The fundamental group of 2-dimensional random cubical complexes

Matthew Kahle, Elliot Paquette, Érika Roldán

January 23, 2020

Abstract

We study the fundamental group of certain random 2-dimensional cubical complexes which contain the complete 1-skeleton of the $d$-dimensional cube, and where every 2-dimensional square face has probability $p$. These are cubical analogues of Linial–Meshulam random simplicial complexes, and also simultaneously are 2-dimensional versions of bond percolation on the hypercube. Our main result is that if $p \leq 1/2$, then with high probability the fundamental group contains a nontrivial free group, and if $p > 1/2$ then with high probability it is trivial. As a corollary, we get the same result for homology with any coefficient ring. We also study the structure of the fundamental group below the transition point, especially its free factorization.

1 Introduction

In this article, we study a model of random 2-dimensional cubical complexes. Denote the $n$-dimensional cube by $Q^n = [0,1]^n$, and the set of vertices of the $n$-dimensional cube by $Q^n_0$. This makes $Q^n_0 = \{0,1\}^n$, which is the set of all $n$-tuples with binary entries. More generally, denote by $Q^n_k$ the $k$-skeleton of $Q^n$. For example, $Q^n_1$ is the graph with vertex set $Q^n_0$ and an edge (a 1-face) between two vertices if and only if they differ by exactly one coordinate. Define the random 2-dimensional cubical complex $Q_2(n,p)$ as having 1-skeleton $Q^n_1$ and including each 2-dimensional face of $Q^n$ independently with probability $p$. 
The space $Q_2(n, p)$ is a cubical analogue of the random simplicial complex $Y_2(n, p)$ introduced by Linial and Meshulam in [22], whose theory is well-developed. The random complex $Y_2(n, p)$ is defined by taking the complete 1-skeleton of the $n$-dimensional simplex $\Delta^n$, and including into it each 2-face independently and with probability $p$. In this way, $Q_2(n, p)$ is constructed in exactly the same way as $Y_2(n, p)$, except that the underlying polytope $\Delta^n$ is replaced by $Q^n$.

The space $Q_2(n, p)$ is also a 2-dimensional version of the random graph studied by Burtin [5], Erdős and Spencer [13], and others; see [21] for a 1992 survey on random cubical graphs. More precisely, let $Q(n, p)$ denote the random subgraph defined by including all vertices of $Q^n$, i.e. $Q^n_0$, and including each edge in $Q^n_1$ independently with probability $p$. One can view $Q(n, p)$ as a natural cubical analogue of $G(n, p)$, the Bernoulli or Erdős-Rényi random graph.

The first major result on $Q(n, p)$ concerns its connectivity:

**Theorem 1.1** (Burtin [5] and Erdős–Spencer [13]). For $Q \sim Q(n, p)$ and for any fixed $p \in [0, 1]$,

$$\lim_{n \to \infty} \mathbb{P}[Q \text{ is connected}] = \begin{cases} 0 & \text{if } p < 1/2, \\ e^{-1} & \text{if } p = 1/2, \\ 1 & \text{if } p > 1/2. \end{cases}$$ (1)

Specifically, Burtin proved that if $p < 1/2$, then $\lim_{n \to \infty} f_n(p) = 0$ and if $p > 1/2$ then $\lim_{n \to \infty} f_n(p) = 1$. Later, Erdős and Spencer refined this argument to show what happens if $p = 1/2$. Moreover they show that for $p \geq 1/2$, the only connected components of $Q(n, p)$ are either isolated or form a giant component (meaning any two vertices that are not isolated are in the same connected component). They then show the number of isolated vertices has a limiting Poisson distribution with mean 1, and as a consequence

$$\mathbb{P}[\beta_0 = k + 1] \to e^{-1}/k!$$

for every integer $k \geq 0$, where $\beta_0$ is the 0-th Betti number, i.e. the number of connected components of $Q(n, p)$.

This picture strongly mirrors what is seen for Erdős-Rényi graphs, with the only major difference being that $p$ should be taken as a function of $n$ to
see interesting connectivity behavior. With \( p = \frac{\log n + c}{n} \) for \( c \) fixed and with \( G \sim G(n, p) \)
\[
\lim_{n \to \infty} \mathbb{P}[G \text{ is connected}] = e^{-e^{-c}}
\]
(see [12] or [4]). Furthermore the same conclusion holds with \( c = \pm \infty \) if we take \( p = \frac{\log n \pm f(n)}{n} \) where \( f(n) \to \infty \). Moreover, one has the same type of description of the components of \( G(n, p) \) for \( p = p(n) \) satisfying \( p(n)n/\log n \to x \in (\frac{1}{2}, 1) \): all components are either isolated vertices or part of a giant component. Even the proofs share a strong similarity, in that the method is to enumerate potential cutsets and show that they are rare by making a first moment estimate of the number of cutsets.

It is therefore perhaps reasonable to speculate that the topological phenomenology of the higher dimensional process \( Q_2(n, p) \) mirrors that of \( Y_2(n, p) \) after appropriately adjusting how \( p \) is chosen as a function of \( n \). We shall show, however, that there are major differences between the topology of \( Q_2(n, p) \) and \( Y_2(n, p) \).

Before discussing our results and these differences, we introduce some common terminology. Our focus is on typical behavior of random objects for large values of \( n \). So, we will say that a sequence of statements \( \mathcal{P}_n \) holds with high probability (abbreviated whp) if
\[
\lim_{n \to \infty} \mathbb{P}[\mathcal{P}_n] = 1.
\]

We will make use of the Landau notations \( O, o, \omega, \Omega, \Theta \) in the asymptotic sense, so that \( f = O(g) \) means \( f/g \) is eventually bounded above as \( n \to \infty \) and \( f = o(g) \) means \( f/g \to 0 \) as \( n \to \infty \). Also, \( f = \omega(g) \) means \( g = o(f) \) and \( f = \Omega(g) \) means \( g = O(f) \). Finally, we will use \( f = \Theta(g) \) to mean \( f = O(g) \) and \( f = \Omega(g) \). We occasionally display parameters like \( O_{a, b, c}(\cdot) \), emphasizing that the implied constants depend on \( a, b, c \).

We will also make use of the notion of thresholds. A function \( f = f(n) \) is said to be a threshold for a property \( \mathcal{P} \) of a sequence of random objects \( G = G_{n, p} \) if \( p = \omega(f) \) implies \( G \in \mathcal{P} \) w.h.p. and \( p = o(f) \) implies \( G \notin \mathcal{P} \) w.h.p. Such a threshold is only defined up to nindependent scalar multiples. If there is a function \( g = o(f) \) so that \( p \geq f + g \) implies \( G \in \mathcal{P} \) w.h.p. and \( p \leq f - g \) implies \( G \notin \mathcal{P} \) w.h.p. the threshold is sharp. If no such \( g \) exists, the threshold is coarse.

In this paper we study the fundamental group \( \pi_1(Q) \) for \( Q \sim Q_2(n, p) \).

The fundamental group can be given a purely combinatorial representation
for a space such as $Q_2(n, p)$, which we discuss in Section 2.1. Our first result establishes the threshold for $\pi_1(Q) = 0$, i.e. for $Q \sim Q_2(n, p)$ to be simply connected. Here, we formulate the theorem for $p$ fixed independently of $n$.

**Theorem 1.2.** With $Q \sim Q_2(n, p)$ if $p > 1/2$, then $\pi_1(Q) = 0$ asymptotically almost surely. Conversely, if $p \leq 1/2$, then whp there are finitely generated groups $G$ and $F$ so that $\pi_1(Q) \cong G \ast F$ and where $F$ is a free group of rank at least 2.

We recall that for any groups $G$ and $H$, the free product $G \ast H$ is the group of all words $\{g_1 h_1 g_2 \cdots g_k h_k\}$ with the operation of concatenation and where $\{g_i\}$ and $\{h_i\}$ are any elements of their respective groups. Two words are considered equivalent if and only if they are equal after removing all those $\{g_i\}$ and $\{h_i\}$ equal to the identity.

This theorem marks a substantial difference between $Q_2(n, p)$ and $Y_2(n, p)$. The threshold (in a suitably coarse sense) for $\pi_1(Y) = 0$ for $Y \sim Y_2(n, p)$ is $p = n^{-1/2}$, which is proven by Babson, Hoffman and Kahle [3]. This threshold is subsequently sharpened by Luria and Peled [25]. However, for $p = p_n$ below this threshold also satisfying $p \geq (2 + \epsilon) \log n/n$, it is shown in [19] that $\pi_1(Y_2(n, p))$ has Kazhdan’s property (T) (see [19] for the discussion therein). This property precludes the possibility of having any nontrivial free subgroup.

Moreover, the homology vanishing threshold $Q \sim Q_2(n, p)$ coincides with the threshold for simple-connectedness. The first homology group $H_1(Q; \mathbb{Z})$ is the abelianization $\pi_1(Q)/[\pi_1(Q), \pi_1(Q)]$. Therefore from Theorem 1.2, $p = \frac{1}{2}$ is a sharp threshold for $H_1(Q; \mathbb{Z}) = 0$.

In $Y_2(n, p)$, the homology vanishing threshold is $2 \log n/n$ (due to [22] over $\mathbb{F}_2$, to [26] over general fields, and finally to $\mathbb{Z}$ coefficients by [24]; see also [28] for the extension to higher dimensions). Hence in $Y_2(n, p)$, there is a wide range of $p$ below the simple-connectivity threshold in which $Y_2(n, p)$ has nontrivial fundamental group but trivial homology. Again, in $Q_2(n, p)$ the threshold for the vanishing of homology coincides with the threshold for the vanishing of $\pi_1$.

Moreover the fundamental group of $Q_2(n, p)$ below the vanishing threshold displays remarkably different structure. Recall that any finitely generated group $G$ has a free product decomposition [16, Chapter 1.2], i.e. a unique representation

$$G = G_1 \ast G_2 \ast \cdots \ast G_k,$$
where the \( \{G_i\} \) are subgroups of \( G \). For \( p > 1 - (\frac{1}{2})^{1/2} \), we are able to completely characterize the fundamental group \( \pi_1(Q) \) for \( Q \sim Q_2(n,p) \) :

**Theorem 1.3.** For \( p > 1 - (\frac{1}{2})^{1/2} \), with high probability, for \( Q \sim Q_2(n,p) \)

\[
\pi_1(Q) \cong (\mathbb{Z} * \mathbb{Z} * \cdots * \mathbb{Z})_N,
\]

where \( N \) denotes the number of isolated 1-faces in \( Q \).

An isolated 1-face is one which is not contained in any 2-face. Hence, just below the \( p = \frac{1}{2} \) threshold, only isolated 1-faces contribute to the fundamental group. This is strongly reminiscent of the homology of the simplicial complex \( Y \sim Y_2(n,p) \) just below its homology vanishing threshold. From [19], for \( p = t \log n/n \) with \( t \in (1, 2) \), the homology group \( H_1(Y; \mathbb{Z}) \) is a free abelian group with rank given by the number of isolated 1-faces of \( Y \). The fact that in \( Q_2(n,p) \) this holds for the fundamental group is a substantially stronger statement.

**Remark.** We could also say more about \( N \), and hence about the probability \( \pi_1(Q) = 0 \) for \( Q \sim Q_2(n,p) \). If we take for fixed \( c \in \mathbb{R} \),

\[
p = \frac{1}{2} \left( 1 + \frac{\log n + c}{n} \right),
\]

then \( \mathbb{P}(N = 0) \to e^{-e^{-c}} \) from a standard Poisson approximation. It will also follow that \( \mathbb{P}(\pi_1(Q) = 0) \to e^{-e^{-c}} \). See [1] for details on such Poisson approximations.

It is also possible to formulate a process version of this statement. Here one couples all \( Q_2(n,p) \) together for all \( p \in [0, 1] \) in a monotone fashion, so the 2-faces of \( Q_2(n,p_1) \) are a subset of \( Q_2(n,p_2) \) whenever \( p_1 \leq p_2 \). Let \( (Q_p : p \in [0, 1]) \) have this distribution, and let \( N_p \) be the number of isolated 1-faces in \( Q_p \). Then we can formulate the stopping times \( T_{sc} \) and \( T_{2d} \) as

\[
T_{sc} = \inf\{p : \pi_1(Q_p) = 0\} \quad \text{and} \quad T_{2d} = \inf\{p : N_p = 0\}.
\]

Then from Theorem 1.3, \( T_{sc} = T_{2d} \) whp.

For general \( p \leq 1 - \frac{1}{\sqrt{2}} \), we are not able to completely describe the fundamental group \( \pi_1(Q) \) for \( Q \sim Q_2(n,p) \), but we nonetheless give a partial characterization of the free factors that arise.
Definition 1.4. For a cubical sub-complex $T$ of $Q^n$, define its edge complexity $e(T)$ as the number of edges in $T$. Let $\mathcal{S}_p$ be the set of pure 2-dimensional strongly connected cubical complexes $T$ that are subcomplexes of $Q_2^n$ for some $n$ and so that $(1 - (\frac{1}{2})^{1/e(T)}) < p$.

While we do not characterize all free factors, we are able to characterize some:

Theorem 1.5. For any $p \in (0, 1)$, and for $Q \sim Q_2(n, p)$, let the free product decomposition of $\pi_1(Q)$ be given by

$$\pi_1(Q) \cong F * \pi_1(X_1) * \pi_1(X_2) * \cdots * \pi_1(X_\ell),$$

with $F$ a free group. With high probability, any $T \in \mathcal{S}_p$ appears as a factor $\pi_1(X_j)$ for some $1 \leq j \leq \ell$.

Remark. It is possible to say more about the number of factors of $\pi_1(T)$ in Theorem 1.5, for generic $p$ the number of copies of $\pi_1(T)$ grows exponentially in $n$ whp. Finding the exact asymptotic is an open problem.

As a highlight of what is possible, for $0 < p < 1 - (\frac{1}{2})^{1/18} \approx 0.037776$ and $Q \sim Q_2(n, p)$ we show that $\pi_1(Q)$ has a $\mathbb{Z}/(2\mathbb{Z})$ free factor whp (see Corollary 5.1). This in particular shows that $H_1(Q; \mathbb{Z})$ has torsion elements already for all $p \in (0, p_c)$, where $p_c$ is some critical value in $(0, 1)$. The question of when torsion appears in $H_1(Y; \mathbb{Z})$ for $Y \sim Y_2(n, p)$ is a question of major interest [20]. As another possible factor, for $0 < p < 1 - (\frac{1}{2})^{1/32} \approx 0.021428$, $\pi_1(Q)$ has a $\mathbb{Z} \times \mathbb{Z}$ free factor whp (see Corollary 5.2).

Conjecture 1.6. For $p > 1 - (\frac{1}{2})^{1/18}$ and $Q \sim Q_2(n, p)$, $\pi_1(Q)$ is torsion-free whp.

We further believe it is possible that for $p$ above this threshold, $\pi_1(Q)$ is free whp. For the case of $\pi_1(Y)$ with $Y \sim Y_2(n, p)$ the sharp threshold is found by Newman in [27], improving on previous work of [6].

Discussion

We have not addressed many results about the fundamental group of $\pi_1(Y)$ for $Y \sim Y_2(n, p)$ which may have interesting analogues in $Q_2(n, p)$, which could further elucidate what appear to be deep differences between simplicial and cubical random complexes. As the body of literature on $Y_2(n, p)$, is
substantial, we discuss possible directions of interest for questions about $Q_2(n, p)$.

A major such result on $Y_2(n, p)$ is that of Costa and Farber [10] addressing the threshold for $Y_2(n, p)$ to be asphericable, meaning that some 2-faces can be removed in such a way that $\pi_1(Y)$ is unchanged but $\pi_2(Y) = 0$. This could prove an interesting direction for $Q_2(n, p)$.

Many interesting topological phases of $\pi_1(Q)$ are likely to exist when $p$ tends to 0 with $n$. For $Y_2(n, p)$, a particularly rich regime of $p$ is when the mean degree of an edge $np$ tends to a constant, and we would expect this regime to be similarly rich for $Q_2(n, p)$: to name a few transitions that should appear in this regime, the collapsibility threshold [2], the threshold for a giant shadow [23], and the threshold for $\pi_1(Q)$ to have an irreducible factor in its free product decomposition with growing rank.

A natural direction is to consider higher dimensional complexes $Q_d(n, p)$, built in an analogous way to $Y_d(n, p)$. For $Q_d(n, p)$ it may be possible to analyze the higher homotopy groups in a similar fashion to what is done here.

In a different direction, we mention that all of the results we present are about the $n \to \infty$ limit, but have some content for some large $n$. These could provide useful results for understanding 2-dimensional percolation on a sufficiently high dimensional lattice $\mathbb{Z}^n$. There are some recent related results for such higher dimensional cubical percolation [11, 17, 18].

**Multiparameter generalizations**

In the random cubical graph literature, there is a 2-parameter model $Q(n; p_0, p_1)$ (see [21] for a survey of some results). First, we take a random induced subgraph of the $n$-cube, where every vertex with probability $p_0$ independently. Then we include each of the remaining edges with probability $p_1$ independently. Bond percolation on the hypercube is the random cubical graph where $p_0 = 1$ and $p_1$ varies, and site percolation is where $p_1 = 1$ and $p_0$ varies.

It seems natural to form a higher dimensional generalization of this model $Q_2(n; p_0, p_1, p_2)$. Indeed, Costa and Farber have made a detailed study of the analogous model $Y_2(n; p_0, p_1, p_2)$ (see [7–9]), including many interesting results on the fundamental group. See also [14] wherein new questions about the fundamental group are discussed for this multiparameter model. Our discussion has been about the special case where $p_0 = p_1 = 1$ and $p_2$ varies.
For, $Q_2(n; p_0, p_1, p_2)$, it is natural to ask if there is a critical surface for homology $H_1$ vanishing in the unit cube. The case of setting $p_1 = p_2 = 1$, and letting $p_0$ vary looks particularly interesting, analogous to the site percolation model. Higher homology is no longer monotone, as in, for example, a random clique complex or Vietoris–Rips complex. Are there separate thresholds for $H_0$ vanishing, $H_1$ appearing, $H_1$ vanishing, and $H_2$ appearing?

**Overview and organization**

We begin with Section 2 where we define some key notions for working with $Q_2(n, p)$, and we make some elementary estimates about it. In Section 2.1 we give a combinatorial definition of $\pi_1$. In Section 2.2 we introduce notation to work with subcomplexes of $Q^n$, and we introduce the notion of parallel faces. Here we make an explicit connection between $Q_2(n, p)$ and $Q(n, p^4)$. In Section 2.3 we summarize some estimates from [13] that we need about the random graph $Q(n, p)$. In Section 2.4 we make estimates for the existence of uncovered 1-faces, and we use this to deduce Theorem 1.2 from Theorem 1.3.

In Section 3, we introduce an algorithm for identifying contractible 4-cycles in $\pi_1$. This algorithm reduces the analysis of $\pi_1$ to determining the topology of small subcomplexes. In this section, we finish the proof of Theorems 1.2 and 1.3.

In Section 4, we show a general structure theorem that describes the free product decomposition of $\pi_1(Q_2(n, p))$, and we then prove Theorem 1.5. In Section 5, we construct specific complexes which show that certain interesting free factors appear.

**Acknowledgements**

We would like to thank Yuval Peled for interesting conversations that helped launch this work.
2 Preliminaries

2.1 The edge path group

For subcomplexes $Q$ of $Q^n$, the fundamental group $\pi_1(Q)$ has a nice combinatorial definition as the edge path group, which we now define.

Say that two edges (1-faces) of $Q^n$ are adjacent if they intersect at a vertex. An edge-path in $Q^n$ is defined to be a sequence of edges, for which every consecutive pair are adjacent.

In $Q^n$ any 2-face has a 4-cycle as its boundary. Conversely, every 4-cycle in $Q^n$ is the boundary of a 2-face. Hence, any two adjacent edges are contained in a unique 2-face and therefore in a unique 4-cycle in $Q^n$.

Two edge-paths in $Q$ are said to be edge-equivalent if one can be obtained from the other by successively doing one of the following moves:

1) replacing two consecutive adjacent edges by the two opposite edges of the 4-cycle $x$ that contains them, if the 2-face that bounds $x$ is in $Q$;

2) replacing one edge contained in a 4-cycle $x$ with the other three consecutive edges in $x$, if the 2-face that bounds $x$ is in $Q$;

3) replacing three consecutive edges in a 4-cycle $x$ with the complementary edge in $x$, if the 2-face that bounds $x$ is in $Q$;

4) removing an edge that appears twice consecutively or adding an edge that appears twice consecutively.

Define $\overrightarrow{0}$ to be the vector with only zero entries in $Q^n_0$. An edge-loop at $\overrightarrow{0}$ is an edge-path starting and ending at $\overrightarrow{0}$. The random edge-path group $\pi_1(Q)$ is defined as the set of edge-equivalence classes of edge-loops at $\overrightarrow{0}$ (with product and inverse defined by concatenation and reversal of edge-loop).

We explore first the extremal cases. If $p = 0$ then any $Q \sim Q_2(n, p)$ is equal to $Q^n_0$, that is, with a probability of one the complex $Q$ has not a single 2-face included. Observing that in any graph $G$, the number of independent generators in $\pi_1(G)$ is equal to $E(G) - V(G) + 1$, in the case of an element $Q \sim Q_2(n, p)$, we get

$$E(C) = 2^{n-1}n, \text{ and } V(G) = 2^n,$$

which implies that the number of independent generators in $\pi_1(Q)$ will be at most $2^{n-1}(n - 2) + 1$. Thus, when $p = 0$ we have that $\pi_1(Q)$ is a free
group with \(2^{n-1}(n-2)+1\) independent generators, and this is the maximum number of independent generators that the edge-path group of a random 2-cubical complex can attain. This number of independent generators is less than the total number of 4-cycles in \(Q^n\) which is \(2^{n-3}n(n-1)\). If \(p = 1\) then any \(Q \sim Q_2(n, p)\) will have all the 2-faces included, which implies that \(\pi_1(Q) = 0\) with a probability of 1.

### 2.2 Star notation and the parallel relation

In \(Q^n\), the four vertices belonging to a 4-cycle have \(n-2\) equal entries and two coordinate entries that are not equal in all of them. Denote these non-equal coordinate entries as \(i\) and \(j\), then we can uniquely represent a 4-cycle using an \(n\)-tuple with \(n-2\) fixed binary values and two \(*\). One \(*\) will be located on coordinate \(i\), and the other will be located on coordinate \(j\). As an example, the 4-cycles of \(Q^3\) are \(\{(0, *, *), (*, 1, *), (*, *, 1), (*, 0, *), (1, *, *), (*, *, 0)\}\), with, for instance, \((*, *, 0) = \{(1, 1, 0), (1, 0, 0), (0, 1, 0), (0, 0, 0)\}\).

For dice, physical realizations of the cube \(Q^3\), we have a physical intuitive notion of parallel faces; there are three pairs of parallel faces, and if the die is fair each pair should add up to 7. Using the \(*\) notation of 4-cycles introduce above, we extend this notion of parallel faces to \(Q^n\).

**Definition 2.1.** Two 4-cycles in \(Q^n\) are parallel if they have the two \(*\) in the same entries, and if their Hamming distance is 1.

Thus, in \(Q^3\) (the cube), there three pairs of parallel 4-cycles: \((0, *, *)\) and \((1, *, *)\), the 4-cycles \((0, *, *)\) and \((1, *, *)\), and the 4-cycles \((0, *, *)\) and \((1, *, *)\). With this parallel notion, we are able to define a binary relation in the set of 4-cycles of a random 2-cubical complex.

We represent a 3-dimensional cube in \(Q^n\) with a vector with \(n\) entries, three of which have a fixed \(*\), and the rest of which are binary numbers. If we have two parallel faces \(x\) and \(y\) that have \(*\) in the \(i\) and \(j\) coordinates and which differ (only) in the binary value of the \(k\) coordinate, then the cube that contains them is represented by a vector with the three fixed \(*\) in entries \(i\), \(j\), and \(k\), and with the rest of the \(n-3\) entries equal to the entries of \(x\) (which are also equal to the entries of \(y\)).

Observe that the other four 4-cycles of the cube will have all the entries that are not \(i\), \(j\), or \(k\) equal to the entries of \(x\), two \(*\) in positions either \(\{i, k\}\) or \(\{j, k\}\), and a binary number in the remaining coordinate. Observe that
given two parallel 2-faces in $Q^n$ there is a unique 3-dimensional cube in $Q^n$ that contains them.

**Definition 2.2.** Given $Q \sim Q_2(n,p)$, two parallel 4-cycles $x,y \in Q$ are related if the 3-dimensional cube that contains them has a 2-face attached to each one of the other four 4-cycles in the cube. We represent this by $x \parallel y$.

**Lemma 2.3.** If two 4-cycles, $x$ and $y$, are related ($x \parallel y$), then they are edge-equivalent.

**Definition 2.4 (Graph of parallel related 4-cycles).** Given $Q \sim Q_2(n,p)$, we define its graph of parallel cycles, that we represent by $G[Q]$, as the graph with set of vertices $V^n$ whose elements are all the the 4-cycles in $Q^n$ (there are $2^{n-3}n(n-1)$ 4-cycles), and an edge between two of them if they are related by $\parallel$.

Observe that $\parallel$ is reflexive but not transitive. This implies, for instance, that $G[Q_2^n]$ (remember that $Q_2^n$ is the 2-skeleton of $Q^n$) is not the complete graph. We can completely characterize $G[Q_2^n]$.

**Lemma 2.5.** The graph $G[Q_2^n]$ has $\binom{n}{2}$ components, and each one of these components is a $Q_1^{n-2}$ graph.

**Proof.** Fix an $n > 0$. By definition of the relation $\parallel$, a necessary condition for two 4-cycles to be related is to have their two stars in the same position. There are $\binom{n}{2}$ ways of choosing the positions of two stars in a vector of size $n$, which implies that there are at most $\binom{n}{2}$ components in $G[Q_2^n]$. This gives us a partition of the set of vertices that we represent by

$$V^n = \bigcup_{i=1}^{\binom{n}{2}} V_i^n.$$

Let $1 \leq i \leq \binom{n}{2}$, in what follows we prove that the induced subgraph of $V_i^n$ in $G[Q_2^n]$ is isomorphic to $Q_1^{n-2}$. Any element in $V_i^n$ has the two stars in the same position, and the rest of the $n - 2$ entries have all possible binary entries. This implies that $|V_i^n| = |Q_0^{n-2}|$. Let $\phi : V_i^n \to Q_0^{n-2}$ be the natural bijection between these two sets of vertices. It is clear from Definition 2.3 that two 4-cycles $x$ and $y$ in $V_i^n$ are connected in $G[Q_2^n]$ if and only if the Hamming distance of $\phi(x)$ and $\phi(y)$ is 1. This implies that $G[V_i^n] \equiv Q_1^{n-2}$. ■
For a $Q \sim Q_2(n, p)$ the graph $G[Q]$ is a subgraph of $G[Q_2^n]$. We say that a vertex in $V^n$ is colored if the 4-cycle that it represents has its 2-face present in $Q$, and we say that it is not colored otherwise. We use the previous established partition of $V^n$ in $\binom{n}{2}$ sets to denote accordingly the induced subgraphs $G_1, ..., G_{\binom{n}{2}}$ of the graph $G[Q_2^n]$. For a $Q \sim Q_2(n, p)$, this partition defines $\binom{n}{2}$ random subgraphs, that we represent by $G_1[Q], ..., G_{\binom{n}{2}}[Q]$. Then, the edges that are included in $G_i[Q]$, for $1 \leq i \leq \binom{n}{2}$ depend on the 2-faces included in $Q$. The next lemma characterizes the probability distribution of each $G_i[Q]$.

**Lemma 2.6.** Let $p^* = p^4$ and $Q \sim Q_2(n, p)$, then, for $1 \leq i \leq \binom{n}{2}$, each random graph $G_i[Q]$ is a random graph on $Q_1^{n-2}$, with each edge included independently with probability $p^*$ and with each vertex colored independently with probability $p$. Moreover, the vertex colorings are independent of the edge set. Using the notation established in the Introduction, the uncolored graph $G_i[Q]$ has the same distribution as $Q(n-2, p^*)$ for all $1 \leq i \leq \binom{n}{2}$.

**Proof.** Let $Q \sim Q_2(n, p)$. From Lemma 2.5 we know that each $G_i[Q]$ is a random graph on $Q_1^{n-2}$. Let $x$ and $y$ be two vertices in $G_i[Q]$ that are connected in $G[Q_2^n]$ and represent this edge by $\overline{xy}$. This implies in particular that $x$ and $y$ are 4-cycles that have, with the star notation previously defined, the * in the same entries and Hamming distance equal to 1. Let $c_{\overline{xy}} \in Q^n$ be the unique 3-dimensional cube that contains $x$ and $y$.

The probability of $\overline{xy}$ being an edge in $G_i[Q]$ is equal to the probability of the other 4-cycles in $c_{\overline{xy}}$ being covered by 2-faces in $Q$. This event happens with probability $p^* = p^4$ because in $Q_2(n, p)$ each 2-face is added independently with probability $p$.

Moreover, observe that any of these 4-cycles in $c_{\overline{xy}}$ are not vertices in $G_i[Q]$ because they do not have the two stars in the same location as $x$ (or $y$). This implies the independence between the coloring of the vertices and the inclusion of the edges in $G_i[Q]$.

Let $C$ be the set of all 3-dimensional cubes $c_{\overline{xy}}$ with $x$ and $y$ varying among all unordered pairs of vertices in $G_i[Q]$ that are connected in $G[Q_2^n]$. Then, by uniqueness of the cube $c_{\overline{xy}} \in Q^n$ we have that $|C|$ is equal to the number of edges in $G_i[Q]$. Finally, edges in $G_i[Q]$ are added independently with probability $p^* = p^4$ because each 4-face in a cube in $C$ only appears in one 3-dimensional cube in $C$. ■
Remark. If $p > (1/2)^{1/4}$, then $p^* = [(1/2)^{1/4}]^4 > 1/2$ and from Theorem 1.1 and Lemma 2.6, if $Q \sim Q_2(n, p)$, for a fixed $i$ such that $1 \leq i \leq \binom{n}{2}$ we have that $G_i[Q]$ is connected. Define $H_i$ as the event that the $i$-th graph $G_i[Q]$ is connected and let $H^*$ be the event defined by

$$H^* = \bigcap_{i=1}^{\binom{n}{2}} H_i.$$ 

Observe that $H_i$ is not independent from $H_j$ if $i \neq j$, but if the probability of $\mathbb{P}(H_1^C) = o(1/d^2)$, then from a union bound $H^*$ holds whp. It is natural to expect that the proof of Theorem 1.1 from [13] gives this stronger statement for $p^* > \frac{1}{2}$.

### 2.3 The sizes of the components of $Q(n, p)$

For a given $p \in (0, 1)$, we want to better understand the structure of $Q(n, p)$. In particular, we want to study the sizes of the components of $Q(n, p)$ and how these sizes change with $p$. We will need an argument from [13] that rules out components of small sizes from appearing in $Q(n, p)$. As we slightly adapt those lemmas, we give proofs below.

Denote by $\mathcal{Q}_s$ the set of all subsets of vertices in $Q^n$ which are connected and have cardinality $s$. Given a subset $S$ of vertices in $Q^n$ that are connected, define

$$b(S) = | \{(u, v) \in Q^n \mid (u, v) \text{ is an edge in } Q^n, \ u \in S, \text{ and } v \notin S \} | . \quad (3)$$

Let

$$g(s) = \sum_{S \in \mathcal{Q}_s} (1 - p)^{b(S)}. \quad (4)$$

Then from a union bound, $g(s)$ is an upper bound for the probability of the existence of a connected component on $s$ vertices appearing in $Q(n, p)$.

**Lemma 2.7.**

$$g(s) \leq 2^n (ns)^s (1 - p)^{s(n - \lfloor \log_2(s) \rfloor)} \quad (5)$$

**Proof.** Let $s \geq 1$, then for any $S \in \mathcal{Q}_s$ we have from [15],

$$b(S) \geq s(n - \lfloor \log 2(s) \rfloor). \quad (6)$$
Also, by using that the degree of each vertex of $Q^n$ is at most $n$,

$$|Q_s| \leq 2^n(n)(2n)(3n) \cdots ((s-1)n) \leq 2^n(ns)^s.$$

(7)

Hence,

$$g(s) \leq \sum_{S \in Q_s} (1-p)^{s(n-\lfloor \log_2(s) \rfloor)} \leq 2^n(ns)^s(1-p)^{s(n-\lfloor \log_2(s) \rfloor)}.$$  

(8)

Lemma 2.8. For any $p \in (0, 1)$, there is a number $T_p \in \mathbb{N}$ and there exists $\delta, \epsilon > 0$ such that

$$\sum_{s} g(s) < 2^{-\delta n}$$

with the sum over all $s$ so that $T_p \leq s \leq 2^\epsilon n$.

Proof. Let $T_p$ be defined by

$$T_p = \inf_{T \in \mathbb{N}} \{2 \cdot (1-p)^T \} < 1.$$

Then for $s \leq 2^\epsilon n$ by Lemma 2.7,

$$g(s) \leq 2^n(ns)^s(1-p)^{s(n-\lfloor \log_2(s) \rfloor)} \leq 2^n(1-p)^{-2\epsilon n}(n2^\epsilon n(1-p)^n)^s$$

for all $n$ sufficiently large. Then

$$\sum_{s=T_p}^{[2^\epsilon n]} g(s) \leq 2^n(1-p)^{-2\epsilon n} \sum_{s=T_p}^{\infty} (n2^\epsilon n(1-p)^n)^s \leq 2^n(1-p)^{-2\epsilon n}(n2^\epsilon n(1-p)^n)^{T_p}(1+o(1)),$$

provided $\epsilon$ is chosen so that $2^\epsilon (1-p) < 1$ and $n$ is taken large. By taking $\epsilon$ sufficiently small

$$\alpha = 2^{1+\epsilon T_p}(1-p)^{-2\epsilon} < 1.$$

Hence, in terms of $\alpha$,

$$\sum_{s=T_p}^{[2^\epsilon d]} g(s) \leq \alpha^n n^{T_p}(1+o(1)) \leq 2^{-\delta n}$$

for some $\delta > 0$ sufficiently small and all $n$ sufficiently large.
2.4 The threshold for isolated edges

Any element \( Q \sim Q_2(n, p) \) has \( 2^{n-1}n \) edges, that we represent by

\[ e_1, e_2, \ldots, e_{2^{n-1}n}, \]

with each one of these edges being in \( n - 1 \) different 4-cycles. We represent by \( I_i \) the indicator function of the event that the edge \( e_i \) is isolated, that is, that none of the \( (n - 1) \) 4-cycles that contain \( e_i \) have an attached 2-face. Then,

\[ \mathbb{E}[I_i] = (1 - p)^{n-1}. \]

Let \( \mathcal{I}(Q) \) be the random variable that counts the number of isolated edges in \( Q \), i.e.

\[ \mathcal{I} = \sum_{i=1}^{2^{n-1}n} I_i. \]

Then

\[ \mathbb{E}[\mathcal{I}] = 2^{n-1}n(1 - p)^{n-1}. \]  \( \text{(10)} \)

Observe that if \( p = 1/2 \), then \( \mathbb{E}[\mathcal{I}] = n. \)

We now prove that Theorem 1.2 follows from Theorem 1.3.

**Proof of Theorem 1.2.** We first establish that for \( p > \frac{1}{2} \), \( \mathcal{I} = 0 \) whp and for \( p \leq \frac{1}{2} \), \( \mathcal{I} \geq 2 \) whp. For the first claim, the expectation (10) tends to 0. For the second, again from (10), if \( (1 - p) \geq 1/2 \) for a random 2-cubical complex, \( Q \sim Q_2(n, p) \),

\[ \mathbb{E}[\mathcal{I}] = 2^{n-1}d(1 - p)^{n-1} \geq n. \]

Thus \( \mathbb{E}[\mathcal{I}] \to \infty \) as \( n \to \infty \). Now, we use a second moment argument (see Corollary 4.3.5 of [1]) to prove that \( \mathbb{P}[\mathcal{I} \geq 2] \to 1 \) as \( n \to \infty \).

Fix an edge \( e_i \). Any other edge \( e_j \) such that \( I_j \) is not independent from \( I_i \), we represent this non-independence relation between edges \( e_i \) and \( e_j \) by \( j \sim i \), will be an edge of one and only one of the \( (n - 1) \) 4-cycles that contain \( e_i \). There are \( 3(n - 1) \) such edges and \( \mathbb{P}[I_j \mid I_i] = (1 - p)^{n-2} \). If we define

\[ \Delta_i^* = \sum_{j \sim i} \mathbb{P}[I_j \mid I_i], \]

then

\[ \Delta_i^* = \sum_{j \sim i} \mathbb{P}[I_j \mid I_i] = 3(n - 1)(1 - p)^{n-2}. \]

Thus \( \Delta_i^* = o(\mathbb{E}[I_i]) \) which implies that \( \mathbb{P}[\mathcal{I} > n/2] \to 1 \) as \( n \to \infty \).
Hence from Theorem 1.3, we have that for $p > \frac{1}{2}$, $\pi_1(Q_2(n,p)) = 0$ whp and for any $p = \frac{1}{2}$, $\pi_1(Q_2(n,p)) = G \ast \mathbb{Z} \ast \mathbb{Z}$ for some group $G$ whp. The event that $Q_2(n,p)$ has such a free factorization is a decreasing event, in that for any complex $Q$ that satisfies $\pi_1(Q) = G \ast \mathbb{Z} \ast \mathbb{Z}$ for some group $G$, removing any 2-face (i.e. removing relations from $\pi_1(Q)$) yields a complex $Q'$ so that $\pi_1(Q') = G' \ast \mathbb{Z} \ast \mathbb{Z}$ for some other group $G'$. It follows that for any $p \leq \frac{1}{2}$,

$$\mathbb{P}[\exists \ G : \pi_1(Q_2(n,p)) = G \ast \mathbb{Z} \ast \mathbb{Z}] \geq \mathbb{P}[\exists \ G : \pi_1(Q_2(n,\frac{1}{2})) = G \ast \mathbb{Z} \ast \mathbb{Z}] \to 1,$$

as $n \to \infty$, which completes the proof.

\section*{3 Parallel homotopy algorithm}

In this section, we introduce a simple iterative algorithm for finding contractible 4-cycles. For $Q_2(n,p)$ with $p > 0$, this algorithm rapidly and dramatically simplifies the fundamental group to its nontrivial parts.

We begin by introducing the algorithm. We have defined $V^n$ as the set of all 4-cycles in $Q^n$. For any subset $V \subset V^n$ we define the graph of parallel related 4-cycles denoted by $G(V)$ in a similar fashion to 2.4: the vertex set of $G(V)$ is given by the $V^n$ and two 4-cycles $x$ and $y$ are connected if they have stars in the same positions, are contained in a 3-cube $c$, and all other 4-cycles in $c$ are in $V$.

Given a $Q \sim Q_2(n,p)$ we denote by $V^n_t$ the subset of $V^n$ that are boundaries of 2-faces in $Q$. We then iteratively run the following procedure, with $t \in \mathbb{N}$.

**Stage $t$:** Build the graph of parallel related 4-cycles $G(V^n_t)$. Define the set of 4-cycles $V^n_{t+1}$ as the set of 4-cycles that are connected in $G(V^n_t)$ to a 4-cycle that is in $V^n_t$.

The algorithm stops at the first $t$ for which $V^n_{t+1} = V^n_t$.

As an aside, we observe that half of Theorem 1.2 follows from the following result:

**Theorem 3.1.** For $p > 1/2$

$$\lim_{n \to \infty} \mathbb{P}[V^n_3 = V^n] = 1. \quad (11)$$
3.1 Stage 1: explosive growth

For any set of 4-cycles \( V \subset V^4 \), say that a set of vertices \( S \) in \( G(V) \) is a quasicomponent if \( S \) is connected in \( G(V^n) \) and \( S \) is disconnected from its complement in \( G(V) \).

**Theorem 3.2.** Let \( Q \sim Q_2(n,p) \), and let \( A_s \) be the event that there exists a quasicomponent of size \( s \) in \( G(V^n) \). Then for any \( p \in (0,1) \), there is an integer \( T_p \) and \( \epsilon, \delta > 0 \) so that for all \( n \) sufficiently large

\[
P\left[ \bigcup_{s=T_p}^{2^n} A_s \right] < 2^{\delta n}. \tag{12}
\]

Also, the probability that there exists a component of \( G(V^n) \) bigger than \( T_p \) with no vertex in \( V^n \) tends to zero with \( n \).

**Proof.** The first part of the statement, inequality (12), follows by a union bound and Lemma 2.8 by observing that \( P[A_s] \leq g(s) \) -See equation (4).

By virtue of (12), it remains to show that there are no components of \( G(V^n) \) bigger than \( 2^n \) which do not intersect \( V^n \). Let \( W \) be the event that there exists a component in \( G(V^n) \) of size bigger or equal than \( 2^n \) that does not intersect \( V^n \). We show in what follows that \( P[W] \to 0 \) as \( n \to \infty \). First, we observe that \( G(V^n) = G[Q] \) and that the vertices in \( V^n \) are precisely the colored vertices in \( G[Q] \), which by Lemma 2.6 are colored independently with probability equal to \( p \). For \( 1 \leq i \leq \binom{n}{3} \), define \( W_i \) as the event that there exists in \( G_i[Q] \) a component of size bigger or equal to \( 2^n \) that has all its vertices uncolored. Thus, by Lemma 2.6,

\[
W = \bigcup_{i=1}^{\binom{n}{3}} W_i. \tag{13}
\]

Let \( 1 \leq i \leq \binom{n}{3} \). Conditioned on knowing \( G_i[Q] \), in particular on knowing that there are exactly \( l \) components with uncolored vertices and with sizes \( s_1, s_2, \ldots, s_l \), bigger than \( 2^n \) in \( G_i[Q] \) we get

\[
P[W_i | G_i] \leq \sum_{k=1}^{l} (1 - p)^{s_k}. \tag{14}
\]

Observing that \( s_1 + s_2 + \cdots s_l \leq 2^{n-2} \), it has to be the case that \( l \leq 2^{n-2} \), and because \( (1 - p) < 1 \) we have that \( (1 - p)^{s_k} \leq (1 - p)^{2^n} \) for all \( 1 \leq k \leq l \).
Thus, from equation (14) we get
\[ P[W_i \mid G_i] \leq \sum_{k=1}^{l} (1 - p)^{2^k n} \leq 2^{n-2}(1 - p)^{2^n}. \] (15)

This implies that \( E[P[W_i \mid G_i]] \leq 2^{n-2}(1 - p)^{2^n} \), and thus
\[ P[W_i] \leq 2^{n-2}(1 - p)^{2^n} \] (16)
for all \( 1 \leq i \leq \binom{n}{2} \). Finally, by a union bound argument on (13) and inequality (16) we have that
\[ P[W] \leq \left( \frac{n}{2} \right)^{2^{n-2}(1 - p)^{2^n}}, \] (17)
with
\[ \lim_{n \to \infty} \left( \frac{n}{2} \right)^{2^{n-2}(1 - p)^{2^n}} = 0. \] (18)

### 3.2 Stage 2: Only local defects remain

Let \( \mathcal{F} \) be the event that there is no quasicomponent of \( G(V_1^n) \) bigger than \( T_p \) that is disjoint from \( V_1^n \). This event was shown to hold whp by Theorem 3.2.

**Lemma 3.3.** On the event \( \mathcal{F} \), any 4-cycle \( v \) with at least \( T_p \) neighbors in \( G(V_2^n) \) is in \( V_3^n \). Likewise, any 4-cycle \( v \) with at least \( T_p \) neighbors in \( G(V_1^n) \) is in \( V_2^n \).

**Proof.** Suppose \( \mathcal{F} \) holds, and let \( v \) be any 4-cycle. Suppose that \( v \) has at least \( T_p \) neighbors in \( G(V_1^n) \). Then the connected component of \( v \) in \( G(V_1^n) \) has at least \( T_p \) neighbors, and therefore this connected component intersects \( V_1^n \). It follows by the definition of \( V_2^n \) that \( v \in V_2^n \).

Suppose now that \( v \) has at least \( T_p \) neighbors in \( G(V_2^n) \). We may suppose that \( v \) is not in a component of \( G(V_1^n) \) that intersects \( V_1^n \), for if it were, then \( v \in V_2^n \) and we are done. If none of these neighbors are in \( V_2^n \), then each is in a component of \( G(V_1^n) \) disjoint from \( V_1^n \). Hence, the union of these components and the component of \( G(V_1^n) \) containing \( v \) is a quasicomponent of \( G(V_1^n) \) that is disjoint from \( V_1^n \). Moreover, it is a quasicomponent which is larger than \( T_p \), which is disjoint from \( V_1^n \). This does not exist on \( \mathcal{F} \), and therefore \( v \) has a neighbor in \( V_2^n \). Hence \( v \in V_3^n \).
We will show that as a consequence of Lemma 3.3, in Stage 2, all those 4-cycles whose every constituent edge has high enough degree will be collapsed. For any $p$, define
\[
M_p = \inf_{M > 0} \mathbb{P}(\text{Binomial}([M/4], p^3) < T_p) < (\frac{1}{2})^{1/4}. \tag{19}
\]
For any 1-face $f$ in $Q^n$, define $\deg(f)$ as the number of 2-faces in $Q$ containing $f$. Call a 1-face of $Q \sim Q_2(n, p)$ light if its degree is less than or equal to $M_p$. Otherwise, call it heavy. We show that 4-cycles made from heavy edges are contracted in the second stage of the algorithm:

**Lemma 3.4.** For any $p \in (0,1)$, with probability tending to 1 as $n \to \infty$, every 4-cycle whose every 1-face is heavy is contained in $V_3^n$.

We will introduce some additional notation for working with faces of $Q$. For two disjoint sets $U, W \subset [n]$, let $(U^*, W^1)$ denote the $|U|$-dimensional face of $Q$ with *s in the positions given by $U$, and 1s exactly in the positions given by $W$.

Using symmetry it will be enough to analyze the 4-cycle $(\{1,2\}^*, \emptyset^1)$. With the $M_p$ from (19), define $\mathcal{E}$ as the event that all the 1-faces in the 4-cycle $(\{1,2\}^*, \emptyset^1)$ are heavy, i.e.
\[
\mathcal{E} = \{ \deg((\{1\}^*, \emptyset^1)) > M_p, \deg((\{1\}^*, \{2\}^1)) > M_p, \deg((\{2\}^*, \emptyset^1)) > M_p, \deg((\{2\}^*, \{1\}^1)) > M_p \}.
\]

To prove Lemma 3.4, it suffices to show that

**Lemma 3.5.** For any $p \in (0,1)$, there is an $\epsilon > 0$ so that
\[
\mathbb{P}(\mathcal{E} \cap \mathcal{F} \cap \{\text{the degree of } (\{1,2\}^*, \emptyset^1) \text{ in } G(V_2^n) \text{ is less than } T_p\}) \leq n^{O(1)2^{-(1+\epsilon)n}}.
\]

*Proof.* The possible neighbors of $(\{1,2\}^*, \emptyset^1)$ in $G(V^n)$ all have the form $(\{1,2\}^*, \{j\}^1)$ for some $3 \leq j \leq n$. To have an edge between these 4-cycles in $G(V_2^n)$, we must have that
\[
(\{1,j\}^*, \emptyset^1) \in V_2^n, \quad (\{1,j\}^*, \{2\}^1) \in V_2^n, \\
(\{2,j\}^*, \emptyset^1) \in V_2^n, \quad (\{2,j\}^*, \{1\}^1) \in V_2^n.
\]

On the event $\mathcal{F}$, we must only lower bound the degree of these 4-cycles in $G(V_1^n)$ to ensure they are in $V_2^n$. Hence, define
\[
Y_{ij} = 1\{\deg((\{1,j\}^*, \emptyset^1)) \geq T_p\}, \quad Y_{2j} = 1\{\deg((\{1,j\}^*, \{2\}^1)) \geq T_p\}, \\
Y_{3j} = 1\{\deg((\{2,j\}^*, \emptyset^1)) \geq T_p\}, \quad Y_{4j} = 1\{\deg((\{2,j\}^*, \{1\}^1)) \geq T_p\}. \tag{20}
\]
The degree above refers to the degree of the 4-cycle in $G(V^n_1)$. We would like to show there are at least $T_p$ choices $j$ for which all $Y_{lj}$ for $l \in \{1, 2, 3, 4\}$ are 1.

On the event $\mathcal{E}$, there are 4 disjoint sets $R_\ell \subset \{3, 4, \ldots, d\}$ for $\ell \in \{1, 2, 3, 4\}$ of size $[M_p/4]$ so that

$$(\{1, k\}^*, 0^1) \in V_1^n, \quad (\{1, k\}^*, \{2\}^1) \in V_1^n, \quad (\{2, k\}^*, 0^1) \in V_1^n, \quad (\{2, k\}^*, \{1\}^1) \in V_1^n.$$  

Observe that the possible neighbors of $((\{1, j\}^*, \emptyset^1)$, for $j \in \{3, 4, \ldots, n\}$ are given by $(\{1, j\}^*, \{k\}^1)$ for $k \notin \{1, j\}$. For simplicity, we will also discard the case $k = 2$. To have this edge in $G(V^n_1)$, we would need that

$$(\{1, k\}^*, \emptyset^1) \in V_1^n, \quad (\{1, j\}^*, \{j\}^1) \in V_1^n, \quad (\{j, k\}^*, \emptyset^1) \in V_1^n, \quad (\{j, k\}^*, \{1\}^1) \in V_1^n.$$  

In particular, for $k \in R_1$, the first of these requirements is guaranteed. Hence we can define

$$Z_{1jk} = 1 \{(\{1, k\}^*, \{j\}^1) \in V_1^n, (\{j, k\}^*, \emptyset^1) \in V_1^n, (\{j, k\}^*, \{1\}^1) \in V_1^n\},$$  

and define

$$Z_{ij} = \sum_{k \in R_1} Z_{1jk}.$$  

Then $Z_{1j}$ is a lower bound for $\deg((\{1, j\}^*, \emptyset^1))$, and so if $Z_{1j}$ is at least $T_p$, then $Y_{1j} = 1$.

We do a similar construction for $\ell \in \{2, 3, 4\}$, making appropriate modifications. We list these for clarity below:

$$Z_{2jk} = 1 \{(\{1, k\}^*, \{2, j\}^1) \in V_1^n, (\{j, k\}^*, \{2\}^1) \in V_1^n, (\{j, k\}^*, \{1, 2\}^1) \in V_1^n\},$$  

$$Z_{3jk} = 1 \{(\{2, k\}^*, \{j\}^1) \in V_1^n, (\{j, k\}^*, \emptyset^1) \in V_1^n, (\{j, k\}^*, \{1\}^1) \in V_1^n\},$$  

$$Z_{4jk} = 1 \{(\{2, k\}^*, \{1, j\}^1) \in V_1^n, (\{j, k\}^*, \emptyset^1) \in V_1^n, (\{j, k\}^*, \{1, 2\}^1) \in V_1^n\}.$$  

In terms of these, we set $Z_{\ell j} = \sum_{k \in R_\ell} Z_{\ell jk}$. Let $J = \{3, 4, \ldots, d\} \setminus (\cup_\ell R_\ell)$. Then the family

$$\{Z_{\ell jk} : \ell \in \{1, 2, 3, 4\}, j \in J, k \in R_\ell\}$$  

are independent random variables. Moreover for any $\ell \in \{1, 2, 3, 4\}$ and $j \in J$, from (19),

$$\mathbb{P}(Z_{\ell j} < T_p) \leq \mathbb{P}(\text{Binomial}([M_p/4], p^3) < T_p) \leq \left(\frac{1}{2}\right)^{1/4}.$$  

20
It follows that with

\[ Z = \sum_{j \in J} \prod_{\ell=1}^{4} 1\{Z_{\ell j} \geq T_p\}, \]

and with \( q = \mathbb{P}(Z_{\ell j} \geq T_p)^4 > \frac{1}{2} \),

\[ \mathbb{P}(Z < T_p) \leq \mathbb{P}(\text{Binomial}(n - 3 - M_p, q) < T_p) = n^{O(1)}(1 - q)^n, \]

which completes the proof.

\[ \blacksquare \]

3.3 Stage 3: The final squeeze

In this section we draw conclusions on what remains non-contracted in the complex in the third stage.

3.3.1 The simply connected regime, \( p > \frac{1}{2} \)

We begin by showing that for \( p > 1/2 \), there are simply no light 1-faces. Hence in fact for \( p > \frac{1}{2} \), \( V_3^n = V_n \) with high probability (proving Theorem 3.1).

Lemma 3.6. For any \( p > 1/2 \), there is an \( \epsilon > 0 \) so that with probability tending to 1 with \( n \), for every 1-face \( f \) of \( Q \sim Q_2(n, p) \), \( \deg(f) > M_p \).

Proof. The degree of a 1-face is distributed as \( \text{Binomial}(n - 2, p) \). For \( p > \frac{1}{2} \), the probability this is less than any fixed constant \( M \) is \( n^{O(1)}(1 - p)^n \). Hence by a union bound, the lemma follows.

3.3.2 Completely shielded 1-faces

Call a 1-face \( f \in Q \sim Q_2(n, p) \) completely shielded if every 3-face \( c \in Q^n \) that contains \( f \) only contains heavy 1-faces of \( Q \), besides possibly \( f \). Completely shielded 1-faces modify the fundamental group of \( Q \) in a simple way, contributing exactly one free factor of \( \mathbb{Z} \) if \( f \) is isolated.

To see this we begin with the following definition:

Definition 3.7. Let \( f \) be any 1-face of \( Q^n \). Define the \( n \)-bubble around \( f \) to be the subcubical complex of \( Q^n \) given by the union of the complete 1-skeletons of all 3-faces containing \( f \), and every 2-face on this skeleton which does not contain \( f \).
A $n$-bubble has fundamental group $\mathbb{Z}$.

**Lemma 3.8.** For any $n \geq 3$, and any $n$-bubble $X$ around $f$,

$$\pi_1(X) \cong \mathbb{Z}.$$  

Furthermore, the complex $X \setminus \{f\}$ and the complex $X \cup \{e\}$, where $e$ is any 2-face containing $f$, are simply connected.

**Proof.** Without loss of generality, suppose that $f$ is the face ($\{1\}^*, \emptyset^1$). The 3-faces containing $f$ all have the form ($\{1, i, j\}^*, \emptyset^1$), and so the 1-skeleton of $X$ is

$$\{\{(i)^*, A^1) : A \subset \{1, 2, \ldots, n\}, i \notin A, |A \cup \{i\}| \leq 3\}.$$  

We claim that all the 4-cycles containing $f$ are homotopic. As all other 4-cycles are contractible from the definition of $X$, the statements in the lemma follow.

The 4-cycles that contain $f$ are boundaries of the 2-faces of $Q^n$ of the form

$$\{\{(1)^*, \emptyset^1) : 2 \leq i \leq n\}.$$  

For any $2 \leq i < j \leq n$, the 3-face $c = (\{1, i, j\}^*, \emptyset)$ intersected with $X$ contains 4 2-faces. Moreover, the 2-faces ($\{1, i\}^*, \emptyset$) and ($\{1, j\}^*, \emptyset$) are adjacent in this cube. Hence, these cycles can be deformed through $c$ to one another. As this held for any such $i$ and $j$, the proof follows.

**Lemma 3.9.** For any $p \in (0, 1)$, let $\hat{Q}$ be the cubical complex that results from deleting from $Q \sim Q_2(n, p)$ every completely shielded 1-face $f$ and any 2-face of $Q$ containing $f$. Then with high probability,

$$\pi_1(Q) \cong \pi_1(\hat{Q}) \ast \underbrace{(\mathbb{Z} \ast \mathbb{Z} \ast \cdots \ast \mathbb{Z})}_{N}$$  

where $N$ denotes the number of completely shielded 1-faces in $Q$ that are isolated.

**Proof.** From Lemma 3.4, all 4-cycles whose every 1-face is heavy are contractible. In particular we do not modify the fundamental group of $Q$ if we include all those 2-faces into $Q$ whose boundary is in $V_3^n$. Let $\tilde{Q}$ be this cube complex.

We now remove completely shielded 1-faces from $\tilde{Q}$ one at a time, tracking the changes to the fundamental group. We will show what happens after
removing the first. It will be clear that by using induction, a similar analysis would give the claim in the lemma.

Let \( f \) be a completely shielded 1-face of \( \tilde{Q} \). Let \( Q_1 \) be the complex that results after removing \( f \) from \( \tilde{Q} \) and any 2-face containing \( f \). Let \( Q_2 \) be the union of all the complete 2-skeletons of all 3-faces that contain \( f \). Then \( Q_2 \) contains a \( n \)-bubble, and it is exactly a \( n \)-bubble if \( f \) is isolated.

As \( Q_2 \cup Q_1 = \tilde{Q} \) and \( Q_1 \cap Q_2 \) is open and path connected (c.f. Lemma 3.8, as this complex is a \( n \)-bubble with its central 1-face deleted). Moreover, every 4-cycle in \( Q_1 \cap Q_2 \) is contractible, and so \( \pi_1(Q_1 \cap Q_2) \) is trivial. From the Siefert-van Kampen theorem, we therefore have that

\[
\pi_1(\tilde{Q}) \cong \pi_1(Q_1) \ast \pi_1(Q_2).
\]

If \( f \) is isolated then from Lemma 3.8, the fundamental group \( \pi_1(Q_2) \) is isomorphic \( \mathbb{Z} \). \( \blacksquare \)

### 3.3.3 The velvety bubble phase

For \( p > 1 - (\frac{1}{2})^{1/2} \approx 0.292893 \), we further show that the fundamental group completely reduces to its isolated 1-faces. In this phase, while light 1-faces may exist in \( Q_2(n,p) \) (for \( p \leq \frac{1}{2} \)), they are well separated.

**Lemma 3.10.** For \( p > 1 - (\frac{1}{2})^{1/2} \), with high probability, there are no 3-faces \( c \in Q^n \) that contain more than one light 1-face of \( Q \sim Q_2(n,p) \).

**Proof.** For a fixed \( c \) and a fixed choice of two 1-faces \( f_1, f_2 \), for the degrees of \( f_1 \) and \( f_2 \) are both light with probability at most \( n^{O(1)}(1-p)^{2n-1} \). Hence for any \( p \) as in the statement of the lemma, there is an \( \epsilon > 0 \) so that the probability this occurs is \( 2^{-(1+\epsilon)n+O(\log n)} \). As there are \( 2^n n^{O(1)} \) many ways to pick a 3-face with two designated edges, the lemma follows from a first moment estimate. \( \blacksquare \)

We now give the proof of Theorem 1.3, which we recall for convenience.

**Theorem 3.11.** For \( p > 1 - (\frac{1}{2})^{1/2} \), with high probability, for \( Q \sim Q_2(n,p) \)

\[
\pi_1(Q) \cong (\mathbb{Z} \ast \mathbb{Z} \ast \cdots \ast \mathbb{Z}),
\]

where \( N \) denotes the number of isolated 1-faces in \( Q \).
Proof. From Lemma 3.4, with high probability every 4-cycle containing only heavy 1-faces is in $V^3_n$. From Lemma 3.10, with high probability no 3-faces $c \in Q^n$ contain more than one light face. Hence taking $\tilde{Q}$ as $Q$ together with all 2-faces bounded by some element of $V^3_n$ (so that $\pi_1(\tilde{Q}) = \pi_1(Q)$) every light 1-face $f$ of $\tilde{Q}$ is completely shielded in $\tilde{Q}$. Moreover, every 4-cycle of $\tilde{Q}$ either intersects a light 1-face, or it is the boundary of a 2-face. Hence in the notation of Lemma 3.9, $\pi_1(\tilde{Q}) = 0$. It follows that from Lemma 3.9 is a free group on $N'$ generators, with $N'$ the number of completely shielded isolated 1-faces. As every light 1-face is completely shielded w.h.p, it follows that $N' = N$ with high probability. $\blacksquare$

4 Structure theorem for general $p$

In this section we prove Theorem 1.5. For convenience, we recall some definitions from the introduction. Recall Definition 1.4:

**Definition 4.1.** For a cubical sub-complex $T$ of any cube $Q^n$, define its *edge complexity* $e(T)$ as the number of edges in $T$. Let $\mathcal{T}_p$ be the set of pure 2-dimensional strongly connected cubical complexes $T$ that are subcomplexes of $Q^n_2$ for some $n$ and so that $(1 - (\frac{1}{2})^{1/e(T)}) < p$.

We will prove Theorem 1.5, which we recall below:

**Theorem 4.2.** For any $p \in (0, 1)$, and for $Q \sim Q_2(n, p)$, let the free product decomposition of $\pi_1(Q)$ be given by

$$\pi_1(Q) \cong F \ast \pi_1(X_1) \ast \pi_1(X_2) \ast \cdots \ast \pi_1(X_\ell),$$

with $F$ a free group. With high probability, any $T \in \mathcal{T}_p$ appears as a factor $\pi_1(X_j)$ for some $1 \leq j \leq \ell$.

Our main technical tool will be the following:

**Definition 4.3.** For a cubical sub-complex $T$ of a cubical complex $W \subset Q^n$ denote by $h(T)$ the minimal cubical sub-complex of $W$ so that

1. the 1–skeleton of $h(T)$ is the 1–skeleton of a $k$–dimensional hypercube
2. every 2–face of $W$ that is incident to $T$ is contained in $h(T)$

24
3. every 2–face of \( Q^n \) with 1–skeleton in \( h(T) \) which is not incident to \( T \) is in \( h(T) \)

Also denote by \( H(T) \subset W \) as a complete 2–skeleton of a \( k \)–dimensional hypercube which is parallel to \( h(T) \) so that any 2–face that has an edge in \( h(T) \setminus T \) and another edge in \( H(T) \) is contained in \( W \).

We emphasize that \( T \) need not be connected in any sense.

**Lemma 4.4.** For any \( p \in (0, 1) \), with \( Q \sim Q_2(n, p) \) there exists a number \( k_p \) so that whp every \(((M_p + 2) \times k_p)^2\)-dimensional cube in \( Q^n \) contains fewer than \( k_p \) light edges of \( Q \).

**Proof.** We argue by a first moment estimate. For any \( \ell \), the number of \( \ell \)–dimensional cubes in \( Q^n \) is given by \( 2^{n-\ell} \binom{n}{\ell} \). The probability that any such a cube contains \( k \) light edges is \( O(k^{2\ell}p^{n\ell}) \). Hence taking \( \ell = ((M_p + 2)k)^{2} \), if we pick \( k \) sufficiently large that \((1-p)^k < \frac{1}{2}\), then the expected number of \( \ell \)–dimensional cubes containing more than \( k \) light edges tends to 0 exponentially in \( n \). □

**Lemma 4.5.** Let \( p \in (0, 1) \) and let \( \ell \in \mathbb{N} \) be fixed, then whp for \( Q \sim Q_2(n, p) \), every \( \ell \)-dimensional cube \( X \) has a parallel cube \( Y \) that has no light edges in \( Q \) and for which there are no light edges in \( Q \) between \( X \) and \( Y \).

**Proof.** This is similar to Lemma 4.4. We argue by a first moment estimate. For any \( \ell \), the number of \( \ell \)–dimensional cubes in \( Q^n \) is given by \( 2^{n-\ell} \binom{n}{\ell} \). For a fixed \( \ell \)–dimensional cube \( X \subset Q^n \), the probability that every parallel \( \ell \)–dimensional cube \( Y \) either

(i) contains at least one light edge or

(ii) contains the endpoint of a light edge between \( X \) and \( Y \)

is at most

\[ ((\ell + 2)^{2-1})^{2-\ell} (1-p)^{\ell} \binom{n-\ell}{\ell} = o(n^{\ell}2^{-n}). \]

Hence from a first moment estimate, for any fixed \( \ell \) and for any \( p \in (0, 1) \), whp every \( \ell \)–dimensional cube \( X \) has a parallel cube \( Y \) that contains no light edges of \( Q \) and shares no endpoint of a light edge between \( X \) and \( Y \). □

**Theorem 4.6.** Let \( p \in (0, 1) \) and \( k_p \) as in Lemma 4.4. Let \( Q \sim Q_2(n, p) \). Let \( \overline{Q} \) be \( Q \) with all the 2–faces bounded by 4–cycles having no light edges. Whp, there are disjoint cubical complexes \( \{\tau_1, \tau_2, \ldots, \tau_\ell\} \) in \( Q \) so that
(i) the union of 1-faces over all \( \{ \tau_j : 1 \leq j \leq \ell \} \) is the set of all light 1-faces,

(ii) for each \( 1 \leq j \leq \ell \), both \( h(\tau_j) \) and \( H(\tau_j) \) exist in \( \overline{Q} \),

(iii) for each \( 1 \leq i \neq j \leq \ell \), the Hamming distance between the 0-skeleta of \( h(\tau_j) \) and \( h(\tau_i) \) is at least 2.

We need the next definition for proving Theorem 4.6.

**Definition 4.7.** Let \( X \) and \( Y \) be two subcomplexes of \( Q^n \). Define \( X \Box Y \) to be the face of smallest dimension \( Q^m \) such that \( Q^m \subset Q^n \) and \( X \cup Y \subset Q^m \). Observe that \( m \leq n \). More in general, let \( X_1, X_2, \ldots, X_l \) be any finite collection of subcomplexes of \( Q^n \) and \( I = \{1, 2, \ldots, l\} \). We define

\[ \Box_{i \in I} X_i \]

as the face of smallest dimension \( Q^m \) such that \( Q^m \subset Q^n \) and such that

\[ [X_1 \cup X_2 \ldots \cup X_l] \subset Q^m. \]

In this case, \( m \leq n \) as well.

**Proof of Theorem 4.6.** We first show that for every light 1-face \( e \) there is a cubical complex \( \sigma_e \) containing \( e \) and having all its 1-faces light so that \( h(\sigma_e) \) exists in \( \overline{Q} \). We will then merge these \( h(\sigma_e) \) to form the partition claimed to exist in the theorem.

Let \( e_1 := e \) be any light edge of \( Q \). Let \( T_1 \) be the cube complex which is the down closure of \( e_1 \). Let \( X_1 \) be the smallest induced complex in \( \overline{Q} \) which contains \( T_1 \), which contains all 2-faces of \( \overline{Q} \) incident to \( e_1 \) and whose 1-skeleton is a hypercube. If \( X_1 = h(T_1) \), we are finished. Otherwise, by definition, there must be a 2-face \( f \) of \( Q^n_2 \) with 1-skeleton in \( X_1 \) but which is not itself in \( X_1 \). Then, there must be at least one light edge \( e_2 \in X_1 \setminus T_1 \).

We then define \( T_2 \) as the induced subcomplex of \( Q \) on edges \( e_1, e_2 \). Let \( X_2 \) be the smallest induced complex in \( \overline{Q} \) which contains \( T_2 \), which contains all 2-faces of \( \overline{Q} \) incident to \( T_2 \) and whose 1-skeleton is a hypercube. Once more, if \( X_2 = h(T_2) \), we are done. Otherwise, we proceed inductively by the same argument.

This produces a nested sequence of complexes \( \{T_k\} \) each having \( k \) edges. It also produces a sequence of complexes \( \{X_k\} \) such that each \( X_k \supset T_k \),
each $X_k$ contains at least $k$ light edges, and such that $X_k$ has the 1-skeleton of a hypercube of dimension at most $k \times M_p$. By Lemma 4.4, with high probability, this sequence must terminate at some $k^* \leq k_p$. The complex $X_{k^*} = h(T_{k^*})$ by definition, and we define $\sigma_e = T_{k^*}$.

We define a graph $G$ with vertex set given by the collection of $\sigma_e$. Two vertices $\sigma_{e_1}, \sigma_{e_2}$ in this graph are connected if the hamming distance between $h(\sigma_{e_1})$ and $h(\sigma_{e_2})$ is less than two. Let $\{\tau_1, \tau_2, \ldots, \tau_\ell\}$ be the unions of the connected components in $G$. Then for each $1 \leq j \leq \ell$, we construct the hypercube

$$\Sigma_j = \Box_{e \in \tau_j} h(\sigma_e).$$

It is easy to see that $\Sigma_j = h(\tau_j)$, which implies that $h(\tau_j)$ exists and is exactly $\Sigma_j$.

The dimension of $\Sigma_j$ is at most

$$\sum_{\sigma_e} \dim(h(\sigma_e)) + 2,$$

where the sum is over all $\sigma_e$ contained in $\tau_j$. Therefore by Lemma 4.4, each $\tau_j$ has at most $k_p$ edges. Hence by Lemma 4.5, each $H(\tau_j)$ exists as well. ■

**Lemma 4.8.** Let $W$ be a subcomplex of $Q_2^n$ and $T$ a subcomplex of $W$. Suppose that $h(T)$ and $H(T)$ exist in $W$. Let $\hat{W}$ be the complex formed by adding to $W$ the complete 1-skeleton of $h(T) \Box H(T)$ and any 2-face of $Q_2^n$ with 1-skeleton in $h(T) \Box H(T)$. Then

$$\pi_1(W) \cong \pi_1(\hat{W}) \ast \pi_1(W \cap (h(T) \Box H(T))).$$

Recall that for a disconnected cube complex $X$, we define $\pi(X)$ as the free product of its connected components.

**Proof.** Let $\hat{T}$ be all the 1-faces in $T$ and any 2-face of $W$ incident to $T$. Let $\hat{T}$ be the down closure of $\hat{T}$. Let $S = W \cap (h(T) \Box H(T))$.

Let $X = (W \setminus \hat{T}) \cap S$. We claim that $\pi_1(X) \cong 1$. The 1-faces of $X$ that are in $h(T)$ are not in $T$. Therefore, by Definition 4.3, for every edge $e \in X \cap h(T)$, the unique 4-cycle connecting $e$ to $H(T)$ is the boundary of a 2-face in $X$. Hence, every closed curve in $X$ is homotopic to a curve in $H(T)$. Since $\pi_1(H(T)) \cong 1$, it follows that $\pi_1(X) \cong 1$. Therefore, the Siefert-van Kampen theorem states that

$$\pi_1(W) \cong \pi_1(W \setminus \hat{T}) \ast \pi_1(S).$$

27
We now show that $\pi_1(W \setminus \hat{T}) \cong \pi_1(\hat{W})$. Define a complex $S^*$ as the down closure of all 2-faces in $Q_2^n$ incident to $T$ with 1-skeleton in $h(T) \square H(T)$, union with $H(T)$. Any 1-face $e$ of $S^* \cap (W \setminus \hat{T})$ that is in $h(T)$ must be in $h(T) \setminus T$. In particular, there is a 2-face $f$ containing $e$ which has a 1-face in $H(T)$. Hence, any closed curve in $S^* \cap (W \setminus \hat{T})$ is homotopic to one in $H(T)$, which is simply connected. Therefore by the Siefert-van Kampen theorem,

$$\pi_1(\hat{W}) = \pi_1(S^* \cup (W \setminus \hat{T})) \cong \pi_1(W \setminus \hat{T}) \ast \pi_1(S^*).$$

It remains to evaluate the fundamental group of $S^*$. Any edge in $S^* \cap h(T)$ has a 4-cycle that has an edge in $H(T)$. By construction we know that this 4-cycle has a 2-face added. Therefore any closed curve in $S^*$ is homotopic to a closed curve in $H(T)$. Thus $S^*$ is simply connected because $H(T)$ is by definition.

**Lemma 4.9.** Let $W$ be a subcomplex of $Q_2^n$ and $T$ a subcomplex of $W$. Suppose that $h(T)$ and $H(T)$ exist in $W$. Let $P_1, \ldots, P_m$ be all the pure 2-dimensional strongly connected components completely contained in $T$, such that any 2-face adjacent to the 1-skeleton of any $P_i$ is also contained in $P_i$. Suppose that $T = \bigcup_{i=1}^m P_i$. Then there is a free group $F$ so that

$$\pi_1(W \cap (h(T) \square H(T))) \cong \pi_1(P_1) \ast \pi_1(P_2) \ast \cdots \ast \pi_1(P_m) \ast F.$$ 

**Proof.** Let $S = W \cap (h(T) \square H(T))$. Suppose we fill $H(T)$ by taking the flag cubical complex of $H(T)$. The fundamental group of $S$ is unchanged and we can contract $H(T)$ to a point $x$. We denote this complex by $\hat{S}$. If $e$ is an edge in $T$ then $e$ forms an unfilled triangle with $x$ in $\hat{S}$. Let $T_x \subset \hat{S}$ be the union of $T$, $x$ and all the edges between $T$ and $x$. Any edge $f \in h(T)$ which is not contained in $T_x$ is the base of a filled triangle with $x$ in $\hat{S}$, and so any closed curve in $\hat{S}$ is homotopic to a closed curve in $T_x$. Hence

$$\pi_1(S) = \pi_1(\hat{S}) = \pi_1(T_x) = \pi_1(P_1) \ast \pi_1(P_2) \ast \cdots \ast \pi_1(P_m) \ast F$$

where $F$ is a free group.

**Theorem 4.10.** Fix a $p \in (0,1)$. For a $Q \sim Q_2(n, p)$, whp if $\tau_1, \tau_2, \ldots, \tau_\ell$ are as constructed in Theorem 4.6, then with $S_j = C \cap (h(\tau_j) \square H(\tau_j))$ for all $1 \leq j \leq \ell$,

$$\pi_1(Q) \cong \pi_1(S_1) \ast \pi_1(S_2) \ast \cdots \ast \pi_1(S_\ell).$$
Proof. Let $\overline{Q}$ be $Q$ with all the 2-faces bounded by 4-cycles having no light edges. By Lemma 3.4, all 4-cycles with no light edges are in $V_3^n$ whp, and so $\pi_1(\overline{Q}) = \pi_1(Q)$. We apply Lemma 4.8 inductively to each of the complexes $\tau_j$. As a result, we have that

$$\pi_1(\overline{Q}) = \pi_1(J) \ast \pi_1(S_1) \ast \pi_1(S_2) \ast \cdots \ast \pi_1(S_\ell),$$

where $J$ is the complex $\overline{Q}$ together with all 2-faces in $Q_2^n$ having 1-skeleton contained in some $h(\tau_j) \Box H(\tau_j)$ for some $1 \leq j \leq \ell$.

It just remains to prove that $\pi_1(J) \cong 1$. The 1-skeleton of $J$ is $Q_1^n$, and so it suffices to show that every 4-cycle in $J$ is contractible. The only 4-cycles $x$ in $J$ that do not bound a 2-face are those that contain a 1-face $e$ of some $\tau_j$ for $1 \leq j \leq \ell$ but which were not contained in $h(\tau_j) \Box H(\tau_j)$. However, as $e$ is in $h(\tau_j)$, it has a parallel 1-face $f$ in $H(\tau_j)$. In the unique cube $c = x \Box e$ that contains $x$ and $e$ has all 2-faces except for the face bounded by $x$ which implies that $x$ is contractible. ■

Proof of Theorem 1.5. Let $T \in \mathcal{T}$ be fixed. By assumption there is a $k$-dimensional cube $X$ so that $T$ is a subcomplex of $X$, and we may choose $k$ minimal. We do not take the full 2-skeleton for $X$, but instead we choose exactly those 2-faces which are either in $T$ or share no edge with $T$.

Let $\phi$ be a cubical embedding of the 2-skeleton of $X$ into $Q_2^n$. Define the event $E_\phi$, for $Q \sim Q_2^n(\cdot, \cdot)$:

1. The 2-faces of $Q$ that are contained in the 1-skeleton of $\phi(X)$ are exactly the 2-faces of $\phi(X)$.
2. No other 2-face in $Q$ contains a 1-face of $\phi(T)$.
3. There are no light 1-faces in $\phi(X \setminus T)$ and no light 1-faces within Hamming distance $2k_p + 2$ of $\phi(T)$, except possibly those in $\phi(T)$. Here, $k_p$ is defined as in Lemma 4.4.

We now estimate the probability of $E_\phi$ under the law of $Q_2^n(\cdot, \cdot)$. Note that this probability does not depend on $\phi$, and so these estimates will be uniform in $\phi$. First, observe that each edge of $\phi(T)$ has degree bounded independently of $n$ on this event, and so there are $(e(T) * d) - O(1)$ 2-faces which must be absent for $E_\phi$ to hold. There are $O(1)$ 2-faces that must be present for $E_\phi$ to hold, also. There are also $O(d^{2k_p + 2})$ 1-faces which are contained in the Hamming distance $(2k_p + 2)$-neighborhood of $\phi(X)$ which we
require to be not light. As the probability that a 1-face is light is \( O((1 - p)^n) \), we conclude that

\[
P(\mathcal{E}_\phi) = \Theta((1 - p)^e(T)d) = \Omega(2^{-(1-\epsilon)T}),
\]

for some \( \epsilon > 0 \), where the second equality follows from Definition 1.4. So the expected number of occurrences of \( \mathcal{E}_\phi \) goes to infinity exponentially fast as \( n \to \infty \).

We can now show that some \( \mathcal{E}_\phi \) now occurs with high probability by using a second moment computation (see Corollary 4.3.5 of [1]). Observe that if the Hamming distance of \( \phi(X) \) to \( \psi(X) \) is greater than 4, then the events \( \mathcal{E}_\phi \) and \( \mathcal{E}_\psi \) are independent. Let \( \psi \sim \phi \) if \( \mathcal{E}_\phi \) and \( \mathcal{E}_\psi \) are not independent. Then,

\[
\Delta^*_\phi = \sum_{\psi \sim \phi} P[\mathcal{E}_\psi|\mathcal{E}_\phi] \leq \sum_{\psi \sim \phi} 1 = O(d^{O(1)}),
\]

which is much smaller than the expected number of \( \mathcal{E}_\phi \) that occur (which grows exponentially in \( n \)).

Hence, with the decomposition given by Theorem 4.10,

\[
\pi_1(Q) \cong \pi_1(S_1) \ast \pi_1(S_2) \ast \cdots \ast \pi(S_\ell),
\]

where \( S_j = Q \cap (h(\tau_j) \square H(\tau_j)) \) and where \( \tau_j \) are the complexes from Theorem 4.6. For any embedding \( \phi \), if \( \mathcal{E}_\phi \) occurs, then \( \phi(X) \in Q \) is such that \( \phi(X) = h(\phi(T)) \). Moreover, \( \phi(T) = \tau_j \) for some \( j \) with \( 1 \leq j \leq \ell \) as the Hamming distance of \( \phi(T) \) to any other light 1-face is at least \( 2k_p + 2 \). By Lemma 4.9, \( \pi_1(S_j) \cong F \ast \pi_1(T) \) for some free group \( F \).

5 Below the threshold for isolated edges

Corollary 5.1. Let

\[
T_2 = \{(0,*,0,*),(*,*,0,1),(*,0,*,1),(*,0,1,*),(0,*,1,*),(0,*,*,0)\},
\]

and \( T \) the down closure of \( T_2 \) which is a strongly connected pure 2-dimensional subcomplex of \( Q_2^4 \) and has \( e(T) = 18 \). Then \( \pi_1(T) \cong \mathbb{Z}/(2\mathbb{Z}) \); in particular it has torsion. Hence, for \( p \neq 0 \), if \( p < (1 - (1/2)^{1/18}) \approx 0.037776 \), the fundamental group of \( Q \sim Q_2(n,p) \) has a torsion group in its free product decomposition with high probability.
Proof of Corollary. We depict $T$ in Figure 1. Let $\{a, b, c, \ldots, k, l\}$ be the vertices of $T$. The vertex $a$ is $(0,0,0,1)$. The vertex $b$ is $(0,1,0,1)$. The remainder can be determined from adjacency, using Figure 1. All edges in the figure are distinct, save for the edge $\overline{ab} = (0,*0,1)$, which is depicted twice. It is clear that $\pi_1(T) \cong \mathbb{Z}/(2\mathbb{Z})$, and that $e(T) = 18$. The corollary now follows from Theorem 1.5.

![Figure 1: Observe that $a = (0,0,0,1)$, $b = (0,1,0,1)$, and $\overline{ab} = (0,*0,1)$.](image)

Corollary 5.2. For $0 < p < (1 - (1/2)^{1/5}) \approx 0.021428$, $\pi_1(Q)$ for $Q \sim Q_2(n,p)$ has a $\mathbb{Z} \times \mathbb{Z}$ factor with high probability.

Proof. Let $T_1$ be the subcomplex of $Q_2^4$ depicted in Figure 2. Observe that $T_1$ is a strongly connected pure 2-dimensional subcomplex of $Q_2^4$ and has $e(T_1) = 32$. From Figure 2, it is easy to see that $\pi_1(T_1) \cong \mathbb{Z} \times \mathbb{Z}$. Hence, from Theorem 1.5, for $0 < p \neq 0$, if $p < (1 - (1/2)^{1/32})$, the fundamental group of $Q \sim Q_2(n,p)$ has a copy of $\mathbb{Z} \times \mathbb{Z}$ in its free product decomposition with high probability.

References

[1] N. Alon and J. H. Spencer. The probabilistic method. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, fourth edition, 2016. ISBN 978-1-119-06195-3.

[2] L. Aronshtam, N. Linial, T. Łuczak, and R. Meshulam. Collapsibility and vanishing of top homology in random simplicial complexes. Discrete & Computational Geometry, 49(2):317–334, 2013.
Figure 2: Here we depict the subcomplex $T_1 \subset Q_4^2$. It has 16 2-dimensional faces, 32 1-dimensional faces, and 16 0-dimensional faces. The labeled 1-dimensional faces are: $a = (1,0,0,0), b = (0,0,0,0), c = (0,1,0,0), d = (1,1,0,0), e = (1,0,1,0), f = (1,0,1,1)$, and $g = (1,0,0,1)$. Observe that the edges that are in the boundary of the figure are identify in such a way that $T_1$ creates a torus in $Q_4^2$.

[3] E. Babson, C. Hoffman, and M. Kahle. The fundamental group of random 2-complexes. *J. Amer. Math. Soc.*, 24(1):1–28, 2011. ISSN 0894-0347. doi: 10.1090/S0894-0347-2010-00677-7. URL http://dx.doi.org/10.1090/S0894-0347-2010-00677-7.

[4] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001. ISBN 0-521-80920-7; 0-521-79722-5. doi: 10.1017/CBO9780511814068. URL https://doi.org/10.1017/CBO9780511814068.

32
[5] J. D. Burtin. The probability of connectedness of a random subgraph of an n-dimensional cube. *Problemy Peredači Informacii*, 13(2):90–95, 1977. ISSN 0555-2923. (Russian).

[6] D. Cohen, A. Costa, M. Farber, and T. Kappeler. Topology of random 2-complexes. *Discrete Comput. Geom.*, 47(1):117–149, 2012. ISSN 0179-5376. doi: 10.1007/s00454-011-9378-0. URL https://doi.org/10.1007/s00454-011-9378-0.

[7] A. Costa and M. Farber. Large random simplicial complexes, I. *J. Topol. Anal.*, 8(3):399–429, 2016. ISSN 1793-5253. doi: 10.1142/S179352531650014X. URL https://doi.org/10.1142/S179352531650014X.

[8] A. Costa and M. Farber. Large random simplicial complexes, II; the fundamental group. *J. Topol. Anal.*, 9(3):441–483, 2017. ISSN 1793-5253. doi: 10.1142/S1793525317500170. URL https://doi.org/10.1142/S1793525317500170.

[9] A. Costa and M. Farber. Large random simplicial complexes, III: the critical dimension. *J. Knot Theory Ramifications*, 26(2):1740010, 26, 2017. ISSN 0218-2165. doi: 10.1142/S0218216517400107. URL https://doi.org/10.1142/S0218216517400107.

[10] A. E. Costa and M. Farber. The asphericity of random 2-dimensional complexes. *Random Structures Algorithms*, 46(2):261–273, 2015. ISSN 1042-9832. doi: 10.1002/rsa.20499. URL https://doi.org/10.1002/rsa.20499.

[11] K. Dowling and E. Lundberg. Homotopy Types of Random Cubical Complexes. *arXiv e-prints*, art. arXiv:1910.12803, Oct 2019.

[12] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

[13] P. Erdős and J. Spencer. Evolution of the n-cube. *Comput. Math. Appl.*, 5(1):33–39, 1979. ISSN 0898-1221. doi: 10.1016/0898-1221(81)90137-1. URL https://doi-org.proxy.lib.ohio-state.edu/10.1016/0898-1221(81)90137-1.
[14] M. Farber and T. Nowik. Topological embeddings into random 2-complexes. *arXiv e-prints*, art. arXiv:1912.03939, Dec 2019.

[15] S. Hart. A note on the edges of the n-cube. *Discrete Mathematics*, 14 (2):157–163, 1976.

[16] A. Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002. ISBN 0-521-79160-X; 0-521-79540-0.

[17] Y. Hiraoka and T. Shirai. Tutte polynomials and random-cluster models in Bernoulli cell complexes. *arXiv e-prints*, art. arXiv:1602.04561, Feb 2016.

[18] Y. Hiraoka and K. Tsunoda. Limit theorems for random cubical homology. *Discrete Comput. Geom.*, 60(3):665–687, 2018. ISSN 0179-5376. doi: 10.1007/s00454-018-0007-z. URL https://doi-org.proxy.lib.ohio-state.edu/10.1007/s00454-018-0007-z.

[19] C. Hoffman, M. Kahle, and E. Paquette. Spectral Gaps of Random Graphs and Applications. *International Mathematics Research Notices*, 05 2019. ISSN 1073-7928. doi: 10.1093/imrn/rnz077. URL https://doi.org/10.1093/imrn/rnz077.

[20] M. Kahle, F. H. Lutz, A. Newman, and K. Parsons. Cohen–lenstra heuristics for torsion in homology of random complexes. *Experimental Mathematics*, 2018.

[21] A. V. Kostochka, A. A. Sapozhenko, and K. Weber. On random cubical graphs. In *Fourth Czechoslovakian Symposium on Combinatorics, Graphs and Complexity* (Prachatice, 1990), volume 51 of *Ann. Discrete Math.*, pages 155–160. North-Holland, Amsterdam, 1992. doi: 10.1016/S0167-5060(08)70621-6. URL https://doi-org.proxy.lib.ohio-state.edu/10.1016/S0167-5060(08)70621-6.

[22] N. Linial and R. Meshulam. Homological connectivity of random 2-complexes. *Combinatorica*, 26(4):475–487, 2006. ISSN 0209-9683. doi: 10.1007/s00493-006-0027-9. URL https://doi-org.proxy.lib.ohio-state.edu/10.1007/s00493-006-0027-9.

[23] N. Linial and Y. Peled. On the phase transition in random simplicial complexes. *Ann. of Math. (2)*, 184(3):745–773,
[24] T. Łuczak and Y. Peled. Integral homology of random simplicial complexes. *Discrete & Computational Geometry*, 59(1):131–142, 2018.

[25] Z. Luria and Y. Peled. On simple connectivity of random 2-complexes. *arXiv preprint arXiv:1806.03351*, 2018.

[26] R. Meshulam and N. Wallach. Homological connectivity of random $k$-dimensional complexes. *Random Structures Algorithms*, 34(3):408–417, 2009. ISSN 1042-9832. doi: 10.1002/rsa.20238. URL http://dx.doi.org/10.1002/rsa.20238.

[27] A. Newman. Freeness of the random fundamental group. *Journal of Topology and Analysis*, pages 1–7, 2018.

[28] A. Newman and E. Paquette. The integer homology threshold in $Y_d(n, p)$. *arXiv e-prints*, art. arXiv:1808.10647, Aug. 2018.