

1 Fraud Detection for Random Walks

2 **Varsha Dani** ✉🏠

3 Rochester Institute of Technology, Rochester NY, U.S.A

4 **Thomas P. Hayes** ✉🏠

5 University at Buffalo, Buffalo NY U.S.A

6 **Seth Pettie** ✉🏠

7 University of Michigan, Ann Arbor MI, U.S.A

8 **Jared Saia** ✉🏠

9 University of New Mexico, Albuquerque NM, U.S.A

10 — Abstract —

11 Traditional fraud detection is often based on finding statistical anomalies in data sets and transaction
12 histories. A sophisticated fraudster, aware of the exact kinds of tests being deployed, might be
13 difficult or impossible to catch. We are interested in paradigms for fraud detection that are *provably*
14 *robust* against any adversary, no matter how sophisticated. In other words, the detection strategy
15 should rely on signals in the data that are inherent in the goals the adversary is trying to achieve.

16 Specifically, we consider a fraud detection game centered on a random walk on a graph. We
17 assume this random walk is implemented by having a player at each vertex, who can be honest or
18 not. In particular, when the random walk reaches a vertex owned by an honest player, it proceeds
19 to a uniformly random neighbor at the next timestep. However, when the random walk reaches a
20 dishonest player, it instead proceeds to an arbitrary neighbor chosen by an omniscient Adversary.

21 The game is played between the Adversary and a Referee who sees the trajectory of the random
22 walk. At any point during the random walk, if the Referee determines that a *specific* vertex is
23 controlled by a dishonest player, the Referee accuses that player, and therefore wins the game. The
24 Referee is allowed to make the occasional incorrect accusation, but must follow a policy that makes
25 such mistakes with small probability of error. The goal of the adversary is to make the cover time
26 large, ideally infinite, i.e., the walk should *never* reach at least one vertex. We consider the following
27 basic question: how much can the omniscient Adversary delay the cover time without getting caught?
28 Our main result is a tight upper bound on this delay factor.

29 We also discuss possible applications of our results to settings such as Rotor Walks, Leader
30 Election, and Sybil Defense.

31 **2012 ACM Subject Classification** Theory of computation → Random walks and Markov chains;
32 Mathematics of computing → Probability and statistics; Security and privacy → Intrusion detection
33 systems

34 **Keywords and phrases** Fraud detection, random processes, Markov chains

35 **Digital Object Identifier** 10.4230/LIPIcs.ITCS.2024.??

36 **Funding** *Seth Pettie*: Supported by NSF Grants CCF-1815316 and CCF-2221980.

37 *Jared Saia*: Supported by NSF Grants CNS-2210299 and NSF NRI-2024520.

38 **1** Introduction

39 Many modern fraud detection efforts look for *statistical* features of data that do not fit a known
40 probabilistic model, or are intrinsically implausible or internally inconsistent. The Newcomb–
41 Benford (“first digit”) Law [29, 30, 26, 28] is a well known filter for detecting fabricated data
42 in financial records, which can be applied to detecting fraud in other numerical data, e.g.,
43 manipulated images [14, 44]. Recently uncovered frauds in social science research [35, 36, 37]
44 can also be seen as *distribution testing* against known or unknown distributions.



© Varsha Dani, Thomas P. Hayes, Seth Pettie and Jared Saia;
licensed under Creative Commons License CC-BY 4.0

15th Innovations in Theoretical Computer Science Conference (ITCS 2024).

Editor: Venkatesan Guruswami; Article No. ??; pp. ??:1–??:22

Leibniz International Proceedings in Informatics



Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

45 One weakness of this variety of fraud detection is that it preys on relatively unsophisticated
 46 fraudsters, who could easily evade detection if they were just aware of the statistical tests in
 47 advance. This critique could also be leveled against most fraud detection efforts in machine
 48 learning and information retrieval, which treat it as a pattern-matching problem [9, 13, 40,
 49 38, 42, 31, 39].

50 In this paper we advance a perspective on fraud detection that differs sharply from
 51 all the work cited above. First, rather than begin with an *application domain* or a single
 52 empirical *instance* of fraud, we want to build a more general theory of fraud detection. In
 53 the most fundamental examples cited above, fraud manifests as corruption of a random
 54 process. Thus, we focus our study on abstract random processes that can be perturbed by an
 55 adversary. Furthermore, we adopt the norms of theoretical computer science, cryptography,
 56 and game theory in our adversarial model. In particular, a fraud detection mechanism
 57 should be evaluated in a *worst case* fashion, ideally against a computationally unbounded
 58 and omniscient adversary. Following Kerckhoffs’ principle [23], its success should *not* depend
 59 on assuming the adversary is ignorant of the statistical tests it will be subject to.

60 **1.1 Fraud Detection for Random Walks**

61 Let $G = (V, E)$ be a connected, undirected graph. A random walk $(v_i)_{i \geq 0}$ is generated by
 62 placing a token at some v_0 and, in each step, letting v_i pass the token to a uniformly random
 63 neighbor $v_{i+1} \in N(v_i)$. The cover time for this walk is the time until all the vertices have
 64 been visited by the token.

65 Now suppose an adversary *corrupts* a set $B \subseteq V$ of up to b vertices, who may pass the
 66 token as they like. The adversary wishes to delay the cover time as much as possible, without
 67 being detected.

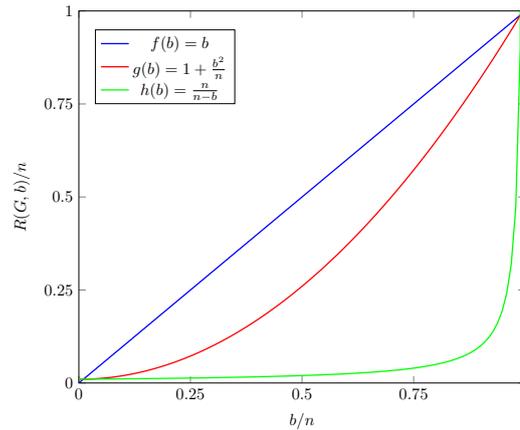
68 It is well known [2] that for any graph, the cover time is $O(mn \log n)$ with high probability.
 69 So if, after this many steps, there are vertices that have not been reached, the *existence* of
 70 corruption will be evident. However, we require a stronger form of fraud detection: a *specific*
 71 vertex must be accused. We formalize this process as the following game.

72 ► **Definition 1 (The Random Walk Game).** *Let T be a fixed time horizon and $b \leq n$*
 73 *a fixed number. The game is played between two players, the Referee and the Adversary.*
 74 *The Adversary picks a starting vertex $v_0 \in V$ and a subset $B \subseteq V$ of (corrupt) vertices with*
 75 *$|B| \leq b$. A walk (v_0, v_1, \dots, v_T) is constructed iteratively, with each move from an honest*
 76 *vertex being random, and each move from a corrupt vertex being chosen by the Adversary.*
 77 *If $\{v_0, v_1, \dots, v_T\} = V$, the Referee immediately wins (the vertex set has been covered).*
 78 *Otherwise, the Referee must specify one “accused” vertex; the Referee wins if and only if this*
 79 *vertex is in B .*

80 We are interested in the *threshold time*, $T(G, b)$, which is the minimum time T such that,
 81 with best play, the Referee wins the T -step Random Walk Game with probability at least
 82 $1 - 1/n^5$.

83 We note that there is nothing particularly special about the exponent 5 in the allowed
 84 error probability above, and could instead make the error probability $1/n^C$. However, for our
 85 lower bounds, we do require that C be large enough to avoid pathological examples where
 86 the Referee could accuse a random vertex of being in B and be correct just by chance.

87 When $b = 0$, the threshold time $T(G, 0)$ is essentially the expected cover time. More
 88 precisely, if τ is the maximum, over all starting locations, of the expected cover time of G ,



■ **Figure 1** Comparison of the bounds on $R(G, b)$ from our main results. All the log terms have been dropped, and n has been set to 100. The blue curve is b , the general upper bound on R from Theorem 3. We note that, for every b , there is a graph for which this upper bound is tight (up to log factors). The red curve is $1 + b^2/n$, which is $\Theta(R)$ in the special case of the path, as stated in Theorem 6. Note that $1 + b^2/n$ is also the right value of $R(b)$ in the special case of the clique, if the referee is restricted to purely local strategies that make accusations only a function of the particular player's choices. The green curve gives the correct value of $R(b)$ for the clique, when the referee is allowed to make accusations based on the entire transcript. This result is given in Theorem 13.

89 then

$$90 \quad \tau/2 \leq T(G, 0) \leq (10 \log n) \tau$$

91 with the actual value depending on the specific graph.

92 We now introduce our main object of study.

93 ► **Definition 2.** We define the price of corruption as the ratio

$$94 \quad R(G, b) = \frac{T(G, b)}{T(G, 0)}$$

95 Informally, this is the factor by which an adversary with up to b corrupt vertices can increase
96 the cover time, before the referee will be able to reliably accuse a bad player.

97 Our goal in this paper is to understand how much $T(G, b)$, and therefore $R(G, b)$, can
98 depend on b . Our main result is that this dependence is at most nearly linear

99 ► **Theorem 3** (Price of Corruption is at most nearly linear). Let G be any graph on n vertices,
100 and let $0 \leq b \leq n$. Then,

$$101 \quad R(G, b) = O(b \log n).$$

102 Moreover, there exists a family of graphs $G = G(n, b)$ for which

$$103 \quad R(G, b) = \Omega(b / \log n).$$

104 The lower bound in Theorem 3 does not apply to all graphs. For instance, we will see
105 that the behavior of $R(G, b)$ is more nuanced in the cases when G is a path or a clique; we
106 examine these special cases in Sections 2 and 4

107 This suggests a related question: for which graphs is the Price of Corruption, $R(G, b)$,
108 smallest? Knowing this might be helpful in applications where we have some choice about
109 the graph on which the random walk takes place. Small-degree expander graphs seem like
110 particularly good candidates for bounds of this type.

111 **1.2 Related Work**

112 *Biased* random walks are a mainstay of introductory courses in random processes. Azar,
 113 Broder, Karlin, Linial, and Phillips [8] studied the adversarial biasing of random walks to
 114 maximize the time spent among some target set. In their model the token moves randomly a
 115 $(1 - \epsilon)$ -fraction of the time, and is controlled by the adversary an ϵ -fraction of the time. Azar
 116 et al. [8] did not consider the problem of *detecting* such interventions or evading detection.

117 Our problem is inspired by the Byzantine Agreement protocols of King and Saia [24]
 118 and Huang, Pettie, and Zhu [21, 22], which achieved polynomial latency with $f = \Theta(n)$ and
 119 optimal $f < n/3$ resiliency (Byzantine corruptions), respectively. These protocols attempt
 120 to flip a fair coin via a natural distributed coin-flipping protocol. However, the adversary
 121 may interfere with the protocol by choosing coin-flip outcomes strategically, and by inducing
 122 subtle disagreements among the non-corrupt players. If such an adversary continually foils
 123 attempts to flip a fair coin, an individual Byzantine player can be identified and *blacklisted*,
 124 removing its influence over the coin flipping protocol.¹

125 The notion of fraud detection seems to be “in the air” these days. This year Alon, Gunby,
 126 He, Shmaya, and Solan [3] also proposed a fraud detection-type game for random walks. In
 127 their model a walk on \mathbb{Z} begins near the origin and is run in perpetuity but never reaches
 128 the origin, or does not reach it infinitely often. The movement of the walk is controlled by
 129 two players, Alice and Bob, who alternate (purportedly) flipping fair coins and announcing
 130 outcomes in $\{-1, 1\}$ —but exactly one of them is a fraud. The question is how to detect
 131 which of Alice or Bob is not behaving correctly. Their fraud detection mechanism is not an
 132 “algorithm” *per se*, as it requires evaluating functions of infinitely long walks. Although our
 133 setup and the setup of [3] have some syntactic similarities, the mathematical structure of the
 134 two problems are different and lead, in some ways, to opposite conclusions.²

135 **1.2.1 Random Walks and Dynamic Networks**

136 Several recent results make use of random walks to solve classic problems in distributed
 137 computing over dynamically changing networks in the presence of Byzantine nodes. Problems
 138 addressed include Byzantine agreement [5]; information dissemination [34]; and leader
 139 election [6]. See also [7] for a survey of results.

140 The type of random walk problem considered in these results is more general than ours
 141 in that the network topology may change from step-to-step. The problem is more specific
 142 than ours in that the network is assumed to always be a regular expander; and the number
 143 of Byzantine nodes is always $O(\sqrt{n}/\log^k n)$ for some constant k .

144 Central to these results is a technical lemma showing that if good nodes generate random
 145 walk tokens at a certain rate, then there is a large set of nodes that have access to many
 146 well-mixed random-walk tokens. The random-walk algorithms are simple: no attempt is
 147 made to detect or identify Byzantine behavior, and the algorithms are fully distributed and
 148 scalable in terms of latency and message cost.

¹This application illustrates why it is important to distinguish between global detection — *something* has gone wrong — and specific detection, namely, a *specific* player is corrupt w.h.p.

²Specifically, to make the cover time infinite in our model, the corrupt vertices must have some measurable bias, and the question is how long it takes to detect that bias. In the infinite Alice & Bob game [3], any biases are trivially detected (in the limit); the detector must also pay attention to negative correlations between Alice and Bob’s moves.

1.3 Organization

In Section 1.4 we review Bernstein’s and Freedman’s concentration inequalities. In Section 2 we analyze the random walk game on the simplest topology, an n -path P_n , and obtain nearly sharp bounds on $T(P_n, b)$. In Section 3 we generalize the detection method to work on an arbitrary graph G , and bound the price of corruption by $R(G, b) = O(b \log n)$. In Section 4 we design a fraud detection method specific to the n -clique K_n , and give nearly tight upper and lower bounds on $T(K_n, b)$. In Section 5 we discuss some possible applications of our results. We conclude with some open problems in Section 6.

1.4 Concentration Inequalities

The Referee’s task is to observe the random walk, and identify vertices that are not behaving as they should. In order to do this, we need a fairly accurate idea of what the local behavior of such a random walk *should* look like. To get a handle on this, we will make use of the following concentration inequalities.

The following version of Bernstein’s inequality (see [16]) will be useful in analyzing the random walk games on the path (Section 2) and the clique (Section 4).

► **Theorem 4.** (*Bernstein’s Inequality*) Let X_1, \dots, X_n be independent random variables with $|X_i - E(X_i)| \leq b$ for each $i \in [n]$, and each with variance σ_i^2 . Let $X = \sum_i X_i$, and $\sigma^2 = \sum_i \sigma_i^2$ be the variance of X . Then for all $t > 0$,

$$\Pr(X \leq E(X) - t) \leq \exp\left(-\frac{t^2}{2\sigma^2 + (2/3)bt}\right)$$

When dealing with general graphs (Section 3) we will instead need the following extension of Freedman’s inequality for martingales.

► **Theorem 5.** ([10, Lem. 2]) Suppose X_1, \dots, X_T is a martingale difference sequence with $|X_t| \leq \rho$. Let $\mathbf{Var}_t X_t = \mathbf{Var}(X_t \mid X_1, \dots, X_{t-1})$. Let $V = \sum_{t=1}^T \mathbf{Var}_t X_t$ be the sum of conditional variances and $\bar{\sigma} = \sqrt{V}$. Then for any $\delta < 1/e$ and $T \geq 4$,

$$\mathbb{P}\left(\left|\sum_{t=1}^T X_t\right| \leq 2\sqrt{\ln(1/\delta)} \max\{2\bar{\sigma}, \rho\sqrt{\ln(1/\delta)}\}\right) \geq 1 - \delta \log T.$$

2 The Path

Consider the path graph $G = (V, E)$ with vertices numbered 1 through n . Without loss of generality we can assume the token is initially at vertex 1 and never reaches vertex n . How long must a corrupted random walk be until we may accuse a corrupt vertex?

Theorem 6 gives nearly sharp bounds for this class of graphs and illustrates two qualitative features of this fraud detection model. First, although *one* corrupt vertex can make the cover time infinite it cannot do so without detection, and in fact any coalition of $b = O(\sqrt{n})$ corrupt vertices is powerless to increase the cover time by more than a constant factor, without detection. Second, there is a significant gap between the moment we detect likely corruption ($\Theta(n^2 \log n)$ time) and the moment we can confidently level an accusation at one vertex ($\tilde{\Theta}(n^3)$ time when $b = \Omega(n)$).

► **Theorem 6.** Let G be the path of length n . Suppose the Random Walk Game on G is played for T timesteps and the adversary is allowed to corrupt up to b vertices. Then

6.6 Fraud Detection for Random Walks

1. If $T = \Omega((n^2 + nb^2) \log n)$, then there is a strategy that enables the Referee to win with probability at least $1 - \frac{1}{n^5}$. In other words,

$$R(G, b) = O\left(1 + \frac{b^2}{n}\right)$$

2. If $T < n^2 + nb^2$, there is an adversarial strategy such that one vertex is never visited, and no detection mechanism can identify any corrupt vertex with high probability. In other words,

$$R(G, b) = \Omega\left(\left(1 + \frac{b^2}{n}\right) / \log(n)\right)$$

The remainder of this section constitutes a proof of Theorem 6.

Part 1 of Theorem 6. Suppose we pass the token for T time steps. For each vertex j , let X_j denote the number of times that vertex j passes the token, and Y_j the number of times j passes the token to the left. We will accuse vertex j if the number of left passes, Y_j substantially exceeds the number of right passes, $X_j - Y_j$, or more specifically, if

$$\Delta_j \stackrel{\text{def}}{=} 2Y_j - X_j \geq \sqrt{CX_j \log n}.$$

A standard application of Chernoff's bound ensures that this criterion almost certainly does not falsely accuse any good vertex. In other words after T timesteps, for each good player j ,

$$\Delta_j < \sqrt{CX_j \log n} \tag{1}$$

holds with high probability $1 - n^{-\Omega(C)}$. We may assume that (1) also holds for all bad players as well, since otherwise the algorithm will make a correct accusation.

Let $v^* = \arg \max_v X_v$ be the *mode* vertex. Since the token is passed around for T timesteps, by the pigeonhole principle, $X_{v^*} \geq T/n$. Each time v^* receives the token, it passes it to a neighbor and the token makes a round-trip excursion back to v^* . We may assume that at least $\frac{1}{3}X_{v^*}$ of these round-trip excursions are to the right of v^* , for otherwise Δ_{v^*} would already be large enough to justify accusing v^* . Define G and B to be the sets of good and bad vertices among $\{v^* + 1, v^* + 2, \dots, n - 1\}$. Since vertex n is never reached, every round-trip excursion from v^* to the right entails $G \cup B$ passing the token one more time to the left than the right.

Let Δ_G and Δ_B be the sum of this left-excess associated with G and B , respectively. Then we know $\Delta_G + \Delta_B \geq \frac{1}{3}X_{v^*}$. Let X_G and X_B denote the total number of times the token is passed by vertices in G and B , respectively.

Applying Chernoff's bound to the good vertices as a group, since $T \leq nX_{v^*}$, we have, with high probability $1 - n^{-\Omega(C)}$,

$$\Delta_G \leq \sqrt{CX_G \log n} \leq \sqrt{CT \log n} \leq \sqrt{CnX_{v^*} \log n}. \tag{2}$$

To estimate $\Delta_B = \sum_{i \in B} \Delta_i$, note that each bad vertex satisfies Eqn. (1) to avoid detection.

$$\Delta_B \leq \sum_{i \in B} \sqrt{CX_i \log n} \leq |B| \sqrt{CX_{v^*} \log n} \leq b \sqrt{CX_{v^*} \log n}. \tag{3}$$

where the second inequality follows from the choice of v^* as the mode. Combining Eqns.(2,3), we have

$$\frac{1}{3}X_{v^*} \leq \Delta_G + \Delta_B \leq \sqrt{CX_{v^*} \log n}(\sqrt{n} + b).$$



■ **Figure 2** A biased random walk on the line graph.

221 Squaring and rearranging terms,

$$222 \quad X_{v^*} \leq 9C \log n (\sqrt{n} + b)^2.$$

223 Finally, since $(\sqrt{n} + b)^2 \leq 2(n + b^2)$, we have

$$224 \quad T \leq nX_{v^*} \leq 18Cn(n + b^2) \log n,$$

225 which completes the proof of the upper bound. ◀

226 **Part 2 of Theorem 6.** For this part it is more convenient to number the vertices in reverse
227 order: vertex 0 is the rightmost vertex and the token begins at vertex $n - 1$ and never reaches
228 0. See Figure 2.

229 In adversarial strategy \mathcal{S} , the adversary corrupts vertices in $[b] = \{1, \dots, b\}$ and gives
230 vertex j a left-bias of $1/j$. Specifically, vertex j passes left with probability p_j , where

$$231 \quad p_j = \begin{cases} \frac{1}{2} \left(1 + \frac{1}{j}\right) & \text{if } j \in [b] \\ 1/2 & \text{if } j > b. \end{cases}$$

232 This process corresponds to a reversible Markov chain on the states $\{n - 1, \dots, 1\}$ where, for
233 $j \in [b]$, the edge between $j + 1$ and j has weight $\binom{j+1}{2}$ and all the edges to the left of b have
234 weight $\binom{b+1}{2}$. Note that the cover time is infinite as vertex 0 is unreachable.

235 It follows that, for $j \in [b]$, vertex j has stationary probability proportional to j^2 , while
236 the vertices in $\{n - 2, \dots, b + 1\}$ to the left of b all have probability proportional to $b(b + 1)$.
237 The leftmost vertex $n - 1$ has stationary probability proportional to $b(b + 1)/2$, and vertex 0
238 is unreachable. See Figure 2. Summing these terms, we obtain the normalization factor N
239 to be

$$240 \quad N = b \frac{(b+1)}{2} + b(b+1)(n-b-2) + \sum_{j=1}^b j^2 = b(b+1) \left(n - b - \frac{3}{2} + \frac{2b+1}{6} \right) = \Theta(nb^2).$$

242 Thus, for $j \in [b]$, the stationary probability of vertex j is $j^2/N = \Theta(nj^2/b^2)$.

243 Define \mathcal{S}_{-j} to be identical to strategy \mathcal{S} except that vertex j is not corrupt, i.e., it passes
244 left and right with probability $1/2$. We want to argue that if a corrupted random walk is too
245 short, the false positive rate of *any* detection strategy will be intolerably large. Lemma 7
246 lower bounds this error.

► **Lemma 7.** *Assume the adversary picks a strategy from $\{\mathcal{S}, \mathcal{S}_{-1}, \dots, \mathcal{S}_{-b}\}$ uniformly at random. Abusing notation, let \mathcal{S}_{-i} also refer to the event that strategy \mathcal{S}_{-i} is chosen. Let W be the resulting corrupted random walk. Define $q = \min_{i \in [b]} \min(\rho_i, 1 - \rho_i)$, where $\rho_i = \Pr(\mathcal{S}_{-i} \mid (\mathcal{S}_{-i} \cup \mathcal{S}), W)$. Then,*

$$\forall i \in [b], \Pr(\mathcal{S}_{-i} \mid W) \geq q^2/b.$$

247 **Proof.** Note that for all $i \in [b]$:

$$248 \quad \Pr(\mathcal{S}_{-i} \mid W) = \Pr(\mathcal{S}_{-i} \mid (\mathcal{S}_{-i} \cup \mathcal{S}), W) \cdot (\Pr(\mathcal{S}_{-i} \mid W) + \Pr(\mathcal{S} \mid W)).$$

8.8 Fraud Detection for Random Walks

250 Letting $\rho_i = \Pr(\mathcal{S}_{-i} \mid (\mathcal{S}_{-i} \cup \mathcal{S}), W)$, and solving for $\Pr(\mathcal{S}_{-i} \mid W)$ in the above, we get

$$251 \quad \Pr(\mathcal{S}_{-i} \mid W) = \frac{\rho_i}{1 - \rho_i} \cdot \Pr(\mathcal{S} \mid W).$$

252 Note that $q \leq \frac{\rho_i}{1 - \rho_i} \leq 1/q$. Letting $i^* = \arg \min_{i \in [b]} \Pr(\mathcal{S}_{-i} \mid W)$, we have:

$$253 \quad \Pr(\mathcal{S}_{-i^*} \mid W) \geq q \Pr(\mathcal{S} \mid W),$$

254 and for all $i \in [b] \setminus \{i^*\}$,

$$255 \quad \Pr(\mathcal{S}_{-i} \mid W) \leq (1/q) \Pr(\mathcal{S} \mid W).$$

256 Hence,

$$257 \quad 1 = \Pr(\mathcal{S}_{-i^*} \mid W) + \Pr(\mathcal{S} \mid W) + \sum_{i \in [b], i \neq i^*} \Pr(\mathcal{S}_{-i} \mid W)$$

$$258 \quad \leq \Pr(\mathcal{S}_{-i^*} \mid W) + (1/q) \Pr(\mathcal{S}_{-i^*} \mid W) + ((b-1)/q^2) \Pr(\mathcal{S}_{-i^*} \mid W)$$

$$259 \quad \leq \frac{b}{q^2} \Pr(\mathcal{S}_{-i^*} \mid W),$$

260 where the last inequality follows since $1 + 1/q \leq 1/q^2$. ◀

261 Lemma 7 says that we can assume the detector accuses the vertex $j \in [b]$ that minimizes
262 $\Pr(\mathcal{S}_{-j} \mid (\mathcal{S}_{-j} \cup \mathcal{S}), W)$. In order to make the false positive rate small, we need the likelihood
263 ratio

$$264 \quad \frac{\Pr(W \mid \mathcal{S}_{-j})}{\Pr(W \mid \mathcal{S})} = \frac{(1/2)^{X_j}}{p_j^{Y_j} (1 - p_j)^{X_j - Y_j}} = \left(1 - \frac{1}{j}\right)^{Y_j} \left(1 + \frac{1}{j-1}\right)^{X_j - Y_j}$$

$$265 \quad < \exp\left(-\frac{Y_j}{j} + \frac{X_j - Y_j}{j-1}\right) = \exp\left(-\frac{\Delta_j}{j-1} + \frac{Y_j}{j(j-1)}\right)$$

266 to be $n^{-\Omega(C)}$. By Chernoff bounds, the likelihood ratio never gets this small until $Y_j \geq$
267 $Cj^2 \log n$, so we may use this as a proxy prerequisite for accusing vertex j .

268 Once j is visited for the first time, the expected return time is $\Theta(nb^2/j^2)$, so *in expectation*,
269 the criterion $Y_j \geq Cj^2 \log n$ is satisfied after another $\Theta(nb^2 \log n)$ steps. However, these
270 return times have large variances so it is not clear that this random variable is sufficiently
271 concentrated around its mean.³

272 We may assume without loss of generality that $b \in [\Omega(\sqrt{n}), n/2]$. Let W_j be the length of
273 a random walk that begins and ends at vertex j , conditioned on moving left initially and let
274 $E_j = \mathbb{E}(W_j)$ and $V_j = \mathbb{E}(W_j^2)$.⁴ Such a walk moves to $j+1$, makes zero or more roundtrips
275 from $j+1$, and then returns to j . The number of roundtrips from $j+1$ is distributed
276 geometrically, so by linearity of expectation,

$$277 \quad E_j = 2 + \left(\frac{1}{1 - p_{j+1}} - 1\right) E_{j+1} = \begin{cases} 2 + \frac{j+2}{j} E_{j+1} & \text{if } j+1 \leq b, \\ 2 + E_{j+1} & \text{if } j+1 > b. \end{cases}$$

³(Lemma 8 implies that in any graph, the visitation rate of a vertex is, with high probability, at most twice its stationary probability after a sufficiently long (corrupted) random walk, which on the line would be $\tilde{\Theta}(bn^2)$ steps. Since we are looking for a tight bound of $\tilde{\Theta}(n^2 + nb^2)$ we require a more careful analysis.)

⁴The condition that $b \leq n/2$ implies that starting at a corrupt vertex, a roundtrip to the left is longer in expectation than a roundtrip in general.

283 Thus $E_b = 2(n - b - 1)$ and writing E_j , $j < b$, in terms of E_b we have a telescoping
284 product, $E_j = \Theta(nb^2/j^2)$.

285 To bound the second moment V_j , suppose that a leftward roundtrip from j makes k
286 (leftward) roundtrips from $j + 1$ before returning to j , i.e., it has length $2 + W_{j+1}^{(1)} + \dots + W_{j+1}^{(k)}$,
287 where the $W_{j+1}^{(i)}$ are independent copies of W_{j+1} . This would contribute $kV_{j+1} + (k^2 -$
288 $k)E_{j+1}^2 + 2kE_{j+1} + 4$ to V_j . Thus, we can express V_j recursively as

$$\begin{aligned} 289 \quad V_j &= (1 - p_{j+1}) \sum_{k \geq 0} p_{j+1}^k (kV_{j+1} + (k^2 - k)E_{j+1}^2 + 2kE_{j+1} + 4) \\ 290 \quad &= (1 - p_{j+1}) \frac{p_{j+1}}{(1 - p_{j+1})^2} \cdot V_{j+1} + \Theta(E_{j+1}^2) \\ 291 \quad &= \begin{cases} \frac{j+2}{j} \cdot V_{j+1} + \Theta(E_{j+1}^2) & \text{if } j + 1 \leq b, \\ V_{j+1} + \Theta(E_{j+1}^2) & \text{if } j + 1 > b. \end{cases} \\ 292 \end{aligned}$$

293 Then $V_b = \Theta(n^3)$, and expressing V_j , $j < b$, in terms of V_b we have another telescoping
294 product with $V_j = \Theta(n^3b^2/j^2)$.

295 We claim that it is not possible to reliably accuse any vertex j in less than $M = n^2 + nb^2$
296 steps. In particular, once j is first visited, the length of the next $Y_j = K = Cj^2 \log n$ leftward
297 roundtrips is not less than M . Let $W_j^{(i)}$ be the length of the i th leftward roundtrip. In
298 expectation $W_j^{(1)} + \dots + W_j^{(K)}$ is $K \cdot E_j = \Theta(Cnb^2 \log n) = \mu$.

299 We may assume each $|W_j^{(i)}| \leq M$, for otherwise there's nothing to prove. Thus, by
300 Bernstein's inequality,

$$\begin{aligned} 301 \quad \Pr \left(\sum_{i=1}^K W_j^{(i)} < M \right) &< \exp \left(- \frac{(\mu - M)^2}{2 \sum_{i=1}^K \mathbb{E}((W_j^{(i)})^2) + (2/3)M(\mu - M)} \right) \\ 302 \quad &= \exp \left(- \frac{(1 - o(1))(KE_j)^2}{2KV_j + (2/3)(1 + o(1))KE_j \cdot M} \right) \\ 303 \quad &= \exp \left(- \frac{\Theta(Cnb^2 \log n)^2}{\Theta(Cn^3b^2 \log n) + \Theta(Cnb^2 \log n \cdot (n^2 + nb^2))} \right) \\ 304 \end{aligned}$$

305 and since $b = \Omega(\sqrt{n})$, $n^2b^4 = \Omega(n^3b^2)$,

$$306 \quad = \exp(-\Omega(C \log n)) = n^{-\Omega(C)}. \quad \blacktriangleleft$$

308 Theorem 6 gives a nearly tight characterization for the fraud detection time on paths.
309 Qualitatively speaking, Theorem 6 shows that tracking *individual* deviations suffices to
310 achieve near-optimal fraud detection, i.e., a vertex v is judged *solely* on the distribution of
311 token passes to $N(v)$. Section 3 extends this type of analysis to general graphs, and obtains
312 strong bounds for all graphs.

313 However, tracking *individual deviations* alone is, on some graph topologies, insufficient
314 for optimal fraud detection. The *clique* is one such topology, which we analyze in detail in
315 Section 4.

316 **3 Fraud Detection on General Graphs**

317 In this section we consider the random walk game played on an *arbitrary* connected graph G
318 on n vertices. The Adversary can corrupt any set $\mathcal{B} \subset V$ consisting of up to b of the vertices.
319 As in the case of the path, the Referee will watch the individual vertices and track their

320 apparent deviation from uniformly random behaviour. We will prove the upper bound in
 321 Theorem 3, by showing that there is a Referee strategy that guarantees that if the Adversary
 322 tries to delay the cover time by more than an $O(b \log n)$ factor, the Referee has a $1 - 1/n^5$
 323 chance to win. To prove the lower bound we will demonstrate a family of graphs and an
 324 Adversarial strategy for which a $O(b/\log n)$ factor is achieved.

325 Although we eventually want to bound how much the adversary can delay the cover time,
 326 it will be convenient analyze the Random Walk Game in terms of *hitting times*. In the next
 327 subsection we discuss some concepts and terminology relating to random walks and hitting
 328 times.

329 3.1 Notation

330 G is an undirected connected graph on n vertices. For a random walk on G , given vertices
 331 v and y , the *hitting time* from v to y is the (random) first time at which the walk, having
 332 started at v , arrives at y . The expected hitting times between all the pairs of vertices in G
 333 will be of particular interest in designing our Referee strategy..

334 The following quantities are solely a function of the structure graph G , not strategic
 335 considerations of the random walk game.

- 336 ■ π is the stationary distribution, i.e., $\pi(v) = \deg(v)/2m$.
- 337 ■ $H(v, y)$ is the expected hitting time to y starting from v . Let $H_{\max}(y) = \max_v H(v, y)$
 338 and $H_{\max} = \max_y H_{\max}(y)$ be the maximum hitting times when only y is fixed, and
 339 when neither is fixed.
- 340 ■ For $w \in N(v)$, define $h_y(v, w)$ to be

$$341 \quad h_y(v, w) = H(w, y) - H(v, y) + 1 - \frac{\mathbb{1}(v=y)}{\pi(y)}.$$

342 Here $\mathbb{1}(\mathcal{E})$ is the indicator variable for event \mathcal{E} . For $v \neq w$, this definition ensures that
 343 $h_y(v, w) - 1$ equals the change in the expected hitting time to y that results from moving
 344 across the edge $\{v, w\}$. The definition ensures that, when w is a randomly chosen neighbor
 345 of v , the quantity $h_y(v, w) - 1$ has an expected value of -1 , corresponding to the elapsing
 346 of the first time step in a random walk from v to y . The extra term, $\frac{\mathbb{1}(v=y)}{\pi(y)}$, which, when
 347 $v = y$, equals the expected excursion time from y , ensures that the expected value of
 348 $h_y(v, w)$ is zero for all $v \in V$.

- 349 ■ Let $\rho_y(v) = \max_{w \in N(v)} |h_y(v, w)|$, $\rho_y = \max_v \rho_y(v)$ and $\rho = \max_y \rho_y$.
- 350 ■ Let $\sigma_y^2(v) = \frac{1}{\deg(v)} \sum_{w \in N(v)} h_y(v, w)^2$ be the conditional variance of $H(w, y)$, conditioned
 351 on v , where as before we assume that w is a randomly chosen neighbor of v .
- 352 ■ Let $\mathcal{V}_\pi^y = \mathbb{E}_\pi \sigma_y^2(v) = \sum_v \sigma_y^2(v) \pi(v)$ be the average conditional variance when v is chosen
 353 from the stationary distribution. Let $\mathcal{V}_\pi = \max_y \mathcal{V}_\pi^y$ be the maximum of this average over
 354 all target vertices y .

355 Now consider the Random Walk Game. Let \mathcal{G} be the set of good vertices and \mathcal{B} be the
 356 set of bad vertices, with $|\mathcal{B}| \leq b$. Let T be the number of time steps for which the random
 357 walk game will be played, and for $t \in [0, T]$ let v_t denote the vertex holding the token at
 358 time t . Then for each t , v_{t+1} is a neighbor of v_t , and it is a uniformly random neighbor if v_t
 359 is a good vertex (*i.e.* $v_t \in \mathcal{G}$).

360 For each $v \in V$, let S_v denote the set of times when the token is at v , and let $S_{\mathcal{G}}$ denote
 361 the times when the token is with a good vertex in \mathcal{G} and $S_{\mathcal{B}}$ denote the times when the token

is with a bad vertex in \mathcal{B} . Also, let T_v , $T_{\mathcal{G}}$ and $T_{\mathcal{B}}$ denote the sizes of the corresponding sets of times. That is

$$\begin{aligned} S_v &= \{t \mid v_t = v\}, & T_v &= |S_v|, \\ S_{\mathcal{G}} &= \{t \mid v_t \text{ is a good vertex}\}, & T_{\mathcal{G}} &= |S_{\mathcal{G}}|, \\ S_{\mathcal{B}} &= \{t \mid v_t \text{ is a bad vertex}\}, & \text{and } T_{\mathcal{B}} &= |S_{\mathcal{B}}|. \end{aligned}$$

For a target vertex $y \in V$, we want to track the evolution of the values $H(v_t, y)$. Let

$$\Delta_t = H(v_{t+1}, y) - H(v_t, y)$$

be the change in expected hitting time at step t . Observe that $\mathbb{E}(\Delta_t \mid v_t \neq y, v_t \in \mathcal{G}) = -1$ and $\mathbb{E}(\Delta_t \mid v_t = y, v_t \in \mathcal{G}) = 1/\pi(y) - 1$. This motivates the definition of the sequence D_t^y :

$$D_t^y = \Delta_t + 1 - \frac{\mathbb{1}(v_t = y)}{\pi(y)} = h_y(v_t, v_{t+1})$$

It follows that if $v \in \mathcal{G}$ is any fixed good vertex and $y \in V$ any target, that $\mathbb{E}(D_t^y \mid v_t = v) = 0$ and moreover,

- The subsequence $(D_t^y : v_t = v)$ is a martingale difference sequence with step sizes bounded by $\rho_y(v)$,
- The subsequence $(D_t^y : v_t \in \mathcal{G})$ is a martingale difference sequence with step sizes bounded by ρ_y .

The above sequences are martingale difference sequences because, at timesteps when the token is controlled by good players, the next player is chosen fairly, and cannot be predicted in advance by the Adversary. The specific martingale difference sequence depends on the Adversary's strategy for the bad players' moves.

3.2 Referee Strategy

The referee's strategy will be based on Theorem 5 ([10, Lemma 2]), which is a version of Freedman's inequality for martingales.

Since $(D_t^y : v_t = v)$ is a martingale difference sequence with step sizes bounded by $\rho_y(v)$ whenever v is a good vertex, applying Freedman's inequality with $\delta = 1/n^C$, we know that for each good vertex v and target y ,

$$\Pr \left(\left| \sum_{t \in S_v} D_t^y \right| \geq \max \left\{ 4\sqrt{C\sigma_y^2(v)T_v \ln n}, 2C\rho_y(v) \ln n \right\} \right) \leq \frac{\log T_v}{n^C}. \quad (4)$$

With this in mind, we will accuse vertex v if $|\sum_{t \in S_v} D_t^y|$ is suspiciously large. Specifically, we will accuse v if

$$\exists y \in V. \quad \left| \sum_{t \in S_v} D_t^y \right| \geq \max \left\{ 4\sqrt{C\sigma_y^2(v)T_v \ln n}, 2C\rho_y(v) \ln n \right\}$$

By a union bound over all v, y , the probability any good vertex is mistakenly accused is at most $n^{-C+2} \log T$.

395 **3.3 Analysis**

396 Suppose the token passing game is played for T time steps and no player is accused by the
 397 referee of Section 3.2. Let v^* be the “stationary mode,” *i.e.*, the vertex that is visited most
 398 frequently relative to its stationary probability. In particular, for all v ,

399
$$\frac{T_v}{\pi(v)} \leq \frac{T_{v^*}}{\pi(v^*)}.$$

400 We will denote by α the ratio between the number of times v^* is visited and the number of
 401 times it expects to be visited at stationarity. That is

402
$$\alpha = \frac{T_{v^*}}{T\pi(v^*)}.$$

403 Since v^* has been chosen to maximize the right hand side, and *some* vertex must be visited
 404 at least as often as expected, it follows that $\alpha \geq 1$. Also, we have for all v ,

405
$$T_v \leq \alpha T \pi(v) \tag{5}$$

406 Note that both v^* and α depend on the actual run of the game, so that they depend on
 407 T , the good players’ randomness and the adversarial strategy. Nevertheless, we can show
 408 that when T is sufficiently large, the adversary has only a limited ability to skew who gets
 409 the token. Recall that b is the number of bad players.

► **Lemma 8.** *If $T \geq \max\{6H_{max}, 144C\mathcal{V}_\pi(1+b)\ln n, 12C\rho(1+b)\ln n\}$ then $\alpha \leq 2$. That
 is, for every vertex y ,*

$$T_y \leq 2T\pi(y).$$

410 **Proof.** Since the bad vertices want to avoid getting accused, based on the referee’s strategy,
 411 we may assume that:

412
$$\forall v, y \in V. \left| \sum_{t \in S_v} D_t^y \right| \leq \max \left\{ 4\sqrt{C\sigma_y^2(v)T_v \ln n}, 2C\rho_y(v) \ln n \right\}. \tag{6}$$

413 Consider the sum $\sum_{t=0}^{T-1} D_t^y$. As $\sum_{t=0}^{T-1} \Delta_t$ telescopes to $H(v_T, y) - H(v_0, y)$ we have

414
$$\sum_{t=0}^{T-1} D_t^y = \sum_{t=0}^{T-1} \left(\Delta_t + 1 - \frac{\mathbb{1}(v_t = y)}{\pi(y)} \right) = H(v_T, y) - H(v_0, y) + T - \frac{T_y}{\pi(y)}$$

415 so that

416
$$T - \frac{T_y}{\pi(y)} \leq H(v_0, y) - H(v_T, y) + \sum_{t=0}^{T-1} D_t^y. \tag{7}$$

417 On the other hand, we can write

418
$$\left| \sum_{t=0}^{T-1} D_t^y \right| \leq \left| \sum_{t \in S_{\mathcal{G}}} D_t^y \right| + \left| \sum_{v \in \mathcal{B}} \sum_{t \in S_v} D_t^y \right|.$$

419 Of the two sums on the right, we can deal with the first one by directly applying Freedman’s
 420 inequality, since the subsequence $(D_t^y : v_t \in \mathcal{G})$ is actually a martingale difference sequence

421 with step sizes bounded by $\rho_y = \max_v \rho_y(v)$. Thus, by Theorem 5, with error probability
 422 $(\log T)/n^C$, we have:

$$423 \quad \left| \sum_{t \in S_G} D_t^y \right| \leq \max \left\{ 4\sqrt{C\mathcal{V}_G^y \ln n}, 2C\rho_y \ln n \right\} \quad (8)$$

$$424 \quad \leq 4\sqrt{C\mathcal{V}_G^y \ln n} + 2C\rho_y \ln n, \quad (9)$$

426 where \mathcal{V}_G^y is the sum of the conditional variances of the steps of the martingale. It is bounded
 427 by

$$428 \quad \mathcal{V}_G^y = \sum_{t \in S_G} \mathbf{Var}_t D_t^y = \sum_{v \in \mathcal{G}} \sum_{t \in S_v} \mathbf{Var}(D_t^y | v_t = v)$$

$$429 \quad = \sum_{v \in \mathcal{G}} \sum_{t \in S_v} \sigma_y^2(v)$$

$$430 \quad = \sum_{v \in \mathcal{G}} \sigma_y^2(v) T_v \leq \alpha T \sum_{v \in \mathcal{G}} \sigma_y^2(v) \pi(v),$$

432 where the last line follows from equation (5). Plugging this back into (9), we get

$$433 \quad \left| \sum_{t \in S_G} D_t^y \right| \leq 4\sqrt{2\alpha T \ln n \sum_{v \in \mathcal{G}} \sigma_y^2(v) \pi(v)} + 4\rho_y \ln n. \quad (10)$$

435 To bound the corresponding term of the bad players we use apply Eqns. (5) and (6).

$$436 \quad \left| \sum_{v \in \mathcal{B}} \sum_{t \in S_v} D_t^y \right| \leq \sum_{v \in \mathcal{B}} \max \left\{ 4\sqrt{C\sigma_y^2(v) T_v \ln n}, 2C\rho_y(v) \ln n \right\}$$

$$437 \quad \leq \sum_{v \in \mathcal{B}} \left(4\sqrt{C\sigma_y^2(v) T_v \ln n} + 2C\rho_y(v) \ln n \right)$$

$$438 \quad \leq \sum_{v \in \mathcal{B}} \left(4\sqrt{C\sigma_y^2(v) T \alpha \pi(v) \ln n} + 2C\rho_y(v) \ln n \right)$$

$$439 \quad \leq 4\sqrt{C\alpha T \ln n} \left(\sum_{v \in \mathcal{B}} \sqrt{\sigma_y^2(v) \pi(v)} \right) + 2Cb\rho_y \ln n$$

$$440 \quad \leq 4\sqrt{C\alpha T \ln n} \sqrt{b \sum_{v \in \mathcal{B}} \sigma_y^2(v) \pi(v)} + 2Cb\rho_y \ln n \quad (11)$$

442 where (11) follows from the Cauchy-Schwarz inequality. By Cauchy-Schwarz again,

$$443 \quad \sqrt{\sum_{v \in \mathcal{G}} \sigma_y^2(v) \pi(v)} + \sqrt{b \sum_{v \in \mathcal{B}} \sigma_y^2(v) \pi(v)} \leq \sqrt{(1+b) \sum_v \sigma_y^2(v) \pi(v)} = \sqrt{\mathcal{V}_\pi^y (1+b)}. \quad (12)$$

444 Combining (10), (11), and (12), we obtain

$$445 \quad \left| \sum_{t=0}^{T-1} D_t^y \right|$$

$$446 \quad \leq 4\sqrt{C\alpha T \ln n \sum_{v \in \mathcal{G}} \sigma_y^2(v) \pi(v)} + 2C\rho_y \ln n + 4\sqrt{C\alpha T b \ln n \sum_{v \in \mathcal{B}} \sigma_y^2(v) \pi(v)} + 2Cb\rho_y \ln n$$

$$447 \quad \leq 4\sqrt{C\alpha T \mathcal{V}_\pi^y (1+b) \ln n} + 2C\rho_y (1+b) \ln n. \quad (13)$$

??:14 Fraud Detection for Random Walks

449 Now, plugging Eqn. (13) back into (7), and noting that $|H(v_0, y) - H(v_T, y)| \leq H_{\max}(y)$, we
 450 have that for every target y ,

$$\begin{aligned}
 451 \quad \left| T - \frac{T_y}{\pi(y)} \right| &\leq |H(v_1, y) - H(v_{T+1}, y)| + \left| \sum_{t=0}^{T-1} D_t^y \right| \\
 452 \quad &\leq H_{\max}(y) + 4\sqrt{C\alpha T \mathcal{V}_\pi^y(1+b) \ln n} + 2C\rho_y(1+b) \ln n. \tag{14} \\
 453
 \end{aligned}$$

Recall that for the stationary mode vertex v^* , $T_{v^*} = \alpha T \pi(v^*)$, where $\alpha > 1$. Dividing by T
 454 and fixing $y = v^*$, we have

$$\begin{aligned}
 455 \quad \alpha - 1 &\leq \frac{H_{\max}(v^*)}{T} + 4\sqrt{\frac{C\alpha \mathcal{V}_\pi^{v^*}(1+b) \ln n}{T}} + \frac{2C\rho_{v^*}(1+b) \ln n}{T} \\
 456 \quad &\leq \frac{H_{\max}}{T} + 4\sqrt{\frac{C\alpha \mathcal{V}_\pi(1+b) \ln n}{T}} + \frac{2C\rho(1+b) \ln n}{T} \\
 457 \quad &\leq \frac{1}{6} + \frac{\sqrt{\alpha}}{3} + \frac{1}{6} \tag{15} \\
 458
 \end{aligned}$$

459 where (15) follows because $T \geq \max\{6H_{\max}, 144C\mathcal{V}_\pi(1+b) \ln n, 12C\rho(1+b) \ln n\}$. Thus α
 460 satisfies the quadratic inequality

$$461 \quad \alpha - \frac{\sqrt{\alpha}}{3} - \frac{4}{3} \leq 0,$$

462 which implies $\sqrt{\alpha} \leq 4/3$ and hence $\alpha \leq 16/9 < 2$. Substituting this back into Eqn. (5) proves
 463 the lemma. \blacktriangleleft

464 The proof of Lemma 8 shows that for sufficiently large T we can drive α arbitrarily close
 465 to 1. Moreover the proof actually shows something even stronger. Let $\hat{\pi}$ denote the empirical
 466 distribution of how often each vertex is visited. By definition, for all y ,

$$467 \quad \hat{\pi}(y) = \frac{T_y}{T}.$$

468 Using the fact that $\alpha < 2$ in Eqn. (14), we see that for all y ,

$$469 \quad \left| T - \frac{T_y}{\pi(y)} \right| \leq H_{\max}(y) + 4\sqrt{2CT \mathcal{V}_\pi^y(1+b) \ln n} + 2C\rho_y(1+b) \ln n$$

471 Dividing by T ,

$$472 \quad \left| 1 - \frac{\hat{\pi}(y)}{\pi(y)} \right| \leq \frac{H_{\max}(y)}{T} + 4\sqrt{\frac{2C\mathcal{V}_\pi^y(1+b) \ln n}{T}} + \frac{2C\rho_y(1+b) \ln n}{T}.$$

474 We restate this as a Corollary of Lemma 8.

475 **► Corollary 9.** *If $T \geq \max\{6H_{\max}, 144C\mathcal{V}_\pi(1+b) \ln n, 12C\rho(1+b) \ln n\}$ then for every*
 476 *vertex y ,*

$$477 \quad \left| 1 - \frac{\hat{\pi}(y)}{\pi(y)} \right| \leq \frac{H_{\max}(y)}{T} + 4\sqrt{\frac{2C\mathcal{V}_\pi^y(1+b) \ln n}{T}} + \frac{2C\rho_y(1+b) \ln n}{T}.$$

478 Corollary 9 actually implies that it is impossible for the adversary to prolong the game
 479 for this many time steps without detection. If some vertex $x \in V$ has never passed the

480 token, then $\hat{\pi}(x) = 0$ and $1 - \hat{\pi}(x)/\pi(x) = 1$. By Corollary 9, if no accusations yet have been
481 leveled, then

$$482 \quad 1 \leq \frac{H_{\max}(x)}{T} + 4\sqrt{\frac{2C\mathcal{V}_\pi^y(1+b)\ln n}{T}} + \frac{2C\rho_y(1+b)\ln n}{T}$$

$$483 \quad \leq \frac{1}{6} + \frac{\sqrt{2}}{3} + \frac{1}{6}$$

$$484 \quad < 1.$$

485 which is a contradiction. Thus we have shown that

486 **► Theorem 10.** *If $T = \Omega(H_{\max} + b(\mathcal{V}_\pi + \rho) \log n)$ then the Referee of Section 3.2 wins
487 with probability at least $1 - 1/n^5$. In other words,*

$$488 \quad T(G, b) = O(H_{\max} + b(\mathcal{V}_\pi + \rho) \log n)$$

489 We can simplify the expression of Theorem 10 as follows. Since for all y , $\rho_y \leq H_{\max}(y)$
490 it follows that $\rho \leq H_{\max}$. Furthermore, for any y the stationary conditional variance can
491 be bounded by $\mathcal{V}_\pi^y \leq 2\mathbb{E}_\pi H(\cdot, y) \leq 2H_{\max}(y)$. (For completeness, these last inequalities are
492 proved in Appendix A.) Thus, $\mathcal{V}_\pi + \rho \leq 3H_{\max}$. Also, for any graph, $H_{\max} = O(mn)$. This
493 establishes the following Corollary.

494 **► Corollary 11.** *For any graph G and any number b of bad players,*

$$495 \quad T(G, b) = O(bH_{\max} \log n) = O(bmn \log n).$$

496 The following corollary also follows directly from the above bounds and Corollary 9.

497 **► Corollary 12.** *For any graph G , any number b of bad players, and a walk that lasts for T
498 steps with no accusations, if*

$$499 \quad T = \Omega(mnb \log n)$$

500 then, for every vertex y ,

$$501 \quad \left| 1 - \frac{\hat{\pi}(y)}{\pi(y)} \right| = O\left(\sqrt{\frac{mnb \log n}{T}}\right).$$

502 To relate these results back to the *price of corruption*, we note that the maximum expected
503 cover time is clearly at least H_{\max} , and therefore $T(G, 0) \geq H_{\max}$. Moreover, repeatedly
504 applying Markov's inequality shows that regardless of the starting vertex, after $6H_{\max} \log n$
505 steps, the probability that a particular vertex is unreached, is at most $1/n^6$. Taking a union
506 bound over all the vertices, after $6H_{\max} \log n$ steps, the probability that there is an unreached
507 vertex is at most $1/n^5$ and therefore $T(G, 0) = O(H_{\max} \log n)$, and $R(G, b) = O(b \log n)$.

508 This bound on $R(G, b)$ is close to tight, as witnessed by the class of *Ball & Chain* graphs.
509 Let $BC_{n,b}$ consist of an $(n-b)$ -clique attached to a b -path; we assume $b \leq n/2$. Starting
510 from a vertex in the “ball,” the cover time is $H_{\max} = \Theta(n^2b)$, thus, the zero-corruption game
511 threshold is $T(BC_{n,b}, 0) = \Theta(n^2b \log n)$. We now need to lower bound $T(BC_{n,b}, b) = \Omega(n^2b^2)$.
512 The corrupt vertices will lie only on the chain, and bias the walk slightly towards the ball,
513 as in the proof of Theorem 6. Let u be the common vertex of the ball and chain. By
514 Theorem 6, the walk restricted to the chain takes $\Omega(b^3)$ time steps, with high probability.
515 Vertex u sees the token at least as often as in a truly random walk, which would be at least
516 $\Omega(b^2)$ times. Each time u takes the token from the chain, it returns it to the chain after
517 $\Theta(n^2)$ steps, in expectation, walking around the ball. Hence $T(BC_{n,b}, b) = \Omega(n^2b^2)$ and
518 $R(BC_{n,b}, b) = \Omega(b/\log n)$.

519 Putting this all together, we have established Theorem 3.

521 **4 The Clique**

522 In this section, we analyze the Random Walk Game on the clique K_n . Here, every starting
 523 vertex is equivalent, and the hitting time to any vertex is a geometric random variable with
 524 mean $n - 1$. Thus $H_{\max} = n - 1$. Moreover, the cover time for the clique is essentially the
 525 coupon collector problem, and therefore the maximum expected cover time is $O(n \log n)$,
 526 and the cover time is at most $\beta n \log n$ with probability $1 - 1/n^{\beta-1}$. The results of Section 3
 527 tell us that $T(K_n, b) = O(bn \log n)$. When $b = \Omega(n)$, that is an upper bound of $O(n^2 \log n)$.
 528 In this section, we show that in fact, we can use the structure of the clique to get a much
 529 better bound. To get a sense of why fraud detection is faster for the clique, consider the
 530 game from the Adversary's perspective, and suppose the adversary wants to select a vertex
 531 that will not be reached. In a graph where there are low degree vertices, the adversary can
 532 surround such a vertex with corrupted vertices, all of whom always pass the token to one
 533 of their other neighbors. But in the clique, unless the Adversary takes over $n - 1$ vertices,
 534 every vertex has some good neighbors, who will pass it the token every n th time they get it,
 535 on average. This makes the Adversary's task much more difficult.

536 Theorem 13 gives near-tight bounds on the fraud detection time for cliques. The Referee's
 537 strategy differs from the strategy for the path or for a general graph, in that we take the
 538 entire trajectory of the walk into account when judging how a vertex v passes the token.

539 **► Theorem 13.** *Consider the Random Walk Game played on an n -clique, in which the*
 540 *adversary controls b vertices.*

- 541 1. *There is a Referee strategy that enables the Referee to win with high probability after*
 542 *$T(K_n, b) = O\left(\frac{n^2 \log n \log(n/(n-b))}{n-b}\right)$ steps.*
- 543 2. *Moreover, there is an adversarial strategy for b corrupted players such that any accusation*
 544 *within $O\left(\frac{n^2 \log n}{n-b}\right)$ steps cannot be correct with probability $1 - n^{-5}$, so that $T(G, b) =$*
 545 *$\Omega\left(\frac{n^2 \log n}{n-b}\right)$.*

546 The remainder of this section constitutes a proof of Theorem 13.

547 Let C be a sufficiently large constant and G, B be the sets of good and bad players. If
 548 the G -players collectively pass the token $Cn \ln n$ times then the game will end naturally with
 549 high probability $1 - n^{-C+1}$, regardless of what other actions are taken by B .

550 Suppose the path taken by the token in T steps is $P = (v_1, v_2, \dots, v_T)$. When the token
 551 is at v_i , define L_i ("low" vertices) to be the set of vertices that have passed the token less
 552 than $2C|G|^{-1}n \ln n$ times. In the beginning $|L_1| = n$ and once $|L_i| \leq |G|/2$ at least $|G|/2$
 553 vertices are not in L and the game has already ended, with high probability.

554 We partition time into stages, where stage $j \in [0, \log(2n/|G|)]$ covers the time that
 555 $|L_i| \in (n/2^{j+1}, n/2^j]$. Fix some stage j and let X_v be the number of times v passes the token
 556 in stage j and Y_v be the number of times v passes it to an L -vertex. Note that if v is good,
 557 Y_v is the sum of X_v indicator variables each with mean at least $2^{-(j+1)}$ and variance less
 558 than 2^{-j} . By Bernstein's inequality, $\Pr(Y_v < 2^{-(j+1)}X_v - t) < \exp\left(-\frac{t^2}{2 \cdot 2^{-j}X_v + (2/3)t}\right)$. We
 559 will accuse v whenever $Y_v \leq 2^{-(j+1)}X_v - \sqrt{C2^{-(j+1)}X_v \ln n}$. Thus, with probability $n^{-\Omega(C)}$
 560 no good vertex is accused. Suppose that stage j lasts for $T_j = 4C|G|^{-1}n^2 \ln n$ steps without

561 any vertex being accused. Then:

$$\begin{aligned}
562 \quad \sum_{v \in V} Y_v &\geq \sum_{v \in V} \left(2^{-(j+1)} X_v - \sqrt{C 2^{-(j+1)} X_v \ln n} \right) \\
563 \quad &\geq 2^{-(j+1)} T_j - \sqrt{2^{-(j+1)} T_j \cdot C n \ln n} && \text{(Cauchy-Schwarz)} \\
564 \quad &\geq 2^{-(j+2)} T_j && \text{(Since: } C n \ln n = T_j |G| / (4n) \leq T_j 2^{-(j+1)} / 4) \\
565 \quad &= (n/2^{j+1}) \cdot (2C|G|^{-1} n \ln n). \\
566
\end{aligned}$$

567 However, if this were true then the number of L -vertices would have already shrunk
568 to less than $n/2^{j+1}$, ending stage j . Thus, stage j cannot last for $4Cn^2|G|^{-1} \ln n$ steps
569 without accusing a vertex of corruption. In total the number of steps before an accusation is
570 $O\left(\frac{n^2 \log n \log(n/|G|)}{|G|}\right) = O\left(\frac{n^2 \log n \log(n/(n-b))}{n-b}\right)$.

571 Turning to the lower bound, suppose we are aiming to make a correct accusation with
572 probability $1 - n^{-C}$. Suppose the adversary picks a set $B \subseteq V$ uniformly at random with
573 $|B| = b$. Under strategy \mathcal{S} it corrupts B and under strategy \mathcal{S}_{-j} , $j \in B$, it corrupts $B - \{j\}$.
574 In either case, whenever a corrupt vertex v has the token it passes it to a neighbor in
575 B uniformly at random. The adversary chooses its strategy uniformly at random from
576 $\{\mathcal{S}\} \cup \{\mathcal{S}_{-j}\}_{j \in B}$. Let \mathcal{E} be the event that, after a walk of length $T = n^2 \ln n / (n - b)$, every
577 vertex in B has only passed the token to others in B . Since, by Chernoff bounds, each vertex
578 in B sees the token less than $3n \ln n / (n - b)$ times with probability $1 - o(1)$, we have

$$579 \quad \Pr(\mathcal{E}) \geq (1 - o(1))(1 - (n - b)/n)^{3n \log n / (n - b)} = \Omega(n^{-3}).$$

580 Moreover, $\Pr(\mathcal{S}_{-j} \mid (\mathcal{S}_{-j} \cup \mathcal{S}), \mathcal{E}) = q = 1/2$ since once we condition on \mathcal{E} , \mathcal{S}_{-j} and \mathcal{S} behave
581 identically. By Lemma 7, the probability of error is at least $q^2/b = 1/(4b)$ after conditioning
582 on \mathcal{E} , hence at least $\Omega(n^{-3}b^{-1})$ with no conditioning. For $C > 4$ this bound does not meet
583 the desired n^{-C} error bound.

584 This concludes the proof of Theorem 13. It says that a coalition of $(1 - \epsilon)n$ bad vertices
585 can delay the hitting time by $\Omega(\epsilon^{-1} n \ln n)$ and $O(\epsilon^{-1} \log \epsilon^{-1} n \ln n)$, i.e., no asymptotic delay
586 at all when ϵ is constant. This is quite different than the line graph, in which a tiny minority
587 of $\omega(\sqrt{n})$ bad vertices can asymptotically delay the hitting time.

588 **5 Applications**

589 **5.1 Rotor Walks and Derandomization.**

590 Our results show that if all nodes pass the token in a way that is *locally* balanced across their
591 neighbors, then the resulting *global* random walk has good cover time. The local balance
592 condition can be ensured even without making any random choices. For example, in the
593 *rotor walk* algorithm [32, 17, 20], every node passes the token to each of its neighbors in a
594 round-robin fashion whenever it receives the token. A rotor walk ensures that every node
595 satisfies the referee of Section 3.2.

596 Thus, Corollaries 11 and 12 directly apply to rotor walks when we set $b = n$. They
597 give results analogous to Theorems 2 and 3 of [20]. In particular, Corollary 11 bounds
598 the cover time of rotor walks, and Corollary 12 bounds the occupation frequencies. Our
599 results are weaker than Theorems 2 and 3 of [20] in that they only apply to walks on
600 unweighted, undirected graphs. But, they are stronger in that they apply to a broader class
601 of derandomization techniques: for example, any routing works that ensures token passing is
602 locally balanced across neighbors as specified by the referee of Section 3.2.

603 **5.2 Leader Election.**

604 Leader election is a fundamental problem in distributed computing [12, 11, 27, 33, 25,
 605 45]. Consider a simple communication model common to blockchains: there is a public
 606 key infrastructure (PKI) over the players, and communication occurs synchronously via
 607 a broadcast primitive that enables any player to send to all other players in the network
 608 (See [18, 15, 19]). Further, assume there is a publicly-known connected and regular graph G
 609 that has m edges and n nodes.

610 Corollary 12 enables us to perform repeated leader elections such that after $T =$
 611 $O(mn^2 \log n)$ elections, the fraction of good players elected approximates the fraction of good
 612 players, or else at least one bad player is caught.

613 The algorithm to achieve this is simple. First, the player with the token is the leader for
 614 that turn. This leader chooses one of its neighbors in G to pass the token to, and broadcasts
 615 a cryptographically signed message giving their choice. The PKI prevents equivocation, and
 616 synchronous communication forces some choice to be made, or else the current leader is
 617 known to be bad. Since all players learn all choices of the other players, each player can
 618 individually enforce the referee strategy of Section 3.2.

619 **5.3 Sybil Defense**

620 Consider a graph G with n nodes and $m = \Theta(n)$ edges, where (1) the bad and good nodes
 621 are separated by a cut with only α crossing edges; and (2) the subgraph induced by the
 622 good nodes is an expander. We want for almost all good nodes v , that node v learns a set of
 623 players S_v such that (1) S_v contains almost all of the good nodes; and (2) S_v contains “few”
 624 bad nodes. A simple algorithm is for each node to start a random walk at each of its edges,
 625 and for each of these walks to continue for $\Theta(\sqrt{n \ln n})$ steps. Then, for each node v , S_v is
 626 the set of all nodes w such that there is some node in the intersection of the nodes visited by
 627 walks starting at v and the nodes visited by walks starting at w .

628 This problem and algorithm is inspired by random-walk based Sybil defenses prevalent in
 629 the academic literature [43, 41, 4, 1], particularly the work of Yu, Kaminsky, Gibbons and
 630 Flaxman [43]. The graph represents a social network where the good nodes and the Sybil
 631 nodes are typically separated by a “small” number of crossing edges.

632 We can extend our referee and Corollary 12 to handle the algorithm described here that
 633 creates many random walks. Each edge has probability $1/m$ in the stationary distribution,
 634 and the initial steps in the algorithm above are also distributed uniformly over the edges.
 635 Thus, as the number of steps increases, each edge is visited $\Theta(\sqrt{n \ln n})$ times. This is true
 636 no matter what choices are made by the Sybil nodes, provided none of them are caught by
 637 the referee.

638 Thus, the total number of steps on the α crossing edges should be $\Theta(\alpha\sqrt{n \ln n})$. Call
 639 a random walk *bad* if it crosses one of the crossing edges and *good* otherwise. Then, there
 640 are at most $\Theta(\alpha\sqrt{n \ln n})$ bad walks. In particular, assuming $\alpha = o(\sqrt{n/\log(n)})$, the vast
 641 majority of the random walks starting on good nodes visit only good nodes.

642 Since the graph induced by the good nodes is an expander, with high probability, each
 643 pair of good random walks starting at two good nodes will intersect. Let \mathcal{G} be the set of
 644 good nodes and \mathcal{B} be the set of bad nodes. Then, by the above, there is a set $\mathcal{G}' \subseteq \mathcal{G}$ such
 645 that $|\mathcal{G}'| = \Omega(n - \alpha\sqrt{n \ln n})$ and for all $v \in \mathcal{G}'$, $\mathcal{G}' \subseteq S_v$ and $|S_v \cap \mathcal{B}| = O(\alpha\sqrt{n \ln n})$.

646 Thus, if $\alpha = o(\sqrt{n/\ln n})$, and we say that node u trusts node v if $v \in S_u$, we can say
 647 the following. There is a set, \mathcal{G}' of all but a $o(1)$ fraction of the good nodes such that: all
 648 nodes in \mathcal{G}' mutually trust each other; and every node in \mathcal{G}' has a $o(1)$ fraction of Sybil nodes

649 among the nodes it trusts.

650 **6 Conclusion and Open Problems**

651 It is well known that real-world fraud can sometimes be discovered by looking for statistical
652 anomalies in data sets or transaction records. However, these statistical tests [29, 30, 26, 28,
653 35, 36, 37] work best on *unsophisticated* fraudsters, and may not work against adversaries
654 who operate with full knowledge of the specific statistical tests.

655 In this paper we advocated for an approach to fraud detection that is *abstract, robust*
656 against sophisticated adversaries, and *rigorous* in its quantitative guarantees.⁵ We illustrated
657 how rigorous fraud detection against powerful adversaries can work in a simple abstract
658 setting, namely *random walks on undirected graphs* in which *vertices* can be corrupted by
659 the adversary; cf. [3, 22]. There are several directions for future research.

- 660 ■ One of our findings is that there can be a large delay between the time to detect the
661 existence of fraud, w.h.p., and the time to make an accurate *accusation*, w.h.p. One could
662 explore less strict notions of “accurate” accusation. In some contexts it may be fine to
663 accuse a set $S \subseteq V$, such that 90% of S is corrupt, w.h.p.
- 664 ■ Given a specific graph G , we may be interested in the gap between its cover time and
665 the fraud detection time against an adversary controlling b vertices. Up to log-factors we
666 understand this gap on the path and clique, and know the extremal bound for arbitrary
667 graphs, which is attained by the Ball and Chain graph. However, it is an open problem
668 to efficiently *compute* this gap-factor for a specific G , or to bound it in terms of natural
669 parameters of G , e.g., diameter.
- 670 ■ There are several algorithmic problems from the adversary’s perspective. Given a graph
671 G and budget b , which b vertices should be corrupted to maximize the time of detection?
672 To lower bound the detection time, we considered adversaries that corrupt vertices by
673 simply changing the transition probabilities for their incident edges; such adversarial
674 strategies are *Markovian*. Is there a specific graph for which all Markovian strategies are
675 *suboptimal*? If so, it would be interesting to see what a superior non-Markovian strategy
676 would look like.
- 677 ■ A natural direction is to consider random walks on directed, strongly connected graphs.

678 In general, the fraud detection paradigm can be introduced into the analysis of essentially
679 any random process where it is conceivable that some or all of the randomness is being
680 controlled by an adversary to achieve an unlikely outcome.

681 **References**

- 682 1 Muhammad Al-Qurishi, Mabrook Al-Rakhami, Atif Alamri, Majed Alrubaian, Sk Md Mizanur
683 Rahman, and M Shamim Hossain. Sybil defense techniques in online social networks: a survey.
684 *IEEE Access*, 5:1200–1219, 2017.
- 685 2 Romas Aleliunas, Richard M. Karp, Richard J. Lipton, László Lovász, and Charles Rackoff.
686 Random walks, universal traversal sequences, and the complexity of maze problems. In
687 *Proceedings of the 20th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*,
688 pages 218–223, 1979. doi:10.1109/SFCS.1979.34.

⁵In the context of some purportedly random process, *fraud* is defined as effecting a particular outcome that is statistically unlikely by corrupting elements of the random process.

- 689 3 Noga Alon, Benjamin Gunby, Xiaoyu He, Eran Shmaya, and Eilon Solan. Identifying the deviator. *CoRR*, abs/2203.03744, 2022. arXiv:2203.03744, doi:10.48550/arXiv.2203.03744.
- 690
- 691 4 Lorenzo Alvisi, Allen Clement, Alessandro Epasto, Silvio Lattanzi, and Alessandro Panconesi. Communities, random walks, and social sybil defense. *Internet Mathematics*, 10(3-4):360–420, 2014.
- 692
- 693
- 694 5 John Augustine, Gopal Pandurangan, and Peter Robinson. Fast byzantine agreement in dynamic networks. In *Proceedings of the 2013 ACM symposium on Principles of distributed computing*, pages 74–83, 2013.
- 695
- 696
- 697 6 John Augustine, Gopal Pandurangan, and Peter Robinson. Fast byzantine leader election in dynamic networks. In *International Symposium on Distributed Computing*, pages 276–291. Springer, 2015.
- 698
- 699
- 700 7 John Augustine, Gopal Pandurangan, and Peter Robinson. Distributed algorithmic foundations of dynamic networks. *ACM SIGACT News*, 47(1):69–98, 2016.
- 701
- 702 8 Yossi Azar, Andrei Z. Broder, Anna R. Karlin, Nathan Linial, and Steven J. Phillips. Biased random walks. *Combinatorica*, 16(1):1–18, 1996. doi:10.1007/BF01300124.
- 703
- 704 9 Nikesh Bajaj, Tracy Goodluck Constance, Marvin Rajwadi, Julie A. Wall, Mansour Moniri, Cornelius Glackin, Nigel Cannings, Chris Woodruff, and James Laird. Fraud detection in telephone conversations for financial services using linguistic features. *CoRR*, abs/1912.04748, 2019. URL: <http://arxiv.org/abs/1912.04748>, arXiv:1912.04748.
- 705
- 706
- 707
- 708 10 Peter Bartlett, Varsha Dani, Thomas Hayes, Sham Kakade, Alexander Rakhlin, and Ambuj Tewari. High-probability regret bounds for bandit online linear optimization. In *Proceedings of the 21st Annual Conference on Learning Theory-COLT 2008*, pages 335–342. Omnipress, 2008.
- 709
- 710
- 711
- 712 11 Michael Ben-Or, M Linial, and Michael Saks. *Collective coin flipping and other models of imperfect randomness*. IBM Thomas J. Watson Research Division, 1989.
- 713
- 714 12 Michael Ben-Or and Nathan Linial. Collective coin flipping, robust voting schemes and minima of banzhaf values. In *26th Annual Symposium on Foundations of Computer Science (sfcs 1985)*, pages 408–416. IEEE, 1985.
- 715
- 716
- 717 13 Romain Bertrand, Petra Gomez-Krämer, Oriol Ramos Terrades, Patrick Franco, and Jean-Marc Ogier. A system based on intrinsic features for fraudulent document detection. In *Proceedings 12th International Conference on Document Analysis and Recognition (ICDAR)*, pages 106–110, 2013. doi:10.1109/ICDAR.2013.29.
- 718
- 719
- 720
- 721 14 Nicolò Bonettini, Paolo Bestagini, Simone Milani, and Stefano Tubaro. On the use of Benford’s law to detect GAN-generated images. In *Proceedings 25th International Conference on Pattern Recognition (ICPR)*, pages 5495–5502, 2020. doi:10.1109/ICPR48806.2021.9412944.
- 722
- 723
- 724 15 Jing Chen and Silvio Micali. Algorand: A secure and efficient distributed ledger. *Theoretical Computer Science*, 777:155–183, 2019.
- 725
- 726 16 Devdatt P. Dubhashi and Alessandro Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press, 2009. URL: <http://www.cambridge.org/gb/knowledge/isbn/item2327542/>.
- 727
- 728
- 729 17 Ioana Dumitriu, Prasad Tetali, and Peter Winkler. On playing golf with two balls. *SIAM Journal on Discrete Mathematics*, 16(4):604–615, 2003.
- 730
- 731 18 Juan Garay, Aggelos Kiayias, and Nikos Leonardos. The Bitcoin backbone protocol: Analysis and applications. In *Annual international conference on the theory and applications of cryptographic techniques*, pages 281–310. Springer, 2015.
- 732
- 733
- 734 19 Yossi Gilad, Rotem Hemo, Silvio Micali, Georgios Vlachos, and Nikolai Zeldovich. Algorand: Scaling byzantine agreements for cryptocurrencies. In *Proceedings of the 26th symposium on operating systems principles*, pages 51–68, 2017.
- 735
- 736
- 737 20 Alexander E Holroyd and James Propp. Rotor walks and markov chains. *Algorithmic probability and combinatorics*, 520:105–126, 2010.
- 738

- 739 21 Shang-En Huang, Seth Pettie, and Leqi Zhu. Byzantine agreement in polynomial time with
740 near-optimal resilience. In *Proceedings of the 54th Annual ACM Symposium on Theory of*
741 *Computing (STOC)*, pages 502–514, 2022. doi:10.1145/3519935.3520015.
- 742 22 Shang-En Huang, Seth Pettie, and Leqi Zhu. Byzantine agreement with optimal resilience via
743 statistical fraud detection. *CoRR*, abs/2206.15335, 2022. arXiv:2206.15335, doi:10.48550/
744 arXiv.2206.15335.
- 745 23 A. Kerckhoffs. La cryptographie militaire. *Journal des Sciences Militaires*, pages 161–191,
746 1883.
- 747 24 Valerie King and Jared Saia. Byzantine agreement in expected polynomial time. *J. ACM*,
748 63(2):13:1–13:21, 2016. doi:10.1145/2837019.
- 749 25 Valerie King, Jared Saia, Vishal Sanwalani, and Erik Vee. Scalable leader election. In *SODA*,
750 volume 6, pages 990–999, 2006.
- 751 26 Alex Ely Kossovsky. *Benford’s Law: Theory, the General Law of Relative Quantities, and*
752 *Forensic Fraud Detection Applications*. World Scientific, 2014. doi:https://doi.org/10.
753 1142/9089.
- 754 27 Nathan Linial. *Games computers play: Game-theoretic aspects of computing*. Citeseer, 1992.
- 755 28 Steven J. Miller. *Benford’s Law: Theory and Applications*. Princeton University Press,
756 Princeton, N.J., 2015.
- 757 29 Mark J. Nigrini. *Digital analysis using Benford’s Law*. Global Audit Publications, 2000.
- 758 30 Mark J. Nigrini. *Benford’s Law: Applications for Forensic Accounting, Auditing, and Fraud*
759 *Detection*. Wiley, Hoboken, N.J., 2012.
- 760 31 Shashank Pandit, Duen Horng Chau, Samuel Wang, and Christos Faloutsos. Netprobe:
761 a fast and scalable system for fraud detection in online auction networks. In *Proceedings*
762 *of the 16th International Conference on World Wide Web (WWW)*, pages 201–210, 2007.
763 doi:10.1145/1242572.1242600.
- 764 32 Vyatcheslav B Priezzhev, Deepak Dhar, Abhishek Dhar, and Supriya Krishnamurthy. Eulerian
765 walkers as a model of self-organized criticality. *Physical Review Letters*, 77(25):5079, 1996.
- 766 33 Alexander Russell and David Zuckerman. Perfect information leader election in $\log^* n + o(1)$
767 rounds. *Journal of Computer and System Sciences*, 63(4):612–626, 2001.
- 768 34 Atish Das Sarma, Anisur Rahaman Molla, and Gopal Pandurangan. Distributed computation
769 in dynamic networks via random walks. *Theoretical Computer Science*, 581:45–66, 2015.
- 770 35 Uri Simonsohn. Just post it: The lesson from two cases of fabricated data detected by statistics
771 alone. *Psychological Science*, 24(10):1875–1888, 2013.
- 772 36 Uri Simonsohn, Joseph P. Simmons, and Leif D. Nelson. Better p -curves: Making p -curve
773 analysis more robust to errors, fraud, and ambitious p -hacking, a reply to Ulrich and Miller
774 (2015). *Journal of Experimental Psychology*, 144(6):1146–1152, 2015.
- 775 37 Uri Simonsohn, Joseph P. Simmons, and Leif D. Nelson. Datacolada 98: Evidence of fraud in
776 an influential field experiment about dishonesty, 2021. URL: <http://datacolada.org/98>.
- 777 38 Niek Tax, Kees Jan de Vries, Mathijs de Jong, Nikoleta Dosoula, Bram van den Akker,
778 Jon Smith, Olivier Thuong, and Lucas Bernardi. Machine learning for fraud detection in
779 e-commerce: A research agenda. *CoRR*, abs/2107.01979, 2021. URL: <https://arxiv.org/abs/2107.01979>, arXiv:2107.01979.
- 780
- 781 39 Tian Tian, Jun Zhu, Fen Xia, Xin Zhuang, and Tong Zhang. Crowd fraud detection in internet
782 advertising. In *Proceedings of the 24th International Conference on World Wide Web (WWW)*,
783 pages 1100–1110. ACM, 2015. doi:10.1145/2736277.2741136.
- 784 40 Chen Wang, Yingtong Dou, Min Chen, Jia Chen, Zhiwei Liu, and Philip S. Yu. Deep fraud
785 detection on non-attributed graph. *CoRR*, abs/2110.01171, 2021. URL: <https://arxiv.org/abs/2110.01171>, arXiv:2110.01171.
- 786
- 787 41 Wei Wei, Fengyuan Xu, Chiu C Tan, and Qun Li. Sybildefender: A defense mechanism for
788 sybil attacks in large social networks. *IEEE transactions on parallel and distributed systems*,
789 24(12):2492–2502, 2013.

- 790 42 Chang Xu and Jie Zhang. Collusive opinion fraud detection in online reviews: A probabilistic
791 modeling approach. *ACM Trans. Web*, 11(4):25:1–25:28, 2017. doi:10.1145/3098859.
- 792 43 Haifeng Yu, Michael Kaminsky, Phillip B Gibbons, and Abraham Flaxman. Sybilguard:
793 defending against sybil attacks via social networks. In *Proceedings of the 2006 conference on*
794 *Applications, technologies, architectures, and protocols for computer communications*, pages
795 267–278, 2006.
- 796 44 João G. Zago, Fabio L. Baldissera, Eric A. Antonelo, and Rodrigo T. Saad. Benford’s
797 law: what does it say on adversarial images? *CoRR*, abs/2102.04615, 2021. URL: <https://arxiv.org/abs/2102.04615>, arXiv:2102.04615.
- 798 45 Mahdi Zamani, Mahnush Movahedi, and Mariana Raykova. Rapidchain: Scaling blockchain
800 via full sharding. In *Proceedings of the 2018 ACM SIGSAC conference on computer and*
801 *communications security*, pages 931–948, 2018.

802 **A Stationary Conditional Variances**

803 Fix any target y and let $H(v)$ be short for $H(v, y)$. The stationary conditional variance \mathcal{V}_π^y is

804
$$\mathcal{V}_\pi^y = \sum_{v \in V} \pi(v) \left(\frac{1}{\deg(v)} \sum_{w \in N(v)} (H(w) - H(v))^2 - \left(\frac{1}{\deg(v)} \sum_{w \in N(v)} H(w) - H(v) \right)^2 \right).$$

805

806 This is a centered second moment, and is therefore always less than the corresponding
807 uncentered second moment, which is better known as the Dirichlet form.

808
$$\mathcal{E}(H, H) = \sum_{v \in V} \sum_{w \in N(v)} \frac{\pi(v)}{\deg(v)} (H(v) - H(w))^2.$$

809 Since what we are about to say applies to arbitrary reversible Markov chains, we will switch
810 notations accordingly. Let P be the transition matrix for any reversible Markov chain on
811 state space V , with stationary distribution π . Reversible means that every pair of states v, w
812 satisfies the *detailed balance* condition,

813
$$\pi(v)P(v, w) = \pi(w)P(w, v).$$

814 In this setting, the Dirichlet form \mathcal{E} can be defined by either of the expressions below. Here,
815 $f, g : V \rightarrow \mathbb{R}$.

816
$$\mathcal{E}(f, g) = \sum_{v, w \in V} \pi(v)P(v, w)(f(v) - f(w))^2 = 2 \cdot \sum_{v, w \in V} \pi(v)P(v, w)f(v)(f(v) - f(w)).$$

817 Specializing to the case where $f = g = H$, where recall that H is the hitting time to a fixed
818 target state $y \in V$, we find that

819
$$\begin{aligned} \mathcal{E}(H, H) &= 2 \sum_{v, w \in V} \pi(v)P(v, w)H(v)(H(v) - H(w)) \\ &= 2 \sum_{v \in V} \pi(v)H(v) \left(1 - \frac{\mathbb{1}(v=y)}{\pi(y)} \right) \\ &= 2 \sum_{v \in V} \pi(v)H(v) && \text{since } H(y) = 0 \\ &= 2 \mathbb{E}_\pi H. \end{aligned}$$

820
821
822
823

824 Hence, for any y , $\mathcal{V}_\pi^y \leq 2 \mathbb{E}_\pi H$, and so $\mathcal{V}_\pi \leq 2H_{\max}$.