Spectrum Enforcement and Liability Assignment in Cognitive Radio Systems

George Atia†, Anant Sahai‡ and Venkatesh Saligrama†

Abstract— **The advent of frequency-agile radios holds the potential for improving the utilization of spectrum by allowing wireless systems to dynamically adapt their spectral footprint based on the local conditions. Whether this is done using market mechanisms or opportunistic approaches, the gains result from shifting some responsibility for avoiding harmful interference from the static "regulatory layer" to layers that can adapt at runtime. However, this leaves open the major problem of how to enforce/incentivize compliance and what the structure of "light-handed" regulation should be. This paper examines this and focuses on two specific technical problems: (a) determining whether harmful interference is occurring and (b) assigning liability by detecting the culprits.**

"Light-handed regulation" is interpreted as making unambiguous (and easily certified) requirements on the behavior of individual devices themselves while still preserving significant freedom to innovate at both the device and the system level. The basic idea explored here is to require the PHY/MAC layers of a cognitive radio to guarantee silence during certain timeslots where the exact sequence of required silences is given by a device/system-specific code. Thus, if a system is a source of harmful interference, the interference pattern itself contains the signature of the culprit. Nevertheless, identifying the unique interference pattern becomes challenging as both the number of cognitive radios and the number of harmful interferers increases.

The key tradeoffs are explored in terms of the "regulatory overhead" (amount of enforced silence) needed to make guarantees. The quality of regulatory guarantees is expressed by the time required to convict the guilty, the number of *potential* cognitive systems that can be supported, and the number of simultaneously guilty parties that can be resolved. *We show that the time to conviction need only scale logarithmically in the potential number of cognitive users. The base of the logarithm is determined by the amount of overhead that we will tolerate and how many guilty parties we want to be able to resolve.*

I. Introduction

A. Spectrum sharing — an overview

The idea that there are deficiencies in our current model of spectrum allocation originated in the economics/law/public-policy literature in the seminal papers by Coase [1] in 1959 and de Vany, et. al.[2] in 1969 and more recently by Goodman [3]. Their focus was mostly on making sure that spectrum was efficiently allocated to the socially most important uses. Mitola [4] introduced the idea of "cognitive radios" that are more intelligent and autonomous than the dumb radios of yesterday. The FCC's subsequent Spectrum Policy Task Force report [5] generated technical interest in the topic. This report revealed that the inefficiency of the current system went beyond the issue of assigning bands to inefficient uses — the current system has a dramatic underuse of spectrum since much of it is simply not used at all in most places/times. It has been argued that this waste is an inevitable consequence of the current static approach to spectrum access [6].

Dynamic spectrum access thus has the potential for improving the overall utilization of spectrum. Much of the work has focused on formalizing the concept of spectrum holes and discovering how to utilize them while avoiding harmful interference to those primary users that are actually active. This concept complements the ultrawideband philosophy where a strict spectral mask is employed since the constant presence of primary users is presumed [7].

Since cognitive radios are more intelligent, the novel idea was to replace static spectral masks as the interface between regulation and implementation and instead to directly deal with limiting interference itself. This is a shift in the very foundations of the field of wireless communication. Whereas earlier information-theoretic formulations assumed spectral mask constraints, in [8], Gastpar investigated the behavior of capacity under explicit spectrum-sharing constraints and showed that things can scale in a qualitatively different manner. Other work considered the capacity regions for cognitive radios using dirty-paper coding and ideas of interference cancellation [9], [10], although later such approaches were shown to be severely non-robust to wireless uncertainties [11].

Meanwhile, the dominant stream of research has focused on opportunistic spectrum sharing wherein the cognitive radios only use a band if they can verify that doing so will not disturb the existing primary users. Much work has focused on the spectrum sensing aspect of the problem [12]. One key challenge that emerges is the phenomenon of SNR walls that limit the ability of a single cognitive user to detect severely faded primary users [13]. Even if the SNR walls could be evaded, it turns out that single-user detection inevitably incurs a large overhead in the spatial domain — effectively giving up a significant amount of real estate where opportunistic use would have been safely possible but for an unreasonable fear of causing interference [6]. Cooperative sensing allows both of these barriers to be overcome [14], [15].

Although these efforts have led to a new understanding of the technical aspects of opportunistic spectrum usage, fundamental gaps remain in translating these into practice. Indeed the very prospect of cooperative spectrum use in turn raises regulatory problems. How should cognitive radios be regulated? When single-user sensing is all that is contem-

† G. Atia and V. Saligrama are at the Department of Electrical and Computer Engineering, Boston University, Boston, MA, Emails: {geokamal,srv}@bu.edu

‡ A. Sahai is at the department of Electrical Engineering and Computer Sciences, University of California Berkeley, CA, Email: sahai@eecs.berkeley.edu

plated, then it is easy to see how the current ultrawideband and unlicensed device paradigm can be extended to cover it: as long as a device senses at an appropriate sensitivity, it is allowed to communicate within an appropriate mask. (See [16] for how sensitivity trades off with power control.) This is an *a priori* rule that can be relatively easily enforced. But how can we certify the behavior of a network of cooperating users when that network forms dynamically in the field? Similar questions plague market-oriented approaches. How can we enforce that users have the right to use the spectrum that they are using?

What is really missing is a means for *a posteriori* **Spectrum Enforcement** that works in conjunction with some amount of *a priori* certification of the devices themselves. Such a perspective has always been present on the policy side in [1], [2], [3], but has to our knowledge, never been explored technically. A companion paper [17] and [18] develop a toy game-theoretic model to discern how effective spectrum enforcement mechanisms must be to properly incentivize cognitive systems to not cheat. Even without such a quantitative model, it is qualitatively obvious that *if there is no chance of being caught, then there is very little incentive to invest serious engineering effort in complying with the regulations*, especially in cases where compliance might result in lower quality of service to the cognitive-radio system's own users.[1] The core problem of identity was vividly captured by Faulhaber in [19] with his evocative phrase "hit and run radios."

B. Motivation

The main idea in this paper is to design a way to trace who is violating a protocol. In this sense, this is the opposite of privacy. The analogy is to cars. There are rules designed to protect the safety of others. However, not all of these rules are unbreakable at the device level (e.g. cars do not sense stop signs and force you to stop). Instead, cars are required to have visible license plates that allow violators to be identified and penalized. Our goal is to impose the minimal number of rules that can presumably be checked at device certification time.[2] Such minimalist rules at device certification time allows substantial room for innovation as new technologies develop.

A philosophy of hierarchical punishment is proposed wherein the offended entity should be able to identify that it is being interfered with and then identify a subset of users who are responsible for causing the interference. There are three potential approaches to 'identity.' In the most straightforward approach, identity is explicitly transmitted by the physical layer as a separate wireless signal in a mandated format. If a primary user experiences harmful interference, then it merely has to decode this signal to learn the identities of all the potential interferers so that they can be penalized. Such an "identification pilot" signal would necessarily impose an overhead on the secondary users and one could analyze the tradeoffs possible. However, while this approach is conceptually simple, it has four major shortcomings:

1) It forces us to mandate a standard PHY-layer format for transmission of this identity information. This adds additional complexity to systems that want to use another format/modulation for their own signals. Moreover, this format would have to be tamper-proof and thus might impose costly recertification requirements for changes that could be handled as a simple software update.
2) It imposes a decoder PHY burden on the primary user to implement a way to decode this identity information so that it can penalize secondary systems that are cheating in the vicinity. This is in addition to the primary's own PHY layer for decoding its own data. Sadly, if the regulation is successful and the threat of punishment is enough to prevent cheating, then this particular part of the primary system will be unexercised and hence is likely to suffer "bit-rot" as the primary system evolves.
3) It does not allow the primary user to distinguish between harmful interference and unfortunate fading or bad luck. As Hatfield points out in [20], wireless environments are notoriously unpredictable. A primary user might just be out of range of its transmitter or it might be drowning in harmful interference. There is no way to tell them apart if the secondary identity information was just carried by a separate broadcast signal.
4) More subtly, a broadcast identity does not distinguish between the innocent and the guilty. Thus it greatly reduces the incentive to deploy innovative approaches to reduce interference. For example, a cognitive-radio network might be able to use beamforming to null out its transmissions at the primary receiver. However, if any other cognitive radio causes harmful interference, the careful radios will also be punished since they were also in the neighborhood at the time of the incident.

The second approach to identity trades beacon overhead for reporting overhead. Cognitive radios could be required to keep extensive and detailed records of their trajectories and operations. These logs could then be regularly uploaded to the authorities and searched to find the culprits whenever a credible report of harmful interference is filed. This avoids the first two shortcomings, but does nothing about the second two. It is also a privacy/security nightmare since once such logs exist, it will be very tempting for malicious parties[3] to target them for reasons that have nothing to do with preventing harmful interference.

[1]This is far from a hypothetical concern. Consider the example of DVD players. The license officially requires players to obey region-coding commands on the disk even when the owner of the disk would prefer that the disk will play. Many manufacturers of cheap DVD players ship players that do not comply with this license requirement and instead act in the interest of the paying customers who would prefer that their disks play, no matter where those disks might have been purchased.

[2]This issue did not arise in de Vany, et al's models from [2] because at the time of that work, wireless transmission was considered to be limited to only a few big players operating steadily from fixed positions at fixed frequencies. Identity could piggy-back on physical location. Finding the offending transmitter would resolve the issue and the fixed transmitter could be found by a man with a van and an antenna. Tomorrow, the story will be entirely different with frequency-agile mobile devices potentially operating in ad-hoc cooperative modes.

[3]Such as hackers or oppressive governments.

Because of these issues, we explore a different approach in which the identity of a device is implicitly announced by the pattern of use/interference itself. Because it is simpler to analyze, we will focus here on the assumption that all wireless nodes have access to a common sense of time and can divide both time and frequency into slots of moderate length.[4] The complete system consists of a set of frequency-specific binary codes on users that govern their medium access to that frequency band (See [21] for a perspective on why similar codes can be useful from the perspective of *ad-hoc* MAC protocols). This restriction is hard-coded into the device and verified as such during the device certification process.[5] In addition, there might be a certified feedback path by which nodes can be told to "cease and desist" their interfering activities, but this is not studied here.

The focus here is entirely on being able to detect the presence of harmful interference and to identify the culprits. We adopt a "no harm no foul" philosophy towards interference. If no primary receiver is disturbed, then there is no problem. In Section II, codes are arranged so that the primary can rule out "natural causes" (e.g. shadowing) for degraded performance and decide that foul play must be at work. In this section, we consider the tradeoff between the two kinds of errors (false accusations and missed accusations), the overhead of silence slots, and the time required to detect foul play.

Section III then deals with liability assignment using the idea of superimposed codes. Bounds on fundamental tradeoffs between the various system parameters are provided within an even more idealized context wherein there is effectively no noise (See Figs. 8 and 9). Finally conclusions are presented in Section IV.

The main result of this paper is the quantification of the fundamental tradeoffs between the various system parameters showing that:

1) Supporting a larger number of potential[6] users and increasing robustness come at the expense of increased time till conviction. (See Fig. 8)
2) A fundamental tradeoff exists between efficiency, in terms of achievable utilization rates, and timeliness (See Fig. 9). For example, with time-slots of four milliseconds, we can support more than two-hundred potential cognitive users each having access to more than 80% of the time-slots and still be able to resolve harmful interference to a guilty pair of users within two seconds! This means that even if one user is malfunctioning, there is still no incentive for another user to start cheating since it will fear being identified and punished.
3) Allowing a very large number of distinct identities for potential cognitive users becomes prohibitively expensive suggesting the need for a gradual punishment mechanism wherein innocent bystander systems might incur short periods of false conviction.

[4]For example, we assume that all systems are directly or indirectly synchronized to GPS. The length of a slot is assumed to be long enough that the propagation lag from any possible harmful interferer is much shorter than the slot length. The frequency width of the slot is considered to be large enough that the secondary user will not leak significantly out of its frequency slot. The case of asynchronous secondary users remains open and techniques lifted from feature-detection might prove useful in determining the tradeoffs in that context.

[5]Of course, in addition cognitive radios are not supposed to transmit if they will cause harmful interference to primary users. But the point of this whole approach is that this restriction need not be completely verified during device certification.

[6]It is critical to understand that for regulatory purposes, it is the number of potential identities that is important rather than the number of actual users operating in a particular environment. This is like the difference between how many license plates are pressed vs how many cars are on a particular road.

II. Detecting harmful interference

In this section we address the first part of the problem that deals with the identification of the source of performance degradation. For simplicity, we assume that the primary user has a packetized system wherein each packet is shorter than a time-slot for the code. The goal is to decide what is the cause for the observed packet losses to avoid unsubstantiated false accusations. This is because the primary's QOS can degrade due to a variety of "natural causes" in addition to the "foul play" represented by harmful interference. For example: the primary receiver is too far away to get a strong enough signal, the primary receiver walked into a shadowed area, the primary antenna became disconnected, etc.

The goal is for the primary user to be able to distinguish between the uncertain background losses and the presence of harmful secondary users. In this sense, the problem is the exact inversion of the usual cognitive-radio perspective of secondary users trying to detect the primary. As discussed in [13], this can be hard and generally, there exist SNR Walls that prevent the detection of weak users. The reason for hope comes from [22] wherein it is shown that while microscale features of weak signals can be blurred by environmental uncertainties, it is possible to construct signals that have macroscale[7] features and thus impose no SNR Walls. Whereas designing the primary signal is a dubious proposition, it is completely reasonable to demand some features from the secondary transmissions.

Our idea is to introduce silence slots during which no secondary transmission is allowed and compliance of this is enforced at the device certification level. These silence periods serve as "inverse canaries in a coal mine" and hence we refer to them as canaries or canary slots. Whereas canaries would stop singing in mines if the gas levels were too high, these inverse canary slots serve as examples of clean channels where there can be no interference. Intuitively, if the source of interference is anything other than spectrum violators, then the observed degradation should be time invariant. Hence, these silence periods can be efficiently used for testing and discrimination.

Based on packet-drop data, the primary user must decide among two competing hypothesis:
(a) Null Hypothesis (H_0): The packet-drop data pattern is attributable to environmental factors, such as fading and attenuation losses commonly encountered in wireless systems;

[7]Roughly, microscale features are things like pilot tones and cyclostationary features that are at the same scale of frequency-selective fading, coherence times, and timing uncertainty. Macroscale features persist on a much bigger scale and thus cannot be masked by the uncertainty.

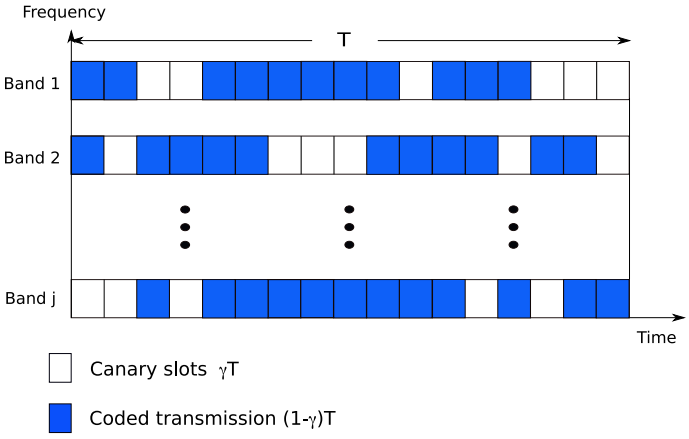

Fig. 1. Canary slots introduced for identification. There are certain code positions that always contain a zero for a particular frequency. Canary slots serve as examples of clean channels where there is no interference. This allows the primary to rule out "natural causes" for degraded performance and decide that interference must be due to spectrum violation. Canary slots are different in different bands. This allows systems to hop among different frequency bands to maintain stable links.

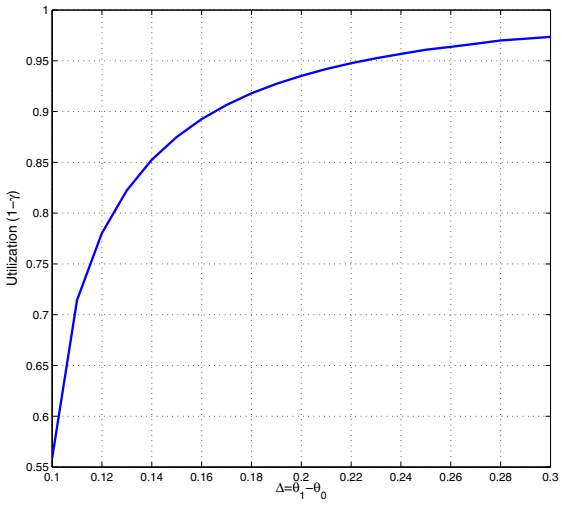

Fig. 2. Utilization versus $\Delta = \theta_1 - \theta_0$ at $P_M = 10\%$, $P_F = 20\%$ and $T = 450$ slots

(b) Significant Hypothesis (H_1): The packet-drop pattern suggests existence of secondary violators.

To formalize this approach, we let γ denote the fraction of canary slots as shown in Fig. 1. We also define the Time-to-Identification T, as the total number of slots required to achieve a target detection criteria.

Let X_t and Y_t be i.i.d. Bernoulli random variables with unknown parameters θ_0 and θ_1 representing the packet-drop probabilities during canary and non-canary slots, respectively. In other words:

$$\begin{aligned} X_1, X_2, \ldots, X_{\gamma T} &\sim B(\theta_0) \\ Y_1, Y_2, \ldots, Y_{(1-\gamma)T} &\sim B(\theta_1). \end{aligned} \tag{1}$$

Then the hypothesis testing problem reduces to:

$$H_0 : \theta_0 = \theta_1 \qquad H_1 : \theta_0 \neq \theta_1, \theta_0 < \theta_1$$

which is a classical problem known in the statistics literature as a test for the equality of two proportions [23]. For target missed-detection and false-alarm probabilities $P_M \leq \epsilon$ and $P_F \leq \delta$, respectively, it is not hard to show for sufficiently large sample sizes, when suitable Gaussian approximations are valid (see Appendix A for details), that:

$$T \approx \frac{\left[\sqrt{\bar{\theta}(1-\bar{\theta})(1+\frac{1}{\kappa})}z_\delta + \sqrt{\theta_1(1-\theta_1) + \frac{\theta_0(1-\theta_0)}{\kappa}}z_\epsilon\right]^2}{(1-\gamma)\Delta^2} \tag{2}$$

where $\kappa = \frac{\gamma}{1-\gamma}$ is the ratio between the total number of canary and non-canary slots, $\bar{\theta} = \frac{\theta_1 + \kappa\theta_0}{1+\kappa}$ and $z_w = \Phi^{-1}(1-w)$, with $\Phi^{-1}(.)$ denoting the inverse CDF of a normal Gaussian distribution. $\Delta = \theta_1 - \theta_0$ is the difference between the packet-drop probabilities during canary and regular slots.

Since secondary users can only transmit during non-canary slots, it is natural to define utilization as the fraction of non-canary slots, i.e. $1 - \gamma$ (See [24] for another context of code construction in which maintaining a high utilization is important). Fig. 2 plots utilization as function of the hypothesis separation Δ. It is shown that whenever packet-drop probabilities are sufficiently distinct, only a small fraction of canary slots is needed and high utilization rates are thus achievable. Fig. 3 shows the required time till identification versus the fraction of canary slots γ at missed-detection $P_M = 10\%$ and false-alarm $P_F = 20\%$. The true unknown parameters are $\theta_0 = 0.4$ and $\theta_1 = 0.6$. Again this demonstrates that high utilization rates are achievable if we are willing to wait long enough. Figure 4 shows the missed-detection probability P_M versus the fraction of canary slots γ for different values of Δ. In this setup, the time till identification is 150 time-slots and the target probability of false alarm is $P_F = 20\%$.

For convenience, the plots were made using these relatively high values for the probabilities of error. However it is easy to see how these same formulas can also give bounds for much stricter performance requirements. However, large-deviations theory tells us that very low probabilities of error will come at the cost of either a moderate increase in the time till identification or a dramatic reduction in the utilization factor.

If there is only one system then there is a fear of being caught. However when there are many coexisting systems there is an incentive to cheat and hide in the crowd. Hence there is a need for identifiability of culprits which we address in the next section.

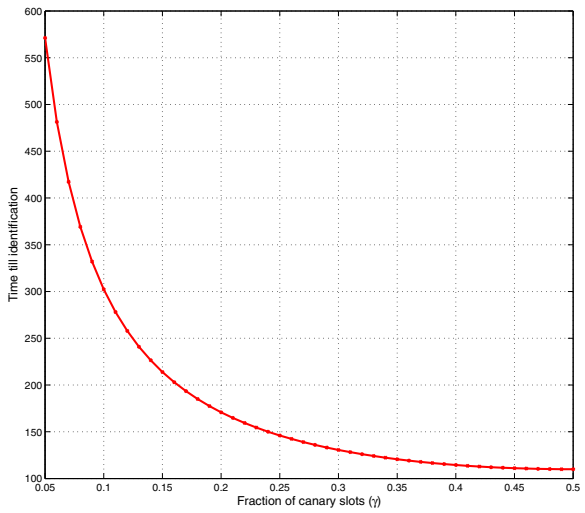

Fig. 3. Time till identification versus fraction of canary slots γ at $P_M = 10\%$, $P_F = 20\%$ and $\Delta = 0.2$

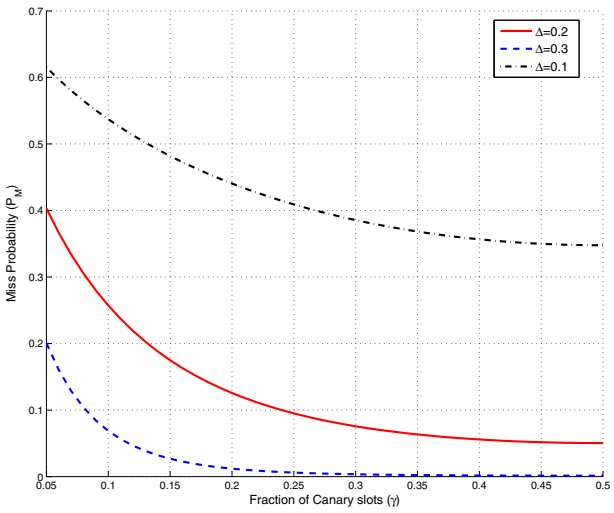

Fig. 4. Miss Probability P_M versus fraction of canary slots γ for different values of Δ.

III. Blame Assignment

So far, we devised a hierarchical coding architecture with periods of silence that allows foul play detection. If the cause of interference is identified as harmful interference, then it remains to identify the guilty parties. The most important quantity of interest is how long does it take to identify the guilty parties. We define this as the time to conviction T_c. The second important aspect is how powerful the code is in terms of how many cognitive systems it supports and how large the size of the guilty set could be. In this section, we identify fundamental tradeoffs between the various system parameters. We also show how superimposed codes could be used for culprits identification.

Every cognitive user is assigned a different binary codeword that defines its allowable transmission slots and this code is known to the enforcement system.[8] The goal is to identify the guilty parties by matching the interference pattern to the known collection of codewords. Here, we consider the best-case scenario where it is assumed that the background level of noise is zero and that all the interfering transmissions are therefore detected. Furthermore no gaming or adversarial transmission protocol is attempted by the secondary users. In other words:

1) A culprit transmits whenever he is allowed to transmit, i.e. the transmission times of the guilty parties coincide with the instants when the code of *any* of its users is equal to 1. An example is shown in Fig. 6 where packets are dropped whenever user 2 transmits.

2) The packet-drop process is deterministic in the following sense:

- Whenever there is an illegal transmission, a primary packet is dropped.
- During periods of no interference, packets are received with probability 1 (negligible wireless losses).

One obvious choice would be to allow for exactly one user to transmit at each time-slot (TDMA) as shown in Fig. 5.a. This way, packet drops in specific slots could be exclusively assigned to illegal transmission from the corresponding user. However, the major drawback of such a scheme is reduced utilization since each user would be only allowed to transmit for $1/N$ fraction of the time. This amounts to a large overhead in terms of time during which individual cognitive radios are denied access to an otherwise free band. Alternatively, one could design a code that permits higher utilization rates while still being able to catch spectrum violators. After all, there is no guarantee that all cognitive users are actually active in this geographic area and so it is important to allow any individual cognitive user to *a priori* be able to use as much of a band as possible.

As an example, see Fig. 5.b where the code design achieves an average per-user utilization rate of 35.71% while being able to identify up to any 2 users out of a total of 6 users. In each case we show the codebook (rows representing the codewords for each user) and their open transmission slots. The example also highlights a tradeoff between utilization and the code strength, i.e. how many culprits are uniquely identifiable. Note that there are two kinds of overhead that are causing utilization to be less than full. First, as pointed out in the previous section, are the inverse-canary slots which are

[8]It is clear that this code could be constructed in a hierarchical manner by ANDing together various codes. Common to all systems, we have the canary and non-canary code for this frequency. Beyond that, there might be different codes corresponding to different operators, device-manufacturers, and individual devices themselves.

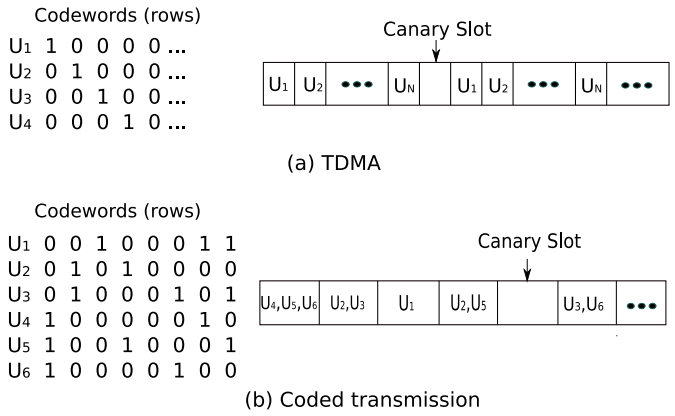

Fig. 5. TDMA versus coded transmission. Using the coding idea higher utilization rates are achievable. In the given example with $N = 6$ users the average per-user utilization with TDMA is 16.67% in contrast to 35.71% with coded transmission. In each case, we show the code structure (left) (identity matrix in case of TDMA) and the open transmission slots for the different users (right).

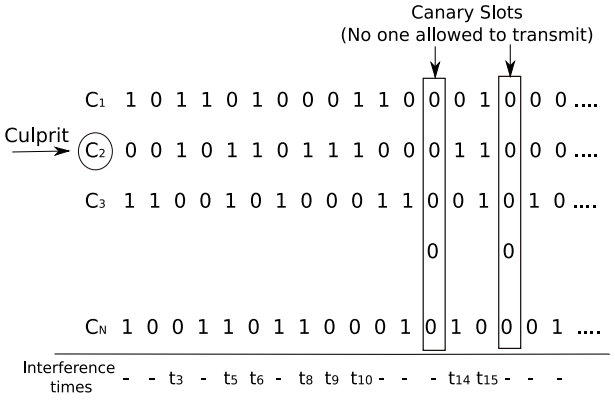

Fig. 6. Codes and interference pattern. The i-th user is only allowed transmission during periods where its code $c_i = 1$. This restriction is hard-coded into the device. The interference pattern can then be used to identify the culprits as it has the signature of the culprit set.

common to all users. Second are the silence periods that are different among different users. This section focuses on the second aspect of utilization.

A. Basics and the connection to superimposed codes

In what follows we identify fundamental tradeoffs between the necessary time to conviction, the size of the culprits' set and the per-user potential utilization. First, we introduce the model and define our notation:

- N: number of cognitive systems.
- G: the unknown set of culprits (guilty parties) of size K.
- g_i: binary indicator function for the i-th user. $g_i = 1$ if i-th user is an actual culprit, i.e. $i \in G$.
- $y(t)$: primary packet-drop observation at time t, $y(t) \in \{0, 1\}$ where 1 is a drop.
- c_i: binary code of the i-th system. (**a column vector**)

Assume that the primary system makes its accusation after observing y for T_c time-slots (time to conviction). Given these assumptions the problem simplifies to a solution of the following set of simultaneous Boolean equations:

$$y(t) = \bigvee_{i=1}^{N} (g_i \wedge c_i(t)), \ t = 1, \dots T_c \tag{3}$$

where $\vee$ denotes the (OR) operation and $\wedge$ the (AND) operation. The above equation simply states that a packet is dropped if any user from the culprits set is allowed to transmit (i.e. its code=1). Defining $\mathbf{y}$ as the observation vector of length T_c, Eq. (3) can be also rewritten as the Boolean sum of the codes of the guilty users:

$$\mathbf{y} = \bigvee_{i \in G} \mathbf{c}_i. \tag{4}$$

If we do not care about computational complexity in this simple deterministic framework, there is a simple "guilty until proven innocent" algorithm to detect a set of spectrum violators by simply looking at the times of "no interference" and comparing them to the codes. Any user with a 1 in such slots is deemed innocent since no interference was detected there. All others are convicted as guilty.

The key questions are: 1) what is the minimum time to conviction T_c to uniquely identify the set G; 2) How to design efficient codes to identify the guilty users without sacrificing secondary utilization? Is random coding sufficient?

Before we get to answer the aforementioned questions, we show how the classical concept of superimposed codes is relevant to our identification problem.

Definition 3.1: **Superimposed codes:** An $M \times t$ matrix $\mathbf{X}$ is called a superimposed (s, t, L)-code of length M if the Boolean sum of any $\leq s$-subset of its columns can cover[9] not more than $L-1$ columns that are not components of the given Boolean sum [25].

In the case where $L = 1$, i.e. $(s, t, 1)$ codes, the Boolean sum of up to s different codewords covers no codeword other than those used to form the Boolean sum. This class of codes is also known as Zero-False-Drop of order s (ZFD_s) in [26].

Superimposed codes have been used in numerous applications such as file retrieval, multiple access communication, screening [27], [28], [29] and here we show that they could be used for the culprits identification problem.

Theorem 3.1: If the matrix of codewords $C = [\mathbf{c}_i], i = 1 \dots N$, of size $T_c \times N$ is a ZFD_K code then it guarantees *unique* identification of any culprits set G of size less than or equal to K if packets are dropped according to the deterministic model in Eq. (3).

Proof: See Appendix B. ■

In fact, a less constrained class of codes known as $\widetilde{(s, t)}$-codes is both necessary and sufficient for unique identification of the guilty set as their properties provide necessary and sufficient conditions to be uniquely decipherable [25]. These codes are defined as codes where the Boolean sums composed of s different columns are different.

[9] We say that a column x is covered by a column y iff $x \vee y = y$.

The use of superimposed codes provides a significant improvement in utilization in comparison to TDMA. The maximum cardinality K of the guilty set G, represents the code strength which can be a design parameter for the system. If $|G| > K$, i.e. more than K users are violating the spectrum regulation, unique identification might not be possible and a fraction of innocent users (those whose codes are covered by the actual measurement vector) have to be sanctioned along with the guilty parties. In the following, we provide both upper and lower bounds (as function of N and K) on the time-till-conviction T_c, under various conditions, for unique identification of the set G.

B. Necessary conditions on time to conviction T_c (Lower bounds)

To guarantee that any set G of size at most K is uniquely identifiable, an obvious lower bound on T_c can be written as:

$$T_c \geq \log\left(1 + \sum_{i=1}^{K} \binom{N}{i}\right). \tag{5}$$

Proof: This follows from the fact that the total number of different vectors obtained from the superposition of i vectors, $i = 1 \ldots K$, cannot exceed the number of nonzero T_c-digit binary numbers. Otherwise, unique decipherability is impossible. ■

Thus, asymptotically when N is large and K is constant,

$$T_c \geq K \log\left(N(1 + o(1))\right). \tag{6}$$

Note that this could also be used to provide an upper bound on the total number of allowable cognitive systems N subject to conviction-time constraints.

If the code matrix C is chosen to be a ZFD_K code then other lower bounds on T_c could be also obtained from well-known bounds on minimum lengths of such codes [25]. Among these bounds [25]:

$$T_c \geq \min\{\frac{(K+1)(K+2)}{2}; N\} \text{ if } C \in ZFD_K. \tag{7}$$

This means that if it is required to uniquely detect $K \geq \sqrt{2N}$ then one cannot beat the TDMA scheme in terms of shortest time to conviction (but at the expense of low utilization $\frac{1}{N}$). Also using results from [25],

$$T_c \geq \beta(K) \log(N(1 + o(1)) \text{ as } N \to \infty, \quad K \text{ const.} \tag{8}$$

where $\beta(K) \geq \frac{K^2}{2\log(e(K+1)/2)}$ for $K \geq 2$ and $\beta(K) = \frac{K^2}{2\log K}(1 + o(1))$ if $K \to \infty$.

As argued before, $(\widetilde{s,t})$ codes (also known as UD codes) are less restrictive than ZFD codes and provide necessary and sufficient conditions on unique detection and hence a lower bound on their length could serve as a lower bound on T_c. Using proposition 5 in [25] one can show that:

$$T_c(ZFD_{K;N}) \geq T_c(UD_{K;N}) \geq T_c(ZFD_{\lfloor K-1/2 \rfloor; \lfloor N/2 \rfloor}).$$

Hence a lower bound on T_c of any code could be obtained by replacing the factor multiplied by the $\log$ in Eq. (8) by $\beta(\lfloor \frac{K-1}{2} \rfloor)$, i.e.

$$T_c \geq \frac{K^2}{8 \log K} \log(N(1 + o(1))) \text{ if } N \to \infty, \quad K \text{ const.} \tag{9}$$

It is worth mentioning that all the bounds above provide constraints on the minimum required T_c as a function of K and N even without utilization being considered. However, high utilization is important and so next we derive another information-theoretic lower bound on T_c that incorporates utilization. We consider the case where the $T_c \times N$ code matrix C has the same number of ones (ℓ) per row. In other words, the code is designed such that ℓ users are allowed to transmit whenever the primary is absent. The average per-user potential utilization p is thus $\frac{\ell}{N}$.

Theorem 3.2: If the code matrix C has the same number ℓ of ones per row, then a lower bound on the time to conviction T_c for a culprit set of size K is given by:

$$T_c \geq \frac{\log \binom{N}{K}}{H(\eta)} \tag{10}$$

where $H(\cdot)$ is the binary entropy function and $\eta = \frac{\binom{N-K}{Np}}{\binom{N}{Np}} = \frac{\binom{N(1-p)}{K}}{\binom{N}{K}}$.

Proof: See Appendix C. ■

When N is large, $\eta \approx (1-p)^K$. The bound thus has an intuitive explanation as the number of bits to be identified divided by the entropy assuming that K users are cheating at random. It is possible to prove a similar bound using other information-theoretic techniques as well. These allow a generalization to the case of noisy observations [30].

C. Sufficient conditions for time to conviction (Upper bounds)

In this section we prove sufficient conditions, i.e. achievable bounds, on the time to conviction of the culprits set. Since this is a very simplified model, these sufficient conditions are less informative, but it is important to see that such codes can actually exist. We consider random coding for code generation i.e., we assume that C is a randomly generated code according to a Bernoulli distribution with parameter p, where p denotes the average per-user potential utilization.

1) Connection to bipartite graph covering: Under the simplified assumptions mentioned earlier, it is interesting to see that our problem can be formulated as a problem of finding a minimum cardinality discriminating *code*[10] over a bipartite graph. In Fig. 7, a set of nodes U represents time instants and the other set of nodes V represents the available cognitive

[10]In order to match the literature we would like to point out that here the word *code* is used in a completely different sense where it basically refers to a subset of nodes. To avoid confusion we use the italic word *code* to make the distinction.

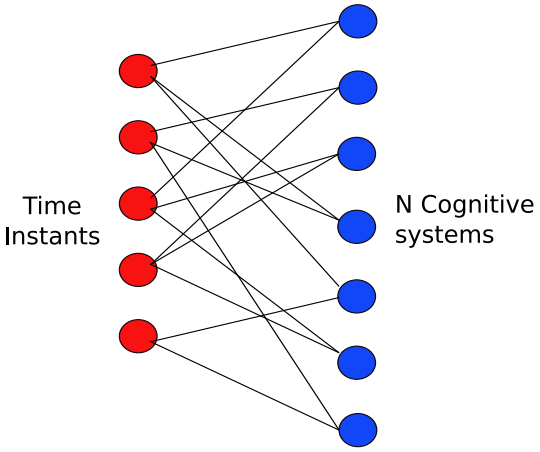

Fig. 7. A bipartite graph to model time instants during which different users are allowed transmission. Graphical approaches could then be used to find a set of nodes providing unique coverage of any set of size K.

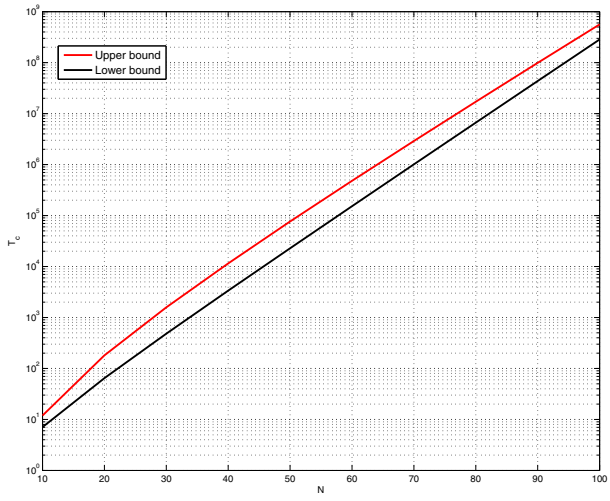

Fig. 8. Sufficient and Necessary conditions for time to conviction with $\alpha = \frac{K}{N} = 0.2$ constant plotted versus total number of potential cognitive systems N. A fundamental tradeoff showing that supporting a larger number of potential users and increasing robustness come at the expense of increased time till conviction requirements.

systems. An edge between nodes t and i means that the i-th system is allowed to transmit at the t-th time-slot, i.e. $c_i(t) = 1$. Let's assume that the code is designed in such a way that it detects any culprit set of size at most K. Each node in U runs a test whose outcome is 1 or 0, where 1 designates a packet drop. Then the question is equivalent to a vertex-covering problem, where the goal is to find the minimal set U (of size T_c) to *uniquely* cover subsets of vertices of V of size at most K. This is a function of the node degree which represents (how many users per slot) and (how often is a user allowed to transmit) from U and V perspective, respectively.

Note that for the simple case where $K = 1$ this becomes a problem of searching for an identifying *code* [31] over a bipartite graph. In this case, for the graph to possess a discriminating *code* (*codewords* belong to the set U) to be able to distinguish the different individuals (of set V) then a necessary and sufficient condition is that the graph is twin free [32].

If the goal is to identify any culprit set of size at most K then the problem reduces to finding a K-identifying *code* over a bipartite random graph where edges are added to the graph with probability p. Finding the minimal K-ID *code* corresponds to the minimum time to conviction T_c. Before defining a K-ID identifying *code* we first introduce some notation. Let $\mathcal{G}(A \bigcup I, p)$ denote a bipartite graph with 2 separate sets of vertices A and I and random edge placement with probability p. Let $N(x)$ denote the neighborhood of any vertex $x \in I$. Obviously, $N(x) \subseteq A$ since the graph is bipartite. The goal is to find a minimal *code* D, where in our case $D = A$, such that it uniquely identifies any set $X \subseteq I$, where $|X| \leq K$. Denote the set that contains all such sets by $\mathcal{S}_K$. By unique identification we mean the following:

1) For any X and $Y \in I \bigcap \mathcal{S}_K$, $I(X, D) \neq I(Y, D)$ where $I(X, D) = \bigcup_{x \in X} N(x) \bigcap D = \bigcup_{x \in X} N(x)$.
2) If $X \neq \emptyset$, then $I(X, D)$ is nonempty.

In [33] it was shown that for random graphs a *code* is K-ID iff the above conditions are satisfied. We use this result to prove a sufficient condition on the time to conviction T_c. Before that we derive a lemma about the probability that a *code* D of size $|D|$ is not a K-ID *code*.

Lemma 3.3: Given bipartite graphs $\mathcal{G}(A \bigcup I, p)$ with $|I| = N$ and edge placement probability p, then the probability $\Pr(D \text{ not a } code)$ that a *code* $D = A$ of size $|D|$ is not a K-ID *code* is upper bounded by $N^{2K}(1 - \min\{p, 2p(1-p)\}(1-p)^{K-1})^{|D|}$.

Proof: See Appendix D for a proof from first principles. In [25] the authors provide an upper bound on the probability that a randomly generated code does not happen to be a proper superimposed code. This is equivalent to this lemma. ■

The lemma immediately implies the following theorem:

Theorem 3.4: Let $0 < p < 1$ be a constant. Then for unique identification of K culprits out of N secondary systems using a random Bernoulli(p) generated code, a time to conviction $T_c = \Omega\left(\frac{2K \log N}{\log(1/q_K)}\right)$, where $q_K = (1 - \min\{p, 2p(1-p)\}(1-p)^{K-1})$, is achievable.

This is easy to show since for the given cardinality the probability that the corresponding set is not discriminating is less than 1 (Lemma 3.3). Since the problem is deterministic, there must exist a good code if the probability of a bad code is less than 1. Notice that $1/\log(1/q_K) = O(2^K)$, thus the minimum time to conviction (minimum cardinality of the K-ID code) is $O(K2^K \log N)$.

For illustration purposes Fig. 8 depicts the matching of the lower and upper bounds derived in Thm. 3.2 and Thm. 3.4, respectively.

It follows from Thm 3.1 that random coding bounds on the length of ZFD_K codes also provide achievable bounds on T_c.

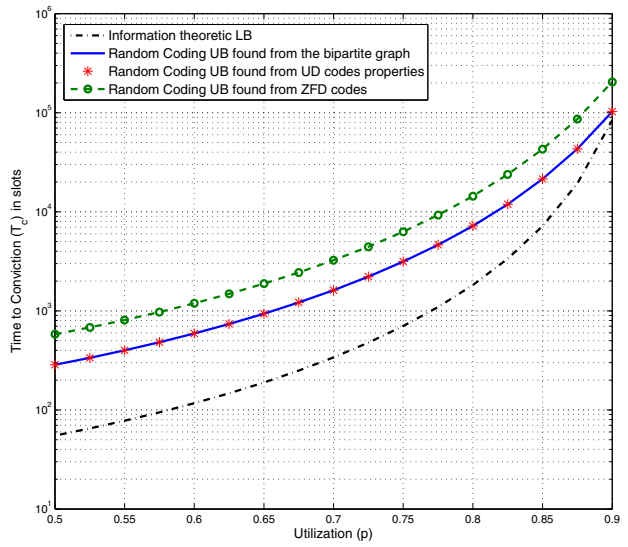

Fig. 9. Utilization (p), Time to conviction (T_c) Tradeoff with random coding, N=40, K=4. This figure shows a fundamental tradeoff between efficiency, in terms of higher utilization rates, and timeliness for fixed N and K.

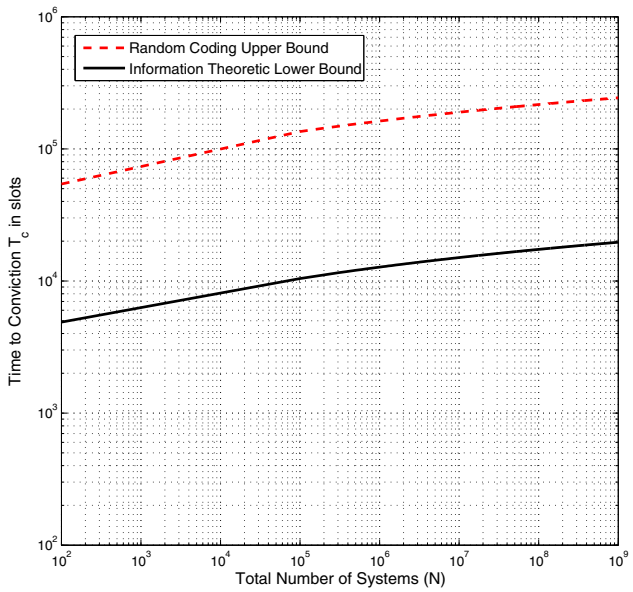

Fig. 10. Time till conviction T_c versus total number of systems N, with $K = 6$. The figure shows a scaling behavior of the time to conviction as N is scaled and K held constant. Supporting a large number of users becomes prohibitively expensive.

More specifically for per-user utilization p:

$$T_c(K,N) \leq \frac{K+1}{\log(1/q_K)} \log N + \frac{\log K!}{\log q_K} \tag{11}$$

where $q_K = 1 - p(1-p)^K$. The minimum time to conviction T_c is thus achievable at utilization $p = \frac{1}{K+1}$ (better than the TDMA utilization of $1/N$).

Hence,

$$T_c \leq \frac{K+1}{\log(1/q_K)} \log(N(1+o(1))) \quad \text{as } N \to \infty, \quad K \quad const. \tag{12}$$

and as $K \to \infty$,

$$T_c \leq \frac{e}{\log e} K^2 (1+o(1)) \log(N(1+o(1))). \tag{13}$$

2) Kautz-Singleton constant-weight codes: Existing codes like Kautz-Singleton [26] are a good example of constructed superimposed codes. As they have this property of a fixed number of ones (w) for each codeword, they could provide a fixed per-user utilization $\frac{w}{T_c} \leq \frac{1}{K} + \frac{1}{N}$. Also those codes are guaranteed to be ZFD_K $\forall K \leq \lfloor \frac{w-1}{\lambda} \rfloor$ where λ is the maximum correlation of any pair of codewords. Again bounds on the length of such codes translate to bounds on the time to conviction T_c.

D. Tradeoffs at a glance

Ideally, one would like to see high utilization (hence efficiency p), robustness (large K hence catch-ability), small T_c for timeliness (system responsiveness) and the ability of the system to support a large number of users N. However, a key message of this paper is the identification of tradeoffs between the various system parameters p, K, T_c and N.

Fig. 9 shows a fundamental tradeoff between secondary utilization p and the time to conviction T_c when K and N are held constant. It is shown that high values of utilization (e.g. 90% of the slots being available to any given user) are achievable at the expense of longer time till conviction. Intuitively, one could think of a scenario where more slots are added during which all users are allowed to transmit. This leads to an increased utilization but effectively these slots will not help with the conviction process.

As pointed out earlier, increasing N and K simultaneously ($\alpha = \frac{K}{N}$ fixed) would also lead to an increased T_c (See Fig. 8). However, allowing N to scale, while holding a fixed high utilization p, is only moderately expensive as demonstrated in Fig. 10 as long as K stays moderately small.

IV. Conclusions and Discussion

In this paper we considered spectrum enforcement in cognitive-radio systems. We developed a hierarchical coding architecture with enforced silences that allows a primary user to both identify the source of interference and assign the blame to the guilty set of cognitive radios if harmful interference turns out to be due to spectrum violation. This architecture only uses a minimal set of rules by certifying each cognitive user with a particular code which is published before/as the user joins the network. The goal is to understand what fraction of potential spectrum use must be given up to support enforceability in the form of being able to find violators.

It was shown that given a willingness to wait long enough, a high potential secondary utilization can be achieved since even a small proportion of silence slots would be sufficient to discern whether the source of interference was secondary users. Then it was shown that the identities of the specific violators could be identified within detection-time constraints and necessary conditions were provided that must hold even under idealized conditions. We identified fundamental tradeoffs between the various system parameters: secondary utilization, time to conviction, the size of the culprits set and the total number of cognitive systems. For an idealized deterministic model, it was shown that superimposed codes play a major role for unique identification of guilty parties. In future work we consider stochastic models for packet drops and users' transmissions. For further details we refer the reader to [30].

This is only the beginning of this line of investigation, but it is an important first step. The results here show it might not be possible to achieve everything we want simultaneously. This suggests that the usage of a gradual punishment hierarchical scheme might be necessary to satisfy primary QOS concerns. In such a scheme, technical conviction might start with automatic short time-scale technical penalties — allowing a higher false conviction rate — and then extend to legal sanctions[11] when a node is proven to be malfunctioning over longer durations. By allowing false alarms, this will in turn decrease the effective number of users over the short term epochs. Users who are not guilty will prove innocent and would only incur short-term false penalties. However, it should still be possible to satisfy a high legal standard of conviction for long-term violators.

V. Acknowledgements

This work was supported in part by PECASE grant no. N00014-02-100362, NSF CAREER award ECS-0449194, the MIT-Portugal program, Sumitomo Electric, and NSF Grants CCF-0430983, CNS-0435353, ANI-0326503, CCF-0635372, and CNS-0627161. The authors would also like to thank the anonymous reviewers for their comments, Prof. Mark Karpovsky at Boston University for helpful discussions, and the students (especially Kristen Ann Woyach, Rahul Tandra, Pulkit Grover, and Hari Palaiyanur) at Wireless Foundations and the Berkeley Wireless Research Center for their useful feedback.

Appendix

Appendix A: Extra details for Section II

For a sufficiently large sample size, in particular for $\gamma T\theta_0(1-\theta_0) \geq 5$ and $(1-\gamma)T\theta_1(1-\theta_1) \geq 5$ the sample proportions are closely approximated by Gaussian distributions [34]. We then obtain a Gaussian approximation for the difference, Δ, between the two sample populations:

$$\widehat{\Delta} = \widehat{\theta}_1 - \widehat{\theta}_0 \sim N\left(\theta_1 - \theta_0, \frac{\theta_1(1-\theta_1)}{(1-\gamma)T} + \frac{\theta_2(1-\theta_2)}{\gamma T}\right). \quad \text{(A.1)}$$

[11] Injunctions and lawsuits

Conditioning on H_0, i.e. when $\theta_1 = \theta_0 = \theta$:

$$p(\widehat{\Delta}|H_0) \approx N\left(0, \theta(1-\theta)\left(\frac{1}{\gamma T} + \frac{1}{(1-\gamma)T}\right)\right). \quad \text{(A.2)}$$

However, since θ is unknown we plug-in its maximum-likelihood(ML) estimate under the H_0. The ML estimate is given by:

$$\widehat{\theta} = \frac{\sum_{i=1}^{\gamma T} X_i + \sum_{j=1}^{(1-\gamma)T} Y_i}{T} \quad \text{(A.3)}$$

where X_i and Y_i are Bernoulli random variables representing packet drops during canary and non canary slots, respectively. This results in the following normalized test statistic under H_0, i.e.

$$\frac{\widehat{\Delta}}{\sqrt{\widehat{\theta}(1-\widehat{\theta})\left(\frac{1}{\gamma T} + \frac{1}{(1-\gamma)T}\right)}}. \quad \text{(A.4)}$$

Standard computations following [34] lead to Eq. (2) and the results in Section II.

Appendix B: Proof of Theorem 3.1

Define S_ℓ to be the set of all possible observation vectors if the size of the culprits set is exactly ℓ, i.e. $S_\ell = \{\mathbf{y} : |G| = \ell\}$. This is equivalent to the Boolean sum of exactly ℓ distinct vectors from the superimposed code matrix. From the coverage property of superimposed codes one can show that:

$$S_i \cap S_k = \phi \quad \forall\ i,k : 1 \leq i \leq k \leq K+1. \quad \text{(B.1)}$$

Hence, it must be that the sequence of sets $S_k, k = 1 \ldots K+1$ are disjoint. It is also true that $S_k, k = 1 \ldots K$ have exactly $\binom{N}{k}$ *d*ifferent codewords. This can be argued by considering a situation where the set S_K has a duplicate codeword. Then there exist two sets of vectors $x_1, \ldots x_K$ and $y_1, \ldots y_K$ such that their Boolean sum is the same. but then one could form a codeword from the Boolean sum of $K+1$ vectors $y_i \& (x_1 \ldots x_K)$ (and hence $\in S_{K+1}$) which is equal to a codeword in S_K (**Contradiction**). Given disjointness between sets of size less than or equal to K and differentiability within the set it is clear that these codes could be used for unique identification of the culprits set as long as $|G| \leq K$. In other words, if the used code C is ZFD_K then the packet-drop function in Eq. 3 is a bijection function whose domain is all the possible choices for the guilty set G ($\binom{N}{k}, k = 1, \ldots, K$) and with range represented by the extended sum sets $S_1, \ldots, S_K$.

Appendix C: Proof of Theorem 3.2

For exact recovery of the vector $g = \{g_i\}_{i=1}^N$ the probability of error, defined as $P_e = \Pr[\hat{g(y)} \neq g]$, must be zero[12] for all sets G of size K or equivalently for all vectors of indices g which are K-sparse. To account for the worst case, the probability of

[12] This is a consequence of our simplistic deterministic model of interference in which it is meaningless to accept anything less. For more stochastic models of interference, Fano's inequality must be invoked.

error must then be zero for all possible probability distributions π_g of the vector g, i.e.,

$$\max_{\pi_g} \sum_g \pi_g \Pr[\hat{g(y)} \neq g] = 0. \tag{C.1}$$

In the following we derive a necessary condition on the time to conviction T_c when g is uniformly distributed over K-sparse vectors. This is consequently a necessary condition for the maximizing distribution in Eq. (C.1) as well. For any fixed C (with ℓ ones per row) the following holds if g is uniformly distributed:

$$\begin{aligned} H(\mathbf{g}) &= \log \binom{N}{K} \\ &\overset{(a)}{=} H(\mathbf{g}|C) - H(\mathbf{g}|\mathbf{y}, C) \\ &= I(\mathbf{g}; \mathbf{y}|C) \\ &= H(\mathbf{y}|C) - H(\mathbf{y}|\mathbf{g}, C) \overset{(b)}{=} H(\mathbf{y}|C). \end{aligned} \tag{C.2}$$

(a) follows from the fact that the code generation is independent of g. Moreover, the requirement of exact recovery of g given the observation vector y and the code matrix C implies that $H(\mathbf{g}|\mathbf{y}, C) = 0$.
(b) Since $H(\mathbf{y}|\mathbf{g}, C) = 0$ by the determinism of this simplified interference model.

Now we compute the quantity $H(\mathbf{y}|C)$ for any given choice of the code C with exactly ℓ ones per row:

$$\begin{aligned} &H(\mathbf{y}|C) = H(y_1, \ldots, y_{T_c}|C) \\ &\overset{(1)}{\leq} \sum_{i=1}^{T_c} H(y_i|C) \\ &\overset{(2)}{=} \sum_{i=1}^{T_c} H(y_i|c_i) \\ &\overset{(3)}{=} T_c H(y_1|c_1) \\ &= T_c[-p(y_1{=}0|c_1) \log p(y_1{=}0|c_1) - p(y_1{=}1|c_1) \log p(y_1{=}1|c_1)] \\ &= T_c[-\Pr(c_1 g{=}0) \log \Pr(c_1 g{=}0) - \Pr(c_1 g{>}0) \log \Pr(c_1 g{>}0)] \\ &= T_c \left[-\frac{\binom{N-\ell}{K}}{\binom{N}{K}} \log \frac{\binom{N-\ell}{K}}{\binom{N}{K}} - \left(1 - \frac{\binom{N-\ell}{K}}{\binom{N}{K}}\right) \log \left(1 - \frac{\binom{N-\ell}{K}}{\binom{N}{K}}\right) \right] \end{aligned} \tag{C.3}$$

where c_i is the i-th row of the code matrix C.
(1) is obtained by the independence bound of entropy
(2) follows from the fact that y_i depends on C only through the corresponding row c_i since $y_i = u(c_i g)$, where $u(.)$ is the unit step function.
(3) The probability mass function of $y_i, i = 1, \ldots, T_c$ conditioned on c_i is independent of i since g is uniformly distributed and c_i has exactly ℓ ones $\forall i$.[13]

[13]Note that $H(y_i|c_i)$ only depends on the conditional probability function $p(y_i|c_i)$ since c_i is fixed not random, namely $H(y_i|c_i) = -\sum_{y_i} p(y_i|c_i) \log p(y_i|c_i)$.

The final line comes from a simple calculation of the number of ways that a 0 can occur out of the total number of possibilities. From the equations above it is clear that the total time to conviction is lower-bounded by:

$$\begin{aligned} T_c &\geq \frac{\log \binom{N}{K}}{-\frac{\binom{N-\ell}{K}}{\binom{N}{K}} \log \frac{\binom{N-\ell}{K}}{\binom{N}{K}} - \left(1 - \frac{\binom{N-\ell}{K}}{\binom{N}{K}}\right) \log \left(1 - \frac{\binom{N-\ell}{K}}{\binom{N}{K}}\right)} \\ &= \frac{\log \binom{N}{K}}{H(\eta)} \doteq \frac{NH(\alpha)}{2^{-N\xi}} = N \cdot 2^{N\xi} H(\alpha) \end{aligned} \tag{C.4}$$

where $\alpha = \frac{K}{N}$ and $\eta = \frac{\binom{N-K}{\ell}}{\binom{N}{\ell}} = \frac{\binom{N-\ell}{K}}{\binom{N}{K}}$ and $\xi = H(\frac{\ell}{N}) - (1-\alpha)H(\frac{\ell}{N(1-\alpha)})$. This is a necessary condition on the time to conviction for any code matrix C with ℓ non zero entries per row since it has to hold for the uniform distribution and consequently for the maximizing distribution in Eq. (C.1).

Appendix D: Proof of Lemma 3.3

The proof follows closely the proof in [33] but in this case for bipartite graphs.

$$\begin{aligned} &\Pr(D \text{ is not a code}) \leq \sum_{X,Y \in I \bigcap \mathcal{M}} Pr[N(X) = N(Y)] \\ &\leq \sum_{X,Y} \prod_{z \in A} \Pr[\{z \in N(X) \bigcap N(Y)\} \bigcup \{z \notin N(X) \bigcup N(Y)\}] \end{aligned} \tag{D.1}$$

WLOG $|X| \leq |Y|$. Since we need only consider maximal pairs in $\mathcal{M}$ we only need to look at the following two cases:
Case 1: $X \subset Y$. Then $|X| = K-1$, $|Y| = K$ and obviously $N(X) \bigcap N(Y) = N(X \bigcap Y) = N(X)$. In this case:

$$\begin{aligned} &\Pr[\{z \in N(X) \bigcap N(Y)\} \bigcup \{z \notin N(X) \bigcup N(Y)\} / X \subset Y, \\ &X \& Y \in \mathcal{M}] = 1 - (1-p)^{K-1} + (1-p)^K = 1 - p(1-p)^{K-1}. \end{aligned} \tag{D.2}$$

Case 2: X is not a subset Y. Then $|X| = |Y| = K$. In this case the upper bound is maximized if the intersection of the sets is maximal, i.e. $X \cap Y = K-1$. Thus to satisfy $N(X) = N(Y)$, it must be that any $z \in A$ is whether in $N(X \cap Y)$, or otherwise in $N(X \backslash Y)$ as well as $N(Y \backslash X)$ or not in $N(X \cup Y)$. The probability of this happening is:

$$\begin{aligned} &\Pr[\{z \in N(X) \bigcap N(Y)\} \bigcup \{z \notin N(X) \bigcup N(Y)\} / |X| = |Y| = K] \\ &\leq 1 - (1-p)^{K-1} + (1-p)^{K+1} + (1-p)^{K-1}(1-(1-p))^2 \\ &= 1 - 2p(1-p)^K. \end{aligned} \tag{D.3}$$

Thus replacing in Eq. (D.1) we get the upper bound:

$$\begin{aligned} &\Pr(D \text{ is not a code}) \\ &\leq N^{2K} (\max\{1 - 2p(1-p)^K, 1 - p(1-p)^{K-1}\})^{|D|} \\ &= N^{2K} (1 - \min\{p, 2p(1-p)\}(1-p)^{K-1})^{|D|}. \end{aligned} \tag{D.4}$$

REFERENCES

[1] R. H. Coase, "The federal communications commision," *The Journal of Law and Economics*, vol. 2, pp. 1–40, Oct. 1959.

[2] A. De Vany, R. D. Eckert, C. T. Meyers, D.J. O'Hara, and R. C. Scott, "A property system for market allocation of the electromagnetic spectrum: A legal-economic-engineering study," *Stanford Law Review*, vol. 3, pp. 145–162, 1969.

[3] E. Goodman, "Spectrum rights in the telecosm to come," *San Diego Law Review*, vol. 41, pp. 269–404, 2004.

[4] Joseph Mitola, *Cognitive Radio: an integrated agent architecture for software defined radio*, Ph.d. thesis, KTH Royal Inst. of Tech., Stockholm, Sweden, 2000.

[5] FCC, "FCC spectrum policy task force report 04-113," *[Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/*, May 2004.

[6] Rahul Tandra, Shridhar Mubaraq Mishra, and Anant Sahai, "What is a spectrum hole and what does it take to recognize one?," *Accepted to the Proceedings of the IEEE, special issue on Cognitive Radio*, 2008.

[7] Moe Z. Win and Robert A. Scholtz, "Ultra-wide bandwidth time-hopping spread-spectrum impulse radio for wireless multiple-access communications," *IEEE Trans. Commun.*, vol. 48, no. 4, pp. 679–689, Apr. 2000.

[8] Michael Gastpar, "On capacity under receive and spatial spectrum-sharing constraints," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 471–487, Feb. 2007.

[9] Natasha Devroye, Patrick Mitran, and Vahid Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[10] Natasha Devroye, Patrick Mitran, and Vahid Tarokh, "Limits on communications in a cognitive radio channel," *IEEE Commun. Mag.*, vol. 44, no. 6, pp. 44–49, June 2006.

[11] Pulkit Grover and Anant Sahai, "On the need for knowledge of the phase in exploiting known primary transmissions," in *Proceedings of the IEEE DySpAN*, Dublin, Ireland, Apr. 2007, pp. 462–471.

[12] Anant Sahai, Niels Hoven, and Rahul Tandra, "Some fundamental limits on cognitive radio," in *Forty-second Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Oct. 2004, IEEE.

[13] Rahul Tandra and Anant Sahai, "SNR walls for signal detection," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 2, no. 1, pp. 4–17, Feb. 2008.

[14] Shridhar M. Mishra, Anant Sahai, and Robert W. Brodersen, "Cooperative sensing among cognitive radios," in *Proceedings of the International Conference on Communications (ICC)*. IEEE, June 2006, vol. 4, pp. 1658–1663.

[15] George Atia, Erhan Ermis, Shuchin Aeron, and Venkatesh Saligrama, "Robust energy efficient cooperative spectrum sensing in cognitive radios," in *Proceedings of the forty-fifth Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept. 2007.

[16] Rahul Tandra and Anant Sahai, "Fundamental limits on detection in low SNR under noise uncertainty," in *WirelessCom 05 Symposium on Signal Processing*, June 2005, vol. 1, pp. 464–469.

[17] Anant Sahai, Kristen A. Woyach, George Atia, and Venkatesh Saligrama, "Crime and punishment for cognitive radios," in *Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept. 2008.

[18] Kristen A. Woyach, "Crime and punishement for cognitive radios," Masters thesis, University of California, Berkeley, CA, 2008.

[19] Gerald R. Faulhaber, "The future of wireless telecommunications: spectrum as a critical resource," *Information Economics and Policy*, vol. 18, no. 3, pp. 256–271, Sept. 2006.

[20] Dale N. Hatfield, "Measures of spectral efficiency in land mobile radio," *IEEE Trans. Electromagn. Compat.*, vol. EMC-19, no. 3, pp. 266–268, Aug. 1977.

[21] R. Rozovsky and P. R. Kumar, "SEEDEX: a MAC protocol for ad hoc networks," in *MobiHoc '01: Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing*, New York, NY, USA, 2001, pp. 67–75, ACM.

[22] Rahul Tandra and Anant Sahai, "Overcoming SNR walls through macroscale features," in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept. 2008.

[23] K. R. Eberhardt and M. A. Fligner, "A comparison of two tests for equality of two proportions," *The American Statistician*, vol. 31, no. 4, pp. 151–155, Nov. 1977.

[24] Lav R Varshney, "Transporting information and energy simultaneously," in *Proceedings of the IEEE International Symposium on Information Theory*, Toronto, ON, July 2008, pp. 1612–1616.

[25] A. G. Dyachkov and V. V. Rykov, "A survey of superimposed code theory," *Problems of Control and Information Theory*, vol. 12, no. 4, pp. 1–13, 1983.

[26] W. H. Kautz and R.C. Singleton, "Nonrandom binary superimposed codes," *IEEE Trans. Inf. Theory*, vol. 10, no. 4, pp. 363–377, Oct. 1964.

[27] A. G. Dyachkov, "Error probability bounds for two models of randomized design of screening experiments," *Problems of Information Transmission*, vol. 15, no. 4, pp. 17–31, 1979.

[28] M. Taube, "Superimposed coding for data storage," Tech. Rep. 15, Documentation, Inc., Washington, D. C, Sept. 1956.

[29] Shakir Abdul-Jabbar and Peter de Laval, "Constant weight codes for multiaccess channels without feedback," in *8th European Conference on Area Communication, EUROCON 88*, Stockholm, Sweden, June 1988, pp. 150–153.

[30] George Atia, Venkatesh Saligrama, and Anant Sahai, "Codes to unmask spectrum violators," in *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Oct. 2008.

[31] Mark G. Karpovsky, Krishnendu Chakrabarty, and Lev B. Levitin, "On a new class of codes for identifying vertices in graphs," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 599–611, Mar 1998.

[32] E. Charbit, I. Charon, G. Cohen, and O. Hudry, "Discriminating codes in bipartite graphs," *Electronic Notes in Discrete Mathematics*, vol. 26, pp. 29–35, 2006.

[33] A. Frieze, R. Martin, J. Moncel, M. Ruszinko, and C. Smyth, "Codes identifying sets of vertices in random networks," *Discrete Mathematics*, vol. 307, no. 9-10, pp. 1094–1107, May 2007.

[34] E. L. Lehmann, *Testing statistical hypothesis*, John Wiley and Sons, 1986.