

Fault Tolerant Key Generation and Secure Spread Spectrum Communication

Hussein Moradi, Arslan Majid, Behrouz
Farhang-Boroujeny

August 2017



The INL is a U.S. Department of Energy National Laboratory
operated by Battelle Energy Alliance

Fault Tolerant Key Generation and Secure Spread Spectrum Communication

Hussein Moradi, Arslan Majid, Behrouz Farhang-Boroujeny

August 2017

**Idaho National Laboratory
Idaho Falls, Idaho 83415**

<http://www.inl.gov>

**Prepared for the
U.S. Department of Energy
Office of Nuclear Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517**

Fault Tolerant Key Generation and Secure Spread Spectrum Communication

Arslan Majid, Hussein Moradi, and Behrouz Farhang-Boroujeny

Abstract—This paper presents a secure information transmission system compromised of two main parts. The first component makes use of the principle of reciprocity in frequency-selective wireless channels to derive a pair of keys for two legitimate parties. The proposed key generation algorithm allows for two asynchronous transceivers to derive a pair of similar keys. Moreover, we apply a unique augmentation - called *strongest path cancellation* (SPC) - to the keys and have validated through simulation and real-world measurements that this technique significantly boosts the security level of the overall design. In the second part of the secure information transmission system, we introduce the concept of artificial noise to multi-carrier spread-spectrum (MC-SS) systems. The keys generated in the first part are used as spreading code sequences for MC-SS and artificial noise is added to further enhance the security of this communication setup. Two different attacks on our proposed system are evaluated. First, a passive adversary who follows the exact same steps as the legitimate users to detect confidential information is considered. In the second attack, we evaluate the design in presence of an adversary with significant blind detection capabilities.

Index Terms—Spread-Spectrum, Reciprocal Channel Key Exchange, Secure Information Transmission, Physical Layer Security

I. INTRODUCTION

We consider the security issues involved in spread spectrum (SS) wireless communication systems. Such systems are used in applications which require resilience to harsh environments, resistance to channel fading through frequency diversity, low probability of detection (LPD), and low probability of interception (LPI) [1]. However, as with any wireless communication system, due to the broadcast nature of the communication, a passive eavesdropper (Eve) within range of broadcast can obtain the transmitted signal between a pair of legitimate users (Alice and Bob) and given a sufficient number of signal samples may be able to identify the spreading sequence and, hence, recover the transmitted information.

As of today, the security in SS has been limited mostly to their spreading sequence. It is often simply stated that a SS message signal transmitted by Alice cannot be recovered without the right spreading code. However, very little has been said on true security of the SS systems and most security solutions in the implementations of SS are limited. In fact,

surveys in [2] and [3] confirms that research in this area is open.

Real-world implementations of SS systems, e.g., IS-95 and IS-2000 standards [4], have used long-periodic pseudo-noise (PN) sequences in combination with a mask for physical-layer security. The mask is shared between mobile and base station, while the long-code PN sequence is defined by a 42-bit linear feedback shift register with a *publicly* known characteristic polynomial. Despite the long period of the PN sequence, it has been shown that an adversary with reasonable computational resources can implement a brute force attack in as little as 2.2 seconds [2]. Li et. al. [5] showed that an adversary with knowledge of the characteristic polynomial need only intercept 42 continuous long-code PN sequence bits to regenerate the entire long-code sequence. A solution proposed in [5] uses a combination of cryptography and physical-layer techniques to aid in scrambling the long PN sequence. However, the security of this method is reliant on the assumption of a computationally bounded adversary and is limited by secrecy of the encryption session key, assumed to be known *a-priori* in [5] between Alice and Bob.

Cryptography based solutions require key establishment, and key establishment for wireless networks is generally handled through public key cryptography (PKC) [6]. PKC comprises of a set of protocols including the well-known Diffie-Hellman [7]. Here, session keys are generated with the help of *trapdoor one-way functions* - functions that are computationally difficult to compute without a special code - the code being provided to the legitimate users by the certificate authority. PKC based methods require a lower bound assumption on the computational power of the adversary and are mathematically unproven to be secure [8]. Additionally, they are computationally expensive, hindering the application in devices with limited battery power.

An interesting alternative to the cryptography based approach to increase physical-layer security for Alice and Bob, is to make use of the following properties of the wireless channel:

- *Channel reciprocity*: The wireless channel between any pair of transceivers using the same wireless link experience the same fading properties (gains, phase shifts, and multipath delays).
- *Channel randomness*: Channel fading across time and frequency benefits from randomness due to Doppler spread and multipath delay spread, respectively.
- *Channel independence over space*: An adversary located more than a few wavelengths away from the legitimate

users experiences another random and uncorrelated channel.

These properties of the wireless channel allow for a pair of users to effectively share a secret - the secret being a realization of the channel - which is statistically uncorrelated for a third party located more than a few wavelengths away from the two users.

Application of wireless channels in physical-layer security is not new. In fact, there have been many papers that consider this topic. A good set of surveys for this area are [8]–[10]. In this line of work, there are at least three prominent research areas: 1) physical-layer key generation, 2) secure information transmission, and 3) theoretical bounds of secret key and secrecy capacity.

In physical-layer key generation, the wireless channel is used to obtain a secret key. In general, the procedure to generate a key from the fading channel requires the following steps: 1) randomness sharing, 2) information reconciliation, 3) privacy amplification, and 4) secure communication [11]. In *randomness sharing*, legitimate parties probe the reciprocal wireless channel between them. *Information reconciliation* requires the two nodes to communicate with one another to reconcile differences, or non-reciprocities, between their channel measurements. *Privacy amplification* is a process that maps reconciled channel measurements to a key whose maximum size depends on the randomness of the measurements and the amount of information leaked to the eavesdropper. Finally, in *secure communication* the parties transmit messages using the key either as a one-time pad or for use with a symmetric encryption algorithm.

Secure information transmission methods, which also rely on the channel reciprocity between Alice and Bob, make use of the channel state information (CSI) to degrade Eve's channel. Researchers in this field, e.g. [12]–[15], use the CSI between Alice and Bob to encrypt/pre-code the information bits B_i before transmission. When these pre-coded information bits are broadcast by Alice to Bob and Eve, the channel between Alice and Bob acts as a decryptor which (ideally) allows Bob to see the transmitted bits B_i and Eve to obtain $B'_i \neq B_i$.

One secure information transmission solution proposed by Goel et al. [15] introduces the concept of artificial noise. In this method, the degrees of freedom available in multiple antenna communication systems are utilized to generate the artificial noise. The produced artificial noise lies in the null-space of the legitimate user's channel while the information is transmitted in the range space of said channel. Hence, when the channel state information (CSI) is *perfectly* known at both the transmitter and receiver, the legitimate user's channel removes the artificial noise completely. Moreover, for eavesdroppers in different locations who experience their own unique channels, the artificial noise leaks into the eavesdropper's range-space causing a significant toll on the link quality of these users.

In this paper, we propose a secure information transmission system for SS communications that has two parts. First, we make use of the reciprocal wireless channel to derive a pair of spreading gain vectors as keys for the legitimate parties (Alice and Bob). Second, we extend the concept of artificial noise to spread-spectrum systems.

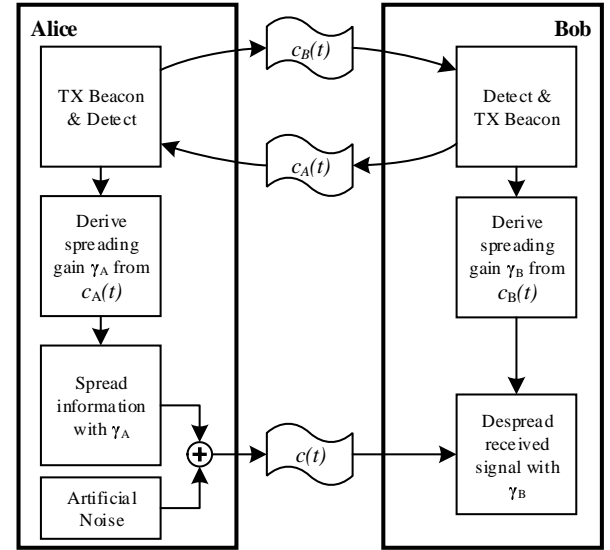


Fig. 1. Block diagram of proposed key generation method and secure communication.

For the secure key generation, we make use of the large bandwidth available in SS systems, hence, variation of the channel gain with frequency, to develop a novel method that leads to similar keys at Alice and Bob's nodes, but a significantly different key at Eve's node. The proposed method is also designed to take into account the fact that the Alice and Bob's nodes may be time asynchronous with respect to each other. Practicality of the proposed method is confirmed through a vast set of experimental works.

Discussion of artificial noise in the literature - e.g. [15]–[17] - has largely been in the context of multiple antenna systems. Rather than using multiple antenna to obtain the necessary dimensionality with which to transmit artificial noise, we introduce the use of chips in spread-spectrum for this purpose. The artificial noise is produced in the null-space of the spreading gain vector generated in the secure key generation step. To show our idea's robustness against eavesdroppers, we also study the attack scenario where an intelligent eavesdropper may take advantage of advanced blind detection methods to break the secret key. We show that by proper design of the system parameters, it is possible to avoid such attacks.

Figure 1 shows a block diagram of the secure information transmission system that we propose. In the first step, Alice transmits a beacon to Bob who detects it and transmits a beacon back to Alice. Both parties process their received beacons to derive spreading codes from the CSI. The processing that takes place here is detailed in a later section in this paper. In the final step, Alice encodes confidential information symbols using her spreading code and adds artificial noise before transmission. Bob applies his own spreading code to the received data for detection of the information symbols and to remove artificial noise. Eavesdroppers with significantly different spreading codes fail to recover the information symbols due to the artificial noise.

By the end of this paper, we hope to not only show that our

TABLE I
MATHEMATICAL NOTATIONAL CONVENTIONS

Symbol	Description
\mathbf{x}, \mathbf{X}	Vector and matrix
$\ \mathbf{x}\ $	Euclidean of vector \mathbf{x}
$\lfloor x \rfloor$	Rounding operation
$x(t)$	Continuous-time signal x
$x[n]$	Discrete-time signal x
$\{\cdot\}^H$	Matrix Hermitian
A, B, E	Subscripts for Alice, Bob, and Eve

approach can adequately secure SS systems, but also that the proposed SS-based technique is a robust solution for physical-layer security applications. At this point, we also note that our study will be focused on multi-carrier spread spectrum (MC-SS) techniques. MC-SS has been known in the literature for some time [1], [18], [19] and a version of MC-SS called filter-bank MC-SS (FB-MC-SS) has recently been successfully implemented by our group [20], [21].

This paper is organized as follows. Our adversary model is discussed in Section II. The proposed steps taken to convert the channel measurements to a key are in Section III. Section IV describes how the key is utilized for secure information transmission. Simulation and experimental results are presented in Section VI and concluding remarks are made in Section VII. Notation conventions for this paper are presented in Table I.

II. ADVERSARY MODEL

We assume Eve is a computationally unbounded passive eavesdropper and she can estimate the channel between herself and the legitimate parties. Eve performs the same steps as Alice and Bob in order to obtain her own key to detect the communicated data transmitted by Alice. Eve can be near the legitimate users (i.e. in our experiments, her antenna is placed 1/3 meter away from Bob) but she cannot be in the *exact* same location as Alice or Bob. Eve does not transmit channel probes and our current security solution does not authenticate the nodes. Eve does not jam the parties during channel probing, nor does she modify the transmitted messages before they are received by a legitimate party. Moreover, it is assumed a sufficiently clean channel is available between Alice and Bob so that they can obtain reciprocal estimates.

III. SECURE KEY GENERATION

In this section, we discuss the steps that Alice and Bob will go through to set up a secure key for information transmission. The first step is to measure the channel impulse response (CIR). The measured response is subsequently used for the key generation. Here, we proposed a key generation procedure whose result is a random vector that will be used as a spreading gain vector in a spread spectrum system. In particular, we emphasize on the measures that should be taken to assure the dissimilarity of the key generated by Eve with those of Alice and Bob, assuming that Eve is aware of the steps used by Alice and Bob to set their keys.

A. Channel Model

The wireless channel model of interest to us is the commonly used frequency-selective wideband channel model [22]

$$c(t) = \sum_{i \in \mathcal{M}} \alpha_i p(t - \tau_i) \quad (1)$$

where $\mathcal{M} = \{0, 1, \dots, M-1\}$ and M is the number of paths. The parameters α_i and τ_i are the complex gain and delay associated with the i^{th} path and $p(t)$ is the combined responses of the transmit and receive filters. In this paper, we call $p(t)$ the *probing pulse*, because of obvious reasons that will become clear as we proceed.

We also use (1) to represent the reciprocal channel between Alice and Bob. Eve's channel is represented by

$$c'(t) = \sum_{i \in \mathcal{M}'} \alpha'_i p(t - \tau'_i). \quad (2)$$

Considering *channel independence over space*, the channel parameters in (1) and (2) are assumed to be independent of each other.

B. Channel Estimation

In this preliminary step, Alice and Bob probe and estimate the wireless channel link that connect them together. To avoid interfering with one another, they resort to a time-division duplex (TDD) method in which Alice transmits a beacon packet to Bob who, upon receiving the beacon, immediately transmits the same packet back to Alice.

After the probing stage, channel estimation is carried out using the cyclic channel estimation procedure mentioned in [23]. The transmit beacon of choice is a length N Zadoff-Chu (ZC) sequence [24], [25]. A few periods of the ZC sequence is transmitted, and signal averaging is performed at the receiver for an accurate estimation of the channel. It is worth noting that the ZC sequence has seen widespread use in LTE and UMTS systems [26] due mostly to its special signal processing properties.

The received signal, after demodulation to baseband, is oversampled to a rate which is L times faster than the beacon symbols in ZC sequence. After averaging across multiple periods of the received signal, the L polyphase components of the signal sequence are separated, and the result is passed to a channel estimator following the least squares channel estimator of [23]. This leads to polyphase components of the channel estimates. These estimates are then interleaved to obtain the samples of channel impulse response at a sample interval $T_s = T_b/L$, where T_b is the time interval between the beacon symbols in the ZC sequence. This process which is performed by Alice, Bob, and Eve leads to the respective CIR estimates that we denote by $c_A[n]$, $c_B[n]$, $c_E[n]$. This channel estimation technique is advantageous in that it allows us to obtain the samples of CIR at a high resolution in time with a relatively low complexity, [23]. This, as will be found later, will become instrumental in development of an effective key generation algorithm. For a more thorough discussion of the channel estimation step, see [27].

C. Timing and Phase Synchronization

At this point, assuming that the channel is static within the probing interval, Alice and Bob both have their own discrete sample estimates $c_A[n]$ and $c_B[n]$ with three notable differences: 1) Because of TDD nature of probing, $c_A[n]$ and $c_B[n]$ are subject to a time misalignment; 2) $c_A[n]$ and $c_B[n]$ are affected by different phase errors, arising from the unsynchronized local oscillator (LO) of Alice and Bob, respectively; 3) $c_A[n]$ and $c_B[n]$ are affected differently by the channel noise. We can represent these differences in equation form as

$$c_A[n] = c(nT_s - \mu_A)e^{j\theta_A} + \eta_A[n] \quad (3a)$$

$$c_B[n] = c(nT_s - \mu_B)e^{j\theta_B} + \eta_B[n] \quad (3b)$$

where μ_A and μ_B are time delays, θ_A and θ_B are phase errors, and $\eta_A[n]$ and $\eta_B[n]$ are noise terms that arise from the channel noise. The parameters (μ_A, θ_A) and (μ_B, θ_B) are, in general, different between Alice and Bob and, hence, if uncompensated for can lead to a pair of dissimilar keys

To convert $c_A[n]$ and $c_B[n]$ to a pair of time and phase aligned CIRs, we proceed as follows. For time alignment, the center of the strongest path in both $c_A[n]$ and $c_B[n]$ are identified and shifted to a predefined location. Subsequently, for phase alignment, the elements of both CIRs are normalized with a pair of phase rotations that equalizes the phase of the samples that correspond to the center of the strongest path of both.

An early attempt to align $c_A[n]$ and $c_B[n]$ based on the strongest path has been reported in [27]. In this work, the location of the strongest path in $c_A[n]$ and $c_B[n]$ is found independently by both nodes and time aligned. This procedure may fail in the following scenario. When the CIR of the wireless link between Alice and Bob contains two or more strong paths with similar amplitudes, the presence of the noise terms $\eta_A[n]$ and $\eta_B[n]$ may lead to different locations for the strongest path in $c_A[n]$ and $c_B[n]$. As a result, the generated keys by Alice and Bob may be significantly different. In this paper, we solve this problem by taking the following approach. Our solution allows Bob to time align to Alice's strongest path by using some limited information that he gets (through a public channel) from Alice.

To start, Alice and Bob interpolate their respective CIRs $c_A[n]$ and $c_B[n]$ by a factor of L_2 , to further increase the time resolution of the available samples to them. The remaining steps are performed on these interpolated CIRs which we call $\tilde{c}_A[n]$ and $\tilde{c}_B[n]$ and assume to be a good approximations to the respective continuous time functions. In the subsequent discussions, we refer to the length of the interpolated CIRs $\tilde{c}_A[n]$ and $\tilde{c}_B[n]$ as N_c .

After interpolation, both nodes estimate the "path candidates" in their respective CIRs. Path candidates are considered to be a combination of estimated path gains and delays, which for Alice are respectively denoted by $\tilde{a}_{i,A}$ and $\tilde{k}_{i,A}$, for $i = 0, 1, \dots, M-1$, and are similarly defined for Bob. These parameters are determined by taking the following steps. Here, we have removed the subscripts A and B for simplicity, but it

should be understood that the presented steps are applied to both $\tilde{c}_A[n]$ and $\tilde{c}_B[n]$.

Step 0: Initialize $i = 0$ and $\tilde{c}_i[n] = \tilde{c}[n]$.

Step 1: Let

$$[\tilde{k}_i, \tilde{a}_i] = \arg \min_{\tilde{k}_i, \tilde{a}_i} ||\tilde{c}_i[n] - \tilde{a}_i p[n - \tilde{k}_i]||^2. \quad (4)$$

Step 2: Remove path candidate i from the CIR by taking

$$\tilde{c}_{i+1}[n] = \tilde{c}_i[n] - \tilde{a}_i p[n - \tilde{k}_i] \quad (5)$$

Step 3: Increment i by one and repeat **Step 1** and **Step 2** until $i = M$.

Once path candidates $(\tilde{a}_i, \tilde{k}_i)$ are calculated, the interpolated CIRs are time-shifted such that their largest path gain falls to the middle point of respective sequences. Note that this requires adjustment of the delay parameters \tilde{k}_i . Subsequently, the mean delay parameter

$$\bar{k} = \left[\frac{\sum_i \tilde{k}_i |\tilde{a}_i|^2}{\sum_i \tilde{k}_i} \right]. \quad (6)$$

is also calculated at both Alice's and Bob's nodes.

Next, Alice calculates the relative time difference between her estimate of the instantaneous mean delay and the location of her strongest path - which had been time aligned to the middle of the CIR. This results in a new delay parameter $k_D = \frac{N_c}{2} - \bar{k}_A$. Alice then transmits k_D to Bob. Note that this transmission does not need to be secure as this information has no value to Eve, whose channel has no similarity to Alice's or Bob's channel. Upon receiving k_D , Bob calculates the reference delay

$$k_{\text{ref}} = \bar{k}_B + k_D \quad (7)$$

At this point, k_{ref} should be a time location in Bob's CIR near the strongest path of Alice's CIR. However, non-reciprocities in the CIR along with estimation error muddle the location of Alice's strongest path relative to Bob's. To handle this problem, we propose that Bob solves the equation

$$[\tilde{k}_{\text{sp},B}, \tilde{a}_{\text{sp},B}] = \arg \min_{i \in [0, M-1]} ||p_i[n] - p[n - k_{\text{ref}}]||^2. \quad (8)$$

where $p_i[n] = \tilde{a}_{i,B} p[n - \tilde{k}_{i,B}]$. Note that (8) searches for the path candidate of Bob's channel which maximally correlates to $p[n - k_{\text{ref}}]$. The corresponding output $\tilde{k}_{i,B}$ in (8) is then used as Bob's reference point and is thus time aligned to the middle of the respective sequence. This procedure finalizes the time alignment of $\tilde{c}_A[n]$ and $\tilde{c}_B[n]$.

Once $\tilde{c}_A[n]$ and $\tilde{c}_B[n]$ are time aligned, they are circularly shifted so that the center of the strongest path of Alice's CIR and the matching strong path of Bob's CIR will be located at the time index $n = 0$. The results are subsequently decimated L_2 fold to obtain a pair of channel estimates of length NL . Lastly, the channel estimates are phase aligned by introducing a phase shift to the elements of each CIR such that the path located at time index $n = 0$ has phase of zero. We call the final time and phase aligned CIRs $\hat{c}_A[n]$ and $\hat{c}_B[n]$.

D. Strongest Path Cancellation

Now that the timing and phase offsets have been resolved, let us consider a passive adversary, Eve, who follows the exact same synchronization steps as Alice for her own estimated CIR to obtain $\hat{c}_E[n]$. Without getting into the detail, we note that by following Alice's time alignment steps, Eve can better synchronize with the legitimate users than she could by following Bob's alignment procedure.

For simplicity, we ignore the channel noise term and, thus, note that the final CIR estimates for Alice and Eve can be expressed as

$$\hat{c}_A[n] = |\tilde{\alpha}_{sp,A}|p[nL_2] + \sum_{i \in \mathcal{M}_{\setminus sp}} \tilde{\alpha}_{i,A}p[nL_2 - \tilde{k}_{i,A}] \quad (9)$$

$$\hat{c}_E[n] = |\tilde{\alpha}_{sp,E}|p[nL_2] + \sum_{i \in \mathcal{M}'_{\setminus sp}} \tilde{\alpha}_{i,E}p[nL_2 - \tilde{k}_{i,E}]. \quad (10)$$

where $\mathcal{M}_{\setminus sp}$ and $\mathcal{M}'_{\setminus sp}$ contains the set of all paths excluding the strongest path for Alice and Eve, respectively.

Next, we define the length NL CIR vectors $\hat{\mathbf{c}}_A = \{\hat{c}_A[n]\}$, $\hat{\mathbf{c}}_B = \{\hat{c}_B[n]\}$, and $\hat{\mathbf{c}}_E = \{\hat{c}_E[n]\}$. Also, we let $\mathbf{p} = \{p[n]\}$. Given (9) and (10), the cross-correlation between Alice and Eve's CIR estimates can be expressed as

$$\varrho_{AE} = \frac{\hat{\mathbf{c}}_A^H \hat{\mathbf{c}}_E}{\|\hat{\mathbf{c}}_A\| \|\hat{\mathbf{c}}_E\|}. \quad (11)$$

Evaluation of (11) using (9) and (10) gives

$$\varrho_{AE} = \varrho_{sp,AE} + \varrho_{\setminus sp,AE} \quad (12)$$

where

$$\varrho_{sp,AE} = \frac{|\tilde{\alpha}_{sp,A}| |\tilde{\alpha}_{sp,E}| \|\mathbf{p}\|^2}{\|\hat{\mathbf{c}}_A\| \|\hat{\mathbf{c}}_E\|} \quad (13)$$

is a positive and relatively large term arising from the time and phase synchronized strongest paths of Alice and Eve, and $\varrho_{\setminus sp,AE}$ is the residual cross-correlation arising from the remaining paths. Since these remaining paths are not synchronized, their cross-correlations are usually a set of zero-mean, low variance random variables that add-up to a statistically small value. This observation leads us to the following proposal.

To minimize the similarity of the keys generated by Alice and Bob with the key that Eve generates, Alice or Bob should remove the strongest paths of their respective synchronized CIR estimates and use the residual responses to set the keys. We call this method strongest path cancellation (SPC) and use $\bar{c}_A[n]$, $\bar{c}_B[n]$, and $\bar{c}_E[n]$ to denote the residual CIRs for Alice, Bob, and Eve, respectively. For instance, Alice's CIR after removal of the strongest path is obtained as

$$\bar{c}_A[n] = \hat{c}_A[n] - |\tilde{\alpha}_{sp,A}|p[nL_2]. \quad (14)$$

Similarly equations are used to obtain the residual CIRs of Bob and Eve.

Our assumption, here, which has been validated through an extensive set of 32.5 MHz wide indoor wireless channel measurements has confirmed that the residual CIRs have sufficient information to assure highly correlated keys for Alice and Bob, while leading to a dissimilar key for Eve.

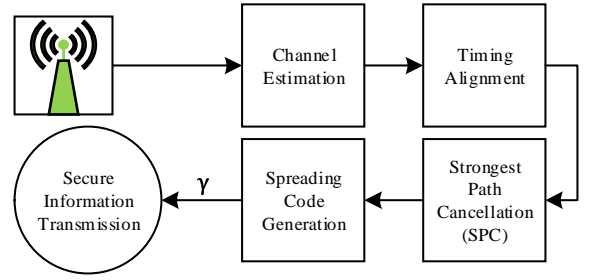


Fig. 2. Block diagram of proposed key generation method in which spreading gains are generated from the channel impulse response.

E. Key Generation

The final step in obtaining a spreading code sequence from the reciprocal wireless channel is outlined here. All parties follow the same procedure as Alice who first takes the DFT of $\bar{c}_A[n]$ and stores it in $\bar{\mathbf{C}}_A$. Next, a key is constructed as

$$\left\{ \gamma_A = \frac{\bar{\mathbf{C}}_A[\mathbf{m}]}{\|\bar{\mathbf{C}}_A[\mathbf{m}]\|} \mid \mathbf{m} \in [\text{Passband of } \bar{\mathbf{C}}_A] \right\} \quad (15)$$

by Alice, and γ_B and γ_E are generated similarly by Bob and Eve, respectively.

At this point, we note that further steps can be taken to build a more secure key. For instance, in [27] a method is brought up in which the users take the key to be the summation of the phase of the frequency response of the channel with a shuffled version of the same signal. This key has a nice property which further decorrelates Eve's key from the legitimate users'. However, after consideration of the artificial noise discussed in the following section, we have found that by simply adopting the passband response of the channel, we get a better overall performance than the key from [27].

The key generation procedure discussed in this section is summarized in Fig. 2. Once obtained, these sequences are used as an integral part of the secure information transmission communication system that is discussed in the following section.

IV. SECURE INFORMATION TRANSMISSION

In this section, the proposed secure information transmission system is detailed. First, the mathematical model of the solution is given. Next, we propose our artificial noise transmit strategy and finally, the security level of the proposed solution is analyzed.

Alice is taken to be the node that wishes to transmit confidential information to Bob while Eve is a passive eavesdropper listening to Alice's transmit signal. A simplified approach to the physical-layer will be taken so that analysis is straightforward.

A. MC-SS with Artificial Noise

The secure information transmission system proposed in this section makes use of the chips available to SS systems to produce artificial noise and, thus it is easily adoptable to single-antenna systems. This concept differs significantly from

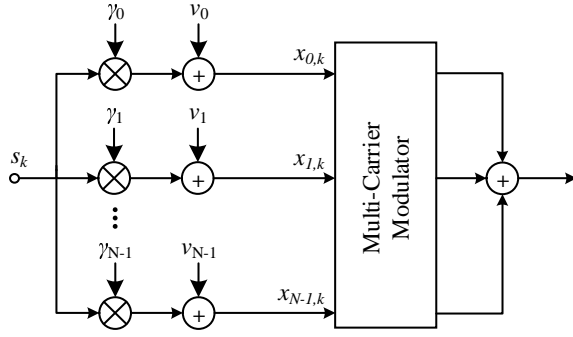


Fig. 3. MC-SS transmitter with artificial noise block diagram

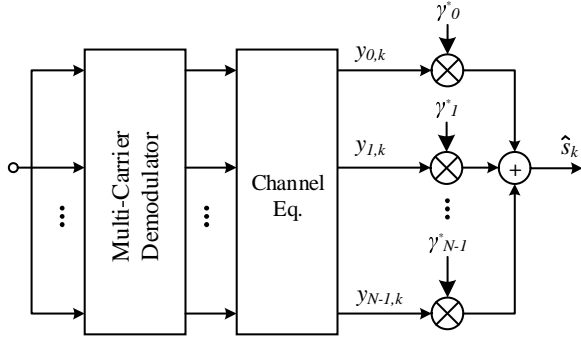


Fig. 4. MC-SS receiver block diagram

those in the literature - e.g. [15]–[17] - which use multiple antennas as means of obtaining the necessary dimensionality with which to produce artificial noise.

At this point, we note that we will only consider MC-SS as the transmit waveform for reasons explained later in this section. A block diagram for the proposed MC-SS transmitter with the addition of artificial noise is shown in Figure 3 and its corresponding receiver is in Figure 4. Note that the receiver does not need additional circuitry to account for the artificial noise since it is removed by the despreader and any residual error due to artificial noise leaking into the information space is taken as additive noise. Additionally, the "Multi-Carrier Modulator" and "Multi-Carrier Demodulator" blocks in Figure 3 and 4 are meant to allow for any MC-SS based waveform design such as OFDM, FB-MC-SS [20], etc. We purposely do not restrict our study to any particular MC-SS based method as the following system model can be easily extended to any MC-SS waveform.

Following the transmitter in Figure 3, Alice constructs the transmit signal using the key from (15) as

$$\mathbf{x}_k = \gamma_A s_k + \mathbf{v}_k \quad (16)$$

where $k = 0, 1, \dots, K-1$ and \mathbf{v}_k is artificial noise vector, added to increase security in presence of an eavesdropper. The artificial noise \mathbf{v}_k is selected to lie in the null-space of γ_A , so that $\gamma_A^H \mathbf{v}_k = 0$. More explicitly, \mathbf{v}_k is generated as the residual between an i.i.d circularly symmetric complex Gaussian noise vector - \mathbf{w}_k - and its projection onto the space of γ_A as follows

$$\mathbf{v}_k = \mathbf{w}_k - (\gamma_A^H \mathbf{w}_k) \gamma_A \quad (17)$$

The total transmit power across the entire occupied bandwidth can be obtained by combining (16) and (17). This gives

$$\begin{aligned} P &= E[\mathbf{x}_k^H \mathbf{x}_k] \\ &= \sigma_s^2 + \frac{N-1}{N} \sigma_w^2 \end{aligned} \quad (18)$$

where

$$\sigma_w^2 = E[\mathbf{w}^H \mathbf{w}] \quad (19)$$

and the term $\frac{N-1}{N}$ arises from the fact that artificial noise is generated from a complete N -dimensional vector space with one of its dimensions removed. We denote the fraction of power allocated to the information signal as ϕ . This implies that

$$\sigma_s^2 = \phi P \quad (20)$$

and

$$\sigma_w^2 = \frac{(1-\phi)NP}{N-1}. \quad (21)$$

Following (16), the signal received by Bob and Eve *after* multi-carrier demodulation and application of a zero-forcing channel equalizer are respectively given by

$$\mathbf{y}_k = \mathbf{x}_k + \boldsymbol{\eta}_k \quad (22)$$

$$\mathbf{z}_k = \mathbf{x}_k + \boldsymbol{\epsilon}_k \quad (23)$$

where the components of $\boldsymbol{\eta}_k$ and $\boldsymbol{\epsilon}_k$ arise from channel noise. Note that the elements of $\boldsymbol{\eta}_k$ and $\boldsymbol{\epsilon}_k$ may not be i.i.d due to frequency selectivity of the channel. The SNR at the Alice-Bob link and Alice-Eve link, thus, can be expressed as

$$\text{SNR}_B^i = \frac{P}{\sigma_\eta^2} \quad (24)$$

and

$$\text{SNR}_E^i = \frac{P}{\sigma_\epsilon^2} \quad (25)$$

where $\sigma_\eta^2 = E[\boldsymbol{\eta}_k^H \boldsymbol{\eta}_k]$ and $\sigma_\epsilon^2 = E[\boldsymbol{\epsilon}_k^H \boldsymbol{\epsilon}_k]$. Note that the superscript 'i' is added to the SNR terms to emphasize that these are at the receiver input.

Next, Bob and Eve despread their received signals from (22) and (23) with their own spreading gains to get

$$\begin{aligned} \gamma_B^H \mathbf{y}_k &= \gamma_B^H (\mathbf{x}_k + \boldsymbol{\eta}_k) \\ &= \gamma_B^H \gamma_A s_k + \gamma_B^H \mathbf{v}_k + \gamma_B^H \boldsymbol{\eta}_k \end{aligned} \quad (26)$$

and

$$\begin{aligned} \gamma_E^H \mathbf{z}_k &= \gamma_E^H (\mathbf{x}_k + \boldsymbol{\epsilon}_k) \\ &= \gamma_E^H \gamma_A s_k + \gamma_E^H \mathbf{v}_k + \gamma_E^H \boldsymbol{\epsilon}_k. \end{aligned} \quad (27)$$

The SNR at Bob's node after the despreader is derived in Appendix A and is found to be

$$\text{SNR}_B^o = \frac{N\phi\rho_{AB}\text{SNR}_B^i}{\frac{N}{N-1}(1-\phi)(1-\rho_{AB})\text{SNR}_B^i + 1} \quad (28)$$

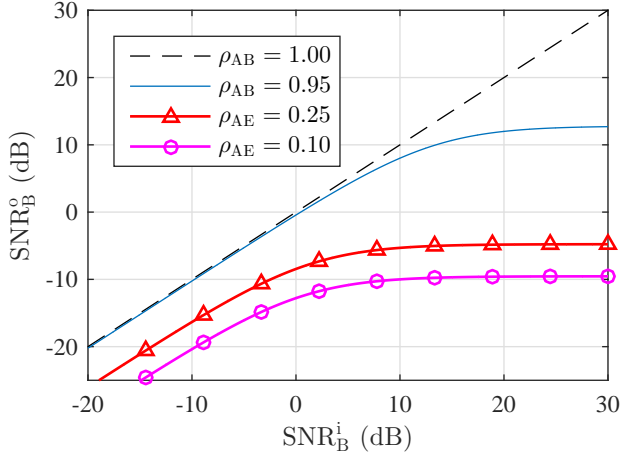


Fig. 5. Plot of the SNR after despreading versus receiver SNR for $\phi = 1/N$ and selected values of ρ_{AB} and ρ_{AE} .

where

$$\rho_{AB} = |\gamma_B^H \gamma_A|^2 \quad (29)$$

and the superscript ‘o’ is added to the SNR terms to emphasize that it is at the output, i.e., after the despreader. Equations (28) and (29) are defined similarly for the Alice-Eve link with the appropriate substitutions.

Note when there is no artificial noise, i.e. $\phi = 1$, (28) reduces to $\text{SNR}_B^o = N\rho_{AB}\text{SNR}_B^i$. This shows that the despreading procedure, through coherent linear combination of the received signal vector allows Bob to achieve an SNR up to N times the link SNR given in (24). On the other hand, if Eve can gain access to the spreading code γ_A , she can decode the information symbols sent by Alice. This highlights the necessity of the algorithm discussed in Section III where Alice and Bob make use of the reciprocal wireless channel to generate a pair of similar keys while Eve, despite following the same steps as the legitimate nodes, generates a significantly different key.

To enlighten the reader on how artificial noise can be used to boost security, we plot SNR_B^o and SNR_E^o as a function of the received signal SNR at $\phi = 1/N$ in Figure 5. We use the terms SNR^i and SNR^o to respectively describe the SNR before and after despreading at either Bob or Eve’s node, depending on the context. The values of ρ_{AB} and ρ_{AE} were chosen arbitrarily to show that when Alice and Bob share nearly identical keys, they have a significant SNR advantage compared to the adversary who generates a different key.

B. Artificial Noise Transmit Strategy

An important parameter for artificial noise transmission systems is the signal-to-artificial noise ratio ϕ . In [15], this parameter is found through maximization of the secrecy capacity - defined as the maximum rate at which Alice-Bob can communicate a message without an eavesdropper being able to decode it. This power allocation strategy is applicable only when Alice has perfect knowledge of the full CSI - e.g. the CSI between herself and Bob and herself and Eve. Since the knowledge of perfect CSI between Bob and Eve is not

possible in practice, this method of selection of the artificial noise power may be only of interest from a theoretical point of view.

Here, we are interested in practical scenarios where Eve’s CSI is not known to Alice. In this case, secrecy cannot be guaranteed as Alice does not know how much artificial noise to inject to Eve’s channel. With the assumption of Eve’s CSI remaining unknown, an approach taken by [16] distributes ϕ to meet a target SNR at Bob’s node, assuming that the Alice-Bob CSI is perfectly known. The rest of the available transmit power is devoted to artificial noise, hoping this will sufficiently deteriorate Eve’s channel such that she will not be able to decode the transmit message.

Our power allocation strategy follows the same idea of dedicating enough power to the information subspace to ensure a certain link quality between Alice and Bob. However, we do not make the assumption that perfect knowledge of Alice-Bob CSI is available. Instead, we propose an artificial noise power allocation strategy in which Alice dedicates enough power to the information to ensure a target SNR is met for Bob so long as the similarity of their keys - quantified by ρ_{AB} - is larger than a threshold ρ_{\min} .

The parameter ϕ for our signal-to-artificial noise power allocation strategy can be determined by replacing SNR_B^o in (28) with a target SNR - SNR_T^o - and ρ_{AB} with threshold ρ_{\min} . Solving for ϕ with these substitutions in place gives

$$\phi = \frac{\frac{N}{N-1}(1 - \rho_{\min})\text{SNR}_B^i + 1}{\frac{N}{N-1}(1 - \rho_{\min})\text{SNR}_B^i + N\rho_{\min}\frac{\text{SNR}_B^i}{\text{SNR}_T^i}}. \quad (30)$$

As a check, note that if the spreading codes are assumed perfectly known by Alice and Bob, hence, $\rho_{\min} = 1$, the above reduces to

$$\phi = \frac{\text{SNR}_T^o}{N\text{SNR}_B^i}. \quad (31)$$

This is the same as the result reported in [16].

One point to be noted here is that there is a limit to how low ρ_{\min} can be set for a minimal signal-to-artificial noise ratio ϕ_{\min} that a designer would consider, given a target SNR. To determine this limit, we set $\phi = \phi_{\min}$ in (30) and solve for ρ_{\min} to get

$$\lim_{\text{SNR}_B^i \rightarrow \infty} \rho_{\min} = \frac{\text{SNR}_T^o(1 - \phi_{\min})}{\text{SNR}_T^o(1 - \phi_{\min}) + (N - 1)\phi_{\min}} \quad (32)$$

V. SECURITY LEVEL OF PROPOSED SOLUTION

In this section, two different attack scenarios are evaluated to justify the security of the proposed solution.

A. Scenario 1: The Passive Eavesdropper

For the first attack scenario, Eve is a passive eavesdropper who tries to decode Alice’s transmitted information symbols using the key she generates. Eve is equipped with the same receiver as Bob and the only difference among them is their despreading codes.

To evaluate the security level of this attack, we first consider use of the secrecy characterization from [28], where the notion of secrecy outage probability is expressed as

$$\mathcal{P}_{\text{out}}(R_s) = \mathcal{P}(C_B - C_E < R_s). \quad (33)$$

where $C_B = \log_2(1 + \text{SNR}_B^0)$, $C_E = \log_2(1 + \text{SNR}_E^0)$.

The definition in (33) makes an assumption that the transmitter chooses a strategy that leads to the main-link communicating near capacity of the channel. In this way, while SNR_B^0 will allow for sufficient information recovery at Bob's node, SNR_E^0 will be inadequate in decoding information at Eve's node. The parameter R_s is effectively a margin that when chosen larger, increases the secrecy outage probability. An outage is said to occur when a message is either unreliable for Bob to decode or insecure, i.e. there will be a possibility that Eve decode the message.

A known weakness of the secrecy characterization by (33) was first discussed in [29]. It was noted that (33) does not distinguish between reliability and security. The secrecy outage probability may be minimized for a given set of design parameters, but it is not obvious from (33) whether this is due to an information leak or a reliability issue.

Accordingly, the following alternative definition of the secrecy outage probability was proposed. The outage was defined for when the difference between the target capacity and Eve's capacity is lower than R_s conditioned on the event that a message was transmitted. In our model, we assume that message transmission always occurs since Alice and Bob are operating independent of one another. The secrecy outage probability definition from [29] is thus modified as

$$\mathcal{P}_{\text{out}}(R_s) = \mathcal{P}(R_T - C_E < R_s) \quad (34)$$

where

$$R_T = \log_2(1 + \text{SNR}_T^0) \quad (35)$$

The definition of (34) states that an outage occurs when the SNR after despreading at Eve's node is within the margin of R_s from the target rate R_T . The characterization in (34) is useful from a practical standpoint in which Alice and Bob are operating independent of one another. In such a case, Alice chooses a code that optimally works (i.e. error-free) for target rate R_T . This is different from the characterization in (33) where it is implied that if $\text{SNR}_B^0 > \text{SNR}_T^0$, then Alice chooses a different code to work at rate C_B rather than the target rate R_T . It also follows that if Alice and Bob are communicating near target rate R_T , an appropriate definition for an information leak is the scenario in which C_E is near R_T rather than the case where C_E is close to C_B .

With regards to the reliability of the main-link, we note that a nice feature of our artificial noise power allocation strategy is that it assures a target SNR is met so long as the similarity between the keys generated by Alice and Bob is above a threshold. In other words, it can be easily verified that if ϕ is obtained from (30), then $\mathcal{P}(\text{SNR}_B^0 < \text{SNR}_T^0)$ is equal to $\mathcal{P}(\rho_{AB} < \rho_{\min})$.

B. Scenario 2: The Sophisticated Eavesdropper

In this attack scenario, Eve is given a significant advantage in decoding the transmitted data. In traditional artificial noise systems, e.g. [15]–[17], the assumption of a block-fading channel model limits the number of symbols that can be transmitted confidentially. In our system, we can transmit as many symbols as needed with a given spreading code since the channel is not used to directly decrypt the information. The consequence of encoding many symbols with the same spreading gain sequence is that a sophisticated adversary may use a *blind method* to identify the information signal subspace and subsequently use that knowledge to decode the communicated data.

Without artificial noise (i.e. $\phi = 1$), secrecy against knowledgeable adversaries could *only* exist for our system model if 1) Eve is at an SNR disadvantage compared to the main-link or 2) Alice only transmits one symbol per key, effectively applying a one-time pad to the solution provided the key is generated at channel coherence time intervals. Both of these assumptions are considerably strong to impose on the security of a wireless communication network. Here, we study the use of artificial noise as a way to increase throughput of the secure communication system without assuming Eve to be at a disadvantage.

This attack scenario considers the situation where Alice transmits K symbols with the same spreading code sequence and Eve seeks to estimate γ_A from her received signal. To facilitate this study, we redefine (23) as a matrix of concatenated received signal vectors spread with the same key

$$\begin{aligned} \mathbf{Z} &= [\mathbf{z}_0 \quad \mathbf{z}_1 \quad \dots \quad \mathbf{z}_{K-1}] \\ &= [\mathbf{x}_0 + \mathbf{e}_0 \quad \mathbf{x}_1 + \mathbf{e}_1 \quad \dots \quad \mathbf{x}_{K-1} + \mathbf{e}_{K-1}] \end{aligned} \quad (36)$$

where $\mathbf{Z} \in \mathbb{C}^{N \times K}$.

The columns of \mathbf{Z} are a set of random vectors. Each of these vectors have the average energy/power of P and the form of (23). There is a fixed direction γ_A that carries data symbols with the power ϕP . The rest of the power is in a random direction perpendicular to γ_A . When $\phi = 1/N$, the signal power is equally distributed in all directions, including the direction γ_A . In this scenario, the signal space will appear to be white with respect to all directions, including the data direction, making it hard for an observer that wishes to find γ_A . The situation will be different when $\phi \neq 1/N$. In such cases, the intruder can search for the direction that carries a different power than the remaining directions.

The standard solution to find the signal direction, i.e., the spreading gain vector γ_A , when $\phi > 1/N$, is the following.

- 1) Construct the $N \times N$ matrix

$$\mathbf{R}_{ZZ} = \frac{1}{K} \mathbf{Z} \mathbf{Z}^H \quad (37)$$

- 2) Invoking the Rayleigh-Ritz Theorem [30] an estimate of γ_A is obtained by solving the following maximization problem

$$\hat{\gamma}_A = \arg \max_{\|\gamma\|=1} \gamma^H \mathbf{R}_{ZZ} \gamma \quad (38)$$

For this procedure to give an accurate estimate, the number of signal samples (i.e., the parameter K) should be sufficiently

large. To give an idea of how large K should be to obtain a reasonable estimate of γ_A , we resort to some numerical results which are presented in the next section.

VI. RESULTS

The proposed key generation and secure information transmission system are tested in this section. Simulation results are presented so that results can be repeated, confirmed, and numerically evaluated. Additionally, experimentation results are obtained to validate the performance of the proposed system in real-world environments.

A. Key Generation

1) *Simulation Results:* To show how well the proposed key generation algorithm uses both time and amplitude to its advantage, we run a simulation. The parameters chosen for the simulation match the experiment. For the simulation, we assume a block fading channel model and no fractional timing offset for all three parties. Given this setup, Monte Carlo simulations were processed according to the following procedure.

- 1) Alice and Bob generate a probing beacon consisting of 25 periods of a length $N = 64$ ZC sequence. This is interpolated by a factor of $L = 4$ using a square-root raised-cosine filter with a roll-off factor of 1/2 and transmitted at a sampling rate of $\frac{1}{T_s} = 130\text{MHz}$.
- 2) The beacon is transmitted across a simulated wireless channel. The channel follows an exponential power delay profile with delay spread of 50 ns [22], sampled at uniform intervals of LT_s . The complex gain of each multipath component is Rayleigh faded and the total number of effective multipaths is set to $M = \lfloor 10 \times \frac{50\text{ ns}}{LT_s} \rfloor = 16$. Alice and Bob share the same channel and the only difference between the Alice-Bob and Alice-Eve's channels are the M complex-valued gains.
- 3) Noise is independently added to the signal received by Alice, Bob, and Eve.
- 4) Channel estimation, time alignment, and key generation are processed according to the procedure described in Section III. Note that for time alignment, Eve time aligns according to her own strongest path as our experiments have shown more similarity to the Alice-Bob channel is observed with this approach.

Fig. 6 shows the cumulative distribution function (CDF) of ρ_{AB} and ρ_{AE} before and after SPC from 10,000 runs of the described simulation. Alice and Bob have an SNR of 10 dB, while Eve has zero additive noise in her received signal. In addition to these curves, the cross-correlation between the M complex gains of the Alice-Bob and Alice-Eve channels are also plotted and this is denoted by ρ_R .

A few interesting aspects of the proposed key generation algorithm can be found in Figure 6. First, as expected, a slight decorrelation occurs between Alice and Bob's keys after SPC due to removal of the strongest path. However, this decorrelation is not quite as large as one would expect. In fact, the average value of ρ_{AB} before and after SPC at the present SNR of 10 dB is 0.994 and 0.985, respectively.

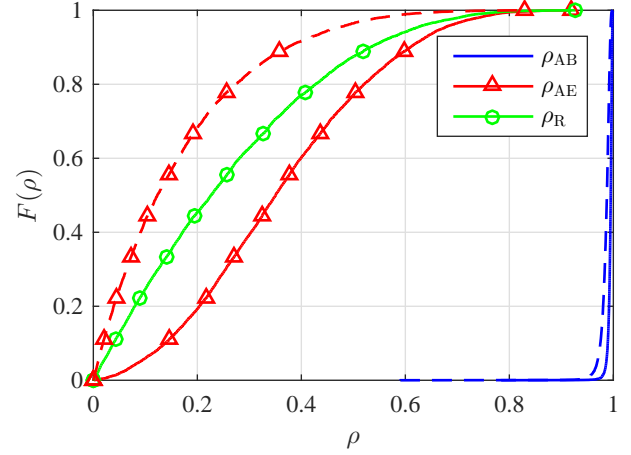


Fig. 6. CDF of cross-correlation between Alice and Bob's keys ρ_{AB} and Alice and Eve's keys ρ_{AE} from simulation results using a channel model in which path gains are derived from an exponential power delay profile with RMS delay spread of 50 ns. Dashed lines show results after SPC is applied. Additionally, ρ_R shows the cross-correlation between the M complex-valued gains of Alice-Bob and Alice-Eve channels

Next, consider the curves depicted in Figure 6 which show the cross-correlation between Alice and Eve's keys before and after SPC, as well as the parameter ρ_R . First, it can be seen that before SPC, timing and phase recovery causes Alice and Eve's keys to be relatively strongly correlated. However, after SPC, the cross-correlation bias due to time and phase synchronization (13) is removed and thus the similarity between the keys is significantly less.

Perhaps the most interesting aspect here is that after SPC, ρ_{AE} is statistically much less than ρ_R despite the fact that 1) the time-delays of each multipath component are the same for all channels and 2) there is one less source of randomness due to removal of the strongest path. The reason for this is because after time alignment according to the strongest path, the time delays relative to the strongest path are different between the Alice-Bob and Alice-Eve channels. After removing the strongest path, the remaining multipaths add an additional secret - the time delays of the residuals paths relative to the strongest.

Ultimately Figure 6 shows that the time delays, in combination with the complex-valued gains of the channel, allow Alice and Bob to share a stronger secret than they would if *only* the complex gains of the channel were for key generation.

2) *Experimental Results:* The experiment is run on a transceiver based on the National Instruments (NI) platform. The transceiver consists of an NI FlexRIO FPGA Module (NI PXIe-7975R). This module is connected to an NI FlexRIO RF Transceiver (NI 5791R), which has a sampling rate of 130 MHz. The FPGA and Transceiver module are both connected to an NI real-time controller (NI PXIe-1082), which is used as a host PC and is programmed using NI LabVIEW Real-Time. The FlexRIO RF Transceiver is connected to a circulator (Model No. CS-0.900) that is fed to an RF amplifier (NI PXI-5691) and then to a single antenna. All three parties in our experiment (Alice, Bob, and Eve) use identical transceiver setups and Eve's transmitter is turned off. Experiments are

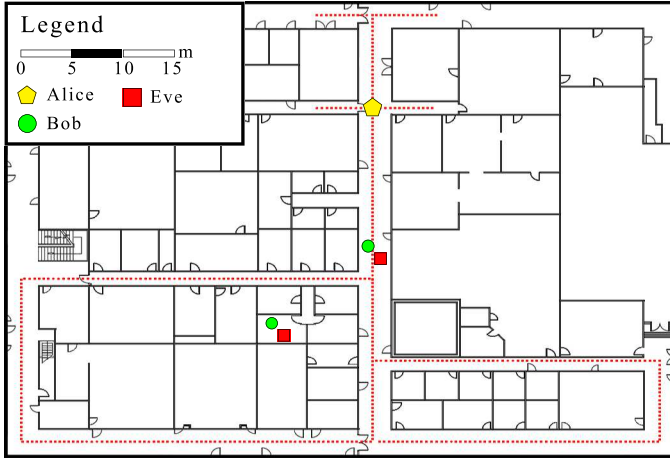


Fig. 7. Experimental setup on the third floor of Merrill Engineering Building at the University of Utah. The position of Alice was varied across the dotted lines while Bob and Eve remained stationary in one of the two displayed locations. Eve was synchronized to Bob's clock, while Alice and Bob operated on asynchronous clocks. The antenna of Bob and Eve were placed approximately 1/3 meter apart.

run at a carrier frequency of 900 MHz.

Time-division duplexing is used for channel probing. The duration of each transmitted packet, consisting of multiple repetitions of the ZC sequence is $118\mu\text{s}$. A few extra ZC sequences are prepended to the packet for packet detection purposes. The time duration between the time it takes for Alice to measure Bob's channel and vice versa is ~ 1 ms.

Details of the experimental setup are explained in Figure 7. In total, 6500 channel measurements are captured. Prior to obtaining each measurement, the environment around Alice is varied, either by moving Alice or having an experimenter move around the node to ensure variation between measurements. Data is collected from over-the-air measurements and subsequently used to generate keys offline.

Figure 8 shows the CDF of ρ_{AB} and ρ_{AE} before and after SPC from the experimental data. Similar to Fig. 6, we see a slight decorrelation between Alice and Bob's keys after SPC as well as a significant increase of dissimilarity between Alice and Eve's keys. Results shown in Fig. 6 and Fig. 8 are fairly similar, though over-the air measurements show Alice and Eve's keys to be slightly more correlated in the experiment than in the simulation. A possible cause for this is that the channel model used in our simulation contains more randomness and/or paths than observed in the measurements. Figures 6 and 8 indicate that Eve's key has been decorrelated through SPC. However, the question looming at this point is whether this is worth the reduction in similarity between Alice and Bob's keys. In the following section, we examine this very point. In addition, we recognize that the correlation between the elements within N -length key is high. We find, remarkably, that this matters less when a large amount of artificial noise is used.

B. Secure Information Transmission

The secure information transmission system that we propose adds artificial noise to the traditional MC-SS as a means of

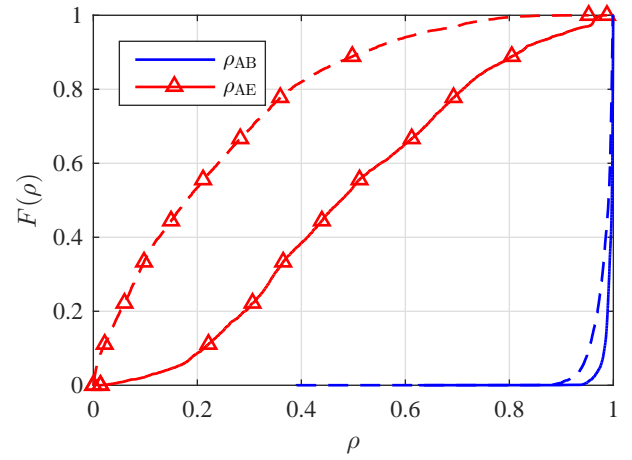


Fig. 8. CDF of cross-correlation between Alice and Bob's keys ρ_{AB} and Alice and Eve's keys ρ_{AE} from over-the-air data obtained from the experiment whose setup is shown in Fig. 7. Dashed lines show results after SPC is applied.

enhancing physical layer secrecy. At this point, we turn our attention to validating the security of the proposed system and its expected performance for the legitimate parties. We take $N = 64$ since it is the value of N that was used in our spreading gain generation experiments as well as in our current implementation of the FB-MC-SS system in [21].

We consider the limit of (28) as the SNR approaches infinity. To simplify the discussion here, we define SNR^o in (28) as the despread signal SNR at Bob or Eve's link. Similarly, SNR^i and ρ respectively from (24) and (29) are defined in this way. Using L'Hospital's Rule,

$$\lim_{\text{SNR}^i \rightarrow \infty} \text{SNR}^o = \frac{(N-1)\phi\rho}{(1-\phi)(1-\rho)}. \quad (39)$$

It is trivial to see from (39) that if no artificial noise is used (i.e. $\phi = 1$) and as $\text{SNR}_E^i \rightarrow \infty$, SNR_E^o will also increase to infinity, hence, no secrecy can be guaranteed regardless of how dissimilar Alice and Eve's keys are. However, the addition of artificial noise (i.e. when ϕ drops below one) provides an intriguing opportunity to securely transmit confidential information despite Eve having a significant SNR advantage. Fig. 9 plots (39) as a function of ρ for different values of ϕ .

In previous literature of artificial noise, where perfect CSI knowledge is assumed (i.e. $\rho_{AB} = 1$), high values of secrecy data rates can be achieved. However, as it can be seen in Fig. 9, if $\rho_{AB} < 1$, there is an exponential drop in SNR_B^o which can lead to a significant data reliability issue that get exacerbated as artificial noise power is increased. This highlights the main advantage of our artificial noise power allocation strategy in (30) because it compensates for dissimilarity between Alice and Bob's keys by strategically introducing enough artificial noise such that a target rate is hit for a given ρ_{\min} that fits the criteria of (32).

In Fig. 9, it can also be seen that the limit of the despread signal's SNR linearly increases for $\{\rho \mid 0.2 < \rho < 0.8\}$. Moreover, (39) asymptotically approaches infinity as $\rho \rightarrow 1$ and approaches zero as $\rho \rightarrow 0$. In short, this trend is encouraging as it shows that the introduction of artificial noise allows for nodes with the "right" key to reliably decode

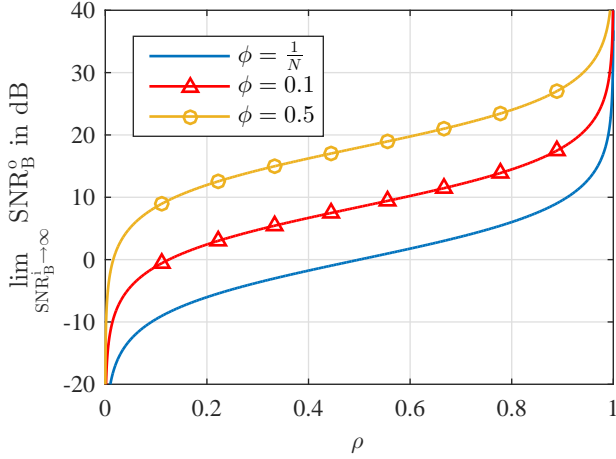


Fig. 9. Plot of the SNR after despreading - SNR_B^0 - as link SNR approaches infinity as a function of ρ - a measure of similarity between Alice and Bob's keys - for selected values of ϕ .

the confidential information while it hampers the decoding ability of users with different keys, even when they have a considerable SNR advantage.

Next, we discuss the two scenarios of the "passive eavesdropper" and "sophisticated eavesdropper".

1) *Scenario 1: The Passive Eavesdropper:* Results from the simulations and experiments in Section VI-A are used to evaluate the key generation procedure in the context of the proposed secure information transmission system. Two sets of keys will be compared: the key generated before SPC and the key after applying SPC.

As discussed before, Eve follows the same steps as the main-link in retrieving the transmitted symbols sent by Alice. The effectiveness of this attack is evaluated using the secrecy outage probability in (34). To ensure fair comparison between the two sets of keys, ρ_{\min} is set so that 95% of the keys used will meet the target SNR after despreading, i.e., $\mathcal{P}(\text{SNR}_B^0 < \text{SNR}_T^0) = \mathcal{P}(\rho_{AB} < \rho_{\min}) = 5\%$.

In this way, ρ_{\min} will be smaller for the keys that use SPC due to the slight decorrelation effect that SPC has on the keys. In turn, this means less artificial noise can be added for the keys obtained using SPC. Note that this formulation uses *a-priori* knowledge of ρ_{AB} to determine ρ_{\min} , but this is only used to ensure a fair comparison between the two sets of keys. In practice, when *a-priori* knowledge is not available, ρ_{\min} should be set differently. A method that we propose for this practical scenario is detailed in [31].

The secure information transmission strategy we propose is one in which the target rate is adaptive to SNR_B^i . For the adaptive target rate strategy, when SNR_B^i is too low to meet a minimum target rate $R_{T\min}$ at $\phi = 1/N$, the signal to artificial noise ratio is calculated using (30). When the SNR at Bob's receiver is high enough and thus a large amount of artificial noise power can be added (i.e. $\phi = \phi_{\min} = 1/N$), then the target rate is increased. To find the target rate in this scenario, we first solve for SNR_T^0 in (30) at $\phi = 1/N$ to obtain

$$\text{SNR}_T^0 = \frac{\text{SNR}_B^i \rho_{\min}}{1 + \text{SNR}_B^i (1 - \rho_{\min})} \quad (40)$$

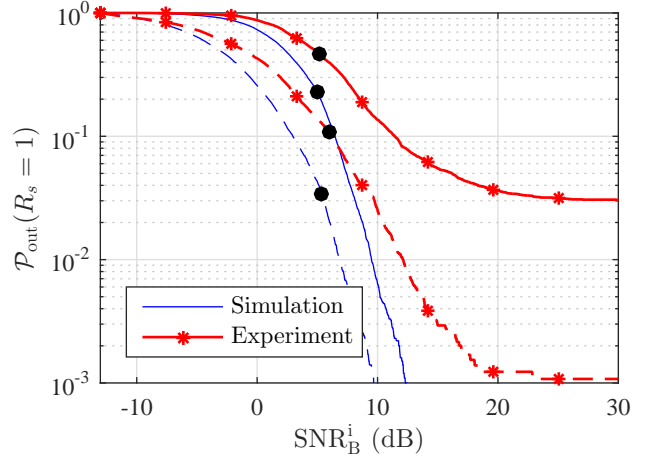


Fig. 10. Plot of the outage probability of secrecy evaluated at $R_s = 1$ as a function of receiver SNR for the adaptive target rate strategy. The plot was obtained using the ρ_{AB} and ρ_{AE} values from keys obtained through simulation and experiment. Dashed lines show keys after SPC is applied. Black circles indicate the transition point at which the target rate R_T starts increasing

and then calculate R_T using (35).

Fig. 10 shows evaluation of (34) at $R_s = 1$ for the passive eavesdropper attack for simulation and experimentation results when using the adaptive target rate strategy. The solid lines correspond to before SPC and the dashed lines correspond to after SPC. To generate this figure, we assume a worst-case scenario where Eve has zero additive noise at the receiver. The minimum target rate $R_{T\min}$ is set to 2 bits and is incremented according to SNR_B^i . Additionally, a reliability of 95% at the main-link is met for all SNR values in Fig. 10. Note that to guarantee this reliability, the smallest value for SNR_B^i corresponds to $\phi = 1$ for the keys applied with SPC. Below the minimum value of SNR_B^i , there is not enough transmit power at Alice's node to allow for the minimum target rate of 2 bits.

Results from Fig. 10 indicate that the keys derived using SPC provide a significant boost to the security of the system. This is despite the fact that less artificial noise is being broadcast at lower values of SNRs (i.e. the SNR_B^i values to the left of the black circles) as a result of the way ρ_{\min} was obtained. It can also be seen that when SNR_B^i is high and $\phi = 1/N$, the probability of secrecy outage approaches a steady state. For the experiment data set, the steady state value for the secrecy outage probability is $\approx 3\%$ for keys that do not use SPC, while keys applied with SPC are $\approx 0.1\%$. Therefore, the minimum amount of security is better with SPC than without. Finally, we note that this transmit strategy allows us to use high levels of artificial noise power ($\phi = 1/N$), which is not only good in securing communications from a secrecy outage standpoint but also beneficial in thwarting the efforts of the more sophisticated adversary that we discuss next.

2) *Scenario 2: The Sophisticated Eavesdropper:* In this section, we show two different sets of results pertaining to the case wherein multiple symbols are transmitted with the same spreading code sequence. First, we examine the transmitted sequence correlation matrix and show that although there is correlation between elements in the key, it matters less as artificial noise power is increased. Next, we examine the

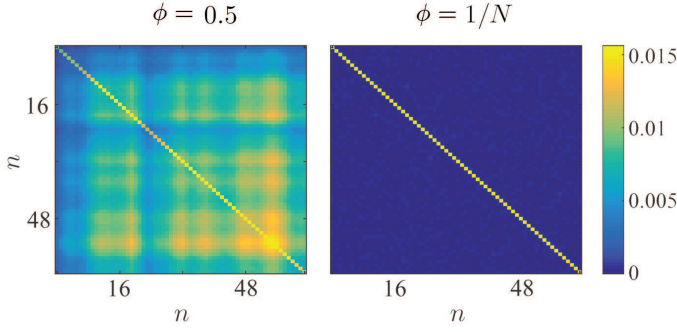


Fig. 11. Plot of the magnitude of covariance matrix of \mathbf{x}_k for one γ_A from the experiment data set for $\phi = 0.5$ and $\phi = 1/N$.

amount of symbols that can be transmitted with one key as artificial noise power is varied.

First, consider the transmitted data vector \mathbf{x}_k from (16). It is shown in Appendix B that the covariance matrix of this signal, for a given γ_A can be expressed as

$$\mathbb{E}[\mathbf{x}_k \mathbf{x}_k^H] = \frac{1}{N} \sigma_w^2 \mathbf{I}_N + \left(\sigma_s^2 - \frac{1}{N} \sigma_w^2 \right) \gamma_A \gamma_A^H \quad (41)$$

where \mathbf{I}_N is an $N \times N$ identity matrix.

Figure 11 plots the magnitude of the covariance matrix of an example γ_A coming from our experimental data set for two extreme cases of ϕ . Here, it is interesting to see that when $\phi = 1/N$, the second term in (42) vanishes and, hence, the covariance matrix of the transmit sequence will be identity. The significance of this finding is that when $\phi = 1/N$, the signal direction γ_A will not be observable in the correlation matrix $\mathbb{E}[\mathbf{x}_k \mathbf{x}_k^H]$ and, thus, any method that seeks to estimate γ_A by exploring the second order moments of \mathbf{x}_k will be unsuccessful.

Next, we use numerical results to evaluate the effectiveness of the blind attack from a sophisticated adversary as discussed in Section V-B. The goal of this simulation is to examine the number of symbols K that can be sent with a given key γ_A . To start, Alice transmits K symbols with a given spreading code sequence obtained from the experimental data set. The information symbols are encoded with binary phase-shift keying (BPSK) so that $s_k = \pm \sigma_s$ and we assume the worst-case eavesdropper who has zero channel noise, and thus $\mathbf{z}_k = \mathbf{x}_k$.

Once Eve receives K symbols, she constructs the matrix \mathbf{Z} as in (36). Next, the Rayleigh-Ritz theorem (38) is applied to obtain a blind estimate of the spreading code sequence. This process is run for increasing values of K from 4 to 256, and different choices of ϕ . To evaluate the effectiveness of this attack, we calculate the similarity between $\hat{\gamma}_A$ and γ_A using (29).

Figure 12 plots the 99th percentile of ρ_{AE} , when γ_E is set equal to $\hat{\gamma}_A$, as a function of K for varying values of ϕ . The 99th percentile shows the line where 99% of the time, ρ_{AE} remains below it. As observed for larger values of ϕ , i.e., when the level of artificial noise is relatively low, the eavesdropper may be able to obtain a reasonable estimate of γ_A within a relatively small number of observed samples. However, as

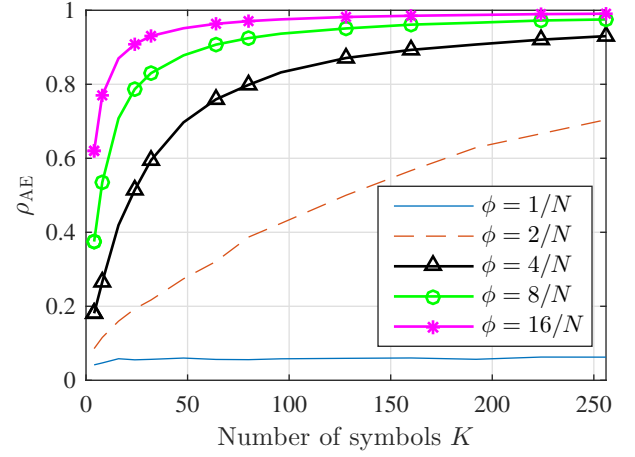


Fig. 12. Plot of the 99th percentile of ρ_{AE} as a function of K . Note that the 99th percentile indicates that 99% of the time, ρ_{AE} is below the lines indicated in the graph

ϕ increases, it becomes more difficult to estimate γ_A . For the golden ratio of $\phi = 1/N$, since $\mathbb{E}[\mathbf{x}_k \mathbf{x}_k^H]$ become the identity matrix, almost all the estimates of γ_A remain nearly orthogonal to γ_A , hence, an almost sure secure communication can be guaranteed.

VII. CONCLUSION

In this paper, we proposed and studied a secret-key enabled secure information transmission system for spread spectrum communication. We presented a method through which two asynchronous radios can exchange a set of spreading codes through the use of the reciprocal wireless channel. The key itself is shown to have been made stronger through the use of a method that we named SPC (Strongest Path Cancellation). We validated our approach through both simulation and experimentation and showed that the use of SPC greatly aids our proposed secure information transmission solution.

The secure information transmission solution proposed here introduces the concept of artificial noise to multi-carrier spread-spectrum systems as a means of enhancing the physical-layer security in wide band communications. The solution was tested against both a passive eavesdropper who follows the same procedure of Alice and Bob as well as the more sophisticated adversary who seeks to blindly detect the key transmitted by Alice. In the first situation, it was shown that despite SPC introducing a slight decorrelation between keys of Alice and Bob, the probability of a secrecy outage remains in favor of the key generated using SPC. For the more sophisticated adversary, we made the following observation. When Alice and Bob have enough SNR to take advantage of, by adding sufficient artificial noise, they will be able to communicate many information symbols securely.

APPENDIX

A. Derivation of SNR after despreading

Using (26), we note that Bob's received signal after despreading is

$$\gamma_B^H y_k = \gamma_B^H \gamma_A s_k + \gamma_B^H \mathbf{v}_k + \gamma_B^H \boldsymbol{\eta}_k. \quad (42)$$

The SNR after despreading is taken to be

$$\text{SNR}_B^o = \frac{\text{Var}[\gamma_B^H \gamma_A s_k]}{\text{Var}[\gamma_B^H \mathbf{v}_k + \gamma_B^H \boldsymbol{\eta}_k]}. \quad (43)$$

The numerator in (43) is evaluated as

$$\text{Var}[\gamma_A^H \gamma_B s_k] = \rho_{AB} \sigma_s^2 \quad (44)$$

For the denominator, we know that the noise and artificial noise are uncorrelated and consequently the variance of the two terms can be separated. Using (17), we get

$$\begin{aligned} \text{Var}[\gamma_B^H \mathbf{v}_k] &= \text{E}[\|\gamma_B^H \mathbf{v}_k\|^2] \\ &= \text{E}[(\mathbf{w}_k - (\gamma_A^H \mathbf{w}_k) \gamma_A)^H \gamma_B \gamma_B^H (\mathbf{w}_k - (\gamma_A^H \mathbf{w}_k) \gamma_A)] \\ &= \sigma_w^2 (1 - \rho_{AB}). \end{aligned} \quad (45)$$

and since γ_B and $\boldsymbol{\eta}_k$ are uncorrelated, one will find that

$$\text{Var}(\gamma_B^H \boldsymbol{\eta}_k) = \frac{\sigma_\eta^2}{N} \quad (46)$$

where $\sigma_\eta^2 = \text{E}[\boldsymbol{\eta}_k^H \boldsymbol{\eta}_k]$ is the total noise power across the occupied bandwidth. Finally, by combining (44), (45), and (46), and recalling (24) and (29), one can obtain (28).

B. Derivation of covariance matrix of transmit sequence

The covariance matrix of the transmit signal \mathbf{x}_k is evaluated in this section. We start by using the definition of the transmit signal equation from (16) to obtain

$$\text{E}[\mathbf{x}_k \mathbf{x}_k^H] = \text{E}[(\gamma_A s_k + \mathbf{v}_k)(\gamma_A s_k + \mathbf{v}_k)^H] \quad (47)$$

The terms in (47) can be evaluated as follows

$$\text{E}[\gamma_A s_k s_k^H \gamma_A^H] = \sigma_s^2 \gamma_A \gamma_A^H \quad (48)$$

$$\text{E}[\gamma_A s_k \mathbf{v}_k^H] = \text{E}[\mathbf{v}_k s_k^H \gamma_A^H] = 0 \quad (49)$$

$$\text{E}[\mathbf{v}_k \mathbf{v}_k^H] = \frac{1}{N} \sigma_w^2 \mathbf{I}_N - \frac{1}{N} \sigma_w^2 \gamma_A \gamma_A^H \quad (50)$$

where (49) follows from the assumption that artificial noise is uncorrelated to the information symbols s_k . Also, in writing (50) we have recalled the definition of σ_w^2 from (19). Combining the results in (48), (49), and (50) lead to (41).

REFERENCES

- [1] K. Cheun, K. Choi, H. Lim, and K. Lee, "Antijamming performance of a multicarrier direct-sequence spread-spectrum system," *Communications, IEEE Transactions on*, vol. 47, no. 12, pp. 1781–1784, 1999.
- [2] Y.-S. Shiu, S. Y. Chang, H.-C. Wu, S. C.-H. Huang, and H.-H. Chen, "Physical layer security in wireless networks: a tutorial," *Wireless Communications, IEEE*, vol. 18, no. 2, pp. 66–74, 2011.
- [3] T. Kang, X. Li, C. Yu, and J. Kim, "A survey of security mechanisms with direct sequence spread spectrum signals," *Journal of Computing Science and Engineering*, vol. 7, no. 3, pp. 187–197, 2013.
- [4] M. A. Abu-Rgheff, *Introduction to CDMA wireless communications*. Academic Press, 2007.
- [5] T. Li, Q. Ling, and J. Ren, "Physical layer built-in security analysis and enhancement algorithms for cdma systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, no. 3, p. 7, 2007.
- [6] H. Imai, *Wireless communications security*. Artech House, Inc., 2005.
- [7] W. Diffie and M. E. Hellman, "New directions in cryptography," *Information Theory, IEEE Transactions on*, vol. 22, no. 6, pp. 644–654, 1976.
- [8] J. Zhang, T. Q. Duong, A. Marshall, and R. Woods, "Key generation from wireless channels: A review," *IEEE Access*, vol. 4, pp. 614–626, 2016.
- [9] K. Ren, H. Su, and Q. Wang, "Secret key generation exploiting channel characteristics in wireless communications," *Wireless Communications, IEEE*, vol. 18, no. 4, pp. 6–12, 2011.
- [10] Y. E. H. Shehadeh and D. Hogrefe, "A survey on secret key generation mechanisms on the physical layer in wireless networks," *Security and Communication Networks*, vol. 8, no. 2, pp. 332–341, 2015.
- [11] M. Bloch, J. Barros, M. R. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *Information Theory, IEEE Transactions on*, vol. 54, no. 6, pp. 2515–2534, 2008.
- [12] H. Koorapaty, A. Hassan, and S. Chennakeshu, "Secure information transmission for mobile radio," in *Information Theory, 1998. Proceedings. 1998 IEEE International Symposium on*. IEEE, 1998, p. 381.
- [13] A. O. Hero III, "Secure space-time communication," *Information Theory, IEEE Transactions on*, vol. 49, no. 12, pp. 3235–3249, 2003.
- [14] G. R. Tsouri and D. Wulich, "Reverse piloting protocol for securing time varying wireless channels," in *Wireless Telecommunications Symposium, 2008. WTS 2008*. IEEE, 2008, pp. 125–131.
- [15] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 6, pp. 2180–2189, 2008.
- [16] A. L. Swindlehurst, "Fixed sinr solutions for the mimo wiretap channel," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2009, pp. 2437–2440.
- [17] X. Zhou and M. R. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 8, pp. 3831–3842, 2010.
- [18] K. Fazel and S. Kaiser, *Multi-Carrier and Spread Spectrum Systems: From OFDM and MC-CDMA to LTE and WiMAX: Second Edition*. Wiley, 2008.
- [19] G. K. Kaleh, "Frequency-diversity spread-spectrum communication system to counter bandlimited gaussian interference," *Communications, IEEE Transactions on*, vol. 44, no. 7, pp. 886–893, 1996.
- [20] D. L. Wasden, H. Moradi, and B. Farhang-Boroujeny, "Design and implementation of an underlay control channel for cognitive radios," *Selected Areas in Communications, IEEE Journal on*, vol. 30, no. 10, pp. 1875–1889, 2012.
- [21] T. Haddadin, A. Laraway, A. Majid, T. Sibbet, B. Lo, D. Couch, L. Lloyd, H. Moradi, D. Wasden, and B. Farhang-Boroujeny, "An underlay control channel for dynamic spectrum access: Packet design, implementation, analysis, and experimental results," in *IEEE International Conference on Computer Communications (Pending Review)*, 2016.
- [22] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. Prentice Hall PTR New Jersey, 1996, vol. 2.
- [23] B. Farhang-Boroujeny, *Signal processing techniques for software radios*. Lulu publishing house, 2008.
- [24] R. Frank, S. Zadoff, and R. Heimiller, "Phase shift pulse codes with good periodic correlation properties (corresp.)," *IRE Transactions on Information Theory*, vol. 6, no. 8, pp. 381–382, 1962.
- [25] D. Chu, "Polyphase codes with good periodic correlation properties (corresp.)," *IEEE Transactions on Information Theory*, pp. 531–532, 1972.
- [26] S. Sesia, I. Toufik, and M. Baker, *LTE: the UMTS long term evolution*. Wiley Online Library, 2009.
- [27] A. Majid, H. Moradi, and B. Farhang-Boroujeny, "Secure information transmission for filter bank multi-carrier spread spectrum systems," in *Military Communications Conference, 2015, IEEE (pending publishing)*, 2015.
- [28] J. Barros and M. R. D. Rodrigues, "Secrecy capacity of wireless channels," in *Proc. IEEE Int. Symp. Information Theory*, Jul. 2006, pp. 356–360.
- [29] X. Zhou, M. R. McKay, B. Maham, and A. Hjørungnes, "Rethinking the secrecy outage formulation: A secure transmission design perspective," *IEEE Communications Letters*, vol. 15, no. 3, pp. 302–304, Mar. 2011.
- [30] F. Zhang, *Matrix theory: basic results and techniques*. Springer Science & Business Media, 2011.
- [31] A. Majid, "Secure communications in filter-bank multi-carrier spread spectrum systems," Ph.D. dissertation, The University of Utah, 2016.