# Intrinsic Information of Wideband Channels

Yuan Shen, *Student Member, IEEE*, and Moe Z. Win, *Fellow, IEEE*

***Abstract*—The ability to exchange secret messages and protect against security attacks becomes increasingly important for providing information superiority and confidentiality in modern information systems. These systems require shared secret keys, which can be generated from common random sources with known distributions. However, the assumption on the distribution of the sources may not hold in many realistic scenarios. In this paper, we establish a mathematical framework for secret-key generation using common *unknown deterministic* sources (UDSs). In particular, we propose a new information measure called *intrinsic information* to characterize the achievable length of the secret key that can be generated from a UDS. As a case study, we consider a wideband propagation medium in mobile wireless networks as a UDS and derive its intrinsic information as a function of various network parameters. Our results provide a non-Bayesian perspective for secret-key generation as well as practical implications of this new perspective.**

***Index Terms*—Secret-key generation, physical-layer security, wideband channels, non-Bayesian inference.**

## I. Introduction

SECURE COMMUNICATION is essential for providing information superiority and confidentiality in modern information systems. It is a challenging task in wireless systems, due to the broadcast nature of radio frequency (RF) transmission. Contemporary wireless security methods, employing cryptographic protocols (e.g., RSA) at upper layers of a network, rely on the computational difficulty of some number-theoretic problems (e.g., factorization) [1]–[3]. Another branch of secure communication techniques, known as physical-layer security or information-theoretic security, exploits properties of the physical medium (e.g., wireless channels) [4]–[9]. Such security techniques require no assumption on the adversary's computational capability, and they can complement existing upper-layer security algorithms for enhancing communication secrecy. Moreover, there is an emerging research trend in network secrecy that leverages intrinsic properties of large wireless networks for secure communication [10]–[14].

Information-theoretic security originates from the notion of perfect secrecy [4], which can only be achieved when the secret key is at least as long as the message. This work was then extended by the seminal work [5]–[7], in which the adversary on a wiretap channel receives degraded messages compared to the legitimate nodes. It was shown later in [8], [9] that a positive secrecy capacity can still be achieved even when the adversary has a better reception, provided that the legitimate nodes can also communicate over an insecure but authenticated channel. Physical-layer security research has followed two main directions, secret-key agreement [15]–[20] and secrecy capacity [21]–[27]. Their analysis can be unified by the source and channel type models [8]. In essence, all the work on physical-layer security is built upon the availability of the randomness with a known distribution, either for the source or the channel, shared between the legitimate nodes. For example, in secret-key agreement, two legitimate nodes aim to agree on a key that is secret to the eavesdropping node by using correlated observations via public discussion [8], [9].

The reciprocity[1] of wireless propagation channels [28]–[31] has been considered as a source of common randomness for secret-key generation using both narrowband transmission [32]–[34] and ultra-wide bandwidth (UWB) transmission [35]–[37]. Propagation of UWB signals in multipath environments [38]–[42] provides a rich common source between two legitimate nodes. This source is ideal for generating secret keys, since it is difficult for eavesdropping nodes to intercept or observe the channel between legitimate nodes. Based on the random source model [9], the secret-key rate was derived for UWB channels where the channel gains are assumed to be Gaussian random variables (RVs) [35]. However, the Gaussian assumption of the channel gains often does not hold, especially for UWB signals in multipath environments. In many scenarios, it may even be unrealistic to statistically model the unknown parameter of the source [43]. This leads to a fundamental question: what is the achievable length of the secret key that can be generated from common unknown deterministic sources (UDSs)?

To the best of the authors' knowledge, no research has addressed the use of common UDSs for secret-key generation from an information-theoretic perspective. Note that existing information measures are either not well-defined or not suited for unknown deterministic parameters (UDPs) in the context of secret-key generation. For example, mutual information only deals with RVs and is undefined without a known distribution [44], Fisher information quantifies the bound for the estimation error and is often independent of the parameter value [43], and quantization only gives an approximate number of reliable bits in the presence of observation noises in high SNR regimes [45].

In this paper, we establish a mathematical framework for secret-key generation using common UDSs. We determine the *secret-key length* via an information-theoretic analysis, starting from binary representation of real-valued parameters and then extending to the general case. As a case study, we consider the wideband channels in mobile wireless networks as a common UDS between two legitimate nodes for secret-key generation.

Manuscript received September 15, 2012; revised March 1, 2013. This research was supported, in part, by the National Science Foundation under Grants CCF-1116501, the Office of Naval Research under Grant N00014-11-1-0397, and MIT Institute for Soldier Nanotechnologies.

The authors are with the Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139 USA (e-mail: shenyuan@mit.edu, moewin@mit.edu).

Digital Object Identifier 10.1109/JSAC.2013.130919.

[1] The channel reciprocity principle between two transceivers refers to the scenario in which the channel impulse responses from each transceiver to the other are identical except for observation noises.

The main contributions of the paper are as follows.

- We propose a new information measure called *intrinsic information* and characterize the achievable length of the secret key that can be generated from a UDS.
- We derive the intrinsic information of real-valued parameters represented in different bases, obtain its functional properties, and quantify the information loss due to finite-base representations.
- We determine the intrinsic information of a wideband propagation medium in mobile wireless networks, providing insights into the potential of time-varying wideband channels for secret-key generation.

The outcome of this work can be used to complement contemporary cryptography systems and foster new secure and authentication applications in mobile wireless networks.

The rest of the paper is organized as follows. In Section II, we present the preliminaries including the problem statement and a process of secret-key generation. In Section III, we give the main results for the intrinsic information. In Section IV, we determine the intrinsic information of wideband channels in a mobile wireless network. Finally, numerical and simulation results are given in Section V, and the conclusions are drawn in the last section.

*Notations:* All logarithms are base 2 and the information is measured in bits unless otherwise specified; $\mathbb{P}\{\cdot\}$ denotes the probability of an event; $H(\cdot)$ and $h(\cdot)$ denote the discrete and differential entropy functions, respectively; $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$; $\mathcal{U}(\mathcal{G})$ denotes the uniform distribution on the set $\mathcal{G}$; $|\mathcal{G}|$ denotes the cardinality of the set $\mathcal{G}$; $\mathbb{E}_{\mathbf{x}}\{\cdot\}$ denotes the expectation operator with respect to the random vector $\mathbf{x}$; $\mathbf{A}^{\mathrm{T}}$, $\mathbf{A}^{\dagger}$, $\mathrm{tr}\{\mathbf{A}\}$, and $|\mathbf{A}|$ denote the transpose, Hermitian transpose, trace, and determinant of $\mathbf{A}$, respectively; matrices $\mathbf{A} \succeq \mathbf{B}$ denotes that $\mathbf{A} - \mathbf{B}$ is positive semidefinite; $\otimes$ denotes the Kronecker product; $\mathbf{E}_{i,j}^{N}$ denotes the $N \times N$ matrix with all zeros except a 1 on the $i$th row and $j$th column; $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ denotes an identity matrix, where the subscript $n$ is omitted if clear in the context; $[x]^+$ denotes $\max\{x, 0\}$ for $x \in \mathbb{R}$; and $\lfloor x \rfloor$ and $\lceil x \rceil$ denote the largest integer not greater than $x$ and the smallest integer not less than $x$, respectively.

## II. Preliminaries

In this section, we present the problem statement, introduce the notions of intrinsic digits and uncertainty sets, and describe an enciphering and deciphering process.

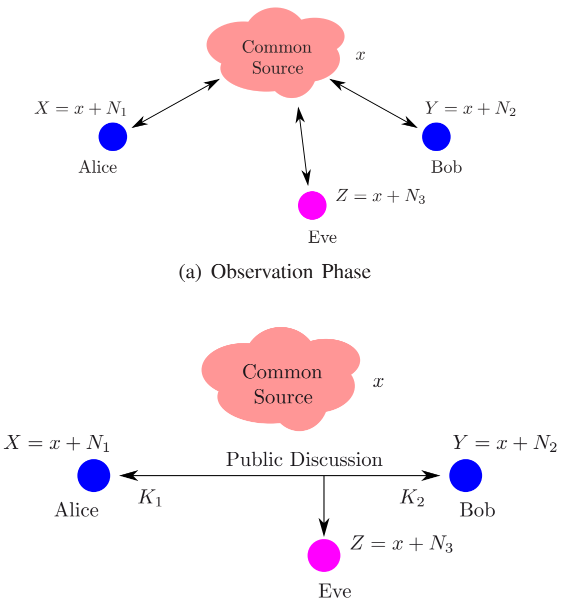
### A. Problem Statement

Consider a scenario where two legitimate nodes 1 and 2 (Alice and Bob) and an eavesdropping node 3 (Eve) have noisy observations of the common source with parameter $x \in \mathbb{R}$, given by

$$X = x + N_1\,, \quad Y = x + N_2\,, \quad Z = x + N_3 \tag{1}$$

where $N_k$'s are additive Gaussian noises. Alice and Bob aim to generate a key that is secret to Eve using their respective observations together with public discussion (see Fig. 1).

In contrast to the conventional setting where $x$ is a RV with a known distribution, we consider $x$ to be a UDP as in classic detection and estimation theory. The goal of this work is to characterize the achievable length of the secret key that Alice and Bob can generate from their respective observations $X$ and $Y$ of the parameter $x$.[2]



(a) Observation Phase

(b) Generation Phase

Fig. 1. Secret-key generation: (a) Alice, Bob, and Eve have noisy observations of a common source; (b) Alice and Bob generate a key that is secret to Eve using their respective observations together with public discussion.

In this paper, we mainly focus on the case in which Eve has access to the public channel but no observation of the common source.[3] We will give a brief discussion on the general case in which Eve also has an observation of the source in Section III-D.

### B. Intrinsic Digits and Uncertainty Sets

We next introduce the notions of the intrinsic digits and the uncertainty set of a quantized real-valued parameter.

For a nonzero parameter $x \in \mathbb{R}_{\neq 0}$, let $s_x$ be the sign of $x$,

$$l_x := \begin{cases} \lfloor \log_M |x| \rfloor\,, & x > 0 \\ \lceil \log_M |x| \rceil - 1\,, & x < 0 \end{cases}$$

and

$$x_{(i)} := \lfloor M^i \,\breve{x} \rfloor - M \lfloor M^{i-1} \,\breve{x} \rfloor\,, \qquad i \in \mathbb{Z}$$

where $\breve{x} := x + (M+1)\, M^{l_x} \cdot \mathbb{1}_{\{s_x = -1\}}$.[4] Then, the $M$-ary representation of $x$ quantized to the $m$th digit after the decimal point ($m \in \mathbb{Z}$) can be written as

$$\lfloor x \rfloor_m = -(M+1)\, M^{l_x} \cdot \mathbb{1}_{\{s_x = -1\}} + \sum_{i=-l_x}^{m} x_{(i)} \cdot M^{-i} \tag{2}$$

[2] Analogous to the rate of the secret key generated from a random source using sufficiently many independent RVs [9], the length of the secret key generated from a UDS is achievable in the information-theoretic sense of using sufficiently many different UDPs.

[3] For instance, the wireless channel between Alice and Bob serves as a common source, which cannot be observed by Eve at a position away from Alice or Bob.

[4] The term $(M+1)\, M^{l_x}$ translates $x$ to the interval $[M^{l_x}, M^{l_x+1})$ when $x < 0$.

where $s_x \in \{-1, 1\}$, the most significant bit (MSB) $x_{(-l_x)} \in \mathcal{M} \setminus \{0\}$, and $x_{(i)} \in \mathcal{M}$ for $-l_x + 1 \le i \le m$, in which $\mathcal{M} := \{0, 1, \ldots, M-1\}$. Note that the digits to the left of MSB $x_{(i)} = 0$ for $i \le -l_x - 1$ are dummy digits.

*Definition 1 (Intrinsic Digits):* The *intrinsic digits* of $\lfloor x \rfloor_m$ with $M$-ary representation are defined to be $\{s_x, x_{(i)} : -l_x \le i \le m\}$ given in (2).

*Definition 2 (Uncertainty Set):* The *uncertainty set* of $\lfloor x \rfloor_m$ is defined to be the set generated by its intrinsic digits over all possible values, given by

$$\mathcal{S}_M(\lfloor x \rfloor_m) := \Big\{ -(M+1)\, M^{l_x} \mathbb{1}_{\{\tilde{s}_x = -1\}} + \sum_{i=-l_x}^{m} \tilde{x}_{(i)} \cdot M^{-i} : \tilde{s}_x \in \{-1, 1\}, \tilde{x}_{(-l_x)} \in \mathcal{M} \setminus \{0\}, \tilde{x}_{(i)} \in \mathcal{M}, -l_x + 1 \le i \le m \Big\}.$$

*Remark 1:* Note that the parameters $x$ with the same $l_x$ generate the same uncertainty set $\mathcal{S}_M(\lfloor x \rfloor_m)$. Moreover, when partial knowledge of $x$ is available such that $x$ is nonnegative, the sign $s_x \equiv 1$ is no longer an intrinsic digit and $\tilde{s}_x$ in $\mathcal{S}_M(\lfloor x \rfloor_m)$ is always equal to one, thus reducing the cardinality of the uncertainty set by half (see Appendix A for details).

We now give the following perspective for UDPs: since the quantized parameter $\lfloor x \rfloor_m$ is unknown before observation, all parties can at best assume $\lfloor x \rfloor_m$ to be one of the elements in its uncertainty set without favoring any particular one. Even after the observation, Eve still does not have access to the intrinsic digits of $\lfloor x \rfloor_m$. Thus, the uncertainty set $\mathcal{S}_M(\lfloor x \rfloor_m)$ form a basis for generating a secret key from the common UDS between Alice and Bob. However, the dummy digits to the left of the MSB of $x$ are not safe for secret-key generation since they are zeros by default.[5] The notion of secrecy will build upon this perspective.

### C. Enciphering and Deciphering Process

We now introduce an enciphering and deciphering process that uses the intrinsic digits of the quantized parameter $\lfloor x \rfloor_m$, based on which we give the notion of the secret-key length. To facilitate the discussion, we start with the case in which Alice and Bob's observations are $X = x$ and $Y = x + N$, respectively. In Section III-D, we will extend the results to the case of (1).

Let $W$ be a confidential message following a uniform distribution on a finite set $\mathcal{W}$. The enciphering process is a deterministic mapping $e : \mathcal{W} \times \mathcal{S}_M(\lfloor x \rfloor_m) \to \mathcal{W}'$ that transforms the plain message $W$ to the cryptogram $e(W, \lfloor x \rfloor_m)$. The deciphering process is a deterministic mapping $d : \mathcal{W}' \times \mathcal{S}_M(\lfloor x \rfloor_m) \to \mathcal{W}$ that transforms the cryptogram back to the plain message. Note that the enciphering process is reversible so that the deciphering is unique if $\lfloor x \rfloor_m$ is known, i.e., $d(e(w, \lfloor x \rfloor_m), \lfloor x \rfloor_m) = w$ for any $(w, \lfloor x \rfloor_m) \in \mathcal{W} \times \mathcal{S}_M(\lfloor x \rfloor_m)$. We consider the following secure communication scheme:

- Alice enciphers the message $W$ using $\lfloor x \rfloor_m$ to the cryptogram $e(W, \lfloor x \rfloor_m)$ and broadcasts the cryptogram on the public channel;
- Bob deciphers the message $W$ from the cryptogram using his noisy observation
$$Y_m := \lfloor x \rfloor_m + \lfloor N \rfloor_m$$
where $\lfloor N \rfloor_m$ denotes the quantized noise.

While Eve has access to the cryptogram $e(W, \lfloor x \rfloor_m)$ from the public channel, she cannot favor $\lfloor x \rfloor_m$ to be any particular element in $\mathcal{S}_M(\lfloor x \rfloor_m)$. This implies that secrecy of $W$ with respect to Eve can be guaranteed by proper enciphering if $|\mathcal{W}| \le |\mathcal{S}_M(\lfloor x \rfloor_m)|$.[6] Thus, we consider $|\mathcal{W}| = |\mathcal{S}_M(\lfloor x \rfloor_m)|$ and $\mathcal{W}' = \mathcal{W}$ to achieve the largest possible message set with secrecy.

Meanwhile, Bob deciphers the cryptogram using $Y_m$, from which he can obtain the likelihood of $\lfloor x \rfloor_m$. Such deciphering creates a conceptual secure channel between Alice and Bob. The secret-key length $\mathcal{L}_{\mathrm{S}}(\lfloor x \rfloor_m; Y_m)$, generated from the common UDS between Alice and Bob, is defined as the rate of this conceptual channel, i.e.,

$$\begin{aligned} \mathcal{L}_{\mathrm{S}}(\lfloor x \rfloor_m; Y_m) &:= I\big(W; Y_m, e(W, \lfloor x \rfloor_m)\big) \\ &= H(W) - H\big(W | Y_m, e(W, \lfloor x \rfloor_m)\big) \end{aligned} \tag{3}$$

where the UDP $\lfloor x \rfloor_m \in \mathcal{S}_M(\lfloor x \rfloor_m)$.

While finer quantization of $x$ gives more intrinsic digits, the lower digits of $Y_m$ are less reliable for representing those of $\lfloor x \rfloor_m$ due to the observation noise. We next characterize the secret-key length for the real-valued parameter $x$ by taking the limit of $\mathcal{L}_{\mathrm{S}}(\lfloor x \rfloor_m; Y_m)$ as $m \to \infty$.

## III. Intrinsic Information

In this section, we first derive the secret-key length for the standard case with binary representation and no scaling.[7] We then introduce the notion of intrinsic information and determine such information for the general case with $M$-ary representation and scaling.

### A. Standard Case

We first consider the standard case where binary representation ($M = 2$) is used and no scaling is applied before secret-key generation. If the parameter $x \in \mathcal{G}_{\mathrm{S}} := [-2, -1) \cup [1, 2)$, then the quantized parameter $\lfloor x \rfloor_m$ by (2) belongs to $\mathcal{G}_{\mathrm{S},m} := \{k/2^m : k \in \mathcal{K}_m\}$, where

$$\mathcal{K}_m := \{2^m + k', -2^{m+1} + k' : 0 \le k' \le 2^m - 1, k' \in \mathbb{Z}\}.$$

Note that the uncertainty set $\mathcal{S}_2(\lfloor x \rfloor_m) = \mathcal{G}_{\mathrm{S},m}$ for any $x \in \mathcal{G}_{\mathrm{S}}$, and the cardinality of the sets $|\mathcal{G}_{\mathrm{S},m}| = |\mathcal{K}_m| = 2^{m+1}$. The quantized noise is given by $\lfloor N \rfloor_m \in \{k/2^m : k \in \mathbb{Z}\}$ with distribution

$$P_{k,m} := \mathbb{P}\Big\{ \lfloor N \rfloor_m = \frac{k}{2^m} \Big\} = \mathbb{P}\Big\{ \frac{k}{2^m} \le N < \frac{k+1}{2^m} \Big\}.$$

[5] More specifically, if $\mathcal{X} \subseteq \mathbb{R}$ is the set of parameters for which a secret-key generation protocol uses some dummy digits, then Eve would have partial knowledge of the secret key generated from those $x \in \mathcal{X}$ by assuming the first a few digits of $x$ to be zeros.

[6] For example, one-time pad employing the intrinsic digits of $\lfloor x \rfloor_m$ can be used for the enciphering and deciphering process.

[7] Note that Alice and Bob can scale their observations together by a known constant $t \in \mathbb{R}_{\neq 0}$ before generating a secret key. Such scaling operations may increase or decrease the secret-key length as will be shown in Section III-B.

We next derive the secret-key length generated from the quantized parameter $\lfloor x \rceil_m$ and the observation $Y_m$ for the standard case.

*Proposition 1:* For $\lfloor x \rceil_m = k_0/2^m \in \mathcal{G}_{\mathrm{S},m}$ where $k_0 \in \mathcal{K}_m$, the secret-key length generated from $\lfloor x \rceil_m$ and $Y_m$ is

$$\mathcal{L}_{\mathrm{S}}(\lfloor x \rceil_m; Y_m) = H(U_m) - \sum_{k_1 \in \mathbb{Z}} P_{k_1 - k_0, m} \cdot H\Big(U_m \Big| U_m + \lfloor N \rceil_m = \frac{k_1}{2^m}\Big) \tag{4}$$

where the auxiliary RV $U_m \sim \mathcal{U}(\mathcal{G}_{\mathrm{S},m})$.

*Proof:* See Appendix B. □

To facilitate further analysis, we now introduce a new information measure called *intrinsic information* between $\lfloor x \rceil_m$ and $Y_m$, defined as the average of the secret-key length $\mathcal{L}_{\mathrm{S}}(\lfloor x \rceil_m; Y_m)$ over the uncertainty set of $\lfloor x \rceil_m$, i.e.,[8]

$$\mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m) := \frac{1}{|\mathcal{S}_2(\lfloor x \rceil_m)|} \sum_{\lfloor \tilde{x} \rceil_m \in \mathcal{S}_2(\lfloor x \rceil_m)} \mathcal{L}_{\mathrm{S}}(\lfloor \tilde{x} \rceil_m; \tilde{Y}_m) \tag{5}$$

where $\tilde{Y}_m = \lfloor \tilde{x} \rceil_m + \lfloor N \rceil_m$. The intrinsic information has good information-theoretic properties as shown in the next proposition, and hence, we will use intrinsic information to characterize the secret-key length in this paper.[9]

*Proposition 2:* For a parameter $x \in \mathcal{G}_{\mathrm{S}}$, the intrinsic information between $\lfloor x \rceil_m$ and $Y_m$ for the standard case is

$$\mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m) = I(U_m; U_m + \lfloor N \rceil_m)\,. \tag{6}$$

*Proof:* See Appendix C. □

*Remark 2:* Proposition 2 shows that the intrinsic information between $\lfloor x \rceil_m$ and $Y_m$ is equal to the mutual information between $U_m$ and $U_m + \lfloor N \rceil_m$. This implies that the unknown parameter $\lfloor x \rceil_m$ plays the role of the RV $U_m$ in the derivation of the secret-key rate in [9]. Note that $\lfloor x \rceil_m$ is a *single* UDP, and in practice one need to utilize different UDPs to achieve the secret-key length.

The next proposition gives some functional properties of the intrinsic information $\mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m)$ with respect to the quantization level $m$.

*Proposition 3:* For a parameter $x \in \mathcal{G}_{\mathrm{S}}$, the intrinsic information (6) has the following properties:

(i) $0 \le \mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m) < m + 1$;

(ii) $\mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m) \le \mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_{m+1}; Y_{m+1})$.

*Proof:* See Appendix D. □

*Remark 3:* The proposition shows that the intrinsic information $\mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m)$ is positive, finite, and nondecreasing with the quantization level $m$. In particular, each new quantization bit of the parameter, although corrupted by the observation noise, increases the intrinsic information.

The monotonicity property of the intrinsic information implies that the limit of $\mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m)$ exists when $m \to \infty$, defined as

$$\mathcal{I}_{\mathrm{S}}^{(2)}(x; Y) := \lim_{m \to \infty} \mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m)$$

which is given in the next theorem.

*Theorem 1:* For a parameter $x \in \mathcal{G}_{\mathrm{S}}$, the intrinsic information between $x$ and $Y$ for the standard case is

$$\mathcal{I}_{\mathrm{S}}^{(2)}(x; Y) = I(U_{\mathrm{S}}; U_{\mathrm{S}} + N) \tag{7}$$

where the RV $U_{\mathrm{S}} \sim \mathcal{U}(\mathcal{G}_{\mathrm{S}})$.

*Proof:* (Outline) First, the limit of $\mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m)$ exists by Proposition 3 and it can be shown that $\mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m) \le I(U_{\mathrm{S}}; U_{\mathrm{S}} + N)$. Since $\lfloor x \rceil_m \to x$, we can construct a sequence of auxiliary RVs $U_m \to U_{\mathrm{S}}$ a.s. Moreover, $N_m \to N$ a.s. Hence, the limit converges to $I(U_{\mathrm{S}}; U_{\mathrm{S}} + N)$ by the dominated convergence theorem. □

*Remark 4:* Note that the limit in (7) can be infinite if the probability distribution of $N$ does not satisfy some regularity condition, e.g., $h(N) = -\infty$ or the distribution contains discrete components.

We now extend the intrinsic information (7) for $x \in \mathcal{G}_{\mathrm{S}}$ to general nonzero parameter $x \in \mathbb{R}_{\neq 0}$. Note that the case of nonzero parameters can be converted to the standard case by shifting the position of the decimal point of both $x$ and $N$ (and hence $Y$) to the left by $l_x$ digits, since such an operation does not affect the intrinsic digits. By a similar derivation of Proposition 2 and Theorem 1, we can derive the intrinsic information for nonzero parameters.

*Corollary 1:* For a parameter $x \in \mathbb{R}_{\neq 0}$, the intrinsic information between $x$ and $Y$ for the standard case is

$$\mathcal{I}_{\mathrm{S}}^{(2)}(x; Y) = I(U_{\mathrm{S}}; U_{\mathrm{S}} + 2^{-l_x} \cdot N)\,.$$

Moreover, $\mathcal{I}_{\mathrm{S}}^{(2)}(2^l \cdot x; 2^l \cdot Y) = \mathcal{I}_{\mathrm{S}}^{(2)}(x; Y)$ for any $l \in \mathbb{Z}$.

### B. General Case

Since Alice and Bob can scale their observations by any nonzero constant before generating a secret key, we now consider the general case with $M$-ary representation and (nonzero) scaling. First, the intrinsic information for the general case with binary representation is defined as the average of those for the standard case over all nonzero scalings, i.e.,

$$\mathcal{I}^{(2)}(x; Y) := \liminf_{T \to \infty} \frac{1}{2T} \int_{[-T,0) \cup (0,T]} \mathcal{I}_{\mathrm{S}}^{(2)}(tx; tY)\, dt\,. \tag{8}$$

Note that for a known constant $t \in \mathbb{R}_{\neq 0}$, the scaled parameter $tx$ is unknown but deterministic. Hence we apply the same analysis described in the previous subsection for the unknown parameter $tx$. Let the set $\mathcal{G}^{(M)} := [-1, -1/M\,] \cup [\,1/M, 1]$ for $M \in \{2, 3, \ldots\}$ and the RV $U^{(M)} \sim \mathcal{U}(\mathcal{G}^{(M)})$.

*Theorem 2:* For a parameter $x \in \mathbb{R}$, the intrinsic information between $x$ and $Y$ for the general case with binary representation is

$$\mathcal{I}^{(2)}(x; Y) = 2 \int_{\frac{1}{2}}^{1} I(U_x^{(2)}; U_x^{(2)} + tN)\, dt \tag{9}$$

where the RV $U_x^{(2)} = x \cdot U^{(2)}$.

*Proof:* Note that the intrinsic information $\mathcal{I}_{\mathrm{S}}^{(2)}(tx; tY)$ is a periodic function of $\log t$ by Corollary 1. Within each

[8] We use $\mathcal{I}_{\mathrm{S}}^{(2)}(\cdot, \cdot)$ to denote the intrinsic information for the standard case, where the superscript 2 denotes binary representation.

[9] Numerical results showed that the secret-key length as a function of $\lfloor x \rceil_m$ is relatively flat over $\lfloor x \rceil_m \in \mathcal{G}_{\mathrm{S},m}$. Thus the intrinsic information serves as a good approximation for the secret-key length.

period $t \in [\,2^n/|x|, 2^{n+1}/|x|)$ for $n \in \mathbb{Z}$, it can be shown that $\mathcal{I}_{\mathrm{S}}^{(2)}(tx; tY)$ is a decreasing function of $t$. Hence, the $\liminf$ of the average in (8) is equal to the average in one period, given by

$$\mathcal{I}^{(2)}(x;Y) = \frac{|x|}{2^{n+1}-2^n}\int_{2^n/|x|}^{2^{n+1}/|x|} \mathcal{I}_{\mathrm{S}}^{(2)}(tx;tY)\, dt$$

which leads to (9) after some algebra. □

*Remark 5:* The intrinsic information $\mathcal{I}^{(2)}(x;Y)$ is also applicable to the case $x = 0$, i.e.,

$$\mathcal{I}^{(2)}(x=0; Y=N) = \lim_{x\to 0} \mathcal{I}^{(2)}(x;Y) = 0$$

since the distribution of $U_x^{(2)}$ converges to the step function as $x \to 0$. This result agrees with the intuition since Eve can generate statistically the same observation as $Y = N$.

We now extend the results for the binary representation to $M$-ary representation of the parameters. Due to space constraints, we will omit some of the proofs for brevity in the following.

*Theorem 3:* For a parameter $x \in \mathbb{R}$, the intrinsic information between $x$ and $Y$ for the general case with $M$-ary representation is

$$\mathcal{I}^{(M)}(x;Y) = \frac{1}{1-M^{-1}}\int_{M^{-1}}^{1} I(U_x^{(M)}; U_x^{(M)} + tN)\, dt \quad (10)$$

where the RV $U_x^{(M)} = x \cdot U^{(M)}$.

*Proof:* The proof follows from a similar derivation as that for the binary representation case. □

*Proposition 4:* The intrinsic information $\mathcal{I}^{(M)}(x;Y)$ is a non-decreasing function with $M$, i.e., $\mathcal{I}^{(M_1)}(x;Y) \le \mathcal{I}^{(M_2)}(x;Y)$ for $M_1 \le M_2 \in \{2,3,\ldots\}$.

*Remark 6:* Recall that the $M$-ary representation of a parameter given in Definition 1, the MSB contains $\log(M-1)$ bits while each of the rest contains $\log M$ bits. This proposition can be interpreted as that larger $M$ gives a better resolution of each digit.

Since the intrinsic information $\mathcal{I}^{(M)}(x;Y)$ is non-decreasing with $M$, the limit exists as $M \to \infty$ and we define such a limit as the general intrinsic information, i.e.,

$$\mathcal{I}(x;Y) := \lim_{M\to\infty} \mathcal{I}^{(M)}(x;Y)$$

which is derived in the next theorem.

*Theorem 4:* For a parameter $x \in \mathbb{R}$, the intrinsic information between $x$ and $Y$ for the general case with $\infty$-ary representation is[10]

$$\mathcal{I}(x;Y) = \int_0^1 I(U_x; U_x + tN)\, dt \quad (11)$$

where the RV $U_x \sim \mathcal{U}([\,-|x|, |x|\,])$.

*Proof:* We can show that for any $t \in [\,0,\,1\,]$,

$$I(U_x^{(M)}; U_x^{(M)} + tN) \le I(U_x; U_x + N) - \log t$$

and thus by Theorem 3 we have

$$\mathcal{I}^{(M)}(x;Y) \le I(U_x; U_x + N) + \log e\,.$$

Moreover, since $U_x^{(M)} \to U_x$ in distribution and $M^{-1} \to 0$ as $M \to \infty$, the limit in (11) can be obtained by the dominated convergence theorem. □

In the rest of the paper, we will refer the notion of intrinsic information to (11) if not otherwise specified.

### C. Properties of Intrinsic Information

We next present some properties of the intrinsic information (11). The following propositions derive upper and lower bounds for the intrinsic information as well as the information loss due to finite-base representations.

*Proposition 5:* The intrinsic information $\mathcal{I}(x;Y)$ is bounded above and below as

$$\max\bigl\{I(U_x;U_x+N),\, h(U_x) - h(N) + \log e\bigr\} \le \mathcal{I}(x;Y) \le I(U_x;U_x+N) + \log e\,. \quad (12)$$

Moreover, the upper and lower bounds coincide when $|x| \to \infty$.

*Proof:* Note that for any $t \in [\,0,\,1\,]$, we have

$$I(U_x; U_x + tN) \ge I(U_x; U_x + N)$$

and

$$\begin{aligned} I(U_x; U_x + tN) &\ge h(U_x) - h(tN) \\ &= h(U_x) - h(N) - \log t \end{aligned}$$

which lead to the lower bound by taking the integration over $t$ on both sides. On the other hand, the upper bound is due to the fact that for any $t \in [\,0,\,1\,]$,

$$\begin{aligned} I(U_x; U_x + tN) &\le h(U_x + N) - h(tN) \\ &= I(U_x; U_x + N) - \log t\,. \end{aligned}$$

Moreover, the gap between the upper and lower bounds is less than or equal to

$$\bigl[\,I(U_x;U_x+N) + \log e\,\bigr] - \bigl[\,h(U_x) - h(N) + \log e\,\bigr] = I(N; U_x + N)$$

which goes to zero as $|x| \to \infty$. □

*Remark 7:* The upper and lower bounds for the intrinsic information are more tractable since they do not involve the integral over $t \in (0, 1\,]$ as in (11). Moreover, the gap between the upper and lower bounds is shown to be at most $\log e \approx 1.4$ bits, and such a gap diminishes when $|x| \to \infty$. Hence, these bounds can be used for further analysis, especially in high SNR scenarios where $|x|^2/\sigma_N^2 \gg 1$ in which $\sigma_N^2$ is the variance of $N$.

*Proposition 6:* The loss of intrinsic information due to $M$-ary representation is

$$0 \le \mathcal{I}(x;Y) - \mathcal{I}^{(M)}(x;Y) \le \frac{1}{M-1}\log\frac{M^M}{(M-1)^{M-1}}$$

and the upper bound can be achieved asymptotically when $|x| \to \infty$, i.e.,

$$\lim_{|x|\to\infty}\bigl[\,\mathcal{I}(x;Y) - \mathcal{I}^{(M)}(x;Y)\,\bigr] = \frac{1}{M-1}\log\frac{M^M}{(M-1)^{M-1}}\,.$$

In particular, the maximum loss is 2 bits, obtained when $M = 2$ and $|x| \to \infty$.

[10] The superscript $\infty$ is omitted for brevity.

*Remark 8:* Proposition 6 provides the upper bound for the information loss due to $M$-ary representation of the parameters. For example, the maximum loss is 2, $0.52$, and $0.36$ bits for binary, decimal, and hexadecimal representations, respectively. Moreover, the loss of intrinsic information increases with the parameter value $|x|$, in which case the intrinsic information itself also increases. Numerical examples will be given in Section V to quantify the information loss due to finite-base representations.

### *D. Extensions and Discussion*

In this section, we derive the intrinsic information between two noisy observations of the same unknown parameter, and then extend the results to vector parameter cases. We will also provide a brief discussion on the case in which Eve also has an observation of the parameter and a remark on the secret-key generation using deterministic and random sources.

*1) Two Noisy Observations:* We next present the intrinsic information between two noisy observations, following an analogous derivation as those in Section III.

*Theorem 5:* For a parameter $x \in \mathbb{R}$, the intrinsic information between $X = x + N_1$ and $Y = x + N_2$ is

$$\mathcal{I}(X;Y) = \int_0^1 I(U_x + tN_1; U_x + tN_2)\, dt\,. \tag{13}$$

*Remark 9:* The intrinsic information between two observations of a UDP reduces to (11) in Theorem 2 if $N_1 = 0$.

*2) Vector Parameters:* Let $\mathbf{T} := \text{diag}\{t_1, t_2, \ldots, t_d\}$, and define the integral

$$\int_{\mathbf{0}}^{\mathbf{I}_d} g(\cdot)\, d\mathbf{T} := \int_{t_1=0}^{1} \cdots \int_{t_d=0}^{1} g(\cdot)\, dt_1 dt_2 \cdots dt_d\,.$$

The intrinsic information for vector parameter cases is given in the next theorem.

*Theorem 6:* For a parameter vector $\mathbf{x} \in \mathbb{R}^d$, the intrinsic information between $\mathbf{X} = \mathbf{x} + \mathbf{N}_1$ and $\mathbf{Y} = \mathbf{x} + \mathbf{N}_2$ is

$$\mathcal{I}(\mathbf{X};\mathbf{Y}) = \int_{\mathbf{0}}^{\mathbf{I}_d} I(\mathbf{U}_\mathbf{x} + \mathbf{T}\mathbf{N}_1; \mathbf{U}_\mathbf{x} + \mathbf{T}\mathbf{N}_2)\, d\mathbf{T} \tag{14}$$

where $\mathbf{U}_\mathbf{x} = [\, U_{x_1} \;\; U_{x_2} \;\; \cdots \;\; U_{x_d}\,]^\dagger$, in which the RV $U_{x_i} \sim \mathcal{U}([\,-|x_i|, |x_i|\,])$.

*Proposition 7:* The intrinsic information $\mathcal{I}(\mathbf{X};\mathbf{Y})$ is bounded below as

$$\begin{aligned}\mathcal{I}(\mathbf{X};\mathbf{Y}) &\geq d\,(1 + \log e) + \sum_{i=1}^{d} \log|x_i| - h(\mathbf{N}_1 - \mathbf{N}_2)\\ &=: \mathcal{I}_\text{H}(\mathbf{X};\mathbf{Y})\,.\end{aligned} \tag{15}$$

Moreover, the lower bound can be achieved asymptotically when $|x_i| \to \infty$ for all $i = 1, 2, \ldots, d$.

*Proof:* See Appendix E. □

*Remark 10:* The proposition gives a lower bound for the intrinsic information and shows that the bound is asymptotically tight in high SNR regimes where $|x_i|^2/\sigma_N^2$ is large. Since the lower bound does not involve integral over $\mathbf{T}$, it is more appealing than the exact expression (14) and we will use it for analysis in high SNR scenarios in the next section.

*3) Observation at Eve:* We next briefly discuss the case in which Eve also has an observation of the unknown parameter. By a similar derivation leading to Theorem 4 and the results in [9], one can show that the secret-key length between Alice and Bob, conditioned on Eve's observation, is bounded between $\max\big\{\mathcal{I}(X;Y) - \mathcal{I}(X;Z),\, \mathcal{I}(X;Y) - \mathcal{I}(Y;Z)\big\}$ and $\min\big\{\mathcal{I}(X;Y),\, \mathcal{I}(X;Y|Z)\big\}$, where

$$\mathcal{I}(X;Y|Z) = \int_0^1 I(U_x + tN_1; U_x + tN_2 | U_x + tN_3)\, dt\,.$$

Moreover, for some special cases [9], the upper and lower bounds coincide.

*4) Deterministic v.s. random sources:* We close this section with a remark on the difference between secret-key generation using UDSs developed in this work and that using random sources studied in literature. The main difference is whether the source is considered deterministic or random. In the former case, the unknown parameter $x$ is nonrandom (i.e., it does not follow any distribution), and the secrecy is derived from the ambiguity of $x$ according to its uncertainty set; the secret-key length is equal to the intrinsic information between $x$ and $Y = x + N$, which depends on the particular value of $x$. However, in the latter case, the unknown parameter $X$ is random (i.e., it follows a presumed distribution), and the secrecy is derived from the randomness of $X$ according to its distribution; the secret-key length is equal to the mutual information between $X$ and $Y = X + N$, which depends on the distribution of $X$.

## IV. Case Study: Wideband Channels in Mobile Wireless Networks

In this section, we first introduce the models for time-varying wideband channels in mobile wireless networks and then derive the intrinsic information of such wideband channels.

### *A. Wideband Channel Models*

Let $\mathcal{T} := \{1, 2, \ldots, N\}$ be the set of time indexes. The received wideband waveform by multipath propagation at time $n \in \mathcal{T}$ can be modeled as [39]

$$r^{(n)}(t) = \sum_{l=1}^{L^{(n)}} \alpha_l^{(n)} \cdot \sqrt{P}\, s\big(t - \tau_l^{(n)}\big) + z^{(n)}(t)\,, \quad t \in [\,0\,, T_\text{ob})\tag{16}$$

where $s(t)$ is a normalized known wideband waveform, $P$ is the transmit power of the signal measured at 1 m away from the transmitter, $L^{(n)}$ is the number of multipath components (MPCs), $\alpha_l^{(n)} \in \mathbb{R}$ and $\tau_l^{(n)} \in \mathbb{R}_{\geq 0}$ are the amplitude and delay of the $l$th path, respectively, $z^{(n)}(t)$ is the observation noise modeled as a real-valued white Gaussian processes with spectral density $N_0/2$, and $[\,0\,, T_\text{ob})$ is the observation interval. The effective bandwidth of the signal is defined as $\beta := \big(\int_{-\infty}^{\infty} f^2 S(f)^2\, df\big)^{1/2}$, where $S(f)$ is the Fourier transform of $s(t)$.

In the following, we consider two simple channel models for low and high mobility scenarios, aiming to provide insights into the intrinsic information of time-varying wideband channels.[11] Our methods and analysis for the intrinsic information in the case study are applicable to general scenarios.

*Low Mobility:* Consider low mobility scenarios in which the path amplitudes and delays experience slow variations over time. In particular, there is no removal or addition of MPCs in the channels at different time instants, i.e., $L^{(n)} = L$ for all $n \in \mathcal{T}$, and the variations of the path amplitudes and delays are modeled as

$$\alpha_l^{(n+1)} = \alpha_l^{(n)} + \delta_\alpha^{(n,l)}(v^{(n)}) \tag{17}$$

$$\tau_l^{(n+1)} = \tau_l^{(n)} + \delta_\tau^{(n,l)}(v^{(n)}) \tag{18}$$

where $v^{(n)}$ is the relative speed of the two nodes at time $n$, and $\delta_\alpha^{(n,l)}(\cdot)$ and $\delta_\tau^{(n,l)}(\cdot)$ are functions of the node speed.[12] For instance, based on ray-tracing multipath propagation, $\delta_\tau^{(n,l)}(v^{(n)}) \in [-v^{(n)}/c\,, +v^{(n)}/c]$, where $c$ is the speed of propagation.

*High Mobility:* Consider high mobility scenarios in which the path amplitudes and delays experience fast variations over time. In particular, the channel impulse responses of different measurements are completely independent, i.e., the number of MPCs, path amplitudes, and path delays have no relationship among channels at different time instants. For simplicity, we assume that the number of MPCs for different time instants are equal, i.e., $L^{(n)} = L$ for all $n \in \mathcal{T}$.

### B. Intrinsic Parameters

The unknown parameters of the wideband channels during time 1 to $N$ include all the multipath parameters of the waveforms, given by

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}^{(1)\dagger} & \boldsymbol{\theta}^{(2)\dagger} & \cdots & \boldsymbol{\theta}^{(N)\dagger} \end{bmatrix}^\dagger \tag{19}$$

where $\boldsymbol{\theta}^{(n)} = \begin{bmatrix} \boldsymbol{\alpha}^{(n)\dagger} & \boldsymbol{\tau}^{(n)\dagger} \end{bmatrix}^\dagger$ includes the channel parameters at time $n$, in which $\boldsymbol{\alpha}^{(n)} = [\,\alpha_1^{(n)} \;\; \alpha_2^{(n)} \;\; \cdots \;\; \alpha_L^{(n)}\,]^\dagger$ and $\boldsymbol{\tau}^{(n)} = [\,\tau_1^{(n)} \;\; \tau_2^{(n)} \;\; \cdots \;\; \tau_L^{(n)}\,]^\dagger$.

These multipath parameters fully characterize the time-varying channels from time 1 to $N$, and they can serve as a common UDS between two nodes according to the channel reciprocity principle. Note that since the multipath parameters at different time instants follow the mobility models, not all the parameters in $\boldsymbol{\theta}$ are intrinsic. We next give the intrinsic parameters based on the channel models in low and high mobility scenarios.

*Low Mobility:* Without loss of generality, we assume $\tau_1^{(1)} < \tau_2^{(1)} < \cdots < \tau_L^{(1)}$ for time 1. Since the path delay $\tau_l^{(1)}$ increases with $l$, the first arrival time and inter-arrival times are intrinsic for the channel. Moreover, based on the multipath parameter variation models (17) and (18), the variations of the path amplitudes and delays are intrinsic. Therefore, the intrinsic parameter set is

$$\begin{aligned}&\big\{\alpha_l^{(1)}, \tau_l^{(1)} - \tau_{l-1}^{(1)} : l \in \mathcal{L}\big\}\\ &\quad \cup \big\{\alpha_l^{(n+1)} - \alpha_l^{(n)}, \, \tau_l^{(n+1)} - \tau_l^{(n)} : l \in \mathcal{L}, \, n \in \mathcal{T}\backslash\{N\}\big\}\end{aligned}$$

where $\mathcal{L} := \{1, 2, \ldots, L\}$ and $\tau_0^{(n)} := 0$ for notational convenience. We can then represent the intrinsic parameters by the vector $\boldsymbol{\eta} = \mathbf{H}_1\,\boldsymbol{\theta}$, where $\mathbf{H}_1$ is a $2NL \times 2NL$ matrix given by

$$\mathbf{H}_1 = \mathbf{I}_{2NL} - \sum_{n=3}^{2N} \mathbf{E}_{n,n-2}^{2N} \otimes \mathbf{I}_L - \mathbf{E}_{2,2}^{2N} \otimes \mathbf{D}$$

in which $\mathbf{D}$ is a $L \times L$ matrix given by $\mathbf{D} = \sum_{l=2}^{L} \mathbf{E}_{l,l-1}^{L}$.

*High Mobility:* Without loss of generality, we assumed $\tau_1^{(n)} < \tau_2^{(n)} < \cdots < \tau_L^{(n)}$ for all $n \in \mathcal{T}$. Since the multipath parameters are uncorrelated among different measurements, the intrinsic parameters can be written as the vector $\boldsymbol{\eta} = \mathbf{H}_2\,\boldsymbol{\theta}$, where $\mathbf{H}_2$ is a $2NL \times 2NL$ matrix given by

$$\mathbf{H}_2 = \mathbf{I}_{2NL} - \sum_{n=1}^{N} \mathbf{E}_{2n,2n}^{2N} \otimes \mathbf{D}\,.$$

We consider that the waveform $s(t)$ is a continuous function and define functions $\phi_1(\tau) := s(t-\tau)$ and $\phi_2(\tau) := -\dot{s}(t-\tau)/2\pi\beta$. We also introduce the kernel matrix

$$\boldsymbol{\Phi}(\boldsymbol{\tau}, \boldsymbol{\tau}) := \sum_{i=1}^{L}\sum_{j=1}^{L} \mathbf{E}_{i,j}^{L} \otimes \boldsymbol{\Phi}(\tau_i, \tau_j)$$

for $\boldsymbol{\tau} = [\,\tau_1 \;\; \tau_2 \;\; \cdots \;\; \tau_L\,]^\dagger$, where

$$\boldsymbol{\Phi}(\tau_1, \tau_2) := \begin{bmatrix} \langle \phi_1(\tau_1), \phi_1(\tau_2)\rangle & \langle \phi_1(\tau_1), \phi_2(\tau_2)\rangle \\ \langle \phi_2(\tau_1), \phi_1(\tau_2)\rangle & \langle \phi_2(\tau_1), \phi_2(\tau_2)\rangle \end{bmatrix}$$

in which the inner product is given by $\langle g(\tau_1), h(\tau_2)\rangle := \int g(\tau_1)\, h^\dagger(\tau_2)\, dt$. We next define the notion of resolvable MPCs in terms of the kernel matrix.

*Definition 3 (Resolvable Multipath Channel):* The multipath channel is called *resolvable*, if $\boldsymbol{\Phi}(\tau_i, \tau_j) = \mathbf{0}$ for all pairs of $\{(i, j) : i, j \in \mathcal{L}, i \neq j\}$.

*Remark 11:* When the multipath channel with the delay vector $\boldsymbol{\tau}$ is resolvable, one can show that $\boldsymbol{\Phi}(\boldsymbol{\tau}, \boldsymbol{\tau}) = \mathbf{I}$. As an example, a multipath channel is resolvable if the supports of $s(t-\tau_i)$ and $s(t-\tau_j)$ do not overlap for all pairs $i \neq j \in \mathcal{L}$.

### C. MSE Bounds and Intrinsic Information

We next obtain the information inequality for the mean-squared error (MSE) matrix of the intrinsic parameters, and then characterize the intrinsic information of the wideband channels in high SNR regimes.

The MSE matrix of the unbiased estimate $\hat{\boldsymbol{\eta}}$ of the intrinsic parameter vector is given by [46], [47]

$$\mathbb{E}_{\mathbf{r}}\big\{(\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})(\boldsymbol{\eta} - \hat{\boldsymbol{\eta}})^\dagger\big\} \succeq \mathbf{J}_{\boldsymbol{\eta}}^{-1} \tag{20}$$

where $\mathbf{J}_{\boldsymbol{\eta}}$ is the Fisher information matrix (FIM) for the parameter $\boldsymbol{\eta}$ based on the observations $\{r^{(n)}(t)\,, n \in \mathcal{T}\}$. In high SNR regimes, the information inequality (20) is achievable asymptotically, and the error vector $\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}$ follows a joint Gaussian distribution with zero mean and covariance matrix $\mathbf{J}_{\boldsymbol{\eta}}^{-1}$, i.e., $\hat{\boldsymbol{\eta}} \sim \mathcal{N}(\boldsymbol{\eta}, \mathbf{J}_{\boldsymbol{\eta}}^{-1})$ [43].

[11] Note that the models characterize two extreme scenarios in mobile networks, i.e., low and high mobility. The channels for general scenarios are expected to behave somewhere between the two extreme scenarios. The time-varying channel models need to be verified by experimentation, which is beyond the scope of this paper.

[12] The functions for the amplitude and delay variations can account for a wide range of channel behaviors in various scenarios.

We consider that the wireless systems operate in high SNR regimes, and let $\mathbf{r}_{\mathrm{A}}^{(1:N)}$ and $\mathbf{r}_{\mathrm{B}}^{(1:N)}$ be the set of waveforms received during time 1 to $N$ at Alice and Bob, respectively. The channel between them is assumed to be reciprocal so that $\mathbf{r}_{\mathrm{A}}^{(n)}$ and $\mathbf{r}_{\mathrm{B}}^{(n)}$ are the same except for observation noises. Based on the information inequality (20), the estimates of the intrinsic parameter vector at Alice and Bob can be written, respectively, as $\hat{\boldsymbol{\eta}}_{\mathrm{A}} = \boldsymbol{\eta} + \widetilde{\boldsymbol{\eta}}_{\mathrm{A}}$ and $\hat{\boldsymbol{\eta}}_{\mathrm{B}} = \boldsymbol{\eta} + \widetilde{\boldsymbol{\eta}}_{\mathrm{B}}$, where the error vectors $\widetilde{\boldsymbol{\eta}}_{\mathrm{A}}$ and $\widetilde{\boldsymbol{\eta}}_{\mathrm{B}}$ can be modeled as independent $\mathcal{N}(\mathbf{0}, \mathbf{J}_{\boldsymbol{\eta}}^{-1})$ RVs.

By using (15), we next derive the intrinsic information of the wideband channels in high SNR regimes.

*Theorem 7:* In high SNR regimes, the intrinsic information of the wideband channels $\mathcal{I}_{\mathrm{H}}(\mathbf{r}_{\mathrm{A}}^{(1:N)}; \mathbf{r}_{\mathrm{B}}^{(1:N)})$ is given by (21) and (22) shown at the bottom of the page for the low and high mobility scenarios, respectively.

*Proof:* See Appendix F. □

*Proposition 8:* The path overlapping effect satisfies $\frac{1}{2}\log\left|\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)})\right| \leq 0$, with the equality iff the multipath channels are resolvable.

*Proof:* Sufficiency: If the multipath channel is resolvable, then $\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)}) = \mathbf{I}$ by Definition 3 and hence $|\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)})| = 1$.

Necessity: If $|\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)})| = 1$, since $\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)}) \succeq 0$ and all its diagonal elements of are 1's, then $\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)}) = \mathbf{I}$, i.e., the multipath channel is resolvable. □

*Remark 12:* The theorem gives the intrinsic information of the wideband channels for both low and high mobility scenarios. We make several remarks as follows.

First, the intrinsic information is proportional to the number of observations $N$ and MPCs $L$, since each MPC in an observation has two parameters, i.e., amplitude and delay, that provides intrinsic information.

Second, the intrinsic information increases logrithmatically with the effective bandwidth $\beta$ and transmit power $P$. Roughly speaking, this is because in high SNR regimes, doubling the effective bandwidth reduces the *standard derivation* of the delay estimation errors by half, and hence increases one intrinsic information bit for each delay parameter. Similarly, doubling the power reduces the *variances* of both the amplitude and delay estimation errors by half, and hence increases one intrinsic information bit for each pair of amplitude and delay parameters.

Third, for the low mobility scenarios, the intrinsic information increases logrithmatically with the path gains, inter-arrival times, and multipath parameter variations, where the variations depend on the node speed. Each of these parameters is a UDS that provides intrinsic information, and the information increases logrithmatically in high SNR regimes because more intrinsic digits are available. Similarly, for the high mobility scenarios, the intrinsic information increases logrithmatically with the path gains and inter-arrival times. The node speed is not involved because the multipath parameters in different channel measurements are not correlated.

Lastly, the intrinsic information also depends on the path overlapping effect characterized by $\frac{1}{2}\log|\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)})|$. Since the determinant is less than or equal to 1, path overlapping always decreases the intrinsic information unless the multipath channels are resolvable, in which case $\left|\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)})\right| = 1$ and the last terms in (21) and (22) vanish. Hence, resolvable channels are desirable for secret-key generation because they provide better channel parameter estimation, and consequently larger intrinsic information.

## V. Numerical and Simulation Results

In this section, we illustrate applications of our analytical results by numerical examples and simulation. We first give some numerical examples of the intrinsic information, and then quantify the intrinsic information of wideband channels based on the standard IEEE model by simulation.

### A. Numerical Properties

We first investigate the intrinsic information as a function of the parameter value $x$ where the observation noise is a Gaussian RV with zero mean and unit variance. Figure 2(a) depicts the intrinsic information given by (11), its upper and lower bounds given by (12), as well as the intrinsic information with $M$-ary representation given by (10). First, the intrinsic information is monotonically increasing with $x$ for a given noise distribution, since larger parameter values provide more intrinsic digits for secret-key generation. Second, the upper and lower bounds are tight for asymptotically large $x$, which is proven in Proposition 5. For instance, the gap between the bounds is less than 0.2 bits when $x \geq 10$, which is less than 5% of the intrinsic information. Third, the intrinsic information increases as $\log x$ for asymptotically large $x$, since both the upper and lower bounds increase logarithmically. This is consistent with the quantization theory that doubling the parameter results in one more representation bit in high SNR regimes. Fourth, the loss of intrinsic information due to $M$-ary representation decreases with $M$, but such a loss is not negligible when $M$ is small. For example, the loss is 0.36 bits

$$\mathcal{I}_{\mathrm{H}}(\mathbf{r}_{\mathrm{A}}^{(1:N)}; \mathbf{r}_{\mathrm{B}}^{(1:N)}) = L\Big(2N + N\log\frac{e\,\beta\,P}{N_0} - 1\Big) + \sum_{l=1}^{L}\big(2\log|\alpha_l^{(1)}| + \log|\tau_l^{(1)} - \tau_{l-1}^{(1)}|\big)$$
$$+ \sum_{n=1}^{N-1}\sum_{l=1}^{L}\big(\log|\delta_{\alpha}^{(n,l)}(v^{(n)})| + \log|\delta_{\tau}^{(n,l)}(v^{(n)})| + \log|\alpha_l^{(n+1)}|\big) + \frac{1}{2}\sum_{n=1}^{N}\log\left|\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)})\right| \tag{21}$$

$$\mathcal{I}_{\mathrm{H}}(\mathbf{r}_{\mathrm{A}}^{(1:N)}; \mathbf{r}_{\mathrm{B}}^{(1:N)}) = NL\Big(1 + \log\frac{e\,\beta\,P}{N_0}\Big) + \sum_{n=1}^{N}\sum_{l=1}^{L}\big(2\log|\alpha_l^{(n)}| + \log|\tau_l^{(n)} - \tau_{l-1}^{(n)}|\big) + \frac{1}{2}\sum_{n=1}^{N}\log\left|\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)})\right| \tag{22}$$

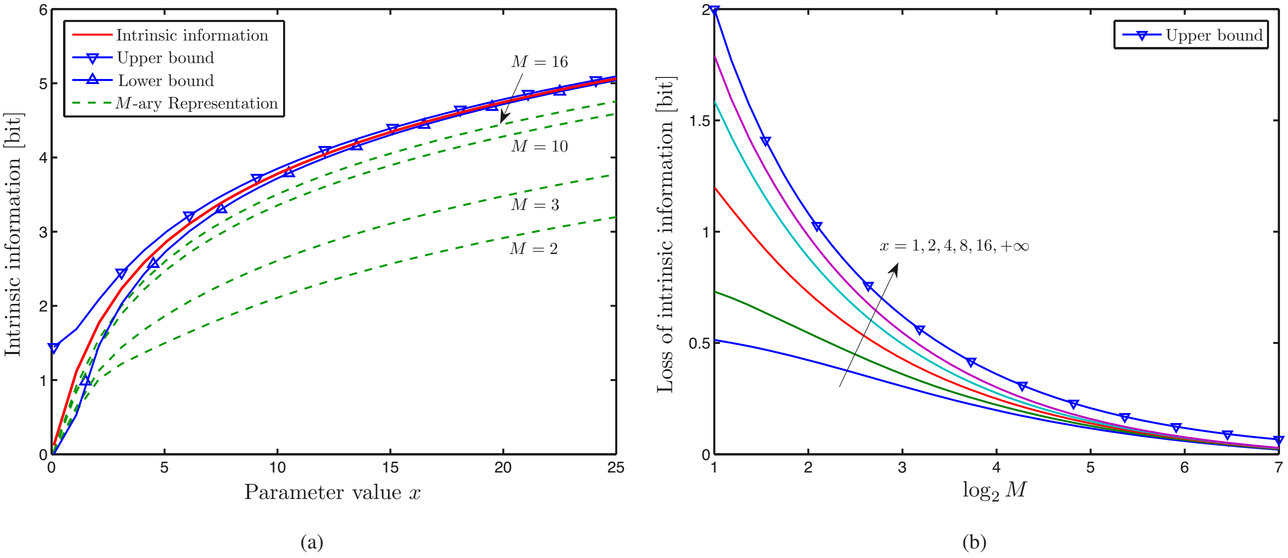


Fig. 2. (a) Intrinsic information, its upper and lower bounds, and the intrinsic information with $M$-ary representation as a function of the parameter value $x$. (b) Loss of intrinsic information due to $M$-ary representation and its upper bound as a function of the representation base $M$.

when $M = 16$ and 2 bits when $M = 2$ for large $x$, as shown in Proposition 6.

We then examine the loss of intrinsic information due to $M$-ary representation of the parameter value $x$. Figure 2(b) shows the loss of intrinsic information as well as its upper bound. First, the loss decreases with $M$ as shown in Theorem 4. This is because larger $M$ gives a better resolution of the intrinsic digits of the parameter. Second, the loss increases with $x$, approaching the upper bound derived in Proposition 6 when $x \to \infty$. For example, with $M = 2$, the gaps to the upper bound are about 1.5 and 0.25 bits for $x = 1$ and $x = 16$, respectively.

### B. Wideband Channels

We define the *intrinsic information rate* as the average intrinsic information per channel measurement, i.e.,

$$\mathcal{R}(r_{\mathrm{A}}(t); r_{\mathrm{B}}(t)) = \frac{1}{N}\,\mathcal{I}_{\mathrm{H}}(\mathbf{r}_{\mathrm{A}}^{(1:N)}; \mathbf{r}_{\mathrm{B}}^{(1:N)})$$

and next evaluate such an information rate of wideband channels by simulation.

We first give a numerical example for the intrinsic information rate of the wideband channels as a function of the node speed in low mobility scenarios described in Section IV-A. The network parameters are given as follows: the transmit signal is a sinc function with bandwidth $W$ GHz; the path arrival times follow the Poisson model with a fixed rate $2\,\mathrm{ns}^{-1}$; the power dispersion profile is an exponential function with delay time constant $15\,\mathrm{ns}$; the received SNR is 30 dB; and the variations of multipath parameters are proportional to speed, modeled as uniform RVs $\delta_\tau(v) \sim \mathcal{U}([-v/c, +v/c])$ and $\delta_\alpha(v) \sim \mathcal{U}([-|\alpha|v/10, +|\alpha|v/10])$.

Figure 3 shows the intrinsic information rate as a function of the node speed for different transmission bandwidths and numbers of MPCs based on (21). First, the information rate increases logarithmically with the speed and it is proportional to the number of MPCs. For example, doubling the speed increases the information by 50 and 100 bits for $L = 50$ and $L = 100$, respectively. Second, larger transmission bandwidth gives a larger information rate due to two factors: it (i) provides more accurate path delay estimates, and (ii) reduces the effect of path overlapping. As an example for $L = 50$, the information rate increases by 57.2 bits when the bandwidth increases from 5 to 10 GHz, where the former and latter factors contribute 50 and 7.2 bits, respectively.

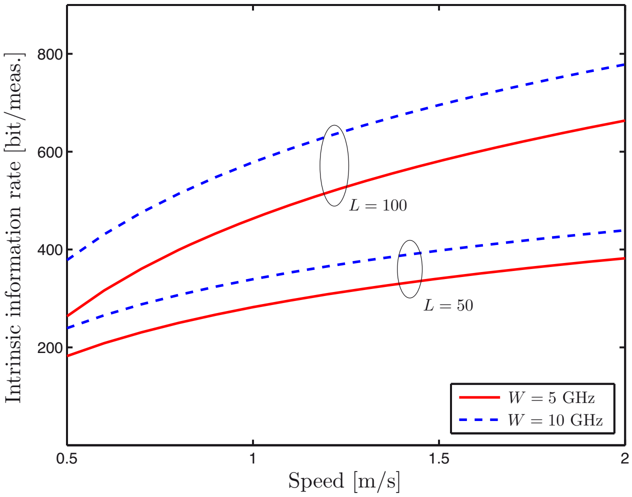
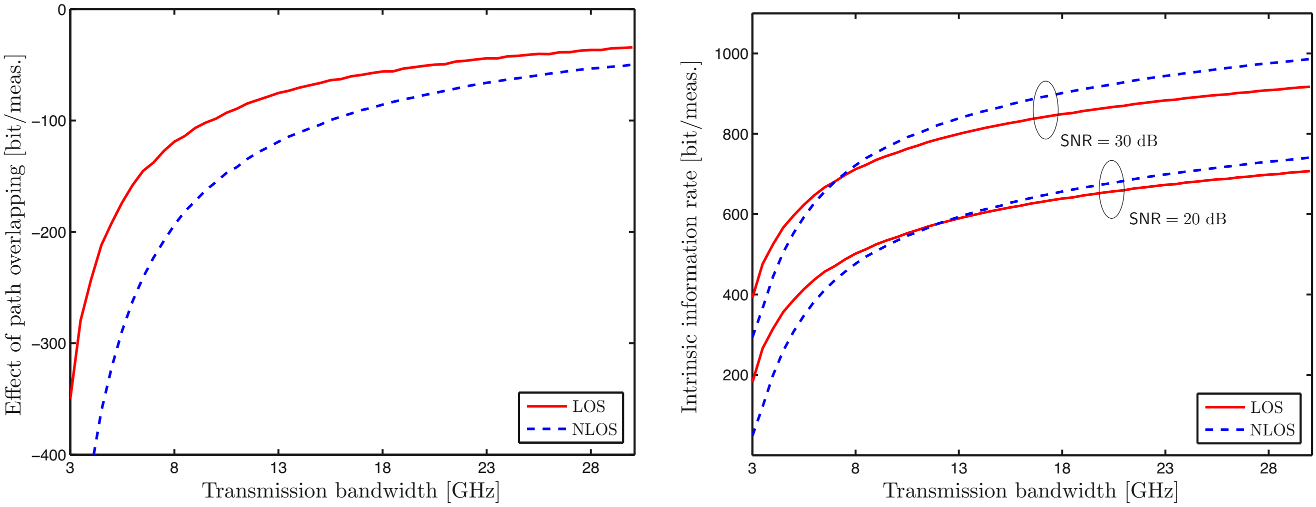


Fig. 3. Intrinsic information rates as a function of the node speed for different transmission bandwidth and number of MPCs. (SNR = 30 dB)

We then evaluate the intrinsic information rate of the wideband channels according to the IEEE 802.15.4a channel model [40] by Monte-Carlo simulation. In particular, we consider residential line-of-sight (LOS) CM1 and non-LOS (NLOS) CM2 environments. The transmit signal is a sinc function with bandwidth $W$ GHz. Figure 4(a) shows the effect



(a)

(b)

Fig. 4. (a) Effect of path overlapping on intrinsic information, $\frac{1}{2}\log|\mathbf{\Phi}|$, as a function of the transmission bandwidth in LOS and NLOS environments. (b) Intrinsic information rates as a function of the transmission bandwidth in LOS and NLOS environments.

of path overlapping on the intrinsic information rate as a function of the transmission bandwidth based on (22). The path overlapping always reduces the intrinsic information rate since $\log|\mathbf{\Phi}| \leq 0$, and the amount of this reduction decreases with the transmission bandwidth since larger bandwidth gives better channel resolvability. Moreover, the information rate loss due to the path overlapping is larger in NLOS scenarios than in LOS scenarios, because there are more MPCs in the former scenarios.[13]

Finally, Fig. 4(b) shows the intrinsic information rate as a function of the transmission bandwidth for different SNRs. First, a higher SNR gives a larger information rate, which is intuitive since it yields smaller errors in estimating the multipath parameters. Second, the information rate increases with the transmission bandwidth, because larger bandwidth increases the estimation accuracy of path delays as well as reduces the effect of path overlapping. In particular, the contribution of the former increases logarithmically with the bandwidth and is proportional to the number of MPCs, as shown in Fig. 4(a). Third, the information rate increases faster, with respect to the bandwidth, in NLOS scenarios than in LOS scenarios. This is due to the fact that more MPCs are present in NLOS scenarios. Therefore, for a given transmission bandwidth, the number of MPCs plays a dual role for the information rate: more MPCs increase the rate because of more parameters available for generating a secret key, but decrease the rate due to the path overlapping effect that degrades the estimation accuracy. When the bandwidth is large, the former factor dominates and thus the rate in NLOS scenarios is larger than that in the LOS scenarios, and vice versa when the bandwidth is small since the latter factor dominates. For example, as shown in Fig. 4(b), the curves of the information rates in the LOS and NLOS scenarios intersect at the bandwidth around 8 GHz for SNR = 30 dB.

## VI. Conclusion

In this paper, we developed a framework for secret-key generation using common UDSs and proposed a new information measure called intrinsic information to characterize the achievable secret-key length. In contrast to the existing work based on random sources, we considered a non-Bayesian scenario where the parameter of the source is unknown but deterministic. Through an information-theoretic analysis, we derived the intrinsic information of real-valued parameters represented in different bases as a function of the parameter value and noise distribution. Moreover, we determined the intrinsic information of a wideband propagation medium to characterize the potential of such a medium for secret-key generation. We also quantified the effect of the network parameters such as transmit power, transmission bandwidth, mobility, etc. on the intrinsic information. These results provide a new information-theoretic perspective and its practical implications for secret-key generation, which will foster new secure and authentication applications in mobile wireless networks.

## Appendix A
## Cardinality of Uncertainty Set

*Proposition 9:* The cardinality of the uncertainty set of $\lfloor x \rfloor_m$ with $M$-ary representation is given by

$$|\mathcal{S}_M(\lfloor x \rfloor_m)| = 2 \cdot (M-1)^{\mathbb{1}_{\{m+l_x \geq 0\}}} \cdot M^{[m+l_x]^+} \qquad (23)$$

for $x \in \mathbb{R}$. Moreover, when the partial knowledge of $x$ is available such that $x$ is nonnegative, the cardinality is

$$|\mathcal{S}_M^{+}(\lfloor x \rfloor_m)| = (M-1)^{\mathbb{1}_{\{m+l_x \geq 0\}}} \cdot M^{[m+l_x]^+}.$$

*Proof:* Recall the $M$-ary representation of $x$ quantized to the $m$th digit after the decimal point given by (2). There are three cases:

- if $m < -l_x$, $\mathcal{S}_M(\lfloor x \rfloor_m)$ involves only $\tilde{s}_x$, and hence $|\mathcal{S}_M(\lfloor x \rfloor_m)| = 2$ since there are two possible values for $\tilde{s}_x$;

[13] In the simulation, the average number of MPCs are 65 and 102 for LOS scenarios and NLOS scenarios, respectively.

- if $m = -l_x$, $\mathcal{S}_M(\lfloor x \rceil_m)$ involves $\tilde{s}_x$ and $\tilde{x}_{(-l_x)}$, and hence $|\mathcal{S}_M(\lfloor x \rceil_m)| = 2 \cdot (M-1)$ since there are $M-1$ possible values for $\tilde{x}_{(-l_x)}$;
- if $m > -l_x$, $\mathcal{S}_M(\lfloor x \rceil_m)$ involves $\tilde{s}_x$ and $\{\tilde{x}_{(-l_x)}, \cdots, \tilde{x}_{(m)}\}$, and hence $|\mathcal{S}_M(\lfloor x \rceil_m)| = 2 \cdot (M-1) \cdot M^{m+l_x}$ since there are $M$ possible values for each $\tilde{x}_{(i)}$ where $-l_x + 1 \le i \le m$.

In summary, the cardinality of $\mathcal{S}_M(\lfloor x \rceil_m)$ is given by (23). Moreover, for nonnegative parameters, the sign $s_x \equiv 1$ and hence the cardinality of the uncertainty set is reduced by half. □

## Appendix B
## Proof of Proposition 1

*Proof:* The likelihood function of the parameter $\lfloor x \rceil_m = k_2/2^m$ given the observation $Y_m = k_1/2^m$ is

$$\mathcal{L}\Big(\lfloor x \rceil_m = \frac{k_2}{2^m} \big| Y_m = \frac{k_1}{2^m}\Big) = \mathbb{P}\Big(Y_m = \frac{k_1}{2^m} \big| \lfloor x \rceil_m = \frac{k_2}{2^m}\Big)$$

for $k_2 \in \mathcal{K}_m$. We introduce a RV $V_m \in \mathcal{G}_{\mathrm{S},m}$ with a distribution characterizing the likelihood of $\lfloor x \rceil_m$ conditioned on $Y_m$, i.e., the conditional distribution of $V_m$ follows

$$\begin{aligned}
&\mathbb{P}\Big\{V_m = \frac{k_2}{2^m} \big| Y_m = \frac{k_1}{2^m}\Big\} \\
&= \frac{\mathcal{L}\big(\lfloor x \rceil_m = \frac{k_2}{2^m} \big| Y_m = \frac{k_1}{2^m}\big)}{\sum_{k_2 \in \mathcal{K}_m} \mathcal{L}\big(\lfloor x \rceil_m = \frac{k_2}{2^m} \big| Y_m = \frac{k_1}{2^m}\big)} \\
&= \begin{cases} \frac{P_{k_1-k_2,m}}{\sum_{k_2 \in \mathcal{K}_m} P_{k_1-k_2,m}}, & k_2 \in \mathcal{K}_m \\ 0, & \text{otherwise} \end{cases}
\end{aligned} \tag{24}$$

where the denominator is a normalization for the distribution.

For a given observation $Y_m$, Bob deciphers the cryptogram based on the likelihood function of $\lfloor x \rceil_m$, or equivalently $V_m$. Note that the message $W$ is uniformly distributed in the set $\mathcal{W}$ with cardinality $|\mathcal{W}| = |\mathcal{G}_{\mathrm{S},m}|$, and thus $H(W) = H(U_m)$. Moreover, since the enciphering and deciphering process is a reversible one-to-one mapping from $\mathcal{W}$ to itself, the conditional entropy of $W$ given $Y_m$ and the cryptogram is equal to the conditional entropy of $V_m$, i.e.,

$$H\Big(W \big| Y_m = \frac{k_1}{2^m}, e(W, \lfloor x \rceil_m) = w_i\Big) = H\Big(V_m \big| Y_m = \frac{k_1}{2^m}\Big)$$

where $\mathcal{W} := \{w_i : i = 1, 2, \ldots, |\mathcal{W}|\}$.

Finally, let $Y'_m := U_m + \lfloor N \rceil_m$, and one can verify that the conditional distribution of $U_m$ is equal to (24), i.e.,

$$\mathbb{P}\Big\{U_m = \frac{k_2}{2^m} \big| Y'_m = \frac{k_1}{2^m}\Big\} = \mathbb{P}\Big\{V_m = \frac{k_2}{2^m} \big| Y_m = \frac{k_1}{2^m}\Big\}$$

and therefore, it follows that

$$\begin{aligned}
&H(W | Y_m, e(W, \lfloor x \rceil_m)) \\
&= \sum_{i=1}^{|\mathcal{W}|} \sum_{k_1 \in \mathbb{Z}} \mathbb{P}\{e(W, \lfloor x \rceil_m) = w_i\} \mathbb{P}\Big\{Y_m = \frac{k_1}{2^m}\Big\} \\
&\qquad \cdot H\Big(W \big| Y_m = \frac{k_1}{2^m}, e(W, \lfloor x \rceil_m) = w_i\Big) \\
&= \sum_{k_1 \in \mathbb{Z}} \mathbb{P}\Big\{Y_m = \frac{k_1}{2^m}\Big\} H\Big(U_m \big| Y'_m = \frac{k_1}{2^m}\Big)
\end{aligned}$$

where the last equality is due to $\mathbb{P}\{e(W, \lfloor x \rceil_m) = w_i\} = 1/|\mathcal{W}|$. This leads to (4) since $H(W) = H(U_m)$. □

## Appendix C
## Proof of Proposition 2

*Proof:* Let $Y'_m := U_m + \lfloor N \rceil_m$. Based on the definition (5) and the secret-key length (4), we have

$$\begin{aligned}
&\mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m) \\
&= H(U_m) - \frac{1}{2^{m+1}} \sum_{k_0 \in \mathcal{K}_m} \sum_{k_1 \in \mathbb{Z}} P_{k_1-k_0,m} H\Big(U_m \big| Y'_m = \frac{k_1}{2^m}\Big) \\
&= H(U_m) - \sum_{k_1 \in \mathbb{Z}} \mathbb{P}\{Y'_m = k_1/2^m\} H\Big(U_m \big| Y'_m = \frac{k_1}{2^m}\Big) \\
&= H(U_m) - H(U_m | Y'_m) = I(U_m; U_m + \lfloor N \rceil_m)
\end{aligned}$$

where the second equality follows from $\mathbb{P}\{Y'_m = k_1/2^m\} = \frac{1}{2^{m+1}} \sum_{k_0 \in \mathcal{K}_m} P_{k_1-k_0,m}$. □

## Appendix D
## Proof of Proposition 3

*Proof:* (i) The intrinsic information $\mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m)$ is nonnegative because it can be represented as the mutual information between two RVs in (6), and it is bounded above because

$$\begin{aligned}
\mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m) &= H(U_m) - H(U_m | U_m + \lfloor N \rceil_m) \\
&\le H(U_m) = m + 1\,.
\end{aligned}$$

(ii) The distribution of $Y'_m = U_m + \lfloor N \rceil_m$ is given by

$$Q_{k,m} := \mathbb{P}\Big\{Y'_m = \frac{k}{2^m}\Big\} = \frac{1}{K_m} \sum_{j \in \mathcal{K}_m} P_{k-j,m} \tag{25}$$

where $K_m := |\mathcal{K}_m| = 2^{m+1}$, and the joint distribution of $Y'_m$ and $U_m$ is given by

$$\mathbb{P}\Big\{Y'_m = \frac{k}{2^m}, U_m = \frac{j}{2^m}\Big\} = \begin{cases} P_{k-j,m}/K_m, & j \in \mathcal{K}_m \\ 0, & \text{otherwise}. \end{cases}$$

Hence, the intrinsic information between $\lfloor x \rceil_m$ and $Y'_m$ in (6) can be obtained as

$$\mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_m; Y_m) = \sum_{k=-\infty}^{\infty} \Big[ g(Q_{k,m}) - \frac{1}{K_m} \sum_{j \in \mathcal{K}_m} g(P_{k-j,m}) \Big] \tag{26}$$

where $g(t) := -t \log t$. Similarly, we can obtain

$$\begin{aligned}
&\mathcal{I}_{\mathrm{S}}^{(2)}(\lfloor x \rceil_{m+1}; Y_{m+1}) \\
&= \sum_{k=-\infty}^{\infty} \Big[ g(Q_{2k,m+1}) + g(Q_{2k+1,m+1}) \qquad\qquad (27) \\
&\qquad - \frac{1}{K_m} \sum_{j \in \mathcal{K}_{m+1}} \frac{g(P_{2k-j,m+1}) + g(P_{2k+1-j,m+1})}{2} \Big].
\end{aligned}$$

Next, we denote

$$\begin{aligned}
Q_{k,m}^{(1)} &:= \frac{1}{K_m} \sum_{j \in \mathcal{K}_m} P_{2k-2j,m+1} \\
Q_{k,m}^{(2)} &:= \frac{1}{K_m} \sum_{j \in \mathcal{K}_m} P_{2k-2j+1,m+1}\,.
\end{aligned}$$

Since $P_{k,m} = P_{2k,m+1} + P_{2k+1,m+1}$, we have $Q_{k,m} = Q_{k,m}^{(1)} + Q_{k,m}^{(2)}$, $Q_{2k,m+1} = \big(Q_{k,m}^{(1)} + Q_{k,m}^{(2)}\big)/2$, and

$Q_{2k+1,m+1} = \big(Q^{(1)}_{k+1,m} + Q^{(2)}_{k,m}\big)/2$. Combining (26) and (27), we have

$$\mathcal{I}^{(2)}_{\mathrm{S}}(\lfloor x \rfloor_{m+1}; Y_{m+1}) - \mathcal{I}^{(2)}_{\mathrm{S}}(\lfloor x \rfloor_{m}; Y_{m}) = A - B \tag{28}$$

where $A$ is given by

$$\begin{aligned} A = & \sum_{k=-\infty}^{\infty} \Big[ g\Big(\frac{Q^{(1)}_{k,m} + Q^{(2)}_{k,m}}{2}\Big) + g\Big(\frac{Q^{(1)}_{k+1,m} + Q^{(2)}_{k,m}}{2}\Big) \\ & \qquad - g\big(Q^{(1)}_{k,m} + Q^{(2)}_{k,m}\big)\Big] \\ \geq & \sum_{k=-\infty}^{\infty} \Big[ \frac{g\big(Q^{(1)}_{k,m}\big) + g\big(Q^{(2)}_{k,m}\big)}{2} + \frac{g\big(Q^{(1)}_{k+1,m}\big) + g\big(Q^{(2)}_{k,m}\big)}{2} \\ & \qquad - g\big(Q^{(1)}_{k,m} + Q^{(2)}_{k,m}\big)\Big] \\ = & \sum_{k=-\infty}^{\infty} \tilde{g}(Q^{(1)}_{k,m}, Q^{(2)}_{k,m}) \end{aligned} \tag{29}$$

in which $\tilde{g}(t_1, t_2) := g(t_1) + g(t_2) - g(t_1 + t_2)$, the inequality follows from the concavity of $g(\cdot)$, and the last equality follows from reordering; and similarly $B$ is given by

$$B = \frac{1}{K_m} \sum_{k=-\infty}^{\infty} \sum_{j \in \mathcal{K}_m} \tilde{g}(P_{2k-2j,m+1}, P_{2k+1-2j,m+1}) \tag{30}$$

Therefore, by substituting (29) and (30) in (28), we can conclude that

$$\mathcal{I}^{(2)}_{\mathrm{S}}(\lfloor x \rfloor_{m+1}; Y_{m+1}) - \mathcal{I}^{(2)}_{\mathrm{S}}(\lfloor x \rfloor_{m}; Y_{m}) \geq 0$$

because $\tilde{g}(t_1, t_2)$ is a concave function, and $Q^{(1)}_{k,m}$ and $Q^{(2)}_{k,m}$ are weighted average of $P_{2k-2j,m+1}$ and $P_{2k+1-2j,m+1}$, respectively. □

## Appendix E
## Proof of Proposition 7

*Proof:* The lower bound can be derived as

$$\begin{aligned} \mathcal{I}(\mathbf{X}; \mathbf{Y}) &= \int_{\mathbf{0}}^{\mathbf{I}_d} \big[ h(\mathbf{U}_{\mathbf{x}} + \mathbf{T}\,\mathbf{N}_1) \\ &\qquad - h(\mathbf{T}(\mathbf{N}_1 - \mathbf{N}_2) | \mathbf{U}_{\mathbf{x}} + \mathbf{T}\,\mathbf{N}_1) \big]\, d\mathbf{T} \\ &\geq \int_{\mathbf{0}}^{\mathbf{I}_d} \big[ h(\mathbf{U}_{\mathbf{x}}) - h(\mathbf{T}(\mathbf{N}_1 - \mathbf{N}_2)) \big]\, d\mathbf{T} \\ &= h(\mathbf{U}_{\mathbf{x}}) - \int_{\mathbf{0}}^{\mathbf{I}_d} \log |\mathbf{T}|\, d\mathbf{T} - h(\mathbf{N}_1 - \mathbf{N}_2) \\ &= \mathcal{I}_{\mathrm{H}}(\mathbf{X}; \mathbf{Y}) \end{aligned}$$

where the last equation follows from $\int_0^1 \log t\, dt = -\log e$ and $h(\mathbf{U}_{\mathbf{x}}) = d + \sum_{i=1}^{d} \log |x_i|$.

Moreover, when $|x_i| \to \infty$ for all $i = 1, 2, \ldots, d$, we have $h(\mathbf{T}(\mathbf{N}_1 - \mathbf{N}_2) | \mathbf{U}_{\mathbf{x}} + \mathbf{T}\,\mathbf{N}_1) \nearrow h(\mathbf{T}(\mathbf{N}_1 - \mathbf{N}_2))$ and $h(\mathbf{U}_{\mathbf{x}} + \mathbf{T}\,\mathbf{N}_1) - h(\mathbf{U}_{\mathbf{x}}) \searrow 0$, leading to $\mathcal{I}(\mathbf{X}; \mathbf{Y}) - \mathcal{I}_{\mathrm{H}}(\mathbf{X}; \mathbf{Y}) \searrow 0$, i.e., the lower bound is asymptotically tight. □

## Appendix F
## Proof of Theorem 7

*Proof:* For the low mobility scenarios, note that the delays are always positive, but the amplitudes as well as the variations in delays and amplitudes can be negative. Hence, by the expression (15), we have the intrinsic information for the parameter vector $\boldsymbol{\eta}$ as

$$\begin{aligned} \mathcal{I}_{\mathrm{H}}&(\mathbf{r}_{\mathrm{A}}^{(1:N)}; \mathbf{r}_{\mathrm{B}}^{(1:N)}) = \mathcal{I}_{\mathrm{H}}(\hat{\boldsymbol{\eta}}_{\mathrm{A}}; \hat{\boldsymbol{\eta}}_{\mathrm{B}}) \\ &= (2N-1)L + 2NL \log e + \sum_{n=1}^{N} \sum_{l=1}^{2L} \log |\boldsymbol{\eta}_{2L(n-1)+l}| \\ &\quad - h(\widetilde{\boldsymbol{\eta}}_{\mathrm{A}} - \widetilde{\boldsymbol{\eta}}_{\mathrm{B}}) \\ &= (2N - 1 + 2N \log e) L \\ &\quad + \sum_{l=1}^{L} \big( \log |\alpha_l^{(1)}| + \log |\tau_l^{(1)} - \tau_{l-1}^{(1)}| \big) \\ &\quad + \sum_{n=1}^{N-1} \sum_{l=1}^{L} \big( \log |\delta_{\alpha}^{(n,l)}(v^{(n)})| + \log |\delta_{\tau}^{(n,l)}(v^{(n)})| \big) \\ &\quad - h(\widetilde{\boldsymbol{\eta}}_{\mathrm{A}} - \widetilde{\boldsymbol{\eta}}_{\mathrm{B}})\,. \end{aligned} \tag{31}$$

Since $\widetilde{\boldsymbol{\eta}}_{\mathrm{A}}$ and $\widetilde{\boldsymbol{\eta}}_{\mathrm{B}}$ are independent, $\widetilde{\boldsymbol{\eta}}_{\mathrm{A}} - \widetilde{\boldsymbol{\eta}}_{\mathrm{B}}$ follows $\mathcal{N}(\mathbf{0}, 2 \cdot \mathbf{J}_{\boldsymbol{\eta}}^{-1})$ and hence

$$\begin{aligned} h(\widetilde{\boldsymbol{\eta}}_{\mathrm{A}} - \widetilde{\boldsymbol{\eta}}_{\mathrm{B}}) &= \frac{1}{2} \log \big[ (2\pi e)^{2NL} |2 \cdot \mathbf{J}_{\boldsymbol{\eta}}^{-1}| \big] \\ &= 2NL + NL \cdot \log(\pi e) - \frac{1}{2} \log \big|\mathbf{J}_{\boldsymbol{\eta}}\big|\,. \end{aligned} \tag{32}$$

We next derive the expression for $\log \big|\mathbf{J}_{\boldsymbol{\eta}}\big|$. Since $\boldsymbol{\eta} = \mathbf{H}_1\, \boldsymbol{\theta}$, we have $\mathbf{J}_{\boldsymbol{\eta}} = \mathbf{H}_1\, \mathbf{J}_{\boldsymbol{\theta}}\, \mathbf{H}_1^{\dagger}$, where $\mathbf{J}_{\boldsymbol{\theta}}$ is the FIM for the parameter $\boldsymbol{\theta}$. Based on the channel model (16), the FIM $\mathbf{J}_{\boldsymbol{\theta}}$ can be derived after some algebra as $\mathbf{J}_{\boldsymbol{\theta}} = \mathrm{diag}\{\mathbf{J}_{\boldsymbol{\theta}^{(1)}}, \mathbf{J}_{\boldsymbol{\theta}^{(2)}}, \ldots, \mathbf{J}_{\boldsymbol{\theta}^{(N)}}\}$ where

$$\mathbf{J}_{\boldsymbol{\theta}^{(n)}} = \frac{2}{N_0}\, \mathbf{U} \cdot \boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)}) \cdot \mathbf{U}^{\dagger}$$

in which $\mathbf{U}$ is a $2L \times 2L$ diagonal matrix given by

$$\mathbf{U} = \sqrt{P} \sum_{l=1}^{L} \mathbf{E}_{l,l}^{2L} \otimes 1 + 2\pi\beta\, \sqrt{P} \sum_{l=L+1}^{2L} \mathbf{E}_{l,l}^{2L} \otimes \alpha_l^{(n)}\,.$$

It is easy to check $|\mathbf{H}_1| = 1$, and hence

$$\begin{aligned} \log \big|\mathbf{J}_{\boldsymbol{\eta}}\big| = \log \big|\mathbf{J}_{\boldsymbol{\theta}}\big| &= \sum_{n=1}^{N} \log \big|\mathbf{J}_{\boldsymbol{\theta}^{(n)}}\big| \\ &= \sum_{n=1}^{N} \log \Big( \frac{2^{2L}}{N_0^{2L}}\, |\mathbf{U}|^2\, |\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)})| \Big) \\ &= 2NL \cdot \log \frac{4P\pi\beta}{N_0} + \sum_{n=1}^{N} \sum_{l=1}^{L} 2 \log |\alpha_l^{(n)}| \\ &\quad + \sum_{n=1}^{N} \log \big|\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)})\big|\,. \end{aligned}$$

Substituting the above into (32) and then (31), we can obtain (21).

For the high mobility scenarios, the intrinsic information for the parameter vector $\boldsymbol{\eta}$ can be derived as

$$\begin{aligned}\mathcal{I}_{\mathrm{H}}(\mathbf{r}_{\mathrm{A}}^{(1:N)};\mathbf{r}_{\mathrm{B}}^{(1:N)}) &= (2N - N + 2N\log e)L \\ &\quad + \sum_{n=1}^{N}\sum_{l=1}^{L}\big(\log|\alpha_l^{(n)}| + \log|\tau_l^{(n)} - \tau_{l-1}^{(n)}|\big) \\ &\quad - h(\widetilde{\boldsymbol{\eta}}_{\mathrm{A}} - \widetilde{\boldsymbol{\eta}}_{\mathrm{B}})\end{aligned}$$

where $h(\widetilde{\boldsymbol{\eta}}_{\mathrm{A}} - \widetilde{\boldsymbol{\eta}}_{\mathrm{B}})$ can be obtained in the same way as (32) for the low mobility scenarios, leading to (22).

Moreover, $\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)}) \succeq \mathbf{0}$ since it is a kernel matrix, and its diagonal elements of are all equal to one. Hence, $\big|\boldsymbol{\Phi}(\boldsymbol{\tau}^{(n)}, \boldsymbol{\tau}^{(n)})\big| = 1$ when the off-diagonal elements are all zero, corresponding to the case in which the MPCs are resolvable. □

## Acknowledgment

The authors would like to thank M. Chiani, A. Conti, J. Lee, S. Mazuelas, A. Rabbachin, C. Chan, W. Dai, L. Ruan, A. J. Young, and T. Wang for helpful discussion and for the careful reading of the manuscript. They would also like to thank the anonymous reviewers for their constructive comments.

## References

[1] W. Diffie and M. E. Hellman, “New directions in cryptography,” *IEEE Trans. Inf. Theory*, vol. 22, no. 6, pp. 644–652, Nov. 1976.

[2] W. Diffie and M. E. Hellman, “Privacy and authentication: An introduction to cryptography,” *Proc. IEEE*, vol. 67, no. 3, pp. 397–427, Mar. 1979.

[3] J. L. Massey, “An introduction to contemporary cryptology,” *Proc. IEEE*, vol. 76, no. 5, pp. 533–549, May 1988.

[4] C. E. Shannon, “Communication theory of secrecy systems,” *Bell Syst. Tech. J.*, vol. 28, no. 4, pp. 656–715, Oct. 1949.

[5] A. D. Wyner, “The wire-tap channel,” *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1367, Oct. 1975.

[6] S. Leung-Yan-Cheong and M. E. Hellman, “The Gaussian wire-tap channel,” *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451–456, Jul. 1978.

[7] I. Csiszár and J. Korner, “Broadcast channels with confidential messages,” *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[8] R. Ahlswede and I. Csiszár, “Common randomness in information theory and cryptography - Part I: Secret sharing,” *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, Jul. 1993.

[9] U. M. Maurer, “Secret key agreement by public discussion from common information,” *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[10] P. C. Pinto and M. Z. Win, “Percolation and connectivity in the intrinsically secure communications graph,” *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1716–1730, Mar. 2012.

[11] M. Haenggi, “The secrecy graph and some of its properties,” in *Proc. IEEE Int. Symp. on Inf. Theory*, Toronto, ON, Jul. 2008, pp. 539–543.

[12] O. O. Koyluoglu, C. E. Koksal, and H. El Gamal, “On secrecy capacity scaling in wireless networks,” *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3000–3015, May 2012.

[13] J. Lee, A. Conti, A. Rabbachin, and M. Z. Win, “Distributed network secrecy,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, Sep. 2013.

[14] A. Rabbachin, A. Conti, and M. Z. Win, “The role of aggregate interference on intrinsic network secrecy,” in *Proc. IEEE Int. Conf. Commun.*, Ottawa, ON, Jun. 2012, pp. 3548–3553.

[15] U. M. Maurer and S. Wolf, “Unconditionally secure key agreement and intrinsic conditional information,” *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 499–514, Mar. 1999.

[16] U. M. Maurer and S. Wolf, “Information-theoretic key agreement: From weak to strong secrecy for free,” in *Advances in Cryptology – EUROCRYPT 2000*, vol. 1807. Springer-Verlag, 2000, pp. 351–368.

[17] I. Csiszár and P. Narayan, “Secrecy capacities for multiple terminals,” *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3047–3061, Dec. 2004.

[18] ——, “Secrecy capacities for multiterminal channel models,” *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2437–2452, Jun. 2008.

[19] S. Nitinawarat, C. Ye, A. Barg, P. Narayan, and A. Reznik, “Secret key generation for a pairwise independent network model,” in *Proc. IEEE Int. Symp. on Inf. Theory*, Toronto, ON, Jul. 2008, pp. 1015–1019.

[20] L. Lai, Y. Liang, and H. V. Poor, “A unified framework for key agreement over wireless fading channels,” *IEEE Trans. Inf. Forens. Security*, vol. 7, no. 2, pp. 480–490, Apr. 2012.

[21] A. O. Hero, “Secure space-time communication,” *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3235–3249, Dec. 2003.

[22] S. Goel and R. Negi, “Guaranteeing secrecy using artificial noise,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[23] E. Ekrem and S. Ulukus, “Secrecy in cooperative relay broadcast channels,” in *Proc. IEEE Int. Symp. on Inf. Theory*, Toronto, ON, Jul. 2008, pp. 2217–2221.

[24] Y. Liang, H. V. Poor, and S. Shamai, “Secure communication over fading channels,” *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, Jun. 2008.

[25] Y. Liang, A. Somekh-Baruch, H. V. Poor, S. Shamai, and S. Verdu, “Capacity of cognitive interference channels with and without secrecy,” *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 604–619, Feb. 2009.

[26] H. Weingarten, T. Liu, S. Shamai, Y. Steinberg, and P. Viswanath, “The capacity region of the degraded multiple-input multiple-output compound broadcast channel,” *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5011–5023, Nov. 2009.

[27] O. O. Koyluoglu, H. El Gamal, L. Lai, and H. Poor, “Interference alignment for secrecy,” *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3323–3332, Jun. 2011.

[28] W. J. Welch, “Reciprocity theorems for electromagnetic fields whose time dependence is arbitrary,” *IRE Trans. Antennas Propagat.*, vol. 8, no. 1, pp. 68–73, Jan. 1960.

[29] C. A. Balanis, *Antenna Theory: Analysis and Design*, 2nd ed. New York: Wiley, 1997.

[30] G. S. Smith, “A direct derivation of a single-antenna reciprocity relation for the time-domain,” *IEEE Trans. Antennas Propag.*, vol. 52, no. 6, pp. 1568–1577, Jun. 2004.

[31] R. C. Qiu, C. Zhou, N. Guo, and J. Q. Zhang, “Time reversal with MISO for ultrawideband communications: Experimental results,” *IEEE Antennas Wireless Propag. Lett.*, vol. 5, no. 1, pp. 269–273, Dec. 2006.

[32] A. A. Hassan, W. E. Stark, J. E. Hershey, and S. Chennakeshu, “Cryptographic key agreement for mobile radio,” *Digital Signal Process.*, vol. 6, no. 4, pp. 207–212, Oct. 1996.

[33] J. E. Hershey, A. A. Hassan, and R. Yarlagadda, “Unconventional cryptographic keying variable management,” *IEEE Trans. Commun.*, vol. 43, no. 1, pp. 3–6, Jan. 1995.

[34] M. A. Tope and J. C. McEachen, “Unconditionally secure communications over fading channels,” in *Proc. Military Commun. Conf.*, Vienna, VA, Oct. 2001, pp. 54–58.

[35] R. Wilson, D. Tse, and R. A. Scholtz, “Channel identification: Secret sharing using reciprocity in ultrawideband channels,” *IEEE Trans. Inf. Forens. Security*, vol. 2, no. 3, pp. 364–375, Sep. 2007.

[36] B. Azimi-Sadjadi, A. Kiayias, A. Mercado, and B. Yener, “Robust key generation from signal envelopes in wireless networks,” in *Proc. ACM Comput. Commun. Security*, Alexandria, VA, Oct. 2007, pp. 401–410.

[37] M. G. Madiseh, S. He, M. L. McGuire, S. W. Neville, and X. Dong, “Verification of secret key generation from UWB channel observations,” in *Proc. IEEE Int. Conf. Commun.*, Dresden, Germany, Jun. 2009, pp. 1–5.

[38] M. Z. Win and R. A. Scholtz, “Characterization of ultra-wide bandwidth wireless indoor communications channel: A communication-theoretic view,” *IEEE J. Sel. Areas Commun.*, vol. 20, no. 9, pp. 1613–1627, Dec. 2002.

[39] A. F. Molisch, “Ultrawideband propagation channels-theory, measurement, and modeling,” *IEEE Trans. Veh. Technol.*, vol. 54, no. 5, pp. 1528–1545, Sep. 2005.

[40] A. F. Molisch, D. Cassioli, C.-C. Chong, S. Emami, A. Fort, B. Kannan, J. Karedal, J. Kunisch, H. Schantz, K. Siwiak, and M. Z. Win, “A comprehensive standardized model for ultrawideband propagation channels,” *IEEE Trans. Antennas Propag.*, vol. 54, no. 11, pp. 3151–3166, Nov. 2006.

[41] D. Cassioli, M. Z. Win, and A. F. Molisch, “The ultra-wide bandwidth indoor channel: From statistical model to simulations,” *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1247–1257, Aug. 2002.

[42] A. F. Molisch, “Ultra-wide-band propagation channels,” *Proc. IEEE*, vol. 97, no. 2, pp. 353–371, Feb. 2009.

[43] H. L. Van Trees, *Detection, Estimation and Modulation Theory, Part 1*. New York, NY: Wiley, 1968.

[44] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.

[45] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[46] Y. Shen and M. Z. Win, "Fundamental limits of wideband localization – Part I: A general framework," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4956–4980, Oct. 2010.

[47] Y. Shen, H. Wymeersch, and M. Z. Win, "Fundamental limits of wideband localization – Part II: Cooperative networks," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4981–5000, Oct. 2010.

**Yuan Shen** (S'05) received the B.S. degree (with highest honor) in electrical engineering from Tsinghua University, China, in 2005 and the S.M. degree in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, MA, in 2008.

Since 2005, he has been with Wireless Communications and Network Science Laboratory at MIT, where he is currently a Ph.D. candidate. He was with the Electronics Engineering Department at Tsinghua University in summer 2013, the Wireless Communications Laboratory at The Chinese University of Hong Kong in summer 2010, the Hewlett-Packard Labs in winter 2009, the Corporate R&D of Qualcomm Inc. in summer 2008, and the Intelligent Sensing Laboratory at Tsinghua University from 2003 to 2005. His research interests include statistical inference, network science, communication theory, and information theory. His current research focuses on network localization and navigation, network inference techniques, intrinsic wireless secrecy, localization network optimization, and cooperative networks.

Mr. Shen served as a member of the Technical Program Committee for the IEEE Globecom in 2010–2013, the IEEE ICC in 2010–2013, the IEEE WCNC in 2009–2013, the IEEE ICUWB in 2011–2013, and the IEEE ICCC in 2012. He was a recipient of the Marconi Society Paul Baran Young Scholar Award (2010), the MIT EECS Ernst A. Guillemin Best S.M. Thesis Award (2008), the Qualcomm Roberto Padovani Scholarship (2008), and the MIT Walter A. Rosenblith Presidential Fellowship (2005). His papers received the IEEE Communications Society Fred W. Ellersick Prize (2012) and three Best Paper Awards from the IEEE Globecom (2011), the IEEE ICUWB (2011), and the IEEE WCNC (2007).

**Moe Z. Win** (S'85-M'87-SM'97-F'04) received both the Ph.D. in Electrical Engineering and the M.S. in Applied Mathematics as a Presidential Fellow at the University of Southern California (USC) in 1998. He received the M.S. in Electrical Engineering from USC in 1989 and the B.S. (*magna cum laude*) in Electrical Engineering from Texas A&M University in 1987.

He is a Professor at the Massachusetts Institute of Technology (MIT). Prior to joining MIT, he was with AT&T Research Laboratories for five years and with the Jet Propulsion Laboratory for seven years. His research encompasses fundamental theories, algorithm design, and experimentation for a broad range of real-world problems. His current research topics include network localization and navigation, network interference exploitation, intrinsic wireless secrecy, adaptive diversity techniques, and ultra-wide bandwidth systems.

Professor Win is an elected Fellow of the AAAS, the IEEE, and the IET, and he was an IEEE Distinguished Lecturer. He was honored with two IEEE Technical Field Awards: the IEEE Kiyo Tomiyasu Award (2011) and the IEEE Eric E. Sumner Award (2006, jointly with R. A. Scholtz). Together with students and colleagues, his papers have received numerous awards, including the IEEE Communications Society's Stephen O. Rice Prize (2012), the IEEE Aerospace and Electronic Systems Society's M. Barry Carlton Award (2011), the IEEE Communications Society's Guglielmo Marconi Prize Paper Award (2008), and the IEEE Antennas and Propagation Society's Sergei A. Schelkunoff Transactions Prize Paper Award (2003). Highlights of his international scholarly initiatives are the Copernicus Fellowship (2011), the Royal Academy of Engineering Distinguished Visiting Fellowship (2009), and the Fulbright Fellowship (2004). Other recognitions include the Laurea Honoris Causa from the University of Ferrara (2008), the Technical Recognition Award of the IEEE ComSoc Radio Communications Committee (2008), and the U.S. Presidential Early Career Award for Scientists and Engineers (2004).

Dr. Win is an elected Member-at-Large on the IEEE Communications Society Board of Governors (2011–2013). He was the Chair (2004–2006) and Secretary (2002–2004) for the Radio Communications Committee of the IEEE Communications Society. Over the last decade, he has organized and chaired numerous international conferences. He is currently an Editor-at-Large for the IEEE WIRELESS COMMUNICATIONS LETTERS and is serving on the Editorial Advisory Board for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He served as Editor (2006–2012) for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and also as Area Editor (2003–2006) and Editor (1998–2006) for the IEEE TRANSACTIONS ON COMMUNICATIONS. He was Guest-Editor for the PROCEEDINGS OF THE IEEE (2009) and IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (2002).