One-Shot Information Hiding and Compound Wiretap Channels

Yanxiao Liu, *Member, IEEE* and Cheuk Ting Li, *Member, IEEE*

***Abstract*—In recent years, due to the growing reliance on large amounts of data that are communicated, analyzed, and utilized which inherently contain sensitive and personal information, secrecy and privacy in communication become increasingly important. We present nonasymptotic information-theoretic analyses of two fundamental secrecy problems under channel uncertainties: the information hiding problem and the compound wiretap channel. The former admits a game-theoretic formulation, where one party (an information hider and a decoder) seeks to embed secret messages into a host signal for later reconstruction, while the opposing party (an attacker) attempts to remove or degrade the embedded information. The latter generalizes Wyner's wiretap channel by allowing multiple potential channel states. The information hiding problem concerns *active* attacks during data transmission, while the compound wiretap channel addresses *passive* eavesdropping and information leakage. The two problems, both of which consider channels with uncertainties, can be studied under a unified framework by utilizing a covering argument and the Poisson matching lemma. We derive novel one-shot achievability results for both problems that are applicable to any source distribution and any class of channels (not necessarily memoryless or ergodic), and that apply to both discrete and continuous cases. We also show that existing asymptotic results can be recovered by applying our results to discrete memoryless channels.**

***Index Terms*—One-shot achievability, finite blocklength analysis, information hiding, watermarking, wiretap channels.**

I. Introduction

In data science and wireless communication, secrecy and privacy are becoming increasingly important due to the growing dependence on large amounts of data being communicated, analyzed, and utilized, which inherently contain sensitive and personal information and, therefore, should be well protected from leakage. Over the past few decades, the fundamental information-theoretic limits of various secrecy and privacy problems have been extensively studied. During information transmission, there are two types of concerns regarding information leakage: *active* attacks and *passive* wiretapping. In this paper, we consider two fundamental problems in information theory that address these two concerns, respectively: the information hiding problem [1] and the compound wiretap channel [2].

For the information hiding problem [1], it can be formulated as a communication system from a game-theoretic perspective, where an encoder-decoder team seeks to transmit a confidential message *embedded* in a host data source, while the opposing side is an attacker, modeled as a noisy channel, attempts to destroy or degrade the message. The information-theoretic limits of different variations of information hiding have been extensively investigated over the past two decades [1], [3]–[5], due to its wide range of applications, including watermarking, fingerprinting, steganography, and copyright protection. Existing analyses of information hiding problems borrow techniques from various fields, including wireless communication, signal processing, cryptography, and game theory. For the compound wiretap channel, [2] modeled the problem as a generalization of Wyner's conventional wiretap channel [6], where the communication channel can take multiple potential states. The objective is to ensure reliable transmission and minimize information leakage regardless of which state occurs. This model is more general and better suited to practical wireless communication scenarios where the transmitter may not have knowledge of the channel conditions or where channel characteristics change rapidly, yet communication performance and security must still be guaranteed.

In all existing studies on these two problems, the information-theoretic limits have been analyzed in the *asymptotic* regime, assuming that the signal has a blocklength approaching infinity. However, this assumption does not hold in practice, as packets have bounded lengths, which can be quite short in certain applications [7]. Over the past decade, *finite-blocklength information theory* has been extensively studied [8]–[11], leading to the derivation of nonasymptotic limits, where the law of large numbers does not apply and conventional typicality-based tools are inapplicable. More generally, we are interested in *one-shot* achievability results, where the channel or source is arbitrary and used only *once*. Various one-shot coding techniques have been proposed [12]–[14], yielding one-shot bounds that recover existing (first-order and second-order) asymptotic results when applied to memoryless sources or channels. See Section II-C for a review.

In this paper, which extends the conference version [15],[1] we study the one-shot achievability results of the information

This work was partially supported by two grants from the Research Grants Council of the Hong Kong Special Administrative Region, China [Project No.s: CUHK 24205621 (ECS), CUHK 14209823 (GRF)].

This paper was presented in part at the 2024 IEEE Information Theory Workshop (ITW 2024).

Yanxiao Liu and Cheuk Ting Li are with the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong SAR of China. Email: yanxiaoliu@link.cuhk.edu.hk, ctli@ie.cuhk.edu.hk.

[1] The conference version [15] studies the information hiding setting only. In comparison, this journal version not only provides more detailed discussions and generalizations of the information hiding problem, but also extends the framework to the compound wiretap channel and derives novel one-shot achievability results, as well as the recovery of the asymptotic achievable rates, which were not covered in [15].

hiding problem and the compound wiretap channel. Compared to the asymptotic information hiding [1], we remove the assumption that the decoder knows the attacker's strategy, considering instead that the only available knowledge is that the attack channel belongs to a set (which may have infinite cardinality). For both problems, our goal is to provide *distributionally robust* coding strategies (see [16]) due to the channel uncertainty, and we hence study them under a unified framework. We briefly summarize our contributions as follows.

- We derive novel one-shot achievability results for the information hiding problem [1] and the compound wiretap channel [2], under the same framework.
- For both problems, most of the existing asymptotic analyses assume the decoder knows the channels.[2] We derive one-shot results without such an assumption, by utilizing the Poisson matching lemma [14] and a covering argument [17].
- We show the one-shot results recover existing asymptotic bounds when applied to memoryless channels (possibly subject to distortion constraints), hence providing alternative proofs that can be simpler.
- Additionally, unlike [1], [2], our results apply to both continuous and discrete cases.

The paper is organized as follows. We begin with a literature review on information hiding, watermarking, compound wiretap channels, and one-shot information theory in Section II. Next, we present the one-shot information hiding problem in Section IV and recover existing asymptotic results in Section V. Within the same framework, we study the one-shot compound wiretap channel in Section VI. Finally, in Section VII, we discuss how our analysis on information hiding can benefit the design of better AI-based watermarking tools, which are becoming increasingly important in the era of generative models.

II. Related Literature

A. Information Hiding

The information hiding problem has been studied since [1], [3]–[5], due to its wide range of applications, including watermarking, fingerprinting, audio/image/video processing, copyright protection, and steganography. The goal is to *hide* a message into some host signal (by introducing a certain level of distortion), so that the message can be correctly reconstructed after suffering *attacks* (which introduce another level of distortion). This problem was modeled as a communication problem, and asymptotic information-theoretic capacity was derived in [1]. In general, information hiding is closely related to the Gelfand–Pinsker problem [18], [19], and various extensions have been studied, e.g., the case where side information is available to the encoder, decoder, and adversary [20], and the case where the decoder has rate-limited side information [21]. See [22] for its duality with the Wyner–Ziv problem, and [23] for a comprehensive survey. We discuss its applications and related settings with different objectives as follows.

*1) **Watermarking, Fingerprinting, and Steganography**:* The setting in [1] can be viewed as *public watermarking* [4], where the host signal is available only at the encoder. In contrast, when it is also available at the decoder, *private watermarking* has been studied in [3], [24]. In the Gaussian case, public and private watermarking have the same capacity [5], but this is not true in general. Watermarking problems consider messages containing personal identification information to be protected from attacks, but secrecy is not always required. In comparison, *digital fingerprinting* [1], [25] embeds fingerprints into the host data to uniquely identify users for tracing illegal data usage, which can be more challenging due to potential collusion. A provably good data embedding strategy was introduced by [26]. Random coding error exponents have been investigated in these problems [20], [27], [28]. In [1, Sec. VII.C] it was indicated that information hiding was applicable to steganography, and this was later studied by [29] for the capacity of perfectly secure steganographic systems. See [30] for trellis codes and [31] for polar codes, which are used in steganographic code designs.

*2) **Host, Stegotext, and Reversibility**:* In the conventional information hiding [1], the message is embedded into host data by producing an encoded signal ("stegotext"), with the goal of recovering the message only. Other objectives have been considered later, such as conveying the host [32] or reconstructing the stegotext [33], [34]. *Reversible* information embedding has also been investigated [35]–[37], where the host signal needs to be decoded. However, this can incur a high cost when the host has high entropy [35], making perfect reversibility even impossible for continuous host signals [37]. Nevertheless, in practice, the goal is often to enable retransmission of the stegotext, and codes for stegotext recovery have been studied [33], [34]. For this setting, single-letter capacity-distortion tradeoffs are known only for logarithmic distortion [32] and quadratic distortion in the Gaussian case [38]. We discuss these generalizations on our setting in Section IV-C.

B. Compound Wiretap Channels

Compound wiretap channels [2] generalize the conventional wiretap channel model [6] by allowing both the legitimate channel and the eavesdropper's channel to have multiple possible states. The objective is to guarantee reliable and secure signal transmission regardless of which state occurs. This is a practical model for channel uncertainty, where the transmitter may have no knowledge of the channel (due to the dynamic nature of the wireless medium or unavoidable implementation/estimation inaccuracies), but zero performance outage is still required (e.g., for ultra-reliable communications [7]). [2] proposed achievable and converse results, with the converse bounds shown to be tight in certain cases by [39]. They also studied the achievable secrecy degrees of freedom (s.d.o.f.) region for a multi-input multi-output (MIMO) model, which was later extended to the case of two confidential messages in [40]. The s.d.o.f. of compound wiretap parallel channels

[2] This assumption can be reasonable when the bolcklength is large, but should be dropped in the one-shot scenarios. See Section IV-B for discussions.

were also studied in [41]. See [42], [43] for discussions on Gaussian MIMO compound wiretap channels.

In [2], [39], the results focused on discrete memoryless channels with a countably finite uncertainty set (i.e., the set from which the exact channel realization is drawn). This was later extended to *arbitrary uncertainty sets*, including continuous alphabets, by [44], which is also one of the contributions of our paper. Moreover, [45] showed that the secrecy capacity is a continuous function of the uncertainty set.

C. One-shot Information Theory

For all the literature discussed in this section so far, the information-theoretic bounds are investigated *asymptotically* in the large blocklength limit, based on the law of large numbers. However, this assumption is impractical, as packets have bounded lengths, which can be very short in low-latency communications [7]. Over the past decade, *finite blocklength* [8]–[10] and *one-shot* [11]–[14], [46], [47] information theory have been widely studied, leading to nonasymptotic results. In the one-shot setting, we assume the channel or source is used only *once* (i.e., it need not be memoryless or ergodic), and the blocklength is 1. These results are expected to recover existing (first-order and second-order) asymptotic bounds when applied to memoryless channels or sources (e.g., asymptotic channel coding capacity by Shannon [48] can be implied by the one-shot bounds of Feinstein [49], Shannon [50], or [10], [11]).

In this paper, we consider the one-shot scenario and derive achievability results using the *Poisson matching lemma* [14], which has been shown to improve upon previously known one-shot bounds in various settings with simpler analyses [14], with recent applications including hypothesis testing [51], secret key generation [52], unequal message protection [53] and oblivious relaying [54]. It is based on the Poisson functional representation [55], which has also recently been applied to various fields in combination with other techniques, such as neural estimation [56] and differential privacy [57]. A refined version of the Poisson matching lemma has been used to analyze general (multi-hop) noisy networks in [58].

Notations

We assume logarithm and entropy are to the base 2. For a statement S, we use $\mathbf{1}\{S\}$ to denote its indicator, i.e., $\mathbf{1}\{S\} = 1$ if S holds, and otherwise $\mathbf{1}\{S\} = 0$. We use δ_a to denote the degenerate distribution $\mathbf{P}\{X = a\} = 1$. For two random variables X, Y, the information density is defined as $\iota_{X;Y}(x; y) = \log((\mathrm{d}P_{X|Y}(\cdot|y)/\mathrm{d}P_X)(x))$, where $\mathrm{d}P_{X|Y}(\cdot|y)/\mathrm{d}P_X$ denotes the Radon-Nikodym derivative. We sometimes omit the subscript and write $\iota(x; y)$ if the random variables are clear from the context. The total variation (TV) distance between two distributions P, Q over $\mathcal{X}$ is $\|P - Q\|_{\mathrm{TV}} := \sup_{A\subseteq\mathcal{X} \text{ measurable}} |P(A) - Q(A)|$. The set of positive integers is denoted by $\mathbb{N}_+$. We write $[i] := \{1, \ldots, i\}$ for $i \in \mathbb{N}_+$. We write $P \ll Q$ to denote that P is absolutely continuous with respect to Q.

III. Poisson Matching Lemma

In this section, we introduce our main technique, called the Poisson matching lemma [14]. The techniques in [1], [4] (e.g., the tools based on the *typical sets*, which have resemblances to the Gelfand-Pinsker coding [18], [19]) are not suitable for the one-shot setting. The Poisson matching lemma has been shown to be useful in proving one-shot achievability results of network information theory [14], [58], see Section IV-B for a detailed discussion.

The Poisson matching lemma is rooted in the Poisson functional representation [55] that is reviewed as follows. Fix a probability distribution $\bar{P}$ over $\mathcal{U}$. Let $(T_i)_{i=1,2,\ldots}$ be a Poisson process with rate 1, i.e., $T_1, T_2 - T_1, T_3 - T_2 \overset{\mathrm{iid}}{\sim} \mathrm{Exp}(1)$. Let $(\bar{U}_i)_i$ be an independent i.i.d. sequence with distribution $\bar{P}$. This "marked" Poisson process $(\bar{U}_i, T_i)_i$ supports a "query operation" given by the Poisson functional representation, where one can input a distribution P over $\mathcal{U}$, and obtain one sample $\tilde{U}_P$ with distribution P.

The Poisson functional representation is given by

$$\tilde{U}_P := \bar{U}_K, \text{ where } K := \arg\min_i T_i \cdot \Big(\frac{\mathrm{d}P}{\mathrm{d}\bar{P}}(\bar{U}_i)\Big)^{-1}.$$

The way this Poisson process is used in communication settings (e.g., in [14], [58]) is that the encoder would query the process using the prior distribution of the signal to obtain the signal to be sent, and the decoder would query using the posterior distribution of the signal given the noisy observation to obtain the message. There is no error in the communication if the two queries return the same sample. The probability of error can be bounded by the Poisson matching lemma [14], shown as follows.

Lemma 1 (Poisson matching lemma [14])**.** *Consider two distributions $P, Q \ll \bar{P}$. Almost surely, we have*

$$\mathbf{P}\big(\tilde{U}_Q \neq \tilde{U}_P \,\big|\, \tilde{U}_P\big) \leq 1 - \Big(1 + \frac{\mathrm{d}P}{\mathrm{d}Q}(\tilde{U}_P)\Big)^{-1}.$$

IV. One-shot Information Hiding

In this section, we formulate the one-shot information hiding problem, discuss its connections to related works, present our novel one-shot achievability results, and also explore some generalizations.

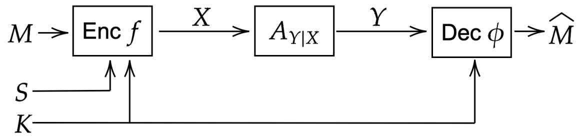

Fig. 1. The information hiding problem.

The one-shot information hiding problem is described in Figure 1. A message M is uniformly chosen from the set $[1 : \mathsf{L}]$, where L is the *message size*. The goal is to *hide* the message M into a host data source $S \in \mathcal{S}$ to produce X. The stegotext X is sent through the attack channel $A_{Y|X}$, which is chosen by an attacker that attempts to remove or degrade the embedded message during the signal transmission. We allow the *common randomness* $K \in \mathcal{K}$ to be available at both the encoder and the decoder, but not the attacker, and the decoder attempts to reconstruct the original message M after observing Y, K. The common randomness K reveals information about S to the decoder, which may be correlated with K according to the joint distribution $P_{S,K}$.

Given the random variables, the information hiding problem can be viewed as a *game* between two parties: the first party consists of the encoder (information hider) and the decoder, who are cooperatively transmitting the message M; the second party is an attacker, who is trying to remove or degrade the hidden message M in S so that the decoder cannot correctly reconstruct it. See [1], [3]–[5] for more discussions on the game-theoretic formulation. Their roles and assumptions are elaborated as follows.

- **Encoder**: The encoder observes a message M that is uniformly chosen from the set $[1:\mathsf{L}]$, and the goal of encoding is to *hide* M into a host data source $S \in \mathcal{S}$ by introducing some tolerable level of distortions. Given S, K, M, the encoder outputs $X = f(S, K, M)$, where $f : \mathcal{S} \times \mathcal{K} \times [1:\mathsf{L}] \to \mathcal{X}$. It is expected that X is close to S, in the sense that $d_1(S, X)$ is small, where $d_1 : \mathcal{S} \times \mathcal{X} \to [0, \infty)$ is a distortion measure. We want $d_1(S, X) \le \mathsf{D}_1$ with high probability. This will be elaborated later. The encoded steogtext X is then transmitted through the a channel $A_{Y|X} \in \mathcal{A}$.
- **Attacker**: The attacker is formulated as a noisy channel $A_{Y|X}$, called the *attack channel*. With input X, it performs data processing attacks by introducing another level of distortion and produces Y, a corrupted version of X. Unlike the asymptotic study [1], we do not assume the attacker's strategy is known by the encoder and the decoder (See IV-B for discussions). The attacker is free to choose $A_{Y|X}$ from a class of channels $\mathcal{A}$ (e.g., the class of channels satisfying some distortion constraint between X and Y, or the class of memoryless channels where X and Y are sequences), whose cardinality can be infinite. Both deterministic and randomized attacks can be performed. We assume the attacker has knowledge of the distributions (but not the values) of S, M, K, and also knows the code that the encoder-decoder team uses.
- **Decoder**: Upon observing the attacker's output Y and the side information K, the decoder wishes to recover the message M. It outputs $\hat{M} = \phi(K, Y)$, where $\phi : \mathcal{K} \times \mathcal{Y} \to [1:\mathsf{L}]$. The decoder is uninformed of the attacker's strategy. We require the following worst case failure probability to be small:
$$P_e := \sup_{A_{Y|X} \in \mathcal{A}} \mathbf{P}\big(d_1(S, X) > \mathsf{D}_1 \ \text{ OR } \ M \neq \hat{M}\big), \quad (1)$$
where $(S, K, M) \sim P_{S,K} \times \mathrm{Unif}[1:\mathsf{L}]$, $X = f(S, K, M)$, $Y|X \sim A_{Y|X}$ and $\hat{M} = \phi(K, Y)$ in the probability.[3]

A. One-shot Achievability Results

We then provide one-shot achievability results of the information hiding problem. Since we let the encoder-decoder team account for all possible attack channels in a set $\mathcal{A}$, the achievability results have to suffer a penalty depending on the "size" of $\mathcal{A}$. Though the cardinality of $\mathcal{A}$ could be infinite, we can often find a finite subset $\tilde{\mathcal{A}}$ such that every attack channel $A \in \mathcal{A}$ is close enough to some $\tilde{A} \in \tilde{\mathcal{A}}$. We capture this notion of size by the ϵ-covering number defined below (see similar covering arguments in [1], [17]).

Definition 1. Given a set of channels $\mathcal{A}$ from $\mathcal{X}$ to $\mathcal{Y}$, its *ϵ-covering number* is defined as

$$N_\epsilon(\mathcal{A}) := \min\Big\{|\tilde{\mathcal{A}}| : \tilde{\mathcal{A}} \subseteq \mathcal{A}, \sup_{A \in \mathcal{A}} \min_{\tilde{A} \in \tilde{\mathcal{A}}} \sup_{x \in \mathcal{X}} \Big\| A_{Y|X}(\cdot|x) - \tilde{A}_{Y|X}(\cdot|x) \Big\|_{\mathrm{TV}} \le \epsilon \Big\},$$

where $\|A_{Y|X}(\cdot|x) - \tilde{A}_{Y|X}(\cdot|x)\|_{\mathrm{TV}} \in [0, 1]$ denotes the total variation distance between $A_{Y|X}(\cdot|x)$ (the distribution of Y if $X = x$, and Y follows $A_{Y|X}$) and $\tilde{A}_{Y|X}(\cdot|x)$.

We now present the main result, which is a one-shot achievability result with a bound on the error probability in terms of $N_\epsilon(\mathcal{A})$ and the information density terms.

Theorem 2. *Fix any $P_{U,X|S,K}$ and channel $\hat{A}_{Y|X}$. For any $\epsilon \ge 0$, there exists an information hiding scheme satisfying*

$$P_e \le N_\epsilon(\mathcal{A}) \sup_{A_{Y|X} \in \mathcal{A}} \mathbf{E}_{Y|X \sim A_{Y|X}} \bigg[1 - \mathbf{1}\{d_1(S, X) \le \mathsf{D}_1\} \cdot \Big(1 + \mathsf{L} 2^{-\hat{\iota}(U;Y|K) + \iota(U;S|K)}\Big)^{-1} \bigg] + \epsilon,$$

where we assume $(S, K, U, X, Y) \sim P_{S,K} P_{U,X|S,K} A_{Y|X}$ in the expectation, and $\hat{\iota}(U;Y|K)$ is the information density computed by the joint distribution $P_{S,K} P_{U,X|S,K} \hat{A}_{Y|X}$ (instead of $A_{Y|X}$), assuming that $\iota(U;S|K), \hat{\iota}(U;Y|K)$ are almost surely finite for every $A_{Y|X} \in \mathcal{A}$.

Proof. The idea is that we design the decoder assuming that the attack channel is fixed to $A_{Y|X}$, and hope that this decoder works for every attack channel $A_{Y|X} \in \mathcal{A}$. Let $\mathcal{C} := ((\bar{U}_i, \bar{M}_i), T_i)_i$ where $(T_i)_i$ is a Poisson process, $\bar{U}_i \overset{\text{iid}}{\sim} P_U$, and $\bar{M}_i \overset{\text{iid}}{\sim} P_M$ (where $P_M = \mathrm{Unif}[1:\mathsf{L}]$). This will act as a random codebook shared between the encoder and the decoder (and this codebook will be fixed later).

The encoder observes the message $M \sim P_M$, the host signal S and side information K, by the Poisson functional representation [14], [55] on the distribution $P_{U|S,K}(\cdot|S, K) \times \delta_M$ over

[3] Note that [1] imposes a constraint on the expected distortion $\mathbf{E}[d_1(S, X)]$, which is reasonable in the context of [1] because the memoryless assumption and the law of large numbers ensure that the actual distortion is close to the expected distortion. Since we are considering a one-shot setting where we only assume the attack channel is chosen from a set $\mathcal{A}$, if constraint need to be specified, it might be more reasonable to consider $d_1(S, X) > \mathsf{D}_1$ as a failure event and bound the probability of failure, i.e., the excess distortion probability instead, compared to expected distortion.

$\mathcal{U} \times [1 : \mathsf{L}]$ it produces $U = \tilde{U}_{P_{U|S,K}(\cdot|S,K)\times\delta_M}$,[4] and sends the generated $X|(S,K,U) \sim P_{X|S,K,U}$. The decoder observes Y, K and outputs $\hat{M} = \tilde{M}_{\hat{P}_{U|Y,K}(\cdot|Y,K)\times P_M}$ by the Poisson functional representation, where $\hat{P}_{U|Y,K}$ is the conditional probability distribution computed by the joint distribution $P_{S,K}P_{U,X|S,K}\hat{A}_{Y|X}$. When the attack channel is $A_{Y|X} \in \mathcal{A}$, the error probability is bounded as follows:

$$
\begin{aligned}
&P_e(A) := 1 - \mathbf{P}_{Y|X\sim A_{Y|X}}\big(d_1(S,X) \le \mathsf{D}_1 \;\text{ AND }\; M = \hat{M}\big) \\
&= \mathbf{E}\Big[1 - \mathbf{1}\{d_1(S,X) \le \mathsf{D}_1\}\cdot\mathbf{1}\{M = \hat{M}\}\Big] \\
&= \mathbf{E}\Big[1 - \mathbf{1}\{d_1(S,X) \le \mathsf{D}_1\}\cdot\mathbf{P}\big(M = \hat{M}|M,S,U,Y,K\big)\Big] \\
&\le \mathbf{E}\Big[1 - \mathbf{1}\{d_1(S,X) \le \mathsf{D}_1\} \\
&\qquad \cdot\mathbf{P}\big((U,M) = (\tilde{U},\tilde{M})_{\hat{P}_{U|Y,K}(\cdot|Y,K)\times P_M}|M,S,U,Y,K\big)\Big] \\
&\overset{(a)}{\le} \mathbf{E}\Big[1 - \mathbf{1}\{d_1(S,X) \le \mathsf{D}_1\} \\
&\qquad \cdot\Big(1 + \frac{\mathrm{d}P_{U|S,K}(\cdot|S,K)\times\delta_M}{\mathrm{d}\hat{P}_{U|Y,K}(\cdot|Y,K)\times P_M}(U,M)\Big)^{-1}\Big] \\
&= \mathbf{E}\Big[1 - \mathbf{1}\{d_1(S,X)\le\mathsf{D}_1\}\Big(1 + \mathsf{L}2^{-\hat{\imath}(U;Y|K)+\iota(U;S|K)}\Big)^{-1}\Big] \\
&\le \sup_{A_{Y|X}\in\mathcal{A}} \mathbf{E}_{Y|X\sim A_{Y|X}}\Big[1 - \mathbf{1}\{d_1(S,X) \le \mathsf{D}_1\} \\
&\qquad \cdot\Big(1 + \mathsf{L}2^{-\hat{\imath}(U;Y|K)+\iota(U;S|K)}\Big)^{-1}\Big] \quad =: \overline{P_e},
\end{aligned}
$$

where (a) is by the Poisson matching lemma.[5] If we allow the encoder and the decoder to share unlimited additional common randomness, we can assume the codebook $\mathcal{C} = ((\bar{U}_i, \bar{M}_i), T_i)_i$ is actually shared, and conclude that $P_e = \sup_{A\in\mathcal{A}} P_e(A) \le \overline{P_e}$. Nevertheless, the only actual common randomness between the encoder and the decoder is K, which we cannot control. Therefore, we have to fix the codebook.

Let $P_e(A, c)$ be the probability of error when the attack channel is A and the codebook is $\mathcal{C} = c$. We have $P_e(A) = \mathbf{E}_{\mathcal{C}}[P_e(A,\mathcal{C})]$. Let $\tilde{\mathcal{A}} \subseteq \mathcal{A}$ attain the minimum in $N_\epsilon(\mathcal{A})$.

Consider any $A \in \mathcal{A}$, and let $\tilde{A} \in \tilde{\mathcal{A}}$ satisfy

$$\sup_{x\in\mathcal{X}} \Big\|A_{Y|X}(\cdot|x) - \tilde{A}_{Y|X}(\cdot|x)\Big\|_{\mathrm{TV}} \le \epsilon.$$

The total variation distance between the joint distribution of M, S, K, U, X, Y under the attack channel A conditional on $\mathcal{C} = c$ and the joint distribution under the attack channel $\tilde{A}$ conditional on $\mathcal{C} = c$ is also bounded by ϵ. Hence $|P_e(A,c) - P_e(\tilde{A},c)| \le \epsilon$ and

$$
\begin{aligned}
P_e(A,c) &\le P_e(\tilde{A},c) + \epsilon \\
&\le \sum_{\tilde{A}\in\tilde{\mathcal{A}}} P_e(\tilde{A},c) + \epsilon.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\mathbf{E}_{\mathcal{C}}\Big[\sup_{A\in\mathcal{A}} P_e(A,\mathcal{C})\Big] \\
&\le \mathbf{E}_{\mathcal{C}}\Big[\sum_{\tilde{A}\in\tilde{\mathcal{A}}} P_e(\tilde{A},\mathcal{C}) + \epsilon\Big] \\
&= \sum_{\tilde{A}\in\tilde{\mathcal{A}}} P_e(\tilde{A}) + \epsilon \\
&\le |\tilde{\mathcal{A}}|\cdot\overline{P_e} + \epsilon.
\end{aligned}
$$

The proof is completed by the existence of a codebook c such that $\sup_{A\in\mathcal{A}} P_e(A,c) \le |\tilde{\mathcal{A}}|\cdot\overline{P_e} + \epsilon$. □

Remark 1. Note that when $K = \emptyset$, $d_1(s,x) = 0$, and $\mathcal{A} = \{A_{Y|X}\}$ is a singleton set, taking $\hat{A}_{Y|X} = A_{Y|X}$, Theorem 2 recovers the one-shot Gelfand-Pinsker coding result in [14], which is the only known one-shot bound that attains the best known second order result in [59]. The asymptotic Gelfand-Pinsker capacity [18], [19] can be readily recovered.

Remark 2. The one-shot achievability results can be converted to a finite blocklength result where n is a fixed number using the Berry-Esseen theorem [60]–[62]. For example, see [14, Section IV] for converting the one-shot result of the Gelfand-Pinsker problem to a second-order result by the Berry-Esseen theorem [60]–[62], which can also be recovered by Theorem 2.

B. Discussions

In [1], it is assumed that the attack channel must be memoryless, and the decoder can obtain full knowledge of the attack channel. This assumption can be reasonably justified by the large blocklength, as one can use training symbols at the beginning of transmission, with their size becoming negligible compared to the blocklength. However, this assumption becomes questionable in the one-shot scenario, where the attack channel can be arbitrary and is only used once. We drop this assumption and allow the decoder to be totally uninformed of the attack channel. In [4] this assumption is also dropped, and an asymptotic hiding capacity expressed as the limit of a sequence of single-letter expressions has been derived using constant composition codes. The key difference between [4] and our setting is that the side information in [4] is a shared key of unlimited size independent of M, S that can be chosen as a part of the coding scheme, whereas in our paper and [1] the K is a given side information that may be correlated with S (where the dependence is from the joint distribution $P_{S,K}$), and cannot be changed. In some watermarking problems [24], [63] certain components can be further constrained, e.g., there may exist a mapping from the message M to a codeword $V(M)$ which is independent of S, and then composite data are obtained by a mapping from S, K and $V(M)$.

The information hiding setting can be regarded as a variant of Gelfand-Pinsker coding for channels with side information at the encoder [18], [19], where the channel is fixed and not

[4] The Poisson functional representation produces a pair $(\tilde{U}, \tilde{M})$, and U is set to be the first component of the pair.

[5] The Poisson matching lemma is applied on the conditional distributions given M, S, U, Y, K. Also see the conditional Poisson matching lemma [14].

chosen by the attacker, and there is no shared side information between the encoder and the decoder. Since the encoder and the decoder have to account for all possible attack channels, this can be regarded as a combination of Gelfand-Pinsker coding and compound channel [17], [64], [65]. The analyses in [1], [4] utilize techniques such as random binning, joint typicality decoding and constant composition codes, which are also commonly utilized in the asymptotic analyses of Gelfand-Pinsker coding [18], [59]. These techniques may not be suitable for our one-shot setting. Strong typicality and constant composition codes are inapplicable when the blocklength is 1. While random binning can be applied to one-shot Gelfand-Pinsker coding [12], [13], [47], it produces weaker results compared to the Poisson matching lemma [14]. To obtain tight one-shot bounds for information hiding, we utilize the Poisson matching lemma instead, which has been shown to perform well in various one-shot settings [14], [58].

C. Generalizations

We discuss some generalizations of our information hiding setting.

As shown in Figure 1, we have a host signal S available to the encoder and a side information K available to both the encoder and the decoder. This can be generalized by letting the encoder have a side information source S_{Enc}, the decoder have another side information source S_{Dec}, and they possibly share a codebook $\mathcal{C}$. In this case, our setting (also [1] and public watermarking [4]) correspond to $S_{\mathrm{Enc}} = (S, K)$ and $S_{\mathrm{Dec}} = K$, while private watermarking [3] corresponds to $S_{\mathrm{Enc}} = S_{\mathrm{Dec}} = S$. This setting has been investigated in [20]. Moreover, in [33], [34], the decoder not only recovers the message M, but also lossily reconstructs the stegotext X, within another level of distortion $d_1'(X, \hat{X})$. See Figure 2 for an illustration.

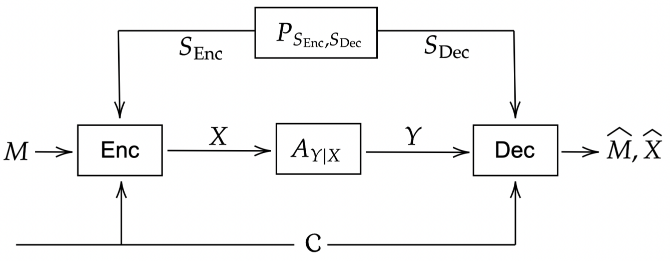
$P_{S_{\mathrm{Enc}},S_{\mathrm{Dec}}}$
S_{Enc}
S_{Dec}
M
Enc
X
$A_{Y|X}$
Y
Dec
$\widehat{M}, \widehat{X}$
C

Fig. 2. The generalized information hiding problem.

By employing these generalizations, the information hiding setting can be extended as follows: Upon observing side information S_{Enc} and the message M, the encoding function $f : \mathcal{S}_{\mathrm{Enc}} \times [1 : \mathsf{L}] \to \mathcal{X}$ outputs $X = f(S_{\mathrm{Enc}}, M)$; the decoder observes Y, the output of the attacker, together with another side information S_{Dec}, and computes $(\hat{M}, \hat{X}) = \phi(S_{\mathrm{Dec}}, Y)$, the distorted versions of M and X, where $\phi : \mathcal{S}_{\mathrm{Dec}} \times \mathcal{Y} \to [1 : \mathsf{L}] \times \mathcal{X}$. To highlight the potential advantage of *randomized* coding schemes, we allow both the encoding function f and decoding function g to depend on another random variable $\mathcal{C}$, which is shared between the encoder and decoder but unknown to the attacker. This random variable $\mathcal{C}$ can serve as a source of common randomness (or equivalently, as a codebook shared between the encoder and decoder).

The distortion of the stegotext X is measured by $d_1'(X, \hat{X})$, where $d_1' : \mathcal{X} \times \mathcal{X} \to [0, \infty)$ is another distortion measure. The decoder is uninformed of the attack channel, and in this generalized case we also intend to bound the worst case failure probability

$$P_e := \sup_{A_{Y|X} \in \mathcal{A}} \mathbf{P}\big(d_1(S, X) > \mathsf{D}_1 \ \text{ OR } \ d_1'(X, \hat{X}) > \mathsf{D}_1' \ \text{ OR } \ M \neq \hat{M}\big), \qquad (2)$$

where we assume $(S_{\mathrm{Enc}}, S_{\mathrm{Dec}}, M) \sim P_{S_{\mathrm{Enc}},S_{\mathrm{Dec}}} \times \mathrm{Unif}[1 : \mathsf{L}]$, $X = f(S_{\mathrm{Enc}}, M)$, $Y|X \sim A_{Y|X}$ and $(\hat{M}, \hat{X}) = \phi(S_{\mathrm{Dec}}, Y)$ in the probability.

Based on this setting, we can derive the following theorem, which is a generalized version of Theorem 2. The proof is similar to that of Theorem 2 and is therefore omitted.

Theorem 3. *Fix any $P_{U,X|S_{\mathrm{Enc}},S_{\mathrm{Dec}}}$ and channel $\hat{A}_{Y|X}$. For any $\epsilon \geq 0$, there exists a scheme for the information hiding problem satisfying*

$$P_e \leq N_\epsilon(\mathcal{A}) \cdot \sup_{A \in \mathcal{A}} \mathbf{E}_{Y|X \sim A}\Bigg[1 - \mathbf{1}\{d_1(S_{\mathrm{Enc}}, X) \leq \mathsf{D}_1\} \cdot \mathbf{1}\{d_1'(X, \hat{X}) \leq \mathsf{D}_1'\} \cdot \Big(1 + \mathsf{L} \cdot 2^{-\hat{\iota}(U;Y,S_{\mathrm{Dec}}) + \iota(U;S_{\mathrm{Enc}})}\Big)^{-1}\Bigg] + \epsilon,$$

where we assume $(S_{\mathrm{Enc}}, S_{\mathrm{Dec}}, U, X, Y) \sim P_{S_{\mathrm{Enc}},S_{\mathrm{Dec}}} P_{U,X|S_{\mathrm{Enc}}} A_{Y|X}$ in the expectation, and $\hat{\iota}(U; Y, S_{\mathrm{Dec}})$ is the information density computed by the joint distribution $P_{S,K} P_{U,X|S,K} \hat{A}_{Y|X}$ (instead of $A_{Y|X}$), assuming that $\iota(U; S_{\mathrm{Enc}}), \hat{\iota}(U; Y, S_{\mathrm{Dec}})$ are almost surely finite for every $A_{Y|X} \in \mathcal{A}$.

V. Recovery of the Asymptotic Information Hiding Capacity

In this section, we show that our Theorem 2 recovers the asymptotic information hiding capacity [1] when applied to the discrete and memoryless setting, with the attack channels subject to a distortion constraint. This gives an alternative proof to [1]. Note that, from a similar procedure, we can recover the achievable bound for information hiding with stegotext reconstruction [34, Theorem 1] (which, in turn, is an extension of [33] and [37]) by using Theorem 3. We choose to present the details of Theorem 2 recovering the information hiding capacity [1] for simplicity, as it already captures the core idea.

We first provide a simple bound on the ϵ-covering number in the case that X and Y are discrete and finite.

Proposition 4. *If $\mathcal{X}$ and $\mathcal{Y}$ are finite, then*

$$N_\epsilon(\mathcal{A}) \le \Big(\frac{1}{2\epsilon} + \frac{|\mathcal{Y}|+1}{2}\Big)^{|\mathcal{X}|\cdot|\mathcal{Y}|}.$$

Proof. Write

$$d(A, \tilde{A}) := \sup_{x\in\mathcal{X}} \|A_{Y|X}(\cdot|x) - \tilde{A}_{Y|X}(\cdot|x)\|_{\mathrm{TV}}.$$

We use the standard method to bound the covering number, where we start with $\tilde{\mathcal{A}} = \emptyset$, and add $A \in \mathcal{A}$ not currently covered by $\tilde{\mathcal{A}}$ (i.e., $\min_{\tilde{A}\in\tilde{\mathcal{A}}} d(A,\tilde{A}) > \epsilon$) to $\tilde{\mathcal{A}}$ one by one until all of $\mathcal{A}$ is covered. Note that every two different $\tilde{A}, \tilde{A}' \in \tilde{\mathcal{A}}$ produced this way must satisfy $d(\tilde{A},\tilde{A}') > \epsilon$, and hence the $(\epsilon/2)$-balls $\{A : d(A,\tilde{A}) \le \epsilon/2\}$ must be disjoint for $\tilde{A} \in \tilde{\mathcal{A}}$.

We now treat $A_{Y|X}$ as a transition probability matrix $A \in \mathbb{R}^{|\mathcal{Y}|\times|\mathcal{X}|}$. We have

$$\begin{aligned} d(A,\tilde{A}) &= \frac{1}{2}\|A - \tilde{A}\|_1 \\ &= \frac{1}{2}\max_x \sum_y |A_{y,x} - \tilde{A}_{y,x}|. \end{aligned}$$

The volume of the ball $\{A \in \mathbb{R}^{|\mathcal{Y}|\times|\mathcal{X}|} : d(A,\tilde{A}) \le \epsilon/2\}$ (i.e., its Lebesgue measure in the space $\mathbb{R}^{|\mathcal{Y}|\cdot|\mathcal{X}|}$) is $((2\epsilon)^{|\mathcal{Y}|}/(|\mathcal{Y}|!))^{|\mathcal{X}|}$, and all these balls are subsets of $\{A \in \mathbb{R}^{|\mathcal{Y}|\times|\mathcal{X}|} : \min_{x,y} A_{y,x} \ge -\epsilon,\ \max_x \sum_y A_{y,x} \le 1+\epsilon\}$, which has a volume $((1+(|\mathcal{Y}|+1)\epsilon)^{|\mathcal{Y}|}/(|\mathcal{Y}|!))^{|\mathcal{X}|}$. Hence, the size of $\tilde{\mathcal{A}}$ is upper-bounded by

$$\frac{\Big((1+(|\mathcal{Y}|+1)\epsilon)^{|\mathcal{Y}|}/(|\mathcal{Y}|!)\Big)^{|\mathcal{X}|}}{\big((2\epsilon)^{|\mathcal{Y}|}/(|\mathcal{Y}|!)\big)^{|\mathcal{X}|}} = \Big(\frac{1}{2\epsilon} + \frac{|\mathcal{Y}|+1}{2}\Big)^{|\mathcal{X}|\cdot|\mathcal{Y}|}.$$

□

We now show that Theorem 2 recovers the asymptotic result in [1] when S, K, X, Y are finite and discrete, and the attack channel must be memoryless and is subject to a distortion constraint.

Consider sequences $S^n = (S_1, \ldots, S_n)$, K^n, X^n, Y^n where $(S_i, K_i) \overset{\mathrm{iid}}{\sim} P_{S,K}$. Consider a channel input distribution P_X. The class of attack channels $\mathcal{A}_n = \mathcal{A}_n(P_X)$ (which depends on P_X) is taken to be

$$\mathcal{A}_n(P_X) := \big\{A^n_{Y|X} : A_{Y|X} \in \mathcal{A}(P_X)\big\},$$

and we let

$$\mathcal{A}(P_X) := \big\{A_{Y|X} : \mathbf{E}_{(X,Y)\sim P_X A_{Y|X}}[d_2(X,Y)] \le \mathsf{D}_2\big\},$$

where $d_2 : \mathcal{X}\times\mathcal{Y} \to [0,\infty)$ is a distortion measure, and D_2 is the allowed distortion level. In other words, the attacker can only use memoryless channels $A^n_{Y|X}$ that satisfy the expected distortion constraint $\mathbf{E}[d_2(X,Y)] \le \mathsf{D}_2$. The asymptotic hiding capacity given in [1] is

$$C = \max_{P_{U,X|S,K}} \min_{A_{Y|X} : \mathbf{E}[d_2(X,Y)]\le\mathsf{D}_2} \big(I(U;Y|K) - I(U;S|K)\big).$$

where the maximum is over $P_{U,X|S,K}$ with $\mathbf{E}[d_1(S,X)] \le \mathsf{D}_1$.

We now show the achievability of the above asymptotic rate as a direct corollary of Theorem 2. Fix $P_{U,X|S,K}$ which achieves the above maximum subject to $\mathbf{E}[d_1(S,X)] \le \mathsf{D}_1'$ where $\tilde{\mathsf{D}}_1 < \mathsf{D}_1$. Take $\hat{A}_{Y|X}$ to be the minimizer of the rate-distortion function $\min_{A_{Y|X} : \mathbf{E}[d_2(X,Y)]\le\mathsf{D}_2} I(U;Y|K)$, and assume $(S,K,U,X,Y) \sim P_{S,K}P_{U,X|S,K}\hat{A}_{Y|X}$. Write the information density and mutual information obtained from this distribution as $\hat{\iota}_{U;Y|K}$ and $\hat{I}(U;Y|K)$, respectively. Fix a coding rate $R < \hat{I}(U;Y|K) - I(U;S|K)$. We want to show that this rate is achievable.

Consider any attack channel $A_{Y|X}$ with $\mathbf{E}[d_2(X,Y)] \le \mathsf{D}_2$. Let $A^\lambda_{Y|X} := (1-\lambda)\hat{A}_{Y|X} + \lambda A_{Y|X}$ for $0 \le \lambda \le 1$. Write $I_\lambda(U;Y|K)$ for the mutual information computed assuming $Y|X \sim A^\lambda_{Y|X}$. It is straightforward to check that

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} I_\lambda(U;Y|K)\Big|_{\lambda=0} = \mathbf{E}_{Y|X\sim A_{Y|X}}[\hat{\iota}(U;Y|K)] - \hat{I}(U;Y|K).$$

By the optimality of $\hat{A}$, the above derivative is nonnegative, and hence $\mathbf{E}_{Y|X\sim A_{Y|X}}[\hat{\iota}(U;Y|K)] \ge \hat{I}(U;Y|K)$. Therefore, when we have i.i.d. sequences $(S^n, K^n, U^n, X^n, Y^n) \sim P^n_{S,K}P^n_{U,X|S,K}A^n_{Y|X}$ and $\mathsf{L} = \lfloor 2^{nR} \rfloor$, by the law of large numbers,

$$\begin{aligned} &\mathsf{L}2^{-\hat{\iota}(U^n;Y^n|K^n)+\iota(U^n;S^n|K^n)} \\ &\le 2^{nR-\sum_{i=1}^n(\hat{\iota}(U_i;Y_i|K_i)-\iota(U_i;S_i|K_i))} \\ &\to 0 \end{aligned}$$

exponentially as $n \to \infty$ since $\mathbf{E}[\hat{\iota}(U_i;Y_i|K_i) - \iota(U_i;S_i|K_i))] \ge \hat{I}(U;Y|K) - I(U;S|K) > R$. We also have $d_1(S^n,X^n) = \sum_{i=1}^n d_1(S_i,X_i) > n\mathsf{D}_1$ with probability approaching 0 exponentially since $\tilde{\mathsf{D}}_1 < \mathsf{D}_1$. These convergences are uniform over all such attack channels $A_{Y|X}$ since the random variables are discrete and finite.

Therefore, to bound P_e using Theorem 2, it is left to bound the ϵ-covering number $N_\epsilon(\mathcal{A}_n(P_X))$. Note that

$$\begin{aligned} &\Big\|A^n_{Y|X}(\cdot|x^n) - \tilde{A}^n_{Y|X}(\cdot|x^n)\Big\|_{\mathrm{TV}} \\ &\le \sum_{i=1}^n \Big\|A_{Y|X}(\cdot|x_i) - \tilde{A}_{Y|X}(\cdot|x_i)\Big\|_{\mathrm{TV}}, \end{aligned}$$

and hence we can construct a ϵ-cover of $\mathcal{A}_n(P_X)$ using an (ϵ/n)-cover of $\mathcal{A}(P_X)$. Therefore,

$$\begin{aligned} N_\epsilon(\mathcal{A}_n(P_X)) &\le N_{\epsilon/n}(\mathcal{A}(P_X)) \\ &= O((n/\epsilon)^{|\mathcal{X}|\cdot|\mathcal{Y}|}) \end{aligned}$$

by Proposition 4, which grows much slower than the exponential decrease of the expectation in Theorem 2. Therefore, taking $\epsilon = 1/n$, we have $P_e \to 0$ as $n \to \infty$. Taking $\tilde{\mathsf{D}}_1 \to \mathsf{D}_1$ completes the proof.

VI. One-shot Compound Wiretap Channels

In this section, we consider the compound wiretap channel [2] in the one-shot setting. We utilize the Poisson matching lemma [14], under a framework similar to the one-shot information hiding discussed in Section IV. We provide novel one-shot achievablity results for the compound wiretap channel [2]. To the best of our knowledge, the one-shot results of this problem has not been discussed in literature, though finite-blocklength bounds on the single wiretap channel (without uncertainties/multiple-states of the channel) can be found in [66]–[68].

Unlike the asymptotic analysis of the discrete memoryless compound wiretap channel [2], our one-shot results also apply to *continuous* scenarios. Note that [44] also studied the continuous case of compound wiretap channels, but the focus in [44] was mainly on the compound Gaussian MIMO wiretap channels, and the analysis was not one-shot.

A. Problem Formulation

The one-shot compound wiretap channel setting is described as follows. A message M is uniformly chosen from $\mathrm{Unif}[\mathsf{L}]$. Upon observing $M \sim \mathrm{Unif}[\mathsf{L}]$, the encoder produces $X = f(M)$, where $f : [\mathsf{L}] \to \mathcal{X}$ is a randomized encoding function. Then X is sent through a channel $P_{Y,Z|X}$ that is unknown to the encoder and the decoder but known to the eavesdropper. A legitimate decoder observes Y and recovers $\hat{M} = g(Y)$, where $g : \mathcal{Y} \to [\mathsf{L}]$ is a decoding function. The eavesdropper observes $Z \in \mathcal{Z}$. Justified by [2] and [69, Lemma 1], we can assume the transition probability distribution is $P_{Y|X}P_{Z|X}$ by decomposing $P_{Y,Z|X}$ without loss of optimality.

We do not fix $P_{Y|X}$, but rather allow $P_{Y|X}$ to be any element in $\mathcal{D}$, a set of channels to the legitimate **d**ecoder. Similarly, we allow $P_{Z|X}$ to be any element in $\mathcal{E}$, a set of channels to the **e**avesdropper. In this paper, we assume the encoder and decoder have no knowledge of $P_{Y|X}$ and $P_{Z|X}$, making them *universal* in the sense that their performance guarantees do not depend on specific channel realizations. However, for secrecy applications, we make the conservative assumption that the eavesdropper knows both channel distributions $P_{Y|X}$ and $P_{Z|X}$.

Unlike [2] but similar to our discussions on the information hiding problem, the set of channels to the legitimate decoder $\mathcal{D}$ and the set of channels to the eavesdropper $\mathcal{E}$ can be infinite, which captures the infinite variability of real-world signals and their propagation characteristics in practical applications. While the cardinalities of $\mathcal{D}, \mathcal{E}$ can be infinite, we can often find finite subsets $\tilde{\mathcal{D}}$, $\tilde{\mathcal{E}}$ such that every channel to the decoder (or eavesdropper) in $\tilde{\mathcal{D}}$ (or $\tilde{\mathcal{E}}$) would be close enough to some $\tilde{P}_{Y|X} \in \tilde{\mathcal{D}}$ (or $\tilde{P}_{Z|X} \in \tilde{\mathcal{E}}$). This idea has appeared in Section IV and also in [17], [44].

The objective is to bound the worst case error probability

$$P_e := \sup_{P_{Y|X} \in \mathcal{D}} \mathbf{P}\big(M \neq \hat{M}\big), \tag{3}$$

while also ensure the secrecy is guaranteed, which is measured by the total variation distance

$$\gamma := \sup_{P_{Z|X} \in \mathcal{E}} \|P_{M,Z} - P_M \times P_Z\|_{\mathrm{TV}} \tag{4}$$

being small.

B. One-Shot Achievability

We then provide one-shot achievability results of the compound wiretap channel. Note the result can be viewed as a combination of the covering argument appeared in Section IV and the one-shot soft covering lemma in [14, Proposition 3]. Other existing one-shot wiretap channel results [66], [67] might also be utilized in a similar framework as well.

Theorem 5. *Fix any $P_{U,X}$ and any wiretap channel $\hat{P}_{Y,Z|X} = \hat{P}_{Y|X}P_{Z|X}$. For any $\nu \geq 0$, any $\epsilon_1, \epsilon_2 \geq 0$ and $\mathsf{A}, \mathsf{B} \in \mathbb{N}$, there exists a code for the compound wiretap channel setting, with message $M \sim \mathrm{Unif}[\mathsf{L}]$, satisfying*

$$\begin{aligned}
&P_e + \nu \cdot \gamma \leq \\
&N_{\epsilon_1}(\mathcal{D}) \sup_{P_{Y|X} \in \mathcal{D}} \mathbf{E}_{Y|X \sim P_{Y|X}}\Big[\min\Big\{\mathsf{LA}2^{-\hat{\iota}(U;Y)}, 1\Big\}\Big] + \epsilon_1 \\
&\quad + \nu \cdot N_{\epsilon_2}(\mathcal{E})\left(\sup_{P_{Z|X} \in \mathcal{E}} 2 \cdot \mathbf{E}_{Z|X \sim P_{Z|X}}\Big[\big(1 + 2^{-\iota(U;Z)}\big)^{-\mathsf{B}}\Big]\right. \\
&\qquad\qquad \left. + \sqrt{\mathsf{BA}^{-1}}\right) + \nu \cdot \epsilon_2,
\end{aligned}$$

where we assume $(U, X, Y, Z) \sim P_{U,X}P_{Y|X}P_{Z|X}$ in the expectation, and $\hat{\iota}(U;Y)$ is the information density computed for compound channels by the joint distribution $P_{U,X}\hat{P}_{Y|X}P_{Z|X}$, assuming that $\hat{\iota}(U;Y)$ is almost surely finite for every $P_{Y|X} \in \mathcal{D}$.

Proof. Since the encoder and the decoder are uninformed of the transmission channel, we first design our coding strategy assuming that the transmission channel to the legitimate decoder is fixed to $\hat{P}_{Y|X} \in \mathcal{D}$. However, note that the eavesdropper is aware of both the transmission channel and the eavesdropping channel $P_{Z|X} \in \mathcal{E}$ that are in use.

Let $\mathcal{C} := ((\bar{U}_i, \bar{M}_i), T_i)_i$ where $(T_i)_i$ is a Poisson process, $\bar{U}_i \overset{\text{iid}}{\sim} P_U$, and $\bar{M}_i \overset{\text{iid}}{\sim} P_M$ (where $P_M = \mathrm{Unif}[\mathsf{L}]$). This is a random codebook that is known to both the encoder and the decoder. Conditioned on using this codebook, we analyze the error probability and secrecy as follows. We will fix the codebook later.

Let $A \sim \mathrm{Unif}[\mathsf{A}]$ be independent of $(M, \mathcal{C})$. The encoder observes the message $M \sim P_M$, computes $U = \tilde{U}_{P_U \times \delta_M}(A)$ by the Poisson functional representation [55], and sends the generated $X|U \sim P_{X|U}$. The decoder observes the channel ouptut Y and recovers $\hat{M} = \tilde{M}_{\hat{P}_{U|Y}(\cdot|Y) \times P_M}$ where $\hat{P}_{U|Y}$ is the conditional distribution computed by the joint distribution $P_{U,X}\hat{P}_{Y|X}$. We have $(M, A, U, X, Y, Z) \sim P_M \times P_A \times P_{U,X}\hat{P}_{Y|X}\hat{P}_{Z|X}$. For the case of a fixed $\hat{P}_{Y|X} \in \mathcal{D}$, the error probability $\mathbf{P}\{M \neq \hat{M}\}$ in this case, denoted by $P_e(\hat{P}_{Y|X})$, can be bounded as follows:

$$\begin{aligned}
&P_e(\hat{P}_{Y|X}) \\
&\leq \mathbf{E}\Big[\mathbf{P}\big((U, M) \neq (\tilde{U}, \tilde{M})_{\hat{P}_{U|Y}(\cdot|Y) \times P_M} | M, A, U, Y\big)\Big] \\
&\overset{(a)}{\leq} \mathbf{E}\Big[\min\Big\{\mathsf{A}\frac{dP_U \times \delta_M}{d\hat{P}_{U|Y}(\cdot|Y) \times P_M}(U, M), 1\Big\}\Big] \\
&= \mathbf{E}\Big[\min\Big\{\mathsf{LA}2^{-\hat{\iota}_{U;Y}(U;Y)}, 1\Big\}\Big]
\end{aligned}$$

$$
\begin{aligned}
&\leq \sup_{P_{Y|X}\in\mathcal{D}} \mathbf{E}_{Y|X\sim A_{Y|X}}\left[\min\left\{\mathsf{LA}2^{-\hat{\iota}_{U;Y}(U;Y)},1\right\}\right] \qquad (5)\\
&=: \overline{P_e}
\end{aligned}
$$

where (a) is by the conditional generalized Poisson matching lemma [14], and we define (5) to be $\overline{P_e}$.

For the secrecy measure γ, note again that the wiretap channel $P_{Z|X}$ is known to the eavesdropper, and for this fixed choice of $P_{Z|X}$, for the total variation distance $\gamma(P_{Z|X})$, we write

$$
\begin{aligned}
&\mathbf{E}\left[\left\|P_{M,Z|\mathcal{C}}(\cdot,\cdot|\mathcal{C}) - P_M \times P_{Z|\mathcal{C}}(\cdot|\mathcal{C})\right\|_{\mathrm{TV}}\right]\\
&= \mathbf{E}\left[\left\|P_{Z|M,\mathcal{C}}(\cdot,\cdot|\mathcal{C}) - P_{Z|\mathcal{C}}(\cdot|\mathcal{C})\right\|_{\mathrm{TV}}\right]\\
&\leq \mathbf{E}\left[\left\|P_{Z|M,\mathcal{C}}(\cdot,\cdot|\mathcal{C}) - P_{Z}(\cdot)\right\|_{\mathrm{TV}}\right] +\\
&\qquad\qquad \mathbf{E}\left[\left\|P_{Z|\mathcal{C}}(\cdot|\mathcal{C}) - P_{Z}(\cdot)\right\|_{\mathrm{TV}}\right]\\
&\overset{(a)}{\leq} 2\cdot\mathbf{E}\left[\left\|P_{Z|M,\mathcal{C}}(\cdot,\cdot|\mathcal{C}) - P_{Z}(\cdot)\right\|_{\mathrm{TV}}\right]\\
&= 2\cdot\mathbf{E}\left[\left\|\mathsf{A}^{-1}\sum_{a=1}^{\mathsf{A}} P_{Z|U}(\cdot|\tilde{U}_{P_U\times\delta_M}(a)) - P_{Z}(\cdot)\right\|_{\mathrm{TV}}\right]\\
&\overset{(b)}{\leq} 2\cdot\mathbf{E}\left[\left(1+2^{-\iota(U;Z)}\right)^{-\mathsf{B}}\right] + \sqrt{\mathsf{BA}^{-1}}\\
&\leq \sup_{P_{Z|X}\in\mathcal{E}} 2\cdot\mathbf{E}\left[\left(1+2^{-\hat{\iota}(U;Z)}\right)^{-\mathsf{B}}\right] + \sqrt{\mathsf{BA}^{-1}} \qquad (6)\\
&=: \bar{\gamma}
\end{aligned}
$$

where (a) is by the convexity of total variation distance, (b) is by [14, Proposition 3] since $\left\{\tilde{U}_{P_U\times\delta_m}(a)\right\}_{a\in[\mathsf{A}]} \overset{\mathrm{iid}}{\sim} P_U$ for any m, and we define (6) to be $\overline{\gamma}$.

Let $P_e\left(P_{Y|X}, c\right)$ denote be the probability of error when the legitimate channel is $P_{Y|X}$ and the codebook is $\mathcal{C} = c$ and also let $\gamma\left(P_{Z|X}, c\right)$ denote the total variation distance γ when the wiretap channel is $P_{Z|X}$ and the codebook is $\mathcal{C} = c$.

Let $\tilde{\mathcal{D}} \subseteq \mathcal{D}$ attain the minimum in $N_{\epsilon_1}(\mathcal{D})$ and $\tilde{\mathcal{E}} \subseteq \mathcal{E}$ attain the minimum in $N_{\epsilon_2}(\mathcal{E})$. Consider any $P_{Y|X} \in \mathcal{D}$ and any $P_{Z|X} \in \mathcal{E}$, and let $\tilde{P}_{Y|X} \in \tilde{\mathcal{D}}, \tilde{P}_{Z|X} \in \tilde{\mathcal{E}}$ satisfy

$$
\begin{aligned}
\sup_{x\in\mathcal{X}} \left\|P_{Y|X}(\cdot|x) - \tilde{P}_{Y|X}(\cdot|x)\right\|_{\mathrm{TV}} \leq \epsilon_1,\\
\sup_{x\in\mathcal{X}} \left\|P_{Z|X}(\cdot|x) - \tilde{P}_{Z|X}(\cdot|x)\right\|_{\mathrm{TV}} \leq \epsilon_2.
\end{aligned}
$$

The total variation distance between the joint distribution of M, A, U, X, Y, Z under the channel $P_{Y|X}$ (or $P_{Z|X}$) conditional on $\mathcal{C} = c$ and the joint distribution under the channel $\tilde{P}_{Y|X}$ (or $\tilde{P}_{Z|X}$) conditional on $\mathcal{C} = c$ is also bounded by ϵ_1 (or ϵ_2). Therefore, we have $\left|P_e(P_{Y|X}, c) - P_e(\tilde{P}_{Y|X}, c)\right| \leq \epsilon_1$ and $\left|\gamma(P_{Z|X}, c) - \gamma(\tilde{P}_{Z|X}, c)\right| \leq \epsilon_2$.

Hence,

$$
\begin{aligned}
&P_e(P_{Y|X}, \mathcal{C}) + \nu\cdot\gamma(P_{Z|X}, \mathcal{C})\\
&\leq P_e\left(\tilde{P}_{Y|X}, c\right) + \epsilon_1 + \nu\cdot\gamma\left(\tilde{P}_{Z|X}, c\right) + \nu\cdot\epsilon_2\\
&\leq \sum_{\tilde{P}_{Y|X}\in\tilde{\mathcal{D}}} P_e\left(\tilde{P}_{Y|X}, c\right) + \epsilon_1\\
&\qquad + \nu\cdot\sum_{\tilde{P}_{Z|X}\in\tilde{\mathcal{E}}} \gamma\left(\tilde{P}_{Z|X}, c\right) + \nu\cdot\epsilon_2.
\end{aligned}
$$

Combining with our previous analyses on the error probability and secrecy, we have

$$
\begin{aligned}
&\mathbf{E}_{\mathcal{C}}\left[\sup_{P_{Y|X}\in\mathcal{D}, P_{Z|X}\in\mathcal{E}} \left(P_e(P_{Y|X}, \mathcal{C}) + \nu\cdot\gamma(P_{Z|X}, \mathcal{C})\right)\right]\\
&\leq \mathbf{E}_{\mathcal{C}}\left[\sum_{\tilde{P}_{Y|X}\in\tilde{\mathcal{D}}} P_e\left(\tilde{P}_{Y|X}, \mathcal{C}\right) + \epsilon_1\right.\\
&\qquad\qquad \left. + \nu\cdot\sum_{\tilde{P}_{Z|X}\in\tilde{\mathcal{E}}} \gamma\left(\tilde{P}_{Z|X}, \mathcal{C}\right) + \nu\cdot\epsilon_2\right]\\
&\leq |\tilde{\mathcal{D}}|\cdot\overline{P_e} + \nu\cdot|\tilde{\mathcal{E}}|\cdot\overline{\gamma} + \epsilon_1 + \nu\cdot\epsilon_2,
\end{aligned}
$$

and there exists a fixed set of points $((\bar{u}_i, \bar{m}_i), t_i)_i$ such that conditioned on codebook $\mathcal{C} = ((\bar{U}_i, \bar{M}_i), T_i)_i = ((\bar{u}_i, \bar{m}_i), t_i)_i$ the desired bound in our theorem is attained, and hence the proof is completed.

□

C. Recovery of the Asymptotic Results

In this section, we recover the existing asymptotic results [2]. The procedure is generally similar to Section V where we recover the asymptotic information hiding capacity [1]. In [2], it was assumed that all the random variables are discrete, the channels are memoryless, and $\mathcal{D} := \{P^n_{Y_1|X}, \ldots, P^n_{Y_{\mathsf{J}}|X}\}$ and $\mathcal{E} := \{P^n_{Z_1|X}, \ldots, P^n_{Z_{\mathsf{K}}|X}\}$ for some finite $\mathsf{J}, \mathsf{K} \in \mathbb{N}_+$. The setting [2] can be understood as Figure 3.

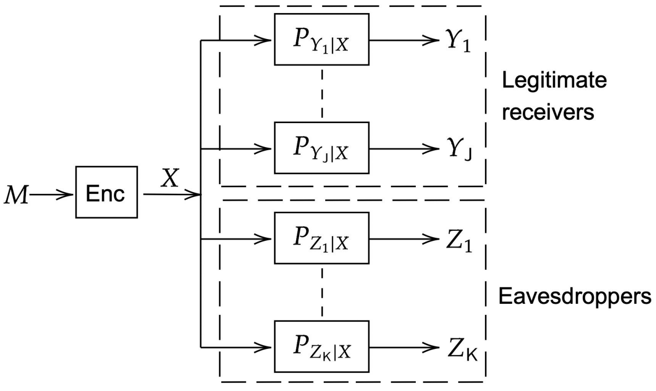

Fig. 3. Discrete memoryless compound wiretap channel setting in [2].

By [2], for discrete memoryless channels, an achievable secrecy rate is

$$
\begin{aligned}
R &= \max\left[\min_j I(U;Y_j) - \max_k I(U;Z_k)\right]\\
&= \max\min_{j,k}\left[I(U;Y_j) - I(U;Z_k)\right], \qquad (7)
\end{aligned}
$$

where the maximum is taken over all distributions $P_{U,X}$ such that the auxiliary random variable U satisfies the Markov chain

$U \leftrightarrow X \leftrightarrow (Y_j, Z_k)$ for $j = 1, \ldots, \mathsf{J}$ and $k = 1, \ldots, \mathsf{K}$. Intuitively, this means that we should design the auxiliary random variable U that maximizes the *worst-case* communication rate $I(U; Y_j) - I(U; Z_k)$, where we consider the worst receiver in $\mathcal{D}$ and the most-powerful eavesdropper in $\mathcal{E}$.

We now show the achievability of the above asymptotic rate as a direct corollary of Theorem 5. Fix $P_{U,X}$ that achieves the maximum in (7). Within the choices of $\mathcal{D}' := \{P_{Y_1|X}, \ldots, P_{Y_\mathsf{J}|X}\}$, take $\hat{P}_{Y|X}$ to be the one that minimizes $I(U;Y)$. Assume $(M, U, X, Y, Z) \sim P_M \times P_{U,X}\hat{P}_{Y|X}P_{Z|X}$. Write the information density and mutual information obtained terms from this distribution as $\hat{\imath}(U;Y)$ and $\hat{I}(U;Y)$, respectively. Fix a coding rate $R = \hat{I}(U;Y) - I(U;Z) - \epsilon$, and we are left to show that this rate is achievable for any $\epsilon > 0$.

Consider any channel $P'_{Y|X}$ from $\mathcal{D}'$. Let $P_\lambda := (1 - \lambda)\hat{P}_{Y|X} + \lambda P'_{Y|X}$ for $0 \le \lambda \le 1$. Write $I_\lambda(U;Y)$ for the mutual information computed assuming $Y|X \sim P_\lambda$. It is straightforward to check that

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} I_\lambda(U;Y)\Big|_{\lambda=0} = \mathbf{E}_{Y|X\sim P'_{Y|X}}[\hat{\imath}(U;Y)] - \hat{I}(U;Y).$$

By the optimality of $\hat{P}_{Y|X}$, the above derivative is nonnegative, and hence

$$\mathbf{E}_{Y|X\sim P'_{Y|X}}[\hat{\imath}(U;Y)] \ge \hat{I}(U;Y). \tag{8}$$

Therefore, having i.i.d. sequences $(U^n, X^n, Y^n, Z^n) \sim P^n_{U,X}P^n_{Y|X}P^n_{Z|X}$ and taking $\mathsf{L} = \lfloor 2^{nR} \rfloor$, $\mathsf{A} = 2^{n(I(U;Z)+\epsilon/2)}$ and $\mathsf{B} = 2^{n(I(U;Z)+\epsilon/3)}$, by the law of large numbers, the following terms in Theorem 5

$$\begin{aligned}
\mathsf{LA}2^{-\hat{\imath}(U;Y)} &\overset{(a)}{\le} 2^{nR+n(I(U;Z)+\epsilon/2)-\sum_{i=1}^n \hat{\imath}(U_i;Y_i)} \to 0,\\
\left(1 + 2^{-\hat{\imath}(U;Z)}\right)^{-\mathsf{B}} &\overset{(b)}{\le} 2^{\sum_{i=1}^n \hat{\imath}(U_i;Z_i)-n(I(U;Z)+\epsilon/3)} \to 0,\\
\sqrt{\mathsf{BA}^{-1}} &= \sqrt{2^{n(I(U;Z)+\epsilon/3)-n(I(U;Z)+\epsilon/2)}} \to 0,
\end{aligned}$$

exponentially as $n \to \infty$, where (a) is by (8) and (b) used $(1 + 2^{-x})^{-2^y} \le 2^{x-y}$.

Similar to the discussions on Theorem 2 recovering asymptotic hiding capacity in Section V, it is left to bound the ϵ-covering numbers $N_{\epsilon_1}(\mathcal{D}), N_{\epsilon_2}(\mathcal{E})$ in Theorem 5. By constructing ϵ-covers of them and by Proposition 4, we know $N_{\epsilon_1}(\mathcal{D}_n) \le N_{\epsilon_1/n}(\mathcal{D}) = O((n/\epsilon)^{|\mathcal{X}|\cdot|\mathcal{Y}|})$ and similarly $N_{\epsilon_2}(\mathcal{E}_n) \le O((n/\epsilon)^{|\mathcal{X}|\cdot|\mathcal{Z}|})$, which grow much slower than the exponential decrease of above terms. Therefore we have $P_e + \nu \cdot \gamma \to 0$ as $n \to \infty$.

VII. Future Applications

We briefly discuss some future directions related to the information hiding problem.

In recent years, with the great success of artificial intelligence, software based on generative models is now able to produce content as realistic as works by human creators. However, the issue of copyright has become controversial: generative models can potentially be trained on public data without obtaining permission from human authors. To protect intellectual property and prohibit the unauthorized use of original works in the training of generative models, practical watermarking tools have been developed [70]–[72] to embed information into human-created works before they are published. However, most existing implementations rely on ad hoc designs that may introduce vulnerabilities [73]. For better design of such tools, it is crucial to develop a solid understanding of the information hiding problem. As an example, a principled design for text watermarking was proposed in [73], following the theoretical analysis of the information hiding problem [1]. Since our one-shot achievability results generalize the asymptotic information hiding problem [1], and can recover the asymptotic hiding capacity even if certain assumptions are dropped, we are thus providing an alternative understanding of the problem that is possibly more fundamental, and it is of interests to explore how can our nonasymptotic information hiding results contribute to potentially improved design of modern watermarking technologies.

Moreover, steganography is closely related to the information hiding problem (see [29]), and an interesting connection between steganography and minimum entropy coupling [74]–[76] has been discussed in [77]. Considering that the coding scheme in this paper is based on the Poisson functional representation [55], which is also related to the minimum entropy coupling (see discussions in [74]), it is worth investigating whether there are deeper connections between minimum entropy coupling and the one-shot information hiding problem studied in this paper.

VIII. Conclusion

In this paper, we derive novel one-shot achievability results for two fundamental secrecy problems in information theory: the information hiding problem [1] and the compound wiretap channel [2]. Our bounds are based on the Poisson matching lemma combined with a covering argument, and they apply to both continuous and discrete settings. For both problems, unlike most existing studies, we do not assume that the decoder knows the channel state in the one-shot setting. The proposed one-shot results hold for arbitrary source distributions and general channel classes (not necessarily memoryless or ergodic), and they recover existing asymptotic results when applied to discrete memoryless channels, potentially under distortion constraints. Thus, our results offer alternative, and potentially simpler, proofs of known capacity theorems. We have also discussed possible generalizations and future applications.

References

[1] P. Moulin and J. A. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Transactions on information theory*, vol. 49, no. 3, pp. 563–593, 2003.

[2] Y. Liang, G. Kramer, and H. V. Poor, "Compound wiretap channels," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, pp. 1–12, 2009.

[3] A. Somekh-Baruch and N. Merhav, "On the error exponent and capacity games of private watermarking systems," *IEEE Transactions on Information Theory*, vol. 49, no. 3, pp. 537–562, 2003.

[4] ——, "On the capacity game of public watermarking systems," *IEEE Transactions on Information Theory*, vol. 50, no. 3, pp. 511–524, 2004.

[5] A. S. Cohen and A. Lapidoth, "The gaussian watermarking game," *IEEE transactions on Information Theory*, vol. 48, no. 6, pp. 1639–1667, 2002.

[6] A. D. Wyner, "The wire-tap channel," *Bell system technical journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[7] H. Ji, S. Park, J. Yeo, Y. Kim, J. Lee, and B. Shim, "Ultra-reliable and low-latency communications in 5g downlink: Physical layer aspects," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 124–130, 2018.

[8] V. Kostina and S. Verdú, "Lossy joint source-channel coding in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2545–2575, 2013.

[9] V. Y. Tan and O. Kosut, "On the dispersions of three network information theory problems," *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 881–903, 2013.

[10] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[11] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 4947–4966, 2009.

[12] S. Verdú, "Non-asymptotic achievability bounds in multiuser information theory," in *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2012, pp. 1–8.

[13] M. H. Yassaee, M. R. Aref, and A. Gohari, "A technique for deriving one-shot achievability results in network information theory," in *2013 IEEE International Symposium on Information Theory*. IEEE, 2013, pp. 1287–1291.

[14] C. T. Li and V. Anantharam, "A unified framework for one-shot achievability via the poisson matching lemma," *IEEE Transactions on Information Theory*, vol. 67, no. 5, pp. 2624–2651, 2021.

[15] Y. Liu and C. T. Li, "One-shot information hiding," in *accepted at the IEEE Information Theory Workshop*. IEEE, 2024.

[16] V. Malik, T. Kargin, V. Kostina, and B. Hassibi, "A distributionally robust approach to shannon limits using the wasserstein distance," in *2024 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2024, pp. 861–866.

[17] D. Blackwell, L. Breiman, A. Thomasian *et al.*, "The capacity of a class of channels," *The Annals of Mathematical Statistics*, vol. 30, no. 4, pp. 1229–1241, 1959.

[18] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Contr. and Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[19] C. Heegard and A. El Gamal, "On the capacity of computer memory with defects," *IEEE transactions on Information theory*, vol. 29, no. 5, pp. 731–739, 1983.

[20] P. Moulin and Y. Wang, "Capacity and random-coding exponents for channel coding with side information," *IEEE Transactions on Information Theory*, vol. 53, no. 4, pp. 1326–1347, 2007.

[21] Y. Steinberg, "Coding for channels with rate-limited side information at the decoder, with applications," *IEEE transactions on information theory*, vol. 54, no. 9, pp. 4283–4295, 2008.

[22] R. J. Barron, B. Chen, and G. W. Wornell, "The duality between information embedding and source coding with side information and some applications," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1159–1180, 2003.

[23] G. Keshet, Y. Steinberg, N. Merhav *et al.*, "Channel coding in the presence of side information," *Foundations and Trends® in Communications and Information Theory*, vol. 4, no. 6, pp. 445–586, 2008.

[24] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE transactions on image processing*, vol. 6, no. 12, pp. 1673–1687, 1997.

[25] D. Boneh and J. Shaw, "Collusion-secure fingerprinting for digital data," *IEEE Transactions on Information Theory*, vol. 44, no. 5, pp. 1897–1905, 1998.

[26] B. Chen and G. W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Transactions on Information theory*, vol. 47, no. 4, pp. 1423–1443, 2001.

[27] N. Merhav, "On random coding error exponents of watermarking systems," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 420–430, 2000.

[28] P. Moulin, "Universal fingerprinting: Capacity and random-coding exponents," in *2008 IEEE International Symposium on Information Theory*. IEEE, 2008, pp. 220–224.

[29] Y. Wang and P. Moulin, "Perfectly secure steganography: Capacity, error exponents, and code constructions," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2706–2722, 2008.

[30] Q. Guan, P. Liu, W. Zhang, W. Lu, and X. Zhang, "Double-layered dual-syndrome trellis codes utilizing channel knowledge for robust steganography," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 501–516, 2022.

[31] W. Li, W. Zhang, L. Li, H. Zhou, and N. Yu, "Designing near-optimal steganographic codes in practice based on polar codes," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 3948–3962, 2020.

[32] Y.-H. Kim, A. Sutivong, and T. M. Cover, "State amplification," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 1850–1859, 2008.

[33] P. Grover, A. B. Wagner, and A. Sahai, "Information embedding and the triple role of control," *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1539–1549, 2015.

[34] Y. Xu, J. Lu, X. Guang, and W. Xu, "Information embedding with stegotext reconstruction," *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 1415–1428, 2023.

[35] T. Kalker and F. M. Willems, "Capacity bounds and constructions for reversible data-hiding," in *2002 14th International Conference on Digital Signal Processing Proceedings. DSP 2002 (Cat. No. 02TH8628)*, vol. 1. IEEE, 2002, pp. 71–76.

[36] Y. Steinberg, "Reversible information embedding with compressed host at the decoder," in *2006 IEEE International Symposium on Information Theory*. IEEE, 2006, pp. 188–191.

[37] O. Sumszyk and Y. Steinberg, "Information embedding with reversible stegotext," in *2009 IEEE International Symposium on Information Theory*. IEEE, 2009, pp. 2728–2732.

[38] A. Sutivong, M. Chiang, T. M. Cover, and Y.-H. Kim, "Channel capacity and state estimation for state-dependent gaussian channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1486–1495, 2005.

[39] I. Bjelaković, H. Boche, and J. Sommerfeld, "Secrecy results for compound wiretap channels," *Problems of Information Transmission*, vol. 49, no. 1, pp. 73–98, 2013.

[40] M. Kobayashi, Y. Liang, S. Shamai, and M. Debbah, "On the compound mimo broadcast channels with confidential messages," in *2009 IEEE International Symposium on Information Theory*. IEEE, 2009, pp. 1283–1287.

[41] T. Liu, V. Prabhakaran, and S. Vishwanath, "The secrecy capacity of a class of parallel gaussian compound wiretap channels," in *2008 IEEE International Symposium on Information Theory*. IEEE, 2008, pp. 116–120.

[42] E. Ekrem and S. Ulukus, "On gaussian mimo compound wiretap channels," in *2010 44th Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2010, pp. 1–6.

[43] ——, "Degraded compound multi-receiver wiretap channels," *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 5681–5698, 2012.

[44] R. F. Schaefer and S. Loyka, "The secrecy capacity of compound gaussian mimo wiretap channels," *IEEE Transactions on Information Theory*, vol. 61, no. 10, pp. 5535–5552, 2015.

[45] H. Boche, R. F. Schaefer, and H. V. Poor, "On the continuity of the secrecy capacity of compound and arbitrarily varying wiretap channels," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2531–2546, 2015.

[46] E. C. Song, P. Cuff, and H. V. Poor, "The likelihood encoder for lossy compression," *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 1836–1849, 2016.

[47] S. Watanabe, S. Kuzuoka, and V. Y. F. Tan, "Nonasymptotic and second-order achievability bounds for coding with side-information," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1574–1605, April 2015.

[48] C. E. Shannon, "A mathematical theory of communication," *Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[49] A. Feinstein, "A new basic theorem of information theory," *IRE Trans. Inf. Theory*, no. 4, pp. 2–22, 1954.

[50] C. E. Shannon, "Certain results in coding theory for noisy channels," *Information and control*, vol. 1, no. 1, pp. 6–25, 1957.

[51] Y. Guo, S. Salehkalaibar, S. C. Draper, and W. Yu, "One-shot achievability region for hypothesis testing with communication constraint," in *accepted at the IEEE Information Theory Workshop*. IEEE, 2024.

[52] H. Hentilä, Y. Y. Shkel, and V. Koivunen, "Communication-constrained secret key generation: Second-order bounds," *IEEE Transactions on Information Theory*, 2024.

[53] A. Khisti, A. Behboodi, G. Cesa, and P. Kumar, "Unequal message protection: One-shot analysis via poisson matching lemma," in *2024 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2024, pp. 629–634.

[54] Y. Liu, S. H. Advary, and C. T. Li, "Nonasymptotic oblivious relaying and variable-length noisy lossy source coding," in *2025 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2025.

[55] C. T. Li and A. El Gamal, "Strong functional representation lemma and applications to coding theorems," *IEEE Transactions on Information Theory*, vol. 64, no. 11, pp. 6967–6978, 2018.

[56] E. Lei, H. Hassani, and S. S. Bidokhti, "Neural estimation of the rate-distortion function with applications to operational source coding," *IEEE Journal on Selected Areas in Information Theory*, vol. 3, no. 4, pp. 674–686, 2022.

[57] Y. Liu, W.-N. Chen, A. Özgür, and C. T. Li, "Universal exact compression of differentially private mechanisms," *Advances in Neural Information Processing Systems*, 2024.

[58] Y. Liu and C. T. Li, "One-shot coding over general noisy networks," in *2024 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2024.

[59] J. Scarlett, "On the dispersions of the gel'fand–pinsker channel and dirty paper coding," *IEEE Transactions on Information Theory*, vol. 61, no. 9, pp. 4569–4586, 2015.

[60] A. C. Berry, "The accuracy of the Gaussian approximation to the sum of independent variates," *Transactions of the American Mathematical Society*, vol. 49, no. 1, pp. 122–136, 1941.

[61] C.-G. Esseen, "On the Liapunov limit error in the theory of probability," *Ark. Mat. Astr. Fys.*, vol. 28, pp. 1–19, 1942.

[62] W. Feller, *An introduction to probability theory and its applications*, 2nd ed. Wiley, New York, 1971, vol. 2.

[63] F. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1079–1107, 1999.

[64] R. Dobrushin, "Optimum information transmission through a channel with unknown parameters," *Radio Eng. Electron*, vol. 4, no. 12, pp. 1–8, 1959.

[65] J. Wolfowitz, *Simultaneous Channels*. New York: Springer-Verlag, 1980.

[66] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1562–1575, 2006.

[67] J. Liu, P. Cuff, and S. Verdú, "e_γ-resolvability," *IEEE Transactions on Information Theory*, vol. 63, no. 5, pp. 2629–2658, 2016.

[68] W. Yang, R. F. Schaefer, and H. V. Poor, "Wiretap channels: Nonasymptotic fundamental limits," *IEEE Transactions on Information Theory*, vol. 65, no. 7, pp. 4069–4093, 2019.

[69] Y. Liang and H. V. Poor, "Multiple-access channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 976–1002, 2008.

[70] S. Shan, J. Cryan, E. Wenger, H. Zheng, R. Hanocka, and B. Y. Zhao, "Glaze: Protecting artists from style mimicry by text-to-image models," *arXiv preprint arXiv:2302.04222*, 2023.

[71] J. Zhu, R. Kaplan, J. Johnson, and L. Fei-Fei, "Hidden: Hiding data with deep networks," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 657–672.

[72] C. Zhang, C. Lin, P. Benz, K. Chen, W. Zhang, and I. S. Kweon, "A brief survey on deep learning based data hiding," *arXiv preprint arXiv:2103.01607*, 2021.

[73] Z. Ji, Q. Hu, Y. Zheng, L. Xiang, and X. Wang, "A principled approach to natural language watermarking," in *ACM Multimedia 2024*.

[74] C. T. Li, "Efficient approximate minimum entropy coupling of multiple probability distributions," *arXiv preprint arXiv:2006.07955*, 2020.

[75] F. Cicalese, L. Gargano, and U. Vaccaro, "Minimum-entropy couplings and their applications," *IEEE Trans. Inf. Theory*, vol. 65, no. 6, pp. 3436–3451, 2019.

[76] S. Compton, D. Katz, B. Qi, K. Greenewald, and M. Kocaoglu, "Minimum-entropy coupling approximation guarantees beyond the majorization barrier," in *International Conference on Artificial Intelligence and Statistics*. PMLR, 2023, pp. 10 445–10 469.

[77] C. S. de Witt, S. Sokota, J. Z. Kolter, J. Foerster, and M. Strohmeier, "Perfectly secure steganography using minimum entropy coupling," *arXiv preprint arXiv:2210.14889*, 2022.