Key-Space Complexity of Information-Theoretic Target-Distribution Semantic Security

Pengcheng Su[1], Haibo Cheng[2], and Ping Wang[3]

[1] School of Computer Science, Peking University pcs@pku.edu.cn,
[2] Peking University & Minzu University of China hbcheng@pku.edu.cn
[3] National Engineering Research Center of Software Engineering, Peking University
pwang@pku.edu.cn

Abstract. We study the uniform key-space complexity of information-theoretic target-distribution semantic security (TDSS) for symmetric encryption. Given a message distribution P, we ask for the minimum size of a uniform key space such that observing the ciphertext improves the optimal prediction probability of any Boolean predicate of the message by at most ε.

Our main construction is a sample-based encryption scheme with κ keys: for a secret key $k \in [\kappa]$, the message is placed in the k-th position, and each of the remaining $\kappa - 1$ positions is filled with an independent sample from a dummy distribution P_D. When $P_D = P$, the posterior distribution of the message conditional on the ciphertext is exactly the empirical distribution of κ independent samples from P. This gives a universal semantic-security bound of $1/(2\sqrt{\kappa})$, so $\kappa = O(\varepsilon^{-2})$ keys suffice for every message distribution, independently of the message-space size. We prove a complementary lower bound in terms of $\|P\|_2$, matching this rate up to constants for sufficiently flat distributions. In particular, for the uniform distribution on $|\mathcal{M}|$ messages, the optimal key-space size is $\Theta\big(\min\{|\mathcal{M}|, \varepsilon^{-2}\}\big)$.

The same framework recovers and generalizes the classical key-length scaling of entropic security: taking P_D uniform on $\{0,1\}^n$, a mixture-and-cloning argument protects every source of min-entropy at least h using $n - h + 2\log(1/\varepsilon) + O(1)$ secret-key bits. It also yields a broader source-family view of information-theoretic semantic security, showing that short-key TDSS is possible for any family of sources with bounded second-order Rényi divergence from a common reference distribution.

We further study zero-advantage TDSS (i.e., $\varepsilon = 0$) for restricted predicate classes $\mathcal{F}$. We show that a perfectly correct cipher achieving zero-advantage TDSS for $\mathcal{F}$ with a uniform κ-element key space exists if and only if P admits a decomposition into κ-sparse posteriors that preserve the prior Bayes-optimal prediction for every $f \in \mathcal{F}$. This sparse-convex characterization implies that $|\mathcal{F}| + 1$ keys always suffice. Surprisingly, point, threshold, and interval predicates can all be protected exactly with only two keys for every message distribution.

Our results connect information-theoretic encryption with sparse convex decompositions and the shuffle model, and provide a systematic view of zero-advantage and approximate TDSS with short secret keys.

1 Introduction

Syntax of symmetric encryption. We use the following standard formalism for symmetric encryption. An encryption scheme E is a tuple $\mathsf{E} = (\mathcal{M}, \mathcal{K}, \mathcal{C}, \mathcal{E}, \mathcal{D})$, where $\mathcal{M}, \mathcal{K}, \mathcal{C}$ denote the message, key, and ciphertext spaces, respectively. $\mathcal{E} = \{\mathsf{Enc}_k : \mathcal{M} \to \mathcal{C}\}_{k\in\mathcal{K}}, \mathcal{D} = \{\mathsf{Dec}_k : \mathcal{C} \to \mathcal{M}\}_{k\in\mathcal{K}}$, where the encryption algorithm Enc_k can be randomized, while the decryption algorithm Dec_k is deterministic. Throughout, the key is independent of the message, and we focus on perfectly correct encryption, meaning that for every $k \in \mathcal{K}$ and $m \in \mathcal{M}$,

$$\Pr[\mathsf{Dec}_k(C) = m : \quad C \leftarrow \mathsf{Enc}_k(m)] = 1.$$

where the probability is over the internal randomness of Enc_k.

Throughout the paper, all message, key, and ciphertext spaces are finite.

Information-theoretic symmetric encryption. Shannon introduced the classical information-theoretic security notion for symmetric encryption [14]. Perfect secrecy requires that the ciphertext C reveal no information about the message M; that is, observing C does not change the distribution of M. Equivalently, for every ciphertext c such that $\Pr[C = c] > 0$,

$$P_{M|C=c} = P_M.$$

For a fixed message distribution P with support $\mathcal{M}$, perfect correctness and perfect secrecy require $|\mathcal{K}| \geq |\mathcal{M}|$, and the one-time pad meets this bound with $|\mathcal{M}|$ keys [14].

Perfect secrecy simultaneously for every message distribution is equivalent to ideal semantic security: for every message distribution P over $\mathcal{M}$, $M \sim P$, and every Boolean predicate f,

$$P_{\text{guess}}(f(M) \mid C) = P_{\text{guess}}(f(M)).$$

Throughout this paper, $P_{\text{guess}}(X)$ and $P_{\text{guess}}(X \mid C)$ denote the maximum probabilities of correctly guessing X without and with the ciphertext, respectively, where the maximum ranges over all adversaries[4].

Computational semantic security. Modern cryptography commonly circumvents this barrier by restricting the adversary's computational power [6]. Fix a computational model, and let $P^{\leq t}_{\text{guess}}(X \mid C)$ and $P^{\leq t}_{\text{guess}}(X)$ denote the maximum probabilities of correctly guessing X, with and without observing C, respectively, over all adversaries of computational cost at most t (e.g., Turing-machine running time or circuit size).

In the predicate-guessing formulation relevant here, computational semantic security requires that, for every message distribution P, every Boolean predicate f, and every polynomial resource bound $t(\lambda)$,

$$P^{\leq t(\lambda)}_{\text{guess}}\big(f(M) \mid C\big) - P^{\leq t(\lambda)}_{\text{guess}}\big(f(M)\big) \leq \text{negl}(\lambda),$$

[4] Explicitly, $P_{\text{guess}}(X) = \max_x \Pr[X = x]$ and $P_{\text{guess}}(X \mid C) = \mathbb{E}_C[\max_x \Pr[X = x \mid C]]$.

where λ is the security parameter and negl denotes a negligible function.

Thus, computational semantic security relies on computational hardness to achieve short secret keys, whereas the notions studied in this paper place no computational restriction on the adversary.

Entropic security. Entropic security studies semantic security in the information-theoretic setting and shortens the key by restricting the message source [3, 13]. For every message distribution of min-entropy at least h (equivalently, $P_{\text{guess}}(M) \leq 2^{-h}$), entropic security requires a single cipher to satisfy, for every predicate f fixed in advance,

$$P_{\text{guess}}(f(M) \mid C) - P_{\text{guess}}(f(M)) \leq \varepsilon.$$

For n-bit messages and $0 < \varepsilon < 1/8$, every entropically secure cipher requires at least $n - h$ key bits [3], while existing constructions achieve key length[5] [3]

$$n - h + 2\log(1/\varepsilon) + O(1).$$

Entropic security shows that uncertainty inherent in the message source can substitute for some shared secret randomness. It describes the admissible sources only through an entropy threshold, however, without exploiting the full structure of a particular distribution.

Target-distribution security. Honey encryption takes a different approach: it fixes a target message distribution P and allows the cipher to use that distribution, so that decryptions under incorrect keys still resemble plausible messages [9]. Its early security analyses focused primarily on recovering the entire message. In the notation above, for $M \sim P$, the message-recovery advantage of a cipher E can be written as

$$\text{Adv}_P^{\text{MR}}(\mathsf{E}) := P_{\text{guess}}(M \mid C) - P_{\text{guess}}(M).$$

A small message-recovery advantage does not preclude the leakage of partial information: the ciphertext may improve the probability of guessing the entire message only slightly while making some property of the message substantially more predictable. To address this issue, Jaeger, Ristenpart, and Tang introduced target-distribution semantic security (TDSS) [8]. For a Boolean predicate $f : \mathcal{M} \to \{0, 1\}$ fixed in advance, define

$$\text{Adv}_{P,f}(\mathsf{E}) := P_{\text{guess}}(f(M) \mid C) - P_{\text{guess}}(f(M)), \qquad M \sim P.$$

The existing notion of TDSS requires this advantage to be small for every Boolean predicate:

$$\text{TDSS}_P(\mathsf{E}) := \sup_{f:\mathcal{M}\to\{0,1\}} \text{Adv}_{P,f}(\mathsf{E}).$$

[5] Throughout this paper, all logarithms are taken to base 2.

The predicate may depend on the target distribution and the encryption scheme, but it must be fixed before the particular ciphertext is observed. After observing the ciphertext, the adversary may decide how to predict that predicate, but may not choose a different predicate for each ciphertext.

Entropic security and TDSS both measure security by the predicate-guessing advantage of an unbounded adversary, but they quantify over message sources differently. Entropic security requires a fixed cipher to protect an entire class of high-entropy sources, whereas TDSS evaluates security for a given distribution P and allows the cipher to exploit that distribution.

Our problem. We consider the following question in the information-theoretic setting: given a target distribution P and an allowable TDSS advantage ε, what is the minimum size of a uniform key space needed to achieve TDSS? We call this quantity the uniform key-space complexity of TDSS. More precisely, let the secret key K be uniform over

$$[\kappa] = \{1, \ldots, \kappa\}$$

and independent of M. We define $\kappa^*_P(\varepsilon)$ as the smallest κ for which there exists a perfectly correct randomized symmetric cipher E satisfying

$$\mathrm{TDSS}_P(\mathsf{E}) \le \varepsilon.$$

To study the key complexity when only specified types of message properties need to be protected, we further introduce a predicate class $\mathcal{F} \subseteq \{0,1\}^{\mathcal{M}}$ and define

$$\mathrm{TDSS}_{P,\mathcal{F}}(\mathsf{E}) := \sup_{f \in \mathcal{F}} \mathrm{Adv}_{P,f}(\mathsf{E}).$$

Accordingly, let $\kappa^*_{P,\mathcal{F}}(\varepsilon)$ denote the minimum number of uniform keys needed to achieve $\mathrm{TDSS}_{P,\mathcal{F}}(\mathsf{E}) \le \varepsilon$. When $\mathcal{F}$ is the class of all Boolean predicates, this definition reduces to $\kappa^*_P(\varepsilon)$.

For zero-advantage TDSS, we abbreviate $\kappa^*_{P,\mathcal{F}} := \kappa^*_{P,\mathcal{F}}(0)$.

1.1 Our results and technical overview

Approximate security for all predicates We first study the original notion of TDSS, which requires protection of all Boolean predicates. Our main quantitative result is the following: if $U_{\mathcal{M}}$ denotes the uniform distribution over the message space $\mathcal{M}$, then

$$\kappa^*_{U_{\mathcal{M}}}(\varepsilon) = \Theta\big(\min\{|\mathcal{M}|, \varepsilon^{-2}\}\big).$$

This result follows from a general lower bound valid for every target distribution and a universal upper bound independent of the message-space size.

Lower bound. The key observation is that perfect correctness constrains the support of every ciphertext posterior. Once a ciphertext c is fixed, each key can decrypt it to at most one message, and hence the posterior distribution

$$\mu_c(m) := \Pr[M = m \mid C = c]$$

is supported on at most κ messages. Using Rademacher averaging, we translate this posterior sparsity into the bound

$$\mathrm{TDSS}_P(\mathsf{E}) \geq \frac{1}{2}\left(\frac{1}{\sqrt{2\kappa}} - \|P\|_2\right)_+.$$

Consequently,

$$\kappa_P^*(\varepsilon) \geq \frac{1}{2(2\varepsilon + \|P\|_2)^2}.$$

For sufficiently flat P, this gives an $\Omega(\varepsilon^{-2})$ lower bound on the number of keys.

Upper bound. We construct a natural sample-based cipher. To encrypt a message m under a key $k \in [\kappa]$, place m in the k-th position of the ciphertext tuple and independently sample each of the remaining $\kappa - 1$ positions from a dummy distribution P:

$$C = (S_1, \ldots, S_\kappa), \qquad S_k = m, \qquad S_i \leftarrow P \quad (i \neq k).$$

Decryption returns the k-th entry. Because $M \sim P$, the κ entries of the ciphertext are independent samples from P, and the secret position remains uniform after conditioning on the ciphertext. The posterior distribution of the message is therefore exactly the empirical distribution of these samples:

$$\mu_C = \frac{1}{\kappa}\sum_{i=1}^{\kappa} \delta_{S_i},$$

where δ_{S_i} denotes the Dirac distribution concentrated at S_i. Applying the variance bound for a binomial random variable to each fixed predicate gives

$$\mathrm{TDSS}_P(\mathsf{SampleEnc}_P^{(\kappa)}) \leq \frac{1}{2\sqrt{\kappa}}.$$

This estimate holds separately for every predicate fixed in advance, so taking the supremum over all predicates introduces no additional factor depending on the number of predicates[6]. It follows that $\kappa = O(\varepsilon^{-2})$ keys suffice for every finite distribution P, independently of $|\mathcal{M}|$.

For the uniform distribution, combining this upper bound with the preceding lower bound and the $|\mathcal{M}|$-key one-time pad yields the tight result stated at the beginning of this subsection.

[6] The price is ciphertext length rather than key length: the ciphertext contains κ messages. Our focus here is the minimum number of possible keys; compression and efficient ciphertext representations are orthogonal questions.

Sample-Based Encryption

$$\mathcal{M}, \qquad \mathcal{K} = [\kappa], \qquad \mathcal{C} = \mathcal{M}^{\kappa}.$$

Encryption $\mathsf{SampleEnc}_{P_D}^{(\kappa)}(m, k)$: for each $i \neq k$, independently sample $s_i \leftarrow P_D$; set $s_k = m$ and output

$$(s_1, \ldots, s_\kappa).$$

Decryption $\mathsf{SampleDec}_{P_D}^{(\kappa)}(c, k) := c[k]$.

From a target distribution to a class of message sources The same construction can instead use a public reference distribution P_D to generate dummy samples. A single cipher can then protect a class of actual message sources, rather than only a target source identical to the dummy distribution. Specifically, we prove that

$$\mathrm{TDSS}_P(\mathsf{SampleEnc}_{P_D}^{(\kappa)}) \leq \frac{1}{2}\sqrt{\frac{2^{D_\infty(P\|P_D)}}{\kappa}}, \tag{1}$$

where

$$D_\infty(P\|P_D) = \log \max_{m\in\mathcal{M}} \frac{P(m)}{P_D(m)},$$

with $D_\infty(P\|P_D) = \infty$ if $P \not\ll P_D$ (i.e., $\mathrm{supp}(P) \not\subseteq \mathrm{supp}(P_D)$).

Therefore, for

$$\mathcal{P}_\infty^\alpha(P_D) := \{P : D_\infty(P\|P_D) \leq \alpha\},$$

a fixed cipher simultaneously protects every message source in the class using key length

$$\alpha + 2\log(1/\varepsilon) + O(1).$$

This result also quantifies the security loss caused by a mismatch between the reference distribution and the actual distribution. Taking P_D to be uniform over $\{0,1\}^n$ recovers the classical entropic-security key length [3]

$$n - h + 2\log(1/\varepsilon) + O(1).$$

The proof uses a mixture-and-cloning reduction. The condition $D_\infty(P\|P_D) \leq \alpha$ allows us to express P_D as a mixture containing P with weight $2^{-\alpha}$ and then reduce the security experiment to sample-based encryption with a random number of identically distributed samples.

A more refined analysis, inspired by the privacy-blanket[7] technique in the shuffle model [1, 15–17], shows that the dependence on $D_\infty(P\|P_D)$ in Equa-

[7] The term *privacy blanket* vividly describes a setting in which a message drawn from P is shuffled together with many dummy messages drawn independently from P_D. Fortunately, in this setting, the posterior distribution of the distinguished message admits a closed-form expression, which enables further analyses of quantities such as key-recovery advantage [15], differential privacy [1, 16], and mutual information [17].

tion (1) can be sharpened to a dependence on

$$D_2(P\|P_D) = \log \sum_m \frac{P(m)^2}{P_D(m)},$$

with $D_2(P\|P_D) = \infty$ if $P \not\ll P_D$.

Since $D_2(P\|P_D) \le D_\infty(P\|P_D)$, this yields a potentially much sharper guarantee. More importantly, it replaces worst-case pointwise likelihood-ratio control by an averaged second-moment condition.

Zero-advantage security for restricted predicate classes The preceding results consider all Boolean predicates and allow a positive security error ε. We next study the case $\varepsilon = 0$ in which only a specified predicate class $\mathcal{F}$ must be protected. For $S_f = f^{-1}(1)$, the predicate advantage satisfies

$$\mathrm{Adv}_{P,f}(\mathsf{E}) = \mathbb{E}_C \left| \mu_C(S_f) - \frac{1}{2} \right| - \left| P(S_f) - \frac{1}{2} \right|.$$

Thus, zero advantage does not require the posterior to equal the prior; it requires only that the posterior probability never cross the $1/2$ Bayes decision boundary determined by the prior. This distinction explains why zero-advantage TDSS for a restricted predicate class can use fewer keys than Shannon perfect secrecy.

Let $\mathcal{P}_{\mathcal{F}}(P)$ denote the polytope of posterior distributions that preserve all these Bayes decisions (see Definition 1 for the formal definition), and let $\Delta^{(\le\kappa)}$ denote the set of distributions supported on at most κ messages. We prove that

$$\kappa^*_{P,\mathcal{F}} \le \kappa \quad \Longleftrightarrow \quad P \in \mathrm{conv}\Big(\mathcal{P}_{\mathcal{F}}(P) \cap \Delta^{(\le\kappa)}\Big),$$

where $\mathrm{conv}(\cdot)$ denotes the convex hull.

In other words, the uniform key-space complexity is exactly the complexity of a sparse, safe posterior decomposition of the prior: every posterior in the decomposition is supported on at most κ messages, preserves the Bayes decision for every predicate in $\mathcal{F}$, and averages to P. This characterization further yields

$$\kappa^*_{P,\mathcal{F}} \le \min\{|\operatorname{supp}(P)|,\, r_P(\mathcal{F}) + 1\} \le \min\{|\operatorname{supp}(P)|,\, |\mathcal{F}| + 1\}.$$

where

$$r_P(\mathcal{F}) := \dim \operatorname{span}\big(\{\mathbf{1}\} \cup \{\mathbf{1}_{S_f} : f \in \mathcal{F}\}\big) - 1,$$

and the vectors are restricted to $\operatorname{supp}(P)$.

For point, threshold, and interval predicates over an ordered message space, we prove that two uniform keys always suffice; except in degenerate cases, two keys are also necessary. The corresponding construction arranges the probability mass of P consecutively around the unit circle and pairs a uniform point with its antipode, producing posterior distributions that both preserve the relevant Bayes decisions and are supported on at most two messages.

Posterior sparsity provides a common structure underlying these results. Perfect correctness forces κ uniform keys to induce only κ-sparse posteriors. For approximate security against all predicates, we construct empirical sparse posteriors; for zero-advantage security over restricted predicate classes, we characterize sparse posterior decompositions that preserve Bayes decisions.

1.2 Contributions

In summary, our contributions are fourfold.

- First, we introduce a distribution-aware sample-based encryption scheme and establish both a universal upper bound and a general lower bound on the uniform key-space complexity of target-distribution semantic security.
- Second, we establish source-family guarantees relative to an arbitrary reference distribution P_D, with the classical entropic-security scaling recovered as the special case of a uniform reference distribution. We give both a conceptually simple mixture-and-cloning analysis and a refined privacy-blanket-inspired analysis, the latter yielding a stronger bound in terms of second-order Rényi divergence.
- Third, we give an exact geometric characterization of predicate-restricted zero-advantage TDSS in terms of sparse, Bayes-plausible posterior decompositions.
- Finally, we instantiate this characterization for several natural predicate classes, showing that a single secret bit suffices to protect point, threshold, and interval predicates over an ordered message space.

1.3 Relation to prior work

Our work differs from entropic security [3, 13] in how the message source is parameterized. Entropic security fixes a class of sources defined by a min-entropy constraint and requires a single cipher to protect the entire class. We first optimize the uniform key space for a given target distribution P and then use a reference distribution P_D to extend the result to source families characterized by bounded divergence from a common public reference distribution; our refined bound is governed by second-order Rényi divergence rather than worst-case pointwise likelihood ratios. Recovering classical entropic security by taking P_D to be uniform is only one special case of this source-class result.

Our work differs from honey encryption and prior work on TDSS [8, 9, 12] in the resource being optimized. Prior work introduced and analyzed distribution-aware security notions and concrete constructions, focusing on weak keys, message recovery, and partial-information security. We keep the secret key strictly uniform, allow arbitrary randomized encryption, and study the minimum key-space cardinality required to achieve TDSS. Our goal is to establish information-theoretic complexity bounds and structural characterizations, rather than to design a practical honey-encryption system.

Khouzani, Malacaria, and others studied optimal encryption channels under generalized entropy measures [2, 10, 11]. Because TDSS measures leakage through the prediction advantage for Boolean predicates, it does not fall directly within these generalized conditional-entropy frameworks. Our posterior-sparsity approach yields a different optimization problem and geometric characterization tailored to TDSS.

Technically, the mixture-and-cloning argument and the blanket analysis technique for our reference-distribution results are inspired by methods developed for privacy amplification by shuffling [1, 4, 5, 15–17].

1.4 Organization

Section 2 proves the upper and lower bounds for security against all predicates. Section 3 establishes the reference-distribution guarantee and its entropic-security corollary. Section 4 develops the sparse-convex characterization of zero-advantage predicate-restricted TDSS. Section 5 treats point, threshold, and interval predicates. Section 6 concludes the paper.

Proofs of Lemmas are deferred to the appendix.

2 Approximate TDSS for All Predicates

We begin with the original all-predicate notion of TDSS and study the approximate regime $\varepsilon > 0$. The main theme of this section is that the size of the uniform key space manifests itself as a sparsity constraint on ciphertext posteriors. We first analyze the sample-based encryption scheme, whose posterior has an especially simple empirical form, and use it to obtain a distribution-independent upper bound. We then show that posterior sparsity is also an inherent consequence of perfect correctness and use a Rademacher-averaging argument to derive a general lower bound. Together, these bounds give a tight characterization for uniform message distributions.

Lemma 1 (Empirical-posterior principle). *If $M \sim P$, $K \sim U_{[\kappa]}$, and $C = \mathsf{SampleEnc}_P^{(\kappa)}(M, K)$, then the coordinates of C are i.i.d. from P, $K \mid C$ is uniform on $[\kappa]$, and $M \mid C$ is*

$$\mu_C(m) = \frac{1}{\kappa} \sum_{i=1}^{\kappa} \delta_{C_i}(m) = \frac{1}{\kappa} \sum_{i=1}^{\kappa} \mathbf{1}\{C_i = m\}.$$

Lemma 2 (Bayes advantage identity). *For every cipher E and every Boolean predicate f (with $S_f = f^{-1}(1)$),*

$$\mathrm{Adv}_{P,f}(\mathsf{E}) = \mathbb{E}_C \left| \mu_C(S_f) - \frac{1}{2} \right| - \left| P(S_f) - \frac{1}{2} \right|.$$

2.1 Upper bounds

Theorem 1 (Universal upper bound). *For every message distribution P and every $\kappa \geq 1$, there exists a perfectly correct cipher E with a uniform κ-element key space such that* $\mathrm{TDSS}_P(\mathsf{E}) \leq \frac{1}{2\sqrt{\kappa}}$.

Consequently, $\kappa_P^*(\varepsilon) \leq \left\lceil \frac{1}{4\varepsilon^2} \right\rceil$, independently of the message-space size $|\mathcal{M}|$ and of the distribution P.

Proof. Consider the sample-based encryption scheme $\mathsf{SampleEnc}_P^{(\kappa)}$. Fix f and put $p_f = P(S_f)$. Lemma 1 gives $\mu_C(S_f) \sim \frac{1}{\kappa}\mathrm{Bin}(\kappa, p_f)$. Denote $X \sim \mathrm{Bin}(\kappa, p_f)$, Lemma 2 implies

$$\begin{aligned}
\mathrm{Adv}_{P,f}(\mathsf{SampleEnc}_P^{(\kappa)}) &= \mathbb{E}_C \left| \mu_C(S_f) - \frac{1}{2} \right| - \left| P(S_f) - \frac{1}{2} \right| \\
&= \mathbb{E}_X \left| \frac{X}{\kappa} - \frac{1}{2} \right| - \left| p_f - \frac{1}{2} \right| \\
&\leq \mathbb{E} \left| \frac{X}{\kappa} - p_f \right| \leq \sqrt{\mathrm{Var}\left(\frac{X}{\kappa}\right)} = \sqrt{\frac{p_f(1-p_f)}{\kappa}} \leq \frac{1}{2\sqrt{\kappa}},
\end{aligned}$$

where the first inequality follows from the reverse triangle inequality and the second from Cauchy–Schwarz. The estimate holds separately for every fixed f, so taking the supremum adds no class-size factor. □

Remark 1 (Exponential decay away from the decision boundary). The analysis can be substantially improved when the prior probability of the predicate is bounded away from the Bayes decision boundary $1/2$. Let $X \sim \mathrm{Bin}(\kappa, p_f)$. If $p_f \neq 1/2$, then

$$\begin{aligned}
\mathbb{E} \left| \frac{X}{\kappa} - \frac{1}{2} \right| - \left| p_f - \frac{1}{2} \right| &\leq \Pr \left[\left| \frac{X}{\kappa} - p_f \right| > \left| p_f - \frac{1}{2} \right| \right] \\
&\leq \exp\left(-2\kappa \left(p_f - \frac{1}{2} \right)^2 \right),
\end{aligned}$$

by Hoeffding's inequality. Thus, for every predicate whose prior probability is bounded away from $1/2$, the semantic advantage of the sample-based scheme decays exponentially in κ. In contrast, when $p_f = 1/2$, the advantage is of order $O(\kappa^{-1/2})$.

Consequently, define the *Bayes gap* of P by $g_P := \min_{S \subseteq \mathcal{M}} \left| P(S) - \frac{1}{2} \right|$. If $g_P > 0$, then every Boolean predicate is uniformly separated from the Bayes decision boundary, and $\mathrm{TDSS}_P\left(\mathsf{SampleEnc}_P^{(\kappa)}\right) \leq \exp(-2\kappa g_P^2)$. Hence $\kappa_P^*(\varepsilon) = O\left(\frac{\log(1/\varepsilon)}{g_P^2}\right)$.

2.2 Lower bound

The preceding argument shows that κ keys are always sufficient to reduce the semantic advantage to order $\kappa^{-1/2}$. A natural question is whether this rate is

merely an artifact of the sample-based construction. We next show that it is inherent for sufficiently flat message distributions. The key observation is that perfect correctness alone forces every ciphertext posterior to be supported on at most κ messages; averaging over Boolean predicates then turns this support constraint into a semantic-leakage lower bound.

Theorem 2 (General lower bound). *Every perfectly correct cipher* E *with a uniform* κ*-element key space satisfies*

$$\mathrm{TDSS}_P(\mathsf{E}) \geq \frac{1}{2}\left(\frac{1}{\sqrt{2\kappa}} - \|P\|_2\right)_+,$$

where $\|P\|_2 := \left(\sum_{m\in\mathcal{M}} P(m)^2\right)^{1/2}$. *It follows that*

$$\kappa_P^*(\varepsilon) \geq \frac{1}{2(2\varepsilon + \|P\|_2)^2}.$$

Proof. Fix a ciphertext c with positive probability. If $\mu_c(m) > 0$, then $\Pr[M = m, C = c] > 0$. Hence, for some key $k \in [\kappa]$, $\Pr[M = m, K = k, C = c] > 0$, and perfect correctness implies $\mathsf{Dec}(c, k) = m$. Thus every message in $\mathrm{supp}(\mu_c)$ is the decryption of c under some key, so $|\operatorname{supp}(\mu_c)| \leq \kappa$.

Since μ_c is a probability distribution, Cauchy–Schwarz gives

$$1 = \|\mu_c\|_1 \leq \sqrt{|\operatorname{supp}(\mu_c)|}\,\|\mu_c\|_2 \leq \sqrt{\kappa}\,\|\mu_c\|_2,$$

and therefore

$$\|\mu_c\|_2 \geq \kappa^{-1/2}.$$

For a Boolean predicate f, define its signed representation

$$z_f(m) := 2f(m) - 1 \in \{\pm 1\}.$$

By Lemma 2,

$$\mathrm{Adv}_{P,f}(\mathsf{E}) = \frac{1}{2}\left(\mathbb{E}_C\left|\langle z_f, \mu_C\rangle\right| - \left|\langle z_f, P\rangle\right|\right).$$

Since Boolean predicates are in one-to-one correspondence with vectors $z \in \{\pm 1\}^{\mathcal{M}}$, we have

$$2\,\mathrm{TDSS}_P(\mathsf{E}) = \max_{z\in\{\pm1\}^{\mathcal{M}}} \left(\mathbb{E}_C|\langle z, \mu_C\rangle| - |\langle z, P\rangle|\right).$$

Let $Z = (Z(m))_{m\in\mathcal{M}}$ be a Rademacher random vector; that is, the coordinates $Z(m)$ are independent and satisfy

$$\Pr[Z(m) = 1] = \Pr[Z(m) = -1] = \frac{1}{2}.$$

The maximum over z is at least the average over the uniformly random choice Z, and hence

$$\begin{aligned} 2\,\mathrm{TDSS}_P(\mathsf{E}) &\geq \mathbb{E}_Z\left[\mathbb{E}_C|\langle Z, \mu_C\rangle| - |\langle Z, P\rangle|\right] \\ &= \mathbb{E}_C\mathbb{E}_Z|\langle Z, \mu_C\rangle| - \mathbb{E}_Z|\langle Z, P\rangle|. \end{aligned}$$

By the Khintchine inequality (see, e.g., [7]), for every vector a,

$$\mathbb{E}_Z|\langle Z,a\rangle| \geq \frac{1}{\sqrt{2}}\|a\|_2, \qquad \mathbb{E}_Z|\langle Z,a\rangle| \leq \|a\|_2.$$

Therefore,

$$2\,\mathrm{TDSS}_P(\mathsf{E}) \geq \frac{1}{\sqrt{2}}\mathbb{E}_C\|\mu_C\|_2 - \|P\|_2 \geq \frac{1}{\sqrt{2\kappa}} - \|P\|_2.$$

Together with the nonnegativity of semantic advantage, this gives

$$\mathrm{TDSS}_P(\mathsf{E}) \geq \frac{1}{2}\left(\frac{1}{\sqrt{2\kappa}} - \|P\|_2\right)_+.$$

Finally, if $\mathrm{TDSS}_P(\mathsf{E}) \leq \varepsilon$, then $\frac{1}{\sqrt{2\kappa}} \leq 2\varepsilon + \|P\|_2$, and hence

$$\kappa \geq \frac{1}{2(2\varepsilon + \|P\|_2)^2}.$$

This proves the second claim. □

Corollary 1. *Let $U_{\mathcal{M}}$ be uniform on a message space $\mathcal{M}$. For every $0 < \varepsilon \leq 1/2$,*

$$\kappa^*_{U_{\mathcal{M}}}(\varepsilon) = \Theta\big(\min\{|\mathcal{M}|, \varepsilon^{-2}\}\big).$$

Proof. For the upper bound, Theorem 1 gives $\kappa^*_{U_{\mathcal{M}}}(\varepsilon) \leq \left\lceil \frac{1}{4\varepsilon^2} \right\rceil$, while perfect secrecy is achievable with $|\mathcal{M}|$ keys, e.g., by a one-time pad over a group of size $|\mathcal{M}|$. Hence

$$\kappa^*_{U_{\mathcal{M}}}(\varepsilon) = O\big(\min\{|\mathcal{M}|, \varepsilon^{-2}\}\big).$$

For the lower bound, since $\|U_{\mathcal{M}}\|_2 = |\mathcal{M}|^{-1/2}$, Theorem 2 implies

$$\kappa^*_{U_{\mathcal{M}}}(\varepsilon) \geq \frac{1}{2(2\varepsilon + |\mathcal{M}|^{-1/2})^2}.$$

Using $2\varepsilon + |\mathcal{M}|^{-1/2} \leq 2\max\{2\varepsilon, |\mathcal{M}|^{-1/2}\}$, we obtain

$$\kappa^*_{U_{\mathcal{M}}}(\varepsilon) = \Omega\big(\min\{|\mathcal{M}|, \varepsilon^{-2}\}\big).$$

Combining the two bounds proves the claim. □

Hence, in the regime $\varepsilon \gtrsim |\mathcal{M}|^{-1/2}$, the optimal key-space size for $P = U_{\mathcal{M}}$ is $\Theta(\varepsilon^{-2})$; below this scale, the complexity saturates at $\Theta(|\mathcal{M}|)$. In particular, when $\mathcal{M} = \{0,1\}^n$ and hence $|\mathcal{M}| = 2^n$, every inverse-polynomial security parameter $\varepsilon = 1/\mathsf{poly}(n)$ lies in the former regime, giving $\kappa^*_{U_{\mathcal{M}}}(\varepsilon) = \Theta(\varepsilon^{-2})$ and therefore only $O(\log(1/\varepsilon))$ secret-key bits.

The results above assume that the dummy distribution used by the sample-based cipher coincides with the actual message distribution. This is natural when the target distribution is known exactly, but it also raises a more robust question: can one fixed sample-based cipher protect several different message sources? We address this question next by allowing the dummy samples to come from an arbitrary public reference distribution.

3 Sample-Based Encryption for a Family of Distributions

We now decouple the distribution used by the encryption scheme from the actual message distribution. Fix a public reference distribution P_D and consider the single cipher $\mathsf{SampleEnc}_{P_D}^{(\kappa)}$. We first show that its security for an arbitrary source P can be controlled by the max-divergence $D_\infty(P\|P_D)$. The proof is based on a mixture-and-cloning reduction: samples from P_D are viewed as containing a random number of "clones" drawn from P, after which the analysis reduces to the matched-distribution case of Section 2. As a special case, taking P_D to be uniform recovers the classical key-length scaling of entropic security.

We then give a more refined, though slightly more involved, analysis inspired by the privacy-blanket technique in the shuffle model [1, 15, 16]. This argument improves the dependence from $D_\infty(P\|P_D)$ to the smaller second-order Rényi divergence $D_2(P\|P_D)$, and can therefore yield substantially sharper guarantees.

We present the conceptually simpler clone-based analysis first in Section 3.1, followed by the refined blanket-inspired analysis in Section 3.2.

3.1 A clone-based warm-up: a Max-Divergence bound

Lemma 3 (Post-processing for guessing advantage). *Let X be any random variable, and suppose*

$$X \longrightarrow \widetilde{C} \longrightarrow C$$

forms a Markov chain. Then

$$P_{\text{guess}}(X \mid C) \le P_{\text{guess}}(X \mid \widetilde{C}).$$

We will use this elementary monotonicity property to analyze a stronger experiment in which the adversary is additionally told which dummy samples arise from the P component of a suitable mixture decomposition of P_D. Revealing this information can only increase the guessing advantage, while exposing the clone structure needed for the reduction.

Theorem 3 (TDSS for a family of distributions). *For every message distribution P,*

$$\mathrm{TDSS}_P(\mathsf{SampleEnc}_{P_D}^{(\kappa)}) \le \frac{1}{2}\sqrt{\frac{2^{D_\infty(P\|P_D)}}{\kappa}}.$$

Consequently, $\mathsf{SampleEnc}_{P_D}^{(\kappa)}$ simultaneously protects every source in $\mathcal{P}_\infty^\alpha(P_D) = \{P \mid D_\infty(P\|P_D) \le \alpha\}$ using $\alpha + 2\log(\frac{1}{\varepsilon}) + O(1)$ secret-key bits.

Proof. If $D_\infty(P\|P_D) = \infty$, the claimed bound is trivial. Otherwise, let

$$\alpha := D_\infty(P\|P_D) < \infty, \qquad \gamma := 2^{-\alpha}.$$

Then

$$P(m) \le 2^\alpha P_D(m)$$

Algorithm 1 Augmented mixture-and-clone experiment for the analysis

Require: $P_D = \gamma P + (1-\gamma)Q$, message $m \in \mathcal{M}$, and secret key $k \in [\kappa]$.

1: **for** $i \in [\kappa] \setminus \{k\}$ **do**

2: Sample $\ell_i \leftarrow \mathrm{Bern}(\gamma)$.

3: Sample

$$s_i \leftarrow \begin{cases} P, & \ell_i = 1, \\ Q, & \ell_i = 0. \end{cases}$$

4: **end for**

5: Set $s_k \leftarrow m$ and

$$S \leftarrow \{k\} \cup \{i \in [\kappa] \setminus \{k\} : \ell_i = 1\}.$$

6: **return**

$$\widetilde{C} = \big((s_1, \ldots, s_\kappa), S\big).$$

for every m, and hence

$$P_D = \gamma P + (1-\gamma)Q$$

for some distribution Q.

Thus, each independent sample from P_D can be generated by first drawing a mixture label: with probability $\gamma := 2^{-\alpha}$, it is sampled from P, and with probability $1-\gamma$ it is sampled from Q. We may reveal all mixture labels (whether this sample is from P) to the adversary. Providing this additional side information can only increase its prediction advantage (Lemma 3), so it suffices to analyze the resulting, stronger experiment (shown in Algorithm 1).

Let $S \subseteq [\kappa]$ denote the set consisting of the true-message coordinate together with all dummy coordinates whose mixture labels select the P component, and let

$$R := |S| = 1 + B, \qquad B \sim \mathrm{Bin}(\kappa - 1, \gamma).$$

Conditioned on the revealed side information and on S, the true-message position remains uniform[8] over the R coordinates in S, and the values in these coordinates are i.i.d. samples from P.

After discarding the revealed Q-coordinates, the remaining experiment is therefore exactly $\mathsf{SampleEnc}_P^{(R)}$. By the universal bound,

$$\mathrm{TDSS}_P\Big(\mathsf{SampleEnc}_{P_D}^{(\kappa)}\Big) \le \frac{1}{2}\,\mathbb{E}\Big[R^{-1/2}\Big].$$

Applying Jensen's inequality to the concave function $x \mapsto \sqrt{x}$ gives

$$\mathbb{E}\Big[R^{-1/2}\Big] = \mathbb{E}\Big[\sqrt{R^{-1}}\Big] \le \sqrt{\mathbb{E}[R^{-1}]}.$$

Since $R = 1 + B$ with $B \sim \mathrm{Bin}(\kappa - 1, \gamma)$,

$$\mathbb{E}\left[\frac{1}{R}\right] = \mathbb{E}\left[\frac{1}{B+1}\right] = \frac{1-(1-\gamma)^\kappa}{\kappa\gamma} \le \frac{1}{\kappa\gamma}.$$

[8] Indeed, for every $s \subseteq [\kappa]$ and $k \in s$, $\Pr[K = k, S = s] = \frac{1}{\kappa}\gamma^{|s|-1}(1-\gamma)^{\kappa-|s|}$, which is independent of k. Hence $K \mid S = s$ is uniform on s.

Combining the above inequalities yields

$$\mathrm{TDSS}_P\Big(\mathsf{SampleEnc}_{P_D}^{(\kappa)}\Big) \le \frac{1}{2\sqrt{\kappa\gamma}} = \frac{1}{2}\sqrt{\frac{2^{\alpha}}{\kappa}},$$

which proves the claimed bound. □

Theorem 3 applies to an arbitrary reference distribution and thus goes beyond the usual high-min-entropy formulation. The classical entropic-security regime is recovered by choosing the uniform distribution as the reference, since max-divergence from the uniform distribution is exactly the corresponding min-entropy deficiency.

Corollary 2 (Recovering the entropic-security bound). *For every distribution P over $\mathcal{M} = \{0,1\}^n$ satisfying $H_\infty(P) \ge h$,*

$$\mathrm{TDSS}_P\big(\mathsf{SampleEnc}_{U_{\mathcal{M}}}^{(\kappa)}\big) \le \frac{1}{2}\sqrt{\frac{2^{n-h}}{\kappa}}.$$

Consequently, a single encryption scheme simultaneously protects every source of min-entropy at least h using $n - h + 2\log(1/\varepsilon) + O(1)$ secret-key bits.

Proof. Let $U_{\mathcal{M}}$ be the uniform distribution on $\mathcal{M} = \{0,1\}^n$. Since $H_\infty(P) \ge h$,

$$\max_{m\in\mathcal{M}} P(m) \le 2^{-h}.$$

Hence

$$D_\infty(P\|U_{\mathcal{M}}) = \log \max_m \frac{P(m)}{U_{\mathcal{M}}(m)} = \log\Big(2^n \max_m P(m)\Big) \le n - h.$$

Applying Theorem 3 with $P_D = U_{\mathcal{M}}$ gives

$$\mathrm{TDSS}_P\big(\mathsf{SampleEnc}_{U_{\mathcal{M}}}^{(\kappa)}\big) \le \frac{1}{2}\sqrt{\frac{2^{D_\infty(P\|U_{\mathcal{M}})}}{\kappa}} \le \frac{1}{2}\sqrt{\frac{2^{n-h}}{\kappa}}.$$

Therefore, it suffices to take $\kappa \ge \frac{2^{n-h}}{4\varepsilon^2}$, which corresponds to

$$\log_2 \kappa = n - h + 2\log_2(1/\varepsilon) + O(1)$$

secret-key bits. The scheme depends only on $U_{\mathcal{M}}$, so the same cipher simultaneously protects every source with min-entropy at least h. □

3.2 Blanket-based analysis and the refined D_2 bound

Theorem 4 (TDSS for a family of distributions, refined). *For every message distribution P,*

$$\mathrm{TDSS}_P(\mathsf{SampleEnc}_{P_D}^{(\kappa)}) \le \frac{1}{2}\sqrt{\frac{2^{D_2(P\|P_D)}}{\kappa}}.$$

Consequently, $\mathsf{SampleEnc}_{P_D}^{(\kappa)}$ simultaneously protects every source in $\mathcal{P}_2^{\alpha}(P_D) = \{P \mid D_2(P\|P_D) \le \alpha\}$ using $\alpha + 2\log(\frac{1}{\varepsilon}) + O(1)$ secret-key bits.

Proof. If $D_2(P\|P_D) = \infty$, the claimed bound is trivial. We may therefore assume that P is absolutely continuous with respect to P_D, and define the likelihood ratio

$$w(m) := \begin{cases} P(m)/P_D(m), & P_D(m) > 0, \\ 0, & P_D(m) = 0. \end{cases}$$

Let

$$Q := P_D^{\otimes\kappa}$$

denote the distribution of κ independent samples from P_D.

For $c = (c_1, \ldots, c_\kappa) \in \mathcal{M}^\kappa$, the ciphertext distribution induced by $\mathsf{SampleEnc}_{P_D}^{(\kappa)}$ satisfies

$$\begin{aligned} P_C(c) &= \frac{1}{\kappa} \sum_{k=1}^{\kappa} P(c_k) \prod_{i \neq k} P_D(c_i) \\ &= Q(c)\, \frac{1}{\kappa} \sum_{k=1}^{\kappa} w(c_k). \end{aligned}$$

Hence,

$$\frac{P_C(c)}{Q(c)} = \frac{1}{\kappa} \sum_{k=1}^{\kappa} w(c_k).$$

Moreover, by Bayes' rule, for every c with $P_C(c) > 0$,

$$\Pr[K = k \mid C = c] = \frac{w(c_k)}{\sum_{i=1}^{\kappa} w(c_i)}.$$

Therefore, for any $S \subseteq \mathcal{M}$,

$$\mu_c(S) := \Pr[M \in S \mid C = c] = \frac{\sum_{i=1}^{\kappa} w(c_i) \mathbf{1}_S(c_i)}{\sum_{i=1}^{\kappa} w(c_i)},$$

where $\mathbf{1}_S$ denotes the indicator function of S. Write

$$W(c) := \sum_{i=1}^{\kappa} w(c_i).$$

Then

$$\frac{P_C(c)}{Q(c)} = \frac{W(c)}{\kappa},$$

and, whenever $W(c) > 0$,

$$\mu_c(S) - \frac{1}{2} = \frac{\sum_{i=1}^{\kappa} w(c_i)\big(\mathbf{1}_S(c_i) - \frac{1}{2}\big)}{W(c)}.$$

Consequently, changing measure from P_C to Q yields

$$\begin{aligned}\mathbb{E}_{C\sim P_C}\left|\mu_C(S)-\frac{1}{2}\right| &= \sum_{c\in\mathcal{M}^\kappa} P_C(c)\left|\mu_c(S)-\frac{1}{2}\right| \\ &= \sum_{c\in\mathcal{M}^\kappa} Q(c)\,\frac{W(c)}{\kappa}\left|\frac{\sum_{i=1}^{\kappa} w(c_i)\big(\mathbf{1}_S(c_i)-\frac{1}{2}\big)}{W(c)}\right| \\ &= \sum_{c\in\mathcal{M}^\kappa} Q(c)\left|\frac{1}{\kappa}\sum_{i=1}^{\kappa} w(c_i)\big(\mathbf{1}_S(c_i)-\frac{1}{2}\big)\right|.\end{aligned}$$

If $W(c)=0$, then $P_C(c)=0$, so such c makes no contribution to the expectation.

Since under Q the coordinates $C_1,\dots,C_\kappa$ are i.i.d. according to P_D, we obtain

$$\mathbb{E}_{C\sim P_C}\left|\mu_C(S)-\frac{1}{2}\right| = \mathbb{E}_{C_1,\dots,C_\kappa \overset{\text{i.i.d.}}{\sim} P_D}\left|\frac{1}{\kappa}\sum_{i=1}^{\kappa} w(C_i)\left(\mathbf{1}_S(C_i)-\frac{1}{2}\right)\right|. \tag{2}$$

Now fix a Boolean predicate f and $S_f=f^{-1}(1)$. Let $X\sim P_D$ and define

$$Y := w(X)\left(\mathbf{1}_{S_f}(X)-\frac{1}{2}\right).$$

Then

$$\begin{aligned}\mathbb{E}[Y] &= \sum_m P_D(m)w(m)\left(\mathbf{1}_{S_f}(m)-\frac{1}{2}\right) \\ &= P(S_f)-\frac{1}{2}.\end{aligned}$$

Let $Y_1,\dots,Y_\kappa$ be i.i.d. copies of Y. By Lemma 2 and (2),

$$\begin{aligned}\mathrm{Adv}_{P,f}\big(\mathsf{SampleEnc}_{P_D}^{(\kappa)}\big) &= \mathbb{E}\left|\frac{1}{\kappa}\sum_{i=1}^{\kappa} Y_i\right| - |\mathbb{E}Y| \\ &\le \mathbb{E}\left|\frac{1}{\kappa}\sum_{i=1}^{\kappa}(Y_i-\mathbb{E}Y)\right| \\ &\le \sqrt{\frac{\mathrm{Var}(Y)}{\kappa}}.\end{aligned}$$

Since

$$\left(\mathbf{1}_{S_f}(X)-\frac{1}{2}\right)^2 = \frac{1}{4},$$

we have

$$\begin{aligned}\mathrm{Var}(Y) &= \mathbb{E}[Y^2]-(\mathbb{E}Y)^2 \\ &= \frac{1}{4}\mathbb{E}_{X\sim P_D}[w(X)^2]-\left(P(S_f)-\frac{1}{2}\right)^2 \\ &= \frac{1}{4}\,2^{D_2(P\|P_D)}-\left(P(S_f)-\frac{1}{2}\right)^2.\end{aligned}$$

Thus, more precisely,

$$\mathrm{Adv}_{P,f}\big(\mathsf{SampleEnc}_{P_D}^{(\kappa)}\big) \le \frac{1}{\sqrt{\kappa}}\sqrt{\frac{1}{4}\,2^{D_2(P\|P_D)} - \left(P(S_f) - \frac{1}{2}\right)^2}.$$

Discarding the subtractive term inside the square root and taking the supremum over all Boolean predicates f gives

$$\mathrm{TDSS}_P\big(\mathsf{SampleEnc}_{P_D}^{(\kappa)}\big) \le \frac{1}{2}\sqrt{\frac{2^{D_2(P\|P_D)}}{\kappa}}.$$

The claimed key-length bound follows immediately. □

4 Geometric Characterization of Zero-Advantage TDSS

This section gives an exact characterization of uniform key-space complexity in the zero-advantage regime. The argument separates the problem into two independent ingredients. First, zero semantic advantage imposes a collection of linear constraints requiring every ciphertext posterior to preserve the prior Bayes-optimal decision for each protected predicate. Second, perfect correctness with κ uniform keys constrains every posterior to have support size at most κ. We show that, apart from Bayes plausibility, this sparsity condition is also sufficient for realizability. Combining the two ingredients reduces optimal cipher design to a sparse convex-decomposition problem.

Definition 1 (Safe posterior polytope). *For a message distribution P over $\mathcal{M}$ and a class of predicates $\mathcal{F} \subseteq \{0,1\}^{\mathcal{M}}$, define the* safe posterior polytope

$$\mathcal{P}_{\mathcal{F}}(P) := \left\{ \mu \in \Delta(\mathcal{M}) : \begin{array}{rcl} P(S_f) < \frac{1}{2} & \Longrightarrow & \mu(S_f) \le \frac{1}{2}, \\ P(S_f) = \frac{1}{2} & \Longrightarrow & \mu(S_f) = \frac{1}{2}, \\ P(S_f) > \frac{1}{2} & \Longrightarrow & \mu(S_f) \ge \frac{1}{2} \end{array} \quad \forall f \in \mathcal{F} \right\}. \tag{3}$$

$\Delta(\mathcal{M})$ denotes the probability simplex over $\mathcal{M}$.

Thus, every posterior in $\mathcal{P}_{\mathcal{F}}(P)$ preserves the prior Bayes-optimal prediction for all predicates in $\mathcal{F}$.

Lemma 4 (Zero-advantage TDSS). *A cipher satisfies $\mathrm{TDSS}_{P,\mathcal{F}}(\mathsf{E}) = 0$ if and only if $\mu_C \in \mathcal{P}_{\mathcal{F}}(P)$ almost surely.*

Lemma 4 characterizes which posteriors are security-compatible, but it does not yet determine which such posterior laws can arise from a cipher with a bounded key space. We next isolate this independent realizability question. The following representation lemma shows that, for a uniform key, posterior sparsity is essentially the only additional constraint.

Lemma 5 (Uniform-key posterior representation). *Let Π be a finite-support distribution over posterior distributions on $\mathcal{M}$. There exists a perfectly correct randomized cipher with*

$$K \sim U_{[\kappa]}, \qquad K \perp M,$$

whose induced posterior μ_C has law Π if and only if

$$\mathbb{E}_{\mu\sim\Pi}[\mu] = P, \quad \text{and} \quad \Pi\big(\Delta^{(\le\kappa)}\big) = 1.$$

The preceding two lemmas separate zero-advantage security from uniform-key realizability. Their combination immediately yields the following sparse-convex characterization.

Theorem 5 (Sparse-convex characterization). *For every finite target distribution P, every predicate class $\mathcal{F}$, and every $\kappa \ge 1$,*

$$\kappa^*_{P,\mathcal{F}} \le \kappa \quad\Longleftrightarrow\quad P \in \operatorname{conv}\Big(\mathcal{P}_{\mathcal{F}}(P) \cap \Delta^{(\le\kappa)}\Big),$$

where $\operatorname{conv}(\cdot)$ *denotes the convex hull. Equivalently,*

$$\kappa^*_{P,\mathcal{F}} = \min\Big\{\kappa : P \in \operatorname{conv}\Big(\mathcal{P}_{\mathcal{F}}(P) \cap \Delta^{(\le\kappa)}\Big)\Big\}.$$

Proof. Suppose first that $\kappa^*_{P,\mathcal{F}} \le \kappa$, and let E be a perfectly correct cipher with $\kappa^*_{P,\mathcal{F}}$ uniform keys and $\mathrm{TDSS}_{P,\mathcal{F}}(\mathsf{E}) = 0$. By Lemma 4, every induced posterior μ_C lies in $\mathcal{P}_{\mathcal{F}}(P)$; by Lemma 5, it is also $\kappa^*_{P,\mathcal{F}}$-sparse and satisfies $\mathbb{E}[\mu_C] = P$. Hence

$$P \in \operatorname{conv}\Big(\mathcal{P}_{\mathcal{F}}(P) \cap \Delta^{(\le\kappa^*_{P,\mathcal{F}})}\Big) \subseteq \operatorname{conv}\Big(\mathcal{P}_{\mathcal{F}}(P) \cap \Delta^{(\le\kappa)}\Big).$$

Conversely, suppose

$$P \in \operatorname{conv}\Big(\mathcal{P}_{\mathcal{F}}(P) \cap \Delta^{(\le\kappa)}\Big).$$

Then this convex decomposition defines a finite law over κ-sparse posteriors with barycenter P. By Lemma 5, it is realizable by a perfectly correct cipher with κ uniform keys; since every posterior lies in $\mathcal{P}_{\mathcal{F}}(P)$, Lemma 4 gives $\mathrm{TDSS}_{P,\mathcal{F}} = 0$. Thus $\kappa^*_{P,\mathcal{F}} \le \kappa$.

The equivalent minimization formula follows immediately. □

Theorem 5 is an exact characterization, but its optimization over sparse posterior decompositions may still be difficult for a general predicate class. We next derive a simple universal upper bound by exploiting the linear structure of the safe posterior polytope. In fact, the relevant parameter is not the cardinality of $\mathcal{F}$ itself, but the rank of its defining indicator vectors.

Theorem 6 (General class-size and rank bounds). *For every finite target distribution P and every finite predicate class $\mathcal{F}$,*

$$\kappa^*_{P,\mathcal{F}} \le \min\big\{|\operatorname{supp}(P)|,\, r_P(\mathcal{F}) + 1\big\} \le \min\big\{|\operatorname{supp}(P)|,\, |\mathcal{F}| + 1\big\},$$

where $r_P(\mathcal{F}) = \dim \operatorname{span}\big(\{\mathbf{1}\} \cup \{\mathbf{1}_{S_f} : f \in \mathcal{F}\}\big) - 1$*, and the vectors are restricted to* $\operatorname{supp}(P)$*.*

Proof. Let $\Omega := \operatorname{supp}(P)$, and restrict all distributions and indicator vectors below to Ω. Write $r := r_P(\mathcal{F})$.

Consider the polytope

$$\mathcal{Q} := \mathcal{P}_{\mathcal{F}}(P) \cap \Delta(\Omega).$$

By definition, P satisfies every constraint defining $\mathcal{P}_{\mathcal{F}}(P)$. Hence $P \in \mathcal{Q}$.

We first show that every vertex of $\mathcal{Q}$ is supported on at most $r+1$ points. Let μ be a vertex of $\mathcal{Q}$, and suppose for contradiction that

$$S := \operatorname{supp}(\mu) \qquad \text{satisfies} \qquad |S| > r+1.$$

Restrict the vectors

$$\mathbf{1}, \qquad \{\mathbf{1}_{S_f} : f \in \mathcal{F}\}$$

to the coordinates in S. Their span has dimension at most $r+1$. Since $|S| > r+1$, there exists a nonzero vector $v \in \mathbb{R}^S$ such that

$$\langle \mathbf{1}, v \rangle = 0$$

and

$$\langle \mathbf{1}_{S_f}, v \rangle = 0 \qquad \text{for every } f \in \mathcal{F}.$$

Extend v by zero outside S.

Because $\mu(m) > 0$ for every $m \in S$, for sufficiently small $\eta > 0$ both

$$\mu^+ := \mu + \eta v, \qquad \mu^- := \mu - \eta v$$

are nonnegative. Moreover,

$$\sum_m \mu^{\pm}(m) = \sum_m \mu(m) \pm \eta \langle \mathbf{1}, v \rangle = 1,$$

so $\mu^{\pm} \in \Delta(\Omega)$. For every $f \in \mathcal{F}$,

$$\mu^{\pm}(S_f) = \mu(S_f) \pm \eta \langle \mathbf{1}_{S_f}, v \rangle = \mu(S_f).$$

Hence all defining constraints of $\mathcal{P}_{\mathcal{F}}(P)$ are preserved, and therefore

$$\mu^+, \mu^- \in \mathcal{Q}.$$

Since $v \neq 0$,

$$\mu^+ \neq \mu^-, \qquad \mu = \frac{1}{2}(\mu^+ + \mu^-),$$

contradicting that μ is a vertex. Thus every vertex μ of $\mathcal{Q}$ satisfies

$$|\operatorname{supp}(\mu)| \leq r+1.$$

Since $\mathcal{Q}$ is a finite-dimensional polytope, it is the convex hull of its vertices. As $P \in \mathcal{Q}$, we obtain

$$P \in \operatorname{conv}\Big(\mathcal{P}_{\mathcal{F}}(P) \cap \Delta^{(\leq r+1)}\Big).$$

The sparse-convex characterization, Theorem 5, therefore gives

$$\kappa^*_{P,\mathcal{F}} \le r + 1.$$

On the other hand, P itself belongs to $\mathcal{P}_{\mathcal{F}}(P)$ and is supported on $|\operatorname{supp}(P)|$ points. Hence

$$P \in \mathcal{P}_{\mathcal{F}}(P) \cap \Delta^{(\le |\operatorname{supp}(P)|)},$$

and another application of Theorem 5 yields

$$\kappa^*_{P,\mathcal{F}} \le |\operatorname{supp}(P)|.$$

Combining the two bounds,

$$\kappa^*_{P,\mathcal{F}} \le \min\left\{|\operatorname{supp}(P)|,\, r_P(\mathcal{F}) + 1\right\}.$$

Finally,

$$\dim \operatorname{span}\left(\{\mathbf{1}\} \cup \{\mathbf{1}_{S_f} : f \in \mathcal{F}\}\right) \le |\mathcal{F}| + 1,$$

and hence

$$r_P(\mathcal{F}) \le |\mathcal{F}|.$$

Therefore

$$\min\left\{|\operatorname{supp}(P)|,\, r_P(\mathcal{F}) + 1\right\} \le \min\left\{|\operatorname{supp}(P)|,\, |\mathcal{F}| + 1\right\},$$

as claimed. □

The rank bound is completely general, but it need not reflect the true key complexity of a structured predicate class. For some natural families, the safe posterior polytope admits much sparser decompositions than the generic vertex bound suggests. We illustrate this phenomenon next with point, threshold, and interval predicates.

5 One-Bit Keys Suffice for Several Predicate Classes

We now instantiate the geometric framework on several simple but expressive predicate classes over an ordered message space. Although these classes may contain up to $\Theta(N^2)$ predicates, we show that their zero-advantage key complexity is completely independent of N: two uniform keys always suffice. The proof uses a direct two-point posterior construction based on a half-rotation of the probability mass of P on the unit circle.

Let

$$\mathcal{M} = [N] = \{1, \dots, N\},$$

and consider

$$\begin{aligned}
\mathcal{F}_{\mathrm{pt}} &= \{\mathbf{1}\{M = i\} : i \in [N]\},\\
\mathcal{F}_{\mathrm{th}} &= \{\mathbf{1}\{M \le i\} : i \in [N]\},\\
\mathcal{F}_{\mathrm{int}} &= \{\mathbf{1}\{a \le M \le b\} : 1 \le a \le b \le N\}.
\end{aligned}$$

These classes capture different forms of partial information about an ordered message. Point predicates ask whether the message equals a particular value, and hence protect against membership or identification tests for individual messages. Threshold predicates protect ordinal information, such as whether the message lies below a prescribed cutoff. Interval predicates are more expressive ($\mathcal{F}_{\mathrm{pt}} \subseteq \mathcal{F}_{\mathrm{int}},\ \mathcal{F}_{\mathrm{th}} \subseteq \mathcal{F}_{\mathrm{int}}$): they protect membership in every contiguous range and contain both point predicates and one-sided thresholds as special cases.

The same classes are canonical examples in statistical learning theory. Point functions and thresholds have VC dimension one, while intervals have VC dimension two. Their cardinalities, however, can be large:

$$|\mathcal{F}_{\mathrm{pt}}| = N, \qquad |\mathcal{F}_{\mathrm{th}}| = N, \qquad |\mathcal{F}_{\mathrm{int}}| = \Theta(N^2).$$

Thus the general bound

$$\kappa^*_{P,\mathcal{F}} \le |\mathcal{F}| + 1$$

is far from tight for these families. In fact, their exact key complexity is independent of both N and the size of the predicate class.

In other words, a single uniformly random secret bit is sufficient to achieve zero semantic advantage simultaneously for all point, threshold, or interval queries, for *every* target distribution P.

Theorem 7 (Key complexity of several predicate classes). *For every distribution P on $[N]$,*

$$\kappa^*_{P,\mathcal{F}_{\mathrm{pt}}} \le 2, \qquad \kappa^*_{P,\mathcal{F}_{\mathrm{th}}} \le 2, \qquad \kappa^*_{P,\mathcal{F}_{\mathrm{int}}} \le 2.$$

For each of the three classes, the complexity is exactly two whenever $|\operatorname{supp}(P)| \ge 2$.

Proof. Since $\mathcal{F}_{\mathrm{pt}},\ \mathcal{F}_{\mathrm{th}} \subseteq \mathcal{F}_{\mathrm{int}}$, it therefore suffices to construct a Bayes-plausible law over two-sparse posteriors that is safe for every interval predicate.

We use the half-rotation construction sketched in the introduction. Let

$$F_i := \sum_{j \le i} P(j), \qquad F_0 := 0,$$

and identify each message i with the interval $I_i = [F_{i-1}, F_i)$ on the unit circle. Draw $U \sim U[0,1)$ and let

$$V = U + \frac{1}{2} \pmod 1.$$

Let X and Y be the messages whose intervals contain U and V, respectively, and define

$$\mu_U := \frac{1}{2}(\delta_X + \delta_Y).$$

Since both U and V are marginally uniform, $X, Y \sim P$, and hence

$$\mathbb{E}_U[\mu_U] = P.$$

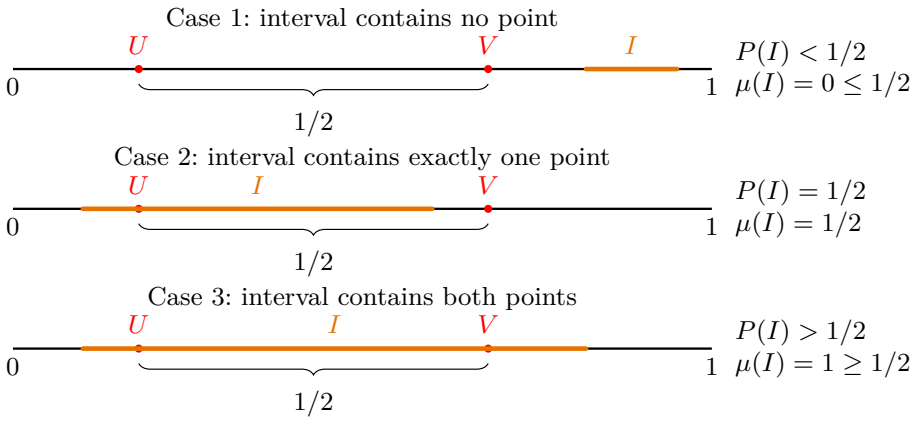


Fig. 1. Half-rotation construction for interval predicates. The antipodal points U, V are separated by CDF-distance $1/2$. An interval of prior mass below, equal to, or above $1/2$ contains respectively at most one, exactly one, or at least one of the two points, so the Bayes-optimal prediction is preserved.

Thus the posterior law is Bayes-plausible, and every μ_U is supported on at most two messages.

Now fix an interval predicate $f = \mathbf{1}\{a \le M \le b\}$ and let $S_f = \{a, a+1, \dots, b\}$. Under the CDF representation, S_f corresponds to an arc of length $P(S_f)$. As illustrated in Figure 1, an arc of length less than $1/2$ cannot contain both antipodal points U, V, while an arc of length greater than $1/2$ cannot exclude both; if its length is exactly $1/2$, it contains exactly one of them almost surely. Therefore

$$P(S_f) < \frac{1}{2} \implies \mu_U(S_f) \le \frac{1}{2},$$

$$P(S_f) = \frac{1}{2} \implies \mu_U(S_f) = \frac{1}{2},$$

and

$$P(S_f) > \frac{1}{2} \implies \mu_U(S_f) \ge \frac{1}{2}.$$

Hence

$$\mu_U \in \mathcal{P}_{\mathcal{F}_{\mathrm{int}}}(P) \cap \Delta^{(\le 2)} \qquad \text{almost surely.}$$

Together with $\mathbb{E}[\mu_U] = P$, Theorem 5 gives $\kappa^*_{P,\mathcal{F}_{\mathrm{int}}} \le 2$.

If some predicate in the class is nonconstant on $\mathrm{supp}(P)$, then one key cannot achieve zero advantage: with a single key and perfect correctness, the ciphertext determines M, and hence determines that predicate, while its prior guessing probability is strictly less than one. Thus the complexity is exactly two whenever $|\mathrm{supp}(P)| \ge 2$. □

Remark 2. The explicit construction of this encryption scheme can be restated as follows. Let $\mathcal{K} = \{1, 2\}$ and $\mathcal{C} = \mathcal{M}^2$. To encrypt a message m_j under key $k \in \mathcal{K}$, sample

$$U \sim U_{I_j},$$

set

$$c_k = m_j,$$

and let

$$c_{3-k} = m_{j'} \quad \text{where} \quad \left(U + \frac{1}{2}\right) \bmod 1 \in I_{j'}.$$

The ciphertext is then

$$C \leftarrow \mathsf{Enc}_k(m_j) = (c_1, c_2).$$

Remark 3. This example also highlights a distinction from standard learning-theoretic uniform convergence. Learning theory typically asks whether an empirical distribution approximates $p(f)$ uniformly over all $f \in \mathcal{F}$, with sample complexity controlled by quantities such as VC dimension. Here we ask for something different: every posterior may be far from the prior, as long as it remains on the correct side of the Bayes decision boundary $1/2$ for every protected predicate. The half-rotation construction exploits precisely this weaker geometric requirement, allowing zero-advantage TDSS with only one secret bit.

6 Conclusion

We studied the uniform key-space complexity of information-theoretic target-distribution semantic security. Our results reveal two complementary structural perspectives on short-key TDSS.

For approximate security, the sample-based encryption scheme yields a universal $O(\epsilon^{-2})$ upper bound on the key-space size, together with matching lower bounds for sufficiently flat source distributions. More generally, by introducing a public reference distribution P_D, we obtain simultaneous security guarantees for an entire family of message sources. The resulting bounds recover classical entropic security as the special case of a uniform reference distribution, and a refined analysis improves the dependence from the max-divergence $D_\infty(P\|P_D)$ to the second-order Rényi divergence $D_2(P\|P_D)$.

For zero-advantage security against a restricted predicate class, we show that the optimal uniform key-space size is characterized by the sparsity of Bayes-plausible posterior decompositions. This geometric viewpoint yields general rank-based upper bounds and, for point, threshold, and interval predicates, shows that a single secret bit suffices.

Several questions remain open. A natural direction is to develop an approximate analogue of the sparse-convex characterization, with the goal of obtaining sharper bounds on $\kappa^*_{P,\mathcal{F}}(\epsilon)$ that simultaneously capture the fine structure of the source distribution and the geometry of the predicate class. Another direction is to relax the uniform-key assumption. In that setting, posterior support size alone no longer captures the available secret randomness, suggesting a richer connection among the key distribution, key entropy, and semantic leakage.

References

1. Balle, B., Bell, J., Gascón, A., Nissim, K.: The privacy blanket of the shuffle model. In: Proceedings of CRYPTO 2019. pp. 638–667 (2019)
2. Biondi, F., Given-Wilson, T., Legay, A.: Universal Optimality of Apollonian Cell Encoders. In: 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE). pp. 424–435 (Aug 2018)
3. Dodis, Y., Smith, A.: Entropic security and the encryption of high entropy messages. In: Proceedings of the Second International Conference on Theory of Cryptography. pp. 556–577 (2005). https://doi.org/10.1007/978-3-540-30576-7_30
4. Erlingsson, Ú., Feldman, V., Mironov, I., Raghunathan, A., Talwar, K., Thakurta, A.: Amplification by shuffling: from local to central differential privacy via anonymity. In: Proceedings of the Thirtieth Annual ACM-SIAM SODA. pp. 2468–2479 (2019)
5. Feldman, V., McMillan, A., Talwar, K.: Hiding among the clones: A simple and nearly optimal analysis of privacy amplification by shuffling. In: IEEE FOCS 2021. pp. 954–964 (2022). https://doi.org/10.1109/FOCS52979.2021.00096
6. Goldwasser, S., Micali, S.: Probabilistic encryption. J. Comput. Syst. Sci. **28**, 270–299 (1984)
7. Haagerup, U.: The best constants in the khintchine inequality. Studia Mathematica **70**(3), 231–283 (1981)
8. Jaeger, J., Ristenpart, T., Tang, Q.: Honey encryption beyond message recovery security. In: Advances in Cryptology – EUROCRYPT 2016. pp. 758–788 (2016). https://doi.org/10.1007/978-3-662-49890-3_29
9. Juels, A., Ristenpart, T.: Honey encryption: Security beyond the brute-force bound. In: Advances in Cryptology – EUROCRYPT 2014. pp. 293–310 (2014)
10. Khouzani, M., Malacaria, P.: Relative perfect secrecy: Universally optimal strategies and channel design. In: 2016 IEEE 29th Computer Security Foundations Symposium (CSF). pp. 61–76 (2016). https://doi.org/10.1109/CSF.2016.12
11. Khouzani, M., Malacaria, P.: Generalized entropies and metric-invariant optimal countermeasures for information leakage under symmetric constraints. IEEE Transactions on Information Theory **65**(2), 888–901 (2019). https://doi.org/10.1109/TIT.2018.2883705
12. Li, X., Tang, Q., Zhang, Z.: Fooling an Unbounded Adversary with a Short Key, Repeatedly: The Honey Encryption Perspective. In: 2nd Conference on Information-Theoretic Cryptography (ITC 2021). vol. 199, pp. 23:1–23:21 (2021). https://doi.org/10.4230/LIPIcs.ITC.2021.23
13. Russell, A., Wang, H.: How to fool an unbounded adversary with a short key. IEEE Transactions on Information Theory **52**(3), 1130–1140 (2006). https://doi.org/10.1109/TIT.2005.864438
14. Shannon, C.E.: Communication theory of secrecy systems. The Bell System Technical Journal **28**(4), 656–715 (1949). https://doi.org/10.1002/j.1538-7305.1949.tb00928.x
15. Su, P., Cheng, H., Wang, P.: Bayesian Advantage of Re-Identification Attack in the Shuffle Model. In: 2026 IEEE 39th Computer Security Foundations Symposium (CSF). pp. 49–63 (Jul 2026). https://doi.org/10.1109/CSF68417.2026.00004
16. Su, P., Cheng, H., Wang, P.: Decomposition-Based Optimal Bounds for Privacy Amplification via Shuffling. In: 2026 IEEE Symposium on Security and Privacy (SP). pp. 2191–2210 (May 2026). https://doi.org/10.1109/SP63933.2026.00151

17. Su, P., Cheng, H., Wang, P.: Which privacy blanket is optimal in the shuffle model? Cryptology ePrint Archive, Paper 2026/667 (2026), https://eprint.iacr.org/2026/667

A Proof of Lemmas

Lemma 1 (Empirical-posterior principle). *If $M \sim P$, $K \sim U_{[\kappa]}$, and $C = \mathsf{SampleEnc}_P^{(\kappa)}(M, K)$, then the coordinates of C are i.i.d. from P, $K \mid C$ is uniform on $[\kappa]$, and $M \mid C$ is*

$$\mu_C(m) = \frac{1}{\kappa}\sum_{i=1}^{\kappa} \delta_{C_i}(m) = \frac{1}{\kappa}\sum_{i=1}^{\kappa} \mathbf{1}\{C_i = m\}.$$

Proof. For $c = (c_1, \ldots, c_\kappa)$ and $k \in [\kappa]$,

$$\Pr[C = c, K = k] = \frac{1}{\kappa} P(c_k) \prod_{i \neq k} P(c_i) = \frac{1}{\kappa} \prod_{i=1}^{\kappa} P(c_i),$$

which is independent of k. Bayes' rule proves $K \mid C$ is uniform on $[\kappa]$. Since $M = C_K$, the posterior mass of m is $\kappa^{-1} \sum_k \mathbf{1}\{c_k = m\}$. □

Lemma 2 (Bayes advantage identity). *For every cipher E and every Boolean predicate f (with $S_f = f^{-1}(1)$),*

$$\mathrm{Adv}_{P,f}(\mathsf{E}) = \mathbb{E}_C \left| \mu_C(S_f) - \frac{1}{2} \right| - \left| P(S_f) - \frac{1}{2} \right|.$$

Proof. If a Bernoulli random variable has mean q, then the optimal constant prediction succeeds with probability

$$\max\{q, 1-q\} = \frac{1}{2} + \left| q - \frac{1}{2} \right|.$$

Applying this identity first to the prior distribution of $f(M)$ and then to its posterior distribution conditioned on C, and averaging over the ciphertext, gives the claim. □

Lemma 3 (Post-processing for guessing advantage). *Let X be any random variable, and suppose*

$$X \longrightarrow \widetilde{C} \longrightarrow C$$

forms a Markov chain. Then

$$P_{\mathrm{guess}}(X \mid C) \leq P_{\mathrm{guess}}(X \mid \widetilde{C}).$$

Proof. Any predictor of X from C can be simulated from $\widetilde{C}$ by first generating C according to the post-processing channel $\widetilde{C} \mapsto C$ and then applying the predictor. Hence access to $\widetilde{C}$ cannot decrease the optimal guessing probability.

Lemma 4 (Zero-advantage TDSS). *A cipher satisfies* $\mathrm{TDSS}_{P,\mathcal{F}}(\mathsf{E}) = 0$ *if and only if* $\mu_C \in \mathcal{P}_{\mathcal{F}}(P)$ *almost surely.*

Proof. Fix $f \in \mathcal{F}$, and write $p_f := P(S_f),\ X := \mu_C(S_f)$. By Bayes plausibility,

$$\mathbb{E}[X] = p_f.$$

Using Lemma 2, we characterize when $\mathrm{Adv}_{P,f}(\mathsf{E}) = 0$.

If $p_f < 1/2$, then

$$\begin{aligned}\mathrm{Adv}_{P,f}(\mathsf{E}) &= \mathbb{E}\left|X - \frac{1}{2}\right| - \left(\frac{1}{2} - p_f\right) \\ &= \mathbb{E}\left[\left|X - \frac{1}{2}\right| + X - \frac{1}{2}\right] \\ &= 2\mathbb{E}\left[\left(X - \frac{1}{2}\right)_+\right].\end{aligned}$$

Hence $\mathrm{Adv}_{P,f}(\mathsf{E}) = 0$ if and only if $X \le \frac{1}{2}$ almost surely. The other cases $p_f = 1/2$ and $p_f > 1/2$ are similar. □

Lemma 5 (Uniform-key posterior representation). *Let* Π *be a finite-support distribution over posterior distributions on* $\mathcal{M}$*. There exists a perfectly correct randomized cipher with*

$$K \sim U_{[\kappa]}, \qquad K \perp M,$$

whose induced posterior μ_C *has law* Π *if and only if*

$$\mathbb{E}_{\mu\sim\Pi}[\mu] = P, \quad \text{and} \quad \Pi\big(\Delta^{(\le\kappa)}\big) = 1.$$

Proof. Necessity is immediate by the law of total probability and the support bound from perfect-correctness. It remains to show sufficiency.

Write the finite-support law Π as

$$\Pi = \sum_{a\in\mathcal{A}} \lambda_a \delta_{\mu_a}, \qquad P = \sum_{a\in\mathcal{A}} \lambda_a \mu_a,$$

where

$$S_a := \mathrm{supp}(\mu_a), \qquad s_a := |S_a| \le \kappa.$$

For each a, let Σ_a be the set of injections

$$\sigma : S_a \hookrightarrow [\kappa],$$

and write

$$N_a := |\Sigma_a| = (\kappa)_{s_a} = \frac{\kappa!}{(\kappa - s_a)!}.$$

We construct a ciphertext for every pair (a, σ) and define the joint distribution of (M, K, C) by

$$\Pr[M = m, K = k, C = (a, \sigma)] := \begin{cases} \dfrac{\lambda_a \mu_a(m)}{N_a}, & \sigma(m) = k, \\ 0, & \text{otherwise.} \end{cases}$$

For fixed a, m, k with $m \in S_a$, exactly $\frac{N_a}{\kappa}$ injections in Σ_a satisfy $\sigma(m) = k$. Hence

$$\begin{aligned} \Pr[M = m, K = k] &= \sum_a \frac{\lambda_a \mu_a(m)}{N_a} \cdot \frac{N_a}{\kappa} \\ &= \frac{1}{\kappa} \sum_a \lambda_a \mu_a(m) = \frac{P(m)}{\kappa}. \end{aligned}$$

Thus $M \sim P$, $K \sim U_{[\kappa]}$, and $K \perp M$.

Define decryption by

$$\mathsf{Dec}((a, \sigma), k) := \sigma^{-1}(k)$$

whenever $k \in \sigma(S_a)$, and arbitrarily otherwise. Every positive-probability triple $(m, k, (a, \sigma))$ satisfies $\sigma(m) = k$, so decryption returns m; hence the cipher is perfectly correct. The above joint distribution therefore defines a randomized encryption algorithm by conditioning on (M, K).

Finally,

$$\Pr[C = (a, \sigma)] = \frac{\lambda_a}{N_a},$$

and, conditioned on $C = (a, \sigma)$,

$$\Pr[M = m \mid C = (a, \sigma)] = \mu_a(m).$$

Thus every ciphertext (a, σ) induces posterior μ_a, and the total probability of ciphertexts inducing μ_a is

$$N_a \cdot \frac{\lambda_a}{N_a} = \lambda_a.$$

Therefore the law of μ_C is exactly Π. □