This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2015.2421477, IEEE Journal of Selected Topics in Signal Processing

Achieving Undetectable Communication

Seonwoo Lee*, Robert J. Baxley^, Mary Ann Weitnauer*, Brett Walkenhorst†
*Georgia Institute of Technology †Georgia Tech Research Institute ^Bastille
Atlanta, GA 30332

***Abstract*—In this paper we consider the problem of achieving a positive error-free communications rate without being detected by an eavesdropper—we coin this the *privacy rate*. Specifically, we analyze the privacy rate over Additive White Gaussian Noise (AWGN) channels with finite and infinite number of samples and Rayleigh Single Input-Single Output (SISO) and Multiple Input-Multiple Output (MIMO) channels with infinite samples when an eavesdropper employs a radiometer detector and has uncertainty about his noise variance. Leveraging recent results on the phenomenon of a Signal to Noise Ratio (SNR) wall when there is eavesdropper noise power measurement uncertainty, we show that a non-zero privacy rate is possible. We also show that in this scenario, the detector should not necessarily take as many samples as possible.**

I. Introduction

In wireless communications there are several situations where a user would want to communicate such that his emissions are undetectable to other users—that is, transmit with privacy. One emerging example is underlay cognitive radio (CR) [1], where a secondary user seeks to communicate with such low power as to not interfere with or be detected by primary users. Another example is secure communications where a wireless user does not want to reveal his presence in the spectrum to an eavesdropper. Many attacks on wireless networks are predicated on an attacker's ability to determine that a target is transmitting, e.g. [2], [3]. By transmitting with sufficiently low power, we can avoid potential network attacks, and also politely use the spectrum in the presence of primary users. In this paper we determine the achievable communications rate afforded by the privacy constraint under a variety of eavesdropper and channel assumptions.

To formalize our objective, consider a scenario where two users, Alice and Bob, would like to communicate over a wireless channel without being detected by a detector, Dave. Dave's objective is not to decode Alice's transmissions, but merely to detect the presence of Alice's transmissions. If Alice does not want to reveal her position or even her existence, encrypting her communications is not enough. Bash, Goeckel, and Towsley [4] found that if Alice knows a lower bound on the noise power Dave sees, $O(\sqrt{N})$ bits can be sent in N channel uses while guaranteeing that Dave's sum of probability of false alarm P_{FA} and missed detection P_{MD} is asymptotically arbitrarily close to one.

To make this more clear, we define two terms. $I(N)$, which behaves as $O(\sqrt{N})$, is the number of undetected error-free bits that can be sent in N channel uses. Likewise, $C_{pr} = \lim_{N\to\infty} I(N)/N$ is the error-free privacy channel capacity. The result in [4] means that $C_{pr} = 0$ in AWGN channels. While the asymptotic rate is zero, this does not mean no information can be communicated—$I(N)$ is positive so long as the probability of detection is nonzero. Bash, Goeckel, and Towsley's work is the first work that we are aware of that puts information theoretic bounds on low probability of detection communication.

The square root law found in [4] relates to problems in steganography where a fixed-size, finite-alphabet covertext object can be changed to hide a message. Because the covertext object is transmitted noiselessly in steganography, $O(\sqrt{N}\log N)$ bits can be transmitted by modifying $O(\sqrt{N})$ symbols in covertext of size N [5, Ch. 8, Ch. 13]. If we put this in information theory terms of rate over a channel, where covertext of size N is analogous to N channel uses, this is still asymptotically zero rate despite the noiseless transmission because $\lim_{N\to\infty} O(\sqrt{N}\log N)/O(N) = 0$.

However, it is possible to achieve a positive rate when we assume that Dave is uncertain of his noise level, and he uses a radiometer as his detection test. This improves upon the AWGN case with noise power certainty, where positive privacy rate is not possible with a radiometer detector. However, it is important to note that while a radiometer is the optimal detector for AWGN systems where Dave knows his noise variance, a radiometer is not optimal when Dave does not know his noise variance [6]. Thus, the result we present is not as strong as the one in [4], but our result does demonstrate that in practical situations, a positive rate is possible while still guaranteeing that Dave's $P_{MD} + P_{FA} \to 1$.

It is important to distinguish privacy capacity from secrecy capacity, which is the maximum error free rate that Alice can talk to Bob, while preventing an eavesdropper from decoding Alice's transmissions. These two quantities have different constraints, and unfortunately because of this we have not been able to devise a fair metric for which to compare power constrained, secrecy constrained, and privacy constrained capacity.

In this paper we delve in greater detail into the notion of an SNR wall [4], and how Alice can use it to her advantage to communicate without being detected. We also try to estimate what kinds of uncertainty we can reasonably expect and the resultant communications rates that Alice and Bob can achieve over SISO and MIMO AWGN and Rayleigh channels. We use several assumptions of channel information: channel state information (CSI) on the Alice-Bob and Alice-Dave channels and CSI on the Alice-Bob channel and channel distribution information (CDI) on the Alice-Dave channel, as seen in our previous works [7], [8]. Unlike our previous work, in this paper we incorporate for the SISO AWGN channel the fact that Dave does not necessarily want to take into account as many samples as possible.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2015.2421477, IEEE Journal of Selected Topics in Signal Processing

II. Privacy Capacity

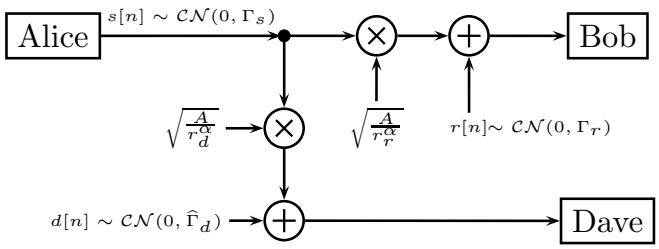

Fig. 1. System block diagram. $\widehat{\Gamma}_d \in [1/\rho\Gamma_d, \rho\Gamma_d]$

A. Parameters and Notation Used

n	Time Index
N	Number of samples Dave observes
$s[n]$	Alice's Signal to Bob
$r[n]$	Bob's noise
$d[n]$	Dave's noise
r_d	Distance from Alice to the detector, Dave
r_r	Distance from Alice to the receiver, Bob
A	Proportionality constant in free space path loss
α	Path loss exponent
Γ_s	Variance of $s[n]$
Γ_d	True variance of $d[n]$
I	The interval $[1/\rho\Gamma_d, \rho\Gamma_d]$
$\widehat{\Gamma}_d$	Uncertain variance of $d[n]$. $\widehat{\Gamma}_d \in I$.
ρ	Characterizes Dave's uncertainty about Γ_d
$\mathcal{CN}(\mu, \Gamma)$	Denotes circularly symmetric complex Gaussian distribution with mean μ and variance Γ
P_D	Probability of detection
P_{MD}	Probability of missed detection
P_{FA}	Probability of false alarm
ξ	Sum of probabilities of detection errors P_{MD} and P_{FA}
R_{pr}	Privacy rate
C_{pr}	Privacy capacity
$Q_{\chi^2_N}(\cdot)$	Tail probability of a X^2_N distribution
$Q(\cdot)$	Tail probability of a Gaussian distribution
γ'	Dave's detection threshold

Consider a communications scenario where Alice has one antenna and transmits a signal $s[n]$ to Bob. At the same time, Dave is a passive detector, listening for Alice's transmissions and trying to determine her presence: whether or not she is transmitting.

We initially assume SISO AWGN Alice-Bob and Alice-Dave channels with complex valued symbols as depicted in Fig. 1. In Section V we assume SISO Rayleigh channels and in Section VI we assume MIMO Rayleigh channels. Bob and Dave experience noise $r[n]$ and $d[n]$, respectively. All of our signals—that is, $s[n]$, $r[n]$, and $d[n]$, are mutually independent. We assume the received signal power, P, is a scaled monomial function of the distance, which is consistent with the free space path loss model where $P \propto 1/r^2$ [9, p. 107], as well as multipath path loss models, where $P \propto 1/r^\alpha$ with α, the path loss exponent, as low as 1.2 and as high as 6.2 [10]. We let $P = A/r^\alpha$ for some proportionality constant A. The uncertainty in Dave's measurement is given by $\widehat{\Gamma}_d \in [(1/\rho)\Gamma_d, \rho\Gamma_d]$, $\rho > 1$, where Γ_d is the true noise and ρ characterizes the uncertainty, as done in [6]. As discussed in section VIII, one source of Dave's noise uncertainty Alice can expect and put reasonable bounds on is thermal noise.

Dave needs to establish the noise level, as he will see in (1). This noise comes from several sources, which include but are not limited to thermal noise in his receiver and environmental noise from his surroundings. The thermal noise can be modeled purely based on temperature. However, the temperature also needs to be measured, and even the most accurate thermometers have an uncertainty range. Also, the environmental noise can be unpredictable, and the only way Dave can attempt to establish the environmental noise is to gather samples. However, he can never be certain that the samples he collected were in the absence of Alice's transmissions. Hence, he may believe that the environmental noise is higher than the true value.

We chose the uncertainty model as $\widehat{\Gamma}_d \in [(1/\rho)\Gamma_d, \rho\Gamma_d]$– the same model that Tandra and Sahai use [6]. We will see later in Section VIII that this geometrically symmetric model does not mesh well with temperature uncertainty from thermometers, which is arithmetically symmetric. However, an arithmetically symmetric model does not allow for the width of the uncertainty interval to ever be greater than 2 Γ_d, making the geometrically symmetric model a more natural choice.

To define privacy capacity, we assume that Dave is trying to distinguish between the following two signal hypotheses,

$$H_0 : x[n] = d[n], \tag{1}$$

$$H_1 : x[n] = \sqrt{\tfrac{A}{r_d^\alpha}} s[n] + d[n], \tag{2}$$

with $n \in \{1, ..., N\}$ and associated probability distributions $P_0(\mathbf{x})$ and $P_1(\mathbf{x})$, respectively.

The privacy capacity C_{pr} is defined as the maximum error free rate at which Alice can talk to Bob, while guaranteeing that

$$\xi = P_{MD} + P_{FA} > 1 - \epsilon \tag{3}$$

for Dave for some arbitrarily small ϵ. It is possible to bound ξ by bounding the total variation distance between $P_0(x)$ and $P_1(x)$, defined as

$$||P_1 - P_0||_1 = \int |P_1(\mathbf{x}) - P_0(\mathbf{x})| d\mathbf{x}. \tag{4}$$

Under the optimal detector for distinguishing $P_1(x)$ from $P_0(x)$ [11, Ch. 13],

$$\xi = 1 - \frac{1}{2}||P_1 - P_0||_1. \tag{5}$$

Hence, if we force $||P_1 - P_0|| < 2\epsilon$, then Dave's $\xi > 1 - \epsilon$.

To find the privacy capacity under noise uncertainty, we would have to find the input distribution such that the rate of information between Alice and Bob is maximized when Dave uses the optimal detector for the input distribution that Alice chooses, while keeping the total variation distance between P_1 and P_0 less than 2ϵ. We leave this as an open problem, and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2015.2421477, IEEE Journal of Selected Topics in Signal Processing

instead solve the problem when we fix the detection test to be an energy detector (radiometer), described by

$$T(x) = \frac{1}{N}x^H x = \frac{1}{N}\sum_{n=1}^{N} x[n]^* x[n] > \gamma', \tag{6}$$

where γ' is the detection threshold of Dave's choosing and N is the number of samples.

Because we have assumed a suboptimal detector for Dave, we are no longer solving for C_{pr} and instead are solving for an achievable privacy rate R_{pr}. The capacity of an AWGN channel is maximized with a Gaussian input distribution, so we assume Gaussian signaling for Alice.

III. SNR Wall

As mentioned in Sec. II, we assume that Dave only knows his noise $\widehat{\Gamma}_d$ is contained to the interval $I = [1/\rho\Gamma_d, \rho\Gamma_d]$ (where ρ would equal one if Dave knew the noise exactly). In this scenario, Tandra and Sahai showed that robust detection of Alice is impossible [6], even if Dave takes an infinite number of samples. In their proof, they derive

$$P_{FA} = \max_{\widehat{\Gamma}_d \in I} Q\left(\frac{\gamma' - \widehat{\Gamma}_d}{\sqrt{\frac{2}{N}}\widehat{\Gamma}_d}\right) \tag{7}$$

$$P_{MD} = 1 - \min_{\widehat{\Gamma}_d \in I} Q\left(\frac{\gamma' - \Gamma_s - \widehat{\Gamma}_d}{\sqrt{\frac{2}{N}}(\widehat{\Gamma}_d + \Gamma_s)}\right), \tag{8}$$

where they have used the Central Limit Theorem (CLT) on the chi square distribution of the test statistic. From this they conclude that Dave faces an SNR wall: if Alice transmits with an SNR below $\rho - 1/\rho$, Dave cannot detect Alice, even if he gathers an infinite number of samples. By maximizing P_{FA} and P_{MD} independently, Dave's true performance is no worse, and with probability 1 better, than if he did not maximize over I. If Dave were to instead just assume one value of $\widehat{\Gamma}_d \in I$, then with probability 1, Dave's assumption about $\widehat{\Gamma}_d$ is incorrect, and his P_{MD} and P_{FA} will be higher than what he calculates.

In Tandra and Sahai's work, when P_{FA} and P_{MD} are maximized independently, $\rho\Gamma_d$ maximizes (7) and $1/\rho\Gamma_d$ maximizes (8). Because it is impossible for $\widehat{\Gamma}_d$ to be $1/\rho\Gamma_d$ and $\rho\Gamma_d$ simultaneously, Dave's detection performance can be improved and remain robust. We instead analyze the scenario that Dave maximizes their sum,

$$\xi' = \min_{\gamma'} \max_{\widehat{\Gamma}_d \in I} \left[P_{FA}(\widehat{\Gamma}_d, \gamma') + P_{MD}(\widehat{\Gamma}_d, \gamma')\right]. \tag{9}$$

Dave performs a min max—for every threshold, he has to maximize ξ' over the uncertainty interval.

It is important to note that ξ' is bounded between 0 and 1, which follows from the fact that $P_D \geq P_{FA}$, where P_D is the probability of detection. A detector can always achieve $P_D = P_{FA}$ by ignoring the input data and flipping a coin with probability of heads being P_D, and declaring a detection when it is heads [12]. Hence any algorithm the detector uses should be able to achieve $P_D \geq P_{FA}$. Additionally, if $P_D < P_{FA}$, then the detector can simply switch what he declares a detection and a non-event.

IV. AWGN Channel Privacy Rate with Measurement Uncertainty

Alice can achieve a rate of $\log_2(1+\Gamma_s/\Gamma_r)$ bits per channel use when her power is constrained to Γ_s and Bob's noise power is Γ_r [13, Ch 9]. Using this we define the privacy rate

$$R_{pr} = \max_{\Gamma_s:\ \lim_{N\to\infty}\xi'(N,\Gamma_s)=1} \log_2(1 + \tfrac{A}{r_r^\alpha}\tfrac{\Gamma_s}{\Gamma_r}), \tag{10}$$

where $\xi'(N, \Gamma_s)$ is the sum of P_{FA} and P_{MD} after N observations.

A. Privacy Rate

From [7], we know that when Dave uses (9) as his detection metric,

$$\begin{aligned} P_{FA} &= Pr(T(x) > \gamma'; H_0) \\ &= Q_{\chi^2_{2N}}\left(\frac{2N\gamma'}{\widehat{\Gamma}_d}\right) \end{aligned} \tag{11}$$

$$\lim_{N\to\infty} P_{FA} = \begin{cases} 0, & \text{if } \gamma' > \widehat{\Gamma}_d \\ 1, & \text{if } \gamma' < \widehat{\Gamma}_d \end{cases} \tag{12}$$

$$\begin{aligned} P_D &= Pr(T(x) > \gamma'; H_1) \\ &= Q_{\chi^2_{2N}}\left(\frac{2N\gamma'}{\widehat{\Gamma}_d + \frac{A}{r_d^\alpha}\Gamma_s}\right) \end{aligned} \tag{13}$$

$$\lim_{N\to\infty} P_D = \begin{cases} 0, & \text{if } \gamma' > \widehat{\Gamma}_d + \dfrac{A}{r_d^\alpha}\Gamma_s \\ 1, & \text{if } \gamma' < \widehat{\Gamma}_d + \dfrac{A}{r_d^\alpha}\Gamma_s, \end{cases} \tag{14}$$

for some choice of $\widehat{\Gamma}_d \in [1/\rho\Gamma_d, \rho\Gamma_d]$ (where (11) and (13) are the equations for an energy detector). We want to maximize Alice's signal power while forcing $\xi \to 1$, so we can either force $P_D \to 0$ or $P_{FA} \to 1$. To do this we need to satisfy

$$\gamma' < \widehat{\Gamma}_d \tag{15}$$

or

$$\gamma' > \widehat{\Gamma}_d + \frac{A}{r_d^\alpha}\Gamma_s \tag{16}$$

for all γ' and some $\widehat{\Gamma}_d \in [1/\rho\Gamma_d, \rho\Gamma_d]$ while maximizing Γ_s. For $\gamma' < \rho\Gamma_d$, we can choose $\widehat{\Gamma}_d = \rho\Gamma_d$ to satisfy (15). For $\gamma' \geq \rho\Gamma_d$, we can't satisfy (15) but we can satisfy (16) by choosing $\widehat{\Gamma} = 1/\rho\Gamma_d$ and constraining

$$\frac{A}{r_d^\alpha}\Gamma_s < (\rho - 1/\rho)\Gamma_d. \tag{17}$$

Hence, the SNR wall to force $\xi \to 1$ is

$$\Gamma_s = \Gamma_d r_d^\alpha(\rho - \tfrac{1}{\rho})/A. \tag{18}$$

Given this, for Alice to achieve privacy, she should emit less power than (18), resulting in

$$\begin{aligned} R_{pr} &= \lim_{N\to\infty} \log_2(1 + \tfrac{A}{r_r^\alpha}\tfrac{\Gamma_s}{\Gamma_r}) \\ &= \log_2\left(1 + \tfrac{\Gamma_d}{\Gamma_r}\left(\tfrac{r_d}{r_r}\right)^\alpha(\rho - \tfrac{1}{\rho})\right). \end{aligned} \tag{19}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2015.2421477, IEEE Journal of Selected Topics in Signal Processing

B. Lower Bound on SNR Wall

Alice can communicate with a positive rate given by (19) while forcing Dave's detector to have all errors so long as she talks below the SNR wall in (18). Unfortunately, Alice does not know what Dave's uncertainty is, so Alice cannot know with certainty if she is communicating just below the SNR wall to maximize her rate. However, she can lower bound all of the SNR wall parameters under some assumptions.

In most situations there is at least some area in which Alice can be certain that there is no eavesdropper, such as her immediate vicinity or her building. She can use this to lower bound r_d. Dave's noise level depends on the temperature, so Alice can also lower bound ρ by assuming a temperature uncertainty that is less than what is available in highly-accurate thermometers. The noise level Γ_d can also be lower bounded by assuming a temperature in Dave's receiver and some noise figure. The path loss exponent α can be lower bounded as well based on the propagation environment characteristics.

With these lower bounds, Alice can achieve private communication—that is, she can pick a rate $R < R_{pr} = \log_2(1 + \frac{\Gamma_d}{\Gamma_r}\left(\frac{r_d}{r_r}\right)^{\alpha}(\rho - \frac{1}{\rho}))$. Numerical results are discussed in Section VIII.

V. Rayleigh Fading Channel Privacy Rate with Measurement Uncertainty

A. Problem Statement

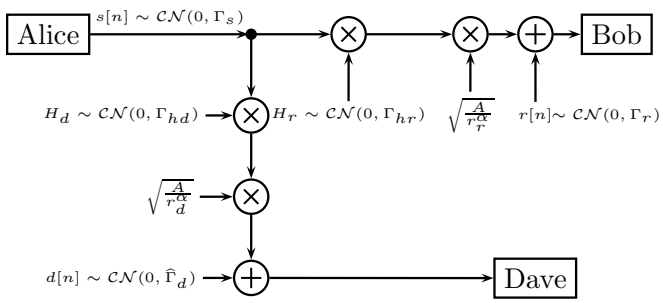

Fig. 2. System block diagram. $\widehat{\Gamma}_d \in [1/\rho\Gamma_d, \rho\Gamma_d]$

We can also apply similar analysis to Rayleigh fading channels with complex valued symbols as depicted in Fig. 2. All other aspects of the scenario are the same as the SISO setup. For simplicity, the channel gains H_d and H_r are assumed to be static over the signaling period.

For detection the two hypotheses are:

$$H_0: \; x[n] = d[n] \tag{20}$$

$$H_1: \; x[n] = \sqrt{\frac{A}{r_d^{\alpha}}} H_d s[n] + d[n]. \tag{21}$$

When Alice has channel state information (CSI) for the Alice-Dave channel, Dave and Alice's objectives are the same as the AWGN case. We use the same strategy to analyze the privacy rate in this scenario. This is an unlikely scenario in practice, but the resulting privacy rate gives us an idea of the best case Alice can hope to achieve. When Alice only has channel distribution information (CDI) for the Alice-Dave channel, Dave's objective is the same as the AWGN case. However, Alice can no longer guarantee that $\xi' \to 1$ because she will not know the instantaneous value of the channel fade. Accordingly, we have to change the constraint in the privacy rate definition to be $\lim_{N\to\infty} E[\xi'(\Gamma_s, N)] \geq 1 - \epsilon$, which is equivalent to requiring that $\lim_{N\to\infty} Pr(\xi'(\Gamma_s, N) = 1) \geq 1 - \epsilon$.

B. Privacy Rate Under Alice-Dave CSI

Under CSI with a static channel gain, the channel is still characterized as a AWGN channel with a known scalar multiplier, so we assume Gaussian signaling for Alice. Dave uses the same detection test as the AWGN case and hence the same detection threshold. The probability of detection is now

$$\begin{aligned} P_D &= Pr(T(x) > \gamma'; H_1), \\ &= Q_{\chi^2_{2N}}\left(\frac{2N\gamma'}{\widehat{\Gamma}_d + \frac{A}{r_d^{\alpha}}|H_d|^2\Gamma_s}\right). \end{aligned} \tag{22}$$

We quickly see that aside from the addition of a new scale factor $|H_d|^2$ everywhere there is $\frac{A}{r_d^{\alpha}}$, our equations for the Rayleigh fading CSI case will be the same as the AWGN case (note the similarity between (13) and (22)). Hence Alice should talk below

$$\Gamma_s|H_d = \frac{\Gamma_d r_d^{\alpha}}{|H_d|^2 A}(\rho - 1/\rho). \tag{23}$$

Using the power level from (23),

$$\begin{aligned} R_{pr}|H_d, H_r &= \lim_{N\to\infty} \log_2(1 + \frac{A}{r_r^{\alpha}}\frac{\Gamma_s}{\Gamma_r}) \\ &= \log_2(1 + \frac{\Gamma_d}{\Gamma_r}\left(\frac{r_d}{r_r}\right)^{\alpha}\frac{|H_r|^2}{|H_d|^2}(\rho - \tfrac{1}{\rho})). \end{aligned} \tag{24}$$

Because $H_d \sim \mathcal{CN}(0, \Gamma_{hd})$ and $H_r \sim \mathcal{CN}(0, \Gamma_{hr})$, we have

$$R_{pr} = \log_2(1 + \frac{\Gamma_d}{\Gamma_r}\left(\frac{r_d}{r_r}\right)^{\alpha}\psi\frac{\Gamma_{hr}}{\Gamma_{hd}}(\rho - 1/\rho)), \tag{25}$$

where $\psi \sim F(2,2)$, that is, an F-distribution. With this, the ergodic rate is

$$\begin{aligned} R_{pr,erg} &= \int_0^{\infty} \log_2(1 + \phi x)\, f_x(x) dx \\ &= \tfrac{\phi}{\phi - 1}\log_2(\phi), \end{aligned} \tag{26}$$

where $\phi = \frac{\Gamma_{hr}}{\Gamma_{hd}}(\rho - \frac{1}{\rho})\frac{\Gamma_d}{\Gamma_r}\left(\frac{r_d}{r_r}\right)^{\alpha}$. This ergodic rate represents the weighted average rate, with the weight being the pdf $f_x(x)$ of the F distributed ψ.

We can also find the outage rate

$$\begin{aligned} Pr(R_{pr} < c) &= Pr(F(2,2) \leq (2^c - 1)\tfrac{1}{\phi}) \\ &= \frac{(2^c - 1)}{(2^c - 1) + \phi}. \end{aligned} \tag{27}$$

C. Analysis of Privacy Rate under CSI

Alice can communicate with a positive rate with zero probability of detection so long as she talks below the power in (23). If we compare the privacy rates of the Rayleigh fading and AWGN channels,

$$Pr(\text{Privacy Rate}_{\text{Rayleigh}} < \text{Privacy Rate}_{\text{AWGN}}) = \tfrac{1}{1 + \frac{\Gamma_{hr}}{\Gamma_{hd}}}.$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2015.2421477, IEEE Journal of Selected Topics in Signal Processing

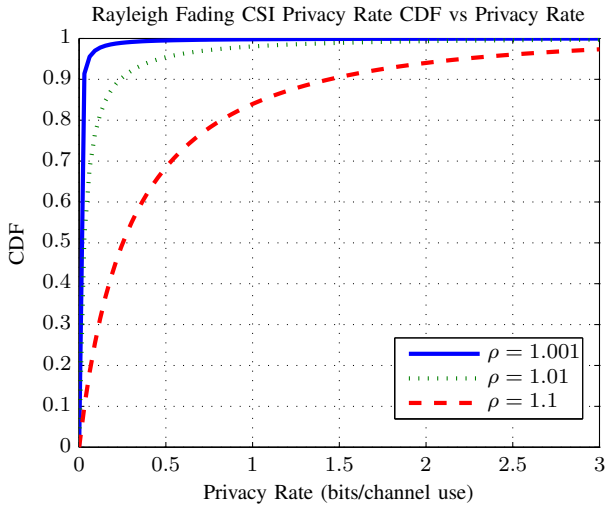

Fig. 3. Rayleigh Outage Rate under CSI for various ρ's. All parameter ratios are 1.

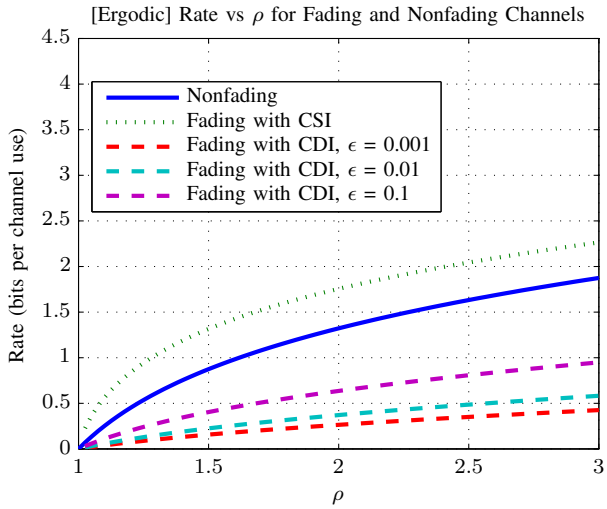

Fig. 4. Comparison of Rate or Ergodic Rate vs ρ. All parameter ratios are 1. The ergodic rate under CDI increases with ϵ.

we see that if the channel gains have identical distributions, then the probability that the Rayleigh fading channel under CSI has a greater privacy rate than the AWGN channel is one half. There is a small probability that the channel gain ratio will be very large in Alice's favor, causing the ergodic privacy rate under CSI to increase over the rate of the AWGN channel. This phenomenon is similar to what occurs in physical layer security - by sending at a high rate when the channel is in Alice's favor, Alice can achieve a higher ergodic secrecy capacity under fading channels than under a AWGN channel [14]. A plot of the outage rate can be found in Fig. 3.

D. Privacy Rate under Alice-Dave CDI

Next we study the privacy rate when only channel distribution information is known about the Alice-Dave channel. We assume CSI for the Alice-Bob channel and $E[s[n]^* H_d] = 0\ \forall n$. Otherwise the system setup is the same as the CSI case. However, Alice can no longer guarantee that $\xi' \to 1$ because she no longer knows the exact value of H_d when she transmits. Hence, we have to modify our definition of privacy rate to

$$\tilde{R}_{pr,\epsilon} = \max_{\lim_{N\to\infty} E[\xi'(\Gamma_s,N)]\geq 1-\epsilon} \log_2(1+\tfrac{A}{r_r^\alpha}\tfrac{\Gamma_s}{\Gamma_r}). \quad (28)$$

While P_{FA} remains the same, P_D now changes to

$$P_D = \begin{cases} 0, & \text{with probability } 1 - Q_{\chi_2^2}\left(\dfrac{\gamma' - \widehat{\Gamma}_d}{\frac{A}{r_d^\alpha}\Gamma_s\Gamma_{hd}/2}\right) \\ 1, & \text{with probability } Q_{\chi_2^2}\left(\dfrac{\gamma' - \widehat{\Gamma}_d}{\frac{A}{r_d^\alpha}\Gamma_s\Gamma_{hd}/2}\right). \end{cases} \quad (29)$$

When we analyze ξ' we can see that the worst case scenario for Alice is when Dave picks $\gamma' = \rho\Gamma_d$, which maximizes P_D. For any $\gamma' < \rho\Gamma_d$, we choose $\rho\Gamma_d$ for the value of $\widehat{\Gamma}_d$. Hence to have $\lim_{N\to\infty} \mathrm{E}[\xi'(N,\Gamma_s)] \geq 1-\epsilon$ or $\lim_{N\to\infty} Pr(\xi'(N,\Gamma_s)=1) \geq 1-\epsilon$, we need

$$\begin{aligned} P_{FA} + P_{MD} &\geq 1-\epsilon \\ 1 - Q_{\chi_2^2}\left(\frac{(\rho-\frac{1}{\rho})\Gamma_d}{\frac{A}{r_d^\alpha}\Gamma_s\Gamma_{hd}/2}\right) &\geq 1-\epsilon. \end{aligned} \quad (30)$$

Thus, to maximize rate under the constraint, Alice should transmit with power

$$\Gamma_s = \frac{(\rho-\frac{1}{\rho})\Gamma_d}{\frac{A}{r_d^\alpha}Q^{-1}_{\chi_2^2}(\epsilon)\Gamma_{hd}/2}.$$

Assuming CSI on the Alice-Bob channel, we have

$$\tilde{R}_{pr,\epsilon}|H_r = \log_2\left(1 + \frac{|H_r|^2}{\Gamma_{hd}/2}\frac{\Gamma_d}{\Gamma_r}\left(\frac{r_d}{r_r}\right)^\alpha \frac{\rho-\frac{1}{\rho}}{Q^{-1}_{\chi_2^2}(\epsilon)}\right). \quad (31)$$

Because $|H_r|^2 \sim \frac{\Gamma_{hr}}{2}\chi_2^2$, we can find the ergodic rate

$$\begin{aligned} \tilde{R}_{pr,\epsilon,erg} &= \int_0^\infty \log_2(1+Bx)\frac{e^{-x/2}}{2}dx \\ &= \frac{1}{\ln(2)}\exp\left(\frac{1}{2B}\right)\mathrm{E}_1\left(\frac{1}{2B}\right) \end{aligned} \quad (32)$$

where $E_1(x) = \int_x^\infty \frac{e^{-t}}{t}dt$ and $B = \left(\frac{r_d}{r_r}\right)^\alpha \frac{\Gamma_d}{\Gamma_r}\frac{\Gamma_{hr}}{\Gamma_{hd}}\frac{\rho-\frac{1}{\rho}}{Q^{-1}_{\chi_2^2}(\epsilon)}$.

We can also find the outage rate

$$\begin{aligned} Pr(\tilde{R}_{pr,\epsilon} \leq c) &= Pr\left(\chi_2^2 \leq (2^c-1)\frac{1}{B}\right) \\ &= 1 - Q_{\chi_2^2}\left((2^c-1)\frac{1}{B}\right). \end{aligned} \quad (33)$$

E. Comparison of Privacy Rates Under Different Channels

A plot of the privacy rates can be found in Fig. 4 with all parameter ratios set to one (that is, $\frac{\Gamma_{hr}}{\Gamma_{hd}} = \frac{\Gamma_r}{\Gamma_d} = \frac{r_r}{r_d} = 1$). As we previously observed the ergodic privacy rate of a Rayleigh channel under CSI is greater than that of a AWGN channel because of the small probability of having a channel gain ratio in Alice's favor. We also observe that the ergodic privacy rate for a Rayleigh channel under CDI is lower than that of an AWGN channel, with only small increases in privacy rate for orders of magnitude increases in ϵ.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2015.2421477, IEEE Journal of Selected Topics in Signal Processing

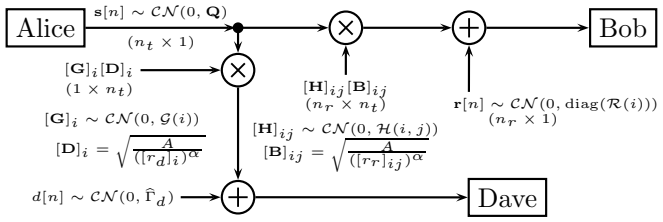

Fig. 5. The circle multiplication symbols denote matrix multiplication.

VI. MIMO Rayleigh Privacy Rate

We also extend our results to MIMO Rayleigh fading channels with complex valued symbols (Fig. 5). We also assume that while Alice and Bob have multiple antennas, Dave has only one antenna.

Let a bolded quantity represent a vector or matrix. Let $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Xi})$ denote a vector of circularly symmetric complex jointly Gaussian random variables with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Xi}$. Let n_t and n_r denote the number of transmit and receive antennas, respectively. Let the $[v]_i$ operator be the ith entry of a vector v, and let the $[M]_{ij}$ operator be the row i, column j entry of a matrix M. Let $\mathcal{H}$ denote the set of all variances of the matrix $\mathbf{H}$, with $\mathrm{Var}([\mathbf{H}]_{ij}) = \mathcal{H}(i,j)$. Let the diag operator denote a diagonal matrix with the diagonal entries given by the argument.

Alice sends signal $\mathbf{s}[n]$ at time index n. Bob and Dave experience noise $\mathbf{r}[n]$ and $d[n]$, respectively. Bob's jth antenna is located $[r_r]_{ij}$ away from Alice's ith antenna, and Dave's antenna is located $[r_d]_i$ away from Alice's ith antenna. Bob and Dave experience channel gains $\mathbf{H}$ and $\mathbf{G}$, respectively. We denote the diagonal entries of $\mathbf{Q}$, the covariance matrix of our signal $\mathbf{s}[n]$, as $\mathcal{S}(i)$. For simplicity, the channel gains $\mathbf{H}$ and $\mathbf{G}$ are assumed to be static over the signaling period.

Dave's detection hypotheses are

$$H_0: \; x[n] = d[n] \tag{34}$$

$$H_1: \; x[n] = \sum_{i=1}^{n_t} \sqrt{\frac{A}{([r_d]_i)^\alpha}} [\mathbf{G}]_i [\mathbf{s}[n]]_i + d[n]. \tag{35}$$

Alice's objective is to find the maximum error-free rate at which she can communicate to Bob while forcing $\xi \geq 1 - \epsilon$.

A. Privacy Rate under Alice-Dave Channel Distribution Information (CDI)

We assume Dave uses the same detection test as before (6). Let $L[n] = \sum_{i=1}^{n_t} \sqrt{\frac{A}{([r_d]_i)^\alpha}} [\mathbf{G}]_i [\mathbf{s}[n]]_i + d[n]$, and let $l = \sum_{i=1}^{n_t} \frac{A}{([r_d]_i)^\alpha} |[\mathbf{G}]_i|^2 \Gamma_{si} + \widehat{\Gamma}_d$. Therefore $L[n] \sim \mathcal{CN}(0, l)$. Then we can find Dave's detection probability

$$\begin{aligned} P_D &= Pr\left(\frac{1}{N} \sum_{n=1}^{N} (L[n]^* L[n]) > \gamma' \right) \\ &= Q_{\chi^2_{2N}} \left(\frac{2N\gamma'}{\sum_{i=1}^{n_t} \frac{A}{([r_d]_i)^\alpha} |[\mathbf{G}]_i|^2 \mathcal{S}(i) + \widehat{\Gamma}_d} \right). \end{aligned} \tag{36}$$

Dave's asymptotic P_D and P_{FA} are [7]

$$\lim_{N\to\infty} P_{FA} = \begin{cases} 0, & \gamma' > \widehat{\Gamma}_d \\ 1, & \gamma' < \widehat{\Gamma}_d \end{cases} \tag{37}$$

$$\lim_{N\to\infty} P_D = \begin{cases} 0, & \gamma' > \sum_{i=1}^{n_t} \frac{A}{([r_d]_i)^\alpha} |[\mathbf{G}]_i|^2 \mathcal{S}(i) + \widehat{\Gamma}_d \\ 1, & \gamma' < \sum_{i=1}^{n_t} \frac{A}{([r_d]_i)^\alpha} |[\mathbf{G}]_i|^2 \mathcal{S}(i) + \widehat{\Gamma}_d \end{cases} \tag{38}$$

For Dave to robustly detect Alice Dave should choose the γ' that maximizes ξ'. Forcing $\xi' \to 1$ is equivalent to forcing $P_D \to 0$ for $\widehat{\Gamma}_d = \rho \Gamma_d$ [7]. However, we can only lower bound $Pr(\xi' \to 1)$ because under CDI the $[\mathbf{G}]_i$'s are random. Hence

$$Pr\left(\sum_{i=1}^{n_t} \frac{A}{([r_d]_i)^\alpha} |[\mathbf{G}]_i|^2 \mathcal{S}(i) < (\rho - \tfrac{1}{\rho}) \Gamma_d \right) \geq 1 - \epsilon. \tag{39}$$

Ideally we would use a generalized chi square distribution ($|[\mathbf{G}]_i|^2$ are χ^2_2 distributed) and calculate the set $\mathcal{S}$ that satisfies (39). However, we are unable to find an analytical solution. Instead, we use the Lyapunov Central Limit Theorem (LCLT) for an approximate analytical solution (see Section VI-A1), and also compute the constraint numerically (see Section VI-A2).

Once we have the set of valid power allocations,

$$R_{pr} = \max_{\substack{\mathbf{Q}: \; \mathcal{S} \text{ satisfies (39)}, \; [\mathbf{Q}]_{ii} \leq S(i) \; \forall i, \\ \mathbf{Q} \text{ positive semidefinite}}} \log_2 |\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^H|. \tag{40}$$

1) Analytic Solution to Privacy Rate under Alice-Dave CDI: For this solution we assume $[r_d]_i = r_d, [\mathbf{G}]_i = \Gamma_g, \mathcal{R}(i) = \Gamma_r, [r_r]_{ij} = r_r, \mathcal{H}(i,j) = \Gamma_h \; \forall i,j$. These parameter uniformity assumptions allow us to use the Marchenko-Pastur (MP) law [15]. We also assume $n_t = n_r = \tilde{n}$, but these results can be generalized to $n_t \neq n_r$.

We use the LCLT, which unlike the classical CLT allows for the random variables to not be identically distributed but requires some extra bounds on their means and variances. The LCLT allows us to avoid the problem of writing the inverse tail of a generalized chi square distribution $Q^{-1}_{\chi^2_2;\mathcal{S}}(\cdot)$, where the function itself depends on $\mathcal{S}$, the values we are trying to solve for. By applying the LCLT to (39),

$$\sum_{i=1}^{\tilde{n}} \frac{A}{r_d^\alpha} \Gamma_g \mathcal{S}(i) + Q^{-1}(\epsilon) \sqrt{\sum_{i=1}^{\tilde{n}} \left(\frac{A}{r_d^\alpha} \Gamma_g \mathcal{S}(i) \right)^2} \leq (\rho - \tfrac{1}{\rho}) \Gamma_d. \tag{41}$$

The combination of the LCLT's $\tilde{n} \to \infty$ assumption with the following constraint results in a good approximation of privacy rate, as we will see in Section VI-A2.

To simplify (41), we use the norm property that for $a_i \geq 0$,

$$\sqrt{\textstyle\sum_i a_i^2} \leq \textstyle\sum_i a_i, \tag{42}$$

giving us the new constraint function

$$\sum_{i=1}^{n_t} \mathcal{S}(i) \leq \frac{(\rho - 1/\rho)\Gamma_d}{(1 + Q^{-1}(\epsilon)) \frac{A}{r_d^\alpha} \Gamma_g}. \tag{43}$$

To see (42), observe that the unit ball described by setting the right hand side (RHS) of (42) to one is a strict subset of the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2015.2421477, IEEE Journal of Selected Topics in Signal Processing

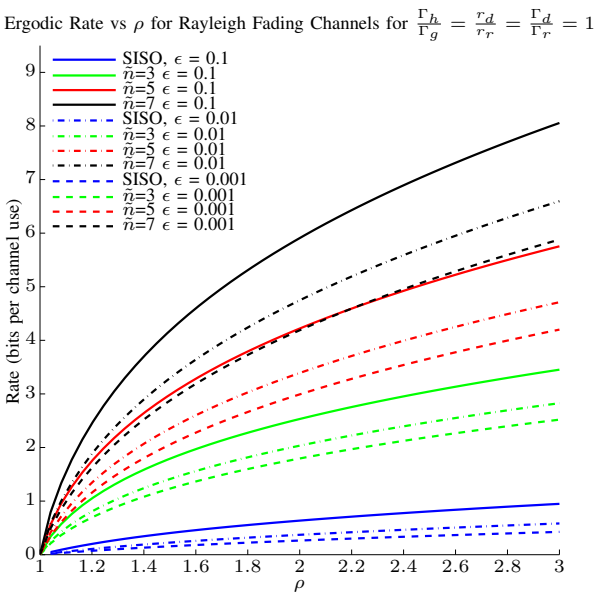

Fig. 6. Privacy rates vs ρ under the LCLT model. Rates increase with ϵ and number of antennas

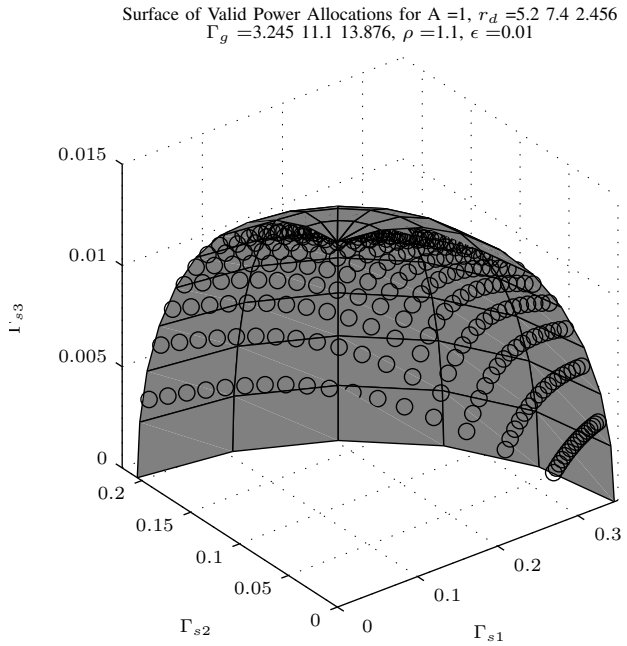

Fig. 7. Boundary surface of valid power allocations for arbitrary A, ϵ, ρ, Γ_g, r_d. All points below the surface in the first octant are also valid. The discrete points are the true boundary, whereas the background surface is an ellipsoid.

unit ball described by setting the left hand side of (42) to one. Hence by using the RHS, we have restricted the valid set of power allocations that we are maximizing over.

From this point forward we use the MP distribution. The MP law tells us the distribution of the eigenvalues of a matrix $\mathbf{JJ^H}$ when $[\mathbf{J}]_{ij} \sim \mathcal{CN}(0,1)$ [15]. The parameter uniformity assumptions allow us to write our new channel matrix $\tilde{\mathbf{H}} = \sqrt{\Gamma_h}\mathbf{JJ}^H$. If the distribution of the eigenvalues for the general $\mathbf{H}$ were known, that distribution could be used.

If we take the singular value decomposition (SVD) of $\tilde{\mathbf{H}} = \sqrt{\Gamma_h}\mathbf{U\Sigma V}^H$ where $\mathbf{\Sigma} = \mathrm{diag}(\sigma_i)$ and let $\mathbf{Q} = \mathbf{VSV}^H$ where $\mathbf{S} = \mathrm{diag}(\frac{\mathcal{S}(i)}{\Gamma_r})$ then our privacy rate approximation is

$$R_{pr,CLT} = \max_{\substack{\mathcal{S}: \ \mathcal{S}(i)\geq 0 \ \forall i, \\ \mathcal{S} \text{ satisfies (43)}}} \sum_{i=1}^{\tilde{n}} \log_2\left(1 + \sigma_i^2 \frac{\Gamma_h A \mathcal{S}(i)}{r_r^\alpha \Gamma_r}\right) \quad (44)$$

where $\sigma_i^2 = \lambda_i$ are the eigenvalues of $\mathbf{JJ}^H$. Our numerical solution in VI-A2 considers the off diagonal elements of $\mathbf{Q}$.

By using Lagrange multipliers and Kuhn-Tucker conditions, we can find the optimal power allocation as

$$\mathcal{S}(i) = \left(\frac{\theta}{(1+Q^{-1}(\epsilon))\frac{A}{r_d^\alpha}\Gamma_g} - \frac{\Gamma_r r_r^\alpha}{A\Gamma_h \lambda_i}\right)^+ \quad \forall i \quad (45)$$

where θ is a water filling parameter. The familiar water filling solution follows from the fact that applying the LCLT and (42) changes our constraint function (43) to a total power constraint.

While the eigenvalue distribution of $\tilde{\mathbf{H}}$ converges asymptotically with the number of antennas, it converges very quickly. By using Equations 15, 19, 20, and 21 in [15] with

$$P_0 = \frac{\rho - 1/\rho}{1+Q^{-1}(\epsilon)} \frac{\Gamma_d}{\Gamma_r} \frac{\Gamma_h}{\Gamma_g} \left(\frac{r_d}{r_r}\right)^\alpha \quad (46)$$

we get an analytical approximation of the privacy rate.

If $n_t \neq n_r$, the rate bound can be found numerically by evaluating (15) in [15]. The privacy rates are plotted in Fig. 6 for 1, 3, 5, and 7 antennas, with $\epsilon = 0.001, 0.01,$ and 0.1.

It is important to note that there is no SNR term in the privacy rate. By solving for the maximum allowable power given the detector noise uncertainty and the ratios of parameters, we eliminate SNR.

This privacy rate differs to that found in [16]. Hero derives $C_{pr} = \mathrm{E}\left[\log\left(\frac{1}{2}\sqrt{1+\mu\lambda_i^2}\right)\right]$ where λ_i are the eigenvalues of $\mathbf{H}^H\mathbf{H}$ and μ is a water-filling parameter. However, the low probability of detection (LPD) constraint in [16] is different from ours— [16] constrains the Chernoff exponent, which limits how quickly Dave's detection errors decay exponentially to zero. This constraint acknowledges that while Dave's detection will be asymptotically perfect with noise power certainty, it is still possible to transmit a finite amount of data with a reasonably high ξ for Dave. Our result differs because we assume noise power uncertainty and a radiometer for Dave.

2) Numerical Solution to Privacy Rate under Alice-Dave CDI:

Again, we are interested in maximizing Alice's rate under the constraint given by (39). By using the generalized χ_2^2 distribution [17], we plot valid power allocations for arbitrary values of $A, \mathcal{G}, \Gamma_d$ and r_d and $\tilde{n} = 3$. Because the rate monotonically increases in power, we are only interested in power allocations at the boundary of our constraint function. The discrete points in Fig. 7 come from (39), whereas the surface plot is that of an ellipsoid, as $(\frac{x}{\mathcal{S}(1)})^2 + (\frac{y}{\mathcal{S}(2)})^2 + (\frac{z}{\mathcal{S}(3)})^2 = 1$, where $\mathcal{S}$ can be found by solving $\mathcal{S}(i) = \frac{(\rho - 1/\rho)\Gamma_d}{\frac{A}{([r_d]_i)^\alpha} Q^{-1}_{\chi_2^2}(\epsilon)\mathcal{G}(i)/2}$, the maximum power allowed for that antenna if only that antenna were used [7]. The model match can be evaluated by calculating the average of $(\frac{x}{\mathcal{S}(1)})^2 + (\frac{y}{\mathcal{S}(2)})^2 + (\frac{z}{\mathcal{S}(3)})^2$,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2015.2421477, IEEE Journal of Selected Topics in Signal Processing

which is approximately 0.9 for the plotted values and for thirty other sets of arbitrarily chosen parameters. Additionally, all the points are strictly interior to the corresponding ellipsoid. As a side note, consider that the constraint surface for total power-constrained MIMO is a plane in the first hyperoctant.

When just accounting for noise from temperature, we have a low resultant transmit power, as we will see in Section VIII. Under the traditional sum power constraint, maximizing MIMO capacity at low SNR involves beamforming. The optimal beamforming covariance matrix is $\mathbf{Q} = P\mathbf{v}\mathbf{v}^H$, where P is the power constraint and $\mathbf{v}$ is the right singular vector of $\mathbf{H}$ that corresponds to its largest singular value. We can employ this same method for the MIMO privacy rate. However it is important to note that while precoding with the right singular vectors is optimal under the sum power constraint, it is not optimal under a per-antenna power constraint [18]. Finding the privacy rate can be reformulated as finding the maximum of the capacities with per-antenna power constraints for each valid power allocation in Fig. 7. The advantage of using the right singular vectors is that it is computationally inexpensive - it only involves finding the SVD of $\mathbf{H}$ and then scaling the vector out to the boundary of the valid power allocation surface. Additionally, the beamforming approach does not require the parameter uniformity assumptions, unlike the LCLT approach.

By using only one eigenchannel and sending only one symbol $x \sim \mathcal{CN}(0, \sigma_x^2)$, we precode $\hat{\mathbf{x}} = \mathbf{v}x$. Defining $\mathbf{\Gamma_s} = (\mathcal{S}(1), \mathcal{S}(2), \ldots, \mathcal{S}(n_t))^T$, our power allocation is $\mathbf{\Gamma_s} = \sigma_x^2 \tilde{\mathbf{v}}$, where $\tilde{\mathbf{v}}$ is the vector such that $[\tilde{\mathbf{v}}]_i = |[\mathbf{v}]_i|^2$. We find the scalar σ_x^2 such that $\mathbf{\Gamma_s}$ is at the boundary of the set of valid power allocations. Having found σ_x^2 and λ_1, the largest eigenvalue of $\mathbf{H}\mathbf{H}^H$,

$$R_{pr,\text{beamforming}} = \log_2(1 + \tfrac{\sigma_x^2 \lambda_1}{\Gamma_r}). \tag{47}$$

We then use a Monte-Carlo simulation to find the ergodic rate. Additionally, we do a brute force search to find the true ergodic privacy rate. In our Monte Carlo simulation, for every realization of $\mathbf{H}$ we discretize the space of valid power allocations, calculate the per antenna power constrained (PAPC) capacity at each allocation [18], and then pick the maximum across all power allocations. Because calculating the PAPC capacity is computationally expensive at low power allocations [18], we also present a lower bound which sets the channel covariance matrix as the diagonal matrix with the per antenna power constraints along the diagonal.

We compare the LCLT, beamforming, grid search, and grid search lower bound methods under the parameter uniformity assumptions (as required by the LCLT) in Fig. 9. We see that at 3 antennas the computationally fast LCLT method provides a good approximation of the privacy rate. However, we see increasing the number of antennas increases the error in the LCLT approximation. There are three factors affecting the error approximation: the use of the LCLT which assumes $\tilde{n} \to \infty$, the use of the MP law which also assumes $\tilde{n} \to \infty$ but converges quite rapidly, and the use of inequality (42). All three factors together combine to result in an approximation that lower bounds the true privacy rate, and becomes worse as the number of antennas increases.

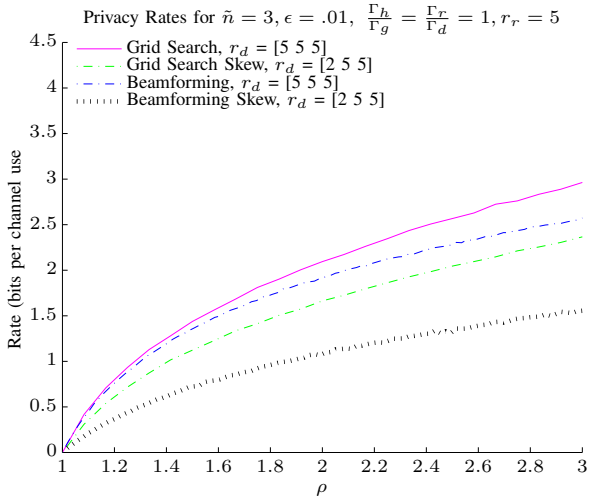

Fig. 8. Privacy rates vs ρ. in skewed vs not skewed

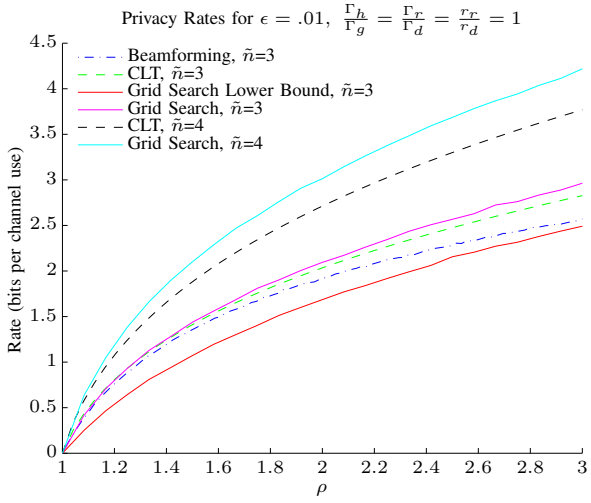

Fig. 9. Comparison of privacy rates vs ρ under different models

The beamforming solution performs well with parameter uniformity, but as the privacy constraint region becomes skewed, the approximation error grows (Fig. 8). With parameter uniformity, the privacy constraint region is symmetric and represents the best case scenario for the beamforming solution, allowing it to perform well despite using only one eigenchannel and the wrong precoding matrix.

We can apply all these results to look at some hypothetical numbers on privacy rates.

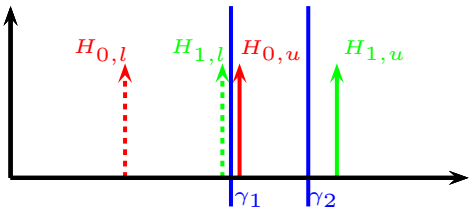

Fig. 10. PDFs of test statistics at infinite samples. The dotted PDFs are those at the lower end of the uncertainty interval

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2015.2421477, IEEE Journal of Selected Topics in Signal Processing

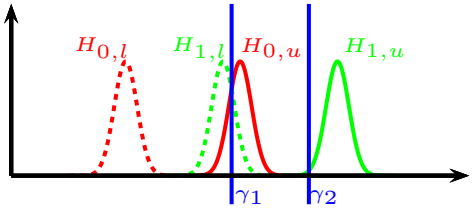

Fig. 11. PDFs of test statistics at finite samples, The dotted PDFs are those at the lower end of the uncertainty interval.

VII. Dave's Optimal Number of Samples

A. Worst case scenario

So far in this paper, we assumed that Dave's detection performance increases with the number of samples he takes, despite being forced to have $\xi' > 1-\epsilon$ [7]. We showed that asymptotically, Dave's ξ' is either 0 or 1, depending on Alice's transmit power. However, the assumption that more samples is better is actually incorrect, given constraint (9). To see this, we calculate ξ'' at 1 sample,

$$\xi'' = \min_{\gamma'} \max_{\widehat{\Gamma}_d \in I} \left[\exp\left(\frac{-\gamma'}{\widehat{\Gamma}_d + \frac{A}{r_d^\alpha}\Gamma_s} \right) - \exp\left(\frac{-\gamma'}{\widehat{\Gamma}_d} \right) \right] . \quad (48)$$

We can plainly see that because $\frac{A}{r_d^\alpha}\Gamma_s > 0$, $\xi' > 0$ at $N = 1$. Also, $\xi' < 1$ because an exponential with a negative exponent must be less than 1. However, we showed that at an infinite number of samples, Dave's $\xi' = 1$, provided Alice's transmit power is low enough. Because ξ' is a continuous function over N, there must exist some finite N where ξ' is minimized for Dave, and we revise (9) to

$$\xi''' = \min_{N} \min_{\gamma'} \max_{\widehat{\Gamma}_d \in I} \left[P_{FA}(\widehat{\Gamma}_d, \gamma', N) + P_{MD}(\widehat{\Gamma}_d, \gamma', N) \right] \quad (49)$$

This result initially seems counterintuitive. In any detection scenario, the detector is at worst no better off, and almost always better off by gathering more samples. In this scenario, the detector is actually better off ignoring samples past the optimal number. However, if we look at (49), we see that Dave is trying to account for the worst case scenario when he maximizes $\widehat{\Gamma}_d$ over I. As Dave collects more samples, there is a chance that he will observe a rare event that represents his worst case. Because his test statistic is cumulative, he will eventually accumulate enough rare events that decrease his detection performance.

Another way to analyze this situation is with the test statistic probability density function (pdf)'s themselves.

At a large number of samples, by the CLT, the pdf of the test statistic under each hypothesis converges to a Gaussian distribution, and with an infinite number of samples the Gaussian distribution converges to a delta function at the mean of the distribution. The red deltas in Figures 10 and 11 represent the null hypothesis of Alice not transmitting. $H_{0,l}$ represents the null hypothesis pdf with $\widehat{\Gamma}_d = 1/\rho\Gamma_d$, and $H_{0,u}$ represents the null hypothesis pdf with $\widehat{\Gamma}_d = \rho\Gamma_d$. Conversely, the green deltas represent the alternate hypothesis that Alice is transmitting, with $H_{1,l}$ representing the alternate hypothesis pdf with $\widehat{\Gamma}_d = 1/\rho\Gamma_d$, and $H_{1,u}$ representing the alternate hypothesis pdf with $\widehat{\Gamma}_d = \rho\Gamma_d$.

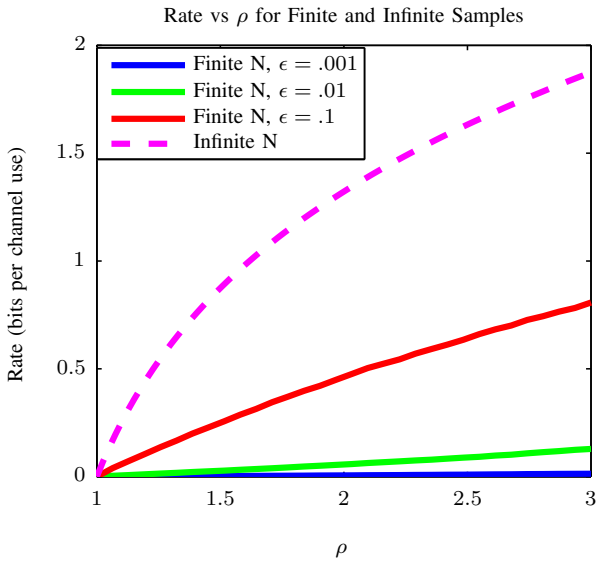

Fig. 12. Rate vs ρ for finite and infinite samples. Rate increases with ϵ

If Dave chooses any detection threshold less than $H_{0,u}$, such as γ_1, by the robustness criterion in (49), he chooses $\widehat{\Gamma}_d = \rho\Gamma_d$. This means the pdfs of his test statistic are $H_{0,u}$ and $H_{1,u}$, implying he will have 100% false alarms. If Dave chooses any detection threshold greater than $H_{1,u}$, such as γ_2, he chooses $\widehat{\Gamma}_d = 1/\rho\Gamma_d$ to satisfy (49). This means the pdfs of his test statistic are $H_{0,l}$ and $H_{1,l}$, implying that he will have 100% misses. There does not exist a threshold that Dave can choose that will allow his asymptotic ξ to be less than 1, and hence it would seem that Alice can communicate while forcing Dave's ξ'' to asymptotically approach 1.

However, if we analyze the finite sample case, we see that the strategy employed in the infinite sample case no longer works because the pdfs now have a support that is not infinitesimal. Dave can choose any detection threshold he desires, as long as he satisfies (49). Dave could choose the threshold γ_1, in which case there is no choice of $\widehat{\Gamma}_d$ and the corresponding pdfs that will force his ξ''' to be arbitrarily close to 1. From Fig. 11, with γ_1 as Dave's choice of detection threshold, his worst case ξ over the choice of $\widehat{\Gamma}_d$ would be on the order of .15 as opposed to asymptotically 1.

In order for Dave to actually compute the optimal N, γ', and $\widehat{\Gamma}_d$ in (49), Dave needs to know Alice's Γ_s. Because we are assuming that Alice and Dave are not cooperative, it is not realistic to assume that Dave has this information. However, if Alice assumes that Dave knows Γ_s, then Alice will be assuming the best case detection performance for Dave under the constraints he is given, and Alice will be guaranteed to communicate with privacy.

As we can see in Fig. 12, there is a dramatic decrease in privacy rate when we assume Dave uses the optimal number of samples.

Interestingly, the optimal number of samples decreases as ρ increases, as seen in Fig. 13.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2015.2421477, IEEE Journal of Selected Topics in Signal Processing

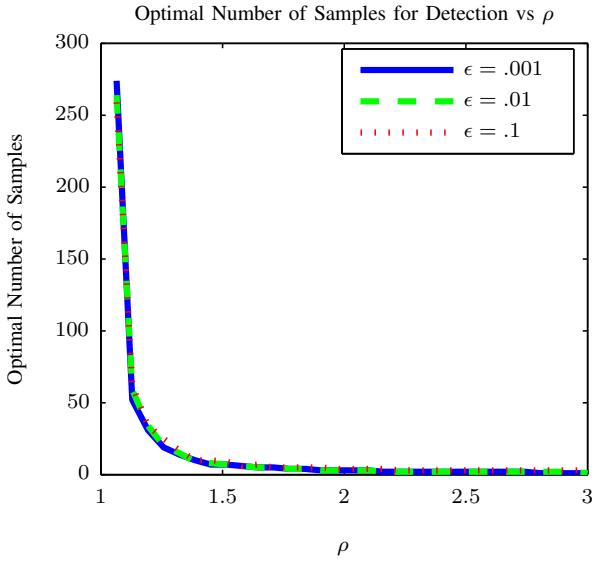

Fig. 13. Optimal number of samples for Dave. ϵ has little effect. Becomes difficult to compute at low values of ρ

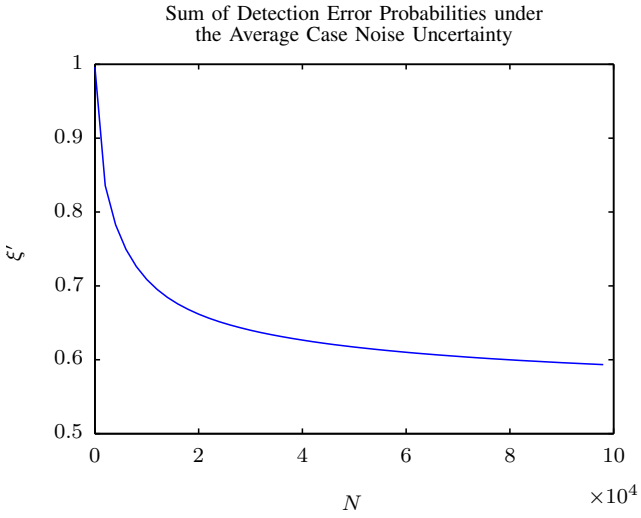

Fig. 14. Sum of detection error probabilities for the average case, vs number of samples

B. Average Case Scenario

If instead of maximizing his worst case to be robust, Dave targeted the average case, we see that more samples is better.

While we were not able to show this analytically, we did numerically compute the average case

$$\xi'' = \min_{N} \min_{\gamma'} \int_{1/\rho\Gamma_d}^{\rho\Gamma_d} \Big[P_{FA}(\widehat{\Gamma}_d, \gamma', N) + P_{MD}(\widehat{\Gamma}_d, \gamma', N) \Big] f_{\widehat{\Gamma}_d}(\widehat{\Gamma}_d) d\widehat{\Gamma}_d \quad (50)$$

where $f_{\widehat{\Gamma}_d}(\cdot)$ is the pdf of $\widehat{\Gamma}_d$. Assuming a uniform distribution on $\widehat{\Gamma}_d$, we can see in Fig. 14 that ξ'' decreases monotonically with respect to N when the optimal threshold γ'_{opt} is used.

This scenario is not entirely realistic because we are assuming that Dave has knowledge about $f_{\widehat{\Gamma}_d}(\cdot)$, and more importantly that he is not trying to be robust about his detection method. This detection metric would allow for his calculated ξ'' to be not be his true sum of detection errors because $\widehat{\Gamma}_d$ can only be one value in I. However, analyzing this scenario does provide intuition for the initially counterintuitive result seen previously.

TABLE I
PRIVACY RATES

Bandwidth	MIMO R_{pr}	SISO R_{pr}
1 Mhz	98.1 bits/s	9.07 bit/s
10 Mhz	981.2 bits/s	90.7 bit/s
20 Mhz	1962.3 bits/s	181.4 bits/s

C. Further Extension of Finite Sample Results

We leave this finite sample analysis as an open problem for the other types of channels we consider: SISO and MIMO Rayleigh channels. While we didn't conduct the finite sample analysis on such channels, the fact that we can overcome the square root law in [4] by assuming noise uncertainty, radiometer use, and Dave's taking of a finite number of samples, is the important aspect in this work.

VIII. PRACTICAL RATES

One concern in achieving these rates in practice is that Alice will not be certain of where the SNR wall is, especially in the Rayleigh case as the SNR wall is random. We make some practical assumptions in [7], leading to the results in Table I.

While these bitrates found in Table I are low, if Alice can obtain better estimates of the noise uncertainty by taking into account interference sources or other factors, this privacy rate can increase. Fig. 6 shows the privacy rate versus ρ. Additionally, if Bob gets closer, the bitrates can increase significantly because the received power is inversely proportional to distance squared.

The MIMO privacy rates are 2.7 times greater than having four individual SISO channels. It is important to remember that the search space for power allocation is not a plane like the standard total power constrained capacity problem—the search space is ellipsoidal in nature, as seen in Fig. 7. This non-planar shape allows us to increase our capacity by a factor beyond the number of antennas.

IX. OTHER SOURCES OF NOISE UNCERTAINTY

Up to this point we assumed that Dave's noise uncertainty is not affected by Alice's behavior. However, Alice could set up interference sources that turn on and off at random intervals. This interference can create more noise uncertainty for Dave and increase Alice's privacy rate. Also, because Alice set up the interference sources herself, she can estimate Dave's uncertainty from these sources.

Additionally, there can be other noise sources present that are not in collusion with Alice. In the extreme underlay scenario, the primary user could be seen as an interference source that increases Dave's noise uncertainty. However, Bob has to be able to reject the noise for this to increase his rate, because otherwise his noise increases as well and offsets the gain in allowable transmit power.

We leave further study into these areas as an open problem.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2015.2421477, IEEE Journal of Selected Topics in Signal Processing

X. Conclusion

It is possible to overcome the square root law of private communication if two assumptions are made: the detector is uncertain of its noise level and the detector uses a radiometer. We showed that in it's attempt to maximize the worst case scenario in the pursuit of robustness, the detector should only take into account a finite number of samples. While the detector cannot actually calculate the optimal number of samples without knowing the transmitter's power, the detector does know that the optimal number of samples decreases as its uncertainty about the noise increases.

Further work would be to analyze Rayleigh SISO and MIMO channels to confirm that a finite number of samples is optimal in those cases as well.

References

[1] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinvasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[2] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The feasibility of launching and detecting jamming attacks in wireless networks," in *Mobile Ad Hoc Networking and Computing, Proc. of the 6th ACM International Symposium on*, 2005, pp. 46–57.

[3] L. Lazos, S. Liu, and M. Krunz, "Mitigating control-channel jamming attacks in multi-channel ad hoc networks," in *Wireless Network Security, Proc. of the Second ACM Conference on*, 2009, pp. 169–180.

[4] B. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 9, pp. 1921–1930, September 2013.

[5] J. Fridrich, *Steganography in Digital media: Principles, Algorithms, and Applications*, 2nd ed. MIT Press, 2001.

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

[7] S. Lee and R. J. Baxley, "Achieving positive rate with undetectable communication over AWGN and Rayleigh channels," in *Communications (ICC), 2014 IEEE International Conference on*, Jun. 2014, pp. 780–785.

[8] S. Lee, R. J. Baxley, J. B. McMahon, and R. S. Frazier, "Achieving positive rate with undetectable communication over AWGN and Rayleigh channels," in *Sensor Array and Multichannel Signal Processing Workshop (SAM), 2014 IEEE 8th*, Jun. 2014, pp. 257–260.

[9] T. Rappaport, *Wireless Communications*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2002.

[10] S. Alexander, "Characterising buildings for propagation at 900 MHz," *Electronics Letters*, vol. 19, no. 20, pp. 860–, September 1983.

[11] E. Lehmann and J. Romano, *Testing Statistical Hypotheses*, 3rd ed. NY: Springer, 2005.

[12] K. S, *Fundamentals of Statistical Signal Processing Detection Theory*, 2nd ed. Prentice Hall, Inc., Upper City River, New Jersey, 1998.

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, Hoboken, NJ, 2002.

[14] M. Bloch, J. Barros, M. Rodrigues, and S. McLaughlin, "Wireless information-theoretic security," *Information Theory, IEEE Transactions on*, vol. 54, no. 6, pp. 2515–2534, June 2008.

[15] D. Bliss, K. Forsythe, and A. Yegulalp, "MIMO communication capacity using infinite dimension random matrix eigenvalue distributions," in *Signals, Systems and Computers, 2001. Conference Record of the Thirty-Fifth Asilomar Conference on*, vol. 2, Nov 2001, pp. 969–974 vol.2.

[16] A. Hero, "Secure space-time communication," *Information Theory, IEEE Transactions on*, vol. 49, no. 12, pp. 3235–3249, Dec 2003.

[17] D. Hammarwall, M. Bengtsson, and B. Ottersten, "Acquiring partial CSI for spatially selective transmission by instantaneous channel norm feedback," *Signal Processing, IEEE Transactions on*, vol. 56, no. 3, pp. 1188–1204, March 2008.

[18] M. Vu, "MIMO capacity with per-antenna power constraint," in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, Dec 2011, pp. 1–5.

Seonwoo Lee receieved the B.S. degree in electrical engineering from Cornell University, Ithaca, NY, in 2012. He is currently pursuing the Ph.D. degree in electrical engineering at the Georgia Institute of Technology (Georgia Tech), Atlanta, USA.

He is a member of the Software Defined Radio Lab at the Georgia Tech Research Institute, Atlanta. His current research interests include signal processing and wireless communications. He is particularly interested in applications to private communications.

Brett Walkenhorst is a Principal Research Engineer, principal investigator, program manager, and technical contributor on multiple government- and industry-funded efforts, and Acting Associate Chief of GTRIs Systems Technology and Analysis Division (STAD). He received his B.S. and M.S. degrees in electrical engineering from Brigham Young University and his Ph.D. degree in electrical engineering from Georgia Tech. His research interests include signal processing, cognitive RF systems, detection and estimation theory, machine learning, and physical layer security. Dr. Walkenhorst is a member of EE honor society Eta Kappa Nu, a Senior member of IEEE, and the Chair of the Atlanta Chapter of the IEEE Communications Society. He has taught communications and signal processing at the graduate and undergraduate level and has taught professional short courses on communication theory including modulation, demodulation, forward error correction, and advanced topics including MIMO communication, interference suppression, OFDM, spread spectrum, etc.

Mary Ann Weitnauer (formerly Mary Ann Ingram) received the B.E.E. and Ph.D. degrees from the Georgia Institute of Technology (Georgia Tech), Atlanta, USA, in 1983 and 1989, respectively. In 1989, she joined the faculty of the School of Electrical and Computer Engineering at Georgia Tech, where she is currently Professor. Her early research areas were optical communications and radar systems. In 1997, she established the Smart Antenna Research Laboratory (SARL) at Georgia Tech, which applies real and virtual array antennas to wireless networks and radar systems. She held the Georgia Tech ADVANCE Professorship for the College of Engineering from 2006-2012. She was a Visiting Professor at Aalborg University, Aalborg, Denmark in the summers of 2006-2008 and at Idaho National Labs in 2010. The SARL performs system analysis and design, channel measurement, and prototyping relating primarily to wireless local area, ad hoc and sensor networks, with focus on the lower three layers of the protocol stack. SARL has also recently developed signal processing algorithms for impulse radio ultrawideband (IR-UWB) for non-contact vital signs measurement. Dr. Weitnauer has authored or co-authored over 160 refereed journal and conference papers, including four conference papers that have won Best Paper awards. She served as a Guest Editor for the EURASIP Special Issue on Cross-Layered Design for Physical/MAC/Link layers in Wireless Systems in 2007 and Co-Chair for America for the IEEE Wireless VITAE Conference in 2009. She was an associate editor for the IEEE TRANSACTIONS ON MOBILE COMPUTING from 2009-2012. Dr. Weitnauer is a Senior Member of the IEEE.

Robert J. Baxley Dr. Bob Baxley received the BS, MS, and PhD degrees all in Electrical Engineering from Georgia Tech. In the course of his graduate work, he was the recipient of the Sigma Xi Award, the National Science Foundation Graduate Research Fellowship Program award, the Grand Prize in the SAIC student paper competition, and the Georgia Tech Center for Signal & Image Processing Outstanding Research Award. From 2008 to 2014, Dr. Baxley was employed at the Georgia Tech Research Institute where he served as the Director of the Software Defined Radio Laboratory among other roles. While at GTRI, he led the GTRI team that placed second out of 90 teams in the DARPA Spectrum Challenge. Dr. Baxley is currently the Chief Engineer at Bastille. He also holds an Adjunct Faculty appointment in the Georgia Tech School of Electrical and Computer Engineering.

Dr. Baxley's research interests are in the areas of cognitive radio, wireless security, and signal processing for communications systems. From 2012 to 2014 Dr. Baxley served as an Associate Editor for Digital Signal Processing. Dr. Baxley is a member of the Association of Old Crows and is a Senior