DIAMETER OF RANDOM CUBIC SUM GRAPHS

STEPHEN J. YOUNG

ABSTRACT. A random cubic sum graph is formed by selecting each element of $\mathbb{Z}_2^k - \{\mathbf{0}\}$ with probability p and defining two elements u and v to be adjacent if $u+v$ has also been selected. Beveridge recently showed that the connectivity threshold for this random graph occurs at $p = \sqrt{\frac{\ln(n)+\ln\ln(n)}{n}}$, where $n = 2^k - 1$. We show that for p slightly above the connectivity threshold the diameter is $\Theta\left(\frac{\ln(np)}{\ln(np^2)}\right)$ and if np^2 is sufficiently large, then the diameter is concentrated on at most 3 values.

1. Introduction

The concept of a sum graph dates back to Harary [20] who considered the structure of graphs where the vertices could be labelled by elements of $\mathbb{Z}^+$ in such a manner that the edges respected the addition on $\mathbb{Z}^+$. That is, a graph G is a sum graph if there is a labelling of the vertices $\ell(\cdot)$ by elements of $\mathbb{Z}^+$ such that $u \sim V$ if and only if $\ell(u) + \ell(v) = \ell(w)$ for some vertex w. Clearly such graphs are in general disconnected and thus much of the work on sum graphs has focused on determining the number of isolated vertices it is necessary to add to a graph in order than it has a valid sum graph labelling [4, 14, 19, 25, 16]. Since Harary's work there have been several generalizations of the idea of sum graphs, most notably mod sum graphs where the addition takes place over $\mathbb{Z}_m$ [6] and f-sum graphs where the addition operation is replaced by some symmetric two variable symmetric polynomial f (that is, $u \sim v$ if and only if there is a vertex w where $f(\ell(u), \ell(v)) = \ell(w)$) [2].

Recently, Beveridge considered a randomized version of sum graphs, termed random cubic sum graphs, in [5]. In these graphs, the vertex set is a random subset S of $\mathbb{Z}_2^k - \{\mathbf{0}\}$ and two vertices $u, v \in S$ are adjacent if and only if $u + v \in S$, where the addition is over the group $\mathbb{Z}_2^k$. If each element of $\mathbb{Z}_2^k - \mathbf{0}$ is chosen independently with probability p, such a graph is denoted by $\mathcal{SG}(Q_k, p)$. In [5], Beveridge primarily studies the emergence of and disappearance of small subgraphs, providing sharp threshold for the emergence of a triangle $\{x, y, x+y\} \in S$, a butterfly $\{x, y, z, x+y, x+z\} \in S$, and the disappearance of isolated vertices. Additionally, he shows that if $n = 2^k - 1$ (the maximum possible number of vertices) and $p = \sqrt{\frac{\ln(n)}{n}}$, then with high probability $\mathcal{SG}(Q_k, p)$ consists of isolated vertices and a constant number of components of size at least $\frac{\sqrt{n\ln(n)}}{100}$. That is, it consists of a isolated vertices and a collection of components containing at least a constant fraction of the vertices. As p increases from $\sqrt{\frac{\ln(n)}{n}}$ to $\sqrt{\frac{\ln(n)+\ln(\ln(n))+c}{n}}$, the graph consists of a small number of isolated vertices and a single connected component. Finally, if $p = \sqrt{\frac{\log(n)+\ln\ln(n)+c(n)}{n}}$ with $c(n) \to \infty$ and $c(n)$ being $o(\ln\ln(n)))$ then $\mathcal{SG}(Q_k, p)$ is connected with high probability. Thus $np^2 = \ln(n) + \ln\ln(n)$ provides the sharp threshold for the connectivity $\mathcal{SG}(Q_k, p)$. In this work we use two expansion type results in order to provide upper and lower bounds for the diameter of $\mathcal{SG}(Q_k, p)$ away from this threshold. Additionally, we will show that if np^2 is sufficiently large, then the diameter is concentrated on one of a finite number of values, potentially as few as three. Before proving these results, however, we focus on the case when the diameter is at most 2.

Theorem 1. *If $p \to 0$ and $np^3 = \frac{4}{3}\ln(n) + \frac{2}{3}\ln\ln(n) + \zeta$ for any $\zeta > 0$, then the diameter of $\mathcal{SG}(Q_k, p)$ is at most 2 with probability at least $1 - e^{-\zeta(1-o(1))}$.*

Proof. For any two elements $\alpha, \beta \in \mathbb{Z}_2^k - \{\mathbf{0}\}$, we can partition $\mathbb{Z}_2^k - \{\mathbf{0}, \alpha, \beta, \alpha+\beta\}$ into sets of the form $\{x, x+\alpha, x+\beta, x+\alpha+\beta\}$. Notice that these sets are closed under addition by α and β, and hence for any element $y \in \{x, x+\alpha, x+\beta, x+\alpha+\beta\}$ the remaining elements are $y+\alpha, y+\beta, y+\alpha+\beta$. Thus if at most one element, say z, of $\{x, x+\alpha, x+\beta, x+\alpha+\beta\}$ is not a vertex, and α and β are vertices, then α and β

are connected by a path of length 2 through $z+\alpha+\beta$. As a consequence the probability that any pair of elements α and β are witnesses to the $\operatorname{diam}(\mathcal{SG}(Q_k,p)) > 2$ is $p^2(1-p)\left(1-4p^3(1-p)-p^4\right)^{\frac{n-2}{4}}$. Hence the probability that a $\operatorname{diam}(\mathcal{SG}(Q_k,p)) > 2$ is at most

$$\begin{aligned} n^2p^2\left(1-4p^3+3p^4\right)^{\frac{n-2}{4}} &= e^{2\ln(np)+\frac{n-2}{4}\ln\left(1-4p^3+3p^4\right)} \\ &\leq e^{\ln(n^2p^2)-\left(4p^3-3p^4\right)\frac{n-3}{4}} \\ &= e^{\ln(n^2p^2)-np^3+\frac{3}{4}np^4+2p^3(1-\frac{3}{4}p)} \end{aligned}$$

Now as $np^3 = \frac{4}{3}\ln(n)+\frac{2}{3}\ln\ln(n)+\zeta$ we have that $p=\sqrt[3]{\frac{\frac{4}{3}\ln(n)+\frac{2}{3}\ln\ln(n)+\zeta}{n}}$ and thus

$$\ln(n^2p^2) = \frac{4}{3}\ln(n)+\frac{2}{3}\ln\left(\frac{4}{3}\ln(n)+\frac{2}{3}\ln\ln(n)+\zeta\right) = \frac{4}{3}\ln(n)+\frac{2}{3}\ln\ln(n)+\frac{2}{3}\ln\left(\frac{4}{3}+\frac{2}{3}\frac{\ln\ln(n)}{\ln(n)}+\frac{\zeta}{\ln(n)}\right).$$

Hence the probability that the diameter is greater than 2 is at most

$$e^{\frac{2}{3}\ln\left(\frac{4}{3}+\mathcal{O}\left(\frac{\ln\ln(n)+\zeta}{\ln(n)}\right)\right)-\zeta+\mathcal{O}((\ln(n)+\zeta)p)}.$$

Now if $\zeta\in\mathcal{O}(\ln(n))$ then $\mathcal{O}\left(\frac{\ln\ln(n)+\zeta}{\ln(n)}\right)\in\mathcal{O}(1)$ and $\mathcal{O}((\ln(n)+\zeta)\,p)=\mathcal{O}(\ln(n)p)\in\mathcal{O}(1)$. Thus the probability is $e^{-\zeta+\mathcal{O}(1)}=e^{-\zeta(1-o(1))}$. On the other hand if $\zeta\gg\ln(n)$, then the probability is

$$e^{-\zeta+\frac{2}{3}\ln\left(\mathcal{O}\left(\frac{\zeta}{\ln(n)}\right)\right)+\mathcal{O}(\zeta p)} = e^{-\zeta\left(1+\frac{\frac{2}{3}\ln\left(\mathcal{O}\left(\frac{\zeta}{\ln(n)}\right)\right)}{\zeta}+\mathcal{O}(p)\right)} = e^{-\zeta(1-o(1))}.$$

□

As an immediate corollary of the proof, we have

Corollary 2. *If there is some $c>0$ such that $p\geq c$, then the diameter of $\mathcal{SG}(Q_k,p)$ is at most 2 with probability at least $1-e^{-\Theta(n)}$.*

2. Diameter Greater than 2

In analyzing the situation for diameter at most 3, it is natural to attempt to generalize the argument in Theorem 1. That is, when looking for a path between α and β, consider partitioning the elements $\mathbb{Z}_2^k-\{\mathbf{0},\alpha,\beta,\alpha+\beta\}$ into sets of the form $x,x+\alpha,x+\beta,x+\alpha+\beta$. The condition that α and β are connected by a path of length exactly 3 then transforms into the condition that there exist $x,y\in\mathbb{Z}_2^k$ such that $x+y\notin\{\mathbf{0},\alpha,\beta,\alpha+\beta\}$, where either $\{x,x+\alpha\}$ or $\{x+\beta,x+\alpha+\beta\}$ are present, and $\{y,y+\beta\}$ or $\{y+\alpha,y+\alpha+\beta\}$ are present, and one of $x+y, x+y+\alpha, x+y+\beta$, or $x+y+\alpha+\beta$ is present. Now for an appropriate range of p, it can be shown that with high probability every quartet of elements in either satisfies the first condition (we can think of these as α-type), the second condition (β-type), has only one element present, or has no elements present. Then, with an appropriate isomorphism and viewing α and β as colors, the question can be translated into whether $\mathbb{Z}_2^{k-2}-\mathbf{0}$ has any triangles satisfying a chromatic condition on the vertices. In principle, this can then be analyzed following the techniques used in [5]. However, it is clear that this approach will rapidly become infeasible because of the complicated dependencies present in longer paths, thus we turn to alternative approach modeled after the connectivity of Erdős-Rényi random graphs. Specifically, in Lemma 4 we show an expansion-type result for small subsets of vertices of $\mathcal{SG}(Q_k,p)$, then in Lemma 6 we show that any two sufficiently large subsets of vertices have a common neighbor with high probability, and the combined analysis of these results in Theorem 8 gives the desired bounds on the diameter.

In order to prove these results, and in particular Lemma 4 and Lemma 6, it is helpful to take an alternative view of $\mathcal{SG}(Q_k,p)$. Rather than viewing $\mathcal{SG}(Q_k,p)$ as a random subset S of $\mathbb{Z}_2^k-\mathbf{0}$ with $x\sim y$ if and only if $x,y,x+y\in S$, we instead view $\mathcal{SG}(Q_k,p)$ as random subgraph of a random Cayley graph. Recall that if G is a group and $S\subset G$ is such that for all $s\in S$, $s^{-1}\in S$, then the Cayley graph of G generated by S, denoted $\mathcal{C}(G,S)$, is a graph with vertex set G and $g,h\in G$ adjacent if $g+s=h$ for some $s\in S$. As we are concerned only with the commutative group $\mathbb{Z}_2^k$, we will ignore the issues of whether S acts on G on the left or right and use addition to represent the group operation. We also note that since every element of $\mathbb{Z}_2^k$ is of order 2, every collection of elements $S \subseteq \mathbb{Z}_2^k$ is automatically closed under inversion. We now note that if S is a random subset of $\mathbb{Z}_2^k - \mathbf{0}$ where each element is chosen with probability p, then $\mathcal{SG}(Q_k, p) = \mathcal{C}\big(\mathbb{Z}_2^k, S\big)[S]$ where for a graph G, $G[S]$ denotes the subgraph induced by S. In particular $x \sim y$ in $\mathcal{C}\big(\mathbb{Z}_2^k, S\big)[S]$ if and only if $x, y \in S$ and $x \sim y$ in $\mathcal{C}\big(\mathbb{Z}_2^k, S\big)$ and thus $x + y \in S$.

This viewpoint leads to the following natural observation.

Observation 1. Fix $V \subset S \subset \mathbb{Z}_2^k - \mathbf{0}$. The neighbors of V in $\mathcal{C}\big(\mathbb{Z}_2^k, S\big)[S]$ are precisely the set of non-isolated vertices in $\mathcal{C}\big(Z_2^k, V\big)[S]$.

Proof. For a vertex $v \in V$ and $x \in S$, $v \sim x$ if and only if $v + x \in S$. But x is not isolated in $\mathcal{C}\big(\mathbb{Z}_2^k, V\big)[S]$ if $v + x \in S$. □

At this point it is worth mentioning that the structural properties of the related random graph formed by random induced subgraphs of the hypercube, or alternatively, the random graph $\mathcal{C}\big(\mathbb{Z}_2^k, \{e_1, \ldots, e_k\}\big)[S]$ where e_i is the standard basis for $\mathbb{Z}_2^k$, has been extensively studied especially from the viewpoint of random Boolean functions (see [23] for a survey of the early results in the field). Of particular interest is the result of Bollobás, Kohayakawa, and Łuczak, showing that there is a phase transition for the giant component in $\mathcal{C}\big(\mathbb{Z}_2^k, \{e_1, \ldots, e_k\}\big)[S]$. Recently, there has been significant work on the component and distance structure of a variety of generalizations including induced subgraphs of generalized k-cubes and Cayley graphs with a minimal set of generators [22, 26, 27, 28, 29]. Additionally, the connectivity, diameter, and t-connectivity of the related random graph formed by percolating the edges of $\mathcal{C}\big(\mathbb{Z}_2^k, \{e_1, \ldots, e_k\}\big)$ has also been studied [1, 9, 10, 11, 12, 15]. Further, in [24] the connectivity and the diameter (up to a small constant) of an edge percolation on an induced subgraph of $\mathcal{C}\big(\mathbb{Z}_2^k, \{e_1, \ldots, e_k\}\big)$ were determined. However, as we are dealing with a large set of random (almost surely) generators of $\mathbb{Z}_2^k$, these results give little insight in to the structure of $\mathcal{SG}(Q_k, p)$.

Since the edges of $\mathcal{SG}(Q_k, p)$ only appear as part of an induced K_3, there is a significant amount of dependency between the various edges. In order to control this dependency we will use the following version of Janson's inequality (see [3, 21]).

Theorem 3 (Janson's Inequality)**.** *Let $X_1, \ldots, X_k$ be a collection of indicator random variables for the events $\mathcal{B}_1, \ldots, \mathcal{B}_k$, and let $X = \sum_i X_i$. If $\mu = \mathbb{E}[X]$ and $\Delta = \sum_{\mathcal{B}_i \sim \mathcal{B}_j} \mathbb{P}(\mathcal{B}_i \cup \mathcal{B}_j)$, where $\mathcal{B}_i \sim \mathcal{B}_j$ if the events $\mathcal{B}_i$ and $\mathcal{B}_j$ are dependent, then*

$$\mathbb{P}(X \le (1-\lambda)\mu) \le e^{-\lambda^2 \frac{\mu^2}{2\mu + \Delta}}.$$

Additionally, since at various points of the proof we will be interested in whether large collections of elements of $\mathbb{Z}_2^k$ are all not present in S, it will be helpful to have the following easy bounds on $(1-p)^t$.

Observation 2. Let $p \in [0, 1]$, then for any positive integers t and k,

$$\sum_{i=0}^{k} (-1)^j \binom{t}{j} p^j - t^{k+1} p^{k+1} \le (1-p)^t \le \sum_{i=0}^{k} (-1)^j \binom{t}{j} p^j + t^{k+1} p^{k+1}.$$

In particular,

$$1 - tp - t^2p^2 \le (1-p)^t \le 1 - tp + t^2p^2$$

and

$$1 - tp + \binom{t}{2} p^2 - t^3p^3 \le (1-p)^t \le 1 - tp + \binom{t}{2} p^2 + t^3p^3.$$

For two functions $f, g \colon \mathbb{Z}_2^k - \mathbf{0} \to \mathbb{R}$, we define the convolution $f * g(x) = \sum_{y \in \mathbb{Z}_2^k - \mathbf{0}} f(y)g(x-y) = \sum_{y \in \mathbb{Z}_2^k - \mathbf{0}} f(x-y)g(y)$. Then if $\mathbb{1}_A$ is the indicator functions for the set $A \subseteq \mathbb{Z}_2^k - \mathbf{0}$ we have $\mathbb{1}_B * \mathbb{1}_A(\omega) = |\{(\alpha, \beta) \in A \times B \mid \alpha + \beta = \omega\}|$

2.1. **Small Set Expansion.** The primary goal of this section is to show that, similarly to Erdős-Rényi random graphs, every small set of vertices expands by a factor that is roughly proportional to the average degree in $\mathcal{SG}(Q_k, p)$. The natural approach to this problem would be to use the extensive literature studying the structural properties of sum-sets (see for instance [17, 18]) to show that with high probability for any $A \subset S$, $|(A+A) \cap S|$ is large relative to A. However, such an approach will fail to capture a significant number of the neighbors of A. Specifically, if $a_1, a_2 \in A$ and $x \notin A$, then this approach will only discover neighbors of the form a_1+a_2 and miss all the neighbors arriving in pairs $\{x, x+a_1\}$. Thus, rather than attack the problem directly, we use Observation 1 to reformulate the expansion idea and avoid the difficulties posed by a random vertex set. Specifically, rather than consider the set of neighbors of an small arbitrary subset of the vertices $V \subseteq S$ in $\mathcal{C}\left(\mathbb{Z}_2^k, S\right)[S]$, we examine an small arbitrary subset of group elements $A \subseteq \mathbb{Z}_2^k - \mathbf{0}$ and consider their neighbors in $\mathcal{C}\left(\mathbb{Z}_2^k, S \cup A\right)[S \cup A]$. The following result shows that with reasonably large probability (depending on the size of A), the number of neighbors of A in $\mathcal{C}\left(\mathbb{Z}_2^k, S \cup A\right)[S \cup A]$ is $\Theta\left(np^2 |A|\right)$.

Lemma 4. *Let S be a random subset of $\mathbb{Z}_2^k$ formed by choosing each element independently with probability p. For any fixed set A not containing $\mathbf{0}$ such that $2|A|p \leq 1$, the probability that there are less than $(1-\lambda)(n+1)\left(|A|p^2 - |A|^2 p^3\right)$ non-isolated vertices in $\mathcal{C}\left(\mathbb{Z}_2^k, A\right)[S]$ is at most $e^{-\lambda^2 \frac{(n+1)|A|p^2}{18+24|A|^3p^2}}$.*

Proof. For every $x \in \mathbb{Z}_2^k$ let $\mathcal{E}_x$ be the event that the vertex x is not isolated in $\mathcal{C}\left(\mathbb{Z}_2^k, A\right)[S]$ (and hence is present in S). Thus $\mu = \sum_{x \in Z_2^k} \mathbb{P}(\mathcal{E}_x) = (n+1)p(1-(1-p)^{|A|})$, and in particular $(n+1)p\left(|A|p - |A|^2 p^2\right) \leq \mu \leq (n+1)p\left(|A|p + |A|^2 p^2\right)$. Now consider the dependency graph for the events $\mathcal{E}_x$. If the distance between x and y in $\mathcal{C}\left(\mathbb{Z}_2^k, A\right)$ is at least 3, then $\mathcal{E}_x$ and $\mathcal{E}_y$ are independent. Thus we will evaluate $\mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_y)$ in the cases where the distance between x and y is either 1 or 2 in $\mathcal{C}\left(\mathbb{Z}_2^k, A\right)$. If $x \sim y$ in $\mathcal{C}\left(\mathbb{Z}_2^k, A\right)$, then the event $\mathcal{E}_x \cap \mathcal{E}_y$ is precisely the event that $x, y \in S$, and so $\mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_y) = p^2$.

Now consider the case where x and y are at distance 2 in $\mathcal{C}\left(\mathbb{Z}_2^k, A\right)$. Then $y = x+\omega$ where $\omega \in (A+A)-A$ and thus,

$$\begin{aligned}
\mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_{x+\omega}) &= p^2\left((1-(1-p)^{\mathbb{1}_A * \mathbb{1}_A(\omega)}) + (1-p)^{\mathbb{1}_A*\mathbb{1}_A(\omega)}(1-(1-p)^{|A|-\mathbb{1}_A*\mathbb{1}_A(\omega)})^2\right) \\
&= p^2\left(1-2(1-p)^{|A|} + (1-p)^{2|A|-\mathbb{1}_A*\mathbb{1}_A(\omega)}\right) \\
&\leq p^2\left(1-2\left(1-|A|p-|A|^2p^2\right)+1-(2|A|-\mathbb{1}_A*\mathbb{1}_A(\omega))p + (2|A|-\mathbb{1}_A*\mathbb{1}_A(\omega))^2p^2\right) \\
&\leq p^2\left(\mathbb{1}_A*\mathbb{1}_A(\omega)p^3 + 2|A|^2p^2 + (2|A|)^2p^2\right) \\
&= \mathbb{1}_A*\mathbb{1}_A(\omega)p^3 + 6|A|^2p^4,
\end{aligned}$$

where the first inequality comes from Observation 2. Using this bound, we can bound find an upper bound for Δ, namely

$$\begin{aligned}
\Delta &= \sum_{z \in \mathbb{Z}_2^k} \sum_{\substack{y \\ \mathcal{E}_x \sim \mathcal{E}_y}} \mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_y) \\
&= \sum_{x \in \mathbb{Z}_2^k} \sum_{\alpha \in (A+A)\cup A} \mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_{x+\alpha}) \\
&= \sum_{x \in \mathbb{Z}_2^k}\left(\sum_{\alpha \in A}\mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_{x+\alpha}) + \sum_{\omega \in (A+A)-A}\mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_{x+\omega})\right) \\
&\leq \sum_{x \in \mathbb{Z}_2^k}\left(|A|p^2 + \sum_{\omega \in (A+A)-A}\mathbb{1}_A*\mathbb{1}_A(\omega)p^3 + 6|A|^2p^4\right) \\
&\leq (n+1)\left(|A|p^2 + |A|^2p^3 + 6|A|^4p^4\right) \\
&= (n+1)|A|p^2\left(1+|A|p+6|A|^3p^2\right).
\end{aligned}$$

Now noting that, $2|A|p \leq 1$, we have that

$$\begin{aligned}
\frac{\mu^2}{2\mu+\Delta} &\geq \frac{(n+1)^2\left(|A|p^2-|A|^2p^3\right)^2}{2(n+1)\left(|A|p^2+|A|^2p^3\right)+(n+1)|A|p^2(1+|A|p+6|A|^3p^2)} \\
&= (n+1)|A|p^2\frac{(1-|A|p)^2}{2(1+|A|p)+(1+|A|p+6|A|^3p^2)} \\
&\geq (n+1)|A|p^2\frac{\frac{1}{4}}{3+3|A|p+6|A|^3p^2} \\
&\geq \frac{(n+1)|A|p^2}{18+24|A|^3p^2} \\
&\geq \frac{n|A|p^2}{18+24|A|^3p^2}.
\end{aligned}$$

Applying Janson's Inequality yields the desired result. ☐

Now by limiting the size of A somewhat, we can extend this result to hold for all small sets A.

Corollary 5. *Let $\delta \in (0,1)$ and let S be a random subset of $\mathbb{Z}_2^k$ formed by choosing each element independently with probability $8p \leq 1$. If $\lambda^2 np^2 \geq 42\ln\left(\frac{2n}{\delta}\right)$, then with probability at least $1-\delta$ every set A not containing* **0** *where $|A|^2p^3 \leq 1$ is such that there are at least $(1-\lambda)np^2|A|(1-p^{1/3})$ non-isolated elements in $\mathcal{C}(\mathbb{Z}_2^k, A)[S]$.*

Proof. Let t be the maximum integer such that $t^3p^2 \leq 1$, and thus $t = \left\lfloor p^{-\frac{2}{3}}\right\rfloor = p^{-\frac{2}{3}} - \zeta$ for some $\zeta \geq 0$. Now $2tp = 2\left(p^{-\frac{2}{3}}-\zeta\right)p \leq 2p^{-\frac{2}{3}}p = 2p^{\frac{1}{3}} \leq 1$. Thus by Lemma 4, for an arbitrary set A with $|A| \leq t$, the probability that there are less than $(1-\lambda)np^2|A|(1-|A|p) \geq (1-\lambda)np^2|A|(1-p^{1/3})$ non-isolated elements in $\mathcal{C}(\mathbb{Z}_2^k, A)[S]$ is at most $e^{-\lambda^2\frac{np^2|A|}{42}}$. Hence the probability that any set A not containing **0** and with $|A| \leq t$ has less than $(1-\lambda)np^2|A|(1-p^{1/3})$ non-isolated elements in $\mathcal{C}(\mathbb{Z}_2^k, A)[S]$ is at most

$$\sum_{s=1}^{t}\binom{n}{s}e^{-\lambda^2\frac{np^2s}{42}} \leq \sum_{s=1}^{t}n^se^{-\lambda^2\frac{np^2s}{42}} \leq \sum_{s=1}^{t}n^se^{-\ln\left(\frac{2n}{\delta}\right)s} = \sum_{s=1}^{t}\left(\frac{\delta}{2}\right)^s \leq \sum_{s=1}^{\infty}\left(\frac{\delta}{2}\right)^s \leq \delta. \quad \square$$

It is worth noting at this point, that the condition that $|A|^3p^2 \leq 1$ is not tight, and in fact can be relaxed to $|A| \in \mathcal{O}\left(\sqrt[4]{\frac{n}{\ln\left(\frac{n}{\delta}\right)}}\right)$ for constant λ.

2.2. **Connections Between Large Sets.** We note at this point that the restriction from Lemma 4 that $2|A|p \leq 1$ is best possibly up to constants, as if $|A|p \to \infty$, then the number of neighbors would be $|A|np^2 = (np)|A|p \in \Omega(np)$, more than the total number of vertices. Thus, again similarly to Erdős-Rényi random graphs, we must find a means of bridging the gap between two large subsets of vertices. We will once again use the viewpoint of Observation 1 and note that two disjoint sets of vertices A and B have a common neighbor if and only if there is a vertex which is non-isolated in both $\mathcal{C}(Z_2^k, A)[S]$ and $\mathcal{C}(\mathbb{Z}_2^k, B)[S]$. Thus we desire a control on the number of such vertices, which is provided by the following lemma.

Lemma 6. *Let S be a random subset of $\mathbb{Z}_2^k$ formed by choosing each element independently with probability p. For any set two fixed disjoint sets A, B neither containing* **0** *each such that $2|A|p, 2|B|p \leq 1$, the probability that there are at least $(1-\lambda)np^3|A||B|(1-|A|p)(1-|B|p)$ vertices that are not isolated in $\mathcal{C}(\mathbb{Z}_2^k, A)[S]$ and $\mathcal{C}(\mathbb{Z}_2^k, B)[S]$ is at least $1-e^{-\lambda^2\frac{n|A|^2|B|^2p^3}{152|A||B|+640(|A|+|B|)^5p^2}}$*

Proof. Similarly to the proof of Lemma 4, we will proceed by using Janson's inequality. For every $x \in \mathbb{Z}_2^k$, let $\mathcal{E}_x$ be the event that x is not isolated in both $\mathcal{C}(\mathbb{Z}_2^k, A)[S]$ and $\mathcal{C}(\mathbb{Z}_2^k, B)[S]$. Then $\mu = (n+1)p(1-(1-$

$p)^{|A|})(1-(1-p)^{|B|})$. Before calculating Δ we make a few observations about the value of $\mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_{x+\omega})$. If $\omega \in A$ then

$$\begin{aligned}\mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_{x+\omega}) &= p^2\left((1-(1-p)^{\mathbb{1}_B*\mathbb{1}_B(\omega)}) + (1-p)^{\mathbb{1}_B*\mathbb{1}_B(\omega)}(1-(1-p)^{|B|-\mathbb{1}_B*\mathbb{1}_B(\omega)})^2\right)\\ &= p^2\left(1-2(1-p)^{|B|}+(1-p)^{2|B|-\mathbb{1}_B*\mathbb{1}_B(\omega)}\right)\\ &\leq p^2\left(1-2\left(1-|B|\,p-|B|^2\,p^2\right)+\left(1-(2\,|B|-\mathbb{1}_B*\mathbb{1}_B(\omega))\,p+(2\,|B|-\mathbb{1}_B*\mathbb{1}_B(\omega))^2\,p^2\right)\right)\\ &= \mathbb{1}_B*\mathbb{1}_B(\omega)\,p^3+6\,|B|^2\,p^4.\end{aligned}$$

Similarly, if $\omega \in B$, then

$$\mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_{x+\omega}) = p^2\left(1-2(1-p)^{|A|}+(1-p)^{2|A|-\mathbb{1}_A*\mathbb{1}_A(\omega)}\right) \leq \mathbb{1}_A*\mathbb{1}_A(\omega)\,p^3+6\,|A|^2\,p^4.$$

In order to understand the behavior of $\mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_{x+\omega})$ when $\omega \in ((A\cup B)+(A\cup B))-(A\cup B)$, let $A_\omega \subset A$ and $B_\omega \subset B$ be the minimal (potentially empty) subsets so that $\mathbb{1}_B*\mathbb{1}_A(\omega) = \mathbb{1}_{B_\omega}*\mathbb{1}_{A_\omega}(\omega)$.

Now let $\mathcal{E}_x^{A_\omega}$ be the event that x is not isolated in $\mathcal{C}(\mathbb{Z}_2^k, A_\omega)\,[S]$ and let $\mathcal{E}_x^{B_\omega}$ be the event that x is not isolated in $\mathcal{C}(\mathbb{Z}_2^k, B_\omega)\,[S]$. We notice that $\mathcal{E}_x^{A_\omega}\cap\mathcal{E}_x^{B_\omega}, \overline{\mathcal{E}_x^{A_\omega}}\cap\mathcal{E}_x^{B_\omega}, \mathcal{E}_x^{A_\omega}\cap\overline{\mathcal{E}_x^{B_\omega}}$, and $\overline{\mathcal{E}_x^{A_\omega}}\cap\overline{\mathcal{E}_x^{B_\omega}}$ form a partition of the space, and thus

$$\begin{aligned}\mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_{x+\omega}) = {} & \mathbb{P}\big(\mathcal{E}_x \cap \mathcal{E}_{x+\omega} \mid \mathcal{E}_x^{A_\omega}\cap\mathcal{E}_x^{B_\omega}\big)\,\mathbb{P}\big(\mathcal{E}_x^{A_\omega}\cap\mathcal{E}_x^{B_\omega}\big)\\ &+\mathbb{P}\Big(\mathcal{E}_x \cap \mathcal{E}_{x+\omega} \mid \overline{\mathcal{E}_x^{A_\omega}}\cap\mathcal{E}_x^{B_\omega}\Big)\,\mathbb{P}\Big(\overline{\mathcal{E}_x^{A_\omega}}\cap\mathcal{E}_x^{B_\omega}\Big)\\ &+\mathbb{P}\Big(\mathcal{E}_x \cap \mathcal{E}_{x+\omega} \mid \mathcal{E}_x^{A_\omega}\cap\overline{\mathcal{E}_x^{B_\omega}}\Big)\,\mathbb{P}\Big(\mathcal{E}_x^{A_\omega}\cap\overline{\mathcal{E}_x^{B_\omega}}\Big)\\ &+\mathbb{P}\Big(\mathcal{E}_x \cap \mathcal{E}_{x+\omega} \mid \overline{\mathcal{E}_x^{A_\omega}}\cap\overline{\mathcal{E}_x^{B_\omega}}\Big)\,\mathbb{P}\Big(\overline{\mathcal{E}_x^{A_\omega}}\cap\overline{\mathcal{E}_x^{B_\omega}}\Big).\end{aligned}$$

Now observing that $|A_\omega| = |B_\omega| = \mathbb{1}_B*\mathbb{1}_A(\omega)$ for compactness of notation we let $\tau_\omega = \mathbb{1}_B*\mathbb{1}_A(\omega)$ and have that

$$\begin{aligned}\mathbb{P}\big(\mathcal{E}_x^{A_\omega}\cap\mathcal{E}_x^{B_\omega}\big) &= p\,(1-(1-p)^{\tau_\omega})^2\\ \mathbb{P}\Big(\overline{\mathcal{E}_x^{A_\omega}}\cap\mathcal{E}_x^{B_\omega}\Big) = \mathbb{P}\Big(\mathcal{E}_x^{A_\omega}\cap\overline{\mathcal{E}_x^{B_\omega}}\Big) &= p\,(1-(1-p)^{\tau_\omega})\,(1-p)^{\tau_\omega}\\ \mathbb{P}\Big(\overline{\mathcal{E}_x^{A_\omega}}\cap\overline{\mathcal{E}_x^{B_\omega}}\Big) &= p(1-p)^{2\tau_\omega}\end{aligned}$$

We also note that $\mathbb{P}\big(\mathcal{E}_x \cap \mathcal{E}_{x+\omega} \mid \mathcal{E}_x^{A_\omega}\cap\mathcal{E}_x^{B_\omega}\big) = \mathbb{P}\big(\mathcal{E}_{x+\omega} \mid \mathcal{E}_x^{A_\omega}\cap\mathcal{E}_x^{B_\omega}\big) = p$ and

$$\left.\begin{aligned}\mathbb{P}\Big(\mathcal{E}_x \cap \mathcal{E}_{x+\omega} \mid \mathcal{E}_x^{A_\omega}\cap\overline{\mathcal{E}_x^{B_\omega}}\Big)\\ \mathbb{P}\Big(\mathcal{E}_x \cap \mathcal{E}_{x+\omega} \mid \overline{\mathcal{E}_x^{A_\omega}}\cap\mathcal{E}_x^{B_\omega}\Big)\end{aligned}\right\} = p\left(1-(1-p)^{|A|-\tau_\omega}\right)\left(1-(1-p)^{|B|-\tau_\omega}\right).$$

Noting that $\mathbb{1}_A*\mathbb{1}_A(\omega) = \mathbb{1}_{A-A_\omega}*\mathbb{1}_{A-A_\omega}(\omega)$ and $\mathbb{1}_B*\mathbb{1}_B(\omega) = \mathbb{1}_{B-B_\omega}*\mathbb{1}_{B-B_\omega}(\omega)$ by the disjointness of A and B combined with the definition of A_ω and B_ω, we have

$$\begin{aligned}\mathbb{P}\Big(\mathcal{E}_x \cap \mathcal{E}_{x+\omega} \mid \overline{\mathcal{E}_x^{A_\omega}}\cap\overline{\mathcal{E}_x^{B_\omega}}\Big) &= p\left(\left(1-(1-p)^{\mathbb{1}_A*\mathbb{1}_A(\omega)}\right)+(1-p)^{\mathbb{1}_A*\mathbb{1}_A(\omega)}\left(1-(1-p)^{|A|-\tau_\omega-\mathbb{1}_A*\mathbb{1}_A(\omega)}\right)^2\right)\\ &\quad\times\left(\left(1-(1-p)^{\mathbb{1}_B*\mathbb{1}_B(\omega)}\right)+(1-p)^{\mathbb{1}_B*\mathbb{1}_B(\omega)}\left(1-(1-p)^{|B|-\tau_\omega-\mathbb{1}_B*\mathbb{1}_B(\omega)}\right)^2\right)\\ &= p\left(1-2(1-p)^{|A|-\tau_\omega}+(1-p)^{2|A|-2\tau_\omega-\mathbb{1}_A*\mathbb{1}_A(\omega)}\right)\\ &\quad\times\left(1-2(1-p)^{|B|-\tau_\omega}+(1-p)^{2|B|-2\tau_\omega-\mathbb{1}_B*\mathbb{1}_B(\omega)}\right)\end{aligned}$$

Combining and expanding these results, we have that

$$\begin{aligned}\mathbb{P}(\mathcal{E}_x \cup \mathcal{E}_{x+\omega}) = p^2 \Big(&1 + (1-p)^{2|A| - \mathbb{1}_A * \mathbb{1}_A(\omega)} + (1-p)^{2|B| - \mathbb{1}_B*\mathbb{1}_B(\omega)} + (1-p)^{2|A|+2|B| - \mathbb{1}_A*\mathbb{1}_A(\omega) - \mathbb{1}_B*\mathbb{1}_B(\omega) - 2\mathbb{1}_B*\mathbb{1}_A(\omega)}\Big)\\ &- 2p^2\left((1-p)^{2|A|+|B| - \mathbb{1}_B*\mathbb{1}_A(\omega) - \mathbb{1}_A*\mathbb{1}_A(\omega)} + (1-p)^{|A|+2|B| - \mathbb{1}_B*\mathbb{1}_A(\omega) - \mathbb{1}_B*\mathbb{1}_B(\omega)}\right)\\ &+ 2p^2\left((1-p)^{|A|+|B|} - (1-p)^{|A|} - (1-p)^{|B|} + (1-p)^{|A|+|B| - \mathbb{1}_B*\mathbb{1}_A(\omega)}\right).\end{aligned}$$

Now, using the second order approximation from Observation 2, we have that

$$\mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_{x+\omega}) \le p^2\left(c_0 + c_1 p + c_2 p^2 + c_3 p^3\right).$$

It is easy to see that in this case $c_0 = c_1 = 0$. Now observing that $\binom{t}{2} = \frac{1}{2}t^2 - \frac{1}{2}t$, we note that c_2 is half the weighted sum of the squares of the exponents minus half c_1, which is 0. That is,

$$\begin{aligned}2c_2 &= \left(2|A| - \mathbb{1}_A*\mathbb{1}_A(\omega)\right)^2 + \left(2|B| - \mathbb{1}_B*\mathbb{1}_B(\omega)\right)^2 + \left(2|A| - \mathbb{1}_A*\mathbb{1}_A(\omega) + 2|B| - \mathbb{1}_B*\mathbb{1}_B(\omega) - 2\tau_\omega\right)^2\\ &\quad - 2\left(2|A| + |B| - \tau_\omega - \mathbb{1}_A*\mathbb{1}_A(\omega)\right)^2 - 2\left(|A| + 2|B| - \tau_\omega - \mathbb{1}_B*\mathbb{1}_B(\omega)\right)^2\\ &\quad + 2\left(|A|+|B|\right)^2 - 2|A|^2 - 2|B|^2 + 2\left(|A|+|B|-\tau_\omega\right)^2\\ &= \tau_\omega^2 + \mathbb{1}_A*\mathbb{1}_A(\omega)\,\mathbb{1}_B*\mathbb{1}_B(\omega).\end{aligned}$$

Similarly we have that

$$\begin{aligned}c_3 &= \left(2|A| - \mathbb{1}_A*\mathbb{1}_A(\omega)\right)^3 + \left(2|B| - \mathbb{1}_B*\mathbb{1}_B(\omega)\right)^3 + \left(2|A| - \mathbb{1}_A*\mathbb{1}_A(\omega) + 2|B| - \mathbb{1}_B*\mathbb{1}_B(\omega) - 2\tau_\omega\right)^3\\ &\quad + 2\left(2|A| + |B| - \tau_\omega - \mathbb{1}_A*\mathbb{1}_A(\omega)\right)^3 + 2\left(|A| + 2|B| - \tau_\omega - \mathbb{1}_B*\mathbb{1}_B(\omega)\right)^3\\ &\quad + 2\left(|A|+|B|\right)^3 + 2|A|^3 + 2|B|^3 + 2\left(|A|+|B|-\tau_\omega\right)^3\\ &\le \left(2|A|\right)^3 + \left(2|B|\right)^3 + \left(2|A|+2|B|\right)^3 + 2\left(2|A|+|B|\right)^3 + 2\left(|A|+2|B|\right)^3\\ &\quad + 2\left(|A|+|B|\right)^3 + 2|A|^3 + 2|B|^3 + 2\left(|A|+|B|\right)^3\\ &\le 40|A|^3 + 72|A|^2|B| + 72|A||B|^2 + 40|B|^3\\ &\le 40\left(|A|+|B|\right)^3.\end{aligned}$$

Thus we have that

$$\mathbb{P}(\mathcal{E}_x \cap \mathcal{E}_{x+\omega}) \le \mathbb{1}_B*\mathbb{1}_A(\omega)^2 + \mathbb{1}_A*\mathbb{1}_A(\omega)\,\mathbb{1}_B*\mathbb{1}_B(\omega)\,p^4 + 40\left(|A|+|B|\right)^3 p^5.$$

We notice that $\sum_\omega \mathbb{1}_B*\mathbb{1}_A(\omega)^2$ counts the number of solutions to $\alpha_1 + \beta_1 + \beta_2 + \alpha_2 = 0$, with $\alpha_1, \alpha_2 \in A$ and $\beta_1, \beta_2 \in B$. Similarly, $\sum_\omega \mathbb{1}_A*\mathbb{1}_A(\omega)\,\mathbb{1}_B*\mathbb{1}_B(\omega)$, counts the solutions to $\alpha_1 + \alpha_2 + \beta_2 + \beta_1 = 0$ with $\alpha_1, \alpha_2 \in A$ and $\beta_1, \beta_2 \in B$. Thus

$$\sum_{\omega \in ((A\cup B)+(A\cup B))-(A\cup B)} \mathbb{1}_B*\mathbb{1}_A(\omega)^2 + \mathbb{1}_A*\mathbb{1}_A(\omega)\,\mathbb{1}_B*\mathbb{1}_B(\omega) \le |A||B|\left(|A|+|B|\right).$$

Combining these results we have that

$$\Delta \le (n+1)\left(2|A||B|p^3 + 7|A||B|\left(|A|+|B|\right)p^4 + 40\left(|A|+|B|\right)^5 p^5\right).$$

Now observing that $\mu = (n+1)p\left(1-(1-p)^{|A|}\right)\left(1-(1-p)^{|B|}\right)$, we have that

$$\begin{aligned}
\frac{\Delta}{\mu} &\leq \frac{(n+1)\left(2\,|A|\,|B|\,p^3 + 7\,|A|\,|B|\,(|A|+|B|)\,p^4 + 40\,(|A|+|B|)^5\,p^5\right)}{(n+1)p\left(|A|\,p - |A|^2\,p^2\right)\left(|B|\,p - |B|^2\,p^2\right)} \\
&= \frac{2\,|A|\,|B| + 7\,|A|\,|B|\,(|A|+|B|)\,p + 40\,(|A|+|B|)^5\,p^2}{|A|\,|B|\,(1-|A|\,p)\,(1-|B|\,p)} \\
&\leq \frac{2\,|A|\,|B| + 7\,|A|\,|B|\,(|A|+|B|)\,p + 40\,(|A|+|B|)^5\,p^2}{|A|\,|B|\,1/4} \\
&\leq 36 + 160\frac{(|A|+|B|)^5}{|A|\,|B|}p^2
\end{aligned}$$

Thus we have that

$$\begin{aligned}
\frac{\mu^2}{2\mu+\Delta} &= \frac{\mu}{2+\frac{\Delta}{\mu}} \\
&\geq \frac{\mu}{38+160\frac{(|A|+|B|)^5}{|A||B|}p^2} \\
&\geq \frac{(n+1)p\left(|A|\,p - |A|^2\,p^2\right)\left(|B|\,p - |B|^2\,p^2\right)}{38+160\frac{(|A|+|B|)^5}{|A||B|}p^2} \\
&= \frac{(n+1)\,|A|\,|B|\,p^3\,(1-|A|\,p)\,(1-|B|\,p)}{38+160\frac{(|A|+|B|)^5}{|A||B|}p^2} \\
&\geq \frac{n\,|A|^2\,|B|^2\,p^3}{152\,|A|\,|B| + 640\,(|A|+|B|)^5\,p^2}
\end{aligned}$$

Applying Janson's Inequality completes the proof. □

In order to apply Lemma 6 to bound the diameter, we extend it to collections of a sets of the same size.

Corollary 7. *Let S be a random subset of $\mathbb{Z}_2^k$ formed by choosing each element independently with probability p, let $A_1, \ldots, A_t$ be subsets of $\mathbb{Z}_2^k - \mathbf{0}$ with $|A_i| = s$ for all i, and let $\delta \in (0,1)$. If $50\sqrt{\frac{\ln\left(\frac{t^2}{\delta}\right)}{np^3}} \leq s \leq \min\left\{\frac{1}{2p}, \frac{np}{2500\ln\left(\frac{t^2}{\delta}\right)}\right\}$, then with probability at least $1-\delta$ every pair A_i and A_j is such that either $A_i \cap A_j \neq \varnothing$ or there are vertices that are not isolated in both $\mathcal{C}\left(\mathbb{Z}_2^k, A_i\right)[S]$ and $\mathcal{C}\left(\mathbb{Z}_2^k, A_j\right)[S]$.*

Proof. Consider an arbitrary pair A_i and A_j. If $A_i \cap A_j = \varnothing$, then since $2\,|A_i|\,p, 2\,|A_j|\,p \leq 1$ by Lemma 6 for any $\lambda \in (0,1)$ the probability that there are fewer than

$$(1-\lambda)np^3s^2\,(1-sp)^2 \geq \frac{1-\lambda}{4}np^3s^2 \geq \frac{2500(1-\lambda)}{4}\ln\left(\frac{t^2}{\delta}\right) > 0$$

vertices that are not isolated in both $\mathcal{C}\left(\mathbb{Z}_2^k, A_i\right)[S]$ and $\mathcal{C}\left(\mathbb{Z}_2^k, A_j\right)[S]$ is at most $e^{-\lambda^2\frac{ns^2p^3}{152+2048s^3p^2}}$. In particular, the probability that there no vertex with is not isolated in both $\mathcal{C}\left(\mathbb{Z}_2^k, A_i\right)[S]$ and $\mathcal{C}\left(\mathbb{Z}_2^k, A_j\right)[S]$ is at most $e^{-\lambda^2\frac{ns^2p^3}{152+2048s^3p^2}}$. We note that if $s^3p^2 > 1$, then

$$\frac{ns^2p^3}{152+2048s^3p^2} \geq \frac{ns^2p^3}{2200s^3p^2} = \frac{np}{2200s} \geq \frac{25}{22}\ln\left(\frac{t^2}{\delta}\right).$$

Similarly, if $s^3p^2 \leq 1$, we have

$$\frac{ns^2p^3}{1522+2048s^3p^2} \geq \frac{ns^2p^3}{2200} \geq \frac{25}{22}\ln\left(\frac{t^2}{\delta}\right).$$

Thus choosing $\lambda = \sqrt{\frac{22}{25}}$ and observing that there are at most t^2 pairs to consider gives the result. $\square$

2.3. **Diameter.** In this section we combine Corollaries 5 and 7, to provide an upper bound on the diameter of $\mathcal{SG}(Q_k, p)$. Additionally, we will combine the upper bound with a natural lower bound to show that with probability approach one, the diameter is in general concentrated on a constant width range of values.

Theorem 8. *Let $\delta, \lambda, p \in (0,1)$ with $8\lambda \leq 1$. If $\frac{42}{\lambda^4}\ln\left(\frac{4n^2}{\delta}\right) \leq np^2$ and $np^3 \leq \ln\left(\frac{4n^2}{\delta}\right)$, then with probability at least $1-\delta$,*

$$\left\lceil \frac{\ln\left((1-\sqrt{p}\lambda)np\right)}{\ln\left((1+\lambda)np^2\right)} \right\rceil \leq \operatorname{diam}\left(\mathcal{SG}(Q_k,p)\right) \leq 2\left\lceil \frac{\ln\left(64\sqrt{\frac{\ln\left(\frac{4n^2}{\delta}\right)}{np^3}}\right)}{\ln\left((1-\lambda)np^2(1-p^{\frac{1}{3}})\right)} \right\rceil + 2 \leq 2\left\lceil \frac{\ln\left(64(1-\lambda)\sqrt{\ln\left(\frac{4n^2}{\delta}\right)np}\right)}{\ln\left((1-\lambda)np^2(1-p^{\frac{1}{3}})\right)} \right\rceil.$$

Proof. We first consider the lower bound. By Chernoff bounds, the probability that there are less than $(1-\sqrt{p}\lambda)np$ vertices is at most $e^{-p\lambda^2\frac{np}{3}} = e^{-\lambda^2\frac{np^2}{3}}$. Similarly, there is a some vertex with degree greater than $(1+\lambda)np^2$ with probability at most $npe^{-\lambda^2\frac{np^2}{3}}$. Thus with probability at least $1-2ne^{-\lambda^2\frac{np^2}{3}} \geq 1-2ne^{\frac{14}{\lambda^2}\ln\left(\frac{4n^2}{\delta}\right)} \geq 1-\frac{\delta}{2}$, we have $\left((1+\lambda)np^2\right)^{\operatorname{diam}(\mathcal{SG}(Q_k,p))} \geq (1-\sqrt{p}\lambda)np$, and thus $\operatorname{diam}\left(\mathcal{SG}(Q_k,p)\right) \geq \left\lceil \frac{\ln\left((1-\sqrt{p}\lambda)np\right)}{\ln((1+\lambda)np^2)} \right\rceil$.

For the upper bound we will consider a variant of $\mathcal{SG}(Q_k,p)$ where the vertices are additionally assigned a color, either red or blue. Specifically, we independently color each element $\mathbb{Z}_2^k - \mathbf{0}$ red with probability p_R and blue with probability p_B. Let $\mathcal{R}$ be the collection of red elements and let $\mathcal{B}$ be the collection of blue elements. Since the colors are independent, with probability p_Rp_B an element of $\mathbb{Z}_2^k - \mathbf{0}$ belongs to both $\mathcal{R}$ and $\mathcal{B}$. We note that if $(1-p_R)(1-p_B) = (1-p)$ then the probability of a element not being colored is the same as an element not being a vertex of $\mathcal{SG}(Q_k,p)$. Hence, the collection of colored elements of $\mathbb{Z}_2^k - \mathbf{0}$ are the vertices of $\mathcal{SG}(Q_k,p)$. We will proceed by building a neighborhood of red vertices of size s around each element of $\mathbb{Z}_2^k - \mathbf{0}$, where s is chosen so that there is a blue vertex between each of these sets.

Now if $p_R \leq \frac{1}{8}$ and $\lambda^2 np_R^2 \geq 42\ln\left(\frac{8n}{\delta}\right)$, then Corollary 5 applies. Thus, with probability at least $1-\frac{\delta}{4}$ every set A such that $|A|^2 p_R^3 \leq 1$ is such that there are at least $(1-\lambda)np_R^2|A|(1-p_R^{\frac{1}{3}})$ non-isolated elements in $\mathcal{C}\left(\mathbb{Z}_2^k, A\right)[\mathcal{R}]$. Now for each $v \in \mathbb{Z}_2^k - \mathbf{0}$, we define $N_v^{(0)} = \{v\}$ and a nested sequence $N_v^{(0)} \subset N_v^{(1)} \subset \cdots$ where $N_v^{(i+1)}$ is the collection of non-isolated vertices in $\mathcal{C}\left(Z_2^k, N_n^{(i)}\right)[\mathcal{R}]$. Now by observation 1, if v is colored then $N_v^{(k)}$ is a subset of the vertices that are distance at most k from v. Notice that if $s \leq (1-\lambda)np_R^2(1-p_R^{\frac{1}{3}})p_R^{-\frac{2}{3}}$, then there is some index j such that $\left|N_v^{(j)}\right| \geq s$. Let j_v be the first such index and let N_v be a set of size s such that $N_v^{(j_v-1)} \subset N_v \subseteq N_v^{(j_v)}$. Furthermore, we have that

$$j_v \leq \left\lceil \frac{\ln(s)}{\ln\left((1-\lambda)np_R^2(1-p_R^{\frac{1}{3}})\right)} \right\rceil.$$

Consider then the collection $\{N_v\}_{v\in\mathbb{Z}_2^k-\mathbf{0}}$. By Corollary 7, if $50\sqrt{\frac{\ln\left(\frac{4n^2}{\delta}\right)}{np_B^3}} \leq s \leq \min\left\{\frac{1}{2p_B}, \frac{np_B}{2500\ln\left(\frac{4n^2}{\delta}\right)}\right\}$, then with probability at least $1-\frac{\delta}{4}$ every pair N_v and N_u either intersects or there is at least one vertex that is non-isolated in $\mathcal{C}\left(\mathbb{Z}_2^k, N_v\right)[\mathcal{B}]$ and $\mathcal{C}\left(\mathbb{Z}_2^k, N_u\right)[\mathcal{B}]$. Thus, if v and u are colored then the distance between them is at most $2\max_v j_v + 2$ with probability at least $1-\frac{\delta}{2}$.

Thus it suffices to show that there is some choice of an integer s, $p_R, p_B \in (0,1)$ such that

$$\begin{aligned} p_R + p_B - p_R p_B = &\quad p \\ \sqrt{\frac{42\ln\left(\frac{8n}{\delta}\right)}{\lambda^2 n}} \leq &\quad p_R \quad \leq \tfrac{1}{8} \\ 50\sqrt{\frac{\ln\left(\frac{4n^2}{\delta}\right)}{np_B^3}} \leq &\quad s \quad \leq \min\left\{\frac{1}{2p_B}, \frac{np_B}{2500\ln\left(\frac{4n^2}{\delta}\right)}, (1-\lambda)np_R^{\frac{4}{3}}\left(1-p_R^{\frac{1}{3}}\right)\right\}. \end{aligned}$$

To that end let $p_R = \lambda p$ and $s = \left\lfloor 64\sqrt{\frac{\ln\left(\frac{4n^2}{\delta}\right)}{np^3}} \right\rfloor$. Note that our choice of p_B fixes $(1-\lambda)p < p_B < p$.

In order to clarify the verification of the remaining constraints, we define $\tau = \ln\left(\frac{4n^2}{\delta}\right)$. Noting that $\ln\left(\frac{8n}{\delta}\right) - \ln\left(\frac{4n^2}{\delta}\right) = \ln\left(\frac{2}{n}\right) < 0$, we have

$$\sqrt{\frac{42\ln\left(\frac{8n}{\delta}\right)}{\lambda^2 n}} \leq \sqrt{\frac{42\tau}{\lambda^2 n}} = \sqrt{\frac{42\tau}{\lambda^2 np^2}}p \leq \sqrt{\frac{42\tau\lambda^2}{42\tau}}p = \lambda p = p_R.$$

Hence, the first two requirements are satisfied by this choice of s and p_R.

Since

$$64\sqrt{\frac{\tau}{np^3}} - 50\sqrt{\frac{\tau}{np_B^3}} = \left(64\,(1-\lambda)^{\frac{3}{2}} - 50\right)\sqrt{\frac{\tau}{np_B^3}} \geq \left(64\,(1-\lambda)^{\frac{3}{2}} - 50\right) \geq \left(64\left(\frac{7}{8}\right)^{\frac{3}{2}} - 50\right) \geq 1,$$

we have that $50\sqrt{\frac{\tau}{np_B^3}} \leq s$. Additionally, since $\frac{np^3}{\tau} \leq 1$ and $\frac{42\tau}{\lambda^4 np^2} \leq 1$ we have

$$\begin{aligned} &\min\left\{\frac{1}{2p_B}, \frac{np_B}{2500\tau}, (1-\lambda)np_R^{\frac{4}{3}}(1-p_R^{\frac{1}{3}})\right\} \\ &\geq \min\left\{\frac{1}{2p}, \frac{(1-\lambda)np}{2500\tau}, (1-\lambda)^{\frac{7}{3}}np^{\frac{4}{3}}(1-p^{\frac{1}{3}})\right\} \\ &\geq \min\left\{\frac{1}{2p}\sqrt{\frac{np^3}{\tau}}\left(\frac{42\tau}{\lambda^4 np^2}\right), \frac{(1-\lambda)np}{2500\tau}\sqrt{\frac{np^3}{\tau}}\left(\frac{42\tau}{\lambda^4 np^2}\right)^2, (1-\lambda)^{\frac{7}{3}}np^{\frac{4}{3}}(1-p^{\frac{1}{3}})\right\} \\ &= \min\left\{\frac{21}{\lambda^4}\sqrt{\frac{\tau}{np^3}}, \frac{1764}{2500\lambda^8}\sqrt{\frac{\tau}{np^3}}, (1-\lambda)^{\frac{7}{3}}np^{\frac{4}{3}}(1-p^{\frac{1}{3}})\right\} \\ &\geq \min\left\{64\sqrt{\frac{\tau}{np^3}}, (1-\lambda)^{\frac{7}{3}}np^{\frac{4}{3}}(1-p^{\frac{1}{3}})\right\} \\ &\geq \min\left\{64\sqrt{\frac{\tau}{np^3}}, (1-\lambda)^{\frac{7}{3}}np^{\frac{4}{3}}(1-p^{\frac{1}{3}})\left(\frac{np^3}{\tau}\right)^{1/6}\left(\frac{42\tau}{\lambda^4 np^2}\right)^{5/3}\right\} \\ &= \min\left\{64\sqrt{\frac{\tau}{np^3}}, \frac{42^{\frac{5}{3}}(1-\lambda)^{\frac{7}{3}}}{\lambda^{\frac{20}{3}}}(1-p^{\frac{1}{3}})\tau\sqrt{\frac{\tau}{np^3}}\right\} \\ &\geq 64\sqrt{\frac{\tau}{np^3}}. \end{aligned}$$

Thus our choice of s and p_B, yields that the diameter of $\mathcal{SG}(Q_k, p)$ is at most

$$2\left\lceil \frac{\ln\left(64\sqrt{\frac{\ln\left(\frac{4n^2}{\delta}\right)}{np^3}}\right)}{\ln\left((1-\lambda)np^2(1-p^{\frac{1}{3}})\right)} \right\rceil + 2$$

with probability at least $1 - \frac{\delta}{2}$, completing the proof. □

As a consequence to this result, and similarly to the case with the Erdős-Rényi random graph [13, 7, 8], we have that the diameter concentrates on a small number of values.

Corollary 9. *If* $np^2 \geq 348169\ln(n)$*, then* $\operatorname{diam}(\mathcal{SG}(Q_k,p)) = (1+o(1))\frac{\ln(np)}{\ln(np^2)}$ *asymptotically almost surely. Let*

$$L = \limsup_{n\to\infty} \frac{\ln(np)}{\ln^2(np^2)}\sqrt[4]{\frac{85\ln(n)}{np^2}}.$$

If L *is finite, then asymptotically almost surely we have*

$$\left\lceil \frac{\ln(np)}{\ln(np^2)} \right\rceil - \lceil L \rceil - 1 \leq \operatorname{diam}(\mathcal{SG}(Q_k,p)) \leq \left\lceil \frac{\ln(np)}{\ln(np^2)} \right\rceil + \lceil L \rceil + 3$$

and the diameter is concentrated on one of $2\lceil L \rceil + 4$ *values. If in addition* $\ln(np^2) \gg \ln\ln(n)$*, then asymptotically almost surely*

$$\left\lceil \frac{\ln(np)}{\ln(np^2)} \right\rceil - 1 \leq \operatorname{diam}(\mathcal{SG}(Q_k,p)) \leq \left\lceil \frac{\ln(np)}{\ln(np^2)} \right\rceil + 2$$

and the diameter is concentrated on 3 values.

Proof. Let $\delta \in (0,1)$ go to zero arbitrarily slowing with n and thus $\ln\left(\frac{4n^2}{\delta}\right) = (2+o(1))\ln(n)$. We also note that if $np^3 \geq 2\ln(n)$, then the diameter is either 2 or 1 by Theorem 1, thus we may assume that $np^3 \leq 2\ln(n)$. Hence, for sufficiently large n Theorem 8 applies for any choice of $\lambda \in \left(\sqrt[4]{\frac{85\ln(n)}{np^2}}, \frac{1}{8}\right)$. Before choosing λ we consider the difference of the upper and lower bounds from $\frac{\ln(np)}{\ln(np^2)}$.

First we have that

$$\begin{aligned}
\frac{\ln(np)}{\ln(np^2)} - \frac{\ln\left((1-\sqrt{p}\lambda)np\right)}{\ln\left((1+\lambda)np^2\right)} &= \frac{\ln(np)\ln(np^2) + \ln(1+\lambda)\ln(np) - \ln(np)\ln(np^2) - \ln(1-\sqrt{p}\lambda)\ln(np^2)}{\ln^2(np^2) + \ln(1+\lambda)\ln(np^2)} \\
&\leq \frac{\lambda\ln(np) + \sqrt{p}\ln(np^2)}{\ln^2(np^2)} \\
&= \frac{\lambda\ln(np)}{\ln^2(np^2)} + o\left(\frac{1}{\ln(np^2)}\right)
\end{aligned}$$

Now considering the form of the upper bound and recalling that $8\lambda \leq 1$ we have

$$\begin{aligned}
\frac{\ln\left(64^2(1-\lambda)^2\ln\left(\frac{4n^2}{\delta}\right)np\right)}{\ln\left((1-\lambda)(1-p^{\frac{1}{3}})np^2\right)} - \frac{\ln(np)}{\ln(np^2)} &= \frac{\ln\left(64^2(1-\lambda)^2\ln\left(\frac{4n^2}{\delta}\right)\right)\ln(np^2) - \ln\left((1-\lambda)(1-p^{\frac{1}{3}})\right)\ln(np)}{\ln^2(np^2) + \ln\left((1-\lambda)(1-p^{\frac{1}{3}})\right)\ln(np^2)} \\
&\leq \frac{\ln\ln(n)\ln(np^2) + \left(\lambda + p^{\frac{1}{3}}\right)\ln(np)}{(1-o(1))\ln^2(np^2)} + \mathcal{O}\left(\frac{1}{\ln(np^2)}\right) \\
&\leq \frac{\ln\ln(n)\ln(np^2) + \lambda\ln(np)}{(1-o(1))\ln^2(np^2)} + \mathcal{O}\left(\frac{1}{\ln(np^2)}\right) \\
&= (1+o(1))\left(\frac{\lambda\ln(np)}{\ln^2(np^2)} + \frac{\ln\ln(n)}{\ln(np^2)}\right) + \mathcal{O}\left(\frac{1}{\ln(np^2)}\right)
\end{aligned}$$

In the case that $np^2 \geq 348160\ln(n)$, we may let $\lambda = \frac{1}{8}$ and thus the upper and lower bounds from Theorem 8 give that $\operatorname{diam}(\mathcal{SG}(Q_k,p)) = (1+o(1))\frac{\ln(np)}{\ln(np^2)}$.

Now in the case L is finite and positive, we choose $\lambda = \sqrt[4]{\frac{85\ln(n)}{np^2}}$ and thus the conditions of Theorem 8 hold. Noting that $\lim_{n\to\infty}\frac{\lambda\ln(np)}{\ln^2(np^2)} \leq \lceil L \rceil$ we have that

$$\operatorname{diam}(\mathcal{SG}(Q_k,p)) \geq \left\lceil \frac{\ln(np)}{\ln(np^2)} - \lceil L \rceil - o(1) \right\rceil \geq \left\lceil \frac{\ln(np)}{\ln(np^2)} \right\rceil - \lceil L \rceil - 1.$$

In order to analyze the upper bound let c be a positive integer such that $2c < \frac{\ln(np)}{\ln(np^2)} \leq 2c + 2$ and let $\zeta = \frac{\ln(np)}{\ln(np^2)} - 2c$. Thus we have that

$$\begin{aligned}\operatorname{diam}(\mathcal{SG}(Q_k, p)) &\leq 2\left\lceil \frac{1}{2}\left(\frac{\ln(np)}{\ln(np^2)} + (1 + o(1))(\lceil L\rceil + 1) + o(1)\right)\right\rceil \\ &\leq 2\left\lceil \frac{2c + \zeta + \lceil L\rceil + 1 + o(1)}{2}\right\rceil \\ &= 2c + 2\left\lceil \frac{\zeta + \lceil L\rceil + 1 + o(1)}{2}\right\rceil .\end{aligned}$$

Now this upper bound is $2c + \lceil L\rceil + 4$ if $\lceil L\rceil$ is even, while if $\lceil L\rceil$ is odd and $\zeta \leq 1$ it is $2c + \lceil L\rceil + 3$ and if $\zeta > 1$ it is $2c + \lceil L\rceil + 5$. Observing that $\lceil 2c + \zeta\rceil = 2c + 1$ if $\zeta \leq 1$ and $2c + 2$ if $\zeta > 1$, gives the bounds. The fact that the ζ is the same in both the upper and lower bounds gives that the diameter is concentrated on at most $2\lceil L\rceil + 4$ values. Finally, if $\ln(np^2) \gg \ln\ln(n)$, then $\frac{\ln\ln(n)}{\ln(np^2)}, \frac{\ln^C(n)}{np^2} \in o(1)$ for any fixed constant C and thus

$$\left\lceil \frac{\ln(np)}{\ln(np^2)}\right\rceil - 1 \leq \operatorname{diam}(\mathcal{SG}(Q_k, p)) \leq \left\lceil \frac{\ln(np)}{\ln(np^2)}\right\rceil + 2,$$

and the diameter is concentrated on at most 3 values. □

References

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

[2] Noga Alon and Edward R. Scheinerman, *Generalized sum graphs*, Graphs Combin., 8 (1992), pp. 23–29.

[3] Noga Alon and Joel H. Spencer, *The probabilistic method*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience [John Wiley & Sons], New York, second ed., 2000. With an appendix on the life and work of Paul Erdős.

[4] Deborah Bergstrand, Ken Hodges, George Jennings, Lisa Kuklinski, Janet Wiener, and Frank Harary, *The sum number of a complete graph*, Bull. Malaysian Math. Soc. (2), 12 (1989), pp. 25–28.

[5] Andrew Beveridge, *Connectivity of random cubic sum graphs*, SIAM J. Discrete Math., 24 (2010), pp. 895–909.

[6] J. Boland, R. Laskar, C. Turner, and G. Domke, *On mod sum graphs*, in Proceedings of the Twentieth Southeastern Conference on Combinatorics, Graph Theory, and Computing (Boca Raton, FL, 1989), vol. 70, 1990, pp. 131–135.

[7] Béla Bollobás, *The diameter of random graphs*, Trans. Amer. Math. Soc., 267 (1981), pp. 41–52.

[8] ———, *The evolution of random graphs*, Trans. Amer. Math. Soc., 286 (1984), pp. 257–274.

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

[10] ———, *On the diameter and radius of random subgraphs of the cube*, Random Structures Algorithms, 5 (1994), pp. 627–648.

[11] ———, *Connectivity properties of random subgraphs of the cube*, in Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993), vol. 6, 1995, pp. 221–230.

[12] Ju. D. Burtin, *The probability of connectedness of a random subgraph of an n-dimensional cube*, Problemy Peredači Informacii, 13 (1977), pp. 90–95.

[13] Fan Chung and Linyuan Lu, *The diameter of sparse random graphs*, Adv. in Appl. Math., 26 (2001), pp. 257–279.

[14] M. N. Ellingham, *Sum graphs from trees*, Ars Combin., 35 (1993), pp. 335–349.

[15] Paul Erdős and Joel Spencer, *Evolution of the n-cube*, Comput. Math. Appl., 5 (1979), pp. 33–39.

[16] Joseph A. Gallian, *A dynamic survey of graph labeling*, 2011.

[17] Ben Green, *Counting sets with small sumset, and the clique number of random Cayley graphs*, Combinatorica, 25 (2005), pp. 307–326.

[18] B. Green and I. Z. Ruzsa, *Counting sumsets and sum-free sets modulo a prime*, Studia Sci. Math. Hungar., 41 (2004), pp. 285–293.

[19] Tianbao Hao, *On sum graphs*, J. Combin. Math. Combin. Comput., 6 (1989), pp. 207–212.

[20] Frank Harary, *Sum graphs and difference graphs*, in Proceedings of the Twentieth Southeastern Conference on Combinatorics, Graph Theory, and Computing (Boca Raton, FL, 1989), vol. 72, 1990, pp. 101–108.

[21] Svante Janson, *Poisson approximation for large deviations*, Random Structures Algorithms, 1 (1990), pp. 221–229.

[22] Emma Yu Jin and Christian M. Reidys, *Random induced subgraphs of Cayley graphs induced by transpositions*, Discrete Math., 311 (2011), pp. 2496–2511.

[23] A. V. Kostochka, A. A. Sapozhenko, and K. Weber, *On random cubical graphs*, in Fourth Czechoslovakian Symposium on Combinatorics, Graphs and Complexity (Prachatice, 1990), vol. 51 of Ann. Discrete Math., North-Holland, Amsterdam, 1992, pp. 155–160.

[24] ———, *Radius and diameter of random subgraphs of the hypercube*, Random Structures Algorithms, 4 (1993), pp. 215–229.

[25] Mirka Miller, Slamin, Joseph Ryan, and William F. Smyth, *Labelling wheels for minimum sum number*, J. Combin. Math. Combin. Comput., 28 (1998), pp. 289–297. Papers in honour of Anne Penfold Street.
[26] Christian M. Reidys, *Random induced subgraphs of generalized n-cubes*, Adv. in Appl. Math., 19 (1997), pp. 360–377.
[27] C. M. Reidys, *Random subgraphs of Cayley graphs over p-groups*, European J. Combin., 21 (2000), pp. 1057–1066.
[28] ———, *Distances in random induced subgraphs of generalized n-cubes*, Combin. Probab. Comput., 11 (2002), pp. 599–605.
[29] Christian M. Reidys, *Large components in random induced subgraphs of n-cubes*, Discrete Math., 309 (2009), pp. 3113–3124.

University of Louisville, Department of Mathematics, Louisville, KY, 40292
E-mail address: stephen.young@louisville.edu