Rumor Source Detection With Multiple Observations Under Adaptive Diffusions

Miklós Z. Rácz and Jacob Richey

***Abstract*—Recent work, motivated by anonymous messaging platforms, has introduced adaptive diffusion protocols which can obfuscate the source of a rumor: a "snapshot adversary" with access to the subgraph of "infected" nodes can do no better than randomly guessing the entity of the source node. What happens if the adversary has access to multiple independent snapshots? We study this question when the underlying graph is the infinite d-regular tree. We show that (1) a weak form of source obfuscation is still possible in the case of two independent snapshots, but (2) already with three observations there is a simple algorithm that finds the rumor source with constant probability, regardless of the adaptive diffusion protocol. We also characterize the tradeoff between local spreading and source obfuscation for adaptive diffusion protocols (under a single snapshot). These results raise questions about the robustness of anonymity guarantees when spreading information in social networks.**

***Index Terms*—Information diffusion, social networks, source detection, source obfuscation.**

I. Introduction

Detecting the source of information diffusion on a network is an important problem in network science, with applications such as finding the source of a virus epidemic or finding the source of a rumor on Twitter. A prototypical graph on which source detection is studied is the infinite d-regular tree $\mathbb{T}_d$ (with $d \geq 3$), which is our focus in this paper as well.

Rumor Source Detection. Perhaps the simplest and most natural model of information diffusion on a network is the susceptible-infected (SI) model, where the rumor is spread along each edge of the network at a constant rate, and once a node is infected it remains infected forever. Shah and Zaman studied detecting the source in this model [1], [2]. Formally, at time $t = 0$ a vertex $v^* \in \mathbb{T}_d$ is "infected" and the information propagates on the network according to the SI model; one then observes the subset V_t of infected vertices at time t, which consists of $N_t := |V_t|$ vertices. We assume that the underlying graph (in this case $\mathbb{T}_d$) is known and hence the subgraph G_t induced by the vertices in V_t is also known. The goal is to find the rumor source v^*.

The maximum likelihood estimator (MLE) $\widehat{v}_{\mathrm{ML}} := \arg\max_{v \in V_t} \mathbb{P}(G_t \,|\, v^* = v)$ has particularly nice properties in this setting [1], [2]. In particular, Shah and Zaman showed that it is computable in linear time and that it detects the source with constant probability. More precisely, they show (in [3]) that there exists a universal constant $\alpha_d > 0$ such that $\lim_{t\to\infty} \mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^*) = \alpha_d$ (when $d \geq 3$). Many results extend to more general settings such as random trees [3].

Wang *et al.* [4] studied rumor source detection in the same setting but now with multiple independent observations; that is, observing the infected nodes $V_t^{(1)}, \ldots, V_t^{(k)}$ of k independent diffusions started from the same source v^*. They show that the detection probability increases with k and that it goes to 1 exponentially as $k \to \infty$.

Rumor Source Obfuscation. The results above show that if information propagates according to the SI model, then the source can be found efficiently and with good probability (that is, with at least constant probability). In certain applications, such as anonymous messaging apps,[1] this is undesirable. Motivated by these applications, Fanti *et al.* [9] asked whether it is possible to devise messaging protocols that can *obfuscate* the rumor source, while at the same time still spreading information widely and quickly.

They devised a family of messaging protocols, termed *adaptive diffusions*, for this purpose; see Section I-A for a detailed description. Their main result shows that a specific messaging protocol within this family achieves *perfect obfuscation:* under this spreading model a "snapshot adversary" can do no better than randomly guessing the source node:

$$\mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^* \,|\, N_t = n) = \frac{1 + o(1)}{n}. \tag{1}$$

Many results extend to more general settings such as irregular trees [10], [11].

Our results. We study the source obfuscation guarantees that adaptive diffusion protocols can provide, in a couple of settings. First, we do this in the context of the adversary having multiple independent observations. We show that when an adversary has access to two independent observations then a

Manuscript received June 26, 2020; revised August 26, 2020; accepted August 31, 2020. Date of publication September 7, 2020; date of current version March 17, 2021. The work of Miklos Z. Rácz was supported in part by NSF under Grant DMS 1811724 and in part by a Princeton SEAS Innovation Award. Recommended for acceptance by Prof. Xianbin Cao. *(Corresponding author: Miklos Racz.)*

Miklós Z. Rácz is with Princeton University, Princeton, NJ 08544 USA (e-mail: mracz@princeton.edu).

Jacob Richey is with the University of Washington, Seattle, WA 98195 USA (e-mail: jfrichey@uw.edu).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSE.2020.3022621, provided by the authors.

Digital Object Identifier 10.1109/TNSE.2020.3022621

[1] Examples include Whisper [5] Blind [6], and the now-defunct Yik Yak [7] and Secret [8].

weak form of obfuscation is still possible. However, when it has access to three or more independent snapshots, then source detection with constant probability is always possible, regardless of the adaptive diffusion protocol.

We also do this in the context of spreading information *locally* around the source. We introduce a natural quantitative measure of local spreading, and characterize the tradeoff between local spreading and source obfuscation for adaptive diffusion protocols (under a single snapshot).

Put together, these results raise questions about the robustness of possible anonymity guarantees when spreading information in social networks. In order to precisely state our results, we first describe in Section I-A the setting of information diffusion processes in general and adaptive diffusions in particular. We then state our results in Sections I-B and I-C.

A. Information Diffusion and Adaptive Diffusion

We define a (discrete time) information diffusion process on a graph $G = (V, E)$ as a (potentially random) increasing sequence of subgraphs $G_0 \subseteq G_1 \subseteq G_2 \subseteq \dots$, where $G_t = (V_t, E_t)$ is the subgraph induced by the vertices V_t who have the information at time t. Throughout the paper we assume that G_0 consists of a single vertex $v^* \in G$, which we term the *source*. We also assume that the information spreads along the edges of the graph and hence $V_{t+1} \subseteq V_t \cup \partial G_t$, where $\partial G_t := \{v \in V : v \notin V_t, \exists w \in V_t : (v, w) \in E\}$ denotes the (outer) vertex boundary of G_t, consisting of vertices that are not in G_t but which are connected to a vertex in G_t.

A simple example is when every vertex who obtains the information spreads it to all its neighbors in the next time step. In this case $G_t = B_t(v^*)$ for every $t \geq 0$, where $B_r(v) := \{u \in V : \delta_G(u, v) \leq r\}$ denotes the (closed) ball of radius r around vertex $v \in V$ (here δ_G denotes graph distance in G). The SI model mentioned above[2] can be defined inductively as follows: given G_t, let v_{t+1} be a uniformly randomly chosen vertex from ∂G_t and let $V_{t+1} := V_t \cup \{v_{t+1}\}$.

The *source detection* problem is the following: given the underlying graph G, the distribution of the sequence $\{G_t\}_{t\geq 0}$, and a single observation G_t at some time $t > 0$, the goal is to estimate the source v^*. This is also known as the "snapshot adversary" model, since we get to observe G_t, a single snapshot in time.

Adaptive diffusion, introduced by Fanti *et al.* [9], is a family of information diffusion processes designed with *source obfuscation* in mind. We now introduce and define adaptive diffusion on $\mathbb{T}_d =: G$; we refer the reader to [11] for a comprehensive introduction more generally. The notation and definitions that follow match those in [9]–[11].

Adaptive diffusion is defined via an auxiliary process, the path $\{vs_t\}_{t\geq 0}$ of a so-called *virtual source*. This is a time-inhomogeneous Markov chain, which we now define. Initially, the virtual source is the same as the true source: $vs_0 := v^*$. Next, it moves to a uniformly random neighbor of v^*:

$$\mathbb{P}(vs_1 = w) = \frac{1}{d}\mathbf{1}_{\{(w,v^*)\in E\}}.$$

For the remainder of the path, assuming that vs_t is given, vs_{t+1} is defined as follows. If t is odd, then $vs_{t+1} = vs_t$; that is, the virtual source stays put. If t is even, then the virtual source either stays put or it moves to one of its $d-1$ neighbors that it has not visited before; in the latter case, it chooses the neighbor to move to uniformly at random. Note that if the virtual source moves then it moves *away* from the source v^*. The probability of choosing one action or the other is a function of time t and also the distance of vs_t from v^* (hence the name *adaptive*). Specifically, let $h_t := \delta_G(vs_t, v^*)$ denote the graph distance between vs_t and v^*. Then

- with probability $\alpha(t, h_t)$ we have that $vs_{t+1} = vs_t$, that is, the virtual source stays put;
- and with probability $1 - \alpha(t, h_t)$ the virtual source moves to one of its $d-1$ neighbors that it has not visited before, chosen uniformly at random.

The probabilities $\alpha(t, h) \in [0, 1]$, with $t \in \{2, 4, 6, \dots\}$ and $h \in \{1, 2, 3, \dots, t/2\}$, are parameters that fully describe the distribution of the path $\{vs_t\}_{t\geq 0}$ of the virtual source. Each choice of parameters defines a particular Markov chain and thus a particular adaptive diffusion protocol.

Having defined the path of the virtual source, we are now ready to define the associated adaptive diffusion protocol, given $\{vs_t\}_{t\geq 0}$. When t is even, the set of infected nodes is defined as

$$V_t := \{v \in V : \delta_G(v, vs_t) \leq t/2\}.$$

That is, V_t is a ball of radius $t/2$—equivalently, a balanced tree of depth $t/2$—around the virtual source vs_t. For t odd, the set of infected nodes V_t is chosen so that $\{G_t\}_{t\geq 0}$ satisfies $V_{t+1} \subseteq V_t \cup \partial G_t$.[3] The resulting information diffusion process is called an *adaptive diffusion*; see Figure 1 for an illustration. Note that by construction adaptive diffusion spreads the information to $N_t \asymp (d-1)^{t/2}$ nodes at time t, which is only a factor of two slower than the fastest possible spread.

Fanti *et al.* [9] show that a particular adaptive diffusion protocol—specifically, the process with $\alpha(t, h) := ((d-1)^{t/2-h+1} - 1)/((d-1)^{t/2+1} - 1)$—perfectly obfuscates the source from an adversary who sees a snapshot of a single diffusion. The key property of this construction is that, for t even, all vertices in $V_t \setminus \{vs_t\}$ are equally likely to be the original source v^* and hence an adversary can do no better than randomly guess among them. A similar statement holds also for t odd, showing that the MLE satisfies (1).

B. Results: Adaptive Diffusion With Multiple Independent Observations

In many applications it is common for individuals to send not just one but multiple messages over time, each one

[2] Note that the SI model is often defined in continuous time. Viewing this continuous-time process at the times when a new vertex obtains the information, we obtain the described discrete time information diffusion process.

[3] Specifically, we have the following. First, $V_1 := \{vs_0, vs_1\}$. Next, for $t \geq 3$ such that t is odd, we distinguish two cases. If $vs_t = vs_{t-1}$, then $V_t := V_{t-1}$; that is, if the virtual source stays put (instead of moving), then the set of infected nodes is unchanged. If $vs_t \neq vs_{t-1}$, then $V_t := V_{t-1} \cup \{w \in \partial G_{t-1} : \delta_G(w, vs_t) = (t-1)/2\}$; in other words, if the virtual source moves, then the information is spread in the same direction.

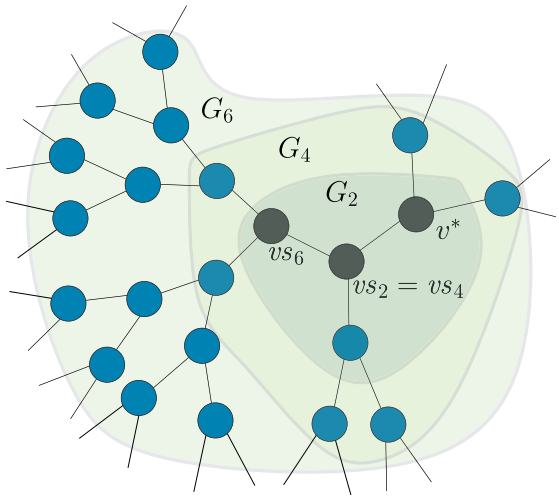


Fig. 1. An example of an adaptive diffusion spreading on the infinite 3-regular tree $\mathbb{T}_3$. Here the virtual source moves to a uniformly randomly chosen neighbor of the source v^* at time 1, then it stays put for several time steps, and moves again at time 5. The shaded regions show the infected subgraphs G_t for $t \in \{2, 4, 6\}$; note that they are all balanced trees of depth $t/2$, centered at the virtual source vs_t.

spreading over the same underlying network. If an adversary has access to a snapshot of each such diffusion, then they are in a much better position to find the source. Is it still possible to obfuscate the source with some form of information diffusion? We investigate this question in the context of adaptive diffusion protocols.

We show that when an adversary has access to two independent observations, a weak form of obfuscation is still possible with adaptive diffusion. However, when three or more independent observations are available, detection with constant probability is always possible, regardless of which adaptive diffusion protocol is used. This is the content of Theorems 1 and 2.

Theorem 1 (Two independent observations): Suppose that information is spread according to an adaptive diffusion protocol on $\mathbb{T}_d$, $d \geq 3$, and that an adversary has two independent observations of infected subgraphs, $G^1_{t_1}$ and $G^2_{t_2}$, started from a fixed source v^*.

(1) There exists a computationally efficient estimator $\widehat{v}$, which is agnostic to the adaptive diffusion protocol, such that if $t_1, t_2 \geq 2$ then

$$\mathbb{P}(\widehat{v} = v^*) \geq \frac{d-1}{d} \cdot \frac{2}{\min\{t_1, t_2\}}.$$

(2) There exists an adaptive diffusion protocol such that the maximum likelihood estimator $\widehat{v}_{\mathrm{ML}}$ satisfies for all $t_1, t_2 \geq 1$ that

$$\mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^*) \leq \frac{d-1}{d} \cdot \frac{7}{\min\{t_1, t_2\}}. \tag{2}$$

A few comments are in order. First, the bounds in parts (1) and (2) above match up to a small constant factor, hence this is best possible within the family of adaptive diffusions. Next, the detection probability in (2) still vanishes as $t = \min\{t_1, t_2\} \to \infty$, but only very slowly—exponentially more slowly than in the case of one observation (see (1) and recall that $N_t \asymp (d-1)^{t/2}$ is exponential in t). We also note that the adaptive diffusion protocol in part (2) is different from the one used by Fanti *et al.* [9] to achieve perfect obfuscation in the case of a single observation; in fact, if this latter adaptive diffusion protocol is used to independently spread two diffusions, then the estimator $\widehat{v}$ in part (1) succeeds at finding the source with constant probability. Finally, we mention that the estimator in part (1) is essentially the same as the MLE in part (2) when t_1 and t_2 are both even—see Section III for details.

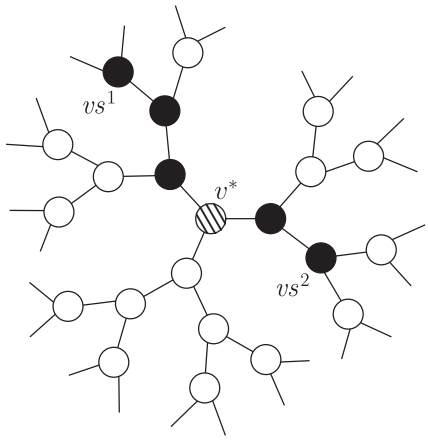


Fig. 2. Detecting the source from two observations. If the two virtual sources are in different subtrees, then the path connecting them contains the source v^*.

Figure 2 illustrates the basic idea behind the estimator in part (1) of Theorem 1; we refer to Section III for details.

Once the adversary has three independent observations, not even weak obfuscation is possible with adaptive diffusion. In fact, the detection probability converges to one exponentially quickly in the number of observations (see (3) below), extending the results of Wang *et al.* [4] for the SI model to the family of adaptive diffusions.

Theorem 2 (Three or more independent observations): Suppose that information is spread according to an adaptive diffusion protocol on $\mathbb{T}_d$, $d \geq 3$, and that an adversary has $k \geq 3$ independent observations of infected subgraphs, $G^i_{t_i}$ for $i \in \{1, \ldots, k\}$, started from a fixed source v^*.

When $k = 3$, there is a computationally efficient estimator $\widehat{v}$, which is agnostic to the adaptive diffusion protocol, satisfying

$$\mathbb{P}(\widehat{v} = v^*) \geq \frac{(d-1)(d-2)}{d^2}.$$

More generally, there exists a computationally efficient estimator $\widehat{w} = \widehat{w}(k)$, which is agnostic to the adaptive diffusion protocol, such that

$$\mathbb{P}(\widehat{w} = v^*) \geq 1 - d \times \exp\left(-\frac{(d-2)^2}{2\,d^2}k\right). \tag{3}$$

Theorem 2 follows from basic symmetry properties of adaptive diffusion; the basic idea is illustrated in Figure 3 (see Section II for further details). Comparing Figure 2 and Figure 3

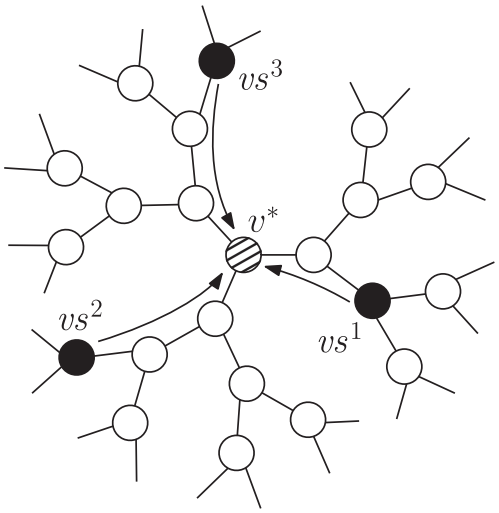


Fig. 3. Detecting the source from three observations. If the three virtual sources are in different subtrees, then the paths connecting them intersect in a single node: the source v^*.

provides intuition into why the dramatic shift from two snapshots to three snapshots occurs.

We also note that Theorem 2 extends, with essentially the same proof, to adaptive diffusions on any irregular tree with minimum degree 3. This is because the proof in Section II is based on basic symmetry properties; we leave the details to the reader.

C. Results: Local Spreading vs. Source Obfuscation

It is often desirable to not only spread information widely and quickly, but also to spread it *locally* around the source. Indeed, the local neighborhood of the source typically consists of nodes that are closely related to the source, and the information that the source is spreading is often most relevant to this local neighborhood. In particular, this is true for scenarios where source obfuscation is relevant and important, for instance, spreading information about a local protest. At the same time, local spreading is at odds with source obfuscation. Here we introduce a natural way to quantify local spreading, and characterize the tradeoff between local spreading and source obfuscation for adaptive diffusion protocols (under a single snapshot).

Formally, define for an adaptive diffusion the quantity

$$R_t := \max\{r \geq 0 : B_r(v^*) \subseteq G_t\}.$$

In words, R_t is the radius of the largest ball of infected nodes centered at the rumor source at time t. Since R_t is (in general) a random quantity, we may use $\mathbb{E}[R_t]$ as a deterministic measure of local spreading of an adaptive diffusion protocol. Observe that $0 \leq R_t \leq t/2$ and hence also $0 \leq \mathbb{E}[R_t] \leq t/2$.

Ideally for local spreading we would like $\mathbb{E}[R_t]$ to grow linearly with t; at the very least, local spreading requires $\mathbb{E}[R_t] \to \infty$ as $t \to \infty$. However, the adaptive diffusion protocol that achieves perfect source obfuscation (see the end of Section I-A) does not have local spreading: in fact, $\mathbb{E}[R_t] \leq 1$ for all t and, moreover, $\sup_t R_t$ is finite almost surely.

This shows that source obfuscation guarantees have to be relaxed in order to have local spreading. It turns out that it is still possible to have reasonable source obfuscation guarantees—we refer to this as "polynomial obfuscation," see (4) below—and local spreading at the same time. The following theorem characterizes this tradeoff for adaptive diffusion protocols (under a single snapshot). For simplicity, we focus here on even times t.

Theorem 3 (Tradeoff between local spreading and source obfuscation): Suppose that information is spread according to an adaptive diffusion protocol on $\mathbb{T}_d$, $d \geq 3$, and that an adversary observes, at an even time t, an infected subgraph, G_t, started from a fixed source v^*.

(1) Suppose that the adaptive diffusion protocol achieves "polynomial obfuscation," that is, the following holds:

$$\mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^*) \leq \frac{C}{N_t^{\gamma}} \tag{4}$$

for some $\gamma \in (0,1)$ and $C < \infty$, where recall that

$$N_t = |V_t| = \frac{d}{d-2}\Big((d-1)^{t/2}-1\Big) + 1 \asymp (d-1)^{t/2}.$$

Then

$$\mathbb{E}[R_t] \leq (1-\gamma)\frac{t}{2} + \frac{\log(Ct)}{\log(d-1)} + 2.$$

(2) For every $\gamma \in (0,1)$ there exists an adaptive diffusion protocol that satisfies (4) with $C = 2(d-1)$ and also

$$\mathbb{E}[R_t] \geq (1-\gamma)\frac{t}{2}$$

for all even $t > 2/\gamma$ (and for all even $t \leq 2/\gamma$ we have $\mathbb{E}[R_t] = t/2 - 1$).

In particular, we see from Theorem 3 that the power γ in polynomial obfuscation (see (4)) and the speed $(1-\gamma)/2$ of local spreading are directly related. This precisely quantifies the tradeoff between local spreading and source obfuscation guarantees: the faster local spreading is—that is, the smaller γ is—the weaker the source obfuscation guarantee.

D. Organization

The rest of the paper is organized as follows. We first prove Theorem 2 in Section II, since the proof relies only on a simple symmetry property of adaptive diffusion protocols on $\mathbb{T}_d$ and provides good intuition for the subsequent proofs. We then prove Theorem 1 in Section III; the proof of part (1) is similar to the proof of Theorem 2 in Section II, while the proof of part (2) requires understanding the maximum likelihood estimator in the case of two observations. (Some cases in the proof of part (2) of Theorem 1 are deferred to the supplementary material.) In Section IV we turn to studying local spreading and prove Theorem 3. Finally, we conclude in Section V by discussing some implications and limitations of our results,

how they relate to other works, as well as further questions for future research.

II. Adaptive Diffusion With $k \geq 3$ Independent Observations

In this section we prove Theorem 2. The main idea is simple and relies on a symmetry property of adaptive diffusion protocols on $\mathbb{T}_d$: that they send the virtual source in a uniformly random direction. First, note that if we remove the source v^* from the tree $\mathbb{T}_d$ then it breaks into d subtrees. The main observation is that the virtual source of an adaptive diffusion is equally likely to be in each subtree. This symmetry property alone guarantees a constant probability of detection when there are at least three independent observations, as we now explain.

Assume for now that $t_1, \ldots, t_k$ are even; the proof is cleaner in this case, though not much changes in the general case. Recall that for an adaptive diffusion protocol the infected tree G_t is a ball with center vs_t when t is even. Hence from the infected tree G_t we may determine the virtual source vs_t. We may thus assume that the adversary is given k independent virtual sources $vs^1, vs^2, \ldots, vs^k$ (the time stamps of the virtual sources are not relevant for what follows). The main observation is that if vs^1, vs^2, and vs^3 are in different subtrees, then v^* is the unique vertex at the intersection of the three shortest paths connecting vs^1 and vs^2, vs^1 and vs^3, and vs^2 and vs^3; see Figure 3 for an illustration.

This immediately leads to a source detection algorithm: if the three shortest paths connecting vs^1 and vs^2, vs^1 and vs^3, and vs^2 and vs^3 intersect at a single vertex, the algorithm outputs this vertex; if not, pick a vertex from the intersection uniformly at random. Since each virtual source is equally likely to be in each subtree, there is a constant probability that vs^1, vs^2, and vs^3 are in different subtrees and therefore the algorithm successfully detects the source.

The proof that follows makes this formal and also presents an improved algorithm when the number of observations k is large, in order to show that the detection probability goes to 1 as $k \to \infty$.

Proof of Theorem 2: We start with some notational preliminaries. For distinct nodes $x, y \in \mathbb{T}_d$, let T_x^y denote the subtree of $\mathbb{T}_d$ away from y in the direction of x. In other words, if y were removed from $\mathbb{T}_d$ then the tree would break into a forest of d trees and T_x^y is the tree that contains x. Formally, if n_x^y is the neighbor of y that is closest to x, then

$$T_x^y := \big\{ z \in \mathbb{T}_d : \delta\big(z, n_x^y\big) \, < \, \delta(z, y) \big\}.$$

We first assume, for simplicity, that $t_1, \ldots, t_k$ are all even; this simplifies the proof and we explain at the end what changes if some of these times are odd. Then for every $i \in \{1, \ldots, k\}$ we have that $G_{t_i}^i$ is a ball (of radius $t_i/2$) with center $vs_{t_i}^i$, the virtual source at time t_i. Thus we may assume that the adversary observes $k \geq 3$ independent virtual sources $vs^1, \ldots, vs^k \in \mathbb{T}_d$; as the time indices do not play a role in what follows, we drop them for notational convenience. We first define an estimator $\widehat{v}$ using only the first three samples (vs^1, vs^2, and vs^3) and show that it detects the source with constant probability. For $i, j \in \{1, 2, 3\}$ let P_{ij} denote the set of vertices in the unique path in $\mathbb{T}_d$ between vs^i and vs^j. If the three paths P_{12}, P_{13}, and P_{23} intersect in a single vertex, let $\widehat{v}$ be this vertex. If the intersection of P_{12}, P_{13}, and P_{23} contains more than one vertex, let $\widehat{v}$ pick a vertex from this intersection uniformly at random.

Consider the event A where the three virtual sources take different first steps away from the source. By the construction of adaptive diffusion, this is the same as the virtual sources being in different subtrees for all positive times; that is,

$$A = \Big\{ T_{vs^1}^{v^*} \cap T_{vs^2}^{v^*} = \emptyset \Big\} \cap \Big\{ T_{vs^1}^{v^*} \cap T_{vs^3}^{v^*} = \emptyset \Big\} \cap \Big\{ T_{vs^2}^{v^*} \cap T_{vs^3}^{v^*} = \emptyset \Big\}.$$

On the event A we have that $P_{12} \cap P_{13} \cap P_{23} = \{v^*\}$ and, hence, $\widehat{v} = v^*$. That is, on the event A, the estimator correctly detects the source of the diffusion. Since the direction of the first step of a virtual source is uniformly random among the d choices and the different samples are independent, we have that $\mathbb{P}(A) = \frac{(d-1)(d-2)}{d^2}$, which concludes this part of the proof.

We now explain how more samples can be used to achieve a detection probability that converges to 1 exponentially in k as $k \to \infty$. For any vertex $v \in \mathbb{T}_d$ and w a neighbor of v, define

$$N_w(v) := \#\big\{ j \in [k] : vs^j \in T_w^v \big\}.$$

That is, $N_w(v)$ counts the number of virtual sources in the subtree of $\mathbb{T}_d$ away from v in the direction of w. Using these quantities we define the following estimator:

$$\widehat{w} := \arg\min_{v \in \mathbb{T}_d} \max\nolimits_{w:(w,v)\in E} N_w(v), \tag{5}$$

provided that this is well-defined (i.e., the minimum is attained at a single vertex); if this is not well-defined, let $\widehat{w}$ be an arbitrary vertex. Let $w_1, \ldots, w_d$ denote the neighbors of v^* in $\mathbb{T}_d$ and let $Y := (N_{w_1}(v^*), \ldots, N_{w_d}(v^*))$. We now argue that if $\|Y\|_\infty < k/2$, then $\widehat{w} = v^*$, that is, the estimator correctly detects the source of the diffusion.

First, observe that $\max_{w:(w,v^*)\in E} N_w(v^*) = \|Y\|_\infty$, which is less than $k/2$ under the assumption. Second, if $v \neq v^*$, then there must exist w' a neighbor of v and $i \in [d]$ such that

$$T_{w'}^v \supseteq \bigcup_{j \in [d]\setminus\{i\}} T_{w_j}^{v^*}.$$

This implies that

$$\begin{aligned} N_{w'}(v) &\geq \sum_{j \in [d]\setminus\{i\}} N_{w_j}(v^*) = k - N_{w_i}(v^*) \\ &\geq k - \|Y\|_\infty > \, k/2, \end{aligned}$$

where we used that $\|Y\|_1 = k$, as well as the assumption that $\|Y\|_\infty < k/2$. Consequently, $\max_{w:(w,v)\in E} N_w(v) \geq N_{w'}(v) > k/2$ and, hence, $\widehat{w} \neq v$. We have thus shown that $\|Y\|_\infty < k/2$ implies that $\widehat{w} = v^*$.

To conclude, we estimate from below the probability that $\|Y\|_\infty < k/2$, or rather, we estimate from above the complimentary event that $\|Y\|_\infty \geq k/2$. First, by a union bound and symmetry we have that $\mathbb{P}(\|Y\|_\infty \geq k/2) \leq d \times \mathbb{P}(N_{w_1}(v^*) \geq k/2)$. Now since $N_{w_1}(v^*) \sim \mathrm{Bin}(k, 1/d)$, we have by a Chernoff bound that

$$\mathbb{P}\big(N_{w_1}(v^*) \geq k/2\big) = \mathbb{P}\bigg(N_{w_1}(v^*) - \mathbb{E}\big[N_{w_1}(v^*)\big] \geq \frac{d-2}{2\ d} k\bigg) \leq \exp\bigg(-\frac{(d-2)^2}{2\ d^2} k\bigg).$$

Finally, we return to our simplifying assumption that the observation times $t_1, \ldots, t_k$ are all even. If t_i is odd, then there are two cases. If $G^i_{t_i}$ is a ball, then it is a ball with center $vs^i_{t_i}$, so the adversary can again determine the virtual source at time t_i and everything is unchanged. If $G^i_{t_i}$ is not a ball, then it is symmetric about the edge connecting $vs^i_{t_i-1}$ and $vs^i_{t_i}$. Thus the adversary can determine the set $\{vs^i_{t_i-1}, vs^i_{t_i}\}$. Picking either element of the set as the virtual source, the remainder of the proof goes through unchanged. ■

At first glance, it may appear that computing the estimator $\widehat{w}$ requires solving a minimization problem over the entire infinite tree $\mathbb{T}_d$, but this is not the case. For every vertex v that is not on a shortest path between two virtual sources we have that $\max_{w:(w,v)\in E} N_w(v) = k$ and therefore $\widehat{w}$ must lie on a shortest path between two virtual sources. Moreover, the distance between any two virtual sources is at most $2\max_{i\in[k]} t_i$. Thus the minimization problem in (5) is over a set of size $O(k^2 \max_{i\in[k]} t_i)$. For each node v in this set, one can efficiently compute $\max_{w:(w,v)\in E} N_w(v)$ as follows. For every virtual source vs^j, connect v and vs^j, and let w be the neighbor of v on this path. We then have that $vs^j \in T^v_w$. By doing this for every virtual source, we can compute the quantities $\{N_w(v)\}_{w:(w,v)\in E}$ and hence also the quantity $\max_{w:(w,v)\in E} N_w(v)$. In short, the estimator $\widehat{w}$ can be computed efficiently.

III. Adaptive Diffusion With Two Independent Observations

In this section we prove Theorem 1. We start with the proof of part (1) in Section III-A, which builds on similar ideas as the proof of Theorem 2 in Section II. Then, in order to prove part (2) of Theorem 1, we need to understand the maximum likelihood estimator—this is done in Section III-B. Due to the nature of adaptive diffusion, we have to deal with even and odd times separately. To focus on the key insights and computations, we first prove Theorem 1(2) when t_1 and t_2 are both even—this is in Section III-C. The cases when one or both of t_1 and t_2 are odd are similar but more complicated, while not adding anything conceptually—hence we defer the proof in these cases to the supplementary material.

A. Source Detection

The proof of Theorem 1(1) builds on similar ideas as the proof of Theorem 2 in Section II. Recall the notation that we introduced in Section II, which we use here.

Proof of Theorem 1(1): Assume first that t_1 and t_2 are even; this simplifies the proof and we explain at the end what changes if either time is odd. Then for $i \in \{1, 2\}$ we have that $G^i_{t_i}$ is a ball of radius $t_i/2$ with center $vs^i_{t_i}$. The adversary can thus determine the two virtual sources $vs^1 \equiv vs^1_{t_1}$ and $vs^2 \equiv vs^2_{t_2}$.

By definition we always have that $v^* \in V^1_{t_1} \cap V^2_{t_2}$, that is, the source v^* is contained in both sets of infected nodes. Let P_{12} denote the set of vertices that are on the path in $\mathbb{T}_d$ between vs^1 and vs^2, excluding vs^1 and vs^2. Furthermore, define the set $S := P_{12} \cap V^1_{t_1} \cap V^2_{t_2}$. Let A_{12} denote the event that vs^1 and vs^2 are in different subtrees away from v^*; that is,

$$A_{12} := \Big\{T^{v^*}_{vs^1} \cap T^{v^*}_{vs^2} = \emptyset\Big\}. \tag{6}$$

Since the two diffusions are independent and the first step of the virtual source is to a uniformly random neighbor of v^*, we have that $\mathbb{P}(A_{12}) = (d-1)/d$. The main observation is that, on the event A_{12}, we have that $v^* \in P_{12}$; see Figure 2 for an illustration.[4] Consequently, on the event A_{12} we also have that $v^* \in S$.

This suggests a natural estimator: if $S \neq \emptyset$, let $\widehat{v}$ be a uniformly randomly chosen node from S (note that S is a measurable function of $G^1_{t_1}$ and $G^2_{t_2}$); if $S = \emptyset$ (this occurs when $\delta(vs^1, vs^2) \leq 1$), let $\widehat{v}$ be an arbitrary node.[5] Then, given A_{12} and S, the conditional probability that $\widehat{v} = v^*$ is $1/|S|$ (note that A_{12} implies that $|S| \geq 1$, as we argued above). We have thus shown that

$$\mathbb{P}(\widehat{v} = v^*) \geq \mathbb{P}(\widehat{v} = v^* \,|\, A_{12})\mathbb{P}(A_{12}) = \mathbb{E}[1/|S| \,|\, A_{12}]\frac{d-1}{d}.$$

To conclude, it suffices to show that $|S| \leq \min\{t_1/2, t_2/2\}$ whenever A_{12} holds. To see this, note that the intersection $P_{12} \cap V^1_{t_1}$ contains at most $t_1/2$ nodes, since $G^1_{t_1}$ is a (closed) ball of radius $t_1/2$ centered at vs^1, the path P_{12} starts at the virtual source vs^1, and vs^1 is not included in P_{12}. Thus $|S| \leq |P_{12} \cap V^1_{t_1}| \leq t_1/2$. Similarly, $P_{12} \cap V^2_{t_2}$ contains at most $t_2/2$ nodes, and the claim follows.

Finally, we explain what changes when t_i is odd for $i = 1$ and/or $i = 2$. If $G^i_{t_i}$ is a ball, then its center is $vs^i_{t_i}$, so the adversary can again determine the virtual source at time t_i and everything is unchanged. If $G^i_{t_i}$ is not a ball, then it is symmetric about the edge connecting vs_{t_i-1} and vs_{t_i}. Thus the adversary can determine the set $\{vs_{t_i-1}, vs_{t_i}\}$. Connecting *both* of these virtual sources with the other virtual source(s), we again obtain a path, where now at both ends of the path we have either one or two virtual sources. In any case, we can define P_{12} analogously, where again the known virtual sources are

[4] The two virtual sources can indeed be excluded from P_{12} and we still have that $v^* \in P_{12}$ on the event A_{12}. This is because the virtual source can never be the true source, by construction. This assumes that $t_1, t_2 \geq 1$—which holds, since we assume in the proof that $t_1, t_2 \geq 2$. In any case, if $\min\{t_1, t_2\} < 2$, then one of the observed snapshots contains at most two vertices, so a random guess succeeds in identifying the source with probability at least $1/2$.

[5] We note that the two virtual sources, vs^1 and vs^2, can be determined efficiently, and thus so can S, and hence also the estimator $\widehat{v}$.

not considered as part of P_{12}. The rest of the proof is unchanged. ■

B. Maximum Likelihood Source Estimation

In order to prove Theorem 1(2), we need to understand maximum likelihood source estimation. Here we discuss this for adaptive diffusions in general. Recall that an adaptive diffusion protocol is given by the probabilities $\alpha(t,h) \in [0,1]$, with $t \in \{2,4,6,\ldots\}$ and $h \in \{1,2,3,\ldots,t/2\}$, which determine the distribution of the path of the virtual source $\{vs_t\}_{t\geq 0}$. Let $h_t := \delta(vs_t, v^*)$ denote the graph distance between vs_t and v^*, and let $p(t,h) := \mathbb{P}(h_t = h)$ denote the distribution of h_t.

When determining the likelihood function $L(v) = \mathbb{P}(G_t \mid v^* = v)$ we have to specify whether the value of t is known or not (since it is not always possible to infer the value of t from the observation G_t). We assume in the following that t is *known*. Knowing t can only help the adversary and hence any upper bounds on the success probability of the MLE under this assumption still hold without this assumption. Furthermore, the rumor source detection results (Theorem 1(1) and Theorem 2) hold regardless of whether we assume this or not. Finally, this assumption is also what is used in previous works [9]–[11].

We now determine the likelihood function $L(v) = \mathbb{P}(G_t \mid v^* = v)$ for even t; it is similar for odd t, but we leave this for later. First, we always have that $v^* \in V_t \setminus \{vs_t\}$, so $L(v) = 0$ if $v \notin V_t \setminus \{vs_t\}$. Next, since G_t is a ball of radius $t/2$ with center vs_t, it is fully determined by the position of the virtual source, together with the time t. It is important to note a key symmetry property of adaptive diffusion: all nodes at a particular distance from the virtual source are equally likely to have been the source. This is because the virtual source always moves to a uniformly randomly chosen neighbor away from the source. Thus the distribution of the virtual source is completely determined by the distribution of h_t. Altogether, since there are $d(d-1)^{h-1}$ nodes at distance $h \geq 1$ from a particular vertex, we obtain that

$$L(v) = \frac{1}{d(d-1)^{\delta(v,vs_t)-1}} p(t, \delta(v, vs_t)) \mathbf{1}_{\{v \in V_t \setminus \{vs_t\}\}}. \tag{7}$$

Now assume that we have k independent observations of infected subgraphs, $G^i_{t_i} = (V^i_{t_i}, E^i_{t_i})$ for $i \in \{1,\ldots,k\}$, started from a fixed source v^*. Assume also, for now, that all the times $t_1,\ldots,t_k$ are even. Then, by independence, the likelihood function is

$$L(v) = \left(\frac{d-1}{d}\right)^k \times \\ \times \prod_{i=1}^{k} p(t_i, X_i(v)) \cdot (d-1)^{-X_i(v)} \mathbf{1}_{\left\{v \in \bigcap_{i=1}^k \left(V^i_{t_i} \setminus \left\{vs^i_{t_i}\right\}\right)\right\}},$$

where we have introduced

$$X_i(v) := \delta\left(v, vs^i_{t_i}\right) \tag{8}$$

for convenience (and recall that we can determine $vs^i_{t_i}$, and thus also $X_i(v)$, from $G^i_{t_i}$). By taking logarithms, we obtain that the MLE satisfies

$$\widehat{v}_{\mathrm{ML}} \in \\ \underset{v \in \bigcap_{i=1}^k \left(V^i_{t_i} \setminus \left\{vs^i_{t_i}\right\}\right)}{\arg\max} \sum_{i=1}^{k} \{\log p(t_i, X_i(v)) - X_i(v) \log(d-1)\}. \tag{9}$$

We now turn to determining the likelihood function $L(v) = \mathbb{P}(G_t \mid v^* = v)$ for odd t. This is similar to the case of even t, but there are slight differences. Specifically, there are two cases to distinguish: when t is odd, the observed graph G_t is either a ball or it is not (in which case it consists of two balanced rooted trees of depth $(t-1)/2$, whose roots are connected by an edge).

The former case occurs when the virtual source does not move at time $t-1$, that is, when $vs_{t-1} = vs_t$. In this case, we know that $G_{t-1} = G_t$, we know the likelihood of G_{t-1} (which is given by (7) with t replaced by $t-1$), and in order to obtain the likelihood of G_t we have to multiply this by the probability that $vs_{t-1} = vs_t$, which is $\alpha(t-1, X(v))$, where $X(v) = \delta(v, vs_{t-1}) = \delta(v, vs_t)$.

In the latter case, when G_t is not a ball, we know that the virtual source moved at time $t-1$. Furthermore, we can determine the set $\{vs_{t-1}, vs_t\}$, as these two vertices are connected by the central edge of G_t. In this case, we define $X(v) := \min\{\delta(v, vs_{t-1}), \delta(v, vs_t)\}$ (note that $X(v)$ can be determined from G_t). In order to obtain the likelihood of G_t we have to multiply the expression in (7) (with t replaced by $t-1$ and $\delta(v, vs_t)$ replaced with $\min\{\delta(v, vs_{t-1}), \delta(v, vs_t)\}$) with the probability that $vs_{t-1} \neq vs_t$, which is $1 - \alpha(t-1, X(v))$.

Altogether, when t is odd we have that the likelihood function is

$$L(v) = \frac{1}{d(d-1)^{X(v)-1}} p(t-1, X(v)) \\ \times \alpha(t-1, X(v)) \mathbf{1}_{\{v \in V_t \setminus \{vs_t\}\}} \tag{10}$$

if G_t is a ball, and

$$L(v) = \frac{1}{d(d-1)^{X(v)-1}} p(t-1, X(v)) \\ \times \{1 - \alpha(t-1, X(v))\} \mathbf{1}_{\{v \in V_t \setminus \{vs_{t-1}, vs_t\}\}}$$

otherwise. Here $X(v) := \min\{\delta(v, vs_{t-1}), \delta(v, vs_t)\}$ (note that this definition of $X(v)$ works for both cases; when G_t is a ball then $vs_{t-1} = vs_t$ and hence $X(v) = \delta(v, vs_{t-1}) = \delta(v, vs_t)$).

C. Source Obfuscation — Even Times

We are now ready to prove Theorem 1(2). We first prove this when both t_1 and t_2 are even. This is done in order to highlight the key insights and computations. The remaining cases (when one or both of t_1 and t_2 are odd) are similar but more complicated and hence are deferred to the supplementary material.

Proof of Theorem 1(2) when t_1 and t_2 are both even: We may assume in the following that $t_1, t_2 \geq 4$, since when $\min\{t_1, t_2\} = 2$ then the right hand side of (2) is greater than 1 and thus the statement is vacuously true.

Consider the adaptive diffusion protocol—which we term the *uniform protocol* $\mathcal{U}$ for reasons to become clear—given by the probabilities

$$\alpha_{\mathcal{U}}(t,h) := \frac{t-2h+2}{t+2} \tag{11}$$

for $t \in \{2,4,6,\ldots\}$ and $h \in \{1,2,\ldots,t/2\}$. This is the same protocol introduced by Fanti *et al.* [9] to achieve perfect obfuscation from a single snapshot on $\mathbb{Z}$—the difference is that here we use this protocol regardless of the degree d. The important property of this protocol is that the distance $h_t := \delta(vs_t, v^*)$ between the virtual source vs_t and the true source v^* is *uniformly distributed* over the set of possible values $\{1,2,\ldots,t/2\}$, for all even t. That is, for all even t we have that

$$p_{\mathcal{U}}(t,h) = \frac{2}{t}\mathbf{1}_{\{h\in\{1,2,\ldots,t/2\}\}}. \tag{12}$$

This can be shown by induction; we leave the details to the reader.

We now turn to analyzing the maximum likelihood estimator of the source, $\widehat{v}_{\mathrm{ML}}$, given two independent snapshots $G^1_{t_1}$ and $G^2_{t_2}$. Recall that we assume now that t_1 and t_2 are both even. The adversary can thus determine the two virtual sources $vs^1 \equiv vs^1_{t_1}$ and $vs^2 \equiv vs^2_{t_2}$. By plugging in (12) into (9), we obtain that the MLE satisfies

$$\widehat{v}_{\mathrm{ML}} \in \operatorname*{arg\,min}_{v\in V^1_{t_1}\cap V^2_{t_2}\setminus\{vs^1,vs^2\}} (X_1(v)+X_2(v)),$$

where recall from (8) that $X_i(v) = \delta(v, vs^i)$ for $i \in \{1,2\}$. In words, the maximum likelihood estimator minimizes the sum of the distances to the two virtual sources, over all nodes that were infected in both diffusions, excluding the two virtual sources.

To understand the MLE better we distinguish three cases, the last one being the most important:

(1) If $vs^1 = vs^2$, then $\widehat{v}_{\mathrm{ML}}$ chooses a neighbor of $vs^1 = vs^2$ uniformly at random.
(2) If $\delta(vs^1, vs^2) = 1$, then $\widehat{v}_{\mathrm{ML}}$ chooses a neighbor of the set $\{vs^1, vs^2\}$ uniformly at random.[6]
(3) If $\delta(vs^1, vs^2) \geq 2$, then $X_1(v) + X_2(v)$ is minimized when v is on the shortest path between vs^1 and vs^2. Let P_{12} denote the set of vertices that are on the shortest path between vs^1 and vs^2, excluding vs^1 and vs^2. Furthermore, define the set $S := P_{12} \cap V^1_{t_1} \cap V^2_{t_2}$ and note that when $\delta(vs^1, vs^2) \geq 2$, then S is nonempty, because the vertex in P_{12} that is closest to v^* is always in S. We have thus argued that the likelihood function is maximized at the nodes in S and thus the maximum likelihood estimator $\widehat{v}_{\mathrm{ML}}$ chooses a node from S uniformly at random.

[6] Here we use that $t_1 t_2 \geq 4$, to ensure that all neighbors of the set $\{vs^1, vs^2\}$ are in $V^1_{t_1} \cap V^2_{t_2}$.

Note that $\widehat{v}_{\mathrm{ML}}$ is (essentially) the same as the estimator $\widehat{v}$ introduced in the proof of part (1) of Theorem 1.

Let A_{12} denote the event that vs^1 and vs^2 are in different subtrees away from v^* (see (6)), and note that $\mathbb{P}(A_{12}) = (d-1)/d$. Observe that if the event A_{12} holds, then necessarily $\delta(vs^1, vs^2) \geq 2$, and hence the first two cases above imply that A_{12} does not hold. To compute the probability that the MLE $\widehat{v}_{\mathrm{ML}}$ is correct, we may condition on whether or not A_{12} holds:

$$\begin{aligned}
&\mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^*) \\
&= \mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^* \mid A_{12})\mathbb{P}(A_{12}) + \mathbb{P}\big(\widehat{v}_{\mathrm{ML}} = v^* \mid A^C_{12}\big)\mathbb{P}\big(A^C_{12}\big) \\
&= \mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^* \mid A_{12}) \cdot \frac{d-1}{d} + \mathbb{P}\big(\widehat{v}_{\mathrm{ML}} = v^* \mid A^C_{12}\big) \cdot \frac{1}{d}.
\end{aligned} \tag{13}$$

Let us now turn to computing $\mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^* \mid A^C_{12})$. There are two cases when the MLE can be correct, given that A_{12} does not hold. First, corresponding to Case (1) above: if $vs^1 = vs^2$ and $\delta(v^*, vs^1) = \delta(v^*, vs^2) = 1$, then the MLE is correct with probability $1/d$. Second, corresponding to Case (2) above: if $\delta(v^*, vs^1) = \delta(vs^1, vs^2) = 1$ or if $\delta(v^*, vs^2) = \delta(vs^1, vs^2) = 1$, then the MLE is correct with probability $1/(2d-2)$. If $\delta(vs^1, vs^2) \geq 2$ and A_{12} does not hold, then $\widehat{v}_{\mathrm{ML}} \neq v^*$. Putting these together and using (12) we obtain that

$$\begin{aligned}
\mathbb{P}\big(\widehat{v}_{\mathrm{ML}} = v^* \mid A^C_{12}\big) &= \frac{2}{t_1}\cdot\frac{2}{t_2}\cdot\frac{1}{d} + 2\cdot\frac{2}{t_1}\cdot\frac{2}{t_2}\cdot\frac{1}{2d-2} \\
&= \frac{4}{t_1 t_2}\left(\frac{1}{d} + \frac{1}{d-1}\right).
\end{aligned} \tag{14}$$

We now turn to computing $\mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^* \mid A_{12})$. Given A_{12} and S, the conditional probability that $\widehat{v}_{\mathrm{ML}} = v^*$ is $1/|S|$. We thus have that

$$\mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^* \mid A_{12}) = \mathbb{E}[1/|S| \mid A_{12}]. \tag{15}$$

On the event A_{12}, we can express $|S|$ as a function of $X_1(v^*)$ and $X_2(v^*)$ as follows. First, the set S always contains v^* when A_{12} holds. Next, there are $X_1(v^*) - 1$ nodes on the path P_{12} between v^* and vs^1. However, only the $t_2/2 - X_2(v^*)$ nodes of these that are closest to v^* are in $V^2_{t_2}$ as well. Similarly, there are $X_2(v^*) - 1$ nodes on the path P_{12} between v^* and vs^2, but only the $t_1/2 - X_1(v^*)$ nodes of these that are closest to v^* are in $V^1_{t_1}$ as well. Altogether, on the event A_{12} we have that

$$\begin{aligned}
|S| = 1 &+ \min\{X_1(v^*) - 1, t_2/2 - X_2(v^*)\} \\
&+ \min\{X_2(v^*) - 1, t_1/2 - X_1(v^*)\}.
\end{aligned} \tag{16}$$

Recall from (12) that $X_i(v^*)$ is uniformly distributed on $\{1,2,\ldots,t_i/2\}$, for $i \in \{1,2\}$. Moreover, $X_1(v^*)$ and $X_2(v^*)$ are independent. Both of these statements hold conditioned on A_{12}. Therefore, plugging in (16) into (15) and writing out the expectation we obtain that

$$\mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^* \,|\, A_{12}) = \frac{1}{st}\sum_{j=1}^{s}\sum_{\ell=1}^{t}\frac{1}{1+\min\{j-1,t-\ell\}+\min\{\ell-1,s-j\}}, \tag{17}$$

where we have introduced $s := \min\{t_1,t_2\}/2$ and $t := \max\{t_1,t_2\}/2$ in order to abbreviate notation. With this notation, we can write $|S|$ from (16) more succintly by breaking things into three cases, as follows:

- If $X_1(v^*) + X_2(v^*) \le s+1$, then $|S| = X_1(v^*) + X_2(v^*) - 1$.
- If $s+1 < X_1(v^*) + X_2(v^*) \le t+1$, then $|S| = s$.
- If $t+1 < X_1(v^*) + X_2(v^*)$, then $|S| = 1 + s + t - (X_1(v^*) + X_2(v^*))$.

Accordingly, we can break the sum in (17) into three parts. Let $\mathcal{I} := \{(j,\ell) : 1 \le j \le s, 1 \le \ell \le t\}$ denote the index set over which we take the sum in (17). We can write it as the disjoint union $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2 \cup \mathcal{I}_3$, where $\mathcal{I}_1 := \{(j,\ell) \in \mathcal{I} : j+\ell \le s+1\}$, $\mathcal{I}_2 := \{(j,\ell) \in \mathcal{I} : s+1 < j+\ell \le t+1\}$, and $\mathcal{I}_3 := \{(j,\ell) \in \mathcal{I} : t+1 < j+\ell\}$. We now consider the index sets $\mathcal{I}_1, \mathcal{I}_2$, and $\mathcal{I}_3$ separately.

First, suppose that $m \in \{2,3,\dots,s+1\}$. There are $m-1$ pairs of indices $(j,\ell) \in \mathcal{I}_1$ such that $j+\ell = m$. For each such index pair, the fraction in (17) is equal to $1/(m-1)$. Since there are s different values of m, the sum over the index set $\mathcal{I}_1$ is equal to s.

Next, observe that $|\mathcal{I}_2| = s(t-s)$. For every $(j,\ell) \in \mathcal{I}_2$, the fraction in (17) is $1/s$. Therefore the sum over the index set $\mathcal{I}_2$ is equal to $s(t-s)/s = t-s$.

Finally, suppose that $m \in \{t+2,\dots,t+s\}$. There are $1+s+t-m$ pairs of indices $(j,\ell) \in \mathcal{I}_3$ such that $j+\ell = m$. For each such index pair, the fraction in (17) is equal to $1/(1+s+t-m)$. Since there are $s-1$ different values of m, the sum over the index set $\mathcal{I}_3$ is equal to $s-1$.

Putting together the previous three paragraphs, we thus have that

$$\sum_{j=1}^{s}\sum_{\ell=1}^{t}\frac{1}{1+\min\{j-1,t-\ell\}+\min\{\ell-1,s-j\}} = s+t-1.$$

Plugging this back into (17), and returning to the notation of t_1 and t_2, we obtain that

$$\mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^* \,|\, A_{12}) = \frac{s+t-1}{st} = \frac{2t_1+2t_2-4}{t_1t_2}. \tag{18}$$

Putting together (13), (14), and (18), we have obtained that

$$\begin{aligned}\mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^*) &= \frac{d-1}{d}\cdot\frac{2t_1+2t_2-4}{t_1t_2} + \frac{1}{d}\cdot\frac{4}{t_1t_2}\left(\frac{1}{d}+\frac{1}{d-1}\right)\\ &< \frac{d-1}{d}\cdot\frac{2t_1+2t_2}{t_1t_2},\end{aligned}$$

where we used that $1/d + 1/(d-1) < 1$. Using that $2t_1 + 2t_2 \le 4\max\{t_1,t_2\}$, we obtain the bound in (2), when t_1 and t_2 are both even.

IV. Local Spreading vs. Source Obfuscation

In this section we prove Theorem 3. Recall the notation we introduced in previous sections, which we use here as well. In particular, $h_t := \delta(vs_t, v^*)$ denotes the graph distance between vs_t and v^*, and $p(t,h) := \mathbb{P}(h_t = h)$. We will also use the elementary inequalities

$$(d-1)^{t/2} \le N_t \le \frac{d}{d-2}(d-1)^{t/2}. \tag{19}$$

Proof of Theorem 3: Our starting observation is that, due to the definition of adaptive diffusion protocols, we have that

$$R_t = \frac{t}{2} - h_t. \tag{20}$$

Thus in order to understand R_t it is equivalent to understand h_t.

We first turn to part (a) of the theorem. We described the likelihood function in Section III-B, see (7) in particular, from which it follows that

$$\mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^*) = \max_{1\le h\le t/2}\frac{p(t,h)}{d(d-1)^{h-1}}. \tag{21}$$

The assumption (4) thus implies that

$$p(t,h) \le \frac{Cd(d-1)^{h-1}}{N_t^{\gamma}} \le Cd(d-1)^{h-\gamma t/2-1} \tag{22}$$

for all $1 \le h \le t/2$, where in the second inequality we used (19). Now define

$$m_t := \frac{\gamma t}{2} - \frac{\log(Ct)}{\log(d-1)} - 1.$$

We then have that

$$\begin{aligned}\mathbb{P}(h_t \le m_t) &\le \sum_{h=1}^{\lfloor m_t\rfloor} Cd(d-1)^{h-\gamma t/2-1}\\ &= Cd(d-1)^{-\gamma t/2}\frac{(d-1)^{\lfloor m_t\rfloor}-1}{d-2}\\ &\le Cd(d-1)^{m_t-\gamma t/2} = \frac{d}{d-1}\cdot\frac{1}{t} \le \frac{2}{t}.\end{aligned}$$

In particular, we thus have that $\mathbb{P}(h_t > m_t) \ge 1 - 2/t$. Therefore

$$\begin{aligned}\mathbb{E}[h_t] &\ge m_t\mathbb{P}(h_t > m_t) \ge m_t\left(1-\frac{2}{t}\right)\\ &\ge \frac{\gamma t}{2} - \frac{\log(Ct)}{\log(d-1)} - 1 - \gamma.\end{aligned}$$

Now using (20) we have that

$$\mathbb{E}[R_t] = \frac{t}{2} - \mathbb{E}[h_t] \leq (1-\gamma)\frac{t}{2} + \frac{\log(Ct)}{\log(d-1)} + 1 + \gamma,$$

which concludes the proof of part (a) of the theorem.

We now turn to part (b) of the theorem. Consider the adaptive diffusion protocol defined as follows:

- For $t \leq 2/\gamma$, let $\alpha(t,h) = 1$ for all $1 \leq h \leq t/2$.
- For $t > 2/\gamma$, let $\alpha(t,h) = 1$ if $\lfloor \gamma t/2 \rfloor = \lfloor \gamma(t/2+1) \rfloor$ and let $\alpha(t,h) = 0$ otherwise.

This construction guarantees that for all even t we have that $h_t = 1$ if $t \leq 2/\gamma$, while for even $t > 2/\gamma$ we have that

$$h_t = \lfloor \gamma t/2 \rfloor$$

deterministically. Thus by (20) we have, for all even t satisfying $t > 2/\gamma$, that

$$R_t = t/2 - h_t = t/2 - \lfloor \gamma t/2 \rfloor \geq (1-\gamma)t/2.$$

On the other hand, by (21) we have, for all even t satisfying $t > 2/\gamma$, that

$$\mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^*) = \frac{1}{d(d-1)^{\lfloor \gamma t/2 \rfloor - 1}} \leq \frac{1}{d(d-1)^{\gamma t/2 - 2}}.$$

From (19) it follows that $(d-1)^{-t/2} \leq (d/(d-2))/N_t$ and so

$$\begin{aligned}\mathbb{P}(\widehat{v}_{\mathrm{ML}} = v^*) &\leq \frac{(d-1)^2}{d}\left(\frac{d}{d-2}\right)^{\gamma}\frac{1}{N_t^{\gamma}} \\ &\leq \frac{(d-1)^2}{d-2}\cdot\frac{1}{N_t^{\gamma}} \leq \frac{2(d-1)}{N_t^{\gamma}},\end{aligned}$$

where in the second inequality we used that $\gamma \leq 1$ and in the third inequality we used that $d \geq 3$.

V. Discussion

The main message of this work is that while adaptive diffusion protocols can hide the source from a snapshot adversary, they are ineffective when the adversary has access to multiple independent snapshots. The main question raised by our work is whether there exist other diffusion protocols that can obfuscate the source against such an adversary.

We make several simplifying assumptions in this work, which are important to discuss and study further. First, we assume throughout that the underlying graph is the infinite d-regular tree $\mathbb{T}_d$ (with $d \geq 3$), which is not a realistic model of real-world (social) networks. It is therefore important to study the questions we consider here on other underlying graphs, for instance, on more realistic models as well as on real-world social networks. We conjecture that our qualitative conclusions will carry over to more realistic settings, which motivates studying such a simplified setting.

We also assume that the adversary observes multiple *independent* snapshots. Previous work has considered multiple *sequential* snapshots (in time): Wang *et al.* [4] show that additional sequential snapshots cannot improve detectability under the SI model, while Fanti *et al.* [11] show that they can improve the detection probability at most logarithmically for adaptive diffusions. On the other hand, Cai *et al.* [12] show that multiple sequential snapshots can help detection when the spreading rates are heterogeneous, both theoretically and on Twitter data. As mentioned before, Wang *et al.* [4] show that multiple independent snapshots help significantly with detection under the SI model, and our results extend this to the family of adaptive diffusions. An interesting question is what happens in between, when the adversary observes multiple correlated snapshots (that are not necessarily sequential observations of the same diffusion). In particular, can spreading protocols take advantage of correlation in order to obfuscate the source against an adversary who observes multiple snapshots?

This question is related to local spreading as follows. An adversary who observes multiple snapshots can always use the following simple source estimator: pick a node uniformly at random among those which are infected in each snapshot. The probability of success of this estimator is the inverse of the size of the set of nodes which are infected in each snapshot. To minimize this, a spreading protocol should aim to maximize the size of this set. This can be done by having highly correlated snapshots, or by having a large amount of local spreading (which we have discussed in Sections I-C and IV). In any case, we conjecture that if there is a reasonable amount of independence among the observed snapshots, then the results will be qualitatively similar to those which we have obtained.

There are also many natural variations on what information the adversary has access to. For instance, Fanti *et al.* [11], [13] consider a spy-based model, where a fraction of nodes are corrupted and continuously monitor metadata such as message timestamps; they also consider a mixed model using both spies and a snapshot. Other information models include having a snapshot and additional relative information about the infection times of a fraction of node pairs [14], having partial infection timestamps [15], and having a noisy time series of observations [16], [17]. Understanding how our results change under these different information models of adversaries is a natural question for future work.

Further avenues to explore related to our work include game-theoretic formulations [18], optimal sensor/spy placement [19], confidence sets for the source [20], and multiple rumor sources [21]. We refer the reader to the position paper by Fanti and Viswanath [22] for further discussion of anonymous communication over networks.

We also note the importance of validating the main message of this work via real-world data sets. While obtaining data from anonymous messaging apps (such as Whisper [5], Blind [6], Yik Yak [7], or Secret [8]) is likely not feasible, an alternative option is to take a graph from an online social network (such as Facebook) as the underlying graph and to run simulations of adaptive diffusions and detection algorithms. We leave this for future work.

In conclusion, most results in this space—including ours in this work—are positive in terms of rumor source detection, and thus highlight major difficulties with guaranteeing anonymity for the source of a message in a social network. As

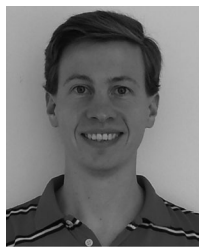
surveillance techniques grow ever more prominent in society, this emphasizes the need for further research, with the hope of ultimately providing robust anonymity guarantees.

ACKNOWLEDGMENTS

We thank two anonymous reviewers for their careful reading of the paper and their helpful questions and suggestions that helped improve the paper.

REFERENCES

[1] D. Shah and T. Zaman, "Detecting sources of computer viruses in networks: Theory and experiment," in *Proc. ACM SIGMETRICS Perform. Eval. Rev.*, 2010, vol. 38, pp. 203–214.
[2] D. Shah and T. Zaman, "Rumors in a network: Who's the culprit?" *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5163–5181, Aug. 2011.
[3] D. Shah and T. Zaman, "Finding rumor sources on random trees," *Operations Res.*, vol. 64, no. 3, pp. 736–755, 2016.
[4] Z. Wang, W. Dong, W. Zhang, and C. W. Tan, "Rumor source detection with multiple observations: Fundamental limits and algorithms," in *Proc. ACM SIGMETRICS Perform. Eval. Rev.*, 2014, vol. 42, pp. 1–13.
[5] Whisper, Accessed: Jun. 15, 2020. [Online]. Available: http://whisper.sh/.
[6] Blind, Accessed: Jun. 15, 2020. [Online]. Available: https://www.teamblind.com.
[7] Y. Yak, Accessed: Jun. 15, 2020. [Online]. Available: https://en.wikipedia.org/wiki/Yik_Yak.
[8] Secret, Accessed: Jun. 15, 2020. [Online]. Available: https://en.wikipedia.org/wiki/Secret_(app).
[9] G. Fanti, P. Kairouz, S. Oh, and P. Viswanath, "Spy vs. spy: Rumor source obfuscation," in *Proc. ACM SIGMETRICS Perform. Eval. Rev.*, 2015, vol. 43, pp. 271–284.
[10] G. Fanti, P. Kairouz, S. Oh, K. Ramchandran, and P. Viswanath, "Rumor source obfuscation on irregular trees," in *Proc. ACM SIGMETRICS Perform. Eval. Rev.*, 2016, vol. 44, pp. 153–164.
[11] G. Fanti, P. Kairouz, S. Oh, K. Ramchandran, and P. Viswanath, "Hiding the Rumor Source," *IEEE Trans. Inf. Theory*, vol. 63, no. 10, pp. 6679–6713, Oct. 2017.
[12] K. Cai, H. Xie, and J. C. Lui, "Information spreading forensics via sequential dependent snapshots," *IEEE/ACM Trans. Netw.*, vol. 26, no. 1, pp. 478–491, Feb. 2018.
[13] G. Fanti, P. Kairouz, S. Oh, K. Ramchandran, and P. Viswanath, "Metadata-conscious anonymous messaging," in *Proc. 33rd Int. Conf. on Mach. Learn.*, 2016, vol. 33, pp. 108–116.
[14] A. Kumar, V. S. Borkar, and N. Karamchandani, "Temporally agnostic rumor-source detection," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 3, no. 2, pp. 316–329, Jun. 2017.
[15] W. Tang, F. Ji, and W. P. Tay, "Estimating infection sources in networks using partial timestamps," *IEEE Trans. Inf. Forensics Secur.*, vol. 13, no. 12, pp. 3035–3049, Dec. 2018.
[16] A. Sridhar and H. V. Poor, "Sequential estimation of network cascades," 2020. [Online]. Available: http://arxiv.org/abs/1912.03800
[17] A. Sridhar and H. V. Poor, "Bayes-optimal methods for finding the source of a cascade," 2020. [Online]. Available: http://arxiv.org/abs/2001.11942
[18] W. Luo, W. P. Tay, and M. Leng, "Infection spreading and source identification: A hide and seek game," *IEEE Trans. Signal Process.*, vol. 64, no. 16, pp. 4228–4243, Aug. 2016.
[19] B. Spinelli, E. Celis, and P. Thiran, "A general framework for sensor placement in source localization," *IEEE Trans. Netw. Sci. Eng.*, vol. 6, no. 2, pp. 86–102, Apr.–Jun. 2019.
[20] J. Khim and P.-L. Loh, "Confidence sets for the source of a diffusion in regular trees," *IEEE Trans. Netw. Sci. Eng.*, vol. 4, no. 1, pp. 27–40, Jan.–Mar. 2017.
[21] S. Spencer and R. Srikant, "On the impossibility of localizing multiple rumor sources in a line graph," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 43, no. 2, pp. 66–68, 2015.
[22] G. Fanti and P. Viswanath, "Algorithmic advances in anonymous communication over networks," in *Proc. IEEE Annu. Conf. Inf. Sci. Syst.*, 2016, pp. 133–138.



Miklós Z. Rácz is an assistant professor at Princeton University in the ORFE department, as well as an affiliated faculty member at the Center for Statistics and Machine Learning (CSML). Before coming to Princeton, he received his Ph.D. degree in Statistics from UC Berkeley and was then a postdoc in the Theory Group at Microsoft Research, Redmond. His research focuses on probability, statistics, and their applications, and he is particularly interested in network science.

Jacob Richey is currently working toward the graduate degree with the University of Washington, Seattle, WA, USA, studying combinatorial probability. His research interests include random processes on graphs and interacting particle systems.