ADDENDUM TO RANDOM NILPOTENT GROUPS I

MOON DUCHIN, MENG-CHE HO, ANDREW SÁNCHEZ

1. Intro

The paper *Random Nilpotent Groups I* [1] makes various claims about random walks, and this note fills in some details for those claims. We will focus here on the simple random walk (SRW) on $\mathbb{Z}^m$ but also give indications of how to extend these arguments to non-backtracking simple random walk (NBSRW), colored blue so they are easy to find. This document will be continually updated, and once it contains a full proof of a local central limit theorem for NBSRW we plan to publish that separately.

1.1. **Fixing notation.** We fix a natural number $m \geq 2$. Consider the $\mathbb{Z}^m$-valued **simple** random walk (each step equals $\pm e_i$ with probability $1/2m$). Let us define Y_ℓ to be the random variable which with probability one-half gives the outcome of SRW after ℓ steps, and with probability one-half gives the outcome after $\ell - 1$ steps. Based on this we can define several more random variables.

Random variable	Valued in	Definition (for fixed m and k)
Y_ℓ	$\mathbb{Z}^m$	SRW after ℓ or $\ell-1$ steps, described above
R_ℓ	$\mathbb{Z}$	projection of Y_ℓ to a single coordinate
M_ℓ	$m \times m$ matrices	columns are Y_ℓ, independently sampled
Δ_ℓ	$\mathbb{Z}$	$\det(M_\ell)$
$d_\ell^{(k)}$	$\mathbb{Z}$	$\gcd(\Delta_{\ell,1}, \dots, \Delta_{\ell,k})$, with $\Delta_{\ell,i}$ independently sampled from Δ_ℓ
$M_\ell^{(k)}$	$k \times k$ matrices	$k \times k$ minor of M_ℓ (for $k \leq m$)

If E_ℓ is an event that depends on a parameter ℓ, we use the symbol $\mathbb{P}(E_\ell)$ for the probability for fixed ℓ and write $\overline{\mathbb{P}}(E_\ell) := \lim_{\ell\to\infty} \mathbb{P}(E_\ell)$ for the asymptotic probability. If E is an event with respect to a matrix-valued random variable, we use the notation $\mathbb{P}'(E)$ to denote the conditional probability of E given that no matrix entries are zero.

Note that for any $\epsilon > 0$ and for large enough ℓ, we have $\overline{\mathbb{P}}(R_\ell < \ell^{\frac{1}{2}+\epsilon}) = 1$. Indeed, the expectation for $|R_\ell|$ is well known to be $\sqrt{\frac{\ell}{m}}$; for NBSRW, the expectation is $\sqrt{\frac{\ell}{m-1}}$ by the central limit theorem from [2].

We will analyze primes by their size relative to ℓ, so we fix a small $\epsilon > 0$ and define size ranges:

$$\mathcal{P}_1 := \{p \leq \log\log \ell\} \qquad \mathcal{P}_2 := \{\log\log \ell \leq p \leq \ell^{\frac{1}{2}-\epsilon}\} \qquad \mathcal{P}_3 := \{\ell^{\frac{1}{2}-\epsilon} \leq p \leq \ell^{m+1}\} \qquad \mathcal{P}_4 := \{p \geq \ell^{m+1}\}.$$

1.2. **Result statements.** The first lemma we will discuss below is an estimate for the distribution of our $\mathbb{Z}$-valued random variables modulo n for any n.

RNG Lemma 15 (Arithmetic uniformity)**.** *With the notation defined above,*

$$\exists c_1, c_2 > 0 \text{ such that } \forall n < \ell^{\frac{1}{2}-\epsilon}, \quad \forall k, \qquad \mathbb{P}\Big(R_\ell \equiv k \mod n\Big) < \frac{1}{n} + c_1 e^{-c_2 \ell^{2\epsilon}}.$$

Furthermore, the distributions are asymptotically independent in different coordinates: For $1 \leq i \leq m$, let $Y_\ell^{(i)}$ be the projection of Y_ℓ to the ith coordinate, so that the $Y_\ell^{(i)}$ are each distributed like R_ℓ but are not a priori independent. Then $\forall n, \exists c_1, c_2 > 0$ s.t.$\forall 1 \leq s \leq m, \forall k_1, \dots, k_s$,

$$\mathbb{P}\Big(Y_\ell^{(1)} \equiv k_1, \cdots, Y_\ell^{(s)} \equiv k_s \mod n\Big) < \frac{1}{n^s} + c_1 e^{-c_2 \ell^{2\epsilon}}.$$

Using RNG Lemma 15 and a few more estimates we will conclude the following:

RNG Corollary 17 (Probability of primitivity)**.** *For a random freely reduced word in F_m, the probability that it is primitive in abelianization tends to $1/\zeta(m)$, where ζ is the Riemann zeta function. In particular, for $m = 2$, the probability is $6/\pi^2$.*

In this note we will carefully establish the following proposition.

Date: Last updated November 12, 2016.

Proposition 8 (Common divisors of random determinants)**.** *Fixing m and any $k > 4m+4$, we have*

$$\overline{\mathbb{P}}(d_\ell^{(k)} = 1) = \prod_{primes\ p} 1 - \left[1 - \left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{p^2}\right)\cdots\left(1 - \frac{1}{p^m}\right)\right]^k.$$

In the paper, this is what is needed to prove one of the main theorems.

RNG Theorem 36 (Collapsing abelian quotients)**.** *For random abelian groups $H = \mathbb{Z}^m/\langle R\rangle$ with $|R|$ random relators, if $|R| \to \infty$ as a function of ℓ, then $H = \{0\}$ with probability one (a.a.s.). If $|R|$ is bounded as a function of ℓ, then there is a positive probability of a nontrivial quotient, both for each ℓ and asymptotically.*

2. Expanded arguments

For SRW, R_ℓ is just a lazy simple random walk on $\mathbb{Z}$: at each step, it advances left or right with probability $1/2m$, and otherwise it stands still. A similar statement is true for NBSRW, but the probabilities depend on the previous step. As mentioned above, classical central limit theorems tell us that R_ℓ is asymptotically normally distributed, and this is true for the NBSRW case as well [2]. To get more detail about the rate of convergence to the Gaussian distribution, we have local central limit theorems (LCLT) which give upper bounds on the difference between the probability that $R_\ell = x$ and the estimate derived from the Gaussian, in terms of x and ℓ [3, Chapter 2].

First, we present the crucial result that random walks equidistribute mod n.

RNG Lemma 15 (Arithmetic uniformity)**.** *With the notation defined above,*

$$\exists c_1, c_2 > 0 \text{ such that } \forall n < \ell^{\frac{1}{2}-\epsilon}, \quad \forall k, \qquad \mathbb{P}\Big(R_\ell \equiv k \mod n\Big) < \frac{1}{n} + c_1 e^{-c_2 \ell^{2\epsilon}}.$$

Furthermore, the distributions are asymptotically independent in different coordinates: For $1 \le i \le m$, let $Y_\ell^{(i)}$ be the projection of Y_ℓ to the ith coordinate, so that the $Y_\ell^{(i)}$ are each distributed like R_ℓ but are not a priori independent. Then $\forall n, \exists c_1, c_2 > 0$ s.t.$\forall 1 \le s \le m, \forall k_1, \dots, k_s$,

$$\mathbb{P}\Big(Y_\ell^{(1)} \equiv k_1, \cdots, Y_\ell^{(s)} \equiv k_s \mod n\Big) < \frac{1}{n^s} + c_1 e^{-c_2 \ell^{2\epsilon}}.$$

Proof. We will consider the residues mod n by studying the position of the random walk Y_ℓ on the torus $(\mathbb{Z}/n\mathbb{Z})^m$. Furthermore, we will just consider the last statement in the case $s = m$, since all of the results for $s < m$ can be derived from this by summing: for instance, the positions satisfying $Y_\ell^{(i)} \equiv k_i$ for $i = 1, 2$ are represented by n^{m-2} positions in the torus, so the bound can be added polynomially many times to get the right main term, at the cost of slightly enlarging the constant c_2.

A theorem from Saloff-Coste [5, Theorem 7.8] controls the distance from a lazy symmetric generating random walk to the uniform distribution on any family of finite graphs which satisfies a uniform bound on volume growth. (In our case, the growth $\#B_r$ is bounded on $(\mathbb{Z}/n\mathbb{Z})^m$ by $(2r+1)^m$, independent of n, and satisfies the doubling hypothesis.) First we will explain how this theorem provides the needed bound, then we will explain how to modify Y_ℓ to satisfy the hypotheses.

In our notation, the theorem says that

$$\mathbb{P}\Big(Y_\ell^{(1)} \equiv k_1, \cdots, Y_\ell^{(m)} \equiv k_m \mod n\Big) < \frac{1}{n^m} + \frac{c_1 \cdot n^m}{\ell^{m/2}} e^{-c_2 \ell/(mn)^2},$$

for the following reasons: the L^2 distance upper-bounds the difference in probabilities at any single point, and the diameter of $(\mathbb{Z}/n\mathbb{Z})^m$ is less than mn. Since $n < \ell^{\frac{1}{2}-\epsilon}$, we have $n^2 < \ell^{1-2\epsilon}$, and by enlarging the constants we obtain

$$\mathbb{P}\Big(Y_\ell^{(1)} \equiv k_1, \cdots, Y_\ell^{(m)} \equiv k_m \mod n\Big) < \frac{1}{n^m} + c_1 e^{-c_2 \ell^{2\epsilon}},$$

as desired.

In order to use this theorem on our (non-lazy) walk, we apply the following technique: we replace the simple random walk P with the two-step walk $P * P$ which is lazy and symmetric. If n is odd, the support of $P * P$ is a generating set, and we can proceed. If n is even, $P * P$ is supported on the sublattice of torus points where the sum of the coordinates is even, which does not generate. But in that case the random variable Y_ℓ that we are studying (which takes either ℓ or $\ell - 1$ steps) lives on the even or odd sublattice with equal probability; Saloff-Coste's statement will ensure equidistribution on the even sublattice, and by symmetry, taking one more step will equidistribute on the odds. (To be precise, we should use $\ell/2$ rather than ℓ on the right-hand side because of the parity fix, but this gets absorbed in the constants.) □

Lemma 1 (Divisibility of coordinate projections)**.** *For every $m, n \geq 2$, $1 \leq s \leq m$, and $\ell \gg 1$, there is a conditional probability bound given by*

$$\mathbb{P}'(Y_\ell^{(1)} \equiv \cdots \equiv Y_\ell^{(s)} \equiv 0 \mod n) < 1/n^s.$$

In particular, $\mathbb{P}(R_\ell \equiv 0 \mod n \mid R_\ell \neq 0) < 1/n$.

Proof. We give the detailed argument for $s = 1$. Let $p_\ell(x) = \mathbb{P}(R_\ell = x)$. This result will follow from monotonicity of the distribution of R_ℓ, i.e., $p_\ell(x) > p_\ell(x+1)$ for $x \geq 0$. We proceed by induction on ℓ. For $\ell = 1$, we have $p_1(0) = \frac{2m-1}{m}$ and $p_1(1) = \frac{1}{4m}$, which establishes the base case. For $\ell > 1$, we have

$$p_\ell(x) = \frac{1}{2m}p_{\ell-1}(x-1) + \frac{m-1}{m}p_{\ell-1}(x) + \frac{1}{2m}p_{\ell-1}(x+1);$$

$$p_\ell(x+1) = \frac{1}{2m}p_{\ell-1}(x) + \frac{m-1}{m}p_{\ell-1}(x+1) + \frac{1}{2m}p_{\ell-1}(x+2).$$

Now we know that $\frac{m-1}{m} > \frac{1}{2m}$ (since $m \geq 2$), and this means

$$\frac{m-1}{m}p_{\ell-1}(x) + \frac{1}{2m}p_{\ell-1}(x+1) > \frac{1}{2m}p_{\ell-1}(x) + \frac{m-1}{m}p_{\ell-1}(x+1),$$

since the LHS has a larger coefficient on the larger term. This compares two of the terms of $p_\ell(x)$ with two of the terms of $p_\ell(x+1)$, so it only remains to compare the remaining terms. Since $x \geq 0$, we have $|x-1| \leq x+1$. Thus, by repeatedly applying the inductive hypothesis, we have $p_{\ell-1}(x-1) > p_{\ell-1}(x+2)$, which completes the proof for all ℓ. In particular, we have shown: if the positive integers $\mathbb{Z}_{>0}$ are partitioned into intervals $[kn+1, kn+n]$, then the farthest point in each interval from 0 (the value divisible by n) has the lowest probability.

For NBSRW, we would need to inspect the LCLT bounds to establish monotonicity rigorously, though it is intuitively clear for $\ell \gg 1$.

The argument for general s runs along exactly the same lines: $\mathbb{Z}_{>0}^m$ is cut up into $n \times \cdots \times n$ boxes which are obtained as products of the intervals described above, then in each box, the point farthest from the origin (which satisfies the congruence condition in the statement of the lemma) has the lowest probability in the random walk. □

Lemma 2 (Values of coordinate projections)**.** *There is a constant c such that for any $\alpha \in \mathbb{Z}$ and any $\epsilon > 0$,*

$$\mathbb{P}(R_\ell = \alpha) < \frac{c}{\sqrt{\ell}} \text{ for } \ell \gg 1.$$

Proof. Bounding $\mathbb{P}(R_\ell = \alpha)$ by a multiple of $\ell^{-1/2}$ follows from the standard LCLT and could be extended to NBSRW from its LCLT. □

With this, we can establish the probability that a random relator is primitive in abelianization.

Lemma 3 (RNG Corollary 17)**.** *Let δ_ℓ be the greatest common divisor of the entries of Y_ℓ. Then*

$$\overline{\mathbb{P}}(\delta_\ell > 1) = 1 - \frac{1}{\zeta(m)}.$$

Proof. Recall that for an event expressed in terms of a matrix-valued random variable, $\mathbb{P}'(E)$ denotes the conditional probability of E given that the entries of the matrix are nonzero (and this definition makes sense for vectors in particular). Since

$$\mathbb{P}(\delta_\ell > 1) \leq \mathbb{P}'(\text{some prime divides } \delta_\ell) + \mathbb{P}(\text{some entry of } Y_\ell \text{ is zero}),$$

we have

$$\mathbb{P}(\text{some } p \in \mathcal{P}_1 \text{ divides } \delta_\ell) \leq \mathbb{P}(\delta_\ell > 1) \leq \mathbb{P}'(\text{some } p \in \mathcal{P}_1 \text{ divides } \delta_\ell) + \mathbb{P}'(\text{some } p \in \mathcal{P}_1^c \text{ divides } \delta_\ell) + \mathbb{P}(\text{some entry is zero}).$$

Now $\mathbb{P}'(\text{some } p \in \mathcal{P}_1 \text{ divides } \delta_\ell) \to 1 - \frac{1}{\zeta(m)}$ by arithmetic uniformity and independence (Lemma 15), via the Euler product formula for the zeta function. By Lemma 1,

$$\mathbb{P}'(\text{some } p \in \mathcal{P}_1^c \text{ divides } \delta_\ell) < \sum_{p \notin \mathcal{P}_1} \frac{1}{p^m} \longrightarrow 0,$$

where the inequality is just the sum-bound $\mathbb{P}(\bigcup_i E_i) \leq \sum_i \mathbb{P}(E_i)$ and it converges to zero as the tail of a convergent sequence. Lastly, $\mathbb{P}(\text{some entry is zero}) \to 0$ and the lemma follows. □

To prove RNG Theorem 36, we now build up a series of Lemmas regarding divisibility with regards to our partition of the primes.

Lemma 4 (Divisibility of determinants by small primes)**.** *Let* $\mathbf{P}_m(p) := 1 - \left(1 - \frac{1}{p}\right)\left(1 - \frac{1}{p^2}\right)\cdots\left(1 - \frac{1}{p^m}\right)$. *There exist constants* $c_1, c_2 > 0$ *such that for all* ℓ, k, *and* $p < \ell^{\frac{1}{2}-\epsilon}$ *(i.e.,* $p \in \mathcal{P}_1 \cup \mathcal{P}_2$*),*

$$\left| \mathbb{P}(p \mid d_\ell^{(k)}) - [\mathbf{P}_m(p)]^k \right| < c_1 e^{-c_2 \ell^{2\epsilon}}.$$

Furthermore,

$$\mathbb{P}(\text{no } p \in \mathcal{P}_1 \text{ divides } d_\ell^{(k)}) = \prod_{\mathcal{P}_1} \left(1 - [\mathbf{P}_m(p)]^k\right) + c_1 e^{-c_2 \ell^{2\epsilon}}.$$

Proof. The number of nonsingular matrices with $\mathbb{F}_p$ entries is

$$\left|GL_m(\mathbb{F}_p)\right| = (p^m - 1)\,(p^m - p)\cdots\left(p^m - p^{m-1}\right)$$

out of p^{m^2} total matrices [4], so the ratio of singular matrices is $\mathbf{P}_m(p)$. Thus the lemma follows from the fact that each entry of M_ℓ approaches a uniform distribution with the error term decaying exponentially fast in ℓ. Summing the error over the km^2 entries appearing in k $m \times m$ matrices only worsens the constant c_2 that appeared in RNG Lemma 15.

For the second statement we use the (well-known) fact that the product of primes up to N is asymptotically e^N, so that $\prod_{\mathcal{P}_1} p < \ell^{\frac{1}{2}-\epsilon}$, which allows us to apply RNG Lemma 15 with $n = \prod_{\mathcal{P}_1} p$. The Chinese Remainder Theorem ensures that for any $m \times m$ matrices A with entries in $\mathbb{Z}/p\mathbb{Z}$ and B with entries in $\mathbb{Z}/q\mathbb{Z}$, there is a unique matrix C with entries in $\mathbb{Z}/pq\mathbb{Z}$ that agrees with both in the respective projections. Using this repeatedly, with $n = \prod_{\mathcal{P}_1} p$, we count that the number of matrices over $\mathbb{Z}/n\mathbb{Z}$ such that no $p \in \mathcal{P}_1$ divides the determinant must equal $\prod_{\mathcal{P}_1} |GL_m(\mathbb{F}_p)|$. The statement follows. □

To get a similar bound for large primes, we prove two lemmas on the divisibility of the determinants of the submatricies $M_\ell^{(k)}$, and then combine them for a bound that works on $\mathcal{P}_3$ and $\mathcal{P}_4$.

Lemma 5 (Divisibility of determinants by large primes)**.** *Fix* $\epsilon > 0$. *Then there is a constant* c *such that for sufficiently large* ℓ *and any prime* $p \geq \ell^{1/2-\epsilon}$ *(i.e.,* $p \in \mathcal{P}_3 \cup \mathcal{P}_4$*), we have*

$$\mathbb{P}'(\det M_\ell^{(k)} \equiv 0 \mod p) < \frac{c}{\ell^{\frac{1}{2}-2\epsilon}} + \frac{c}{\sqrt{\ell}} + \frac{1}{p},$$

where $\mathbb{P}'$ *denotes conditional probability given that the matrix entries are nonzero. It follows that there is a constant* c *such that* $\mathbb{P}'(p \mid \Delta_{\ell,i}) < c\ell^{2\epsilon-\frac{1}{2}}$ *for* $p \in \mathcal{P}_3 \cup \mathcal{P}_4$, *and* $\mathbb{P}'(p \mid \Delta_{\ell,i}i) < cp^{\frac{4\epsilon-1}{2m+2}}$ *for* $p \in \mathcal{P}_3$.

Proof. For fixed m, we induct on k. When $k = 1$, the statement follows from Lemma 1. For $k > 1$, we introduce the equivalence relation $A \sim B \iff a_{ij} = b_{ij}$ for all $(i,j) \neq (k,k)$; that is, we declare two $k \times k$ matrices equivalent if they agree in all entries except possibly the bottom right. Then there is a constant C_M for each matrix M such that

$$\det A = a_{kk} \det N + C_M \qquad \forall A \in [M],$$

where N is the upper-left-hand $(k-1) \times (k-1)$ minor. Now if $p \nmid \det N$, then solving for a_{kk} gives $(\det A - C_M)(\det N)^{-1} \bmod p$, so at most $1/p$ of the a_{kk} values in $\mathbb{Z}$ give a possible solution. Thus there are at most $2\ell^{1/2+\epsilon}/p + 1$ matrices $A \in [M]$ with determinant divisible by p in this case, and since $\ell^{\frac{1}{2}-\epsilon} \leq p$ this has a conditional probability at most $(\frac{2\ell^{1/2+\epsilon}}{p} + 1)\frac{c}{\sqrt{\ell}} < \frac{2c}{\ell^{1/2-2\epsilon}} + \frac{c}{\sqrt{\ell}}$ (given that the matrix falls in the equivalence class).

The induction hypothesis says that the probability that $\det N$ is divisible by p is $\frac{2(m-2)c}{\ell^{\frac{1}{2}-2\epsilon}} + \frac{(m-2)c}{\sqrt{\ell}} + \frac{1}{p}$. After enlarging c, the first statement follows by induction. Next, we want to combine the three terms on the right-hand side. Since $p > \ell^{\frac{1}{2}-\epsilon}$, we first observe that $\frac{1}{p} < \frac{1}{\ell^{1/2-\epsilon}} < \frac{1}{\ell^{1/2-2\epsilon}}$, and clearly $\frac{1}{\sqrt{\ell}} < \frac{1}{\ell^{1/2-2\epsilon}}$ as well.

Note that if $p \leq \ell^{m+1}$, then $\ell \geq p^{\frac{1}{m+1}}$, and we get the final statement. □

Lemma 6 (Nonsingularity)**.** $\overline{\mathbb{P}}(\Delta = 0) = 0$.

Proof. The idea is that determinant zero is a codimension-one condition. To show it rigorously, we prove the following stronger result: for fixed m, we will show that $\overline{\mathbb{P}}(\det M_\ell^{(k)} = 0) = 0$ for each $1 \leq k \leq m$. For $k = 1$, we note that $M_\ell^{(1)} = R_\ell$, so the statement follows from Lemma 2. Let's show that if it is true for $M_\ell^{(k-1)}$, then

it is true for $M = M_\ell^{(k)}$. Let us write q_ℓ to denote the lower-right entry of $M_\ell^{(k)}$ and μ_ℓ to denote the list of the other k^2-1 entries $(M_{1,1},\dots,M_{k,k-1})$. The induction hypothesis tells us the probability that $\det N = 0$ tends to zero for N the upper left-hand $k-1\times k-1$ minor. Assuming that minor is nonsingular, there is exactly one value of q_ℓ making $\det M_\ell^{(k)} = 0$ for each μ; call it $q(\mu)$. But, recalling that 0 is the most likely value for q_ℓ and that the different $\mu_\ell = \mu$ are disjoint events, we have

$$\mathbb{P}(\det M_\ell^{(k)} = 0) = \sum_\mu \mathbb{P}(q_\ell = q(\mu)) \le \sum_\mu \mathbb{P}(q_\ell = 0) = \mathbb{P}(q_\ell = 0).$$

But q_ℓ is distributed like R_ℓ, so by Lemma 2, this tends to zero. □

Proposition 7 (Common divisors of random determinants)**.** *Fixing m and any $k > 10m$, we have*

$$\overline{\mathbb{P}}(d_\ell^{(k)} = 1) = \prod_{\text{primes } p} 1 - \left[1 - \left(1-\frac{1}{p}\right)\left(1-\frac{1}{p^2}\right)\cdots\left(1-\frac{1}{p^m}\right)\right]^k.$$

Proof. We'll break down the probability by dividing the primes into the size ranges $\mathcal{P}_1$, $\mathcal{P}_2$, $\mathcal{P}_3$, and $\mathcal{P}_4$. As above, let $\mathbf{P}_m(p) := 1-\left(1-\frac{1}{p}\right)\left(1-\frac{1}{p^2}\right)\cdots\left(1-\frac{1}{p^m}\right)$, and note that $\mathbf{P}_m(p) \le \frac{2^m}{p}$ because there are at most 2^m nonzero terms with denominators at least p. We clearly have the following bounds:

$$\begin{aligned}\mathbb{P}'(p|d_\ell^{(k)} \text{ for some } p\in\mathcal{P}_1) \;<\; \mathbb{P}(d_\ell^{(k)} > 1) \;<\; &\mathbb{P}'(p|d_\ell^{(k)} \text{ for some } p\in\mathcal{P}_1) + \mathbb{P}'(p|d_\ell^{(k)} \text{ for some } p\in\mathcal{P}_2)\\ &+ \mathbb{P}'(p|d_\ell^{(k)} \text{ for some } p\in\mathcal{P}_3) + \mathbb{P}'(p|d_\ell^{(k)} \text{ for some } p\in\mathcal{P}_4)\\ &+ \mathbb{P}(\text{some entry is zero}).\end{aligned}$$

We apply Lemma 4 and take a limit to get

$$\mathbb{P}(p \mid d_\ell^{(k)} \text{ for some } p\in\mathcal{P}_1) = 1 - \prod_{\mathcal{P}_1}\left(1-[\mathbf{P}_m(p)]^k\right) + O(e^{-\ell^{2\epsilon}}) \longrightarrow 1 - \prod_{\text{primes } p}\left(1-[\mathbf{P}_m(p)]^k\right).$$

Note that $\mathbb{P}'$ conditions on an event whose probability tends to 1, thus $\lim_{\ell\to\infty}\mathbb{P}'(E) = \overline{\mathbb{P}}(E)$ if the limits exist. We have thus shown that $\overline{\mathbb{P}}(d_\ell^{(k)} > 1) \ge 1 - \prod_{\text{primes } p}\left(1-[\mathbf{P}_m(p)]^k\right)$ which implies that $\overline{\mathbb{P}}(d_\ell^{(k)} = 1) \le \prod_{\text{primes } p}\left(1-[\mathbf{P}_m(p)]^k\right)$.

To finish the theorem we must use the other four terms that bound $\mathbb{P}(d_\ell^{(k)} > 1)$ limit to zero, starting with the primes in $\mathcal{P}_2$. We have

$$\mathbb{P}(p \mid d_\ell^{(k)} \text{ for some } p\in\mathcal{P}_2) < \sum_{\mathcal{P}_2}\mathbb{P}(p|d_\ell^{(k)}) = \sum_{\mathcal{P}_2}\left(\mathbf{P}_m(p) + O(e^{-\ell^{2\epsilon}})\right)^k \longrightarrow 0,$$

where the $\mathbf{P}_m(p)$ term appears because $p < \ell^{\frac{1}{2}-\epsilon}$ means we can apply Lemma 4. To justify the convergence to zero, recall that $\mathbf{P}_m(p) \le \frac{2^m}{p}$ and $k \ge 2$.

We now compute the case of $\mathcal{P}_3$, applying Lemma 5 (and recalling that $k > 10m$ and ϵ is small) to get

$$\mathbb{P}'(p \mid d_\ell^{(k)} \text{ for some } p\in\mathcal{P}_3) \le \sum_{\mathcal{P}_3}\mathbb{P}'(p \mid d_\ell^{(k)}) = \sum_{\mathcal{P}_3}\left(\mathbb{P}'(p \mid \Delta_i)^k\right) \le \sum_{\mathcal{P}_3} c\cdot p^{\frac{4\epsilon-1}{2m+2}k} \le \sum_{\mathcal{P}_3}\frac{c}{p^2}.$$

Since the sum over all primes of p^{-2} converges, this certainly converges to zero as $\ell\to\infty$.

In the range $\mathcal{P}_4$, since all coordinates of the random walk vector are $\le \ell$, we have $|\Delta_\ell| \le m!\ell^m < \ell^{m+1}$ for $\ell \gg 1$. Since $\Delta_\ell = 0$ is an asymptotically negligible event (Lemma 6), we have $\mathbb{P}'(p|d_\ell^{(k)} \text{ for some } p) \longrightarrow 0$. Finally, the probability of a zero entry also goes to zero (Lemma 2), which completes the proof. □

If we take k to be a function of ℓ such that $k\to\infty$, then $\overline{\mathbb{P}}(d_\ell^{(k)} = 1) = 1$. This proves Theorem 36.

References

[1] Cordes, Duchin, Duong, Ho, Sánchez, *Random Nilpotent Groups I*, IMRN, to appear.

[2] R. Fitzner and R. van der Hofstad, *Non-backtracking random walk*. Journal of Statistical Physics 150(2): 264-284, (2013).

[3] Gregory F. Lawler and Vlada Limic, *Random walk: a modern introduction*. Cambridge Studies in Advanced Mathematics, 123. Cambridge University Press, Cambridge, 2010.

[4] D. J. S. Robinson, *A course in the theory of groups*, vol. 80 of Graduate Texts in Mathematics, Springer-Verlag, New York, second ed., 1996.

[5] L. Saloff-Coste, *Random walks on finite groups*. Probability on discrete structures, 263–346, Encyclopaedia Math. Sci., 110, Springer, Berlin, 2004.