J. Eur. Math. Soc. 18, 931–995

DOI 10.4171/JEMS/606

Pietro Caputo · Eyal Lubetzky · Fabio Martinelli · Allan Sly ·
Fabio Lucio Toninelli

Scaling limit and cube-root fluctuations in SOS surfaces above a wall

Received February 27, 2013 and in revised form February 20, 2014

Abstract. Consider the classical $(2+1)$-dimensional Solid-On-Solid model above a hard wall on an $L \times L$ box of $\mathbb{Z}^2$. The model describes a crystal surface by assigning a nonnegative integer height η_x to each site x in the box and 0 heights to its boundary. The probability of a surface configuration η is proportional to $\exp(-\beta\mathcal{H}(\eta))$, where β is the inverse-temperature and $\mathcal{H}(\eta)$ sums the absolute values of height differences between neighboring sites.

We give a full description of the shape of the SOS surface for low enough temperatures. First we show that with high probability (w.h.p.) the height of almost all sites is concentrated on two levels, $H(L) = \lfloor (1/4\beta) \log L \rfloor$ and $H(L) - 1$. Moreover, for most values of L the height is concentrated on the single value $H(L)$. Next, we study the ensemble of level lines corresponding to the heights $(H(L), H(L)-1, \ldots)$. We prove that w.h.p. there is a unique macroscopic level line for each height. Furthermore, when taking a diverging sequence of system sizes L_k, the rescaled macroscopic level line at height $H(L_k) - n$ has a limiting shape if the fractional parts of $(1/4\beta)\log L_k$ converge to a noncritical value. The scaling limit is an explicit convex subset of the unit square Q and its boundary has a flat component on the boundary of Q. Finally, the highest macroscopic level line has $L_k^{1/3+o(1)}$ fluctuations along the flat part of the boundary of its limiting shape.

Keywords. SOS model, scaling limits, loop ensembles, random surface models

1. Introduction

The $(d+1)$-dimensional *Solid-On-Solid* model is a crystal surface model whose definition goes back to Temperley [38] in 1952 (also known as the Onsager–Temperley sheet). At low temperatures, the model approximates the interface between the plus and minus

P. Caputo: Dipartimento di Matematica, Università Roma Tre,
Largo S. Murialdo 1, 00146 Roma, Italy; e-mail: caputo@mat.uniroma3.it

E. Lubetzky: Courant Institute, New York University, New York, NY 10012-1185, USA;
e-mail: eyal@courant.nyu.edu

F. Martinelli: Dipartimento di Matematica, Università Roma Tre,
Largo S. Murialdo 1, 00146 Roma, Italy; e-mail: martin@mat.uniroma3.it

A. Sly: Department of Statistics, UC Berkeley, Berkeley, CA 94720, USA;
e-mail: sly@stat.berkeley.edu

F. L. Toninelli: Université de Lyon, CNRS and Institut Camille Jordan, Université Lyon 1,
43 bd du 11 novembre 1918, 69622 Villeurbanne, France; e-mail: toninelli@math.univ-lyon1.fr

Mathematics Subject Classification (2010): 60K35, 82B41, 82C24

phases in the $(d+1)$D Ising model, with particular interest stemming from the study of 3D Ising.

The configuration space of the model on a finite box $\Lambda \subset \mathbb{Z}^d$ with boundary conditions at height zero is the set of all height functions η on $\mathbb{Z}^d$ such that $\Lambda \ni x \mapsto \eta_x \in \mathbb{Z}$ whereas $\eta_x = 0$ for all $x \notin \Lambda$. The probability of η is given by the Gibbs distribution proportional to

$$\exp\Big(-\beta \sum_{x\sim y} |\eta_x - \eta_y|\Big), \tag{1.1}$$

where $\beta > 0$ is the inverse-temperature and $x \sim y$ denotes a nearest-neighbor bond in $\mathbb{Z}^d$.

Numerous works have studied the rich random surface phenomena, e.g. roughening, localization/delocalization, layering and wetting to name but a few, exhibited by the SOS model and some of its many variants. These include the *discrete Gaussian* (replacing $|\eta_x - \eta_y|$ by $|\eta_x - \eta_y|^2$ for the integer analogue of the Gaussian free field), *restricted* SOS (nearest neighbor gradients restricted to $\{0, \pm 1\}$), *body centered* SOS [6], etc. (for more on these flavors see e.g. [1, 5, 9]).

Of special importance is SOS with $d = 2$, the only dimension featuring a *roughening transition*. For $d = 1$, it is well known [38, 39, 23] that the SOS surface is *rough* (delocalized) for any $\beta > 0$, i.e., the expected height at the origin diverges (in absolute value) in the thermodynamic limit $|\Lambda| \to \infty$. However, for $d \geq 3$ it is known that the surface is *rigid* (localized) for any $\beta > 0$ (see [14]), i.e., $|\eta_0|$ is uniformly bounded in expectation. A simple Peierls argument shows that this is also the case for $d = 2$ and large enough β [12, 28]. That the surface is rough for $d = 2$ at high temperatures was established in seminal works of Fröhlich and Spencer [25, 26, 27]. Numerical estimates for the critical inverse-temperature β_r where the roughening transition occurs suggest that $\beta_r \approx 0.806$. When the $(2+1)$D SOS surface is constrained to stay above a hard wall (or floor), i.e. η is constrained to be nonnegative in (1.1), Bricmont, El-Mellouki and Fröhlich [13] showed in 1986 the appearance of *entropic repulsion*: for large enough β, the floor pushes the SOS surface to diverge even though $\beta > \beta_r$. More precisely, using Pirogov–Sinaĭ theory (see the review [37]), the authors of [13] showed that the SOS surface on an $L \times L$ box rises, amid the penalizing zero boundary conditions, to an average height in the interval $[(1/C\beta)\log L, (C/\beta)\log L]$ for some absolute constant $C > 0$, in favor of freedom to create spikes downwards. In a companion paper [15], focusing on the dynamical evolution of the model, we established that the average height is in fact $(1/4\beta)\log L$ up to an additive $O(1)$-error.

Entropic repulsion is one of the key features of the physics of random surfaces. This phenomenon has been rigorously analyzed mainly for some continuous-height variants of the SOS model in which the interaction potential $|\eta_x - \eta_y|$ is replaced by a *strictly convex* potential $V(\eta_x - \eta_y)$; see, e.g., [11, 8, 17, 19, 10, 40, 41], and also [4] for a recent analysis of the wetting transition in the SOS model. It was shown in the companion paper [15] that entropic repulsion drives the evolution of the surface under the natural single-site dynamics. Started from a flat configuration, the surface rises to an average height of $(1/4\beta)\log L - O(1)$ through a sequence of metastable states, corresponding roughly to plateaux at heights $0, 1, \ldots, (1/4\beta)\log L$.

Despite the recent progress on understanding the typical height of the surface, little was known on its actual 3D shape. The fundamental problem is the following:

Question. *Consider the ensemble of all level lines of the low temperature* $(2+1)$D *SOS on an* $L \times L$ *box with floor and boundary conditions at height zero, rescaled to the unit square.*

(i) *Do these jointly converge to a scaling limit as* $L \to \infty$*, e.g., in Hausdorff distance?*
(ii) *If so, can the limit be explicitly described?*
(iii) *For finite large* L*, what are the fluctuations of the level lines around their limit?*

In this work we fully resolve parts (i) and (ii) and partially answer part (iii). En route, we also establish that for most values of L the surface height concentrates on the single level $\lfloor (1/4\beta) \log L \rfloor$.

1.1. Main results. We now state our three main results. As we will see, two parameters rule the macroscopic behavior of the SOS surface:

$$H(L) = \left\lfloor \frac{1}{4\beta} \log L \right\rfloor, \qquad \alpha(L) = \frac{1}{4\beta} \log L - H(L), \tag{1.2}$$

that is, integer part and the fractional part of $(1/4\beta) \log L$. The first result states that with probability tending to one as $L \to \infty$ the SOS surface has a large fraction of sites at height equal either to $H(L)$ or to $H(L) - 1$. Moreover only one of the two possibilities holds, depending on whether $\alpha(L)$ is above or below a critical threshold that can be expressed in terms of an explicit critical parameter $\lambda_c = \lambda_c(\beta)$. The second result describes the macroscopic shape for large L of any finite collections of level lines at height $H(L), H(L) - 1, \ldots$. The third result establishes cube root fluctuations of the level lines along the flat part of its macroscopic shape.

In what follows we consider boxes Λ of the form $\Lambda = \Lambda_L = [1, L] \times [1, L]$, $L \in \mathbb{N}$. We write π_Λ^0 for the SOS distribution on Λ with floor and boundary conditions at height

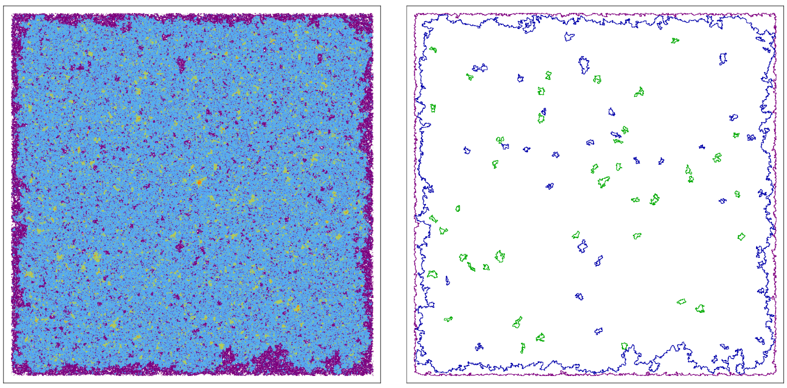

Fig. 1. Loop ensemble formed by the level lines of an SOS configuration on a box of side-length 1000 with floor (showing loops longer than 100).

zero, and let $\hat\pi_\Lambda^0$ be its analog without a floor. The function $L \mapsto \lambda(L) \in (0,\infty)$ appearing below is explicitly given in terms of $\alpha(L)$ (see Definition 2.5 below). For large β it satisfies $\lambda(L)e^{-4\beta\alpha(L)} \simeq 1$.

Theorem 1 (Height concentration). *Fix $\beta > 0$ sufficiently large and define*

$$E_h = \left\{\eta : \#\{x : \eta_x = h\} \ge \tfrac{9}{10}L^2\right\}.$$

Then the SOS measure π_Λ^0 on the box $\Lambda = \Lambda_L$ with floor, at inverse-temperature β, satisfies

$$\lim_{L\to\infty} \pi_\Lambda^0(E_{H(L)-1} \cup E_{H(L)}) = 1. \tag{1.3}$$

Furthermore, the typical height of the configuration is governed by $L \mapsto \lambda(L)$ as follows. Let Λ_k be a diverging sequence of boxes with side-lengths L_k. For an explicit constant $\lambda_c > 0$ (given by (3.5)) we have:

(i) *If* $\liminf_{k\to\infty} \lambda(L_k) > \lambda_c$ *then* $\lim_{k\to\infty} \pi_{\Lambda_k}^0(E_{H(L_k)}) = 1$.
(ii) *If* $\limsup_{k\to\infty} \lambda(L_k) < \lambda_c$ *then* $\lim_{k\to\infty} \pi_{\Lambda_k}^0(E_{H(L_k)-1}) = 1$.

Remark 1.1. The constant $9/10$ in the definition of E_h can be replaced by $1-\epsilon$ for any arbitrarily small $\epsilon > 0$ provided that β is large enough. As shown in Remark 3.7, for large enough fixed β, $\lambda_c \simeq 4\beta$ whereas $\lambda(L) \simeq e^{4\beta\alpha(L)}$, and hence most values of $L \in \mathbb{N}$ will yield $\pi_\Lambda^0(E_{H(L)}) = 1 - o(1)$.

It is interesting to compare these results to the 2D Gaussian free field (GFF) conditioned to be nonnegative, qualitatively akin to high-temperature SOS. It is known [8] that the height of the GFF in the box $\Lambda = \Lambda_L$ with floor and boundary conditions at height zero, at any point $x \in \Lambda$ such that $\mathrm{dist}(x, \partial\Lambda) \ge \delta L$, $\delta > 0$, is asymptotically the same as the maximal height in the unconditioned GFF in Λ. (Here, $\mathrm{dist}(\cdot,\cdot)$ denotes Euclidean distance and $\partial\Lambda$ is the external boundary of Λ.) On the other hand, our results show that the SOS surface is lifted to height $H(L)$ or $H(L)-1$, which is asymptotically only one half of the SOS unconditioned maximum. Moreover, on the comparison of the maxima of the fields with and without wall we obtain the following:

Corollary 1.2. *Fix $\beta > 0$ large enough, let X_L^* be the maximum of the SOS surface on the box $\Lambda = \Lambda_L$ with floor, and let $\widehat{X}_L^*$ be its analog in the SOS model without floor. Then for any diverging sequence $\varphi(L)$ one has*

$$\lim_{L\to\infty} \hat\pi_\Lambda^0\left(\left|\widehat{X}_L^* - \tfrac{1}{2\beta}\log L\right| \ge \varphi(L)\right) = 0, \tag{1.4}$$

$$\lim_{L\to\infty} \pi_\Lambda^0\left(\left|X_L^* - \tfrac{3}{4\beta}\log L\right| \ge \varphi(L)\right) = 0. \tag{1.5}$$

We now address the scaling limit of the ensemble of level lines. The latter is described as follows (see Section 3 for the full details). As for the 2D Ising model (see e.g. [20, 7]), for the low-temperature SOS model without floor there is a natural notion of surface tension $\tau(\cdot)$ satisfying the strict convexity property. We emphasize that the surface tension we consider here is constructed in the usual way, namely by imposing Dobrushin type conditions (between height zero and height one) around a box. Consider the associated Wulff shape, namely the convex body with support function τ, and let $\mathcal{W}_1$ denote the Wulff shape rescaled to enclose area 1. For a given $s > 0$, define the shape $\mathcal{L}_c(s)$ by taking the union of all possible translates of $\ell_c(s)\mathcal{W}_1$ within the unit square, with an explicit dilation parameter $\ell_c(s)$ (defined in (3.4)) which satisfies $\ell_c(s) \sim 2\beta/s$ for large β. (Of course, $\mathcal{L}_c(s)$ is defined only if $\ell_c(s)\mathcal{W}_1$ fits inside a unit square.) Next, for a fixed $\lambda_\star > 0$, consider the nested shapes $\{\mathcal{L}_c(\lambda_\star^{(n)})\}_{n\geq 0}$ obtained by taking s equal to $\lambda_\star^{(n)} := e^{4\beta n}\lambda_\star$, $n = 0, 1, \ldots$ as shown in Figure 2. See Claim 3.8 for the interpretation of $\mathcal{L}_c$ as the solution of a surface tension variational problem.

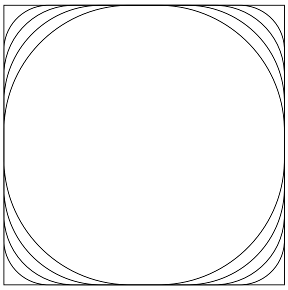

Fig. 2. The nested limiting shapes $\{\mathcal{L}_c(\lambda_\star^{(n)})\}$ of the rescaled loop ensemble $(1/L)\{\Gamma_n\}$.

The next theorem gives a necessary and sufficient condition for the existence of the scaling limit of the ensemble of level lines in terms of the above defined shapes. As a convention, we often write "with high probability", or w.h.p., whenever the probability of an event is at least $1 - e^{-c(\log L)^2}$ for some constant $c > 0$, uniformly in L.

Theorem 2 (Shape Theorem). *Fix $\beta > 0$ sufficiently large and let L_k be a diverging sequence of side-lengths. Set $H_k = H(L_k)$. For an SOS surface on the box Λ_k with side L_k, let $(\Gamma_0^{(k)}, \Gamma_1^{(k)}, \ldots)$ be the collections of loops with length at least $(\log L_k)^2$ belonging to the level lines at heights $(H_k, H_k - 1, \ldots)$, respectively. Then:*

(a) *W.h.p. the level lines of every height $h > H_k$ consist of loops shorter than $(\log L_k)^2$, while $\Gamma_0^{(k)}$ is either empty or contains a single loop, and $\Gamma_n^{(k)}$ consists of exactly one loop for each $n \geq 1$.*

(b) *If $\lambda_\star := \lim_{k\to\infty} \lambda(L_k)$ exists and differs from λ_c (as given by (3.5)) then the rescaled loop ensemble $(1/L_k)(\Gamma_0^{(k)}, \Gamma_1^{(k)}, \ldots)$ converges to a limit in the Hausdorff distance: for any $\varepsilon > 0$, w.h.p.*

- *if $\lambda_\star > \lambda_c$ then*
$$\sup_{n\geq 0} d_{\mathcal{H}}\Big(\tfrac{1}{L_k}\Gamma_n^{(k)}, \mathcal{L}_c(\lambda_\star^{(n)})\Big) \leq \varepsilon,$$
where $d_{\mathcal{H}}$ denotes the Hausdorff distance;
- *if instead $\lambda_\star < \lambda_c$ then $\Gamma_0^{(k)}$ is empty while*
$$\sup_{n\geq 1} d_{\mathcal{H}}\Big(\tfrac{1}{L_k}\Gamma_n^{(k)}, \mathcal{L}_c(\lambda_\star^{(n)})\Big) \leq \varepsilon.$$

As for the critical behavior, from the above theorem we immediately read that it is possible to have $\lambda(L_k) \to \lambda_c$ without admitting a scaling limit for the loop ensemble (consider a sequence that oscillates between the subcritical and supercritical regimes). However, understanding the critical window around λ_c and the limiting behavior there remains an interesting open problem.

The fluctuations of the loop $\Gamma_0^{(k)}$ (the macroscopic plateau at level $H(L_k)$ if it exists) from its limit $\mathcal{L}_c(\lambda_\star)$ along the side-boundaries are now addressed. As shown in Figure 2, the boundary of the limit shape $\mathcal{L}_c(\lambda_\star)$ coincides with the boundary of the unit square Q except for a neighborhood of the four corners of Q. Let the interval $[a, 1-a]$, $a = a(\lambda_\star) > 0$, denote the horizontal projection of the intersection of the shape $\mathcal{L}_c(\lambda_\star)$ with the bottom side of the unit square Q.

Theorem 3 (Cube-root fluctuations). *In the setting of Theorem 2 suppose $\lambda_\star > \lambda_c$. Then for any $\epsilon > 0$, w.h.p. the vertical fluctuation of $\Gamma_0^{(k)}$ from the boundary interval*

$$I_\epsilon^{(k)} = [a(1+\epsilon)L_k, (1-a(1+\epsilon))L_k]$$

is of order $L_k^{1/3+o(1)}$. More precisely, let $\rho(x) = \max\{y \le L_k/2 : (x, y) \in \Gamma_0^{(k)}\}$ (and set for instance $\rho(x) = 0$ if there is no y satisfying $(x, y) \in \Gamma_0^{(k)}$) be the vertical fluctuation of $\Gamma_0^{(k)}$ from the bottom boundary of Λ_k at coordinate x. Then w.h.p.

$$L_k^{1/3-\epsilon} < \sup_{x \in I_\epsilon^{(k)}} \rho(x) < L_k^{1/3+\epsilon}.$$

Remark 1.3. We will actually prove the stronger fact that w.h.p. a fluctuation of at least $L_k^{1/3-\epsilon}$ is attained in every subinterval of $I_\epsilon^{(k)}$ of length $L_k^{2/3-\epsilon}$ (cf. Section 6.4).

As a direct corollary of Theorem 3 it was deduced in [16] that the following upper bound on the fluctuations of *all* level lines $\Gamma_n^{(k)}$ ($n \ge 1$) holds.

Corollary 1.4 (Cascade of fluctuation exponents). *In the same setting of Theorem 3, let $\rho(n, x)$ be the vertical fluctuation of $\Gamma_n^{(k)}$ from the bottom boundary at coordinate x. Let $0 < t < 1$ and let $n = \lfloor t H_k \rfloor$ with H_k defined in Theorem 2. Then for any $\epsilon > 0$,*

$$\lim_{k\to\infty} \pi_{\Lambda_k}^0 \Big(\sup_{x \in I_\epsilon^{(k)}} \rho(n, x) > L_k^{(1-t)/3+\epsilon} \Big) = 0.$$

1.2. Related work. In the two papers [35, 36] Schonmann and Shlosman studied the limiting shape of a droplet of the low temperature 2D Ising with minus boundary under a prescribed small positive external field, proportional to the inverse of the side-length L. The behavior of the droplet of plus spins in this model is qualitatively similar to the behavior of the top loop Γ_0 in our case. Here, instead of an external field, it is the entropic repulsion phenomenon that induces the surface to rise to level $H(L)$ producing the macroscopic loop Γ_0. In line with this connection, the shape $\mathcal{L}_c(s)$ appearing in Theorem 2 is constructed in the same way as the limiting shape of the plus droplet in the aforementioned works, although with a Wulff shape generated by the surface tension of the SOS model. In particular, as in [35, 36], the shape $\mathcal{L}_c(s)$ arises as the solution to a variational problem (see Section 3 below).

An important difference between the two models, however, is that in our case there exist $H(L)$ loops (rather than just one), which are interacting in two nontrivial ways. First, by definition, they cannot cross each other. Second, they can weakly either attract or repel one another depending on the local geometry and height. Moreover, the box boundary itself can attract or repel the level lines. A prerequisite to proving Theorem 2 is to overcome these "pinning" issues. We remark that at times such pinning issues have been overlooked in the relevant literature.

As for the fluctuations of the plus droplet from its limiting shape, it was argued in [35] that these should be normal (i.e., of order $\sqrt{L}$). However, due to the analogy mentioned above between the models, it follows from our proof of Theorem 3 that these fluctuations are in fact $L^{1/3+o(1)}$ along the flat pieces of the limiting shape, while it seems natural to conjecture that normal fluctuations appear along the curved portions, where the limiting shapes corresponding to distinct levels are macroscopically separated (see Figure 2).

There is a rich literature on contour models featuring similar cube root fluctuations. In some of these works (e.g., [3, 22, 32, 41]) the phenomenon is induced by an externally imposed constraint (by conditioning on the event that the contour contains a large area and/or by adding an external field); see in particular [41] for the above mentioned case of the 2D Ising model in a weak external field, and the recent works [29, 30] for refined bounds in the case of FK percolation. In other works, modeling ordered random walks (e.g., [18, 33] to name but a few), the exact solvability of the model (e.g., via determinantal representations) plays an essential role in the analysis. In our case, the phenomenon is again a consequence of the tilting of the distribution of contours induced by the entropic repulsion. The lack of exact solvability for the $(2+1)$D SOS forces us to resort to cluster expansion techniques and contour analysis as in the framework of [20].

We conclude by mentioning some problems that remain unaddressed by our results. The first is to establish the exponents for the fluctuations of all intermediate level lines from the side-boundaries. We believe the upper bound in Corollary 1.4 features the correct cascade of exponents. The second problem is to find the correct fluctuation exponent of the level lines $\{\Gamma_n\}$ around the curved part of their limiting shapes $\{\mathcal{L}_c(\lambda_\star^{(n)})\}$. Third, we expect that, as in [3, 29, 30], the fluctuation exponent of the highest level line around its convex envelope is $1/3$, while, as mentioned above, there should be normal fluctuations around the curved parts of the deterministic limiting shape.

1.3. On the ensemble of macroscopic level lines. We turn to a high-level description of the statistics of the level lines of the SOS interface. Given a *closed contour* γ (i.e., a closed loop of dual edges as for the standard Ising model), a positive integer h and a surface configuration η we say that γ is an *h-contour* (or *h-level line*) for η if the surface height jumps from being at least h along the internal boundary of γ to at most $h-1$ along the external boundary of γ. Clearly an h-contour γ is energetically penalized proportionally to its length $|\gamma|$ because of the form of the SOS energy function. As in many spin models admitting a contour representation, with high probability contours are either all small, say $|\gamma| = O((\log L)^2)$, or there exist macroscopically large ones (i.e. $|\gamma| \propto L$), if L is the size of the system. An instance of the first situation is the SOS model *without* a wall and boundary conditions at height zero (see [12]). On the contrary, macroscopic contours appear in the low temperature 2D Ising model with *negative* boundary conditions and a *positive* external field $\mathcal{H}$ of the form $\mathcal{H} = B/L$, $B > 0$, as in [35, 36]. In this case the probabilistic weight of a contour separating the inside plus spins from the outside minus spins is roughly given by $\exp(-\beta|\gamma| + \Psi(\gamma) + m^*_\beta \mathcal{H} A(\gamma))$, where $A(\gamma)$ denotes the area enclosed by γ, m^*_β is the spontaneous magnetization and $\Psi(\gamma)$ is a "decoration" term which is not essential for the present discussion. If the parameter B is above a certain threshold then the area term dominates the boundary term and a macroscopic contour appears with high probability. Moreover, by simple isoperimetric arguments, the macroscopic contour is unique in this case.

The SOS model with a wall shares some similarities with the Ising example above but has a richer structure that can be roughly described as follows.

Suppose $\{\gamma_1, \dots, \gamma_n\}$ are macroscopic h-contours corresponding to heights $h = 1, \dots, n$ and no other macroscopic contour exists. Then necessarily the collection $\{\gamma_i\}_{i=1}^n$ must consist of *nested* contours, with γ_n and γ_1 being the innermost and outermost contour respectively. If we denote by Λ_i the region enclosed by γ_i, and by A_i the annulus $\Lambda_i \setminus \Lambda_{i+1}$, then the partition function of all the surfaces satisfying the above requirements can be written as

$$Z(\gamma_1, \dots, \gamma_n) = \exp\Big(-\beta \sum_i |\gamma_i|\Big) \prod_i Z_{A_i}$$

where Z_{A_i} is the partition function of the SOS model in A_i, with a wall at height zero, boundary conditions at height $h = i$ and restricted to configurations without macroscopic contours.[1]

Usually (see, e.g., [20]) in these cases one tries to exponentiate the partition functions Z_{A_i} using cluster expansion techniques. However, because of the presence of the wall, one cannot apply this approach directly and it is instead more convenient to compare Z_{A_i} with $\hat{Z}_{A_i}$, where $\hat{Z}_{A_i}$ is as Z_{A_i} but *without* the wall. One then observes that the ratio $Z_{A_i}/\hat{Z}_{A_i}$ is simply the probability that the surface is nonnegative computed for the Gibbs distribution of the SOS model in A_i with boundary conditions at height $h = i$, no wall, and conditioned to have no macroscopic contours. The key point, which was already noted in [15], is that with respect to the above Gibbs measure the random variables $\{\mathbf{1}_{\eta_x \geq 0}\}_{x \in A_i}$ behave approximately as i.i.d. with

$$\mathbb{P}(\eta_x \geq 0) \simeq 1 - c_\infty e^{-4\beta(i+1)},$$

where is c_∞ a computable constant. Therefore,

$$Z_{A_i} \simeq \exp(-c_\infty e^{-4\beta(i+1)}|A_i|)\hat{Z}_{A_i}.$$

In conclusion, rewriting $|A_i| = |\Lambda_i| - |\Lambda_{i+1}|$, we get

$$Z(\gamma_1, \dots, \gamma_n) \propto \exp\Big(\sum_i [-\beta|\gamma_i| + c_\infty e^{-4\beta i}(1 - e^{-4\beta})|\Lambda_i|]\Big) \prod_i \hat{Z}_{A_i}.$$

[1] Strictly speaking, one should also require that the height is at most i [at least i] along the inner [outer] boundary of the annulus, but we skip these details for the present discussion.

The terms proportional to the area encode the effect of the entropic repulsion and play the same role as the magnetic field term in the Ising example. Cluster expansion techniques can now be applied to the partition functions $\hat{Z}_{A_i}$ without wall. As in many other similar cases their net result is the appearance of a decoration term $\Psi(\gamma_i) = O(\varepsilon_\beta|\gamma_i|)$ for each contour, ε_β small for large β, and an effective many-body *interaction* $\Phi(\gamma_1, \ldots, \gamma_n)$ among the contours, which however is very rapidly decaying with their mutual distance. Thus the probability of the above macroscopic contours should then be proportional to

$$\exp\Big(\sum_i [-\beta|\gamma_i| + c_\infty e^{-4\beta i}(1 - e^{-4\beta})|\Lambda_i| + \Psi(\gamma_i)] + \Phi(\gamma_1, \ldots, \gamma_n)\Big). \tag{1.6}$$

Note that in each term of the above sum the area part is dominant up to height $i \simeq (1/4\beta)\log L$. In other words, macroscopic h-contours are sustained by the entropic repulsion up to a height $h \simeq (1/4\beta)\log L$ while higher contours are exponentially suppressed. More precisely, if we measure heights relative to $H(L) = \lfloor (1/4\beta)\log L \rfloor$, then the i^{th} area term can be rewritten as

$$\lambda \frac{e^{4\beta(H(L)-i)}}{L}|\Lambda_i| \quad \text{with} \quad \lambda = \lambda(L) = c_\infty e^{4\beta\alpha(L)}(1 - e^{-4\beta}).$$

The quantity $\lambda(L)$ is exactly the key parameter appearing in the main theorems. Notice that the loop Γ_0 appearing in Theorem 2 would correspond to the contour γ_n if $n = H(L)$.

Summarizing, the macroscopic contours behave like nested random loops with an area bias and with some interaction potential Φ. While the latter is in many ways a weak perturbation, in principle delicate pinning effects may occur among the different level lines, as emphasized earlier.

Although one could try to implement the above line of reasoning directly, we found it more convenient to combine the above ideas with monotonicity properties of the model (with respect to the height of the boundary conditions and/or the height of the wall) to reduce ourselves always to the analysis of one single macroscopic contour at a time. That allowed us to partially overcome the above mentioned pinning problem. On the other hand, a more detailed control of the interaction between level lines should be crucial in order to derive finer results.

2. General tools

In this section we collect some preliminary definitions together with basic results which will be used several times throughout the paper. Once combined together with standard cluster expansion methods, they precisely quantify the effect of the entropic repulsion from the floor.

2.1. Preliminaries. In order to formulate our first tools we need a bit of extra notation.

Boundary conditions and infinite volume limit. Given a height function $\mathbb{Z}^2 \ni x \mapsto \tau_x \in \mathbb{Z}$ (the boundary conditions) and a finite set $\Lambda \subset \mathbb{Z}^2$ we denote by $\Omega_\Lambda^{(\tau)}$ (resp. $\hat{\Omega}_\Lambda^{(\tau)}$) all the height functions η on $\mathbb{Z}^2$ such that $\Lambda \ni x \mapsto \eta_x \in \mathbb{Z}_+$ (resp. $\Lambda \ni x \mapsto \eta_x \in \mathbb{Z}$)

whereas $\eta_x = \tau_x$ for all $x \notin \Lambda$. The corresponding Gibbs measure given by (1.1) will be denoted by π_Λ^τ (resp. $\hat\pi_\Lambda^\tau$). In other words π_Λ^τ describes the SOS model in Λ with boundary conditions τ and floor at zero while $\hat\pi_\Lambda^\tau$ describes the SOS model in Λ with boundary conditions τ and no floor. The corresponding partition functions will be denoted by Z_Λ^τ and $\hat Z_\Lambda^\tau$ respectively. If τ is constant and equal to $j \in \mathbb{Z}$ we will simply replace τ by j in all the notation. We will denote by $\hat\pi$ the infinite volume Gibbs measure obtained as the thermodynamic limit of the measure $\hat\pi_\Lambda^0$ along an increasing sequence of boxes. The limit exists and does not depend on the sequence of boxes (see [12]).

Monotonicity. On the set $\Omega_\Lambda^{(\tau)}$ (or on $\hat\Omega_\Lambda^{(\tau)}$) one can define the partial order such that $\sigma \le \eta$ if $\sigma_x \le \eta_x$ for every $x \in \Lambda$. A function f is said to be increasing (resp. decreasing) if $\sigma \le \eta$ implies $f(\sigma) \le f(\eta)$ (resp. $f(\sigma) \ge f(\eta)$). An event E is called increasing (resp. decreasing) if the indicator function $f = \mathbf{1}_E$ is increasing (resp. decreasing). The SOS Gibbs measure enjoys the following crucial properties.

- Monotonicity with respect to boundary conditions: If $\tau \le \tau'$, then $\pi_\Lambda^\tau(f) \le \pi_\Lambda^{\tau'}(f)$ for all increasing functions. The same holds for $\hat\pi_\Lambda^\tau$.
- Monotonicity with respect to floor constraints: If $V \subset \Lambda$ and $\xi \in \mathbb{Z}^V$, then for all increasing functions f one has $\hat\pi_\Lambda^\tau(f) \le \hat\pi_\Lambda^\tau(f \mid E_\xi)$, where E_ξ denotes the increasing event $\eta\restriction_V \ge \xi$. Moreover, $\hat\pi_\Lambda^\tau(f \mid E_\xi)$ is an increasing function of ξ.

The above items are a consequence of the more general FKG inequality [24]: for all $\Lambda \subset \mathbb{Z}^2$ and all boundary conditions τ, one has

$$\hat\pi_\Lambda^\tau(fg) \ge \hat\pi_\Lambda^\tau(f)\hat\pi_\Lambda^\tau(g)$$

for all bounded increasing functions f, g. In turn the FKG inequality follows from the so-called *FKG lattice condition* which is easily verified for the SOS Gibbs measure. We shall make repeated use of such relations, and will often refer to them simply as "monotonicity".

Contours and level lines. The level lines of the SOS surface, and the corresponding loop ensemble they give rise to, are formally defined as follows.

Definition 2.1 (Geometric contour). We let $(\mathbb{Z}^2)^*$ be the dual lattice of $\mathbb{Z}^2$ and we define a *bond* to be any segment joining two neighboring sites in $(\mathbb{Z}^2)^*$. Two sites x, y in $\mathbb{Z}^2$ are said to be *separated by a bond* e if their distance (in $\mathbb{R}^2$) from e is $1/2$. A pair of orthogonal bonds which meet in a site $x^* \in (\mathbb{Z}^2)^*$ is said to be a *linked pair of bonds* if both bonds are on the same side of the forty-five degrees line (in the north-east direction) across x^*. A *geometric contour* (for short a *contour*) is a sequence $e_0, \dots, e_n$ of bonds such that:

(1) $e_i \ne e_j$ for $i \ne j$, except for $i = 0$ and $j = n$ where $e_0 = e_n$,
(2) for every i, e_i and e_{i+1} have a common vertex in $(\mathbb{Z}^2)^*$,
(3) if $e_i, e_{i+1}, e_j, e_{j+1}$ intersect at some $x^* \in (\mathbb{Z}^2)^*$, then e_i, e_{i+1} and e_j, e_{j+1} are linked pairs of bonds.

We denote the length of a contour γ by $|\gamma|$, its interior (the sites in $\mathbb{Z}^2$ it surrounds) by Λ_γ and its interior area (the number of such sites) by $|\Lambda_\gamma|$. Moreover we let Δ_γ be the set of sites in $\mathbb{Z}^2$ such that either their distance (in $\mathbb{R}^2$) from γ is $1/2$, or their distance from the set of vertices in $(\mathbb{Z}^2)^*$ where two nonlinked bonds of γ meet equals $1/\sqrt{2}$. Finally, we let $\Delta_\gamma^+ = \Delta_\gamma \cap \Lambda_\gamma$ and $\Delta_\gamma^- = \Delta_\gamma \setminus \Delta_\gamma^+$.

Definition 2.2 (h-contour). Given a contour γ we say that γ is an *h-contour* (or an *h-level line*) for the configuration η if the height function restricted to Δ_γ^- (resp. to Δ_γ^+) is pointwise smaller than $h-1$ (resp. larger than h), in formulas

$$\eta\restriction_{\Delta_\gamma^-} \le h-1, \qquad \eta\restriction_{\Delta_\gamma^+} \ge h.$$

We will say that γ is a *contour for the configuration* η if there exists h such that γ is an h-contour for η. Contours longer than $(\log L)^2$ will be called *macroscopic contours*.[2] Finally $\mathscr{C}_{\gamma,h}$ will denote the event that γ is an h-contour.

Note that γ can be at the same time an h-contour and an h'-contour with $h \neq h'$. In general, contours are not disjoint but they cannot cross.

Definition 2.3 (Negative h-contour). We say that a closed contour γ is a *negative h-contour* if the height on the external boundary of γ is at least h and the height on the internal boundary of γ is at most $h-1$. That is to say, denoting this event by $\mathscr{C}_{\gamma,h}^-$, we have $\eta \in \mathscr{C}_{\gamma,h}^-$ iff $\eta\restriction_{\Delta_\gamma^+} \le h-1$ and $\eta\restriction_{\Delta_\gamma^-} \ge h$.

Entropic repulsion parameters. In order to define key parameters measuring the entropic repulsion we need the following lemma whose proof is postponed to Appendix A.1.

Lemma 2.4. *For β large enough the limit $c_\infty := \lim_{h\to\infty} e^{4\beta h}\hat{\pi}(\eta_0 \ge h)$ exists and*

$$|c_\infty - e^{4\beta h}\hat{\pi}(\eta_0 \ge h)| = O(e^{-2\beta h}).$$

Moreover $\lim_{\beta\to\infty} c_\infty = 1$.

Definition 2.5. Given an integer $L > 1$ we define

$$\lambda := \lambda(L) = e^{4\beta\alpha(L)} c_\infty (1 - e^{-4\beta}) \tag{2.1}$$

where $\alpha(L)$ denotes the fractional part of $(1/4\beta)\log L$. Also, for $n \ge 0$, we let $\lambda^{(n)} := \lambda^{(n)}(L) = \lambda e^{4\beta n}$.

[2] This convention is slightly abusive, since the term *macroscopic* is usually reserved to objects with size comparable to the system size L. However, as we will see, it is often the case in our context that with overwhelming probability, there are no contours at intermediate scales between $(\log L)^2$ and L.

2.2. An isoperimetric inequality for contours. The following simple lemma will prove useful in establishing the existence of macroscopic loops.

Lemma 2.6. *For all $\delta' > 0$ there exists a $\delta > 0$ such that the following holds. Let $\{\gamma_i\}$ be a collection of closed contours enclosing areas $A(\gamma_i)$ satisfying $A(\gamma_1) \geq A(\gamma_2) \geq \cdots$, and suppose that*

$$\sum_i |\gamma_i| \leq (1+\delta)4L \quad \textit{and} \quad \sum_i A(\gamma_i) \geq (1-2\delta)L^2.$$

Then the interior of γ_1 contains a square of area at least $(1-\delta')L^2$.

Proof. Define $\alpha_i = A(\gamma_i)[(1-2\delta)L^2]^{-1}$ so that $\sum_i \alpha_i \geq 1$. Then (see, e.g., [20, Section 2.9]) $\sum_i \sqrt{\alpha_i} \geq 1/\sqrt{\alpha_1}$. From the isoperimetric bound $|\gamma_i| \geq 4\sqrt{A(\gamma_i)}$, it follows that

$$\begin{aligned}(1+\delta)4L &\geq \sum_i |\gamma_i| \geq 4\sqrt{(1-2\delta)L^2}\sum_i \sqrt{\alpha_i} \\ &\geq 4\sqrt{(1-2\delta)L^2}\frac{1}{\sqrt{\alpha_1}} = 4(1-2\delta)L^2\frac{1}{\sqrt{A(\gamma_1)}}.\end{aligned}$$

This implies, for δ small enough,

$$A(\gamma_1) \geq \frac{(1-2\delta)^2}{(1+\delta)^2}L^2 \geq (1-8\delta)L^2.$$

Noting that the unit square is the unique shape with area at least 1 and L^1-boundary length at most 4, it follows by continuity that for all $\delta' > 0$ there exists $\delta > 0$ such that if a curve has length at most $(1+\delta)4$ and encloses an area of at least $1-8\delta$ then it contains a square of side-length $1-\delta'$, implying the last assertion of the lemma. □

2.3. Peierls estimates and entropic repulsion. Our first result is an upper bound on the probability of encountering a given h-contour; it is a refinement of [15, Proposition 3.6]. Recall the definition (1.2) of the height $H(L)$, of the parameters $\lambda^{(n)}$ (Definition 2.5) and of the events $\mathscr{C}_{\gamma,h}$, $\mathscr{C}^-_{\gamma,h}$ (Definitions 2.2 and 2.3).

Proposition 2.7. *Fix $j \geq 0$ and consider the SOS model in a finite connected subset V of $\mathbb{Z}^2$ with floor at height 0 and boundary conditions at height $j \geq 0$. There exist δ_h and ε_β with $\lim_{h\to\infty}\delta_h = \lim_{\beta\to\infty}\varepsilon_\beta = 0$ such that, for all $h \in \mathbb{N}$,*

$$\pi_V^j(\mathscr{C}_{\gamma,h}) \leq \exp\big(-\beta|\gamma| + c_\infty(1+\delta_h)|\Lambda_\gamma|e^{-4\beta h}\big)\exp(\varepsilon_\beta\, e^{-4\beta h}|\gamma|\log|\gamma|), \tag{2.2}$$

$$\pi_V^j(\mathscr{C}^-_{\gamma,h}) \leq e^{-\beta|\gamma|}. \tag{2.3}$$

Remark 2.8. Notice that if $h = H(L) - n$ then $c_\infty e^{-4\beta h} = (1-e^{-4\beta})^{-1}\lambda^{(n)}/L$.

Proof of (2.2). Let $Z_{\text{in}}^{+,n}$ (resp. $\hat{Z}_{\text{in}}^{+,n}$) be the partition function of the SOS model in Λ_γ with floor at height 0 (resp. no floor), boundary conditions at height n and $\eta\restriction_{\Delta_\gamma^+} \geq n$.

Similarly let $Z^{-,n}_{\rm out}$ be the partition function of the SOS model in $V \setminus \Lambda_\gamma$ with floor at height 0, boundary conditions at height j along ∂V, at height n along γ and satisfying $\eta\restriction_{\Delta^-_\gamma} \le n$. One has

$$\pi^j_\Lambda(\mathscr{C}_{\gamma,h}) = e^{-\beta|\gamma|}\,\frac{Z^{-,h-1}_{\rm out} Z^{+,h}_{\rm in}}{Z^j_V} \le e^{-\beta|\gamma|}\,\frac{Z^{-,h-1}_{\rm out} Z^{+,h}_{\rm in}}{Z^{-,h-1}_{\rm out} Z^{+,h-1}_{\rm in}} = e^{-\beta|\gamma|}\,\frac{Z^{+,h}_{\rm in}}{Z^{+,h-1}_{\rm in}}$$
$$\le e^{-\beta|\gamma|}\,\frac{\hat Z^{+,h}_{\rm in}}{Z^{+,h-1}_{\rm in}} = e^{-\beta|\gamma|}\,\frac{\hat Z^{+,h-1}_{\rm in}}{Z^{+,h-1}_{\rm in}},$$

where in the last equality we have used the fact that $\hat Z^{+,n}_{\rm in}$ is independent of n.

Let now $\hat\pi^n_{\Lambda_\gamma}$ be the Gibbs measure in Λ_γ with boundary conditions at height n and no floor. Then, using the FKG inequality (all the events involved are increasing and therefore positively correlated), we get

$$\frac{Z^{+,h-1}_{\rm in}}{\hat Z^{+,h-1}_{\rm in}} = \hat\pi^{h-1}_{\Lambda_\gamma}(\eta\restriction_{\Lambda_\gamma} \ge 0 \mid \eta\restriction_{\Delta^+_\gamma} \ge h-1) \ge \hat\pi^{h-1}_{\Lambda_\gamma}(\eta\restriction_{\Lambda_\gamma} \ge 0)$$
$$\ge \prod_{x\in\Lambda_\gamma} \hat\pi^{h-1}_{\Lambda_\gamma}(\eta_x \ge 0) = \prod_{x\in\Lambda_\gamma}[1-\hat\pi^0_{\Lambda_\gamma}(\eta_x \ge h)]. \tag{2.4}$$

In the last step we have used the symmetry of $\hat\pi^0_\Lambda$ under $\eta \leftrightarrow -\eta$. It follows from [15, Proposition 3.9] that $\max_{x\in\Lambda_\gamma}\hat\pi^0_{\Lambda_\gamma}(\eta_x \ge h) \le c\exp(-4\beta h)$ for some constant c independent of β. Moreover, using the exponential decay of correlations of the SOS measure without floor (cf. [12]), we obtain

$$\hat\pi^0_{\Lambda_\gamma}(\eta_x \ge h) \le \begin{cases} ce^{-4\beta h} & \text{if } \mathrm{dist}(x,\gamma) \le \varepsilon_\beta \log|\Lambda_\gamma|, \\ \hat\pi(\eta_x \ge h) + 1/|\Lambda_\gamma|^2 & \text{otherwise,} \end{cases}$$

with $\lim_{\beta\to\infty}\varepsilon_\beta = 0$. If we now use Lemma 2.4 to write $\hat\pi(\eta_x \ge h) = c_\infty(1+\delta_h)e^{-4\beta h}$ with $\lim_{h\to\infty}\delta_h = 0$, we get

$$\prod_{x\in\Lambda_\gamma}[1-\hat\pi^0_{\Lambda_\gamma}(\eta_x \ge h)] = \prod_{\substack{x\in\Lambda_\gamma\\ \mathrm{dist}(x,\gamma)\le\varepsilon_\beta\log|\Lambda_\gamma|}}[1-\hat\pi^0_{\Lambda_\gamma}(\eta_x \ge h)]$$
$$\times \prod_{\substack{x\in\Lambda_\gamma\\ \mathrm{dist}(x,\gamma)>\varepsilon_\beta\log|\Lambda_\gamma|}}[1-\hat\pi^0_{\Lambda_\gamma}(\eta_x \ge h)]$$
$$\ge \exp(-c\varepsilon_\beta e^{-4\beta h}|\gamma|\log|\Lambda_\gamma|)\exp(-c_\infty(1+\delta_h)e^{-4\beta h}|\Lambda_\gamma|),$$

possibly for slightly modified values of ε_β and δ_h. The proof is completed by using $|\Lambda_\gamma| \le |\gamma|^2/16$. □

Proof of (2.3). With the same notation as before (see proof of (2.2)) we write

$$\pi^j_\Lambda(\mathscr{C}^-_{\gamma,h})e^{-\beta|\gamma|}\,\frac{Z^{+,h}_{\rm out} Z^{-,h-1}_{\rm in}}{Z^j_V} \le e^{-\beta|\gamma|}\,\frac{Z^{+,h}_{\rm out} Z^{-,h-1}_{\rm in}}{Z^{+,h}_{\rm out} Z^{-,h}_{\rm in}} \le e^{-\beta|\gamma|}. \qquad \square$$

The next result is a simple geometric criterion to exclude certain large contours.

Lemma 2.9. *Fix $n \in \mathbb{Z}$ and consider the measure π_V^{h-1} of the SOS model in a finite connected subset V of $\mathbb{Z}^2$, with floor at height* 0 *and boundary conditions at height $h-1$ where $h := H(L) - n$. Let $c_0 = 2\log 3$. If*

$$|V| \le \left[\frac{4(\beta - c_0)(1 - e^{-4\beta})L}{\lambda^{(n)}}\right]^2, \tag{2.5}$$

then w.h.p. there are no macroscopic contours.

An immediate consequence of the above bound is that for any sufficiently large β one can exclude the existence of macroscopic $(H(L) + 1)$-contours.

Corollary 2.10. *Let Λ be the square of side-length L and let β be large enough. Then w.h.p. the SOS measure π_Λ^0 does not admit any $(H(L)+1)$-contours of length larger than* $(\log L)^2$.

Proof. The statement is just a special case of Lemma 2.9 with $n = -1$ and $V = \Lambda$. In this case the inequality (2.5) is obvious since $\lambda^{(-1)} = e^{-4\beta}\lambda \le c_\infty \le 2$ for β large. □

Proof of Lemma 2.9. The statement is an easy consequence of Proposition 2.7, applied with $j = h - 1$. Let us first show that w.h.p. there are no macroscopic h-contours. First of all we observe that the error term $\exp(\varepsilon_\beta e^{-4\beta h}|\gamma| \log|\gamma|)$ on the r.h.s. of (2.2) is at most $\exp(c' L^{-1}|\gamma| \log L)$ for some constant $c' = c'(\beta, n)$ because $|\gamma| \le c_1|V| \le c_2 L^2$. Hence it is negligible with respect to the main term $\exp(-\beta|\gamma| + c_\infty(1+\delta_h)|\Lambda_\gamma|e^{-4\beta h})$. If we now use the inequality $|\Lambda_\gamma| \le |V|^{1/2}|\gamma|/4$ (that just comes from $|\Lambda_\gamma| \le |V|$ and $|\gamma| \ge 4\sqrt{|\Lambda_\gamma|}$), we see immediately that, under the stated assumption on the cardinality of V, the area term satisfies

$$c_\infty(1 + \delta_h)|\Lambda_\gamma|e^{-4\beta h} = \frac{(1 + \delta_h)\lambda^{(n)}}{(1 - e^{-4\beta})L}|\Lambda_\gamma| \le (1 + \delta_h)(\beta - c_0)|\gamma|.$$

Hence the probability that a macroscopic h-contour exists can be bounded from above by

$$\sum_{\gamma:\, |\gamma| \ge (\log L)^2} e^{-(c_0 - c' L^{-1}\log L + \beta\delta_h)|\gamma|} = O(e^{-c(\log L)^2})$$

for some positive constant c. We have used the fact that h tends to infinity with L (at β fixed), so that $\beta\delta_h$ is negligible with respect to c_0, and the fact that the number of contours of length m is bounded by $\exp(am)$ for some absolute constant a. Clearly, if no macroscopic h-contour exists then there is no macroscopic j-contour for $j \ge h$. It remains to rule out macroscopic j-contours with $j \le h - 1$. However, the existence of such a contour would imply the existence of a negative macroscopic contour, and such an event has probability $O(e^{-c(\log L)^2})$ because of Proposition 2.7. □

Fix $n \in \mathbb{Z}$ (independent of L) and consider the SOS model in a finite connected subset V of $\mathbb{Z}^2$, with floor at height 0 and boundary conditions at height $h-1$ where $h := H(L)-n$.

Let $\partial_* V$ denote the set of $y \in V$ either at distance 1 from ∂V or at distance $\sqrt{2}$ from ∂V in the south-west or north-east direction. In particular, if V is the set Λ_γ corresponding to a contour γ, then $\partial_* V = \Delta_\gamma^+$. For a fixed $U \subset \partial_* V$, define the partition function $Z_{V,U}^{h-1,+}$ (resp. $Z_{V,U}^{h-1,-}$) of the SOS model on V with boundary conditions at height $h-1$ on ∂V, with floor at height 0 and with the further constraint that $\eta_y \geq h-1$ (resp. $\eta_y \leq h-1$) for all $y \in U$. We write $\hat{Z}_{V,U}^{h-1,\pm}$ for the same partition functions *without* the floor constraint. By translation invariance, $\hat{Z}_{V,U}^{h-1,\pm}$ does not depend on h. We let $\pi_{V,U}^{h-1,\pm}$ and $\hat{\pi}_{V,U}^{h-1,\pm}$ be the Gibbs measures associated to the partition functions $Z_{V,U}^{h-1,\pm}$ and $\hat{Z}_{V,U}^{h-1,\pm}$ respectively.

Remark 2.11. Exactly the same argument given above shows that Lemma 2.9 applies as is to the measures $\pi_{V,U}^{h-1,\pm}$ for any $U \subset \partial_* V$.

The next proposition quantifies the effect of the floor constraint.

Proposition 2.12. *In the above setting, fix $\varepsilon \in (0, 1/10)$ and assume that $|\partial V| \leq L^{1+\varepsilon}$. Then*

$$Z_{V,U}^{h-1,\pm} \geq \hat{Z}_{V,U}^{h-1,\pm} \exp(-c_\infty e^{-4\beta h}|V| + O(L^{1/2+2\varepsilon})). \tag{2.6}$$

If, in addition, (2.5) *holds, then*

$$Z_{V,U}^{h-1,\pm} \leq \hat{Z}_{V,U}^{h-1,\pm} \exp(-c_\infty e^{-4\beta h}|V| + O(L^{1/2+c(\beta)})), \tag{2.7}$$

where $c(\beta) \to 0$ as $\beta \to \infty$.

Remark 2.13. In Section 4 we will apply the above result to sets V with area of order L^2. In this case the error terms in (2.6)–(2.7) will be negligible (recall that $e^{-4\beta h} \propto L^{-1}$, since $h = H(L) - n$ with n fixed independent of L). In Section 5 we will instead apply it to sets with area of order $L^{4/3}$, and then it will be necessary to refine it and show that, in this case, the error term becomes $o(1)$.

The core of the argument is to show that, with respect to the measure $\hat{\pi}_{V,U}^{h-1,\pm}$, the Bernoulli variables $\{\mathbf{1}_{\eta_x \geq 0}\}_{x\in V}$ behave essentially as i.i.d. random variables with $\mathbb{P}(\mathbf{1}_{\eta_x \geq 0} = 1) \approx 1 - \hat{\pi}(\eta_0 \geq h)$ where $\hat{\pi}$ is the infinite volume SOS model without floor.

Proof of (2.6). From the FKG inequality (it is immediate to see that the modified model with partition function $Z_{V,U}^{h-1,\pm}$ still satisfies the FKG lattice condition)

$$\frac{Z_{V,U}^{h-1,\pm}}{\hat{Z}_{V,U}^{h-1,\pm}} = \hat{\pi}_{V,U}^{h-1,\pm}(\eta_x \geq 0\ \forall x \in V) \geq \prod_{x\in V} \hat{\pi}_{V,U}^{h-1,\pm}(\eta_x \geq 0).$$

At this point one can proceed exactly as in the proof of Proposition 2.7 (see (2.4) and its sequel). Indeed, using $|V| \leq |\partial V|^2 \leq L^{2+2\varepsilon}$, $\delta_h = O(e^{-2\beta h}) = O(L^{-1/2})$ (see Lemma 2.4) and $e^{-4\beta h} = O(L^{-1})$, one sees that

$$\max(\delta_h e^{-4\beta h}|V|, e^{-4\beta h}|\partial V| \log|V|) = O(L^{1/2+2\varepsilon}). \qquad \square$$

Proof of (2.7). The upper bound is more involved and it is here that the area constraint plays a role. Without it, the entropic repulsion could push up the whole surface and the product $\prod_{x\in V} \mathbf{1}_{\eta_x \geq 0}$ would no longer behave (under $\hat\pi^{h-1,\pm}_{V,U}$) as a product of i.i.d. variables.

Let $\mathcal{S}$ denote the event that there are no macroscopic contours, and use the identity

$$\hat\pi^{h-1,\pm}_{V,U}(\eta_x \geq 0\ \forall x \in V) = \frac{\hat\pi^{h-1,\pm}_{V,U}(\mathcal{S})}{\pi^{h-1,\pm}_{V,U}(\mathcal{S})}\, \hat\pi^{h-1,\pm}_{V,U}(\eta_x \geq 0\ \forall x \in V \mid \mathcal{S}).$$

Thanks to Lemma 2.9 (see Remark 2.11) and our area constraint, one has $\pi^{h-1,\pm}_{V,U}(\mathcal{S}) = 1 - o(1)$. Hence, it is enough to show that

$$\hat\pi^{h-1,\pm}_{V,U}(\eta_x \geq 0\ \forall x \in V \mid \mathcal{S}) \leq \exp\bigl(-\hat\pi(\eta_0 \geq h)|V| + O(L^{1/2+c(\beta)})\bigr), \tag{2.8}$$

where $c(\beta)$ is a constant that can be made small if β is large, since then one can appeal to Lemma 2.4 to write $\hat\pi(\eta_0 \geq h)|V| = c_\infty e^{-4\beta h}|V| + O(L^{-3/2}|V|)$. The estimate (2.8) has been essentially already proved in [15, Section 7]. For the reader's convenience, we give the details in Appendix A.3. □

Remark 2.14. For technical reasons, later in the proofs we will need Proposition 2.7, Lemma 2.9 and Proposition 2.12 in a slightly more general case, referred to as the "partial floor setting", in which the SOS model in V has the floor constraint $\eta_x \geq 0$ only for those vertices x inside a certain subset W of V. Exactly the same proofs show that in this new setting the very same statements hold with Λ_γ replaced by $|\Lambda_\gamma \cap W|$ in (2.2), and with $|V|$ replaced by $|V \cap W|$ in (2.5) and in the exponent at (2.6)–(2.7).

We conclude by describing a monotonicity trick to upper bound the probability of an increasing event A, under the SOS measure π^0_Λ in some domain Λ with boundary conditions at height zero and floor at height zero.

Lemma 2.15 (Domain-enlarging procedure). *Let $\Lambda \subset \Lambda'$, let $V \subset \Lambda'$ such that $(\Lambda \cup \partial\Lambda) \cap \Lambda' \subset V$, let τ be nonnegative (but otherwise arbitrary) boundary conditions on $\partial\Lambda'$ and let $\pi^\tau_{\Lambda',V}$ denote the SOS measure on Λ' with boundary conditions τ and floor at height zero in V. Let A be an increasing event in Ω_Λ. Then*

$$\pi^0_\Lambda(A) \leq \pi^\tau_{\Lambda',V}(A). \tag{2.9}$$

Proof. Note first of all that $\pi^{\tau'}_{\Lambda',V}(A) \leq \pi^\tau_{\Lambda',V}(A)$ where τ' is obtained from τ by setting $\tau'_x = 0$ for every $x \in \partial\Lambda' \cap \partial\Lambda$. Then π^0_Λ can be seen as the marginal in Λ of the measure $\pi^{\tau'}_{\Lambda',V}$ conditioned on the decreasing event that $\eta = 0$ on $\partial\Lambda$. By FKG, removing the conditioning can only increase the probability of A. □

2.4. Cluster expansion. In order to write down precisely the law of certain macroscopic contours we shall use a cluster expansion for partition functions of the SOS with partial or no floor. Given a finite connected set $V \subset \mathbb{Z}^2$ and $U \subset \partial_* V$ (the set $\partial_* V$ has been defined before Remark 2.11), we write $\hat{Z}_{V,U}$ for the SOS partition function with the sum over η restricted to those $\eta \in \hat{\Omega}_V^0$ such that $\eta_x \geq 0$ for all $x \in U$. Notice that $\hat{Z}_{V,U}$ coincides with the partition function $\hat{Z}_{V,U}^{h,+}$ appearing in Proposition 2.12 (the latter does not depend on h). We refer the reader to [15, Appendix A] for a proof of the following expansion.

Lemma 2.16. *There exists $\beta_0 > 0$ such that for all $\beta \geq \beta_0$, for all finite connected $V \subset \mathbb{Z}^2$ and $U \subset \partial_* V$,*

$$\log \hat{Z}_{V,U} = \sum_{V' \subset V} \varphi_U(V'), \tag{2.10}$$

where the potentials $\varphi_U(V')$ satisfy:

(i) $\varphi_U(V') = 0$ *if V' is not connected.*

(ii) $\varphi_U(V') = \varphi_0(V')$ *if* $\operatorname{dist}(V', U) \neq 0$ *for some shift invariant potential $V' \mapsto \varphi_0(V')$, that is,*

$$\varphi_0(V') = \varphi_0(V' + x) \quad \forall x \in \mathbb{Z}^2.$$

(iii) *For all $V' \subset V$,*

$$\sup_{U \subset \partial_* V} |\varphi_U(V')| \leq \exp(-(\beta - \beta_0)\, d(V'))$$

where $d(V')$ is the cardinality of the smallest connected set of bonds of the dual lattice $(\mathbb{Z}^2)^*$ separating points of V' from points of its complement.

3. Surface tension and variational problem

In this section we first collect all the necessary information about surface tension and associated Wulff shapes. We then consider the variational problem of maximizing a certain functional which will play a key role in our main results, and describe its solution.

We begin by defining the surface tension of the SOS model *without* the wall (see also Appendix A.4). We assume β is large enough in order to enable cluster expansion techniques [12, 20].

Definition 3.1. Let $\Lambda_{n,m} = \{-n, \dots, n\} \times \{-m, \dots, m\}$ and let $\xi(\theta)$, $\theta \in [0, \pi/2)$, be the boundary conditions given by

$$\xi(\theta)_y = \begin{cases} +1 & \text{if } \vec{n} \cdot y \geq 0, \\ 0 & \text{if } \vec{n} \cdot y < 0, \end{cases} \qquad \forall y \in \partial \Lambda_{n,m}$$

where $\vec{n}$ is the unit vector orthogonal to the line forming an angle θ with the horizontal axis.

The *surface tension* $\tau(\theta)$ in the direction θ is defined by

$$\tau(\theta) = \lim_{n \to \infty} \lim_{m \to \infty} -\frac{\cos\theta}{2\beta n} \log \frac{\hat{Z}^{\xi(\theta)}_{\Lambda_{n,m}}}{\hat{Z}^0_{\Lambda_{n,m}}}. \tag{3.1}$$

Using the symmetry of the SOS model we finally extend τ to an even, $\pi/2$-periodic function on $[0, 2\pi]$. Finally, if one extends $\tau(\cdot)$ to $\mathbb{R}^2$ as $x \mapsto \tau(x) := |x|\tau(\theta_x)$, θ_x being the direction of x, then $\tau(\cdot)$ becomes (strictly) convex and analytic. See [20, Ch. 1 and 2] for additional information,[3] and Appendix A.4 for an equivalent definition of $\tau(\cdot)$ in the cluster expansion language.

Next we proceed to define the Wulff shape.

Definition 3.2. Given a closed rectifiable curve γ in $\mathbb{R}^2$, let $A(\gamma)$ be the area of its interior and let $W(\gamma)$ be the Wulff functional $\gamma \mapsto \int_\gamma \tau(\theta_s)\, ds$, with θ_s the direction of the normal with respect to the curve γ at the point s and ds the length element. The convex body with support function $\tau(\cdot)$ (see e.g. [21]) is denoted by $\mathcal{W}_\tau$. The rescaled set

$$\mathcal{W}_1 = \sqrt{\frac{2}{W(\partial \mathcal{W}_\tau)}}\,\mathcal{W}_\tau$$

is called the *Wulff shape* and it has unit area (see e.g. [20, Ch. 2]). $\mathcal{W}_1$ is also the subset of $\mathbb{R}^2$ of unit area that minimizes the Wulff functional. We set $w_1 := W(\partial \mathcal{W}_1)$.

Now, given $\lambda > 0$, consider the problem of maximizing the functional

$$\gamma \mapsto \mathcal{F}_\lambda(\gamma) := -\beta W(\gamma) + \lambda A(\gamma) \tag{3.2}$$

among all curves contained in the square $Q = [0, 1] \times [0, 1]$. In order to solve this variational problem we proceed as follows.

We first observe that if ℓ_τ denotes the side of the smallest square with sides parallel to the coordinate axes into which $\mathcal{W}_1$ can fit, then

$$\ell_\tau = 2\sqrt{\frac{2}{W(\partial \mathcal{W}_\tau)}}\,\tau(0) = 4\frac{\tau(0)}{w_1}. \tag{3.3}$$

Remark 3.3. As β tends to ∞, one has $\tau(\theta) \to |\cos\theta| + |\sin\theta|$ (analyticity is lost in this limit) and the Wulff shape converges to the unit square.

We now set

$$\hat\lambda = 2\beta\tau(0), \qquad \ell_c(\lambda) = \beta w_1/2\lambda. \tag{3.4}$$

Definition 3.4. For r, t, λ such that $0 < t\ell_c\ell_\tau \le 1$ and $r \in (-1, 1)$ we define the convex body $\mathcal{L}(\lambda, t, r)$ as the $(1 + r)$-dilation of the set formed by the union of all possible translates of $t\ell_c\mathcal{W}_1$ contained inside Q. When $t = 1$ and $r = 0$ we write $\mathcal{L}_c(\lambda)$ for $\mathcal{L}(\lambda, 1, 0)$.

Remark 3.5. We point out two useful properties of the parameters ℓ_c and $\hat\lambda$. The first one is that, by construction, the rescaled droplet $\ell_c\mathcal{W}_1$ can fit inside the unit square Q iff $\lambda \ge \hat\lambda$. The second one, proved in Section 6.1, goes as follows. Consider the SOS model with floor in a box of side L with boundary conditions at height zero and assume the existence of an $(H(L)-n)$-contour containing the rescaled Wulff body $L\ell_c(\lambda^{(n)})\mathcal{W}_1$. Necessarily that requires $\lambda^{(n)} \ge \hat\lambda$. Then w.h.p. the $(H(L) - n)$-contour actually contains the whole region $L\mathcal{L}_c(\lambda^{(n)})$ up to $o(L)$ corrections.

[3] Strictly speaking, [20] deals with the nearest-neighbor two-dimensional Ising model, but their proofs immediately extend to our case. Also in the following, whenever a result of [20] can be adapted straightforwardly to our context, we just cite the relevant chapter without an explicit caveat.

Claim 3.6. *Set*

$$\lambda_c = \inf\{\lambda \geq \hat{\lambda} : \mathcal{F}_\lambda(\mathcal{L}_c(\lambda)) > 0\}. \tag{3.5}$$

Then $\lambda_c = \hat{\lambda} + \beta w_1/2$.

Proof. Using the definitions of ℓ_c, $\mathcal{L}_c$ and ℓ_τ, we can write

$$W(\mathcal{L}_c(\lambda)) = \ell_c w_1 + 4\tau(0)(1 - \ell_c \ell_\tau) = \frac{\beta}{2\lambda} w_1^2 + 4\tau(0)\left(1 - 2\beta \frac{\tau(0)}{\lambda}\right),$$
$$A(\mathcal{L}_c(\lambda)) = 1 + \frac{\beta^2 w_1^2}{4\lambda^2} - \frac{4\beta^2 \tau(0)^2}{\lambda^2}.$$

Hence

$$\mathcal{F}_\lambda(\mathcal{L}_c(\lambda)) = -4\beta\tau(0) + \lambda - \frac{\beta^2 w_1^2}{4\lambda} + 4\frac{\beta^2 \tau(0)^2}{\lambda}.$$

Solving the quadratic equation $\mathcal{F}_\lambda(\mathcal{L}_c(\lambda)) = 0$ gives the solutions

$$\lambda_\pm = 2\beta\tau(0) \pm \beta w_1/2 = \hat{\lambda} \pm \beta w_1/2.$$ □

Remark 3.7. In the limit $\beta \to \infty$ we have $\hat{\lambda}/\beta \to 2$, $\lambda_c/\beta \to 4$ and $\ell_c(\lambda_c) \to 1/2$.

Going back to the variational problem of maximizing $\mathcal{F}_\lambda(\gamma)$, the following holds (cf. [36, Sects. 3 and 4, in particular Lemma 1]):

Claim 3.8. (i) *If* $\lambda < \lambda_c$ *then the supremum of* $\mathcal{F}_\lambda(\gamma)$ *corresponds to a sequence of curves* γ_n *that shrinks to a point, so that* $\sup_\gamma \mathcal{F}_\lambda(\gamma) = 0$*; moreover for any* $\delta > 0$ *there exists* $\epsilon > 0$ *such that* $\mathcal{F}_\lambda(\gamma) \leq -\epsilon$ *for any curve* γ *enclosing an area larger than* δ.

(ii) *If* $\lambda > \lambda_c$ *then the maximum is attained for* $\gamma = \partial\mathcal{L}_c(\lambda)$, *and* $\mathcal{F}_\lambda(\partial\mathcal{L}_c(\lambda)) > 0$.

The area (or perimeter) of the optimal curve has therefore a discontinuity at λ_c.

We conclude with a last observation on the geometry of the Wulff shape $\mathcal{W}_1$ which will be important in the proof of Theorems 2 and 3.

Lemma 3.9. *Fix* $\theta \in [-\pi/4, \pi/4]$ *and* $d \ll 1$. *Let* $I(d, \theta)$ *be the segment of length* d *and angle* θ *with respect to the* x*-axis such that its endpoints lie on the boundary of* $\mathcal{W}_1$. *Let* $\Delta(d, \theta)$ *be the vertical distance between the midpoint of* $I(d, \theta)$ *and* $\partial\mathcal{W}_1$. *Then*

$$\Delta(d, \theta) = \frac{w_1}{16(\tau(\theta) + \tau''(\theta))\cos\theta} d^2(1 + O(d^2)) \quad \text{as } d \to 0.$$

Proof. Let x be the midpoint of $I(d, \theta)$ and let h be the distance between x and $\partial\mathcal{W}_1$. Clearly $\Delta(d, \theta) = \frac{h}{\cos\theta}(1 + O(h))$. From elementary considerations, as $d \to 0$,

$$h = \frac{d^2}{8R(\theta)}(1 + O(d^2))$$

where $R^{-1}(\theta)$ is the curvature of the Wulff shape $\mathcal{W}_1$ at angle θ. It is known (see, e.g., [2, Sec. 5]) that

$$R(\theta) = \sqrt{\frac{2}{W(\partial\mathcal{W}_\tau)}}(\tau(\theta) + \tau''(\theta)) = \frac{2}{w_1}(\tau(\theta) + \tau''(\theta))$$

where we have used also (3.3). □

4. Proof of Theorem 1

4.1. An intermediate step: existence of a supercritical $(H(L)-1)$-contour. Our first goal is to show that w.h.p. there exists a large droplet at level $H(L)-1$.

Proposition 4.1. *Let Λ be a square of side-length L. If β is large enough, the SOS measure π_Λ^0 admits an $(H(L)-1)$-contour γ whose interior contains a square of side-length $\frac{9}{10}L$ w.h.p.*

Proof. The first ingredient is a bound addressing the contribution of microscopic contours to the height profile.

Lemma 4.2. *Let $V\subset\Lambda$ where Λ is a square of side-length L with boundary condition $\xi\le h-1$, where $h=H(L)-n$ for some fixed $n\ge 0$. Denote by $\mathscr{B}_h$ the event that there is no h-contour of length at least $(\log L)^2$. Then for any $\delta>0$ there are constants $C_1,C_2>0$ such that for any $\beta\ge C_1$,*

$$\pi_\Lambda^\xi(\#\{v:\eta_v\ge h\}>\delta L^2,\mathscr{B}_h)\le\exp(-C_2(\log L)^2),\tag{4.1}$$

and for any closed contour γ (see Definition 2.2 for $\mathscr{C}_{\gamma,h}$),

$$\pi_\Lambda^\xi(\#\{v\in\Lambda_\gamma:\eta_v\le h-1\}>\delta L^2\mid\mathscr{C}_{\gamma,h})\le\exp(-C_2(\log L)^2).\tag{4.2}$$

Proof. For a configuration η let $\mathcal{N}_k(\eta)$ denote the number of h-contours of length $k\le(\log L)^2$. As there are at most $L^2 4^k$ possible such contours, Proposition 2.7 shows that for some constant $C_0>0$, for any m,

$$\begin{aligned}\pi_\Lambda^\xi(\mathcal{N}_k(\eta)\ge m)&\le\sum_{r\ge m}\binom{4^kL^2}{r}e^{r(-\beta k+C_0e^{-4\beta h}k^2)}\le\sum_{r\ge m}\binom{4^kL^2}{r}e^{-r\beta k/2}\\&\le\frac{\mathbb{P}(\mathrm{Bin}(4^kL^2,e^{-\beta k/2})\ge m)}{(1-e^{-\beta k/2})^{4^kL^2}}\\&\le\exp(2e^{-\beta k/2}4^kL^2)\mathbb{P}(\mathrm{Bin}(4^kL^2,e^{-\beta k/2})\ge m),\end{aligned}$$

where we have used the fact that $1-x\ge e^{-2x}$ for $0\le x\le 1/2$ as well as that $e^{-\beta k/2}\le 1/2$ for β large. For each $1\le k\le(\log L)^2$ we now wish to apply the above inequality for a choice of

$$m(k)=7\cdot 4^kL^2e^{-\beta k/2}+(\log L)^2.$$

By the well-known fact that $\mathbb{P}(X\ge\mu+t)\le\exp[-t^2/(2(\mu+t/3))]$ for any $t>0$ and binomial variable X with mean μ, which in our setting of $t\ge 6\mu$ implies a bound of $\exp(-t)$, we get

$$\pi_\Lambda^\xi(\mathcal{N}_k(\eta)\ge m)\le\exp(-4e^{-\beta k/2}4^kL^2-(\log L)^2)\le e^{-(\log L)^2}.$$

Each h-contour counted by $\mathcal{N}_k(\eta)$ encapsulates at most k^2 sites of height larger than h, thus setting $M(L)=\sum_{k=1}^{(\log L)^2}k^2m(k)$ we get

$$\pi_\Lambda^\xi(\#\{v:\eta_v\ge h\}>M(L),\mathscr{B}_h)\le e^{-(1-o(1))(\log L)^2}.$$

The proof is concluded by the fact that $M(L) = O(e^{-\beta/2}L^2) + L^{1+o(1)}$ for any β large enough, where the $O(L^2)$-term is easily seen to be less than δL^2 for large enough β.

To prove (4.2), observe that by monotonicity

$$\pi_\Lambda^\xi(\#\{v \in \Lambda_\gamma : \eta_v \le h-1\} > \delta L^2 \mid \mathscr{C}_{\gamma,h}) \le \pi_{\Lambda_\gamma}^h(\#\{v \in \Lambda_\gamma : \eta_v \le h-1\} > \delta L^2)$$

(on the l.h.s. heights on Δ_γ^+ are constrained to be $\ge h$ while on the r.h.s. they are lowered to exactly h; this makes the heights in Λ_γ independent of those outside Λ_γ). Thus, if no large negative contours are present, the argument above for (4.1) will imply (4.2). On the other hand, Proposition 2.7 and a simple Peierls bound immediately imply that w.h.p. there exists no macroscopic negative contour. □

We now need to introduce the notion of external h-contours.

Definition 4.3. Given a configuration $\eta \in \Omega_\Lambda$ we say that γ is an *external h-contour* of η if γ is a macroscopic h-contour and there exists no other h-contour γ' containing it. We say that $\{\gamma_i\}_{i=1}^n$ forms the *collection of external h-contours of η* if $\{\gamma_i\}$ is the set of all external h-contours.

With this notation we have

Lemma 4.4. *Let $h = H(L) - 1$ and $\delta > 0$. If β is sufficiently large then the collection $\{\gamma_i\}$ of external h-contours satisfies*

$$\pi_\Lambda^0\Big(\sum_i |\gamma_i| \le (1+\delta)4L\Big) \ge 1 - e^{-\beta\delta L/2}. \tag{4.3}$$

Proof. Let $A = \bigcup \Lambda_{\gamma_i}$ and let $R = \sum_i |\gamma_i|$. Let $U_A : \Omega \to \Omega$ denote the map that increases each $v \notin A$ by 1 (retaining the remaining configuration as is). Then U_A increases the Hamiltonian by at most $|\partial\Lambda| - R$ and so

$$\pi_\Lambda^0(U_A\eta) \ge \exp(-4\beta L + \beta R)\pi_\Lambda^0(\eta).$$

Since U_A is bijective, the probability of having a given configuration $\{\gamma_i\}$ of external contours is bounded by $e^{-\beta(R-4L)}$. Given $R = \ell$, the number of possible external contours is at most $\ell/(\log L)^2$, and the number of their arrangements is easily bounded from above by C^ℓ for some constant $C > 0$, for L large enough. Therefore, if we sum over configurations for which $R \ge (1+\delta)4L$, we obtain

$$\pi_\Lambda^0(R \ge (1+\delta)4L) \le \sum_{\ell \ge (1+\delta)4L} C^\ell e^{-\beta(\ell-4L)} \le e^{-\beta\delta L/2}$$

for large enough β. □

The next ingredient in the proof of Proposition 4.1 is to establish that most of the sites have height at least $H(L) - 1$ with high probability.

Lemma 4.5. *Let Λ be a square of side-length L. For any $\delta > 0$ there exist constants $C_1, C_2 > 0$ such that for any $\beta \ge C_1$,*

$$\pi_\Lambda^0(\#\{v : \eta_v \le H(L) - 2\} > \delta L^2) \le \exp(-C_2 L).$$

Proof. Let $\mathcal{S}_h(\eta) = \{v \in \Lambda : \eta_v = h\}$ for $h = H(L) - k$. Define $U_A : \Omega \to \Omega$ for each $A \subseteq \mathcal{S}_h(\eta)$ as

$$(U_A\eta)_v = \begin{cases} \eta_v + 1, & v \notin A, \\ 0, & v \in A. \end{cases}$$

Since U_A is equivalent to increasing each height by 1 followed by decreasing the sites in A by $h+1$, the Hamiltonian is increased by at most $|\partial\Lambda| + 4(h+1)|A|$ and so

$$\pi_\Lambda^0(U_A\eta) \ge \exp(-4\beta L - 4\beta(h+1)|A|)\pi_\Lambda^0(\eta).$$

Therefore,

$$\begin{aligned}\sum_{A\subseteq \mathcal{S}_h(\eta)} \pi_\Lambda^0(U_A\eta) &\ge \exp(-4\beta L)(1+e^{-4\beta(h+1)})^{|\mathcal{S}_h(\eta)|}\pi_\Lambda^0(\eta), \\ &\ge \exp\big(-4\beta L + \tfrac{1}{2}e^{-4\beta(h+1)}|\mathcal{S}_h(\eta)|\big)\pi_\Lambda^0(\eta),\end{aligned}$$

where we have used $1 + x \ge e^{x/2}$ for $x \in (0,1]$. By definition $U_A\eta \ne U_{A'}\eta$ for any $A \ne A'$ with $A, A' \subseteq \mathcal{S}_h(\eta)$. In addition, if $A \subseteq \mathcal{S}_h(\eta)$ and $A' \subseteq \mathcal{S}_h(\eta')$ for some $\eta \ne \eta'$ then $U_A\eta \ne U_{A'}\eta'$ (one can read the set A from $U_A\eta$ by looking at the sites at level 0, and then proceed to reconstruct η). Using the fact that $e^{-4\beta(h+1)} \ge e^{4\beta(k-1)}/L$ we see that

$$\begin{aligned}1 &\ge \sum_{\eta:\, |\mathcal{S}_h(\eta)| \ge \delta e^{-2\beta(k-1)}L^2}\ \sum_{A\subseteq \mathcal{S}_h(\eta)} \pi_\Lambda^0(U_A\eta) \\ &\ge \exp\big(-4\beta L + \tfrac{1}{2}\delta e^{2\beta(k-1)}L\big)\pi_\Lambda^0(|\mathcal{S}_h(\eta)| \ge \delta e^{-2\beta(k-1)}L^2),\end{aligned}$$

and so, for $k \ge 1$,

$$\pi_\Lambda^0(|\mathcal{S}_h(\eta)| \ge \delta e^{-2\beta(k-1)}L^2) \le \exp\big(4\beta L - \tfrac{1}{2}\delta e^{2\beta(k-1)}L\big).$$

Summing over $k \ge 2$ establishes the required estimate for any sufficiently large β. □

We now complete the proof of Proposition 4.1. Fix $0 < \delta \ll 1$. By Lemma 4.5, the number of sites with height less than $H(L) - 1$ is at most δL^2. Condition on the external macroscopic $(H(L)-1)$-contours $\{\gamma_i\}$ and consider the region obtained by deleting those contours as well as their interiors and immediate external neighborhood, i.e., $V = \Lambda \setminus \bigcup_i(\Lambda_{\gamma_i} \cup \Delta_{\gamma_i}^-)$. An application of Lemma 4.2 to π_V^ξ where ξ is the boundary condition induced by $\partial\Lambda$ and $\{\gamma_i\}$ (in particular at most $H(L) - 1$ everywhere) shows that w.h.p. there are at most δL^2 sites of height larger than $H(L) - 1$ in V. Altogether,

$$\sum_i |\Lambda_{\gamma_i}| \ge (1 - 2\delta)L^2,$$

and therefore, by an application of Lemma 4.4 followed by Lemma 2.6, we can conclude that w.h.p. one of the γ_i contains a square with side-length at least $\frac{9}{10}L$ as required. □

4.2. Absence of macroscopic $H(L)$-contours when $\lambda < \lambda_c$. In this section we prove:

Proposition 4.6. *Consider the SOS measure π_Λ^0 on the box $\Lambda = \Lambda_L$ with floor, at inverse-temperature β. Fix $\delta > 0$ and assume that $\lambda < \lambda_c - \delta$. W.h.p., there are no macroscopic $H(L)$-contours.*

Proof. The strategy of the proof is the following:

- Step 1: Via a simple isoperimetric argument, we show that if a macroscopic $H(L)$-contour exists, then it must contain a square of area almost L^2.
- Step 2: Using the "domain-enlarging procedure" (see Lemma 2.15) we reduce the proof of the nonexistence of a macroscopic $H(L)$-contour as in Step 1 to the proof of the same fact in a larger square Λ' of side $5L$ with boundary conditions at height $H(L) - 1$. That allows us to avoid any pinning issues with the boundary of the original square Λ. Using Proposition 2.12 we write precisely the law of such a contour (assuming it exists) and we show that it satisfies a certain "regularity property" w.h.p.
- Step 3: Using the exact form of the law of the macroscopic $H(L)$-contour in Λ' we are able to bring in the functional $\mathcal{F}_\lambda$ defined in Section 3 and to show, via a precise area vs. surface tension comparison, that the probability that an $H(L)$-contour contains such a square is exponentially (in L) unlikely. This implies that no macroscopic $H(L)$-contour exists and Proposition 4.6 is proven.

For lightness of notation, throughout this section we will write h for $H(L)$.

Step 1. We apply Proposition 2.7 with $V = \Lambda$ and $j = 0$. Noting that $e^{-4\beta h} \log|\gamma| = O((\log L)/L)$ and recalling (2.1) defining λ, we have

$$\pi_\Lambda^0(\mathcal{C}_{\gamma,h}) \le \exp\left(-(\beta + o(1))|\gamma| + (1 + o(1))\frac{\lambda}{L(1 - e^{-4\beta})}|\Lambda_\gamma|\right) \tag{4.4}$$

where $o(1)$ vanishes with L. This has two easy consequences. From $|\Lambda_\gamma| \le L^2$ we see that w.h.p. there are no h-contours γ with $|\gamma| \ge a_1 L := (1 + \epsilon_\beta)L\lambda/\beta$. Here and in the following, ϵ_β denotes some positive constant (not necessarily the same at each occurrence) that vanishes for $\beta \to \infty$ and does not depend on δ. From $|\Lambda_\gamma| \le |\gamma|^2/16$ (isoperimetry) together with standard Peierls counting of contours we see that w.h.p. there are no h-contours γ with

$$(\log L)^2 \le |\gamma| \le a_2 L := \frac{16}{\lambda}\beta L(1 - \epsilon_\beta). \tag{4.5}$$

If $\lambda < 4\beta(1 - \epsilon_\beta)$ then $a_1 < a_2$ and we have excluded the occurrence of h-contours longer than $(\log L)^2$—Proposition 4.6 is proven. The remaining case is

$$4\beta(1 - \epsilon_\beta) \le \lambda < \lambda_c - \delta \tag{4.6}$$

and it remains to exclude h-contours γ with

$$\frac{16}{\lambda}\beta L(1 - \epsilon_\beta) \le |\gamma| \le L(1 + \epsilon_\beta)\frac{\lambda}{\beta}. \tag{4.7}$$

Recall from Remark 3.7 that $\lambda_c/4\beta$ tends to 1 for β large so that under condition (4.6) we have $4\beta(1 - \epsilon_\beta) \le \lambda \le 4\beta(1 + \epsilon_\beta)$. Then condition (4.7) implies

$$4L(1 - \epsilon_\beta) \le |\gamma| \le 4L(1 + \epsilon_\beta).$$

For all such γ, (4.4) implies that $\mathcal{C}_{\gamma,h}$ is extremely unlikely, unless $|\Lambda_\gamma| \geq L^2(1-\epsilon_\beta)$. But, as in the proof of Lemma 2.6, a contour in Λ that has perimeter at most $4L(1+\epsilon_\beta)$ and encloses an area of at least $L^2(1-\epsilon_\beta)$ necessarily contains a square of area $(1-\epsilon_\beta)L^2$, for a different value of ϵ_β.

Step 2. We are left with the task of proving

$$\pi^0_\Lambda(A) := \pi^0_\Lambda(\exists\, h\text{-contour containing a square } Q \subset \Lambda \text{ with area } (1-\epsilon_\beta)L^2) \leq e^{-c(\log L)^2}. \tag{4.8}$$

Observe that the event A is increasing. We apply Lemma 2.15 with $V = \Lambda \cup \partial\Lambda$, Λ' a square of side $5L$ and concentric to Λ and boundary condition $h-1$, to write $\pi^0_\Lambda(A) \leq \pi^{h-1}_{\Lambda',V}(A)$. From now on, for lightness of notation, we write $\tilde\pi^{h-1}_{\Lambda'}$ instead of $\pi^{h-1}_{\Lambda',V}$.

Let γ denote a contour enclosing a square $Q \subset \Lambda$ of area $(1-\epsilon_\beta)L^2$. As in the proof of (2.2),

$$\tilde\pi^{h-1}_{\Lambda'}(\mathcal{C}_{\gamma,h}) = e^{-\beta|\gamma|}\frac{Z^{-,h-1}_{\text{out}}Z^{+,h}_{\text{in}}}{\tilde Z^{h-1}_{\Lambda'}}. \tag{4.9}$$

Here, $\tilde Z^{h-1}_{\Lambda'}$ is the partition function corresponding to the Gibbs measure $\tilde\pi^{h-1}_{\Lambda'}$. In the partition functions $Z^{-,h-1}_{\text{out}}$, $Z^{+,h}_{\text{in}}$, and $\tilde Z^{h-1}_{\Lambda'}$, the floor constraint, imposing nonnegative heights in $\Lambda \cup \partial\Lambda$, is implicit.

Now we can apply Proposition 2.12 (see also Remark 2.13) to the two partition functions in the numerator. For $Z^{+,h}_{\text{in}}$, we have $V = \Lambda_\gamma$ (as usual Λ_γ is the interior of γ and $\Lambda^c_\gamma = \Lambda' \setminus \Lambda_\gamma$), $U = \partial_*\Lambda_\gamma$ and $n = -1$ (recall that $\lambda^{(n)} = \lambda e^{4\beta n}$ and λ is around 4β by (4.6)). Since

$$|\Lambda_\gamma \cap \Lambda| \leq L^2 \ll \left(\frac{4\beta L}{\lambda e^{-4\beta}}\right)^2 \approx L^2 e^{8\beta},$$

condition (2.5) is satisfied, and for some $a \in (0,1)$ we have[4]

$$Z^{+,h}_{\text{in}} = \hat Z^{+,h}_{\text{in}} \exp\left[-\frac{c_\infty}{L}e^{4\beta\alpha(L)}e^{-4\beta}|\Lambda_\gamma \cap \Lambda| + O(L^a)\right]. \tag{4.10}$$

To expand $Z^{-,h-1}_{\text{out}}$ we apply the same argument on the region $\Lambda' \setminus \Lambda_\gamma$. Since by assumption γ contains a square $Q \subset \Lambda$ with area $(1-\epsilon_\beta)L^2$, we have

$$|\Lambda \setminus \Lambda_\gamma| \leq \epsilon_\beta L^2 \ll \left(\frac{4\beta}{\lambda}L\right)^2 \approx L^2.$$

Therefore,

$$Z^{-,h-1}_{\text{out}} = \hat Z^{-,h-1}_{\text{out}} \exp\left[-\frac{c_\infty}{L}e^{4\beta\alpha(L)}|\Lambda \setminus \Lambda_\gamma| + O(L^a)\right]. \tag{4.11}$$

[4] In principle we should have $|\Lambda_\gamma \cap (\Lambda \cup \partial\Lambda)|$ instead of $|\Lambda_\gamma \cap \Lambda|$, but since $|\partial\Lambda|/L = O(1)$, the difference can be absorbed into the error $O(L^a)$.

As for the denominator $\tilde{Z}^{h-1}_{\Lambda'}$, via (2.6) we get

$$\tilde{Z}^{h-1}_{\Lambda'} \geq \hat{Z}^{h-1}_{\Lambda'} \exp\left[-\frac{c_\infty}{L} e^{4\beta\alpha(L)} |\Lambda| + O(L^a)\right]. \tag{4.12}$$

Putting together (4.10)–(4.12) and recalling that $\lambda = c_\infty e^{4\beta\alpha(L)}(1 - e^{-4\beta})$, we get

$$\tilde{\pi}^{h-1}_{\Lambda'}(\mathcal{C}_{\gamma,h}) \leq e^{-\beta|\gamma|} \frac{\hat{Z}^{-,h-1}_{\text{out}} \hat{Z}^{+,h}_{\text{in}}}{\hat{Z}^{h-1}_{\Lambda'}} \exp\left[\frac{\lambda}{L}|\Lambda \cap \Lambda_\gamma| + O(L^a)\right]. \tag{4.13}$$

Finally, the partition functions $\hat{Z}^{-,h-1}_{\text{out}}$, $\hat{Z}^{+,h}_{\text{in}}$ and $\hat{Z}^{h-1}_{\Lambda'}$ can be expanded using Lemma 2.16. The net result is that

$$\frac{\hat{Z}^{-,h-1}_{\text{out}} \hat{Z}^{+,h}_{\text{out}}}{\hat{Z}^{h-1}_{\Lambda'}} = \exp(\Psi_{\Lambda'}(\gamma)) \tag{4.14}$$

where, for every $V \subset \mathbb{Z}^2$ and γ contained in V,

$$\Psi_V(\gamma) = -\sum_{\substack{W \subset V \\ W \cap \gamma \neq \emptyset}} \varphi_0(W) + \sum_{\substack{W \subset \Lambda_\gamma \\ W \cap \gamma \neq \emptyset}} \varphi_{\Delta^+_\gamma}(W) + \sum_{\substack{W \subset V \setminus \Lambda_\gamma \\ W \cap \gamma \neq \emptyset}} \varphi_{\Delta^-_\gamma}(W) \tag{4.15}$$

(see also [15, App. A.3]). Here the notation $W \cap \gamma \neq \emptyset$ means $W \cap (\Delta^+_\gamma \cup \Delta^-_\gamma) \neq \emptyset$.

Altogether, we have obtained

$$\tilde{\pi}^{h-1}_{\Lambda'}(\mathcal{C}_{\gamma,h}) = \exp\left[-\beta|\gamma| + \Psi_{\Lambda'}(\gamma) + \frac{\lambda}{L}|\Lambda \cap \Lambda_\gamma| + O(L^a)\right]. \tag{4.16}$$

Let Σ denote the collection of all possible contours that enclose a square $Q \subset \Lambda$ with area $(1 - \epsilon_\beta)L^2$.

A first observation is that the event that there exists an h-contour $\gamma \in \Sigma$ at distance less than $(\log L)^2$ from $\partial\Lambda'$ (the boundary of the square of side $5L$) has negligible probability. Indeed, such contours have necessarily $|\gamma| \geq 5L$. Then the area term $\lambda|\Lambda \cap \Lambda_\gamma|/L \leq \lambda L \approx 4\beta L$ cannot compensate for $-\beta|\gamma|$, and from the properties of the potentials φ in Lemma 2.16, we see that $|\Psi_{\Lambda'}(\gamma)| \leq \epsilon_\beta|\gamma|$. As a consequence, we can safely replace $\Psi_{\Lambda'}(\gamma)$ with $\Psi_{\mathbb{Z}^2}(\gamma)$ in (4.16): indeed, thanks to Lemma 2.16(iii),

$$|\Psi_{\Lambda'}(\gamma) - \Psi_{\mathbb{Z}^2}(\gamma)| \leq |\gamma| \exp(-(\log L)^2) = O(\exp(-(\log L)^2/2))$$

if γ has distance at least $(\log L)^2$ from $\partial\Lambda'$.

Secondly, we want to exclude contours with long "button-holes". Choose $a' \in (a, 1)$. For any contour γ and any pair of bonds $b, b' \in \gamma$ we let $d_\gamma(b', b)$ denote the number of bonds in Γ between b and b' (along the shortest of the two portions of γ connecting b, b'). Finally, we define the set of *contours with button-holes* as the subset $\Sigma' \subset \Sigma$ such that there exist $b, b' \in \gamma$ with $d_\gamma(b, b') \geq L^{a'}$ and $|x(b) - x(b')| \leq (1/2)d_\gamma(b, b')$, where $x(b), x(b')$ denote the centers of b, b', and $|\cdot|$ is the ℓ^1 distance. The next result states that contours with button-holes are unlikely:

Lemma 4.7. *For any $c > 0$ and β large enough,*

$$\tilde{\pi}^{h-1}_{\Lambda'}(\exists\gamma \in \Sigma' \text{ such that } \mathcal{C}_{\gamma,h} \text{ holds}) \leq e^{-cL^{a'}}.$$

Proof. The proof is based on standard Peierls arguments, so we will be extremely concise. Suppose that $\gamma \in \Sigma'$; that implies the existence of two bonds $b, b' \in \gamma$ with $d_\gamma(b, b') \geq L^{a'}$ and $|x(b) - x(b')| \leq (1/2)d_\gamma(b, b')$. One can then short-cut the button-hole, to obtain a new contour γ' that is at least $(1/2)d_\gamma(b, b') \geq (1/2)L^{a'}$ shorter than γ and at the same time contains the same large square $Q \subset \Lambda$ of area $(1-\epsilon_\beta)L^2$. The basic observation is then that the area variation satisfies

$$\big||\Lambda_\gamma \cap \Lambda| - |\Lambda_{\gamma'} \cap \Lambda|\big| \leq \min(d_\gamma(b, b')^2, \epsilon_\beta L^2),$$

so that

$$-\beta|\gamma| + \Psi_{\mathbb{Z}^2}(\gamma) + \frac{\lambda}{L}|\Lambda \cap \Lambda_\gamma| \leq -\beta|\gamma'| + \Psi_{\mathbb{Z}^2}(\gamma') + \frac{\lambda}{L}|\Lambda \cap \Lambda'_\gamma| - (\beta/4)L^{a'}.$$

At this point, (4.16) together with routine Peierls arguments allows us to sum over all possible shapes of the part of the contour between b, b' and get the claim (recall that $a' > a$). □

The important property of contours without button-holes is that the interaction between two portions of the contour is at most of order $L^{a'}$:

Claim 4.8. *If γ has no button-holes, then for every decomposition of γ into a concatenation $\gamma_1 \circ \cdots \circ \gamma_n$ we have*[5] $|\Psi_{\mathbb{Z}^2}(\gamma) - \sum_{i=1}^n \Psi_{\mathbb{Z}^2}(\gamma_i)| \leq (n-1)L^{a'}$.

Proof. It is sufficient to prove the claim for $n = 2$. Let P be the junction point of the two paths γ_1, γ_2. From the representation (4.15), one has

$$|\Psi_{\mathbb{Z}^2}(\gamma) - \Psi_{\mathbb{Z}^2}(\gamma_1) - \Psi_{\mathbb{Z}^2}(\gamma_2)| \leq \sum_{\substack{V \cap \gamma_1 \neq \emptyset \\ V \cap \gamma_2 \neq \emptyset}} \bar{\varphi}(V),$$

where we write $\bar{\varphi}(V) = \sup_U |\varphi_U(V)|$. The latter sum can be bounded by

$$\sum_{\substack{b \in \gamma_1 : d_{\gamma_1}(b, P) \leq L^{a'} \\ b' \in \gamma_2 : d_{\gamma_2}(b', P) \leq L^{a'}}} \sum_{V \ni \{b, b'\}} \bar{\varphi}(V) + \sum_{\substack{b \in \gamma_1 : d_{\gamma_1}(b, P) > L^{a'} \\ b' \in \gamma_2}} \sum_{V \ni \{b, b'\}} \bar{\varphi}(V)$$
$$+ \sum_{\substack{b' \in \gamma_2 : d_{\gamma_2}(b', P) > L^{a'} \\ b \in \gamma_1}} \sum_{V \ni \{b, b'\}} \bar{\varphi}(V).$$

Using the decay properties of the potentials $\varphi(\cdot)$ (see Lemma 2.16(iii)), one has

$$\sum_{V \ni \{b, b'\}} \bar{\varphi}(V) \leq \sum_{V \ni \{b, b'\}} \exp(-(\beta - \beta_0)d(V)) \leq \exp(-c_\beta |x(b) - x(b')|)$$

[5] Strictly speaking, in (4.15) we have defined $\Psi_\Lambda(\gamma)$ for a closed contour. For an open portion γ' of a closed contour γ, one can define for instance

$$\Psi_{\mathbb{Z}^2}(\gamma') = -\sum_{\substack{W \subset \mathbb{Z}^2 \\ W \cap \gamma' \neq \emptyset}} \varphi_0(W) + \sum_{\substack{W \subset \Lambda_\gamma \\ W \cap \gamma' \neq \emptyset}} \varphi_{\Delta_\gamma^+}(W) + \sum_{\substack{W \subset \mathbb{Z}^2 \setminus \Lambda_\gamma \\ W \cap \gamma' \neq \emptyset}} \varphi_{\Delta_\gamma^-}(W).$$

for a constant $c_\beta > 0$ with $c_\beta \to \infty$ as $\beta \to \infty$. For β large enough, this implies that the first term above contributes at most $\frac{1}{2}L^{a'}$. For the remaining two terms we use the no-button-hole assumption, which implies that $|x(b) - x(b')| \geq \frac{1}{2}d_\gamma(b, b')$, so that each term contributes at most $\frac{1}{4}L^{a'}$. □

Step 3. We are now in a position to conclude the proof of Proposition 4.6. Let $\mathcal{M}$ denote the set of contours in Λ', of length at most $5L$, that do not come too close to the boundary of Λ', that include a square $Q \subset \Lambda$ of side $(1 - \epsilon_\beta)L^2$ and finally that have no button-holes. In view of the previous discussion, it will be sufficient to upper bound the $\tilde{\pi}^{h-1}_{\Lambda'}$-probability of the event $\bigcup_{\gamma \in \mathcal{M}} \mathscr{C}_{\gamma,h}$. Let $\mathcal{V}_s = \{\underline{v} = (v_1, \ldots, v_s, v_{s+1} = v_1) : v_i \in \Lambda'\}$ denote a sequence of points in Λ'. We write $\gamma \in \mathcal{M}_{\underline{v}}$ if $\gamma \in \mathcal{M}$, all the v_i appear along γ in that order, and for each $i \geq 2$, v_i is the first point x on γ after v_{i-1} such that $|x - v_{i-1}| \geq \epsilon L$ where $\epsilon > 0$ is a fixed parameter that at the end we will need to take small enough independently of L. Note that since we are considering $|\gamma| \leq 5L$ we have $s \leq 5/\epsilon$. Then

$$\tilde{\pi}^{h-1}_{\Lambda'}(\exists \gamma \in \mathcal{M}, \mathscr{C}_{\gamma,h}) \leq \sum_{s=1}^{5/\epsilon} \sum_{\underline{v} \in \mathcal{V}_s} \sum_{\gamma \in \mathcal{M}_{\underline{v}}} \tilde{\pi}^{h-1}_{\Lambda'}(\mathscr{C}_{\gamma,h})$$
$$\leq \sum_{s=1}^{5/\epsilon} \sum_{\underline{v} \in \mathcal{V}_s} \sum_{\gamma \in \mathcal{M}_{\underline{v}}} \exp\Big(-\beta|\gamma| + \Psi_{\mathbb{Z}^2}(\gamma) + \frac{\lambda}{L}|\Lambda_\gamma \cap \Lambda| + O(L^a)\Big),$$

where we have used (4.16) (with $\Psi_{\Lambda'}$ replaced by $\Psi_{\mathbb{Z}^2}$). Now let $K_{\underline{v}}$ denote the convex hull of the set of points $\underline{v}$. Since the contour γ is never more than ϵL away from a point in $\mathcal{V}$ (by definition of $\mathcal{M}_{\underline{v}}$), we have

$$|\Lambda_\gamma \cap \Lambda| \leq |K_{\underline{v}} \cap \Lambda| + 4s\epsilon^2 L^2 \leq |K_{\underline{v}} \cap \Lambda| + 20\epsilon L^2.$$

Also, from Claim 4.8 we have, if $\gamma_{i,i+1}$ is the portion of γ between v_i and v_{i+1},

$$\Big|\Psi_{\mathbb{Z}^2}(\gamma) - \sum_{i=1}^{s} \Psi_{\mathbb{Z}^2}(\gamma_{i,i+1})\Big| \leq sL^{a'}.$$

Now note that, by standard estimates of [20, Ch. 4],

$$\sum_{\gamma \in \mathcal{M}_{\underline{v}}} \exp(-\beta|\gamma| + \Psi_{\mathbb{Z}^2}(\gamma)) \leq e^{O(L^{a'})} \prod_{i=1}^{s} \sum_{\gamma_{i,i+1}} e^{-\beta|\gamma_{i,i+1}| + \Psi_{\mathbb{Z}^2}(\gamma_{i,i+1})}$$
$$\leq e^{O(L^{a'})} \prod_{i=1}^{s} \exp\big(-(\beta + o(1))\tau(v_{i+1} - v_i)\big)$$
$$= \exp\Big(-(\beta + o(1)) \int_{\gamma_{\underline{v}}} \tau(\theta_s)\, ds + O(L^{a'})\Big)$$

with $o(1)$ vanishing as $L \to \infty$, the sum is over all contours $\gamma_{i,i+1}$ from v_i to v_{i+1}, $\gamma_{\underline{v}}$ denotes the piecewise linear curve joining $v_1, v_2, \ldots, v_1$ and we applied Appendix A.4 to reconstruct the surface tension $\tau(v_{i+1} - v_i)$ from the sum over $\gamma_{i,i+1}$ (cf. Definition 3.1).

By convexity of the surface tension,

$$\int_{\gamma_{\underline{v}}} \tau(\theta_s)\, ds \geq \int_{\partial K_{\underline{v}}} \tau(\theta_s)\, ds \geq \int_{\partial [K_{\underline{v}} \cap \Lambda]} \tau(\theta_s)\, ds$$

where $\partial K_{\underline{v}}$ denotes the boundary of $K_{\underline{v}}$ and $K_{\underline{v}} \cap \Lambda$ is the intersection of $K_{\underline{v}}$ with $[1, L]^2 \subset \mathbb{R}^2$. By combining the above inequalities we get

$$\sum_{\gamma \in \mathcal{M}_{\underline{v}}} \exp\Big(-\beta|\gamma| + \Psi_{\mathbb{Z}^2}(\gamma) + \frac{\lambda}{L}|\Lambda_\gamma \cap \Lambda|\Big) \leq \exp\Big(-\beta \int_{\partial[K_{\underline{v}} \cap \Lambda]} \tau(\theta_s)\, ds + \frac{\lambda}{L}|K_{\underline{v}} \cap \Lambda| + c\epsilon L\Big)$$

for some constant $c > 0$ and L large enough. After rescaling $K_{\underline{v}} \cap \Lambda$ to the unit square we have a shape with area at least $1 - \epsilon_\beta$. Since $\lambda < \lambda_c$ we know from Claim 3.8 that for all curves $\gamma^\dagger$ in $[0, 1]^2$ enclosing such an area,

$$\mathcal{F}_\lambda(\gamma^\dagger) = -\beta \int_{\gamma^\dagger} \tau(\theta_s)\, ds + \lambda A(\gamma^\dagger) \leq -\theta < 0$$

for some θ (depending on λ, β). Hence

$$\sum_{\gamma \in \mathcal{M}_{\underline{v}}} \exp\Big(-\beta|\gamma| + \Psi_{\mathbb{Z}^2}(\gamma) + \frac{\lambda}{L}|\Lambda_\gamma \cap \Lambda|\Big) \leq \exp(-\theta L/2)$$

provided $\epsilon > 0$ is sufficiently small and L is sufficiently large. Now since $s \leq 5/\epsilon$ we have $|\mathcal{V}_s| \leq |\Lambda'|^{5/\epsilon}$ and so

$$\widetilde{\pi}^{h-1}_{\Lambda'}(\exists \gamma \in \mathcal{M}, \mathscr{C}_{\gamma,h}) \leq (5/\epsilon)|\Lambda'|^{5/\epsilon} \exp(-\theta L/2) \leq c_1 e^{-c_2(\log L)^2},$$

which completes the proof of Proposition 4.6. □

4.3. Existence of a macroscopic $H(L)$-contour when $\lambda > \lambda_c$. In the special case where $\lambda \geq (1+a)\lambda_c$ for any (arbitrarily small) absolute constant $a > 0$ independent of β, one can prove the existence of a macroscopic $H(L)$-contour w.h.p. by following (with some more care) the same line of arguments used to establish a supercritical $H(L) - 1$ droplet in Section 4.1. To deal with the more delicate case where λ is arbitrarily close to λ_c we provide the following proposition. Fix $a > 0$ small enough but independent of β. Then we have

Proposition 4.9. *Let β be sufficiently large. For any $\delta > 0$ there exist constants c_1, c_2 such that if $(1 + \delta)\lambda_c \leq \lambda \leq \lambda_c(1 + a)$ then*

$$\pi^0_\Lambda\big(\exists \gamma : \mathscr{C}_{\gamma,H(L)}, |\Lambda_\gamma| \geq (9/10)L^2\big) \geq 1 - c_1 e^{-c_2(\log L)^2}.$$

Proof. First of all, from (4.4) and (4.5) we see that if $(1 + \delta)\lambda_c \leq \lambda \leq \lambda_c(1 + a)$ (and recalling that $\lambda_c/\beta \sim 4$), w.h.p. there are no $H(L)$-contours of length at least $(\log L)^2$ and enclosing an area of at most $(9/10)L^2$ (if a was chosen sufficiently small). Let $\mathcal{S}_0$ denote the event that there does not exist an $H(L)$-contour γ enclosing an area larger than $(9/10)L^2$. Thus, on the event $\mathcal{S}_0$ w.h.p. the largest $H(L)$-contour has length at most $(\log L)^2$.

Using Proposition 4.1 we find that the assumption of Theorem 6.2 holds. The latter implies that w.h.p. for any $\epsilon > 0$ there exists an external $(H(L)-1)$-contour Γ containing$(1-\epsilon)L\mathcal{L}_c(\lambda)$. We condition on this Γ. Since the event $\mathcal{S}_0$ is decreasing, by monotonicity with respect to the boundary conditions on the interior boundary Δ^+_Γ of Γ,

$$\pi^0_\Lambda(\mathcal{S}_0 \mid \Gamma) \le \pi^{H(L)-1}_{\Lambda_\Gamma}(\mathcal{S}_0),$$

so it suffices to work under $\pi^{H(L)-1}_{\Lambda_\Gamma}$. Let $\mathcal{S}$ denote the event that there are no macroscopic contours (of any height, positive or negative). The same arguments used in the proof of Lemma 2.9 show that w.h.p. (with respect to $\pi^{H(L)-1}_{\Lambda_\Gamma}$) there are no macroscopic contours on the event $\mathcal{S}_0$. Thus, $\pi^{H(L)-1}_{\Lambda_\Gamma}(\mathcal{S}^c \cap \mathcal{S}_0)$ is negligible and it suffices to upper bound the probability $\pi^{H(L)-1}_{\Lambda_\Gamma}(\mathcal{S})$.

To this end, we compare $\pi^{H(L)-1}_{\Lambda_\Gamma}(\mathcal{S})$ with the probability of a specific contour γ approximating the optimal curve and then sum over the choices of γ. Let

$$\mathcal{K} = \partial[(L(1-\epsilon) - L^{3/4})\mathcal{L}_c(\lambda)] \tag{4.17}$$

be the suitably dilated solution to the variational problem of maximizing $\mathcal{F}_\lambda$. By our choice of dilation factor, $\mathcal{K}$ is at distance at least $L^{3/4}$ from Γ. Then for some s growing slowly to infinity with L let $v_1, \dots, v_s$ be a sequence of vertices in clockwise order along $\mathcal{K}$ with $3L/s \le |v_i - v_{i+1}| \le 5L/s$ for $1 \le i \le s$ where $v_{s+1} = v_1$.

Let W be the bounded region delimited by the two curves

$$x \mapsto \xi^\pm(x) := \pm(x(1-x))^{3/5}, \qquad x \in [0,1].$$

We define the cigar shaped region W_i between points v_i and v_{i+1} as in [34, Section 1.4.6] to be given by W modulo a translation/rotation/dilation that brings $(0,0)$ to v_i and $(1,0)$ to v_{i+1}. Now let $\gamma = \gamma_1 \circ \cdots \circ \gamma_s$ be a closed contour where each γ_i is a curve from v_i to v_{i+1} inside the region W_i. Note that by construction $|\Lambda_\gamma| \ge |\Lambda_\mathcal{K}| - s(5L/s)^2 = |\Lambda_\mathcal{K}| - o(L^2)$ and γ is at least at distance $\frac{1}{2}L^{3/4}$ from $\partial\Lambda_\Gamma$.

In analogy with (4.9) we have

$$\frac{\pi^{H(L)-1}_{\Lambda_\Gamma}(\mathscr{C}_{\gamma,H(L)})}{\pi^{H(L)-1}_{\Lambda_\Gamma}(\mathcal{S})} = e^{-\beta|\gamma|}\frac{Z^{+,H(L)}_{\text{in}} Z^{-,H(L)-1}_{\text{out}}}{Z^{H(L)-1}_{\Lambda_\Gamma}(\mathcal{S})}$$

where: $Z^{+,H(L)}_{\text{in}}$ (resp. $Z^{-,H(L)-1}_{\text{out}}$) is the partition function in Λ_γ (resp. $\Lambda_\Gamma \setminus \Lambda_\gamma$) with floor at zero, boundary conditions at height $H(L)$ (resp. $H(L)-1$) and constraint $\eta \ge H(L)$ in $\Delta^+\gamma$ (resp. $\eta \le H(L)-1$ on Δ^-_γ); $Z^{H(L)-1}_{\Lambda_\Gamma}(\mathcal{S})$ is the partition function in Λ_Γ, boundary conditions at height $H(L)-1$, floor at zero and constraint $\eta \in \mathcal{S}$. As in Section 4.2, one can apply Proposition 2.12 to the numerator to get

$$\begin{aligned} &Z^{+,H(L)}_{\text{in}} Z^{-,H(L)-1}_{\text{out}} \\ &\quad = \hat{Z}^{+,H(L)}_{\text{in}} \hat{Z}^{-,H(L)-1}_{\text{out}} \exp\Big[-\frac{c_\infty}{L} e^{4\beta\alpha(L)}(e^{-4\beta}|\Lambda_\gamma| + |\Lambda_\Gamma \setminus \Lambda_\gamma|) + o(L)\Big] \end{aligned}$$

where the partition functions with the "hat" have no floor. As for the denominator,

$$Z_{\Lambda_\Gamma}^{H(L)-1}(\mathcal{S}) \le \hat{Z}_{\Lambda_\Gamma}^{H(L)-1}\hat{\pi}_{\Lambda_\Gamma}^{H(L)-1}(\eta\restriction_{\Lambda_\Gamma} \ge 0 \mid \mathcal{S})$$
$$\le \hat{Z}_{\Lambda_\Gamma}^{H(L)-1}\exp\Big[-\frac{c_\infty}{L}e^{4\beta\alpha(L)}|\Lambda_\Gamma| + o(L)\Big]$$

where we have applied (2.8) in the second step. Together with (4.14), the Definition 2.5 of λ and the fact that $|\Lambda_\gamma| \ge |\Lambda_{\mathcal{K}}| - o(L^2)$, this yields

$$\frac{\pi_{\Lambda_\Gamma}^{H(L)-1}(\mathscr{C}_{\gamma,H(L)})}{\pi_{\Lambda_\Gamma}^{H(L)-1}(\mathcal{S})} \ge \exp\Big(-\beta|\gamma| + \Psi_{\mathbb{Z}^2}(\gamma) + \frac{\lambda}{L}|\Lambda_{\mathcal{K}}| + o(L)\Big),$$

where we have replaced $\Psi_{\Lambda_\Gamma}(\gamma)$ with $\Psi_{\mathbb{Z}^2}(\gamma)$ (cf. the discussion after (4.16)), since by construction γ stays at distance at least $(1/2)L^{3/4}$ from $\partial\Lambda_\Gamma$.

At this point we can sum over γ, with the constraint that each portion $\gamma_{i,i+1}$ from v_i to v_{i+1} is in W_i as specified before. Since the cigar W_i is close to $W_{i\pm1}$ only at its tips, we have, from the decay properties of the potentials φ that define Ψ,

$$\Big|\Psi_{\mathbb{Z}^2}(\gamma) - \sum_{i=1}^{s}\Psi_{\mathbb{Z}^2}(\gamma_i)\Big| = O(s).$$

Also, by Appendix A.4,

$$\sum_{\gamma_i\in W_i} e^{-\beta|\gamma_{i,i+1}|+\Psi_{\mathbb{Z}^2}(\gamma_{i,i+1})} = \exp\big(-(\beta+o(1))\tau(v_{i+1}-v_i)\big).$$

Summing over all such contours we obtain

$$\frac{\sum_\gamma \pi_{\Lambda_\Gamma}^{H(L)-1}(\mathscr{C}_{\gamma,H(L)})}{\pi_{\Lambda_\Gamma}^{H(L)-1}(\mathcal{S})} = e^{\frac{\lambda}{L}|\Lambda_{\mathcal{K}}|+o(L)}\prod_{i=1}^{s}\sum_{\gamma_{i,i+1}}\exp(-\beta|\gamma_{i,i+1}| + \Psi_{\mathbb{Z}^2}(\gamma_{v_i,v_{i+1}}))$$
$$= e^{\frac{\lambda}{L}|\Lambda_{\mathcal{K}}|+o(L)}\prod_{i=1}^{s}\exp(-\beta\tau(v_i - v_{i+1}))$$
$$= \exp\Big(-\beta\int_{\partial\mathcal{K}}\tau(\theta_s)\,ds + \frac{\lambda}{L}|\Lambda_{\mathcal{K}}| + o(L)\Big)$$
$$= \exp(L\mathcal{F}_\lambda(L^{-1}\mathcal{K}) + o(L)),$$

with $\mathcal{F}_\lambda(\cdot)$ the functional in (3.2). Since $L^{-1}\mathcal{K}$ is a close approximation to $\mathcal{L}_c(\lambda)$, by Claim 3.8(ii) it follows that $\mathcal{F}_\lambda(L^{-1}\mathcal{K}) > 0$. Hence $\pi_{\Lambda_\Gamma}^{H(L)-1}(\mathcal{S}) \le e^{-cL}$, which concludes the proof. □

4.4. Conclusion: Proof of Theorem 1. Assume that $\lambda(L_k)$ has a limit (otherwise it is sufficient to work on converging subsequences). The results established thus far show that w.h.p.:

- By Proposition 4.1 there exists an $(H(L_k)-1)$-contour enclosing an area of at least $(9/10)L_k^2$.
- By Corollary 2.10 there are no macroscopic $(H(L_k)+1)$-contours.

- When $\lim_{k\to\infty} \lambda(L_k) < \lambda_c$, by Proposition 4.6 there is no macroscopic $H(L_k)$-contour.
- When $\lim_{k\to\infty} \lambda(L_k) > \lambda_c$, by Proposition 4.9 there exists an $H(L_k)$-contour enclosing an area of at least $(9/10)L_k^2$.

Combining these statements with (4.1) and (4.2) completes the proof when $\lim_{k\to\infty} \lambda(L_k) \neq \lambda_c$. Whenever $\lambda(L_k) \to \lambda_c$ we want to prove that $\pi^0_{\Lambda_k}(E_{H(L_k)-1} \cup E_{H(L_k)}) \to 1$, i.e., we want to exclude, say, that half of the sites have height $H(L_k)$ and the other half have height $H(L_k) - 1$. This is a simple consequence of (4.4) and (4.5) which say that, when $\lambda \approx 4\beta$, either there are no macroscopic $H(L_k)$-contours, or there exists one enclosing an area of $(1-\epsilon_\beta)L^2$. The proof is then concluded also for $\lim_{k\to\infty} \lambda(L_k) = \lambda_c$, invoking again (4.1) and (4.2). ◻

4.5. Proof of Corollary 1.2. Consider first the case with no floor. With a union bound the probability that $\widehat{X}^*_L \geq \varphi(L) + \frac{1}{2\beta}\log L$ can be bounded by $L^2\hat{\pi}^0_\Lambda(\eta_x \geq \varphi(L) + \frac{1}{2\beta}\log L)$, which is $O(e^{-4\beta\varphi(L)})$ by Lemma 2.4. For the other direction, let $\mathcal{A}$ denote the set of $x \in \Lambda$ belonging to the even sublattice of $\mathbb{Z}^2$ and such that $\eta_y = 0$ for all neighbors y of x. Then, by conditioning on $\mathcal{A}$, and using the Markov property, one finds that the probability of $\widehat{X}^*_L \leq -\varphi(L) + \frac{1}{2\beta}\log L$ is bounded by the expected value $\hat{\pi}^0_\Lambda(\exp(-e^{4\beta\varphi(L)}|\mathcal{A}|/L^2))$. Using Chebyshev's bound and the exponential decay of correlations [12] it is easily established that the event $|\mathcal{A}| < \delta L^2$ has vanishing $\hat{\pi}^0_\Lambda$-probability as $L \to \infty$ if δ is small enough. Since $\varphi(L) \to \infty$ as $L \to \infty$, this ends the proof of (1.4).

For the proof of (1.5) we proceed as follows. Consider the π^0_Λ-probability that $X^*_L \leq \frac{3}{4\beta}\log L - \varphi(L)$. Condition on the largest $(H(L)-1)$-contour γ, which contains a square of side-length $\frac{9}{10}L$ w.h.p. thanks to Proposition 4.1. By monotonicity we may remove the floor and fix the height of the internal boundary condition on Λ_γ to $H(L) - 1$. At this point the argument given above for the proof of (1.4) yields

$$\pi^0_\Lambda\left(X^*_L \leq \frac{3}{4\beta}\log L - \varphi(L)\right) = o(1),$$

since $H(L) + \frac{1}{2\beta}\log L = \frac{3}{4\beta}\log L + O(1)$. To show that $\pi^0_\Lambda(X^*_L \geq \frac{3}{4\beta}\log L + \varphi(L)) = o(1)$, recall that w.h.p. there are no macroscopic $(H(L)+1)$-contours thanks to Corollary 2.10. Condition therefore on $\{\gamma_i\}$, all the external microscopic $(H(L)+1)$-contours. The area term in (2.2) is negligible for these, thus it suffices to treat each γ_i without a floor and with an external boundary height $H(L)$. The probability that a given $x \in \Lambda_{\gamma_i}$ sees an additional height increase of k is then at most $ce^{-4\beta k}$, and a union bound completes the proof. ◻

5. Local shape of macroscopic contours

In this section we establish the following result. Given $n \in \mathbb{Z}_+$, consider the SOS model in a domain of linear size $\ell = L^{2/3+\epsilon}$, with floor at zero and Dobrushin's boundary conditions around it at height $\{j-1, j\}$, where $j = H(L) - n$ (see below for the precise definition). We show that the entropic repulsion from the floor forces the unique open

j-contour to have height $(1 + o(1))c(j, \theta)\ell^{1/2+3\epsilon/2}$ above the straight line $\mathbb{L}$ joining its end points. The constant $c(j, \theta)$ is explicitly determined in terms of the contour index j and of the surface tension computed at the angle θ describing the tilting of $\mathbb{L}$ with respect to the coordinate axes. This result will be the key element in proving the scaling limit for the level lines as well as the $L^{1/3}$-fluctuations around the limit.

5.1. Preliminaries. We define a *domino* to be any rectangle in $\mathbb{Z}^2$ of short and long sides $(\log L)^2$ and $2(\log L)^2$ respectively. A subset $\mathcal{C} = \{x_1, \dots, x_k\}$ of the domino will be called a *spanning chain* if

(i) $x_i \neq x_j$ if $i \neq j$;
(ii) $\mathrm{dist}(x_i, x_{i+1}) = 1$ for all $i = 1, \dots k - 1$;
(iii) $\mathcal{C}$ connects the two opposite short sides of the domino.

Remark 5.1. Let R_k be a rectangle with short side $2(\log L)^2$ and long side $k(\log L)^2$, $k \in \mathbb{N}$, $k \geq 2$. Consider any two coverings of R_k, one with horizontal dominos and one with vertical ones, and fix a choice of a spanning chain for each domino in these coverings. The union of all these spanning chains necessarily contains a chain $\widetilde{\mathcal{C}} = \{y_1, \dots, y_n\} \subset R_k$ connecting the opposite short sides of R_k. Any chain constructed in this way will be called a *regular chain*.

Given $j \geq 0$ consider now the SOS model in a subset V of the $L \times L$ box Λ with boundary conditions at height j and floor at height 0.

Definition 5.2. Given a SOS-configuration η, a domino entirely contained in V will be called of *positive type* if there exists a spanning chain $\mathcal{C}$ inside it such that $\eta_x \geq j$ for all $x \in \mathcal{C}$. Similarly, if there exists a spanning chain $\mathcal{C}$ such that $\eta_x \leq j$ for all $x \in \mathcal{C}$, then the domino will be said to be of *negative type*.

Lemma 5.3. *W.h.p. all dominos in V are of positive type.*

Proof. A given domino is not of positive type iff there exists a *-chain $\{y_1, \dots, y_n\}$ connecting the two long opposite sides and such that $\eta_{y_i} < j$ for all i. Such an event is decreasing and therefore its probability is bounded from above by the probability with respect to the SOS model *without* the floor. Moreover, the above event implies the existence of a $(j-1)$-contour longer than $(\log L)^2$. The standard cluster expansion shows that the probability of the latter is $O(e^{-c(\log L)^2})$. A union bound over all possible choices of the domino completes the proof. □

Under the assumption that $|V|$ is not too large depending on j, we can also show that all dominos are of negative type. Recall the definition of c_∞ and δ_j from Lemma 2.4.

Lemma 5.4. *In the same setting of Lemma 5.3 with $j = H(L) - n$ for fixed n, assume V satisfies $|V|^{1/2} \leq 2e^{4\beta(j+1)}[(1 + \delta_j)c_\infty]^{-1}$. Then w.h.p. all dominos in V are of negative type.*

Proof. It follows from Proposition 2.7 that

$$\pi_V^j(\mathscr{C}_{\gamma,j+1}) \le e^{-\beta|\gamma|+(1+\delta_j)c_\infty e^{-4\beta(j+1)}|\Lambda_\gamma|}\, e^{\varepsilon_\beta e^{-4\beta(j+1)}|\gamma|\log|\gamma|}$$

with $\lim_{\beta\to\infty}\varepsilon_\beta = 0$. Clearly $|\Lambda_\gamma| \le |V|^{1/2}|\gamma|/4$ and $|\gamma| \le |V|$. Hence,

$$(1+\delta_j)c_\infty e^{-4\beta(j+1)}|\Lambda_\gamma| \le \beta|\gamma|/2, \qquad \varepsilon_\beta e^{-4\beta(j+1)}|\gamma|\log(|\gamma|) \ll \beta|\gamma|,$$

and a standard Peierls bound proves that w.h.p. there are no macroscopic $(j+1)$-contours. Hence w.h.p. all dominos are of negative type. □

5.2. Main result

Definition 5.5 (Regular circuit $\mathcal{C}_*$). Let $0 < \epsilon \ll 1$ and let Q (resp. $\tilde{Q}$) be the rectangle of horizontal side $L^{2/3+\epsilon}$ and vertical side $2L^{2/3+\epsilon}$ (resp. $L^{2/3+\epsilon} + 4(\log L)^2$ and $2L^{2/3+\epsilon} + 4(\log L)^2$) centered at the origin. Write $\tilde{Q} \setminus Q$ as the union of four thin rectangles (two vertical and two horizontal) of shorter side $2(\log L)^2$ and pick a regular chain for each of them as in Remark 5.1. Consider the shortest (self-avoiding) circuit $\mathcal{C}_*$ surrounding Q, contained in the union of the four chains. We call $\mathcal{C}_*$ a *regular circuit*.

Definition 5.6 (Boundary conditions on $\mathcal{C}_*$). Given a regular circuit $\mathcal{C}_*$ and integers a, b, j with $j > 0$ and $-L^{2/3+\epsilon} \le a \le b \le L^{2/3+\epsilon}$, we define a height configuration $\xi = \xi(\mathcal{C}_*, j, a, b)$ on $\mathcal{C}_*$ as follows. Choose a point P in $\mathcal{C}_*$, with zero horizontal coordinate and positive vertical coordinate. Follow the circuit anti-clockwise (resp. clockwise) until you hit for the first time the vertical coordinate a (resp. b), and let A (resp. B) be the corresponding point of $\mathcal{C}_*$. Set $\xi_x = j-1$ on the portion of $\mathcal{C}_*$ between A and B (not including extremes) that includes P, and set $\xi_x = j$ on the rest of the circuit; see Figure 3. It is easy to check that, given the regularity properties (cf. Remark 5.1) of the four chains composing the circuit, the construction of ξ is independent of the choice of P as above.

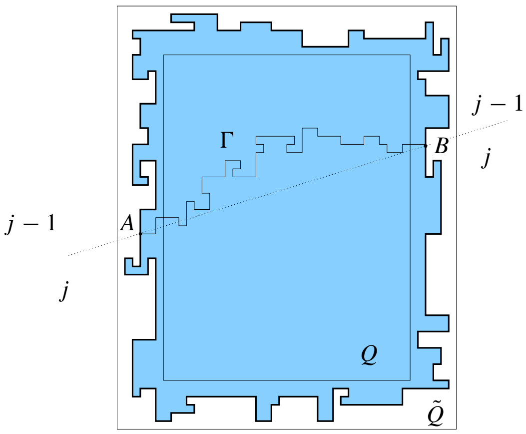

Fig. 3. A sketch of the region Λ delimited by the circuit $\mathcal{C}_*$, with the boundary conditions from Definition 5.6, and the open j-contour Γ induced by the boundary conditions.

With the above definitions let π_Λ^ξ be the SOS measure in the finite subset Λ of $\tilde{Q}$ delimited by $\mathcal{C}_*$, with boundary conditions ξ on $\mathcal{C}_*$ and floor at height 0. Note that the boundary conditions ξ induce a unique open j-contour Γ from A to B. Let also $\theta_{A,B} \in [0, \pi/4]$ be the angle formed with the horizontal axis by the segment AB, let $\ell_{A,B}$ be the Euclidean distance between A, B, and let $d_{A,B} = x_B - x_A$ where x_A, x_B denote the horizontal coordinates of A, B respectively.

Theorem 5.7. *Fix $n \geq 0$ and recall the Definition 2.5 of $\lambda^{(n)}$. Let $x \in [x_A, x_B]$ be such that $(x - x_A) \wedge (x_B - x) \geq \frac{1}{10} d_{A,B}$, let $X^\pm(n, x)$ be the points with horizontal coordinate x and vertical coordinate*

$$Y^\pm(n, x) = Y(n, x) \pm \sigma(x, \theta_{A,B}) L^\epsilon,$$

where

$$\begin{aligned} Y(n, x) &= \frac{a(x - x_A) + b(x_B - x)}{d_{A,B}} + \frac{\lambda^{(n)} (x - x_A)(x_B - x)}{2\beta L(\tau(\theta_{A,B}) + \tau''(\theta_{A,B}))(\cos\theta_{A,B})^3}, \\ \sigma^2(x, \theta_{A,B}) &= \frac{1}{\beta(\tau(\theta_{A,B}) + \tau''(\theta_{A,B}))(\cos\theta_{A,B})^3} \frac{(x - x_A)(x_B - x)}{d_{A,B}}. \end{aligned} \tag{5.1}$$

If the integer j that enters the definition of the boundary conditions ξ is equal to $H(L) - n$, then:

(1) *if $-\frac{1}{2} L^{2/3+\epsilon} \leq a \leq b \leq L^{2/3+\epsilon} - L^{1/3+3\epsilon}$ then w.h.p. $X^-(n, x)$ lies below Γ.*
(2) *if $-L^{2/3+\epsilon} + L^{1/3+3\epsilon} \leq a \leq b \leq \frac{1}{2} L^{2/3+\epsilon}$ then w.h.p. $X^+(n, x)$ lies above Γ.*

Remark 5.8. Since n is a fixed parameter, it is clear here that w.h.p. means that there exists $L_0 = L_0(n)$ such that for all $L \geq L_0$ the required probability is greater than $1 - \exp(-c(\log L)^2)$ for some $c > 0$. In the above the fraction $1/10$ could be replaced by an arbitrarily small constant independent of L. The core of the argument behind Theorem 5.7 is that the height of Γ above x is approximately a Gaussian $\mathcal{N}(Y(n, x), \sigma^2(x, \theta_{A,B}))$. Not surprisingly $\sigma^2(x, \theta_{A,B})$ has the form of the variance of a Brownian bridge. In concrete applications (see Section 6) we will only need the above statement for $\bar{x} = (x_A + x_B)/2$. Note that while $Y(n, \bar{x}) - (a + b)/2$ is of order $L^{1/3+2\epsilon}$, the fluctuation term $\sigma(\bar{x}, \theta_{A,B}) L^\epsilon$ is only $O(L^{1/3+3\epsilon/2})$.

Following [15, Sec. 7 and App. A], we begin by deriving an expression for the law of the open contour Γ. We refer to (4.15) for the definition of the decoration term $\Psi_\Lambda(\Gamma)$ (see also [15, App. A.3]).

Lemma 5.9. *In the setting of Theorem 5.7:*

(i) $\pi_\Lambda^\xi(|\Gamma| \geq 2L^{2/3+\epsilon}) \leq e^{-cL^{2/3+\epsilon}}$.
(ii) *Assume $|\Gamma| \leq 2L^{2/3+\epsilon}$. Then*

$$\pi_\Lambda^\xi(\Gamma) \propto \exp\left(-\beta|\Gamma| + \Psi_\Lambda(\Gamma) + \frac{\lambda^{(n)}}{L} |\Lambda_-| + \varepsilon_n(L) \right),$$

where $|\Lambda_-|$ denotes the number of sites in Λ below Γ, and $\varepsilon_n(L) = o(1)$ for any given n.

Proof. We first establish (i). Denote by $\hat{\pi}_\Lambda^\xi$ the SOS measure in Λ with boundary conditions ξ and no floor. Then

$$\pi_\Lambda^\xi(|\Gamma| \geq 2L^{2/3+\epsilon}) \leq \frac{\hat{\pi}_\Lambda^\xi(|\Gamma| \geq 2L^{2/3+\epsilon})}{\hat{\pi}_\Lambda^\xi(\eta_x \geq 0\ \forall x \in \Lambda)}.$$

The FKG inequality together with the simple bound $\min_{x\in\Lambda} \hat{\pi}_\Lambda^\xi(\eta_x \geq 0) \geq 1 - c/L$ implies that the denominator is larger than $\exp(-c'|\Lambda|/L)$ for some constant $c' = c'(n) > 0$. A simple Peierls estimate shows that the numerator is smaller than $\exp(-cL^{2/3+\epsilon})$. Since $|\Lambda|/L \leq cL^{1/3+2\epsilon}$ the result follows.

We now turn to part (ii). Given Γ, the region Λ is partitioned into two connected regions Λ_+, Λ_- separated by Γ (say that Λ_- is the one below Γ). Thus, if $Z_\Lambda^\xi(\Gamma)$ denotes the partition function restricted to all surfaces whose open contour is Γ, we have

$$Z_\Lambda^\xi(\Gamma) = e^{-\beta|\Gamma|} Z_{\Lambda_-}^{(j)} Z_{\Lambda_+}^{(j-1)}, \tag{5.2}$$

where $Z_{\Lambda_-}^{(j)}$ is the partition function of the SOS model in Λ_- with boundary conditions at height zero, floor at height $-j$ and the additional constraint that $\eta_x \geq 0$ for all $x \in \Lambda_-$ adjacent to Γ; $Z_{\Lambda_+}^{(j-1)}$ is defined similarly except that $\eta_x \leq 0$ for all $x \in \Lambda_+$ adjacent to Γ.

Let $\hat{Z}_{\Lambda_-}$ be defined as $Z_{\Lambda_-}^{(j)}$ but without floor and similarly for $\hat{Z}_{\Lambda_+}$. From Proposition A.1 we know that

$$\frac{Z_{\Lambda_-}^{(j)} Z_{\Lambda_+}^{(j-1)}}{\hat{Z}_{\Lambda_-} \hat{Z}_{\Lambda_+}} = \exp\big(-\hat{\pi}(\eta_0 > j)|\Lambda_-| - \hat{\pi}(\eta_0 \geq j)|\Lambda_+| + \varepsilon_n(L)\big)$$

where $|\Lambda_-|$ denotes the cardinality of Λ_- and $\varepsilon_n(L) = o(1)$ for any finite n. Since $j = H(L) - n$, using Lemma 2.4 we obtain

$$\hat{\pi}(\eta_0 \geq j) - \hat{\pi}(\eta_0 > j) = \frac{\lambda^{(n)}}{L}(1 + O(L^{-1/2})).$$

In conclusion, using $|\Lambda| = |\Lambda_-| + |\Lambda_+|$, (5.2) can be rewritten as

$$Z_\Lambda^\xi(\Gamma) \propto \exp\left(-\beta|\Gamma| + \frac{\lambda^{(n)}}{L}|\Lambda_-| + \varepsilon_n(L)\right)\left(\frac{\hat{Z}_{\Lambda_-}\hat{Z}_{\Lambda_+}}{\hat{Z}_\Lambda}\right)\hat{Z}_\Lambda$$

where $\hat{Z}_\Lambda$ is the partition function in Λ with no floor and boundary conditions at height zero. Using Lemma 2.16, as in (4.14) we have

$$\frac{\hat{Z}_{\Lambda_-}\hat{Z}_{\Lambda_+}}{\hat{Z}_\Lambda} = \exp(\Psi_\Lambda(\Gamma)),$$

and the result follows. □

Proof of Theorem 5.7. The fact that the circuit $\mathcal{C}_*$ enclosing Λ is wiggled introduces a number of inessential technical nuisances. In order not to hide the main ideas we will prove the theorem in the case when Λ coincides with Q and we refer to Appendix A.5 for a discussion covering the general case. In what follows we will drop the subscript A, B from $\ell_{A,B}$ and $\theta_{A,B}$. For simplicity we will only discuss the case $x = \frac{1}{2}(x_A + x_B)$ and we will drop x from $Y(n, x)$, $Y^{\pm}(n, x)$. The case of general x at distance at least $d_{A,B}/10$ from x_A, x_B can be treated similarly.

Proof of (1). We observe that the event, denoted by U, that the point $X^{-}(n)$ is above Γ is decreasing. Thus, by FKG, if G_+ denotes the decreasing event that Γ does not touch a $(\log L)^2$-neighborhood of the top side of Λ then

$$\pi_\Lambda^\xi(U) \le \frac{\pi_\Lambda^\xi(U; G_+)}{\pi_\Lambda^\xi(G_+)}.$$

The reason for conditioning on G_+ will be explained at the end of the proof. Thanks to Lemma 5.9 we can write

$$\frac{\pi_\Lambda^\xi(U; G_+)}{\pi_\Lambda^\xi(G_+)} = \frac{\sum_{\Gamma \in U \cap G_+} e^{-\beta|\Gamma| + \Psi_\Lambda(\Gamma) + \frac{\lambda^{(n)}}{L} A_-(\Gamma)}}{\sum_{\Gamma \in G_+} e^{-\beta|\Gamma| + \Psi_\Lambda(\Gamma) + \frac{\lambda^{(n)}}{L} A_-(\Gamma)}}(1 + o(1)). \tag{5.3}$$

We observe that $A_-(\Gamma)$ is, apart from an additive constant, the *signed area* $A(\Gamma)$ of the contour Γ with respect to the straight line joining A, B. Thus we can safely replace $A_-(\Gamma)$ with $A(\Gamma)$ in the above ratio.

Upper bound of the numerator. Let G_- denote the event that Γ does not touch the $(\log L)^2$-neighborhood of the bottom side of Λ. A simple Peierls argument shows that

$$\sum_{\Gamma \in U \cap G_+} e^{-\beta|\Gamma| + \Psi_\Lambda(\Gamma) + \frac{\lambda^{(n)}}{L} A(\Gamma)} = (1 + o(1)) \sum_{\Gamma \in U \cap G_+ \cap G_-} e^{-\beta|\Gamma| + \Psi_\Lambda(\Gamma) + \frac{\lambda^{(n)}}{L} A(\Gamma)},$$

since getting close to the bottom side of Λ implies an anomalous contour excess length. If now S denotes the infinite vertical strip through the points A, B, for any $\Gamma \in G_+ \cap G_-$ the decoration term $\Psi_\Lambda(\Gamma)$ satisfies

$$|\Psi_\Lambda(\Gamma) - \Psi_S(\Gamma)| = o(1)$$

thanks to (4.15) and Lemma 2.16. Therefore we can upper bound the numerator by

$$(1 + o(1)) \sum_{\Gamma \in U} e^{-\beta|\Gamma| + \Psi_S(\Gamma) + \frac{\lambda^{(n)}}{L} A(\Gamma)}.$$

Although we replaced the finite volume decorations Ψ_Λ with the decorations Ψ_S associated to the infinite strip S, we emphasize that the contours will always be constrained within the original box Λ.

It will be convenient in what follows to define, for an arbitrary event E,

$$Z_{\lambda^{(n)}}(E) := \sum_{\Gamma \in E} e^{-\beta|\Gamma| + \Psi_S(\Gamma) + \frac{\lambda^{(n)}}{L} A(\Gamma)}, \quad Z_{\lambda^{(n)}} := \sum_{\Gamma} e^{-\beta|\Gamma| + \Psi_S(\Gamma) + \frac{\lambda^{(n)}}{L} A(\Gamma)}.$$

Let now $Y \in [-L^{2/3+\epsilon}, L^{2/3+\epsilon}]$ be the height of the first (following Γ from A) horizontal bond of Γ crossing the middle vertical line of S and let $U = U_1 \cup U_2$ where

$$U_1 = U \cap \{Y \geq \hat{Y}^-(n)\}, \quad U_2 = U \cap \{Y \leq \hat{Y}^-(n)\}$$

and $\hat{Y}^-(n) = Y(n) - \frac{1}{2}\sigma(x,\theta)L^\epsilon$. In order to estimate $Z_{\lambda^{(n)}}(U_1)$, $Z_{\lambda^{(n)}}(U_2)$ we will apply the bounds of Section 5.3 below with $\mu = \lambda^{(n)}$.

We start from $Z_{\lambda^{(n)}}(U_1)$. Multiplying and dividing by Z_0 gives

$$Z_{\lambda^{(n)}}(U_1) = Z_0 \times \mathbb{E}_{\lambda=0}(U_1; e^{\frac{\lambda^{(n)}}{L}A(\Gamma)}) \leq Ce^{-\beta\tau(\theta)\ell}(Z_{2\lambda^{(n)}}/Z_0)^{1/2}\sqrt{\mathbb{P}_0(U_1)} \tag{5.4}$$

where we used Cauchy–Schwarz together with [20, Sec. 4.12] (or Corollary 5.13 below at $\mu = 0$) to upper bound Z_0. Again by Corollary 5.13 below,

$$Z_{2\lambda^{(n)}}/Z_0 \leq e^{cL^{3\epsilon}}.$$

Finally, we observe that the event U_1 implies that the contour Γ touches the middle vertical line in two points separated by a distance larger than $\frac{1}{2}\sigma(x,\theta)L^\epsilon \geq cL^{1/3+3\epsilon/2}$. Here $c = c(\beta) > 0$ for all $\beta < \infty$ by the strict convexity and analyticity of the surface tension. From [20, Ch. 4] such an event has probability smaller than $\exp(-cL^{1/3})$ under the $\lambda = 0$ measure. In conclusion

$$Z_{\lambda^{(n)}}(U_1) \leq Ce^{-\beta\tau(\theta)\ell}e^{-cL^{1/3}}. \tag{5.5}$$

We now turn our attention to the term $Z_{\lambda^{(n)}}(U_2)$. We first decompose $\Gamma = \Gamma_1 \circ \Gamma_2$ into the piece Γ_1 from A to $C = (0, Y)$ and Γ_2 from C to B. Then we write

$$A(\Gamma) = A_0 + A_1(\Gamma_1) + A_2(\Gamma_2)$$

where A_0, $A_1(\Gamma_1)$, $A_2(\Gamma_2)$ are the signed areas of the triangle (ACB) and of the contours Γ_1, Γ_2 with respect to the segments AC, CB respectively. Thus

$$\begin{aligned} Z_{\lambda^{(n)}}(U_2) \leq &\sum_{y \leq \hat{Y}^-(n)} e^{\frac{\lambda^{(n)}}{L}A_0} \\ &\times \sum_{\Gamma:\, Y=y} e^{-\beta|\Gamma_1| + \Psi_S(\Gamma_1) + \frac{\lambda^{(n)}}{L}A_1(\Gamma_1)} e^{-\beta|\Gamma_2| + \Psi_S(\Gamma_2) - \Delta\Psi_S(\Gamma_1,\Gamma_2) + \frac{\lambda^{(n)}}{L}A_2(\Gamma_2)} \\ =: &\sum_{y \leq \hat{Y}^-(n)} e^{\frac{\lambda^{(n)}}{L}A_0} \sum_{\Gamma_1:\, Y=y} e^{-\beta|\Gamma_1| + \Psi_S(\Gamma_1) + \frac{\lambda^{(n)}}{L}A_1(\Gamma_1)} Z_{\lambda,y,\Gamma_1} \end{aligned}$$

where

$$\Delta\Psi_S(\Gamma_1, \Gamma_2) = \Psi_S(\Gamma_1) + \Psi_S(\Gamma_2) - \Psi_S(\Gamma_1 \circ \Gamma_2). \tag{5.6}$$

It now follows from Corollary 5.13 that

$$\sup_{\Gamma_1} Z_{\lambda^{(n)},y,\Gamma_1} \leq \exp(\mathcal{G}_{\lambda^{(n)}}(\ell_2, \theta_2) + L^{3\epsilon/2}),$$

$$\sum_{\Gamma_1:\, Y=y} e^{-\beta|\Gamma_1| + \Psi_S(\Gamma_1) + \frac{\lambda^{(n)}}{L}A_1(\Gamma_1)} \leq \exp(\mathcal{G}_{\lambda^{(n)}}(\ell_1, \theta_1) + L^{3\epsilon/2}),$$

where ℓ_1, ℓ_2 denote the distances between A, C and C, B respectively, θ_1, θ_2 the angles with respect to the horizontal direction of the segments AC and CB, and we define

$$\mathcal{G}_\lambda(\ell, \theta) = -\beta\tau(\theta)\ell + \frac{\lambda^2 \ell^3}{24\beta(\tau(\theta) + \tau''(\theta))L^2}. \tag{5.7}$$

Putting all together we get

$$\begin{aligned} Z_{\lambda^{(n)}}(U_2) &\leq e^{2L^{3\epsilon/2}} \sum_{y \leq \hat{Y}^-(n)} \exp\left[\frac{\lambda^{(n)}}{L} A_0 + \mathcal{G}_{\lambda^{(n)}}(\ell_1, \theta_1) + \mathcal{G}_{\lambda^{(n)}}(\ell_2, \theta_2)\right] \\ &=: e^{2L^{3\epsilon/2}}[\Sigma_1 + \Sigma_2] \end{aligned} \tag{5.8}$$

where

$$\Sigma_1 = \sum_{y \leq (a+b)/2 - L^{1/3+3\epsilon}} \exp\left[\frac{\lambda^{(n)}}{L} A_0 + \mathcal{G}_{\lambda^{(n)}}(\ell_1, \theta_1) + \mathcal{G}_{\lambda^{(n)}}(\ell_2, \theta_2)\right], \tag{5.9}$$

and Σ_2 is the remaining sum. Using $\mathcal{G}_{\lambda^{(n)}}(\ell, \theta) \leq -\beta\tau(\theta)\ell + O(L^{3\epsilon})$ together with the strict convexity of the surface tension, we find that

$$\Sigma_1 \leq \exp(-\beta\tau(\theta)\ell - cL^{5\epsilon})$$

for some constant $c > 0$ where we have used that $A_0 \leq 0$ for $y \leq (a+b)/2$.

In order to bound Σ_2 we observe that for all $y \in [(a+b)/2 - L^{1/3+3\epsilon}, \hat{Y}^-(n)]$,

$$\varphi := \theta_1 - \theta = O(L^{-1/3+2\epsilon}).$$

Thus it suffices to expand $\mathcal{G}_{\lambda^{(n)}}(\ell_i, \theta_i)$ in φ up to second order. A little trigonometry shows that

$$\ell_1 = L^{2/3+\epsilon}/\cos(\theta+\varphi), \quad \ell_2 = L^{2/3+\epsilon}/\cos(\theta-\psi(\varphi)), \quad \psi(\varphi) = \varphi + 2(\tan\theta)\varphi^2 + O(\varphi^3).$$

Moreover

$$\frac{1}{24}\ell_i^3 \frac{(\lambda^{(n)})^2}{\beta(\tau(\theta_i) + \tau''(\theta_i))L^2} = \frac{1}{8}\ell^3 \frac{(\lambda^{(n)})^2}{24\beta(\tau(\theta) + \tau''(\theta))L^2} + o(1), \qquad i = 1, 2,$$

while

$$\tau(\theta_1)\ell_1 + \tau(\theta_2)\ell_2 = \tau(\theta)\ell + 2(\tau(\theta) + \tau''(\theta))\frac{[y - (a+b)/2]^2(\cos\theta)^2}{\ell} + o(1).$$

In conclusion

$$\begin{aligned} \mathcal{G}_{\lambda^{(n)}}(\ell_1, \theta_1) + \mathcal{G}_{\lambda^{(n)}}(\ell_2, \theta_2) &= \mathcal{G}_{\lambda^{(n)}}(\ell, \theta) - \frac{1}{32}\frac{(\lambda^{(n)})^2\ell^3}{\beta(\tau(\theta) + \tau''(\theta))L^2} \\ &\quad - 2\beta(\tau(\theta) + \tau''(\theta))\frac{[y - (a+b)/2]^2(\cos\theta)^2}{\ell} + o(1) \end{aligned}$$

and

$$\Sigma_2 \le (1+o(1)) \exp\left(\mathcal{G}_{\lambda^{(n)}}(\ell,\theta) - \frac{1}{32}\frac{(\lambda^{(n)})^2\ell^3}{\beta(\tau(\theta)+\tau''(\theta))L^2}\right)$$
$$\times \sum_{y\in[(a+b)/2-L^{1/3+3\epsilon},\hat{Y}^-(n)]} \exp\left(\frac{\lambda^{(n)}}{L}A_0 - 2\beta(\tau(\theta)+\tau''(\theta))\frac{[y-(a+b)/2]^2(\cos\theta)^2}{\ell}\right). \tag{5.10}$$

Since $A_0 = \frac{1}{2}L^{2/3+\epsilon}(y-(a+b)/2)$ and $\ell\cos\theta = L^{2/3+\epsilon}$, one finds that

$$\frac{\lambda^{(n)}}{L}A_0 - 2\beta(\tau(\theta)+\tau''(\theta))\frac{[y-(a+b)/2]^2(\cos\theta)^2}{\ell}$$
$$= -\frac{(y-Y(n))^2}{2\bar{\sigma}^2} + \frac{1}{32}\frac{(\lambda^{(n)})^2\,\ell^3}{\beta(\tau(\theta)+\tau''(\theta))L^2}, \tag{5.11}$$

where

$$Y(n) = \frac{a+b}{2} + \frac{1}{8}\frac{\lambda^{(n)}L^{1/3+2\epsilon}}{\beta(\tau(\theta)+\tau''(\theta))(\cos\theta)^3},$$

as in (5.1) for $x = (x_A + x_B)/2$, and

$$\bar{\sigma}^2 = \frac{\ell}{4\beta(\tau(\theta)+\tau''(\theta))(\cos\theta)^2}$$

Therefore, apart from the factor $\exp(\mathcal{G}_{\lambda^{(n)}}(\ell,\theta))$, the summand in (5.10) is proportional to a Gaussian density with mean $Y(n)$ and variance $\bar{\sigma}^2$. Using $\hat{Y}^-(n) = Y(n) - \frac{1}{2}\sigma(x,\theta_{A,B})L^\epsilon \le Y(n) - cL^{1/3+3\epsilon/2}$, and $\bar{\sigma}^2 = O(L^{2/3+\epsilon})$, one finds, for some $c>0$,

$$\Sigma_2 \le \exp(\mathcal{G}_{\lambda^{(n)}}(\ell,\theta) - cL^{2\epsilon}).$$

In conclusion, using (5.5) and (5.8), the numerator $Z_{\lambda^{(n)}}(U;G_+)$ appearing on the r.h.s. of (5.3) satisfies

$$Z_{\lambda^{(n)}}(U;G_+) \le \exp(\mathcal{G}_{\lambda^{(n)}}(\ell,\theta) - cL^{2\epsilon}) \tag{5.12}$$

for some new constant $c>0$.

Lower bound on the denominator. We consider the restricted class of contours defined as the set of Γ that stay within the neighborhood of size $(L^{2/3+\epsilon})^{1/2+\epsilon/3}$ around the optimal curve $\Gamma_{\text{opt}}^{\lambda^{(n)}}$ defined by

$$\Gamma_{\text{opt}}^{\mu}(x) = \Gamma_{\text{opt}}^{(1)}(x) + \Gamma_{\text{opt}}^{(2)}(x), \qquad x\in[x_A,x_B], \tag{5.13}$$

where $x\mapsto\Gamma_{\text{opt}}^{(1)}(x)$ describes the straight line segment AB and

$$\Gamma_{\text{opt}}^{(2)}(x) = \frac{\mu\ell_{A,B}^3}{2\beta L(\tau(\theta_{A,B})+\tau''(\theta_{A,B}))d_{A,B}^3}(x-x_A)(x_B-x), \qquad x\in[x_A,x_B],$$

where $d_{A,B} = x_B - x_A$. Note that, thanks to the assumption $a \le b \le L^{2/3+\epsilon} - L^{1/3+3\epsilon}$, the curve $\Gamma_{\text{opt}}^{\lambda^{(n)}}$ is well within the domain Q. For such contours Γ one has

$$A(\Gamma) = A(\Gamma_{\text{opt}}^{\lambda^{(n)}}) + O(L^{1+11\epsilon/6}).$$

Thus

$$Z_{\lambda^{(n)}}(G_+) \ge e^{O(L^{11\epsilon/6})} e^{\frac{\lambda^{(n)}}{L} A(\Gamma_{\text{opt}}^{\lambda^{(n)}})} \sum_{\Gamma:\, \text{dist}(\Gamma, \Gamma_{\text{opt}}^{\lambda^{(n)}}) \le (L^{2/3+\epsilon})^{1/2+\epsilon/3}} e^{-\beta|\Gamma| + \Psi_\Lambda(\Gamma)}.$$

As in [34, proof of Lemma A.6], the latter sum is lower bounded by

$$\exp\Big(-\beta \int_{\Gamma_{\text{opt}}^{\lambda^{(n)}}} ds\, \tau(\theta_s)\Big) \exp(-(\log L)^c).$$

Using (5.17) we finally get

$$Z_{\lambda^{(n)}}(G_+) \ge \exp\big(\mathcal{G}_{\lambda^{(n)}}(\ell, \theta) + O(L^{11\epsilon/6})\big) \tag{5.14}$$

where $\mathcal{G}_{\lambda^{(n)}}(\ell, \theta)$ is as in (5.7).

Conclusion. By combining (5.12) with (5.14) we finally get point (1) of the theorem. □

Proof of (2). The proof of (2) follows exactly the same pattern. Using FKG one first conditions on G_- and then, using Peierls, one restricts to paths in $G_- \cap G_+$. □

This concludes the proof of Theorem 5.7. □

5.3. Iterative upper bound on the partition function. This is a key technical section whose main object is a certain contour partition function $Z_{A,B}$ for open contours joining two points A, B at distance $\sim L^{2/3+\epsilon}$. The exponential weight of a contour contains, besides the familiar length term with decorations as in [20], an additional term proportional to L^{-1} times the signed area of the contour with respect to the segment AB. The main output is a precise upper bound on $Z_{A,B}$. The bound indicates that the main contribution to $Z_{A,B}$ comes from contours close to a deterministic curve (approximately a parabola) joining A, B and satisfying a variational principle (cf. (5.16)).

Setting. Recall that Q is a rectangle of horizontal side $D := L^{2/3+\epsilon}$ and vertical side $2D$ centered at the origin. Set $\ell_0 = L^{2/3}$ and $\delta = 1/10$ and define $\mathcal{R}_n$ as the set of pairs (A, B), $A, B \in Q \cap \mathbb{Z}^{2*}$, whose Euclidean distance $\ell_{A,B}$ satisfies $1 \le \ell_{A,B} \le 2^{n(1-\delta)}\ell_0$. Denote by $\theta_{A,B}$ the angle formed with the horizontal axis by the straight line through A, B and assume that $\theta_{A,B} \in [0, \pi/4]$. Without loss of generality we assume that $x_A < x_B$ if x_A, x_B denote the horizontal coordinates of A, B.

Choose two open contours $\Gamma_{\text{left}}, \Gamma_{\text{right}}$ such that Γ_{left} joins A to the left vertical side of Q without ever going to the right of A, and Γ_{right} joins B to the right vertical side of Q without ever going to the left of B.

Define now the contour ensemble Ξ consisting of all contours Γ joining A, B within Q such that the concatenation $\Gamma_{\text{left}} \circ \Gamma \circ \Gamma_{\text{right}}$ is an admissible open contour. Let $A(\Gamma)$ be the signed area of Γ with respect to the path obtained as concatenation of Γ_{left}, the segment AB and Γ_{right}. Fix a parameter $\mu \geq 0$ and to each $\Gamma \in \Xi$ assign the weight

$$w(\Gamma) = \exp\Big(-\beta|\Gamma| + \Psi_{\mathbb{Z}^2}(\Gamma) + e^{-\beta}|\Gamma \cap Q_{A,B}| + \frac{\mu}{L}A(\Gamma)\Big) \tag{5.15}$$

where $\Psi_{\mathbb{Z}^2}(\Gamma)$ has been defined in (4.15), and $Q_{A,B} \subset Q$ consists of all those points whose horizontal coordinate x satisfies either $x \leq x_A + (\log L)^2$ or $x \geq x_B - (\log L)^2$. The term $e^{-\beta}|\Gamma \cap Q_{A,B}|$ has been added only for technical convenience and, in practice, it will be $O((\log L)^2)$ for the "relevant" contours.

Definition 5.10. Let $Z_{A,B} := \sum_{\Gamma\in\Xi} w(\Gamma)$. We say that statement H_n holds if

$$\sup_{\Gamma_{\text{left}},\Gamma_{\text{right}}} Z_{A,B} \leq z_n e^{\mathcal{G}_\mu(\ell_{A,B},\theta_{A,B})} \quad \forall A, B \in \mathcal{R}_n$$

where

$$\mathcal{G}_\mu(\ell,\theta) = -\beta\tau(\theta)\ell + \ell^3 \frac{\mu^2}{24\beta(\tau(\theta)+\tau''(\theta))L^2}$$

and $z_1 = e^{c(\log L)^2}$, $z_n = (Lz_1)^{2^{n-1}}$, $n \geq 2$.

With the above notation the following holds.

Proposition 5.11. *For any large L, statement H_1 holds. Moreover, for all $n \leq n_f \equiv \epsilon(1-\delta)^{-1}\log_2(L)$, H_n implies H_{n+1}. In particular H_n holds for all $n \leq n_f$.*

Remark 5.12. It is important to observe that $z_{n_f} = O(\exp(L^{3\epsilon/2}))$.

The reason why H_n holds is that the main contribution to the partition function $Z_{A,B}$ comes from contours Γ which are close to the curve from A to B maximizing the functional

$$\mathcal{C} \mapsto -\beta \int_{\mathcal{C}} \tau(\theta_s)\, ds + \frac{\mu}{L} A(\mathcal{C}). \tag{5.16}$$

By expanding the functional up to second order around the straight line from A to B one finds easily that such an optimal curve is approximately the parabola given by (5.13). A short computation shows that

$$-\beta \int_{\Gamma^\mu_{\text{opt}}} \tau(\theta_s)\, ds + \frac{\mu}{L} A(\Gamma^\mu_{\text{opt}}) = \mathcal{G}_\mu(\ell_{A,B},\theta_{A,B}) + o(1). \tag{5.17}$$

Before proving the proposition let us state a simple corollary which formalizes a useful consequence of the result.

Corollary 5.13. *Consider the two partition functions* $Z^{(i)}_{A,B} := \sum_{\Gamma\in\Xi} w^{(i)}(\Gamma)$, $i = 1, 2$, *corresponding to the weights*

$$w^{(1)}(\Gamma) = \exp\Bigl(-\beta|\Gamma| + \Psi_S(\Gamma) + \frac{\mu}{L}A(\Gamma)\Bigr),$$
$$w^{(2)}(\Gamma) = \exp\Bigl(-\beta|\Gamma| + \Psi_S(\Gamma) - \Delta\Psi_S(\Gamma, \Gamma_{\text{left}}) + \frac{\mu}{L}A(\Gamma)\Bigr),$$

where S *is any vertical strip containing the strip through the points* A, B, *and* $\Delta\Psi_S(\Gamma, \Gamma_{\text{left}})$ *has been defined in* (5.6). *Then, uniformly in* $\Gamma_{\text{left}}, \Gamma_{\text{right}}$,

$$\max(Z^{(1)}_{A,B}, Z^{(2)}_{A,B}) \le \exp\bigl(\mathcal{G}_\mu(\ell_{A,B}, \theta_{A,B}) + O(L^{3\epsilon/2})\bigr).$$

Proof. This follows immediately from Proposition 5.11 together with Remark 5.12 and the bounds

$$|\Psi_{\mathbb{Z}^2}(\Gamma) - \Psi_S(\Gamma)| + \sup_{\Gamma_{\text{left}}} |\Delta\Psi_S(\Gamma, \Gamma_{\text{left}})| \le e^{-\beta}|\Gamma \cap Q_{A,B}|. \qquad \square$$

Proof of Proposition 5.11. We use induction on n.

Proof of the base case H_1. Fix $(A, B) \in \mathcal{R}_1$ with $x_A < x_B$, mutual distance $\ell \le 2\ell_0 = 2L^{2/3}$ and angle θ, together with $\Gamma_{\text{left}}, \Gamma_{\text{right}}$, and denote by h_Γ the maximal vertical distance (with respect to the line containing A, B) reached by the contour Γ. Then

$$Z_{A,B} = \hat{Z}_{A,B}\hat{\mathbb{E}}_{A,B}\bigl(e^{\frac{\mu}{L}A(\Gamma)+e^{-\beta}|\Gamma\cap Q_{A,B}|}\bigr) \tag{5.18}$$

where $\hat{Z}_{A,B}$ is defined as $Z_{A,B}$ but with modified weights $\hat{w}(\Gamma)$ in which the area parameter μ is equal to zero and the term $e^{-\beta}|\Gamma \cap Q_{A,B}|$ is absent. The area $A(\Gamma)$ clearly satisfies $|A(\Gamma)| \le |\Gamma|h_\Gamma$. Because of [20, Ch. 4.15],

$$\begin{aligned} \hat{\mathbb{P}}_{A,B}(h_\Gamma = j) &\le c\ell e^{-\min(j, j^2/\ell)/c}, \\ \hat{Z}_{A,B} &\le c e^{-\beta\tau(\theta)\ell}, \\ \hat{\mathbb{P}}_{A,B}(|\Gamma \cap Q_{A,B}| \ge q) &\le e^{-q+c(\log L)^2}, \end{aligned} \tag{5.19}$$

for a suitable constant c and β large enough. From (5.19) it follows that

$$\hat{\mathbb{E}}_{A,B}(e^{2e^{-\beta}|\Gamma\cap Q_{A,B}|}) \le e^{c'(\log L)^2}$$

for some constant $c' > 0$. Moreover, using Peierls, the excess length $(|\Gamma| - 2\ell)^+$ has exponential tail with parameter $\beta - O(1)$. This, combined with the first estimate in (5.19), implies that

$$\hat{\mathbb{E}}_{A,B}(e^{2\frac{\mu}{L}|\Gamma|h_\Gamma}) \le c' + \hat{\mathbb{E}}_{A,B}(e^{c'\mu\frac{1}{\sqrt{\ell}}h_\Gamma}) \le c''\ell^{3/2}$$

for some constants $c', c'' > 0$. The claim with $z_1 := e^{c(\log L)^2}$ then follows from the Cauchy–Schwarz inequality applied to the r.h.s. of (5.18).

Proof of the inductive step $H_n \Rightarrow H_{n+1}$. Fix $(A, B) \in \mathcal{R}_{n+1}$ with $x_A < x_B$, mutual distance ℓ and angle θ, together with $\Gamma_{\text{left}}, \Gamma_{\text{right}}$. Let C be the midpoint between A and B, and define $\mathbb{L}$ as the vertical line through C. Write Γ as $\Gamma = \Gamma_1 \circ \Gamma_2$ where Γ_1 is the contour from A until the first contact X with $\mathbb{L}$, and Γ_2 is the remaining part. Let also ℓ_{AX}, θ_{AX} be the length and angle of AX, and similarly for ℓ_{XB}, θ_{XB}.

Let j be the vertical coordinate of X minus the vertical coordinate of C. We distinguish two cases.

Case 1: $|j| \geq L^{1/3+3\epsilon}$. With the same notation as in the proof of H_1 and using (5.19) we get

$$\sum_{\Gamma:\, |j| \geq L^{1/3+3\epsilon}} w(\gamma) \leq c e^{-\beta\tau(\theta)\ell - c' L^{5\epsilon}} \tag{5.20}$$

for some positive c, c', which is clearly negligible compared with the target $z_{n+1} \exp(\mathcal{G}_\mu(\ell, \theta))$.

Case 2: $|j| \leq L^{1/3+3\epsilon}$. Simple geometry shows that both (A, X) and (X, B) belong to $\mathcal{R}_n$ and we can use induction. Also, a Peierls argument shows that we can safely assume that Γ_2 does not reach horizontal coordinate $x_A + (\log L)^2$, with x_A the horizontal coordinate of A, otherwise this would imply an extremely unlikely large deviation of the length $|\Gamma|$.

Note that the area $A(\Gamma)$ can be written as

$$A(\Gamma) = A_1(\Gamma_1) + A_2(\Gamma_2) + A_0$$

with $A_1(\Gamma_1)$ (resp. $A_2(\Gamma_2)$) the signed area of Γ_1 (resp. Γ_2) with respect to the segment AX (resp. XB), while

$$A_0 = \frac{\ell}{2} j \cos\theta$$

is the signed area with respect to AB of the triangle AXB.

Next, remark that

$$\Psi_{\mathbb{Z}^2}(\Gamma) = \Psi_{\mathbb{Z}^2}(\Gamma_1) + \Psi_{\mathbb{Z}^2}(\Gamma_2) - \Delta\Psi_{\mathbb{Z}^2}(\Gamma_1, \Gamma_2).$$

Using the decay properties of the potentials φ (see Lemma 2.16(iii)), we can bound

$$|\Delta\Psi_{\mathbb{Z}^2}(\Gamma_1, \Gamma_2)| \leq e^{-\beta}(|\Gamma_1 \cap Q_{A,X}| + |\Gamma_2 \cap Q_{X,B}|) \tag{5.21}$$

where $Q_{X,B}$ was defined just after (5.15) (with A replaced by X). As a consequence,

$$\sum_{\Gamma:\, |j| \leq L^{1/3+3\epsilon}} w(\gamma) \leq \sum_{\Gamma:\, |j| \leq L^{1/3+3\epsilon}} \exp\left[\frac{\mu}{L}(A_1(\Gamma_1) + A_2(\Gamma_2)) + \frac{\mu\ell}{2L} j \cos\theta\right]$$
$$\times \exp[e^{-\beta}(|\Gamma_1 \cap Q_{A,X}| + |\Gamma_2 \cap Q_{X,B}|) + \Psi_{\mathbb{Z}^2}(\Gamma_1) + \Psi_{\mathbb{Z}^2}(\Gamma_2)].$$

At last we can use induction: with $\Gamma_{\text{left}} \circ \Gamma_1$ playing the role of Γ_{left} we have (uniformly in Γ_1)

$$\sum_{\Gamma_2} \exp\left[\frac{\mu}{L} A_2(\Gamma_2) + \Psi_{\mathbb{Z}^2}(\Gamma_2) + e^{-\beta}|\Gamma_2 \cap Q_{X,B}|\right] \leq z_n \exp(\mathcal{G}_\mu(\ell_{XB}, \theta_{XB}))$$

and similarly, with e.g. horizontal contour from X to the right vertical boundary of Q playing the role of Γ_{right},

$$\sum_{\Gamma_1} \exp\left[\frac{\mu}{L} A_1(\Gamma_1) + \Psi_{\mathbb{Z}^2}(\Gamma_1) + e^{-\beta}|\Gamma_1 \cap Q_{A,X}|\right] \leq z_n \exp(\mathcal{G}_\mu(\ell_{AX}, \theta_{AX})).$$

To estimate

$$\Sigma := z_n^2 \sum_{|j| \leq L^{1/3+3\epsilon}} \exp\left[\frac{\mu\ell}{2L} j \cos\theta + \mathcal{G}_\mu(\ell_{AX}, \theta_{AX}) + \mathcal{G}_\mu(\ell_{XB}, \theta_{XB})\right]$$

we proceed as in the estimate of the sum Σ_2 appearing in (5.9). Using the restriction $|j| \leq L^{1/3+3\epsilon}$ we can expand up to second order the exponent in e.g. $\theta - \theta_{AX} = O(L^{-1/3+3\epsilon})$. The net result is that

$$\begin{aligned}\Sigma &\leq (1+o(1)) z_n^2 \exp\left(\mathcal{G}_\mu(\ell,\theta) - \frac{1}{32}\frac{\mu^2\ell^3}{\beta(\tau(\theta)+\tau''(\theta))L^2}\right) \\ &\quad \times \sum_j \exp\left(\frac{\mu\ell}{2L} j\cos\theta - 2\beta(\tau(\theta)+\tau''(\theta))\frac{j^2(\cos\theta)^2}{\ell}\right) \\ &\leq c(\beta)\sqrt{\ell}\, z_n^2 \exp(\mathcal{G}_\mu(\ell,\theta))\end{aligned}$$

where we have used a standard Gaussian summation. In conclusion, using (5.20), we have shown that

$$Z_{A,B} \leq c z_n^2 \sqrt{\ell} \exp(\mathcal{G}_\mu(\ell,\theta)) \leq z_{n+1} \exp(\mathcal{G}_\mu(\ell,\theta))$$

thanks to the definition of the constants $\{z_n\}_{n \leq n_f}$. The inductive step is complete. □

6. Proof of Theorems 2 and 3

In this section we show that for all fixed $n \in \mathbb{Z}_+$ (independent of L), if there exists a macroscopic $(H(L)-n)$-contour Γ_n containing the rescaled Wulff body $L\ell_c(\lambda^{(n)})\mathcal{W}_1$, then with high probability it is unique and it is contained in the annulus $(1+\varepsilon_0)L\mathcal{L}_c(\lambda^{(n)}) \setminus (1-\varepsilon_0)L\mathcal{L}_c(\lambda^{(n)})$ for any $\varepsilon_0 > 0$. This result will follow from a bootstrap procedure that involves proving that, roughly, if the macroscopic $(H(L)-n)$-contour Γ_n contains a large enough droplet (i.e. a rescaled Wulff body) then it must contain w.h.p. a slightly (depending on L) larger droplet. We refer to this phenomenon as *growth of droplets*. Combined with the results of Section 4, this will prove Theorem 2. Moreover, we prove that along the flat part of $L\mathcal{L}_c(\lambda^{(n)})$ the contour Γ_n has fluctuations on the scale $L^{1/3}$ up to $O(L^\epsilon)$ corrections. That covers Theorem 3.

6.1. Growth of droplets. Recall the Definition 3.4 of the sets $\mathcal{L}(\lambda, t, r)$ and the sets $\mathcal{L}_c(\lambda)$. Recall also the Definition 2.5 of the parameters $\lambda^{(n)}$. To fix ideas, the Wulff shape $\mathcal{W}_1$ appearing below is assumed to be centered at the origin. To simplify the exposition, we introduce the following notation.

Definition 6.1. Given a subset $\mathcal{A} \subset \mathbb{Z}^2$, we denote by $\mathcal{E}_n(\mathcal{A})$ the event that there exists an $(H(L)-n)$-contour Γ_n that contains $\mathcal{A}$.

Theorem 6.2 (Growth of the critical droplet). *Fix $\varepsilon \in (0, 1/10)$ and set $\delta_L = L^{-\varepsilon/8}$. Let Λ be the square of side L centered at the origin. Consider the SOS model on Λ with boundary conditions at height zero and floor at height zero. For any fixed $n \in \mathbb{Z}_+$, as $L \to \infty$, if*

$$\mathcal{E}_n(L(1+\delta_L)\ell_c(\lambda^{(n)})\mathcal{W}_1) \quad \textit{holds w.h.p.}, \tag{6.1}$$

then w.h.p.

(a) $\mathcal{E}_n(L\mathcal{L}(\lambda^{(n)}, 1+\delta_L, -L^{-2/3+4\varepsilon}))$ *holds;*
(b) *there exists a unique macroscopic $(H(L)-n)$-contour.*

Remark 6.3. Recall from Remark 3.5 that $L\ell_c(\lambda^{(n)})\mathcal{W}_1$ can fit inside the box Λ iff $\lambda^{(n)} \geq \hat{\lambda}$. Since $\hat{\lambda} \sim 2\beta$ (see Remark 3.7), we see that for $n \geq 1$ this condition is always satisfied (for β large enough) while if $n = 0$ we need to require $\lambda \geq \hat{\lambda}$. However, the results of Section 4 show that a macroscopic $H(L)$-contour exists w.h.p. iff $\lambda > \lambda_c > \hat{\lambda}$.

Remark 6.4 (Growth up to $L^{1/3}$ from flat boundary). An immediate corollary of Theorem 6.2 is that assuming (6.1), the unique macroscopic $(H(L)-n)$-contour is at distance $O(L^{1/3+4\varepsilon})$ from the target region $L\mathcal{L}_c(\lambda^{(n)})$, uniformly along most of the flat boundary of $L\mathcal{L}_c(\lambda^{(n)})$. Indeed, $L\mathcal{L}(\lambda^{(n)}, 1+\delta_L, -L^{-2/3+4\varepsilon})$ is uniformly at a distance $L^{1/3+4\varepsilon}$ from the critical region $L\mathcal{L}(\lambda^{(n)}, 1+\delta_L, 0)$, which overlaps with $L\mathcal{L}_c(\lambda^{(n)})$ along the flat boundary of $L\mathcal{L}(\lambda^{(n)}, 1+\delta_L, 0)$. On the other hand, concerning the curved portions of $L\mathcal{L}_c(\lambda^{(n)})$, the above theorem does not allow us to infer an approximation error better than $O(\delta_L L)$, since already the region $L\mathcal{L}(\lambda^{(n)}, 1+\delta_L, 0)$ has radial distance from $L\mathcal{L}_c(\lambda^{(n)})$ of that order at a corner.

The proof of Theorem 6.2(a) will be based on an inductive argument, but first (b) will be shown to be a consequence of (a).

Proof of Theorem 6.2(b) assuming (a). Thanks to (a), for any fixed n, assuming (6.1), w.h.p. there exists an outermost $(H(L)-n)$-contour, which we denote Γ_n, containing the set $\Lambda_n := (1-o(1))L\mathcal{L}_c(\lambda^{(n)})$ for a suitable error term $o(1)$. Proposition 2.7 and a union bound imply that there are no macroscopic negative contours w.h.p. and hence there are no positive macroscopic $(H(L)-n)$-contours nested inside Γ_n. Thus the interior of any other macroscopic contour must be contained inside $\Lambda \setminus \Lambda_n$ and

$$|\Lambda \setminus \Lambda_n| = (1+o(1))L^2\ell_c^2(\lambda^{(n)})(\ell_\tau^2 - 1) \leq \varepsilon_\beta \frac{\beta^2}{(\lambda^{(n)})^2}L^2$$

where $\varepsilon_\beta \to 0$ as $\beta \to \infty$. Here we have used the fact that $\ell_c(\lambda^{(n)}) \sim 2\beta/\lambda^{(n)}$ for $\beta \to \infty$ and that $\lim_{\beta\to\infty} \ell_\tau = 1$ because, in the same limit, the Wulff shape becomes a square.

Now a closed contour γ with $\Lambda_\gamma \subseteq \Lambda \setminus \Lambda_n$ satisfies

$$|\Lambda_\gamma| \leq \left(\frac{|\gamma|^2}{16}\right)^{1/2} |\Lambda \setminus \Lambda_n|^{1/2} \leq \frac{|\gamma|}{4}\frac{\beta}{\lambda^{(n)}}\sqrt{\epsilon_\beta}\,L.$$

Hence the area term $\lambda^{(n)}|\Lambda_\gamma|/L$ appearing in the exponential weight in (2.2) is negligible compared to the length term $\beta|\gamma|$ and the probability that there exists any such macroscopic contour is $O(e^{-c(\log L)^2})$ by a simple counting argument. In conclusion, w.h.p. Γ_n is the unique macroscopic $(H(L)-n)$-contour. □

The proof of Theorem 6.2(a) will be based on the following argument.

Proposition 6.5. *Fix $n \in \mathbb{Z}_+$. Let $1 \le m \le \log L$, let $\Lambda' \subset \mathbb{Z}^2$ be a region containing $L\mathcal{L}(\lambda^{(n)}, 1+\delta_L, -(m-1)\gamma_L)$, $\gamma_L := L^{-2/3+3\varepsilon}\log L$, and consider the SOS model on Λ' with boundary conditions at height $H(L)-n-1$ and floor at height zero. Conditionally on $\mathcal{E}_n(L(1+\delta_L)\ell_c(\lambda^{(n)})\mathcal{W}_1)$, the event $\mathcal{E}_n(L\mathcal{L}(\lambda^{(n)}, 1+\delta_L, -m\gamma_L))$ holds w.h.p.*

We start with the case $n = 0$. When $n = 0$ it is assumed that Λ' contains $\mathcal{A} := L\mathcal{L}(\lambda, 1+\delta_L, -(m-1)\gamma_L)$ and we condition on the event $\mathcal{E}_0(L(1+\delta_L)\ell_c(\lambda)\mathcal{W}_1)$ that there is an $H(L)$-contour containing the Wulff body $L(1+\delta_L)\ell_c(\lambda)\mathcal{W}_1$. We show that w.h.p. this initial droplet grows until it invades the whole region $L\mathcal{L}(\lambda, 1+\delta_L, -m\gamma_L)$. The proof is divided into two main steps.

Step 1. For $x \in \mathcal{A}$, $\ell > 0$, let $\mathcal{W}(x,\ell)$ denote the rescaled Wulff shape $L\ell\mathcal{W}_1$ centered at x. Also, let ℓ_x denote the maximal value of ℓ such that $\mathrm{dist}(\mathcal{W}(x,\ell), \mathcal{A}^c) > L^{1/3+3\varepsilon}$. The next lemma shows that at any $x \in \mathcal{A}$ such that $\ell_x > \ell_c(\lambda)$ one can let an initial droplet $\mathcal{W}(x,\ell)$, $\ell_c(\lambda) < \ell < \ell_x$, grow until it touches the boundary of $\mathcal{A}$ up to $O(L^{1/3+3\varepsilon})$; see Figure 4.

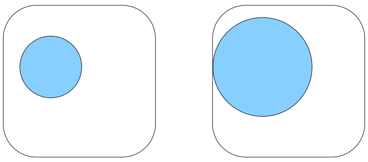

Fig. 4. Growth of the initial droplet as described in Step 1: from $\mathcal{W}(x,\ell)$ to $\mathcal{W}(x,\ell_x)$.

Lemma 6.6. *Fix $x \in \mathcal{A}$ and $\ell_c(\lambda)(1+\delta_L) \le \ell < \ell_x$. Conditionally on $\mathcal{E}_0(\mathcal{W}(x,\ell))$, the event $\mathcal{E}_0(\mathcal{W}(x,\ell_x))$ holds w.h.p.*

Proof. By simple recursion, it suffices to show that conditionally on $\mathcal{E}_0(\mathcal{W}(x,\ell))$, the event $\mathcal{E}_0(\mathcal{W}(x,\ell'))$ holds w.h.p., with $\ell' = \ell(1+L^{-2/3})$, as long as $\ell' \le \ell_x$. Next, we shall use the growth gadget of Theorem 5.7 along the boundary of $\mathcal{W}(x,\ell)$ to show that conditionally on $\mathcal{E}_0(\mathcal{W}(x,\ell))$, w.h.p. there is a circuit $\mathcal{C}$ surrounding $\mathcal{W}(x,\ell')$ such that $\eta_y \ge H(L)$ for all $y \in \mathcal{C}$. The latter event implies $\mathcal{E}_0(\mathcal{W}(x,\ell'))$.

By symmetry, we may restrict our analysis to the north-west corner of the droplet $\mathcal{W}(x,\ell)$. Moreover, using symmetry with respect to reflections along the north-west diagonal we may restrict to the upper half of the north-west corner. Let $\theta \in [0,\pi/4]$ and consider the chord of $\mathcal{W}(x,\ell)$ forming an angle θ with the x axis and whose horizontal projection has length $L^{2/3+\varepsilon}$. Let $z = (x_z, y_z)$ be the midpoint of this chord and denote by (x_a, y_a) and (x_b, y_b) the intersection points of the chord with $\partial\mathcal{W}(x,\ell)$, the boundary

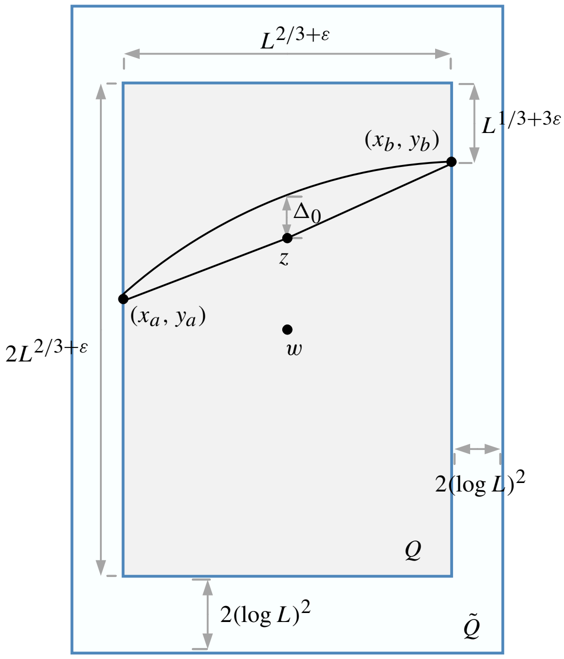

Fig. 5. Analysis of a chord of the droplet $\mathcal{W}(x, \ell)$ as described in Step 1 (Lemma 6.6).

of $\mathcal{W}(x, \ell)$ (see Figure 5). From a natural rescaling of the function $\Delta(d, \theta)$ appearing in Lemma 3.9, one finds that the vertical distance Δ_0 from z to $\partial\mathcal{W}(x, \ell)$ is given by

$$\Delta_0 = \ell L \Delta\left(\frac{L^{2/3+\varepsilon}}{\ell L \cos\theta}, \theta\right) = \frac{w_1}{16(\tau(\theta) + \tau''(\theta))(\cos\theta)^3} \frac{L^{1/3+2\varepsilon}}{\ell}(1 + o(1)).$$

Since $\beta w_1/2 = \lambda \ell_c(\lambda)$, for any $\lambda > 0$, one can rewrite

$$\Delta_0 = \frac{\lambda \ell_c(\lambda)/\ell}{8\beta(\tau(\theta) + \tau''(\theta))(\cos\theta)^3} L^{1/3+2\varepsilon}(1 + o(1)). \tag{6.2}$$

Consider the rectangle Q with horizontal side $L^{2/3+\varepsilon}$ and vertical side $2L^{2/3+\varepsilon}$ centered at the point $w = (x_w, y_w)$ such that $x_w = x_z$ and $y_w = y_b + L^{1/3+3\varepsilon} - L^{2/3+\varepsilon}$, and let $\tilde{Q}$ denote the enlarged rectangle with the same center, horizontal side $L^{2/3+\varepsilon} + 4(\log L)^2$ and vertical side $2L^{2/3+\varepsilon} + 4(\log L)^2$ (as illustrated in Figure 5). Observe that the assumption that $\ell \le \ell_x$, or equivalently $\operatorname{dist}(\mathcal{W}(x, \ell), \mathcal{A}^c) > L^{1/3+3\varepsilon}$, guarantees that the rectangles Q, $\tilde{Q}$ are indeed contained in our region $\Lambda' \supset \mathcal{A}$.

Notice that, setting $n = y_a - y_w$, $m = y_b - y_w$, one has $-\frac{1}{2}L^{2/3+\varepsilon} \le n \le m \le L^{2/3+\varepsilon} - L^{1/3+3\varepsilon}$, as required in Theorem 5.7(1). To ensure that we can indeed apply that statement we now check that w.h.p. there exists a regular circuit $\mathcal{C}_*$ in $\tilde{Q} \setminus Q$ with the required properties, namely that one has w.h.p.: (1) heights at least $H(L) - 1$ in the upper path along $\mathcal{C}_*$ connecting A and B and (2) heights at least $H(L)$ in the lower path along $\mathcal{C}_*$ connecting A and B, where A, B are defined in Definition 5.6; see Figure 3. Point (1) follows from Lemma 5.3 and the fact that we have boundary conditions at height $H(L) - 1$. Point (2) follows from the assumption that $\mathcal{E}_0(\mathcal{W}(x, \ell))$ holds: Indeed, on this event, using monotonicity, one may condition on the outermost $H(L)$-contour in order to obtain boundary conditions at height $H(L)$ outside a region enclosing the set $(\tilde{Q} \setminus Q) \cap \mathcal{W}(x, \ell)$, so that Lemma 5.3 implies the desired claim. By monotonicity in the boundary conditions, we can reduce to the case of constant boundary conditions equal to $H(L) - 1$ and $H(L)$ in the upper and lower parts of the circuit $\mathcal{C}_*$ respectively. Then, an application of Theorem 5.7(1) shows that the point $v = (x_v, y_v)$ with $x_v = x_w$ and $y_v = y_w + K$, with

$$K = \frac{a+b}{2} + \frac{\lambda L^{1/3+2\varepsilon}}{8\beta(\tau(\theta) + \tau''(\theta))(\cos\theta)^3} - c(\beta, \theta) L^{1/3+\varepsilon},$$

for a suitable constant $c(\beta, \theta) > 0$, lies w.h.p. below a chain $\mathcal{C}(v)$ connecting A and B with $\eta_y \geq H(L)$ for all $y \in \mathcal{C}(v)$. Call this event $\mathcal{F}(v)$. Next, observe that the point v lies above $\partial\mathcal{W}(x, \ell)$ and has a vertical distance h at least $L^{1/3+\varepsilon}$ from $\partial\mathcal{W}(x, \ell)$. Indeed, by (6.2),

$$h = K - \frac{a+b}{2} - \Delta_0 = \frac{\lambda(1 - \ell_c(\lambda)/\ell)}{8\beta(\tau(\theta) + \tau''(\theta))(\cos\theta)^3} L^{1/3+2\varepsilon} - c(\beta, \theta) L^{1/3+\varepsilon} \geq L^{1/3+\varepsilon},$$

where we use the assumption $1 - \ell_c(\lambda)/\ell \geq \delta_L$ and we take L large enough. In particular, it follows that v lies outside the enlarged shape $\mathcal{W}(x, \ell')$, $\ell' = \ell(1 + L^{-2/3})$, since this is larger than $\mathcal{W}(x, \ell)$ by an additive $O(L^{1/3})$ only.

Repeating the above argument for all $\theta \in [0, \pi/4]$ (of course $O(L)$ values of θ in this range suffice) and using symmetry to cover the other corners of the droplet, considering the intersection of all events $\mathcal{F}(v(\theta))$, one finds that w.h.p. there exists a chain $\mathcal{C}$ surrounding $\mathcal{W}(x, \ell')$ such that $\eta_y \geq H(L)$ for all $y \in \mathcal{C}$ as desired. □

Step 2. By assumption we can pretend that $\mathcal{E}_0(\mathcal{W}(x_0, \ell))$ holds, where x_0 is the center of the region A, which we identify with the origin. Thus, by Step 1, $\mathcal{E}_0(\mathcal{W}(x_0, \ell_{x_0}))$ holds w.h.p. Next, we establish that this is enough to invade the whole region $L\mathcal{L}(\lambda, 1 + \delta_L, -m\gamma_L)$.

Lemma 6.7. *Conditionally on $\mathcal{E}_0(\mathcal{W}(x_0, \ell_{x_0}))$, $\mathcal{E}_0(L\mathcal{L}(\lambda, 1 + \delta_L, -m\gamma_L))$ holds w.h.p.*

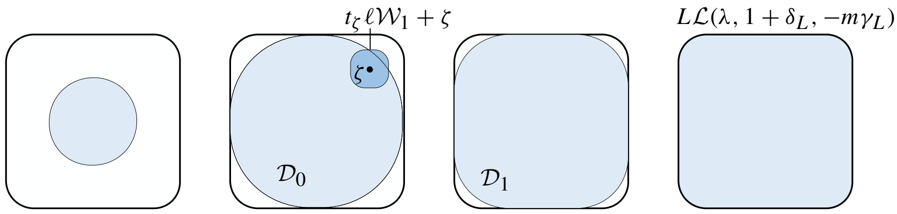

Fig. 6. Growth of the initial droplet as described in Step 2: from $\mathcal{W}(x_0, \ell)$ to $\mathcal{W}(x_0, \ell_{x_0})$ and from $\mathcal{W}(x_0, \ell_{x_0})$ to $L\mathcal{L}(\lambda, 1 + \delta_L, -m\gamma_L)$.

Before proving Lemma 6.7, we need the following deterministic lemma concerning the enlargement of squeezed Wulff shapes. Fix $\lambda > 0$ and $\ell_c(\lambda) < \ell < 1/\ell_\tau$. The Wulff body $\ell\mathcal{W}_1$ is strictly contained in the unit square Q; see Section 3 for the notation. Setting $\ell_* = 1/\ell_\tau$ one finds that $\mathcal{D}_0 := \ell_*\mathcal{W}_1$ is tangent to all four sides of Q, i.e. it is the maximal Wulff shape inside Q. For any $\zeta \in \mathcal{D}_0$ such that $\ell\mathcal{W}_1 + \zeta \subset \mathcal{D}_0$, define

$$t_\zeta = \max\{t \geq 1 : t\ell\mathcal{W}_1 + \zeta \subset Q\},$$

with the convention that $t_\zeta = 0$ if there is no such t. We define $\mathcal{D}_1 = \bigcup_{\zeta\in\mathcal{D}_0}\{t_\zeta\ell\mathcal{W}_1+\zeta\}$. We then repeat the above enlargement procedure. Namely, given the set $\mathcal{D}_k$, we define

$$\mathcal{D}_{k+1} = \bigcup_{\zeta\in\mathcal{D}_k}\{t_\zeta\ell\mathcal{W}_1 + \zeta\},$$

where $t_\zeta = \max\{t \geq 1 : t\ell\mathcal{W}_1 + \zeta \subset Q\}$ for $\zeta \in \mathcal{D}_k$ with $\ell\mathcal{W}_1 + \zeta \subset \mathcal{D}_k$ and with $t_\zeta = 0$ if $\ell\mathcal{W}_1 + \zeta \not\subset \mathcal{D}_k$. The sequence $\{\mathcal{D}_k\}_k$ consists of nested convex subsets of Q.

Lemma 6.8. *The sequence $\mathcal{D}_k$ converges to $\mathcal{D}_\infty := \mathcal{L}(\lambda, \ell/\ell_c(\lambda), 1)$. Moreover, the Hausdorff distance between $\partial\mathcal{D}_k$ and $\partial\mathcal{D}_\infty$ is upper bounded by c^k for some constant $c \in (0, 1)$.*

Proof. The set $\mathcal{D}_k$ has four symmetric flat pieces where it is tangent to the sides of Q. Let v_k denote the length of one flat piece and write $r_k = (1 - v_k)/2$. Moreover, notice that $2r_k$ is the side of the smallest square one can put around the Wulff body $s_k\mathcal{W}_1$ with $s_k = 2r_k/\ell_\tau$. Simple geometric considerations then show that the sequence r_k satisfies

$$r_{k+1} = r_k(1 - \sqrt{2}\,y/\ell_\tau) + \frac{\ell}{\sqrt{2}}\,y, \qquad r_0 = \frac{1}{2},$$

where y is the radius of the Wulff body $\mathcal{W}_1$ in the direction $\theta = \pi/4$. Set $a = 1-\sqrt{2}\,y/\ell_\tau$ and note that $a \in (0, 1)$. It follows that $r_k = \frac{1}{2}a^k + \frac{\ell}{\sqrt{2}}\,y\sum_{j=0}^{k-1} a^j$. As $k \to \infty$, this converges to $\ell\ell_\tau/2$, which is the value corresponding to the limiting shape $\mathcal{L}(\lambda, \ell/\ell_c(\lambda), 1)$. The Hausdorff distance between $\partial\mathcal{D}_k$ and $\partial\mathcal{D}_{k+1}$ is then of order a^k and the desired conclusion follows. □

Proof of Lemma 6.7. Consider the sets $\mathcal{D}_k$, $k = 0, 1, \ldots,$ defined above. By assumption we know that $L(1-(m-1)\gamma_L)\mathcal{D}_0 \sim \mathcal{W}(x_0, \ell_{x_0})$ is contained w.h.p. in an $H(L)$-contour. We now prove that conditionally on $\mathcal{E}_0(L(1 - (m - 1)\gamma_L - kL^{-2/3+3\varepsilon})\mathcal{D}_k)$, the event $\mathcal{E}_0(L(1 - (m - 1)\gamma_L - (k + 1)L^{-2/3+3\varepsilon})\mathcal{D}_{k+1})$ holds w.h.p.

Fix $\ell = \ell_c(\lambda_0)(1 + \delta_L)$, and consider a droplet $\mathcal{W}(x, \ell)$ such that $\mathcal{W}(x, \ell) \subset \mathcal{D}_k$. From Lemma 6.6, we can let $\mathcal{W}(x, \ell)$ grow up to $\mathcal{W}(x, \ell_x)$. Repeating this at every x as above yields the desired claim since the parameter t_ζ in the definition of $\mathcal{D}_k$ can be identified with ℓ_x/ℓ for $x = \zeta L$. This establishes that under the assumptions of Lemma 6.7, for any k, the set $L(1 - (m - 1)\gamma_L - kL^{-2/3+3\varepsilon})\mathcal{D}_k$ is contained w.h.p. in an $H(L)$-contour. From Lemma 6.8, we know that a number $k = O(\log L)$ of steps suffices to attain a distance of order $1/L$ between $\mathcal{D}_k$ and $\mathcal{D}_\infty$, and therefore w.h.p. $L(1 - m\gamma_L)\mathcal{D}_\infty$ is contained in an $H(L)$-contour. This proves Lemma 6.7. □

Proof of Proposition 6.5. The above two steps provide a proof in the case $n = 0$. The other cases are obtained with exactly the same argument, provided one uses $\lambda^{(n)}$ instead of $\lambda^{(0)} = \lambda$ (we recall that n is fixed as $L \to \infty$). □

Proof of Theorem 6.2(a). Fix $n \in \mathbb{Z}_+$. Suppose that the event

$$\mathcal{E}_{n+1}(L\mathcal{L}(\lambda^{(n)}, 1 + \delta_L, -(H(L) - n - 1)\gamma_L)) \tag{6.3}$$

holds w.h.p. On the latter event, conditioning on the outermost $(H(L)-n-1)$-contour Γ and using monotonicity (the event appearing in Theorem 6.2(a) is increasing), one can assume that there are boundary conditions at height $H(L) - n - 1$ outside the region $\Lambda' = \Lambda_\Gamma \setminus \Delta_\Gamma^+$ which contains the set $L\mathcal{L}(\lambda^{(n)}, 1 + \delta_L, -(H(L) - n - 1)\gamma_L)$. It follows from Proposition 6.5 that w.h.p. $\mathcal{E}_n(L\mathcal{L}(\lambda^{(n)}, 1 + \delta_L, -(H(L) - n)\gamma_L)$ holds. Since $(H(L) - n)\gamma_L \le (\log L)^2 L^{-2/3+3\varepsilon} \le L^{-2/3+4\varepsilon}$ the desired conclusion follows.

Thus, it suffices to prove that (6.3) holds w.h.p. assuming (6.1). We use recursion, starting from the case of the 1-contour. Here one has boundary conditions at height zero outside the $L \times L$ square Λ and floor at 0. By monotonicity one can lower the floor down to height $-(H(L) - n - 1)$. Once this is done the statistic of the 1-contours coincides with the statistic of the $(H(L) - n)$-contours with floor at 0 and boundary conditions at height $H(L) - n - 1$. By Proposition 6.5, with $m = 1$, one infers that there is a 1-contour in the original problem that contains $L\mathcal{L}(\lambda^{(n)}, 1+\delta_L, -\gamma_L)$. Recursively, assume that w.h.p. there exists a k-contour containing $L\mathcal{L}(\lambda^{(n)}, 1 + \delta_L, -k\gamma_L)$. Conditioning on the outermost such contour, using monotonicity one can assume boundary conditions at height k on a set Λ' that contains $L\mathcal{L}(\lambda^{(n)}, 1+\delta_L, -k\gamma_L)$. Repeating the above argument (lowering the floor and using Proposition 6.5) one sees that w.h.p. there exists a $(k + 1)$-contour containing $L\mathcal{L}(\lambda^{(n)}, 1+\delta_L, -(k+1)\gamma_L)$. Once we reach the height $k = H(L) - n - 1$, the proof is complete. □

6.2. Retreat of droplets. We recall that, from Section 4, w.h.p. a macroscopic $(H(L) - n)$-contour exists iff $\lambda^{(n)} > \lambda_c$. In that case it is unique w.h.p. by Theorem 6.2(b).

Theorem 6.9. *Fix $\epsilon, \hat\epsilon \in (0, 1/10)$ and let Λ be a square of side L. Consider the SOS model on Λ with boundary conditions at height zero and floor at height zero. Fix $n \in \mathbb{Z}_+$ and assume $\lambda^{(n)} \ge \lambda_c + \hat\epsilon$. Then w.h.p. as $L \to \infty$ the unique macroscopic $(H(L)-n)$-contour is contained in $L\mathcal{L}(\lambda^{(n)}, t_L, \delta_L)$ (cf. Definition 3.4) with $t_L = 1-\delta_L$ and $\delta_L = L^{-\epsilon/8}$.*

Proof. We use induction on n.

(i) We begin by treating the base case $n = 0$.

Definition 6.10. Given $s \in [0, 1]$ we say that $\mathcal{H}(s)$ holds if w.h.p. there exists a unique macroscopic $H(L)$-contour Γ_0 and it is contained in $L\mathcal{L}(\lambda, s, \delta_L)$.

With this definition the statement of the theorem for $n = 0$ follows from the next two lemmas.

Lemma 6.11 (Base case). *For any s small enough $\mathcal{H}(s)$ holds.*

Lemma 6.12 (Inductive step). *Fix $s \le t_L$. Then $\mathcal{H}(s)$ implies $\mathcal{H}(s + L^{-2/3})$.*

Proof of Lemma 6.11. We actually prove that w.h.p. Γ_0 is contained in $L\mathcal{L}(\lambda, s, 0)$ for s small enough. Fix $s \in (0, 1/4)$ and define T_i, $i = 1, \dots, 4$, as the "curved triangle" delimited by the curved portion of the boundary of $L\mathcal{L}(\lambda, 4s, 0)$ facing the i^{th} corner v_i of Λ and $\partial\Lambda$. Let A, B be the end points of the curved portion of the boundary of T_1 (both at distance $2s\ell_c\ell_\tau L$ from v_1). Let now E_i be the event that inside $T_i \setminus L\mathcal{L}(\lambda, 2s, 0)$ there exists a macroscopic chain where the height of the surface is at least $H(L)$. If the macroscopic $H(L)$-contour Γ_0—which, under π^0_Λ, exists w.h.p. by Proposition 4.9 and is unique by Theorem 6.2—is not contained in $L\mathcal{L}(\lambda, s, 0)$, then necessarily one of the four events E_i occurred. By symmetry, it is therefore enough to show that $\pi^0_\Lambda(E_1) = O(e^{-c(\log L)^2})$ for any s small enough.

For this purpose let us introduce boundary conditions τ on $\partial\Lambda$ as follows:

$$\tau_x = \begin{cases} H(L) & \text{if } x \in \partial\Lambda \setminus \partial T_1, \\ H(L) - 1 & \text{if } x \in \partial\Lambda \cap \partial T_1. \end{cases} \tag{6.4}$$

By monotonicity we can bound $\pi^0_\Lambda(E_1)$ from above by $\pi^\tau_\Lambda(E_1)$. Let Γ be the open $H(L)$-contour Γ joining A, B and denote by G the event that Γ does not get out of $L\mathcal{L}(\lambda, 2s, 0)$. We can once again appeal to [15, Lemma A.2] to get

$$\pi^\tau_\Lambda(E_1 \mid G) \le e^{-c(\log L)^2}.$$

Thus we are left with the proof that, for all s small enough, G occurs w.h.p. As in Lemma 5.9, we can write

$$\pi^\tau_\Lambda(\Gamma) \propto \exp\left(-\beta|\Gamma| + \Psi_\Lambda(\Gamma) + \frac{\lambda}{L}A(\Gamma) + o(L)\right) \tag{6.5}$$

where $A(\Gamma)$ is the signed area of Γ with respect to the segment AB with the obvious choice of the signs. Clearly

$$A(\Gamma) \le 2s^2\ell_c^2\ell_\tau^2 L^2 = 2s^2(\hat{\lambda}/\lambda)^2 L^2 \le 2s^2(\hat{\lambda}/\lambda_c)^2 L^2 \le s^2 L^2$$

for β large enough (cf. Remark 3.7). Thus

$$\begin{aligned} \pi^\tau_\Lambda(G^c) &\le e^{s^2(L+o(L))} \frac{\sum_{\Gamma\in G^c} e^{-\beta|\Gamma|+\Psi_\Lambda(\Gamma)}}{\sum_\Gamma e^{-\beta|\Gamma|+\Psi_\Lambda(\Gamma)+\frac{\lambda}{L}A(\Gamma)}} \\ &\le e^{s^2(L+o(L))} \frac{\sum_{\Gamma\in G^c} e^{-(\beta-e^{-\beta})|\Gamma|+\Psi_{\mathbb{Z}^2}(\Gamma)}}{\sum_\Gamma e^{-(\beta+e^{-\beta})|\Gamma|+\Psi_{\mathbb{Z}^2}(\Gamma)+\frac{\lambda}{L}A(\Gamma)}} \\ &= e^{s^2(L+o(L))} \frac{\sum_{\Gamma\in G^c} e^{-(\beta-e^{-\beta})|\Gamma|+\Psi_{\mathbb{Z}^2}(\Gamma)}}{\sum_\Gamma e^{-(\beta-e^{-\beta})|\Gamma|+\Psi_{\mathbb{Z}^2}(\Gamma)}} \times \frac{\sum_\Gamma e^{-(\beta-e^{-\beta})|\Gamma|+\Psi_{\mathbb{Z}^2}(\Gamma)}}{\sum_\Gamma e^{-(\beta+e^{-\beta})|\Gamma|+\Psi_{\mathbb{Z}^2}(\Gamma)+\frac{\lambda}{L}A(\Gamma)}} \end{aligned} \tag{6.6}$$

where we have used $|\Psi_\Lambda(\Gamma) - \Psi_{\mathbb{Z}^2}(\Gamma)| \le e^{-\beta}|\Gamma|$. By [20, Sec. 4.14] the first ratio on the r.h.s. of (6.6) is bounded from above by $\exp(-csL)$ with c independent of β, since the event G^c implies an excess length of order sL for the contour. By Jensen's inequality with respect to the measure ν on Γ corresponding to the weight $e^{-(\beta-e^{-\beta})|\Gamma|+\Psi_{\mathbb{Z}^2}(\Gamma)}$, the second ratio is bounded from above by

$$\exp\left(2e^{-\beta}\nu(|\Gamma|) - \frac{\lambda}{L}\nu(A(\Gamma))\right).$$

Using once again [20, Ch. 4] we obtain $\nu(|\Gamma|) \le 2sL$ and $\nu(A(\Gamma)) = O((sL)^{3/2})$. Hence $\pi^\tau_\Lambda(G^c) = O(e^{-csL/2})$ for any s small enough independent of L. $\square$

Proof of Lemma 6.12. Let us fix some notation. Let $s' = s + L^{-2/3}$ and, referring to Figure 7 and centering the box Λ at the origin, let

$$f_s : (-(1+\delta_L)L/2, 0] \to [-(1+\delta_L)L/2, 0]$$

be the decreasing convex function whose graph is the south-west quarter of $\partial(L\mathcal{L}(\lambda, s, \delta_L))$. Let $\hat{x}(s)$ be the unique solution of $f_s(x) = x$. We will denote by $x_*(s)$ (resp. $x^*(s)$) the point after which $f_s(\cdot)$ is smaller than $-L/2$ (resp. after which $f_s(\cdot)$ is flat and equals $-L(1+\delta_L)/2$).

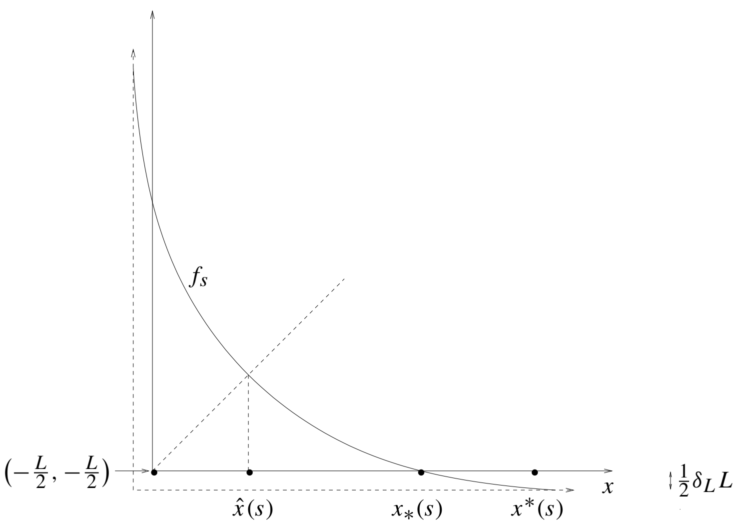

Fig. 7. The graph of the function f_s describing the south-west quarter of $\partial(L\mathcal{L}(\lambda, s, \delta_L))$.

For $x \in [\hat{x}(s), x_*(s')]$ let $x^\pm := x \pm \frac{1}{2}L^{2/3+\epsilon}$ and define

$$Z_s(x) = \frac{1}{2}[f_s(x^-) + f_s(x^+)] - \frac{\lambda}{8\beta(\tau(\theta)+\tau''(\theta))(\cos\theta)^3}L^{1/3+2\epsilon} - \sigma(x,\theta)L^\epsilon \qquad (6.7)$$

with $\theta = \theta_x \in [0, \pi/4]$ such that $\tan\theta = |f'_s(x)|$ and $\sigma^2(x,\theta) = O(L^{2/3+\epsilon})$ is given in Theorem 5.7.

We first observe that, if $s \le t_L$,

$$Z_s(x) \ge f_{s'}(x) \tag{6.8}$$

where the r.h.s. is larger than $-L/2$ because $x \le x_*(s')$. Indeed,

$$x^*(s) - x_*(s') \ge c\sqrt{\delta_L}\, L \quad \text{so that} \quad x^+ < x^*(s),$$

and using Lemma 3.9 together with simple trigonometry we obtain

$$f_s(x) = \frac{1}{2}[f_s(x^-) + f_s(x^+)] - \frac{1}{s(1+\delta_L)}\left[\frac{\lambda}{8\beta(\tau(\theta)+\tau''(\theta))(\cos\theta)^3} L^{1/3+2\epsilon}\right] + o(1). \tag{6.9}$$

Moreover,

$$f_{s'}(x) \le f_s(x) + c(s'-s)L = f_s(x) + cL^{1/3}$$

for some positive constant c. Therefore, if $s \le 1 - \delta_L = 1 - L^{-\epsilon/8}$,

$$\begin{aligned} Z_s(x) - f_{s'}(x) &\ge \frac{\lambda}{8\beta(\tau(\theta)+\tau''(\theta))(\cos\theta)^3} L^{1/3+2\epsilon}\left(\frac{1}{s(1+\delta_L)} - 1\right) - cL^{1/3} \\ &\ge \frac{\lambda}{8\beta(\tau(\theta)+\tau''(\theta))(\cos\theta)^3} L^{1/3+2\epsilon}\delta_L^2 - cL^{1/3}, \end{aligned}$$

which is positive.

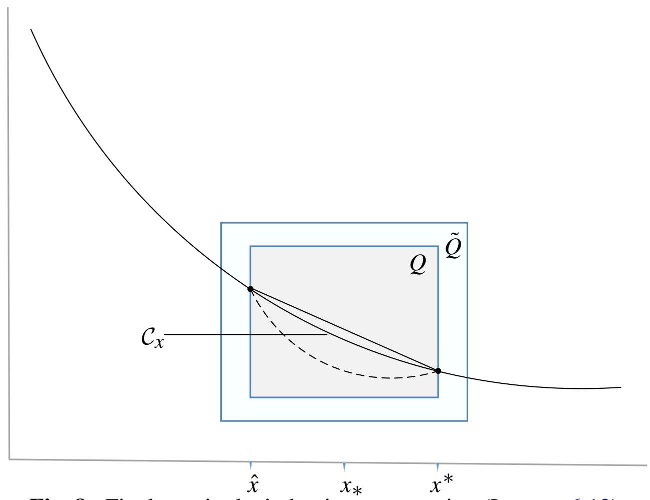

Fig. 8. Final step in the inductive construction (Lemma 6.12).

We are now going to apply the results of Theorem 5.7 (see Definitions 5.6, 5.5 and Figure 3). Consider the rectangle Q_x of horizontal side $L^{2/3+\epsilon}$ and vertical side $2L^{2/3+\epsilon}$ centered at $(x, \frac{1}{2}[f_s(x^-)+f_s(x^+)])$. For simplicity we initially assume that $x \in [\hat{x}(s), x_*(s')]$ is such that $Q_x \subset \Lambda$. Later on we will explain how to treat the general case. Let $\tilde{Q}_x$ denote the $2(\log L)^2$-neighborhood of Q_x and let G_x be the event that there exists a regular circuit $\mathcal{C}_* \in \tilde{Q}_x \setminus Q_x$ such that the height $\eta\restriction_{\mathcal{C}_*}$ on $\mathcal{C}_*$ is not greater than the height $\xi(\mathcal{C}_*, j, a, b)$ given in Definition 5.6 with $a = f_s(x^-) - (\log L)^2$, $b = f_s(x^+) - (\log L)^2$, $j = H(L)$. Here the roles of j, $j-1$ have been interchanged with respect to the setting of Theorem 5.7, i.e. in the present application the boundary conditions are at height j above A, B and $j-1$ below. Define E_x as the event that there exists a chain $\mathcal{C}_x$ of lattice sites satisfying the following conditions:

(i) $\mathcal{C}_x$ connects the points A, B;
(ii) the point $(x, Z_s(x))$ lies below $\mathcal{C}_x$;
(iii) $\eta_y \le H(L) - 1$ for all $y \in \mathcal{C}_x$.

Claim 6.13. *W.h.p. the event E_x occurs.*

Before proving the claim let us conclude the proof of Lemma 6.12. If E_x occurs for all $x \in [\hat{x}(s), x_*(s')]$, then necessarily there exists a chain $\hat{\mathcal{C}}$ (obtained by patching together the individual chains $\mathcal{C}_x$) joining the vertical lines through the points with horizontal coordinates $\hat{x}(s) - \frac{1}{2}L^{2/3+\epsilon}$ and $x_*(s') + \frac{1}{2}L^{2/3+\epsilon}$ and staying above the curve $\mathcal{Z}_s := \{(x, Z_s(x)) : x \in [\hat{x}(s), x_*(s')]\}$ where the surface height is at most $H(L) - 1$. Notice that (6.8) implies that $\mathcal{Z}_s$ is above the curve $f_{s'}$ on the same interval $[\hat{x}(s), x_*(s')]$. By the claim, the above event occurs w.h.p. In view of symmetry with respect to reflection across the south-west diagonal of Λ, the corresponding event occurs in the upper half of the south-west corner of Λ. By patching together the two chains constructed in this way, we have shown that w.h.p. the left vertical boundary of Λ and the bottom horizontal boundary of Λ are connected by a chain of sites $\mathcal{C}$ which stays above the curve $f_{s'}$ such that $\eta_x \le H(L) - 1$ for all $x \in \mathcal{C}$. Since we are assuming $\mathcal{H}(s)$, we know that w.h.p. there exists a unique $H(L)$-contour Γ_0. The contour Γ_0 cannot cross $\mathcal{C}$ so that either $\mathcal{C}$ is contained in Λ_{Γ_0}, the interior of Γ_0, or it is contained in $\Lambda \setminus \Lambda_{\Gamma_0}$. The first case can be excluded since it would produce a macroscopic negative contour in Λ, which has negligible probability by Proposition 2.7. The second case implies that the curve $f_{s'}$ lies outside of Γ_0. The same argument can be repeated for the remaining three corners of the box Λ. That implies that w.h.p. the macroscopic $H(L)$-contour Γ_0 is contained inside $\mathcal{L}(\lambda, s', \delta_L)$. □

Proof of Claim 6.13. To compute the probability of the event G_x defined above, let Ω_s be the event that the unique macroscopic $H(L)$-contour Γ_0 is contained in $L\mathcal{L}(\lambda, s, \delta_L)$. Since we assume $\mathcal{H}(s)$, the event Ω_s occurs w.h.p. On the other hand, conditionally on Ω_s, G_x occurs w.h.p. (cf. Lemma 5.4). In conclusion, G_x occurs w.h.p. We will denote by $\mathcal{C}_*$ the most external circuit characterizing the event G_x.[6]

[6] Of all this complicated construction the reader should just keep in mind the following simplified picture: $\mathcal{C}_* = \partial Q_x$ and the height of the surface is at most $H(L) - 1$ on the portion of ∂Q_x below the curve $f_s(\cdot)$ and at least $H(L)$ above it.

Putting all together we have

$$\pi^0_\Lambda(E_x) \geq \pi^0_\Lambda(E_x \mid G_x) - O(e^{-c(\log L)^2})$$
$$\geq \min_{\mathcal{C}_*} \pi^0_\Lambda\big(E_x \mid \eta\restriction_{\mathcal{C}_*} \leq \xi(\mathcal{C}_*, j, a, b)\big) - O(e^{-c(\log L)^2})$$

with (j, a, b) as above and the minimum taken over all possible regular circuits in $\tilde{Q}_x \setminus Q_x$. Since the event E_x is decreasing, monotonicity gives

$$\min_{\mathcal{C}_*} \pi^0_\Lambda\big(E_x \mid \eta\restriction_{\mathcal{C}_*} \leq \xi(\mathcal{C}_*, j, a, b)\big) \geq \min_{\mathcal{C}_*} \pi^0_\Lambda\Big(E_x \mid \eta\restriction_{\mathcal{C}_*} = \xi(\mathcal{C}_*, j, a, b)b\Big).$$

At this stage we are exactly in the setting of Theorem 5.7 which states that the open $H(L)$-contour inside the region enclosed by $\mathcal{C}_*$ and induced by the boundary conditions $\xi(\mathcal{C}_*, j, a, b)$ goes above the point $(x, Y_s(x))$ w.h.p. The latter event implies the event E_x by the very definition of the $H(L)$-contour. In conclusion,

$$\min_{\mathcal{C}_*} \pi^0_\Lambda\big(E_x \mid \eta\restriction_{\mathcal{C}_*} = \xi(\mathcal{C}_*, j, a, b)\big) = 1 - O(e^{-c(\log L)^2})$$

and E_x occurs w.h.p. □

It remains to consider the case where the rectangle Q_x exits the lower side of Λ, which happens if x is close to $x_*(s')$. In this case we repeat the same reasoning, except that in order to estimate $\pi^0_\Lambda(E_x \mid \eta\restriction_{\mathcal{C}_*} \leq \xi(\mathcal{C}_*, j, a, b))$ from below, we use the domain-enlarging procedure of Lemma 2.15 and we replace Λ with $\Lambda' = \Lambda \cup Q_x$, again with boundary conditions at height zero on $\partial\Lambda'$. The proof then proceeds identically as before and in this case by construction the regular circuit $\mathcal{C}_*$ coincides with ∂Q_x in the portion of ∂Q_x that exits Λ.

(ii) Here we briefly discuss the case $n \geq 1$. As before we denote by Γ_n the (unique w.h.p.) macroscopic $(H(L) - n)$-contour. Assume inductively that

$$L\mathcal{L}(\lambda_{n-1}, t^+_L, -\delta_L) \subset \Gamma_{n-1} \subset L\mathcal{L}(\lambda_{n-1}, t^-_L, \delta_L) \quad \text{w.h.p.} \tag{6.10}$$

where $t^\pm_L = 1 \pm \delta_L$, $\delta_L = L^{-\epsilon/8}$ and $\lambda_0 = \lambda$. Notice that the first inclusion has been proved in Theorem 6.2.

For brevity denote by $\mathcal{G}_{n-1}$ the set of all possible realizations of Γ_{n-1} satisfying (6.10) and, for each $\Gamma \in \mathcal{G}_{n-1}$, denote by V_Γ the interior of Γ. In the base case $n = 1$ the claim (6.10) follows from point **(i)** together with Theorem 6.2. Define $\mathcal{H}^{(n)}(s)$ exactly as $\mathcal{H}(s)$ in Definition 6.10 but with the macroscopic $H(L)$-contour Γ_0 replaced by Γ_n. In order to get the analog of Lemma 6.11 for $\mathcal{H}^{(n)}(s)$, i.e., that $\mathcal{H}^{(n)}(s)$ holds for s small enough, we write

$$\pi^0_\Lambda(\Gamma_n \nsubseteq L\mathcal{L}(\lambda^{(n)}, s, 0)) \leq \max_{\Gamma \in \mathcal{G}_{n-1}} \pi^0_\Lambda\big(\Gamma_n \nsubseteq L\mathcal{L}(\lambda^{(n)}, s, 0) \mid \Gamma_{n-1} = \Gamma\big) + O(e^{-c(\log L)^2}).$$

By monotonicity

$$\pi^0_\Lambda\big(\Gamma_n \nsubseteq L\mathcal{L}(\lambda^{(n)}, s, 0) \mid \Gamma_{n-1} = \Gamma\big) \leq \pi^\xi_{\Lambda^{\text{ext}}_\Gamma}(\Gamma_n \nsubseteq L\mathcal{L}(\lambda^{(n)}, s, 0))$$

where $\Lambda_\Gamma^{\rm ext} = \Lambda \setminus V_\Gamma$ and the boundary height ξ is equal to zero on $\partial\Lambda$ and equal to $H(L) - n$ on $\partial\Lambda_\Gamma^{\rm ext} \setminus \partial\Lambda$. We then proceed as in the proof of Lemma 6.11 for the new setting $(\Lambda_\Gamma^{\rm ext}, \xi)$ with the following modifications:

(i) in the definition of $A, B, \{T_i\}_{i=1}^4$ and $\{E_i\}_{i=1}^4$ the parameter λ is replaced by $\lambda^{(n)}$;
(ii) in the definition (6.4) of the auxiliary boundary conditions τ the height $H(L)$ is replaced by the height $H(L) - n$.

Thus we get

$$\pi^\xi_{\Lambda_\Gamma^{\rm ext}}(\Gamma_n \nsubseteq L\mathcal{L}(\lambda^{(n)}, s, 0)) \le 4\pi^\tau_{\Lambda_\Gamma^{\rm ext}}(E_1)$$

where the measure $\pi^\tau_{\Lambda_\Gamma^{\rm ext}}$ describes the SOS model on $\Lambda_\Gamma^{\rm ext}$, with floor at height zero and boundary conditions τ equal to $H(L)-n-1$ on $\partial\Lambda_\Gamma^{\rm ext} \cap \partial T_1$ and equal to $H(L)-n$ elsewhere. Under $\pi^\tau_{\Lambda_\Gamma^{\rm ext}}$ w.h.p. there is no macroscopic $(H(L)-n+1)$-contour inside $\Lambda_\Gamma^{\rm ext}$ (the argument is as in the proof of Theorem 6.2(b)). We can therefore again write down the distribution of the open $(H(L) - n)$-contour joining the points A, B exactly as in (6.5). The rest of the proof follows step by step the proof of Lemma 6.11; one uses the fact that the distance between the "internal" boundary Γ of $\Lambda_\Gamma^{\rm ext}$ and the segment AB is proportional to L to disregard the possible "pinning interaction" between the open contour and Γ.

Similarly one proves the analog of Lemma 6.12 for $\mathcal{H}^{(n)}(s)$. In conclusion, (6.10) follows for Γ_n and the induction can proceed. □

6.3. Conclusion: proof of Theorem 2. Assume that, along a subsequence L_k, $\lambda(L_k) \to \lambda_\star$. Consider first the case $\lambda_\star > \lambda_c$. Then by Proposition 4.9 one finds that the event $\mathcal{E}_0(L(1+\delta_L)\ell_c(\lambda(L))\mathcal{W}_1)$ holds w.h.p. (see also Remark 6.3). Therefore, from Theorem 6.2, for any $\varepsilon_0 > 0$, the unique macroscopic $H(L)$-contour, say Γ_0, contains the region $L(1-\varepsilon_0)\mathcal{L}_c(\lambda_\star)$ w.h.p. Similarly, Theorem 6.9 shows that Γ_0 is contained in the region $L(1+\varepsilon_0)\mathcal{L}(\lambda_\star)$ w.h.p. Analogous statements hold for the unique macroscopic $(H(L)-n)$-contours Γ_n for all fixed values of $n \in \mathbb{Z}_+$ provided we replace $\lambda_\star$ by $\lambda_\star e^{4\beta n}$. Since the nested limiting shapes $\mathcal{L}_c(\lambda_\star e^{4\beta n})$ converge to the unit square Q as $n \to \infty$, the above statements imply the theorem in the case $\lambda_\star > \lambda_c$.

The case $\lambda_\star < \lambda_c$ uses exactly the same argument, except that here one knows that w.h.p. there is no macroscopic $H(L)$-contour by Proposition 4.6, and the results of Theorems 6.2 and 6.9 can be applied to the macroscopic $(H(L) - n)$-contours with $n \ge 1$ only.

6.4. Proof of Theorem 3. Thanks to Proposition 4.9 and Theorem 6.2 (see Remark 6.4) we know that the distance of the unique macroscopic $H(L)$-contour Γ_0 from the boundary along the flat piece of the limiting shape is w.h.p. not larger than $O(L^{1/3+\epsilon})$ for any fixed $\epsilon > 0$. It remains to prove the lower bound.

Let us write I for the interval $I_\epsilon^{(k)}$ of Theorem 3 and write L instead of L_k. Let $x_i, i = 1, \dots, K$, be a mesh of equally spaced points on I, $x_{i+1} - x_i = 2L^{2/3-\epsilon}$, with x_1 (resp. x_K) the leftmost (resp. rightmost) point in I. Note that K is of order $L^{1/3+\epsilon}$. Let

Q_i, $1 \le i \le K$, be the collection of disjoint squares with side-length $2L^{2/3-\epsilon}$ centered at x_i, let $R_i = Q_i \cap \Lambda$ be the upper half of Q_i and set $\partial^+ R_i = \partial R_i \cap \Lambda$.

Then, under π^0_Λ, w.h.p. the following event A occurs: for every $1 \le i \le K$ there exists a connected chain of sites, within distance $(\log L)^2$ from $\partial^+ R_i$ and touching $\partial\Lambda$ both on the left and on the right of x_i, where the height function η is at most $H(L)$. On the event A, denote by $\mathcal{C}_i$ the most internal such chain and set $\mathcal{C} = \bigcup_i \mathcal{C}_i$.

That A occurs w.h.p. is true by monotonicity: A is decreasing, so we can lift the boundary conditions around Λ to $H(L)$ and then apply Lemma 5.4 to deduce that all dominos in Λ are of negative type (see Definition 5.2 with $j = H(L)$). The existence of the chains $\mathcal{C}_i$ then follows easily (a similar argument was used in the proof of Lemma 6.6).

We want to show that $\max_{x \in I} \rho(x) \ge L^{1/3-\epsilon/2}$ w.h.p. This is a decreasing event. Then, condition on the realization of $\mathcal{C}$ and by monotonicity lift all heights along $\mathcal{C}$ to exactly $H(L)$. This way, the height function in each wiggled rectangle $\tilde{R}_i$ enclosed by $\mathcal{C}_i$ is independent. We therefore concentrate on a single i.

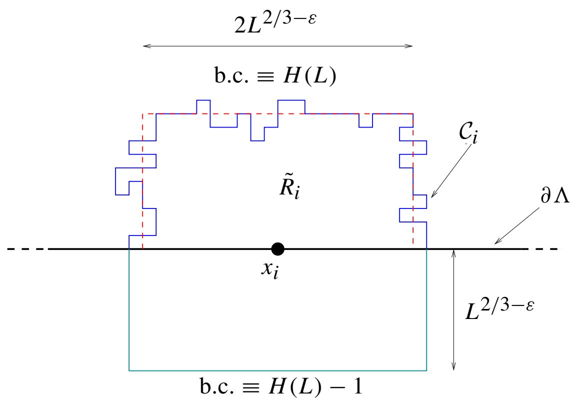

Fig. 9. A point x_i, the rectangle R_i (dashed line), the chain $\mathcal{C}_i$ (wiggled line) enclosing the wiggled rectangle $\tilde{R}_i$. The thick horizontal line is the boundary of Λ. The wiggled square $\tilde{Q}_i$ is obtained by joining to $\tilde{R}_i$ a rectangle of height $L^{2/3-\epsilon}$ and width $2L^{2/3+\epsilon} + O((\log L)^2)$, with two corners coinciding with the endpoints of $\mathcal{C}_i$.

Again by monotonicity, we can apply the domain-enlarging procedure of Lemma 2.15. For this, we enlarge the domain $\tilde{R}_i$ outside Λ as in Figure 9: this way, $\tilde{R}_i$ has been turned into a wiggled square $\tilde{Q}_i$ of side $L^{2/3-\epsilon} + O((\log L)^2)$. Then we consider the SOS measure $\pi^\tau_{\tilde{Q}_i}$ in $\tilde{Q}_i$ with floor at zero and boundary conditions τ that are $\tau \equiv H(L)$ in $\partial\tilde{Q}_i \cap \Lambda$ and $\tau \equiv H(L) - 1$ on $\partial\tilde{Q}_i \setminus \Lambda$. In this situation, we have exactly one open $H(L)$-contour γ. Its law can be written by applying Proposition A.1 to the partition function below and above γ:

$$\pi^\tau_{\tilde{Q}_i}(\gamma) \propto \exp\left[-\beta|\gamma| + \Psi_{\tilde{Q}_i}(\gamma) - \frac{\lambda}{L} A(\gamma) + o(1)\right]$$

with $A(\gamma)$ the signed area of γ with respect to the bottom boundary of Λ (the minus sign in front of the area is due to the fact that we are looking at the area below γ and not above it).

Denote by $P(\cdot)$ the law on γ given by

$$P(\gamma) \propto \exp[-\beta|\gamma| + \Psi_{\tilde{Q}_i}(\gamma)], \tag{6.11}$$

without the area term. From [31] we know that for $L \to \infty$ and rescaling the horizontal (resp. vertical) space direction by $L^{-(2/3-\epsilon)}$ (resp. $L^{-(1/3-\epsilon/2)}$), the law $P(\cdot)$ converges weakly to that of a Brownian bridge on [0, 1], with a suitable diffusion constant. Now let U_i be the event that γ has maximal height less than $L^{1/3-\epsilon/2}$ above the midpoint x_i, and we want to prove

$$\limsup_{L\to\infty} \pi^{\tau}_{\tilde{Q}_i}(U_i) < 1 - \delta \tag{6.12}$$

with $\delta > 0$.

If this is the case, then it follows that w.h.p. there is some $1 \le i \le K \sim cL^{1/3+\epsilon}$ such that the complementary event U_i^c happens, and the proof of the theorem is concluded. Actually, note that (6.12) implies that, if we consider $L^{\epsilon/2}$ "adjacent" points $x_i, x_{i+1}, \dots, x_{i+L^{\epsilon/2}}$, w.h.p. the event U_j^c happens for at least one such j. Since $x_{i+L^{\epsilon/2}} - x_i = O(L^{2/3-\epsilon/2})$, we have proven the stronger version of the theorem, as in Remark 1.3.

It remains to prove (6.12). Write

$$\pi^{\tau}_{\tilde{Q}_i}(U_i^c) = \frac{E(U_i^c; e^{-\frac{\lambda}{L}A(\gamma)+o(1)})}{E(e^{-\frac{\lambda}{L}A(\gamma)+o(1)})}, \tag{6.13}$$

with E the expectation with respect to the law P of (6.11). The denominator is $1 + o(1)$ (just use standard techniques [20, Sec. 4.14] to see that is very unlikely that $|A(\gamma)|$ is much larger than $L^{2/3-\epsilon} \times L^{1/3-\epsilon/2}$, as it should be for a random walk). As for the numerator, Cauchy–Schwarz gives

$$E(U_i^c; e^{-\frac{\lambda}{L}A(\gamma)}) \le \sqrt{P(U_i^c)}\sqrt{E(e^{-2\frac{\lambda}{L}A(\gamma)+o(1)})}.$$

The second term is $1 + o(1)$ like the denominator, while the first factor is uniformly bounded away from 1 since, in the Brownian scaling mentioned above, the event U_i^c becomes the event that the Brownian bridge on [0, 1] is lower than 1 at time 1/2, an event that clearly does not have full probability.

Appendix

A.1. This section contains the proof of Lemma 2.4, showing that for large enough β the $\hat{\pi}$-probability that the height at 0 would exceed h is $(c_\infty + \delta_h)e^{-4\beta h}$ for some $c_\infty = c_\infty(\beta) > 0$ tending to 1 as $\beta \to \infty$.

Proof of Lemma 2.4. Let $a_h = e^{4\beta h}\hat{\pi}(\eta_0 \geq h)$ and define

$$\epsilon_1(\beta, h) = \frac{\hat{\pi}(\eta_0 \geq h,\ S \geq h)}{\hat{\pi}(\eta_0 \geq h,\ S \leq h-1)}, \quad \epsilon_2(\beta, h) = \frac{\hat{\pi}(\eta_0 \geq h-1,\ S \geq h)}{\hat{\pi}(\eta_0 \geq h-1,\ S \leq h-1)},$$

where $S = \max\{\eta_x : x \sim 0\}$ is the maximum height of a neighbor of the origin. With this notation,

$$\frac{a_h}{a_{h-1}} = e^{4\beta} \cdot \frac{1+\epsilon_1(\beta, h)}{1+\epsilon_2(\beta, h)} \cdot \frac{\hat{\pi}(\eta_0 \geq h,\ S \leq h-1)}{\hat{\pi}(\eta_0 \geq h-1,\ S \leq h-1)} = \frac{1+\epsilon_1(\beta, h)}{1+\epsilon_2(\beta, h)}, \tag{A.1}$$

with the last equality due to the fact that $\hat{\pi}(\eta_0 \geq h,\ S \leq h-1)/\hat{\pi}(\eta_0 \geq h-1,\ S \leq h-1) = e^{-4\beta}$. Indeed, this fact easily follows from considering the bijective map T which decreases η_0 by 1 and deducts a cost of 4β from the Hamiltonian of every configuration associated with the numerator compared to its image in the denominator.

In order to bound ϵ_1 and ϵ_2, we note that there exists some absolute constant $c_1 > 0$ independent of β such that, for any $h \geq 1$ and any large enough β,

$$\hat{\pi}(\eta_0 \geq h,\ S \geq h) \leq c_1 e^{-6\beta h}.$$

(This follows from [15, Sec. 7], or alternatively from the proof of [15, Proposition 3.9].) On the other hand, by [15, Proposition 3.9] we know that $\hat{\pi}(\eta_0 \geq h) \geq \frac{1}{2}\exp(-4\beta h)$ for any $h \geq 0$ and $\beta \geq 1$. By combining these with an analogous argument for ϵ_2 we deduce that

$$0 \leq \epsilon_1(\beta, h) \leq c_2 e^{-2\beta h}, \quad 0 \leq \epsilon_2(\beta, h) \leq c_2 \min(e^{-2\beta(h-1)}, e^{-4\beta}),$$

where $c_2 > 0$ is some absolute constant independent of β. Revisiting (A.1) then gives

$$1 - c_2 \min(e^{-2\beta(h-1)}, e^{-4\beta}) \leq a_h/a_{h-1} \leq 1 + c_2 e^{-2\beta h},$$

which readily implies that $c_\infty = \lim_{h\to\infty} a_h$ exists together with $|c_\infty - a_h| \leq c_3 e^{-2\beta h}$. Finally, if we write $c_\infty = a_0 \prod_{h=1}^{\infty}(a_h/a_{h-1})$ and use $\lim_{\beta\to\infty} a_0 = 1$ together with the above bounds we immediately see that $\lim_{\beta\to\infty} c_\infty(\beta) = 1$. □

A.2. Fix $L \in \mathbb{N}$, $\varepsilon > 0$, and consider $\Lambda \subset \mathbb{Z}^2$ with area and external boundary such that

$$|\Lambda| \leq L^{4/3+2\varepsilon}, \quad |\partial\Lambda| \leq L^{2/3+2\varepsilon}. \tag{A.2}$$

Let $Z^{h,\pm}_{\Lambda,U}$ and $\hat{Z}^{\pm}_{\Lambda,U}$ be defined as in Proposition 2.12.

Proposition A.1. *Fix $\beta \geq \beta_0$ and $\varepsilon \in (0, 1/20)$, and assume (A.2) holds. Set $H(L) = \lfloor \frac{1}{4\beta}\log L \rfloor$ and $h = H(L) - n$, $n = 0, 1, \ldots$. Then for all fixed n,*

$$Z^{h,\pm}_{\Lambda,U} = \hat{Z}^{\pm}_{\Lambda,U} \exp(-\hat{\pi}(\eta_0 > h)|\Lambda| + o(1)), \tag{A.3}$$

where $\hat{\pi}$ is the probability obtained as infinite volume limit with boundary conditions at height zero of the SOS model at inverse-temperature β.

Proof. By shifting the heights by $-h$ one can pretend that there are boundary conditions at height zero, that the floor is at $-h$ and that on U the heights are all ≥ 0 (resp. ≤ 0). Let $\omega_\pm$ denote the associated probability measure with no floor, so that

$$Z^{h,\pm}_{\Lambda,U}/\hat{Z}^{\pm}_{\Lambda,U} = \omega_\pm(\eta_x \geq -h,\ x \in \Lambda). \tag{A.4}$$

Consider first the lower bound. Notice that $\omega_\pm$ satisfies the FKG property. Thus

$$\omega_\pm(\eta_x \geq -h,\ x \in \Lambda) \geq \prod_{x\in\Lambda} \omega_\pm(\eta_x \geq -h).$$

As in [15, Sec. 7, (7.19)], for some constant $C > 0$ and all $j \in \mathbb{N}$ one has

$$C^{-1}e^{-4\beta j} \leq \omega_\pm(\eta_x \geq j) \leq Ce^{-4\beta j}. \tag{A.5}$$

In particular, $\omega_\pm(\eta_x < -h) = O(L^{-1})$ for all fixed n, and

$$\omega_\pm(\eta_x \geq -h) = 1 - \omega_\pm(\eta_x < -h) = \exp[-\omega_\pm(\eta_x < -h) + O(L^{-2})].$$

For all $x \in \Lambda$ at distance L^ε from $\partial\Lambda$, one can write $\hat{\pi}(\eta_0 > h)$ instead of $\omega_\pm(\eta_x < -h)$ with an additive error that is $O(L^{-p})$ for any $p > 1$ [15, (7.47)]. By (A.2), this proves the lower bound

$$\omega_\pm(\eta_x \geq -h,\ x \in \Lambda) \geq \exp\bigl(-\hat{\pi}(\eta_0 > h)|\Lambda| + O(L^{-1/3+4\varepsilon})\bigr). \tag{A.6}$$

For the upper bound, we will use essentially the same argument in the proof of [15, Claim 7.7]. We sketch the main steps below. From (A.4), setting $\psi_x = \mathbf{1}_{\{\eta_x < -h\}}$, one writes

$$Z^{h,\pm}_{\Lambda,U}/\hat{Z}^{\pm}_{\Lambda,U} = \omega_\pm\Bigl(\prod_{x\in\Lambda}(1-\psi_x)\Bigr). \tag{A.7}$$

Partition $\mathbb{Z}^2$ into squares P with side $r = L^u + 2L^\kappa$, where $0 < \kappa < u$ will be fixed later (we assume for simplicity that L^u, L^κ are both integers). Consider squares Q of side L^u centered inside the squares P in such a way that each square Q is surrounded within P by a shell of thickness L^κ. The set $\mathcal{S}$ of dual bonds associated to a nonzero height gradient is decomposed into connected components (clusters) S. We let $\mathcal{I}(\kappa)$ be the collection of clusters S in $\mathcal{S}$ such that $|S| \geq L^\kappa$, where $|S|$ denotes the number of edges in S. A point $x \in \Lambda$ can be of four types: 1) those whose distance from the boundary $\partial\Lambda$ is less than L^ε; 2) those that belong to a shell in some $P \setminus Q$; 3) those belonging to a square $Q \subset P$ such that P intersects one of the clusters in $\mathcal{I}(\kappa)$ or its interior; 4) all other $x \in \Lambda$.

We now spell out the contribution of each type of points to (A.7). We estimate by 1 the indicator $1 - \psi_x$ for all x of type 1, 2, or 3:

$$\omega_\pm\Bigl(\prod_{x\in\Lambda}(1-\psi_x)\Bigr) \leq \omega_\pm\Bigl(\prod_{x\in\Lambda'}(1-\psi_x)\Bigr),$$

where Λ' denotes the (random) set of points of type 4. Next, we claim that

$$\omega_\pm\Bigl(\prod_{x\in\Lambda'}(1-\psi_x)\Bigr) \leq \omega_\pm\bigl(\exp[-\hat{\pi}(\eta_0 > h)|\Lambda'| + o(1)]\bigr). \tag{A.8}$$

Once this bound is available one can conclude by showing that

$$\hat{\pi}(\eta_0 > h)\omega_\pm(|\Lambda \setminus \Lambda'|) = o(1). \tag{A.9}$$

Let us first prove (A.9). Recall that $\hat{\pi}(\eta_0 > h) = O(L^{-1})$. Using (A.2), one finds that type 1 points contribute $O(L^{-1/3+4\varepsilon})$ to the l.h.s. of (A.9). Moreover, there are $O(L^{\kappa+u})$ points in each shell and $O(L^{4/3+2\varepsilon-2u})$ shells, therefore type 2 points contribute $O(L^{1/3+2\varepsilon-u+\kappa})$. Exactly as in [15], using the fact that contours in $\mathcal{I}(\kappa)$ have area at least L^κ and at most $L^{4/3+2\varepsilon} = o(L^2)$, one finds that type 3 points only contribute $o(1)$ for any fixed $\kappa > 0$ (cf. [15, (7.53)]). We see that choosing e.g. $u = 1/3 + 4\varepsilon$, $\kappa = \varepsilon$, one deduces that the total contribution of all points of type 1, 2, or 3 is $o(1)$. This proves (A.9).

It remains to prove (A.8). We follow [15]. Any point of type 4 must belong to some square Q that is surrounded by a circuit $\mathcal{C}$ within the shell around Q inside P such that all heights are equal to zero on $\mathcal{C}$. Then, by conditioning on the circuits $\mathcal{C}$ one can proceed by expanding separately the different squares Q. Fix a square Q and let $\hat{\pi}^0_{\mathcal{C}}$ denote the SOS measure with boundary conditions at height zero on the circuit $\mathcal{C}$ surrounding Q. Then [15, (7.56)] yields

$$\hat{\pi}^0_{\mathcal{C}}\Big(\prod_{x\in Q}(1-\psi_x)\Big) \le \exp\Big(-\sum_{x\in Q}\hat{\pi}^0_{\mathcal{C}}(\psi_x) + O(L^{-3/2+2u+c(\beta)}) + O(L^{6u-3})\Big), \tag{A.10}$$

where $c(\beta) \to 0$ as $\beta \to \infty$. Using exponential decay of correlations [12], one can replace $\sum_{x\in Q}\hat{\pi}^0_{\mathcal{C}}(\psi_x)$ by $|Q|\hat{\pi}(\eta_0 > h) + O(L^{u+\delta-1})$ for any $\delta > 0$. Therefore, taking the product over all squares Q containing points of type 4, and taking the average over the realizations of Λ', one finds (A.8), since there are at most $O(L^{4/3+2\varepsilon-2u})$ squares Q in Λ' and $u = 1/3 + 4\varepsilon$, and one can absorb all the errors in the $o(1)$ term if β is large enough. This ends the proof of Proposition A.1. □

A.3. Proof of (2.8). Let $\omega_\pm$ be defined as in Section A.2 above. Let also $\bar{\omega}_\pm$ denote the probability measure $\omega_\pm$ conditioned on the event that there are no macroscopic contours. Then (2.8) becomes equivalent to

$$\bar{\omega}_\pm(\eta_x \ge -h,\ x \in \Lambda) \le \exp\big(-\hat{\pi}(\eta_0 > h)|\Lambda| + O(L^{1/2+c(\beta)})\big). \tag{A.11}$$

We proceed as in the proof of Proposition A.1. The proof of (A.8) now yields

$$\bar{\omega}_\pm\Big(\prod_{x\in\Lambda}(1-\psi_x)\Big) \le \bar{\omega}_\pm\big(\exp[-\hat{\pi}(\eta_0 > h)|\Lambda'| + CL^{1-u+\delta} + CL^{1/2+c(\beta)} + CL^{4u-1}]\big), \tag{A.12}$$

where $C > 0$ is a constant, and $\delta > 0$ is arbitrary. The error terms above are explained as follows: there are at most $O(L^{2-2u})$ squares Q in Λ_0 and for each of those one has a term $O(L^{u+\delta-1})$ coming from the boundary of Q, and a term $O(L^{-3/2+2u+c(\beta)}) + O(L^{6u-3})$ coming from the expansion (A.10).

Next, we need the statement corresponding to (A.9). Thanks to the assumption on the absence of large contours, here there are no points of type 3. Thus the argument behind (A.9) here gives

$$\bar{\omega}_{\pm}(\exp[\hat{\pi}(\eta_0 > h)|\Lambda \setminus \Lambda'|]) \leq \exp[O(L^{-1+\varepsilon}|\partial\Lambda|) + O(L^{2-u+\kappa})]. \tag{A.13}$$

The error terms above are explained as follows: the first term is the worst case contribution of boundary terms (points of type 1, i.e. those at distance at most L^{ε} from $|\partial\Lambda|$); the second term is due to the points of type 2, which are at most $O(L^{2-u+\kappa})$.

Finally, we can combine (A.12) and (A.13). Taking $\kappa = \delta$ sufficiently small, with e.g. $u = 3/5$, using $|\partial\Lambda| = O(L^{1+\varepsilon})$, one finds that the dominant error term is $O(L^{1/2+c(\beta)})$. This implies the desired upper bound.

A.4. Given A and B in $\mathbb{Z}^{2*}$, let $\Xi_{A,B}$ be the set of open contours from A to B that stay within $S_{A,B}$, the infinite strip delimited by the vertical lines going through A and B. Contours are self-avoiding paths, with the usual south-west splitting rule (see e.g. [15, Definition 3.3]), where closed contours are defined). For $\Gamma \in \Xi_{A,B}$, let

$$w(\Gamma) = \exp(-\beta|\Gamma| + \Psi_{S_{A,B}}(\Gamma))$$

where $|\Gamma|$ is the geometric length of Γ. The "decoration term" $\Psi_{S_{A,B}}(\gamma)$ was defined in (4.15) for closed contours: in the present case of an open contour $\Gamma \in \Xi_{A,B}$, it is understood that Λ_γ is the subset of $S_{A,B}$ above Γ, and $\Delta^+_\gamma = \Delta_\gamma \cap \Lambda_\gamma$ (resp. $\Delta^-_\gamma = \Delta_\gamma \cap (S_{A,B} \setminus \Lambda_\gamma)$); see Definition 2.1 for Δ_γ.

The following limit exists [20]:

$$\tau(\theta) = -\lim \frac{1}{\beta|A-B|} \log \sum_{\Gamma \in \Xi_{A,B}} w(\Gamma)$$

where the limit consists in letting $|A - B|$ (the Euclidean distance between A and B) diverge while the angle formed by the segment AB with the horizontal axis tends to $\theta \in (-\pi/2, \pi/2)$.

Remark A.2. As remarked in [20], there is some arbitrariness in the choice of $\Xi_{A,B}$: for instance, one could replace $\Psi_{S_{A,B}}(\Gamma)$ with $\Psi_V(\Gamma)$ for some set V containing an $|A-B|^{1/2+\epsilon}$-neighborhood of the segment AB, and the resulting surface tension would be unchanged.

A.5. We briefly discuss the missing details in the proof of Theorem 5.7 given by the wiggling of the regular circuit $\mathcal{C}_*$. Referring to Figure 10, let $\hat{A}$, $\hat{B}$ be the first and last intersections of Γ with the two vertical lines $\mathbb{L}_1, \mathbb{L}_2$ going through the vertical sides of Q, and let A, B be the points as in Definition 5.6 (see also Figure 3). The contribution to the area term $A(\Gamma)/L$ coming from the parts of Γ before $\hat{A}$ and after $\hat{B}$ is $o(1)$. Moreover the probability that the height difference between $\hat{A}$, A or $\hat{B}$, B is larger that $L^{1/4}$ is smaller than $\exp(-cL^{1/4})$. In fact the circuit $\mathcal{C}_*$, being a regular one, cannot deterministically force such height differences to be larger than $O((\log L)^2)$. Thus a large (of order $L^{1/4}$) height

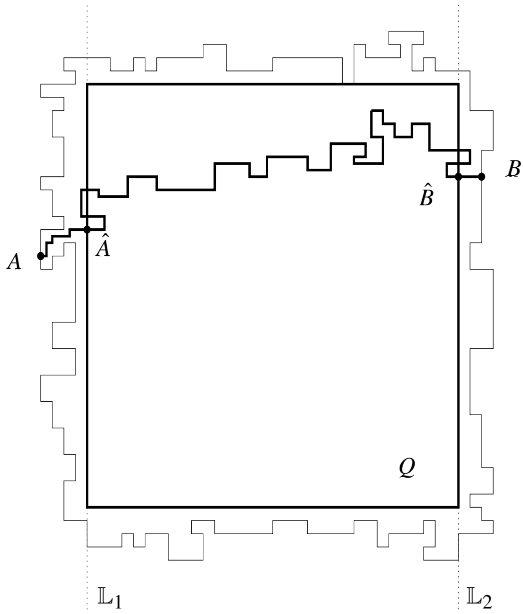

Fig. 10. The rectangle Q, the regular circuit $\mathcal{C}_*$ (closed wiggled line), the contour Γ (thick wiggled line) and the points $\hat{A}$, $\hat{B}$, A, B defined in the text.

difference can only be produced by a large "spontaneous" deviation of the contour Γ. Without the area term such a deviation has probability $O(\exp(-cL^{1/4}))$ (see (5.19)). The area term can be taken care of via an application of the Cauchy–Schwarz inequality: using again (5.19), we see that the average of $\exp(2\lambda A(\Gamma)/L)$ is of order $\exp(L^{3\epsilon})$. Notice that $L^{1/4}$ is negligible with respect to $Y(n) - (a+b)/2 \geq cL^{1/3+2\epsilon}$ defined in Theorem 5.7.

Thus, conditioning on the parts $\Gamma_{\text{left}}, \Gamma_{\text{right}}$ of Γ before $\hat{A}$ and after $\hat{B}$, with $\hat{A}$ at distance $O(L^{1/4})$ from A and similarly for $\hat{B}$, we have reduced ourselves to a geometry to which we can directly apply Corollary 5.13 as in the case where the circuit $\mathcal{C}_*$ is the boundary of Q.

Acknowledgements. We are grateful to Senya Shlosman for valuable discussions on related cluster expansion questions, as well as to Ofer Zeitouni for illuminating discussions on entropic repulsion in the GFF. We thank the referees for a very careful reading and constructive comments. Major parts of this work were carried out during visits to ENS Lyon, Università Roma Tre and the Theory Group of Microsoft Research, Redmond. PC, FM, AS and FLT would like to thank the Theory Group for their hospitality and for creating a stimulating research environment.

This work was supported by the European Research Council through the "Advanced Grant" PTRELSS 228032.

References

[1] Abraham, D. B.: Surface structures and phase transitions—exact results. In: Phase Transitions and Critical Phenomena, Vol. 10, Academic Press, London, 1–74 (1986) MR 0942669

[2] Akutzu, Y., Akutzu, N.: Relationship between the anisotropic interface tension, the scaled interface width and the equilibrium shape in two dimensions. J. Phys. A **19**, 2813–2820 (1986)

[3] Alexander, K. S.: Cube-root boundary fluctuations for droplets in random cluster models. Comm. Math. Phys. **224**, 733–781 (2001) Zbl 0996.82034 MR 1871907

[4] Alexander, K. S., Dunlop, F., Miracle-Solé, S.: Layering and wetting transitions for an SOS interface. J. Statist. Phys. **142**, 524–576 (2011) Zbl 1209.82011 MR 2771044

[5] Baxter, R. J.: Exactly Solved Models in Statistical Mechanics. Academic Press, London (1989) (reprint of the 1982 original) Zbl 0723.60120 MR 0998375

[6] van Beijeren, H.: Exactly solvable model for the roughening transition of a crystal surface. Phys. Rev. Lett. **38**, 993–996 (1977)

[7] Bodineau, T., Ioffe, D., Velenik, Y.: Rigorous probabilistic analysis of equilibrium crystal shapes. J. Math. Phys. **41**, 1033–1098 (2000) Zbl 0977.82013 MR 1757951

[8] Bolthausen, E.: Random walk representations and entropic repulsion for gradient models. In: Infinite Dimensional Stochastic Analysis (Amsterdam, 1999), Verh. Afd. Natuurkd. 1. Reeks. K. Ned. Akad. Wet. 52, R. Neth. Acad. Arts Sci., Amsterdam, 55–83 (2000) Zbl 0991.60093 MR 1831411

[9] Bolthausen, E.: Localization-delocalization phenomena for random interfaces. In: Proc. Int. Congress of Mathematicians, Vol. III (Beijing, 2002), Higher Ed. Press, Beijing, 25–39 (2002) Zbl 1006.60099 MR 1957516

[10] Bolthausen, E., Deuschel, J.-D., Giacomin, G.: Entropic repulsion and the maximum of the two-dimensional harmonic crystal. Ann. Probab. **29**, 1670–1692 (2001) Zbl 1034.82018 MR 1880237

[11] Bolthausen, E., Deuschel, J.-D., Zeitouni, O.: Entropic repulsion of the lattice free field. Comm. Math. Phys. **170**, 417–443 (1995) Zbl 0821.60040 MR 1737994

[12] Brandenberger, R., Wayne, C. E.: Decay of correlations in surface models. J. Statist. Phys. **27**, 425–440 (1982) MR 0659803

[13] Bricmont, J., El Mellouki, A., Fröhlich, J.: Random surfaces in statistical mechanics: roughening, rounding, wetting,. . . . J. Statist. Phys. **42**, 743–798 (1986) MR 0833220

[14] Bricmont, J., Fontaine, J.-R., Lebowitz, J. L.: Surface tension, percolation, and roughening. J. Statist. Phys. **29**, 193–203 (1982) MR 0684948

[15] Caputo, P., Lubetzky, E., Martinelli, F., Sly, A., Toninelli, F. L.: Dynamics of 2+1 dimensional SOS surfaces above a wall: slow mixing induced by entropic repulsion. Ann. Probab. **42**, 1516–1589 (2014) Zbl 1311.60114 MR 3262485

[16] Caputo, P., Lubetzky, E., Martinelli, F., Sly, A., Toninelli, F. L.: The shape of the $(2+1)$-dimensional SOS surface above a wall. C. R. Math. Acad. Sci. Paris **350**, 703–706 (2012) Zbl 1257.82019 MR 2971384

[17] Caputo, P., Velenik, Y.: A note on wetting transition for gradient fields. Stochastic Process. Appl. **87**, 107–113 (2000) Zbl 1045.60102 MR 1751167

[18] Corwin, I., Hammond, A.: Brownian Gibbs property for Airy line ensembles. Invent. Math. **195**, 441–508 (2014) Zbl 06261669 MR 3152753

[19] Deuschel, J.-D., Giacomin, G.: Entropic repulsion for massless fields. Stochastic Process. Appl. **89**, 333–354 (2000) Zbl 1045.60103 MR 1780295

[20] Dobrushin, R., Kotecký, R., Shlosman, S.: Wulff Construction. Transl. Math. Monogr. 104, Amer. Math. Soc., Providence, RI (1992) Zbl 0917.60103 MR 1181197

[21] Eggleston, H. G.: Convexity. Cambridge Tracts in Math. Math. Phys. 47, Cambridge Univ. Press, New York (1958) Zbl 0086.15302 MR 0124813

[22] Ferrari, P. L., Spohn, H.: Constrained Brownian motion: fluctuations away from circular and parabolic barriers. Ann. Probab. **33**, 1302–1325 (2005) Zbl 1082.60071 MR 2150190

[23] Fisher, M. E.: Walks, walls, wetting, and melting. J. Statist. Phys. **34**, 667–729 (1984) Zbl 0589.60098 MR 0751710

[24] Fortuin, C. M., Kasteleyn, P. W., Ginibre, J.: Correlation inequalities on some partially ordered sets. Comm. Math. Phys. **22**, 89–103 (1971) Zbl 0346.06011 MR 0309498

[25] Fröhlich, J., Spencer, T.: Kosterlitz–Thouless transition in the two-dimensional plane rotator and Coulomb gas. Phys. Rev. Lett. **46**, 1006–1009 (1981) MR 0607429

[26] Fröhlich, J., Spencer, T.: The Kosterlitz–Thouless transition in two-dimensional abelian spin systems and the Coulomb gas. Comm. Math. Phys. **81**, 527–602 (1981) MR 0634447

[27] Fröhlich, J., Spencer, T.: The Berežinskiĭ–Kosterlitz–Thouless transition (energy-entropy arguments and renormalization in defect gases). In: Scaling and Self-Similarity in Physics (Bures-sur-Yvette, 1981/1982), Progr. Phys. 7, Birkhäuser Boston, Boston, MA, 29–138 (1983) MR 0733469

[28] Gallavotti, G., Martin-Löf, A., Miracle-Solé, S.: Some problems connected with the description of coexisting phases at low temperatures in the Ising model. In: A. Lenard (ed.), Statistical Mechanics and Mathematical Problems, Lecture Notes in Phys. 20, Springer, 162–204 (1973)

[29] Hammond, A.: Phase separation in random cluster models I: Uniform upper bounds on local deviation. Comm. Math. Phys. **310**, 455–509 (2012) Zbl 1242.82016 MR 2890306

[30] Hammond, A.: Phase separation in random cluster models II: The droplet at equilibrium, and local deviation lower bounds. Ann. Probab. **40**, 921–978 (2012) Zbl 1271.60021 MR 2962083

[31] Hryniv, O.: On local behaviour of the phase separation line in the 2D Ising model. Probab. Theory Related Fields **110**, 91–107 (1998) Zbl 0897.60038 MR 1602044

[32] Hryniv, O., Velenik, Y.: Universality of critical behaviour in a class of recurrent random walks. Probab. Theory Related Fields **130**, 222–258 (2004) Zbl 1078.60034 MR 2093763

[33] Johansson, K.: Discrete polynuclear growth and determinantal processes. Comm. Math. Phys. **242**, 277–329 (2003) Zbl 1031.60084 MR 2018275

[34] Martinelli, F., Toninelli, F. L.: On the mixing time of the 2D stochastic Ising model with "plus" boundary conditions at low temperature. Comm. Math. Phys. **296**, 175–213 (2010) Zbl 1191.82012 MR 2606632

[35] Schonmann, R. H., Shlosman, S. B.: Complete analyticity for 2D Ising completed. Comm. Math. Phys. **170**, 453–482 (1995) Zbl 0821.60097 MR 1334405

[36] Schonmann, R. H., Shlosman, S. B.: Constrained variational problem with applications to the Ising model. J. Statist. Phys. **83**, 867–905 (1996) Zbl 1081.82547 MR 1392417

[37] Sinaĭ, Ya. G.: Theory of Phase Transitions: Rigorous Results. Int. Ser. Natural Philosophy 108, Pergamon Press, Oxford (1982) Zbl 0537.60097 MR 0691854

[38] Temperley, H. N. V.: Statistical mechanics and the partition of numbers. II. The form of crystal surfaces. Proc. Cambridge Philos. Soc. **48**, 683–697 (1952) MR 0053036

[39] Temperley, H. N. V.: Combinatorial problems suggested by the statistical mechanics of domains and of rubber-like molecules. Phys. Rev. (2) **103**, 1–16 (1956) Zbl 0075.20502 MR 0078256

[40] Velenik, Y.: Entropic repulsion of an interface in an external field. Probab. Theory Related Fields **129**, 83–112 (2004) Zbl 1078.60088 MR 2052864

[41] Velenik, Y.: Localization and delocalization of random interfaces. Probab. Surv. **3**, 112–169 (2006) Zbl 1189.82051 MR 2216964