Random Latin square graphs

Demetres Christofides Klas Markström

Abstract

In this paper we introduce new models of random graphs, arising from Latin squares and including random Cayley graphs as a special case. We investigate some properties of these graphs including their clique, independence and chromatic numbers, their expansion properties as well as their connectivity and Hamiltonicity. The results obtained are compared with other models of random graphs and several similarities and differences are pointed out. For many properties our results for the general case are as strong as the known results for random Cayley graphs and sometimes improve the previously best result for the Cayley case.

1 Introduction

The concept of random graphs is a very important notion in combinatorics. Although there are several models of random graphs, by a random graph one usually refers to the model $\mathscr{G}(n,p)$, the probability space of all graphs on $[n]$ in which every edge appears independently with probability p. For standard results on random graphs we refer the reader to the textbooks of Bollobás [7] and Janson, Łuczak and Ruciński [14].

In this paper, we introduce new models of random graphs which generalise the random Cayley graphs and study some of their properties with particular interest in their relation to the model $\mathscr{G}(n,p)$. Our models arise from Latin squares. Given a group, one can obtain Latin squares by considering its multiplication table or its division table. It turns out that the random graph obtained by the division table of a group G, is exactly the random Cayley graph of G (with respect to a random subset S of G.)

Before defining our models, let us recall that a ***Latin square*** of order n is an $n \times n$ matrix L with entries from a set of n elements, such that in each row and in each column, every element appears exactly once. Given a Latin square L with entries in a set A of size n, and a subset S of A, we define the ***Latin square graph*** $G(L,S)$ on vertex set $[n]$, by joining i to j if and only if either $L_{ij} \in S$, or $L_{ji} \in S$.

Suppose we are given a sequence (L_n) of Latin squares of order n, with entries in $[n]$, say. Choosing $S \subseteq [n]$ by picking its elements independently at random with probability p, we obtain a random Latin square graph $G(L_n,S)$. We denote this model of random Latin square graphs by $\mathscr{G}(L_n,p)$. A related model is obtained by choosing a multiset S of k elements of $[n]$ by picking its elements independently and unifomly at random (with replacement). We denote this model by $\mathscr{G}(L_n, k)$. Note that our underlying graphs are simple. However, for the model $\mathscr{G}(L_n, k)$ it will be convenient for some of our results to retain multiple edges and loops. When we do this, we will denote this new model by $\mathscr{G}_m(L_n, k)$. To be more explicit, in this model the number of edges joining i to j is exactly the total number of times that L_{ij} and L_{ji} appear in S. In particular, every $G \in \mathscr{G}_m(L_n, k)$ is a $2k$-regular multigraph.

A similar model is obtained by looking at the complement of the graph $G \in \mathscr{G}(L_n, p)$. We denote this model by $\bar{\mathscr{G}}(L_n, p)$. In general, this model is not the same as $\mathscr{G}(L_n, 1-p)$, the reason being that L_{ij} is not necessarily equal to L_{ji}. However, usually it is not too difficult to translate results from one model into the other, so we will only concentrate on $G \in \mathscr{G}(L_n, p)$.

Note that, as mentioned above, our models include random Cayley graphs as a special case. Indeed, given a group G, consider the Latin square L defined by $L_{xy} = xy^{-1}$. Then, given any subset S of elements of G, the Latin square graph $G(L, S)$ is exactly the Cayley graph of G with respect to S. The multiplication table of a group is also a Latin square, giving rise to what is usually known (motivated by the abelian case) as a ***Cayley sum graph***. So this model includes random Cayley sum graphs as well.

We should mention here that there are several differences between random Cayley graphs and our more general models of random Latin square graphs. For example, random Cayley graphs are always vertex transitive. On the other hand a random Latin square graph, even if it arises from the multiplication table of a (non-abelian) group, might not even be regular. However, it is easy to see that random Latin square graphs are not far from being regular in the sense that the ratio of maximum to minimum degree is bounded above by 2.

The fact that random Latin square graphs are almost regular (in the above sense) motivates also the comparison of our models with $\mathscr{G}_{n,r}$, the probability space of all r-regular graphs on n vertices taken with the uniform measure. (As usual, it is always assumed that rn is even.)

Sometimes, it is easier to work with random Cayley graphs or random Cayley sum graphs for abelian groups, rather than random Latin square graphs. This is because we always have $L_{ij} = L_{ij}^{-1}$ in the case of Cayley graphs, and $L_{ij} = L_{ji}$ in the case of Cayley sum graphs, and so dependences between the edges can be easier to deal with. This sometimes leads to sharper results for the first two families of random graphs than for general random Latin square graphs; however we have opted to state our results only in the general case of random Latin square graphs.

It seems that the general class of random graphs arising from Latin squares have not been studied before. However there has been much interest in random Cayley graphs and random Cayley sum graphs. For example, Agarwal, Alon, Aronov and Suri [1] established an upper bound on the clique number of random Cayley graphs arising from cyclic groups and used it to construct visibility graphs of line segments in the plane which need many cliques and complete bipartite graphs to represent them. In their study of a communication problem, Alon and Orlitsky [5] proved a similar upper bound for random Cayley graphs arising from abelian groups of odd order. Green [12], using number theoretic tools, studied the clique number of various Cayley sum graphs and showed that some of them are good examples of Ramsey graphs while others are not. The diameter of random Cayley graphs with logarithmic degree was studied by Alon, Barak and Manber in [2]. Alon and Roichman [6] proved that random Cayley graphs (on sufficiently many generators) are almost surely expanders, a result which was later improved by several authors [18, 19, 8]. The fact that random Cayley graphs are expanders has several consequences for the diameter, connectivity and Hamiltonicity of such graphs. Finally, some other aspects of the diameter, connectivity and Hamiltonicity of random Cayley graphs and random Cayley digraphs were studied in [23, 22, 20, 21].

In this paper we extend many of these resutls to the general case of random Latin square graphs and show that the structure of the Latin squares have a non-trivial influence on many properties of random Latin square graphs. In Section 2 we state and discuss our main results regarding random Latin square graphs. We prove these results in Section 3, Section 4 and Section 5. In Section 6 we give further examples and open problems.

2 Statements and discussion of the results

In this section, we list our main results and make a few comments about them, comparing them with the corresponding results in the $\mathscr{G}(n,p)$ and $\mathscr{G}_{n,r}$ models. In Subsection 2.1, we will be interested in the maximum size of cliques and independent sets in $\mathscr{G}_{n,p}$, as well as the chromatic number of $\mathscr{G}_{n,p}$ and its complement. In Subsection 2.2, we will be interested in the expansion properties of random Latin square graphs as well as several consequences of these properties regarding connectivity and Hamiltonicity. For the results of this subsection it will be easier to work in the the models $\mathscr{G}_m(L_n,k)$ and $\mathscr{G}(L_n,k)$.

2.1 Cliques, independent sets and colouring

We begin with an upper bound on the clique number of random Latin square graphs. It is well known that the clique number of $\mathscr{G}(n,1/2)$, is **whp** asymptotic to $2\log_2 n$. For the case of dense random regular graphs, it was proved in [17] that the clique number of $\mathscr{G}_{n,n/2}$ is **whp** asymptotic to $2\log_2 n$.

Guided by the above results, one might hope to prove that the clique number of $\mathscr{G}(L_n,1/2)$ is **whp** $\Theta(\log n)$. However, it turns out that this is not the case. Green [12] proved that the clique number of the random Cayley sum graph on $\mathbb{Z}_2^m$, with $p=1/2$, is **whp** $\Theta(\log n \log\log n)$, where $n=2^m=|\mathbb{Z}_2^m|$. In the same paper, Green proved that the clique number of the random Cayley sum graph on $\mathbb{Z}_n$, with $p = 1/2$, is **whp** $\Theta(\log n)$. This shows that, in general, results about the model $\mathscr{G}(L_n, p)$ can depend on the actual sequence of Latin squares chosen.

To the best of our knowledge, the best known general result on the clique number is due to Alon and Orlitsky [5], which says that the clique number of a random Cayley graph arising from an abelian group of odd order n is **whp** $O((\log n)^2)$. Using similar methods, we have managed to show that the same bound is in fact true for random Latin square graphs. In particular, it is also true for random Cayley graphs arising from non-abelian groups. We believe but cannot prove that the 2 in the exponent can be reduced further. and that $c \log n \log\log n$ is possibly the right bound.

Theorem 1 (Clique number; upper bound)**.** *Let $0 < p < 1$ be a fixed constant and let $d = 1/(2p - p^2)$. Then, for almost every $G \in \mathscr{G}(L_n, p)$, we have*

$$\mathrm{cl}(G) \leqslant 27 \left(\log_d n\right)^2 .$$

Since the model $\bar{\mathscr{G}}(L_n, p)$ is different from $\mathscr{G}(L_n, 1-p)$, we cannot immediately deduce a corresponding upper bound for the independence number. One way to find such a bound is to couple the model $\bar{\mathscr{G}}(L_n, p)$ with $\mathscr{G}(L_n, 1-p)$, and use Theorem 1 to deduce that for almost every $G \in \mathscr{G}(L_n, p)$,

$$\alpha(G) = \mathrm{cl}(\bar{G}) \leqslant 27 \left(\log_{1/(1-p^2)} n\right)^2 .$$

In fact, using an argument similar to the one used in the proof of Theorem 1, we can obtain a slightly better result.

Theorem 2 (Independence number; upper bound)**.** *Let $0 < p < 1$ be a fixed constant and let $d = 1/(1 - p)$. Then, for almost every $G \in \mathscr{G}(L_n, p)$, we have*

$$\alpha(G) \leqslant 27 \left(\log_d n\right)^2 .$$

Recall that the ***(vertex) clique cover number*** $\theta(G)$ of a graph G is the smallest integer k such that the vertex set of G can be partitioned into k cliques. I.e. $\theta(G) = \chi(\bar{G})$. So an immediate corollary of Theorem 1 is:

Corollary 3 (Clique cover number; lower bound)**.** *Let $0 < p < 1$ be a fixed constant and let $d = 1/(2p - p^2)$. Then, for almost every $G \in \mathscr{G}(L_n, p)$, we have*

$$\theta(G) \geqslant \frac{n}{27 \left(\log_d n\right)^2} . \qquad \square$$

Similarly, Theorem 2 implies:

Corollary 4 (Chromatic number; lower bound)**.** *Let $0 < p < 1$ be a fixed constant and let $d = 1/(1 - p)$. Then, for almost every $G \in \mathscr{G}(L_n, p)$, we have*

$$\chi(G) \geqslant \frac{n}{27 \left(\log_d n\right)^2} . \qquad \square$$

We now move to our upper bound on the chromatic number of random Latin square graphs. Recall that for constant p, the chromatic number of $\mathscr{G}(n,p)$ is **whp** asymptotic to $\frac{n}{2\log_b n}$, where $b = 1/(1-p)$. A similar behaviour was proved in [17] for the case of random regular graphs of high degree. More specifically, it was proved that for any $\varepsilon > 0$, if $\varepsilon n \leqslant r \leqslant 0.9n$, then the chromatic number of $\mathscr{G}_{n,r}$ is **whp** asymptotic to $\frac{n}{2\log_b n}$, where $b = n/(n-r)$.

For the case of random Latin square graphs, we prove an upper bound of the same order of magnitude. However, since our lower bound is only of order $\frac{n}{(\log_b n)^2}$, we still do not have a sharp asymptotic result for the chromatic number. In fact, as in the case of the clique and independence numbers, we know that the chromatic number can depend on the sequence of Latin squares chosen. For example, the result of Green [12] mentioned above, that the clique number of the random Cayley sum graph on $\mathbb{Z}_n$ (with $p = 1/2$) is **whp** $\Theta(\log n)$, provides a lower bound for the chromatic number of these graphs which is of the same order of magnitude as our corresponding upper bound. On the other hand, we claim that the chromatic number of the random Cayley sum graph on $\mathbb{Z}_2^m$ is **whp** $\Theta(\frac{n}{\log n \log\log n})$, where $n = 2^m = |\mathbb{Z}_2^m|$. The lower bound follows immediately from the result of Green [12] mentioned above for the independence number of these graphs. The upper bound does not follow directly from that result, however it follows from the proof in [12] that in fact there is **whp** a $\lfloor \log m + \log\log m - 1 \rfloor$-dimensional subspace of $\mathbb{Z}_2^m$ which is an independent set. Indeed, given this result, it follows that **whp**, a random Cayley sum graph on $\mathbb{Z}_2^m$ can be partitioned into at most $\frac{4n}{\log n \log\log n}$ independent sets of this form.

In fact, our upper bound on the chromatic number will be an immediate consequence of an upper bound on the list-chromatic number. Recall that the ***list-chromatic number*** $\chi_l(G)$ of a graph G is the smallest positive integer k such that for any assignment of k-element sets $L(v)$ to the vertices of G, there is a proper vertex colouring c of G with $c(v) \in L(v)$ for every vertex v of G.

Theorem 5 (List-chromatic number; upper bound)**.** *Let $0 < p < 1$ be a fixed constant and let $d = 1/(1-p)$. Then, for almost every $G \in \mathscr{G}(L_n, p)$, we have*

$$\chi_l(G) \leqslant \frac{n}{\frac{1}{4}\log_d n - \frac{1}{2}\log_d \log_d n - 2}.$$

Corollary 6 (Chromatic number; upper bound)**.** *Let $0 < p < 1$ be a fixed constant and let $d = 1/(1-p)$. Then, for almost every $G \in \mathscr{G}(L_n, p)$, we have*

$$\chi(G) \leqslant \frac{n}{\frac{1}{4}\log_d n - \frac{1}{2}\log_d \log_d n - 2}. \quad \square$$

With similar methods we will show the following upper bound for the clique cover number.

Theorem 7 (Clique cover number; upper bound)**.** *Let $0 < p < 1$ be a fixed constant and let $d = 1/p$. Then, for almost every $G \in \mathscr{G}(L_n, p)$, we have*

$$\theta(G) \leqslant \frac{n}{\frac{1}{2}\log_d n - \log\log_d n - 6}.$$

From Theorem 6 and Theorem 7 we deduce corresponding lower bounds on the independence and clique numbers.

Corollary 8 (Independence number; lower bound)**.** *Let $0 < p < 1$ be a fixed constant and let $d = 1/(1-p)$. Then, for almost every $G \in \mathscr{G}(L_n, p)$, we have*

$$\alpha(G) \geqslant \frac{1}{2}\log_d n - \log\log_d n - 6. \quad \square$$

Corollary 9 (Clique number; lower bound)**.** *Let $0 < p < 1$ be a fixed constant and let $d = 1/p$. Then, for almost every $G \in \mathscr{G}(L_n, p)$, we have*

$$\mathrm{cl}(G) \geqslant \frac{1}{4}\log_d n - \frac{1}{2}\log_d\log_d n - 2. \quad \square$$

2.2 Expansion and related properties

Alon and Roichman [6] proved that random Cayley graphs on logarithmic number of generators are expanders **whp**. Our main result of this subsection states that a similar result holds in the case of random Latin square graphs.

Before stating our result we need to introduce some notation. Given a multigraph G, its ***adjacency matrix*** is the 0,1 matrix $A = A(G)$ with rows and columns indexed by the vertices of G, in which A_{xy} is the number of edges in G joining x to y. If G is d-regular then its ***normalised adjacency matrix*** $T = T(G)$ is defined by $T = \frac{1}{d}A$. Note that T is a real symmetric matrix, so it has an orthonormal basis of real eigenvectors. We will write $\lambda_0 \geqslant \lambda_1 \geqslant \ldots \geqslant \lambda_{n-1}$ for the eigenvalues of T. It is easy to check that $\lambda_0 = 1$ and that $\lambda_{n-1} \geqslant -1$. We will write μ for the second largest eigenvalue in absolute value, i.e. $\mu = \max\{|\lambda_1|, |\lambda_{n-1}|\}$.

Finally, for $0 < x < 1$, we define

$$H(x) = x\log(2x) + (1-x)\log(2(1-x)),$$

where we use the convention that all logarithms are natural.

We can now state our main theorem.

Theorem 10 (Second eigenvalue)**.** *Let L be an $n \times n$ Latin square with entries in $[n]$ and let $G \in \mathscr{G}_m(L, k)$. Then, for every $0 < \varepsilon < 1$,*

$$\Pr(\mu(G) \geqslant \varepsilon) \leqslant 2n\exp\left\{-kH\left(\frac{1+\varepsilon}{2}\right)\right\} \leqslant 2n\exp\left\{-\frac{k\varepsilon^2}{2}\right\}.$$

We remark that if L is the difference table of a group, then the above theorem is similar to the result of Alon and Roichman mentioned in the beginning of this subsection. The only difference is that the bounds appearing in the above theorem, are the same as the bounds appearing in the authors' proof [8] of the Alon-Roichman theorem and are slightly better than the original bounds of the Alon-Roichman theorem.

Recall that a graph G is an (n, d, ε)-***expander*** if it is a graph on n vertices with maximum degree d such that for every subset W of its vertices of size at most $n/2$ we have $|N(W) \setminus W| \geqslant \varepsilon|W|$, where $N(W)$ denotes the neighbourhood of W. Note that for this definition we may ignore any multiple edges or loops that G may have. For more on expander graphs and their applications, we refer the reader to the recent survey of Hoory, Linial and Wigderson [13].

It is well known [25, 4] that a small second eigenvalue implies good expansion properties. The following corollary is an immediate consequence of Theorem 10 together with this fact.

Corollary 11 (Expansion). *For every $\delta > 0$, there is a $c(\delta) > 0$ depending only on δ, such that almost every $G \in \mathscr{G}_m(L_n, c(\delta)\log n)$ is an $(n, 2c(\delta)\log n, \delta)$-expander.* □

The fact that the second eigenvalue of the graph is small implies that such a graph has several properties that many 'random-like' graphs possess. Informally, a graph of density p is pseudorandom if its edge distribution resembles the edge distribution of $\mathscr{G}(n, p)$. The study of pseudorandom graphs was initiated by Thomason in [26, 27]. Chung, Graham and Wilson [10] showed that many properties that a graph may possess, including the property of having small second eigenvalue, are in some sense equivalent to pseudorandomness.

Here we list just a few of these consequences, mostly taken from the recent survey of Krivelevich and Sudakov [16]. We omit some of the proofs, but we note that some care needs to be takne since our graphs are multigraphs, while the result in the survey are stated only for simple graphs.

To begin with, let us consider what value of k guarantees that almost every $G \in \mathscr{G}(L_n, k)$ is connected. Let us first recall the corresponding results in $\mathscr{G}(n, p)$ and $\mathscr{G}_{n,r}$. It is well known that for any fixed $\delta > 0$, if $p \leqslant (1-\delta)\log n/n$, then $\mathscr{G}(n, p)$ is **whp** disconnected, while if $p \geqslant (1+\delta)\log n/n$, then $\mathscr{G}(n, p)$ is **whp** connected. On the other hand, $\mathscr{G}_{n,r}$ is **whp** connected provided that $r \geqslant 3$.

So what is the right threshold for the connectivity of random Latin square graphs? Once again this depends on the sequence (L_n) of Latin squares chosen. For example, the Cayley graph of $\mathbb{Z}_q$ for q prime, with respect to any set S containing a non-trivial element is connected. On the other hand, the Cayley graph of $G = \mathbb{Z}_2^m$ with respect to any set of size less than $m = \log_2 |G|$ is disconnected. Here, we prove that choosing slightly more elements are enough to guarantee **whp** the connectedness not only of the random Cayley graph of $\mathbb{Z}_2^m$ but in fact the connectedness of any random Latin square graph.

Theorem 12 (Connectedness). *For any fixed $\delta > 0$, almost every $G \in \mathscr{G}(L_n, (1+\delta)\log_2 n)$ is* **whp** *connected.*

Let us now move to the vertex connectivity of random Latin square graphs. Recall that the ***vertex connectivity*** $\kappa(G)$ of a graph G is the minimal number of vertices that we need to remove in order to disconnect G. Clearly the vertex connectivity of any graph is at most its minimum degree $\delta(G)$. It is well known that for $G \in \mathscr{G}(n,p)$ we have $\kappa(G) = \delta(G)$. Recently, it was shown in [17, 11] that the same holds for random r-regular graphs provided $3 \leqslant r \leqslant n-4$. In our case, the Cayley graphs on $\mathbb{Z}_2^m$ show that no such result can hold if the generating set S has size less than $\log_2 n$. Can we expect that such a result holds if the size of S is large enough? As the following example shows the answer is no.

Example. Define a Latin square L on $\{0,1,\dots,r-1\} \times \{0,1\}$ with entries in $\{0,1,\dots,2r-1\}$ as follows:

$$L_{(x,0),(y,0)} = \begin{cases} x+y & \text{if } x \leqslant y \\ x+y+r & \text{if } x > y \end{cases}$$

$$L_{(x,0),(y,1)} = \begin{cases} x+y+r & \text{if } x \leqslant y \\ x+y & \text{if } x > y \end{cases}$$

$$L_{(x,1),(y,0)} = \begin{cases} x+y+r & \text{if } x \leqslant y \\ x+y & \text{if } x > y \end{cases}$$

$$L_{(x,1),(y,1)} = \begin{cases} x+y & \text{if } x \leqslant y \\ x+y+r & \text{if } x > y \end{cases}$$

Here, addition is done modulo $2r$. It can be easily checked that L is indeed a Latin square. Pick any $S \subseteq \{0,1,\dots,2r-1\}$ and let $G = G(L,S)$. Note that G is d-regular for some d. Note also that for any $x \in \{0,1,\dots,r-1\}$, we have that $N_G((x,0)) \setminus \{(x,1)\} = N_G((x,1)) \setminus \{(x,0)\}$, where N_G denotes the neigbourhood of a vertex in G. But then, if $(x,0)$ is adjacent to $(x,1)$ for some x, and G is not complete, we have that $\kappa(G) \leqslant d-1$. Indeed, $N_G((x,0)) \setminus \{(x,1)\}$ is a disconnecting set of size $d-1$. Now $(x,0)$ is adjacent to $(x,1)$ if and only if $2x+r \in S$. Let $p = p(r) \in (0,1)$ be chosen such that $pr \to \infty$ and $(1-p)r \to \infty$ as $r \to \infty$ and choose S by picking its elements independently at random with probability p. Then **whp** G is not complete and there is an x such that $(x,0)$ is adjacent to $(x,1)$ and so $\kappa(G) \leqslant \delta(G) - 1$.

The above example shows that even if the size of S is large enough the vertex connectivity of a random Latin square qraph can be **whp** strictly smaller than its minimum degree. However, our next theorem shows that if S is large enough then the vertex connectivity of a random Latin square graph is **whp** at most one less than its minimum degree.

Theorem 13 (Vertex connectivity)**.** *There is an absolute constant $C \leqslant 168$ such that whenever $C \log n \leqslant k \leqslant n/4$, then $\delta(G) - 1 \leqslant \kappa(G) \leqslant \delta(G)$ for almost every $G \in \mathscr{G}(L_n, k)$.*

The example of $\mathbb{Z}_2^m$ shows that we cannot take C to be equal to 1. It would be interesting to know whether every C strictly larger than 1 works or not. It seems that our proof cannot bring the value of C down to $1+\delta$ for any $\delta > 0$, so we have not tried to optimize the value of C that our proof gives.

It should be noted that above result is not a direct consequence of the expansion properties of random Latin square graphs. From Theorem 10, we can only deduce that $\mu = O(\sqrt{\log n/k})$. However one can construct examples of d-regular graphs on n vertices, with $d = \Omega(\log n)$, $\mu = \Omega(\sqrt{\log n/d})$ but $\kappa(G) \leqslant d - \Omega(\log n)$. We refer the reader to the discussion following [16, Theorem 4.1] for more details about how one can construct such a graph.

Similar to the vertex connectivity, the ***edge connectivity*** $\lambda(G)$ of a graph G is the minimal number of edges that we need to remove in order to disconnect G. It is easy to show that $\kappa(G) \leqslant \lambda(G) \leqslant \delta(G)$. Hence, Theorem 13 applies with $\kappa(G)$ replaced by $\lambda(G)$. In fact, our next theorem shows that we can do a bit more. If $|S| \geqslant (1+\delta)\log_2 n$ then **whp** the edge connectivity is equal to the minimum degree of G. In view of random Cayley graphs on $\mathbb{Z}_2^m$, this is in fact best possible.

Theorem 14 (Edge connectivity)**.** *For any $\delta > 0$, if L is an $n \times n$ Latin square with entries in $[n]$ and S is a set of $(1+\delta)\log_2 n$ elements of $[n]$, chosen independently and uniformly at random, then* **whp***, $\lambda(G(L,S)) = \delta(G(L,S))$.*

Another graph property which follows from pseudorandomness is that of Hamiltonicity. Again, this property depends on the structure of the Latin square. For example, the Cayley graph of $\mathbb{Z}_q$ for q prime, with respect to any non-trivial element is Hamiltonian. On the other hand, as it was mentioned earlier, the Cayley graph of $G = \mathbb{Z}_2^m$ with respect to any set of size less than $m = \log_2 |G|$ is not even connected. A very appealing conjecture attributed to Lovasz, states that every connected Cayley graph is Hamiltonian. This would imply for example that every random Cayley graph on $(1+\delta)\log_2 n$ generators is Hamiltonian. However, even this consequence is still not known. Recently, Krivelevich and Sudakov [15] proved that every d-regular graph on n vertices satisfying

$$\mu \leqslant \frac{(\log\log n)^2}{1000\log n(\log\log\log n)},$$

is Hamiltonian, provided n is large enough. Using this, together with the proof technique of the Alon-Roichman theorem, they proved that a random Cayley graph on $O((\log n)^5)$ generators is **whp** Hamiltonian. Here, we extend this result to random Latin square graphs as well. Moreover, using Theorem 10 directly, we can in fact replace $(\log n)^5$ by $(\log n)^3$.

Theorem 15 (Hamiltonicity)**.** *Let $k = \frac{(\log n)^3(\log\log\log n)^2}{(\log\log n)^2}\omega(n)$, where $\omega(n)$ is any function of n which tends to infinity as n tends to infinity. Then almost every $G \in \mathscr{G}(L_n, k)$ is Hamiltonian.*

3 Cliques and independent sets

We begin by finding upper bounds for the clique number of random Latin square graphs. Naturally, one would like to find a good upper bound for the expected number of d-cliques of a random Latin square graph, and from this deduce a corresponding upper bound for the clique number. Given $A \subseteq [n]$ let $A' = \{L_{ij} : i, j \in A, i \neq j\}$. If $|A| = d$, then $|A'|$ can be as large as $\binom{d}{2}$ and as small as $d-1$. In the former case, the probability that A forms a clique in $\mathscr{G}(L_n, p)$ is $(2p - p^2)^{\binom{d}{2}}$. However, in the latter case, this probability is at least p^{d-1}. So, unless one is able to bound the number of $A \subseteq [n]$ for which $|A'|$ is relatively small, then this approach cannot give any good bounds. Our approach will be to show that any $A \subseteq [n]$ of size d, has a subset B of size $\Omega(\sqrt{d})$, such that $|B'|$ is relatively large, i.e. $\Omega(|B|^2)$. By standard arguments it will then follow that **whp** (if d is large enough,) no such B forms a clique, and hence no $A \subseteq [n]$ of size d forms a clique. Before stating our main lemma, we need to introduce some more notation.

$$n_2(A) = |\{\{i,j\} : i, j \in A \text{ distinct and } L_{ij} = L_{ji}\}|;$$

$$n_3(A) = |\{(i,j,k) : i, j, k \in A \text{ distinct and } L_{ij} = L_{jk}\}|;$$

$$n_4(A) = |\{\{(i,j),(k,l)\} : i, j, k, l \in A \text{ distinct and } L_{ij} = L_{kl}\}|.$$

If $x \in A'$ appears exactly r_x times as L_{ij} for distinct $i, j \in A$, then, with the above notation, we have

$$n_2(A) + n_3(A) + n_4(A) = \sum_{x \in A'} \binom{r_x}{2}.$$

We are now ready to state and prove our main lemma.

Lemma 16. *Let A be a set of elements of X of size a. Then for every $b \leqslant a$, A contains a subset B of size b such that*

$$|B'| \geqslant b(b-1)\left(1 - \frac{b-2}{a-2} - \frac{(b-2)(b-3)}{2(a-3)}\right) - n_2(B).$$

Proof. For any $B \subseteq A$ of size b, we have

$$\begin{aligned} |B'| &= b(b-1) - \sum_{x \in B'} (r_x - 1) \\ &\geqslant b(b-1) - \sum_{x \in B'} \binom{r_x}{2} \\ &= b(b-1) - n_2(B) - n_3(B) - n_4(B). \end{aligned}$$

Picking B at random from all b element subsets of A, we have

$\mathbb{E}(n_3(B)) = n_3(A)\frac{b(b-1)(b-2)}{a(a-1)(a-2)}$; and

$\mathbb{E}(n_4(B)) = n_4(A)\frac{b(b-1)(b-2)(b-3)}{a(a-1)(a-2)(a-3)}$.

Fixing distinct $i, j \in A$, there is exactly one $k \in [n]$ such that $L_{ij} = L_{jk}$, hence $n_3(A) \leqslant a(a-1)$. Similarly, fixing distinct $i, j, k \in A$, there is exactly one $L \in [n]$ such that $L_{ij} = L_{kl}$, hence $n_4(A) \leqslant \frac{a(a-1)(a-2)}{2}$. It follows that

$$\mathbb{E}(|B'| + n_2(B)) \geqslant b(b-1)\left(1 - \frac{b-2}{a-2} - \frac{(b-2)(b-3)}{2(a-3)}\right),$$

and hence there is a choice of B satisfying the requirements of the lemma. □

We can now prove Theorem 1.

Proof of Theorem 1. Let $d = 1/(2p - p^2)$, let $b = 3\log_d n$ and let $a = 3b^2$. Pick any $A \subseteq [n]$ of size a. By Lemma 16, there is a $B \subseteq A$ of size b, such that

$$|B'| \geqslant \frac{5}{6}b^2 - n_2(B) + O(b).$$

Pick $|B'|$ pairs (i, j) in $B \times B$, with $i \neq j$, such that all L_{ij} are distinct. Suppose that for exactly k of the pairs we have $L_{ij} = L_{ji}$. It follows that there are at least

$$(|B'| - k) - \left(\binom{b}{2} - n_2(B)\right) = \frac{1}{3}b^2 - k + O(b),$$

sets $\{i, j\}$, such that both (i, j) and (j, i) have been chosen (and so $L_{ij} \neq L_{ji}$). Therefore, the probability that B is a clique is at most

$$p^k \left(2p - p^2\right)^{\frac{1}{3}b^2 - k + O(b)} \leqslant \left(2p - p^2\right)^{\frac{1}{3}b^2 + O(b)}.$$

So the expected number of cliques $B \subseteq [n]$ of size b with $|B'| \geqslant \frac{5}{6}b^2 - n_2(B) + O(b)$ is at most

$$\binom{n}{b}\left(2p - p^2\right)^{\frac{1}{3}b^2 + O(b)} \leqslant \frac{1}{b!}\left(n\left(2p - p^2\right)^{\frac{1}{3}b + O(1)}\right)^b = o(1).$$

Thus, by Markov's Inequality, we deduce that **whp**, no such B exists. By Theorem 16, it now follows that **whp**, there is no clique of size $3b^2$, as required. □

In a similar way, we can prove the upper bound for the independence number.

Proof of Theorem 2. Let $b = 3\log_{1/(1-p)} n$, and let $a = 3b^2$. Pick any $A \subseteq [n]$ of size a. By Lemma 16, there is a $B \subseteq A$ of size b, such that

$$|B'| \geqslant \frac{5}{6}b^2 - n_2(B) + O(b) \geqslant \frac{1}{3}b^2 + O(b).$$

Therefore, the probability that B is an independent set, is at most $(1-p)^{b^2/3 + O(b)}$, and so the expected number of independent sets is

$$\binom{n}{b}(1-p)^{b^2/3 + O(b)} \leqslant \frac{1}{b!}\left(n(1-p)^{b/3 + O(1)}\right)^b = o(1).$$

Thus, by Markov's Inequality, we deduce that **whp**, no such B exists. By Theorem 16 it now follows that **whp**, there is no independent set of size $3b^2$, as required. □

4 Colouring

We now move to the proof of the upper bounds on the chromatic number. Before presenting our proof, let us see why a standard approach from the theory of random graphs does not seem to generalise in a straightforward manner.

Suppose we could show that **whp**, every induced subgraph of $G \in \mathscr{G}(L_n, 1/2)$ on $n_1 = n/(\log n)^2$ vertices has an independent set of size at least $s_1 = (2-\varepsilon)\log_2 n$. It then follows immediately that **whp**, the chromatic number is at most $n/s_1 + n_1 \sim n/2(\log_2 n)$. To do this, one usually shows that the probability that a given induced subgraph on n_1 vertices does not contain an independent set of size s_1 is $O(\exp\left\{-n^{1+\delta}\right\})$, for some $\delta > 0$. However in our model, this is far from being true. In fact, the probability that $G \in \mathscr{G}(L_n, 1/2)$ is empty is 2^{-n}, which is much larger than $O(\exp\left\{-n^{1+\delta}\right\})$. It turns out that this problem can be rectified by using the expansion properties of the graph G. We refer the reader to [3] to see how one can do this. Here, we will use a different approach from which we can obtain a better constant in the bound.

Another approach for finding an upper bound for the chromatic number, is to analyse the greedy algorithm. This is the approach that we are going to use. This approach will in fact give an upper bound on the list-chromatic number as well. However, we need to modify the standard argument, because of the dependencies in the appearance of edges. In our modification we will make use of Talagrand's Inequality [24]. We will use the following version taken (essentially) from [14].

Talagrand's Inequality. Let X be a non-negative integer valued random variable, not identically 0, which is determined by n independent random variables and let M be a median of X. Suppose also that there exist K and r such that

1. X is K-Lipschitz. I.e. changing the outcome of one of the variables, changes the value of X by at most K.

2. For any s, if $X \geqslant s$, then there is a set of at most rs of the variables, whose outcome certifies that $X \geqslant s$.

Then

$$\Pr(|X - M| \geqslant t) \leqslant \begin{cases} 4\exp\left\{-\frac{t^2}{8rK^2M}\right\} & \text{if } 0 \leqslant t \leqslant M; \\ 2\exp\left\{-\frac{t}{8rK^2}\right\} & \text{if } t > M. \end{cases}$$

In particular, it follows that,

$$\begin{aligned} |\mathbb{E}X - M| \leqslant \mathbb{E}|X - M| &= \int_0^\infty \Pr(|X - M| > t)\, dt \\ &\leqslant 4\int_0^M \exp\left\{-\frac{t^2}{8rK^2M}\right\} dt + 2\int_M^\infty \exp\left\{-\frac{t}{8rK^2}\right\} dt \\ &\leqslant 2K\sqrt{8\pi rM} + 16rK^2. \end{aligned}$$

Since also $M = 2M \Pr(X \geqslant M) \leqslant 2\mathbb{E}X$, we deduce that for $0 \leqslant t \leqslant \mathbb{E}X$

$$\Pr\left(|X - \mathbb{E}X| \geqslant t + 16rK^2 + 16K\sqrt{r\mathbb{E}X}\right) \leqslant 4\exp\left\{-\frac{t^2}{16rK^2\mathbb{E}X}\right\}.$$

This is the form of Talagrand's Inequality that we will be using.

Let us now proceed to the proof of Theorem 5.

Proof of Theorem 5. Let $d = 1/(1-p)$ and let $u = \frac{1}{4}\log_d n - \frac{1}{2}\log_d\log_d n - 2$. Suppose every vertex v has a list $L(v)$ of size $\lfloor n/u \rfloor$. Fix an ordering $v_1, \ldots, v_n$ of the vertices. Suppose we are given a (not necessarily proper) colouring c of vertices $v_1, \ldots, v_m$, such that $c(v_i) \in L(v_i)$ for each $1 \leqslant i \leqslant m$. Suppose $L(v_{m+1}) = \{x_1, \ldots, x_{\lfloor n/u \rfloor}\}$, let $c_i = c_i(m)$ be the number of times that colour x_i is used on vertices $v_1, \ldots, v_m$ and let A_{m+1} be the event that v_{m+1} has an earlier neighbour in every colour of the list $L(v_{m+1})$. We claim that $\Pr(A_{m+1}) = o(1/n)$. Having proved this, we proceed by list-colouring the graph greedily. The probability that this fails is at most $\sum_{m=1}^{n} \Pr(A_m) = o(1)$, so by Markov, we have **whp** $\chi_l(G) \leqslant n/u$.

To prove our claim, let $B_i = B_i(m)$ be the event that v_{m+1} is joined with an earlier vertex of colour x_i. Then clearly $\Pr(B_i) \leqslant 1 - (1-p)^{2c_i}$. Let Y be the number of colours in $L(v_{m+1})$ appearing on earlier neighbours of v_{m+1}. Then

$$\mathbb{E}Y \leqslant \frac{n}{u} - \sum (1-p)^{2c_i} \leqslant \frac{n}{u} - \frac{n}{u}(1-p)^{2mu/n} \leqslant \frac{n}{u}\left(1 - (1-p)^{2u}\right),$$

where the second inequality follows from the Arithmetic-Geometric Mean Inequality. Let $X = Y - \mathbb{E}Y + \frac{n}{u}(1 - (1-p)^{2u})$ and let $t = c\frac{n}{u}(1-p)^{2u}$, for some $0 < c < 1$ to be determined later. Then X satisfies the conditions of Talagrand's Inequality with $K = 2$ and $r = 1$. Note that, for n large enough, $0 \leqslant t \leqslant \mathbb{E}X$, so

$$\begin{aligned}\Pr\left(|X - \mathbb{E}X| \geqslant c\frac{n}{u}(1-p)^{2u} + 64 + 32\sqrt{\frac{n}{u}(1 - (1-p)^{2u})}\right) &\leqslant \\ 4\exp\left\{-\frac{c^2 n(1-p)^{4u}}{64u}\right\} = 4\exp\left\{-\frac{c^2 n\left(\log_d n\right)^2}{64u(1-p)^4}\right\} &\leqslant \\ 4\exp\left\{-\frac{c^2 \log n}{16(1-p)^8\log\left(1/(1-p)\right)}\right\}&.\end{aligned}$$

By elementary calculus, it is easy to show that $16x^8\log(1/x) \leqslant 2/e$ whenever $0 < x < 1$. Hence, choosing any c with $\sqrt{2/e} < c < 1$, we deduce that

$$\Pr\left(|X - \mathbb{E}X| \geqslant c\frac{n}{u}(1-p)^{2u} + 64 + 32\sqrt{\frac{n}{u}(1 - (1-p)^{2u})}\right) = o(1/n).$$

In particular, since $Y \leqslant X$,

$$\Pr\left(Y \geqslant \frac{n}{u} - c\frac{n}{u}(1-p)^{2u} + 64 + 32\sqrt{\frac{n}{u}}\right) = o(1/n).$$

Since

$$\frac{n}{u}(1-p)^{4u} = \frac{(\log_d n)^2}{u(1-p)^8} \to \infty,$$

we deduce that (for n large enough,)

$$\Pr(A_{m+1}) = \Pr(Y \geqslant \lfloor n/u \rfloor) = o(1/n). \qquad \square$$

Similarly, we can give an upper bound to the clique cover number.

Proof of Theorem 7. Let $d = 1/p$ and let $u = \frac{1}{2}\log_d n - \log_d \log_d n - 6$. Fix an ordering of the vertices. Suppose we are given a not necessarily proper colouring of the first m vertices of $\bar{G}$, using colours 1 up to $\lfloor n/u \rfloor$. Let $c_i = c_i(m)$ be the number of times colour i is used and let A_{m+1} be the event that the $(m+1)$-th vertex has a neighbour in every colour. We claim that $\Pr(A_{m+1}) = o(1/n)$. Having proved this, we colour the graph greedily. The probability that we need more than $\lfloor n/u \rfloor$ colours is at most $\sum_{m=1}^{n} \Pr(A_m) = o(1)$, so by Markov, we have **whp** $\theta(G) = \chi(\bar{G}) \leqslant n/u$.

To prove our claim, let $B_i = B_i(m)$ be the event that the $(m+1)$-th vertex is joined (in $\bar{G}$,) with an earlier vertex of colour i. Then clearly $\Pr(B_i) \leqslant 1 - p^{c_i}$. Let Y be the number of colours appearing on earlier neighbours of the $(m+1)$-th vertex. Then

$$\mathbb{E}Y \leqslant \frac{n}{u} - \sum p^{c_i} \leqslant \frac{n}{u} - \frac{n}{u}p^{mu/n} \leqslant \frac{n}{u}\left(1 - p^u\right).$$

Let $X = Y - \mathbb{E}Y + \frac{n}{u}(1-p^u)$ and let $t = c\frac{n}{u}p^u$, for some $0 < c < 1$ to be determined later. Then X satisfies the conditions of Talagrand's Inequality with $K = 2$ and $r = 1$. Note that, for n large enough, $0 \leqslant t \leqslant \mathbb{E}X$, so

$$\Pr\left(|X - \mathbb{E}X| \geqslant c\frac{n}{u}p^u + 64 + 32\sqrt{\frac{n}{u}(1-p^u)}\right) \leqslant 4\exp\left\{-\frac{c^2 n p^{2u}}{64u}\right\} \leqslant 4\exp\left\{-\frac{c^2 \log n}{32 p^{12} \log(1/p)}\right\}.$$

But $32x^{12}\log(1/x) \leqslant 8/(3e) < 1$ whenever $0 < x < 1$. Hence, choosing any c with $\sqrt{8/3e} < c < 1$, we deduce that

$$\Pr\left(|X - \mathbb{E}X| \geqslant c\frac{n}{u}p^u + 64 + 32\sqrt{\frac{n}{u}(1-p^u)}\right) = o(1/n).$$

In particular, since $Y \leqslant X$,

$$\Pr\left(Y \geqslant \frac{n}{u} - c\frac{n}{u}p^{2u} + 64 + 32\sqrt{\frac{n}{u}}\right) = o(1/n).$$

Since

$$\frac{n}{u}p^{2u} = \frac{(\log_d n)^2}{up^{12}} \to \infty,$$

we deduce that (for n large enough,)

$$\Pr(A_{m+1}) = \Pr(Y \geqslant \lfloor n/u \rfloor) = o(1/n). \qquad \square$$

5 Expansion and consequences

We now proceed to the expansion properties of random Latin square graphs and to the proof of Theorem 10 on the second eigenvalue of such graphs. In [8], we generalized Hoeffding's inequality, to an inequality where the random variables do not necessarily take real values, but instead take their values in the set of (self-adjoint) operators of a (finite dimensional) Hilbert space. We then used this inequality to give a new proof of the Alon-Roichman theorem. The main tool in the proof of Theorem 10 will be this Operator Hoeffding Inequality. Before stating the inequality, we need to introduce some more notation.

Let V be a Hilbert space of dimension d, let $A(V)$ be the set of self adjoint operators on V and let $P(V)$ be the be the cone of positive operators on V, i.e.

$$P(V) = \{A \in A(V) : \text{all eigenvalues of } A \text{ are nonnegative}\}.$$

This defines a partial order on $A(V)$ by $A \leqslant B$ iff $B - A \in P(V)$. We denote by $[A, B]$ the set of all $C \in A(V)$ such that $A \leqslant C \leqslant B$. We also denote by $\|A\|$ the largest eigenvalue of A in absolute value.

We can now state our Operator Hoeffding Inequality. We refer the reader to [8] for its proof.

Theorem 17 ([8] Operator Hoeffding Inequality)**.** *Let V be a Hilbert space of dimension d and let $X_i = \mathbb{E}(X|\mathscr{F}_i)$ be a martingale, taking values in $A(V)$, whose difference sequence satisfies $Y_i \in [-\frac{1}{2}I, \frac{1}{2}I]$. Then*

$$\Pr(\|X - \mathbb{E}X\| \geqslant nh) \leqslant 2d \exp\{-nH(1/2 + h)\}.$$

Note that the case $d = 1$ of this inequality is exactly Hoeffding's inequality.

We now proceed to show that random Latin square graphs have small second eigenvalue and thus good expansion properties.

Proof of Theorem 10. Let $s_1, \ldots, s_k$ be elements of $[n]$ chosen independently and uniformly at random. For $s \in [n]$ let $L(s)$ be the 0,1 matrix in which $L(s)_{ij} = 1$ if and only if $L_{ij} = s$. So the normalised adjacency matrix of the multigraph G generated by these elements is $T = \frac{1}{2k}\sum_{i=1}^{k}(L(s_i) + L(s_i)^T)$. Let $B = T - \frac{1}{n}J$, where J is the n by n matrix having '1' in every entry. We claim that $\mu(G) = \|B\|$, where J is the having '1' in every entry. Indeed, if $\{v_0, v_1, \ldots, v_{n-1}\}$ is an orthonormal basis of T, with each v_i having eigenvalue λ_i, and $v_0 = \frac{1}{\sqrt{n}}(1, \ldots, 1)$, then $Bv_0 = 0$ and $Bv_i = \lambda_i v_i$, so $\mu(G) = \|B\|$ as required. Let Y_i be the operator whose matrix is $\frac{1}{4}\left(L(s_i) + L(s_i)^T - \frac{2}{n}J\right)$. It is easy to check that $X_i = Y_1 + \ldots + Y_i$ is a martingale satisfying the conditions of the theorem. It follows that

$$\begin{aligned}\Pr(\mu(G) \geqslant \varepsilon) &= \Pr\left(\left\|\frac{1}{k}\sum_{i=1}^{k} Y_i\right\| \geqslant \frac{\varepsilon}{2}\right)\\ &= \Pr\left(\|X - \mathbb{E}X\| \geqslant \frac{\varepsilon k}{2}\right)\\ &\leqslant 2n \exp\left\{-kH\left(\frac{1+\varepsilon}{2}\right)\right\},\end{aligned}$$

as required. □

Proof of Theorem 12. It is enough to prove the result for $0 < \delta < 1/2$. Note that $H(x)$ is continuous in $(0, 1)$ and tends to $\log 2$ as x tends to 1. Pick an x such that $H(x) \geqslant (1 - \delta/2) \log 2$. Then, for $k = |S| = (1 + \delta) \log_2 n$, we have $kH(x) \geqslant (1 + \delta/4) \log n$. Thus,

$$\Pr(\mu(G(L, S)) \geqslant 2x - 1) \leqslant 2n \exp\{-(1 + \delta/4) \log n\} = 2n^{-\delta/4} = o(1).$$

Thus **whp**, $\mu(G(L, S)) < 2x - 1 < 1$. It is well known that if $\mu(G) < 1$ then G is connected, so the result follows. □

Proof of Theorem 13. Let T be a minimial disconnecting set, so $|T| \leqslant 2k$. Let U be the smallest component of $G \setminus T$ and let $W = V(G) \setminus (U \cup T)$. Note that $|W| \geqslant n/4$. We claim that **whp**, $|U| \leqslant 128 \log n$. Our proof of this claim is very similar to [16, Theorem 4.1]. Firstly, we deduce from Theorem 10 that **whp** $\mu(G) \leqslant 4\sqrt{k \log n}$. Since there are no edges from U to W, it follows from the edge distribution bound for pseudorandom graphs (see e.g. [16, Theorem 2.11]) that

$$\frac{2k}{n}|U||W| < \mu\sqrt{|U||W|}$$

and so

$$|U| < \frac{\mu^2 n^2}{4k^2|W|} \leqslant \frac{\mu^2 n}{k^2} \leqslant \frac{16n \log n}{k}.$$

Using [16, Theorem 2.11] again, we deduce that the number of edges having both endpoints in U (counted with multiplicity) satisfy

$$e(U, U) \leqslant \frac{2k}{n}|U|^2 + \mu|U| < (32 \log n + 4\sqrt{k \log n})|U| \leqslant \frac{k}{2}|U|,$$

provided C is large enough. It follows that the number of edges with exactly one endpoint in U and one in T satisfy

$$e(U, T) = 2k|U| - e(U, U) > \frac{3k}{2}|U|.$$

On the other hand, using [16, Theorem 2.11] once more, we have

$$e(U, T) \leqslant \frac{2k}{n}|U||S| + \mu\sqrt{|U||S|} \leqslant \left(1 + 4\sqrt{\frac{2 \log n}{|U|}}\right) k|U| \leqslant \frac{3k}{2}|U|,$$

a contradiction. So we may assume that $|U| \leqslant 128 \log n$.

We now claim that **whp**, the following holds: For any 3 distinct vertices x, y, z of G, $|(N(x) \cup N(y)) \setminus N(z)| > 128 \log n$, where $N(x)$ denotes the neighbourhood of the vertex x. Having proved this, it will follow that **whp** $|U| \leqslant 2$ and so $|T| \geqslant \delta(G) - 1$.

So, let x, y, z be distinct vertices of G. Let $s_1, s_2, \ldots, s_k$ be the elements of S chosen uniformly at random and let $X_i = \mathbb{E}(|(N(x) \cup N(y)) \setminus N(z)||s_1, \ldots, s_i)$. Then $X_0, X_1, \ldots, X_k$ is a martingale with Lipschitz constant 4 and $X_0 \geqslant k(1 - 2/n)^k$. It follows by the Hoffding-Azuma inequality that

$$\Pr(|(N(x) \cup N(y)) \setminus N(z)| \leqslant k(1 - 2/n)^k - t) \leqslant \exp\left\{-\frac{t^2}{2k}\right\}.$$

Letting $t = \sqrt{8k \log n}$ we obtain that

$$\Pr(|(N(x) \cup N(y)) \setminus N(z)| \leqslant 128 \log n) = O(n^{-4})$$

provided C is large enough. Our claim now follows from the union bound. This completes the proof of the theorem. (It can be checked that $C = 168$ works. □

We omit the proof of Theorem 14, as it can be proved using a similar argument as in [16, Theorem 4.3]

Sketch proof of Theorem 15. Firstly, one needs to check that the result of Krivelevich and Sudakov [15] mentioned before the statement of the theorem, also holds for d-regular multigraphs. We omit the details of this check. Then the result follows directly from Theorem 10 □

6 Conclusion and open problems

We have introduced new models of random graphs arising from Latin squares and studied some of their properties. There is still a lot of research that needs to be done even for many of the properties that we have considered here.

Regarding the clique and independence numbers it would be interesting to know if the upper bound can be reduced further. In particular, we believe (but cannot prove) that the 2 in the exponent can be reduced further. It would be also interesting to know whether there are examples of random Latin square graphs whose clique/independence number is significantly larger than $\Theta(\log n \log \log n)$.

Similar remarks hold for the lower bound on the chromatic and clique cover numbers. Any improvement on the upper bound of the independence/clique numbers would give a corresponding improvement on the chromatic/clique cover numbers but it might be possible (or even easier) to get such improvements directly.

Another interesting question which we have not been able to answer so far is the determination of the Hadwiger number of random Latin square graphs, i.e. the largest integer k such that the graph can be contracted into a K_k. We do not even know, for $p = 1/2$ say, whether this number depends on the sequence of Latin squares or not.

We have not studied at all the girth of random Latin square graphs. The reason is that it depends a lot on the structure of the Latin squares chosen. For example, almost every $G \in \mathscr{G}(\mathbb{Z}_3^m, p)$ has **whp** girth 3, provided $pn \to \infty$, where $n = 3^m$. On the other hand, we claim that almost every $G \in \mathscr{G}(\mathbb{Z}_2^m, p)$ has **whp** girth strictly greater than 3 provided that $pn^{2/3} \to 0$, where $n = 2^m$. Indeed, the expected number of triangles containing a fixed vertex x is $\binom{n-1}{2}p^3$ which tends to 0. By Markov's inequality x is **whp** not contained in any triangle. But since the graph is vertex transitive, our claim follows.

The expansion properties of random Latin squares imply that almost every $G \in \mathscr{G}(L_n, c\log_2 n)$, with $c > 1$, has logarithmic diameter. An interesting question here is the threshold for the diameter becoming equal to 2. It turns out that there are constants c_1 and c_2 such that if $p < c_1 \log n/n$, then almost every $G \in \mathscr{G}(L_n, p)$ has diameter greater than 2, while if $p > c_2 \log n/2$, then almost every $G \in \mathscr{G}(L_n, p)$ has diameter less than or equal to 2. The values of c_1 and c_2 depend on the sequence of Latin squares chosen. Our results regarding the diameter will appear in a forthcoming paper [9].

Acknowledgements

This work started when the second author was visiting the Department of Pure Mathematics and Mathematical Statistics of the University of Cambridge and continued when the first author was visiting the Deparment of Mathematics and Mathematical Statistics of Umeå University. The authors would like to thank both departments for their hospitality.

The first author would like to thank the Engineering and Physical Sciences Research Council and the Cambridge Commonwealth Trust which supprted him while he was at the University of Cambridge and also the Deparment of Mathematics and Mathematical Statistics of Umeå University for the grant making his visit there possible. The second author would like to thank the Royal Physiographic Society in Lund for the grant making his visit to the Univesity of Cambridge possible.

References

[1] P. K. Agarwal, N. Alon, B. Aronov and S.Suri, Can visibility graphs be represented compactly?, Discrete Comput. Geom. **12** (1994), 347–365.

[2] N. Alon, A. Barak and U. Mabner, On disseminating information reliably without broadcasting, in *Proc. of the 7th International Conference on Distributed Computing Systems*, 74–81, Berlin, Germany, 1987.

[3] N. Alon, M. Krivelevich and B. Sudakov, List coloring of random and pseudo-random graphs, Combinatorica **19** (1999), 453–472.

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

[5] N. Alon and A. Orlitsky, Repeated communication and Ramsey graphs, IEEE Trans. Inform. Theory **41** (1995), 1276–1289.

[6] N. Alon and Y. Roichman, Random Cayley graphs and expanders, Random Structures Algorithms **5** (1994), 271–284.

[7] B. Bollobás, *Random graphs*, Second edition, Cambridge Univ. Press, Cambridge, 2001.

[8] D. Christofides and K. Markström, Expansion properties of random Cayley graphs and vertex transitive graphs, Random Structures Algorithms **1** (2008), 88-100.

[9] D. Christofides and K. Markström, The diameter of random Latin square graphs, in preparation.

[10] F. R. K. Chung, R. L. Graham and R. M. Wilson, Quasi-random graphs, Combinatorica **9** (1989), 345–362.

[11] C. Cooper, A. Frieze and B. Reed, Random regular graphs of non-constant degree: connectivity and Hamiltonicity, Combin. Probab. Comput. **11** (2002), 249–261.

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

[13] S. Hoory, N. Linial and A. Wigderson, Expander graphs and their applications, Bull. Amer. Math. Soc. (N.S.) **43** (2006), 439–561.

[14] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley, New York, 2000.

[15] M. Krivelevich and B. Sudakov, Sparse pseudo-random graphs are Hamiltonian, J. Graph Theory **42** (2003), 17–33.

[16] M. Krivelevich and B. Sudakov, Pseudo-random graphs, in *More sets, graphs and numbers*, 199–262, Springer, Berlin.

[17] M. Krivelevich, B. Sudakov, V. H. Vu and N. C. Wormald, Random regular graphs of high degree, Random Structures Algorithms **18** (2001), 346–363.

[18] Z. Landau and A. Russell, Random Cayley graphs are expanders: a simple proof of the Alon-Roichman theorem, Electron. J. Combin. **11** (2004), Research Paper 62.

[19] P.-S. Loh and L. J. Schulman, Improved expansion of Random Cayley Graphs, Discrete Math. Theor. Comput. Sci. **6** (2004), 523–528.

[20] J. Meng, Connectivity of random Cayley digraphs, Acta Math. Sinica (Chin. Ser.) **40** (1997), 493–498.

[21] J. Meng and Q. Huang, Almost all Cayley graphs are Hamiltonian, Acta Math. Sinica (N.S.) **12** (1996), 151–155.

[22] J. Meng and Q. Huang, Almost all Cayley graphs have diameter 2, Discrete Math. **178** (1998), 267–269.

[23] J. Meng and X. Liu, The diameters of almost all Cayley digraphs, Acta Math. Appl. Sinica (English Ser.) **13** (1997), 410–413.

[24] M. Talagrand, Concentration of measure and isoperimetric inequalities in product spaces, Inst. Hautes Études Sci. Publ. Math. No. 81 (1995), 73–205.

[25] R. M. Tanner, Explicit concentrators from generalized N-gons, SIAM J. Algebraic Discrete Methods **5** (1984), 287–293.

[26] A. Thomason, Pseudorandom graphs, in *Random graphs '85 (Poznań, 1985)*, 307–331, North-Holland, Amsterdam.

[27] A. Thomason, Random graphs, strongly regular graphs and pseudorandom graphs, in *Surveys in combinatorics 1987 (New Cross, 1987)*, 173–195, Cambridge Univ. Press, Cambridge.

Demetres Christofides

School of Mathematics
University of Birmingham
Birmingham, B15 2TT
United Kingdom

christod@maths.bham.ac.uk

Klas Markström

Department of Mathematics and Mathematical Statistics
Umeå universitet
Umeå, SE-901 87
Sweden

klas.markstrom@math.umu.se