# The Chaining Lemma and Its Application

Ivan Damgård[1,⋆], Sebastian Faust[2,⋆⋆], Pratyay Mukherjee[1,⋆⋆⋆], and Daniele Venturi[3]

[1] Aarhus University
[2] EPFL
[3] Sapienza University of Rome

**Abstract.** We present a new information-theoretic result which we call the Chaining Lemma. It considers a so-called "chain" of random variables, defined by a source distribution $X^{(0)}$ with high min-entropy and a number (say, $t$ in total) of arbitrary functions $(T_1, \ldots, T_t)$ which are applied in succession to that source to generate the chain $X^{(0)} \xrightarrow{T_1} X^{(1)} \xrightarrow{T_2} X^{(2)} \cdots \xrightarrow{T_t} X^{(t)}$. Intuitively, the Chaining Lemma guarantees that, if the chain is not too long, then either (i) the entire chain is "highly random", in that every variable has high min-entropy; or (ii) it is possible to find a point $j$ $(1 \leq j \leq t)$ in the chain such that, conditioned on the end of the chain i.e. $X^{(j)} \xrightarrow{T_{j+1}} X^{(j+1)} \cdots \xrightarrow{T_t} X^{(t)}$, the preceding part $X^{(0)} \xrightarrow{T_1} X^{(1)} \cdots \xrightarrow{T_j} X^{(j)}$ remains highly random. We think this is an interesting information-theoretic result which is intuitive but nevertheless requires rigorous case-analysis to prove.

We believe that the above lemma will find applications in cryptography. We give an example of this, namely we show an application of the lemma to protect essentially any cryptographic scheme against memory-tampering attacks. We allow several tampering requests, the tampering functions can be arbitrary, however, they must be chosen from a bounded size set of functions that is fixed a priori.

## 1 Introduction

Assume that we have a uniform random distribution over some finite set $\mathcal{X}$, represented by a discrete random variable $X$. Let us now apply an arbitrary (deterministic) function $T$ to $X$ and denote the output random variable by $X' = T(X)$. Since $T$ is an arbitrary function, the variable $X'$ can also be arbitrarily distributed. Consider now the case where $X'$ is "easy to predict", or more concretely where $X'$ has "low" min-entropy. A natural question, in this case, is *how much information can $X'$ reveal about $X$?* or more formally, *how much min-entropy can $X$ have if we condition on $X'$?*

⋆ Partially supported by Danish Council for Independent Research via DFF Starting Grant 10-081612.
⋆⋆ Supported by the Marie Curie IEF/FP7 project GAPS, grant number: 626467.
⋆⋆⋆ Partially supported by Danish Council for Independent Research via DFF Starting Grant 10-081612. Partially supported by the ERC Starting Grant 279447.


A. Lehmann and S. Wolf (Eds.): ICITS 2015, LNCS 9063, pp. 181–196, 2015.
DOI: 10.1007/978-3-319-17470-9_11

Intuitively, one might expect that since $X'$ has low entropy, it cannot tell us much about $X$, so $X$ should still be "close to random" and hence have high entropy. While this would be true for Shannon entropy, it turns out to be completely false for min-entropy. This may seem a bit counter-intuitive at first, but is actually easy to see from an example: Let $T$ be the function which maps half of the elements in $\mathcal{X}$ to one "heavy" point but is injective on all the other elements. For this $T$, the variable $X'$ has very small min-entropy (namely 1) because the heavy point occurs with probability 1/2. But on the other hand, $X'$ reveals everything about $X$ half the time, and so the entropy of $X$ in fact decreases very significantly (on average) when $X'$ is given. So despite having very low min-entropy, $X' = T(X)$ does reveal a lot about $X$.

There is, however, a more refined statement that will be true for min-entropy: Let $E$ be the event that $X$ takes one of the values that are *not* mapped to the "heavy point" by $T$, while $\bar{E}$ is the event that $X$ is mapped to the heavy point. Now, conditioned on $E$, both $X_{|E}$ and $X'_{|E}$ have high min-entropy. On the other hand, conditioned on $\bar{E}$, $X_{|\bar{E}}$ will clearly have the same (high) min-entropy whether we are given $X'_{|E}$ or not.

This simple observation leads to the following conjecture: there always exists an event $E$ such that: (i) Conditioned on $E$, both $X$ and $X'$ have "high" min-entropy, (ii) conditioned on $\bar{E}$, $X'$ reveals "little" about $X$. In this paper, from a very high-level, we mainly focus into settling (a generalization of) this conjecture, which results in our main contribution: the information-theoretic lemma which we call the Chaining Lemma.

*Main Question.* Towards generalizing the above setting let us rename, for notational convenience, the above symbols as follows: $X^{(0)} \equiv X$, $T_1 \equiv T$ and $X^{(1)} \equiv X'$. We consider $t$ (deterministic) functions $T_1, T_2, \ldots, T_t$ which are applied to the the variables sequentially starting from $X^{(0)}$. In particular, each $T_i$ is applied to $X^{(i-1)}$ to produce a new variable $X^{(i)} = T_i(X^{(i-1)})$ for $i \in [t]$. We call the sequence of variables $(X^{(0)}, \ldots, X^{(t)})$ a "chain" which is completely defined by the "source" distribution $X^{(0)}$ and the sequence of $t$ functions $(T_1, \ldots, T_t)$. It can be presented more vividly as follows: $X^{(0)} \xrightarrow{T_1} X^{(1)} \xrightarrow{T_2} X^{(2)} \ldots \xrightarrow{T_t} X^{(t)}$.

We are now interested in the min-entropy of $X^{(1)}, \ldots, X^{(t)}$. Of course, each variable $X^{(i)}$ has min-entropy less than (or equal to) the preceding variable $X^{(i-1)}$ (as a deterministic function can not generate randomness). Assume now that we fix some threshold value $u$ and consider any value of min-entropy less than $u$ to be "low". Assume further that the source has min-entropy much larger than $u$. As a motivation, one may think of a setting where each $X^{(i)}$ is used as key in some cryptographic application, where, as long $X^{(i)}$ has high min-entropy we are fine and the adversary will not learn something he should not. But if $X^{(i)}$ has low min-entropy, things might go wrong and the adversary might learn $X^{(i)}$.

Now, there are two possible scenarios for the above chain: either (i) all the variables (hence the last variable $X^{(t)}$) in the chain have high min-entropy; or (ii) one or more variable (obviously including the last variable $X^{(t)}$) has low min-entropy. In case (i), everything is fine. But in case (ii), things might go wrong

at a certain point. We now want to ask if we can at least "save" some part of the chain, i.e., *can we find a point in the chain such that if we condition on all the variables after that point, all the preceding variables (obviously including the source $X^{(0)}$) would still have high min-entropy?* This hope might be justified if $t$ is small enough compared to the entropy of $X^{(0)}$: since the entropy drops below $u$ after a small number of steps, there must be a point (say $j$) where the entropy falls "sharply", i.e., $X^{(j)}$ has much smaller min-entropy than $X^{(j-1)}$. However, as the above example shows, even if there is a large gap in min-entropy between two successive variables ($X^{(j)}$ and $X^{(j-1)}$ in this case), the succeeding one ($X^{(j)}$) might actually reveal a lot about the preceding one ($X^{(j-1)}$) on average. So it is not clear that we can use $j$ as the point we are looking for. However, one could hope that a generalised version of the above conjecture might be true, namely there might exist some event, further conditioning on which, all variables would have high min-entropy, and on the other hand, conditioning on the complement, $X^{(j-1)}$ (and hence the entire preceding chain) would have high min-entropy. Essentially that is what our Chaining Lemma says, which we present next although in an informal way. We give the formal statement and proof of the lemma in Section 3.

**Lemma 1 (The Chaining Lemma, Informal).** *Let $X^{(0)}$ be a uniform random variable over $\mathcal{X}$ and $(T_0, \ldots, T_t)$ be arbitrary functions mapping $\mathcal{X} \to \mathcal{X}$ and defining a chain $X^{(0)} \xrightarrow{T_1} X^{(1)} \xrightarrow{T_2} X^{(2)} \cdots \xrightarrow{T_t} X^{(t)}$. If the chain is "sufficiently short", there exists an event $E$ such that (i) if $E$ happens, then all the variables $(X^{(0)}, \ldots, X^{(t)})$ (conditioned on $E$) have "high" min-entropy; otherwise (ii) if $E$ does not happen there is an index $j$ such that conditioning on $X^{(j)}$ (and also on $\bar{E}$) all the previous variables namely $X^{(0)}, \ldots, X^{(j-1)}$ have "high" min-entropy.*

*Application to Tamper-resilient Cryptography.* Although we think that the Chaining Lemma is interesting in its own right, in this paper we provide an application in cryptography, precisely in tamper-resilient cryptography. In tamper-resilient cryptography the main goal is to "theoretically" protect cryptographic schemes against so-called fault attacks which are found to be devastating (as shown by [5,12] and many more). In this model, the adversary, in addition to standard black-box access to a primitive, is allowed to change its secret state [9,28,23,32,8], or its internals [30,27,18,19], and observes the effect of such changes at the output. In this paper we restrict ourselves to the model where the adversary is not allowed to alter the computation, but only the secret state (i.e. only the memory of the device, but not the circuitry, is subject to tampering).

To illustrate such memory tampering, consider a digital signature scheme $\mathsf{Sign}$ with public/secret key pair $(pk, sk)$. The tampering adversary obtains $pk$ and can replace $sk$ with $T(sk)$ for arbitrary tampering function $T$. Then, the adversary gets access to an oracle $\mathsf{Sign}(T(sk), \cdot)$, i.e., to a signing oracle running with the tampered key $T(sk)$. As usual the adversary wins the game by

outputting a valid forgery with respect to the original public key $pk$.[1] In the most general setting, the adversary is allowed to ask an arbitrary polynomial number of tampering queries. However, a general impossibility result by Gennaro *et al.* [28] shows that the above flavour of tamper resistance is unachievable without further assumptions. To overcome this impossibility one usually relies on self-destruct (e.g., [23,15,1,14,13,24,25,26,17,2,3,4,16,31]), or limits the power of the tampering function (e.g., [9,33,7,6,29,34,36,10,11,31,35]).

Recently Damgård *et al.* [20] proposed a different approach where, instead of limiting the type of allowed modifications, one assumes an upper bound on the number of tampering queries that the adversary can ask, so that now the attacker can issue some a-priori fixed number $t$ of *arbitrary* tampering queries. As argued by [20], this limitation is more likely to capture realistic tampering attacks. They also show how to construct public key encryption and identification schemes secure against bounded leakage[2] and tampering (BLT) attacks.

The above model fits perfectly with the setting of the Chaining Lemma, as we consider a limited number of tampering functions $(T_1, \ldots, T_t)$, for some fixed bound $t$, applied on a uniform (or close to uniform) secret-state $X^{(0)}$. Now recall that Lemma 1 guarantees that, for "small enough" $t$, the source distribution stays unpredictable in essentially "any" case. Therefore, the source can be used as a "highly unpredictable" secret-key resisting $t$ arbitrary tampering attacks. As a basic application of the Chaining Lemma, we show in Section 4 that *any* cryptographic scheme can be made secure in the BLT model. To the best of our knowledge, this is the first such general result that holds for arbitrary tampering functions and multiple tampering queries. The price we pay for this is that the tampering functions must be chosen from a bounded-size set that is fixed a priori.

Previous work by Faust *et al.* [26], shows how to protect generically against tampering using a new primitive called *non-malleable key-derivation*. This result also works for arbitrary tampering functions, does not require that a small set of functions is fixed in advance, but works only for one-time tampering.

## 2 Preliminaries

### 2.1 Notation

For $n \in \mathbb{N}$, we write $[n] := \{1, \ldots, n\}$. Given a set $\mathcal{S}$, we write $s \leftarrow \mathcal{S}$ to denote that element $s$ is sampled uniformly from $\mathcal{S}$. If $\mathsf{A}$ is an algorithm, $y \leftarrow \mathsf{A}(x)$ denotes an execution of $\mathsf{A}$ with input $x$ and output $y$; if $\mathsf{A}$ is randomized, then $y$ is a random variable.

We denote with $k$ the security parameter. A machine $\mathsf{A}$ is called *probabilistic polynomial time* (PPT) if for any input $x \in \{0,1\}^*$ the computation of $\mathsf{A}(x)$ terminates in at most $poly(|x|)$ steps and $\mathsf{A}$ is probabilistic (i.e., it uses randomness

[1] Notice that $T$ may be the identity function, in which case we get the standard security notion of digital signature scheme as a special case.

[2] The adversary is also allowed to leak a bounded—yet arbitrary—amount of information on the secret key; we refer the reader to Section 4 for the details.

as part of its logic). Random variables are usually denoted by capital letters. We sometimes abuse notation and denote a distribution and the corresponding random variable with the same capital letter, say $X$. We write $\mathsf{sup}(X)$ for the support of $X$. Given an event $E$, we let $X_{|E}$ be the conditional distribution of $X$ conditioned on $E$ happening. The statistical distance of two random variables $X$ and $Y$, defined over a common set $\mathcal{S}$ is $\Delta(X;Y) = \frac{1}{2}\sum_{s\in\mathcal{S}}|\Pr[X=s] - \Pr[Y=s]|$. Given a random variable $Z$, the statistical distance of $X$ and $Y$ conditioned on $Z$ is defined as $\Delta(X;Y|Z) = \Delta((X,Z);(Y,Z))$.

### 2.2 Information Theory Basics

The min-entropy of a random variable $X$ over a set $\mathcal{X}$ is defined as $\mathbf{H}_\infty(X) := -\log\max_x \Pr[X=x]$, and measures how $X$ can be predicted by the best (unbounded) predictor. The conditional average min-entropy [22] of $X$ given a random variable $Z$ (over a set $\mathcal{Z}$) possibly dependent on $X$, is defined as

$$\widetilde{\mathbf{H}}_\infty(X|Z) := -\log \mathbb{E}_{z\leftarrow Z}[2^{-\mathbf{H}_\infty(X|Z=z)}] = -\log\sum_{z\in\mathcal{Z}}\Pr[Z=z]\cdot 2^{-\mathbf{H}_\infty(X|Z=z)}.$$

We say that a distribution $X$ over a set $\mathcal{X}$ of size $|\mathcal{X}| = 2^n$ is $(\alpha, n)$-good if $\mathbf{H}_\infty(X) \geq \alpha$ and $\Pr[X=x] \geq 2^{-n}$ for all $x \in \mathsf{sup}(X)$.

We will rely on the following basic property (see [22, Lemma 2.2]).

**Lemma 2.** *For all random variables $X, Z$ and $\Lambda$ over sets $\mathcal{X}$, $\mathcal{Z}$ and $\{0,1\}^\lambda$ such that $\widetilde{\mathbf{H}}_\infty(X|Z) \geq \alpha$, we have that*

$$\widetilde{\mathbf{H}}_\infty(X|Z,\Lambda) \geq \widetilde{\mathbf{H}}_\infty(X|Z) - \lambda \geq \alpha - \lambda.$$

The above lemma can be easily extended to the case of random variables $\Lambda$ with bounded support, i.e., $\widetilde{\mathbf{H}}_\infty(X|Z,\Lambda) \geq \widetilde{\mathbf{H}}_\infty(X|Z) - \log|\mathsf{sup}(\Lambda)|$.

## 3 The Chaining Lemma

Before presenting the statement and proof of the Chaining Lemma, we state and prove two sub-lemmas. We do not provide any intuitions at this point regarding the whole proof of the Chaining Lemma due to involvement of rigorous case-analysis. Instead, we take a modular approach presenting intuitions step-by-step for each of the sub-lemmas and finally providing an intuition of the Chaining Lemma after the proof of these sub-lemmas.

The first lemma states that if the support of a distribution is sufficiently large then there always exists an event $E$ such that, conditioned on $E$, the conditional distribution has high min-entropy.

**Lemma 3.** *For $n \in \mathbb{N}_{>1}$ let $c$ be some parameter such that $\sqrt{n} < c < n$. Let $\mathcal{X}$ be a set of size $2^n = |\mathcal{X}|$ and $X$ be a distribution over $\mathcal{X}$ with $|\mathsf{sup}(X)| > 2^c$ such that for all $x \in \mathsf{sup}(X)$ we have $\Pr[X=x] \geq \frac{1}{2^n}$. There exists an event $E$ such that:*

*(i)* $\mathbf{H}_\infty(X_{|E}) > c - 2\sqrt{n}$, *and*
*(ii)* $|\mathsf{sup}(X_{|\overline{E}})| < |\mathsf{sup}(X)|$.

*Proof.* Intuitively, the lemma is proven by showing that if a distribution has sufficiently large support, then over a large subset of the support the distribution must be "almost" flat. We will describe below what it means for a distribution to be "almost flat". We then define an event $E$ that occurs when $X$ takes some value in the almost flat area. Clearly, $X$ conditioned on $E$ must be "almost" uniformly distributed, and if furthermore the support of $X$ conditioned on $E$ is still sufficiently large, we get that $\mathbf{H}_\infty(X_{|E})$ must be large. We proceed with the formal proof.

We introduce a parameter $b$ which is a positive integer such that $c > n/b$. We explain how to set the value of $b$ later. For ease of description we assume that $n$ is a multiple of $b$. We start by defining what it means for an area to be flat. For some probability distribution $X$ we define $k \in [2^{n/b} - 1]$ sets as follows:

- $I_k := \left\{ x \in \mathsf{sup}(X) : \frac{k^b}{2^n} \leq \Pr[X = x] < \frac{(k+1)^b}{2^n} \right\}$, for $k \in [2^{n/b} - 1]$ and
- $I_{2^{n/b}} := \{x \in \mathsf{sup}(X) : \Pr[X = x] = 1\}$.

These sets characterize the (potential) flat areas in the distribution $X$ as the probability of all values in some set $I_k$ lies in a certain range that is bounded from below and above. Clearly, the sets $I_k$ are pairwise disjoint and cover the whole space between $1/2^n$ and 1. Therefore, each $x \in \mathsf{sup}(X)$ with some probability $\Pr[X = x]$ must fall into some unique set $I_k$.

We denote by $I_m$ the set that contains the most elements among all sets $I_k$, and define the event $E$ as the event that occurs when $x \in \mathsf{sup}(X)$ falls into $I_m$, i.e., $X$ takes a value that falls in the largest set $I_m$. We now lower bound the probability that $E$ occurs.

$$\Pr[E] \geq |I_m| \frac{m^b}{2^n} \tag{1}$$

$$\geq 2^{c-n/b} \frac{m^b}{2^n}. \tag{2}$$

Inequality (1) holds as for all $x \in I_m$ we have $\Pr[X = x] \geq \frac{m^b}{2^n}$. (2) follows from the fact that $I_m$ must have size at least $2^{c-n/b}$, as there are $2^{n/b}$ sets and there are at least $2^c$ elements in the support of $X$.

As $\mathbf{H}_\infty(X_{|E}) = -\log \max_x \Pr[X = x|E]$, we can give a lower bound for the min entropy of $X_{|E}$ by upper bounding $\Pr[X = x|E]$. More precisely,

$$\Pr[X = x|E] = \frac{\Pr[X = x \wedge E]}{\Pr[E]}$$

$$< \frac{(m+1)^b/2^n}{2^{(c-n/b)} m^b/2^n} \tag{3}$$

$$= \left(1 + \frac{1}{m}\right)^b 2^{-c+n/b}$$

$$\leq 2^{b-c+n/b}. \tag{4}$$

Inequality (3) uses (2) and the fact that $\Pr[X = x \wedge E] < \frac{(m+1)^b}{2^n}$ by definition of $I_m$. (4) follows from $m \geq 1$. This implies that $\mathbf{H}_\infty(X_{|E}) > c - n/b - b$. Now we observe that the loss in min-entropy, given by $(b + n/b)$ is minimum when $b = \sqrt{n}$. Since $b$ is a free parameter, we fix $b := \sqrt{n}$ (note that, since $c > \sqrt{n}$, the constraint $c > n/b$ holds) to get $\mathbf{H}_\infty(X_{|E}) > n - 2\sqrt{n}$ as stated in part (i) of the lemma.

For part (ii), it is easy to see from the definition of $E$ that the support of the conditional probability distribution $X_{|\overline{E}}$ decreases by at least $2^{(c-n/b)}$ points (as these points belong to $E$). Clearly, $|\mathsf{sup}(X_{|\overline{E}})| \leq |\mathsf{sup}(X)| - 2^{c-n/b} < |\mathsf{sup}(X)|$ as stated in the lemma. □

In the following lemma we consider an arbitrary distribution $X$ with sufficiently high min-entropy and some arbitrary function $T$. We show that if the support of $Y = T(X)$ is sufficiently large, then there exists an event $E$ such that one of the following happens:

(i) The min-entropy of $Y$ conditioned on the event $E$ is high, i.e., $Y$ conditioned on $E$ has an almost flat area with large support;
(ii) If $\overline{E}$ happens, then the average min-entropy of $X$ given $Y$ is high. Intuitively, this means that $Y$ conditioned on $\overline{E}$ has small support as then it does not "reveal" too much about $X$.

We formalize this statement in the lemma below.

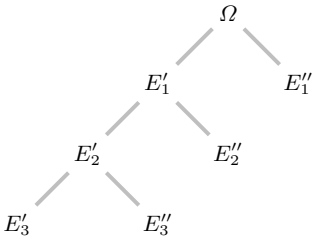
**Lemma 4.** *For $n \in \mathbb{N}_{>1}$ let $c, \alpha$ be some parameters such that $\sqrt{n} < c < \alpha \leq n$. Let $\mathcal{X}$ be some set of size $2^n = |\mathcal{X}|$ and $X$ be an $(\alpha, n)$-good distribution over $\mathcal{X}$. For any function $T : \mathcal{X} \to \mathcal{X}$, let $Y = T(X)$ be such that $|\mathsf{sup}(Y)| > 2^c$. There exists an event $E$ such that the following holds:*

*(i)* $\mathbf{H}_\infty(Y_{|E}) > c - 2\sqrt{n}$.
*(ii)* $\widetilde{\mathbf{H}}_\infty(X_{|\overline{E}}|Y_{|\overline{E}}) \geq \alpha - c - \log \frac{1}{1-\Pr[E]}$.

*Proof.* Intuitively, in the proof below we apply Lemma 3 iteratively to the distribution $Y$ to find flat areas in $Y$. We "cut off" these flat areas until we have a distribution (derived from $Y$) which has sufficiently small support. Clearly such restricted $Y$ cannot reveal too much information about $X$. To formalize this approach, we construct iteratively an event $E$ by combining the events $E_i$ obtained by applying Lemma 3 to $Y$. If $E$ happens then $Y$ takes values that lie in a large flat area. On the other hand $\overline{E}$ characterizes only a relatively small support, and hence giving such $Y$ does not reveal much information (on average) about $X$. The formal proof with an explicit calculation of the parameters follows.

We will define the event $E$ depending on events $\{E_i, E'_i, E''_i\}_{i \in \{0,\ldots,m-1\}}$ (for some integer $m$) which we will specify later. These events partition the probability space as follows (cf. Figure 1):

$$E'_i := \bigwedge_{j=0}^{i} \overline{E}_j = \overline{E}_i \wedge E'_{i-1} \qquad E''_i := E_i \wedge \left( \bigwedge_{j=0}^{i-1} \overline{E}_j \right) = E_i \wedge E'_{i-1}. \quad (5)$$



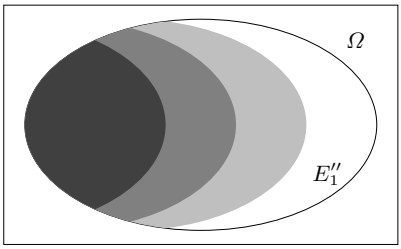




**Fig. 1.** Events covering the probability space in the proof of Lemma 4 and Lemma 5

We will rely on some properties of the above partition. In particular, note that for all $i \in \{0, \ldots, m-1\}$ we have

$$E'_i \vee E''_i = E'_{i-1} \qquad E'_i \wedge E''_i = \emptyset. \tag{6}$$

We start by constructing the events $\{E_i, E'_i, E''_i\}$ and conditional probability distributions $Y^{(i)}$ that are derived from $Y$ by applying Lemma 3. Lemma 3 requires the following two conditions:

- $|\mathsf{sup}(Y^{(i)})| > 2^c$, and
- $\Pr[Y^{(i)} = y] \geq 2^{-n}$, for all $y \in \mathsf{sup}(Y^{(i)})$.

Clearly these two conditions are satisfied by $Y^{(0)} = Y$, since $Y^{(0)}$ is computed from $X$ by applying a function $T$ and for all $x \in \mathsf{sup}(X)$ the statement assumes $\Pr[X = x] \geq 2^{-n}$. Hence, Lemma 3 gives us an event $E_0$. We set and we define $Y^{(1)} = Y^{(0)}_{|\overline{E}_0}$. For all $i \geq 1$ we proceed to construct events $E_i$ and conditional distributions $Y^{(i+1)} = Y^{(i)}_{|\overline{E}_i}$ as long as the requirements from above are satisfied. Notice that by applying Lemma 3 to distribution $Y^{(i)}$ we get for each event $E_i$:

- $\mathbf{H}_\infty(Y^{(i)}_{|E_i}) > c - 2\sqrt{n}$, and
- $|\mathsf{sup}(Y^{(i+1)})| < |\mathsf{sup}(Y^{(i)})|$.

Clearly, there are only finitely many (say $m$) events before we stop the iteration as the size of the support is strictly decreasing. At the stopping point we have $|\mathsf{sup}(Y^{(m-1)})| > 2^c$ and $|\mathsf{sup}(Y^{(m)})| \leq 2^c$. We define $E = \bigvee_{i=0}^{m-1} E_i = \bigvee_{i=0}^{m-1} E''_i$ and $\overline{E} = \bigwedge_{i=0}^{m-1} \overline{E}_i = E'_{m-1}$ and show in the claims below that they satisfy conditions (i) and (ii) of the lemma.

*Claim.* $\mathbf{H}_\infty(Y_{|E}) > c - 2\sqrt{n}$.

*Proof.* Recall that for each $0 \leq i \leq m-1$ we have

$$Y^{(i)}_{|E_i} = Y_{|E_i \wedge \overline{E}_{i-1} \ldots \wedge \overline{E}_0} \tag{7}$$

$$= Y_{|E''_i} \tag{8}$$

Eq. (7) follows from the definition of the conditional probability distribution $Y^{(i)}_{|E_i}$. Eq. (8) from the definition of the constructed events. From Eq. (8) and Lemma 3 we have for each $0 \leq i \leq m-1$ that $\mathbf{H}_\infty(Y_{|E_i''}) > c - 2\sqrt{n}$. As for each $0 \leq i \leq m-1$ we have $|\mathsf{sup}(Y_{|E}) | \geq |\mathsf{sup}(Y_{|E_i''})|$ we get that $\mathbf{H}_\infty(Y_{|E}) > c - 2\sqrt{n}$. This concludes the proof of this claim. □

*Claim.* $\widetilde{\mathbf{H}}_\infty(X_{|\overline{E}}|Y_{|\overline{E}}) \geq \alpha - c - \log \frac{1}{1-\Pr[E]}$.

*Proof.* We first lower bound $\mathbf{H}_\infty(X_{|\overline{E}})$.

$$\mathbf{H}_\infty(X_{|\overline{E}}) = -\log\left(\max_x \frac{\Pr[X = x \wedge \overline{E}]}{\Pr[\overline{E}]}\right) \tag{9}$$

$$\geq -\log\left(\frac{1}{\Pr[\overline{E}]} \max_x \Pr[X = x]\right) \tag{10}$$

$$= \mathbf{H}_\infty(X) - \log \frac{1}{\Pr[\overline{E}]} \geq \alpha - \log \frac{1}{1-\Pr[E]}. \tag{11}$$

Eq. (9) follows from the definition of min-entropy and the definition of conditional probability. Eq. (10) follows from the basic fact that for any two events $\Pr[E \wedge E'] \leq \Pr[E]$. Finally, we get Eq. (11) from our assumption that $\mathbf{H}_\infty(X) \geq \alpha$. To conclude the claim we compute:

$$\widetilde{\mathbf{H}}_\infty(X_{|\overline{E}}|Y_{|\overline{E}}) \geq \mathbf{H}_\infty(X_{|\overline{E}}, Y_{|\overline{E}}) - \log |\mathsf{sup}(Y_{|\overline{E}})| \tag{12}$$

$$= \mathbf{H}_\infty(X_{|\overline{E}}) - \log |\mathsf{sup}(Y_{|\overline{E}})| \tag{13}$$

$$\geq \alpha - \log \frac{1}{1-\Pr[E]} - c = \alpha - c - \log \frac{1}{1-\Pr[E]}. \tag{14}$$

Eq. (12) follows from Lemma 2 and (13) from the fact that $Y_{|\overline{E}}$ is computed as a function from $X_{|\overline{E}}$. Inequality (14) follows from (11) and the fact that the size of $\mathsf{sup}(Y_{|\overline{E}})$ is at most $c$. The latter follows from the definition of the event $\overline{E} = E'_{m-1}$ which in turn implies that $|\mathsf{sup}(Y_{|\overline{E}})| = |\mathsf{sup}(Y_{|E'_{m-1}})| = |\mathsf{sup}(Y^{(m-1)}_{|\overline{E}_{m-1}})| = |\mathsf{sup}(Y^{(m)})| \leq 2^c$, which concludes the proof. □

The above two claims finish the proof. □

We now turn to state and prove the Chaining Lemma.

**Lemma 5 (The Chaining Lemma).** *For $n \in \mathbb{N}_{>1}$ let $\alpha, \beta, t, \epsilon$ be some parameters where $t \in \mathbb{N}$, $0 < \alpha \leq n$, $\beta > 0$, $\epsilon \in (0, 1]$ and $t \leq \frac{\alpha-\beta}{\beta+2\sqrt{n}}$. Let $\mathcal{X}$ be some set of size $|\mathcal{X}| = 2^n$ and let $X^{(0)}$ be a $(\alpha, n)$-good distribution over $\mathcal{X}$. For $i \in [t]$ let $T_i : \mathcal{X} \to \mathcal{X}$ be arbitrary functions and $X^{(i)} = T_i(X^{(i-1)})$. There exists an event $E$ such that:*

*(i) If* $\Pr[E] > 0$*, for all* $i \in [t]$, $\mathbf{H}_\infty(X^{(i)}_{|E}) \geq \beta$.
*(ii) If* $\Pr[\overline{E}] \geq \epsilon$ *there exists an index* $j \in [t]$ *such that*

$$\widetilde{\mathbf{H}}_\infty(X^{(j-1)}_{|\overline{E}} | X^{(j)}_{|\overline{E}}) \geq \beta - \log \frac{t}{\epsilon}.$$

*Proof.* Consider the chain of random variables $X^{(0)} \xrightarrow{T_1} X^{(1)} \xrightarrow{T_2} \ldots \xrightarrow{T_t} X^{(t)}$. Given a pair of random variables in the chain, we refer to $X^{(i-1)}$ as the "source distribution" and to $X^{(i)}$ as the "target distribution". The main idea is to consider different cases depending on the characteristics of the target distribution. In case the min-entropy of $X^{(i)}$ is high enough to start with, we get immediately property (i) of the statement and we can immediately move to the next pair of random variables in the chain. In case the min-entropy of $X^{(i)}$ is small, we further consider two different sub-cases depending on some bound on the support of the variable. If the support of $X^{(i)}$ happens to be "small", intuitively we can condition on the target distribution since this cannot reveal much about the source; roughly this implies property (ii) of the statement. On the other hand, if the support happens to be not small enough, we are not in a position which allows us to condition on $X^{(i)}$.

In the latter case, we will invoke Lemma 4. Roughly this guarantees that there exists some event such that, conditioned on this event happening, the target lies in a large "flat" area and the conditional distribution has high min-entropy; this yields property (i) of the statement. If instead the event does not happen, then conditioning on the event not happening we get a "restricted" distribution with small enough support which leads again to property (ii) of the statement.

Whenever we are in those cases where (possibly conditioning on some event) the target distribution has high min-entropy, we move forward in the chain by considering $X^{(i)}$ as the source and $X^{(i+1)}$ as the target. However, when we reach a situation where we can "reveal" the target distribution we do not proceed further, since the remaining values can be computed as a deterministic function of the revealed distribution and, as such, do not constrain the min-entropy further. We now proceed with the formal proof.

Similar to Lemma 4, we will define the event $E$ depending on events $\{E_i, E'_i, E''_i\}_{i\in[t]}$ which we will specify later. These events partition the probability space as follows (cf. Figure 1):

$$E'_i := \bigwedge_{j=1}^{i} E_j = E_i \wedge E'_{i-1} \qquad E''_i := \overline{E}_i \wedge \left( \bigwedge_{j=1}^{i-1} E_j \right) = \overline{E}_i \wedge E'_{i-1}. \quad (15)$$

We will rely on some properties of the above partition. In particular, note that for all $i \in [t]$ we have

$$E'_i \vee E''_i = E'_{i-1} \qquad E'_i \wedge E''_i = \emptyset. \quad (16)$$

For all $i \in [t+1]$, define the following parameters:

$$s_i = (t - i + 1)(\beta + 2\sqrt{n}) \quad (17)$$

$$\alpha_{i-1} = \beta + s_i. \quad (18)$$

Note that using the bound on $t$ from the statement of the lemma, we get $\alpha \geq \alpha_0$; moreover, it is easy to verify that $\alpha_{i-1} > s_i > \sqrt{n}$ for all $i \in [t]$.

In the next claim we construct the events $\{E_i, E'_i, E''_i\}_{i\in[t]}$.

*Claim.* For all $i = 0, \ldots, t-1$, there exist events $E'_{i+1}$ and $E''_{i+1}$ (as given in Eq. (16)) such that the following hold:

(*) If $\Pr\left[E'_{i+1}\right] > 0$, $\mathbf{H}_\infty(X^{(i+1)}_{|E'_{i+1}}) \geq \alpha_{i+1}$.

(**) If $\Pr\left[E''_{i+1}\right] \geq \epsilon'$, $\widetilde{\mathbf{H}}_\infty(X^{(i)}_{|E''_{i+1}} | X^{(i+1)}_{|E''_{i+1}}) \geq \beta - \log\frac{1}{\epsilon'}$. where $0 < \epsilon' \leq 1$.

*Proof.* We prove the claim by induction.

*Base Case:* In this case we let $E_0$ denote the whole probability space and thus $\Pr[E_0] = 1$. Note that $\mathbf{H}_\infty(X^{(0)}_{|E_0}) = \mathbf{H}_\infty(X^{(0)}) = \alpha \geq \alpha_0$. The rest of the proof for the base case is almost the same to that of the inductive step except the use of the above property instead of the induction hypothesis. Therefore we only prove the induction step in detail here. The proof details for the base case are a straightforward adaptation, with some notational changes.

*Induction Step:* The following holds by the *induction hypothesis*:

(*) If $\Pr[E'_i] > 0$, then $\mathbf{H}_\infty(X^{(i)}_{|E'_i}) \geq \alpha_i$.

(**) If $\Pr[E''_i] \geq \epsilon'$ then, $\widetilde{\mathbf{H}}_\infty(X^{(i-1)}_{|E''_i} | X^{(i)}_{|E''_i}) \geq \beta - \log\frac{1}{\epsilon'}$ where $0 < \epsilon' \leq 1$.

By construction of the events, $E'_i$ is partitioned into two sub-events $E'_{i+1}$ and $E''_{i+1}$ (cf. Eq. 16). From the statement of the claim we observe that, since we are assuming $\Pr\left[E'_{i+1}\right] > 0$ in (*) and $\Pr\left[E''_{i+1}\right] \geq \epsilon' > 0$ in (**), in both cases we have $\Pr[E'_i] > 0$. Hence, property (*) from the induction hypothesis holds: $\mathbf{H}_\infty(X^{(i)}_{|E'_i}) \geq \alpha_i$, which we use to prove the inductive step. We will define the events $E'_{i+1}$ and $E''_{i+1}$ differently depending on several (complete) cases. For each of these cases we will show that property (*) and (**) hold.

Suppose first that $\mathbf{H}_\infty(X^{(i+1)}_{|E'_i}) \geq \alpha_{i+1}$. In this case we define $E'_{i+1}$ to be $E'_i$, which implies $E''_{i+1} = \emptyset$ by Eq. (16). Moreover property (*) holds since, if $\Pr\left[E'_{i+1}\right] > 0$, then $\Pr[E'_i] > 0$ and $\mathbf{H}_\infty(X^{(i+1)}_{|E'_{i+1}}) = \mathbf{H}_\infty(X^{(i+1)}_{|E'_i}) \geq \alpha_{i+1}$; as for property (**) there is nothing to prove, since $\Pr\left[E''_{i+1}\right] = 0$ in this case.

Consider now the case that $\mathbf{H}_\infty(X^{(i+1)}_{|E'_i}) < \alpha_{i+1}$. Here we consider two subcases, depending on the support size of $X^{(i+1)}$.

1. $|\mathsf{sup}(X^{(i+1)}_{|E'_i})| \leq 2^{s_{i+1}}$. We define $E''_{i+1} = E'_i$, which implies $E'_{i+1} = \emptyset$ by Eq. (16). As for property (*) there is nothing to prove, since $\Pr\left[E'_{i+1}\right] = 0$. To prove property (**) we observe that if $\Pr\left[E''_{i+1}\right] \geq \epsilon' > 0$, then

$\Pr[E'_i] > 0$. Hence,

$$\widetilde{\mathbf{H}}_\infty(X^{(i)}_{|E''_{i+1}} | X^{(i+1)}_{|E''_{i+1}}) = \widetilde{\mathbf{H}}_\infty(X^{(i)}_{|E'_i} | X^{(i+1)}_{|E'_i}) \tag{19}$$

$$\geq \mathbf{H}_\infty(X^{(i)}_{|E'_i}, X^{(i+1)}_{|E'_i}) - \log(|\mathsf{sup}(X^{(i+1)}_{|E'_i})|) \tag{20}$$

$$\geq \alpha_i - s_{i+1} \tag{21}$$

$$= \beta + s_{i+1} - s_{i+1} = \beta.$$

Eq. (19) follows as $E''_{i+1} = E'_i$. Eq. (20) follows from Lemma 2. Eq. (21) follows from two facts: (i) $X^{(i+1)}$ is a deterministic function of $X^{(i)}$, which means $\mathbf{H}_\infty(X^{(i)}_{|E'_i}, X^{(i+1)}_{|E'_i}) = \mathbf{H}_\infty(X^{(i)}_{|E'_i}) \geq \alpha_i$ (plugging-in the value from induction hypothesis), and (ii) $|\mathsf{sup}(X^{(i+1)}_{|E'_i})| \leq 2^{s_{i+1}}$.

2. $|\mathsf{sup}(X^{(i+1)}_{|E'_i})| > 2^{s_{i+1}}$. By the induction hypothesis $\mathbf{H}_\infty(X^{(i)}_{|E'_i}) \geq \alpha_i$; we now invoke Lemma 4 on the distribution $X^{(i+1)}_{|E'_i}$ (recall that $\alpha_i > s_{i+1} > \sqrt{n}$), to obtain the event $E_{i+1}$ such that:

$$\mathbf{H}_\infty(X^{(i+1)}_{|E'_i \wedge E_{i+1}}) > s_{i+1} - 2\sqrt{n} \tag{22}$$

$$\widetilde{\mathbf{H}}_\infty(X^{(i)}_{|E'_i \wedge \overline{E}_{i+1}} | X^{(i+1)}_{|E'_i \wedge \overline{E}_{i+1}}) > \alpha_i - s_{i+1} - \log \frac{1}{1 - \Pr[E_{i+1}]}. \tag{23}$$

Note that by our definitions of the events $E'_i, E''_i$ (cf. Eq. (15)), we have $E'_i \wedge E_{i+1} = E'_{i+1}$ and $E'_i \wedge \overline{E}_{i+1} = E''_{i+1}$.
To prove (*) we consider that if $\Pr\left[E'_{i+1}\right] > 0$, then $\Pr[E'_i] > 0$ and $\Pr[E_{i+1}] > 0$. Plugging the values of $\alpha_i$ and $s_{i+1}$ from Eq. (18) and (17) into Eq. (22), we get

$$\begin{aligned}
\mathbf{H}_\infty(X^{(i+1)}_{|E'_{i+1}}) &> s_{i+1} - 2\sqrt{n} \\
&= (t-i)(\beta + 2\sqrt{n}) - 2\sqrt{n} \\
&= \beta + (t-i-1)(\beta + 2\sqrt{n}) \\
&= \beta + s_{i+2} = \alpha_{i+1},
\end{aligned}$$

Similarly, to prove (**), we consider that if $\Pr\left[E''_{i+1}\right] \geq \epsilon'$, then $\Pr[E'_i] \geq \epsilon' > 0$ and $\Pr\left[\overline{E}_{i+1}\right] \geq \epsilon'$. Using Eq. (23), we obtain:

$$\begin{aligned}
\widetilde{\mathbf{H}}_\infty(X^{(i)}_{|E''_{i+1}} | X^{(i+1)}_{|E''_{i+1}}) &> \alpha_i - s_{i+1} - \log \frac{1}{\Pr\left[\overline{E}_{i+1}\right]} \\
&= \beta - \log \frac{1}{\Pr\left[\overline{E}_{i+1}\right]} \\
&\geq \beta - \log \frac{1}{\epsilon'},
\end{aligned}$$

This concludes the proof of the claim. □

We define the event $E$ to be $E = E'_t = \bigwedge_{i=1}^{t} E_i = \bigwedge_{i=1}^{t} E'_i$. It is easy to verify that this implies $\overline{E} = \bigvee_{i=1}^{t} E''_i$. We distinguish two cases:

- If $\Pr[E] > 0$, by definition of $E$ we get that $\Pr[E'_i] > 0$ for all $i \in [t]$. In particular, $\Pr[E'_t] > 0$. Hence, $\mathbf{H}_\infty(X^{(t)}_{|E}) = \mathbf{H}_\infty(X^{(t)}_{|E'_t}) \geq \alpha_t = \beta$, where the last inequality follows from property (*) of the above Claim, using $i = t-1$. Also, we observe that for all $i \in [t]$, $\mathbf{H}_\infty(X^{(i-1)}_{|E}) \geq \mathbf{H}_\infty(X^{(i)}_{|E})$. This proves property (i) of the lemma.
- If $\Pr\left[\overline{E}\right] \geq \epsilon$, then we get

$$\Pr\left[\bigvee_{i=1}^{t} E''_i\right] \geq \epsilon. \tag{24}$$

$$\sum_{i=1}^{t} \Pr[E''_i] \geq \epsilon. \tag{25}$$

Eq. (24) follows from the definition of $E$ and Eq. (25) follows applying union bound. Clearly, from Eq. (25), there must exists some $j$ such that $\Pr\left[E''_j\right] \geq \epsilon/t$.

Hence, putting $i = j-1$ and $\epsilon' = \epsilon/t$ in property (**) of the above Claim, we get:

$$\widetilde{\mathbf{H}}_\infty(X^{(j-1)}_{|E''_j} | X^{(j)}_{|E''_j}) \geq \beta - \log\frac{t}{\epsilon}.$$

From the definition of $E$, $E''_j$ implies $\overline{E}$ and hence property (ii) of the lemma follows.

□

## 4 Application to Tamper-Resilient Cryptography

We show that *any* cryptographic primitive where the secret key can be chosen as a uniformly random string can be made secure in the BLT model of [20] by a simple and efficient transformation. Our result therefore covers pseudorandom functions, block ciphers, and many encryption and signature schemes. However, the result holds in a restricted model of tampering: the adversary first selects an arbitrary set of tampering functions of bounded size and, as he interacts with the scheme, he must choose every tampering function from the set that was specified initially. We call this the *semi-adaptive* BLT model. Our result holds only when the set of functions is "small enough".[3]

The basic intuition behind the construction using the Chaining Lemma is easy to explain. We use a random string $X_0$ as secret key, and a universal hash function $h$ as public (and tamper proof) parameter. The construction then computes

[3] In particular, the adversary can choose a "short enough" sequence of tampering functions, from a set containing polynomially many such sequences.

$K_0 = h(X_0)$, and uses $K_0$ as secret key for the original primitive. The intuitive reason why one might hope this would work is as follows: each tampering query changes the key, so we get a chain of keys $X_0, X_1, \ldots, X_t$ where $X_i = T_i(X_{i-1})$ for some tampering function $T_i$. Recall that the chaining lemma guarantees that for such a chain, there exists an event $E$ such that: (i) when $E$ takes place then all $X_i$ have high min-entropy, and, by a suitable choice of $h$, all the hash values $K_0 = h(X_0), K_1 = h(X_1), \ldots, K_t = h(X_t)$ are statistically close to uniformly and independently chosen keys; (ii) when $E$ does not happen, for some index $j \in [t]$ we are able to reveal the value of $X_j$ to the adversary as the $X_i$'s with $i < j$ still have high entropy, and hence hash to independent values. On the other hand the $X_i$'s with $i \geq j$ are a deterministic function of $X_j$ and hence the tampering queries corresponding to any subsequent key can be simulated easily.

Due to its generality the above result suffers from two limitations. First, as already mentioned above, the tampering has to satisfy a somewhat limited form of adaptivity. Second, the number of tampering queries one can tolerate is upper bounded by the length $n$ of the secret key. While this is true in general for schemes without key update, for our general result the limitation is rather strong. More concretely, with appropriately chosen parameters our transformation yields schemes that can tolerate up to $O(\sqrt[3]{n})$ tampering queries. We discuss the application in full detail in the full version of this paper [21].

*Comparison with Faust* et al. *[26].* Very recently, Faust *et al.* [26] introduced the concept of non-malleable key derivation which is similar in spirit to our application of the Chaining Lemma. Intuitively a function $h$ is a non-malleable key derivation function if $h(X)$ is close to uniform even given the output of $h$ applied to a related input $T(X)$, as long as $T(X) \neq X$. They show that a random $t$-wise independent hash function already meets this property, and moreover that such a function can be used to protect arbitrary cryptographic schemes (with a uniform key) against "one-time" tampering attacks (i.e., the adversary is allowed a single tampering query) albeit against a much bigger class of functions.[4]

We stress that the novelty of our result is in discovering the Chaining Lemma rather than this application, which can be instead thought of as a new technique, fundamentally different from that of [26], to achieve security in the BLT model. We believe that the Chaining Lemma is interesting in its own right, and might find more applications in cryptography in the future.

## References

1. Aggarwal, D., Dodis, Y., Lovett, S.: Non-malleable codes from additive combinatorics. Electronic Colloquium on Computational Complexity (ECCC) 20, 81 (2013), To appear in STOC 2014
2. Agrawal, S., Gupta, D., Maji, H.K., Pandey, O., Prabhakaran, M.: Explicit non-malleable codes resistant to permutations. IACR Cryptology ePrint Archive, 2014:316 (2014)

[4] It might be possible to extend the analysis of [26] to bounded tampering, but this seems not straightforward.

3. Agrawal, S., Gupta, D., Maji, H.K., Pandey, O., Prabhakaran, M.: Explicit non-malleable codes resistant to permutations and perturbations. IACR Cryptology ePrint Archive, 2014:841 (2014)
4. Agrawal, S., Gupta, D., Maji, H.K., Pandey, O., Prabhakaran, M.: Explicit optimal-rate non-malleable codes against bit-wise tampering and permutations. IACR Cryptology ePrint Archive, 2014:842 (2014)
5. Anderson, R., Kuhn, M.: Tamper resistance: A cautionary note. In: WOEC 1996: Proceedings of the 2nd Conference on Proceedings of the Second USENIX Workshop on Electronic Commerce, pp. 1–1. USENIX Association, Berkeley (1996)
6. Applebaum, B., Harnik, D., Ishai, Y.: Semantic security under related-key attacks and applications. In: ICS, pp. 45–60 (2011)
7. Bellare, M., Cash, D.: Pseudorandom functions and permutations provably secure against related-key attacks. In: Rabin, T. (ed.) CRYPTO 2010. LNCS, vol. 6223, pp. 666–684. Springer, Heidelberg (2010)
8. Bellare, M., Cash, D., Miller, R.: Cryptography secure against related-key attacks and tampering. In: Lee, D.H., Wang, X. (eds.) ASIACRYPT 2011. LNCS, vol. 7073, pp. 486–503. Springer, Heidelberg (2011)
9. Bellare, M., Kohno, T.: A theoretical treatment of related-key attacks: RkA-PRPs, RkA-PRFs, and applications. In: Biham, E. (ed.) EUROCRYPT 2003. LNCS, vol. 2656, pp. 491–506. Springer, Heidelberg (2003)
10. Bellare, M., Paterson, K.G., Thomson, S.: RKA security beyond the linear barrier: IBE, encryption and signatures. In: Wang, X., Sako, K. (eds.) ASIACRYPT 2012. LNCS, vol. 7658, pp. 331–348. Springer, Heidelberg (2012)
11. Bhattacharyya, R., Roy, A.: Secure message authentication against related key attack. In: Moriai, S. (ed.) FSE 2013. LNCS, vol. 8424, pp. 305–324. Springer, Heidelberg (2014)
12. Boneh, D., DeMillo, R.A., Lipton, R.J.: On the importance of eliminating errors in cryptographic computations. J. Cryptology 14(2), 101–119 (2001)
13. Cheraghchi, M., Guruswami, V.: Capacity of non-malleable codes. In: ITCS, pp. 155–168 (2014)
14. Cheraghchi, M., Guruswami, V.: Non-malleable coding against bit-wise and split-state tampering. In: Lindell, Y. (ed.) TCC 2014. LNCS, vol. 8349, pp. 440–464. Springer, Heidelberg (2014)
15. Choi, S.G., Kiayias, A., Malkin, T.: BiTR: Built-in tamper resilience. In: Lee, D.H., Wang, X. (eds.) ASIACRYPT 2011. LNCS, vol. 7073, pp. 740–758. Springer, Heidelberg (2011)
16. Coretti, S., Dodis, Y., Tackmann, B., Venturi, D.: Self-destruct non-malleability. IACR Cryptology ePrint Archive, 2014:866 (2014)
17. Coretti, S., Maurer, U., Tackmann, B., Venturi, D.: From single-bit to multi-bit public-key encryption via non-malleable codes. IACR Cryptology ePrint Archive, 2014:324 (2014)
18. Dachman-Soled, D., Kalai, Y.T.: Securing circuits against constant-rate tampering. In: Safavi-Naini, R., Canetti, R. (eds.) CRYPTO 2012. LNCS, vol. 7417, pp. 533–551. Springer, Heidelberg (2012)
19. Dachman-Soled, D., Kalai, Y.T.: Securing circuits and protocols against $1/\mathrm{poly}(k)$ tampering rate. In: Lindell, Y. (ed.) TCC 2014. LNCS, vol. 8349, pp. 540–565. Springer, Heidelberg (2014)
20. Damgård, I., Faust, S., Mukherjee, P., Venturi, D.: Bounded tamper resilience: How to go beyond the algebraic barrier. In: Sako, K., Sarkar, P. (eds.) ASIACRYPT 2013, Part II. LNCS, vol. 8270, pp. 140–160. Springer, Heidelberg (2013)

21. Damgård, I., Faust, S., Mukherjee, P., Venturi, D.: The chaining lemma and its application. IACR Cryptology ePrint Archive, 2014:979 (2014)
22. Dodis, Y., Ostrovsky, R., Reyzin, L., Smith, A.: Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. SIAM J. Comput. 38(1), 97–139 (2008)
23. Dziembowski, S., Pietrzak, K., Wichs, D.: Non-malleable codes. In: ICS, pp. 434–452 (2010)
24. Faust, S., Mukherjee, P., Nielsen, J.B., Venturi, D.: Continuous non-malleable codes. In: Lindell, Y. (ed.) TCC 2014. LNCS, vol. 8349, pp. 465–488. Springer, Heidelberg (2014)
25. Faust, S., Mukherjee, P., Nielsen, J.B., Venturi, D.: A tamper and leakage resilient von Neumann architecture. IACR Cryptology ePrint Archive, 2014:338 (2014)
26. Faust, S., Mukherjee, P., Venturi, D., Wichs, D.: Efficient non-malleable codes and key-derivation for poly-size tampering circuits. In: Nguyen, P.Q., Oswald, E. (eds.) EUROCRYPT 2014. LNCS, vol. 8441, pp. 111–128. Springer, Heidelberg (2014)
27. Faust, S., Pietrzak, K., Venturi, D.: Tamper-proof circuits: How to trade leakage for tamper-resilience. In: Aceto, L., Henzinger, M., Sgall, J. (eds.) ICALP 2011, Part I. LNCS, vol. 6755, pp. 391–402. Springer, Heidelberg (2011)
28. Gennaro, R., Lysyanskaya, A., Malkin, T., Micali, S., Rabin, T.: Algorithmic tamper-proof (ATP) security: Theoretical foundations for security against hardware tampering. In: Naor, M. (ed.) TCC 2004. LNCS, vol. 2951, pp. 258–277. Springer, Heidelberg (2004)
29. Goyal, V., O'Neill, A., Rao, V.: Correlated-input secure hash functions. In: Ishai, Y. (ed.) TCC 2011. LNCS, vol. 6597, pp. 182–200. Springer, Heidelberg (2011)
30. Ishai, Y., Prabhakaran, M., Sahai, A., Wagner, D.: Private circuits II: Keeping secrets in tamperable circuits. In: Vaudenay, S. (ed.) EUROCRYPT 2006. LNCS, vol. 4004, pp. 308–327. Springer, Heidelberg (2006)
31. Jafargholi, Z., Wichs, D.: Tamper detection and continuous non-malleable codes. Cryptology ePrint Archive, Report 2014/956 (2014), `http://eprint.iacr.org/`
32. Kalai, Y.T., Kanukurthi, B., Sahai, A.: Cryptography with tamperable and leaky memory. In: Rogaway, P. (ed.) CRYPTO 2011. LNCS, vol. 6841, pp. 373–390. Springer, Heidelberg (2011)
33. Lucks, S.: Ciphers secure against related-key attacks. In: Roy, B., Meier, W. (eds.) FSE 2004. LNCS, vol. 3017, pp. 359–370. Springer, Heidelberg (2004)
34. Pietrzak, K.: Subspace LWE. In: Cramer, R. (ed.) TCC 2012. LNCS, vol. 7194, pp. 548–563. Springer, Heidelberg (2012)
35. Qin, B., Liu, S., Yuen, T.H., Deng, R.H., Chen, K.: Continuous non-malleable key derivation and its application to related-key security. Cryptology ePrint Archive, Report 2015/003 (2015), `http://eprint.iacr.org/`
36. Wee, H.: Public key encryption against related key attacks. In: Fischlin, M., Buchmann, J., Manulis, M. (eds.) PKC 2012. LNCS, vol. 7293, pp. 262–279. Springer, Heidelberg (2012)