Continuously Non-Malleable Codes against Bounded-Depth Tampering

Gianluca Brian[1] * Sebastian Faust[2] † Elena Micheli[2] † Daniele Venturi[1] ‡

[1] Sapienza University of Rome, Rome, Italy
[2] Technische Universität Darmstadt, Darmstadt, Germany

September 16, 2022

Abstract

Non-malleable codes (Dziembowski, Pietrzak and Wichs, ICS 2010 & JACM 2018) allow protecting arbitrary cryptographic primitives against related-key attacks (RKAs). Even when using codes that are guaranteed to be non-malleable against a *single* tampering attempt, one obtains RKA security against poly-many tampering attacks at the price of assuming perfect memory erasures. In contrast, *continuously* non-malleable codes (Faust, Mukherjee, Nielsen and Venturi, TCC 2014) do not suffer from this limitation, as the non-malleability guarantee holds against *poly-many* tampering attempts. Unfortunately, there are only a handful of constructions of continuously non-malleable codes, while standard non-malleable codes are known for a large variety of tampering families including, e.g., NC0 and decision-tree tampering, AC0, and recently even bounded polynomial-depth tampering. We change this state of affairs by providing the first constructions of continuously non-malleable codes in the following natural settings:

- Against decision-tree tampering, where, in each tampering attempt, every bit of the tampered codeword can be set arbitrarily after adaptively reading up to d locations within the input codeword. Our scheme is in the plain model, can be instantiated assuming the existence of one-way functions, and tolerates tampering by decision trees of depth $d = O(n^{1/8})$, where n is the length of the codeword. Notably, this class includes NC0.
- Against bounded polynomial-depth tampering, where in each tampering attempt the adversary can select any tampering function that can be computed by a circuit of bounded polynomial depth (and unbounded polynomial size). Our scheme is in the common reference string model, and can be instantiated assuming the existence of time-lock puzzles and simulation-extractable (succinct) non-interactive zero-knowledge proofs.

*Supported by grant SPECTRA from Sapienza University of Rome. This work was partly done while G. Brian was visiting the University of Warsaw, Poland, supported by the Copernicus Award (agreement no. COP/01/2020) from the Foundation for Polish Science and by the Premia na Horyzoncie grant (agreement no. 512681/PnH2/2021) from the Polish Ministry of Education and Science.

†This work has been funded by the German Research Foundation (DFG) CRC 1119 CROSSING (project S7), by the German Federal Ministry of Education and Research and the Hessen State Ministry for Higher Education, Research and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE.

‡Supported by grant SPECTRA from Sapienza University of Rome.

Contents

1 Introduction

Related-key attacks (RKAs) allow an adversary to break security of a cryptographic primitive by invoking it under one or more keys that satisfy known relations. Such attacks were first introduced as a tool for the cryptanalysis of blockciphers [Knu93, Bih94], but can also be mounted in practice thanks to the ability of attackers to influence secret keys via tampering attacks [BDL97, BS97, GLM+04].

Theoretically, we can model $\mathcal{F}$-RKA security of a given cryptographic primitive as follows: The attacker can choose multiple tampering functions $f_1, f_2, \ldots$ within a family of allowed

manipulations $\mathcal{F}$ of the secret key, and later observe the effect of such changes at the output by invoking the primitive on chosen inputs. An elegant solution to the problem of constructing $\mathcal{F}$-RKA-secure cryptoschemes is provided by *non-malleable codes* [DPW10]. Intuitively, an $\mathcal{F}$-non-malleable code allows us to encode a message so that a modified codeword via a function $f \in \mathcal{F}$ either decodes to the same message or to a completely unrelated value. In the application to RKA security, we simply encode the secret key κ and store the corresponding codeword γ in memory. A RKA changes the memory content to $\tilde{\gamma} = f(\gamma)$ for some function $f \in \mathcal{F}$. Hence, at each invocation, we decode the codeword stored in memory and run the corresponding cryptographic primitive using the obtained key. Since the decoded key is either equal to the original or unrelated to it, we obtain $\mathcal{F}$-RKA security.

Unfortunately, there are two important caveats to the above general solution: (i) Since non-malleable codes are only secure against a *single* tampering attempt $f \in \mathcal{F}$, at each invocation we must completely erase the memory and re-encode the key; (ii) In case the modified codeword is invalid, and thus cannot be decoded, we must self-destruct and stop using the underlying primitive. It turns out that limitation (ii) is inherent, in that Gennaro, Lysyanskaya, Malkin, Micali and Rabin [GLM+04] established that RKA security is impossible without self-destruct.[1] On the other hand, it would be desirable to remove limitation (i) as perfect erasures of the memory are notoriously hard to implement in practice [CEGL08]. Another drawback of limitation (i) is that it makes the cryptoscheme stateful (even if it was stateless to start with) and requires fresh randomness for re-encoding the key.

The stronger notion of *continuously* non-malleable codes [FMNV14] allows us to overcome limitation (i): Since such codes guarantee $\mathcal{F}$-non-malleability even against poly-many tampering attempts, one immediately obtains $\mathcal{F}$-RKA security without assuming perfect erasures.

1.1 Our Contribution

A nice feature of the above compiler is its generality: In order to achieve $\mathcal{F}$-RKA security all we need to do is to design an $\mathcal{F}$-non-malleable code. In recent years, there has been a tremendous progress in the design of non-malleable codes for several tampering families $\mathcal{F}$ of practical interest, including: bit-wise independent and split-state tampering [DPW10, LL12, ADL14, AO20, CZ14, Li17, Li18, KOS17, KOS18, AKO+21], space-bounded tampering [FHMV17], small-locality and small-depth circuits [AGM+15, BDKM16, BDG+18, GMW19], decision-tree and AC0 tampering [BDKM18, BGW19], and very recently even bounded polynomial-depth tampering [BDK+19, DKP21, BDL22]. In contrast, continuous non-malleability is only known for bit-wise independent tampering [CMTV15, CDTV16], tampering functions with few fixed points and high entropy [JW15], constant-state tampering [ADN+19], split-state tampering [FMNV14, AKO17, OPVV18, DK19] and space-bounded tampering [CCHM19], leaving open the following intriguing question:

> *Can we construct* continuously *non-malleable codes against natural non-compartmentalized tampering families, such as decision trees, AC0 or even bounded polynomial-depth circuits?*

We answer the above question in the affirmative:

- In the setting of decision-tree tampering, we construct a code which resists continuous tampering attacks from the family of functions that modify every bit of the tampered codeword arbitrarily after adaptively reading up to d locations from the input codeword.

[1] Their attack is simple: The j-th tampering function tries to set the j-th bit of the secret key to 0: If the device returns an invalid output, the next function f_{j+1} additionally sets the j-th bit of the key to 1 and otherwise it sets it to 0.

Our scheme is in the plain model, assumes the existence of one-way functions, and tolerates tampering by decision trees of depth $d = O(n^{1/8})$, where n is the length of the codeword. Notably, this class includes NC0.

- In the setting of bounded polynomial-depth tampering, we construct a code that resists continuous tampering attacks, where the adversary can select any tampering function that can be computed by a circuit of bounded polynomial depth (and unbounded polynomial size). Notably, this class includes non-uniform NC. Our scheme is in the common reference string (CRS) model, and assumes the existence of time-lock puzzles and simulation-extractable (succinct) non-interactive zero-knowledge (NIZK) proofs.

We remark that both our constructions rely on computational assumptions. Although we don't know whether they are necessary for decision-tree or bounded-depth continuous tampering, achieving information-theoretic guarantees in the continuous scenario turned out to be challenging for even more well-studied families [CDTV16, CFV19, CMTV15, JW15, ADN+19, DKO+18, CCHM19, FHMV17, CGL16, AKO17]. We leave this problem open for future work.

1.2 Technical Overview

Let us start by reviewing different flavors of *one-time* non-malleability (see Section 3 for formal definitions).

- *Non-malleability w.r.t. message/codeword:* A code is non-malleable *w.r.t. message* (resp. *w.r.t. codeword*) if a tampered codeword either decodes to the original message (resp. is identical to the original codeword) or decodes to a completely unrelated value.

- *Super non-malleability:* A code is *super non-malleable* [FMNV14, FMVW14] if the tampered codeword itself (when valid) is unrelated to the original message. Note that the distinction between w.r.t. message and w.r.t. codeword also applies here.

Persistent tampering. The above flavors can be naturally extended to the setting of continuous non-malleability. Our first observation is that, in the setting of non-compartmentalized tampering, continuous non-malleability is only achievable in the case of *persistent* tampering, where the j-th tampering function f_j is applied to the output of the previous function f_{j-1}.

The latter can be seen as follows. Consider an adversary that computes offline a valid encoding of two different messages, for simplicity say $\mu_0 = 0^k$ and $\mu_1 = 1^k$. Call γ_0 and γ_1 the corresponding codewords. Next, the attacker prepares a tampering query that hard-wires γ_0, γ_1 and proceeds as follows: It reads the first bit $\gamma[1]$ of the target codeword; if $\gamma[1] = 0$ it overwrites the target codeword with γ_0, while if $\gamma[1] = 1$ it overwrites the target codeword with γ_1. As a result, the adversary learns the first bit of the target codeword. Now, if tampering is non-persistent, the attacker can repeat this procedure to efficiently recover the entire codeword, which clearly violates continuous non-malleability.[2]

In light of the above attack, in what follows, and without loss of generality, when we refer to continuous non-malleability, we implicitly refer to the case of persistent tampering.

Decision-tree tampering. To show our first result, we revisit the recent construction of non-malleable codes against decision-tree tampering by Ball, Guo and Wichs [BGW19]. On a high-level, this construction first encodes the message μ using a *leakage-resilient* non-malleable

[2] To the best of our knowledge, this observation is new. Previous work in the setting of non-compartmentalized tampering implicitly circumvented the above attack by requiring each tampering function to have high min-entropy and few fixed points, or by assuming that the number of tampering queries is a-priori bounded [JW15].

code in the split-state model, resulting in a codeword $(\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}})$ consisting of a right and a left part. Then, each part γ_i for $i \in \{\mathsf{L}, \mathsf{R}\}$ is encoded independently as follows: we sample a random small set (whose size is that of the underlying codeword) in a much larger array, plant the input in these locations and zero everything else out. Finally, we use a ramp secret sharing with relatively large secrecy threshold to encode a description of the small set (which can be represented by a seed ζ_i). To decode, we can simply extract the seed and output what is in the corresponding locations of the array. This allows us to recover both parts $\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}}$ and thus obtain the initial message.

Ball, Guo and Wichs [BGW19] show how to simulate the decoded message corresponding to one decision-tree tampering query using bounded split-state leakage and one split-state tampering query on the underlying non-malleable code. Although our construction is similar to theirs, proving continuous non-malleability is non-trivial and requires significant new ideas. We discuss some of them below.

First, in the original construction, the positions of the codeword that are not indexed by ζ_i are ignored, since they are not useful for the reconstruction. In our case, however, an attacker could copy the original codeword into the zero bits and overwrite the rest with a valid encoding of an unrelated message, which would allow it to retrieve the original encoding, thus breaking continuous non-malleability. We avoid this by requiring such positions to be 0 for the codeword to be valid. Second, we must ensure that the adversary cannot modify the other parts of the outer codeword without touching the inner codeword: this is because otherwise the adversary could use some tampering queries to save a state inside the codeword, and then use another tampering query to actually tamper with the codeword using more information than he should. We avoid this attack, by relying on computational assumptions. The idea is to sample verification keys $vk_{\mathsf{L}}, vk_{\mathsf{R}}$ for a one-time signature scheme, generate $(\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}})$ as an encoding of the string $\mu||vk_{\mathsf{L}}||vk_{\mathsf{R}}$, and finally append signatures $\sigma_{\mathsf{L}}, \sigma_{\mathsf{R}}$ to the left and right part of the above described final encoding. In Section 3.3, we also show that this trick works generically to compile any super non-malleable code w.r.t. message into a super non-malleable code w.r.t. codeword, so long as the tampering family $\mathcal{F}$ allows us to evaluate the signing algorithm of the signature scheme. Intuitively, our code against decision-tree tampering removes this assumption thanks to the fact that the split-state model allows us to run arbitrary polynomial-time functions (independently on the two parts of the codeword).

In a nutshell, our scheme uses as building blocks a split-state nmc, a signature scheme and a simple procedure transforming states into their sparse versions. The latter takes as input a length-c-string γ, samples a random set $\mathcal{I}$ of c indices in $[n]$ with $n > c$, and outputs the sparse codeword $\gamma^* = (\gamma_1^*, \gamma_2^*)$, where γ_1^* is a RSS encoding of $\mathcal{I}$, and γ_2^* is a length-n-string that has γ in the positions indexed by $\mathcal{I}$, and zeros elsewhere. To extract the original string from the sparse one, it suffices to use the RSS decoding algorithm on the first part, and return the corresponding bits of the second part.
The design of our scheme follows.

Algorithm $\mathsf{Enc}^*(\mu)$**.** Proceed as follows:

1. Sample two pairs of keys $(sk_{\mathsf{L}}, vk_{\mathsf{L}}), (sk_{\mathsf{R}}, vk_{\mathsf{R}})$ for the signature scheme
2. Compute the split-state codeword $(\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}})$ for the message $(\mu||vk_{\mathsf{L}}||vk_{\mathsf{R}})$
3. Compute the sparse strings γ_{L}^* and γ_{R}^* for γ_{L} and γ_{R}.
4. Sign γ_{L}^* and γ_{R}^* with, respectively, sk_{L} and sk_{R}, to get σ_{L} and σ_{R}.
5. The final codeword is $(\sigma_{\mathsf{L}}, \gamma_{\mathsf{L}}^*, \sigma_{\mathsf{R}}, \gamma_{\mathsf{R}}^*)$.

The decoding algorithm extracts γ_{L} and γ_{R} from their sparse versions γ^*_{L} and γ^*_{R} and checks that in the remaining positions there are only zeros, decodes the split-state codeword $(\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}})$ to get $\mu||vk_{\mathsf{L}}||vk_{\mathsf{R}}$, verifies the signatures and outputs $\perp$ if verification fails, μ otherwise.

Unfortunately, even with the above modifications, it is unclear how to extend the original proof of security to the setting of continuous tampering, even if one assumes the underlying split-state non-malleable code to be continuously non-malleable. The reason is that the reduction needs to leak some bits from the codeword for each tampering query, therefore having a large number of tampering queries would lead to leaking too much information from the split-state codeword. Instead, we exploit the power of *super* non-malleability: Assume the underlying split-state code is super non-malleable w.r.t. codeword.[3] Then, the reduction only needs to know the index q^* of the first tampering query which actually modifies the inner codeword. In case the tampered inner codeword $(\tilde{\gamma}_{\mathsf{L}}, \tilde{\gamma}_{\mathsf{R}})$ is invalid, the experiment stops and we are done. Otherwise, if $(\tilde{\gamma}_{\mathsf{L}}, \tilde{\gamma}_{\mathsf{R}})$ is valid, the reduction obtains it in full. At this point, the reduction is able to simulate the answer to all subsequent tampering queries on its own, as tampering is persistent, which allows us to conclude continuous non-malleability.[4]

It remains to be seen how the reduction can obtain the index q^*. A possible strategy would be to simulate the outcome of all the tampering queries inside the leakage oracle, and then return the index of the first tampering query which actually modifies the codeword; however, each bit of a tampering query can depend on bits of both the left and right part of the inner codeword, while a split-state leakage query is only allowed to see one of these parts. Our strategy is to guess the index q^*, and then check at the end of the experiment if the guess was correct or wrong. Here, we additionally exploit the fact that the underlying split-state super non-malleable code is information-theoretically secure, which essentially allows the reduction to run many instances of the experiment inside the leakage oracle, and check that the adversary does not try to cancel its advantage (due to a wrong simulation). A similar strategy was already used in [OPVV18, BFO+20, BFV21]. The formal proof appears in Section 4.1.

Bounded polynomial-depth tampering. Our second construction exploits the observation that, for certain tampering families, continuous non-malleability w.r.t. codeword can be reduced to one-time super non-malleability w.r.t. codeword plus logarithmic (in the security parameter) leakage on the codeword. Indeed, this is the case as long as the leakage family allows us to run polynomially-many tampering functions in parallel, and return the index of the first query that actually modifies the codeword (if any). We formalize this observation in Section 3.3 (see Theorem 3). Note that the latter clearly holds true in the setting of bounded polynomial-depth leakage and tampering.[5]

In light of the above, it suffices to construct a one-time super non-malleable code w.r.t. codeword against bounded polynomial-depth tampering. We do so, by looking at a slightly more general question. Namely, in Section 4.2, we show how to compile a *leakage-resilient* non-malleable code into a super non-malleable code in the CRS model, using simulation-extractable NIZK proofs. The idea is to encode a message μ using the underlying code, and then append to the resulting encoding γ a NIZK proof of knowledge π of the randomness ρ used by the encoder. The decoder outputs $\perp$ if the NIZK proof does not verify correctly.

In the reduction, we can simulate the NIZK proof π and then use a leakage query in order to obtain the tampered proof $\tilde{\pi}$ (so long as the proof $\tilde{\pi}$ is valid), along with the extracted witness $\tilde{\rho}$ corresponding to a tampered codeword $(\tilde{\gamma}, \tilde{\pi}) = f(\gamma, \pi)$ in the experiment defining super

[3] We can take, e.g., the non-malleable code of [AKO17] for a concrete instantiation.

[4] As a bonus, we actually prove continuous *super* non-malleability.

[5] The same observation holds true for the setting of AC0 tampering, but not for decision-tree tampering.

non-malleability. Unfortunately, the randomness $\tilde{\rho}$ is too long[6] for being obtained via a leakage query. However, this issue can be resolved by generating ρ using a pseudorandom generator G and letting the corresponding λ-bit seed σ be the witness. This allows the overall leakage to depend only on the security parameter, either assuming simulation-extractable SNARKs [BPR20] (which inherently require non-falsifiable assumptions [GW11]), or by making the size of the proof depend only on the size of the witness (which can be achieved using fully-homomorphic encryption [GGI+15]).

More in detail, our compiler builds on a leakage-resilient one-time non-malleable code $(\mathsf{Enc}, \mathsf{Dec})$, a pseudorandom generator G, and a simulation-extractable proof system. The relation $\mathcal{R}$ for the proof system is satisfied by every couple statement-witness (γ, σ) where $\gamma = \mathsf{Enc}(\mu; \mathsf{G}(\sigma))$ for some message μ. Our encoding (and decoding) algorithm takes as input a CRS ω for the underlying proof system, and is described below.

Algorithm $\mathsf{Enc}^*(\omega, \mu)$**:** Proceed as follows:

1. Generate a random seed σ for the PRG.
2. Use the underlying non-malleable encoding algorithm Enc with randomness $\mathsf{G}(\sigma)$) to compute a codeword γ for μ
3. Generate a proof π for the couple (γ, σ)
4. Output (γ, π).

The decoding algorithm verifies the proof, returns $\perp$ if verification fails, and the message μ underlying γ otherwise.

A subtlety in the above proof sketch is that the leakage family supported by the underlying code must allow simulating the proof π, applying the tampering function f on (γ, π), verifying the tampered proof $\tilde{\pi}$, and extracting the corresponding tampered seed $\tilde{\sigma}$. Similarly, the tampering family supported by the underlying code must allow simulating the proof π and applying the tampering function f on (γ, π). Hence, this compiler does not work for all tampering families. Fortunately, it clearly works for the setting of bounded polynomial-depth tampering.

Our final result is then achieved by adapting a recent construction of Dachman-Soled, Komargodski and Pass [DKP21], who showed how to obtain one-time non-malleability w.r.t. message against bounded polynomial-depth tampering assuming the existence of key-less hash functions and time-lock puzzles (along with other standard assumptions); in the CRS model, we show that their construction can be simplified and proven leakage-resilient one-time non-malleable assuming the existence of time-lock puzzles and simulation-extractable NIZKs. We refer the reader to Section 4.2 and Appendix B for more details.

Necessity of super non-malleability for the inner split-state code. In our construction against decision-tree tampering, we require the inner split-state encoding to be a super non-malleable code, thus allowing for the simulation of the whole codeword. We argue that this is not an artifact of our proof, but rather a necessity for our construction to achieve security. Indeed, by using a contrived instance of a non-malleable code which is not super non-malleable, and contrived instances of the ramp secret sharing and the signature scheme, we are able to instantiate our scheme so that the adversary becomes able to retrieve the message in full. We consider here a simplified version of our scheme in which we remove the signature scheme, and we point the reader to Appendix D for the detailed explanation and for how to reintroduce back the signatures.

[6]Note that we cannot extract the proof outside the leakage function, as the corresponding statement is the tampered modified codeword $\tilde{\gamma}$ inside the leakage oracle.

First of all, we need a split-state non-malleable code which has a good amount of spare bits, initially set to 0, and a secondary mode of operation which uses the spare bits to reconstruct the message instead of the actual relevant bits. Then, we need a malleable RSS encoding which allows to only replace a part of the encoded value leaving everything else intact.

The attack then proceeds as follows: the adversary is now able to tamper with the RSS encoding so that the spare bits of the split-state codeword are in a known location (while keeping the other positions untouched), and he is also able to replace those spare bits with some encoding of either 0 or 1 depending on some bit that the adversary wants to leak, leaving everything else untouched. Finally, the adversary uses multiple queries to leak every bit he left untouched, thus recovering all the bits that are necessary to reconstruct the original message.

Application to RKA security without erasures. It is well known that a continuously $\mathcal{F}$-non-malleable code allows us to obtain a natural notions of $\mathcal{F}$-RKA security for arbitrary cryptographic primitives. This was proven by Faust, Mukherjee, Nielsen, and Venturi [FMNV14] for the case of non-persistent tampering. In Appendix C, we show that the same works for the case of persistent tampering.

1.3 Related Work

In recent work, Freitag *et al.* [EFKP20] investigate non-malleable time-lock puzzles in the concurrent setting. Their definition generalizes continuous non-persistent non-malleable codes against bounded depth tampering, but requires that the adaptive choice of tampering functions runs in bounded depth too. They provide an impossibility result for the latter, which we extend to all the continuous non-persistent non-malleable codes against global tampering. Given that, they introduce the weaker notion of functional concurrent non-malleable time-lock puzzles, present a construction assuming the existence of (plain) time-lock puzzles in the auxiliary-input random oracle model, and provide interesting applications in coin tossing and electronic auctions.

Dachman-Soled and Kulkarni [DK19] show that *continuous* super non-malleability in the split-state model inherently requires setup. This impossibility, instead, does not hold for continuous super non-malleability against *persistent* tampering attacks, which can be achieved information-theoretically in the split-state model.

Leakage-Resilient Locally Decodable and Updatable Non-Malleable Codes [DLSZ15] are a fine-grained tool for protecting RAM machines against leakage and tampering. In literature, there are constructions in the split-state and continuous setting [DLSZ15], with information theoretic security [CKR16], as well as tight upper and lower bounds [DKS17].

An alternative approach for obtaining generic RKA-security is to rely on non-malleable key derivation [FMVW14, QLY+15, CQZ+16]. The difference with non-malleable codes is that in this case one stores a uniformly random string in memory which is used to derive a key for the underlying cryptoscheme at each invocation. Continuously non-malleable key derivation can essentially be achieved only for tampering via polynomials or functions with high entropy.

Another line of research seeks direct constructions of RKA-secure cryptographic primitives, including, e.g., pseudorandom functions [BK03, BC10, ABPP14] public-key encryption [AHI11, Wee12], identity-based encryption and signatures [BPT12]. RKA security has become a de-facto standard for block-ciphers, and systems are often designed while implicitly relying on the RKA-security of the underlying block-cipher (see, e.g., [BCM11] and references therein).

2 Preliminaries

We start by setting up some basic notation and by recalling the definitions of the necessary standard cryptographic primitives.

2.1 Notation

We denote by $[n]$ the set $\{1, \ldots, n\}$. For a string $x \in \{0,1\}^*$, we denote its length by $|x|$; if $i \in [|x|]$ and $\mathcal{I} \subseteq [|x|]$, we denote by $x[i]$ the i-th bit of x and by $x[\mathcal{I}]$ the substring of x obtained by only considering the bits indexed by $\mathcal{I}$.

If $\mathcal{X}$ is a set, $|\mathcal{X}|$ represents the number of elements in $\mathcal{X}$. When x is chosen randomly in $\mathcal{X}$, we write $x \leftarrow\!\!\$\ \mathcal{X}$. When A is a randomized algorithm, we write $y \leftarrow\!\!\$\ \mathsf{A}(x)$ to denote a run of A on input x (and implicit random coins ρ) and output y; the value y is a random variable and $\mathsf{A}(x;\rho)$ denotes a run of A on input x and randomness ρ. An algorithm A is *probabilistic polynomial-time* (PPT for short) if A is randomized and for any input $x, \rho \in \{0,1\}^*$, the computation of $\mathsf{A}(x;\rho)$ terminates in a polynomial number of steps (in the size of the input).

Asymptotics. We denote by $\lambda \in \mathbb{N}$ the security parameter. A function p is *polynomial* (in the security parameter), if $p(\lambda) = O(\lambda^c)$ for some constant $c > 0$. A function $\nu : \mathbb{N} \to [0,1]$ is *negligible* (in the security parameter) if it vanishes faster than the inverse of any polynomial in λ, i.e. $\nu(\lambda) = O(1/p(\lambda))$ for all positive polynomials $p(\lambda)$. Unless stated otherwise, we implicitly assume that the security parameter is given as input (in unary) to all algorithms.

Random variables. For a random variable $\mathbf{X}$, we write $\mathbb{P}[\mathbf{X} = x]$ for the probability that $\mathbf{X}$ takes on a particular value $x \in \mathcal{X}$, with $\mathcal{X}$ being the set over which $\mathbf{X}$ is defined. The statistical distance between two random variables $\mathbf{X}$ and $\mathbf{Y}$ over $\mathcal{X}$ is defined as $\Delta(\mathbf{X};\mathbf{Y}) := \frac{1}{2}\sum_{x\in\mathcal{X}} |\mathbb{P}[\mathbf{X} = x] - \mathbb{P}[\mathbf{Y} = x]|$. Given two ensembles $\mathbf{X} = \{\mathbf{X}_\lambda\}_{\lambda\in\mathbb{N}}$ and $\mathbf{Y} = \{\mathbf{Y}_\lambda\}_{\lambda\in\mathbb{N}}$, we write $\mathbf{X} \equiv \mathbf{Y}$ to denote that $\mathbf{X}_\lambda$ and $\mathbf{Y}_\lambda$ are identically distributed, $\mathbf{X} \overset{\text{s}}{\approx} \mathbf{Y}$ to denote that they are *statistically close*, i.e. $\Delta(\mathbf{X}_\lambda;\mathbf{Y}_\lambda) \leq \nu(\lambda)$ for some negligible function $\nu : \mathbb{N} \to [0,1]$, and $\mathbf{X} \overset{\text{c}}{\approx} \mathbf{Y}$ to denote that they are *computationally indistinguishable*, i.e. for all PPT distinguishers D there is a negligible function $\nu : \mathbb{N} \to [0,1]$ such that:

$$\Delta_{\mathsf{D}}(\mathbf{X}_\lambda;\mathbf{Y}_\lambda) := |\mathbb{P}[\mathsf{D}(X_\lambda) = 1] - \mathbb{P}[\mathsf{D}(Y_\lambda) = 1]| \leq \nu(\lambda).$$

The notion of computational/statistical indistinguishability generalizes immediately to ensembles of interactive experiments $\{\mathbf{G}_{\mathsf{A}}(\lambda)\}_{\lambda\in\mathbb{N}}$ where the adversary A outputs a bit at the end of the interaction.

2.2 Signature Schemes

A *signature scheme* is a triple of polynomial-time algorithms $\Sigma = (\mathsf{Gen}, \mathsf{Sign}, \mathsf{SigVer})$ specified as follows.

Key Generation: The probabilistic key generation algorithm Gen takes as input the security parameter $\lambda \in \mathbb{N}$ (in unary) and returns a pair of keys $(vk, sk) \in \mathcal{V} \times \mathcal{K}$. We call vk the *verification key* and sk the *signing key*.

Signature Generation: The probabilistic signing algorithm Sign takes as input a message $\mu \in \mathcal{M}$ and a signing key sk, and outputs a signature $\sigma \leftarrow\!\!\$\ \mathsf{Sign}(sk, m)$ with $\sigma \in \mathcal{S}$.

Signature Verification: The deterministic verification algorithm SigVer takes as input a pair (μ, σ) and a verification key vk, and outputs a decision bit. A signature σ is called *valid* with respect to μ and vk iff $\mathsf{SigVer}(vk, \mu, \sigma) = 1$.

Correctness requires that for all $\lambda \in \mathbb{N}$, all $(vk, sk) \in \mathsf{Gen}(1^\lambda)$, and all $\mu \in \mathcal{M}$, it holds that $\mathsf{SigVer}(vk, \mu, \mathsf{Sign}(sk, \mu)) = 1$ with probability one over the randomness of the signing algorithm. As for security, we require that no efficient attacker can forge a valid signature without knowing the secret key.

Definition 1 (Strong one-time signature)**.** *A signature scheme Σ is* one-time strongly unforgeable *if for all PPT adversaries $\mathsf{A} = (\mathsf{A}_0, \mathsf{A}_1)$ there is a negligible function $\nu : \mathbb{N} \to [0, 1]$ such that:*

$$\mathbb{P}\left[\mathsf{SigVer}(vk, \mu^*, \sigma^*) = 1 \wedge (\mu^*, \sigma^*) \neq (\mu, \sigma) : \begin{array}{l}(vk, sk) \leftarrow\!\!\$\ \mathsf{Gen}(1^\lambda) \\ (\mu, \alpha) \leftarrow\!\!\$\ \mathsf{A}_0(vk) \\ \sigma \leftarrow\!\!\$\ \mathsf{Sign}(sk, \mu) \\ (\mu^*, \sigma^*) \leftarrow\!\!\$\ \mathsf{A}_1(\alpha, \mu, \sigma)\end{array}\right] \leq \nu(\lambda).$$

consider the following game $\mathbf{G}^{\mathsf{ots}}_{\Sigma,\mathsf{A}}(\lambda)$ depending on security parameter $\lambda \in \mathbb{N}$ and attacker A.

Key generation: The game runs $(vk, sk) \leftarrow\!\!\$\ \mathsf{Gen}(1^\lambda)$, and outputs vk.

Signature queries: Upon input a message m, the game computes and returns $\sigma \leftarrow\!\!\$\ \mathsf{Sign}(sk, \mu)$.

Forgery: Upon input (μ^*, σ^*), the game outputs 1 iff $\mathsf{SigVer}(vk, \mu^*, \sigma^*) = 1$ and $(\mu^*, \sigma^*) \neq (\mu, \sigma)$ for all pairs (μ, σ) corresponding to signature queries.

2.3 Non-Interactive Zero Knowledge

Let $\mathcal{L} \subseteq \{0,1\}^*$ be an NP language, with corresponding relation $\mathcal{R} \subseteq \{0,1\}^* \times \{0,1\}^*$ such that $\mathcal{L} = \{y : \exists x \text{ s.t. } (y, x) \in \mathcal{R}\}$. A *non-interactive argument system* for $\mathcal{R}$ is a tuple of polynomial-time algorithms $\Pi = (\mathsf{CRSGen}, \mathsf{Prove}, \mathsf{ProofVer})$ specified as follows.

CRS Generation: The CRS generation algorithm CRSGen is a probabilistic algorithm that takes as input the security parameter $\lambda \in \mathbb{N}$ (in unary) and outputs a common reference string $\omega \in \{0,1\}^*$.

Proof Generation: The prover algorithm Prove is a probabilistic algorithm that takes as input the CRS ω and a pair $(y, x) \in \mathcal{R}$ and returns an argument $\pi \in \{0,1\}^*$.

Verification: The verification algorithm $\mathsf{ProofVer}$ is a deterministic algorithm that takes as input the CRS ω and a pair (y, π) and returns a bit.

We say that Π is *correct* if for every $\lambda \in \mathbb{N}$, all $\omega \in \mathsf{CRSGen}(1^\lambda)$, and all pairs $(y, x) \in \mathcal{R}$, it holds that

$$\mathbb{P}\left[\mathsf{ProofVer}(\omega, y, \mathsf{Prove}(\omega, y, x)) = 1\right] = 1$$

where the probability is over the randomness of the prover algorithm.

A non-interactive argument system typically satisfies two properties, which we recall below. The first property informally states that honestly computed arguments for true instances y reveal nothing beyond the fact that $y \in \mathcal{L}$.

Definition 2 (Single-theorem zero knowledge)**.** *A non-interactive argument system Π for a relation $\mathcal{R}$ satisfies* single-theorem zero knowledge *if there exists a PPT simulator $(\mathsf{S}_0, \mathsf{S}_1)$ such that the following conditions are met:*

$\mathbf{Real}_{\Pi,\mathsf{A}}(\lambda)$	$\mathbf{Ideal}_{\Pi,\mathsf{A},\mathsf{S}}(\lambda)$
1: $\omega \leftarrow\!\!\$\ \mathsf{CRSGen}(1^\lambda)$	1: $(\omega,\tau,\xi) \leftarrow\!\!\$\ \mathsf{S}_0(1^\lambda)$
2: $(y,x) \leftarrow\!\!\$\ \mathsf{A}(\omega)$	2: $(y,x) \leftarrow\!\!\$\ \mathsf{A}(\omega)$
3: **if** $(y,x) \in \mathcal{R}$ **then**	3: **if** $(y,x) \in \mathcal{R}$ **then**
4: $\quad \pi \leftarrow\!\!\$\ \mathsf{Prove}(\omega,y,x)$	4: $\quad \pi \leftarrow\!\!\$\ \mathsf{S}_1(\tau,y)$
5: **else**	5: **else**
6: $\quad \pi \leftarrow \bot$	6: $\quad \pi \leftarrow \bot$
7: **return** ω,π	7: **return** ω,π

Figure 1: Experiments defining non-interactive zero knowledge.

- *Simulator* S_0 *outputs a CRS* $\omega \in \{0,1\}^*$*, a simulation trapdoor* $\tau \in \{0,1\}^*$ *and an extraction trapdoor* $\xi \in \{0,1\}^*$*.*
- *For all PPT adversaries* A*, it holds* $\{\mathbf{Real}_{\Pi,\mathsf{A}}(\lambda)\}_{\lambda\in\mathbb{N}} \overset{c}{\approx} \{\mathbf{Ideal}_{\Pi,\mathsf{A},\mathsf{S}}(\lambda)\}_{\lambda\in\mathbb{N}}$*, where the experiments* $\mathbf{Real}_{\Pi,\mathsf{A}}(\lambda)$ *and* $\mathbf{Ideal}_{\Pi,\mathsf{A},\mathsf{S}}(\lambda)$ *are defined in Fig. 1.*

The second property implies that no efficient adversary can prove a false statement. In fact, we require a stronger property saying that accepting proofs can be "extracted" in polynomial time, given a trapdoor, even when the attacker can see a single simulated proof of a (possibly false) statement.

Definition 3 (Simulation extractability)**.** *Let* Π *be a non-interactive argument system for a relation* $\mathcal{R}$*, satisfying single-theorem zero-knowledge with simulator* $(\mathsf{S}_0,\mathsf{S}_1)$*. We say that* Π *is* one-time simulation extractable *if there exists a PPT algorithm* K *such that for all PPT adversaries* $\mathsf{A} = (\mathsf{A}_0,\mathsf{A}_1)$ *the following quantity is negligible:*

$$\mathbb{P}\left[\begin{array}{l|l} \mathsf{ProofVer}(\omega,\tilde{y},\tilde{\pi})) = 1 & (\omega,\tau,\xi) \leftarrow\!\!\$\ \mathsf{S}_0(1^\lambda); (y,\alpha) \leftarrow\!\!\$\ \mathsf{A}_0(\omega) \\ \wedge (\tilde{y},\tilde{\pi}) \neq (y,\pi) \wedge (\tilde{y},\tilde{x}) \notin \mathcal{R} & : \ \pi \leftarrow\!\!\$\ \mathsf{S}_1(\tau,y); (\tilde{y},\tilde{\pi}) \leftarrow\!\!\$\ \mathsf{A}_1(\alpha,\pi) \\ & \tilde{x} \leftarrow\!\!\$\ \mathsf{K}(\xi,\tilde{y},\tilde{\pi}) \end{array}\right].$$

2.4 Pseudorandom Generators

A pseudorandom generator (PRG) is a deterministic polynomial-time algorithm G taking as input a seed of true randomness and expanding it into a much longer sequence of pseudorandom bits.

Definition 4 (Pseudorandom generator)**.** *We call* $\mathsf{G} : \{0,1\}^{s(\lambda)} \to \{0,1\}^{r(\lambda)}$ *a secure PRG if for all* $\lambda \in \mathbb{N}$ *it holds that* $r(\lambda) > s(\lambda)$*, and additionally*

$$\{\mathsf{G}\left(\mathbf{U}_{s(\lambda)}\right)\}_{\lambda\in\mathbb{N}} \overset{c}{\approx} \{\mathbf{U}_{r(\lambda)}\}_{\lambda\in\mathbb{N}},$$

where $\mathbf{U}_n$ *is the uniform distribution over* $\{0,1\}^n$*.*

2.5 Time-lock Puzzles

A time-lock puzzle Π is a pair of algorithms $(\mathsf{PGen},\mathsf{PSol})$ specified as follows.

Puzzle Generation: The puzzle generation algorithm is a probabilistic algorithm that takes as input a security parameter λ, a message $\mu \in \mathcal{M}$, a difficulty parameter T and returns a puzzle $\zeta \in \mathcal{Z}$.

$\mathbf{TLP}_{\Pi,\mathsf{A}}(\lambda, b)$

1: $\zeta \leftarrow\!\!\$\ \mathsf{PGen}(1^\lambda, \mu_b, T)$
2: **return** $\mathsf{A}(\zeta)$

Figure 2: Experiment defining time-lock puzzles, with an adversary A

Puzzle Solution: The puzzle solution algorithm is a deterministic algorithm that takes as input a puzzle $\zeta \in \mathcal{Z}$ and outputs either a message $\mu \in \mathcal{M}$ or $\perp$.

We will consider time-lock puzzles satisfying the properties below.

- *Correctness:* For every security parameter λ, $\mu \in \mathcal{M}$ and difficulty parameter T, it holds

$$\mathbb{P}\left[\mathsf{PSol}(\mathsf{PGen}(1^\lambda, \mu, T)) = \mu\right] = 1.$$

- *Efficiency:* The puzzle generation algorithm runs in time $poly(\log T, \lambda)$, while the puzzle solution algorithm can be computed in time $T \cdot poly(\lambda)$.

- *(S, ϵ)-hardness:* There exists a polynomial $\overline{T}$ such that, for every $T > \overline{T}$, every T^ϵ-time S-size distinguisher A, every security parameter λ and couple of messages $\mu_0, \mu_1 \in \mathcal{M}$, the following holds for the experiments defined in Fig. 2:

$$\mathbf{TLP}_{\Pi,\mathsf{A}}(\lambda, 0) \overset{c}{\approx} \mathbf{TLP}_{\Pi,\mathsf{A}}(\lambda, 1)$$

In their work, Dachman-Soled *et al.* [DKP21] present an explicit construction for time-lock puzzles against exponential-size adversaries. Their scheme is based on the repeated squaring assumption.

Definition 5 (Repeated Squaring Assumption)**.** *For some $\epsilon, \epsilon' \in (0, 1)$ and any large enough t, the following holds: any $2^{\lambda^{\epsilon'}}$-size t^ϵ-time algorithm cannot distinguish $(g, N, t, g^{2^t} mod N)$ from (g, N, t, g'), where g, g' are uniform elements in $\mathbb{Z}^*_{p\cdot q}$, $N = p \cdot q$ and p, q are two λ-bit primes.*

Given a message μ, and parameters $p, q, N, g, \epsilon, \epsilon'$ as above, define the time-lock puzzle as

$$\zeta = (g, N, t, \mu + g^{2^t} mod N).$$

The repeated squaring assumption guarantees hardness for T-time S-size adversaries whenever $t = T^{1/\epsilon}$ and both p and q have $(\log S)^{1/\epsilon'}$ bits. Efficiency holds for $\log S \in poly(\log T, \lambda)$. Note that S can also be a function of the puzzle length z itself, as long as

$$z \geq 6(\log S)^{1/\epsilon'} + \log T^{1/\epsilon}. \tag{1}$$

2.6 Coding Schemes

A (k, n)-code is a pair of algorithms $\Gamma = (\mathsf{Init}, \mathsf{Enc}, \mathsf{Dec})$ specified as follows.

Initialization: The initialization algorithm Init is a randomized algorithm that takes as input the security parameter $\lambda \in \mathbb{N}$ (in unary) and outputs a CRS $\omega \in \{0, 1\}^*$.

Encoding: The encoding algorithm Enc is a randomized algorithm that takes as input a CRS $\omega \in \{0, 1\}^*$, a message $\mu \in \{0, 1\}^k$ and outputs a codeword $\gamma \in \{0, 1\}^n$.

Decoding: The decoding algorithm Dec is a deterministic algorithm that takes as input a CRS $\omega \in \{0,1\}^*$, a codeword $\gamma \in \{0,1\}^n$ and outputs either a value in $\{0,1\}^k$ or $\perp$ (denoting an invalid codeword).

We say that Γ satisfies correctness if for all $\omega \in \mathsf{Init}(1^\lambda)$ and all messages $\mu \in \{0,1\}^k$ it holds that $\mathbb{P}\left[\mathsf{Dec}(\omega, \mathsf{Enc}(\omega, \mu)) = \mu\right] = 1$, where the probability is over the randomness of the encoding algorithm.

Remark 1 (Coding schemes in the plain model)**.** *A code in the plain model can be obtained by restricting* Init *to output the empty string. In that case, we simply write* $\Gamma = (\mathsf{Enc}, \mathsf{Dec})$ *and omit the string* ω *as an input of the encoding and decoding algorithm.*

Ramp secret sharing. A ramp secret sharing is a coding scheme satisfying the additional property that any subset of the bits of a codeword with size at most $\lfloor t \cdot n \rfloor$, for some $t \in (0,1)$, reveals nothing about the message.

Definition 6 (Ramp secret sharing)**.** *We say that* Γ *is a binary* (k, n, t)*-ramp secret sharing if* Γ *is a* (k, n)*-code satisfying the following property: For every* $\mu \in \{0,1\}^k$*, and for every non-empty subset* $\mathcal{I} \subseteq \{0,1\}^n$ *of size at most* $\lfloor t \cdot n \rfloor$*, we have that* $\mathsf{Enc}(\mu)|_{\mathcal{I}}$ *is identically distributed to the uniform distribution over* $\{0,1\}^{|\mathcal{I}|}$*.*

As shown by Ball *et al.* [BDG+18], any binary linear error correcting code is a binary ramp secret sharing with suitable secrecy. In particular, every binary linear error correcting code correcting at most d errors is a binary ramp secret sharing with secrecy $(d-1)/n$.

Lemma 1 ([BDG+18])**.** *For any message length* $k \in \mathbb{N}$ *there exist parameters* $n \in \mathbb{N}$ *and* $t \in (0,1)$ *such that there is a binary* (k, n, t)*-ramp secret sharing.*

3 Non-Malleable Codes

In this section, we revisit the definition of non-malleable codes and establish relations among different flavors of non-malleability.

3.1 Non-Malleability

Let Γ be a (k, n)-code, and $\mathcal{F} = \{f : \{0,1\}^n \to \{0,1\}^n\}$ be a family of functions. Informally, Γ is non-malleable against tampering in $\mathcal{F}$ if decoding a codeword tampered via functions in $\mathcal{F}$ yields either the *original message* or a completely unrelated value. In this paper, we refer to the above flavor of security as *non-malleability w.r.t. message*. Instead, when a tampered codeword (always via functions in $\mathcal{F}$) is either identical to the *original codeword* or decodes to a completely unrelated value, we speak of *non-malleability w.r.t. codeword*.[7]

A stronger (as the name suggests) flavor of non-malleability is the so-called *super non-malleability*, introduced implicitly in [FMNV14] (and explicitly in [FMVW14]). This property requires that not only the output of the decoding, but *the codeword itself*, is independent of the message, as long as the tampered codeword is valid and either different from the original codeword (yielding super non-malleability w.r.t. codeword) or decoding to something different than the original message (yielding super non-malleability w.r.t. message).

[7] In the literature, the latter flavor of non-malleability is sometimes known as *strong non-malleability* whereas the former flavor is also known as *weak non-malleability*. However, we find this terminology rather confusing due to the fact that a code can be at the same time weakly non-malleable and super non-malleable (as defined below).

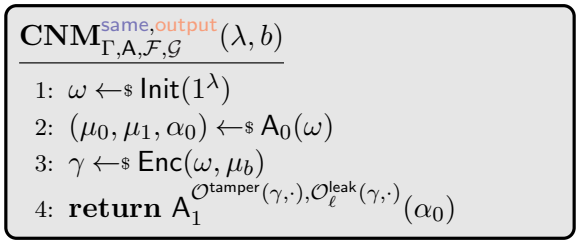

Figure 3: Experiment defining leakage-resilient (super) non-malleable codes, with an adversary A consisting of subroutines $(\mathsf{A}_0, \mathsf{A}_1)$.

The definition below formalizes continuous (super) non-malleability w.r.t. message/codeword. For readability, it will be useful to introduce the following predicates depending on a code Γ, a CRS ω, two messages μ_0, μ_1, two codewords $\gamma, \tilde{\gamma}$ and a tampering function $f \in \mathcal{F}$:

- $\mathsf{msg}(\omega, \mu_0, \mu_1, \gamma, \tilde{\gamma})$: outputs 1 if and only if $\mathsf{Dec}(\omega, \tilde{\gamma}) \in \{\mu_0, \mu_1\}$;
- $\mathsf{cdw}(\omega, \mu_0, \mu_1, \gamma, \tilde{\gamma})$: outputs 1 if and only if $\tilde{\gamma} = \gamma$;
- $\mathsf{standard}(\tilde{\mu}, \tilde{\gamma})$: outputs $\tilde{\mu}$;
- $\mathsf{super}(\tilde{\mu}, \tilde{\gamma})$: outputs $\tilde{\mu}$ if $\tilde{\mu} \in \{\diamond, \perp\}$, and $\tilde{\gamma}$ otherwise.

The above algorithms are called inside the tampering oracle $\mathcal{O}^{\mathsf{tamper}}(\gamma, \cdot)$, which initializes[8] $\hat{\gamma} = \gamma$ and self-destruct parameter $\delta = 0$, and behaves as follows:

1. if $\delta = 1$, output $\perp$;
2. compute $\tilde{\gamma} = f(\hat{\gamma})$ and $\tilde{\mu} = \mathsf{Dec}(\omega, \tilde{\gamma})$;
3. if $\mathsf{same}(\omega, \mu_0, \mu_1, \hat{\gamma}, \tilde{\gamma}) = 1$, set $\tilde{\mu} = \diamond$;
4. if $\tilde{\mu} = \perp$, set $\delta = 1$;
5. set $\hat{\gamma} = \tilde{\gamma}$ and return $\mathsf{output}(\tilde{\mu}, \tilde{\gamma})$;

We model leakage resilience by an oracle $\mathcal{O}^{\mathsf{leak}}_\ell(\gamma, \cdot)$ that accepts as input functions $g \in \mathcal{G}$ and returns $g(\gamma)$ (or $\perp$ if $\delta = 1$), for a total of at most ℓ bits.

Definition 7 (Continuously non-malleable codes)**.** *Let Γ be a (k, n)-code, and $\mathcal{F} \subseteq \{f : \{0,1\}^n \to \{0,1\}^n\}$ and $\mathcal{G} \subseteq \{g : \{0,1\}^n \to \{0,1\}^*\}$ be family of functions. For flags $\mathsf{same} \in \{\mathsf{msg}, \mathsf{cdw}\}$ and $\mathsf{output} \in \{\mathsf{standard}, \mathsf{super}\}$ we say that Γ is a $(\mathcal{G}, \ell)$-*leakage-resilient persistent continuously $\mathcal{F}$-non-malleable *code if the following holds for the experiments defined in Fig. 3:*

$$\left\{\mathbf{CNM}^{\mathsf{same},\mathsf{output}}_{\Gamma,\mathsf{A},\mathcal{F},\mathcal{G}}(\lambda, 0)\right\}_{\lambda\in\mathbb{N}} \overset{\mathrm{c}}{\approx} \left\{\mathbf{CNM}^{\mathsf{same},\mathsf{output}}_{\Gamma,\mathsf{A},\mathcal{F},\mathcal{G}}(\lambda, 1)\right\}_{\lambda\in\mathbb{N}}. \tag{2}$$

In particular:

- *When Eq. (2) holds for* $\mathsf{same} = \mathsf{msg}$ *(resp.* $\mathsf{same} = \mathsf{cdw}$*) we speak of persistent continuous non-malleability* w.r.t. message *(resp.* w.r.t. codeword*);*

[8] The oracle additionally takes as input all the values that are required to evaluate the above predicates. We omit them for clarity.

- *When Eq. (2) holds for* output = super*, we refer to persistent continuous super non-malleability w.r.t. message/codeword. When* output = standard*, we speak of persistent continuous non-malleability w.r.t. message/codeword.*
- *When Eq. (2) holds in the information-theoretic setting with statistical distance at most* $\epsilon \in [0,1]$*, we say that* Γ *is leakage-resilient persistent continuously super non-malleable with* statistical security ϵ*.*

One-time non-malleability. When we restrict the adversary by only allowing one tampering query, we obtain the weaker notion of *one-time non-malleability*. To formalize the latter, it suffices to replace Item 4 with an instruction which sets $\delta = 1$ regardless of the value of $\tilde{\mu}$. We denote the resulting experiment as $\mathbf{1NM}^{\mathsf{same},\mathsf{output}}_{\Gamma,\mathsf{A},\mathcal{F},\mathcal{G}}(\lambda, b)$, and in Definition 7 it only suffices to replace Eq. (2) with

$$\left\{\mathbf{1NM}^{\mathsf{same},\mathsf{output}}_{\Gamma,\mathsf{A},\mathcal{F},\mathcal{G}}(\lambda,0)\right\}_{\lambda\in\mathbb{N}} \overset{\mathrm{c}}{\approx} \left\{\mathbf{1NM}^{\mathsf{same},\mathsf{output}}_{\Gamma,\mathsf{A},\mathcal{F},\mathcal{G}}(\lambda,1)\right\}_{\lambda\in\mathbb{N}} \tag{3}$$

to obtain the new notion.

3.2 Families of Tampering Functions

Below, we define a few tampering families of interest for this paper.

Split-state tampering. Let Γ be a $(k, n_{\mathsf{L}} + n_{\mathsf{R}})$-code. In the split-state model, we think of a codeword $\gamma \in \{0,1\}^n$ as consisting of two parts $\gamma_{\mathsf{L}} \in \{0,1\}^{n_{\mathsf{L}}}, \gamma_{\mathsf{R}} \in \{0,1\}^{n_{\mathsf{R}}}$. Hence, we consider the following families of tampering and leakage functions:

$$\begin{aligned}\mathcal{F}_{\mathsf{split}}(n_{\mathsf{L}}, n_{\mathsf{R}}) &:= \{f = (f_{\mathsf{L}}, f_{\mathsf{R}}) : f_{\mathsf{L}} : \{0,1\}^{n_{\mathsf{L}}} \to \{0,1\}^{n_{\mathsf{L}}}, f_{\mathsf{R}}\{0,1\}^{n_{\mathsf{R}}} \to \{0,1\}^{n_{\mathsf{R}}}\}\\ \mathcal{G}_{\mathsf{split}}(n_{\mathsf{L}}, n_{\mathsf{R}}) &:= \left\{g = (g_{\mathsf{L}}, g_{\mathsf{R}}) : g_{\mathsf{L}} : \{0,1\}^{n_{\mathsf{L}}} \to \{0,1\}^{\ell}, g_{\mathsf{R}} : \{0,1\}^{n_{\mathsf{R}}} \to \{0,1\}^{\ell}\right\}.\end{aligned}$$

In this case, we simply say that Γ is ℓ-leakage-resilient super non-malleable w.r.t. message/codeword in the split-state model.

Decision trees. Let Γ be a (k, n)-code and $d \in \mathbb{N}$. Consider a binary tree of depth d whose internal nodes are labelled by numbers in $[n]$ and whose leaves contain values in $\{0,1\}$. Given a binary tree as above, we define a decision tree of depth d for $\{0,1\}^n$ as a Boolean function that takes as input a string $\gamma \in \{0,1\}^n$ and is described as follows:

- it starts from the root;
- it reads the label $i \in [n]$ of the node, and observes the i-th bit of the codeword $\gamma_i \in \{0,1\}$: if $\gamma_i = 0$, it descends to the left subtree, while if $\gamma_i = 1$, it moves to the right subtree;
- it outputs the value of the leaf at the end of the path.

We denote with $\mathcal{DT}^d(n)$ the set of all decision trees for $\{0,1\}^n$ with depth at most d. Hence, we consider the tampering family:

$$\mathcal{F}^d_{\mathsf{dtree}}(n) := \left\{f := (f_1, \ldots, f_n) : \forall i \in [n], f_i \in \mathcal{DT}^d(n)\right\},$$

and the leakage family $\mathcal{G}^d_{\mathsf{dtree}}(n) := \mathcal{DT}^d(n)$. In this case, we simply say that Γ is ℓ-leakage-resilient super non-malleable w.r.t. message/codeword against depth-d decision-tree tampering and leakage.

Bounded polynomial-time tampering. Let $S(\lambda), T(\lambda)$ be polynomials in the security parameter. A non-uniform algorithm A is described by a family of algorithms $\{\mathsf{A}_\lambda\}_{\lambda\in\mathbb{N}}$ (i.e., a different algorithm for each choice of the security parameter). Each A_λ corresponds to an algorithm whose input size is $n(\lambda)$, where $n : \mathbb{N} \to \mathbb{N}$. We say that a non-uniform algorithm A is S-size T-time if, for every input of size $n(\lambda)$ for some $\lambda \in \mathbb{N}$, the total work of the algorithm is at most $S(\lambda)$ and its parallel running time is upper bounded by $T(\lambda)$. We denote the family of non-uniform S-size T-time algorithms as $\mathcal{F}^{S,T}_{\mathsf{non\text{-}uni}}(n)$, and let

$$\mathcal{F}^{T}_{\mathsf{non\text{-}uni}}(n) := \bigcup_{S\in poly(\lambda)} \mathcal{F}^{S,T}_{\mathsf{non\text{-}uni}}(n).$$

3.3 Simple Facts

It is not hard to show that (super) non-malleability w.r.t. message is strictly weaker than (super) non-malleability w.r.t. codeword (e.g., consider a contrived code where we append a dummy bit to each codeword which is ignored by the decoding algorithm). It is also easy to see that non-malleability w.r.t. message/codeword is strictly weaker than super non-malleability w.r.t. message/codeword (e.g., consider a contrived code where we encode the message twice and where the decoding algorithm ignores the second copy of the codeword).

Below, we formalize three simple observations. (i) Assuming one-way functions, one can transform any (super) non-malleable code w.r.t. message into one w.r.t. codeword. (ii) For any (super) non-malleable code w.r.t. message/codeword there is a natural tradeoff between security and leakage resilience. (iii) In some cases, one-time super non-malleability w.r.t. codeword, along with leakage resilience, are sufficient to imply continuous non-malleability (in the setting of persistent tampering). All the above statements hold as long as the tampering family $\mathcal{F}$ and the leakage family $\mathcal{G}$ supported by the code are large enough (as detailed below). For simplicity, we stick to the plain model (but similar statements hold true in the CRS model).

Adding super non-malleability w.r.t. codeword. Let $\Gamma = (\mathsf{Enc}, \mathsf{Dec})$ be a code and $\Sigma = (\mathsf{Gen}, \mathsf{Sign}, \mathsf{SigVer})$ be a signature scheme. Consider the following derived code $\Gamma^* = (\mathsf{Enc}^*, \mathsf{Dec}^*)$.

Encoding: The encoding algorithm Enc^* takes as input a message $\mu \in \{0,1\}^k$, samples $(sk, vk) \leftarrow\!\!\$\ \mathsf{Gen}(1^\lambda)$, computes $\gamma \leftarrow\!\!\$\ \mathsf{Enc}(vk||\mu)$ and $\sigma \leftarrow\!\!\$\ \mathsf{Sign}(sk, \gamma)$, and outputs $\gamma^* = (\gamma, \sigma)$.

Decoding: The decoding algorithm Dec^* takes as input a codeword $\gamma^* = (\gamma, \sigma)$, and computes $\mu^* = vk||\mu = \mathsf{Dec}(\gamma)$. If either $\mu^* = \perp$ or $\mathsf{SigVer}(vk, \gamma, \sigma) = 0$, output $\perp$. Else output μ.

Let $\mathcal{F} \subseteq \{f : \{0,1\}^{n+s} \to \{0,1\}^{n+s}\}$, $\mathcal{G} \subseteq \{g : \{0,1\}^{n+s} \to \{0,1\}^*$ be families of functions. In the theorem below, for any function $f \in \mathcal{F}$, and any $\gamma \in \{0,1\}^n$ and $\sigma \in \{0,1\}^s$, we write $f(\gamma, \sigma)_1$ (resp. $f(\gamma, \sigma)_2$) for the function that outputs the first n bits (resp. the last s bits) of $f(\gamma, \sigma)$.

Theorem 1. *Assume that Σ is a strongly one-time unforgeable signature scheme with $\mathcal{M} = \{0,1\}^n$, $\mathcal{S} = \{0,1\}^s$ and $\mathcal{V} = \{0,1\}^v$, and that Γ is a $(\mathcal{G}(n), \ell + s)$-leakage-resilient persistent continuously $\mathcal{F}(n)$-super-non-malleable $(k + v, n)$-code w.r.t. message. Then, the above defined $(k, n + s)$-code Γ^* is $(\mathcal{G}(n + s), \ell)$-leakage-resilient persistent continuously $\mathcal{F}(n + s)$-super-non-malleable w.r.t. codeword, so long as for all $g \in \mathcal{G}(n + s)$, all $f \in \mathcal{F}(n + s)$, and all $(sk, vk) \in \mathsf{Gen}(1^\lambda)$ and $\rho \in \{0,1\}^*$, it holds that*

$$\mathcal{G}(n) \supseteq \{g(\cdot, \mathsf{Sign}(sk, \cdot; \rho)), f(\cdot, \mathsf{Sign}(sk, \cdot; \rho))_2, \} \tag{4}$$

$$\mathcal{F}(n) \supseteq \{f(\cdot, \mathsf{Sign}(sk, \cdot; \rho))_1, \mathsf{SigVer}(vk, f(\cdot, \mathsf{Sign}(sk, \cdot; \rho)))\}. \tag{5}$$

Intuitively, if the signature scheme is strongly unforgeable, then a tampering attacker cannot

maul γ^* while preserving vk. On the other hand, the security of the underlying non-malleable code guarantees that every change to vk makes the mauled message independent. The formal proof is in Appendix A.1.

Remark 2 (On compartmentalized tampering)**.** *Note that Theorem 1 does not immediately apply in the split-state setting where $\mathcal{F} = \mathcal{F}_{\mathsf{split}}(n,n)$ and $\mathcal{G} = \mathcal{G}_{\mathsf{split}}(n,n)$, because the conditions of Eq. (4) and Eq. (5) are not satisfied in general. However, we can slightly modify the code Γ^* by signing the left part γ_{L} and the right part γ_{R} of a codeword $\gamma = (\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}}) \in \{0,1\}^{2n}$ separately, yielding signatures σ_{L} and σ_{R}, and letting $\gamma^* = ((\gamma_{\mathsf{L}}, \sigma_{\mathsf{L}}), (\gamma_{\mathsf{R}}, \sigma_{\mathsf{R}}))$ to obtain the above result for the families $\mathcal{G}_{\mathsf{split}}(n+s, n+s)$ and $\mathcal{F}_{\mathsf{split}}(n+s, n+s)$.*

Adding leakage resilience. Next, we show how to use complexity leveraging in order to add leakage resilience to any strong-enough non-malleable code. The latter was already shown by Brian *et al.* [BFO+20] for the case of split-state tampering, who proved how leakage can be simulated by guessing and later verifying the accuracy of the guess. In particular, the security loss is exponential in the number of bits leaked, as the reduction correctly simulates the leakage only when the guess is exact. We observe that this can be generalized to the case where tampering via $\mathcal{F}$ can reveal whether the answer to a leakage query in $\mathcal{G}$ is correct. We call this property $\mathcal{G}$-friendliness.

Definition 8 (Leakage-friendly tampering)**.** *Let $\mathcal{F} \subseteq \{f : \{0,1\}^n \to \{0,1\}^n\}$ and $\mathcal{G} \subseteq \{g : \{0,1\}^n \to \{0,1\}^\ell\}$ be families of functions. We say that $\mathcal{F}$ is $\mathcal{G}$-leakage friendly if for all $g \in \mathcal{G}$, all $f \in \mathcal{F}$, and all strings $y \in \{0,1\}^\ell$ it holds that $\hat{f} \in \mathcal{F}$ where $\hat{f}$ is the function that upon input $\gamma \in \{0,1\}^n$ outputs $f(\gamma)$ if and only if $y = g(\gamma)$ (and outputs $\perp$ otherwise).*

Theorem 2. *Let $\mathcal{F} \subseteq \{f : \{0,1\}^n \to \{0,1\}^n\}$ and $\mathcal{G} \subseteq \{g : \{0,1\}^n \to \{0,1\}^\ell\}$ be families of functions such that $\mathcal{F}$ is $\mathcal{G}$-leakage friendly. Assume that Γ is persistent continuously $\mathcal{F}$-(super)-non-malleable w.r.t. message/codeword, with statistical security $\epsilon \in (0,1)$. Then, Γ is $\mathcal{G}$-leakage-resilient persistent continuously $\mathcal{F}$-(super)-non-malleable w.r.t. message/codeword, with statistical security $2^\ell \cdot \epsilon$, assuming that all the leakage is done before the first tampering query.*

Proof. For concreteness, we only prove the theorem for the case of continuous persistent super non-malleability w.r.t. codeword. The proof for the other cases is analogous. Let $\mathbf{G}(\lambda, b)$ be the experiment defining continuous persistent leakage-resilient super non-malleability of Γ, and let $\hat{\mathbf{G}}(\lambda, b)$ be the experiment defining continuous persistent super non-malleability of Γ (without leakage). By contradiction, assume there exists an unbounded attacker A such that

$$|\mathbb{P}[\mathbf{G}(\lambda,0) = 1] - \mathbb{P}[\mathbf{G}(\lambda,1) = 1]| > 2^\ell \cdot \epsilon.$$

Consider the following adversary $\hat{\mathsf{A}}$ attacking super non-malleability of Γ (without leakage).

1. Run $\mathsf{A}(1^\lambda)$.
2. Upon receiving (μ_0, μ_1) from A, forward (μ_0, μ_1) to the challenger.
3. Upon receiving the leakage query $g \in \mathcal{G}$ from A, return $y \leftarrow\!\!\$\ \{0,1\}^\ell$.
4. Upon receiving the first tampering query $f \in \mathcal{F}$, prepare the tampering function $\hat{f}(\gamma)$ that hard-wires f, g and the simulated leakage y and:
   - if $g(\gamma) = y$ returns $f(\gamma)$;
   - else, returns $\perp$.

Forward $\hat{f}$ to the challenger and send the answer back to A.

5. Process the following tampering query by returning $\perp$ if self-destruction occurred or forwarding them to the oracle and returning the corresponding answer otherwise.

6. Output whatever A does.

Let $\mathbf{W}_b$ be the event that $y = g(\gamma)$ in experiment $\mathbf{G}(\lambda, b)$. Note that since y is uniform and independent of b, it holds that $\mathbb{P}[\mathbf{W}_0] = \mathbb{P}[\mathbf{W}_1] = 2^{-\ell}$. Hence, we can write:

$$
\begin{aligned}
&\left|\mathbb{P}\left[\hat{\mathbf{G}}(\lambda,0)=1\right]-\mathbb{P}\left[\hat{\mathbf{G}}(\lambda,1)=1\right]\right| \\
&=\Big|\mathbb{P}[\mathbf{W}_0]\cdot\mathbb{P}\left[\hat{\mathbf{G}}(\lambda,0)=1|\mathbf{W}_0\right]-\mathbb{P}[\mathbf{W}_1]\cdot\mathbb{P}\left[\hat{\mathbf{G}}(\lambda,1)=1|\mathbf{W}_1\right] \\
&\qquad+\mathbb{P}\left[\overline{\mathbf{W}}_0\right]\cdot\mathbb{P}\left[\hat{\mathbf{G}}(\lambda,0)=1|\overline{\mathbf{W}}_0\right]-\mathbb{P}\left[\overline{\mathbf{W}}_1\right]\cdot\mathbb{P}\left[\hat{\mathbf{G}}(\lambda,1)=1|\overline{\mathbf{W}}_1\right]\Big| \\
&=\left|\mathbb{P}[\mathbf{W}_0]\cdot\mathbb{P}\left[\hat{\mathbf{G}}(\lambda,0)=1|\mathbf{W}_0\right]-\mathbb{P}[\mathbf{W}_1]\cdot\mathbb{P}\left[\hat{\mathbf{G}}(\lambda,1)=1|\mathbf{W}_1\right]\right| && (6)\\
&=2^{-\ell}\cdot\left|\mathbb{P}[\mathbf{G}(\lambda,0)=1]-\mathbb{P}[\mathbf{G}(\lambda,1)=1]\right| && (7)\\
&>\epsilon. && (8)
\end{aligned}
$$

In the above derivation, Eq. (6) follows because $\mathbb{P}\left[\overline{\mathbf{W}}_0\right] = \mathbb{P}\left[\overline{\mathbf{W}}_1\right]$ and moreover when we condition on $\overline{\mathbf{W}}_0$ and $\overline{\mathbf{W}}_1$ the view of $\hat{\mathsf{A}}$ is independent of b (and thus $\mathbb{P}\left[\hat{\mathbf{G}}(\lambda,0)=1|\overline{\mathbf{W}}_0\right] = \mathbb{P}\left[\hat{\mathbf{G}}(\lambda,1)=1|\overline{\mathbf{W}}_1\right] = 1/2$). Eq. (7) follows because $\mathbb{P}[\mathbf{W}_0] = \mathbb{P}[\mathbf{W}_1] = 2^{-\ell}$ and moreover when we condition on $\mathbf{W}_0$ and $\mathbf{W}_1$ the reduction $\hat{\mathsf{A}}$ perfectly simulates the view of A in $\mathbf{G}(\lambda, b)$. Finally, Eq. (8) follows by our assumption on A. The theorem follows. □

Remark 3 (On computational security)**.** *Theorem 2 also holds in the computational setting, so long as $\ell = O(\log \lambda)$. In fact, it even holds for $\ell = \omega(\log \lambda)$ assuming Γ has sub-exponential security.*

Remark 4 (On adaptive leakage)**.** *We can extend Theorem 2 to leakage families $\mathcal{G} \subseteq \{g : \{0,1\}^n \to \{0,1\}^*\}$, so long as the notion of leakage friendliness holds for up to q leakage functions. In this case, leakage resilience holds against adversaries making at most[9] q leakage queries.*

Achieving persistent continuous super non-malleability. Finally, we establish a connection between one-time super non-malleability and persistent continuous super non-malleability. Intuitively, one can simulate continuous tampering by leaking the index of the first tampering query that modifies the codeword and then obtaining the corresponding mauled codeword via a single tampering query. This connection was first outlined in [JW15], and later proven formally in [AKO17] in the split-state setting. We generalize this observation to general tampering families.

Theorem 3. *Let Γ be a $(\mathcal{G}(n), \ell + 1)$-leakage-resilient $\mathcal{F}(n)$-super-non-malleable (k, n)-code w.r.t. codeword. Assume that for every $q(\lambda) \in poly(\lambda)$, and every tuple of tampering functions $f^{(1)}, \ldots, f^{(q)} \in \mathcal{F}(n)$, the leakage family $\mathcal{G}(n)$ contains the function $\hat{g}(\gamma)$ that computes $(f^{(1)}(\gamma), \ldots, f^{(q)}(\gamma)))$ and returns 1 if and only if $f^{(1)}(\gamma) = \cdots f^{(q-1)}(\gamma) = \gamma$, but $f^{(q)}(\gamma) \neq \gamma$. Then, Γ is also a $(\mathcal{G}(n), \ell)$-leakage-resilient persistent continuously $\mathcal{F}(n)$-super-non-malleable (k, n)-code w.r.t. codeword.*

[9] Note that, e.g., $\mathcal{F}_{\mathsf{split}}$ is $\mathcal{G}_{\mathsf{split}}$-leakage friendly for any $q \in poly(\lambda)$.

The proof can be found in Appendix A.2.

Remark 5 (On super non-malleability w.r.t. codeword)**.** *Theorem 3 holds even starting with a super non-malleable code w.r.t. message, so long as the family $\mathcal{F}$ is closed under composition of poly-many functions (i.e., for all $q(\lambda) \in poly(\lambda)$ and all $f^{(1)}, \ldots, f^{(q)} \in \mathcal{F}$ the function $f^{(q)} \circ f^{(q-1)} \circ \cdots \circ f^{(1)}$ is contained in $\mathcal{F}$).*

4 Our Constructions

4.1 Decision-Tree Tampering

Our construction is inspired by [BGW19], with a few modifications that are necessary in order to prove persistent continuous super non-malleability w.r.t. codeword. To facilitate the description, let us introduce the following auxiliary function. For $n, c \in \mathbb{N}$, let $\phi : \{0,1\}^{c \log n} \to \mathcal{P}([n])$ be the function that, upon input a string $\zeta \in \{0,1\}^{c \log n}$ corresponding to c binary representations of distinct numbers in $[n]$, outputs the corresponding set of indices $\mathcal{I} \subseteq [n]$.

Our scheme is made of two layers, where the outer layer takes as input a split-state encoding of the message. Let $n, c, t \in \mathbb{N}$ be such that $t \geq c \log n$. Let $(\mathsf{Enc}_{\mathsf{RSS}}, \mathsf{Dec}_{\mathsf{RSS}})$ be a binary ramp secret sharing with messages in $\{0,1\}^t$. Consider the coding scheme $(\mathsf{Enc}^*_{n,c,t}, \mathsf{Dec}^*_{n,c,t})$ described below.

Algorithm $\mathsf{Enc}^*_{n,c,t}(\gamma)$**.** Upon input $\gamma \in \{0,1\}^c$:

1. Sample a random string ζ over the set of all strings of length $c \log n$ corresponding to c binary representations of distinct numbers in $[n]$.
2. Let $\mathcal{I} = \phi(\zeta)$ and let $\overline{\mathcal{I}} = [n] \setminus \mathcal{I}$.
3. Let γ^* be such that $\gamma^*[\mathcal{I}] = \gamma$ and $\gamma^*[\overline{\mathcal{I}}] = 0^{n-c}$.
4. Output $(\mathsf{Enc}_{\mathsf{RSS}}(\zeta), \gamma^*)$.

Algorithm $\mathsf{Dec}^*_{n,c,t}(\gamma_{\mathsf{RSS}}, \gamma^*)$**.** Proceed as follows:

1. Decode $\zeta = \mathsf{Dec}_{\mathsf{RSS}}(\gamma_{\mathsf{RSS}})$ and let $\mathcal{I} = \phi(\zeta)$.
2. If there exists $i \in [n] \setminus \mathcal{I}$ such that $\gamma^*[i] = 1$, return $\perp$.
3. Let $\gamma := \gamma^*[\mathcal{I}]$.
4. Output γ (or (γ, ζ) when ζ is explicitly needed).

We observe that the only difference between our version of the $\mathsf{Dec}^*_{n,c,t}$ algorithm and the one in [BGW19] is the check we perform in Item 2. This modification is required in order to obtain super non-malleability because, otherwise, an attacker could copy the original codeword into the 0 bits and then overwrite it with a constant valid codeword, and this would allow for the retrieval of the original encoding, and thus of the underlying message, in full.

We are now ready to define the final encoding scheme $\Gamma^* = (\mathsf{Enc}^*, \mathsf{Dec}^*)$ with security against decision-tree leakage and tampering. Let $m, n_{\mathsf{L}}, n_{\mathsf{R}}, c, t_{\mathsf{L}}, t_{\mathsf{R}}, s, v \in \mathbb{N}$ be parameters. Let $\Sigma = (\mathsf{Gen}, \mathsf{Sign}, \mathsf{SigVer})$ be a signature scheme with message space $\mathcal{M} = \{0,1\}^*$, signature space $\mathcal{S} = \{0,1\}^s$ and verification keys in $\mathcal{V} = \{0,1\}^v$. Let $\Gamma = (\mathsf{NMEnc}, \mathsf{NMDec})$ be a $(m + 2v, 2c)$-code. Let $\mathsf{Enc}_{\mathsf{L}} := \mathsf{Enc}^*_{n_{\mathsf{L}},c,t_{\mathsf{L}}}$, $\mathsf{Enc}_{\mathsf{R}} := \mathsf{Enc}^*_{n_{\mathsf{R}},c,t_{\mathsf{R}}}$, $\mathsf{Dec}_{\mathsf{L}} := \mathsf{Dec}^*_{n_{\mathsf{L}},c,t_{\mathsf{L}}}$, $\mathsf{Dec}_{\mathsf{R}} := \mathsf{Dec}^*_{n_{\mathsf{R}},c,t_{\mathsf{R}}}$.

Algorithm $\mathsf{Enc}^*(\mu)$**.** Upon input $\mu \in \{0,1\}^m$:

1. Sample $(sk_{\mathsf{L}}, vk_{\mathsf{L}}) \leftarrow\!\!\$\ \mathsf{Gen}(1^\lambda)$ and $(sk_{\mathsf{R}}, vk_{\mathsf{R}}) \leftarrow\!\!\$\ \mathsf{Gen}(1^\lambda)$.

2. Run $(\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}}) \leftarrow\!\!\$\ \mathsf{NMEnc}(\mu||vk_{\mathsf{L}}||vk_{\mathsf{R}})$.
3. Run $\gamma^*_{\mathsf{L}} \leftarrow\!\!\$\ \mathsf{Enc}_{\mathsf{L}}(\gamma_{\mathsf{L}})$ and $\gamma^*_{\mathsf{R}} \leftarrow\!\!\$\ \mathsf{Enc}_{\mathsf{R}}(\gamma_{\mathsf{R}})$.
4. Compute $\sigma_{\mathsf{L}} \leftarrow\!\!\$\ \mathsf{Sign}(sk_{\mathsf{L}}, \gamma^*_{\mathsf{L}})$ and $\sigma_{\mathsf{R}} \leftarrow\!\!\$\ \mathsf{Sign}(sk_{\mathsf{R}}, \gamma^*_{\mathsf{R}})$.
5. Output $\gamma^* := (\sigma_{\mathsf{L}}, \gamma^*_{\mathsf{L}}, \sigma_{\mathsf{R}}, \gamma^*_{\mathsf{R}})$.

Algorithm $\mathsf{Dec}^*(\gamma^*)$**.** Proceed as follows:

1. Parse $\gamma^* = (\sigma_{\mathsf{L}}, \gamma^*_{\mathsf{L}}, \sigma_{\mathsf{R}}, \gamma^*_{\mathsf{R}})$
2. Run $\gamma_{\mathsf{L}} = \mathsf{Dec}_{\mathsf{L}}(\gamma^*_{\mathsf{L}})$ and $\gamma_{\mathsf{R}} = \mathsf{Dec}_{\mathsf{R}}(\gamma^*_{\mathsf{R}})$.
3. Run $\mu||vk_{\mathsf{L}}||vk_{\mathsf{R}} = \mathsf{NMDec}(\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}})$.
4. Check that $\mathsf{SigVer}(vk_{\mathsf{L}}, \gamma^*_{\mathsf{L}}, \sigma_{\mathsf{L}}) = 1$ and $\mathsf{SigVer}(vk_{\mathsf{R}}, \gamma^*_{\mathsf{R}}, \sigma_{\mathsf{R}}) = 1$
5. Output μ, or $\perp$ if the above check fails.

We establish the following theorem.

Theorem 4. *Let Σ, Γ, and Γ^* be as above. Assume that Σ is a strongly one-time unforgeable signature scheme with signature length $s = \beta c$ for some $\beta \in (0,1)$, that Γ is an αc-leakage-resilient super non-malleable $(k + 2v, 2c)$-code w.r.t. codeword in the split-state model for some constant $\alpha < 1$, and that the privacy thresholds $t_{\mathsf{L}}, t_{\mathsf{R}}$ of the ramp secret sharing satisfy $t_{\mathsf{L}} \geq d$ and $t_{\mathsf{R}} \geq (4t_{\mathsf{L}} + c)d$. Then, the code Γ^* described above is a persistent continuously super non-malleable (m, n)-code against depth-d decision-tree tampering for $d = O(c^{1/4})$ and $n = O(c^2)$.*

Remark 6 (On simulating persistent continuous tampering)**.** *The definition of persistent continuous tampering states that the adversary A has unlimited access to the tampering oracle, unless A fails to produce a valid codeword, thus receiving $\perp$ in all subsequent tampering queries. However, we observe that this is equivalent to asking that A cannot send any more queries to the tampering oracle after receiving, as a result to a tampering query, a codeword $\tilde{\gamma} \in \{0,1\}^n \cup \{\perp\}$ which is different from $\diamond$. This is because, once obtained a tampered codeword which is either $\perp$ or a valid codeword, the adversary can simulate all the other queries on its own. Notice that this only holds in the case of super non-malleability, since A needs the tampered codeword in order to simulate the subsequent queries.*

The remainder of the section is dedicated to the proof of Theorem 4.

Establishing useful notation and procedures. For ease of notation, let $\mathcal{F}^d_{\mathsf{dtree}}(n) := \mathcal{F}^d_{\mathsf{dtree}}, \mathcal{F}_{\mathsf{split}}(c,c) := \mathcal{F}_{\mathsf{split}}, \mathcal{G}_{\mathsf{split}}(c,c) := \mathcal{G}_{\mathsf{split}}$, $\mathbf{G}^*_{\mathsf{A}}(\lambda, b) := \mathbf{CNM}^{\mathsf{cdw},\mathsf{super}}_{\Gamma^*,\mathsf{A},\mathcal{F}^d_{\mathsf{dtree}},\emptyset}(\lambda, b)$ and $\mathbf{G}_{\mathsf{A}}(\lambda, b) := \mathbf{1NM}^{\mathsf{cdw},\mathsf{super}}_{\Gamma,\mathsf{A},\mathcal{F}_{\mathsf{split}}}(\lambda, b)$. For all adversaries A against $\mathbf{G}^*_{\mathsf{A}}(\lambda, b)$, we can assume, without loss of generality, that A performs at most $p = poly(\lambda)$ tampering queries. Finally, let $\mathsf{Enc}^{\mathsf{L}}_{\mathsf{RSS}}$ be the instantiantion of $\mathsf{Enc}_{\mathsf{RSS}}$ used in $\mathsf{Enc}_{\mathsf{L}}$ and let $\mathsf{Enc}^{\mathsf{R}}_{\mathsf{RSS}}$ be the instantiation of $\mathsf{Enc}_{\mathsf{RSS}}$ used in $\mathsf{Enc}_{\mathsf{R}}$.

The proof is by reduction. In order to simplify the exposition, we define a template procedure $\mathsf{ObtainBits}$ which we will invoke several times in the actual proof. Informally, $\mathsf{ObtainBits}$ tries to evaluate the decision trees corresponding to the positions in $\mathcal{I}$ of the codeword tampered via f using the information $\mathcal{L}$ already known to the reduction itself; if some information is missing, it leaks it from the codeword. Since the reduction uses $\mathsf{ObtainBits}$ both inside and outside the leakage oracle, different sub-procedures are needed to leak these bits, depending on when $\mathsf{ObtainBits}$ is invoked. The formal definition follows.

Procedure $\mathsf{ObtainBits}_{\mathsf{Bit},\mathsf{Return}}(f, \mathcal{I}, \mathcal{L})$**:**

- **Instantiation:** A sub-procedure Bit taking as input an index $i \in [n]$ and returning a bit $b_i \in \{0,1\}$, and a sub-procedure Return taking as input the set $\mathcal{L}$ and a string $x \in \{0,1\}^*$ and returning some value.
- **Input:** A collection of decision trees f, a set of indices $\mathcal{I} \subseteq [n]$, a set $\mathcal{L} = \{(i,b) : i \in [n], b \in \{0,1\}\}$ such that if $(i_1,b_1),(i_2,b_2) \in \mathcal{L}$ and $i_1 = i_2$ then $b_1 = b_2$. Informally, $\mathcal{I}$ is the set of decision trees the procedure should compute and $\mathcal{L}$ is the prior knowledge of the algorithm invoking the procedure.

1. Let x be an initially empty string.
2. For all $i \in \mathcal{I}$, let initially $\mathsf{T} = f[i]$ and compute T as follows.
   (a) Let r be the label on the root of T.
   (b) If r is a leaf (i.e. $\mathsf{T} = \mathsf{r}$), then append r to x and step to the next index $i \in \mathcal{I}$ (or break the loop if all decision trees have been computed).
   (c) If there exists $(\mathsf{r}, b) \in \mathcal{L}$ for a $b \in \{0,1\}$, let $b^* = b$; otherwise, run $b^* \leftarrow \mathsf{Bit}(\mathsf{r})$ and replace $\mathcal{L} \leftarrow \mathcal{L} \cup \{(\mathsf{r}, b^*)\}$.
   (d) Replace T with its left subtree if $b^* = 0$ and with its right subtree if $b^* = 1$.
   (e) Go to Item 2a.
3. Run $y \leftarrow \mathsf{Return}(\mathcal{L}, x)$ and output y.

As for the sub-procedures, we define the following possibilities for the template argument Bit.

- **Procedure** $\mathsf{Leak}(i)$**:** Use the leakage oracle to leak the bit i from the split-state codeword.
- **Procedure** $\mathsf{Await}(i)$**:** Abort the procedure $\mathsf{ObtainBits}$, returning (await, i).

Finally, we define the following possibilities for the template argument Return.

- **Procedure** $\mathsf{Ready}(\mathcal{L}, x)$**:** Return (ready).
- **Procedure** $\mathsf{Check}(\mathcal{L}, x)$**:** Return 1 if x is the all 0 string and return 0 if x contains at least one 1.
- **Procedure** $\mathsf{Update}(\mathcal{L}, x)$**:** Return the updated set $\mathcal{L}$.

Notice that, when using the sub-procedure Leak, algorithm $\mathsf{ObtainBits}$ presents an undefined behaviour whenever there exists $(i,b) \notin \mathcal{L}$ such that i does not refer to any position on the split-state codeword; however, our reduction only invokes $\mathsf{ObtainBits}$ with sets $\mathcal{L}$ such that the only missing indices are indices which belong to the split-state codeword.

Ruling out signature forgeries. For any adversary A against $\mathbf{G}^*_{\mathsf{A}}(\lambda, b)$, $q \in [p], j \in \{\mathsf{L}, \mathsf{R}\}$, let $\mathbf{W}^{(q)}_j$ be the event in which the first $q-1$ tampering queries from A do not modify the codeword and the q-th tampering query $f^{(q)}$ is such that $f^{(q)}(\gamma^*) = (\tilde{\sigma}_{\mathsf{L}}, \tilde{\gamma}^*_{\mathsf{L}}, \tilde{\sigma}_{\mathsf{R}}, \tilde{\gamma}^*_{\mathsf{R}})$ satisfies (i) $\mathsf{NMDec}(\mathsf{Dec}_{\mathsf{L}}(\tilde{\gamma}^*_{\mathsf{L}}), \mathsf{Dec}_{\mathsf{R}}(\tilde{\gamma}^*_{\mathsf{R}})) = \tilde{\mu} || \tilde{vk}_{\mathsf{L}} || \tilde{vk}_{\mathsf{R}}$ with $\tilde{vk}_j = vk_j$ and (ii) $\mathsf{SigVer}(vk_j, \tilde{\gamma}^*_j, \tilde{\sigma}_j) = 1$ and $(\tilde{\gamma}^*_j, \tilde{\sigma}_j) \neq (\gamma^*_j, \sigma_j)$. Let $\mathbf{W} := \mathbf{W}_{\mathsf{L}} \cup \mathbf{W}_{\mathsf{R}}$, where $\mathbf{W}_{\mathsf{L}} := \bigcup_{q \in [p]} \mathbf{W}^{(q)}_{\mathsf{L}}$ and $\mathbf{W}_{\mathsf{R}} := \bigcup_{q \in [p]} \mathbf{W}^{(q)}_{\mathsf{R}}$. Informally, $\mathbf{W}$ is the event in which the adversary A against $\mathbf{G}^*_{\mathsf{A}}(\lambda, b)$ modifies the message but not the codeword, thus successfully forging a signature.

For $b \in \{0,1\}$, let $\mathbf{H}^*_{\mathsf{A}}(\lambda, b)$ be the experiment $\mathbf{G}^*_{\mathsf{A}}(\lambda, b)$ in which the challenger aborts whenever $\mathbf{W}$ happens. Clearly, the two experiments $\mathbf{G}^*$ and $\mathbf{H}^*$ are only distinguishable when $\mathbf{W}$ happens, therefore, if we show that $\mathbf{W}$ happens with negligible probability, it follows that $\mathbf{G}^*(\lambda, b)$ and $\mathbf{H}^*(\lambda, b)$ are statistically close.

Lemma 2. *For all PPT adversaries* A *there is a negligible function* $\nu : \mathbb{N} \to [0,1]$ *such that* $\Pr[\mathbf{W}] \leq \nu(\lambda)$.

Proof. We only prove that, when $b = 0$, $\mathbf{W}_\mathsf{L}$ happens with negligible probability, the proofs for $\mathbf{W}_\mathsf{R}$ and for $b = 1$ being the same; the lemma then follows by the union bound, as $\Pr[\mathbf{W}] \leq \Pr[\mathbf{W}_\mathsf{L}] + \Pr[\mathbf{W}_\mathsf{R}]$.

By contradiction, assume that there exists a PPT adversary A and a polynomial $p(\lambda) \in poly(\lambda)$ such that, when $b = 0$, A provokes the event $\mathbf{W}_\mathsf{L}$ with probability at least $1/p(\lambda)$. We construct a PPT attacker $\hat{\mathsf{A}}$ against one-time strong unforgeability of Σ as follows.

- Upon receiving vk from the challenger, run $\mathsf{A}(1^\lambda)$ and fix any $b \in \{0,1\}$.
- Upon receiving μ_0, μ_1 from A, let $vk_\mathsf{L} := vk$, sample $(sk_\mathsf{R}, vk_\mathsf{R}) \leftarrow\!\!\$\ \mathsf{Gen}(1^\lambda)$, compute $(\gamma_\mathsf{L}, \gamma_\mathsf{R}) \leftarrow\!\!\$\ \mathsf{NMEnc}(\mu_0 || vk_\mathsf{L} || vk_\mathsf{R})$, $\gamma^*_\mathsf{L} \leftarrow\!\!\$\ \mathsf{Enc}_\mathsf{L}(\gamma_\mathsf{L})$, $\gamma^*_\mathsf{R} \leftarrow\!\!\$\ \mathsf{Enc}_\mathsf{R}(\gamma_\mathsf{R})$ and $\sigma_\mathsf{R} \leftarrow\!\!\$\ \mathsf{Sign}(sk_\mathsf{R}, \gamma^*_\mathsf{R})$ and forward γ^*_L to the challenger, obtaining a signature $\sigma_\mathsf{L} \in \{0,1\}^s$.
- Let $\gamma^* := (\sigma_\mathsf{L}, \gamma^*_\mathsf{L}, \sigma_\mathsf{R}, \gamma^*_\mathsf{R})$.
- Upon receiving a tampering query $f \in \mathcal{F}^d_{\mathsf{dtree}}$ from A such that $f(\gamma^*) = \gamma^*$, do nothing.
- Upon receiving the first tampering query $f \in \mathcal{F}^d_{\mathsf{dtree}}$ from A such that $f(\gamma^*) \neq \gamma^*$, let $\tilde{\gamma}^* = f(\gamma^*)$, parse $(\tilde{\sigma}_\mathsf{L}, \tilde{\gamma}^*_\mathsf{L}, \tilde{\sigma}_\mathsf{R}, \tilde{\gamma}^*_\mathsf{R})$ and forward $(\tilde{\gamma}^*_\mathsf{L}, \tilde{\sigma}_\mathsf{L})$ to the challenger. If no such tampering query arrives (i.e. all tampering queries leave the codeword intact), abort.

For the analysis, note that the codeword γ^* is distributed exactly as in the experiment $\mathbf{G}^*(\lambda, 0)$ and thus the simulation of leakage queries is perfect. Moreover, when A provokes the event $\mathbf{W}_\mathsf{L}$, it holds that $(\tilde{\gamma}^*_\mathsf{L}, \tilde{\sigma}_\mathsf{L}) \neq (\gamma^*_\mathsf{L}, \sigma_\mathsf{L})$ and $\mathsf{SigVer}(vk_\mathsf{L}, \tilde{\gamma}^*_\mathsf{L}, \tilde{\sigma}_\mathsf{L}) = 1$. This concludes the proof of the lemma. □

Reducing to split-state non-malleability with augmented adversaries. Now we want to perform the reduction to the split-state super non-malleable code. Unfortunately, we cannot convert each decision-tree tampering query to a split-state tampering query because each conversion needs some leakage and we would end up leaking too many bits. However, the reduction only needs to simulate the first tampering query which actually modifies the codeword, because the answer to all previous queries is $\diamond$.

In order to apply this idea, we first define an experiment $\mathbf{H}^{\mathsf{aug}}$ which, informally, is the same of $\mathbf{G}$ except that the adversary is given one-time oracle access to the needed information, i.e. the index of the first tampering query actually modifying the codeword. The only remaining problem is that the reduction $\hat{\mathsf{A}}^{\mathsf{aug}}$ still performs too much leakage to verify whether the tampered codeword is valid or not; this is because the reduction should check that the padding string inside the codeword γ^* does not contain any 1. The solution is to give the reduction $\hat{\mathsf{A}}^{\mathsf{aug}}$ oracle access to this information too, so that now $\hat{\mathsf{A}}^{\mathsf{aug}}$ is able to simulate the experiment to the adversary A without performing too much leakage.

More formally, let A be an adversary telling apart $\mathbf{H}^*(\lambda, 0)$ and $\mathbf{H}^*(\lambda, 1)$ with non-negligible advantage and let $\mathsf{RunInfo}_\mathsf{A}(\tau)$ be a function which takes as input the transcript τ of the execution of $\mathsf{A}(1^\lambda; \rho_\mathsf{A})$ when the codeword is γ^* and outputs the index $q_\diamond$ of the first tampering query that is not answered $\diamond$ and a bit $b_\perp$ which is 1 if the output of such tampering query is $\perp$ and 0 otherwise. Notice that τ is uniquely determined by the random coins ρ_A of A and the decision-tree codeword $\gamma^* = \mathsf{comp}(\gamma)$ which is compiled by some deterministic compilation instructions comp from the split-state codeword γ. Therefore, we can define the oracle $\mathcal{O}^{\mathsf{aug}}_\mathsf{A}(\gamma, \mathsf{comp}, \rho_\mathsf{A})$ which is initialized with the split-state codeword γ and, upon receiving the query containing the instructions comp and the random coins ρ_A, computes τ and outputs $\mathsf{RunInfo}_\mathsf{A}(\tau)$. Consider the

experiment $\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}},\mathsf{A}}(\lambda, b)$ which is exactly the same as $\mathbf{G}_{\hat{\mathsf{A}}}$ except that $\hat{\mathsf{A}}$ is given one-time oracle access to $\mathcal{O}^{\mathsf{aug}}_{\mathsf{A}}(\gamma, \cdot, \cdot)$ before the tampering query.

Let $\mathcal{I}^{\mathsf{sgn}}_{\mathsf{L}}, \mathcal{I}^{\mathsf{rss}}_{\mathsf{L}}, \mathcal{I}^{\mathsf{str}}_{\mathsf{L}}, \mathcal{I}^{\mathsf{sgn}}_{\mathsf{R}}, \mathcal{I}^{\mathsf{rss}}_{\mathsf{R}}, \mathcal{I}^{\mathsf{str}}_{\mathsf{R}}$ be a partition of $[n]$ such that, given an encoding γ^*, $\mathcal{I}^{\mathsf{sgn}}_{\mathsf{L}}$ (resp. $\mathcal{I}^{\mathsf{sgn}}_{\mathsf{R}}$) contains the positions of γ^* in which is stored the left (resp. right) signature, $\mathcal{I}^{\mathsf{rss}}_{\mathsf{L}}$ (resp. $\mathcal{I}^{\mathsf{rss}}_{\mathsf{R}}$) contains the positions of γ^* in which is stored the left (resp. right) ramp secret sharing and $\mathcal{I}^{\mathsf{str}}_{\mathsf{L}}$ (resp. $\mathcal{I}^{\mathsf{str}}_{\mathsf{R}}$) contains the positions of γ^* in which is stored the string containing the left (resp. right) part of the codeword and the left (resp. right) zeroes. For $j \in \{\mathsf{L}, \mathsf{R}\}$, let $\mathcal{I}_j = \mathcal{I}^{\mathsf{sgn}}_j \cup \mathcal{I}^{\mathsf{rss}}_j \cup \mathcal{I}^{\mathsf{str}}_j$. We now show a reduction $\hat{\mathsf{A}}^{\mathsf{aug}}$ which is able to tell apart $\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\mathsf{A}}(\lambda, 0)$ and $\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\mathsf{A}}(\lambda, 1)$ with non-negligible advantage.

1. Sample random coins $\rho_{\mathsf{A}}, \rho^{\mathsf{enc}}_{\mathsf{L}}, \rho^{\mathsf{enc}}_{\mathsf{R}}, \rho^{\mathsf{sgn}}_{\mathsf{L}}, \rho^{\mathsf{sgn}}_{\mathsf{R}}$ and random strings $\zeta_{\mathsf{L}}, \zeta_{\mathsf{R}}$.
2. For $j \in \{\mathsf{L}, \mathsf{R}\}$, let $\mathcal{I}^{\mathsf{cdw}}_j = \phi(\zeta_j)$ and $\mathcal{I}^{\mathsf{zero}}_j = \mathcal{I}^{\mathsf{str}}_j \setminus \mathcal{I}^{\mathsf{cdw}}_j$.
3. Compute $\omega_{\mathsf{L}} = \mathsf{Enc}^{\mathsf{L}}_{\mathsf{RSS}}(\zeta_{\mathsf{L}}; \rho^{\mathsf{enc}}_{\mathsf{L}})$ and $\omega_{\mathsf{R}} = \mathsf{Enc}^{\mathsf{R}}_{\mathsf{RSS}}(\zeta_{\mathsf{R}}; \rho^{\mathsf{enc}}_{\mathsf{R}})$.
4. Sample $(sk_{\mathsf{L}}, vk_{\mathsf{L}}) \leftarrow\!\!\$\ \mathsf{Gen}(1^\lambda)$ and $(sk_{\mathsf{R}}, vk_{\mathsf{R}}) \leftarrow\!\!\$\ \mathsf{Gen}(1^\lambda)$.
5. Run $\mathsf{A}(1^\lambda; \rho_{\mathsf{A}})$, obtaining the challenge messages μ_0, μ_1; then construct the challenge messages $\mu^*_0 := \mu_0 || vk_{\mathsf{L}} || vk_{\mathsf{R}}$ and $\mu^*_1 := \mu_1 || vk_{\mathsf{L}} || vk_{\mathsf{R}}$ and send μ^*_0, μ^*_1 to the challenger.
6. For $j \in \{\mathsf{L}, \mathsf{R}\}$, construct the leakage function g^{sgn}_j which hard-wires the values $\rho^{\mathsf{enc}}_j, \rho^{\mathsf{sgn}}_j, \zeta_j$, ω_j, sk_j and, upon input the codeword part γ_j, computes $\gamma^*_j = \mathsf{Enc}_j(\gamma_j; \rho^{\mathsf{enc}}_j, \zeta_j)$ and $\sigma_j = \mathsf{Sign}(sk_j, \gamma^*_j; \rho^{\mathsf{sgn}}_j)$ and outputs σ_j.
7. Send $(g^{\mathsf{sgn}}_{\mathsf{L}}, g^{\mathsf{sgn}}_{\mathsf{R}})$ to the leakage oracle, thus obtaining the signatures $(\sigma_{\mathsf{L}}, \sigma_{\mathsf{R}})$.
8. Let $\mathcal{L}$ be a set which, initially, contains all the pairs (i, b) such that $i \in [n] \setminus \mathcal{I}^{\mathsf{cdw}}_{\mathsf{L}} \cup \mathcal{I}^{\mathsf{cdw}}_{\mathsf{R}}$ and $b = \gamma^*[i]$. Notice that the only bits unknown to $\hat{\mathsf{A}}$ are the ones belonging to the split-state codeword, namely, the ones in $\mathcal{I}^{\mathsf{cdw}}_{\mathsf{L}}$ and in $\mathcal{I}^{\mathsf{cdw}}_{\mathsf{R}}$; therefore, $\hat{\mathsf{A}}$ is able to construct the set $\mathcal{L}$.
9. Using the information in $\mathcal{L}$, construct the compilation information comp which is used to compile the split-state codeword γ into the decision-tree codeword $\gamma^* = \mathsf{comp}(\gamma)$.
10. Send $(\mathsf{comp}, \rho_{\mathsf{A}})$ to the augmented oracle $\mathcal{O}^{\mathsf{aug}}_{\mathsf{A}}$, thus receiving a pair $(q_\diamond, b_\perp)$.
11. Upon receiving the q-th tampering query $f^{(q)} \in \mathcal{F}^d_{\mathsf{dtree}}$ from A, return $\diamond$ if $q < q_\diamond$; otherwise let $f = f^{(q)}$ and do the following.
    (a) *Obtaining the necessary bits:* run the procedure
    $$\mathcal{L} \leftarrow \mathsf{ObtainBits}_{\mathsf{Leak},\mathsf{Update}}(f, \mathcal{I}^{\mathsf{sgn}}_{\mathsf{L}} \cup \mathcal{I}^{\mathsf{rss}}_{\mathsf{L}} \cup \mathcal{I}^{\mathsf{sgn}}_{\mathsf{R}} \cup \mathcal{I}^{\mathsf{rss}}_{\mathsf{R}}, \mathcal{L}),$$
    then use the set $\mathcal{L}$ to compute the tampered signatures $\tilde{\sigma}_{\mathsf{L}}, \tilde{\sigma}_{\mathsf{R}}$ and the tampered encodings $\tilde{\omega}_{\mathsf{L}}, \tilde{\omega}_{\mathsf{R}}$.
    (b) *Obtaining the new positions:* for $j \in \{\mathsf{L}, \mathsf{R}\}$, compute $\tilde{\zeta}_j = \mathsf{Dec}_{\mathsf{RSS}}(\tilde{\omega}_j)$ and let $\tilde{\mathcal{I}}^{\mathsf{cdw}}_j = \phi(\tilde{\zeta}_j)$ and $\tilde{\mathcal{I}}^{\mathsf{zero}}_j = \tilde{\mathcal{I}}^{\mathsf{str}}_j \setminus \tilde{\mathcal{I}}^{\mathsf{cdw}}_j$.
    (c) *Leaking the remaining bits for the left part:* construct the leakage function $\hat{g}^{\mathcal{L}}_{\mathsf{L}}$ which hard-wires (a description of) the tampering function f, the sets $\mathcal{I}^{\mathsf{cdw}}_{\mathsf{L}}$ and $\tilde{\mathcal{I}}^{\mathsf{cdw}}_{\mathsf{L}}$ and

the set $\mathcal{L}$ and, upon input the left part γ_{L} of the codeword, constructs the set $\mathcal{L}_{\mathsf{L}} = \{(i, \gamma^*[i]) : i \in \mathcal{I}^{\mathsf{cdw}}_{\mathsf{L}}\}$, runs the procedure

$$y \leftarrow \mathsf{ObtainBits}_{\mathsf{Await},\mathsf{Ready}}(f, \tilde{\mathcal{I}}^{\mathsf{cdw}}_{\mathsf{L}}, \mathcal{L} \cup \mathcal{L}_{\mathsf{L}}) \tag{9}$$

and returns y. Then, send $(\hat{g}^{\mathcal{L}}_{\mathsf{L}}, \epsilon)$ to the leakage oracle, thus obtaining a value y. If $y = (\mathsf{await}, i)$ for some $i \in \mathcal{I}^{\mathsf{cdw}}_{\mathsf{R}}$, leak $\gamma^*[i]$ from the right part of the codeword, update $\mathcal{L} \leftarrow \mathcal{L} \cup \{(i, \gamma^*[i])\}$ and repeat this step.

(d) *Leaking the remaining bits for the right part:* construct the leakage function $\hat{g}^{\mathcal{L}}_{\mathsf{R}}$ which hard-wires (a description of) the tampering function f, the sets $\mathcal{I}^{\mathsf{cdw}}_{\mathsf{R}}$ and $\tilde{\mathcal{I}}^{\mathsf{cdw}}_{\mathsf{R}} \cup \tilde{\mathcal{I}}^{\mathsf{zero}}_{\mathsf{R}}$ and the set $\mathcal{L}$ and, upon input the right part γ_{R} of the codeword, constructs the set $\mathcal{L}_{\mathsf{R}} = \{(i, \gamma^*[i]) : i \in \mathcal{I}^{\mathsf{cdw}}_{\mathsf{R}}\}$, runs the procedure

$$y \leftarrow \mathsf{ObtainBits}_{\mathsf{Await},\mathsf{Ready}}(f, \tilde{\mathcal{I}}^{\mathsf{cdw}}_{\mathsf{R}} \cup \tilde{\mathcal{I}}^{\mathsf{zero}}_{\mathsf{R}}, \mathcal{L} \cup \mathcal{L}_{\mathsf{R}})$$

and returns y. Then, send $(\epsilon, \hat{g}^{\mathcal{L}}_{\mathsf{R}})$ to the leakage oracle, thus obtaining a value y. If $y = (\mathsf{await}, i)$ for some $i \in \mathcal{I}^{\mathsf{cdw}}_{\mathsf{L}}$, leak $\gamma^*[i]$ from the left part of the codeword, update $\mathcal{L} \leftarrow \mathcal{L} \cup \{(i, \gamma^*[i])\}$ and repeat this step.

(e) *Validating the right part:* construct the leakage function $\hat{h}^{\mathsf{chk}}_{\mathsf{R}}$ which hard-wires (a description of) the tampering function f, the sets $\mathcal{I}^{\mathsf{cdw}}_{\mathsf{R}}$ and $\tilde{\mathcal{I}}^{\mathsf{cdw}}_{\mathsf{R}} \cup \tilde{\mathcal{I}}^{\mathsf{zero}}_{\mathsf{R}}$ and the set $\mathcal{L}$ and, upon input the right part γ_{R} of the codeword, constructs the set $\mathcal{L}_{\mathsf{R}} = \{(i, \gamma^*[i]) : i \in \mathcal{I}^{\mathsf{cdw}}_{\mathsf{R}}\}$, runs the procedure

$$b_{\mathsf{valid}} \leftarrow \mathsf{ObtainBits}_{\mathsf{Await},\mathsf{Check}}(f, \tilde{\mathcal{I}}^{\mathsf{cdw}}_{\mathsf{R}} \cup \tilde{\mathcal{I}}^{\mathsf{zero}}_{\mathsf{R}}, \mathcal{L} \cup \mathcal{L}_{\mathsf{R}})$$

and returns b_{valid}. Then, send $(\epsilon, \hat{h}^{\mathsf{chk}}_{\mathsf{R}})$ to the leakage oracle, thus obtaining the bit b_{valid}. If $b_{\mathsf{valid}} = 0$, abort the simulation and return a random guess.

(f) *Tampering with the codeword:* for $j \in \{\mathsf{L}, \mathsf{R}\}$, construct the tampering function $\hat{f}_j$ which hard-wires the strings $\sigma_{\mathsf{L}}, \sigma_{\mathsf{R}}, \omega_{\mathsf{L}}, \omega_{\mathsf{R}}$, the sets $\mathcal{I}^{\mathsf{cdw}}_{\mathsf{L}}, \mathcal{I}^{\mathsf{cdw}}_{\mathsf{R}}, \mathcal{L}$ and (a description of) the tampering query f and, upon input the codeword part γ_j, computes the tampered codeword part $\tilde{\gamma}_j$ by using the additional bits given by $\mathcal{L}$ and then returns $\tilde{\gamma}_j$. Send the query $(\hat{f}_{\mathsf{L}}, \hat{f}_{\mathsf{R}})$ to the tampering oracle, thus obtaining a codeword $\tilde{\gamma} \in \{0,1\}^{2c} \cup \{\diamond, \perp\}$. If $\tilde{\gamma} \in \{\diamond, \perp\}$, abort the simulation and return a random guess. Otherwise, let $\tilde{\gamma} = (\tilde{\gamma}_{\mathsf{L}}, \tilde{\gamma}_{\mathsf{R}})$, reconstruct $\tilde{\gamma}^*_{\mathsf{L}}$ (resp. $\tilde{\gamma}^*_{\mathsf{R}}$) using the value $\tilde{\gamma}_{\mathsf{L}}$ and the set $\mathcal{I}^{\mathsf{cdw}}_{\mathsf{L}}$ (resp. the value $\tilde{\gamma}_{\mathsf{R}}$ and the set $\mathcal{I}^{\mathsf{cdw}}_{\mathsf{R}}$) and set $\tilde{\gamma}^* = (\tilde{\sigma}_{\mathsf{L}}, \tilde{\gamma}^*_{\mathsf{L}}, \tilde{\sigma}_{\mathsf{R}}, \tilde{\gamma}^*_{\mathsf{R}})$.

(g) *Checking the signature:* reconstruct the tampered message $\tilde{\mu}||\tilde{vk}_{\mathsf{L}}||\tilde{vk}_{\mathsf{R}}$. Then, check that $\tilde{vk}_{\mathsf{L}} \neq vk_{\mathsf{L}}$ and $\tilde{vk}_{\mathsf{R}} \neq vk_{\mathsf{R}}$, compute $b_{\mathsf{L}} = \mathsf{SigVer}(vk_{\mathsf{L}}, \tilde{\sigma}_{\mathsf{L}}, \tilde{\gamma}^*_{\mathsf{L}})$ and $b_{\mathsf{R}} = \mathsf{SigVer}(vk_{\mathsf{R}}, \tilde{\sigma}_{\mathsf{R}}, \tilde{\gamma}^*_{\mathsf{R}})$, check that $b_{\mathsf{L}} = b_{\mathsf{R}} = 1$ and abort the simulation returning a random guess if one of the previous checks fails.

Finally, set $\tilde{\gamma}^* = \perp$ if $b_{\perp} = 1$, return $\tilde{\gamma}^*$ to A and output the same distinguishing bit as A.

For the analysis, notice that the reduction $\hat{\mathsf{A}}^{\mathsf{aug}}$ perfectly simulates $\mathbf{H}^*(\lambda, b)$ to A unless the leakage performed exceeds the admissible leakage $\alpha c - 1$; therefore, when the leakage is within the bounds, $\hat{\mathsf{A}}^{\mathsf{aug}}$ has the same advantage of A.

The following lemma allows us to conclude that $\mathbf{G}_{\hat{\mathsf{A}}}(\lambda, 0)$ and $\mathbf{G}_{\hat{\mathsf{A}}}(\lambda, 1)$ are computationally close.

Lemma 3. *If there exists an adversary $\hat{\mathsf{A}}^{\mathsf{aug}}$ which is able to distinguish between $\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\mathsf{A}}(\lambda, 0)$ and $\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\mathsf{A}}(\lambda, 1)$ with non-negligible advantage, then there exists an adversary $\hat{\mathsf{A}}$ which is able to distinguish between $\mathbf{G}_{\hat{\mathsf{A}}}(\lambda, 0)$ and $\mathbf{G}_{\hat{\mathsf{A}}}(\lambda, 1)$ with non-negligible advantage.*

Proof. Consider an adversary $\hat{\mathsf{A}}^{\mathsf{aug}}$ telling apart the experiments $\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\mathsf{A}}(\lambda,0)$ and $\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\mathsf{A}}(\lambda,1)$ with non-negligible advantage. In particular, assume that, without loss of generality, $\hat{\mathsf{A}}^{\mathsf{aug}}$ outputs 1 more often when playing experiment $\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\mathsf{A}}(\lambda,1)$. Let $(\mathbf{q}_\diamond, \mathbf{b}_\perp)$ be the random variable for the output of $\mathsf{RunInfo}_{\mathsf{A}}$.

Consider the following reduction $\hat{\mathsf{A}}$ playing the experiment $\mathbf{G}_{\hat{\mathsf{A}}}(\lambda, b)$.

1. Sample random coins $\rho_{\hat{\mathsf{A}}^{\mathsf{aug}}}$, a random index $\hat{q}_\diamond \in [p+1]$ and a random bit $\hat{b}_\perp \in \{0,1\}$.
2. Run $\hat{\mathsf{A}}^{\mathsf{aug}}(1^\lambda; \rho_{\hat{\mathsf{A}}^{\mathsf{aug}}})$ and, upon receiving the challenge messages μ_0^*, μ_1^*, forward them to the challenger.
3. Keep track of all the leakage queries g_q from $\hat{\mathsf{A}}^{\mathsf{aug}}$, along with their respective answers τ_q.
4. Upon receiving a query $(\mathsf{comp}, \rho_{\mathsf{A}})$ for $\mathcal{O}^{\mathsf{aug}}$, answer with $(\hat{q}_\diamond, \hat{b}_\perp)$.
5. Upon receiving the tampering query f, if any:
   (a) Construct the leakage function $\hat{g}_{\mathsf{chk}}$ which hard-wires $\rho_{\hat{\mathsf{A}}^{\mathsf{aug}}}$, (a description of) the adversary $\hat{\mathsf{A}}^{\mathsf{aug}}$, (a description of) the leakage queries g_q and their respective answers τ_q and, upon input the left codeword γ_{L}, constructs the set $\hat{\mathcal{S}}_{\mathsf{R}}$ of all the possible $\hat{\gamma}_{\mathsf{R}}$ such that $(\gamma_{\mathsf{L}}, \hat{\gamma}_{\mathsf{R}})$ is either an encoding of μ_0^* or an encoding of μ_1^* and also it is compatible with the answer to the first $\hat{q}_\diamond - 1$ queries resulting in $\diamond$ and the answer to the $\hat{q}_\diamond$-th query being a valid codeword if $b_\perp = 0$ or an invalid codeword if $b_\perp = 1$. If $\hat{\mathcal{S}}_{\mathsf{R}} = \emptyset$, the function returns $b_{\mathsf{keep}} = 0$; otherwise, g_{chk} runs $\hat{\mathsf{A}}^{\mathsf{aug}}(1^\lambda; \rho_{\hat{\mathsf{A}}^{\mathsf{aug}}})$ once for each codeword in $\{\gamma_{\mathsf{L}}\} \times \hat{\mathcal{S}}_{\mathsf{R}}$ and returns $b_{\mathsf{keep}} = 1$ if A returns 0 more often when the target codeword $(\gamma_{\mathsf{L}}, \hat{\gamma}_{\mathsf{R}})$ is an encoding of μ_0 or returns $b_{\mathsf{keep}} = 0$ otherwise.
   (b) Send $(\hat{g}_{\mathsf{chk}}, \epsilon)$ to the leakage oracle, thus receiving the bit b_{keep}. If $b_{\mathsf{keep}} = 0$, abort the simulation and return a random guess.
   (c) Send f to the tampering oracle and forward the answer to $\hat{\mathsf{A}}^{\mathsf{aug}}$.
6. Return the same distinguishing bit of $\hat{\mathsf{A}}^{\mathsf{aug}}$.

The reduction runs in exponential time and performs the same amount of leakage of $\hat{\mathsf{A}}^{\mathsf{aug}}$, plus 1 bit for the output of the leakage function $\hat{g}_{\mathsf{chk}}$. For the analysis, notice that, when the reduction guesses correctly the index $\mathbf{q}_\diamond$ and the bit $\mathbf{b}_\perp$, the simulation is perfect. We have that, for $b \in \{0,1\}$,

$$
\begin{aligned}
&\Pr\left[\mathbf{G}_{\hat{\mathsf{A}}}(\lambda,b)=1\right] \\
&\quad= \Pr\left[\mathbf{G}_{\hat{\mathsf{A}}}(\lambda,b)=1 \wedge b_{\mathsf{keep}}=1\right] + \frac{1}{2}\Pr\left[b_{\mathsf{keep}}=0\right] && (10)\\
&\quad= \Pr\left[\mathbf{G}_{\hat{\mathsf{A}}}(\lambda,b)=1 \wedge \mathbf{q}_\diamond=\hat{q}_\diamond \wedge \mathbf{b}_\perp=\hat{b}_\perp\right] + \frac{1}{2}\Pr\left[b_{\mathsf{keep}}=0\right] && (11)\\
&\qquad+ \Pr\left[\mathbf{G}_{\hat{\mathsf{A}}}(\lambda,b)=1 \wedge b_{\mathsf{keep}}=1 \wedge \left(\mathbf{q}_\diamond\neq\hat{q}_\diamond \vee \mathbf{b}_\perp\neq\hat{b}_\perp\right)\right] \\
&\quad= \frac{1}{2(p+1)}\Pr\left[\mathbf{G}_{\hat{\mathsf{A}}}(\lambda,b)=1 \middle| \mathbf{q}_\diamond=\hat{q}_\diamond \wedge \mathbf{b}_\perp=\hat{b}_\perp\right] + \frac{1}{2}\Pr\left[b_{\mathsf{keep}}=0\right] && (12)\\
&\qquad+ \Pr\left[\mathbf{G}_{\hat{\mathsf{A}}}(\lambda,b)=1 \wedge b_{\mathsf{keep}}=1 \wedge \left(\mathbf{q}_\diamond\neq\hat{q}_\diamond \vee \mathbf{b}_\perp\neq\hat{b}_\perp\right)\right] \\
&\quad= \frac{1}{2(p+1)}\Pr\left[\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\hat{\mathsf{A}}}(\lambda,b)=1\right] + \frac{1}{2}\Pr\left[b_{\mathsf{keep}}=0\right] && (13)\\
&\qquad+ \Pr\left[\mathbf{G}_{\hat{\mathsf{A}}}(\lambda,b)=1 \wedge b_{\mathsf{keep}}=1 \wedge \left(\mathbf{q}_\diamond\neq\hat{q}_\diamond \vee \mathbf{b}_\perp\neq\hat{b}_\perp\right)\right],
\end{aligned}
$$

where Eq. (10) holds because, when $b_{\mathsf{keep}} = 0$, $\hat{\mathsf{A}}$ makes a random guess, in Eq. (11) we split the event $b_{\mathsf{keep}} = 1$ in two parts, namely, when the reduction guesses the correct values for $\mathbf{q}_\diamond = \hat{q}_\diamond$ and $\mathbf{b}_\perp = \hat{b}_\perp$ and when the advantage is preserved even if the reduction fails to simulate correctly the experiment, Eq. (12) follows from the fact that $(\mathbf{q}_\diamond, \mathbf{b}_\perp)$ is uniformly sampled in a set of $2(p+1)$ possible choices and, finally, Eq. (13) holds because, when the guess $\mathbf{q}_\diamond = \hat{q}_\diamond, \mathbf{b}_\perp = \hat{b}_\perp$ is correct, the simulation is perfect and $\hat{\mathsf{A}}$ outputs the same distinguishing bit of $\hat{\mathsf{A}}^{\mathsf{aug}}$. By putting it together,

$$
\begin{aligned}
&\left|\Pr\left[\mathbf{G}_{\hat{\mathsf{A}}}(\lambda,0)=1\right]-\Pr\left[\mathbf{G}_{\hat{\mathsf{A}}}(\lambda,1)=1\right]\right| \\
&\quad=\left|\frac{1}{2(p+1)}\left(\Pr\left[\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\mathsf{A}}(\lambda,0)=1\right]-\Pr\left[\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\mathsf{A}}(\lambda,1)=1\right]\right)\right. \qquad (14)\\
&\qquad+\Pr\left[\mathbf{G}_{\hat{\mathsf{A}}}(\lambda,0)=1\wedge b_{\mathsf{keep}}=1\wedge\left(\mathbf{q}_\diamond\neq\hat{q}_\diamond\vee\mathbf{b}_\perp\neq\hat{b}_\perp\right)\right]\\
&\qquad\left.-\Pr\left[\mathbf{G}_{\hat{\mathsf{A}}}(\lambda,1)=1\wedge b_{\mathsf{keep}}=1\wedge\left(\mathbf{q}_\diamond\neq\hat{q}_\diamond\vee\mathbf{b}_\perp\neq\hat{b}_\perp\right)\right]\right| \\
&\quad\geq\frac{1}{2(p+1)}\left|\Pr\left[\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\mathsf{A}}(\lambda,0)=1\right]-\Pr\left[\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\mathsf{A}}(\lambda,1)=1\right]\right|, \qquad (15)
\end{aligned}
$$

where Eq. (14) directly follows from Eq. (13), in which the terms $\Pr[b_{\mathsf{keep}}=0]$ cancel each other and Eq. (15) follows from the fact that, when $b_{\mathsf{valid}} = 1$, A does not invert the advantage, hence the expressions

$$\Pr\left[\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\mathsf{A}}(\lambda,0)=1\right]-\Pr\left[\mathbf{H}^{\mathsf{aug}}_{\hat{\mathsf{A}}^{\mathsf{aug}},\mathsf{A}}(\lambda,1)=1\right]$$

and

$$
\begin{aligned}
&\Pr\left[\mathbf{G}_{\hat{\mathsf{A}}}(\lambda,0)=1\wedge b_{\mathsf{keep}}=1\wedge\left(\mathbf{q}_\diamond\neq\hat{q}_\diamond\vee\mathbf{b}_\perp\neq\hat{b}_\perp\right)\right]\\
&\quad-\Pr\left[\mathbf{G}_{\hat{\mathsf{A}}}(\lambda,1)=1\wedge b_{\mathsf{keep}}=1\wedge\left(\mathbf{q}_\diamond\neq\hat{q}_\diamond\vee\mathbf{b}_\perp\neq\hat{b}_\perp\right)\right]
\end{aligned}
$$

have the same sign. □

Bounding the leakage. It remains to bound the leakage made by the reduction.

Proposition 1 ([BGW19, Proposition 1])**.** *Let $n, c, t \in \mathbb{N}$ such that $t \geq c \log n$. Let A be an arbitrary algorithm that reads adaptively at most t bits of $(\mathsf{Enc}_{\mathsf{RSS}}(\zeta), \phi(\zeta))$. Let $\mathbf{Y}$ denote the number of distinct 1's in $\phi(\zeta)$ which are read by A. Then, over the randomness of ζ and $\mathsf{Enc}_{\mathsf{RSS}}$,*

$$\Pr\left[\mathbf{Y}\geq\frac{2tc}{n}\right]\leq\exp\left(-\frac{tc}{3n}\right).$$

Lemma 4. *Suppose $t_{\mathsf{R}} \geq (4t_{\mathsf{L}} + c + 2s)d$. Let $\ell^{\mathsf{bit}}_{\mathsf{R}}$ be the amount of positions b leaked from γ_{R}. Then, for any $\gamma \in \{0,1\}^{2c}$, the event that $\ell^{\mathsf{bit}}_{\mathsf{R}} \geq 2(4t_{\mathsf{R}} + 4t_{\mathsf{L}} + c + 2s)dc/n_{\mathsf{R}}$ happens with probability at most $(4d+1)\exp(-t_{\mathsf{R}}c/3n_{\mathsf{R}})$.*

Proof. Fix $\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}}$ and the randomness for $\mathsf{Enc}_{\mathsf{L}}(\gamma_{\mathsf{L}})$. Note that the reduction $\hat{\mathsf{A}}$ needs to produce $4t_{\mathsf{R}} + 4t_{\mathsf{L}} + c + 2s$ coordinates for the left part, namely $4t_{\mathsf{R}}$ to reconstruct the tampered "seed" $\tilde{\zeta}_{\mathsf{R}}$, $4t_{\mathsf{L}}$ to reconstruct the tampered "seed" $\tilde{\zeta}_{\mathsf{L}}$, $2s$ to obtain the signatures $(\tilde{\sigma}_{\mathsf{L}}, \tilde{\sigma}_{\mathsf{R}})$ and, finally, the at most c locations specified by $\tilde{\mathcal{I}}^{\mathsf{cdw}}_{\mathsf{L}} = G(\tilde{\zeta}_{\mathsf{L}})$. All these bits can be partitioned into $4d+1$ subsets of at most t_{R}/d bits each, and each decision tree tampering one of these bits makes at most d queries to $\mathsf{Enc}_{\mathsf{R}}(\gamma_{\mathsf{R}})$ (since it makes at most d queries in total). Therefore, the reduction, in order to simulate the tampering of the bits required to obtain $(\tilde{\sigma}_{\mathsf{L}}, \tilde{\gamma}^*_{\mathsf{L}})$, makes at most t_{R} queries to $\mathsf{Enc}_{\mathsf{R}}(\gamma_{\mathsf{R}})$ for each subset. By Proposition 1, for each subset, the event that $\hat{\mathsf{A}}$ queries

more than $2t_{\mathsf{R}}c/n_{\mathsf{R}}$ locations happens with probability at most $\exp(-t_{\mathsf{R}}c/3n_{\mathsf{R}})$. Finally, by a union bound over these subsets, the probability that $\ell_{\mathsf{R}}^{\mathsf{bit}} \geq 2(4t_{\mathsf{R}}+4t_{\mathsf{L}}+c+2s)dc/n_{\mathsf{R}}$ is at most $(4d+1)\exp(-t_{\mathsf{R}}c/3n_{\mathsf{R}})$. The lemma follows. □

Lemma 5. *Suppose $t_{\mathsf{L}} \geq d$. Let $\ell_{\mathsf{L}}^{\mathsf{bit}}$ be the amount of positions b leaked from γ_{L}. Then, for any $\gamma \in \{0,1\}^{2c}$, the event that $\ell_{\mathsf{L}}^{\mathsf{bit}} \geq 2(4t_{\mathsf{L}}+4t_{\mathsf{R}}+n_{\mathsf{R}}+2s)dc/n_{\mathsf{L}}$ happens with probability at most $(4t_{\mathsf{L}}+4t_{\mathsf{R}}+n_{\mathsf{R}}+2s)d/t_{\mathsf{L}}\exp(-t_{\mathsf{L}}c/3n_{\mathsf{L}})$.*

Proof. Fix $\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}}$ and the randomness for $\mathsf{Enc}_{\mathsf{R}}(\gamma_{\mathsf{R}})$. Note that the reduction $\hat{\mathsf{A}}$ needs to produce $4t_{\mathsf{L}}+4t_{\mathsf{R}}+n_{\mathsf{R}}+2s$ coordinates for the right part, namely $4t_{\mathsf{L}}$ to reconstruct the tampered "seed" $\tilde{\zeta}_{\mathsf{L}}$, $4t_{\mathsf{R}}$ to reconstruct the tampered "seed" $\tilde{\zeta}_{\mathsf{R}}$, $2s$ to obtain the signatures $(\tilde{\sigma}_{\mathsf{L}}, \tilde{\sigma}_{\mathsf{R}})$ and, finally, the at most n_{R} locations for the string which combines the right codeword $\tilde{\zeta}_{\mathsf{L}}$ and the string of zeroes. All these bits can be partitioned into $(4t_{\mathsf{L}}+4t_{\mathsf{R}}+n_{\mathsf{R}}+2s)d/t_{\mathsf{L}}$ subsets of at most t_{L}/d bits each, and each decision tree tampering one of these bits makes at most d queries to $\mathsf{Enc}_{\mathsf{L}}(\gamma_{\mathsf{L}})$ (since it makes at most d queries in total). Therefore, the reduction, in order to simulate the tampering of the bits required to obtain $(\tilde{\sigma}_{\mathsf{R}}, \tilde{\gamma}_{\mathsf{R}}^*)$, makes at most t_{L} queries to $\mathsf{Enc}_{\mathsf{L}}(\gamma_{\mathsf{L}})$ for each subset. By Proposition 1, for each subset, the event that $\hat{\mathsf{A}}$ queries more than $2t_{\mathsf{L}}c/n_{\mathsf{R}}$ locations happens with probability at most $\exp(-t_{\mathsf{L}}c/3n_{\mathsf{L}})$. Finally, by a union bound over these subsets, the probability that $\ell_{\mathsf{L}}^{\mathsf{bit}} \geq 2(4t_{\mathsf{L}}+4t_{\mathsf{R}}+n_{\mathsf{R}}+2s)dc/n_{\mathsf{L}}$ is at most $(4t_{\mathsf{L}}+4t_{\mathsf{R}}+n_{\mathsf{R}}+2s)d/t_{\mathsf{L}}\exp(-t_{\mathsf{L}}c/3n_{\mathsf{L}})$. The lemma follows. □

Let

$$\begin{aligned}\ell^{\mathsf{tamp}} &= \left(\ell_{\mathsf{L}}^{\mathsf{bit}} + \ell_{\mathsf{R}}^{\mathsf{bit}}\right)(1+\log(c)) \\ &= 2dc\left(\frac{4t_{\mathsf{L}}+4t_{\mathsf{R}}+n_{\mathsf{R}}+2s}{n_{\mathsf{L}}} + \frac{4t_{\mathsf{R}}+4t_{\mathsf{L}}+c+2s}{n_{\mathsf{R}}}\right)(1+\log(c)).\end{aligned}$$

By the above lemmas, the event that the amount of leakage performed by $\hat{\mathsf{A}}$ exceeds $\ell^{\mathsf{tamp}}+2s+2$ (recall that the reduction also leaks $2s$ bits for the signatures and 2 bits for checking the simulation) happens with probability at most

$$(4d+1)\exp(-t_{\mathsf{R}}c/3n_{\mathsf{R}}) + (4t_{\mathsf{L}}+4t_{\mathsf{R}}+n_{\mathsf{R}}+2s)d/t_{\mathsf{L}}\exp(-t_{\mathsf{L}}c/3n_{\mathsf{L}}). \tag{16}$$

Lemma 6. *Fix $\alpha \in (0,1)$. Then, there exist constants $\eta_1, \eta_2, \eta_3, \eta_4$ (only dependent on α) such that, if $t_{\mathsf{L}} = \eta_1 c\log n$, $t_{\mathsf{R}} = \eta_2 dc\log n\log c$, $n_{\mathsf{L}} = \eta_3 d^3 c\log n\log^3 c$, $n_{\mathsf{R}} = \eta_4 d^2 c\log n\log^2 c$, then $\ell^{\mathsf{tamp}} \leq \alpha c$ with overwhelming probability.*

Proof. Since $t_{\mathsf{R}} \geq (4t_{\mathsf{L}}+c+2s)d$,

$$4t_{\mathsf{L}}+4t_{\mathsf{R}}+c+2s \leq 5t_{\mathsf{R}},$$

thus simplifying both the Eq. (16) and the expression for ℓ^{tamp}. Indeed, for ℓ^{tamp} we have that

$$\begin{aligned}\ell^{\mathsf{tamp}} &\leq 2d\left(\frac{5t_{\mathsf{R}}+n_{\mathsf{R}}}{n_{\mathsf{L}}} + \frac{5t_{\mathsf{R}}}{n_{\mathsf{R}}}\right)(1+\log c)\,c \\ &= \left(10\frac{\eta_2}{\eta_3 d\log^2 c} + 2\frac{\eta_4}{\eta_3\log c} + 10\frac{\eta_2}{\eta_4\log c}\right)(1+\log c)\,c \leq \alpha c.\end{aligned}$$

Similarly, for some constant η_5, we have that Eq. (16) is bounded by

$$\begin{aligned}&(4d+1)\exp(-t_{\mathsf{R}}c/3n_{\mathsf{R}}) + (5t_{\mathsf{R}}+n_{\mathsf{R}})d/t_{\mathsf{L}}\exp(-t_{\mathsf{L}}c/3n_{\mathsf{L}}) \\ &\quad\leq (4d+1)\exp\left(-\frac{\eta_2 c}{3\eta_4 d\log c}\right) + \eta_5 d^3\log^2 c\exp\left(-\frac{\eta_1 c}{3\eta_3 d^3\log^3 c}\right).\end{aligned}$$

Hence, by letting $d = O(c^{1/4})$, the above quantity is negligible in c. □

By choosing the parameters as in Lemma 6, the length of the final codeword satisfies

$$n = 2s + 4t_{\mathsf{L}} + 4t_{\mathsf{R}} + n_{\mathsf{L}} + n_{\mathsf{R}} = O(d^3 c \log n \log^3 c),$$

which can be rewritten as $n/\log n = O(c^{7/4} \log^3 c)$, thus making $n = O(c^2)$ a good approximation, and the total amount of leakage is $\ell = \ell^{\mathsf{tamp}} + 2\beta c + 2$, which, with a good choice of the parameters $\eta_1, \ldots, \eta_4$ and α, β, can simply be rewritten as $\ell \leq \alpha c$. This concludes the proof of Theorem 4. □

4.2 Bounded Polynomial-Depth Tampering

Our construction for bounded polynomial-depth tampering, works in three steps.

(i) First, we show a compiler for turning any *leakage-resilient* non-malleable code into a leakage-resilient *super* non-malleable code; the compiler is non-black-box, as it relies on NIZK proofs, and thus yields a code in the CRS model (even if the initial code is in the plain model).

(ii) Second, we show how to instantiate the above compiler by simplifying the non-malleable code for bounded polynomial-depth tampering of Dachman-Soled *et al.* [DKP21] (thanks to the fact that we rely on trusted setup).

(iii) Third, we argue that the family of bounded polynomial-depth tampering satisfies the conditions of Theorem 3, so that persistent continuous non-malleability follows by steps (i) and (ii).

Let $\Gamma = (\mathsf{Enc}, \mathsf{Dec})$ be a (k, n)-code with randomness space $\{0,1\}^r$, let $\mathsf{G} : \{0,1\}^s \to \{0,1\}^r$ be a PRG, and let $\Pi = (\mathsf{CRSGen}, \mathsf{Prove}, \mathsf{ProofVer})$ be a non-interactive argument system with proof space $\mathcal{P} = \{0,1\}^m$ for the relation:

$$\mathcal{R} = \left\{ (\gamma, \sigma) \in \{0,1\}^n \times \{0,1\}^s : \exists \mu \in \{0,1\}^k \text{ s.t. } \gamma = \mathsf{Enc}(\mu; \mathsf{G}(\sigma)) \right\}. \tag{17}$$

Consider the following $(k, n+m)$-code $\Gamma^* = (\mathsf{Init}^*, \mathsf{Enc}^*, \mathsf{Dec}^*)$ in the CRS model.

Initialization: The initialization algorithm Init^* outputs $\omega \leftarrow\!\!\$\ \mathsf{CRSGen}(1^\lambda)$.

Encoding: The encoding algorithm Enc^* proceeds as follows:

- sample a uniformly random seed $\sigma \leftarrow\!\!\$\ \{0,1\}^s$ and compute $\rho = \mathsf{G}(\sigma)$;
- let $\gamma = \mathsf{Enc}(\mu; \rho)$;
- run $\pi \leftarrow\!\!\$\ \mathsf{Prove}(\omega, \gamma, \sigma)$;
- return $\gamma^* = (\gamma, \pi)$.

Decoding: The decoding algorithm Dec^*, upon input a codeword $\gamma^* = (\gamma, \pi)$, proceeds as follows:

- run $\mathsf{ProofVer}(\omega, \gamma, \pi)$, and output $\perp$ if the verification fails;
- compute $\mu = \mathsf{Dec}(\gamma)$, and output $\perp$ if the decoding fails;
- else, return μ.

Let $\mathcal{F} \subseteq \{f : \{0,1\}^{n+m} \to \{0,1\}^{n+m}\}$ be a family of functions. In the theorem below, for any function $f \in \mathcal{F}$, and any $\gamma \in \{0,1\}^n$ and $\pi \in \{0,1\}^m$, we write $f(\gamma, \pi)_1$ (resp. $f(\gamma, \pi)_2$) for the function that outputs the first n bits (resp. the last m bits) of $f(\gamma, \pi)$.

Theorem 5. *Assume that* Π *is a one-time simulation extractable non-interactive zero-knowledge argument system for the relation of Eq. (17), with proof space* $\mathcal{P} = \{0,1\}^m$*, with zero-knowledge simulator* $\mathsf{S} = (\mathsf{S}_0, \mathsf{S}_1)$ *and with extractor* K*. Let* Γ *be a* $(\mathcal{G}(n), \ell + s + m)$*-leakage-resilient* $\mathcal{F}(n)$*-non-malleable* (k, n)*-code w.r.t. message/codeword.*

Then, the above defined $(k, n+m)$*-code* Γ^* *is* $(\mathcal{G}(n+m), \ell)$*-leakage-resilient* $\mathcal{F}(n+m)$*-super-non-malleable w.r.t. message/codeword, so long as for every* $g \in \mathcal{G}(n+m)$ *and every* $f \in \mathcal{F}(n+m)$*, all* $\gamma \in \{0,1\}^n$*, all* $\pi \in \{0,1\}^m$*, and all* $(\omega, \zeta, \xi) \in \mathsf{S}_0(1^\lambda)$*, it holds that:*

$$\mathcal{G} \supseteq \{g(\cdot, \mathsf{S}_1(\zeta, \cdot))_1, \mathsf{ProofVer}(\omega, f(\cdot, \mathsf{S}_1(\zeta, \cdot))), \mathsf{K}(\xi, f(\cdot, \mathsf{S}_1(\zeta, \cdot)))\} \tag{18}$$

$$\mathcal{F} \ni f(\cdot, \mathsf{S}_1(\zeta, \cdot))_1. \tag{19}$$

The proof of the above theorem appears in Appendix A.3.

Instantiating the proof system. Since the underlying code Γ needs to tolerate at least m bits of leakage, where m is the size of a proof under Π, Theorem 5 implicitly requires proofs that are sub-linear in the size of the statement (which is a codeword), but not of the witness (which is a seed for the PRG). In the literature, such proofs are referred as Succinct Non-interactive Arguments of Knowledge (SNARKs). In [BPR20], the authors present a simulation-extractable SNARK whose proofs consist of 4 group elements. The security proof relies on both the generic group model (GGM) and the random oracle model (ROM).

Alternatively, we can use [GGI+15], where fully-homomorphic encryption (FHE) and NIZK argument systems are used to achieve succinct-proof NIZK argument systems for all of NP. The succinct proof for an NP relation $\mathcal{R}$ is built as follows:

- The witness x is encrypted with key σ into a ciphertext u of the same length by means of a symmetric-key encryption scheme (namely, one-time pad with a pseudorandom key generated from σ via a PRG G).
- The key generation algorithm of the FHE is called with randomness ρ to get (pk, sk). Next, the FHE scheme is used with keys (pk, sk) and randomness τ to encrypt the symmetric key σ into a ciphertext z. Then, the FHE evaluation algorithm takes as input the ciphertext z, and the NP relation $\mathcal{R}$ over statement y and witness $u \oplus \mathsf{G}(\cdot)$, and returns a ciphertext v.
- The underlying prover provides an argument π for the statement (pk, z, v) and witness (ρ, σ, τ), proving that (pk, sk) are generated according to ρ, that z is an encryption of σ according to pk, τ and that v decrypts to 1.
- The succinct proof is given by (pk, z, u, π).

Since $|u| = |x|$ and (pk, z, π) are polynomial in the security parameter, the proof size is $|x| + poly(\lambda)$. Also note that (pk, z, u, π) is sufficient to verify the proof, as one can obtain v and then call the underlying verification algorithm.

In their work, Gentry *et al.* [GGI+15] show that this transformation preserves the soundness and the zero-knowledge property of the underlying NIZK argument system. However, their result also applies to simulation extractability. For a high-level idea, call A an adversary against the simulation-extractability of the succinct-proof scheme. Assume that, given a simulated proof (pk, z, u, π) for a statement y of its choice, A manages to produce an accepting and fresh pair $(\tilde{y}, (\tilde{pk}, \tilde{u}, \tilde{z}, \tilde{\pi}))$. Consider the extractor that takes as input $(\tilde{y}, (\tilde{pk}, \tilde{u}, \tilde{z}, \tilde{\pi}))$ and as trapdoor (pk, sk), and does the following. If $\tilde{pk} \neq pk$, it computes the homomorphic evaluation $\tilde{v}$ of the circuit $\mathcal{R}(\tilde{y}, \tilde{u} \oplus \mathsf{G}(\cdot))$ on ciphertext $\tilde{z}$ with $\tilde{pk}$. Then, it runs the underlying extractor over $((\tilde{pk}, \tilde{z}, \tilde{v}), \tilde{\pi})$ to get $(\tilde{\rho}, \tilde{\sigma}, \tilde{\tau})$. If $\tilde{pk} = pk$, the extractor only needs to decrypt $\tilde{z}$ to get $\tilde{\sigma}$. In both cases, it outputs $\tilde{x} = \tilde{u} \oplus \mathsf{G}(\tilde{\sigma})$.

Instantiating the underlying code. To instantiate the underlying code, we start from the construction of Dachman-Soled *et al.* [DKP21] which is in the plain model and relies on key-less hash functions, time-lock puzzles, as well as other standard assumptions. In the CRS model, their construction can be simplified as follows: The encoding of a message μ consists of a time-lock puzzle ζ computed using μ (with some fixed difficulty parameter) and a simulation-extractable NIZK proof of knowledge π of the message μ inside the puzzle. We refer the reader to Appendix B for the formal description and the security analysis in the CRS model.

Proving continuous non-malleability. Finally, we invoke Theorem 3 to conclude persistent continuous non-malleability. To do that, we need to check that the leakage family of bounded polynomial-depth circuits contains the function $\hat{g}$ in the statement of the theorem. In our case, it suffices to consider leakage resilience against circuits of depth $\leq T + c$ for a small constant c, and compute the leakage function $\hat{g}$ as follows. Upon input the codeword γ, consider q parallel sub-circuits, where the i-th circuit computes $f^{(i)}(\gamma)$, and outputs $b_i = 1$ if $f^{(i)}(\gamma) = \gamma$, $b_i = 0$ otherwise. The circuit will then output 1 if $b_1 = \cdots = b_{q-1} = 0$ and $b_q = 1$, and 0 otherwise. By inspection, every sub-circuit has depth $\leq T + c$, as it computes a tampering function and a bit-wise comparison (feasible in constant depth). To check if $b_1 = \cdots = b_{q-1} = 0$, it suffices to compute $b = \mathbf{OR}(b_1, \ldots, b_{q-1})$. The leakage function finally outputs $\mathbf{AND}(\mathbf{NOT}(b), b_q)$.

5 Conclusions

We have shown how to achieve continuous non-malleability in two natural settings: (i) decision-tree tampering, and (ii) bounded polynomial-depth tampering. The first result is in the plain model; the second result requires trusted setup. Both constructions rely on computational assumptions (one-way functions in (i), and time-lock puzzles and simulation-extractable succinct-proof NIZKs in (ii)). Natural open problems include: removing computational assumptions from our construction in (i), and weakening the assumptions from our construction in (ii). We leave these as interesting directions for future research.

Our paper provides the first crucial insights for constructing continuously non-malleable codes against non-compartmentalized tampering. In particular:

- We prove for the first time that security against non-persistent global tampering is impossible in the continuous setting.
- We prove for the first time that, when the target tampering family is powerful enough, continuous non-malleability follows from one-time super non-malleability with log bits of leakage resilience. The latter, in particular, is true for bounded-depth tampering and for AC0 tampering.
- We show a generic transform to reduce one-time *super* non-malleability to one-time non-malleability using NIZK proofs; this transform requires the underlying tampering family to satisfy certain properties, which are met in the setting of bounded polynomial-depth tampering.

We believe the above observations are important, and will turn useful for future constructions of continuously non-malleable codes against other non-compartmentalized tampering families (e.g., AC0 tampering), possibly under weaker assumptions.

References

[ABPP14] Michel Abdalla, Fabrice Benhamouda, Alain Passelègue, and Kenneth G. Paterson. Related-key security for pseudorandom functions beyond the linear barrier. In Juan A. Garay and Rosario Gennaro, editors, *CRYPTO 2014, Part I*, volume 8616 of *LNCS*, pages 77–94. Springer, Heidelberg, August 2014.

[ADL14] Divesh Aggarwal, Yevgeniy Dodis, and Shachar Lovett. Non-malleable codes from additive combinatorics. In David B. Shmoys, editor, *46th ACM STOC*, pages 774–783. ACM Press, May / June 2014.

[ADN+19] Divesh Aggarwal, Nico Döttling, Jesper Buus Nielsen, Maciej Obremski, and Erick Purwanto. Continuous non-malleable codes in the 8-split-state model. In Yuval Ishai and Vincent Rijmen, editors, *EUROCRYPT 2019, Part I*, volume 11476 of *LNCS*, pages 531–561. Springer, Heidelberg, May 2019.

[AGM+15] Shashank Agrawal, Divya Gupta, Hemanta K. Maji, Omkant Pandey, and Manoj Prabhakaran. Explicit non-malleable codes against bit-wise tampering and permutations. In Rosario Gennaro and Matthew J. B. Robshaw, editors, *CRYPTO 2015, Part I*, volume 9215 of *LNCS*, pages 538–557. Springer, Heidelberg, August 2015.

[AHI11] Benny Applebaum, Danny Harnik, and Yuval Ishai. Semantic security under related-key attacks and applications. In Bernard Chazelle, editor, *ICS 2011*, pages 45–60. Tsinghua University Press, January 2011.

[AKO17] Divesh Aggarwal, Tomasz Kazana, and Maciej Obremski. Inception makes non-malleable codes stronger. In Yael Kalai and Leonid Reyzin, editors, *TCC 2017, Part II*, volume 10678 of *LNCS*, pages 319–343. Springer, Heidelberg, November 2017.

[AKO+21] Divesh Aggarwal, Bhavana Kanukurthi, Sai Lakshmi Bhavana Obbattu, Maciej Obremski, and Sruthi Sekar. Rate one-third non-malleable codes. Cryptology ePrint Archive, Report 2021/1042, 2021. https://eprint.iacr.org/2021/1042.

[AO20] Divesh Aggarwal and Maciej Obremski. A constant rate non-malleable code in the split-state model. In *61st FOCS*, pages 1285–1294. IEEE Computer Society Press, November 2020.

[BC10] Mihir Bellare and David Cash. Pseudorandom functions and permutations provably secure against related-key attacks. In Tal Rabin, editor, *CRYPTO 2010*, volume 6223 of *LNCS*, pages 666–684. Springer, Heidelberg, August 2010.

[BCM11] Mihir Bellare, David Cash, and Rachel Miller. Cryptography secure against related-key attacks and tampering. In Dong Hoon Lee and Xiaoyun Wang, editors, *ASIACRYPT 2011*, volume 7073 of *LNCS*, pages 486–503. Springer, Heidelberg, December 2011.

[BDG+18] Marshall Ball, Dana Dachman-Soled, Siyao Guo, Tal Malkin, and Li-Yang Tan. Non-malleable codes for small-depth circuits. In Mikkel Thorup, editor, *59th FOCS*, pages 826–837. IEEE Computer Society Press, October 2018.

[BDK+19] Marshall Ball, Dana Dachman-Soled, Mukul Kulkarni, Huijia Lin, and Tal Malkin. Non-malleable codes against bounded polynomial time tampering. In Yuval Ishai

and Vincent Rijmen, editors, *EUROCRYPT 2019, Part I*, volume 11476 of *LNCS*, pages 501–530. Springer, Heidelberg, May 2019.

[BDKM16] Marshall Ball, Dana Dachman-Soled, Mukul Kulkarni, and Tal Malkin. Non-malleable codes for bounded depth, bounded fan-in circuits. In Marc Fischlin and Jean-Sébastien Coron, editors, *EUROCRYPT 2016, Part II*, volume 9666 of *LNCS*, pages 881–908. Springer, Heidelberg, May 2016.

[BDKM18] Marshall Ball, Dana Dachman-Soled, Mukul Kulkarni, and Tal Malkin. Non-malleable codes from average-case hardness: AC^0, decision trees, and streaming space-bounded tampering. In Jesper Buus Nielsen and Vincent Rijmen, editors, *EUROCRYPT 2018, Part III*, volume 10822 of *LNCS*, pages 618–650. Springer, Heidelberg, April / May 2018.

[BDL97] Dan Boneh, Richard A. DeMillo, and Richard J. Lipton. On the importance of checking cryptographic protocols for faults (extended abstract). In Walter Fumy, editor, *EUROCRYPT'97*, volume 1233 of *LNCS*, pages 37–51. Springer, Heidelberg, May 1997.

[BDL22] Marshall Ball, Dana Dachman-Soled, and Julian Loss. (Nondeterministic) hardness vs. non-malleability. Cryptology ePrint Archive, Report 2022/070, 2022. https://eprint.iacr.org/2022/070.

[BFO+20] Gianluca Brian, Antonio Faonio, Maciej Obremski, Mark Simkin, and Daniele Venturi. Non-malleable secret sharing against bounded joint-tampering attacks in the plain model. In Daniele Micciancio and Thomas Ristenpart, editors, *CRYPTO 2020, Part III*, volume 12172 of *LNCS*, pages 127–155. Springer, Heidelberg, August 2020.

[BFV21] Gianluca Brian, Antonio Faonio, and Daniele Venturi. Continuously non-malleable secret sharing: Joint tampering, plain model and capacity. In Kobbi Nissim and Brent Waters, editors, *TCC 2021, Part II*, volume 13043 of *LNCS*, pages 333–364. Springer, Heidelberg, November 2021.

[BGW19] Marshall Ball, Siyao Guo, and Daniel Wichs. Non-malleable codes for decision trees. In Alexandra Boldyreva and Daniele Micciancio, editors, *CRYPTO 2019, Part I*, volume 11692 of *LNCS*, pages 413–434. Springer, Heidelberg, August 2019.

[Bih94] Eli Biham. New types of cryptanalytic attacks using related keys (extended abstract). In Tor Helleseth, editor, *EUROCRYPT'93*, volume 765 of *LNCS*, pages 398–409. Springer, Heidelberg, May 1994.

[BK03] Mihir Bellare and Tadayoshi Kohno. A theoretical treatment of related-key attacks: RKA-PRPs, RKA-PRFs, and applications. In Eli Biham, editor, *EUROCRYPT 2003*, volume 2656 of *LNCS*, pages 491–506. Springer, Heidelberg, May 2003.

[BPR20] Karim Baghery, Zaira Pindado, and Carla Ràfols. Simulation extractable versions of groth's zk-SNARK revisited. In Stephan Krenn, Haya Shulman, and Serge Vaudenay, editors, *CANS 20*, volume 12579 of *LNCS*, pages 453–461. Springer, Heidelberg, December 2020.

[BPT12] Mihir Bellare, Kenneth G. Paterson, and Susan Thomson. RKA security beyond the linear barrier: IBE, encryption and signatures. In Xiaoyun Wang and Kazue Sako, editors, *ASIACRYPT 2012*, volume 7658 of *LNCS*, pages 331–348. Springer, Heidelberg, December 2012.

[BS97] Eli Biham and Adi Shamir. Differential fault analysis of secret key cryptosystems. In Burton S. Kaliski Jr., editor, *CRYPTO'97*, volume 1294 of *LNCS*, pages 513–525. Springer, Heidelberg, August 1997.

[CCHM19] Binyi Chen, Yilei Chen, Kristina Hostáková, and Pratyay Mukherjee. Continuous space-bounded non-malleable codes from stronger proofs-of-space. In Alexandra Boldyreva and Daniele Micciancio, editors, *CRYPTO 2019, Part I*, volume 11692 of *LNCS*, pages 467–495. Springer, Heidelberg, August 2019.

[CDTV16] Sandro Coretti, Yevgeniy Dodis, Björn Tackmann, and Daniele Venturi. Non-malleable encryption: Simpler, shorter, stronger. In Eyal Kushilevitz and Tal Malkin, editors, *TCC 2016-A, Part I*, volume 9562 of *LNCS*, pages 306–335. Springer, Heidelberg, January 2016.

[CEGL08] Ran Canetti, Dror Eiger, Shafi Goldwasser, and Dah-Yoh Lim. How to protect yourself without perfect shredding. In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfsdóttir, and Igor Walukiewicz, editors, *ICALP 2008, Part II*, volume 5126 of *LNCS*, pages 511–523. Springer, Heidelberg, July 2008.

[CFV19] Sandro Coretti, Antonio Faonio, and Daniele Venturi. Rate-optimizing compilers for continuously non-malleable codes. In Robert H. Deng, Valérie Gauthier-Umaña, Martín Ochoa, and Moti Yung, editors, *ACNS 19*, volume 11464 of *LNCS*, pages 3–23. Springer, Heidelberg, June 2019.

[CGL16] Eshan Chattopadhyay, Vipul Goyal, and Xin Li. Non-malleable extractors and codes, with their many tampered extensions. In Daniel Wichs and Yishay Mansour, editors, *48th ACM STOC*, pages 285–298. ACM Press, June 2016.

[CKR16] Nishanth Chandran, Bhavana Kanukurthi, and Srinivasan Raghuraman. Information-theoretic local non-malleable codes and their applications. In Eyal Kushilevitz and Tal Malkin, editors, *TCC 2016-A, Part II*, volume 9563 of *LNCS*, pages 367–392. Springer, Heidelberg, January 2016.

[CMTV15] Sandro Coretti, Ueli Maurer, Björn Tackmann, and Daniele Venturi. From single-bit to multi-bit public-key encryption via non-malleable codes. In Yevgeniy Dodis and Jesper Buus Nielsen, editors, *TCC 2015, Part I*, volume 9014 of *LNCS*, pages 532–560. Springer, Heidelberg, March 2015.

[CQZ+16] Yu Chen, Baodong Qin, Jiang Zhang, Yi Deng, and Sherman S. M. Chow. Non-malleable functions and their applications. In Chen-Mou Cheng, Kai-Min Chung, Giuseppe Persiano, and Bo-Yin Yang, editors, *PKC 2016, Part II*, volume 9615 of *LNCS*, pages 386–416. Springer, Heidelberg, March 2016.

[CZ14] Eshan Chattopadhyay and David Zuckerman. Non-malleable codes against constant split-state tampering. In *55th FOCS*, pages 306–315. IEEE Computer Society Press, October 2014.

[DK19] Dana Dachman-Soled and Mukul Kulkarni. Upper and lower bounds for continuous non-malleable codes. In Dongdai Lin and Kazue Sako, editors, *PKC 2019, Part I*, volume 11442 of *LNCS*, pages 519–548. Springer, Heidelberg, April 2019.

[DKO+18] Ivan Damgård, Tomasz Kazana, Maciej Obremski, Varun Raj, and Luisa Siniscalchi. Continuous NMC secure against permutations and overwrites, with applications to CCA secure commitments. In Amos Beimel and Stefan Dziembowski, editors, *TCC 2018, Part II*, volume 11240 of *LNCS*, pages 225–254. Springer, Heidelberg, November 2018.

[DKP21] Dana Dachman-Soled, Ilan Komargodski, and Rafael Pass. Non-malleable codes for bounded parallel-time tampering. In Tal Malkin and Chris Peikert, editors, *CRYPTO 2021, Part III*, volume 12827 of *LNCS*, pages 535–565, Virtual Event, August 2021. Springer, Heidelberg.

[DKS17] Dana Dachman-Soled, Mukul Kulkarni, and Aria Shahverdi. Tight upper and lower bounds for leakage-resilient, locally decodable and updatable non-malleable codes. In Serge Fehr, editor, *PKC 2017, Part I*, volume 10174 of *LNCS*, pages 310–332. Springer, Heidelberg, March 2017.

[DLSZ15] Dana Dachman-Soled, Feng-Hao Liu, Elaine Shi, and Hong-Sheng Zhou. Locally decodable and updatable non-malleable codes and their applications. In Yevgeniy Dodis and Jesper Buus Nielsen, editors, *TCC 2015, Part I*, volume 9014 of *LNCS*, pages 427–450. Springer, Heidelberg, March 2015.

[DPW10] Stefan Dziembowski, Krzysztof Pietrzak, and Daniel Wichs. Non-malleable codes. In Andrew Chi-Chih Yao, editor, *ICS 2010*, pages 434–452. Tsinghua University Press, January 2010.

[EFKP20] Naomi Ephraim, Cody Freitag, Ilan Komargodski, and Rafael Pass. Non-malleable time-lock puzzles and applications. Cryptology ePrint Archive, Report 2020/779, 2020. https://eprint.iacr.org/2020/779.

[FHMV17] Sebastian Faust, Kristina Hostáková, Pratyay Mukherjee, and Daniele Venturi. Non-malleable codes for space-bounded tampering. In Jonathan Katz and Hovav Shacham, editors, *CRYPTO 2017, Part II*, volume 10402 of *LNCS*, pages 95–126. Springer, Heidelberg, August 2017.

[FMNV14] Sebastian Faust, Pratyay Mukherjee, Jesper Buus Nielsen, and Daniele Venturi. Continuous non-malleable codes. In Yehuda Lindell, editor, *TCC 2014*, volume 8349 of *LNCS*, pages 465–488. Springer, Heidelberg, February 2014.

[FMVW14] Sebastian Faust, Pratyay Mukherjee, Daniele Venturi, and Daniel Wichs. Efficient non-malleable codes and key-derivation for poly-size tampering circuits. In Phong Q. Nguyen and Elisabeth Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 111–128. Springer, Heidelberg, May 2014.

[GGI+15] Craig Gentry, Jens Groth, Yuval Ishai, Chris Peikert, Amit Sahai, and Adam D. Smith. Using fully homomorphic hybrid encryption to minimize non-interative zero-knowledge proofs. *Journal of Cryptology*, 28(4):820–843, October 2015.

[GLM+04] Rosario Gennaro, Anna Lysyanskaya, Tal Malkin, Silvio Micali, and Tal Rabin. Algorithmic tamper-proof (ATP) security: Theoretical foundations for security

against hardware tampering. In Moni Naor, editor, *TCC 2004*, volume 2951 of *LNCS*, pages 258–277. Springer, Heidelberg, February 2004.

[GMW19] Divya Gupta, Hemanta K. Maji, and Mingyuan Wang. Explicit rate-1 non-malleable codes for local tampering. In Alexandra Boldyreva and Daniele Micciancio, editors, *CRYPTO 2019, Part I*, volume 11692 of *LNCS*, pages 435–466. Springer, Heidelberg, August 2019.

[GW11] Craig Gentry and Daniel Wichs. Separating succinct non-interactive arguments from all falsifiable assumptions. In Lance Fortnow and Salil P. Vadhan, editors, *43rd ACM STOC*, pages 99–108. ACM Press, June 2011.

[JW15] Zahra Jafargholi and Daniel Wichs. Tamper detection and continuous non-malleable codes. In Yevgeniy Dodis and Jesper Buus Nielsen, editors, *TCC 2015, Part I*, volume 9014 of *LNCS*, pages 451–480. Springer, Heidelberg, March 2015.

[Knu93] Lars R. Knudsen. Cryptanalysis of LOKI91. In Jennifer Seberry and Yuliang Zheng, editors, *AUSCRYPT'92*, volume 718 of *LNCS*, pages 196–208. Springer, Heidelberg, December 1993.

[KOS17] Bhavana Kanukurthi, Sai Lakshmi Bhavana Obbattu, and Sruthi Sekar. Four-state non-malleable codes with explicit constant rate. In Yael Kalai and Leonid Reyzin, editors, *TCC 2017, Part II*, volume 10678 of *LNCS*, pages 344–375. Springer, Heidelberg, November 2017.

[KOS18] Bhavana Kanukurthi, Sai Lakshmi Bhavana Obbattu, and Sruthi Sekar. Non-malleable randomness encoders and their applications. In Jesper Buus Nielsen and Vincent Rijmen, editors, *EUROCRYPT 2018, Part III*, volume 10822 of *LNCS*, pages 589–617. Springer, Heidelberg, April / May 2018.

[Li17] Xin Li. Improved non-malleable extractors, non-malleable codes and independent source extractors. In Hamed Hatami, Pierre McKenzie, and Valerie King, editors, *49th ACM STOC*, pages 1144–1156. ACM Press, June 2017.

[Li18] Xin Li. Non-malleable extractors and non-malleable codes: Partially optimal constructions. Cryptology ePrint Archive, Report 2018/353, 2018. https://eprint.iacr.org/2018/353.

[LL12] Feng-Hao Liu and Anna Lysyanskaya. Tamper and leakage resilience in the split-state model. In Reihaneh Safavi-Naini and Ran Canetti, editors, *CRYPTO 2012*, volume 7417 of *LNCS*, pages 517–532. Springer, Heidelberg, August 2012.

[OPVV18] Rafail Ostrovsky, Giuseppe Persiano, Daniele Venturi, and Ivan Visconti. Continuously non-malleable codes in the split-state model from minimal assumptions. In Hovav Shacham and Alexandra Boldyreva, editors, *CRYPTO 2018, Part III*, volume 10993 of *LNCS*, pages 608–639. Springer, Heidelberg, August 2018.

[QLY+15] Baodong Qin, Shengli Liu, Tsz Hon Yuen, Robert H. Deng, and Kefei Chen. Continuous non-malleable key derivation and its application to related-key security. In Jonathan Katz, editor, *PKC 2015*, volume 9020 of *LNCS*, pages 557–578. Springer, Heidelberg, March / April 2015.

[Wee12] Hoeteck Wee. Public key encryption against related key attacks. In Marc Fischlin, Johannes Buchmann, and Mark Manulis, editors, *PKC 2012*, volume 7293 of *LNCS*, pages 262–279. Springer, Heidelberg, May 2012.

A Omitted Proofs

A.1 Proof of Theorem 1

For simplicity, we prove the theorem for the case of continuous persistent super non-malleability. The proof for the one-time case is analogous. Let A be an adversary in the experiment $\mathbf{CNM}^{\mathsf{cdw},\mathsf{super}}_{\Gamma^*,\mathsf{A},\mathcal{F},\mathcal{G}}(\lambda,b) := \mathbf{G}^*(\lambda,b)$ defining leakage-resilient continuous persistent super non-malleability w.r.t. codeword of Γ^*.

Consider the event $\mathbf{W}$ that becomes true whenever, for some $q \in poly(\lambda)$, the first $q-1$ tampering queries do not modify the codeword and the q-th tampering query output by A is such that $f(\gamma^*) = (\tilde{\gamma},\tilde{\sigma})$ satisfies the following: (i) $\mathsf{Dec}(\gamma) = \tilde{vk}||\tilde{\mu}$ with $\tilde{vk} = vk$; (ii) $\mathsf{SigVer}(\tilde{vk},\tilde{\gamma},\tilde{\sigma}) = 1$ and $(\tilde{\gamma},\tilde{\sigma}) \neq (\gamma,\sigma)$. We define, for $b \in \{0,1\}$, the hybrid experiment $\mathbf{H}^*(\lambda,b)$ which is the same of $\mathbf{G}^*(\lambda,b)$ except that, when $\mathbf{W}$ happens, the challenger returns $\diamond$. Clearly, the two experiments are only distinguishable when $\mathbf{W}$ happens, therefore, if we show that $\mathbf{W}$ happens with negligible probability, it follows that $\mathbf{G}^*(\lambda,b)$ and $\mathbf{H}^*(\lambda,b)$ are computationally close.

Lemma 7. *For all PPT adversaries* A *there is a negligible function* $\nu : \mathbb{N} \to [0,1]$ *such that* $\mathbb{P}[\mathbf{W}] \leq \nu(\lambda)$.

Proof. By contradiction, assume that there exists a PPT adversary A and a polynomial $p(\lambda) \in poly(\lambda)$ such that A provokes event $\mathbf{W}$ with probability at least $1/p(\lambda)$. We construct a PPT attacker $\hat{\mathsf{A}}$ attacking one-time strong unforgeability of Σ as follows:

1. Upon receiving vk from the challenger, run $\mathsf{A}(1^\lambda)$ and fix any $b \in \{0,1\}$.
2. Upon receiving $\mu_0,\mu_1 \in \{0,1\}^k$ from A, compute $\gamma \leftarrow\!\!\$\ \mathsf{Enc}(vk||\mu_b)$ and forward γ to the challenger obtaining back a signature $\sigma \in \{0,1\}^s$.
3. Upon receiving a leakage query $g \in \mathcal{G}$ from A, return $g(\gamma,\sigma)$.
4. Upon receiving a tampering query f such that $f((\gamma,\sigma)) = (\gamma,\sigma)$, return $\diamond$. Else, let $(\tilde{\gamma},\tilde{\sigma}) = f(\gamma,\sigma)$ and forward $(\tilde{\gamma},\tilde{\sigma})$ to the challenger. If the latter never happens, abort.

For the analysis, note that the codeword (γ,σ) is distributed exactly as in the experiment $\mathbf{G}(\lambda,b)$ and thus the simulation of leakage queries is perfect. Moreover, whenever A provokes event $\mathbf{W}$ it holds that $(\tilde{\gamma},\tilde{\sigma}) \neq (\gamma,\sigma)$ and $\mathsf{SigVer}(vk,\tilde{\gamma},\tilde{\sigma}) = 1$ so that $\hat{\mathsf{A}}$ breaks one-time strong unforgeability with probability at least $\mathbb{P}[\mathbf{W}] \geq 1/p(\lambda)$. This concludes the proof of the lemma. □

Now that we ruled out the event in which the adversary modifies the codeword but not the message, we can conclude the proof via a reduction $\hat{\mathsf{A}}$ to the leakage-resilient continuous persistent non-malleability w.r.t. message of Γ. Towards this, suppose that there exists an adversary A which is able to tell apart $\mathbf{H}^*(\lambda,0)$ and $\mathbf{H}^*(\lambda,1)$ with non-negligible advantage. Consider the following description of $\hat{\mathsf{A}}$:

1. Set $\delta = \iota^* = 0$, sample a randomness for the signing algorithm $\rho \leftarrow\!\!\$\ \{0,1\}^*$ and a couple of keys $(sk,vk) \leftarrow\!\!\$\ \mathsf{Gen}(1^\lambda)$.
2. Run $\mathsf{A}(1^\lambda)$ and, upon receiving $\mu_0,\mu_1 \in \{0,1\}^k$ from A, forward $(vk||\mu_0, vk||\mu_1)$ to the challenger.
3. Upon receiving a leakage query $g \in \mathcal{G}$ from A, output $\perp$ if $\delta = 1$. Else, define the leakage function $\hat{g}(\gamma)$ which outputs the same as $g(\gamma,\mathsf{Sign}(sk,\gamma;\rho))$. Forward $\hat{g}$ to the challenger and send the obtained leakage back to A.

4. Upon receiving a tampering function f, output $\perp$ if $\delta = 1$. Else, do the following:
   (a) If $\iota^* = 1$, compute $(\tilde{\gamma}, \tilde{\sigma}) = f((\hat{\gamma}, \hat{\sigma}))$. Update $(\hat{\gamma}, \hat{\sigma}) = (\tilde{\gamma}, \tilde{\sigma})$ and output it.
   (b) Else, submit to the oracle the tampering function $\hat{f}(\gamma)$ which computes the complete mauled codeword $(\tilde{\gamma}, \tilde{\sigma}) = f((\gamma, \mathsf{Sign}(sk, \gamma; \rho)))$ and outputs $\tilde{\gamma}$ whenever $(\tilde{\gamma}, \tilde{\sigma})$ correctly verifies, or returns an invalid codeword otherwise. Process the answer $\tilde{\gamma}$ as follows. If $\tilde{\gamma} = \perp$, set $\delta = 1$ and return $\perp$. If $\tilde{\gamma} = \diamond$, return $\diamond$. Else, make a leakage query g^* to get $\tilde{\sigma}$, set $\iota^* = 1$, update $(\hat{\gamma}, \hat{\sigma}) = (\tilde{\gamma}, \tilde{\sigma})$ and return it.
5. Return the same bit as A.

For the analysis, note that by assumption the leakage functions $\hat{g}, g^*$ defined above satisfy $\hat{g}, g^* \in \mathcal{G}(n)$, leak at most $\ell + s$ bits and the simulation of A's leakage queries is perfect. As for A's tampering queries, we claim that the simulation is perfect conditioning on $\mathbf{W}$ not happening. Indeed, letting $f(\gamma, \sigma) = (\tilde{\gamma}, \tilde{\sigma})$, the reduction:

- Returns $\perp$ whenever either $\tilde{\gamma}$ is invalid or $\tilde{\sigma}$ does not verify correctly.
- Returns $(\tilde{\gamma}, \tilde{\sigma}) = f(\hat{\gamma}, \hat{\sigma})$ whenever $\tilde{\gamma} \neq \gamma$ is valid and $\tilde{\sigma}$ verifies correctly, with $(\hat{\gamma}, \hat{\sigma})$ being equal to (γ, σ) as long as $\delta = 0$, and the last mauled codeword otherwise.
- Returns $\diamond$ whenever $\mathsf{Dec}(\tilde{\gamma}) \in \{vk||\mu_0, vk||\mu_1\}$. However, conditioning on $\overline{\mathbf{W}}$, the latter happens if and only if $(\tilde{\gamma}, \tilde{\sigma}) = (\gamma, \sigma)$.

Since, by Lemma 7, $\mathbf{W}$ happens with negligible probability, this concludes the proof.

A.2 Proof of Theorem 3

Let $\mathbf{G}(\lambda, b) = \mathbf{CNM}^{\mathsf{cdw},\mathsf{super}}_{\Gamma,\mathsf{A},\mathcal{F},\mathcal{G}}(\lambda, b)$ be the experiment defining continuous persistent super non-malleability of Γ, and let $\hat{\mathbf{G}}(\lambda, b) = \mathbf{1CNM}^{\mathsf{cdw},\mathsf{super}}_{\Gamma^*,\mathsf{A},\mathcal{F},\mathcal{G}}(\lambda, b)$ be the experiment defining one-time leakage-resilient super non-malleability of Γ. Let A be an adversary for $\mathbf{G}(\lambda, b)$. For every $q \in poly(\lambda)$, call $\mathbf{W}^{(q)}(b)$ the event where, in $\mathbf{G}(\lambda, b)$, q is the index of the first tampering query whose output is not $\diamond$. Since

$$\begin{aligned}&|\mathbb{P}\left[\mathbf{G}(\lambda, 0) = 1\right] - \mathbb{P}\left[\mathbf{G}(\lambda, 1) = 1\right]| \\ &\leq \textstyle\sum_{q \in poly(\lambda)} \left|\mathbb{P}\left[\mathbf{G}(\lambda, 0) = 1 \wedge \mathbf{W}^{(q)}(0)\right] - \mathbb{P}\left[\mathbf{G}(\lambda, 1) = 1 \wedge \mathbf{W}^{(q)}(1)\right]\right|,\end{aligned}$$

then the theorem holds if we show a uniform bound such that, for every $q \in poly(\lambda)$,

$$\left|\mathbb{P}\left[\mathbf{G}(\lambda, 0) = 1 \wedge \mathbf{W}^{(q)}(0)\right] - \mathbb{P}\left[\mathbf{G}(\lambda, 1) = 1 \wedge \mathbf{W}^{(q)}(1)\right]\right| \in negl(\lambda).$$

By contradiction, assume that there exists $q \in poly(\lambda)$ and a polynomial $p(\lambda) \in poly(\lambda)$ such that,

$$\left|\mathbb{P}\left[\mathbf{G}(\lambda, 0) = 1 \wedge \mathbf{W}^{(q)}(0)\right] - \mathbb{P}\left[\mathbf{G}(\lambda, 1) = 1 \wedge \mathbf{W}^{(q)}(1)\right]\right| \geq \frac{1}{p(\lambda)}.$$

Consider the following adversary $\hat{\mathsf{A}}$ attacking super non-malleability of Γ.

1. Run $\mathsf{A}(1^\lambda)$.
2. Upon receiving (μ_0, μ_1) from A, forward (μ_0, μ_1) to the challenger.
3. Upon receiving the leakage query $g \in \mathcal{G}$ from A, submit it to the challenger and return the output to A.

4. Upon receiving the j-th tampering function $f^{(j)}$ with $j < q$, answer $\diamond$.

5. Upon receiving $f^{(q)}$ from A, submit the leakage function $g^{f^{(1)},\ldots,f^{(q)}}$ to the challenger, and abort if the answer is 0. Otherwise, submit the tampering query $f^{(q)}$ to the challenger, save the answer $\tilde{\gamma}$ as $\hat{\gamma}$ and return it to A.

6. Upon receiving the j-th tampering function with $j > q$, compute $\tilde{\gamma} = f^{(j)}(\hat{\gamma})$ and output $\perp$ if it is invalid, $\diamond$ if it equals $\hat{\gamma}$. Otherwise, update $\hat{\gamma}$ with $\tilde{\gamma}$ and return it.

7. Output whatever A does.

Since $\hat{\mathsf{A}}$ aborts whenever $\mathbf{W}^{(q)}(b)$ doesn't happen and makes a perfect simulation otherwise, it holds

$$\begin{aligned}
&\left|\mathbb{P}\left[\hat{\mathbf{G}}(\lambda,0)=1\right]-\mathbb{P}\left[\hat{\mathbf{G}}(\lambda,1)=1\right]\right| \\
&=\left|\mathbb{P}\left[\mathbf{G}(\lambda,0)=1\wedge\mathbf{W}^{(q)}(0)\right]-\mathbb{P}\left[\mathbf{G}(\lambda,1)=1\wedge\mathbf{W}^{(q)}(1)\right]\right| \\
&\geq\frac{1}{p(\lambda)}.
\end{aligned}$$

This concludes the proof.

A.3 Proof of Theorem 5

For simplicity, we prove the theorem only for the case of super non-malleability w.r.t. codeword. The proof for the other flavors of non-malleability is analogous. Let $\mathbf{G}(\lambda,b) := \mathbf{1NM}^{\mathsf{cdw},\mathsf{super}}_{\Gamma^*,\mathsf{A},\mathcal{F},\mathcal{G}}(\lambda,b)$ be the original experiment defining leakage-resilient super non-malleability of Γ^*. We consider a sequence of hybrids experiments, as described below.

$\mathbf{H}_1(\lambda,b)$: This hybrid is identical to $\mathbf{G}(\lambda,b)$, except that we replace the proof π with a simulated one. Namely, at the beginning of the experiment, the challenger now runs $(\omega,\zeta,\xi)\leftarrow\!\!\$\ \mathsf{S}_0(1^\lambda)$. Moreover, the target codeword is of the form $\gamma^*=(\gamma,\pi)$ where $\pi\leftarrow\!\!\$\ \mathsf{S}_1(\zeta,\gamma)$.

$\mathbf{H}_2(\lambda,b)$: This hybrid is identical to the previous one, except that we replace ρ with a uniformly random string when computing the target encoding.

$\mathbf{H}_3(\lambda,b)$: We change the way the tampering query $f\in\mathcal{F}(n+m)$ is answered. Namely, let $\tilde{\gamma}^*=(\tilde{\gamma},\tilde{\pi})=f(\gamma,\pi)$. Then:

- If $\tilde{\gamma}^*=\gamma^*$, return $\diamond$.
- Else, if either $\mathsf{ProofVer}(\omega,\tilde{\gamma},\tilde{\pi})=0$ or $\mathsf{Dec}(\tilde{\gamma})=\tilde{\mu}=\perp$ return $\perp$.
- Else return $(\tilde{\gamma},\tilde{\pi})$, where $\tilde{\gamma}=\mathsf{Enc}(\tilde{\mu};\mathsf{G}(\tilde{\sigma}))$ for $\tilde{\sigma}\leftarrow\!\!\$\ \mathsf{K}(\xi,\tilde{\gamma},\tilde{\pi})$.

Lemma 8. *For every $b\in\{0,1\}$, it holds that $\{\mathbf{G}(\lambda,b)\}_{\lambda\in\mathbb{N}}\overset{\mathrm{c}}{\approx}\{\mathbf{H}_1(\lambda,b)\}_{\lambda\in\mathbb{N}}$.*

Proof. We reduce to the zero knowledge property of Π. Fix any $b\in\{0,1\}$. By contradiction, assume that there exists a PPT adversary A and a polynomial $p(\lambda)\in poly(\lambda)$ such that

$$|\mathbb{P}\left[\mathbf{G}(\lambda,b)=1\right]-\mathbb{P}\left[\mathbf{H}_1(\lambda,b)=1\right]|\geq 1/p(\lambda).$$

Consider the following PPT attacker $\hat{\mathsf{A}}$ against Π:

1. Upon receiving CRS ω from the challenger, forward it to A.

2. Upon receiving (μ_0, μ_1) from A, sample $\sigma \leftarrow\!\!\$\ \{0,1\}^s$, compute $\rho = \mathsf{G}(\sigma)$ and $\gamma = \mathsf{Enc}(\mu_b; \rho)$, and forward (γ, σ) to the challenger, obtaining a proof π (either real or simulated). Let $\gamma^* = (\gamma, \pi)$.

3. Upon receiving a leakage query $g \in \mathcal{G}$ from A, answer it as in the original experiment. Namely, return $g(\gamma, \pi)$.

4. Upon receiving the tampering query $f \in \mathcal{F}$, answer it as in the original experiment. Namely, let $\tilde{\gamma}^* = (\tilde{\gamma}, \tilde{\pi}) = f(\gamma, \pi)$. Then:

   - If $\tilde{\gamma}^* = \gamma^*$, return $\diamond$.
   - Else, if either $\mathsf{ProofVer}(\omega, \tilde{\gamma}, \tilde{\pi}) = 0$ or $\mathsf{Dec}(\tilde{\gamma}) = \perp$ return $\perp$.
   - Else return $\tilde{\gamma}^*$.

5. Output whatever A outputs.

By inspection, the reduction $\hat{\mathsf{A}}$ perfectly simulates the view of the adversary. In particular, assuming the CRS is generated using $\omega \leftarrow\!\!\$\ \mathsf{CRSGen}(1^\lambda)$ and the proof π is computed by running $\pi \leftarrow\!\!\$\ \mathsf{Prove}(\omega, \gamma, \sigma)$ (resp. $(\omega, \zeta, \xi) \leftarrow\!\!\$\ \mathsf{S}_0(1^\lambda)$ and $\pi \leftarrow\!\!\$\ \mathsf{S}_1(\zeta, \gamma)$), the simulated codeword γ^* is distributed identically to the target codeword in $\mathbf{G}(\lambda, b)$ (resp. $\mathbf{H}_1(\lambda, b)$). Hence, $\hat{\mathsf{A}}$ retains the same distinguishing advantage as that of A. The lemma follows. □

Lemma 9. *For every $b \in \{0,1\}$, it holds that $\{\mathbf{H}_1(\lambda, b)\}_{\lambda\in\mathbb{N}} \overset{\text{c}}{\approx} \{\mathbf{H}_2(\lambda, b)\}_{\lambda\in\mathbb{N}}$.*

Proof. We reduce to the security of G. Fix any $b \in \{0,1\}$. By contradiction, assume that there exists a PPT adversary A and a polynomial $p(\lambda) \in \mathit{poly}(\lambda)$ such that

$$|\mathbb{P}[\mathbf{H}_1(\lambda, b) = 1] - \mathbb{P}[\mathbf{H}_2(\lambda, b) = 1]| \geq 1/p(\lambda).$$

Consider the following PPT attacker $\hat{\mathsf{A}}$ against G:

1. Upon receiving $\rho \in \{0,1\}^r$ from the challenger, sample $(\omega, \zeta, \xi) \leftarrow\!\!\$\ \mathsf{S}_0(1^\lambda)$ and forward ω to A.

2. Upon receiving (μ_0, μ_1) from A, compute $\gamma = \mathsf{Enc}(\mu_b; \rho)$ and $\pi \leftarrow\!\!\$\ \mathsf{S}_1(\zeta, \gamma)$, and let $\gamma^* = (\gamma, \pi)$.

3. Upon receiving a leakage query $g \in \mathcal{G}$ from A answer it as in the original experiment. Namely, return $g(\gamma, \pi)$.

4. Upon receiving the tampering query $f \in \mathcal{F}$, answer it as in the original experiment. Namely, let $\tilde{\gamma}^* = (\tilde{\gamma}, \tilde{\pi}) = f(\gamma, \pi)$. Then:

   - If $\tilde{\gamma}^* = \gamma^*$, return $\diamond$.
   - Else, if either $\mathsf{ProofVer}(\omega, \tilde{\gamma}, \tilde{\pi}) = 0$ or $\mathsf{Dec}(\tilde{\gamma}) = \perp$ return $\perp$.
   - Else return $\tilde{\gamma}^*$.

5. Output whatever A outputs.

By inspection, the reduction $\hat{\mathsf{A}}$ perfectly simulates the view of the adversary. In particular, assuming the string ρ is generated via $\mathsf{G}(\sigma)$ for random $\sigma \leftarrow\!\!\$\ \{0,1\}^s$ (resp. $\rho \leftarrow\!\!\$\ \{0,1\}^r$), the simulated codeword γ^* is distributed identically to the target codeword in $\mathbf{H}_1(\lambda, b)$ (resp. $\mathbf{H}_2(\lambda, b)$). Hence, $\hat{\mathsf{A}}$ retains the same distinguishing advantage as that of A. The lemma follows. □

Lemma 10. *For every $b \in \{0,1\}$, it holds that* $\{\mathbf{H}_2(\lambda,b)\}_{\lambda\in\mathbb{N}} \overset{c}{\approx} \{\mathbf{H}_3(\lambda,b)\}_{\lambda\in\mathbb{N}}$.

Proof. We reduce to the one-time simulation extractability property of Π. Fix any $b \in \{0,1\}$. Define the event $\mathbf{W}$, over the probability space of $\mathbf{H}_3(\lambda,b)$, which becomes true whenever: (i) $(\tilde{\gamma},\tilde{\pi}) \neq (\gamma,\pi)$; (ii) $\mathsf{ProofVer}(\omega,\tilde{\gamma},\tilde{\pi}) = 1$; (iii) $\mathsf{Enc}(\tilde{\mu};\mathsf{G}(\tilde{\sigma})) \neq \tilde{\gamma}$. Clearly, $\mathbf{H}_2(\lambda,b)$ and $\mathbf{H}_3(\lambda,b)$ are identical conditioning on $\overline{\mathbf{W}}$, and thus $|\mathbb{P}[\mathbf{H}_2(\lambda,b)=1] - \mathbb{P}[\mathbf{H}_3(\lambda,b)=1]| \leq \mathbb{P}[\mathbf{W}]$.

Next, we show that $\mathbf{W}$ only happens with negligible probability. By contradiction, assume that there exists a PPT adversary A and a polynomial $p(\lambda) \in \mathit{poly}(\lambda)$ such that A provokes $\mathbf{W}$ with probability at least $1/p(\lambda)$. Consider the following PPT attacker $\hat{\mathsf{A}}$ against Π:

1. Upon receiving the CRS ω from the challenger, forward ω to A.
2. Upon receiving (μ_0,μ_1) from A, compute $\gamma = \mathsf{Enc}(\mu_b;\rho)$ for $\rho \leftarrow\!\!\$\ \{0,1\}^r$ and forward γ to the challenger obtaining a proof π. (Note that γ may be a false statement, as the randomness ρ may be outside the range of the PRG.) Let $\gamma^* = (\gamma,\pi)$.
3. Upon receiving a leakage query $g \in \mathcal{G}$ from A answer it as in the original experiment. Namely, return $g(\gamma,\pi)$.
4. Upon receiving the tampering query $f \in \mathcal{F}$, output $(\tilde{\gamma},\tilde{\pi}) = f(\gamma,\pi)$.

By inspection, the reduction $\hat{\mathsf{A}}$ perfectly simulates the view of the adversary. In particular, the simulated codeword γ^* is distributed identically to the target codeword in experiment $\mathbf{H}_3(\lambda,b)$. Hence, A will provoke event $\mathbf{W}$ with probability at least $1/p(\lambda)$ which in turn means that $\hat{\mathsf{A}}$ breaks one-time simulation extractability of Π with the same probability. The lemma follows. □

Lemma 11. $\{\mathbf{H}_3(\lambda,0)\}_{\lambda\in\mathbb{N}} \overset{c}{\approx} \{\mathbf{H}_3(\lambda,1)\}_{\lambda\in\mathbb{N}}$

Proof. We reduce to leakage-resilient non-malleability of Γ. By contradiction, assume that there exists a PPT adversary A and a polynomial $p(\lambda) \in \mathit{poly}(\lambda)$ such that

$$|\mathbb{P}[\mathbf{H}_3(\lambda,0)=1] - \mathbb{P}[\mathbf{H}_3(\lambda,1)=1]| \geq 1/p(\lambda).$$

Consider the following PPT attacker $\hat{\mathsf{A}}$ against Γ:

1. Run $(\omega,\zeta,\xi) \leftarrow\!\!\$\ \mathsf{S}_0(1^\lambda)$ and forward ω to A. Sample $\rho_{\mathsf{S}} \leftarrow\!\!\$\ \{0,1\}^*$.
2. Upon receiving (μ_0,μ_1) from A, forward (μ_0,μ_1) to the challenger.
3. Upon receiving a leakage query $g \in \mathcal{G}(n+m)$ from A, define the leakage function $\hat{g}$ that upon input $\gamma \in \{0,1\}^n$:
   - computes $\pi = \mathsf{S}_1(\zeta,\gamma;\rho_{\mathsf{S}})$;
   - returns $g(\gamma,\pi)$.

   Query $\hat{g}$ to the target leakage oracle and forward the answer to A.
4. Upon receiving the tampering query $f \in \mathcal{F}(n+m)$, define the leakage function $\hat{g}'$ that upon input $\gamma \in \{0,1\}^n$:
   - computes $\pi = \mathsf{S}_1(\zeta,\gamma;\rho_{\mathsf{S}})$;
   - lets $(\tilde{\gamma},\tilde{\pi}) = f(\gamma,\pi)$;
   - runs $\mathsf{ProofVer}(\omega,\tilde{\gamma},\tilde{\pi})$ and outputs $\perp$ if the verification fails;

- else, runs $\tilde{\sigma} \leftarrow\!\!\$\ \mathsf{K}(\xi, \tilde{\gamma}, \tilde{\pi})$ and outputs $(\tilde{\sigma}, \tilde{\pi})$.

Query $\hat{g}'$ to the target leakage oracle. If the output is $\perp$, return $\perp$ to A. Else, forward to the challenger the tampering function $\hat{f}(\cdot)$ that upon input $\gamma \in \{0,1\}^n$:

- computes $\pi = \mathsf{S}_1(\zeta, \gamma; \rho_{\mathsf{S}})$;
- outputs $\tilde{\gamma} = f(\gamma, \pi)_1$.

Upon receiving $\tilde{\mu} \in \{0,1\}^k \cup \{\diamond, \perp\}$, if $\tilde{\mu} \in \{0,1\}^k$ return $(\tilde{\gamma}, \tilde{\pi})$ to A where $\tilde{\gamma} = \mathsf{Enc}(\tilde{\mu}; \mathsf{G}(\tilde{\sigma}))$. Else, return $\tilde{\mu}$.

5. Output whatever A does.

By inspection, the reduction $\hat{\mathsf{A}}$ perfectly simulates the view of the adversary. In particular, assuming the target codeword γ is an encoding of μ_b, the answer to both leakage and tampering queries is identically distributed to that in $\mathbf{H}_3(\lambda, b)$. Furthermore, by Eq. (18) and Eq. (19), it holds that $\hat{g}, \hat{g}' \in \mathcal{G}(n)$ and $\hat{f} \in \mathcal{F}(n)$. Hence, $\hat{\mathsf{A}}$ retains the same distinguishing advantage as that of A. This concludes the proof of the lemma. □

Theorem 5 now follows by combining the above lemmas.

B A Light Version of [DKP21] in the CRS Model

In this section, we show a simplified version of the non-malleable code of [DKP21] in the CRS model. The resulting scheme achieves non-malleability w.r.t. codeword and leakage resilience for the class of bounded polynomial-depth functions, and can therefore be used to instantiate Theorem 5.

Let $\Pi' = (\mathsf{PGen}, \mathsf{PSol})$ be a time-lock puzzle, and $\Pi = (\mathsf{CRSGen}, \mathsf{Prove}, \mathsf{ProofVer})$ be a non-interactive argument system for the relation:

$$\mathcal{R} = \left\{ (\zeta, \mu) : \exists T \text{ s.t. } \zeta \leftarrow\!\!\$\ \mathsf{PGen}(1^\lambda, \mu, T) \right\}. \tag{20}$$

The description of the scheme $\Gamma = (\mathsf{Init}, \mathsf{Enc}, \mathsf{Dec})$ follows:

Initialization: The initialization algorithm Init outputs $\omega \leftarrow\!\!\$\ \mathsf{CRSGen}(1^\lambda)$.

Encoding: The encoding algorithm Enc proceeds as follows:

- get $\zeta \leftarrow\!\!\$\ \mathsf{PGen}(1^\lambda, \mu, T)$;
- compute $\pi \leftarrow\!\!\$\ \mathsf{Prove}(\omega, \zeta, \mu)$;
- return $\gamma = (\zeta, \pi)$.

Decoding: The decoding algorithm Dec, upon input a codeword $\gamma = (\zeta, \pi)$, proceeds as follows:

- run $\mathsf{ProofVer}(\omega, \zeta, \pi)$, and output $\perp$ if the verification fails;
- else, return $\mu = \mathsf{PSol}(\zeta)$.

Theorem 6. *Let $\ell, T_1, T_2 \in \mathit{poly}(\lambda)$. Assume that Π is a one-time simulation extractable non-interactive zero-knowledge argument system for the relation of Eq. (20), with CRS length in $\mathit{poly}(\lambda)$, and that Π' is a (S, ϵ)-hard time-lock puzzle for some $S \in n \cdot 2^{\mathit{poly}(\lambda)}$. Then, the (k, n)-code described above is $(\mathcal{F}^{T_1}_{\mathsf{non\text{-}uni}}(n), \ell)$-leakage-resilient $(\mathcal{F}^{T_2}_{\mathsf{non\text{-}uni}}(n))$-non-malleable w.r.t. codeword.*

Remark 7 (Instantiating the scheme)**.** *For this construction, we can use as building blocks the NIZK proof system of [GGI+15] and the time-lock puzzle of [DKP21] (respectively described in Section 4.2 and Section 2.5). The first one yields a valid candidate because it achieves simulation extractability and has CRS length in poly(λ). For the time-lock puzzle, note that using these building blocks we get $n = z + poly(\lambda)$, where z is the length of the time-lock puzzle ζ. This implies that our choice for S satisfies Eq. (1).*

of Theorem 6. Let $\mathbf{G}(\lambda, b) := \mathbf{1NM}^{\mathsf{cdw},\mathsf{standard}}_{\Gamma,\mathsf{A},\mathcal{F},\mathcal{G}}(\lambda, b)$. To prove the theorem, we consider the following hybrid experiments.

$\mathbf{H}_1(\lambda, b)$**:** This hybrid is identical to $\mathbf{G}(\lambda, b)$, except that we replace the proof π with a simulated one. Namely, at the beginning of the experiment, the challenger now runs $(\omega, \tau, \xi) \leftarrow\!\!\$\ \mathsf{S}_0(1^\lambda)$. Moreover, the target codeword is of the form $\gamma^* = (\zeta, \pi)$ where $\pi \leftarrow\!\!\$\ \mathsf{S}_1(\tau, \zeta)$.

$\mathbf{H}_2(\lambda, b)$**:** This hybrid is defined as $\mathbf{H}_1(\lambda, b)$, but it answers tampering queries by extracting the message from the proof instead of solving the puzzle. More in detail, given $(\tilde{\zeta}, \tilde{\pi}) = f(\zeta, \pi)$, the answer to the tampering query is computed as follows:

- If $(\tilde{\zeta}, \tilde{\pi}) = (\zeta, \pi)$, return $\diamond$.
- If $\mathsf{ProofVer}(\omega, \tilde{\zeta}, \tilde{\pi}) = 0$, return $\perp$.
- Else, return $\tilde{\mu} \leftarrow\!\!\$\ \mathsf{K}(\xi, \tilde{\zeta}, \tilde{\pi})$.

Lemma 12. *For every $b \in \{0, 1\}$, it holds that $\{\mathbf{G}(\lambda, b)\}_{\lambda\in\mathbb{N}} \overset{c}{\approx} \{\mathbf{H}_1(\lambda, b)\}_{\lambda\in\mathbb{N}}$.*

Proof. We reduce to the zero knowledge property of Π. Fix any $b \in \{0, 1\}$. By contradiction, assume that there exists a PPT adversary A and a polynomial $p(\lambda) \in poly(\lambda)$ such that

$$|\mathbb{P}[\mathbf{G}(\lambda, b) = 1] - \mathbb{P}[\mathbf{H}_1(\lambda, b) = 1]| \geq 1/p(\lambda).$$

Consider the following PPT attacker $\hat{\mathsf{A}}$ against Π:

1. Upon receiving CRS ω from the challenger, forward it to A.
2. Upon receiving (μ_0, μ_1) from A, compute $\mu \leftarrow\!\!\$\ \mathsf{PGen}(\mu_b)$, and forward (ζ, μ) to the challenger, obtaining a proof π (either real or simulated). Let $\gamma = (\zeta, \pi)$.
3. Upon receiving a leakage query $g \in \mathcal{G}$ from A, answer it as in the original experiment. Namely, return $g(\gamma)$.
4. Upon receiving the tampering query $f \in \mathcal{F}$, answer it as in the original experiment. Namely, let $\tilde{\gamma} = f(\gamma)$. Then:
   - If $\tilde{\gamma} = \gamma$, return $\diamond$.
   - If $\mathsf{ProofVer}(\omega, \tilde{\zeta}, \tilde{\pi}) = 0$, return $\perp$.
   - Else, return $\mu = \mathsf{PSol}(\zeta)$.
5. Output whatever A outputs.

By inspection, the reduction $\hat{\mathsf{A}}$ perfectly simulates the view of the adversary. In particular, assuming the CRS is generated using $\omega \leftarrow\!\!\$\ \mathsf{CRSGen}(1^\lambda)$ and the proof π is computed by running $\pi \leftarrow\!\!\$\ \mathsf{Prove}(\omega, \zeta, \mu_b)$ (resp. $(\omega, \tau, \xi) \leftarrow\!\!\$\ \mathsf{S}_0(1^\lambda)$ and $\pi \leftarrow\!\!\$\ \mathsf{S}_1(\tau, \zeta)$), the simulated codeword γ is distributed identically to the target codeword in $\mathbf{G}(\lambda, b)$ (resp. $\mathbf{H}_1(\lambda, b)$). Hence, $\hat{\mathsf{A}}$ retains the same distinguishing advantage as that of A. The lemma follows. □

Lemma 13. *For every $b \in \{0,1\}$, it holds that* $\{\mathbf{H}_1(\lambda, b)\}_{\lambda\in\mathbb{N}} \overset{c}{\approx} \{\mathbf{H}_2(\lambda, b)\}_{\lambda\in\mathbb{N}}$.

Proof. Let $\hat{\mathsf{A}}$ be an attacker able to tell apart the hybrids with non-negligible probability. First, we remark that $\mathbf{H}_1$ and $\mathbf{H}_2$ compute the codeword the same way. This guarantees that the answer to the leakage query and the choice of the tampering function are identically distributed in both experiments, as well as the answer $\diamond$ to the tampering query. The latter holds because we are proving non-malleability w.r.t. codeword. This implies that $\hat{\mathsf{A}}$ provokes $(\tilde{\zeta}, \tilde{\pi}) \neq (\zeta, \pi)$ with non-negligible probability.

For the same reason, the attacker cannot distinguish by producing a couple $(\tilde{\zeta}, \tilde{\pi})$ that doesn't correctly verify, as the output would be $\perp$ in both experiments. Last, we remark that the mauled codeword must satisfy $\mathsf{K}(\xi, \tilde{\zeta}, \tilde{\pi}) \neq \mathsf{PSol}(\tilde{\zeta})$ with non-negligible probability. In other words, this means that $(\tilde{\zeta}, \mathsf{K}(\xi, \tilde{\zeta}, \tilde{\pi})) \notin \mathcal{R}$.

Thus, $\hat{\mathsf{A}}$ mauls with non-negligible probability (ζ, π) into $(\tilde{\zeta}, \tilde{\pi})$ such that:

- $(\tilde{\zeta}, \tilde{\pi}) \neq (\zeta, \pi)$;
- $\mathsf{ProofVer}(\omega, \tilde{\zeta}, \tilde{\pi}) = 1$;
- $(\tilde{\zeta}, \mathsf{K}(\xi, \tilde{\zeta}, \tilde{\pi})) \notin \mathcal{R}$.

This is a contradiction to the simulation extractability of Π. □

Lemma 14. *For every $b \in \{0,1\}$, it holds that* $\{\mathbf{H}_2(\lambda, 0)\}_{\lambda\in\mathbb{N}} \overset{c}{\approx} \{\mathbf{H}_2(\lambda, 1)\}_{\lambda\in\mathbb{N}}$.

Proof. First, note that the distinguisher A for the two hybrids goes through the following phases:

(i) It takes the crs ω and outputs μ_0, μ_1 and a query (either the first leakage query or the tampering function).

(ii) It chooses each subsequent query based on the crs ω and on the answers to the previous queries.

(iii) It computes the decision bit according to the crs, the leaked values and the mauled message.

In each of these phases, the attacker can be represented as a table where each slot describes the subsequent query, based on the collected data. In particular, each table's size is the multiplication of the length of the string describing the query with the space size of the data needed to compute it. The largest data space is the one including all the common reference strings, all the possible leaked values and tampered messages. Our assumptions guarantee that this space has size $2^{poly(\lambda)}$. On the other hand, the query description that requires more space is that of the tampering function. Since the set of tampering functions has at most 2^{2n} elements, where n is the codeword length, every tampering functions can be described via a $2n$-long string. In a nutshell, we can represent the attacker as a sequence of tables, whose total depth is a fixed polynomial in the security parameter, and the size is $n \cdot 2^{poly(\lambda)}$.
Consider the attacker $\hat{\mathsf{A}}$ for the time-lock puzzle described below:

1. Generate $(\omega, \tau, \xi) \leftarrow\!\!\$\ \mathsf{S}_0(1^\lambda)$ and share the crs ω with A.
2. Upon receivig a pair of messages μ_0, μ_1 from A, submit it to the oracle.
3. Upon receiving the puzzle ζ from the oracle, compute a simulated proof $\pi \leftarrow\!\!\$\ \mathsf{S}_1(\tau, \zeta)$ and let $\gamma = (\zeta, \pi)$.

4. Upon receiving a leakage query g, answer with $g(\gamma)$.
5. Upon receiving the tampering function f, compute $\tilde{\gamma} = f(\gamma)$ and answer
   - with $\diamond$ if $\tilde{\gamma} = \gamma$;
   - with $\perp$ if $\mathsf{ProofVer}(\omega, \tilde{\zeta}, \tilde{\pi})) = 0$;
   - with $\mathsf{K}(\xi, \tilde{\zeta}, \tilde{\pi})$ otherwise
6. Output the same as A.

This attacker perfectly simulates the hybrids, runs in fixed polyomial time and requires $n \cdot 2^{poly(\lambda)}$ size. The lemma follows. □

The theorem follows by combining the above lemmas. □

C Related-Key Attacks Security

In this section, we show that continuous persistent non-malleable codes remove the need for perfect erasures when protecting cryptographic primitives in the presence of persistent related-key attacks.

In what follows, we will consider public and possibly randomized functionalities Π taking as input a secret key $\kappa \in \mathcal{K}$ (fixed once at all) and a value $x \in \mathcal{X}$, and returning a value $y \in \mathcal{Y}$. In literature, this kind of functionality is called *stateless*. It is possible to interact with Π in two ways.

$(\mathtt{Eval}, x)$: Upon receiving an evaluation query with value $x \in \mathcal{X}$, it computes and outputs $y \leftarrow\!\!\$\ \Pi(\kappa, x)$.

$(\mathtt{Tamper}, f)$: Upon receving a tampering query with function f, updates the current key κ to $f(\kappa)$.

Our purpose is to make the functionality secure against an active adversary that tampers with the secret key in order to achieve information about it. For this reason we show how to use a coding scheme Γ to transform a functionality Π into the so called hardened functionality $\Psi_\Gamma(\Pi)$.

Definition 9 (Stateless hardened functionality)**.** *Let $k, n \in \mathbb{N}$, $\delta \in \{0,1\}$. Let $\Gamma = (\mathsf{Enc}, \mathsf{Dec})$ be a (k,n)-code. Let Π be a stateless functionality whose inputs are a secret key in $\mathcal{K} = \{0,1\}^k$ and a value in $\mathcal{X}$. We define a stateless hardened functionality $\Pi^* = \Psi_\Gamma(\Pi)$ as the algorithm that, upon input $\delta \in \{0,1\}$, $\gamma \in \{0,1\}^n$ and $x \in \mathcal{X}$, is defined as follows.*

- *If $\delta = 1$, stop with output $\perp$.*
- *Else, compute $\kappa = \mathsf{Dec}(\gamma)$. If $\kappa = \perp$, set $\delta = 1$ and output $\perp$. Else, compute and output $y \leftarrow\!\!\$\ \Pi(\kappa, x)$.*

Note that Π^* never re-encodes the key or erases the memory.

As for security, we restrict tampering to be *persistent*, meaning that tampering functions are applied to the output of the last tampering attempt. We make this formal below.

Given an algorithm S and a family $\mathcal{F}$ of tampering functions for Γ, we define security for the functionality Π^* by means of the games $\mathbf{Real}_{\Pi^*,\mathcal{F}}(\kappa, \lambda)$ and $\mathbf{Ideal}_{\Pi^*,\mathcal{F},\mathsf{S}}(\kappa, \lambda)$:

Experiment $\mathbf{Real}_{\Pi^*,\mathcal{F}}(\kappa, \lambda)$: The experiment first sets $\delta = 0$ and runs $\gamma \leftarrow\!\!\$\ \mathsf{Enc}(\kappa)$. Then it accepts the following queries $poly(\lambda)$ many times (in any order):

- $(\texttt{Tamper}, f)$: Upon receiving as input a tampering query $f \in \mathcal{F}$, the experiment computes $\tilde{\gamma} = f(\gamma)$ and then $\gamma = \tilde{\gamma}$.
- $(\texttt{Eval}, x)$: Upon receiving as input an evaluation query $x \in \mathcal{X}$, the experiment computes and outputs $y \leftarrow\!\!\$\ \Pi_\Gamma(\delta, \gamma, x)$.

Experiment $\mathbf{Ideal}_{\Pi^*,\mathcal{F},\mathsf{S}}(\kappa, \lambda)$: The simulator S, given black-box access to the functionality $\Pi(\kappa, \cdot)$, is used in order to deal with evaluation and tampering queries.

Definition 10 (Tamper simulatability against persistent tampering for stateless functionalities)**.** *Let $\Gamma = (\mathsf{Enc}, \mathsf{Dec})$ be a (k, n)-coding scheme. We say that Γ is tamper simulatable against persistent tampering for stateless functionalities w.r.t. the family of functions $\mathcal{F}$, if for all stateless functionalities Π and for all PPT adversaries A asking a polynomial amount of tampering queries there exists a simulator S such that for any key $\kappa \in \mathcal{K}$ it holds*

$$\Delta^{\mathsf{A}}\left(\mathbf{Real}_{\Pi^*,\mathcal{F}}(\kappa, \lambda), \mathbf{Ideal}_{\Pi^*,\mathcal{F},\mathsf{S}}(\kappa, \lambda)\right) \in \mathit{negl}(\lambda)$$

Remark 8 (On self-destruct)**.** *The public value $\delta \in \{0, 1\}$ implements the so-called* self-destruct feature *which is well known to be necessary in order to obtain RKA-security against continuous* persistent *related-key attacks [GLM+04].*

The theorem below states that if Γ is $\mathcal{F}$-continuous persistent non-malleable, then the hardened primitive Π^* is $\mathcal{F}$-RKA-secure.

Theorem 7. *Let $\mathcal{F}$ be a family of functions, and let Γ be any computationally $\mathcal{F}$-continuous persistent non-malleable (k, n)-code w.r.t. message. Let Π be a stateless functionality with keys in $\mathcal{K} = \{0, 1\}^k$. Then the hardened functionality Π^* is tamper simulatable against tampering (for stateless functionalities) w.r.t. the family of functions $\mathcal{F}$.*

To define the simulator S, we make use of a new flag ι^*, which is initialized to 0 and becomes 1 after the first tampering query mauling the codeword. The description of S is as follows.

- First, initialize $\delta = 0$, $\iota^* = 0$, sample a random key $\kappa' \leftarrow\!\!\$\ \{0, 1\}^k$ and compute an encoding $\gamma \leftarrow\!\!\$\ \mathsf{Enc}(\kappa')$.
- For every evaluation query $(\texttt{Eval}, x)$ such that $\iota^* = 0$, run the keyed functionality on x and output $y \leftarrow\!\!\$\ \Pi(\kappa, x)$. If, instead, $\iota^* = 1$, compute and output $\Pi^*(\delta, \gamma, x)$.
- Upon receiving a tampering query $(\texttt{Tamper}, f)$, compute $\tilde{\gamma} = f(\gamma)$. If $\mathsf{Dec}(\tilde{\gamma}) \neq \mathsf{Dec}(\gamma)$, set $\iota^* = 1$. Replace and γ with $\tilde{\gamma}$.

Proof. Assume by contradiction that there exists a PPT adversary A and a key κ such that

$$\Delta^{\mathsf{A}}\left(\mathbf{Real}_{\Pi^*,\mathcal{F}}(\kappa, \lambda), \mathbf{Ideal}_{\Pi^*,\mathcal{F},\mathsf{S}}(\kappa, \lambda)\right) \notin \mathit{negl}(\lambda)$$

We define a reduction $\hat{\mathsf{A}}$ to the continuous persistent non-malleability of Γ as follows.

Pre-processing: The reduction sets $\delta = \iota^* = 0$, samples $\kappa' \leftarrow\!\!\$\ \{0, 1\}^k$ and submits κ, κ' to the oracle.

Evauation queries: Upon receiving an evaluation query $(\texttt{Eval}, x)$ such that $\iota^* = 0$, output $y \leftarrow\!\!\$\ \Pi(\kappa, x)$. Else, output $\perp$ if $\delta = 1$, and $y \leftarrow\!\!\$\ \Pi(\hat{\kappa}, x)$ otherwise.

Tampering queries: Upon receiving a tampering query $(\texttt{Tamper}, f)$ with $\iota^* = 0$, submit it to the oracle, and then:

- If the answer is $\tilde{\kappa} = \diamond$, do nothing.
- If it is $\perp$, set $\delta = 1$ and $\iota^* = 1$.
- Otherwise, set $\iota^* = 1$ and save $\hat{\kappa} = \tilde{\kappa}$.

If, instead, $\iota^* \neq 0$, forward f to the oracle and update $\hat{\kappa}$ with the answer $\tilde{\kappa}$.

Guess: Output the same as A.

We remark that, until $\iota^* = 0$, then the original key is preserved. This means, in particular, that the reduction is authorized to make further tampering queries. When $\iota^* = 1$ for the first time, then $\hat{\mathsf{A}}$ either learns the mauled key $\tilde{\kappa}$ or knows that the last tampering was invalid. In both cases, the reduction gets all the information needed in order to continue the simulation. This makes the simulation perfect, and concludes the proof. □

Remark 9 (On leakage resilience)**.** *Assuming the underlying code Γ is leakage-resilient non-malleable, we can extend Theorem 7 to the setting where the attacker is allowed to also leak bounded information from the memory.*

Remark 10 (On stateful functionalities)**.** *Theorem 7 readily extends to stateful functionalities, which, taken as input the usual parameters κ and x, also output a fresh key $\tilde{\kappa} \in \mathcal{K}$ together with the value $y \in \mathcal{Y}$. However, in this case the construction is less efficient, as hardening a stateful functionality inherently requires re-encoding at each step.*

D Necessity of super non-malleability in the decision-tree tampering construction

In this section, we show that our construction against decision-tree tampering strictly needs the inner split-state encoding to be *super* non-malleable, and therefore our requirement is not only an artifact of our proof.

First of all, we need to instantiate our construction with contrived primitives which allow us to perform the attack. Then, we proceed to describe the attack on a simplified version of our construction. Finally, we show how to remove the simplifications without invalidating the attack.

D.1 The contrived primitives

The primitives used by our construction are the following.

- A super non-malleable code w.r.t. codeword in the split-state model; since we are showing the attack, we drop the *super* requirement and we only require regular non-malleability w.r.t. codeword.
- A binary ramp secret sharing with $4t$ shares, privacy threshold t, reconstruction threshold $4t$.
- A signature scheme.

We drop the signature scheme for now and focus on the other two primitives.

Split-state non-malleable code. Let $(\mathsf{NMEnc}, \mathsf{NMDec})$ be a (possibly super) split-state non-malleable code w.r.t. codeword with codeword length $2c$ and consider the following encoding scheme.

Algorithm $\mathsf{NMEnc}'(\mu)$**.** Upon input $\mu \in \{0,1\}^m$:

1. Run $(\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}}) \leftarrow\!\!\$\ \mathsf{NMEnc}(0||\mu)$.
2. For $i \in \{\mathsf{L}, \mathsf{R}\}$, let $\gamma_i' := 0||0^c||\gamma_i$.
3. Output $(\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}})$.

Algorithm $\mathsf{NMDec}'(\gamma_{\mathsf{L}}', \gamma_{\mathsf{R}}')$**.** Proceed as follows:

1. For $i \in \{\mathsf{L}, \mathsf{R}\}$, parse $\gamma_i' := b_i||\gamma_i^{(1)}||\gamma_i^{(0)}$ and let $\gamma_i := \gamma_i^{(b_i)}$.
2. Run $b^*||\mu = \mathsf{NMDec}(\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}})$.
3. If $b^* = 0$ and $b_i||\gamma_i^{(1)} \neq 0^{c+1}$ for any $i \in \{0,1\}$, output $\perp$.
4. Output μ.

Lemma 15. *Let* $(\mathsf{NMEnc}, \mathsf{NMDec})$ *be a split-state* ℓ'*-leakage resilient non-malleable code w.r.t. codeword with* $\ell' = \ell + 1$*. Then, the above scheme* $(\mathsf{NMEnc}', \mathsf{NMDec}')$ *is a split-state* ℓ*-leakage resilient non-malleable code w.r.t. codeword.*

Proof. By reduction to split-state non-malleability w.r.t. codeword of $(\mathsf{NMEnc}, \mathsf{NMDec})$. Suppose that there exists an unbounded adversary A breaking non-malleability w.r.t. codeword of $(\mathsf{NMEnc}', \mathsf{NMDec}')$ and consider the following reduction $\hat{\mathsf{A}}$.

- Run A to obtain the challenge messages μ_0, μ_1.
- Construct, for $b \in \{0,1\}$, $\mu_b^* := 0||\mu_b$; then, send μ_0^*, μ_1^* to the challenger.
- Upon receiving a leakage query $(g_{\mathsf{L}}, g_{\mathsf{R}})$, construct, for $i \in \{\mathsf{L}, \mathsf{R}\}$, the leakage function g_i^* as follows.
  - Upon receiving state γ_i, let $\gamma_i^* := 0||0^c||\gamma_i$.
  - Compute and output $g_i(\gamma_i^*)$.

  Then, send the leakage query $(g_{\mathsf{L}}^*, g_{\mathsf{R}}^*)$ to the adversary and forward the result to the adversary.
- Upon receiving a leakage query $(g_{\mathsf{L}}, g_{\mathsf{R}})$, construct, for $i \in \{\mathsf{L}, \mathsf{R}\}$, the leakage function g_i^* as follows.
  - Upon receiving state γ_i, let $\gamma_i^* := 0||0^c||\gamma_i$.
  - Compute and output $g_i(\gamma_i^*)$.

  Then, send the leakage query $(g_{\mathsf{L}}^*, g_{\mathsf{R}}^*)$ to the challenger and forward the result to the adversary.
- Upon receiving a tampering query $(f_{\mathsf{L}}, f_{\mathsf{R}})$, construct, for $i \in \{\mathsf{L}, \mathsf{R}\}$, the leakage function h_i as follows.
  - Upon receiving state γ_i, let $\gamma_i^* := 0||0^c||\gamma_i$.
  - Compute $\tilde{b}_i||\tilde{\gamma}_i^{(1)}||\tilde{\gamma}_i^{(0)} = f_i(\gamma_i^*)$.

– Output 0 if $\tilde{b}_i||\tilde{\gamma}_i^{(1)} = 0^{c+1}$ and output 1 otherwise.

Then, for $i \in \{\mathsf{L}, \mathsf{R}\}$, construct the tampering function f_i^* as follows.

– Upon receiving state γ_i, let $\gamma_i^* := 0||0^c||\gamma_i$.
– Compute $\tilde{b}_i||\tilde{\gamma}^{(1)}||\tilde{\gamma}^{(0)} = f_i(\gamma_i^*)$.
– Output $\tilde{\gamma}^{(\tilde{b}_i)}$.

Finally, send the leakage query $(h_{\mathsf{L}}, h_{\mathsf{R}})$ to the challenger, thus receiving two bits $(b_{\mathsf{L}}^*, b_{\mathsf{R}}^*)$, and send the tampering query $(f_{\mathsf{L}}^*, f_{\mathsf{R}}^*)$, thus receiving either $\perp$ or some message $\tilde{b}^*||\tilde{\mu} \in \{0,1\} \times \mathcal{M}$.

- If the answer to the tampering query was $\perp$, forward $\perp$ to A.
- Otherwise, if $\tilde{b}^* = 0$ and $(b_{\mathsf{L}}^*, b_{\mathsf{R}}^*) \neq (0,0)$, forward $\perp$ to A; else, forward $\tilde{\mu}$ to A.
- Return the same distinguishing bit of A.

For the analysis, notice that the simulation is perfect. In particular, $\hat{\mathsf{A}}$ correctly constructs the codeword of NMEnc' for the leakage and tampering queries and then uses 1 bit of additional leakage from each state to check that the condition $\tilde{b}_i||\tilde{\gamma}_i^{(1)} = 0^{c+1}$ is satisfied when the first bit of the message of the inner scheme is 0. This means that, if A wants to modify the first $c+1$ bits of each state, A should also modify the first bit of the message, which is prevented by $(\mathsf{NMEnc}, \mathsf{NMDec})$ being non-malleable. The lemma follows. □

Ramp secret sharing. The only additional property we need for the ramp secret sharing is the ability to replace some of the indices without touching the other indices. Towards this, we use the fact that the string ζ represents c distinct numbers in $[n]$. Additionally, our construction only uses ζ to reconstruct a set $\mathcal{I} \in [n]$ and the ordering of the numbers in ζ is not needed.

The first idea is to split the sharing algorithm in three parts: the first part will encode the first position, the second part will encode the next c positions and the third part will encode the last c positions. In this way, it is possible to modify the first $c+1$ positions leaving untouched the last c positions. However, the positions in ζ are not sorted and, therefore, it is not guaranteed that modifying the first $c+1$ positions will affect the first $c+1$ bits of the codeword state.

The second idea is to sort the string ζ, which is not a problem since our construction does not need the ordering anyway. There are two possible ways to apply this solution.

1. We could relegate the process of sorting to the construction itself; however, in this case we would show an attack towards a construction which is different from ours.
2. We could relegate the process of sorting to the ramp secret sharing; however, in the process of sorting, the secret sharing algorithm loses some information which, although being not useful for our construction, breaks completely its correctness.

The final idea is to relegate the process of sorting to the ramp secret sharing and then additionally store a string which allows to revert the sorting. Given a ramp secret sharing $(\mathsf{Enc}_{\mathsf{RSS}}, \mathsf{Dec}_{\mathsf{RSS}})$ as a base, we construct our new ramp secret sharing as follows.

Algorithm $\mathsf{Enc}'_{\mathsf{RSS}}(\zeta)$**.** Upon input $\zeta \in \{0,1\}^{c' \log(n)}$:

1. For all $i \in [c']$, let pos_i the position of the i-th smallest value in ζ.
2. Parse $(\mathsf{pos}_1, \ldots, \mathsf{pos}_{c'})$ as a string $\sigma \in \{0,1\}^{c' \log(c')}$ corresponding to c' binary representations of distinct numbers in $[c']$.

3. Let ζ' be a sorted version of ζ.
4. Parse $\zeta' = (\zeta'_1, \zeta'_2, \zeta'_3) \in \{0,1\}^{\log(n)} \times \{0,1\}^{c\log(n)} \times \{0,1\}^{c\log(n)}$.
5. Output $(\mathsf{Enc}_{\mathsf{RSS}}(\zeta'_1), \mathsf{Enc}_{\mathsf{RSS}}(\zeta'_2), \mathsf{Enc}_{\mathsf{RSS}}(\zeta'_3), \mathsf{Enc}_{\mathsf{RSS}}(\sigma))$.

Algorithm $\mathsf{Dec}'_{\mathsf{RSS}}(\gamma_1^{\mathsf{RSS}}, \gamma_2^{\mathsf{RSS}}, \gamma_3^{\mathsf{RSS}}, \gamma_4^{\mathsf{RSS}})$**.** Proceed as follows:

1. Decode $\sigma = \mathsf{Dec}_{\mathsf{RSS}}(\gamma_4^{\mathsf{RSS}})$ and, for all $i \in [3]$, $\zeta'_i = \mathsf{Dec}_{\mathsf{RSS}}(\gamma_i^{\mathsf{RSS}})$.
2. Put back $\zeta' = (\zeta'_1, \zeta'_2, \zeta'_3)$.
3. For all $i \in [c']$, obtain ζ by putting $\zeta'[i]$ in position $\sigma[i]$.
4. Output ζ.

We now proceed to show that this is a ramp secret sharing with the desired properties.

Correctness. Correctness follows by observing that, while we secret share a sorted version ζ' of ζ, we also secret share the necessary information σ to recover ζ from ζ'.

Privacy. Intuitively, the scheme inherits the privacy threshold of the inner ramp secret sharing $(\mathsf{Enc}_{\mathsf{RSS}}, \mathsf{Dec}_{\mathsf{RSS}})$. Indeed, if the adversary learns at most t bits, it means that it learns at most t bits from each instance of $(\mathsf{Enc}_{\mathsf{RSS}}, \mathsf{Dec}_{\mathsf{RSS}})$, which is below the privacy threshold.

More formally, let, for $i \in [4]$, Enc_i be the i-th instance of $\mathsf{Enc}_{\mathsf{RSS}}$, and let, for $b \in \{0,1\}$, $\zeta_1^{(b)}, \zeta_2^{(b)}, \zeta_3^{(b)}, \sigma^{(b)}$ be the encoding of $\zeta^{(b)}$ used inside $\mathsf{Enc}'_{\mathsf{RSS}}$. Finally, let $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2 \cup \mathcal{I}_3 \cup \mathcal{I}_4$ be a set of at most t indices, where, for each $i \in [4]$, $\mathcal{I}_i$ is a set of indices in the codeword of Enc_i. Then,

$$
\begin{aligned}
\mathsf{Enc}'_{\mathsf{RSS}}(\zeta^{(0)})|_{\mathcal{I}} &\equiv \left(\mathsf{Enc}_1(\zeta_1^{(0)})|_{\mathcal{I}_1}, \mathsf{Enc}_2(\zeta_2^{(0)})|_{\mathcal{I}_2}, \mathsf{Enc}_3(\zeta_3^{(0)})|_{\mathcal{I}_3}, \mathsf{Enc}_4(\sigma^{(0)})|_{\mathcal{I}_4}\right) \\
&\equiv \left(\mathsf{Enc}_1(\zeta_1^{(1)})|_{\mathcal{I}_1}, \mathsf{Enc}_2(\zeta_2^{(0)})|_{\mathcal{I}_2}, \mathsf{Enc}_3(\zeta_3^{(0)})|_{\mathcal{I}_3}, \mathsf{Enc}_4(\sigma^{(0)})|_{\mathcal{I}_4}\right) \\
&\equiv \left(\mathsf{Enc}_1(\zeta_1^{(1)})|_{\mathcal{I}_1}, \mathsf{Enc}_2(\zeta_2^{(1)})|_{\mathcal{I}_2}, \mathsf{Enc}_3(\zeta_3^{(0)})|_{\mathcal{I}_3}, \mathsf{Enc}_4(\sigma^{(0)})|_{\mathcal{I}_4}\right) \\
&\equiv \left(\mathsf{Enc}_1(\zeta_1^{(1)})|_{\mathcal{I}_1}, \mathsf{Enc}_2(\zeta_2^{(1)})|_{\mathcal{I}_2}, \mathsf{Enc}_3(\zeta_3^{(1)})|_{\mathcal{I}_3}, \mathsf{Enc}_4(\sigma^{(0)})|_{\mathcal{I}_4}\right) \\
&\equiv \left(\mathsf{Enc}_1(\zeta_1^{(1)})|_{\mathcal{I}_1}, \mathsf{Enc}_2(\zeta_2^{(1)})|_{\mathcal{I}_2}, \mathsf{Enc}_3(\zeta_3^{(1)})|_{\mathcal{I}_3}, \mathsf{Enc}_4(\sigma^{(1)})|_{\mathcal{I}_4}\right) \\
&\equiv \mathsf{Enc}'_{\mathsf{RSS}}(\zeta^{(1)})|_{\mathcal{I}},
\end{aligned}
$$

where by $\mathbf{X} \equiv \mathbf{Y}$ we denote that the two random variables $\mathbf{X}$ and $\mathbf{Y}$ are identically distributed, and in each step we gradually replaced $\mathsf{Enc}_i(\zeta_i^{(0)})$ with $\mathsf{Enc}_i(\zeta_i^{(1)})$ (or $\mathsf{Enc}_i(\sigma^{(0)})$ with $\mathsf{Enc}_i(\sigma^{(1)})$ if $i = 4$) since they are identically distributed and independent of any other value in the tuple.

D.2 The simplified attack.

We now show the attack in the simplified setting in which there are no signature schemes. First of all, we instantiate our construction with the non-malleable code $(\mathsf{NMEnc}', \mathsf{NMDec}')$ and the ramp secret sharing $(\mathsf{Enc}'_{\mathsf{RSS}}, \mathsf{Dec}'_{\mathsf{RSS}})$ from Appendix D.1. The attacker also needs access to the underlying primitives, respectively the super split-state non-malleable code $(\mathsf{NMEnc}, \mathsf{NMDec})$ and the ramp secret sharing $(\mathsf{Enc}_{\mathsf{RSS}}, \mathsf{Dec}_{\mathsf{RSS}})$.

In what follows, we parse the whole decision-tree codeword as

$$
\left(\gamma_{1,\mathsf{L}}^{\mathsf{RSS}}, \gamma_{2,\mathsf{L}}^{\mathsf{RSS}}, \gamma_{3,\mathsf{L}}^{\mathsf{RSS}}, \gamma_{4,\mathsf{L}}^{\mathsf{RSS}}, \gamma_{1,\mathsf{L}}, \gamma_{2,\mathsf{L}}, \gamma_{3,\mathsf{L}}, \gamma_{1,\mathsf{R}}^{\mathsf{RSS}}, \gamma_{2,\mathsf{R}}^{\mathsf{RSS}}, \gamma_{3,\mathsf{R}}^{\mathsf{RSS}}, \gamma_{4,\mathsf{R}}^{\mathsf{RSS}}, \gamma_{1,\mathsf{R}}, \gamma_{2,\mathsf{R}}, \gamma_{3,\mathsf{R}}\right),
$$

where, for $i \in \{\mathsf{L}, \mathsf{R}\}$, $(\gamma_{1,i}^{\mathsf{RSS}}, \gamma_{2,i}^{\mathsf{RSS}}, \gamma_{3,i}^{\mathsf{RSS}}, \gamma_{4,i}^{\mathsf{RSS}})$ is the output of $\mathsf{Enc}_{\mathsf{RSS},i}$, i.e. the instantiation of $\mathsf{Enc}_{\mathsf{RSS}}$ for encoding the state i, and $(\gamma_{1,i}, \gamma_{2,i}, \gamma_{3,i})$ is the encoding containing zeroes and the state i of the split-state codeword of NMEnc'; in particular, $\gamma_{1,i}$ is the first bit and $\gamma_{2,i}$ is the sequence of the next c bits. A description of the attacker A follows.

1. Sample two different messages $\mu_0, \mu_1 \in \mathcal{M}$ and send them for the challenge.
2. Sample two different messages $\hat{\mu}_0, \hat{\mu}_1 \in \mathcal{M} \setminus \{\mu_0, \mu_1\}$.
3. For $b \in \{0,1\}$, compute the super split-state non-malleable encoding $(\hat{\gamma}_{\mathsf{L},b}, \hat{\gamma}_{\mathsf{R},b}) = \mathsf{NMEnc}(1||\hat{\mu}_b)$.
4. For $b \in \{0,1\}, i \in \{\mathsf{L}, \mathsf{R}\}$, construct the strings $\hat{\zeta}_{1,i,b}, \zeta_{2,i,b}$ encoding the indices in $[c+1]$ in ascending order.
5. For $b \in \{0,1\}, i \in \{\mathsf{L}, \mathsf{R}\}, j \in [2]$, compute $\hat{\gamma}_{j,i,b}^{\mathsf{RSS}} \leftarrow\!\!\$\ \mathsf{Enc}_{\mathsf{RSS}}(\hat{\zeta}_{j,i,b})$.
6. For every position k in the codeword, construct the tampering function f_k as follows.
   - Trees for tampering with $\gamma_{j,i}^{\mathsf{RSS}}$ for $i \in \{\mathsf{L}, \mathsf{R}\}$ and $j \in [2]$: read bit b in position k and replace $\gamma_{j,i}^{\mathsf{RSS}}$ with $\hat{\gamma}_{j,i,b}^{\mathsf{RSS}}$.
   - Trees for tampering with $\gamma_{1,i}$ for $i \in \{\mathsf{L}, \mathsf{R}\}$: output constant 1.
   - Trees for tampering with $\gamma_{2,i}$ for $i \in \{\mathsf{L}, \mathsf{R}\}$: read bit b in position k and replace $\gamma_{2,i}$ with $\hat{\gamma}_{i,b}$.
   - All the other trees: act as identity (i.e. do not modify the target bit).
7. For $i \in \{\mathsf{L}, \mathsf{R}\}$ and for all positions k in which are stored $\gamma_{3,i}^{\mathsf{RSS}}, \gamma_{4,i}^{\mathsf{RSS}}, \gamma_{3,i}$, send the tampering query f_k, receiving either $\hat{\mu}_0$ or $\hat{\mu}_1$ depending on the value of bit b in position k.
8. For $i \in \{\mathsf{L}, \mathsf{R}\}$, reconstruct $\zeta_{3,i} = \mathsf{Dec}_{\mathsf{RSS}}(\gamma_{3,i}^{\mathsf{RSS}})$.
9. For $i \in \{\mathsf{L}, \mathsf{R}\}$, use $\zeta_{3,i}$ to recover the split-state codeword γ_i.
10. Reconstruct $\mu_{b^*} = \mathsf{NMDec}(\gamma_{\mathsf{L}}, \gamma_{\mathsf{R}})$.
11. Output b^*.

Notice that, by construction, the ramp secret sharing is malleable and allows to modify the first $c+1$ positions without affecting the other ones. Moreover, the first $c+1$ bits of the encoded state are 0, because the encoding doesn't modify the order of the bits and, by construction, the first $c+1$ bits of the split-state codeword are 0. Finally, A modifies the first bit of the state so that NMDec' reconstructs using the next c bits instead of the last c bits, and modifies the relevant bits to an encoding of $\hat{\mu}_b$, depending on the value b of position k. This allows A to learn the bit b in position k. Iterating such attack for all the positions necessary to recover the second half of the split-state codeword, the adversary is able to fully recover the original codeword and, therefore, the original message.

D.3 How to extend the attack to the original construction.

It only remains to show how to deal with the signature scheme.

Let $(\mathsf{Gen}, \mathsf{Sign}, \mathsf{SigVer})$ be a signature scheme, let $(\overline{vk}, \overline{sk})$ be a fixed key pair and consider the following scheme.

Algorithm Gen'. Upon input 1^λ, run $(vk, sk) \leftarrow\!\!\$\ \mathsf{Gen}(1^\lambda)$ until $vk \neq \overline{vk}$; then output (vk, sk).

Algorithm $\mathsf{Sign}'(sk, \mu)$**.** Run $\sigma \leftarrow\!\!\$\ \mathsf{Sign}(sk, \mu)$ and output σ.

Algorithm $\mathsf{SigVer}'(vk, \mu, \sigma)$**.** If $vk = \overline{vk}$, return 1; otherwise, run $b = \mathsf{SigVer}(vk, \mu, \sigma)$ and output b.

For the analysis, notice that $(\mathsf{Gen}', \mathsf{Sign}', \mathsf{SigVer}')$ and $(\mathsf{Gen}, \mathsf{Sign}, \mathsf{SigVer})$ are exactly the same scheme conditioned on $\overline{vk} \neq vk$; moreover, since Gen' never outputs $\overline{vk}$, after the key generation the two schemes behave identically and, therefore, have the same security.

Using this signature scheme, the attack in Appendix D.2 can be extended to the general case as follows.

- We replace the decision-tree codeword with

$$\Big(\sigma_{\mathsf{L}}, \gamma^{\mathsf{RSS}}_{1,\mathsf{L}}, \gamma^{\mathsf{RSS}}_{2,\mathsf{L}}, \gamma^{\mathsf{RSS}}_{3,\mathsf{L}}, \gamma^{\mathsf{RSS}}_{4,\mathsf{L}}, \gamma_{1,\mathsf{L}}, \gamma_{2,\mathsf{L}}, \gamma_{3,\mathsf{L}}, \\ \sigma_{\mathsf{R}}, \gamma^{\mathsf{RSS}}_{1,\mathsf{R}}, \gamma^{\mathsf{RSS}}_{2,\mathsf{R}}, \gamma^{\mathsf{RSS}}_{3,\mathsf{R}}, \gamma^{\mathsf{RSS}}_{4,\mathsf{R}}, \gamma_{1,\mathsf{R}}, \gamma_{2,\mathsf{R}}, \gamma_{3,\mathsf{R}}\Big),$$

- In step Item 3, we replace

  > For $b \in \{0,1\}$, compute the super split-state non-malleable encoding $(\hat{\gamma}_{\mathsf{L},b}, \hat{\gamma}_{\mathsf{R},b}) = \mathsf{NMEnc}(1||\hat{\mu}_b)$.

  with

  > For $b \in \{0,1\}$, compute the super split-state non-malleable encoding $(\hat{\gamma}_{\mathsf{L},b}, \hat{\gamma}_{\mathsf{R},b}) = \mathsf{NMEnc}(1||\hat{\mu}_b||\overline{vk}_{\mathsf{L}}||\overline{vk}_{\mathsf{R}})$, where $\overline{vk}_{\mathsf{L}}$ (resp. $\overline{vk}_{\mathsf{R}}$ comes from the fixed key pair $(\overline{vk}, \overline{sk})$ of the left (resp. right) instance of the signature scheme.

Notice that, since the adversary A replaces the verification keys with the special keys $\overline{vk}_{\mathsf{L}}, \overline{vk}_{\mathsf{R}}$, the signature always verifies and, therefore, no other modification is needed.