

# MAXIMUM AND SHAPE OF INTERFACES IN 3D ISING CRYSTALS

REZA GHEISSARI AND EYAL LUBETZKY

ABSTRACT. Dobrushin (1972) showed that the interface of a 3D Ising model with minus boundary conditions above the  $xy$ -plane and plus below is rigid (has  $O(1)$ -fluctuations) at every sufficiently low temperature. Since then, basic features of this interface—such as the asymptotics of its maximum—were only identified in more tractable random surface models that approximate the Ising interface at low temperatures, e.g., for the (2+1)D Solid-On-Solid model. Here we study the large deviations of the interface of the 3D Ising model in a cube of side-length  $n$  with Dobrushin’s boundary conditions, and in particular obtain a law of large numbers for  $M_n$ , its maximum: if the inverse-temperature  $\beta$  is large enough, then  $M_n/\log n \rightarrow 2/\alpha_\beta$  as  $n \rightarrow \infty$ , in probability, where  $\alpha_\beta$  is given by a large deviation rate in infinite volume.

We further show that, on the large deviation event that the interface connects the origin to height  $h$ , it consists of a 1D spine that behaves like a random walk, in that it decomposes into a linear (in  $h$ ) number of asymptotically-stationary weakly-dependent increments that have exponential tails. As the number  $T$  of increments diverges, properties of the interface such as its surface area, volume, and the location of its tip, all obey CLTs with variances linear in  $T$ . These results generalize to every dimension  $d \geq 3$ .

## 1. INTRODUCTION

We study the plus-minus Ising interface in  $d$ -dimensions at sufficiently low temperatures, where for  $d \geq 3$  the interface is known to be rigid and yet its large deviations, including the asymptotic behavior of its maximum, were unknown. The Ising model on a finite subgraph  $\Lambda \subset \mathbb{Z}^d$  is an assignment of  $\pm 1$  to the  $d$ -dimensional *cells* of  $\mathbb{Z}^d$  (faces when  $d = 2$  and cubes of side-length 1 when  $d = 3$ ), collected in the set  $\mathcal{C}(\Lambda)$ . These cells are identified with their midpoints, corresponding to the vertices of the dual graph  $(\mathbb{Z} + \frac{1}{2})^d$ , and  $u, v \in \mathcal{C}(\Lambda)$  are considered adjacent (denoted  $u \sim v$ ) if their midpoints are at Euclidean distance 1. The Ising model on  $\Lambda$  is then the Gibbs distribution  $\mu_\Lambda = \mu_{\Lambda, \beta}$  over configurations in  $\Omega = \{\pm 1\}^{\mathcal{C}(\Lambda)}$  given by

$$\mu_\Lambda(\sigma) \propto \exp[-\beta \mathcal{H}(\sigma)], \quad \text{for } \mathcal{H}(\sigma) = \sum_{u \sim v} \mathbf{1}\{\sigma_u \neq \sigma_v\},$$

where  $\beta > 0$  is the inverse temperature. Placing boundary condition  $\eta$  on the model,  $\mu_\Lambda^\eta$ , refers to the conditional distribution of  $\mu_H$ , for some larger given graph  $H \supset \Lambda$ , where the configuration of  $\mathcal{C}(H) \setminus \mathcal{C}(\Lambda)$  coincides with  $\eta$ . These definitions extend to infinite graphs via weak limits, and in the low temperature regime studied here, different boundary conditions  $\eta$  on boxes in  $\mathbb{Z}^d$  lead to distinct limiting Gibbs distributions [32, §6.2].

Here, we consider  $\beta > \beta_0$  for some fixed  $\beta_0$  and  $\Lambda = \Lambda_n$ , the infinite cylinder of side-length  $2n$  in  $\mathbb{Z}^d$ ,

$$\Lambda_n = \llbracket -n, n \rrbracket^{d-1} \times \llbracket -\infty, \infty \rrbracket = \{-n, \dots, n\}^{d-1} \times \{-\infty, \dots, \infty\},$$

with boundary conditions that are (+) in the lower half-space  $\mathcal{C}(\mathbb{Z}^{d-1} \times \{-\infty, \dots, 0\})$  and (−) elsewhere, called *Dobrushin’s boundary conditions*. Let  $\mu_n = \mu_{\Lambda_n, \beta}^\mp$  denote the Ising model with these boundary conditions, and note that every  $\sigma \sim \mu_n$  defines a set of  $(d-1)$ -cells separating disagreeing spins, which in turn give rise to an *interface*  $\mathcal{I}$  separating the minus and plus phases: in 2D, it is a (maximal) connected component of such separating edges connecting  $(-n, 0)$  and  $(n, 0)$ ; in three dimensions, it is the (maximal) connected component of such separating faces containing  $\partial\Lambda_n \cap (\mathbb{Z}^{d-1} \times \{0\})$  (we defer more detailed definitions to §2.1).

The classical argument of Peierls, which established the phase transition in the Ising model for  $d \geq 2$ , shows that in the above described setting, the size of “bubbles” (finite connected components of plus or minus spins) has an exponential tail. One thus looks to determine the behavior of the interface  $\mathcal{I}$ .

In the 2D Ising model, the properties of this random interface between plus/minus phases in  $\mu_n$  is very well-understood: for  $\beta > \beta_c$ , the critical point of the Ising model, this interface converges to a Brownian bridge as  $n \rightarrow \infty$ , and detailed quantitative estimates are available for its fluctuations and large deviations for large  $n$ , mimicking those of a random walk (see, e.g., [24, 25, 35–37, 42, 43]). In view of its height fluctuations that diverge with  $n$  (in this case, with variance  $C_\beta n$  in the bulk), the interface is referred to as *rough*.

For the 3D Ising model (and in fact extending to every dimension  $d \geq 3$ ), Dobrushin [28] famously showed that, for large enough  $\beta$ , the plus/minus interface  $\mathcal{I}$  is *rigid* (localized) around height 0: the height fluctuations are  $O(1)$  everywhere. Namely, Dobrushin established that the probability that the interface  $\mathcal{I}$  reaches height at least  $h$  above any given  $xy$ -coordinate in  $\llbracket -n, n \rrbracket^2$  is  $O(\exp(-\frac{1}{3}\beta h))$ . An important consequence of rigidity is that the Gibbs distribution  $\mu_{\mathbb{Z}^3}^\mp$  arising as the weak limit of  $\mu_n$  is not translation-invariant in its  $z$ -coordinate. It is believed that the interface becomes rigid only after a roughening threshold

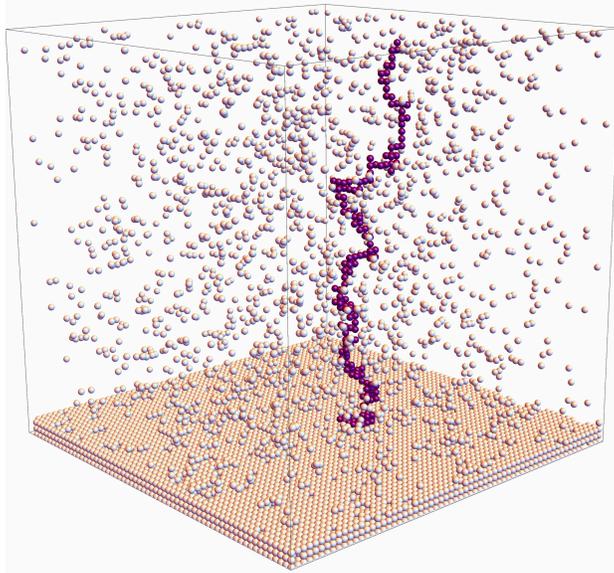


FIGURE 1. The plus/minus interface  $\mathcal{I}$  in the 3D Ising model  $\mu_n$  (side length  $n = 64$ ) with Dobrushin's boundary conditions, when conditioning on  $\mathcal{I}$  reaching height  $h = 64$ .

$\beta_r > \beta_c$ , with this roughening phase transition being exclusive to dimension 3. Interfaces of tilted Dobrushin boundary conditions are, unlike the flat ones, believed to always be rough; see §1.4 for more details.

Since Dobrushin's work showing that the interface  $\mathcal{I}$  is typically a flat surface at height 0, basic features of this interface—such as the asymptotics of its maximum, the shape of the surface near the maximum and the effect of entropic repulsion—were only identified in more tractable random surface models that approximate the Ising interface at low temperatures, e.g., the (2+1)D Solid-On-Solid model by Bricmont, El-Mellouki and Fröhlich [9] and Caputo et al. [16, 17], and the Discrete Gaussian and  $|\nabla\phi|^p$ -models in [39] (in these, the surfaces are height functions, with no overhangs or interacting bubbles that do exist in the Ising model).

In what follows, for the sake of the exposition, we state our new results on the interface  $\mathcal{I}$  in the context of the 3D Ising model, noting that they extend to any dimension  $d \geq 3$  (see Remark 1.1).

**1.1. Maximum height.** Let  $M_n$  be the maximum height ( $z$ -coordinate) of a face in  $\mathcal{I}$ . Dobrushin's estimate that  $\mu_n(\mathcal{I} \ni (y_1, y_2, h)) = O(\exp(-\frac{1}{3}\beta h))$  shows, by a union bound, that  $M_n/\log n \leq C_\beta$  in probability as  $n \rightarrow \infty$  for some  $C_\beta > 0$ . As we later explain, a lower bound of matching order,  $M_n/\log n \geq c_\beta$  in probability for some other  $c_\beta > 0$ , can also be deduced from those methods via decorrelation estimates. Our main goals here are obtaining the asymptotics of  $M_n$  (law of large numbers (LLN) for the maximum) and characterizing the typical structure of the surface around points conditioned to achieve large deviations. The first result establishes the LLN and expresses the limit in terms of a large deviation (LD) rate function of having the origin be  $*$ -connected to height  $h$  via  $(+)$ -spins (denoted  $\xrightarrow{+}$ ) within  $\mathcal{C}(\mathbb{Z}^2 \times \llbracket 0, h \rrbracket)$  in the measure  $\mu_{\mathbb{Z}^3}^\mp$ .

**Theorem 1** (LLN for the maximum). *There exists  $\beta_0$  such that, for all  $\beta > \beta_0$ , the maximum  $M_n$  of the interface  $\mathcal{I}$  in the 3D Ising model with Dobrushin's boundary conditions  $\mu_{\Lambda_n, \beta}^\mp$  satisfies*

$$\lim_{n \rightarrow \infty} \frac{M_n}{\log n} = \frac{2}{\alpha_\beta} \quad \text{in probability,} \quad (1.1)$$

where the constant  $\alpha_\beta > 0$  is given by

$$\alpha_\beta = \lim_{h \rightarrow \infty} -\frac{1}{h} \log \mu_{\mathbb{Z}^3}^\mp \left( \left( \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right) \xrightarrow{+}_{\mathcal{C}(\mathbb{Z}^2 \times \llbracket 0, h \rrbracket)} \left( \mathbb{Z} + \frac{1}{2} \right)^2 \times \left\{ h - \frac{1}{2} \right\} \right), \quad (1.2)$$

and satisfies  $\alpha_\beta/\beta \rightarrow 4$  as  $\beta \rightarrow \infty$ .

Note that the existence of the limit in (1.2) is both nontrivial and essential, and its proof (see §6.2 and in particular Proposition 6.7) relies on our results on the structure of the interface  $\mathcal{I}$  conditioned on large deviations in  $\mu_n$ , which drive an approximate sub-additivity argument.

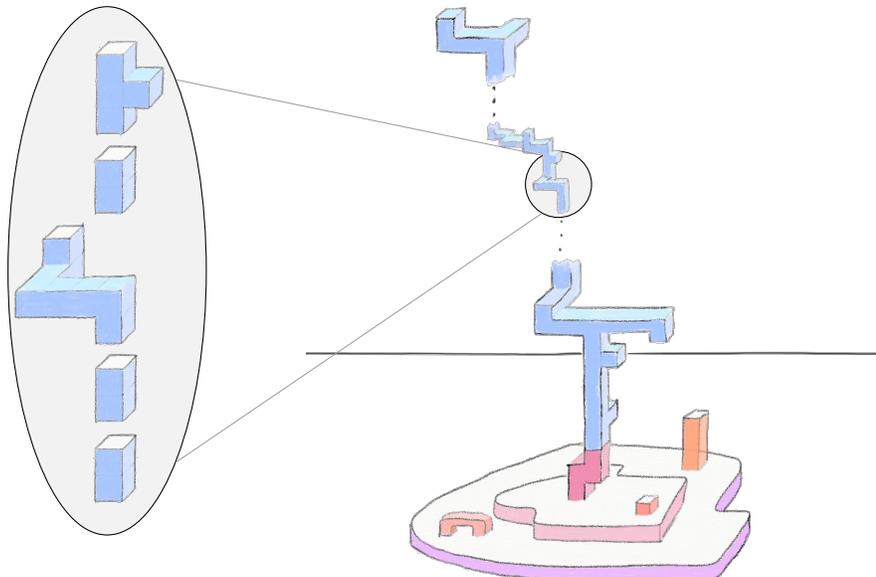


FIGURE 2. The pillar above a point  $x$ , denoted  $\mathcal{P}_x$ ; in blue, the *spine*, partitioned into *increments*.

**1.2. Structure of tall pillars.** To formalize the notion of  $\mathcal{I}$  achieving a large deviation above a point  $x$ , define the *pillar* associated to a point  $x \in \llbracket -n + \frac{1}{2}, n - \frac{1}{2} \rrbracket^2 \times \{0\}$  (we defer detailed definitions to §2.3): from a configuration  $\sigma \sim \mu_n$ , repeatedly delete every finite cluster of (+) or (-) by flipping its spins (thus eliminating all bubbles), then discard  $\mathcal{C}(\llbracket -n, n \rrbracket^2 \times \mathbb{Z}_-)$ ; the pillar of  $x$ , denoted  $\mathcal{P}_x$ , is the resulting (possibly empty) \*-connected component of (+) cells containing  $x + (0, 0, \frac{1}{2})$ , along with all faces of  $\mathcal{I}$  that bound it.

The height of the pillar  $\mathcal{P}_x$ , denoted  $\text{ht}(\mathcal{P}_x)$ , is the maximal  $y_3$  such that some  $(y_1, y_2, y_3) \in \mathcal{P}_x$ . The proof of (1.1) in Theorem 1 hinges on a large deviation estimate for  $\text{ht}(\mathcal{P}_x)$  stating (see Proposition 6.1) that

$$\lim_{h \rightarrow \infty} -\frac{1}{h} \log \mu_n(\text{ht}(\mathcal{P}_x) \geq h) = \alpha_\beta.$$

(Observe that the upper bound on  $M_n/\log n$  in (1.1) readily follows from this by a union bound over  $x$ .)

A key step in the analysis of the typical structure of  $\mathcal{P}_x$  conditioned on  $\{\text{ht}(\mathcal{P}_x) \geq h\}$  is to decompose the pillar into increments: define the *cut-points* of  $\mathcal{P}_x$  to be every  $y = (y_1, y_2, y_3) \in \mathcal{P}_x$  such that  $y$  is the unique cell in the horizontal slab with height  $y_3$  belonging to  $\mathcal{P}_x$ . Ordering the cut-points as  $v_1, \dots, v_{\mathcal{T}}$  with an increasing third coordinate, their role mimics regeneration points of random walks (though the increment sequence is far from Markovian); thus we refer to the subset of  $\mathcal{P}_x$  delimited by  $v_i, v_{i+1}$  (including these two cells) as a *pillar increment* (see Figure 2). Let  $\mathfrak{X}$  be the (countable infinite) set of possible increments, and let  $A(X)$  be the surface area (number of bounding dual-faces) of an increment  $X$ . Our next result is a central limit theorem (CLT) for averages of a function along the pillar increment sequence.

**Theorem 2** (CLT for the increments). *There exist  $\beta_0, \kappa_0 > 0$  so that the following holds for all  $\beta > \beta_0$ : for every sequence  $T = T_n$  with  $1 \ll T \ll n$ , every non-constant observable on increments  $f : \mathfrak{X} \rightarrow \mathbb{R}$  such that*

$$f(X) \leq e^{\kappa_0 A(X)} \quad \text{for every } X \in \mathfrak{X},$$

*and every  $x = (x_1, x_2, 0)$  with  $(x_1, x_2) \in \llbracket -n + \Delta_n + \frac{1}{2}, n - \Delta_n - \frac{1}{2} \rrbracket^2$  for some  $\Delta_n \gg T$ , if  $(\mathcal{X}_1, \dots, \mathcal{X}_{\mathcal{T}})$  is the random increment sequence of  $\mathcal{P}_x$ , then conditional on the event  $\{\mathcal{T} \geq T\}$ , one has that*

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T (f(\mathcal{X}_t) - \mathbb{E}[f(\mathcal{X}_t)]) \implies \mathcal{N}(0, \sigma^2) \quad \text{for some } \sigma(\beta, f) > 0.$$

The variance  $\sigma^2$  and asymptotic behavior of  $\frac{1}{\sqrt{T}} \sum_{t=1}^T \mathbb{E}[f(\mathcal{X}_t)]$  in Theorem 2 are expressed in terms of a stationary distribution on increments (see Theorem 4(iv), and Proposition 9.1 for their explicit expressions). While the above is only conditional on  $\{\mathcal{T} \geq T\}$ , we find that  $\mathcal{T}$  and the height of  $\mathcal{P}_x$  are typically comparable (see Lemma 3.3):  $\lim_{h \rightarrow \infty} \mu_n(\mathcal{T} \geq (1 - \delta_\beta)h \mid \text{ht}(\mathcal{P}_x) \geq h) = 1$  (and  $\text{ht}(\mathcal{P}_x) \geq \mathcal{T}$  deterministically). In fact, we establish (see Theorem 4) that, conditioned on  $\{\mathcal{T} \geq T\}$ , the first cut-point typically appears at height  $O(\log T)$ , and the increment sequence captures all but a negligible portion of the pillar  $\mathcal{P}_x$ .



FIGURE 3. Two views of a pillar  $\mathcal{P}_x$  with  $\mathcal{T} = 20$  increments and its cut-points highlighted. On left: every pillar  $\mathcal{P}_y$  whose “shadow” (its projection on  $\mathbb{R}^2 \times \{0\}$ ) intersects that of  $\mathcal{P}_x$  will belong to the *same wall* in Dobrushin’s interface decomposition into walls and ceilings.

A special case of the above CLT is that the distribution of the “tip” of the pillar conditioned on having at least  $T$  increments is asymptotically Gaussian, as are its volume  $V(\mathcal{P}_x)$  and surface area  $A(\mathcal{P}_x)$ .

**Corollary 3.** *There exists  $\beta_0$  such that, for every  $\beta > \beta_0$  and sequences  $T = T_n$  with  $1 \ll T \ll n$  and  $x = (x_1, x_2, 0)$  where  $(x_1, x_2) \in \llbracket -n + \Delta_n + \frac{1}{2}, n - \Delta_n - \frac{1}{2} \rrbracket^2$  for some  $\Delta_n \gg T$ , the pillar at  $x$  has that its number of increments  $\mathcal{T} = \mathcal{T}(\mathcal{P}_x)$  and height  $\text{ht}(\mathcal{P}_x)$  satisfy, for some  $\lambda(\beta) > 1$ ,*

$$\text{ht}(\mathcal{P}_x)/\mathcal{T} \xrightarrow{\mathbb{P}} \lambda \text{ conditional on } \{\mathcal{T} \geq T\}. \quad (1.3)$$

Furthermore, conditional on  $\{\mathcal{T} \geq T\}$ , the height of  $\mathcal{P}_x$  is asymptotically Gaussian, and moreover:

(1) distribution of the tip: the variables  $(Y_1, Y_2, \text{ht}(\mathcal{P}_x)) \in \mathcal{P}_x$  (arbitrarily chosen if ambiguous) satisfy

$$\frac{(Y_1, Y_2, \text{ht}(\mathcal{P}_x)) - (x_1, x_2, \lambda T)}{\sqrt{T}} \implies \mathcal{N}\left(0, \begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & (\sigma')^2 \end{pmatrix}\right) \text{ for some } \lambda(\beta) > 1 \text{ and } \sigma(\beta), \sigma'(\beta) > 0.$$

(2) volume and surface area: there exist  $\lambda_i(\beta) > 1$  and  $\sigma_i(\beta) > 0$  ( $i = 1, 2$ ) such that

$$\frac{V(\mathcal{P}_x) - \lambda_1 T}{\sqrt{T}} \implies \mathcal{N}(0, \sigma_1^2), \quad \text{and} \quad \frac{A(\mathcal{P}_x) - \lambda_2 T}{\sqrt{T}} \implies \mathcal{N}(0, \sigma_2^2).$$

In order to establish the above results, one must control the behavior of the pillar below its first cut-point. But, it is precisely this part of the pillar where the effect of neighboring pillars is the most difficult to control: the abundance of nearby pillars around height 0 might in principal cause a pillar, conditioned to contain  $T$  increments, to have a large (diverging with  $T$ ) segment preceding its first increment. We account for this via a novel decomposition of the pillar into a *base* and a *spine*: the next result shows that the former’s total size is typically negligible, while the latter admits a detailed characterization in terms of its increment sequence.

**Theorem 4** (pillar structure). *There exists  $\beta_0 > 0$  such that the following holds for all  $\beta > \beta_0$ : for every sequence  $T = T_n$  with  $1 \ll T \ll n$  and  $x = (x_1, x_2, 0)$  with  $(x_1, x_2) \in \llbracket -n + \Delta_n + \frac{1}{2}, n - \Delta_n - \frac{1}{2} \rrbracket^2$  for some  $\Delta_n \gg T$ , there exist  $c, C > 0$  such that, conditional on  $\mathcal{T} \geq T$ , the pillar  $\mathcal{P}_x$  has the following structure:*

- (i) [Base] There is a cut-point  $v_{\tau_{\text{sp}}}$  so that the base of  $\mathcal{P}_x$ , defined as  $\mathcal{B}_x = \{y \in \mathcal{P}_x : \text{ht}(y) \leq \text{ht}(v_{\tau_{\text{sp}}})\}$ , satisfies  $\text{diam}(\mathcal{B}_x) \leq r$  except with probability  $O(\exp(-c\beta r))$  for every  $C \log T \leq r \leq T$ .
- (ii) [Spine] The increments  $\mathcal{X}_{\tau_{\text{sp}}+1}, \dots, \mathcal{X}_{\mathcal{T}}$  of the spine  $\mathcal{S}_x := \mathcal{P}_x \setminus \mathcal{B}_x$  satisfy, for every  $k, r \leq h$ , that the probability that  $A(\mathcal{X}_{\tau_{\text{sp}}+k}) \geq r$  is  $O(\exp(-c\beta r))$  (letting  $A(\mathcal{X}_t) := 0$  for  $t > \mathcal{T}$ ).
- (iii) [ $\alpha$ -mixing] For every  $k(T) > j(T)$ , if  $A_1 \in \mathcal{F}_1 := \sigma((\mathcal{X}_i)_{i=C \log T}^j)$  and  $A_2 \in \mathcal{F}_2 := \sigma((\mathcal{X}_i)_{i=k}^T)$  then the probability of  $A_1 \cap A_2$  differs from the product of the probabilities of  $A_i$  by  $O((k-j)^{-10})$ .
- (iv) [Asymptotic stationarity] There exists a stationary distribution  $\nu$  on  $\mathfrak{X}^{\mathbb{Z}}$  so that the conditional law of the increments  $(\dots, \mathcal{X}_{T/2-1}, \mathcal{X}_{T/2}, \mathcal{X}_{T/2+1}, \dots)$  given  $\mathcal{T} \geq T$  converges weakly to  $\nu$ .

(These are special cases of stronger statements, which do require additional definitions; for those results implying Items (i)–(iv), see Prop. 5.1, Prop. 4.1, Prop. 7.1 and Cor. 7.3, respectively.) As mentioned, each of these require delicately designed maps on interfaces, for which we can control both the change in probability under the map, and its multiplicity; the maps for Items (i)–(iv) are depicted in Figures 9–12 respectively.

**Remark 1.1.** Theorems 1–4 generalize naturally to all dimensions  $d \geq 3$ ; the main changes will be that the maximum  $M_n$  will have  $\frac{M_n}{\log n} \rightarrow \frac{(d-1)}{\alpha_\beta}$  in probability, and  $\alpha_\beta/\beta \rightarrow 2(d-1)$  as  $\beta \rightarrow \infty$ . The results and proofs are otherwise unchanged except that the constants will depend on the dimension  $d$ , and the lattice notation would be changed, e.g., the interface will be a connected set of  $(d-1)$ -cells, or plaquettes. For the sake of clarity of exposition and visualization we present all proofs in the most physical  $d = 3$  setting.

**Remark 1.2.** While Theorems 1–4 are w.r.t. the measure  $\mu_n$  (which is the Ising model on the infinite cylinder  $\Lambda_n$  with Dobrushin boundary conditions), the fact that the same results hold on the box  $\llbracket -n, n \rrbracket^d$  follows from a standard coupling argument. Indeed, by the exponential tails on interface fluctuations and on bubbles, the interfaces on  $\Lambda_n$  and  $\llbracket -n, n \rrbracket^d$  can be coupled to match with probability  $1 - O(e^{-cn})$ ; likewise their pillars  $\mathcal{P}_x$  conditioned on having at least  $T$  increments, agree with probability  $1 - O(e^{-cn})$  since  $T \ll n$ .

### 1.3. Tools and key ideas.

*Cluster expansion vs. Peierls maps under mixed boundary conditions.* The classical Peierls map—an injection from configurations with a specified bubble (a connected set of  $(d-1)$ -cells homeomorphic to a  $(d-1)$ -sphere) to ones without it, demonstrating that the energetic cost of such a bubble outweighs its entropy at large enough  $\beta$ —is a strikingly effective and robust tool for handling low-temperature behavior under homogeneous boundary conditions. There (within the plus or minus phase) it implies that for any dimension  $d \geq 2$ , bubbles are microscopic (and their size obeys an exponential tail) at low enough temperature. However, Peierls maps are insufficient to address the rigidity of the interface in the presence of Dobrushin’s boundary conditions: the natural attempt to define a Peierls map on configurations which would “flatten” the interface is hindered by (a) the interaction of the interface with nearby bubbles, and (b) its self-interactions due to overhangs.

To overcome this obstacle, Dobrushin used *cluster expansion* (cf. also [40]), a robust machinery that, in this case, allows one to disregard the floating bubbles and move to a distribution over interfaces  $\mathcal{I}$  given by

$$\mu_n(\mathcal{I}) \propto \exp \left[ -\beta|\mathcal{I}| + \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) \right], \quad (1.4)$$

where  $\mathbf{g}$  is a function (over interfaces  $\mathcal{I}$  with a marked face  $f$ ) which is uniformly bounded and local in the sense that  $|\mathbf{g}(f, \mathcal{I}) - \mathbf{g}(f', \mathcal{I}')|$  decays exponentially in the radius  $\mathbf{r}$  about which the balls  $B_{\mathbf{r}}(f)$  in  $\mathcal{I}$  and the local neighborhoods of  $f$  in  $\mathcal{I}$  and  $f'$  in  $\mathcal{I}'$  are isomorphic (see Theorem 2.21 in §2.5 for the full statement). N.b. that by moving to distributions on random interfaces, hiding the interacting bubbles in the Ising model, one loses several useful features of the Ising model: the law of  $\mathcal{I}$  does not have the domain Markov property, and there are long range interactions between faces in  $\mathcal{I}$ .

With this representation, properties of the Ising interface can be deduced from Peierls-like maps. The general strategy for utilizing such maps is as follows. Suppose we wish to show that some set of interfaces  $\mathcal{A}_r$  (e.g., those with height oscillations of at least  $r$  above the origin) is exponentially in  $r$  rare at  $\beta$  large. Then we construct a map  $\Psi$  sending  $\mathcal{A}_r$  to a subset  $\Psi(\mathcal{A}_r)$  of interfaces for which we have the following control:

- (1) energy gain: for every  $\mathcal{I} \in \mathcal{A}_r$ , the map  $\Psi$  induces an energy gain  $|\mathcal{I}| - |\Psi(\mathcal{I})| \geq r$ .
- (2) weight modification: for every  $\mathcal{I} \in \mathcal{A}_r$ , we obtain  $\frac{\mu_n(\mathcal{I})}{\mu_n(\Psi(\mathcal{I}))} \leq e^{-c\beta(|\mathcal{I}| - |\Psi(\mathcal{I})|)}$  from (1.4).
- (3) multiplicity: for all  $\ell \geq r$ , every  $\mathcal{J}$  in the image of  $\Psi$  has at most  $C^\ell$  pre-images  $\mathcal{I}$  with  $|\mathcal{I}| - |\mathcal{J}| = \ell$ .

(If we wish to show  $\mathcal{A}_r$  has small probability conditionally on some set  $\mathcal{B}$ , we further require  $\Psi(\mathcal{A}_r) \subset \mathcal{B}$ .) The complication in carrying this out is, of course, the function  $\mathbf{g}$ , which captures the very same obstacles that hindered the basic Peierls approach—the (hidden in the cluster expansion framework) bubbles in the Ising model and self-interactions of the interface. Ideally, one would be able to bound the effect of  $\mathbf{g}$  by comparing the faces  $f \in \mathcal{I}$  which were modified under  $\Psi$  to faces  $f' \in \mathcal{J}$  with isomorphic local neighborhoods.

*Dobrushin’s walls and ceilings decomposition and why it fails for LLN.* Dobrushin was able to carry out the above approach via a clever combinatorial decomposition of the interface, which reduced the analysis of the maps on the 3D interface to two-dimensional interactions. This decomposition is based on the following partition of  $\mathcal{I}$  tailored to view it as a perturbation of the flat interface  $\mathcal{L}_0 := \mathcal{F}(\mathbb{R}^2 \times \{0\})$ :

- A *ceiling face*  $f \in \mathcal{I}$  is a horizontal face whose projection on the  $xy$ -plane is unique among all faces of the interface  $\mathcal{I}$ . A *ceiling* of  $\mathcal{I}$  is a maximal connected component of ceiling faces.
- A *wall face*  $f \in \mathcal{I}$  is a non-ceiling face. A *wall* of  $\mathcal{I}$  is a maximal connected component of wall faces.

Consequently, one can “disregard” the ceilings as well as the vertical positions of every wall, and “standardize” each wall by moving it down to height zero, obtaining a *standard wall representation* of the 3D Ising interface. Importantly, this yields a bijection between collections of standard walls, and interfaces (see Lemma 2.12), akin to the contour representation of the 2D Ising configuration.

The natural attempt at a map  $\Psi$  is then to have it delete a specific wall  $W$  rooted at a face  $x \in \mathcal{L}_0$ , from the standard wall representation of the interface  $\mathcal{I}$ , then recover from the resulting standard wall collection, the interface  $\Psi(\mathcal{I})$ . The difficulty is, as usual, due to the function  $\mathbf{g}$ , and specifically due to *non-deleted* faces whose local neighborhoods would be vertically shifted by  $\Psi$ . To circumvent this, one may further delete any wall that is “too close” to  $W$ ; formally, one defines a *group of walls* according to some criterion of proximity, while relying on the fact that when walls are sufficiently far apart, the exponential decay of  $\mathbf{g}$  will negate their interaction. However, deleting too many additional walls can forfeit the second requirement from the map—control over its multiplicity. Dobrushin’s criterion was a carefully chosen middle-ground, importantly based solely on two-dimensional distances in the  $xy$  directions (see also Definition 2.23):

- Two walls  $W$  and  $W'$  are said to be “close” if the interface  $\mathcal{I}$  contains at least  $\text{dist}(x, x')^2$  faces above  $x$  or above  $x'$  for some  $x, x' \in \mathcal{L}_0$  in the projections of  $W$  and  $W'$  onto  $\mathcal{L}_0$  respectively.
- A *group of walls* if a maximal component of pairwise close walls.

(Note that “tall” walls are easier to group with, and the seemingly arbitrary threshold  $\text{dist}(x, x')^2$  plays a special role, via an isoperimetric inequality, in the analysis of faces deleted vs. ones that are only shifted.) The advantage in Dobrushin’s combinatorial decomposition is then that under the map  $\Psi$ , faces only undergo *vertical* shifts, and  $xy$ -distances between faces are preserved: as such the radius  $\mathbf{r}$  coming from  $\mathbf{g}$  can be expressed in terms of an  $xy$ -distance to the nearest deleted wall, so that the above definition of closeness enables the desired control on the contribution from the  $\mathbf{g}$  terms in (1.4) in terms of  $\beta(|\mathcal{I}| - |\mathcal{J}|)$ .

This argument showed that the group of walls adjacent to a fixed face  $x \in \mathcal{L}_0$  in  $\mathcal{I}$  has an exponential tail, implying the rigidity of  $\mathcal{I}$  and that its maximum height is  $O(\log n)$  with probability tending to 1. However, it is far too crude to handle subtle quantities of interest such as the asymptotics of the maximum (LLN) and the structure of the interface in a local neighborhood surrounding it (e.g., results à la Corollary 3):

1. The classification of faces into walls and ceilings does not relate well to the local spin configuration—as it depends on the behavior of the interface far above/below a face. But, the LLN (Theorem 1) does embed local spin-spin correlation: the leading order term of the maximum of  $\mathcal{I}$  is given in terms of a connective constant of spin agreement in infinite volume, which operations of walls are too coarse to reflect.
2. Recall that treating connected sets of wall faces as a single wall means any two connected wall-sets with intersecting shadows on the  $xy$ -plane are one and the same. While crucial to Dobrushin’s reduction of the problem to 2D, this comes in the way of analyzing the connected component of plus spins emanating from a fixed face  $x \in \mathcal{L}_0$ ; the taller this component is, the more pronounced this issue is (see Fig. 3, left).
3. Further bundling of walls into groups of walls attaches an extra layer of walls to a connected component of plus spins; moreover, the criterion for this bundling says that if the wall  $W_x$  of some face  $x \in \mathcal{L}_0$  has  $h$  faces above  $x$ , then it will collect every distinct wall  $W_y$  for  $y$  within a circle of area  $h$  centered about  $x$  (and so on, in a cascading manner). This would make it impossible to use this framework for more delicate questions such as tightness for the centered maximum (Problem 1.4).
4. Analyzing the effect of operations on walls (beyond simply *deleting the entire* group of walls of  $x \in \mathcal{L}_0$ ) is problematic: the collection of walls does not enjoy monotonicity / FKG inequalities, nor a domain Markov property (these properties are critical in the proof of sub-multiplicativity, as explained below).

*Maps on the increment sequence and base.* Unlike Dobrushin’s proof of the rigidity of  $\mathcal{I}$  which used maps to compare  $\mathcal{I}$  to flatten interfaces, in this work we construct Peierls-type arguments with reference interfaces that, rather than flat, have a three-dimensional large deviation above a point  $x \in \mathcal{L}_0$ :

1. At a high level, we would like our maps to “straighten” the pillar in the input interface  $\mathcal{I}$ , namely we would like to replace an increment in the pillar by a straight column of singleton boxes. The potential interactions of the pillar with its base, whose size and shape are much more difficult to control, necessitates that every map should first “flatten” the base as well. Consequently, we wish to use a map with a reference

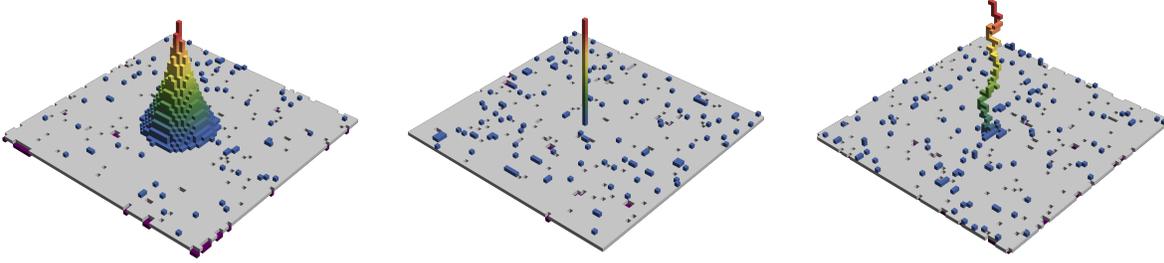


FIGURE 4. Typical pillars of the Discrete Gaussian model (left), the SOS model (middle), and the 3D Ising model (right) conditioned on the large deviation event  $\text{ht}(\mathcal{P}_x) \geq h$ .

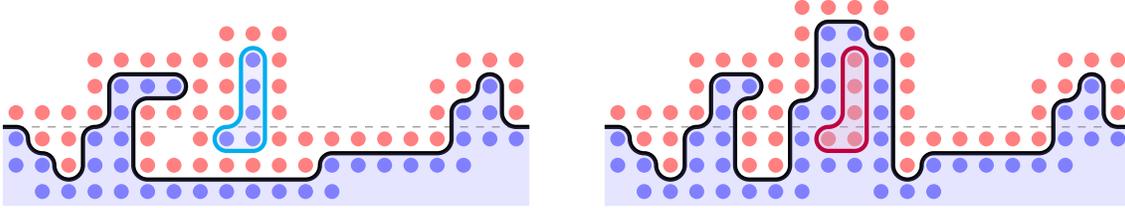


FIGURE 5. The pillar  $\mathcal{P}_x$  vs. the (+)-component  $\mathcal{S}$  above  $x$ : on left,  $\mathcal{P}_x = \emptyset$  whereas  $\mathcal{S} \neq \emptyset$  (the interface tunnels underneath  $\mathcal{S}$ ); on right,  $\mathcal{P}_x \neq \emptyset$  whereas  $\mathcal{S} = \emptyset$  (a minus bubble).

interface consisting of a flat plane appended to a modification of the random pillar  $\mathcal{P}_x$  (altered at its base and the designated increment we wish to control). This is achieved in two steps:

- (i) A map  $\Psi_i$  to straighten the increment  $\mathcal{X}_i$  (see §4.1 for its definition, and §4.2 for its proof strategy).
  - (ii) A map  $\Phi_{\mathcal{B}}$  to flatten the base (see §5.1 for its definition, and §5.2 for its proof strategy).
2. Whereas Dobrushin proofs only had vertical shifts, and thus interaction distances were controlled by 2D distances, in the above maps we must account for both horizontal and vertical shifts and their interplay. The subtle choice of  $v_{\tau_{\text{sp}}}$ , the “source point” for the spine as given in Theorem 4, serves as a key ingredient: in a sense it protects the pillar from interaction with neighboring ones (whose analysis is essential in the LLN for the maximum—see below) and isolates the effects of horizontal and vertical shifts: below  $v_{\tau_{\text{sp}}}$  faces will only be shifted vertically by our maps, and above it they will only undergo horizontal shifts.

*Establishing the limiting LD rate function.* As the leading order constant of the maximum of the interface is given by a solution to the LD problem of plus connectivity in infinite volume (much like the maximum of the surface in approximating models for the 3D Ising model such as the  $(2+1)$ D SOS and DG models were governed by LD problems; see Figure 4), a prerequisite to the proof of Theorem 1 is to establish existence of the limit given in (1.2). A standard approach to accomplish this would be to establish sub-multiplicativity or super-multiplicativity for  $a_h := \mu_{\mathbb{Z}^3}^{\mp}(A_h)$ , where  $A_h$  is the event in the right-hand of (1.2):

- One may expect  $(a_h)$  to be super-multiplicative, just like other increasing connection events in the Ising model and other monotone spin systems. However, if we reveal the  $+$  connection up to height  $h_1$  due to  $A_{h_1}$  in hope that only positive information is given on  $A_{h_1+h_2}$  (whereby FKG would provide the sought estimate), we find that at height  $h_1$  the measure is more negative than at height 0—the non-translation invariance of the boundary conditions makes a connection from  $h_1$  to  $h_2$  exponentially less likely than one from height 0 to  $h_1$ .
- Instead, we prove approximate sub-multiplicativity via a crucial application of Theorem 4(i). The notion of a pillar is well-suited to describe the (+)-component of  $x$  above height 0—which we may reveal up to height  $h_1$ . The  $(-)$  spins on its boundary yield negative information, which we may discard via monotonicity and domain Markov; however, this reveals additional (+)-spins at height 0, which encompass positive information. Yet these are part of the *base*  $\mathcal{B}_x$ , which Theorem 4 shows has size at most  $C \log^2 h_1$  with probability  $1 - o(1)$ . Tilting the measure by these (+)-spins thus costs a factor of  $e^{O(\log^2 h_1)} = e^{o(h_1)}$ , which does not affect the sought sub-multiplicativity bound.

A subtle point worthwhile stressing is that, despite the close connection between the pillar  $\mathcal{P}_x$  and the (+)-component above  $x$  in  $\mathbb{R}^2 \times [0, \infty)$ , neither one necessarily contains the other (see Figure 5).

*Maps on pairs of interfaces for mixing and stationarity.* In order to prove the more refined  $\alpha$ -mixing and stationarity properties of the increment sequence, we introduce 2-to-2 maps that act not on a single interface, but on a pair of interfaces. Importantly, with mixing and stationarity, our aim is not to show some set of interfaces is unlikely, but rather that some set of interfaces have roughly equal probability to some other set of interfaces: e.g., the pair  $(\mathcal{X}_j, \mathcal{X}_k) = (X_j, X_k)$  is roughly equally likely as  $(\mathcal{X}_j, \mathcal{X}_k) = (X_j, X'_k)$  in the case of mixing, and  $\mathcal{X}_j = X$  is roughly equally likely as  $\mathcal{X}_k = X$  in the case of stationarity. There is no relative energy gain here, so we need the cost in the exponent coming from the function  $\mathbf{g}$  in (1.4) to be  $o(1)$ . To resolve this, we instead pair up interfaces, and apply the map to pairs of interfaces, performing a swapping operation to be able to identify each face in the original pair of interfaces, with some face in the image pair of interfaces. We explain the subtleties in carrying this through in more detail in §7.1.1 and §7.2.1.

*Stein's method argument for the CLT.* The proof of the CLT in Proposition 9.1 (which implies Corollary 3) uses a Stein's method type argument which was used by Bolthausen [4] to handle stationary, mixing sequences of random variables (appealing to the new results on  $\alpha$ -mixing and stationarity obtained via the 2-to-2 maps). We explain the complications in our setting compared to that of [4] in §9.1.

*Comparison to Ornstein–Zernike theory.* We pause to compare our proof approach above to the well-known Ornstein–Zernike (OZ) theory of which the results of Theorem 4 may be reminiscent. Since the pioneering works [13, 14], there has been a remarkable line of work analyzing the structure of “long connections” in the high-temperature Ising model (all  $\beta < \beta_c$ ) in all dimensions  $d \geq 2$  using what is known as modernized OZ theory; the analysis was extended to the FK and Potts models (see, e.g., [15, 38]).

Namely, these works have analyzed, in the setting of the Ising model, the shape of a plus cluster connecting the origin to a site  $\vec{x}$  at distance  $\|\vec{x}\|$ . Via a decomposition into cut-points or cone-points, and increments between these, these works have identified a renewal structure in the long finite clusters of the high-temperature Ising model, with diffusive random-walk behavior at cut-points, and microscopic excursions in between.

In  $d = 2$ , by the duality between  $\beta < \beta_c$  and  $\beta > \beta_c$ , OZ theory directly translates to the low-temperature interface under Dobrushin boundary conditions. As such, for all  $\beta > \beta_c$ , the 2D Ising interfaces have been decomposed into cut-points with a renewal structure, and small increments in between with rapid decay of correlations; this was instrumental in pushing convergence of the interface to a Brownian bridge all the way to  $\beta_c$  [34]. In  $d \geq 3$ , there is no correspondence between high-temperature connections and low-temperature interfaces; rather, the more naturally analogous low-temperature event is a truncated connection event of the origin being connected by pluses to some  $x$  under the infinite-volume minus measure—in percolation language, a connection from 0 to  $x$  not connected to the *unique* infinite component.

By contrast, in our setting, the pillars of the plus phase are part of the infinite plus component, and are thinned by the *distinct* infinite minus component whose coexistence is forced by the boundary conditions. By Theorem 4, these pillars appear to have similar behavior beyond their first cut-point to long finite plus clusters in the minus phase. But, below that first cut-point there is a strong influence from the connection to the infinite plus component. The cut-point, increment decomposition is not helpful for dealing with these interactions with other branches of the infinite component (at the base); thus, controlling the base of the pillar is the most delicate part of our analysis.

It is therefore important to stress that, while appearing similar to our cut-point decomposition of pillars, *one cannot hope to characterize the pillars of the low temperature 3D Ising interface via the OZ theory.* Indeed, the OZ behavior is valid for all  $\beta > \beta_c$  in any dimension, whereas rigidity, let alone the results we prove, is conjectured to be false near  $\beta_c$  in dimension  $d = 3$ , as well as under any tilt in dimension  $d = 3$ .

**1.4. Related work and open problems.** In this section, we give a (by no means complete) overview of literature related to the analysis of random interfaces/surfaces describing separation of phases, and highlight some unresolved problems. As discussed, the pioneering work of Dobrushin rigorously established results on such interfaces of the Ising model via cluster expansion, including in particular rigidity at low temperatures in three (and higher) dimensions, and thus the existence of (infinite-volume) Gibbs measures describing the coexistence of phases. The approach of [28], outlined in §2.2–2.6, has been used to show rigidity for various other statistical physics models in  $d \geq 3$ , e.g., for the Widom–Rowlinson model [11, 12], the Falicov–Kimball models [20] and percolation and random-cluster/Potts models [33, 49]. We also mention that Van-Beijeren gave an elegant and simplified proof of the rigidity of the Ising interface using correlation inequalities in [48].

Subsequently, cluster expansion was instrumental in analyzing the analogous interface in two dimensions. This line of work culminated in the seminal monograph [25], showing that the shape of a macroscopic minus droplet in the plus phase takes after the *Wulff shape*, the convex body minimizing the surface energy to volume ratio (where the former is in terms of some explicit, analytic, surface tension  $\tau_\beta > 0$ ). Microscopic properties of an interface of angle  $\theta$  in an  $n \times n$  box are by now also very well-understood, with fluctuations on  $O(\sqrt{n})$  scales, and a scaling limit to a Brownian bridge [24, 25, 34, 35]; these hold up to the critical  $\beta_c$  [36, 37].

In dimensions three and higher, the microscopic features of the interface are only well-understood for approximations to the random surface separating the plus and minus phases, given by integer valued height functions  $\phi : [-n, n]^2 \rightarrow \mathbb{Z}$  on an  $n \times n$  box. Perhaps the most well-studied of these approximations is the Solid-On-Solid (SOS) model, going back to the 1950's (see [47] and [1]); the  $(2 + 1)$ -dimensional SOS model (approximating 3D Ising) is a special case of  $|\nabla\phi|^p$  models: a class of gradient models with Hamiltonians  $\mathcal{H}(\phi) = \sum_x \sum_i |\nabla_i \phi(x)|^p$  ( $p = 1$  is the SOS model, and  $p = 2$  is the discrete Gaussian model (DG)). In particular, the SOS Hamiltonian matches that of Ising with Dobrushin boundary conditions restricted to configurations where the intersection of the plus spins with each column  $\{(x_1, x_2, h) : h \in \mathbb{Z}\}$  is connected (i.e., SOS configurations have no overhangs or bubbles, which are microscopic in Ising in the  $\beta \rightarrow \infty$  limit).

In the setting of the SOS model at low temperatures, the maximum of the surface is typically of order  $\log n$  (see [9]). In [16, 17], its maximum was found to be tight around  $\frac{1}{2\beta} \log n$ , by showing that the probability of a “pillar above a face  $x$  reaching height  $h$ ” is  $\exp[-4\beta h + O(1)]$ ; on this large deviation event the interface looks like a vertical column of height  $h + O(1)$  with an  $O(1)$  “base” (c.f., Corollary 3, where for instance, the tip is delocalized, and see the depiction in Figure 4). Related properties in the presence of a floor inducing entropic repulsion were studied in [17], and extended to the discrete Gaussian and other  $|\nabla\phi|^p$ -models in [39].

**Problem 1.3.** For  $\alpha_\beta$  defined in (1.2), what are the asymptotics of  $\alpha_\beta - 4\beta$  (next order asymptotics of  $\alpha_\beta$ ) as  $\beta \rightarrow \infty$ ? in particular, is it the case that  $\alpha_\beta < 4\beta$ , so that 3D Ising is “rougher” than  $(2 + 1)$ D SOS?

While cluster expansion only converges at sufficiently large  $\beta$ , it is natural to ask if the rigidity of the interface, and our new results, hold for all  $\beta > \beta_c$ . This is not believed to be the case, as the Ising model is widely believed to undergo a *roughening transition* for  $d = 3$  (and no other dimension): much like the SOS and DG approximations, which exhibit phase transitions in  $\beta$ —whereby they roughen and resemble the discrete Gaussian free field [8, 30] for small  $\beta$ —it is conjectured that for the 3D Ising model there exists a point  $\beta_R > \beta_c$  such that, for  $\beta \in (\beta_c, \beta_R)$ , the model has long-range order, yet the typical fluctuations of its horizontal interface diverge with  $n$ ; proving this transition is a longstanding open problem (see, e.g., [1, 10]).

Much progress has been made in recent years on understanding the distribution of the maximum of the 2D discrete Gaussian free field and its local geometry. It is known for instance ([5–7]; see also, e.g., [50]) that this maximum is tight around an expected maximum that is asymptotically  $2\sqrt{2/\pi}(\log n - \frac{3}{8} \log \log n)$ , and that the centered maximum has the law of a randomly shifted Gumbel random variable.

**Problem 1.4.** What are the asymptotics of  $\mathbb{E}[M_n] - \frac{2}{\alpha_\beta} \log n$  (next order asymptotics of  $\mathbb{E}[M_n]$ ) as  $n \rightarrow \infty$ ? Are the fluctuations of the centered maximum  $O(1)$ , i.e., is the sequence  $\{\mu_n(M_n - \mathbb{E}[M_n] \cdot \cdot)\}$  tight?

We end this section with other well-studied perspectives on the 3D Ising model at low temperatures. While the interface-based approach of Dobrushin [28] proved to be extremely fruitful in 2D (where the results hold for interfaces in any angle), in 3D the combinatorics of that argument break down as soon as the ground state is not flat. It remains a well-known open problem to show that there do not exist non-translation invariant Gibbs measures corresponding to interfaces other than those parallel to the coordinate axes. The progress to date on roughness and fluctuations of “tilted interfaces” has been limited either to 1-step perturbations of a flat interface [41], or to results at zero temperature using rich connections to exactly solvable models [18].

In lieu of these approaches, a coarse-graining technique of Pisztor [44] enabled the establishment of surface tension and a Wulff shape scaling limit for the 3D Ising model at low-temperature: Cerf and Pisztor [19] considered an Ising model on an  $n \times n \times n$  box with all-plus boundary conditions, and showed that conditional on having  $(1 + \varepsilon)\mu^+(\sigma_0 = -1)n^3$  minus spins (atypically many), the largest minus cluster macroscopically takes on the corresponding Wulff shape. Results of this sort are focused on the macroscopic behavior of the model (as opposed to the interface-based approach) and do not describe the fluctuations around the limiting shape. In particular, the convergence to the Wulff shape holds all the way up to  $\beta_c$  (when combined with [2, 3]), even though near  $\beta_c$  (above the roughening transition) it is expected that the interface is not only delocalized, but that the minus cluster actually percolates all the way to the boundary of the box [10].

**1.5. Outline of Paper.** In §2, we first overview the notation of the paper and introduce Dobrushin's decomposition of the interface  $\mathcal{I}$  into *walls and ceilings*; then, in §2.6, we recap the proof of rigidity from [28] and the bounds this implies on  $\mu_n(\text{ht}(\mathcal{P}_x) \geq h)$ . In §3, we define increments of  $\mathcal{P}_x$ , and use them to split  $\mathcal{P}_x$  into its base and spine; in §4, we show that spine increments have an exponential tail on their size. In §5, we prove that the base of a pillar consisting of  $T$  increments has an exponential tail on its diameter beyond  $C \log T$ . Then in §6, we use the structural results of §3–5 to prove the existence of the large deviations rate (1.2); with this we prove the law of large numbers for the maximum, Theorem 1. In §7, we analyze finer properties of the increment sequence of  $\mathcal{P}_x$ , showing in §7.1 that correlations between increments decay polynomially in their distance, and in §7.2 that the increment sequences are asymptotically stationary. With these in hand, in §8, we prove a priori estimates on the mean and variance of observables of the increment sequence of  $\mathcal{P}_x$ , and in §9 combine the above to prove the CLT of Theorem 2 and deduce Corollary 3.

## 2. PRELIMINARIES: INTERFACES, CLUSTER EXPANSION AND RIGIDITY

In this section, we introduce key definitions from Dobrushin's decomposition of 3D Ising interfaces into *walls and ceilings* and recap his proof of rigidity of the Ising model interface. We modify the presentation of [28] slightly to track certain constants, and this will serve as a useful indication of the difficulties we will encounter when our reference interface is no longer a flat plane.

**2.1. Notation.** In this section we compile much of the notation used globally throughout the paper.

**2.1.1. Lattice notation.** Since the object of study in the present paper is the interface separating the plus and minus phases, we consider the Ising model as an assignment of spins to the vertices of the dual graph  $(\mathbb{Z}^3)^* = (\mathbb{Z} + \frac{1}{2})^3$  so that spins are assigned to the cells of  $\mathbb{Z}^3$  and interfaces are subsets of the faces of  $\mathbb{Z}^3$ .

Namely, let  $\mathbb{Z}^3$  be the integer lattice graph with vertices at  $(x_1, x_2, x_3) \in \mathbb{Z}^3$  and edges between nearest neighbor vertices (at Euclidean distance one). A *face* of  $\mathbb{Z}^3$  is the open set of points bounded by four edges (or four vertices) forming a square of side-length one, lying parallel to one of the coordinate axes. A face is *horizontal* if its normal vector is  $\pm e_3$ , and is *vertical* if its normal vector is one of  $\pm e_1$  or  $\pm e_2$ .

A *cell* or *site* of  $\mathbb{Z}^3$  is the open set of points bounded by six faces (or eight vertices) forming a cube of side-length one. We will frequently identify edges, faces, and cells with their midpoints, so that points with two integer and one half-integer coordinate are midpoints of edges, points with one integer and two half-integer coordinates are midpoints of faces, and points with three half-integer coordinates are midpoints of cells. A subset  $\Lambda \subset \mathbb{Z}^3$  identifies an edge, face, and cell collection via the edges, faces, and cells whose bounding vertices are all in  $\Lambda$ ; denote this edge set  $\mathcal{E}(\Lambda)$ , its face set  $\mathcal{F}(\Lambda)$  and its cell set  $\mathcal{C}(\Lambda)$ .

Two edges are adjacent if they share a vertex; two faces are adjacent if they share a bounding edge; two cells are adjacent if they share a bounding face. A set of faces (resp., edges, cells) is connected if for any pair of faces (edges, cells), there is a sequence of adjacent faces (edges, cells) starting at one and ending at the other. We will denote adjacency by the notation  $\sim$ .

It will also be useful to have a notion of connectivity in  $\mathbb{R}^3$  (as opposed to  $\mathbb{Z}^3$ ); we say that an edge/face/cell is  $*$ -adjacent to another edge/face/cell if and only if they share a bounding vertex.

Throughout the paper, we will use the notation  $d(x, y) = |x - y|$  to denote the Euclidean distance in  $\mathbb{R}^3$  between two points  $x, y$  (or if they are edges/faces/cells their respective midpoints). Similarly, we will use the notation  $B_r(x)$  to denote the (closed) Euclidean ball of radius  $r$  about the point  $x$ . When these balls are viewed as subsets of edges/faces/cells, we include all those whose midpoint is in  $B_r(x)$ . We further denote by  $A \oplus B$  the symmetric difference of the face sets  $A$  and  $B$ .

*Subsets of  $\mathbb{Z}^3$ .* The main subsets of  $\mathbb{Z}^3$  with which we will be concerned are of the form of cubes and cylinders. In view of that, define the centered  $2n \times 2m \times 2h$  box,

$$\Lambda_{n,m,h} := \llbracket -n, n \rrbracket \times \llbracket -m, m \rrbracket \times \llbracket -h, h \rrbracket \subset \mathbb{Z}^3,$$

where  $\llbracket a, b \rrbracket := \{a, a+1, \dots, b-1, b\}$ . We can then let  $\Lambda_n$  denote the special case of the cylinder  $\Lambda_{n,n,\infty}$ . The (outer) boundary  $\partial\Lambda$  of the cell set  $\mathcal{C}(\Lambda)$  is the set of cells in  $\mathcal{C}(\mathbb{Z}^3) \setminus \mathcal{C}(\Lambda)$  adjacent to a cell in  $\mathcal{C}(\Lambda)$ .

Additionally, for any  $h \in \mathbb{Z}$  let  $\mathcal{L}_h$  be the subgraph of  $\mathbb{Z}^3$  having vertex set  $\mathbb{Z}^2 \times \{h\}$  and correspondingly defined edge and face sets  $\mathcal{E}(\mathcal{L}_h)$  and  $\mathcal{F}(\mathcal{L}_h)$ . For a half-integer  $h \in \mathbb{Z} + \frac{1}{2}$ , let  $\mathcal{L}_h$  collect the faces and cells in  $\mathcal{F}(\mathbb{Z}^3) \cup \mathcal{C}(\mathbb{Z}^3)$  whose midpoints have half-integer  $e_3$  coordinate  $h$ . Finally we occasionally use  $\mathcal{L}_{>0} = \bigcup_{h>0} \mathcal{L}_h$  for the upper half-space and  $\mathcal{L}_{<0} = \bigcup_{h<0} \mathcal{L}_h$  for the lower half-space.

2.1.2. *Ising model.* An Ising configuration  $\sigma$  on  $\Lambda \subset \mathbb{Z}^3$  is an assignment of  $\pm 1$ -valued spins to the cells of  $\Lambda$ , i.e.,  $\sigma \in \{\pm 1\}^{\mathcal{C}(\Lambda)}$ . For a finite connected subset  $\Lambda \subset \mathbb{Z}^3$ , the Ising model on  $\Lambda$  with boundary conditions  $\sigma(\partial\Lambda) = \eta$  is the probability distribution over  $\sigma \in \{\pm 1\}^{\mathcal{C}(\Lambda)}$  given by

$$\mu_\Lambda^\eta(\sigma) \propto \exp[-\beta\mathcal{H}(\sigma)], \quad \text{where} \quad \mathcal{H}(\sigma) = \sum_{\substack{v,w \in \mathcal{C}(\Lambda) \\ v \sim w}} \mathbf{1}\{\sigma_v \neq \sigma_w\} + \sum_{\substack{v \in \mathcal{C}(\Lambda), w \in \partial\Lambda \\ v \sim w}} \mathbf{1}\{\sigma_v \neq \eta_w\}.$$

Throughout this paper, we will be considering the boundary conditions  $\eta_w = -1$  if  $w$  is in the upper half-space ( $w_3 > 0$ ) and  $\eta_w = +1$  if  $w$  is in the lower half-space ( $w_3 < 0$ ). We refer to these boundary conditions as *Dobrushin boundary conditions*, and denote them by  $\eta = \mp$ ; for ease of notation, let  $\mu_{n,m,h}^\mp = \mu_{\Lambda_{n,m,h}}^\mp$ .

*Domain Markov and FKG properties.* The Ising model is said to satisfy the *domain Markov property*, meaning that for any two finite subsets  $A \subset B \subset \mathcal{C}(\mathbb{Z}^3)$ , and every configuration  $\eta$  on  $B \setminus A$ ,

$$\mu_B(\sigma_A \in \cdot \mid \sigma_{B \setminus A} = \eta_{B \setminus A}) = \mu_A^{\eta \circ A}(\sigma_A \in \cdot),$$

where we use  $\sigma_A$  to denote the restriction of the configuration to the set  $A$ . It also satisfies an important consequence of its monotonicity, known as the *FKG inequality*. That is, for any two increasing (in the natural partial order on configurations) functions  $f, g : \{\pm 1\}^{\mathcal{C}(\Lambda)}$ , we have

$$\mathbb{E}_{\mu_\Lambda} [f(\sigma)g(\sigma)] \geq \mathbb{E}_{\mu_\Lambda} [f(\sigma)]\mathbb{E}_{\mu_\Lambda} [g(\sigma)].$$

A special case of this inequality, is when  $f$  and  $g$  are indicator functions of increasing events  $A$  and  $B$  (meaning that if  $\sigma \leq \sigma'$  and  $\sigma \in A$ , then  $\sigma' \in A$ , and similarly for  $B$ ), yielding  $\mu_\Lambda(A, B) \geq \mu_\Lambda(A)\mu_\Lambda(B)$ .

An increasing event that will appear in the proof of Theorem 1, is a connection event. Namely, we call a cell set a connected set of plus sites in  $\sigma$ , if it is a connected set of cells such that all the cells are assigned  $+1$  under  $\sigma$ . A *plus cluster* in  $\sigma$  is a maximal connected set of plus sites. If we denote by  $\{v \overset{+}{\leftrightarrow} w\}$  the event that  $v, w \in \mathcal{C}(\Lambda)$  are in the same plus cluster, we see that this is an increasing event. Finally, for a subset  $\Lambda' \subset \Lambda$ , denote by  $\{v \overset{+}{\leftrightarrow}_{\Lambda'} w\}$  the event that  $v, w$  are part of the same plus cluster using only cells of  $\Lambda'$ .

*Infinite-volume measures.* Care is needed to define the Ising model on infinite graphs, as the partition function becomes infinite; infinite-volume Gibbs measures are therefore defined via what is known as the *DLR conditions*; namely, for an infinite graph  $G$ , a measure  $\mu_G$  on  $\{\pm 1\}^G$ , defined in terms of its finite dimensional distributions, satisfies the DLR conditions if for every finite subset  $\Lambda \subset G$ ,

$$\mathbb{E}_{\mu_G(\sigma_{G \setminus \Lambda} \in \cdot)} [\mu_G(\sigma_\Lambda \in \cdot \mid \sigma_{G \setminus \Lambda})] = \mu_G(\sigma_\Lambda \in \cdot).$$

On  $\mathbb{Z}^d$ , infinite-volume Gibbs measures arise as weak limits of finite-volume measures, say  $n \rightarrow \infty$  limits of the Ising model on boxes of side-length  $n$  with certain prescribed boundary conditions. At low temperatures  $\beta > \beta_c(d)$ , the Ising model on  $\mathbb{Z}^d$  admits multiple infinite-volume Gibbs measures; taking plus and minus boundary conditions on boxes of side-length  $n$  yield the distinct infinite-volume measures  $\mu_{\mathbb{Z}^3}^+$  and  $\mu_{\mathbb{Z}^3}^-$  [40].

**2.2. Interfaces under Dobrushin boundary conditions.** We begin with the key combinatorial decomposition from [28] describing the interface separating the minus and plus phases under the Dobrushin boundary conditions. We refer the reader to [28] for more details.

**Definition 2.1** (Interfaces). For a domain  $\Lambda_{n,m,h}$  with Dobrushin boundary conditions, and an Ising configuration  $\sigma$  on  $\mathcal{C}(\Lambda_{n,m,h})$ , the *interface*  $\mathcal{I} = \mathcal{I}(\sigma)$  is defined as follows:

- (1) Extend  $\sigma$  to a configuration on  $\mathcal{C}(\mathbb{Z}^3)$  by taking  $\sigma_v = +1$  (resp.,  $\sigma_v = -1$ ) if  $v \in \mathcal{L}_{<0} \setminus \mathcal{C}(\Lambda_{n,m,h})$  (resp.,  $v \in \mathcal{L}_{>0} \setminus \mathcal{C}(\Lambda_{n,m,h})$ ).
- (2) Let  $F(\sigma)$  be the set of faces in  $\mathcal{F}(\mathbb{Z}^3)$  separating cells with differing spins under  $\sigma$ .
- (3) Call the (maximal)  $*$ -connected component of  $\mathcal{L}_0 \setminus \mathcal{F}(\Lambda)$  in  $F(\sigma)$ , the *extended interface*. (This is also the unique infinite  $*$ -connected component in  $F(\sigma)$ .)
- (4) The interface  $\mathcal{I}$  is the restriction of the extended interface to  $\mathcal{F}(\Lambda_{n,m,h})$ .

It is easily seen (by Borel–Cantelli) that taking the  $h \rightarrow \infty$  limit  $\mu_{n,m,h}$  to obtain the infinite-volume measure  $\mu_{n,m,\infty}^\mp$ , the interface defined above stays finite almost surely. Thus,  $\mu_{n,m,\infty}^\mp$ -almost surely, the above process also defines the interface for configurations on all of  $\mathcal{C}(\Lambda_{n,m,\infty})$ .

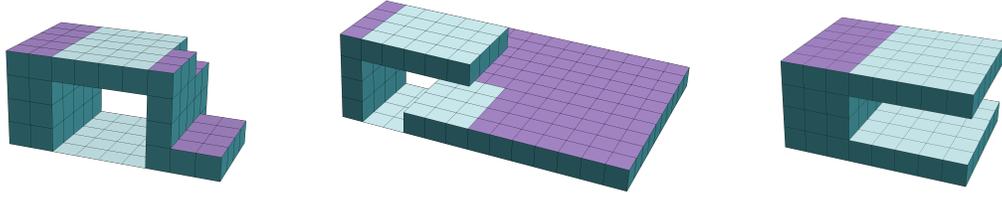


FIGURE 6. Three distinct standard walls, with their interior ceiling faces (purple) and wall faces (vertical in teal, horizontal in light blue) as per Definition 2.4. Middle example features two distinct (+) components in  $\mathbb{R}^2 \times \mathbb{R}_+$  which correspond to two distinct pillars but form a single wall (consistent with the fact that projections of distinct walls on  $\mathcal{L}_0$  are disjoint).

**Remark 2.2.** In lieu of the above definition of the interface due to [28], one could consider other flavors, e.g., letting  $\mathcal{I}$  be a *minimal* connected set of faces separating differing spins (or following some splitting rule which singles out a unique connected set of such faces, e.g., along the northeast diagonal in 2D). A simple Peierls argument implies that the set difference between that definition and Dobrushin’s definition consists of finite connected sets of faces with exponential tails on their size.

**Remark 2.3.** Just as Ising configurations with Dobrushin boundary conditions define an interface, every interface uniquely defines a configuration with exactly one \*-connected plus component and exactly one \*-connected minus component. For every  $\mathcal{I}$ , we can obtain this configuration  $\sigma(\mathcal{I})$  by iteratively assigning spins to  $\mathcal{C}(\Lambda_{n,m,h})$ , starting from the boundary and proceeding inwards, in such a way that adjacent sites have differing spins if and only if they are separated by a face in  $\mathcal{I}$ . Informally,  $\sigma(\mathcal{I})$  is distinguishing the sites that are in the “plus phase” and “minus phase” given the interface  $\mathcal{I}$ .

(Note that the extended interface also splits  $\mathcal{C}(\mathbb{Z}^3)$  into precisely two infinite connected (as opposed to \*-connected) components, along with possibly additional finite connected components.)

Following [28], we can decompose the faces in  $\mathcal{I}$  and define certain useful subsets of  $\mathcal{I}$ . For a face  $f \in \mathcal{F}(\mathbb{Z}^3)$ , its *projection*  $\rho(f)$  is the edge or face of  $\mathcal{L}_0$  given by  $\{(x_1, x_2, 0) : (x_1, x_2, s) \in f \text{ for some } s \in \mathbb{R}\} \subset \mathbb{R}^2 \times \{0\}$ . Specifically, the projection of a *horizontal* face (a face that is parallel to the plane  $\mathcal{L}_0$ ) is a face in  $\mathcal{F}(\mathcal{L}_0)$ , while the projection of a *vertical* face (one that is not parallel to  $\mathcal{L}_0$ ) is an edge in  $\mathcal{E}(\mathcal{L}_0)$ . The projection of a collection of faces  $F$  is  $\rho(F) := \bigcup_{f \in F} \rho(f)$ , which may consist both of edges and faces of  $\mathcal{L}_0$ .

**Definition 2.4** (Ceilings and walls). A face  $f \in \mathcal{I}$  is a *ceiling face* if it is horizontal and there is no  $f' \in \mathcal{I} \setminus \{f\}$  such that  $\rho(f) = \rho(f')$ . A face  $f \in \mathcal{I}$  is a *wall face* if it is not a ceiling face. A *wall* is a (maximal) \*-connected set of wall faces. A *ceiling* of  $\mathcal{I}$  is a (maximal) \*-connected set of ceiling faces.

**Definition 2.5** (Floors of walls). For a wall  $W$ , the complement of its projection (a subset of  $\mathbb{R}^2$ )

$$\rho(W)^c := (\mathcal{E}(\mathcal{L}_0) \cup \mathcal{F}(\mathcal{L}_0)) \setminus \rho(W)$$

splits into one infinite component, and some finite ones. Any ceiling adjacent to the wall  $W$  projects into one of these components; the one that projects into the infinite component is called the *floor* of  $W$ .

This can be reinterpreted with the following notion of nesting of walls and ceilings.

**Definition 2.6.** We say an edge or face  $u \in \mathcal{E}(\mathcal{L}_0) \cup \mathcal{F}(\mathcal{L}_0)$  is interior to a wall  $W$  if  $u$  is not in the infinite component of  $\rho(W)^c$ .

A wall  $W$  is *interior* to (or nested in) a wall  $W'$  if every element of  $\rho(W)$  is interior to  $W'$ . Similarly, a ceiling  $C$  is *interior* to a wall  $W$  if every element of  $\rho(C)$  is interior to  $W$ .

Observe that of the ceilings  $C_0, C_1, \dots, C_l$  adjacent to a wall  $W$ , one of them is the floor of  $W$ —say  $C_0$ —and the rest are interior to  $W$ . For any admissible pair of standard walls, as their projections are disjoint, either one wall is nested in the other, or  $\rho(W_x)$  is contained in the infinite component of  $\rho(W_y)^c$  and vice versa.

**Definition 2.7** (Standard walls). A wall  $W$  is a *standard wall* if there exists an interface  $\mathcal{I}_W$  such that  $\mathcal{I}_W$  has exactly one wall,  $W$ —as such it must have as its unique floor a subset of  $\mathcal{L}_0$ . A collection of standard walls is *admissible* if they are all disjoint and have pairwise disjoint projections (see Figure 6).

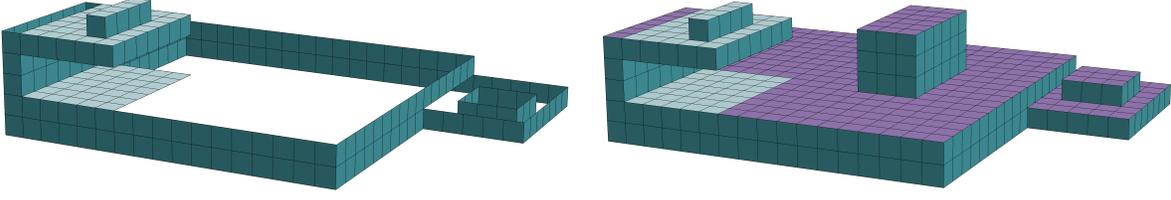


FIGURE 7. Correspondence between an interface and its standard wall representation (Lemma 2.12): three distinct standard walls (left) and their corresponding interface (right).

**Lemma 2.8** ([28]). *For a projection of the walls of an interface, each connected component of that projection (as a subset of edges and faces) corresponds to a single wall. Moreover, there is a 1-1 correspondence between the ceilings adjacent to a standard wall  $W$  and the connected components of  $\rho(W)^c$ . Similarly, for a wall  $W$ , all other walls  $W' \neq W$  can be identified to the connected component of  $\rho(W)^c$  they project into, and in that manner they can be identified to the ceiling of  $W$  to which they are interior.*

**Definition 2.9** (Standardization of walls). To each ceiling  $C$ , we can identify a unique *height*  $\text{ht}(C)$  since all faces in the ceiling have the same  $x_3$  coordinate. For every wall  $W$ , we can define its *standardization*  $\theta_{\text{ST}}(W)$  which is the translate of the wall by  $(0, 0, -s)$  where  $s$  is the height of its floor.

**Remark 2.10.** We can index walls as follows: assign an ordering of the faces of  $\mathcal{L}_0$ , and index  $W$  by the minimal face in  $\mathcal{L}_0$  that shares an edge with  $\rho(W)$ , and lies either in  $\mathcal{F}(\rho(W))$  or in one of the finite connected components of  $\rho(W)^c$ . For any admissible collection of standard walls, the indices of the walls are distinct.

We then have the following important bijection between interfaces and their standard wall representation.

**Definition 2.11.** Let the *standard wall representation* of an interface  $\mathcal{I}$  be the collection of standard walls given by standardizing all walls of  $\mathcal{I}$ .

**Lemma 2.12** ([28]). *There is a 1-1 correspondence between the set of interfaces and the set of admissible collections of standard walls. In particular, the standardization  $\theta_{\text{ST}}(W)$  of a wall  $W$  is a standard wall.*

*Proof.* From an interface, the standard wall representation is an admissible collection of standard walls as projections of distinct walls are disjoint. To obtain an interface from an admissible collection of standard walls, it suffices to take the standard wall representation of an interface  $\mathcal{I}$  and describe how the addition of one standard wall  $\theta_{\text{ST}}(W_{t_0})$ , compatible with the standard walls of  $\mathcal{I}$  and not interior to any walls in  $\mathcal{I}$ , changes  $\mathcal{I}$  to  $\mathcal{I}'$ . (One could then construct an interface  $\mathcal{I}$  from its standard wall representation by beginning with the interface  $\mathcal{L}_0$  with empty standard wall representation, and iterating the above procedure, adding the standard walls from innermost outward).

Consider an interface  $\mathcal{I}$  with standard wall collection  $(\theta_{\text{ST}}W_t)_{t \neq t_0}$  such that  $((\theta_{\text{ST}}W_t)_{t \neq t_0}) \cup \theta_{\text{ST}}(W_{t_0})$  is admissible; suppose further that  $\theta_{\text{ST}}$  is not interior to any wall of  $\mathcal{I}$ . Let  $\mathcal{J}_{W_{t_0}}$  be the interface whose only wall is the standard wall  $\theta_{\text{ST}}W_{t_0}$ , and denote its floor by  $C_0$  and non-floor ceilings by  $C_1, \dots, C_l$ .

Construct a face set from  $\mathcal{I}$  and  $\theta_{\text{ST}}W_{t_0}$  as follows:

- (1) Remove all horizontal faces of  $\mathcal{I}$  in  $\rho(W_{t_0})$ .
- (2) Vertically shift every face of  $\mathcal{I}$  projecting into one of  $\rho(C_i)_{1 \leq i \leq l}$  by  $\text{ht}(C_i)$ .
- (3) Add all faces of  $\theta_{\text{ST}}W_{t_0}$ .

The resulting face set is evidently a valid interface  $\mathcal{I}'$  and one can check that it has standard wall representation  $(\theta_{\text{ST}}W_t)_{t \neq t_0} \cup \theta_{\text{ST}}W_{t_0}$ . ■

We note the following important observation based on the above bijection.

**Observation 2.13.** *Consider interfaces  $\mathcal{I}$  and  $\mathcal{J}$ , such that the standard wall representation of  $\mathcal{I}$  contains that of  $\mathcal{J}$  (and additionally has the standardizations  $\mathbf{W} = W_1, \dots, W_r$ ). By the construction in Lemma 2.12, there is a 1-1 map between the faces of  $\mathcal{I} \setminus \mathbf{W}$  and the faces of  $\mathcal{J} \setminus \mathbf{H}$  where  $\mathbf{H}$  is the set of faces in  $\mathcal{J}$*

projecting into  $\rho(\mathbf{W})$ . Moreover, this bijection can be encoded into a map  $f \mapsto \tilde{f}$  that only consists of vertical shifts, and such that all faces projecting into the same component of  $\rho(\mathbf{W})^c$  undergo the same vertical shift.

Finally, we introduce a notion of nested walls which will prove useful to bounding the base of tall pillars.

**Definition 2.14.** To any edge/face/cell  $x$ , we can assign a *nested sequence of walls*  $\mathfrak{W}_x = \bigcup_s W_{u_s}$  that is composed of all walls that  $\rho(x)$  is interior to (by Definition 2.6, this forms a nested sequence of walls).

**Observation 2.15.** For  $u \in \mathcal{L}_0$ , for a nested sequence of walls  $\mathfrak{W}_u$ , one can read off the height of the face(s) of  $\mathcal{I}$  projecting onto  $u$ . In particular, if a face  $f \in \mathcal{I}$  has height  $h$ , its nested sequence of walls must be such that the sum of the heights of the walls in  $\mathfrak{W}_{\rho(f)}$  exceeds  $h$ .

**2.3. Interface pillars.** The above definitions were all from [28] and, informally, they reduce the analysis of 3D Ising interfaces to that of a low-temperature 2D polymer model given by the projections of walls. For us, this is insufficient as we aim to study the structure of tall walls, wherein the projection does not carry much information about the shape and height. As such, we define the notion of a pillar above  $x \in \mathcal{L}_0$ .

**Definition 2.16 (Pillars).** For every interface  $\mathcal{I}$ , consider the restriction  $\sigma(\mathcal{I})|_{\mathcal{L}_{>0}}$  of the Ising configuration  $\sigma(\mathcal{I})$  to the upper half-space. For any face  $x \in \mathcal{L}_0$ , the cell-set  $\sigma(\mathcal{P}_x)$  of the *pillar*  $\mathcal{P}_x = \mathcal{P}_x(\mathcal{I})$  above  $x$  will be the (possibly empty)  $*$ -connected plus component in  $\sigma(\mathcal{I})|_{\mathcal{L}_{>0}}$  containing  $x + (0, 0, \frac{1}{2})$ . The pillar  $\mathcal{P}_x$  will have face-set consisting of the bounding faces of  $\sigma(\mathcal{P}_x)$  in the upper half-space, so that it is a subset of  $\mathcal{I}$ .

Pillars can be viewed as a subset of some collection of nested walls along with their ceilings as follows:

**Observation 2.17.** The pillar  $\mathcal{P}_x$  is described by  $\mathfrak{W}_x$  together with all walls that are nested in some wall of  $\mathfrak{W}_x$ ; namely, if we index walls by enumerating faces of  $\mathcal{L}_0$  in terms of distance to  $x$ , then the set of walls  $\bigcup_{y: d(y,x) \leq \text{diam}(\mathcal{B}_x)} W_y$  contain all the information about the pillar  $\mathcal{P}_x$ . Moreover,  $\mathcal{P}_x \cap (\mathbb{R}^2 \times ([\text{ht}(v_1)], \infty))$  (possibly with the exception of one upper delimiting face) is all a subset of a single wall.

Much of this paper is interested in the large deviations regime for the height of such pillars, so we formally define heights of interface subsets.

**Definition 2.18.** For a point  $(x_1, x_2, x_3) \in \mathbb{R}^3$ , we say its height is  $\text{ht}(x) = x_3$ . The height of a cell is the height of its midpoint. For a pillar  $\mathcal{P}_x \subset \mathcal{I}$ , its height is given by

$$\text{ht}(\mathcal{P}_x) = \sup\{x_3 : (x_1, x_2, x_3) \in f, f \in \mathcal{P}_x\}.$$

It is important to distinguish between situations where  $\mathcal{P}_x$  is empty because the interface lies exactly at face  $x$ , and when it goes below face  $x$ ; in view of this, if  $\mathcal{P}_x = \emptyset$ , we say that  $\text{ht}(\mathcal{P}_x) = 0$  if  $x - (0, 0, \frac{1}{2})$  is in the plus phase (i.e., is plus in  $\sigma(\mathcal{I})$ ), and  $\text{ht}(\mathcal{P}_x) < 0$  if  $x - (0, 0, \frac{1}{2})$  is in the minus phase.

**2.4. Excess area.** For a pair of interfaces, we need to quantify the energy cost/gain of having one interface over the other. The competition of this energy cost with respect to the interface  $\mathcal{L}_0$  with the entropy gain from additional fluctuations governs the behavior of the Dobrushin interface.

**Definition 2.19 (Excess area).** For two interfaces  $\mathcal{I}, \mathcal{I}'$ , the *excess area* of  $\mathcal{I}$  with respect to  $\mathcal{I}'$ , denoted  $\mathbf{m}(\mathcal{I}; \mathcal{I}')$ , is given by

$$\mathbf{m}(\mathcal{I}; \mathcal{I}') := |\mathcal{I}| - |\mathcal{I}'|,$$

where these are the cardinalities of the face-sets of  $\mathcal{I}$  and  $\mathcal{I}'$  respectively. Evidently, for any Dobrushin interface  $\mathcal{I}$ , we have that  $\mathbf{m}(\mathcal{I}; \mathcal{L}_0 \cap \Lambda) \geq 0$ .

We can also define excess areas for subsets of interfaces, and interpret these as the “excess area of the interface that contains the subset with respect to a reference one that does not.” For instance, for a standard wall  $W$ , if we denote by  $\mathcal{I}_W$  the interface whose only wall is  $W$ , then  $\mathbf{m}(W) = \mathbf{m}(\mathcal{I}_W; \mathcal{L}_0 \cap \Lambda)$ . For a wall  $W$ , its excess area is given by the excess area of the standard wall  $\theta_{\text{ST}}(W)$ . The excess area of a collection of walls  $F$  is analogously defined, and one can easily see that  $\mathbf{m}(F) = \sum_{W \in F} \mathbf{m}(W)$ .

Finally, define the excess area of a pillar  $\mathbf{m}(\mathcal{P}_x)$ , and of one pillar with respect to another,  $\mathbf{m}(\mathcal{P}_x; \mathcal{P}'_x)$ , via the excess areas of the unique interfaces consisting only of the faces in  $\mathcal{P}_x$  (resp.,  $\mathcal{P}'_x$ ) along with faces of  $\mathcal{L}_0$ .

**Remark 2.20.** Notice that for a wall  $W_x$ , its excess area is exactly given by

$$\mathbf{m}(W_x) = \mathbf{m}(\theta_{\text{ST}}(W_x)) = |W_x| - |\mathcal{F}(\rho(W_x))|$$

where  $\mathcal{F}(\rho(W_x))$  is the face set of the projection  $\rho(W_x)$ . Moreover, for an interface  $\mathcal{I}$  having standard wall representation  $(W_t)_{t \in \mathcal{L}_0}$  per Lemma 2.12, we have that

$$\mathfrak{m}(\mathcal{I}) = \mathfrak{m}(\mathcal{I}; \mathcal{L}_0 \cap \Lambda) = \sum_{W \in (W_x)_{x \in \mathcal{L}_0}} \mathfrak{m}(W).$$

As observed in [28], this form of the excess area makes a few key properties clear:

$$\mathfrak{m}(W_x) \geq \frac{1}{2}|W_x| \quad \text{and} \quad \mathfrak{m}(W_x) \geq |\mathcal{E}(\rho(W_x))| + |\mathcal{F}(\rho(W_x))|. \quad (2.1)$$

Moreover, any two faces  $x, y \in \mathcal{L}_0$  interior to the projection  $\rho(W_x)$  satisfy  $|x - y| \leq \mathfrak{m}(W_x)$ .

**2.5. Cluster expansion for interfaces describing phase coexistence.** Cluster expansion is a classical tool for expressing the partition function of a spin system on a domain as a product of polymer weights (in an appropriate polymer representation of the model) rather than as a sum over weights of configurations. Crucially, this product is an infinite product that only converges in perturbative regimes (e.g., for us  $\beta \gg 1$ ).

In our setting of the Ising model, these polymers are minimal connected sets of faces which separate differing spins, and are the bounding face-set of a connected set of cells. The associated weight of such a face-set  $\gamma$  is given by  $e^{-\beta|\gamma|}$ . The polymers are then endowed with hard-core interaction rules encoding the admissibility of a collection of polymers, so that it in fact encodes uniquely, an Ising spin configuration. For a full derivation of the validity of cluster expansion, we refer the reader to the book [29, Chapter 5]. In our setting, the hard-core polymer interactions preclude distinct polymers from sharing any edges or vertices.

Using this cluster expansion, [40] proved properties of the single-phase Ising measures  $\mu_{\mathbb{Z}^3}^-$  and  $\mu_{\mathbb{Z}^3}^+$  at low temperatures. An easy implication of this cluster expansion is that one can take a limit of  $\mu_{\Lambda_{n,n,h}}^\mp$  as  $h \rightarrow \infty$  and obtain an infinite-volume Gibbs measure on the cylinder  $\Lambda_n = \Lambda_{n,n,\infty}$  whose interface is finite almost surely (for each fixed  $n$ ), and this limit does not depend on the boundary conditions taken at the top and bottom of  $\Lambda_{n,n,h}$ : see [28, (2.7) as well as Lemma 3]. Denote this limiting measure  $\mu_n = \mu_{\Lambda_{n,n,\infty}}^\mp$ .

Applying the cluster expansion one can compute probabilities of interfaces under this  $\mu_n$  measure.

**Theorem 2.21** ([28, Lemma 1]). *Consider the Ising measure  $\mu_n = \mu_n^\mp$  on the cylinder  $\Lambda_{n,n,\infty}$ . There exists  $\beta_0 > 0$  and a function  $\mathbf{g}$  such that for every  $\beta > \beta_0$  and any two interfaces  $\mathcal{I}$  and  $\mathcal{I}'$ ,*

$$\frac{\mu_n(\mathcal{I})}{\mu_n(\mathcal{I}')} = \exp \left( -\beta \mathfrak{m}(\mathcal{I}; \mathcal{I}') + \left( \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \mathcal{I}'} \mathbf{g}(f', \mathcal{I}') \right) \right)$$

and  $\mathbf{g}$  satisfies the following for some  $\bar{c}, \bar{K} > 0$  independent of  $\beta$ : for all  $\mathcal{I}, \mathcal{I}'$  and  $f \in \mathcal{I}$  and  $f' \in \mathcal{I}'$ ,

$$|\mathbf{g}(f, \mathcal{I})| \leq \bar{K} \quad (2.2)$$

$$|\mathbf{g}(f, \mathcal{I}) - \mathbf{g}(f', \mathcal{I}')| \leq \bar{K} e^{-\bar{c}\mathbf{r}(f, \mathcal{I}; f', \mathcal{I}')} \quad (2.3)$$

where  $\mathbf{r}(f, \mathcal{I}; f', \mathcal{I}')$  is the largest radius around the origin on which  $\mathcal{I} - f$  ( $\mathcal{I}$  shifted by the midpoint of the face  $f$ ) is congruent to  $\mathcal{I}' - f'$ . That is to say,

$$\mathbf{r}(f, \mathcal{I}; f', \mathcal{I}') := \sup \{ r : (\mathcal{I} - f) \cap B_r(0) \equiv (\mathcal{I}' - f') \cap B_r(0) \}$$

where the congruence relation  $\equiv$  is equality as subsets of  $\mathbb{R}^3$ , up to, possibly, reflections and  $\pm \frac{\pi}{2}$  rotations in the horizontal plane.

Throughout the rest of the paper, the constants  $\bar{c}$  and  $\bar{K}$  will be reserved for those of (2.2)–(2.3).

**Remark 2.22.** In [28] and other works, the congruence above is written only as a congruence up to translation. However, one can see by following the derivation of Theorem 2.21, that this congruence can also be up to reflections and  $\pm \frac{\pi}{2}$  rotations in the  $xy$ -plane (under which the Ising Hamiltonian is invariant). More precisely, for polymer weights  $w(\gamma)$  and interactions  $\delta(\gamma_i, \gamma_j')$ , we can define the *Ursell functions* as

$$\varphi(\gamma_1, \dots, \gamma_m) = \frac{1}{m!} \sum_{G \subset K_m} \prod_{(i,j) \in G} [\delta(\gamma_i, \gamma_j) - 1],$$

where the sum is over connected subgraphs of the complete graph on  $m$  vertices. The cluster expansion formally expresses the partition function of the Ising model on a graph as

$$\mathcal{Z} = \exp \left[ \sum_{m \geq 1} \sum_{\gamma_1, \dots, \gamma_m} \varphi(\gamma_1, \dots, \gamma_m) \prod_{i \leq m} w(\gamma_i) \right].$$

Theorem 2.21 arises from viewing  $\mu_n^\mp$  with interface  $\mathcal{I}$  as a cost from the disagreements along  $\mathcal{I}$ , along with one Ising model above  $\mathcal{I}$  with minus boundary conditions, and one below  $\mathcal{I}$  with plus boundary conditions. The function  $\mathbf{g}$  is therefore given by simple algebraic manipulations from the Ursell functions and polymer weights, all of which are invariant under reflections and rotations in the  $xy$ -plane.

We end this section with a piece of terminology that we will use frequently. We will say that the radius  $\mathbf{r}(f, \mathcal{I}; f', \mathcal{I}')$  is *attained* by a face  $g \in \mathcal{I}$  (resp.,  $g' \in \mathcal{I}'$ ) of minimal distance to  $f$  (resp.,  $f'$ ) whose presence prevents  $\mathbf{r}(f, \mathcal{I}; f', \mathcal{I}')$  from being any larger.

**2.6. Rigidity of Dobrushin interfaces.** For the benefit of the reader, we include Dobrushin's proof of rigidity for 3D interfaces from [28], namely that the walls corresponding to horizontal interfaces have exponential tails on their excess areas. This will straightforwardly imply that the probability that the pillar above a face  $x \in \mathcal{L}_0$  reaches a height  $h$  has an exponentially decaying tail. We will need the following definition of [28] that collects walls that are close, and therefore excessively interact with one another, together.

**Definition 2.23.** For a wall  $W$ , for every edge or face  $u \in \rho(W)$ , let  $N_\rho(u) = \#\{f \in W : \rho(f) = u\}$ . We say that two walls  $W_1$  and  $W_2$  are *close* if there exist  $u_1 \in \rho(W_1)$  and  $u_2 \in \rho(W_2)$  such that

$$|u_1 - u_2| \leq \sqrt{N_\rho(u_1)} + \sqrt{N_\rho(u_2)}.$$

Then an admissible set of standard walls  $F = \bigcup_i W_i$  is a *group of walls* if it is a maximal connected component (via the adjacency relation induced by closeness) of walls i.e., every wall in  $F$  is close to some other wall in  $F$  and no wall not in  $F$  is close to a wall of  $F$ . Index a group of walls by the minimal index of its walls, and let  $(\mathcal{F}_x)_{x \in \mathcal{L}_0}$  be the admissible group of wall collection of  $\mathcal{I}$ .

Following the definition of admissible sets of standard walls and Lemma 2.12, it should be clear how admissible collections of groups of walls would be defined, and that the set of all admissible collections of groups of walls are in 1-1 correspondence with the set of all possible Dobrushin interfaces (see §5 of [28]).

**Remark 2.24.** The procedure for sorting the faces of  $\mathcal{L}_0$  and using this ordering to identify each group of walls by the appropriate minimal face in  $\mathcal{L}_0$  that can be used to identify the group of walls, will be called an *indexing* of  $\mathcal{I}$ . Our results will easily be seen to hold uniformly over this indexing (i.e., uniformly over all orderings of the faces in  $\mathcal{L}_0$ ).

**Lemma 2.25** ([28, Lemma 8]). *There exists  $\beta_0$  and a universal  $C$  such that for  $\beta > \beta_0$ , for any admissible collection of groups of walls  $(\mathcal{F}_y)_{y \neq x}, \mathcal{F}_x$ , we have*

$$\frac{\mu_n(\mathcal{F}_x = F_x, (\mathcal{F}_y)_{y \neq x} = (F_y)_{y \neq x})}{\mu_n(\mathcal{F}_x = \emptyset, (\mathcal{F}_y)_{y \neq x} = (F_y)_{y \neq x})} \leq \exp[-(\beta - C)\mathbf{m}(F_x)].$$

The above readily implies an exponential tail on the size of the group of walls indexed by face  $x \in \mathcal{L}_0$ . In fact, it can easily be used to show that the probability that the interface intersects the column  $\{(x_1, x_2, s) : s \in \mathbb{R}\}$  above a height  $H$  decays exponentially in  $H$ , and with our definition of pillars, we can also use it to show that it implies an exponential tail on  $\text{ht}(\mathcal{P}_x)$ .

**Theorem 2.26** ([23, 27, 28], see also [12]). *There exists  $C > 0$  such that for every  $\beta > \beta_0$ , for every  $x \in \mathcal{L}_0 \cap \Lambda$ , and every  $r \geq 1$ ,*

$$\mu_n(\mathbf{m}(\mathcal{F}_x) \geq r) \leq \exp[-(\beta - C)r].$$

Furthermore, we have that for every  $h \geq 1$ ,

$$\mu_n(\text{ht}(\mathcal{P}_x) \geq h) \leq \exp[-4(\beta - C)h].$$

**Proof of Lemma 2.25.** Let  $\Phi_x$  be the map that takes an interface  $\mathcal{I}$  and eliminates its group of walls  $\mathcal{F}_x$  (if such a group of walls is nonempty), generating the new interface as per Lemma 2.12. Now for ease of notation, let  $\mathcal{I}$  be the interface with the collection of groups of walls  $(\mathcal{F}_y)_y = (F_y)_y$  and let  $\mathcal{I}'$  be the one with the collection of groups of walls  $(\mathcal{F}'_y)_y$  where  $\mathcal{F}'_y = F_y$  for  $y \neq x$  whereas  $F'_x = \emptyset$ , so that  $\mathcal{I}' = \Phi_x(\mathcal{I})$  and  $\mathbf{m}(\mathcal{I}; \mathcal{I}') = |\mathcal{I}| - |\mathcal{I}'| = \mathbf{m}(F_x)$ . By Theorem 2.21, we have

$$\frac{\mu_n(\mathcal{F}_x = F_x, (\mathcal{F}_y)_{y \neq x} = (F_y)_{y \neq x})}{\mu_n(\mathcal{F}_x = \emptyset, (\mathcal{F}_y)_{y \neq x} = (F_y)_{y \neq x})} = \frac{\mu_n(\mathcal{I})}{\mu_n(\mathcal{I}')} = \exp\left(-\beta \mathbf{m}(F_x) + \left(\sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \mathcal{I}'} \mathbf{g}(f', \mathcal{I}')\right)\right).$$

We wish to bound the absolute value of the difference of the sums in the right-hand side. Denote the walls constituting  $F_x$  by  $W_{x_1}, W_{x_2}, \dots, W_{x_l}$  for some  $l$ . Recall from Observation 2.13, the 1-1 correspondence between  $\mathcal{I} \setminus F_x$ , and the faces of  $\mathcal{I}'$  that do not project in to  $\mathcal{F}(\rho(F_x))$  and encode it with the notation  $f \mapsto \tilde{f}$ . Then, we have

$$\begin{aligned} \left| \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \mathcal{I}'} \mathbf{g}(f', \mathcal{I}') \right| &\leq \sum_{f \in F_x} |\mathbf{g}(f, \mathcal{I})| + \sum_{f' \in \mathcal{I}' : \rho(f') \in \mathcal{F}(\rho(F_x))} |\mathbf{g}(f', \mathcal{I}')| + \sum_{f \notin F_x} |\mathbf{g}(f, \mathcal{I}) - \mathbf{g}(\tilde{f}, \mathcal{I}')| \\ &\leq 3\bar{K} \mathbf{m}(F_x) + \sum_{f \notin F_x} \bar{K} \exp[-\bar{c} \mathbf{r}(f, \mathcal{I}; \tilde{f}, \mathcal{I}')] . \end{aligned}$$

It is clear by construction, that for every  $f, \tilde{f}$  the distance  $\mathbf{r}(f, \mathcal{I}; \tilde{f}, \mathcal{I}')$  is attained by the distance to a wall face. Since the distance between two faces is at least the distance between their projections, and projections of distinct walls are distinct,

$$\sum_{f \notin F_x} \bar{K} \exp[-\bar{c} \mathbf{r}(f, \mathcal{I}; \tilde{f}, \mathcal{I}')] \leq \sum_{f \notin F_x} \bar{K} \max_{u \in \rho(F_x)} \exp[-\bar{c} d(\rho(f), u)] .$$

Then by the definition of groups of walls and closeness of walls, for a ceiling face  $f$ ,  $N_\rho(\rho(f)) = 1$ , and for a wall face  $f \notin F_x$ ,  $N_\rho(\rho(f)) \leq |\rho(f) - \rho(g)|^2$  for all  $g \in F_x$ . Thus this is at most

$$\sum_{u \in \rho(F_x)^c} \bar{K} N_\rho(u) \max_{u' \in \rho(F_x)} \exp[-\bar{c} d(u, u')] \leq \sum_{u' \in \rho(F_x)} \sum_{u \in \rho(F_x)^c} \bar{K} (|u - u'|^2 + 1) \exp[-\bar{c} |u - u'|] .$$

which by integrability of exponential tails is easily seen to be at most  $\bar{C}(|\mathcal{E}(\rho(F_x))| + |\mathcal{F}(\rho(F_x))|)$  for some constant  $\bar{C}$ , which is in turn at most  $\bar{C} \mathbf{m}(F_x)$  by (2.1).  $\blacksquare$

It is also important for us to control the number of interfaces that get mapped to the same interface under application of the map  $\Phi_x$ . We begin with the following geometric observation.

**Observation 2.27** (e.g., Lemma 2 in [28]). *The number of  $\ast$ -connected collections of  $k$  faces in  $\mathbb{Z}^d$  containing a specified face  $f_\star$  is at most  $s^k$  for some universal (only lattice-dependent)  $s > 0$ .*

The following follows from Observation 2.27 and Definition 2.23; we do not include the proof here, but it can be found as part of the proof of the more complicated combinatorial estimate in Proposition 5.7.

**Lemma 2.28** ([28, Lemma 9]). *There exists  $s$  such that for any  $x \in \mathcal{L}_0 \cap \Lambda$ , the number of possible groups of walls  $F_x$  with excess area  $\mathbf{m}(F_x) = k$  is at most  $s^k$ . Likewise, there exists  $s'$  such that the number of possible groups of walls  $F$  containing  $x$  in their interior, with  $\mathbf{m}(F) = k$  is at most  $(s')^k$ .*

Together, Lemmas 2.25 and 2.28 imply an exponential tail on groups of walls. In various papers [12, 27, 33] proving rigidity for such models, they were used to show that the height of the interface above a face  $x = (x_1, x_2, 0) \in \mathcal{L}_0$ , defined there as  $\max\{h : (x_1, x_2, h) \in \mathcal{I}\}$ , has an exponential tail. Since that definition of height above  $x$  differs from the pillar-based perspective we take in the present paper, we modify the argument therein slightly to prove an exponential tail on the height of the pillar  $\text{ht}(\mathcal{P}_x)$ .

**Proof of Theorem 2.26.** We begin with the first estimate. Let  $\mathbf{I}_{F_x = \emptyset} = \text{Im}(\Phi_x)$  be the set of interfaces where the group of walls  $\mathcal{F}_x$  is empty. By Lemma 2.28 (and the definition of the map  $\Phi_x$  defined above, relying on Lemma 2.12), we see that for every  $\mathcal{I}' \in \mathbf{I}_{F_x = \emptyset}$ , the pre-image

$$\{\mathcal{I} \in \Phi_x^{-1}(\mathcal{I}') : \mathbf{m}(\mathcal{I}; \mathcal{I}') = k\}$$

has cardinality at most  $s^k$ . Then by Lemma 2.25, for every  $r \geq 1$ ,

$$\mu_n(\mathbf{m}(\mathcal{F}_x) \geq r) \leq \sum_{k \geq r} \sum_{\mathcal{I}' \in \mathbf{I}_{F_x = \emptyset}} \sum_{\mathcal{I} \in \Phi_x^{-1}(\mathcal{I}') : \mathbf{m}(\mathcal{I}; \Phi_x(\mathcal{I})) = k} \mu_n(\mathcal{I}) \leq \sum_{k \geq r} \sum_{\mathcal{I}' \in \mathbf{I}_{F_x = \emptyset}} \mu_n(\mathcal{I}') s^k \exp[-(\beta - C)k]$$

from which we obtain by summability of exponential tails, that for some  $C'$ , for  $\beta > \beta_0$ ,

$$\mu_n(\mathbf{m}(\mathcal{F}_x) \geq r) \leq C' \exp[-(\beta - C')r] \mu_n(\mathbf{I}_{F_x = \emptyset}) \leq C' \exp[-(\beta - C')r].$$

We now turn to bounding the probability of  $\text{ht}(\mathcal{P}_x) \geq h$ .

In order for  $\text{ht}(\mathcal{P}_x) \geq h$ , by Observation 2.17, there must be one sequence of nested walls  $\mathfrak{W}_x = (W_{x_s})_s$  all of which contain  $x$  in their interior, with  $\sum_s \mathbf{m}(W_{x_s}) = h_1$ , along with a sequence of nested walls  $\mathfrak{W}_y = (W_{y_t})_t$  with  $y_t \neq x_s$  for any  $t, s$ , containing some  $y$  in the interior ceilings of  $\mathfrak{W}_x$ , such that  $\sum_t \mathbf{m}(W_{y_t}) \geq 4h - h_1$ . In order to bound this, we can therefore write

$$\begin{aligned} \mu_n(\text{ht}(\mathcal{P}_x) \geq h) &\leq \mu_n(\mathbf{m}(\mathfrak{W}_x) \geq 4h) + \sum_{h_1 \leq 4h} \mu_n(\mathbf{m}(\mathfrak{W}_x) = h_1) \mu_n(\exists y : |y - x| \leq h_1, \mathbf{m}(\mathfrak{W}_y) \geq 4h - h_1 \mid \mathbf{m}(\mathfrak{W}_x) = h_1) \\ &\leq \mu_n(\mathbf{m}(\mathfrak{W}_x) \geq 4h) + \sum_{h_1 \leq 4h} \mu_n(\mathbf{m}(\mathfrak{W}_x) \geq h_1) \sum_{y : |y - x| \leq h_1} \sup_{(W_{x_s})_s : \mathbf{m}(\mathfrak{W}_x) = h_1} \mu_n(\mathbf{m}(\mathfrak{W}_y) \geq 4h - h_1 \mid (W_{x_s})_s). \end{aligned}$$

To bound the probabilities expressed above, let us turn to groups of walls instead of walls, denoting by  $\mathfrak{F}_x$  the group of walls of the nested sequence  $\mathfrak{W}_x$  and  $\mathfrak{F}_y$  corresponding to the nested sequence of walls of  $\mathfrak{W}_y$ . Following [12, 27], for a group of walls  $F_z$ , set

$$\phi_z^x(F_z) = \mathbf{m}(F_z) \mathbf{1}_{\{\mathbf{m}(F_z) \geq |z - x|\}}$$

and notice that  $\mathbf{m}(\mathfrak{W}_x) \leq \sum_z \phi_z^x(\mathcal{F}_z)$ . Indeed, every wall  $W_y \in \mathfrak{W}_x$  must nest  $x$  and therefore must have excess area at least  $d(y, x)$ , from which it follows that its group of walls  $F_y$  in turn has  $\mathbf{m}(F_y) \geq d(y, x)$ . By the tail estimates of part (1) of Theorem 2.26, we can bound

$$\mathbb{E}_{\mu_n}[e^{(\beta - 2C)\phi_z^x(\mathcal{F}_z)} \mid (\mathcal{F}_{z'})_{z' \neq z}] \leq 1 + Ce^{-C|z - x|}$$

(where  $\mathbb{E}_{\mu_n}$  is expectation with respect to  $\mu_n$ ) which implies (by iteratively revealing  $(F_z)$  for all  $z$ ) that

$$\mathbb{E}[e^{(\beta - 2C)\sum_z \phi_z^x(\mathcal{F}_z)}] \leq \prod_z (1 + Ce^{-C|z - x|}) \leq K < \infty.$$

By Markov's inequality, then,

$$\mu_n(\mathbf{m}(\mathfrak{F}_x) \geq r) \leq \mu_n\left(e^{(\beta - 2C)\sum \phi_z^x(F_z)} \geq e^{(\beta - 2C)r}\right) \leq Ke^{-(\beta - 2C)r}. \quad (2.4)$$

By the same reasoning, for any collection  $\mathfrak{F}_x = (F_{x_s})_s$ , since the exponential tail of Theorem 2.26 on  $F_y$  for  $y \neq x_s$  for any  $s$  holds conditionally on  $(F_{x_s})_s$ , we see that similarly,

$$\sup_{(F_{x_s})_s : \mathbf{m}(\mathfrak{F}_x) = h_1} \mu_n(\mathbf{m}(\mathfrak{F}_y) \geq r \mid (F_{x_s})_s) \leq Ke^{-(\beta - 2C)r}.$$

Using that  $\mathbf{m}(\mathfrak{W}_x) \leq \mathbf{m}(\mathfrak{F}_x)$  deterministically, we can plug in these estimates to see that

$$\begin{aligned} \mu_n(\text{ht}(\mathcal{P}_x) \geq h) &\leq Ke^{-4(\beta - 2C)h} + K \sum_{h_1 \leq 4h} e^{-\beta h_1 + 2Ch_1} h_1^2 e^{-(\beta - 2C)(4h - h_1)} \\ &\leq K \sum_{h_1 \leq 4h} h_1^2 e^{-4\beta h + 8Ch} \leq 16Kh^4 e^{-4\beta h + 8Ch} \end{aligned}$$

which for  $h$  large, is at most  $\exp[-4(\beta - C')h]$  for some other universal constant  $C'$ .  $\blacksquare$

By pairing the Dobrushin result with a straightforward forcing argument, we see the following.

**Proposition 2.29.** *There exist  $\beta_0, C > 0$  such that for every  $\beta > \beta_0$ , every  $n, h$  large and  $x \in \mathcal{L}_0 \cap \Lambda_n$ ,*

$$-4\beta - e^{-4\beta} \leq \frac{1}{h} \log \mu_n(\text{ht}(\mathcal{P}_x) \geq h) \leq -4\beta + C.$$

*Proof.* The upper bound here was given by the second part of Theorem 2.26. It remains to prove the lower bound; this proof will follow a more traditional coupling argument. First of all, with probability  $1 - \varepsilon_\beta$  for some  $\varepsilon_\beta$  vanishing as  $\beta \rightarrow \infty$ , we have that  $\text{ht}(\mathcal{P}_x) \geq 0$  using e.g., the reflected version of Theorem 2.26; (also notice that the event  $\{\text{ht}(\mathcal{P}_x) \geq 0\}$  is an increasing event).

Let  $\mathcal{P}_\emptyset$  be the set of all sites  $\{x + (0, 0, \ell - \frac{1}{2}) : \ell = 1, \dots, h\}$ . On the intersection of  $\text{ht}(\mathcal{P}_x) \geq 0$  with  $\sigma(\mathcal{P}_\emptyset) \equiv +1$ , the interface has  $\text{ht}(\mathcal{P}_x) \geq h$ , so that by the FKG inequality, it suffices to show the lower bound

$$\frac{1}{h} \log \mu_n(\sigma(\mathcal{P}_\emptyset) \equiv +1) > -4\beta h - e^{-4\beta} h.$$

In order to show this estimate, we can expose the spins of  $\mathcal{P}_\emptyset$  from bottom up, starting with the one at  $x + (0, 0, \frac{1}{2})$ . With probability at least  $\frac{1}{2}$ ,  $\sigma_{x-(0,0,\frac{1}{2})} = +1$ , and by monotonicity, at worst, all other spins in  $\sigma(\mathcal{P}_\emptyset^c)$  are minus; by the domain Markov property and an elementary calculation, the probability of the spin at  $x + (0, 0, \frac{1}{2})$  being plus is at least  $\frac{\exp(-4\beta)}{1+\exp(-4\beta)} = \frac{1}{2}(1 - \tanh(2\beta))$ . Continuing on to the next site in  $\mathcal{P}_\emptyset$ , conditional on the first one being plus, the same lower bound applies. As such, we can lower bound

$$\mu_n \left( \sigma(\mathcal{P}_\emptyset) = +1 \mid \sigma_{x-(0,0,\frac{1}{2})} = +1 \right) \geq \frac{e^{-4\beta h}}{(1 + e^{-4\beta})^h} > e^{-4\beta h - e^{-4\beta} h},$$

concluding the proof as long as  $h$  is sufficiently large. ■

### 3. INCREMENTS AND THE SHAPE OF TALL PILLARS

In this section, we give a structural decomposition of a pillar, in the large deviation regime where it reaches a height of  $h$ . We prove that it is composed of a base—shown in §5 to have an exponential tail beyond height  $O(\log h)$ —and a spine protruding from this base up to a height of  $h$ . This spine is further decomposed into a sequence of *increments* between *cut-points* where the spine is one-dimensional and vertical. In the remainder of this section, we give preliminary bounds regarding this decomposition, showing that the total number of increments is comparable to  $h$ , and has an exponential tail beyond that. In the following §4, we analyze individual increments, showing that they each have an exponential tail on their excess area.

**3.1. Increments of the pillar.** We begin by defining the building blocks of the pillar where the 3D Ising interface undergoes an atypical fluctuation.

**Definition 3.1** (Cut-points). Call a height  $h \in \mathbb{Z} + \frac{1}{2}$  a *cut-height* of the pillar  $\mathcal{P}_x$  if the intersection of the slab  $\mathcal{L}_h$  with  $\sigma(\mathcal{P}_x)$  consists of exactly one (midpoint of a) cell. We can call that single plus site  $v \in \sigma(\mathcal{P}_x)$  a *cut-point* and identify it with its midpoint.

**Definition 3.2** (Increments of the pillar). For a pillar  $\mathcal{P}$ , we define its increment collection  $(\mathcal{X}_i)_i$  as follows. Enumerate the cut-points of  $\mathcal{P}$  as  $v_1, v_2, \dots, v_{\mathcal{T}}, v_{\mathcal{T}+1}$  in order of increasing height, for some  $\mathcal{T}$ . The  $k$ -th increment of the pillar  $\mathcal{P}$  is the set of all plus sites in  $\sigma(\mathcal{P})$  centered at heights between  $\text{ht}(v_k)$  and  $\text{ht}(v_{k+1})$ , inclusively (this is also identified with the bounding sets of faces in  $\mathcal{P} \cap \mathbb{R}^2 \times ([\text{ht}(v_k)], [\text{ht}(v_{k+1})])$ , as before). Denote by  $\mathbf{I}_{x,T}$  the set of interfaces which have  $\mathcal{T} \geq T$ .

Since the pillar does not necessarily end at a cut-point, there may be a remainder of plus sites in the pillar above the height  $\text{ht}(v_{\mathcal{T}+1})$ . We can call this the *remainder* and denote it by  $\mathcal{X}_{>\mathcal{T}}$ ; in fact for any  $t \leq \mathcal{T}$ , we could denote the remainder beyond the  $t$ -th increment  $\mathcal{X}_{>t}$  which consists of  $(\mathcal{X}_{t+1}, \dots, \mathcal{X}_{\mathcal{T}}, \mathcal{X}_{>\mathcal{T}})$ .

**3.2. Comparability of height and number of increments.** In this section, we show that the number of increments (as defined in the preceding subsection) serves as a good proxy for the height of a pillar. We remark that the converse part of the next lemma would have readily followed had we had an exponential tail for  $\mathbf{m}(\mathcal{P}_x)$  (when added to Proposition 2.29)—however, this is false, since  $\mathcal{P}_x$  may contain a wall with surface area  $r^2$  and  $\varepsilon r^2$  nested thermal fluctuations (resulting in  $\mathbf{m}(\mathcal{P}_x) \geq cr^2$ ) at a cost of only  $\exp(-cr)$ .

**Lemma 3.3.** *One always has  $\mathbf{I}_{x,k} \subset \{\text{ht}(\mathcal{P}_x) \geq k+1\}$  for every  $k$ . Conversely, there exist absolute constants  $C, c > 0$  such that, if  $\beta > \beta_0$  and  $T = \lfloor (1 - C/\beta)h \rfloor$  then*

$$\mu_n(\mathbf{I}_{x,T} \mid \text{ht}(\mathcal{P}_x) \geq h) \geq 1 - O(e^{-ch}).$$

*Proof.* The first assertion follows from the fact that, by definition, each increment increases the height of the pillar by at least 1 and the extremal increment contributes two to the height.

The lower bound is substantially more involved, and requires the use of a map that replaces a pillar of height  $h$  and fewer than  $\lfloor(1 - C/\beta)h\rfloor$  increments, by a straight column of height  $h$  (consisting of  $h - 1$  total increments). This will combine the proof of Lemma 2.25 with some new ideas that will serve as a warm-up for the more sophisticated maps on pillars used in Section 4 and especially Section 5.

Let  $\Phi_{x,h}$  be the map that takes an interface  $\mathcal{I}$  and generates an interface  $\mathcal{J}$  as follows:

- (1) Let  $(W_z)_{z \in \mathcal{L}_0}$  be the standard wall representation of  $\mathcal{I}$  per Lemma 2.12.
- (2) If  $[x] := \{x\} \cup \bigcup_{f \in \mathcal{L}_0: f \sim x} \{f\}$  delete from the collection  $(W_z)_{z \in \mathcal{L}_0}$ ,  $\mathfrak{F}_{[x]} := \bigcup_{f \in [x]} \mathfrak{F}_f$  as well as  $\mathfrak{F}_{\rho(v_1)}$ .
- (3) If the interface  $\mathcal{I}'$  whose standard wall representation equals  $\mathfrak{F}_{[x]} \cup \mathfrak{F}_{\rho(v_1)}$  has a cut-height below  $\text{ht}(v_1)$ , let  $h^\dagger$  be the highest such cut-height and let  $y^\dagger$  be the index of a wall of  $\mathcal{P}_x$  that attains height  $h^\dagger$  and is not included in  $\mathfrak{F}_{[x]} \cup \mathfrak{F}_{\rho(v_1)}$ . Delete  $\mathfrak{F}_{y^\dagger}$  from the standard wall representation obtained after step (2). (The existence of such a  $y^\dagger$  is guaranteed by the definition of  $v_1$  as the lowest cut-point.)
- (4) Add to this standard wall representation the bounding vertical faces of a straight column of  $h$  cells above  $x$ , centered at  $x + (0, 0, \ell - \frac{1}{2}) : \ell = 1, \dots, h$ .
- (5) Let  $\mathcal{J}$  be the interface with the standard wall representation resulting from step (4) as per Lemma 2.12.

The map is well-defined because after step (2), there are no walls incident to  $x$  nor its bounding edges and the addition of the standard wall in step (4) maintains the admissibility of the standard wall collection. The resulting  $\mathcal{J}$  therefore has a pillar  $\mathcal{P}_x^{\mathcal{J}}$  consisting of exactly a column of  $h - 1$  increments, attaining height  $h$ .

We next claim that if  $\mathcal{I}$  is such that  $\{\text{ht}(\mathcal{P}_x) \geq h\}$  but  $\mathcal{S} < (1 - \delta)h$  (for  $\delta$  to be chosen later), then

$$\mathbf{m}(\mathcal{I}; \mathcal{J}) = |\mathfrak{F}_{[x]} \cup \mathfrak{F}_{\rho(v_1)} \cup \mathfrak{F}_{y^\dagger}| - 4h \geq 2\delta h. \quad (3.1)$$

By Observation 2.17, the entirety of the pillar above  $\text{ht}(v_1)$  is deleted and therefore, for each height between  $\text{ht}(v_1)$  and  $h$  that is not a cut-height, there is an excess area contribution of 2 faces (due to 6 faces bounding two cells vs. 4 faces bounding one cell), totaling to  $2(h - \frac{1}{2} - \text{ht}(v_1) - \mathcal{S})$ . For heights between 0 and  $\text{ht}(v_1)$ , we claim that the interface having standard wall representation  $\mathfrak{F}_{[x]} \cup \mathfrak{F}_{\rho(v_1)} \cup \mathfrak{F}_{y^\dagger}$  has no cut-heights, in which case it would follow that those heights together contribute at least  $2(\text{ht}(v_1) - \frac{1}{2})$  to the excess area and (3.1) would follow. Indeed, if no  $y^\dagger$  is chosen in step (3), then by definition there were no cut-heights of the interface corresponding to  $\mathfrak{F}_{[x]} \cup \mathfrak{F}_{\rho(v_1)}$  below  $\text{ht}(v_1)$ , so suppose there was a highest such cut-height at  $h^\dagger$  and a corresponding  $y^\dagger$  was selected (noting that then  $W_{y^\dagger}$  must be distinct from  $W_{\rho(v_1)}$ ). Then since the walls  $W_{y^\dagger}$  and  $W_{\rho(v_1)}$  must each attain the height  $h^\dagger$ , there can be no cut-heights at or below  $h^\dagger$  in the interface corresponding to  $\mathfrak{W}_{\rho(v_1)} \cup \mathfrak{W}_{y^\dagger}$  and therefore there also cannot be any at or below  $h^\dagger$  in the interface corresponding to  $\mathfrak{F}_{[x]} \cup \mathfrak{F}_{\rho(v_1)} \cup \mathfrak{F}_{y^\dagger}$ .

Having constructed the map  $\Phi_{x,h}$ , the proof now proceeds in two parts: (1) we show that the relative weight  $\mu_n(\mathcal{I})/\mu_n(\mathcal{J})$  is exponentially decaying in  $(\beta - C)\mathbf{m}(\mathcal{I}; \mathcal{J})$  and (2) we show that the multiplicity of the map  $\Phi_{x,h}$  is at most exponentially growing in  $\mathbf{m}(\mathcal{I}; \mathcal{J})$ .

To begin with the first, it suffices for us to show the bound

$$\left| \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \mathcal{J}} \mathbf{g}(f', \mathcal{J}) \right| \leq C[\mathbf{m}(\mathcal{I}; \mathcal{J}) + h] \leq C[1 + \delta^{-1}]h. \quad (3.2)$$

To establish such a bound, we decompose  $\mathcal{I}$  and  $\mathcal{J}$  into different subsets of faces as follows:

- Let  $\mathbf{E}$  be the set of all faces in the groups of walls  $\mathfrak{F}_{[x]} \cup \mathfrak{F}_{\rho(v_1)} \cup \mathfrak{F}_{y^\dagger}$ .
- Let  $\mathbf{F}$  be the set of  $f \in \mathcal{J}$  such that  $\rho(f) \in \rho(\mathcal{F}(\mathbf{E}))$ , added in place of a removed horizontal wall face in  $\rho(\mathbf{E})$  to “fill in” the interface.
- Let  $\mathbf{G}$  be the bounding vertical faces of a column of  $h$  cells above  $x$ , added in step (4) of  $\Phi_{x,h}$ .

Under this decomposition, there is a 1-1 correspondence between  $\mathcal{I} \setminus \mathbf{E}$  and  $\mathcal{J} \setminus (\mathbf{F} \cup \mathbf{G})$  via vertical shifts as determined by Observation 2.13; encode this into  $f \mapsto \tilde{f}$ . Then,

$$\left| \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \mathcal{J}} \mathbf{g}(f', \mathcal{J}) \right| \leq \sum_{f \in \mathbf{E}} |\mathbf{g}(f, \mathcal{I})| + \sum_{f' \in \mathbf{F}} |\mathbf{g}(f', \mathcal{J})| + \sum_{f' \in \mathbf{G}} |\mathbf{g}(f', \mathcal{J})| + \sum_{f \in \mathcal{I} \setminus \mathbf{E}} \left| \mathbf{g}(f, \mathcal{I}) - \mathbf{g}(\tilde{f}, \mathcal{J}) \right|.$$

The first quantity is bounded by  $\bar{K}|\mathbf{E}| = \bar{K}|\mathfrak{F}_{[x]} \cup \mathfrak{F}_{\rho(v_1)} \cup \mathfrak{F}_{y^\dagger}| \leq 2\bar{K}(1 + 2\delta^{-1})\mathbf{m}(\mathcal{I}; \mathcal{J})$  by (2.1) and (3.1). Similarly, the second term is at most  $\bar{K}(1 + 2\delta^{-1})\mathbf{m}(\mathcal{I}; \mathcal{J})$  and the third term is at most  $4\bar{K}h$  which is in turn at most  $2\bar{K}\delta^{-1}\mathbf{m}(\mathcal{I}; \mathcal{J})$ . The last term is bounded similarly to the proof of Lemma 2.25. By construction,

for every  $f$ , the radius  $\mathbf{r}(f, \mathcal{I}; \tilde{f}, \mathcal{J})$  is attained by the distance to a wall face, and as before, moving to the distance between projections,

$$\sum_{f \in \mathcal{I} \setminus \mathbf{E}} \bar{K} \exp[-\bar{c}\mathbf{r}(f, \mathcal{I}; \tilde{f}, \mathcal{J})] \leq \sum_{f \in \mathcal{I} \setminus \mathbf{E}} \bar{K} \max_{u \in \rho(\mathbf{E} \cup \mathbf{G})} \exp[-\bar{c}d(\rho(f), u)]$$

and using the definition of closeness of walls, there is a  $\bar{C} > 0$  such that this is at most

$$\sum_{u' \in \rho(\mathbf{E} \cup \mathbf{G})} \sum_{u \in \rho(\mathbf{E})^c} \bar{K} (|u - u'|^2 + 1) \exp[-\bar{c}|u - u'|] \leq \bar{C} \bar{K} |\mathbf{E} \cup \mathbf{G}|,$$

yielding (3.2) by applying (2.1) and (3.1) as above.

We next wish to bound the multiplicity of the map, i.e., we wish to show that for every  $M$  and every  $\mathcal{J}$  in the image of  $\Phi_{x,h}$ , the size of the set  $\{\mathcal{I} \in \Phi_{x,h}^{-1}(\mathcal{J}) : \mathbf{m}(\mathcal{I}; \mathcal{J}) = M\}$ . Every such  $\mathcal{I}$  can be identified with the choices of the standard walls in  $\mathfrak{F}_{[x]} \cup \mathfrak{F}_{\rho(v_1)} \cup \mathfrak{F}_{y^\dagger}$ , so that it suffices to bound the number of possible such choices leading to excess area  $M$ . By iteratively applying Lemma 2.28 (first choosing how many walls constitute  $\mathfrak{F}_{[x]}$ , and then the size of each such that the total size is at most  $2M$ ), then enumerating over the at most  $M^2$  choices of where to place  $\rho(v_1)$  and  $\rho(y^\dagger)$  (using Observation 2.17 these must be interior to some wall of  $\mathfrak{W}_x$ ), and then enumerating over the choices for  $\mathfrak{F}_{\rho(v_1)}$  and  $\mathfrak{F}_{y^\dagger}$ , we see that for some universal  $\bar{s} > 0$ ,

$$\{\mathcal{I} \in \Phi_{x,h}^{-1}(\mathcal{J}) : \mathbf{m}(\mathcal{I}; \mathcal{J}) = M\} \leq \bar{s}^M.$$

(See e.g., the proof of Proposition 5.7 for more details on a similar enumeration process.)

We now combine the two parts above to conclude the desired. Expressing  $\mu_n(\mathbf{I}_{x,(1-\delta)h}^c, \text{ht}(\mathcal{P}_x) \geq h)$  as

$$\begin{aligned} \sum_{\mathcal{I} \in \mathbf{I}_{x,(1-\delta)h}^c, \text{ht}(\mathcal{P}_x) \geq h} \mu_n(\mathcal{I}) &\leq \sum_{M \geq 2\delta h} \sum_{\mathcal{J} : \text{ht}(\mathcal{P}_x^{\mathcal{J}}) \geq h} \sum_{\mathcal{I} \in \Phi_{x,h}^{-1}(\mathcal{J}) : \mathbf{m}(\mathcal{I}; \mathcal{J}) = M} \mu_n(\mathcal{J}) e^{-\beta M + C(1+\delta^{-1})M} \\ &\leq \sum_{M \geq 2\delta h} \bar{s}^M e^{-\beta M + C(1+\delta^{-1})M + M \log \bar{s}} \mu_n(\text{ht}(\mathcal{P}_x) \geq h), \end{aligned}$$

from which the lemma follows by dividing through by  $\mu_n(\text{ht}(\mathcal{P}_x) \geq h)$  and taking  $\beta$  large and  $\delta = C'/\beta$  for some sufficiently large  $C'$ .  $\blacksquare$

**3.3. Spine and base of pillars.** The fundamental difficulty in understanding the structure of pillars conditionally on reaching a height  $h$ , or on having  $T$  increments, is the interactions of the pillar with nearby oscillations of the interface, particularly at low heights, where these are plentiful. Towards this, it will be important to us to isolate the portion of the pillar which interacts most strongly with other pillars near it—called the *base*—and the rest of the pillar, which climbs above all oscillations in some ball of radius  $O(h \vee T)$  about  $x$ , called the *spine*. The ball of proximity grows with the number of increments  $T$  we are conditioning on having, as the pillar's  $(xy)$ -coordinates diffuse as  $T$  grows; this creates the complication that the definitions of the spine and base must be  $T$  dependent.

For a set  $A \subset \mathcal{F}(\mathbb{Z}^3)$ , let  $\mathcal{C}_r(A)$  denote the points in  $\mathbb{R}^3$  whose projection is distance at most  $r$  from  $\rho(A)$ :

$$\mathcal{C}_r(A) = \{y : \min_{f \in A} |\rho(y) - \rho(f)| \leq r\}.$$

Let  $R_0$  be some sufficiently large constant, e.g., to be chosen in Lemma 3.15 to be 100. For two faces  $x, y \in \mathcal{L}_0$ , let  $\langle\langle x, y \rangle\rangle$  be a minimal connected set of faces of  $\mathcal{L}_0$  connecting  $x$  to  $y$ . For ease of notation, for a cut-point  $v$ , we'll define

$$\mathcal{C}_{v,x,T} := \mathcal{C}_{R_0 T}(\langle\langle \rho(v), x \rangle\rangle).$$

**Definition 3.4 (Spine).** Consider an interface  $\mathcal{I}$  with pillar  $\mathcal{P}_x = \mathcal{P}_x(\mathcal{I})$ . For each  $T$ , let  $\tau_{\text{SP}} = \tau_{\text{SP}}(T)$  be the minimal index  $i \geq 1$  such that the cut-point  $v_i$  of  $\mathcal{P}_x$  lies above the largest height attained by walls in  $\mathcal{I} \setminus \mathcal{P}_x$  indexed by faces in  $\mathcal{C}_{v_i, x, T}$  (in every possible ordering of  $\mathcal{F}(\mathcal{L}_0)$ ). We then call  $v_{\tau_{\text{SP}}}^T$  the  $T$ -*source-point*. When  $T$  is understood from the context (e.g., for  $\mathcal{I} \in \mathbf{I}_{x,T}$ ) we drop it from the notation and write  $v_{\tau_{\text{SP}}}$ . With respect to that  $T$  the *spine*  $\mathcal{S}_x$  will then be the  $*$ -connected component of  $\mathcal{P}_x$  consisting of all sites/faces above  $\text{ht}(v_{\tau_{\text{SP}}}) - \frac{1}{2}$ , i.e., consisting of the increments  $\mathcal{X}_{\tau_{\text{SP}}}, \dots, \mathcal{X}_{\mathcal{J}}, \mathcal{X}_{>\mathcal{J}}$ .

**Definition 3.5 (Base).** For an interface  $\mathcal{I}$  in  $\mathbf{I}_{x,T}$  with pillar  $\mathcal{P}_x$ , let the *base*  $\mathcal{B}_x$  of the pillar be given by the entirety of the pillar below the height  $\text{ht}(v_{\tau_{\text{SP}}}) + \frac{1}{2}$ . In general, a base can be identified with the set-difference  $\mathcal{P}_x \setminus \mathcal{S}_x$ , along with the four bounding faces of the  $T$ -source-point  $v_{\tau_{\text{SP}}}$ .

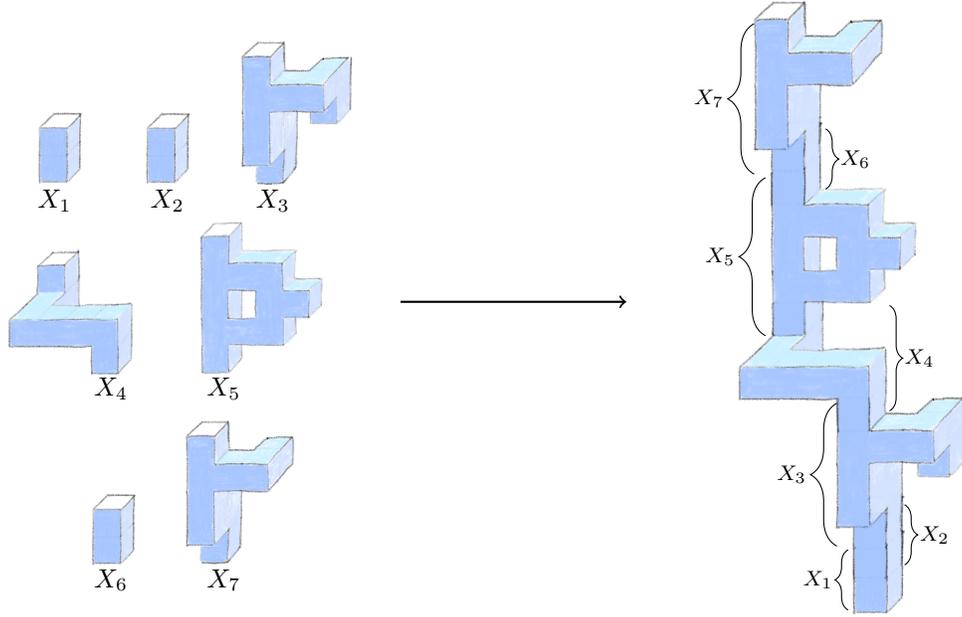


FIGURE 8. A stretch of the spine (right) decomposes into a collection of rooted increments which are members of  $\mathfrak{X}$  (left). The increments  $X_1, X_2$  and  $X_6$  are trivial increments  $X_\emptyset$ .

(For reference, in Figures 2 and 10, the spine is shaded in blue, and the base is shaded in pink and orange.) We will show in Section 5 that, with high probability, most of the increments of a pillar belong to the spine.

**3.4. Properties of increments.** Let  $\mathfrak{X}$  be the set of all possible rooted increments, where an increment is identified with a  $*$ -connected subset of plus sites in the upper half-space of  $\mathbb{Z}^3$  consisting of a cut-point plus site at  $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ , as well as a cut-point plus site at its largest height, and such that no height in between these is a cut-height. As usual, we also identify such increments with a  $*$ -connected collection of faces that bound its plus sites; however, in this face set  $\mathcal{F}(X)$  we exclude the bottom-most and top-most delimiting faces (since, viewing this increment as a subset of an interface, those faces would not be present in the interface).

Let  $\mathfrak{X}_{\text{REM}}$  be the set of rooted remainders i.e.,  $*$ -connected subsets of plus sites in the upper half-space of  $\mathbb{Z}^3$  where we only impose that they have exactly one cell in the slab  $\mathcal{L}_{\frac{1}{2}}$  at  $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ . Correspondingly, its face set is the set of faces that bound it, now excluding only the bottom-most delimiting face (at  $(\frac{1}{2}, \frac{1}{2}, 0)$ ).

**Definition 3.6.** For each increment  $\mathcal{X}_i$  in the pillar  $\mathcal{P}_x$ , recall that its bottom-most and top-most cells are  $v_i$  and  $v_{i+1}$  respectively. We define the height of an increment  $X_i \in \mathfrak{X}$  by  $\text{ht}(X_i) = \text{ht}(v_{i+1}) - \text{ht}(v_i)$ .

**Lemma 3.7.** *There is a 1-1 correspondence between the triplets of  $v_{\tau_{\text{sp}}}$ , the collection of  $T$  rooted increments  $(X_i)_{i \leq T}$  for  $X_i \in \mathfrak{X}$ , and remainder increment  $X_{>T} \in \mathfrak{X}_{\text{REM}}$ , and the set of spines of at least  $T$  increments.*

*Proof.* Identifying the increment sequence given a spine was described by the definition of increments. Obtaining from this increment sequence, the rooted increments, consists only of shifting each by the vector  $-v_i + (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ ; the rooted remainder is similarly recovered.

Given a sequence of  $T$  rooted increments, a source point  $v_{\tau_{\text{sp}}}$  and a remainder  $X_{>T}$ , we can reconstruct the cell-set of the spine by taking the union over  $i$  of the translates of  $X_i$  by the vectors  $-(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}) + v_j$  where  $v_j$  are defined inductively as increments are stacked. (Naturally, the rooted remainder  $X_{>T}$  is shifted by  $-(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}) + v_{T+1}$ .) As a consequence, we can identify the set of all rooted spines of at least  $T$  increments with the set  $\mathfrak{X}^T \times \mathfrak{X}_{\text{REM}}$ . See Figure 8 for a visualization of this scheme.  $\blacksquare$

We will always use the notation  $X_\emptyset$  to denote the trivial increment that consists of exactly two plus cells, one on top of the other (the rooted one has the plus sites centered at  $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$  and  $(\frac{1}{2}, \frac{1}{2}, \frac{3}{2})$ ). The trivial remainder increment consists of exactly one plus cell at  $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ .

**Definition 3.8.** The excess area of an increment  $X \in \mathfrak{X}$  is given by its excess area as compared to the trivial increment  $X_\emptyset$  so that

$$\mathbf{m}(X) := |\mathcal{F}(X)| - |\mathcal{F}(X_\emptyset)| = |\mathcal{F}(X)| - 8.$$

The excess area of a remainder  $X_{>T}$  is measured with respect to the trivial remainder increment, so that  $\mathbf{m}(X_{>T}) := |\mathcal{F}(X_{>T})| - 5$  (recall that the remainder increment includes its upper delimiting face(s)).

**Remark 3.9.** Notice that  $\mathbf{m}(\mathcal{X}_i) \geq \sqrt{2}|\rho(v_{i+1}) - \rho(v_i)|$  (the nontrivial increment  $X$  of height 2 consisting of two \*-adjacent cubes has  $|\mathcal{F}(X)| = 10$  and  $\mathbf{m}(X) = 2$ ) and for every  $X \neq X_\emptyset$ ,

$$\mathbf{m}(X) \geq \frac{1}{5}|\mathcal{F}(X)| \quad \text{and} \quad |\mathcal{F}(X)| \geq 6(\text{ht}(X) - 1) + 8.$$

**Definition 3.10.** For a spine  $\mathcal{S}_x$  of  $\mathcal{T} - \tau_{\text{sp}}$  increments  $(\mathcal{X}_i)_{\tau_{\text{sp}} \leq i \leq \mathcal{T}}$  and remainder  $\mathcal{X}_{>\mathcal{T}}$ , the excess area of the spine  $\mathbf{m}(\mathcal{S}_x)$  with respect to a trivial increment sequence of height  $t - \tau_{\text{sp}}$ ,

$$\mathbf{m}_t(\mathcal{S}_x) = \mathbf{m}(\mathcal{I}; \mathcal{I}_{\emptyset, t}) = \mathbf{m}(\mathcal{X}_{>t}) + \sum_{\tau_{\text{sp}} \leq i \leq t} \mathbf{m}(\mathcal{X}_i),$$

and if  $t > \mathcal{T}$ , we set  $\mathbf{m}_t(\mathcal{S}_x) = \mathbf{m}(\mathcal{S}_x)$ . The excess area of the spine (dropping the index  $t$ ) is  $\mathbf{m}(\mathcal{S}_x) := \mathbf{m}_{\mathcal{T}}(\mathcal{S}_x)$ .

The height of a spine  $\mathcal{S}_x$  is  $\text{ht}(\mathcal{S}_x) = \text{ht}(\mathcal{P}_x) - \frac{1}{2} - \text{ht}(v_{\tau_{\text{sp}}}) = \text{ht}(\mathcal{X}_{>\mathcal{T}}) + \sum_{\tau_{\text{sp}} \leq i \leq \mathcal{T}} \text{ht}(\mathcal{X}_i)$ .

**Definition 3.11.** For any interface  $\mathcal{I} \in \mathbf{I}_{x, T}$ , let  $\mathcal{I}_{\text{TR}}$  be its *truncation*, with cell-set  $\sigma(\mathcal{I}_{\text{TR}}) := (\sigma(\mathcal{I}) \setminus \sigma(\mathcal{S}_x)) \cup \{v_{\tau_{\text{sp}}}\}$  where we have removed all plus sites of the spine besides  $v_{\tau_{\text{sp}}}$  from  $\mathcal{I}$ , and face-set consisting of the faces in  $\mathcal{I}$  that bound cells in  $\sigma(\mathcal{I}_{\text{TR}})$ . A truncation is  $T$ -admissible if its pillar  $\mathcal{P}_x(\mathcal{I}_{\text{TR}})$  has a  $T$ -source point  $v_{\tau_{\text{sp}}}$  and nothing above  $\text{ht}(v_{\tau_{\text{sp}}}) + \frac{1}{2}$ . (Recall that the property of being a  $T$ -source point is independent of the increment sequence of the spine above it).

**3.5. Exponential tail on the number of increments.** Here, we show that a spine  $\mathcal{S}_x$  of an interface in  $\mathbf{I}_{x, T}$  has an exponential tail on the surface area (as well as excess area) of its remainder  $\mathcal{X}_{>T}$ . This implies an exponential tail on the number of increments beyond  $T$  in a spine conditioned on having at least  $T$  increments. Since we are only looking at a portion of the increment above a cut-point, it is droplet-like, and the proof does not involve any of the more delicate issues we will encounter in later sections.

**Lemma 3.12.** *There exists  $C > 0$  such that for every  $\beta > \beta_0$ , every  $T$  and every  $r > 0$ ,*

$$\mu_n(\mathbf{m}(\mathcal{X}_{>T}) \geq r \mid \mathbf{I}_{x, T}) \leq \exp[-(\beta - C)r].$$

*In particular,  $\mu_n(\mathbf{I}_{x, T+k} \mid \mathbf{I}_{x, T}) \leq \exp[-4k(\beta - C)]$ . Moreover, these estimates also hold conditionally on any  $T$ -admissible truncation  $\mathcal{I}_{\text{TR}}$  and spine increment sequence  $(\mathcal{X}_i)_{\tau_{\text{sp}} \leq i \leq T} = (X_i)_{\tau_{\text{sp}} \leq i \leq T}$ .*

*Proof.* Let  $\Phi_T : \mathbf{I}_{x, T} \mapsto \mathbf{I}_{x, T}$  be the map that, for each  $\mathcal{I} \in \mathbf{I}_{x, T}$ , generates the interface  $\Phi_T(\mathcal{I})$  by replacing  $\mathcal{X}_{>T}$  with the trivial remainder, and agrees with  $\mathcal{I}$  otherwise. It should be clear that  $\Phi_T(\mathcal{I}) \in \mathbf{I}_{x, T} \setminus \mathbf{I}_{x, T+1}$ ; moreover, the pillar of  $\Phi_T(\mathcal{I})$  will have height equal to  $\text{ht}(v_{T+1}) + \frac{1}{2}$ . By Theorem 2.21, for any  $\mathcal{I} \in \mathbf{I}_{x, T}$ ,

$$\frac{\mu_n(\mathcal{I})}{\mu_n(\Phi_T(\mathcal{I}))} = \exp\left(-\beta \mathbf{m}(\mathcal{I}; \Phi_T(\mathcal{I})) + \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \Phi_T(\mathcal{I})} \mathbf{g}(f', \Phi_T(\mathcal{I}))\right).$$

By definition of excess areas of remainders,  $\mathbf{m}(\mathcal{I}; \Phi_T(\mathcal{I})) = \mathbf{m}(\mathcal{X}_{>T})$ . Suppose without loss of generality that  $\mathcal{I}$  has remainder  $X_{>T}$  such that  $\mathbf{m}(X_{>T}) \geq 1$  as the lemma is trivially satisfied for  $r = 0$ . For ease of notation, let  $\mathcal{I}' = \Phi_T(\mathcal{I})$  and consider the difference of the sums in the exponential. By (2.2)–(2.3),

$$\begin{aligned} \left| \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \mathcal{I}'} \mathbf{g}(f', \mathcal{I}') \right| &\leq \sum_{f \in \mathcal{I} \cap \mathcal{I}'} |\mathbf{g}(f, \mathcal{I}) - \mathbf{g}(f, \mathcal{I}')| + \sum_{f \in \mathcal{I} \setminus \mathcal{I}'} |\mathbf{g}(f, \mathcal{I})| + \sum_{f \in \mathcal{I}' \setminus \mathcal{I}} |\mathbf{g}(f, \mathcal{I}')| \\ &\leq \sum_{f \in \mathcal{I} \cap \mathcal{I}'} \sum_{f' \in \mathcal{I} \oplus \mathcal{I}'} \bar{K} \exp[-\bar{c}d(f, f')] + \sum_{f \in \mathcal{I} \oplus \mathcal{I}'} \bar{K} \leq (\bar{C} + \bar{K})|\mathcal{I} \oplus \mathcal{I}'|, \end{aligned}$$

for some constant  $\bar{C}$ . But by construction, we have that  $|\mathcal{I} \oplus \mathcal{I}'| = \mathbf{m}(X_{>T}) + 2$ , where the additive 2 comes from the upper-bounding face of the remainder, which is shifted between  $\mathcal{I}$  and  $\mathcal{I}'$ . Consequently, we have that for some universal  $C$  independent of  $\beta$ , for every  $\mathcal{I} \in \mathbf{I}_{x, T}$ ,

$$\frac{\mu_n(\mathcal{I})}{\mu_n(\mathcal{I}')} \leq \exp[-(\beta - C)\mathbf{m}(\mathcal{I}; \mathcal{I}')] = \exp[-(\beta - C)\mathbf{m}(X_{>T})].$$

At the same time, we claim that for every  $\mathcal{I}' \in \mathbf{I}_{x,T} \setminus \mathbf{I}_{x,T+1}$ , there are at most  $s^k$  elements in the pre-image  $\Phi_T^{-1}(\mathcal{I}')$  with excess area  $k$ , for some universal  $s > 0$ . Since  $\mathfrak{m}(\mathcal{I}; \mathcal{I}') = |\mathcal{I} \oplus \mathcal{I}'| - 2$ , to every  $\mathcal{I} \in \Phi_T^{-1}(\mathcal{I}')$  of excess area  $\mathfrak{m}(\mathcal{I}; \mathcal{I}') = k$ , we can uniquely identify the connected set of faces constituting  $\mathcal{I} \oplus \mathcal{I}'$  of cardinality  $k + 2$  containing the upper bounding face of  $v_{T+1}$ . By Observation 2.27, the number of such sets is at most  $s^{k+2}$ . We now can expand the probability  $\mu_n(\mathfrak{m}(X_{>T}) \geq r, \mathbf{I}_{x,T})$

$$\sum_{\mathcal{I} \in \mathbf{I}_{x,T} : \mathfrak{m}(X_{>T}) \geq r} \mu_n(\mathcal{I}) = \sum_{k \geq r} \sum_{\mathcal{I} \in \mathbf{I}_{x,T} : \mathfrak{m}(X_{>T}) = k} \mu_n(\mathcal{I}) \leq \sum_{k \geq r} \sum_{\mathcal{I}' \in \Phi_T(\mathbf{I}_{x,T})} e^{-(\beta-C)k} s^{k+2} \mu_n(\mathcal{I}'),$$

At this point, since  $\Phi_T(\mathbf{I}_{x,T}) \subset \mathbf{I}_{x,T}$ , we see that for  $\beta > \beta_0$ , this is at most

$$\sum_{\mathcal{I}' \in \Phi_T(\mathbf{I}_{x,T})} C' e^{-(\beta-C')r} \mu_n(\mathcal{I}') \leq C' e^{-(\beta-C')r} \mu_n(\mathbf{I}_{x,T}),$$

for some other constant  $C' > 0$  independent of  $\beta$ ; dividing both sides by  $\mu_n(\mathbf{I}_{x,T})$  implies the first inequality. The second inequality follows because  $\mathcal{I} \in \mathbf{I}_{x,T+1}$  implies that  $\mathfrak{m}(\mathcal{I}; \Phi_T(\mathcal{I})) = \mathfrak{m}(X_{>T}) \geq 4$ .

To see the analogous conditional estimates, fix a  $T$ -admissible truncation  $\mathcal{I}_{\text{TR}}$  and first  $T$  increments of the spine  $(X_i)_{i \leq T}$ , and let  $\hat{\mathbf{I}}_{x,T}$  be the set of interfaces in  $\mathbf{I}_{x,T}$  having  $\mathcal{I}_{\text{TR}}$  and  $(\mathcal{X}_i)_{i \leq T} = (X_i)_{i \leq T}$ . Repeating the argument above, we see that  $\mu_n(\mathfrak{m}(X_{>T}) \geq r, \hat{\mathbf{I}}_{x,T})$  can be expressed as

$$\sum_{\mathcal{I} \in \mathbf{I}_{x,T} : \mathfrak{m}(X_{>T}) \geq r} \mu_n(\mathcal{I}) = \sum_{k \geq r} \sum_{\mathcal{I} \in \mathbf{I}_{x,T} : \mathfrak{m}(X_{>T}) = k} \mu_n(\mathcal{I}) \leq \sum_{k \geq r} \sum_{\mathcal{I}' \in \Phi_T(\mathbf{I}_{x,T})} e^{-(\beta-C)k} s^{k+2} \mu_n(\mathcal{I}').$$

Observing that  $\Phi_T(\hat{\mathbf{I}}_{x,T}) \subset \hat{\mathbf{I}}_{x,T}$ , we see that the right-hand side is at most  $C' e^{-(\beta-C')r} \mu_n(\hat{\mathbf{I}}_{x,T})$  and dividing through by  $\mu_n(\hat{\mathbf{I}}_{x,T})$  yields the desired conditional estimate.  $\blacksquare$

**3.6. Increment sequences are typically tame.** Before turning to the tail estimates on the increments themselves, we prove an easy preliminary estimate, showing that under the event  $\mathbf{I}_{x,T}$ , the probability that  $\mathcal{S}_x$  is not contained in a ball of radius of order  $T$  centered at  $v_{\tau_{\text{sp}}}$  is exponentially small in  $T$ .

Let  $r_0$  be a large constant, say 20, and let  $R_0 := 5r_0$ ; we will reserve these letters for these specific constants. We now define a notion of *tameness* for spines, and subsequently in Lemma 3.15 demonstrate that with high probability, a spine is tame.

**Definition 3.13.** Fix  $T$ ; for every  $t$ , a spine in  $\mathcal{S}_x \in \mathfrak{X}^t \times \mathfrak{X}_{\text{REM}}$  is *tame* with respect to  $\mathbf{I}_{x,T}$  if

$$\mathfrak{m}(\mathcal{S}_x) \leq r_0 T, \quad \text{and} \quad \text{ht}(\mathcal{S}_x) \leq r_0 T.$$

Call an interface  $\mathcal{I} \in \mathbf{I}_{x,T}$  *tame* if its spine  $\mathcal{S}_x$  is tame, and denote by  $\bar{\mathbf{I}}_{x,T}$  the set of tame interfaces in  $\mathbf{I}_{x,T}$ .

Before turning to the proof that spines are typically tame, we pause to comment on the usefulness of restricting to tame spines going forward.

**Remark 3.14.** First of all, notice that the tameness of a spine is only a property of the increment sequence constituting the spine  $\mathfrak{X}^{\mathcal{J}-\tau_{\text{sp}}} \times \mathfrak{X}_{\text{REM}}$  and does not depend on the truncation below it. Moreover, note that any spine  $\mathcal{S}_x$  with source point  $v_{\tau_{\text{sp}}}$  that is tame is such that the spine  $\mathcal{S}_x$  is contained entirely in a cylinder of radius  $r_0 T$  and height  $r_0 T$  above (and centered at)  $v_{\tau_{\text{sp}}}$ . This is in turn confined to the cylinder  $\mathcal{C}_{2r_0 T}(v_{\tau_{\text{sp}}}) \subset \mathcal{C}_{v_{\tau_{\text{sp}}}, x, T}$ , so that for any  $x$  such that  $d(x, \partial\Lambda_n) \geq 100T$ , adjoining to any  $T$ -admissible interface  $\mathcal{I}_{\text{TR}}$  any tame spine (identified with an element of  $\mathfrak{X}^{\mathcal{J}-\tau_{\text{sp}}} \times \mathfrak{X}_{\text{REM}}$ ), yields a valid interface in  $\bar{\mathbf{I}}_{x,T}$ .

Additionally, notice that, by construction, if  $\mathcal{I} \in \bar{\mathbf{I}}_{x,T}$ , for any face  $f \in \mathcal{S}_x$ , the distance  $d(f, \mathcal{I}_{\text{TR}})$  is attained by a face in  $\mathcal{I}_{\text{TR}} \cap \mathcal{C}_{R_0 T}(v_{\tau_{\text{sp}}}) \subset \mathcal{C}_{v_{\tau_{\text{sp}}}, x, T}$ , as the distance to  $\mathcal{I}_{\text{TR}} \setminus \mathcal{C}_{R_0 T}(v_{\tau_{\text{sp}}})$  is at least  $3r_0 T$  while the distance to  $v_{\tau_{\text{sp}}}$  is at most  $2r_0 T$ .

We now prove that spines of interfaces in  $\mathbf{I}_{x,T}$  are exponentially unlikely in  $T$  to not be tame.

**Lemma 3.15.** *There exists  $C, \beta_0 > 0$  such that for every  $\beta > \beta_0$ , such that for every  $T$ , every  $T$ -admissible truncated interface  $\mathcal{I}_{\text{TR}}$ , we have that for every  $r \geq 8T$*

$$\mu_n(\mathfrak{m}_T(\mathcal{S}_x) \geq r \mid \mathcal{I}_{\text{TR}}, \mathbf{I}_{x,T}) \leq \exp[-(\beta - C)r].$$

*In particular,  $\mu_n(\bar{\mathbf{I}}_{x,T} \mid \mathcal{I}_{\text{TR}}, \mathbf{I}_{x,T}) \geq 1 - O(e^{-(\beta-C)r_0 T})$ , and hence also  $\mu_n(\bar{\mathbf{I}}_{x,T} \mid \mathbf{I}_{x,T}) \geq 1 - O(e^{-(\beta-C)r_0 T})$ .*

*Proof.* The second statement follows from the fact that  $\mathbf{m}(\mathcal{S}_x) \leq \mathbf{m}_T(\mathcal{S}_x)$  and  $\text{ht}(\mathcal{S}_x) \leq T + \frac{1}{4}\mathbf{m}_T(\mathcal{S}_x)$  and  $T \leq \frac{3}{4}r$  when  $r \geq 8T$ . Therefore, we focus on proving the bound on  $\{\mathbf{m}_T(\mathcal{S}_x) \geq r\}$ . Let  $\mathcal{I} \in \mathbf{I}_{x,T}$  be such that it has  $T$ -admissible truncation  $\mathcal{I}_{\text{TR}}$  with source-point index  $\tau_{\text{SP}}$ , and spine  $\mathcal{S}_x$  with increment collection  $(X_i)_{\tau_{\text{SP}} \leq i \leq T}$  and  $X_{>T}$ , such that  $\mathbf{m}_T(\mathcal{S}_x) = \mathbf{m}(X_{>T}) + \sum_{\tau_{\text{SP}} \leq i \leq T} \mathbf{m}(X_i) \geq r$ . Let  $\mathcal{I}_{\emptyset,T}$  be the interface with the same  $T$ -admissible truncation and spine of exactly  $T - \tau_{\text{SP}}$  increments that are all  $X_{\emptyset}$ . By Theorem 2.21,

$$\frac{\mu_n(\mathcal{I})}{\mu_n(\mathcal{I}_{\emptyset,T})} = \exp\left(-\beta \mathbf{m}(\mathcal{I}; \mathcal{I}_{\emptyset,T}) + \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \mathcal{I}_{\emptyset,T}} \mathbf{g}(f', \mathcal{I}_{\emptyset,T})\right),$$

and we recall that  $\mathbf{m}(\mathcal{I}; \mathcal{I}_{\emptyset,T}) = \mathbf{m}_T(\mathcal{S}_x)$ . Denote by  $\mathcal{S}_x(\mathcal{I}_{\emptyset,T})$  the spine of  $\mathcal{I}_{\emptyset,T}$ . We can bound the difference,

$$\left| \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \mathcal{I}_{\emptyset,T}} \mathbf{g}(f', \mathcal{I}_{\emptyset,T}) \right| \leq \sum_{f \notin \mathcal{S}_x} |\mathbf{g}(f, \mathcal{I}) - \mathbf{g}(f, \mathcal{I}_{\emptyset,T})| + \sum_{f \in \mathcal{S}_x} |\mathbf{g}(f, \mathcal{I})| + \sum_{f' \in \mathcal{S}_x(\mathcal{I}_{\emptyset,T})} |\mathbf{g}(f', \mathcal{I}_{\emptyset,T})|.$$

By (2.2), the latter two terms contribute at most  $\bar{K}(|\mathcal{S}_x| + 4T + 1) = \bar{K}(\mathbf{m}(\mathcal{I}; \mathcal{I}_{\emptyset,T}) + 8T + 2)$ . By (2.3), the first term is bounded as

$$\sum_{f \notin \mathcal{S}_x} \bar{K} \exp[-\bar{c}r(f, \mathcal{I}; f, \mathcal{I}_{\emptyset,T})] \leq \sum_{f \notin \mathcal{S}_x} \sum_{f' \in \mathcal{I} \oplus \mathcal{I}_{\emptyset,T}} \bar{K} e^{-\bar{c}d(f, f')} \leq \sum_{f' \in \mathcal{I} \oplus \mathcal{I}_{\emptyset,T}} \sum_{f \in \mathbb{Z}^3} \bar{K} e^{-\bar{c}d(f, f')},$$

which by integrability of exponential tails is at most  $\bar{C}|\mathcal{I} \oplus \mathcal{I}_{\emptyset,T}| \leq \bar{C}(\mathbf{m}(\mathcal{I}; \mathcal{I}_{\emptyset,T}) + 8T + 2)$  for some universal  $\bar{C}$ . As such, once  $\mathbf{m}(\mathcal{I}; \mathcal{I}_{\emptyset,T}) \geq 8T$ , say, this is comparable up to a universal constant to  $\mathbf{m}_T(\mathcal{S}_x) = \mathbf{m}(\mathcal{I}; \mathcal{I}_{\emptyset,T})$ . Also, notice that the number of possible spines  $\mathcal{S}_x$  of excess area  $\mathbf{m}_T(\mathcal{S}_x) = k$  is at most the number of connected sets of faces of size  $k + 1$  incident to the upper-delimiting face of  $v_{\tau_{\text{SP}}}$ , which is at most  $s^{k+1}$ , for some universal  $s$  by Observation 2.27. Thus, there is a universal  $C$  such that for any  $r \geq r_0T$ , we have

$$\begin{aligned} \mu_n(\mathbf{m}_T(\mathcal{S}_x) \geq r \mid \mathcal{I}_{\text{TR}}, \mathbf{I}_{x,T}) &\leq \sum_{k \geq r} \sum_{\substack{\mathcal{S}_x: \mathcal{S}_x \cup \mathcal{I}_{\text{TR}} \in \mathbf{I}_{x,T} \\ \mathbf{m}_T(\mathcal{S}_x) = k}} \mu_n(\mathcal{S}_x \mid \mathcal{I}_{\text{TR}}) \leq \sum_{k \geq r} \sum_{\substack{\mathcal{S}_x: \mathcal{S}_x \cup \mathcal{I}_{\text{TR}} \in \mathbf{I}_{x,T} \\ \mathbf{m}_T(\mathcal{S}_x) = k}} \frac{\mu_n(\mathcal{S}_x, \mathcal{I}_{\text{TR}})}{\mu_n(\mathcal{S}_x, \mathcal{I}_{\emptyset,T}, \mathcal{I}_{\text{TR}})} \\ &\leq \sum_{k \geq r} s^{k+1} \exp[-(\beta - C)k], \end{aligned}$$

at which point, absorbing the  $s^k$  into the exponential, yields the desired bound for some different  $C$ .  $\blacksquare$

**Remark 3.16.** If  $T$  is comparable to  $h$ , we can attain a version of Lemma 3.15 that also conditions on the event  $\{\text{ht}(\mathcal{P}_x) \geq h\}$ . Namely, for any  $\mathcal{I}_{\text{TR}}$ , if we set  $T' = T \vee (\text{ht}(\mathcal{P}_x) - \frac{1}{2} - \text{ht}(v_{\tau_{\text{SP}}}) + \tau_{\text{SP}})$ , and apply the proof of Lemma 3.15 with respect to  $\mathcal{I}_{\emptyset, T'}$ , we would see that for every  $\mathcal{I} \in \mathbf{I}_{x,T} \cap \{\text{ht}(\mathcal{P}_x) \geq h\}$ , we have

$$\frac{\mu_n(\mathcal{I})}{\mu_n(\mathcal{I}_{\emptyset, T'})} \leq \exp[-(\beta - C)\mathbf{m}_{T'}(\mathcal{S}_x) + 8CT'].$$

As long as  $r \geq 8T'$ , this would imply that  $\mu_n(\mathbf{m}_{T'}(\mathcal{S}_x) \geq r \mid \mathcal{I}_{\text{TR}}, \mathbf{I}_{x,T}, \text{ht}(\mathcal{P}_x) \geq h) \leq \exp(-(\beta - C)r)$ ; therefore, as long as  $8T' \leq 8(T \vee h)$  is less than  $20T$ , e.g., as long as  $\frac{h}{2} \leq T \leq h$ , we have for every  $\mathcal{I}_{\text{TR}}$ ,

$$\mu_n(\bar{\mathbf{I}}_{x,T}^c \mid \mathbf{I}_{x,T}, \text{ht}(\mathcal{P}_x) \geq h, \mathcal{I}_{\text{TR}}) \leq \exp[-4(\beta - C)h]. \quad (3.3)$$

#### 4. EXPONENTIAL TAIL ON INCREMENT EXCESS AREAS

In this section, we control the excess areas of the increments that constitute the spine of a tall pillar. Of course it could be that the source point of the spine is itself an order  $T$  distance from  $x$  and the base contributes macroscopically to the surface area, but this is ruled out in Section 5. Henceforth, take  $T$  to be large and take  $x$  to be any point in the ‘‘bulk’’ of  $\mathcal{L}_0 \cap \Lambda_{n,n,\infty}$  relative to  $T$ , e.g.,  $d(x, \partial\Lambda_n) \geq 100T$ .

We show an exponential tail on the excess area of the  $i$ -th increment of the spine of an interface  $\mathcal{I} \in \bar{\mathbf{I}}_{x,T}$ ; the bound will be uniform over both the truncated interface and all the increments below the  $i$ -th one.

**Proposition 4.1.** *There exists  $c_0 > 0$  such that for every  $\beta > \beta_0$ , every  $T$ , and every  $i \leq T$ , we have that*

$$\mu_n(\mathbf{m}(\mathcal{X}_{\tau_{\text{SP}}+i}^* \geq r \mid \bar{\mathbf{I}}_{x,T}) \leq \exp[-c_0\beta r],$$

where if  $\tau_{\text{SP}} + i > \mathcal{T}$ , we define  $\mathbf{m}(\mathcal{X}_{\tau_{\text{SP}}+i}^*) = 0$ . In fact, for every  $T$ -admissible truncation  $\mathcal{I}_{\text{TR}}$  and every sequence of increments  $(X_{\tau_{\text{SP}}+j})_{j < i} \in \mathfrak{X}$ , we have the same estimate:

$$\mu_n(\mathbf{m}(\mathcal{X}_{\tau_{\text{SP}}+i}^* \geq r \mid \mathcal{I}_{\text{TR}}, (X_{\tau_{\text{SP}}+j})_{j < i} = (X_{\tau_{\text{SP}}+j})_{j < i}, \bar{\mathbf{I}}_{x,T}) \leq \exp[-c_0\beta r].$$

A useful corollary of the above proposition is the following tail estimate on a quantity measuring the interaction of the spine with the truncated interface  $\mathcal{I}_{\text{TR}}$ .

**Corollary 4.2.** *Let  $c_0 > 0$  be the constant from Proposition 4.1. There exists some  $C > 0$  such that for every  $\beta > \beta_0$  and every  $T$ , for each  $T$ -admissible truncation  $\mathcal{I}_{\text{TR}}$  and every  $r > 0$ ,*

$$\mu_n \left( \sum_{i \geq 1} |\mathcal{F}(\mathcal{X}_{\tau_{\text{sp}}+i})| e^{-\bar{c}i} + |\mathcal{F}(\mathcal{X}_{>T})| e^{-\bar{c}(T+1-\tau_{\text{sp}})} \geq r \mid \mathcal{I}_{\text{TR}}, \bar{\mathbf{I}}_{x,T} \right) \leq C \exp \left[ -\frac{1}{2} c_0 \beta (r - C) \right].$$

Similarly, for every  $T$ -admissible truncation  $\mathcal{I}_{\text{TR}}$  and increment sequence  $(X_{\tau_{\text{sp}}+i})_{i \leq i_0}$ , we have

$$\begin{aligned} \mu_n \left( |\mathcal{F}(\mathcal{X}_{>T})| e^{-\bar{c}(T+1-\tau_{\text{sp}}-i_0)} + \sum_{i \geq i_0} |\mathcal{F}(\mathcal{X}_{\tau_{\text{sp}}+i})| e^{-\bar{c}(i-i_0)} \geq r \mid \mathcal{I}_{\text{TR}}, (\mathcal{X}_{\tau_{\text{sp}}+i})_{i < i_0} = (X_{\tau_{\text{sp}}+i})_{i < i_0}, \bar{\mathbf{I}}_{x,T} \right) \\ \leq C \exp \left[ -\frac{1}{2} c_0 \beta (r - C) \right]. \end{aligned}$$

*Proof.* By Proposition 4.1, and in particular its second assertion, for any  $T$ -admissible truncated interface  $\mathcal{I}_{\text{TR}}$  with source-point index  $\tau_{\text{sp}}$ , the sequence  $(\mathbf{m}(\mathcal{X}_{\tau_{\text{sp}}+i}))_{i \geq 1}$  is dominated by a sequence of i.i.d. exponential random variables  $\xi_i$  with rate  $c_0 \beta$  (as seen by revealing the increments one at a time from bottom to top). Noting that for every  $0 < \lambda \leq \frac{1}{2} c_0 \beta$  and every  $i \geq 1$ ,

$$\mathbb{E}_{\mu_n} [\exp(\lambda e^{-\bar{c}i} \xi_i)] = [1 - \lambda e^{-\bar{c}i} / (c_0 \beta)]^{-1} \leq 1 + 2(c_0 \beta)^{-1} \lambda e^{-\bar{c}i} \leq \exp(e^{-\bar{c}i}),$$

we set  $\lambda = \frac{1}{2} c_0 \beta$  and obtain that

$$\begin{aligned} \mathbb{E}_{\mu_n} \left[ \exp \left( \lambda \sum_{i \geq 1} |\mathcal{F}(\mathcal{X}_{\tau_{\text{sp}}+i})| e^{-\bar{c}i} \right) \mid \mathcal{I}_{\text{TR}}, \bar{\mathbf{I}}_{x,T} \right] &\leq \prod_{i \geq 1} \mathbb{E} \left[ \exp \left( \lambda (\xi_i + 4) e^{-\bar{c}i} \right) \right] \leq \exp \left( \sum_{i \geq 1} (1 + 4\lambda) e^{-\bar{c}i} \right) \\ &\leq \exp \left( \frac{1 + 2c_0 \beta}{1 - e^{-\bar{c}}} \right). \end{aligned}$$

Letting  $\gamma = 1/(1 - e^{-\bar{c}})$ , this implies by Markov's inequality that

$$\mu_n \left( \sum_{i \geq 1} |\mathcal{F}(\mathcal{X}_{\tau_{\text{sp}}+i})| e^{-\bar{c}i} \geq r \mid \mathcal{I}_{\text{TR}}, \bar{\mathbf{I}}_{x,T} \right) \leq e^{(1+2c_0\beta)\gamma - \lambda r} = e^\gamma \exp \left[ -\frac{1}{2} c_0 \beta (r - 4\gamma) \right].$$

The matching conditional bounds follow from the analogous conditional estimates in Proposition 4.1.  $\blacksquare$

We prove Proposition 4.1 by constructing a map for shrinking increments of the pillar. In order to do so, we define a map between collections of pillars that replaces increments of the pillar with  $X_\emptyset$ , decreasing the excess area of the increment and, in turn the pillar—the complication is that unlike the map  $\Phi_x$  of [28], the effect of this removal is not localized and translates the entirety of the pillar above that increment.

**4.1. The increment reduction map  $\Psi_i$ .** For each  $T$  and  $i \leq T$ , we define a map  $\Psi_i$  that replaces the  $i$ -th increment of a spine with a stretch of trivial increments  $X_\emptyset$ .

**Definition 4.3.** For every  $i \leq T$ , we will define the map  $\Psi_i : \bar{\mathbf{I}}_{x,T} \rightarrow \bar{\mathbf{I}}_{x,T}$ . Suppose  $\mathcal{I} \in \bar{\mathbf{I}}_{x,T}$ , consists of a  $T$ -admissible truncated interface  $\mathcal{I}_{\text{TR}}$  with source point index  $\tau_{\text{sp}}$ , an increment sequence  $(X_{\tau_{\text{sp}}+j})_{j \leq T-\tau_{\text{sp}}}$ , and remainder  $X_{>T} \in \mathfrak{X}_{\text{REM}}$ . Then  $\Psi_i(\mathcal{I})$  will have the same truncated interface  $\mathcal{I}_{\text{TR}}$ , and its spine will have increment sequence  $(X'_{\tau_{\text{sp}}+j})_{j \leq T-\tau_{\text{sp}}}$  and  $X'_{>T}$  constructed as follows. If  $X_i = X_\emptyset$  or if  $\tau_{\text{sp}} + i > T$ , then let  $(X'_j)_{j \leq T} = (X_j)_{j \leq T}$  and  $X'_{>T} = X_{>T}$ ; otherwise, construct the increment sequence of  $\Psi_i(\mathcal{I})$  by taking the increment sequence  $(X_j)_{\tau_{\text{sp}} \leq j \leq T}$  and

- (1) Mark the index  $\tau_{\text{sp}} + i$ , as well as every index  $j > \tau_{\text{sp}} + i$  having the property that

$$\mathbf{m}(X_j) \geq \mathbf{m}(X_{\tau_{\text{sp}}+i}) e^{\frac{1}{2} \bar{c}(j-\tau_{\text{sp}}-i)}.$$

Also mark the remainder if it has  $\mathbf{m}(X_{>T}) \geq \mathbf{m}(X_{\tau_{\text{sp}}+i}) e^{\frac{1}{2} \bar{c}(T+1-\tau_{\text{sp}}-i)}$ .

- (2) Label the sequence of marked indices  $j_0 = \tau_{\text{sp}} + i$ , and  $j_0 < j_1 < \dots < j_\kappa$ , where, if the remainder is marked,  $j_\kappa$  is  $> T$ .
- (3) For each marked index  $j_k$ , replace  $X_{j_k}$  in the increment sequence  $(X_j)_j$  by a stretch of  $\text{ht}(X_{j_k})$  consecutive trivial increments  $(X_\emptyset, \dots, X_\emptyset)$ , to obtain  $(X'_j)_j$ .

We refer the reader to Figure 9 for a visualization of the map  $\Psi_i$ .

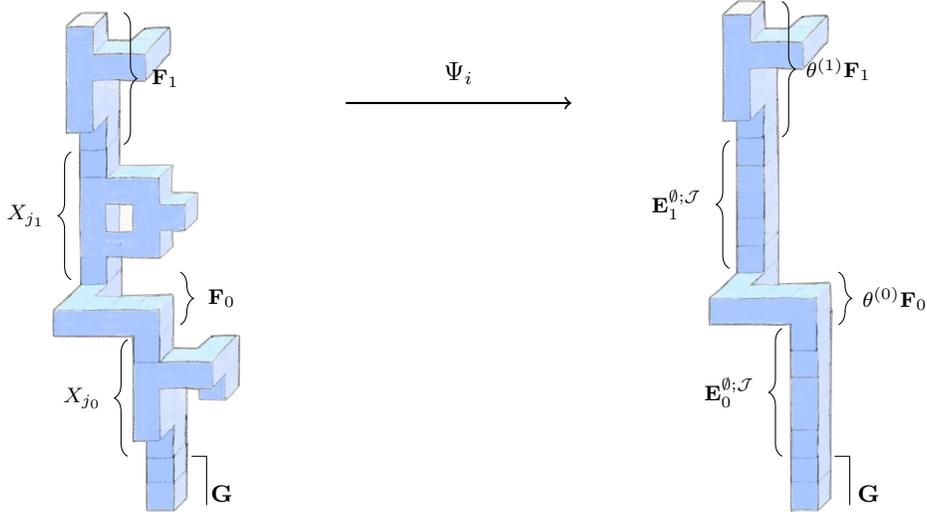


FIGURE 9. The increment map  $\Psi_i$  sends the stretch of increments on the left to the stretch on the right. The increment  $X_{j_0} = X_{\tau_{\text{sp}}+i}$  is replaced by a stretch of five trivial increments; the increment  $X_{j_1} = X_{\tau_{\text{sp}}+i+2}$  is also replaced by trivial increments as  $\mathbf{m}(X_{j_1}) \geq \mathbf{m}(X_{j_0})e^{\tilde{c}}$ .

**4.2. Strategy of the map  $\Psi_i$ .** Let us briefly describe the strategy behind the construction of the map above. Our goal is to show an exponential tail on the excess area of the  $(\tau_{\text{sp}} + i)$ 'th increment conditionally either on having at least  $T$  increments—Proposition 4.1—or on having height at least  $h$  and  $T$  increments—Proposition 4.7. Towards this, we wish to construct a map  $\Psi_i$  having that

- (1) For every  $\mathcal{I} \in \bar{\mathbf{I}}_{x,T} \cap \{\text{ht}(\mathcal{P}_x) \geq h\}$ , the interface  $\Psi_i(\mathcal{I}) \in \bar{\mathbf{I}}_{x,T} \cap \{\text{ht}(\mathcal{P}_x) \geq h\}$
- (2)  $\mu_n(\bar{\mathbf{I}}_{x,T} \cap \{\mathbf{m}(\mathcal{X}_{\tau_{\text{sp}}+i}) > r\}) \leq e^{-(\beta-C)r} \mu_n(\Psi_i(\bar{\mathbf{I}}_{x,T}))$  via the steps (1)–(3) of (1)–(3) in §1.3 as well as the analogue of this inequality, with both events also intersected with  $\{\text{ht}(\mathcal{P}_x) \geq h\}$ .

Towards this, our map replaces the  $(\tau_{\text{sp}} + i)$ 'th increment by a sequence of  $\text{ht}(\mathcal{X}_{\tau_{\text{sp}}+i})$  trivial increments, yielding an energy gain that is comparable to  $\mathbf{m}(\mathcal{X}_{\tau_{\text{sp}}+i})$ . N.b. replacing it by just one trivial increment would not ensure that the resulting pillar also attains the same height as the original pillar.

Unlike changes in the standard wall representation, changes in the increment sequence subsequently induce a *horizontal* shift of all increments above the  $(\tau_{\text{sp}} + i)$ 'th one. These horizontally shifted increments  $\mathcal{X}_j$  can then interact with increments below  $\mathcal{X}_{\tau_{\text{sp}}+i}$  via the term  $\mathbf{g}(f, \mathcal{I}; f', \Psi_i(\mathcal{I}))$  of (2.3). By (2.3), this quantity decays exponentially in the distance to  $\mathcal{X}_{\tau_{\text{sp}}+i}$ , so that if the excess area  $\mathbf{m}(\mathcal{X}_j)$  is larger than  $e^{-d(\mathcal{X}_j, \mathcal{X}_{\tau_{\text{sp}}+i})}$ , we cannot compare the contribution of the perturbative  $\mathbf{g}$  term to the energy gain of the map. For this reason, we additionally delete all increments whose excess areas are greater than some exponential factor times their distance to  $\mathcal{X}_{\tau_{\text{sp}}+i}$ . Iterating this procedure up the spine yields the map  $\Psi_i$ .

The following remark summarizes the properties of the map  $\Psi_i$  that we will use in its analysis.

**Remark 4.4.** By construction, the excess area of the spine of  $\Psi_i(\mathcal{I})$  is at most the excess area of the spine of  $\mathcal{I}$ , the map  $\Psi_i(\mathcal{I})$  keeps the height of  $\mathcal{S}_x$ , and thus also  $\mathcal{P}_x$ , fixed, and the map  $\Psi_i$  only increases the number of increments of the spine. Therefore, for every  $\mathcal{I} \in \bar{\mathbf{I}}_{x,T}$ , we have  $\Psi_i(\mathcal{I}) \in \bar{\mathbf{I}}_{x,T}$ . Moreover, notice that the truncated interfaces of  $\mathcal{I}$  and  $\Psi_i(\mathcal{I})$ , and their first  $\tau_{\text{sp}} + i - 1$  increments, agree.

**4.3. Analysis of the map  $\Psi_i$ .** We will bound the effect of the map on the energy in Proposition 4.5, and its multiplicity in Lemma 4.6. Combining these will imply Lemma 4.1, the main result in this section.

**Proposition 4.5.** *There exists  $C > 0$  such that for every  $\beta > \beta_0$  and every  $i \leq T$ , if  $\mathcal{I} \in \bar{\mathbf{I}}_{x,T}$ ,*

$$\left| \log \frac{\mu_n(\mathcal{I})}{\mu_n(\Psi_i(\mathcal{I}))} + \beta \mathbf{m}(\mathcal{I}; \Psi_i(\mathcal{I})) \right| \leq C \mathbf{m}(\mathcal{I}; \Psi_i(\mathcal{I})).$$

*Proof.* For ease of notation, fix any such  $i$  and let  $\mathcal{J} = \Psi_i(\mathcal{I})$ . Suppose that  $\mathcal{I}$  has  $T$ -admissible truncation  $\mathcal{I}_{\text{TR}}$  with increment sequence  $(X_{\tau_{\text{sp}}+i})_{i \leq T - \tau_{\text{sp}}}$  and remainder  $X_{>T}$ . If  $\mathcal{J} = \mathcal{I}$ , the inequality trivially holds,

so let us assume that  $\mathcal{I}$  is such that  $\mathfrak{m}(X_{\tau_{\text{sp}}+i}) > 0$ . By Theorem 2.21, we can express

$$\frac{\mu_n(\mathcal{I})}{\mu_n(\mathcal{J})} = \exp\left(-\beta\mathfrak{m}(\mathcal{I}; \mathcal{J}) + \sum_{f \in \mathcal{I}} \mathfrak{g}(f, \mathcal{I}) - \sum_{f' \in \mathcal{J}} \mathfrak{g}(f', \mathcal{J})\right).$$

Let  $\{j_k\}_{k \leq \kappa}$  be as in Definition 4.3. To allow us to consider the increments and remainder in a uniform manner, let  $j_\kappa := T + 1$  if  $j_\kappa$  is “ $> T$ ”, so that  $X_{j_\kappa}$  refers to  $X_{>T}$  if  $j_\kappa = T + 1$ . We have

$$\mathfrak{m}(\mathcal{I}; \Psi_i(\mathcal{I})) \geq \frac{1}{3} \sum_{k \leq \kappa} \mathfrak{m}(X_{j_k}),$$

since every nontrivial increment with height bigger than 1 must have at least six faces at each height between  $\lceil \text{ht}(v_j) \rceil$  and  $\lceil \text{ht}(v_{j+1}) \rceil$ , whereas the stretch of trivial increments would have four faces at those heights. Now let us split the set of faces in  $\mathcal{I}$  into the following sets (refer to Figure 9):

- For each  $k \leq \kappa$ , let  $\mathbf{E}_k$  be the set of faces  $\mathcal{F}(X_{j_k})$  in  $\mathcal{I}$ .
- For each  $k \leq \kappa$ , let  $\mathbf{F}_k$  be the (possibly empty if  $j_{k+1} = j_k + 1$  or if  $j_k = T + 1$ ) set of all faces between  $v_{j_{k+1}}$  and  $v_{j_k}$  (not-inclusive), with  $\mathbf{F}_\kappa$  defined as the set of all faces above  $v_{j_\kappa}$ .
- Let  $\mathbf{G}$  be the set of all remaining faces in  $\mathcal{I}$

Also, for notation, let  $\mathbf{E}_k^\emptyset \subset \mathbf{E}_k$  be the bounding faces of  $v_{j_k}$  and  $v_{j_{k+1}}$  in  $\mathcal{I}$  (so  $|\mathbf{E}_k^\emptyset| = 8$ ) and if  $j_\kappa = T + 1$ , then  $\mathbf{E}_\kappa^\emptyset$  will only be the four bounding faces of  $v_{j_\kappa}$  in  $\mathcal{I}$ . Let  $\mathbf{E}_k^{\emptyset; \mathcal{J}}$  be the corresponding faces in  $\mathcal{J}$ , i.e., the faces of the  $\text{ht}(X_{j_k})$  consecutive trivial increments, so that  $|\mathbf{E}_k^{\emptyset; \mathcal{J}}| = 4 \text{ht}(X_{j_k})$  (if  $j_\kappa = T + 1$ , also include the top-most bounding face of the spine in  $\mathcal{J}$ ).

By definition, the faces in  $\mathbf{G}$  are shared between  $\mathcal{I}$  and  $\mathcal{J}$ , the faces  $\bigcup_k \mathbf{E}_k \setminus \mathbf{E}_k^\emptyset$  are precisely those that are removed by the map  $\Psi_i$ , and the faces in  $\mathbf{F}_k$  can be translated to correspond in a one-to-one fashion to the faces in  $\mathcal{J} \setminus (\mathbf{G} \cup \bigcup_k \mathbf{E}_k^{\emptyset; \mathcal{J}})$ . Namely, if for every  $k \leq \kappa$ , we set  $\theta^{(k)}$  to be the shift map by the vector

$$-\sum_{0 \leq \ell \leq k} \rho(v_{j_{\ell+1}}) - \rho(v_{j_\ell})$$

then every face  $f \in \mathbf{F}_k$ , is identified with the face  $\theta^{(k)}f$  in  $\mathcal{J}$ , and for  $k < \kappa$ , each stretch  $\theta^{(k)}\mathbf{F}_k$  is delimited from below by the upper-bounding face of  $\theta^{(k)}\mathbf{E}_k^\emptyset \subset \mathbf{E}_k^{\emptyset; \mathcal{J}}$  and from above by the lower-bounding face of  $\theta^{(k+1)}\mathbf{E}_{k+1}^\emptyset$ . By construction, we have

$$\mathcal{I} = \mathbf{G} \cup \bigcup_{k \leq \kappa} \mathbf{E}_k \cup \mathbf{F}_k, \quad \text{and} \quad \mathcal{J} = \mathbf{G} \cup \bigcup_{k \leq \kappa} \mathbf{E}_k^{\emptyset; \mathcal{J}} \cup \theta^{(k)}\mathbf{F}_k.$$

We can therefore split up the sum

$$\begin{aligned} \left| \sum_{f \in \mathcal{I}} \mathfrak{g}(f, \mathcal{I}) - \sum_{f' \in \mathcal{J}} \mathfrak{g}(f', \mathcal{J}) \right| &\leq \sum_k \sum_{f \in \mathbf{E}_k} |\mathfrak{g}(f, \mathcal{I})| + \sum_k \sum_{f \in \mathbf{E}_k^{\emptyset; \mathcal{J}}} |\mathfrak{g}(f, \mathcal{J})| \\ &\quad + \sum_k \sum_{f \in \mathbf{F}_k} |\mathfrak{g}(f, \mathcal{I}) - \mathfrak{g}(\theta^{(k)}f, \mathcal{J})| + \sum_{f \in \mathbf{G}} |\mathfrak{g}(f, \mathcal{I}) - \mathfrak{g}(f, \mathcal{J})|. \end{aligned} \quad (4.1)$$

We bound these sums one at a time. By (2.2), along with  $|\mathcal{F}(X_{j_k})| \leq 5\mathfrak{m}(X_{j_k})$  (by Remark 3.9 and  $\mathfrak{m}(X_{j_k}) > 0$ ), the first and second sums in (4.1) are bounded above as

$$\sum_k \sum_{f \in \mathbf{E}_k} |\mathfrak{g}(f, \mathcal{I})| + \sum_k \sum_{f \in \mathbf{E}_k^{\emptyset; \mathcal{J}}} |\mathfrak{g}(f, \mathcal{J})| \leq 2\bar{K} \sum_k |\mathcal{F}(X_{j_k})| \leq 10\bar{K}\mathfrak{m}(\mathcal{I}; \mathcal{J}).$$

Let us now turn to the third term of (4.1), which we can bound as follows: first of all, notice that for any face  $f \in \mathbf{F}_k$  in increment  $\mathcal{F}(X_j)$  for  $j_k < j < j_{k+1}$ , the radius  $\mathfrak{r}(f, \mathcal{I}; \theta^{(k)}f, \mathcal{J})$  is attained either by some face in a spine (belonging to precisely one of  $\mathcal{I}$  or  $\mathcal{J}$ ), in which case its value is at least  $(j - j_k) \wedge (j_{k+1} - j)$ , or by some face in  $\mathbf{G}$  (the faces in  $\mathbf{G}$  are the same in both  $\mathcal{I}$  and  $\mathcal{J}$ , but will be at different relative locations to  $f$  vs.  $\theta^{(k)}f$ ). Let us take any  $k \leq \kappa$ , fix a  $j_k < j < j_{k+1}$  ( $j > j_\kappa$  if  $k = \kappa$  and  $j_\kappa \neq T + 1$ ) and a face  $f \in \mathcal{F}(X_j) \subset \mathbf{F}_k$ , and expand

$$\begin{aligned} |\mathfrak{g}(f, \mathcal{I}) - \mathfrak{g}(\theta^{(k)}f, \mathcal{J})| &\leq \bar{K} \exp[-\bar{c}\mathfrak{r}(f, \mathcal{I}; \theta^{(k)}f, \mathcal{J})] \\ &\leq \bar{K} \exp[-\bar{c}(j - j_k)] + \bar{K} \exp[-\bar{c}(j_{k+1} - j)] + \bar{K} \exp[-\bar{c}d(\{f, \theta^{(k)}f\}, \mathbf{G})]. \end{aligned}$$

Notice that since both  $\mathcal{I}$  and  $\mathcal{J}$  are in  $\bar{\mathbf{I}}_{x,T}$ , their spines are contained in  $\mathcal{C}_{2r_0T}(v_{\tau_{\text{sp}}})$ . As a consequence,  $d(f, \mathbf{G})$  is attained by a face in  $\mathcal{C}_{R_0T}(v_{\tau_{\text{sp}}})$ , and is at least  $j - \tau_{\text{sp}} - i \geq j - j_k$  (n.b. there are  $j - \tau_{\text{sp}} - i$  cut-points separating  $\mathbf{G}$  and  $f$ ), and the same holds for  $d(\theta^{(k)}f, \mathbf{G})$ ; therefore, the above becomes

$$|\mathbf{g}(f, \mathcal{I}) - \mathbf{g}(\theta^{(k)}f, \mathcal{J})| \leq 2\bar{K} \exp[-\bar{c}(j - j_k)] + \bar{K} \exp[-\bar{c}(j_{k+1} - j)]. \quad (4.2)$$

Now summing the first term in the right-hand side in (4.2) over all  $k$  and  $f \in \mathbf{F}_k$ ,

$$\bar{K} \sum_{k \leq \kappa} \sum_{j_k < j < j_{k+1}} |\mathcal{F}(X_j)| e^{-\bar{c}(j-j_k)} \leq \bar{K} \sum_{k \leq \kappa} \sum_{j_k < j < j_{k+1}} \left[8 + \mathbf{m}(X_{j_k})\right] e^{\frac{1}{2}\bar{c}(j-j_k)} e^{-\bar{c}(j-j_k)},$$

using that  $|\mathcal{F}(X_j)| = 8 + \mathbf{m}(X_j)$  for  $j_k < j < j_{k+1}$ , and the facts that the reduction map was not applied at index  $j$  and was applied at  $j_k$ , so that

$$\mathbf{m}(X_j) \leq \mathbf{m}(X_{\tau_{\text{sp}}+i}) e^{\frac{1}{2}\bar{c}(j-\tau_{\text{sp}}-i)} \leq \mathbf{m}(X_{j_k}) e^{\frac{1}{2}\bar{c}(j-j_k)}.$$

Then, by integrability of exponential tails, we see that the right-hand side above is in turn bounded by  $\bar{C} \sum_{k \leq \kappa} \mathbf{m}(X_{j_k}) = \bar{C} \mathbf{m}(\mathcal{I}; \mathcal{J})$  for some universal  $\bar{C}$ . The second term in (4.2) can similarly be bounded as

$$\bar{K} \sum_{k < \kappa} \sum_{j_k < j < j_{k+1}} \mathbf{m}(X_j) e^{-\bar{c}(j_{k+1}-j)} \leq \bar{K} \sum_{k < \kappa} \sum_{j_k < j < j_{k+1}} \mathbf{m}(X_j) \leq \bar{C} \sum_{k < \kappa} \mathbf{m}(X_{j_{k+1}}).$$

for some constant  $\bar{C}$ , where we used that

$$\sum_{j_k < j < j_{k+1}} \mathbf{m}(X_j) \leq \sum_{j_k < j < j_{k+1}} \mathbf{m}(X_{\tau_{\text{sp}}-i}) e^{\frac{1}{2}\bar{c}(j-\tau_{\text{sp}}-i)} \leq \bar{C} \mathbf{m}(X_{\tau_{\text{sp}}-i}) e^{\frac{1}{2}\bar{c}(j_{k+1}-\tau_{\text{sp}}-i)} \leq \bar{C} \mathbf{m}(X_{j_{k+1}}).$$

Again, by integrability of exponential tails, we see that for some other  $\bar{C} > 0$ , the contribution of this term is bounded by  $\bar{C} \sum_{k < \kappa} \mathbf{m}(X_{j_k}) \leq \bar{C} \mathbf{m}(\mathcal{I}; \mathcal{J})$ . It remains to bound the fourth sum in (4.1): for faces  $f \in \mathbf{G}$ , the radius  $\mathbf{r}(f, \mathcal{I}; f, \mathcal{J})$  must be attained by a face in  $\mathcal{I} \oplus \mathcal{J}$ , so that

$$\sum_{f \in \mathbf{G}} |\mathbf{g}(f, \mathcal{I}) - \mathbf{g}(f, \mathcal{J})| \leq \sum_{f \in \mathbf{G} \cap \mathcal{C}_{R_0T}(v_{\tau_{\text{sp}}})} \sum_{g \in \mathcal{I} \oplus \mathcal{J}} \bar{K} e^{-\bar{c}d(f,g)} + \sum_{f \in \mathbf{G}, f \notin \mathcal{C}_{R_0T}(v_{\tau_{\text{sp}}})} \sum_{g \in \mathcal{I} \oplus \mathcal{J}} \bar{K} e^{-\bar{c}d(f,g)}.$$

Since  $\mathcal{I} \oplus \mathcal{J} \subset \mathcal{C}_{2r_0T}(v_{\tau_{\text{sp}}})$ , integrating the exponential tail, the second sum above is at most  $O(T^2 e^{-\bar{c}r_0T})$ . On the other hand, by definition of the spine and the fact that it is tame,

$$\begin{aligned} \sum_{f \in \mathbf{G} \cap \mathcal{C}_{R_0T}(v_{\tau_{\text{sp}}})} \sum_{g \in (X_j)_{j \geq j_0} \cup (X'_j)_{j \geq j_0}} e^{-\bar{c}d(f,g)} &\leq 2 \sum_{j \geq j_0} |\mathcal{F}(X_j)| e^{-\bar{c}(j-\tau_{\text{sp}}-i)} \\ &\leq 8\bar{C} + 2 \sum_{k \leq \kappa} \mathbf{m}(X_{j_k}) + 2 \sum_{j \geq j_0: j \notin \{j_k\}_{k \leq \kappa}} \mathbf{m}(X_{\tau_{\text{sp}}+i}) e^{-\frac{1}{2}\bar{c}(j-\tau_{\text{sp}}-i)}, \end{aligned}$$

which we again find to be bounded by  $\bar{C} \mathbf{m}(\mathcal{I}; \mathcal{J})$  for some other universal constant  $\bar{C}$ . Plugging all the above bounds into (4.1), we see that for some universal  $C$  (independent of  $\beta$ ), we have

$$\left| \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \mathcal{J}} \mathbf{g}(f', \mathcal{J}) \right| \leq C \mathbf{m}(\mathcal{I}; \mathcal{J}). \quad \blacksquare$$

We now bound the multiplicity of the map  $\Psi_i$  for a fixed excess area  $\mathbf{m}(\mathcal{I}; \Psi_i(\mathcal{I}))$ .

**Lemma 4.6.** *For every  $i \leq T$  and every  $\mathcal{J} \in \Psi_i(\bar{\mathbf{I}}_{x,T})$ , there exists an  $s > 0$  such that for every  $K$ ,*

$$|\{\mathcal{I} \in \Psi_i^{-1}(\mathcal{J}) : \mathbf{m}(\mathcal{I}; \mathcal{J}) = K\}| \leq s^K.$$

*Proof.* Fix any  $i$  and  $\mathcal{J} \in \Psi_i(\bar{\mathbf{I}}_{x,T})$ . For a fixed spine  $\mathcal{S}_x(\mathcal{J})$  it suffices to bound the number of spines  $\mathcal{S}_x(\mathcal{I})$  for which the map  $\Psi_i$  sends  $\mathcal{S}_x(\mathcal{I})$  to  $\mathcal{S}_x(\mathcal{J})$  with  $\mathbf{m}(\mathcal{S}_x(\mathcal{I}); \mathcal{S}_x(\mathcal{J})) = K$  (as the map  $\Psi_i$  fixes all faces of the interface in  $\mathcal{I}_{\text{TR}}$ ). We first observe a few basic facts.

By definition of  $\Psi_i$ , any spine  $\mathcal{S}_x(\mathcal{I})$  that gets mapped to  $\mathcal{S}_x(\mathcal{J})$  by  $\Psi_i$  is such that their increment sequences  $(X_j)_{j < i}$  coincide, and therefore the spines agree up to the  $i$ -th increment of  $\mathcal{S}_x(\mathcal{J})$ , which will satisfy  $X'_{\tau_{\text{sp}}+i} = X_\emptyset$ . In particular, for a given  $\mathcal{J}$ , the interface  $\mathcal{I}$  is uniquely identified by the collection of increments  $(X_{j_k})_{k \leq \kappa}$  and the indices of those increments  $\{j_k\}_{k \leq \kappa}$ , since the rest of its spine is given by increments which are the same in both  $\mathcal{I}$  and  $\mathcal{J}$ . Therefore, starting from  $v_{\tau_{\text{sp}}+i}$  for the interface  $\mathcal{J}$  (which coincides with the same cell for  $\mathcal{I}$ ), we can build a set of faces that uniquely identify the interface  $\mathcal{I}$  by taking the union of all the increments  $(X_j)$  between  $X_{j_0} = X_{\tau_{\text{sp}}+i}$  and the final  $X_{j_\kappa}$  (inclusive).

The union of these increments, viewed as a subset of the spine  $\mathcal{S}_x$ , clearly forms a  $*$ -connected set of faces in  $\mathcal{F}(\mathbb{Z}^3)$  that are  $*$ -adjacent to the upper-bounding face of the marked cell  $v_i$ . We claim that this subset of  $\mathcal{S}_x$  has cardinality bounded above by  $CK$  for some universal  $C$ . This follows from the fact that the cardinality of the face set of an increment is at most  $4\mathbf{m}(X_j)$  so long as  $X_j \neq X_\emptyset$ , so that the total cardinality of the face set of  $\bigcup_{j_0 \leq j \leq j_\kappa} \mathcal{F}(X_j)$  is at most four times

$$\begin{aligned} \sum_{k \leq \kappa} \mathbf{m}(X_{j_k}) + \sum_{k < \kappa} \sum_{j_k < j < j_{k+1}} [1 + \mathbf{m}(X_j)] &\leq \mathbf{m}(\mathcal{I}; \mathcal{J}) + \sum_{j \notin \{j_k\}, j_0 < j < j_\kappa} 2\mathbf{m}(X_j) e^{\frac{1}{2}c(j - \tau_{\text{sp}} - i)} \\ &\leq \mathbf{m}(\mathcal{I}; \mathcal{J}) + 2\bar{C} \sum_{k \leq \kappa} \mathbf{m}(X_{j_k}), \end{aligned}$$

which is, in turn, bounded above by  $C\mathbf{m}(\mathcal{I}; \mathcal{J})$  for some large enough, universal  $C$ . Since this rooted face-set uniquely identifies  $\mathcal{I} \in \Psi_i^{-1}(\mathcal{J})$ , the result then follows immediately from Observation 2.27.  $\blacksquare$

**Proof of Proposition 4.1.** Since  $\mathbf{m}(\mathcal{I}; \Psi_i(\mathcal{I})) \geq \frac{1}{3}\mathbf{m}(\mathcal{X}_{\tau_{\text{sp}}+i})$ , it will suffice for us to show the upper bound on  $\mu_n(\mathbf{m}(\mathcal{I}; \Psi_i(\mathcal{I})) \geq r \mid \mathcal{I}_{\text{TR}}, \bar{\mathbf{I}}_{x,T})$ . Fix a  $T$ -admissible truncation  $\mathcal{I}_{\text{TR}}$  and an  $i \leq T$ , and express  $\mu_n(\mathbf{m}(\mathcal{I}; \Psi_i(\mathcal{I})) \geq r, \mathcal{I}_{\text{TR}}, \bar{\mathbf{I}}_{x,T})$  as

$$\begin{aligned} \sum_{\mathcal{I} \in \bar{\mathbf{I}}_{x,T} \cap \mathcal{I}_{\text{TR}}, \mathbf{m}(\mathcal{I}; \Psi_i(\mathcal{I})) \geq r} \mu_n(\mathcal{I}) &= \sum_{k \geq r} \sum_{\mathcal{J} \in \Psi_i(\bar{\mathbf{I}}_{x,T} \cap \mathcal{I}_{\text{TR}})} \sum_{\mathcal{I} \in \Psi_i^{-1}(\mathcal{J}): \mathbf{m}(\mathcal{I}; \mathcal{J}) = k} \frac{\mu_n(\mathcal{I})}{\mu_n(\Psi_i(\mathcal{I}))} \mu_n(\mathcal{J}) \\ &\leq \sum_{\mathcal{J} \in \Psi_i(\bar{\mathbf{I}}_{x,T} \cap \mathcal{I}_{\text{TR}})} \mu_n(\mathcal{J}) \sum_{k \geq r} s^k e^{-(\beta-C)k}, \end{aligned}$$

where we used the shorthand  $\bar{\mathbf{I}}_{x,T} \cap \mathcal{I}_{\text{TR}}$  to denote the set of interfaces in  $\bar{\mathbf{I}}_{x,T}$  with that truncation, and the inequality followed from Proposition 4.5 and Lemma 4.6. Since  $\Psi_i(\bar{\mathbf{I}}_{x,T} \cap \mathcal{I}_{\text{TR}}) \subset \bar{\mathbf{I}}_{x,T} \cap \mathcal{I}_{\text{TR}}$ , by integrability of exponential tails, the sum over  $k \geq r$  is at most  $Ce^{-(\beta-C)r}$  for some universal constant  $C$ , leaving

$$\mu_n(\mathbf{m}(\mathcal{I}; \Psi_i(\mathcal{I})) \geq r, \mathcal{I}_{\text{TR}}, \bar{\mathbf{I}}_{x,T}) \leq C\mu_n(\bar{\mathbf{I}}_{x,T}, \mathcal{I}_{\text{TR}}) \exp[-(\beta-C)r],$$

for some universal constant  $C$ . Dividing both sides by  $\mu_n(\mathcal{I}_{\text{TR}}, \bar{\mathbf{I}}_{x,T})$  yields the first estimate. The matching estimate conditional also on  $(\mathcal{X}_{\tau_{\text{sp}}+j})_{j < i}$  follows by repeating the argument, additionally restricting our sum to interfaces with that increment sequence, as the map  $\Psi_i$  fixes all increments before the  $(\tau_{\text{sp}} + i)$ -th one.  $\blacksquare$

**4.4. Exponential tail conditionally on  $\text{ht}(\mathcal{P}_x) \geq h$ .** In the proofs of the existence of a limiting large deviation rate and the law of large numbers for the maximum of the interface, it will be important to work with the monotone event  $\{\text{ht}(\mathcal{P}_x) \geq h\}$  rather than  $\bar{\mathbf{I}}_{x,T}$ . The fact that the map  $\Psi_i$  keeps the height of a spine fixed allows us to also deduce the analogous exponential tails conditional on  $\{\text{ht}(\mathcal{P}_x) \geq h\}$ . In fact, if one were only interested in estimates conditional on  $\bar{\mathbf{I}}_{x,T}$  (as are relevant to the shape theorem and central limit theorem), the map  $\Psi_i$  could be simplified to replace each increment in  $(X_{j_k})_k$  by  $X_\emptyset$ , keeping the number of increments fixed, but shrinking the height.

**Proposition 4.7.** *There exists  $C > 0$  such that for every  $\beta > \beta_0$ , every  $T \leq h$ , every half-integer  $h_1 \leq h$ , and every  $T$ -admissible truncation  $\mathcal{I}_{\text{TR}}$  having  $\text{ht}(v_{\tau_{\text{sp}}}) \leq h_1$ , we have*

$$\mu_n(|\mathcal{F}(\mathcal{S}_x \cap \mathcal{L}_{h_1})| \geq 4 + r \mid \text{ht}(\mathcal{P}_x) \geq h, \mathcal{I}_{\text{TR}}, \bar{\mathbf{I}}_{x,T}) \leq \exp[-\beta r/C].$$

Similarly, for every  $i \leq T$ , and every  $T$ -admissible truncation  $\mathcal{I}_{\text{TR}}$  with  $\tau_{\text{sp}} < i$ , and sequence  $(X_j)_{j < i}$ ,

$$\mu_n(\mathbf{m}(\mathcal{X}_i) \geq r \mid (X_j)_{j < i}, \mathcal{I}_{\text{TR}}, \text{ht}(\mathcal{P}_x) \geq h, \bar{\mathbf{I}}_{x,T}) \leq \exp[-\beta r/C]. \quad (4.3)$$

This latter estimate also implies the analogue of Corollary 4.2, also conditioned on  $\{\text{ht}(\mathcal{P}_x) \geq h\}$  for  $h \geq T$ .

*Proof.* The proof of (4.3) goes similarly to the proof of Proposition 4.1. Namely, if we restrict the proof therein to interfaces additionally having  $\text{ht}(\mathcal{P}_x) \geq h$ , and notice that for all such interfaces, their image under  $\Psi_i$  is also a subset of  $\{\text{ht}(\mathcal{P}_x) \geq h\}$ . With this observation, the natural modifications yield the desired.

The proof of the first inequality in Proposition 4.7 is more subtle as the increment intersecting  $\mathcal{L}_{h_1}$  is random. For each interface  $\mathcal{I}$  having  $|\mathcal{F}(\mathcal{S}_x \cap \mathcal{L}_{h_1})| \geq 4 + r$ , let  $\tau_1(\mathcal{I})$  denote its increment index such that  $\mathcal{X}_{\tau_{\text{sp}}+\tau_1}$  intersects  $\mathcal{L}_{h_1}$  non-trivially. Then let  $\Psi_{h_1}$  denote the map that for each  $\mathcal{I}$  having  $|\mathcal{F}(\mathcal{S}_x \cap \mathcal{L}_{h_1})| \geq 4 + r$ ,

is defined by  $\Psi_{h_1}(\mathcal{I}) = \Psi_{\tau_1(\mathcal{I})}(\mathcal{I})$ . Fix a  $T$ -admissible truncation  $\mathcal{I}_{\text{TR}}$  and express  $\mu_n(|\mathcal{F}(\mathcal{S}_x) \cap \mathcal{L}_{h_1}| \geq 4+r, \text{ht}(\mathcal{P}_x) \geq h, \mathcal{I}_{\text{TR}}, \bar{\mathbf{I}}_{x,T})$  as

$$\sum_{\substack{\mathcal{I} \in \bar{\mathbf{I}}_{x,T} \cap \text{ht}(\mathcal{P}_x) \geq h \cap \mathcal{I}_{\text{TR}} \\ |\mathcal{F}(\mathcal{S}_x \cap \mathcal{L}_{h_1})| \geq 4+r}} \mu_n(\mathcal{I}) = \sum_{k \geq r} \sum_{\mathcal{J} \in \Psi_{h_1}(\bar{\mathbf{I}}_{x,T} \cap \mathcal{I}_{\text{TR}} \cap \text{ht}(\mathcal{P}_x) \geq h)} \mu_n(\mathcal{J}) \sum_{\tau_1} \sum_{\substack{\mathcal{I} \in \Psi_{\tau_1}^{-1}(\mathcal{J}) \\ \mathcal{X}_{\tau_{\text{sp}}+\tau_1} \cap \mathcal{L}_{h_1} \neq \emptyset \\ \mathbf{m}(\mathcal{I}; \mathcal{J}) = k}} \frac{\mu_n(\mathcal{I})}{\mu_n(\Psi_{\tau_1}(\mathcal{I}))}$$

At this point, we notice that for each  $\mathcal{J}$ , and  $k \geq r$ , there are at most  $k$  possible choices of  $\tau_1$  such that  $\{\mathcal{I} : \mathcal{I} \in \Psi_{\tau_1}^{-1}(\mathcal{J}), \mathbf{m}(\mathcal{I}; \mathcal{J}) = k\}$  is non-empty. This is because, if  $X_{\tau_{\text{sp}}+\tau_1} \cap \mathcal{L}_{h_1} \neq \emptyset$  it must be that  $h_1 - v_{\tau_{\text{sp}}+\tau_1} \leq k$  (the excess area of the map is at least  $\text{ht}(v_{\tau_{\text{sp}}+\tau_1+1}) - \text{ht}(v_{\tau_{\text{sp}}+\tau_1})$ ). Reading off from  $\mathcal{J}$ , the increment index intersecting  $\mathcal{L}_{h_1-k-1}$ , one of the next  $k$  increment indices must be  $\tau_1$ . Combining this with Propositions 4.5 and Lemma 4.6, we see that this is at most

$$\sum_{\mathcal{J} \in \Psi_{h_1}(\bar{\mathbf{I}}_{x,T} \cap \mathcal{I}_{\text{TR}} \cap \text{ht}(\mathcal{P}_x) \geq h)} \mu_n(\mathcal{J}) \sum_{k \geq r} k s^k e^{-(\beta-C)k} \leq C \mu_n(\bar{\mathbf{I}}_{x,T} \cap \mathcal{I}_{\text{TR}} \cap \text{ht}(\mathcal{P}_x) \geq h) e^{-(\beta-C)r}.$$

since  $\Psi_{h_1}(\bar{\mathbf{I}}_{x,T} \cap \mathcal{I}_{\text{TR}} \cap \{\text{ht}(\mathcal{P}_x) \geq h\}) \subset (\bar{\mathbf{I}}_{x,T} \cap \mathcal{I}_{\text{TR}} \cap \{\text{ht}(\mathcal{P}_x) \geq h\})$ . Dividing both sides out by  $\mu_n(\bar{\mathbf{I}}_{x,T} \cap \mathcal{I}_{\text{TR}} \cap \text{ht}(\mathcal{P}_x) \geq h)$  then yields the desired.  $\blacksquare$

## 5. EXPONENTIAL TAILS ON THE BASE OF A PILLAR

In Section 4, we showed that the increments of the spine each have exponential tails on their excess areas. Here, we show that the groups of walls that constitute the interface apart from the spine, but are “near” the spine have excess area at most order  $O(\log T)$ . As a consequence, we see that  $|v_{\tau_{\text{sp}}} - x|$  is at most  $O(\log T)$  with high probability; this difference, and the base more generally, will be negligible as far as any  $T \rightarrow \infty$  limit theorems are concerned.

As before, take  $T$  to be large and take  $x$  to be a point in the “bulk” of  $\mathcal{L}_0 \cap \Lambda_{n,n,\infty}$ , e.g.,  $d(x, \partial\Lambda_n) \geq 100T$ .

**Proposition 5.1.** *There exists  $K, c > 0$  such that for every  $\beta > \beta_0$ , we have for every  $r \geq K \log T$ ,*

$$\mu_n(\text{ht}(v_{\tau_{\text{sp}}}) \geq r \mid \bar{\mathbf{I}}_{x,T}) \leq \exp[-c\beta r], \quad (5.1)$$

$$\mu_n(\text{diam}(\rho(\mathcal{B}_x)) \geq r \mid \bar{\mathbf{I}}_{x,T}) \leq \exp[-c\beta r]. \quad (5.2)$$

where the diameter  $\text{diam}(\rho(\mathcal{B}_x)) := \max_{x,y \in \rho(\mathcal{B}_x)} |x - y|$ . Finally, we can also deduce that for  $r \geq K \log T$ ,

$$\mu_n\left(\sum_{i \leq \tau_{\text{sp}}} \mathbf{m}(\mathcal{X}_i) \geq r \mid \bar{\mathbf{I}}_{x,T}\right) \leq \exp[-c\beta r]. \quad (5.3)$$

**Remark 5.2.** In particular, (5.1) implies that for every  $r \geq K \log T$ , with probability  $1 - e^{-c\beta r}$ , we have  $\tau_{\text{sp}} \leq r$ ; the bound (5.2) immediately implies the analogous bound on  $|\rho(v_{\tau_{\text{sp}}}) - x|$ .

**Remark 5.3.** Since we prove that the projection of the spine attains an order  $\sqrt{T}$  distance from  $x$ , there should be groups of walls of  $\mathcal{I} \setminus \mathcal{P}_x$  onto which  $\mathcal{P}_x$  projects, that attain an excess area  $c \log T$ ; thus the order of the bounds on  $\text{ht}(v_{\tau_{\text{sp}}})$  and  $\text{diam}(\rho(\mathcal{B}_x))$  is correct. On the other hand, we expect that  $\text{ht}(v_1)$  is order one, and already the increments starting from  $\mathcal{X}_1$  have exponential tails; the difficulty in proving this is in controlling the interactions of  $\mathcal{X}_1, \dots, \mathcal{X}_{\tau_{\text{sp}}}$  with nearby groups of walls which attain a higher height.

**5.1. The base reduction map  $\Phi_{\mathcal{B}}$ .** We first define a map that eliminates at least one group of walls of excess area larger than  $K \log T$  in  $\mathcal{C}_{v_{\tau_{\text{sp}}}, x, T}$ , and in doing so, allows one to lower the height of the source point for the spine. The map also shifts  $v_{\tau_{\text{sp}}}$  in the  $xy$ -plane, to lie above  $x$  if  $|\rho(v_{\tau_{\text{sp}}}) - x| \geq K \log T$ .

In order to study the impact of this map on  $\mu$ , it will help to formulate the base  $\mathcal{B}_x$  in terms of the definitions outlined in Section 2.2. In view of this, for any interface  $\mathcal{I} \in \mathbf{I}_{x,T}$ , recall from Definition 3.11 that it has a truncated interface  $\mathcal{I}_{\text{TR}}$ . We can define the groups of walls corresponding to  $\mathcal{I}_{\text{TR}}$  via Lemma 2.12. Fix some  $K$  sufficiently large with respect to all other constants that are independent of  $\beta$ . We refer to Figure 10 for a visualization of the map  $\Phi_{\mathcal{B}}$ .

**Definition 5.4.** Let  $\Phi_{\mathcal{B}} : \bar{\mathbf{I}}_{x,T} \rightarrow \mathbf{I}_{x,T}$  generate an interface  $\Phi_{\mathcal{B}}(\mathcal{I})$  from  $\mathcal{I}$  as follows. Suppose  $\mathcal{I}_{\text{TR}}$  has standard wall representation  $(W_z)_{z \in \mathcal{L}_0}$ , base  $\mathcal{B}_x$  and source point  $v_{\tau_{\text{sp}}}$ , and further suppose that its spine  $\mathcal{S}_x$  has increment collection  $(X_i)_{\tau_{\text{sp}} \leq i \leq T}$  and remainder  $X_{>T}$ . If  $\text{ht}(v_{\tau_{\text{sp}}}) \leq K \log T$  and  $\text{diam}(\rho(\mathcal{B}_x)) \leq K \log T$ ,

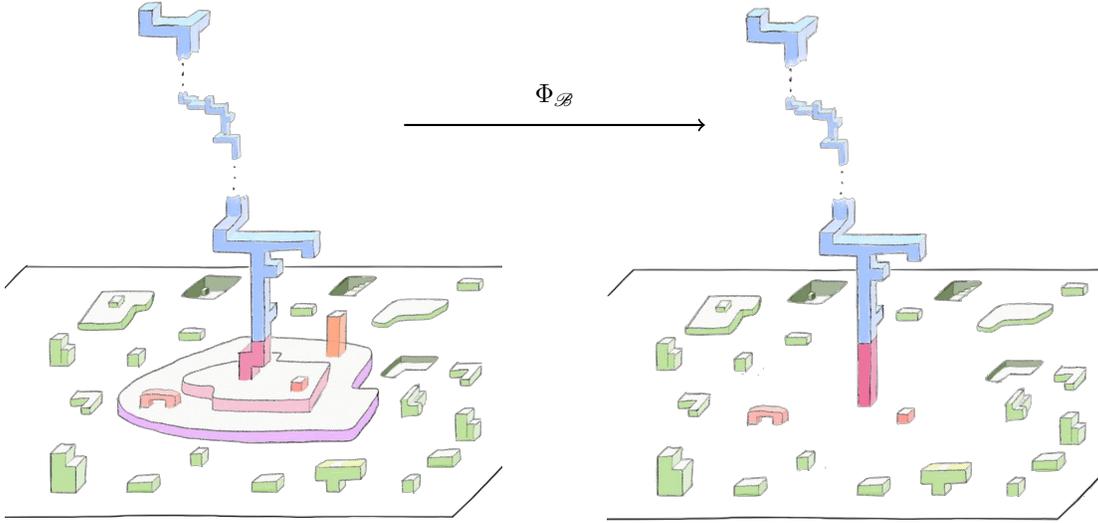


FIGURE 10. The base reduction map  $\Phi_{\mathcal{B}}$  sends the interface  $\mathcal{I}$  (left) to  $\Phi_{\mathcal{B}}(\mathcal{I})$  (right) by eliminating the nested sequence of walls  $\mathfrak{W}_{v_{\tau_{\text{sp}}}}$  (pink, left), along with the vertical column which attains the height  $v_{\tau_{\text{sp}}}$  (orange, left), replacing it by a straight vertical column above  $x$  (pink, right) and shifting the spine  $\mathcal{S}_x$  (blue, left) to lie above that column.

then the map is set to be trivial,  $\Phi_{\mathcal{B}}(\mathcal{I}) = \mathcal{I}$ . Otherwise, if  $\text{ht}(v_{\tau_{\text{sp}}}) > K \log T$  and/or  $\text{diam}(\rho(\mathcal{B})) > K \log T$ , then construct  $\mathcal{J} = \Phi_{\mathcal{B}}(\mathcal{I})$ , by

- (1) Mark in the standard wall representation  $(W_z)_{z \in \mathcal{L}_0}$  the groups of walls  $\mathfrak{F}_{v_{\tau_{\text{sp}}}}$ , as well as  $\mathfrak{F}_{[x]} = \bigcup_{f \in [x]} \mathfrak{W}_f$ , where  $[x] = \{x\} \cup \bigcup_{f \in \mathcal{L}_0: f \sim x} \{f\}$ .
- (2) If there exists some  $h < \text{ht}(v_{\tau_{\text{sp}}})$  such that the interface with standard wall representation  $\mathfrak{W}_{v_{\tau_{\text{sp}}}} \cup \mathfrak{W}_{[x]}$  intersects  $\mathcal{L}_h$  in exactly one plus cell, take the largest such height  $h^\dagger$  and let  $y^\dagger$  be an index in  $\mathcal{L}_0 \cap \mathcal{C}_{v_{\tau_{\text{sp}}}, x, T}$  of a wall that attains height  $h^\dagger$  and is not included in  $\mathfrak{W}_{[x]} \cup \mathfrak{W}_{v_{\tau_{\text{sp}}}}$ , and then mark  $\mathfrak{F}_{y^\dagger}$ . (If  $h < \text{ht}(v_{\tau_{\text{sp}}})$ , such a wall must exist by the definition of the source point  $v_{\tau_{\text{sp}}}$ .)
- (3) Remove all the marked walls, i.e.,  $\mathfrak{F}_{v_{\tau_{\text{sp}}}} \cup \mathfrak{F}_{[x]} \cup \mathfrak{F}_{y^\dagger}$  from the standard wall representation of  $\mathcal{I}_{\text{TR}}$ .
- (4) Let  $\mathcal{I}'$  be the interface with the resulting standard wall representation, let  $\mathbf{h} - 1$  be the height of a highest wall indexed by  $\mathcal{I}' \cap \mathcal{C}_{v_{\tau_{\text{sp}}}, x, T}$  and let  $h_\star := \mathbf{h} \vee (\text{ht}(v_{\tau_{\text{sp}}}) + \frac{1}{2})$ .
- (5) To that new interface, add the new standard wall consisting of the vertical bounding faces of a column of height  $\text{ht}(h_\star)$  above  $x$ : i.e., the cells  $x + (0, 0, \ell - \frac{1}{2}) : \ell = 1, \dots, h_\star$ . The resulting interface has a  $T$ -source point, which we will denote  $v_{\tau_{\text{sp}}}(\mathcal{J})$ .
- (6) Shift the spine  $\mathcal{S}_x$  by the vector  $x + (0, 0, h_\star - \frac{1}{2}) - v_{\tau_{\text{sp}}}$ ; i.e., the increment sequence of the new spine  $\mathcal{S}_x(\mathcal{J})$  sourced at  $v_{\tau_{\text{sp}}}(\mathcal{J})$  will be  $h_\star - \frac{1}{2} - \text{ht}(v_{\tau_{\text{sp}}}(\mathcal{J}))$  trivial increments, followed by the increment sequence of  $\mathcal{S}_x$ .

**5.2. Strategy of the map  $\Phi_{\mathcal{B}}$ .** As in §4.2, to obtain exponential tail bounds conditionally on  $\mathcal{P}_x$  having  $T$  increments and/or attaining height  $h$ , we require that  $\Phi_{\mathcal{B}}$  send  $\bar{\mathbf{I}}_{x, T} \cap \{\text{ht}(\mathcal{P}_x) \geq h\}$  to  $\mathbf{I}_{x, T} \cap \{\text{ht}(\mathcal{P}) \geq h\}$ .

Ideally, the map would replace the base  $\mathcal{B}_x$  with a single column of height  $\text{ht}(v_{\tau_{\text{sp}}})$  above  $x$ , and have an energy gain proportional to  $\text{diam}(\rho(\mathcal{B}_x))$  and  $\text{ht}(v_{\tau_{\text{sp}}})$ . Since the spine of  $\mathcal{I}$  starts at  $v_{\tau_{\text{sp}}}$  and not above  $x$ , we must additionally shift the spine to lie above  $x$ , so that together with the added column, it forms the new pillar  $\mathcal{P}_x^{\mathcal{J}}$ . We summarize the role the different steps above play in constructing such a map.

- (1) Step (1) above marks the nested sequence of walls supporting  $v_{\tau_{\text{sp}}}$ , and therefore attaining  $\text{ht}(v_{\tau_{\text{sp}}})$ , for deletion. It additionally marks the nested sequence of walls of  $x$ , to clear out space for the spine to be shifted horizontally and reattached above  $x$ .
- (2) If we only performed this first step of deletions, and then added back a column of height  $\text{ht}(v_{\tau_{\text{sp}}})$  above  $x$ , the map would have an excess area proportional to  $\text{diam}(\mathcal{B}_x)$  but not necessarily to  $\text{ht}(v_{\tau_{\text{sp}}})$ . In particular, if the groups of walls  $\mathfrak{F}_{v_{\tau_{\text{sp}}}}$  were mostly composed of trivial increments, so that its excess area is not much more than  $4 \text{ht}(v_{\tau_{\text{sp}}})$ , the energy gain would not be proportional to  $\text{ht}(v_{\tau_{\text{sp}}})$ .

In order to obtain an energy gain proportional to  $\text{ht}(v_{\tau_{\text{sp}}})$ , we note that by definition of  $v_{\tau_{\text{sp}}}$ , every height below  $\text{ht}(v_{\tau_{\text{sp}}})$  must have been intersected by at least two plus sites in  $\sigma(\mathcal{I}) \cap (\mathcal{L}_0 \cap \mathcal{C}_{v_{\tau_{\text{sp}}}, x, T})$ . Therefore, if  $\mathfrak{F}_{v_{\tau_{\text{sp}}}} \cup \mathfrak{F}_{[x]}$  did not contain excess energy larger than, say,  $6 \text{ht}(v_{\tau_{\text{sp}}})$ , there must exist another group of walls indexed by  $\mathcal{L}_0 \cap \mathcal{C}_{v_{\tau_{\text{sp}}}, x, T}$  whose excess energy is also comparable to  $\text{ht}(v_{\tau_{\text{sp}}})$ . Step 2 finds one such groups of walls and additionally marks it for deletion.

- (3) The remaining steps (3)–(6) then shift the spine to lie above  $x$ , and reconnect the shifted spine to  $x$  via a column of  $h_*$  plus sites. The reason  $\text{ht}(v_{\tau_{\text{sp}}})$  is possibly increased to height  $h_*$ , is the following: even though the map  $\Phi_{\mathcal{B}}$  only deletes walls of  $\mathcal{I}_{\text{TR}}$ , it could be that the deletion of such a wall *increases* the height of some  $W$  nested in a wall of  $\mathfrak{F}_{[x]} \cup \mathfrak{F}_{\rho(v_{\tau_{\text{sp}}})} \cup \mathfrak{F}_{y^\dagger}$ . The effect of this would be that placing the spine at  $x + (0, 0, \text{ht}(v_{\tau_{\text{sp}}}))$  could take the spine very close to the vertical shift of  $W$ , and their interactions could be large: in order to maintain the separation between the spine and the new truncated interface, we therefore place the spine at the higher height  $\mathbf{h}$ .

**5.3. Properties of the map  $\Phi_{\mathcal{B}}$ .** With the above remarks in place, we now establish, formally, the fact that  $\Phi_{\mathcal{B}}$  is well-defined, and show that it has an energy gain that is comparable to each of the quantities we prove exponential tails on, in (5.1)–(5.3).

**Lemma 5.5.** *The map  $\Phi_{\mathcal{B}} : \bar{\mathbf{I}}_{x, T} \rightarrow \mathbf{I}_{x, T}$  is well-defined, keeps the height of  $\mathcal{P}_x$  fixed, and,*

$$\mathbf{m}(\mathcal{I}; \Phi_{\mathcal{B}}(\mathcal{I})) = \mathbf{m}(\mathfrak{F}_{[x]} \cup \mathfrak{F}_{v_{\tau_{\text{sp}}}} \cup \mathfrak{F}_{y^\dagger}) - 4h_* \geq 2((h_* - 1) \vee \text{diam}(\rho(\mathcal{B}_x))) \vee \frac{1}{3} \sum_{i \leq \tau_{\text{sp}}} \mathbf{m}(X_i),$$

as long as the map  $\Phi_{\mathcal{B}}$  was nontrivial ( $\Phi_{\mathcal{B}}(\mathcal{I}) \neq \mathcal{I}$ ).

*Proof.* For the fact that  $\Phi_{\mathcal{B}}$  is well-defined, we observe that by definition of the source-point, for every height  $h < h(v_{\tau_{\text{sp}}})$ , the set of walls indexed by faces in  $\mathcal{C}_{v_{\tau_{\text{sp}}}, x, T}$  intersect that height in at least two plus sites; by Definition 2.14, any face  $f \in \mathcal{I}_{\text{TR}}$  is also attained by the interface corresponding to the nested sequence of walls  $\mathfrak{W}_{\rho(f)}$ . Thus, the sequence(s) described by steps (1)–(2) of Definition 5.4 exist. At step (5), since all walls projecting onto  $[x]$  have been removed, the standard wall being added by the vertical column of plus sites above  $x$ , maintains the admissibility of the remaining standard wall collection.

Since  $\mathcal{I} \in \bar{\mathbf{I}}_{x, T}$ , and  $\mathcal{C}_{R_0 T}(x) \subset \mathcal{C}_{v_{\tau_{\text{sp}}(x)}, x, T}$ , the resulting  $T$ -source point  $v_{\tau_{\text{sp}}}(\mathcal{J})$  has height at most  $h_*$ ; moreover, when we shift the spine of  $\mathcal{I}$  in (6), the resulting spine will be confined to  $\mathcal{C}_{R_0 T}(x) = \mathcal{C}_{v_{\tau_{\text{sp}}(\mathcal{J})}, x, T}$ ; finally, the new pillar  $\mathcal{P}_x(\Phi_{\mathcal{B}}(\mathcal{I}))$  has first  $h_* - 1 \geq \tau_{\text{sp}} - 1$  increments that are trivial, followed by  $T - \tau_{\text{sp}} + 1$  increments from  $\mathcal{S}_x$ , so it has at least  $T$  increments total. Thus, the map yields a valid interface  $\Phi_{\mathcal{B}}(\mathcal{I}) \in \mathbf{I}_{x, T}$ .

The lower bounds on the excess area follow from the following considerations. Since the marked sequences of walls in items (1)–(2) of Definition 5.4 intersected each height below  $\text{ht}(v_{\tau_{\text{sp}}})$  in at least two sites, we removed an excess of  $6(\text{ht}(v_{\tau_{\text{sp}}}) - \frac{1}{2})$  vertical faces from  $\mathcal{I}$ , and added back at most  $4(\text{ht}(v_{\tau_{\text{sp}}}) + \frac{1}{2})$  faces in step (5); the excess area  $\mathbf{m}(\mathcal{I}; \Phi_{\mathcal{B}}(\mathcal{I}))$  is at least the difference between these. If  $h_* = \text{ht}(v_{\tau_{\text{sp}}}) + \frac{1}{2}$ , then this implies  $\mathbf{m}(\mathcal{I}; \Phi_{\mathcal{B}}(\mathcal{I})) \geq 2(h_* - 1)$ , as desired. Now suppose otherwise that  $h_* = \mathbf{h} > \text{ht}(v_{\tau_{\text{sp}}}) + \frac{1}{2}$ . This could only have happened if there had been a wall nested inside one of  $\mathfrak{F}_{[x]} \cup \mathfrak{F}_{v_{\tau_{\text{sp}}}} \cup \mathfrak{F}_{y^\dagger}$  that was shifted vertically *upward* by at least  $\mathbf{h} - \frac{1}{2} - \text{ht}(v_{\tau_{\text{sp}}})$  upon deletion of  $\mathfrak{F}_{[x]} \cup \mathfrak{F}_{v_{\tau_{\text{sp}}}} \cup \mathfrak{F}_{y^\dagger}$  (the maximum height of that wall must have been below  $\text{ht}(v_{\tau_{\text{sp}}})$  in  $\mathcal{I}$ ). For such a vertical shift to be possible, its nesting sequence of walls must have had height at least  $(\mathbf{h} - \text{ht}(v_{\tau_{\text{sp}}}) - \frac{1}{2})$  and therefore excess area at least  $9(\mathbf{h} - \text{ht}(v_{\tau_{\text{sp}}}) - \frac{1}{2})$  faces (with the extreme case being a  $3 \times 3$  column with the nested wall in its center). On the other hand,  $4(\mathbf{h} - \text{ht}(v_{\tau_{\text{sp}}}) - \frac{1}{2})$  faces were added, so  $\mathbf{m}(\mathcal{I}; \Phi_{\mathcal{B}}(\mathcal{I})) \geq 2(h_* - 1)$  still holds.

Let us turn to the bound w.r.t.  $\text{diam}(\mathcal{B})$ . In order for  $v_{\tau_{\text{sp}}}$  to be in the pillar of  $x$ , there must be a wall in  $\mathcal{I}_{\text{TR}}$  containing both  $x$  and  $\rho(v_{\tau_{\text{sp}}})$  in its interior, which will be marked and removed in item (1) of Definition 5.4; in fact the maximal nested wall containing both  $v_{\tau_{\text{sp}}}$  and  $x$  bounding the entirety of the base is marked by item (1), resulting in an excess area of at least  $2 \text{diam}(\rho(\mathcal{B}_x)) \geq 2|\rho(v_{\tau_{\text{sp}}}) - x|$ .

Lastly, the fact that  $\mathbf{m}(\mathcal{I}; \Phi_{\mathcal{B}}(\mathcal{I})) \geq \frac{1}{3} \sum_{i \leq \tau_{\text{sp}}} \mathbf{m}(X_i)$  follows from the facts that all the increments  $(\mathcal{X}_1, \dots, \mathcal{X}_{\tau_{\text{sp}}})$  are part of the same wall, which contains  $v_{\tau_{\text{sp}}}$  in its interior, so it is removed, and replaced by a straight vertical column of the same height.  $\blacksquare$

**5.4. Estimating the effect of  $\Phi_{\mathcal{B}}$ .** We bound the change in probability under application of the map  $\Phi_{\mathcal{B}}$ . For ease of notation, locally in these sections we will simply denote this map by  $\Phi = \Phi_{\mathcal{B}}$ .

**Proposition 5.6.** *There exists  $C > 0$  such that the following holds for all  $\beta > \beta_0$ . For every  $\mathcal{I} \in \bar{\mathbf{I}}_{x,T}$  with spine increment sequence  $(X_i)_{\tau_{\text{sp}} \leq i \leq T}$  and remainder  $X_{>T}$ ,*

$$\left| \log \frac{\mu_n(\mathcal{I})}{\mu_n(\Phi(\mathcal{I}))} + \beta \mathbf{m}(\mathcal{I}; \Phi(\mathcal{I})) \right| \leq C \left[ \mathbf{m}(\mathcal{I}; \Phi(\mathcal{I})) + |\mathcal{F}(X_{>T})| e^{-\bar{c}(T+1-\tau_{\text{sp}})} + \sum_{\tau_{\text{sp}} \leq i \leq T} |\mathcal{F}(X_i)| e^{-\bar{c}(i-\tau_{\text{sp}})} \right].$$

*Proof.* Suppose that  $\mathcal{I} \in \bar{\mathbf{I}}_{x,T}$  with truncation  $\mathcal{I}_{\text{TR}}$ , increment sequence  $(X_i)_{i \leq T}$  and remainder  $X_{>T}$  and suppose that one of  $\text{ht}(v_{\tau_{\text{sp}}})$  or  $\text{diam}(\rho(\mathcal{B}_x))$  are at least  $K \log T$ , as otherwise the inequality is trivially satisfied. Set  $\mathcal{J} = \Phi(\mathcal{I})$  for ease of notation; by Theorem 2.21,

$$\frac{\mu_n(\mathcal{I})}{\mu_n(\mathcal{J})} = \exp \left( -\beta \mathbf{m}(\mathcal{I}; \mathcal{J}) + \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \mathcal{J}} \mathbf{g}(f', \mathcal{J}) \right).$$

We wish to bound the absolute difference between the sums above by the right-hand side of Proposition 5.6. We will decompose  $\mathcal{I}$  and  $\mathcal{J}$  into different subsets of faces, in order to pair up faces of  $\mathcal{I}$  with faces of  $\mathcal{J}$  that locally do not feel the effect of  $\Phi_{\mathcal{B}}$ . Let

$$\mathfrak{F} = \mathfrak{F}_{[x]} \cup \mathfrak{F}_{v_{\tau_{\text{sp}}}} \cup \mathfrak{F}_{y^\dagger} = (F_{x_s})_s \cup (F_{\rho(v_{\tau_{\text{sp}}})_s})_s \cup (F_{y_s^\dagger})_s,$$

be the nested sequences of groups of walls marked in steps (1)–(2) that were eliminated in step (3) of Definition 5.4 (indexed by  $[x], y^\dagger, v_{\tau_{\text{sp}}} \in \mathcal{L}_0 \cap \mathcal{C}_{v_{\tau_{\text{sp}}}, x, T}$ ). Sets  $\mathbf{E}, \mathbf{F}, \mathbf{G} \subset \mathcal{I} \cup \mathcal{J}$  will consist of all those faces that were removed from  $\mathcal{I}$  or added to  $\mathcal{J}$ :

- Let  $\mathbf{E}$  be the set of all faces in the groups of walls  $\mathfrak{F}$ ; these were removed in step (3) of  $\Phi_{\mathcal{B}}$ .
- Let  $\mathbf{F}$  be the set of  $f$  in  $\mathcal{J}$  such that  $\rho(f) \in \rho(\mathcal{F}(\mathbf{E}))$ , added in place of a removed face in  $\rho(\mathfrak{F})$  to “fill in” the interface.
- Let  $\mathbf{G}$  be the set of all other faces added to form  $\mathcal{J}$ , namely the single wall consisting of the bounding faces of a vertical column above  $x$  added in step (5) of  $\Phi_{\mathcal{B}}$ .

Also, for any  $f \in \mathcal{I}_{\text{TR}} \setminus \mathbf{E}$ , we set  $\tilde{f}$  to be the vertical translation of  $f$  as governed by the interface corresponding to the remaining walls after  $\mathbf{E}$  have been removed: see Observation 2.13. Finally, for every  $f \in \mathcal{S}_x = \mathcal{S}_x(\mathcal{I})$ , let  $\theta_{v_{\tau_{\text{sp}}}, x} f$  be its translation by  $x + (0, 0, h_\star - \frac{1}{2}) - v_{\tau_{\text{sp}}}$ . This decomposition allows us to expand,

$$\left| \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \mathcal{J}} \mathbf{g}(f', \mathcal{J}) \right| \leq \sum_{f \in \mathbf{E}} |\mathbf{g}(f, \mathcal{I})| + \sum_{f \in \mathbf{F}} |\mathbf{g}(f, \mathcal{J})| + \sum_{f \in \mathbf{G}} |\mathbf{g}(f, \mathcal{J})| \\ + \sum_{f \in \mathcal{S}_x(\mathcal{I})} \left| \mathbf{g}(f, \mathcal{I}) - \mathbf{g}(\theta_{v_{\tau_{\text{sp}}}, x} f, \mathcal{J}) \right| + \sum_{f \in \mathcal{I}_{\text{TR}} \setminus \mathbf{E}} \left| \mathbf{g}(f, \mathcal{I}) - \mathbf{g}(\tilde{f}, \mathcal{J}) \right|. \quad (5.4)$$

Let us begin with the first three terms, for which crude bounds suffice. By (2.2) and Lemma 5.5, there is a universal constant  $C > 0$  such that they are at most

$$\bar{K} (|\mathbf{E}| + |\mathbf{F}| + |\mathbf{G}|) \leq \bar{K} (2\mathbf{m}(\mathfrak{F}_{[x]} \cup \mathfrak{F}_{v_{\tau_{\text{sp}}}} \cup \mathfrak{F}_{y^\dagger}) + \mathbf{m}(\mathfrak{F}_{[x]} \cup \mathfrak{F}_{v_{\tau_{\text{sp}}}} \cup \mathfrak{F}_{y^\dagger}) + 4h_\star) \leq C\mathbf{m}(\mathcal{I}; \mathcal{J}).$$

Now, let us turn to the fourth term in (5.4), which encodes the contributions from the spine. Since the entire spine  $\mathcal{S}_x(\mathcal{I})$  is translated by the same vector  $x + (0, 0, h_\star - \frac{1}{2}) - v_{\tau_{\text{sp}}}$ , for every  $f \in \mathcal{S}_x(\mathcal{I})$ , the radius  $\mathbf{r}(f, \mathcal{I}; \theta_{v_{\tau_{\text{sp}}}, x} f, \mathcal{J})$  is attained either by a face at height at most  $\text{ht}(v_{\tau_{\text{sp}}}) - \frac{1}{2}$  in  $\mathcal{I}$  or at most  $h_\star - 1$  in  $\mathcal{J}$ , or by a face outside of  $\mathcal{C}_{v_{\tau_{\text{sp}}}, x, T}$ . However, since the increment sequence  $(X_i)_{\tau_{\text{sp}} \leq i \leq T}, X_{>T}$  is tame, and the height of the pillar is fixed by the map  $\Phi_{\mathcal{B}}$ , it must in fact be attained by a face in  $\mathcal{C}_{v_{\tau_{\text{sp}}}, x, T} \supset \mathcal{C}_{v_{\tau_{\text{sp}}(\mathcal{J})}, x, T}$  of height at most  $\text{ht}(v_{\tau_{\text{sp}}}) - \frac{1}{2}$  in  $\mathcal{I}$  or  $h_\star - 1$  in  $\mathcal{J}$ . The contribution from the fourth term in (5.4) is at most

$$\sum_{\tau_{\text{sp}} \leq i \leq T+1} \sum_{f \in \mathcal{F}(X_i)} \bar{K} \exp[-\bar{c}\mathbf{r}(f, \mathcal{I}; \theta_{v_{\tau_{\text{sp}}}, x} f, \mathcal{J})] \leq \sum_{\tau_{\text{sp}} \leq i \leq T+1} \sum_{f \in \mathcal{F}(X_i)} \bar{K} e^{-\bar{c}(\text{ht}(f) - \text{ht}(v_{\tau_{\text{sp}}}) - \frac{1}{2})} \\ \leq \sum_{\tau_{\text{sp}} \leq i \leq T+1} \bar{K} |\mathcal{F}(X_i)| e^{-\bar{c}(i-\tau_{\text{sp}})},$$

where we again used subscript “ $T+1$ ” to indicate  $> T$  here. (Notice that the radius  $\mathbf{r}(f, \mathcal{I}; \theta_{v_{\tau_{\text{sp}}}, x} f, \mathcal{J})$  is attained by a face whose height is at most  $\text{ht}(h_\star)$  and if  $h_\star \neq \text{ht}(v_{\tau_{\text{sp}}}) + \frac{1}{2}$ , then  $\theta_{v_{\tau_{\text{sp}}}, x}$  shifts the spine vertically accordingly, so that  $\text{ht}(\theta_{v_{\tau_{\text{sp}}}, x} f) - h_\star = \text{ht}(\theta_{v_{\tau_{\text{sp}}}, x} f) - \text{ht}(v_{\tau_{\text{sp}}}) - \frac{1}{2}$ .)

It remains to control the contribution from the interactions of the truncated pillar  $\mathcal{I}_{\text{TR}}$  with the application of the map  $\Phi_{\mathcal{B}}$ . The key idea here is that either they interact through the spine, in which case the contribution

is bounded as the above term, or they interact through the groups of walls in  $\mathcal{I}_{\text{TR}}$ , in which case they are controlled as in the proof of Lemma 2.25. To this end, let  $\tilde{\mathcal{J}}_{\text{TR}}$  be the image of the truncated interface under steps (1)–(3) of  $\Phi_{\emptyset}$ , prior to the addition of the faces in  $\mathbf{G}$ . Then, we can bound the difference

$$\sum_{f \in \mathcal{I}_{\text{TR}}, f \notin \mathbf{E}} \left| \mathbf{g}(f, \mathcal{I}) - \mathbf{g}(\tilde{f}, \mathcal{J}) \right| \leq \sum_{f \in \mathcal{I}_{\text{TR}}, f \notin \mathbf{E}} \bar{K} \exp \left[ -\bar{c} \mathbf{r}(f, \mathcal{I}; \tilde{f}, \mathcal{J}) \right]$$

by noticing that the distance  $\mathbf{r}(f, \mathcal{I}; \tilde{f}, \mathcal{J})$  is either attained by a face in  $\mathcal{S}_x$ , a face in  $\theta_{v_{\tau_{\text{SP}}}, x} \mathcal{S}_x$ , a face in the set  $\mathbf{G}$ , or is equal to  $\mathbf{r}(f, \mathcal{I}_{\text{TR}}; \tilde{f}, \tilde{\mathcal{J}}^{\text{tr}})$ . This lets us bound

$$\begin{aligned} \sum_{f \in \mathcal{I}_{\text{TR}} \setminus \mathbf{E}} \bar{K} e^{-\bar{c} \mathbf{r}(f, \mathcal{I}; \tilde{f}, \mathcal{J})} &\leq \sum_{f \in \mathcal{I}_{\text{TR}} \setminus \mathbf{E}} \sum_{g \in \mathcal{S}_x} \bar{K} [e^{-\bar{c} d(f, g)} + e^{-\bar{c} d(\tilde{f}, \theta_{v_{\tau_{\text{SP}}}, x} g)}] + \sum_{f \in \mathcal{I}_{\text{TR}} \setminus \mathbf{E}} \sum_{g \in \mathbf{G}} \bar{K} [e^{-\bar{c} d(f, g)} + e^{-\bar{c} d(\tilde{f}, g)}] \\ &+ \sum_{f \in \mathcal{I}_{\text{TR}} \setminus \mathbf{E}} \bar{K} e^{-\bar{c} \mathbf{r}(f, \mathcal{I}_{\text{TR}}; \tilde{f}, \tilde{\mathcal{J}}_{\text{TR}})}. \end{aligned} \quad (5.5)$$

As argued for  $f \in \mathcal{S}_x$ , the first term in the right-hand side of (5.5) can be bounded from above by

$$\begin{aligned} 2\bar{K} \left[ \sum_{\tau_{\text{SP}} \leq i \leq T} \sum_{g \in \mathcal{F}(X_i)} \sum_{f: d(f, g) \geq i - \tau_{\text{SP}}} e^{-\bar{c} d(f, g)} + \sum_{g \in \mathcal{F}(X_{>T})} \sum_{f: d(f, g) > T + 1 - \tau_{\text{SP}}} e^{-\bar{c} d(f, g)} \right] \\ \leq 2\bar{C} \left[ |\mathcal{F}(X_{>T})| e^{-\bar{c}(T + 1 - \tau_{\text{SP}})} + \sum_{\tau_{\text{SP}} \leq i \leq T} |\mathcal{F}(X_i)| e^{-\bar{c}(i - \tau_{\text{SP}})} \right], \end{aligned}$$

for some universal  $\bar{C}$ , where we used that  $\mathcal{S}_x$  is tame and the definition of the source point. The second term in (5.5) is trivially bounded above by

$$\sum_{f \in \mathcal{I}_{\text{TR}} \setminus \mathbf{E}} \sum_{g \in \mathbf{G}} \bar{K} [e^{-\bar{c} d(f, g)} + e^{-\bar{c} d(\tilde{f}, g)}] \leq 2 \sum_{g \in \mathbf{G}} \sum_{f \in \mathcal{F}(\mathbb{Z}^3)} \bar{K} e^{-\bar{c} d(g, f)} \leq 2\bar{K}\bar{C}|\mathbf{G}|.$$

By Lemma 5.5, we have that  $|\mathbf{G}| = 4h_* \leq 2\mathbf{m}(\mathcal{I}; \mathcal{J})$ . Finally, we bound the last term of (5.5) as in the proof of Lemma 2.25 and 3.3: by construction, for every  $f, \tilde{f}$  pair, the distance  $\mathbf{r}(f, \mathcal{I}_{\text{TR}}; \tilde{f}, \tilde{\mathcal{J}}^{\text{tr}})$  is attained by the distance to a wall face, and therefore, moving to the distance between projections,

$$\begin{aligned} \sum_{f \in \mathcal{I}_{\text{TR}} \setminus \mathbf{E}} \bar{K} \max_{u \in \rho(\mathfrak{F})} \exp[-\bar{c} d(\rho(f), u)] &\leq \sum_{u' \in \rho(\mathfrak{F})^c} \bar{K} N_\rho(u) \max_{u \in \rho(\mathfrak{F})} \exp[-\bar{c} d(u, u')] \\ &\leq \sum_{u' \in \rho(\mathfrak{F})^c} \sum_{u \in \rho(\mathfrak{F})} \bar{K} (|u - u'|^2 + 1) \exp[-\bar{c}|u - u'|]. \end{aligned}$$

By integrability of exponential tails, this is at most  $\bar{C}|\mathcal{F}(\rho(\mathfrak{F}))| + \bar{C}|\mathcal{E}(\rho(\mathfrak{F}))| \leq 2\bar{C}\mathbf{m}(\mathcal{I}; \mathcal{J})$  for some universal constant  $C$ . Combining all the above estimates concludes the proof.  $\blacksquare$

**5.5. Bounding the multiplicity of  $\Phi_{\emptyset}$ .** Here, we bound the multiplicity of the map  $\Phi_{\emptyset}$ . This is where we use the fact that the nested sequences of groups of walls we eliminated had excess area at least  $K \log T$ .

**Proposition 5.7.** *There exists  $s$  independent of  $\beta$  such that for every  $T$ , every  $\mathcal{J} \in \Phi_{\emptyset}(\bar{\mathbf{I}}_{x, T})$  and every  $k$ ,*

$$|\{\mathcal{I} \in \Phi_{\emptyset}^{-1}(\mathcal{J}) : \mathbf{m}(\mathcal{I}; \mathcal{J}) = k\}| \leq s^k.$$

*Proof.* If  $k \leq K \log T$ , the map  $\Phi_{\emptyset}$  must be the identity map and we must have  $k = 0$ , so the bound is trivially satisfied; therefore, suppose  $k \geq K \log T$ . In order to bound the multiplicity of the map, we will uniquely identify any pre-image  $\mathcal{I}$  with several collections of admissible groups of walls, indicating the nested sequence(s) of walls that are marked in steps (1)–(2) of Definition 5.4, along with their groups of walls. The requirement of  $k \geq K \log T$  will allow us to pick the centers of the nested sequence of walls from step (1), amongst the faces in  $\mathcal{L}_0$  that were in the cylinder  $\mathcal{C}_{v_{\tau_{\text{SP}}}, x, T}$ .

**Claim 5.8.** *Given  $\mathcal{J} \in \Phi_{\emptyset}(\bar{\mathbf{I}}_{x, T})$ , one can uniquely identify  $\mathcal{I} \in \Phi_{\emptyset}^{-1}(\mathcal{J})$  from*

- (1) the site  $v_{\tau_{\text{SP}}}$ ,
- (2) the groups of nested walls  $\mathfrak{F}^1 = \mathfrak{F}_{[x]}$ ,  $\mathfrak{F}^2 = \mathfrak{F}_{v_{\tau_{\text{SP}}}}$ ,
- (3) a group of nested walls  $\mathfrak{F}^3$  which is either empty if  $\mathfrak{F}^1 \cup \mathfrak{F}^2$  intersect every height below  $h(v_{\tau_{\text{SP}}})$  in more than one cell, or  $\mathfrak{F}_{y^\dagger}$  for some  $y^\dagger \in \mathcal{C}_{v_{\tau_{\text{SP}}}, x, T}$ .

**Proof of Claim 5.8.** To prove the claim, we reconstruct  $\mathcal{I}$  given this collection. First, in order to read-off  $h_*$ , we need to read  $\mathbf{h}$  by removing the pillar  $\mathcal{P}_x(\mathcal{J})$  from  $\mathcal{J}$  and finding the height of a highest wall in  $\mathcal{C}_{v_{\tau_{\text{sp}}}, x, T}$ . With  $h_*$  in hand, take the interface  $\mathcal{J}$ , and remove the set of faces  $\mathbf{G}$  from it by eliminating the column wall above  $x$  up to height  $h_*$ . This leaves a truncated interface  $\bar{\mathcal{J}}_{\text{TR}}$  along with a spine  $\mathcal{S}$ . By construction, this spine is exactly the spine  $\mathcal{S}_x$  of  $\mathcal{I}$  up to the translation  $v_{\tau_{\text{sp}}} - x - (0, 0, h_* - \frac{1}{2})$ . The truncated interface  $\bar{\mathcal{J}}_{\text{TR}}$  has a standard wall representation, to which we can add all standard walls in  $\mathfrak{F}^1 \cup \mathfrak{F}^2 \cup \mathfrak{F}^3$ . The resulting collection of standard walls is admissible and can then be mapped back to an interface by Lemma 2.12, which is exactly the truncated interface  $\mathcal{I}_{\text{TR}}$ . Appending the spine  $\mathcal{S}_x$  at  $v_{\tau_{\text{sp}}}$  yields  $\mathcal{I}$ .  $\blacksquare$

With the claim in hand, we begin by enumerating the number of choices we have for  $v_{\tau_{\text{sp}}}$ ; since  $\mathbf{m}(\mathcal{I}; \mathcal{J}) \geq 2\text{ht}(v_{\tau_{\text{sp}}})$  and  $2|\rho(v_{\tau_{\text{sp}}}) - x|$  by Lemma 5.5, there are at most  $k^3$  possible choices of  $v_{\tau_{\text{sp}}}$ .

We now wish to bound the number of possible pairs of (a) collections of groups of nested walls  $\mathfrak{F}_{[x]}$  and  $\mathfrak{F}_{v_{\tau_{\text{sp}}}}$ , and (b) collections of groups of walls corresponding to the nested walls  $\mathfrak{F}_{y^\dagger}$ , indexed by some face  $y^\dagger$  in  $\mathcal{L}_0 \cap \mathcal{C}_{v_{\tau_{\text{sp}}}, x, T}$ . Take the at most three sequences of nested walls identified by steps (1)–(2) in Definition 5.4 of  $\mathcal{I}$ , denote them by  $(W_i^a)_i$  with groups of walls  $\mathfrak{F}^a = (F_j^a)_j$  for  $a \in \{1, 2, 3\}$  (so that  $W_i^a$  is nested in  $W_{i+1}^a$ ). One can generate a  $*$ -connected set of faces out of each such sequence as follows:

- (1) Assign to each point in  $u \in \rho(\bigcup_j F_j^a)$  the set  $R_u$  of faces in  $\mathcal{L}_0$  a distance at most  $\sqrt{N_\rho(u)}$  from  $u$ ,
- (2) For every wall  $W_i^a$  nested in  $W_{i+1}^a$ , assign to it the set  $R_{i+1}^a$  of minimal collection of faces in  $\mathcal{L}_0$  connecting  $W_i^a$  to  $W_{i+1}^a$ . In the case of  $\mathfrak{F}_{[x]}$  (resp.,  $\mathfrak{F}_{v_{\tau_{\text{sp}}}}$ , or  $\mathfrak{F}_{y^\dagger}$ ) include also the faces of  $[x]$  (resp.,  $(\rho(v_{\tau_{\text{sp}}}))$  and  $y^\dagger$ ) and connect them via a shortest path of faces  $R_1^a$  to  $W_1^a$ .

The union of the groups of walls  $\mathfrak{F}^a$ , along with the face sets  $\bigcup_{u \in \rho(\mathfrak{F}^a)} R_u$  and  $\bigcup_i R_i^a$  is a  $*$ -connected set of faces by the definition of groups of walls, and by construction. Moreover, given this union, one can recover the set  $\mathfrak{F}^a$  because any face  $f \in \mathcal{L}_0$  in this union is in  $\mathfrak{F}^a$  if and only if another face in the union projects onto it (otherwise it couldn't be a wall face). The cardinality of this union of faces is bounded above by

$$\sum_{a \in \{1, 2, 3\}} \left[ |\mathfrak{F}^a| + \sum_{u \in \rho(\mathfrak{F}^a)} |R_u^a| + \sum_i |R_i^a| \right].$$

By Remark 2.20,  $|\mathfrak{F}^a| \leq 2\mathbf{m}(\mathfrak{F}^a)$  for each  $i, a$ . By construction, and the nesting of walls,  $\sum_i |R_i^a| \leq \sum_i \mathbf{m}(W_i^a) \leq \mathbf{m}(\mathfrak{F}^a)$ . Finally, by definition of groups of walls,

$$\sum_i \sum_{u \in \rho(F_i^a)} |R_u^a| \leq \sum_i \sum_{u \in \rho(F_i^a)} N_\rho(u) \leq 2\mathbf{m}(\mathfrak{F}^a).$$

Because  $\mathbf{m}(\mathfrak{F}^a) \leq 2\mathbf{m}(\mathcal{I}; \mathcal{J})$ , we see that for each  $a \in \{1, 2, 3\}$ , the union described above is a connected collection of at most  $12\mathbf{m}(\mathcal{I}; \mathcal{J})$  faces rooted at some specific face ( $x$  or  $\rho(v_{\tau_{\text{sp}}})$  in the cases  $a = 1, 2$ ). Therefore, the number of possible such collections of groups of walls of nested walls associated to  $\mathcal{I} \in \Phi^{-1}(\mathcal{J}) : \mathbf{m}(\mathcal{I}; \mathcal{J}) = k$  is bounded as follows: pick an origin  $y^\dagger \in \mathcal{L}_0 \cap \mathcal{C}_{v_{\tau_{\text{sp}}}, x, T}$ , pick  $0 \leq k_1, k_2, k_3 \leq k$  and then finally, to each of  $[x], v_{\tau_{\text{sp}}}, y^\dagger$ , associate a connected group of faces of size at most  $k_a$ . The number of choices of origin  $y^\dagger$  is at most the size of  $\mathcal{L}_0 \cap \mathcal{C}_{v_{\tau_{\text{sp}}}, x, T}$ , which is at most  $(R_0 T + k)^2$ . The number of total such choices is then easily seen to be at most  $(R_0 T + k)^2 k^3 s^k$ , which is at most exponential in  $k$  as long as  $k \geq K \log T$  for some large  $K$  to make the  $O(T^2)$  term negligible.  $\blacksquare$

**Proof of Proposition 5.1.** By Lemma 5.5, the event that  $\mathcal{I}$  has  $\text{ht}(v_{\tau_{\text{sp}}}) \geq r$  implies that  $\mathbf{m}(\mathcal{I}; \Phi_{\mathcal{B}}(\mathcal{I})) \geq r$ , and similarly, the event that  $\text{diam}(\mathcal{B}_x) \geq r$  implies that  $\mathbf{m}(\mathcal{I}; \Phi_{\mathcal{B}}(\mathcal{I})) \geq r$ . As such, let us fix an  $r \geq K \log T$ ; for ease of notation, let us denote, for the rest of this section,

$$\Gamma_r = \left\{ \mathcal{I} \in \bar{\mathcal{I}}_{x, T} : |\mathcal{F}(\mathcal{X}_{>T})| e^{-\bar{c}(T+1-\tau_{\text{sp}})} + \sum_{\tau_{\text{sp}} \leq i \leq T} |\mathcal{F}(\mathcal{X}_i^c)| e^{-\bar{c}(i-\tau_{\text{sp}})} < r \right\}$$

Then we can write

$$\mu_n(\mathbf{m}(\mathcal{I}; \Phi(\mathcal{I})) \geq r \mid \bar{\mathcal{I}}_{x, T}) \leq \mu_n(\mathbf{m}(\mathcal{I}; \Phi(\mathcal{I})) \geq r, \Gamma_r^c \mid \bar{\mathcal{I}}_{x, T}) + \mu_n(\mathbf{m}(\mathcal{I}; \Phi(\mathcal{I})) \geq r, \Gamma_r \mid \bar{\mathcal{I}}_{x, T}).$$

By Corollary 4.2, the first quantity on the right-hand side is at most  $C \exp(-\beta r/C)$  for some universal  $C$ . The latter quantity can be bounded as follows by Propositions 5.6–5.7:

$$\begin{aligned} \sum_{\substack{\mathcal{I} \in \bar{\mathbf{I}}_{x,T} \cap \Gamma_r \\ \mathbf{m}(\mathcal{I}; \Phi(\mathcal{I})) \geq r}} \mu_n(\mathcal{I}) &= \sum_{\mathcal{J} \in \Phi(\bar{\mathbf{I}}_{x,T} \cap \Gamma_r)} \sum_{k \geq r} \sum_{\substack{\mathcal{I} \in \Phi^{-1}(\mathcal{J}) \\ \mathbf{m}(\mathcal{I}; \Phi(\mathcal{I})) = k}} \frac{\mu_n(\mathcal{I})}{\mu_n(\Phi(\mathcal{I}))} \mu_n(\mathcal{J}) \\ &\leq \sum_{\mathcal{J} \in \Phi(\bar{\mathbf{I}}_{x,T} \cap \Gamma_r)} \mu_n(\mathcal{J}) \sum_{k \geq r} s^k \exp[-(\beta - C)k]. \end{aligned}$$

for some other universal constant  $C$  (where we absorbed the contribution from the increments to the right-hand side of Proposition 5.6, which was at most an extra  $r$ , into the  $C$ ). By integrability of exponential tails, and the fact that  $\mu_n(\Phi(\bar{\mathbf{I}}_{x,T} \cap \Gamma_r)) \leq \mu_n(\mathbf{I}_{x,T}) \leq 2\mu_n(\bar{\mathbf{I}}_{x,T})$  by Lemma 3.15, we have

$$\mu_n(\mathbf{m}(\mathcal{I}; \Phi(\mathcal{I})) \geq r \mid \bar{\mathbf{I}}_{x,T}) \leq 2e^{-(\beta-C)r} \mu_n(\bar{\mathbf{I}}_{x,T}).$$

Dividing both sides by  $\mu_n(\bar{\mathbf{I}}_{x,T})$  concludes the proof.  $\blacksquare$

As we did in Proposition 4.7, since the map  $\Phi_{\mathcal{B}}$  keeps the height of the pillar  $\mathcal{P}_x$  fixed, we can also prove the estimates of Proposition 5.1.

**Proposition 5.9.** *There exists  $K, c > 0$  such that for every  $\beta > \beta_0$ , we have for every  $r \geq K \log h$ , every  $T \in [\frac{h}{2}, h]$ ,*

$$\begin{aligned} \mu_n(\text{ht}(v_{\tau_{\text{sp}}}) \geq r \mid \text{ht}(\mathcal{P}_x) \geq r, \bar{\mathbf{I}}_{x,T}) &\leq \exp(-c\beta r), \\ \mu_n(\text{diam}(\rho(\mathcal{B}_x)) \geq r \mid \text{ht}(\mathcal{P}_x) \geq r, \bar{\mathbf{I}}_{x,T}) &\leq \exp(-c\beta r). \end{aligned}$$

*Proof.* The proof is again analogous to the proof of Proposition 4.7 and we therefore do not include all the details. For any  $h$  and any  $T \leq h$ , we can expand as above,

$$\begin{aligned} \mu_n(\mathbf{m}(\mathcal{I}; \Phi(\mathcal{I})) \geq r \mid \text{ht}(\mathcal{P}_x) \geq h, \bar{\mathbf{I}}_{x,T}) &\leq \mu_n(\mathbf{m}(\mathcal{I}; \Phi(\mathcal{I})) \geq r, \Gamma_r^c \mid \text{ht}(\mathcal{P}_x) \geq h, \bar{\mathbf{I}}_{x,T}) \\ &\quad + \mu_n(\mathbf{m}(\mathcal{I}; \Phi(\mathcal{I})) \geq r, \Gamma_r \mid \text{ht}(\mathcal{P}_x) \geq h, \bar{\mathbf{I}}_{x,T}). \end{aligned}$$

The same estimate on the first term on the right-hand side holds from the conditional estimate (4.3) of Proposition 4.7, and the analogue of Corollary 4.2 under the measure that also conditions on  $\{\text{ht}(\mathcal{P}_x) \geq h\}$ , by taking a supremum over all truncated interfaces  $\mathcal{I}_{\text{TR}}$ , and noting that the exponential tails on spine increments are uniform in  $\mathcal{I}_{\text{TR}}$ . The second term on the right-hand side, we also bound as in the proof of Proposition 5.6, summing only over interfaces that also have the property that  $\{\text{ht}(\mathcal{P}_x) \geq h\}$ , and using that  $\mu_n(\Phi(\bar{\mathbf{I}}_{x,T})) \leq \mu_n(\text{ht}(\mathcal{P}_x) \geq h, \mathbf{I}_{x,T}) \leq 2\mu_n(\text{ht}(\mathcal{P}_x) \geq h, \bar{\mathbf{I}}_{x,T})$  as long as  $T \geq \frac{h}{2}$  by (3.3) of Remark 3.16. Following the rest of the proof with these modifications yields the desired estimates.  $\blacksquare$

## 6. LARGE DEVIATION RATE AND LAW OF LARGE NUMBERS FOR THE MAXIMUM

In this section, we use the results of Sections 3–5, to prove Theorem 1. We begin, in Section 6.1, with a rough equivalence between pillars and groups of walls, and recall early decorrelation estimates of Dobrushin [23] for groups of walls in the bulk of  $\Lambda_n$ . In Section 6.2, we show the existence of the limiting large deviation rate for the event  $\{\text{ht}(\mathcal{P}_x) \geq h\}$  and relate it to an infinite-volume large deviation rate under the measure  $\mu_{\mathbb{Z}^3}^{\mp}$ . The key estimate there will be the following:

**Proposition 6.1.** *The limit  $\alpha_\beta$  given by (1.2) exists and moreover, there exists  $\beta_0$  such that for all  $\beta > \beta_0$ , for every sequence  $h = h_n$  such that  $1 \ll h \ll n$  and every  $x = x_n \in \mathcal{L}_0 \cap \Lambda_{h_n}$  such that  $d(x_n, \partial\Lambda_n) \gg h_n$ ,*

$$\lim_{n \rightarrow \infty} -\frac{1}{h_n} \log \mu_n(\text{ht}(\mathcal{P}_{x_n}) \geq h_n) = \alpha_\beta. \quad (6.1)$$

*As a consequence, the quantity  $\alpha_\beta \in [4\beta - C, 4\beta + e^{-4\beta}]$  for a universal constant  $C$ .*

In Section 6.3, we use the decorrelation estimates for pillars and the existence of this large deviation rate to show that the maximum height of an interface satisfies a law of large numbers.

**6.1. Decorrelation estimates for groups of walls and pillars.** In this section, we use the decomposition of pillars into a base and a spine, and in particular, the exponential tail on the size of the base proved in Section 5, to show that the structure of a pillar is, with high probability, captured by the groups of walls indexed by faces within a  $o(T)$  neighborhood of  $x$ . We use this to translate decorrelation estimates for groups of walls into decorrelation estimates for pillars.

The following is then an immediate corollary of Eq. (2.4).

**Proposition 6.2.** *With  $\mu_n$ -probability  $1 - O(e^{-c\beta r})$ , the nested sequence of walls  $\mathfrak{W}_x$  is indexed by faces a distance at most  $r$  from  $x$ , (and therefore so are all walls nested in a wall of  $\mathfrak{W}_x$ ).*

With these equivalences in mind, we recall some decorrelation estimates for groups of walls proved by Dobrushin in [23, 27]. The first of these says that the dependence of the law of a group of walls  $F_x$  on the containing box size  $n$  decays exponentially fast in the distance between  $x$  and  $\partial\Lambda_n$ . When combined with Proposition 6.2, this will imply that the law of the pillar above a face in  $\mathcal{L}_0$  (approximately) does not depend on the side-length  $n$  or on the position of  $x$ , as long as  $x$  is sufficiently far from  $\partial\Lambda_n$ .

**Proposition 6.3** ([27], [23, Lemma 5], as well as [12, Prop. 2.3]). *There is a  $C > 0$  such that for every  $\beta > \beta_0$ , every  $n \leq m$ , for a sequence of  $x = x_n \in \mathcal{L}_0 \cap \Lambda_n$ ,*

$$\left\| \mu_n((\mathcal{F}_y)_{y:|y-x| \leq r} \in \cdot) - \mu_m((\mathcal{F}_y)_{y:|y-x| \leq r} \in \cdot) \right\|_{\text{TV}} \leq C \exp(- (d(x, \partial\Lambda_n) - r)/C).$$

*In particular, sending  $m$  to  $\infty$ , and using tightness of  $(\mathcal{F}_y)_y$ , this estimate holds if we replace  $\mu_m$  by  $\mu_{\mathbb{Z}^3}^\mp$ .*

**Corollary 6.4.** *There is a  $C > 0$  such that for every  $\beta > \beta_0$ , every  $n, m$  and two sequences  $x = x_n$  and  $y = y_m$  such that  $d(x, \partial\Lambda_n) \wedge d(y, \partial\Lambda_m) \geq r$ ,*

$$\left\| \mu_n(\mathcal{P}_x \in \cdot) - \mu_m(\mathcal{P}_y \in \cdot) \right\|_{\text{TV}} \leq C \exp[-r/C].$$

*Proof.* For any interface, with a standard wall representation  $(F_y)_{y \in \mathcal{L}_0}$ , we can set  $\mathcal{I}^{(R), x}$  which is the interface having only groups of walls indexed by  $y : |y - x| < R$  and let  $\mathcal{P}_x^{(R)}$  be the pillar of  $x$  in the interface  $\mathcal{I}^{(R), x}$ . By Observation 2.17 and Proposition 6.2, with probability  $1 - O(e^{-c\beta R})$ , the pillars  $\mathcal{P}_x$  and  $\mathcal{P}_x^{(R)}$  are equal. Take an  $N$  large which we will send to infinity, and expand the difference

$$\begin{aligned} \left\| \mu_n(\mathcal{P}_x \in \cdot) - \mu_m(\mathcal{P}_y \in \cdot) \right\|_{\text{TV}} &\leq \left\| \mu_n(\mathcal{P}_x \in \cdot) - \mu_N(\mathcal{P}_x \in \cdot) \right\|_{\text{TV}} + \left\| \mu_N(\mathcal{P}_x \in \cdot) - \mu_N(\mathcal{P}_y \in \cdot) \right\|_{\text{TV}} \\ &\quad + \left\| \mu_m(\mathcal{P}_y \in \cdot) - \mu_N(\mathcal{P}_y \in \cdot) \right\|_{\text{TV}}. \end{aligned}$$

The first term above is bounded as follows: there exists  $C > 0$  such that for all  $\beta > \beta_0$ ,

$$\begin{aligned} \left\| \mu_n((\mathcal{F}_y)_{y:|y-x| \leq r} \in \cdot) - \mu_N((\mathcal{F}_y)_{y:|y-x| \leq r} \in \cdot) \right\|_{\text{TV}} + \mu_n(\mathcal{P}_x \neq \mathcal{P}_x^{(r)}) + \mu_N(\mathcal{P}_x \neq \mathcal{P}_x^{(r)}) \\ \leq C e^{-(d(x, \partial\Lambda_n) - r)/C} + e^{-\beta r/C}, \end{aligned}$$

as if  $\mathcal{P}_x$  is contained in the ball of radius  $r$  around  $x$ , then the pillar  $\mathcal{P}_x$  is a marginal of the collection of groups of walls  $(\mathcal{F}_y)_{y:|y-x| \leq r}$ . The third term is bounded analogously. In order to bound the second term,

$$\left\| \mu_N(\mathcal{P}_x \in \cdot) - \mu_N(\mathcal{P}_y \in \cdot) \right\|_{\text{TV}} \leq \left\| \mu_N(\mathcal{P}_x^{(R)} \in \cdot) - \mu_N(\mathcal{P}_y^{(R)} \in \cdot) \right\|_{\text{TV}} + 2e^{-R\beta/C}.$$

Taking  $N \rightarrow \infty$ , the first term on the right-hand side here vanishes as the infinite-volume measure  $\mu_{\mathbb{Z}^3}^\mp$  is invariant under translations in the  $xy$ -directions [23]. Sending  $N \rightarrow \infty$  first, then  $R \rightarrow \infty$ , and replacing  $r$  by say  $2r$ , we obtain the desired inequality.  $\blacksquare$

We also mention a result of Dobrushin showing that groups of walls decorrelate exponentially fast in their distance. That they decorrelate exponentially fast conditionally on the other groups of walls of the interface follows relatively straightforwardly from the cluster expansion and definition of groups of walls—however, a powerful bound of Dobrushin from [23, 26] allows one to translate conditional decorrelation estimates for random fields to unconditional ones. This estimate, together with the equivalence of groups of walls and pillars, greatly simplifies the second moment estimate in Section 6.3.

**Proposition 6.5** ([23], see also [12, Proposition 2.1]). *There is a  $C > 0$  such that for every  $\beta > \beta_0$ , every  $n$  and two sequences  $x = x_n$  and  $y = y_n$ ,*

$$\left\| \mu_n((\mathcal{F}_s)_{|s-x| < r} \in \cdot, (\mathcal{F}_t)_{|t-y| < r} \in \cdot) - \mu_n((\mathcal{F}_s)_{|s-x| < r} \in \cdot) \mu_n((\mathcal{F}_t)_{|t-y| < r} \in \cdot) \right\|_{\text{TV}} \leq C e^{-(|x-y| - 2r)/C}.$$

**Corollary 6.6.** *There is a  $C > 0$  such that for every  $\beta > \beta_0$ , every  $n$  and every two sequences  $x = x_n$  and  $y = y_n$  such that  $d(x, y) \geq r$ , we have*

$$\|\mu_n(\mathcal{P}_x \in \cdot, \mathcal{P}_y \in \cdot) - \mu_n(\mathcal{P}_x \in \cdot)\mu_n(\mathcal{P}_y \in \cdot)\|_{\text{TV}} \leq C \exp[-r/C].$$

*Proof.* Fix an  $r$ , recall the definition of  $\mathcal{P}_x^{(r)}$  and  $\mathcal{P}_y^{(r)}$ , and use the shorthand  $(\mathcal{F}_s)$  and  $(\mathcal{F}_t)$  for  $(\mathcal{F}_s)_{|s-x|<r}$  and  $(\mathcal{F}_t)_{|t-y|<r}$ . Then,

$$\begin{aligned} \|\mu_n(\mathcal{P}_x \in \cdot, \mathcal{P}_y \in \cdot) - \mu_n(\mathcal{P}_x \in \cdot)\mu_n(\mathcal{P}_y \in \cdot)\|_{\text{TV}} &\leq \|\mu_n(\mathcal{P}_x \in \cdot, \mathcal{P}_y \in \cdot) - \mu_n(\mathcal{P}_x^{(r)} \in \cdot, \mathcal{P}_y^{(r)} \in \cdot)\|_{\text{TV}} \\ &\quad + \|\mu_n((\mathcal{F}_s) \in \cdot, (\mathcal{F}_t) \in \cdot) - \mu_n((\mathcal{F}_s) \in \cdot)\mu_n((\mathcal{F}_t) \in \cdot)\|_{\text{TV}} \\ &\quad + \|\mu_n(\mathcal{P}_x^{(r)} \in \cdot)\mu_n(\mathcal{P}_y^{(r)} \in \cdot) - \mu_n(\mathcal{P}_x \in \cdot)\mu_n(\mathcal{P}_y \in \cdot)\|_{\text{TV}} \end{aligned}$$

where the second term is as it is because  $\mathcal{P}_x^{(r)}$  is a marginal of  $(\mathcal{F}_s)$  and  $\mathcal{P}_y^{(r)}$  is a marginal of  $(\mathcal{F}_t)$ . The second term above, then, is exactly the quantity bounded by Proposition 6.5. The first and third terms are bounded by  $\exp(-r/C)$  by Observation 2.17 and Proposition 6.2, yielding the desired.  $\blacksquare$

**6.2. Limiting large deviation rate.** In this section, we use an approximate sub-additivity argument to demonstrate the existence of a limiting large deviation rate for the probability that  $\text{ht}(\mathcal{P}_x)$  exceeds  $h$  as  $h \rightarrow \infty$ . We will first show how Proposition 6.1 follows from the existence of the limit in (1.2); we then prove the existence of the limit in (1.2) leveraging the fact that connection events are increasing, to use the monotonicity and FKG property of the Ising model.

Let us begin by proving Proposition 6.1 given the existence of the limit in (1.2). Without loss, we will change from sequences indexed by  $n$  to sequences indexed by  $h$ , so that  $n_h$  is any sequence having  $n_h \gg h$  and  $x_h$  is such that  $d(x_h, \partial\Lambda_{n_h}) \gg h$ . Recall that  $x \xrightarrow{+}_A y$  denotes that there is a  $*$ -connected path of  $+$  sites in  $\mathcal{C}(A)$  between  $x$  and  $y$ . Let us denote by  $A_h$  the event, measurable with respect to the configuration on  $\mathcal{C}(\mathbb{Z}^2 \times \llbracket 0, h \rrbracket)$ , defined by

$$A_h = A_h^x = \left\{ \sigma : x + (0, 0, \frac{1}{2}) \xrightarrow{+}_{\mathbb{Z}^2 \times \llbracket 0, h \rrbracket} \mathcal{L}_{h-\frac{1}{2}} \right\}.$$

We will show that the limit in (6.1) is equal to the following limit

$$\lim_{h \rightarrow \infty} -\frac{1}{h} \log \mu_{n_h}(A_h^{x_h}), \quad (6.2)$$

which we will show exists and equals the infinite-volume limit  $\alpha_\beta$  defined in (1.2).

**Proof of Proposition 6.1, given existence of (1.2).** For every  $n$  large, every  $x \in \mathcal{L}_0 \cap \Lambda_n$ , we claim that we have the comparability of events: there exists  $\varepsilon_\beta$  vanishing as  $\beta \rightarrow \infty$  such that

$$(1 - \varepsilon_\beta)\mu_n(A_h^x) \leq \mu_n(\text{ht}(\mathcal{P}_x) \geq h) \leq (1 + \varepsilon_\beta)\mu_n(A_h^x). \quad (6.3)$$

(This indicates that the connectivity event  $A_h^x$  serves as a good proxy for the relevant event  $\{\text{ht}(\mathcal{P}_x) \geq h\}$ : refer to Figure 5 for examples of configurations in  $\{\text{ht}(\mathcal{P}_x) \geq h\}^c \cap A_h^x$  (left) and  $\{\text{ht}(\mathcal{P}_x) \geq h\} \cap (A_h^x)^c$ .)

Letting  $A_h = A_h^x$ , on the one hand, by Definitions 2.16–2.18, we have

$$A_h \cap \{\text{ht}(\mathcal{P}_x) \geq 0\} \subset \{\text{ht}(\mathcal{P}_x) \geq h\};$$

since  $\mathfrak{F}_x = \emptyset$  implies  $\text{ht}(\mathcal{P}_x) \geq 0$ , by (2.4),  $\mu_n(\text{ht}(\mathcal{P}_x) \geq 0) \geq 1 - \varepsilon_\beta$ , and the FKG inequality implies the left-hand side of (6.3). On the other hand, given  $\{\text{ht}(\mathcal{P}_x) \geq h\}$ , the event  $(A_h^x)^c$  implies that there is a (nearest-neighbor) connected component of minuses separating  $x + (0, 0, \frac{1}{2})$  from the inner boundary of the pillar, and in particular, from height  $h$ , in the slab  $\mathcal{C}(\mathbb{Z}^2 \times \llbracket 0, h \rrbracket)$ . If  $\text{ht}(\mathcal{P}_x) \geq h \geq 1$ , this is in the plus phase of the Ising model with interface  $\mathcal{I}$ , and thus the probability of such a half-bubble of minuses is at most the probability that  $x + (0, 0, \frac{1}{2})$  is not  $*$ -connected by plus sites to  $\infty$  in  $\mathcal{C}(\mathbb{Z}^2 \times \llbracket 0, \infty \rrbracket)$  under  $\mu_{\mathbb{Z}^3}^+$ ; this probability is in turn at most  $\varepsilon_\beta$  by a classical Peierls argument. Thus, we can express

$$\begin{aligned} \mu_n(\text{ht}(\mathcal{P}_x) \geq h) &= \mu_n(\text{ht}(\mathcal{P}_x) \geq h, A_h^x) + \mu_n(\text{ht}(\mathcal{P}_x) \geq h, (A_h^x)^c) \\ &\leq \mu_n(A_h^x) + \varepsilon_\beta \mu_n(\text{ht}(\mathcal{P}_x) \geq h), \end{aligned}$$

from which the right-hand side of (6.3) follows. It remains to show that the limit (6.2) is given by  $\alpha_\beta$ .

By Corollary 6.4 and the fact that the distance from  $x_h$  to the boundary grows faster than  $h$ , if we show (6.1) for one such sequence of  $x_h$ , it implies it for every such sequence (the error  $e^{-cd(x_h, \partial\Lambda_{n_h})}$  vanishes

after taking a logarithm, dividing by  $h$  and sending  $h \rightarrow \infty$ ). Now take a fixed  $x$ , say  $(\frac{1}{2}, \frac{1}{2}, 0)$  and any two sequences  $n_h$  and  $m_h$  such that  $\frac{n_h}{h}$  and  $\frac{m_h}{h}$  go to infinity. By (6.3), and Corollary 6.4, the following limits are equal (if they exist),

$$\lim_{h \rightarrow \infty} -\frac{1}{h} \log \mu_{n_h}(A_h) = \lim_{h \rightarrow \infty} -\frac{1}{h} \log \mu_{m_h}(A_h),$$

and since this holds for every sequence  $m_h \gg h$ , both are equal to the limit in (1.2). Finally, since the upper and lower bounds of Proposition 2.29 on  $\mu_n(\text{ht}(\mathcal{P}_x) \geq h)$  hold for all sufficiently large  $h$  and are both independent of  $n$ , it is clear that for every  $\beta > \beta_0$ , we have  $\alpha_\beta \in [4\beta - C, 4\beta + e^{-4\beta}]$ . ■

Both (1.2) and Proposition 6.1 would follow if we show for a fixed  $x$ , say  $(\frac{1}{2}, \frac{1}{2}, 0)$ , and some sequence  $n_h \gg h$ , that the limit (6.2) exists, and call it  $\alpha_\beta$ . To see the existence of (6.2), we rely on the fact that  $A_h$  is an increasing event; we would like to leverage the monotonicity and FKG property of the Ising model to show sub/super multiplicativity of  $\mu_{n_h}(A_h)$ . The problem with this is that on the one hand, the event of reaching a height  $h_1$  gives positive information towards the event of going from height  $h_1$  to  $h_2$ , while on the other hand, the Ising measure  $\mu_{n_h}^\mp$  near height  $h_1$  is much more negative than it is near height 0. We overcome this by a careful revealing procedure, that exposes the plus connected component of  $x + (0, 0, \frac{1}{2})$  and controls the amount of positive information obtained by this revealing via the estimates of Sections 4–5.

**Proposition 6.7.** *For every  $\beta > \beta_0$ , for every  $h_1$  large, and  $h_2 \in [\frac{1}{2}h_1, 2h_1]$  if  $x$  is such that  $\frac{d(x, \partial\Lambda_{n_h})}{h} \rightarrow \infty$ , then*

$$\log \mu_{n_{h_1+h_2}}(A_{h_1+h_2}) \leq \log \mu_{n_{h_1}}(A_{h_1}) + \log \mu_{n_{h_2}}(A_{h_2}) + O(\log^2 [h_1 + h_2]).$$

Let us first conclude the proof of (1.2) and in turn, Proposition 6.1, by applying an approximate version of Fekete's sub-additivity lemma.

**Proof of (1.2) in Theorem 1.** By an approximate version of Fekete's Lemma ([22, Theorem 23], also, see [46, Theorem 1.9.2]), since  $\int t^{-2}(\log t)^2 dt < \infty$ , Proposition 6.7 implies that

$$\lim_{h \rightarrow \infty} \frac{1}{h} \log \mu_{n_h}(A_h) = \alpha_\beta$$

for some  $\alpha_\beta \in [-\infty, \infty]$ , and by the above proof of Proposition 6.1, this  $\alpha_\beta$  is also the same limit as in (6.1). As argued above, this implies that  $\alpha_\beta \in [4\beta - C, 4\beta + e^{-4\beta}]$  for some universal constant  $C$  given in Proposition 2.29, so  $\alpha_\beta/\beta \rightarrow 4$  as  $\beta \rightarrow \infty$ . ■

We now turn to proving the approximate sub-additivity of the sequence  $(\log \mu_{n_h}(A_h))_h$ .

**Proof of Proposition 6.7.** Recall that we may fix  $x = (\frac{1}{2}, \frac{1}{2}, 0)$  and set  $A_h = A_h^x$ . We will also be interested in the vertical shift of  $A_h$ , defined by

$$\theta_{h_1} A_{h_2} = \theta_{h_1} A_{h_2}^x = \left\{ \sigma : x + (0, 0, h_1 + \frac{1}{2}) \xrightarrow[\mathbb{Z}^2 \times \llbracket h_1, h_1+h_2 \rrbracket]{+} \mathcal{L}_{h_1+h_2-\frac{1}{2}} \right\}.$$

By translation, it is evident that  $\mu_n^\mp(\theta_{h_1} A_{h_2}) = \mu_n^{\mp, -h_1}(A_{h_2})$ , where  $(\mp, -m)$  boundary conditions are those that are plus on  $\partial\Lambda \cap \mathcal{L}_{<-m}$  and minus on  $\partial\Lambda \cap \mathcal{L}_{>m}$ . For every  $n$ , by monotonicity in boundary conditions,

$$\mu_n^{\mp, -h_1}(A_{h_2}) \leq \mu_n(A_{h_2}).$$

Finally, denote by  $\mathcal{P}$  the  $*$ -connected plus component of  $x + (0, 0, \frac{1}{2})$  in  $\llbracket -n_{h_1+h_2}, n_{h_1+h_2} \rrbracket^2 \times \llbracket 0, h_1 \rrbracket$ , and notice, crucially, that on the event that  $\text{ht}(\mathcal{P}_x) \geq 0$ , this plus component satisfies  $\mathcal{P} \subset \sigma(\mathcal{P}_x)$ .

Our goal is to say that conditionally on a connected plus component  $\mathcal{P}$  reaching height  $h_1$ , the probability of reaching a further height  $h_1 + h_2$  is at most  $\mu_n^{\mp, -h_1}(A_{h_2}) \leq \mu_n(A_{h_2})$ . This does not hold true, as the fact that  $\mathcal{P}$  reached height  $h_1$  contains positive information. We define a set  $\Gamma_{x, h_1}$  of plus components  $\mathcal{P}$  which, due to our structural results on tall pillars, has positive probability on the event  $A_{h_1+h_2}$ , such that for every  $\mathcal{P} \in \Gamma_{x, h_1}$ , the positive information obtained from revealing  $\mathcal{P}$  is not too large.

More precisely, let  $\Gamma_{x, h_1}$  be the event that  $\mathcal{P}$  satisfies (for  $C, K$  to be chosen sufficiently large later)

- (1) its intersection with  $\mathcal{L}_{h_1-\frac{1}{2}}$  is at most a single cell,
- (2) its bounding face-set has size at most  $Ch_1$ ,
- (3) its intersection with  $\mathcal{L}_{\frac{1}{2}}$  has diameter at most  $K \log h_1$ .

(Notice that  $\Gamma_{x, h_1}$  is a decreasing event.) The proposition will follow from the following two claims.

**Claim 6.8.** *For every  $\beta > \beta_0$ , there exist choices of  $C, K$  above, such that for every  $h_1$  and  $h_2 \in [\frac{1}{2}h_1, 2h_1]$  sufficiently large, as long as  $\lim_{h \rightarrow \infty} \frac{nh}{h} = \infty$ , we have*

$$\mu_{n_{h_1+h_2}}(A_{h_1+h_2}, \Gamma_{x, h_1}) \leq e^{C\beta(K \log h_1)^2} \mu_{n_{h_1}}(A_{h_1}) \mu_{n_{h_2}}(A_{h_2}).$$

**Claim 6.9.** *For every  $\beta > \beta_0$ , there exists a constant  $c > 0$  and choices of  $C, K$  above, such that for every  $h_1$  and  $h_2 \in [\frac{1}{2}h_1, 2h_1]$  sufficiently large, as long as  $\lim_{h \rightarrow \infty} \frac{nh}{h} = \infty$ , we have*

$$\mu_{n_{h_1+h_2}}(A_{h_1+h_2}, \Gamma_{x, h_1}) \geq c \mu_{n_{h_1+h_2}}(A_{h_1+h_2}).$$

Clearly, combining Claims 6.8–6.9 and taking logarithms on both sides concludes the proof. For ease of notation, set  $h = h_1 + h_2$ .

**Proof of Claim 6.8:** Since  $h_1$  and  $h_2$  are comparable, and  $n_h$  is such that it diverges faster than  $h$ , by Corollary 6.4 and the equivalence (6.3), as argued before, incurring errors that are decaying faster than any exponential in  $h$ , we can switch from  $\mu_{n_h}$  to  $\mu_{n_{h_1}}$  and  $\mu_{n_{h_2}}$ ; thus it will suffice for us to show the inequality

$$\mu_{n_h}(A_h, \Gamma_{x, h_1}) \leq e^{C\beta(K \log h_1)^2} \mu_{n_h}(A_{h_1}) \mu_{n_h}(A_{h_2}).$$

We begin by using the domain Markov property, the containment  $A_{h_1} \supset A_h$ , and the measurability of  $A_{h_1} \cap \Gamma_{x, h_1}$ , and in particular  $\mathcal{P}$ , with respect to  $\sigma_{[-n_h, n_h]^2 \times [0, h_1]}$  to express

$$\begin{aligned} \mu_{n_h}(A_h, \Gamma_{x, h_1}) &= \mu_{n_h}(A_{h_1}, \Gamma_{x, h_1}) \mathbb{E}_{\mathcal{P}} [\mu_{n_h}(A_h \mid \mathcal{P}) \mid A_{h_1} \cap \Gamma_{x, h_1}] \\ &\leq \mu_{n_h}(A_{h_1}) \mathbb{E}_{\mathcal{P}} [\mu_{n_h}(\theta_{h_1} A_{h_2}^Y \mid \mathcal{P}) \mid A_{h_1} \cap \Gamma_{x, h_1}], \end{aligned} \quad (6.4)$$

where  $Y \in \mathcal{L}_0$  is the projection of the singleton dictated by item (1) of  $\Gamma_{x, h_1}$  when  $A_{h_1} \cap \Gamma_{x, h_1}$  occurs. The expectations are with respect to the law of  $\mathcal{P}$  under  $\mu_{n_h}$ .

We need to bound the latter term on the right-hand side of (6.4) by the quantity  $e^{C\beta(K \log h_1)^2} \mu_{n_h}(\theta_{h_1} A_{h_2})$  to obtain the claim. We investigate this latter term as follows: notice that since  $A_{h_1} \cap \Gamma_{x, h_1}$  are measurable with respect to the plus  $*$ -connected component  $\mathcal{P}$ , we can condition on  $\mathcal{P} \in A_{h_1} \cap \Gamma_{x, h_1}$  by starting from the site at  $x + (0, 0, \frac{1}{2})$  and only revealing its  $*$ -connected plus-component in  $\Lambda_{n_h, n_h, h_1} \cap \mathcal{L}_{>0}$ .

This revealing process exposes  $\mathcal{P}$ , along with minus vertices along its entire boundary (sites in  $\mathcal{C}(\mathbb{Z}^3 \setminus \mathcal{P})$  that are  $*$ -adjacent to  $\mathcal{P}$ ) inside  $\Lambda_{n_h, n_h, h_1} \cap \mathcal{L}_{>0}$ . Let  $\sigma(\mathcal{P}) \subset \Lambda_{n_h, n_h, h_1} \cap \mathcal{L}_{>0}$  be the set of sites “interior to”  $\mathcal{P}$ , so that if the revealing procedure revealed a finite (nearest-neighbor) connected component of minus spins, corresponding to a minus bubble in  $\Lambda_{n_h, n_h, h_1} \cap \mathcal{L}_{>0}$ , set them to plus and continue revealing their interior; in this manner,  $\sigma(\mathcal{P})$  are the sites which we know to be in the plus phase given  $\mathcal{P}$ .

Let  $\mathcal{P}^\mp$  boundary conditions on  $\Lambda_{n_h, n_h, \infty} \setminus \sigma(\mathcal{P})$  be the  $\mp$  boundary conditions that additionally have plus spins in all of  $\sigma(\mathcal{P})$ , and minus spins along the boundary of  $\sigma(\mathcal{P})$  in  $\Lambda_{n_h, n_h, h_1} \cap \mathcal{L}_{>0}$ . By domain Markov, these boundary conditions are equivalent to those that have the same minus spins, but only set  $\sigma(\mathcal{P}) \cap \mathcal{L}_{\frac{1}{2}}$  and  $Y + (0, 0, h_1 - \frac{1}{2})$  to plus. Then by monotonicity and the FKG inequality, we have that

$$\mathbb{E}_{\mathcal{P}} [\mu_{n_h}(\theta_{h_1} A_{h_2}^Y \mid \mathcal{P}) \mid A_{h_1} \cap \Gamma_{x, h_1}] \leq \mathbb{E}_{\mathcal{P}} [\mu_{n_h}^{\mathcal{P}^\mp}(\theta_{h_1} A_{h_2}^Y) \mid A_{h_1} \cap \Gamma_{x, h_1}]. \quad (6.5)$$

But then, we are able to express for any such  $\mathcal{P}$  in  $A_{h_1} \cap \Gamma_{x, h_1}$ ,

$$\mu_{n_h}^{\mathcal{P}^\mp}(\theta_{h_1} A_{h_2}^Y) \leq \mu_{n_h}^{\mathcal{P}^+}(\theta_{h_1} A_{h_2}^Y) \leq e^{C\beta(|\sigma(\mathcal{P}) \cap \mathcal{L}_{\frac{1}{2}}| + 1)} \mu_{n_h}(\theta_{h_1} A_{h_2}^Y)$$

where  $\mathcal{P}^+$  boundary conditions are  $\mp$  boundary conditions that additionally have plus spins in  $\sigma(\mathcal{P}) \cap \mathcal{L}_{\frac{1}{2}}$  and  $Y + (0, 0, h_1 - \frac{1}{2})$ ; here, the first inequality was by monotonicity, and the second inequality holds for some universal constant  $C$ , by application of the *finite energy property* of the Ising model to set all spins at height  $\frac{1}{2}$  in  $\sigma(\mathcal{P})$  and the spin at  $Y + (0, 0, h_1 - \frac{1}{2})$  to be plus. As noted earlier, for every  $Y \in \mathcal{L}_0$ ,

$$\mu_{n_h}(\theta_{h_1} A_{h_2}^Y) \leq \mu_{n_h}(A_{h_2}^Y).$$

Since  $\mathcal{P} \in \Gamma_{x, h_1}$ , the distance  $|Y - x| \leq Ch_1$ ; then deterministically  $d(Y, \partial\Lambda_{n_h})$  is proportional to  $d(x, \partial\Lambda_{n_h})$  so that by the coupling of Corollary 6.4 and the comparison (6.3), up to an additive error of  $\exp[-cd(Y, \partial\Lambda_{n_h})]$ , which goes to zero faster than any exponential decay in  $h$ , we can replace  $\mu_{n_h}(A_{h_2}^Y)$  by

$\mu_{n_h}(A_{h_2}^x) = \mu_{n_h}(A_{h_2})$ . Plugging this into (6.4), and using the inequality (6.5), we see that

$$\begin{aligned} \mu_{n_h}(A_h, \Gamma_{x, h_1}) &\leq \mu_{n_h}(A_{h_1}) \cdot \mu_{n_h}(A_{h_2}) \cdot \mathbb{E}_{\mathcal{P}} \left[ e^{C\beta(|\sigma(\mathcal{P}) \cap \mathcal{L}_{\frac{1}{2}}| + 1)} \mid A_{h_1} \cap \Gamma_{x, h_1} \right] \\ &\leq \mu_{n_h}(A_{h_1}) \mu_{n_h}(A_{h_2}) \sup_{\mathcal{P} \in A_{h_1} \cap \Gamma_{x, h_1}} \exp \left[ C\beta(|\sigma(\mathcal{P}) \cap \mathcal{L}_{\frac{1}{2}}| + 1) \right]. \end{aligned}$$

By definition of  $\Gamma_{x, h_1}$ , any plus component  $\mathcal{P}$  in  $\Gamma_{x, h_1}$  has  $\text{diam}(\mathcal{P} \cap \mathcal{L}_{\frac{1}{2}}) \leq K \log h_1$  for some sufficiently large (but independent of other parameters)  $K$ , so that  $|\sigma(\mathcal{P}) \cap \mathcal{L}_{\frac{1}{2}}| \leq (K \log h_1)^2$ , concluding the proof.

**Proof of Claim 6.9:** We wish to lower bound the probability  $\mu_{n_h}(\Gamma_{x, h_1} \mid A_h)$ . We will use the equivalence (6.3) to translate the conditioning on  $A_h$  to conditioning on  $\{\text{ht}(\mathcal{P}_x) \geq h\}$ . Let us express,

$$\begin{aligned} \mu_{n_h}(\Gamma_{x, h_1}, \text{ht}(\mathcal{P}_x) \geq h) &= \mu_{n_h}(\Gamma_{x, h_1}, A_h, \text{ht}(\mathcal{P}_x) \geq h) + \mu_{n_h}(\Gamma_{x, h_1}, A_h^c, \text{ht}(\mathcal{P}_x) \geq h) \\ &\leq \mu_{n_h}(\Gamma_{x, h_1}, A_h) + \mu_{n_h}(\Gamma_{x, h_1}, \text{ht}(\mathcal{P}_x) \geq h) \mu_{n_h}(A_h^c \mid \Gamma_{x, h_1}, \text{ht}(\mathcal{P}_x) \geq h) \\ &\leq \mu_{n_h}(\Gamma_{x, h_1}, A_h) + \mu_{n_h}(\Gamma_{x, h_1}, \text{ht}(\mathcal{P}_x) \geq h) \frac{\mu_{n_h}(A_h^c, \text{ht}(\mathcal{P}_x) \geq h)}{\mu_{n_h}(\text{ht}(\mathcal{P}_x) \geq h, \Gamma_{x, h_1})}. \end{aligned}$$

Assume for the moment that we also have that for every  $\delta > 0$  there is  $h$  large enough such that for the appropriate choice of sufficiently large  $C$  and  $K$ ,

$$\mu_{n_h}(\Gamma_{x, h_1} \mid \text{ht}(\mathcal{P}_x) \geq h) \geq 1 - \delta. \quad (6.6)$$

Then, using also that  $\mu_{n_h}(A_h^c \mid \text{ht}(\mathcal{P}_x) \geq h) \leq 1/2$  for  $\beta$  large enough (as mentioned above, this is at most  $\varepsilon_\beta$  by the classical Peierls argument), we would obtain

$$\begin{aligned} \mu_{n_h}(\Gamma_{x, h_1}, A_h) &\geq \left[1 - \frac{1}{2(1-\delta)}\right] \mu_{n_h}(\Gamma_{x, h_1}, \text{ht}(\mathcal{P}_x) \geq h) \geq (1 - \delta) \left[1 - \frac{1}{2(1-\delta)}\right] \mu_{n_h}(\text{ht}(\mathcal{P}_x) \geq h) \\ &\geq \frac{1}{2} (1 - \delta - \frac{1}{2}) \mu_{n_h}(A_h) \end{aligned}$$

by (6.3). Therefore, it suffices for us to show (6.6). For the choice of  $T = \frac{1}{2}h$ , we can express,

$$\begin{aligned} \mu_{n_h}(\Gamma_{x, h_1}^c \mid \text{ht}(\mathcal{P}_x) \geq h) &\leq \mu_{n_h}(\mathbf{I}_{x, T}^c \mid \text{ht}(\mathcal{P}_x) \geq h) + \mu_{n_h}(\bar{\mathbf{I}}_{x, T}^c \mid \mathbf{I}_{x, T}, \text{ht}(\mathcal{P}_x) \geq h) \\ &\quad + \mu_{n_h}(\Gamma_{x, h_1}^c \mid \bar{\mathbf{I}}_{x, T}, \text{ht}(\mathcal{P}_x) \geq h). \end{aligned}$$

The first term on the right-hand side is  $o(1)$  as  $h \rightarrow \infty$  as long as  $\beta$  is sufficiently large by Lemma 3.3. The second term is bounded from above by  $O(e^{-c\beta h})$  for some universal  $c > 0$  by Remark 3.16. For the third term, we can union bound by the conditional probabilities of violating each of the three events constituting  $\Gamma_{x, h_1}$ ; moreover, it suffices to bound the corresponding probabilities for  $\mathcal{P}_x$  since  $\sigma(\mathcal{P}_x) \supset \mathcal{P}$ .

The conditional probability of violating item (2) of the definition of  $\Gamma_{x, h_1}$  is simply the probability of the base having surface area at least  $(C - 2R_0)h_1$ , since we are conditioning on the spine being tame; by Proposition 5.9 and the fact that  $h_1$  and  $h$  are comparable, this is  $O(\exp(-c\beta h_1^{1/3}))$  as long as  $C$  sufficiently large. The conditional probability of item (3) of  $\Gamma_{x, h_1}$  is bounded by the conditional probability of the base having diameter at least  $K \log h$  which is also  $o(1)$  in  $h$  as long as  $K$  is large enough, by Proposition 5.9. Finally, the conditional probability of item (1) is bounded by the probability of the base having height at least  $h_1$ , which is again  $o(1)$  since  $h$  is comparable to  $h_1$ , or the increment intersecting height  $h_1$  in the spine being non-trivial, which is at most  $\delta/2$  for  $\beta > \beta_0$  by Proposition 4.1.

Combining these estimates, one obtains the desired for  $\beta > \beta_0$  once  $h$  is large enough.  $\blacksquare$

**6.3. Law of large numbers for the maximum.** In this section we use Proposition 6.1 to obtain a law of large numbers for the maximum of the 3D interface on  $\Lambda_{n, n, \infty}$ . The proof follows from a simple second moment method; the fact that the correlations between large deviations of the pillar above  $x$  and  $y$  decays exponentially in  $|x - y|$ , follows from the equivalence between groups of walls and pillars, and the decay of correlations between groups of walls shown in Section 6.1.

**Proof of (1.1) in Theorem 1.** Fix  $\alpha_\beta$  to be that given by (6.1) of Theorem 6.1, equal to (1.2). We need to show that for every  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \mu_n \left( \frac{1}{\log n} \max_{x \in \mathcal{L}_0} \text{ht}(\mathcal{P}_x) \leq \frac{2}{\alpha_\beta} + \varepsilon \right) = 1, \quad \text{and} \quad \lim_{n \rightarrow \infty} \mu_n \left( \frac{1}{\log n} \max_{x \in \mathcal{L}_0} \text{ht}(\mathcal{P}_x) \geq \frac{2}{\alpha_\beta} - \varepsilon \right) = 1.$$

**Upper bound:** To see an upper bound on the maximum of the interface, we use a union bound as follows: for any two sequences  $a_n$  and  $K_n$  going to  $\infty$  as  $n \rightarrow \infty$ , we can write

$$\mu_n\left(\max_{x \in \mathcal{L}_0} \text{ht}(\mathcal{P}_x) \geq K_n\right) \leq \sum_{x: d(x, \partial\Lambda_n) \leq a_n K_n} \mu_n(\text{ht}(\mathcal{P}_x) \geq K_n) + \sum_{x: d(x, \partial\Lambda_n) \geq a_n K_n} \mu_n(\text{ht}(\mathcal{P}_x) \geq K_n).$$

For all of the first summands, we use the estimate of Theorem 2.26, that for every  $x \in \mathcal{L}_0 \cap \Lambda_n$  (including those close to the boundary  $\partial\Lambda_n$ ), the first sum is bounded above by

$$\sum_{d(x, \partial\Lambda_n) \leq a_n K_n} \mu_n(\text{ht}(\mathcal{P}_x) \geq K_n) \leq 4a_n n K_n \exp(-(4\beta - C)K_n),$$

for some universal constant  $C$ . For the second summands, since  $x$  is such that  $\frac{d(x, \Lambda_n)}{K_n} \geq a_n$  which goes to infinity as  $n \rightarrow \infty$ , the conditions of (6.1) of Proposition 6.1 are met, so as long as  $a_n K_n \ll n$ ,

$$\sum_{x: d(x, \partial\Lambda_n) \geq a_n K_n} \mu_n(\text{ht}(\mathcal{P}_x) \geq K_n) \leq (n - a_n K_n)^2 \exp(-\alpha_\beta K_n + T_n),$$

for some sequence  $T_n = o(K_n)$ . Taking

$$K_n = \frac{2}{\alpha_\beta} \log n + \kappa_n,$$

for some sequence  $\kappa_n = o(K_n)$  to be chosen subsequently in terms of  $T_n$ , we see that

$$\mu_n\left(\max_{x \in \mathcal{L}_0} \text{ht}(\mathcal{P}_x) \geq K_n\right) \leq C a_n n \log n e^{-\frac{2(4\beta - C)}{\alpha_\beta} \log n} + n^2 e^{-2 \log n - \alpha_\beta \kappa_n + T_n}.$$

Since  $\alpha_\beta - 4\beta \in [-C, e^{-4\beta}]$ , as long as  $\beta$  is sufficiently large, we have that  $\frac{2(4\beta - C)}{\alpha_\beta} > 1$ , in which case the first term is  $o(1)$  so long as, say,  $a_n < n^{\frac{4\beta - C}{\alpha_\beta} - \frac{1}{2}}$ . At the same time, if we take  $\kappa_n$  proportional to  $T_n$  such that  $\frac{\kappa_n}{T_n} > \frac{1}{\alpha_\beta}$  uniformly in  $n$ , the latter term is also  $o(1)$ . Since  $T_n = o(K_n)$ , also  $\kappa_n = o(K_n)$ , so that  $K_n \leq (\frac{2}{\alpha_\beta} + \varepsilon) \log n$  for every  $\varepsilon > 0$  for large enough  $n$ .

**Lower bound:** In order to obtain the matching lower bound, we use an easy second moment argument. Fix any small  $\varepsilon > 0$  and take

$$K_n = \left(\frac{2}{\alpha_\beta} - \varepsilon\right) \log n.$$

Now, begin by defining the subset of faces in  $\mathcal{L}_0 \cap \Lambda_n$ ,

$$\bar{\mathcal{L}}_0 = \{(x_1, x_2, 0) \in \mathcal{L}_0 : x_1 = \frac{1}{2} + i \lfloor K_n^3 \rfloor, x_2 = \frac{1}{2} + j \lfloor K_n^3 \rfloor \text{ where } (i, j) \in \llbracket 1 - \frac{n}{2K_n^3}, \frac{n}{2K_n^3} - 1 \rrbracket^2\}.$$

Then, we can define the random variable,

$$Z = Z_{K_n} = \sum_{x \in \bar{\mathcal{L}}_0} \mathbf{1}_{\{E_x\}}, \quad \text{where} \quad E_x = \{\text{ht}(\mathcal{P}_x) \geq K_n\}.$$

First of all, notice that for the above choice of  $K_n$ , we have that for  $n$  sufficiently large,

$$\mathbb{E}[Z] \geq \left(\frac{n}{K_n^3}\right)^2 e^{-\alpha_\beta K_n - T_n} \geq \left(\frac{n}{K_n^3}\right)^2 e^{-2 \log n + \varepsilon \alpha_\beta \log n - T_n} \geq n^\delta,$$

for some  $\delta > 0$  (small depending on  $\varepsilon$ ), since  $T_n = o(K_n) = o(\log n)$ .

We now wish to do a second moment estimate for  $Z$  and use the fact that the events therein are weakly correlated (exponentially decaying in their distance), to show that for  $K_n$  as above with any  $\varepsilon > 0$ ,

$$\mathbb{P}(Z > 0) \geq \frac{(\mathbb{E}[Z])^2}{\mathbb{E}[(Z)^2]} \geq 1 - o(1). \quad (6.7)$$

Expanding out  $\mathbb{E}[Z^2] = \mathbb{E}[Z] + \sum_{x \neq y \in \bar{\mathcal{L}}_0} \mu_n(E_x, E_y)$ , by Corollary 6.6, we have

$$|\mu_n(E_x) \mu_n(E_y) - \mu_n(E_x, E_y)| \leq \|\mu_n(\mathcal{P}_x \in \cdot) \mu_n(\mathcal{P}_y \in \cdot) - \mu_n(\mathcal{P}_x \in \cdot, \mathcal{P}_y \in \cdot)\|_{\text{TV}} \leq e^{-cK_n^2},$$

which is smaller than any polynomially decaying function of  $n$  by the choice of  $K_n$  and the fact that  $|x - y| \geq \frac{1}{4}K_n^3$ . Therefore, we can bound

$$\sum_{x,y \in \bar{\mathcal{L}}_0: x \neq y} \mu_n(E_x, E_y) \leq \sum_{x,y \in \bar{\mathcal{L}}_0, x \neq y} [\mu_n(E_x)\mu_n(E_y) + e^{-cK_n^2}] \leq (\mathbb{E}[Z])^2 + O(n^2 e^{-cK_n^2}).$$

Plugging this bound in, we see that as  $n \rightarrow \infty$ ,

$$\frac{(\mathbb{E}[Z])^2}{\mathbb{E}[Z^2]} \geq \frac{(\mathbb{E}[Z])^2}{\mathbb{E}[Z] + (\mathbb{E}[Z])^2 + o(1)} \rightarrow 1,$$

as long as  $\mathbb{E}[Z]$  is diverging, which as noted earlier, is indeed the case for our choice of  $K_n$ .  $\blacksquare$

## 7. FINER PROPERTIES OF THE INCREMENT SEQUENCE OF THE SPINE

In this section, we begin to analyze the shape of the pillars of the interface that attain the maximum of Section 6.3. We show that for tall pillars consisting of  $T$  increments, their spine can be decomposed into an asymptotically (as you get further from the base or tip) stationary sequence of weakly mixing increments. In particular, the increment sequence, viewed from the  $(T/2)$ -th increment converges weakly to a stationary bi-infinite sequence of increments, with polynomially decaying bounds on its mixing rate.

Since this section (and most of the remainder of the paper) is concerned with the properties of pillars under the event  $\bar{\mathbf{I}}_{x,T}$ , let us henceforth take any sequence  $n = n_T$  and  $x = x_T$  satisfying  $n_T \gg T$  and  $d(x_T, \partial\Lambda_{n_T}) \gg T$  and denote the Ising measure on  $\Lambda_{n_T}$  conditional on  $\bar{\mathbf{I}}_{x_T,T}$  by

$$\pi_T(\cdot) := \mu_n(\cdot | \bar{\mathbf{I}}_{x,T}).$$

In §7.1, we prove a spatial mixing estimate for the increment sequence  $(\mathcal{X}_1, \dots, \mathcal{X}_T)$ :

**Proposition 7.1.** *For every  $\gamma$ , there exist  $\beta_0, K$ , and  $C$  such that for every  $\beta > \beta_0$  and  $K \log T \leq j < k \leq T$ ,*

$$\sup_{E_j \subset \bar{\mathfrak{X}}^{j-K \log T}, E_k \subset \bar{\mathfrak{X}}^{T-k}} \left| \pi_T((\mathcal{X}_{K \log T}, \dots, \mathcal{X}_j) \in E_j, (\mathcal{X}_k, \dots, \mathcal{X}_T) \in E_k) - \pi_T((\mathcal{X}_{K \log T}, \dots, \mathcal{X}_j) \in E_j) \pi_T((\mathcal{X}_k, \dots, \mathcal{X}_T) \in E_k) \right| \leq C|k - j|^{-\gamma}.$$

In §7.2, we prove that the increment sequence is asymptotically stationary away from the base and the tip.

**Proposition 7.2.** *For every  $\gamma$ , there exist  $\beta_0, K$ , and  $C$  such that for every  $\beta > \beta_0$ , every  $K \log T \leq j \leq T$  and  $K \log T' \leq j' \leq T'$ , and every  $s \leq (T - j) \wedge (T' - j')$ ,*

$$\sup_{E \in \bar{\mathfrak{X}}^s} |\pi_T((\mathcal{X}_j, \dots, \mathcal{X}_{j+s}) \in E) - \pi_{T'}((\mathcal{X}_{j'}, \dots, \mathcal{X}_{j'+s}) \in E)| \leq C[(j - K \log T) \wedge (j' - K \log T')]^{-\gamma} \vee [(T - j - s) \wedge (T' - j' - s)]^{-\gamma}.$$

In §7.3 we combine Propositions 7.1–7.2, to define a limiting distribution on increment sequences.

**Corollary 7.3.** *For every  $\gamma$  large, let  $\beta > \beta_0$  where  $\beta_0$  is the one given by Propositions 7.1–7.2 for that  $\gamma$ . There exists a stationary distribution  $\nu = \nu_\beta$  on  $\bar{\mathfrak{X}}^{\mathbb{Z}}$  so that, if  $a_T$  has  $(a_T \vee (T - a_T))/\log T \rightarrow \infty$  as  $T \rightarrow \infty$ , then the law of  $(\dots, \mathcal{X}_{a_T-1}, \mathcal{X}_{a_T}, \mathcal{X}_{a_T+1}, \dots)$  under  $\pi_T$  converges weakly to  $\nu((\dots, \mathcal{X}_{-1}, \mathcal{X}_0, \mathcal{X}_1, \dots) \in \cdot)$ . In particular, the distribution  $\nu$  satisfies*

- (1) *There exists  $c > 0$  (independent of  $\beta$ ) such that  $\nu(\mathbf{m}(\mathcal{X}_0) \geq r) \leq \exp[-c\beta r]$  for every  $r$ .*
- (2) *There exists  $C = C_\beta$  such that  $\|\nu(\mathcal{X}_0 \in \cdot, \mathcal{X}_k \in \cdot) - \nu(\mathcal{X}_0 \in \cdot)\nu(\mathcal{X}_k \in \cdot)\|_{TV} \leq Ck^{-\gamma}$  for every  $k$ .*

The key step in the proofs of Propositions 7.1–7.2 is the use of what we call “two-to-two” maps, which are bijections on the set of pairs of interfaces  $\bar{\mathbf{I}}_{x,T} \times \bar{\mathbf{I}}_{x,T}$ , in contrast to all the maps we have applied up to this point. The reason for this is that any “one-to-one” map  $\Phi$  that changes an increment  $\mathcal{X}_i$  sustains a multiplicative cost of  $e^{\pm C\mathbf{m}(\mathcal{X}_i)}$  in the ratio  $\mu_n(\mathcal{I})/\mu_n(\Phi(\mathcal{I}))$ , which would overwhelm the upper bounds we wish to attain. “Two-to-two” maps give us a mechanism of avoiding any such costs, and ensuring all faces in the pair  $(\mathcal{I}, \mathcal{I}')$  are identified with faces in  $\Phi(\mathcal{I}, \mathcal{I}')$  with which they have congruent local neighborhoods. We explain this in more detail in Sections 7.1.1–7.2.1.

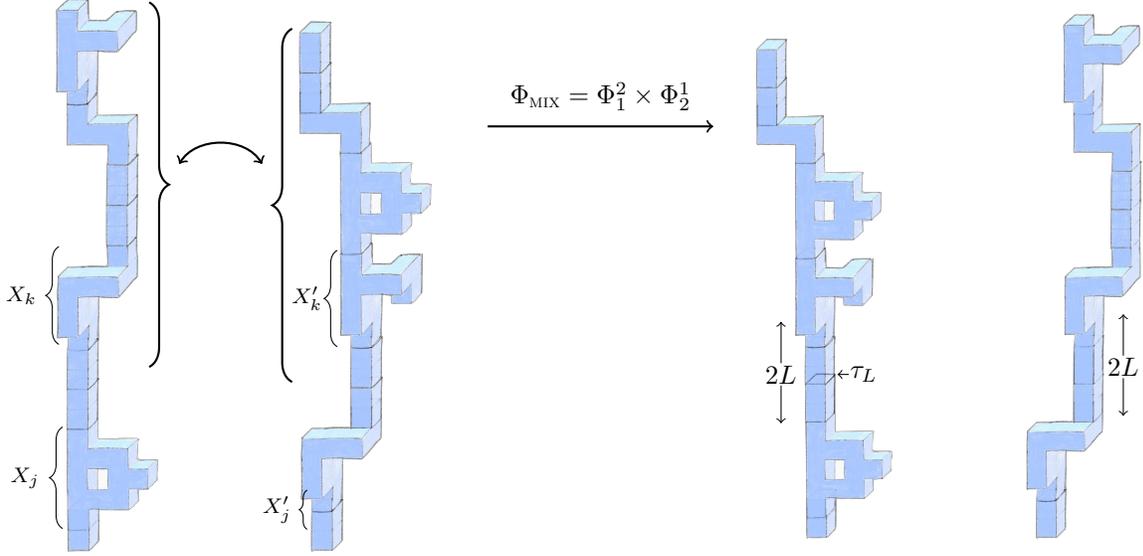


FIGURE 11. The map  $\Phi_{\text{MIX}} = \Phi_1^2 \times \Phi_2^1$  acting on a pair of increment sequences in  $\Gamma_{\emptyset, L}$

**7.1. Proof of Proposition 7.1: mixing properties of the increment sequence.** We wish to show that the correlations between the  $j$ -th and  $k$ -th increments decay polynomially fast in their distance, with the exponent of the polynomial increasing with  $\beta$ .

Fix any  $\gamma$  and let  $K$  be such that if  $c_{\emptyset}$  is the constant from Proposition 5.1,  $c_{\emptyset}\beta K > \gamma$ . Next, fix  $K \log T < j < k < T$ , and let  $L = \lceil \frac{4\gamma}{\epsilon} \log |k - j| \rceil$ ; due to our freedom to take  $C$  as desired, we may assume without loss that  $|k - j|$  is sufficiently large. Fix any  $E_j \in \mathfrak{X}^{j - K \log T}$ ,  $E_k \in \mathfrak{X}^{T - k}$ , and, in order to simplify notation, let us denote the tuples  $\mathcal{Z}_j = (\mathcal{X}_{K \log T}, \dots, \mathcal{X}_j)$  and  $\mathcal{Z}_k = (\mathcal{X}_k, \dots, \mathcal{X}_T)$ , with fixed instantiations  $Z_j = (X_{K \log T}, \dots, X_j)$  and  $Z_k = (X_k, \dots, X_T)$ .

Let  $\mathcal{A}_{jk}$  denote the set of all  $T$ -admissible truncated interfaces, increment sequences  $(X_i)_{i \in [j+1, k-1]}$ , and remainder increment  $X_{>T}$ . For any triplet  $(A, Z_j, Z_k)$  where  $A = A_{jk} \in \mathcal{A}_{jk}$ , we write  $\pi_T(Z_j, Z_k, A)$  to denote the probability that the random interface under  $\pi_T$  has  $\mathcal{Z}_j = Z_j$ ,  $\mathcal{Z}_k = Z_k$  and has  $\mathcal{I}_{\text{TR}}$ ,  $(\mathcal{X}_i)_{i \in [j+1, k-1]}$  and  $\mathcal{X}_{>T}$  agreeing with  $A_{jk}$ .

We begin by expressing the left-hand side in the proposition as

$$\sum_{\substack{Z_j \in E_j, Z_k \in E_k \\ Z'_j \in \mathfrak{X}^{j - K \log T}, Z'_k \in \mathfrak{X}^{T - k}}} \left[ \sum_{A, A' \in \mathcal{A}_{jk}} \pi(Z_j, Z_k, A) \pi_T(Z'_j, Z'_k, A') - \sum_{\tilde{A}, \tilde{A}' \in \mathcal{A}_{jk}} \pi_T(Z_j, Z'_k, \tilde{A}) \pi_T(Z'_j, Z_k, \tilde{A}') \right]$$

Define a set of nice interfaces  $\Gamma_L$  for which we have good decorrelation between  $\mathcal{Z}_j$  and  $\mathcal{Z}_k$ . First, let  $\Gamma_{\emptyset, L} = \Gamma_{\emptyset, L}(j, k)$  be the set of pairs of increment sequences  $\mathfrak{X}^T \times \mathfrak{X}^T$  for which there is a stretch of  $2L$  consecutive indices between  $j$  and  $k$  on which both  $A$  and  $A'$  have trivial increments. That is,

$$\Gamma_{\emptyset, L} = \{((X_i)_{i \leq T}, (X'_i)_{i \leq T}) : \exists \tau_L \in [j, k], (X_{\tau_L - L}, \dots, X_{\tau_L + L}) = (X'_{\tau_L - L}, \dots, X'_{\tau_L + L}) = (X_{\emptyset}, \dots, X_{\emptyset})\}.$$

Abusing notation, under the event  $\Gamma_{\emptyset, L}$ , let  $\tau_L$  be the smallest index greater than  $j$  such that the stretch of length  $2L$  centered at  $\tau$  satisfies the condition of  $\Gamma_{\emptyset, L}$ . We can now define a map  $\Phi_{\text{MIX}}$  on pairs of increment sequences, that swaps the increment stretches above  $\tau_L$ : see Figure 11 for a visualization.

**Definition 7.4.** For each  $j, k$ , let  $\Phi_{\text{MIX}} = \Phi_{\text{MIX}}(j, k) : (\mathfrak{X}^T \times \mathfrak{X}_{\text{REM}})^2 \rightarrow (\mathfrak{X}^T \times \mathfrak{X}_{\text{REM}})^2$  be given as follows. For any pair of increment sequences  $(X^{(1)}, X^{(2)}) = ((X_i^{(1)})_{i \leq T}, X_{>T}^{(1)}, (X_i^{(2)})_{i \leq T}, X_{>T}^{(2)})$  let  $\Phi_{\text{MIX}}(X^{(1)}, X^{(2)}) = (\Phi_1^2(X^{(1)}, X^{(2)}), \Phi_2^1(X^{(1)}, X^{(2)}))$  be the pair of increment sequences attained as follows: if  $(X^{(1)}, X^{(2)}) \notin \Gamma_{\emptyset, L}$ , let  $\Phi_{\text{MIX}}(X^{(1)}, X^{(2)}) = (X^{(1)}, X^{(2)})$ ; otherwise

- (1) Find the first run of  $2L$  consecutive indices between  $j$  and  $k$  on which both  $X^{(1)}$  and  $X^{(2)}$  are trivial increments, and call the middle index of this run  $\tau_L \in [j, k]$ .
- (2) Let  $\Phi_1^2(X^{(1)}, X^{(2)})$  have increment sequence given by

$$\Phi_1^2(X^{(1)}, X^{(2)}) = (X_1^{(1)}, \dots, X_j^{(1)}, \dots, X_{\tau_L}^{(1)}, X_{\tau_L+1}^{(2)}, \dots, X_k^{(2)}, \dots, X_{>T}^{(2)}).$$

(3) Let  $\Phi_2^1(X^{(1)}, X^{(2)})$  have increment sequence given by

$$\Phi_2^1(X^{(1)}, X^{(2)}) = (X_1^{(2)}, \dots, X_j^{(2)}, \dots, X_{\tau_L}^{(2)}, X_{\tau_L+1}^{(1)}, \dots, X_k^{(1)}, \dots, X_{>T}^{(1)}).$$

Abusing notation, we define  $\Phi_{\text{MIX}}$  on  $\bar{\mathbf{I}}_{x,T} \times \bar{\mathbf{I}}_{x,T}$  that uses the same truncations of the pair  $(\mathcal{I}, \mathcal{I}')$  and applies the map  $\Phi$  to their respective pairs of increment sequences in their pillars  $\mathcal{P}_x$  and  $\mathcal{P}'_x$ . If the two interfaces are both tame and also satisfy  $\tau_{\text{SP}} \vee \tau_{\text{SP}'} \leq K \log T$ , the resulting pair of interfaces would be in  $\mathbf{I}_{x,T} \times \mathbf{I}_{x,T}$ .

**7.1.1. Strategy of the map  $\Phi_{\text{MIX}}$ .** We briefly motivate the construction of the map  $\Phi_{\text{MIX}}$ . We first describe the complications that would arise if we used a map that sent one interface to another interface, instead of acting on pairs of interfaces. In order to prove a mixing property on the increment sequence, one would want to construct a map which maps an interface with an increment  $X_j$  and an increment  $X_k$ , to an interface with some other increment  $X'_j$  and the same  $X_k$ , say having  $\mathbf{m}(X_j) = \mathbf{m}(X'_j)$ . If the weight distortion of such a map is  $o(|k-j|^{-c})$  for some  $c > 0$ , we will have shown that that conditioning on the presence of the increment  $X_j$  vs.  $X'_j$  does not influence the conditional probability of  $X_k$ . Unlike the maps in Sections 4–5 there is no energy gain in such a map; however, the replacement of  $X_j$  by  $X'_j$  inevitably costs an  $e^{C(|X_j| \vee |X'_j|)}$  in the weight ratio, coming from the uniform bound on  $\mathbf{g}$  (2.2).

In order to obtain ratios of weights that are  $o(1)$  in  $|k-j|$ , we use a second interface, whose increment sequence has  $X'_j$  and  $X'_k$ , and we demonstrate that the sequence is mixing by showing that the probabilities of a pair of interfaces having increment pairs  $\{(X_j, X_k), (X'_j, X'_k)\}$  is close to the probability of the pair having  $\{(X_j, X'_k), (X'_j, X_k)\}$ . Then, in the control of the  $\mathbf{g}$  term, we could identify faces from  $X_j, X_k$  with one another and  $X'_j, X'_k$  with one another across the pairs of interfaces. However, a naive application of this kind of map would lead to a  $1 \pm \varepsilon_\beta$  weight distortion, rather than one that is  $1 + o(1)$  in  $|k-j|$ . More precisely, every increment would feel the change in the  $\mathbf{g}$  term in terms of its distance to the increment where we spliced the interface to perform the swap—in particular, the increments near the splicing location, if they disagree between the pair of interfaces, will contribute a constant, but not  $o(1)$  to the weight distortion.

To improve this to something decaying polynomially in  $|k-j|$ , the map  $\Phi_{\text{MIX}}$  relies on the existence of a sequence of consecutive increments of logarithmic length in  $|k-j|$ , that are trivial in both interfaces. In that case, after the splicing, for every face in either of the interfaces, the radius of congruence is bounded by half the length of the consecutive sequence of interfaces, and by (2.3), the weight distortion is at most polynomially decaying in  $|k-j|$  for large enough  $\beta$ , as desired. Refer to Figure 11 for a visualization.

**7.1.2. Analysis of the map  $\Phi_{\text{MIX}}$ .** We now use the map  $\Phi_{\text{MIX}}$  to define a good set of pairs of increment sequences, refining the set  $\Gamma_{\emptyset, L}$ , on which we will have good control on the ratio of probabilities under  $\Phi_{\text{MIX}}$ . Let  $\Gamma_L$  be the set of  $(\mathcal{I}, \mathcal{I}') \in \bar{\mathbf{I}}_{x,T} \times \bar{\mathbf{I}}_{x,T}$  such that its pair of increment sequences are in  $\Gamma_{\emptyset, L}$ , and additionally having

- (1) Their source point indices  $\tau_{\text{SP}}, \tau_{\text{SP}'}$  are both less than  $K \log T$ ; denote this event  $\Gamma_{\tau_{\text{SP}}}$ .
- (2) The pair of interfaces  $(\mathcal{I}, \mathcal{I}')$  are such that  $\Phi_{\text{MIX}}(\mathcal{I}, \mathcal{I}')$  are both tame; denote this event  $\bar{\Gamma}$ .
- (3) Their increment sequences  $(X_i)$  and  $(X'_i)$  satisfy

$$|\mathcal{F}(X_{>T})| e^{-\bar{c}(T+1-\tau_L+L)} + \sum_i |\mathcal{F}(X_{\tau_L+i})| e^{-\bar{c}(L+i)} \leq e^{-\bar{c}L/2}.$$

and analogously for  $(X'_i)$ ; denote this event  $\Gamma_{i>\tau_L}$ .

We will separately consider the cases where  $(A, Z_j, Z_k) \times (A', Z'_j, Z'_k)$  and  $(\tilde{A}, Z'_j, Z_k) \times (\tilde{A}', Z'_k, Z_j)$  are in  $\Gamma_L$  and the cases when they aren't: without loss of generality, let us consider the former pairs (the latter estimate would hold after swapping  $Z'_j$  with  $Z_j$ ). The contribution from pairs of interfaces where one is not in  $\Gamma_L$  are bounded above by the sum of

$$\sum_{\substack{Z_j \in E_j, Z_k \in E_k, Z'_j, Z'_k, A, A': \\ (A, Z_j, Z_k), (A', Z'_j, Z'_k) \in \Gamma_L^c}} \pi_T(Z_j, Z_k, A) \pi_T(Z'_j, Z'_k, A') \leq \pi_T^{\otimes 2}(\Gamma_L^c).$$

**Lemma 7.5.** *For the choices of  $\gamma, K, L$  above, for  $\beta > \beta_0$ , we have  $\pi_T^{\otimes 2}(\Gamma_L^c) \leq |k-j|^{-\gamma}$ .*

**Proof of Lemma 7.5.** By a union bound, we can express

$$\pi_T^{\otimes 2}(\Gamma_L^c) \leq 2\pi_T(\Gamma_{\tau_{\text{SP}}}^c) + \pi_T^{\otimes 2}(\bar{\Gamma}^c \mid \Gamma_{\tau_{\text{SP}}}) + \pi_T^{\otimes 2}(\Gamma_{\emptyset, L}^c \mid \Gamma_{\tau_{\text{SP}}}) + \pi_T^{\otimes 2}(\Gamma_{i>\tau_L} \mid \Gamma_{\emptyset, L}, \Gamma_{\tau_{\text{SP}}}).$$

By Proposition 5.1 and the fact that  $\tau_{\text{SP}} \leq \text{ht}(v_{\tau_{\text{SP}}}) + \frac{1}{2}$ , we have that  $\pi_T(\Gamma_{\tau_{\text{SP}}}^c) \leq e^{-c\beta K \log T} \leq T^{-\gamma}$ . In order for  $\Phi_{\text{MIX}}(\mathcal{I}, \mathcal{I}')$  to not be tame, one of  $\mathfrak{m}_T(\mathcal{S}_x)$  or  $\mathfrak{m}_T(\mathcal{S}'_x)$  must be at least  $r_0/2$ ; by Lemma 3.15 and the fact that  $r_0/2 > 8T$ , then,  $\pi_T^{\otimes 2}(\bar{\Gamma} \mid \Gamma_{\tau_{\text{SP}}})$  is at most  $2e^{-c\beta T}$  for some  $c > 0$ . We now turn to the latter two terms above. By Proposition 4.1, in particular the conditional version of it, given any  $\mathcal{I}_{\text{TR}}$  (in particular any  $\tau_{\text{SP}}$ ) the sequence  $(\mathbf{1}\{\mathcal{X}_i = X_{\emptyset}\})_{i \geq \tau_{\text{SP}}}$  (which includes the increments between indices  $j$  and  $k$  by  $\Gamma_{\tau_{\text{SP}}}$ ) stochastically dominates a sequence of independent  $\text{Ber}(1 - \varepsilon_{\beta})$  coin tosses, for some  $\varepsilon_{\beta}$  going to zero as  $\beta \rightarrow \infty$ . As a consequence,  $\pi_T^{\otimes 2}(\Gamma_{\emptyset, L}^c)$  is at most the probability that a set of  $|k - j|$  i.i.d.  $\text{Ber}((1 - \varepsilon_{\beta})^2)$  coin flips has no sequence of  $2L$  consecutive ones. Thus, for large enough  $\beta$ , (depending on  $\gamma, \bar{c}$ )

$$\pi_T^{\otimes 2}(\Gamma_{\emptyset, L}^c \mid \Gamma_{\tau_{\text{SP}}}) \leq (1 - (1 - \varepsilon_{\beta})^{4L})^{|k-j|/(2L)} \leq \exp[-(1 - \varepsilon_{\beta})^{4L}|k-j|/(2L)] \leq \exp[-|k-j|^{3/4}].$$

By Corollary 4.2, conditional on the entire increment sequence up to  $\tau_L + L$  (which contains the information of  $\tau_{\text{SP}}, \Gamma_{\tau_{\text{SP}}}, \Gamma_{\emptyset, L}$  the concentration estimate on the excess areas of subsequent increments holds (uniformly in the choice of  $\tau_L$ ). Combining these, one sees the bound (where the conditioning on  $\Gamma_{\emptyset, L}(\tau_L)$  is to say that  $\Gamma_{\emptyset, L}$  happens for that specific  $\tau_L$ ),

$$\begin{aligned} & \pi_T^{\otimes 2}(\Gamma_{i > \tau_L} \mid \Gamma_{\emptyset, L}, \Gamma_{\tau_{\text{SP}}}) \\ & \leq 2 \sup_{\mathcal{I}_{\text{TR}} \in \Gamma_{\tau_{\text{SP}}, \tau_L}} \pi_T \left( |\mathcal{F}(\mathcal{X}_{>T})| e^{-\bar{c}(T+1-\tau_L+L)} + \sum_i |\mathcal{F}(X_{\tau_L+i})| e^{-\bar{c}(L+i)} \geq e^{-\bar{c}L/2} \mid \mathcal{I}_{\text{TR}}, \Gamma_{\emptyset, L}(\tau_L) \right), \end{aligned}$$

which is at most  $2 \exp(-c\beta \bar{c}L/2)$ ; therefore,  $\pi_T^{\otimes 2}(\Gamma_L^c) \leq \exp[-|k-j|^{3/4}] + 2 \exp[-c\beta \bar{c}L/2]$ .

Our choice of  $L$  was precisely such that as long as  $\beta \bar{c} > 1$ , the latter quantity is at most  $|k-j|^{-2\gamma}$ , which dominates the first term.  $\blacksquare$

On the other hand, when both pairs of triplets  $((A, Z_j, Z_k), (A', Z'_j, Z'_k))$  and  $((\tilde{A}, Z'_j, Z_k), (\tilde{A}', Z'_k, Z_j))$  are in  $\Gamma_L$ , we are left to control

$$\begin{aligned} & \sum_{\substack{Z_j \in E_j, Z_k \in E_k \\ Z'_j, Z'_k, A, A' \\ (A, Z_j, Z_k), (A', Z'_j, Z'_k) \in \Gamma_L}} \pi_T(Z_k, Z_j, A) \pi_T(Z'_k, Z'_j, A') - \\ & \sum_{\substack{Z_j \in E_j, Z_k \in E_k \\ Z'_j, Z'_k, \tilde{A}, \tilde{A}' \\ (\tilde{A}, Z_j, Z'_k), (\tilde{A}', Z'_j, Z_k) \in \Gamma_L}} \pi(Z_k, Z'_j, \tilde{A}') \pi(Z'_k, Z_j, \tilde{A}). \end{aligned}$$

Now that we have restricted to tame interfaces, with well-behaved increment sequences, we can naturally view  $\Phi_{\text{MIX}}$  as a map on  $(\bar{\mathbf{I}}_{x, T} \times \bar{\mathbf{I}}_{x, T}) \cap \Gamma_L$ . This restriction yields the following correspondence.

**Claim 7.6.** *The restriction of  $\Phi_{\text{MIX}}$  to  $(\bar{\mathbf{I}}_{x, T} \times \bar{\mathbf{I}}_{x, T}) \cap \Gamma_L$  is a bijection from  $(\bar{\mathbf{I}}_{x, T} \times \bar{\mathbf{I}}_{x, T}) \cap \Gamma_L$  to itself.*

**Proof of Claim 7.6.** Since  $\Phi_{\text{MIX}} = \Phi_{\text{MIX}}^{-1}$ , it suffices to show that for every pair  $(\mathcal{I}, \mathcal{I}') \in (\bar{\mathbf{I}}_{x, T} \times \bar{\mathbf{I}}_{x, T}) \cap \Gamma_L$ , we have  $\Phi_{\text{MIX}}(\mathcal{I}, \mathcal{I}') \in (\bar{\mathbf{I}}_{x, T} \times \bar{\mathbf{I}}_{x, T}) \cap \Gamma_L$ . Indeed, as mentioned, the fact that  $(\mathcal{I}, \mathcal{I}') \in \Gamma_{\tau_{\text{SP}}}$  ensures that  $\Phi_{\text{MIX}}(\mathcal{I}, \mathcal{I}') \in \mathbf{I}_{x, T} \cap \mathbf{I}_{x, T}$ ; the fact that  $(\mathcal{I}, \mathcal{I}') \in \bar{\Gamma}$ , by definition, guarantees that  $\Phi_{\text{MIX}}(\mathcal{I}, \mathcal{I}')$  are both tame. Finally, the fact that  $\Phi_{\text{MIX}}(\mathcal{I}, \mathcal{I}')$  remains in  $\Gamma_L$  holds for the following reasons: (1) holds as the source points are unchanged by the map; (2) holds as  $\Phi_{\text{MIX}} = \Phi_{\text{MIX}}^{-1}$ ; (3) holds for  $\Phi_{\text{MIX}}(\mathcal{I}, \mathcal{I}')$  since the pair of increment sequences above  $\tau_L$  in  $\Phi_{\text{MIX}}(\mathcal{I}, \mathcal{I}')$  are exactly the pair of increment sequences above  $\tau_L$  of  $(\mathcal{I}, \mathcal{I}')$ .  $\blacksquare$

With the claim in hand, notice that  $\Phi_{\text{MIX}}$  preserves the  $\tau_L$  at which  $\Gamma_{\emptyset, L}$  is attained and, we have

$$\Phi_{\text{MIX}} \left( (Z_j, Z_k, A), (Z'_j, Z'_k, A') \right) = \left( (Z_j, Z'_k, \Phi_1^2(A, A')), (Z'_j, Z_k, \Phi_2^1(A, A')) \right)$$

in the sense that the  $Z_k$  and  $Z'_k$  get swapped by application of the map in the manner desired, as does everything else in the spine above index  $\tau$ . Using this bijection, we rewrite the difference above as

$$\begin{aligned} & \sum_{\substack{Z_j \in E_j, Z_k \in E_k \\ Z'_j, Z'_k, A, A' \\ (A, Z_j, Z_k), (A', Z'_j, Z'_k) \in \Gamma_L}} \pi_T(Z_j, Z_k, A) \pi_T(Z'_j, Z'_k, A') - \pi_T(Z_j, Z'_k, \Phi_1^2(A, A')) \pi_T(Z'_j, Z_k, \Phi_2^1(A, A')). \end{aligned}$$

Now fix  $Z_j \in E_j, Z_k \in E_k, Z'_j, Z'_k$  and  $(A, A')$  such that the above triplets are in  $\Gamma_L$ ; for ease of notation, let  $\tilde{A} = \Phi_1^2(A, A')$  and  $\tilde{A}' = \Phi_2^1(A, A')$ . Consider the quantity

$$\begin{aligned} & |\pi_T(Z_k, Z_j, A)\pi_T(Z'_k, Z'_j, A') - \pi_T(Z_j, Z'_k, \tilde{A})\pi_T(Z'_j, Z_k, \tilde{A}')| \\ &= \pi_T(Z_j, Z'_k, \tilde{A})\pi_T(Z'_j, Z_k, \tilde{A}') \left| \frac{\mu_n(Z_j, Z_k, A)\mu_n(Z'_j, Z'_k, A')}{\mu_n(Z_j, Z'_k, \tilde{A})\mu_n(Z'_j, Z_k, \tilde{A}')} - 1 \right|; \end{aligned} \quad (7.1)$$

since each of the triplets are in  $\bar{\mathbf{I}}_{x,T}$ , expressing e.g.,  $\pi_T(Z_j, Z_k, A) = \frac{\mu_n(Z_j, Z_k, A)}{\mu_n(\bar{\mathbf{I}}_{x,T})}$ , the contributions from  $\mu_n(\bar{\mathbf{I}}_{x,T})$  cancel out. Let us now focus on the difference in the absolute value in (7.1), and in particular the ratio of the probabilities of the two pairs of interfaces. This formulation allows us to apply the machinery of Theorem 2.21 to the pair of interfaces: for ease of notation, let us denote the interface given by  $(Z_j, Z_k, A)$  by  $\mathcal{I}^{jk}$ , and denote  $\mathcal{I}^{j'k'} = (Z_j, Z'_k, \tilde{A})$ ,  $\mathcal{I}^{j'k} = (Z'_j, Z_k, \tilde{A}')$  and  $\mathcal{I}^{jk'} = (Z'_j, Z'_k, A')$  analogously. Express

$$\begin{aligned} \frac{\mu_n(Z_j, Z_k, A)\mu_n(Z'_j, Z'_k, A')}{\mu_n(Z_j, Z'_k, \tilde{A})\mu_n(Z'_j, Z_k, \tilde{A}')} &= \frac{e^{\beta(|\mathcal{I}^{jk}|+|\mathcal{I}^{j'k'}|)}}{e^{\beta(|\mathcal{I}^{j'k'}|+|\mathcal{I}^{j'k}|)}} \cdot \frac{e^{\sum_{f \in \mathcal{I}^{j'k'}} \mathbf{g}(f, \mathcal{I}^{j'k'}) + \sum_{f \in \mathcal{I}^{jk}} \mathbf{g}(f, \mathcal{I}^{jk})}}{e^{\sum_{f \in \mathcal{I}^{j'k}} \mathbf{g}(f, \mathcal{I}^{j'k}) + \sum_{f \in \mathcal{I}^{jk'}} \mathbf{g}(f, \mathcal{I}^{jk'})}} \\ &= \exp \left[ \sum_{f \in \mathcal{I}^{jk}} \mathbf{g}(f, \mathcal{I}^{jk}) + \sum_{f \in \mathcal{I}^{j'k'}} \mathbf{g}(f, \mathcal{I}^{j'k'}) - \sum_{f \in \mathcal{I}^{j'k}} \mathbf{g}(f, \mathcal{I}^{j'k}) - \sum_{f \in \mathcal{I}^{jk'}} \mathbf{g}(f, \mathcal{I}^{jk'}) \right]. \end{aligned}$$

We now turn to bounding the absolute value of the exponent. Recall that  $\tau_L$  denotes the first index of the increment run of  $2L$  trivial increments in both  $(A, A')$  (and consequently also  $(\tilde{A}, \tilde{A}')$  for the same  $\tau_L$ ).

**Lemma 7.7.** *There is a universal  $\bar{C}$  such that for any pair  $(\mathcal{I}^{jk}, \mathcal{I}^{j'k'}) \in (\bar{\mathbf{I}}_{x,T} \times \bar{\mathbf{I}}_{x,T}) \cap \Gamma_L$ , we have*

$$\left| \sum_{f \in \mathcal{I}^{jk}} \mathbf{g}(f, \mathcal{I}^{jk}) + \sum_{f \in \mathcal{I}^{j'k'}} \mathbf{g}(f, \mathcal{I}^{j'k'}) - \sum_{f \in \mathcal{I}^{j'k}} \mathbf{g}(f, \mathcal{I}^{j'k}) - \sum_{f \in \mathcal{I}^{jk'}} \mathbf{g}(f, \mathcal{I}^{jk'}) \right| \leq \bar{C} \exp[-\bar{c}L/2].$$

Let us first conclude the proof of Proposition 7.1 given Lemma 7.7. By our choice  $L$ , the right-hand side above is at most  $\bar{C}|k-j|^{-2\gamma}$ , from which we would deduce that

$$\left| \frac{\mu_n(Z_j, Z_k, A)\mu_n(Z'_j, Z'_k, A')}{\mu_n(Z_j, Z'_k, \tilde{A})\mu_n(Z'_j, Z_k, \tilde{A}')} - 1 \right| \leq 2\bar{C}|k-j|^{-2\gamma}.$$

Since this upper bound is independent of  $Z_j, Z_k, Z'_j, Z'_k, A, A'$ , when we sum (7.1), it factors out, and the sum of the probabilities over some subset of interfaces in  $\bar{\mathbf{I}}_{x,T}$  is of course at most one. Combining this with the contribution from terms not in  $\Gamma_L$  yields an additional  $|k-j|^{-\gamma}$ , implying the desired estimate.

**Proof of Lemma 7.7.** It will be important to use the structure of the map  $\Phi_{\text{MIX}}$  to choose the right pairing of summands in the different sums on the left-hand side above. To that end, let us define the following subsets of faces of the interfaces we consider: let  $\mathcal{I}_{\text{TR}}$  and  $\mathcal{I}'_{\text{TR}}$  be the respective truncations of  $A$  and  $A'$ . Let

$$F_-^{jk} = \left\{ f \in \mathcal{I}_{\text{TR}} \cup \bigcup_{i \leq \tau_L} \mathcal{F}(X_i) \right\}, \quad \text{and} \quad F_+^{jk} = \left\{ f \in \bigcup_{i > \tau_L} \mathcal{F}(X_i) \cup \mathcal{F}(X_{>T}) \right\}.$$

be the sets of all faces “below”  $X_{\tau_L}$ , and all faces “above”  $X_{\tau_L}$  respectively. In this manner,  $\mathcal{I}^{jk} = F_-^{jk} \cup F_+^{jk}$ , and we can define  $F_{\pm}^{j'k'}, F_{\pm}^{j'k}, F_{\pm}^{jk'}$  analogously (where whether or not  $j$  is primed indicates whether  $\mathcal{I}_{\text{TR}}$  or  $\mathcal{I}'_{\text{TR}}$  is used in  $F_-$ ). By definition, we have the equalities

$$F_-^{jk} = F_-^{j'k'}, \quad \text{and} \quad F_-^{j'k'} = F_-^{j'k}.$$

Let  $\theta_{A,A'}$  be the shift map by the vector  $-v_{\tau_L+1} + v'_{\tau_L+1}$  (where  $v_{\tau_L+1}$  is that cut-point in  $(Z_j, A)$  and  $v'_{\tau_L+1}$  is that in  $(Z'_j, A')$ ) and let  $\theta_{A',A}$  be the shift by  $-v'_{\tau_L+1} + v_{\tau_L+1}$ . Then observe that

$$\theta_{A',A} F_+^{j'k'} = F_+^{jk'}, \quad \text{and} \quad \theta_{A,A'} F_+^{jk} = F_+^{j'k}.$$

Using this decomposition of the faces in the four interfaces, we can express

$$\begin{aligned}
 & \left| \sum_{f \in \mathcal{I}^{j'k'}} \mathbf{g}(f, \mathcal{I}^{j'k'}) + \sum_{f \in \mathcal{I}^{jk}} \mathbf{g}(f, \mathcal{I}^{jk}) - \sum_{f \in \mathcal{I}^{j'k}} \mathbf{g}(f, \mathcal{I}^{j'k}) - \sum_{f \in \mathcal{I}^{jk'}} \mathbf{g}(f, \mathcal{I}^{jk'}) \right| \\
 & \leq \sum_{f \in F_+^{jk}} |\mathbf{g}(f, \mathcal{I}^{jk}) - \mathbf{g}(\theta_{A,A'} f, \mathcal{I}^{j'k})| + \sum_{f \in F_+^{j'k'}} |\mathbf{g}(\theta_{A',A} f, \mathcal{I}^{j'k'}) - \mathbf{g}(f, \mathcal{I}^{jk'})| \\
 & \quad + \sum_{f \in F_-^{jk}} |\mathbf{g}(f, \mathcal{I}^{jk}) - \mathbf{g}(f, \mathcal{I}^{j'k})| + \sum_{f \in F_-^{j'k'}} |\mathbf{g}(f, \mathcal{I}^{j'k'}) - \mathbf{g}(f, \mathcal{I}^{jk'})|. \quad (7.2)
 \end{aligned}$$

We begin by bounding the contributions of faces above  $X_{\tau_L}$ , i.e., the first line of (7.2); we write the bound for one of the sums as the other will evidently be analogous:

$$\sum_{f \in F_+^{jk}} |\mathbf{g}(f, \mathcal{I}^{jk}) - \mathbf{g}(\theta_{A,A'} f, \mathcal{I}^{j'k})| \leq \sum_{f \in F_+^{jk}} \bar{K} e^{-\bar{c}\mathbf{r}(f, \mathcal{I}^{jk}; \theta_{A,A'} f, \mathcal{I}^{j'k})}.$$

By tameness of all of  $\mathcal{I}^{jk}, \mathcal{I}^{j'k}, \mathcal{I}^{j'k'}$ , and  $\mathcal{I}^{jk'}$ , the radius  $\mathbf{r}$  is either attained by a face below  $v_{\tau_{\text{SP}}}$ , in which case for a face  $f \in \mathcal{F}(X_{\tau_L} + i)$ , the radius is at least  $\tau_L - \tau_{\text{SP}} + i \geq L + i$  or, it is attained in the differences between the spines  $(X_{\tau_{\text{SP}}}, \dots, X_{\tau_L})$  and  $(X'_{\tau_{\text{SP}}}, \dots, X'_{\tau_L})$ —but since all the increments between  $\tau_L - L$  and  $\tau_L$  are trivial both in  $\mathcal{I}^{jk}$  and  $\mathcal{I}^{j'k}$ , this distance would be at least  $L + i$ . The above is at most  $\bar{K}$  times

$$\sum_{i \geq 1} \sum_{f \in \mathcal{F}(X_{\tau_L+i})} e^{-\bar{c}(L+i)} + \sum_{f \in \mathcal{F}(X_{>T})} e^{-\bar{c}(T+1-\tau_L+L)} \leq \sum_{i \geq 1} |\mathcal{F}(X_{\tau_L+i})| e^{-\bar{c}(L+i)} + |\mathcal{F}(X_{>T})| e^{-\bar{c}(T+1-\tau_L+L)},$$

and the fact that the interfaces are both in  $\Gamma_L \subset \Gamma_{i > \tau_L}$  implies this is at most  $\bar{K} e^{-\bar{c}L/2}$ . The sum over  $f \in F_+^{j'k'}$  in the first line of (7.2) is handled identically. Next, we consider the contributions from the increments below  $\tau_L$  as well as in the truncated interface, say the faces in  $F_-^{jk}$  (the sum over  $f \in F_-^{j'k'}$  is again identical). Notice that for these faces, the radius  $\mathbf{r}$  is attained by a face in  $F_+^{jk} \oplus F_+^{j'k'}$  with increment index at least  $\tau_L + L$ , so that

$$\begin{aligned}
 \sum_{f \in F_-^{jk}} |\mathbf{g}(f, \mathcal{I}^{jk}) - \mathbf{g}(f, \mathcal{I}^{j'k})| & \leq \sum_{i \leq L} 4\bar{K} e^{-\bar{c}(L+i)} + \sum_{\tau_{\text{SP}} \leq i \leq \tau_L - L} \sum_{f \in X_i} \sum_{g \in F_+^{jk} \cup F_+^{j'k'}} \bar{K} e^{-\bar{c}d(f,g)} \\
 & \quad + \sum_{f \in \mathcal{I}_{\text{TR}}} \sum_{g \in F_+^{jk} \cup F_+^{j'k'}} \bar{K} e^{-\bar{c}d(f,g)}.
 \end{aligned}$$

By tameness, the distance between any face in  $\mathcal{I}_{\text{TR}}$  to a face  $g \in F_+^{jk} \cup F_+^{j'k'}$  that is in the  $(\tau_L + i')$ -th increment, is at least,  $\tau_L + i' - \tau_{\text{SP}} \geq i' + L$ , so that

$$\begin{aligned}
 \sum_{f \in F_-^{jk}} |\mathbf{g}(f, \mathcal{I}^{jk}) - \mathbf{g}(f, \mathcal{I}^{j'k})| & \leq 4\bar{C} e^{-\bar{c}L} + \sum_{i' \geq 1} \sum_{g \in X_{\tau_L+i'} \cup X'_{\tau_L+i'}} \sum_{\substack{f \in \mathcal{F}(\mathbb{Z}^3) \\ \text{ht}(f) \leq \text{ht}(v_{\tau_L})}} \bar{K} e^{-\bar{c}d(g,f)} \\
 & \leq 4\bar{C} e^{-\bar{c}L} + \sum_{i' \geq 1} |\mathcal{F}(X_{\tau+i'})| \vee |\mathcal{F}(X'_{\tau+i'})| \bar{C} e^{-\bar{c}(L+i')}.
 \end{aligned}$$

which is at most  $5\bar{C} e^{-\bar{c}L/2}$  by our assumption that the interfaces are in  $\Gamma_L$ . Combining all of these in to (7.2) and using our choice of  $L$  yields Lemma 7.7. This completes the proof of Proposition 7.1.

**7.2. Proof of Proposition 7.2: spine increments are asymptotically stationary.** In this section, we prove Proposition 7.2, showing that spine increments are asymptotically stationary in the sense that changing the conditioning from  $T$  to  $T'$  and the location of an increment stretch from  $j = j_T$  to  $j' = j'_{T'}$  does not change the law much as long as  $j$  and  $j'$  are in the bulks of their respective spines. Up to the choice of the two-to-two map, which is tailored to proving stationarity estimates here, much of the proof will match that of the mixing and we therefore omit some repeated details.

Fix any  $\gamma$ , and let  $K$  be such that if  $c_{\mathcal{B}}$  is the constant from Proposition 5.1,  $c_{\mathcal{B}}\beta K > 2\gamma$ . Next fix  $j, j'$  and  $s$  satisfying the required conditions, and let

$$L = \left\lceil \frac{4\gamma}{\bar{c}} \log D \right\rceil, \quad \text{where} \quad D = (j - K \log T) \wedge (j' - K \log T') \wedge (T' - (j' + s)) \wedge (T - (j + s)).$$

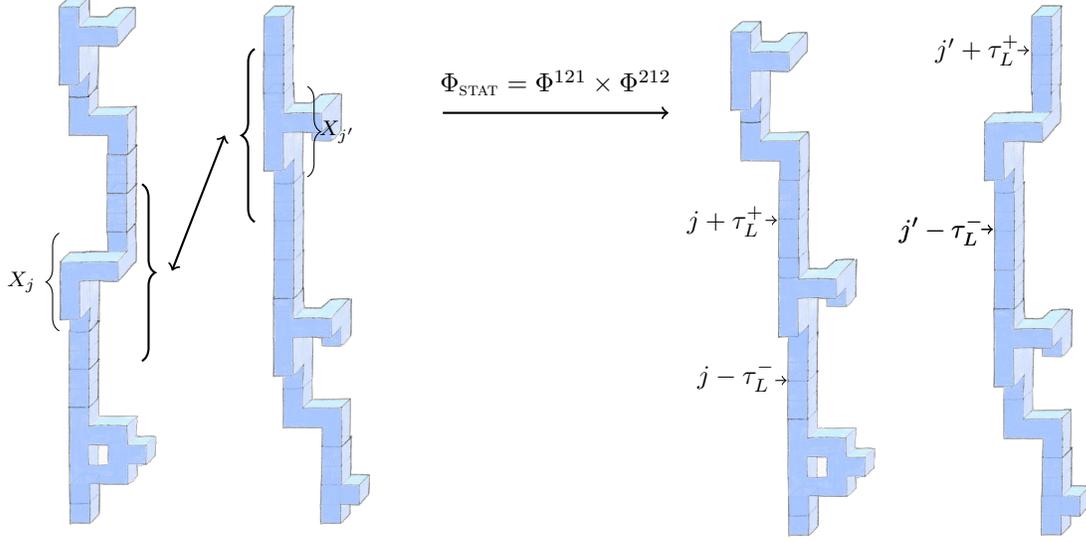


FIGURE 12. The map  $\Phi_{\text{STAT}} = \Phi^{121} \times \Phi^{212}$  acting on a pair of increment sequences in  $\Gamma_{\emptyset, L}$ .

Due to our freedom to take  $C$  as desired, we may assume without loss that  $D$  is sufficiently large. Let us denote the tuples  $\mathcal{X}_j = (\mathcal{X}_j, \dots, \mathcal{X}_{j+s})$  and  $\mathcal{X}_{j'} = (\mathcal{X}_{j'}, \dots, \mathcal{X}_{j'+s})$ , with fixed instantiations  $Z_j$  and  $Z_{j'}$  in  $\mathfrak{X}^s$ . Let  $\mathcal{A}_T$  denote the set of all  $T$ -admissible truncated interfaces along with increment sequences  $(X_i)_{i \notin \llbracket j, j+s \rrbracket}$  and remainder increment  $X_{>T}$ . Begin by expressing the left-hand side in Proposition 7.2 as

$$\left| \sum_{Z_j \in E} \sum_{A_T \in \mathcal{A}_T} \pi(Z_j, A_T) \sum_{\tilde{Z}_{j'} \in \mathfrak{X}^s, A_{T'} \in \mathcal{A}_{T'}} \pi_{T'}(\tilde{Z}_{j'}, A_{T'}) - \sum_{\tilde{A}_T \in \mathcal{A}_T, \tilde{Z}_j \in \mathfrak{X}^s} \pi_{T'}(\tilde{Z}_j, \tilde{A}_T) \sum_{Z_{j'} \in E} \sum_{\tilde{A}_{T'} \in \mathcal{A}_{T'}} \pi_T(Z_{j'}, \tilde{A}_{T'}) \right|$$

We follow the same strategy of the proof of Proposition 7.1. Namely, define the events  $\Gamma_{\emptyset, L}^-$  and  $\Gamma_{\emptyset, L}^+$  as the following subsets of pairs of increment sequences  $(\mathfrak{X}^T \times \mathfrak{X}_{\text{REM}}) \times (\mathfrak{X}^{T'} \times \mathfrak{X}_{\text{REM}})$ ,

$$\Gamma_{\emptyset, L}^- = \{((X_i)_{i \leq T}, (X'_i)_{i \leq T'}) : \exists \tau_L^- \in \llbracket L, D-L \rrbracket \text{ s.t. } X_{j-\tau_L^-+i} = X'_{j'-\tau_L^-+i} = X_{\emptyset} \text{ for all } i = -L, \dots, L\},$$

$$\Gamma_{\emptyset, L}^+ = \{((X_i)_{i \leq T}, (X'_i)_{i \leq T'}) : \exists \tau_L^+ \in \llbracket L+s, D-L \rrbracket \text{ s.t. } X_{j+\tau_L^++i} = X'_{j'+\tau_L^++i} = X_{\emptyset} \text{ for all } i = -L, \dots, L\}.$$

We can now define a map  $\Phi$  that takes a pair of interfaces and swaps the increment stretch between  $j - \tau_L^-$  and  $j + \tau_L^+$  in  $\mathcal{I}$  with the stretch between  $j' - \tau_L^-$  and  $j' + \tau_L^+$  in  $\mathcal{I}'$ : refer to Figure 12 for a visualization.

**Definition 7.8.** For any  $j, j'$ , let  $\Phi_{\text{STAT}} = \Phi_{\text{STAT}}(j, j') : (\mathfrak{X}^T \times \mathfrak{X}_{\text{REM}})^2 \rightarrow (\mathfrak{X}^{T'} \times \mathfrak{X}_{\text{REM}})^2$  be given as follows. For any pair of increment sequences  $(X^{(1)}, X^{(2)}) = ((X_i^{(1)})_i, X_{>T}^{(1)}, (X_i^{(2)})_i, X_{>T'}^{(2)})$ , let

$$\Phi_{\text{STAT}}(X^{(1)}, X^{(2)}) = (\Phi^{121}(X^{(1)}, X^{(2)}), \Phi^{212}(X^{(1)}, X^{(2)}))$$

be attained as follows. If  $(X^{(1)}, X^{(2)}) \notin \Gamma_{\emptyset, L}^- \cap \Gamma_{\emptyset, L}^+$ , let  $\Phi_{\text{STAT}}(X^{(1)}, X^{(2)}) = (X^{(1)}, X^{(2)})$ ; otherwise

(1) Find the smallest indices  $\tau_L^-$  and  $\tau_L^+$  for which the events  $\Gamma_{\emptyset, L}^-$  and  $\Gamma_{\emptyset, L}^+$  are satisfied.

(2) Let  $\Phi^{121}(X^{(1)}, X^{(2)})$  is the pair of increment sequences given by

$$\Phi^{121}(X^{(1)}, X^{(2)}) = (X_1^{(1)}, \dots, X_{j-\tau_L^-}^{(1)}, X_{j'-\tau_L^-+1}^{(2)}, \dots, X_{j'}^{(2)}, \dots, X_{j'+\tau_L^+}^{(2)}, X_{j+\tau_L^++1}^{(1)}, \dots, X_T^{(1)}, X_{>T}^{(1)}).$$

(3) Let  $\Phi^{212}(X^{(1)}, X^{(2)})$  is the pair of increment sequences given by

$$\Phi^{212}(X^{(1)}, X^{(2)}) = (X_1^{(2)}, \dots, X_{j'-\tau_L^-}^{(2)}, X_{j-\tau_L^-+1}^{(1)}, \dots, X_j^{(1)}, \dots, X_{j+\tau_L^+}^{(1)}, X_{j'+\tau_L^++1}^{(2)}, \dots, X_{T'}^{(2)}, X_{>T'}^{(2)}).$$

Abusing notation, we can define  $\Phi_{\text{STAT}}$  on  $\bar{\mathbf{I}}_{x, T} \times \bar{\mathbf{I}}_{x, T}$  that uses the same truncations of the pair  $(\mathcal{I}, \mathcal{I}')$  but applies the map  $\Phi_{\text{STAT}}$  to their increment sequences in the pillars  $(\mathcal{P}_x, \mathcal{P}'_x)$ . If the two interfaces are both tame and additionally satisfy  $\tau_{\text{SP}} \leq K \log T$  and  $\tau_{\text{SP}}' \leq K \log T'$ , the resulting pair would be in  $\mathbf{I}_{x, T} \times \mathbf{I}_{x, T}$ .

7.2.1. *Strategy of the map  $\Phi_{\text{STAT}}$ .* Similarly to the mixing map, if one were to take a naive approach of constructing a map that sends a single interface to a single interface, a possible choice would be a map that e.g., inserts an increment  $X_0$  at the bottom of the increment sequence, shifting the remainder of the increment sequence and showing that the weights of interfaces with  $\mathcal{X}_i = X$  are close to those with  $\mathcal{X}_{i+1} = X$ . (Notice that any map we construct must increase the number of increments as we wish to show not only that the law is close to stationary in shifts for fixed  $T$ , but that it remains stationary as  $T \rightarrow \infty$ .) Similar to the explanation in Section 7.1.1, however, the addition of an increment means that, the best one could hope for is a ratio of weights that is  $1 \pm \varepsilon_\beta$ , rather than  $1 + o_T(1)$ .

Instead, we use the two-to-two map which shifts an increment  $X_j$  in a spine of  $T$  increments, to a position  $j'$  in a spine of  $T'$  increments. As with  $\Phi_{\text{MIX}}$ , we use that  $j, j'$  are far from  $1, T \wedge T'$  to find paired stretches of trivial increments equal distances above and below  $X_j$  and  $X_{j'}$ . We then splice in the middle of these trivial increment sequences, and use them to decorrelate  $X_j, X_{j'}$  from the rest of their respective interfaces, showing that the relative weight of the pair of interfaces is almost unchanged by the map  $\Phi_{\text{STAT}}$ . Refer to Figure 12 for a visualization of this map.

7.2.2. *Analysis of the map  $\Phi_{\text{STAT}}$ .* We now define, analogously to the proof of mixing, a good set of pairs of increment sequences, denoted  $\Gamma_L$  on which we can control the ratio of probabilities under the map  $\Phi$ . Let  $\Gamma_L$  be the set of  $(\mathcal{I}, \mathcal{I}') \in \bar{\mathbf{I}}_{x,T} \times \bar{\mathbf{I}}_{x,T'}$  such that its pair of increment sequences are in  $\Gamma_{\emptyset,L}^- \cap \Gamma_{\emptyset,L}^+$ , and have

- (1) Their source point indices satisfy  $\tau_{\text{sp}} \leq K \log T$  and  $\tau_{\text{sp}}' \leq K \log T'$ ; denote this event  $\Gamma_{\tau_{\text{sp}}}$ .
- (2) The pair of interfaces  $(\mathcal{I}, \mathcal{I}')$  are such that  $\Phi_{\text{STAT}}(\mathcal{I}, \mathcal{I}')$  are both tame; denote this event  $\bar{\Gamma}$ .
- (3) The increment sequence  $(X_i)$  satisfies the events (denoted  $\Gamma_{i>\tau_L^-}$  and  $\Gamma_{i>\tau_L^+}$ )

$$|\mathcal{F}(X_{>T})| e^{-\bar{c}(T+1-j+\tau_L^-+L)} + \sum_{i \geq 1} (|\mathcal{F}(X_{j-\tau_L^-+i})|) e^{-\bar{c}(L+i)} \leq e^{-\bar{c}L/2}, \quad \text{as well as} \quad (7.3)$$

$$|\mathcal{F}(X_{>T})| e^{-\bar{c}(T+1-j-\tau_L^++L)} + \sum_{i \geq 1} (|\mathcal{F}(X_{j+\tau_L^++i})|) e^{-\bar{c}(L+i)} \leq e^{-\bar{c}L/2}. \quad (7.4)$$

and  $(X'_i)$  satisfies the analogous events with respect to  $T'$  (denoted  $\Gamma'_{i>\tau_L^-}$  and  $\Gamma'_{i>\tau_L^+}$ ).

As in the proof of Proposition 7.1, we can bound the contribution from pairs of interfaces not in  $\Gamma_L$  by

$$\begin{aligned} \pi_T^{\otimes 2}(\Gamma_L^c) &\leq 2\pi_T(\Gamma_{\tau_{\text{sp}}}^c) + \pi_T^{\otimes 2}(\bar{\Gamma}^c | \Gamma_{\tau_{\text{sp}}}) + \pi_T^{\otimes 2}((\Gamma_{\emptyset,L}^-)^c | \Gamma_{\tau_{\text{sp}}}) + \pi_T^{\otimes 2}((\Gamma_{\emptyset,L}^+)^c | \Gamma_{\tau_{\text{sp}}}) + \pi_T(\Gamma_{i>\tau_L^-}^c | \Gamma_{\emptyset,L}, \Gamma_{\tau_{\text{sp}}}) \\ &\quad + \pi_T(\Gamma_{i>\tau_L^+}^c | \Gamma_{\emptyset,L}, \Gamma_{\tau_{\text{sp}}}) + \pi_T((\Gamma'_{i>\tau_L^-})^c | \Gamma_{\emptyset,L}, \Gamma_{\tau_{\text{sp}}}) + \pi_T((\Gamma'_{i>\tau_L^+})^c | \Gamma_{\emptyset,L}, \Gamma_{\tau_{\text{sp}}}). \end{aligned}$$

The bounds on the first two terms above are identical to those in the proof of Lemma 7.5, so that their contribution is at most  $T^{-2\gamma}$ . The bounds on the third and fourth terms are as in the proof of Lemma 7.5, noticing that on  $\Gamma_{\tau_{\text{sp}}}$ , the sequence of indicator functions  $(\mathbf{1}_{\{X_{j-i}=X_\emptyset\}} \mathbf{1}_{\{X'_{j'-i}=X_\emptyset\}})_{i \leq D}$  stochastically dominate i.i.d.  $\text{Ber}((1 - \varepsilon_\beta)^2)$  random variables; therefore, their contribution is at most  $\exp[-D^{3/4}]$  once  $\beta$  is sufficiently large (independently of  $j, j'$ ). The sixth and eighth terms above are also bounded as in Lemma 7.5 by  $2D^{-2\gamma}$  using the conditional version of Corollary 4.2.

A crucial difference arises in the bounds on the fifth and seventh terms, since knowledge of  $\tau_L^-$  gives information regarding the increment sequence above index  $\tau_L^- + L$  (namely that there is no possible smaller choice of  $\tau_L^-$ ), so Corollary 4.2 does not immediately bound  $\pi_T(\Gamma_{i>\tau_L^-}^c | \Gamma_{\emptyset,L}, \Gamma_{\tau_{\text{sp}}})$ . Instead, we union bound over the  $D$  possible choices of  $\tau_L^-$  and sustaining this union bound, see that

$$\pi_T(\Gamma_{i>\tau_L^-}^c | \Gamma_{\emptyset,L}, \Gamma_{\tau_{\text{sp}}}) \leq D e^{-c\beta\bar{c}L/2} \leq D^{-2\gamma+1}$$

as long as  $\beta c > 1$ , and likewise for  $\pi_T((\Gamma'_{i>\tau_L^-})^c | \Gamma_{\emptyset,L}, \Gamma_{\tau_{\text{sp}}})$ . Combining all these estimates yields the desired bound of  $\pi_T^{\otimes 2} \leq CD^{-\gamma}$  for  $\beta$  sufficiently large (depending on  $\gamma$  and  $\bar{c}$ ).

Now that we've restricted to nice pairs of increment sequences, we can naturally view  $\Phi_{\text{STAT}}$  as a map on  $(\bar{\mathbf{I}}_{x,T} \times \bar{\mathbf{I}}_{x,T'}) \cap \Gamma_L$ : as in Claim 7.6, we arrive at the following claim.

**Claim 7.9.** *The restriction of  $\Phi_{\text{STAT}}$  to  $(\bar{\mathbf{I}}_{x,T} \times \bar{\mathbf{I}}_{x,T'}) \cap \Gamma_L$  is a bijection from  $(\bar{\mathbf{I}}_{x,T} \times \bar{\mathbf{I}}_{x,T'}) \cap \Gamma_L$  to itself.*

We are therefore left to bound

$$\sum_{Z_j \in E} \sum_{\tilde{Z}_{j'}, A_T, A_{T'}} \left| \pi_T(Z_j, A_T) \pi_T(\tilde{Z}_{j'}, A_{T'}) - \pi_T(\Phi^{121}((Z_j, A_T), (\tilde{Z}_{j'}, A_{T'}))) \pi_T(\Phi^{212}((Z_j, A_T), (\tilde{Z}_{j'}, A_{T'}))) \right|.$$

In order to bound the summands above, as before, let us focus on the ratio of the probabilities under application of  $\Phi_{\text{STAT}}$ , and use the machinery of Theorem 2.21.

We will use the short-hand  $\mathcal{I}^{ZA} = (Z_j, A_T)$  for the interface in  $\bar{\mathbf{I}}_{x,T}$  with  $\mathcal{Z}_j = Z_j$  and  $A_T$  elsewhere, and  $\mathcal{I}^{Z'A'} = (\tilde{Z}_{j'}, A_{T'})$  for the interface in  $\bar{\mathbf{I}}_{x,T'}$  that has  $\mathcal{Z}_{j'} = \tilde{Z}_{j'}$  and  $A_{T'}$  elsewhere. Moreover, let

$$\mathcal{I}^{ZA'} = \Phi^{212}(\mathcal{I}^{ZA}, \mathcal{I}^{Z'A'}), \quad \text{and} \quad \mathcal{I}^{Z'A} = \Phi^{121}(\mathcal{I}^{ZA}, \mathcal{I}^{Z'A'}).$$

(In particular,  $A'$  (resp.,  $A$ ) are interfaces of  $T'$  (resp.,  $T$ ) increments and the  $Z$  or  $Z'$  in the superscript indicates whether the increments in indices  $j' - \tau_L^-, \dots, j' + \tau_L^+$  (resp.,  $j - \tau_L^-, \dots, j + \tau_L^+$ ) are those coming from  $j - \tau_L^-, \dots, j + \tau_L^+$  in  $(Z, A)$  or  $j' - \tau_L^-, \dots, j' + \tau_L^+$ .) Then, for any such interfaces in  $(\bar{\mathbf{I}}_{x,T} \times \bar{\mathbf{I}}_{x,T'}) \cap \Gamma_L$ ,

$$\begin{aligned} \frac{\pi_T(\mathcal{I}^{ZA}) \pi_T(\mathcal{I}^{Z'A'})}{\pi_T(\mathcal{I}^{ZA'}) \pi_T(\mathcal{I}^{Z'A})} &= \frac{\mu_n(\mathcal{I}^{ZA}) \mu_n(\mathcal{I}^{Z'A'})}{\mu_n(\mathcal{I}^{ZA'}) \mu_n(\mathcal{I}^{Z'A})} \\ &= \exp \left( \sum_{f \in \mathcal{I}^{ZA}} \mathbf{g}(f, \mathcal{I}^{ZA}) + \sum_{f \in \mathcal{I}^{Z'A'}} \mathbf{g}(f, \mathcal{I}^{Z'A'}) - \sum_{\tilde{f} \in \mathcal{I}^{Z'A}} \mathbf{g}(\tilde{f}, \mathcal{I}^{Z'A}) + \sum_{f \in \mathcal{I}^{Z'A'}} \mathbf{g}(f, \mathcal{I}^{Z'A'}) \right). \end{aligned}$$

Proposition 7.2 then follows from the following lemma, just as in the proof of Proposition 7.1.

**Lemma 7.10.** *There is a universal constant  $\bar{C}$  such that for any pair  $(\mathcal{I}, \mathcal{I}') \in \Gamma_L$ , we have*

$$\left| \sum_{f \in \mathcal{I}^{ZA}} \mathbf{g}(f, \mathcal{I}^{ZA}) + \sum_{f \in \mathcal{I}^{Z'A'}} \mathbf{g}(f, \mathcal{I}^{Z'A'}) - \sum_{f \in \mathcal{I}^{Z'A}} \mathbf{g}(f, \mathcal{I}^{Z'A}) - \sum_{f \in \mathcal{I}^{Z'A'}} \mathbf{g}(f, \mathcal{I}^{Z'A'}) \right| \leq \bar{C} \exp(-\bar{c}L/2).$$

We wish to bound the absolute value of the quantity in the exponential by pairing various subsets of the different interfaces together in a manner that they look locally alike. We denote by  $F_{\text{INT}}^{ZA}$  the face set of the increments in  $\mathcal{I}^{ZA}$  between index  $j - \tau_L^-$  and  $j + \tau_L^+$  and denote the two connected components of  $\mathcal{I}^{ZA} \setminus F_{\text{INT}}^{ZA}$  by  $F_-^{ZA}$  and  $F_+^{ZA}$  respectively. Likewise define the  $F_{\text{INT}}, F_-$  and  $F_+$  for  $\mathcal{I}^{Z'A'}, \mathcal{I}^{Z'A}, \mathcal{I}^{Z'A}$ , where if the superscript is  $A'$ , the interior will have indices  $j' - \tau_L^-$  and  $j' + \tau_L^+$  (instead of  $j - \tau_L^-$  and  $j + \tau_L^+$ ).

Notice that  $F_-^{ZA} = F_-^{Z'A}$  and  $F_-^{Z'A'} = F_-^{Z'A}$ . We can then define the shift maps  $\theta_{A,A'}^{(1)}$  which is the shift by the vector  $-v_{j-\tau_L^-+1} - v'_{j'-\tau_L^-+1}$  (where  $v_{j-\tau_L^-+1}$  is the cut-point in  $\mathcal{I}^{ZA}$  and  $v'_{j'-\tau_L^-+1}$  is the cut-point in  $\mathcal{I}^{Z'A'}$ ) and  $\theta_{A,A'}^{(2)}$  which is the shift by the vector  $-v_{j+\tau_L^+} + v'_{j'+\tau_L^+}$ . With these definitions, we see that

$$\begin{aligned} \theta_{A,A'}^{(1)} F_{\text{INT}}^{ZA} &= F_{\text{INT}}^{Z'A'}, & \theta_{A',A}^{(1)} F_{\text{INT}}^{Z'A'} &= F_{\text{INT}}^{Z'A}, & \text{and} \\ \theta_{A,A'}^{(2)} \theta_{A',A}^{(1)} F_+^{ZA} &= F_+^{Z'A'}, & \theta_{A',A}^{(2)} \theta_{A',A}^{(1)} F_+^{Z'A'} &= F_+^{Z'A}. \end{aligned}$$

With this decomposition, we see that

$$\begin{aligned} & \left| \sum_{f \in \mathcal{I}^{ZA}} \mathbf{g}(f, \mathcal{I}^{ZA}) + \sum_{f \in \mathcal{I}^{Z'A'}} \mathbf{g}(f, \mathcal{I}^{Z'A'}) - \sum_{f \in \mathcal{I}^{Z'A}} \mathbf{g}(f, \mathcal{I}^{Z'A}) - \sum_{f \in \mathcal{I}^{Z'A'}} \mathbf{g}(f, \mathcal{I}^{Z'A'}) \right| \\ & \leq \sum_{f \in F_+^{ZA}} \left| \mathbf{g}(f, \mathcal{I}^{ZA}) - \mathbf{g}(\theta_{A,A'}^{(2)} \theta_{A',A}^{(1)} f, \mathcal{I}^{Z'A'}) \right| + \sum_{f \in F_+^{Z'A'}} \left| \mathbf{g}(f, \mathcal{I}^{Z'A'}) - \mathbf{g}(\theta_{A',A}^{(2)} \theta_{A',A}^{(1)} f, \mathcal{I}^{Z'A'}) \right| \\ & \quad + \sum_{f \in F_{\text{INT}}^{ZA}} \left| \mathbf{g}(f, \mathcal{I}^{ZA}) - \mathbf{g}(\theta_{A,A'}^{(1)} f, \mathcal{I}^{Z'A'}) \right| + \sum_{f \in F_{\text{INT}}^{Z'A'}} \left| \mathbf{g}(f, \mathcal{I}^{Z'A'}) - \mathbf{g}(\theta_{A',A}^{(1)} f, \mathcal{I}^{Z'A'}) \right| \\ & \quad + \sum_{f \in F_-^{ZA}} \left| \mathbf{g}(f, \mathcal{I}^{ZA}) - \mathbf{g}(f, \mathcal{I}^{Z'A}) \right| + \sum_{f \in F_-^{Z'A'}} \left| \mathbf{g}(f, \mathcal{I}^{Z'A'}) - \mathbf{g}(f, \mathcal{I}^{Z'A}) \right|. \end{aligned} \quad (7.5)$$

The first two terms are bounded above by  $O(e^{-\bar{c}L/2})$  analogously to the contribution of faces in  $F_+^{jk}$  in (7.2); by construction for a face in the  $(j + \tau_L^+ + i)$ -th increment, the radius  $\mathbf{r}(f, \mathcal{I}^{ZA}; \theta_{A,A'}^{(2)} \theta_{A',A}^{(1)} f, \mathcal{I}^{Z'A'})$  is at least  $L + i$ ; the first  $L$  such increments have exactly four faces, and their contribution is thus at most  $4\bar{C}e^{-\bar{c}L}$ , while the contributions of increments above  $j + \tau_L^+ + L$  is bounded by  $\bar{K}e^{-\bar{c}L/2}$  by (7.4).

The last two terms in (7.5) are bounded in the same manner as the term  $F_-^{jk}$  in (7.2); for these faces, the radius of congruence is attained by some face in  $F_{\text{INT}}^{ZA} \cup \theta_{A,A'}^{(1)} F_{\text{INT}}^{ZA} \cup F_+^{ZA} + \theta_{A,A'}^{(2)} \theta_{A',A}^{(1)} F_+^{ZA}$ . Then the set of faces can be split into those faces that are between increment  $j - \tau_L^-$  and  $j - \tau_L^- - L$ , whose contribution is easily seen to be at most  $4\bar{C}e^{-\bar{c}L}$ , and those that are below increment  $j - \tau_L^- - L$  along with the truncation  $\mathcal{I}_{\text{TR}}$ . The contribution of these latter faces is bounded as in the bound of (7.2), by additionally summing over the possible faces that attain the radius of convergence, and using integrability of exponential tails to reduce this to a multiple of the quantities (7.3)–(7.4).

It remains to bound the contribution of the middle two terms, say that of faces in  $F_{\text{INT}}^{ZA}$ . These terms can be bounded by decomposing into the event that the radius of congruence is attained by a face in  $F_-^{ZA} \cup F_-^{Z'A'}$  and the event that it is attained by a face in  $F_+^{ZA} \cup \theta_{A',A}^{(2)} \theta_{A,A'}^{(1)} F_+^{Z'A'}$ . In the former case, these terms are treated analogously to the first two terms, and therefore their contribution is at most  $\bar{C}e^{-\bar{c}L/2}$  by (7.3). In the latter case, they are treated analogously to the last two terms, swapping the summation into one over faces in  $F_+^{ZA} \cup \theta_{A',A}^{(2)} \theta_{A,A'}^{(1)} F_+^{Z'A'}$ , and the contribution is at most  $O(e^{-\bar{c}L/2})$ , by (7.4). ■

**7.3. Proof of Corollary 7.3: existence of a limiting measure.** We first claim that for each  $k$ , the subsequence of measures

$$\left( \pi_T((\mathcal{X}_{\frac{T}{2}-k}, \dots, \mathcal{X}_{\frac{T}{2}}, \dots, \mathcal{X}_{\frac{T}{2}+k}) \in \cdot) \right)_T$$

is a Cauchy sequence in the total-variation metric: indeed for every  $T' \geq T$ , we have by Proposition 7.2 that

$$\|\pi_T((\mathcal{X}_{\frac{T}{2}-k}, \dots, \mathcal{X}_{\frac{T}{2}}, \dots, \mathcal{X}_{\frac{T}{2}+k}) \in \cdot) - \pi_{T'}((\mathcal{X}_{\frac{T'}{2}-k}, \dots, \mathcal{X}_{\frac{T'}{2}}, \dots, \mathcal{X}_{\frac{T'}{2}+k}) \in \cdot)\|_{\text{TV}} \leq C \left(\frac{T}{2} - k\right)^{-\gamma}.$$

By completeness of the space of probability measures on  $\mathfrak{X}^{2k}$  with respect to the total-variation distance, this implies that for each  $k$ , there exists a limiting measure  $\nu_k$  on  $\mathfrak{X}^{2k}$  such that the marginals above converge to  $\nu_k$ . If the family  $(\nu_k)$  is viewed as marginals on  $\mathcal{X}_{-k}, \dots, \mathcal{X}_0, \dots, \mathcal{X}_k$  of a limiting law  $\nu$  on  $\mathfrak{X}^{\mathbb{Z}}$ , the Kolmogorov consistency criterion is trivially satisfied as these finite-dimensional distributions are arising as limits of marginals of a single consistent distribution (the law of  $\mathcal{X}_1, \dots, \mathcal{X}_T$  under  $\pi_T$  viewed about  $\mathcal{X}_{\frac{T}{2}}$ ).

To see that any other sequence  $a_T$  satisfying  $(a_T \vee T - a_T) \gg \log T$  has the same limit, take any such  $a_T$  (without loss of generality  $a_T \leq T/2$ ) as well as any  $k$ , and bound

$$\begin{aligned} & \|\pi_T((\mathcal{X}_{a_T-k}, \dots, \mathcal{X}_{a_T}, \dots, \mathcal{X}_{a_T+k}) \in \cdot) - \nu((\mathcal{X}_{-k}, \dots, \mathcal{X}_0, \dots, \mathcal{X}_k) \in \cdot)\|_{\text{TV}} \\ & \leq \|\pi_T((\mathcal{X}_{\frac{T}{2}-k}, \dots, \mathcal{X}_{\frac{T}{2}}, \dots, \mathcal{X}_{\frac{T}{2}+k}) \in \cdot) - \nu((\mathcal{X}_{-k}, \dots, \mathcal{X}_0, \dots, \mathcal{X}_k) \in \cdot)\|_{\text{TV}} \\ & \quad + \|\pi_T((\mathcal{X}_{a_T-k}, \dots, \mathcal{X}_{a_T}, \dots, \mathcal{X}_{a_T+k}) \in \cdot) - \pi_T((\mathcal{X}_{\frac{T}{2}-k}, \dots, \mathcal{X}_{\frac{T}{2}}, \dots, \mathcal{X}_{\frac{T}{2}+k}) \in \cdot)\|_{\text{TV}}. \end{aligned}$$

The first term on the right-hand side above is  $o(1)$  as  $T \rightarrow \infty$  by the convergence of  $\pi_T$  to  $\nu$  in total-variation. The second term on the right-hand side above is at most  $C(a_T - K \log T)^{-\gamma}$  for  $K$  satisfying  $c_{\emptyset} \beta K > 2\gamma$  for some  $\gamma > 2$  by Proposition 7.2; this is also  $o(1)$  as  $T \rightarrow \infty$ . The two consequences of this follow immediately from the definition of weak convergence and Proposition 4.1 and Proposition 7.1. ■

## 8. MEAN AND VARIANCE OF OBSERVABLES OF THE INCREMENT SEQUENCE

In this section, we prove estimates for the mean and variance of running sums of increment observables  $f : \mathfrak{X} \rightarrow \mathbb{R}$  (these appear in e.g., Theorem 2). In Section 8.1 we prove that any function  $f$  with rotational or reflective symmetries, has mean zero under  $\nu$ . In Section 8.2, we prove that non-constant functions will have a variance that diverges linearly in  $T$  variance in the central limit theorem. These will, in particular, imply the choices of the mean and covariances in items (1)–(2) of Corollary 3.

**8.1. Anti-symmetric observables have mean zero.** In this section, we prove that for any observable  $f$  that is anti-symmetric in reflections or rotations in the  $xy$ -plane, its mean under  $\nu$  is zero as long as  $\beta$  is sufficiently large. In particular, its central limit theorem, holds without any recentering. The proof follows by applying a reflection map above an atypically long stretch of trivial increments, and seeing that this map essentially leaves the probability distribution over the increment sequence invariant.

We say a map  $\varphi : \mathfrak{X} \rightarrow \mathfrak{X}$  is a *reflection* map if it is a reflection about one of the two planes with normal vector  $e_1$  or  $e_2$  going through the point  $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ . We say it is a *rotation map* if it is a  $\frac{\pi n}{4}$  rotation about the  $e_3$  axis through  $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ . (Notice that the trivial increment  $X_{\emptyset}$  is fixed by any of these maps.) The same  $\varphi$  can naturally also be viewed as a map on remainder increments. (Note that  $\varphi(X_{\emptyset}) = X_{\emptyset}$ .)

**Proposition 8.1.** *There exists  $\beta_0 > 0$  such that the following holds for every  $\beta > \beta_0$ . If  $f : \mathfrak{X} \rightarrow \mathbb{R}$  satisfies  $|f(X)| \leq M$  and  $f(X) = -f(\varphi(X))$  for all  $X \in \mathfrak{X}$ , for some  $M \in \mathbb{R}$  and reflection or rotation map  $\varphi$ , then*

- (1)  $\mathbb{E}_\nu[f(\mathcal{X}_0)] = 0$ ;
- (2) if  $g : \mathfrak{X} \rightarrow \mathbb{R}$  has  $|g(X)| \leq M$  and  $g(\varphi(X)) = g(X)$  for every  $X$ , then

$$\mathbb{E}_\nu \left[ \sum_{i \in \mathbb{Z}} f(\mathcal{X}_0) g(\mathcal{X}_i) \right] = \sum_{i \in \mathbb{Z}} \text{Cov}_\nu(f(\mathcal{X}_0), g(\mathcal{X}_i)) = 0.$$

The proof of Proposition 8.1 relies crucially on bounding the effect of a map that reflects or rotates the pillar above some stretch of  $O(\log T)$  consecutive trivial increments. To that end, let  $\varphi$  be a reflection or rotation map and define the map  $\Phi_\varphi : \bar{\mathbf{I}}_{x,T} \rightarrow \bar{\mathbf{I}}_{x,T}$  as follows.

**Definition 8.2.** For a given  $L$ , we can denote  $\tau_L = \tau_L(\mathcal{I})$  as the smallest index greater than  $\tau_{\text{SP}} + L$  such that all of  $X_{\tau_L-L}, \dots, X_{\tau_L} = X_\emptyset$ . Then, for an interface  $\mathcal{I}$ , let  $\Phi_\varphi$  agree with  $\mathcal{I}$  on its truncation and its increments up to the  $(\tau_L - \frac{L}{2})$ -th increment, then apply the map  $\varphi$  to all increments with index ranging from  $\tau_L - \frac{L}{2}$  to  $T$ , as well as the remainder increment. Notice that this is the same as applying the map  $\varphi$  to the entire subset of the pillar above the  $(\tau_L - \frac{L}{2})$ -th increment, by correspondingly reflecting/rotating it about the  $e_3$  axis going through  $v_{\tau_L - \frac{L}{2} + 1}$ . (If  $\tau_L$  does not exist, then let  $\Phi_\varphi$  be the identity.)

**Claim 8.3.** *For every reflection or rotation map  $\varphi : \mathfrak{X} \rightarrow \mathfrak{X}$ , every  $\mathcal{I} \in \bar{\mathbf{I}}_{x,T}$ , and every  $\gamma > 1$ , there exists some  $\beta_0 > 0$  such that, for all  $\beta > \beta_0$ ,*

$$\left| \frac{\pi_T(\mathcal{I})}{\pi_T(\Phi_\varphi(\mathcal{I}))} - 1 \right| \leq O(T^{-\gamma+1}).$$

*Proof.* Fix any  $\gamma$  and let  $L = \lceil \frac{2\gamma}{\epsilon} \log T \rceil$ . By Theorem 2.21 and Definition 8.2, for  $\mathcal{I} \in \bar{\mathbf{I}}_{x,T}$ ,

$$\frac{\pi_T(\mathcal{I})}{\pi_T(\Phi_\varphi(\mathcal{I}))} = \frac{\mu_n(\mathcal{I})}{\mu_n(\Phi_\varphi(\mathcal{I}))} = \exp \left( \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \Phi_\varphi(\mathcal{I})} \mathbf{g}(f', \Phi_\varphi(\mathcal{I})) \right).$$

For every interface  $\mathcal{I}$ , let us split its faces up as  $\mathcal{I}^-$  denoting the union of  $\mathcal{I}_{\text{TR}}$  and the increment sequence up to  $X_{\tau_L - \frac{L}{2}}$ , and  $\mathcal{I}^+$  denoting the union of the increments above  $X_{\tau_L - \frac{L}{2}}$  along with the remainder  $X_{>T}$ . Moreover, for a face  $f \in \mathcal{S}_x$ , let  $\varphi(f)$  be the image of that face  $f$  under the reflection/rotation map  $\varphi$ , viewed as a face in  $\Phi_\varphi(\mathcal{I})$ . Then,

$$\left| \sum_{f \in \mathcal{I}} \mathbf{g}(f, \mathcal{I}) - \sum_{f' \in \Phi_\varphi(\mathcal{I})} \mathbf{g}(f', \Phi_\varphi(\mathcal{I})) \right| \leq \sum_{f \in \mathcal{I}^-} |\mathbf{g}(f, \mathcal{I}) - \mathbf{g}(f, \Phi_\varphi(\mathcal{I}))| + \sum_{f \in \mathcal{I}^+} |\mathbf{g}(f, \mathcal{I}) - \mathbf{g}(\varphi(f), \Phi_\varphi(\mathcal{I}))|.$$

It is clear that if  $f \in X_i$  for  $i > \tau_L - \frac{L}{2}$ , the radius  $\mathbf{r}(f, \mathcal{I}; \varphi(f), \Phi_\varphi(\mathcal{I}))$  is attained by a face a distance at least  $i - (\tau_L - L)$ , because the spine is tamed. We used crucially that in Theorem 2.21, the radius of congruence is congruence up to rotation and reflection in the  $xy$ -plane, and the increments between  $\tau_L - L$  and  $\tau_L$  are fixed by such reflection and rotations. Consequently,

$$\sum_{f \in \mathcal{I}^+} e^{-\bar{c}\mathbf{r}(f, \mathcal{I}; \varphi(f), \Phi_\varphi(\mathcal{I}))} \leq 4\bar{C}e^{-\bar{c}L/2} + \sum_{i \geq 1} \bar{K} |\mathcal{F}(X_{\tau_L+i})| e^{-\bar{c}(L+i)}$$

which is at most  $O(T^{-\gamma+1})$  by the tameness of  $\mathcal{I}$  and the choice of  $L$ .

At the same time, for each  $f \in \mathcal{I}^-$ , the radius  $\mathbf{r}(f, \mathcal{I}; f, \Phi_\varphi(\mathcal{I}))$  is attained by a face in  $\mathcal{I}^+ \cup \Phi_\varphi(\mathcal{I}^+)$ , so that proceeding as usual with these terms, their contribution is bounded by  $\bar{K}$  times

$$\begin{aligned} \sum_{f \in \mathcal{I}^-} e^{-\bar{c}\mathbf{r}(f, \mathcal{I}; f, \Phi_\varphi(\mathcal{I}))} &\leq \sum_{i \leq L/2} 4e^{-\bar{c}(i + \frac{L}{2})} + \sum_{g \in \mathcal{I}^+ \cup \Phi_\varphi(\mathcal{I}^+)} \left[ \sum_{f \in \mathcal{I}_{\text{TR}}} e^{-\bar{c}d(f, g)} + \sum_{\tau_{\text{SP}} \leq i \leq \tau_L - L} \sum_{f \in \mathcal{F}(X_i)} e^{-\bar{c}d(f, g)} \right] \\ &\leq 4\bar{C}e^{-\bar{c}L/2} + \sum_{i' \geq 1} \sum_{g \in \mathcal{F}(X_{\tau_L+i'}) \cup \mathcal{F}(\Phi_\varphi(X_{\tau_L+i'}))} \sum_{f \in \mathcal{F}(\mathbb{Z}^3): d(f, g) \geq L+i'} e^{-\bar{c}d(g, f)} \\ &\leq 4\bar{C}e^{-\bar{c}L/2} + \sum_{i' \geq 1} 2|\mathcal{F}(X_{\tau_L+i'})| \bar{C}e^{-\bar{c}(L+i')} \end{aligned}$$

which is at most  $O(T^{-\gamma+1})$  since  $\mathcal{I}$  is tame. Putting these together implies that for every tame interface  $\mathcal{I}$  (otherwise  $\pi_T(\mathcal{I}) = 0$ ), the log of the ratio of probabilities is  $O(T^{-\gamma+1})$  as desired.  $\blacksquare$

**Proof of Proposition 8.1.** By Corollary 7.3, and boundedness of  $f$ ,

$$\mathbb{E}_{\pi_T} \left[ \sum_{i \leq T} f(\mathcal{X}_i) \right] = O(\log T) + T(1 - o(1))\mathbb{E}_\nu[f(\mathcal{X}_i)].$$

Consequently, if we prove that the left-hand side is  $o(T)$ , it will imply that  $\mathbb{E}_\nu[f(\mathcal{X}_0)] = 0$ . We can split up

$$\mathbb{E}_{\pi_T} \left[ \sum_{i \leq T} f(\mathcal{X}_i) \right] = \mathbb{E}_{\pi_T} \left[ \sum_{i=1}^{\tau_L} f(\mathcal{X}_i) + \sum_{i=\tau_L+1}^T f(\mathcal{X}_i) \right],$$

for  $\tau_L$  as in Definition 8.2, and begin by bounding the first of these sums. Recall that by Corollary 4.2, the sequence  $(\mathbf{1}_{\{\mathcal{X}_i = X_\emptyset\}})_{i \geq \tau_{\text{sp}}}$  stochastically dominates a sequence of i.i.d.  $\text{Ber}(1 - \varepsilon_\beta)$  for some  $\varepsilon_\beta > 0$  satisfying  $\varepsilon_\beta \rightarrow 0$  as  $\beta \rightarrow \infty$ . Using this, we can estimate

$$\mathbb{E}_{\pi_T} \left[ \sum_{i \leq \tau_L} |f(\mathcal{X}_i)| \right] \leq MT\pi_T(\tau_L \geq T^{1/4}) + \mathbb{E}_{\pi_T} \left[ \sum_{i \leq T^{1/4}} |f(\mathcal{X}_i)| \right].$$

In order for  $\tau_L \geq T^{1/4}$ , either  $\tau_{\text{sp}} \geq K \log T = O(T^{-\gamma})$  (for  $K$  large enough), which by Proposition 5.1 has probability  $e^{-c\beta K \log T}$ , or there is no stretch of  $L$  consecutive  $X_\emptyset$  increments in the first  $T^{1/4} - K \log T$  spine increments; as argued in the proof of Lemma 7.5, for large enough  $\beta$  (depending on  $\gamma, \bar{c}$ ) this latter probability is at most  $\exp[-T^{3/16}]$ . The second term above is at most  $MT^{1/4}$  by the bound on  $f$ .

Let us now turn to  $|\mathbb{E}_{\pi_T}[\sum_{i=\tau_L+1, \dots, T} f(\mathcal{X}_i)]| \leq \sup_{\tau_L} \mathbb{E}_{\pi_T}[\sum_{i=\tau_L+1, \dots, T} f(\mathcal{X}_i) \mid \tau_L]$ . For each instantiation of  $\tau_L$ , we can expand,

$$\mathbb{E}_{\pi_T} \left[ \sum_{i=\tau_L+1, \dots, T} f(\mathcal{X}_i) \mid \tau_L \right] \leq MT\mathbb{E}_{\pi_T} \left[ \left| \frac{\pi_T(\mathcal{I})}{\pi_T(\Phi_\varphi(\mathcal{I}))} - 1 \right| \mid \tau_L \right].$$

By Claim 8.3 the quantity in the expectation is  $O(T^{-\gamma+1})$  for every tame interface, and therefore also in expectation under  $\mathbb{E}_{\pi_T}[\cdot \mid \tau_L]$  for every  $\tau_L$ . All in all, we have

$$\mathbb{E}_{\pi_T} \left[ \sum_{i \leq T} f(\mathcal{X}_i) \right] = O(T^{1/4}) + O(T^{-\gamma+1}) + O(T^{-\gamma+2}),$$

which is  $o(T)$  as long as  $\gamma > 1$  implying item (1).

Let us now turn to the proof of item (2). The proof is analogous and we therefore do not include all details. Suppose by way of contradiction that  $\mathbb{E}_\nu[\sum_i f(\mathcal{X}_0)g(\mathcal{X}_i)] \neq 0$ . We claim that it suffices, in order to obtain a contradiction, to show that for sufficiently large  $K$ , the following is  $o(T)$ :

$$\begin{aligned} \mathbb{E}_{\pi_T} \left[ \sum_{i \leq T} \sum_{j \leq T} f(\mathcal{X}_i)g(\mathcal{X}_j) \right] &= \mathbb{E}_{\pi_T} \left[ \sum_{i, j \in [1, T^{1/4}] \cup [T - T^{1/4}, T]} |f(\mathcal{X}_i)||g(\mathcal{X}_j)| \right] \\ &+ \mathbb{E}_{\pi_T} \left[ \sum_{i, j \in [T^{1/4}, T - T^{1/4}]: d(i, j) \geq K \log T} f(\mathcal{X}_i)g(\mathcal{X}_j) \right] \\ &+ \sum_{i \in [T^{1/4}, T - T^{1/4}]} \mathbb{E}_{\pi_T} \left[ \sum_{j: d(i, j) \leq K \log T} f(\mathcal{X}_i)g(\mathcal{X}_j) \right]. \end{aligned} \quad (8.1)$$

To see that this is sufficient, notice that the first term of (8.1) is at most  $16M^2K^2 \log^2 T$ . Arguing as in item (1) above, by Claim 8.3,  $\mathbb{E}_{\pi_T}[f(\mathcal{X}_i)] = O(T^{-\gamma})$  if  $i \geq T^{1/4}$ . Using Proposition 7.1, for each  $i, j$  at least  $T^{1/4}$  away from 1 and  $T$ , with  $d(i, j) \geq K \log T$ , if  $K$  is sufficiently large we have

$$\mathbb{E}_{\pi_T} [f(\mathcal{X}_i)g(\mathcal{X}_j)] \leq \mathbb{E}_{\pi_T}[f(\mathcal{X}_i)]\mathbb{E}_{\pi_T}[g(\mathcal{X}_j)] + M^2 \|\pi_T(\mathcal{X}_i \in \cdot, \mathcal{X}_j \in \cdot) - \pi_T(\mathcal{X}_i \in \cdot)\pi_T(\mathcal{X}_j \in \cdot)\|_{\text{TV}},$$

which is at most  $MT^{-\gamma} + M^2T^{-\gamma}$  for  $\gamma > 2$ , so that the second term in (8.1) is  $o(1)$ . Finally, by Proposition 7.2, specifically Corollary 7.3, together with Proposition 7.1, we deduce that

$$\mathbb{E}_{\pi_T} \left[ \sum_{i \leq T} \sum_{j \leq T} f(\mathcal{X}_i)g(\mathcal{X}_j) \right] = O(\log^2 T) + O(T^{-\gamma+2}) + T(1 - o(1))\mathbb{E}_\nu \left[ \sum_{j \in \mathbb{Z}} f(\mathcal{X}_0)g(\mathcal{X}_j) \right].$$

Thus, if we showed that the left-hand side of (8.1) is  $o(T)$ , we would deduce that  $\mathbb{E}_\nu[\sum_{j \in \mathbb{Z}} f(\mathcal{X}_0)g(\mathcal{X}_j)] = 0$ . Proceeding as in item (1), it suffices to show that the following is  $o(T)$  as  $T \rightarrow \infty$ :

$$\mathbb{E}_{\pi_T} \left[ \sum_{i,j \geq T^{1/4}} f(\mathcal{X}_i)g(\mathcal{X}_j) \right] = M^2 T^2 \pi_T(\tau_L \geq T^{1/4}) + \sup_{\tau_L \leq T^{1/4}} \mathbb{E}_{\pi_T} \left[ \sum_{\tau_L+1 \leq i,j \leq T} f(\mathcal{X}_i)g(\mathcal{X}_j) \mid \tau_L \right].$$

As before, the first term is  $O(\exp[-T^{3/16}])$  for large enough  $\beta$  (depending on  $\gamma, \bar{c}$ ). The second term is bounded by

$$M^2 T^2 \sup_{\tau_L \leq T^{1/4}} \mathbb{E}_{\pi_T} \left[ \left| \frac{\pi_T(\mathcal{I})}{\pi_T(\Phi_\varphi(\mathcal{I}))} - 1 \right| \mid \tau_L \right] \leq O(M^2 T^2 T^{-\gamma+1}),$$

for some  $\gamma > 2$  as long as  $\beta$  is large enough, which is in turn  $o(T)$  in  $T$ .  $\blacksquare$

**8.2. Linearity of variances.** In this section, we prove that the running sum of  $T$  increment observables  $f$ , for every  $f$  that is non-constant on the set of possible increments  $\mathfrak{X}$ , will have a variance of order  $T$ . This will in particular imply such a scaling for the variance of the total surface area of a pillar, the excess area of a pillar, and its  $xyz$ -displacements, conditional on having  $T$  increments.

**Proposition 8.4.** *There exists  $\beta_0$  such that for every  $\beta > \beta_0$  the following holds. If  $f : \mathfrak{X} \rightarrow \mathbb{R}$  is bounded,  $|f(X)| \leq M$  for all  $X \in \mathfrak{X}$ , and moreover, there exist distinct  $X_A, X_B \in \mathfrak{X}$  such that  $f(X_A) \neq f(X_B)$ , then*

$$\sum_{j \in \mathbb{Z}} \text{Cov}_\nu(f(\mathcal{X}_0), f(\mathcal{X}_j)) = \sigma_f^2 > 0.$$

*Proof.* It will suffice for us to prove that the following variance simply diverges as a function of  $t$ :

$$\text{Var}_\nu \left( \sum_{i \in [-t, t]} f(\mathcal{X}_i) \right) \rightarrow \infty \quad \text{as } t \rightarrow \infty.$$

Indeed, this will follow from the next well-known claim; we include its short proof for completeness.

**Claim 8.5.** *Let  $(Z_k)_{k \in \mathbb{Z}}$  be stationary with  $\sum_{k \in \mathbb{Z}} |\text{Cov}(Z_0, Z_k)| < \infty$ , and let  $V_t = \text{Var}(\sum_{k=-t}^t Z_k)$ . Then*

$$\exists \lim_{t \rightarrow \infty} \frac{V_t}{2t} =: \sigma^2 \geq 0 \quad \text{and} \quad \sigma^2 = \sum_{k \in \mathbb{Z}} \text{Cov}(Z_0, Z_k).$$

*Furthermore, if  $\sum_{k \in \mathbb{Z}} |k| |\text{Cov}(Z_0, Z_k)| < \infty$  then  $\sigma = 0$  iff  $\sup_t V_t < \infty$ .*

*Proof.* Let  $S_n = \sum_{k=1}^n Z_k$ , so  $\text{Var}(S_n) = \sum_{t=1}^n \mathfrak{C}_t$  for  $\mathfrak{C}_t = \text{Var}(Z_t) + 2 \sum_{k=1}^{t-1} \text{Cov}(Z_k, Z_t)$ . By the stationarity,  $\mathfrak{C}_t = \sum_{|k| \leq t-1} \text{Cov}(Z_0, Z_k)$ , and since  $\lim_{t \rightarrow \infty} \mathfrak{C}_t$  exists (by the absolute convergence hypothesis for this sum), Césaro's lemma shows this limit is equal to  $\lim_{n \rightarrow \infty} \frac{1}{n} \text{Var}(S_n) = \sigma^2 \geq 0$ . For the last statement, if  $\sigma = 0$  then  $\mathfrak{C}_t = -\sum_{|k| \geq t} \text{Cov}(Z_0, Z_k)$ , whence  $|\text{Var}(S_n)| = |\sum_{t=1}^n \mathfrak{C}_t| \leq \sum_{t=1}^n |\mathfrak{C}_t| \leq \sum_{k \in \mathbb{Z}} |k| |\text{Cov}(Z_0, Z_k)|$ .  $\blacksquare$

In fact, it suffices for us to show that the following diverges as  $T \rightarrow \infty$ :

$$\text{Var}_{\pi_T} \left( \sum_{i \leq T} f(\mathcal{X}_i) \right) = \sum_{i,j: d(i,j) > K \log T} \text{Cov}_{\pi_T}(f(\mathcal{X}_i), f(\mathcal{X}_j)) + \sum_{i,j: d(i,j) \leq K \log T} \text{Cov}_{\pi_T}(f(\mathcal{X}_i), f(\mathcal{X}_j)).$$

This is because by Proposition 7.1, the first sum on the right-hand side is  $O(T^{-\gamma+1})$  for  $\gamma > 1$  as long as  $\beta$  is large enough, and by Corollary 7.3, the second sum is  $(1 + o(1)) \text{Var}_\nu(\sum_{i \in [-\frac{T}{2}, \frac{T}{2}]} f(\mathcal{X}_i))$ .

The strategy to show this will be to find long stretches of trivial increments, which serve to decorrelate increments, and inject variance coming from either an  $X_A$  or  $X_B$  increment, into their centers. These injections will behave essentially independently, and therefore, will add some amount of variance proportional to the number of long stretches of trivial increments found. Fix  $L = \lceil \frac{5}{\epsilon} \log T \rceil$ . For any interface  $\mathcal{I} \in \bar{\mathbf{I}}_{x,T}$ , mark the first  $T^{3/4}$  indices  $j_k$  in increasing order, that satisfy  $j_k \geq \tau_{\text{SP}}$  and have  $j_k \geq \tau_{\text{SP}}$ ,  $\mathcal{X}_{j_k-L}, \dots, \mathcal{X}_{j_k-1}, \mathcal{X}_{j_k+1}, \dots, \mathcal{X}_{j_k+L} = X_\emptyset$  along with  $\mathcal{X}_{j_k} \in \{X_A, X_B\}$ . Let  $\mathcal{G} = \mathcal{G}(L)$  be the  $\sigma$ -algebra generated by the truncated interface  $\mathcal{I}_{\text{TR}}$ , the sequence  $(j_k)_{k \leq T^{3/4}}$  and all increments  $(\mathcal{X}_i)_{i \notin \{j_k\}}$ . By the law of total variance, we can express

$$\text{Var}_{\pi_T} \left( \sum_{i \in [1, T]} f(\mathcal{X}_i) \right) \geq \mathbb{E}_{\pi_T} \left[ \text{Var}_{\pi_T} \left( \sum_{i \in [1, T]} f(\mathcal{X}_i) \mid \mathcal{G} \right) \right].$$

However, conditionally on  $\mathcal{G}$ , the only contributions to the variance come from the increments  $\mathcal{X}_{j_k}$ , so that this quantity is the same as

$$\mathbb{E}_{\pi_T} \left[ \text{Var}_{\pi_T} \left( \sum_{k \leq T^{3/4}} f(\mathcal{X}_{j_k}) \mid (\mathcal{X}_{j_k+\ell})_{\ell \in [-L, L] \setminus \{0\}} = X_{\emptyset}, (\mathcal{X}_{j_k}) \in \{X_A, X_B\}, (\mathcal{X}_j)_{d(j, \cup_k \{j_k\}) > L}, \mathcal{I}_{\text{TR}} \right) \right].$$

Now fix any set of indices  $(j_k)_{k \leq T^{3/4}}$  which identify the trivial increments surrounding them, as well as the fact that  $\mathcal{X}_{j_k}$  is either  $X_A$  or  $X_B$ , and also fix all the other increments  $(\mathcal{X}_j)_{d(j, \cup_k \{j_k\}) > L}$ . We will show that for most such choices, the sum  $\sum_k f(\mathcal{X}_{j_k})$  has a variance that diverges in  $T$ .

Let us define a good set  $\Gamma_L$  in  $\mathcal{G}$  on which we can prove the variance above is at least  $T^{3/4}$ , say, as follows: an element of  $\mathcal{G}$ , given by  $\mathcal{I}_{\text{TR}}, \{j_k\}_k, (\mathcal{X}_j)_{d(j, \cup_k \{j_k\}) > L}$  is in  $\Gamma_L$  if there are indeed  $T^{3/4}$  many  $\{j_k\}$  and for every assignment of  $X_A, X_B$  to  $\{j_k\}_k$ , the resulting interface is tame. We will prove that  $\pi_T(\Gamma_L) \geq 1 - o(1)$ , and then that for any element of  $\Gamma_L$ , the variance of  $\sum_{k \leq T^{3/4}} f(\mathcal{X}_{j_k})$  goes to infinity with  $T$ .

**Claim 8.6.** *For every  $i \leq T$ , every  $\mathcal{I}_{\text{TR}}$  with  $\tau_{\text{SP}} \leq i$ , and every sequence of increments  $(X_j)_{j < i}$ , for any fixed increment  $X_{\star} \in \mathcal{X}$ , we have*

$$\mu_n(\mathcal{X}_i = X_{\star} \mid \mathcal{I}_{\text{TR}}, (X_j)_{j < i}, \bar{\mathbf{I}}_{x, T}) \geq \exp \left[ -(\beta + C)\mathfrak{m}(X_{\star}) \right].$$

The same estimate holds if we condition, e.g., on the increments above  $X_i$  as long as the first  $L$  are trivial: in particular, for every  $k \leq T^{3/4}$  and  $X_{\star} \in \{X_A, X_B\}$ ,

$$\inf_{G \in \Gamma_L} \pi_T(\mathcal{X}_{j_k} = X_{\star} \mid G) \geq \exp \left[ -(\beta + C)\mathfrak{m}(X_{\star}) \right].$$

**Proof of Claim 8.6.** In the interest of brevity we do not include a full proof. The first bound can be shown via a similar (simplified) version to the proof of Proposition 4.1, with the following modifications. Define a map  $\Psi_{\star} : \bar{\mathbf{I}}_{x, T} \rightarrow \bar{\mathbf{I}}_{x, T}$  which replaces the  $i$ -th increment of a pillar by  $X_{\star}$ ; one can readily see that for every  $\mathcal{I}$  with  $T$ -admissible truncation  $\mathcal{I}_{\text{TR}}$  with  $\tau_{\text{SP}} < i$  and increment sequence  $(X_i)_{i \leq T}$ , we have

$$\left| \log \frac{\pi_T(\mathcal{I})}{\pi_T(\mathcal{I}_{\star})} + \beta \mathfrak{m}(X_i; X_{\star}) \right| \leq \bar{K} (|\mathcal{F}(X_i)| + |\mathcal{F}(X_{\star})|) + \sum_{j > i} \bar{C} \bar{K} |\mathcal{F}(X_j)| e^{-\bar{c}(j-i)}.$$

We can bound the latter term on the right-hand side above by Corollary 4.2, and we can bound the multiplicity of the map for interfaces with  $\mathfrak{m}(\mathcal{I}, \Phi_{\star} \mathcal{I}) = k$  by  $s^k$  via Observation 2.27. Together these would imply the desired estimate, as the bound of Corollary 4.2 holds uniformly over all increment sequences below the  $i$ -th one, and the map  $\Psi_{\star}$  leaves those increments fixed.

The second part (where we may condition also on the increments sequence above  $i$  whilst in  $\Gamma_L$ ) is similar: for an interface in  $\Gamma_L$ , as the first  $L$  increments above  $\mathcal{X}_{j_k}$  are trivial and the increment sequence is tame,

$$\sum_{j > i} |\mathcal{F}(X_j)| e^{-\bar{c}(j-i)} \leq 4\bar{C} + \bar{C}T e^{-\bar{c}L},$$

and this is  $O(1)$  by our choice of  $L$ . Therefore, applying the map  $\Psi_{\star}$  for the  $j_k$ -th increment, we see that the probabilities of having  $X_A, X_B$  at marked indices  $\{j_k\}$  are comparable.  $\blacksquare$

First, by Proposition 5.1, with high probability the truncated interface  $\mathcal{I}_{\text{TR}}$  is such that  $\tau_{\text{SP}} \leq T^{1/4}$ , so let us work only with truncated interfaces that satisfy that bound. By Proposition 4.1 and Claim 8.6, for any stretch of  $2L + 1$  increments, the probability of the first and last  $L$  being trivial increments, and the middle element being in  $\{X_A, X_B\}$  is at least  $e^{-(\beta+C)\mathfrak{m}(X_A)}(1 - \varepsilon_{\beta})^{2L}$  for some  $\varepsilon_{\beta}$  going to zero as  $\beta \rightarrow \infty$ . There are  $\frac{T - T^{1/4}}{2L + 1}$  stretches for which this lower bound holds independently of the others; as before, a simple calculation yields that for  $\beta$  sufficiently large, the probability of having  $T^{3/4}$  such increment stretches of  $L$  trivial increments, an element of  $\{X_A, X_B\}$  then another  $L$  trivial increments, is  $1 - o(1)$ . Now, let us lower bound the following quantity by something diverging as  $T \rightarrow \infty$ :

$$\inf_{G \in \Gamma_L} \text{Var}_{\pi_T} \left( \sum_{k \leq T^{3/4}} f(\mathcal{X}_{j_k}) \mid G \right) = \inf_{G \in \Gamma_L} \sum_{k, k' \leq T^{3/4}} \text{Cov}_{\pi_T} \left( f(\mathcal{X}_{j_k}), f(\mathcal{X}_{j_{k'}}) \mid G \right).$$

By the second item in Claim 8.6, for every  $G \in \Gamma_L$ , the contribution of the diagonals, satisfies

$$\sum_{k \leq T^{3/4}} \text{Var}_{\pi_T}(f(\mathcal{X}_{j_k}) \mid G) \geq c_{f, \beta} T^{3/4},$$

for some  $c_{f,\beta} > 0$ , as  $f$  takes on different values on  $X_A$  and  $X_B$ , and both have strictly positive probability under  $\pi_T(\cdot | G)$ . On the other hand, we claim that the contribution from any off-diagonal term,  $\text{Cov}_{\pi_T}(f(\mathcal{X}_{j_k}), f(\mathcal{X}_{j_{k'}}) | G)$  is at most  $\bar{C}T e^{-\bar{c}L/2}$  for every pair  $k, k'$ . This can be shown via a straightforward modification of the map  $\Phi_{\text{MIX}}$  of Proposition 7.1; namely, the map would use  $j_k + \frac{L}{2}$  as the index above which it swaps the increment sequences. Following this through would imply that

$$\|\pi_T(\mathcal{X}_{j_k} \in \cdot, \mathcal{X}_{j_{k'}} \in \cdot) - \pi_T(\mathcal{X}_{j_k} \in \cdot)\pi_T(\mathcal{X}_{j_{k'}} \in \cdot)\|_{\text{TV}} \leq \bar{C}T e^{-\bar{c}L/2},$$

thus  $\text{Cov}_{\pi_T}(f(\mathcal{X}_{j_k}), f(\mathcal{X}_{j_{k'}})) \leq \bar{C}M^2T e^{-\bar{c}L/2}$  which is  $o(T^{-2})$  by our choice of  $L$ . Therefore, we see that

$$\text{Var}_{\pi_T}\left(\sum_{i \in \llbracket 1, T \rrbracket} f(\mathcal{X}_i)\right) \geq \inf_{G \in \Gamma_L} \text{Var}_{\pi_T}\left(\sum_{k \leq T^{3/4}} f(\mathcal{X}_{j_k}) | G\right) + o(1) \geq c_{f,\beta} T^{3/4} + o(1),$$

which diverges as  $T \rightarrow \infty$ , yielding the desired.  $\blacksquare$

## 9. CENTRAL LIMIT THEOREM FOR OBSERVABLES OF THE INCREMENT SEQUENCE

In the section we prove the following proposition, which, combined with Remark 9.3 yields the CLT from Theorem 2 as well as the expressions for the mean and variance of the limiting distribution in terms of the measure  $\nu$  on bi-infinite sequences of increments that was derived in §7, Corollary 7.3.

**Proposition 9.1.** *There exist  $\beta_0, \kappa_0 > 0$  such that the following holds. For every  $\beta > \beta_0$ , every non-constant function of the increments  $f : \mathfrak{X} \rightarrow \mathbb{R}$  such that*

$$f(X) \leq \exp(\kappa_0 |\mathcal{F}(X)|) \quad \text{for all } X \in \mathfrak{X}, \quad (9.1)$$

every sequence  $1 \ll T_n \ll n$ , and every  $x_n \in \llbracket -n + \Delta_n, n - \Delta_n \rrbracket^2 \times \{0\}$  for  $\Delta_n \gg T_n$ , the increment sequence  $\{\mathcal{X}_i\}$  of  $\mathcal{P}_{x_n}$  under  $\pi_T$ , the Ising measure conditioned on  $\bar{\mathbf{I}}_{x_n, T_n}$ , satisfies

$$\frac{1}{\sqrt{T_n}} \sum_{t=1}^{T_n} (f(\mathcal{X}_t) - \lambda) \implies \mathcal{N}(0, \sigma_f^2) \quad \text{as } n \rightarrow \infty,$$

where

$$\lambda = \mathbb{E}_\nu[f(\mathcal{X}_0)], \quad \sigma_f^2 = \sum_{j=-\infty}^{\infty} \text{Cov}_\nu(f(\mathcal{X}_0), f(\mathcal{X}_j)) > 0$$

for the measure  $\nu$  on bi-infinite sequences of increments  $(\mathcal{X}_i)_{-\infty}^{\infty}$  given by Corollary 7.3.

Modulo this result, the CLT readily extends to  $f : \mathfrak{X} \rightarrow \mathbb{R}^d$  for any  $d$ :

**Corollary 9.2.** *In the setting of Proposition 9.1, if  $f : \mathfrak{X} \rightarrow \mathbb{R}^d$  for some fixed  $d \geq 2$ , where  $f = (f_1, \dots, f_d)$  is such that each  $f_i$  is non-constant and satisfies (9.1), then under  $\pi_T$*

$$\frac{1}{\sqrt{T_n}} \sum_{t=1}^{T_n} (f(\mathcal{X}_t) - (\lambda_1, \dots, \lambda_d)) \implies \mathcal{N}(0, \Sigma),$$

where  $\lambda_i = \mathbb{E}_\nu[f_i(\mathcal{X}_0)]$  and  $\Sigma_{ij} = \Sigma_{ji} = \sum_{k=-\infty}^{\infty} \text{Cov}_\nu(f_i(\mathcal{X}_0), f_j(\mathcal{X}_k))$  for  $1 \leq i, j \leq d$ .

**Proof of Corollary 9.2.** First note that the fact that the matrix  $\Sigma$  is symmetric follows from the stationarity of the sequence  $(\mathcal{X}_i)$  under  $\nu$ . From the expression for  $\sigma^2$  given by Proposition 9.1, we see that for every linear combination  $f = \sum_i a_i f_i$  for  $a \in \mathbb{R}^d$  of functions centered w.r.t.  $\mathbb{E}_\nu$  and satisfying (9.1), one has  $\sum_{i=1}^T f(\mathcal{X}_i)/\sqrt{T} \implies \mathcal{N}(0, a^T \Sigma a)$  as  $n \rightarrow \infty$ . The proof is concluded via the Cramér–Wold device.  $\blacksquare$

**Remark 9.3.** Both Proposition 9.1 and Corollary 9.2 hold identically under the measure  $\mu_n(\cdot | \mathbf{I}_{x,T})$  (as opposed to  $\pi_T = \mu_n(\cdot | \bar{\mathbf{I}}_{x,T})$ ). To see this, let  $S_n = \frac{1}{\sqrt{T_n}} \sum_{t=1}^{T_n} (f(\mathcal{X}_t) - \lambda)$ ; for every Borel set  $B \subset \mathbb{R}$ ,

$$\mu_n(S_n \in B | \mathbf{I}_{x,T}) \leq \frac{\mu_n(S_n \in B, \bar{\mathbf{I}}_{x,T})}{\mu_n(\bar{\mathbf{I}}_{x,T})} + \mu_n(\bar{\mathbf{I}}_{x,T}^c | \mathbf{I}_{x,T}) = \pi_T(S_n \in B) \mu_n(\bar{\mathbf{I}}_{x,T} | \mathbf{I}_{x,T}) + \mu_n(\bar{\mathbf{I}}_{x,T}^c | \mathbf{I}_{x,T}),$$

which is at most  $\pi_T(S_n \in B) + O(\exp(-(\beta - C)r_0 T))$  by Lemma 3.15. Hence,  $\limsup_{n \rightarrow \infty} \mu_n(S_n \in B | \mathbf{I}_{x,T})$  is at most  $\lim_{n \rightarrow \infty} \pi_T(S_n \in B) = \mathbb{P}(\mathcal{N}(0, \sigma_f^2) \in B)$ , so (by Portmanteau)  $S_n \implies \mathcal{N}(0, \sigma_f^2)$  under  $\mu_n(\cdot | \mathbf{I}_{x,T})$ .

**9.1. Strategy of proof of the CLT.** We prove Proposition 9.1 by adapting a useful Stein's method type argument by Bolthausen [4] for treating stationary, mixing sequences of random variables. Our setting has several complications compared to [4]:

- (1) Our sequence of random variables, rather than being infinite and stationary, is a triangular array, where the individual laws  $\pi_T$  change due to the conditioning on  $\{\text{ht}(\mathcal{P}_x) \geq T\}$ .
- (2) Our  $\alpha$ -mixing estimates are invalid for base increments, and instead hold (see Proposition 7.1) only beyond a prefix of  $K \log T$  increments.
- (3) The increments are not stationary, and only become asymptotically stationary (see Proposition 7.2) away from the base and from the tip.

The asymptotic stationarity obstacle was handled by slight modifications of Bolthausen's argument in [31]; our proof follows a similar route, yet becomes somewhat simpler thanks to the nature of our  $\alpha$ -mixing estimates and control over higher moments of functions of the increment sequence.

**9.2. Proof of Proposition 9.1.** The first step in establishing the CLT is a standard truncation argument, using our control on  $\alpha$ -mixing and on moments of the increment sequence. Take  $b > 0$  to be a large enough constant, in particular larger than the constant  $K$  as given by Proposition 7.1 w.r.t.  $\gamma = 20$ . Our first step is to truncate the prefix and suffix of the increment sequence, as well as individual increment contributions. In what follows, recall  $\lambda = \mathbb{E}_\nu[f(\mathcal{X}_0)]$ , and let  $\sigma_f^2 = \sum_{j=-\infty}^{\infty} \text{Cov}_\nu(f(\mathcal{X}_0), f(\mathcal{X}_j))$  for any function  $f : \mathfrak{X} \rightarrow \mathbb{R}$ .

**Claim 9.4.** *In the setting of Proposition 9.1, let*

$$\ell = \lceil T^{1/5} \rceil \quad \text{and} \quad f_M(X) := f(X) \mathbf{1}_{\{|f(X)| \leq M\}}.$$

*If  $\frac{1}{\sqrt{T}} \sum_{j=\ell}^{T-\ell} (f_M(\mathcal{X}_j) - \mathbb{E}_{\pi_T}[f_M(\mathcal{X}_j)]) \implies \mathcal{N}(0, \sigma_{f_M}^2)$  for every  $M$  then  $\frac{1}{\sqrt{T}} \sum_{j=1}^T (f(\mathcal{X}_j) - \lambda) \implies \mathcal{N}(0, \sigma_f^2)$ .*

*Proof.* Let us first look at the effect of omitting the  $\ell_0$ -prefix and  $\ell_0$ -suffix of the summation over

$$Y_j = f(\mathcal{X}_j) - \mathbb{E}_{\pi_T} f(\mathcal{X}_j),$$

where

$$\ell_0 := \lceil b \log T \rceil$$

for some large  $b > 0$ , taken to be at least  $K$  from Proposition 7.1 for a choice of  $\gamma = 20$ . Following this step, we will be able to truncate the  $Y_j$ 's, and thereafter omit the  $\ell$ -prefix and  $\ell$ -suffix of the sum.

Proposition 5.1 (specifically, the exponential tail in (5.3)) implies that, for a sufficiently small  $\kappa$ , we have

$$\mathbb{E}_{\pi_T} \left[ e^{\kappa \sum_{i < \tau_{\text{sp}}} |\mathcal{F}(\mathcal{X}_i)|} \right] \leq T^{1/3}.$$

For the spine increments, Proposition 4.1 (together with Lemma 3.15 on the tameness of the spine) shows that, conditioned on  $\{\mathcal{X}_{\tau_{\text{sp}}+j} : j < i\}$ , the variable  $\mathbf{m}(\mathcal{X}_{\tau_{\text{sp}}+i})$  is dominated by an exponential variable with parameter  $c_0\beta$  (for  $c_0 > 0$  from that proposition). In particular,  $\sum_{i=\tau_{\text{sp}}}^{\ell_0} \mathbf{m}(\mathcal{X}_i) + \sum_{i>T-\ell_0} \mathbf{m}(\mathcal{X}_i)$  is stochastically dominated by a gamma-distributed random variable with parameters  $(2\ell_0, c_0\beta)$ , which again satisfies

$$\mathbb{E}_{\pi_T} \left[ e^{\kappa \left( \sum_{i=\tau_{\text{sp}}}^{\ell_0} |\mathcal{F}(\mathcal{X}_i)| + \sum_{i>T-\ell_0} |\mathcal{F}(\mathcal{X}_i)| \right)} \right] \leq T^{1/3}$$

(e.g., take  $\kappa = (1 - e^{-1/(8b)})(c_0\beta)^{-1}$ ). Overall, the hypothesis (9.1) implies, for a small enough  $\kappa_0$ , that

$$\frac{1}{\sqrt{T}} \sum_{j \leq \ell_0} \left( |f(\mathcal{X}_j)| + |f(\mathcal{X}_{T+1-j})| \right) \xrightarrow{L^1} 0 \quad \text{under } \pi_T, \quad (9.2)$$

so  $\frac{1}{\sqrt{T}} \sum_{j \leq \ell_0} (Y_j + Y_{T+1-j}) \rightarrow 0$  in probability, and hence does not affect the limiting law of  $\frac{1}{\sqrt{T}} \sum_j Y_j$ .

Again recalling Proposition 5.1, each variable  $Y_j$  is a function of a spine increment except with probability  $\exp(-c \log T) = O(T^{-5})$  for a large enough choice of  $b$ . Consequently, as per the exponential tail on spine increments established by Proposition 4.1 and the hypothesis  $|f(X)| \leq \exp[\kappa_0 |\mathcal{F}(X)|]$  for all  $X \in \mathfrak{X}$ ,

$$\pi_T(|Y_j| \geq a) \leq a^{-c_0\beta\kappa_0} + O(T^{-5}). \quad (9.3)$$

Moreover, on the event that the index  $j$  is not a spine index,  $\mathbf{m}(\mathcal{X}_j)$  has an exponential tail beyond  $K \log T$  by (5.3). Combining these two implies that  $Y_j$  has uniformly bounded moments of  $k$ -th order for small enough  $\kappa_0(k)$ . Namely, on the event  $j > \tau_{\text{sp}}$ , Proposition 4.1 implies that its  $k$ -th moment is finite as long as

$\kappa < \kappa_0(k)$ ; the event  $j \leq \tau_{\text{SP}}$ , has probability  $O(T^{-5})$ , and in that case, we can bound  $\mathbb{E}[|Y_j|^k] \leq O(e^{\kappa K \log T})$ , so that an application of Cauchy–Schwarz implies that for each  $k$ , there exists  $\kappa_0(k)$  such that for  $\kappa < \kappa_0$ ,

$$\max_j \mathbb{E}_{\pi_T}[|Y_j|^k] < C(\kappa, \beta, k).$$

For random variables  $Z_1, Z_2$ , let  $\sigma(Z_i)$  be the  $\sigma$ -algebra generated by  $Z_i$  and define the  $\alpha$ -mixing coefficient

$$\alpha(Z_1, Z_2) := \max_{A_1 \in \sigma(Z_1), A_2 \in \sigma(Z_2)} |\mathbb{P}(Z_1 \in A_1, Z_2 \in A_2) - \mathbb{P}(Z_1 \in A_1)\mathbb{P}(Z_2 \in A_2)|.$$

Write  $Y_j = Y'_j + Y''_j$  where  $Y''_j = Y_j \mathbf{1}_{\{|Y_j| > M\}}$ ; noting that  $\mathbb{E}_{\pi_T}(Y''_j)^3 \leq (\mathbb{E}_{\pi_T} Y_j^4)^{3/4} \mathbb{P}(|Y_j| > M)^{1/4}$  by Hölder's inequality, and that every two random variables  $Z_1, Z_2$  satisfy

$$\text{Cov}(Z_1, Z_2) \leq 2 (\alpha(Z_1, Z_2) \mathbb{E}[Z_1^3] \mathbb{E}[Z_2^3])^{1/3} \quad (9.4)$$

(see, e.g., [45, §1] for this inequality, originally by Davydov [21] with a larger constant pre-factor), one has

$$\begin{aligned} \mathbb{E}_{\pi_T} \left( \frac{1}{\sqrt{T}} \sum_{j=\ell_0}^{T-\ell_0} (Y''_j - \mathbb{E}_{\pi_T} Y''_j) \right)^2 &\leq \frac{C}{T} \max_{j \geq \ell_0} \mathbb{E}_{\pi_T} [(Y''_j)^3]^{\frac{2}{3}} \sum_{j, k \geq \ell_0} \alpha(\mathcal{X}_j, \mathcal{X}_k)^{1/3} \\ &\leq C' \max_{j \geq \ell_0} \sqrt{\mathbb{P}(|Y_j| > M)}, \end{aligned}$$

using  $\alpha(\mathcal{X}_j, \mathcal{X}_k) \leq C|k-j|^{-\gamma}$  for  $k > j \geq \ell_0$  and  $\gamma = 20$  by the above application of Proposition 7.1. As the expression on the right can be made arbitrarily small as a function of  $M$ , uniformly over  $n$  (it is at most  $CM^{-c_0 \kappa_0/2} + o(1)$  by (9.3)), we see that showing  $\frac{1}{\sqrt{T}} \sum_{j=\ell_0}^{T-\ell_0} (Y'_j - \mathbb{E}_{\pi_T} Y'_j) \Rightarrow \mathcal{N}(0, \sigma_{f_M}^2)$  as  $n \rightarrow \infty$  for every fixed  $M$ , as well as  $\sigma_{f_M}^2 \rightarrow \sigma_f^2$ , will imply that  $\frac{1}{\sqrt{T}} \sum_{j=1}^T (Y_j - \mathbb{E}_{\pi_T} Y_j) \Rightarrow \mathcal{N}(0, \sigma_f^2)$ .

To verify that  $\sigma_{f_M}^2 \rightarrow \sigma_f^2$  as  $M \rightarrow \infty$ , recall from Corollary 7.3 that  $\nu(\mathbf{m}(\mathcal{X}_0) \geq r) \leq \exp(-c_0 \beta r)$ , so for some  $C, c > 0$  we get

$$\mathbb{E}_{\nu} |f(\mathcal{X}_0) - f_M(\mathcal{X}_0)|^3 \leq \sqrt{\mathbb{E}_{\nu} [f(\mathcal{X}_0)^6] \nu(|f(\mathcal{X}_0)| \geq M)} \leq CM^{-c}$$

using Corollary 7.3 and (9.1) to uniformly bound  $\mathbb{E}_{\nu} [f(\mathcal{X}_0)^6]$  and the  $\nu$ -probability of  $|f(\mathcal{X}_0)| \geq M$ . Writing

$$\sigma_{f_M}^2 = \sigma_f^2 + \sum_{k=-\infty}^{\infty} \text{Cov}(f_M(\mathcal{X}_0) - f(\mathcal{X}_0), f(\mathcal{X}_k)) + \sum_{k=-\infty}^{\infty} \text{Cov}(f_M(\mathcal{X}_0), f_M(\mathcal{X}_k) - f(\mathcal{X}_k)),$$

we can infer from the fact  $\alpha(\mathcal{X}_0, \mathcal{X}_k) \leq Ck^{-\gamma}$  under  $\nu$ , and another application of (9.4), that

$$\sum_{k=-\infty}^{\infty} |\text{Cov}(f_M(\mathcal{X}_0) - f(\mathcal{X}_0), f(\mathcal{X}_k))| \leq C' M^{-c'} \sum_k k^{-\gamma} \leq C'' M^{-c'},$$

and the same holds for  $\sum \text{Cov}(f_M(\mathcal{X}_0), f_M(\mathcal{X}_k) - f(\mathcal{X}_k))$  in the same manner, implying  $\sigma_{f_M}^2 \rightarrow \sigma_f^2$ .

Thus far we established that it suffices to show  $\frac{1}{\sqrt{T}} \sum_{j=\ell_0}^{T-\ell_0} Y'_j \Rightarrow \mathcal{N}(0, \sigma_{f_M}^2)$  for every  $M > 0$ . Note that

$$\frac{1}{T} \mathbb{E}_{\pi_T} \left[ \left( \sum_{j=\ell_0}^{\ell} (Y'_j + Y'_{T+1-j}) \right)^2 \right] \leq \frac{(2\ell)^2}{T} \max_{\ell_0 \leq j, k \leq T-\ell_0} |\mathbb{E}_{\pi_T} [Y'_j Y'_k]| = O(T^{-3/5})$$

since  $|Y'_j| \leq M$  for all  $j$ . Hence, we may indeed replace  $\ell_0$  by  $\ell$  when considering  $\sum Y'_j$ , as the contributions to the limiting law by the  $\ell$ -prefix and  $\ell$ -suffix in  $\frac{1}{\sqrt{T}} \sum_{j=\ell_0}^{T-\ell_0} Y'_j$  are negligible.

Finally, we wish to replace centering term  $\mathbb{E}_{\pi_T} f(\mathcal{X}_j)$  by  $\lambda = \mathbb{E}_{\nu} f(\mathcal{X}_0)$  for each  $j = 1, \dots, T$ . Recall from (9.2) that  $\frac{1}{\sqrt{T}} \sum_{j \leq \ell_0} (\mathbb{E}_{\pi_T} |f(\mathcal{X}_j)| + \mathbb{E}_{\pi_T} |f(\mathcal{X}_{T+1-j})|) \rightarrow 0$ . For each  $j \geq \ell_0$ , we have  $\mathbb{E}_{\pi_T} |f(\mathcal{X}_j)| \leq C$ , as we had established above (following (9.3)). Therefore, we can neglect the  $\ell$ -prefix and  $\ell$ -suffix of the sequence of expectations, as  $\frac{1}{\sqrt{T}} \sum_{j \leq \ell} (\mathbb{E}_{\pi_T} |f(\mathcal{X}_j)| + \mathbb{E}_{\pi_T} |f(\mathcal{X}_{T+1-j})|) = O(\ell/\sqrt{T}) = o(1)$ . For each of the remaining indices  $\ell \leq j \leq T - \ell$ , by Proposition 7.2 (as used in the proof of Corollary 7.3), we have that  $\|\pi_T(\mathcal{X}_j \in \cdot) - \nu(\mathcal{X}_j \in \cdot)\|_{\text{TV}} \leq C\ell^{-\gamma}$ . Hence, looking at the truncated function  $f_{\ell}$ , we have  $|\mathbb{E}_{\pi_T} f_{\ell}(\mathcal{X}_j) - \mathbb{E}_{\nu} f_{\ell}(\mathcal{X}_j)| = O(\ell^{1-\gamma}) = O(T^{-2})$  (so the sum of these over all  $j$  is  $o(1)$ ), whereas  $\mathbb{E}_{\pi_T} |(f - f_{\ell})(\mathcal{X}_j)|$  and  $\mathbb{E}_{\nu} |(f - f_{\ell})(\mathcal{X}_j)|$  are each  $O(\ell^{-c\beta\kappa_0}) = o(1/T)$  by Cauchy–Schwarz, the  $O(1)$  bounds on the means of  $f(\mathcal{X}_j)$  under  $\pi_T$  and  $\nu$ , and the exponential tails of  $\mathbf{m}(\mathcal{X}_j)$  together with (9.1) and (5.3). ■

Through the remainder of the proof, let  $M > 0$  and  $Y_j = f_M(\mathcal{X}_j) - \mathbb{E}_{\pi_T} f_M(\mathcal{X}_j)$  (so that  $|Y_j| \leq M$ ) for each  $j = \ell, \dots, T - \ell$ , with the goal of showing that  $\frac{1}{\sqrt{T}} \sum_{j=\ell}^{T-\ell} Y_j \implies \mathcal{N}(0, \sigma_{f_M}^2)$ . We further assume w.l.o.g. that  $f_M$  is non-constant (as this holds for all  $M > M_0$  for some  $M_0 = M_0(f) > 0$ ), whence

$$\sigma_{f_M}^2 = \sum_{j=-\infty}^{\infty} \text{Cov}_{\nu}(f_M(\mathcal{X}_0), f_M(\mathcal{X}_j)) = \lim_{t \rightarrow \infty} \frac{1}{2t} \text{Var}_{\nu} \left( \sum_{j=-t}^t f_M(\mathcal{X}_j) \right) > 0,$$

by Proposition 8.4, applied to the bounded non-constant function  $f_M$ . Defining

$$\mathfrak{S}_T := \sum_{j=\ell}^{T-\ell} \mathbb{E}_{\pi_T} \left[ Y_j \sum_{k=\ell}^{T-\ell} \mathbf{1}_{\{|k-j| \leq \ell\}} Y_k \right],$$

recall from Propositions 7.1–7.2 and Corollary 7.3 that  $\sum_{|j| \leq \ell} \text{Cov}_{\nu}(f_M(\mathcal{X}_0), f_M(\mathcal{X}_j)) = \sigma_{f_M}^2 + o(1)$  and

$$\begin{aligned} \sum_{k=\ell}^{T-\ell} |\text{Cov}_{\pi_T}(f_M(\mathcal{X}_j), f_M(\mathcal{X}_k))| \mathbf{1}_{\{|k-j| > \ell\}} &= O(T\ell^{-\gamma}) = o(1), \\ \sum_{k=\ell}^{T-\ell} |\text{Cov}_{\pi_T}(f_M(\mathcal{X}_j), f_M(\mathcal{X}_k)) - \text{Cov}_{\nu}(f_M(\mathcal{X}_0), f_M(\mathcal{X}_k))| \mathbf{1}_{\{|k-j| \leq \ell\}} &= O(\ell^{1-\gamma}) = o(1) \end{aligned}$$

for all  $2\ell \leq j \leq T - 2\ell$  (whereas both terms are  $O(1)$  for  $\ell \leq j \leq 2\ell$  and  $T - 2\ell \leq j \leq T - \ell$ ), we get

$$\mathfrak{S}_T = (1 + o(1)) \text{Var}_{\pi_T} \left( \sum_{k=\ell}^{T-\ell} Y_j \right) \quad \text{as well as} \quad \mathfrak{S}_T = (1 + o(1)) \sigma_{f_M}^2 T.$$

The following simple argument of Bolthausen [4] gives a convenient approach for establishing CLTs for mixing random fields, even in the situation where (unlike the original setting of [4]) the sequence of increments is only asymptotically stationary. At the heart of the argument is the following observation:

**Lemma 9.5** ([4, Lemma 2]). *If  $(Z_n)$  is a sequence of real-valued random variables with  $\sup_n \mathbb{E} Z_n^2 < \infty$  and*

$$\lim_{n \rightarrow \infty} \mathbb{E} [(i\lambda - Z_n) e^{i\lambda Z_n}] = 0 \quad \text{for every } \lambda \in \mathbb{R}, \quad (9.5)$$

*then  $Z_n$  converges weakly to the standard Gaussian  $\mathcal{N}(0, 1)$ .*

(Indeed, tightness is implied by the uniform bound on the  $\mathbb{E} Z_n^2$ , and verifying that every subsequential limit point is standard Gaussian can be derived from (9.5), as a variable  $Z$  having the law of such a limit point has  $\mathbb{E}[f'(Z) - Zf(Z)] = 0$  for every  $f \in C^1(\mathbb{R})$ , hence must be standard Gaussian by Stein's characterization.) Define the random variables  $Z_T$  and for  $r \in [\ell, T - \ell]$ ,  $Z_{r,T}$  by

$$Z_T = \frac{1}{\sqrt{\mathfrak{S}_T}} \sum_{j=\ell}^{T-\ell} Y_j, \quad \text{and} \quad Z_{r,T} = \frac{1}{\sqrt{\mathfrak{S}_T}} \sum_{j=\ell}^{T-\ell} \mathbf{1}_{\{|j-r| \leq \ell\}} Y_j,$$

We aim to verify (9.5) for the random variables  $(Z_T)$  via the following useful decomposition of the term  $(i\lambda - Z_n) e^{i\lambda Z_n}$  given in [4]: for every  $\lambda \in \mathbb{R}$ ,

$$(i\lambda - Z_T) e^{i\lambda Z_T} = \Xi_1 - \Xi_2 - \Xi_3,$$

for

$$\Xi_1 = i\lambda e^{i\lambda Z_T} \left( 1 - \frac{1}{\sqrt{\mathfrak{S}_T}} \sum_{r=\ell}^{T-\ell} Y_r Z_{r,T} \right), \quad (9.6)$$

$$\Xi_2 = \frac{1}{\sqrt{\mathfrak{S}_T}} e^{i\lambda Z_T} \sum_{r=\ell}^{T-\ell} Y_r \left( 1 - e^{-i\lambda Z_{r,T}} - i\lambda Z_{r,T} \right), \quad (9.7)$$

$$\Xi_3 = \frac{1}{\sqrt{\mathfrak{S}_T}} \sum_{r=\ell}^{T-\ell} Y_r e^{i\lambda(Z_T - Z_{r,T})}, \quad (9.8)$$

(where the equality used only that  $Z_T = \frac{1}{\sqrt{\mathfrak{S}_T}} \sum_{j=\ell}^{T-\ell} Y_j$ , irrespective of the definitions of  $Z_{r,T}$  and  $\ell$ ). Thus, it will suffice to show that  $\mathbb{E}_{\pi_T}[\Xi_i] \rightarrow 0$  as  $n \rightarrow \infty$  for each  $i = 1, 2, 3$  in order to verify (9.5) for  $(Z_T)$ .

For the first of these terms, recall that by the definition of  $\mathfrak{S}_T$  and  $Z_{r,T}$  one has that

$$\mathbb{E}_{\pi_T} \left[ \frac{1}{\sqrt{\mathfrak{S}_T}} \sum_{r=\ell}^{T-\ell} Y_r Z_{r,T} \right] = \frac{1}{\mathfrak{S}_T} \sum_{r=\ell}^{T-\ell} \mathbb{E}_{\pi_T} \left[ Y_r \sum_{j=\ell}^{T-\ell} \mathbf{1}_{\{|j-r|\leq\ell\}} Y_j \right] = 1,$$

whence

$$\begin{aligned} \mathbb{E}_{\pi_T} |\Xi_1|^2 &= \lambda^2 \text{Var}_{\pi_T} \left( \frac{1}{\sqrt{\mathfrak{S}_T}} \sum_{r=\ell}^{T-\ell} Y_r Z_{r,T} \right) = \frac{\lambda^2}{\mathfrak{S}_T^2} \text{Var}_{\pi_T} \left( \sum_{r=\ell}^{T-\ell} Y_r \sum_{\substack{j \in [\ell, T-\ell] \\ |j-r| \leq \ell}} Y_j \right) \\ &\leq \frac{\lambda^2}{\mathfrak{S}_T^2} \sum_{r, r' \in [\ell, T-\ell]} \sum_{\substack{j, j' \in [\ell, T-\ell] \\ |j-r| \leq \ell, |j'-r'| \leq \ell}} |\text{Cov}_{\pi_T}(Y_j Y_r, Y_{j'} Y_{r'})|. \end{aligned}$$

Splitting the sum over  $r, r'$  according to  $|r - r'|$ , we see that if  $|r' - r| \geq 3\ell$  then each of the  $O(T^2\ell^2)$  summands satisfies

$$|\text{Cov}_{\pi_T}(Y_j Y_r, Y_{j'} Y_{r'})| \leq M^4 c \ell^{-\gamma} = O(T^{-3})$$

for some  $c > 0$ , by Proposition 7.1; on the other hand, there are  $O(T\ell^3)$  summands  $r, r', j, j'$  with  $|r - r'| \leq 3\ell$ , each of which is uniformly bounded by  $M^4$ . Altogether, recalling that  $\mathfrak{S}_T = (1 + o(1))\sigma_{f_M}^2 T$ , we deduce

$$\mathbb{E}_{\pi_T} |\Xi_1|^2 = O(\ell^3/T) = O(T^{-2/5}).$$

For the second term, observe that

$$|\Xi_2| \leq M \frac{1}{\sqrt{\mathfrak{S}_T}} \sum_{r=\ell}^{T-\ell} |1 - e^{-i\lambda Z_{r,T}} - i\lambda Z_{r,T}| \leq M\lambda^2 \frac{1}{\sqrt{\mathfrak{S}_T}} \sum_{r=\ell}^{T-\ell} |Z_{r,T}|^2,$$

where the last inequality was obtained by Taylor expanding  $1 - \cos(\lambda Z_{r,T})$  and  $\sin(\lambda Z_{r,T}) - \lambda Z_{r,T}$  (the real and imaginary parts of of each summand, respectively). Since  $|Z_{r,T}| \leq (2\ell + 1)M/\sqrt{\mathfrak{S}_T}$  by its definition (and the truncation bound on the  $Y_j$ 's), it follows that

$$|\Xi_2| = O\left(\frac{T\ell^2}{\mathfrak{S}_T^{3/2}}\right) = O\left(T^{-1/10}\right).$$

Finally, when treating  $\Xi_3$ , we can use Proposition 7.1 to decompose  $\mathbb{E}_{\pi_T} [Y_r e^{i\lambda(Z_T - Z_{r,T})}]$  as follows:

$$\begin{aligned} \left| \mathbb{E}_{\pi_T} \left[ Y_r e^{i\lambda(Z_T - Z_{r,T})} \right] \right| &= \left| \mathbb{E}_{\pi_T} \left[ Y_r \prod_{j=\ell}^{r-\ell} e^{i\lambda\mathfrak{S}_T^{-1/2} Y_j} \prod_{j=r+\ell}^{T-\ell} e^{i\lambda\mathfrak{S}_T^{-1/2} Y_j} \right] \right| \\ &\leq \left| \mathbb{E}_{\pi_T} \left[ Y_r \prod_{j=\ell}^{r-\ell} e^{i\lambda\mathfrak{S}_T^{-1/2} Y_j} \right] \mathbb{E}_{\pi_T} \left[ \prod_{j=r+\ell}^{T-\ell} e^{i\lambda\mathfrak{S}_T^{-1/2} Y_j} \right] \right| + M c \ell^{-\gamma}, \end{aligned}$$

(using that the variables in the two expectations in the last line are at most  $M$  and 1 in absolute value, respectively). Since  $|\mathbb{E}_{\pi_T} [\prod_{j=\ell}^{r-\ell} e^{i\lambda\mathfrak{S}_T^{-1/2} Y_j}]| \leq 1$ , we can apply Proposition 7.1 again to obtain that the last expression is, in turn, at most

$$\left| \mathbb{E}_{\pi_T} \left[ Y_r \prod_{j=\ell}^{r-\ell} e^{i\lambda\mathfrak{S}_T^{-1/2} Y_j} \right] \right| + M c \ell^{-\gamma} \leq \left( \left| \mathbb{E}_{\pi_T} [Y_r] \mathbb{E}_{\pi_T} \left[ \prod_{j=\ell}^{r-\ell} e^{i\lambda\mathfrak{S}_T^{-1/2} Y_j} \right] \right| + M c \ell^{-\gamma} \right) + M c \ell^{-\gamma} = 2M c \ell^{-\gamma},$$

using the fact that  $\mathbb{E}_{\pi_T} Y_r = 0$ . This concludes the proof.  $\blacksquare$

**9.3. Proof of Corollary 3.** We wish to apply Theorem 2 with specific choices of observables, that contain the information about the distribution of the tip and volume and surface area of the pillar. We begin with item (1), regarding the distribution of the tip,  $(Y_1, Y_2, Y_3)$ . Define observables  $f_i : \mathfrak{X} \rightarrow \mathbb{R}$  for  $i = 1, 2, 3$ , by

$$f_i(X) = (v_+(X) - v_-(X)) \cdot e_i,$$

where  $v_+(X)$  is the midpoint of the highest cell of a rooted increment  $X \in \mathfrak{X}$ , and  $v_-(X)$  is the midpoint of its lowest cell, i.e.,  $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ . Then, we can express, for  $(Y_1, Y_2, Y_3)$  the tip of  $\mathcal{P}_x$ ,

$$\left| \frac{Y_i - x_i}{\sqrt{T}} - \frac{1}{\sqrt{T}} \sum_{j=1}^T f_i(\mathcal{X}_j) \right| \leq \frac{1}{\sqrt{T}} (\text{diam}(\mathcal{B}_x) + \text{ht}(v_{\tau_{\text{sp}}}) + \frac{1}{2} + |\mathcal{F}(\mathcal{X}_{>T})|).$$

By Proposition 5.1, both  $\mathbb{E}_{\pi_T}[\text{diam}(\mathcal{B}_x)^2]$  and  $\mathbb{E}[\text{ht}(v_{\tau_{\text{sp}}})^2]$  are  $O(\log^2 T)$ , and by Lemma 3.12,  $\mathbb{E}[|\mathcal{F}(\mathcal{X}_{>T})|^2]$  is  $O(1)$ . Hence, the right-hand side goes to 0 in probability as  $T \rightarrow \infty$ , and a CLT for  $\frac{1}{\sqrt{T}} \sum_{j=1}^T f_i(\mathcal{X}_j)$  yields the same CLT for  $\frac{Y_i - x_i}{\sqrt{T}}$ . Since  $|f_i(X)| \leq \mathfrak{m}(X)$  for every  $X \in \mathfrak{X}$ , by Corollary 9.2,

$$\frac{(Y_1, Y_2, Y_3) - (x_1, x_2, 0) - (\lambda_1, \lambda_2, \lambda_3)T}{\sqrt{T}} \implies \mathcal{N}(0, \Sigma),$$

where  $\lambda_i = \mathbb{E}_\nu[f_i(\mathcal{X}_0)]$  and  $\Sigma_{i,j} = \sum_{k \in \mathbb{Z}} \text{Cov}(f_i(\mathcal{X}_0), f_j(\mathcal{X}_k))$ .

The observables  $f_1, f_2$  are anti-symmetric with respect to reflections about the plane with outward normal  $e_1, e_2$ , so by item (1) Proposition 8.1, they have  $\mathbb{E}_\nu[f_i(\mathcal{X}_0)] = \lambda_i = 0$  (though  $f_i$  are not bounded, this follows by truncating  $f_i$ , and using Corollary 7.3 to deduce that the truncated means converge to the true means). The heights  $f_3$  are at least one, and thus  $\lambda_3 \geq 1$ . Since  $f_i$  are invariant under reflection about the plane with outward normal  $e_j$  (for  $j \neq i, j \in \{1, 2\}$ ), by item (2) of Proposition 8.1, the off-diagonals of  $\Sigma$  are 0 (again truncating the observables and noticing that the truncated covariances converge to the true covariances). By Proposition 8.2 and the observation that  $f_i$  are non-constant on  $\mathfrak{X}$ , the diagonals of  $\Sigma$  are positive, say  $\sigma_{f_i}^2 > 0$ . It remains to verify that  $\sigma_{f_1}^2 = \sigma_{f_2}^2$ . Note that for every  $\tau_{\text{sp}} \leq j, k \leq T$  the observable  $g(\mathcal{X}_j, \mathcal{X}_k) = f_1(\mathcal{X}_j)f_1(\mathcal{X}_k) - f_2(\mathcal{X}_j)f_2(\mathcal{X}_k)$  is anti-symmetric in application of the map that rotates the increments above some  $\tau_L$  by  $\frac{\pi}{2}$ , as long as  $j > \tau_L$ . Analogously to Proposition 8.1, we would then see that  $\mathbb{E}_\nu[g(\mathcal{X}_j, \mathcal{X}_k)] = 0$  for every  $T^{\frac{1}{4}} \leq j < k \leq T$  as long as  $|k - j| \leq T^{\frac{1}{4}}$ . By linearity and the decay estimate of item (2) of Corollary 7.3, we see that  $\mathbb{E}_\nu[\sum_{j \in \mathbb{Z}} f_1(\mathcal{X}_0)f_1(\mathcal{X}_j)] = \mathbb{E}_\nu[\sum_{j \in \mathbb{Z}} f_2(\mathcal{X}_0)f_2(\mathcal{X}_j)]$ .

Item (2) follows in a similar fashion. Let  $f_V(X) = |\mathcal{C}(X)| - 1$  and  $f_A(X) = |\mathcal{F}(X)| - 4$ . We can bound

$$\left| |\mathcal{C}(\mathcal{P}_x)| - \sum_{i \leq T} f_V(\mathcal{X}_i) \right| \leq (\text{diam}(\mathcal{B}_x))^2 (\text{ht}(v_{\tau_{\text{sp}}}) + \frac{1}{2}) + |\mathcal{F}(\mathcal{X}_{>T})|^3, \quad \text{and}$$

$$\left| |\mathcal{F}(\mathcal{P}_x)| - \sum_{i \leq T} f_A(\mathcal{X}_i) \right| \leq 4(\text{diam}(\mathcal{B}_x))^2 (\text{ht}(v_{\tau_{\text{sp}}}) + \frac{1}{2}) + |\mathcal{F}(\mathcal{X}_{>T})|.$$

Thus, as before, by Proposition 5.1 and Lemma 3.12, a CLT for  $\frac{1}{\sqrt{T}} \sum_i f_V(\mathcal{X}_i)$  implies the same CLT for  $\frac{1}{\sqrt{T}} |\mathcal{C}(\mathcal{P}_x)|$  and a CLT for  $\frac{1}{\sqrt{T}} \sum_i f_A(\mathcal{X}_i)$  implies the same CLT for  $\frac{1}{\sqrt{T}} |\mathcal{F}(\mathcal{P}_x)|$ . Since both  $f_V(X)$  and  $f_A(X)$  are positive, are at most  $4 + \mathfrak{m}(X)^2$ , and are non-constant, they satisfy central limit theorems with positive means and variances, implying the same for the volume and surface area of  $\mathcal{P}_x$ .

**Acknowledgment.** We thank the referees for valuable suggestions. E.L. was supported in part by NSF grant DMS-1812095.

## REFERENCES

- [1] D. B. Abraham. Solvable model with a roughening transition for a planar ising ferromagnet. *Phys. Rev. Lett.*, 44:1165–1168, May 1980.
- [2] T. Bodineau. The Wulff construction in three and more dimensions. *Comm. Math. Phys.*, 207(1):197–229, 1999.
- [3] T. Bodineau. Slab percolation for the Ising model. *Probab. Theory Related Fields*, 132(1):83–118, 2005.
- [4] E. Bolthausen. On the central limit theorem for stationary mixing random fields. *Ann. Probab.*, 10(4):1047–1050, 1982.
- [5] E. Bolthausen, J.-D. Deuschel, and G. Giacomin. Entropic repulsion and the maximum of the two-dimensional harmonic crystal. *Ann. Probab.*, 29(4):1670–1692, 2001.
- [6] M. Bramson, J. Ding, and O. Zeitouni. Convergence in law of the maximum of the two-dimensional discrete Gaussian free field. *Comm. Pure Appl. Math.*, 69(1):62–123, 2016.
- [7] M. Bramson and O. Zeitouni. Tightness of the recentered maximum of the two-dimensional discrete Gaussian free field. *Comm. Pure Appl. Math.*, 65(1):1–20, 2012.
- [8] R. Brandenberger and C. E. Wayne. Decay of correlations in surface models. *J. Statist. Phys.*, 27(3):425–440, 1982.
- [9] J. Bricmont, A. El Mellouki, and J. Fröhlich. Random surfaces in statistical mechanics: roughening, rounding, wetting, . . . *J. Statist. Phys.*, 42(5-6):743–798, 1986.
- [10] J. Bricmont, J.-R. Fontaine, and J. L. Lebowitz. Surface tension, percolation, and roughening. *J. Statist. Phys.*, 29(2):193–203, 1982.

- [11] J. Bricmont, J. L. Lebowitz, E. Olivieri, and C. E. Pfister. Nontranslation-invariant Gibbs states with coexisting phases. I. Existence of sharp interface for Widom-Rowlinson type lattice models in three dimensions. *Comm. Math. Phys.*, 66(1):1–20, 1979.
- [12] J. Bricmont, J. L. Lebowitz, and C. E. Pfister. Nontranslation-invariant Gibbs states with coexisting phases. II. Cluster properties and surface tension. *Comm. Math. Phys.*, 66(1):21–36, 1979.
- [13] M. Campanino and D. Ioffe. Ornstein-Zernike theory for the Bernoulli bond percolation on  $\mathbb{Z}^d$ . *Ann. Probab.*, 30(2):652–682, 2002.
- [14] M. Campanino, D. Ioffe, and Y. Velenik. Ornstein-Zernike theory for finite range Ising models above  $T_c$ . *Probab. Theory Related Fields*, 125(3):305–349, 2003.
- [15] M. Campanino, D. Ioffe, and Y. Velenik. Fluctuation theory of connectivities for subcritical random cluster models. *Ann. Probab.*, 36(4):1287–1321, 2008.
- [16] 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.
- [17] 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.
- [18] R. Cerf and R. Kenyon. The low-temperature expansion of the Wulff crystal in the 3D Ising model. *Comm. Math. Phys.*, 222(1):147–179, 2001.
- [19] R. Cerf and A. Pisztor. On the Wulff crystal in the Ising model. *Ann. Probab.*, 28(3):947–1017, 2000.
- [20] N. Datta, A. Messenger, and B. Nachtergaele. Rigidity of interfaces in the Falicov-Kimball model. *J. Statist. Phys.*, 99(1-2):461–555, 2000.
- [21] Y. A. Davydov. The convergence of distributions generated by stationary stochastic processes. *Theory of Probability & Its Applications*, 13:691–696, 1968.
- [22] N. G. de Bruijn and P. Erdős. Some linear and some quadratic recursion formulas. II. *Indagationes Math.*, 14:152–163, 1952.
- [23] L. R. Dobrushin. An investigation of Gibbs states for three-dimensional lattice systems. *Teor. Veroyatnost. i Primenen.*, 18:261–279, 1973.
- [24] R. Dobrushin and O. Hryniv. Fluctuations of the phase boundary in the 2D Ising ferromagnet. *Comm. Math. Phys.*, 189(2):395–445, 1997.
- [25] R. Dobrushin, R. Kotecký, and S. Shlosman. *Wulff construction*, volume 104 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1992.
- [26] R. L. Dobrushin. Description of a random field by means of conditional probabilities and conditions for its regularity. *Theory of Probability & Its Applications*, 13(2):197–224, 1968.
- [27] R. L. Dobrushin. Asymptotical behavior of Gibbsian distributions for lattice systems and their dependence on the form of the container. *Teoret. Mat. Fiz.*, 12(1):115–134, 1972.
- [28] R. L. Dobrushin. The Gibbs state that describes the coexistence of phases for a three-dimensional Ising model. *Teor. Veroyatnost. i Primenen.*, 17:619–639, 1972.
- [29] S. Friedli and Y. Velenik. *Statistical mechanics of lattice systems*. Cambridge University Press, Cambridge, 2018. A concrete mathematical introduction.
- [30] J. Fröhlich and T. Spencer. The Kosterlitz-Thouless transition in two-dimensional abelian spin systems and the Coulomb gas. *Comm. Math. Phys.*, 81(4):527–602, 1981.
- [31] S. Ganguly, E. Lubetzky, and F. Martinelli. Cutoff for the east process. *Comm. Math. Phys.*, 335(3):1287–1322, 2015.
- [32] H.-O. Georgii. *Gibbs measures and phase transitions*, volume 9 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, second edition, 2011.
- [33] G. Gielis and G. Grimmett. Rigidity of the interface in percolation and random-cluster models. *J. Statist. Phys.*, 109(1-2):1–37, 2002.
- [34] L. Greenberg and D. Ioffe. On an invariance principle for phase separation lines. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(5):871–885, 2005.
- [35] O. Hryniv. On local behaviour of the phase separation line in the 2D Ising model. *Probab. Theory Related Fields*, 110(1):91–107, 1998.
- [36] D. Ioffe. Large deviations for the 2D Ising model: a lower bound without cluster expansions. *J. Statist. Phys.*, 74(1-2):411–432, 1994.
- [37] D. Ioffe. Exact large deviation bounds up to  $T_c$  for the Ising model in two dimensions. *Probab. Theory Related Fields*, 102(3):313–330, 1995.
- [38] D. Ioffe. Ornstein-Zernike behaviour and analyticity of shapes for self-avoiding walks on  $\mathbf{Z}^d$ . *Markov Process. Related Fields*, 4(3):323–350, 1998.
- [39] E. Lubetzky, F. Martinelli, and A. Sly. Harmonic pinnacles in the discrete Gaussian model. *Comm. Math. Phys.*, 344(3):673–717, 2016.
- [40] R. A. Minlos and J. G. Sinaĭ. The phenomenon of “separation of phases” at low temperatures in certain lattice models of a gas. I. *Mat. Sb. (N.S.)*, 73 (115):375–448, 1967.
- [41] S. Miracle-Sole. Surface tension, step free energy, and facets in the equilibrium crystal. *J. Statist. Phys.*, 79(1):183–214, Apr 1995.
- [42] C.-E. Pfister and Y. Velenik. Large deviations and continuum limit in the 2D Ising model. *Probab. Theory Related Fields*, 109(4):435–506, 1997.
- [43] C.-E. Pfister and Y. Velenik. Interface, surface tension and reentrant pinning transition in the 2D Ising model. *Comm. Math. Phys.*, 204(2):269–312, 1999.

- [44] A. Pisztora. Surface order large deviations for Ising, Potts and percolation models. *Probab. Theory Related Fields*, 104(4):427–466, 1996.
- [45] E. Rio. *Asymptotic theory of weakly dependent random processes*, volume 80 of *Probability Theory and Stochastic Modelling*. Springer, Berlin, 2017.
- [46] J. M. Steele. *Probability theory and combinatorial optimization*, volume 69 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.
- [47] H. N. V. Temperley. Statistical mechanics and the partition of numbers. II. The form of crystal surfaces. *Proc. Cambridge Philos. Soc.*, 48:683–697, 1952.
- [48] H. van Beijeren. Interface sharpness in the Ising system. *Comm. Math. Phys.*, 40(1):1–6, 1975.
- [49] J. Černý and R. Kotecký. Interfaces for random cluster models. *J. Statist. Phys.*, 111(1-2):73–106, 2003.
- [50] O. Zeitouni. Branching random walks and Gaussian fields. In *Probability and statistical physics in St. Petersburg*, volume 91 of *Proc. Sympos. Pure Math.*, pages 437–471. Amer. Math. Soc., Providence, RI, 2016.

R. GHEISSARI

DEPARTMENT OF STATISTICS, UC BERKELEY, 367 EVANS HALL, BERKELEY, CA 94720, USA.

*E-mail address:* [gheissari@berkeley.edu](mailto:gheissari@berkeley.edu)

E. LUBETZKY

COURANT INSTITUTE, NEW YORK UNIVERSITY, 251 MERCER STREET, NEW YORK, NY 10012, USA.

*E-mail address:* [eyal@courant.nyu.edu](mailto:eyal@courant.nyu.edu)