Article:
Björnberg, J.E. and Lees, B. (2023) Correlation inequalities for the uniform eight-vertex model and the toric code model. Reviews in Mathematical Physics, 35 (10). 2350029. ISSN 0129-055X

https://doi.org/10.1142/s0129055x23500290

CORRELATION INEQUALITIES FOR THE UNIFORM 8-VERTEX MODEL AND THE TORIC CODE MODEL

J. E. BJÖRNBERG AND B. LEES

Abstract. We investigate connections between four models in statistical physics and probability theory: (1) the toric code model of Kitaev, (2) the uniform eight-vertex model, (3) random walk on a hypercube, and (4) a classical Ising model with four-body interaction. As a consequence of our analysis (and of the GKS-inequalities for the Ising model) we obtain correlation inequalities for the toric code model and the uniform eight-vertex model.

1. Introduction

Recent years have seen a rapid development of research on vertex models, primarily the six-vertex model. The six-vertex model was originally introduced by Pauling as a simple model of hydrogen bonding in water ice. In simple terms, the model is defined by taking a subgraph of $\mathbb{Z}^2$ with nearest neighbour edges and imposing a direction on each edge so that every vertex has precisely two edges (or arrows) directed towards it. This gives six possibilities illustrated by vertex types $\texttt{I}, \dots, \texttt{VI}$ in Figure 1.1. Lieb [26, 27, 28] carried out pioneering work on the model using the Bethe Ansatz originally developed for the Heisenberg spin chain [33]. Recently, the model has received growing attention from the probability community, and progress on the rigorous understanding of aspects of these models such as the (de)localisation of its associated height function for various values of its parameters has been immense, e.g. [10, 11, 12, 18].

The eight-vertex model is a generalisation of the six-vertex model that allows two extra types of vertices: *sources* and *sinks*, illustrated by VII and VIII in Figure 1.1. Both the six- and eight-vertex models are integrable lattice models [15, 32]. Using methods coming from integrability, Baxter [4] computed the free energy per site. For information on integrable models we direct the reader to the book by Baxter [5].

In this work we consider the eight-vertex model from a different perspective than solvability/integrability. We consider a simple dynamics which preserve eight-vertex configurations: select a *plaquette* (face) of the square lattice at random and reverse the direction of its four bounding edges. Because every vertex has an even number (either zero or two) of its incident edges reversed, this dynamics preserve the set of allowed configurations. (A

Date: July 4, 2023.

JEB: University of Gothenburg and Chalmers University of Technology, Sweden.

BL: School of Mathematics, University of Leeds, Leeds, UK.

This manuscript has no associated data.

variant of the dynamics preserves the six-vertex model and is in fact in detailed balance with the six-vertex model's probability distribution [1].)

In the case of uniform weights for vertices, we find that consideration of these simple dynamics allows us to determine its communicating classes and their sizes, as well as a simple way to sample from its limiting distribution – the uniform measure on configurations. These useful considerations assist us in determining the *emptiness formation probability* of the model with *domain wall boundary conditions* and the *entropy* of the model, in particular we show (Proposition 4.7) that the entropy of the uniform eight-vertex model is realised for any fixed, valid, boundary conditions. These results use simple arguments. See Sections 4.2 and 4.3 for precise definitions and statements. We also derive a correlation inequality for the uniform eight-vertex model and for other related models, this is discussed in detail below.

The domain wall boundary condition means that we take a box in $\mathbb{Z}^2$ and fix a boundary condition so that arrows at the top and bottom of the boundary point in towards the box and arrows on the left and right boundary point away from the box, or vice-versa. See Figure 4.2 for an illustration. This boundary condition was introduced in [23] for the six-vertex model to study correlation functions for exactly solvable models. There is also a connection to enumeration of alternating sign matrices that was used to find expressions for the free energy density in the disordered and ferroelectric phase and a connection to domino tilings [24]. These findings showed that the behaviour is quite different to the case of periodic boundary conditions, a feature of the model that one might not expect in models that are not as tightly constrained.

Emptiness formation is the event that an entire column of horizontal edges has the same direction (left, say), see Figure 4.2 for an illustration. This non-local correlation function allowed the investigation of limit shapes in the six-vertex model [8].

There is a close connection between this dynamics and Kitaev's *toric code model*. Kitaev's model (at zero temperature) is an important example of a *quantum code*, which are motivated by problems in quantum computing. While quantum computers would allow certain computations to be performed much faster than a conventional computer (e.g. factoring large numbers [30] or searching unstructured databases [20, 21]), a major barrier in realising this potential is the proclivity for the quantum bits (qubits) of the computer to have errors. One way to try to overcome this problem is to use multiple physical qubits to encode a single *logical* qubit that performs better than its individual physical qubit components, primarily due to the ability to recover the intended state of the logical qubit even after one or more of the component physical qubits have experienced an error. We direct the reader to [29] for an accessible treatment of major topics in quantum information and quantum computation. Surface codes, such as Kitaev's toric code, are examples of such a scheme that have enjoyed considerable attention, in part due to their relatively high tolerance for local errors [31]. For an introduction to the various aspects of the toric/surface code and how it operates, we direct the reader to the review [14]. A brief summary may be found in Appendix A.

Using the dynamics indicated above facilitates explicit computations of certain thermodynamic quantities for the toric code model at positive temperature. Guided by these calculations and their consequences, we establish a connection between certain correlations in the toric code model and expectations in a many-body classical Ising model. The ground state of the latter Ising model also gives the eight-vertex model with uniform vertex weights. Using well-known correlation inequalities for the Ising model (GKS) we obtain correlation inequalities for the uniform eight-vertex model and the toric code model. The following is an example of our correlation inequality for the uniform eight-vertex model: given any two sets of vertices, consider the events that they are all sources or sinks. Then these events are positively correlated, i.e. the probability of both occurring is at least as large as the product of the probabilities of the two individual events. The full statement appears in Theorem 1.1.

Correlation inequalities are major tools in the analysis of classical and quantum systems, as well as in probability more widely. The GKS inequalites, first proven by Griffiths for the Ising model [17] and then extended to a more general framework by Ginibre [16], were used in the construction of infinite volume Gibbs states for the Ising model in dimension at least 2. The construction of infinite volume Gibbs states is essential for the rigorous understanding of phase transitions, arguably the most important phenomena of systems studied in statistical mechanics. The Ising model also enjoys other correlation inequalities, such as the FKG inequality. The FKG inequality has been invaluable for the study of the Ising model as well at the *random cluster* model [19]. The literature making use of correlation inequalities for the Ising model and various related models is too vast to do justice to here, we direct the reader to [9] and references therein for a glimpse of this enormous and interesting story. For other models, such as the classical and quantum XY model, GKS inequalities are also major tools for constructing infinite volume Gibbs states. Additionally, in the classical case, they have been used to show monotonicity of the *spontaneous magnetisation* with respect to temperature and dimension. In the classical and quantum case they allow a rigorous comparison of the critical temperature for phase transition in the model with the critical temperature for the Ising model. We direct the reader to [6, 7] for statements and proofs of these inequalities and overview of the GKS inequality and its applications.

1.1. **Setting and results.**

For $m, n \geq 1$, let $V = \{1, \dots, m\} \times \{1, \dots, n\} \subseteq \mathbb{Z}^2$. We view V as the vertex set of the $m \times n$ torus Λ. (It is possible to work with Λ as a subset of $\mathbb{Z}^2$ with a boundary; we do this in Section 4.) Let E denote the set of edges of Λ and F the set of faces. In keeping with common terminology in the area, faces will often be referred to as *plaquettes* and denoted $p \in F$, and vertices referred to as *stars* and denoted $s \in V$.

The first model we consider is the *uniform eight-vertex model*, defined as follows. Let Δ denote the set of assignments of directions to the edges E, meaning that each horizontal edge is directed either $\rightarrow$ or $\leftarrow$, and each vertical edge either $\uparrow$ or $\downarrow$. Elements of Δ will sometimes be referred to as *arrow-configurations*. Let $\Delta_{8vx} \subseteq \Delta$ denote the set of arrow-configurations such that the number of arrows pointing towards any vertex $s \in V$ is even. At each vertex $s \in V$, there are eight possible configurations of arrows, depicted in Figure 1.1. We refer to these eight local configurations by the roman numerals $\texttt{I}, \texttt{II}, \dots, \texttt{VIII}$. The first six of these are the allowed configurations for the six-vertex model; $\texttt{VII}$ is called a *sink* and $\texttt{VIII}$ a *source*. We let $\mu(\cdot)$ denote the uniform probability measure on $\Delta_{8\text{vx}}$; this is the probability measure governing the uniform 8-vertex model.

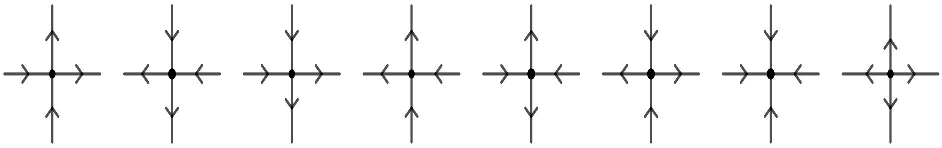

FIGURE 1. The eight vertex types $\texttt{I}, \texttt{II}, \dots, \texttt{VIII}$ for arrow-configurations in $\Delta_{8\text{vx}}$.

The next model we consider is Kitaev's *toric code model* [22], defined as follows. For E the edge-set of the torus Λ as above, let $(\mathbb{C}^2)^{\otimes E}$ be the tensor product of one copy of $\mathbb{C}^2$ for each edge $e \in E$. Using the standard basis $|+\rangle = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$, $|-\rangle = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ for $\mathbb{C}^2$, one obtains a (product) basis for $(\mathbb{C}^2)^{\otimes E}$ with elements $|\omega\rangle$ for $\omega \in \Omega = \{-1, +1\}^E$. (Note that we use the statistical physics convention of labelling this basis $|\pm\rangle$, while in quantum information theory it would be more common to use $|0\rangle$ and $|1\rangle$.) Consider the Pauli matrices

$$\sigma^{(1)} = \left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right), \qquad \sigma^{(2)} = \left(\begin{smallmatrix}0&-i\\i&0\end{smallmatrix}\right), \qquad \sigma^{(3)} = \left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right), \tag{1}$$

and write $\sigma_e^{(j)}$ for the linear operator on $(\mathbb{C}^2)^{\otimes E}$ which acts as $\sigma^{(j)}$ on the e factor and the identity elsewhere. Since each $\sigma_e^{(3)}$ acts diagonally on the basis elements $|\omega\rangle$, we refer to this basis as the $\sigma^{(3)}$ product basis.

Introduce the operators

$$Z_s = \prod_{e\sim s} \sigma_e^{(3)}, \qquad X_p = \prod_{e\sim p} \sigma_e^{(1)}, \qquad s \in V, \quad p \in F, \tag{2}$$

where $e \sim s$ and $e \sim p$ mean that the edge e is adjacent to the vertex s or the face p, respectively. (The operators Z_s and X_p are more commonly denoted A_s and B_p, respectively.) For $\mathbf{J}^{\mathrm{x}} = (J_p^{\mathrm{x}} : p \in F)$ and $\mathbf{J}^{\mathrm{z}} = (J_s^{\mathrm{z}} : s \in V)$ vectors of constants satisfying $J_p^{\mathrm{x}}, J_s^{\mathrm{z}} \geq 0$, the *Hamiltonian operator* is defined as

$$H = H(\mathbf{J}^{\mathrm{x}}, \mathbf{J}^{\mathrm{z}}) = -\sum_{p\in F} J_p^{\mathrm{x}} X_p - \sum_{s\in V} J_s^{\mathrm{z}} Z_s. \tag{3}$$

For all $J_p^{\mathrm{x}} = J_s^{\mathrm{z}} = 1$ this is Kitaev's H_0, see [22]. All terms in H commute, since for any star and plaquette the number of edges belonging to both of them is even.

The Hamiltonian (3) defines the toric code model through the equilibrium states defined as follows: for an operator A on $(\mathbb{C}^2)^{\otimes E}$,

$$\langle A\rangle = \langle A\rangle_{\mathbf{J}^{\mathrm{x}},\mathbf{J}^{\mathrm{z}}} = \frac{\operatorname{tr}[A\, e^{-H}]}{\operatorname{tr}[e^{-H}]}. \tag{4}$$

If we take $J_p^{\mathrm{x}} = J_s^{\mathrm{z}} = \beta \to \infty$ we obtain the ground state which is important in the theory of quantum codes; see Appendix A. We will see below that the limit of all $J_p^{\mathrm{x}} \to \infty$, with all $J_s^{\mathrm{z}} = \beta$ fixed, essentially gives the uniform 8-vertex model.

By further relating the uniform 8-vertex model and the toric code model to a classical Ising model, defined in the next subsection, we will prove certain correlation inequalities. The setting is as follows. First, for the 8-vertex model, let $\eta \in \{\mathtt{I}, \dots, \mathtt{VIII}\}^V$ be any 8-vertex configuration. For $C \subseteq V$, let $\mathcal{A}_C^\eta$ be the event that, for each $s \in C$, the arrows around s are either all equal to those of η or all the opposite to those of η. Additionally, for $\nu \in \{\mathtt{I}, \dots, \mathtt{VIII}\}^V$ with $\nu \neq \eta$ let $\mathcal{A}_C^{\eta;\nu}$ be the event that, for each $s \in C$, the arrows around s are either all equal to those of η or ν or all the opposite to those of η or ν (this trivially includes the event $\mathcal{A}_C^\eta$ when ν has all arrows opposite to those of η).

For the toric code the setting is similar but more complex to state. For $s \in V$, consider the four edges e_1, e_2, e_3 and e_4 adjacent to s and write $\sigma_i^{(3)}(s)$ for $\sigma_{e_i}^{(3)}$. Let $\varepsilon \in \{-1, +1\}^4$, satisfying $\prod_{i=1}^4 \varepsilon_i = +1$, be a local sign-configuration. Define the operators

$$P_s^{(3)}(\varepsilon) = \tfrac{1}{16} \prod_{i=1}^4 \big(1 + \varepsilon_i \sigma_i^{(3)}(s)\big), \qquad \bar{P}_s^{(3)}(\varepsilon) = \tfrac{1}{16} \prod_{i=1}^4 \big(1 - \varepsilon_i \sigma_i^{(3)}(s)\big). \tag{5}$$

These are projection operators onto the subspaces of $(\mathbb{C}^2)^{\otimes E}$ spanned by elements $|\omega\rangle$ satisfying $\omega_{e_i} = \varepsilon_i$, respectively $\omega_{e_i} = -\varepsilon_i$, for $i \in \{1, 2, 3, 4\}$. For $C \subseteq V$ let

$$Q_C^{(3)}(\varepsilon) = \prod_{s \in C} (P_s^{(3)}(\varepsilon) + \bar{P}_s^{(3)}(\varepsilon)). \tag{6}$$

Note that each factor $(P_s^{(3)}(\varepsilon) + \bar{P}_s^{(3)}(\varepsilon))$ respects the symmetry of reversing all the signs around s and that $Q_C^{(3)}(\varepsilon)$ is an operator version of the event $\mathcal{A}_C^\eta$ for appropriately paired ε, η. For $\delta \in \{-1, +1\}^4$ with $\delta \neq \pm\varepsilon$ define the analog of $\mathcal{A}_C^{\eta;\nu}$ by

$$Q_C^{(3)}(\varepsilon, \delta) = \prod_{s \in C} (P_s^{(3)}(\varepsilon) + \bar{P}_s^{(3)}(\varepsilon) + P_s^{(3)}(\delta) + \bar{P}_s^{(3)}(\delta)). \tag{7}$$

We also define similar operators for faces, but working in the $\sigma^{(1)}$-basis rather than the $\sigma^{(3)}$-basis. For $p \in F$ consider the four edges e_1, e_2, e_3 and e_4 surrounding p and write $\sigma_i^{(1)}(p) = \sigma_{e_i}^{(1)}$. Let $\varepsilon \in \{-1, +1\}^4$ be a local sign-configuration. Define the operators

$$P_p^{(1)}(\varepsilon) = \tfrac{1}{16} \prod_{i=1}^4 \big(1 + \varepsilon_i \sigma_i^{(1)}(p)\big), \qquad \bar{P}_p^{(1)}(\varepsilon) = \tfrac{1}{16} \prod_{i=1}^4 \big(1 - \varepsilon_i \sigma_i^{(1)}(p)\big), \tag{8}$$

and for $D \subseteq F$ let

$$Q_D^{(1)}(\varepsilon) = \prod_{p \in D} (P_p^{(1)}(\varepsilon) + \bar{P}_p^{(1)}(\varepsilon)). \tag{9}$$

For $\delta \in \{-1, +1\}^4$ with $\delta \neq \pm\varepsilon$ define

$$Q_D^{(1)}(\varepsilon, \delta) = \prod_{p \in D} (P_p^{(1)}(\varepsilon) + \bar{P}_p^{(1)}(\varepsilon) + P_p^{(1)}(\delta) + \bar{P}_p^{(1)}(\delta)). \tag{10}$$

Finally, we say that $C \subseteq V$ is *contractible* if C does not contain a set of nearest-neighbour vertices that forms a non-contractible loop on the torus

Λ and that $D \subseteq F$ is *contractible* if D does not contain a set of nearest-neighbour faces that forms a non-contractible loop on the torus Λ. Note that a non-contractible loop which wraps around the torus horizontally or vertically necessarily has length which is of the same parity as the corresponding side-length m or n, thus the loop can only have even length if the corresponding side-length is even.

THEOREM 1.1. *Let η, ν, ε, δ, ε', δ' with $\nu \neq \eta$, $\delta \neq \pm\varepsilon$, and $\varepsilon' \neq \delta'$ be as above and let $C_1, C_2 \subseteq V$ be sets of vertices such that*

- *either $C_1 \cup C_2$ is contractible,*
- *or any non-contractible loop in $C_1 \cup C_2$ has even length.*

Further let $C \subset V$ and $D \subset F$ be such that both C and D are either contractible or any non-contractible loop in them has even length. Then

(1) For the uniform 8-vertex model: $\mu(\mathcal{A}^{\eta}_{C_1} \cap \mathcal{A}^{\eta}_{C_2}) \geq \mu(\mathcal{A}^{\eta}_{C_2})\mu(\mathcal{A}^{\eta}_{C_2})$ and $\mu(\mathcal{A}^{\eta;\nu}_{C_1} \cap \mathcal{A}^{\eta;\nu}_{C_2}) \geq \mu(\mathcal{A}^{\eta;\nu}_{C_2})\mu(\mathcal{A}^{\eta;\nu}_{C_2})$.

(2) For the toric code model: $\langle Q^{(3)}_{C_1}(\varepsilon) Q^{(3)}_{C_2}(\varepsilon)\rangle \geq \langle Q^{(3)}_{C_1}(\varepsilon)\rangle\langle Q^{(3)}_{C_2}(\varepsilon)\rangle$, and $\langle Q^{(3)}_{C_1}(\varepsilon,\delta) Q^{(3)}_{C_2}(\varepsilon,\delta)\rangle \geq \langle Q^{(3)}_{C_1}(\varepsilon,\delta)\rangle\langle Q^{(3)}_{C_2}(\varepsilon,\delta)\rangle$.

(3) For the toric code model: $\langle Q^{(3)}_{C}(\varepsilon) Q^{(1)}_{D}(\delta)\rangle \leq \langle Q^{(3)}_{C}(\varepsilon)\rangle\langle Q^{(1)}_{D}(\delta)\rangle$, and $\langle Q^{(3)}_{C}(\varepsilon,\delta) Q^{(1)}_{D}(\varepsilon',\delta')\rangle \leq \langle Q^{(3)}_{C}(\varepsilon,\delta)\rangle\langle Q^{(1)}_{D}(\varepsilon',\delta')\rangle$.

The interpretation of the first part of the theorem is that the two events $\mathcal{A}^{\eta}_{C_1}, \mathcal{A}^{\eta}_{C_2}$ are more likely to occur at the same time, than if they had been independent. This is usually referred to as *positive correlation*. Such monotonicity properties are useful in many situations, for example when establishing existence of limits.

As an example of the correlation inequality for the 8-vertex model, we can take η to be the constant vector of all `VII`'s, i.e. only *sinks*. Then $\mathcal{A}^{\eta}_{C}$ is the event that every vertex in C is either a source or a sink, and the Theorem says that $\mathcal{A}^{\eta}_{C_1}$ and $\mathcal{A}^{\eta}_{C_2}$ are positively correlated. This is primarily interesting when C_1 and C_2 are not too far apart: we will see in Section 4.1 that if C_1 and C_2 are not adjacent to any common faces, then the states of the vertices in C_1 and C_2 are in fact independent under $\mu(\cdot)$.

1.2. **Relations between the models.** As mentioned above, Theorem 1.1 is built on relating the two models (uniform 8-vertex and toric code) to a classical Ising model. We now define the latter. Let $\mathbf{J} = (J_s : s \in V)$ be a vector with all $J_s \geq 0$. Recalling that $\Omega = \{-1,+1\}^E$, define the probability measure $\mathbb{P}_{\mathbf{J}}(\cdot)$ on Ω by

$$\mathbb{P}_{\mathbf{J}}(\sigma) = \frac{\exp\left(\sum_{s\in V} J_s \prod_{e\sim s}\sigma_e\right)}{\sum_{\omega\in\Omega}\exp\left(\sum_{s\in V} J_s \prod_{e\sim s}\omega_e\right)}. \tag{11}$$

This is a ferromagnetic Ising model with four-body interaction, and the restriction of the toric code model to a certain class of observables is equivalent to this model, as stated in the next result. Note that we take $\mathbf{J} = \mathbf{J}^{\mathrm{z}}$ in the statement.

Proposition 1.2. *Let Q be any observable diagonal in the $\sigma^{(3)}$ product basis, with $Q|\sigma\rangle = q(\sigma)|\sigma\rangle$ for any $\sigma \in \Omega$. Then*

$$\langle Q\rangle = \mathbb{E}_{\mathbf{J}^{z}}[q(\sigma)], \tag{12}$$

where $\mathbb{E}_{\mathbf{J}}[\cdot]$ denotes expectation with respect to $\mathbb{P}_{\mathbf{J}}(\cdot)$. In particular, $\langle Q\rangle$ does not depend on $\mathbf{J}^{x}$.

Note also that if $J_s \to \infty$ for all $s \in V$ then $\lim_{J\to\infty} \mathbb{P}_J(\cdot)$ is supported on the set $\Omega_{8\mathrm{vx}} \subseteq \Omega$ of configurations σ satisfying $\prod_{e\sim s}\sigma_e = +1$ for all $s \in V$, since $\sum_{s\in V}\prod_{e\sim s}\sigma_e$ attains its maximum value $|V|$ for such σ. This observation will allow us to (essentially) identify the uniform 8-vertex measure $\mu(\cdot)$ with $\lim_{J\to\infty}\mathbb{P}_J(\cdot)$.

The key fact we use about the Ising model (11) is that it satisfies GKS-inequalities (see e.g. [13, Theorem 3.49]): for any sets $A, B \subseteq E$ we have

$$\mathbb{E}_{\mathbf{J}}\Big[\prod_{e\in A}\sigma_e\prod_{e\in B}\sigma_e\Big] \geq \mathbb{E}_{\mathbf{J}}\Big[\prod_{e\in A}\sigma_e\Big]\mathbb{E}_{\mathbf{J}}\Big[\prod_{e\in B}\sigma_e\Big], \qquad \mathbb{E}_{\mathbf{J}}\Big[\prod_{e\in A}\sigma_e\Big] \geq 0. \tag{13}$$

This will be used in conjunction with Propositon 1.2 in order to establish the first two parts of Theorem 1.1. For the third part of the Theorem, we will use [6, Theorem 1], which implies that for any two subsets $A, B \subseteq E$ we have

$$\Big\langle \prod_{e\in A}\sigma_e^{(3)}\prod_{e\in B}\sigma_e^{(1)}\Big\rangle \leq \Big\langle \prod_{e\in A}\sigma_e^{(3)}\Big\rangle\Big\langle \prod_{e\in B}\sigma_e^{(1)}\Big\rangle. \tag{14}$$

The fourth model which we use in our analysis is *simple random walk on the hypercube*. To define this, let $\mathbb{F}_2^{+}$ denote the two-element group with elements $\{0,1\}$ satisfying $1+1=0$, and consider $G = (\mathbb{F}_2^{+})^F$ with generators g^p given by

$$g_j^p = \begin{cases} 1, & \text{if } j = p,\\ 0, & \text{otherwise.}\end{cases} \tag{15}$$

Then G is a hypercube of dimension $|F|$. A random walk on G is given by selecting, independently at random, faces $p_1, p_2, \dots, p_k$ and letting $X(k) = X(0) + \sum_{i=1}^{k} g^{p_i}$ where $X(0)$ is a starting position. Here at each step we choose face p with probability $J_p/\sum_{q\in F}J_q$.

The relation to the previous models comes by letting G act on the sets Ω and Δ. The intuition is simple: g^p acts on elements of Δ by reversing all the arrows around the face p, and on Ω by negating the $\pm$ signs on the edges surrounding p. It is easy to see that this action leaves $\Delta_{8\mathrm{vx}} \subseteq \Delta$ invariant, and also leaves the measures $\mu(\cdot)$ and $\mathbb{P}_{\mathbf{J}}(\cdot)$ invariant. The connection to the toric code model is slightly more subtle: we will see in Section 2.1 that the term $\sum_{p\in F}J_p^{x}X_p$ in the Hamiltonian (3) is essentially the generator matrix of (the continuous-time version of) simple random walk on G. This allows us to express the equilibrium states (4) for H using transition probabilities for the random walk, which will be instrumental in understanding some basic features of the toric code model, including the relation to the Ising model, Proposition 1.2.

1.3. **Outline.** In Section 2 we develop the tools which we will use to prove our results, including the dynamics of plaquette-flipping and the connection to the Ising model. In particular, we prove Proposition 1.2 at the end of Section 2. In Section 3 we focus on correlation inequalities, and in particular prove Theorem 1.1. In Section 4 we prove the results about emptiness formation probabilities and entropy described in Section 1, which are natural applications of our methods.

1.4. **Acknowledgements.** JEB gratefully acknowledges financial support from Vetenskapsrådet, grant 2019-04185, from *Ruth och Nils Erik Stenbäcks stiftelse*, and from the Sabbatical Program at the Faculty of Science, University of Gothenburg, as well as kind hospitality at the University of Warwick and the University of Bristol. Both authors thank the referee for useful comments.

2. Properties of the toric code model

2.1. **Dynamic description.** Recall that (below (15)) we defined random walk on the hypercube $G = (\mathbb{F}_2^+)^F$ as the process obtained by sampling independently at random faces $p_1, p_2, \dots$ and successively adding them to an initial element of G. We make this into a continuous-time random walk using a Poisson process $N(t)$ of rate $|F|$ and defining

$$X(t) = X(0) + \sum_{i=1}^{N(t)} g^{p_i}. \tag{16}$$

An equivalent description is that each face $p \in F$ is assigned an independent exponentially distributed random 'clock' of rate J_p and that we add 1 in the position of p when the corresponding clock rings.

We next map the random walk (16) onto a process in $\Omega = \{-1, +1\}^E$ as well as onto a process in the set Δ of arrow-configurations. As already described for the discrete-time case, the process in Ω proceeds by 'flipping' (negating) all signs on the edges around the selected face p, and the process in Δ similarly reverses the arrows around p. We may describe this more formally as follows. The group G acts on the set Ω by

$$g^p : \omega \mapsto x^p \cdot \omega, \qquad \text{where} \quad x_e^p = \begin{cases} -1, & \text{if } e \sim p, \\ +1, & \text{otherwise,} \end{cases} \tag{17}$$

and $\cdot$ denotes component-wise multiplication. (We are implicitly presenting $\mathbb{F}_2^+$ as the multiplicative group with elements $\{-1, +1\}$.) Then we define $\omega(t)$ as the result of applying (17) for the randomly selected faces p_i in (16):

$$\omega(t) = \Big(\prod_{i=1}^{N(t)} x^{p_i} \Big) \cdot \omega(0). \tag{18}$$

Here $\omega(0) \in \Omega$ is an initial configuration, and since the group G is commutative we do not need to specify an order of multiplication in (18).

It will be useful to record here the generator-matrix for the random walk (18). Indeed, regarding the x^p of (17) as a matrix with rows and columns indexed by Ω, the generator matrix can be written as

$$\sum_{p \in F} J_p (x^p - 1), \tag{19}$$

where 1 denotes the identity matrix. It follows that the transition-probabilities at time t are given by the matrix

$$\exp\big(t \textstyle\sum_{p\in F} J_p(x^p - 1)\big). \tag{20}$$

To formally define the process in Δ, we note that we may think of $\omega \in \Omega$ as encoding an arrow-configuration in Δ by indicating which edges are reversed (or not) with respect to some a-priori configuration of arrows. More precisely, fix a reference-configuration $\rho \in \Delta$ and define, for $\omega \in \Omega$, the arrow-configuration $\omega \cdot \rho$ such that the arrow at e has the *same* orientation as in ρ if $\omega_e = +1$, respectively the *opposite* orientation if $\omega_e = -1$. We then obtain a random walk $\delta(t) \in \Delta$ by

$$\delta(t) = \omega(t) \cdot \rho. \tag{21}$$

Note that $\delta(t)$ depends both on the reference-configuration $\rho \in \Delta$ and the initial sign-configuration $\omega(0) \in \Omega$.

We next describe the relevance of the random walk $\omega(t)$ for the toric code model. Recall that the elements $\omega \in \Omega$ are in one-to-one correspondence with basis-vectors $|\omega\rangle$ for $(\mathbb{C}^2)^{\otimes E}$, given as the product-basis obtained from the basis $|+\rangle = \big(\begin{smallmatrix}1\\0\end{smallmatrix}\big)$, $|-\rangle = \big(\begin{smallmatrix}0\\1\end{smallmatrix}\big)$ for $\mathbb{C}^2$. Thus the action of G on Ω carries over to a representation of G on $(\mathbb{C}^2)^{\otimes E}$. Moreover, from the explicit form of the Pauli-matrices (1) we see that $\sigma^{(1)}|\pm\rangle = |\mp\rangle$. Thus the operator X_p in (2) acts precisely as x^p in (17). It follows from (19) that the term $\sum_{p\in F} J_p^{\mathrm{x}} X_p$ in the Hamiltonian (3) is, up to adding a multiple of the identity, the generator for a random walk $|\omega(t)\rangle$ on the set of basis-vectors.

We note here that the random walk on $\Delta_{8\mathrm{vx}}$ is not irreducible, i.e. the set decomposes into several disjoint communicating classes. See Proposition 4.1.

2.2. **Duality.** There are two useful notions of duality for the toric code model: that of the lattice Λ as a planar graph, as well as a duality between the operators Z_s and X_p. In this subsection we aim to make precise these dualities as well as the connection between them.

We start by looking at the operators Z_s and X_p. Recall that we have been using a basis for $\mathbb{C}^2$ in which $\sigma^{(3)}$ is diagonal and the other Pauli matrices are given as in (1). For the rest of this subsection, we denote this basis by $|+\rangle_{\mathrm{z}} = \big(\begin{smallmatrix}1\\0\end{smallmatrix}\big)$, $|-\rangle_{\mathrm{z}} = \big(\begin{smallmatrix}0\\1\end{smallmatrix}\big)$ where the subscript z serves to indicate that the third Pauli matrix is diagonal. The corresponding product basis for $(\mathbb{C}^2)^{\otimes E}$ will be denoted $|\omega\rangle_{\mathrm{z}}$ for $\omega \in \Omega$. Recall that, in this basis, the operator X_p negates the signs on all edges e surrounding $p \in F$, while Z_s is diagonal.

We can also consider a basis for $\mathbb{C}^2$ in which $\sigma^{(1)}$ is diagonal: define

$$|+\rangle_{\mathrm{x}} = \frac{|+\rangle_{\mathrm{z}} + |-\rangle_{\mathrm{z}}}{\sqrt{2}} \qquad \text{and} \qquad |-\rangle_{\mathrm{x}} = \frac{|+\rangle_{\mathrm{z}} - |-\rangle_{\mathrm{z}}}{\sqrt{2}}. \tag{22}$$

The corresponding product basis for $(\mathbb{C}^2)^{\otimes E}$ will be denoted $|\omega\rangle_{\mathrm{x}}$ for $\omega \in \Omega$. The basis-change matrix for going between $|\pm\rangle_{\mathrm{z}}$ and $|\pm\rangle_{\mathrm{x}}$ is the symmetric, orthogonal matrix (the Hadamard matrix)

$$U = \tfrac{1}{\sqrt{2}}\big(\begin{smallmatrix}1 & 1\\1 & -1\end{smallmatrix}\big), \qquad \text{for which} \quad U|\pm\rangle_{\mathrm{z}} = |\pm\rangle_{\mathrm{x}}. \tag{23}$$

To go between the bases for $(\mathbb{C}^2)^{\otimes E}$ one uses the E-fold tensor product $U^{\otimes E}$. This change of basis maps $\sigma^{(1)}$ to $U\sigma^{(1)}U = \sigma^{(3)}$ and $\sigma^{(3)}$ to $U\sigma^{(3)}U = \sigma^{(1)}$, meaning that in the $|\cdot\rangle_{\mathrm{X}}$-basis, Z_s negates all the signs on the edges adjacent $s \in V$, while X_p is diagonal.

What we have described so far dovetails with the planar duality of Λ, as follows. Define the dual Λ^* to be the graph with vertex set $V^* = F(\Lambda)$, edge-set $E^* = E(\Lambda)$, and faces $F^* = V(\Lambda)$. One obtains Λ^* by placing a vertex $s^*(p)$ in the middle of each face p of Λ and drawing edges perpendicularly across those of Λ. In this way, any vertex s of the orignal lattice Λ lies in the middle of a unique face $p^*(s)$ of the dual Λ^*. We can then write, for $s \in V$ and $p \in F$,

$$\begin{aligned} U^{\otimes E} X_p U^{\otimes E} &= \prod_{e\sim p} U\sigma_e^{(1)}U = \prod_{e\sim s^*(p)} \sigma_e^{(3)} =: Z_{s^*}, \text{ and} \\ U^{\otimes E} Z_s U^{\otimes E} &= \prod_{e\sim s} U\sigma_e^{(3)}U = \prod_{e\sim p^*(s)} \sigma_e^{(1)} =: X_{p^*}. \end{aligned} \tag{24}$$

As noted above, the operators Z_{s^*} and X_{p^*} act on the $\sigma^{(1)}$-basis $|\omega\rangle_{\mathrm{X}}$ in the exact same way as Z_s and X_p act on the $\sigma^{(3)}$-basis $|\omega\rangle_{\mathrm{Z}}$. Namely, Z_{s^*} is diagonal while X_{p^*} negates the signs around p^*. In this sense, going between the lattice Λ and its dual Λ^* is equivalent to changing basis between $|\cdot\rangle_{\mathrm{Z}}$ and $|\cdot\rangle_{\mathrm{X}}$. Some particular instances of this are the identities:

$$\begin{aligned} {}_{\mathrm{Z}}\langle\tau|\exp\big(t\textstyle\sum_{p\in F} J_p^{\mathrm{x}} X_p\big)|\sigma\rangle_{\mathrm{Z}} &= {}_{\mathrm{X}}\langle\tau|\exp\big(t\textstyle\sum_{s^*\in V^*} J_{s^*}^{\mathrm{x}} Z_{s^*}\big)|\sigma\rangle_{\mathrm{X}}, \\ {}_{\mathrm{Z}}\langle\tau|\exp\big(t\textstyle\sum_{s\in V} J_s^{\mathrm{z}} Z_s\big)|\sigma\rangle_{\mathrm{Z}} &= {}_{\mathrm{X}}\langle\tau|\exp\big(t\textstyle\sum_{p^*\in F^*} J_{p^*}^{\mathrm{z}} X_{p^*}\big)|\sigma\rangle_{\mathrm{X}}, \end{aligned} \tag{25}$$

and for any $A \subseteq V(\Lambda)$, $B \subseteq F(\Lambda)$ with corresponding dual sets $A^* \subseteq F(\Lambda^*)$, $B^* \subseteq V(\Lambda^*)$,

$${}_{\mathrm{Z}}\langle\tau|\textstyle\prod_{s\in A} Z_s \prod_{p\in B} X_p|\sigma\rangle_{\mathrm{Z}} = {}_{\mathrm{X}}\langle\tau|\prod_{p^*\in A^*} X_{p^*} \prod_{s^*\in B^*} Z_{s^*}|\sigma\rangle_{\mathrm{X}}. \tag{26}$$

These identities may all be obtained by conjugating with $U^{\otimes E}$.

2.3. **Consequences.** We now provide some applications of the random-walk dynamics and of the duality. Write $\mathbf{P}_\sigma(\cdot)$ for the probability measure governing the process $\omega(t)$ of (18) started at $\omega(0) = \sigma$. We revert to the notation $|\omega\rangle$ without a subscript for the usual basis defined above (1) and referred to as $|\omega\rangle_{\mathrm{Z}}$ above.

LEMMA 2.1. *For any $\sigma, \tau \in \Omega$ and $t > 0$, we have the following matrix entries:*

$$\begin{aligned} \langle\tau|\exp\big(t\textstyle\sum_{p\in F} J_p^{\mathrm{x}} X_p\big)|\sigma\rangle &= e^{t\sum_{p\in F} J_p^{\mathrm{x}}} \mathbf{P}_\sigma(\omega(t) = \tau), \\ \langle\tau|\exp\big(t\textstyle\sum_{s\in V} J_s^{\mathrm{z}} Z_s\big)|\sigma\rangle &= \exp\big(t\textstyle\sum_{s\in V} J_s^{\mathrm{z}} \prod_{e\sim s}\sigma_e\big)\mathbb{1}_{\sigma=\tau}. \end{aligned} \tag{27}$$

Moreover,

$$e^{t\sum_{p\in F} J_p^{\mathrm{x}}} \mathbf{P}_\sigma(\omega(t) = \sigma) = \prod_{p\in F}\cosh(tJ_p^{\mathrm{x}}) + \prod_{p\in F}\sinh(tJ_p^{\mathrm{x}}). \tag{28}$$

In particular, $\exp\big(t\sum_{p\in F} J_p^{\mathrm{x}} X_p\big)$ *is constant on the diagonal.*

Proof. The first identity in (27) follows from the discussion at the end of Section 2.1, since $\exp\big(t\sum_{p\in F} J_p^{\mathrm{x}}(X_p - 1)\big)$ is the matrix of transition-probabilities of $\omega(t)$. The second identity in (27) follows from the fact that

$\exp\big(t\sum_{s\in V} J^{\mathrm{z}}_s Z_s\big)$ is diagonal in the given basis, with (σ,σ) entry precisely $\exp\big(t\sum_{s\in V} J^{\mathrm{z}}_s \prod_{e\sim s}\sigma_e\big)$.

For (28), we see from (18) that $\omega(t)=\omega(0)$ if and only if either (i) each plaquette is flipped an even number of times, or (ii) each plaquette is flipped an odd number of times. The number of times a given plaquette p is flipped by time t is a Poisson random variable with parameter $J^{\mathrm{x}}_p t$, and they are independent. This gives the stated expression. □

Lemma 2.1 allows for straightforward, explicit calculation of some thermodynamic quantities:

Proposition 2.2. *The partition function is given explicitly as*

(29)
$$\mathrm{tr}[e^{-H}] = 2^{|E|}\Big(\prod_{s\in V}\cosh(J^{\mathrm{z}}_s)+\prod_{s\in V}\sinh(J^{\mathrm{z}}_s)\Big)\Big(\prod_{p\in F}\cosh(J^{\mathrm{x}}_p)+\prod_{p\in F}\sinh(J^{\mathrm{x}}_p)\Big).$$

Moreover, for any $C\subset F$ and $D\subset V$ we have

(30)
$$\Big\langle\prod_{p\in C}X_p\Big\rangle = \frac{\prod_{p\in F\smallsetminus C}\cosh(J^{\mathrm{x}}_p)\prod_{p\in C}\sinh(J^{\mathrm{x}}_p)+\prod_{p\in F\smallsetminus C}\sinh(J^{\mathrm{x}}_p)\prod_{p\in C}\cosh(J^{\mathrm{x}}_p)}{\prod_{p\in F}\cosh(J^{\mathrm{x}}_p)+\prod_{p\in F}\sinh(J^{\mathrm{x}}_p)}$$

and

(31)
$$\Big\langle\prod_{s\in D}Z_s\Big\rangle = \frac{\prod_{s\in V\smallsetminus D}\cosh(J^{\mathrm{z}}_s)\prod_{s\in D}\sinh(J^{\mathrm{z}}_s)+\prod_{s\in V\smallsetminus D}\sinh(J^{\mathrm{z}}_s)\prod_{s\in D}\cosh(J^{\mathrm{z}}_s)}{\prod_{s\in V}\cosh(J^{\mathrm{z}}_s)+\prod_{s\in V}\sinh(J^{\mathrm{z}}_s)}.$$

Lastly, we have that

$$\Big\langle\prod_{p\in C}X_p\prod_{s\in D}Z_s\Big\rangle = \Big\langle\prod_{p\in C}X_p\Big\rangle\Big\langle\prod_{s\in D}Z_s\Big\rangle. \tag{32}$$

In particular, (30), (31), *and* (32) *are positive.*

Proof. For (29), note that

$$\begin{aligned}\mathrm{tr}[e^{-H}] &= \mathrm{tr}\big[\exp\big(\textstyle\sum_{s\in V}J^{\mathrm{z}}_s Z_s\big)\exp\big(\sum_{p\in F}J^{\mathrm{x}}_p X_p\big)\big]\\ &= 2^{-|E|}\mathrm{tr}\big[\exp\big(\textstyle\sum_{s\in V}J^{\mathrm{z}}_s Z_s\big)\big]\mathrm{tr}\big[\exp\big(\sum_{p\in F}J^{\mathrm{x}}_p X_p\big)\big].\end{aligned} \tag{33}$$

The first equality follows from the fact that the terms commute, the second equality from Lemma 2.1 and the elementary fact that $\mathrm{tr}[AB]=\mathrm{tr}[A]\mathrm{tr}[B]/r$ if A,B are $r\times r$ matrices such that A is diagonal and B is constant on the diagonal. Next, from Lemma 2.1 we have

$$\mathrm{tr}\big[\exp\big(\textstyle\sum_{p\in F}J^{\mathrm{x}}_p X_p\big)\big] = 2^{|E|}\Big(\displaystyle\prod_{p\in F}\cosh(J^{\mathrm{x}}_p)+\prod_{p\in F}\sinh(J^{\mathrm{x}}_p)\Big). \tag{34}$$

Also, from the duality-relations (25) we have

$$\mathrm{tr}\big[\exp\big(\textstyle\sum_{s\in F}J^{\mathrm{z}}_s Z_s\big)\big] = 2^{|E|}\Big(\displaystyle\prod_{s\in V}\cosh(J^{\mathrm{z}}_s)+\prod_{s\in V}\sinh(J^{\mathrm{z}}_s)\Big) \tag{35}$$

This follows by taking the trace in the $|\omega\rangle_{\mathrm{x}}$-basis rather than the $|\omega\rangle_{\mathrm{z}}$-basis.

For (30), we note that

(36)
$$\begin{aligned}\mathrm{tr}\big[\big(\textstyle\prod_{p\in C} X_p\big)e^{-H}\big] &= \mathrm{tr}\big[\exp\big(\textstyle\sum_{s\in V} J^{\mathrm{z}}_s Z_s\big)\big(\prod_{p\in C} X_p\big)\exp\big(\sum_{p\in F} J^{\mathrm{x}}_p X_p\big)\big]\\ &= \sum_{\sigma,\tau\in\Omega}\langle\sigma|\exp\big(\textstyle\sum_{s\in V} J^{\mathrm{z}}_s Z_s\big)|\tau\rangle\langle\tau|\big(\prod_{p\in C} X_p\big)\exp\big(\sum_{p\in F} J^{\mathrm{x}}_p X_p\big)|\sigma\rangle\\ &= \sum_{\sigma\in\Omega}\langle\sigma|\exp\big(\textstyle\sum_{s\in V} J^{\mathrm{z}}_s Z_s\big)|\sigma\rangle\langle\sigma|\big(\prod_{p\in C} X_p\big)\exp\big(\sum_{p\in F} J^{\mathrm{x}}_p X_p\big)|\sigma\rangle,\end{aligned}$$

where we used the fact that $\exp\big(\sum_{s\in V} J^{\mathrm{z}}_s Z_s\big)$ is diagonal. Next, using the notation of (17),

$$\langle\sigma|\big(\textstyle\prod_{p\in C} X_p\big)\exp\big(\sum_{p\in F} J^{\mathrm{x}}_p X_p\big)|\sigma\rangle = e^{\sum_{p\in F} J^{\mathrm{x}}_p}\mathbf{P}_\sigma\big(\omega(1) = \prod_{p\in C} x^p\cdot\sigma\big). \tag{37}$$

For the event in the probability to occur, either (i) all $p\in C$ are flipped an odd number of times and all $p\in F\smallsetminus C$ an even number of times, or (ii) vice versa. Computing this probability and using (35) and (29) gives (30). We have that (31) follows from the duality (25).

Lastly, we will prove (32). To begin, note that, because Z_p's are diagonal we have that

(38)
$$\begin{aligned}&\mathrm{tr}\big[\big(\textstyle\prod_{p\in C} X_p\big)\big(\prod_{s\in D} Z_s\big)e^{-H}\big]\\ &= \sum_{\sigma,\tau\in\Omega}\langle\sigma|\big(\textstyle\prod_{s\in D} Z_s\big)\exp\big(\sum_{s\in V} J^{\mathrm{z}}_s Z_s\big)|\tau\rangle\langle\tau|\big(\prod_{p\in C} X_p\big)\exp\big(\sum_{p\in F} J^{\mathrm{x}}_p X_p\big)|\sigma\rangle\\ &= \sum_{\sigma\in\Omega}\langle\sigma|\big(\textstyle\prod_{s\in D} Z_s\big)\exp\big(\sum_{s\in V} J^{\mathrm{z}}_s Z_s\big)|\sigma\rangle\langle\sigma|\big(\prod_{p\in C} X_p\big)\exp\big(\sum_{p\in F} J^{\mathrm{x}}_p X_p\big)|\sigma\rangle.\end{aligned}$$

As we saw in (37) and the remarks directly below it, we have

(39)
$$\begin{aligned}&\langle\sigma|\big(\textstyle\prod_{p\in C} X_p\big)\exp\big(\sum_{p\in C} J^{\mathrm{x}}_p X_p\big)|\sigma\rangle\\ &\qquad= \prod_{p\in F\smallsetminus C}\cosh(J^{\mathrm{x}}_p)\prod_{p\in C}\sinh(J^{\mathrm{x}}_p) + \prod_{p\in F\smallsetminus C}\sinh(J^{\mathrm{x}}_p)\prod_{p\in C}\cosh(J^{\mathrm{x}}_p)\end{aligned}$$

which, in particular, is independent of σ. The remaining sum over $\sigma\in\Omega$ is $\mathrm{tr}\big[\big(\prod_{s\in D} Z_s\big)e^{\sum_{s\in V} J^{\mathrm{z}}_s Z_s}\big]$ which by duality is

$$\begin{aligned}&\mathrm{tr}\big[\big(\textstyle\prod_{s\in D} Z_s\big)e^{\sum_{s\in V} J^{\mathrm{z}}_s Z_s}\big]\\ &=2^{|E|}\Big(\prod_{s\in V\smallsetminus D}\cosh(J^{\mathrm{z}}_s)\prod_{s\in D}\sinh(J^{\mathrm{z}}_s) + \prod_{s\in V\smallsetminus D}\sinh(J^{\mathrm{z}}_s)\prod_{s\in D}\cosh(J^{\mathrm{z}}_s)\Big).\end{aligned} \tag{40}$$

Now by combining (39), (40), and (29) we obtain (32). □

From (29) we immediately obtain an expression for the *free energy* of the system in the case when all $J^{\mathrm{x}}_p = \beta_{\mathrm{x}}$ and all $J^{\mathrm{z}}_s = \beta_{\mathrm{z}}$: using that $\frac{\sinh\beta}{\cosh\beta} < 1$ for all $\beta > 0$ we get

$$f(\beta_{\mathrm{x}},\beta_{\mathrm{z}}) = \lim_{|V|\to\infty}\frac{1}{|V|}\log\big(\mathrm{tr}\big[e^{-H}\big]\big) = \log(\cosh\beta_{\mathrm{x}}) + \log(\cosh\beta_{\mathrm{z}}) + 2\log 2. \tag{41}$$

In particular, $f(\beta_{\mathrm{x}},\beta_{\mathrm{z}})$ is analytic for $\beta_{\mathrm{x}},\beta_{\mathrm{z}} > 0$, indicating that the system does not undergo a phase transition, as pointed out earlier in e.g. [2].

Next we record a result complementary to Proposition 2.2 that will be useful in the proof of the correlation inequalities. The result is proven in [2], we give another argument using our dynamic picture.

PROPOSITION 2.3. *For $A, B \subset E$ define the operators*

$$X_A = \prod_{e\in A} \sigma_e^{(1)}, \qquad Z_B = \prod_{e\in B} \sigma_e^{(3)}. \tag{42}$$

Then $\langle X_A Z_B\rangle = 0$ unless X_A is a (possibly empty) product of plaquette operators X_p, and Z_B is a (possibly empty) product of star operators Z_s, i.e. unless $X_A = \prod_{p\in C} X_p$ for some $C \subseteq F$ and $Z_B = \prod_{s\in D} Z_s$ for some $D \subseteq V$.

Proof. The expansion (36) remains valid with $\prod_p X_p$ replaced by X_A, and in place of (37) we find that

$$\langle\sigma|X_A \exp\big(\textstyle\sum_{p\in F} J_p^{\mathrm{x}} X_p\big)|\sigma\rangle = e^{\sum_{p\in F} J_p^{\mathrm{x}}}\mathbf{P}_\sigma\big(\omega(1) = \prod_{e\in A} x^e\cdot\sigma\big), \tag{43}$$

where x^e is -1 in position e and $+1$ elsewhere. Note that this expression is independent of σ and hence can be pulled out of the sum over σ in (36). Now if X_A is not a product of plaquette operators then there is no realisation of the dynamics with the property that $\omega(1) = \prod_{e\in A} x^e\cdot\sigma$. Then $\mathbf{P}_\sigma\big(\omega(1) = \prod_{e\in A} x^e\cdot\sigma\big) = 0$, as claimed. Additionally, using the duality (25) it is easy to show that the correlation is 0 unless Z_B is a product of star operators. □

We now turn to the Ising model (11) and Proposition 1.2. Recall that Q is assumed to be an operator on $(\mathbb{C}^2)^{\otimes E}$ diagonal in the $|\cdot\rangle = |\cdot\rangle_{\mathrm{z}}$-basis with $Q|\sigma\rangle = q(\sigma)|\sigma\rangle$.

Proof of Proposition 1.2. We use Lemma 2.1 and orthonormality to expand:

$$\begin{aligned}\mathrm{tr}[Qe^{-H}] &= \sum_{\sigma,\tau,\varphi\in\Omega} \langle\sigma|Q|\tau\rangle\langle\tau|\exp\big(\textstyle\sum_{s\in V} J_s^{\mathrm{z}} Z_s\big)|\varphi\rangle\langle\varphi|\exp\big(\textstyle\sum_{p\in F} J_p^{\mathrm{x}} X_p\big)|\sigma\rangle\\ &= \sum_{\sigma,\tau,\varphi\in\Omega} q(\tau)\delta_{\sigma,\tau}\exp\big(\textstyle\sum_{s\in V} J_s^{\mathrm{z}}\prod_{e\sim s}\varphi_e\big)\delta_{\tau,\varphi} e^{\sum_{p\in F} J_p^{\mathrm{x}}}\mathbf{P}_\sigma(\omega(\beta) = \varphi)\\ &= \sum_{\sigma\in\Omega} q(\sigma)\exp\big(\textstyle\sum_{s\in V} J_s^{\mathrm{x}}\prod_{e\sim s}\sigma_e\big) e^{\sum_{p\in F} J_p^{\mathrm{x}}}\mathbf{P}_\sigma(\omega(\beta) = \sigma).\end{aligned} \tag{44}$$

In particular,

$$\mathrm{tr}[e^{-H}] = \sum_{\sigma\in\Omega}\exp\big(\textstyle\sum_{s\in V} J_s^{\mathrm{z}}\prod_{e\sim s}\sigma_e\big) e^{\sum_{p\in F} J_p^{\mathrm{x}}}\mathbf{P}_\sigma(\omega(\beta) = \sigma). \tag{45}$$

By Lemma 2.1, $e^{\sum_{p\in F} J_p^{\mathrm{x}}}\mathbf{P}_\sigma(\omega(\beta) = \sigma)$ does not depend on σ, thus these factors cancel in $\langle Q\rangle = \mathrm{tr}[Qe^{-H}]/\mathrm{tr}[e^{-H}]$, giving the result. □

REMARK 2.4.

(1) *Recall that the probability measure $\mathbb{P}_{\mathbf{J}}(\cdot)$ is invariant under reversing all the spin-values around any fixed plaquette p. This immediately gives that $\mathbb{E}_{\mathbf{J}}[\prod_{e\in A}\sigma_e] = 0$ if there is any plaquette p such that A contains an odd number of the edges surrounding p, which is in line with Proposition 2.3.*

(2) *Clearly the analog of Proposition 1.2 is true also for operators Q diagonal in the $|\omega\rangle_{\mathrm{x}}$-basis, by the duality (25). Indeed, the same argument carried out in that basis gives $\langle Q\rangle = \mathbb{E}^*_{\mathbf{J}^{\mathrm{x}}}[q(\sigma)]$, where $\mathbb{E}^*_{\mathbf{J}}[\cdot]$ denotes expectation under the measure*

$$\mathbb{P}^*_{\mathbf{J}}(\sigma) = \frac{\exp\big(\sum_{p\in F} J_p\prod_{e\sim p}\sigma_e\big)}{\sum_{\omega\in\Omega}\exp\big(\sum_{p\in F} J_p\prod_{e\sim s}\omega_e\big)}, \qquad \sigma\in\Omega.$$

In this sense one can see the toric code model as two coupled classical Ising models.

3. Correlation inequalities

The proof of Theorem 1.1 will proceed by first proving a similar statement for the Ising model (11). To state the latter, we introduce the following notation. For $s \in V$ consider the 4 edges e_1, e_2, e_3 and e_4 adjacent to s, to be definite ordered as in Figure 2. For $\sigma \in \Omega$, write $\sigma_i(s) = \sigma_{e_i}$. We define the following quantities similar to (5) and (6): for $\varepsilon \in \{-1,+1\}^4$ satisfying $\prod_{i=1}^4 \varepsilon_i = +1$, let

$$I_s^\varepsilon(\sigma) = \tfrac{1}{16}\prod_{i=1}^{4}\big(1+\varepsilon_i\sigma_i(s)\big), \qquad \bar I_s^\varepsilon(\sigma) = \tfrac{1}{16}\prod_{i=1}^{4}\big(1-\varepsilon_i\sigma_i(s)\big), \tag{46}$$

and for $C \subseteq V$,

$$\mathcal{I}_C^\varepsilon(\sigma) = \prod_{s\in C}\big(I_s^\varepsilon(\sigma)+\bar I_s^\varepsilon(\sigma)\big). \tag{47}$$

Then $\mathcal{I}_C^\varepsilon(\sigma)$ is the indicator of the event that, for each $s \in C$, the values $\sigma_i(s)$ for $i = 1,\dots,4$ either all agree with ε_i, or are all the opposite. Similarly to the eight-vertex model and toric code, for $\delta \in \{-1,+1\}^4$ with $\delta \neq \pm\varepsilon$ we also define

$$\mathcal{I}_C^{\varepsilon,\delta}(\sigma) = \prod_{s\in C}\big(I_s^\varepsilon(\sigma)+\bar I_s^\varepsilon(\sigma)+I_s^\delta(\sigma)+\bar I_s^\delta(\sigma)\big). \tag{48}$$

Lemma 3.1. *Let ε,δ, with $\delta \neq \pm\varepsilon$, be as above and let $C_1, C_2 \subseteq V$ be sets of vertices such that*

- *either $C_1 \cup C_2$ is contractible,*
- *or any non-contractible loop in $C_1 \cup C_2$ has even length.*

Then for any $\mathbf{J} = (J_s : s \in V)$ such that all $J_s \geq 0$,

$$\mathbb{E}_{\mathbf{J}}\big[\mathcal{I}_{C_1}^\varepsilon \mathcal{I}_{C_2}^\varepsilon\big] \geq \mathbb{E}_{\mathbf{J}}\big[\mathcal{I}_{C_1}^\varepsilon\big]\mathbb{E}_{\mathbf{J}}\big[\mathcal{I}_{C_2}^\varepsilon\big]. \tag{49}$$

and

$$\mathbb{E}_{\mathbf{J}}\big[\mathcal{I}_{C_1}^{\varepsilon,\delta} \mathcal{I}_{C_2}^{\varepsilon,\delta}\big] \geq \mathbb{E}_{\mathbf{J}}\big[\mathcal{I}_{C_1}^{\varepsilon,\delta}\big]\mathbb{E}_{\mathbf{J}}\big[\mathcal{I}_{C_2}^{\varepsilon,\delta}\big]. \tag{50}$$

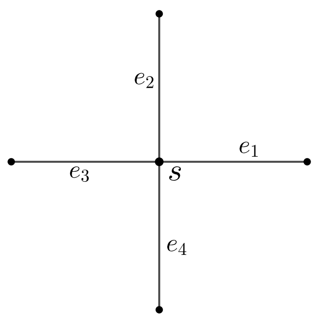

Figure 2. A vertex s with its four incident edges labelled in a counter-clockwise fashion.

The method of proof will be to expand the products defining $\mathcal{I}_C^\varepsilon$ and then apply the GKS-inequalities (13) to the terms of the expansion. The subtlety is that the terms come with signs; we show, using Proposition 2.3, that all terms with negative sign actually vanish in expectation.

Proof. We will prove the first inequality, the proof of the second inequality is almost identical. Since $\prod_{i=1}^4 \varepsilon_i = +1$, we have that $I_s^\varepsilon + \bar{I}_s^\varepsilon = \frac{1}{8}(1 + A_s + R_s)$ where $A_s = \prod_{i=1}^4 \sigma_i(s)$ and (with the argument s suppressed for readability)

$$R_s = \sum_{1 \le i < j \le 4} \varepsilon_i \varepsilon_j \sigma_i \sigma_j. \tag{51}$$

Let C be any of C_1, C_2 or $C_1 \cup C_2$. By expanding the product over $s \in C$ we get that

$$\mathcal{I}_C^\varepsilon = \prod_{s \in C} \tfrac{1}{8}(1 + A_s + R_s) = \frac{1}{8^{|C|}} \sum_{D \subseteq C} \sum_{D_1 \subseteq D} \prod_{s \in D_1} R_s \prod_{t \in D \setminus D_1} A_t. \tag{52}$$

In this expression we expand the product over $s \in D_1$. To write the expansion, we use the following notation. We let $\Sigma(D_1)$ denote the set of sequences $(\{i(s), j(s)\})_{s \in D_1}$ of two-element subsets of $\{1, 2, 3, 4\}$, indexed by $s \in D_1$. We may represent such a subset $\{i(s), j(s)\}$ pictorially as an element of the set $\{\llcorner, -, \lrcorner, \ulcorner, |, \urcorner\}$, indicating the orientations of the two edges selected. Then,

$$\mathcal{I}_C^\varepsilon = \frac{1}{8^{|C|}} \sum_{\substack{D_1, D_2 \subseteq C \\ D_1 \cap D_2 = \varnothing}} \sum_{\Sigma(D_1)} \prod_{s \in D_1} \varepsilon_{i(s)} \varepsilon_{j(s)} \sigma_{i(s)} \sigma_{j(s)} \prod_{t \in D_2} A_t. \tag{53}$$

Now consider an arbitrary term T in the latter expansion, corresponding to a choice of D_1, D_2 and sequence in $\Sigma(D_1)$. We write this term as

$$T = \prod_{s \in D_1} \varepsilon_{i(s)} \varepsilon_{j(s)} \prod_{s \in D_1} \sigma_{i(s)} \sigma_{j(s)} \prod_{t \in D_2} A_t. \tag{54}$$

The following is the key claim: if T can be written as a 'product of stars', then it comes with a positive sign. That is, if there is a set $D_3 \subseteq V$ such that

$$\prod_{s \in D_1} \sigma_{i(s)} \sigma_{j(s)} \prod_{t \in D_2} A_t = \prod_{u \in D_3} A_u, \tag{55}$$

then

$$\prod_{s \in D_1} \varepsilon_{i(s)} \varepsilon_{j(s)} = +1. \tag{56}$$

Before proving the claim (i.e. that (55) implies (56)) we show how to deduce the result. We may write

$$\mathcal{I}_{C_1}^\varepsilon \mathcal{I}_{C_2}^\varepsilon = \frac{1}{8^{|C_1| + |C_2|}} \sum_{T_1, T_2} T_1 T_2, \tag{57}$$

where T_1 and T_2 are terms of the form (54) for $C = C_1$ and for $C = C_2$, respectively, and $T = T_1 T_2$ is of the form (54) for $C = C_1 \cup C_2$. We will use the GKS-inequality (13) to see that all terms satisfy $\mathbb{E}_{\mathbf{J}}[T_1 T_2] \ge \mathbb{E}_{\mathbf{J}}[T_1] \mathbb{E}_{\mathbf{J}}[T_2]$. By retracing the steps of the expansion, this implies the result. Now, if $T = T_1 T_2$ does *not* satsify (55) then at least one of T_1 and T_2 does not satisfy it either. By Propositions 1.2 and 2.3, then $\mathbb{E}_{\mathbf{J}}[T_1 T_2] = \mathbb{E}_{\mathbf{J}}[T_1] \mathbb{E}_{\mathbf{J}}[T_2] = 0$. If $T = T_1 T_2$

does satisfy (55) but one of T_1 and T_2 does not, then by Propositions 2.2 and 2.3, $\mathbb{E}_{\mathbf{J}}[T_1T_2] \geq 0 = \mathbb{E}_{\mathbf{J}}[T_1]\mathbb{E}_{\mathbf{J}}[T_2]$. Finally, if all three of $T_1, T_2, T = T_1T_2$ satisfy (55) then the desired inequality $\mathbb{E}_{\mathbf{J}}[T_1T_2] \geq \mathbb{E}_{\mathbf{J}}[T_1]\mathbb{E}_{\mathbf{J}}[T_2]$ follows from the GKS-inequality (13) and the fact that the terms have positive coefficients, thanks to the claim.

We now prove the claim that (55) implies (56). If (55) holds, then since the A_t satisfy $A_t^2 = 1$, we get

$$\prod_{s\in D_1} \sigma_{i(s)}\sigma_{j(s)} \prod_{t\in D_2\triangle D_3} A_t = 1. \tag{58}$$

Recall that we represent each factor of the product over $s \in D_1$ as an element of the set of shapes $\{\llcorner, -, \lrcorner, \ulcorner, |, \urcorner\}$, indicating the two selected edges. Similarly, each factor of the product over t may be represented as a +, indicating that all four edges are selected. We think of the latter as consisting of two superimposed shapes, such as $\llcorner$ and $\urcorner$ (the precise choice does not matter).

We can assume that $D_1 \cap (D_2 \triangle D_3) = \varnothing$. Then we conclude from (58) that each edge $e \in E$ appears in the product either exactly twice, or not at all. Consider now the subset of edges which appear exactly twice (counting each such edge once). Due to the possible choices of 'shapes' at each vertex, this subset forms an even subgraph (each vertex has even degree) and thus decomposes as a collection of closed loops. We claim that each such loop has the following property: the number of vertices where it exhibits a shape from $\{\llcorner, \urcorner\}$ is even, similarly the number of vertices where it exhibits a shape from $\{\lrcorner, \ulcorner\}$ is even, and finally the number of vertices where it exhibits a shape from $\{-, |\}$ is even. The latter claim can be seen by induction, as follows. We may take the loops to be non-crossing. For any non-crossing contractible loop enclosing at least two squares, we can decrease its enclosed area one square by removing a corner. As we do so, the number of shapes from each of the four sets changes by an even amount. Eventually the loop reduces to a single square, for which the claim holds by inspection. If the loop is not contractible, then a similar reduction can be used to reduce it to a 'straight' loop, which by our assumption on $C_1 \cup C_2$ has even length, meaning again that the claim holds by inspection.

Finally, the sign on the left-hand-side of (56) can be written as

$$\prod_{s\in D_1} \varepsilon_{i(s)}\varepsilon_{j(s)} = \prod_{s\in D_1} \varepsilon_{i(s)}\varepsilon_{j(s)} \prod_{t\in D_2\triangle D_3} \varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4. \tag{59}$$

Here we used that $\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4 = +1$. For any choice of an even number of shapes from each of the sets $\{\llcorner, \urcorner\}$, $\{\lrcorner, \ulcorner\}$ and $\{-, |\}$, the product of the corresponding ε's in (59) is $= +1$, again using $\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4 = +1$. The result follows. □

Proof of Theorem 1.1. For the uniform 8-vertex model we deduce the result by fixing a reference configuration $\rho \in \Delta_{8\text{vx}}$ and letting all $J_s^{\text{z}} \to \infty$. Since the limit of the Ising-measure (11) is uniform on $\sigma \in \Omega_{8\text{vx}}$, the distribution of $\sigma \cdot \rho$ is uniform on $\Delta_{8\text{vx}}$.

For the first claim for the toric code model, note that the operators $Q_C^{(3)}(\varepsilon)$ and $Q_C^{(3)}(\varepsilon, \delta)$ are diagonal in the $|\omega\rangle = |\omega\rangle_{\text{z}}$-basis and satisfy $Q_C^{(3)}(\varepsilon)|\omega\rangle = \mathcal{I}_C^{\varepsilon}(\omega)|\omega\rangle$ and $Q_C^{(3)}(\varepsilon,\delta)|\omega\rangle = \mathcal{I}_C^{\varepsilon,\delta}(\omega)|\omega\rangle$. Then, using Proposition 1.2 for the first identity, Lemma 3.1 for the inequality, and Proposition 1.2 again for the last identity,

(60)
$$\langle Q^{(3)}_{C_1}(\varepsilon)Q^{(3)}_{C_2}(\varepsilon)\rangle = \mathbb{E}_{\mathbf{J}^z}\big[\mathcal{I}^\varepsilon_{C_1}\mathcal{I}^\varepsilon_{C_2}\big] \geq \mathbb{E}_{\mathbf{J}^z}\big[\mathcal{I}^\varepsilon_{C_1}\big]\mathbb{E}_{\mathbf{J}^z}\big[\mathcal{I}^\varepsilon_{C_2}\big] = \langle Q^{(3)}_{C_1}(\varepsilon)\rangle\langle Q^{(3)}_{C_2}(\varepsilon)\rangle,$$

as claimed. The argument for $Q^{(3)}_C(\varepsilon,\delta)$ is similar.

The second claim for the toric code model uses similar arguments, but in place of the GKS-inequality uses a correlation inequality for quantum XY-models in [6]. The main step is to check that the argument of Lemma 3.1 goes through. Indeed, by expanding as in (51)–(53) we find that

$$Q^{(3)}_C(\varepsilon)Q^{(1)}_D(\delta) = \frac{1}{8^{|C|+|D|}}\sum_{T^3,T^1} T^3T^1 \tag{61}$$

where T^3 are operators of the form

$$T^3 = \sum_{s\in C'}\varepsilon_{i(s)}\varepsilon_{j(s)}\prod_{s\in C'}\sigma^{(3)}_{i(s)}\sigma^{(3)}_{j(s)}\prod_{t\in\tilde{C}\smallsetminus C'} Z_t \tag{62}$$

with $i(s), j(s)$ two elements of $\{1,2,3,4\}$ indexing the edges incident to s and $\tilde{C}, C' \subset C$. The operators T^1 are of the form

$$T^1 = \sum_{p\in D'}\delta_{i(s)}\delta_{j(s)}\prod_{p\in D'}\sigma^{(1)}_{i(p)}\sigma^{(1)}_{j(p)}\prod_{q\in\tilde{D}\smallsetminus D'} X_q \tag{63}$$

with $i(s), j(s)$ two elements of $\{1,2,3,4\}$ indexing the edges around p and $\tilde{D}, D' \subset D$.

Now by Proposition 2.3, $\langle T^3T^1\rangle = 0$ unless T^3 is a product of star operators and T^1 is a product of plaquette operators, in which case T^3 and T^1 both come with a positive sign (i.e. $\prod_{s\in C'}\varepsilon_{i(s)}\varepsilon_{j(s)} = \prod_{p\in D'}\delta_{i(s)}\delta_{j(s)} = +1$) as in (56). Now by [6, Theorem 1] we have that $\langle T^3T^1\rangle \leq \langle T^3\rangle\langle T^1\rangle$ and by tracing back the calculation leading to (61) we obtain the result. □

REMARK 3.2. *It is possible, as in* [3]*, to construct a quantum system generalising the Kitaev model, whose ground state configurations correspond to eight-vertex configurations with the general weights*

$$\mu_{a,b,c,d}(\omega) \propto a^{\#\mathtt{I}+\#\mathtt{II}}b^{\#\mathtt{III}+\#\mathtt{IV}}c^{\#\mathtt{V}+\#\mathtt{VI}}d^{\#\mathtt{VII}+\#\mathtt{VIII}} \tag{64}$$

depending on the numbers $\#\mathtt{I},\dots,\#\mathtt{VIII}$ of vertices of the different types. However, the operators that must be added to the hamiltonian to achieve this do not satisfy the required non-negativity. This suggests that the inequalities in Theorem 1.1 may not hold for non-uniform weights.

Using almost the same argument as for Lemma 3.1 we can also prove the following:

PROPOSITION 3.3. *For any $C, D \subseteq V$ and any $\varepsilon \in \{-1,+1\}^4$ satisfying $\prod_{i=1}^4 \varepsilon_i = +1$, as above,*

$$\langle Q^{(3)}_C(\varepsilon)\textstyle\prod_{u\in D} Z_u\rangle \geq \langle Q^{(3)}_C(\varepsilon)\rangle\langle\prod_{u\in D} Z_u\rangle. \tag{65}$$

In particular, for any $s \in V$,

$$\frac{\partial}{\partial J_s}\langle Q^{(3)}_C(\varepsilon)\rangle = \langle Q^{(3)}_C(\varepsilon)Z_s\rangle - \langle Q^{(3)}_C(\varepsilon)\rangle\langle Z_s\rangle \geq 0. \tag{66}$$

Similarly, for $C \subseteq F$ and $D \subseteq V$,

$$\langle Q_C^{(1)}(\varepsilon) \textstyle\prod_{u\in D} Z_u\rangle \leq \langle Q_C^{(1)}(\varepsilon)\rangle\langle \textstyle\prod_{u\in D} Z_u\rangle, \tag{67}$$

and hence

$$\tfrac{\partial}{\partial J_s}\langle Q_C^{(1)}(\varepsilon)\rangle = \langle Q_C^{(1)}(\varepsilon) Z_s\rangle - \langle Q_C^{(1)}(\varepsilon)\rangle\langle Z_s\rangle \leq 0. \tag{68}$$

Proof. For (65), we have

$$\langle Q_C^{(3)}(\varepsilon) \textstyle\prod_{u\in D} Z_u\rangle = \mathbb{E}_{\mathbf{J}}[\mathcal{I}_C^{\varepsilon}(\sigma) \textstyle\prod_{u\in D} A_u(\sigma)], \tag{69}$$

where $A_u(\sigma) = \prod_{i=1}^4 \sigma_i(s)$ as before. The same working as for (53) combined with the fact that $A_u(\sigma)^2 = +1$ gives

$$\mathcal{I}_C^{\varepsilon} \prod_{u\in D} A_u(\sigma) = \frac{1}{8^{|C|}} \sum_{\substack{D_1, D_2 \subseteq C \\ D_1\cap D_2 = \varnothing}} \sum_{\Sigma(D_1)} \prod_{s\in D_1} \varepsilon_{i(s)}\varepsilon_{j(s)}\sigma_{i(s)}\sigma_{j(s)} \prod_{t\in D_2 \triangle D} A_t. \tag{70}$$

From there the same argument as in Lemma 3.1 applies.

Similarly, for (67) we may write

$$Q_C^{(1)}(\varepsilon) \prod_{u\in D} Z_u = \frac{1}{8^{|C|}} \sum_{\substack{\tilde{C}, C' \subseteq C \\ \tilde{C}\cap C' = \varnothing}} \sum_{\Sigma(\tilde{C})} \prod_{p\in\tilde{C}} \varepsilon_{i(p)}\varepsilon_{j(p)}\sigma^{(1)}_{i(p)}\sigma^{(1)}_{j(p)} \prod_{p\in C'} A_p^{(1)} \prod_{u\in D} A_p^{(3)} \tag{71}$$

and as before, only terms which are products of stars and plaquettes survive the expectation and they come with positive sign. □

A standard consequence of (66) is the existence of infinite-volume limits of correlation functions of the form $\langle Q_C^{(3)}(\varepsilon)\rangle$; see e.g. [6, Proposition 4] for arguments of this nature.

4. Further results for the uniform 8-vertex model

In this section we discuss further properties of the uniform eight-vertex model that are natural to consider from the viewpoint of 'plaquette-flipping'. While not directly related to the correlation inequalities that were the main focus of the previous section, they serve as further illustration of the usefulness of the methods used in the proofs of the correlation inequalitites.

4.1. **Communicating classes.** We start by considering the communicating classes of the dynamics (21) in the case when all $J_p > 0$. Recall that Λ is an $m \times n$ torus.

Proposition 4.1. *We have that $|\Delta_{8\mathrm{vx}}| = 2^{|V|+1}$, and there are four communicating classes for the dynamics (21), each of size $2^{|V|-1}$. We may move between them by reversing arrows along a non-contractible path in Λ.*

Proof. We start by showing that $|\Delta_{8\mathrm{vx}}| = 2^{|V|+1}$. First, clearly $|\Delta| = 2^{|E|} = 2^{2|V|}$. By fixing a reference-configuration ρ as in (21) we can encode each element of Δ using an element $v \in (\mathbb{F}_2)^E$ where $\mathbb{F}_2 = \{0, 1\}$ is the two-element field. Then, for each $s \in V$, the constraint that s has an even number of incoming arrows becomes a linear constraint over $\mathbb{F}_2$, namely $g^s \cdot v = 0$ where (similarly to (15))

$$g_e^s = \begin{cases} 1, & \text{if } e \sim s, \\ 0, & \text{otherwise,} \end{cases} \tag{72}$$

and $\cdot$ is the scalar product. These constraints are *not* linearly independent, since $(\sum_{s\in V} g^s)_e = 0$ for each $e \in E$. However, this is the only linear relation: any other linear relation would have to be of the form $\sum_{s\in A} g^s \equiv 0$ for some proper subset A of V. But then there must be some edge $e \in E$ with precisely one end point in A and then $(\sum_{s\in A} g^s)_e = 1 \neq 0$. From this and the rank-nullity theorem we get that $|\Delta_{8\mathrm{vx}}| = 2^{2|V|-(|V|-1)} = 2^{|V|+1}$, as claimed.

Now it is simple to see that there are four communicating classes for the dynamics. There are $|V|$ plaquettes in F and hence $2^{|V|}$ possible sums $\sum_{p\in A} g^p$ for $A \subseteq V$. However, we have that $\sum_{p\in A} g^p = \sum_{p\in V\setminus A} g^p$ (since $\sum_{p\in V} g^p \equiv 0$). Hence, starting from any reference configuration $\rho \in \Delta_{8\mathrm{vx}}$, there are $2^{|V|-1}$ configurations reachable by flipping plaquettes. This shows that there are four communicating classes.

Lastly, let $P \subset E$ be a non-contractible path of edges in Λ, it is clear that such a path is not the boundary of a collection of plaquettes $A \subset F$. This means that the configuration obtained by reversing arrows on P is not reachable from the reference configuration via the dynamics. There are two homotopy classes of non-contractible paths on the torus and reversing arrows along a path of edges in one or two of the classes moves us between the four communicating classes. □

Remark 4.2.

(1) *Because our dynamics for the uniform eight-vertex model correspond to a random walk on the hypercube* $(\mathbb{F}_2^+)^F$*, we can import results about mixing times from the literature. For example, the mixing time for the dynamics is* $\log(|V|)$ *(recall that the Poisson process of steps has rate* $|F|$ *rather than the usual rate 1). The interested reader can consult* [25] *and references therein for detailed statements.*

(2) *The limiting distribution of random walk on the hypercube* $(\mathbb{F}_2^+)^F$ *is uniform. This leads to a method for sampling from the uniform eight-vertex distribution* $\mu(\cdot)$*: first sample uniformly a representative* $\rho \in \Delta_{8\mathrm{vx}}$ *from each of the four communicating classes to act as reference-configuration, then toss independent coins for each of the plaquettes for whether to 'flip' the plaquette or not.*

(3) *If* $s_1, s_2 \in V$ *are vertices not adjacent to any common plaquette* p*, then by the previous item, the vertex-types at* s_1 *and* s_2 *can be written as functions of independent random variables (states of the plaquettes surrounding them). Thus their states are independent.*

4.2. **Emptiness formation probability.** In this subsection we allow to view Λ not only as a torus, but also as a subset of $\mathbb{Z}^2$ with a boundary. In that case we replace each of the edges of the form $\{(m,y),(1,y)\}$ with two boundary-edges $\{(m,y),(m+1,y)\}$ and $\{(0,y),(1,y)\}$, and similarly replace each $\{(x,n),(x,1)\}$ with two boundary-edges $\{(x,n),(x,n+1)\}$ and $\{(x,0),(x,1)\}$.

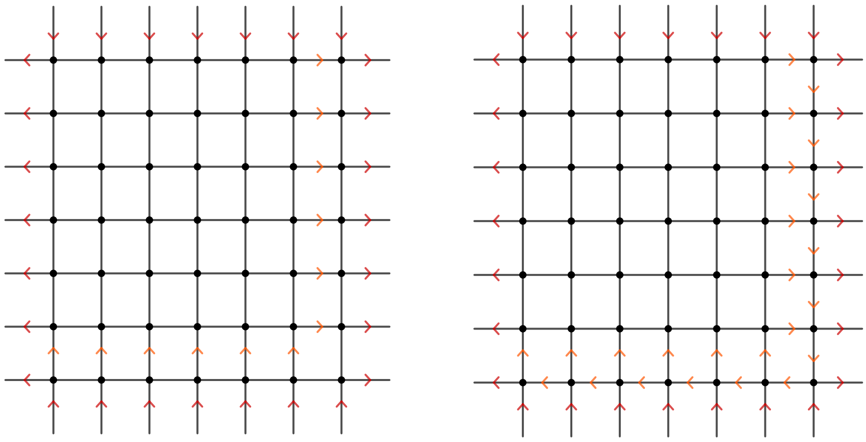

FIGURE 3. Left: extending $\Lambda_{m-1,n-1}$ to $\Lambda_{m,n}$ with both having domain wall boundary conditions. Right: $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m-1,n-1}) \neq \varnothing$ implies $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n}) \neq \varnothing$.

First we consider the *domain wall boundary condition*. This means that the top and bottom boundary edges (of the form $\{(x,0),(x,1)\}$ or $\{(x,m),(x,m+1)\}$) receive the fixed orientation pointing *in* towards Λ, and that the left and right boundary-edges (of the form $\{(0,y),(1,y)\}$ or $\{(m,y),(m+1,y)\}$) are fixed to point *out* of Λ. See Figure 3 for an illustration.

We emphasise m and n by using the notation $\Lambda_{m,n}$. We denote by $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n})$ the set of eight-vertex configurations on $\Lambda_{m,n}$ with domain wall boundary conditions as described above.

PROPOSITION 4.3. *We have that*

$$|\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n})| = \begin{cases} 0, & \text{if } m+n \text{ is odd,} \\ 2^{(m-1)(n-1)}, & \text{if } m+n \text{ is even.} \end{cases} \tag{73}$$

When non-empty, $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n})$ consists of a single communicating class for the dynamics (21).

Proof. We begin by showing that $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n}) \neq \varnothing$ if and only if $m+n$ is even. This is done in three steps.

Step 1: if $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m-1,n-1}) \neq \varnothing$ then $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n}) \neq \varnothing$. Indeed, Figure 3 illustrates how an element of $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m-1,n-1})$ can be 'extended' to an element of $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n})$.

Step 2: conversely, if $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n}) \neq \varnothing$ then $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m-1,n-1}) \neq \varnothing$. Indeed, fix an element $\delta \in \Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n})$ and consider the bottom row of vertical edges and the rightmost column of horizontal edges in $\Lambda_{m,n}$ (see Figure 4). If all those vertical edges are oriented *in*, and all the horizontal ones are oriented *out*, then the restriction of δ to $\Lambda_{m-1,n-1}$ is an element of $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m-1,n-1})$. Otherwise, the edges pointing the 'wrong way' (i.e. out for vertical edges on the bottom, in for horizontal edges on the right side) will be called *faults*. It is not hard to check that the number of faults must be even. Then, pair up the successive faults and mark the plaquettes between the pairs, as in

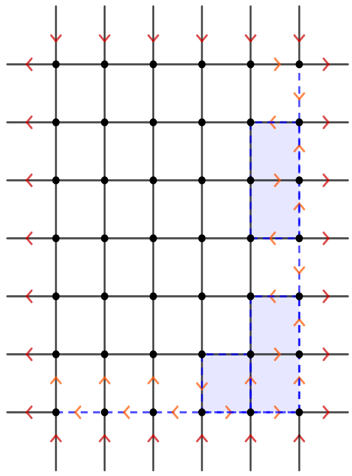

FIGURE 4. An arbitrary element of $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n})$ can be mapped to one whose restriction to $\Lambda_{m-1,n-1}$ also has domain wall boundary condition. The highlighted plaquettes separate pairs of *faults*, i.e. arrows pointing the opposite way to the boundary condition. Flipping the highlighted plaquettes gives the desired element of $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m-1,n-1})$.

Figure 4. Flipping all the marked plaquettes maps δ to a configuration whose restriction to $\Lambda_{m-1,n-1}$ is an element of $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m-1,n-1})$.

Step 3: from the previous two steps we see that $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n}) \neq \varnothing$ if and only if $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m-1,n-1}) \neq \varnothing$. If $m \geq n$ this holds if and only if $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m-n+1,1}) \neq \varnothing$; if $n \geq m$ it holds if and only if $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{1,n-m+1}) \neq \varnothing$. It is straightforward to check that the latter sets are non-empty if and only if $m-n$ is even, which is equivalent to $m+n$ being even.

To determine the number of configurations when $m+n$ is even, we use a simple counting argument. First, let us count the number of choices at each vertex if we start at the top left corner $(1,n)$ and proceed left to right and then top to bottom. As we proceed, each time we 'arrive' at a new vertex *except* the rightmost column or bottom row, exactly two incident arrows are already fixed, leaving us with exactly two choices. In the rightmost column and bottom row, *three* incident arrows are fixed when we arrive, leaving us at most one (possibly no) choice per vertex. This gives that

$$|\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n})| \leq 2^{(m-1)(n-1)}. \tag{74}$$

On the other hand, if $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n}) \neq \varnothing$, fix some $\delta \in \Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n})$. An argument similar to that of Proposition 4.1 shows that, in this case, the number of configurations reachable by flipping plaquettes equals the total number of configurations. Since there are $(m-1)(n-1)$ plaquettes, this gives that

$$|\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n})| \geq 2^{(m-1)(n-1)}. \tag{75}$$

Combining (74) and (75) shows that, when non-empty, $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n})$ has size $2^{(m-1)(n-1)}$. The claim about irreducibility follows from the proof of (75). □

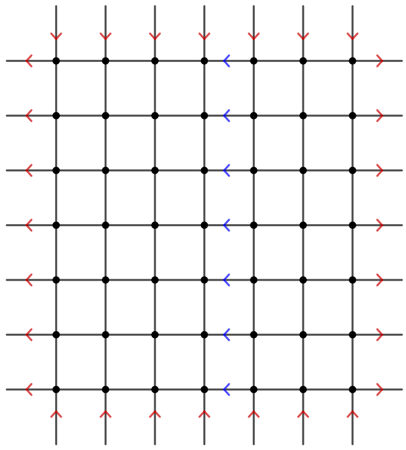

FIGURE 5. Example of an empty column in a box with domain wall boundary condition. Here $m = n = 7$ and $r = 4$.

We now turn our attention to the so-called *emptiness formation probability* of the model. This is the probability that a fixed column of horizontal edges has all of its arrows pointing to the left (say). In this case we say that the column is *empty*. Denote by EF_r the event that the column of horizontal edges whose left end-points have first coordinate r is empty, see Figure 5. In what follows, $\mu^{\mathrm{DW}}(\cdot)$ denotes the uniform probability measure on $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n})$. We calculate the probability of EF_r in the case of the torus and the case of domain wall boundary conditions. We begin with the domain wall case.

PROPOSITION 4.4. *Assume that $m + n$ is even so that $\Delta^{\mathrm{DW}}_{8\mathrm{vx}}(\Lambda_{m,n}) \neq \varnothing$. Then $\mu^{\mathrm{DW}}(\mathrm{EF}_r) = 2^{-(m-1)} \mathbb{1}_{\{r \in 2\mathbb{N}\}}$.*

Proof. The proof is similar to that of Proposition 4.3. First we show that $\mathrm{EF}_r \neq \varnothing$ if and only if r is even. Indeed, by Proposition 4.3, the part of $\Lambda_{m,n}$ to the right of the empty column can be 'filled in' if and only if $n - r + m$ is even. Since $m + n$ is even, this is equivalent to r being even. To fill in the part to the left of the empty column, we can make the top row alternate between types `VI` and `VII` (this is also possible if and only if r is even) and the rest all type `IV`. (Recall Figure 1.1 for the vertex types.) This shows that $\mathrm{EF}_r \neq \varnothing$ if and only if r is even.

Assuming then that r is even, let us count the number of choices at each vertex in the two parts, going from left to right and then top to bottom in each. We see that $(r-1)(m-1) + (n-r-1)(m-1) = (m-1)(n-2)$ vertices have two choices of vertex type and the remaining vertices have at most one choice. From this and Proposition 4.3 we see that

$$\mu^{\mathrm{DW}}(\mathrm{EF}_r) \le 2^{(m-1)(n-2)} / 2^{(m-1)(n-1)} = 2^{-(m-1)}. \tag{76}$$

On the other hand, consider the $(m-1)(n-1) - (m-1)$ plaquettes that may be reversed without affecting the empty column. Similarly to the proof of

Proposition 4.3 we see that each distinct choice gives a different configuration. From this we have that

$$\mu^{\mathrm{DW}}(\mathrm{EF}_r) \geq 2^{(m-1)(n-1)-(m-1)}/2^{(m-1)(n-1)} = 2^{-(m-1)}, \tag{77}$$

which completes the proof. □

In the case when $\Lambda_{m,n}$ is viewed as a torus, the emptiness formation probability is independent of the position of the empty column. We hence simply denote by EF the event that an arbitrary, fixed, column of the torus is empty. Recall that $\mu(\cdot)$ is the uniform distribution on eight-vertex configurations.

Proposition 4.5. *Consider $\Lambda_{m,n}$ with sides identified to form a torus. Then $\mu(\mathrm{EF}) = 2^{-m}$.*

Proof. Let us fix the column that is to be empty as the rightmost column. First, it is simple to see that $\mathrm{EF} \neq \varnothing$ as the configuration consisting of all vertices being type `II` has all columns empty. For an upper bound on $\mu(\mathrm{EF})$ we again use a simple counting argument. By fixing the vertex types starting from left to right and then top to bottom, we see that the leftmost column has one fixed incident arrow, hence these $n-1$ vertices have four choices. The 'bulk' $(m-2)(n-1)+1$ vertices have two choices and the remaining vertices have at most one choice. We therefore have that

$$\mu(\mathrm{EF}) \leq 2^{2(n-1)+(m-2)(n-1)+1}/2^{mn+1} = 2^{-m}. \tag{78}$$

On the other hand, consider the $m(n-1)$ plaquettes whose bounding arrows can be reversed without affecting the empty column. Note that distinct choices of which plaquettes to reverse give different configurations (we can now only reverse plaquettes in one of $A, A^c \subset F$ and not both because one of A, A^c contains the plaquettes of the empty column). Thus we can obtain $2^{m(n-1)}$ distinct configuration in EF from a fixed reference configuration. Next, note that if we reverse all arrows along a straight, vertical, non-contractible path of vertices away from the rightmost column, then we obtain a configuration in EF that can not be reached from the reference configuration by reversing arrows around plaquettes. By now reversing around plaquettes from this new eight-vertex configuration we find another $2^{m(n-1)}$ distinct configuration with the empty column. This gives that

$$\mu(\mathrm{EF}) \geq 2 \cdot 2^{m(n-1)}/2^{mn+1} = 2^{-m}, \tag{79}$$

which matches our upper bound. □

4.3. **Entropy.** Now we consider the *entropy* of the model, meaning that we compare the number of eight-vertex configurations for various boundary conditions. We will suppose that m, n are fixed such that $m+n$ is even and we revert to the notation Λ without subindices. Let η be a fixed assignment of arrows to the boundary edges (i.e. the edges $\{(m,y),(m+1,y)\}$, $\{(0,y),(1,y)\}$, $\{(x,n),(x,n+1)\}$ and $\{(x,0),(x,1)\}$) and denote by $\Delta^{\eta}_{8\mathrm{vx}}(\Lambda)$ the set of eight-vertex configurations on Λ with the boundary configuration η. We call η *valid* if $\Delta^{\eta}_{8\mathrm{vx}}(\Lambda) \neq \varnothing$.

Proposition 4.6. *A boundary condition η is valid if and only if the number of arrows around the boundary of Λ that point into Λ is even. In this case $|\Delta^{\eta}_{8\mathrm{vx}}(\Lambda)| = 2^{(m-1)(n-1)}$.*

Proof. If $\Delta^{\eta}_{8\mathrm{vx}} \neq \varnothing$, then the same argument as in the proof of Proposition 4.3 shows that all configurations in $\Delta^{\eta}_{8\mathrm{vx}}$ are reachable by flipping plaquettes, and thus $|\Delta^{\eta}_{8\mathrm{vx}}| = 2^{(m-1)(n-1)}$. It remains to determine when $\Delta^{\eta}_{8\mathrm{vx}} \neq \varnothing$.

Let η_0 be the boundary condition such that all arrows are $\rightarrow$ or $\uparrow$. By taking the configuration on Λ with every vertex of type I we see that $\Delta^{\eta_0}_{8\mathrm{vx}} \neq \varnothing$. This boundary condition does indeed have an even number $(m+n)$ of inward arrows. Now, a similar argument as for Proposition 4.1 shows that any valid boundary condition η is obtainable by flipping the exterior boundary plaquettes. Then η differs from η_0 at an even number of boundary edges, and hence still has an even number of inward pointing arrows. Indeed, if η differs from the reference boundary condition in an even number of places then we can pair off differing edges into neighbouring pairs and reverse arrow around plaquettes between the pairs, similarly to Figure 4. On the other hand, if η differs from η_0 at an odd number of edges then there is no set of plaquettes that can be flipped to obtain η. □

The next result can be interpreted as saying that the entropy of the uniform eight-vertex model on Λ is realised (up to a factor that is exponential only in the size of the boundary) for any fixed valid boundary condition. It is an immediate consequence of Propositions 4.1 and 4.6.

Proposition 4.7. *Let η be a valid boundary condition on Λ. Then*

$$\frac{|\Delta_{8\mathrm{vx}}|}{|\Delta^{\eta}_{8\mathrm{vx}}|} = 2^{m+n-1} = 2^{O(\partial\Lambda)}.$$

Appendix A. The toric code

In this appendix we summarise some of the original motivation for the toric code model, starting with basic properties of quantum codes.

Recall that classical codes store information in sequences of numbers 0 or 1, each of which is called a bit. For quantum codes, the bits 0 or 1 are replaced by so-called *qubits* which are more elaborate objects. A qubit may be defined as *an irreducible two-dimensional representation of* $\mathfrak{su}_2(\mathbb{C})$ and a quantum code as *a sub-representation of a tensor product of qubits.* Let us now unpack these definitions.

Consider as before the two-dimensional vector space $\mathbb{C}^2$ and with standard basis $|+\rangle = \binom{1}{0}$ and $|-\rangle = \binom{0}{1}$ (as mentioned before, these are often denoted $|1\rangle$ and $|0\rangle$ in quantum information theory, while $|+\rangle$ and $|-\rangle$ is more common in statistical physics). The Pauli matrices (1) generate a Lie-algebra of two-by-two Hermitian matrices with trace 0. Together with the underlying vector space $\mathbb{C}^2$ on which the matrices act, this algebra of matrices is a *representation* of $\mathfrak{su}_2(\mathbb{C})$. That is, they form a single qubit by our definition.

It is a crucial point that there are other ways to represent a qubit than the specific construction above. That is, one may find other two-dimensional vector spaces $\mathbb{V}$ (over $\mathbb{C}$), together with matrices X, Y, Z acting on $\mathbb{V}$, that have "all relevant properties" of $\mathbb{C}^2$ together with the $\sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)}$. More precisely, there are other representations of $\mathfrak{su}_2(\mathbb{C})$ which are *isomorphic* to the one above and therefore also form a single qubit. To check that $\mathbb{V}$ together with X, Y, Z form a qubit, one needs to check that X, Y, Z satisfy the following relations:

$$[X, Z] := XZ - ZX = -2iY = -2ZX, \qquad \text{and} \qquad X^2 + Y^2 + Z^2 = 3\mathbb{1}. \tag{80}$$

For classical codes, one obtains protection from errors by using *redundance*; that is, a single bit is encoded in a sequence of several bits of length $n > 1$, where the extra bits are used to detect possible errors of transmission. For quantum codes, the analogous setting is obtained using the n-fold tensor product $(\mathbb{C}^2)^{\otimes n}$. For each $\ell \in \{1, \dots, n\}$ one then has a *physical qubit* consisting of copies $\sigma_\ell^{(1)}, \sigma_\ell^{(2)}, \sigma_\ell^{(3)}$ of the Pauli matrices which act only on the ℓ entry. To obtain a quantum code with desirable properties, one sets this up in such a way that $(\mathbb{C}^2)^{\otimes n}$ has at least one two-dimensional subspace carrying a representation of $\mathfrak{su}_2(\mathbb{C})$, that is, forms a qubit (other than the ones obtained using the factors $\sigma_\ell^{(1)}, \sigma_\ell^{(2)}, \sigma_\ell^{(3)}$). Such a sub-representation is then called a *logical qubit*.

For the toric code, we use the setting described above with $n = |E|$ physical qubits, that is the physical qubits are indexed by the edge-set E of the torus. For simplicity, and consistency with notation in the literature, we will consider vertex set $\{1, \dots, d\}^2 \subset \mathbb{Z}^2$ with edge set E making it a torus. This means we have $2d^2$ physical qubits. To identify logical qubits we use the operators

$$A_s = \prod_{e \sim s} \sigma_e^{(3)}, \qquad B_p = \prod_{e \sim p} \sigma_e^{(1)}, \qquad s \in V, \quad p \in F. \tag{81}$$

(These are the same as X_s and Z_p in (2) but here we use the more common notation (81).) The relevant subspace of $(\mathbb{C}^2)^{\otimes E}$ is

$$\mathcal{L} = \big\{ |\xi\rangle \in (\mathbb{C}^2)^{\otimes E} : A_s|\xi\rangle = B_p|\xi\rangle = |\xi\rangle, \text{ for all } s \in V, p \in F \big\}, \tag{82}$$

i.e. the subspace stabilised by all A_s and B_p. We will see that $\mathcal{L}$ carries *two* logical qubits.

Let us look more closely at the condition $A_s|\xi\rangle = |\xi\rangle$. Note that if $|\xi\rangle = \otimes_{e \in E} |\xi_e\rangle$ is a tensor product of eigenvectors of the $\sigma^{(3)}$-basis then

$$A_s|\xi\rangle = \big(\prod_{e \sim s} \xi_e \big) |\xi\rangle. \tag{83}$$

Then the condition $A_s|\xi\rangle = |\xi\rangle$ is identical to the constraint defining $\Omega_{8\text{vx}} \subseteq \Omega$ as the set of configurations σ satisfying $\prod_{e \sim s} \sigma_1 = +1$ for all $s \in V$. In particular, we may identify $\mathcal{L}$ with a subset of $\Omega_{8\text{vx}}$.

Translating the multiplicative constraints defining $\mathcal{L}$ into linear constraints, one may use the same reasoning (rank-nullity) as for Proposition 4.1 to conclude that $\dim(\mathcal{L}) = 4$. Thus $\mathcal{L}$ has the correct dimension for carrying two logical qubits. To see that it indeed does, we need to define operators X_1, Z_1 and X_2, Z_2 on $\mathcal{L}$ which commute (for differing indices) and satisfy (80) (for matching indices). Let $L_1 \subseteq E$ be the set of edges of the form $\{(x, 1), (x+1, 1)\}$ for $1 \le x \le d$. Thus L_1 is a horizontal path that wraps around the torus. Also let L_1' be the set of edges of the form $\{(1, y), (2, y)\}$

for $1 \le y \le d$ forming a vertical 'ladder' around the torus. Define

$$X_1 = \prod_{e \in L_1} \sigma_e^{(1)}, \qquad Z_1 = \prod_{e \in L_1'} \sigma_e^{(3)}. \tag{84}$$

Similarly let L_2 be the set of edges of the form $\{(1,y),(1,y+1)\}$ for $1 \le y \le d$ and let L_2' be the set of edges of the form $\{(x,1),(x,2)\}$ for $1 \le x \le d$, respectively forming a vertical path and a horizontal 'ladder', and define

$$X_2 = \prod_{e \in L_2} \sigma_e^{(1)}, \qquad Z_2 = \prod_{e \in L_2'} \sigma_e^{(3)}. \tag{85}$$

One may check that these operators indeed have the desired properties, the key observation being that the paths share either zero or exactly one edge. We call these operators *logical operators* as they have the effect of applying a Pauli operator to one of the logical qubits.

Let us briefly describe the error-correction properties of the code $\mathcal{L}$ on an intuitive level.

Imagine that data is encoded as an element of the space $\mathcal{L}$. In this case the data is encoded by two qubits (because $\dim(\mathcal{L}) = 4$) but various methods (creating a boundary and removing qubits to create "holes" [14] or simply taking tensor products of multiple copies of $\mathcal{L}$) allow for a greater number of qubits and hence more data. This data may represent some information we wish to retrieve, or the state of our quantum system (quantum computer) at an intermediate step of an algorithm. An error may occur in the data, due to imperfect implementation of a step of the algorithm, or otherwise, so that the data is no longer an element of $\mathcal{L}$. This means that one or more qubits are in a state that the algorithm did not intend. Assuming the hardware implementing the algorithm is well built, we would hope that the most likely cause for this fault results in a *minimal* (non-zero) number of the constraints $A_s|\xi\rangle = |\xi\rangle$ and $B_p|\xi\rangle = |\xi\rangle$ being broken. By parity constraints, this minimal number is two (since $\prod_{s\in V} A_s = 1$ and $\prod_{p\in F} B_p = 1$). Suppose that $A_{s_1}|\xi\rangle \neq |\xi\rangle$ and $A_{s_2}|\xi\rangle \neq |\xi\rangle$, the case of constraints $\prod_{p\in F} B_p = 1$ being broken is analogous by duality, as is the case of more constraints being broken. Thinking in terms of arrow configurations $\omega \in \Omega$, $|\xi\rangle$ can be thought of as a linear combination of arrow configurations. This means that s_1 and s_2 are *faults*, i.e. the total number of in- (or out-) pointing arrows is odd in each term of this linear combination. A moments thought will reveal that there must be a path π of adjacent vertices and edges whose end vertices are s_1 and s_2 such that each arrow along edges of this path has been reversed (flipped). We say the corresponding qubits have picked up *faults* and note that there are at least $\|s_1 - s_2\|$ of these *faulty qubits*. By flipping arrows (i.e. applying $\sigma^{(1)}$) at every edge along this path we will correct this fault. In other words, we have that $|\xi\rangle = \prod_{e\in\pi} \sigma_e^{(1)}|\xi'\rangle$ for some $|\xi'\rangle \in \mathcal{L}$ (the element of $\mathcal{L}$ we would have if there had been no fault) and hence that $|\xi'\rangle = \prod_{e\in\pi} \sigma_e^{(1)}|\xi\rangle$. Notice, however, that we only know the end points of this path and there is no way to know the path itself because $A_s|\xi\rangle = |\xi\rangle$ for every s in the interior of the path (in terms of arrows, each such s has had two adjacent arrows flipped, preserving the even parity of incoming arrows at s). Thankfully this is often not a problematic feature of the model (in fact, in some sense, it is an *essential* feature of the model for error correction). We can flip arrows (i.e. apply $\sigma^{(1)}$) to all the qubits along *any* path π' with end points s_1 and s_2 and often obtain $|\xi'\rangle \in \mathcal{L}$. Indeed, $\prod_{e'\in\pi'} \sigma^{(1)}_{e'}|\xi\rangle = \prod_{e\in\pi\sqcup\pi'} \sigma^{(1)}_{e'}|\xi'\rangle$ and if $\pi\cup\pi'$ is a contractible (closed) path then $\prod_{e\in\pi\sqcup\pi'} \sigma^{(1)}_{e'} = \prod_{p\in P} B_p$ where P is the set of plaquettes enclosed by $\pi\cup\pi'$. By noting that $P = P^{-1}$ and $P|\xi'\rangle = |\xi'\rangle$ (as $|\xi'\rangle$ is an element of $\mathcal{L}$) we see that flipping arrows along π' returns us to $\mathcal{L}$. The element, $|\xi'\rangle$, that we obtain in this way is the element we would have obtained if no fault occurred and we say we have successfully *corrected the fault.* If $\pi\cup\pi'$ is not contractible then $\prod_{e\in\pi\sqcup\pi'} \sigma^{(1)}_{e'}$ is not a product of B_p operators, so applying it to $|\xi\rangle$ has the effect of applying a logical operator.

In order to avoid the case where $\pi\cup\pi'$ is not contractible as much as possible, a sensible choice of π' is a shortest path between s_1 and s_2. This will result in $\pi\cup\pi'$ being contractible whenever $|s_1(j) - s_2(j)| < d/2$ for $j = 1, 2$, where $s_i(j)$ is the j^{th} coordinate of s_i. It is common to select d to be odd to avoid a "tie-breaking" situation when $|s_1(j) - s_2(j)| = d/2$. In this case we are guaranteed to successfully correct the fault whenever $\|s_1 - s_2\| \le (d-1)/2$ (we will also successfully correct cases where $\|s_1 - s_2\| > (d-1)/2$ but $|s_1(j) - s_2(j)| < d/2$ for $j = 1, 2$) and we say the code has *distance d*. Notice that in our case (and indeed generally) d is the minimum distance (i.e. number of qubits that need to be acted upon) between *code words* (elements of $\mathcal{L}$). For the toric code the distance between neighbouring code words is constant, but this need not be the case in general. Also note that n, k, d do not necessarily uniquely identify the code. In the literature the parameters are often written together as $[[n, k, d]]$ to describe a code with n physical qubits encoding k logical qubits and with the guaranteed ability to correct $(d-1)/2$ faults.

References

[1] D. Allison and N. Reshetikhin, *Numerical study of the 6-vertex model with domain wall boundary conditions*, Ann. Inst. Fourier, 55(6), 2005

[2] R. Alicki, M. Fannes, and M. Horodecki, *A statistical mechanics view on Kitaev's proposal for quantum memories*, J. Phys. A: Math. Theor. 40 6451

[3] E. Ardonne, P. Fendley, and E. Fradkin, *Topological order and conformal quantum critical points*, Ann. Phys. 310(2) pp. 493-551, 2004

[4] R. J. Baxter, *Eight-Vertex Model in Lattice Statistics*, Phys. Rev. Lett. vol. 26 pp. 832-833, 1971

[5] R. J. Baxter, *Exactly solved models in statistical mechanics*, London Academic Press Limited, 1982

[6] C. Benassi, B. Lees, and D. Ueltschi, *Correlation Inequalities for the Quantum XY Model*, J. Stat. Phys. 164 1157-1166, 2016

[7] C. Benassi, B. Lees, and D. Ueltschi: Correlation Inequalities for classical and quantum XY models, In: Advances in Quantum mechanics. *Springer INdAM Series*, Vol. 18, pp.15-31. 2017

[8] F. Colomo and A. G. Pronko *The arctic circle revisited*, Contemp. Math. 458 pp. 361-376, 2008

[9] H. Duminil-Copin *100 Years of the (Critical) Ising Model on the Hypercubic Lattice*, Proc. Int. Cong. Math. 2022 Vol. 1

[10] H. Duminil-Copin, M. Gagnebin, M. Harel, I. Manolescu, and V. Tassion. *Discontinuity of the phase transition for the planar random-cluster and Potts models with* $q > 4$, Annales Scientifiques ENS 54(6), 1363-1413, 2021

[11] H. Duminil-Copin, A. Karrila, I. Manolescu, and M. Oulamara, *Delocalization of the height function of the six-vertex model*, arXiv2012.13750, 2020

[12] H. Duminil-Copin, K. K. Kozlowski, D. Krachun, I. Manolescu, and Tikhonovskaia, *On the six-vertex model's free energy*, arXiv preprint arXiv:2012.11675, 2020.

[13] S. Friedli and Y. Velenik, *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction*, Cambridge University Press, 2017

[14] A. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, *Surface codes: Towards practical large-scale quantum computation*, Physical Review A, 86(3), p.032324, 2012

[15] C. Fan and F- Y- Wu, *General Lattice Model of Phase Transitions*, Phys. Rev. B 2, 723, 1970

[16] J. Ginibre *General formulation of Griffiths' inequalities*, Comm. Math. Phys. Vol 16(4), pp.310-328, 1970

[17] R. B. Griffiths *Correlations in Ising ferromagnets. I*, J. Math. Phys. Vol. 8, pp.478-483, 1967

[18] A. Glazman and R. Peled, *On the transition between the disordered and antiferroelectric phases of the 6-vertex model*, arXiv preprint arXiv:1909.03436

[19] G. Grimmett, *The Random-Cluster Model* Vol. 333 Berlin: Springer, 2006

[20] L. K. Grover, *Quantum Mechanics Helps in Searching for a Needle in a Haystack*, Phys. Rev. Lett. Vol. 79 325, 1997

[21] L. K. Grover, *A Fast Quantum Mechanical Algorithm for Database Search*, Ann. ACM Symp. Th. Comp. pp. 212-219, 1996

[22] A. Yu. Kitaev, *Fault-tolerant quantum computation by anyons*, Ann. Phys. 303(1) pp. 2-30, 2003

[23] V. E. Korepin, *Calculations of norms of Bethe wave functions*, Commum. Math. Phys. 86 pp. 391-418, 1982

[24] V. E. Korepin and P. Zinn-Justin, *Thermodynamic limit of the six-vertex model with domain wall boundary conditions*, J. Phys. A: Math. Gen. 33 7053, 2000

[25] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov Chains and Mixing Times*, AMS, 2009

[26] E. H. Lieb, *Residual Entropy of Square Ice*, Phys. Rev. vol. 162 pp. 162-172, 1967

[27] E. H. Lieb, *Exact Solution of the F Model of An Antiferroelectric*, Phys. Rev. Lett. vol. 18 pp. 1046-1048, 1967

[28] E. H. Lieb, *Exact Solution of the Two-Dimensional Slater KDP Model of a Ferroelectric*, Phys. Rev. Lett. vol. 19 pp. 108-110, 1967

[29] M. A. Nielsen and I. L. Chuang *Quantum Computation and Quantum Information*, Cam. Uni. Press, 2010

[30] P. W. Shor, *Algorithms for quantum computation: discrete logarithms and factoring*, Proc. 35th Ann. Symp. Found. Comp. Sci., pp. 124-134, 1994

[31] C. Wang, J. Harrington, and J. Preskill, *Confinement-Higgs transition in a disordered gauge theory and the accuracy threshold for quantum memory*, Ann. Phys. 303 31, 2003

[32] B. Sutherland, *Two-Dimensional Hydrogen Bonded Crystals without the Ice Rule*, J. Math. Phys. 11 3183, 1970

[33] C. N. Yang and C. P. Yang. *Ground-state energy of a Heisenberg-Ising lattice*, Physical Review, 147(1), p.303, 1966