Exploiting the Non-Commutativity of Nonlinear Operators for Information-Theoretic Security in Disadvantaged Wireless Environments

Azadeh Sheikholeslami, Dennis Goeckel and Hossein Pishro-Nik

Electrical and Computer Engineering Department, University of Massachusetts, Amherst, MA
{sheikholesla,goeckel,pishro}@ecs.umass.edu

Abstract—**Information-theoretic security guarantees that a message is kept secret from potential eavesdroppers regardless of their current or future computational abilities. But current information-theoretic security approaches generally rely on an advantage of the channel of the desired recipient over the adversary, and such an advantage can be difficult to guarantee in a wireless network where an eavesdropper might be very near the transmitter. This paper initiates an approach to everlasting security for wireless communication links by exploiting a fundamental concept from systems theory: that nonlinear systems are not (necessarily) commutative. This property is exploited by employing a short-term cryptographic key to force the eavesdropper's signal to be subjected to nonlinear operations in the reverse order of that of the signal at the desired recipient. After introducing the idea and providing analysis for the general case, we next consider a simple (and practical) instantiation where the transmitter uses the ephemeral cryptographic key to rapidly power modulate the transmitted signal. Secrecy rates with this rapid power modulation under various assumptions establish the promise of the approach, even in the case of an eavesdropper with uniformly better conditions (channel and receiver quality) than the intended recipient.**

I. Introduction

Wireless networks are vulnerable to being eavesdropped and hence security is a primary concern to be addressed. The standard method of providing security against eavesdroppers is to encrypt the information so that it is beyond the eavesdropper's computational capabilities to decrypt the message [1]; however, the vulnerability shown by many implemented cryptographic schemes, the lack of a fundamental proof establishing the difficulty of the problem presented to the adversary, and the potential for transformative changes in computing motivate forms of security that are provably everlasting. In particular, when a cryptographic scheme is employed, the adversary can record the clean cypher and recover it later when the cryptographic algorithm is broken [2], which is not acceptable in sensitive applications requiring everlasting secrecy. The desire for such everlasting security motivates considering emerging information-theoretic approaches, where the eavesdropper is unable to extract from the received signal any information about the secret message.

In 1946, Shannon introduced information theoretic (or unconditional) secrecy [3]. If the uncertainty of the message after seeing the cypher is equal to the uncertainty of the message before seeing the cypher, we have perfect secrecy without any condition on the eavesdropper's capabilities. However, in these schemes, the length of the code must be at least as long as the length of the message (e.g. one-time pad), which makes them difficult to implement and often impractical. However, Wyner later showed that if the eavesdropper's channel is degraded with respect to the main channel, adding some randomness to the codebook allows perfect secrecy to be achieved [4]. Csiszar and Korner extended the idea to more general cases, where the eavesdropper's channel is not necessarily degraded with respect to the main channel, but it must be "more noisy" or "less capable" than the main channel [5]. When such an advantage does not exist, one can turn to approaches based on "public discussion" [6], [7], but these approaches, while they could be used to generate an information-theoretically secure one-time pad, are generally envisioned for secret key agreement to support a cryptographic approach [8, Chapter 7.4] rather than efficient one-way secret communication. Consequently, the desirable situation for achieving information theoretic secrecy is to have a better channel from the transmitter to the intended receiver than that from the transmitter to the eavesdropper. However, this is not always guaranteed, particularly in wireless systems where the eavesdropper can have a large advantage over the intended receiver. In the case of a passive adversary, the eavesdropper can be very close to the transmitter or it can use a directional antenna to improve its received signal, while there is no way for the legitimate nodes to know the eavesdropper's location or its channel state information. Recent authors have considered approaches that relax the need for assumptions on Eve's location or channel in one-way systems. For cases when the eavesdropper location is unknown (which means the case of a "near Eve" must be considered), approaches largely based on the cooperative jamming approach of [9] and [10] have been considered [11], [12]. However, all of these approaches require either multiple antennas, helper nodes, and/or fading (see, for example, [13]), and many are susceptible to attacks such as pointing directive antennas at one or both communicating parties.

This work was supported by the National Science Foundation under grants CNS-0905349, ECCS-0725616, and CIF-1249275.

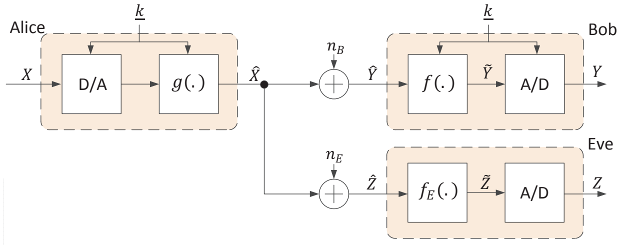

Fig. 1. The message X is observed at Bob and Eve through the transmitter, the AWGN channels with different noise variances, and their respective receivers with (possibly nonlinear) functions $g(.)$, $f(.)$, and $f_E(.)$. The sequence $\underline{k}$ is a cryptographic key shared by Alice and Bob, which is assumed to be obtained by Eve immediately *after* she has recorded Z.

For a one-way scenario with a single antenna where Bob's channel is worse than Eve's, Cachin and Maurer [14] exploited the realizability of hardware to consider the case of everlasting security, as is our interest. In particular, they introduced the "bounded memory model" - signal in such a way that the receiver cannot store the information it would need to eventually break the cypher. This novel approach suffers from two shortcomings: (1) by Moore's Law (see NAND scaling plot at [15]), the density of memories increases at an exponential rate; (2) memories can be stacked arbitrarily subject only to (very) large space limitations. Hence, although the bounded memory model is a viable approach to everlasting security, it is difficult to pick a memory size beyond which it will be effective, making its employment for secret wireless communication difficult. Rather than attacking the memory in the receiver back-end, our contention is that one should instead consider attacking the receiver front-end and analog-to-digital (A/D) conversion process, where technology progresses slowly and there exist well-known techniques for severely handicapping the component. And, unlike memory, A/D's cannot be stacked arbitrarily, as clock jitter prevents the timing required; in fact, high-quality A/D's already employ parallelization to the limit of the jitter. And, importantly from a long-term perspective, there is a *fundamental bound* on the ability to perform A/D conversion [16], [17]. Consider the channel model shown in Figure 1, which reflects the understanding that in an adversarial game in modern communication systems, it is the interference effects on wideband receiver front-ends rather than the baseband processing that is the significant detriment [18]. In particular, the signal is subject to a variety of distortions due to the RF front-end of the receiver and the analog-to-digital (A/D) conversion. A large interferer, even if it is orthogonal to the signal of interest and thus (supposedly) easily rejected by baseband processing, can saturate the receiver front-end, leading to nonlinearities, and, of particular interest here, reducing the receiver's dynamic range (i.e. resolution) significantly.

The primary focus of this paper is to exploit the receiver processing effects for security. In particular, based on a pre-shared key between Alice and Bob that only needs to be kept secret for the duration of the wireless transmission (i.e. it can be given to Eve immediately afterward), we consider how inserting intentional (but known to Bob) distortion on the transmitted signal can provide information-theoretic security. In particular, since Bob knows the distortion, he can undo its effect before his A/D, whereas Eve must store the signal and try to compensate for the distortion after her A/D. Since the A/D is necessarily non-linear, the operations are not commutative and there is the potential for information-theoretic security. This paper introduces this idea and initiates its investigation.

As a first example, we perform a rapid power modulation between two distinct far apart power levels at the transmitter and put the reciprocal of that power gain before Bob's A/D. Cellular (and other) networks usually provide a large power "headroom" to ensure the quality of the received signal. In short range communication, even when the transmitted signal is modulated with a very small gain, this headroom guarantees that Bob will receive the signal with an acceptable signal-to-noise ratio. Moreover, the power gain before Bob's A/D ensures that the received signal is in an appropriate range. Modern power amplifiers can easily have their power switched at high bandwidths [19] [20, Chapter 7]. Since the power can be changed every symbol, Eve cannot use any type of automatic gain control (AGC) loop and is left trying to select a gain that trades off resolution and the probability of overflow of her A/D, and hence information theoretic secrecy is obtained.

The rest of paper is as follows. Section II describes the system model, metrics, and the proposed idea in detail. In Section III, the proposed method is applied to settings with noisy channels and noiseless channels, respectively, to find achievable secrecy rates in each case, and an asymptotic analysis of the proposed method is provided. In Section IV, the results of numerical examples for various realizations of the system are presented. Conclusions and ideas for future work are discussed in Section V.

II. System Model and Approach

A. System Model and Metric

We consider a simple wiretap channel, which consists of a transmitter, Alice, a receiver, Bob, and an eavesdropper, Eve. Eve is a passive eavesdropper, i.e. she just tries to obtain as much information as possible to recover the message that Alice sends and she does not attempt to actively thwart (i.e. via jamming, signal insertion) the legitimate nodes. Therefore, the location and channel state information of Eve can be difficult to obtain and thus is assumed unknown to the legitimate nodes.

We assume that Alice and Bob either pre-share a cryptographic key or that they employ a standard key agreement scheme (e.g. Diffie-Hellman [21]) to generate a shared key and then expand the number of key bits aggressively using a linear feedback shift register (LFSR) to generate a long key sequence. In general the output of an LFSR, even when the initial state is unknown, is easily predictable and thus inappropriate for cryptography. However, our system design only employs the key ephemerally; in fact, we assume (pessimistically) that

Eve is handed the full key as soon as transmission is complete. Since Eve only views the transmitted signal through a very noisy process (see below), we assume she cannot hope to recover the key during the (very) short transmission period.

We consider a memoryless one-way communication system, and assume that both Bob and Eve are at a unit distance from the transmitter by including variations in the path-loss in the noise variance. Thus, the channel gain of both channels is unity and both channels experience additive white Gaussian noise (AWGN). Let n_B and n_E denote the zero-mean noise processes at Bob's and Eve's receivers with variances N_B and N_E, respectively. Let $\hat{X}$ denote the input of both channels, $\hat{Y}$ denote the received signal at Bob's receiver, and $\hat{Z}$ denote the received signal at Eve's receiver. The signal at Bob's receiver is:

$$\hat{Y} = \hat{X} + n_B$$

and the signal at Eve's receiver is:

$$\hat{Z} = \hat{X} + n_E$$

Both Bob and Eve employ high precision uniform analog-to-digital converters. The effect of the A/D on the received signal (quantization error) is modeled by a quantization noise due to the limitation in the size of each quantization level, and a clipping function due to the quantizer's overflow. The quantization noise in this case is (approximately) uniformly distributed [22], so we will assume it is uniformly distributed throughout the paper. For an m-bit quantizer ($b = 2^m$ gray levels) over the full dynamic range $(-a, a)$, two adjacent quantization levels are spaced by $\delta = 2a/b$, and thus the quantization noise is uniformly distributed over an interval of length δ. Quantizer overflow happens when the amplitude of the received signal is greater than the quantizer's dynamic range, which can be modeled by a clipping function.

Let X denote the current code symbol, which we assume is taken from a standard Gaussian codebook where each entry has variance P, i.e. $X \sim \mathcal{N}(0, P)$. Note that although the Gaussian codebook is optimal to achieve the secrecy capacity in the case of AWGN wiretap channels, because we consider quantization errors in our model, the Gaussian codebook is no longer optimum, implying that our results represent achievable rates but not upper bounds.

From [23], for an arbitrary stationary memoryless wiretap channel with arbitrary input and output alphabets, any secrecy rate

$$R_s < \max_{X \to YZ} [I(X;Y) - I(X;Z)]$$

is achievable. Hence, the average secrecy rate that can be achieved is :

$$R_s(\mathcal{S}, \mathcal{S}') = E\left[I(X;Y) - I(X;Z) | \mathcal{S}, \mathcal{S}'\right]$$

where $\mathcal{S}$ is the strategy taken by Alice, $\mathcal{S}'$ is the strategy taken by Eve, and expectation is over the (potential) randomness of the strategy. The eavesdropper (Eve) tries to pick a strategy that minimizes R_s. On the other hand, the transmitter (Alice) tries to choose a strategy to maximize the worst case (minimum) R_s. To formulate this, we use the following maximin criteria:

$$R_s^* = \max_{\mathcal{S}} \min_{\mathcal{S}'} R_s(\mathcal{S}, \mathcal{S}') \tag{1}$$

Here, R_s^* is the minimum secrecy rate that can be guaranteed. In other words, no matter which strategy $\mathcal{S}'$ Eve takes, choosing the appropriate strategy $\mathcal{S}$ will guarantee the secrecy rate R_s^*.

B. General Nonlinearity: Rough Analysis

Our goal is to consider how Alice and Bob can employ bits of the shared cryptographic key to modify their radios as shown in Figure 1 to gain (or maximize) an information theoretic advantage. For now, assume that they insert general memoryless nonlinearities $g(.)$ at the transmitter and $f(.) = g^{-1}(.)$ at the receiver based on the key. Suppose that Eve is able to obtain the key just after the transmission is finished; considering for the moment that she applies $g^{-1}(.)$ to Z, one sees how the security is (potentially) obtained: Bob sees $g(X)$ through $g^{-1}(.)$ and the A/D, whereas Eve sees those operations in reverse. Since nonlinear operations are not (necessarily) commutative, the signals are not the same and there is the potential for some form of information-theoretic security.

Now, stepping back to allow Eve to use the key sequence $\underline{k}$ in whatever manner she wants after she has recorded the transmission yields an illustrative information-theoretic model. In particular, using the same random coding arguments as for fading channels, consider a collection $\mathcal{G}$ of functions $g(.)$ from which $\underline{k}$ selects; then, the secrecy rate is:

$$R_s = E_{\mathcal{G}}[I(X;Y|\underline{k}) - I(X;Z|\underline{k})]$$

Let us be pessimistic and assume $\sigma_E^2 = 0$. Furthermore, to get some insight, assume temporarily that $\sigma_B^2 = 0$, corresponding to a short-range situation which is not power-limited. For $\sigma_B^2 = 0$, Y does not depend on $\underline{k}$ and thus using the approach for analyzing quantizers of [24, pg. 251], which is accurate at high resolution:

$$\begin{aligned} R_s &= E_{\mathcal{G}}[I(X;Y) - I(X;Z|\underline{k})] \\ &= E_{\mathcal{G}}[H(Y) - H(Y|X) - (H(Z|\underline{k}) - H(Z|X,\underline{k}))] \\ &\approx E_{\mathcal{G}}[h(\tilde{Y}) - \log(\delta) - (h(\tilde{Z}|\underline{k}) - \log(\delta))] \\ &= E_{\mathcal{G}}[h(\tilde{Y}) - h(\tilde{Z}|\underline{k})] \\ &= E_{\mathcal{G}}[h(X) - h(g(X))] \end{aligned}$$

where $\tilde{Y}$ and $\tilde{Z}$ are the inputs to Bob and Eve's A/D converters, respectively. It then becomes apparent that the gain observed here for high-resolution A/D's at both Bob and Eve is a shaping gain between X and $g(X)$. Whereas we think of shaping gains as tending to be relatively small (1.53 dB on the Gaussian channel [25]), that is because the generally considered gains are between the optimal (Gaussian) shaping and a standard but reasonable (uniform) shaping. In our design scenario, if we are able to severely distort $g(X)$, the gains can become enormous. We quickly caveat

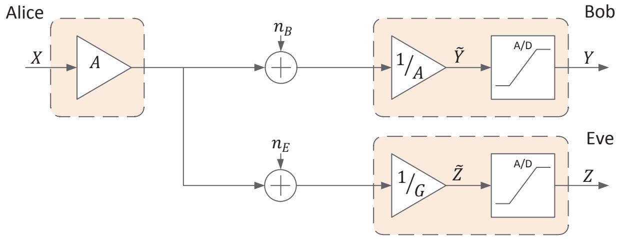

Fig. 2. Alice and Bob share a cryptographic key that determines the value of A at each time instance. Eve puts a (possibly variable) gain before her A/D to decrease the A/D erasures and/or overflows and hence increase the information leakage.

this conclusion by noting that the assumption $\sigma_B^2 = 0$ is critical, since those $g(.)$ which are most distorting can also cause significant "noise enhancement" on the channel from Alice to Bob. Hence, unless the noise is truly negligible (i.e. very short range communication), judgment should be reserved on the applicability of the technique until $\sigma_B^2 \neq 0$ is considered in Section III.

C. Rapid power modulation for secrecy

For the rest of the paper, we simplify the operator $g(.)$ to a random gain to consider a practical architecture easily implemented and discuss specific operating scenarios. Our goal is to achieve a positive secrecy rate by confusing the eavesdropper's A/D. Throughout this paper we assume that Eve is able to employ just one A/D, reserving discussion of the multiple A/D case to Section V. Here, the strategy that Alice takes to confuse Eve's A/D is to apply a random gain from a fixed probability distribution to the signal amplitude of each symbol that she transmits. Suppose that A denotes the random variable associated with this random gain, and the probability density function (pdf) of this gain is $p_A(u)$ where $u \in \mathcal{A}$ (see Fig. 2). The pdf of A is known to all nodes, but only legitimate nodes know the exact sequence of values of A (i.e. $u_1, u_2, u_3, \cdots$) that has been applied to the symbol sequence.

We want to find a probability distribution for A that maximizes this secrecy rate. We choose this gain such that it does not change the average power of the transmitted signal, i.e. E$[|A|^2] = 1$. To control the number of key bits required, we consider that $|A|$ is drawn from one of two levels A_1 and A_2 with random polarity (i.e. $\mathcal{A} = \{A_1, -A_1, A_2, -A_2\}$):

$$Pr(A = u) = \begin{cases} p, & u = A_1 \\ 1 - p, & u = A_2 \end{cases}$$

and $Pr\{A > 0\} = Pr\{A < 0\} = 1/2$. Suppose that A_1 is the large gain and A_2 is the small gain that the transmitter applies. We define the ratio between the large gain and the small gain $r = \frac{A_1}{A_2}$.

Since Bob shares the (long) key with Alice, he easily "inverts" the gain A to operate his A/D properly, whereas Eve will struggle with such. In essence, we are inducing a fading channel at Bob that he is able to equalize *before* his A/D, whereas Eve cannot. Bob applies the reciprocal of A before his A/D and thus given A, the signal that Bob's A/D sees is:

$$\tilde{Y} = X + \frac{n_B}{A} \tag{2}$$

To cancel the effect of this gain, Eve also applies an arbitrary (possibly random) gain, $1/G$. So, the signal at Eve's A/D given A and G is:

$$\tilde{Z} = \frac{A}{G}X + \frac{n_E}{G} \tag{3}$$

Suppose that Eve knows the pdf of A; hence, she tries to find a probability density function $p_G(g)$ for G such that it minimizes the secrecy rate R_s or equivalently maximizes the information leakage $I(X; Z)$. On the other hand, Alice sets the pdf parameters to maximize R_s. So, Alice chooses a probability distribution for A such that no matter what $p_G(g)$ Eve chooses, some secrecy rate R_s is always guaranteed, and she tries to maximize this R_s. Hence, the maximin criteria in (1) turns into:

$$R_s^* = \max_{p, A_1, A_2} \min_{p_G} R_s(p_G, A_1, A_2, p) \tag{4}$$

Obviously, larger $r = \frac{A_1}{A_2}$ leads to more eavesdropper confusion. However, because $E[|A|^2] = 1$, $r \gg 1$ leads to a small A_2, and Bob then suffers noise enhancement. We talk about the choice of r in the next paragraph.

Recall the potential operating scenario from Section I, and assume that system radios are operating in a scenario where they have adequate power amplifier headroom, as in the "near" situation in cellular systems [26], and the user's noise is relatively negligible. However, an Eve at the same range can also intercept the signal. By changing the power of the transmitters between the power-controlled level (e.g. A_2), where it meets the receiver requirements and its maximum power (e.g. A_1), Bob, knowing the sequence, obtains a signal that is at least equivalent to operating at its power controlled level and thus sees little degradation in information transmission. The ratio between the large gain and the small gain, r, can be chosen such that in the case of $A = A_2$ (small gain), the minimum acceptable signal level at Bob's receiver is satisfied. On the other hand, Eve's A/D struggles even to record a reasonable form of the signal; hence, she sees significant degradation, and information-theoretic security is obtained. Also, because the power level is changed very fast (at every symbol), the automatic gain control (AGC) at the eavesdropper's receiver cannot follow the deep fades that cause erasures and/or strong signals that cause A/D saturation.

To choose optimum values for A_1, A_2, and p, note that the following constraints must be met:

$$\frac{A_1}{A_2} = r \quad \text{and} \quad pA_1^2 + (1-p)A_2^2 = 1 \tag{5}$$

Hence, two of these values are constrained by the system parameter r and conservation of transmission power, and the transmitter is free to choose only one (e.g. p). So, equation (4) reduces to:

$$R_s^* = \max_{p} \min_{p_G} R_s(p_G, p) \tag{6}$$

Eve can employ a number of countermeasures to decrease R_s. She can find an optimum probability density function that minimizes R_s, or she can employ a better A/D to decrease erasures and/or overflows of her A/D. In the sequel, we will consider these scenarios and consider the secrecy rate R_s that can be achieved by the method proposed in this paper in each case.

III. Achievable Secrecy Rates

In this section the secrecy rates that can be achieved considering the non-idealities of the A/D's at the front-ends of Bob and Eve's receivers are studied. In the first part, the channel between Alice and Bob and the channel between Alice and Eve are considered to be AWGN channels. In the second part, to get more insight into the problem, the noise is removed from the channels and only the effect of A/D's on the signals will be considered.

A. Noisy channels

Consider the derivation of $I(X;Y) - I(X;Z) = h(Y) - h(Y|X) - (h(Z) - h(Z|X))$. Clearly, each of $h(Y)$, $h(Y|X)$, $h(Z)$, and $h(Z|X)$ are required. Recall that throughout this paper the non-idealities of the A/D's are modeled by an additive uniformly distributed quantization noise and a clipping function; hence, the signal Y after Bob's A/D with input $\tilde{Y}$ is:

$$Y = \begin{cases} \tilde{Y} + n_q, & |\tilde{Y}| < a \\ +a, & \tilde{Y} > a \\ -a, & \tilde{Y} < -a \end{cases}$$

where $\tilde{Y} = X + \frac{n_B}{A}$ and a is determined by the span $[-a, a]$ of the A/D. Thus, $\tilde{Y}$ has a zero-mean Gaussian distribution with variance $P + N_B/A^2$, i.e. $\tilde{Y} \sim \mathcal{N}(0, P + N_B/A^2)$. Let's define the event $E_1 = \{|\tilde{Y}| < a\}$ which corresponds to the case that no clipping occurs, and the events $E_2 = \{\tilde{Y} > a\}$ and $E_3 = \{\tilde{Y} < -a\}$ to correspond to clipping (A/D overflow); thus,

$$h(Y) = \sum_{i=1}^{3} h(Y|E_i)p(E_i) \tag{7}$$

In the case of clipping we have $h(Y|E_2) = h(Y|E_3) = 0$. Calculations of $h(Y|E_1)$ and $E(E_1)$ are presented in [27] due to lack of space.

Similarly, for $h(Y|X)$ we have,

$$\begin{aligned} h(Y|X) &= \sum_{i=1}^{3} h(Y|E_i, X)p(E_i|X) \\ &= \int_{-\infty}^{\infty} \sum_{i=1}^{3} h(Y|E_i, X = x)p(E_i|X = x)f_X(x)dx \end{aligned} \tag{8}$$

where $h(Y|E_2, X = x) = h(Y|E_3, X = x) = 0$. For calculations of $h(Y|E_1, X = x)$ and $E(E_1|X = x)$ please see [27]. By using $h(Y)$ from (7) and $h(Y|X)$ from (8), the mutual information between X and Y can be found:

$$I(X;Y) = h(Y) - h(Y|X) \tag{9}$$

The signal that Eve after her A/D sees is,

$$Z = \begin{cases} \tilde{Z} + n_q, & |\tilde{Z}| < a \\ +a, & \tilde{Z} > a \\ -a, & \tilde{Z} < -a \end{cases}$$

where $\tilde{Z} = \frac{AX}{G} + \frac{n_E}{G}$ and thus $\tilde{Z} \sim \mathcal{N}(0, \frac{A^2P+N_E}{G^2})$. Similar to the previous case, the event that the signal before Eve's A/D falls in its dynamic range is $E_1' = \{|\tilde{Z}| < a\}$, and the events $E_2' = \{\tilde{Z} > a\}$ and $E_3' = \{\tilde{Z} < -a\}$ correspond to the cases that Eve's A/D overflows. Consequently,

$$h(Z) = \sum_{i=1}^{3} h(Z|E_i')p(E_i') \tag{10}$$

In the case that clipping occurs $h(Z|E_2') = h(Z|E_3') = 0$. Again due to lack of space, the calculation of $h(Z|E_1')$ and $p(E_1')$ are omitted here (see [27]). Similarly,

$$\begin{aligned} h(Z|X) &= \sum_{i=1}^{3} h(Z|E_i', X)p(E_i'|X) \\ &= \int_{-\infty}^{\infty} \sum_{i=1}^{3} h(Z|E_i', X = x)p(E_i'|X = x)f_X(x)dx \end{aligned} \tag{11}$$

where $h(Z|E_2', X = x) = h(Z|E_3', X = x) = 0$, and, $h(Z|E_1', X = x)$ and $p(E_1', X = x)$ can be substituted from [27].

By substituting $h(Z)$ from (10) and $h(Z|X)$ from (11) in the following equation,

$$I(X;Z) = h(Z) - h(Z|X) \tag{12}$$

the mutual information between Alice and Eve can be found.

Finally, the achievable secrecy rate can be found by substituting the mutual informations from (9) and (12) into the following equation:

$$\begin{aligned} R_s &= E_{G,A}\left[I(X;Y) - I(X;Z)\right] \\ &= E_A\left[I(X;Y)\right] - E_{G,A}\left[I(X;Z)\right] \end{aligned} \tag{13}$$

Alice is able to choose p to maximize the R_s that can be achieved by this method; on the other side, Eve tries to minimize R_s by choosing an appropriate $p_G(g)$. The following lemma shows that for an arbitrary discrete alphabet for G, choosing a single value (which depends on the value of p) with probability one minimizes the secrecy rate, and thus is the optimal strategy for Eve.

Lemma 1. The gain $1/G$ that Eve applies before her A/D should take a single value with probability one to minimize the secrecy rate.

Proof: Suppose G has the following probability distribution:

$$Pr(G = g) = \begin{cases} \alpha_i, & g = G_i, i = 1, \cdots, n \\ 0, & \text{else} \end{cases}$$

such that $\sum_{i=1}^{n} \alpha_i = 1$. Without loss of generality, assume that for a specific p, the maximum information leakage occurs at $G = G_1$, i.e. for any gain $G_i, i = 2, \cdots, n$ we have $I(X;Z|G = G_1) \geq I(X;Z|G = G_i)$; hence,

$$\begin{aligned} I(X;Z) &= \sum_{i=1}^{n} \alpha_i I(X;Z|G = G_i) \\ &\leq \sum_{i=1}^{n} \alpha_i I(X;Z|G = G_1) \\ &= I(X;Z|G = G_1) \end{aligned}$$

■

The above lemma can easily be generalized to continuous random variables. Numerical results are given in Section IV.

B. Noiseless Channels

In the case Bob has a noiseless channel, $h(Y)$ can be found by setting $N_B = 0$. To calculate $h(Y|X)$, using (8) and the fact that $h(Y|E_2, X = x) = h(Y|E_3, X = x) = 0$ we have,

$$h(Y|X) = \log(\delta)\left(1 - 2Q\left(\frac{a}{\sqrt{P}}\right)\right) \qquad (14)$$

Similarly, in the case that the channel between Alice and Eve is noiseless, $h(Z)$ can be found by setting $N_E = 0$. To calculate $h(Y|X)$, using (11) and the fact that $h(Z|E'_2, X = x) = h(Z|E'_3, X = x) = 0$ we have,

$$h(Z|X) = \log(\delta)\left(1 - 2Q\left(\frac{Ga}{A\sqrt{P}}\right)\right) \qquad (15)$$

The details are omitted here due to lack of space and can be found in [27]. In each case, the secrecy rate can be found by substituting (14) and (15) in (9) and (12), respectively. Numerical results are shown in Section IV and in [27, Section VII].

Clearly, considering the noiseless channels makes the results less complicated and hence more insightful. Hence, we continue our investigation by studying the asymptotic behavior of the proposed method (as $r \to \infty$) in the noiseless regime, which will help us to achieve some intuition regarding this scheme. We assume that Bob and Eve use A/D's of the same quality for this analysis.

Since in the noiseless regime, $I(X;Y)$ does not depend on A, it does not change with r and thus we just evaluate $I(X;Z)$ for our asymptotic analysis.

From (5) we have,

$$A_1 = \frac{r}{\sqrt{pr^2 + (1-p)}} \quad \text{and} \quad A_2 = \frac{1}{\sqrt{pr^2 + (1-p)}}$$

Let $G(r)$ be the inverse of the gain that Eve employs as a function of r. Recall that from Lemma 1, $G(r)$ will take a single value with probability one for a given r, but that value can depend on r. Since $A_1 \to 1/\sqrt{p}$ and $A_2 \to 0$, we claim that in the limit (as $r \to \infty$), the best strategy that Eve can take is to choose either $G(r) = \Theta(1)$ or $G(r) = \Theta(r^{-1})$; otherwise, she will get no information (see Appendix A in [27]). In [27, Section III] it is shown that the secrecy rate that can be achieved in the asymptotic case (as $r \to \infty$) is:

$$R_s = (1 - \epsilon)I(X;Y) \qquad (16)$$

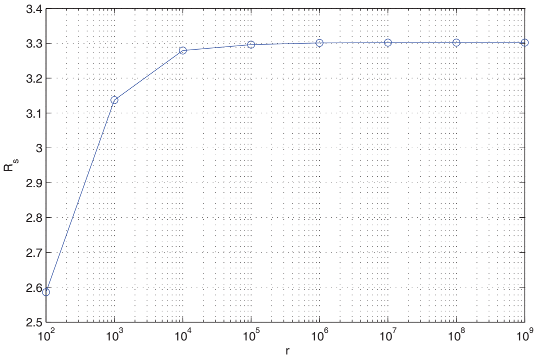

Fig. 3. Achievable secrecy rate versus r (the ratio between the large and the small gain) when both the main and eavesdropper's channels are noiseless and both Bob and Eve apply 10-bit A/D's with the dynamic range $a = 2.5$.

where ϵ is probability of erasure. We can interpret these results as follows; when $A/G(r) = A_1/\Theta(r^{-1})$, the total gain that Eve's A/D sees approaches infinity as $r \to \infty$; hence, even if Eve uses an A/D with larger range than Bob's A/D, her quantizer overflows. When $A/G(r) = A_2/\Theta(1)$, the total gain goes to zero as r approaches infinity and thus even if Eve uses an A/D with better precision, the received signal amplitude is less than one quantization level. In both cases, the eavesdropper receives no information about the transmitted signal and thus the eavesdropper's channel can be modeled by an erasure channel, where for $G(r) = \Theta(r^{-1})$, the probability of erasure $\epsilon = 1 - p$ and for $G(r) = \Theta(1)$, $\epsilon = p$.

To maximize the achievable secrecy rate, it is reasonable for Alice to choose $p = 0.5$. In Section IV-A it is shown that for a 10-bit A/D and the transmitter power $P = 1$, the optimum span of the A/D is $a = 2.5$, and the corresponding mutual information between Alice and Bob (when the channel between them is noiseless) is $I(X;Y) = 6.597$. Hence, using (16) $R_s \to 0.5 \times 6.597 = 3.2985$. Figure 3 shows the achievable secrecy rate versus r when both main and eavesdropper's channels are noiseless. It can be seen that as r gets larger, the achievable secrecy rate goes to a constant which is similar to what anticipated. Furthermore, for larger r's ($r \geq 10^3$) the actual optimum probability that maximizes the worst case secrecy rate is $p = 0.5$. These show that in the limit, our results are consistent to expectations.

IV. Numerical Results

A. Noiseless Channels: Eve with the same A/D as Bob

In this section, we begin our investigation by considering only the effect of A/D's on the signals. Hence, we assume that the eavesdropper's channel is noiseless. i.e. $n_E = 0$ (which benefits the eavesdropper). However, we also assume the system nodes are working in a very high SNR regime and thus the channel noise at Bob can be neglected ($n_B = 0$).

Now suppose that both Bob and Eve use 10-bit quantizers ($b = 2^{10}$) and the transmitter power is $P = 1$. Since $\delta = 2a/b$, for a fixed number of quantization bits, $I(X;Y)$ is a function

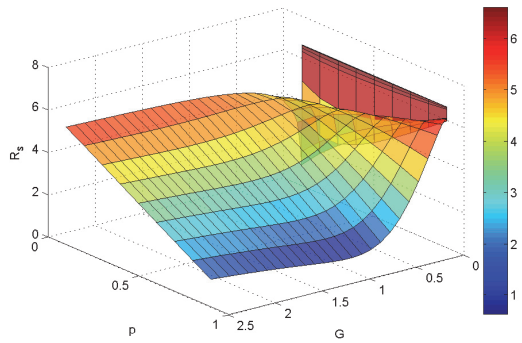

Fig. 4. Achievable secure rate vs. the probability p and the gain G at Eve's receiver. Both Bob's and Eve's channels are noiseless and they use identical 10-bit A/D's. The ratio between the two power levels at the transmitter is $r = 10^3$ (i.e. 30 dB) and the average transmitting power is $P = 1$. A maximin rate of $R_s = 3.1372$ is achieved.

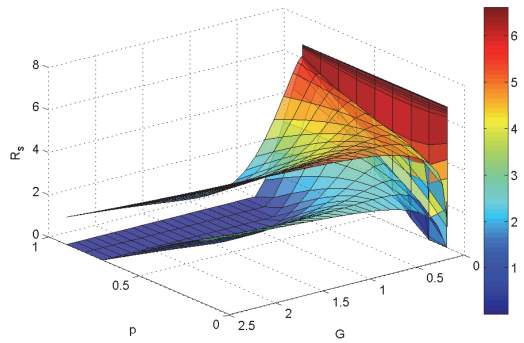

Fig. 5. Achievable secure rate vs. the probability p and the gain at Eve's receiver, G for the case of noiseless channels. The ratio between the two power levels at the transmitter is $r = 10^3$ (i.e. 30 dB) and the average transmitting power is $P = 1$. In the upper curve, both Bob and Eve have the same 10-bit A/D's. In the lower curve, Bob uses a 10-bit A/D while Eve uses a 14-bit A/D (Eve's A/D is 24 dB better than Bob's A/D) and a maximin rate of $R_s = 1.2478$ is achieved (for $p = 0.4$).

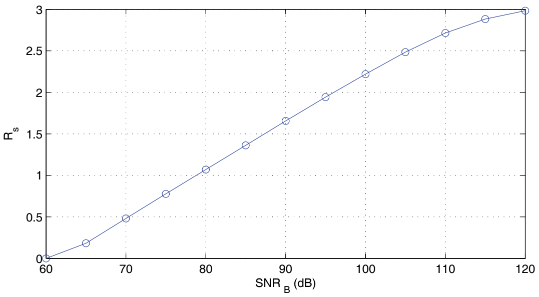

Fig. 6. Achievable secure rate vs. SNR at Bob's receiver while the SNR at Eve's receiver is infinity (Eve has perfect access to the transmitted signal) for $r = 10^3$, $P = 1$ and Bob and Eve applying 10-bit A/D's. Note that the assumption of Eve having a noiseless channel is the extreme case when the eavesdropper has perfect access to the transmitter's output.

of the span of the A/D (a), and the optimal quantization range that maximizes $I(X;Y)$ can be found. Since $I(X;Y)$ is an intricate function in terms of a, we find the optimum a numerically. In this case, the optimum quantization range that maximizes $I(X;Y)$ is $a = 2.5$, and the corresponding mutual information between Alice and Bob is $I(X;Y) = 6.597$. From now on, we use $a = 2.5$ in our calculations. Suppose that Eve has the same A/D as Bob. From the lemma above, putting a random gain is undesirable for Eve; hence, she chooses a fixed gain G that minimizes R_s. Because Alice is not aware of Eve's choice, she has to choose a probability p that maximizes the worst case R_s. The plot of R_s versus p and G for $P = 1$ and $r = 10^3$ (i.e 30 dB)where both Bob and Eve are using 10-bit A/D's is shown in Figure 4. This function is complicated and hence the optimum value of p cannot be derived analytically. Numerical analysis shows that $p \approx 0.5$ maximizes the worst case R_s, and the maxi-min value is $R_s = 3.1372$. Hence, choosing $p = 0.5$ guarantees that at least the secrecy rate $R_s = 3.1372$ can be achieved.

B. Noiseless Channels: Eve with a Better A/D than Bob

Now suppose that Eve has access to a better A/D than Bob. Depending on the gain that Eve applies before her A/D, a better A/D results in less erasures and/or less A/D overflows. Hence, the mutual information between Alice and Eve increases and consequently, the achievable secrecy rate decreases. Figure 5 shows the effect of using a better A/D on the achievable secrecy rate versus p and G. It can be seen that even if Eve uses an A/D which is 24 dB (4 bits) better than Bob's A/D (Eve has a 14-bit A/D while Bob has a 10-bit A/D), by choosing an appropriate value for p, a positive secure rate can be achieved. In this example, by choosing $p = 0.4$, a secure rate $R_s = 1.2478$ is achievable. Even if we do not change the probability p from the previous section ($p = 0.5$), assuming that Alice is not aware of Eve's better A/D, a secure rate $R_s = 0.6023$ is achievable. In spite of having a better A/D, Eve will still lose some symbols and hence a positive secrecy rate is available. This is because the ratio between the large and the small gain, A_1 and A_2, is 10^3, while Eve's A/D has only 16 times better resolution; thus, she still needs to compromise between resolution and overflow. To cancel the effect of these gains completely, Eve has to use an A/D that has an effective resolution after taking into account jamming, interference, etc. on the order of 10^3 times (10 bits) better than Bob's A/D, which would be very difficult in an adversarial environment.

C. Noisy Main Channel, Noiseless Eavesdropper's channel

Now we look at the extreme case that Eve is able to receive exactly what Alice transmits and receives (e.g. the adversary is able to pick up the transmitter's radio and hook directly to the antenna), but the channel between Alice and Bob is noisy and hence no other technique works. In other words, the channel between Alice and Bob experiences an additive white Gaussian noise ($n_B \sim \mathcal{N}(0, N_B)$), while Eve's channel is noiseless ($n_E = 0$). Figure 6 shows the secrecy rate R_s that can be achieved using the proposed scheme versus the signal to noise ratio (SNR) at Bob's receiver. In this case, the transmitted power $P = 1$ and the ratio between the large and the small gain is 30 dB. Both Bob and Eve use 10-bit A/D's and Alice sets $p = 0.5$. It can be seen that, although the eavesdropper's channel is much better than the main channel, when the SNR at Bob's receiver is greater than 60 dB, which is quite common in short-range communication, a positive secrecy rate is available.

V. Conclusion

In this paper, we introduce a new approach that exploits a short-term cryptographic key to force different orderings at Bob and Eve of two operators, one of which is necessarily non-linear, to obtain the desired advantage for information-theoretic security in a wireless communication system regardless of the location of Eve. We then investigate a simple power modulation instantiation of the approach. It is shown in [27] that when Eve's channel condition is better than the main channel, the secrecy rates that can be achieved using our proposed method are substantially higher than other methods designed to work in situations with an advantaged eavesdropper (such as public discussion). In particular, it is shown that in contrast to public discussion, even in the case that the adversary is able to pick up the transmitter's radio (i.e. Eve has perfect access to the output of the transmitter), a positive secrecy rate is achievable at high SNRs which might apply to a short-range wireless system. For example, one might use the transmission power of typical cellular systems with the corresponding excess power at short ranges to establish a secure radio system in a limited area.

Although we have considered the case of Eve with a better A/D than Bob, the clear risk to the approach is still that of asymmetric capabilities at the receivers. For example, if we employ the simple power modulation approach studied extensively here, Eve may employ multiple A/D's with different gain settings in front of each. Hence, Eve would be able to record two signals independently and decode them later when she gets the key or extracts the key based on the pattern of erasures and overflows at each A/D. A simple approach to combat this problem is, rather than applying a gain A with a discrete pdf at the transmitter, the transmitter can apply a gain with a continuous pdf. More promising, however, is to consider drawing the signal warping from a class of nonlinearities and adding memory to the signal warping process.

Broadly considering potential techniques for everlasting security in wireless systems, including that proposed here, yields that each approach still holds some risk. In the case of cryptographic security, assumptions must be made on both the hardness of the problem and the current/future computational capabilities of the adversary. In the case of information-theoretic security, assumptions must be made on the quality of the channel to Eve, generally corresponding to limitations on her location, In the method proposed here, assumptions must be made on Eve's current conversion hardware capabilities, but, as in information-theoretic secrecy, there is no assumption on future capabilties. All three approaches thus have different applicability.

References

[1] D. Stinson, *Cryptography: theory and practice*. CRC press, 2006.

[2] R. Benson, "The verona story," *National Security Agency Central Security Service, Historical Publications (available via WWW)*.

[3] C. Shannon, "Communication theory of secrecy systems," *Bell System Technical Journal*, vol. 28, no. 4, pp. 656–715, 1949.

[4] A. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[5] I. Csiszár and J. Korner, "Broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, 1978.

[6] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 733–742, 1993.

[7] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography. i. secret sharing," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1121–1132, 1993.

[8] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[9] R. Negi and S. Goel, "Secret communication using artificial noise," in *IEEE Vehicular Technology Conference, 2005*, vol. 62, p. 1906.

[10] S. Goel and R. Negi, "Secret communication in presence of colluding eavesdroppers," in *IEEE Military Communications Conference, 2005*, pp. 1501–1506.

[11] X. He and A. Yener, "Two-hop secure communication using an untrusted relay: A case for cooperative jamming," in *IEEE GLOBECOM 2008*, pp. 1–5, 2008.

[12] L. Lai and H. El Gamal, "Cooperative secrecy: The relay-eavesdropper channel," in *ISIT 2007*, pp. 931–935, 2007.

[13] P. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4687–4698, 2008.

[14] C. Cachin and U. Maurer, "Unconditional security against memory-bounded adversaries," *Advances in Cryptology*, pp. 292–306, 1997.

[15] R. Kuchibhatla, "Imft 25-nm mlc nand: technology scaling barriers broken," *EE Times News and Analysis*, 2010.

[16] S. Krone and G. Fettweis, "Fundamental limits to communications with analog-to-digital conversion at the receiver," in *IEEE 10th Workshop on Signal Processing Advances in Wireless Communications*, pp. 464–468, 2009.

[17] S. Krone and G. Fettweis, "A fundamental physical limit to data transmission and processing," *Signal Processing Letters, IEEE*, vol. 17, no. 3, pp. 305–307, 2010.

[18] R. Harjani, B. Sadler, H. Hashemi, and J. Rudell (Organizers), "Systems and circuits for sensing, co-existence, and interference mitigation in sdr and cognitive radios," in *IEEE RFIC Symposium*, 2011.

[19] L. Kahn, "Single-sideband transmission by envelope elimination and restoration," *Proceedings of the IRE*, vol. 40, no. 7, pp. 803–806, 1952.

[20] P. Kenington, *High linearity RF amplifier design*. Artech House, Inc., 2000.

[21] W. Diffie and M. Hellman, "New directions in cryptography," *IEEE Transactions on Information Theory*, vol. 22, no. 6, pp. 644–654, 1976.

[22] B. Widrow and I. Kollár, *Quantization noise*. Cambridge University Press, 2008.

[23] M. Bloch and J. Laneman, "On the secrecy capacity of arbitrary wiretap channels," in *46th Annual Allerton Conference*, pp. 818–825, 2008.

[24] T. Cover, J. Thomas, J. Wiley, *et al.*, *Elements of information theory*. Wiley, 2006.

[25] A. Calderbank and L. Ozarow, "Non-equiprobable signaling on the gaussian channel," *IEEE Transactions on Information Theory*, vol. 36, pp. 726–740, 1990.

[26] R. Kohno, R. Meidan, and L. Milstein, "Spread spectrum access methods for wireless communications," *IEEE Communications Magazine*, vol. 33, no. 1, pp. 58–67, 1995.

[27] A. Sheikholeslami, D. Goeckel, and H. Pishro-Nik, "Everlasting secrecy by exploiting non-idealities of the eavesdropper's receiver," *Arxiv preprint*, 2012.