The Annals of Applied Probability
2020, Vol. 30, No. 5, 2439–2464
https://doi.org/10.1214/20-AAP1562

RARITY OF EXTREMAL EDGES IN RANDOM SURFACES AND OTHER THEORETICAL APPLICATIONS OF CLUSTER ALGORITHMS

BY OMRI COHEN-ALLORO[*] AND RON PELED[†]

School of Mathematical Sciences, Tel Aviv University, [*]*omrialloro@gmail.com;* [†]*peledron@tauex.tau.ac.il*

Motivated by questions on the delocalization of random surfaces, we prove that random surfaces satisfying a Lipschitz constraint rarely develop extremal gradients. Previous proofs of this fact relied on reflection positivity and were thus limited to random surfaces defined on highly symmetric graphs, whereas our argument applies to general graphs. Our proof makes use of a cluster algorithm and reflection transformation for random surfaces of the type introduced by Swendsen–Wang, Wolff and Evertz et al. We discuss the general framework for such cluster algorithms, reviewing several particular cases with emphasis on their use in obtaining theoretical results. Two additional applications are presented: A reflection principle for random surfaces and a proof that pair correlations in the spin $O(n)$ model have monotone densities, strengthening Griffiths' first inequality for such correlations.

1. Introduction. Our purpose in this paper is two-fold. First, we consider random surface models satisfying a Lipschitz constraint, that is, random surfaces whose gradients are constrained to be at most 1. For such surfaces we prove that extremal gradients (close to 1 in magnitude) are very unlikely to occur on any given set of edges. This is established for all Lipschitz random surface models whose interaction potential is monotone. The question of controlling the extremal gradients of random surfaces was explicitly asked in [27], Section 6, where such control was a key ingredient in proving that Lipschitz (and more general) random surfaces delocalize in two dimensions. Such a control was achieved in [27] via the use of reflection positivity (through the chessboard estimate) and as such was limited to random surfaces defined on a torus graph. In contrast, our result applies to random surfaces defined on an arbitrary, bounded degree, graph. New delocalization results may be obtained as a consequence as briefly discussed in Section 6. Our proof makes use of a cluster algorithm and reflection transformation for random surfaces of the type introduced by Swendsen–Wang [34], Wolff [40] (see also Brower and Tamayo [8]) and Evertz, Hasenbusch, Lana, Marcu, Pinn and Solomon [19, 24].

Our second goal is to discuss cluster algorithms of the above type in some generality. Such cluster algorithms, commonly used in Monte Carlo simulation of the models, rely on finding a discrete, Ising-type, symmetry in the spin space of the corresponding model (unlike the symmetries used in the reflection-positivity method which are symmetries of the underlying graph on which the model is defined). In Section 2 we discuss their general framework, reviewing in detail the cases of the Potts model, random surfaces, spin $O(n)$ model, Sheffield's cluster-swapping method and reversible Markov chains. Our review emphasizes the use of the algorithms in obtaining theoretical results and we demonstrate such use in two additional applications whose proof via the algorithms is relatively straightforward: A reflection principle for random surfaces and a proof that pair correlations in the spin $O(n)$ model have monotone densities, strengthening Griffiths' first inequality [21, 22] for such correlations.

Received February 2018; revised December 2019.
MSC2020 subject classifications. 82B05, 82B20, 82B41.
Key words and phrases. Random surfaces, extremal gradients, cluster algorithms, Swendsen–Wang algorithm, Wolff algorithm.

We are by no means the first to discuss theoretical applications of cluster algorithms. Many such results are known in the literature including a work of Aizenman [1], following Patrascioiu and Seiler [28], on decay of correlations in Lipschitz spin $O(2)$ models, a work of Burton and Steif [9], Section 2, on characterizing the translation-invariant Gibbs states of a certain subshift of finite type, works of Chayes–Machta [14, 15], Chayes [13] and Campbell–Chayes [10] relating phase transitions of spin systems with percolative properties of the graphical representation defined by their cluster algorithm, Sheffield's cluster swapping algorithm [31], Chapter 8, used in the characterization of translation-invariant gradient Gibbs states of random surfaces (see also van den Berg [37] for a related swapping idea used to study uniqueness of Gibbs measures) and a recent work of Armendáriz, Ferrari and Soprano-Loto [3] on phase transition in the dilute clock model. However, these works mostly make use of ad-hoc transformations suitable to the task at hand and we feel that further emphasis of the unifying framework may still be of interest.

1.1. *Random surfaces.* We begin by introducing the random surface model. Let $G = (V, E)$ be a finite connected graph (all our graphs will be simple, undirected and without self-loops or multiple edges) and $V_0 \subseteq V$ be a nonempty subset of the vertices. Let U be a *potential*, defined to be a measurable function $U : \mathbb{R} \to (-\infty, \infty]$ satisfying $U(x) < \infty$ on a set of positive Lebesgue measure and $U(x) = U(-x)$ for all x. The random surface model with potential U, normalized to be 0 at the subset V_0, is the probability measure μ_{U,G,V_0} on functions $\varphi : V \to \mathbb{R}$ defined by

$$d\mu_{U,G,V_0}(\varphi) := \frac{1}{Z_{U,G,V_0}} \exp\Bigl(-\sum_{\{v,w\}\in E} U(\varphi_v - \varphi_w)\Bigr) \prod_{v\in V_0} \delta_0(d\varphi_v) \prod_{v\in V\setminus V_0} d\varphi_v, \tag{1}$$

where $d\varphi_v$ denotes the Lebesgue measure on φ_v, δ_0 is a Dirac delta measure at 0 and

$$Z_{U,G,V_0} := \int \exp\Bigl(-\sum_{\{v,w\}\in E} U(\varphi_v - \varphi_w)\Bigr) \prod_{v\in V_0} \delta_0(d\varphi_v) \prod_{v\in V\setminus V_0} d\varphi_v$$

which we shall assume satisfies

$$Z_{U,G,V_0} < \infty \tag{2}$$

for μ_{U,G,V_0} to be well defined (the fact that $Z_{U,G,V_0} > 0$ follows from our definition of potential).

For our applications we restrict attention to monotone potentials, when U satisfies

$$U(x) \le U(y), \quad 0 \le x \le y. \tag{3}$$

This assumption implies that the density of a surface increases when its gradients are decreased (in absolute value). In addition, we often consider finitely-supported potentials in the sense that

$$U(x) = \infty, \quad x > 1. \tag{4}$$

This assumption implies that a random surface configuration φ sampled from μ_{U,G,V_0} is a *Lipschitz function*, almost surely, in the sense that

$$|\varphi_v - \varphi_w| \le 1 \quad \text{for all adjacent } v, w. \tag{5}$$

We note that assumption (2), that μ_{U,G,V_0} is well defined, is a consequence of (3) and (4).

An important example of a random surface satisfying (3) and (4) is the case that U is given by the hammock potential,

$$U_{\text{hammock}}(x) = \begin{cases} 0 & |x| \le 1, \\ \infty & |x| > 1. \end{cases} \tag{6}$$

In this case, the random surface is sampled uniformly among all Lipschitz functions normalized to be 0 on V_0.

1.1.1. *Extremal gradients.* Our main result deals with random Lipschitz functions in the sense of (5). How rare are extremal gradients in such surfaces, edges $\{v, w\} \in E$ on which $|\varphi_v - \varphi_w| \geq 1 - \varepsilon$ for some small ε? In [27], Section 6, it was asked whether, under mild assumptions on the potential U, such gradients are exponentially suppressed ("controlled gradients property") in the sense that for each $\delta > 0$ there exists $\varepsilon > 0$, depending only on δ and U (not on the graph G), such that for any distinct edges $\{\{v_i, w_i\}\}_{1 \leq i \leq k} \subseteq E$,

$$\mathbb{P}\big(|\varphi_{v_i} - \varphi_{w_i}| \geq 1 - \varepsilon \text{ for all } i\big) \leq \delta^k.$$

A similar formulation was given for random surfaces with more general potentials. The controlled gradients property was established in [27], for rather general potential functions, when the graph G is a torus in $\mathbb{Z}^d$ (a box with periodic boundary conditions), using reflection positivity (via the chessboard estimate). This property was a key ingredient in showing that two-dimensional random surfaces delocalize for a large class of potential functions including the hammock potential (6). The work [27] continues the delocalization results of Brascamp, Lieb and Lebowitz [5], Section V, and extends the class of potentials treated there, so it is interesting to note that the arguments of [5] relied on a related property [5], inequality (16). Our main result establishes the controlled gradients property for Lipschitz random surfaces with monotone potentials on general, bounded degree, graphs.

THEOREM 1.1. *Let $G = (V, E)$ be a finite connected graph with maximal degree Δ, let $V_0 \subseteq V$ be nonempty, let U be a potential satisfying (3) and (4) and let φ be randomly sampled from μ_{U,G,V_0}. Then for any $0 < \varepsilon \leq \frac{1}{8}$, $k \in \mathbb{N}$ and distinct $\{v_1, w_1\}, \dots, \{v_k, w_k\} \in E$,*

$$\mathbb{P}\big(\{|\varphi_{v_i} - \varphi_{w_i}| \geq 1 - \varepsilon : 1 \leq i \leq k\}\big) \leq \big(C(\Delta)\delta(U, \varepsilon)\big)^{\frac{k}{C(\Delta)}}, \tag{7}$$

where

$$\delta(U, \varepsilon) := \varepsilon \cdot \exp\biggl(-U(1-\varepsilon) + U(0) + \Delta\biggl(U\biggl(\frac{3}{4}\biggr) - U(0)\biggr)\biggr),$$

and where $C(\Delta)$ depends only on Δ.

To illustrate the result we note that when $U = U_{\text{hammock}}$ we have $\delta(U, \varepsilon) = \varepsilon$ and, in addition, that if G is a tree then the probability in (7) exactly equals ε^k.

We note that the dependence on the maximal degree Δ in (7) cannot be completely removed. Indeed, suppose $G = K_{n,n}$ is a complete bipartite graph with partite classes V_1, V_2. Take the boundary set $V_0 = \{v_0\}$ for some $v_0 \in V_1$, take $U = U_{\text{hammock}}$ and let φ be randomly sampled from μ_{U,G,V_0}. It is straightforward to check that for any $0 < \varepsilon < 1$,

$$\begin{aligned} &\mathbb{P}\big(\{|\varphi_{v_i} - \varphi_{w_i}| \geq 1 - \varepsilon : v_i \in V_1, w_i \in V_2\}\big) \\ &\quad \geq \mathbb{P}\biggl(\varphi(V_1) \subseteq \biggl[0, \frac{\varepsilon}{2}\biggr], \varphi(V_2) \subseteq \biggl[1 - \frac{\varepsilon}{2}, 1\biggr]\biggr) \geq \biggl(\frac{\varepsilon}{4}\biggr)^{2n-1}, \end{aligned}$$

with the exponent $2n - 1$ significantly smaller than the amount n^2 of edges between V_1 and V_2 in G.

Notwithstanding the above, we point out that Theorem 1.1 remains true if Δ is replaced by the maximal degree over all vertices other than the vertices of V_0. In fact, inequality (7) holds for a given set of edges $\{v_1, w_1\}, \dots, \{v_k, w_k\}$ when Δ is replaced by the maximal degree of vertices in $\{v_1, w_1, \dots, v_k, w_k\} \setminus V_0$, and this is the phrasing that we shall establish in the proof. This fact will allow us to work with the graph in which the set V_0 is contracted to a single vertex.

In Section 6 we briefly discuss the consequences of Theorem 1.1 to the delocalization of random surfaces.

1.1.2. *Reflection principle for random surfaces*. Let us first remind the reflection principle for simple random walk, a fundamental relation between the distributions of the maximum of the walk and the value at its endpoint. Let (X_j), $0 \le j \le n$, be a simple random walk. That is, $X_0 = 0$ and $\{X_j - X_{j-1}\}$, $1 \le j \le n$, are independent increments, each uniformly distributed on $\{-1, 1\}$. Then the reflection principle states that for all integer k, m satisfying $m \ge \max\{0, k\}$,

$$\mathbb{P}\big(\max\{X_j : 0 \le j \le n\} \ge m, X_n = k\big) = \mathbb{P}(X_n = 2m - k). \tag{8}$$

The law of the maximum of the walk is obtained as an immediate consequence,

$$\begin{aligned} &\mathbb{P}\big(\max\{X_j : 0 \le j \le n\} \ge m\big) \\ &\quad = 2\mathbb{P}(X_n \ge m) - \mathbb{P}(X_n = m) = \mathbb{P}\big(|X_n| \ge m\big) - \mathbb{P}(X_n = m). \end{aligned} \tag{9}$$

In the standard proof of the reflection principle one reflects the final segment of the walk around height m and observes that this is a one-to-one transformation between the events in the two sides of (8). As our main tool in this work is a reflection transformation for random surfaces, one may naturally wonder whether it yields an analogue of (8). This turns out to be the case, as we now proceed to describe. We mention that while our main interest is in random surfaces, the result applies equally well to random walks (having symmetric increments with monotone densities), as these can be seen as random surfaces on a line segment graph, and yields a bound similar to that obtained from Doob's maximal inequality.

We first describe what replaces the maximum in (8). Let $G = (V, E)$ be a graph, $V_0 \subseteq V$ be nonempty and $\varphi : V \to \mathbb{R}$. Let us write $\{V_0 \xleftrightarrow{\varphi<m} v\}$ for the event that there exists a path $v_0, v_1, \dots, v_k$ in G such that $v_0 \in V_0$, $v_k = v$ and $\varphi_{v_i} < m$ for all i. We write $\{V_0 \overset{\varphi<m}{\not\longleftrightarrow} v\}$ for the complementary event, that the "height barrier' between V_0 and v is at least m, meaning that on any path from V_0 to v there is some vertex w with $\varphi_w \ge m$. We similarly define $\{V_0 \xleftrightarrow{\varphi>m} v\}$, etcetera.

Observe that in the one-dimensional case, when $V = \{0, 1, \dots, n\}$ with $E = \{\{i, i+1\} : 0 \le i < n\}$ and $V_0 = \{0\}$, we have $\{V_0 \overset{\varphi<m}{\not\longleftrightarrow} n\} = \{\max\{\varphi_i : 0 \le i \le n\} \ge m\}$, so that our definition generalizes that of the maximum in (8).

THEOREM 1.2. *Let $G = (V, E)$ be a finite connected graph, let $V_0 \subseteq V$ be nonempty, let U be a potential satisfying the monotonicity condition* (3) *and the assumption* (2) *that μ_{U,G,V_0} is well defined. Let φ be randomly sampled from μ_{U,G,V_0}. Then*

$$\frac{1}{2}\mathbb{P}\big(|\varphi_v| \ge m\big) \le \mathbb{P}(V_0 \overset{\varphi<m}{\not\longleftrightarrow} v) \le \mathbb{P}\big(|\varphi_v| \ge m\big) \quad \textit{for all } v \in V, m \ge 0. \tag{10}$$

If, additionally, U satisfies the finite-support condition (4) *then*

$$\mathbb{P}(V_0 \overset{\varphi<m}{\not\longleftrightarrow} v) \ge \mathbb{P}\big(|\varphi_v| \ge m\big) - \mathbb{P}\big(\varphi_v \in (m, m+1)\big) \quad \textit{for all } v \in V, m \ge 0. \tag{11}$$

The above theorem gives an analogue of (9) for random surfaces and our proof proceeds by first establishing an analogue of (8); see Section 3. We remark that the lower bound in (10) is trivial, as $\frac{1}{2}\mathbb{P}(|\varphi_v| \ge m) = \mathbb{P}(\varphi_v \ge m)$ and $\{\varphi_v \ge m\} \subseteq \{V_0 \overset{\varphi<m}{\not\longleftrightarrow} v\}$. The improved lower bound (11) does not hold without additional assumptions (such as (4)) as one can check on the example of the single-edge graph, $V = \{0, 1\}$, $E = \{\{0, 1\}\}$, $V_0 = \{0\}$ and $v = 1$, taking, for example, $U(x) = x^2$ and m large.

A discussion of the relation of the above results to the study of excursion-set percolation of random surfaces appears in Section 6.

1.2. *Spin systems.* Our final result concerns the monotonicity of densities in spin $O(n)$ models and is closely related to an inequality of Armendáriz, Ferrari and Soprano-Loto [3], Lemma 2.4, and Soprano-Loto [33], Section 1.3.6.

Let $G = (V, E)$ be a finite graph and $V_0 \subseteq V$ be a (possibly empty) subset of its vertices. Let $U : [-1, 1] \to (-\infty, \infty]$ be a measurable function. The spin $O(n)$ model with integer $n \geq 1$, potential function U and normalized to equal $e_1 := (1, 0, \dots, 0) \in \mathbb{S}^{n-1}$ at the subset V_0, is the probability measure μ_{U,G,n,V_0} on functions $\varphi : V \to \mathbb{S}^{n-1}$ defined by

$$\begin{aligned} &d\mu_{U,G,n,V_0}(\varphi) \\ &\quad := \frac{1}{Z_{U,G,n,V_0}} \exp\Bigl(-\sum_{\{v,w\}\in E} U\bigl(\langle \varphi_v, \varphi_w\rangle\bigr)\Bigr) \prod_{v\in V\setminus V_0} d\mu_{\mathbb{S}^{n-1}}(\varphi_v) \prod_{v\in V_0} d\delta_{e_1}(\varphi_v), \end{aligned} \tag{12}$$

where $\langle\cdot,\cdot\rangle$ denotes the standard inner product in $\mathbb{R}^n$, $\mu_{\mathbb{S}^{n-1}}$ denotes the uniform measure on $\mathbb{S}^{n-1}$, δ_{e_1} is a Dirac delta measure at e_1 and

$$Z_{U,G,n,V_0} := \int \exp\Bigl(-\sum_{\{v,w\}\in E} U\bigl(\langle \varphi_v, \varphi_w\rangle\bigr)\Bigr) \prod_{v\in V\setminus V_0} d\mu_{\mathbb{S}^{n-1}}(\varphi_v) \prod_{v\in V_0} d\delta_{e_1}(\varphi_v)$$

which we shall assume satisfies

$$0 < Z_{U,G,n,V_0} < \infty \tag{13}$$

for μ_{U,G,n,V_0} to be well defined. The standard spin $O(n)$ model is obtained as the special case where $U(r) = -\beta r$, with β representing the inverse temperature. Special cases of the standard spin $O(n)$ model have names of their own: The case $n = 1$ is the Ising model, the case $n = 2$ is the XY model, or plane rotator model, and the case $n = 3$ is the Heisenberg model.

Observe that when φ is randomly sampled from μ_{U,G,n,V_0}, the distribution of φ_v is absolutely continuous with respect to $\mu_{\mathbb{S}^{n-1}}$ for each $v \in V \setminus V_0$. Denote its density function by d_v, so that $d_v : \mathbb{S}^{n-1} \to [0, \infty)$. Note that d_v is only defined up to a $\mu_{\mathbb{S}^{n-1}}$-null set and that, by symmetry, there is a version of d_v in which $d_v(b)$ is a function of $\langle b, e_1\rangle$. The next theorem states that monotonicity of the potential function implies monotonicity of the densities d_v.

THEOREM 1.3. *Let $G = (V, E)$ be a finite graph and $V_0 \subseteq V$ be a (possibly empty) subset of its vertices. Let $n \geq 1$ be an integer. Suppose that $U : [-1, 1] \to (-\infty, \infty]$ is non-increasing in the sense that*

$$U(r) \geq U(s) \quad \textit{for } r \leq s \tag{14}$$

and that (13) holds. Let φ be randomly sampled from μ_{U,G,n,V_0}. Then for any $v \in V \setminus V_0$, there exists a version of the density d_v satisfying

$$d_v(b_1) \geq d_v(b_2) \quad \textit{when } \langle b_1, e_1\rangle \geq \langle b_2, e_1\rangle. \tag{15}$$

We make a few remarks regarding the theorem: The conclusion of the theorem implies that $\mathbb{E}(\langle \sigma_v, e_1\rangle) \geq 0$ for all $v \in V$, as in Griffiths first inequality [21, 22]. However, we are not aware that the monotonicity of the density has been noted in earlier works, even for the standard ferromagnetic spin $O(n)$ model (when $U(r) = -\beta r$ with $\beta > 0$) with $n \geq 2$.

The result need not hold without the monotonicity condition (14). Indeed, monotonicity is a necessary condition when G is the single-edge graph $V = \{0, 1\}$, $E = \{\{0, 1\}\}$ with $x = 0$, $y = 1$.

An analogous result holds for random surface models of the type (1) as we now state. This result, however, follows easily from convexity considerations as explained in Section 4.

THEOREM 1.4. *Let $G = (V, E)$ be a finite connected graph, let $V_0 \subseteq V$ be nonempty, let U be a potential satisfying the monotonicity condition* (3) *and the assumption* (2) *that μ_{U,G,V_0} is well defined. Let φ be randomly sampled from μ_{U,G,V_0}. Then for any $x \in V$,*

$|\varphi_x|$ *has a nonincreasing density with respect to Lebesgue measure on* $[0, \infty)$.

We mention that analogous results to Theorem 1.3 and Theorem 1.4 remain valid for clock models (spins on equally spaced points of $\mathbb{S}^1$) and *integer-valued* height functions, respectively, and that our proofs apply to this setup without change.

2. Cluster algorithms and reflection transformations. In this section we describe the cluster algorithms and reflection transformations on which our results are based. As mentioned in the Introduction, such ideas are not new, starting with the pioneering works of Swendsen–Wang [34] and Wolff [40], they have been developed by many authors, mostly in the context of fast simulation algorithms (cluster algorithms) but also in theoretical contexts; see [11, 14, 15, 25, 32, 33] for surveys and some recent results. Nevertheless, we believe that there is still room for presenting the special case that we rely upon in some generality, highlighting connections with previous works, to raise further awareness to the general framework and its potential theoretical use. The general description below is followed by specific examples.

Let $(S, \mathcal{S})$ be a measurable space and let $G = (V, E)$ be a finite graph. Let $\vec{E}$ be an arbitrary orientation of the edges, that is, $\vec{E}$ consists of either (v, w) or (w, v), but not both, for each edge $\{v, w\} \in E$. For each vertex $v \in V$, let λ_v be a (finite or infinite) measure on $(S, \mathcal{S})$ and for each directed edge $(v, w) \in \vec{E}$, let $h_{(v,w)} : S \times S \to [0, \infty)$ be a measurable function. The model is defined by the probability measure $\mu_{\lambda,h,G}$ on configurations $\varphi : V \to S$ given by

$$d\mu_{\lambda,h,G}(\varphi) = \frac{1}{Z_{\lambda,h,G}} \prod_{(v,w)\in\vec{E}} h_{(v,w)}(\varphi_v, \varphi_w) \prod_{v\in V} d\lambda_v(\varphi_v), \tag{16}$$

where

$$Z_{\lambda,h,G} = \int \prod_{(v,w)\in\vec{E}} h_{(v,w)}(\varphi_v, \varphi_w) \prod_{v\in V} d\lambda_v(\varphi_v)$$

and we make the assumption that

$$0 < Z_{\lambda,h,G} < \infty. \tag{17}$$

In many of our examples h will be specified on the undirected edge set E but the possibility to define it on directed edges gives the model extra flexibility.

The reflection transformation is based on a function $\tau : S \to S$ with the following properties: For some set $V_0 \subseteq V$,

$$\tau \text{ is an involution:} \quad \tau(\tau(s)) = s \text{ for all } s \in S, \tag{18}$$

$$\tau \text{ preserves } \lambda \text{ for } v \notin V_0\text{:} \quad \lambda_v \circ \tau^{-1} = \lambda_v \text{ for all } v \in V \setminus V_0, \tag{19}$$

$$\tau \text{ preserves } h\text{:} \quad h_{(v,w)}(\tau(a), \tau(b)) = h_{(v,w)}(a, b) \quad \text{for all } (v, w) \in \vec{E}, a, b \in S. \tag{20}$$

Here V_0 plays the role of the "boundary" of G in the sense that we think of (λ_v), $v \in V_0$, which are possibly concentrated on a single value, as prescribing boundary conditions for the measure $\mu_{\lambda,h,G}$. We allow the possibility that $V_0 = \varnothing$ corresponding to free boundary conditions (but in any case we require (17)). We call a τ satisfying the above properties a *reflection*.

The reflection τ identifies "embedded Ising spins" in the model in the following sense. Suppose φ is randomly sampled from $\mu_{\lambda,h,G}$ and define $\psi_v := \{\varphi_v, \tau(\varphi_v)\}$, $v \in V$. Then, conditioned on ψ, each φ_v has at most two possible values, and the joint distribution of these new "binary spins" is that of an Ising model (with general coupling constants, not necessarily of one sign, and zero magnetic field).

Fix a reflection τ. We define a joint probability distribution, the *τ-Edwards–Sokal coupling*, on pairs (φ, ω) where $\varphi : V \to S$ is a configuration and $\omega : E \to \{0, 1\}$ may be thought of as a set of edges, by the following prescription: The marginal distribution of φ is $\mu_{\lambda,h,G}$.

$$\text{Given } \varphi, \text{ the } (\omega_{\{v,w\}}) \text{ are independent and satisfy} \\ \mathbb{P}(\omega_{\{v,w\}} = 1 \mid \varphi) = p_{\{v,w\}}(\varphi),\ \{v, w\} \in E, \tag{21}$$

where, letting $(v, w) \in \vec{E}$ be the directed version of $\{v, w\}$,

$$p_{\{v,w\}}(\varphi) := \max\left(1 - \frac{h_{(v,w)}(\tau(\varphi_v), \varphi_w)}{h_{(v,w)}(\varphi_v, \varphi_w)}, 0\right), \tag{22}$$

where we set $\frac{0}{0} := 1$, $\frac{t}{0} = \infty$ for $t > 0$ and we note that $h_{(v,w)}(\tau(\varphi_v), \varphi_w)$ equals $h_{(v,w)}(\varphi_v, \tau(\varphi_w))$ due to the assumptions (18) and (20) so the latter expression can be used instead of the former in (22). Unfortunately, the marginal distribution of ω does not seem to have a simple formula in general. We note for later use the following immediate property,

$$\text{Conditioned on } \varphi,\ \omega_{\{v,w\}} = 0 \text{ almost surely for each } (v, w) \in \vec{E} \\ \text{with } h_{(v,w)}(\tau(\varphi_v), \varphi_w) \geq h_{(v,w)}(\varphi_v, \varphi_w). \tag{23}$$

In particular, if $\tau(\varphi_v) = \varphi_v$ then $\omega_{\{v,w\}} = 0$ for all edges $\{v, w\}$ incident to v. We also observe that if $h_{(v,w)}$ takes values in $\{0, 1\}$, as in the case of the hammock potential (see (6)), then $p_{\{v,w\}}(\varphi)$ also belongs to $\{0, 1\}$, almost surely, so that ω is a deterministic function of φ.

We proceed to describe the reflection transformation. Let $(\varphi, \omega) \in S^V \times \{0, 1\}^E$ and let $x \in V$. Write $x \overset{\omega}{\longleftrightarrow} v$ if there is a path from x to v with $\omega_{\{u,u'\}} = 1$ for all edges $\{u, u'\}$ along the path. Similarly write $x \overset{\omega}{\longleftrightarrow} V_0$ if there is some $v_0 \in V_0$ with $x \overset{\omega}{\longleftrightarrow} v_0$ and write $x \overset{\omega}{\not\longleftrightarrow} v$ or $x \overset{\omega}{\not\longleftrightarrow} V_0$ for the nonexistence of such paths. The reflected configuration $\varphi^{\omega,x} : V \to S$ is defined by:

$$\text{If } x \overset{\omega}{\longleftrightarrow} V_0 \text{ then } \varphi^{\omega,x} := \varphi. \text{ Otherwise, } \varphi^{\omega,x}_v := \begin{cases} \tau(\varphi_v) & x \overset{\omega}{\longleftrightarrow} v, \\ \varphi_v & x \overset{\omega}{\not\longleftrightarrow} v. \end{cases} \tag{24}$$

That is, $\varphi^{\omega,x}$ is formed by applying τ to all vertices in the ω-connected component of x, unless this connected component intersects V_0 in which case $\varphi^{\omega,x} = \varphi$. The following lemma shows that this transformation preserves the distribution of the τ-Edwards–Sokal coupling.

LEMMA 2.1. *Let (φ, ω) be randomly sampled from the τ-Edwards–Sokal coupling. For each $x \in V$,*

$$\left(\varphi^{\omega,x}, \omega\right) \text{ has the same distribution as } (\varphi, \omega). \tag{25}$$

Of course, the equality in distribution (25) implies also that $\varphi^{\omega,x}$ has the same distribution as φ, leading to a natural Markov chain on configurations. In the context of the spin $O(n)$ model (see also below), the fact that τ is applied to a *single* connected component of ω in each update is one of the innovations introduced by Wolff in his pioneering work [40]. We remark, however, that the equality in distribution (25) remains true when the vertex x is chosen as a function of ω. That is, for any function $x : \{0, 1\}^E \to V$,

$$\left(\varphi^{\omega,x(\omega)}, \omega\right) \text{ has the same distribution as } (\varphi, \omega).$$

More generally, (φ, ω) has the same distribution as (ψ, ω) where ω determines whether $\psi = \varphi$ or $\psi = \varphi^{\omega, x(\omega)}$. These facts can be deduced in a straightforward manner from (25) itself. By composing several operations of this type one may define various other measure-preserving transformations. For instance, in a Swendsen–Wang-type update, one applies τ with probability $1/2$ independently to each of the ω-connected components which do not intersect V_0. Another possibility, when V_0 is a singleton, is to either apply τ to the ω-connected component of x, or to the complement of this connected component, according to whether the component intersects V_0. We use a variant of this latter choice in Section 5 below.

One typical use of the reflection transformation described above and Lemma 2.1 is to define a reversible Markov chain (a "cluster algorithm") on the set of configurations with stationary distribution $\mu_{\lambda,h,G}$. A step of this chain starting at φ is conducted by deciding on some $x \in V$ (possibly randomly), then sampling $\omega : E \to \{0, 1\}$ from the conditional distribution (21) and finally outputting $\varphi^{\omega,x}$. Such Markov chains sometimes mix faster than the more traditional Glauber dynamics, especially near critical points of the model, and are used in practice in Monte Carlo simulations of the model (e.g., for Ising and spin $O(n)$ models following Swendsen–Wang [34] and Wolff [40]. See [23, 36] for recent polynomial-time mixing bounds). Our emphasis, however, will be on theoretical applications.

The reflection transformation is defined above on a finite graph. It is also natural to work on infinite graphs, with the configuration φ sampled from a *Gibbs measure*, which is specified on finite graphs by distributions of the form (16), and with the distribution of ω given φ specified by (21) and (22). In this case it may be shown, using Lemma 2.1, that reflections of *finite* ω-connected components preserve the joint distribution of (φ, ω). This may fail, however, when reflecting *infinite* ω-connected components. Still, it may be shown that reflections of infinite ω-connected components transform the distribution of φ to that of another Gibbs measure, with the same underlying specification, while the distribution of ω given φ continues to be specified by (21) and (22) (in the context of the cluster swapping reflection applied to random surfaces, see below, this was shown by Sheffield [31]; see also [12], Lemma 4.2). We do not discuss this further here.

PROOF OF LEMMA 2.1. It is sufficient to prove that

$$\mathbb{P}\Big(\bigcap_{v\in V}\{\varphi_v^{\omega,x} \in A_v\} \cap \{\omega = \omega_0\}\Big) = \mathbb{P}\Big(\bigcap_{v\in V}\{\varphi_v \in A_v\} \cap \{\omega = \omega_0\}\Big), \tag{26}$$

for all choices of $A_v \in \mathcal{S}$ for $v \in V$ and $\omega_0 : E \to \{0, 1\}$. Fix such (A_v) and ω_0. For brevity, we introduce the notation

$$f_{(v,w)}(\varphi) := h_{(v,w)}(\varphi_v, \varphi_w) p_{\{v,w\}}(\varphi)^{\omega_0(\{v,w\})}\big(1 - p_{\{v,w\}}(\varphi)\big)^{1-\omega_0(\{v,w\})}, \quad (v, w) \in \vec{E},$$

with $0^0 := 1$. Our definitions yield the following formula for the right-hand side of (26),

$$\mathbb{P}\Big(\bigcap_{v\in V}\{\varphi_v \in A_v\} \cap \{\omega = \omega_0\}\Big) = \frac{1}{Z_{\lambda,h,G}} \int \prod_{v\in V} d\lambda_v(\varphi_v)\mathbf{1}_{\varphi_v\in A_v} \prod_{(v,w)\in\vec{E}} f_{(v,w)}(\varphi). \tag{27}$$

We proceed to evaluate the left-hand side of (26). Define $V_{\omega_0,x}$ to be the set of vertices on which τ is applied in the definition of $\varphi^{\omega_0,x}$. Precisely,

$$\text{If } x \overset{\omega}{\leftrightarrow} V_0 \text{ then } V_{\omega_0,x} := \varnothing. \text{ Otherwise, } V_{\omega_0,x} := \{v \in V : x \overset{\omega}{\leftrightarrow} v\}.$$

With this definition,

$$\bigcap_{v\in V}\{\varphi_v^{\omega,x} \in A_v\} \cap \{\omega = \omega_0\} = \bigcap_{v\in V_{\omega_0,x}} \{\tau(\varphi_v) \in A_v\} \bigcap_{v\in V\setminus V_{\omega_0,x}} \{\varphi_v \in A_v\} \cap \{\omega = \omega_0\},$$

whence the left-hand side of (26) satisfies

$$\mathbb{P}\Big(\bigcap_{v\in V}\{\varphi_v^{\omega,x}\in A_v\}\cap\{\omega=\omega_0\}\Big)$$
$$=\frac{1}{Z_{\lambda,h,G}}\int\prod_{v\in V}d\lambda_v(\varphi_v)\prod_{v\in V_{\omega_0,x}}\mathbf{1}_{\tau(\varphi_v)\in A_v}\prod_{v\in V\setminus V_{\omega_0,x}}\mathbf{1}_{\varphi_v\in A_v}\prod_{(v,w)\in\vec{E}}f_{(v,w)}(\varphi).$$

To simplify the last integral, we make the change of variables $\varphi\mapsto\psi$ where

$$\psi_v:=\begin{cases}\tau(\varphi_v) & v\in V_{\omega_0,x},\\ \varphi_v & v\notin V_{\omega_0,x}.\end{cases}$$

This mapping is one-to-one as τ is invertible by (18). In addition, it preserves the measure $\prod_{v\in V}\lambda_v$ by (19) and the fact that $V_0\cap V_{\omega_0,x}=\varnothing$. Thus,

$$\mathbb{P}\Big(\bigcap_{v\in V}\{\varphi_v^{\omega,x}\in A_v\}\cap\{\omega=\omega_0\}\Big)=\frac{1}{Z_{\lambda,h,G}}\int\prod_{v\in V}d\lambda_v(\psi_v)\mathbf{1}_{\psi_v\in A_v}\prod_{(v,w)\in\vec{E}}f_{(v,w)}(\varphi).\tag{28}$$

Comparing (27) and (28) we see that (26) is a consequence of

$$f_{(v,w)}(\varphi)=f_{(v,w)}(\psi)\tag{29}$$

for all $(v,w)\in\vec{E}$. The equality (29) is trivial in the case that $v,w\notin V_{\omega_0,x}$. In the case that $v,w\in V_{\omega_0,x}$ it follows from (20) by using that

$$h_{(v,w)}(\psi_v,\psi_w)=h_{(v,w)}\big(\tau(\varphi_v),\tau(\varphi_w)\big)=h_{(v,w)}(\varphi_v,\varphi_w),$$
$$p_{\{v,w\}}(\psi)=\max\Big(1-\frac{h_{(v,w)}(\tau(\tau(\varphi_v)),\tau(\varphi_w))}{h_{(v,w)}(\tau(\varphi_v),\tau(\varphi_w))},0\Big)$$
$$=\max\Big(1-\frac{h_{(v,w)}(\tau(\varphi_v),\varphi_w)}{h_{(v,w)}(\varphi_v,\varphi_w)},0\Big)=p_{\{v,w\}}(\varphi).$$

Lastly, in the case that, say, $v\in V_{\omega_0,x}$ and $w\notin V_{\omega_0,x}$, using now the involution property (18),

$$f_{(v,w)}(\psi)=h_{(v,w)}(\psi_v,\psi_w)\big(1-p_{\{v,w\}}(\psi)\big)$$
$$=h_{(v,w)}\big(\tau(\varphi_v),\varphi_w\big)\min\Big(\frac{h_{(v,w)}(\tau(\tau(\varphi_v)),\varphi_w)}{h_{(v,w)}(\tau(\varphi_v),\varphi_w)},1\Big)$$
$$=h_{(v,w)}\big(\tau(\varphi_v),\varphi_w\big)\min\Big(\frac{h_{(v,w)}(\varphi_v,\varphi_w)}{h_{(v,w)}(\tau(\varphi_v),\varphi_w)},1\Big)$$
$$\overset{(*)}{=}h_{(v,w)}(\varphi_v,\varphi_w)\min\Big(\frac{h_{(v,w)}(\tau(\varphi_v),\varphi_w)}{h_{(v,w)}(\varphi_v,\varphi_w)},1\Big)$$
$$=h_{(v,w)}(\varphi_v,\varphi_w)\big(1-p_{\{v,w\}}(\varphi)\big)=f_{(v,w)}(\varphi),$$

where the equality $(*)$ follows by separately considering the two cases $h_{(v,w)}(\tau(\varphi_v),\varphi_w)\geq h_{(v,w)}(\varphi_v,\varphi_w)$ and $h_{(v,w)}(\tau(\varphi_v),\varphi_w)<h_{(v,w)}(\varphi_v,\varphi_w)$. This finishes the proof of the lemma. □

We now illustrate the general construction above with specific examples:

Potts model: The q-state Potts model (with free boundary conditions) is obtained by taking $S=\{1,2,\dots,q\}$ (with the discrete sigma algebra), all λ_v equal to the counting measure, $V_0=\varnothing$ and

$$h_{\{v,w\}}(a,b)=\exp(\beta\delta_{a,b}),\quad \{v,w\}\in E,a,b\in S,$$

where $\delta_{a,b}$ is the Kronecker delta function. Let $\tau : S \to S$ be any involution. One checks simply that τ satisfies (19) and (20) so that τ is a reflection.

The model is called ferromagnetic when the parameter β is nonnegative. In this case, the general prescription (21) for the distribution of ω becomes

$$\mathbb{P}(\omega_{\{v,w\}} = 1 \mid \varphi) = \begin{cases} 1 - \exp(-\beta) & \varphi_v = \varphi_w \text{ and } \tau(\varphi_v) \neq \varphi_v, \\ 0 & \text{otherwise.} \end{cases}$$

In particular, the value of φ is constant on each ω-connected component. Applying τ to the values of φ on ω-connected components leads to a variant of the Swendsen–Wang Markov chain (the original Markov chain is obtained by setting $\omega_{\{v,w\}} = 1$ with probability $1 - \exp(-\beta)$, independently, for each $\{v, w\} \in E$ with $\varphi_v = \varphi_w$ and $\omega_{\{v,w\}} = 0$ for other edges. Then updating φ by assigning it a uniform value in S on each ω-connected component, independently. In this case the marginal distribution of ω is explicit and given by the q-random cluster model).

The model is called anti-ferromagnetic when β is negative. In this case, the general prescription (21) becomes

$$\mathbb{P}(\omega_{\{v,w\}} = 1 \mid \varphi) = \begin{cases} 1 - \exp(\beta) & \tau(\varphi_v) = \varphi_w \text{ and } \varphi_v \neq \varphi_w, \\ 0 & \text{otherwise.} \end{cases}$$

In particular, on each path in the subgraph given by $\omega^{-1}(1)$ the value of φ alternates between two distinct values $a, b \in S$ with $b = \tau(a)$. We note in passing that in the limiting case $\beta = -\infty$, corresponding to φ being a uniformly sampled proper q-coloring, the ω-connected components are exactly the Kempe chains of the pairs $a, b \in S$ with $b = \tau(a)$. In the anti-ferromagnetic case, applying τ to the values of φ on ω-connected components leads to the Wang–Swendsen–Kotecký [39] Markov chain.

Random surfaces: The random surface measure μ_{U,G,V_0} defined in (1), having potential U and normalized to be 0 on V_0, fits the framework (16) by taking $S = \mathbb{R}$ (with the Borel sigma algebra), λ_v to be Lebesgue measure for $v \notin V_0$, $\lambda_v = \delta_0$ for $v \in V_0$ and $h_{\{v,w\}}(a, b) = \exp(-U(a - b))$ for all $\{v, w\} \in E$.

For each $m \in \mathbb{R}$ let $\tau_m : \mathbb{R} \to \mathbb{R}$ be the "reflection around m" mapping. That is,

$$\tau_m(a) = 2m - a. \tag{30}$$

It is straightforward to check the conditions (18), (19) and (20), using that $U(x) = U(-x)$ for all x, and conclude that τ_m is a reflection for any random surface measure μ_{U,G,V_0}. In this case, the general prescription (21) for the distribution of ω becomes

$$\mathbb{P}(\omega_{\{v,w\}} = 1 \mid \varphi) = \max\bigl(1 - \exp\bigl(U(\varphi_v - \varphi_w) - U(2m - \varphi_v - \varphi_w)\bigr), 0\bigr). \tag{31}$$

The following consequence plays a central role in the proofs of our main theorems:

$$\begin{aligned} &\text{If } U \text{ is monotone in the sense of (3) then, almost surely,} \\ &\text{on each } \omega\text{-connected component } C \\ &\text{either } \varphi_v > m \text{ for all } v \in C, \text{ or } \varphi_v < m \text{ for all } v \in C, \text{ or } C = \{v\} \text{ and } \varphi_v = m. \end{aligned} \tag{32}$$

This follows from (31) by noting that if U is monotone and $\varphi_v \geq m \geq \varphi_w$ then $\varphi_v - \varphi_w \geq |2m - \varphi_v - \varphi_w|$ whence $U(\varphi_v - \varphi_w) \geq U(2m - \varphi_v - \varphi_w)$.

Extensions of the above ideas to integer-valued random surfaces (when a measure on $\varphi : S \to \mathbb{Z}$ is defined analogously) follow in a similar manner, with the reflection height m restricted to $\mathbb{Z} \cup (\mathbb{Z} + \frac{1}{2})$. The reflection principle for simple random walk (see (8)) can be seen as a reflection transformation of an integer-valued random surface on a one-dimensional graph (see also the discussion of reflection transformations for Markov chains below). Markov chain algorithms for simulating random surfaces based on the above ideas were developed by Evertz, Hasenbusch, Lana, Marcu, Pinn and Solomon [19, 24].

We are not aware of a simple expression for the marginal distribution of ω.

Spin $O(n)$ model: The spin $O(n)$ measure μ_{U,G,n,V_0} defined in (12), having integer $n \geq 1$, potential U and normalized to equal $e_1 = (1, 0, \dots, 0)$ on V_0, fits the framework (16) by taking $S = \mathbb{S}^{n-1} \subseteq \mathbb{R}^n$ (with the sigma algebra inherited from $\mathbb{R}^n$), $d\lambda_v = d\varphi_v$ for $v \in V \setminus V_0$, $\lambda_v = \delta_{e_1}$ for $v \in V_0$ and $h_{\{v,w\}}(a, b) = \exp(-U(\langle a, b\rangle))$ for all $\{v, w\} \in E$.

For each $a \in \mathbb{S}^{n-1}$ let $\tau_a : \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$ be the "reflection around the hyperplane orthogonal to a" mapping. That is,

$$\tau_a(b) = b - 2\langle a, b\rangle \cdot a. \tag{33}$$

It is straightforward to check the conditions (18), (19) and (20), verifying along the way that τ_a is an isometry, and conclude that τ_a is a reflection for any spin $O(n)$ measure μ_{U,G,n,V_0}. In this case, the general prescription (21) for the distribution of ω becomes

$$\mathbb{P}(\omega_{\{v,w\}} = 1 \mid \varphi) = \max\bigl(1 - \exp\bigl(U\bigl(\langle\varphi_v, \varphi_w\rangle\bigr) - U\bigl(\langle\varphi_v - 2\langle a, \varphi_v\rangle a, \varphi_w\rangle\bigr)\bigr), 0\bigr). \tag{34}$$

Again, the following consequence plays a central role in our proof of Theorem 1.3:

$$\begin{aligned} &\text{If } U \text{ is nonincreasing in the sense of (14) then, almost surely,}\\ &\text{on each } \omega\text{-connected component } C\\ &\text{either } \langle a, \varphi_v\rangle > 0 \text{ for all } v \in C, \text{ or } \langle a, \varphi_v\rangle < 0\\ &\text{for all } v \in C, \text{ or } C = \{v\} \text{ and } \langle a, \varphi_v\rangle = 0. \end{aligned} \tag{35}$$

This follows from (34) by noting that if U is nonincreasing and $\langle a, \varphi_v\rangle\langle a, \varphi_w\rangle \leq 0$ then $\langle\varphi_v, \varphi_w\rangle \leq \langle\varphi_v - 2\langle a, \varphi_v\rangle a, \varphi_w\rangle$ whence $U(\langle\varphi_v, \varphi_w\rangle) \geq U(\langle\varphi_v - 2\langle a, \varphi_v\rangle a, \varphi_w\rangle)$.

Extensions of the above ideas to clock models (when φ takes values in a set of q equally-spaced marks on $\mathbb{S}^1$) follow in a similar manner, with the vector a restricted to be one of the marks or exactly in between two marks. Wolff's algorithm [40] pioneered the use of the above ideas to fast simulation algorithms for the spin $O(n)$ model.

We are not aware of a simple expression for the marginal distribution of ω. Nevertheless, Chayes [13], Chayes–Campbell [10], following Chayes–Machta [14, 15], have considered the standard spin $O(n)$ model with $n \in \{2, 3\}$ and proved that the distribution of ω has positive association (every two monotone increasing functions of ω are nonnegatively correlated), that an infinite ω-connected component (in a suitable infinite-volume limit) arises if and only if there is positive magnetization in the spin model and related results.

Cluster swapping: The term cluster swapping was coined by Sheffield [31], Chapter 8, for the following setup, in the special setting of random surfaces. A related swapping idea was used by van den Berg [37] to study uniqueness of Gibbs measures (see also van den Berg and Steif [38], Proof of Theorem 2.4). Let $V_0 \subseteq V$. Let $\mu_{\lambda^1,h,G}$, $\mu_{\lambda^2,h,G}$ be general measures of the type (16), with the same h and with $\lambda^1_v = \lambda^2_v$ for all $v \in V \setminus V_0$. Let φ^1, φ^2 be independent samples from $\mu_{\lambda^1,h,G}$ and $\mu_{\lambda^2,h,G}$ respectively. We regard the pair (φ^1, φ^2) as a configuration $(\varphi^1, \varphi^2) : V \to S \times S$ which is sampled from the measure $\mu_{\lambda^1\times\lambda^2,h\times h,G}$, where

$$\bigl(\lambda^1 \times \lambda^2\bigr)_v := \lambda^1_v \times \lambda^2_v, \quad v \in V$$

and

$$\begin{aligned} &(h \times h)_{(v,w)}\bigl((a_1, a_2), (b_1, b_2)\bigr)\\ &\quad := h_{(v,w)}(a_1, b_1) \cdot h_{(v,w)}(a_2, b_2), \quad (v, w) \in \vec{E}, a_1, a_2, b_1, b_2 \in S. \end{aligned}$$

Let $\tau : S \times S \to S \times S$ be the "swap" mapping, defined by

$$\tau\big((a_1, a_2)\big) = (a_2, a_1). \tag{36}$$

It is straightforward to check the conditions (18), (19) and (20) and conclude that τ is a reflection for $\mu_{\lambda^1 \times \lambda^2, h \times h, G}$. In this case, the general prescription (21) for the distribution of ω becomes

$$\mathbb{P}(\omega_{\{v,w\}} = 1 \mid \varphi) = \max\left(1 - \frac{h_{(v,w)}(\varphi_v^2, \varphi_w^1) \cdot h_{(v,w)}(\varphi_v^1, \varphi_w^2)}{h_{(v,w)}(\varphi_v^1, \varphi_w^1) \cdot h_{(v,w)}(\varphi_v^2, \varphi_w^2)}, 0\right). \tag{37}$$

We see that, as remarked before, if $\varphi_v^1 = \varphi_v^2$ then $\omega_{\{v,w\}} = 0$ for all w adjacent to v. Thus, an ω-connected component is 'blocked' by places where the two coordinates of (φ^1, φ^2) are equal. This observation is closely related to the theorem of van den Berg [37], Theorem 1, showing the equality of two Gibbs measures under the assumption that there is no 'disagreement percolation' between independent samples from the measures.

Sheffield [31], Lemma 8.1.3, made an additional important observation in the context of random surfaces. The setup considered there allows for both integer-valued and real-valued surfaces, and also for surfaces whose potential U is not required to satisfy the restriction $U(x) = U(-x)$ (allowing to introduce a slope to the surface). We explain Sheffield's observation in the real-valued case: Take $S = \mathbb{R}$, λ_v^1, λ_v^2 to be Lebesgue measure for $v \in V \setminus V_0$ and $h_{(v,w)}(a, b) = \exp(-U(a - b))$ for a measurable $U : \mathbb{R} \to (-\infty, \infty]$ and all $(v, w) \in \vec{E}$. Then,

if U is convex then, almost surely, on each ω-connected component C

either $\varphi_v^1 > \varphi_v^2$ for all $v \in C$, or $\varphi_v^1 < \varphi_v^2$ for all $v \in C$, or $C = \{v\}$ and $\varphi_v^1 = \varphi_v^2$.

This follows from (37) by noting that,

$$\begin{aligned} &\frac{h_{(v,w)}(\varphi_v^2, \varphi_w^1) \cdot h_{(v,w)}(\varphi_v^1, \varphi_w^2)}{h_{(v,w)}(\varphi_v^1, \varphi_w^1) \cdot h_{(v,w)}(\varphi_v^2, \varphi_w^2)} \\ &\quad = \exp\big(U(\varphi_v^1 - \varphi_w^1) + U(\varphi_v^2 - \varphi_w^2) - U(\varphi_v^2 - \varphi_w^1) - U(\varphi_v^1 - \varphi_w^2)\big). \end{aligned}$$

Writing

$$\begin{aligned} \varphi_v^2 - \varphi_w^1 &= p(\varphi_v^1 - \varphi_w^1) + (1 - p)(\varphi_v^2 - \varphi_w^2), \\ \varphi_v^1 - \varphi_w^2 &= (1 - p)(\varphi_v^1 - \varphi_w^1) + p(\varphi_v^2 - \varphi_w^2), \end{aligned}$$

where

$$p := \frac{\varphi_w^2 - \varphi_w^1}{\varphi_v^1 - \varphi_v^2 + \varphi_w^2 - \varphi_w^1},$$

so that if either $\varphi_v^1 \ge \varphi_v^2$ and $\varphi_w^1 \le \varphi_w^2$, or $\varphi_v^1 \le \varphi_v^2$ and $\varphi_w^1 \ge \varphi_w^2$ (but at least one inequality is strict) then $p \in [0, 1]$ whence convexity of U implies that

$$U(\varphi_v^2 - \varphi_w^1) + U(\varphi_v^1 - \varphi_w^2) \le U(\varphi_v^1 - \varphi_w^1) + U(\varphi_v^2 - \varphi_w^2).$$

Sheffield [31] used the cluster swapping idea in his investigation of the translation-invariant gradient Gibbs measures of random surfaces, in the real- and integer-valued cases, with convex potentials (following Funaki–Spohn [20] for the real-valued case). Along the way, he uses cluster swapping in a beautifully simple manner to obtain monotnicity in boundary conditions and log-concavity of the single-site marginal distributions in random surfaces with convex potentials [31], Section 8.2.

Lastly, we remark that cluster swapping may sometimes be given a fruitful interpretation as the swapping of disagreement sets between configurations of an extended model. Such an interpretation was provided in [2] for the random-field Ising model where, rather than defining ω, configurations of the model were extended to continuous functions on the metric graph—the metric space obtained by replacing edges with continuous segments—and swaps were made on connected components of the disagreement set of two extended configurations.

Markov chains: We point out that Markov chains in discrete time also fit the framework (16). For simplicity, we discuss the case where the state space of the chain is countable. Let S be a finite or countable set (with the discrete sigma algebra) and let $P = (P(a,b))_{a,b\in S}$ be the transition probabilities for a Markov chain on S. Thus we assume that

$$P(a,b) \geq 0 \quad \text{for all } a,b \in S \quad \text{and} \quad \sum_{b\in S} P(a,b) = 1 \quad \text{for all } a \in S.$$

A finite sequence $X_0, X_1, \dots, X_n$ of S-valued random variables is a sample of the Markov chain if

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \mu(x_0)P(x_0,x_1)\cdots P(x_{n-1},x_n), \tag{38}$$

for some initial probability distribution μ on S. The distribution of the sequence thus fits the framework (16) with the graph $G = (V,E)$ having $V = \{0,1,\dots,n\}$ and $E = \{\{0,1\},\dots,\{n-1,n\}\}$, with the single-site measures $\lambda_0 := \mu$ and λ_j being the counting measure on S for $1 \leq j \leq n$ and the interaction functions $h_{(j-1,j)} = P$ for $1 \leq j \leq n$.

Thus, if one has in hand a reflection $\tau : S \to S$, one may apply the general reflection transformation procedure to the Markov chain. In this setting, letting $V_0 := \{0\}$, the conditions (18), (19) and (20) defining a reflection translate to $\tau(\tau(a)) = a$ for $a \in S$ and $P(\tau(a),\tau(b)) = P(a,b)$ for $a,b \in S$. As already mentioned, the reflection principle for simple random walk (see (8)) can be seen as a special case of this setup.

In most of the discussion above, the functions $h_{(v,w)}$ were chosen the same for all $(v,w) \in \vec{E}$. We point out that nonhomogeneous setups, with $h_{(v,w)}$ depending on (v,w), may arise even in the investigation of homogeneous models, as in the discussion at the end of Section 4. The use of nonhomogeneous (λ_v) may also be natural in some, otherwise homogeneous, contexts. For instance, the height function of the dimer model on the triangular lattice has the modulo 3 of the heights fixed on each of the three sub-lattices, and this restriction may be imposed with a suitable choice of (λ_v) (or, alternatively, with a suitable choice of $(h_{(v,w)})$).

The usefulness of the above discussion is demonstrated in the next sections where we prove our main theorems, with the proofs in Section 3 and Section 4 being particularly short.

3. Sublevel set connectivity for random surfaces.

In this section we prove Theorem 1.2. As remarked there, the lower bound in (10) is trivial so we focus here on proving the upper bound and (11).

As in the theorem, let $G = (V,E)$ be a finite connected graph, let $V_0 \subseteq V$ be nonempty, let U be a potential satisfying the monotonicity condition (3) and the assumption (2) that μ_{U,G,V_0} is well defined. Let φ be randomly sampled from μ_{U,G,V_0}. We first prove a more general inequality than the upper bound in (10),

$$\mathbb{P}(V_0 \overset{\varphi<m}{\not\longleftrightarrow} v, \varphi_v \in D) \leq \mathbb{P}(\varphi_v \in 2m - D) \quad \text{for all } v \in V, m \geq 0 \text{ and Borel } D \subseteq \mathbb{R}, \tag{39}$$

where $2m - D := \{2m - t : t \in D\}$. The upper bound in (10) follows since

$$\mathbb{P}(V_0 \overset{\varphi<m}{\not\longleftrightarrow} v) = \mathbb{P}(V_0 \overset{\varphi<m}{\not\longleftrightarrow} v, \varphi_v < m) + \mathbb{P}(\varphi_v \geq m) \leq 2\mathbb{P}(\varphi_v \geq m) = \mathbb{P}(|\varphi_v| \geq m).$$

We proceed to prove (39). Let τ_m be the reflection introduced in (30). Let $\omega : E \to \{0, 1\}$ be sampled according to (31) and note that, as a consequence of (32), we have

$$\{V_0 \overset{\varphi<m}{\not\longleftrightarrow} v, \varphi_v \in D\} \subseteq \{v \overset{\omega}{\not\leftrightarrow} V_0\} \qquad (\mathrm{mod}\ \mathbb{P}),$$

where we write $A_1 \subseteq A_2 \pmod{\mathbb{P}}$, for events A_1, A_2, to indicate that $\mathbb{P}(A_1 \setminus A_2) = 0$. Thus

$$\{V_0 \overset{\varphi<m}{\not\longleftrightarrow} v, \varphi_v \in D\} \subseteq \{\varphi_v^{\omega,v} \in 2m - D\} \qquad (\mathrm{mod}\ \mathbb{P}),$$

where we recall the definition of $\varphi^{\omega,v}$ from (24). Thus,

$$\mathbb{P}(V_0 \overset{\varphi<m}{\not\longleftrightarrow} v, \varphi_v \in D) \le \mathbb{P}(\varphi_v^{\omega,v} \in 2m - D) = \mathbb{P}(\varphi_v \in 2m - D),$$

where the last equality follows from Lemma 2.1. This establishes (39).

Now suppose, additionally, that U satisfies the finite-support condition (4). We again prove a more general inequality than needed,

$$\mathbb{P}(V_0 \overset{\varphi<m}{\not\longleftrightarrow} v, \varphi_v \in D) \ge \mathbb{P}(\varphi_v \in 2m + 1 - D) \\ \text{for all } v \in V, m \ge 0 \text{ and Borel } D \subseteq (-\infty, m]. \tag{40}$$

The inequality (11) is a consequence of (40) and the symmetry of φ_v as

$$\begin{aligned}\mathbb{P}(V_0 \overset{\varphi<m}{\not\longleftrightarrow} v) &= \mathbb{P}(V_0 \overset{\varphi<m}{\not\longleftrightarrow} v, \varphi_v < m) + \mathbb{P}(\varphi_v \ge m) \ge \mathbb{P}(\varphi_v > m + 1) + \mathbb{P}(\varphi_v \ge m) \\ &= \mathbb{P}(|\varphi_v| \ge m) - \mathbb{P}(\varphi_v \in (m, m + 1)).\end{aligned}$$

To see (40), we now consider the reflection $\tau_{m+\frac{1}{2}}$. Again, we let $\omega : E \to \{0, 1\}$ be sampled according to (31) (with respect to $\tau_{m+\frac{1}{2}}$) and note that, by (32),

$$\{\varphi_v \ge m + 1\} \subseteq \{v \overset{\omega}{\not\leftrightarrow} V_0\} \qquad (\mathrm{mod}\ \mathbb{P}).$$

Additionally, the finite-support condition (4) implies that on the event $\{\varphi_v \ge m + 1\}$ any path connecting V_0 and v must pass through a vertex w on which $\varphi_w \in [m, m + 1]$. Since $\tau_{m+\frac{1}{2}}$ preserves the interval $[m, m + 1]$ we conclude that

$$\{\varphi_v \ge m + 1\} \subseteq \{V_0 \overset{\varphi^{\omega,v}<m}{\not\longleftrightarrow} v\} \qquad (\mathrm{mod}\ \mathbb{P}).$$

Altogether, we conclude that, for each Borel $D \subseteq (-\infty, m]$,

$$\mathbb{P}(\varphi_v \in 2m + 1 - D) \le \mathbb{P}(V_0 \overset{\varphi^{\omega,v}<m}{\not\longleftrightarrow} v, \varphi_v^{\omega,v} \in D) = \mathbb{P}(V_0 \overset{\varphi<m}{\not\longleftrightarrow} v, \varphi_v \in D),$$

where we used Lemma 2.1 in the last step. This concludes the proof of (40).

4. Monotonicity of pair correlations in spin $O(n)$ models. In this section we prove Theorem 1.3. Let $v \in V \setminus V_0$. Let $b_1, b_2 \in \mathbb{S}^{n-1}$ be such that $\langle b_1, e_1\rangle \ge \langle b_2, e_1\rangle$ and $b_1 \ne b_2$. Define

$$a := \frac{b_2 - b_1}{\|b_2 - b_1\|}$$

and consider the reflection along the hyperplane orthogonal to a, as defined in (33),

$$\tau_a(b) = b - 2\langle a, b\rangle \cdot a.$$

As explained, τ_a is a reflection for the spin $O(n)$ model with potential U. In addition $\tau_a(b_2) = b_1$. Moreover, we observe that

$$\langle e_1, a\rangle = \|b_2 - b_1\|^{-1}(\langle e_1, b_2\rangle - \langle e_1, b_1\rangle) \le 0, \tag{41}$$

and that

$$\langle b_2, a\rangle = \|b_2 - b_1\|^{-1}\big(\langle b_2, b_2\rangle - \langle b_2, b_1\rangle\big) > 0.$$

Choose $r > 0$ sufficiently small that

$$\langle b, a\rangle > 0 \quad \text{for all } b \in B(b_2, r), \tag{42}$$

where we write $B(b_2, r) := \{b \in \mathbb{R}^n \mid \|b - b_2\|_2 \le r\}$ for the closed Euclidean ball of radius r around b.

As before, we let $\omega : E \to \{0, 1\}$ be sampled according to (34) (with respect to τ_a). The relations (41) and (42) combined with the property (35) imply that

$$\{\varphi_v \in B(b_2, r)\} \subseteq \{v \overset{\omega}{\not\leftrightarrow} V_0\} \qquad (\text{mod } \mathbb{P}).$$

Consequently, recalling the definition of $\varphi^{\omega,v}$ from (24) and the fact that $\tau_a(b_2) = b_1$ and τ_a is an isometry,

$$\{\varphi_v \in B(b_2, r)\} \subseteq \{\varphi_v^{\omega,v} \in B(b_1, r)\} \qquad (\text{mod } \mathbb{P}).$$

Since $\varphi^{\omega,v}$ has the same distribution as φ, we have

$$\mathbb{P}\big(\varphi_v \in B(b_2, r)\big) \le \mathbb{P}\big(\varphi_v^{\omega,v} \in B(b_1, r)\big) = \mathbb{P}\big(\varphi_v \in B(b_1, r)\big).$$

Fix a version d_v of the density of φ_v with respect to $\mu_{\mathbb{S}^{n-1}}$, which we assume is rotationally symmetric in the sense that $d_v(b)$ is a function of $\langle b, e_1\rangle$. As

$$\frac{\mathbb{P}(\varphi_v \in B(b, r))}{\mu_{\mathbb{S}^{n-1}}(B(b, r))} \to d_v(b) \quad \text{as } r \downarrow 0 \tag{43}$$

for $\mu_{\mathbb{S}^{n-1}}$-almost every b, we conclude that $d_v(b)$ is a monotone increasing function of $\langle b, e_1\rangle$ except on the $\mu_{\mathbb{S}^{n-1}}$-null set of b's for which the convergence in (43) fails. We may then redefine $d_v(b)$ on this null set to make $d_v(b)$ monotone for all b, as we wanted to show.

We may prove Theorem 1.4, regarding the monotonicity of marginal densities in random surfaces, in exactly the same manner, replacing the reflection τ_a on $\mathbb{S}^{n-1}$ by the reflection τ_m on $\mathbb{R}$. For variety (and possible interest in other contexts), we briefly explain an alternative route to the proof of Theorem 1.4 via convexity considerations (we are not aware of such an alternative for proving Theorem 1.3).

When the potential U is a convex function, the distribution of μ_{U,G,V_0} is log-concave and centrally symmetric (invariant to a global sign flip), whence the marginal distribution of φ_x is also log-concave and centrally symmetric so that $|\varphi_x|$ has a nonincreasing density. While our assumption that U is monotone does not imply that U is convex, it allows us to decompose the distribution of μ_{U,G,V_0} as a mixture of centrally-symmetric log-concave distributions and deduce Theorem 1.4 as before. The decomposition is a special case of the Edwards–Sokal decomposition [18] and we briefly describe it next.

Let $t = (t_e)_{e\in E} \in (0,\infty)^E$ and define the measure μ_{t,G,V_0} by

$$d\mu_{t,G,V_0}(\varphi) := \frac{1}{Z_{t,G,V_0}} \prod_{\{v,w\}\in E} \mathbf{1}_{[-t_{\{v,w\}}, t_{\{v,w\}}]}(\varphi_v - \varphi_w) \prod_{v\in V_0} \delta_0(d\varphi_v) \prod_{v\in V\setminus V_0} d\varphi_v, \tag{44}$$

where Z_{t,G,V_0} is a normalizing constant. In other words, the measure μ_{t,G,V_0} is uniform on the set of all t-Lipschitz functions, functions changing by at most $t_{\{v,w\}}$ on the edge $\{v, w\}$, normalized to be 0 at V_0. In particular, μ_{t,G,V_0} is log-concave and centrally-symmetric. We say that μ_{t,G,V_0} is the measure of a *random surface with inhomogeneous hammock potentials*.

Recalling our assumption that U is monotone in the sense (3), let us suppose, for convenience, that e^{-U} is right continuous on $[0, \infty)$ (noting that the measure μ_{U,G,V_0} is invariant to changing the value of U at countably many points). Define the measure λ_U with the Lebesgue–Stieltjes differential, $d\lambda_U(t) := -d\exp(-U(t))$ on $[0, \infty)$, that is,

$$\int_{(a,b]} d\lambda_U(s) := \exp(-U(a)) - \exp(-U(b)), \quad 0 < a < b \text{ and } \lambda_U(\{0\}) := 0.$$

Since U is symmetric, we can write

$$\exp(-U(x)) = \int_{(|x|,\infty)} d\lambda_U(s) = \int_{[0,\infty)} \mathbf{1}_{[-s,s]}(x)\, d\lambda_U(s).$$

Substituting this expression for $\exp(-U(x))$ in the density of μ_{U,G,V_0} given in (1) shows, after a short calculation, that μ_{U,G,V_0} is a mixture, with respect to t, of measures of the form μ_{t,G,V_0}.

As remarked in Section 1.2, analogs of Theorem 1.3 and Theorem 1.4 continue to hold for clock models and integer-valued random surfaces, with the same proofs. To prove the analog of Theorem 1.4 via the convexity approach, one needs to know that if the potential U, now defined on integers, is convex in a suitable sense then the marginal distribution of φ_x is log-concave (in a suitable sense). Such a result was proved by Sheffield [31], Lemma 8.2.4, (using the cluster-swapping method described in Section 2).

5. Estimating the probability of having extremal gradients. In this section we prove Theorem 1.1.

Let $G = (V, E)$ be a finite connected graph and let $V_0 \subseteq V$ be nonempty. Let U be a potential satisfying assumptions (3) and (4), and recall from Section 1 the random surface measure μ_{U,G,V_0}. Observe that if any of the given edges $\{v_1, w_1\}, \dots, \{v_k, w_k\}$ has both endpoints in V_0 then the conclusion (7) of Theorem 1.1 follows trivially as the gradient of φ on that edge is zero almost surely. We assume henceforth that each of the given edges has at most one endpoint in V_0. Without loss of generality, we now assume that V_0 is a singleton, that is,

$$V_0 = \{v_0\},$$

as we can replace our graph with the graph in which all vertices in V_0 have been identified to a single vertex v_0, erasing self-loops and keeping a single representative of each multiple edge, and note that the random surface measure is naturally preserved under this operation. This identification operation may substantially increase the degree of v_0, possibly beyond the maximal degree in the original graph. Thus, we will take care not to rely on the degree of v_0 in our proofs. The degrees of all other vertices cannot increase under the identification operation. To address these issues it is convenient to define, for a given set of edges $F \subseteq E$,

$$\Delta(F) := \max\{\deg(w) : \exists\{v, w\} \in F, w \neq v_0\},$$

where we write $\deg(w)$ for the degree of w in the graph G, so that $\Delta(F)$ is the maximal degree of a vertex other than v_0 in one of the edges of F. For brevity we write

$$\mu_{U,G,v_0} := \mu_{U,G,\{v_0\}}.$$

We start with several definitions which will be used throughout the section. For a given set of edges $H \subseteq E$, let

$$\mathrm{Orient}(H) := \{(v, w) : \{v, w\} \in H\}$$

stand for all orientations of the edges of H (each undirected edge appears with both orientations in $\mathrm{Orient}(H)$). For brevity, we denote the set of all oriented edges by

$$\vec{E} := \mathrm{Orient}(E).$$

We will frequently reference the event that the random surface has extremal gradients on a given set of edges. This event will be used both for oriented and for unoriented sets of edges and thus we define, for each $0 < \varepsilon < 1$,

$$\begin{aligned} \operatorname{Ext}(H, \varepsilon) &:= \{|\varphi_v - \varphi_w| \geq 1 - \varepsilon \text{ for all } \{v, w\} \in H\}, \quad H \subseteq E, \\ \operatorname{Ext}(\overrightarrow{H}, \varepsilon) &:= \{|\varphi_v - \varphi_w| \geq 1 - \varepsilon \text{ for all } (v, w) \in \overrightarrow{H}\}, \quad \overrightarrow{H} \subseteq \overrightarrow{E}. \end{aligned} \tag{45}$$

Theorem 1.1 is proved as a consequence of three lemmas which we now proceed to describe.

Diluting the given edge set. We partition the real line into 9 sets, each of which is an arithmetic progression of intervals, as follows

$$D_j := \frac{j}{4} + \left[-\frac{1}{8}, \frac{1}{8}\right) + 2\frac{1}{4}\mathbb{Z} = \left\{\frac{j}{4} + x + 2\frac{1}{4}k : -\frac{1}{8} \leq x < \frac{1}{8}, k \in \mathbb{Z}\right\}, \quad 1 \leq j \leq 9, \tag{46}$$

(the shorthand $k\frac{\ell}{m}$ means $k + \frac{\ell}{m}$). The main property of these domains which we shall make use of is that, for each j, the set D_j is invariant to reflection with respect to numbers in $\frac{j}{4} + 1\frac{1}{8} + 2\frac{1}{4}\mathbb{Z}$. In other words, for any $1 \leq j \leq 9$,

$$\text{if } y \in D_j \text{ then also } 2m - y \in D_j \text{ for any } m \in \frac{j}{4} + 1\frac{1}{8} + 2\frac{1}{4}\mathbb{Z}. \tag{47}$$

The sets D_j are in fact invariant to reflections with respect to numbers in the larger set $\frac{j}{4} + 1\frac{1}{8}\mathbb{Z}$ but this will not be used in our proof.

Given a set of oriented edges, we define the event that the value of the surface on the first vertex of each oriented edge belongs to D_j,

$$\Omega_j(\overrightarrow{H}) := \{\varphi_v \in D_j \text{ for all } (v, w) \in \overrightarrow{H}\}, \quad 1 \leq j \leq 9$$

and also

$$\Omega(\overrightarrow{H}) := \bigcup_{j=1}^{9} \Omega_j(\overrightarrow{H}).$$

We will make use of certain separation properties between oriented edges as given in the following definition.

DEFINITION 5.1 (Separated set). A subset of oriented edges $\overrightarrow{H} \subseteq \overrightarrow{E}$ is said to be *separated* if:

(i) for every distinct $(v_1, w_1), (v_2, w_2) \in \overrightarrow{H}$, $w_1 \neq w_2$, $v_1 \neq w_2$ and $v_2 \neq w_1$,
(ii) for all $(v, w) \in \overrightarrow{H}$, $w \neq v_0$.

In words, a separated set of oriented edges is a set in which every two edges are either disjoint or coincide in their first vertex and in which no edge is oriented towards v_0.

Our first lemma shows that the probability of $\operatorname{Ext}(F, \varepsilon)$, for a given $F \subseteq E$, may be bounded in terms of the probability of $\operatorname{Ext}(\overrightarrow{H}, \varepsilon) \cap \Omega(\overrightarrow{H})$, for some large separated $\overrightarrow{H} \subseteq \operatorname{Orient}(F)$.

LEMMA 5.2. *Let $0 < \varepsilon < 1$, let $F \subseteq E$ be a nonempty set of edges and let φ be randomly sampled from μ_{U,G,v_0}. Then there exists a separated set $\overrightarrow{H} \subseteq \operatorname{Orient}(F)$ satisfying $|\overrightarrow{H}| \geq \frac{|F|}{9\Delta(F)}$ and*

$$\mathbb{P}(\operatorname{Ext}(F, \varepsilon)) \leq 2^{|F|} \cdot \mathbb{P}(\operatorname{Ext}(\overrightarrow{H}, \varepsilon) \cap \Omega(\overrightarrow{H})). \tag{48}$$

Unlocking edges. We now describe the key step in our proof. Suppose that the random surface φ has an extremal gradient on the oriented edge $\vec{e} = (v, w)$, oriented towards w, say, in the sense that $\varphi_w - \varphi_v \geq 1 - \varepsilon$. If φ_u is not much higher than φ_w for all neighbors u of w, then we may change the value of the surface on w, thereby reducing the gradient on (v, w) without reducing significantly the density under the random surface measure (see Lemma 5.5 below). Therefore, the difficulty in showing that extremal gradients are rare lies in the possibility that the edge with the extremal gradient is being "locked" into this extreme position by a neighbor of one of its endpoints. Our next definition quantifies the notion that an edge is not locked in such a manner.

DEFINITION 5.3. An oriented edge $\vec{e} = (v, w) \in \vec{E}$ is called *unlocked* in $\varphi \in \mathbb{R}^V$ if

$$\max_{u:\{u,w\}\in E} |\varphi_u - \varphi_v| \leq 1\frac{1}{4}. \tag{49}$$

We define the corresponding event as

$$\mathrm{U}_{\vec{e}} := \{\vec{e} \text{ is unlocked in } \varphi\}.$$

The following key lemma will allow us to reduce our study of extremal edges to the case that these edges are unlocked.

LEMMA 5.4. *Let $0 < \varepsilon < 1$, let $\vec{H} \subseteq \vec{E}$ be a nonempty, separated set of oriented edges, let $\vec{e} = (v, w) \in \vec{H}$ and let φ be randomly sampled from μ_{U,G,v_0}. Then*

$$\mathbb{P}\big(\mathrm{Ext}(\vec{H}, \varepsilon) \cap \Omega(\vec{H})\big) \leq 2^{\deg(w)-1} \cdot \mathbb{P}\big(\mathrm{Ext}(\vec{H}, \varepsilon) \cap \Omega(\vec{H}) \cap \mathrm{U}_{\vec{e}}\big).$$

The proof of the lemma uses the reflection transformation described in Section 2.

The probability that an unlocked edge is extremal. As noted above, if an edge $(v, w) \in \vec{E}$ has an extremal gradient and is unlocked in the surface φ, then we may change φ_w to reduce the gradient on the edge while controlling the change in the density of the surface under the measure μ_{U,G,v_0}. This idea is quantified by the next lemma.

LEMMA 5.5. *Let $0 < \varepsilon \leq \frac{1}{8}$, let $\vec{e} = (v, w) \in \vec{E}$ with $w \neq v_0$ and let φ be randomly sampled from μ_{U,G,v_0}. Then*

$$\mathbb{P}\big(|\varphi_w - \varphi_v| \geq 1 - \varepsilon \mid (\varphi_u)_{u\in V\setminus\{w\}}\big) \cdot \mathbf{1}_{\mathrm{U}_{\vec{e}}} \leq \delta(U, \vec{e}, \varepsilon) \quad \textit{almost surely},$$

where $\mathbf{1}_A$ denotes the indicator random variable of the event A and where we write

$$\delta(U, \vec{e}, \varepsilon) := 8\varepsilon \cdot \exp\left(-U(1-\varepsilon) + U(0) + \deg(w)\left(U\left(\frac{3}{4}\right) - U(0)\right)\right).$$

The above three lemmas are proved in the next section. We now explain how Theorem 1.1 follows as a consequence of them.

PROOF OF THEOREM 1.1. Let $0 < \varepsilon \leq \frac{1}{8}$, let $F \subseteq E$ be a nonempty set of edges and let φ be randomly sampled from μ_{U,G,v_0}. Our goal is to estimate the probability of $\mathrm{Ext}(F, \varepsilon)$.

Using Lemma 5.2, let $\vec{H} \subseteq \mathrm{Orient}(F)$ be a separated set satisfying $|\vec{H}| \geq \frac{|F|}{9\Delta(F)}$ and

$$\mathbb{P}\big(\mathrm{Ext}(F, \varepsilon)\big) \leq 2^{|F|}\mathbb{P}\big(\mathrm{Ext}(\vec{H}, \varepsilon) \cap \Omega(\vec{H})\big). \tag{50}$$

Let $\vec{e} = (v, w) \in \vec{H}$. By Lemma 5.4, we have

$$\begin{aligned}&\mathbb{P}(\mathrm{Ext}(\vec{H}, \varepsilon) \cap \Omega(\vec{H}))\\ &\le 2^{\deg(w)-1}\mathbb{P}(\mathrm{Ext}(\vec{H}, \varepsilon) \cap \Omega(\vec{H}) \cap \mathrm{U}_{\vec{e}})\\ &\le 2^{\deg(w)-1}\mathbb{P}(\{|\varphi_v - \varphi_w| \ge 1-\varepsilon\} \cap \mathrm{Ext}(\vec{H} \setminus \{\vec{e}\}, \varepsilon) \cap \Omega(\vec{H} \setminus \{\vec{e}\}) \cap \mathrm{U}_{\vec{e}}).\end{aligned} \tag{51}$$

We note that, as $\vec{H}$ is a separated set, the events $\mathrm{Ext}(\vec{H} \setminus \{\vec{e}\}, \varepsilon)$, $\Omega(\vec{H} \setminus \{\vec{e}\})$ and $\mathrm{U}_{\vec{e}}$ are measurable with respect to the random variables $(\varphi_u)_{u \in V\setminus\{w\}}$. Thus, using Lemma 5.5, we may estimate

$$\begin{aligned}&\mathbb{P}(\{|\varphi_v - \varphi_w| \ge 1-\varepsilon\} \cap \mathrm{Ext}(\vec{H} \setminus \{\vec{e}\}, \varepsilon) \cap \Omega(\vec{H} \setminus \{\vec{e}\}) \cap \mathrm{U}_{\vec{e}})\\ &= \mathbb{E}[\mathbb{P}(|\varphi_w - \varphi_v| \ge 1-\varepsilon \mid (\varphi_u)_{u\in V\setminus\{w\}})\mathbf{1}_{\mathrm{Ext}(\vec{H}\setminus\{\vec{e}\},\varepsilon)\cap\Omega(\vec{H}\setminus\{\vec{e}\})\cap \mathrm{U}_{\vec{e}}}]\\ &\le \delta(U, \vec{e}, \varepsilon)\mathbb{P}(\mathrm{Ext}(\vec{H} \setminus \{\vec{e}\}, \varepsilon) \cap \Omega(\vec{H} \setminus \{\vec{e}\})).\end{aligned} \tag{52}$$

Recall from the statement of Theorem 1.1 that

$$\delta(U, \varepsilon) = \varepsilon \cdot \exp\biggl(-U(1-\varepsilon) + U(0) + \Delta(F)\biggl(U\biggl(\frac{3}{4}\biggr) - U(0)\biggr)\biggr)$$

and observe that, as $w \ne v_0$ since $\vec{H}$ is separated and as $U(3/4) \ge U(0)$ by our assumption that U is nondecreasing on $[0, \infty)$,

$$\deg(w) \le \Delta(F) \quad \text{and} \quad \delta(U, \vec{e}, \varepsilon) \le 8\delta(U, \varepsilon). \tag{53}$$

Putting together (51), (52) and (53), we conclude that

$$\mathbb{P}(\mathrm{Ext}(\vec{H}, \varepsilon) \cap \Omega(\vec{H})) \le 2^{\Delta(F)-1} \cdot 8\delta(U, \varepsilon)\mathbb{P}(\mathrm{Ext}(\vec{H} \setminus \{\vec{e}\}, \varepsilon) \cap \Omega(\vec{H} \setminus \{\vec{e}\})).$$

Iterating this estimate over all edges in $\vec{H}$ shows that

$$\mathbb{P}(\mathrm{Ext}(\vec{H}, \varepsilon) \cap \Omega(\vec{H})) \le (2^{\Delta(F)+2}\delta(U, \varepsilon))^{|\vec{H}|}.$$

Substituting this estimate into (50) and using the fact that $|\vec{H}| \ge \frac{|F|}{9\Delta(F)}$, we have

$$\mathbb{P}(\mathrm{Ext}(F, \varepsilon)) \le \min(2^{|F|}(2^{\Delta(F)+2}\delta(U, \varepsilon))^{|\vec{H}|}, 1) \le \min((2^{10\Delta(F)+2}\delta(U, \varepsilon))^{\frac{|F|}{9\Delta(F)}}, 1).$$

This concludes the proof of Theorem 1.1, given the above lemmas, with the constant $C(\Delta) = 2^{10\Delta+2}$. $\square$

5.1. *Proof of Lemma* 5.2.

PROOF. Assign an orientation to the edges of F, chosen arbitrarily except for the rule that edges having v_0 as an endpoint are oriented to have v_0 as their first vertex. Denote the resulting set of oriented edges by $\vec{F}$.

Let φ be randomly sampled from μ_{U,G,v_0}. Observe that there is a $1 \le j \le 9$ and a (random) subset $\vec{F}' \subseteq \vec{F}$ satisfying $|\vec{F}'| \ge \frac{1}{9}|\vec{F}|$ such that $\varphi_v \in D_j$ for all $(v, w) \in \vec{F}'$ (as each edge satisfies $\varphi_v \in D_j$ for a unique j). In addition, we may choose a (random) separated subset $\vec{H} \subseteq \vec{F}'$ satisfying $|\vec{H}| \ge \frac{1}{\Delta(F)}|\vec{F}'|$. Indeed, this may be done in a greedy manner: Sequentially, for each edge (v, w) still in $\vec{F}'$, we discard from $\vec{F}'$ all edges (x, y) with either $x = w$ or $y = w$, discarding in this way at most $\deg(w) - 1 \le \Delta(F) - 1$ edges. In conclusion, defining

$$\mathcal{H} := \biggl\{\vec{H} \subseteq \vec{F} : |\vec{H}| \ge \frac{|F|}{9\Delta(F)}, \vec{H} \text{ is separated}\biggr\} \tag{54}$$

we have shown that, almost surely, $\varphi \in \Omega(\overrightarrow{H})$ for some $\overrightarrow{H} \in \mathcal{H}$. Thus, using that $|\mathcal{H}| \le 2^{|\overrightarrow{F}|} = 2^{|F|}$,

$$\mathbb{P}(\mathrm{Ext}(F,\varepsilon)) \le \sum_{\overrightarrow{H}\in\mathcal{H}} \mathbb{P}(\mathrm{Ext}(\overrightarrow{H},\varepsilon)\cap\Omega(\overrightarrow{H})) \le 2^{|F|} \max_{\overrightarrow{H}\in\mathcal{H}} \mathbb{P}(\mathrm{Ext}(\overrightarrow{H},\varepsilon)\cap\Omega(\overrightarrow{H})). \qquad \square$$

5.2. *Proof of Lemma* 5.4. Fix $0 < \varepsilon < 1$, a nonempty, separated set of oriented edges $\overrightarrow{H} \subseteq \overrightarrow{E}$ and an oriented edge $\vec{e} = (v, w) \in \overrightarrow{H}$. Let φ be randomly sampled from μ_{U,G,v_0}. As the events $\Omega_j(\overrightarrow{H})$ which comprise $\Omega(\overrightarrow{H})$ are disjoint, it suffices to prove that

$$\mathbb{P}(\mathrm{Ext}(\overrightarrow{H},\varepsilon)\cap\Omega_j(\overrightarrow{H})) \le 2^{\deg(w)-1}\cdot\mathbb{P}(\mathrm{Ext}(\overrightarrow{H},\varepsilon)\cap\Omega_j(\overrightarrow{H})\cap \mathrm{U}_{\vec{e}}), \quad 1\le j\le 9. \tag{55}$$

Thus we also fix $1 \le j \le 9$. Define the subset M_j of the real line by

$$M_j := \frac{j}{4} + 1\frac{1}{8} + 2\frac{1}{4}\mathbb{Z}.$$

Now define

$$m(\varphi) := \begin{cases} \min(m \in M_j : m \ge \varphi_v) & \varphi_w \ge \varphi_v, \\ \max(m \in M_j : m \le \varphi_v) & \varphi_w < \varphi_v, \end{cases}$$

so that $m(\varphi)$ is on the same side of φ_v as φ_w and $m(\varphi) \ne \varphi_v$ on the event $\Omega_j(\overrightarrow{H})$. Define also

$$W(\varphi) := \{u \in V : \{u, w\} \in E, \mathrm{sign}(\varphi_u - m(\varphi)) = \mathrm{sign}(m(\varphi) - \varphi_v)\}, \tag{56}$$

where, as usual, $\mathrm{sign}(x) = 1$ if $x > 0$, $\mathrm{sign}(x) = -1$ if $x < 0$ and $\mathrm{sign}(x) = 0$ if $x = 0$. We shall prove that for each $W \subseteq V$ and $m \in M_j$,

$$\begin{aligned} &\mathbb{P}(\mathrm{Ext}(\overrightarrow{H},\varepsilon)\cap\Omega_j(\overrightarrow{H})\cap\{W(\varphi)=W\}\cap\{m(\varphi)=m\}) \\ &\quad\le \mathbb{P}(\mathrm{Ext}(\overrightarrow{H},\varepsilon)\cap\Omega_j(\overrightarrow{H})\cap \mathrm{U}_{\vec{e}}\cap\{m(\varphi)=m\}). \end{aligned} \tag{57}$$

This relation implies (55), and hence the lemma, by summing over all possible values of W and m, and using the fact that $\mathbb{P}(W(\varphi) = W) = 0$ unless W is a subset of $\{u \in V : \{u, w\} \in E, u \ne v\}$.

We proceed to prove (57) and it is here that we make use of the reflection transformation described in Section 2. Fix $W \subseteq V$ and $m \in M_j$ satisfying

$$\mathbb{P}(W(\varphi) = W, m(\varphi) = m) > 0$$

(as the relation (57) is trivial if this probability is zero). Recall the "reflection around m" mapping $\tau_m : \mathbb{R} \to \mathbb{R}$ given by $\tau_m(a) = 2m - a$, as in (30). Let (φ, ω) be randomly sampled from the τ_m-Edwards–Sokal coupling defined in (31). We define the reflected configuration $\varphi^{\omega,W} : V \to \mathbb{R}$ as follows: If $W = \varnothing$ then we set $\varphi^{\omega,W} := \varphi$. Otherwise,

$$\begin{aligned} &\text{if } W \overset{\omega}{\not\leftrightarrow} v_0 \text{ then } \varphi^{\omega,W}_v := \begin{cases} \tau_m(\varphi_v) & W \overset{\omega}{\leftrightarrow} v, \\ \varphi_v & W \overset{\omega}{\not\leftrightarrow} v. \end{cases} \\ &\text{If } W \overset{\omega}{\leftrightarrow} v_0 \text{ then } \varphi^{\omega,W}_v := \begin{cases} \varphi_v & W \overset{\omega}{\leftrightarrow} v, \\ \tau_m(\varphi_v) & W \overset{\omega}{\not\leftrightarrow} v. \end{cases} \end{aligned} \tag{58}$$

It then follows from the discussion after Lemma 2.1 that $(\varphi^{\omega,W}, \omega)$ has the same distribution as (φ, ω). The equality in distribution shows that (57) is a consequence of the following relation:

$$\begin{aligned} &\{\varphi \in \mathrm{Ext}(\overrightarrow{H},\varepsilon)\cap\Omega_j(\overrightarrow{H})\}\cap\{W(\varphi)=W\}\cap\{m(\varphi)=m\} \\ &\quad\subseteq \{\varphi^{\omega,W} \in \mathrm{Ext}(\overrightarrow{H},\varepsilon)\cap\Omega_j(\overrightarrow{H})\cap \mathrm{U}_{\vec{e}}\}\cap\{m(\varphi^{\omega,W})=m\} \qquad (\mathrm{mod}\ \mathbb{P}) \end{aligned} \tag{59}$$

(as before, we write (mod $\mathbb{P}$) to indicate that the containment is in the sense of the difference having zero probability), where, with a slight abuse of notation, we consider the events $\text{Ext}(\overrightarrow{H}, \varepsilon)$, $\Omega_j(\overrightarrow{H})$ and $\text{U}_{\vec{e}}$ as subsets of the space $\mathbb{R}^V$ of configurations. Thus, it remains to prove (59), which is a consequence of the following three claims.

CLAIM 5.6. *Almost surely, if $\varphi \in \text{Ext}(\overrightarrow{H}, \varepsilon) \cap \Omega_j(\overrightarrow{H})$ then $\varphi^{\omega,W} \in \text{Ext}(\overrightarrow{H}, \varepsilon)$.*

PROOF. We will prove the stronger consequence that, under the given assumptions,

$$|\varphi_x - \varphi_y| = |\varphi_x^{\omega,W} - \varphi_y^{\omega,W}| \quad \text{for all } (x, y) \in \overrightarrow{H}. \tag{60}$$

Fix $(x, y) \in \overrightarrow{H}$. Observe that, by definition of the reflection operation,

$$|\varphi_x^{\omega,W} - \varphi_y^{\omega,W}| \in \{|\varphi_x - \varphi_y|, |2m - \varphi_x - \varphi_y|\}.$$

Suppose that

$$|\varphi_x^{\omega,W} - \varphi_y^{\omega,W}| = |2m - \varphi_x - \varphi_y| \tag{61}$$

as in the other case (60) is clearly satisfied. Since $\varphi \in \Omega_j(\overrightarrow{H})$ it follows that $\varphi_x \in D_j$. Thus, recalling the definition (46) of D_j and the fact that $m \in M_j$ we see that

$$|m - \varphi_x| \geq \text{dist}(M_j, D_j) = 1. \tag{62}$$

Since $|\varphi_x - \varphi_y| \leq 1$ it implies that either $m \geq \max(\varphi_x, \varphi_y)$ or $m \leq \min(\varphi_x, \varphi_y)$. In both cases,

$$|2m - \varphi_x - \varphi_y| = |m - \varphi_x| + |m - \varphi_y| \geq 1. \tag{63}$$

Since $\varphi^{\omega,W}$ is a Lipschitz function almost surely, we conclude from (61) that equality must hold in (63). Taking into account (62), this implies that $\varphi_y = m$ and $|m - \varphi_x| = 1$, in which case $|2m - \varphi_x - \varphi_y| = |\varphi_x - \varphi_y|$ so that (60) holds. □

CLAIM 5.7. *If $\varphi \in \Omega_j(\overrightarrow{H})$ then $\varphi^{\omega,W} \in \Omega_j(\overrightarrow{H})$.*

PROOF. The claim follows from the fact that $\tau_m(D_j) \subseteq D_j$ since $m \in M_j$, as noted in (47). □

CLAIM 5.8. *Almost surely, if $\varphi \in \text{Ext}(\overrightarrow{H}, \varepsilon) \cap \Omega_j(\overrightarrow{H})$, $W(\varphi) = W$ and $m(\varphi) = m$ then $m(\varphi^{\omega,W}) = m$ and $\varphi^{\omega,W} \in \text{U}_{\vec{e}}$.*

PROOF. Let $k \in \mathbb{Z}$ be such that $\varphi_v \in \frac{j}{4} + [-\frac{1}{8}, \frac{1}{8}) + 2\frac{1}{4}k$, using that $\varphi \in \Omega_j(\overrightarrow{H})$. As $\varphi \in \text{Ext}(\overrightarrow{H}, \varepsilon)$ we have that $\varphi_w \neq \varphi_v$. For concreteness, assume that $\varphi_w > \varphi_v$ with the other case being treated similarly. Thus $m = \frac{j}{4} + 1\frac{1}{8} + 2\frac{1}{4}k$ and note that $\varphi_w \leq m$ as φ is a Lipschitz function.

The definition of $\varphi^{\omega,W}$ implies that $\varphi_v^{\omega,W} \in \{\varphi_v, 2m - \varphi_v\}$ and $\varphi_w^{\omega,W} \in \{\varphi_w, 2m - \varphi_w\}$. The fact that both $\varphi_v \leq m$ and $\varphi_w \leq m$ imply that in all four possibilities for the values of $\varphi_v^{\omega,W}$ and $\varphi_w^{\omega,W}$ we have $m(\varphi^{\omega,W}) = m$.

Fix u with $\{u, w\} \in E$. If $u \in W$, that is $\varphi_u > m$, the definition of $\varphi^{\omega,W}$ and property (32) imply that, almost surely,

$$\text{if } W \overset{\omega}{\not\leftrightarrow} v_0 \text{ then } \varphi_v^{\omega,W} = \varphi_v, \varphi_u^{\omega,W} = 2m - \varphi_u \quad \text{and}$$

$$\text{if } W \overset{\omega}{\leftrightarrow} v_0 \text{ then } \varphi_v^{\omega,W} = 2m - \varphi_v, \varphi_u^{\omega,W} = \varphi_u.$$

In both cases

$$|\varphi_v^{\omega,W} - \varphi_u^{\omega,W}| = |\varphi_v + \varphi_u - 2m| \leq \max(m - \varphi_v, \varphi_u - m) \leq 1\frac{1}{4},$$

where we used that $\varphi_u - m \leq \varphi_u - \varphi_w \leq 1$ as φ is a Lipschitz function.

If $u \notin W$, that is $\varphi_u \le m$, then $\varphi_u - \varphi_v \in (-1, 1\frac{1}{4}]$ as φ is a Lipschitz function. As $\varphi_v, \varphi_u \le m$ we conclude from the definition of $\varphi^{\omega,W}$, the fact that $\varphi_x > m$ for all $x \in W$ and property (32) that, almost surely, $|\varphi_u^{\omega,W} - \varphi_v^{\omega,W}| = |\varphi_u - \varphi_v|$, whence $|\varphi_u^{\omega,W} - \varphi_v^{\omega,W}| \le 1\frac{1}{4}$.

Thus, $\max_{u:\{u,w\}\in E} |\varphi_u^{\omega,W} - \varphi_v^{\omega,W}| \le 1\frac{1}{4}$, from which it follows that $\varphi^{\omega,W} \in \mathrm{U}_{\vec{e}}$. □

5.3. *Proof of Lemma* 5.5. Fix $0 < \varepsilon \le \frac{1}{8}$ and $\vec{e} = (v, w) \in \vec{E}$ with $w \neq v_0$. Let φ be randomly sampled from μ_{U,G,v_0}. The conditional density of φ_w given $(\varphi_u)_{u\in V\setminus\{w\}}$ equals

$$\frac{\exp(-\sum_{u:\{u,w\}\in E} U(\varphi_u - \varphi_w))}{\int_{-\infty}^{\infty} \exp(-\sum_{u:\{u,w\}\in E} U(\varphi_u - x))\,dx}.$$

Thus, the lemma will follow by showing that

$$\int_{(-\infty,-(1-\varepsilon)]\cup[1-\varepsilon,\infty)} e^{-\sum_{u:\{u,w\}\in E} U(\varphi_u-(\varphi_v+t))}\,dt \cdot \mathbf{1}_{U_{\vec{e}}} \le \delta(U,\vec{e},\varepsilon) \int_{-\infty}^{\infty} e^{-\sum_{u:\{u,w\}\in E} U(\varphi_u-(\varphi_v+t))}\,dt.$$

Taking into account the Lipschitz assumption (4), we see it suffices to prove the pair of inequalities,

$$\int_{-1}^{-(1-\varepsilon)} e^{-\sum_{u:\{u,w\}\in E} U(\varphi_u-(\varphi_v+t))}\,dt \cdot \mathbf{1}_{U_{\vec{e}}} \le \delta(U,\vec{e},\varepsilon) \int_{-(3/4-\varepsilon)}^{-1/2} e^{-\sum_{u:\{u,w\}\in E} U(\varphi_u-(\varphi_v+t))}\,dt, \tag{64}$$

$$\int_{1-\varepsilon}^{1} e^{-\sum_{u:\{u,w\}\in E} U(\varphi_u-(\varphi_v+t))}\,dt \cdot \mathbf{1}_{U_{\vec{e}}} \le \delta(U,\vec{e},\varepsilon) \int_{1/2}^{3/4-\varepsilon} e^{-\sum_{u:\{u,w\}\in E} U(\varphi_u-(\varphi_v+t))}\,dt. \tag{65}$$

We prove only inequality (65) as inequality (64) follows from it by applying a global sign change to φ. We assume that $U_{\vec{e}}$ holds as (65) is trivially verified otherwise. In addition, we assume that

$$\min\{\varphi_u : \{u,w\} \in E\} \ge \varphi_v - \varepsilon \tag{66}$$

as otherwise, using the Lipschitz assumption (4), the left-hand side of (65) is zero, again verifying (65) trivially. We proceed to estimate separately the two integrals in (65). First, using the assumption (3) that U is nondecreasing on $[0, \infty)$,

$$\int_{1-\varepsilon}^{1} \exp\Bigl(-\sum_{u:\{u,w\}\in E} U\bigl(\varphi_u - (\varphi_v + t)\bigr)\Bigr)\,dt \le \varepsilon \exp\bigl(-U(1-\varepsilon) - \bigl(\deg(w) - 1\bigr)U(0)\bigr). \tag{67}$$

Second, observe that $\max_{u:\{u,w\}\in E} |\varphi_u - \varphi_v| \le 1\frac{1}{4}$ as $\mathrm{U}_{\vec{e}}$ holds and therefore, using also (66),

$$\varphi_w - \varphi_v \in [1/2, 3/4 - \varepsilon] \text{ implies that } \max_{u:\{u,w\}\in E} |\varphi_u - \varphi_w| \le \frac{3}{4}.$$

Using again the nondecreasing assumption (3) and the assumption that $\varepsilon \le \frac{1}{8}$, we obtain

$$\int_{1/2}^{3/4-\varepsilon} \exp\Bigl(-\sum_{u:\{u,w\}\in E} U\bigl(\varphi_u - (\varphi_v + t)\bigr)\Bigr)\,dt \ge \frac{1}{8} \exp\Bigl(-\deg(w)U\Bigl(\frac{3}{4}\Bigr)\Bigr). \tag{68}$$

Plugging the inequalities (67) and (68) into (65) and comparing with the definition of $\delta(U,\vec{e},\varepsilon)$ verifies the inequality (65) and finishes the proof of the lemma.

6. Discussion and open questions. *Extremal gradients.* Theorem 1.1 provides quantitative estimates on the rarity of extremal gradients in random surfaces satisfying a Lipschitz constraint and having a monotone interaction potential. Its proof makes use of a cluster algorithm for random surfaces and thus provides an alternative to a previous approach via reflection positivity [27], Theorem 3.2. The main advantage of the cluster algorithm approach is that it applies to random surfaces defined on general graphs and thus removes a chief limitation of the reflection positivity proof which is restricted to torus graphs. However, the proof presented in this paper introduces new limitations which it would be desirable to remove. Specifically, the current proof applies only to Lipschitz surfaces with monotone interaction potential whereas one may expect, as put forth explicitly in [27], Section 6, that results analogous to Theorem 1.1 should hold almost without restriction on the potential function (some integrability conditions are required for the model to be well defined) and such a result indeed holds on torus graphs as shown with the reflection positivity method [27], Theorem 3.2.

Theorem 1.1 may be used together with the arguments of [27] to prove the delocalization of random surfaces in cases not previously known. For instance, delocalization would follow for random surfaces whose potential satisfies (3) and (4) (such as the hammock potential (6)) on finite connected domains of $\mathbb{Z}^2$ with Dirichlet boundary conditions (when V_0 is the set of boundary vertices of the domain), or any other choice of the nonempty normalization set V_0. The same techniques should apply to show delocalization on many (finite domains in) infinite graphs on which simple random walk is recurrent.

Extremal gradients in spin systems. Similarly to the previous point, it is also of interest to extend the control of extremal gradients to the spin system setting. Bricmont and Fontaine [6] show that extremal gradients are unlikely in the XY (spin $O(2)$) model (allowing even for multi-body interactions and external magnetic fields). Their proof makes use of Ginibre's extension of Griffiths' inequalities [21, 22] and thus does not extend to the spin $O(n)$ model with $n \geq 3$, where they obtain somewhat weaker results instead. As cluster algorithms are available in some generality for spin $O(n)$ models (as reviewed in Section 2), it is possible that our approach to the control of extremal gradients may be extended to the spin system setting and provide additional results for models in which Ginibre's inequality is unavailable.

Excursion-set percolation. The reflection principle for random surfaces given in Theorem 1.2 may remind the reader of the study of excursion-set percolation in random surfaces. Initiated by Lebowitz–Saleur [26] and Bricmont–Lebowitz–Maes [7], this line of investigation focuses on the percolative properties of the set $\{v \in V : \varphi_v \geq h\}$. Triggered by its relations with the random interlacement model [30, 35] introduced by Sznitman, the subject has recently seen significant activity; see [17] and references within. In these studies, one starts with the infinite-volume limit of a random surface φ on $\mathbb{Z}^d$, typically the Gaussian free field with Dirichlet boundary conditions (see [7, 29] for an exception), and aims to study the set of $h \in \mathbb{R}$ for which there is, almost surely, percolation in the set of vertices v with $\varphi_v \geq h$ (i.e., there is an infinite connected component of vertices v with $\varphi_v \geq h$). For the Gaussian free field, it is known [7, 30] that for each $d \geq 3$ there is an $h_*(d) \in \mathbb{R}$ such that percolation occurs if $h < h_*(d)$ and does not occur if $h > h_*(d)$. Moreover, $h_*(d) > 0$ in any dimension $d \geq 3$; for high dimensions this was shown in [30] and was strengthened to every $d \geq 3$ in the very recent [17].

Theorem 1.2 seems far from the state-of-the-art of these studies but does provide an alternative approach to one of the basic, simple, results in this direction. It was shown in [7] that for any strictly convex potential U and any $\varepsilon > 0$, in the infinite-volume limit on $\mathbb{Z}^d$ of the random surface measure with Dirichlet boundary conditions, the set of vertices x with $\varphi_x \geq -\varepsilon$ percolates almost surely. We discuss this result in the context of Theorem 1.2. Let

$G=(V,E)$ be a finite connected graph, let $V_0 \subseteq V$ be nonempty, let U be a potential satisfying the monotonicity condition (3) and the assumption (2) that μ_{U,G,V_0} is well defined and let φ be randomly sampled from μ_{U,G,V_0}. Then, for any $\varepsilon > 0$ and $v \in V \setminus V_0$, by (10),

$$\mathbb{P}(V_0 \xleftrightarrow{\varphi>-\varepsilon} v) = \mathbb{P}(V_0 \xleftrightarrow{\varphi<\varepsilon} v) = 1 - \mathbb{P}(V_0 \overset{\varphi\leq\varepsilon}{\not\longleftrightarrow} v) \geq \mathbb{P}(|\varphi_v| < \varepsilon). \tag{69}$$

If, in addition, U satisfies the finite-support condition (4) then, relying now on (11), we obtain an inequality in the opposite direction,

$$\mathbb{P}(V_0 \xleftrightarrow{\varphi>-\varepsilon} v) \leq \mathbb{P}(\varphi_v \in (\varepsilon, \varepsilon+1)). \tag{70}$$

When the potential U is strictly convex, one has available the Brascamp–Lieb inequality [4] which bounds the variance of φ_x by the variance of the Gaussian free field on the same graph. Together with (69) this can be used to show that the probability that in a discrete cube graph $\{-L,\dots,L\}^d$ in $\mathbb{Z}^d$, $d \geq 3$, the origin is connected to the boundary of the cube via vertices v with $\varphi_v \geq -\varepsilon$ is uniformly positive as L tends to infinity. Conversely, the same probability in dimension $d=2$ necessarily tends to zero as L increases when U satisfies the finite-support condition and is twice-continuously differentiable on its support, by (70) and the delocalization results of [27].

Correlation of gradients. One approach to bounding the fluctuations of random surfaces proceeds via control of the correlations of gradients of the surface. Consider a random surface φ sampled from the measure μ_{U,G,V_0} (see (1)) with $G=(V,E)$ a finite, connected graph and $V_0 \subseteq V$ a nonempty set on which φ is set to zero. One may then express the height of the surface at some vertex $x \notin V_0$ as a linear combination of the gradients of the surface. That is, if $\vec{E}$ denotes the set of oriented edges of G (both orientations of each edge appear in $\vec{E}$) one writes

$$\varphi_x = \sum_{(u,v)\in\vec{E}} c_{(u,v)}(\varphi_u - \varphi_v)$$

for a suitable choice of coefficients $c_{(u,v)}$. Among the many possible choices of these coefficients, Brascamp, Lieb and Lebowitz [5], Section VII, consider the one obtained from the Green's function g of the graph G by writing

$$\varphi_x = \frac{1}{2}\sum_{(u,v)\in\vec{E}} (g_u - g_v)(\varphi_u - \varphi_v), \tag{71}$$

where g_y is the expected number of visits to x of a simple random walk on G started at y and stopped when it first hits V_0 (the factor $\frac{1}{2}$ is needed since each edge is taken with both orientations). The equality (71) is a consequence of the discrete Green's identity $\sum_{\{u,v\}\in E}(a_u - a_v)(b_u - b_v) = -\sum_{u\in V} a_u(\Delta b)_u$, valid for any two functions a, b on G (with $(\Delta b)_u := \sum_{v\in V:\,\{u,v\}\in E}(b_v - b_u)$), and the fact that $(\Delta g)_y = -\delta_{x,y}$. The identity (71) shows that

$$\begin{aligned}\operatorname{Var}(\varphi_x) = &\sum_{\{u,v\}\in E}(g_u - g_v)^2 \operatorname{Var}(\varphi_u - \varphi_v)\\ &+ \frac{1}{4}\sum_{\substack{(u,v),(z,w)\in\vec{E}\\ \{u,v\}\neq\{z,w\}}} (g_u - g_v)(g_z - g_w)\operatorname{Cov}(\varphi_u - \varphi_v, \varphi_z - \varphi_w)\end{aligned} \tag{72}$$

and thus highlights how bounding the covariances $\operatorname{Cov}(\varphi_u - \varphi_v, \varphi_z - \varphi_w)$ implies an upper bound on the fluctuations of the surface. Indeed, as pointed out in [5], when G is a discrete cube $\{-L,\dots,L\}^d$ in the lattice $\mathbb{Z}^d$, $d \geq 3$, and V_0 is the boundary of this cube, a decay of the covariances faster than $\|u-z\|_2^{-(2+\varepsilon)}$ for large $\|u-z\|$ would imply that the fluctuations of the surface at the origin remain bounded uniformly in L. Such decay is expected but seems quite difficult to establish. Recently, Conlon and Fahim [16] used PDE tools to establish asymptotic formulas for the covariance which imply such decay when the potential function U satisfies $0 < \inf U''(x) \le \sup U''(x) < \infty$ and certain additional assumptions.

The above discussion provides motivation for studying the gradient-gradient covariances of random surfaces as appear in (72). It is interesting whether the cluster algorithms used in this work (see Section 2) can provide additional tools for controlling such covariances. Indeed, one may try to reflect the value of φ at v around the height φ_u, thus reversing the gradient on the edge (u, v). The contribution to the covariance on the event that this reflection leaves the gradient on the edge (z, w) unchanged is exactly zero as the reflection is a measure preserving one-to-one mapping. Thus, this approach connects the problem of estimating the gradient-gradient covariances to the problem of controlling properties of the reflected cluster (of v, when reflecting around the height φ_u) in the cluster algorithm. Related connections in the spin system settings were found by Chayes [13] and Campbell–Chayes [10]. It is unclear whether this connection can simplify the problem.

Acknowledgements. Both authors were partially supported by ISF grant 861/15 and ERC starting grant 678520 (LocalOrder). We thank Martin Tassy for helping to clarify the connection between Sheffield's cluster-swapping algorithm and the general cluster algorithms discussed in Section 2. We also thank Nishant Chandgotia and Yinon Spinka for several useful discussions.

REFERENCES

[1] AIZENMAN, M. (1994). On the slow decay of O(2) correlations in the absence of topological excitations: Remark on the Patrascioiu–Seiler model. *J. Stat. Phys.* **77** 351–359. MR1300539 https://doi.org/10.1007/BF02186846

[2] AIZENMAN, M., HAREL, M. and PELED, R. (2019). Exponential decay of correlations in the 2D random field Ising model. *J. Stat. Phys.* 1–28.

[3] ARMENDÁRIZ, I., FERRARI, P. A. and SOPRANO-LOTO, N. (2015). Phase transition for the dilute clock model. *Stochastic Process. Appl.* **125** 3879–3892. MR3373307 https://doi.org/10.1016/j.spa.2015.05.010

[4] BRASCAMP, H. J. and LIEB, E. H. (1976). On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *J. Funct. Anal.* **22** 366–389. MR0450480 https://doi.org/10.1016/0022-1236(76)90004-5

[5] BRASCAMP, H. J., LIEB, E. H. and LEBOWITZ, J. L. (1975). The statistical mechanics of anharmonic lattices. In *Statistical Mechanics* 379–390. Springer.

[6] BRICMONT, J. and FONTAINE, J.-R. (1981). Correlation inequalities and contour estimates. *J. Stat. Phys.* **26** 745–753. MR0648992 https://doi.org/10.1007/BF01010936

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

[8] BROWER, R. C. and TAMAYO, P. (1989). Embedded dynamics for φ^4 theory. *Phys. Rev. Lett.* **62** 1087.

[9] BURTON, R. and STEIF, J. E. (1995). New results on measures of maximal entropy. *Israel J. Math.* **89** 275–300. MR1324466 https://doi.org/10.1007/BF02808205

[10] CAMPBELL, M. and CHAYES, L. (1998). The isotropic O(3) model and the Wolff representation. *J. Phys. A* **31** L255–L259. MR1621490 https://doi.org/10.1088/0305-4470/31/13/002

[11] CARACCIOLO, S., EDWARDS, R. G., PELISSETTO, A. and SOKAL, A. D. (1993). Wolff-type embedding algorithms for general nonlinear σ-models. *Nuclear Phys. B* **403** 475–541.

[12] CHANDGOTIA, N., PELED, R., SHEFFIELD, S. and TASSY, M. (2018). Delocalization of uniform graph homomorphisms from $\mathbb{Z}^2$ to $\mathbb{Z}$. arXiv preprint arXiv:1810.10124.

[13] CHAYES, L. (1998). Discontinuity of the spin-wave stiffness in the two-dimensional XY model. *Comm. Math. Phys.* **197** 623–640. MR1652795 https://doi.org/10.1007/s002200050466

[14] CHAYES, L. and MACHTA, J. (1997). Graphical representations and cluster algorithms I. Discrete spin systems. *Phys. A* **239** 542–601.

[15] CHAYES, L. and MACHTA, J. (1998). Graphical representations and cluster algorithms II. *Phys. A* **254** 477–516.

[16] CONLON, J. G. and FAHIM, A. (2015). Long range correlation inequalities for massless Euclidean fields. *Illinois J. Math.* **59** 143–187. MR3459632

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

[18] EDWARDS, R. G. and SOKAL, A. D. (1988). Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm. *Phys. Rev. D* **38** 2009–2012. MR0965465 https://doi.org/10.1103/PhysRevD.38.2009

[19] EVERTZ, H. G., HASENBUSCH, M., MARCU, M., PINN, K. and SOLOMON, S. (1991). Stochastic cluster algorithms for discrete Gaussian (SOS) models. *Phys. Lett. B* **254** 185–191.

[20] FUNAKI, T. and SPOHN, H. (1997). Motion by mean curvature from the Ginzburg–Landau $\nabla\phi$ interface model. *Comm. Math. Phys.* **185** 1–36. MR1463032 https://doi.org/10.1007/s002200050080

[21] GINIBRE, J. (1970). General formulation of Griffiths' inequalities. *Comm. Math. Phys.* **16** 310–328. MR0269252

[22] GRIFFITHS, R. B. (1967). Correlation in Ising ferromagnets I, II. *J. Math. Phys.* **8** 478–489.

[23] GUO, H. and JERRUM, M. (2017). Random cluster dynamics for the Ising model is rapidly mixing. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms* 1818–1827. SIAM, Philadelphia, PA. MR3627847 https://doi.org/10.1137/1.9781611974782.118

[24] HASENBUSCH, M., LANA, G., MARCU, M. and PINN, K. (1992). Cluster algorithm for a solid-on-solid model with constraints. *Phys. Rev. B* **46** 10472.

[25] KANDEL, D. and DOMANY, E. (1991). General cluster Monte Carlo dynamics. *Phys. Rev. B* **43** 8539.

[26] LEBOWITZ, J. L. and SALEUR, H. (1986). Percolation in strongly correlated systems. *Phys. A* **138** 194–205. MR0865243 https://doi.org/10.1016/0378-4371(86)90180-9

[27] MIŁOŚ, P. and PELED, R. (2015). Delocalization of two-dimensional random surfaces with hard-core constraints. *Comm. Math. Phys.* **340** 1–46. MR3395146 https://doi.org/10.1007/s00220-015-2419-4

[28] PATRASCIOIU, A. and SEILER, E. (1992). Phase structure of two-dimensional spin models and percolation. *J. Stat. Phys.* **69** 573–595. MR1193851 https://doi.org/10.1007/BF01050426

[29] RODRIGUEZ, P.-F. (2016). Decoupling inequalities for the Ginzburg–Landau $\nabla\varphi$ models. arXiv preprint arXiv:1612.02385.

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

[31] SHEFFIELD, S. (2005). Random surfaces. *Astérisque* **304** vi+175. MR2251117

[32] SOKAL, A. (1997). Monte Carlo methods in statistical mechanics: Foundations and new algorithms. In *Functional Integration* (*Cargèse*, 1996). *NATO Adv. Sci. Inst. Ser. B Phys.* **361** 131–192. Plenum, New York. MR1477456

[33] SOPRANO LOTO, N. (2015). Transición de fase para modelos diluidos con espines discretos y medidas de Young-Gibbs para el modelo de Ising. Ph.D. thesis, Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. http://cms.dm.uba.ar/academico/carreras/doctorado/tesis_sopranoloto.pdf.

[34] SWENDSEN, R. H. and WANG, J.-S. (1987). Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.* **58** 86.

[35] SZNITMAN, A.-S. (2010). Vacant set of random interlacements and percolation. *Ann. of Math.* (2) **171** 2039–2087. MR2680403 https://doi.org/10.4007/annals.2010.171.2039

[36] ULLRICH, M. (2014). Swendsen–Wang is faster than single-bond dynamics. *SIAM J. Discrete Math.* **28** 37–48. MR3148642 https://doi.org/10.1137/120864003

[37] VAN DEN BERG, J. (1993). A uniqueness condition for Gibbs measures, with application to the 2-dimensional Ising antiferromagnet. *Comm. Math. Phys.* **152** 161–166. MR1207673

[38] VAN DEN BERG, J. and STEIF, J. E. (1994). Percolation and the hard-core lattice gas model. *Stochastic Process. Appl.* **49** 179–197. MR1260188 https://doi.org/10.1016/0304-4149(94)90132-5

[39] WANG, J.-S., SWENDSEN, R. H. and KOTECKÝ, R. (1989). Antiferromagnetic Potts models. *Phys. Rev. Lett.* **63** 109.

[40] WOLFF, U. (1989). Collective Monte Carlo updating for spin systems. *Phys. Rev. Lett.* **62** 361.