On Optimal Information-Theoretic Security in Symmetric Encryption under Low-Entropy Keys

Haibo Cheng[1,2], Haijie Su[2], Dongyi Li[2], Wenting Li[3], Ping Wang[2], Kaitai Liang[4]

[1]Minzu University of China [2]Peking University
[3]Beijing Institute of Graphic Communication [4]University of Turku

Abstract. Low-entropy keys such as passwords and biometrics are pervasive, yet classical information-theoretic guarantees, such as perfect secrecy and entropic security, cannot generally be achieved in this regime. Honey encryption (HE) and homophonic ciphers (HC) tailor encryption to the message distribution to provide meaningful security, but their analyses do not establish optimal guarantees for arbitrary distributions. We ask: for messages and keys drawn independently from arbitrary distributions $(p_{\mathrm{m}}, p_{\mathrm{k}})$, what is the optimal information-theoretic security achievable by any symmetric encryption scheme, and which conditions characterize its attainment?

For key confidentiality (KC), the optimal guarantee is that the ciphertext C reveals only negligible information about the key K, i.e., $I(K;C) = \mathrm{negl}(\ell)$. For uniform schemes, this is achieved if and only if decrypting under every fixed key yields a distribution statistically close to p_{m}. HC and HE achieve this guarantee. For message-recovery security (MRS), the optimal guarantee is that no adversary guesses the message with probability above $p_{\max} + \mathrm{negl}(\ell)$, while a trivial adversary achieves at least $p_{\max}$ by guessing the most likely message or decrypting under the most likely key. We construct MRS-OE, a uniform scheme tailored to $(p_{\mathrm{m}}, p_{\mathrm{k}})$ that attains $p_{\max} + O(2^{-\ell})$, and prove that p_{k}-agnostic schemes, including HC and HE, cannot achieve optimal MRS in general. We also show that optimal KC and optimal MRS generally cannot be achieved simultaneously. Technically, we establish the KC characterization through flow and cut analysis on a weighted directed acyclic graph. We also introduce a continuous-ciphertext framework that separates structural constraints from discretization error.

1 Introduction

Low-entropy keys arise pervasively in practice, from user-chosen passwords [33] to biometrically derived secrets [19, 24], and they enable powerful offline guessing once ciphertexts leak. Empirical studies estimate password min-entropy at only 7 bits [34, 9], and deep learning-based methods further improve password guessing [21, 44]. Recent attacks also show that even large biometric dictionaries can succeed with only tens of thousands guesses [12]. In long-term confidentiality settings, this problem only worsens: advances in computation can turn "today's

high-entropy keys" into "tomorrow's low-entropy keys", enabling harvest-now-decrypt-later attacks.

Computationally intensive key-derivation functions, including memory-hard constructions, can increase the cost of each key guess [33, 3, 5], while fresh salts [33] and unpredictable nonces [25] can limit the reuse of precomputed work. These measures do not prevent offline guessing once an attacker obtains a ciphertext, so conventional symmetric encryption does not address this threat in an information-theoretic sense. When decrypting under an incorrect key typically yields an explicit failure symbol or obvious garbage, an offline attacker can enumerate candidate keys and identify the unique key that produces a plausible message. Exhaustive-guessing resistance therefore requires *ambiguity*: a ciphertext should admit many plausible decryptions, making correct and wrong-key outputs statistically hard to distinguish.

We study this question in a distributional model where messages M and keys K are drawn independently from given distributions $(p_{\mathrm{m}}, p_{\mathrm{k}})$ over finite spaces $\mathcal{M}$ and $\mathcal{K}$. Consider the extreme case $|\mathcal{K}| = 2$: any scheme that hopes to provide information-theoretic security must ensure that decrypting under the wrong key yields a *plausible* message. Consequently, no universal, distribution-agnostic encryption scheme, such as a conventional AES-based design, can provide satisfactory guarantees for every message distribution.

Prior work on homophonic ciphers (HC) [36, 41, 30] and honey encryption (HE) [31, 29, 35] provides an important clue: by randomizing encryption in a way tailored to the message distribution, *wrong-key decryption can be made to resemble sampling from* p_{m}. At a high level, these schemes can be viewed under a unified construction. *Encryption* first encodes a non-uniform message into a nearly uniform seed using a distribution-dependent randomized encoder, and then applies a conventional cipher to encrypt the seed. *Decryption* first recovers a seed under the given key, and then decodes it by using the seed as the randomness for sampling from p_{m}. The main difference between HC and HE is that HC typically relies on variable-length randomized encodings to achieve perfect uniformity of the seed, whereas HE works with fixed-length ciphertexts for near-uniformity.

Existing analyses of HC focus primarily on key confidentiality (KC) via Shannon-style key equivocation $H(K \mid C)$ [36, 41, 30], while HE studies focus on message confidentiality (MC) notions such as message-recovery security (MRS) [31] and target-distribution semantic security / non-malleability (TDSS/TDNM) [29, 35]. However, HC analyses are carried out in an infinite-sequence setting, which can obscure leakage through the (variable) ciphertext length when one encrypts a single message, and they do not quantify the impact of using finite-length (hence imperfect) uniformization on KC. On the HE side, existing bounds for MRS/TDSS/TDNM are derived for particular constructions under *specific* message and key distributions and are not tight for arbitrary $(p_{\mathrm{m}}, p_{\mathrm{k}})$. Moreover, HE notions largely emphasize MC, while KC—crucial when keys are sensitive or reused across systems, as with passwords and

biometrics—has received comparatively less attention. A detailed analysis is deferred to Section 2.

In this study, we investigate both KC and MC in a general framework and ask, for arbitrary message and key distributions $(p_{\mathrm{m}}, p_{\mathrm{k}})$, what is the *optimal achievable* information-theoretic security over all (possibly probabilistic) symmetric encryption schemes—beyond the HC/HE paradigm—and which necessary and/or sufficient conditions characterize when this level can be achieved?

Symmetric encryption. We use the following simple formalism for (probabilistic) symmetric encryption. A scheme E is a tuple $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \mathcal{E}, \mathcal{D})$, where $\mathcal{M}, \mathcal{K}$ are the given message and key spaces, and $\mathcal{C}$ is the finite ciphertext space; $\mathcal{E} := \{E_k \mid k \in \mathcal{K}\}$, where each E_k is a (possibly randomized) encryption algorithm that takes $m \in \mathcal{M}$ and outputs $c \in \mathcal{C}$; and $\mathcal{D} := \{D_k \mid k \in \mathcal{K}\}$, where each D_k is a deterministic decryption algorithm that takes $c \in \mathcal{C}$ and outputs either a message in $\mathcal{M}$ or a failure symbol $\perp \notin \mathcal{M}$ (we write $\mathcal{M}' := \mathcal{M} \cup \{\perp\}$). Correctness requires that for every $k \in \mathcal{K}$ and $m \in \mathcal{M}$,

$$\Pr[\, D_k(C) = m \;:\; C \leftarrow E_k(m) \,] = 1,$$

where the probability is over the internal randomness of E_k.

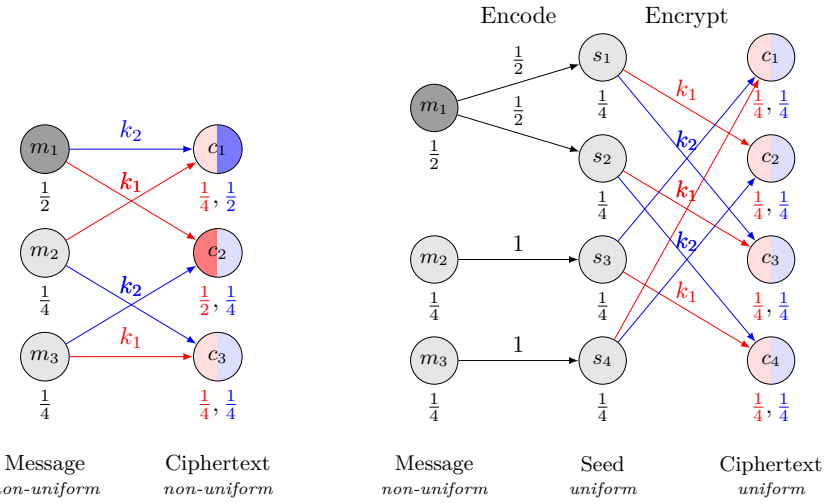
For each message–key pair (m, k), define the (m, k)-*ciphertext region* and its size by

$$\mathcal{C}(m, k) := \{c \in \mathcal{C} \mid D_k(c) = m\}, \qquad \mu(m, k) := |\mathcal{C}(m, k)|.$$

A probabilistic encryption algorithm E_k induces a conditional distribution $p_{\mathrm{c|m,k}}(\cdot \mid m, k)$ over $\mathcal{C}(m, k)$. In general, $p_{\mathrm{c|m,k}}(\cdot \mid m, k)$ may be pathological or degenerate; for instance, its probability mass may concentrate on a single ciphertext[1]. We focus mainly on *uniform* probabilistic schemes, which are uniform over each ciphertext region, i.e., $p_{\mathrm{c|m,k}}(c \mid m, k) = 1/\mu(m, k)$ for every $c \in \mathcal{C}(m, k)$. Uniform sampling is straightforward to implement over interval regions and entails no loss of optimality for the main security notions studied here, as we later show that uniform schemes attain the global optima.

HC and HE through ciphertext regions. With this formalization, the common mechanism behind HC and HE becomes clear. As illustrated in Fig. 1, HC and HE employ a *uniform* construction that spreads the probability mass $p_{\mathrm{m}}(m)$ over $\mathcal{C}(m, k)$. In deterministic encryption and conventional message-independent nonce/IV randomization, $\mu(m, k)$ is typically independent of m—equal to 1 in the deterministic case, or to the size of the nonce/IV space—so the non-uniformity of p_{m} is preserved. HC/HE instead choose $\mu(m, k) \propto p_{\mathrm{m}}(m)$, assigning larger ciphertext regions to more likely messages and making the per-ciphertext probability $p_{\mathrm{m}}(m)/\mu(m, k)$ nearly constant. Thus, $p_{\mathrm{c|k}}(\cdot \mid k)$ becomes nearly uniform for every key k, and decrypting an HC/HE ciphertext under a wrong key yields a message approximately distributed according to p_{m}.

[1] Correctness only guarantees $\mathrm{Supp}(E_k(m)) \subseteq \mathcal{C}(m, k)$. Since D_k may be set to $\perp$ outside the encryption supports without affecting honest decryption, KC, or MRS, we assume throughout that $\mathcal{C}(m, k) = \mathrm{Supp}(E_k(m))$ without loss of generality.



(a) Deterministic cipher with $\mu(m,k) = 1$.

(b) Homophonic cipher (HC) / honey encryption (HE) with $\mu(m,k) \propto p_{\mathrm{m}}(m)$.

Fig. 1: **Probability-mass spreading in HC and HE.** Node intensity indicates probability mass; red and blue denote the ciphertext distributions under k_1 and k_2, respectively. A deterministic cipher permutes the message mass, whereas HC/HE spread each message over a seed region proportional to $p_{\mathrm{m}}(m)$, inducing a uniform ciphertext distribution under either key.

Ciphertext length as the security parameter. Once $(\mathcal{M}, \mathcal{K}, p_{\mathrm{m}}, p_{\mathrm{k}})$ are fixed, the parameter controlling the approximation to the information-theoretic optimum is the ciphertext length, as in HE/HC. We consider fixed-length ciphertexts $C \in \mathcal{C}^{(\ell)} = \{0,1\}^{\ell}$, where ℓ is the ciphertext length and $|\mathcal{C}^{(\ell)}| = 2^{\ell}$.[2] Increasing ℓ does not enlarge the key space or increase the cost of exhaustive key search; rather, it provides more ciphertext points over which probability mass can be spread, reducing the discretization error to $O(2^{-\ell})$. It incurs the corresponding cost of longer ciphertexts, more randomness, and higher arithmetic precision. We use ℓ as the security parameter and write negligible terms as functions of ℓ.

Optimal attainable security and tight conditions. With this formalism in place, we analyze the optimal information-theoretic security achievable by *any* (possibly probabilistic) symmetric encryption scheme under given message and key distributions $(p_{\mathrm{m}}, p_{\mathrm{k}})$. Since finite-bit sampling may introduce approximation error, throughout the paper, "optimal" means achieving the optimum up to a negligible additive term in ℓ. We consider both key confidentiality (KC) and

[2] Classical HC is often formulated with an infinite ciphertext space $\mathcal{C} = \{0,1\}^{+} = \bigcup_{n \geq 1} \{0,1\}^{n}$, and thus does not literally fit our definition. Nevertheless, it can be viewed as the limit of a sequence of fixed-length schemes as $\ell \to \infty$.

message confidentiality (MC), and derive *tight* (necessary-and-sufficient) structural conditions under which these optima are achievable.

Key confidentiality (KC). We quantify key leakage via mutual information $I(K;C) = H(K) - H(K\,|\,C)$, which is the standard information-theoretic counterpart of Shannon-style key equivocation [39, 30]. We show that the optimal attainable level is $I(K;C) = \mathrm{negl}(\ell)$, i.e., the ciphertext reveals only negligible information about the key; in particular, this implies strong guessing-based notions of KC. Moreover, we prove a tight characterization: $I(K;C) = \mathrm{negl}(\ell)$ holds if and only if ciphertext distributions are statistically indistinguishable across keys. Equivalently, $C \overset{s}{\equiv} E_k(M)$ for every $k \in \mathcal{K}$, where $M \leftarrow_{p_\mathrm{m}} \mathcal{M}$, $C \leftarrow_{p_\mathrm{c}} \mathcal{C}$ (equivalently, $M, M' \leftarrow_{p_\mathrm{m}} \mathcal{M}$, $K' \leftarrow_{p_\mathrm{k}} \mathcal{K}$, $C \leftarrow E_{K'}(M')$), and $X \overset{s}{\equiv} Y$ denotes that the statistical distance between random variables X and Y is negligible in the security parameter ℓ. For *uniform* schemes, this ciphertext indistinguishability is further equivalent to the message indistinguishability, i.e., $M \overset{s}{\equiv} D_k(C)$ for every $k \in \mathcal{K}$, which is the HC/HE design goal that decrypting under any key yields messages distributed (approximately) as p_m. This demonstrates that the structural principle behind HC/HE gives *tight* conditions for achieving optimal KC. We also bound key leakage from multiple ciphertexts encrypted under the same key, allowing correlations among messages.

The main technical challenge is to prove that message indistinguishability implies ciphertext indistinguishability for uniform schemes. Message-distribution errors measure only the net probability difference in each decryption region, where positive and negative differences can cancel. We represent density differences between overlapping regions as flows on a weighted acyclic graph, with message errors corresponding to net node flows. A cut-counting argument then bounds the total edge flow and hence the ciphertext distance, with a coefficient independent of the ciphertext space size.

Message confidentiality (MC). We focus on message-recovery security (MRS), measuring the adversary's success probability in recovering M from C. We show that the optimal bound over all schemes is $p_{\max} + \mathrm{negl}(\ell)$, where $p_\mathrm{m}^{\max} := \max_m p_\mathrm{m}(m)$, $p_\mathrm{k}^{\max} := \max_k p_\mathrm{k}(k)$, and $p_{\max} := \max\{p_\mathrm{m}^{\max}, p_\mathrm{k}^{\max}\}$. The trivial guess $m_{\mathrm{triv}}(c)$ outputs the most likely message if $p_\mathrm{m}^{\max} \geq p_\mathrm{k}^{\max}$, and otherwise decrypts under the most likely key, succeeding with probability at least $p_{\max}$.

We construct a uniform scheme, called MRS-optimal encryption (MRS-OE) and denoted by OE for brevity, achieving $p_{\mathrm{guess}}(M \mid C) = p_{\max} + O(2^{-\ell})$. Using the cumulative distribution functions (CDFs) of p_m and p_k, it arranges ciphertext regions with $\mu(m,k) \propto \min\{p_\mathrm{m}(m), p_\mathrm{k}(k)\}$, up to finite-precision loss. The sizes dilute each pair's probability mass, while the placement ensures correctness and avoids *decryption collisions*, where distinct keys decrypt the same ciphertext to the same message. This gives $p_{\mathrm{m,c}}(m,c) \leq 2^{-\ell} p_{\max} + O(2^{-2\ell})$, and hence $p_{\mathrm{guess}}(M \mid C) = \sum_c \max_m p_{\mathrm{m,c}}(m,c) \leq p_{\max} + O(2^{-\ell})$.

OE computes its ciphertext regions directly from the message and key CDFs. Its storage and per-encryption/decryption costs are therefore additive in those of the two CDF representations, apart from simple interval arithmetic and sampling. With explicit sorted CDF tables, this gives $O(\ell|\mathcal{M}| + \ell|\mathcal{K}|)$ bits of storage,

$O(\ell + \log|\mathcal{M}| + \log|\mathcal{K}|)$ encryption time, and $O(\ell \log|\mathcal{M}| + \log|\mathcal{K}|)$ decryption time. For structured probability models, the tables can instead be replaced by model-provided CDF and inverse-CDF operations.

For the stronger notion TDSS (capturing partial-information leakage), negligible advantage cannot be guaranteed for arbitrary message and key distributions; determining the optimal attainable bound in general remains open.

A continuous-ciphertext framework. To cleanly propagate discretization-induced negligible errors throughout our bounds and tight conditions, we introduce a continuous-ciphertext abstraction over $[0, 1)$ that separates structural constraints (ciphertext allocation and collision avoidance) from discretization error. We establish two fundamental transformations: (i) any continuous scheme can be discretized to $\{0,1\}^\ell$ with KC/MC degradation at most $O(2^{-\ell})$, and (ii) any sequence of discrete schemes admits subsequences that converge to a continuous limit with convergent security metrics. This enables us to reason in the continuous domain without precision artifacts, while retaining tight discrete guarantees up to negligible error.

Within this framework, we derive several necessary conditions for optimal security, implying three impossibility results for general message and key distributions: (1) except for trivial schemes that reveal the message, optimal KC is impossible for p_{m}-agnostic schemes; (2) optimal MRS is impossible for p_{k}-agnostic schemes (including HC/HE schemes); and (3) simultaneously achieving optimal KC and optimal MRS is impossible. For (3), we give a variant of HC/HE that achieves optimal KC and, when $p_{\mathrm{m}}^{\max} \le 1/|\mathcal{K}|$, also achieves optimal MRS.

Practicality of distribution-tailored encryption. Tailoring to an exact pair $(p_{\mathrm{m}}, p_{\mathrm{k}})$ may appear impractical when these distributions are complex or only partially known. Nevertheless, prior HE applications—including honey password vaults [11, 23, 14, 13] and long-term genetic protection [26]—show that tailored encryption can be deployed with approximate statistical models of the underlying data. More recently, work on honey password vaults [13] shows that modern generative models (e.g., Transformer-based models [42]) can learn realistic low-entropy password distributions, and that even imperfect models can yield practical and meaningful security gains. In practice, the message and key distributions used to design a scheme are estimates and generally differ from the true distributions. Our continuous analysis bounds the resulting security loss in terms of the statistical distances between the estimated and true distributions. More broadly, as deep generative models continue to improve, we expect increasingly accurate learning of complex distributions across domains (text, images, and structured data), and deployment settings where the data owner specifies message/key types while the encryption mechanism selects an appropriate learned model.

2 Related Works

Information-theoretic notions and entropy requirements. Shannon introduced key and message equivocation, $H(K \mid C)$ and $H(M \mid C)$, as funda-

mental measures for information-theoretic security of encryption schemes [39]. In general, $H(K \mid C) \geq H(M \mid C)$, and equality holds for schemes without decryption collisions, i.e., for every pair (m, c) there is at most one key k such that $D_k(c) = m$. He defined *perfect secrecy* by $H(M \mid C) = H(M)$, equivalently $I(M; C) = 0$, capturing the ideal guarantee that C leaks no information about M. However, perfect secrecy requires key entropy comparable to message entropy, namely $H(K) \geq H(M)$. More generally, even relaxing perfect secrecy to $I(M; C) \leq \epsilon$ still requires $H(K) \geq H(M) - \epsilon$ [28]. Entropic security [38, 20] further relaxes perfect secrecy by bounding the adversary's advantage in predicting any function of the message from the ciphertext by the corresponding trivial (prior-only) baseline plus ϵ. For an n-bit message distribution with min-entropy t, achieving entropic security for *arbitrary* such distributions still requires essentially $H(K) \geq n - t$ [20]. These entropy requirements make perfect secrecy and entropic security typically unattainable under low-entropy keys.

Equivocation bounds for deterministic ciphers. A large body of work derives bounds on key/message equivocation for deterministic encryption with uniform keys, including Shannon's pure and random ciphers and classical substitution ciphers [39, 6, 22, 7, 10, 2]. These works typically consider a single key K used to encrypt a sequence of messages M^n and quantify how $H(K \mid C^n)$ or $H(M^n \mid C^n)$ decays as $n \to \infty$, yielding precise impossibility results for maintaining large equivocation in these deterministic schemes. More broadly, such schemes are distribution-agnostic (i.e., not tailored to a particular message distribution), and as argued in the Introduction, such designs generally cannot provide satisfactory information-theoretic security under low-entropy keys.

Distribution-tailored probabilistic encryption. In contrast, homophonic ciphers (HC) [41, 36, 30] and honey encryption (HE) [31, 29] exploit *probabilistic encryption* and tailor the randomization to the message distribution so as to create many plausible decryptions. At a high level, both share an *encode-then-encrypt* structure: a distribution-dependent randomized encoder maps m to a (nearly) uniform seed, a conventional cipher encrypts the seed, and a corresponding decoder maps the decrypted seed back to a message so that decrypting under wrong keys resembles sampling from the target message distribution.

HC. Classical HC analyses often allow variable-length and infinite ciphertext spaces to achieve perfect uniformization, and then argue strong key-equivocation guarantees such as $H(K \mid C^n) = H(K)$ for long/infinite message sequences [36, 30]. However, in a single-message setting (as considered in this paper), ciphertext length is directly observable; these works do not formally analyze $H(K \mid C)$ in this regime and do not quantify the key leakage induced by variable ciphertext length. In particular, they cannot in general guarantee perfect key secrecy, i.e., $I(K; C) = 0$ (equivalently $H(K \mid C) = H(K)$). When the idealized infinite-length construction is approximated by finite-length ciphertexts, they also do not quantify the resulting deviation from perfect uniformization (see Appendix A).

HE. HE was proposed with fixed-length ciphertexts explicitly for low-entropy keys [31]. Its cryptographic analysis focuses on three message-confidentiality notions: (1) message-recovery security (MRS) [31], measuring the adversary's suc-

cess probability of recovering the *entire* message; (2) target-distribution semantic security (TDSS) [29], measuring leakage of *partial information* about the message (via predicates); and (3) target-distribution non-malleability (TDNM) [29], measuring resistance to *meaningful ciphertext manipulation*. Corresponding bounds for this fixed-length construction are proved in the random-oracle / ideal-cipher models [31, 29], under specific distributional assumptions. TDSS bounds have also been proved in the standard model using pairwise-independent hashing [35]. These bounds are not tight for general message and key distributions (see Appendix B).

Limitations of prior HC/HE work. Overall, prior HC and HE studies leave open: (1) an information-theoretic characterization of the *optimal attainable* guarantees (and tight conditions) across *all* probabilistic symmetric encryption schemes under arbitrary $(p_{\mathrm{m}}, p_{\mathrm{k}})$; and (2) a formal analysis of key confidentiality, which is important when keys are sensitive or reused across systems (as with passwords and biometrics), but has received comparatively less attention.

3 Notation and Definitions

Notation. Let $\mathcal{M}$ denote the message space with distribution p_{m} and $\mathcal{K}$ the key space with distribution p_{k}. For a distribution p_{x} over a set $\mathcal{X}$, we write $X \leftarrow_{p_{\mathrm{x}}} \mathcal{X}$ to sample X from $\mathcal{X}$ according to p_{x}, and write $X \leftarrow_{\$} \mathcal{X}$ when p_{x} is uniform. For $\mathcal{X}' \subseteq \mathcal{X}$, we define $p(\mathcal{X}') \coloneqq \sum_{x \in \mathcal{X}'} p(x)$. Throughout, script letters (e.g., $\mathcal{M}$) denote spaces, italic capitals (e.g., $M \leftarrow_{p_{\mathrm{m}}} \mathcal{M}$) denote random variables, and italic lowercase letters (e.g., m) denote concrete values. Fix a specified order $\preceq$ on each finite set $\mathcal{X}$. We define $p_{\mathrm{x}}^{\min} \coloneqq \min_x p_{\mathrm{x}}(x)$ and $p_{\mathrm{x}}^{\max} \coloneqq \max_x p_{\mathrm{x}}(x)$, and let $x^\star \in \arg\max_{x \in \mathcal{X}} p_{\mathrm{x}}(x)$ denote an arbitrary maximum-probability element. Thus $p_{\mathrm{x}}^{\max} = p_{\mathrm{x}}(x^\star)$. Whenever a unique representative is required, ties are resolved by the specified order $\preceq$. We define $p_{\mathrm{x,y}}(x, y) \coloneqq \Pr[X = x, Y = y]$ and $p_{\mathrm{x|y}}(x \mid y) \coloneqq \Pr[X = x \mid Y = y]$, and analogously for multiple variables. For p_{x} over $\mathcal{X}$, we denote the cumulative distribution function (CDF) by $F_{\mathrm{x}}(x) \coloneqq \sum_{x' \preceq x} p_{\mathrm{x}}(x')$ and let x^- denote the immediate predecessor of x. If x is the first element, set $x^- \coloneqq \perp_{\mathrm{s}}$ and extend the CDF by $F_{\mathrm{x}}(\perp_{\mathrm{s}}) \coloneqq 0$. Then, $F_{\mathrm{x}}(x^-) \coloneqq F_{\mathrm{x}}(x) - p_{\mathrm{x}}(x) = \sum_{x' \prec x} p_{\mathrm{x}}(x')$. The *CDF interval* associated with x is $\mathcal{I}_{\mathrm{x}}(x) \coloneqq [F_{\mathrm{x}}(x^-), F_{\mathrm{x}}(x))$, whose length is $p_{\mathrm{x}}(x)$. For convenience, we assume that p_{m} and p_{k} are strictly positive over $\mathcal{M}$ and $\mathcal{K}$.

Entropy and statistical distances. For a random variable X with distribution p, its Shannon entropy is $H(X) \coloneqq -\sum_x p(x) \log p(x)$. For jointly distributed X, Y, the mutual information is $I(X;Y) \coloneqq \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} = H(X) - H(X \mid Y)$. We use base-2 log and natural ln. For random variables X_0, X_1 over the same space with distributions p_0, p_1, the statistical distance is $\mathrm{SD}(X_0, X_1) \coloneqq \frac{1}{2} \sum_x |p_0(x) - p_1(x)|$, and the Kullback–Leibler divergence is $\mathrm{KL}(X_0 \| X_1) \coloneqq \sum_x p_0(x) \log \frac{p_0(x)}{p_1(x)}$.

Negligibility and statistical indistinguishability. A function $\varepsilon(\ell)$ is *negligible* in ℓ, denoted $\varepsilon(\ell) = \mathrm{negl}(\ell)$, if for every positive polynomial $poly(\cdot)$

there exists ℓ_0 such that for all $\ell > \ell_0$, $|\varepsilon(\ell)| < 1/poly(\ell)$. Two sequences of random variables $\{X_0^{(\ell)}\}_\ell$ and $\{X_1^{(\ell)}\}_\ell$ are *statistically indistinguishable* if $\mathrm{SD}(X_0^{(\ell)}, X_1^{(\ell)}) = \mathrm{negl}(\ell)$, denoted $X_0^{(\ell)} \overset{s}{\equiv} X_1^{(\ell)}$; X and Y are *perfectly indistinguishable*, denoted $X \equiv Y$, if the X and Y have identical distributions.

Quantization. The negligible terms throughout this paper arise primarily from the finite-precision loss introduced by quantization. For an ℓ-bit discrete space, define $Q_\ell(u) \coloneqq \lfloor 2^\ell u \rfloor$[3] For a CDF F_{x}, define

$$\widetilde{F}_{\mathrm{x}}^{(\ell)}(x) \coloneqq Q_\ell(F_{\mathrm{x}}(x)), \qquad F_{\mathrm{x}}^{(\ell)}(x) \coloneqq 2^{-\ell}\widetilde{F}_{\mathrm{x}}^{(\ell)}(x),$$
$$\widetilde{p}_{\mathrm{x}}^{(\ell)}(x) \coloneqq \widetilde{F}_{\mathrm{x}}^{(\ell)}(x) - \widetilde{F}_{\mathrm{x}}^{(\ell)}(x^-), \quad p_{\mathrm{x}}^{(\ell)}(x) \coloneqq 2^{-\ell}\widetilde{p}_{\mathrm{x}}^{(\ell)}(x).$$

The tilded quantities are integer boundary indices and point counts; the untilded quantities are their normalized counterparts. Quantizing the endpoints of $\mathcal{I}_{\mathrm{x}}(x)$ gives the *discrete CDF interval* $\mathcal{I}_{\mathrm{x}}^{(\ell)}(x) \coloneqq [\widetilde{F}_{\mathrm{x}}^{(\ell)}(x^-), \widetilde{F}_{\mathrm{x}}^{(\ell)}(x)) \cap \mathbb{Z}$, with $|\mathcal{I}_{\mathrm{x}}^{(\ell)}(x)| = \widetilde{p}_{\mathrm{x}}^{(\ell)}(x)$. Moreover, $|F_{\mathrm{x}}^{(\ell)}(x) - F_{\mathrm{x}}(x)| < 2^{-\ell}$ and $|p_{\mathrm{x}}^{(\ell)}(x) - p_{\mathrm{x}}(x)| < 2^{-\ell}$. If $X \leftarrow_{p_{\mathrm{x}}} \mathcal{X}$ and $X^{(\ell)} \leftarrow_{p_{\mathrm{x}}^{(\ell)}} \mathcal{X}$, then $\mathrm{SD}(X^{(\ell)}, X) = O(|\mathcal{X}|2^{-\ell}) = \mathrm{negl}(\ell)$, so $X^{(\ell)} \overset{s}{\equiv} X$.

Encryption schemes indexed by ℓ. We treat the ciphertext length ℓ as the security parameter and view an encryption scheme as a family $(\mathcal{M}, \mathcal{K}, \mathcal{C}^{(\ell)}, \mathcal{E}^{(\ell)}, \mathcal{D}^{(\ell)})$ with $\mathcal{C}^{(\ell)} \coloneqq \{0,1\}^\ell = \mathbb{Z}_{2^\ell}$, $\mathcal{E}^{(\ell)} \coloneqq \{E_k^{(\ell)}\}_{k\in\mathcal{K}}$ and $\mathcal{D}^{(\ell)} \coloneqq \{D_k^{(\ell)}\}_{k\in\mathcal{K}}$. When the context is clear, we omit the superscript (ℓ).

Ciphertext regions and uniformity. We use the ciphertext regions $\mathcal{C}(m,k)$ and their sizes $\mu(m,k)$ as defined in the Introduction. A probabilistic encryption algorithm E_k induces a conditional distribution $p_{\mathrm{c|m,k}}(\cdot \mid m,k)$ on $\mathcal{C}(m,k)$. We call a scheme *uniform* if it is uniform on each ciphertext region, i.e., $p_{\mathrm{c|m,k}}(c \mid m,k) = 1/\mu(m,k)$ for all $c \in \mathcal{C}(m,k)$.

Structural properties and classes of schemes. We further distinguish (probabilistic) schemes by how the ciphertext-region measures $\mu(m,k)$ vary:

1. **Message independence (MID).** For each key k, the value $\mu(m,k)$ is constant over $m \in \mathcal{M}$.
2. **Key independence (KID).** For each message m, the value $\mu(m,k)$ is constant over $k \in \mathcal{K}$.

We write **MD** (message-dependent) and **KD** (key-dependent) for the negations of MID and KID, respectively.

Deterministic schemes have $\mu(m,k) = 1$ for all (m,k), and thus satisfy both MID and KID. Standard probabilistic schemes that append a fixed-length nonce/IV also satisfy both MID and KID, since $\mu(m,k)$ equals the nonce/IV space size. In contrast, HC and HE schemes are MD and KID, as they tailor $\mu(m,k)$ to p_{m} but not p_{k}. To the best of our knowledge, no prior work has studied KD schemes.

We use the following shorthand for spaces of schemes: Ω_{E} for all symmetric encryption schemes; Ω_{DE} and Ω_{PE} for deterministic and probabilistic schemes,

[3] Other works use alternative quantizers, such as rounding to the nearest integer [31]. The precise error constants differ slightly, but the normalized error remains $O(2^{-\ell})$.

respectively (we treat deterministic schemes as a special case of probabilistic schemes such that $\Omega_{\mathrm{DE}} \subseteq \Omega_{\mathrm{PE}} = \Omega_{\mathrm{E}}$); and Ω_{EE} and Ω_{HE} for encode-then-encrypt constructions and HE schemes (defined below). For probabilistic encryption, we also consider spaces Ω^{P}_{PE} defined by attributes $P \in \{\mathrm{U}, \mathrm{MID}, \mathrm{MD}, \mathrm{KID}, \mathrm{KD}\}$ (or combinations thereof), where **U** denotes uniformity of encryption algorithms. Appendix N summarizes the optimal attainable security in different spaces of encryption schemes and compares these optima with the global optimum over the whole space Ω_{E}.

Encode-then-encrypt (EE) constructions. We formalize the *EE* construction as a combination of an encoder $(\mathtt{encode}, \mathtt{decode})$ and a standard cipher $(\mathtt{encrypt}, \mathtt{decrypt})$ with $E_k \coloneqq \mathtt{encrypt}_k \circ \mathtt{encode}$ and $D_k \coloneqq \mathtt{decode} \circ \mathtt{decrypt}_k$. Here $\mathtt{encode}$ is possibly randomized and maps $m \in \mathcal{M}$ to $s \in \mathcal{S}$. The decoder $\mathtt{decode}$ maps every seed in $\mathcal{S}$ to a message in $\mathcal{M}$. Correctness requires $\mathtt{decode}(\mathtt{encode}(m)) = m$ with probability 1. Moreover, we focus on *uniform* encoders: for each m, the random variable $\mathtt{encode}(m)$ is supported on $\mathtt{decode}^{-1}(m) \coloneqq \{s \mid \mathtt{decode}(s) = m\}$ and uniform over this set, i.e., $\Pr[\mathtt{encode}(m) = s] = 1/|\mathtt{decode}^{-1}(m)|$ for all $s \in \mathtt{decode}^{-1}(m)$. For each key k, the cipher is modeled as a bijection $\mathtt{encrypt}_k : \mathcal{S} \to \mathcal{C}$, with $\mathtt{decrypt}_k = \mathtt{encrypt}_k^{-1}$. Fixed-length, message- and key-independent randomness r, such as a nonce or IV included in the ciphertext, can be incorporated into an augmented seed (s, r): the modified encoder samples r independently, while $\mathtt{decode}'(s, r) = \mathtt{decode}(s)$. Thus, standard nonce/IV-based randomized ciphers are represented as deterministic bijections on the augmented seed space.

In fact, an encryption scheme can be presented as an EE construction if and only if it is uniform and KID, and $D_k(c) \neq \perp$ for every $k \in \mathcal{K}$ and $c \in \mathcal{C}$. Its region sizes may depend on the message or degenerate to the MID case (see Appendix C).

Homophonic ciphers and honey encryption. Homophonic ciphers (HC) and honey encryption (HE) can be viewed as EE constructions whose encoder aims to (approximately) *uniformize* a non-uniform message distribution by mapping messages to (nearly) uniform seeds. As discussed before, HC with infinite ciphertext space can be viewed as a limit of ℓ-length HE schemes as $\ell \to \infty$. In the rest of the paper, we do not distinguish HC and HE.

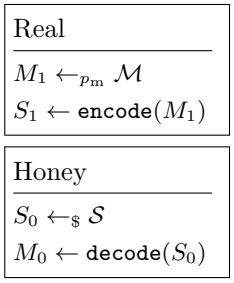


Fig. 2: Indistinguishability of real and honey message-seed pairs (**MS-IND**) for encoders: $(M_1, S_1) \overset{s}{\equiv} (M_0, S_0)$

To formalize this encoder requirement, we recall message–seed indistinguishability (MS-IND) [31]: the *real* pair (M_1, S_1) and the *honey* pair (M_0, S_0) in Fig. 2 are statistically indistinguishable. When their statistical distance is zero, we call this *perfect* MS-IND. The induced honey distribution [31] is $p_{\mathrm{d}}(m) \coloneqq |\mathtt{decode}^{-1}(m)|/|\mathcal{S}|$ for $m \in \mathcal{M}'$.

Because S is uniform over the finite seed space $\mathcal{S}$, every probability $p_{\mathrm{d}}(m)$ is an integer multiple of $1/|\mathcal{S}|$ (in particular, of $2^{-\ell}$ when $|\mathcal{S}| = 2^{\ell}$). Thus, exact MS-IND may be unattainable for an arbitrary p_{m} at a fixed ℓ. HE instead quantizes p_{m}, achieving $p_{\mathrm{d}} = p_{\mathrm{m}}^{(\ell)}$, and consequently $(M_1, S_1) \overset{s}{\equiv} (M_0, S_0)$. Correctness further requires $\ell \geq \lceil \log_2(1/p_{\mathrm{m}}^{\min}) \rceil$, so that every message in the support of p_{m} is assigned at least one seed.

4 Key Confidentiality

Key confidentiality (KC) is crucial for user-based keys such as passwords and biometrics: since such keys are often reused across systems, any information about a reused key leaked in one setting may compromise security in another.

We quantify key leakage by the mutual information $I(K;C)$ between the key K and ciphertext C. Our main result is that the optimal attainable level of KC is $I(K;C) = \mathrm{negl}(\ell)$. Moreover, we derive *tight* characterizations of when this optimum is achieved: for general schemes, $I(K;C) = \mathrm{negl}(\ell)$ holds if and only if ciphertext distributions are statistically indistinguishable across keys; for *uniform* schemes, it is further equivalent to a message-indistinguishability condition that matches the HE design goal (informally, decrypting under any key yields messages distributed almost as real ones). Finally, we show that HE with IS-DTE achieves $I(K;C) = O(2^{-2\ell})$.

4.1 Security Definition

We begin with an idealized notion of key confidentiality.

Definition 1. *A symmetric encryption scheme* $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \mathcal{E}, \mathcal{D})$ *achieves* perfect KC *if, for* $M \leftarrow_{p_{\mathrm{m}}} \mathcal{M}$, $K \leftarrow_{p_{\mathrm{k}}} \mathcal{K}$, *and* $C \leftarrow E_K(M)$, *it holds that* $I(K;C) = 0$.

Perfect KC is an information-theoretic ideal in the spirit of Shannon's framework. It is trivially attainable by message-revealing schemes; for example, $E_k(m) = m$ ignores the key, making C independent of K while completely revealing M. Excluding such schemes, when encryption uses a bounded number of fair random bits, perfect KC need not be attainable for arbitrary distributions, since exact sampling may require unbounded randomness. We collect the corresponding conditions and impossibility statements in Appendix D. Accordingly, we focus on the asymptotic relaxation of *statistical KC* defined next.

For optimal achievable KC, we define statistical KC for an encryption scheme consisting of a family $(\mathcal{M}, \mathcal{K}, \mathcal{C}^{(\ell)}, \mathcal{E}^{(\ell)}, \mathcal{D}^{(\ell)})$ with finite ℓ and $\mathcal{C}^{(\ell)} = \{0,1\}^{\ell} = \mathbb{Z}_{2^{\ell}}$. Unless otherwise stated, KC refers to statistical KC in this paper.

Definition 2. *An encryption scheme achieves* statistical KC *if* $I(K;C^{(\ell)}) = \mathrm{negl}(\ell)$, *where* $M \leftarrow_{p_{\mathrm{m}}} \mathcal{M}$, $K \leftarrow_{p_{\mathrm{k}}} \mathcal{K}$, *and* $C^{(\ell)} \leftarrow E_K^{(\ell)}(M)$.

We note that requiring $I(K;C) = \mathrm{negl}(\ell)$ is strong and in particular implies guessing-based notions of KC (see Appendix I).

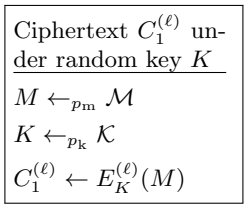
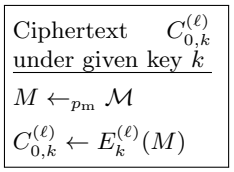
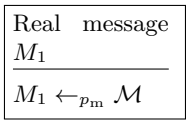


Fig. 3: Indistinguishability of ciphertexts (**C-IND**): $C_1^{(\ell)} \stackrel{s}{\equiv} C_{0,k}^{(\ell)}$ for each k





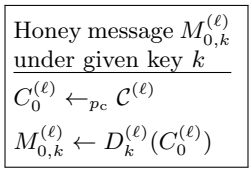


Fig. 4: Indistinguishability of real and honey messages (**M-IND**) for encryption: $M_1 \stackrel{s}{\equiv} M_{0,k}^{(\ell)}$ for each key k

4.2 Necessary and/or Sufficient Conditions for Statistical KC

We begin with a high-level summary of our characterizations of statistical KC, and relate them to three indistinguishability notions: ciphertext indistinguishability across keys (C-IND in Fig. 3), message indistinguishability under wrong-key decryption (M-IND in Fig. 4), and message–seed indistinguishability for EE encoders (MS-IND in Fig. 2). The formal theorems appear below.

For general schemes, statistical KC holds if and only if C-IND holds. Moreover, C-IND implies M-IND: for every key k, decrypting a real ciphertext under k yields a message distribution statistically close to that of the real message (and hence close to p_m), matching the decoy-message intuition behind HE-style designs. The converse (M-IND $\Rightarrow$ C-IND) fails in general, but it does hold for *uniform* schemes. In that case, M-IND is equivalent to statistical KC and further implies a structural near-uniformity property: for each key, the key-conditioned ciphertext distribution is nearly uniform within each equivalence class of the decryption-induced partition of the ciphertext space.

For EE constructions, MS-IND of the encoder is sufficient for statistical KC but not necessary. In addition, MS-IND implies that the marginal ciphertext distribution p_c is nearly uniform over the entire ciphertext space.

In summary:

1. General schemes: Statistical KC $\Leftrightarrow$ C-IND (Fig. 3) $\not\Rightarrow\not\Leftarrow$ M-IND (Fig. 4).
2. Uniform schemes: Statistical KC $\Leftrightarrow$ M-IND $\Rightarrow$ local uniformity of ciphertexts.
3. EE constructions: Statistical KC $\not\Rightarrow\not\Leftarrow$ MS-IND (of encoders) $\Rightarrow$ global uniformity of ciphertexts.

Theorem 1. *A symmetric encryption scheme achieves statistical KC if and only if it satisfies C-IND in Fig. 3, i.e., for every key k we have $C_1^{(\ell)} \stackrel{s}{\equiv} C_{0,k}^{(\ell)}$. Specifically,*

$$\mathrm{SD}(C_{0,k}^{(\ell)}, C_1^{(\ell)}) \le \sqrt{\frac{\ln 2}{2p_k^{\min}} I(K; C^{(\ell)})} = O\left(\sqrt{I(K; C^{(\ell)})}\right), \qquad \forall k \in \mathcal{K};$$

$$I(K; C^{(\ell)}) \le \frac{\log(1/p_k^{\min})}{1 - p_k^{\min}} \max_k \mathrm{SD}(C_{0,k}^{(\ell)}, C_1^{(\ell)}) = O\left(\max_k \mathrm{SD}(C_{0,k}^{(\ell)}, C_1^{(\ell)})\right).$$

The proof given in Appendix F combines Pinsker's inequality [18] with a suitable reverse Pinsker's inequality (Theorem 7 in [43]), relating statistical distance and KL divergence. Intuitively, this shows that for each k, $p_{\mathrm{c|k}}(\cdot\,|\,k)$ of $E_k^{(\ell)}(M)$ is close to p_{c}. Thus, for every $k \in \mathcal{K}$, $D_k^{(\ell)}$ maps p_{c} back to nearly the same message distribution p_{m}, yielding Corollary 1. This reflects the central design principle of HE schemes: generating indistinguishable honey messages for each key.

Definition 3. *An encryption scheme is* M-IND *if for every key k, the indistinguishability in Fig. 4 holds, i.e., $M_1 \overset{s}{\equiv} M_{0,k}^{(\ell)}$.*

Corollary 1. *If a symmetric encryption scheme achieves statistical KC, then it is M-IND. Specifically, for each $k \in \mathcal{K}$,*

$$\mathrm{SD}(M_{0,k}^{(\ell)}, M_1) \le \mathrm{SD}(C_1^{(\ell)}, C_{0,k}^{(\ell)}) \le \sqrt{\frac{\ln 2}{2p_{\mathrm{k}}^{\min}} I(K; C^{(\ell)})} = O\left(\sqrt{I(K; C^{(\ell)})}\right).$$

The converse of Corollary 1 does not hold in general (see Appendix H). Fortunately, if the encryption is uniform, the converse does hold.

Theorem 2. *A* uniform *encryption scheme achieves statistical KC if and only if it is M-IND. Specifically,*

$$\max_k \mathrm{SD}(M_{0,k}^{(\ell)}, M_1) \le \max_k \mathrm{SD}(C_{0,k}^{(\ell)}, C_1^{(\ell)}) \le \frac{|\mathcal{M}||\mathcal{K}|}{2p_{\mathrm{k}}^{\min}} \max_k \mathrm{SD}(M_{0,k}^{(\ell)}, M_1).$$

Proof sketch. Correctness and data processing give the first inequality. For the converse, M-IND bounds the error in each region's total mass, whereas C-IND must control absolute errors at individual ciphertexts; cancellation prevents a direct converse to data processing. Uniformity makes each conditional density constant on a message–key region. We introduce a weighted directed acyclic graph to capture density differences between regions. Each node represents a nonempty message–key region, including failure regions with density zero. An edge joins two overlapping regions of different densities and points from higher to lower density. Its weight is the overlap size times the two key probabilities, and its flow is this weight times the density difference. A node's net outgoing flow equals its prior-weighted message-mass error. Ordering nodes by density gives at most $|\mathcal{M}||\mathcal{K}|$ cuts after the failure regions are merged; each cut is bounded by the total positive error. Summing over cuts bounds the total flow, which controls pairwise ciphertext distances and hence distance from their mixture. The full proof is in Appendix F.

Local uniformity of ciphertexts. For a uniform encryption scheme, the ciphertext density under each key is constant within each decryption region. By Theorem 2, M-IND makes the ciphertext marginal close to such a piecewise-uniform distribution for every key. Densities may differ between regions, so this local uniformity does not imply uniformity over the entire ciphertext space. The formal statement is Proposition 11 in Appendix F. In the next section, we show that EE constructions achieve global uniformity by allocating seeds in proportion to message probabilities.

4.3 Security Analysis

The insufficiency of traditional encryption. For uniformly random messages, a deterministic cipher can satisfy Theorem 1. However, for non-uniform p_{m}, deterministic ciphers fail: E_k produces ciphertext distributions that vary with k. Fig. 1b shows an example with $(p_{\mathrm{m}}(m_i))_i = (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$, yielding $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ for k_1 and $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ for k_2.

Define $c_k^\star := E_k(m^\star)$. If a deterministic cipher achieves KC, Theorem 1 requires $p_{\mathrm{c}}(c_k^\star) = p_{\mathrm{c|k}}(c_k^\star \mid k) = p_{\mathrm{m}}(m^\star)$ for all k, and further $p_{\mathrm{m}}(D_{k_2}(c_{k_1}^\star)) = p_{\mathrm{m}}(m^\star)$ for all k_1, k_2. These equalities may extend over large subsets of $\mathcal{M}$, but they rarely hold except when p_{m} is uniform.

Some prior studies quantify key leakage in certain deterministic encryption schemes [6, 10, 7] under multiple-message models. In contrast, we derive a general bound for the ideal cipher in our single-message model, for arbitrary p_{m} and p_{k}; see Theorem 20 in Appendix F. However, our bound is typically loose for specific $p_{\mathrm{m}}, p_{\mathrm{k}}$.

To illustrate the magnitude of key leakage in practice, we consider a password-vault example, a typical HE application [13], and estimate $H(K \,|\, C)$ by empirical sampling in the ideal-cipher model. Both the master password (key) and the stored password (message) are modeled as 4-digit numbers following a Zipf distribution with $p(\mathrm{pin}_i) \propto i^{-r}$ and $r = 1.0$. This estimate gives $H(K \,|\, C) \approx 3.3$ bits. For two 4-digit passwords as the message, the leakage rises to nearly 6 bits. Given that the key entropy $H(K)$ is only 9.5 bits, this leakage is substantial.

Security of HE schemes and EE constructions. Unlike traditional encryption, EE constructions incorporate message dependency through a randomized encoder. The encoder assigns $|\mathcal{S}| \cdot p_{\mathrm{d}}(m)$ seeds to each message m, which translates into $|\mathcal{C}| \cdot p_{\mathrm{d}}(m)$ ciphertexts under a deterministic cipher.

If the encoder achieves perfect MS-IND (i.e., $p_{\mathrm{m}} = p_{\mathrm{d}}$), then the ciphertext distribution is uniform under every key. By Theorem 1, such a construction achieves perfect KC. As shown in Fig. 1b with a three-message distribution $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$: the maximum-probability message $m^\star = m_1$ (probability $\frac{1}{2}$) is assigned two seeds, while the less probable messages m_2 and m_3 (probability $\frac{1}{4}$ each) are assigned one seed each.

Lemma 1. *For any EE construction, let* $M_1 \leftarrow_{p_{\mathrm{m}}} \mathcal{M}$, $M_0^{(\ell)} \leftarrow_{p_{\mathrm{d}}} \mathcal{M}$, $C_1^{(\ell)} \leftarrow_{p_{\mathrm{c}}} \mathcal{C}^{(\ell)}$, *and* $U^{(\ell)} \leftarrow_{\$} \mathcal{C}^{(\ell)}$. *Then*

$$I(K; C^{(\ell)}) + \mathrm{KL}(C_1^{(\ell)} \| U^{(\ell)}) = \mathrm{KL}(M_1 \| M_0^{(\ell)}).$$

Theorem 3. *An EE construction with an encoder and a cipher achieves statistical KC if the encoder is MS-IND. Specifically, if* $\mathrm{SD}(M_0^{(\ell)}, M_1) < p_{\mathrm{m}}^{\min}$, *then*

$$I(K; C^{(\ell)}) \le \frac{2\big(\mathrm{SD}(M_0^{(\ell)}, M_1)\big)^2}{(p_{\mathrm{m}}^{\min} - \mathrm{SD}(M_0^{(\ell)}, M_1)) \ln 2} = O\Big(\big(\mathrm{SD}(M_0^{(\ell)}, M_1)\big)^2\Big).$$

The converse of Theorem 3 does not hold (see Appendix H).

Proposition 1 (Global uniformity of ciphertexts). *For an EE construction, if its encoder is MS-IND, then* $C_{0,k}^{(\ell)} \overset{s}{\equiv} C_1^{(\ell)} \overset{s}{\equiv} U^{(\ell)}$ *for every* $k \in \mathcal{K}$. *Specifically,* $\mathrm{SD}(C_1^{(\ell)}, U^{(\ell)}) \le \mathrm{SD}(C_{0,k}^{(\ell)}, U^{(\ell)}) = \mathrm{SD}(M_0^{(\ell)}, M_1)$ *for every* $k \in \mathcal{K}$.

Theorem 3 provides a bound for the EE construction by the statistical distance. However, with carefully designed encoders, such as IS-DTE [31] and IS-PMTE [15], it is possible to directly constrain the disparity $|p_{\mathrm{d}}(m) - p_{\mathrm{m}}(m)|$ for each $m \in \mathcal{M}$, thereby achieving a more precise bound.

Proposition 2. *For an EE construction, if* $|p_{\mathrm{d}}(m) - p_{\mathrm{m}}(m)| \le \epsilon$ *for every* $m \in \mathcal{M}$, *where* $0 < \epsilon < p_{\mathrm{m}}^{\min}$, *then*

$$I(K; C^{(\ell)}) \le \frac{|\mathcal{M}|\epsilon^2}{(p_{\mathrm{m}}^{\min} - \epsilon)\ln 2} = O(\epsilon^2).$$

Proposition 2 gives concrete leakage bounds for practical encoders such as IS-DTE and IS-PMTE, including $I(K; C^{(\ell)}) = O(2^{-2\ell})$ for IS-DTE (see Appendix G).

p_{m}-agnostic encryption with optimal KC. The trivial encryption $E_k(m) = m$ achieves KC because its ciphertext is independent of the key, but it provides no message confidentiality: $p_{\mathrm{guess}}(M \mid C) = 1$.

More generally, call an encryption scheme p_{m}-agnostic if the same encryption and decryption algorithms are used for every message distribution on $\mathcal{M}$. The following theorem applies to such p_{m}-agnostic encryption schemes.

Theorem 4. *For any* p_{m}*-agnostic encryption scheme* $\mathsf{E} = \{\mathsf{E}^{(\ell)}\}_\ell$, *if* E *achieves statistical KC under* p_{k} *for every message distribution* p_{m}, *then, for every* p_{m},

$$p_{\mathrm{guess}}^{\mathsf{E}^{(\ell)}, p_{\mathrm{m}}, p_{\mathrm{k}}}(M \mid C^{(\ell)}) = 1 - \mathrm{negl}(\ell).$$

Therefore, except for such trivial encryption schemes that reveal the message, achieving KC for arbitrary message distributions requires tailoring the encryption scheme to p_{m}.

Key confidentiality in multiple ciphertexts.

Proposition 3. *For* n *ciphertexts* $C_1, \ldots, C_n$ *encrypted with the same key* K *(possibly under different encryption schemes), suppose the messages* $M_1, \ldots, M_n$ *are jointly independent of* K *and the encryption randomness is independent across encryptions and of the sources. Then*

$$I(K; C_1, \ldots, C_n) \le \sum_{i=1}^{n} I(K; C_i) + \mathrm{TC}(M_1, \ldots, M_n),$$

where $\mathrm{TC}(M_1, \ldots, M_n) = \sum_{i=1}^{n} H(M_i) - H(M_1, \ldots, M_n)$ *is the total correlation. In particular, if the messages are mutually independent, then*

$$I(K; C_1, \ldots, C_n) \le \sum_{i=1}^{n} I(K; C_i).$$

5 Message Confidentiality

We study message confidentiality primarily via message-recovery security (MRS). We show that the best achievable information-theoretic guarantee is to bound any adversary's message-recovery success probability by $p_{\max} + \mathrm{negl}(\ell)$, while a trivial strategy that guesses the most likely message or decrypts under the most likely key succeeds with probability at least $p_{\max}$.

MRS Game $\mathrm{MRS}^{\mathcal{A}}_{\mathsf{E},p_{\mathrm{m}},p_{\mathrm{k}}}$

$M \leftarrow_{p_{\mathrm{m}}} \mathcal{M}$
$K \leftarrow_{p_{\mathrm{k}}} \mathcal{K}$
$C \leftarrow E_K(M)$
$M' \leftarrow \mathcal{A}(C)$
return $M' = M$

Fig. 5: Game for message recovery security (MRS)

We then construct a scheme OE that attains $p_{\max} + O(2^{-\ell})$ for general $(p_{\mathrm{m}}, p_{\mathrm{k}})$. OE uses the cumulative distribution functions (CDFs) of p_{m} and p_{k} to allocate ciphertext regions, controlling each message–ciphertext joint probability and thereby bounding the overall recovery probability. Finally, we show that HE schemes and more general EE constructions cannot in general achieve the $p_{\max} + \mathrm{negl}(\ell)$ bound for arbitrary distributions, revealing that achieving optimal MRS security requires principles for using randomness in probabilistic encryption that are fundamentally different from those underlying HC/HE.

5.1 Security Definition

Message-recovery security (MRS) captures attacks that attempt to recover the entire message from the ciphertext. For a symmetric-encryption scheme E, MRS is defined by the game in Fig. 5. An adversary $\mathcal{A}$ wins if it outputs the correct message M given C. Its success probability[4] is $\Pr[\mathrm{MRS}^{\mathcal{A}}_{\mathsf{E},p_{\mathrm{m}},p_{\mathrm{k}}} \Rightarrow \mathrm{true}]$, and the optimal MRS guessing probability is

$$p^{\mathsf{E},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) \coloneqq \max_{\mathcal{A}} \Pr\left[\mathrm{MRS}^{\mathcal{A}}_{\mathsf{E},p_{\mathrm{m}},p_{\mathrm{k}}} \Rightarrow \mathrm{true}\right].$$

Given a ciphertext c, the optimal strategy is to output a message maximizing the posterior $p_{\mathrm{m}|\mathrm{c}}(m\,|\,c)$. Therefore,

$$p^{\mathsf{E},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) = \sum_{c} p_{\mathrm{c}}(c) \max_{m} p_{\mathrm{m}|\mathrm{c}}(m\,|\,c) = \sum_{c} \max_{m} p_{\mathrm{m,c}}(m, c).$$

In particular, a trivial adversary succeeds with probability at least $p_{\max}$ by outputting $m_{\mathrm{triv}}(c) \coloneqq m^{\star}$ if $p^{\max}_{\mathrm{m}} \geq p^{\max}_{\mathrm{k}}$, and $m_{\mathrm{triv}}(c) \coloneqq D_{k^{\star}}(c)$ otherwise. Therefore, for an arbitrary encryption scheme E,

$$p^{\mathsf{E},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) \geq p_{\max}.$$

Beyond MRS, we also consider target-distribution semantic security (TDSS); see Appendix M.

[4] Juels and Ristenpart [31] refer to the probability as $\mathcal{A}$'s MRS advantage $\mathbf{Adv}^{\mathrm{mrs}}_{\mathsf{HE},p_{\mathrm{m}},p_{\mathrm{k}}}(\mathcal{A})$.

5.2 MRS-Optimal Encryption (MRS-OE)

Circular intervals on the ciphertext space. We identify the ciphertext space $\mathcal{C}^{(\ell)} = \{0,1\}^\ell$ with $\{0,\dots,2^\ell-1\}$ and view it as a circle modulo 2^ℓ. For integers a, t with $0 \le t \le 2^\ell$, write $[a, a+t) \coloneqq \{(a+r) \bmod 2^\ell \mid r \in \mathbb{Z},\ 0 \le r < t\}$ for the length-t circular interval starting at a; it may wrap around zero. For an integer s, define its translation by $s + [a, a+t) \coloneqq [a+s, a+s+t)$.

Design idea. To control the recovery probability $\sum_c \max_m p_{\mathrm{m,c}}(m,c)$, we seek to reduce each joint probability $p_{\mathrm{m,c}}(m,c)$. For uniform encryption, avoiding *decryption collisions*—distinct keys decrypting the same ciphertext to the same message—leaves at most one key contributing to each guess, with contribution $p_{\mathrm{m}}(m)p_{\mathrm{k}}(k)/\mu(m,k)$. Our aim is therefore to enlarge the ciphertext regions to dilute these contributions, while keeping them disjoint across messages for each key (correctness) and across keys for each message (no collisions). If we achieve $\mu(m,k) = 2^\ell \min\{p_{\mathrm{m}}(m), p_{\mathrm{k}}(k)\} + O(1)$, then $p_{\mathrm{m,c}}(m,c) \le 2^{-\ell} p_{\max} + O(2^{-2\ell})$ for every (m,c), and hence $p_{\mathrm{guess}}(M \mid C) \le p_{\max} + O(2^{-\ell})$.

OE meets these requirements by jointly using the message and key CDF intervals on the ciphertext circle. For each key k, we shift the message CDF intervals $\mathcal{I}_{\mathrm{m}}^{(\ell)}(m)$ by the common offset $\widetilde{F}_{\mathrm{k}}^{(\ell)}(k^-)$; for each message m, we shift the key CDF intervals $\mathcal{I}_{\mathrm{k}}^{(\ell)}(k)$ by $\widetilde{F}_{\mathrm{m}}^{(\ell)}(m^-)$. These shifts preserve both disjointness properties and align the two intervals for (m,k) at their left endpoints, so their intersection retains the entire shorter interval. Specifically, we define

$$\begin{aligned}\mathcal{C}^{(\ell)}(m,k) &\coloneqq \big(\widetilde{F}_{\mathrm{m}}^{(\ell)}(m^-) + \mathcal{I}_{\mathrm{k}}^{(\ell)}(k)\big) \cap \big(\widetilde{F}_{\mathrm{k}}^{(\ell)}(k^-) + \mathcal{I}_{\mathrm{m}}^{(\ell)}(m)\big) \\ &= \widetilde{F}_{\mathrm{m}}^{(\ell)}(m^-) + \widetilde{F}_{\mathrm{k}}^{(\ell)}(k^-) + \big[0, \min\{\widetilde{p}_{\mathrm{m}}^{(\ell)}(m), \widetilde{p}_{\mathrm{k}}^{(\ell)}(k)\}\big).\end{aligned}$$

Thus, OE satisfies the above constraints with $\mu(m,k) = 2^\ell \min\{p_{\mathrm{m}}(m), p_{\mathrm{k}}(k)\} + O(1)$, achieving the desired recovery bound.

The OE architecture. For each (m,k), OE uses the circular interval $\mathcal{C}^{(\ell)}(m,k)$ defined above as its ciphertext region (see Fig. 6). Encryption samples $C \leftarrow_{\$} \mathcal{C}^{(\ell)}(m,k)$. Decryption under k locates an interval containing the ciphertext and returns its message, or $\perp$ if no such interval exists. To ensure that all regions are nonempty, we require $\ell \ge \ell_{\mathsf{OE}}^{\min}$, where $\mu_{\min} \coloneqq \min\{p_{\mathrm{m}}^{\min}, p_{\mathrm{k}}^{\min}\}$ and $\ell_{\mathsf{OE}}^{\min} \coloneqq \lfloor \log_2(1/\mu_{\min}) \rfloor + 1$. The preceding analysis yields the following lemma and theorem.

Lemma 2. *The ciphertext regions of* OE *are both keywise and messagewise disjoint: for each fixed message, they are pairwise disjoint across keys; for each fixed key, they are pairwise disjoint across messages.*

Theorem 6. *For the discrete scheme* OE*, if* $\ell \ge \ell_{\mathsf{OE}}^{\min}$*, then*

$$p_{\mathrm{guess}}^{\mathsf{OE},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) \le \frac{2^\ell \mu_{\min}}{2^\ell \mu_{\min} - 1} p_{\max} = p_{\max} + O(2^{-\ell}).$$

Formal proofs of the lemma and theorem appear in Appendix J.

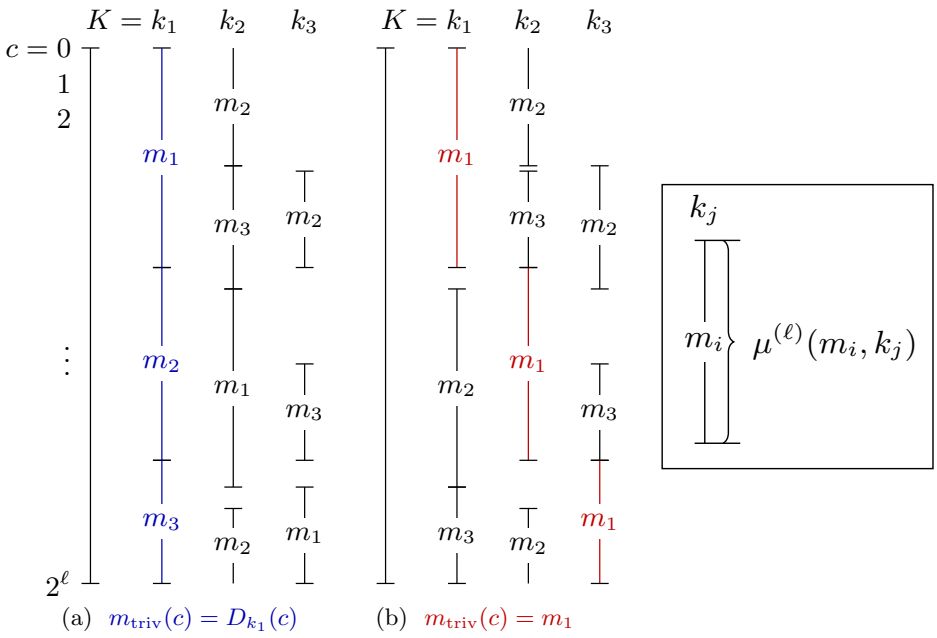


Fig. 6: Ciphertext regions in OE.

5.3 Computational Complexity

To encrypt (m, k), OE computes the region size $\mu^{(\ell)}(m, k)$ from the message and key CDFs, samples $R \leftarrow_{\$} [0, \mu^{(\ell)}(m, k))$, and outputs $\widetilde{F}_{\mathrm{m}}^{(\ell)}(m^-) + \widetilde{F}_{\mathrm{k}}^{(\ell)}(k^-) + R \pmod{2^\ell}$. To decrypt (c, k), it computes the key-CDF coordinate $y \coloneqq c - \widetilde{F}_{\mathrm{k}}^{(\ell)}(k^-) \pmod{2^\ell}$ and uses the inverse message CDF to locate the CDF interval containing y. Let m be its message and $r \coloneqq y - \widetilde{F}_{\mathrm{m}}^{(\ell)}(m^-)$ its offset. Return m if $r < \widetilde{p}_{\mathrm{k}}^{(\ell)}(k)$, and return $\perp$ otherwise. Thus, apart from interval arithmetic and sampling, encryption requires one message-CDF and one key-CDF computation, whereas decryption requires one key-CDF and one inverse message-CDF computation.

With explicitly stored CDF tables, the storage is $O(\ell|\mathcal{M}| + \ell|\mathcal{K}|)$ bits. With binary-search queries, encryption takes $O(\ell + \log|\mathcal{M}| + \log|\mathcal{K}|)$ time, whereas decryption takes $O(\ell \log|\mathcal{M}| + \log|\mathcal{K}|)$ time. For structured probability models, the CDF and inverse-CDF computations can be further optimized using the model's specific representation.

5.4 Insufficiency of HE Schemes and EE Constructions for Optimal Security

In this section, we show that HE schemes and general EE constructions cannot achieve optimal security for certain key and message distributions.

Insufficiency of HE schemes with MS-IND encoders. We ignore finite-precision errors, treating MS-IND encoders as perfectly MS-IND. Prior work [31, 29, 11, 26, 23, 15, 14] focused only on MS-IND encoders, implicitly treating them as optimal. We show this is not the case.

In HE schemes, each message m_i is assigned $p_{\mathrm{m}}(m_i) \cdot |\mathcal{C}|$ ciphertexts.

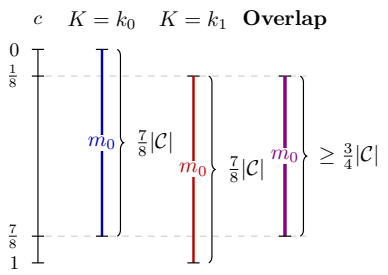


Fig. 7: Unavoidable collision in HE under $p_{\mathrm{m}} = (\frac{7}{8}, \frac{1}{8})$.

Consider $\mathcal{M} = \{m_0, m_1\}$ with distribution $(\frac{7}{8}, \frac{1}{8})$ and $\mathcal{K} = \{k_0, k_1\}$ with distribution $(\frac{7}{8}, \frac{1}{8})$. For $\ell \geq 3$, the MS-IND encoder assigns exactly $\frac{7}{8}|\mathcal{C}|$ seeds to m_0, resulting in $\frac{7}{8}|\mathcal{C}|$ ciphertexts under both k_0 and k_1. Thus, as illustrated in Fig. 7, at least $(2 \cdot \frac{7}{8} - 1)|\mathcal{C}| = \frac{3}{4}|\mathcal{C}|$ ciphertexts decrypt to m_0 under both keys. This yields $p_{\mathrm{guess}}^{\mathsf{HE},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) \geq p_{\mathrm{k}}(k_0) + p_{\mathrm{k}}(k_1) \cdot \frac{3}{4} = \frac{31}{32} > p_{\max} = \frac{7}{8}$. In contrast, a deterministic encoder achieves optimal MRS for these distributions. Let $\mathcal{C} \coloneqq \{0,1\}$, define $\mathtt{encode}_{\mathrm{t}}(m_i) \coloneqq i$, and set $E_{k_j}(m_i) \coloneqq i \oplus j$. Then $\max_m p_{\mathrm{m,c}}(m,0) = \frac{49}{64}$ and $\max_m p_{\mathrm{m,c}}(m,1) = \frac{7}{64}$, giving $p_{\mathrm{guess}}^{\mathsf{EE}_{\mathrm{t}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) = \frac{7}{8} = p_{\max}$.

In this example, the MS-IND requirement forces decryption collisions and prevents optimal MRS. In contrast, EE_{t} avoids collisions and achieves optimal MRS. This suggests new directions for encoder design: *which encoder is optimal for EE constructions?*

Insufficiency of EE constructions with general encoders. Even with general encoders, EE constructions cannot achieve optimal MRS for all message and key distributions. The analysis is more delicate for general encoders, especially as $\ell \to \infty$.

Theorem 7. *For $\mathcal{M} = \{a, b\}$ with $p_{\mathrm{m}} = (\frac{1}{2}, \frac{1}{2})$ and any $\mathcal{K}$ such that $p_{\mathrm{k}}^{\max} \leq \frac{1}{2}$ and $p_{\mathrm{k}}(\mathcal{K}_1) \neq p_{\mathrm{k}}(\mathcal{K}_2)$ for every two distinct subsets $\mathcal{K}_1, \mathcal{K}_2 \subseteq \mathcal{K}$, it holds that*

$$\inf_{\mathsf{EE} \in \Omega_{EE}} p_{\mathrm{guess}}^{\mathsf{EE},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) > p_{\max} = \tfrac{1}{2}.$$

From the bound in the proof, if $p_{\mathrm{k}} = (\frac{1}{4}, \frac{1}{3}, \frac{5}{12})$, then $p_{\mathrm{guess}}^{\mathsf{EE},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) \geq \frac{1}{2} + \frac{1}{48}$ for any EE construction EE, even as $\ell \to \infty$. EE constructions are key-independent probabilistic encryption schemes, unlike our key-dependent OE. This independence prevents them from reaching optimal security for general distributions.

In Theorem 7, the MRS guessing probability is computed for an EE construction with a fixed deterministic cipher. In the random oracle and ideal cipher models [31, 29], the cipher is modeled as a random oracle or ideal cipher, and the MRS guessing probability is taken in expectation over the choice of the cipher. Since $\inf_x \mathbb{E}_y f(x,y) \geq \inf_{x,y} f(x,y)$, the insufficiency of EE constructions still holds in these models.

5.5 Other Security Properties of Our Scheme OE

Key confidentiality. OE is designed for optimal MRS and generally does not guarantee optimal KC. Ignoring quantization errors, the regions for key k cover a fraction $\sum_m \min\{p_{\mathrm{m}}(m), p_{\mathrm{k}}(k)\}$ of the ciphertext space and may leave gaps. Since these gaps and the conditional ciphertext densities generally depend on k, the ciphertext can leak information about the key.

For general encryption schemes, KC and MRS cannot be simultaneously optimal under arbitrary distributions. Proposition 4 illustrates this obstruction.

Semantic security. In the information-theoretic setting, low-entropy keys do not generally guarantee semantic security. A weaker notion, TDSS, has been studied for HE schemes, but the known bounds are specific to HE constructions and are not tight for general message and key distributions (see Appendices B and M). Determining the optimal achievable TDSS bound for arbitrary $(p_{\mathrm{m}}, p_{\mathrm{k}})$ remains open. The TDSS security of our scheme OE also remains an open question.

Computational security. The design of OE focuses solely on information-theoretic security against unbounded adversaries, and does not incorporate mechanisms for increasing the computational cost of guessing attacks. In particular, the assignment of $\mathcal{C}^{(\ell)}(m, k)$ is deterministic, allowing adversaries to exploit it at low computational cost.

A natural idea is to incorporate computationally intensive hash functions H (e.g., iterative or memory-hard hashes [8, 3]). Following [31], hashing the key together with a random salt s would force the adversary to evaluate H for every key candidate, thereby multiplying its cost. In the random-oracle model, however, key-dependent hashing can introduce decryption collisions and increase the MRS guessing probability. As an alternative, applying H only to a random salt s and using $H(s)$ to uniformly shift all intervals $\mathcal{C}^{(\ell)}(m, k)$ preserves the information-theoretic guarantees, but increases the adversary's cost by only a fixed constant.

How to amplify the computational hardness of OE remains an open problem.

6 Encryption Schemes with Continuous Ciphertext Spaces

To simplify the analysis and design of HE schemes, we extend encryption schemes to a continuous ciphertext space $[0, 1)$. On the one hand, continuous schemes can be discretized onto $\{0, 1\}^\ell$ with only $O(2^{-\ell})$ precision loss, enabling the construction of practical HE schemes from secure continuous ones. On the other hand, after uniformization inside decryption-equivalence classes, the finite vector of class masses of any sequence of discrete schemes over $\{0, 1\}^\ell$ has convergent subsequences. Each limit defines a canonical continuous scheme over $[0, 1)$; if the original KC or MRS metric approaches its optimum, every such canonical limit attains that optimum. This allows optimal-security questions that depend only on these class masses to be analyzed on asymptotic continuous schemes.

In addition, we identify necessary and sufficient conditions under which continuous schemes achieve optimal key or message confidentiality.

6.1 Definition and Convergence

Definition. A continuous encryption scheme over $\mathcal{C} \coloneqq [0,1)$ is defined in the same way as a discrete scheme, except that ciphertexts are points of $[0,1)$. We identify $\mathcal{C}$ with the circle $\mathbb{R}/\mathbb{Z}$. Thus, a circular interval with real endpoints $a \le b \le a+1$ is the image of $[a,b)$ under reduction modulo 1, and any ciphertext-coordinate sum is interpreted modulo 1. An interval wrapping around 0 is represented as the union of two half-open intervals in $[0,1)$. For a message m and a key k, the encryption algorithm $E_k(m)$ samples c from the ciphertext region $\mathcal{C}(m,k)$ according to the conditional density $p_{\mathrm{c|m,k}}(c \mid m,k)$, and decryption returns $D_k(c) = m$. We write $\mu(m,k)$ for the Lebesgue measure of $\mathcal{C}(m,k)$. Throughout the continuous model, all intervals are half-open. For every (m,k), $\mathcal{C}(m,k)$ is the union of the intervals on which $p_{\mathrm{c|m,k}}(\cdot\,|\,m,k)$ is positive. Thus, if encryption is uniform on its ciphertext region,

$$p_{\mathrm{c|m,k}}(c\,|\,m,k) = \frac{1}{\mu(m,k)} \qquad \text{for } c \in \mathcal{C}(m,k).$$

More general schemes on $[0,1)$ could allow arbitrary probability measures. Throughout this section, however, every conditional ciphertext distribution is absolutely continuous with respect to Lebesgue measure and has a density that is piecewise constant on a finite partition into nonempty intervals; ciphertext regions are finite unions of intervals. Thus every singleton has probability zero. Zero-length pieces are omitted, and intervals on which $p_{\mathrm{c}} = 0$ may be removed and the remaining intervals rescaled to $[0,1)$.

Convergence of schemes discretized from continuous ones. A continuous scheme E_{c} can be discretized naturally by quantizing its ciphertext space, giving a discrete scheme $\mathsf{E}^{(\ell)}$ over the finite grid $\mathbb{Z}_{2^\ell}$ that converges to E_{c} as $\ell \to \infty$.

More precisely, for every (m,k), consider the maximal intervals in $\mathcal{C}(m,k)$ on which $p_{\mathrm{c|m,k}}(\cdot\,|\,m,k)$ is positive and constant. Apply the quantizer Q_ℓ defined in Section 3 to the two endpoints of each such interval, so that $[a,b)$ becomes $[Q_\ell(a), Q_\ell(b))$. Preserve the conditional probability mass of the entire original interval and distribute it uniformly over the ciphertexts in its quantized image. The same endpoint is always quantized in the same way, and decryption on the quantized interval is inherited from the original interval. This defines the discrete scheme $\mathsf{E}^{(\ell)}$.

Conversely, the discrete scheme $\mathsf{E}^{(\ell)}$ can be continuized on $[0,1)$ by spreading the mass of every discrete ciphertext c uniformly over its interval $[c/2^\ell, (c+1)/2^\ell)$, with the same decryption output throughout that interval. Denote the resulting continuous scheme by $\mathsf{E}_{\mathrm{c}}^{(\ell)}$. The containing interval, and hence c, is recoverable almost surely from the continuous ciphertext, so $\mathsf{E}^{(\ell)}$ and $\mathsf{E}_{\mathrm{c}}^{(\ell)}$ have the same correctness and security metrics.

This makes the convergence transparent. In $\mathsf{E}_{\mathrm{c}}^{(\ell)}$, the interval $[a, b)$ is replaced by $[2^{-\ell}Q_\ell(a), 2^{-\ell}Q_\ell(b))$ with the same probability mass. Each endpoint moves by less than $2^{-\ell}$, and there are only finitely many endpoints. Hence the affected boundary strips have total length $O(2^{-\ell})$, and the density on every remaining interval changes by only $O(2^{-\ell})$. Under the half-open interval convention, the convergence is pointwise. A fixed point that is not an endpoint eventually lies in the quantized copy of the same interval; at an endpoint, the floor quantizer places the point in the same right-hand interval in both schemes. Thus $\mathsf{E}_{\mathrm{c}}^{(\ell)}$ converges to E_{c} in statistical distance, equivalently in L^1 for the conditional densities. The full calculation appears in Appendix L.

Theorem 8. *For any continuous scheme* E_{c} *and its discretization* $\mathsf{E}^{(\ell)}$,

$$I_{\mathsf{E}^{(\ell)}}(K; C^{(\ell)}) = I_{\mathsf{E}_{\mathrm{c}}}(K; C) + O(2^{-\ell}),$$
$$p_{\mathrm{guess}}^{\mathsf{E}^{(\ell)}, p_{\mathrm{m}}, p_{\mathrm{k}}}(M \mid C^{(\ell)}) = p_{\mathrm{guess}}^{\mathsf{E}_{\mathrm{c}}, p_{\mathrm{m}}, p_{\mathrm{k}}}(M \mid C) + O(2^{-\ell}).$$

Convergence of general discrete schemes. A general sequence $\mathsf{E} = (\mathsf{E}^{(\ell)})_\ell$ need not converge after continuization. A ciphertext region may become increasingly fragmented as ℓ grows: the number of its component intervals may grow without bound, while their locations also vary with ℓ. To obtain a stable representation, we first merge ciphertexts having the same decryption behavior under every key.

Write $\mathfrak{M} := (\mathcal{M}')^{|\mathcal{K}|}$ for the set of all decryption-output vectors $\boldsymbol{m} = (m_k)_{k \in \mathcal{K}}$. For each ℓ, define

$$c_1 \; R_{\mathrm{a}}^{(\ell)} \; c_2 \quad \Longleftrightarrow \quad D_k^{(\ell)}(c_1) = D_k^{(\ell)}(c_2) \quad \text{for every } k \in \mathcal{K}.$$

The equivalence classes are indexed by $\boldsymbol{m} \in \mathfrak{M}$; namely,

$$S_{\boldsymbol{m}}^{(\ell)} = \bigcap_{k \in \mathcal{K}} \mathcal{C}^{(\ell)}(m_k, k).$$

Fix a partition $(I_{\boldsymbol{m}})_{\boldsymbol{m} \in \mathfrak{M}}$ of $[0, 1)$ into $|\mathfrak{M}| = (|\mathcal{M}| + 1)^{|\mathcal{K}|}$ equal half-open intervals. For every ℓ, uniformize the conditional distributions within each nonempty $S_{\boldsymbol{m}}^{(\ell)}$ and rearrange that class onto $I_{\boldsymbol{m}}$, preserving its total conditional probability mass. An empty class contributes zero mass to its assigned interval. Denote the resulting continuous scheme by $\mathsf{E}_{\mathrm{r}}^{(\ell)}$. This transformation preserves correctness and cannot weaken security: its KC leakage and MRS guessing probability do not increase (Theorem 22).

The number and locations of the intervals are now independent of ℓ. On every interval, the conditional density is constant and determined by the total mass of the corresponding equivalence class. Since there are only finitely many such masses, there is a subsequence along which all of them converge. The corresponding conditional densities then converge pointwise to those of a continuous scheme E_{a}, and the two security metrics converge along the same subsequence. We call every such limit an *asymptotic scheme* of E.

Theorem 9. *Every asymptotic scheme* E_{a} *of* E *satisfies*

$$I_{\mathsf{E}^{(\ell)}}(K; C^{(\ell)}) = o(1) \Longrightarrow I_{\mathsf{E}_{\mathrm{a}}}(K; C) = 0,$$
$$p_{\mathrm{guess}}^{\mathsf{E}^{(\ell)}, p_{\mathrm{m}}, p_{\mathrm{k}}}(M \mid C^{(\ell)}) = p_{\max} + o(1) \Longrightarrow p_{\mathrm{guess}}^{\mathsf{E}_{\mathrm{a}}, p_{\mathrm{m}}, p_{\mathrm{k}}}(M \mid C) = p_{\max}.$$

6.2 Sufficiency and Necessity of Continuous Schemes

As a secure continuous scheme naturally induces a secure discrete scheme, any sufficient condition for continuous schemes also applies to discrete schemes. Conversely, a secure discrete sequence has one or more secure canonical continuous asymptotic schemes. This converse transfers conditions expressed in terms of the joint law of $(M, K, [C]_{R_{\mathrm{a}}})$; a condition involving additional geometry of ciphertext regions requires a separate closure argument.

We next give the continuous analogues of Theorems 1 and 2.

Theorem 1′. *A continuous scheme* E_{c} *achieves* $I(K; C) = 0$, if and only if $p_{\mathrm{c}|\mathrm{k}}(c \,|\, k) = p_{\mathrm{c}}(c)$ *for every* $k \in \mathcal{K}$ *and every* $c \in \mathcal{C}$.

Theorem 2′. *A* uniform *continuous scheme* E_{c} *achieves* $I(K; C) = 0$, if and only if $\forall m \in \mathcal{M}, k \in \mathcal{K}$, $\Pr[D_k(C) = m \,|\, C \leftarrow_{p_{\mathrm{c}}} \mathcal{C}] = p_{\mathrm{m}}(m)$.

Unlike in the discrete setting, the continuous framework allows us to give a clean necessary-and-sufficient characterization of optimal MRS.

Theorem 10. *A continuous encryption scheme* E_{c} *achieves* $p_{\mathrm{guess}}^{\mathsf{E}_{\mathrm{c}}, p_{\mathrm{m}}, p_{\mathrm{k}}}(M \mid C) = p_{\max}$ *if and only if, for every choice of maximizers* $m^\star$ *and* $k^\star$*:*

1. *If* $p_{\mathrm{k}}^{\max} \geq p_{\mathrm{m}}^{\max}$*: 1)* $D_{k^\star}(c)$ *maximizes* $p_{\mathrm{m,c}}(\cdot, c)$ *for every* c*, and 2)* $\mathcal{C}(m, k^\star) \cap \mathcal{C}(m, k') = \emptyset$ *for every* m *and every* $k' \neq k^\star$.
2. *If* $p_{\mathrm{m}}^{\max} \geq p_{\mathrm{k}}^{\max}$*:* $m^\star$ *maximizes* $p_{\mathrm{m,c}}(\cdot, c)$ *for every* c.

Corollary 2. *For a continuous encryption scheme* E_{c} *with* $p_{\mathrm{guess}}^{\mathsf{E}_{\mathrm{c}}, p_{\mathrm{m}}, p_{\mathrm{k}}}(M \mid C) = p_{\max}$*, the following conditions hold for every choice of maximizers* $m^\star$ *and* $k^\star$*:*

1. *If* $p_{\mathrm{k}}^{\max} \geq p_{\mathrm{m}}^{\max}$*: 1)* $\{\mathcal{C}(m, k^\star)\}_m$ *is a partition of* $\mathcal{C}$*, and 2)* $\mathcal{C}(m, k^\star) \cap \mathcal{C}(m, k') = \emptyset$ *for every* m *and every* $k' \neq k^\star$.
2. *If* $p_{\mathrm{m}}^{\max} \geq p_{\mathrm{k}}^{\max}$*:* $\bigcup_k \mathcal{C}(m^\star, k) = \mathcal{C}$.

Formal proofs of Theorem 10 and Corollary 2 appear in Appendix L.

These conditions are natural. When $p_{\mathrm{k}}^{\max} \geq p_{\mathrm{m}}^{\max}$, the conditions mean: 1) there does not exist a ciphertext c such that $D_{k^\star}(c) = \perp$; and 2) there does not exist a ciphertext c such that $D_{k^\star}(c) = D_{k'}(c) \in \mathcal{M}$ for some $k' \neq k^\star$. If such ciphertexts exist, the guessing probability exceeds $p_{\mathrm{k}}(k^\star) = p_{\max}$. When $p_{\mathrm{m}}^{\max} \geq p_{\mathrm{k}}^{\max}$, the condition means that there does not exist a ciphertext c such that $D_k(c) \neq m^\star$ for every k. If such a ciphertext exists, the guessing probability exceeds $p_{\mathrm{m}}(m^\star) = p_{\max}$.

6.3 Analysis and Designs of Continuous Schemes

Building on the above theorems on continuous encryption, we obtain more concise insights and streamlined designs for encryption schemes.

Necessary tailoring to the key distribution for optimal MRS. Using the necessary conditions, we demonstrate that tailoring the scheme to the key distribution is essential to achieving MRS in general cases with $|\mathcal{M}| \geq 2$, except in the cases characterized by the following theorem.

Theorem 11. *For a message distribution p_{m}, there exists a p_{k}-agnostic continuous encryption scheme E_{c} achieving $p_{\mathrm{guess}}^{\mathsf{E}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) = p_{\max}$ for every key distribution p_{k} (under a given $|\mathcal{K}|$) if and only if one of the following three conditions holds:*

1. $|\mathcal{M}| = |\mathcal{K}| = 2$ *(or $|\mathcal{M}| = 1$).*
2. $|\mathcal{M}| \geq 3$, $|\mathcal{K}| = 2$, *and* $p_{\mathrm{m}}(m^\star) \geq 0.5$.
3. $p_{\mathrm{m}}^{\max} \leq \frac{1}{|\mathcal{K}|}$.

When $|\mathcal{M}| = 1$, every correct scheme is optimal. Fig. 8 illustrates the constructions for the three nontrivial cases. In Case 1 ($|\mathcal{M}| = |\mathcal{K}| = 2$), the two messages partition the ciphertext space into two equal parts, and the two keys swap the assignments of these parts. In Case 2 ($|\mathcal{K}| = 2$ and $p_{\mathrm{m}}(m^\star) \geq 0.5$), the maximum-probability message $m^\star$ occupies half of the ciphertext space, while the remaining messages jointly partition the other half. The construction used in Case 3 is also useful more generally. For each pair (m_i, k_j), define one circular interval with endpoints

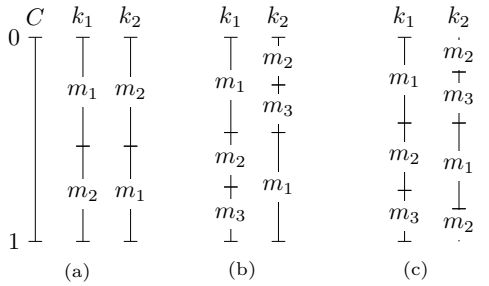


Fig. 8: p_{k}-agnostic scheme for any key distribution p_{k}: (a) $|\mathcal{M}| = |\mathcal{K}| = 2$; (b) $|\mathcal{M}| \geq 3$, $|\mathcal{K}| = 2$, $p_{\mathrm{m}}(m^\star) \geq 0.5$; (c) $p_{\mathrm{m}}(m^\star) \leq \frac{1}{|\mathcal{K}|}$.

$$a_{\mathsf{HES}}(m_i, k_j) \coloneqq F_{\mathrm{m}}(m_i^-) + (j-1)p_{\mathrm{m}}^{\max}, \qquad b_{\mathsf{HES}}(m_i, k_j) \coloneqq F_{\mathrm{m}}(m_i) + (j-1)p_{\mathrm{m}}^{\max},$$

and set $\mathcal{C}_{\mathsf{HES}}(m_i, k_j) \coloneqq [a_{\mathsf{HES}}(m_i, k_j), b_{\mathsf{HES}}(m_i, k_j))$. Encryption samples uniformly from this interval. We call the resulting HE-style shift construction HES. When $p_{\mathrm{m}}^{\max} \leq 1/|\mathcal{K}|$, all key shifts fit without decryption collisions and the construction achieves optimal MRS, as stated in Proposition 5.

Analysis of EE constructions. The following continuous counterpart of Theorem 7 shows that the obstruction persists even without discretization error.

Theorem 7′. *There exist distributions p_{m} and p_{k}, such that no continuous EE construction EE_{c} with a general encoder can achieve $p_{\mathrm{guess}}^{\mathsf{EE}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) = p_{\max}$.*

By leveraging continuous encryption, we explicitly compute the MRS guessing probability $p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C)$ for general EE constructions under a simple message distribution, thereby illustrating their insecurity. We leave the details in Appendix K.

Design and security of our OE. The continuous counterpart OE_{c} on $[0,1)$ is given by the region below, where $\mu(m,k) \coloneqq \min\{p_{\text{m}}(m), p_{\text{k}}(k)\}$:

$$\mathcal{C}(m,k) \coloneqq F_{\text{m}}(m^-) + F_{\text{k}}(k^-) + [0, \mu(m,k)).$$

Theorem 6′. *For arbitrary p_{m} and p_{k},*

$$p_{\text{guess}}^{\mathsf{OE}_{\text{c}},p_{\text{m}},p_{\text{k}}}(M \mid C) = p_{\max}.$$

Proof. The CDF-interval arguments of Lemma 2 apply without quantization, so the continuous regions are messagewise and keywise disjoint. Consequently, each nonzero $p_{\text{m,c}}(m,c)$ has a unique contributing key k. Uniformity and the definition of $\mu(m,k)$ therefore give

$$p_{\text{m,c}}(m,c) = \frac{p_{\text{m}}(m)p_{\text{k}}(k)}{\mu(m,k)} = \max\{p_{\text{m}}(m), p_{\text{k}}(k)\} \le p_{\max}.$$

Further,

$$p_{\max} \le p_{\text{guess}}^{\mathsf{OE}_{\text{c}},p_{\text{m}},p_{\text{k}}}(M \mid C) = \int_0^1 \max_m p_{\text{m,c}}(m,c)\,\mathrm{d}c \le p_{\max}.$$

Impossibility of simultaneously optimal KC and MRS. Some message and key distributions preclude simultaneous optimal KC and MRS; the following bound quantifies this obstruction.

Proposition 4. *For $p_{\text{m}} = (1/2, 1/2)$ and $p_{\text{k}} = (1/3, 1/3, 1/3)$, every encryption scheme satisfies*

$$p_{\text{guess}}(M \mid C) \ge \max\left\{\frac{1}{2},\ \frac{2}{3} - \sqrt{\frac{\ln 2}{2} I(K;C)}\right\}.$$

In particular, perfect KC forces $p_{\text{guess}}(M \mid C) \ge 2/3$, whereas $p_{\text{guess}}(M \mid C) = p_{\max} = 1/2$ requires $I(K;C) \ge 1/(18 \ln 2) \approx 0.0801$ bits. The latter lower bound is not asserted to be tight.

As an HE scheme, HES achieves optimal KC without depending on p_{k}. If $p_{\text{m}}^{\max} \le \frac{1}{|\mathcal{K}|}$, its shifts also avoid decryption collisions, and it achieves optimal MRS. This situation arises in many practical settings, such as when the message is a random secret (e.g., an RSA private key) and the key is a password.

Proposition 5. *For any $(p_{\text{m}}, p_{\text{k}})$ such that $p_{\text{m}}^{\max} \le 1/|\mathcal{K}|$, HES satisfies*

$$I(K;C) = 0, \qquad p_{\text{guess}}^{\mathsf{HES},p_{\text{m}},p_{\text{k}}}(M \mid C) = p_{\text{k}}^{\max} = p_{\max}.$$

6.4 Robustness to Distributional Mismatch

The preceding analysis assumes that the message and key distributions are known exactly. In practice, distributions of data such as passwords and natural-language text are typically estimated with some error. The following results quantify the resulting security loss. Let $p_{\mathrm{m}}, p_{\mathrm{k}}$ be the true distributions and $\hat{p}_{\mathrm{m}}, \hat{p}_{\mathrm{k}}$ their full-support estimates. The scheme is designed for $(\hat{p}_{\mathrm{m}}, \hat{p}_{\mathrm{k}})$, while messages and keys are independently drawn from $(p_{\mathrm{m}}, p_{\mathrm{k}})$.

Proposition 6. *For a continuous scheme achieving perfect KC under $(\hat{p}_{\mathrm{m}}, \hat{p}_{\mathrm{k}})$, the key leakage under $(p_{\mathrm{m}}, p_{\mathrm{k}})$ satisfies*

$$I(K;C) \le \mathrm{KL}(p_{\mathrm{m}} \| \hat{p}_{\mathrm{m}}) \le \frac{2\left(\mathrm{SD}(p_{\mathrm{m}}, \hat{p}_{\mathrm{m}})\right)^2}{\ln 2 \, \min_m \hat{p}_{\mathrm{m}}(m)}.$$

Proposition 7. *For the continuous scheme OE_{c} designed for $(\hat{p}_{\mathrm{m}}, \hat{p}_{\mathrm{k}})$, the excess recovery probability under $(p_{\mathrm{m}}, p_{\mathrm{k}})$ satisfies*

$$p_{\mathrm{guess}}(M \mid C) - p_{\max} \le \mathrm{SD}(p_{\mathrm{m}}, \hat{p}_{\mathrm{m}}) + \mathrm{SD}(p_{\mathrm{k}}, \hat{p}_{\mathrm{k}}) + \max\{\mathrm{SD}(p_{\mathrm{m}}, \hat{p}_{\mathrm{m}}), \mathrm{SD}(p_{\mathrm{k}}, \hat{p}_{\mathrm{k}})\}.$$

These bounds show how accurate distribution estimates can limit the security loss. Advances in statistical learning, particularly deep learning, have made it possible to model many practical data distributions accurately. Such models could extend the use of distribution-tailored encryption to a wider range of data, with the above bounds accounting for the security loss due to estimation errors.

6.5 Discussion

Randomness for information-theoretic security. Our results not only reaffirm the fundamental role of randomness in information-theoretic security but also clarify *how* it should be used: our uniform construction tailors ciphertext regions to the message and key distributions and samples uniformly from each region. For optimal KC, the ciphertext distribution must be independent of the key; for HE-style global uniformization, it suffices to choose $|\mathcal{C}(m,k)|$ proportional to $p_{\mathrm{m}}(m)$. For optimal MRS, choosing $|\mathcal{C}^{(\ell)}(m,k)| = 2^{\ell} \min\{p_{\mathrm{m}}(m), p_{\mathrm{k}}(k)\} + O(1)$ and avoiding decryption collisions yields $p_{\mathrm{guess}}(M \mid C) = p_{\max} + O(2^{-\ell})$.

Decryption collisions and message confidentiality. Decryption collisions can prevent optimal MRS, as illustrated in Fig. 7. As another example, for uniform message and key distributions with sufficiently large $|\mathcal{K}| = |\mathcal{M}|$, the random-oracle HE analysis gives an upper bound of $\frac{1}{|\mathcal{K}|} + \frac{3 \ln |\mathcal{K}|}{|\mathcal{K}| \ln \ln |\mathcal{K}|} + O(2^{-\ell})$ on recovery success [31], which likewise reflects the effect of decryption collisions (see Appendix B). By contrast, in this case, the one-time pad (OTP) achieves Shannon's perfect security with MRS guessing probability exactly $p_{\max}$. Our OE can be seen as a variant of the OTP that uses addition instead of XOR, thereby avoiding collisions. Similarly, when $p_{\mathrm{m}}^{\max} \le 1/|\mathcal{K}|$, the shift cipher in HES avoids collisions and achieves optimal MRS.

The advantage of our OE and HES over HE schemes. Existing analyses of HE provide meaningful bounds only under specific distributional regimes (see Appendix B), e.g., when $|\mathcal{M}| \gg |\mathcal{K}|$ and $p_{\mathrm{m}}^{\max} \ll p_{\mathrm{k}}^{\max}$. For general $(p_{\mathrm{m}}, p_{\mathrm{k}})$, these bounds can become loose or uninformative—for instance, when encrypting a 3-digit CVV under a 6-digit PIN. In contrast, OE achieves optimal MRS for arbitrary $(p_{\mathrm{m}}, p_{\mathrm{k}})$, while HES achieves both optimal KC and optimal MRS when $p_{\mathrm{m}}^{\max} \leq 1/|\mathcal{K}|$.

7 Conclusion and Future Work

We initiated a systematic study of the information-theoretic limits of *all* (possibly probabilistic) symmetric-encryption schemes under low-entropy keys, in a distributional model where messages and keys are drawn from arbitrary given distributions $(p_{\mathrm{m}}, p_{\mathrm{k}})$. We identified the optimal attainable levels of key and message confidentiality, measured by $I(K;C)$ and the MRS guessing probability, respectively, and tightly characterized when they are achievable. We also established impossibility results for distribution-agnostic schemes and for simultaneously achieving optimal KC and MRS under arbitrary distributions. We further quantified the security loss caused by estimation errors in p_{m} and p_{k}.

Technically, our KC characterization uses flows and cut counting on a weighted directed acyclic graph to derive ciphertext indistinguishability from message indistinguishability. We also introduced a continuous-ciphertext abstraction over $[0, 1)$ that separates structural constraints from discretization artifacts and may serve as a basic tool for future research on information-theoretic encryption. Our main message is that in this regime, the optimal attainable levels and the appropriate design principles are governed by *how a scheme uses randomness*.

Several directions remain open. One is to determine the optimal guarantees for stronger notions, such as TDSS and TDNM, and construct encryption schemes that attain them. Another is to incorporate computational hardness mechanisms into encryption schemes, including OE, to increase the adversary's computational cost while preserving their information-theoretic security guarantees.

References

1. Ali, S.M., Silvey, S.D.: A general class of coefficients of divergence of one distribution from another. Journal of the Royal Statistical Society: Series B (Methodological) **28**(1), 131–142 (1966)
2. Biondi, F., Given-Wilson, T., Legay, A.: Attainable unconditional security for shared-key cryptosystems. In: IEEE TrustCom (2015)
3. Biryukov, A., Dinu, D., Khovratovich, D.: Argon2: the memory-hard function for password hashing and other applications. In: EuroS&P 2016. pp. 292–303. IEEE (2016). https://doi.org/10.1109/EuroSP.2016.31
4. Blackwell, D.: Equivalent comparisons of experiments. Annals of Mathematical Statistics **24**(2), 265–272 (1953)

5. Blocki, J., Holman, B.: Data-dependent memory-hard functions: Sustained space and cumulative complexity trade-offs in the parallel random oracle model. In: CRYPTO 2026. pp. 35–64. Springer (2026). https://doi.org/10.1007/978-3-032-35412-9_2
6. Blom, R.J.: Bounds on key equivocation for simple substitution ciphers. IEEE Transactions on Information Theory (1979)
7. Blom, R.J.: An upper bound on the key equivocation for pure ciphers. IEEE Transactions on Information Theory (1984)
8. Boneh, D., Corrigan-Gibbs, H., Schechter, S.: Balloon hashing: A memory-hard function providing provable protection against sequential attacks. In: ASIACRYPT 2016, pp. 220–248. Springer (2016)
9. Bonneau, J.: The science of guessing: Analyzing an anonymized corpus of 70 million passwords. In: IEEE S&P 2012. pp. 538–552 (2012)
10. Borissov, Y.L., Lee, M.H.: Bounds on key appearance equivocation for substitution ciphers. IEEE Transactions on Information Theory (2007). https://doi.org/10.1109/TIT.2007.896865
11. Chatterjee, R., Bonneau, J., Juels, A., Ristenpart, T.: Cracking-resistant password vaults using natural language encoders. In: IEEE S&P 2015. pp. 481–498 (2015)
12. Chen, Y., Yu, Y., Zhai, L.: InfinityGauntlet: Expose smartphone fingerprint authentication to brute-force attack. In: USENIX Security 2023. pp. 2027–2041 (2023)
13. Cheng, H., Huang, F., Yang, J., Li, W., Wang, P.: Practically secure honey password vaults: New design and new evaluation against online guessing. In: USENIX Security 2025. pp. 7780–7798 (2025)
14. Cheng, H., Li, W., Wang, P., Chu, C.H., Liang, K.: Incrementally updateable honey password vaults. In: USENIX Security 2021. pp. 857–874 (2021)
15. Cheng, H., Zheng, Z., Li, W., Wang, P., Chu, C.H.: Probability model transforming encoders against encoding attacks. In: USENIX Security 2019. pp. 1573–1590 (2019)
16. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley-Interscience, Hoboken, New Jersey, 2nd edn. (July 2006)
17. Csiszár, I.: Information-type measures of difference of probability distributions and indirect observations. Studia Scientiarum Mathematicarum Hungarica **2**, 299–318 (1967)
18. Csiszár, I., Körner, J.: Information theory: coding theorems for discrete memoryless systems. Cambridge University Press (2011)
19. Dodis, Y., Reyzin, L., Smith, A.: Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In: EUROCRYPT 2004. pp. 523–540. Springer (2004)
20. Dodis, Y., Smith, A.: Entropic security and the encryption of high entropy messages. In: TCC 2005. pp. 556–577. Springer (2005)
21. Duan, M., Xu, M., Zhang, S., Han, W.: MoPE: A mixture of password experts for improving password guessing. In: IEEE S&P 2026 (2026)
22. Dunham, J.G.: Bounds on message equivocation for simple substitution ciphers. IEEE Transactions on Information Theory (1980)
23. Golla, M., Beuscher, B., Dürmuth, M.: On the security of cracking-resistant password vaults. In: ACM CCS 2016. pp. 1230–1241 (2016)
24. Hao, F., Anderson, R., Daugman, J.: Combining crypto with biometrics effectively. IEEE transactions on computers **55**(9), 1081–1088 (2006)
25. Housley, R.: Using advanced encryption standard (AES) counter mode with IPsec encapsulating security payload (ESP). RFC 3686, RFC Editor (Jan 2004). https://doi.org/10.17487/RFC3686, https://www.rfc-editor.org/rfc/rfc3686

26. Huang, Z., Ayday, E., Fellay, J., Hubaux, J.P., Juels, A.: Genoguard: Protecting genomic data against brute-force attacks. In: IEEE S&P 2015. pp. 447–462 (2015)
27. Issa, I., Kamath, S., Wagner, A.B.: An operational measure of information leakage. In: 2016 Annual Conference on Information Science and Systems (CISS). pp. 234–239. IEEE, Princeton, NJ, USA (March 2016)
28. Iwamoto, M., Ohta, K., Shikata, J.: Security formalizations and their relationships for encryption and key agreement in information-theoretic cryptography. IEEE Transactions on Information Theory (2018). https://doi.org/10.1109/TIT.2017.2744650
29. Jaeger, J., Ristenpart, T., Tang, Q.: Honey encryption beyond message recovery security. In: EUROCRYPT 2016. pp. 758–788 (2016)
30. Jendal, H.N., Kuhn, Y.J.B., Massey, J.L.: An information-theoretic treatment of homophonic substitution. In: EUROCRYPT 1989 (1989)
31. Juels, A., Ristenpart, T.: Honey encryption: Security beyond the brute-force bound. In: EUROCRYPT 2014. pp. 293–310. Springer (2014)
32. Kahn, D.: The Codebreakers: The Story of Secret Writing. Macmillan, New York, 1st edn. (1967)
33. Kaliski, B.: PKCS #5: Password-based cryptography specification version 2.0 (2000)
34. Kelley, P.G., Komanduri, S., Mazurek, M.L., Shay, R., Vidas, T., Bauer, L., Christin, N., Cranor, L.F., Lopez, J.: Guess again (and again and again): Measuring password strength by simulating password-cracking algorithms. In: IEEE S&P 2012. pp. 523–537 (2012)
35. Li, X., Tang, Q., Zhang, Z.: Fooling an unbounded adversary with a short key, repeatedly: The honey encryption perspective. In: International Conference on Information-Theoretic Cryptography (ITC 2021). LIPIcs (2021). https://doi.org/10.4230/LIPIcs.ITC.2021.23
36. Massey, J.L.: Probabilistic encipherment. E. & M. (Elektrotechnik & Maschinenbau, Austria) **104**(12), 561–562 (Dec 1987)
37. Rényi, A.: On measures of entropy and information. In: Neyman, J. (ed.) Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics. pp. 547–561. University of California Press, Berkeley and Los Angeles (1961)
38. Russell, A., Wang, H.: How to fool an unbounded adversary with a short key. In: EUROCRYPT 2002. pp. 133–148. Springer (2002)
39. Shannon, C.E.: Communication theory of secrecy systems. The Bell system technical journal **28**(4), 656–715 (1949)
40. Sibson, R.: Information radius. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **14**(2), 149–160 (1969)
41. Stahl, F.A.: A homophonic cipher for computational cryptography. In: AFIPS 1973. pp. 565–568. ACM (1973). https://doi.org/10.1145/1499586.1499729
42. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. In: NIPS 2017. vol. 30 (2017)
43. Verdú, S.: Total variation distance and the distribution of relative information. In: 2014 information theory and applications workshop (ITA). pp. 1–3. IEEE (2014)
44. Zou, Y., An, M., Wang, D.: Password guessing using large language models. In: USENIX Security 2025. pp. 7799–7818 (2025)

A Homophonic Ciphers

Homophonic substitution is a classical cryptographic technique, documented as early as 1401 and employed by the Duchy of Mantua to thwart frequency analysis [41, 32]. In its traditional setting, each plaintext symbol is mapped to one of several ciphertext symbols, with the mappings specifically chosen to uniformize the resulting ciphertext frequencies.

In the modern information-theoretic formulation of homophonic substitution [30], a message is first *encoded* into a perfectly uniform bit string before a conventional cipher is applied. To achieve *perfect* uniformization, the encoder utilizes variable-length codewords, which inherently yields ciphertexts with message-dependent lengths. For instance, given the toy example in Fig. 1 with a message space $\{m_1, m_2, m_3\}$ distributed as $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$, a variable-length homophonic encoder assigns codewords as follows: $m_1 \mapsto \mathtt{0}$, $m_2 \mapsto \mathtt{10}$, and $m_3 \mapsto \mathtt{11}$.

Furthermore, for general message distributions that cannot be expressed as finite dyadic fractions—such as the uniform distribution $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$—perfect uniformization requires codewords of potentially unbounded length. As illustrated by the infinite prefix tree in Fig. 9, the encoder must expand the mapping infinitely:

- $m_1 \mapsto (\mathtt{11})^k\mathtt{00}$ (e.g., 00, 1100, 111100, ...)
- $m_2 \mapsto (\mathtt{11})^k\mathtt{01}$ (e.g., 01, 1101, 111101, ...)
- $m_3 \mapsto (\mathtt{11})^k\mathtt{10}$ (e.g., 10, 1110, 111110, ...)

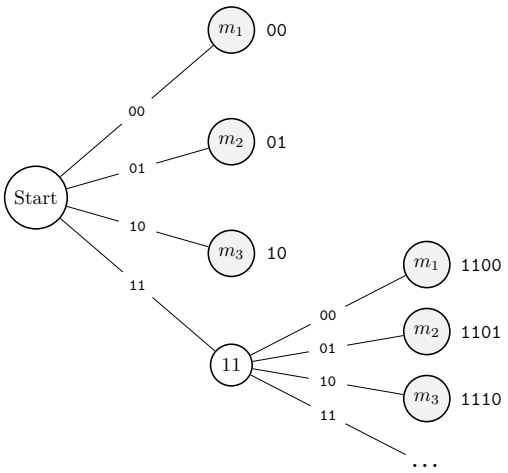


Fig. 9: The infinite prefix tree illustrating how a variable-length homophonic encoder maps codewords to messages m_1, m_2, m_3 under a uniform distribution $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. The tree continues to expand infinitely along the 11 branches.

Classical analyses often work in a long/infinite-message-sequence model, where subsequent ciphertext blocks can be viewed as padding to mask such length leakage, and argue that strong Shannon-style key equivocation is achieved in that idealized setting. In our single-message setting, however, variable ciphertext length constitutes an explicit side channel and prevents perfect key confidentiality in general.

B Existing Provable Bounds of HE Schemes

For MRS, Juels and Ristenpart [31] analyzed the HE scheme in the random oracle model only under several specific message and key distributions:

1. Uniform message and key distribution with $|\mathcal{K}| = |\mathcal{M}|$:
$$p_{\mathrm{guess}}^{\mathsf{HE},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) \leq \frac{1}{|\mathcal{K}|} + \frac{3\ln|\mathcal{K}|}{|\mathcal{K}|\ln\ln|\mathcal{K}|} + O(2^{-\ell}).$$
2. Uniform message distribution with $|\mathcal{K}| \ll |\mathcal{M}|$:
$$p_{\mathrm{guess}}^{\mathsf{HE},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) \leq p_{\mathrm{k}}^{\max}(1+\delta) + O(2^{-\ell}),$$
where $\alpha = \lceil 1/p_{\mathrm{k}}^{\max} \rceil$, $\delta = \frac{\alpha^2}{2|\mathcal{M}|} + \frac{e\alpha^4}{27|\mathcal{M}|^2} \cdot \left(1 - \frac{e\alpha^2}{|\mathcal{M}|^2}\right)^{-1}$, and e is the base of the natural logarithm. Here, δ is small when $|\mathcal{M}| \gg \alpha^2$.
3. General message and key distributions with $|\mathcal{K}| \ll |\mathcal{M}|$: If $p_{\mathrm{m}}^{\max} \leq 3 - \sqrt{5}$, then
$$p_{\mathrm{guess}}^{\mathsf{HE},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) \leq p_{\mathrm{k}}^{\max}(1+\delta) + O(2^{-\ell}),$$
where $\delta = \frac{\alpha^2}{2b} + \frac{e\alpha^4}{27b^2} \cdot \left(1 - \frac{e\alpha^2}{b^2}\right)^{-1}$, $\alpha = \lceil 3/p_{\mathrm{k}}^{\max} \rceil$, and $b = \lfloor 2/p_{\mathrm{m}}^{\max} \rfloor$.

The MRS bounds are useful in many practical cases, but they do not achieve the optimal $p_{\max} + \mathrm{negl}(\ell)$ even in the simplest setting with uniform messages and keys of equal length. In this setting, the one-time pad naturally satisfies Shannon's perfect secrecy, which in turn implies optimal MRS. For typical low-entropy keys such as passwords with $p_{\mathrm{k}}^{\max} = 0.01$, the MRS bound is 0.0529 when $p_{\mathrm{m}}^{\max} = 10^{-4}$ and 0.0125 when $p_{\mathrm{m}}^{\max} = 10^{-5}$.

For TDSS, Jaeger et al. [29] has analyzed the HE scheme in the random oracle model for general message and key distributions. For any s satisfying $\alpha^{s-1} \leq \frac{1}{2} - O(2^{-\ell})$,
$$\mathbf{Adv}_{\mathsf{HE},p_{\mathrm{m}},p_{\mathrm{k}}}^{\mathrm{tdss}} \leq p_{\mathrm{k}}^{\max} + \alpha^{s-1} + 2e^{-\alpha^{2s-1}/3} + O(2^{-\ell}),$$
where $\alpha \coloneqq \lceil 1/p_{\mathrm{k}}^{\max} \rceil$.

The TDSS bound vanishes as $p_{\mathrm{k}}^{\max} \to 0$, but it remains unsuitable for low-entropy keys such as passwords. For instance, when $p_{\mathrm{k}}^{\max} = 0.01$ (a typical value for passwords), the bound is 0.5579. Yet by definition, $\frac{1}{2}$ already serves as a trivial bound for $\mathbf{Adv}_{\mathsf{E},p_{\mathrm{m}},p_{\mathrm{k}}}^{\mathrm{tdss}}$.

Li, Tang, and Zhang [35] construct an HE scheme in the standard model using a pairwise-independent hash and prove the TDSS bound
$$\mathbf{Adv}_{\mathsf{HE},p_{\mathrm{m}},p_{\mathrm{k}}}^{\mathrm{tdss}} \leq 2\sqrt{p_{\mathrm{k}}^{\max}} + O(2^{-\ell}).$$

C Space of EE constructions.

EE constructions are inherently uniform and key-independent, and every ciphertext decrypts under every key; their ciphertext-region sizes may either depend on the message or be constant over messages. Conversely, the following proposition shows that any encryption scheme with these properties can be represented as an EE construction. This fully characterizes the space of EE constructions.

Proposition 8. *A probabilistic encryption scheme* $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \mathcal{E}, \mathcal{D})$ *can be represented as an EE construction* $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \mathcal{E}', \mathcal{D}')$ *with the same transition probabilities if and only if it is uniform and key-independent and satisfies*

$$D_k(c) \neq \perp \qquad \text{for every } k \in \mathcal{K} \text{ and } c \in \mathcal{C}.$$

In particular, for all $m \in \mathcal{M}$, $k \in \mathcal{K}$, *and* $c \in \mathcal{C}$,

$$\Pr[E_k(m) = c] = \Pr[E'_k(m) = c].$$

Proof. Every EE construction is uniform. Moreover, for each message m, its encoder preimage has a fixed size, and the cipher maps this preimage bijectively to $\mathcal{C}(m,k)$ under every key. Hence $|\mathcal{C}(m,k)|$ is independent of k, so every EE construction is KID. Moreover, its underlying cipher decrypts every ciphertext under every key, so $D_k(c) \neq \perp$ for all $k \in \mathcal{K}$ and $c \in \mathcal{C}$.

Conversely, suppose the scheme is uniform and KID and satisfies $D_k(c) \neq \perp$ for every k and c. For each key k, the regions $\{\mathcal{C}(m,k)\}_{m\in\mathcal{M}}$ therefore partition $\mathcal{C}$. Put $\mu_m = \mu(m,k) = |\mathcal{C}(m,k)|$, which is independent of k. Choose an encoder with $|\mathtt{decode}^{-1}(m)| = \mu_m$, so that $p_{\mathrm{d}}(m) = \mu_m/2^{\ell}$. For every key k, choose a bijection from $\mathtt{decode}^{-1}(m)$ to $\mathcal{C}(m,k)$ for each m; together these bijections define the deterministic cipher because both families partition their respective spaces. Uniformity then gives

$$\Pr[E_k(m) = c] = \Pr[E'_k(m) = c] = \begin{cases} 1/\mu_m, & c \in \mathcal{C}(m,k), \\ 0, & c \notin \mathcal{C}(m,k). \end{cases}$$

Together with the insufficiency results for EE constructions, this proposition shows that achieving optimal security necessarily requires schemes *tailored* to the key distribution. We establish this formally in Section 6 with Theorem 11.

D Results on Perfect KC

We explore the concept of perfect KC, which requires that the ciphertext C reveals no information about the key K. This concept is analogous to the perfect security defined by Shannon for message confidentiality [39].

Theorem 16. *A symmetric encryption scheme* $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \mathcal{E}, \mathcal{D})$ *achieves perfect KC,* if and only if *for each* $k \in \mathcal{K}$ *and* $c \in \mathcal{C}$, $p_{\mathrm{c|k}}(c\,|\,k) = p_{\mathrm{c}}(c)$, *i.e.,* $p_{\mathrm{c|k}}(c\,|\,k)$ *is independent of* k.

Proof. Echoing Shannon's proof for perfect security [39], $I(K;C) = 0$ if and only if K and C are independent, i.e., for each $k \in \mathcal{K}$ and $c \in \mathcal{C}$, $p_{\mathrm{c|k}}(c\,|\,k) = p_{\mathrm{c}}(c)$.

This means that, for each $k \in \mathcal{K}$, E_k transforms the message distribution p_{m} into the same ciphertext distribution p_{c}. Accordingly, for each $k \in \mathcal{K}$, D_k transforms the ciphertext distribution p_{c} back into the message distribution p_{m}. Thus, the honey-message distribution under each key coincides with the real distribution. Thus, we have the following proposition.

Definition 4. *A symmetric encryption scheme is* perfect M-IND*, if for each $k \in \mathcal{K}$, $M_1 \equiv M_{0,k}$, where $C_0 \leftarrow_{p_{\mathrm{c}}} \mathcal{C}, M_{0,k} \leftarrow D_k(C_0)$ and $M_1 \leftarrow_{p_{\mathrm{m}}} \mathcal{M}$.*

Proposition 9. *If a symmetric encryption scheme achieves perfect KC, then it satisfies perfect M-IND.*

However, unlike Theorem 16, the converse of Proposition 9 does not hold, because E_k can be randomized whereas D_k is deterministic. A randomized E_k may map p_{m} to different ciphertext distributions $p_{\mathrm{c|k}}(\cdot\,|\,k)$ under different keys, even though the corresponding D_k still maps the overall ciphertext distribution p_{c} back to p_{m}. Fig. 11 provides such a counterexample.

Fortunately, if E_k is uniform, meaning it assigns equal probability to all possible ciphertexts in $\mathcal{C}(m,k)$ for each message m and key k, the converse of Proposition 9 holds.

Theorem 17. *A* uniform *symmetric encryption scheme achieves perfect KC, if and only if it is perfect M-IND.*

More interestingly, a stronger result holds: $p_{\mathrm{c|k}}(c\,|\,k)$ and $p_{\mathrm{c}}(c)$ are not only equal but also uniformly distributed over certain subspaces of $\mathcal{C}$. These subspaces are defined by the partition $\mathcal{P} \coloneqq \mathcal{C}/R$ induced by an equivalence relation R. Here, R is the equivalence closure where $c_1, c_2 \in \mathcal{C}$ are related if and only if there exist $m \in \mathcal{M}$ and $k \in \mathcal{K}$ such that $c_1, c_2 \in \mathcal{C}(m,k)$.

Proposition 10. *If a uniform symmetric encryption scheme is perfect M-IND, then for every equivalence class S of R, the distributions p_{c} and $p_{\mathrm{c|k}}$ are identical and uniform over S, i.e., $p_{\mathrm{c}}(c_1) = p_{\mathrm{c}}(c_2) = p_{\mathrm{c|k}}(c_1 \mid k) = p_{\mathrm{c|k}}(c_2 \mid k)$ for all $c_1, c_2 \in S$ and $k \in \mathcal{K}$.*

Proof (Proof of Proposition 10). Since

$$\sum_m \Pr[D_k(C) = m] = \sum_m p_{\mathrm{m}}(m) = 1,$$

it follows that for each k, $D_k(C)$ never fails ($D_k(C) \neq \perp$).

The condition is equivalent to the following identity: for all $m \in \mathcal{M}, k \in \mathcal{K}$,

$$\begin{aligned}\sum_{c \in \mathcal{C}(m,k)} p_{\mathrm{c}}(c) &= p_{\mathrm{m}}(m) \\ &= \sum_{c \in \mathcal{C}(m,k)} p_{\mathrm{m}}(m) p_{\mathrm{c|m,k}}(c\,|\,m,k) \\ &= \sum_{c \in \mathcal{C}(m,k)} p_{\mathrm{c|k}}(c\,|\,k).\end{aligned}$$

Here, $p_{\mathrm{c|k}}(c\,|\,k)$ is constant over $\mathcal{C}(m,k)$, given by

$$p_{\mathrm{c|k}}(c\,|\,k) = \frac{p_{\mathrm{m}}(m)}{|\mathcal{C}(m,k)|}.$$

Intuitively, $p_{\mathrm{c}}(c)$ is the weighted average of $p_{\mathrm{c|k}}(c\,|\,k)$ over all k, with weights $p_{\mathrm{k}}(k)$. Typically, the maximum of $p_{\mathrm{c|k}}(c\,|\,k)$ is greater than the average $p_{\mathrm{c}}(c)$. However, the equation above implies that the maximum is equal to the average, forcing $p_{\mathrm{c|k}}(c\,|\,k) = p_{\mathrm{c}}(c)$. Based on this intuition, we now present a formal proof.

Formally, consider any equivalence class S of R, and let c_1, k_1 be the pair that maximizes $p_{\mathrm{c|k}}(c\,|\,k)$ over all $k \in \mathcal{K}, c \in S$. Define $m_1 \coloneqq D_{k_1}(c_1)$.

By the uniformity of encryption, for all $c \in \mathcal{C}(m_1,k_1)$,

$$p_{\mathrm{c|k}}(c\,|\,k_1) = p_{\mathrm{c|k}}(c_1\,|\,k_1) = \frac{p_{\mathrm{m}}(m_1)}{|\mathcal{C}(m_1,k_1)|}.$$

Therefore, we derive the following:

$$\begin{aligned}
\sum_{c\in\mathcal{C}(m_1,k_1)} p_{\mathrm{c|k}}(c\,|\,k_1) &= \sum_{c\in\mathcal{C}(m_1,k_1)} p_{\mathrm{c}}(c) \\
&= \sum_{c\in\mathcal{C}(m_1,k_1)} \sum_{k} p_{\mathrm{k}}(k) p_{\mathrm{c|k}}(c\,|\,k) \\
&\le \sum_{c\in\mathcal{C}(m_1,k_1)} \sum_{k} p_{\mathrm{k}}(k) p_{\mathrm{c|k}}(c\,|\,k_1) \\
&= \sum_{c\in\mathcal{C}(m_1,k_1)} p_{\mathrm{c|k}}(c\,|\,k_1).
\end{aligned}$$

Since both sides are equal, the inequality must hold with equality, forcing

$$p_{\mathrm{c|k}}(c\,|\,k) = p_{\mathrm{c|k}}(c\,|\,k_1) = p_{\mathrm{c|k}}(c_1\,|\,k_1)$$

for all $k \in \mathcal{K}$ and $c \in \mathcal{C}(m_1,k_1)$. Furthermore, for all $c \in \mathcal{C}(m_1,k_1)$,

$$p_{\mathrm{c}}(c) = \sum_{k} p_{\mathrm{k}}(k) p_{\mathrm{c|k}}(c\,|\,k) = p_{\mathrm{c|k}}(c_1\,|\,k_1).$$

This result shows that the maximum value $p_{\mathrm{c|k}}(c_1\,|\,k_1)$ propagates to all $p_{\mathrm{c}}(c)$ and $p_{\mathrm{c|k}}(c\,|\,k)$ within $\mathcal{C}(m_1,k_1)$. By the definition of equivalence classes, this propagation extends to the entire S through the same argument. In other words, p_{c} and $p_{\mathrm{c|k}}$ are identical and uniform over S.

Impossibility of perfect KC. For a general message distribution, such as the two-message distribution $\left(\frac{\sqrt{2}}{2}, 1-\frac{\sqrt{2}}{2}\right)$ illustrated in Fig. 10, it is impossible to achieve perfect KC with a finite ciphertext length ℓ and uniform encryption, except for trivial encryption schemes. Specifically, in these trivial schemes, E_k is the identity mapping id or any mapping independent of k, ensuring perfect KC. However, such trivial schemes fail to provide message confidentiality, as

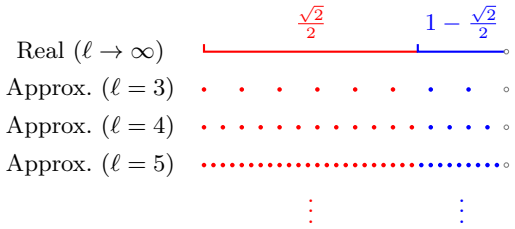


Fig. 10: Approximating the probability distribution $\left(\frac{\sqrt{2}}{2}, 1-\frac{\sqrt{2}}{2}\right)$ over $\{m_1, m_2\}$ by random ℓ bits (i.e., ℓ-bit ciphertexts/seeds).

the ciphertext completely determines the message without requiring the key. Except for these trivial schemes, we establish by contradiction that no uniform encryption scheme can achieve perfect KC.

Assume that a uniform encryption scheme achieves perfect KC. Then, by Proposition 10, $p_{\mathrm{c}} = p_{\mathrm{c|k}}(\cdot \mid k)$ must be uniform over each equivalence class $S \in \mathcal{C}/R$. Thus, for any $\mathcal{C}(m,k)$ and $\mathcal{C}(m',k')$ belonging to the same S,

$$\frac{p_{\mathrm{m}}(m)}{p_{\mathrm{m}}(m')} = \frac{|\mathcal{C}(m,k)|}{|\mathcal{C}(m',k')|} \in \mathbb{Q}.$$

However, since $\frac{p_{\mathrm{m}}(m_1)}{p_{\mathrm{m}}(m_2)} = 1+\sqrt{2} \notin \mathbb{Q}$, it follows that $\mathcal{C}(m_1,k)$ and $\mathcal{C}(m_2,k')$ cannot belong to the same S for any k,k'. Thus, $\mathcal{C}(m_1,k) \cap \mathcal{C}(m_2,k') = \emptyset$. Further, by $\mathcal{C}(m_1,k) \cup \mathcal{C}(m_2,k) = \mathcal{C}$, $\mathcal{C}(m_1,k)$ and $\mathcal{C}(m_2,k)$ remain unchanged across different keys, which are the only two equivalence classes of $\mathcal{C}$. This implies that the encryption scheme is trivial.

A non-uniform (and non-trivial) encryption scheme may achieve perfect KC, but it must assign *irrational* probabilities to ciphertexts, which cannot be sampled with any fixed number of fair random bits; worst-case randomness may be unbounded.

E Auxiliary Information-Theoretic Inequalities

Theorem 18 (Pinsker's inequality [18]). *For two random variables X_0 and X_1 on the space $\mathcal{X}$ with the distributions p_0 and p_1,*

$$\mathrm{SD}(X_0, X_1) \;\le\; \sqrt{\frac{\ln 2}{2}\,\mathrm{KL}(X_0 \| X_1)}\,.$$

Theorem 19 (Reverse Pinsker inequality; Theorem 7 in [43]). *For two random variables X_0 and X_1 on the space $\mathcal{X}$ with the distributions p_0 and p_1 such that for each $x \in \mathcal{X}$, $p_0(x) \le c \cdot p_1(x)$, where $c \ge 1$ is a constant:*

$$\mathrm{KL}(X_0 || X_1) \le \frac{c \log c}{c-1} \mathrm{SD}(X_0, X_1).$$

F Proofs in Section 4

Theorem 1 (restated). *A symmetric encryption scheme achieves statistical KC if and only if it satisfies C-IND in Fig. 3, i.e., for every key k, $C_1^{(\ell)} \overset{s}{\equiv} C_{0,k}^{(\ell)}$. Specifically,*

$$\mathrm{SD}(C_{0,k}^{(\ell)}, C_1^{(\ell)}) \le \sqrt{\frac{\ln 2}{2p_{\mathrm{k}}^{\min}} I(K; C^{(\ell)})} = O\left(\sqrt{I(K; C^{(\ell)})}\right), \qquad \forall k \in \mathcal{K},$$

and

$$I(K; C^{(\ell)}) \le \frac{\log(1/p_{\mathrm{k}}^{\min})}{1 - p_{\mathrm{k}}^{\min}} \max_k \mathrm{SD}(C_{0,k}^{(\ell)}, C_1^{(\ell)}) = O\left(\max_k \mathrm{SD}(C_{0,k}^{(\ell)}, C_1^{(\ell)})\right).$$

Proof (Proof of Theorem 1).

$$\begin{aligned} I(K; C^{(\ell)}) &= \sum_{k \in \mathcal{K}, c \in \mathcal{C}^{(\ell)}} p_{\mathrm{k}}(k) p_{\mathrm{c|k}}(c) \log \frac{p_{\mathrm{k}}(k) p_{\mathrm{c|k}}(c \,|\, k)}{p_{\mathrm{k}}(k) p_{\mathrm{c}}(c)} \\ &= \sum_{k \in \mathcal{K}} p_{\mathrm{k}}(k) \sum_{c \in \mathcal{C}^{(\ell)}} p_{\mathrm{c|k}}(c) \log \frac{p_{\mathrm{c|k}}(c \,|\, k)}{p_{\mathrm{c}}(c)} \\ &= \sum_{k \in \mathcal{K}} p_{\mathrm{k}}(k) \mathrm{KL}(C_{0,k}^{(\ell)} || C_1^{(\ell)}) \\ &= \mathop{\mathbb{E}}_{K \leftarrow_{p_{\mathrm{k}}} \mathcal{K}} [\mathrm{KL}(C_{0,K}^{(\ell)} || C_1^{(\ell)})]. \end{aligned}$$

With the equation, we can easily prove the necessity (i.e., the negligibility of $I(K; C^{(\ell)})$ implies that of $\mathrm{SD}(C_{0,k}^{(\ell)}, C_1^{(\ell)})$ for each $k \in \mathcal{K}$):

Mutual information is the prior-weighted average of the conditional KL divergences. Each divergence is therefore at most $I(K; C^{(\ell)})/p_{\mathrm{k}}^{\min}$. Pinsker's inequality [18] gives

$$\begin{aligned} \mathrm{SD}(C_{0,k}^{(\ell)}, C_1^{(\ell)}) &\le \sqrt{\frac{\ln 2}{2} \mathrm{KL}(C_{0,k}^{(\ell)} || C_1^{(\ell)})} \\ &\le \sqrt{\frac{\ln 2}{2p_{\mathrm{k}}^{\min}} \cdot I(K; C^{(\ell)})}. \end{aligned}$$

Thus, the necessity holds.

By a reverse Pinsker's inequality (Theorem 7 in [43]), we naturally get the sufficiency (i.e., the negligibility of $\mathrm{SD}(C_{0,k}^{(\ell)}, C_1^{(\ell)})$ for each $k \in \mathcal{K}$ implies that of $I(K; C^{(\ell)})$):

For each $k \in \mathcal{K}$, $p_{\mathrm{c}}(c) = \sum_{k \in \mathcal{K}} p_{\mathrm{k}}(k) p_{\mathrm{c|k}}(c \,|\, k)$, and

$$p_{\mathrm{c|k}}(c \,|\, k) \le p_{\mathrm{c}}(c) \cdot \frac{1}{p_{\mathrm{k}}^{\min}},$$

where $\frac{1}{p_{\mathrm{k}}^{\min}}$ is a constant. Applying this reverse Pinsker's inequality (X_0 and X_1 correspond to $C_{0,k}^{(\ell)}$ and $C_1^{(\ell)}$, meanwhile p_0 and p_1 correspond to $p_{\mathrm{c}|\mathrm{k}}$ and p_{c}),

$$\mathrm{KL}(C_{0,k}^{(\ell)} || C_1^{(\ell)}) \leq \frac{\log \frac{1}{p_{\mathrm{k}}^{\min}}}{1 - p_{\mathrm{k}}^{\min}} \mathrm{SD}(C_{0,k}^{(\ell)}, C_1^{(\ell)})$$

Further,

$$I(K; C^{(\ell)}) \leq \frac{\log \frac{1}{p_{\mathrm{k}}^{\min}}}{1 - p_{\mathrm{k}}^{\min}} \max_k \mathrm{SD}(C_{0,k}^{(\ell)}, C_1^{(\ell)})$$

Thus, the sufficiency holds.

Theorem 2 (restated). *A* uniform *encryption scheme achieves statistical KC if and only if it is M-IND. Specifically,*

$$\max_k \mathrm{SD}(M_{0,k}^{(\ell)}, M_1) \leq \max_k \mathrm{SD}(C_{0,k}^{(\ell)}, C_1^{(\ell)}) \leq \frac{|\mathcal{M}||\mathcal{K}|}{2p_{\mathrm{k}}^{\min}} \max_k \mathrm{SD}(M_{0,k}^{(\ell)}, M_1).$$

Proof (Proof of Theorem 2). Fix ℓ and omit its superscript. For this proof and its consequences, write $B = |\mathcal{M}||\mathcal{K}|$ and

$$\varepsilon_k = \mathrm{SD}(M_{0,k}, M_1), \quad \delta_k = \mathrm{SD}(C_{0,k}, C_1),$$
$$\overline{\varepsilon} = \sum_k p_{\mathrm{k}}(k)\varepsilon_k, \qquad \overline{\delta} = \sum_k p_{\mathrm{k}}(k)\delta_k.$$

We prove the stronger prior-weighted inequalities

$$\overline{\varepsilon} \leq \overline{\delta} \leq B\overline{\varepsilon}, \qquad \delta_k \leq \frac{B\overline{\varepsilon}}{2p_{\mathrm{k}}(k)}.$$

Set $\mathcal{M}' = \mathcal{M} \cup \{\perp\}$. To simplify notation, write $q_k(c) \coloneqq p_{\mathrm{c}|\mathrm{k}}(c \,|\, k)$ and $q(c) \coloneqq p_{\mathrm{c}}(c) = \sum_k p_{\mathrm{k}}(k) q_k(c)$. We construct a weighted directed acyclic graph whose nodes represent regions and whose edge flows measure weighted density differences on their overlaps. The net outgoing flow at each node equals its prior-weighted message-mass error, so the total positive node imbalance is $\overline{\varepsilon}$. Ordering the nodes by density and counting consecutive cuts bounds the total edge flow by $B\overline{\varepsilon}$. We then express this flow in terms of pairwise ciphertext distances and use the mixture identity to obtain the bounds on δ_k and $\overline{\delta}$.

Region graph and node imbalances. For every nonempty region $\mathcal{C}(m, k)$, $m \in \mathcal{M}$, make a node $v = (m, k)$ with region $A_v = \mathcal{C}(m, k)$ and density

$$a_v \coloneqq \frac{p_{\mathrm{m}}(m)}{\mu(m, k)}.$$

Uniformity gives $q_k(c) = a_v$ on A_v. Include a node for each nonempty failure region $A_{(\perp,k)} = \{c : D_k(c) = \perp\}$ and give it density zero; correctness implies $q_k = 0$ there. Thus the regions for each key partition $\mathcal{C}$. Ciphertexts with $q(c) = 0$

contribute no positive flow and need not be removed. For nodes $v = (m, k)$ and $u = (m', k')$ with $k \neq k'$, set

$$w_{vu} \coloneqq p_{\mathrm{k}}(k)p_{\mathrm{k}}(k')|A_v \cap A_u|;$$

set $w_{vu} = 0$ for nodes belonging to the same key. These weights are symmetric. Direct expansion yields the signed node imbalance

$$\begin{aligned} b_v \coloneqq \sum_u w_{vu}(a_v - a_u) &= p_{\mathrm{k}}(k) \sum_{c \in A_v} \big(q_k(c) - q(c)\big) \\ &= p_{\mathrm{k}}(k)\big(p_{\mathrm{m}}(m) - \Pr[D_k(C_1) = m]\big). \end{aligned} \tag{1}$$

The identity also holds for $m = \perp$ with $p_{\mathrm{m}}(\perp) = 0$. Both distributions being compared have total mass one on $\mathcal{M}'$. Consequently, writing $(x)_+ = \max\{x, 0\}$,

$$\sum_v (b_v)_+ = \sum_k p_{\mathrm{k}}(k)\varepsilon_k = \overline{\varepsilon}, \qquad \sum_v b_v = 0. \tag{2}$$

Acyclic flow and cut counting. Give an edge $v \to u$ flow $w_{vu}(a_v - a_u)$ whenever $a_v > a_u$; zero-flow edges are omitted. The net outgoing flow at v is b_v. All flows strictly decrease density, so the directed graph is acyclic. Every failure node has density zero and nonpositive imbalance. Merge these nodes into a single zero-density sink, summing parallel edge weights. Their mutual flows are zero, and their imbalances have the same sign; hence merging preserves total edge flow and the positive-imbalance sum in Eq. (2). The merged graph has $r \leq B + 1$ nodes; retain the notation w_{vu}, a_v, and b_v for this graph.

Order its nodes as $v_1, \ldots, v_r$ by nonincreasing density, breaking ties arbitrarily. For a prefix S of this order, define its outgoing cut flow by

$$F(S) \coloneqq \sum_{v \in S,\, u \notin S} w_{vu}(a_v - a_u).$$

No positive flow enters S from its complement, so

$$F(S) = \sum_{v \in S} b_v \leq \sum_v (b_v)_+ = \overline{\varepsilon}.$$

Let $S_i = \{v_1, \ldots, v_i\}$ for $1 \leq i < r$. Every positive-flow edge crosses at least one of these $r - 1 \leq B$ consecutive cuts. Thus, writing $T = \sum_{\{v,u\}} w_{vu}|a_v - a_u|$ for the total edge flow, counted once per unordered pair before merging, we obtain

$$T \leq \sum_{i=1}^{r-1} F(S_i) \leq (r-1)\overline{\varepsilon} \leq B\overline{\varepsilon}. \tag{3}$$

The argument uses neither a smallest positive edge weight nor a bound on the number of ciphertext points inside a region.

From edge flow to statistical distance. Each pair of key-indexed partitions covers $\mathcal{C}$ by their intersections. Before merging the failure nodes, this gives

$$T = \sum_{k \prec k'} p_{\mathrm{k}}(k) p_{\mathrm{k}}(k') \|q_k - q_{k'}\|_1. \tag{4}$$

Here $\prec$ is the fixed key order, so each pair is counted once. For each k define its incident flow, retaining the original key labels,

$$T_k := \sum_{k' \neq k} p_{\mathrm{k}}(k) p_{\mathrm{k}}(k') \|q_k - q_{k'}\|_1.$$

Then $T_k \leq T$ and $\sum_k T_k = 2T$. The mixture identity and the triangle inequality show that

$$\begin{aligned} \delta_k &= \tfrac{1}{2} \Big\| q_k - \sum_{k'} p_{\mathrm{k}}(k') q_{k'} \Big\|_1 \\ &\leq \tfrac{1}{2} \sum_{k'} p_{\mathrm{k}}(k') \|q_k - q_{k'}\|_1 = \frac{T_k}{2p_{\mathrm{k}}(k)} \leq \frac{B}{2p_{\mathrm{k}}(k)} \overline{\varepsilon}. \end{aligned}$$

Averaging before discarding the incident-flow information yields

$$\overline{\delta} \leq \tfrac{1}{2} \sum_k T_k = T \leq B \overline{\varepsilon}.$$

Conversely, correctness means that $D_k(C_{0,k})$ has the law of M_1, so data processing gives $\varepsilon_k \leq \delta_k$ for every k. Taking maxima proves the first inequality of the theorem, and averaging proves the weighted lower bound. Since $\overline{\varepsilon} \leq \max_k \varepsilon_k$ and $p_{\mathrm{k}}(k) \geq p_{\mathrm{k}}^{\min}$, the stated maximum-error bound follows. For fixed finite spaces and a full-support key distribution, M-IND therefore implies C-IND and hence statistical KC by Theorem 1. The other direction follows from Corollary 1. The singleton-key case has $\varepsilon_k = \delta_k = I(K; C) = 0$ throughout.

Proposition 11. *If a uniform encryption scheme is M-IND, then for each $k' \in \mathcal{K}$ there exists a random variable $U_{k'}^{(\ell)}$ on $\mathcal{C}^{(\ell)}$ whose probability mass function is constant within every decryption region under k', such that $C_{0,k}^{(\ell)} \overset{s}{\equiv} C_1^{(\ell)} \overset{s}{\equiv} U_{k'}^{(\ell)}$ for every $k \in \mathcal{K}$.*

Proof. Take $U_{k'}^{(\ell)}$ to have the same distribution as $C_{0,k'}^{(\ell)}$. Uniformity makes its probability mass function constant within every message region, and correctness makes it zero on the failure region. Theorem 2 gives C-IND, so each $C_{0,k}^{(\ell)}$ and $U_{k'}^{(\ell)}$ is statistically indistinguishable from $C_1^{(\ell)}$.

Lemma 1 (restated). *For any EE construction, let $M_1 \leftarrow_{p_{\mathrm{m}}} \mathcal{M}$, $M_0^{(\ell)} \leftarrow_{p_{\mathrm{d}}} \mathcal{M}$, $C_1^{(\ell)} \leftarrow_{p_{\mathrm{c}}} \mathcal{C}^{(\ell)}$, and $U^{(\ell)} \leftarrow_{\$} \mathcal{C}^{(\ell)}$. Then*

$$I(K; C^{(\ell)}) + \mathrm{KL}(C_1^{(\ell)} \| U^{(\ell)}) = \mathrm{KL}(M_1 \| M_0^{(\ell)}).$$

Proof (Proof of Lemma 1). Write $N = |\mathcal{C}^{(\ell)}| = |\mathcal{S}|$. Uniform encoding and the bijective outer cipher imply $|\mathcal{C}(m,k)| = Np_{\mathrm{d}}(m)$ and

$$p_{\mathrm{c|k}}(c\,|\,k) = \frac{p_{\mathrm{m}}(D_k(c))}{Np_{\mathrm{d}}(D_k(c))}.$$

Correctness and full support ensure $p_{\mathrm{d}}(m) > 0$ for every m. Grouping ciphertexts by their decryption output gives, for every k,

$$\mathrm{KL}(C_{0,k}^{(\ell)} \| U^{(\ell)}) = \sum_m p_{\mathrm{m}}(m) \log \frac{p_{\mathrm{m}}(m)}{p_{\mathrm{d}}(m)} = \mathrm{KL}(M_1 \| M_0^{(\ell)}).$$

On the other hand, splitting the logarithm through $p_{\mathrm{c}}(c)$ yields

$$\sum_k p_{\mathrm{k}}(k)\mathrm{KL}(C_{0,k}^{(\ell)} \| U^{(\ell)}) = I(K; C^{(\ell)}) + \mathrm{KL}(C_1^{(\ell)} \| U^{(\ell)}).$$

This proves the identity.

Theorem 3 (restated). *An EE construction with an encoder and a cipher achieves statistical KC if the encoder is MS-IND. Specifically, if* $\mathrm{SD}(M_0^{(\ell)}, M_1) < p_{\mathrm{m}}^{\min}$*, then*

$$I(K; C^{(\ell)}) \le \frac{2\big(\mathrm{SD}(M_0^{(\ell)}, M_1)\big)^2}{(p_{\mathrm{m}}^{\min} - \mathrm{SD}(M_0^{(\ell)}, M_1)) \ln 2} = O\Big(\big(\mathrm{SD}(M_0^{(\ell)}, M_1)\big)^2\Big).$$

Proof (Proof of Theorem 3). By Lemma 1 and nonnegativity of KL divergence, $I(K; C^{(\ell)}) \le \mathrm{KL}(M_1 \| M_0^{(\ell)})$. Using $\ln x \le x - 1$ gives

$$\mathrm{KL}(M_1 \| M_0^{(\ell)}) \le \frac{1}{\ln 2} \sum_m \frac{(p_{\mathrm{m}}(m) - p_{\mathrm{d}}(m))^2}{p_{\mathrm{d}}(m)}.$$

If $d = \mathrm{SD}(M_0^{(\ell)}, M_1) < p_{\mathrm{m}}^{\min}$, then $p_{\mathrm{d}}(m) \ge p_{\mathrm{m}}^{\min} - d$. The positive and negative parts of $p_{\mathrm{m}} - p_{\mathrm{d}}$ each sum to d, so the sum of their squared entries is at most $2d^2$. This gives the stated quadratic bound. For a uniform encoder the statistical distance between the real and honey message–seed pairs equals d, by the same grouping calculation. Thus MS-IND makes d negligible.

Proposition 1 (restated). *For an EE construction, if its encoder is MS-IND, then* $C_{0,k}^{(\ell)} \overset{s}{\equiv} C_1^{(\ell)} \overset{s}{\equiv} U^{(\ell)}$ *for every* $k \in \mathcal{K}$*. Specifically,* $\mathrm{SD}(C_1^{(\ell)}, U^{(\ell)}) \le \mathrm{SD}(C_{0,k}^{(\ell)}, U^{(\ell)}) = \mathrm{SD}(M_0^{(\ell)}, M_1)$ *for every* $k \in \mathcal{K}$*.*

Proof (Proof of Proposition 1). Write $N = |\mathcal{C}^{(\ell)}| = |\mathcal{S}|$ as in the proof of Lemma 1. Then

$$\mathrm{SD}(C_{0,k}^{(\ell)}, U^{(\ell)}) = \frac{1}{2} \sum_m Np_{\mathrm{d}}(m) \left| \frac{p_{\mathrm{m}}(m)}{Np_{\mathrm{d}}(m)} - \frac{1}{N} \right| = \frac{1}{2} \sum_m |p_{\mathrm{m}}(m) - p_{\mathrm{d}}(m)|.$$

The bound for $C_1^{(\ell)}$ follows by convexity of statistical distance under the mixture over k.

Proposition 2 (restated). *For an EE construction, if* $|p_{\mathrm{d}}(m) - p_{\mathrm{m}}(m)| \le \epsilon$ *for every* $m \in \mathcal{M}$*, where* $0 < \epsilon < p_{\mathrm{m}}^{\min}$*, then*

$$I(K; C^{(\ell)}) \le \frac{|\mathcal{M}|\epsilon^2}{(p_{\mathrm{m}}^{\min} - \epsilon)\ln 2} = O(\epsilon^2).$$

Proof (Proof of Proposition 2). Apply the KL bound in the proof of Theorem 3 and use $p_{\mathrm{d}}(m) \ge p_{\mathrm{m}}^{\min} - \epsilon$ and $\sum_m (p_{\mathrm{m}}(m) - p_{\mathrm{d}}(m))^2 \le |\mathcal{M}|\epsilon^2$.

Theorem 4 (restated). *For any* p_{m}*-agnostic encryption scheme* $\mathsf{E} = \{\mathsf{E}^{(\ell)}\}_\ell$*, if* E *achieves statistical KC under* p_{k} *for every message distribution* p_{m}*, then, for every* p_{m}*,*

$$p_{\mathrm{guess}}^{\mathsf{E}^{(\ell)}, p_{\mathrm{m}}, p_{\mathrm{k}}}(M \mid C^{(\ell)}) = 1 - \mathrm{negl}(\ell).$$

Proof (Proof of Theorem 4). Fix $m \in \mathcal{M}$ and $k, k_0 \in \mathcal{K}$. Apply statistical KC to the following two message distributions. The first selects m with probability $2/3$ and otherwise samples uniformly from $\mathcal{M}$; the second selects m with probability $1/3$ and otherwise samples uniformly from $\mathcal{M}$. Since E is p_{m}-agnostic, the same conditional ciphertext distributions are used in both cases. By Theorem 1, statistical KC implies

$$\mathrm{SD}\Bigg(a\, p^{(\ell)}_{\mathrm{c|m,k}}(\cdot \,|\, m, k) + \frac{1-a}{|\mathcal{M}|} \sum_{x \in \mathcal{M}} p^{(\ell)}_{\mathrm{c|m,k}}(\cdot \,|\, x, k), \\ a\, p^{(\ell)}_{\mathrm{c|m,k}}(\cdot \,|\, m, k_0) + \frac{1-a}{|\mathcal{M}|} \sum_{x \in \mathcal{M}} p^{(\ell)}_{\mathrm{c|m,k}}(\cdot \,|\, x, k_0) \Bigg) = \mathrm{negl}(\ell), \qquad a \in \left\{ \frac{1}{3}, \frac{2}{3} \right\}.$$

The coefficient matrix $\left(\begin{smallmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{smallmatrix}\right)$ is invertible. Applying its inverse to these two relations and using the triangle inequality for signed measures gives

$$\mathrm{SD}\big(p^{(\ell)}_{\mathrm{c|m,k}}(\cdot \,|\, m, k), p^{(\ell)}_{\mathrm{c|m,k}}(\cdot \,|\, m, k_0)\big) = \mathrm{negl}(\ell).$$

By correctness, decrypting with k_0 returns m with probability one for a ciphertext sampled from $p^{(\ell)}_{\mathrm{c|m,k}}(\cdot \,|\, m, k_0)$. The statistical distance bound shows that the same decoder returns m with probability $1 - \mathrm{negl}(\ell)$ for a ciphertext sampled from $p^{(\ell)}_{\mathrm{c|m,k}}(\cdot \,|\, m, k)$. Since m and k were arbitrary, always decrypting with k_0 recovers M with probability $1 - \mathrm{negl}(\ell)$.

Proposition 3 (restated). *For* n *ciphertexts* $C_1, \ldots, C_n$ *encrypted with the same key* K *(possibly under different encryption schemes), suppose the messages* $M_1, \ldots, M_n$ *are jointly independent of* K *and the encryption randomness is independent across encryptions and of the sources. Then*

$$I(K; C_1, \ldots, C_n) \le \sum_{i=1}^{n} I(K; C_i) + \mathrm{TC}(M_1, \ldots, M_n),$$

where $\mathrm{TC}(M_1, \ldots, M_n) = \sum_{i=1}^{n} H(M_i) - H(M_1, \ldots, M_n)$ *is the total correlation. In particular, if the messages are mutually independent, then*

$$I(K; C_1, \ldots, C_n) \leq \sum_{i=1}^{n} I(K; C_i).$$

Proof (Proof of Proposition 3). Write M^n and C^n for the message and ciphertext vectors, and define $\mathrm{TC}(C^n \mid K) = \sum_i H(C_i \mid K) - H(C^n \mid K)$. Expanding the entropies gives the exact identity

$$I(K; C^n) = \sum_i I(K; C_i) + \mathrm{TC}(C^n \mid K) - \mathrm{TC}(C^n).$$

For each fixed k, the channel from M^n to C^n is a product channel. Applying data processing to the divergence between the joint message law and the product of its marginals gives

$$\mathrm{TC}(C^n \mid K = k) \leq \mathrm{TC}(M^n).$$

Here $M^n \perp K$ ensures that the same message law applies to every k. Average this inequality and use $\mathrm{TC}(C^n) \geq 0$. If the messages are mutually independent, $\mathrm{TC}(M^n) = 0$, which gives the independent-message bound. For a first-order Markov source, the entropy chain rule gives $H(M^n) = H(M_1) + \sum_{i=2}^{n} H(M_i \mid M_{i-1})$, proving the last claim.

Theorem 20. *In the ideal-cipher model where $\mathcal{C} = \mathcal{M}$ and each E_k is an independent uniformly random permutation over $\mathcal{M}$, the expected key leakage satisfies*

$$\mathbb{E}[I(K; C)] \geq \frac{|\mathcal{M}| \sum_m p_{\mathrm{m}}(m)^2 - 1}{2 \ln 2 \cdot |\mathcal{M}| \cdot p_{\mathrm{m}}^{\max}} \left(1 - \sum_k p_{\mathrm{k}}(k)^2\right),$$

where the expectation is over the random choice of the ideal cipher.

Proof (Proof of Theorem 20). In the ideal cipher model, the randomness arises from the fact that each encryption function E_k is an independently and uniformly chosen permutation over $\mathcal{M}$. For any fixed set of permutations $\{E_k\}_k$, the cipher becomes deterministic.

Define $c \, R_{\mathrm{a}} \, c'$ when $D_k(c) = D_k(c')$ for every key k. Its classes are indexed by vectors in $\mathcal{M}^{|\mathcal{K}|}$, some of which may correspond to empty classes. For each class $S_{\boldsymbol{m}}$ with $\boldsymbol{m} \coloneqq (m_{j_i})_i$, we have $D_{k_i}(c) = m_{j_i}$ for each key k_i and each $c \in S_{\boldsymbol{m}}$.

Since the cipher is deterministic and $E_{k_i}(m_{j_i}) = c$, we have:

$$\begin{aligned} p_{\mathrm{c|k}}(c \,|\, k_i) &= p_{\mathrm{m}}(m_{j_i}) p_{\mathrm{c|m,k}}(c \,|\, m_{j_i}, k_i) \\ &= p_{\mathrm{m}}(m_{j_i}). \end{aligned}$$

Thus $I(K; C)$ can be rewritten as:

$$I(K; C) = \sum_c \sum_i p_{\mathrm{k}}(k_i) p_{\mathrm{c|k}}(c \,|\, k_i) \log \frac{p_{\mathrm{c|k}}(c \,|\, k_i)}{p_{\mathrm{c}}(c)}$$

$$= \sum_{\boldsymbol{m}} \sum_{c \in S_{\boldsymbol{m}}} \sum_{i} p_{\mathrm{k}}(k_i) p_{\mathrm{c|k}}(c \,|\, k_i) \log \frac{p_{\mathrm{c|k}}(c \,|\, k_i)}{\sum_{i'} p_{\mathrm{k}}(k_{i'}) p_{\mathrm{c|k}}(c \,|\, k_{i'})}$$
$$= \sum_{\boldsymbol{m}} |S_{\boldsymbol{m}}| \sum_{i} p_{\mathrm{k}}(k_i) p_{\mathrm{m}}(m_{j_i}) \log \frac{p_{\mathrm{m}}(m_{j_i})}{\sum_{i'} p_{\mathrm{k}}(k_{i'}) p_{\mathrm{m}}(m_{j_{i'}})}.$$

Since the randomness of the ideal cipher model only affects the value of $|S_{\boldsymbol{m}}|$, so

$$\mathbb{E}[I(K;C)] = \sum_{\boldsymbol{m}} \mathbb{E}[|S_{\boldsymbol{m}}|] \sum_{i} p_{\mathrm{k}}(k_i) p_{\mathrm{m}}(m_{j_i}) \cdot \log \frac{p_{\mathrm{m}}(m_{j_i})}{\sum_{i'} p_{\mathrm{k}}(k_{i'}) p_{\mathrm{m}}(m_{j_{i'}})}.$$

For the former part, noticed that $|S_{\boldsymbol{m}}|$ is binary, thus

$$\begin{aligned} \mathbb{E}[|S_{\boldsymbol{m}}|] &= \Pr[|S_{\boldsymbol{m}}| = 1] \\ &= \Pr[E_{k_i}(m_{j_i}) = E_{k_1}(m^{(1)}), \forall i \geq 2] \\ &= \frac{1}{|\mathcal{M}|^{|\mathcal{K}|-1}}. \end{aligned}$$

For the latter part, it can be transformed into

$$\begin{aligned} & \sum_{i} p_{\mathrm{k}}(k_i) p_{\mathrm{m}}(m_{j_i}) \log \frac{p_{\mathrm{m}}(m_{j_i})}{\sum_{i'} p_{\mathrm{k}}(k_{i'}) p_{\mathrm{m}}(m_{j_{i'}})} \\ =& \sum_{i} p_{\mathrm{k}}(k_i) p_{\mathrm{m}}(m_{j_i}) \log p_{\mathrm{m}}(m_{j_i}) \\ & - \sum_{i} p_{\mathrm{k}}(k_i) p_{\mathrm{m}}(m_{j_i}) \log \left(\sum_{i'} p_{\mathrm{k}}(k_{i'}) p_{\mathrm{m}}(m_{j_{i'}}) \right) \\ =& \sum_{i} \lambda_i f(x_i) - f(\overline{x}), \end{aligned}$$

where $\lambda_i \coloneqq p_{\mathrm{k}}(k_i)$, $x_i \coloneqq p_{\mathrm{m}}(m_{j_i})$, $\overline{x} \coloneqq \sum_i \lambda_i x_i$, and $f(x) \coloneqq x \log x$. According to Taylor's expansion,

$$\begin{aligned} \sum_{i} \lambda_i f(x_i) - f(\overline{x}) &= \sum_{i} \lambda_i \left(f(\overline{x}) + f'(\overline{x})(x_i - \overline{x}) + \frac{1}{2} f''(\xi_i)(x_i - \overline{x})^2 \right) - f(\overline{x}) \\ &= f(\overline{x}) + f'(\overline{x}) \sum_{i} \lambda_i (x_i - \overline{x}) + \frac{1}{2} \sum_{i} \lambda_i f''(\xi_i)(x_i - \overline{x})^2 - f(\overline{x}) \\ &= \frac{1}{2} \sum_{i} \lambda_i f''(\xi_i)(x_i - \overline{x})^2, \end{aligned}$$

where ξ_i lies between x_i and $\overline{x}$. It's clear that both of them can be upper-bounded by $p_{\mathrm{m}}^{\max}$. So $f''(\xi_i)$ has the following lower bound:

$$f''(\xi_i) = \frac{1}{\xi_i \ln 2} \geq \frac{1}{p_{\mathrm{m}}^{\max} \ln 2}.$$

Based on the above results, we have

$$\mathbb{E}[I(K;C)] \geq \frac{\sum_{\boldsymbol{m}} \sum_i p_{\mathrm{k}}(k_i) \left(p_{\mathrm{m}}(m_{j_i}) - \sum_{i'} p_{\mathrm{k}}(k_{i'}) p_{\mathrm{m}}(m_{j_{i'}})\right)^2}{2 \ln 2 |\mathcal{M}|^{|\mathcal{K}|-1} p_{\mathrm{m}}^{\max}}.$$

For the numerator part,

$$\begin{aligned}
&\sum_{\boldsymbol{m}} \sum_i p_{\mathrm{k}}(k_i) \left(p_{\mathrm{m}}(m_{j_i}) - \sum_{i'} p_{\mathrm{k}}(k_{i'}) p_{\mathrm{m}}(m_{j_{i'}})\right)^2 \\
=&\sum_{\boldsymbol{m}} \left(\sum_i p_{\mathrm{k}}(k_i) p_{\mathrm{m}}(m_{j_i})^2 - \left(\sum_i p_{\mathrm{k}}(k_i) p_{\mathrm{m}}(m_{j_i}) \right)^2 \right) \\
=&\sum_{\boldsymbol{m}} \left(\sum_i \left(p_{\mathrm{k}}(k_i) - p_{\mathrm{k}}(k_i)^2\right) p_{\mathrm{m}}(m_{j_i})^2 \right. \\
&\left. - \sum_{i \neq i'} p_{\mathrm{k}}(k_i) p_{\mathrm{k}}(k_{i'}) p_{\mathrm{m}}(m_{j_i}) p_{\mathrm{m}}(m_{j_{i'}}) \right) \\
=&\sum_i \left(p_{\mathrm{k}}(k_i) - p_{\mathrm{k}}(k_i)^2\right) \sum_{\boldsymbol{m}} p_{\mathrm{m}}(m_{j_i})^2 \\
&- \sum_{i \neq i'} p_{\mathrm{k}}(k_i) p_{\mathrm{k}}(k_{i'}) \sum_{\boldsymbol{m}} p_{\mathrm{m}}(m_{j_i}) p_{\mathrm{m}}(m_{j_{i'}}) \\
=&\left(1 - \sum_i p_{\mathrm{k}}(k_i)^2\right) |\mathcal{M}|^{|\mathcal{K}|-1} \sum_m p_{\mathrm{m}}(m)^2 \\
&- \left(1 - \sum_i p_{\mathrm{k}}(k_i)^2\right) |\mathcal{M}|^{|\mathcal{K}|-2} \left(\sum_m p_{\mathrm{m}}(m) \right)^2 \\
=&|\mathcal{M}|^{|\mathcal{K}|-2} \left(1 - \sum_i p_{\mathrm{k}}(k_i)^2\right) \left(|\mathcal{M}| \sum_m p_{\mathrm{m}}(m)^2 - 1 \right).
\end{aligned}$$

Therefore,

$$\mathbb{E}[I(K;C)] \geq \frac{|\mathcal{M}| \sum_m p_{\mathrm{m}}(m)^2 - 1}{2 \ln 2 |\mathcal{M}| p_{\mathrm{m}}^{\max}} \left(1 - \sum_k p_{\mathrm{k}}(k)^2\right).$$

G Bounds for HE schemes with Specific Well-designed Encoders

IS-DTE. Juels and Ristenpart [31] introduced a general approach to build encoders for arbitrary non-uniform distributions, resulting in what are known as IS-DTEs. For the public message order, let F_{m} be the ordinary CDF, with

$$F_{\mathrm{m}}(m_i^-) \coloneqq \sum_{t<i} p_{\mathrm{m}}(m_t), \qquad F_{\mathrm{m}}(m_i) \coloneqq F_{\mathrm{m}}(m_i^-) + p_{\mathrm{m}}(m_i).$$

If U is uniform on $[0,1)$, then $F_{\mathrm{m}}^{-1}(U)$ follows p_{m}. IS-DTE discretizes $[0,1)$ to $\mathcal{S} \coloneqq \{0,1,\ldots,2^{\ell}-1\} = \{0,1\}^{\ell}$ and assigns to m_i an integer seed chosen uniformly from

$$\big[\mathrm{round}(2^{\ell}F_{\mathrm{m}}(m_i^{-})), \mathrm{round}(2^{\ell}F_{\mathrm{m}}(m_i))\big).$$

Here $\ell \geq \lceil -\log \min_m p_{\mathrm{m}}(m) \rceil$ and $\lceil x \rceil$ is the least integer not smaller than x. This gives $|p_{\mathrm{d}}(m) - p_{\mathrm{m}}(m)| \leq 2^{-\ell}$ for every $m \in \mathcal{M}$.

Corollary 3. *An HE scheme built from a cipher and an IS-DTE with MS-IND achieves statistical KC. Specifically, if* $2^{-\ell} < p_{\mathrm{m}}^{\min}$*, then*

$$I(K;C^{(\ell)}) \leq \frac{|\mathcal{M}|\,2^{-2\ell}}{(p_{\mathrm{m}}^{\min} - 2^{-\ell})\ln 2} = O(2^{-2\ell}).$$

For messages with complex distributions, such as natural language texts, directly applying IS-DTE often leads to prohibitive time and space complexity. To address this, Cheng et al. [15] proposed a method to construct an efficient encoder based on a probability model, termed the probability-model-transforming encoder (PMTE). Similar to IS-DTE, IS-PMTE is designed using inverse sampling. During encoding, IS-PMTE parses the message into a sequence of generating rules and encodes each rule individually. Fix the message distribution and probability model independently of ℓ. Since discretization error vanishes as ℓ grows, MS-IND then requires the model's induced message distribution to equal p_{m}. Under this condition, $|p_{\mathrm{d}}(m) - p_{\mathrm{m}}(m)| \leq \frac{\ell_{\mathrm{seq}} n_{\mathrm{r}}}{2^{\ell/\ell_{\mathrm{seq}}}}$ holds for each message m, where ℓ_{seq} represents the maximum sequence length and n_{r} denotes the number of rules.

Corollary 4. *An HE scheme built from a cipher and an IS-PMTE with MS-IND achieves statistical KC. Specifically, let* $\epsilon_\ell = \ell_{\mathrm{seq}} n_{\mathrm{r}} 2^{-\ell/\ell_{\mathrm{seq}}}$*. If* $\epsilon_\ell < p_{\mathrm{m}}^{\min}$*, then*

$$I(K;C^{(\ell)}) \leq \frac{|\mathcal{M}|\epsilon_\ell^2}{(p_{\mathrm{m}}^{\min} - \epsilon_\ell)\ln 2} = O(2^{-2\ell/\ell_{\mathrm{seq}}}).$$

H Counterexamples to the Converse Implications

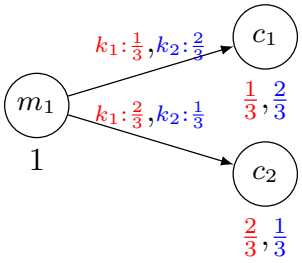


Fig. 11: Counterexample to the converse of Corollary 1

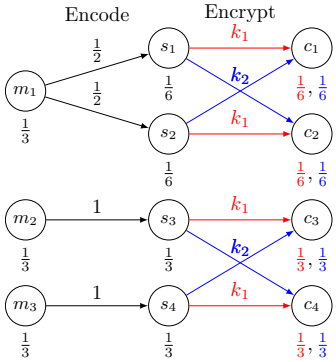


Fig. 12: Counterexample to the converse of Theorem 3: $p_{\mathrm{m}} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, and $p_{\mathrm{d}} = (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$

The converse of Corollary 1 does not hold in general. As illustrated in Fig. 11, a simple counterexample arises when, for the same message, E_k assigns different probabilities to different ciphertexts, resulting in distinct ciphertext distributions.

The converse of Theorem 3 does not hold. For certain ciphers with small key sizes, the ciphertext distribution $p_{\mathrm{c|k}}(\cdot \,|\, k)$ may remain consistent across different keys k, even when there are significant differences between p_{d} and p_{m}. Fig. 12 presents a trivial counterexample, which naturally leads to more complex ones.

I Implications of Mutual Information for Guessing-Based Security Notions

n-GS Game $\mathrm{GS}^{\mathcal{A}_n}_{p_{\mathrm{x,y}}}$

$X, Y \leftarrow_{p_{\mathrm{x,y}}} \mathcal{X} \times \mathcal{Y}$

$S \leftarrow \mathcal{A}_n(Y)$, where $|S| = n$

return $X \in S$

Fig. 13: n-Guessing Security Game (GS): guessing X given Y with n attempts

Security notions defined via mutual information inherently bound guessing-based security. We formalize this connection for general random variables X and Y, rather than restricting to specific cryptographic instantiations such as keys and ciphertexts.

Consider the n-guessing security game $\mathrm{GS}^{\mathcal{A}_n}_{p_{\mathrm{x,y}}}$ in Fig. 13. Given Y, the adversary $\mathcal{A}_n$ outputs a set S of n candidate values for X. The success probability of $\mathcal{A}_n$ is

$$\Pr\left[\mathrm{GS}^{\mathcal{A}_n}_{p_{\mathrm{x,y}}} \Rightarrow \mathrm{true}\right],$$

and the optimal n-guessing probability is

$$p^{p_{\mathrm{x,y}}}_{n\text{-guess}}(X \mid Y) \coloneqq \max_{\mathcal{A}_n} \Pr\left[\mathrm{GS}^{\mathcal{A}_n}_{p_{\mathrm{x,y}}} \Rightarrow \mathrm{true}\right].$$

The optimal strategy of $\mathcal{A}_n$ is to output the n most likely values under the posterior distribution $p_{\mathrm{x|y}}(\cdot \,|\, y)$. Let $S_{\mathrm{o}}(y, n)$ denote this optimal set for a given y, and let $S_{\mathrm{o}}(n)$ denote the set of n most likely values under the prior $p_{\mathrm{x}}(\cdot)$.

Define

$$p_{n\text{-guess}}(X \mid Y) \coloneqq \Pr[X \in S_{\mathrm{o}}(Y, n)], \quad p_{n\text{-guess}}(X) \coloneqq \Pr[X \in S_{\mathrm{o}}(n)].$$

The following theorem quantifies the gain from observing Y.

Theorem 21. *For any random variables X, Y and any integer $n \geq 1$,*

$$p_{n\text{-guess}}(X \mid Y) \leq p_{n\text{-guess}}(X) + \sqrt{\frac{\ln 2}{2} I(X;Y)}.$$

Proof. By Pinsker's inequality, for any y,

$$\mathrm{KL}\big(p_{\mathrm{x|y}}(\cdot \,|\, y) \,\|\, p_{\mathrm{x}}(\cdot)\big) \geq \frac{2}{\ln 2} \mathrm{SD}\big(p_{\mathrm{x|y}}(\cdot \,|\, y), p_{\mathrm{x}}(\cdot)\big)^2.$$

Hence,

$$\mathrm{SD}\big(p_{\mathrm{x|y}}(\cdot \,|\, y), p_{\mathrm{x}}(\cdot)\big) \leq \sqrt{\frac{\ln 2}{2} \mathrm{KL}\big(p_{\mathrm{x|y}}(\cdot \,|\, y) \,\|\, p_{\mathrm{x}}(\cdot)\big)}.$$

Moreover, by the definition of statistical distance,

$$\sum_{x \in S_{\mathrm{o}}(y,n)} \big(p_{\mathrm{x|y}}(x \,|\, y) - p_{\mathrm{x}}(x)\big) \leq \mathrm{SD}\big(p_{\mathrm{x|y}}(\cdot \,|\, y), p_{\mathrm{x}}(\cdot)\big).$$

Combining the above, we obtain

$$\begin{aligned}
&p_{n\text{-guess}}(X \mid Y) \\
=&\sum_y p_{\mathrm{y}}(y) \sum_{x \in S_{\mathrm{o}}(y,n)} p_{\mathrm{x|y}}(x \,|\, y) \\
\leq&\sum_y p_{\mathrm{y}}(y) \sum_{x \in S_{\mathrm{o}}(y,n)} p_{\mathrm{x}}(x) + \sum_y p_{\mathrm{y}}(y) \sqrt{\frac{\ln 2}{2} \mathrm{KL}\big(p_{\mathrm{x|y}}(\cdot \,|\, y) \,\|\, p_{\mathrm{x}}(\cdot)\big)} \\
\leq& \sum_{x \in S_{\mathrm{o}}(n)} p_{\mathrm{x}}(x) + \sqrt{\frac{\ln 2}{2} \sum_y p_{\mathrm{y}}(y) \mathrm{KL}\big(p_{\mathrm{x|y}}(\cdot \,|\, y) \,\|\, p_{\mathrm{x}}(\cdot)\big)} \\
=&p_{n\text{-guess}}(X) + \sqrt{\frac{\ln 2}{2} I(X;Y)},
\end{aligned}$$

where the second inequality follows from Jensen's inequality.

We next extend the result to partial-information recovery. Many guessing-based notions (e.g., TDSS) aim to recover a function of X rather than X itself.

Corollary 5. *For any random variables X, Y and any function $f : \mathcal{X} \to \mathcal{Z}$, let $\omega_f := \max_z \Pr[f(X) = z]$. Then*

$$p_{\text{guess}}(f(X) \mid Y) \le \omega_f + \sqrt{\frac{\ln 2}{2} I(X;Y)}.$$

Proof. Applying Theorem 21 with $n = 1$ and replacing X by $f(X)$ gives

$$p_{\text{guess}}(f(X) \mid Y) \le \omega_f + \sqrt{\frac{\ln 2}{2} I(f(X);Y)}.$$

Since $Y \to X \to f(X)$ forms a Markov chain, the data processing inequality implies $I(f(X);Y) \le I(X;Y)$.

Theorem 21 and Corollary 5 together show that statistical key confidentiality, expressed via mutual information, automatically bounds all guessing-based key recovery attacks, including those targeting arbitrary partial information and allowing multiple guesses.

J Proofs in Section 5

Lemma 2 (restated). *The ciphertext regions of* OE *are both keywise and messagewise disjoint: for each fixed message, they are pairwise disjoint across keys; for each fixed key, they are pairwise disjoint across messages.*

Proof (Proof of Lemma 2). By construction,

$$\mathcal{C}^{(\ell)}(m,k) = \big[\widetilde{F}^{(\ell)}_{\mathrm{m}}(m^-) + \widetilde{F}^{(\ell)}_{\mathrm{k}}(k^-), \widetilde{F}^{(\ell)}_{\mathrm{m}}(m^-) + \widetilde{F}^{(\ell)}_{\mathrm{k}}(k^-) + \mu^{(\ell)}(m,k)\big).$$

For a fixed message m, since $\mu^{(\ell)}(m,k) \le \widetilde{p}^{(\ell)}_{\mathrm{k}}(k)$,

$$\mathcal{C}^{(\ell)}(m,k) \subseteq \widetilde{F}^{(\ell)}_{\mathrm{m}}(m^-) + \mathcal{I}^{(\ell)}_{\mathrm{k}}(k).$$

The key CDF intervals $\mathcal{I}^{(\ell)}_{\mathrm{k}}(k)$ are pairwise disjoint across k, and a common circular translation preserves disjointness. Hence, for every $k \neq k'$,

$$\mathcal{C}^{(\ell)}(m,k) \cap \mathcal{C}^{(\ell)}(m,k') = \emptyset,$$

which proves keywise disjointness.

Similarly, for a fixed key k, since $\mu^{(\ell)}(m,k) \le \widetilde{p}^{(\ell)}_{\mathrm{m}}(m)$,

$$\mathcal{C}^{(\ell)}(m,k) \subseteq \widetilde{F}^{(\ell)}_{\mathrm{k}}(k^-) + \mathcal{I}^{(\ell)}_{\mathrm{m}}(m).$$

The message CDF intervals $\mathcal{I}^{(\ell)}_{\mathrm{m}}(m)$ are pairwise disjoint across m. Thus, for every $m \neq m'$,

$$\mathcal{C}^{(\ell)}(m,k) \cap \mathcal{C}^{(\ell)}(m',k) = \emptyset,$$

which proves messagewise disjointness.

Theorem 6 (restated). *For the discrete scheme* OE*, if* $\ell \geq \ell_{\mathsf{OE}}^{\min}$*, then*

$$p_{\text{guess}}^{\mathsf{OE},p_{\text{m}},p_{\text{k}}}(M \mid C) \leq \frac{2^{\ell}\mu_{\min}}{2^{\ell}\mu_{\min}-1}p_{\max} = p_{\max} + O(2^{-\ell}).$$

Proof (Proof of Theorem 6). Let $\mu(m,k) \coloneqq \min\{p_{\text{m}}(m), p_{\text{k}}(k)\}$. Coherent quantization of adjacent CDF boundaries gives, for $\text{x} \in \{\text{m},\text{k}\}$,

$$\widetilde{p}_{\text{x}}^{(\ell)}(x) = \lfloor 2^{\ell}F_{\text{x}}(x)\rfloor - \lfloor 2^{\ell}F_{\text{x}}(x^{-})\rfloor > 2^{\ell}p_{\text{x}}(x) - 1.$$

It follows that

$$\mu^{(\ell)}(m,k) > 2^{\ell}\mu(m,k) - 1.$$

By Lemma 2, for each fixed (m,c) at most one key contributes to $p_{\text{m,c}}(m,c)$. Therefore,

$$\begin{aligned} p_{\text{guess}}^{\mathsf{OE},p_{\text{m}},p_{\text{k}}}(M \mid C) &= \sum_{c}\max_{m} p_{\text{m,c}}(m,c) \\ &\leq 2^{\ell}\max_{m,k}\frac{p_{\text{m}}(m)p_{\text{k}}(k)}{\mu^{(\ell)}(m,k)}. \end{aligned}$$

For every (m,k),

$$\begin{aligned} \frac{p_{\text{m}}(m)p_{\text{k}}(k)}{\mu^{(\ell)}(m,k)} &< \frac{p_{\text{m}}(m)p_{\text{k}}(k)}{\mu(m,k)}\frac{\mu(m,k)}{2^{\ell}\mu(m,k)-1} \\ &\leq p_{\max}\frac{\mu_{\min}}{2^{\ell}\mu_{\min}-1}. \end{aligned}$$

The first factor equals $\max\{p_{\text{m}}(m), p_{\text{k}}(k)\} \leq p_{\max}$. The second is at most its value at $\mu_{\min}$ because $x/(2^{\ell}x-1)$ is decreasing for $x > 2^{-\ell}$ and $\mu(m,k) \geq \mu_{\min}$. Hence,

$$p_{\text{guess}}^{\mathsf{OE},p_{\text{m}},p_{\text{k}}}(M \mid C) \leq \frac{2^{\ell}\mu_{\min}}{2^{\ell}\mu_{\min}-1}p_{\max} = p_{\max} + O(2^{-\ell}),$$

as claimed.

Theorem 7 (restated). *For* $\mathcal{M} = \{a,b\}$ *with* $p_{\text{m}} = (\frac{1}{2},\frac{1}{2})$ *and any* $\mathcal{K}$ *such that* $p_{\text{k}}^{\max} \leq \frac{1}{2}$ *and* $p_{\text{k}}(\mathcal{K}_1) \neq p_{\text{k}}(\mathcal{K}_2)$ *for every two distinct subsets* $\mathcal{K}_1, \mathcal{K}_2 \subseteq \mathcal{K}$*, it holds that*

$$\inf_{\mathsf{EE}\in\Omega_{EE}} p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C) > p_{\max} = \tfrac{1}{2}.$$

Proof (Proof of Theorem 7). Let $\mu_m \coloneqq \mu(m,k)$ as in EE constructions, $\mu(m,k)$ is independent of k. Without loss of generality, assume that $\mu_a \geq \mu_b$. Then

$$\begin{aligned} p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C) &= \sum_{c\in\mathcal{C}}\max_{m}\left\{\sum_{k:D_k(c)=m}\frac{p_{\text{k}}(k)p_{\text{m}}(m)}{\mu(m,k)}\right\} \\ &= \sum_{c\in\mathcal{C}}\max\left\{\sum_{k:D_k(c)=a}\frac{p_{\text{k}}(k)p_{\text{m}}(a)}{\mu(a,k)}, \sum_{k:D_k(c)=b}\frac{p_{\text{k}}(k)p_{\text{m}}(b)}{\mu(b,k)}\right\} \end{aligned}$$

$$= \sum_{c\in\mathcal{C}} \max\left\{ \frac{p_{\mathrm{m}}(a)}{\mu_a} \sum_{k:D_k(c)=a} p_{\mathrm{k}}(k), \frac{p_{\mathrm{m}}(b)}{\mu_b} \sum_{k:D_k(c)=b} p_{\mathrm{k}}(k) \right\}.$$

Note that each term of the summation is related to its decryption result under each key. Define $f : \mathcal{C} \to \mathcal{M}^{|\mathcal{K}|}$ by $f(c) \coloneqq (D_{k_i}(c))_{i=1}^{|\mathcal{K}|}$. Using f, the ciphertext space is partitioned into distinct equivalence classes $\mathcal{C}_x$, defined for each $x \in \mathcal{M}^{|\mathcal{K}|}$ by $\mathcal{C}_x \coloneqq \{c \mid f(c) = x\}$. Further, define $g(x) \coloneqq \sum_{i:x_i=a} p_{\mathrm{k}}(k_i)$. Then

$$\begin{aligned}
p^{\mathsf{EE},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) &= \sum_{c\in\mathcal{C}} \max\left\{ \frac{p_{\mathrm{m}}(a)}{\mu_a} \sum_{k:D_k(c)=a} p_{\mathrm{k}}(k), \frac{p_{\mathrm{m}}(b)}{\mu_b} \sum_{k:D_k(c)=b} p_{\mathrm{k}}(k) \right\} \\
&= \sum_{x\in\mathcal{M}^{|\mathcal{K}|}} \sum_{c\in\mathcal{C}_x} \max\left\{ \frac{p_{\mathrm{m}}(a)}{\mu_a} \sum_{i:x_i=a} p_{\mathrm{k}}(k_i), \frac{p_{\mathrm{m}}(b)}{\mu_b} \sum_{i:x_i=b} p_{\mathrm{k}}(k_i) \right\} \\
&= \sum_{x\in\mathcal{M}^{|\mathcal{K}|}} |\mathcal{C}_x| \cdot \max\left\{ \frac{p_{\mathrm{m}}(a)}{\mu_a} g(x), \frac{p_{\mathrm{m}}(b)}{\mu_b}(1-g(x)) \right\} \\
&= \frac{1}{2} \sum_{x\in\mathcal{M}^{|\mathcal{K}|}} |\mathcal{C}_x| \cdot \left(\frac{p_{\mathrm{m}}(a)}{\mu_a} g(x) + \frac{p_{\mathrm{m}}(b)}{\mu_b}(1-g(x)) \right. \\
&\qquad\qquad \left. + \left| \frac{p_{\mathrm{m}}(a)}{\mu_a} g(x) - \frac{p_{\mathrm{m}}(b)}{\mu_b}(1-g(x)) \right| \right) \\
&= \frac{1}{2} + \frac{1}{2} \sum_{x\in\mathcal{M}^{|\mathcal{K}|}} |\mathcal{C}_x| \cdot \left| \left(\frac{p_{\mathrm{m}}(a)}{\mu_a} + \frac{p_{\mathrm{m}}(b)}{\mu_b} \right) g(x) - \frac{p_{\mathrm{m}}(b)}{\mu_b} \right| \\
&= \frac{1}{2} + \frac{1}{2} \sum_{x\in\mathcal{M}^{|\mathcal{K}|}} |\mathcal{C}_x| \cdot |(\alpha+\beta)\, g(x) - \beta| \,,
\end{aligned}$$

here $\alpha \coloneqq \frac{p_{\mathrm{m}}(a)}{\mu_a} = \frac{1}{2\mu_a}$ and $\beta \coloneqq \frac{p_{\mathrm{m}}(b)}{\mu_b} = \frac{1}{2\mu_b}$. Notice that $\alpha \le \beta$ and

$$\alpha + \beta = \frac{1}{2\mu_a} + \frac{1}{2\mu_b} \ge \frac{2}{\mu_a + \mu_b} = \frac{2}{|\mathcal{C}|}.$$

Denote $\mathcal{T} \coloneqq \{g(x) \mid x \in \mathcal{M}^{|\mathcal{K}|}\}$. Since $p_{\mathrm{k}}(\mathcal{K}_1) \neq p_{\mathrm{k}}(\mathcal{K}_2)$ for any subsets $\mathcal{K}_1 \neq \mathcal{K}_2$, $\mathcal{T}$ consists of $|\mathcal{M}|^{|\mathcal{K}|}$ distinct elements. Arrange the elements of $\mathcal{T}$ in ascending order as $\{t_i\}_{i=0}^{|\mathcal{T}|-1}$. Each $t_i \in \mathcal{T}$ corresponds to a unique element $x_i \in \mathcal{M}^{|\mathcal{K}|}$. In particular, $x_0 = (b, b, \ldots, b)$ and $t_0 = 0$, while $x_{|\mathcal{T}|-1} = (a, a, \ldots, a)$ and $t_{|\mathcal{T}|-1} = 1$. Denoting $\delta \coloneqq \min_{\substack{t,t'\in\mathcal{T} \\ t\neq t'}} |t - t'| > 0$, we provide a lower bound of $p^{\mathsf{EE},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C)$ based on δ. We discuss the bound in two cases: 1) $\mu_a \le t_{|\mathcal{T}|-2}|\mathcal{C}|$ and 2) $\mu_a > t_{|\mathcal{T}|-2}|\mathcal{C}|$.

Case 1: $\mu_a \le t_{|\mathcal{T}|-2}|\mathcal{C}|$:

Define i by $t_i < \mu_a/|\mathcal{C}| \le t_{i+1}$, and let $\mathcal{S} \coloneqq \mathcal{M}^{|\mathcal{K}|} \backslash \{x_i, x_{i+1}\}$. Since $\frac{\beta}{\alpha+\beta} = \mu_a/|\mathcal{C}| \in (t_i, t_{i+1}]$, we have $(\alpha+\beta)t_i < \beta$ and $(\alpha+\beta)t_{i+1} \ge \beta$

If $|\mathcal{C}_{x_i}|+|\mathcal{C}_{x_{i+1}}| < 3|\mathcal{C}|/4$: $\sum_{x\in\mathcal{S}} |\mathcal{C}_x| \geq |\mathcal{C}|/4$. For $j > i+1$, $|(\alpha+\beta)g(x_j)-\beta| = (\alpha+\beta)t_j - \beta \geq (\alpha+\beta)(t_{i+1}+\delta) - \beta \geq (\alpha+\beta)\delta \geq 2\delta/|\mathcal{C}|$. Similarly, for $j < i$, $|(\alpha+\beta)g(x_j)-\beta| = \beta - (\alpha+\beta)t_j \geq \beta - (\alpha+\beta)(t_i-\delta) \geq (\alpha+\beta)\delta \geq 2\delta/|\mathcal{C}|$. Combining the inequalities together yields

$$\begin{aligned}
p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C) &\geq \frac{1}{2} + \frac{1}{2}\sum_{x\in\mathcal{M}^{|\mathcal{K}|}} |\mathcal{C}_x| \cdot |(\alpha+\beta)\, g(x) - \beta| \\
&\geq \frac{1}{2} + \frac{1}{2}\sum_{x\in\mathcal{S}} |\mathcal{C}_x| \cdot |(\alpha+\beta)\, g(x) - \beta| \\
&\geq \frac{1}{2} + \frac{1}{2}\cdot\frac{2\delta}{|\mathcal{C}|}\sum_{x\in\mathcal{S}} |\mathcal{C}_x| \\
&\geq \frac{1}{2} + \frac{\delta}{4}.
\end{aligned}$$

If $|\mathcal{C}_{x_i}|+|\mathcal{C}_{x_{i+1}}| \geq 3|\mathcal{C}|/4$: Since $t_i < \mu_a/|\mathcal{C}| \leq t_{|\mathcal{T}|-2}$, we have $i < |\mathcal{T}|-2$. Besides, $t_{i+1} \geq \mu_a/|\mathcal{C}| \geq \frac{1}{2} > t_1$ implies $i > 1$. Note that for $x \in \mathcal{M}^{|\mathcal{K}|}\backslash\{x_0, x_{|\mathcal{T}|-1}\}$, $|\mathcal{C}_x| \leq |\mathcal{C}|/2$ (otherwise, there exist j, j' such that $x_j = a$ and $x_{j'} = b$, and further $\mu_a > |\mathcal{C}_x| > |\mathcal{C}|/2$ and $\mu_b > |\mathcal{C}_x| > |\mathcal{C}|/2$, leading to contradiction). Therefore, $|\mathcal{C}_{x_i}| \geq 3|\mathcal{C}|/4 - |\mathcal{C}_{x_{i+1}}| \geq |\mathcal{C}|/4$. Similarly, $|\mathcal{C}_{x_{i+1}}| \geq |\mathcal{C}|/4$. Thus,

$$\begin{aligned}
p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C) &\geq \frac{1}{2} + \frac{1}{2}\sum_{x\in\mathcal{M}^{|\mathcal{K}|}} |\mathcal{C}_x| \cdot |(\alpha+\beta)\, g(x) - \beta| \\
&\geq \frac{1}{2} + \frac{1}{2}|\mathcal{C}_{x_i}| \cdot |(\alpha+\beta)\, g(x_i) - \beta| \\
&\quad + \frac{1}{2}|\mathcal{C}_{x_{i+1}}| \cdot |(\alpha+\beta)\, g(x_{i+1}) - \beta| \\
&\geq \frac{1}{2} + \frac{|\mathcal{C}|}{8}\,|(\alpha+\beta)\, g(x_i) - \beta| \\
&\quad + \frac{|\mathcal{C}|}{8}\,|(\alpha+\beta)\, g(x_{i+1}) - \beta| \\
&\geq \frac{1}{2} + \frac{|\mathcal{C}|}{8}\,|(\alpha+\beta)\, g(x_i) - (\alpha+\beta)\, g(x_{i+1})| \\
&\geq \frac{1}{2} + \frac{|\mathcal{C}|}{8}\cdot\delta(\alpha+\beta) \\
&\geq \frac{1}{2} + \frac{\delta}{4}.
\end{aligned}$$

Case 2: $\mu_a > t_{|\mathcal{T}|-2}|\mathcal{C}|$:

Note $\alpha g(x) > \beta(1-g(x))$ only holds if $x = x_{|\mathcal{T}|-1}$. Therefore,

$$\begin{aligned}
p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C) &= \sum_{x\in\mathcal{M}^{|\mathcal{K}|}} |\mathcal{C}_x| \cdot \max\{\alpha g(x), \beta(1-g(x))\} \\
&= \alpha|\mathcal{C}_{x_{|\mathcal{T}|-1}}| + \sum_{x\in\mathcal{M}^{|\mathcal{K}|}\backslash\{x_{|\mathcal{T}|-1}\}} |\mathcal{C}_x| \cdot \beta(1-g(x))
\end{aligned}$$

$$
\begin{aligned}
&= \alpha|\mathcal{C}_{x_{|\mathcal{T}|-1}}| + \sum_{c\in\mathcal{C}} \sum_{k:D_k(c)=b} \beta p_{\mathrm{k}}(k) \\
&= p_{\mathrm{m}}(b) + \alpha|\mathcal{C}_{x_{|\mathcal{T}|-1}}| \\
&= \frac{1}{2} + \frac{|\mathcal{C}_{x_{|\mathcal{T}|-1}}|}{2\mu_a}.
\end{aligned}
$$

On one hand, $\sum_{k\in\mathcal{K}} \sum_{c\in\mathcal{C}} 1_{D_k(c)=a} = \sum_{k\in\mathcal{K}} \mu_a = |\mathcal{K}|\mu_a$. On the other hand,

$$
\begin{aligned}
\sum_{c\in\mathcal{C}} \sum_{k\in\mathcal{K}} 1_{D_k(c)=a} &= \sum_{c\in\mathcal{C}_{x_{|\mathcal{T}|-1}}} \sum_{k\in\mathcal{K}} 1_{D_k(c)=a} + \sum_{c\in\mathcal{C}\setminus\mathcal{C}_{x_{|\mathcal{T}|-1}}} \sum_{k\in\mathcal{K}} 1_{D_k(c)=a} \\
&\le |\mathcal{K}||\mathcal{C}_{x_{|\mathcal{T}|-1}}| + (|\mathcal{K}|-1)(|\mathcal{C}| - |\mathcal{C}_{x_{|\mathcal{T}|-1}}|) \\
&= (|\mathcal{K}|-1)|\mathcal{C}| + |\mathcal{C}_{x_{|\mathcal{T}|-1}}| \\
&< (|\mathcal{K}|-1)\frac{\mu_a}{t_{|\mathcal{T}|-2}} + |\mathcal{C}_{x_{|\mathcal{T}|-1}}| \\
&= \frac{\mu_a(|\mathcal{K}|-1)}{1-p_{\mathrm{k}}^{\min}} + |\mathcal{C}_{x_{|\mathcal{T}|-1}}|.
\end{aligned}
$$

Thus, $\frac{|\mathcal{C}_{x_{|\mathcal{T}|-1}}|}{\mu_a} > \frac{1-|\mathcal{K}|p_{\mathrm{k}}^{\min}}{1-p_{\mathrm{k}}^{\min}} > 0$. Let $\lambda \coloneqq \frac{1-|\mathcal{K}|p_{\mathrm{k}}^{\min}}{1-p_{\mathrm{k}}^{\min}}$, we have

$$
p_{\mathrm{guess}}^{\mathsf{EE},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) \ge \frac{1}{2} + \frac{\lambda}{2}.
$$

With the bounds in the two cases,

$$
\inf_{\mathsf{EE}\in\Omega_{\mathrm{EE}}} p_{\mathrm{guess}}^{\mathsf{EE},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) \ge \frac{1}{2} + \frac{1}{2}\min\{\frac{\delta}{2}, \lambda\} > \frac{1}{2}.
$$

Namely, the theorem holds.

K Explicit Advantage of EE Constructions and HE Schemes

The detailed calculations are provided below.

Example 2. For $p_{\mathrm{m}} = (\frac{1}{2}, \frac{1}{2})$ and $p_{\mathrm{k}} = (p_1, p_2, p_3)$ with $\frac{1}{2} > p_1 \ge p_2 \ge p_3$: $\inf_{\mathsf{EE}\in\Omega_{\mathrm{EE}}} p_{\mathrm{guess}}^{\mathsf{EE},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C)$ are

1. If $p_2 \ge \frac{1}{3}$, then

$$
\frac{1}{2} + \min\left\{ \frac{1-2p_1}{2}, \frac{(2p_1-1)(1-3p_1)}{2(1-p_1)p_1}, \right.
\left. \frac{1-3p_3}{4}, \frac{(1-2p_2)(p_2-p_3)}{2(1-p_2)p_2} \right\}.
$$

2. If $p_2 \le \frac{1}{3}$, then

$$\frac{1}{2} + \min\left\{\frac{1-2p_1}{2}, \frac{3p_1-1}{4}\right\}.$$

Then, $\inf_{\mathsf{EE}\in\Omega_{\mathrm{EE}}} p_{\mathrm{guess}}^{\mathsf{EE},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) \ge \frac{1}{2}$, where the equality holds if and only if $p_1 = p_2 = p_3 = \frac{1}{3}$.

Example 3. For $p_{\mathrm{m}} = (\frac{1}{2}, \frac{1}{2})$ and $p_{\mathrm{k}} = (p_1, p_2, p_3)$ with $\frac{1}{2} > p_1 \ge p_2 \ge p_3$:

$$\inf_{\mathsf{HE}\in\Omega_{\mathrm{HE}}} p_{\mathrm{guess}}^{\mathsf{HE},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) = \frac{1-2p_1}{2} + \frac{1}{2} > \frac{1}{2}.$$

The above analysis studies the optimal security of EE constructions with different ciphers, aiming to minimize decryption collisions. Previous work analyzed their *average* security in the random oracle (RO) or ideal cipher (IC) models, where the advantage is computed as an expectation over random ciphers.

We can naturally associate any discrete EE construction $\mathsf{EE}^{(\ell)}$ with a continuous one EE_{c} by defining $|S_{\boldsymbol{m}}| \coloneqq \prod_i p_{\mathrm{d}}(m_{j_i})$. Under this mapping, EE_{c} characterizes the expected collisions, and the expected security metrics (for both MC and KC) of $\mathsf{EE}^{(\ell)}$ converge to those of EE_{c}. Thus, analyzing EE constructions in the RO or IC models is equivalent to analyzing the corresponding continuous scheme.

Proposition 12. *For any EE construction $\mathsf{EE}^{(\ell)}$ in the random oracle or ideal cipher models and the associated continuous scheme EE_{c}, as $\ell \to \infty$,*

$$\mathbb{E}[I_{\mathsf{EE}^{(\ell)}}(K; C^{(\ell)})] \to I_{\mathsf{EE}_{\mathrm{c}}}(K; C),$$
$$\mathbb{E}[p_{\mathrm{guess}}^{\mathsf{EE}^{(\ell)},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C)] \to p_{\mathrm{guess}}^{\mathsf{EE}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C).$$

Example 4. For $p_{\mathrm{m}} = \left(\frac{1}{2}, \frac{1}{2}\right)$ and $p_{\mathrm{k}} = (p_1, p_2, p_3)$ with $\frac{1}{2} > p_1 \ge p_2 \ge p_3$, in the random-oracle or ideal-cipher model, the continuous HE scheme HE_{c} (with $p_{\mathrm{d}} = p_{\mathrm{m}}$) and EE constructions EE_{c} (with general p_{d}) satisfy the bounds:

$$p_{\mathrm{guess}}^{\mathsf{HE}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) = \frac{3}{4},$$
$$\inf_{\mathsf{EE}_{\mathrm{c}}\in\Omega_{\mathrm{EE}}} p_{\mathrm{guess}}^{\mathsf{EE}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) = \frac{1}{2} + p_1(1-p_1) \ge \frac{13}{18}.$$

L Details of Continuous Encryption Schemes and Proofs

Security of uniformized encryption schemes. Intuitively, uniformizing $p_{\mathrm{c|m,k}}$ over each equivalence class S of R_{a} induces a new ciphertext C' from the original ciphertext C via a post-processing step: first identify the equivalence class S to which C belongs, and then sample C' uniformly from S. As a result, C' reveals only the equivalence class of C. Since C' is obtained by post-processing C, it cannot reveal more information about the underlying variables than C.

To formalize this intuition, we consider a general measure $M(X;Y)$ that quantifies the amount of information about a sensitive variable X revealed by an observation Y. Formally, M is a functional that maps the joint distribution $P_{X,Y}$ to a real value:

$$M : \mathcal{P}(\mathcal{X} \times \mathcal{Y}) \to \mathbb{R}.$$

Depending on the context, M may represent an information-theoretic quantity (e.g., mutual information or divergence) or an operational quantity (e.g., optimal guessing success probability).

Definition 5 (Post-processing monotonicity). *A measure $M(X;Y)$ is said to be monotone under post-processing if for any Markov chain $X \to Y \to Z$, it holds that*

$$M(X;Z) \leq M(X;Y).$$

This property is a standard abstraction of the data processing inequality (DPI) in information theory [16], and is closely related to Blackwell's comparison of statistical experiments [4]. A wide class of measures satisfy this property, including f-divergences [17, 1], Shannon mutual information, Rényi divergence and Sibson mutual information [37, 40], as well as decision-theoretic quantities such as optimal guessing success probability. More recently, maximal leakage has also been shown to satisfy post-processing monotonicity [27]. In this paper, we focus on mutual information and guessing success probability as security measures, both of which satisfy post-processing monotonicity.

Proposition 13. *$p_{n\text{-guess}}(X \mid Y)$ and $I(X;Y)$ are monotone under post-processing.*

Theorem 22. *For any finite-ciphertext scheme E, if E' is obtained by uniformizing the conditional ciphertext distributions within each equivalence class of R_{a}, then E' is correct and*

$$I_{\mathsf{E}'}(K;C') \leq I_{\mathsf{E}}(K;C),$$
$$p_{\mathrm{guess}}^{\mathsf{E}',p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C') \leq p_{\mathrm{guess}}^{\mathsf{E},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C).$$

Proof. For a ciphertext C, write $S(C) = [C]_{R_{\mathrm{a}}}$ for its equivalence class. The uniformized ciphertext C' is generated by first computing $S(C)$ and then sampling uniformly from that class. Since all ciphertexts in the class have the same decryption output under every key, correctness is preserved. Moreover, $(M,K) \longrightarrow C \longrightarrow C'$ is a Markov chain. The two inequalities follow from post-processing monotonicity of mutual information and optimal guessing probability.

Theorem 8 (restated). *For any continuous scheme E_{c} and its discretization $\mathsf{E}^{(\ell)}$,*

$$I_{\mathsf{E}^{(\ell)}}(K;C^{(\ell)}) = I_{\mathsf{E}_{\mathrm{c}}}(K;C) + O(2^{-\ell}),$$
$$p_{\mathrm{guess}}^{\mathsf{E}^{(\ell)},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C^{(\ell)}) = p_{\mathrm{guess}}^{\mathsf{E}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) + O(2^{-\ell}).$$

Proof. For every (m,k), write $\mathcal{I}_{m,k}$ for the finite collection of maximal nonempty half-open intervals in $\mathcal{C}(m,k)$ on which $p_{\mathrm{c|m,k}}(\cdot\,|\,m,k)$ is positive and constant. These intervals come from the original design for that message–key pair; they are not refined using endpoints belonging only to other pairs. For $I=[a_I,b_I)\in\mathcal{I}_{m,k}$, define

$$\widetilde{I}^{(\ell)}=[Q_\ell(a_I),Q_\ell(b_I))\subseteq\mathbb{Z}_{2^\ell}$$

and its normalized counterpart

$$I^{(\ell)}=[2^{-\ell}Q_\ell(a_I),2^{-\ell}Q_\ell(b_I))\subseteq[0,1).$$

Write

$$t=\sum_{m,k}|\mathcal{I}_{m,k}|,\qquad \mu_{\min}=\min_{m,k,I\in\mathcal{I}_{m,k}}\mu(I),\qquad p^{\max}_{\mathrm{c|m,k}}=\max_{m,k,c}p_{\mathrm{c|m,k}}(c\,|\,m,k).$$

All three quantities are finite and $\mu_{\min}>0$. Since

$$\Big|\mu(I^{(\ell)})-\mu(I)\Big|<2^{-\ell},$$

every $\widetilde{I}^{(\ell)}$ is nonempty and $\mu(I^{(\ell)})\geq\mu(I)/2$ for all sufficiently large ℓ.

The discretization preserves the mass of each original maximal interval:

$$p^{(\ell)}_{\mathrm{c|m,k}}(c\,|\,m,k)=\frac{\displaystyle\int_I p_{\mathrm{c|m,k}}(u\,|\,m,k)\,\mathrm{d}u}{2^\ell\mu(I^{(\ell)})}\qquad\big(c\in\widetilde{I}^{(\ell)},\ I\in\mathcal{I}_{m,k}\big).\tag{5}$$

For each (m,k), the masses assigned in Eq. (5) sum to one. For a fixed key, the original intervals belonging to different messages are disjoint, and coherent endpoint quantization preserves this disjointness. Inheriting the decryption output from each original interval therefore gives a normalized and correct discrete scheme.

Continuizing $\mathsf{E}^{(\ell)}$ spreads the mass of each discrete ciphertext c uniformly over $[c/2^\ell,(c+1)/2^\ell)$. Conditioned on $C^{(\ell)}=c$, the resulting ciphertext $C^{(\ell)}_{\mathrm{c}}$ is uniform on this interval, while

$$C^{(\ell)}=\lfloor 2^\ell C^{(\ell)}_{\mathrm{c}}\rfloor.$$

Thus each ciphertext is obtained from the other by post-processing, and $\mathsf{E}^{(\ell)}$ and $\mathsf{E}^{(\ell)}_{\mathrm{c}}$ have the same KC and MRS metrics.

We next compare the conditional densities of $\mathsf{E}^{(\ell)}_{\mathrm{c}}$ and E_{c}. On $I^{(\ell)}$, the continuized density equals the original constant density on I multiplied by

$$\frac{\mu(I)}{\mu(I^{(\ell)})}.$$

This factor converges to one and is at most two for all sufficiently large ℓ. Define the affected boundary region by

$$\mathcal{B}_\ell=\bigcup_{m,k}\ \bigcup_{I\in\mathcal{I}_{m,k}}\big(I\triangle I^{(\ell)}\big),$$

where $\triangle$ denotes symmetric difference. Every endpoint moves by less than $2^{-\ell}$, so

$$\mu(\mathcal{B}_\ell) \le 2t\, 2^{-\ell}.$$

Outside $\mathcal{B}_\ell$, membership in every original maximal interval agrees with membership in its quantized image. On this region, the density scaling above differs from one by $O(2^{-\ell})$. On $\mathcal{B}_\ell$, the original and continuized densities are bounded by $p^{\max}_{\mathrm{c|m,k}}$ and $2p^{\max}_{\mathrm{c|m,k}}$, respectively. Consequently,

$$\max_{m,k} \left\| p^{\mathsf{E}_{\mathrm{c}}^{(\ell)}}_{\mathrm{c|m,k}} - p_{\mathrm{c|m,k}} \right\|_1 = O(2^{-\ell}).$$

The same argument gives pointwise convergence. Every fixed non-endpoint eventually has the same membership in all corresponding intervals. At an endpoint, the floor quantizer and the half-open convention assign the point to the right-hand interval in both schemes.

It remains to compare the security metrics. For MRS, the joint message–ciphertext densities are

$$p_{\mathrm{m,c}}(m,c) = p_{\mathrm{m}}(m) \sum_k p_{\mathrm{k}}(k) p_{\mathrm{c|m,k}}(c \,|\, m,k),$$

and analogously for $\mathsf{E}_{\mathrm{c}}^{(\ell)}$. Hence

$$\begin{aligned} &\left| p^{\mathsf{E}_{\mathrm{c}}^{(\ell)},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) - p^{\mathsf{E}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) \right| \\ &\qquad \le \sum_m \left\| p^{\mathsf{E}_{\mathrm{c}}^{(\ell)}}_{\mathrm{m,c}}(m,\cdot) - p^{\mathsf{E}_{\mathrm{c}}}_{\mathrm{m,c}}(m,\cdot) \right\|_1 = O(2^{-\ell}). \end{aligned}$$

For KC, both the original and continuized densities are piecewise constant on the finite partition induced by their interval endpoints. Outside $\mathcal{B}_\ell$, their positive density values differ by $O(2^{-\ell})$ and have the same zero pattern. The mutual-information integrand is therefore changed by $O(2^{-\ell})$ there. On $\mathcal{B}_\ell$, the integrand for either scheme can be written as $p_{\mathrm{c}}(c) D(p_{\mathrm{k|c}}(\cdot \,|\, c) \| p_{\mathrm{k}})$ and is bounded by

$$p_{\mathrm{c}}(c) \log \frac{1}{p^{\min}_{\mathrm{k}}},$$

and the ciphertext density is uniformly bounded. Since $\mu(\mathcal{B}_\ell) = O(2^{-\ell})$, we obtain

$$\left| I_{\mathsf{E}_{\mathrm{c}}^{(\ell)}}(K;C) - I_{\mathsf{E}_{\mathrm{c}}}(K;C) \right| = O(2^{-\ell}).$$

Finally, $\mathsf{E}^{(\ell)}$ and $\mathsf{E}_{\mathrm{c}}^{(\ell)}$ have identical security metrics, which gives the two claimed bounds.

Theorem 9 (restated). *Every asymptotic scheme* E_{a} *of* E *satisfies*

$$\begin{aligned} I_{\mathsf{E}^{(\ell)}}(K; C^{(\ell)}) = o(1) &\Longrightarrow I_{\mathsf{E}_{\mathrm{a}}}(K;C) = 0, \\ p^{\mathsf{E}^{(\ell)},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C^{(\ell)}) = p_{\max} + o(1) &\Longrightarrow p^{\mathsf{E}_{\mathrm{a}},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) = p_{\max}. \end{aligned}$$

Proof. Recall that $\mathfrak{M} = (\mathcal{M}')^{|\mathcal{K}|}$ is the set of all decryption-output vectors $\boldsymbol{m} = (m_k)_{k\in\mathcal{K}}$. For each ℓ, uniformizing $\mathsf{E}^{(\ell)}$ within every nonempty equivalence class $S_{\boldsymbol{m}}^{(\ell)}$ gives $\mathsf{U}^{(\ell)}$. Define the class masses by

$$\rho_{k,\boldsymbol{m}}^{(\ell)} = \begin{cases} \displaystyle\sum_{c\in S_{\boldsymbol{m}}^{(\ell)}} p_{\mathrm{c|m,k}}^{(\ell)}(c\,|\,m_k,k), & m_k \in \mathcal{M}, \\ 0, & m_k = \perp . \end{cases}$$

Fix any asymptotic scheme E_{a}. By definition, it is obtained from a subsequence $(\ell_t)_t$ along which the vector of class masses converges coordinatewise, say $\rho_{k,\boldsymbol{m}}^{(\ell_t)} \to \rho_{k,\boldsymbol{m}}$. Such subsequences exist because these vectors lie in the fixed finite cube $[0,1]^{|\mathcal{K}|\cdot|\mathfrak{M}|}$. For every $k \in \mathcal{K}$ and $m \in \mathcal{M}$, correctness and normalization give

$$\sum_{\boldsymbol{m}:m_k=m} \rho_{k,\boldsymbol{m}}^{(\ell)} = 1, \qquad \rho_{k,\boldsymbol{m}}^{(\ell)} = 0 \quad \text{if } m_k = \perp .$$

Passing to the limit shows that the same identities hold for $\rho_{k,\boldsymbol{m}}$. On $I_{\boldsymbol{m}}$, define the decryption rule of E_{a} by $D_k(c) = m_k$ and its conditional density by

$$p_{\mathrm{c|m,k}}^{\mathsf{E}_{\mathrm{a}}}(c\,|\,m,k) = \begin{cases} |\mathfrak{M}|\rho_{k,\boldsymbol{m}}, & m = m_k, \\ 0, & m \neq m_k. \end{cases}$$

The preceding identities ensure that E_{a} is normalized and correct. Moreover, for $c \in I_{\boldsymbol{m}}$,

$$p_{\mathrm{c|m,k}}^{\mathsf{E}_{\mathrm{r}}^{(\ell_t)}}(c\,|\,m,k) = \begin{cases} |\mathfrak{M}|\rho_{k,\boldsymbol{m}}^{(\ell_t)}, & m = m_k, \\ 0, & m \neq m_k. \end{cases}$$

Because the half-open intervals $I_{\boldsymbol{m}}$ are fixed, the coordinatewise convergence of the class masses gives pointwise convergence of every conditional density of $\mathsf{E}_{\mathrm{r}}^{(\ell_t)}$ to the corresponding density of E_{a}.

Let $S^{(\ell)} = \boldsymbol{m}$ when $C^{(\ell)} \in S_{\boldsymbol{m}}^{(\ell)}$; thus $S^{(\ell)}$ is the equivalence-class label on the fixed alphabet $\mathfrak{M}$. Conditional on $S^{(\ell)} = \boldsymbol{m}$, the ciphertext of $\mathsf{U}^{(\ell)}$ is uniform on $S_{\boldsymbol{m}}^{(\ell)}$ independently of (M,K). Since $S^{(\ell)}$ is also a function of the ciphertext, it is a sufficient statistic for both security metrics:

$$\begin{aligned} I_{\mathsf{U}^{(\ell)}}(K;C^{(\ell)}) &= I(K;S^{(\ell)}), \\ p_{\mathrm{guess}}^{\mathsf{U}^{(\ell)},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C^{(\ell)}) &= p_{\mathrm{guess}}(M \mid S^{(\ell)}). \end{aligned} \tag{6}$$

More explicitly,

$$\Pr[K = k, S^{(\ell)} = \boldsymbol{m}] = \begin{cases} p_{\mathrm{k}}(k)p_{\mathrm{m}}(m_k)\rho_{k,\boldsymbol{m}}^{(\ell)}, & m_k \in \mathcal{M}, \\ 0, & m_k = \perp, \end{cases} \tag{7}$$

and

$$p_{\mathrm{guess}}(M \mid S^{(\ell)}) = \sum_{\boldsymbol{m}\in\mathfrak{M}} \max_{m\in\mathcal{M}} \left\{ p_{\mathrm{m}}(m) \sum_{k:m_k=m} p_{\mathrm{k}}(k)\rho_{k,\boldsymbol{m}}^{(\ell)} \right\}. \tag{8}$$

The finite joint law in Eq. (7), and hence its mutual information, is continuous in the class masses. The finite maximum-and-sum expression in Eq. (8) is continuous as well.

The interval containing a ciphertext of $\mathsf{E}_{\mathrm{r}}^{(\ell)}$ identifies its label $\boldsymbol{m}$. By construction, the joint law of this label and (M, K) is the same as that of $S^{(\ell)}$ and (M, K), while the position within the interval contains no further information. Thus $\mathsf{E}_{\mathrm{r}}^{(\ell)}$ has the same KC and MRS metrics as $\mathsf{U}^{(\ell)}$. The same holds for E_{a} with the limiting class masses. The preceding continuity therefore gives

$$\begin{aligned} I_{\mathsf{E}_{\mathrm{r}}^{(\ell_t)}}(K;C) &\longrightarrow I_{\mathsf{E}_{\mathrm{a}}}(K;C), \\ p_{\mathrm{guess}}^{\mathsf{E}_{\mathrm{r}}^{(\ell_t)},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) &\longrightarrow p_{\mathrm{guess}}^{\mathsf{E}_{\mathrm{a}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C). \end{aligned} \tag{9}$$

By Theorem 22,

$$0 \le I_{\mathsf{U}^{(\ell)}}(K;C^{(\ell)}) \le I_{\mathsf{E}^{(\ell)}}(K;C^{(\ell)}),$$

while correctness gives

$$p_{\max} \le p_{\mathrm{guess}}^{\mathsf{U}^{(\ell)},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C^{(\ell)}) \le p_{\mathrm{guess}}^{\mathsf{E}^{(\ell)},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C^{(\ell)}).$$

Indeed, an adversary can always guess a most likely message, with success $p_{\mathrm{m}}^{\max}$, or decrypt under a most likely key, with success probability at least $p_{\mathrm{k}}^{\max}$. The two conclusions now follow from these squeeze inequalities and Eq. (9). Since E_{a} was arbitrary, they hold for every asymptotic scheme.

Theorem 2′ (restated). *A* uniform *continuous scheme* E_{c} *achieves* $I(K;C) = 0$, *if and only if* $\forall m \in \mathcal{M}, k \in \mathcal{K}$, $\Pr[D_k(C) = m \,|\, C \leftarrow_{p_{\mathrm{c}}} \mathcal{C}] = p_{\mathrm{m}}(m)$.

Proof. The weighted stability proof of Theorem 2 also applies to uniform finite-region continuous schemes: replace overlap cardinalities by Lebesgue measures and ciphertext sums by integrals. The same finite number of density cuts gives the same coefficient, without a minimum positive overlap measure. Setting the message error to zero gives ciphertext indistinguishability and hence $I(K;C) = 0$. The converse follows from correctness and data processing.

Theorem 10 (restated). *A continuous encryption scheme* E_{c} *achieves*

$$p_{\mathrm{guess}}^{\mathsf{E}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) = p_{\max}$$

if and only if, for every choice of maximizers $m^\star$ *and* $k^\star$*:*

1. *If* $p_{\mathrm{k}}^{\max} \ge p_{\mathrm{m}}^{\max}$*: 1)* $D_{k^\star}(c)$ *maximizes* $p_{\mathrm{m,c}}(\cdot, c)$ *for every* c*, and 2)* $\mathcal{C}(m,k^\star) \cap \mathcal{C}(m,k') = \emptyset$ *for every* m *and every* $k' \neq k^\star$*.*
2. *If* $p_{\mathrm{m}}^{\max} \ge p_{\mathrm{k}}^{\max}$*:* $m^\star$ *maximizes* $p_{\mathrm{m,c}}(\cdot, c)$ *for every* c*.*

Proof (Proof of Theorem 10). Let

$$p_{\mathrm{m,c}}^{\max}(c) \coloneqq \max_{m \in \mathcal{M}} p_{\mathrm{m,c}}(m,c).$$

Pointwise optimal guessing in the continuous setting gives

$$p_{\text{guess}}(M \mid C) = \int_{\mathcal{C}} p^{\max}_{\text{m,c}}(c)\,\mathrm{d}c.$$

In the finite-interval model of Section 6, a nonnegative piecewise constant function has zero integral exactly when it vanishes for every c; likewise, a collision integral below is zero exactly when the corresponding ciphertext-region intersection is empty.

First suppose $p^{\max}_{\text{k}} \geq p^{\max}_{\text{m}}$ and fix a most likely key $k^\star$. With the convention $p_{\text{m,c}}(\perp, c) = 0$, the strategy that decrypts with $k^\star$ and correctness give

$$\begin{aligned}
p_{\text{guess}}(M \mid C) &= \int_{\mathcal{C}} p^{\max}_{\text{m,c}}(c)\,\mathrm{d}c \\
&\geq \int_{\mathcal{C}} p_{\text{m,c}}(D_{k^\star}(c), c)\,\mathrm{d}c \\
&= p_{\text{k}}(k^\star) + \sum_{k' \neq k^\star} \sum_{m \in \mathcal{M}} p_{\text{m}}(m) p_{\text{k}}(k') \\
&\qquad \times \int_{\mathcal{C}(m,k^\star) \cap \mathcal{C}(m,k')} p_{\text{c|m,k}}(c \,|\, m, k')\,\mathrm{d}c \\
&\geq p^{\max}_{\text{k}} = p_{\max}.
\end{aligned}$$

Equality holds exactly when (i) $D_{k^\star}(c)$ maximizes $p_{\text{m,c}}(\cdot, c)$ for every c (which also excludes $D_{k^\star}(c) = \perp$, since $p_{\text{c}}(c) > 0$), and (ii)

$$\mathcal{C}(m, k^\star) \cap \mathcal{C}(m, k') = \emptyset \qquad \text{for every } m \text{ and every } k' \neq k^\star.$$

Now suppose $p^{\max}_{\text{m}} \geq p^{\max}_{\text{k}}$ and fix a most likely message $m^\star$. The strategy that always outputs $m^\star$ gives

$$p_{\text{guess}}(M \mid C) = \int_{\mathcal{C}} p^{\max}_{\text{m,c}}(c)\,\mathrm{d}c \geq \int_{\mathcal{C}} p_{\text{m,c}}(m^\star, c)\,\mathrm{d}c = p^{\max}_{\text{m}} = p_{\max}.$$

Equality holds exactly when $m^\star$ maximizes $p_{\text{m,c}}(\cdot, c)$ for every c. If $p^{\max}_{\text{k}} = p^{\max}_{\text{m}}$, both arguments apply, so both stated conditions are required and together are sufficient.

Corollary 2 (restated). *For a continuous encryption scheme* E_{c} *satisfying*

$$p^{\mathsf{E}_{\text{c}}, p_{\text{m}}, p_{\text{k}}}_{\text{guess}}(M \mid C) = p_{\max},$$

the following conditions hold for every choice of maximizers $m^\star$ *and* $k^\star$*:*

1. *If* $p^{\max}_{\text{k}} \geq p^{\max}_{\text{m}}$*: 1)* $\{\mathcal{C}(m, k^\star)\}_m$ *is a partition of* $\mathcal{C}$*, and 2)* $\mathcal{C}(m, k^\star) \cap \mathcal{C}(m, k') = \emptyset$ *for every* m *and every* $k' \neq k^\star$*.*
2. *If* $p^{\max}_{\text{m}} \geq p^{\max}_{\text{k}}$*:* $\bigcup_k \mathcal{C}(m^\star, k) = \mathcal{C}$*.*

Proof (Proof of Corollary 2). If $p_{\mathrm{k}}^{\max} \geq p_{\mathrm{m}}^{\max}$, Theorem 10 gives $D_{k^\star}(c) \in \mathcal{M}$ for every c. The deterministic decryption regions $\{\mathcal{C}(m, k^\star)\}_m$ therefore partition $\mathcal{C}$, and the second claim is the theorem's no-collision condition.

If $p_{\mathrm{m}}^{\max} \geq p_{\mathrm{k}}^{\max}$ and some $c \notin \bigcup_k \mathcal{C}(m^\star, k)$, then $p_{\mathrm{m,c}}(m^\star, c) = 0$. But $p_{\mathrm{c}}(c) > 0$ implies $p_{\mathrm{m,c}}(m, c) > 0$ for some m, contradicting the theorem's pointwise maximality of $m^\star$. Hence $\bigcup_k \mathcal{C}(m^\star, k) = \mathcal{C}$. In the equality case both arguments apply.

Lemma 3. *Fix a continuous encryption scheme E_{c} over finite $\mathcal{M}$ and $\mathcal{K}$. For any message distributions $p_{\mathrm{m}}, p'_{\mathrm{m}}$ and key distributions $p_{\mathrm{k}}, p'_{\mathrm{k}}$,*

$$\left| p_{\mathrm{guess}}^{\mathsf{E}_{\mathrm{c}}, p_{\mathrm{m}}, p_{\mathrm{k}}}(M \mid C) - p_{\mathrm{guess}}^{\mathsf{E}_{\mathrm{c}}, p'_{\mathrm{m}}, p'_{\mathrm{k}}}(M \mid C) \right| \leq \mathrm{SD}(p_{\mathrm{m}}, p'_{\mathrm{m}}) + \mathrm{SD}(p_{\mathrm{k}}, p'_{\mathrm{k}}).$$

The encryption and decryption algorithms are identical in the two experiments.

Proof. For any fixed adversary, its success probability conditional on (m, k) is a function with values in $[0, 1]$, determined by the fixed encryption channel. Its expectations under the two input distributions differ by at most their statistical distance. This uniform bound remains valid after maximizing over adversaries, so the difference is at most $\mathrm{SD}(p_{\mathrm{m}} \otimes p_{\mathrm{k}}, p'_{\mathrm{m}} \otimes p'_{\mathrm{k}})$. The product-distribution triangle inequality gives

$$\mathrm{SD}(p_{\mathrm{m}} \otimes p_{\mathrm{k}}, p'_{\mathrm{m}} \otimes p'_{\mathrm{k}}) \leq \mathrm{SD}(p_{\mathrm{m}}, p'_{\mathrm{m}}) + \mathrm{SD}(p_{\mathrm{k}}, p'_{\mathrm{k}}).$$

The argument also permits zero-probability source elements.

Lemma 4. *Fix a continuous encryption scheme E_{c} over finite $\mathcal{M}$ and $\mathcal{K}$, a message distribution p_{m} on $\mathcal{M}$, and a key distribution p_{k} on $\mathcal{K}$. Suppose that $p_{\mathrm{k}}(k_0) = 0$ for some $k_0 \in \mathcal{K}$. Define*

$$\mathcal{K}' \coloneqq \mathcal{K} \setminus \{k_0\}, \qquad p'_{\mathrm{k}}(k) \coloneqq p_{\mathrm{k}}(k) \quad \text{for every } k \in \mathcal{K}',$$

and let E'_{c} be the encryption scheme obtained by restricting E_{c} to the key space $\mathcal{K}'$. Then

$$p_{\mathrm{guess}}^{\mathsf{E}_{\mathrm{c}}, p_{\mathrm{m}}, p_{\mathrm{k}}}(M \mid C) = p_{\mathrm{guess}}^{\mathsf{E}'_{\mathrm{c}}, p_{\mathrm{m}}, p'_{\mathrm{k}}}(M \mid C).$$

Proof. Because $p_{\mathrm{k}}(k_0) = 0$, p'_{k} is a probability distribution on $\mathcal{K}'$. Moreover, for every m and c,

$$\begin{aligned} p_{\mathrm{m,c}}^{\mathsf{E}_{\mathrm{c}}, p_{\mathrm{m}}, p_{\mathrm{k}}}(m, c) &= \sum_{k \in \mathcal{K}} p_{\mathrm{m}}(m) p_{\mathrm{k}}(k) p_{\mathrm{c|m,k}}(c \,|\, m, k) \\ &= \sum_{k \in \mathcal{K}'} p_{\mathrm{m}}(m) p'_{\mathrm{k}}(k) p_{\mathrm{c|m,k}}(c \,|\, m, k) \\ &= p_{\mathrm{m,c}}^{\mathsf{E}'_{\mathrm{c}}, p_{\mathrm{m}}, p'_{\mathrm{k}}}(m, c). \end{aligned}$$

Taking the maximum over m and integrating over c proves the claim.

Theorem 11 (restated). *For a message distribution p_{m}, there exists a p_{k}-agnostic continuous encryption scheme E_{c} achieving $p_{\mathrm{guess}}^{\mathsf{E}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) = p_{\max}$ for every key distribution p_{k} (under a given $|\mathcal{K}|$) if and only if one of the following three conditions holds:*

1. $|\mathcal{M}| = |\mathcal{K}| = 2$ *(or* $|\mathcal{M}| = 1$*).*
2. $|\mathcal{M}| \geq 3$, $|\mathcal{K}| = 2$, *and* $p_{\mathrm{m}}(m^\star) \geq 0.5$.
3. $p_{\mathrm{m}}^{\max} \leq \frac{1}{|\mathcal{K}|}$.

Proof (Proof of Theorem 11).

If $|\mathcal{M}| = 1$, every correct scheme satisfies $p_{\mathrm{guess}}(M \mid C) = p_{\max} = 1$ for every key distribution.

If $\mathcal{M} = \mathcal{K} = \{0,1\}$, then $E_k(m) = m \oplus k$ with $\mathcal{C} = \{0,1\}$ achieves the optimal MRS. This scheme can be naturally extended to $[0,1)$, yielding E_2 (as shown in Fig. 8a).

If $|\mathcal{M}| \geq 3$, $|\mathcal{K}| = 2$, and $p_{\mathrm{m}}(m^\star) \geq 0.5$, we use the scheme E_2 for two message classes and assign every $m \neq m^\star$ within its second interval according to the conditional probabilities (as shown in Fig. 8b). In other words, $\mathcal{C}(m^\star, k) = \bigcup_{m \neq m^\star} \mathcal{C}(m, k')$ for each $k \neq k'$.

If $|\mathcal{M}| \geq 3$, $|\mathcal{K}| = 2$, and $p_{\mathrm{m}}(m^\star) < 0.5$, then $p_{\mathrm{m}}(m^\star) < \frac{1}{|\mathcal{K}|}$, which can be reduced to the next case.

If $p_{\mathrm{m}}(m^\star) \leq \frac{1}{|\mathcal{K}|}$, we can use an EE construction with the following interval structure as the general scheme (as shown in Fig. 8c):

$$\mathcal{C}(m_i, k_j) = \left[F_{\mathrm{m}}(m_i^-) + (j-1)p_{\mathrm{m}}^{\max}, F_{\mathrm{m}}(m_i) + (j-1)p_{\mathrm{m}}^{\max}\right).$$

Since $|\mathcal{K}| p_{\mathrm{m}}(m^\star) \leq 1$, the scheme is well-defined and avoids collisions.

It remains to prove necessity when $|\mathcal{M}| \geq 2$. The only remaining case is $|\mathcal{K}| \geq 3$ and $p_{\mathrm{m}}(m^\star) > \frac{1}{|\mathcal{K}|}$. In this case, general schemes do not exist. Otherwise, we derive contradictions by exploiting the arbitrariness of p_{k} and applying Corollary 2.

Fix $k' \in \mathcal{K}$. Choose a full-support key distribution p_{k} such that k' is its unique most likely key and $p_{\mathrm{k}}(k') > p_{\mathrm{m}}(m^\star)$. Since E_{c} is optimal for this p_{k}, Corollary 2 implies that 1) $\{\mathcal{C}(m, k')\}_m$ is a partition of $\mathcal{C}$, and 2) $\mathcal{C}(m, k') \cap \mathcal{C}(m, k'') = \emptyset$ for each $k'' \neq k'$ and each m. Since k' is arbitrary, every c corresponds to different messages under different keys. We next analyze two cases: 1) $p_{\mathrm{m}}(m^\star) > \frac{1}{|\mathcal{K}|-1}$, and 2) $\frac{1}{|\mathcal{K}|} < p_{\mathrm{m}}(m^\star) \leq \frac{1}{|\mathcal{K}|-1}$.

Case 1: $p_{\mathrm{m}}(m^\star) > \frac{1}{|\mathcal{K}|-1}$:

Under the uniform key distribution, since $p_{\mathrm{m}}(m^\star) > \frac{1}{|\mathcal{K}|} = p_{\mathrm{k}}^{\max}$, $\{\mathcal{C}(m^\star, k)\}_k$ covers $\mathcal{C}$. Furthermore, by the above analysis, $\{\mathcal{C}(m^\star, k)\}_k$ forms a partition of $\mathcal{C}$. Now let p_{k} be uniform on $\mathcal{K} \setminus \{k'\}$ and satisfy $p_{\mathrm{k}}(k') = 0$. By Lemma 3 and the continuity of $p_{\max}$, E_{c} remains optimal under this boundary distribution. Lemma 4 then shows that the scheme restricted to $\mathcal{K} \setminus \{k'\}$ remains optimal under the corresponding restricted key distribution. Because $p_{\mathrm{m}}(m^\star) > 1/(|\mathcal{K}| - 1) = p_{\mathrm{k}}^{\max}$ for the reduced key space, $\{\mathcal{C}(m^\star, k)\}_{k \neq k'}$ must still cover $\mathcal{C}$. Since the full

family $\{\mathcal{C}(m^\star, k)\}_k$ is a partition and $\mathcal{C}(m^\star, k')$ has positive measure, the reduced family $\{\mathcal{C}(m^\star, k)\}_{k \neq k'}$ cannot cover $\mathcal{C}$, contradicting Corollary 2. Thus, general schemes do not exist.

Case 2: $\frac{1}{|\mathcal{K}|} < p_{\mathrm{m}}(m^\star) \leq \frac{1}{|\mathcal{K}|-1}$: Fix $k_i \neq k_j$ and let $p_{\mathrm{k}}(k_i) = p_{\mathrm{k}}(k_j) = 1/2$, with all other coordinates zero. By Lemma 3 and the continuity of $p_{\max}$, E_{c} is optimal under this boundary distribution. Applying Lemma 4 successively to every zero-probability key then shows that the scheme restricted to $\{k_i, k_j\}$ is optimal under the corresponding restricted distribution. Since $p_{\mathrm{m}}(m^\star) \leq \frac{1}{|\mathcal{K}|-1} \leq \frac{1}{2} = p_{\mathrm{k}}^{\max}$, Theorem 10 gives $p_{\mathrm{m,c}}(D_{k_i}(c), c) = p_{\mathrm{m,c}}(D_{k_j}(c), c)$ for every c.

Given that each c corresponds to a different m under different k, it holds that $p_{\mathrm{m,c}}(D_k(c), c) = p_{\mathrm{k}}(k) p_{\mathrm{m}}(D_k(c)) p_{\mathrm{c|m,k}}(c \,|\, D_k(c), k)$. Thus, due to the arbitrariness of k_i and k_j, the quantity

$$g(c) \coloneqq p_{\mathrm{m}}(D_k(c)) p_{\mathrm{c|m,k}}(c \,|\, D_k(c), k)$$

is independent of k. After taking a common refinement of the equivalence classes and the constant-density intervals, g is piecewise constant. For every key k,

$$\int_0^1 g(c)\,\mathrm{d}c = \sum_m p_{\mathrm{m}}(m) \int_{c \in \mathcal{C}(m,k)} p_{\mathrm{c|m,k}}(c \,|\, m, k)\mathrm{d}c = 1,$$

so we may reparameterize the ciphertext space by

$$T(c) \coloneqq \int_0^c g(u)\,\mathrm{d}u.$$

After omitting zero-mass pieces, T is an increasing piecewise-linear bijection of $[0, 1)$. It maps every equivalence class S to $T(S)$; on a piece where $g(c) = \gamma$, its length is scaled by the factor γ. Define the transformed scheme by

$$D'_k(T(c)) = D_k(c), \qquad p'_{\mathrm{c|m,k}}(T(c) \,|\, m, k) = \frac{p_{\mathrm{c|m,k}}(c \,|\, m, k)}{g(c)}.$$

This bijective change of ciphertext coordinates preserves correctness and both security metrics. Moreover,

$$p_{\mathrm{m}}(D'_k(T(c))) p'_{\mathrm{c|m,k}}(T(c) \,|\, D'_k(T(c)), k) = 1.$$

Relabeling the transformed ciphertext coordinate as c, we may therefore assume without loss of generality that $p_{\mathrm{m}}(D_k(c)) p_{\mathrm{c|m,k}}(c \,|\, D_k(c), k) = 1$ for every k and c.

Now, let $p_{\mathrm{k}}(k) = \frac{1}{|\mathcal{K}|}$ for each k. Since $p_{\mathrm{m}}(m^\star) > \frac{1}{|\mathcal{K}|} = p_{\mathrm{k}}^{\max}$, Corollary 2 and the previously established disjointness property imply that the family $\{\mathcal{C}(m^\star, k)\}_k$ forms a partition of $\mathcal{C}$.

Hence, for every c, there exists a unique k such that $D_k(c) = m^\star$. Denote this k as $k(c, m^\star)$. Then,

$$p_{\mathrm{m}}(m^\star) p_{\mathrm{c|m,k}}(c \,|\, m^\star, k(c, m^\star)) = 1,$$

which gives

$$p_{\mathrm{c|m,k}}(c\,|\,m^\star, k(c,m^\star)) = \frac{1}{p_{\mathrm{m}}(m^\star)} = \frac{1}{|\mathcal{C}(m^\star, k(c,m^\star))|}.$$

Thus, $|\mathcal{C}(m^\star, k(c,m^\star))| \cdot |\mathcal{K}| > 1$, contradicting that $\{\mathcal{C}(m^\star,k)\}_k$ is a partition of $\mathcal{C}$.

Proposition 4 (restated). *For $p_{\mathrm{m}} = (1/2, 1/2)$ and $p_{\mathrm{k}} = (1/3, 1/3, 1/3)$, every encryption scheme satisfies*

$$p_{\mathrm{guess}}(M \mid C) \geq \max\left\{\frac{1}{2},\ \frac{2}{3} - \sqrt{\frac{\ln 2}{2} I(K;C)}\right\}.$$

In particular, perfect KC forces $p_{\mathrm{guess}}(M \mid C) \geq 2/3$, whereas $p_{\mathrm{guess}}(M \mid C) = p_{\max} = 1/2$ requires $I(K;C) \geq 1/(18 \ln 2) \approx 0.0801$ bits. The latter lower bound is not asserted to be tight.

Proof (Proof of Proposition 4). For each supported ciphertext c, correctness implies that $M = D_K(c)$ under the posterior law of K. A key with $D_k(c) = \perp$ has posterior probability zero. Complete the decryption outputs of any such keys arbitrarily to one of the two messages, without changing the induced posterior message law. Under the uniform *prior* on three keys, one of the two completed message classes has mass at least $2/3$. Changing the key law from the prior to $p_{\mathrm{k|c}}(\cdot \mid c)$ changes the mass of this class by at most the statistical distance. Hence

$$\max_m p_{\mathrm{m|c}}(m \mid c) \geq \frac{2}{3} - \mathrm{SD}(p_{\mathrm{k|c}}(\cdot \mid c), p_{\mathrm{k}}).$$

Average over C and apply Pinsker's inequality and Jensen's inequality:

$$\begin{aligned} p_{\mathrm{guess}}(M \mid C) &\geq \frac{2}{3} - \mathop{\mathbb{E}}_{C} \sqrt{\frac{\ln 2}{2} \mathrm{KL}(p_{\mathrm{k|c}}(\cdot \mid C) \| p_{\mathrm{k}})} \\ &\geq \frac{2}{3} - \sqrt{\frac{\ln 2}{2} I(K;C)}. \end{aligned}$$

Correctness also gives $p_{\mathrm{guess}} \geq p_{\max} = 1/2$. The two special cases follow by substitution.

Lemma 5. *For $f_a(x) = \max\{\frac{a}{x}, \frac{1-a}{1-x}\} + \max\{\frac{a}{1-x}, \frac{1-a}{x}\}$, it always holds*

$$f_{a_1}(x) \geq f_{a_2}(x)$$

for $0 < a_1 < a_2 < 1/2$ and $0 < x < 1$.

Example 2 (restated). *For $p_{\mathrm{m}} = (\frac{1}{2}, \frac{1}{2})$ and $p_{\mathrm{k}} = (p_1, p_2, p_3)$ with $\frac{1}{2} > p_1 \geq p_2 \geq p_3$:*

1. *If* $p_2 \geq \frac{1}{3}$,

$$\inf_{\mathsf{EE}\in\Omega_{EE}} p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C) = \frac{1}{2} + \min\left\{ \frac{1-2p_1}{2}, \frac{(2p_1-1)(1-3p_1)}{2(1-p_1)p_1}, \frac{1-3p_3}{4}, \frac{(1-2p_2)(p_2-p_3)}{2(1-p_2)p_2} \right\}.$$

2. *If* $p_2 \leq \frac{1}{3}$,

$$\inf_{\mathsf{EE}\in\Omega_{EE}} p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C) = \frac{1}{2} + \min\left\{ \frac{1-2p_1}{2}, \frac{3p_1-1}{4} \right\}.$$

Then, $\inf_{\mathsf{EE}\in\Omega_{EE}} p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C) \geq \frac{1}{2}$*, where the equality holds if and only if* $p_1 = p_2 = p_3 = \frac{1}{3}$.

Derivation of Example 2. In an EE construction, encryption is uniform over $\mathcal{C}(m,k)$ for every message m and key k, and the size of $\mathcal{C}(m,k)$ is independent of k. Consequently, the construction is characterized by the sizes of the equivalence classes

$$S_{(m_{i_1},m_{i_2},m_{i_3})}$$

subject to the requirement that their marginals agree with the prescribed region sizes $|\mathcal{C}(m_i,k)|$ for every key.

For notational convenience, write

$$s_{i_1i_2i_3} \coloneqq \left|S_{(m_{i_1},m_{i_2},m_{i_3})}\right|, \qquad \alpha_i \coloneqq \left|\mathcal{C}(m_i,k)\right|.$$

These variables satisfy

$$\begin{aligned} s_{111}+s_{112}+s_{121}+s_{122} &= \alpha_1, \\ s_{111}+s_{112}+s_{211}+s_{212} &= \alpha_1, \\ s_{111}+s_{121}+s_{211}+s_{221} &= \alpha_1, \\ \sum_{i_1,i_2,i_3=1}^{2} s_{i_1i_2i_3} &= 1, \\ \alpha_1+\alpha_2 &= 1, \\ s_{i_1i_2i_3} \geq 0 \quad \forall\, i_1,i_2,i_3 &\in \{1,2\}. \end{aligned}$$

First, observe that setting

$$s_{111} = s_{222} = 0$$

cannot increase the adversary's advantage. Indeed, any positive mass on s_{111} or s_{222} contributes an advantage of 1, and one can redistribute this mass to other classes to lower the overall advantage.

Hence, without loss of generality we set $s_{111} = s_{222} = 0$ and treat α_1, s_{112}, and s_{121} as free parameters. The remaining variables admit the closed-form expressions

$$\begin{aligned} s_{122} &= \alpha_1 - s_{112} - s_{121}, \\ s_{211} &= -1 + 3\,\alpha_1 - s_{112} - s_{121}, \\ s_{212} &= 1 - 2\,\alpha_1 + s_{121}, \\ s_{221} &= 1 - 2\,\alpha_1 + s_{112}, \\ \alpha_2 &= 1 - \alpha_1, \end{aligned}$$

subject to

$$\frac{1}{3} \;\le\; \alpha_1 \;\le\; \frac{2}{3}, \qquad L \;\le\; s_{112} \;\le\; U, \qquad L \;\le\; s_{121} \;\le\; U,$$

where

$$L \coloneqq \max\{0,\; 2\alpha_1 - 1\}, \qquad U \coloneqq \min\{\alpha_1,\; 3\alpha_1 - 1\}.$$

For any $c \in S_{i_1 i_2 i_3}$,

$$\max_m \Pr[M = m,\, C = c] = \frac{1}{2}\;\max\left\{\frac{\sum_{j:i_j=1} p_j}{\alpha_1},\; \frac{\sum_{j:i_j=2} p_j}{\alpha_2}\right\}.$$

Therefore the overall advantage is

$$p^{\mathsf{EE},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) = \sum_{i_1,i_2,i_3} \frac{s_{i_1 i_2 i_3}}{2}\;\max\left\{\frac{\sum_{j:i_j=1} p_j}{\alpha_1},\; \frac{\sum_{j:i_j=2} p_j}{\alpha_2}\right\}.$$

Fix α_1, the feasible region for $(s_{i_1 i_2 i_3})_{i_1 i_2 i_3}$ is a triangle, and $p^{\mathsf{EE},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C)$ is affine in these two variables as

$$\frac{1}{2}\,s_{121}\,A_1 \;+\; \frac{1}{2}\,s_{112}\,A_2 \;+\; \frac{1}{2}\,A_3,$$

where

$$\begin{aligned} A_1 \coloneqq\; & \max\left\{\frac{p_2}{\alpha_1},\, \frac{1-p_2}{1-\alpha_1}\right\} + \max\left\{\frac{p_2}{1-\alpha_1},\, \frac{1-p_2}{\alpha_1}\right\} \\ & - \max\left\{\frac{p_1}{\alpha_1},\, \frac{1-p_1}{1-\alpha_1}\right\} - \max\left\{\frac{p_1}{1-\alpha_1},\, \frac{1-p_1}{\alpha_1}\right\}, \\ A_2 \coloneqq\; & \max\left\{\frac{p_3}{\alpha_1},\, \frac{1-p_3}{1-\alpha_1}\right\} + \max\left\{\frac{1-p_3}{\alpha_1},\, \frac{p_3}{1-\alpha_1}\right\} \\ & - \max\left\{\frac{p_1}{\alpha_1},\, \frac{1-p_1}{1-\alpha_1}\right\} - \max\left\{\frac{p_1}{1-\alpha_1},\, \frac{1-p_1}{\alpha_1}\right\}, \\ A_3 \coloneqq\; & \alpha_1\,\max\left\{\frac{p_1}{\alpha_1},\, \frac{1-p_1}{1-\alpha_1}\right\} + \max\left\{\frac{p_2}{\alpha_1},\, \frac{1-p_2}{1-\alpha_1}\right\} + \max\left\{\frac{p_3}{\alpha_1},\, \frac{1-p_3}{1-\alpha_1}\right\} \end{aligned}$$

$$+3\,\alpha_1\,\max\left\{\frac{p_1}{1-\alpha_1},\,\frac{1-p_1}{\alpha_1}\right\}-2\,\alpha_1\,\max\left\{\frac{p_2}{\alpha_1},\,\frac{1-p_2}{1-\alpha_1}\right\}$$
$$-2\,\alpha_1\,\max\left\{\frac{p_3}{\alpha_1},\,\frac{1-p_3}{1-\alpha_1}\right\}-\max\left\{\frac{p_1}{1-\alpha_1},\,\frac{1-p_1}{\alpha_1}\right\}.$$

As $1 > p_1 \geq p_2 \geq p_3 \geq 0$, by Lemma 5, $A_1 \geq 0$ and $A_2 \geq 0$. Therefore, $p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C)$ achieves its minimum value A_3 when

$$s_{112} = s_{121} = L.$$

At this vertex $s_{i_1i_2i_3}$ take the values

$$\begin{aligned}
s_{111} &= s_{222} = 0,\\
s_{122} &= \alpha_1 - 2L,\\
s_{211} &= 3\,\alpha_1 - 1 - 2L,\\
s_{212} &= 1 - 2\,\alpha_1 + L,\\
s_{221} &= 1 - 2\,\alpha_1 + L.
\end{aligned}$$

Naturally, we can partition the space $\left[\frac{1}{3},\frac{2}{3}\right]$ of α_1 into several intervals such that, within each interval, each expression of the form $\max\{\cdot,\cdot\}$ becomes fixed. Then,

$$p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C) = A + \frac{B}{\alpha_1} + \frac{C}{1-\alpha_1},$$

where A, B, C are constants determined by (p_1, p_2, p_3). *We can compute the minimum of* $p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C)$ *over each interval and then determine the global minimum.*

The interval boundaries consist of all switching points of the max functions, along with the endpoints of the interval $\left[\frac{1}{3},\frac{2}{3}\right]$, forming the set

$$\left\{\frac{1}{3},\,\frac{1}{2},\,\frac{2}{3},\,p_1,\,p_2,\,p_3,\,p_1{+}p_2,\,p_1{+}p_3,\,p_2{+}p_3\right\}\cap\left[\frac{1}{3},\frac{2}{3}\right].$$

Considering the symmetry between m_1 and m_2, we can construct an EE' with $s'_{i'_1i'_2i'_3} = s_{i_1i_2i_3}$, $i'_j = 3 - i_j$ and $\alpha'_i = 1 - \alpha$ such that

$$p_{\text{guess}}^{\mathsf{EE}',p_{\text{m}},p_{\text{k}}}(M \mid C) = p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C).$$

Thus, $p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C)$ achieves the same value at α_1 and $1-\alpha_1$. We only need to consider the minimum on the interval $\left[\frac{1}{3},\frac{1}{2}\right]$ with the vertex satisfying

$$\begin{aligned}
s_{111} &= s_{222} = s_{112} = s_{121} = 0,\\
s_{122} &= \alpha_1,\\
s_{211} &= 3\,\alpha_1 - 1,\\
s_{212} &= s_{221} = 1 - 2\,\alpha_1,
\end{aligned}$$

and the boundary point set

$$P \coloneqq \left\{\frac{1}{3},\, \frac{1}{2},\, p_1,\, p_2,\, p_3\right\} \cap \left[\frac{1}{3}, \frac{1}{2}\right].$$

In fact, for $0 < \alpha_1 < 1$, it always holds that $B \leq 0$ and $C \leq 0$ (see the following calculation). Therefore, $p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C)$ is a concave function of α_1 and its minimum occurs at one of the interval endpoints.

P varies in two cases: 1) $p_2 \geq \frac{1}{3}$ and 2) $p_2 \leq \frac{1}{3}$.

Case 1: If $p_2 \geq \frac{1}{3}$, then

$$p_3 = 1 - p_1 - p_2 \leq \frac{1}{3} \leq p_2 \leq p_1 < \frac{1}{2}.$$

Thus,

$$P = \left\{\frac{1}{3},\, p_2,\, p_1,\, \frac{1}{2}\right\}.$$

On each resulting interval, $p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C)$ simplifies to one of the following expressions:

$$\frac{3}{2} - \frac{1-p_3}{2(1-\alpha_1)} - \frac{p_3}{2\alpha_1},$$
$$\frac{5}{2} - \frac{1-p_1}{2\alpha_1} - \frac{p_1+1}{2(1-\alpha_1)},$$
$$2 - \frac{1-p_1}{2\alpha_1} - \frac{p_1}{1-\alpha_1}.$$

Thus, we compute the values on the endpoints:

$$\frac{1}{4}(1-3p_3) + \frac{1}{2},$$
$$\frac{(2p_2-1)(p_1+2p_2-1)}{2(p_2-1)p_2} + \frac{1}{2},$$
$$\frac{(2p_1-1)(3p_1-1)}{2(p_1-1)p_1} + \frac{1}{2},$$
$$\frac{1}{2}(1-2p_1) + \frac{1}{2}.$$

Finally, the global minimum is given by:

$$\min\left\{\frac{1-2p_1}{2},\, \frac{(2p_1-1)(1-3p_1)}{2(1-p_1)p_1},\, \frac{1-3p_3}{4},\, \frac{(1-2p_2)(p_2-p_3)}{2(1-p_2)p_2}\right\} + \frac{1}{2}.$$

Case 2: If $p_2 \leq \frac{1}{3}$, then

$$p_3 \leq p_2 \leq \frac{1}{3} \leq p_1 \leq \frac{1}{2}.$$

Therefore,

$$P = \left\{ \frac{1}{3},\ p_1,\ \frac{1}{2} \right\}.$$

On the resulting intervals, the expression for $p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C)$ becomes:

$$\frac{5}{2} - \frac{1-p_1}{2\alpha_1} - \frac{p_1+1}{2(1-\alpha_1)},$$
$$2 - \frac{1-p_1}{2\alpha_1} - \frac{p_1}{1-\alpha_1}.$$

We evaluate the expression at the endpoints:

$$\frac{3p_1-1}{4} + \frac{1}{2},$$
$$\frac{(2p_1-1)(3p_1-1)}{2(p_1-1)p_1} + \frac{1}{2},$$
$$\frac{1-2p_1}{2} + \frac{1}{2}.$$

Thus, the minimum value of $p_{\text{guess}}^{\mathsf{EE},p_{\text{m}},p_{\text{k}}}(M \mid C)$ is:

$$\min\left\{ \frac{1-2p_1}{2},\ \frac{3p_1-1}{4} \right\} + \frac{1}{2}.$$

Example 3 (restated). *For $p_{\text{m}} = (\frac{1}{2}, \frac{1}{2})$ and $p_{\text{k}} = (p_1, p_2, p_3)$ with $\frac{1}{2} > p_1 \geq p_2 \geq p_3$:*

$$\inf_{\mathsf{HE} \in \Omega_{HE}} p_{\text{guess}}^{\mathsf{HE},p_{\text{m}},p_{\text{k}}}(M \mid C) = \frac{1-2p_1}{2} + \frac{1}{2} > \frac{1}{2}.$$

Derivation of Example 3. By the derivation of Example 2, setting $\alpha_1 = \frac{1}{2}$ naturally implies this bound.

Proposition 5 (restated). *For any $(p_{\text{m}}, p_{\text{k}})$ such that $p_{\text{m}}^{\max} \leq 1/|\mathcal{K}|$, HES satisfies*

$$I(K;C) = 0 \quad \textit{and} \quad p_{\text{guess}}^{\mathsf{HES},p_{\text{m}},p_{\text{k}}}(M \mid C) = p_{\text{k}}^{\max} = p_{\max}.$$

Proof (Proof of Proposition 5). For every key k, the shifted message intervals partition $[0,1)$ and encryption is uniform on the interval of m. Hence

$$p_{\text{c}|\text{k}}(c\,|\,k) = \sum_{m : c \in \mathcal{C}_{\mathsf{HES}}(m,k)} p_{\text{m}}(m) \frac{1}{p_{\text{m}}(m)} = 1,$$

so the ciphertext is uniform and independent of the key. Thus $I(K;C) = 0$.

Let $r = p_{\text{m}}^{\max}$. Since $|\mathcal{K}|r \leq 1$, the shifts $0, r, \ldots, (|\mathcal{K}|-1)r$ place the intervals associated with any fixed message disjointly across the keys. Therefore, for every m and c, at most one key satisfies $D_k(c) = m$. Moreover,

$$p_{\text{m,c}}(m,c) = \sum_{k : D_k(c) = m} p_{\text{k}}(k),$$

and consequently $\max_m p_{\mathrm{m,c}}(m,c) = p_{\mathrm{k}}^{\max}$ for every c. Integrating over $[0,1)$ gives

$$p_{\mathrm{guess}}^{\mathsf{HES},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) = p_{\mathrm{k}}^{\max}.$$

Finally, $p_{\mathrm{k}}^{\max} \geq 1/|\mathcal{K}| \geq p_{\mathrm{m}}^{\max}$, so $p_{\mathrm{k}}^{\max} = p_{\max}$.

Proposition 12 (restated). *For any EE construction $\mathsf{EE}^{(\ell)}$ in the random oracle or ideal cipher models and the associated continuous scheme EE_{c}, as $\ell \to \infty$,*

$$\mathbb{E}[I_{\mathsf{EE}^{(\ell)}}(K;C^{(\ell)})] \to I_{\mathsf{EE}_{\mathrm{c}}}(K;C) \quad \textit{and} \quad \mathbb{E}[p_{\mathrm{guess}}^{\mathsf{EE}^{(\ell)},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C)] \to p_{\mathrm{guess}}^{\mathsf{EE}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C).$$

Proof (Proof of Proposition 12). In an EE construction, $p_{\mathrm{c|m,k}}(c\,|\,m,k) = \frac{1}{2^{\iota} p_{\mathrm{d}}^{(\ell)}(m)}$, so $p_{\mathrm{c|k}}(c\,|\,k)$, $p_{\mathrm{c}}(c)$, and $p_{\mathrm{m,c}}(m,c)$ are constant for all $c \in S_{\boldsymbol{m}}$. Therefore,

$$\begin{aligned} I(K;C) &= \sum_k p_{\mathrm{k}}(k) \sum_c p_{\mathrm{c|k}}(c\,|\,k) \log \frac{p_{\mathrm{c|k}}(c\,|\,k)}{p_{\mathrm{c}}(c)} \\ &= \sum_{\boldsymbol{m}} |S_{\boldsymbol{m}}| \sum_k p_{\mathrm{k}}(k) p_{\mathrm{c|k}}(c\,|\,k) \log \frac{p_{\mathrm{c|k}}(c\,|\,k)}{p_{\mathrm{c}}(c)}. \end{aligned}$$

$$p_{\mathrm{guess}}^{\mathsf{EE}^{(\ell)},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) = \sum_c \max_m p_{\mathrm{m,c}}(m,c) = \sum_{\boldsymbol{m}} |S_{\boldsymbol{m}}| \max_m p_{\mathrm{m,c}}(m,c).$$

It suffices to compute $\mathbb{E}[|S_{\boldsymbol{m}}|]$ in the RO/IC setting. Define $I_c \coloneqq \mathbf{1}\{D_{k_i}(c) = m_{j_i} \text{ for all } i\}$. Then

$$\mathbb{E}[I_c] = \Pr[D_{k_i}(c) = m_{j_i}, \forall i] = \Pr[\mathrm{decode}(c \oplus H(k_i)) = m_{j_i}, \forall i].$$

Since $c \oplus H(k_i)$ is independently uniform over $\{0,1\}^{\ell}$ in the RO/IC models,

$$\mathbb{E}[|S_{\boldsymbol{m}}|] = \sum_c \mathbb{E}[I_c] = \sum_c \Pr[\mathrm{decode}(c \oplus H(k_i)) = m_{j_i}, \forall i] = 2^{\ell} \prod_i p_{\mathrm{d}}^{(\ell)}(m_{j_i}).$$

As $p_{\mathrm{d}}^{(\ell)}(m_{j_i}) \to p_{\mathrm{d}}(m_{j_i})$, we define the continuous scheme EE_{c} by $|S_{\boldsymbol{m}}| \coloneqq \prod_i p_{\mathrm{d}}(m_{j_i})$ and $p_{\mathrm{c|m,k}}(c\,|\,m,k) \coloneqq 1/p_{\mathrm{d}}(m)$, which yields the claimed convergence.

Example 4 (restated). *For $p_{\mathrm{m}} = \left(\frac{1}{2}, \frac{1}{2}\right)$ and $p_{\mathrm{k}} = (p_1, p_2, p_3)$ with $\frac{1}{2} > p_1 \geq p_2 \geq p_3$, in the random-oracle or ideal-cipher model, the continuous HE scheme HE_{c} (with $p_{\mathrm{d}} = p_{\mathrm{m}}$) and EE constructions EE_{c} (with general p_{d}) satisfy the bounds:*

$$p_{\mathrm{guess}}^{\mathsf{HE}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) = \frac{3}{4} \quad \textit{and} \quad \inf_{\mathsf{EE}_{\mathrm{c}} \in \Omega_{EE}} p_{\mathrm{guess}}^{\mathsf{EE}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) = \frac{1}{2} + p_1(1-p_1) \geq \frac{13}{18}.$$

Derivation of Example 4. We borrow the notation from the derivation of Example 2. By Proposition 12, we have $s_{i_1 i_2 i_3} = \alpha_{i_1}\alpha_{i_2}\alpha_{i_3}$ where $i_1, i_2, i_3 \in \{1,2\}$.

For the HE scheme, we have $\alpha_1 = \alpha_2 = \frac{1}{2}$, so $s_{i_1 i_2 i_3} = \frac{1}{8}$, and:

$$p_{\mathrm{guess}}^{\mathsf{HE}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C)$$

$$
\begin{aligned}
&= \sum_{i_1,i_2,i_3} \frac{s_{i_1 i_2 i_3}}{2} \max\left\{ \frac{\sum_{j:i_j=1} p_j}{\alpha_1}, \frac{\sum_{j:i_j=2} p_j}{\alpha_2} \right\} \\
&= \frac{1}{8}\Big(\max\{1,0\} + \max\{1-p_3,p_3\} + \max\{1-p_2,p_2\} + \max\{1-p_1,p_1\} \\
&\qquad + \max\{p_1,1-p_1\} + \max\{p_2,1-p_2\} + \max\{p_3,1-p_3\} + \max\{0,1\}\Big) \\
&= \frac{1}{8}\left(1+1-p_3+1-p_2+1-p_1+1-p_1+1-p_2+1-p_3+1\right) \\
&= \frac{3}{4}.
\end{aligned}
$$

For EE constructions, suppose $\alpha_1 = \alpha, \alpha_2 = 1-\alpha$ and $0 < \alpha \le \frac{1}{2}$. We have:

$$
\begin{aligned}
&p^{\mathsf{EE}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) \\
&= \sum_{i_1,i_2,i_3} \frac{s_{i_1 i_2 i_3}}{2} \max\left\{ \frac{\sum_{j:i_j=1} p_j}{\alpha}, \frac{\sum_{j:i_j=2} p_j}{1-\alpha} \right\} \\
&= \frac{\alpha^3}{2} \max\left\{\frac{1}{\alpha}, 0\right\} \\
&\qquad + \frac{\alpha^2(1-\alpha)}{2}\left(\max\left\{\frac{1-p_3}{\alpha}, \frac{p_3}{1-\alpha}\right\} + \max\left\{\frac{1-p_2}{\alpha}, \frac{p_2}{1-\alpha}\right\} \right. \\
&\qquad \left. + \max\left\{\frac{1-p_1}{\alpha}, \frac{p_1}{1-\alpha}\right\}\right) \\
&\qquad + \frac{\alpha(1-\alpha)^2}{2}\left(\max\left\{\frac{p_1}{\alpha}, \frac{1-p_1}{1-\alpha}\right\} + \max\left\{\frac{p_2}{\alpha}, \frac{1-p_2}{1-\alpha}\right\} \right. \\
&\qquad \left. + \max\left\{\frac{p_3}{\alpha}, \frac{1-p_3}{1-\alpha}\right\}\right) + \frac{(1-\alpha)^3}{2} \max\left\{0, \frac{1}{1-\alpha}\right\} \\
&= \frac{\alpha^2}{2} + \frac{\alpha^2(1-\alpha)}{2}\left(\frac{1-p_3}{\alpha} + \frac{1-p_2}{\alpha} + \frac{1-p_1}{\alpha}\right) \\
&\qquad + \frac{1-\alpha}{2}\left(\max\{p_1,\alpha\} - p_1\alpha + \max\{p_2,\alpha\} - p_2\alpha + \max\{p_3,\alpha\} - p_3\alpha\right) \\
&\qquad + \frac{(1-\alpha)^2}{2} \\
&= \frac{1}{2} + \frac{1-\alpha}{2}\left(\max\{p_1,\alpha\} + \max\{p_2,\alpha\} + \max\{p_3,\alpha\} - \alpha\right) \\
&= \begin{cases} \frac{\alpha^2}{2} - \alpha + 1, & 0 < \alpha \le p_3 \\ \frac{1+(1-\alpha)(1-p_3)}{2}, & p_3 < \alpha \le p_2 \\ \frac{1+(1-\alpha)(p_1+\alpha)}{2}, & p_2 < \alpha \le p_1 \\ \frac{1}{2} + \alpha(1-\alpha), & p_1 < \alpha \le \frac{1}{2} \end{cases}.
\end{aligned}
$$

By simple analysis of the above piecewise function of α, the minimum value is achieved when $\alpha = p_1$, therefore:

$$\inf_{\mathsf{EE}\in\Omega_{\mathrm{EE}}} p_{\mathrm{guess}}^{\mathsf{EE}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(M \mid C) = \frac{1}{2} + p_1(1-p_1).$$

Moreover, the above results show that the optimal encoder is $p_{\mathrm{d}} = (p_1, 1-p_1)$ *but not* $\left(\frac{1}{2}, \frac{1}{2}\right)$ *as in the HE scheme.*

Proofs of the distributional mismatch bounds. As in Section 6.4, messages and keys are independent, with true distributions $(p_{\mathrm{m}}, p_{\mathrm{k}})$ and full-support estimates $(\hat{p}_{\mathrm{m}}, \hat{p}_{\mathrm{k}})$ used to design the scheme.

Proposition 6 (restated). *For a continuous scheme achieving perfect KC under* $(\hat{p}_{\mathrm{m}}, \hat{p}_{\mathrm{k}})$*, the key leakage under* $(p_{\mathrm{m}}, p_{\mathrm{k}})$ *satisfies*

$$I(K;C) \le \mathrm{KL}(p_{\mathrm{m}} \| \hat{p}_{\mathrm{m}}) \le \frac{2\left(\mathrm{SD}(p_{\mathrm{m}}, \hat{p}_{\mathrm{m}})\right)^2}{\ln 2 \, \min_m \hat{p}_{\mathrm{m}}(m)}.$$

Proof. Analogously to p_{c} and $p_{\mathrm{c|k}}$, define $\hat{p}_{\mathrm{c}}$ and $\hat{p}_{\mathrm{c|k}}$ as the corresponding ciphertext densities under $(\hat{p}_{\mathrm{m}}, \hat{p}_{\mathrm{k}})$, with the encryption scheme fixed. Perfect KC and full support of $\hat{p}_{\mathrm{k}}$ give $\hat{p}_{\mathrm{c|k}}(c\,|\,k) = \hat{p}_{\mathrm{c}}(c)$ for every key k. By correctness, for $c \in \mathcal{C}(m,k)$,

$$\frac{p_{\mathrm{c|k}}(c\,|\,k)}{\hat{p}_{\mathrm{c}}(c)} = \frac{p_{\mathrm{m}}(m)\, p_{\mathrm{c|m,k}}(c\,|\,m,k)}{\hat{p}_{\mathrm{m}}(m)\, p_{\mathrm{c|m,k}}(c\,|\,m,k)} = \frac{p_{\mathrm{m}}(m)}{\hat{p}_{\mathrm{m}}(m)}.$$

Integrating over the disjoint regions $\mathcal{C}(m,k)$ therefore yields

$$\int_{\mathcal{C}} p_{\mathrm{c|k}}(c\,|\,k) \log \frac{p_{\mathrm{c|k}}(c\,|\,k)}{\hat{p}_{\mathrm{c}}(c)} \, \mathrm{d}c = \sum_m p_{\mathrm{m}}(m) \log \frac{p_{\mathrm{m}}(m)}{\hat{p}_{\mathrm{m}}(m)} = \mathrm{KL}(p_{\mathrm{m}} \| \hat{p}_{\mathrm{m}}).$$

Averaging over k with weights $p_{\mathrm{k}}(k)$ and splitting the logarithm through $p_{\mathrm{c}}(c)$ gives

$$I(K;C) + \mathrm{KL}(p_{\mathrm{c}} \| \hat{p}_{\mathrm{c}}) = \mathrm{KL}(p_{\mathrm{m}} \| \hat{p}_{\mathrm{m}}).$$

Nonnegativity of KL divergence proves the first bound. For the second, $\ln x \le x - 1$ yields

$$\mathrm{KL}(p_{\mathrm{m}} \| \hat{p}_{\mathrm{m}}) \le \frac{1}{\ln 2} \sum_m \frac{(p_{\mathrm{m}}(m) - \hat{p}_{\mathrm{m}}(m))^2}{\hat{p}_{\mathrm{m}}(m)}.$$

The positive and negative parts of $p_{\mathrm{m}} - \hat{p}_{\mathrm{m}}$ each sum to $\mathrm{SD}(p_{\mathrm{m}}, \hat{p}_{\mathrm{m}})$, so

$$\sum_m (p_{\mathrm{m}}(m) - \hat{p}_{\mathrm{m}}(m))^2 \le 2\left(\mathrm{SD}(p_{\mathrm{m}}, \hat{p}_{\mathrm{m}})\right)^2.$$

Bounding each denominator below by $\min_m \hat{p}_{\mathrm{m}}(m)$ completes the proof.

Proposition 7 (restated). *For the continuous scheme* OE_{c} *designed for* $(\hat{p}_{\mathrm{m}}, \hat{p}_{\mathrm{k}})$*, the excess recovery probability under* $(p_{\mathrm{m}}, p_{\mathrm{k}})$ *satisfies*

$$p_{\mathrm{guess}}(M \mid C) - p_{\max} \le \mathrm{SD}(p_{\mathrm{m}}, \hat{p}_{\mathrm{m}}) + \mathrm{SD}(p_{\mathrm{k}}, \hat{p}_{\mathrm{k}}) + \max\{\mathrm{SD}(p_{\mathrm{m}}, \hat{p}_{\mathrm{m}}), \mathrm{SD}(p_{\mathrm{k}}, \hat{p}_{\mathrm{k}})\}.$$

TDSS Game $\mathrm{TDSS}^{\mathcal{A},f}_{\mathsf{E},p_\mathrm{m},p_\mathrm{k}}$

$M \leftarrow_{p_\mathrm{m}} \mathcal{M}$
$K \leftarrow_{p_\mathrm{k}} \mathcal{K}$
$C \leftarrow E_K(M)$
$b \leftarrow \mathcal{A}(C)$
return $b = f(M)$

Fig. 14: Game for TDSS security

Proof. Write $\hat{p}_{\max} = \max\{\max_m \hat{p}_\mathrm{m}(m), \max_k \hat{p}_\mathrm{k}(k)\}$, analogously to $p_{\max}$. Under $(\hat{p}_\mathrm{m}, \hat{p}_\mathrm{k})$, Theorem 6′ gives

$$p_{\mathrm{guess}}(M \mid C) = \hat{p}_{\max}.$$

Under $(p_\mathrm{m}, p_\mathrm{k})$, keeping the scheme fixed, Lemma 3 gives

$$p_{\mathrm{guess}}(M \mid C) \le \hat{p}_{\max} + \mathrm{SD}(p_\mathrm{m}, \hat{p}_\mathrm{m}) + \mathrm{SD}(p_\mathrm{k}, \hat{p}_\mathrm{k}).$$

Each individual probability changes by at most the corresponding statistical distance, so

$$|\hat{p}_{\max} - p_{\max}| \le \max\{\mathrm{SD}(p_\mathrm{m}, \hat{p}_\mathrm{m}), \mathrm{SD}(p_\mathrm{k}, \hat{p}_\mathrm{k})\}.$$

Combining the two inequalities proves the result.

M Target-distribution Semantic Security

As a stronger notion of MC [29], target-distribution semantic-security (TDSS) models the leakage of partial information about the message from the ciphertext.

Formally, for a predicate $f : \mathcal{M} \to \{0,1\}$, the TDSS game is shown in Fig. 14. An adversary $\mathcal{A}$ is going to predict $f(M)$ given the ciphertext C of M. Clearly, a trivial adversary can simply output the most likely value of $f(M)$ and achieves the success probability of $p_{\mathrm{guess}}(f(M)) := \max_b \Pr[f(M) = b]$. The maximum success probability of any adversary $\mathcal{A}$ is denoted by

$$p_{\mathrm{guess}}(f(M)|C) := \max_{\mathcal{A}} \Pr[\mathrm{TDSS}^{\mathcal{A},f}_{\mathsf{E},p_\mathrm{m},p_\mathrm{k}} \Rightarrow \mathrm{true}].$$

Then, Jaeger et al. [29] defined the TDSS advantage as:

$$\mathbf{Adv}^{\mathrm{tdss}}_{\mathsf{E},p_\mathrm{m},p_\mathrm{k}} := \max_f \left\{ p_{\mathrm{guess}}(f(M)|C) - p_{\mathrm{guess}}(f(M)) \right\}.$$

Subtracting $p_{\mathrm{guess}}(f(M))$ aims to ensure that the minimum TDSS advantage is zero. However, we show that zero TDSS advantage is generally unattainable for arbitrary message and key distributions.

Theorem 23. *If* $|\mathcal{M}| = |\mathcal{K}| = 2$, *then*

$$\min_{\mathsf{E}\in\Omega_E} \mathbf{Adv}^{\mathrm{tdss}}_{\mathsf{E},p_{\mathrm{m}},p_{\mathrm{k}}} = \max\{0, p_{\mathrm{k}}^{\max} - p_{\mathrm{m}}^{\max}\}.$$

Proof. Every predicate on a two-element message space is either constant or bijective. Constant predicates give zero advantage, while predicting a bijective predicate is equivalent to recovering the message. Hence

$$\mathbf{Adv}^{\mathrm{tdss}}_{\mathsf{E},p_{\mathrm{m}},p_{\mathrm{k}}} = p^{\mathsf{E},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) - p_{\mathrm{m}}^{\max}.$$

The MRS lower bound $p_{\mathrm{guess}}(M \mid C) \geq p_{\max}$ therefore gives the claimed TDSS lower bound.

For attainability, label the most likely message and key by 0, and write $a = p_{\mathrm{m}}^{\max}$ and $b = p_{\mathrm{k}}^{\max}$, so $a, b \geq 1/2$. For the XOR scheme $E_k(m) = m \oplus k$, the optimal guessing formula gives

$$\begin{aligned} p_{\mathrm{guess}}(M \mid C) &= \max\{ab, (1-a)(1-b)\} + \max\{a(1-b), (1-a)b\} \\ &= ab + \max\{a - ab, b - ab\} = \max\{a, b\} = p_{\max}. \end{aligned}$$

Thus its TDSS advantage is exactly $\max\{0, b - a\}$. For any ciphertext length $\ell \geq 1$, append $\ell - 1$ independent uniform bits and ignore them during decryption; this preserves the advantage and realizes the same optimum on $\{0,1\}^{\ell}$.

The TDSS bound $2\sqrt{p_{\mathrm{k}}^{\max}} + O(2^{-\ell})$ of Li, Tang, and Zhang [35] applies to their HE construction (see Appendix B) and is not optimal in general, as shown by Theorem 23. The optimal TDSS bound for general distributions remains open.

N Sufficiency and Necessity for Optimal Security

Based on the analysis in Sections 4, 5, and 6, we summarize the necessary and sufficient conditions for continuous schemes to achieve optimal security. For given p_{m} and p_{k}, define the optimal KC and message-confidentiality values over a class Ω of continuous schemes by

$$\begin{aligned} I^{\Omega,p_{\mathrm{m}},p_{\mathrm{k}}}(K;C) &:= \inf_{\mathsf{E}_{\mathrm{c}}\in\Omega} I^{\mathsf{E}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(K;C), \\ p^{\Omega,p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) &:= \inf_{\mathsf{E}_{\mathrm{c}}\in\Omega} p^{\mathsf{E}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C). \end{aligned}$$

For arbitrary $(p_{\mathrm{m}}, p_{\mathrm{k}})$, the optimal values over the class $\Omega_{\mathrm{E,c}}$ of all continuous schemes are

$$I^{\Omega_{\mathrm{E,c}},p_{\mathrm{m}},p_{\mathrm{k}}}(K;C) = 0, \qquad p^{\Omega_{\mathrm{E,c}},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) = p_{\max}.$$

We also consider the classes $\Omega_{\mathrm{EE,c}}$ and $\Omega_{\mathrm{HE,c}}$ of continuous EE constructions and continuous HE schemes. Since every scheme in the continuous model samples from a density on $[0,1)$, no separate deterministic class is needed. We

distinguish continuous schemes by the same attributes as in the discrete setting: ciphertext uniformity (U), and dependence or independence of the ciphertext region measure on the key (KD or KID) and on the message (MD or MID). Thus, $\Omega^{P}_{\mathrm{E,c}}$ denotes the continuous schemes with attributes $P \in \{\mathrm{U}, \mathrm{KD}, \mathrm{KID}, \mathrm{MD}, \mathrm{MID}\}$, or combinations thereof, where dependence is measured through the Lebesgue measure $\mu(m,k)$. Finally, let $\Omega^{\mathrm{KC}}_{\mathrm{E,c}}$ and $\Omega^{\mathrm{MC}}_{\mathrm{E,c}}$ denote the continuous schemes achieving $I(K;C)=0$ and $p_{\mathrm{guess}}(M \mid C) = p_{\max}$, respectively.

Key confidentiality. The results of Section 4 yield:

1. Combining Theorem 4 with Theorem 8, if a p_{m}-agnostic continuous scheme E_{c} achieves perfect KC for every message distribution p_{m}, then it necessarily reveals the message:

$$\left[\forall p_{\mathrm{m}},\ I^{\mathsf{E}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}(K;C)=0\right] \Longrightarrow \left[\forall p_{\mathrm{m}},\ p^{\mathsf{E}_{\mathrm{c}},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C)=1\right].$$

   Thus, except for such message-revealing schemes, no single p_{m}-agnostic continuous scheme can achieve perfect KC for all message distributions; the scheme must instead be tailored to p_{m}.

2. $$\Omega_{\mathrm{EE,c}} = \{\mathsf{E}_{\mathrm{c}} \in \Omega^{\mathrm{U,KID}}_{\mathrm{E,c}} : D_k(c) \neq \perp \text{ for every } k \in \mathcal{K} \text{ and } c \in \mathcal{C}\}.$$

   This is the continuous version of Proposition 8: uniformity, key independence of $\mu(m,k)$, and decryption of every ciphertext under every key characterize continuous EE constructions.

3. $\Omega^{\mathrm{KC}}_{\mathrm{E,c}} \subseteq \Omega^{\text{M-IND}}_{\mathrm{E,c}}$. Thus, continuous schemes achieving optimal KC must be M-IND.

4. $$I^{\Omega_{\mathrm{HE,c}},p_{\mathrm{m}},p_{\mathrm{k}}}(K;C) = I^{\Omega^{\mathrm{U}}_{\mathrm{E,c}},p_{\mathrm{m}},p_{\mathrm{k}}}(K;C) = I^{\Omega_{\mathrm{E,c}},p_{\mathrm{m}},p_{\mathrm{k}}}(K;C) = 0.$$

   A continuous HE scheme with a perfect MS-IND encoder achieves this optimum.

Message-recovery security. The results of Sections 5 and 6 yield:

1. For general message and key distributions, one may have

$$p^{\Omega_{\mathrm{EE,c}},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) > p_{\max}.$$

   Moreover, outside the special cases in Theorem 11, no p_{k}-agnostic uniform continuous scheme is optimal for all p_{k}. Thus, optimal MRS may require tailoring the scheme to p_{k}.

2. $$p^{\Omega^{\mathrm{U}}_{\mathrm{E,c}},p_{\mathrm{m}},p_{\mathrm{k}}}_{\mathrm{guess}}(M \mid C) = p_{\max},$$

   because the continuous scheme $\mathsf{OE}_{\mathrm{c}} \in \Omega^{\mathrm{U}}_{\mathrm{E,c}}$ achieves optimal message confidentiality.