Intersection Coupling for birthday attacks
and collision of independent walks on $\mathbb{Z}$

Ravi Montenegro *

Abstract

We develop a version of coupling which can be used to study many birthday attacks, simplifying a key step in past analysis and making it possible to extend the analysis to additional problems. This leads us to confirm a conjecture of Pollard’s that the Kangaroo method starting in an interval of size $N$ with steps sizes $\{x^{k-1}\}_{k=1}^d$ for fixed $x$ has the same collision time of $(2 + o(1))\sqrt{N}$ as does $x = 2$. Furthermore we partially answer a question he poses regarding the case when the number of generators $d$ is fixed and the base $x$ is variable, by showing that when $d = 3$ the collision time is order $O(\sqrt{N \ln N})$ and giving strong evidence that it is $\Theta(\sqrt{N \ln N})$, while for $d \geq 4$ the collision time is shown to be order $\Theta(\sqrt{N \ln N})$.

1 Introduction

Birthday attacks are a widely used method for breaking codes, and although they have been in widespread use for some time it is only recently that methods of rigorously analyzing them have appeared [1 2]. While these attacks appear in various guises, in general the principle is similar: run a random walk on some state space and proceed until it intersects itself, or run multiple walks in parallel until some two intersect each other [3 5].

Common random walks used include additive walks, multiplicative walks, and mixtures of both. Our coupling method is sufficiently general that it is likely to apply in all three cases, but we have found additive walks to be particularly amenable to analysis and so that is our focus here. An additive walk is of the following type: Given a set of jump sizes $S = \{s_k\}_{k=1}^d \subset \mathbb{N}$ and a probability distribution $p : S \rightarrow [0, 1]$ on $S$ so that average step size $\bar{S} = \sum_{s \in S} p(s) s$, consider the Markov chain on $\mathbb{Z}$ given by $P(x, y) = p(y - x)$ if $y - x \in S$ and 0 otherwise.

If intersection of two walks is being considered, starting from values (or distributions) $X_0$ and $Y_0$, then the analysis can be broken into two stages. First, the rear walk runs until it catches up with the lead one. This is elementary and requires $|X_0 - Y_0|/\bar{S}$ steps. In the common situation when $X_0 \in_{uar} [a, b]$ is chosen uniformly at random from an interval then the first stage time is minimized when $Y_0 = \frac{a + b}{2}$, so that $\mathbb{E}\frac{|Y_0 - X_0|}{\bar{S}} = \frac{b - a}{4\bar{S}}$ steps are required. The difficult stage in the analysis is in determining the time until intersection of two walks with $|X_0 - Y_0| \leq \frac{1}{2} S_{\text{max}}$ where $S_{\text{max}} = \max_{s \in S} s$.

Following are several common additive walks and what is known about their intersection time. Unless stated otherwise these refer to the case of intersection of two walks, with $X_0 \in_{uar} [a, b]$, $Y_0 = \frac{a + b}{2}$, $d$ chosen so that $\bar{S} \approx \frac{b - a}{2}$ and $N = b - a$.

- $S = \{2^{k-1}\}_{k=1}^d$: Proven that $(2 + o(1))\sqrt{b - a}$ steps are required on average [2].

*Department of Mathematical Sciences, University of Massachusetts Lowell, Lowell, MA 01854, USA. Email: ravi_montenegro@uml.edu; this work was done while the author was at The Tokyo Institute of Technology.
• $S = \{x^{k-1}\}^d_{k=1}$ for $x > 2$ fixed: Pollard conjectures that $(2 + o(1))\sqrt{N}$ steps are required on average [4]. This is proven to be correct in the journal version of [2].

• $S = \{x^2\}^d_{k=1}$: Pollard proved that $\Theta(N^{2/3})$ steps are required with $\bar{S} = N^{1/3}$ [4].

• $S = \{x^d\}^d_{k=1}$ for $d > 2$ fixed: Pollard proposes that this be examined but does not conjecture as to the correct result [4].

• $S = \{s_k^d\}^d_{k=1}$ for $d$ fixed ($d = 20$ is common) with $s_k \in \text{var } Z_N$, considering self-intersection of a walk: Teske gives a convincing but incomplete argument that when $d \geq 5$ then $\Theta(\sqrt{N})$ steps are required for the walk to intersect itself. It is conjectured that this holds when $d = 4$ as well, and is observed that experimental data suggests it does not hold when $d = 3$, but no conjecture as to the correct rate is given. [5]

We can apply our new coupling method to find solutions to all but the last of these problems, thus providing a common perspective from which to understand the full class of additive walks with generating set of the form $S = \{x^d\}^d_{k=1}$. In particular, when there are three generators, i.e. $S = \{x^{k-1}\}^3_{k=1}$, we prove that it takes $O(\sqrt{N} \ln N)$ steps until intersection, while only a minor improvement is required to match the $O(\sqrt{N} \ln N)$ order that experimental data suggests is correct, and with $d \geq 4$ generators we show that $\Theta(\sqrt{N})$ steps suffice. We expect our methods extend to Teske’s problem as well and intend to pursue this in the full version of the paper. From simulation data it appears that the same bounds will hold as for Pollard’s Kangaroo case with $d$ generators, namely $O(\sqrt{N} \ln N)$ when $d = 3$ generators are considered and $\Theta(\sqrt{N})$ when 4 or more generators are used.

One practical consequence of our work is in understanding which features of a birthday attack effect its performance. Heuristically, when two walks start nearby then there is initially a warm-up of some $T$ steps until they become sufficiently randomized to take on their asymptotic behavior, so that each subsequent step has an $\bar{S}^{-1}$ chance of hitting a site the other walk also visits. The expected number of collisions in the next $M$ steps is then $M/\bar{S}$, but once the walks have reached a common state then they are much more likely to intersect before they are re-randomized in $T$ steps. If there are some $B$ collisions during this period then two walks intersecting contribute about $(1 + B)$ intersections to the expected number of collisions, so that the probability of two walks colliding in $N$ steps will be roughly $\frac{M}{\bar{S}(1+B)}$, and so around $\frac{b-a}{4\bar{S}} + T + \bar{S}(1+B)$ steps are required in total.

Our new coupling method simplifies analysis of the warm-up period, so that it remains only to determine the number of collisions after two walks intersect. Contrary to what some have suspected, we find that mixing time is not the primary determinant of efficiency, but rather the rate at which peak probabilities decay from an average point, as this minimizes the number of consecutive intersections. This suggests that a walk which is locally tree-like will converge better than one which is not. Indeed, Lemma 4.3 of [1] shows that for the degree-3 Pollard Rho algorithm probabilities decay exponentially like $P^t(u, v) \lesssim (2/3)^{t/3}$, which explains why it requires only $O(\sqrt{N})$ steps on a group of order $N$. In contrast, in this paper we find that the 3-generator Kangaroo walk has a fairly slow probability decay of $\Theta(1/t)$ and requires a larger $O(\sqrt{N} \log N)$ steps until a collision. At degree 2 this is even more marked, with a degree 2 Rho type walk maintaining $O(\sqrt{N})$ time until collision, while the 2-generator Kangaroo method is seen to have probability decay of only $\Theta(1/\sqrt{t})$ and requires a much larger $\Omega(N^{2/3})$ steps.

This paper is organized as follows. In Section 2 we introduce tools developed in past papers to study Birthday attacks. This is followed in Section 3 by our main innovation, a coupling method for intersection problems. It is applied to general additive walks in Section 4 and then combined
with an additional quantity in Section 5 to give rigorous results for the Birthday attacks mentioned in this introduction.

2 Preliminaries

As mentioned earlier, many of the methods developed in our earlier analysis of Birthday attacks are highly specific, but a few of the tools have broad application. We introduce here results from the journal version of [2], which are similar to those in the conference version but extended to the more general setting of stopping times.

Definition 2.1. A stopping time for a random walk \( \{X_i\}_{i=0}^{\infty} \) is a random variable \( T \in \mathbb{N} \) such that the event \( \{T = t\} \) depends only on \( X_0, X_1, \ldots, X_t \). The average time until stopping is \( T = \mathbb{E}T \).

Definition 2.2. Consider a Markov chain \( P \) on an infinite group \( G \). A nearly uniform intersection time \( T(\epsilon) \) is a stopping time such that for some \( U > 0 \) and \( \epsilon \geq 0 \) the relation

\[
(1 - \epsilon)U \leq \Pr(\exists j : X_{T(\epsilon) + \Delta} = Y_j) \leq (1 + \epsilon)U
\]

holds for every \( \Delta \geq 0 \) and every \( (X_0, Y_0) \) in a designated set of initial states \( \Omega \subset G \times G \).

In general the intersection probability may go to zero in the limit. However, if a walk is transitive on \( \mathbb{Z} \) (i.e. \( P(u, v) = P(0, v - u) \)), increasing (i.e. \( P(u, v) > 0 \) only when \( v > u \)), and aperiodic (i.e. \( \gcd\{k : P(0, k) > 0\} = 1 \)), as are additive walks, then one out of every \( S = \sum_{k=1}^{\infty} kP(0, k) = \sum_{v < 0} P(v, Z \geq 0) \) states is visited and a stopping time will exist satisfying

\[
\frac{1 - \epsilon}{S} \leq \Pr(\exists j : X_{T(\epsilon) + \Delta} = Y_j) \leq \frac{1 + \epsilon}{S}.
\]

An obvious choice of starting states \( \Omega \) are all \( Y_0 \leq X_0 \), but for reasons that will be apparent later it better serves our purposes to expand to the case of \( Y_0 < X_0 + S_{\max} \), where \( S_{\max} = \max_{s \in S} s \) is the largest step size. By transitivity and since no intersection can occur until the first time \( Y_j \geq X_0 \) then it actually suffices to verify the nearly uniform intersection time for the case \( X_0 = 0 \leq Y_0 < S_{\max} \).

The main theorem of [2] requires the quantity \( B_\epsilon \), the worst-case expected number of collisions between two independent walks before the nearly uniform intersection time \( T(\epsilon) \). To be precise:

\[
B_\epsilon = \max_{Y_0 < X_0 + S_{\max}} \mathbb{E} \left[ \sum_{i=1}^{T(\epsilon)} 1_{\{\exists j : X_i = Y_j\}} \right].
\]

The main result of [2] which we will use is a bound on the expected number of steps until a collision.

Theorem 2.3. Given an increasing transitive Markov chain on \( \mathbb{Z} \), if two independent walks have starting states with \( Y_0 < X_0 + S_{\max} \) then

\[
\mathbb{E} \min\{i > 0 : \exists j, X_i = Y_j\} \leq 1 + \left( \frac{\sqrt{S(1 + B_\epsilon)} + \sqrt{T(\epsilon)}}{1 - \epsilon} \right)^2
\]

\[
\mathbb{E} \min\{i > 0 : \exists j, X_i = Y_j\} \geq 1 + S \left( \max\{0, 1 - \sqrt{B_\epsilon}\} \right)^2
\]

In particular, when \( \epsilon \) is close to zero then

\[
\mathbb{E} \min\{i > 0 : \exists j, X_i = Y_j\} \preceq T(\epsilon) + S(1 + B_\epsilon) + 2 \sqrt{T(\epsilon)S(1 + B_\epsilon)},
\]

which makes rigorous the heuristic given in the introduction.
3 Coupling to bound nearly uniform intersection time

As established in the preliminaries, the key steps required to bound collision time are to show existence of a nearly uniform intersection time and to show that one intersection is unlikely to be followed by many others. We prove here the main result of this paper, a coupling theorem for nearly uniform intersection time.

**Definition 3.1.** A coupling of Markov Chains with transition matrix $P$ is a random process $(X_t, Y_t)_{t=0}^\infty$ such that $X_t$ and $Y_t$ are Markov Chains with transition matrix $P$, although possibly with different initial distributions.

Any coupling can be modified so that once the two chains meet then they subsequently stay together, by merely making the same transitions on both after the first time $X_t = Y_t$. This property will be assumed in the remainder; we say the walks are coupled once $X_t = Y_t$.

In the normal application of coupling the goal is to show that after some number of steps the two distributions are nearly the same. However, when studying intersection of two walks then having the same distribution is of little use, as two independent copies of the same walk started from the same state might never intersect other than at their initial state, despite having identical probability distributions at all times. Instead, the goal here is to show that after some point in time or space every later state has nearly the same probability of being hit by both walks, not necessarily at the same time. As such, we can permit more flexible couplings in which occasionally one of the walks does not make a transition, regardless of whether $P$ allows for this, and this decision may even depend on future transition of the other walk.

**Definition 3.2.** An intersection coupling of Markov Chains with transition matrix $P$ is a random process $(X_t, Y_t)_{t=0}^\infty$ such that when $X_{t+1} \neq X_t$ then $X_{t+1}$ is distributed as $\frac{1}{1-P(X_t, X_t)} P(X_t, \cdot)$, and likewise when $Y_{t+1} \neq Y_t$ although possibly with a different initial distribution.

In this more general form $X_t$ is not necessarily the $t$th state visited, but rather the $\tau(t)$th state for some randomized stopping time $\tau(t) \leq t$. Likewise for $Y_t$. As such, an intersection coupling at some time $t$ is similar to coupled stopping times.

Coupling has proven to be extremely useful in studying convergence of finite Markov Chains. We add here another application, using coupling to show a nearly uniform intersection time and thereby bound the expected time until two independent random walks intersect. The approach taken will be to couple the $Y_j$ walk to a random process which has an $\tilde{S}^{-1}$ chance of hitting any specified state, so that after the $Y_j$ walk has coupled with this process then it too will have this property. But then every subsequent state the $X_i$ visits has an $\tilde{S}^{-1}$ chance of being visited by the $Y_j$ walk as well, i.e. a nearly-uniform intersection time.

First, a coupling theorem:

**Theorem 3.3.** Consider an increasing Markov Chain $P$ on $\mathbb{Z}$. Given initial states $Y_0$ and $Z_0$ and an intersection coupling $(Y_t, Z_t)$ then

$$\forall z \in \mathbb{Z} : \left| \Pr(z \in \{Y_j\}) - \Pr(z \in \{Z_k\}) \right| \leq \Pr(\mathcal{L} > z)$$

where $\mathcal{L}$ denotes the location (not time) at which the two walks couple.

**Proof.** Since $Y$ and $Z$ follow the same trajectories after the two walks have coupled then

$$\Pr(z \in \{Z_k\}) = \Pr(\mathcal{L} \leq z) \Pr(z \in \{Z_k\} \mid \mathcal{L} \leq z) + \Pr(\mathcal{L} > z) \Pr(z \in \{Z_k\} \mid \mathcal{L} > z)$$

$$= \Pr(\mathcal{L} \leq z) \Pr(z \in \{Y_j\} \mid \mathcal{L} \leq z) + \Pr(\mathcal{L} > z) \Pr(z \in \{Z_k\} \mid \mathcal{L} > z)$$

$$= \Pr(z \in \{Y_j\})$$

$$\forall z \in \mathbb{Z} : \left| \Pr(z \in \{Y_j\}) - \Pr(z \in \{Z_k\}) \right| \leq \Pr(\mathcal{L} > z)$$
Likewise,
\[
\Pr(z \in \{Y_j\}) = \Pr(\mathcal{L} \leq z) \Pr(z \in \{Y_j\} \mid \mathcal{L} \leq z) + \Pr(\mathcal{L} > z) \Pr(z \in \{Y_j\} \mid \mathcal{L} > z)
\]
And so
\[
|\Pr(z \in \{Z_k\}) - \Pr(z \in \{Y_j\})| = \Pr(\mathcal{L} > z) |\Pr(z \in \{Z_k\} \mid \mathcal{L} > z) - \Pr(z \in \{Y_j\} \mid \mathcal{L} > z)| \\
\leq \Pr(\mathcal{L} > z)
\]

\[\square\]

To show a near uniform intersection time it is necessary to couple with a walk which has a uniform probability $1/\bar{S}$ of hitting any state larger than $Y_0$. It turns out that finding an appropriate distribution for $Z_0$ is a non-issue for the most part:

**Corollary 3.4.** Suppose a Markov kernel $P$ is increasing on $Z$ and doubly-stochastic, i.e. $\forall y : \sum_x P(x, y) = 1$. Fix some $w_0$ and let $W = \{w \geq w_0 : \exists v < w_0, P(v, w) > 0\}$ and $\bar{S} := \sum_{v < w_0} P(v, W)$. Given walk $Y$ with transition kernel $P$, for each state $w \in W$ define an intersection coupling between $Y$ and a walk $\{Z_k\}$ with $Z_0 = w$. Then
\[
\forall z \in Z : |\Pr(z \in \{Y_j\}) - \bar{S}^{-1}| \leq \max_{w \in W} \Pr(\mathcal{L} > z \mid Z_0 = w)
\]
where $\mathcal{L}$ denotes the location (not time) at which the two walks couple.

In particular, when $w_0 \approx Y_0$ then it suffices to couple $Y$ with walks starting nearby, and whichever of these is least likely to couple by the time location $z$ is reached will determine the strength of the bound.

**Proof.** The corollary is trivial when $z < w_0$ since $\bar{S} \geq P(w_0 - 1, W) = 1$ for an increasing walk, so assume $z \geq w_0$.

Consider the initial distribution
\[
\Pr(Z_0 = w) = \frac{\sum_{v < w_0} P(v, w)}{\bar{S}} 1_{\{w \geq w_0\}}
\]
If each state $< w_0$ is visited with probability $1/|\bar{S}|$ then this is the probability that $w$ is the first state visited $\geq w_0$. We show that it and future states will also be hit with uniform probability $1/\bar{S}$.

Observe that
\[
\Pr(z \in \{Z_k\}) = \sum_{w \in W} \Pr(Z_0 = w) \Pr(z \in \{Z_k\} \mid Z_0 = w)
\]
and so by Theorem 3.3
\[
|\Pr(z \in \{Y_j\}) - \Pr(z \in \{Z_k\})| = \left| \sum_{w \in W} \Pr(Z_0 = w) (\Pr(z \in \{Y_j\}) - \Pr(z \in \{Z_k\} \mid Z_0 = w)) \right| \\
\leq \sum_{w \in W} \Pr(Z_0 = w) |\Pr(z \in \{Y_j\}) - \Pr(z \in \{Z_k\} \mid Z_0 = w)| \\
\leq \max_{w \in W} |\Pr(z \in \{Y_j\}) - \Pr(z \in \{Z_k\} \mid Z_0 = w)| \\
\leq \max_{w \in W} \Pr(\mathcal{L} > z \mid Z_0 = w)
\]
If $z = w_0$ the result follows because $\Pr(w_0 \in \{Z_k\}) = \Pr(Z_0 = w_0) = \bar{S}^{-1}$ for an increasing doubly-stochastic walk.

If $z > w_0$ then inductively assume that $\Pr(v \in \{Z_k\}) = \bar{S}^{-1}$ for all $v \in \{w_0, z\}$. Then

$$\Pr(z \in \{Z_k\}) = \Pr(Z_0 = z) + \sum_{v < z} \Pr(v \in \{Z_k\}) \Pr(v, z) = \frac{\sum_{v < w_0} \Pr(v, z)}{\bar{S}} + \frac{\sum_{w_0 \leq v < z} \Pr(v, z)}{\bar{S}} = \frac{1}{\bar{S}}$$

One issue in applying this is that couplings usually involve time, not location. If $T$ denotes the time at which the walks couple, so that $Y_T = L$, then

$$\left| \Pr(z \in \{Y_j\}) - \frac{1}{\bar{S}} \right| \leq \max_{w \in W} \Pr(Y_T > z \mid Z_0 = w)$$

Bounding the probability can be simplified by boosting any non-trivial bound on $\Pr(T > T)$ into an arbitrarily strong result. To do this, let $\Omega$ denote the set of all possible $(Y_i, Z_i)$ that could be visited during a coupling, for instance the coupling might start with states satisfying $\lvert Y_0 - Z_0 \rvert \leq \frac{1}{2}S_{\text{max}}$ and maintain the relation $\lvert Y_t - Z_t \rvert \leq M$ at all times, so $\Omega \subseteq \{(y, z) : \lvert y - z \rvert \leq M\}$. Let $\Pr(T > T) := \max_{(Y_0, Z_0) \in \Omega} \Pr(T > T \mid (Y_0, Z_0))$. If the walks haven’t coupled by $T$ then take $(Y_0, Z_0) \leftarrow (Y_T, Z_T)$ and repeat, obtaining $\Pr(T > 2T \mid T > T) \leq \Pr(T > T)$, and after $k$ rounds $\Pr(T > kT) \leq \Pr(T > T)^k$. So if we define $T(\epsilon) := \min\{T : \Pr(T > T) \leq \epsilon\}$ then for any choice of $T$

$$T(\epsilon) \leq T \log_{1/\Pr(T > T)} \frac{1}{\epsilon} \leq \frac{T}{1 - \Pr(T > T)} \ln \frac{1}{\epsilon} \leq e \mathbb{E}T \ln \frac{1}{\epsilon} \leq \frac{e \mathbb{E}T}{1 - \Pr(T > T)}$$

(2)

A nearly-uniform intersection time can be obtained from this by defining

$$T(\epsilon) = \min \left\{ i : \Pr(X_i < Y_{T(\epsilon/2\bar{S})}) \leq \epsilon/2\bar{S} \right\}$$

(3)

and observing that

$$\Pr(Y_T > X_{T(\epsilon)}) \leq \Pr(Y_T > Y_{T(\epsilon/2\bar{S})}) + \Pr(Y_{T(\epsilon/2\bar{S})} > X_{T(\epsilon)}) \leq \epsilon/\bar{S}$$

The quantity $T(\epsilon)$ can typically be bounded by using the Hoeffding Inequality or a Chernoff Bound to upper bound $Y_{T(\epsilon/2\bar{S})}$ in terms of $T(\epsilon/2\bar{S})$ and lower bound the $X_i$ walk.

Since $T(\epsilon)$ provides no useful information about nearly uniform intersection times unless $\epsilon < \bar{S}^{-1}$, then the average value $\mathbb{T}$ also provides no useful information about the average $T(\epsilon)$ which appears in Theorem 2.3. However, a somewhat stronger method can provide the desired relation:

**Definition 3.5.** A stopping time $T$ is a strong intersection time if

$$\forall z, \Pr(z \in \{Y_t\} \mid z \geq Y_T) = \bar{S}^{-1}.$$
In other words, after the strong intersection time the walk will thereafter hit every vertex with uniform probability, whereas in the coupling case such a statement can only be made if \( \Pr(Y_T > z) = 0 \).

**Theorem 3.6.** Suppose a Markov kernel \( P \) is increasing on \( \mathbb{Z} \). Given a strong intersection time \( T \) for a walk \( Y \) then \( T = \min\{i : X_i \geq Y_T\} \) is a uniform intersection time, while \( T'(\epsilon) = \min\{T, M\} \) is a bounded nearly uniform intersection time for any \( \epsilon = S\Pr(T > M) = S\Pr(Y_T > X_M) \) and \( ET(\epsilon) \leq ET \).

**Proof.** First consider the uniform intersection time. At time \( T + \Delta \) then \( X_{T+\Delta} \geq X_T \geq Y_T \) and so by definition of strong intersection time \( \Pr(X_{T+\Delta} \in \{Y_i\}) = \bar{S}^{-1} \).

Next, at time \( T(\epsilon) + \Delta \) then \( \Pr(Y_T > X_{T(\epsilon)+\Delta}) \leq \Pr(Y_T > X_{T(\epsilon)}) \leq \epsilon S \) and so

\[
\Pr(X_{T+\Delta} \in \{Y_i\}) = \Pr(X_{T(\epsilon)+\Delta} < Y_T) \Pr(X_{T+\Delta} \in \{Y_i\} \mid X_{T(\epsilon)+\Delta} < Y_T) + \Pr(X_{T(\epsilon)+\Delta} \geq Y_T) \Pr(X_{T+\Delta} \in \{Y_i\} \mid X_{T(\epsilon)+\Delta} \geq Y_T) = \Pr(X_{T(\epsilon)+\Delta} < Y_T) \Pr(X_{T+\Delta} \in \{Y_i\} \mid X_{T(\epsilon)+\Delta} < Y_T) + (1 - \Pr(X_{T(\epsilon)+\Delta} < Y_T)) \bar{S}^{-1}
\]

and so

\[
|\Pr(X_{T+\Delta} \in \{Y_i\}) - \bar{S}^{-1}| = \Pr(X_{T(\epsilon)+\Delta} < Y_T) |\Pr(X_{T+\Delta} \in \{Y_i\} \mid X_{T(\epsilon)+\Delta} < Y_T) - \bar{S}^{-1}|
\leq \frac{\epsilon}{\bar{S}} \max\{1 - \bar{S}^{-1}, \bar{S}^{-1}\}
\]

\( \square \)

Unfortunately a strong intersection time is harder to construct than an intersection coupling, somewhat analogous to the case with strong stationary times and coupling. One successful example can be found in the journal version of [2]. The approach taken there was to split the walk into two independent components, show one of them could be used to generate a sample from some distribution \( \sigma \) over an interval, and then rejection sampling can pull out distribution \( \Pi \) from \( \sigma \). This highly specialized construction seems unlikely to extend to walks not considered in that paper, in contrast to our new coupling method which we easily apply to a large class of walks in the following section.

## 4 Coupling additive walks

It is now time to apply the coupling theorem. The versatility of the technique will be apparent in that a fairly simple coupling will suffice to handle all additive walks.

The approach taken is a common one for couplings: Define a discrete distance metric and use this to measure whether the two walks are getting closer or not. In this case we choose the shortest-path metric:

\[
dist(x, y) = \min \left\{ \sum_{s \in S} |n_s| : y - x = \sum_{s \in S} n_s s, n_s \in \mathbb{Z} \right\}
\]

**Theorem 4.1.** Consider an additive walk with step sizes \( S = \{s_k\}_{k=1}^d \subset \mathbb{N}^d \) and transition probabilities \( \frac{\gamma}{S} \geq p(s) \geq \frac{\gamma-1}{S} \). Define \( \tilde{S} = \sum_{s \in S} p(s)s \), \( S_{\max} = \max_k s_k \) and \( L_{\text{half}} = \max_{0 < z \leq S_{\max}/2} \text{dist}(0, z) \). Then

\[
T(1/xd) = 6\epsilon \gamma d L_{\text{half}} (3 + L_{\text{half}}) \frac{S_{\max}}{S} \ln(2xd\bar{S})
\]

7
is a bounded nearly uniform intersection time.

Proof. By Corollary 3.8 with $w_0 = Y_0 - \lfloor \frac{S_{\max}}{2} \rfloor$ it suffices to couple with states satisfying $|Z_0 - Y_0| \leq \frac{1}{2} S_{\max}$, and so in particular $\text{dist}(Y_0, Z_0) \leq \Lhalf$. Our coupling will have $|Y_t - Z_t| \leq \Lhalf S_{\max}$ always, and so to use (3) and (1) requires a weaker initial assumption, that $|Y_0 - Z_0| \leq \Lhalf S_{\max}$.

To prevent the distance from getting too large the coupling will have two types of transitions, those from “normal” states with $\text{dist}(Y_t, Z_t) < \Lhalf$ and those from “abnormal” states with $\text{dist}(Y_t, Z_t) \geq \Lhalf$.

If $(Y_t, Z_t)$ is a normal state then choose a generator $s \in S$ with distribution $p$. If $s$ appears in a shortest path of $\text{dist}(Y_t, Z_t)$ then 50% of the time set $Y_{t+1} = Y_t + s$ and $Z_{t+1} = Z_t$, while 50% of the time set $Y_{t+1} = Y_t$ and $Z_{t+1} = Z_t + s$, so distance changes by $-1$ half the time and by $-1$, $0$ or $+1$ the other half the time. Otherwise $s$ does not appear in a shortest path and 50% of the time set $Y_{t+1} = Y_t + s$ and $Z_{t+1} = Z_t + s$, while 50% of the time remain with $Y_{t+1} = Y_t$ and $Z_{t+1} = Z_t$, so that distance does not change.

At $(Y_0, Z_0)$ or if $(Y_t, Z_t)$ is an abnormal state then repeatedly transition only the rear walk – $Y_t$ if $Y_t < Z_t$ and $Z_t$ otherwise – until eventually $|Z_t - Y_t| \leq S_{\max}/2$, and hence also $\text{dist}(Y_t, Z_t) \leq \Lhalf$. Then proceed with the transition types for normal states.

It remains to analyze the intersection coupling. If $(Y_t, Z_t)$ is a normal state with $Y_t \neq Z_t$ then with probability at least $\frac{1}{\delta}$ a generator in a shortest path was chosen, and when this happens

\[ \Pr(\text{dist}(Y_{t+1}, Z_{t+1}) = \text{dist}(Y_t, Z_t) - 1) \geq \frac{1}{2} \]
\[ \Pr(\text{dist}(Y_{t+1}, Z_{t+1}) = \text{dist}(Y_t, Z_t) + 1) \leq \frac{1}{2} \]

Also, when distance is increased to $2\Lhalf$ then an abnormal state is reached and after some sequence of steps the walk is returned to a new state with distance $\leq \Lhalf$.

The coupling thus induces a random walk on the line $[0, 1, \ldots, 2\Lhalf]$, with $2\Lhalf$ denoting the first abnormal state, and when 0 is reached then the walks have coupled. Assume the first normal state is $\Lhalf$, no bias, and transitions of $\pm 1$, noting that the previous paragraph establishes that any variation from this only speeds the walk towards 0. The walk on a line with $E[\Delta \text{dist}] \leq 0$ appears frequently in the analysis of couplings and when started at state $k$ on a line of length $L$ it requires an average of $\bar{\tau} \leq \frac{k(2L-k)}{E[(\Delta \text{dist})^2]}$ steps to reach 0, so $\bar{\tau} \leq 3\gamma dL_{\text{half}}^2$ for this problem.

And now to the abnormal states. The walk just discussed starts at a state $\leq \Lhalf$ and is biased towards 0, so it has probability $\geq 1/2$ of reaching 0 before $2\Lhalf$. If $2\Lhalf$ is reached then this is an abnormal state and the walk is returned (after some steps) to a normal state with $\text{dist}(Y_t, Z_t) \leq \Lhalf$ once again. It follows that on average an abnormal state will be reached only $\leq \frac{1}{1/2} - 1 = 1$ time, while the sequence initiated at time 0 also behaves like another abnormal state.

The initial abnormal state has $|Z_t - Y_t| \leq 2\Lhalf S_{\max}$ and so $2\Lhalf$ steps of size $S_{\max}$ will suffice to return the coupling to a normal state with $|Z_t - Y_t| \leq S_{\max}/2$. This requires an average of $\leq 2\gamma d\Lhalf$ steps.

The coupling described here required an average of $\bar{\tau} = \mathbb{E} \tau \leq 3\gamma dL_{\text{half}}^2 + 2 \times 2\gamma dL_{\text{half}}$ steps. By equation (2) then $\tau(\epsilon) < 3\epsilon r_d L_{\text{half}}(2 + L_{\text{half}})$ steps. By equation (2), Lemma 1.2 below, and the relation $\bar{S} \geq p(S_{\max})S_{\max} \geq \frac{\gamma^3}{\delta} S_{\max}$ finish the proof. \(\square\)

The following simple application of Hoeffding’s Inequality was used above.

**Lemma 4.2.** Suppose a non-negative random variable has average $\bar{S}$ and maximum $S_{\max}$. If $N$ is a constant and $\delta_1, \delta_2, \ldots, \delta_M$ are some $M$ independent samples then the sum $X = \sum_{i=1}^{M} \delta_i$ satisfies

\[ \Pr(X < (1 + N)S_{\max}) \leq \epsilon \]

8
when
\[ M = 2 ^ {S_{\text{max}}} \max \left\{ \frac{S_{\text{max}}}{S} \ln(1/\epsilon), 1 + N \right\} \]

Proof. By Hoeffding’s Inequality
\[ \Pr \left( X - \mathbb{E}X < -\frac{M}{2} \tilde{S} \right) \leq \exp \left( \frac{-M^2 \tilde{S}^2}{2M S_{\text{max}}^2} \right) \]

Plugging in \( M \) from the Lemma finishes the proof. \( \square \)

5 Collision times for specific additive walks

As discussed in Introduction, and made concrete by Theorem 2.3, it remains only to upper bound \( B_\epsilon \), the number of intersections in \( T(\epsilon) \) steps between two walks started at the same state. This requires some specialization but generally revolves around upper bounding the peak probability after \( t \) steps, i.e. \( \max_{u,v} P_t(u,v) \).

Lemma 5.1. If an additive walk has generating set of the form \( S = \{x^{k-1}\}_{k=1}^d \) with \( \gamma \geq 1 \) fixed then the nearly uniform intersection time of Theorem 4.1 satisfies \( T(1/xd) = O(d^5 x^2 \log x) \).

Furthermore,
\[ B_\epsilon = \begin{cases} O(x \log x) & \text{if } d = 2 \text{ and } x \text{ varies} \\ O(\log^2 x) & \text{if } d = 3 \text{ and } x \text{ varies} \\ O(1) & \text{if } d \geq 4 \text{ and } x \text{ varies} \\ o(1) & \text{if } d \text{ varies and } x \text{ fixed} \end{cases} \]

Sketch of Proof. First apply Theorem 4.1 to determine details of the nearly uniform intersection time. The maximum step size is \( S_{\text{max}} = x^{d-1} \) with average step
\[ \bar{S} = \sum_{s \in S} p(s) s \leq \sum_{k=1}^d \frac{\gamma}{d} x^{k-1} \leq \frac{2\gamma x^{d-1}}{d} \]
\[ \tilde{S} = \sum_{s \in S} p(s) s \geq p(x^{k-1}) x^{k-1} \geq \frac{\gamma^{-1} x^{d-1}}{d} = \frac{S_{\text{max}}}{\gamma d} \]

If \( |Z_0 - Y_0| \leq \frac{S_{\text{max}}}{2} = \frac{1}{2} x^{d-1} \) then \( L_{\text{half}} \leq \frac{(d-1)(x+1)}{2} \) because there are \( (d-1) \) terms to consider in the base-\( x \) expansion and each requires at most \( \frac{x+1}{2} \) generators to write (e.g. a units digit of \( \frac{x+2}{2} = 1 \times x - \frac{x-2}{2} \times 1 \)). Combining these quantities, by Theorem 4.1 the nearly uniform intersection time has
\[ T(1/xd) \leq 6e\gamma^2 d^4 (x+1)^2 \ln(4\gamma x^d) = O(d^5 x^2 \ln x) \]

It remains to bound \( B_\epsilon \).

The case when \( x \) is constant was considered in [2], so the details are not included here.

When \( d \) is fixed instead, observe that when \( X_0 \neq Y_0 \) then at least from the first collision \( X_i = Y_j \) the walk will subsequently behave exactly as in the case when \( X_0 = Y_0 \), so \( B_\epsilon \) is at worst 1 more than the value when \( X_0 = Y_0 \). Under the assumption that \( X_0 = Y_0 \), suppose that for each \( w \) that...
\(P^t(0, w)\) has non-trivial probability for at most a single power of \(t\). Then
\[
\mathbb{E} \sum_{i=1}^{T(c)} 1_{\{X_i \in \{Y_i\}\}} \approx \mathbb{E} \sum_{i=1}^{T(c)} \Pr(X_i = Y_i) = \sum_{i=1}^{T(c)} \Pr(X_i = Y_i) = \sum_{i=1}^{T(c)} P^t(0, w)^2
\]
\[
\leq \mathbb{E} \sum_{i=1}^{T(c)} \max_w P^t(0, w) \sum_w P^t(0, w) = \mathbb{E} \sum_{i=1}^{T(c)} \max_w P^t(0, w)
\]

It remains to show that for each \(w\) that \(P^t(0, w)\) has non-trivial probability only for a single power of \(t\), and then upper bound the maximum \(\max_w P^t(0, w)\). As a start, observe that each integer can be written in at most one way as the sum of fewer than \(x\) values in \(S = \{x^k-1\}_{k=1}^d\), up to rearranging the order in which generators are added, namely by taking the base-\(x\) decomposition. For instance, if \(x > 3\) then \(v - u = 2x + 3 = x + x + 1 + 1 + 1\).

With 2 generators, for a given exponent \(t\) the most likely outcome is to have \(t/2\) generators of each type, so that by Stirling’s formula
\[
P^t(u, v) \leq \binom{t}{t/2} \left(\frac{1}{2}\right) = \frac{(t/e)^t \sqrt{2\pi t} e^{o(1)}}{(t/2e)^{t/2} \sqrt{2\pi t/2} (t/2e)^{t/2} \sqrt{2\pi t/2} 2^t} = \sqrt{\frac{2}{\pi} \frac{1 + o(1)}{t}}
\]

More generally, the Chernoff Bound can be used to show that for fixed \(u, v\) there is at most one value of \(t \ll x^2\) such that \(P^t(u, v)\) is not exponentially small, while for larger values of \(t\) there are at most \(O(\sqrt[4]{x})\) such values of \(t\). Due to space constraints the details are left to the full version.

With 3 generators the most likely outcome has \(t/3\) generators of each type, so that
\[
P^t(u, v) \leq \binom{t}{t/3, t/3, t/3} \frac{1}{3^t} = \frac{3\sqrt{3}}{2\pi} \frac{1 + o(1)}{t}
\]

In this case the Chernoff Bound shows that for fixed \(u, v\) there is at most one value of \(t \ll \frac{x^2}{\log x}\) such that \(P^t(u, v)\) is not exponentially small, while for larger values of \(t\) there are at most \(O\left(\frac{x}{\sqrt{\log x}}\right)^{(d-1)/2}\) such values of \(t\).

More generally, for a fixed number \(d\) of generators
\[
P^t(u, v) = O\left(\frac{1}{t^{(d-1)/2}}\right)
\]

The Chernoff Bound for \(d \geq 4\) shows that for fixed \(u, v\) there is at most one value of \(t \ll \frac{x^2}{\log x}\) such that \(P^t(u, v)\) is not exponentially small, while for larger values of \(t\) there are at most \(O\left(\frac{x}{\sqrt{x^2/\log x}}\right)^{(d-1)/2}\) such values of \(t\).

Applying this to \(B_t\) the above relations finishes the proof. For instance, with 2 generators then
\[
B_t = O\left(\mathbb{E} \sum_{i=1}^{T(c)} \frac{1}{\sqrt{T(c)}} \max \left\{1, \frac{T(c)}{x} \right\}\right) \leq \mathbb{E} \left(1 + \int_1^{T(c)} \frac{1}{\sqrt{t}} dt + \frac{T(c)}{x}\right) = \mathbb{E} \left(1 + 2\sqrt{T(c)} + \frac{T(c)}{x}\right) \leq 1 + 2\sqrt{ET(c)} + \frac{ET(c)}{x}
\]
These results can be used together with Theorem 2.3 and a warm-up of $\frac{N}{32}$ steps as discussed in the introductory heuristic to complete a proof of collision times. See [2] for details worked out when $x = 2$ is fixed.

Note that in our proof the quantity $T = O(x^2)$. If it is possible to show a similar relation for the closely related quantity $T(\epsilon)$, then the bounds of Lemma 5.1 improve to $B = O(x)$ when $d = 2$ and $B = O(\log x)$ when $d = 3$. When used in Theorem 2.3 these lead to optimal bounds on expected collision time: $O(N^{2/3})$ for $d = 2$ and $O(\sqrt{N \log N})$ for $d = 3$.

References


