

# Non-Malleable Encryption: Simpler, Shorter, Stronger

Sandro Coretti\*  
ETH Zurich  
corettis@inf.ethz.ch

Yevgeniy Dodis†  
New York University  
dodis@cs.nyu.edu

Björn Tackmann‡  
UC San Diego  
btackmann@eng.ucsd.edu

Daniele Venturi§  
Sapienza University of Rome  
venturi@di.uniroma1.it

## Abstract

In a seminal paper, Dolev *et al.* [15] introduced the notion of *non-malleable* encryption (NM-CPA). This notion is very intriguing since it suffices for many applications of chosen-ciphertext secure encryption (IND-CCA), and, yet, can be generically built from semantically secure (IND-CPA) encryption, as was shown in the seminal works by Pass *et al.* [29] and by Choi *et al.* [9], the latter of which provided a black-box construction. In this paper we investigate three questions related to NM-CPA security:

1. Can the rate of the construction by Choi *et al.* of NM-CPA from IND-CPA be improved?
2. Is it possible to achieve multi-bit NM-CPA security more efficiently from a single-bit NM-CPA scheme than from IND-CPA?
3. Is there a notion stronger than NM-CPA that has natural applications and can be achieved from IND-CPA security?

We answer all three questions in the positive. First, we improve the rate in the construction of Choi *et al.* by a factor  $\mathcal{O}(\lambda)$ , where  $\lambda$  is the security parameter. Still, encrypting a message of size  $\mathcal{O}(\lambda)$  would require ciphertext and keys of size  $\mathcal{O}(\lambda^2)$  times that of the IND-CPA scheme, even in our improved scheme. Therefore, we show a more efficient domain extension technique for building a  $\lambda$ -bit NM-CPA scheme from a single-bit NM-CPA scheme with keys and ciphertext of size  $\mathcal{O}(\lambda)$  times that of the NM-CPA one-bit scheme. To achieve our goal, we define and construct a novel type of continuous non-malleable code (NMC), called *secret-state NMC*, as we show that standard continuous NMCs are *not enough* for the natural “encode-then-encrypt-bit-by-bit” approach to work.

Finally, we introduce a new security notion for public-key encryption (PKE) that we dub *non-malleability under (chosen-ciphertext) self-destruct attacks* (NM-SDA). After showing that NM-SDA is a *strict* strengthening of NM-CPA and allows for more applications, we nevertheless show that both of our results—(faster) construction from IND-CPA and domain extension from one-bit scheme—also hold for our stronger NM-SDA security. In particular, the notions of IND-CPA, NM-CPA, and NM-SDA security are all equivalent, lying (plausibly, strictly?) below IND-CCA security.

---

\*Supported by SNF project no. 200020-132794.

†Partially supported by gifts from VMware Labs and Google, and NSF grants 1319051, 1314568, 1065288, 1017471.

‡Work partially done while author was at ETH Zurich. Author was partially supported by the SNF Fellowship P2EZF2\_155566 and NSF grants CNS-1228890 and CNS-1116800.

§Partially supported by the European Commission (Directorate-General Home Affairs) under the GAINS project HOME/2013/CIPS/AG/4000005057, and by the European Union’s Horizon 2020 research and innovation programme under grant agreement No 644666.

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Preliminaries</b>	<b>5</b>
<b>3</b>	<b>Non-Malleability under Self-Destruct Attacks</b>	<b>6</b>
<b>4</b>	<b>Domain Extension</b>	<b>7</b>
4.1	A New Flavor of Non-Malleable Codes . . . . .	7
4.2	Combining Single-bit PKE and Non-Malleable Codes . . . . .	8
4.3	Non-Malleable Code Construction . . . . .	10
4.4	Proof of the Non-Malleable Code Construction . . . . .	11
4.4.1	Tamper-Query Types . . . . .	11
4.4.2	Analyzing Query Types . . . . .	11
4.4.3	Hybrids . . . . .	13
4.4.4	Simulation . . . . .	14
4.5	LECSS for the Non-Malleable Code . . . . .	15
4.6	Impossibility for Codes without State . . . . .	16
<b>5</b>	<b>Construction from CPA Security</b>	<b>16</b>
5.1	The CDMW construction. . . . .	17
5.2	Security Proof of the CDMW Construction . . . . .	18
5.2.1	Overview . . . . .	18
5.2.2	Dealing with Forgeries . . . . .	18
5.2.3	Alternative Decryption Algorithm . . . . .	18
5.2.4	Reduction to IND-CPA Security . . . . .	20
5.2.5	Overall Proof . . . . .	21
5.3	LECSS for the CDMW Construction . . . . .	21
<b>6</b>	<b>A General Indistinguishability Paradigm</b>	<b>21</b>
<b>A</b>	<b>Relating IND-SDA and NM-CPA</b>	<b>25</b>
<b>B</b>	<b>Necessity of Codes with Secret State</b>	<b>28</b>
<b>C</b>	<b>Composing Non-Malleable Codes, PKE, and MACs</b>	<b>32</b>
C.1	Replayable NM-SDA Security . . . . .	32
C.2	Non-Malleable Codes and PKE . . . . .	33
C.3	From Replayable to Full NM-SDA Security . . . . .	35
<b>D</b>	<b>Miscellaneous</b>	<b>37</b>
D.1	Chernoff Bound . . . . .	37
D.2	Plotkin Bound . . . . .	37

## 1 Introduction

Several different security notions for public-key encryption (PKE) have been proposed. The most basic one is that of indistinguishability under chosen-plaintext attacks (IND-CPA) [21], which requires that an adversary with no decryption capabilities be unable to distinguish between the encryption of two messages. Although extremely important and useful for a number of applications, in many cases IND-CPA security is not sufficient. For example, consider the simple setting of an electronic auction, where the auctioneer  $U$  publishes a public key  $\text{pk}$ , and invites several participants  $P_1, \dots, P_q$  to encrypt their bids  $b_i$  under  $\text{pk}$ . As was observed in the seminal paper of Dolev *et al.* [15], although IND-CPA security of encryption ensures that  $P_1$  cannot decrypt a bid of  $P_2$  under the ciphertext  $e_2$ , it leaves open the possibility that  $P_1$  can construct a special ciphertext  $e_1$  which decrypts to a *related* bid  $b_1$  (e.g.,  $b_1 = b_2 + 1$ ). Hence, to overcome such “malleability” problems, stronger forms of security are required.

The strongest such level of PKE security is indistinguishability under chosen-ciphertext attacks (IND-CCA), where the adversary is given unrestricted, adaptive access to a decryption oracle (modulo not being able to ask on the “challenge ciphertext”). This notion is sufficient for most natural applications of PKE, and several generic [15, 28, 31, 5, 25] and concrete [13, 14, 24, 22] constructions of IND-CCA secure encryption schemes are known by now. Unfortunately, all these constructions either rely on specific number-theoretic assumptions, or use much more advanced machinery (such as non-interactive zero-knowledge proofs or identity-based encryption) than IND-CPA secure encryption. Indeed, despite numerous efforts (e.g., a partial negative result [20]), the relationship between IND-CPA and IND-CCA security remains unresolved until now. This motivates the study of various “middle-ground” security notions between IND-CPA and IND-CCA, which are sufficient for applications, and, yet, might be constructed from simpler basic primitives (e.g., any IND-CPA encryption).

One such influential notion is non-malleability under chosen-plaintext attacks (NM-CPA), originally introduced by Dolev *et al.* [15] with the goal of precisely addressing the auction example above, by demanding that an adversary not be able to maul ciphertexts to other ciphertexts encrypting related plaintexts. As was later shown by Bellare and Sahai [4] and by Pass *et al.* [30], NM-CPA is equivalent to security against adversaries with access to a *non-adaptive* decryption oracle, meaning that the adversary can only ask one “parallel” decryption query. Although NM-CPA appears much closer to IND-CCA than IND-CPA security, a seminal result by Pass *et al.* [29] showed that one can generically build NM-CPA encryption from any IND-CPA-secure scheme, and Choi *et al.* [9] later proved that this transformation can also be achieved via a black-box construction. Thus, NM-CPA schemes can be potentially based on weaker assumptions than IND-CCA schemes, and yet suffice for important applications.

OUR WORK. In this paper we investigate three questions related to NM-CPA security:

1. Can the efficiency of the construction by Choi *et al.* of NM-CPA from IND-CPA be improved?
2. Is it possible to achieve multi-bit NM-CPA security more efficiently from a single-bit NM-CPA scheme than from IND-CPA?
3. Is there a notion stronger than NM-CPA that has natural applications and can be achieved from IND-CPA security?

We answer all three questions positively. We start with Question 3, as it will also allow us to achieve stronger answers for Questions 1 and 2. In a recent paper, Coretti *et al.* [10] introduced a new middle-ground security notion for encryption—termed indistinguishability under (chosen-ciphertext) self-destruct attacks (IND-SDA) in this paper<sup>1</sup>—where the adversary gets access to an *adaptive* decryption oracle, which, however, stops decrypting after the first *invalid* ciphertext is submitted. Applying this notion to the auction example above, it means that the auctioneer can reuse the secret key for subsequent auctions, as long as all the encrypted bids are valid. Unfortunately, if an invalid ciphertext is submitted, even the results of the *current* auction should be discarded, as IND-SDA security is not powerful enough to argue that the decryptions of the remaining ciphertexts are unrelated w.r.t. prior plaintexts.

---

<sup>1</sup>The original name used in [10] is self-destruct chosen-ciphertext attacks security.

Motivated by the above, we introduce a new security notion that we dub *non-malleability under (chosen-ciphertext) self-destruct attacks* (NM-SDA). This notion (see Definition 3) naturally combines NM-CPA and IND-SDA, by allowing the adversary to ask many adaptive “parallel” decryption queries (i.e., a query consists of many ciphertexts) up to the point when the first invalid ciphertext is submitted. In such a case, the whole parallel decryption query containing an invalid ciphertext is still answered in full, but no future decryption queries are allowed. By being stronger (as we show below) than both NM-CPA and IND-SDA, NM-SDA security appears to be a strongest natural PKE security notion that is still weaker (as we give evidence below) than IND-CCA—together with  $q$ -bounded CCA-secure PKE [12], to which it seems incomparable. In particular, it seems to apply better to the auction example above: First, unlike with basic NM-CPA, the auctioneer can reuse the same public key  $\text{pk}$ , provided no invalid ciphertexts were submitted. Second, unlike IND-SDA, the current auction can be safely completed, even if some ciphertexts are invalid. Compared to IND-CCA, however, the auctioneer will still have to change its public key for *subsequent* auctions if some of the ciphertexts are invalid. Still, one can envision situations where parties are penalized for submitting such malformed ciphertexts, in which case NM-SDA security might be practically sufficient, leading to an implementation under (potentially) lesser computational assumptions as compared to using a full-blown IND-CCA PKE.

Having introduced and motivated NM-SDA security, we provide a comprehensive study of this notion, and its relationship to other PKE security notions. First, we observe that the prior notions of NM-CPA and IND-SDA are incomparable, meaning that there are (albeit contrived) schemes that satisfy the former but not the latter notion and vice versa (cf. Theorem 22 in Appendix A). This also implies that our notion of NM-SDA security is strictly stronger than either of the two other notions.

Next, we turn to Question 2 above and answer it affirmatively even for our *stronger* notion of NM-SDA security; indeed, our security proof is easily seen to carry over to the simpler case of NM-CPA security. Finally, we also *simultaneously* answer Questions 1 and 3, by presenting a generalization of the Choi *et al.* [9] construction from IND-CPA encryption which: (a) allows us to improve the plaintext-length to ciphertext-length rate by a factor linear in the security parameter as compared to the construction of [9] (which is a special case of our abstraction, but with sub-optimal parameters); (b) generically achieves NM-SDA security (with or without the efficiency improvement). We detail these results below.

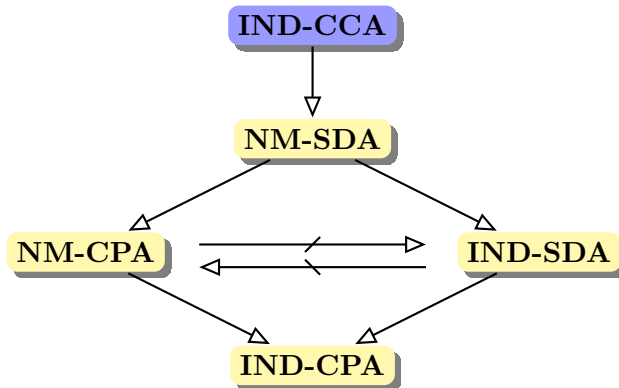
DOMAIN EXTENSION. For several security notions in public-key cryptography, it is known that single-bit public-key encryption implies multi-bit public-key encryption. For IND-CPA, this question is simple [21], since the parallel repetition of a single-bit scheme (i.e., encrypting every bit of a message separately) yields an IND-CPA secure multi-bit scheme. For the other notions considered in this paper, i.e., for NM-CPA, IND-SDA, and NM-SDA, as well as for IND-CCA, the parallel repetition (even using independent public keys) is not a scheme that achieves the same security level as the underlying single-bit scheme. However, Coretti *et al.* [10] provide a single-to-multi-bit transformation for IND-SDA security based on non-malleable codes [17] (see below), and Myers and Shelat [27], as well as Hohenberger *et al.* [23], provide (much) more complicated such transformations for IND-CCA security. To complement these works, we answer the question of domain extension for NM-SDA and NM-CPA in the affirmative. In particular we show the following result:

**Theorem 1** (Informal). *Let  $\lambda$  be the security parameter. Then there is a black-box construction of a  $\lambda$ -bit NM-SDA (resp. NM-CPA) PKE scheme from a single-bit NM-SDA (resp. NM-CPA) PKE scheme, making  $\mathcal{O}(\lambda)$  calls to the underlying single-bit scheme.<sup>2</sup>*

The proof of Theorem 1 can be found in Section 4. Our approach follows that for IND-SDA [10] and combines single-bit PKE with so-called *non-malleable codes (NMCs)*, introduced by Dziembowski *et al.* [17]. Intuitively, NMCs protect encoded messages against a tampering adversary, which tampers with the codeword by means of applying functions  $f$  from a particular function class  $\mathcal{F}$  to it, in the sense that the decoding results in either the original message or a completely unrelated value.

---

<sup>2</sup>For longer than  $\lambda$ -bit messages, one can also use standard hybrid encryption.



**Figure 1:** Diagram of the main relationships between the security notions considered in this paper.  $X \rightarrow Y$  means that  $X$  implies  $Y$ ;  $X \dashrightarrow Y$  indicates a separation between  $X$  and  $Y$ . Notions with the same color are equivalent under black-box transformations; notions with different colors are not known to be equivalent.

Our construction has the following simple structure (see also Figure 4): The plaintext  $m$  is first encoded using an appropriate non-malleable code into an encoding  $c$ , which is in turn encrypted bit-by-bit (under independent public keys) with the single-bit NM-SDA scheme.<sup>3</sup> The fact that NM-SDA security guarantees that an attacker can either leave a ciphertext intact or replace it, which results in an unrelated message, translates to the following capability of an adversary w.r.t. decryption queries: It can either leave a particular bit of the encoding unchanged, or fix it to 0 or to 1. Therefore, the tamper class against which the non-malleable code must be resilient is the class  $\mathcal{F}_{\text{set}}$  of functions that tamper with each bit of an encoding individually and can either leave it unchanged or set it to a fixed value.

The main new challenge for our construction is to deal with the *parallel* decryption queries: in order for the combined scheme to be NM-SDA secure, the NMC needs to be resilient against parallel tamper queries as well. Unfortunately, we show that no standard non-malleable code (as originally defined by Dziembowski *et al.* [17] and Faust *et al.* [18]) can achieve this notion (see Section B). Fortunately, we observe that the NMC concept can be extended to allow the decoder to make use of (an initially generated) secret state, which simply becomes part of the secret key in the combined scheme. This modification of NMCs—called secret-state NMCs—allows us to achieve resilience against parallel tampering and may be of independent interest. This reduces our question to building a secret-state non-malleable code resilient against continuous parallel tampering attacks from  $\mathcal{F}_{\text{set}}$ . We construct such a code in Section 4.3, by combining the notion of linear error-correcting secret sharing (see [17]) with the idea of a secret “trigger set” [9]. This construction forms one of the main technical contributions of our work.

NM-SDA FROM IND-CPA. Next, we show:

**Theorem 2 (Informal).** *There exists a black-box construction of an NM-SDA-secure PKE scheme from an IND-CPA-secure PKE.*

Hence, the notions of IND-CPA, NM-CPA, IND-SDA, and NM-SDA security are all equivalent, lying (plausibly, strictly?) below IND-CCA security. See Figure 1.

The proof of Theorem 2 appears in Section 5. In fact, we show that a generalization of the construction by Choi *et al.* already achieves NM-SDA security (rather than only NM-CPA security). Our proof much follows the pattern of the original one, except for one key step in the proof, where a brand new proof technique is required. Intuitively, we need to argue that no sensitive information about the secret “trigger set” is leaked to the adversary, unless one of the ciphertexts is invalid. This rather general technique (for analyzing security of so called “parallel stateless self-destruct games”) may be interesting in its own right (e.g., it is also used in the security proof of our non-malleable code in Section 4), and is detailed in Section 6.

<sup>3</sup>Technically, this scheme only achieves a slight relaxation of NM-SDA security, called *replayable* NM-SDA security, but the latter can be easily transformed into the former.

Along the way, we also manage to slightly abstract the transformation of [9], and to re-phrase it in terms of certain linear error-correcting secret-sharing schemes (LECSSs) satisfying a special property (as opposed to using Reed-Solomon codes directly as an example of such a scheme). Aside from a more modular presentation (which gives a more intuitive explanation for the elegant scheme of Choi *et al.* [9]), this also allows us to instantiate the required LECSs more efficiently and thereby improve the rate of the transformation of [9] by a factor linear in the security parameter (while also arguing NM-SDA, instead of NM-CPA, security), giving us the positive answer to Question 1.<sup>4</sup>

## 2 Preliminaries

This section introduces notational conventions and basic concepts that we use throughout the work.

**BITS AND SYMBOLS.** Let  $\ell \in \mathbb{N}$ . For any multiple  $m = t\ell$  of  $\ell$ , an  $m$ -bit string  $x = (x[1], \dots, x[m]) = (x_1, \dots, x_t)$  can be seen as composed of its *bits*  $x[j]$  or its *symbols*  $x_i \in \{0, 1\}^\ell$ . For two  $m$ -bit strings  $x$  and  $y$ , denote by  $d_{\text{H}}(x, y)$  their hamming distance as the number of *symbols* in which they differ.

**ORACLE ALGORITHMS.** Oracle algorithms are algorithms that can make special oracle calls. An algorithm  $A$  with an oracle  $O$  is denoted by  $A(O)$ . Note that oracle algorithms may make calls to other oracle algorithms (e.g.,  $A(B(O))$ ).

**DISTINGUISHERS AND REDUCTIONS.** A *distinguisher* is an (possibly randomized) oracle algorithm  $D(\cdot)$  that outputs a single bit. The distinguishing advantage on two (possibly stateful) oracles  $S$  and  $T$  is defined by

$$\Delta^D(S, T) := |\mathbb{P}[D(S) = 1] - \mathbb{P}[D(T) = 1]|,$$

where the probabilities are taken over the randomness of  $D$  as well as  $S$  and  $T$ , respectively.

Reductions between distinguishing problems are modeled as oracle algorithms as well. Specifically, when reducing distinguishing two oracles  $U$  and  $V$  to distinguishing  $S$  and  $T$ , one exhibits an oracle algorithm  $R(\cdot)$  such that  $R(U)$  behaves as  $S$  and  $R(V)$  as  $T$ ; then,  $\Delta^D(S, T) = \Delta^D(R(U), R(V)) = \Delta^{D(R(\cdot))}(U, V)$ .

**LINEAR ERROR-CORRECTING SECRET SHARING.** The following notion of a linear error-correcting secret sharing, introduced by Dziembowski *et al.* [17], is used in several places in this paper.

**Definition 1** (Linear error-correcting sharing scheme). *Let  $n \in \mathbb{N}$  be a security parameter and  $\mathbb{F}$  a field of size  $L = 2^\ell$  for some  $\ell \in \mathbb{N}$ . A  $(k, n, \delta, \tau)$  linear error-correcting secret sharing (LECSS) over  $\mathbb{F}$  is a pair of algorithms  $(\mathbf{E}, \mathbf{D})$ , where  $\mathbf{E} : \mathbb{F}^k \rightarrow \mathbb{F}^n$  is randomized and  $\mathbf{D} : \mathbb{F}^n \times \mathbb{N} \rightarrow \mathbb{F}^k \cup \{\perp\}$  is deterministic, with the following properties:*

- **Linearity:** For any vectors  $w$  output by  $\mathbf{E}$  and any  $c \in \mathbb{F}^n$ ,

$$\mathbf{D}(w + c) = \begin{cases} \perp & \text{if } \mathbf{D}(c) = \perp, \text{ and} \\ \mathbf{D}(w) + \mathbf{D}(c) & \text{otherwise.} \end{cases}$$

- **Minimum distance:** For any two codewords  $w, w'$  output by  $\mathbf{E}$ ,  $d_{\text{H}}(w, w') \geq \delta n$ .
- **Error correction:** It is possible to efficiently correct up to  $\delta n/2$  errors, i.e., for any  $x \in \mathbb{F}^k$  and any  $w$  output by  $\mathbf{E}(x)$ , if  $d_{\text{H}}(c, w) \leq t$  for some  $c \in \mathbb{F}^n$  and  $t < \delta n/2$ , then  $\mathbf{D}(c, t) = x$ .
- **Secrecy:** The symbols of a codeword are individually uniform over  $\mathbb{F}$  and are  $\tau n$ -wise independent (over the randomness of  $\mathbf{E}$ ).

---

<sup>4</sup>Note that Choi *et al.* [9] consider the ciphertext blow-up between the underlying IND-CPA scheme and the resulting scheme as quality measure of their construction, while we consider the rate (number of plaintext bits per ciphertext bit) of the resulting scheme.

This paper considers various instantiations of LECSSs, which are described in Sections 4.5 and 5.3, where they are used.

**ONE-TIME SIGNATURES.** A *digital signature scheme* (DSS) is a triple of algorithms  $\Sigma = (KG, S, V)$ , where the key-generation algorithm  $KG$  outputs a key pair  $(sk, vk)$ , the (probabilistic) signing algorithm  $S$  takes a message  $m$  and a signing key  $sk$  and outputs a signature  $s \leftarrow S_{sk}(m)$ , and the verification algorithm takes a verification key  $vk$ , a message  $m$ , and a signature  $s$  and outputs a single bit  $V_{vk}(m, s)$ . A (*strong*) *one-time signature (OTS) scheme* is a digital signature scheme that is secure as long as an adversary only observes a single signature. More precisely, OTS security is defined using the following game  $G^{\Sigma, \text{ots}}$  played by an adversary  $A$ : Initially, the game generates a key pair  $(sk, vk)$  and hands the verification key  $vk$  to  $A$ . Then,  $A$  can specify a single message  $m$  for which he obtains a signature  $s \leftarrow S_{vk}(m)$ . Then, the adversary outputs a pair  $(m', s')$ . The adversary wins the game if  $(m', s') \neq (m, s)$  and  $V_{vk}(m', s') = 1$ . The *advantage* of  $A$  is the probability (over all involved randomness) that  $A$  wins the game, and is denoted by  $\Gamma^A(G^{\Sigma, \text{ots}})$ .

**Definition 2.** A DSS scheme  $\Sigma$  is a  $(t, \varepsilon)$ -strong one-time signature scheme if for all adversaries  $A$  with running time at most  $t$ ,  $\Gamma^A(G^{\Sigma, \text{ots}}) \leq \varepsilon$ .

### 3 Non-Malleability under Self-Destruct Attacks

A *public-key encryption (PKE) scheme* with message space  $\mathcal{M} \subseteq \{0, 1\}^*$  and ciphertext space  $\mathcal{C}$  is defined as three algorithms  $\Pi = (KG, E, D)$ , where the key-generation algorithm  $KG$  outputs a key pair  $(pk, sk)$ , the (probabilistic) encryption algorithm  $E$  takes a message  $m \in \mathcal{M}$  and a public key  $pk$  and outputs a ciphertext  $e \leftarrow E_{pk}(m)$ , and the decryption algorithm takes a ciphertext  $e \in \mathcal{C}$  and a secret key  $sk$  and outputs a plaintext  $m \leftarrow D_{sk}(e)$ . The output of the decryption algorithm can be the special symbol  $\perp$ , indicating an invalid ciphertext. A PKE scheme is correct if  $m = D_{sk}(E_{pk}(m))$  (with probability 1 over the randomness in the encryption algorithm) for all messages  $m$  and all key pairs  $(pk, sk)$  generated by  $KG$ .

Security notions for PKE schemes in this paper are formalized using the distinguishing game  $G_b^{\Pi, q, p}$ , depicted in Figure 2: The distinguisher (adversary) is initially given a public key and then specifies two messages  $m_0$  and  $m_1$ . One of these, namely  $m_b$ , is encrypted and the adversary is given the resulting challenge ciphertext. During the entire game, the distinguisher has access to a decryption oracle that allows him to make at most  $q$  decryption queries, each consisting of at most  $p$  ciphertexts. Once the distinguisher specifies an invalid ciphertext, the decryption oracle self-destructs, i.e., no additional decryption queries are answered.

The general case is obtained when both  $q$  and  $p$  are arbitrary (denoted by  $q = p = *$ ), which leads to our main definition of non-malleability under (chosen-ciphertext) self-destruct attacks (NM-SDA). For readability, set  $G_b^{\Pi, \text{nm-sda}} := G_b^{\Pi, *, *}$  for  $b \in \{0, 1\}$ . Formally, NM-SDA is defined as follows:

**Definition 3** (Non-malleability under self-destruct attacks). A PKE scheme  $\Pi$  is  $(t, q, p, \varepsilon)$ -NM-SDA-secure if for all distinguishers  $D$  with running time at most  $t$  and making at most  $q$  decryption queries of size at most  $p$  each,

$$\Delta^D(G_0^{\Pi, \text{nm-sda}}, G_1^{\Pi, \text{nm-sda}}) \leq \varepsilon.$$

All other relevant security notions in this paper can be derived as special cases of the above definition, by setting the parameters  $q$  and  $p$  to different values.

**CHOSEN-PLAINTEXT SECURITY (IND-CPA).** In this variant, the distinguisher is not given access to a decryption oracle, i.e.,  $q = p = 0$ . For readability, set  $G_b^{\Pi, \text{ind-cpa}} := G_b^{\Pi, 0, 0}$  for  $b \in \{0, 1\}$  in the remainder of this paper. We say that  $\Pi$  is  $(t, \varepsilon)$ -IND-CPA-secure if it is, in fact,  $(t, 0, 0, \varepsilon)$ -NM-SDA-secure.

**NON-MALLEABILITY (NM-CPA).** A scheme is non-malleable under chosen-plaintext attacks [29] (NM-CPA), if the adversary can make a single decryption query consisting of arbitrarily many ciphertexts, i.e.,

Distinguishing Game $G_b^{\Pi, q, p}$	
<pre> <b>init</b>   ctr ← 0   (pk, sk) ← KG   <b>output</b> pk  <b>on</b> (chall, m<sub>0</sub>, m<sub>1</sub>) with  m<sub>0</sub>  =  m<sub>1</sub>    e ← E<sub>pk</sub>(m<sub>b</sub>)   <b>output</b> e </pre>	<pre> <b>on</b> (dec, e<sup>(1)</sup>, ..., e<sup>(p)</sup>)   ctr ← ctr + 1   <b>for</b> j ← 1 <b>to</b> p     m<sup>(j)</sup> ← D<sub>sk</sub>(e<sup>(j)</sup>)     <b>if</b> e<sup>(j)</sup> = e       m<sup>(j)</sup> ← test   <b>output</b> (m<sup>(1)</sup>, ..., m<sup>(p)</sup>)   <b>if</b> ∃j : m<sup>(j)</sup> = ⊥ <b>or</b> ctr ≥ q     <b>self-destruct</b> </pre>

**Figure 2:** Distinguishing game  $G_b^{\Pi, q, p}$ , where  $b \in \{0, 1\}$ , used to define security of a PKE scheme  $\Pi = (KG, E, D)$ . The numbers  $q, p \in \mathbb{N}$  specify the maximum number of decryption queries and their size, respectively. The command **self-destruct** results in all future decryption queries being answered by  $\perp$ .

$q = 1$  and  $p$  arbitrary (denoted by  $p = *$ ). Similarly to above, set  $G_b^{\Pi, \text{nm-cpa}} := G_b^{\Pi, 1, *}$  for  $b \in \{0, 1\}$ . We say that  $\Pi$  is  $(t, p, \varepsilon)$ -NM-CPA-secure if it is, in fact,  $(t, 1, p, \varepsilon)$ -NM-SDA-secure.<sup>5</sup>

INDISTINGUISHABILITY UNDER SELF-DESTRUCT ATTACKS (IND-SDA). This variant, introduced in [10], allows arbitrarily many queries to the decryption oracle, but each of them may consist of a single ciphertext only, i.e.,  $q$  arbitrary (denoted by  $q = *$ ) and  $p = 1$ . Once more, set  $G_b^{\Pi, \text{ind-sda}} := G_b^{\Pi, *, 1}$ . We say that  $\Pi$  is  $(t, q, \varepsilon)$ -IND-SDA-secure if it is, in fact,  $(t, q, 1, \varepsilon)$ -NM-SDA-secure.

CHOSEN-CIPHERTEXT SECURITY (IND-CCA). The standard notion of IND-CCA security can be obtained as a strengthening of NM-SDA where  $q = *$ ,  $p = 1$ , and the decryption oracle never self-destructs. We do not define this notion formally, as it is not the main focus of this paper.

ASYMPTOTIC FORMULATION. To allow for concise statements, sometimes we prefer to use an asymptotic formulation instead of stating concrete parameters. More precisely, we will say that a PKE scheme  $\Pi$  is  $X$ -secure (for  $X \in \{\text{IND-CPA}, \text{NM-CPA}, \text{IND-SDA}, \text{NM-SDA}\}$ ) if for all efficient adversaries the advantage  $\varepsilon$  in the corresponding distinguishing game is negligible in the security parameter.

NON-MALLEABLE CPA VS. INDISTINGUISHABLE SDA. We provide a separation between the notions of NM-CPA and IND-SDA security; a corresponding theorem and proof can be found in Appendix A. Given such a separation, our notion of NM-SDA security (see Definition 3) is strictly stronger than either of the two other notions.

## 4 Domain Extension

This section contains one of our main technical results. We show how single-bit NM-SDA PKE can be combined with so-called *secret-state non-malleable codes* resilient against *continuous parallel tampering*, which we believe is an interesting notion in its own right, to achieve multi-bit NM-SDA-secure PKE. We construct such a code and prove its security, and, additionally, we show that no code without secret state can achieve security against parallel tampering unconditionally (see Theorem 14 in Appendix B).<sup>6</sup>

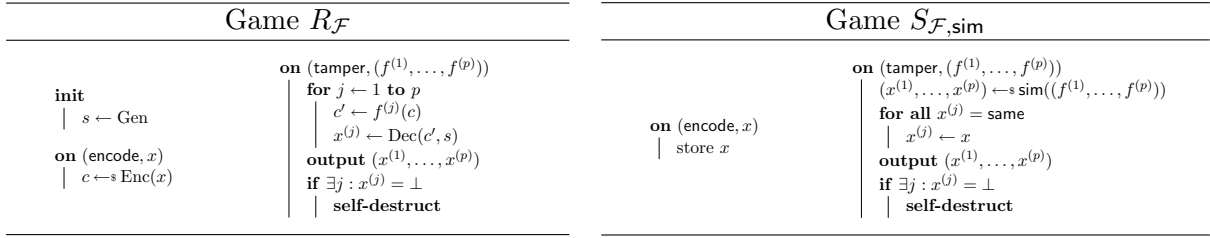
### 4.1 A New Flavor of Non-Malleable Codes

Non-malleable codes were introduced by Dziembowski *et al.* [17]. Intuitively, they protect encoded messages in such a way that any tampering with the codeword causes the decoding to either output the original message or a completely unrelated value. The original notion can be extended to include the aforementioned secret state in the decoder as follows:

<sup>5</sup>Note that the way NM-CPA is defined here is slightly stronger than the normal notion. This is due to the adversary's ability to ask a parallel decryption query at any time—as opposed to only after receiving the challenge ciphertext in earlier definitions (cf., e.g., [29]).

<sup>6</sup>The question whether the notion is achievable by a computationally-secure code remains open for future work.





**Figure 3:** Distinguishing game  $(R_{\mathcal{F}}, S_{\mathcal{F},\text{sim}})$  used to define non-malleability of a secret-state coding scheme  $(\text{Gen}, \text{Enc}, \text{Dec})$ . The command **self-destruct** has the effect that all future queries are answered by  $\perp$ .

**Definition 4** (Code with secret state). A  $(k, n)$ -code with secret state (CSS) is a triple of algorithms  $(\text{Gen}, \text{Enc}, \text{Dec})$ , where the (randomized) state-generation algorithm  $\text{Gen}$  outputs a secret state  $s$  from some set  $\mathcal{S}$ , the (randomized) encoding algorithm  $\text{Enc}$  takes a  $k$ -bit plaintext  $x$  and outputs an  $n$ -bit encoding  $c \leftarrow \text{Enc}(x)$ , and the (deterministic) decoding algorithm  $\text{Dec}$  takes an encoding as well as some secret state  $s \in \mathcal{S}$  and outputs a plaintext  $x \leftarrow \text{Dec}(c, s)$  or the special symbol  $\perp$ , indicating an invalid encoding.

Tampering attacks are captured by functions  $f$ , from a certain function class  $\mathcal{F}$ , that are applied to an encoding. The original definition by [17] allows an attacker to apply only a single tamper function. In order to capture continuous parallel attacks, the definition below permits the attacker to repeatedly specify parallel tamper queries, each consisting of several tamper functions. The process ends as soon as one of the tamper queries leads to an invalid codeword.

The non-malleability requirement is captured by considering a real and an ideal experiment. In both experiments, an attacker is allowed to encode a message of his choice. In the real experiment, he may tamper with an actual encoding of that message, whereas in the ideal experiment, the tamper queries are answered by a (stateful) simulator. The simulator is allowed to output the special symbol **same**, which the experiment replaces by the originally encoded message. In either experiment, if a component of the answer vector to a parallel tamper query is the symbol  $\perp$ , a self-destruct occurs, i.e., all *future* tamper queries are answered by  $\perp$ . The experiments are depicted in Figure 3.

**Definition 5** (Non-malleable code with secret state). Let  $q, p \in \mathbb{N}$  and  $\varepsilon > 0$ . A CSS  $(\text{Gen}, \text{Enc}, \text{Dec})$  is  $(\mathcal{F}, q, p, \varepsilon)$ -non-malleable if the following properties are satisfied:

- Correctness: For each  $x \in \{0, 1\}^k$  and all  $s \in \mathcal{S}$  output by  $\text{Gen}$ ,  $\text{Dec}(\text{Enc}(x), s) = x$  with probability 1 over the randomness of  $\text{Enc}$ .
- Non-Malleability: There exists a (possibly stateful) simulator  $\text{sim}$  such that for any distinguisher  $D$  asking at most  $q$  parallel queries, each of size at most  $p$ ,  $\Delta^D(R_{\mathcal{F}}, S_{\mathcal{F},\text{sim}}) \leq \varepsilon$ .

We remark that for codes without secret state (as the ones considered in [17]), one obtains the standard notion of non-malleability [17] by setting  $q = p = 1$ , and continuous non-malleability [18] by letting  $p = 1$  and  $q$  arbitrary (i.e.,  $q = *$ ).

## 4.2 Combining Single-bit PKE and Non-Malleable Codes

Our construction of a multi-bit NM-SDA-secure PKE scheme  $\Pi'$  from a single-bit NM-SDA-secure scheme  $\Pi$  and a secret-state non-malleable  $(k, n)$ -code follows the approach of [10]: It encrypts a  $k$ -bit message  $m$  by first computing an encoding  $c = (c[1], \dots, c[n])$  of  $m$  and then encrypting each bit  $c[j]$  under an independent public key of  $\Pi$ ; it decrypts by first decrypting the individual components and then decoding the resulting codeword using the secret state of the non-malleable code; the secret state is part of the secret key. The scheme is depicted in detail in Figure 4.

PKE Scheme $\Pi' = (KG', E', D')$		
Key Generation $KG'$ <b>for</b> $i \leftarrow 1$ <b>to</b> $n$   $(pk_i, sk_i) \leftarrow_s KG$ <b>pk</b> $\leftarrow (pk_1, \dots, pk_n)$ <b>sk</b> $\leftarrow (sk_1, \dots, sk_n)$ <b>s</b> $\leftarrow$ Gen <b>return</b> (pk, (sk, s))	Encryption $E'_{pk}(m)$ $c = (c[1], \dots, c[n]) \leftarrow$ Enc( $m$ ) <b>for</b> $i \leftarrow 1$ <b>to</b> $n$   $e_i \leftarrow_s E_{pk_i}(c[i])$ <b>return</b> $e = (e_1, \dots, e_n)$	Decryption $D'_{(sk,s)}(e)$ $(e_1, \dots, e_n) \leftarrow e$ <b>for</b> $i \leftarrow 1$ <b>to</b> $n$   $c[i] \leftarrow_s D_{sk_i}(e_i)$   <b>if</b> $c[i] = \perp$     <b>return</b> $\perp$ <b>m</b> $\leftarrow$ Dec( $c[1] \dots c[n], s$ ) <b>return</b> $m$

**Figure 4:** The  $k$ -bit PKE scheme  $\Pi' = (KG', E', D')$  built from a 1-bit PKE scheme  $\Pi = (KG, E, D)$  and a  $(k, n)$ -coding scheme with secret state (Gen, Enc, Dec).

Intuitively, NM-SDA security (or CCA security in general) guarantees that an attacker can either leave a message intact or replace it by an independently created one. For our construction, which separately encrypts every bit of an encoding of the plaintext, this translates to the following capability of an adversary w.r.t. decryption queries: It can either leave a particular bit of the encoding unchanged or fix it to 0 or to 1. Therefore, the tamper class against which the non-malleable code must be resilient is the class  $\mathcal{F}_{\text{set}} \subseteq \{f \mid f : \{0, 1\}^n \rightarrow \{0, 1\}^n\}$  of functions that tamper with each bit of an encoding individually and can either leave it unchanged or set it to a fixed value. More formally,  $f \in \mathcal{F}_{\text{set}}$  can be characterized by  $(f[1], \dots, f[n])$ , where  $f[j] : \{0, 1\} \rightarrow \{0, 1\}$  is the action of  $f$  on the  $j^{\text{th}}$  bit and  $f[j] \in \{\text{zero}, \text{one}, \text{keep}\}$  with the meaning that it either sets the  $j^{\text{th}}$  bit to 0 (zero) or to 1 (one) or leaves it unchanged (keep).

Before stating the theorem about the security of our construction  $\Pi'$ , it needs to be pointed out that it achieves only the so-called *replayable* variant of NM-SDA security. The notion of replayable CCA (RCCA) security (in general) was introduced by Canetti *et al.* [6] to deal with the fact that for many applications (full) CCA security is unnecessarily strict. Among other things, they provide a MAC-based generic transformation of RCCA-secure schemes into CCA-secure ones, which we can also apply in our setting (as we show) to obtain a fully NM-SDA-secure scheme  $\Pi''$ .

**Theorem 3.** *Let  $q, p \in \mathbb{N}$  and  $\Pi$  be a  $(t + t_{1\text{bit}}, q, p, \varepsilon_{1\text{bit}})$ -NM-SDA-secure 1-bit PKE scheme,  $(T, V)$  a  $(t + t_{\text{mac}}, 1, qp, \varepsilon_{\text{mac}})$ -MAC, and  $(\text{Gen}, \text{Enc}, \text{Dec})$  a  $(\mathcal{F}_{\text{set}}, q, p, \varepsilon_{\text{nmc}})$ -non-malleable  $(k, n)$ -code with secret state. Then,  $\Pi''$  is  $(t, q, p, \varepsilon)$ -NM-SDA-secure PKE scheme with*

$$\varepsilon = 2(3(n\varepsilon_{1\text{bit}} + \varepsilon_{\text{nmc}}) + qp \cdot 2^{-\ell} + \varepsilon_{\text{mac}}),$$

where  $t_{1\text{bit}}$  and  $t_{\text{mac}}$  are the overheads incurred by the corresponding reductions and  $\ell$  is the length of a verification key for the MAC.

The full proof of Theorem 3 can be found in Appendix C; here we only provide a sketch. We stress that an analogous statement as the one of the above theorem works for domain extension of NM-CPA, i.e., for constructing a multi-bit NM-CPA scheme out of a single-bit NM-CPA scheme. The proof is very similar to the one of Theorem 3 and therefore omitted.

*Proof (sketch).* The proof considers a series of  $n$  hybrid experiments. In very rough terms, the  $i^{\text{th}}$  hybrid generates the challenge ciphertext by computing an encoding  $c = (c[1], \dots, c[n])$  of the challenge plaintext and by replacing the first  $i$  bits  $c[i]$  of  $c$  by random values  $\tilde{c}[i]$  before encrypting the encoding bit-wise, leading to the challenge  $(e_1^*, \dots, e_n^*)$ . Moreover, when answering decryption queries  $(e'_1, \dots, e'_n)$ , if  $e'_j = e_j^*$  for  $j \leq i$ , the  $i^{\text{th}}$  hybrid sets the outcome of  $e'_j$ 's decryption to be the corresponding bit  $c[j]$  of the original encoding  $c$ , whereas if  $e'_j \neq e_j^*$ , it decrypts normally (then it decodes the resulting  $n$ -bit string normally). This follows the above intuition that a CCA-secure PKE scheme guarantees that if a decryption query is different from the challenge ciphertext, then the plaintext contained in it must have been created

independently of the challenge plaintext. The indistinguishability of the hybrids follows from the security of the underlying single-bit scheme  $\Pi$ .

In the  $n^{\text{th}}$  hybrid, the challenge consists of  $n$  encryptions of random values. Thus, the only information about the encoding of the challenge plaintext that an attacker gets is that leaked through decryption queries. But in the  $n^{\text{th}}$  hybrid there is a 1-to-1 correspondence between decryption queries and the tamper function  $f = (f[1], \dots, f[n])$  applied to the encoding of the challenge plaintext: The case  $e'_j = e_j^*$  corresponds to  $f[j] = \text{keep}$ , and the case  $e'_j \neq e_j^*$  corresponds to  $f[j] = \text{zero}$  or  $f[j] = \text{one}$ , depending on whether  $e'_j$  decrypts to zero or to one. This allows a reduction to the security of the non-malleable code.  $\square$

### 4.3 Non-Malleable Code Construction

It remains to construct a non-malleable code (with secret state) resilient against parallel tampering. The intuition behind our construction is the following: If a code has the property (as has been the case with previous schemes secure against (non-parallel) bit-wise tampering) that changing a single bit of a valid encoding results in an invalid codeword, then the tamper function that fixes a particular bit of the encoding and leaves the remaining positions unchanged can be used to determine the value of that bit; this attack is parallelizable, and thus a code of this type cannot provide security against parallel tampering. A similar attack is also possible if the code corrects a fixed (known) number of errors. To circumvent this issue, our construction uses a—for the lack of a better word—“dynamic” error-correction bound: The secret state (which is initially chosen at random) is used to determine the positions of the encoding in which (a certain amount of) errors is tolerated.

**CONSTRUCTION.** Let  $\mathbb{F} = \text{GF}(2)$  and  $\alpha > 0$ . Let  $(\text{E}, \text{D})$  be a  $(k, n, \delta, \tau)$ -LECSS (cf. Definition 1 in Section 2) with minimum distance  $\delta$  and secrecy  $\tau$  over  $\mathbb{F}$  such that:<sup>7</sup>

- *Minimum distance:*  $\delta > 1/4 + 2\alpha$  and  $\delta/2 > 2\alpha$ .
- *Constant rate:*  $k/n = \Omega(1)$ .
- *Constant secrecy:*  $\tau = \Omega(1)$ .

In the following, we assume that  $\alpha \geq \tau$ , an assumption that can always be made by ignoring some of the secrecy. Consider the following  $(k, n)$ -code with secret state  $(\text{Gen}, \text{Enc}, \text{Dec})$ :

- **Gen:** Choose a subset  $T$  of  $[n]$  of size  $\tau n$  uniformly at random and output it.
- **Enc( $x$ )** for  $x \in \{0, 1\}^k$ : Compute  $c = \text{E}(x)$  and output it.
- **Dec( $c, T$ )** for  $c \in \{0, 1\}^n$ : Find a codeword  $w = (w[1], \dots, w[n])$  with  $d_{\text{H}}(w, c) \leq \alpha n$ . If no such  $w$  exists, output  $\perp$ . Moreover, if  $w[j] \neq c[j]$  for some  $j \in T$ , output  $\perp$  as well. Otherwise, decode  $w$  to its corresponding plaintext  $x$  and output it.

We prove the following theorem:

**Theorem 4.** *For all  $q, p \in \mathbb{N}$ ,  $(k, n)$ -code  $(\text{Gen}, \text{Enc}, \text{Dec})$  based on a  $(k, n, \delta, \tau)$ -LECSS satisfying the three conditions above is  $(\mathcal{F}_{\text{set}}, q, p, \varepsilon_{\text{nmc}})$ -non-malleable with*

$$\varepsilon_{\text{nmc}} = p(\mathcal{O}(1) \cdot e^{-\tau n/16} + e^{-\tau^2 n/4}) + p e^{-\tau^2 n}.$$

**INSTANTIATING THE CONSTRUCTION.** Section 4.5 details how a LECSS satisfying the above properties can be constructed by combining high-distance binary codes with a recent result by Cramer *et al.* [11] in order to “add” secrecy. The resulting LECSS has secrecy  $\tau = \Omega(1)$  and rate  $\rho = \Omega(1)$  (cf. Corollary 13 in Section 4.5). The secrecy property depends on the random choice of a universal hash function. Thus, the

---

<sup>7</sup>The reasons for these restrictions become apparent in the proof; of course,  $\alpha$  must be chosen small enough in order for these constraints to be satisfiable.

instantiated code can be seen as a construction in the CRS model. When combined with the single-bit PKE as described above, the description of the hash function can be made part of the public key.

By combining Theorem 3, Theorem 4, and Corollary 13, we obtain a 1-to- $k$ -bit black-box domain extension for NM-SDA (and NM-CPA) making  $\mathcal{O}(k)$  calls to the underlying 1-bit scheme, therefore establishing Theorem 1.<sup>8</sup>

#### 4.4 Proof of the Non-Malleable Code Construction

For the proof of Theorem 4, fix  $q, p \in \mathbb{N}$  and a distinguisher  $D$  making at most  $q$  tamper queries of size  $p$  each. Set  $\mathcal{F} := \mathcal{F}_{\text{set}}$  for the rest of the proof. In the following, we assume that  $\alpha \geq \tau$ , an assumption that can always be made by ignoring some of the secrecy. The goal is to show

$$\Delta^D(R_{\mathcal{F}}, S_{\mathcal{F}, \text{sim}}) \leq \varepsilon_{\text{nmcc}} = p(\mathcal{O}(1) \cdot e^{-\tau n/16} + e^{-\tau^2 n/4}) + pe^{-\tau^2 n}$$

for a simulator  $\text{sim}$  to be determined.

On a high level, the proof proceeds as follows: First, it shows that queries that interfere with too many bits of an encoding and at the same time do not fix enough bits (called *middle* queries below) are rejected with high probability. For the remaining query types (called *low* and *high* queries), one can show that their effect on the decoding process can always be determined from the query itself and the bits of the encoding at the positions indexed by the secret trigger set  $T$ . Since the size of  $T$  is  $\tau n$ , these symbols are uniformly random and independent of the encoded message, which immediately implies a simulation strategy for  $\text{sim}$ .

##### 4.4.1 Tamper-Query Types

Recall that  $f \in \mathcal{F}_{\text{set}}$  can be characterized by  $(f[1], \dots, f[n])$ , where  $f[j] : \{0, 1\} \rightarrow \{0, 1\}$  is the action of  $f$  on the  $j^{\text{th}}$  bit, for  $f[j] \in \{\text{zero}, \text{one}, \text{keep}\}$ , with the meaning that it either sets the  $j^{\text{th}}$  bit to 0 (*zero*) or to 1 (*one*) or leaves it unchanged (*keep*). Define  $A(f)$  to be the set of all indices  $j$  such that  $f[j] \in \{\text{zero}, \text{one}\}$ , and let  $q(f) := |A(f)|$ . Moreover, let  $\text{val}(\text{zero}) := 0$  and  $\text{val}(\text{one}) := 1$ .

A tamper query  $f$  is a *low query* if  $q(f) \leq \tau n$ , a *middle query* if  $\tau n < q(f) < (1 - \tau)n$ , and a *high query* if  $q(f) \geq (1 - \tau)n$ .

##### 4.4.2 Analyzing Query Types

The following lemma states that an isolated middle query is rejected with high probability.

**Lemma 5.** *Let  $f \in \mathcal{F}_{\text{set}}$  be a middle query. Then, for any  $x \in \{0, 1\}^k$ ,*

$$\mathbb{P}[\text{Dec}(f(\text{Enc}(x))) \neq \perp] \leq \mathcal{O}(1) \cdot e^{-\tau n/16} + e^{-\tau^2 n/4}$$

where the probability is over the randomness of  $\text{Enc}$  and the choice of the secret trigger set  $T$ .

*Proof.* Fix  $x \in \{0, 1\}^k$  and a middle query  $f = (f[1], \dots, f[n])$ . Suppose first that  $q(f) \geq n/2$ . Define

$$\mathcal{W} := \{w \in \mathbb{F}^n \mid w \text{ is codeword} \wedge \exists r : d_{\mathbb{H}}(f(\mathbf{E}(x; r)), w) \leq \alpha n\},$$

where  $r$  is the randomness of  $\mathbf{E}$ . That is,  $\mathcal{W}$  is the set of all codewords that could possibly be considered while decoding an encoding of  $x$  tampered with via  $f$ . Consider two distinct codewords  $w, w' \in \mathcal{W}$ . From the definition of  $\mathcal{W}$  it is apparent that  $w[j] \neq \text{val}(f[j])$  for at most  $\alpha n$  positions  $j \in A(f)$  (and similarly for  $w'$ ), which implies that  $w$  and  $w'$  differ in at most  $2\alpha n$  positions  $j \in A(f)$ . Therefore,  $w$  and  $w'$  differ in at least  $(\delta - 2\alpha)n$  positions  $j \notin A(f)$ .

<sup>8</sup>Note that for the construction to be secure, it is necessary that  $n = \Omega(\lambda)$  and, therefore, due to the constant rate of the LECSS, the plaintext length is  $k = \Omega(\lambda)$  as well.

For  $w \in \mathcal{W}$ , let  $\tilde{w}$  be the projection of  $w$  onto the unfixed positions  $j \notin A(f)$  and set  $\tilde{\mathcal{W}} := \{\tilde{w} \mid w \in \mathcal{W}\}$ . The above distance argument implies that  $|\mathcal{W}| = |\tilde{\mathcal{W}}|$ . Moreover,  $\tilde{\mathcal{W}}$  is a binary code with block length  $n - q(f)$  and relative distance at least

$$\frac{(\delta - 2\alpha)n}{n - q(f)} \geq \frac{(\delta - 2\alpha)n}{n/2} = 2\delta - 4\alpha > 1/2,$$

where the last inequality follows from the fact that  $\delta$  and  $\alpha$  are such that  $\delta - 2\alpha > 1/4$ . Therefore, by the Plotkin bound (Theorem 42),<sup>9</sup>

$$|\mathcal{W}| = |\tilde{\mathcal{W}}| \leq \mathcal{O}(1).$$

Denote by  $c = (c[1], \dots, c[n])$  and  $\tilde{c} = (\tilde{c}[1], \dots, \tilde{c}[n])$  the (random variables corresponding to the) encoding  $c = \text{Enc}(x)$  and the tampered encoding  $\tilde{c} = f(c)$ , respectively. For an arbitrary ( $n$ -bit) codeword  $w \in \mathcal{W}$ ,

$$\mathbf{E}[d_{\text{H}}(\tilde{c}, w)] = \sum_{j=1}^n \mathbf{E}[d_{\text{H}}(\tilde{c}[j], w[j])] \geq \sum_{j \in J} \mathbf{E}[d_{\text{H}}(\tilde{c}[j], w[j])],$$

where  $J \subseteq [n]$  is the set containing the indices of the first  $\tau n$  bits *not* fixed by  $f$ . Note that by the definition of middle queries, there are at least that many, i.e.,  $|J| = \tau n$ .

Observe that for  $j \in J$ ,  $d_{\text{H}}(\tilde{c}[j], w[j])$  is an indicator variable with expectation  $\mathbf{E}[d_{\text{H}}(\tilde{c}[j], w[j])] \geq \frac{1}{2}$ , since  $c[j]$  is a uniform bit. Thus,  $\mathbf{E}[d_{\text{H}}(\tilde{c}, w)] \geq \frac{\tau n}{2}$ .

Additionally,  $(d_{\text{H}}(\tilde{c}[j], w[j]))_{j \in J}$  are independent. Therefore, using a Chernoff bound (Theorem 41), for  $\varepsilon > 0$

$$\mathbf{P}[d_{\text{H}}(\tilde{c}, w) < (1 - \varepsilon)\tau n/2] \leq e^{-\tau\varepsilon^2 n/4}.$$

Therefore, the probability that there exists  $w \in \mathcal{W}$  for which the above does not hold is at most

$$|\mathcal{W}| \cdot e^{-\tau\varepsilon^2 n/4} \leq \mathcal{O}(1) \cdot e^{-\tau\varepsilon^2 n/4},$$

by a union bound.

Suppose now that  $d_{\text{H}}(\tilde{c}, w) \geq (1 - \varepsilon)\tau n/2$  for all codewords  $w \in \mathcal{W}$ . Then, over the choice of  $T$ ,<sup>10</sup>

$$\mathbf{P}[\forall j \in T : d_{\text{H}}(\tilde{c}[j], w[j]) = 0] \leq (1 - (1 - \varepsilon)\tau/2)^{\tau n} \leq e^{-(1 - \varepsilon)\tau^2 n/2}.$$

The lemma now follows by setting  $\varepsilon := \frac{1}{2}$ .

If  $q(f) < n/2$  an analogous argument can be made for the difference  $d := c - \tilde{c}$  between the encoding and the tampered codeword, as such a query  $f$  fixes at least half of the bits of  $d$  (to 0, in fact) and  $D(d) \neq \perp$  implies  $D(\tilde{c}) \neq \perp$ .  $\square$

It turns out that low and high queries always result in  $\perp$  or one other value.

**Lemma 6.** *Low queries  $f \in \mathcal{F}_{\text{set}}$  can result only in  $\perp$  or the originally encoded message  $x \in \{0, 1\}^k$ . High queries  $f \in \mathcal{F}_{\text{set}}$  can result only in  $\perp$  or one other value  $x_f \in \{0, 1\}^k$ , which solely depends on  $f$ . Furthermore,  $x_f$ , if existent, can be found efficiently given  $f$ .*

*Proof.* The statement for low queries is trivial, since a low query  $f$  cannot change the encoding beyond the error correction bound  $\alpha n$ .

Consider now a high query  $f$  and the following efficient procedure:

1. Compute  $\tilde{c}_f \leftarrow f(0^n)$ .
2. Find a codeword  $w_f$  with  $d_{\text{H}}(w_f, \tilde{c}_f) \leq 2\alpha n$  (which is possible since  $2\alpha < \delta/2$ ).

<sup>9</sup>The size constant absorbed by  $\mathcal{O}(1)$  here depends on how close  $2\delta - 4\alpha$  is to  $1/2$ .

<sup>10</sup>Recall that  $|T| = \tau n$ .

3. Output  $w_f$  or  $\perp$  if none exists.

Consider an arbitrary encoding  $c$  and let  $\tilde{c} \leftarrow f(c)$  be the tampered encoding. Assume there exists  $w$  with  $d_{\mathbb{H}}(w, \tilde{c}) \leq \alpha n$ . Since a high query  $f$  fixes all but  $\tau n$  bits,  $d_{\mathbb{H}}(\tilde{c}, \tilde{c}_f) \leq \tau n \leq \alpha n$ , and, thus,  $d_{\mathbb{H}}(w, \tilde{c}_f) \leq 2\alpha n$ , by the triangle inequality. Hence,  $w = w_f$ .

In other words, if the decoding algorithm Dec on  $\tilde{c}$  finds a codeword  $w = w_f$ , one can find it using the above procedure, which also implies that high queries can only result in  $\perp$  or one other message  $x_f = D(w_f)$ .  $\square$

#### 4.4.3 Hybrids

HANDLING MIDDLE QUERIES. Consider the hybrid game  $H_1$  that behaves as  $R_{\mathcal{F}}$ , except that it answers all middle queries by  $\perp$ .

**Lemma 7.**  $\Delta^D(R_{\mathcal{F}}, H_1) \leq p(\mathcal{O}(1) \cdot e^{-\tau n/16} + e^{-\tau^2 n/4})$ .

The proof of Lemma 7 follows a generic paradigm, at whose core is the so-called *self-destruct lemma*, which deals with the indistinguishability of hybrids with the self-destruct property and is explained in detail in Section 6. Roughly, this lemma applies whenever the first hybrid (in this case  $R_{\mathcal{F}}$ ) can be turned into the second one (in this case  $H_1$ ) by changing (“bending”) the answers to a subset (the “bending set”) of the possible queries to always be  $\perp$ , and when additionally non-bent queries have a unique answer (cf. the statement of Lemma 21). Intuitively, the lemma states that parallelism and adaptivity do not help distinguish (much) in such cases, which allows using Lemma 5.

*Proof.* The lemma is proved conditioned on the message  $x$  encoded by  $D$ . To use the self-destruct lemma, note first that both  $R_{\mathcal{F}}$  and  $H_1$  answer parallel tamper queries in which each component is from the set  $\mathcal{X} := \mathcal{F}$  by vectors whose components are in  $\mathcal{Y} := \{0, 1\}^k \cup \{\perp\}$ . Moreover, both hybrids use as internal randomness a uniformly chosen element from  $\mathcal{R} := \{0, 1\}^{\rho} \times \mathcal{S}$ , where  $\rho$  is an upper bound on the number of random bits used by Enc and  $\mathcal{S}$  is the set of all  $\tau n$ -subsets  $T$  of  $[n]$ .  $R_{\mathcal{F}}$  answers each component of a query  $f \in \mathcal{X}$  by

$$g(f, (r, T)) := \text{Dec}(f(\text{Enc}(x; r)), T).$$

Define  $\mathcal{B} \subseteq \mathcal{X}$  to be the set of all middle queries;  $H_1$  is the  $\mathcal{B}$ -bending of  $R_{\mathcal{F}}$  (cf. Definition 7).

Observe that queries  $f \notin \mathcal{B}$  are either low or high queries. For low queries  $f$ , the unique answer is  $y_f = x$ , and for high queries  $f$ ,  $y_f = x_f$  (cf. Lemma 6). Thus, by Lemmas 21 and 5,

$$\Delta^D(R_{\mathcal{F}}, H_1) \leq p \cdot \max_{f \in \mathcal{B}} \mathbb{P}[g(f, (r, T)) \neq \perp] \leq p(\mathcal{O}(1) \cdot e^{-\tau n/16} + e^{-\tau^2 n/4}),$$

where the probability is over the choice of  $(r, T)$ .  $\square$

HANDLING HIGH QUERIES. Consider the following hybrid game  $H_2$ : It differs from  $H_1$  in the way it decodes high queries  $f$ . Instead of applying the normal decoding algorithm to the tampered codeword  $\tilde{c}$ , it proceeds as follows:

1. Find  $w_f$  (as in the proof of Lemma 6).
2. If  $w_f$  does not exist, return  $\perp$ .
3. If  $\tilde{c}[j] = w_f[j]$  for all  $j \in T$ , return Dec( $w$ ). Otherwise, return  $\perp$ .

**Lemma 8.**  $\Delta^D(H_1, H_2) \leq p e^{-\tau^2 n}$ .

*Proof.* The lemma is proved conditioned on the message  $x$  encoded by  $D$  and the randomness  $r$  of the encoding. For the remainder of the proof,  $r$  is therefore considered fixed inside  $H_1$  and  $H_2$ . The proof, similarly to that of Lemma 7, again uses the self-destruct lemma.

Set  $\mathcal{X} := \mathcal{F}$  and  $\mathcal{Y} := \{0, 1\}^k \cup \{\perp\}$ . However, this time, let  $\mathcal{R} := \mathcal{S}$ . For  $f \in \mathcal{X}$  and  $T \in \mathcal{R}$ , define

$$g(f, T) := \text{Dec}(\tilde{c}, T),$$

where  $\tilde{c} := f(\text{Enc}(x; r))$ . The bending set  $\mathcal{B} \subseteq \mathcal{X}$  is the set of all high queries  $f$  such that  $w_f$  exists and  $d_{\text{H}}(w_f, \tilde{c}) > \alpha n$ .<sup>11</sup> It is readily verified that  $H_2$  is a parallel stateless self-destruct game (cf. Definition 6) that behaves according to  $g$ , and that  $H_1$  is its  $\mathcal{B}$ -bending.

Consider a query  $f \notin \mathcal{B}$ . If  $f$  is a low query, the unique answer is  $y_f = x$ ; if it is a middle query,  $y_f = \perp$ ; if it is a high query,  $y_f = x_f$  (cf. Lemma 6). Therefore,

$$\Delta^D(H_1, H_2) \leq \max_{f \in \mathcal{B}} \mathbb{P}[g(f, T) \neq \perp] \leq pe^{-\tau^2 n},$$

where the first inequality follows from the self-destruct lemma (Lemma 21) and the second one from the fact that  $d_{\text{H}}(x_f, \tilde{c}) > \tau n$  for queries  $f \in \mathcal{B}$ , and therefore the probability over the choice of  $T$  that it is accepted is at most  $(1 - \tau)^{\tau n} \leq e^{-\tau^2 n}$ .  $\square$

#### 4.4.4 Simulation

By analyzing hybrid  $H_2$ , one observes that low and high queries can now be answered knowing only the query itself and the symbols of the encoding indexed by the secret trigger set  $T \in \mathcal{S}$ .

**Lemma 9.** *Consider the random experiment of distinguisher  $D$  interacting with  $H_2$ . There is an efficiently computable function  $\text{Dec}' : \mathcal{F}_{\text{set}} \times \mathcal{S} \times \{0, 1\}^{\tau n} \rightarrow \{0, 1\}^k \cup \{\text{same}, \perp\}$  such that for any low or high query  $f$ , any fixed message  $x$ , any fixed encoding  $c$  thereof, and any output  $T$  of  $\text{Gen}$ ,*

$$[\text{Dec}'(f, T, (c[j])_{j \in T})]_{\text{same}/x} = \text{Dec}(f(c)),$$

where  $[\cdot]_{\text{same}/x}$  is the identity function except that **same** is replaced by  $x$  and where  $(c[j])_{j \in T}$  are the symbols of  $c$  specified by  $T$ .

*Proof.* Consider a low query  $f$ . Due to the error correction,  $\text{Dec}(f(c))$  is the message originally encoded if no bit indexed by  $T$  is changed and  $\perp$  otherwise. Which one is the case can clearly be efficiently computed from  $f$ ,  $T$ , and  $(c[j])_{j \in T}$ .

For high queries  $f$  the statement follows by inspecting the definition of  $H_2$  and Lemma 6.  $\square$

In  $H_2$ , by the  $\tau n$ -secrecy of the LECSS, the distribution of the symbols indexed by  $T$  is independent of the message  $x$  encoded by  $D$ . Moreover, the distribution of  $T$  is trivially independent of  $x$ . This suggests the following simulator  $\text{sim}$ : Initially, it chooses a random subset  $T$  from  $\binom{[n]}{\tau n}$  and chooses  $\tau n$  random symbols  $(c[j])_{j \in T}$ . Every component  $f$  of any tamper query is handled as follows: If  $f$  is a low or a high query, the answer is  $\text{Dec}'(f, T, (c[j])_{j \in T})$ ; if  $f$  is a middle query, the answer is  $\perp$ . This implies:

**Lemma 10.**  $H_2 \equiv S_{\mathcal{F}, \text{sim}}$ .

*Proof of Theorem 4.* Follows from Lemmas 7, 8, and 10 and a triangle inequality.  $\square$

<sup>11</sup>These are queries potentially accepted by  $H_2$  but not by  $H_1$ .

## 4.5 LECSS for the Non-Malleable Code

Let  $\mathbb{F} = \text{GF}(2)$  and  $\alpha > 0$ . In this section we show how to construct a  $(k, n, \delta, \tau)$ -LECSS  $(\text{E}, \text{D})$  (cf. Definition 1 in Section 2) with minimum distance  $\delta$  and secrecy  $\tau$  over  $\mathbb{F}$  and the following properties (as required in Section 4.3):

- *Minimum distance:*  $\delta > 1/4 + 2\alpha$  and  $\delta/2 > 2\alpha$ .
- *Constant rate:*  $k/n = \Omega(1)$ .
- *Constant secrecy:*  $\tau = \Omega(1)$ .

The construction combines high-distance binary codes with a recent result by Cramer *et al.* [11], which essentially allows to “add” secrecy to any code of sufficient rate.

Let  $\mathcal{C}$  be a  $(n, l)$ -code with rate  $R = \frac{l}{n}$  over  $\mathbb{F}$ . In the following we write  $\mathcal{C}(x)$  for the codeword corresponding to  $x \in \mathbb{F}^l$  and  $\mathcal{C}^{-1}(c, e)$  for the output of the efficient error-correction algorithm attempting to correct up to  $e$  errors on  $c$ , provided that  $e < \delta n/2$ ,<sup>12</sup> the output is  $\perp$  if there is no codeword within distance  $e$  of  $c$ .

ADDING SECRECY. Let  $l$  be such that  $k < l < n$ . The construction by [11] combines a surjective linear universal hash function  $\mathbf{h} : \mathbb{F}^l \rightarrow \mathbb{F}^k$  with  $\mathcal{C}$  to obtain a LECSS  $(\text{E}, \text{D})$  as follows:<sup>13</sup>

- $\text{E}(x)$  for  $x \in \{0, 1\}^k$ : Choose  $s \in \{0, 1\}^l$  randomly such that  $\mathbf{h}(s) = x$  and output  $c = \mathcal{C}(s)$ .
- $\text{D}(c, e)$  for  $c \in \{0, 1\}^n$  and  $e < \delta n/2$ : Compute  $s = \mathcal{C}^{-1}(c, e)$ . If  $s = \perp$ , output  $\perp$ . Otherwise, output  $x = \mathbf{h}(s)$ .

The resulting LECSS has rate  $\rho = \frac{k}{ln}$  and retains all distance and error-correction properties of  $\mathcal{C}$ . Additionally, if  $R$  is not too low, the LECSS has secrecy. More precisely, Cramer *et al.* prove the following theorem:

**Theorem 11** ([11]). *Let  $\tau > 0$  and  $\eta > 0$  be constants and  $\mathcal{H}$  be a family of linear universal hash functions  $\mathbf{h} : \mathbb{F}^l \rightarrow \mathbb{F}^k$ . Given that  $R \geq \rho + \eta + \tau + h(\tau)$ , there exists a function  $\mathbf{h} \in \mathcal{H}$  such that  $(\text{E}, \text{D})$  achieves secrecy  $\tau$ . Moreover, such a function  $\mathbf{h}$  can be chosen randomly with success probability  $1 - 2^{-m}$ .*

It should be pointed out that the version of the above theorem in [11] does not claim that any  $\tau n$  bits of an encoding are uniform and independent but merely that they are independent of the message encoded. However, by inspecting their proof, it can be seen that uniformity is guaranteed if  $\tau n \leq l - k$ , which is the case if and only if  $\tau \leq \frac{l}{n} - \frac{k}{n} = R - \rho$ , which is clearly implied by the precondition of the theorem.

ZYABLOV BOUND. For code  $\mathcal{C}$ , we use concatenated codes reaching the Zyablov bound:

**Theorem 12.** *For every  $\delta < 1/2$  and all sufficiently large  $n$ , there exists a code  $\mathcal{C}$  that is*

- *linear,*
- *efficiently encodable,*
- *of distance at least  $\delta n$ ,*
- *allows to efficiently correct up to  $\delta n/2$  errors,*

and has rate

$$R \geq \max_{0 \leq r \leq 1 - h(\delta + \varepsilon)} r \left( 1 - \frac{\delta}{h^{-1}(1 - r) - \varepsilon} \right),$$

for  $\varepsilon > 0$  and where  $h(\cdot)$  is the binary entropy function.

<sup>12</sup>This assumes that  $\mathcal{C}$  is efficiently decodable up to relative distance  $\delta/2$ . However, while the codes we consider here have this property, for our non-malleable code construction, it would be sufficient to have efficient error correction up to distance  $2\alpha$  for whatever particular choice of the constant  $\alpha$ .

<sup>13</sup>Note that we switched the roles of  $l$  and  $k$  here in order to remain consistent with the notation in this paper.



The Zyablov bound is achieved by concatenating Reed-Solomon codes with linear codes reaching the Gilbert-Varshamaov bound (which can be found by brute-force search in this case). Alternatively, Shen [32] showed that the bound is also reached by an explicit construction using algebraic geometric codes.

**CHOICE OF PARAMETERS.** Set  $\alpha := 1/200$  and  $\delta := 1/4 + 2\alpha + \varepsilon$  for  $\varepsilon := 1/500$ , say. Then,  $\delta - 2\alpha > 1/4$ , as required. Moreover, the rate of the Zyablov code with said distance  $\delta$  can be approximated to be  $R \geq 0.0175$ . Setting,  $\tau := 1/1000$  yields  $\tau + h(\tau) \leq 0.0125$ , leaving a possible rate for the LECSS of up to  $\rho \approx 0.005 - \eta$ . Hence:

**Corollary 13.** *For any  $\alpha > 0$  there exists a  $(k, n, \delta, \tau)$ -LECSS  $(E, D)$  with the following properties:*

- Minimum distance:  $\delta > 1/4 + 2\alpha$  and  $\delta/2 > 2\alpha$ .
- Constant rate:  $k/n = \Omega(1)$ .
- Constant secrecy:  $\tau = \Omega(1)$ .

#### 4.6 Impossibility for Codes without State

We show that codes without secret state (as, e.g., the ones in [17, 16, 1, 19, 10, 7, 2]) cannot achieve (unconditional) non-malleability against parallel tampering. Specifically, we prove the following theorem:

**Theorem 14.** *Let  $\mathcal{F} := \mathcal{F}_{\text{set}}$ . Let  $(\text{Enc}, \text{Dec})$  be a  $(k, n)$ -code without secret state and noticeable rate. There exists a distinguisher  $D$  asking a single parallel tampering query of size  $n^6$  such that, for all simulators  $\text{sim}$  and all  $n$  large enough,  $\Delta^D(R_{\mathcal{F}}, S_{\mathcal{F}, \text{sim}}) \geq 1/2$ .*

The above impossibility result requires that the rate of the code not be too small (in fact  $n = o(2^{k/6})$  suffices, see Appendix B for the exact parameters). The distinguisher  $D$  is inefficient, so it might still be possible to construct a non-malleable code against parallel tampering with only computational security. We leave this as an interesting open question for future research.

Here, we outline an attack for the case where  $\text{Dec}$  is deterministic. A full proof and a generalization to the setting where  $\text{Dec}$  uses (independent) randomness for (each) decoding is in Appendix B.

*Proof (sketch).* Roughly, a possible attack works as follows: There exists an (inefficient) extraction algorithm that, by suitably tampering with an encoding in the real experiment  $R_{\mathcal{F}}$ , is able to recover the original plaintext with high probability. Since (modulo some technicalities) this is not possible in the ideal experiment  $S_{\mathcal{F}, \text{sim}}$  (for any simulator  $\text{sim}$ ), this constitutes a distinguishing attack.

For simplicity, suppose that the decoding algorithm  $\text{Dec}$  is deterministic. The extraction relies on the fact that for any position  $i \in [n]$  with relevance in the decoding, there exist two codewords  $c'_i$  and  $c''_i$  with  $\text{Dec}(c'_i) \neq \text{Dec}(c''_i)$  and differing in position  $i$  only. From the result of a tamper query fixing all but the  $i^{\text{th}}$  position to correspond with the bits of  $c'_i$  (or  $c''_i$ ) one can therefore infer the value of the  $i^{\text{th}}$  bit of the encoding. This extraction is an independent process for every (relevant) position and thus parallelizable. In other words, a single parallel tamper query can be used to recover every relevant position of an encoding (from which the original message can be computed by filling the non-relevant positions with arbitrary values and applying the decoding algorithm).  $\square$

## 5 Construction from CPA Security

In this section we show that NM-SDA security can be achieved in a black-box fashion from IND-CPA security. Specifically, we prove that a generalization using LECSS (cf. Section 2) of the scheme by Choi *et al.* [9] (dubbed the CDMW construction in the remainder of this section) is NM-SDA secure. Using a constant-rate LECSS allows to improve the rate of the CDMW construction from  $\Omega(1/\lambda^2)$  to  $\Omega(1/\lambda)$ , where  $\lambda$  is the security parameter. This abstraction might also give a deeper understanding of the result of [9]. The main difficulty in the analysis is to extend their proof to deal with adaptively chosen parallel decryption queries (with self-destruct).

PKE Scheme $\Pi' = (KG', E', D')$		
<p>Key Generation <math>KG'</math></p> <p><b>for</b> <math>(b, i, j) \in \{0, 1\} \times [\kappa] \times [n]</math></p> <p style="padding-left: 20px;">  <math>(\text{pk}_{i,j}^b, \text{sk}_{i,j}^b) \leftarrow KG</math></p> <p>PK <math>\leftarrow (\text{pk}_{i,j}^b)_{b,i,j}</math></p> <p>SK <math>\leftarrow (\text{sk}_{i,j}^b)_{b,i,j}</math></p> <p><math>T \leftarrow_s \binom{[n]}{\tau n}</math></p> <p><b>return</b> (PK, (SK, T))</p>	<p>Encryption <math>E'_{\text{PK}}(m)</math></p> <p><math>(c_1, \dots, c_n) \leftarrow \mathbf{E}(m)</math></p> <p><math>(\text{verk}, \text{sigk}) \leftarrow KG^{\text{ots}}</math></p> <p><math>(v[1], \dots, v[\kappa]) \leftarrow \text{verk}</math></p> <p><b>for</b> <math>(i, j) \in [\kappa] \times [n]</math></p> <p style="padding-left: 20px;">  <math>e_{i,j} \leftarrow E_{\text{pk}_{i,j}^{v[i]}}(c_j)</math></p> <p><math>\mathbf{E} \leftarrow (e_{i,j})_{i,j}</math></p> <p><math>\sigma \leftarrow S_{\text{sigk}}(\mathbf{E})</math></p> <p><b>return</b> <math>(\mathbf{E}, \text{verk}, \sigma)</math></p>	<p>Decryption <math>D'_{(\text{SK}, T)}(\mathbf{E}, \text{verk}, \sigma)</math></p> <p><b>if</b> <math>V_{\text{verk}}(\mathbf{E}, \sigma) = 0</math></p> <p style="padding-left: 20px;">  <b>return</b> <math>\perp</math></p> <p><b>for</b> <math>j \in T</math></p> <p style="padding-left: 20px;">  decrypt <math>j^{\text{th}}</math> column of <math>\mathbf{E}</math></p> <p style="padding-left: 40px;">  <b>if</b> <i>not all entries identical</i></p> <p style="padding-left: 40px;">    <b>return</b> <math>\perp</math></p> <p>decrypt first row of <math>\mathbf{E}</math> to <math>c</math></p> <p><math>(m, w) \leftarrow \mathbf{D}(c, \alpha n)</math></p> <p><b>if</b> <math>w = \perp</math> <i>or</i> <math>\exists j \in T : c_j \neq w_j</math></p> <p style="padding-left: 20px;">  <b>return</b> <math>\perp</math></p> <p><b>return</b> <math>m</math></p>

**Figure 5:** The CDMW PKE scheme  $\Pi'$  constructed from a CPA-secure scheme  $\Pi$  [9].

### 5.1 The CDMW construction.

The CDMW construction uses a randomized Reed-Solomon code, which is captured as a special case by the notion of a linear error-correcting secret sharing (LECSS)  $(\mathbf{E}, \mathbf{D})$  (cf. Section 2). For ease of description, we assume that the decoding algorithm returns not only the plaintext  $x$  but also the corresponding codeword  $w$ , i.e.,  $(x, w) \leftarrow \mathbf{D}(c, e)$ , where  $e \in \mathbb{N}$  specifies the number of errors to correct; moreover, the output is  $(x, w) = (\perp, \perp)$  if  $c$  is not within distance  $e$  of any codeword.

The LECSS has to satisfy an additional property, which is that given a certain number of symbols chosen uniformly at random and independently and a plaintext  $x$ , one can efficiently produce an encoding that matches the given symbols and has the same distribution as  $\mathbf{E}(x)$ . It is described in more detail in the proof of Lemma 20, where it is needed.<sup>14</sup>

Let  $\Pi = (KG, E, D)$  be a PKE scheme with message space  $\mathcal{M} = \{0, 1\}^\ell$  (we assume  $\ell = \Omega(\lambda)$ ), and let  $\Sigma = (KG^{\text{ots}}, S, V)$  be a one-time signature scheme with verification keys of length  $\kappa = \mathcal{O}(\lambda)$ . Moreover, let  $\alpha > 0$  be any constant and  $(\mathbf{E}, \mathbf{D})$  a  $(k, n, \delta, \tau)$ -LECSS over  $\text{GF}(2^\ell)$  with  $\delta > 2\alpha$ .

The CDMW construction (cf. Figure 5), to encrypt a plaintext  $m \in \{0, 1\}^{k\ell}$ , first computes an encoding  $(c_1, \dots, c_n) \leftarrow \mathbf{E}(m)$  and then creates the  $(\kappa \times n)$ -matrix  $\mathbf{C}$  in which this encoding is repeated in every row. For every entry  $\mathbf{C}_{ij}$  of this matrix, there are two possible public keys  $\text{pk}_{i,j}^b$ ; which of them is used to encrypt the entry is determined by the  $i^{\text{th}}$  bit  $v[i]$  of the verification key  $\text{verk} = (v[1], \dots, v[\kappa])$  of a freshly generated key pair for  $\Sigma$ . In the end, the encrypted matrix  $\mathbf{E}$  is signed using  $\text{verk}$ , producing a signature  $\sigma$ . The ciphertext is  $(\mathbf{E}, \text{verk}, \sigma)$ .

The decryption first verifies the signature. Then, it decrypts all columns indexed by a set  $T \subset [n]$ , chosen as part of the secret key, and checks that each column consists of a single value only. Finally, it decrypts the first row and tries to find a codeword with relative distance at most  $\alpha$ . If so, it checks whether the codeword matches the first row in the positions indexed by  $T$ . If all checks pass, it outputs the plaintext corresponding to the codeword; otherwise it outputs  $\perp$ .

In the remainder of this section, we sketch the proof of the following theorem, which implies Theorem 2.

**Theorem 15.** *Let  $t \in \mathbb{N}$  and  $\Pi$  be a  $(t + t_{\text{cpa}}, \varepsilon_{\text{cpa}})$ -IND-CPA-secure PKE scheme,  $\alpha > 0$ ,  $(\mathbf{E}, \mathbf{D})$  a  $(k, n, \delta, \tau)$ -LECSS with  $\delta > 2\alpha$ , and  $\Sigma$  a  $(t + t_{\text{ots}}, \varepsilon_{\text{ots}})$ -secure OTS scheme with verification-key length  $\kappa$ . Then, for any  $q, p \in \mathbb{N}$ , PKE scheme  $\Pi'$  is  $(t, q, p, \varepsilon)$ -NM-SDA-secure with*

$$\varepsilon = (1 - \tau)\kappa n \cdot \varepsilon_{\text{cpa}} + 2 \cdot \varepsilon_{\text{ots}} + 4 \cdot p(1 - \tau)^{\alpha n},$$

where  $t_{\text{cpa}}$  and  $t_{\text{ots}}$  represent the overhead incurred by corresponding reductions.

INSTANTIATING THE CONSTRUCTION. Note that the security proof below does not use the linearity of the LECSS. The CDMW construction can be seen as using a Reed-Solomon-based LECSS with rate  $\mathcal{O}(1/\kappa)$ .

<sup>14</sup>Of course, the Reed-Solomon-based LECSS from [9] has this property.

If the construction is instantiated with a constant-rate LECSS, the final rate improves over CDMW by a factor of  $\Omega(\kappa) = \Omega(\lambda)$ . More concretely, assuming a constant-rate CPA encryption, a ciphertext of length  $\mathcal{O}(\lambda^3)$  can encrypt a plaintext of length  $\Omega(\lambda^2)$  as compared to  $\Omega(\lambda)$  for plain CDMW. As shown in Section 5.3, the LECSS can be instantiated with constructions based on Reed-Solomon or algebraic geometric codes (which also satisfy the additional property mentioned above), both with constant rate. Among the constant-rate codes, algebraic geometric codes allow to choose the parameters optimally also for shorter plaintexts.

## 5.2 Security Proof of the CDMW Construction

### 5.2.1 Overview

The proof follows the original one by [9]. The main change is that one needs to argue that, unless they contain invalid ciphertexts, adaptively chosen parallel queries do not allow the attacker to obtain useful information, in particular on the secret set  $T$ . This is facilitated by using the *self-destruct lemma* (cf. Section 6). The proof proceeds in three steps using two hybrid games  $H_b$  and  $H'_b$ :

- The first hybrid  $H_b$  gets rid of signature forgeries for the verification key used to create the challenge ciphertext. The indistinguishability of the hybrid from  $G_b^{\Pi', \text{nm-sda}}$  follows from the security of the OTS scheme and requires only minor modifications compared to the original proof.
- The second hybrid  $H'_b$  uses an alternative decryption algorithm. The indistinguishability of  $H'_b$  and  $H_b$  holds unconditionally; this step requires new techniques compared to the original proof.
- Finally, the distinguishing advantage between  $H'_0$  and  $H'_1$  is bounded by a reduction to the IND-CPA security of the underlying scheme  $\Pi$ ; the reduction again resembles the one in [9].

### 5.2.2 Dealing with Forgeries

For  $b \in \{0, 1\}$ , hybrid  $H_b$  behaves as  $G_b^{\Pi', \text{nm-sda}}$  but generates the signature key pair  $(\text{sigk}^*, \text{verk}^*)$  used for the challenge ciphertext initially and rejects any decryption query  $(\mathbf{E}', \sigma', \text{verk}')$  if  $\text{verk}' = \text{verk}^*$ .

**Lemma 16.** *For  $b \in \{0, 1\}$ , there exists a reduction  $R'_b(\cdot)$  such that for all distinguishers  $D$ ,*

$$\Delta^D(G_b^{\Pi', \text{nm-sda}}, H_b) \leq \Gamma^{R'_b(D)}(G^{\Sigma, \text{ots}}).$$

*Proof.*  $R'_b(\cdot)$  is a standard reduction to the unforgeability of  $\Sigma$ . □

### 5.2.3 Alternative Decryption Algorithm

For  $b \in \{0, 1\}$ , hybrid  $H'_b$  behaves as  $H_b$  but for the way it answers decryption queries  $(\mathbf{E}', \sigma', \text{verk}')$ : As before, it first verifies the signature  $\sigma'$  and checks that each column of  $\mathbf{E}'$  consists of encryptions of a single value. Then, it determines the first position  $i$  at which  $\text{verk}'$  and  $\text{verk}^*$  differ, i.e., where  $v'[i] \neq v^*[i]$ . It decrypts the  $i^{\text{th}}$  row of  $\mathbf{E}$  and checks if there is a codeword  $w$  within distance  $2\alpha n$ .<sup>15</sup> If such  $w$  does not exist or else if  $w$  does not match the *first* row in a position indexed by  $T$ , the check fails. Otherwise, the plaintext corresponding to  $w$  is output.

**Lemma 17.** *For  $b \in \{0, 1\}$  and all distinguishers  $D$ ,  $\Delta^D(H_b, H'_b) \leq 2 \cdot p(1 - \tau)^{\alpha n}$ .*

The proof of Lemma 17 shows that the original and alternative decryption algorithms are indistinguishable not just for a single parallel query (as is sufficient for NM-CPA) but even against adaptively chosen parallel queries (with self-destruct). It is the main technical contribution of this section.

At the core of the proof is an analysis of how different types of encoding matrices  $\mathbf{C}$  are handled inside the two decryption algorithms. To that end, one can define two games  $B$  and  $B'$  (below) that

<sup>15</sup>Recall that the actual decryption algorithm always decrypts the first row and tries to find  $w$  within distance  $\alpha n$ .

capture the behaviors of the original and the alternative decryption algorithms, respectively. The proof is completed by bounding  $\Delta(B, B')$  (for all distinguishers) and showing the existence of a wrapper  $W_b$  such that  $W_b(B)$  behaves as  $H_b$  and  $W_b(B')$  as  $H'_b$  (also below). This proves the lemma since  $\Delta^D(H_b, H'_b) = \Delta^D(W_b(B), W_b(B')) = \Delta^{D(W_b(\cdot))}(B, B')$ .

The games  $B$  and  $B'$  behave as follows: Both initially choose a random size- $\tau$  subset of  $[n]$ . Then, they accept parallel queries with components of the type  $(\mathbf{C}, i)$  for  $\mathbf{C} \in \mathbb{F}^{\kappa \times n}$  and  $i \in [\kappa]$ . The answer to each component is computed as follows:

1. Both games check that all columns indexed by  $T$  consist of identical entries.
2. Game  $B$  tries to find a codeword  $w$  with distance less than  $\alpha n$  from the *first* row (regardless of  $i$ ), whereas  $B'$  tries to find  $w$  within  $2\alpha n$  of row  $i$ . Then, if such a  $w$  is found, *both* games check that it matches the *first* row of  $\mathbf{C}$  in the positions indexed by  $T$ .
3. If all checks succeed, the answer to the (component) query is  $w$ ; otherwise, it is  $\perp$ .

Both games then output the answer vector and implement the self-destruct, i.e., if any of the answers is  $\perp$ , all *future* queries are answered by  $\perp$ .

**Claim 18.** For  $b \in \{0, 1\}$  and all distinguishers  $D$ ,  $\Delta^D(B, B') \leq 2 \cdot p(1 - \tau)^{\alpha n}$ .

ENCODING MATRICES. Towards a proof of Claim 18, consider the following partition of the set of encoding matrices  $\mathbf{C}$  (based on the classification in [9]):

1. There exists a codeword  $w$  within  $\alpha n$  of the first row of  $\mathbf{C}$ , and all rows have distance at most  $\alpha n$ .
2. (a) There exist two rows in  $\mathbf{C}$  with distance greater than  $\alpha n$ .  
(b) The rest; in this case the first row differs in more than  $\alpha n$  positions from any codeword.

Observe that queries  $(\mathbf{C}, i)$  with  $\mathbf{C}$  of type 1 are treated identically by both  $B$  and  $B'$ : A codeword  $w$  within  $\alpha n$  of the first row of  $\mathbf{C}$  is certainly found by  $B$ ; since all rows have distance at most  $\alpha n$ ,  $w$  is within  $2\alpha n$  of row  $i$  and thus also found by  $B'$ . Furthermore, note that if  $\mathbf{C}$  is of type 2b, it is always rejected by  $B$  (but not necessarily by  $B'$ ).

Consider the hybrids  $C$  and  $C'$  that behave as  $B$  and  $B'$ , respectively, but always reject *all* type-2 queries. Since type-1 queries are treated identically,  $C$  and  $C'$  are indistinguishable. Moreover:

**Claim 19.** For all distinguishers  $D$ ,

$$\Delta^D(B, C) \leq p(1 - \tau)^{\alpha n} \quad \text{and} \quad \Delta^D(C', B') \leq p(1 - \tau)^{\alpha n}.$$

The proof of Claim 19 follows a generic paradigm, at whose core is the so-called *self-destruct lemma*, which deals with the indistinguishability of hybrids with the self-destruct property and is explained in detail in Section 6. Roughly, this lemma applies whenever the first hybrid (in this case  $B$  resp.  $B'$ ) can be turned into the second one (in this case  $C$  resp.  $C'$ ) by changing (“bending”) the answers to a subset (the “bending set”) of the possible queries to always be  $\perp$ , and when additionally non-bent queries have a unique answer (cf. the statement of Lemma 21). Intuitively, the lemma states that parallelism and adaptivity do not help distinguish (much) in such cases.

*Proof.* To use the self-destruct lemma, note that  $B, C, C'$ , and  $B'$  all answer queries from  $\mathcal{X} := \mathbb{F}^{\kappa \times n} \times [\kappa]$  by values from  $\mathcal{Y} := \mathbb{F}^n$ . Moreover, note that they use as internal randomness a uniformly chosen element  $T$  from the set  $\mathcal{R} := \binom{[n]}{\tau n}$  of size- $\tau n$  subsets of  $[n]$ .

Consider first  $B$  and  $C$ . Let  $g : \mathcal{X} \times \mathcal{R} \rightarrow \mathcal{Y}$  correspond to how  $B$  answers queries  $(\mathbf{C}, i)$  (see above). Let  $\mathcal{B}$  be the set  $\mathcal{B}$  of all type-2a-queries. Then,  $C$  is its  $\mathcal{B}$ -bending (cf. Definition 7). Observe that queries  $x = (\mathbf{C}, i) \notin \mathcal{B}$  are either of type 1 or 2b. For the former, the unique answer  $y_x$  is the codeword  $w$  within  $\alpha n$  of the first row of  $\mathbf{C}$ ; for the latter,  $y_x$  is  $\perp$ . Therefore, using the self-destruct lemma (Lemma 21), for all distinguishers  $D$ ,

$$\Delta^D(B, C) \leq p \cdot \max_{(\mathbf{C}, i) \in \mathcal{B}} \mathbb{P}[g((\mathbf{C}, i), T) \neq \perp],$$

where the probability is over the choice of  $T$ . Since type-2a queries have two rows with distance greater than  $\alpha n$ , the probability over the choice of  $T$  that this remains unnoticed is at most  $(1 - \tau)^{\alpha n}$ .

For the second part of the claim, consider  $B'$  and  $C'$ . Now, let  $g : \mathcal{X} \times \mathcal{R} \rightarrow \mathcal{Y}$  correspond to how  $B'$  answers queries  $(\mathbf{C}, i)$  (see above again), and let  $\mathcal{B}$  be the set  $\mathcal{B}$  of all type-2-queries. Then,  $C'$  is the  $\mathcal{B}$ -bending of  $B'$ .

Note that all queries  $x = (\mathbf{C}, i) \notin \mathcal{B}'$  are of type 1, and the unique answer  $y_x$  is the codeword  $w$  within  $2\alpha n$  of row  $i$  of  $\mathbf{C}$ . Therefore, using Lemma 21 again, for all distinguishers  $D$ ,

$$\Delta^D(B', C') \leq p \cdot \max_{(\mathbf{C}, i) \in \mathcal{B}'} \mathbb{P}[g'((\mathbf{C}, i), T) \neq \perp],$$

where the probability is again over the choice of  $T$ . Since type-2a queries have two rows with distance greater than  $\alpha n$  and in type-2b queries the first row differs in more than  $\alpha n$  positions from any codeword, the probability over the choice of  $T$  that this remains unnoticed is at most  $(1 - \tau)^{\alpha n}$ .  $\square$

*Proof (of Claim 18).* The proof follows using the triangle inequality:

$$\Delta^D(B, B') \leq \Delta^D(B, C) + \Delta^D(C, C') + \Delta^D(C', B') \leq 2 \cdot p(1 - \tau)^{\alpha n}.$$

$\square$

WRAPPER. It remains to show that there exists a wrapper  $W_b$  such that  $W_b(B)$  behaves as  $H_b$  and  $W_b(B')$  as  $H'_b$ . The construction of  $W_b$  is straight forward:  $H_b$  and  $H'_b$  generate all keys and the challenge in the identical fashion; therefore,  $W_b$  can do it the same way.  $W_b$  answers decryption queries  $(\mathbf{E}', \text{verk}', \sigma')$  by first verifying the signature  $\sigma'$  and rejecting queries if  $\sigma'$  is invalid or if  $\text{verk}'$  is identical to the verification key  $\text{verk}^*$  chosen for the challenge, decrypting the entire matrix  $\mathbf{E}'$  to  $\mathbf{C}'$  and submitting  $(\mathbf{C}', i)$  to the oracle (either  $B$  or  $B'$ ), where  $i$  is the first position at which  $\text{verk}'$  and  $\text{verk}^*$  differ, and decoding the answer  $w$  and outputting the result or simply forwarding it if it is  $\perp$ . Moreover,  $W_b$  implements the self-destruct. By inspection it can be seen that  $W_b(B)$  implements the original decryption algorithm and  $W_b(B')$  the alternative one.

#### 5.2.4 Reduction to IND-CPA Security

**Lemma 20.** *There exists a reduction  $R(\cdot)$  such that for all distinguishers  $D$ ,*

$$\Delta^D(H'_0, H'_1) = (1 - \tau)\kappa n \cdot \Delta^{D(R(\cdot))}(G_0^{\text{II, ind-cpa}}, G_1^{\text{II, ind-cpa}}).$$

*Proof (sketch).* The proof is a straight-forward generalization of the original proof by [9]; the only difference is that it needs to process multiple parallel decryption queries and implement the self-destruct feature appropriately. For ease of exposition, we describe the reduction to a many-public-key version of the CPA game for  $\Pi$ .<sup>16</sup>

Reduction  $R(\cdot)$  initially chooses the secret set  $T$  and creates the challenge OTS key pair with verification key  $\text{verk}^* = (v^*[1], \dots, v^*[\kappa])$  and all key pairs  $(\text{pk}_{i,j}^b, \text{sk}_{i,j}^b)$  with  $j \in T$  or  $b \neq v^*[i]$ . The remaining  $(1 - \tau)\kappa n$  key pairs are generated by the CPA game.

Recall that the LECSS is assumed to satisfy the following property: Given  $\tau n$  symbols  $(c_i)_{i \in T}$  chosen uniformly at random and independently and any plaintext  $x \in \mathbb{F}^k$ , one can efficiently sample symbols  $(c_i)_{i \notin T}$  such that  $(c_1, \dots, c_n)$  has the same distribution as  $\mathbf{E}(x)$ . Using this fact,  $R(\cdot)$  creates the challenge for  $m_0$  and  $m_1$  as follows: It picks the random symbols  $(c_i)_{i \in T}$  and completes them to two full encodings  $c_{m_0}$  and  $c_{m_1}$  with the above procedure, once using  $m_0$  and once using  $m_1$  as the plaintext. Let  $\mathbf{C}_{m_0}$  and  $\mathbf{C}_{m_1}$  be the corresponding matrices (obtained by copying the encodings  $\kappa$  times). Observe that

<sup>16</sup>In the many-public-key version of the CPA game, an attacker can play the CPA game for several independently generated public keys simultaneously; this is equivalent to the normal formulation by a standard hybrid argument [3].

the two matrices match in the columns indexed by  $T$ . These entries are encrypted by  $R(\cdot)$ , using the public key  $\text{pk}_{i,j}^b$  for entry  $(i, j)$  for which  $b \neq v^*[i]$ . Denote by  $\mathbf{C}'_{m_0}$  and  $\mathbf{C}'_{m_1}$  the matrices  $\mathbf{C}_{m_0}$  and  $\mathbf{C}_{m_1}$  with the columns in  $T$  removed. The reduction outputs  $(\text{chall}, \mathbf{C}'_{m_0}, \mathbf{C}'_{m_1})$  to its oracle and obtains the corresponding ciphertexts, which it combines appropriately with the ones it created itself to form the challenge ciphertext.

Finally, note that since the reduction knows all the secret keys  $\text{pk}_{i,j}^b$  with  $b \neq v^*[i]$ , it can implement the alternative decryption algorithm (and the self-destruct).  $\square$

### 5.2.5 Overall Proof

*Proof (of Theorem 15).* Let  $t_{\text{cpa}}$  be the overhead caused by reduction  $R(\cdot)$  and  $t_{\text{ots}}$  the larger of the overheads caused by  $R'_0(\cdot)$  and  $R'_1(\cdot)$ . Moreover, let  $D$  be a distinguisher with running time at most  $t$ . Using the triangle inequality, and Lemmas 16, 17, and 20,

$$\begin{aligned}
\Delta^D(G_0^{\text{II}', \text{nm-sda}}, G_1^{\text{II}', \text{nm-sda}}) &\leq \Delta^D(G_0^{\text{II}', \text{nm-sda}}, H_0) + \Delta^D(H_0, H'_0) \\
&\quad + \Delta^D(H'_0, H'_1) + \Delta^D(H'_1, H_1) + \Delta^D(H_1, G_1^{\text{II}', \text{nm-sda}}) \\
&\leq \Gamma^{D(R_0(\cdot))}(G^{\Sigma, \text{ots}}) + 2 \cdot p(1 - \tau)^{\alpha n} \\
&\quad + (1 - \tau)\kappa n \cdot \Delta^{D(R(\cdot))}(G_0^{\text{II}, \text{ind-cpa}}, G_1^{\text{II}, \text{ind-cpa}}) \\
&\quad + 2 \cdot p(1 - \tau)^{\alpha n} + \Gamma^{D(R'_1(\cdot))}(G^{\Sigma, \text{ots}}) \\
&\leq \varepsilon_{\text{ots}} + 2 \cdot p(1 - \tau)^{\alpha n} \\
&\quad + (1 - \tau)\kappa n \cdot \varepsilon_{\text{cpa}} + 2 \cdot p(1 - \tau)^{\alpha n} + \varepsilon_{\text{ots}}.
\end{aligned}$$

$\square$

## 5.3 LECSS for the CDMW Construction

In this section we show how to instantiate the LECSS used for the CDMW construction in Section 5. Let  $\mathbb{F}$  be a finite field of size  $L = 2^\ell$ , where  $\ell$  is the plaintext length of the IND-CPA scheme used in the construction. Then, there are the following variants of a  $(k, n, \delta, \tau)$ -LECSS:

- *CDMW Reed-Solomon codes:* The original CDMW construction can be seen as using a Reed-Solomon-based LECSS with rate  $\Theta(1/\lambda)$ , which is suboptimal (see next item).
- *Constant-Rate Reed-Solomon codes:* Cheraghchi and Guruswami [8] provide a LECSS based on a construction by Dziembowski *et al.* [17] and on Reed-Solomon (RS) codes with  $\ell = \Theta(\log n)$ . One can show that it achieves the following parameters (not optimized):  $\alpha = 1/8$ ,  $\tau = 1/8$  and rate  $k/n \geq 1/4$  (i.e., all constant).
- *Algebraic geometric codes:* Using algebraic geometric (AG) codes, Cramer *et al.* [12] provide a LECSS with  $\ell = \mathcal{O}(1)$  and still constant error correction, secrecy, and rate (but with worse concrete constants than Reed-Solomon codes).

Note that asymptotically, RS and AG codes are equally good: both have constant rate, distance, and secrecy. However, since with AG codes  $\ell$  is constant (i.e., they work over an alphabet of constant size), the minimal plaintext length can be shorter than with RS codes.

## 6 A General Indistinguishability Paradigm

A recurring issue in this paper are proofs that certain self-destruct games answering successive parallel decryption/tampering queries are indistinguishable. We formalize such games as *parallel stateless self-destruct games*.

**Definition 6.** An oracle  $U$  is a parallel stateless self-destruct (PSSD) game if

- it accepts parallel queries in which each component is from some set  $\mathcal{X}$  and answers them by vectors with components from some set  $\mathcal{Y}$ ,
- $\perp \in \mathcal{Y}$ ,
- there exists a function  $g : \mathcal{X} \times \mathcal{R} \rightarrow \mathcal{Y}$  such that every query component  $x \in \mathcal{X}$  is answered by  $g(x, r)$ , where  $r \in \mathcal{R}$  is the internal randomness of  $U$ , and
- the game self-destructs, i.e., after the first occurrence of  $\perp$  in an answer vector all further outputs are  $\perp$ .

A PSSD game can be transformed into a related one by “bending” the answers to some of the queries  $x \in \mathcal{X}$  to the value  $\perp$ . This is captured by the following definition:

**Definition 7.** Let  $U$  be a PSSD game that behaves according to  $g$  and let  $\mathcal{B} \subseteq \mathcal{X}$ . The  $\mathcal{B}$ -bending of  $U$ , denoted by  $U'$ , is the PSSD game that behaves according to  $g'$ , where

$$g'(x, r) = \begin{cases} \perp & \text{if } x \in \mathcal{B}, \\ g(x, r) & \text{otherwise.} \end{cases}$$

The *self-destruct lemma* below states that in order to bound the distinguishing advantage between a PSSD and its bending, one merely needs to analyze a single, non-parallel query, provided that all non-bent queries  $x$  can only be answered by a unique value  $y_x$  or  $\perp$ .

**Lemma 21.** Let  $U$  be a PSSD game and  $U'$  its  $\mathcal{B}$ -bending for some  $\mathcal{B} \subseteq \mathcal{X}$ . If for all  $x \notin \mathcal{B}$  there exists  $y_x \in \mathcal{Y}$  such that

$$\{g(x, r) \mid r \in \mathcal{R}\} = \{y_x, \perp\},$$

then, for all distinguishers  $D$ ,

$$\Delta^D(U, U') \leq p \cdot \max_{x \in \mathcal{B}} \mathbb{P}[g(x, R) \neq \perp],$$

where the probability is over the choice of  $R$ .

*Proof.* Fix a distinguisher  $D$  and denote by  $R$  and  $R'$  the random variables corresponding to the internal randomness of  $U$  and  $U'$ , respectively. Call a value  $x \in \mathcal{X}$  *dangerous* if  $x \in \mathcal{B}$  and a query dangerous if it contains a dangerous value.

In the random experiment corresponding to the interaction between  $D$  and  $U$ , define the event  $E$  that the first dangerous query contains a dangerous value  $X$  with  $g(X, R) \neq \perp$  and that the self-destruct has not been provoked yet. Similarly, define the event  $E'$  for the interaction between  $D$  and  $U'$  that the first dangerous query contains a dangerous value  $X'$  with  $g(X', R') \neq \perp$  and that the self-destruct has not been provoked yet.<sup>17</sup>

Clearly,  $U$  and  $U'$  behave identically unless  $E$  resp.  $E'$  occur. Thus, it remains to bound  $\mathbb{P}[E] = \mathbb{P}[E']$ . To that end, note that adaptivity does not help in provoking  $E$ . For any distinguisher  $D$ , there exists a *non-adaptive* distinguisher  $\tilde{D}$  such that whenever  $D$  provokes  $E$ , so does  $\tilde{D}$ .  $\tilde{D}$  proceeds as follows: First, it interacts with  $D$  only. Whenever  $D$  asks a non-dangerous query,  $\tilde{D}$  answers every component  $x \notin \mathcal{B}$  by  $y_x$ . As soon as  $D$  specifies a dangerous query,  $\tilde{D}$  stops its interaction with  $D$  and sends all queries to  $U$ .

Fix all randomness in experiment  $\tilde{D}(U)$ , i.e., the coins of  $D$  (inside  $\tilde{D}$ ) and the randomness  $r$  of  $U$ . Suppose  $D$  would provoke  $E$  in the direct interaction with  $U$ . In such a case, all the answers by  $\tilde{D}$  are equal to the answers by  $U$ , since, by assumption, the answers to components  $x \notin \mathcal{B}$  in non-dangerous queries are  $y_x$  or  $\perp$  and the latter is excluded if  $E$  is provoked. Thus, whenever  $D$  provokes  $E$ ,  $\tilde{D}$  provokes it as well.

The success probability of non-adaptive distinguishers  $D$  is upper bounded by the probability over  $R$  that their first dangerous query provokes  $E$ , which is at most  $p \cdot \max_{x \in \mathcal{B}} \mathbb{P}[g(x, R) \neq \perp]$ .  $\square$

<sup>17</sup>Note that the function  $g$  is the *same* in the definitions of either event.

## References

- [1] Divesh Aggarwal, Yevgeniy Dodis, and Shachar Lovett. Non-malleable codes from additive combinatorics. *ECCC*, 20:81, 2013. To appear in STOC 2014.
- [2] Shashank Agrawal, Divya Gupta, Hemanta K. Maji, Omkant Pandey, and Manoj Prabhakaran. Explicit non-malleable codes resistant to permutations and perturbations. *IACR Cryptology ePrint Archive*, 2014:841, 2014.
- [3] Mihir Bellare, Alexandra Boldyreva, and Silvio Micali. Public-key encryption in a multi-user setting: Security proofs and improvements. In *EUROCRYPT*, pages 259–274, 2000.
- [4] Mihir Bellare and Amit Sahai. Non-malleable encryption: Equivalence between two notions, and an indistinguishability-based characterization. In *CRYPTO*, pages 519–536, 1999.
- [5] Ran Canetti, Shai Halevi, and Jonathan Katz. Chosen-ciphertext security from identity-based encryption. In *EUROCRYPT*, pages 207–222, 2004.
- [6] Ran Canetti, Hugo Krawczyk, and Jesper Buus Nielsen. Relaxing chosen-ciphertext security. In *CRYPTO*, pages 565–582, 2003.
- [7] Eshan Chattopadhyay and David Zuckerman. Non-malleable codes against constant split-state tampering. *Electronic Colloquium on Computational Complexity (ECCC)*, 21:102, 2014.
- [8] Mahdi Cheraghchi and Venkatesan Guruswami. Non-malleable coding against bit-wise and split-state tampering. In *TCC*, pages 440–464, 2014.
- [9] Seung Geol Choi, Dana Dachman-Soled, Tal Malkin, and Hoeteck Wee. Black-box construction of a non-malleable encryption scheme from any semantically secure one. In *TCC*, pages 427–444, 2008.
- [10] Sandro Coretti, Ueli Maurer, Björn Tackmann, and Daniele Venturi. From single-bit to multi-bit public-key encryption via non-malleable codes. *IACR Cryptology ePrint Archive*, 2014:324, 2014.
- [11] Ronald Cramer, Ivan Bjerre Damgård, Nico Döttling, Serge Fehr, and Gabriele Spini. Linear secret sharing schemes from error correcting codes and universal hash functions. In *EUROCRYPT*, pages 313–336, 2015.
- [12] Ronald Cramer, Goichiro Hanaoka, Dennis Hofheinz, Hideki Imai, Eike Kiltz, Rafael Pass, Abhi Shelat, and Vinod Vaikuntanathan. Bounded CCA2-secure encryption. In *ASIACRYPT*, pages 502–518, 2007.
- [13] Ronald Cramer and Victor Shoup. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. In *CRYPTO*, volume 1462 of *LNCS*, pages 13–25, 1998.
- [14] Ronald Cramer and Victor Shoup. Universal hash proofs and a paradigm for adaptive chosen ciphertext secure public-key encryption. In *EUROCRYPT*, pages 45–64, 2002.
- [15] Danny Dolev, Cynthia Dwork, and Moni Naor. Nonmalleable cryptography. *SIAM J. Comput.*, 30(2):391–437, 2000.
- [16] Stefan Dziembowski, Tomasz Kazana, and Maciej Obremski. Non-malleable codes from two-source extractors. In *CRYPTO (2)*, pages 239–257, 2013.
- [17] Stefan Dziembowski, Krzysztof Pietrzak, and Daniel Wichs. Non-malleable codes. In *ICS*, pages 434–452, 2010.



- [18] Sebastian Faust, Pratyay Mukherjee, Jesper Buus Nielsen, and Daniele Venturi. Continuous non-malleable codes. In *TCC*, pages 465–488, 2014.
- [19] Sebastian Faust, Pratyay Mukherjee, Daniele Venturi, and Daniel Wichs. Efficient non-malleable codes and key-derivation for poly-size tampering circuits. In *EUROCRYPT*, pages 111–128, 2014.
- [20] Yael Gertner, Tal Malkin, and Steven Myers. Towards a separation of semantic and CCA security for public key encryption. In *TCC*, pages 434–455, 2007.
- [21] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *J. Comput. Syst. Sci.*, 28(2):270–299, 1984.
- [22] Dennis Hofheinz and Eike Kiltz. Practical chosen ciphertext secure encryption from factoring. In *EUROCRYPT*, pages 313–332, 2009.
- [23] Susan Hohenberger, Allison B. Lewko, and Brent Waters. Detecting dangerous queries: A new approach for chosen ciphertext security. In *EUROCRYPT*, pages 663–681, 2012.
- [24] Kaoru Kurosawa and Yvo Desmedt. A new paradigm of hybrid encryption scheme. In *CRYPTO*, pages 426–442, 2004.
- [25] Yehuda Lindell. A simpler construction of cca2-secure public-key encryption under general assumptions. In *EUROCRYPT*, pages 241–254, 2003.
- [26] F.J. MacWilliams and N.J.A. Sloane. *The Theory of Error-Correcting Codes*. North-holland Publishing Company, 2nd edition, 1978.
- [27] Steven Myers and Abhi Shelat. Bit encryption is complete. In *FOCS*, pages 607–616, 2009.
- [28] Moni Naor and Moti Yung. Public-key cryptosystems provably secure against chosen ciphertext attacks. In *STOC*, pages 427–437, 1990.
- [29] Rafael Pass, Abhi Shelat, and Vinod Vaikuntanathan. Construction of a non-malleable encryption scheme from any semantically secure one. In *CRYPTO*, pages 271–289, 2006.
- [30] Rafael Pass, Abhi Shelat, and Vinod Vaikuntanathan. Relations among notions of non-malleability for encryption. In *ASIACRYPT*, pages 519–535, 2007.
- [31] Amit Sahai. Non-malleable non-interactive zero knowledge and adaptive chosen-ciphertext security. In *FOCS*, pages 543–553, 1999.
- [32] Ba-Zhong Shen. A justesen construction of binary concatenated codes that asymptotically meet the zyablov bound for low rate. *IEEE Transactions on Information Theory*, 39(1):239–242, 1993.

PKE Scheme $\Pi' = (KG', E', D')$		
Key Generation $KG'$ $(pk, sk) \leftarrow KG$ $\rho \leftarrow_{\$} \{0, 1\}^\lambda$ $pk' = pk$ $sk' = (\rho, sk)$ <b>return</b> $(pk', sk')$	Encryption $E'_{pk'}(m)$ $e \leftarrow E_{pk}(m)$ $e' = 0 \parallel e$ <b>return</b> $e$	Decryption $D'_{sk'}(e')$ $e' = \beta \parallel e$ $m \leftarrow D_{sk}(e)$ <b>if</b> $\beta = 1$   <b>return</b> $\rho$ <b>else</b>   <b>if</b> $m = \rho$     <b>return</b> $sk$   <b>else</b>     <b>return</b> $m$

**Figure 6:** PKE scheme  $\Pi'$  in the proof of Theorem 22

## A Relating IND-SDA and NM-CPA

**Theorem 22.** *IND-SDA and NM-CPA are incomparable notions:*

- (i) *Assume there exists a PKE scheme  $\Pi$  that is NM-CPA secure. Then there exists a PKE scheme  $\Pi'$  that is NM-CPA secure but not IND-SDA secure.*
- (ii) *Assume there exists a PKE scheme  $\Pi$  that is IND-SDA secure. Then there exists a PKE scheme  $\Pi''$  that is IND-SDA secure but not NM-CPA secure.*

The proof follows the following intuition: Towards proving (i), an assumed non-malleable PKE scheme is modified as follows: An additional 0-bit is added to every ciphertext. If it is changed to 1 by an adversary, the decryption algorithm outputs a secret random value  $\rho$  (from a sufficiently large space) instead of the normal decryption, and when the decryption algorithm is fed with an encryption of  $\rho$ , it outputs the secret key. The modified scheme is not IND-SDA-secure since an adversary can obtain the secret key by making just two *adaptive* decryption queries. It is, however, still non-malleable since a non-adaptive adversary can only try to guess  $\rho$ .

For (ii), an IND-SDA-secure PKE scheme is modified as follows: Prior to encrypting a message, it is encoded using a code with the property that a sufficiently large fraction of the bits of the encoding are random. Similarly to above, ciphertexts have a format that allows an adversary, using the decryption oracle, to guess a bit of this encoding, where a wrong guess leads to a self-destruct. Since, by definition, an adversary can ask arbitrarily many (non-adaptively chosen) invalid ciphertexts in the NM-CPA game, he can recover the entire encoding of the message. At the same time, due to the randomness of the encoding, an adversary trying to guess the bits of the encoding sequentially is highly likely to produce a self-destruct in the process. Therefore, the modified scheme is not non-malleable, but it is still IND-SDA-secure.

*Proof.* We prove the two statements separately.

(i) Let  $\Pi = (KG, E, D)$  be a PKE scheme that is NM-CPA secure, with message space  $\mathcal{M} = \{0, 1\}^\lambda$  for some value  $\lambda$  that is superpolynomial in the security parameter  $\lambda \in \mathbb{N}$ . Consider the PKE scheme  $\Pi' = (KG', E', D')$ , derived from  $\Pi$ , depicted in Figure 6. A secret key generated via  $KG'$  consists of a valid secret key for  $\Pi$ , together with a random string  $\rho \leftarrow_{\$} \{0, 1\}^\lambda$ . A ciphertext generated via  $E'$  consists of a valid ciphertext generated via  $E$ , to which we append a single bit  $\beta$  that is usually set to 0 in honestly generated ciphertexts. Upon input a ciphertext with  $\beta = 0$ , the decryption algorithm  $D'$  behaves as  $D$  except that it returns the secret key in case the ciphertext decrypts to  $\rho$ ; ciphertexts with  $\beta = 1$  are always decrypted to  $\rho$ .

**Claim 23.** *There exists a distinguisher  $D$  such that  $\Delta^D(G_0^{\Pi', 2, 1}, G_1^{\Pi', 2, 1}) = 1$ .*

*Proof.* Consider the following distinguisher  $D$  that breaks IND-SDA security of  $\Pi'$  with probability 1, using two adaptive decryption queries, and without ever provoking a self-destruct. Given the challenge ciphertext  $e$ , the distinguisher creates  $e_1 = 1 \parallel \bar{e}$  where  $\bar{e} \leftarrow E_{\text{pk}}(\bar{m})$  for a fixed message  $\bar{m} \in \mathcal{M}$ , and queries  $(\text{dec}, e_1)$ ; note that the answer to such query will be the value  $\rho$  (no matter what the value of  $\bar{m}$  is). Given  $\rho$ , the distinguisher encrypts  $e_2 \leftarrow 0 \parallel e_\rho$  where  $e_\rho \leftarrow E_{\text{pk}}(\rho)$  and queries  $(\text{dec}, e_2)$ . Note that, by definition of  $\Pi'$ , the answer to such query will be the secret key  $\text{sk}$  that can be used to decrypt the challenge ciphertext  $e$  with probability 1.  $\square$

**Claim 24.** *For all  $p \in \mathbb{N}$ , and for all distinguishers  $D'$ , there exists a distinguisher  $D$  such that*

$$\Delta^D(G_0^{\Pi,1,p}, G_1^{\Pi,1,p}) \geq \Delta^{D'}(G_0^{\Pi',1,p}, G_1^{\Pi',1,p}) - p \cdot 2^{-\lambda}.$$

*Proof.* We build distinguisher  $D$  (based on  $D'$ ) as follows:

1. At the beginning  $D$  samples a random  $\rho \leftarrow_{\$} \{0, 1\}^\lambda$  (but not  $\text{sk}$ ), and forwards the public key  $\text{pk}$  it receives from game  $G_b^{\Pi, \text{nm-cpa}}$  to  $D'$ .
2. Upon input  $(\text{chall}, m_0, m_1)$  from  $D'$ , distinguisher  $D$  calls  $(\text{chall}, m_0, m_1)$  to its own challenge oracle obtaining a ciphertext  $e$ , and returns  $e' := 0 \parallel e$  to  $D'$ .
3. Upon input the decryption query  $(\text{dec}, e'_1, \dots, e'_p)$  from  $D'$ , distinguisher  $D$  parses  $e'_i = \beta_i \parallel e_i$  for all  $i \in [p]$  and queries  $(\text{dec}, e_1, \dots, e_p)$  to its own decryption oracle, obtaining  $(m_1, \dots, m_p)$ . Hence,  $D$  proceeds as follows for each  $i \in [p]$ :
  - In case  $\beta_i = 1$ , replace  $m_i$  with  $m_i := \rho$ .
  - In case  $m_i = \rho$ , abort.

Return  $(m_1, \dots, m_p)$ .

4. Output whatever  $D'$  outputs.

Note that  $D$  perfectly simulates the key generation of  $\Pi'$  (this is because the value  $\rho$  is chosen uniformly and independently by  $D$ ). The same holds for the simulation of the challenge ciphertext. Furthermore, the simulation of the parallel decryption query is perfect unless  $D$  aborts; thus, it suffices to bound the probability that  $D$  aborts. By the union bound, this probability is at most  $p \cdot 2^{-\lambda}$ . The claim follows.  $\square$

The two claims above conclude the proof of statement (i), by observing that  $\lambda$  and  $p$  are, respectively, super-polynomial and polynomial in  $\lambda$ .

(ii) Let  $\Pi = (KG, E, D)$  be a PKE scheme, with message space  $\mathcal{M} = \{0, 1\}^n$ , that is IND-SDA secure. Moreover let  $(\text{Enc}, \text{Dec})$  be a  $(k, n)$ -encoding scheme with the property that the bits of a codeword are individually uniform and  $\tau n$ -wise independent (over the randomness of the encoding). Consider the PKE scheme  $\Pi'' = (KG'', E'', D'')$ , with message space  $\mathcal{M}'' = \{0, 1\}^k$ , obtained from  $\Pi$  as depicted in Figure 7. The key generation is identical in the two schemes, except that  $KG''$  includes a message  $\bar{m} \leftarrow_{\$} \{0, 1\}^k$  in the public key. The encryption algorithm first encodes the plaintext  $m$  obtaining a codeword  $c$ , encrypts  $c$  via  $\text{pk}$ , and appends to the ciphertext two bits  $\beta_1, \beta_2 \in \{0, 1\}$  and a special value  $\alpha \in [n]$  (normally set to 0). The decryption algorithm simply decrypts the ciphertext and runs the decoding algorithm, in case  $\beta_1 = 0$ . Otherwise it checks that the bit in position  $\alpha$  of  $c$  equals  $\beta_2$ ; if this is the case it returns the message  $\bar{m}$ , else it returns  $\perp$ .

**Claim 25.** *There exists a distinguisher  $D''$  such that  $\Delta^{D''}(G_0^{\Pi'',1,n}, G_1^{\Pi'',1,n}) = 1$ .*

PKE Scheme $\Pi'' = (KG'', E'', D'')$		
Key Generation $KG''$ $(\text{pk}, \text{sk}) \leftarrow KG$ $\bar{m} \leftarrow_{\$} \{0, 1\}^k$ $\text{pk}'' = (\text{pk}, \bar{m})$ $\text{sk}'' = \text{sk}$ <b>return</b> $(\text{pk}'', \text{sk}'')$	Encryption $E''_{\text{pk}''}(m)$ $c \leftarrow \text{Enc}(m)$ $e \leftarrow E_{\text{pk}}(c)$ $e'' = 0 \  0 \  0 \  e$ <b>return</b> $e''$	Decryption $D''_{\text{sk}''}(e'')$ $e'' = \beta_1 \  \beta_2 \  \alpha \  e$ $c \leftarrow D_{\text{sk}}(e)$ $m \leftarrow \text{Dec}(c)$ <b>if</b> $\beta_1 = 0$ <b>if</b> $(\beta_2 = 0) \wedge (\text{ctr} = 0)$   <b>return</b> $m$ <b>else</b>   <b>return</b> $\perp$ <b>else</b>   <b>if</b> $(m'[\alpha] = \beta_2)$   <b>return</b> $\bar{m}$   <b>else</b>   <b>return</b> $\perp$

**Figure 7:** PKE scheme  $\Pi''$  in the proof of Theorem 22

*Proof.* Consider the following distinguisher  $D''$  playing the NM-CPA game. At the beginning  $D''$  receives the public key  $(\text{pk}, \bar{m})$ , and issues  $(\text{chall}, m_0, m_1)$ , for some  $m_0, m_1$ , obtaining a challenge  $e'' = (0 \| 0 \| 0 \| e)$ .

Next,  $D''$  defines a sequence of  $n$  ciphertexts  $e_1, \dots, e_n$  and issues  $(\text{dec}, e_1, \dots, e_n)$ . For all  $i \in [n]$ , the  $i$ -th ciphertext  $e_i$  has a type  $(1 \| \beta_{2,i} \| i \| e)$  where  $\beta_{2,i}$  is a guess for  $c[i]$  (i.e., the  $i$ -th bit of the codeword  $c \leftarrow D_{\text{sk}}(e)$ ).

Let  $(m_1, \dots, m_n)$  be the answer from the decryption oracle. Distinguisher  $D''$  defines a codeword  $c = (c[1], \dots, c[n])$  where  $c[i] = \beta_{2,i}$  if  $m = \bar{m}$  (and  $c[i] = 1 - \beta_{2,i}$  otherwise).<sup>18</sup> By inspection, this is exactly the codeword  $c$  encoding the challenge, and thus  $D''$  wins the game with probability 1.<sup>19</sup>  $\square$

**Claim 26.** For all  $q \in \mathbb{N}$ , and for all distinguishers  $D''$ , there exists a distinguisher  $D$  such that

$$\Delta^D(G_0^{\Pi, q, 1}, G_1^{\Pi, q, 1}) \geq \Delta^{D''}(G_0^{\Pi'', q, 1}, G_1^{\Pi'', q, 1}) - 2^{-\tau n}.$$

*Proof.* We build distinguisher  $D$  (based on  $D''$ ) as follows:

1. At the beginning  $D$  samples  $\bar{m} \leftarrow_{\$} \{0, 1\}^k$ , and forwards  $\text{pk}'' = (\text{pk}, \bar{m})$  to  $D''$ , where  $\text{pk}$  is the public key it receives from its own challenger. Furthermore  $D$  samples  $n$  uniform and independent bits  $b = (b_1, \dots, b_n)$ .
2. Upon input  $(\text{chall}, m_0, m_1)$  at the inside,  $D$  defines  $c_0 \leftarrow \text{Enc}(m_0)$  and  $c_1 \leftarrow \text{Enc}(m_1)$  and issues  $(\text{chall}, c_0, c_1)$  to its own challenger. Given the challenge ciphertext  $e$ , distinguisher  $D$  returns  $e'' := (0 \| 0 \| 0 \| e)$  to  $D''$ .
3. For all  $i \in [q]$ , upon input the  $i$ -th decryption query  $(\text{dec}, e'_i)$  from  $D''$ , distinguisher  $D$  parses  $e'_i = (\beta_{1,i} \| \beta_{2,i} \| \alpha_i \| e_i)$  and proceeds as follows:
  - (a) If  $\beta_{1,i} = \beta_{2,i} = 0$  and  $\alpha_i = 0$ , issue  $(\text{dec}, e_i)$  to the decryption oracle receiving a value  $c_i \in \{0, 1\}^n$  and compute  $m_i \leftarrow \text{Dec}(c_i)$ ; return  $m_i$  to  $D''$ .
  - (b) If  $\beta_{1,i} = 0$  and  $\alpha_i \neq 0$  or  $\beta_{2,i} \neq 0$ , output  $\perp$  and self-destruct.
  - (c) If  $\beta_{1,i} = 1$ , check whether  $b[\alpha_i] = \beta_{2,i}$ ; if that is the case return  $\bar{m}$ , and otherwise return  $\perp$  and self-destruct.

<sup>18</sup>Recall that, by definition of  $\Pi''$ , the values  $\bar{m}$  and  $\perp$  are the only possible outcomes for each of the  $m_i$ .

<sup>19</sup>The latter holds true in case the encryption scheme and the encoding scheme have perfect correctness. However, it is straightforward to generalize the proof to the case of a small correctness error.

4.  $D$  outputs whatever  $D''$  does.

Clearly  $D$  perfectly simulates the public key and the challenge ciphertext. As for decryption queries, the simulation for queries of type 3(a) and 3(b) is also perfect. Define the event  $E$  that  $D''$  issues more than  $\tau n$  decryption queries of type 3(c), such that the guess  $\beta_{2,i}$  for  $b[\alpha_i]$  is correct. It is easy to see that in case  $E$  does not happen, the simulation of decryption queries with type 3(c) is correct, because any subset of dimension up to  $\tau n$  of the codeword  $c$  corresponding to the challenge ciphertext is uniformly distributed. Thus, it suffices to bound the probability of the event  $E$  which is at most  $2^{-\tau n}$ . The claim follows.  $\square$

The two claims above imply statement (ii), by setting  $\tau$  to be super-logarithmic in the security parameter  $\lambda$ .<sup>20</sup>  $\square$

## B Necessity of Codes with Secret State

We show that codes without secret state, as defined below, (e.g., those in [17, 16, 1, 19, 10, 7, 2]) cannot achieve (unconditional) non-malleability against parallel tampering—already for the case  $q = 1$ .

**Definition 8** (Code without secret state). *A  $(k, n)$ -code is a pair of algorithms  $(\text{Enc}, \text{Dec})$ , where the (randomized) encoding algorithm  $\text{Enc}$  takes a  $k$ -bit plaintext  $x$  and outputs an  $n$ -bit encoding  $c \leftarrow \text{Enc}(x)$ , and the (randomized) decoding algorithm  $\text{Dec}$  takes an  $n$ -bit encoding  $c$  and outputs a  $k$ -bit plaintext  $x \leftarrow \text{Dec}(c)$  or the special symbol  $\perp$ , indicating an invalid encoding.*

We prove the following theorem:

**Theorem 14.** *Let  $\mathcal{F} := \mathcal{F}_{\text{set}}$ . Let  $(\text{Enc}, \text{Dec})$  be a  $(k, n)$ -code without secret state and noticeable rate. There exists a distinguisher  $D$  asking a single parallel tampering query of size  $n^6$  such that, for all simulators  $\text{sim}$  and all  $n$  large enough,  $\Delta^D(R_{\mathcal{F}}, S_{\mathcal{F}, \text{sim}}) \geq 1/2$ .*

The above impossibility result requires that the rate of the code not be too small (in fact  $n = o(2^{k/6})$  suffices, see below for the exact parameters). The distinguisher  $D$  is inefficient, so it might still be possible to construct a non-malleable code against parallel tampering with only computational security. We leave this as an interesting open question for future research.

The proof of the above theorem follows directly by the following lemma:

**Lemma 27.** *Let  $\mathcal{F} := \mathcal{F}_{\text{set}}$ . Let  $(\text{Enc}, \text{Dec})$  be a  $(k, n)$ -code without secret state, and with correctness error  $\nu < 1/2 - 1/n$ . There exists a distinguisher  $D$  such that, for all simulators  $\text{sim}$ ,*

$$\Delta^D(R_{\mathcal{F}}, S_{\mathcal{F}, \text{sim}}) \geq 1 - (\nu + 2/n + 2ne^{-1/2n} + (2n^6 + 1) \cdot 2^{-k}).$$

*The distinguisher  $D$  asks a single tampering query (i.e.,  $q = 1$ ) with  $p = n^6$ .*

It remains to prove Lemma 27. The proof is non-constructive, meaning that the distinguisher  $D$  depends on some “auxiliary information” that is fixed once the code is given, but might be hard to compute. However, as we show below, such auxiliary information always exists (for any code) provided that the correctness error of the code is not too large (but a good code would typically have a small correctness error).

We discuss some intuition for the proof of Lemma 27. The main idea is to define an extraction algorithm that (almost) always succeeds to extract the encoded message when it interacts with  $R_{\mathcal{F}}$ , but only does so with a small probability when interacting with  $S_{\mathcal{F}, \text{sim}}$  (for any  $\text{sim}$ ). For simplicity, let us first assume that decoding is deterministic and that the code has perfect correctness (i.e.,  $\nu = 0$ ). Define a position  $i \in [n]$  to be *relevant* if there exists a pair of codewords  $(c'_i, c''_i)$ , differing only at position  $i$ , for which decoding  $c'_i$  and  $c''_i$  leads to different values. One can show that the set of all relevant position

<sup>20</sup>Suitable encoding schemes  $(\text{Enc}, \text{Dec})$  with such  $\tau$  exist unconditionally [17].

is not empty (by correctness of the code), and is always fixed once the coding scheme is given (this is because the code has no secret state). Additionally, in order to decode any codeword  $c \in \{0, 1\}^n$ , one needs to know only the values  $c[i]$  for the relevant positions; all other values can be set to 0 and play no role in decoding a codeword.

Consider now the following distinguisher  $D$  that is given a set of pairs  $(c'_i, c''_i)$ , one for each *relevant* position  $i \in [n]$ . At the beginning  $D$  encodes a value  $x$ , which defines a target encoding  $c$ . Next,  $D$  attempts to extract the  $i$ -th *relevant* bit of  $c$  via a tampering query  $f_i \in \mathcal{F}$  that keeps the bit in position  $i$  and replaces all other values with the bits of  $c'_i$  or  $c''_i$  (recall that the two codewords only differ at position  $i$ ). Since  $c'_i$  and  $c''_i$  decode to different values,  $D$  can determine with a single tampering query (of size at most  $n$ ) all relevant values  $c[i]$  with certainty; the non-relevant bits can be set to 0, as they play no role in decoding  $c$ . Distinguisher  $D$  outputs 1 if and only if the above extraction procedure leads to the chosen value  $x$ . Clearly,  $D$  always outputs 1 when interacting with  $R_{\mathcal{F}}$ . On the other hand, one can show that  $D$  almost never outputs 1 when interacting with  $S_{\mathcal{F}, \text{sim}}$ , which concludes the proof.

The impossibility proof extends to the case where the decoding is probabilistic, with correctness error  $\nu > 0$ . That means, in particular, that for a fixed codeword  $c$ , the value  $\text{Dec}(c)$  is a random variable. For each position  $i$  and a parameter  $\mu \in [0, 1]$ , either of the following cases applies: (i) There exists a pair of codewords  $(c'_i, c''_i)$ , differing only at position  $i$ , such that for the statistical distance it holds that  $\Delta(\text{Dec}(c'_i), \text{Dec}(c''_i)) \geq \mu$ ; (ii) For all codewords  $c'_i$  and  $c''_i$  again differing only at position  $i$ ,  $\Delta(\text{Dec}(c'_i), \text{Dec}(c''_i)) < \mu$ . In case (i), position  $i$  is called *relevant*; in case (ii) it is called *non-relevant*. This distinction can then be exploited as in the deterministic case, but with two adaptations:

1. The extraction has to be repeated roughly  $\Theta(\mu^{-3})$  times in order to get a good estimation of the  $i$ -th relevant bit. Since the individual decoding attempts use independent randomness, by the Hoeffding bound we should get a good estimation of that bit after a polynomial number of queries.
2. As before, the decoding is at the end computed by filling the non-relevant positions with sufficiently many 0s. Since the statistical distance for *all* strings with each of the bits flipped in the non-relevant positions varies by at most  $\mu$ , the triangle inequality allows to (lower) bound the probability of actually computing the correct value.

In order to formalize the above intuition, we state a general lemma that will be useful in the sequel. Let  $(\text{Enc}, \text{Dec})$  be a  $(k, n)$ -coding scheme without secret state. Fix a parameter  $\mu \in [0, 1]$ . Define a position  $i \in [n]$  to be  $\mu$ -*relevant* for the encoding scheme if there exists two codewords  $c'_i, c''_i \in \{0, 1\}^n$  that differ *only* in position  $i$ , for which  $\Delta(\text{Dec}(c'_i), \text{Dec}(c''_i)) \geq \mu$ . Let  $\mathcal{R} = \mathcal{R}(\mu) \subseteq [n]$  be the set of all relevant positions.

**Lemma 28.** *Let  $\mu \in [0, 1]$  and consider a  $(k, n)$ -coding scheme  $(\text{Enc}, \text{Dec})$  without secret state, and with correctness error  $\nu \in [0, 1]$ . The following holds:*

- (i) *The set  $\mathcal{R}(\mu)$  is not empty, whenever  $0 \leq \mu < \frac{1-2\nu}{2n}$ .*
- (ii) *Let  $c' \leftarrow \text{Enc}(x')$  for some  $x' \in \{0, 1\}^k$ . For all  $c''$  such that  $c''[i] = c'[i]$  for all  $i \in \mathcal{R}(\mu)$  we have that*

$$\mathbb{P}[\text{Dec}(c'') = x'] \geq 1 - \nu - 2n \cdot \mu.$$

*Proof.* We start by proving statement (i). For the sake of contradiction, assume that  $\mathcal{R}(\mu)$  is empty. This means that for all  $i \in [n]$ , and for all possible pairs of codewords  $c'_i, c''_i$  (differing only at coordinate  $i$ ) we have that  $\Delta(\text{Dec}(c'_i), \text{Dec}(c''_i)) < \mu$ . By the triangle inequality, this implies that any two codewords  $c', c''$  satisfy  $\Delta(\text{Dec}(c'), \text{Dec}(c'')) \leq n \cdot \mu$ . Fix now two values  $x', x'' \in \{0, 1\}^k$ , such that  $x' \neq x''$ . Let  $c' \leftarrow \text{Enc}(x')$  and  $c'' \leftarrow \text{Enc}(x'')$  be the corresponding encodings. We have:

$$\begin{aligned} 2n \cdot \mu &\geq \sum_{x \in \{0, 1\}^k} |\mathbb{P}[\text{Dec}(c'') = x] - \mathbb{P}[\text{Dec}(c') = x]| \\ &\geq \mathbb{P}[\text{Dec}(c'') = x''] - \mathbb{P}[\text{Dec}(c') \neq x'], \end{aligned}$$

where the first inequality follows by definition of statistical distance, and the second inequality follows by the fact that  $\mathbb{P}[\text{Dec}(c') \neq x'] \geq \mathbb{P}[\text{Dec}(c') = x'']$ . Using the correctness property of the code, and our choice of  $\mu < \frac{1-2\nu}{2n}$ , we have obtained

$$\mathbb{P}[\text{Dec}(c') \neq x'] \geq 1 - \nu - 2n \cdot \mu > \nu,$$

a contradiction.

We now show statement (ii). Let  $c', c''$  be as in the statement of the lemma. First observe that, since the code has no secret state, the event that  $c''$  could decode to something different than  $x'$  depends only on the randomness of the decoding process. Without loss of generality, assume that  $c'' \neq c'$  (otherwise there is nothing to prove). By definition of  $\mathcal{R}(\mu)$  and by using the triangle inequality, we get that  $\Delta(\text{Dec}(c'), \text{Dec}(c'')) \leq |\mathcal{R}(\mu)| \cdot \mu \leq n \cdot \mu$ . Hence,

$$\begin{aligned} 2n \cdot \mu &\geq \sum_{x \in \{0,1\}^k} |\mathbb{P}[\text{Dec}(c') = x] - \mathbb{P}[\text{Dec}(c'') = x]| \\ &\geq \mathbb{P}[\text{Dec}(c') = x'] - \mathbb{P}[\text{Dec}(c'') = x'] \\ &\geq 1 - \nu - \mathbb{P}[\text{Dec}(c'') = x'], \end{aligned}$$

which concludes the proof.  $\square$

We now turn to the proof of Lemma 27.

*Proof of Lemma 27.* Fix parameters  $\mu \in [0, 1]$  and  $\rho \in \mathbb{N}$  to be determined later. Let  $\mathcal{R}(\mu) \subseteq [n]$ , with  $|\mathcal{R}(\mu)| := r \leq n$ , the set of all  $\mu$ -relevant positions, and denote with  $(c'_i, c''_i)_{i \in \mathcal{R}(\mu)}$  the corresponding pairs of codewords that are ensured to exist by definition. Additionally, let  $\mathcal{X}_i := \{x \in \{0, 1\}^k : \mathbb{P}[\text{Dec}(c'_i) = x] > \mathbb{P}[\text{Dec}(c''_i) = x]\}$  be the set of all values  $x$  for which the probability that decoding  $c'_i$  leads to  $x$  is larger than the probability that decoding  $c''_i$  leads to  $x$ ; also let  $p'_i := \mathbb{P}[\text{Dec}(c'_i) \in \mathcal{X}_i]$  and  $p''_i := \mathbb{P}[\text{Dec}(c''_i) \in \mathcal{X}_i]$ . We define the “auxiliary information” of the code to be:

$$\text{aux} := \{\mathcal{R}(\mu), (c'_i, c''_i, \mathcal{X}_i, p'_i, p''_i)_{i \in \mathcal{R}(\mu)}\}.$$

Note that information  $\text{aux}$  is fixed once a particular encoding scheme  $(\text{Enc}, \text{Dec})$  is given. This is because the code has no secret state.

Consider the following extraction algorithm  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  (with the above auxiliary information hard-coded), issuing a single parallel tampering query and outputting a value  $\bar{x} \in \{0, 1\}^k \cup \{\perp\}$ .

1. For all  $i \in \mathcal{R}(\mu)$ , consider a function  $f_i \in \mathcal{F}$  being specified via  $(f_i[1], \dots, f_i[n])$  where  $f_i[i] = \text{keep}$  and  $f_i[j]$  is set to  $c'_i[j]$  or  $c''_i[j]$  for all other positions  $j \in [n] \setminus \{i\}$  (recall that  $c'_i$  and  $c''_i$  differ only at position  $i \in \mathcal{R}(\mu)$ ).
2.  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  defines the following parallel tampering query  $(\text{tamper}, (f_i, \dots, f_i)_{i \in \mathcal{R}(\mu)})$ , where we take  $\rho$  copies of each function  $f_i$ . Let  $(x_{1,1}, \dots, x_{1,\rho}, \dots, x_{r,1}, \dots, x_{r,\rho})$  be the answers corresponding to  $\text{Ext}_{\mu, \rho}^{\text{aux}}$ 's query.
3. For each  $i \in \mathcal{R}(\mu)$ ,  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  checks whether

$$\#\{x_{i,j} \in \mathcal{X}_i : j \in [\rho]\} \geq \left(\frac{p'_i}{2} + \frac{p''_i}{2}\right) \rho;$$

if that is the case it sets  $\bar{c}[i] := c'_i[i]$ , and otherwise  $\bar{c}[i] := 1 - c'_i[i] = c''_i[i]$ . All other values  $\bar{c}[i]$ , with  $i \notin \mathcal{R}(\mu)$ , are set to 0.

4. Finally,  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  outputs  $\bar{x} = \text{Dec}(\bar{c})$ .

**Claim 29.** Let  $0 < \mu < \frac{1-2\nu}{2n}$ , and  $\rho \in \mathbb{N}$ . For all  $x \in \{0, 1\}^k$ , let  $\bar{x}$  be the value returned by  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  after tampering with  $c \leftarrow \text{Enc}(x)$ . Then,  $\mathbb{P}[\text{Ext}_{\mu, \rho}^{\text{aux}}$  outputs  $\bar{x} = x] \geq 1 - (\nu + 2n/\mu + 2ne^{-1/2\rho\mu^2})$ .

*Proof.* Note that by Lemma 28, the range of values allowed for the parameter  $\mu$  ensures that the set  $\mathcal{R}(\mu)$  is not empty. We need to show that  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  extracts the bits  $c[i]$  of the target encoding of  $x$  in all  $\mu$ -relevant positions, with high enough probability (the other bits are set to 0 by  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  and play almost no role in decoding  $\bar{c}$ ).

Let  $\tilde{c}_i = f_i(c)$  be the  $i$ -th tampered codeword as defined by  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  in step 2. For all  $j \in [\rho]$ , and for all  $i \in \mathcal{R}(\mu)$ , let  $X'_{i,j}$  (resp.  $X''_{i,j}$ ) be a binary random variable which equals 1 if and only if  $x_{i,j}$  happens to be in the set  $\mathcal{X}_i$ , where  $x_{i,j}$  is the value obtained by decoding  $\tilde{c}_i = c'_i$  (resp.  $\tilde{c}_i = c''_i$ ). Note that, for each  $i \in \mathcal{R}(\mu)$ , the random variables  $X'_{i,j}$  (resp.  $X''_{i,j}$ ) are independent and follow the Bernoulli distribution with parameter  $p'_i$  (resp.  $p''_i$ ). Let  $X'_i = \sum_{j=1}^{\rho} X'_{i,j}$  and similarly  $X''_i = \sum_{j=1}^{\rho} X''_{i,j}$ . We bound the probability that  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  extracts the  $i$ -th relevant bit of the target encoding incorrectly as follows:

$$\begin{aligned} \mathbb{P}[\text{Ext}_{\mu, \rho}^{\text{aux}} \text{ sets } \bar{c}[i] \text{ incorrectly}] &\leq \mathbb{P}[\text{Ext}_{\mu, \rho}^{\text{aux}} \text{ sets } \bar{c}[i] \text{ incorrectly} | \tilde{c}_i = c'_i] \\ &\quad + \mathbb{P}[\text{Ext}_{\mu, \rho}^{\text{aux}} \text{ sets } \bar{c}[i] \text{ incorrectly} | \tilde{c}_i = c''_i] \\ &= \mathbb{P}[X'_i \leq (p'_i + p''_i)\rho/2] + \mathbb{P}[X''_i \geq (p'_i + p''_i)\rho/2] \\ &\leq \mathbb{P}[X'_i \leq (p'_i - \mu/2)\rho] + \mathbb{P}[X''_i \geq (p''_i + \mu/2)\rho] \\ &\leq 2e^{-1/2\rho\mu^2}, \end{aligned}$$

where the first inequality comes from the fact that  $\mu \leq p'_i - p''_i$ , and the second inequality follows from the Hoeffding inequality.<sup>21</sup> By the union bound, we get that the probability of extracting *at least one position incorrectly* is bounded by  $2ne^{-1/2\rho\mu^2}$ , so the probability that  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  extracts all  $\mu$ -relevant positions correctly is at least  $1 - 2ne^{-1/2\rho\mu^2}$ .

Finally, we observe that the probability that decoding  $\bar{c}$  as defined by  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  in step 3 leads to the correct value  $x$  is at least  $1 - \nu - 2n/\mu$  (see Lemma 28). The claim follows.  $\square$

Define now the following distinguisher  $D$  (depending on  $\text{Ext}_{\mu, \rho}^{\text{aux}}$ ). The distinguisher chooses  $x \in \{0, 1\}^k$  and issues  $(\text{encode}, x)$  to the oracle it has access to. Then it fixes  $\mu = n^{-2}$  and  $\rho = n^5$ , and it lets  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  interact with that oracle. In case at least one of the values  $\{x_{i,j}\}_{i \in \mathcal{R}(\mu), j \in [\rho]}$  seen by  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  in step 2 happens to be equal to  $x$ ,  $D$  outputs 0. Otherwise, whenever  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  returns a value  $\bar{x} \in \{0, 1\}^k \cup \{\perp\}$ ,  $D$  outputs 1 if and only if  $\bar{x} = x$ .

**Claim 30.**  $\mathbb{P}[D(R_{\mathcal{F}}) = 1] \geq 1 - (\nu + 2/n + 2ne^{-1/2n} + n^6 \cdot 2^{-k+1})$ .

*Proof.* The statement follows by definition of  $D$ . Note that  $D$  outputs 1 provided that  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  returns the value  $x$  chosen uniformly at random by  $D$ , except in case it happens that one of the answers  $x_{i,j}$  corresponding to  $\text{Ext}_{\mu, \rho}^{\text{aux}}$ 's tampering query happens to be equal to  $x$ . Denote by  $E$  the latter event. Since we already have a lower bound on the probability of success of  $\text{Ext}_{\mu, \rho}^{\text{aux}}$  (see Claim 29), it suffices to bound the probability of the event  $E$ .

Note that for each tampering function  $f_i \in \mathcal{F}$  specified by  $\text{Ext}_{\mu, \rho}^{\text{aux}}$ , we have that the tampered codeword  $\tilde{c}_i = f_i(c)$  overwrites all the bits with known values but the bit in position  $i$ , which is copied from the original codeword  $c$ . It follows that the output of such tampering query can decrease the entropy of  $x$  by at most one, and hence the probability that each  $x_{i,j}$  happens to be equal to  $x$  is at most  $2^{-k+1}$ . By a union bound, we get that  $\mathbb{P}[E] \leq r\rho \cdot 2^{-k+1} \leq n\rho \cdot 2^{-k+1}$ .

<sup>21</sup>The general form is: Let  $X_1, \dots, X_n$  be i.i.d. with  $X_i \sim \text{Be}(p)$ . Then, for  $X := \sum_i X_i$ , and for any  $\varepsilon \in (0, 1]$ ,

$$\mathbb{P}[X \leq (p - \varepsilon)n] \leq e^{-2\varepsilon^2 n} \quad \text{and} \quad \mathbb{P}[X \geq (p + \varepsilon)n] \leq e^{-2\varepsilon^2 n}.$$



The claim now follows by the choice of  $\mu = n^{-2}$  and  $\rho = n^5$ , which in particular ensures that  $\mathcal{R}(\mu)$  is not empty (see Lemma 28) by our assumption that  $\nu < 1/2 - 1/n$ .  $\square$

**Claim 31.** *For all simulators  $\text{sim}$ , we have that  $\mathbb{P}[D(S_{\mathcal{F},\text{sim}}) = 1] \leq 2^{-k}$ .*

*Proof.* We first observe that we can assume, without loss of generality, that  $\text{sim}$  never outputs `same`. In fact, if  $\text{sim}$  happens to return `same` for any of the queries  $f_i$  defined by  $D$  (via  $\text{Ext}_{\mu,\rho}^{\text{aux}}$ ), then the oracle  $S_{\mathcal{F},\text{sim}}$  would replace the output of the simulator with  $x$ . By definition in this case  $D$  outputs 1 with probability 0.

On the other hand, assume that  $\text{sim}$  never outputs `same`. In such a case the view of  $\text{Ext}_{\mu,\rho}^{\text{aux}}$  is completely independent of  $x$ , and so the extractor will output  $\bar{x} = x$  with probability at most  $2^{-k}$ . The claim follows.  $\square$

Lemma 27 follows by combining Claim 30 and Claim 31.  $\square$

## C Composing Non-Malleable Codes, PKE, and MACs

This section contains the proof of Theorem 3, which is split into two parts. First, we prove that the PKE scheme  $\Pi'$  resulting from combining a single-bit PKE  $\Pi$  and a non-malleable code with secret state  $(\text{Gen}, \text{Enc}, \text{Dec})$  as shown in Section 4.2 (cf. Figure 4) is replayable NM-SDA secure (NM-RSDA); the proof is based on the corresponding one in [10] for IND-SDA security. Then, we show that a MAC-based transformation suggested by [6] to obtain IND-CCA security from IND-RCCA security also works in our setting, i.e., the transformation applied to  $\Pi'$  yields a fully NM-SDA secure PKE scheme  $\Pi''$ . All results in this section also apply to NM-CPA security. Overall, we prove:

**Theorem 3.** *Let  $q, p \in \mathbb{N}$  and  $\Pi$  be a  $(t + t_{\text{1bit}}, q, p, \varepsilon_{\text{1bit}})$ -NM-SDA-secure 1-bit PKE scheme,  $(T, V)$  a  $(t + t_{\text{mac}}, 1, qp, \varepsilon_{\text{mac}})$ -MAC, and  $(\text{Gen}, \text{Enc}, \text{Dec})$  a  $(\mathcal{F}_{\text{set}}, q, p, \varepsilon_{\text{nmc}})$ -non-malleable  $(k, n)$ -code with secret state. Then,  $\Pi''$  is  $(t, q, p, \varepsilon)$ -NM-SDA-secure PKE scheme with*

$$\varepsilon = 2(3(n\varepsilon_{\text{1bit}} + \varepsilon_{\text{nmc}}) + qp \cdot 2^{-\ell} + \varepsilon_{\text{mac}}),$$

where  $t_{\text{1bit}}$  and  $t_{\text{mac}}$  are the overheads incurred by the corresponding reductions and  $\ell$  is the length of a verification key for the MAC.

### C.1 Replayable NM-SDA Security

The notion of *replayable CCA* security was introduced by Canetti *et al.* [6] to deal with the artificial strictness of CCA security. Intuitively, it potentially allows an attacker to maul a target ciphertext into one that decrypts to the *same* message.<sup>22</sup> This idea carries over seamlessly to the definition of NM-SDA security; the corresponding distinguishing game  $G_b^{\Pi, \text{nm-rsda}}$  is obtained by changing  $G_b^{\Pi, \text{nm-sda}}$  (cf. Figure 2) to answer `test` whenever a ciphertext  $e^{(j)}$  decrypts to  $m_0$  or  $m_1$  (instead of only when  $e^{(j)}$  equals the challenge ciphertext).

**Definition 9.** *A PKE scheme  $\Pi$  is replayable  $(t, q, p, \varepsilon)$ -NM-SDA-secure (NM-RSDA) if for all distinguishers  $D$  with running time at most  $t$  and making at most  $q$  decryption queries of size at most  $p$  each,*

$$\Delta^D(G_0^{\Pi, \text{nm-rsda}}, G_1^{\Pi, \text{nm-rsda}}) \leq \varepsilon.$$

<sup>22</sup>In contrast, full CCA security requires that any ciphertext created by the attacker (other than the target ciphertext) decrypt to an independent message.

## C.2 Non-Malleable Codes and PKE

In this section we show that the PKE scheme  $\Pi'$  is NM-RSDA if the underlying single-bit scheme  $\Pi$  is NM-SDA secure. Concretely, we prove:

**Theorem 32.** *Let  $q, p \in \mathbb{N}$  and  $\Pi$  be a  $(t_{\text{rsda}} + t_{1\text{bit}}, q, p, \varepsilon_{1\text{bit}})$ -NM-SDA-secure 1-bit PKE scheme and let  $(\text{Gen}, \text{Enc}, \text{Dec})$  be  $(\mathcal{F}_{\text{set}}, q, p, \varepsilon_{\text{nmc}})$ -non-malleable. Then,  $\Pi'$  is  $(t_{\text{rsda}}, q, p, \varepsilon_{\text{rsda}})$ -NM-RSDA-secure PKE scheme with*

$$\varepsilon_{\text{rsda}} = 2(n\varepsilon_{1\text{bit}} + \varepsilon_{\text{nmc}}),$$

where  $t_{1\text{bit}}$  represents the overhead incurred by the reductions.

The proof follows directly from the following lemma:

**Lemma 33.** *For  $b \in \{0, 1\}$  and  $i \in [n]$ , there exist reductions  $R_{b,i}(\cdot)$  and  $W_b(\cdot)$  such that for all distinguishers  $D$ ,*

$$\Delta^D(G_0^{\Pi', \text{nm-rsda}}, G_1^{\Pi', \text{nm-rsda}}) \leq \sum_{b,i} \Delta^{D(R_{b,i}(\cdot))}(G_0^{\Pi, \text{nm-sda}}, G_1^{\Pi, \text{nm-sda}}) + \sum_b \Delta^{D(W_b(\cdot))}(R_{\mathcal{F}}, S_{\mathcal{F}, \text{sim}}),$$

where  $\text{sim}$  is the simulator for the non-malleable code. Moreover, all reductions preserve the number  $q$  and the size  $p$  of the queries.

*Proof.* Let  $t_{1\text{bit}}$  be the maximal occurring overhead caused by the reductions  $R_{b,i}(\cdot)$ . Fix a distinguisher  $D$  having running time  $t_{\text{rsda}}$  and making at most  $q$  decryption queries of size at most  $p$ . Due to the preservation property of the reductions,  $\Delta^{D(R_{b,i}(\cdot))}(G_0^{\Pi, \text{nm-sda}}, G_1^{\Pi, \text{nm-sda}}) \leq \varepsilon_{1\text{bit}}$  and  $\Delta^{D(W_b(\cdot))}(R_{\mathcal{F}}, S_{\mathcal{F}, \text{sim}}) \leq \varepsilon_{\text{nmc}}$ , which completes the proof.  $\square$

Towards a proof of Lemma 33, consider the following hybrids for  $b \in \{0, 1\}$  and  $i \in [n]$ :  $H_{b,i}$  proceeds as  $G_b^{\Pi', \text{nm-rsda}}$  except that the challenge query  $(\text{chall}, m_0, m_1)$  and decryption queries  $(\text{dec}, e^{(1)}, \dots, e^{(p)})$  are handled differently:

- **Challenge query:** The first  $i$  bits of the encoding  $c = (c[1], \dots, c[n])$  of  $m_b$  are replaced by uniformly random and independent bits. The resulting  $n$ -bit string is then encrypted bit-wise (as done by  $E'$ ). This results in the challenge ciphertext  $e^* = (e_1^*, \dots, e_n^*)$ .
- **Decryption query:** Every component  $e^{(l)} = (e'_1, \dots, e'_n)$  is answered as follows: Hybrid  $H_{b,i}$  computes  $c' = (c'[1], \dots, c'[n])$ , where

$$c'[i] = \begin{cases} c[j] & \text{if } e'_j = e_j^*, \text{ and} \\ D_{\text{sk}_j}(e'_j) & \text{otherwise.} \end{cases}$$

Then,  $H_{b,i}$  outputs  $\text{Dec}(c', s)$  as the answer to the component of the decryption query.<sup>23</sup>

Let  $H_{b,0} := G_b^{\Pi', \text{nm-rsda}}$ .

**Lemma 34.** *For all  $b \in \{0, 1\}$  and  $i \in [n]$ , there exist a reduction  $R_{b,i}(\cdot)$  such that for all  $D$*

$$\Delta^D(H_{b,i-1}, H_{b,i}) = \Delta^{D(R_{b,i}(\cdot))}(G_0^{\Pi, \text{nm-sda}}, G_1^{\Pi, \text{nm-sda}}).$$

*Proof.* Fix  $b$  and  $i$ . Hybrid  $R_{b,i}(\cdot)$  works as follows: Initially, it generates the secret state  $s \leftarrow \text{Gen}$  and  $n-1$  key pairs  $(\text{pk}_j, \text{sk}_j)$  for  $j \in [n] \setminus \{i\}$ , obtains  $\text{pk}_i$  (but not  $\text{sk}_i$ ) from the oracle (from  $G_0^{\Pi, \text{nm-sda}}$  or

<sup>23</sup>Assume here and below that  $\text{Dec}(c') = \perp$  if any of the bits  $c'[j]$  equal  $\perp$ .

$G_0^{\text{II,nm-sda}}$ ), and outputs  $\text{pk} := (\text{pk}_1, \dots, \text{pk}_n)$ . When it receives  $(\text{chall}, m_0, m_1)$ , it computes an encoding  $c = (c[1], \dots, c[n]) \leftarrow \text{Enc}(m_b)$ . Then, it chooses  $i$  random bits  $\tilde{c}[1], \dots, \tilde{c}[i]$  and computes

$$e_j^* = \begin{cases} E_{\text{pk}_j}(\tilde{c}[j]) & \text{for } j < i, \text{ and} \\ E_{\text{pk}_j}(c[j]) & \text{for } j > i. \end{cases}$$

Moreover, it outputs  $(\text{chall}, c[i], \tilde{c}[i])$  to its oracle and obtains a ciphertext  $e_i^*$ . It finally returns  $e^* = (e_1^*, \dots, e_n^*)$ .

When  $R_{b,i}(\cdot)$  receives a (parallel) decryption query, for each component  $e' = (e'_1, \dots, e'_n)$  it proceeds as follows: For  $j \neq i$ , it computes  $c'[j]$  as  $H_{b,i}$  does. Moreover, if  $e'_i = e_i^*$ , it sets  $c'[i] \leftarrow c[i]$ . Otherwise, it outputs  $(\text{dec}, e'_i)$  to its oracle and obtains the answer  $c'[i]$ .<sup>24</sup> Then, it computes  $m' \leftarrow \text{Dec}(c')$ . The answer to the component of the decryption query is  $m'$ , unless  $m' \in \{m_0, m_1\}$ , in which case the it is  $\text{test}$ . If one of the component answers is  $\perp$ ,  $R_{b,i}(\cdot)$  implements the self-destruct mode, i.e., answers all future queries by  $\perp$ .

Consider  $R_{b,i}(G_0^{\text{II,nm-sda}})$  and  $H_{b,i-1}$ . Both generate the public key in the same fashion. As to the challenge ciphertext, the first  $i-1$  ciphertext components  $e_j$  generated by  $R_{b,i}(G_0^{\text{II,nm-sda}})$  are encryptions of random bits  $\tilde{c}[j]$ , whereas the  $i^{\text{th}}$  and the remaining components are encryptions of the corresponding bits of an encoding of  $m_b$  (generated by  $G_0^{\text{II,nm-sda}}$  and  $R_{b,i}(\cdot)$ , respectively). The same is true for  $H_{b,i-1}$ . The answer to a decryption query component sent to  $R_{b,i}(G_0^{\text{II,nm-sda}})$  is  $\text{Dec}(c')$  for  $c' = (c'[1], \dots, c'[n])$ , where  $c'[j] = D_{\text{sk}_j}(e'_j)$  unless  $j < i$  and  $e'_j = e_j$ , in which case  $c'[j] = \tilde{c}[j]$ . Again, the same holds for  $H_{b,i-1}$ . Moreover, both  $R_{b,i}(G_0^{\text{II,nm-sda}})$  and  $H_{b,i-1}$  answer  $\text{test}$  if  $\text{Dec}(c') \in \{m_0, m_1\}$ . Thus, they behave identically.

$R_{b,i}(G_1^{\text{II,nm-sda}})$  and  $H_{b,i}$  are compared similarly. This concludes the proof.  $\square$

**Lemma 35.** *For  $b \in \{0, 1\}$ , there exists a wrapper  $W_b(\cdot)$  such that*

1.  $W_b(R_{\mathcal{F}})$  behaves as  $H_{b,n}$ , and
2.  $W_0(S_{\mathcal{F},\text{sim}})$  and  $W_1(S_{\mathcal{F},\text{sim}})$  behave identically.

*Proof.* Wrapper  $W_b(\cdot)$  works as follows: Initially, it generates  $n$  key pairs  $(\text{pk}_i, \text{sk}_i)$  for  $i \in [n]$  and outputs  $\text{pk} := (\text{pk}_1, \dots, \text{pk}_n)$ . When it receives  $(\text{chall}, m_0, m_1)$ , it picks  $n$  random values  $\tilde{c}[1], \dots, \tilde{c}[n]$ , computes  $e_i^* \leftarrow E_{\text{pk}}(\tilde{c}[i])$  for  $i = 1, \dots, n$ , and returns  $e = (e_1, \dots, e_n)$ . Additionally, it outputs  $(\text{encode}, m_b)$  to its oracle.

When it gets a (parallel) decryption query, for every component  $e' = (e'_1, \dots, e'_n)$ , it proceeds as follows: First, it creates a tamper query  $f = (f[1], \dots, f[n])$  where

$$f[i] = \begin{cases} \text{zero} & \text{if } e'_i \neq e_i^* \text{ and } D_{\text{sk}_i}(e'_i) = 0, \\ \text{one} & \text{if } e'_i \neq e_i^* \text{ and } D_{\text{sk}_i}(e'_i) = 1, \text{ and} \\ \text{keep} & \text{if } e'_i = e_i^*. \end{cases}$$

Then, it outputs  $(\text{tamper}, f)$  to its oracle and obtains an answer  $x'$ . If  $x' \in \{m_0, m_1\}$ , the answer to the component query  $\text{test}$ .<sup>25</sup> Otherwise, it is  $x'$ . If one of the component answers is  $\perp$ ,  $W_b(\cdot)$  implements the self-destruct mode, i.e., answers all future queries by  $\perp$ .

Consider  $W_b(R_{\mathcal{F}})$  and  $H_{b,n}$ . Both generate the public key in the same fashion. Furthermore, in either case, the challenge ciphertext consists of  $n$  encryptions of random bits. Finally, both answer a decryption query by applying the same tamper function to an encoding of  $m_b$  before decoding it. When the decoding of the tampered codeword results in  $m_0$  or  $m_1$ , both answer  $\text{test}$ . Therefore, they behave identically.

<sup>24</sup>In fact, it is important that  $R_{b,i}(\cdot)$  output a single parallel decryption query containing all  $e'_i$  for the individual components; but it is less cumbersome to describe how individual components are handled.

<sup>25</sup>Again,  $W_b(\cdot)$  needs to output a single parallel tamper query containing the tamper functions  $f$  for the individual components.

Due to the fact that `test` is output when a decryption query results in  $m_0$  or  $m_1$ , the observable behavior is the same in  $W_0(S_{\mathcal{F},\text{sim}})$  and  $W_1(S_{\mathcal{F},\text{sim}})$ .<sup>26</sup> □

*Proof (of Lemma 33).* Lemma 33 follows using a triangle inequality. Specifically, for any distinguisher  $D$ ,

$$\begin{aligned} \Delta^D(G_0^{\Pi',\text{nm-rsda}}, G_1^{\Pi',\text{nm-rsda}}) &\leq \sum_i \Delta^D(H_{0,i-1}, H_{0,i}) + \Delta^D(W_0(R_{\mathcal{F}}), W_0(S_{\mathcal{F},\text{sim}})) \\ &\quad + \Delta^D(W_1(S_{\mathcal{F},\text{sim}}), W_1(R_{\mathcal{F}})) + \sum_i \Delta^D(H_{1,i-1}, H_{1,i}) \\ &\leq \sum_{b,i} \Delta^D(R_{b,i}(G_0^{\Pi,\text{nm-sda}}), R_{b,i}(G_1^{\Pi,\text{nm-sda}})) \\ &\quad + \sum_b \Delta^{D(W_b(\cdot))}(R_{\mathcal{F}}, S_{\mathcal{F},\text{sim}}). \end{aligned}$$

□

### C.3 From Replayable to Full NM-SDA Security

MESSAGE AUTHENTICATION CODES. A *message authentication code (MAC)* is a pair of algorithms  $(T, V)$ , where the tagging algorithm  $T$  takes as input a message  $m$  and a key  $K$  and outputs a tag  $t \leftarrow T_K(m)$  and where the verification algorithm  $V$  takes a key  $K$ , a message  $m$ , and a tag  $t$  and outputs a bit  $V_K(m, t)$ .

MAC security is defined using the following game  $G^{\text{mac}}$  played by an adversary  $A$ : Initially, the game chooses a random key  $K$ . Then,  $A$  gets access to a tagging oracle, which returns a tag  $t \leftarrow T_K(m)$  when given a message  $m$ , and to a verification oracle, which outputs  $V_K(m, t)$  when given a message  $m$  and a tag  $K$ . The adversary wins the game if he submits to the verification oracle a pair  $(m, t)$  that is not a query-answer pair for the tagging oracle and for which  $V_K(m, t) = 1$ .

**Definition 10.** A MAC  $\Sigma$  is  $(t, u, v, \varepsilon)$ -secure if for all adversaries  $A$  with running time at most  $t$ , making at most  $u$  tag queries and at most  $v$  verification queries,  $\Gamma^A(G^{\text{mac}}) \leq \varepsilon$ .

An NM-RSDA-secure PKE scheme  $\Pi' = (KG', E', D')$  and a message-authentication code (MAC)  $(V, T)$  can be combined as follows to obtain a fully NM-SDA-secure scheme  $\Pi''$  [6]: The key generation remains unchanged. To encrypt a message  $m$ , the new encryption algorithm first chooses a key  $K$  for the MAC and computes an encryption  $e_1 \leftarrow E'_{\text{pk}}(m \parallel K)$  and  $e_2 \leftarrow T_K(e_1)$ ; the ciphertext is  $(e_1, e_2)$ . The new decryption algorithm decrypts  $e_1$  to  $(m, K)$  and verifies the tag  $e_2$ . If the tag is valid, the decryption algorithm outputs  $m$ ; otherwise, it outputs  $\perp$ .

**Theorem 36.** Let  $\Pi'$  be a  $(t+t_{\text{rsda}}, q, p, \varepsilon_{\text{rsda}})$ -NM-RSDA secure PKE scheme and  $(V, T)$  a  $(t+t_{\text{mac}}, \varepsilon_{\text{mac}})$ -secure MAC. Then,  $\Pi''$  is a  $(t, q, p, \varepsilon)$ -NM-SDA-secure PKE scheme for

$$\varepsilon \leq 2(\varepsilon_{\text{rsda}} + qp \cdot 2^{-\ell} + \varepsilon_{\text{mac}}) + \varepsilon_{\text{rsda}},$$

where  $\ell$  is the length of the MAC key.

The theorem follows from the following lemma:

**Lemma 37.** For  $b \in \{0, 1\}$  there exist reductions  $R_b(\cdot)$ ,  $R'(\cdot)$ , and  $R''_b(\cdot)$ , such that for all distinguishers  $D$ ,

$$\begin{aligned} \Delta^D(G_0^{\Pi'',\text{nm-sda}}, G_1^{\Pi'',\text{nm-sda}}) &\leq \sum_b \left( \Delta^{D(R_b(\cdot))}(G_b^{\Pi',\text{nm-rsda}}, G_1^{\Pi',\text{nm-rsda}}) + qp \cdot 2^{-\ell} + \Gamma^{D(R''_b(\cdot))}(G^{\text{mac}}) \right) \\ &\quad + \Delta^{D(R'(\cdot))}(G_0^{\Pi',\text{nm-rsda}}, G_1^{\Pi',\text{nm-rsda}}). \end{aligned}$$

---

<sup>26</sup>This is where the proof reflects that  $\Pi'$  is only NM-RSDA secure.

where  $\ell$  is the length of the MAC key. Moreover, reductions  $R_b(\cdot)$  and  $R'(\cdot)$  preserve the number  $q$  and the size  $p$  of the queries, and reduction  $R_b''(\cdot)$  asks a single tag query and  $qp$  verification queries.

*Proof.* Let  $t_{\text{rsda}}$  be the maximal occurring overhead caused by the reductions  $R_b(\cdot)$ ,  $R'(\cdot)$  and  $t_{\text{mac}}$  that by the reductions  $R_b''(\cdot)$ . Fix a distinguisher  $D$  having running time  $t_{\text{rsda}}$  and making at most  $q$  decryption queries of size at most  $p$ . Due to the preservation properties of the above reductions, the distinguishing advantages on  $G_b^{\Pi', \text{nm-rsda}}$  are at most  $\varepsilon_{\text{rsda}}$  and  $\Gamma^{D(R_b''(\cdot))}(G^{\text{mac}})$  is at most  $\varepsilon_{\text{mac}}$ .  $\square$

**HYBRID 1.** The first hybrid  $H_b$  captures the fact that the MAC key in the challenge ciphertext is computationally hidden; it differs from  $G_b^{\Pi'', \text{nm-sda}}$  as follows:

- It generates the challenge ciphertext using two independent MAC keys  $K^*$  and  $K$ , i.e.,  $(e_1^*, e_2^*) \leftarrow (E'_{\text{pk}}(m_b \| K^*), T_K(e_1^*))$ .
- When answering (components of parallel) decryption queries  $(e'_1, e'_2) \leftarrow (E'_{\text{pk}}(m_b \| K'), e'_2)$ , if  $K' = K^*$ , the tag is verified using  $K$  instead of  $K^*$ .

**Lemma 38.** *There exists a reduction  $R_b(\cdot)$  such that for all distinguishers  $D$  asking at most  $q$  parallel queries of size at most  $p$  each,*

$$\Delta^D(G_b^{\Pi'', \text{nm-sda}}, H_b) \leq \Delta^{D(R_b(\cdot))}(G_0^{\Pi', \text{nm-rsda}}, G_1^{\Pi', \text{nm-rsda}}) + qp \cdot 2^{-\ell},$$

where  $\ell$  is the length of the MAC key.

*Proof (sketch).* Initially, reduction  $R_b(\cdot)$  outputs (to  $D$ ) the public key obtained from its oracle. When it gets  $(\text{chall}, m_0, m_1)$ , it outputs  $((\text{chall}, m_b \| K, m_b \| K^*))$  to its oracle and gets a response  $e_1^*$ . Then, it computes  $e_2^* \leftarrow T_K(e_1^*)$  and outputs  $(e_1^*, e_2^*)$ . As long as no self-destruct has occurred,  $R_b(\cdot)$  answers (components of parallel) decryption queries  $(e'_1, e'_2)$  (different from the challenge ciphertext) as follows: It outputs  $(\text{dec}, e'_1)$  to its oracle. If the answer is **test**,  $H_b$  verifies the tag  $e'_2$  with  $K$  and returns  $m_b$  to  $D$  if it is valid. If the answer is  $m' \| K'$ ,  $H_b$  verifies the tag with  $K'$  and returns  $m'$  if it is valid.

By inspection one observes that  $R_b(G_0^{\Pi', \text{nm-rsda}})$  behaves as  $G_b^{\Pi'', \text{nm-sda}}$  unless  $D$  asks a query  $(e'_1, e'_2)$  where  $e'_1$  is an encryption of a message concatenated with  $K^*$ ; however, since the view of  $D$  when interacting with  $R_b(G_0^{\Pi', \text{nm-rsda}})$  is independent of  $K^*$ , the probability of this event is bounded by  $2^{-\ell}$ .

On the other hand, observe that  $R_b(G_1^{\Pi', \text{nm-rsda}})$  behaves exactly as hybrid  $H_b$ .  $\square$

**HYBRID 2.** The second hybrid  $H'_b$  behaves as  $H_b$  except that queries  $(e'_1, e'_2)$  where  $e'_1$  contains  $K^*$  are always rejected.

**Lemma 39.** *There exists a reduction  $R_b''(\cdot)$  such that for all distinguishers  $D$ ,*

$$\Delta^D(H_b, H'_b) \leq \Gamma^{D(R_b''(\cdot))}(G^{\text{mac}}).$$

*Proof.*  $R_b''(\cdot)$  is a standard reduction to the strong unforgeability of the MAC.  $\square$

**REDUCTION TO NM-RSDA.** Distinguishing  $G_0^{\Pi', \text{nm-rsda}}$  and  $G_1^{\Pi', \text{nm-rsda}}$  can now be reduced to distinguishing  $H'_0$  and  $H'_1$ .

**Lemma 40.** *There exists a reduction  $R'(\cdot)$  such that for all distinguishers  $D$ ,*

$$\Delta^D(H'_0, H'_1) = \Delta^{DR'(\cdot)}(G_0^{\Pi', \text{nm-rsda}}, G_1^{\Pi', \text{nm-rsda}}).$$

*Proof (sketch).* The reduction translates between the NM-SDA game for  $\Pi''$  and the NM-RSDA game for  $\Pi'$ , using the fact that decryption queries for which the first component contains  $K^*$  can be rejected. In particular, when the NM-RSDA game outputs **test**, a ciphertext can be rejected.  $\square$

## D Miscellaneous

### D.1 Chernoff Bound

We make use of the following Chernoff bound.

**Theorem 41.** *Let  $X_1, \dots, X_n$  be i.i.d. with  $X_i \sim \text{Be}(p_i)$ . Then, for  $X := \sum_i X_i$  and  $\mu := \sum_i p_i$ ,*

$$\mathbb{P}[X \leq (1 - \varepsilon)\mu] \leq e^{-\mu\varepsilon^2/2}$$

for any  $\varepsilon \in (0, 1]$ .

### D.2 Plotkin Bound

The following theorem allows to bound the number of codewords of a code over a binary alphabet with relative minimum distance  $\delta > 1/2$ .

**Theorem 42.** *For a code over a binary alphabet with block length  $n$  and distance  $d \geq \frac{n}{2} + 1$ , the maximum number of codewords is*

$$A(n, d) \leq \frac{d}{d - \frac{n}{2}} \leq 1 + \frac{1}{2\varepsilon}$$

where  $\varepsilon = \frac{d}{n} - \frac{1}{2}$ .

A proof can be found in [26, p. 41].