# Universal Encoding for Provably Irreversible Data Erasing

Marek Klonowski[1], Tomasz Strumiński[2] and Małgorzata Sulkowska[1]
[1]*Faculty of Fundamental Problems of Technology, Department of Computer Science, Wrocław University of Science and Technology, Poland*
[2]*BrightIT, Wrocław, Poland*

Keywords: Data Deletion, Provable Security, Formal Analysis.

Abstract: One of the most important assumptions in computer security research is that one can permanently delete some data in such a way that no party can retrieve it. In real-life systems this postulate is realized dependently on the specific device used for storing data. In some cases (e.g., magnetic discs) the deletion/erasing is done by overwriting the data to be erased by new one. Many evidence suggest that such procedure may be not sufficient and the attacker armed with advanced microscopic technology is capable in many cases of retrieving data overwritten even many times. In this paper we present a method that provides **provable**, permanent and irreversible deletion of stored bits based solely on special encoding and processing of data. More precisely the adversary learns nothing about deleted data whp. The security guarantees hold even if the attacker is capable of getting bit-strings overwritten many times. Moreover, in contrast to some previous research, we do not restrict type of data to be deleted.

## 1 INTRODUCTION

In many security related research it is often (silently) assumed that any data can be deleted on demand. That is, one can perform an action such that locally stored data is instantly removed and no one can learn it anymore. In many systems such deletion is in fact realized in a way that some pointer representing physical region of memory is marked as ready for being reused. Clearly, data deleted in such a way can be easily accessed using a special software as long as the respective memory region has not been used yet. Thus, more aware users aiming at irreversible deletion write some (possibly) meaningless data on the critical region.

In the case of magnetic discs it has been quickly noticed that overwriting the memory to be deleted only once may be not sufficient, since the adversary having access to more advanced techniques (mainly microscopic) is capable of retrieving the original data ((Gomez R. D. et al., 1993; Gutmann Peter, 1996; Hughes G. et al., 2009; Mayergoyz I. D. et al., 2001)). As a consequence some more elaborated methods of removing data has been suggested in (Gutmann Peter, 1996). Bulk of them are heuristics based on the idea of overwriting region of the memory using a special, usually alternating, patterns. The presented ideas seem to be convincing and moreover in some cases authors present some experimental examples showing that after applying their methods retrieving data is not possible given more or less advanced techniques. However there is no formal proof that such approach really works. Moreover, judging by security discussions in (Gutmann Peter, 1996; Hughes G. et al., 2009; US Department of Defense, 1997; U.S. National Institute of Standards and Technology, 2006) we cannot be sure what is the minimal sufficient number of layers overwriting the original one to assure irreversibility of deletion.

To explain the nature of the problem let us recall how data (or a single bit) is represented on the magnetic disc. For contemporary discs the width of the path representing consecutive bits is approximately $150-200$ nm, moreover, there are separating *guard-bands* of width of 20 nm. The overwritten data can be revealed because the new bit is not written to the same physical location as the previous one. That is, the physical mark representing a particular bit can be slightly moved from its expected position whenever we overwrite an old bit. The overview of this situation is presented in Figure 1. This enables the adversary to get **in some cases** the former, overwritten bit if only he has access to a sufficiently sensitive device.

The physical mark has to be placed on a given po-



Klonowski, M., Strumiński, T. and Sulkowska, M.
Universal Encoding for Provably Irreversible Data Erasing.
DOI: 10.5220/0007922101370148
In *Proceedings of the 16th International Joint Conference on e-Business and Telecommunications (ICETE 2019)*, pages 137-148
ISBN: 978-989-758-378-0

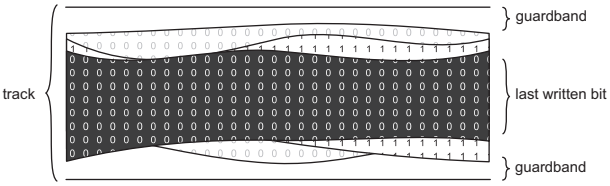


Figure 1: Picture of a magnetic disc: black path represents the current data (bit 0), while the white/grey the previous, overwritten data 1 and 0.

sition in an extremely short time - therefore some inaccuracies appear. This leads to possibility of retrieving some data using technologies like MFM (*magnetic force microscopy*) that is an advanced form of SPM (*scanning probe microscopy*, SPM) and makes it possible to localize even extremely small magnetized scrap of the disc. As pointed in (Gomez R. D. et al., 1993; Rugar D. et al., 1990) in many real-life examples MFM is enough to read the data with sufficient accuracy. There are even more sophisticated methods like *magnetic force scanning tunneling microscopy*, MTM and *spin-stand imaging* (Gutmann Peter, 1996). Moreover one can expect that in foreseeable future the adversary may have access to even stronger diagnostic devices and use it against today's storage devices.

To sum up,

- one cannot be sure if someone could try to retrieve deleted data from today's devices in the future;
- due to dynamic progress in magnetic microscopy techniques one cannot predict what devices can be used to retrieve data in the future.

This does not mean that every single "overwritten" bit can be retrieved after arbitrary long time. It means however that **some** bits can be retrieved after long, unpredictable time and there is still a risk of unwanted and unexpected information leakage. Let us also stress that if a single bit can be retrieved then it is more likely that the neighboring bits can be retrieved as well. Indeed, one may suspect that inaccuracies in positioning of magnetic heads for close places can be correlated. For that reason it is possible that the adversary could get a meaningful part of contiguous bits instead of an isolated (and probably useless) bits. For these reasons the only fully reliable method suggested at hand is the physical destruction of the (magnetic) device. Destroying the devices is not acceptable, however, in most of areas of application.

Considered problem may be also solved using disk encryption. One can always write sensitive data in an encrypted form to the disk and store the appropriate key on, say, cryptographic chip. However, in our approach we want to assume that no additional devices (like chips) are necessary. We suggest an approach to data deletion based solely on a special data encoding that guarantees provable security even in the presence of an adversary capable of recognizing all physical marks (i.e, all history) that were written in a given place of the disc in the past. Note that it is a very strong assumption - indeed, the adversary may learn the data overwritten arbitrary number of times. At first glance it seems that one cannot hide any information from such a strong adversary. We show that it is not true if we apply a special encoding and deletion procedures.

The proposed methods is not restricted to magnetic discs and can be applied to different types of data storage devices. Namely our analysis of security is valid as long as some requirements described in Section 2.1 are fulfilled. That is, our method is based on a modified data representation/organisation and potentially can be applied even in a future-generation devices. In the case of wide class of standard devices just by replacing the firmware.

**On the downside** our method leads to moderate space overhead. More precisely the same device using new encoding can represent less bits (say, two orders of magnitude) when compared to the standard way of using this device. Moreover the data processing (writing and reading stored bits) may be one order of magnitude slower. Let us also note that in some publications authors claim that the investigated concern is overblown. Namely, they suggest that in recently produced (extremely dense) magnetic data storage devices data overwritten even once cannot be retrieved in practice ( Gutmann P., 1996).

In summary, our method is rather not intended for deleting all possible data. We imagine that the proposed technique can be useful in the systems wherein particularly sensitive data is processed but the users can accept a longer processing time and loss of some space. At this price we get a method of removing the data that is **provably secure** even in the presence of a very strong adversary. That is, we imagine that the presented methods can be used for example for data that needs to be particularly secured (e.g., cryptographic material - including secret keys or seeds for pseudo random number generators; data bases with sensitive personal information; descriptions of new, original industrial technologies etc.).

## 1.1 Related Work

There are many papers related to our results. In (Gomez R. D. et al., 1993; Gutmann Peter, 1996; Hughes G. et al., 2009; Mayergoyz I. D. et al., 2001) one can find information about microscopic techniques and properties of magnetic discs in the context

of data retrieval by a physical inspection.

Some practical methods of permanent data deletion can be found in papers (Gutmann Peter, 1996; Hughes G. et al., 2009; US Department of Defense, 1997; U.S. National Institute of Standards and Technology, 2006). All that papers assume overwriting the original data several times using different patterns. The only papers presenting formal analysis of data deletion security we are aware of, are (Klonowski et al., 2008) and (Klonowski et al., 2009) wherein the encoding of the type presented below was introduced (note however that the deletion method is substantially different). Those papers do not present fully formal analysis and are limited to deletion of the data from a very narrow type of distribution in contrast to this paper wherein we present **universal methods** that can be used for **arbitrary** types of data. Moreover some of our methods are much faster while some other require optimal number of operations for getting demanded level of security. In the mentioned papers also a very strong adversary was considered, however the analysis was completely different. Namely the security was considered only for hiding purely random bit-string (that can be a good model of cryptographic material including seeds for pseudo-random bits or secret keys).

In the current paper we analyze security of wiping of much wider classes of bit-strings to be securely removed (including the case of data to be deleted from an arbitrary source). We also use a different security measure that seem to be more adequate for real-life systems based on *differential privacy* notion introduced by Dwork et al. in (Dwork et al., 2006a; Dwork, 2006). Using these definitions gives us strong security guarantees and repels so-called *linkage attacks* as well as some other useful properties including immunity against data post-processing (see e.g (Cynthia Dwork and Aaron Roth, 2013)).

Let us note that there is a well-developed body of recent papers about efficient data deletion that can be seen as somehow related results. They are assuming, however, substantially different models of adversary's acting and data storing methods.

Many recent papers discuss the problem of data deletion in a distributed system (especially in clouds), wherein each piece of data can be stored in several copies and may be processed by various entities (e.g. (Hur J. et al., 2017; Wegberg G. et al., 2017; Ali M. et al., 2017; Bacis E. et al., 2016)). There are also significant recent results about publicly verifiable data deletion in multi-user systems (Yang C. et al., 2018; Ali M. et al., 2017; Hao F. et al., 2016). In contrast to our approach, all that paper are mainly based on cryptographic techniques and assume that the users have access to secure devices, such that the adversary has no access to overwritten data.

Let us stress that the considered problem is substantially different than ORAM (Oblivious Ram) introduced in (Goldreich and Ostrovsky, 1996) which aims at obscuring the operations being performed (read/write) in the past. This problem is orthogonal to our considerations. That is, each ORAM we are aware of, assumes that the adversary has access to the **current state** of memory, only. The same refers to the recent ORAM-related papers discussing distributed deletion (e.g. (Roche et al., 2016).

Finally let us mention also some effort in constructing devices to allowing mitigation of retrieving deleted data or other unexpected information leakages ((Jia et al., 2016; Moritz C.A. et al., 2015)) . The presented method is nevertheless orthogonal to the approach presented in our paper, in particular, do not assume such a strong adversary.

While preparing this paper we were inspired by two other papers not related directly to data deletion - the first one is (Rivest and Shamir, 1982) wherein authors present how a write-once memory (the state representing bit 0 can be changed into bit 1 but the inverse operation is not feasible) can be to some extent re-used by using a special encoding. In paper (Moran T. et al., 2009) authors present deterministic method of storing results of voting preserving privacy of individual voters even in the presence of extremely strong adversary.

### 1.2 Organization of this Paper

In Section 2 we present principles of a special encoding extensively used later. After that we present and justify the assumed mathematical model of the data storage device. Introducing such model allows us to abstract from all physical properties in the security analysis. Finally, we briefly describe how deletion methods work. In Section 3 we begin with analysis of our deletion methods for different types of data. In Section 4 we present and analyze a modified protocol with very fast (optimal) execution time. In Section 5 we conclude and present some future work.

## 2 MODEL AND ENCODING

In this section we describe our methods - we refer to magnetic discs however it can be also used for any other data storing devices as long as they meet assumptions of the presented model. First, let us assume that every piece of the disc can be marked with one of

two states representing 0 or 1. We assume the following four-phase life cycle of a data storage device.

**Preliminary Phase -** this phase covers all actions (marking initial states representing 0 or 1) performed before the storage device is given to the user for storing data; this phase can be performed by regular user or even the manufacturer.

**Regular Usage Phase -** we assume that in this phase the user is able to write (i.e. re-use the space) an arbitrary number of times, always is able to read the data written last time and knows an upper bound for the number of changes she introduced.

**Deletion Phase -** in a given moment the user is asked to perform some actions to make reading the disc impossible.

**Adversarial Inspection Phase -** the adversary is given the disc and techniques to correctly say what states (0 or 1) have been marked in each piece of the disc. For example in the case of the disc depicted in Figure 1 the adversary can learn that the state represents 0 overwriting previous state 1. Moreover the adversary sees that the previous state 1 covers the initial state 0. Clearly, such model is extremely strong.

## 2.1 Special Data Representation

The idea of coding is based on dividing physical space of a magnetic disk into rectangle-like areas that we call *boxes*. Depending on the implementation, each box may contain the space of several to several dozens of regular bits. In our coding each box represents a single logical bit in the new encoding. Each box is divided into two subspaces. The first, inner one, represents the value of the box (which is 0 or 1) and the outer, surrounding the inner one, plays the role of a border and always consists of the state representing 0. This idea is depicted in Figure 3 and compared to regular encoding in Figure 2. A box that in our coding is represented by 1, has got the state representing 1 in the inner part. Analogically a box representing 0 has 0 state in the inner part. Since outer part of the box has always got only 0s, then box representing 0 contains only the state representing 0. To avoid ambiguity of notation we shall call a box that represents 0 a 0-box and a box representing 1 a 1-box.

To change the value of a bit one just checks the value of a box and changes the inner part. We assume that in each box the adversary can learn the changes of the value of the box (i.e. state of the inner part) but **cannot learn anything about one box from another boxes.** This *separation* property is guaranteed by sufficiently large separation area. That is, we assume that

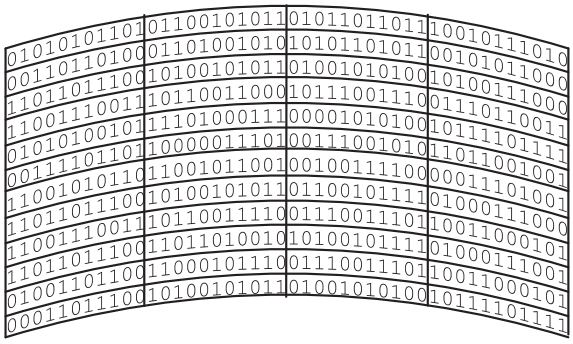

Figure 2: Physical representation of standard layout.

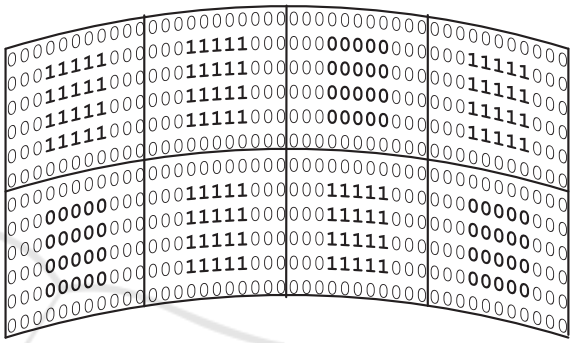

Figure 3: The same disc in the new encoding. It is divided into boxes representing bits $1,1,0,1$ (upper row), $0,1,1,0$ (lower row).

even maximal feasible misalignment during changing the state of one box that may occur does not influence other boxes. This is the only assumption related to the physical nature of the device. The rest is based on special encoding.

## 2.2 Encoding

Each box represents a single logical bit. There are two rules of encoding:

- at the beginning all boxes are set to 0;
- the state of the box is changed only from 0 to 1 or from 1 to 0.

The second rule implies that the algorithm before storing new bit has to check if such operation is necessary. If the current state of a given box is the same as the bit to be stored the algorithm does nothing.

All algorithms in this model work as follows. In the Preliminary Phase the new disc representing no information is filled with 0-boxes. Then possibly each box can be replaced some number of times from 0 to 1 and from 1 to 0 a number of times. Note that each box can experience a different number of such changes.

In Regular Usage Phase the user changes states of boxes to represent the data she needs to store. Read-

ing the data is just checking the inner part of respective box.

In the Deletion Phase each representation of each box is changed a number of times (from 0 to 1, from 1 to 0, and so on). Note that the number of changes **cannot** depend on the number of changes in previous phases. That is, the user (in contrast to the adversary) has no access to overwritten data.

Let us note that such model is simplified and deviates from the current methods of data processing. Namely in the current systems the following holds:

- data buffering – in not a single bits yet a bunches of bits at once,
- data encoding – in real life systems data are encoded before they are written (using for example: MFM, RLL, PRLM or EPRLM encoding (Gutmann Peter, 1996)).

The difference is not negligible, however it is clear that the model can be realized by changing the firmware, only. The price the user has to pay is limited space and slower data processing.

Le us note that this encoding method is based on idea from (Klonowski et al., 2008), however the algorithms presented below are different. Moreover we provide here a formal proves for security declared properties according to a stronger definition.

### 2.3 Adversarial View

We assume that the adversary has an access to all "layers" - that is, for each box it can recognize all bits represented by this box in the past. Due to the assumed encoding the only knowledge the adversary has, is the number of changes. Thus the whole disc can be represented as a vector $(y_1, y_2, \ldots, y_d)$ where $y_i$ is the number of changes in the $i$-th box.

Clearly the assumption that having good microscope one can retrieve the data in every single place is not realistic but one can agree that it is an upper bound for capabilities of any real adversary.

Note that the knowledge of the regular user is significantly smaller - that is, she knows some upper bound for the number of changes she introduced and can only distinguish between 0-box and 1-box on the last written layer. This means that she can only recognize the parity of $y_i$. Indeed the regular user cannot inspect overwritten layers even in the case of significant head's misalignment.

## 3 DELETION BY OBFUSCATION - ANALYSIS

In this section we provide analysis of methods of data obfuscation. As mentioned in previous sections all methods for deletion in the described model are based solely on overwriting alternatively 0 and 1 some number of times. Thus, each box is seen by the adversary as a stack of 0s and 1s - some added by regular usage and some other just added to mislead the adversary. That is - the algorithms are simple, however the problem is to find how many times one has to change the state to get the demanded security level, according to the security definition given below.

Let us introduce some basic notation. For any natural number $K$, the set $\{0,1,\ldots,K\}$ will be denoted by $[K]$. For the set of natural numbers we use symbol $\mathbf{N}$. Let $\mathbb{X} = (X_1, X_2, \ldots, X_d)$ be a $d$-dimensional discrete random variable representing data to be deleted (more precisely the number of bit-flips performed during Regular User Phase). Clearly they do not have to be independent. Let $\mathbb{S} = (S_1, S_2, \ldots, S_d)$ denote *the covering*, i.e., the $d$-dimensional finite discrete random variable whose role is to mask data $\mathbb{X}$. We assume that after the process of *covering* (during Deletion Phase) the adversary can read $\mathbb{Y} = \mathbb{X} + \mathbb{S} = (X_1 + S_1, X_2 + S_2, \ldots, X_d + S_d)$ from the magnetic disk. Thus the random variable $Y_i = X_i + S_i$ denotes the number of changes performed in the $i$-th box till the end of Deletion Phase. We assume that $\mathbb{X}$ and $\mathbb{S}$ are independent and denote their ranges by $\mathcal{X}$ and $\mathcal{S}$, respectively. Additionally, we assume that $S_1, S_2, \ldots, S_d$ are independent and identically distributed. The last assumption is motivated by practical reasons - deletion method has to be simple and does not require any additional storage.

We would like to construct such a covering $\mathbb{S}$ that the adversary knowing already the concrete realization $\mathbf{y}$ of the random variable $\mathbb{Y}$ gets no (or, in some sense, very little) information about the underlying concrete realization $\mathbf{x}$ of the random variable $\mathbb{X}$. We assume that the adversary knows our technique of covering and the distribution of $\mathbb{S}$ we use and may know the distribution of $\mathbb{X}$. The formal definition of secure covering is given below.

**Definition 1.** *For* $\varepsilon \geq 0$ *and* $\delta \in [0,1]$ *we say that* $\mathbb{S}$ $(\varepsilon, \delta)$*-covers data* $\mathbb{X}$ *if*

$$\mathbb{P}[(\mathbb{X}, \mathbb{S}) \in \mathcal{A}_\varepsilon] \geq 1 - \delta,$$

*where*

$$\mathcal{A}_\varepsilon = \left\{(\mathbf{x}, \mathbf{s}) \in \mathcal{X} \times \mathcal{S} : \frac{1}{1+\varepsilon} \leq \frac{\mathbb{P}[\mathbb{X} = \mathbf{x} | \mathbb{Y} = \mathbf{x} + \mathbf{s}]}{\mathbb{P}[\mathbb{X} = \mathbf{x}]} \leq 1 + \varepsilon\right\},$$

*and* $\mathbb{Y} = \mathbb{X} + \mathbb{S}$.

The idea behind this definition is as follows - with probability at least $1-\delta$ the adversary can observe a value $y$ that changes *a priori* distribution of hidden $x$ by at most multiplicative factor $1+\varepsilon$. In the context of our problem there is some unknown value $x$ of layers in the device. The adversary does not know it, however he knows that $x$ layers can be there with some probability $p_x$. After some obfuscating process (adding $s$ layers) the adversary is given $y = x+s$. After gaining some extra knowledge from $y$ the probability that there were $x$ layers deviates from $p_x$ by no more than a multiplicative factor $1+\varepsilon$. Of course, the smaller the values of $\varepsilon$ and $\delta$, the stronger the covering $\mathbb{S}$ of $\mathbb{X}$ is. In our paper we concentrate on algorithms providing **perfect** security, i.e., $\varepsilon = 0$.

Note that this definition is an adaptation of $(\varepsilon,\delta)$-differential privacy introduced by Dwork et al. in (Dwork et al., 2006b). The original definition has been presented for the problem of preserving privacy of individuals when some statistical information from a database is revealed. This definition is *de facto* a standard, formal and very natural method of measuring revealing information. It has been used in various papers from different fields related to information protection (see e.g. (Shi et al., 2011; Golebiewski et al., 2009)). This kind of security definitions is widely accepted since the security guarantees do not depend on any additional knowledge and are immune against so-called *linkage attacks* in contrast to many other methods of defining how well information is hidden (see (Dwork, 2006; Cynthia Dwork and Aaron Roth, 2013)). Moreover the Dwork et al's idea seems to be very intuitive and can be easily formalized. The fact remains, however, that analyzing it can be difficult due to complex calculations. Note that differential privacy notion leads to required properties - i.e. security is independent of adversary's computational power and is immune against any post-processing. Such properties are inherited by our definition ((Cynthia Dwork and Aaron Roth, 2013)). Finally, let us remark that the definition used in our paper needs a little bit subtle treatment since we do not compare a few "neighbouring" states (i.e., database with or without single individual as in the case of original Dwork at al's paper) but all possible states of memory to be concealed.

We start with formulating theorems for $d = 1$, i.e., for one-dimensional (single bit) data and covering denoting them simply by $X$ and $S$, respectively. We will generalize theorems to arbitrary $d$ later on.

**Theorem 1.** *If $X$ has a distribution other than concentrated in a single point, then $S$ which $(0,0)$-covers $X$ does not exist.*

*Proof.* Let us denote the ranges of $X$ and $S$ by $\mathcal{X}$ and $\mathcal{S}$, respectively ($\mathcal{X}$ and $\mathcal{S}$ are subsets of $\mathbf{N}$, the set of natural numbers). Assume by contradiction that $S$ $(0,0)$-covers $X$. Then $\mathbb{P}[(X,S) \in \mathcal{A}_0] = 1$, where $\mathcal{A}_0 = \{(x,s) \in \mathcal{X}\times\mathcal{S} : \mathbb{P}[X = x|Y = x+s] = \mathbb{P}[X = x]\}$ (recall that $Y = X+S$).

Let $x = \min\{i \in \mathcal{X} : \mathbb{P}[X = i] > 0\}$ and $s = \min\{i \in \mathcal{S} : \mathbb{P}[S = i] > 0\}$. Note that then $\mathbb{P}[X = x|Y = x+s] = 1$ and $\mathbb{P}[X = x] < 1$ (since $X$ does not have a single-point distribution). Thus $(x,s) \notin \mathcal{A}_0$ and, by independence of $X$ and $S$, we get
$\mathbb{P}[X = x, S = s] = \mathbb{P}[X = x]\mathbb{P}[S = s] > 0$. Therefore $\mathbb{P}[(X,S) \in \mathcal{A}_0] < 1$. □

**Remark 1.** *Note that when $X$ has a single-point distribution, i.e., $\mathbb{P}[X = x] = 1$ for some $x$, then arbitrary $S$ $(0,0)$-covers $X$, since*

$$\frac{\mathbb{P}[X = x|Y = x+s]}{\mathbb{P}[X = x]} = \frac{\mathbb{P}[X = x, Y = x+s]}{\mathbb{P}[Y = x+s]\cdot 1} =$$
$$= \frac{\mathbb{P}[X = x, S = s]}{\mathbb{P}[S = s]} = \frac{\mathbb{P}[X = x]\mathbb{P}[S = s]}{\mathbb{P}[S = s]} = 1.$$

*(Recall that $Y = X+S$.) Of course, when $X$ has a single-point distribution, known by adversary, nothing can be done to hide the fact that $x$ was stored on the disc.*

**Theorem 2.** *Let $\delta \in (0,1)$ and let $\mathbb{P}[X \in [K]] \geq 1-\delta/2$. Let also $S$ be uniformly distributed on $[N]$, where $N \geq \frac{K+2-\delta/2}{\delta/2}$. Then $S$ $(0,\delta)$-covers $X$.*

*Proof.* We need to prove that $\mathbb{P}[(X,S) \in \mathcal{A}_0] \geq 1-\delta$, where

$$\mathcal{A}_0 = \{(x,s) \in \mathcal{X}\times\mathcal{S} : \mathbb{P}[X = x|Y = x+s] = \mathbb{P}[X = x]\}.$$

Obviously,

$$\begin{aligned}\mathbb{P}[(X,S) \in \mathcal{A}_0] &= \mathbb{P}[(X,S) \in \mathcal{A}_0 | X \in [K]]\mathbb{P}[X \in [K]] \\ &+ \mathbb{P}[(X,S) \in \mathcal{A}_0 | X \notin [K]]\mathbb{P}[X \notin [K]] \\ &\geq \mathbb{P}[(X,S) \in \mathcal{A}_0 | X \in [K]](1-\delta/2).\end{aligned}$$

Note that if we prove that $\mathbb{P}[(X,S) \in \mathcal{A}_0 | X \in [K]] \geq 1-\delta/2$ we are done, since $(1-\delta/2)^2 \geq 1-\delta$. From now on all the calculations are done by the assumption $X \in [K]$.

Note that, by independence of $X$ and $S$, we get

$$\mathbb{P}[X = x|Y = x+s] = \frac{\mathbb{P}[X = x, Y = x+s]}{\mathbb{P}[Y = x+s]} =$$
$$\frac{\mathbb{P}[X = x, S = s]}{\mathbb{P}[Y = x+s]} = \frac{\mathbb{P}[X = x]\mathbb{P}[S = s]}{\mathbb{P}[Y = x+s]},$$

thus the condition $\mathbb{P}[X = x|Y = x+s] = \mathbb{P}[X = x]$ is equivalent to $\mathbb{P}[Y = x+s] = \mathbb{P}[S = s]$. Since $S \sim U\{0,1,\ldots,N\}$ we get

$$\mathcal{A}_0 = \left\{(x,s) \in \mathcal{X}\times\mathcal{S} : \mathbb{P}[Y = x+s] = \frac{1}{N+1}\right\}.$$

By total probability law $\mathbb{P}[Y=x+s]=\sum_{i=0}^{N}\mathbb{P}[X=x+s-i]\mathbb{P}[S=i]$ thus for $(x,s)\in\mathcal{X}\times\mathcal{S}$ such that $K\le x+s\le N$ we have

$$\mathbb{P}[Y=x+s]=\sum_{i=x+s-K}^{x+s}\mathbb{P}[X=x+s-i]\mathbb{P}[S=i]=$$
$$=\frac{1}{N+1}\sum_{i=0}^{K}\mathbb{P}[X=i]=\frac{1}{N+1}.$$

For $(x,s)\in\mathcal{X}\times\mathcal{S}$ such that $x+s<K$ we have

$$\mathbb{P}[Y=x+s]=\sum_{i=0}^{x+s}\mathbb{P}[X=x+s-i]\mathbb{P}[S=i]=$$
$$=\frac{1}{N+1}\sum_{i=0}^{x+s}\mathbb{P}[X=i]\le\frac{1}{N+1},$$

and the above inequality is strict if only there exists $j>x+s$ such that $\mathbb{P}[X=j]>0$. Analogously, for $(x,s)\in\mathcal{X}\times\mathcal{S}$ such that $x+s>N$, we have

$$\mathbb{P}[Y=x+s]=\sum_{i=x+s-K}^{N}\mathbb{P}[X=x+s-i]\mathbb{P}[S=i]=$$
$$\frac{1}{N+1}\sum_{i=x+s-N}^{K}\mathbb{P}[X=i]\le\frac{1}{N+1},$$

and again the above inequality is strict if only there exists $j<x+s-N$ such that $\mathbb{P}[X=j]>0$. Therefore $\{(x,s)\in\mathcal{X}\times\mathcal{S}\colon\ K\le x+s\le N\}\subseteq\mathcal{A}_0$. Thus if we show that $\mathbb{P}[K\le Y\le N]\ge 1-\delta/2$ or, equivalently,

$$\mathbb{P}[Y<K]+\mathbb{P}[Y>N]\le\delta/2,$$

we are done. Let us calculate $\mathbb{P}[Y\le K]$. By equalities from (1) we obtain

$$\mathbb{P}[Y\le K]=\sum_{y=0}^{K}\mathbb{P}[Y=y]=$$
$$=\sum_{y=0}^{K}\frac{1}{N+1}\sum_{x=0}^{y}\mathbb{P}[X=x]=$$
$$=\frac{1}{N+1}\sum_{x=0}^{K}(K+1-x)\mathbb{P}[X=x]$$
$$=\frac{1}{N+1}\left((K+1)\sum_{x=0}^{K}\mathbb{P}[X=x]-\sum_{x=0}^{K}x\mathbb{P}[X=x]\right)$$
$$=\frac{K+1-\mathbb{E}X}{N+1}.$$

On the other hand, by equalities from (1) we get

$$\mathbb{P}[Y\ge N]=\sum_{y=N}^{N+K}\mathbb{P}[Y=y]=\sum_{y=N}^{N+K}\frac{1}{N+1}\sum_{x=y-N}^{K}\mathbb{P}[X=x]=$$
$$=\frac{1}{N+1}\sum_{x=0}^{K}(x+1)\mathbb{P}[X=x]=$$
$$=\frac{1}{N+1}\left(\sum_{x=0}^{K}x\mathbb{P}[X=x]+\sum_{x=0}^{K}\mathbb{P}[X=x]\right)=\frac{\mathbb{E}X+1}{N+1}.$$

Finally, since $N\ge\frac{K+2-\delta/2}{\delta/2}$, we obtain

$$\mathbb{P}[Y<K]+\mathbb{P}[Y>N]\le\mathbb{P}[Y\le K]+\mathbb{P}[Y\ge N]=$$
$$=\frac{K+2}{N+1}\le\delta/2. \tag{1}$$

□

The following theorem generalizes the above one to arbitrary dimension $d$ with the assumption that the coordinates $X_i$ of data $\mathbb{X}=(X_1,X_2,\dots,X_d)$ are independent.

**Theorem 3.** *Let $\delta\in(0,1)$. Let $\mathbb{X}=(X_1,X_2,\dots,X_d)$, where $X_i$'s are independent and let $\mathbb{P}[X_1,\dots,X_d\in[K]]\ge 1-\delta/2$. Let $\mathbb{S}=(S_1,S_2,\dots,S_d)$ with $S_i$'s being independent and uniformly distributed on $[N]$, wherein $N\ge\frac{2K+3+(1-\delta/2)^{1/d}}{1-(1-\delta/2)^{1/d}}$. Assume that $\mathbb{X}$ and $\mathbb{S}$ are independent. Then $\mathbb{S}$ $(0,\delta)$-covers data $\mathbb{X}$.*

*Proof.* Recall that $\mathbb{Y}=\mathbb{X}+\mathbb{S}=(X_1+S_1,\dots,X_d+S_d)$. We need to prove that $\mathbb{P}[(\mathbb{X},\mathbb{S})\in\mathcal{A}_0]\ge 1-\delta$, where

$$\mathcal{A}_0=\{(\mathbf{x},\mathbf{s})\in\mathcal{X}\times\mathcal{S}:\mathbb{P}[\mathbb{X}=\mathbf{x}|\mathbb{Y}=\mathbf{x}+\mathbf{s}]=\mathbb{P}[\mathbb{X}=\mathbf{x}]\}.$$

Let $C$ denote the event that $X_1,\dots,X_d\in[K]$. We have

$$\mathbb{P}[(\mathbb{X},\mathbb{S})\in\mathcal{A}_0]\ge$$
$$\mathbb{P}[(\mathbb{X},\mathbb{S})\in\mathcal{A}_0|C]\mathbb{P}[C]\ge\mathbb{P}[(\mathbb{X},\mathbb{S})\in\mathcal{A}_0|C](1-\delta/2).$$

Note that if we show that $\mathbb{P}[(\mathbb{X},\mathbb{S})\in\mathcal{A}_0|C]\ge 1-\delta/2$ we are done, since $(1-\delta/2)^2\ge 1-\delta$. From now on all the calculations are done assuming that $C$ holds.

The condition $\mathbb{P}[\mathbb{X}=\mathbf{x}|\mathbb{Y}=\mathbf{x}+\mathbf{s}]=\mathbb{P}[\mathbb{X}=\mathbf{x}]$ is equivalent to $\mathbb{P}[\mathbb{Y}=\mathbf{x}+\mathbf{s}]=\mathbb{P}[\mathbb{S}=\mathbf{s}]$ (compare (1)). Since $S_i\sim U\{0,1,\dots,N\}$, $i=1,2,\dots,d$, and $S_i$'s are independent, we get

$$\mathcal{A}_0=\left\{(\mathbf{x},\mathbf{s})\in\mathcal{X}\times\mathcal{S}:\mathbb{P}[\mathbb{Y}=\mathbf{x}+\mathbf{s}]=\frac{1}{(N+1)^d}\right\}.$$

Note that by independence of $X_i$'s and $S_i$'s, $Y_i$'s are also independent. Recall that for $x_i,s_i$ such that $K\le x_i+s_i\le N$ we have

$$\mathbb{P}[Y_i=x_i+s_i]=\frac{1}{N+1}$$

while for $x_i,s_i$ such that $x_i+s_i<K$ or $x_i+s_i>N$ we have

$$\mathbb{P}[Y_i=x_i+s_i]\le\frac{1}{N+1}$$

(compare proof of Theorem 2). Therefore $\mathbb{P}[\mathbb{Y}=\mathbf{x}+\mathbf{s}]=\frac{1}{(N+1)^d}$ if $K\le x_i+s_i\le N$ for all $i=1,\dots,d$. Note that what we actually need is that $S_i$ $(0,1-(1-\delta/2)^{1/d})$-covers $X_i$ for each $i$, since then, by independence of $Y_i$'s

$$\mathbb{P}[(\mathbb{X},\mathbb{S})\in\mathcal{A}_0|C]\ge(1-(1-(1-\delta/2)^{1/d}))^d=1-\delta/2.$$

By Theorem 2 we get that $S_i$ $(0, 1-(1-\delta/2)^{1/d})$-covers $X_i$ for $N \geq \frac{2K+3+(1-\delta/2)^{1/d}}{1-(1-\delta/2)^{1/d}}$. □

Subsequently, we will formulate analogous theorem, this time without the assumption on independence of $X_i$'s. Let us start with proving one lemma that will be helpful later on.

**Lemma 1.** *Let $\mathbb{S} = (S_1, S_2)$, where $S_1$ and $S_2$ are independent random variables, uniformly distributed on $[N]$. Let $\mathbb{X} = (X_1, X_2)$, where $X_1$ and $X_2$ are not necessarily independent random variables with range $[K]$, $K < N$. Assume that $\mathbb{X}$ and $\mathbb{S}$ are independent. Let $\mathbb{Y} = (Y_1, Y_2) = (X_1 + S_1, X_2 + S_2)$. Then*

$$\mathbb{P}[Y_1 \leq K, Y_2 \leq K] + \mathbb{P}[Y_1 \leq K, Y_2 \geq N] +$$

$$+ \mathbb{P}[Y_1 \geq N, Y_2 \leq K] + \mathbb{P}[Y_1 \geq N, Y_2 \geq N] = \left(\frac{K+2}{N+1}\right)^2.$$

*Proof.*

$$\mathbb{P}[Y_1 \leq K, Y_2 \leq K] = \sum_{y_1=0}^{K} \sum_{y_2=0}^{K} \mathbb{P}[Y_1 = y_1, Y_2 = y_2]$$

$$= \sum_{y_1=0}^{K} \sum_{y_2=0}^{K} \sum_{x_1=0}^{y_1} \sum_{x_2=0}^{y_2} \mathbb{P}[\mathbb{X} = (x_1, x_2), \mathbb{S} = (y_1 - x_1, y_2 - x_2)]$$

$$= \frac{1}{(N+1)^2} \sum_{y_1=0}^{K} \sum_{y_2=0}^{K} \sum_{x_1=0}^{y_1} \sum_{x_2=0}^{y_2} \mathbb{P}[X_1 = x_1, X_2 = x_2]$$

$$= \frac{1}{(N+1)^2} \sum_{y_1=0}^{K} \sum_{x_1=0}^{y_1} \sum_{x_2=0}^{K} (K+1-x_2)\mathbb{P}[X_1 = x_1, X_2 = x_2]$$

$$= \frac{1}{(N+1)^2} \sum_{y_1=0}^{K} \sum_{x_1=0}^{y_1} \mathbb{P}[X_1 = x_1](K+1-\mathbb{E}[X_2|X_1 = x_1])$$

$$= \frac{1}{(N+1)^2} \times$$

$$\times \sum_{x_1=0}^{K} (K+1-x_1)\mathbb{P}[X_1 = x_1](K+1-\mathbb{E}[X_2|X_1 = x_1]) =$$

$$= \frac{1}{(N+1)^2} \sum_{x_1=0}^{K} (K+1)\mathbb{P}[X_1 = x_1](K+1-\mathbb{E}[X_2|X_1 = x_1])$$

$$- \frac{1}{(N+1)^2} \sum_{x_1=0}^{K} x_1 \mathbb{P}[X_1 = x_1](K+1-\mathbb{E}[X_2|X_1 = x_1])$$

$$= \frac{(K+1)^2 - (K+1)(\mathbb{E}X_1 + \mathbb{E}X_2) + \mathbb{E}[X_1X_2]}{(N+1)^2}. \tag{2}$$

By analogous calculations we obtain

$$\mathbb{P}[Y_1 \geq N, Y_2 \geq N] = \frac{\mathbb{E}[X_1X_2] + \mathbb{E}X_1 + \mathbb{E}X_2 + 1}{(N+1)^2}, \tag{3}$$

$$\mathbb{P}[Y_1 \geq N, Y_2 \leq K] = \frac{(K+1)\mathbb{E}X_1 - \mathbb{E}X_2 - \mathbb{E}[X_1X_2] + K + 1}{(N+1)^2}, \tag{4}$$

and

$$\mathbb{P}[Y_1 \leq K, Y_2 \geq N] = \frac{(K+1)\mathbb{E}X_2 - \mathbb{E}X_1 - \mathbb{E}[X_1X_2] + K + 1}{(N+1)^2}. \tag{5}$$

Summing up (2), (3), (4), and (5) concludes the proof. □

The formulation of the next theorem is analogous to Theorem 3, however this time we do **not** assume that the coordinates $X_i$'s of data $\mathbb{X} = (X_1, X_2, \ldots, X_d)$ are independent.

**Theorem 4.** *Let $\delta \in (0,1)$. Let $\mathbb{X} = (X_1, X_2, \ldots, X_d)$ and let $\mathbb{P}[X_1, \ldots, X_d \in [K]] \geq 1 - \delta/2$. Let $\mathbb{S} = (S_1, S_2, \ldots, S_d)$ with $S_i$'s being independent and uniformly distributed on $[N]$, wherein $N \geq \frac{d(K+2)-\delta/2}{\delta/2}$. Assume that $\mathbb{X}$ and $\mathbb{S}$ are independent. Then $\mathbb{S}$ $(0, \delta)$-covers data $\mathbb{X}$.*

**Remark 2.** *The proof contains many references to calculations from the proof of Theorem 2.*

*Proof.* Recall that $\mathbb{Y} = \mathbb{X} + \mathbb{S} = (X_1 + S_1, X_2 + S_2, \ldots, X_d + S_d)$ and $\mathcal{X}$ and $\mathcal{S}$ denote the ranges of $\mathbb{X}$ and $\mathbb{S}$, respectively. We need to prove that $\mathbb{P}[(\mathbb{X}, \mathbb{S}) \in \mathcal{A}_0] \geq 1 - \delta$, where

$$\mathcal{A}_0 = \{(\mathbf{x}, \mathbf{s}) \in \mathcal{X} \times \mathcal{S} : \mathbb{P}[\mathbb{X} = \mathbf{x} | \mathbb{Y} = \mathbf{x} + \mathbf{s}] = \mathbb{P}[\mathbb{X} = \mathbf{x}]\}.$$

The condition $\mathbb{P}[\mathbb{X} = \mathbf{x} | \mathbb{Y} = \mathbf{x} + \mathbf{s}] = \mathbb{P}[\mathbb{X} = \mathbf{x}]$ is equivalent to $\mathbb{P}[\mathbb{Y} = \mathbf{x} + \mathbf{s}] = \mathbb{P}[\mathbb{S} = \mathbf{s}] = \frac{1}{(N+1)^d}$ (compare (1)).

Let $\mathcal{B} = \{(\mathbf{x}, \mathbf{s}) \in \mathcal{X} \times \mathcal{S} : K \leq x_i + s_i \leq N, i = 1, \ldots, d\}$. For $(\mathbf{x}, \mathbf{s}) \in \mathcal{B}$ we have (compare 1)

$$\mathbb{P}[\mathbb{Y} = \mathbf{x} + \mathbf{s}] =$$

$$= \sum_{i_1 = x_1 + s_1 - K}^{x_1 + s_1} \cdots \sum_{i_d = x_d + s_d - K}^{x_d + s_d}$$

$$\mathbb{P}[\mathbb{X} = (x_1 + s_1 - i_1, \ldots, x_d + s_d - i_d)]\mathbb{P}[\mathbb{S} = (i_1, \ldots, i_d)]$$

$$= \frac{1}{(N+1)^d} \sum_{x_1=0}^{K} \cdots \sum_{x_d=0}^{K} \mathbb{P}[\mathbb{X} = (x_1, \ldots, x_d)] = \frac{1}{(N+1)^d}.$$

Note that for $(\mathbf{x}, \mathbf{s}) \notin \mathcal{B}$ we always obtain

$$\mathbb{P}[\mathbb{Y} = \mathbf{x} + \mathbf{s}] \leq \frac{1}{(N+1)^d}$$

(compare the calculations for $d = 1$ from the proof of Theorem 3). Thus $\mathcal{B} \subseteq \mathcal{A}_0$ and it is enough to show that $\mathbb{P}[(\mathbb{X}, \mathbb{S}) \in \mathcal{B}] \geq 1 - \delta$. Let $C$ denote the event that $X_1, \ldots, X_d \in [K]$. We have

$$\mathbb{P}[(\mathbb{X}, \mathbb{S}) \in \mathcal{B}] \geq \mathbb{P}[(\mathbb{X}, \mathbb{S}) \in \mathcal{B} | C]\mathbb{P}[C] \geq$$
$$\geq \mathbb{P}[(\mathbb{X}, \mathbb{S}) \in \mathcal{B} | C](1 - \delta/2).$$

Note that if we show that $\mathbb{P}[(\mathbb{X}, \mathbb{S}) \in \mathcal{B} | C] \geq 1 - \delta/2$, or, equivalently, $\mathbb{P}[(\mathbb{X}, \mathbb{S}) \notin \mathcal{B} | C] \leq \delta/2$, we are done, since $(1 - \delta/2)^2 \geq 1 - \delta$. From now on all the calculations are done assuming that $C$ holds.

We have

$$\begin{aligned}
&\mathbb{P}[(\mathbb{X},\mathbb{S}) \notin \mathcal{B}|C] = \\
&\quad = \mathbb{P}[Y_1 < K \cup Y_1 > N \cup \ldots \cup Y_d < K \cup Y_d > N] \\
&\quad \leq \mathbb{P}[Y_1 \leq K \cup Y_1 \geq N \cup \ldots \cup Y_d \leq K \cup Y_d \geq N] \\
&\quad = \sum_{i=1}^{d} (\mathbb{P}[Y_i \leq K] + \mathbb{P}[Y_i \geq N]) \\
&\quad - \sum_{1 \leq i < j \leq d} (\mathbb{P}[Y_i \leq K, Y_j \leq K] + \mathbb{P}[Y_i \leq K, Y_j \geq N]) \\
&\quad - \sum_{1 \leq i < j \leq d} (\mathbb{P}[Y_i \geq N, Y_j \leq K] + \mathbb{P}[Y_i \geq N, Y_j \geq N]) \\
&\quad + \ldots + \\
&(-1)^{d+1} (\mathbb{P}[Y_1 \leq K, \ldots, Y_d \leq K] + \mathbb{P}[Y_1 \geq N, \ldots, Y_d \geq N]).
\end{aligned} \tag{6}$$

Let

$$U = \mathbb{P}[Y_1 < K \cup Y_1 > N \cup \ldots \cup Y_d < K \cup Y_d > N]$$

and

$$\begin{aligned}
L = \sum_{1 \leq i < j \leq d} &(\mathbb{P}[Y_i \leq K, Y_j \leq K] + \mathbb{P}[Y_i \leq K, Y_j \geq N] + \\
&+ \mathbb{P}[Y_i \geq N, Y_j \leq K] + \mathbb{P}[Y_i \geq N, Y_j \geq N]).
\end{aligned}$$

Note that (6) is just a special case of inclusion-exclusion principle, thus

$$U - L \leq \mathbb{P}[(\mathbb{X},\mathbb{S}) \notin \mathcal{B}|C] \leq U.$$

By the equality in 1 we have

$$U = d(\mathbb{P}[Y_1 \leq K] + \mathbb{P}[Y_1 \geq N]) = \frac{d(K+2)}{N+1}$$

and by Lemma 1

$$\begin{aligned}
L &= \binom{d}{2} (\mathbb{P}[Y_1 \leq K, Y_2 \leq K] + \mathbb{P}[Y_1 \leq K, Y_2 \geq N] \\
&\quad + \mathbb{P}[Y_1 \geq N, Y_2 \leq K] + \mathbb{P}[Y_1 \geq N, Y_2 \geq N]) \\
&= \binom{d}{2} \left(\frac{K+2}{N+1}\right)^2.
\end{aligned}$$

Therefore

$$\begin{aligned}
\frac{d(K+2)}{N+1} - \binom{d}{2} \left(\frac{K+2}{N+1}\right)^2 &\leq \mathbb{P}[(\mathbb{X},\mathbb{S}) \notin \mathcal{B}|C] \\
&\leq \frac{d(K+2)}{N+1}.
\end{aligned}$$

Since $N \geq \frac{d(K+2)-\delta/2}{\delta/2}$, the conclusion follows from the righthandside of the above inequality.

Note that if only $dK \ll N$, the above estimation is tight up to the order of $\Theta(dK/N)$. □

**Remark 3.** *Of course, the lower bound on N guaranteeing that* $\mathbb{S}$ $(0,\delta)$*-covers* $\mathbb{X}$ *is bigger when we deal with general data* $\mathbb{X} = (X_1, X_2, \ldots, X_d)$*, i.e., when* $X_i$*'s may be dependent. In Theorem 3 we need*

$$N \geq \frac{2K+3+(1-\delta/2)^{1/d}}{1-(1-\delta/2)^{1/d}} \sim \frac{2K}{1-(1-\delta/2)^{1/d}}$$

*when* $X_i$*'s are independent, while in Theorem 4*

$$N \geq \frac{d(K+2)-\delta/2}{\delta/2} \sim \frac{2dK}{\delta}$$

*when* $X_i$*'s may be dependent. Note, however, that the price we pay in general case (Theorem 4) is not very significant when* $\delta$ *is small. For small* $\delta$ *we may actually estimate* $\frac{1}{1-(1-\delta/2)^{1/d}} \sim \frac{1}{1-e^{-\delta/2d}} \sim \frac{2d}{\delta}$.

## 3.1 Universal Random Variables and Uniform Obfuscation

In this paper we discussed perfect obfuscation, i.e., such that with probability at least $1-\delta$ the observer having access to observable $\mathbb{Y}$, learns no new information about the real data $\mathbb{X}$. The solutions we proposed are *universal* with respect to $K$, i.e., they guarantee (with some strictly controlled probability) obfuscation of any $\mathbb{X}$ concentrated on $[K]$ (except cases of total probability at most $\delta$). Note that all our methods use for obfuscation a random variable with uniform distribution on a fixed set.

One may ask two questions. First, do we really need a universal obfuscation? Clearly, one can construct more efficient methods of obfuscation tailored for a given $\mathbb{X}$. We claim that demanding universal method is fully justified by practical needs. Regular users of data storage devices do not know the distribution of their stored data (possibly except a case, while the disc is used for storing cryptographic materials as in (Klonowski et al., 2008)). Moreover, the adversary having access to the big volume of statistical data can have some knowledge of the distribution to be obfuscated and have another advantage over a regular user.

Second, even if we accept that the universal obfuscation is really needed, one may be tempted to construct substantially different (and possibly more efficient) methods that deviate from using uniform distribution. Below we show that any universal method offering perfect security (i.e., $\varepsilon = 0$) has to use for obfuscating a distribution "close" to uniform. More precisely, if we demand that $\mathbb{S}$ $(0,\delta)$-covers any $\mathbb{X}$, then the set of values of $\mathbb{S}$ having the same probability needs to be of measure at least $1-(\delta/(1-\gamma))$ for arbitrarily small $\gamma \in (0,1)$.

In this section, let $p_s$ denote $\mathbb{P}[\mathbb{S} = s]$ for $s \in \mathbf{N}$ and 0 otherwise. In the theorem below we consider one-dimensional $\mathbb{X}$ and $\mathbb{S}$.

**Theorem 5.** *Let* $\mathbb{S}$ *be a random variable distributed on* $\mathbf{N}$*. If there exists* $\mathcal{S}^* \subset \mathbf{N}$ *such that* $\mathbb{P}[\mathbb{S} \in \mathcal{S}^*] > \delta/(1-\gamma)$ *for arbitrarily small* $\gamma \in (0,1)$ *and for every*

*different $s, s' \in \mathcal{S}^*$ holds $p_s \neq p_{s'}$, then $\mathbb{S}$* ***cannot*** *$(0,\delta)$-cover arbitrarily distributed data $\mathbb{X}$.*

*Proof.* Aiming for a contradiction let us assume that $\mathbb{S}$ $(0,\delta)$-covers any data $\mathbb{X}$ and that there exists $\mathcal{S}^* \subseteq \mathbf{N}$ such that $p_s \neq p_{s'}$ for any different $s, s'$ from $\mathcal{S}^*$ and $\mathbb{P}[\mathbb{S} \in \mathcal{S}^*] > \delta/(1-\gamma)$ for arbitrarily small $\gamma \in (0,1)$.

To every $s \in \mathcal{S}^*$ let us assign $s^{**} = \min\{i \in \mathbf{N} | p_i = p_s\}$. Let $\mathcal{S}^{**} = \{s^{**} | s \in \mathcal{S}^* \wedge p_s > 0\}$. That is, $\mathcal{S}^{**}$ is the set of elements that appear with a given, positive, probability for the first time in the sequence of natural numbers. Clearly, $\mathbb{P}[\mathbb{S} \in \mathcal{S}^{**}] = \mathbb{P}[\mathbb{S} \in S^*] > \delta/(1-\gamma)$ and for any different $s, s'$ from $\mathcal{S}^{**}$ $p_s \neq p_{s'}$ holds.

Now let us look at data $\mathbb{X}$ such that $\mathbb{P}[\mathbb{X} = 0] = 1-\gamma$ and $\mathbb{P}[\mathbb{X} = 1] = \gamma$. Consider the set of pairs $(0, \mathbf{s}) \in \{0,1\} \times \mathcal{S}^{**}$. For those pairs, by the construction of $S^{**}$, we obtain

$$\begin{aligned}\mathbb{P}[\mathbb{Y} = \mathbf{x}+\mathbf{s}] = \mathbb{P}[\mathbb{Y} = \mathbf{s}] &= \mathbb{P}[\mathbb{X} = 0] \cdot p_s + \mathbb{P}[\mathbb{X} = 1] \cdot p_{s-1} \\ &= (1-\gamma) \cdot p_s + \gamma \cdot p_{s-1} \neq p_s = \mathbb{P}[\mathbb{S} = \mathbf{s}].\end{aligned} \tag{7}$$

Recall that $(\mathbf{x}, \mathbf{s}) \in \mathcal{A}_0$ if and only if $\mathbb{P}[\mathbb{Y} = \mathbf{x}+\mathbf{s}] = \mathbb{P}[\mathbb{S} = \mathbf{s}]$. Definitely, by Equation 7, if $\mathbf{s} \in \mathcal{S}^{**}$ then $(0, \mathbf{s}) \notin \mathcal{A}_0$. By independence of $\mathbb{X}$ and $\mathbb{S}$ we get

$$\begin{aligned}\mathbb{P}[(\mathbb{X}, \mathbb{S}) \in \{0\} \times \mathcal{S}^{**}] &= \mathbb{P}[\mathbb{X} = 0] \cdot \mathbb{P}[\mathbb{S} \in \mathcal{S}^{**}] \\ &> \frac{(1-\gamma)\delta}{(1-\gamma)} = \delta.\end{aligned}$$

Thus $\mathbb{P}[(\mathbb{X}, \mathbb{S}) \in \mathcal{A}_0] < 1-\delta$, which contradicts the statement that $\mathbb{S}$ $(0,\delta)$-covers any data $\mathbb{X}$. □

Note that the construction from the proof uses $\mathbb{X}$ concentrated on the set $\{0,1\}$. Thus the proof is valid for **any** $K \geq 2$.

# 4 FAST METHOD FOR DATA OBFUSCATION

The methods presented in previous part of the paper offer a perfect security with fully controllable probability. However, they all have a drawback that prevents them from using in some real-life scenarios. Namely, note that in Deletion Phase the user has to perform a large number of writing operations. Translating it into real-life terms the whole disc has to be overwritten during Deletion Phase the number of times that is greater than the number of times the disc can be used in the Regular User Phase. Even using ultra fast devices the time one needs for all the operations can be prohibitively large in natural application wherein a user has to remove the content possibly quickly.

In this section we present a method that drastically reduces the number of operations during Deletion Phase for the price of a very long Preliminary Phase, while the perfect security property of deletion is still preserved. This means, however, that a storage device has to be prepared for a long time **before** it is used for storing some data. The Deletion Phase needs only linear (with respect to the size of the data to be deleted) number of operations. Note that this number of operations is asymptotically optimal - indeed, one needs $\Omega(d)$ operations to somehow delete data written on the last layer.

## 4.1 Algorithm Description

The protocol uses one parameter - $K$ - the upper bound on the number of times each bit can be changed during Regular User Phase.

**Preliminary Phase -** $i$-th box is flipped $\hat{S}_i$ times, where $\hat{S}_i$'s are independent random variables uniformly distributed on $\{0,2,\ldots,N-3,N-1\}$ ($N$ is an odd number depending on $K$, setting its value will be discussed in the next section).

**Regular Usage Phase -** the user can flip each bit up to $K$ times, we denote those changes by random variable $\mathbb{X} = (X_1, X_2, \ldots, X_d)$ (we mean that $i$-th box was flipped $X_i$ times).

**Deletion Phase -** the state of each box is changed independently with probability $1/2$, which will be denoted by $\mathbb{C} = (C_1, C_2, \ldots, C_d)$, where $C_i$'s are independent random variables such that $\mathbb{P}[C_i = 0] = \mathbb{P}[C_i = 1] = 1/2$ (of course, $C_i$ refers to possible **single** change made in the $i$-th box).

## 4.2 Analysis

First let us note that the Deletion Phase is optimal - it makes only $O(d)$ changes (since each box is changed at most once). Moreover, we believe that the deletion operation can be performed quickly in real-life settings.

Now we discuss security of this approach. Let us analyze the $i$-th box. The adversary knows our technique of obfuscation and is given the state of the disc after the Deletion Phase. Thus what she sees is $\hat{S}_i + X_i + C_i$, where $\hat{S}_i$, $X_i$, and $C_i$ are independent. Note that since $\hat{S}_i$ is uniformly distributed on $\{0,2,\ldots,N-3,N-1\}$ (the set of even numbers not greater than $N-1$) and $C_i$ is a Bernoulli trial with probability $1/2$, we get that $\hat{S}_i + C_i$ (let us denote this sum by $S_i$) is uniformly distributed on $[N]$. Thus what adversary knows is that the number of layers she sees is the sum of independent random variables: $S_i$ (which

is uniformly distributed on $[N]$) and $X_i$ (data with distribution that may be known to the adversary). Thus what we get is the situation exactly analogous to the one from the previous section. Therefore, in order to choose the value of $N$ guaranteeing the desired level of security, one may apply Theorem 4.

Finally let us justify some details of the presented construction that seems to be artificial at first glance. Let us note that the number of flips in the Preliminary Phase, $\hat{S}_i$, is chosen from even numbers to have **always** 0-box on the last layer after Preliminary Phase. Thanks to this trick $X_i$ is independent of $\hat{S}_i$. However $\hat{S}_i + X_i$ still reveals the parity of $X_i$. In particular the adversary can inspect the last layer by just checking the state of the boxes. For that reason we need to add $C_i$ in Deletion Phase to complete obfuscation of $X_i$.

# 5 CONCLUSIONS AND FUTURE WORK

In this paper we have presented and analyzed methods for provable deletion of stored data. We believe that this is a good starting point to broader analysis of provably secure deletion problems. All presented methods offer perfect security (i.e., we set $\varepsilon = 0$). Note that this is a very strong requirement. One can expect that relaxing this assumption will lead to obtaining more efficient algorithms. Note also that throughout the whole paper we were considering concealing data from arbitrary distribution, whereas we hope to construct more practical solutions for special types of data. This issue, as well as the case when the security parameter $\varepsilon$ is greater than 0, are left for a future work.

# ACKNOWLEDGEMENTS

This paper is supported by Polish National Science Center. Preliminary ideas has been supported by the grant UMO-2013/09/B/ST6/02251. Full version with the formal analysis was prepared thanks to grant UMO-2018/29/B/ST6/02969 .

# REFERENCES

Gutmann P. (1996). Epilogue to: Secure deletion of data from magnetic and solid-state.

Ali M., Dhamotharan R., Khan E., Khan S.U., Vasilakos A.V., Li K., and Zomaya A.Y. (2017). Sedasc: Secure data sharing in clouds. *IEEE Systems Journal*, 11(2):395–404.

Bacis E., De Capitani di Vimercati S., Foresti S., Paraboschi S., Rosa M., and Samarati P. (2016). Mix & slice: Efficient access revocation in the cloud. *ACM Conference on Computer and Communications Security*, pages 217–228.

Cynthia Dwork and Aaron Roth (2013). The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3-4):211–407.

Dwork, C. (2006). Differential privacy. *ICALP*.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006a). Calibrating noise to sensitivity in private data analysis. In (Halevi and Rabin, 2006), pages 265–284.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006b). Calibrating noise to sensitivity in private data analysis. In (Halevi and Rabin, 2006), pages 265–284.

Goldreich, O. and Ostrovsky, R. (1996). Software protection and simulation on oblivious rams. *J. ACM*, 43(3):431–473.

Golebiewski, Z., Klonowski, M., Koza, M., and Kutylowski, M. (2009). Towards fair leader election in wireless networks. In Ruiz, P. M. and Garcia-Luna-Aceves, J. J., editors, *Ad-Hoc, Mobile and Wireless Networks, 8th International Conference, ADHOC-NOW 2009, Murcia, Spain, September 22-25, 2009, Proceedings*, volume 5793 of *Lecture Notes in Computer Science*, pages 166–179. Springer.

Gomez R. D., Burke E. R., Adly A. A., Mayergoyz I. D., Gorczyca J. A., and Kryder M. H. (1993). Microscopic investigations of overwritten data. *Journal of Applied Physics*, 73:6001–6003.

Gutmann Peter (1996). Secure deletion of data from magnetic and solid-state memory. In *In Proceedings of the 6th USENIX Security Symposium*, pages 77–89.

Halevi, S. and Rabin, T., editors (2006). *Theory of Cryptography, Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006, Proceedings*, volume 3876 of *Lecture Notes in Computer Science*. Springer.

Hao F., Clarke D., and Zorzo A.F. (2016). Deleting secret data with public verifiability. *IEEE Trans. Dependable Sec. Comput.*, 13(6):617–629.

Hughes G., Coughlin T., and Commins D. (2009). Disposal of disk and tape data by secure sanitization.

Hur J., Koo D., Shin Y., and Kang K. (2017). Secure data deduplication with dynamic ownership management in cloud storage. *IEEE International Conference on Data Engineering*, pages 69–70.

Jia, S., Xia, L., Chen, B., and Liu, P. (2016). NFPS: adding undetectable secure deletion to flash translation layer. In *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security, AsiaCCS 2016, Xi'an, China, May 30 - June 3, 2016*, pages 305–315.

Klonowski, M., Przykucki, M., and Struminski, T. (2008). Data deletion with provable security. In *International Workshop on Information Security Applications*, WISA '08, pages 240–255.

Klonowski, M., Przykucki, M., and Struminski, T. (2009). Data deletion with time-aware adversary model. In

*Proceedings IEEE CSE'09, 12th IEEE International Conference on Computational Science and Engineering, August 29-31, 2009, Vancouver, BC, Canada*, pages 659–664. IEEE Computer Society.

Mayergoyz I. D., Tse C., Krafft C., and Gomez R. D. (2001). Spin-stand imaging of overwritten data and its comparison with magnetic force microscopy. *Journal of Applied Physics*, 89:6772–6774.

Moran T., Naor M., and Segev G. (2009). Deterministic history-independent strategies for storing information on write-once memories. *Theory of Computing*, 5:43–67.

Moritz C.A., Chheda S., and Carver K. (2015). Securing microprocessors against information leakage and physical tampering. Patent US 9940445 B2.

Rivest, R. L. and Shamir, A. (1982). How to reuse a write–once memory. In *Proceedings of the fourteenth annual ACM symposium on Theory of computing*, STOC '82, pages 105–113.

Roche, D. S., Aviv, A. J., and Choi, S. G. (2016). A practical oblivious map data structure with secure deletion and history independence. In *IEEE Symposium on Security and Privacy, SP 2016, San Jose, CA, USA, May 22-26, 2016*, pages 178–197.

Rugar D., Mamin H. J., Guethner P., Lambert S. E., Stern J. E., McFadyen I., and Yogi T. (1990). Magnetic force microscopy: General principles and application to longitudinal recording media. *Journal of Applied Physics*, 68(3):1169–1183.

Shi, E., Chan, T.-H. H., Rieffel, E., Chow, R., and Song, D. (2011). Privacy-preserving aggregation of time-series data. *NDSS*.

US Department of Defense (1997). National industrial security program operating manual NISPOM January 1995. Technical Report DoD 5220.22-M.

U.S. National Institute of Standards and Technology (2006). Nist special publication 800-88: Guidelines for media sanitization.

Wegberg G., Ritzdorf H., and Čapkun S. (2017). Multi-user secure deletion on agnostic cloud storage.

Yang C., Chen X., and Xiang Y. (2018). Blockchain-based publicly verifiable data deletion scheme for cloud storage. *Journal of Network and Computer Applications*, 103:185–193.