}}^{\mathsf{fl}}$ for brevity, denoting the analogous fuzzy Potts model by $\phi_n^{\mathsf{fl}}$. When a boundary condition $\eta$ is understood from context, e.g., when considering $\mu_n^{\mathsf{fl}}$, we use $\partial_{\mathsf{red}} \Lambda_{n,m}$ (respectively, $\partial_{\mathsf{blue}} \Lambda_{n,m}$) to denote the subsets of $\partial \Lambda_{n,m}$ getting color red (resp., blue) under $\eta$.

The FK model corresponding to $\mu_{\Lambda_{n,m}}^{\mathsf{fl}}$, under the FK–Potts coupling, would have us sample $\pi_{\Lambda_{n,m}}^{\mathsf{w,w}}$ with boundary conditions that are *wired* on $\partial_{\mathsf{red}} \Lambda_{n,m}$ (all vertices of $\partial_{\mathsf{red}} \Lambda_{n,m}$ are viewed

---

[1]The definition here is the one from [4], and the special case of the $(1, q-1)$-state fuzzy Potts model from [16].

FIGURE 1. A Potts configuration (left) and its two Potts interfaces $\mathcal{I}_{\mathsf{blue}}$ and $\mathcal{I}_{\mathsf{red}}$ (right).

as being in the same connected component when counting $\kappa(\omega)$ in (2.1)) and wired on $\partial_{\mathsf{blue}}\Lambda_{n,m}$, conditional on these not being in the same connected component of $\omega$, i.e.,

$$\mathfrak{D}_{n,m} := \{\omega : \partial_{\mathsf{red}}\Lambda_{n,m} \not\leftrightarrow \partial_{\mathsf{blue}}\Lambda_{n,m} \text{ in } \omega\} \tag{2.4}$$

One then recovers $\sigma \sim \mu_{\Lambda_{n,m}}^{\mathsf{fl}}$ from $\omega \sim \pi_{\Lambda_{n,m}}^{\mathsf{w},\mathsf{w}}(\cdot \mid \mathfrak{D}_{n,m})$ via giving the cluster of $\partial_{\mathsf{blue}}\Lambda_{n,m}$ the color $\mathsf{blue}$, giving the cluster of $\partial_{\mathsf{red}}\Lambda_{n,m}$ the color $\mathsf{red}$, and then coloring every remaining open cluster of $\omega$ uniformly and independently (cf., e.g., [9, §2.2],[22, Fact 3.4 and Cor. 3.5] where this coupling with disconnection events was previously used). We will use the abbreviated notation $\bar{\pi}_n^{\mathsf{fl}} := \pi_{\Lambda_{n,n}}^{\mathsf{w},\mathsf{w}}(\cdot \mid \mathfrak{D}_{n,n})$, and $\mathbf{P}_n^{\mathsf{fl}}$ for the joint FK–Potts distribution whose marginals are $\mu_n^{\mathsf{fl}}$ and $\bar{\pi}_n^{\mathsf{fl}}$.

We will also need to consider the Potts model on the enlarged $n \times n \times 2m$ box

$$\Lambda'_{n,m} = \left([-\tfrac{n}{2}, \tfrac{n}{2}]^2 \times [-m, m]\right) \cap (\mathbb{Z} + \tfrac{1}{2})^3, \tag{2.5}$$

with boundary conditions that, for an integer $\mathsf{h}$, are $\mathsf{red}$ at height at least $\mathsf{h}$ and $\mathsf{blue}$ at height at most $\mathsf{h}$ (recall our sites are at half-integer heights). We will denote this distribution by $\mu_{\Lambda'_{n,m}}^{\mathsf{h}}$, and write $\mu_n^{\mathsf{h}} = \mu_{\Lambda'_{n,n}}^{\mathsf{h}}$ for brevity. Similarly, we denote by $\bar{\pi}_{\Lambda'_{n,m}}^{\mathsf{h}}$ the FK model on $\Lambda'_{n,m}$ with two wired boundary components, above and below height $\mathsf{h}$, conditioned on no path connecting them in $\omega$, and put $\bar{\pi}_n^{\mathsf{h}} = \bar{\pi}_{\Lambda'_{n,n}}^{\mathsf{h}}$; as before, the FK–Potts measure with marginals $(\mu_n^{\mathsf{h}}, \bar{\pi}_n^{\mathsf{h}})$ is denoted by $\mathbf{P}_n^{\mathsf{h}}$. In the special case $\mathsf{h} = 0$—the setup in Dobrushin's pioneering works—we may also write $\mu_n^{\mathsf{dob}} := \mu_n^0$. In this context, $\partial_{\mathsf{red}}\Lambda'_{n,m}$ and $\partial_{\mathsf{blue}}\Lambda'_{n,m}$ refer to the portions of the outer vertex boundary of $\Lambda'_{n,m}$ above and below height $\mathsf{h}$ respectively.

Under each of the boundary conditions specified above, we can define two Potts interfaces $\mathcal{I}_{\mathsf{blue}}$ and $\mathcal{I}_{\mathsf{red}}$, and two coupled random-cluster interfaces, $\mathcal{I}_{\mathsf{bot}}$ and $\mathcal{I}_{\mathsf{top}}$, defined as follows:

**Definition 2.1** (FK and Potts interfaces). Let $(\sigma, \omega) \sim \mathbf{P}_n^{\mathsf{fl}}$ or $(\sigma, \omega) \sim \mathbf{P}_n^{\mathsf{h}}$. The below definition applies to both, with the sets "blue boundary vertices", and "red boundary vertices" interpreted according to the setting.

(1) Let $\mathcal{V}_{\mathsf{blue}}$ be the connected component (via adjacency in $(\mathbb{Z}+\tfrac{1}{2})^3$) of blue vertices in $\sigma$ incident to the blue boundary vertices; augment it to $\widehat{\mathcal{V}}_{\mathsf{blue}}$ by adding all connected components of $(\mathbb{Z} + \tfrac{1}{2})^3 \setminus \mathcal{V}_{\mathsf{blue}}$ whose outer vertex boundary is a subset of $\mathcal{V}_{\mathsf{blue}}$. (This guarantees that all bubbles in the blue phase are in $\widehat{\mathcal{V}}_{\mathsf{blue}}$.) Define $\mathcal{V}_{\mathsf{red}}$, $\widehat{\mathcal{V}}_{\mathsf{red}}$ analogously. (See Fig. 1.)

(2) Let $\mathcal{V}_{\mathsf{top}}$ denote the vertices of the open cluster of $\omega$ incident to the red boundary vertices; augment it to $\widehat{\mathcal{V}}_{\mathsf{top}}$ by adding all connected components of $(\mathbb{Z}+\tfrac{1}{2})^3 \setminus \mathcal{V}_{\mathsf{top}}$ whose outer vertex boundary is a subset of $\mathcal{V}_{\mathsf{top}}$. Define $\mathcal{V}_{\mathsf{bot}}$, $\widehat{\mathcal{V}}_{\mathsf{bot}}$ analogously. (See Fig. 2.)
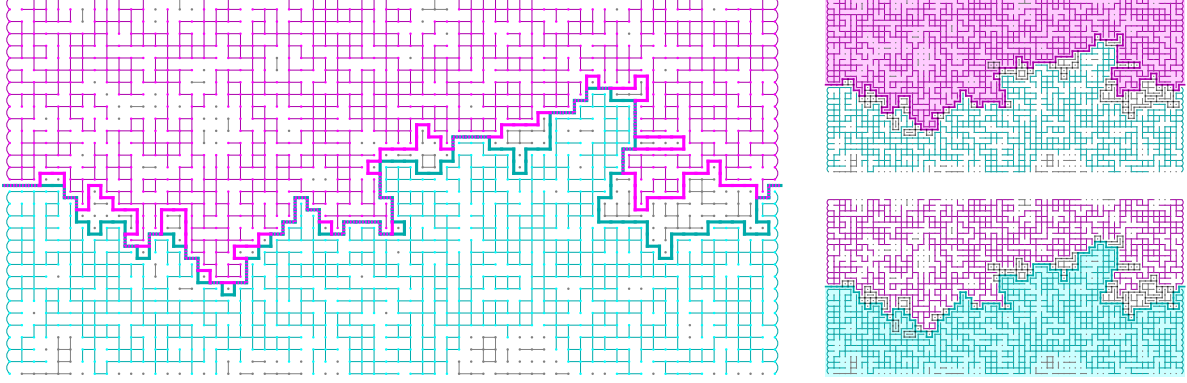
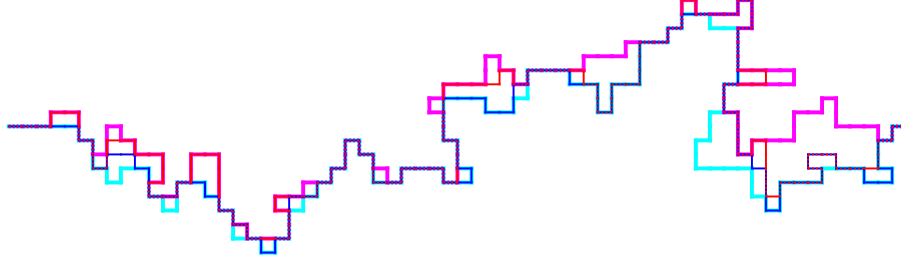FIGURE 2. A random-cluster configuration (left), and its interfaces $\mathcal{I}_{\mathsf{bot}}$ and $\mathcal{I}_{\mathsf{top}}$ (right).



FIGURE 3. The four coupled interfaces, $\mathcal{I}_{\mathsf{top}}$ (pink highlighted), $\mathcal{I}_{\mathsf{red}}$ (red line), $\mathcal{I}_{\mathsf{blue}}$ (blue line) and $\mathcal{I}_{\mathsf{bot}}$ (cyan highlighted), showing the ordering $\mathcal{I}_{\mathsf{top}} \succeq \mathcal{I}_{\mathsf{red}} \succeq \mathcal{I}_{\mathsf{blue}} \succeq \mathcal{I}_{\mathsf{bot}}$.

The interface $\mathcal{I}_{\mathsf{blue}}$ is the set of plaquettes dual to (closed) edges $(x, y)$ of $(\mathbb{Z} + \frac{1}{2})^3$ with $x \in \widehat{\mathcal{V}}_{\mathsf{blue}}$ and $y \in \widehat{\mathcal{V}}_{\mathsf{blue}}^c$. Analogously, define $\mathcal{I}_{\mathsf{red}}$, $\mathcal{I}_{\mathsf{bot}}$ and $\mathcal{I}_{\mathsf{top}}$ w.r.t. $\widehat{\mathcal{V}}_{\mathsf{red}}$, $\widehat{\mathcal{V}}_{\mathsf{bot}}$ and $\widehat{\mathcal{V}}_{\mathsf{top}}$, respectively.

By the FK–Potts coupling, under either set of boundary conditions $(\sigma, \omega) \sim \mathbf{P}_n^{\mathsf{fl}}$ or $(\sigma, \omega) \sim \mathbf{P}_n^{\mathsf{h}}$, one has $\widehat{\mathcal{V}}_{\mathsf{top}} \subset \widehat{\mathcal{V}}_{\mathsf{red}} \subset \widehat{\mathcal{V}}_{\mathsf{blue}}^c$ and $\widehat{\mathcal{V}}_{\mathsf{bot}} \subset \widehat{\mathcal{V}}_{\mathsf{blue}}(\subset \widehat{\mathcal{V}}_{\mathsf{red}}^c)$. Then the four interfaces are ordered (see Fig. 3) via the partial order on subsets of the vertices[2]:

$$\mathcal{I}_{\mathsf{top}} \succeq \mathcal{I}_{\mathsf{red}} \succeq \mathcal{I}_{\mathsf{blue}} \succeq \mathcal{I}_{\mathsf{bot}}. \tag{2.6}$$

The four interfaces $\mathcal{I}_{\mathsf{top}}, \mathcal{I}_{\mathsf{red}}, \mathcal{I}_{\mathsf{blue}}, \mathcal{I}_{\mathsf{bot}}$ are all subsets of a single *full* interface, denoted $\mathcal{I}_{\mathsf{full}}$, which is the object that can be studied via cluster expansion, as done by Gielis and Grimmett [14].

**Definition 2.2** (Full interface). Let $\omega \sim \bar{\pi}_n^{\mathsf{fl}}$ or $\omega \sim \bar{\pi}_n^{\mathsf{h}}$. The below definition applies to both, with the sets "blue boundary vertices", and "red boundary vertices" interpreted according to the setting. Consider the set $F$ of plaquettes dual to closed edges of $\omega$. The full interface $\mathcal{I}_{\mathsf{full}}$ is the maximal 1-connected component of $F$ (plaquettes $f, f'$ are 1-adjacent if they share an edge) incident to the boundary plaquettes (plaquettes dual to edges between the blue and red boundary vertices).

N.B., $\mathcal{I}_{\mathsf{full}}$ is amenable to a low-temperature cluster expansion analysis, because it induces fully wired boundary conditions on both the set of sites above it, and the set of sites below it.

**Proposition 2.3** ([14,15])**.** *Let $q \geq 1$ and $n, m \geq 1$. There exist $p_0 < 1$ so the following holds for all $p \geq p_0$. Denoting by $|I|$ the number of plaquettes in $I$, and by $\kappa(I)$ the number of open clusters*

---

[2]Associate $\mathcal{I}_{\mathsf{blue}}$ and $\mathcal{I}_{\mathsf{bot}}$ with $\widehat{\mathcal{V}}_{\mathsf{blue}}$ and $\widehat{\mathcal{V}}_{\mathsf{bot}}$, resp., and associate $\mathcal{I}_{\mathsf{red}}$ and $\mathcal{I}_{\mathsf{top}}$ with $\widehat{\mathcal{V}}_{\mathsf{red}}^c$ and $\widehat{\mathcal{V}}_{\mathsf{top}}^c$, resp. (i.e., in all cases, use the part incident to the *bottom* face of the boundary conditions) to arrive at this partial order.

*in the configuration where the only closed edges are those dual to plaquettes in $I$, we have*

$$\bar{\pi}_{\Lambda'_{n,m}}^{\mathsf{dob}}(\mathcal{I}_{\mathsf{full}} = I) = \frac{1}{Z_{n,m,p,q}} p^{|\partial I|}(1-p)^{|I|} q^{\kappa(I)} \exp\left(\sum_{f\in\mathcal{I}} \mathbf{g}(f,I)\right), \tag{2.7}$$

*where $\partial I$ is the set of plaquettes that are 1-connected to $I$ but not in $I$, and $\mathbf{g}$ is some function satisfying the following: there exist absolute constants $c, K > 0$ such that*

$$|\mathbf{g}(f,I)| \leq K, \tag{2.8}$$

$$|\mathbf{g}(f,I) - \mathbf{g}(f',I')| \leq K e^{-cr(f,I;f',I')}, \tag{2.9}$$

*for all $f, I, f', I'$, where*

$$r(f,I;f',I') = \sup\{r : I \cap B_r(f) \cap \Lambda'_{n,m} \equiv I' \cap B_r(f') \cap \Lambda'_{n,m}\},$$

*i.e., $r(f,I;f',\mathcal{I}')$ is the largest radius such that the interfaces $I, I'$ agree on the balls of this radius around $f, f'$ (the intersections with a ball of this radius are translates of one another).*

**Remark 2.4.** The above result is stated in [14, Lem. 9],[15, Lem 7.118] for the limit as $m \to \infty$, but its analog was first established for $\bar{\pi}_{\Lambda'_{n,m}}^{\mathsf{h}}(\mathcal{I}_{\mathsf{full}} = I)$ in [14, Lem. 7],[15, Lem. 7.104] for all $n, m \geq 1$. The proof further applies to any box $\Lambda = ([-\frac{n}{2}, \frac{n}{2}]^2 \times [m_1, m_2]) \cap (\mathbb{Z} + \frac{1}{2})^3$ for $m_1 \leq 0$ and $m_2 \geq 1$, with the caveat that the bound on $\mathbf{g}$ in terms of $r(f, I; f', I')$ (Eq. (2.9)) might render Eq. (2.7) useless if $m_1$ or $m_2$ were taken too close to 0 (as $r$ will be capped at the distance to the boundary).

We note in passing that, in Section 4, devoted to studying the limiting large deviation rates of local maxima in the interfaces, we will need to appeal to cluster expansion expressions for more general 1-connected sets of dual-to-closed plaquettes. See Proposition 4.2 and its proof.

Recall from the introduction that if $f$ is a plaquette in $(\mathbb{Z} + \frac{1}{2})^3$—so its midpoint $x = (x_1, x_2, x_3)$ has two half-integer coordinates and one integer coordinate—we write $\mathrm{ht}(f) := x_3$. In what follows, we address the maximal (minimal) height achieved by an interface $I$ above a given $(x_1, x_2)$,

$$\overline{\mathrm{ht}}_{(x_1,x_2)}(I) := \max\{x_3 : (x_1, x_2, x_3) \in I\} \quad \text{and} \quad \underline{\mathrm{ht}}_{(x_1,x_2)}(I) := \min\{x_3 : (x_1, x_2, x_3) \in I\}$$

(where one views $I$ as a subset of $\mathbb{R}^3$, the union of its plaquettes).

**Theorem 2.5** ([14, Thm. 3],[15, Thm. 7.144]). *Let $q \geq 1$, $n \geq 1$ and $m \geq n/2$. There exist $p_0 < 1$ such that, for every $p > p_0$ there is some $a = a_p > 0$ so the following holds,*

$$\bar{\pi}_{\Lambda'_{n,m}}^0(\overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{full}}) \geq k) \leq \exp(-a_p k) \quad \text{for every } x = (x_1, x_2) \in (\mathbb{Z} + \tfrac{1}{2})^2 \text{ and } k = k(n) \geq 1,$$

*and, by symmetry, the same holds for the event $\{\underline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{full}}) \leq -k\}$.*

**Remark 2.6.** The above result was stated for the limit $m \to \infty$, and as explained in Remark 2.4, is further applicable to $\Lambda'_{n,m}$ with some modifications (the potential pinning effects of the hard floor and ceiling are straightforwardly shown to be negligible when they are positioned at height $\pm cn$; we give a detailed proof in Appendix B for completeness). In our context of $p = 1 - e^{-\beta}$, and for $\beta$ large enough the result above holds true for $a = \beta - C$ where $C > 0$ is an absolute constant.

An important ingredient in the proofs is comparison between the hard floor setting of $\mu_n^{\mathsf{fl}}$ and the "soft floor" setting of the conditional distribution of $\mu_n^{\mathsf{h}}$ given that $\mathcal{I}_{\mathsf{blue}}$ lies above height 0, namely

$$\widehat{\mu}_n^{\mathsf{h}} := \mu_n^{\mathsf{h}}(\cdot \mid \mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0}), \tag{2.10}$$

where $\mathcal{L}_{\geq 0}$ denotes the positive half-space $\mathbb{R}^2 \times [0, \infty)$. (Throughout the paper, we use the notation $\mathcal{L}_h = \mathbb{R}^2 \times \{h\}$ for the slab at height $h$, and write $\mathcal{L}_{\geq h} = \bigcup_{z \geq h} \mathcal{L}_z$, and similarly for $\mathcal{L}_{>h}, \mathcal{L}_{\leq h}, \mathcal{L}_{<h}$.)

**Notational comment.** Let us conclude this preliminary subsection with a few notational remarks. Throughout the paper $\beta$ will be taken to be a sufficiently large constant independent of $n$, $p$ will equal $1 - e^{-\beta}$ per the FK–Potts coupling, and $n$ will always be sufficiently large. We use $\varepsilon_\beta$ in different lines to represent (possibly different) small constants that go to zero as $\beta \to \infty$, and $C$ in different lines to represent (possibly different) $\beta$-independent constants.

### 2.2. A logarithmic upper bound on the top interface.

As explained by step (2) in the sketch at the beginning of this section, the lower bound argument requires as an input, a logarithmic upper bound on the maximal height of $\mathcal{I}_{\text{top}}$ in the measure $\bar{\pi}_n^{\text{fl}}$. Recall that $\mathcal{L}_{\geq h}$ is the slab $\mathbb{R}^2 \times [h, \infty)$.

**Theorem 2.7.** *We have*
$$\bar{\pi}_n^{\text{fl}}(\mathcal{I}_{\text{top}} \cap \mathcal{L}_{\geq (14/\beta) \log n} \neq \emptyset) \leq n^{-1+\varepsilon_\beta}. \tag{2.11}$$
*Consequently, by Eq.* (2.6), *the same holds for* $\mathcal{I}_{\text{bot}}$ *under* $\bar{\pi}_n^{\text{fl}}$, *as well for* $\mathcal{I}_{\text{blue}}, \mathcal{I}_{\text{red}}$ *under* $\mu_n^{\text{fl}}$.

While the logarithmic upper bound on $\mathcal{I}_{\text{top}}$ implies one for $\mathcal{I}_{\text{blue}}$, our argument for establishing this bound for $\mathcal{I}_{\text{top}}$ goes through first proving such a result directly on $\mathcal{I}_{\text{blue}}$ (Corollary 2.12 below) and then using the rigidity of the 4 interfaces (Corollary 2.15 below) to conclude the theorem above.

2.2.1. *Upper bound on the blue interface with a soft floor.* We begin with an upper bound on the maximal height of $\mathcal{I}_{\text{blue}}$ in the soft-floor setting of $\hat{\mu}_n^{\text{h}}$, with the boundary conditions h set high enough—larger than an absolute constant multiple of $h_n^*$.

**Lemma 2.8.** *Let* $h_0 = \lfloor (3/\beta) \log n \rfloor$. *We have*
$$\hat{\mu}_n^{h_0}(\mathcal{I}_{\text{blue}} \cap \mathcal{L}_{\geq 2h_0} \neq \emptyset) \leq n^{-1+\varepsilon_\beta}. \tag{2.12}$$

*Proof.* By Theorem 2.5 and Remark 2.6, a union bound over $x = (x_1, x_2) \in [-\frac{n}{2}, \frac{n}{2}]^2 \cap (\mathbb{Z} + \frac{1}{2})^2$, as well as on the two events $\{\overline{\text{ht}}_x(\mathcal{I}_{\text{full}}) \geq h_0\}$ and $\{\underline{\text{ht}}_x(\mathcal{I}_{\text{full}}) \leq -h_0\}$ in Theorem 2.5 (which, in our context here of $\bar{\pi}_n^{h_0}$, address the intersection of $\mathcal{I}_{\text{full}}$ with $\mathcal{L}_{2h_0}$ and $\mathcal{L}_0$, respectively), gives
$$\bar{\pi}_n^{h_0}(\mathcal{I}_{\text{full}} \cap (\mathcal{L}_{\leq 0} \cup \mathcal{L}_{\geq 2h_0}) \neq \emptyset) \leq 2n^2 e^{-(\beta - C)h_0} \leq n^{-1+\varepsilon_\beta},$$
where we absorbed the constant pre-factor into the $\varepsilon_\beta$ term. In particular, as $\mathcal{I}_{\text{blue}} \subset \mathcal{I}_{\text{full}}$, we have
$$\mu_n^{h_0}(\mathcal{I}_{\text{blue}} \cap (\mathcal{L}_{\leq 0} \cup \mathcal{L}_{\geq 2h_0}) \neq \emptyset) \leq n^{-1+\varepsilon_\beta},$$
and thus
$$\hat{\mu}_n^{h_0}(\mathcal{I}_{\text{blue}} \cap \mathcal{L}_{\geq 2h_0} \neq \emptyset) \leq \frac{\mu_n^{h_0}(\mathcal{I}_{\text{blue}} \cap \mathcal{L}_{\geq 2h_0} \neq \emptyset)}{\mu_n^{h_0}(\mathcal{I}_{\text{blue}} \subset \mathcal{L}_{\geq 0})} \leq \frac{n^{-1+\varepsilon_\beta}}{1 - n^{-1+\varepsilon_\beta}} = (1 + o(1))n^{-1+\varepsilon_\beta}$$
(using the preceding display in both the numerator and denominator above), as required. ∎

2.2.2. *Comparing the soft and hard floor measures.* We next wish to move from $\hat{\mu}_n^{\text{h}}$ to $\mu_n^{\text{fl}}$. The next proposition, which may be of independent interest, establishes a comparison between these measures via a subtle revealing and monotonicity argument that utilizes the FKG property of the fuzzy Potts model. It is also key to the sharper $h_n^*$ upper bound proved in Section 3.

**Proposition 2.9.** *Let* h $\geq 0$, *and in the context of both* $\hat{\mu}_n^{\text{h}}$ *and* $\mu_n^{\text{fl}}$, *let* $A$ *be an event on configurations* $\sigma$ *on* $\Lambda_{n,n}'$ *that is increasing in the set of* blue *vertices (that is, the set of configurations in* $A$ *is closed under the operation of coloring any vertex in* blue*) and measurable w.r.t. the set* $\hat{\mathcal{V}}_{\text{blue}}^{\text{c}}$. *Then*
$$\mu_n^{\text{fl}}(A) \leq \hat{\mu}_n^{\text{h}}(A),$$
*where the set* $\hat{\mathcal{V}}_{\text{blue}}^{\text{c}}$ *is understood w.r.t. the corresponding boundary conditions on either side.*
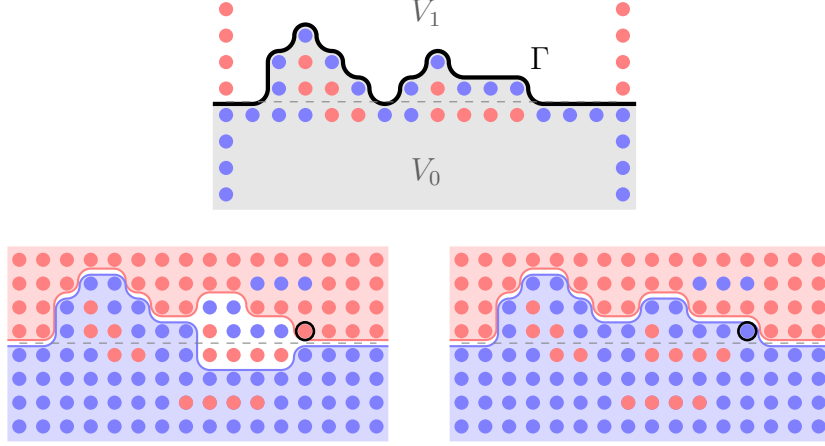
FIGURE 4. Top: the vertex set $\Gamma$ from the proof of Proposition 2.9 for $\mathsf{h} = 0$, revealed with the rest of the vertices of $\mathcal{L}_{<0}$ to form $V_0$ (marked in gray). Bottom: in this example, flipping the color of a single site $v \notin V_0$ (circled in black) is pivotal for the event $\{\mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0}\}$.

*Proof.* Consider $\sigma \sim \mu_n^{\mathsf{h}}$, and let us examine the event $\{\mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0}\}$ delineating $\widehat{\mu}_n^{\mathsf{h}}$. Recalling Definition 2.1, Item 1, we have $v \in \widehat{\mathcal{V}}_{\mathsf{blue}}^c$ if and only if $v \notin \mathcal{V}_{\mathsf{blue}}$ and there exists a path in $(\mathbb{Z} + \frac{1}{2})^3$, connecting $v$ to a red boundary vertex via internal vertices in $\widehat{\mathcal{V}}_{\mathsf{blue}}^c (\subset \mathcal{V}_{\mathsf{blue}}^c)$. Further note that

$$\{\mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0}\} = \bigcap_{v \in \mathcal{L}_{<0}} \{v \in \widehat{\mathcal{V}}_{\mathsf{blue}}\}. \tag{2.13}$$

Indeed, the interface $\mathcal{I}_{\mathsf{blue}}$ is nothing but the set of plaquettes separating nearest-neighbors $u, v \in (\mathbb{Z} + \frac{1}{2})^3$ with $u \in \widehat{\mathcal{V}}_{\mathsf{blue}}$ and $v \in \widehat{\mathcal{V}}_{\mathsf{blue}}^c$ (in fact with $u \in \mathcal{V}_{\mathsf{blue}}$, as $\widehat{\mathcal{V}}_{\mathsf{blue}} \setminus \mathcal{V}_{\mathsf{blue}}$ can never be adjacent to $\widehat{\mathcal{V}}_{\mathsf{blue}}^c$). Thus,

$$\underline{\mathrm{ht}}_{(x_1, x_2)}(\mathcal{I}_{\mathsf{blue}}) = \min \left\{ x_3 : (x_1, x_2, x_3 - \tfrac{1}{2}) \in \widehat{\mathcal{V}}_{\mathsf{blue}}, (x_1, x_2, x_3 + \tfrac{1}{2}) \in \widehat{\mathcal{V}}_{\mathsf{blue}}^c \right\},$$

giving the identity in Eq. (2.13).

Determining whether $\mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0}$ must be done with care; a good example to keep in mind in this context is $\sigma$ that colors every $v \in \mathcal{L}_{<0}$ in blue and every $v \in \mathcal{L}_{>0}$ in red except for $Q = \{\pm \frac{1}{2}\}^3$, where we swap the assignment, coloring $v \in Q \cap \mathcal{L}_{<0}$ in red and $v \in S \cap \mathcal{L}_{>0}$ in blue. In this example, $[0,1]^2 \times \{-1\} \subset \mathcal{I}_{\mathsf{blue}}$ (in particular $\mathcal{I}_{\mathsf{blue}} \not\subset \mathcal{L}_{\geq 0}$) and similarly $[0,1]^2 \times \{1\} \subset \mathcal{I}_{\mathsf{red}}$, since $Q \subset \widehat{\mathcal{V}}_{\mathsf{blue}}^c \cap \widehat{\mathcal{V}}_{\mathsf{red}}^c$. On the other hand, if we were to also set the color of $(\frac{3}{2}, \frac{1}{2}, \frac{1}{2})$ to blue, it would connect $Q \cap \mathcal{L}_{>0}$ to $\mathcal{V}_{\mathsf{blue}}$, thus leading to $Q \subset \widehat{\mathcal{V}}_{\mathsf{blue}}$ and to $\mathcal{I}_{\mathsf{blue}} = \mathcal{I}_{\mathsf{red}} \subset \mathcal{L}_{\geq 0}$.

With this in mind, consider $\sigma \sim \widehat{\mu}_n^{\mathsf{h}} = \mu_n^{\mathsf{h}}(\cdot \mid \mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0})$; let us reveal the color of every $v \in \mathcal{L}_{<0}$, as well as the nonB component $\mathcal{C}_v$ of every $v \in \mathcal{L}_{-1/2}$. Denote by $V_0$ the vertices whose color was revealed, and let $\Gamma$ denote its inner vertex boundary (vertices $x \in V_0$ adjacent to some $y \notin V_0$). By definition, every $x \in \Gamma$ is blue; indeed, either $x \in \mathcal{L}_{-1/2}$, or it is in the external boundary of $\mathcal{C}_v$, and must be blue since no $\mathcal{C}_v$ reaches a red boundary vertex (as this would imply that $v \in \widehat{\mathcal{V}}_{\mathsf{blue}}^c$). Let $V_1$ denote the vertices whose color was not yet revealed. (N.B. every $(\mathbb{Z} + \frac{1}{2})^3$-path from $x \in V_1$ to $y \in \mathcal{L}_{<0}$ must pass through $\Gamma$.) See Fig. 4 for an illustration of the sets $V_0$ and $\Gamma$, as well as the fact that conditional event $\{\mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0}\}$ gives information on the colors of $V_1$, as every path from a nonB site $v \in \mathcal{L}_{-1/2}$ to the boundary must go through some $u \in \mathcal{V}_{\mathsf{blue}}$, or else we will have $v \in \widehat{\mathcal{V}}_{\mathsf{blue}}^c$. (Were we to ignore this conditional event and only look at the colors of $V_0$, then even though the vertices in $\Gamma$ are all blue, some of these may or may not be in $\mathcal{V}_{\mathsf{blue}}$, depending on the colors in $V_1$. This plays a role in the conditioning, since some of nonB vertices $v \in \mathcal{L}_{-1/2}$ must be encapsulated in $\widehat{\mathcal{V}}_{\mathsf{blue}}$ via a boundary of $\mathcal{V}_{\mathsf{blue}}$ vertices using sites also from $V_1$.)

For a realization $\sigma\!\restriction_{V_0} = \eta$ (that is compatible with the event $\mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0}$), let $\mathcal{B}_\eta$ denote the $\sigma\!\restriction_{V_1}$-measurable event $\{\mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0}\}$ under $\mu_n^{\mathsf{h}}$ conditional on $V_0 , \sigma\!\restriction_{V_0} = \eta$. Formally, $B_\eta$ is the event that every vertex of $\Gamma$ not shown to be in $\widehat{\mathcal{V}}_{\mathsf{blue}}$ by $\eta$, is indeed part of $\widehat{\mathcal{V}}_{\mathsf{blue}}$ in $\eta$ concatenated with $\sigma\!\restriction_{V_1}$. By this reasoning,

$$\widehat{\mu}_n^{\mathsf{h}}(\cdot \mid V_0, \, \sigma\!\restriction_{V_0} = \eta) = \mu_n^{\mathsf{h}}(\cdot \mid V_0, \, \sigma\!\restriction_{V_0} = \eta, \, \mathcal{B}_\eta) \,.$$

Note that $\mathcal{B}_\eta$ is increasing in $\mathsf{blue}$ vertices (stemming from the monotonicity of the set $\widehat{\mathcal{V}}_{\mathsf{blue}}$ in the set of $\mathsf{blue}$ vertices).

Having exposed $V_0$ and $\sigma\!\restriction_{V_0}$, we wish to show, uniformly over the realization $\eta$ of the latter, that

$$\mu_n^{\mathsf{h}}(A \mid V_0 , \sigma\!\restriction_{V_0} = \eta , \mathcal{B}_\eta) \geq \mu_n^{\mathsf{fl}}(A) \tag{2.14}$$

(at which point, taking an expectation over $V_0$ and $\eta$ will conclude). To this end, we will establish the following, where here and in what follows, we say that a coloring $\eta$ of a vertex set $S \supseteq \mathcal{L}_{<0}$ realizes $V_0 = S$ if every $\sigma$ that agrees with $\eta$ on $S$ satisfies $V_0(\sigma) = \eta$ (it is important to recall that one can deterministically verify that $V_0(\sigma) = S$ from $\sigma\!\restriction_S$).

**Claim 2.10.** *For every fixed subset of vertices $S \supseteq \mathcal{L}_{<0}$ and configuration $\eta : S \to \{1, \ldots, q\}$ realizing $V_0 = S$, and every $\eta'$ obtained from $\eta$ by coloring an arbitrary subset of $S$ in $\mathsf{blue}$,*

$$\mu_n^{\mathsf{h}}(\widehat{\mathcal{V}}_{\mathsf{blue}}^c \in \cdot \mid \sigma\!\restriction_S = \eta, \mathcal{B}_\eta) = \mu_n^{\mathsf{h}}(\widehat{\mathcal{V}}_{\mathsf{blue}}^c \in \cdot \mid \sigma\!\restriction_S = \eta', \mathcal{B}_\eta) \,. \tag{2.15}$$

*Proof.* We will establish that

$$\forall \sigma : \sigma\!\restriction_S = \eta \text{ and } \sigma\!\restriction_{S^c} \in B_\eta, \text{ changing } \eta \text{ to } \eta' \text{ does not change } \widehat{\mathcal{V}}_{\mathsf{blue}} \tag{2.16}$$

To show this, we first argue that $S \subset \widehat{\mathcal{V}}_{\mathsf{blue}}$ on the events $\sigma\!\restriction_S = \eta$ and $\sigma\!\restriction_{S^c} \in \mathcal{B}_\eta$. Indeed, suppose that $x \in S$ is not in $\widehat{\mathcal{V}}_{\mathsf{blue}}$. If $x \in \mathcal{L}_{<0}$, then since $\mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0}$, then Eq. (2.13) gives a contradiction. If $x \in \mathcal{L}_{>0}$ and $x \notin \widehat{\mathcal{V}}_{\mathsf{blue}}$, there must be a $\widehat{\mathcal{V}}_{\mathsf{blue}}^c$ path connecting $x$ to the $\mathsf{red}$ boundary $\partial \Lambda'_{n,n} \cap \mathcal{L}_{>0}$; at the same time, $x$ must be incident to, or part of, a $\mathsf{nonB}$-component $\mathcal{C}_v$ for some $v \in \mathcal{L}_{<0}$. Appending said path to the path connecting $x$ to $v$ in $\mathcal{C}_v$ gives a $\widehat{\mathcal{V}}_{\mathsf{blue}}^c$ path from the $\mathsf{red}$ boundary to $\mathcal{L}_{<0}$ (the vertices of $\mathcal{C}_v$ are not in $\mathcal{V}_{\mathsf{blue}}$, and every path from $\widehat{\mathcal{V}}_{\mathsf{blue}}$ to $\widehat{\mathcal{V}}_{\mathsf{blue}}^c$ goes through $\mathcal{V}_{\mathsf{blue}}$), contradicting that $\mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0}$. Now, $S \subset \widehat{\mathcal{V}}_{\mathsf{blue}}$ means that changing colors in $S$ to $\mathsf{blue}$ does not change $\widehat{\mathcal{V}}_{\mathsf{blue}}$, and evidently does not affect the configuration on $S^c$ which $\widehat{\mathcal{V}}_{\mathsf{blue}}^c$ is a subset of.

Having established (2.16), this implies that on $\{\sigma\!\restriction_S = \eta, \sigma\!\restriction_{S^c} \in \mathcal{B}_\eta\}$, a realization of $\widehat{\mathcal{V}}_{\mathsf{blue}}^c$ corresponds to the same set of configurations on $\sigma\!\restriction_{S^c}$ as it does on $\{\sigma_S = \eta', \sigma_{S^c} \in \mathcal{B}_\eta\}$. The event $B_\eta$ is by definition measurable with respect to $\sigma\!\restriction_{S^c}$. Therefore, we can apply the domain Markov property to see that the law in the left-hand of Eq. (2.15) only depends on $\sigma\!\restriction_S = \eta$ through the inner vertex boundary of $S$, which is $\mathsf{blue}$ in $\eta$ by definition (recall the set $\Gamma$ defined as the inner vertex boundary of $V_0$), and hence also in $\eta'$. This concludes the proof of Eq. (2.15). ∎

To now show (2.14), take $\eta'$ in Claim 2.10 to be all-$\mathsf{blue}$, and use that $A$ is measurable with respect to $\widehat{\mathcal{V}}_{\mathsf{blue}}^c$, to see that

$$\mu_n^{\mathsf{h}}(A \mid \sigma\!\restriction_S = \eta, \mathcal{B}_\eta) = \mu_n^{\mathsf{h}}(A \mid \sigma\!\restriction_S \equiv \mathsf{blue}, \mathcal{B}_\eta) \,.$$

The right-hand, by Domain Markov, is nothing but a Potts model on $\Lambda_{n,n}$ (as opposed to $\Lambda'_{n,n}$), conditional on $\mathcal{B}_\eta$ as well having the sites of $S$ be all-$\mathsf{blue}$, and having boundary conditions that are $\mathsf{blue}$ in $\mathcal{L}_{<\mathsf{h}}$ and $\mathsf{red}$ above it (the sites of $S$ separate the bottom face from a path to any vertex of $S^c$). Notice that the configuration on $\mathcal{L}_{<0}$ being $\mathsf{blue}$ ensures that interpreting $\widehat{\mathcal{V}}_{\mathsf{blue}}$ w.r.t. these boundary conditions is the same as w.r.t. $\mu_n^{\mathsf{h}}$ on $\Lambda'_{n,n}$. At this point, we wish to use that these are all modifications of $\mu_n^{\mathsf{fl}}$ that are increasing in the set of $\mathsf{blue}$ vertices, so the FKG inequality for the fuzzy Potts model should allow us to drop the conditionings and only decrease the probability.

Still, unlike the Potts model, there is no Domain Markov Property in fuzzy Potts, and so this step must be done with care. First, let

$$\mu_n^\Gamma = \mu_{\Lambda_{n,n}}^{\mathsf h} \left( \sigma{\restriction}_{S^c} \in \cdot \ \big| \ \sigma{\restriction}_S = \eta, \mathcal B_\eta \right)$$

denote the Potts distribution on $S^c$ given $\sigma{\restriction}_S = \eta$ as well as $\mathcal B_\eta$, where the red external boundary vertices (those in $\mathcal L_{>\mathsf h}$) are replaced by a single (high degree) red boundary vertex $v_{\mathsf{red}}$, and the external boundary vertices in $S \cup \mathcal L_{<\mathsf h}$ are all blue. (Replacing the red boundary vertices by a single $v_{\mathsf{red}}$ does not change the Potts model, yet it will facilitate easy transitions to the fuzzy Potts model and back.) As an additional adjustment, let us connect each of the blue external boundary vertices in $\partial \Lambda_{n,n} \cap (\mathcal L_{>0} \cap \mathcal L_{<\mathsf h})$ to $v_{\mathsf{red}}$ with, a bond of strength $J > 0$ (to be taken arbitrarily large). This, again, does not affect $\mu_n^\Gamma$ by the Domain Markov property.

At this point, we move to the fuzzy Potts model $\phi_n^\Gamma$, noting that $\widehat{\mathcal V}_{\mathsf{blue}}^c$ (thus also $A$, by assumption) is measurable under the fuzzy Potts projection $f_{\mathsf{bf}}$. By FKG, and the fact that $A$ is increasing in blue vertices, we can

(a) remove the blue boundary conditions on the sites of $S \cap \mathcal L_{>0}$ and $\partial \Lambda_{n,n} \cap (\mathcal L_{>0} \cap \mathcal L_{<\mathsf h})$, and
(b) remove the conditioning on the event $\mathcal B_\eta$,

arriving at $\phi_n^{\mathsf{fl}}$, a fuzzy Potts measure on $\Lambda_n$ with a hard blue floor at height 0, a single vertex $v_{\mathsf{red}}$ replacing all the red boundary vertices in $\mathcal L_{>\mathsf h}$, and that vertex connected to every formerly blue boundary condition vertex in $\partial \Lambda_{n,n} \cap \mathcal L_{>0} \cap \mathcal L_{<\mathsf h}$ via edges with interaction strength $J$):

$$\mu_n^\Gamma(A) = \phi_n^\Gamma(A) \geq \phi_n^{\mathsf{fl}}(A) = \mu_n^{\mathsf{fl}}(A) + O(n\mathsf h e^{-\beta J}) .$$

Here, the last identity uses that by a union bound, the fuzzy boundary vertices will have the same color except with probability $O(n\mathsf h e^{-\beta J})$; a single fuzzy boundary vertex in fuzzy Potts corresponds to a single nonB boundary vertex in Potts, which by symmetry can be taken to be red. Taking $J$ arbitrarily large completes the proof. ∎

**Remark 2.11.** In Proposition 2.9 we required that $A$, the event that is increasing in the set of blue vertices, should depend only on $\widehat{\mathcal V}_{\mathsf{blue}}^c$. The assertion of the proposition would fail if we omitted that requirement; e.g., if o denotes the origin vertex $(\frac12, \frac12, \frac12)$ and $A = \{\sigma_{\mathsf o} = \mathsf{blue}\}$ then one would have $\widehat\mu_n^0(A) < \mu_n^{\mathsf{fl}}(A)$ (whereas the event $A = \{\sigma_{\mathsf o} \in \widehat{\mathcal V}_{\mathsf{blue}}\}$ would be covered by the proposition and have $\widehat\mu_n^0(A) \geq \mu_n^{\mathsf{fl}}(A)$).

**Corollary 2.12.** *For large enough $n$,*

$$\mu_n^{\mathsf{fl}}(\mathcal I_{\mathsf{blue}} \cap \mathcal L_{\geq (6/\beta)\log n} \neq \emptyset) \leq n^{-1+\varepsilon_\beta} . \tag{2.17}$$

*Proof.* Let $h_0 = \lfloor (3/\beta)\log n \rfloor$, let $A$ be the event $\{\mathcal I_{\mathsf{blue}} \cap \mathcal L_{\geq 2h_0} \neq \emptyset\}$, and recall the upper bound from Lemma 2.8 on $\widehat\mu_n^{h_0}(A)$. Since $A$ is increasing in the set of blue vertices, and is measurable w.r.t. $\widehat{\mathcal V}_{\mathsf{blue}}^c$, we may apply Proposition 2.9 for $\mathsf h = h_0$ to extend this bound to $\mu_n^{\mathsf{fl}}(A)$, as required. ∎

2.2.3. *Rigidity of the four interfaces.* In this subsection we employ a standard Peierls argument to infer the rigidity of the interfaces $\mathcal I_{\mathsf{bot}} \preceq \mathcal I_{\mathsf{blue}} \preceq \mathcal I_{\mathsf{red}} \preceq \mathcal I_{\mathsf{top}}$ from Definition 2.1, which will then allow us to extend the logarithmic upper bound from $\mathcal I_{\mathsf{blue}}$ to the the other three interfaces.

Recall that $\widehat{\mathcal V}_{\mathsf{bot}} \subset \widehat{\mathcal V}_{\mathsf{top}}^c$. The following lemma bounds the tail probability for the boundary size of a $(\mathbb Z + \frac12)^3$-connected component in $\widehat{\mathcal V}_{\mathsf{top}}^c \setminus \widehat{\mathcal V}_{\mathsf{bot}}$.

**Lemma 2.13.** *Let $\omega \sim \bar\pi_n^{\mathsf{fl}}$, consider $v \in \Lambda_{n,n}$. Let $\mathcal C_v$ be the (potentially empty) $(\mathbb Z + \frac12)^3$-connected component of the vertex $v$ in $\Lambda_{n,n} \setminus (\widehat{\mathcal V}_{\mathsf{top}} \cup \widehat{\mathcal V}_{\mathsf{bot}})$. Let $\Gamma_v$ be its plaquette boundary, that is, $f \in \Gamma_v$ if it is dual to the $(\mathbb Z + \frac12)^3$-bond $(u, w)$ with $u \in \mathcal C_v$ and $w \notin \mathcal C_v$. Then*

$$\bar\pi_n^{\mathsf{fl}}(|\Gamma_v| \geq r) \leq e^{-(\beta - C)r/2} .$$

*Proof.* It is convenient for this proof to have the vertex set of $\Lambda_{n,n}$ include its external boundary vertices, letting those in $\mathcal{L}_{>0}$ be part of $\widehat{\mathcal{V}}_{\mathsf{top}}$ in $\omega \sim \bar{\pi}_n^{\mathsf{fl}}$, and those of $\mathcal{L}_{<0}$ be part of $\widehat{\mathcal{V}}_{\mathsf{bot}}$. As such, every plaquette $f \in \Gamma_v$ must separate $x \in \mathcal{C}_v$ from some $y \in \widehat{\mathcal{V}}_{\mathsf{bot}} \oplus \widehat{\mathcal{V}}_{\mathsf{top}}$, whence $f \in \mathcal{I}_{\mathsf{bot}} \oplus \mathcal{I}_{\mathsf{top}}$.

Let
$$A_k := \{ |\Gamma_v| = k \,,\, |\Gamma_v \cap \mathcal{I}_{\mathsf{top}}| \geq k/2 \} \,.$$

The map $\Phi$ that opens every edge dual to $\Gamma_v \cap \mathcal{I}_{\mathsf{top}}$ maintains that $\widehat{\mathcal{V}}_{\mathsf{top}}, \widehat{\mathcal{V}}_{\mathsf{bot}}$ are disconnected in $\omega$ (having only opened edges between $\widehat{\mathcal{V}}_{\mathsf{top}}$ and $\mathcal{C}_v \subset \widehat{\mathcal{V}}_{\mathsf{bot}}^c$), and
$$\frac{\bar{\pi}_n^{\mathsf{fl}}(\Phi(\omega))}{\bar{\pi}_n^{\mathsf{fl}}(\omega)} \geq \left( \frac{p}{(1-p)q} \right)^{|\Gamma_v \cap \mathcal{I}_{\mathsf{top}}|} \geq e^{(\beta - \log q)k/2}$$

uniformly over $\omega \in A_k$ (each edge modification gains a factor of $\frac{p}{1-p}$, perhaps decreasing $\kappa(\omega)$ by 1). Enumerating over the connected set of plaquettes forming $\Gamma_v$ now shows that $\bar{\pi}_n^{\mathsf{fl}}(A_k) \leq \exp(-(\beta - C)k/2)$ for some absolute constant $C > 0$. The analogous argument treating
$$B_k := \{ |\Gamma_v| = k \,,\, |\Gamma_v \cap \mathcal{I}_{\mathsf{bot}}| \geq k/2 \}$$

via opening the edges $\Gamma_v \cap \mathcal{I}_{\mathsf{bot}}$ shows that $\bar{\pi}_n^{\mathsf{fl}}(B_k) \leq \exp(-(\beta - C)k/2)$ (with the same $C > 0$), which, since $\{ |\Gamma_v| \geq r \} \subset \bigcup_{k \geq r}(A_k \cup B_k)$, concludes the proof. $\blacksquare$

Similarly, we can easily bound the differences between the top and full interfaces.

**Lemma 2.14.** *Fix an edge $e \in \Lambda_n$ and consider the dual face $f = f_e$. Let $C_f$ be the (potentially empty) 1-connected component of $f$ in $\mathcal{I}_{\mathsf{full}} \setminus \mathcal{I}_{\mathsf{top}}$. Then,*
$$\bar{\pi}_n^{\mathsf{fl}}(|C_f| \geq r) \leq e^{-(\beta - C)r} \,.$$
*The same statement holds if we define $C_f$ to be the 1-connected component of $f$ in $\mathcal{I}_{\mathsf{full}} \setminus \mathcal{I}_{\mathsf{bot}}$.*

*Proof.* Take the map $\Phi$ that opens every edge dual to $C_f$. Since $C_f \subset \mathcal{I}_{\mathsf{full}} \setminus \mathcal{I}_{\mathsf{top}}$, this operation keeps the resulting configuration in $\mathfrak{D}_{n,n}$. The same short computation as in the proof of Lemma 2.13 immediately implies the desired bound. $\blacksquare$

**Corollary 2.15.** *Fix $\beta$ large enough. There exists a constant $C > 0$ such that for any $k \geq 0$,*
$$\bar{\pi}_n^{\mathsf{fl}} \left( \max_x \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{full}}) > \max_x \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{bot}}) + k \right) \leq C n^3 e^{-(\beta - C)k/3} \,,$$
*and*
$$\bar{\pi}_n^{\mathsf{fl}} \left( \min_x \underline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{full}}) < \min_x \underline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{top}}) - k \right) \leq C n^3 e^{-(\beta - C)k/3} \,.$$
*By the ordering of the interfaces* (2.6)*, the distance between the maximum of any of the interfaces* full, top, red, blue, bot *satisfies the same bound, and likewise for the minimum.*

*Proof.* Every $f \in \mathcal{I}_{\mathsf{top}} \setminus \mathcal{I}_{\mathsf{bot}}$ belongs to the boundary $\Gamma_v$ of some component $\mathcal{C}_v$ as per the definition in Lemma 2.13. Note that said boundary must also contain some plaquette $f' \in \mathcal{I}_{\mathsf{bot}} \setminus \mathcal{I}_{\mathsf{top}}$ (some external boundary vertex of $\mathcal{C}_v$ must be in $\widehat{\mathcal{V}}_{\mathsf{bot}}$, or else $\mathcal{C}_v$ would be part of $\widehat{\mathcal{V}}_{\mathsf{top}}$), whence $|\Gamma_v|$ is an upper bound on the length of a path of plaquettes connecting $f, f'$. Hence, the difference between $\max_x \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{top}})$ and $\max_x \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{bot}})$ is bounded by $\max_v |\Gamma_v|$. Similarly, the difference between $\max_x \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{full}})$ and $\max_x \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{top}})$ is bounded by $\max_f |C_f|$. We conclude by Lemmas 2.13–2.14 and a union bound over the $O(n^3)$ vertices in $\Lambda_n$. $\blacksquare$

**Proof of Theorem 2.7.** Let $(\sigma, \omega) \sim \mathbf{P}_n^{\mathsf{fl}}$. By taking $k = (12/\beta) \log n$ in Corollary 2.15, combined with the ordering of the interfaces in Eq. (2.6), with probability at least $1 - n^{-1+\varepsilon_\beta}$ we have that for every $x$,
$$\overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{top}}) - \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{blue}}) \leq \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{top}}) - \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{bot}}) \leq (12/\beta) \log n \,.$$
The bound on $\max_x \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{blue}})$ from Corollary 2.12 thus concludes the proof. $\blacksquare$

2.3. **Monotonicity of the effect of lifting the top interface with a hard vs. soft floor.**
The basic mechanism for establishing the typical height of the interface above a floor (hard or soft)
is to compare the effect of lifting the entire interface, say by a single $n \times n \times 1$ slab: one wishes to
weigh the gain from entropic repulsion (the ability to drive deeper spikes in the bulk of the surface)
against the loss in energy (due to the additional plaquettes along the perimeter of the newly added
slab). However, for the Potts interfaces (e.g., $\mathcal{I}_{\mathsf{blue}}$), the interactions with the hard floor might
outweigh the entropic repulsion effect (through the change in partition functions induced below
the interface). In fact, curiously, their exponential rates appear to coincide, at least up to the first
order term as a function of $\beta$. Fortunately, the interface $\mathcal{I}_{\mathsf{top}}$ (but not $\mathcal{I}_{\mathsf{full}}$ nor $\mathcal{I}_{\mathsf{bot}}$) exhibits a
nicer interaction with the hard floor; the following lemma shows, roughly, that interactions with
the hard floor make it "only easier to lift the interface" then.

**Lemma 2.16.** *For any* $\mathsf{top}$ *interface* $I \subset \mathcal{L}_{\geq 0}$ *and* $j \geq 1$, *let* $\Theta_j I$ *denote the interface obtained by
shifting* $I$ *up by* $j$ *and adding the* $4jn$ *additional boundary plaquettes of heights in* $[0, j]$, *and set*

$$\Xi_j^{\mathsf{fl}}(I) := \frac{\bar{\pi}_n^{\mathsf{fl}}(\mathcal{I}_{\mathsf{top}} = \Theta_j I)}{\bar{\pi}_n^{\mathsf{fl}}(\mathcal{I}_{\mathsf{top}} = I)} \qquad and \qquad \Xi_j^{\mathsf{dob}}(I) := \frac{\bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{top}} = \Theta_j I)}{\bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{top}} = I)} \,. \tag{2.18}$$

*If* $I$ *and* $j$ *are such that* $\max_x \overline{\mathrm{ht}}_x(I) < j \leq n/2$, *then* $\Xi_j^{\mathsf{fl}}(I) \geq \Xi_j^{\mathsf{dob}}(I)$.

To compare the effects of the map $\Theta_j$ in $\bar{\pi}_n^{\mathsf{fl}}$ vs. $\bar{\pi}_n^{\mathsf{dob}}$, we will need the following consequence of
well-known expressions for derivatives of free energies in edge parameters.

**Claim 2.17.** *For any finite graph* $(V, E)$ *and an edge set* $\tilde{E}$, *fix* $p, q$ *and let* $\tilde{Z}_G^{(\theta)}$ *be the partition
function of the* $(p, q)$-*FK model on* $G = (V, E \cup \tilde{E})$ *modified to have edge probability* $\theta$ *for* $e \in \tilde{E}$.
*Denote by* $\tilde{\pi}_G^{(\theta)}$ *the probability of an edge configuration* $\omega$ *in this model. Then*

$$\log \tilde{Z}_G^{(\theta_1)} - \log \tilde{Z}_G^{(\theta_0)} = \sum_{e \in \tilde{E}} \int_{\theta_0}^{\theta_1} \frac{\tilde{\pi}_G^{(\theta)}(\omega_e = 1)}{\theta(1 - \theta)} \, \mathrm{d}\theta \,.$$

*Proof.* As per Eq. (2.1), we have

$$\tilde{Z}_G^{(\theta)} = \sum_\omega \Big( \frac{p}{1-p} \Big)^{|\omega \cap E|} \Big( \frac{\theta}{1-\theta} \Big)^{|\omega \cap \tilde{E}|} q^{\kappa(\omega)} =: \sum_\omega W_\omega \,.$$

Noting that $\frac{\mathrm{d}}{\mathrm{d}\theta} W_\omega = \big( |\omega \cap \tilde{E}| / [\theta(1-\theta)] \big) W_\omega$, if we consider expectation w.r.t. $\tilde{\pi}_G^{(\theta)}$ , then

$$\frac{\mathrm{d}}{\mathrm{d}\theta} \log \tilde{Z}_G^{(\theta)} = \frac{1}{\tilde{Z}_G^{(\theta)}} \sum_\omega \frac{|\omega \cap \tilde{E}|}{\theta(1-\theta)} W_\omega = \sum_{e \in \tilde{E}} \frac{\tilde{\pi}_G^{(\theta)}(\omega_e = 1)}{\theta(1-\theta)} \,.$$

Integrating over $\theta$ yields the claim. ∎

**Lemma 2.18.** *Let* $I \subset \mathcal{L}_{\geq 0}$ *be a realization for* $\mathcal{I}_{\mathsf{top}}$ *and* $j \geq 1$ *such that* $\max_x \overline{\mathrm{ht}}_x(I) \leq n - j$.
*Let* $G_{I_\downarrow} = (V, E \cup \tilde{E})$ *denote the induced subgraph of* $\Lambda'_{n,n}$ *on* $V = \widehat{\mathcal{V}}_{\mathsf{top}}^{\mathrm{c}}(I)$, *augmented via edges* $\tilde{E}$
*connecting every* $v \in \mathcal{L}_{-1/2} \cap \Lambda'_{n,n}$ *to an (arbitrarily chosen) external boundary vertex* $v_* \in \mathcal{L}_{<0} \backslash \Lambda'_{n,n}$.
*Let* $\tilde{\pi}_{I_\downarrow}^{(\theta)}$ *be the FK model on* $G_{I_\downarrow}$ *with probabilities* $p = 1 - e^{-\beta}$ *in* $E$ *and* $\theta$ *in* $\tilde{E}$, *and boundary
conditions wired on* $\mathcal{L}_{<0} \cap \partial \Lambda'_{n,n}$ *(and free elsewhere). Then* $\Xi_j^{\mathsf{fl}}(I)$ *and* $\Xi_j^{\mathsf{dob}}(I)$ *from* (2.18) *satisfy*

$$\log \Xi_j^{\mathsf{fl}}(I) - \log \Xi_j^{\mathsf{dob}}(I) = \sum_{e \in \tilde{E}} \int_0^1 \frac{\tilde{\pi}_{(\Theta_j I)_\downarrow}^{(\theta)}(\omega_e = 1) - \tilde{\pi}_{I_\downarrow}^{(\theta)}(\omega_e = 1)}{\theta(1 - \theta)} \, \mathrm{d}\theta \,. \tag{2.19}$$

*Proof.* We need to define two counterparts to $\tilde{\pi}_{I_\downarrow}^{(\theta)}$. First, let $G_{I_\uparrow}$ be the induced subgraph of $\Lambda_{n,n}'$ on $\widehat{\mathcal{V}}_{\mathsf{top}}(I)$, and define $\pi_{I_\uparrow}$ to be the FK model on $G_{I_\uparrow}$ (all edge probabilities are $p$, i.e., there is no special edge set $\tilde{E}$) with boundary conditions that are wired on $\mathcal{L}_{>0} \cap \partial\Lambda_{n,n}'$ (and free elsewhere). Second, let $G$ be the graph $\Lambda_{n,n}'$ augmented by the set of edges $\tilde{E}$ (so that $G_{I_\downarrow}$ is its induced subgraph on $\widehat{\mathcal{V}}_{\mathsf{top}}^c(I)$), and let $\tilde{\pi}_G^{(\theta)}$ be its corresponding FK model, with edge probabilities $\theta$ on $\tilde{E}$, and boundary conditions as in $\bar{\pi}_n^{\mathsf{dob}}$, conditioned on no path connecting the $\mathsf{top}$ boundary $\mathcal{L}_{>0} \cap \partial\Lambda_{n,n}'$ to the $\mathsf{bot}$ boundary $\mathcal{L}_{<0} \cap \partial\Lambda_{n,n}'$. Notice that $\tilde{\pi}_G^{(\theta)}$ interpolates between the boundary conditions $\mathsf{dob}$ and $\mathsf{fl}$:

$$\tilde{\pi}_G^{(0)} = \bar{\pi}_{\Lambda_{n,n}'}^{\mathsf{dob}} \quad , \quad \tilde{\pi}_G^{(1)}(\omega\!\restriction_{\Lambda_{n,n}} \in \cdot) = \bar{\pi}_{\Lambda_{n,n}}^{\mathsf{fl}}. \tag{2.20}$$

Letting $\tilde{Z}_G^{(\theta)}$, $\tilde{Z}_{I_\downarrow}^{(\theta)}$ and $Z_{I_\uparrow}$ denote the partition functions of $\tilde{\pi}_G^{(\theta)}$, $\tilde{\pi}_{I_\downarrow}^{(\theta)}$ and $\pi_{I_\uparrow}$, respectively, we infer the following from Domain Markov:

$$\tilde{\pi}_G^{(\theta)}(\mathcal{I}_{\mathsf{top}} = I) = \frac{1}{\tilde{Z}_G^{(\theta)}} (1-p)^{|I|} \tilde{Z}_{I_\downarrow}^{(\theta)} Z_{I_\uparrow}^*,$$

where $Z_{I_\uparrow}^*$ restricts the sum in $Z_{I_\uparrow}$ only to configurations $\omega$ where every $v \in \widehat{\mathcal{V}}_{\mathsf{top}}(I)$ adjacent to $I$ is in $\mathcal{V}_{\mathsf{top}}(\omega)$ (i.e., if $v$ is such that, for some $u$, the edge $(u,v)$ of $(\mathbb{Z}+\frac{1}{2})^3$ is dual to a plaquette $f \in I$, then there must be an open path in $\omega$ connecting $v$ to the $\mathsf{top}$ wired boundary $\mathcal{L}_{>0} \cap \partial\Lambda_{n,n}'$). Indeed, as was the case with the interface $\mathcal{I}_{\mathsf{blue}}$, which consisted of plaquettes dual to $(\mathbb{Z}+\frac{1}{2})^3$-edges between $\mathcal{V}_{\mathsf{blue}}$ and $\widehat{\mathcal{V}}_{\mathsf{blue}}^c$, the interface $\mathcal{I}_{\mathsf{top}}$ examined here consists of plaquettes dual to $(\mathbb{Z}+\frac{1}{2})^3$-edges between $\mathcal{V}_{\mathsf{top}}$ and $\widehat{\mathcal{V}}_{\mathsf{top}}^c$, supporting the above identity.

Comparing the last display for the interfaces $I$ and $\Theta_j I$ (both legal as $\max_x \overline{\mathrm{ht}}_x(I) + j \leq n$) we find that

$$\frac{\tilde{\pi}_G^{(\theta)}(\mathcal{I}_{\mathsf{top}} = \Theta_j I)}{\tilde{\pi}_G^{(\theta)}(\mathcal{I}_{\mathsf{top}} = I)} = (1-p)^{4jn} \frac{\tilde{Z}_{(\Theta_j I)_\downarrow}^{(\theta)} Z_{(\Theta_j I)_\uparrow}^*}{\tilde{Z}_{I_\downarrow}^{(\theta)} Z_{I_\uparrow}^*}.$$

Recalling Eq. (2.20) and the definition of $\Xi_j^{\mathsf{fl}}$ and $\Xi_j^{\mathsf{dob}}$, the values of this ratio at $\theta = 1$ and $\theta = 0$ are nothing but $\Xi_j^{\mathsf{fl}}(I)$ and $\Xi_j^{\mathsf{dob}}(I)$, respectively (for the latter, note that the events $\{\mathcal{I}_{\mathsf{top}} = \Theta_j I\}$ and $\{\mathcal{I}_{\mathsf{top}} = I\}$ are measurable w.r.t. $\omega\!\restriction_{\Lambda_{n,n}}$). As a consequence,

$$\frac{\Xi_j^{\mathsf{fl}}(I)}{\Xi_j^{\mathsf{dob}}(I)} = \frac{\tilde{Z}_{(\Theta_j I)_\downarrow}^{(1)} Z_{(\Theta_j I)_\uparrow}^*}{\tilde{Z}_{I_\downarrow}^{(1)} Z_{I_\uparrow}^*} \cdot \frac{\tilde{Z}_{I_\downarrow}^{(0)} Z_{I_\uparrow}^*}{\tilde{Z}_{(\Theta_j I)_\downarrow}^{(0)} Z_{(\Theta_j I)_\uparrow}^*} = \frac{\tilde{Z}_{(\Theta_j I)_\downarrow}^{(1)}}{\tilde{Z}_{I_\downarrow}^{(1)}} \cdot \frac{\tilde{Z}_{I_\downarrow}^{(0)}}{\tilde{Z}_{(\Theta_j I)_\downarrow}^{(0)}}. \tag{2.21}$$

Therefore,

$$\log \Xi_j^{\mathsf{fl}}(I) - \log \Xi_j^{\mathsf{dob}}(I) = \log \frac{\tilde{Z}_{(\Theta_j I)_\downarrow}^{(1)}}{\tilde{Z}_{(\Theta_j I)_\downarrow}^{(0)}} - \log \frac{\tilde{Z}_{I_\downarrow}^{(1)}}{\tilde{Z}_{I_\downarrow}^{(0)}}. \tag{2.22}$$

The proof is concluded by applying Claim 2.17, once to $G_{(\Theta_j I)_\downarrow}$ and once to $G_{I_\downarrow}$. ∎

We now conclude Lemma 2.16, showing that for the appropriate pairs $(I, j)$, the map $I \mapsto \Theta_j I$ is more costly under the soft floor measure than the hard floor measure.

**Proof of Lemma 2.16.** Observe that the assumption on $(I, j)$ implies that $\max_x \overline{\mathrm{ht}}_x(I) + j < 2j \leq n$, qualifying for an application of Lemma 2.18. It thus suffices to show that, for such $I$ and $j$,

$$\tilde{\pi}_{(\Theta_j I)_\downarrow}^{(\theta)}(\omega_e = 1) \geq \tilde{\pi}_{I_\downarrow}^{(\theta)}(\omega_e = 1) \tag{2.23}$$

holds for all $e \in \tilde{E}$ (implying that the right-hand of Eq. (2.19) is nonnegative). We claim that Eq. (2.23) in fact holds for every edge $e$ in $G_{I_\downarrow}$. To see this, consider $G_{(\Theta_j I)_\downarrow}$ and the associated FK

measure $\tilde{\pi}^{(\theta)}_{(\Theta_j I)_\downarrow}$, where we additionally impose that every edge $(u, v)$ dual to some $f \in I$ should be closed. The hypothesis on $j > \max_x \overline{\mathrm{ht}}(I)$ implies that this set of edges is fully included in $G_{(\Theta_j I)_\downarrow}$, and by the Domain Markov property, its marginal on $G_{I_\downarrow}$ has the same law of $\tilde{\pi}^{(\theta)}_{I_\downarrow}$. The proof is then concluded by the FKG inequality. ■

**Remark 2.19.** Our proof crucially uses the fact that the event $\mathcal{I}_{\mathsf{top}} = I$ is measurable w.r.t. the configuration "above $I$" (on $G_{I_\uparrow}$ in the notation from the proofs above), whereas it is only "below $I$" (on $G_{I_\downarrow}$) where there is a difference between the hard floor and soft floor measures, as $I \subset \mathcal{L}_{\geq 0}$ in both. Thus, the nontrivial portion ("above $I$") cancels in the ratio $\Xi^{\mathsf{fl}}_j / \Xi^{\mathsf{dob}}_j$, and leaves us with a ratio of probabilities in two standard FK models. At this point we may use FKG as $\mathcal{I}_{\mathsf{top}}$ reveals a free boundary below it, c.f. $\mathcal{I}_{\mathsf{full}}$ or $\mathcal{I}_{\mathsf{bot}}$ which would reveal a wired boundary below it, and therefore not go in the right direction. (Indeed, the result cannot hold for $\mathcal{I}_{\mathsf{bot}}$, as it would lead to a lower bound of $(1 + \varepsilon_\beta)h^*_n$ in Theorem 1.1, violating the upper bound $h^*_n$ given there.)

Similarly, this argument would not apply to $\mathcal{I}_{\mathsf{blue}}$ or $\mathcal{I}_{\mathsf{red}}$; e.g., for the Ising model, each of the events $\mathcal{I}_{\mathsf{blue}} = I$ and $\mathcal{I}_{\mathsf{red}} = I$ gives positive information on $G_{I_\downarrow}$ (and negative information on $G_{I_\uparrow}$), whence the inequality analogous to Eq. (2.23) would be reversed.

2.4. **Lower bound for the top interface.** Our aim is now to show the following logarithmic lower bound, in fact $(1 - \varepsilon)h^*_n$, on the typical height of $\mathcal{I}_{\mathsf{top}}$.

**Proposition 2.20.** *There exists an absolute constant $C > 0$ and constant $\varepsilon_\beta$ such that*

$$\bar{\pi}^{\mathsf{fl}}_n(|\{x \in (\mathbb{Z} + \tfrac{1}{2})^2 \ : \ \underline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{top}}) \leq \tfrac{1}{4\beta+C} \log n\}| > \varepsilon_\beta n^2) \leq n^{-1+\varepsilon_\beta} \ .$$

2.4.1. *Map to lift $\mathcal{I}_{\mathsf{top}}$.* We wish to consider the probability of a realization of $\mathcal{I}_{\mathsf{top}}$ versus its shift up by some height. The previous subsection established that this map can be done under the measure $\bar{\pi}^{\mathsf{dob}}_n$ rather than $\bar{\pi}^{\mathsf{fl}}_n$ for the sake of lower bounds on $\mathcal{I}_{\mathsf{top}}$.

**Lemma 2.21.** *Let $I \subset \mathcal{L}_{\geq 0}$ be a realization for $\mathcal{I}_{\mathsf{top}}$, and $j \geq 1$. Then $\Xi^{\mathsf{dob}}_j(I)$ from Eq. (2.18) has*

$$\Xi^{\mathsf{dob}}_j(I) \geq e^{-4(\beta+C)jn} \ .$$

*Proof.* It suffices to prove the lemma for $j = 1$, that is, to show that $\bar{\pi}^{\mathsf{dob}}_n(\mathcal{I}_{\mathsf{top}} = I)/\bar{\pi}^{\mathsf{dob}}_n(\mathcal{I}_{\mathsf{top}} = \Theta_1 I)$ is at most $e^{4(\beta+C)n}$ (which, when iterated $j$ times, would yield for the sought bound on $\Xi^{\mathsf{dob}}_j(I)$). Let $\mathcal{F}(I)$ be the set of full interfaces which have $I$ as the top interface; that is,

$$\mathcal{F}(I) = \{I_0 = \mathcal{I}_{\mathsf{full}}(\omega) \ : \ \omega \text{ satisfies } \mathcal{I}_{\mathsf{top}}(\omega) = I\} \ .$$

**Claim 2.22.** *Let $\omega \sim \bar{\pi}^{\mathsf{dob}}_n$. Then $\mathcal{I}_{\mathsf{top}}(\omega)$ is measurable w.r.t. $\mathcal{I}_{\mathsf{full}}(\omega)$.*

*Proof.* Let $\omega_0$ be the configuration such that $(\omega_0)_e = 0$ if and only if $e$ is dual to a plaquette $f \in I_0$. It suffices to show that for any $\omega$ with $\mathcal{I}_{\mathsf{full}}(\omega) = I_0$, we have $\widehat{\mathcal{V}}_{\mathsf{top}}(\omega) = \widehat{\mathcal{V}}_{\mathsf{top}}(\omega_0)$, as that would imply $\mathcal{I}_{\mathsf{top}}(\omega) = \mathcal{I}_{\mathsf{top}}(\omega_0)$. Since $\omega$ must be pointwise below $\omega_0$, we have immediately that $\mathcal{V}_{\mathsf{top}}(\omega) \subset \mathcal{V}_{\mathsf{top}}(\omega_0)$, and hence $\widehat{\mathcal{V}}_{\mathsf{top}}(\omega) \subset \widehat{\mathcal{V}}_{\mathsf{top}}(\omega_0)$. To show the other inclusion, suppose for contradiction that $\widehat{\mathcal{V}}_{\mathsf{top}}(\omega_0) \setminus \widehat{\mathcal{V}}_{\mathsf{top}}(\omega)$ is nonempty. Then, there must in particular be some $v \in \widehat{\mathcal{V}}_{\mathsf{top}}(\omega_0) \setminus \widehat{\mathcal{V}}_{\mathsf{top}}(\omega)$ such that $v$ is adjacent to some $u \in \widehat{\mathcal{V}}_{\mathsf{top}}(\omega_0)^c$ (if no such $v$ exists, then $\widehat{\mathcal{V}}_{\mathsf{top}}(\omega_0) \setminus \widehat{\mathcal{V}}_{\mathsf{top}}(\omega)$ must consist of finite components surrounded by vertices of $\widehat{\mathcal{V}}_{\mathsf{top}}(\omega_0) \cap \widehat{\mathcal{V}}_{\mathsf{top}}(\omega)$, but this contradicts the construction of $\widehat{\mathcal{V}}_{\mathsf{top}}(\omega)$ which would have included such finite components). Let $C(v)$ be the finite random-cluster component containing $v$. Then $C(v)$ has a 1-connected boundary of dual-to-closed plaquettes which contains the plaquette dual to the edge $(u, v)$, hence the entire 1-connected boundary is a subset of $\mathcal{I}_{\mathsf{full}}$. Thus, $v$ is in a finite component even in $\omega_0$, which together with the fact that $v$ is adjacent to $u \in \widehat{\mathcal{V}}_{\mathsf{top}}(\omega_0)^c$ implies that $v \in \widehat{\mathcal{V}}_{\mathsf{top}}(\omega_0)^c$, concluding the contradiction. ■

As a consequence of the above claim, we can write

$$\bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{top}} = I) = \bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{full}} \in \mathcal{F}(I)). \tag{2.24}$$

We claim that it will therefore suffice to show that there is a variant of the lifting map $\Theta_1$ for full interfaces, call it $\Phi_1$, such that for every $I$ and every $I_0 \in \mathcal{F}(I)$, we have

$$\bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{full}} = I_0) \le e^{4(\beta+C)n} \bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{full}} = \Phi_1 I_0), \tag{2.25}$$

and

$$\sum_{I_0 \in \mathcal{F}(I)} \bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{full}} = \Phi_1 I_0) \le e^{Cn} \bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{full}} \in \mathcal{F}(\Theta_1 I)). \tag{2.26}$$

Indeed, modulo these two inequalities, we can infer from Eq. (2.24) that

$$\bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{full}} \in \mathcal{F}(I)) = \sum_{I_0 \in \mathcal{F}(I)} \bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{full}} = I_0) \le \sum_{I_0 \in \mathcal{F}(I)} e^{4(\beta+C)n} \bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{full}} = \Phi_1 I_0)$$

$$\le e^{4(\beta+2C)n} \bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{full}} \in \mathcal{F}(\Theta_1 I)) = e^{4(\beta+2C)n} \bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{top}} = \Theta_1 I).$$

(The inequality in the first line used Eq. (2.25) and the one in the second line used Eq. (2.26).) Equation (2.25) controls the change in probability incurred by applying $\Phi_1$ (the "lifting up by 1" map) to the full interface, where we have the cluster expansion of Proposition 2.3 at our disposal. Equation (2.26) relates the probability of having a full interface that is lifted by 1 to that of having a top interface that is shifted by 1 (more precisely, an union bound on the former probability: the left-hand is an upper bound on $\bar{\pi}_n^{\mathsf{dob}}(\mathcal{I}_{\mathsf{full}} \in \Phi_1 \mathcal{F}(I))$ via the non-disjoint events $\{\mathcal{I}_{\mathsf{full}} = \Phi_1 I_0\}$). (Note that, unlike the shift map $\Theta_1$ that acts on (a realization of) $\mathcal{I}_{\mathsf{top}}$, working with a full interface must be done with extra care, as it is not a surface—it consists of additional "hairs" of dual-closed plaquettes—the culprit behind the fact that the shift map $\Phi_1$ acting on $\mathcal{I}_{\mathsf{full}}$ is not a bijection.)

Let $\Phi_1$ be the map on a full interface $I_0$ such that $\Phi_1(I_0) = J_0$, where the plaquette set of $J_0$ is the union of the set of plaquettes in $I_0$ shifted up by 1, together with the set of $4n$ vertical plaquettes at height $1/2$ which are adjacent to the boundary vertices in $\partial \Lambda_{n,n}$. Recall that in the context of $\bar{\pi}_n^{\mathsf{dob}}$, we are using $\partial_{\mathsf{red}} \Lambda'_{n,n} = \mathcal{L}_{>0} \cap \partial \Lambda'_{n,n}$ and $\partial_{\mathsf{blue}} \Lambda'_{n,n} = \mathcal{L}_{<0} \cap \partial \Lambda'_{n,n}$.

**Claim 2.23.** *For every realization $I_0$ for $\mathcal{I}_{\mathsf{full}}$ under $\bar{\pi}_n^{\mathsf{dob}}$, the map $\Phi_1$ outputs a valid full interface $J_0 = \Phi_1(I_0)$ separating $\partial_{\mathsf{blue}} \Lambda'_{n,n}$ from $\partial_{\mathsf{red}} \Lambda'_{n,n}$.*

*Proof.* We need to show that $J_0$ is 1-connected and separates $\partial_{\mathsf{blue}} \Lambda'_{n,n}$ from $\partial_{\mathsf{red}} \Lambda'_{n,n}$.

The fact that $J_0$ is 1-connected follows from observing that the translation of $I_0$ up by height 1 is 1-connected, the $4n$ vertical plaquettes at height $1/2$ adjacent to boundary vertices in $\partial \Lambda_{n,n}$ are 1-connected, and these two sets are also 1-adjacent.

To show that $J_0$ is separating, suppose for contradiction that there is a path $P$ from $x \in \partial_{\mathsf{red}} \Lambda'_{n,n}$ to $y \in \partial_{\mathsf{blue}} \Lambda'_{n,n}$ that does not cross any plaquette of $J_0$. Let us first consider the case where $x \in \mathcal{L}_{\ge 3/2}$. Then we can consider the translation $P'$ of the path $P$ by height $-1$. Since $P$ does not cross any plaquette of $J_0$, the path $P'$ does not cross any plaquette of $I_0$; so, the fact $x - \mathfrak{e}_3 \in \partial_{\mathsf{red}} \Lambda'_{n,n}$ implies that $P'$ is a path from $\partial_{\mathsf{red}} \Lambda'_{n,n}$ to $\partial_{\mathsf{blue}} \Lambda'_{n,n}$ that does not cross a plaquette of $I_0$, contradicting the fact that $I_0$ is a valid interface. For the remaining case $x \in \mathcal{L}_{1/2} \cap \partial \Lambda'_{n,n} =: X$, the new (rather than shifted) plaquettes added to $J_0$ (of which there are at most $4n$) ensure that any path in $\Lambda'_{n,n}$ from $X$ to $X^c$ must first pass through a vertex $x' \in \mathcal{L}_{3/2} \cap \partial \Lambda'_{n,n}$, putting us back in the first case. $\blacksquare$

**Claim 2.24.** *We have the following bounds for the terms in the cluster expansion:*

$$|J_0| - |I_0| \leq 4n\,, \qquad |\partial J_0| - |\partial I_0| \leq Cn\,, \qquad \kappa(J_0) \geq \kappa(I_0)\,,$$

$$\left| \sum_{f \in J_0} \mathbf{g}(f, J_0) - \sum_{f \in I_0} \mathbf{g}(f, I_0) \right| \leq Cn\,.$$

*Hence, we have by Proposition 2.3 for the* full *interface that*

$$\bar{\pi}_n^{\mathsf{dob}}(I_0) \leq e^{4(\beta+C)n} \bar{\pi}_n^{\mathsf{dob}}(J_0)\,.$$

*Proof.* The first inequality $|J_0| - |I_0| \leq 4n$ follows by construction of $J_0$. The second inequality similarly follows because the only difference lies in adding the up to $4n$ plaquettes to $J_0$, and each plaquette added can only change $\partial J_0$ by a constant. Adding plaquettes (equivalent to closing edges) can only increase the number of open clusters, so $\kappa(J_0) \geq \kappa(I_0)$. Finally, call $I_1$ the set of plaquettes of $I_0$ shifted up by 1, and for each $f \in I_0$ let $f' \in I_1$ be the copy of $f$ shifted up by 1. We then have

$$\left| \sum_{f \in J_0} \mathbf{g}(f, J_0) - \sum_{f \in I_0} \mathbf{g}(f, I_0) \right| \leq \sum_{f \in J_0 \setminus I_1} |\mathbf{g}(f, J_0)| + \sum_{f : f' \in J_0 \cap I_1} |\mathbf{g}(f', J_0) - \mathbf{g}(f, I_0)|\,.$$

The first term is bounded by $4Kn$, and we can bound the second term since the radius $r(f, I_0; f', J_0)$ is obtained by the distance to either one of the up to $4n$ plaquettes in $J_0 \setminus I_1$ or one of the $4n$ horizontal boundary plaquettes at height 0 in the Dobrushin boundary conditions. If the union of these two plaquette sets is called $F$, then we have

$$\sum_{f : f' \in J_0 \cap I_1} |\mathbf{g}(f', J_0) - \mathbf{g}(f, I_0)| \leq \sum_{f' \in J_0 \cap I_1} \sum_{g \in F} K e^{-cd(f', g)} \leq K|F| \leq K'n\,. \qquad \blacksquare$$

Combining the above two claims proves Eq. (2.25). For Eq. (2.26), we need the following claim.

**Claim 2.25.** *Every $J_0$ in the image of the map $\Phi_1$ has at most $e^{Cn}$ preimages $I_0$.*

*Proof.* Denote by $I_1$, as in the previous proof, the plaquettes of $I_0$ shifted up by 1, and let us decompose $J_0 = I_1 \cup (J_0 \setminus I_1)$. If we know which of the $4n$ vertical plaquettes adjacent to boundary vertices at height $1/2$ are in $I_1$ and which are in $J_0 \setminus I_1$, then we can recover $I_1$ and hence $I_0$ from the image. There are at most $2^{4n}$ subsets of possible plaquettes to attribute to $J_0 \setminus I_1$, as claimed. $\qquad \blacksquare$

We will now argue that

$$\left\{ \mathcal{I}_{\mathsf{full}} \in \bigcup_{I_0 \in \mathcal{F}(I)} \Phi_1 I_0 \right\} \subset \left\{ \mathcal{I}_{\mathsf{full}} \in \mathcal{F}(\Theta_1 I) \right\}, \tag{2.27}$$

which together with Claim 2.25 will establish Eq. (2.26).

Let $\omega_0$ be the configuration that only has closed edges $e$ dual to plaquettes $f \in I_0$. Similarly, let $\omega_1$ be the configuration that only has closed edges $e$ dual to plaquettes $f \in \Phi_1 I_0$. We need to show that $\mathcal{V}_{\mathsf{top}}(\omega_1)$ is nothing but $\Theta_1 \mathcal{V}_{\mathsf{top}}(\omega_0)$ (the set of vertices in $\mathcal{V}_{\mathsf{top}}(\omega_0)$ shifted up by 1).

Notice that if we take $\omega_0$, shift all edges up by 1, and then close the $4n$ edges connecting a boundary vertex in $\partial_{\mathsf{red}} \Lambda'_{n,n}$ at height $1/2$ to a vertex in $\Lambda'_{n,n} \setminus \partial_{\mathsf{red}} \Lambda'_{n,n}$ also at height $1/2$, then we exactly get $\omega_1$. Call $\omega'_1$ the configuration obtained by shifting the edges of $\omega_0$ up by 1 (without closing extra edges).

First we show that $\Theta_1 \mathcal{V}_{\mathsf{top}}(\omega_0) \subset \mathcal{V}_{\mathsf{top}}(\omega_1)$. Indeed, note that for every path $P$ of open edges in $\omega_0$ connecting a vertex $x$ to $\partial_{\mathsf{red}} \Lambda'_{n,n}$, its shift up by 1, call it $P'$, will be a path of open edges in $\omega'_1$ connecting $x + \mathfrak{e}_3$ to $\partial_{\mathsf{red}} \Lambda'_{n,n}$. It remains to show that closing the up to $4n$ edges to obtain $\omega_1$ does not cut off this path. Indeed, if one of those edges were in the path $P'$, this would imply that $P'$ contains a boundary vertex in $\mathcal{L}_{1/2}$; thus, $P$ would contain a boundary vertex in $\mathcal{L}_{-1/2}$, contradicting the fact that there is no path of open edges from $\partial_{\mathsf{blue}} \Lambda'_{n,n}$ to $\partial_{\mathsf{red}} \Lambda'_{n,n}$.

Now we show that $\mathcal{V}_{\text{top}}(\omega_1) \subset \Theta_1 \mathcal{V}_{\text{top}}(\omega_0)$. If $x \in \mathcal{V}_{\text{top}}(\omega_1)$ and $P'$ is a path of open edges in $\omega_1$ connecting $x$ to a boundary vertex $y \in \partial_{\text{red}}\Lambda'_{n,n}$, then it must have such a path to a $y$ at height at least $3/2$ (as all $4n$ faces along $\mathcal{L}_{3/2} \cap \partial\Lambda'_{n,n}$ have been made dual-to-closed). Then $P'$ is still an open path in $\omega'_1$ (since $\omega'_1 \geq \omega_1$), and its shift down by 1, call it $P$, is an open path in $\omega_0$. Thus $P$ is an open path connecting $x - \mathfrak{e}_3$ to $y - \mathfrak{e}_3$, and $y - \mathfrak{e}_3 \in \partial_{\text{red}}\Lambda'_{n,n}$ since $y - \mathfrak{e}_3 \in \mathcal{L}_{\geq 1/2}$.

This establishes Eq. (2.27), thereby concluding the proof of the lemma. ∎

2.4.2. *Injecting entropy via downward spikes.* Now that we have controlled the weight change of lifting $\mathcal{I}_{\text{top}}$ under $\bar{\pi}_n^{\text{fl}}$ via its weight change under $\bar{\pi}_n^{\text{dob}}$, we can return to the hard-floor measure and inject entropic downward spikes to argue that this lifting and the added opportunity for downward spikes to the interface is preferential to having predominantly been at too low a height.

**Proof of Proposition 2.20**. Let $H := \frac{14}{\beta}\log n$ and let $h := \frac{1}{4\beta+C_0}\log n$ for $C_0$ to be chosen later. For $\varepsilon_\beta$, define the set $\mathbf{I}_{\text{bad}}$ to be the set of possible realizations $I$ of $\mathcal{I}_{\text{top}}$ satisfying

$$\left|\left\{x \in (\mathbb{Z} + \tfrac{1}{2})^2 : \underline{\text{ht}}_x(I) \leq h\right\}\right| > \varepsilon_\beta n^2 \qquad \text{and} \qquad \max_x \overline{\text{ht}}_x(I) \leq H.$$

By a union bound with Theorem 2.7, it will suffice for us to show

$$\bar{\pi}_n^{\text{fl}}(\mathcal{I}_{\text{top}} \in \mathbf{I}_{\text{bad}}) \leq e^{-n}. \tag{2.28}$$

Fix any $I \in \mathbf{I}_{\text{bad}}$. Consider the operation $\Theta_H$ which lifts $I$ by $H$ and adds all $4Hn$ vertical plaquettes along the boundary at heights between 0 and $H$. By Lemma 2.16 and Lemma 2.21, we have

$$\frac{\bar{\pi}_n^{\text{fl}}(\mathcal{I}_{\text{top}} = \Theta_H I)}{\bar{\pi}_n^{\text{fl}}(\mathcal{I}_{\text{top}} = I)} \geq e^{-4(\beta+C)Hn}. \tag{2.29}$$

Next, for a subset $A$ of $xy$-plane index points $A \subset (\mathbb{Z} + \frac{1}{2})^2 \times \{0\}$ such that $\underline{\text{ht}}_x(I) \leq h$, consider the following operation $\Phi_A$ on $\{\omega : \mathcal{I}_{\text{top}} = \Theta_H I\}$. Obtain $\Phi_A(\omega)$ by, for every $x \in A$,

(1) opening the $h$ vertical edges (dual to horizontal plaquettes) between vertices in the column

$$v_i^x = x + (\underline{\text{ht}}_x(\Theta_H I) + \tfrac{1}{2} - i)\mathfrak{e}_3, \qquad i = 0, \ldots, h,$$

(if they are not open already).

(2) closing the $4h$ horizontal edges (dual to vertical plaquettes) incident to vertices $(v_i^x)_{i=1,\ldots,h}$, as well as the one vertical edge (dual to a horizontal plaquette) between $v_h^x$ and $v_h^x - \mathfrak{e}_3$ (if they are not closed already).

In $\omega$, all vertices $v_0^x$ are in $\mathcal{V}_{\text{top}}$, and so step (1) above is exactly adding a downward vertical spike of height $h$ to $\mathcal{V}_{\text{top}}$ at every column above $x \in A$. The resulting $\omega$ does not violate the disconnection event $\mathfrak{D}_{n,n}$ because every vertex added to $\mathcal{V}_{\text{top}}$ by step (1) is disconnected from $\partial_{\text{blue}}\Lambda_{n,n}$ by the union of the original $\Theta_H I$ and the edges closed in step (2). By direct comparison of weights, notice that for every $\omega$, we have

$$\frac{\bar{\pi}_n^{\text{fl}}(\Phi_A(\omega))}{\bar{\pi}_n^{\text{fl}}(\omega)} \geq \left(\left(\frac{1-p}{p}\right)^{4h+1} q^{4h+1}\right)^{|A|}. \tag{2.30}$$

From a realization of $\Phi_A(\omega)$, we can read off the set $A$ by first reading off the set $\mathcal{I}_{\text{top}}(\Phi_A(\omega))$, then finding all index points $x \in \mathcal{L}_0$ above which the interface $\mathcal{I}_{\text{top}}(\Phi_A(\omega))$ takes a height below $H$. Indeed, every such $x$ is a member of $A$ because only vertices in columns through $A$ are added to $\mathcal{V}_{\text{top}}$ by the operation, and all members of the original set $\mathcal{V}_{\text{top}}$ were above height $H$ in $\Theta_H I$. Moreover, every member of $A$ is counted in this set of $x$ because $\underline{\text{ht}}_x(\Theta_H I) - h = H + \underline{\text{ht}}_x(I) - h$ is below $H$. As a result, for every $\omega'$ in the image of $\Phi_A$ (there can only one $A$ for which this is the case since $A$ is determined by $\omega'$),

$$\left|\{\omega : \Phi_A(\omega) = \omega'\}\right| \leq 2^{(5h+1)|A|},$$

by reading off $A$, and $\underline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{top}}(\omega'))$ for every $x \in A$ from $\omega'$ as above, reading off from that $v_i^x$ for all $x \in A$, then enumerating over the $2^{5h+1}$ possible values $\omega$ could take on the incident edges for the $h$ vertices $(v_i^x)_{i=1}^h$. From any such $\omega$, we can also read off $I$ (as compared to $\Theta_H I$) because the map $\Theta_H$ is a bijection for $I \subset \mathcal{L}_{\geq 0}$.

Putting the above together, we therefore have that

$$1 \geq \sum_{I \in \mathbf{I}_{\mathsf{bad}}} \sum_{A \subset \{x:\underline{\mathrm{ht}}_x I \leq h\}} \sum_{\omega' \in \Phi_A(\{\omega:\mathcal{I}_{\mathsf{top}}(\omega)=\Theta_H I\})} \bar{\pi}_n^{\mathsf{fl}}(\omega')$$

$$\geq \sum_{I \in \mathbf{I}_{\mathsf{bad}}} \sum_{\omega:\mathcal{I}_{\mathsf{top}}(\omega)=\Theta_H I} \sum_{A \subset \{x:\underline{\mathrm{ht}}_x I \leq h\}} 2^{-(5h+1)|A|} e^{-(4\beta+C)h|A|} \bar{\pi}_n^{\mathsf{fl}}(\omega)$$

where the right-hand side of (2.30) became $e^{-(4\beta+C)h|A|}$. Using that for every $I \in \mathbf{I}_{\mathsf{bad}}$, one has $|\{x : \underline{\mathrm{ht}}_x I \leq h\}| > \varepsilon_\beta n^2$, then up to a change of the constant $C$, one has

$$1 \geq \sum_{I \in \mathbf{I}_{\mathsf{bad}}} \bar{\pi}_n^{\mathsf{fl}}(\mathcal{I}_{\mathsf{top}} = \Theta_H I)\left(1 + e^{-(4\beta+C)h}\right)^{|\{x:\underline{\mathrm{ht}}_x I \leq h\}|} \geq \sum_{I \in \mathbf{I}_{\mathsf{bad}}} \bar{\pi}_n^{\mathsf{fl}}(\mathcal{I}_{\mathsf{top}} = \Theta_H I)\left(1 + e^{-(4\beta+C)h}\right)^{\varepsilon_\beta n^2}.$$

By (2.29), the above becomes

$$1 \geq \left(1 + e^{-(4\beta+C)h}\right)^{\varepsilon_\beta n^2} \sum_{I \in \mathbf{I}_{\mathsf{bad}}} e^{-(4\beta+C)Hn} \bar{\pi}_n^{\mathsf{fl}}(\mathcal{I}_{\mathsf{top}} = I)$$

$$= \left(1 + e^{-(4\beta+C)h}\right)^{\varepsilon_\beta n^2} e^{-(4\beta+C)Hn} \bar{\pi}_n^{\mathsf{fl}}(\mathcal{I}_{\mathsf{top}} \in \mathbf{I}_{\mathsf{bad}}).$$

Using $1 + x \geq e^{x/2}$ for $x \leq 1$ and rearranging, we get

$$\bar{\pi}_n^{\mathsf{fl}}(\mathcal{I}_{\mathsf{top}} \in \mathbf{I}_{\mathsf{bad}}) \leq \exp\left(-\frac{\varepsilon_\beta}{2} n^2 e^{-(4\beta+C)h} + (4\beta+C)Hn\right).$$

By our choices of $h$ and $n$, it becomes clear that there is a choice of $C_0$, say $C_0 = 2C$, such that the negative term in the exponential is larger than $n^{1+\varepsilon'_\beta}$ (regardless of $\varepsilon_\beta$), while the positive term is $O(n \log n)$, implying the requisite bound of (2.28). ∎

**Remark 2.26.** The final step in the proof of Proposition 2.20 was to derive entropic repulsion via comparing the energetic cost of the shift map with the entropy gain from planting spikes consisting of straight columns of height $h$. This part of the argument is suboptimal since $\mathcal{I}_{\mathsf{top}}$ is in fact propelled by the large deviation rate of its downward deviations, which are typically random-walk like as opposed to straight columns. We expect that after lifting $\mathcal{I}_{\mathsf{top}}$ by $C \log n$ (for a large enough $C > 0$) using Lemma 2.21 so that the hard floor is far from the interface, the detailed analysis in [13], which showed entropic repulsion driven by pillar rates for the Ising interface above a soft floor, can be adapted to the interface $\mathcal{I}_{\mathsf{top}}$. This would replace the lower bound $\frac{1}{4\beta+C_0} \log n$ in Proposition 2.20 (and consequently the one in Theorem 1.1) by $(1 - o(1))\frac{1}{\xi'} \log n$, where $\xi'$ is the analog of $\xi$ in $\pi_\infty^{\mathsf{w}}$ (the rate of point-to-plane connections in the complement of the infinite cluster).

2.5. **Extending the lower bound to the blue interface.** We would now like to extend to $\mathcal{I}_{\mathsf{blue}}$ the lower bound of $\frac{1}{4\beta+C} \log n$ obtained in Proposition 2.20 on the height of most plaquettes in $\mathcal{I}_{\mathsf{top}}$. Note that, to that end, the additive error from Corollary 2.15 (whereby $\mathcal{I}_{\mathsf{blue}}$ is sandwiched between $\mathcal{I}_{\mathsf{bot}}, \mathcal{I}_{\mathsf{top}}$) would be far too crude; instead, we will rely on the following result.

**Lemma 2.27.** *There is an absolute constant $C > 0$ such that, if $\delta = C/\beta$, then*

$$\bar{\pi}_n^{\mathsf{fl}}\left(|\mathcal{I}_{\mathsf{full}}| \geq (1+\delta)n^2\right) \leq e^{-n^2}.$$

*In particular, on this event, $\underline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{full}}) = \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{full}})$ for all but at most $\delta n^2$ plaquettes $x \in (\mathbb{Z}+\frac{1}{2})^2$, and under $\mathbf{P}_n^{\mathsf{fl}}$ all four interfaces $\mathcal{I}_{\mathsf{bot}}, \mathcal{I}_{\mathsf{blue}}, \mathcal{I}_{\mathsf{red}}, \mathcal{I}_{\mathsf{top}}$ coincide with $\mathcal{I}_{\mathsf{full}}$ on the corresponding plaquette.*

*Proof.* Let $I_0$ be the "trivial interface," consisting of the $n^2$ horizontal plaquettes of $\mathcal{L}_0 \cap \Lambda_{n,n}$. As mentioned in Remark 2.4, the cluster expansion of Proposition 2.3 for $\mathcal{I}_{\mathsf{full}}$ is valid also for $\bar{\pi}_n^{\mathsf{fl}}$ (taking $m_1 = 0$ and $m_2 = n$ to arrive at $\Lambda_{n,n}$, and $\mathsf{h} = 0$ for the appropriate boundary conditions). Using this (namely, Eq. (2.7)), if $I$ is realization of a full interface with $n^2 + s$ plaquettes, then

$$\frac{\bar{\pi}_n^{\mathsf{fl}}(\mathcal{I}_{\mathsf{full}} = I)}{\bar{\pi}_n^{\mathsf{fl}}(\mathcal{I}_{\mathsf{full}} = I_0)} \leq \left(\frac{(1-p)q}{p^C}\right)^s e^{K(2n^2+s)} \leq e^{Cn^2 - \beta s},$$

with $K > 0$ the absolute constant from Proposition 2.3 and $C > 0$ another absolute constant. Note that, unlike the more careful analysis of the map $\Phi_1$ in Section 2.4.1 (which would correspond to bounding $|\sum_{f \in I} \mathbf{g}(f, I) - \sum_{f' \in I_0} \mathbf{g}(f', I_0)|$ via absorbing $K$ for each $f \in I \oplus I_0$, and another constant contribution for each such $f$ from summing over the exponential decay of its effect on other plaquettes through the $\mathbf{g}$ term) we resorted to the naive bound $(|I| + |I_0|)\|\mathbf{g}\|_\infty \leq (2 + s)Kn^2$; this is due to $I_0$ being adjacent to the boundary of the box, whereby $r(f, I; f', I_0)$ is 1 for all $f' \in I_0$ (demonstrating some of the difficulty in overcoming the potential pinning effect of the hard floor).

There are at most $e^{C(n^2+s)}$ interfaces $I$ with $n^2 + s$ plaquettes, for some other absolute $C > 0$ (viewed as a rooted subgraph of size $n^2 + s$ in a bounded-degree graph). Thus,

$$\bar{\pi}_n^{\mathsf{fl}}(|\mathcal{I}_{\mathsf{full}}| \geq (1+\delta)n^2) \leq \sum_{s \geq \delta n^2} \sum_{I:|I|=n^2+s} \frac{\bar{\pi}_n^{\mathsf{fl}}(\mathcal{I}_{\mathsf{full}} = I)}{\bar{\pi}_n^{\mathsf{fl}}(\mathcal{I}_{\mathsf{full}} = I_0)} \leq e^{C'n^2 - \beta \delta n^2},$$

and taking $\delta = (C' + 1)/\beta$ concludes the proof. ∎

**Proof of lower bound in Theorem 1.1.** By Lemma 2.27, we can assume that all but at most $\delta n^2 = Cn^2/\beta$ plaquettes of $\mathcal{I}_{\mathsf{full}}$ are horizontal plaquettes with no other plaquettes of $\mathcal{I}_{\mathsf{full}}$ sharing the same $xy$-coordinates (simply by the fact that every $x \in (\mathbb{Z} + \frac{1}{2})^2 \times 0$ needs to have at least one plaquette above it, and this already accounts for $n^2$ plaquettes). In particular, each of these plaquettes are common to $\mathcal{I}_{\mathsf{full}}$, $\mathcal{I}_{\mathsf{blue}}$ and $\mathcal{I}_{\mathsf{top}}$ and are one of the minimal height plaquettes of $\mathcal{I}_{\mathsf{top}}$ studied in Proposition 2.20. Hence, on the event $|\mathcal{I}_{\mathsf{full}}| < (1+\delta)n^2$, we have

$$\left\{|\mathcal{I}_{\mathsf{blue}} \cap \mathcal{L}_{\leq \frac{1}{4\beta+C} \log n}| > \varepsilon'_\beta n^2\right\} \subset \left\{|\{x \in (\mathbb{Z} + \tfrac{1}{2})^2 : \underline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{top}}) \leq h\}| > \varepsilon'_\beta n^2 - \delta n^2\right\},$$

which concludes the proof via Proposition 2.20 so long as $\varepsilon'_\beta - \delta$ is bigger than the $\varepsilon_\beta$ there, which is of course attained via $\varepsilon'_\beta = \varepsilon_\beta + C/\beta$. This proves that the lower bound of Theorem 1.1 holds with probability $1 - n^{-1+\varepsilon_\beta}$. ∎

## 3. Sharp upper bound

In this section, we prove the sharp upper bound of $h_n^*$ (1.4) on the typical interface height in Theorem 1.1. By the monotonicity argument of Proposition 2.9 via the fuzzy Potts model, it essentially suffices to establish such an upper bound on the typical interface height under the "soft-floor" Potts measure, i.e., conditioned on $\mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0}$, and with the Dobrushin boundary conditions going from red to blue at height $\mathsf{h} = h_n^*$. Namely, we will show that under $\widehat{\mu}_n^{h_n^*} = \mu_n^{h_n^*}(\cdot \mid \mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0})$ as per Eq. (2.10), the interface has fewer than $\varepsilon_\beta n^2$ many plaquettes above $h_n^*$.

The proof of this will go in two parts, by lower bounding the probability of the event $\{\mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0}\}$ under $\mu_n^{h_n^*}$, and at the same time, upper bounding under $\mu_n^{h_n^*}$ (which is rigid about height $h_n^*$) the large deviation probability of its interface having more than $\varepsilon_\beta n^2$ plaquettes at heights $\geq h_n^* + 1$. The ratio of these will bound the probability of the latter event under the conditional measure $\widehat{\mu}_n^{h_n^*}$.

3.1. **Remark on finite box vs. cylinder.** As a technicality it turns out to be convenient to work with the Potts interfaces in infinite and semi-infinite cylinders (keeping the $xy$-side lengths $n$). The following lemma is standard (similar to e.g., [14, Lem. 8],[15, Lem 7.117]) and shows that these induce essentially the same measure on interfaces. Recall $\Lambda_{n,\ell}$ and $\Lambda'_{n,\ell}$ from (2.3)–(2.5).

**Lemma 3.1.** *Consider the Potts measures* $\mu^{\mathsf{h}}_{\Lambda'_{n,\ell}}$ *and* $\mu^{\mathsf{h}}_{\Lambda'_{n,m}}$ *for* $n \le \ell \le m$ *and with* $|\mathsf{h}| \le \ell/2$. *Then,*

$$\|\mu^{\mathsf{h}}_{\Lambda'_{n,\ell}}(\mathcal{I}_{\mathsf{blue}} \in \cdot) - \mu^{\mathsf{h}}_{\Lambda'_{n,m}}(\mathcal{I}_{\mathsf{blue}} \in \cdot)\|_{\mathrm{TV}} \le \exp(-n)\,.$$

*Similarly, if we consider* $\mu^{\mathsf{fl}}_{\Lambda_{n,\ell}}, \mu^{\mathsf{fl}}_{\Lambda_{n,m}}$ *with* $\ell \le m$, *then*

$$\|\mu^{\mathsf{fl}}_{\Lambda_{n,\ell}}(\mathcal{I}_{\mathsf{blue}} \in \cdot) - \mu^{\mathsf{fl}}_{\Lambda_{n,m}}(\mathcal{I}_{\mathsf{blue}} \in \cdot)\|_{\mathrm{TV}} \le \exp(-n)\,.$$

*Proof.* Sample random-cluster and Potts configurations $(\omega, \sigma) \sim \mathbf{P}^{\mathsf{h}}_{\Lambda'_{n,\ell}}$ and $(\omega', \sigma') \sim \mathbf{P}^{\mathsf{h}}_{\Lambda'_{n,m}}$ independently. We say a set $\mathcal{S}$ of edges is a *separating surface in* $\omega$ *between* $\mathcal{L}_{3\ell/4}$ *and* $\mathcal{L}_\ell$ if $\omega(\mathcal{S}) \equiv 1$ and every 1-connected path of plaquettes in $\Lambda_{n,\ell}$ between $\mathcal{L}_{3\ell/4}$ and $\mathcal{L}_\ell$ must have a plaquette dual to some edge of $\mathcal{S}$. If $\omega \wedge \omega'$ is the pointwise minimal configuration, define the event $\mathcal{G}$ as

(1) there exists a separating set $\mathcal{S}_N$ in $\omega \wedge \omega'$ between $\mathcal{L}_{3\ell/4}$ and $\mathcal{L}_\ell$, and
(2) there exists a separating set $\mathcal{S}_S$ in $\omega \wedge \omega'$ between $\mathcal{L}_{-\ell}$ and $\mathcal{L}_{-3\ell/4}$.

On the event $\mathcal{G}$, expose the highest realizing $\mathcal{S}_N$ and the lowest realizing $\mathcal{S}_S$. Even though there is no order on separating surfaces, the highest (resp., lowest) such is well-defined as the outer boundary of the set generated by exposing the 1-connected component of $\partial_N \Lambda'_{n,\ell}$ (resp., $\partial_S \Lambda'_{n,\ell}$) in the set of plaquettes dual to edges that are closed.

The highest $\mathcal{S}_N$ is measurable with respect to the states of edges above it, and the lowest $\mathcal{S}_S$ with respect to the states of edges below it. Therefore, by the domain Markov property, we can resample the configurations $(\omega, \sigma)$ and $(\omega', \sigma')$ in the region sandwiched between these separating surfaces using the identity coupling (the colorings on $\mathcal{S}_N$ must be red in both, and on $\mathcal{S}_S$ blue in both, so indeed the measures induced on the regions sandwiched by the separating surfaces have the same distribution). Since $\mathcal{I}_{\mathsf{full}}$ and therefore $\mathcal{I}_{\mathsf{blue}}$ must be sandwiched between these two separating surfaces when $|\mathsf{h}| \le \ell/2$, on the event $\mathcal{G}$ we have

$$\mathcal{I}_{\mathsf{blue}}(\omega, \sigma) \equiv \mathcal{I}_{\mathsf{blue}}(\omega', \sigma')\,.$$

It therefore suffices to give the upper bound

$$\mathbf{P}^{\mathsf{h}}_{\Lambda'_{n,\ell}} \otimes \mathbf{P}^{\mathsf{h}}_{\Lambda'_{n,m}}\big((\omega, \omega') \notin \mathcal{G}\big) \le e^{-n}\,.$$

In order to bound this probability, notice that in each of the two measures, by Theorem 2.5 and Remark 2.6, the probability of $\mathcal{I}_{\mathsf{full}} \cap \mathcal{L}_{3\ell/4} \ne \emptyset$ or $\mathcal{I}_{\mathsf{full}} \cap \mathcal{L}_{-3\ell/4}$ is at most $e^{-c\beta n}$. Under both distributions, we can therefore expose the outer boundaries of $\mathcal{I}_{\mathsf{full}}$ above and below, which will induce a pair of wired random-cluster measures (with no conditioning) dominating the infinite-volume super-critical random-cluster measure. By standard Peierls bounds for the super-critical random-cluster measure at large $\beta$, with probability at least $1 - e^{-c\beta n}$, there are common surfaces $\mathcal{S}_N$ and $\mathcal{S}_S$ giving the event $\mathcal{G}$. Indeed, on the complement of the event $\mathcal{G}$, either for north or south (say w.l.o.g. for north) and either for $\omega$ or $\omega'$ (say w.l.o.g. $\omega$), there is a 1-connected path of plaquettes between $\mathcal{L}_{3\ell/4}$ and $\mathcal{L}_\ell$ such that at least half of the edges dual to these plaquettes are closed. This probability is bounded by the exponential tails on sizes of closed components under the infinite-volume random-cluster measure).

The argument for the floor measures $\mu^{\mathsf{fl}}_{\Lambda_{n,\ell}}$ and $\mu^{\mathsf{fl}}_{\Lambda_{n,m}}$ is analogous, only needing the existence of $\mathcal{S}_N$, and no need for $\mathcal{S}_S$ since the southern boundary is already shared at height zero. The only exception is that we cannot apply Theorem 2.5 directly to bound $\mathbf{P}^{\mathsf{fl}}_{\Lambda_{n,\ell}}(\mathcal{I}_{\mathsf{full}} \cap \mathcal{L}_{3\ell/4} \ne \emptyset)$, so we will additionally use a monotonicity argument involving $\mathcal{I}_{\mathsf{blue}}$ (in the same way that we proved an upper bound for $\mathcal{I}_{\mathsf{top}}$ going through $\mathcal{I}_{\mathsf{blue}}$ first in Section 2.2). Compare $\mathbf{P}^{\mathsf{fl}}_{\Lambda_{n,\ell}}$ to $\mathbf{P}^{\ell/2}_{\Lambda_{n,\ell}}$. The latter is just a shift of the measure $\mathbf{P}^{\mathsf{dob}}_{\Lambda'_{n,\ell/2}}$ by $\ell/2$, so we have by Theorem 2.5 that

$$\mathbf{P}^{\ell/2}_{\Lambda_{n,\ell}}(\mathcal{I}_{\mathsf{full}} \cap \mathcal{L}_{3\ell/4} \ne \emptyset) = \mathbf{P}^{\mathsf{dob}}_{\Lambda'_{n,\ell/2}}(\mathcal{I}_{\mathsf{full}} \cap \mathcal{L}_{\ell/4} \ne \emptyset) \le e^{-c\beta n}\,.$$

Since $\mathcal{I}_{\text{blue}} \subset \mathcal{I}_{\text{full}}$, the same bound applies for $\mathcal{I}_{\text{blue}}$. Now looking at the Potts marginals, the fuzzy Potts monotonicity and the fact that $\max_x \overline{\text{ht}}_x(\mathcal{I}_{\text{blue}})$ is increasing in the set of blue vertices implies that we also have

$$\mu_{\Lambda_{n,\ell}}^{\text{fl}}(\mathcal{I}_{\text{blue}} \cap \mathcal{L}_{3\ell/4} \neq \emptyset) \leq e^{-c\beta n} \,.$$

(The same argument is done at the end of the proof of Proposition 2.9, see for details.)  Finally, this implies an upper bound of $\mathbf{P}_{\Lambda_{n,\ell}}^{\text{fl}}(\mathcal{I}_{\text{full}} \cap \mathcal{L}_{7\ell/8} \neq \emptyset) \leq e^{-c\beta n}$ by Corollary 2.15.                    ∎

As such, we perform all arguments in this section on infinite and semi-infinite cylinders, establish the sharp upper bound of Theorem 1.1 under $\mu_{\Lambda_{n,\infty}}^{\text{fl}}$, which we denote in this section by $\mu_n^{\text{fl}}$ for brevity, and then translate it to the upper bound of Theorem 1.1 under $\mu_{\Lambda_{n,n}}^{\text{fl}}$ via Lemma 3.1 with $m$ sent to $\infty$.

We further consider the Potts model on the infinite cylinder $\Lambda'_{n,\infty}$ wherein $\mu_{\Lambda'_{n,\infty}}^{\text{dob}}$ assigns Dobrushin boundary conditions: red in the positive half-space and blue in the negative one. Overriding the notation from Section 2, we define $\mu_n^{\text{h}}$ as the $q$-state Potts model on $\Lambda'_{n,\infty}$ with Dobrushin boundary conditions that transition from blue to red at height h, so that $\mu_{\Lambda'_{n,\infty}}^{\text{dob}} = \mu_n^0$. Let $\widehat{\mu}_n^{\text{h}}$ be the same measure conditioned on $\mathcal{I}_{\text{blue}} \subset \mathcal{L}_{\geq 0}$ as per (2.10). The stochastic domination observation of Proposition 2.9 evidently still holds in the infinite cylinder, by taking the ceiling height to infinity and using the fact that the interface is almost surely finite.

### 3.2. Lower bounding the probability of positivity of the blue interface.

In this subsection, we establish large deviation rates on the maximum height oscillation of a rigid Potts interface. Namely, we establish the following lower bound on the probability of positivity of $\mathcal{I}_{\text{blue}}$ under the no-floor measure $\mu_n^{h_n^*}$.

**Lemma 3.2.** *For every $\beta > \beta_0$, there exists $\varepsilon_\beta$ (going to zero as $\beta \to \infty$) such that*

$$\mu_n^{h_n^*}(\mathcal{I}_{\text{blue}} \subset \mathcal{L}_{\geq 0}) \geq e^{-(1+\varepsilon_\beta)n} \,.$$

Since we are only looking for lower bounds, the proof will go by applying the FKG inequality through the fuzzy Potts model defined in Section 2.1. By vertical translation, it suffices to consider the downwards oscillations of $\mathcal{I}_{\text{blue}}$ from height zero under $\mu_n^{\text{dob}} = \mu_n^0$, and by reflection, these have the same law as the upwards oscillations of $\mathcal{I}_{\text{red}}$. Thus, we focus on the latter, i.e., on bounding

$$\mu_n^{\text{dob}}(\max_x \overline{\text{ht}}_x(\mathcal{I}_{\text{red}}) \leq h_n^*) \geq e^{-(1+\varepsilon_\beta)n} \,. \tag{3.1}$$

We need the appropriate monotone event in the set of nonR sites, for which we have a sharp understanding of the rate as it relates to $\xi$ from Eq. (1.3). Recall the definition of the augmented set of red vertices $\widehat{\mathcal{V}}_{\text{red}}$ from Definition 2.1, whose boundary forms $\mathcal{I}_{\text{red}}$.

**Definition 3.3** (Non-red pillar). Let $x$ be a vertex at height $1/2$. The non-red pillar at $x$, denoted $\mathcal{P}_x^{\text{nonR}}$, is the connected component of vertices in $\widehat{\mathcal{V}}_{\text{red}}^c \cap \mathcal{L}_{\geq 0}$ that contains $x$. The height of the $\mathcal{P}_x^{\text{nonR}}$, denoted $\text{ht}(\mathcal{P}_x^{\text{nonR}})$ is the maximum height of a plaquette $f$ dual to an edge $(y, z)$ where $y \notin \mathcal{P}_x^{\text{nonR}}$ and $z \in \mathcal{P}_x^{\text{nonR}}$.

We can then define the pillar rate under the infinite-volume Dobrushin measure

$$\alpha_h = \alpha_h(\beta, q) := -\log \mu_\infty^{\text{dob}}(\text{ht}(\mathcal{P}_{\mathfrak{o}}^{\text{nonR}}) \geq h) \,. \tag{3.2}$$

where we are using the subscript $\infty$ to mean the infinite-volume measure obtained as the $n \to \infty$ limit of $\mu_n^{\text{dob}}$, and we recall that $\mathfrak{o}$ is the origin vertex $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. Section 4 relates this pillar rate to $\xi_h$ from (1.3) and thus to $h_n^*$, but most of the estimates of this section will be done in terms of $\alpha_h$.

The following lemma explains why we move to expressing the event in terms of pillars.

**Lemma 3.4.** *The quantities $\max_x \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{red}})$ and $\max_x \mathrm{ht}(\mathcal{P}_x^{\mathsf{nonR}})$ are the same. Moreover, the event $\{\mathrm{ht}(\mathcal{P}_x^{\mathsf{nonR}}) \geq h\}$ is decreasing in the set of $\mathsf{red}$ vertices, and is equivalent to the event $\mathcal{A}_{x,h}^{\mathsf{nonR}}$ from [5, Def. 5.1].*

*Proof.* Observe that $\widehat{\mathcal{V}}_{\mathsf{red}}^c$ is connected (see e.g., [5, Remark 2.15] for this) and contains $\partial_{\mathsf{blue}} \Lambda_{n,\infty}'$, and so any vertex of $\widehat{\mathcal{V}}_{\mathsf{red}}^c$ which has height $\geq 1/2$ is part of some non-red pillar $\mathcal{P}_x^{\mathsf{nonR}}$. In the other direction, every vertex in $\mathcal{P}_x^{\mathsf{nonR}}$ for some $x$ is an element of $\widehat{\mathcal{V}}_{\mathsf{red}}^c$. Hence, $\max_x \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{red}})$, which is equal to the maximum height of $\widehat{\mathcal{V}}_{\mathsf{red}}^c$, is equal to the maximum height of $\mathcal{P}_x^{\mathsf{nonR}}$ over all $x$.

For the monotonicity claim, changing a vertex from $\mathsf{nonR}$ to $\mathsf{red}$ can only increase the set $\mathcal{V}_{\mathsf{red}}$, and consequently can only increase $\widehat{\mathcal{V}}_{\mathsf{red}}$.

Finally, note that $\mathcal{A}_{x,h}^{\mathsf{nonR}}$ was defined in [5, Def. 5.1] as the event that there exists a $\widehat{\mathcal{V}}_{\mathsf{red}}^c$-path from $x$ to height $h - 1/2$ using only vertices of a random set $\mathcal{P}_x$ (which was defined as a random-cluster pillar), but this random set $\mathcal{P}_x$ is always a subset of $\mathcal{L}_{>0}$ that contains $\mathcal{P}_x^{\mathsf{nonR}}$, so restricting to $\mathcal{P}_x$ is the same as restricting to $\mathcal{L}_{>0}$ when looking at the connected component of $\widehat{\mathcal{V}}_{\mathsf{red}}^c$ vertices containing $x$. ∎

By [5, Prop. 2.24] with the observation that the random-cluster pillar $\mathcal{P}_x$ defined there necessarily contains $\mathcal{P}_x^{\mathsf{nonR}}$, we have the following weak bound on the height of $\mathcal{P}_x^{\mathsf{nonR}}$.

**Lemma 3.5.** *There exists $C > 0$ such that for all $h \geq 1$, all $x \in \mathcal{L}_{1/2}$, and sufficiently large $\beta$,*

$$\mu_n^{\mathsf{dob}}(\mathrm{ht}(\mathcal{P}_x^{\mathsf{nonR}}) \geq h) \leq \exp(-4(\beta - C)h).$$

We also have the following decorrelation estimate from [5, Cor. A.7], which lets us use the rate $\alpha_h$ of Eq. (3.2) for the pillar deviation rate inside the finite domain $\Lambda_{n,\infty}'$ so long as $x$ is not too close to the boundary.

**Lemma 3.6.** *There exists a constant $C$ such that for all $x, y$ so that $d(x, \partial \Lambda_{n,\infty}') \wedge d(y, \partial \Lambda_{m,\infty}') \geq r$,*

$$|\mu_n^{\mathsf{dob}}(\mathrm{ht}(\mathcal{P}_x^{\mathsf{nonR}}) \geq h) - \mu_m^{\mathsf{dob}}(\mathrm{ht}(\mathcal{P}_x^{\mathsf{nonR}}) \geq h)| \leq e^{-r/C}.$$

With the above two ingredients recalled, and the observations of Lemma 3.4, we can show the following lower bounds on the large deviations of the maximum height of $\mathcal{I}_{\mathsf{red}}$.

**Lemma 3.7.** *For every $1 \leq h \leq \frac{1}{\sqrt{\beta}} \log n$, we have*

$$\mu_n^{\mathsf{dob}}(\max_x \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{red}}) < h) \geq \exp\left(-(1 + o(1))n^2 e^{-\alpha_h}\right).$$

*Consequently by reflection symmetry,*

$$\mu_n^{\mathsf{dob}}(\min_x \underline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{blue}}) > -h) \geq \exp\left(-(1 + o(1))n^2 e^{-\alpha_h}\right).$$

*Proof.* By the first part of Lemma 3.4, we can write

$$\mu_n^{\mathsf{dob}}\left(\max_x \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{red}}) < h\right) = \mu_n^{\mathsf{dob}}\left(\bigcap_{x \in \mathcal{L}_{1/2}} \{\mathrm{ht}(\mathcal{P}_x^{\mathsf{nonR}}) < h\}\right).$$

As in Section 2.1, we look at the fuzzy Potts model this time singling out the color $\mathsf{red}$ and viewing all non-red colors as $\mathsf{fuzzy}$. Since this model has the FKG inequality (2.2) (identifying all $\mathsf{blue}$ boundary vertices as a single $\mathsf{fuzzy}$ vertex to encode that they are obligated to take the same $\mathsf{nonR}$ color) and by the second part of Lemma 3.4, we have that

$$\mu_n^{\mathsf{dob}}\left(\bigcap_x \{\mathrm{ht}(\mathcal{P}_x^{\mathsf{nonR}}) < h\}\right) \geq \prod_x \mu_n^{\mathsf{dob}}(\mathrm{ht}(\mathcal{P}_x^{\mathsf{nonR}}) < h).$$

where the product runs over $x \in (\mathbb{Z} + \frac{1}{2})^2 \times \{0\} \cap \Lambda'_n$. Terms of $x : d(x, \partial\Lambda'_{n,\infty}) \leq \log^2 n$ are lower bounded by the complement of Lemma 3.5. For $x$ such that $d(x, \partial\Lambda'_{n,\infty}) > \log^2 n$, we can apply Lemma 3.6 with $m \to \infty$ to lower bound the right-hand side by

$$\mu_n^{\mathsf{dob}}(\mathrm{ht}(\mathcal{P}_x^{\mathsf{nonR}}) \geq h) \leq e^{-\alpha_h} + e^{-\log^2 n/C} = (1 + o(1))e^{-\alpha_h},$$

where the equality used the facts that $h \leq O(\log n)$ and $\alpha_h \leq 4(\beta + C)h$ per [5, Proposition 1.5] (a simple forcing argument to lower bound the probability of a deviation via the simplest straight column pillar). Hence, using the inequality $1 - x \geq e^{-x/(1-x)}$ for $0 < x < 1$, we get

$$\prod_x \mu_n^{\mathsf{dob}}(\mathrm{ht}(\mathcal{P}_x^{\mathsf{nonR}}) < h) \geq (1 - e^{-4(\beta-C)h})^{4n\log^2 n} \cdot (1 - (1 + o(1))e^{-\alpha_h})^{(n-\log^2 n)^2}$$

$$\geq \exp\left(-e^{-4(\beta-C)h}4n\log^2 n(1 + o(1)) - e^{-\alpha_h}n^2(1 + o(1))\right)$$

$$= \exp\left(-e^{-\alpha_h}n^2(1 + o(1))\right),$$

where the last equality used the upper bound on $h$ and the fact that $\alpha_h \leq 4(\beta + C)h$ (so that $e^{8Ch} \leq n^{\varepsilon_\beta}$ cannot make up for the difference between $n\log^2 n$ and $n^2$). ∎

We now plug in $h = h_n^* + 1$ to obtain Lemma 3.2, given the comparison of $\alpha_h$ and $\xi h$ for $\xi$ from (1.3) (which will separately be established in Section 4).

**Proof of Lemma 3.2.** Theorem 4.1 will prove that $\alpha_h$ is very close (up to an $\varepsilon_\beta$ independent of $h$ exactly related to) to $\xi h$. In particular, by Theorem 4.1 and definition of $h_n^*$ (1.4),

$$\alpha_{h_n^*+1} \geq \xi \cdot (\lfloor \xi^{-1}\log n \rfloor + 1) - \varepsilon_\beta \geq \log n - \varepsilon_\beta. \tag{3.3}$$

By applying the second part of Lemma 3.7 with $h = h_n^* + 1$, and translating vertically by $h_n^*$, we conclude. ∎

3.3. **Large deviations upper bound on number of sites not at $h_n^*$.** The second part to proving the $h_n^*$ upper bound in Theorem 1.1 is establishing that violating an upper bound on the typical height under $\mu_n^{h_n^*}$ is even more atypical than the $e^{-(1+\varepsilon_\beta)n}$ lower bound given on interface positivity in Lemma 3.2. Note that this is an estimate purely on a no-floor rigid interface.

**Lemma 3.8.** *There exists $C > 0$ such that for every $\beta > \beta_0$*

$$\mu_n^{h_n^*}\left(\left|\{x : \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{blue}}) \geq h_n^* + 1\}\right| \geq \frac{C}{\sqrt{\beta}}n^2\right) \leq \exp(-\sqrt{\beta}n/C).$$

We will prove the above by bounding the set of all index points where the full random-cluster interface is not a single horizontal plaquette, under the corresponding Dobrushin random-cluster measure. By vertical translation let us work with $\bar{\pi}_n^{\mathsf{dob}} = \bar{\pi}_n^0$.

We need to recall the definition of interface walls for $\mathcal{I}_{\mathsf{full}}$ from [14]. For a plaquette $f$, let $\rho(f)$ be its projection onto $\mathcal{L}_0$. Recall that two plaquettes are 1-connected if their intersection contains a unit line segment, and 0-connected if their intersection contains a point.

**Definition 3.9** (Walls and Ceilings of $\mathcal{I}_{\mathsf{full}}$). Let $\mathcal{I}_{\mathsf{full}}^*$ be the union of $\mathcal{I}_{\mathsf{full}}$ (defined in Definition 2.2) together with all plaquettes which are horizontal and 1-connected to $\mathcal{I}_{\mathsf{full}}$. If a plaquette $f$ is the unique plaquette in $\mathcal{I}_{\mathsf{full}}^*$ with projection onto $\mathcal{L}_0$ equal to $\rho(f)$, then $f$ is a ceiling plaquette. All other plaquettes of $\mathcal{I}_{\mathsf{full}}^*$ are wall plaquettes. A ceiling is a 0-connected component of ceiling plaquettes. A *wall* is a pair $(A, B)$ where $W$ is a 0-connected component of wall plaquettes, $A = W \cap \mathcal{I}_{\mathsf{full}}$, and $B = W \cap (\mathcal{I}_{\mathsf{full}}^* \setminus \mathcal{I}_{\mathsf{full}})$. (This decomposition is called a standard wall in [15]. By [15, Lem. 7.125(i)], there are no ceiling plaquettes in $\mathcal{I}_{\mathsf{full}}^* \setminus \mathcal{I}_{\mathsf{full}}$, so we do not need such a decomposition for ceilings.)

**Lemma 3.10** ([14, Lem. 12],[15, Lem. 7.127]). *In $\mathcal{I}_{\mathsf{full}}$, there is a 1-1 correspondence between interfaces and admissible families of walls (quotiented out by vertical translations), where two walls are called admissible if their projections are not 0-connected and they are contained in the domain (see items (i)–(ii) on page 211 of [15] for the formal statement).*

**Definition 3.11** (Excess area). The excess area between two interfaces $\mathcal{I}$ and $\mathcal{J}$ is the difference in their number of plaquettes, $\mathfrak{m}(\mathcal{I}; \mathcal{J}) := |\mathcal{I}| - |\mathcal{J}|$. The excess area of a wall $W = (A, B)$ of $\mathcal{I}_{\mathsf{full}}$, is $\mathfrak{m}(W) := |A| - |\rho(W)| := |A| - |\rho(A \cup B)|$. (When we write $|\rho(F)|$ for a set of faces $F$, we mean the area, or number of faces in $\rho(F)$, so projections of vertical faces do not contribute.)

For a wall $W$, the projection of its *hull*, $\rho(\dot{W})$, is the union of $\rho(W)$ with all finite components of $\mathbb{R}^2 \setminus \rho(W)$. Say a wall $W$ of $\mathcal{I}_{\mathsf{full}}$ is *outermost* if there is no wall $W'$ of $\mathcal{I}_{\mathsf{full}}$ such that $\rho(W) \subset \rho(\dot{W}')$. Finally, two walls are called mutually external if neither $\rho(W) \subset \rho(\dot{W}')$ nor $\rho(W') \subset \rho(\dot{W})$.

This combinatorial construction (along with some further constructions we do not need in what follows) were used in [14] to establish rigidity of $\mathcal{I}_{\mathsf{full}}$ under $\mu_n^{\mathsf{dob}}$. In particular, Peierls-type maps based on those in [6] were used to delete walls from the 1-to-1 representation of interfaces, to get exponential tail bounds on the excess areas of walls of $\mathcal{I}_{\mathsf{full}}$. The following is an immediate corollary of [14, Lem. 15] (see also [15, Lem. 7.132]).

**Corollary 3.12.** *There exists $C > 0$ such that for every $\beta > \beta_0$, and every arbitrary set of admissible mutually external walls $W_1, \ldots, W_N$,*

$$\bar{\pi}_n^{\mathsf{dob}}(W_1, \ldots, W_N) \le \exp\left( - (\beta - C) \sum_{i=1}^{N} \mathfrak{m}(W_i) \right),$$

*where we are using $\bar{\pi}_n^{\mathsf{dob}}(W_1, \ldots, W_N)$ to mean the probability that $\{W_1, \ldots, W_N\}$ are walls of $\mathcal{I}_{\mathsf{full}}$.*

Given Corollary 3.12, the following bound uses a union bound over the contribution from walls from diadically growing scales interpolating between those confining interior areas 1 to $n^2$.

**Lemma 3.13.** *Let $\mathfrak{U}$ be the set of all outermost walls of the interface $\mathcal{I}_{\mathsf{full}}$ under the no-floor Dobrushin random-cluster measure $\bar{\pi}_n^{\mathsf{dob}}$. There exists $C > 0$ such that for all $\beta > \beta_0$,*

$$\bar{\pi}_n^{\mathsf{dob}}\left( \sum_{W \in \mathfrak{U}} |\rho(\dot{W})| \ge \frac{C}{\sqrt{\beta}} n^2 \right) \le e^{-\sqrt{\beta} n / C}. \tag{3.4}$$

*Proof.* Partition the set of all outermost walls into the sets $\mathfrak{U}_1, \ldots, \mathfrak{U}_{2L}$ for $L = \log_2 n$, where

$$\mathfrak{U}_k = \{W \in \mathfrak{U} : |\rho(\dot{W})| \in [2^{k-1}, 2^k)\}. \tag{3.5}$$

We will show that for a suitable absolute constant $C_0$, for each $k = 1, \ldots, 2L$,

$$\bar{\pi}_n^{\mathsf{dob}}\left( \sum_{W \in \mathfrak{U}_k} |\rho(\dot{W})| \ge \varepsilon_{\beta,k}^2 n^2 \right) \le e^{-(\beta - C) 2^{-\frac{k}{2}} \varepsilon_{\beta,k}^2 n^2} \quad \text{where} \quad \varepsilon_{\beta,k} = \frac{C_0}{\beta^{1/4} \min\{k, 2L + 1 - k\}}. \tag{3.6}$$

Let us first conclude the proof using (3.6): for an absolute constant $C$, we have,

$$\sum_{k=1}^{2L} \varepsilon_{\beta,k}^2 \le \frac{C}{\sqrt{\beta}},$$

so the union of the above events covers $\sum_{W \in \mathfrak{U}} |\rho(\dot{W})| \ge \frac{C}{\sqrt{\beta}} n^2$. Summing the probability bounds of (3.6), we can bound the probability by $2L$ times the maximum of the probabilities, which is evidently attained at the term $k = 2L$ where it is at most $\exp(-C_0(\sqrt{\beta} - \frac{C}{\sqrt{\beta}})n)$, yielding the claimed (3.4) up to the change of constant.

We now move to proving (3.6). Fix $k$ and note by (3.5) and isoperimetry in $\mathbb{Z}^2$, if $|\mathfrak{U}_k| = N$,

$$N \leq \sum_{W \in \mathfrak{U}_k} |\rho(\dot{W})| 2^{1-k}, \qquad \text{and} \qquad \sum_{W \in \mathfrak{U}_k} \mathfrak{m}(W) \geq 2 \sum_{W \in \mathfrak{U}_k} |\rho(\dot{W})| 2^{-k/2}. \qquad (3.7)$$

By applying Corollary 3.12, for any fixed realization of $\mathfrak{U}_k$, we have

$$\bar{\pi}_n^{\mathsf{dob}}\Big(\{W \in \mathfrak{U}_k\} \text{ are walls in } \mathcal{I}_{\mathsf{full}}\Big) \leq \exp\Big(-(\beta - C)\sum_{W \in \mathfrak{U}_k} \mathfrak{m}(W)\Big).$$

We union bound over collections $\mathfrak{U}_k$ as follows: letting $\chi$ be the fraction $\frac{1}{n^2}\sum_{W \in \mathfrak{U}_k}|\rho(\dot{W})|$,

$$\bar{\pi}_n^{\mathsf{dob}}\Big(\sum_{W \in \mathfrak{U}_k} |\rho(\dot{W})| \geq (\delta n)^2\Big) \leq \sum_{(\delta n)^2 \leq \chi n^2 \leq n^2} \sum_{N \leq \chi n^2 2^{1-k}} \binom{n^2}{N} \sum_{r \geq \chi n^2 2^{-\frac{k}{2}+1}} C^r e^{-(\beta - C)r}$$

$$\leq \sum_{(\delta n)^2 \leq \chi n^2 \leq n^2} \sum_{N \leq \chi n^2 2^{1-k}} \binom{n^2}{N} e^{-2^{1-\frac{k}{2}}(\beta - C)\chi n^2}.$$

The first line used (3.7) for the bounds on the sums, the binomial is for the choice of $N$ points at which to root the walls of $\mathfrak{U}_k$, and the factor of $C^r$ is for the number of possible ways to allocate excess area $r$ to the different root points and enumerate over the realizations of the walls. Using the bound $\sum_{j \leq pA} \binom{A}{j} \leq \exp(\mathsf{H}(p)A)$ where $\mathsf{H}(p)$ is the binary entropy function $-p\log p - (1-p)\log(1-p)$, we get

$$\bar{\pi}_n^{\mathsf{dob}}\Big(\sum_{W \in \mathfrak{U}_k} |\rho(\dot{W})| \geq (\delta n)^2\Big) \leq \sum_{(\delta n)^2 \leq \chi n^2 \leq n^2} \exp\Big(\Big(\mathsf{H}(\chi 2^{1-k}) - 2^{1-\frac{k}{2}}(\beta - C)\chi\Big)n^2\Big).$$

It therefore suffices for us to show (now with the choice of $\delta = \varepsilon_{\beta,k}^2$) that for every $k$, for every $\chi \geq \varepsilon_{\beta,k}^2$, we have $\mathsf{H}(\chi 2^{1-k}) \leq (\beta - C)\chi 2^{-k/2}$ (we would then absorb the prefactor of $n^2$ and be done). To show this, we use the bound $\mathsf{H}(p) \leq p \log \frac{1}{p} + p$ (which holds as long as $p \leq 1/2$, which is the case for all $k \geq 2$ trivially, and for $k = 1$ by the fact that there cannot be distinct walls each of interior size only 1 confining half the area of $\Lambda$). It suffices to show that

$$2^{1-k} \log \frac{2^{k-1}}{\chi} + 2^{1-k} \leq (\beta - C)2^{-\frac{k}{2}},$$

or, equivalently,

$$(k-1)(\log 2) + \log \chi^{-1} + 1 \leq (\beta - C)2^{\frac{k}{2}-1}. \qquad (3.8)$$

Let us consider first the case where $k = 1, \ldots, L$, in which case $\log \chi^{-1} \leq \log \beta + \log(k^2) - \log(C_0^2)$. We can take $C_0$ large (independent of $\beta$ because $2^{\frac{k}{2}-1}\beta$ beats $\log \beta$ for all $k \geq 1$) to only consider large values of $k$, whence the exponential growth on the right-hand side of (3.8) of course dominates the left. Turning to the case where $k = L+1, \ldots, 2L$, now $\log \chi^{-1} \leq \log \beta + 2\log(2L+1-k) - \log(C_0^2)$. For every $k \geq L$, the left-hand side of (3.8) is thus at most $O(\log n)$, whereas the right-hand side is at least $(\beta - C)n^{1/2}/2$. ∎

**Proof of Lemma 3.8.** It is equivalent by vertical translation to establish the upper bound on

$$\mu_n^{\mathsf{dob}}\Big(\big|\{x : \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{blue}}) \geq 1\}\big| \geq \frac{C}{\sqrt{\beta}}n^2\Big) \leq \bar{\pi}_n^{\mathsf{dob}}\Big(\big|\{x : \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{full}}) \geq 1\}\big| \geq \frac{C}{\sqrt{\beta}}n^2\Big),$$

where we used inclusion of $\mathcal{I}_{\mathsf{blue}} \subset \mathcal{I}_{\mathsf{full}}$ and the FK–Potts coupling. Now any $x$ having $\overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{full}}) \neq 0$ or $\underline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{full}}) \neq 0$ must belong to $\rho(\dot{W})$ for some wall $W$. Thus, the right-hand side is bounded by Lemma 3.13, concluding the proof. ∎

We are now in position to combine everything to get the sharp upper bound in Theorem 1.1.

**Proof of upper bound in Theorem 1.1**. By Lemma 3.2 and Lemma 3.8,

$$\widehat{\mu}_n^{h_n^*}\Big(\big|\{x : \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{blue}}) \geq h_n^* + 1\}\big| \geq \tfrac{C}{\sqrt{\beta}}n^2\Big) \leq \frac{\mu_n^{h_n^*}\Big(\big|\{x : \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{blue}}) \geq h_n^* + 1\}\big| \geq \tfrac{C}{\sqrt{\beta}}n^2\Big)}{\mu_n^{h_n^*}(\mathcal{I}_{\mathsf{blue}} \subset \mathcal{L}_{\geq 0})}$$

$$\leq e^{-\sqrt{\beta}n/C + (1+\varepsilon_\beta)n}\,,$$

yielding a bound of $e^{-\sqrt{\beta}n/C}$ for large $\beta$, up to a change of the absolute constant $C$.

To conclude the upper bound of Theorem 1.1, we note that $|\{x : \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{blue}}) \geq h_n^* + 1\}| \geq \frac{C}{\sqrt{\beta}}n^2$ is a monotone event in the set of blue vertices (because it is monotone in the augmented blue vertex set $\widehat{\mathcal{V}}_{\mathsf{blue}}$, and in turn $\mathcal{V}_{\mathsf{blue}}$ which is expressed as a union of vertices connected by blue paths to $\partial_{\mathsf{blue}}\Lambda'_{n,\infty}$), and thus Proposition 2.9 gives the same for $\mu_n^{\mathsf{fl}}$.

Finally, as done in the proof of the lower bound, we observe that on the event $|\mathcal{I}_{\mathsf{full}}| < (1 + \frac{C}{\beta})n^2$ (which by Lemma 2.27 occurs with probability $1 - e^{-n^2}$ and hence can be assumed), we have

$$\big\{|\mathcal{I}_{\mathsf{blue}} \cap \mathcal{L}_{\geq h_n^*+1}| > \tfrac{C}{\sqrt{\beta}}n^2 + \tfrac{C}{\beta}n^2\big\} \subset \big\{|\{x : \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{blue}}) \geq h_n^* + 1\}| > \tfrac{C}{\sqrt{\beta}}n^2\big\}\,,$$

which concludes the proof of the upper bound with the choice of $\varepsilon_\beta = \frac{C}{\sqrt{\beta}}$, holding with probability $1 - e^{-\sqrt{\beta}n/C}$. ∎

## 4. IDENTIFICATION OF THE PILLAR RATES WITH INFINITE-VOLUME CONNECTIVITY RATES

The preceding sections established our Theorem 1.1 up to the deferred identification of the pillar rate $\alpha$ with the with the point-to-plane connectivity rate $\xi$. In this section, our goal is to give the precise relation between the pillar rate and the point-to-plane connectivity rate, and establish the additivity of the latter rate to be able to identify $h_n^*$ with $\lfloor \xi^{-1} \log n \rfloor$ per (1.4).

### 4.1. Comparison of pillar rate to point-to-plane connectivity rate.

Our approach is to work conditional on certain high probability events on the shape of the interface pillar under $\mu_\infty^{\mathsf{dob}}$ and the connected component of $\mathcal{V}_{\mathsf{red}}^c$ under $\mu_\infty^{\mathsf{red}}$ under which we can decorrelate the discrepancies in the boundary conditions using cluster expansion machinery. Recall that we denote the origin by $\mathsf{o} = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, and recall the definitions of the two rates we are interested in comparing,

$$\alpha_h = -\log \mu_\infty^{\mathsf{dob}}\big(\mathrm{ht}(\mathcal{P}_{\mathsf{o}}^{\mathsf{blue}}) \geq h\big)\,, \quad \text{and} \quad \xi_h = -\log \mu_\infty^{\mathsf{red}}\big(\mathsf{o} \longleftrightarrow \mathcal{L}_h \text{ in } \mathcal{V}_{\mathsf{red}}^c\big)\,.$$

For simplicity, define $A_{\mathsf{o},h} = \{\mathsf{o} \longleftrightarrow \mathcal{L}_h \text{ in } \mathcal{V}_{\mathsf{red}}^c\}$. The main goal of this section is to establish the following.

**Theorem 4.1.** *Fix any $q \geq 2$. We have for all $h \geq 1$ that*

$$|\alpha_h - \xi h| \leq \varepsilon_\beta\,.$$

For this, we will need a generalization of Proposition 2.3, which may also be of independent interest. We begin with some preliminary definitions. Recall for any collection of plaquettes $F$, we let $\overline{F}$ denote the set of plaquettes which are 1-connected to $F$ (including $F$ itself), and set $\partial F := \overline{F} \setminus F$. If there is a reference domain $\mathcal{U}$, then $\overline{F}$ is presumed to restrict to $\mathcal{U}$. Let $f_e$ denote the plaquette $f$ dual to an edge $e$. Let $\mathsf{Cl}(F)$ be the event that for all $f_e \in F$, $\omega_e = 0$ and for all $f_e \in \partial F$, $\omega_e = 1$, i.e., that $F$ forms a 1-connected dual-to-closed component.

Now let $\mathcal{U}$ be any subgraph of $\mathbb{Z}^3$ and let $\eta$ be any boundary condition. Let $\mathcal{B}$ be the set of vertices in $\mathcal{U}$ which are also adjacent to a vertex in $\mathbb{Z}^3 \setminus \mathcal{U}$. Viewing $\eta$ in the dual graph as a set of dual-to-open and dual-to-closed plaquettes, we will say a collection of plaquettes $F$ is a compatible set for $(\mathcal{U}, \eta)$ if for every dual-to-closed plaquette $f$ of $\eta$ in $\mathbb{Z}^3 \setminus \mathcal{U}$, every plaquette $g$ which is 1-connected to $f$ and in $\mathcal{U}$ is in $\overline{F}$.

Finally, for $\mathcal{U}$ finite, let $Z^\eta(\mathcal{U})$ denote the partition function for the random-cluster measure on $\mathcal{U}$ with boundary conditions $\eta$.

**Proposition 4.2.** *There exists a constant $K > 0$ and functions $\mathbf{g}(f, F, \mathcal{U})$ taking as inputs $F$ a collection of plaquettes, $f \in F$, and $\mathcal{U}$ a subgraph of the lattice $\mathbb{Z}^3$, such that $\mathbf{g}$ satisfies the following:*

(1) *For any $F$, $f \in F$, and $\mathcal{U} \subset \mathbb{Z}^3$,*

$$|\mathbf{g}(f, F, \mathcal{U})| \leq e^{-\beta/K}\,.$$

(2) *For any $f \in F$, $f' \in F'$, and any domains $\mathcal{U}, \mathcal{U}' \subset \mathbb{Z}^3$, we have*

$$|\mathbf{g}(f, F, \mathcal{U}) - \mathbf{g}(f', F', \mathcal{U}')| \leq e^{-\beta r/K}\,,$$

*where $r$ is the largest radius such that $B_r(f) \cap \mathcal{U} \cong B_r(f') \cap \mathcal{U}'$ and $B_r(f) \cap F \cong B_r(f') \cap F'$, with $B_r(f)$ denoting the $L_\infty$ ball of radius $r$ about the center of $f$, and $\cong$ denoting equivalence (either as graphs or plaquette sets) up to translation.*

(3) *For any $\mathcal{U} \subset \mathbb{Z}^3$ with boundary condition $\eta$, and any compatible $F$ for $(\mathcal{U}, \eta)$,*

$$\pi_{\mathcal{U}}^{\eta}(\mathsf{Cl}(F)) = \frac{Z^{\mathsf{w}}(\mathcal{U})}{Z^{\eta}(\mathcal{U})}(1 - e^{-\beta})^{|\partial F|} e^{-\beta|F|} q^{\kappa(F)-1} \exp\Big( \sum_{f \in F} \mathbf{g}(f, F, \mathcal{U}) \Big)\,,$$

*where $\mathsf{w}$ is the wired boundary conditions, $\kappa(F)$ is the number of connected components of $\mathcal{U}$ after removing the edges $e$ such that $f_e \in F$, and we are viewing vertices which are wired together by open edges of $\eta$ in $\mathbb{Z}^3 \setminus \mathcal{U}$ as connected.*

The proof of Proposition 4.2 is deferred to Appendix A.

**Remark 4.3.** Note that the above proposition can be applied in the setting where $\eta$ is the Dobrushin boundary condition and $F$ is a potential realization of $\mathcal{I}_{\mathsf{full}}$, recovering the law of $\mathcal{I}_{\mathsf{full}}$ as in [15, Lem. 7.104]. Even in this case however, we emphasize that the upper bounds in Items 1 and 2 now decay with $\beta$, whereas they did not in the formulation of Proposition 2.3.

Because we wish to use the above cluster expansion machinery on finite graphs to prove Theorem 4.1, we will work in the finite graph $\Lambda'_{n,n}$. This poses no issues, as we have

$$|\log \mu_{\infty}^{\mathsf{dob}}\big(\mathrm{ht}(\mathcal{P}_{\mathsf{o}}^{\mathsf{nonR}}) \geq h\big) - \log \mu_{n}^{\mathsf{dob}}\big(\mathrm{ht}(\mathcal{P}_{\mathsf{o}}^{\mathsf{nonR}}) \geq h\big)| = o(1)\,,$$

and

$$|\mu_{\infty}^{\mathsf{red}}\big(A_{\mathsf{o},h}\big) - \mu_{n}^{\mathsf{red}}\big(A_{\mathsf{o},h}\big)| = o(1)\,,$$

by weak convergence of measures (and moreover with explicit rates of convergence via Lemma 3.6 and Proposition 4.2). Thus, for each $h$, we can take $n \gg h$ large enough so that the error in using a finite box instead of infinite volume is smaller than $\varepsilon_\beta$. Hence, in the remainder of this section, the lemmas will be stated in terms of the infinite volume quantities $\alpha_h$ and $\xi_h$, but the proofs will work with measures on $\Lambda'_{n,n}$.

Recall that we have the FK–Potts measure on $\Lambda'_{n,n}$ with Dobrushin boundary conditions denoted $\mathbf{P}_n^{\mathsf{dob}}$ with marginals $(\mu_n^{\mathsf{dob}}, \bar{\pi}_n^{\mathsf{dob}})$. We also have the FK–Potts measure with wired red boundary conditions denoted $\mathbf{P}_n^{\mathsf{red}}$ with marginals $(\mu_n^{\mathsf{red}}, \pi_n^{\mathsf{w}})$.

**Lemma 4.4.** *Let $\beta$ be sufficiently large and $q \geq 2$. We have that for all $h \geq 1$,*

$$\big|\alpha_h - (\xi_h - 2\beta + \log(q-1))\big| \leq \varepsilon_\beta\,.$$

*Proof.* The discrepancy of $2\beta - \log(q-1)$ occurs because of differences of the marginal on $\mathsf{o} = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, so we will isolate this by defining $\mathcal{E}^{\mathsf{dob}}$ and $\mathcal{E}^{\mathsf{w}}$ as purely random-cluster events, intending to capture the typical behavior in the random-cluster marginals $\bar{\pi}_n^{\mathsf{dob}}$ and $\pi_n^{\mathsf{w}}$ when $\sigma_{\mathsf{o}}$ is in $\mathcal{V}_{\mathsf{red}}^c$ under $\mathbf{P}_n^{\mathsf{dob}}$ and $\mathbf{P}_n^{\mathsf{red}}$ respectively. First, we define the following cones:

$$\mathbf{C}_{\vee} := ([-\tfrac{1}{2}, \tfrac{1}{2}]^2 \times [0, L_\beta]) \cup \{(z_1, z_2, z_3) : z_3^2 \geq \sqrt{z_1^2 + z_2^2} \text{ for } z_3 \geq L_\beta\}\,,$$

and

$$\mathbf{C}_\Upsilon := \mathcal{L}_0 \cup \{(z_1, z_2, z_3) : z_3 \leq (\log \sqrt{z_1^2 + z_2^2})^2 \text{ for } \sqrt{z_1^2 + z_2^2} > L_\beta\},$$

where $L_\beta$ is chosen such that $L_\beta \to \infty$ as $\beta \to \infty$. Let $\mathcal{C}$ be the outermost 1-connected set of dual-to-closed plaquettes which separates $\mathsf{o}$ from $\partial_{\mathsf{red}}\Lambda'_{n,n}$, i.e., there is no path from $\mathsf{o}$ to $\partial_{\mathsf{red}}\Lambda'_{n,n}$ which does not pass through $\mathcal{C}$. Let $C_1$ be the four vertical plaquettes surrounding $\mathsf{o}$.

- Let $\mathcal{E}_1$ be the event that $C_1 \subset \mathcal{C}$, and $\mathcal{C} \setminus C_1$ consists of two 1-connected components.
- Let $\mathcal{E}_2$ be the event that the 1-connected component of $\mathcal{C} \setminus C_1$ that is above $C_1$ lies inside the cone $\mathbf{C}_\vee$. Call this portion of the component $\mathcal{C}^*$.
- Let $\mathcal{E}_3^{\mathsf{dob}}$ be the event that the rest of $\mathcal{C}$ lies in $\mathbf{C}_\Upsilon$ and $\mathsf{o}$ is $\omega$-connected to $\partial_{\mathsf{blue}}\Lambda'_{n,n}$. Let $\mathcal{E}_3^{\mathsf{w}}$ be that the rest of $\mathcal{C}$ is just the single horizontal plaquette below $\mathsf{o}$.
- Let $\mathcal{E}_4^{\mathsf{dob}}$ be the event that there is no $\omega$-path from $\partial_{\mathsf{blue}}\Lambda'_{n,n}$ to $\partial_{\mathsf{red}}\Lambda'_{n,n}$ (equivalently on $\mathcal{E}_3^{\mathsf{dob}}$, no $\omega$-path from $\mathsf{o}$ to $\partial_{\mathsf{red}}\Lambda'_{n,n}$). Let $\mathcal{E}_4^{\mathsf{w}}$ be that there is no $\omega$-path from $\mathsf{o}$ to $\partial\Lambda'_{n,n}$.

Define

$$\mathcal{E}^{\mathsf{dob}} = \mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3^{\mathsf{dob}} \cap \mathcal{E}_4^{\mathsf{dob}} \qquad \text{and} \qquad \mathcal{E}^{\mathsf{w}} = \mathcal{E}_1 \cap \mathcal{E}_2 \cap \mathcal{E}_3^{\mathsf{w}} \cap \mathcal{E}_4^{\mathsf{w}}.$$

Note that under $\bar{\pi}_n^{\mathsf{dob}}$ (and on $\mathcal{E}^{\mathsf{dob}}$), $\mathcal{C}$ will be equivalent to the full interface $\mathcal{I}_{\mathsf{full}}$. We observe (see [12, Lemma 4.5] for the algebra) that

$$\sum_{f \in \mathbf{C}_\vee} \sum_{g \in \mathbf{C}_\Upsilon} e^{-\beta d(f,g)/K} \leq \varepsilon_\beta. \tag{4.1}$$

We begin by showing that $\mathcal{E}^{\mathsf{dob}}$ and $\mathcal{E}^{\mathsf{w}}$ are indeed typical events, in the sense that

$$\mathbf{P}_n^{\mathsf{red}}(\mathcal{E}^{\mathsf{w}} \mid A_{\mathsf{o},h}) \geq 1 - \varepsilon_\beta, \qquad \text{and} \qquad \mathbf{P}_n^{\mathsf{dob}}(\mathcal{E}^{\mathsf{dob}} \mid \mathrm{ht}(\mathcal{P}_{\mathsf{o}}^{\mathsf{nonR}}) \geq h) \geq 1 - \varepsilon_\beta. \tag{4.2}$$

from which it would follow that

$$\mathbf{P}_n^{\mathsf{red}}(A_{\mathsf{o},h}) \leq (1 + \varepsilon_\beta)\mathbf{P}_n^{\mathsf{red}}(\mathcal{E}^{\mathsf{w}}, A_{\mathsf{o},h}),$$

and

$$\mathbf{P}_n^{\mathsf{dob}}(\mathrm{ht}(\mathcal{P}_{\mathsf{o}}^{\mathsf{nonR}}) \geq h) \leq (1 + \varepsilon_\beta)\mathbf{P}_n^{\mathsf{dob}}(\mathcal{E}^{\mathsf{dob}}, \mathrm{ht}(\mathcal{P}_{\mathsf{o}}^{\mathsf{nonR}}) \geq h).$$

The second (more complicated) of the bounds in (4.2) is a consequence of the "isolated pillar" results of [5]. There, a "top pillar" $\mathcal{P}_{\mathsf{o}}^{\mathsf{top}}$ was defined analogous to Definition 3.3 but using $\mathcal{V}_{\mathsf{top}}$ instead of $\mathcal{V}_{\mathsf{red}}$ (we will not use this definition except to reference the results proved in that paper). If $E_h$ is the event that the top pillar reaches height $h$, then [5, Thm. 3.8] (in the case of the Ising model, [12, Thm. 4.2]) proves that

$$\bar{\pi}_n^{\mathsf{dob}}(\mathcal{E}^{\mathsf{dob}} \mid E_h) \geq 1 - \varepsilon_\beta, \tag{4.3}$$

(the event $\mathsf{Iso}_{x,L_\beta,h}$ defined there implies the event $\mathcal{E}^{\mathsf{dob}}$ above). On the other hand, by [5, Lemmas 5.16, 5.17] (with $\mathcal{A}$ there taken to be $\mathrm{ht}(\mathcal{P}_{\mathsf{o}}^{\mathsf{nonR}}) \geq h$ and $\Omega$ taken to be $\mathcal{E}^{\mathsf{dob}}$), we have

$$\left| \frac{\mathbf{P}_n^{\mathsf{dob}}(\mathrm{ht}(\mathcal{P}_{\mathsf{o}}^{\mathsf{nonR}}) \geq h \mid \mathcal{E}^{\mathsf{dob}})}{\mathbf{P}_n^{\mathsf{dob}}(\mathrm{ht}(\mathcal{P}_{\mathsf{o}}^{\mathsf{nonR}}) \geq h \mid E_h)} - 1 \right| \leq \varepsilon_\beta. \tag{4.4}$$

Combining Eqs. (4.3) and (4.4) with the observation that $\mathrm{ht}(\mathcal{P}_{\mathsf{o}}^{\mathsf{nonR}}) \geq h \subset E_h$ concludes.

We now argue that Eq. (4.2), follows from a simplification of the same arguments of [5] that yielded Eqs. (4.3) and (4.4). The proof of Eq. (4.3) began by defining a cut point as a vertex $v \in \mathcal{P}_{\mathsf{o}}^{\mathsf{top}}$ with no other vertices in $\mathcal{P}_{\mathsf{o}}^{\mathsf{top}}$ having the same height, and such that removing the four vertical plaquettes surrounding $v$ cuts $\mathcal{C}$ into two 1-connected components, with the component above $v$ contained in $\mathcal{L}_{>\mathrm{ht}(v)}$. The portion of $\mathcal{C}$ above the lowest cut point was labeled the "spine" $\mathcal{S}_{\mathsf{o}}$. There was then a map which straightens the spine when it is too large, and deletes walls in the rest of $\mathcal{C}$ when they are too close to the spine. Doing so deletes the portion of $\mathcal{P}_{\mathsf{o}}^{\mathsf{top}}$ below the spine, so the map finishes by appending a column to rejoin the spine to the rest of $\mathcal{C}$. The proof

follows the same way, treating all of $\mathcal{C}$ as the top pillar. Namely, one can define a cut point and the spine of the component $\mathcal{C}$, and apply the same map to the spine portion. Since under $\pi_n^{\mathsf{w}}$ there is no need to preserve the disconnection event $\mathfrak{D}_{n,n}$ (recall Eq. (2.4)) as there was under $\bar{\pi}_n^{\mathsf{dob}}$, the remainder $\mathcal{C} \setminus \mathcal{S}_{\mathsf{o}}$ is handled by the simpler Peierls map which deletes it entirely and replaces it by a column of height $\mathrm{ht}(v)$ up from $\mathsf{o}$. Finally, Eq. (4.4) was shown by observing that when a portion of the spine is straightened, it becomes a column of vertices connected by open edges, and hence under the FK–Potts coupling they are all the same color. Clearly this observation holds regardless of whether the joint measure is $\mathbf{P}_n^{\mathsf{red}}$ or $\mathbf{P}_n^{\mathsf{dob}}$.

Hence, with Eq. (4.2) in hand, it is sufficient to show

$$\frac{\pi_n^{\mathsf{w}}(\mathcal{E}^{\mathsf{w}})}{\bar{\pi}_n^{\mathsf{dob}}(\mathcal{E}^{\mathsf{dob}})} = (1 \pm \varepsilon_\beta) q e^{-2\beta}, \qquad \text{and} \tag{4.5}$$

$$\frac{\mathbf{P}_n^{\mathsf{red}}(A_{\mathsf{o},h} \mid \mathcal{E}^{\mathsf{w}})}{\mathbf{P}_n^{\mathsf{dob}}(\mathrm{ht}(\mathcal{P}_{\mathsf{o}}^{\mathsf{nonR}}) \geq h \mid \mathcal{E}^{\mathsf{dob}})} = (1 \pm \varepsilon_\beta) \frac{q-1}{q}. \tag{4.6}$$

(where we use the shorthand notation $a = b(1 \pm \varepsilon_\beta)$ to mean $a \in [b(1 - \varepsilon_\beta), b(1 + \varepsilon_\beta)])$. We begin with (4.5). It is easy to see (by a Peierls argument) that $\pi_n^{\mathsf{w}}(\mathcal{E}^{\mathsf{w}})$ is dominated by (i.e., $1 - \varepsilon_\beta$ of its mass is on) the event that $\mathcal{C}$ is just the six plaquettes surrounding $\mathsf{o}$, which happens with probability $q e^{-6\beta}(1 \pm \varepsilon_\beta)$. Moreover, by rigidity of $\mathcal{I}_{\mathsf{full}}$ (Theorem 2.5)

$$\bar{\pi}_n^{\mathsf{dob}}(f_{(\mathsf{o}-\mathfrak{e}_3,\mathsf{o})} \text{ is a ceiling plaquette}) \geq 1 - \varepsilon_\beta,$$

and then with a cost of $(1 \pm \varepsilon_\beta) e^{-4\beta}$ we can add one open edge between $\mathsf{o} - \mathfrak{e}_3$ to $\mathsf{o}$ and close the five other incident edges to $\mathsf{o}$. This shows that $\bar{\pi}_n^{\mathsf{dob}}(\mathcal{E}^{\mathsf{dob}}) = (1 \pm \varepsilon_\beta) e^{-4\beta}$.

We turn to establishing (4.6). First note that on $\mathcal{E}^{\mathsf{dob}}$, the event $\mathrm{ht}(\mathcal{P}_{\mathsf{o}}^{\mathsf{nonR}}) \geq h$ is equivalent to $A_{\mathsf{o},h}$. Now, intending to utilize the FK–Potts coupling, we wish to first reveal in the random-cluster model the component $\mathcal{C}$. In order for the event $A_{\mathsf{o},h}$ to be possible, $\mathcal{C}$ must be a set of plaquettes that confines a chain of vertices connecting $\mathsf{o}$ to height $h$. Crucially, the events $\mathcal{E}^{\mathsf{w}}$ and $\mathcal{E}^{\mathsf{dob}}$ were defined so that the set of plaquettes $C^*$ such that $\pi_n^{\mathsf{w}}(\mathcal{C}^* = C^* \mid \mathcal{E}^{\mathsf{w}})$ and $\bar{\pi}_n^{\mathsf{dob}}(\mathcal{C}^* = C^* \mid \mathcal{E}^{\mathsf{dob}})$ both have positive probability is the same, let this set be denoted $\mathbf{C}^*$. In particular, $C^* \in \mathbf{C}^*$ if $C^*$ is a 1-connected component of plaquettes in $\mathbf{C}_\vee$ that forms part of a bounding surface in the sense that there can be no path in $\mathcal{L}_{>0}$ from $\mathsf{o}$ to $\partial \Lambda'_{n,n}$ which does not cross a plaquette of $C^* \cup C_1$. Thus, we can express

$$\mathbf{P}_n^{\mathsf{red}}(A_{\mathsf{o},h} \mid \mathcal{E}^{\mathsf{w}}) = \sum_{C^* \in \mathbf{C}^*} \pi_n^{\mathsf{w}}(\mathcal{C}^* = C^* \mid \mathcal{E}^{\mathsf{w}}) \mathbf{P}_n^{\mathsf{red}}(A_{\mathsf{o},h} \mid \mathcal{C}^* = C^*, \mathcal{E}^{\mathsf{w}}),$$

$$\mathbf{P}_n^{\mathsf{dob}}(A_{\mathsf{o},h} \mid \mathcal{E}^{\mathsf{dob}}) = \sum_{C^* \in \mathbf{C}^*} \bar{\pi}_n^{\mathsf{dob}}(\mathcal{C}^* = C^* \mid \mathcal{E}^{\mathsf{dob}}) \mathbf{P}_n^{\mathsf{dob}}(A_{\mathsf{o},h} \mid \mathcal{C}^* = C^*, \mathcal{E}^{\mathsf{dob}}).$$

Now, first studying the latter terms in the sums concerning the event $A_{\mathsf{o},h}$, the difference between conditioning on $\mathcal{E}^{\mathsf{w}}$ and $\mathcal{E}^{\mathsf{dob}}$ is that on $\mathcal{E}^{\mathsf{dob}}$ we know that the color $\sigma_{\mathsf{o}}$ is blue (because of $\mathcal{E}_3^{\mathsf{dob}}$), whereas $\mathcal{E}^{\mathsf{w}}$ does not provide any information about the color $\sigma_{\mathsf{o}}$, so the conditional probability that $o \in \mathcal{V}_{\mathsf{red}}^c$ (which on $\mathcal{E}^{\mathsf{w}}$ is equivalent to $\sigma_{\mathsf{o}} = \mathsf{nonR}$) is $\frac{q-1}{q}$. Given $\mathcal{C}^* = C^*$, its outer boundary is entirely in $\mathcal{V}_{\mathsf{red}}$ except $\mathsf{o}$ which is not in $\mathcal{V}_{\mathsf{red}}$. Therefore, membership in $\mathcal{V}_{\mathsf{red}}$ interior to $\mathcal{C}^*$ is fully measurable with respect to the coloring on finite components in $\mathcal{C}^*$, and so the event $A_{\mathsf{o},h}$ is measurable with respect to the same. Moreover, the coloring in $A_{\mathsf{o},h}$ is independent of the realization of the rest of $\mathcal{C}$. Hence, we have

$$\mathbf{P}_n^{\mathsf{red}}(A_{\mathsf{o},h} \mid \mathcal{C}^* = C^*, \mathcal{E}^{\mathsf{w}}) = \mathbf{P}_n^{\mathsf{dob}}(A_{\mathsf{o},h} \mid \mathcal{C}^* = C^*, \mathcal{E}^{\mathsf{dob}}) \frac{q-1}{q}.$$

Thus, it suffices to show that the random-cluster Radon–Nikodym derivative satisfies

$$\frac{\pi_n^{\mathsf{w}}(\mathcal{C}^* = C^* \mid \mathcal{E}^{\mathsf{w}})}{\bar{\pi}_n^{\mathsf{dob}}(\mathcal{C}^* = C^* \mid \mathcal{E}^{\mathsf{dob}})} = 1 \pm \varepsilon_\beta \qquad \text{for every } C^* \in \mathbf{C}^*. \tag{4.7}$$

Beginning with the numerator, the event $\mathcal{E}^{\mathsf{w}}$ is equivalent to asking that $\mathcal{C}^* \in \mathbf{C}^*$, and that $\mathcal{C} \setminus \mathcal{C}^*$ is equal to the five plaquettes to the sides and below $\mathsf{o}$. By maximality, this requires that the other plaquettes 1-connected to $f_{(\mathsf{o},\mathsf{o}-\mathfrak{e}_3)}$ are open. Let $F^{\mathsf{w}}$ be the union of these plaquettes with $\mathcal{C} \setminus \mathcal{C}^*$. We can then write $\pi_n^{\mathsf{w}}(\cdot \mid \mathcal{E}^{\mathsf{w}})$ as $\pi_{\Lambda'_{n,n} \setminus F^{\mathsf{w}}}^{\eta}(\cdot \mid \mathcal{C}^* \in \mathbf{C}^*)$, where $\eta$ is the boundary condition induced by the realization of $\mathcal{C} \setminus \mathcal{C}^*$ on $\mathcal{E}^{\mathsf{w}}$.

For the denominator, the event $\mathcal{E}^{\mathsf{dob}}$ is equivalent to asking that $\mathcal{C}^* \in \mathbf{C}^*$, and that the set of plaquettes $\mathcal{C} \setminus \mathcal{C}^*$ satisfies a set of criterion independent of $\mathcal{C}^*$. (Namely, that $\mathcal{C} \setminus \mathcal{C}^*$ contains $C_1$, is a subset of $\mathbf{C}_\Upsilon \cup C_1$, does not cut off $\mathsf{o}$ from $\partial_{\mathsf{blue}}\Lambda'_{n,n}$, and combines with $\mathcal{C}^*$ such that $\mathcal{C}$ is an interface satisfying $\mathfrak{D}_{n,n}$. This last event appears as though the choice of $\mathcal{C} \setminus \mathcal{C}^*$ affects the possible choices of $\mathcal{C}^*$, but the cut-point at $\mathsf{o}$ means it does not.) Exposing $\mathcal{C} \setminus \mathcal{C}^*$ reveals a set of dual-to-open plaquettes around it by maximality. Let $F^{\mathsf{dob}}$ be the union of these plaquettes with $\mathcal{C} \setminus \mathcal{C}^*$. Noting that the separation event $\mathfrak{D}_{n,n}$ is included in $\mathcal{E}^{\mathsf{dob}}$, we can write $\bar{\pi}_n^{\mathsf{dob}}(\cdot \mid \mathcal{E}^{\mathsf{dob}})$ as $\mathbb{E}_{\mathcal{C}\setminus\mathcal{C}^*}[\pi_{\Lambda'_{n,n} \setminus F^{\mathsf{dob}}}^{\gamma}(\cdot \mid \mathcal{C}^* \in \mathbf{C}^*) \mid \mathcal{E}^{\mathsf{dob}}]$ where $\gamma$ is the boundary condition induced by revealing $\mathcal{C} \setminus \mathcal{C}^*$.

It thus suffices to show that for any $C^* \in \mathbf{C}^*$, and any realization of $\gamma$ under $\mathbb{E}_{\mathcal{C}\setminus\mathcal{C}^*}$,

$$\frac{\pi_{\Lambda'_{n,n} \setminus F^{\mathsf{w}}}^{\eta}(\mathcal{C}^* = C^*)}{\pi_{\Lambda'_{n,n} \setminus F^{\mathsf{dob}}}^{\gamma}(\mathcal{C}^* = C^*)} = 1 \pm \varepsilon_\beta, \tag{4.8}$$

as summing over $\mathcal{C}^*$ then implies that $\pi_{\Lambda'_{n,n} \setminus F^{\mathsf{w}}}^{\eta}(\mathcal{C}^* \in \mathbf{C}^*) = \pi_{\Lambda'_{n,n} \setminus F^{\mathsf{dob}}}^{\gamma}(\mathcal{C}^* \in \mathbf{C}^*)(1 \pm \varepsilon_\beta)$, which in turn implies Eq. (4.8) even conditional on the event $\mathcal{C}^* \in \mathbf{C}^*$ up to a change of the $\varepsilon_\beta$. Then, averaging over $\gamma$ and applying the equivalence of measures discussed above implies Eq. (4.7).

Finally, to show Eq. (4.8), observe that each $C^* \in \mathbf{C}^*$ is a compatible set for $(\Lambda'_{n,n} \setminus F^{\mathsf{w}}, \eta)$ and $(\Lambda'_{n,n} \setminus F^{\mathsf{dob}}, \gamma)$, so we can apply Proposition 4.2 to obtain that

$$\left| \log \frac{\pi_{\Lambda'_{n,n} \setminus F^{\mathsf{w}}}^{\eta}(\mathcal{C}^* = C^*)}{\pi_{\Lambda'_{n,n} \setminus F^{\mathsf{dob}}}^{\gamma}(\mathcal{C}^* = C^*)} - \log \frac{Z^{\mathsf{w}}(\Lambda'_{n,n} \setminus F^{\mathsf{w}})}{Z^{\eta}(\Lambda'_{n,n} \setminus F^{\mathsf{w}})} \frac{Z^{\gamma}(\Lambda'_{n,n} \setminus F^{\mathsf{dob}})}{Z^{\mathsf{w}}(\Lambda'_{n,n} \setminus F^{\mathsf{dob}})} \right| \leq \exp\Big( \sum_{f \in C^*} \sum_{g \in F^{\mathsf{w}} \cup F^{\mathsf{dob}}} e^{-\beta d(f,g)/K} \Big).$$

In $\Lambda'_{n,n} \setminus F^{\mathsf{w}}$ with boundary condition $\eta$, there are two components of boundary vertices which are not wired together. The weight of a configuration $\omega$ in $Z^{\mathsf{w}}(\Lambda'_{n,n} \setminus F^{\mathsf{w}})$ is the same as in $Z^{\eta}(\Lambda'_{n,n} \setminus F^{\mathsf{w}})$ when these two components are wired together by $\omega$, and off by a factor of $q$ when they are not. However, the $\pi_{\Lambda'_{n,n} \setminus F^{\mathsf{w}}}^{\eta}$–probability that the two boundary components are not wired together is $\varepsilon_\beta$ (this is true whenever the two boundary components can be connected by a path of constant length, as in this case, as the probability that an edge is open is always $\geq \frac{p}{p+q(1-p)} = 1 - \varepsilon_\beta$ even conditional on a worst case configuration for the rest of the edges). This implies that the ratio of the two partition functions is at least $1 - \varepsilon_\beta$, and similarly, $Z^{\gamma}(\Lambda'_{n,n} \setminus F^{\mathsf{dob}})/Z^{\mathsf{w}}(\Lambda'_{n,n} \setminus F^{\mathsf{dob}}) \leq 1 + \varepsilon_\beta$. Finally, the double summation above is at most $\varepsilon_\beta$ per (4.1), concluding the proof. ∎

4.2. **Exact additivity of the rates.** We finally establish additivity (up to a $1 \pm \varepsilon_\beta$) of the sequence $\xi_h$; this allows us to replace $\xi_h$ by the limiting $\xi h$ without loss in our main theorem. We again truncate our consideration to high probability events on the interface's large deviation, and then establish the additivity on those events. The events will be of a similar cone-form to those in the previous subsection.

It will actually be better to work with $\widetilde{\xi}_h$ which we define exactly as in (1.2), but with an extra conditioning on $\{\sigma_{\mathsf{o}-\mathfrak{e}_3} = \mathsf{nonR}\}$ (or equivalently since it is only one site, by color symmetry, being $\mathsf{blue}$). This will turn out to be exactly additive. We first show the following relating $\xi_h$ to $\widetilde{\xi}_h$.

**Lemma 4.5.** *We have that for all $h \geq 1$,*

$$|\widetilde{\xi}_h - (\xi_h - 2\beta + \log(q-1))| \leq \varepsilon_\beta \qquad \text{and thus by Lemma } 4.4 \quad |\widetilde{\xi}_h - \alpha_h| \leq \varepsilon_\beta.$$

*Proof.* The cone event $\mathcal{E}^{\mathsf{w}}$ implies that in the Potts marginal $\mu_n^{\mathsf{red}}$, on the event $A_{\mathsf{o},h}$, the nonR component of $\mathsf{o}$ contains just a single vertex at height $1/2$. From there, it is not hard to see by a Peierls map that for any $h \geq 1$, under $\mu_n^{\mathsf{red}}(\cdot \mid A_{\mathsf{o},h})$, the probability that $\mathsf{o} - \mathfrak{e}_3$ is nonR is $(1 \pm \varepsilon_\beta)e^{-4\beta}$. A simple Peierls argument implies that unconditionally, the probability that $\mathsf{o} - \mathfrak{e}_3$ is nonR is $(q-1)(1 \pm \varepsilon_\beta)e^{-6\beta}$. We then have

$$\mu_n^{\mathsf{red}}(A_{\mathsf{o},h} \mid \sigma_{\mathsf{o}-\mathfrak{e}_3} = \mathsf{nonR}) = \frac{\mu_n^{\mathsf{red}}(\sigma_{\mathsf{o}-\mathfrak{e}_3} = \mathsf{nonR} \mid A_{\mathsf{o},h})}{\mu_n^{\mathsf{red}}(\sigma_{\mathsf{o}-\mathfrak{e}_3} = \mathsf{nonR})}\mu_n^{\mathsf{red}}(A_{\mathsf{o},h}) = \frac{1}{q-1}(1 \pm \varepsilon_\beta)e^{2\beta}\mu_n^{\mathsf{red}}(A_{\mathsf{o},h}),$$

which when rearranging and taking logarithms implies the claimed bound. ∎

**Lemma 4.6.** *We have for all $h_1, h_2$,*

$$|\widetilde{\xi}_{h_1+h_2} - (\widetilde{\xi}_{h_1} + \widetilde{\xi}_{h_2})| \leq \varepsilon_\beta.$$

*Proof.* Let $\widetilde{\mu}_n^{\mathsf{red}}$ be the distribution $\mu_n^{\mathsf{red}}$ conditioned on $\sigma_{\mathsf{o}-\mathfrak{e}_3}$ being nonR (or equivalently by symmetry, being blue). This implies that in the corresponding joint measure $\widetilde{\mathbf{P}}_n^{\mathsf{red}}$, there is no $\omega$-path from $\mathsf{o} - \mathfrak{e}_3$ to $\partial\Lambda'_{n,n}$, or equivalently there is a confining set of dual-to-closed plaquettes which surrounds $\mathsf{o} - \mathfrak{e}_3$. Let $\mathcal{C}$ be the 1-connected component of dual-to-closed plaquettes which contains the innermost such confining set. Call a vertex $y$ at a height $h$ a cut-point if it is the unique vertex at height $h$ which is separated from $\partial\Lambda'_{n,n}$ by $\mathcal{C}$, and the only plaquettes of $\mathcal{C}$ which have height $h$ are the four vertical plaquettes bounding the sides of $y$. Call $h$ a cut-height if there is a cut-point at height $h$.

We define certain high probability cone events in the random-cluster model, which we will show all have probability $1 - \varepsilon_\beta$ conditional on $A_{\mathsf{o},h_1+h_2}, A_{\mathsf{o},h_1}, A_{\mathsf{o},h_2}$ respectively in $\widetilde{\mathbf{P}}_n^{\mathsf{red}}$. We then show the approximate additivity of the rates under those events. The events are defined as follows:

(1) $\mathcal{E}_{\mathsf{Y}}$: is the event that heights $h_1 - \frac{1}{2}, \ldots, h_1 - \frac{1}{2} + (L_\beta \wedge h_2)$ are cut-heights, and if $y = (Y_1, Y_2, h_1 - \frac{1}{2})$ is the unique vertex at height $h_1 - \frac{1}{2}$ confined by $\mathcal{C}$, then $\mathcal{C} \cap \mathcal{L}_{\geq h_1}$ lies inside the cone

$$\{(z_1, z_2, z_3) : (z_3 - h_1)^2 \geq \sqrt{(z_1 - Y_1)^2 + (z_2 - Y_2)^2}\}.$$

(2) $\mathcal{E}_{\mathsf{V}}$: heights $\frac{1}{2}, \ldots, (L_\beta \wedge h_1) - \frac{1}{2}$ are cut-heights, and $\mathcal{C} \cap \mathcal{L}_{\geq 0}$ lies inside the cone

$$\{(z_1, z_2, z_3) : z_3^2 \geq \sqrt{z_1^2 + z_2^2}\}.$$

(3) $\mathcal{E}_{\|}$: $h_1 - \frac{1}{2}$ is a cut-height.

In the same way that we described Eq. (4.2) following from the "isolated pillar" results of [5] now that we have the generalized cluster expansion for $\mathcal{C}$, we also have the following bounds:

$$\widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\mathsf{Y}} \mid A_{\mathsf{o},h_1+h_2}) \geq 1 - \varepsilon_\beta \qquad \widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\mathsf{V}} \mid A_{\mathsf{o},h_2}) \geq 1 - \varepsilon_\beta \qquad \widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\|} \mid A_{\mathsf{o},h_1}) \geq 1 - \varepsilon_\beta. \quad (4.9)$$

It thus suffices to show

$$\frac{\widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\mathsf{Y}}, A_{\mathsf{o},h_1+h_2})}{\widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\mathsf{V}}, A_{\mathsf{o},h_2})\widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\|}, A_{\mathsf{o},h_1})} = \frac{\widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\mathsf{Y}}, A_{\mathsf{o},h_1+h_2} \mid \mathcal{E}_{\|}, A_{\mathsf{o},h_1})}{\widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\mathsf{V}}, A_{\mathsf{o},h_2})} = 1 \pm \varepsilon_\beta. \quad (4.10)$$

The first equality here is simply the fact that $\mathcal{E}_{\mathsf{Y}} \cap A_{\mathsf{o},h_1+h_2} \subset \mathcal{E}_{\|} \cap A_{\mathsf{o},h_1}$. Now we study the conditional joint measure $\widetilde{\mathbf{P}}_n^{\mathsf{red}}(\cdot \mid \mathcal{E}_{\|}, A_{\mathsf{o},h_1})$. We can expose the component of $\mathcal{C}$ which is below the cut-height at $h_1 - \frac{1}{2}$, which reveals additionally a set of dual-to-open plaquettes around it by maximality of the component. This determines the set of vertices $V$ which are in a finite component of $\mathcal{C}$ and in $\mathcal{L}_{<h_1}$. We furthermore reveal the colors of $V$. Let $\eta$ be the FK–Potts boundary condition

described by the above revealing, together with the wired red boundary at $\partial \Lambda'_{n,n}$. Let $E$ be the set of plaquettes which are fixed by $\eta$. Then, we can write

$$\widetilde{\mathbf{P}}_n^{\mathsf{red}}(\cdot \mid \mathcal{E}_{\parallel}, A_{\mathsf{o},h_1}) = \mathbb{E}_\eta[\mathbf{P}_{\Lambda'_{n,n} \setminus E}^\eta(\cdot) \mid \mathcal{E}_{\parallel}, A_{\mathsf{o},h_1}, \sigma_{0,0,-\frac{1}{2}} = \mathsf{nonR}], \tag{4.11}$$

as the events $\mathcal{E}_{\parallel}$, $A_{\mathsf{o},h_1}$, and $\sigma_{\mathsf{o}-\mathfrak{e}_3} = \mathsf{nonR}$ are implied by the closed edges and colors given by $\eta$.

Now notice that given such a realization $\eta$ satisfying $\mathcal{E}_{\parallel} \cap A_{\mathsf{o},h_1}$, the event $\mathcal{E}_{\mathsf{Y}} \cap A_{\mathsf{o},h_1+h_2}$ is exactly equivalent to the translation by $(Y_1, Y_2, h_1)$ (measurable with respect to $\Gamma$) of the pair of events $\mathcal{E}_{\vee} \cap A_{\mathsf{o},h_2}$. By this translation, and denoting $\eta', E'$ as the shifted versions (by $(-Y_1, -Y_2, -h_1)$) of $\eta, E$, we have

$$\mathbf{P}_{\Lambda'_{n,n} \setminus E}^\eta(\mathcal{E}_{\mathsf{Y}}, A_{0,h_1+h_2}) = \mathbf{P}_{\Lambda'_{n,n} \setminus E'}^{\eta'}(\mathcal{E}_{\vee}, A_{\mathsf{o},h_2}).$$

For an analogous setup in the denominator of Eq. (4.10), we want to expose the component of $\mathcal{C}$ which is below height 0. Let $\mathcal{E}_{\sqcup}$ denote the event that $\mathcal{C} \cap \mathcal{L}_{<0}$ consists only of the five plaquettes which are to the sides of and below $\mathsf{o} - \mathfrak{e}_3$. A standard Peierls map argument in the FK–Potts model shows that we have

$$\widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\sqcup}) \geq 1 - \varepsilon_\beta \qquad \text{and} \qquad \widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\sqcup} \mid \mathcal{E}_{\vee}, A_{\mathsf{o},h_2}) \geq 1 - \varepsilon_\beta.$$

(Indeed, for the first bound, the map would be to delete the outermost 1-connected component of dual-to-closed plaquettes, and if this causes $\mathsf{o} - \mathfrak{e}_3$ to be in the same connected component as $\partial \Lambda'_{n,n}$, then force its six incident edges to all be closed, so that it can be assigned the requisite $\mathsf{nonR}$ color. For the second one, the map would be to delete the 1-connected component in $\mathcal{L}_{<0}$ of dual-to-closed plaquettes incident to $\mathsf{o} - \mathfrak{e}_3$, and replace it by the minimal five dual-to-closed plaquettes incident to $\mathsf{o} - \mathfrak{e}_3$ in $\mathcal{L}_{<0}$.) This implies that

$$\widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\vee}, A_{\mathsf{o},h_2} \mid \mathcal{E}_{\sqcup}) = \widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\vee}, A_{\mathsf{o},h_2}) \frac{\widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\sqcup} \mid \mathcal{E}_{\vee}, A_{\mathsf{o},h_2})}{\widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\sqcup})} = \widetilde{\mathbf{P}}_n^{\mathsf{red}}(\mathcal{E}_{\vee}, A_{\mathsf{o},h_2})(1 \pm \varepsilon_\beta).$$

Note that conditioning on $\mathcal{E}_{\sqcup}$ is equivalent to conditioning on the aforementioned five plaquettes to be dual-to-closed, as well as on the plaquettes that are 1-connected to them and have height $< 0$ to be dual-to-open. Let $\zeta$ be the FK–Potts boundary condition with these dual-to-open and dual-to-closed plaquettes, the wired $\mathsf{red}$ boundary at $\partial \Lambda'_{n,n}$, and $\sigma_{\mathsf{o}-\mathfrak{e}_3} = \mathsf{blue}$ (by symmetry, any $\mathsf{nonR}$ color is equivalent). Let $F$ be the set of plaquettes fixed by $\zeta$. Then we can write

$$\widetilde{\mathbf{P}}_n^{\mathsf{red}}(\cdot \mid \mathcal{E}_{\sqcup}) = \mathbf{P}_{\Lambda'_{n,n} \setminus F}^\zeta(\cdot).$$

Now, we observe that the random-cluster marginal of $\mathbf{P}_{\Lambda'_{n,n} \setminus E'}^{\eta'}$ has the same law as the random-cluster measure on $\Lambda'_{n,n} \setminus E'$ with all the dual-to-open and dual-to-closed plaquettes in $\eta'$ as a boundary condition, further conditional on the event that there is no $\omega$-path from $\mathsf{o} - \mathfrak{e}_3$ to $\partial \Lambda'_{n,n}$. Call this random-cluster measure $\pi_{\Lambda'_{n,n} \setminus E'}^{\eta'(\mathsf{RC})}(\cdot \mid \mathfrak{D}_{\mathsf{o}})$, where $\mathfrak{D}_{\mathsf{o}}$ denotes this separation event for $\mathsf{o} - \mathfrak{e}_3$. Furthermore, by the FK–Potts coupling, the joint configuration under $\mathbf{P}_{\Lambda'_{n,n} \setminus E'}^{\eta'}$ can be sampled by first sampling a random-cluster configuration $\omega$ under $\pi_{\Lambda'_{n,n} \setminus E'}^{\eta'(\mathsf{RC})}(\cdot \mid \mathfrak{D}_{\mathsf{o}})$, assigning vertices in the same $\omega$-component of $\mathsf{o} - \mathfrak{e}_3$ the same $\mathsf{nonR}$ color as given by $\eta'$ (by symmetry, we may assume this is $\mathsf{blue}$), assigning vertices in the $\omega$-component of $\partial \Lambda'_{n,n}$ the color $\mathsf{red}$, and assigning all other $\omega$-components colors chosen independently at random. In the same way (with analogous notation), we can sample a joint configuration under $\mathbf{P}_{\Lambda'_{n,n} \setminus F}^\zeta$ by first sampling a random-cluster configuration $\omega$ under $\pi_{\Lambda'_{n,n} \setminus F}^{\zeta(\mathsf{RC})}(\cdot \mid \mathfrak{D}_{\mathsf{o}})$, assigning vertices in the same $\omega$-component of $\mathsf{o} - \mathfrak{e}_3$ the color $\mathsf{blue}$, assigning vertices in the $\omega$-component of $\partial \Lambda'_{n,n}$ the color $\mathsf{red}$, and assigning all other $\omega$-components colors chosen independently at random.

Now, let $\mathcal{C}^* := \mathcal{C} \cap \mathcal{L}_{\geq 0}$. Let $\mathbf{C}^*$ be the set of possible realizations of $\mathcal{C}^*$ satisfying $\mathcal{E}_\vee$ and $\mathfrak{D}_\mathsf{o}$ when combined with the boundary conditions $\eta'(\mathsf{RC})$ or $\zeta(\mathsf{RC})$ (note that because $\mathsf{o} - \mathfrak{e}_3$ is a cut-point for $\mathcal{C}$ under both boundary conditions, the set $\mathbf{C}^*$ will be the same). We can now write

$$\mathbf{P}^{\eta'}_{\Lambda'_{n,n} \setminus E'}(\mathcal{E}_\vee, A_{\mathsf{o},h_2}) = \sum_{C^* \in \mathbf{C}^*} \pi^{\eta'(\mathsf{RC})}_{\Lambda'_{n,n} \setminus E'}(\mathcal{C}^* = C^* \mid \mathfrak{D}_\mathsf{o})\mathbf{P}^{\eta'}_{\Lambda'_{n,n} \setminus E'}(A_{\mathsf{o},h_2} \mid \mathcal{C}^* = C^*),$$

$$\mathbf{P}^{\zeta}_{\Lambda'_{n,n} \setminus F}(\mathcal{E}_\vee, A_{\mathsf{o},h_2}) = \sum_{C^* \in \mathbf{C}^*} \pi^{\zeta(\mathsf{RC})}_{\Lambda'_{n,n} \setminus F}(\mathcal{C}^* = C^* \mid \mathfrak{D}_\mathsf{o})\mathbf{P}^{\zeta}_{\Lambda'_{n,n} \setminus F}(A_{\mathsf{o},h_2} \mid \mathcal{C}^* = C^*).$$

As $A_{\mathsf{o},h_2}$ now depends only on the coloring of the finite components enclosed by $\mathcal{C}^*$, which is independent of the rest of the FK–Potts model, we have

$$\mathbf{P}^{\eta'}_{\Lambda'_{n,n} \setminus E'}(A_{\mathsf{o},h_2} \mid \mathcal{C}^* = C^*) = \mathbf{P}^{\zeta}_{\Lambda'_{n,n} \setminus F}(A_{\mathsf{o},h_2} \mid \mathcal{C}^* = C^*).$$

Hence, in order to prove Eq. (4.10), it suffices to prove that

$$\frac{\pi^{\eta'(\mathsf{RC})}_{\Lambda'_{n,n} \setminus E'}(\mathcal{C}^* = C^* \mid \mathfrak{D}_\mathsf{o})}{\pi^{\zeta(\mathsf{RC})}_{\Lambda'_{n,n} \setminus F}(\mathcal{C}^* = C^* \mid \mathfrak{D}_\mathsf{o})} = (1 \pm \varepsilon_\beta), \tag{4.12}$$

at which point we can conclude via the sequence of equalities between Eq. (4.11) and Eq. (4.12), and averaging over $\eta$ in Eq. (4.11). Now, since each $\mathcal{C}^* \in \mathbf{C}^*$ is a compatible set for $(\Lambda'_{n,n} \setminus E', \eta')$ and $(\Lambda'_{n,n} \setminus F, \zeta)$, we can apply Proposition 4.2 with the same argument used to show Eq. (4.8), getting that

$$\frac{\pi^{\eta'(\mathsf{RC})}_{\Lambda'_{n,n} \setminus E'}(\mathcal{C}^* = C^*)}{\pi^{\zeta(\mathsf{RC})}_{\Lambda'_{n,n} \setminus F}(\mathcal{C}^* = C^*)} = (1 \pm \varepsilon_\beta).$$

Hence, it remains to show that

$$\frac{\pi^{\eta'(\mathsf{RC})}_{\Lambda'_{n,n} \setminus E'}(\mathfrak{D}_\mathsf{o})}{\pi^{\zeta(\mathsf{RC})}_{\Lambda'_{n,n} \setminus F}(\mathfrak{D}_\mathsf{o})} = (1 \pm \varepsilon_\beta).$$

However, a Peierls type argument shows that if $f$ is the plaquette above $\mathsf{o} - \mathfrak{e}_3$, then

$$\pi^{\eta'(\mathsf{RC})}_{\Lambda'_{n,n} \setminus E'}(\mathfrak{D}_\mathsf{o}) = \pi^{\eta'(\mathsf{RC})}_{\Lambda'_{n,n} \setminus E'}(f \text{ is closed})(1 \pm \varepsilon_\beta) \quad \text{and} \quad \pi^{\zeta(\mathsf{RC})}_{\Lambda'_{n,n} \setminus F}(\mathfrak{D}_\mathsf{o}) = \pi^{\zeta(\mathsf{RC})}_{\Lambda'_{n,n} \setminus F}(f \text{ is closed})(1 \pm \varepsilon_\beta).$$

Finally, the proof concludes by observing that the probability that $f$ is closed under both measures is the same, being equal to $(qe^{-\beta})/(1 - e^{-\beta} + qe^{-\beta})$.  ∎

**Proof of Theorem 4.1.** By Lemma 4.5, we have that $\lim_{h \to \infty} \xi_h/h = \lim_{h \to \infty} \tilde{\xi}_h/h = \xi$, and by Lemma 4.6 together with Fekete's lemma, the latter limit exists so it must be what we defined to be $\xi$ in (1.3). Moreover, by Lemma 4.6, $|\tilde{\xi}_h - \xi h| < \varepsilon_\beta$ uniformly over $h \geq 1$ (indeed, if $a_{h_1+h_2} \leq a_{h_1} + a_{h_2} + \varepsilon$ for all $h_1, h_2$ then $\tilde{a}_h := a_h + \varepsilon$ is sub-additive, thus by Fekete's Lemma, $a = \lim \tilde{a}_h/h = \inf \tilde{a}_h/h$, that is, $a_h \geq ah - \varepsilon$ for all $h$; similarly, $a_h \leq ah + \varepsilon$ for all $h$ whenever we have $a_{h_1+h_2} \geq a_{h_1} + a_{h_2} - \varepsilon$ for all $h_1, h_2$). The bound on $\alpha_h$ then follows from the second part of Lemma 4.5.  ∎

## Appendix A. Generalized cluster expansion

In this section, we provide the proof of Proposition 4.2. The following two lemmas were proved by Grimmett in [15] (with a minor adjustment to the latter, see the proof of Lemma A.2 below).

**Lemma A.1** ([15, Lem. 7.90]). *Let $\mathcal{U} = (V, E)$ be a connected finite graph and $\eta$ a boundary condition on $\Gamma$ where $\mathcal{U} \subset \Gamma$. Then,*

$$\log Z_{\mathcal{U}}^{\eta}(p, q) = \log q + \sum_{e \in E} \mathbf{f}_{\mathcal{U}, p, q}^{\eta}(e),$$

*where*

$$\mathbf{f}_{\mathcal{U}, p, q}^{\eta}(e) := \int_{p}^{1} \frac{s - \pi_{\mathcal{U}, s, q}^{\eta}(\omega_e = 1)}{s(1 - s)} \, ds.$$

*Moreover, we have $0 \le \mathbf{f}_{\mathcal{U}, p, q}^{\eta}(e) \le (1 - p)(q - 1)$ for any $\mathcal{U}, \eta$.*

**Lemma A.2** ([15, Lem. 7.93]). *Let $\mathcal{U}, \mathcal{U}' \subset \mathbb{Z}^3$ be two sub-graphs. Let $e \in \mathcal{U}$, $e' \in \mathcal{U}'$ be two edges. Suppose that there exists $r \ge 1$ such that $B_r(e) \cap \mathcal{U} \cong B_r(e') \cap \mathcal{U}'$. Then, there exists a constant $K > 0$ such that for $\beta$ sufficiently large,*

$$|\mathbf{f}_{\mathcal{U}, p, q}^{\mathsf{w}}(e) - \mathbf{f}_{\mathcal{U}', p, q}^{\mathsf{w}}(e')| \le e^{-\beta r / K}.$$

*Proof.* We note that Lemma A.2 was originally stated with a exponential bound of $e^{-r/K}$ for all $\beta \ge \beta_0$ instead of $e^{-\beta r / K}$ as above. We show here however that the stronger bound can be concluded from their proof. As was done, we can write

$$\frac{s - \pi_{\mathcal{U}, s, q}^{\eta}(\omega_e = 1)}{s(1 - s)} = \frac{(q - 1)(1 - \pi_{\mathcal{U}, s, q}^{\eta}(K_e))}{s + q(1 - s)}$$

where $K_e$ is the event that the endpoints of $e$ are joined by an open path without using the edge $e$. Now let $\rho_s = \frac{s}{s + q(1 - s)}$. It was then shown that for some $c_1, c_2 > 0$, we have

$$|\pi_{\mathcal{U}, s, q}^{\eta}(K_e) - \pi_{\mathcal{U}, s, q}^{\eta}(K_e)| \le c_1^r (1 - \rho_s)^{r / c_2}.$$

Note that $\rho_s$ is increasing for $s < 1$, and hence is minimized at $\rho_p$ on the interval $[p, 1)$. One can compute that $1 - \rho_p \ge \frac{q}{2} e^{-\beta}$. Hence, we can conclude that

$$|\mathbf{f}_{\mathcal{U}, s, q}^{\mathsf{w}}(e) - \mathbf{f}_{\mathcal{U}', s, q}^{\mathsf{w}}(e')| \le c_1^r (\tfrac{q}{2})^{r / c_2} e^{-\beta r / c_2} \int_{p}^{1} \frac{q - 1}{s + q(1 - s)} \, ds.$$

The integral above is a constant, so it and the term $c^r (\frac{q}{2})^{r/c}$ can be absorbed into the term $e^{-\beta r / c_2}$ by increasing the constant $c_2$ to some sufficiently large (but $\beta$ independent) constant $K$. ∎

**Proof of Proposition 4.2.** We follow the proof of [15, Lem. 7.104], with a few minor adjustments. Let $\mathcal{U} \setminus \overline{F}$ denote the graph resulting by removing from $\mathcal{U}$ the edge set $\{e : f_e \in \overline{F}\}$). Then, the fact that $F$ is a compatible set for $(\mathcal{U}, \eta)$ means that conditioning on $\mathsf{Cl}(F)$ induces wired boundary conditions on each component of $\mathcal{U} \setminus \overline{F}$. Thus, as in [15, Eq. 7.102], we can write

$$\pi_{\mathcal{U}}^{\eta}(\mathsf{Cl}(F)) = \frac{Z^{\mathsf{w}}(\mathcal{U})}{Z^{\eta}(\mathcal{U})} \frac{Z^{\mathsf{w}}(\mathcal{U} \setminus \overline{F})}{Z^{\mathsf{w}}(\mathcal{U})} (1 - e^{-\beta})^{|\partial F|} e^{-\beta |F|} q^{\kappa(F) - 1}. \tag{A.1}$$

Applying Lemma A.1, we get that

$$\log \left( \frac{Z^{\mathsf{w}}(\mathcal{U} \setminus \overline{F})}{Z^{\mathsf{w}}(\mathcal{U})} \right) = \sum_{e \in \mathcal{U} \setminus \overline{F}} (\mathbf{f}_{\mathcal{U} \setminus \overline{F}, p, q}^{\mathsf{w}}(e) - \mathbf{f}_{\mathcal{U}, p, q}^{\mathsf{w}}(e)) - \sum_{e : f_e \in \overline{F}} \mathbf{f}_{\mathcal{U}, p, q}^{\mathsf{w}}(e). \tag{A.2}$$

We can group the edges in the above sums according to which plaquette of $F$ they are the closest to. Let $\phi(e, F)$ be the plaquette $f \in F$ which is closest to the edge $e$, with distance $d(e, f)$ given by the $L^{\infty}$ distance between their centers. Ties can be decided in any arbitrary predetermined way, e.g., choosing the plaquette that is earliest in lexicographic ordering. Then, define

$$\mathbf{g}(f, F, \mathcal{U}) = \sum_{\substack{e \in \mathcal{U} \setminus \overline{F}, \\ \phi(e, F) = f}} (\mathbf{f}_{\mathcal{U} \setminus \overline{F}, p, q}^{\mathsf{w}}(e) - \mathbf{f}_{\mathcal{U}, p, q}^{\mathsf{w}}(e)) - \sum_{\substack{e : f_e \in \overline{F}, \\ \phi(e, F) = f}} \mathbf{f}_{\mathcal{U}, p, q}^{\mathsf{w}}(e). \tag{A.3}$$

We claim that $\mathbf{g}$ satisfies the requirements of the lemma. Combining Eqs. (A.1) to (A.3) immediately proves Item 3.

To show Item 1, let $\phi(e, F) = f$ and let $r = d(e, f)$. Then, $B_{r-2}(e)$ does not contain any plaquettes of $\overline{F}$, so by Lemma A.2 we get that for $r \geq 3$,

$$|\mathbf{f}^{\mathsf{w}}_{\mathcal{U} \setminus \overline{F}, p, q}(e) - \mathbf{f}^{\mathsf{w}}_{\mathcal{U}, p, q}(e)| \leq e^{-\beta(r-2)/K} . \tag{A.4}$$

For $r = 1, 2$, we can apply the bound $\mathbf{f}^{\mathsf{w}}_{\mathcal{U}, p, q}(e) \leq (q-1)e^{-\beta}$ from Lemma A.1. Thus, using the fact that the number of edges $e \in \mathcal{U} \setminus \overline{F}$ such that $\phi(e, F) = f$ and $d(e, f) = r$ is at most $c^r$ for some constant $c > 0$, we can sum over the exponential tails to bound the first sum in Eq. (A.3) by $e^{-\beta/K'}$ for some constant $K'$. The second sum in Eq. (A.3) has a constant number of terms, and can thus be bounded by $Ce^{-\beta}$ for some constant $C$, which can then be absorbed for large enough $\beta$ into the constant $K'$.

Finally, the proof of Item 2 follows in the same way as in [15, Lem. 7.104] with similar adjustments as above. Indeed, let $r$ be the largest radius such that $B_r(f) \cap \mathcal{U} \cong B_r(f') \cap \mathcal{U}'$ and $B_r(f) \cap F \cong B_r(f') \cap F'$. The original proof showed that $|\mathbf{g}(f, F, \mathcal{U}) - \mathbf{g}(f', F', \mathcal{U}')| \leq e^{-r/K}$ for $r > 9$. By applying the stronger, $\beta$-decaying bound in Lemma A.2 whenever needed, the same proof now shows that $|\mathbf{g}(f, F, \mathcal{U}) - \mathbf{g}(f', F', \mathcal{U}')| \leq e^{-\beta r/K}$. We can then apply Item 1 to extend this bound to $r \leq 9$ as well, at the cost of increasing the constant $K$. ∎

## APPENDIX B. PROOF OF THEOREM 2.5

Recall the definitions of walls and ceilings, and their excess areas from Section 3.3. Recall from [14] the notion of groups of walls (collections of nearby walls clustered by a threshold on their interactions through $\mathbf{g}$ relative to their excess areas) that are essential to being able to execute the deletion operations and establish rigidity in $\bar{\pi}^0_{\Lambda'_{n,\infty}}$.

The proof of Theorem 2.5 with $m = \infty$, found in [14], goes by considering the Peierls-type map that deletes a group of walls $\mathfrak{W}$ from the standard wall representation of $\mathcal{I}_{\mathsf{full}}$, and then bounds the entropy lost from the operation (controlled by enumerating the number of possible groups of walls of size $k$ nesting a vertex $x$ by $C^k$) by the energetic gain from the deletion, which is given by $\exp(-\beta \sum_{W \in \mathfrak{W}} \mathfrak{m}(W))$.

When the domain is capped at distance $m$ away from height zero, there is the potential concern of interactions with the hard floor (the extreme case of $m = 0$ of course being the same form as the interactions that cause many of the difficulties encountered in this present paper). However, when it is at distance $\Omega(n)$ away in both directions, the interaction is fairly crudely bounded by the energetic cost of reaching up to such a high height.

To begin, notice that for any configuration $\mathcal{I}_{\mathsf{full}}$ having $\overline{\mathrm{ht}}(\mathcal{I}_{\mathsf{full}}) = \max_x \overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{full}}) \leq m/4$, say, the operation of deleting a group of walls can induce a shift, at most, of $m/4$ on its internal walls, so after the operation, the interface will retain distance at least $m/2$ to the boundary conditions at heights $\pm m$. Then, per Remark 2.4, the only difference in the change of weights caused by this map, relative to the same map in $\Lambda'_{n,\infty}$ is an additive $Cn^2 e^{-n/2}$ from the radius $\mathbf{r}(f, I; f', I')$ possibly being attained from faces in the boundary. This is of course negligible compared to the cost of $\beta r$ for deletion of a group of wall of excess area $r$.

Thus, without any change to the argument that takes one from Proposition 2.3 to Theorem 2.5, we get for $m \geq 4k$, for every $x$,

$$\bar{\pi}^0_{\Lambda'_{n,m}}\left(\overline{\mathrm{ht}}_x(\mathcal{I}_{\mathsf{full}}) \geq k, \ \max_y \overline{\mathrm{ht}}_y(\mathcal{I}_{\mathsf{full}}) \leq m/4\right) \leq \exp(-a_p k),$$

It therefore suffices to upper bound the probability of $\max_y \overline{\mathrm{ht}}_y(\mathcal{I}_{\mathsf{full}}) \geq m/4$, for $m \geq n$, by $\exp(-a'_p m)$ as that will be smaller than $\exp(-a_p k)$ up to changing the constant $a_p$. For that purpose, consider any interface with $\max_y \overline{\mathrm{ht}}_y(\mathcal{I}_{\mathsf{full}}) \geq m/4$. There must be a nested sequence

of groups of walls $\overline{\mathfrak{W}}$ (meaning the union of the groups of walls of a nested sequence of walls $W_1, ..., W_N$) whose total excess area satisfies $\mathfrak{m}(\overline{\mathfrak{W}}) \geq m/4$.

Consider the map $\Phi$ that iteratively,

(1) Picks a nested sequence of groups of walls whose total excess area exceeds $m/4$ (if one exists) and deletes it from the standard wall representation;

(2) If no more exist, stops and outputs the resulting interface as $\Phi(\mathcal{I}_{\mathsf{full}})$.

Notice that the resulting interface must be a valid interface in $\Lambda'_{n,m}$ with $\overline{\mathrm{ht}}(\Phi(\mathcal{I}_{\mathsf{full}})) \leq m/4$. When considering the weight gain from this map, as compared to the infinite-cylinder analysis, the only difference is the possibility that the radius of congruence $\mathbf{r}(f, \mathcal{I}_{\mathsf{full}}; f', \Phi(\mathcal{I}_{\mathsf{full}}))$ is attained by interactions between a face $f$ interior of a wall that was deleted, and the boundary conditions at heights $\pm m$. Indeed if $\mathcal{I}_{\mathsf{full}}$ is any such interface, and $\mathfrak{W}$ is the set of walls deleted, then by [14]

$$\left| \log \frac{\bar{\pi}^0_{\Lambda'_{n,\infty}}(\mathcal{I}_{\mathsf{full}})}{\bar{\pi}^0_{\Lambda'_{n,\infty}}(\Phi(\mathcal{I}_{\mathsf{full}}))} + \beta \mathfrak{m}(\mathfrak{W}) \right| \leq C \mathfrak{m}(\mathfrak{W}), \tag{B.1}$$

while at the same time, by Proposition 2.3 and Remark 2.4,

$$\left| \log \frac{\bar{\pi}^0_{\Lambda'_{n,m}}(\mathcal{I}_{\mathsf{full}})}{\bar{\pi}^0_{\Lambda'_{n,m}}(\Phi(\mathcal{I}_{\mathsf{full}}))} - \log \frac{\bar{\pi}^0_{\Lambda'_{n,\infty}}(\mathcal{I}_{\mathsf{full}})}{\bar{\pi}^0_{\Lambda'_{n,\infty}}(\Phi(\mathcal{I}_{\mathsf{full}}))} \right| \leq K n^2 e^{-cm/4} + \sum_{f \in \mathsf{Ceil}(\mathfrak{W}) : |\mathrm{ht}(f)| \geq \frac{m}{2}} K e^{-cd(f, \mathcal{L}_{\pm m})}. \tag{B.2}$$

where $c, K$ are the constants from Proposition 2.3, and the sum runs over $f \in \mathcal{I}_{\mathsf{full}}$ in ceilings interior to walls of $\mathfrak{W}$. At this point, we wish to bound the number of summands in (B.2) by $\mathfrak{m}(\mathfrak{W})$. Notice that if we partition $A = \rho(\{f \in \mathsf{Ceil}(W) : |\mathrm{ht}(f)| \geq m/2\})$ where $\rho(\cdot)$ is used to denote projecting it down to faces of $\mathcal{L}_0$, for any $M$ and any partition of $A$ into $\mathfrak{C}_1, ..., \mathfrak{C}_M$, we have

$$|A| \leq \sum_{i=1}^{M} \mathrm{diam}(\mathfrak{C}_i)^2 \leq \sum_{i=1}^{M} \mathrm{diam}(\mathfrak{C}_i) \cdot n, \tag{B.3}$$

(because at worst, we are overcounting faces with this bound). On the other hand, for every height between $j = 1, ..., m/2$, there must have been a collection of wall faces in $\mathfrak{W}$ nesting these ceilings, and therefore at each height, there was some partition $\mathfrak{C}^j_1, ..., \mathfrak{C}^j_{M_j}$ such that those deleted wall faces at height $j$ contributed at least $\sum_{i=1}^{M_j} \mathrm{diam}(\mathfrak{C}^j_{M_j})$. By (B.3), for each $j$, this is at least $|A|/n$ and we sum over $j = 1, ..., m/2$ to get that $\mathfrak{m}(\mathfrak{W}) \geq |A| \frac{m}{2n}$. Since $m \geq n/2$, this is at least $|A|/4$, so the term in (B.2) the right hand side is bounded by $K n^2 e^{-cn/8} + K \mathfrak{m}(\mathfrak{W})$, which combined with (B.1) implies the same bound as in that equation for $\bar{\pi}^0_{\Lambda'_{n,m}}$ up to a change of $C$ to $C'$ for another $\beta$ independent constant.

The only other thing left to verify is that the entropy lost by deletion of these walls is controlled by $C^{\mathfrak{m}(\mathfrak{W})}$ for a universal constant $C$. To see this, consider the number of pre-images of a $\mathcal{J} = \Phi(\mathcal{I}_{\mathsf{full}})$, having $\mathfrak{m}(\mathcal{I}; \mathcal{J}) = L$. We first pick a number $I \in \{1, ..., L/m\}$ of iterations of step (1) of the map $\Phi$ before it terminated; then pick $I$ many vertices $v_1, ..., v_I$ in $[-\frac{n}{2}, \frac{n}{2}]^2 \times \{0\}$, allocate the total excess area $L$ between them with each being allocated at least $m/4$ into $L_1, ..., L_I$, and then enumerate over nested sequences of groups of walls containing a vertex $v_i$ by $C^{L_i}$ per [14, Proof of Lemma 15]. In total, this bounds the entropy loss by

$$\frac{L}{m} \max_{I \leq L/m} n^{2I} 2^L \prod_{i=1}^{I} C^{L_i} \leq C_1^L$$

where we used the fact that $m \geq n/2$. Since we established that the weight change from such maps with $\mathfrak{m}(\mathcal{I}; \mathcal{J}) = L$ is $\exp(-(\beta - C)L)$, summing over $L \geq m/4$ is at most $\exp(-(\beta - C')n/4)$ concluding the proof. ∎

## References

[1] P. Caddeo, Y. H. Kim, and E. Lubetzky. On level line fluctuations of SOS surfaces above a wall. *Forum Math. Sigma*. To appear.

[2] 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.

[3] 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.

[4] L. Chayes. Percolation and ferromagnetism on $\mathbf{Z}^2$: the $q$-state Potts cases. *Stochastic Process. Appl.*, 65(2):209–216, 1996.

[5] J. Chen and E. Lubetzky. Extrema of 3D Potts interfaces. *Communications in Mathematical Physics*, 405(3):86, 2024.

[6] R. L. Dobrušin. An investigation of Gibbs states for three-dimensional lattice systems. *Teor. Verojatnost. i Primenen.*, 18:261–279, 1973.

[7] J. Fröhlich and C.-E. Pfister. Semi-infinite Ising model. I. Thermodynamic functions and phase diagram in absence of magnetic field. *Comm. Math. Phys.*, 109(3):493–523, 1987.

[8] J. Fröhlich and C.-E. Pfister. Semi-infinite Ising model. II. The wetting and layering transitions. *Comm. Math. Phys.*, 112(1):51–74, 1987.

[9] R. Gheissari and E. Lubetzky. Mixing times of critical two-dimensional Potts models. *Comm. Pure Appl. Math.*, 71(5):994–1046, 2018.

[10] R. Gheissari and E. Lubetzky. Tightness and tails of the maximum in 3D Ising interfaces. *Ann. Probab.*, 49(2):732–792, 2021.

[11] R. Gheissari and E. Lubetzky. Maximum and shape of interfaces in 3D Ising crystals. *Communications on Pure and Applied Mathematics*, 75(12):2575–2684, 2022.

[12] R. Gheissari and E. Lubetzky. Approximate Domain Markov property for rigid Ising interfaces. *J. Stat. Phys.*, 190(5):Paper No. 99, 2023.

[13] R. Gheissari and E. Lubetzky. Entropic repulsion of 3D Ising interfaces conditioned to stay above a floor. *Electronic Journal of Probability*, 28(none):1–44, 2023.

[14] G. Gielis and G. Grimmett. Rigidity of the interface in percolation and random-cluster models. *J. Statist. Phys.*, 109(1-2):1–37, 2002.

[15] G. Grimmett. *The random-cluster model*, volume 333 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.

[16] O. Häggström. Positive correlations in the fuzzy Potts model. *The Annals of Applied Probability*, 9(4):1149 – 1159, 1999.

[17] P. Holický and M. Zahradník. On entropic repulsion in low temperature Ising models. In N. Boccara, E. Goles, S. Martinez, and P. Picco, editors, *Cellular Automata and Cooperative Systems*, pages 275–287. Springer Netherlands, Dordrecht, 1993.

[18] D. Ioffe, S. Ott, Y. Velenik, and V. Wachtel. Invariance principle for a Potts interface along a wall. *Journal of Statistical Physics*, 180(1):832–861, 2020.

[19] D. Ioffe, S. Shlosman, and F. L. Toninelli. Interaction versus entropic repulsion for low temperature Ising polymers. *Journal of Statistical Physics*, 158(5):1007–1050, 2015.

[20] D. Ioffe and Y. Velenik. Low-temperature interfaces: prewetting, layering, faceting and Ferrari-Spohn diffusions. *Markov Process. Related Fields*, 24(3):487–537, 2018.

[21] E. Lubetzky, F. Martinelli, and A. Sly. Harmonic pinnacles in the discrete Gaussian model. *Communications in Mathematical Physics*, 344(3):673–717, 2016.

[22] E. Lubetzky and A. Sly. Critical Ising on the square lattice mixes in polynomial time. *Comm. Math. Phys.*, 313(3):815–836, 2012.

[23] C. Maes and K. V. Velde. The fuzzy potts model. *Journal of Physics A: Mathematical and General*, 28(15):4261, aug 1995.

[24] I. Meličerčík. Entropic repulsion of two interfaces in Ising model, 1991.

J. Chen

Courant Institute, New York University, 251 Mercer Street, New York, NY 10012, USA.
  *Email address*: jlc871@courant.nyu.edu

R. Gheissari

Department of Mathematics, Northwestern University
  *Email address*: gheissari@northwestern.edu

E. Lubetzky

Courant Institute, New York University, 251 Mercer Street, New York, NY 10012, USA.
  *Email address*: eyal@courant.nyu.edu