Authentication Over Noisy Channels

Lifeng Lai, *Member, IEEE*, Hesham El Gamal, *Senior Member, IEEE*, and H. Vincent Poor, *Fellow, IEEE*

***Abstract*—An authentication counterpart of Wyner's study of the wiretap channel is developed in this work. More specifically, message authentication over noisy channels is studied while impersonation and substitution attacks are investigated for both single- and multiple-message scenarios. For each scenario, information-theoretic lower and upper bounds on the opponent's success, or cheating, probability are derived. Remarkably, in both scenarios, the lower and upper bounds are shown to match, and hence, the fundamental limits on message authentication over noisy channels are fully characterized. The opponent's success probability is further shown to be smaller than that derived in the classical noiseless channel model. These results rely on a novel authentication scheme in which shared key information is used to provide simultaneous protection against both types of attacks. Finally, message authentication for the case in which the source and receiver possess only correlated sequences is studied.**

***Index Terms*—Authentication, impersonation attack, noisy channel, substitution attack, wiretapper.**

I. Introduction

SECRECY ENCODING, which ensures that a message is decoded successfully only by its legitimate *receiver*, and *message authentication*, which ensures that an accepted message truly comes from its acclaimed *transmitter*, are among the fundamental building blocks of secure communication systems.

The fundamental limits on communication secrecy have been investigated under the two different models shown in Fig. 1, where a passive opponent O tries to overhear the message sent by the source S to the receiver R. The difference between these two scenarios lies in the channel model. In the case investigated by Shannon [1], i.e., Fig. 1(a), the channel is assumed to be noiseless and the source and intended destination use a common secret key K to encrypt and decrypt the message M. Transmission is said to be perfectly secure if the signal received at the opponent does not reveal any information about M, i.e., $I(M;X) = 0$, where $I(\cdot;\cdot)$ denotes mutual information between its two arguments. Shannon proved that one needs $H(K) \geq H(M)$ to ensure perfect secrecy, where $H(\cdot)$ denotes the entropy of its argument. Taking channel noise into consideration, Wyner developed the wiretap channel model [2], i.e., Fig. 1(b). In this model, the transmitter exploits the independence between the two noise processes impairing the receiver and opponent to secure M. More precisely, in this model, the source and receiver do not share a secret key *a priori* and the secrecy capacity is defined to be the largest transmission rate such that the decoding error at the receiver is arbitrarily small and the *equivocation rate*[1] at the opponent is arbitrarily close to the transmission rate. Csiszár and Körner [3] generalized this model and characterized the capacity of the discrete memoryless channel (DMC) with security constraints. The generalization of these results to multiuser channels and wireless networks has recently emerged as a very active research area (e.g., [4]–[12]).

The model for authentication over noiseless channels, Fig. 2, was developed by Simmons [13]. In this model, the source S and the receiver R share a secret key K, which is used to identify the transmitter. The transmitter and receiver are assumed to be honest, i.e., they will follow the rules and will not attack the system by faking messages. When the transmitter intends to send the message M, it transmits $W = f(K, M)$ over a noiseless public channel, where f is the encoding function. On receiving $\hat{W}$, which might be different from W due to various possible active attacks from the opponent O, the receiver needs to decide whether the packet came from the legitimate transmitter or not. If the receiver accepts the packet (i.e., the receiver believes that the packet is authentic), then it computes an estimate of the source message M; otherwise, it rejects the packet. In this model, the opponent obtains a perfect copy of W. The following two types of attacks are considered. The first one is called an *impersonation attack*, in which the opponent sends W' to the destination before the source sends anything. This attack is successful if W' is accepted by the receiver as authentic, and the success probability of this attack is denoted as P_I. The second attack is referred to a *substitution attack*, in which after receiving W, the opponent modifies it to W' and sends it to the destination. The attack is successful if the receiver accepts W' and decodes this into another erroneous source message. The success probability of this attack is denoted as P_S. Obviously, the opponent will choose the attack that has a higher success probability. Hence the success probability P_D of the opponent (i.e., the *cheating probability*) is $P_D = \max\{P_I, P_S\}$. This model is further extended to the scenario in which neither the source nor the receiver is honest [14]. To solve possible disputes, an honest arbiter is introduced into this model. Desmedt *et al.* [15] further studied the scenario in which the arbiter is also assumed to be dishonest, and hence, can potentially initiate an attack.

Manuscript received February 23, 2008; revised February 23, 2008. Current version published February 04, 2009. This work was supported in part by the National Science Foundation under Grants ANI-03-38807, CNS-06-25637, and CCF-07-28208.The material in this paper was presented in part at the 45th Annual Allerton Conference on Communication, Control and Computing, Monticello, IL, September 2007.

L. Lai and H. V. Poor are with the Department of Electrical Engineering, Princeton University, Engineering Quadrangle, Princeton, NJ 08544 USA (e-mail: llai@princeton.edu; poor@princeton.edu).

H. El Gamal is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus OH 43210 USA. He also serves as the founding Director for the Wireless Intelligent Networks Center (WINC), Nile University, Cairo, Egypt (e-mail: helgamal@ece.osu.edu).

Communicated by T. Fujiwara, Associate Editor for Complexity and Cryptography.

Digital Object Identifier 10.1109/TIT.2008.2009842

[1]Interested readers are referred to [2] for a rigorous definition of the equivocation rate.

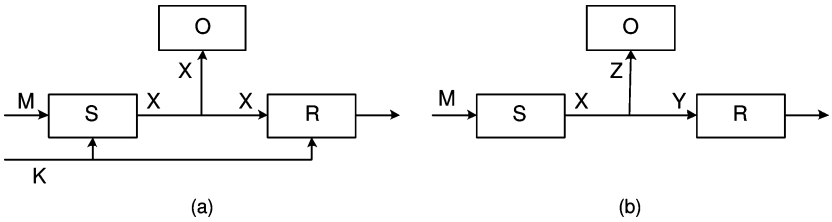

Fig. 1. Two models for secure communications.

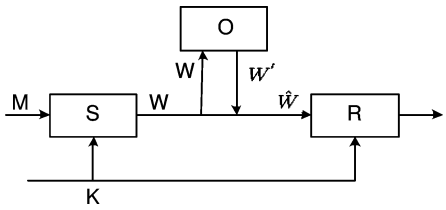

Fig. 2. The authentication channel.

Lower bounds on P_I and P_S have been developed in [13] from first principle and recovered by Maurer [16] from a hypothesis testing perspective. In particular, it has been shown that[2] $P_I \geq 2^{-I(K;W)}$ and $P_S \geq 2^{-H(K|W)}$, where $H(K|W)$ denotes the conditional entropy of K given W. One can easily identify a tradeoff between P_I and P_S. To minimize the probability of a successful impersonation attack, the transmitted ciphertext must contain a sufficient amount of information about the secret key in order to convince the receiver that the transmitted message comes from the legitimate source. That is $I(K;W)$ should be large, which unfortunately decreases $H(K|W)$, since $H(K|W) = H(K) - I(K;W)$. Hence, the attacker can take advantage of the leaked information over its noiseless channel (contained in W) to increase the probability of a successful substitution attack. In fact, the strategy that minimizes the lower bound on $P_D = \max\{P_I, P_S\}$ is to use half of the key information to protect against the impersonation attack and the other half of the key information to protect against the substitution attack, which gives $P_D \geq 2^{-H(K)/2}$. These bounds are of a negative nature, since they give only lower bounds on the cheating probability. The opponent may be able to achieve much better performance.

Simmons's model assumes a noiseless channel. However, since physical transmission systems are noisy, the current common practice is to use channel coding to convert the noisy channel into a noiseless one, and then to design an authentication code on top of channel coding. Liu and Boncelet [18], [19] also considered the situation in which channel coding is not perfect, and hence, there are some residual errors induced by the channel. The main conclusion of these works is that channel noise is *detrimental* to authentication, since it will cause the receiver to reject authentic messages from the transmitter.

In this paper, we take an alternative view of the noisy channel model and design channel and authentication coding jointly. This way, we are able to *exploit* the channel noise to hide the key information from the opponent. The codebook of our channel code is designed such that the conditional distribution of the keys after observing the noisy output at the opponent is very close to a uniform distribution,[3] and hence the opponent is unable to use the noisy observations to increase the success probability of a substitution attack. By using this approach, we derive an upper bound on the cheating probability which is significantly smaller than the existing lower bounds for the noiseless channel model. Moreover, this upper bound is shown to coincide with a simple lower bound on the cheating probability. In particular, we show that $P_D = 2^{-H(K)}$, and thus all the key information can be used to protect against substitution and impersonation attacks simultaneously. We further consider the authentication of multiple messages using the same key K over the noisy channel. Similar to the single-message case, lower and upper bounds on the cheating probability are derived and shown to coincide. Again, all the key information can be used to protect against all the attacks simultaneously. We then investigate a scenario in which the source and receiver possess only correlated sequences, instead of sharing a key. In the same spirit as our earlier results, we show that all of the mutual information between these two sequences can be used to protect against the two attacks considered in this work.

The rest of the paper is organized as follows. In Section II, we introduce our system model and notation. Section III is devoted to the single message authentication scenario. Next, we analyze the authentication of multiple message using the same key in Section IV. Message authentication using correlated sequences is considered in Section V. Finally, in Section VI, we offer some concluding remarks.

II. System Model

Throughout this paper, upper case letters (e.g., X) will denote random variables, lower case letters (e.g., x) will denote realizations of the corresponding random variables, and calligraphic letters (e.g., $\mathcal{X}$) will denote finite alphabet sets over which corresponding variables range. Also, upper case boldface letters (e.g., $\boldsymbol{X}$) will denote random vectors whereas lower case boldface letters (e.g., $\boldsymbol{x}$) will denote realizations of the corresponding random vectors.

Fig. 3 shows the new model under consideration. It differs from Simmons's model only in the channel, which is assumed to be noisy in our model. More specifically, we consider the DMC and assume that when the transmitter sends $\boldsymbol{x}$, the opponent receives $\boldsymbol{z}$ with probability

$$P(\boldsymbol{z}|\boldsymbol{x}) = \prod_{t=1}^{n} P(z(t)|x(t))$$

[2]A slightly better lower bound on the impersonation attack was developed in [17].

[3]Rigorous definitions of distance and closeness will be given in the sequel.

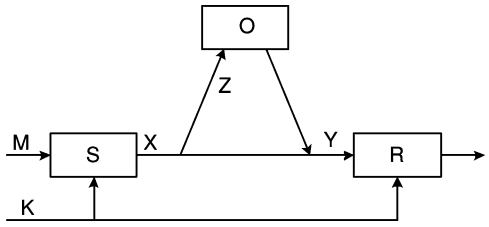

Fig. 3. The authentication channel.

where n is the length of the transmitted vector. If the opponent does not initiate any attack, the legitimate destination will receive $\boldsymbol{y}$ with probability

$$P(\boldsymbol{y}|\boldsymbol{x}) = \prod_{t=1}^{n} P(y(t)|x(t)).$$

If the opponent initiates an attack, the value of $\boldsymbol{y}$ will depend on the strategy of the opponent. Here $P(y(t)|x(t))$ and $P(z(t)|x(t))$ denote the indicated channel transition probabilities, while $x(t)$, $y(t)$, and $z(t)$ range through the finite sets $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, respectively. In order to derive more general bounds, we assume that the channel between the opponent and receiver is noiseless, and that the opponent can send anything over this channel. It is worth noting that this assumption does not incur any loss of generality, and actually gives the opponent an advantage, since any noisy channel can be simulated with this noiseless channel by simply randomizing the transmitted signal.

To identify the transmitter, we assume that the source and the destination have a common secret key K uniformly chosen from a set $\mathcal{K}$ having $|\mathcal{K}|$ possible values. To transmit the message M, the source uses a stochastic encoding function f to convert the message and key into a length n vector $\boldsymbol{X}$, i.e., $\boldsymbol{X} = f(K, M)$. Upon receiving $\boldsymbol{Y}$, which may come from either the source or the opponent, the destination uses a decoding function g to obtain an estimate of the message and key, that is. $(M', K') = g(\boldsymbol{Y})$. If $K' = K$, the receiver accepts the message. Otherwise, the receiver rejects the message. We require that, if the signal is authentic, the decoding error probability at the destination must approach zero as the length of the code increases, i.e., for any $\epsilon > 0$, there is a positive integer n_0, such that for all $n \geq n_0$, we have

$$P_e = \Pr\{M' \neq M | \boldsymbol{Y} \text{ comes from } \boldsymbol{X}\} \leq \epsilon.$$

There are two components of the error probability P_e: P_1 and P_2, where P_1 is the probability of a miss, which is the probability that the receiver wrongly rejects an authentic message, and P_2 is the probability that the decoder correctly accepts the signal as being authentic but incorrectly decodes it.

The opponent is assumed to be aware of the system design, except for the particular realizations k and m of the key K and message M. We consider the two forms of attack described above. That is, we consider the impersonation attack, in which the opponent sends a codeword $\boldsymbol{X}$ to the receiver before the transmitter sends anything. Such an attack is successful if $\boldsymbol{X}$ is accepted as authentic by the receiver, and we denote this probability of success by P_I as noted above. We also consider the substitution attack, in which the opponent blocks the transmission of the main channel while receiving $\boldsymbol{Z}$. After that, the opponent modifies the signal and transmits it to the receiver. This attack is considered to be successful if the modified signal is accepted as authentic by the receiver and is decoded into m' that is not equal to the original message m. Again, the success probability of this attack is denoted by P_S.

III. Authentication of a Single Message

A. The Wiretap Channel

We begin by reviewing some results related to the wiretap channel introduced in [2]. The wiretap channel is defined by two DMCs $\mathcal{X} \rightarrow (\mathcal{Y}, \mathcal{Z})$, where $\mathcal{X}$ is the input alphabet from the transmitter, $\mathcal{Y}$ is the output alphabet at the legitimate receiver, and $\mathcal{Z}$ is the output alphabet at the wiretapper. In the wiretap channel, the wiretapper is assumed to be passive, and the goal is to transmit information to the destination while minimizing the information leakage to the wiretapper. More specifically, to send a message $m \in \mathcal{M}$, the transmitter sends $\boldsymbol{x} = f(m)$, where f is a stochastic encoder. After receiving $\boldsymbol{y}$, the destination obtains an estimate $m' = g(\boldsymbol{y})$. The opponent is fully aware of the system design, and hence, knows the codebook used by the source. A perfectly secure rate R_s is said to be achievable if there exist f and g, such that for each $\epsilon > 0$, there is a positive integer n_0, such that $\forall n > n_0$

$$|\mathcal{M}| \geq 2^{nR_s} \tag{1}$$

$$\Pr\{M' \neq M\} \leq \epsilon \quad \text{and} \tag{2}$$

$$\frac{1}{n} I(M; \boldsymbol{Z}) \leq \epsilon. \tag{3}$$

The perfect secrecy capacity C_s is defined to be the supremum of the set of R_s values that satisfy conditions (1)–(3). It is proved in [3] that the perfect secrecy capacity is given by

$$C_s = \max_{U \rightarrow X \rightarrow YZ} [I(U;Y) - I(U;Z)]^+$$

where U is an auxiliary random variable satisfying the Markov chain relationship $U \rightarrow X \rightarrow (YZ)$.

The source–wiretapper channel is said to be less noisy than the main channel if, for all possible U that satisfy the above Markov chain relationship, one has $I(U;Z) > I(U;Y)$. On the other hand, if the source–wiretapper channel is not less noisy than the main channel, there exists a distribution satisfying $U \rightarrow X \rightarrow (Y, Z)$ such that $I(U;Y) > I(U;Z)$, and thus the perfect secrecy capacity is nonzero. One of the main insights gleaned from the wiretap channel is that perfectly secure communication is possible, without sharing a secret key *a priori* between the source and destination, by using a codebook whose codeword rate is higher than the secret message rate R_s (i.e., one message will correspond to several different codewords). Usually, the codeword rate is set to be the rate that can be supported by the source–destination channel allowing the legitimate receiver to recover the correct codeword while confusing the opponent with the high codeword rate.

B. Proposed Authentication Scheme

In authentication applications, when the source is sending information, the opponent tries to overhear the message and uses the information gained to initiate a substitution attack. This eavesdropping stage corresponds to the wiretap channel model. This observation motivates our approach of using a wiretap channel code to protect our authentication key. More specifically, there exists an input distribution P_X such that $I(X;Y) - I(X;Z) > 0$, then for a given fixed key size $|\mathcal{K}|$ and arbitrarily small $\delta > 0$, there exists a positive integer n_1, such that $\forall n \geq n_1$

$$2^{n(I(X;Y)-I(X;Z)-2\delta)} > |\mathcal{K}|. \tag{4}$$

Also, for a given message size $|\mathcal{M}|$ and key size $|\mathcal{K}|$, there exists a positive integer n_2, such that $\forall n \geq n_2$

$$2^{n(I(X;Y)-\delta)} > |\mathcal{M}||\mathcal{K}|. \tag{5}$$

In our transmission scheme, the source first generates[4] a codebook for the wiretap channel with $2^{n(I(X;Y)-\delta)}$ codewords, whose length n satisfies conditions (4), (5), and a low decoding error probability requirement. The source then partitions the codebook into $|\mathcal{K}|$ subsets, associating one subset with each key. Since the length satisfies (5), there are more than $|\mathcal{M}|$ codewords in each subset. The source then further divides each subset into $|\mathcal{M}|$ bins, each corresponding to a message. There are multiple codewords in each bin. The codebook used in our authentication scheme is shown in Fig. 4. In the transmission, if the intended message is m, and the key is k, the source then randomly chooses a codeword $\boldsymbol{x}$ from the mth bin of the kth subset using a uniform distribution. The source then transmits $\boldsymbol{x}$ over the channel. The opponent receives $\boldsymbol{z}$ with probability

$$P(\boldsymbol{z}|\boldsymbol{x}) = \prod_{t=1}^{n} P(z(t)|x(t)).$$

The receiver receives $\boldsymbol{y}$. If the opponent does not initiate any attack, then

$$P(\boldsymbol{y}|\boldsymbol{x}) = \prod_{t=1}^{n} P(y(t)|x(t)).$$

On the other hand, if the opponent chooses to attack, the value of $\boldsymbol{y}$ depends on the opponent's attack strategy.

After receiving $\boldsymbol{y}$, the destination first obtains an estimate $\hat{\boldsymbol{x}}$ of the transmitted codeword using typical set decoding; that is, the destination decodes $\boldsymbol{y}$ into $\hat{\boldsymbol{x}}$ if $(\hat{\boldsymbol{x}}, \boldsymbol{y})$ are jointly typical. It then obtains an estimate m' of the message and an estimate k' as the corresponding bin index and subset index, respectively, of $\hat{\boldsymbol{x}}$. We denote the decoding process at the destination as $(m', k') = g(\boldsymbol{y})$. If $k' = k$, the receiver accepts the message as authentic; otherwise, it rejects the message.

Note that in the noiseless model, this scheme does not work, since the opponent can obtain a perfect copy of $\boldsymbol{x}$, and hence can determine the values of k and m. Thus, the substitution attack will be successful. In the noisy channel model, if we design the code properly, the output at the opponent will not provide it with such information, as shown in the sequel.

[4]An explicit procedure for generating and partitioning the codebook will be given in the sequel.

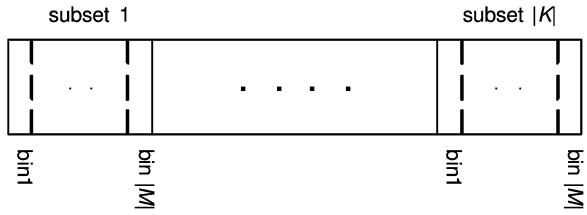

Fig. 4. The codebook used in our authentication scheme. The codebook is divided into $|\mathcal{K}|$ subsets, each of which is further partitioned into $|\mathcal{M}|$ bins. Each subset corresponds to a key k. Each bin in each subset corresponds to a message m.

First, let us consider the impersonation attack. The optimal strategy for the opponent is to transmit a codeword from the subset corresponding to the key that has the largest probability of being accepted by the receiver, i.e., $\boldsymbol{y}_o$, which will be transmitted by the opponent, should be chosen from the subset corresponding to k' such that the following probability is maximized:

$$\sum_{k\in\mathcal{K}} P(k)\gamma_1(k,k')$$

where $\gamma_1(k,k')$ is an indicator function that equals 1 if k' is accepted as authentic, and equals 0 in other cases. In our scheme, $\gamma_1(k,k') = 1$ if $k' = k$; otherwise, $\gamma_1(k,k') = 0$, and hence

$$P_I = \max_{k'\in\mathcal{K}} \left\{ \sum_{k\in\mathcal{K}} P(k)\gamma_1(k,k') \right\}.$$

For a substitution attack, the opponent knows $\boldsymbol{z}$, and hence can choose $\boldsymbol{y}_o$ based on this information. Let h be the transformation employed by the opponent to transform $\boldsymbol{z}$ to $\boldsymbol{y}_o$. Here, h can be any function, either deterministic or stochastic. Also, denote $(m', k') = g(\boldsymbol{y}_o) = g(h(\boldsymbol{z}))$. Note that g is the decoding function at the destination, and hence m' and k' are the decoded message and key at the destination after the opponent's attack. Obviously, for each observation $\boldsymbol{z}$, the opponent should choose h so that

$$\sum_{m,k} P(m,k|\boldsymbol{z})\gamma_2(m,m')\gamma_1(k,k')$$

is maximized. Here $\gamma_2(m,m') = 1$ if $m' \neq m$ and equals 0 otherwise. Meanwhile, as defined above, $\gamma_1(k,k') = 1$ if $k' = k$, and equals 0 otherwise. Hence, the success probability of the substitution attack is

$$P_S = \sum_{\boldsymbol{z}} P(\boldsymbol{z}) \sup_h \left\{ \sum_{m,k} P(m,k|\boldsymbol{z})\gamma_2(m,m')\gamma_1(k,k') \right\}. \tag{6}$$

To simplify the analysis, we have the following lemma.

Lemma 1: For any substitution attack strategy h of the opponent, we have

$$P_S \leq \sum_{\boldsymbol{z}} P(\boldsymbol{z}) \max_{k\in\mathcal{K}} \{P(k|\boldsymbol{z})\}. \tag{7}$$

Proof: We can bound P_S as follows:

$$P_S = \sum_{\boldsymbol{z}} P(\boldsymbol{z}) \sup_h \left\{ \sum_{m,k} P(m,k|\boldsymbol{z})\gamma_2(m,m')\gamma_1(k,k') \right\}$$

$$\begin{aligned}
&\overset{(a)}{\leq} \sum_{\boldsymbol{z}} P(\boldsymbol{z}) \sup_h \left\{ \sum_{m,k} P(m,k|\boldsymbol{z}) \gamma_1(k,k') \right\} \\
&= \sum_{\boldsymbol{z}} P(\boldsymbol{z}) \sup_h \left\{ \sum_k \left(P(k|\boldsymbol{z}) \gamma_1(k,k') \sum_m P(m|k,\boldsymbol{z}) \right) \right\} \\
&\overset{(b)}{=} \sum_{\boldsymbol{z}} P(\boldsymbol{z}) \sup_h \left\{ \sum_k P(k|\boldsymbol{z}) \gamma_1(k,k') \right\} \\
&\overset{(c)}{\leq} \sum_{\boldsymbol{z}} P(\boldsymbol{z}) \max_{k \in \mathcal{K}} \{P(k|\boldsymbol{z})\}. \qquad (8)
\end{aligned}$$

In this expression, inequality (a) follows from the fact that $\gamma_2(m,m') \leq 1$ for any h, m, and m'; inequality (b) comes from the fact that $\sum_m P(m|k,\boldsymbol{z}) = 1$ for any k and $\boldsymbol{z}$; and inequality (c) comes from the fact that

$$\sum_k P(k|\boldsymbol{z}) \gamma_1(k,k') \leq \max_{k \in \mathcal{K}} \{P(k|\boldsymbol{z})\} \qquad (9)$$

since only one term of $\gamma_1(k,k')$ is 1, and all the remaining terms are zero. □

This result shows that, after receiving $\boldsymbol{Z}$, the success probability of any substitution attack is upper-bounded by the probability of the most likely key. With $\boldsymbol{Z}$, the opponent gains an amount $I(K;\boldsymbol{Z})$ of information about the key, and thus can use this information to choose k that maximizes $P(k|\boldsymbol{z})$. From (3), we have that

$$I(K;\boldsymbol{Z}) \leq n\epsilon. \qquad (10)$$

The inequality in (10) is not enough to analyze(8) for the following two reasons. First, though ϵ is small, $n\epsilon$ might go to infinity as n grows, and hence the opponent may eventually gain a sufficient amount of information about the key. This point has been pointed out in [20]–[22]. The second reason is that there is a maximization in the summand in (8), which means that we need to consider the worst case scenario, whereas $I(K;\boldsymbol{Z})$ is an average quantity. Actually, this fact is exploited in [13] and [16] to derive lower bounds by replacing this maximization with an averaging, which readily gives us a lower bound and is more amenable to analysis.

In Section III-C, we borrow techniques from [22] and [23] to analyze this term.

C. Bounds

We begin with some definitions. Let $\mathcal{C}$ be a codebook for the wiretap channel, and let $\tilde{P}(\boldsymbol{x},\boldsymbol{z})$ be the joint distribution on $\mathcal{C} \times \mathcal{Z}^n$. We denote by $Q(\boldsymbol{z})$ the marginal distribution of $\boldsymbol{z}$ when the input distribution is limited to, and is uniform on, $\mathcal{C}$, and by

$$P(\boldsymbol{x}|\boldsymbol{z}) = \tilde{P}(\boldsymbol{x},\boldsymbol{z})/Q(\boldsymbol{z})$$

the conditional distribution of $\boldsymbol{X}$ given $\boldsymbol{Z} = \boldsymbol{z}$.

Let $\{\mathcal{C}_1, \ldots, \mathcal{C}_N\}$ be a partition of $\mathcal{C}$, and denote this partition as a mapping, i.e., $f: \mathcal{C} \to \{\mathcal{C}_1, \ldots, \mathcal{C}_N\}$. Also denote by Q_k the distribution of $\boldsymbol{Z}$ when the input distribution is uniform on $\mathcal{C}_k$, i.e.,

$$Q_k(\boldsymbol{z}) = \sum_{\boldsymbol{x} \in \mathcal{C}_k} \tilde{P}(\boldsymbol{x},\boldsymbol{z})/P(\mathcal{C}_k).$$

Define

$$d_{\text{av}}(f) = \sum_{k=1}^{N} P(\mathcal{C}_k) d(Q_k, Q)$$

with

$$d(Q_k, Q) = \sum_{\boldsymbol{z} \in \mathcal{Z}^n} \left| Q_k(\boldsymbol{z}) - Q(\boldsymbol{z}) \right|.$$

Here $d(Q_k,Q)$ is the $\mathcal{L}_1$ (i.e., variational) distance between the two distributions Q_k and Q. When $d(Q_k,Q)$ is zero, the opponent cannot distinguish between the uniform input distributions on $\mathcal{C}_k$ and $\mathcal{C}$ by observing only the channel output.

Intuitively, if $d_{\text{av}}(f)$ can be made arbitrarily small by appropriate choice of $\mathcal{C}$ and f, the receiver gains no information about the subset $\mathcal{C}_k$ from which the transmitted codeword $\boldsymbol{x}$ comes, given the channel output $\boldsymbol{z}$.

Our main result is the following theorem.

Theorem 1: If the secrecy capacity of the wiretap channel is nonzero, then there exist constants $c > 0$ and $\beta > 0$ so that

$$2^{-H(K)} \leq P_D \leq 2^{-H(K)} + c \exp^{-n\beta}$$

if n is sufficiently large. In particular, if the codeword length n goes to infinity, then $P_I = P_S = 2^{-H(K)}$, and hence, $P_D = \max\{P_I, P_s\} = 2^{-H(K)}$.

Proof: (Outline) To obtain a lower bound, we can consider the situation in which the opponent guesses the value of the key. If the guess is correct, the opponent can invoke any attack and the attack will be successful. The probability that the opponent guesses the value of the key correctly is $2^{-H(K)}$. This provides a lower bound.

To prove the upper bound provided in the theorem, we need to show that the success probability of the opponent's attack with any strategy is upper-bounded by the bound provided in the theorem, if we use the authentication scheme proposed in the current work. To this end, we divide all possible output sequences at the opponent $\boldsymbol{z}$ into two subsets: $\mathcal{O}$ and $\mathcal{O}^c$. If $\boldsymbol{z} \in \mathcal{O}$, $\max_{k \in \mathcal{K}} P(k|\boldsymbol{z})$ is much larger than $1/|\mathcal{K}|$. Hence, if the opponent observes a sequence $\boldsymbol{z} \in \mathcal{O}$, the success probability of a substitution attack will be high. On the other hand, if $\boldsymbol{z} \in \mathcal{O}^c$, $\max_{k \in \mathcal{K}} P(k|\boldsymbol{z})$ is close to $1/|\mathcal{K}|$. Thus, if the opponent observes a sequence $\boldsymbol{z} \in \mathcal{O}^c$, the opponent does not gain any information about the key from the output. We show that if the source uses a code with exponentially small $d_{\text{av}}(f)$, which has shown to exist [22], the probability that the opponent will observe $\boldsymbol{z} \in \mathcal{O}$ is exponentially small. Thus, almost all the sequences $\boldsymbol{z}$ have the property that $\max_{k \in \mathcal{K}} P(k|\boldsymbol{z})$ is close to $1/|\mathcal{K}|$. Simple calculation then shows that P_S is arbitrarily close to $1/|\mathcal{K}|$. For the impersonation attack, the optimal strategy for the opponent is to chooses a codeword at random, and hence P_I is $1/|\mathcal{K}|$.

Please refer to Appendix I for technical details. □

Remark 1: Theorem 1 implies that the opponent is reduced to guessing the key, which essentially means it has been defeated.

IV. Authentication of Multiple Messages

In this section, we consider the situation in which the same key K is used to authenticate a sequence of J messages $M_1, \ldots, M_J$, one per time slot. Here J is finite. The opponent will choose a time slot j in which to initiate either an impersonation attack or a substitution attack. For an impersonation attack at slot j, the opponent sends a message to the receiver before the source sends anything. The opponent will choose the message based on the information it has gained through the last $j-1$ rounds of transmission. The attack will be successful if the opponent's message is accepted as authentic at the receiver. We denote by $P_{I,j}$ the success probability of the impersonation attack at the jth time slot. For a substitution attack at slot j, the opponent will intercept the source's jth packet, modify it, and send the modified packet to the receiver. The opponent can make the modification using the information gathered in the j rounds of transmission that have occurred so far. The attack will be successful if the modified signal is accepted as authentic and the message part is decoded into an incorrect message. We denote by $P_{S,j}$ the success probability of the substitution attack at the jth time slot. Obviously, the opponent will choose the attack that maximizes its cheating probability $P_D = \max\{P_{I,1}, \ldots, P_{I,J}, P_{S,1}, \ldots, P_{S,J}\}$.

The authentication of multiple messages with the same key under the noiseless model has been studied in [16] and [24]–[26]. In these works, to avoid a replay attack, in which the opponent simply resends one of the codewords it has received before, one of the following assumptions is made: 1) the messages in all blocks are distinct (e.g., [24], [25]) or 2) the authentication schemes used in all blocks are distinct (e.g., [26], [26]). Under the second assumption, a lower bound for P_D with the noiseless transmission model was derived in [16], namely

$$P_D \geq 2^{-H(K)/(J+1)}.$$

This bound suggests that after several rounds of authentication, the opponent may be able to obtain almost all the information about the key, and hence, may be able to choose an attack with a high success probability. On the other hand, in the following we show that with a noisy channel model, one can limit the information leaked to the opponent, and thus, the success probability of the opponent will not increase even by observing more packets. In the current work, we do not need to make either of the two assumptions mentioned above (i.e., all messages can be the same, as can all the authentication schemes). The channel noise renders the output at the opponent to be almost independent of the input, and hence the success probability of the replay attack or any other attack is bounded, as we argue next.

We use the same scheme as for the single-message case; that is, the source transmits the message and key using a wiretap channel code. More specifically, the source uses the same codebook discussed in Section III, with $|\mathcal{K}|$ subsets, each corresponding to a key. Also, each subset contains $|\mathcal{M}|$ bins, each corresponding to a message. In block j, if the intended message is m_j, the source randomly chooses a codeword $\boldsymbol{x}_j$ from the m_jth bin in the kth subset using a uniform distribution. The source then transmits $\boldsymbol{x}_j$ over the channel. The opponent receives $\boldsymbol{z}_j$ with probability

$$P(\boldsymbol{z}_j|\boldsymbol{x}_j) = \prod_{t=1}^{n} P(z_j(t)|x_j(t)).$$

The receiver receives $\boldsymbol{y}_j$. If the opponent does not initiate any attack, then

$$P(\boldsymbol{y}_j|\boldsymbol{x}_j) = \prod_{t=1}^{n} P(y_j(t)|x_j(t)).$$

On the other hand, if the opponent chooses to attack, then the value of $\boldsymbol{y}_j$ depends on the opponent's attack strategy. At each time slot, the receiver performs jointly typical set decoding and obtains an estimate $\hat{\boldsymbol{x}}_j$ based only on $\boldsymbol{y}_j$, and then sets m'_j and k' to be the bin index and subset index associated with $\hat{\boldsymbol{x}}_j$. If k' is the same as the key that the receiver knows, then the message is accepted as authentic; otherwise, the message is rejected. As before, we use $(m'_j, k') = g(\boldsymbol{y}_j)$ as the decoding process at the receiver.

To initiate an impersonation attack in block j, the opponent can use the information gained through $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{j-1}$. Let $h_{j,im}$ be the strategy employed by the source that maps $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{j-1}$ to $\boldsymbol{y}_{o,j}$. We also denote by $(m'_j, k') = g(\boldsymbol{y}_{o,j}) = g(h_{j,im}(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{j-1}))$ the decoded message and key at the destination after the opponent's attack. Obviously, for each $(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{j-1})$, the opponent will adapt a strategy $h_{j,im}$ so that the following probability is maximized:

$$\sum_{k \in \mathcal{K}} P(k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{j-1}) \gamma_1(k, k')$$

in which $\gamma_1(k, k')$ is the indicator function defined in Section III. Hence, the success probability of the impersonation attack after receiving $j-1$ rounds of transmission is

$$\begin{aligned} P_{I,j} &= \sum_{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{j-1}} P(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{j-1}) \\ &\quad \times \sup_{h_{j,im}} \left\{ \sum_{k \in \mathcal{K}} P(k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{j-1}) \gamma_1(k, k') \right\} \\ &\leq \sum_{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{j-1}} P(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{j-1}) \max_{k \in \mathcal{K}} \{P(k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{j-1})\}. \end{aligned} \tag{11}$$

The inequality follows from the same reasoning as that used to obtain (9).

The opponent can also choose to invoke a substitution attack after receiving the jth transmission, i.e., it changes the content of the jth package and sends it to the destination. Let $h_{j,sb}$ be the strategy employed by the source that maps $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j$ to $\boldsymbol{y}_{o,j}$. We also denote by $(m'_j, k') = g(\boldsymbol{y}_{o,j}) = g(h_{j,sb}(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j))$ the decoded message and key at the destination after the opponent's attack.

The attack is successful if $m'_j \neq m_j$ and $k' = k$. For each possible observation $(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j)$, the opponent will adopt a strategy $h_{j,sb}$ so that the following probability is maximized:

$$\sum_{m_j, k} P(m_j, k|\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) \gamma_2(m_j, m'_j) \gamma_1(k, k')$$

in which γ_1 and γ_2 are defined as in Section III.

Hence, the success probability of the jth substitution attack $P_{S,j}$ is

$$P_{S,j} = \sum_{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_j} P(\boldsymbol{z}_1,\ldots,\boldsymbol{z}_j) \times \sup_{h_{j,sb}} \left\{ \sum_{m_j,k} P(m_j,k|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_j)\gamma_2(m_j,m_j')\gamma_1(k,k') \right\}.$$

With regard to this quantity, we have the following result.

Lemma 2:

$$P_{S,j} \le \sum_{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_j} P(\boldsymbol{z}_1,\ldots,\boldsymbol{z}_j) \max_{k\in\mathcal{K}} \{P(k|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_j)\}. \quad (12)$$

Proof: The proof follows from the same steps as those used in the proof of Lemma 1. Thus, we omit it here. □

Note that (11) and (12) are valid for any attack strategy of the opponent, including the replay attack mentioned above. Also note that (11) and (12) have similar forms. Hence, in the proof of the following theorem, we derive bounds only for $P_{S,j}$. The bounds for $P_{I,j}$ follow similarly.

Theorem 2: If the secrecy capacity of the wiretap channel is nonzero, then there exist constants $c_m > 0$ and $\beta_1 > 0$ so that

$$2^{-H(K)} \le P_D \le 2^{-H(K)} + c_m \exp^{-n\beta_1}$$

if n is sufficiently large. In particular, if the codeword length n goes to infinity, then $P_D = 2^{-H(K)}$.

Proof: (Outline) We first show that the mutual information between the key and observations at the opponent in the j blocks is exponentially small. Using a result relating divergence and $\mathcal{L}_1$ distance, it is then shown that $d_{\text{av}}(f)$ with a proper definition is exponentially small. We then follow the same steps as that in the proof of Theorem 1 to show the bounds. Please refer to Appendix II for details. □

V. Authentication With Correlated Sequences

In some applications, instead of sharing a key K, the source and the receiver may possess correlated, but not identical, sequences. For example, in biometric authentication systems, where individuals are verified based on their physiological characteristics such as fingerprint, iris, signature, etc., two measurements of a physiological characteristic of the same person are unlikely to be exactly the same, due to measurement noise, injuries, etc. But these two sequences will be correlated.

In this section, we assume that the source and receiver possess length-L correlated sequences $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$, respectively, generated according to a joint distribution $\prod_{l=1}^{L} P(s_1(l), s_2(l))$. We assume that the opponent knows the joint distribution $P(s_1, s_2)$. The marginal distributions of S_1 and S_2 are denoted as $P(s_1)$ and $P(s_2)$, respectively. We will ultimately need L to be sufficiently large. In our authentication scheme, when a source sends a message, it will also send its knowledge of the sequence $\boldsymbol{S}_1$ to the destination to authenticate the message. More specifically, the source first does source coding for sequence $\boldsymbol{S}_1$. That is, the source gives a unique index for any ϵ-strong typical [27]L-sequence, and gives any other nontypical sequence index 0. There are at most $|\mathcal{K}| = 2^{L(H(S_1)+\epsilon)}$ typical sequences, each having probability approximately $2^{-LH(S_1)}$. Then for any message size $|\mathcal{M}|$, the source generates a code for the wiretap channel whose length satisfying the conditions in (4) and (5). As before, the source divides the code into $|\mathcal{K}|$ subset, each corresponding to one strong typical sequence. Also, the source partition each subset into $|\mathcal{M}|$ bins, each corresponding to a message. To send message m with sequence $\boldsymbol{s}_1$, the source randomly chooses a codeword $\boldsymbol{x}$ from the mth bin in the subset corresponding to $\boldsymbol{s}_1$. Upon receiving a signal, the receiver obtains an estimate $\hat{\boldsymbol{x}}$ of the codeword using a typical set decoder. The receiver then obtains an estimate $\hat{\boldsymbol{s}}_1$ of the sequence by choosing it as the sequence corresponding to the subset index of the $\hat{\boldsymbol{x}}$. If $\hat{\boldsymbol{s}}_1$ is strongly jointly typical with $\boldsymbol{s}_2$ with respect to $P(s_1, s_2)$, the receiver accepts the message; otherwise, it rejects the message.

The following theorem shows the bounds for the cheating probability of the opponent in this case.

Theorem 3: Suppose the secrecy capacity of the wiretap channel is nonzero, and there are length-L correlated sequences $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ at the source and receiver, respectively. Then, for any arbitrarily small $\epsilon_0 > 0$, there is an $L_0 > 0$, such that

$$2^{-L(I(S_1;S_2)+\epsilon_0)} \le P_D \le 2^{-L(I(S_1;S_2)-\epsilon_0)}$$

for all $L > L_0$ when n goes to infinity.

Proof: Please refer to Appendix III. □

VI. Conclusion

In this paper, we have laid the foundation for a theory of message authentication over noisy channels. Towards this end, information-theoretic lower and upper bounds on the cheating probability in the single message authentication scenario have been derived. Remarkably, these bounds have been shown to coincide, resulting in a complete characterization of the fundamental limits on authentication over noisy channels. We have also derived the corresponding bounds for the multiple message authentication case and have shown that they match. Interestingly, our results imply that the key information can be used to protect against various attacks simultaneously. We have further shown that, compared with the classical authentication model in which the channel is assumed to be noiseless, the opponent's success probability is largely reduced in both scenarios. We have also extended our study to the message authentication scenario in which the source and receiver only possess correlated sequences. We thus have established the utility of channel noise in message authentication applications.

Exploiting other characteristics of channels, such as multipath fading, to facilitate message authentication is an interesting avenue for further research. Also, developing authentication techniques for the case in which the source–opponent channel is less noisy than the main channel remains an open problem.

APPENDIX A
PROOF OF THEOREM 1

We need the following lemma from [22].

Lemma 3 ([22]): Consider a wiretap channel $\mathcal{X} \to (\mathcal{Y}, \mathcal{Z})$, and choose $\delta > 0$. Suppose $\mathcal{T}_P \subset \mathcal{X}^n$ is a type class with $P(x)$ bounded away from 0, and such that $I(X;Y) > I(X;Z) + 2\delta$. Then, there exist a codebook $\mathcal{C}$ with size $|\mathcal{C}| = 2^{n(I(X;Y)-\delta)}$, drawn from $\mathcal{T}_P$, and equal-size disjoint subsets $\mathcal{C}_1, \ldots, \mathcal{C}_N$ of $\mathcal{C}$ with

$$N \le 2^{n(I(X;Y)-I(X;Z)-2\delta)}$$

such that $\mathcal{C} = \bigcup_{k=1}^{N} \mathcal{C}_k$ is the codeword with exponentially small average probability of error for the main channel $\mathcal{X} \to \mathcal{Y}$. Moreover, the partition function $f : \mathcal{C} \to \{1, \ldots, N\}$ of $\mathcal{C}$ with $f^{-1}(k) = \mathcal{C}_k, k = 1, \ldots, N$ has exponentially small $d_{\mathrm{av}}(f)$ for the distribution $\tilde{P}_C$ defined on $\mathcal{C} \times \mathcal{Z}^n$ by

$$\tilde{P}_C(\boldsymbol{x}, \boldsymbol{z}) = \frac{1}{|\mathcal{C}|} P(\boldsymbol{z}|\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathcal{C},\ \boldsymbol{z} \in \mathcal{Z}^n.$$

Furthermore, $I(N; \boldsymbol{Z})$ is exponentially small.

Proof: Please see [22]. □

With this lemma, we proceed.

1) Lower bound: The opponent can simply guess the value of K, then randomly choose a codeword from $\mathcal{C}_k$ and send the corresponding codeword to the destination. The probability of success is $1/|\mathcal{K}|$. Thus

$$P_D \ge \frac{1}{|\mathcal{K}|} = 2^{-H(K)}.$$

2) Upper bound: We use the following scheme.

Choose $\delta > 0$, and let P_X be a type of $\boldsymbol{X}$ satisfying $I(X;Y) > I(X;Z) + 2\delta$. Denote by $\mathcal{T}_P$ the set of $\boldsymbol{x}$'s having type P_X. Since the source–wiretapper channel is not less noisy than the main channel, such a P_X exists.

Now choose n_1 and n_2 such that

$$|\mathcal{K}| \le 2^{n_1(I(X;Y)-I(X;Z)-2\delta)}$$

and

$$|\mathcal{K}||\mathcal{M}| \le 2^{n_2(I(X;Y)-\delta)}$$

and then choose $n > \max\{n_1, n_2\}$. (n also needs to satisfy other conditions specified later.) Let $\mathcal{C}$ and $\mathcal{C}_k$, $k = 1, \ldots, |\mathcal{K}|$ be the codebook and corresponding partition satisfying the conditions of Lemma 3. That is, for this $\mathcal{C}$ and f, there is an $\alpha > 0$ such that

$$d_{\mathrm{av}}(f) \le \epsilon = \exp\{-n\alpha\}. \tag{13}$$

When the key is k, the transmitted codeword comes from $\mathcal{C}_k$. The receiver will accept any signal $\hat{\boldsymbol{y}}$ that can be decoded into a codeword belonging to the subset corresponding to k. It is easy to see that $P_I = 1/|\mathcal{K}|$.

Based on (8), for an impersonation attack, after receiving $\boldsymbol{z}$, the optimal strategy for the attacker is to choose a codeword $\hat{\boldsymbol{y}}$ from $\mathcal{C}_k$, where k maximizes $P(k|\boldsymbol{z})$.

Let us rewrite $d_{\mathrm{av}}(f)$ as follows:

$$\begin{aligned} d_{\mathrm{av}}(f) &= \sum_{k=1}^{|\mathcal{K}|} \sum_{\boldsymbol{z} \in \mathcal{Z}^n} \left| P(\mathcal{C}_k) Q_k(\boldsymbol{z}) - P(\mathcal{C}_k) Q(\boldsymbol{z}) \right| \\ &= \sum_{\boldsymbol{z} \in \mathcal{Z}^n} Q(\boldsymbol{z}) d(\boldsymbol{z}) \end{aligned}$$

with

$$d(\boldsymbol{z}) = \sum_{k=1}^{|\mathcal{K}|} \left| P(\mathcal{C}_k|\boldsymbol{z}) - P(\mathcal{C}_k) \right|.$$

Notice that in our scheme $P(\mathcal{C}_k) = 1/|\mathcal{K}|$ and $P(\mathcal{C}_k|\boldsymbol{z})$ is the conditional probability that the key value is k after observing $\boldsymbol{z}$. Hence

$$d(\boldsymbol{z}) = \sum_{k \in \mathcal{K}} \left| P(k|\boldsymbol{z}) - \frac{1}{|\mathcal{K}|} \right|.$$

On defining

$$\mathcal{O} = \{\boldsymbol{z} : d(\boldsymbol{z}) > \epsilon^{1/2}\}$$

we have

$$d_{\mathrm{av}}(f) = \sum_{\boldsymbol{z} \in \mathcal{O}} Q(\boldsymbol{z}) d(\boldsymbol{z}) + \sum_{\boldsymbol{z} \in \mathcal{O}^c} Q(\boldsymbol{z}) d(\boldsymbol{z}).$$

Here $\mathcal{O}$ is the set of $\boldsymbol{z}$ such that the $\mathcal{L}_1$ distance between the conditional distribution $P(k|\boldsymbol{z})$ and the uniform distribution is large. Hence, if the opponent observes $\boldsymbol{z}$, it may be able to guess the value of the key correctly with high probability. But

$$\epsilon^{1/2} \sum_{\boldsymbol{z} \in \mathcal{O}} Q(\boldsymbol{z}) \le \sum_{\boldsymbol{z} \in \mathcal{O}} Q(\boldsymbol{z}) d(\boldsymbol{z}) \le d_{\mathrm{av}}(f) \le \epsilon$$

and thus

$$Q(\mathcal{O}) \le \epsilon^{1/2}.$$

For those $\boldsymbol{z} \in \mathcal{O}^c$, we have

$$\max_{k \in \mathcal{K}} \left\{ P(k|\boldsymbol{z}) - \frac{1}{|\mathcal{K}|} \right\} \le \sum_{k \in \mathcal{K}} \left| P(k|\boldsymbol{z}) - \frac{1}{|\mathcal{K}|} \right| = d(\boldsymbol{z}) \le \epsilon^{1/2}.$$

Thus, for $\boldsymbol{z} \in \mathcal{O}^c$, we have

$$\max_{k \in \mathcal{K}} P(k|\boldsymbol{z}) \le \epsilon^{1/2} + 1/|\mathcal{K}|. \tag{14}$$

It follows from Lemma 1 that

$$\begin{aligned} P_S &\le \sum_{\boldsymbol{z}} P(\boldsymbol{z}) \max_{k \in \mathcal{K}} \{P(k|\boldsymbol{z})\} \\ &= \sum_{\boldsymbol{z} \in \mathcal{O}} P(\boldsymbol{z}) \max_{k \in \mathcal{K}} \{P(k|\boldsymbol{z})\} + \sum_{\boldsymbol{z} \in \mathcal{O}^c} P(\boldsymbol{z}) \max_{k \in \mathcal{K}} \{P(k|\boldsymbol{z})\} \\ &\overset{(a)}{\le} Q(\mathcal{O}) + Q(\mathcal{O}^c) \left(\epsilon^{1/2} + 1/|\mathcal{K}| \right) \\ &\le \epsilon^{1/2} + \epsilon^{1/2} + 1/|\mathcal{K}| \\ &\overset{(b)}{\le} 1/|\mathcal{K}| + 2\exp^{-n\alpha/2}. \end{aligned} \tag{15}$$

Here (a) is due to the fact that $\max_{k \in \mathcal{K}}\{P(k|\boldsymbol{z})\} \le 1$ and (14), while (b) is due to (13).

On setting $\beta = \alpha/2$ and $c = 2$, we obtain the desired upper bound.

By increasing n, P_S can be made arbitrarily close to $1/|\mathcal{K}|$, and hence

$$P_S = 1/|\mathcal{K}| = 2^{-H(K)}$$

as $n \to \infty$.

Using the union bound, the error probability at the receiver can be bounded as follows:

$$\begin{aligned} P_e &\le \Pr\{k \neq k' | \text{no attack}\} + \Pr\{m \neq m' | \text{no attack}\} \\ &\le 2\Pr\{\hat{\boldsymbol{x}} \neq \boldsymbol{x} | \text{no attack}\} \end{aligned}$$

in which the first term corresponds to the miss probability, while the second term corresponds to decoding error of the message part. For any message size $|\mathcal{M}|$ and key size $|\mathcal{K}|$, as the codeword length n increases, the probability that $\hat{\boldsymbol{x}} \neq \boldsymbol{x}$ goes to zero, as guaranteed by the code construction.

Appendix B
Proof of Theorem 2

1) Lower bound: For each attack, the opponent can simply ignore available information and guess a value of K, and send a codeword corresponding to this key value for which the probability of success is

$$1/|\mathcal{K}| = 2^{-H(K)}.$$

Thus

$$P_D \ge 2^{-H(K)}.$$

2) Upper bound: As discussed in Section IV, we use the same scheme as for the single-message case. After the jth packet, the opponent gains an amount $I(\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_j; K)$ of information about the key K. From Lemma 3, we know that $I(K; \boldsymbol{Z}_j) \le \epsilon$, where $\epsilon = \exp\{-n\alpha\}$ with $\alpha > 0$. We also have that, given K, the $\boldsymbol{Z}_j$'s are conditionally independent of each other, and thus

$$\boldsymbol{Z}_b \to K \to (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{b-1})$$

forms a Markov chain. Hence

$$I(\boldsymbol{Z}_b; K | \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{b-1}) \le I(\boldsymbol{Z}_b; K).$$

Thus

$$\begin{aligned} I(\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_j; K) &= I(\boldsymbol{Z}_1; K) + \sum_{b=2}^{j} I(\boldsymbol{Z}_b; K | \boldsymbol{Z}_1 \ldots, \boldsymbol{Z}_{b-1}) \\ &\le \sum_{b=1}^{j} I(K; \boldsymbol{Z}_b) \end{aligned} \tag{16}$$

and we therefore have

$$I(K; \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_j) \le j\epsilon. \tag{17}$$

Let $\mathcal{C}$ be the codebook used by the source. Let us further represent the random partition $\{\mathcal{C}_1, \ldots, \mathcal{C}_{|\mathcal{K}|}\}$ of $\mathcal{C}$ by a mapping $f\colon \mathcal{C} \to \{\mathcal{C}_1, \ldots, \mathcal{C}_{|\mathcal{K}|}\}$. Denote by $\tilde{P}(k, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_j)$ the joint distribution on

$$\mathcal{K} \times \boldsymbol{Z}_1 \times \cdots \times \boldsymbol{Z}_j$$

where each $\boldsymbol{Z}_j$ ranges over $\mathcal{Z}^n$, by $Q(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j)$ the marginal distribution of $(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j)$ when the source uses uniform distribution on $\mathcal{K}$, and chooses input codeword uniformly from[5] $\mathcal{C}_k \times \cdots \times \mathcal{C}_k$, and by

$$P(k | \boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) = \tilde{P}(k, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_j)/Q(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j)$$

the conditional distribution of k given the output $(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j)$ at the opponent.

Let us denote by $Q_k(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j)$ the conditional distribution on $\boldsymbol{Z}_1 \times \cdots \times \boldsymbol{Z}_j$ when the key is k. That $Q_k(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j)$ is the distribution of the output at the opponent when the input is uniformly chosen from $\mathcal{C}_k \times \cdots \times \mathcal{C}_k$, and thus

$$Q_k(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) = |\mathcal{C}_k|^{-j} \sum_{(\boldsymbol{x}_1 \times \cdots \times \boldsymbol{x}_j) \in \mathcal{C}_k \times \cdots \times \mathcal{C}_k} \prod_{b=1}^{j} P(\boldsymbol{z}_b | \boldsymbol{x}_b).$$

Now, define

$$d_{\text{av}}(f) = \sum_{k=1}^{|\mathcal{K}|} P(\mathcal{C}_k) d(Q_k, Q),$$

where

$$d(Q_k, Q) = \sum_{(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) \in \boldsymbol{Z}_1 \times \cdots \times \boldsymbol{Z}_j} \left| Q_k(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) - Q(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) \right|.$$

From Pinsker's inequality [28], we have

$$d_{\text{av}}^2(f) \le I(K; \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_j) 2 \log e$$

and hence, using (17), we have

$$d_{\text{av}}^2(f) \le j\epsilon 2 \log e.$$

We write $d_{\text{av}}(f)$ as follows:

$$\begin{aligned} d_{\text{av}}(f) &= \sum_{k=1}^{|\mathcal{K}|} \sum_{(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) \in \boldsymbol{Z}_1 \times \cdots \times \boldsymbol{Z}_j} \\ &\qquad |P(\mathcal{C}_k) Q_k(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) - P(\mathcal{C}_k) Q(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j)| \\ &= \sum Q(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) d(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j), \end{aligned}$$

with

$$d(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) = \sum_{k=1}^{|\mathcal{K}|} \left| P(\mathcal{C}_k | \boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) - P(\mathcal{C}_k) \right|.$$

On defining

$$\mathcal{O} = \{(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) : d(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) > \epsilon^{1/3}\}$$

we have

$$\begin{aligned} d_{\text{av}}(f) = &\sum_{(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) \in \mathcal{O}} Q(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) d(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) \\ &+ \sum_{(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) \in \mathcal{O}^c} Q(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j) d(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_j). \end{aligned}$$

[5]Note that in the transmission, we use the same key for each transmission. Hence, if the key is k, then the transmitted codewords $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j)$ in the j blocks are uniformly distributed on $\mathcal{C}_k \times \cdots \times \mathcal{C}_k$.

It follows that

$$\epsilon^{1/3} \sum_{(\boldsymbol{z}_1,\ldots,\boldsymbol{z}_j)\in\mathcal{O}} Q(\boldsymbol{z}_1,\ldots,\boldsymbol{z}_j) \le d_{\text{av}}(f) \le c_j\epsilon^{1/2}$$

and thus $Q(\mathcal{O}) \le c_j\epsilon^{1/6}$, with $c_j = \sqrt{j2\log e}$.

For any $(\boldsymbol{z}_1,\ldots,\boldsymbol{z}_j) \in \mathcal{O}^c$, we have

$$\max_{k\in\mathcal{K}} \left\{P(k|\boldsymbol{z}_1,\ldots,\boldsymbol{z}_j)) - 1/|\mathcal{K}|\right\} \le d(\boldsymbol{z}_1,\ldots,\boldsymbol{z}_j) \le \epsilon^{1/3}.$$

Thus, on following the same steps as those of (15), we have

$$\begin{aligned} P_{S,j} &\le Q(\mathcal{O}) + \epsilon^{1/3} + 1/|\mathcal{K}| \\ &\le c_j\epsilon^{1/6} + \epsilon^{1/3} + 1/|\mathcal{K}| \\ &\le 1/|\mathcal{K}| + 2c_j \exp^{-n\alpha/6}. \end{aligned}$$

On setting $\beta_1 = \alpha/6$ and $c_m = c_J$, we obtain the desired upper bound. As the length of the code n can be sufficiently long, $P_{S,j}$ is arbitrarily close to $1/|\mathcal{K}|$, hence

$$P_{S,j} = 1/|\mathcal{K}| = 2^{-H(K)}$$

as $n \to \infty$.

Finally, following the same steps as above, we also have $P_{I,j} = 1/|\mathcal{K}| = 2^{-H(K)}$, as the codeword length n goes to infinity.

Similar to the single-message authentication part, the error probability of the message m_j can be bounded as the sum of the probability of a miss and the probability of error in decoding the message. These two terms approach zero as the length of the codeword increases.

Appendix C
Proof of Theorem 3

The proof follows closely that of Theorem 1. In the following, we outline the key steps. For a given $\epsilon > 0$, we have the following.

1) Lower bound:

For the impersonation attack, the opponent can randomly choose one strongly typical sequence $\boldsymbol{s}_1$, and transmit the codeword corresponding to the index of this sequence. The probability that this particular $\boldsymbol{s}_1$ will be strongly jointly typical with $\boldsymbol{s}_2$ with respect to $P(s_1, s_2)$ is larger than [28] $2^{-L(I(S_1;S_2)+\epsilon_1)}$, in which ϵ_1 can be made arbitrarily small by setting $\epsilon \to 0$ and sufficiently large L.

2) Upper bound:

The error probability when the signal truly comes from the source can be bounded as

$$\begin{aligned} P_e \le \Pr\{(\boldsymbol{s}_1, \boldsymbol{s}_2) \text{ are not strongly jointly typical}\} \\ + \Pr\{\boldsymbol{s}_1 \ne \hat{\boldsymbol{s}}_1\} + \Pr\{m \ne m'\}. \end{aligned}$$

It is well known that, for any $\epsilon_e > 0$, there exists a positive number L_2 such that if $L \ge L_2$, then the probability that $(\boldsymbol{s}_1, \boldsymbol{s}_2)$ are not strongly jointly typical is less than ϵ_e. The code for the wiretap channel guarantees that $\Pr\{\boldsymbol{s}_1 \ne \hat{\boldsymbol{s}}_1\} + \Pr\{m \ne m'\} \le 2\epsilon_e$, if n is sufficiently large. Hence, the decoding error at the destination is arbitrarily small, if L and n are sufficiently large.

Now for the impersonation attack, the optimal strategy for the opponent is to send index i, such that the probability that $\boldsymbol{s}_1'$, which is the strongly typical sequence associated with index i, and $\boldsymbol{s}_2$ are strongly jointly typical with respect to $P(s_1, s_2)$ is maximized. Since $\boldsymbol{s}_1'$ and $\boldsymbol{s}_2$ are drawn independently, the success probability of this attack is bounded as [27], [28]

$$P_I \le 2^{-L(I(S_1;S_2)-\epsilon_2)}$$

in which ϵ_2 is arbitrarily small if $\epsilon \to 0$ and L is sufficiently large.

For the substitution attack, following the same steps as that of Lemma 1 and using properties of strongly jointly typical sequences [27], [28], we have

$$P_S \le 2^{L(H(S_1|S_2)+\epsilon_2)} \sum_{\boldsymbol{z}} P(\boldsymbol{z}) \max_{\boldsymbol{s}_1}\{P(\boldsymbol{s}_1|\boldsymbol{z})\}. \tag{18}$$

Following the same arguments as in Theorem 1, we can show that there exist constants $c_c > 0\, \beta_c > 0$ and a partition $\mathcal{O}$ of the output sequences $\boldsymbol{z}$ at the opponent such that

$$P\{\boldsymbol{z} \in \mathcal{O}\} \le c_c \exp^{-n\beta_c}$$

and for any $\boldsymbol{z} \in \mathcal{O}^c$, we have

$$\max_{\boldsymbol{s}_1}\{P(\boldsymbol{s}_1|\boldsymbol{z})\} \le 2^{-L(H(S_1)-\epsilon)} + c_c \exp^{-n\beta_c}.$$

Proceeding from (18), we have

$$\begin{aligned} P_S &\le 2^{L(H(S_1|S_2)+\epsilon_2)} P\{\boldsymbol{z} \in \mathcal{O}\} \\ &\quad + 2^{L(H(S_1|S_2)+\epsilon_2)} \sum_{\boldsymbol{z}\in\mathcal{O}^c} P(\boldsymbol{z}) \max_{\boldsymbol{s}_1}\{P(\boldsymbol{s}_1|\boldsymbol{z})\} \\ &\le 2^{-L(I(S_1;S_2)-\epsilon_2-\epsilon)} + 2c_c \exp^{-n\beta_c}\, 2^{L(H(S_1|S_2)+\epsilon_2)}. \end{aligned}$$

Thus

$$\begin{aligned} P_D &= \max\{P_I, P_S\} \\ &\le 2^{-L(I(S_1;S_2)-\epsilon_0)} + 2c_c \exp^{-n\beta_c}\, 2^{L(H(S_1|S_2)+\epsilon_2)} \end{aligned}$$

in which $\epsilon_0 = \max\{\epsilon_1, \epsilon_2+\epsilon\}$. By properly choosing ϵ and L, ϵ_0 can be made to be arbitrarily small. For any ϵ_0 and L satisfying conditions specified above, this upper bound is arbitrarily close to $2^{-L(I(S_1;S_2)-\epsilon_0)}$, as n goes to infinity.

This completes our proof.

References

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, pp. 656–715, Oct. 1949.

[2] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355–1387, Oct. 1975.

[3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[4] L. Lai and H. El Gamal, "The relay-eavesdropper channel: Cooperation for secrecy," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 4005–4019, Sep. 2008.

[5] L. Lai, H. El Gamal, and H. V. Poor, "The wiretap channel with feedback: Encryption over the channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5059–5067, Nov. 2008.

[6] P. K. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.

[7] Y. Liang and H. V. Poor, "Multiple access channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 976–1002, Mar. 2008.
[8] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, Jun. 2008.
[9] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2515–2534, Jun. 2008.
[10] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, submitted for publication.
[11] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, submitted for publication.
[12] H. Imai, U. Maurer, and Y. Zheng, "Introduction to special issue on information theoretic security," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2405–2407, Jun. 2008.
[13] G. J. Simmons, "Authentication theory/coding theory," in *Proc. CRYPTO 84 on Advances in Cryptology*. New York: Springer-Verlag, 1985, pp. 411–431.
[14] T. Johansson, "Lower bounds on the probability of deception in authentication with arbitration," *IEEE Trans. Inf. Theory*, vol. 40, no. 5, pp. 1573–1585, Sep. 1994.
[15] Y. Desmedt and M. Yung, "Arbitrated unconditionally secure authentication can be unconditionally protected against arbiter's attacks," in *Proc. Int. Cryptology Conf.*, Santa Barbara, CA, 1990, pp. 177–188.
[16] U. M. Maurer, "Authentication theory and hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 1350–1356, Jul. 2000.
[17] R. Johannesson and A. Sgarro, "Strengthening Simmons' bound on impersonation," *IEEE Trans. Inf. Theory*, vol. 37, no. 4, pp. 1182–1185, Jul. 1991.
[18] Y. Liu and C. G. Boncelet, "The CRC-NTMAC for noisy message authentication," *IEEE Trans. Inf. Forensics and Security*, vol. 1, no. 4, pp. 517–523, Dec. 2006.
[19] C. G. Boncelet, "The NTMAC for authentication of noisy messages," *IEEE Trans. Inf. Forensics and Security*, vol. 1, no. 1, pp. 35–42, Mar. 2006.
[20] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer, "Generalized privacy amplification," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1915–1923, Nov. 1995.
[21] U. M. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," in *Proc. Advances in Cryptology-EUROCRYPT*, Bruges (Brugge), Belgium, May 2000, pp. 356–373.
[22] I. Csiszár, "Almost independence and secrecy capacity," *Probl. Inf. Transm.*, vol. 32, pp. 40–47, Jan. 1996.
[23] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography, Part II: CR capacity," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 225–240, Jan. 1998.
[24] V. Fak, "Repeated use of codes which detect deception," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 4, pp. 233–234, Mar. 1979.
[25] U. Rosenbaum, "A lower bound on authentication after having observed a sequence of messages," *J. Cryptol.*, vol. 6, pp. 135–156, Mar. 1993.
[26] B. Smeets, "Bounds on the probability of deception in multiple authentication," *IEEE Trans. Inf. Theory*, vol. 40, no. 5, pp. 1586–1591, Sep. 1994.
[27] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.
[28] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

Lifeng Lai (M'07) received the B.E. and M. E. degrees in information science and electrical engineering from Zhejiang University, Hangzhou, China, in 2001 and 2004, respectively, and the PhD degree in electrical and computer engineering from The Ohio State University, Columbus, in 2007.

He is currently working at Princeton University, Princeton, NJ, as a Postdoctoral Research Associate. His current research interest includes network information theory, wireless networks security, and sequential analysis of wireless networks.

Dr. Lai was a Distinguished University Fellow of The Ohio State University from 2004 to 2007. He coauthored a paper that won the best paper award from IEEE Global Telecommunications Conference (GLOBECOM), 2008.

Hesham El Gamal (M'99–SM'03) received the B.S. and M.S. degrees in electrical engineering from Cairo University, Cairo, Egypt, in 1993 and 1996, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Maryland, College Park, MD, in 1999.

From 1993 to 1996, he served as a Project Manager in the Middle East Regional Office of Alcatel Telecom. From 1996 to 1999, he was a Research Assistant in the Department of Electrical and Computer Engineering, the University of Maryland at College Park. From February 1999 to December 2000, he was with the Advanced Development Group, Hughes Network Systems (HNS), Germantown, MD, as a Senior Member of the Technical Staff. In the Fall of 1999, he served as a Lecturer at the University of Maryland at College Park. In January 2001. he joined the Electrical and Computer Engineering Department at The Ohio State University (OSU), Columbus, where he is now an Associate Professor. He held visiting appointments at UCLA (Fall 2002, Winter 2003) and Institut Eurecom (Summer 2003). He also serves as the Director for the Wireless Intelligent Networks Center (WINC), Nile University, Cairo, Egypt. He holds five U.S. patents and has eight more patent applications pending.

Prof. El Gamal is a recipient of the HNS Annual Achievement Award (2000), the OSU College of Engineering Lumley Research Award (2003), the OSU Electrical Engineering Department FARMER Young Faculty Development Fund (2003–2008), and the National Science Foundation CAREER Award (2004). He currently serves as an Associate Editor for Space-Time Coding and Spread Spectrum for the IEEE TRANSACTIONS ON COMMUNICATIONS.

H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, in 1977.

From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton University, where he is the Dean of Engineering and Applied Science, and the Michael Henry Strater University Professor of Electrical Engineering. His research interests are in the areas of stochastic analysis, statistical signal processing and their applications in wireless networks and related fields. Among his publications in these areas are the recent books *MIMO Wireless Communications* (Cambridge University Press, 2007), coauthored with Ezio Biglieri *et al*., and *Quickest Detection* (Cambridge University Press, 2009), coauthored with Olympia Hadjiliadis.

Dr. Poor is a member of the National Academy of Engineering, a Fellow of the American Academy of Arts and Sciences, and a former Guggenheim Fellow. He is also a Fellow of the Institute of Mathematical Statistics, the Optical Society of America, and other organizations. In 1990, he served as President of the IEEE Information Theory Society, and in 2004–2007 as the Editor-in-Chief of these TRANSACTIONS. He is the recipient of the 2005 IEEE Education Medal. Recent recognition of his work includes the 2007 IEEE Marconi Prize Paper Award of the IEEE Communications Society, the 2007 Technical Achievement Award of the IEEE Signal Processing Society, and the 2008 Aaron D. Wyner Distinguished Service Award of the IEEE Information Theory Society.