THE EVOLUTION OF THE PERMUTAHEDRON

MAURÍCIO COLLARES, JOSEPH DOOLITTLE, AND JOSHUA ERDE

ABSTRACT. In their seminal paper introducing the theory of random graphs, Erdős and Rényi considered the evolution of the structure of a random subgraph of K_n as the density increases from 0 to 1, identifying two key points in this evolution – the *percolation threshold*, where the order of the largest component seemingly jumps from logarithmic to linear in size, and the *connectivity threshold*, where the subgraph becomes connected. Similar phenomena have been observed in many other random graph models, and in particular, works of Ajtai, Komlós and Szemerédi and of Spencer and Erdős determine corresponding thresholds for random subgraphs of the hypercube.

We study similar questions on the *permutahedron*. The permutahedron, like the hypercube, has many different equivalent representations, and arises as a natural object of study in many areas of combinatorics. In particular, as a highly-symmetric simple polytope, like the n-simplex and n-cube, this percolation model naturally generalises the Erdős-Rényi random graph and the percolated hypercube.

We determine the percolation threshold and the connectivity threshold for random subgraphs of the permutahedron. Along the way we develop a novel graph exploration technique which can be used to find exponentially large clusters after percolation in high-dimensional geometric graphs and we initiate the study of the isoperimetric properties of the permutahedron.

1. INTRODUCTION

Given a graph G and a probability $p \in (0,1)$ the *percolated random subgraph* G_p is the random subgraph of G obtained by including each edge of G independently with probability p. In a sense, G_p represents the *typical* structure of a subgraph of G with relative density p. This model was first considered in the context of statistical physics, where it was used to model the flow of liquid through a porous medium whose channels could be randomly blocked [17]. There is a natural way to couple the random subgraphs G_p for $p \in [0,1]$ so that they are increasing (as subgraphs) as p increases, and in this way we can think of the structure of G_p as *evolving* as we increase p from 0 to 1, with G_p gradually growing from an empty graph to the full *host* graph G.

Perhaps the simplest percolation model is when we take the host graph G to be a complete graph K_n, in which case we recover the classical *binomial random graph model* $G(n,p)$. The evolution of the structure of $G(n,p)$ was the topic of one of the earliest papers on random graphs, in which Erdős and Rényi [32] showed that the component structure of $G(n,p)$[1] undergoes a dramatic phase transition when p is around $\frac{1}{n}$. More precisely, given a constant $c > 1$, let us define $\gamma(c)$ to be the unique solution in $(0,1)$ of the equation

$$\gamma = 1 - \exp(-c\gamma). \tag{1}$$

Theorem 1.1 ([32]). *Let $c > 0$ be a constant and let $p = \frac{c}{n}$. Then, with high probability*[2] *(whp)*

(i) if $c < 1$, then the largest component of $G(n,p)$ has order $\Theta(\log n)$; and,

---

[1]In fact, their result was stated in terms of the *uniform* random graph model $G(n,m)$.

[2]with probability tending to one as $n \to \infty$.

(ii) *if $c > 1$, then there exists a unique* giant *component in $G(n,p)$ of order $(1+o(1))\gamma(c)n$, where γ is defined according to* (1)*, and the second largest component has order $\Theta(\log n)$.*

We note that Theorem 1.1 is only a very broad description of the phase transition, and much more is known about the structure of $G(n,p)$ when p is close to the critical point (see, for example, [14, 48, 49]).

It turns out that $G(n,p)$ is in some sense partially *universal* in this class of percolation models, in the sense that for many n-regular[3] host graphs G, the component structure of the percolated subgraph G_p undergoes a quantitatively similar phase transition as that described in Theorem 1.1 when p is around $\frac{1}{n}$. In these models, the broad scale structure of G_p, under the right scaling, seems to be independent of the underlying geometry of the host graph G. A particularly notable example of this phenomenon was shown to occur in the n-dimensional hypercube Q^n by Ajtai, Komlós and Szemerédi [2], answering a well-known conjecture of Erdős and Spencer [33].

Theorem 1.2 ([2, 15])**.** *Let $c > 0$ be a constant and let $p = \frac{c}{n}$. Then, whp*

(i) *if $c < 1$, then the largest component of Q_p^n has order $\Theta(n)$; and,*

(ii) *if $c > 1$, then there exists a unique giant component in Q_p^n of order $(1+o(1))\gamma(c)2^n$, where y is defined according to* (1)*, and the second largest component has order $\Theta(n)$.*

Note here, in comparison to Theorem 1.1, that $\log|Q^n| = n$. Hence, we see a quantitatively similar phase transition in the component structure. When $c < 1$, which we call the *subcritical regime*, whp all the components of the percolated subgraph are small, of logarithmic order, whereas when $c > 1$, which we call the *supercritical regime*, whp a unique giant component of linear order emerges, covering the same asymptotic proportion of the vertices as in Theorem 1.1 (ii), and all other components are again of logarithmic order.

It can be seen that $\gamma \colon (1,\infty) \to \mathbb{R}$ is a continuous, increasing function and $\gamma(c) \to 1$ as $c \to \infty$, and so as c increases, less and less of the percolated subgraph lies outside of the giant component. Hence, already when $p = \frac{\omega(1)}{n}$ the giant component covers all but a vanishing proportion of the vertices. However, another classic result of Erdős and Rényi [32] shows that we have to wait significantly longer until $G(n,p)$ becomes connected.

Indeed, a connected graph cannot have any isolated vertices and hence, since K_n is an $(n-1)$-regular, n-vertex graph, we should not expect the graph to be typically connected until the expected number of isolated vertices $n(1-p)^{n-1}$ becomes small.

Theorem 1.3 ([32])**.** *Let $p = p(n)$ and let $\lambda(n,p) = n(1-p)^{n-1}$. Then*

$$\lim_{n \to \infty} \mathbb{P}(G(n,p) \text{ is connected}) = \begin{cases} 0 & \text{if } \lambda \to \infty; \\ e^{-c} & \text{if } \lambda \to c \geqslant 0. \end{cases}$$

The corresponding result in Q_p^d is originally due to Sapoženko [54] and Burtin [18], where the above heuristic suggests the threshold should occur when the expected number of isolated vertices, which can be seen to be $2^n(1-p)^n$, becomes small.

---

[3]Whilst the complete graph K_n is actually $(n-1)$-regular, we are viewing the phase transition on a coarse enough level that the difference is negligible.

Theorem 1.4 ([33]). *Let $p = p(n)$ and let $\lambda(n, p) = 2^n(1 - p)^n$. Then*

$$\lim_{n \to \infty} \mathbb{P}(Q_p^d \text{ is connected}) = \begin{cases} 0 & \text{if } \lambda(n, p) \to \infty; \\ e^{-c} & \text{if } \lambda(n, p) \to c \geqslant 0. \end{cases}$$

In this paper we study the evolution of the structure of a random subgraph of a different graph, the n-dimensional *permutahedron*. The permutahedron is a well-studied combinatorial object which has multiple equivalent representations, perhaps the simplest of which is its representation as a Cayley graph of the symmetric group on $n + 1$ elements S_{n+1}, generated by the adjacent transpositions.

Definition 1.5. *The n-dimensional permutahedron, denoted by $Perm(n)$, is the graph on vertex set S_{n+1} with*

$$E(Perm(n)) = \big\{ \{\pi, \pi\tau_i\} : \pi \in S_{n+1}, \, 1 \leqslant i \leqslant n \big\},$$

where τ_i is the transposition $(i, i + 1)$.

We note that the generating set of the Cayley graph above are precisely the generators in the presentation of the S_{n+1} as a *Coxeter group* (for background on Coxeter groups, see e.g., [22]).

However, $Perm(n)$ has a number of other equivalent definitions. It is the 1-skeleton of a polytope $P(n)$, which can be described as the convex hull of the points in $\mathbb{R}^{n+1}$ whose coordinates are given by the word representations of all permutations in S_{n+1}, but which can also be seen to be a *zonotope*, a Minkwoski sum of line segments. Note that, since $P(n)$ lives in an affine hyperplane of $\mathbb{R}^{n+1}$, it has intrinsic dimension n. See [57] for more on zonotopes and convex polytopes in general. The permutahedron can also be defined as the covering graph of a particular lattice, the *weak order* or *weak Bruhat lattice*, on S_{n+1}.

A weak analogy can be drawn here with the n-dimensional hypercube, which also has numerous equivalent representations, as the standard Cayley graph of the group $\mathbb{Z}_2^n$, with generators which represent the group as a Coxeter group, as the 1-skeleton of the hypercube polytope, which is also a zonotope, or as the covering graph of the lattice of subsets of $[n]$. As with the hypercube, due to its many equivalent representations, the permutahedron arises naturally in a variety of combinatorial contexts. However, whilst the graph theoretical properties of the hypercube have been well-studied, from both the probabilistic and the extremal viewpoint [36], the permutahedron has mostly been considered in the context of algebraic and enumerative combinatorics, and much less is known about structure of the permutahedron as a graph.

One reason that it is perhaps natural to study bond percolation on the permutahedron, is in the context of percolation on polytopes. Indeed, both K_n and Q^n are 1-skeletons of particularly symmetric polytopes, the n-simplex and the n-cube, respectively, and in some sense these are perhaps the most natural polytopes on which to study this percolation model since they are *simple*, and so the polytope is determined by its 1-skeleton [13, 43], and *regular*, and hence as symmetric as possible. Apart from some sporadic examples in low dimensions, there is only one other infinite family of regular polytopes, the *cross-polytope*, which is not simple, and whose 1-skeleton is the *cocktail party graph*, which is a complete graph with a perfect matching removed. However, since this graph is so close to a clique, similar methods as in [32] can be used to show analogues of Theorems 1.1 and 1.3.

Whilst the permutahedra are not regular, they are a family of *uniform* simple polytopes in each dimension, and so another natural family of highly symmetric polytopes whose combinatorial structure is determined by their 1-skeletons.

Our first main result is that the percolated permutahedron undergoes a quantitatively similar phase transition around the point $\frac{1}{n}$ as the binomial random graph $G(n, p)$.

Theorem 1.6. *Let $c > 0$, let $\gamma(c)$ be the survival probability of a $Po(c)$ branching process and let $p = \frac{c}{n}$. Then, whp,*

(i) *if $c < 1$, then whp the largest component of $Perm(n)_p$ has order $\Theta(n \log n)$; and*

(ii) *if $c > 1$, then there exists a unique giant component in $Perm(n)_p$ of order $(1+o(1))\gamma(c)(n+1)!$, where y is defined according to (1), and the second largest component has order $\Theta(n \log n)$.*

Note that, $\log |Perm(n)| = \Theta(n \log n)$, and so Theorem 1.6 is again quantitatively similar to Theorems 1.1 and 1.2. In the language of [24] we say that the permutahedron exhibits the *Erdős-Rényi component phenomenon*. Here, the fact that the order of $Perm(n)$ is superexponential in its regularity causes significant difficulties compared to the case of the complete graph or hypercube. The proof of Theorem 1.6 utilises a novel exploration process, which we call *projection-first search* (see Section 4), which is applicable to many high-dimensional geometric graphs, and significantly strengthens and simplifies the analysis of the distribution of 'large' clusters after percolation in these graphs. Furthermore, this process can be used to effectively enumerate small subgraphs in such graphs, which is useful for determining the existence of logarithmic sized components in both regimes.

We also consider the connectivity threshold in this model. Note that, since $Perm(n)$ is an n-regular, $(n+1)!$-vertex graph, the expected number of isolated vertices in $Perm(n)_p$ is $(n+1)!(1-p)^n$.

Theorem 1.7. *Let $p = p(n)$ and let $\lambda(n, p) = (n + 1)!(1 - p)^n$. Then*

$$\lim_{n \to \infty} \mathbb{P}(Perm(n)_p \text{ is connected}) = \begin{cases} 0 & \text{if } \lambda(n, p) \to \infty; \\ e^{-c} & \text{if } \lambda(n, p) \to c \geqslant 0. \end{cases}$$

Both the superexponential order of $|Perm(n)|$ and the lack of explicit control of the number of subgraphs and their expansion properties make it hard to argue as in the proof of Theorem 1.4. Here, instead, we give a novel proof of the connectivity threshold which uses instead information about the component structure and the distribution of vertices in the giant component derived in the proof of Theorem 1.6, which again should have applications to similar questions in other percolation models. Using standard methods, it is easy to give a corresponding hitting time result. We note that a similar approach to the connectivity threshold in Q^n and other high-dimensional product graphs was recently developed independently by Diskin and Krivelevich [29] and Diskin and Geisler [27].

As a tool to prove Theorems 1.6 and 1.7, but also as an interesting problem in its own right, we consider the *isoperimetric properties* of $Perm(n)$. In many contexts, the isoperimetric properties of the host graph G have been key to understanding the component structure of the percolated subgraph G_p [2, 15, 24, 34], and in turn the isoperimetric properties of the percolated subgraph G_p have been key to understanding its internal structure [9, 23, 26, 31].

Given a graph G, we define

$$i_k(G) \coloneqq \min_{S \subseteq V(G), |S|=k} \frac{|\partial(S)|}{|S|},$$

where $\partial(S)$ is the edge boundary of S, and the *edge-isoperimetric constant*, sometimes known as the *Cheeger constant*, given by

$$i(G) := \min_{1 \leqslant k \leqslant |V(G)|/2} i_k(G).$$

For a general graph, determining $i_k(G)$ or even $i(G)$ is known to be a computationally hard problem [35] but the exact value, or asymptotics, have been determined for many families of lattice-like graphs [1, 8, 16, 46].

In the case of the hypercube, a classic result of Harper [37] determines the value of $i_k(G)$ for all k. In particular, Harper's result implies the following bound.

Theorem 1.8 ([37], see also [10, 40, 47]). *Let $n \in \mathbb{N}$. For every $k \in [2^n]$,*

$$i_k(Q^n) \geqslant n - \log_2 k.$$

More generally, it is known that many other *high-dimensional* graphs have quantitatively similar expansion properties, for example Cartesian products [26, 55] or abelian Cayley graphs [46], where $i_k(G) = \Omega\left(\log\left(|G|/k\right)\right)$.

However, the isoperimetric problem does not seem to have been well-studied in $\mathrm{Perm}(n)$, despite it being mentioned as an open problem in the monograph of Harper [38]. Nevertheless, due to the many explicit representations of the permutahedron and its high level of symmetry, results in the literature can be used to give an essentially optimal bound for the expansion of small sets, as well as a general bound on the isoperimetric constant of $\mathrm{Perm}(n)$.

Proposition 1.9. *The edge-isoperimetric constant of $Perm(n)$ satisfies*

$$i(Perm(n)) = \Omega\left(\frac{1}{n^2}\right).$$

Moreover, for all $k \leqslant n!$,

$$i_k(Perm(n)) \geqslant n - \log_2 k.$$

The first bound here is a straightforward consequence of known bounds on the spectrum of the permutahedron [7] combined with Cheeger's inequality [19], which relates the expansion of a graph to its spectrum. The second bound is precisely the bound on the expansion of small sets from Harper's Theorem (Theorem 1.8) and follows immediately from the fact that the $\mathrm{Perm}(n)$ is a *partial cube*, an isometric subgraph of a (higher-dimensional) hypercube, which follows for example from its representation as the skeleton of a zonotope or as the Cayley graph of a Coxeter group.

The paper is structured as follows. In Section 2 we introduce some notation and give some preliminary lemmas. In Section 3 we discuss the isoperimetric properties of $\mathrm{Perm}(n)$ and give a proof of Proposition 1.9. In Section 4 we introduce the projection-first search algorithm and give some consequences for percolation on $\mathrm{Perm}(n)$ as well as for small subgraph counts. In Sections 5 and 6 we discuss the percolation and connectivity threshold and prove Theorems 1.6 and 1.7, respectively. Finally, in Section 7 we discuss avenues for future research.

2. Preliminaries

2.1. **Properties of the permutahedron.** We will need to use some basic facts about the structure of the permutahedron and about its various representations. In the paper it will sometimes be convenient to view the permutahedron as a polytope, which we will denote by $P(n)$, in order to

argue geometrically. The most natural way to describe $P(n)$ is as a $\mathcal{V}$-polytope, given by the convex hull of the *permutation vectors* $\{(\pi(1), \pi(2), \ldots, \pi(n+1)) \colon \pi \in S_{n+1}\}$ in $n+1$ dimensions. In fact, it is a simple exercise to show that $P(n)$ is also a *zonotope*, arising as the Minkowski sum

$$\frac{n+2}{2} \cdot \mathbf{1} + \sum_{1 \leqslant j < i \leqslant n+1} [-\boldsymbol{v_{ij}}, \boldsymbol{v_{ij}}],$$

where $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^{n+1}$ and $\boldsymbol{v_{ij}} = (\boldsymbol{e_i} - \boldsymbol{e_j})/2$, where $\{\boldsymbol{e_1}, \ldots, \boldsymbol{e_{n+1}}\}$ is the canonical basis of $\mathbb{R}^{n+1}$.

Remark 2.1. *It is well known that zonotopes are precisely the polytopes that arise as affine projections of hypercubes. Indeed, if $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_k}\} \subseteq \mathbb{R}^d$ then the zonotope $Z = \sum_{i=1}^{k} [-\boldsymbol{v_i}, \boldsymbol{v_i}]$ is the image of the hypercube $[-1, 1]^k \subseteq \mathbb{R}^k$ under the linear mapping $\pi \colon \mathbb{R}^k \to \mathbb{R}^d$ given by the matrix $V = (\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_k})$ whose columns are the summands $\boldsymbol{v_i}$.*

Note that it is apparent that the polytope $P(n)$ is contained in an n-dimensional affine subspace of $\mathbb{R}^{n+1}$, and so we will consider $P(n)$ as an n-dimensional polytope.

It be seen that the *facets*, the $(n-1)$-dimensional faces, of the polytope $P(n)$ are in one to one correspondence with subsets $\emptyset \subsetneq A \subsetneq [n+1]$, where the facet $F(A)$ contains the vertices given by $\{\pi \in S_{n+1} : \pi(A) = \{1, \ldots, |A|\}\}$. Using this, and the fact that $P(n)$ is simple, it is relatively easy to describe the faces of of $P(n)$, and in particular it can be shown that each face of dimension k is a cartesian product of lower dimensional permutahedra whose dimensions sum up to k (and that each such product is realised as a face of dimension k).

If a graph H is isomorphic to the cartesian product $\bigoplus_{i=1}^{t} \mathrm{Perm}(n_i)$ of some family of permutahedra, then we say that H is a *face graph* of dimension $d := d(H) = \sum_{i=1}^{t} n_i$. Note that such a graph is then d-regular.

A second presentation of the permutahedron that will be important to us is as the Cayley graph arising from a *Coxeter system*. The Coxeter groups are those groups Γ which have a representation of the form

$$\langle r_1, r_2, \ldots, r_n \mid (r_i r_j)^{m_{i,j}} = e \rangle$$

where $m_{i,i} = 1$ for all i and $m_{i,j} = m_{j,i} \geqslant 2$ if $i \neq j$. Given such a representation we call $(\Gamma, \{r_1, \ldots, r_n\})$ a *Coxeter system*.

Given a Coxeter system $(\Gamma, \{r_1, \ldots, r_n\})$ and a subset $I \subseteq [n]$ let us write Γ_I for the subgroup of Γ generated by $\{r_i \colon i \notin I\}$. We will need the following elementary fact about Coxeter groups.

Lemma 2.2 ([22, Lemma 4.1.6]). *Let $(\Gamma, \{r_1, \ldots, r_n\})$ be a Coxeter system and let $I, J \subseteq [n]$. Then*

$$\Gamma_I \cap \Gamma_J = \Gamma_{I \cup J}.$$

Given Lemma 2.2, it is an elementary exercise in group theory to show that the intersection of any two cosets $g_i \Gamma_I \cap g_j \Gamma_J$ is then either empty or a coset of $\Gamma_{I \cup J}$. In particular, for any tuple $(g_1, g_2, \ldots, g_n)$ there is at most one element $h \in \bigcap g_i \Gamma_i$, and so we may identify Γ with a particular subset of Γ^n in this fashion, giving a coordinate representation of sorts, which will be useful later in the paper.

2.2. **Auxiliary Lemmas.** We use the following result which bounds the number of m vertex trees in a graph.

Lemma 2.3 ([11, Lemma 2]). *Let G be a graph on n vertices with maximum degree Δ and minimum degree δ. Let $t_m(G)$ be the number of m-vertex rooted trees in G. Then,*

$$n \cdot \frac{m^{m-2}(\delta - m)^{m-1}}{(m-1)!} \leqslant t_m(G) \leqslant n(e\Delta)^{m-1}.$$

We will also need the following two results, which bound the likely diameter and maximum degree in a random m-vertex tree, due to Rényi and Szekeres [53] and Moon [51], respectively.

Lemma 2.4 ([53]). *Let T be a tree chosen uniformly at random from all trees with vertex set $[m]$. Then T has diameter $O_p(\sqrt{m})$.*

Lemma 2.5 ([51]). *Let T be a tree chosen uniformly at random from all trees with vertex set $[m]$. Then $\Delta(T) = (1 + o(1)) \log m / \log \log m$ with probability tending to one as $m \to \infty$.*

We will also need some basic facts about Galton-Watson trees (see for example [39]).

Lemma 2.6. *Let $c > 1$ be a constant and let T be a Galton-Watson tree with child distribution $Bin(n, p)$ with $np = c$ and let $\gamma(c)$ be defined as in (1) Then*

  (i) *$\gamma(c)$ is a continuous function on $(1, \infty)$;*
  (ii) *$\gamma(c) \geqslant c - 1$ for all $1 < c \leqslant 5/4$, and furthermore $\gamma(c) = 2(c-1) + O((c-1)^2)$ as $c \to 1^+$;*
  (iii) *as n tends to infinity, the probability that T is infinite tends to $\gamma(c)$;*
  (iv) *If k is growing with n, then the probability that T is finite and has size at least k is $o_n(1)$; and*
  (v) *if we condition on the fact that T is infinite, then as k tends to infinity*

$$|T_k|^{1/k} \to c$$

  in probability, where $|T_k|$ is the size of the k-th generation.

We will use two concentration bounds. The first is a version of the Chernoff bounds [41].

Lemma 2.7. *Let $n \in \mathbb{N}$, let $p \in [0, 1]$ and let $X \sim Bin(n, p)$.*
  (i) *For every positive t with $t \leqslant \frac{3np}{2}$,*

$$\mathbb{P}\left(|X - np| > t\right) < 2\exp\left(-\frac{t^2}{3np}\right).$$

  (ii) *For every positive b,*

$$\mathbb{P}\left(X \geqslant bnp\right) \leqslant \left(\frac{e}{b}\right)^{bnp}.$$

The second is a version of the Azuma-Hoeffding inequality, [5, Chapter 7].

Lemma 2.8. *Let $X = (X_1, X_2, \ldots, X_n)$ be a random vector with independent entries and range $\Lambda = \prod_{i \in [n]} \Lambda_i$ and let $f \colon \Lambda \to \mathbb{R}$ and $C \in \mathbb{R}$ such that for every $x, x' \in \Lambda$ which differ only in one coordinate,*

$$|f(x) - f(x')| \leqslant C.$$

Then, for every $t \geqslant 0$,

$$\mathbb{P}\left[\left|f(X) - \mathbb{E}\left[f(X)\right]\right| \geqslant t\right] \leqslant 2\exp\left(-\frac{t^2}{2C^2}\right).$$

We will also use the following inequality, which links the isoperimetric properties of a graph to the spectrum of its adjacency matrix, which is sometimes referred to as *Cheeger's inequality*.

Lemma 2.9 ([3, 4])**.** *Let G be an n-regular graph with adjacency matrix A. Then*

$$i(G) \geqslant \lambda_1/2$$

where λ_1 is the second smallest eigenvalue of the Laplacian $nI - A$.

3. The isoperimetric problem

In this section we will discuss the (edge-)isoperimetric problem on $\mathrm{Perm}(n)$, showing in particular how Proposition 1.9 follows from results in the literature.

The first bound in Proposition 1.9 is a relatively straightforward consequence of Cheeger's inequality (Lemma 2.9) together with known bounds on the spectrum of $\mathrm{Perm}(n)$.

Lemma 3.1. *The edge isoperimetric constant of $\mathrm{Perm}(n)$ satisfies*

$$i(\mathit{Perm}(n)) = \Omega\left(\frac{1}{n^2}\right).$$

Proof. Bacher [7] showed that the smallest positive eigenvalue λ_1 of the Laplacian of $\mathrm{Perm}(n)$ is

$$\lambda_1(\mathrm{Perm}(n)) = 2 - 2\cos\left(\frac{\pi}{n+1}\right) = \Theta\left(\frac{1}{n^2}\right),$$

and so the result follows from Lemma 2.9. □

We remark that it seems unlikely that the bound in Lemma 3.1 is optimal in terms of its dependence on n. However, for our application we only require that $i(\mathrm{Perm}(n))$ is not shrinking very fast as a function of n, and in fact any inverse polynomial bound here would be sufficient. We note that *some* inverse polynomial factor here is necessary, as the set

$$S = \left\{\sigma \in S_{n+1} \colon \sigma^{-1}(1) \leqslant \frac{n+1}{2}\right\}$$

of size $\frac{(n+1)!}{2}$ has $\partial(S) = \frac{2|S|}{n+1}$ and so witnesses that $i(\mathrm{Perm}(n)) \leqslant \frac{2}{n+1}$. Indeed, a vertex is incident to an edge in the boundary of S if and only if $\sigma^{-1}(1) = \frac{n+1}{2}$, and in this case the edge is unique, corresponding to the transposition $\tau_{\frac{n+1}{2}}$. Hence $\partial(S) = n! = \frac{2|S|}{n+1}$.

For small sets however, Lemma 3.1 is far from the truth and in fact $\mathrm{Perm}(n)$ satisfies Harper's inequality (Theorem 1.8) and so has almost optimal small-set expansion. Indeed, this follows from the fact that the skeleton of a zonotope is a partial cube. This can be deduced easily from known results in the literature, but we were not able to find a direct statement of this fact.

Lemma 3.2. *Let $\boldsymbol{v_1}, \ldots, \boldsymbol{v}_k \in \mathbb{R}^d$, let $Z = \sum_{i=1}^{k} \boldsymbol{v_k}$ and let G be the 1-skeleton of Z. Then G is a partial cube.*

Proof. There is a known correspondence between zonotopes, hyperplane arrangements and oriented matroids, see [12]. In particular, the skeleton of Z is known to be the *tope graph* of the corresponding oriented matroid and it can be shown that the tope graph of any oriented matroid is a partial cube [12, Lemma 4.2.2]. In fact, it can be seen that the mapping π from Remark 2.1, when restricted to the vertices of Z, is a graph isometry. □

In the particular case of permutahedra, it can be checked that the map from $V(\mathrm{Perm}(n))$ to the hypercube whose vertices encode subsets of $\binom{[n+1]}{2}$ given by

$$\pi \mapsto \{\{i, j\} \subseteq [n+1] : i < j \text{ and } \pi^{-1}(i) > \pi^{-1}(j)\}$$

is a graph isometry.

Let us discuss briefly the (edge)-isoperimetric problem in $\mathrm{Perm}(n)$ more generally. Since $\mathrm{Perm}(n)$ is n-regular, it is clear that Harper's inequality is asymptotically tight for $\log k \ll n$. In fact, it is tight for all k of the form 2^r with $r \leqslant \frac{n}{2}$. Indeed, for all such r, $\mathrm{Perm}(n)$ contains a face F which is a hypercube of dimension r, and since $\mathrm{Perm}(n)$ is n-regular, it follows that $|\partial(F)| = |F|(n-r) = |F|(n - \log_2 k)$.

Here then, as opposed to the case of the hypercube, we see that the optimal sets are not given by lower dimensional permutahedra, as perhaps might be expected at first glance. Indeed, subsets of the form $\mathrm{Perm}(r) \subseteq \mathrm{Perm}(n)$ of size $(r+1)! = k$ have an expansion factor of $n - r \approx n - \frac{\log k}{\log \log k}$, which is significantly larger than the bound given by Theorem 1.8.

The point here is that $\mathrm{Perm}(n)$ contains faces of much higher density than $\mathrm{Perm}(r)$, and in particular the densest faces are copies of the hypercube, the largest of which has dimension $\frac{n}{2}$. It seems likely that, at least for $k \leqslant 2^{\frac{n}{2}}$, the optimal sets for the edge-isoperimetric problem are precisely those which optimise the edge boundary inside a copy of the hypercube, and so in particular can be chosen to be nested.

For larger k it becomes less clear what structure the minimising sets should have. If we restrict ourselves to looking at subsets of faces, then the expansion ratio 'outside' the face is smallest when the face is as dense as possible. So, given $k \in \mathbb{N}$ we might want to choose a face F of size at least k, which is as dense as possible given this restriction, and choose a subset of a face F of size k which minimises the boundary inside F.

Since each face of dimension d induces a regular subgraph of degree d, and each face is product of lower dimensional permutahedra, it is relatively clear that for a fixed dimension d, the largest faces are those in which the dimension of the factors are as evenly split as possible. In particular, if $k = 6^{\frac{n+1}{3}}$ this heuristic would suggest an optimal set is a face which arises as the cartesian product of $\frac{n+1}{3}$ 6-cycles, which has an edge-boundary of order $k \left(\frac{n+1}{3} - 1 \right)$.

Conjecture 3.3. *Let $k = 6^{\frac{n+1}{3}}$. Then $i_k(Perm(n)) = \frac{n+1}{3} - 1$.*

However, since the largest face in $\mathrm{Perm}(n)$ has order $n!$, this intuition cannot extend to larger subsets of $\mathrm{Perm}(n)$, for example linear sized sets. By Lemma 3.1 and the discussion following it we know that

$$\frac{2}{n+1} \geqslant i(\mathrm{Perm}(n)) = \Omega \left(\frac{1}{n^2} \right)$$

and it would be very interesting to determine what the correct order polynomial dependence on n is, and if the optimal sets in the range from $n!$ to $(n+1)!$ are also given, as in the example, by appropriate unions of faces.

Question 3.4. *What is the minimal $c \in \mathbb{R}$ such that*

$$i(Perm(n)) = \Omega \left(n^{-c} \right)?$$

4. The projection lemma and projection-first search

A key property of the permutahedron that will form the basis of our analysis is that it has a certain type of fractal self-symmetry. Roughly, given a small set X of vertices, we can cover X with a family of disjoint subgraphs each of which is in some way a lower-dimensional graph with similar properties to the permutahedra. Similar results have been key to the analysis of percolation in the

hypercube and other high-dimensional product graphs [20, 24, 26, 28] and have been shown to hold in a variety of graphs with high-dimensional geometric structure [20].

A key difference here is that we cannot necessarily cover the set X with disjoint copies of lower dimensional permutahedra, but must allow ourselves more freedom in what we take as our projections, working instead with the more general concept of a face graph.

Lemma 4.1 (Projection Lemma)**.** *Let H be a face graph of dimension m and let $X \subseteq V(G)$ have size $|X| = k \leqslant m$. Then there a disjoint family of subgraphs $\{H(x) \colon x \in X\}$ of H such that*

- *Each $H(x)$ is a face graph of dimension at least $m + 1 - k$;*
- *$x \in V(H(x))$ for all $x \in X$.*

We will often refer to the face graphs guaranteed by Lemma 4.1 as *projections* of the graph H.

We in fact can prove a more genereal statement for zonotopes, but only when the polytope is *simple*, i.e., every vertex figure is a simplex[4], which may be of independent interest.

Lemma 4.2. *Let Z be a zonotope which is a simple polytope and let $S \subseteq V(Z)$ have size $|S| = s$. Then there exists a disjoint family of faces $\{F(x) \colon x \in S\}$ such that*

- *$F(x)$ has codimension at most $s - 1$ for all $x \in S$;*
- *$x \in F(x)$ for all $x \in X$.*

Proof. If $Z = \sum_{i=1}^{k} [-\boldsymbol{v_i}, \boldsymbol{v_i}]$, then it can be seen that every vertex $x \in V(Z)$ has a unique representation of the form

$$x = \sum_{i=1}^{k} \varepsilon_i \boldsymbol{v_i} \tag{2}$$

where $\varepsilon_i \in \{-1, 1\}$, where $(\varepsilon_1, \ldots, \varepsilon_k) = \pi^{-1}(x)$ for the map π from Remark 2.1.

However, the analogue of Lemma 4.1 for the hypercube has already been established, see for example [28, Claim 2.2] or an alternative proof will follow shortly in Lemma 4.3. Applying this to the subset $\pi^{-1}(S)$ of $V\left(Q^k\right)$ of size s, we can conclude the existence of a family $\{Q(x) \colon x \in S\}$ of vertex-disjoint subcubes of codimension at most $s - 1$ such that $\pi^{-1}(x) \in Q(x)$ for each $x \in S$. Let $I(x) \subseteq [k]$ be the set of free coordinates in this subcube, where $|I| \geqslant k - s + 1$.

Since Z is simple, there is a set $J(x) \subseteq [k]$ of size $|J(x)| = d$ such that $\{x, x - 2\varepsilon_j \boldsymbol{v_j}\}$ is an edge of Z if and only if $j \in J(x)$. Since $I(x), J(x) \subseteq [k]$, it follows that $|I(x) \cap J(x)| \geqslant d - s + 1$. Using again that Z is simple, the vertex figure at x is a simplex, and therefore

$$F(x) := x + \sum_{i \in I(x) \cap J(x)} [-2\varepsilon_i \boldsymbol{v_i}, 0]$$

is a face of Z of codimension at most $s - 1$. These faces are disjoint by construction – indeed, if $F(x) \cap F(y)$ meet, they must share some vertex z. By construction, $\pi^{-1}(z)$ must agree with $\pi^{-1}(x)$ outside of the coordinates in $I(x) \cap J(x)$ and agree with $\pi^{-1}(y)$ outside of the coordinates in $I(y) \cap J(y)$, however this implies that $\pi^{-1}(z) \in Q(x) \cap Q(y)$, a contradiction. $\square$

Using the description in Section 2 of the faces of the permutahedron, it is easy to see that each d-dimensional face graph is the skeleton of some d-dimensional zonotope, and the skeleton of every k-dimensional face of this zonotope is a k-dimensional face graph. In this way Lemma 4.1 follows from Lemma 4.2.

---

[4]Not to be confused with a simple zonotope, which is a zonotope where no generators or parallel or zero.

We suspect that the restriction to zonotopes which are simple polytopes in Lemma 4.2 is not necessary, at least in a qualitative manner, and a similar statement, with perhaps a worse bound on the co-dimension of the faces should hold for general zonotopes.

For the specific case of the permutahedron we can also give the following simpler proof using its representation as a Cayley graph of Coxeter group.

Lemma 4.3. *Let $(\Gamma, \{r_1, \ldots, r_m\})$ be a Coxeter system and let $S \subseteq \Gamma$ have size $|S| = s$. Then there is a family of subsets $\{I(x) \colon x \in S\}$ of $[n]$ and a disjoint family of cosets $\{w(x)\Gamma_{I(x)} \colon x \in S\}$ such that*

- *$|I(x)| \leqslant s - 1$ for all $x \in S$;*
- *$x \in w(x)\Gamma_{I(x)}$ for all $x \in S$.*

Proof. We induct on s and m, where the case $s = 1$ is trivial. For each i the cosets of Γ_i partition Γ and by Lemma 2.2 the intersection of any family of cosets $\bigcap_{i=1}^{n} w_i \Gamma_i$ has size at most one. It follows that there must be some i such that the partition of X induced by the cosets of Γ_i is non-trivial. However, since Γ_i is isomorphic to the Coxeter system $(\Gamma_i, \{r_1, \ldots, r_{i-1}, r_{i+1}, \ldots, r_m\})$, and each coset contains at most $s - 1$ points, the claim follows by induction. □

We note however, that it can be shown (see for example [12, Section 2.3]) that every reflection arrangement is simplicial, and so the skeleton of its polar dual, which is isomorphic to the Cayley graph of the Coxeter group corresponding to the reflection arrangement, is a zonotope which is a simple polytope. Hence, Lemma 4.2 implies Lemma 4.3.

To see that Lemma 4.1 follows from Lemma 4.3, we note that $\mathrm{Perm}(n)$ is the Cayley graph of S_{n+1} with respect to a generating set S where (S_{n+1}, S) is a Coxeter system, and that there is a one-to-one correspondence between faces of $\mathrm{Perm}(n)$ of codimension k and cosets of the subgroups of S_{n+1} formed by omitting some set of k generators from S, where the skeleton of the face is isomorphic to the Cayley graph of the coset, which is also a Coxeter group.

4.1. Projection-first search.

A classic technique in the study of percolated subgraphs is to explore the component (cluster) C_v in G_p containing a vertex v using breadth-first search (BFS). When the host graph G is n-regular, then, at least in the early stages, before we have discovered $\Theta(n)$ many vertices, this process behaves like a Galton-Watson branching process with child distribution $\mathrm{Bin}(n, p)$. In this way, we can relate the probability that v lies in a 'large' cluster to the survival probability of this branching process.

In $G(n, p)$ this intuition can be used to show that in the supercritical regime whp the correct proportion of vertices lie in components of linear order, and a simple sprinkling argument shows that these linear components all merge into a unique *giant component*. However, when the order of the host graph is much larger than its regularity, as in many high-dimensional geometric graphs such as the hypercube or permutahedron, we cannot necessarily guarantee that the clusters will grow larger than $\Theta(n)$ with positive probability in this manner, which causes difficulties in the merging step.

In this section, we describe a variant of the depth-first search algorithm, which we call projection-first search (PFS), which grows significantly larger percolation clusters. The idea here is to avoid 'backtracking' by carefully separating the vertices of the BFS tree into disjoint parts using the projection lemma. In this way, we can guarantee that the number of explored neighbours of a vertex will decrease much more slowly during the process: linearly in the depth of the BFS tree,

rather than linearly in the size of the tree. This will ensure that the process can be approximated by a $\text{Bin}(n,p)$ branching process until the tree grows to *depth* $\Theta(n)$, rather than size $\Theta(n)$. Since, in the supercritical regime, we expect the layers of the BFS tree to be growing exponentially quickly, we should then be able to build exponentially large clusters in this manner.

Formally, we will run this algorithm on some face graph H of dimension m, and we will keep track of a number of different subsets and subgraphs of H. The algorithm will run in a number of rounds, and in the t-th round of the algorithm we will have

- A set $W(t) \subseteq V(G)$ of *explored* vertices, such that all edges exposed before the t-th round are incident to $W(t)$;
- A set $U(t) = V(G) \setminus W(t)$ of *unexplored* vertices;
- A *frontier* $A(t) \subseteq W(t)$, such that no edge from $A(t)$ to $U(t)$ has been exposed before the t-th round;
- A (vertex-)disjoint family $\{H(x) \colon x \in A(t)\}$ of face subgraphs of H, such that $H(x) \cap W(t) = \{x\}$ for all $x \in A(t)$.

We initialise by setting $A(1) = W(1) = \{v\}$, $H(v) = H$ and $U(1) = V(H) \setminus \{v\}$.

In the t-th round, for each vertex $x \in A(t)$ we expose the neighbours $N_t(x)$ of x in $H(x)$. Note that since $H(x) \cap W(t) = \{x\}$, the edges incident to x in $H(x)$ have not yet been exposed in the exploration process. We then apply Lemma 4.1 to the set $N_t(x) \cup \{x\}$ inside the face graph $H(x)$ to obtain a disjoint family of projections, and for each $y \in N_t(x)$ we set $H(y)$ to be the projection which contains y, where we note that each $H(y)$ has dimension at least $d(H_t(x)) - |N_t(x)|$. We now update

$$U(t+1) = U(t) \setminus \bigcup_{x \in A(t)} N_t(x),$$

$$W(t+1) = W(t) \cup \bigcup_{x \in A(t)} N_t(x),$$

$$A(t+1) = \bigcup_{x \in A(t)} N_t(x).$$

Note that, since the family $\{H(x) \colon x \in A(t)\}$ is disjoint by assumption and for all $y \in N_t(x)$, $H(y) \subseteq H(x)$, it then follows that $\{H(y) \colon y \in A(t+1)\}$ is a disjoint family by construction. Furthermore, for any $x \in A(t)$ and any $y \in N_t(x)$, since $H(y) \subseteq H(x)$ and $H(y) \cap (N_t(x) \cup \{x\}) = \{y\}$, it follows that $H(y) \cap W(t+1) = \{y\}$.

Let us note some basic facts about the algorithm:

(a) For each t, $W(t)$ spans a tree $T = T(t)$ in H;
(b) For each $x \in A(t)$, $|N_t(x)|$ is distributed as a binomial random variable $\text{Bin}(d(H(x)), p)$;
(c) For each $x \in A(t)$, the dimension of $H(x)$ is given by $d - w(x)$ where, writing $v = v_1 v_2 \ldots v_t = x$ for the path vTx from v to x in T,

$$w(x) := \sum_{i=1}^{t-1} |N_i(v_i)| = d_T(v_1) + \sum_{i=2}^{t-1} (d_T(v_i) - 1). \tag{3}$$

12

Lemma 4.4. *Let H be a face graph of dimension $m \in \mathbb{N}$, let $\beta > 0$ and let $p \geqslant \frac{1+\beta}{m}$. Then for all $v \in V(H)$*

$$\mathbb{P}\left(|C_v| = \exp\left(\Omega\left(\frac{m}{\log^3 m} \right) \right) \right) \geqslant \gamma(1+\beta) - o_m(1)$$

Proof. Let $v \in V(H)$. We first note that, since the event that $|C_v|$ is large is an increasing event, we may assume that $p = \frac{1+\beta}{m}$. We run the PFS algorithm starting at v for $\tau = \frac{m}{\log^3 m}$ rounds.

Deterministically, $|W(\tau)| \leqslant m^\tau = \exp\left(\frac{m}{\log^2 m} \right)$ and by (b) for each $t \leqslant \tau$ and each $x \in A(t)$, the size of the neighbourhood $N_t(x)$ is stochastically dominated by a $\mathrm{Bin}(m, p)$ random variable, and hence, by (3) and the Chernoff bound (Lemma 2.7), whp $w(x) \leqslant \frac{m}{\log m}$ for all $x \in A(\tau)$.

In particular, since $w(y) \leqslant w(x)$ for all $y \in vTx$, for every $y \in W(\tau)$ the dimension of $H(y)$ is at least $m - \frac{m}{\log m}$. Hence, for each $t \leqslant \tau$ and each $x \in A(t)$, the size of the neighbourhood $N_t(x)$ stochastically dominates a $\mathrm{Bin}\left(m - \frac{m}{\log m}, p \right)$ random variable, which has expectation

$$\left(m - \frac{m}{\log m} \right) p \geqslant 1 + \beta + o_m(1).$$

Hence, we can couple the exploration process from below with a Galton-Watson tree with a binomial child distribution with expectation at least $1 + \beta + o_m(1)$, and so by Lemma 2.6(v) by the τ-th round this process grows to size

$$(1 + \beta + o_m(1))^\tau = \exp\left(\Omega\left(\frac{m}{\log^3 m} \right) \right)$$

with probability at least $\gamma(1+\beta) + o_m(1)$. $\qquad\square$

In the above, we have focused on giving a particularly straightforward proof, emphasising the natural coupling with a branching process, without attempting to optimise the size of the clusters, and it is natural to ask about the limit of these methods.

On a heuristic level, each time we project we expect the dimension of the projection to decrease by a constant, and so we expect the process to remain supercritical until we've discovered $\Theta(m)$ many layers. Furthermore, whilst the process remains (strictly) supercritical, for example whilst $p \cdot d(H(x)) \geqslant 1 + \frac{\beta}{2}$ for all $x \in A(t)$, we expect the size of the frontier to grow exponentially quickly, with rate at least $\left(1 + \frac{\beta}{2} \right)$.

In particular, if we are more careful in our analysis, we might hope to build clusters of exponential size in this manner. For example, if we naively truncate our exploration process in each round, to ensure that we only uncover at most K neighbours of each vertex, so that the dimension never reduces by more than K in each step, we can choose $K = K(\beta)$ large enough that this process still stochastically dominates a supercritical branching process for at least $\Theta(m)$ layers. However, the expected number of children in the coupled branching process is then smaller, and hence the probability that the branching process percolates is slightly smaller, leading to a suboptimal bound on the probability that a vertex is contained in a large cluster.

However, if instead we first run a standard BFS process for $\omega(1)$ rounds, then by Lemma 2.6 with probability $\gamma(1+\beta) + o(1)$ not only does the process survive, but the frontier has size $\omega(1)$. At this point we can switch to the truncated PFS process described above, which ensures that the dimension does not drop too much between the layers, which will then remain supercritical for at

least $\Theta(m)$ layers. If we carefully implement this modified process, which we denote by PFS′, we obtain the following strengthening of Lemma 4.4, whose proof we defer to the appendix.

Lemma 4.5. *Let H be a face graph of dimension $m \in \mathbb{N}, \beta > 0$ and let $p \geqslant \frac{1+\beta}{m}$. Then for all $v \in V(H)$*

$$\mathbb{P}\left(|C_v| = \exp\left(\Omega(m)\right)\right) \geqslant \gamma(1+\beta) - o_m(1)$$

We note further that we can apply this method to any class of graphs where a quantitatively similar projection lemma holds, for example the hypercube, arbitrary product graphs [24, Lemma 4.1] or the middle layers or odd graph [20]. We can then compare the strength of this method to previous results on the phase transition in such models.

In the first proof of the existence of a giant component in Q_p^n given by Ajtai, Komlós and Szemerédi [2], they use a two step argument to show that each vertex is contained in a cluster of size $\Omega(n^2)$ with some constant probability $c(\beta)$, although their methods can easily be used inductively to show that the same holds for clusters of size n^k for arbitrary $k \in \mathbb{N}$, and with careful bookkeeping one can take $c(\beta) = \gamma(1+\beta) - o(1)$. The later work of Bollobás, Kohayakawa and Łuczak [15] avoids having to estimate the probability that a vertex lies in a large cluster by instead demonstrating a gap in the order of the components and estimating the number of vertices contained in *small* clusters. However, their proof method relies on particular strong and explicit bounds on the isoperimetric profile of the hypercube, and so cannot be applied in other contexts. In the work of Diskin, Erde, Kang and Krivelevich [24, 25], an inductive argument is used together with a projection lemma to show that each vertex is contained in a cluster of size n^r for any $r = o(n^{1/3})$ with probability $\gamma(1+\beta) - o(1)$, and a natural limit to their method seems to be $r = n^{1/2}$.

In comparison, using for example [24, Lemma 4.1] as a projection lemma, the proof of Lemma 4.5 shows that each vertex in Q_p^n is contained in a cluster of size $2^{\Theta(n)}$ with probability at least $\gamma(1+\beta) - o(1)$. This strengthening seems to have a few uses – beyond improving some quantitative aspects of previous work (for example, it should be possible to improve the bound on the isoperimetric constant in [25, Theorem 2] from $t^{-t^{1/4}}$ to something closer to $2^{-\Theta(t)}$), an immediate application is to percolation in irregular product graphs. In [25] it was observed that, whilst all regular product graphs undergo a quantitatively similar phase transition around the percolation threshold as $G(n, p)$, the same is not necessarily true in the irregular case. In particular, if we consider the product of a number of copies of a star, a third regime in the phase transition appears, where there is still no giant component, but there are components of (almost linear) polynomial size. The authors asked whether this holds in fact for *any* irregular product graph.

Question 4.6 ([25, Question 5.5]). *For all $i \in [t]$, let $G^{(i)}$ be an irregular connected graph of order at most $C > 0$. Let $G = \Box_{i=1}^t G^{(i)}$. Let $\varepsilon > 0$ be a small enough constant, and let $p = \frac{1-\varepsilon}{d}$. Does there exist a $c(\varepsilon, C)$ such that whp the largest component in G_p has order at least $|G|^c$?*

Using the projection lemma from [24, Lemma 4.1], it is relatively straightforward to use the methods of Lemma 4.5 to show that Question 4.6 has a positive answer. As an explicit example, this implies that the phase-transition in a high-dimensional grid $[3]^n$ is qualitatively different to the phase-transition on a high-dimensional torus $\mathbb{Z}_3^n$. It would be very interesting to determine more precisely the size of the largest components in this intermediary regime, and we hope to study this problem in more detail in future work.

4.2. Consequences of PFS for percolation. Given a graph G and $r \in \mathbb{N}$, let us write $V_{\geqslant r}(G)$ for the set of vertices in G contained in a component of size at least r. For our application to the permutahedron it will be convenient to fix

$$r := n^{19}$$

for the rest of the paper.

The first application of Lemma 4.4 is to conclude that the vertices in large clusters are relatively well-distributed throughout $\mathrm{Perm}(n)$, in the sense that every vertex is within distance two of a vertex contained in a large cluster.

Lemma 4.7. *Let $\beta > 0$ and let $p \geqslant \frac{1+\beta}{n}$. Then whp every vertex in $\mathrm{Perm}(n)_p$ is within distance two (in $\mathrm{Perm}(n)$) of $V_{\geqslant r}(\mathrm{Perm}(n)_p)$.*

Proof. Let $s = \frac{n}{\log n}$. Given a vertex $v \in V(\mathrm{Perm}(n))$, let $W = \{w_1, \ldots, w_s\} \subseteq N_{\mathrm{Perm}(n)}(v)$ be an arbitrary subset of the neighbours of v. We apply Lemma 4.1 to $W \cup \{v\}$ to obtain a family $\{H(w_i): i \in [s]\}$ of disjoint face graphs of dimension at least $\left(1 - \frac{1}{\log n}\right)n$. For each w_i, let $\{w_{i,1}, \ldots, w_{i,s}\} \subseteq N_{H(w_i)}(w_i)$ be an arbitrary subset of the neighbours of w_i in $H(w_i)$. Note that each $w_{i,j}$ is at distance two from v. Again, by Lemma 4.1 we can find a family $\{H(w_{i,j}): i, j \in [s]\}$ of disjoint face graphs of dimension at least $m := \left(1 - \frac{2}{\log n}\right)n$.

For each $i, j \in [s]$, since the dimension of each $H(w_{i,j})$ is at least m and $p = \frac{1+\beta}{n} \geqslant \frac{1+\beta+o(1)}{m}$, by Lemma 4.4 the probability that $w_{i,j}$ is contained in a component of order r in $H(w_{i,j})_p$ is at least $\gamma(1 + \beta) - o(1)$, and since the $H(w_{i,j})$ are disjoint, these events are independent for different $i, j \in [s]$. Hence, the probability that no $w_{i,j}$ lies in $V_{\geqslant r}(\mathrm{Perm}(n)_p)$ is at most

$$(1 - \gamma(1 + \beta) + o(1))^{s^2} = o\left(\frac{1}{(n+1)!}\right),$$

and the result follows by taking a union bound. $\qquad\square$

The second consequence is a slightly more technical notion of density, which will be useful later. It implies that whp every large connected set of vertices contains many vertices with many neighbours contained in large clusters.

Lemma 4.8. *Let $\beta > 0$, $\beta' = \min\{\beta, 1\}$, $\alpha \leqslant 2^{-5}(\beta')^2$ and let $p \geqslant \frac{1+\beta}{n}$. Then whp every connected subset $M \subseteq \mathrm{Perm}(n)$ of size at most $n^{3/2}$ contains at most αn vertices with at most αn neighbours in $V_{\geqslant r}(\mathrm{Perm}(n)_p)$.*

Proof. It clearly suffices to consider connected sets $M \subseteq \mathrm{Perm}(n)$ of size exactly $m = n^{3/2}$. For any such set M, and any subset $X \subseteq M$ of size αn, we will upper bound the probability that every $x \in X$ has fewer than αn neighbours in $V_{\geqslant r}(H(x)_p)$. To do so, we start by applying Lemma 4.1 to find a family $\{H(x): x \in X\}$ of disjoint projections of dimension at least $(1 - \alpha)n$ such that $x \in H(x)$ for each $x \in X$.

We now fix a particular $x \in X$ and choose a subset $W = \{w_1, \ldots, w_s\}$ of $s = 8\alpha n/\beta'$ neighbours of x in $H(x)$. Applying Lemma 4.1 again inside $H(x)$, we may find a family $\{H(x,i): i \in [s]\}$ of disjoint projections of dimension at least $(1-\alpha-8\alpha/\beta')n$ such that $w_i \in H(x,i)$. Since $\alpha+8\alpha/\beta' \leqslant \beta'/2$, by Lemma 4.4 each of these neighbours is contained in $V_{\geqslant r}(H(x,i)_p)$ independently with probability at least

$$\gamma\left(\frac{1+\beta}{1-\frac{\beta'}{2}}\right) + o(1) \geqslant \gamma\left(1 + \frac{\beta'}{4}\right) + o(1) \geqslant \frac{\beta'}{4}.$$

15

Therefore, we expect that at least $\frac{\beta'}{4} \cdot s = 2\alpha n$ elements of W are in $V_{\geqslant r}(H(x)_p)$. By Chernoff's inequality (Lemma 2.7), the probability that a particular $x \in X$ has at most αn neighbours in $V_{\geqslant r}(H(x)_p)$ is $\exp(-\Omega(n))$.

Since the $H(x)$ are disjoint, the probability that every $x \in X$ has fewer αn neighbours in $V_{\geqslant r}(H(x)_p)$ is $\exp(-\Omega(n|X|)) = \exp(-\Omega(n^2))$. It only remains to apply the union bound. By Lemma 2.3 there are at most $(n+1)!(en)^m$ connected subsets of $\mathrm{Perm}(n)$ of size m, and given a fixed such subset M there are at most $\binom{m}{\alpha n}$ many choices for the subset X. Hence, the probability that the lemma fails to hold is at most

$$(n+1)!(en)^m \binom{m}{\alpha n} \exp(-\Omega(n^2)) = o(1),$$

as claimed. $\qquad\square$

A third application, of a slightly different flavour, concerns counting small subgraphs. In $G(n,p)$ and Q_p^n, a matching lower bound on the size of the largest component in the subcritical regime and the second largest component in the supercritical regime can be given by using Lemma 2.3 and a second moment argument. However, the lower bound in Lemma 2.3 is ineffective once we start considering trees whose order is linear in n. Using PFS we can give an asymptotically matching lower bound for trees of almost quadratic size.

Lemma 4.9. *Let* $1 \ll m \ll \left(\frac{n}{\log n}\right)^2$ *and let* $t_m(G)$ *be the number of rooted m-vertex trees in* $\mathrm{Perm}(n)$. *Then,*

$$t_m(G) \geqslant (1-o(1))^m (n+1)! \frac{m^{m-2} n^{m-1}}{(m-1)!} = e^{o(m)}(n+1)!(en)^{m-1}$$

Proof. We follow the proof of [11, Lemma 2], replacing their use of BFS with PFS. Call a m-vertex rooted, labelled tree *typical* if its maximum degree is at most $2\frac{\log m}{\log \log m}$ and its depth is at most $\sqrt{m} \log \log m$.

We first note that by Lemmas 2.4 and 2.5, almost every spanning tree of K_m is typical. Therefore, by Cayley's formula, the family $\mathcal{T}$ of typical spanning trees on vertex set $[m]$ rooted at m satisfies

$$|\mathcal{T}| = (1+o(1))m^{m-2}.$$

Fix a vertex v of $\mathrm{Perm}(n)$. We count the number of pairs (T, ϕ) where $T \subset \mathrm{Perm}(n)$ is a typical m-vertex tree rooted at v and $\phi \colon [m] \to V(T)$ is a bijection with $\phi(m) = v$. Clearly each typical tree rooted at v is contained in precisely $(m-1)!$ such pairs, and each pair (T, ϕ) determines a typical tree $T' \subset K_m$ such that $(x,y) \in E(T')$ if and only if $(\phi(x), \phi(y)) \in E(T)$.

We will now fix a typical tree T' with vertex set $[m]$ and construct many pairs (T, ϕ) corresponding to T'. To do so, we will define the map ϕ by exploring the tree T', starting at the root m, layer by layer using projection-first search.

Suppose we have a partial embedding ϕ of the first i layers of T' into $\mathrm{Perm}(n)$, where $F_i \subseteq V(T')$ is the i-th layer, and furthermore we have a disjoint family $\{Q(x) \colon x \in F_i\}$ of faces of $\mathrm{Perm}(n)$ whose dimension is at least $n - w(v)$, where $w(v)$ is defined as in (3). For the initial step we can take $\phi(m) = v$ and $Q(m) = \mathrm{Perm}(n)$.

For each vertex $y \in F_{i+1}$ which is adjacent to $x \in F_i$, we choose an arbitrary neighbour of $\phi(x)$ in $Q(x)$ to assign as $\phi(y)$. Then, for each $x \in F_i$, we apply Lemma 4.1 to find an appropriate family of projections $\{Q(y) \colon y \in N(x) \cap F_{i+1}\}$.

16

At end of this process we see that for each vertex apart from the root we had at least $n - \max_{v \in T} w(v) \geqslant n - 2\sqrt{m} \log m = (1 - o(1))n$ choices for the embedding, and hence the total number of pairs we can build in this way is at least

$$|\mathcal{T}| \cdot \left((1 - o(1))n \right)^{m-1} \geqslant (1 - o(1))^m m^{m-2} n^{m-1}.$$

The claim follows then by summing over all choices of $v \in V(\mathrm{Perm}(n))$. $\qquad \square$

5. The percolation threshold

We first note that the upper bound on the order of the largest component in the subcritical regime follows directly from known results on percolation on regular graphs, see [52, Proposition 1] or [24, Theorem 1]. We now show that, in both regimes, it follows from Lemma 4.9 that there are tree components of order $\Omega(n \log n)$.

Lemma 5.1. *For every $c > 0$ there exists $\alpha > 0$ such that the following holds. For $p = \frac{c}{n}$, whp $\mathrm{Perm}(n)_p$ contains a tree component of order $\alpha n \log n$.*

Proof. The proof follows by a standard second moment calculation. Indeed, letting X be the number of tree components of order $m = \alpha n \log n \ll \left(\frac{n}{\log n} \right)^2$, by Lemma 4.9 for any $\alpha \in (0, 1)$

$$\mathbb{E}(X) \geqslant t_m(G) p^{m-1} (1 - p)^{nm}$$

$$\geqslant (n + 1)! (ec)^{m-1} \exp\left(-\frac{cm}{2} + o(m) \right)$$

$$= \exp\left(n \log n \left(1 + \alpha \left(1 + \log c - \frac{c}{2} \right) + o(1) \right) \right).$$

In particular, we can choose $\alpha > 0$ sufficiently small such that $\alpha \left(1 + \log c - \frac{c}{2} \right) > -\frac{1}{4}$ and so

$$\mathbb{E}(X) \geqslant \exp\left(\frac{3}{4} n \log n \right).$$

On the other hand, if we consider X as a sum of indicator random variables, the only pairs (T_1, T_2) of trees with non-zero covariance are those with $V(T_1)$ and $V(T_2)$ disjoint, but connected by an edge in $\mathrm{Perm}(n)$. In particular, $V(T_1) \cup V(T_2)$ is a connected set of order $2m$ and hence by Lemma 2.3 the variance is at most

$$(n + 1)! (en)^{2m-1} \leqslant \exp\left((1 + 2\alpha) n \log n \right) = o\left(\mathbb{E}(X)^2 \right),$$

and the result follows by Chebyshev's inequality. $\qquad \square$

In particular, to prove Theorem 1.6 it remains to show the existence of a unique component of order $\omega(n \log n)$ in the supercritical regime, and to estimate its order. To this end, let us fix $c > 1$, $\varepsilon = c - 1$ and $p = \frac{c}{n} = \frac{1+\varepsilon}{n}$. We will argue using a multi-round exposure, or sprinkling, argument. Let $p_1 = \frac{1+\varepsilon/2}{n}$, $p_2 = \frac{\varepsilon}{2n}$ and p_3 be such that $(1 - p_1)(1 - p_2)(1 - p_3) = 1 - p$. It is easy to check that $p_3 = \Omega\left(n^{-2} \right)$ and that $(1 - p_1)(1 - p_2) = 1 - \frac{c+o(1)}{n}$. Let us define

$$G_1 = \mathrm{Perm}(n)_{p_1}, \quad G_2 = G_1 \cup \mathrm{Perm}(n)_{p_2} \quad \text{and} \quad G_3 = G_2 \cup \mathrm{Perm}(n)_{p_3},$$

so that $G_3 \sim \mathrm{Perm}(n)_p$. Recalling that $r = n^{19}$, let us further define $W_i = V_{\geqslant r}(G_i)$ and note that $W_1 \subseteq W_2 \subseteq W_3$.

To prove Theorem 1.6 in the supercritical regime, we will show that W_3 has the correct order and is connected, and that all other components of G_3 have size $O(n \log n)$. The first step is the lemma below.

17

Lemma 5.2. *With high probability, $|W_3| = (\gamma(c) - o(1))(n+1)!$.*

Proof. Since $G_3 \sim \mathrm{Perm}(n)_p$, we may apply Lemma 4.4 and obtain that, for $v \in V(\mathrm{Perm}(n))$,

$$\mathbb{P}(v \in W_3) \geqslant \gamma(c) - o(1).$$

Conversely, since $\mathrm{Perm}(n)$ is n-regular, for every $v \in V(\mathrm{Perm}(n))$ we can couple a BFS exploration process in G_3 from above with a Galton-Watson tree with offspring distribution $\mathrm{Bin}(n, p)$. In particular, Lemma 2.6 iv implies that for every $v \in V(G)$, $\mathbb{P}(v \in W_3) \leqslant \gamma(c) + o(1)$. Hence $\mathbb{E}[|W_3|] = (\gamma(c) - o(1))(n+1)!$.

It remains to show that $|W_3|$ is well concentrated about its mean. To this end, let us consider the indicator random variables $\{X_e \colon e \in E(\mathrm{Perm}(n))\}$ for each edge in G_3. Clearly $|W_3|$ is a function of this set of random variables, and it is clear that changing a single one of the X_e can change $|W_3|$ by at most $2r$. Hence, by the Azuma–Hoeffding inequality (Lemma 2.8),

$$\mathbb{P}\left(\big||W_3| - \mathbb{E}\big[|W_3|\big]\big| \geqslant ((n+1)!)^{2/3}\right) \leqslant 2\exp\left(-\frac{((n+1)!)^{4/3}}{8|E(\mathrm{Perm}(n))|r^2}\right).$$

Since $|E(\mathrm{Perm}(n))| = n(n+1)!/2$ and $r = n^{19}$, the right-hand side tends to zero, proving that $|W_3|$ is concentrated around its expected value and finishing the proof of the lemma. $\qquad\square$

It remains to show that W_3 is connected in G_3. We start by proving that the last sprinkling step merges all large components of G_2, that is, that $G_3[W_2]$ is connected. Given a graph G, we say a partition $\{A_1, A_2\}$ of a set $X \subseteq V(G)$ with $A_1 \neq \emptyset \neq A_2$ is G-*component respecting* if no component of G meets both A_1 and A_2.

Lemma 5.3. *Whp for every G_2-component respecting partition $\{A_1, A_2\}$ of W_2 there is a path in G_3 between A_1 and A_2.*

Proof. We first expose G_2 and note that by Lemma 4.7 whp every vertex in $\mathrm{Perm}(n)$ is within distance two of W_2. We will assume in what follows that this holds deterministically.

Let us fix a G_2-component respecting partition $\{A_1, A_2\}$ of W_2, where without loss of generality $a := |A_1| \leqslant |A_2|$. Note that, since all components of $G_2[W_2]$ have size at least r, it follows that $a \geqslant r$. By our assumption, we can extend $\{A_1, A_2\}$ to a partition $\{A_1', A_2'\}$ of $V(\mathrm{Perm}(n))$ such that every vertex in A_i' is within distance two of A_i for $i = 1, 2$, where we note that $|A_i'| \geqslant |A_i| \geqslant r$ for $i = 1, 2$.

By Proposition 1.9 there are at least

$$\Omega\left(\frac{\min\{|A_1'|, |A_2'|\}}{n^2}\right) = \Omega\left(\frac{a}{n^2}\right)$$

edges in $\mathrm{Perm}(n)$ between A_1' and A_2'. By construction we can extend these edges to a (not necessarily disjoint) family of $A_1 - A_2$ paths of length at most five. Then, since $\Delta(\mathrm{Perm}(n)) = n$, each path shares an edge with at most $5n^4$ other paths and so we can naively thin this family out to an edge-disjoint family of size $\Omega\left(\frac{a}{n^6}\right)$. In particular, since $p_3^5 = \Omega\left(n^{-10}\right)$, the probability that none of these paths are contained in $\mathrm{Perm}(n)_{p_3}$ is at most

$$(1 - p_3^5)^{\Omega(a/n^6)} = \exp\left(-\Omega\left(\frac{a}{n^{16}}\right)\right).$$

18

Since W_2 contains at most $(n+1)!$ components, and A_1 is the union of at most $\frac{a}{r}$ of them, there are at most

$$((n+1)!)^{a/r} \leqslant \exp\left(\frac{an^2}{r}\right)$$

G_2-component respecting partitions $\{A_1, A_2\}$ of W_2 with $|A_1| = a$. Hence, since $r = n^{19}$, by the union bound the probability that the statement holds is at most

$$\sum_{a=r}^{|W_2|/2} \exp\left(\frac{an^2}{r}\right) \exp\left(-\Omega\left(\frac{a}{n^{16}}\right)\right) \leqslant \sum_{a \geqslant r} \exp\left(-\Omega\left(\frac{a}{n^{16}}\right)\right) = o(1). \qquad \square$$

Recall that we also want to show that the second largest component has size $O(n \log n)$. In principle, there could exist components larger than that in $G_3[W_1^{\mathsf{c}}]$, either created by the sprinkling steps or pre-existing. However, since p_2 is large, the sprinkled edges on the boundary of W_1 will necessarily merge these with components of W_1. This is made precise by the lemma below.

Lemma 5.4. *There exists a constant $C = C(\varepsilon)$ such that whp every component of G_3 of size at least $Cn \log n$ meets W_1.*

Proof. We first expose G_1, choose $C = C(\varepsilon)$ sufficiently large and $\alpha \ll \varepsilon$ sufficiently small. Let $W_1^{\mathsf{c}} = V(\mathrm{Perm}(n)) \setminus W_1$ and

$$X = \left\{v \in W_1^{\mathsf{c}} : \left|W_1 \cap N_{\mathrm{Perm}(n)}(v)\right| > \alpha n\right\}.$$

By Lemma 4.8 whp every subset $M \subseteq W_1^{\mathsf{c}}$ of size $Cn \log n$ which is connected in $\mathrm{Perm}(n)$ is such that $|M \setminus X| \leqslant \alpha n$. In particular, if $M \subseteq W_1^{\mathsf{c}}$ is a connected subset of $\mathrm{Perm}(n)$ of order at least $Cn \log n$ then

$$\left|E_{\mathrm{Perm}(n)}(M, W_1)\right| \geqslant (Cn \log n - \alpha n) \cdot \alpha n \geqslant \frac{C}{2} \alpha n^2 \log n. \qquad (4)$$

We assume in what follows that this holds deterministically.

We now expose $\mathrm{Perm}(n)_{p_2}$ and $\mathrm{Perm}(n)_{p_3}$ on W_1^{c}. We note that any component of G_3 which avoids W_1 is then a component M of $G_3[W_1^{\mathsf{c}}]$ which is not adjacent to W_1, and we have yet to expose the edges between M and W_1 in $\mathrm{Perm}(n)_{p_2}$ and $\mathrm{Perm}(n)_{p_3}$.

In particular, for each component M of $G_3[W_1^{\mathsf{c}}]$ of size at least $Cn \log n$, since M spans a connected subset of $\mathrm{Perm}(n)$, by (4) the probability that there are no edges from M to W_1 in G_2 is at most

$$(1 - p_2)^{(C\alpha n^2 \log n)/2} \leqslant \exp\left(-\frac{C\alpha\varepsilon}{4} \cdot n \log n\right).$$

However, there are at most $(n+1)!$ components of $G_3[W_1^{\mathsf{c}}]$, and so by the union bound the probability that the conclusion of the lemma fails to holds is at most

$$(n+1)! \exp\left(-\frac{C\alpha\varepsilon}{4} n \log n\right) = o(1),$$

as long as $C = C(\varepsilon)$ is sufficiently large. $\qquad \square$

We now have all the tools to prove Theorem 1.6 in the supercritical case.

Proof of Theorem 1.6 (ii). By Lemma 5.3, whp $G_3[W_2]$ is connected. But by Lemma 5.4 whp every component in W_3 meets $W_1 \subseteq W_2$, and hence contains a component of W_2. Therefore, $L_1 := G_3[W_3]$ is connected. By Lemma 5.2, whp $|L_1| = |W_3| = (\gamma(c) + o(1))(n+1)!$.

Furthermore, by Lemma 5.1 whp G_3 contains a component of order $\alpha n \log n$, whereas by Lemma 5.4 whp there are no components of G_3 of order at least $Cn \log n$ which do not meet $W_1 \subseteq W_3 = V(L_1)$, and hence the second largest component of G_3 has order $\Theta(n \log n)$. $\qquad\square$

Remark 5.5. *We note that in fact, since the proofs of Lemmas 5.3 and 5.4 only require a lower bound on p_3, the same argument will show that for any (potentially growing) $c > 1$ and any $p \geqslant \frac{c}{n}$ whp $\mathrm{Perm}(n)_p$ has a unique component of order at least $Cn \log n$.*

6. The connectivity threshold

We move on to determining the connectivity threshold in $\mathrm{Perm}(n)_p$. We begin by showing that, already well before the threshold in Theorem 1.7, all components in $\mathrm{Perm}(n)_p$ are either isolated vertices, or exponentially large.

Lemma 6.1. *Let $p \geqslant 1 - \frac{3}{n}$. Then whp $\mathrm{Perm}(n)_p$ has no components of order $2 \leqslant k \leqslant 2^{n/4}$.*

Proof. If K is a component of size $k \leqslant 2^{n/4}$, then K contains a tree T of order k, all of whose edges are present in $\mathrm{Perm}(n)_p$ and such that each edge in $\partial(V(K))$ is not present in $\mathrm{Perm}(n)_p$. Note that, by Proposition 1.9

$$|\partial(V(K))| \geqslant k(n - \log k) \geqslant \frac{3kn}{4}.$$

Hence, by Lemma 2.3, the probability that there is a component of order k with $2 \leqslant k \leqslant 2^{n/4}$ is at most

$$\sum_{k=2}^{2^{n/4}} (n+1)!(en)^{k-1} p^{k-1}(1-p)^{3kn/4} \leqslant \sum_{k=2}^{2^{n/4}} (n+1)!(en)^k \left(\frac{3}{n}\right)^{3kn/4}$$

It is easy to check that summand on the right-hand side is decreasing in k. Therefore, since $n! \leqslant n^n$, the right-hand side is at most

$$\exp\left(n \log n - \frac{6n}{4}\log n + O(n)\right),$$

which tends to zero. $\qquad\square$

Proof of Theorem 1.7. We note first that, by Lemma 6.1 and Remark 5.5 whp there is a unique component of order at least 2. In particular,

$$\mathbb{P}(\mathrm{Perm}(n)_p \text{ is connected}) = \mathbb{P}(\mathrm{Perm}(n)_p \text{ contains no isolated vertices}) + o(1).$$

If we let X be the number of isolated vertices in $\mathrm{Perm}(n)_p$, it is a simple exercise to estimate the moments of X. Indeed, given any $r \in \mathbb{N}$ and a subset $S \subseteq \mathrm{Perm}(n)$ of size r, by Proposition 1.9 there are at most rn and at least $r(n - \log r)$ edges meeting S. Furthermore, there are at most $((n+1)!)_{r-1} \cdot rn$ sets of size r meeting fewer than rn edges, where $x_r = x(x-1)\cdots(x-r+1)$ is the r-th falling factorial. It follows that

$$((n+1)!)_r(1-p)^{rn} \leqslant \mathbb{E}[X_r] \leqslant ((n+1)!)_r(1-p)^{rn} + ((n+1)!)_{r-1}rn(1-p)^{r(n-\log r)}.$$

Recalling that $\lambda = (n+1)!(1-p)^n$, elementary estimates lead to

$$(1 - o(1))\lambda^r = \lambda^r\left(1 - \frac{r^2}{(n+1)!}\right) \leqslant \mathbb{E}[X_r] \leqslant \lambda^r\left(1 + \frac{rn}{(n+1)!(1-p)^{r\log r}}\right) = (1 + o(1))\lambda^r.$$

20

In particular, if $\lambda \to c \in \mathbb{R}$, then by the method of moments (see for example [41]) X tends in distribution to a Poisson distribution with mean c and so

$$\mathbb{P}(\mathrm{Perm}(n) \text{ is connected}) = \mathbb{P}(X = 0) \xrightarrow{n \to \infty} \mathbb{P}(\mathrm{Po}(c) = 0) = e^{-c}.$$

Otherwise, in the case $\lim_{n \to \infty} \lambda = \infty$ or $\lim_{n \to \infty} \lambda = 0$, the result follows by a simple first or second moment argument, respectively. $\qquad\square$

It is relatively straightforward to use these ideas to give a corresponding *hitting time* result for the random graph process on $\mathrm{Perm}(n)$. That is, if we build a random subgraph of $\mathrm{Perm}(n)$ by adding one edge at a time, each time choosing uniformly at random from the remaining edges, then whp this subgraph will become connected precisely when the last isolated vertex disappears.

7. Discussion

As we saw in Sections 3 and 4, much of the work in this paper generalises to the case of Cayley graphs of Coxeter groups or zonotopes. However, a key missing part is some control over the large scale isoperimetric properties of these graphs, as in Proposition 1.9. In the case of the permutahedron, spectral methods give a close to optimal bound on the Cheeger constant, up to a small polynomial factor in the dimension/regularity. This is then a key ingredient to the proofs of Theorems 1.6 and 1.7, allowing us to merge the large clusters we grow in a sprinkling step. However, a much weaker bound on the Cheeger constant, for example anything sub-exponential, would be sufficient in this argument.

It is therefore interesting to ask if we can give any general bounds on the Cheeger constant of these classes of graphs. In the case of zonotopes, we make a rather bold conjecture.

Conjecture 7.1. *Let $Z = \sum_{i=1}^{k} \boldsymbol{v_i}$ be a zonotope and let G be its skeleton. Then $i(G)$ is at worst inverse polynomial in k.*

Here, it would seem that the 'worst' examples should come from cycles of length $2k$, which have Cheeger constant around $\frac{2}{k}$. There is a similar, well-known conjecture of Mihail and Vazirani [50] on the expansion of $0/1$ polytopes, where Q^d is conjectured to have the worst expansion. The specific case of matroid basis polytopes was solved in a breakthrough result very recently [6].

More generally, it would be interesting to know how the phase transition at the percolation threshold develops in other classes of polytopes – how the critical probability relates to the degree distribution and the largest cluster sizes in the sub- and super-critical regimes. If we consider the cartesian powers of some fixed low-dimensional polytope with an irregular 1-skeleton, the negative answer to Question 4.6 shows that we do not always have this dichotomy between logarithmic and linear sized clusters. However, it would be very interesting to know if the quantitative behaviour in Theorems 1.1, 1.2 and 1.6 was perhaps universal to *simple* polytopes.

In [26, 31], similar methods to those in Section 5 are used to not only show the existence of a giant component in the supercritical percolated graphs, but to demonstrate that this giant component likely has good expansion properties, from bounds on various structural parameters of the giant component such as its diameter, circumference and mixing time can be derived.

Question 7.2. *What can we say about the structural properties of the giant component L of a supercritical percolated permutahedron?*

(1) Does L whp contain a path of length $\Theta((n + 1)!)$?

(2) What is the likely diameter of L?

(3) What is the likely mixing time of a lazy random walk on L?

Using the methods of [26, 31], it should be relatively straightforward to obtain *some* lower bound on the expansion of the giant component in $\mathrm{Perm}(n)$ using these methods, in particular one which is inverse polynomial in n, which would give answers to Question 7.2 (1)–(3) which are tight up to some power of n. However, it seems unlikely that one could obtain an optimal bound (even up to polylogarithmic factors) as in [26] via these methods without first understanding better the isoperimetric properties of $\mathrm{Perm}(n)$, in particular for large sets. For this reason, we have not made much attempt to optimise the quantitative aspects of the arguments in Section 5. Furthermore, for the diameter and mixing time, it is not obvious what natural lower bounds there are for these quantities. In the case of $G(n,p)$ and Q_p^n, similar arguments as in Lemma 5.1 show the likely existence of a bare path of length $\Omega(\log n)$ and $\Omega(n)$ respectively, leading to natural lower bounds on the diameter and mixing time in these cases. However, the methods of Lemma 5.1 only apply to trees with small depth and maximum degree, and so in particular are not effective for counting paths in $\mathrm{Perm}(n)$. For this reason, it would be interesting to know what the length of the longest bare path in $\mathrm{Perm}(n)_p$ is, and to this end to effectively estimate the number of paths of length $\Theta(n \log n)$ in $\mathrm{Perm}(n)$. From the other side, the diameter of $\mathrm{Perm}(n)$ is clearly $\binom{n}{2}$ and the mixing time of the lazy random walk on $\mathrm{Perm}(n)$ is known to be $\Theta(n^2 \log n)$ [44, 56]. Whilst neither parameter is increasing under taking subgraphs, analogues to the case of $G(n,p)$ and Q_p^n suggest that in the supercritical regime we should expect the giant component to have a diameter and mixing time which is slightly larger (by a logarithmic factor) than that of the host graph.

As mentioned above, whilst we can determine quite precisely the expansion of small sets in $\mathrm{Perm}(n)$, the isoperimetric properties of $\mathrm{Perm}(n)$ in general seem to not be very well-understood. We discussed the edge-isoperimetric problem in Section 3, but the *vertex-isoperimetric problem* in $\mathrm{Perm}(n)$ is also very interesting. A reasonably natural conjecture here would be that vertex-boundary is minimised by Hamming balls in Kendall's τ-metric (the distance metric on S_{n+1} induced by $\mathrm{Perm}(n)$).

Moving on to the connectivity threshold, in both $G(n,p)$ and Q_p^n, it has been shown that the connectivity threshold is asymptotically the same as the threshold for containing a perfect matching and Hamiltonian cycle. In the case of the hypercube, both the connectivity threshold and the threshold for the existence of a perfect matching can be shown using quite delicate but elementary first moment method arguments, which rely on strong and explicit isoperimetric inequalities for the hypercube, in particular Theorem 1.8. On the other hand, the threshold for containing a Hamiltonian cycle in this model was only recently resolved, in groundbreaking work by Condon, Espuny Díaz, Girão, Kühn and Osthus [21].

However, in the case of the permutahedron, even the existence of a Hamiltonian cycle in the host graph is non-trivial (see, e.g., [42]).

Question 7.3. *What is the threshold for the existence of a perfect matching or Hamiltonian cycle in $\mathrm{Perm}(n)_p$? Do they coincide with the connectivity threshold?*

Finally, viewing the permutahedron as a $\mathcal{V}$-polytope rather than as a graph, there is perhaps another natural notion of a random substructure of $\mathrm{Perm}(n)$ given by taking a random subset of the permutations, given by including each permutation independently and with probability p, and considering the convex hull of these points, which we denote by $P(n,p)$. From the other

direction, considering the permutahedron as a $\mathcal{H}$-polytope, we could equally consider the constraints determining the permutahedron as an intersection of hyperplanes and choose a random subset of these constraints.

Question 7.4. *Let $p \in (0,1)$.*

    (1) What is the likely volume of $P(n,p)$?
    (2) What is the likely number of lattice points inside $P(n,p)$?
    (3) What is the likely number of facets of $P(n,p)$?
    (4) What is the likely number of edges of $P(n,p)$?

In the case of $P(n) = P(n,1)$ these parameters have simple combinatorial interpretations, and it would interesting to see if there are combinatorial, or stochastically combinatorial, interpretations of the corresponding parameters in $P(n,p)$. The model is also closely related to the model of random 0/1 polytopes considered in [45], where they showed that a weaker form of Mihail and Vazirani's conjecture holds for almost all 0/1 polytopes. It would interesting to know if the typical expansion in $P(n,p)$ is always smaller than that of $P(n)$, or even just if it can be bounded as an inverse polynomial in n.

REFERENCES

[1] R. Ahlswede and S. L. Bezrukov. Edge isoperimetric theorems for integer point arrays. *Appl. Math. Lett.*, 8(2):75–80, 1995.

[2] M. Ajtai, J. Komlós, and E. Szemerédi. Largest random component of a k-cube. *Combinatorica*, 2(1):1–7, 1982.

[3] N. Alon. Eigenvalues and expanders. *Combinatorica*, 6:83–96, 1986.

[4] N. Alon and V. D. Milman. λ_1, isoperimetric inequalities for graphs, and superconcentrators. *J. Comb. Theory, Ser. B*, 38:73–88, 1985.

[5] N. Alon and J. H. Spencer. *The probabilistic method*. Hoboken, NJ: John Wiley & Sons, fourth edition, 2016.

[6] N. Anari, K. Liu, S. O. Gharan, and C. Vinzant. Log-concave polynomials. II: High-dimensional walks and an FPRAS for counting bases of a matroid. *Ann. Math. (2)*, 199(1):259–299, 2024.

[7] R. Bacher. Valeur propre minimale du laplacien de coxeter pour le groupe symétrique. *J. Algebra*, 167(2):460–472, 1994.

[8] B. Barber and J. Erde. Isoperimetry in integer lattices. *Discrete Anal.*, 7(2018):1–16, 2018.

[9] I. Benjamini, G. Kozma, and N. Wormald. The mixing time of the giant component of a random graph. *Random Structures Algorithms*, 45(3):383–407, 2014.

[10] A. J. Bernstein. Maximally connected arrays on the n-cube. *SIAM J. Appl. Math.*, 15:1485–1489, 1967.

[11] A. Beveridge, A. Frieze, and C. McDiarmid. Random minimum length spanning trees in regular graphs. *Combinatorica*, 18(3):311–333, 1998.

[12] A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. Ziegler. *Oriented matroids.*, volume 46 of *Encycl. Math. Appl.* Cambridge: Cambridge University Press, 2nd ed. edition, 1999.

[13] R. Blind and P. Mani-Levitska. Puzzles and polytope isomorphisms. *Aequationes Math.*, 34:287–297, 1987.

[14] B. Bollobás. The evolution of random graphs. *Trans. Am. Math. Soc.*, 286:257–274, 1984.

[15] B. Bollobás, Y. Kohayakawa, and T. Łuczak. The evolution of random subgraphs of the cube. *Random Structures Algorithms*, 3(1):55–90, 1992.

[16] B. Bollobás and I. Leader. Edge-isoperimetric inequalities in the grid. *Combinatorica*, 11(4):299–314, 1991.

[17] S. B. Broadbent and J. M. Hammersley. Percolation processes. I: Crystals and mazes. *Proc. Camb. Philos. Soc.*, 53:629–641, 1957.

[18] J. D. Burtin. The probability of connectedness of a random subgraph of an n-dimensional cube. *Problemy Peredaci Informacii*, 13(2):90–95, 1977.

[19] J. Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. Probl. Analysis, Sympos. in Honor of Salomon Bochner, Princeton Univ. 1969, 195-199 (1970)., 1970.

[20] M. Collares, J. Erde, A. Geisler, and M. Kang. Majority bootstrap percolation on high-dimensional geometric graphs. in preparation.

[21] P. Condon, A. Espuny Díaz, A. Girão, D. Kühn, and D. Osthus. Hamiltonicity of random subgraphs of the hypercube. *arXiv preprint arXiv:2007.02891*, 2020.

[22] M. W. Davis. *The geometry and topology of Coxeter groups.*, volume 32 of *Lond. Math. Soc. Monogr. Ser.* Princeton, NJ: Princeton University Press, 2008.

[23] J. Ding, E. Lubetzky, and Y. Peres. Anatomy of the giant component: the strictly supercritical regime. *Eur. J. Comb.*, 35:155–168, 2014.

[24] S. Diskin, J. Erde, M. Kang, and M. Krivelevich. Percolation on high-dimensional product graphs. *arXiv preprint arXiv:2211.10139*, 2022.

[25] S. Diskin, J. Erde, M. Kang, and M. Krivelevich. Percolation on irregular high-dimensional product graphs. *Combin. Probab. Comput.*, 33(3):377–403, 2024.

[26] S. Diskin, J. Erde, M. Kang, and M. Krivelevich. Isoperimetric inequalities and supercritical percolation on high-dimensional product graphs. *Combinatorica*, to appear.

[27] S. Diskin and A. Geisler. Perfect matching in product graphs and in their random subgraphs. *arXiv preprint arXiv:2404.14020*, 2024.

[28] S. Diskin and M. Krivelevich. Supercritical site percolation on the hypercube: small components are small. *Comb. Probab. Comput.*, 32(3):422–427, 2023.

[29] S. Diskin and M. Krivelevich. Hitting time of connectedness in the random hypercube process. *arXiv preprint arXiv:2404.09289*, 2024.

[30] J. Erde, M. Kang, and M. Krivelevich. Large complete minors in random subgraphs. *Comb. Probab. Comput.*, 30(4):619–630, 2021.

[31] J. Erde, M. Kang, and M. Krivelevich. Expansion in supercritical random subgraphs of the hypercube and its consequences. *The Annals of Probability*, 51:127–156, 2023.

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

[33] P. Erdös and J. Spencer. Evolution of the n-cube. *Comput. Math. Appl.*, 5(1):33–39, 1979.

[34] A. Frieze, M. Krivelevich, and R. Martin. The emergence of a giant component in random subgraphs of pseudo-random graphs. *Random Structures Algorithms*, 24(1):42–50, 2004.

[35] M. R. Garey, D. S. Johnson, and L. Stockmeyer. Some simplified NP-complete graph problems. *Theor. Comput. Sci.*, 1:237–267, 1976.

[36] F. Harary, J. P. Hayes, and H.-J. Wu. A survey of the theory of hypercube graphs. *Comput. Math. Appl.*, 15(4):277–289, 1988.

[37] L. H. Harper. Optimal assigments of numbers to vertices. *SIAM J. Appl. Math.*, 12:131–135, 1964.

[38] L. H. Harper. *Global methods for combinatorial isoperimetric problems*, volume 90 of *Camb. Stud. Adv. Math.* Cambridge: Cambridge University Press, 2004.

[39] T. E. Harris. *The theory of branching processes*, volume 119 of *Grundlehren Math. Wiss.* Springer, Cham, 1963.

[40] S. Hart. A note on the edges of the n-cube. *Discrete Math.*, 14:157–163, 1976.

[41] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[42] S. M. Johnson. Generation of permutations by adjacent transposition. *Mathematics of computation*, 17(83):282–285, 1963.

[43] G. Kalai. A simple way to tell a simple polytope from its graph. *J. Comb. Theory, Ser. A*, 49(2):381–383, 1988.

[44] H. Lacoin. Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion. *Ann. Probab.*, 44(2):1426–1487, 2016.

[45] B. Leroux and L. Rademacher. Expansion of random 0/1 polytopes, 2022.

[46] V. F. Lev. Edge-isoperimetric problem for Cayley graphs and generalized Takagi functions. *SIAM J. Discrete Math.*, 29(4):2389–2411, 2015.

[47] J. H. Lindsey. Assigment of numbers to vertices. *Amer. Math. Monthly*, 71:508–516, 1964.

[48] T. Łuczak. Component behavior near the critical point of the random graph process. *Random Structures Algorithms*, 1(3):287–310, 1990.

[49] T. Łuczak, B. Pittel, and J. C. Wierman. The structure of a random graph at the point of the phase transition. *Trans. Am. Math. Soc.*, 341(2):721–748, 1994.

[50] M. Mihail. On the expansion of combinatorial polytopes. In *Mathematical foundations of computer science 1992. 17th international symposium, Prague, Czechoslovakia, August 24–28, 1992. Proceedings*, pages 37–49. Berlin: Springer, 1992.

[51] J. W. Moon. On the maximum degree in a random tree. *Michigan Math. J.*, 15(4):429–432, 1968.

[52] A. Nachmias and Y. Peres. Critical percolation on random regular graphs. *Random Structures Algorithms*, 36(2):111–148, 2010.

[53] A. Rényi and G. Szekeres. On the height of trees. *J. Aust. Math. Soc.*, 7:497–507, 1967.

[54] A. A. Sapoženko. Metric properties of almost all functions of the algebra of logic. *Diskret. Analiz*, 10:91–119, 1967.

[55] J. P. Tillich. Edge isoperimetric inequalities for product graphs. *Discrete Math.*, 213(1-3):291–320, 2000.

[56] D. B. Wilson. Mixing times of lozenge tiling and card shuffling Markov chains. *Ann. Appl. Probab.*, 14(1):274–325, 2004.

[57] G. M. Ziegler. *Lectures on polytopes*, volume 152 of *Grad. Texts Math.* Berlin: Springer-Verlag, 1995.

Appendix A. Proof of Lemma 4.5

We start with a simple lemma about the expectation of a truncated binomial random variable, which follows for example from [30, Lemma 10].

Lemma A.1. *Let $\beta > 0$ be sufficiently small, let $K = 4 \ln \frac{1}{\beta}$, let $m' \geqslant \left(1 - \frac{\beta}{2}\right) m$ and let $p = \frac{1+\beta}{m}$. If $X \sim \text{Bin}(m', p)$ is a binomial random variable and $Y = \min\{X, K\}$, then*

$$\mathbb{E}[Y] \geqslant 1 + \frac{\beta}{4}.$$

The algorithm PFS$'$ can then be described as follows: We first run the PFS exploration process starting at v for $\tau_1 := \log \log m$ many rounds and let $A(\tau_1)$ be the frontier.

As in the proof of Lemma 4.4, whp for each $x \in A(\tau_1)$, $w(x) \leqslant \log^2 m$ and hence we can couple this process from below with a Galton-Watson tree with child distribution $\text{Bin}(m - \log^2 m, p)$. In particular, by Lemma 2.6, $|A(\tau_1)| = (c + o(1))^{\log \log m} \geqslant \log \log m$ with probability $\gamma(1 + \beta) + o(1)$.

We now continue with a slightly modified PFS process, starting with frontier $A(\tau_1)$, except when we expose the neighbourhood $N_t(x)$ of a vertex x we stop after we have found $K := 4 \ln \frac{1}{\beta}$ neighbours.

Let us suppose we run this latter step for a further $\tau_2 := \beta^2 m$ rounds. For each $x \in A(\tau_1 + \tau_2)$, we have $w(x) \leqslant \log^2 m + K\tau_2 \leqslant 2K\beta^2 m \leqslant \frac{\beta}{2}m$, since β is small. Hence, during this second round, whenever we exposed the neighbourhood of a vertex y, the size of the neighbourhood $N_t(y)$ we discovered stochastically dominates a random variable distributed as $\min\{\text{Bin}\left(\left(1 - \frac{\beta}{2}\right)m, p\right), K\}$, which by Lemma A.1 has expectation at least $1 + \frac{\beta}{4}$.

In particular, for any $\tau_1 \leqslant t < \tau_1 + \tau_2$, $|A(t+1)|$ stochastically dominates the sum of $|A(t)|$ independent variables, each with mean $1 + \frac{\beta}{4}$ and bounded by K. Therefore, if we let $\alpha := 1 + \frac{\beta}{8}$, we obtain by the Azuma-Hoeffding inequality (Lemma 2.8) that

$$\mathbb{P}\Big(|A(t+1)| \leqslant \alpha k \,\Big|\, |A(t)| = k\Big) \leqslant \exp\left(-\Omega(k)\right).$$

We may therefore bound

$$\mathbb{P}\big(|A(\tau_1 + \tau_2)| \leqslant (\log \log m)\alpha^{\tau_2}\big)$$

$$\leqslant \sum_{t=0}^{\tau_2 - 1} \mathbb{P}\Big(|A(t + \tau_1 + 1)| \leqslant (\log \log m)\alpha^{t+1} \,\Big|\, |A(t + \tau_1)| \geqslant (\log \log m)\alpha^t\Big)$$

$$\leqslant \sum_{t=0}^{\tau_2 - 1} \exp\left(-\Omega\left((\log \log m)\alpha^t\right)\right)$$

$$= o(1).$$

Therefore, whp

$$|C_v| \geqslant |A(\tau_1 + \tau_2)| \geqslant (\log \log m)\alpha^{\tau_2} = \exp\left(\Theta(m)\right),$$

finishing the proof. $\qquad\square$

Institute of Discrete Mathematics, Graz University of Technology, Steyrergasse 30, 8010 Graz, Austria

Email address: mauricio@collares.org, erde@math.tugraz.at

RobotDreams GmbH, Stremayrgasse 16/IV, 8010 Graz, Austria

Email address: nerdyjoe@gmail.com