J. Cryptol. (2014) 27: 109–138
DOI: 10.1007/s00145-012-9139-0

Journal of
CRYPTOLOGY

A New Interactive Hashing Theorem*

Iftach Haitner†
School of Computer Science, Tel Aviv University, Tel Aviv, Israel
iftachh@cs.tau.ac.il

Omer Reingold
Microsoft Research, Silicon Valley Campus, Mountain View, CA, USA
omreing@microsoft.com
and
Weizmann Institute of Science, Rehovot, Israel

Communicated by Goldreich.

Received 16 June 2008
Online publication 4 January 2013

Abstract. Interactive hashing, introduced by Naor, Ostrovsky, Venkatesan, and Yung (J. Cryptol. 11(2):87–108, 1998), plays an important role in many cryptographic protocols. In particular, interactive hashing is a major component in all known constructions of statistically hiding commitment schemes and of statistical zero-knowledge arguments based on general one-way permutations/functions. Interactive hashing with respect to a one-way function f is a two-party protocol that enables a sender who knows $y = f(x)$ to transfer a random hash $z = h(y)$ to a receiver such that the sender is committed to y: the sender cannot come up with x and x' such that $f(x) \neq f(x')$, but $h(f(x)) = h(f(x')) = z$. Specifically, if f is a permutation and h is a two-to-one hash function, then the receiver does not learn which of the two preimages $\{y, y'\} = h^{-1}(z)$ is the one the sender can invert with respect to f. This paper reexamines the notion of interactive hashing, and proves the security of a variant of the Naor et al. protocol, which yields a more versatile interactive hashing theorem. When applying our new proof to (an equivalent variant of) the Naor et al. protocol, we get an alternative proof for this protocol that seems simpler and more intuitive than the original one, and achieves better parameters (in terms of how security preserving the reduction is).

Key words. Cryptography, Interactive hashing, Statistically hiding and computationally binding commitments, Statistical zero-knowledge arguments.

* Preliminary version in [6].

† I. Haitner was supported by The Israeli Centers of Research Excellence (I-CORE) program, (Center No. 4/11). Part of his work was done while at Weizmann Institute of Science.

1. Introduction

In an interactive hashing protocol introduced by Naor et al. [16], the sender S transfers to the receiver R the "hash value" $h(y)$ of y, where the "hash function" h is chosen (at random) by the receiver from a predetermined function family. The protocol is required to be *binding* in the sense that S is bounded by the protocol to at most one value of y. This binding requirement can hold is several ways: clearly, binding holds if after the interaction ends there is only a single element y that is consistent with the hash value. Protocols with this strong binding property are known as "information theoretic" interactive hashing. In contrast, in the computational setting we are interested in the case where a random h *does* induce many collisions. Thus, we only require the binding property to hold against efficient senders. Assuming that h is taken (at random) from a family of collision resistant hash functions,[1] the binding property is immediate. In this paper we do not rely on such families, as we do not want to assume their existence. Following [16], we enforce the binding by asking the sender to provide additional information about y (typically, the honest sender gets this additional information as part of its input).

We formally define the computational binding property in Sect. 3, but in the meanwhile let us consider the following important example: let f be a one-way permutation and view the committed value y as an image of f. For the purpose of binding, we can now require the sender to provide x such that $y = f(x)$ is consistent with the transcript (i.e., $h(y) = z$, where R's output equals (h, z)). Thus, for breaking the binding of the protocol a cheating sender needs not only output $y_1 \neq y_2$ such that $h(y_1) = h(y_2) = z$, but it is also required to output x_1 and x_2 with $f(x_1) = y_1$ and $f(x_2) = y_2$. Indeed, using this additional requirement [16] constructs an interactive hashing protocol that allows collisions, but is nevertheless binding (see Sect. 1.1 for more details).

Connection to Statistically Hiding Commitments Interactive hashing (in the flavor mentioned above) is closely related and to a large extent motivated by the fundamental notion of statistically hiding (and computationally binding) commitments. Statistically hiding commitment schemes are used as building blocks in constructions of statistical zero-knowledge arguments [1,16] and of certain coin-tossing protocols [14]. Naor et al. [16] use their interactive hashing protocol, hereafter the NOVY protocol, in order to construct statistically hiding commitments based on any one-way permutation. Haitner et al. [8] generalized their result by showing that the NOVY protocol can be used to construct statistically hiding commitments based on regular one-way functions (and also on the so called approximable-preimage-size one-way functions). Finally, Haitner et al. [9] constructed statistically hiding commitments based on *any* one-way function.[2] Not surprisingly, interactive hashing is heavily used in the underlying commitment scheme of [9].

A possible drawback of [9] is that their construction is rather inefficient and complex. Indeed, a major motivation for looking into interactive hashing is to simplify [9]. Such

[1] $\mathcal{H}$ is a family of collision resistant hash functions if given a random $h \in \mathcal{H}$, it is infeasible to find $x_1 \neq x_2$ with $h(x_1) = h(x_2)$.

[2] [9] is the full version that corresponds to Nguyen et al. [17] and Haitner and Reingold [5]. Reference [17] is independent of this work and [5] is subsequent to both [17] and this work.

a simplification was recently given by Haitner et al. [10], critically using the results we present here.

Before discussing our results and their applications, let us have a closer look into the notion of interactive hashing.

1.1. *Interactive Hashing in the Setting of One-Way Permutations*

Let $f\colon \{0,1\}^n \mapsto \{0,1\}^n$ be a one-way permutation and consider the following two-party protocol between a sender S, getting as input $x \in \{0,1\}^n$, and a receiver R. The receiver selects a random (almost) pairwise-independent two-to-one hash function $h\colon \{0,1\}^n \mapsto \{0,1\}^{n-1}$,[3] and sends its description to S. In return, S sends $z = h(y)$ back to R, where $y = f(x)$. Note that if both parties follow the protocol, then the following binding property is guaranteed: it is not feasible for S to find $x' \in \{0,1\}^n$ such that $f(x') \neq f(x)$ but $h(f(x')) = h(f(x)) = z$, although (exactly one) such element x' does exist. This is since the task of finding such x' can be shown (as firstly done in [15]) to be equivalent in hardness to inverting f on a random image (whereas the latter task is assumed to be hard by the one-wayness of f).

What happens, however, if S selects x only *after* seeing h? In such a case, it is quite plausible that S would be able to "cheat" by producing $x, x' \in \{0,1\}^n$ such that $f(x) \neq f(x')$ but $h(f(x')) = h(f(x)) = z$.[4] The NOVY interactive hashing protocol prevents such cheating. For that, it employs a specific family of hash functions such that each one of its functions h can be decomposed into $(n-1)$ Boolean functions $h_1, \ldots, h_{n-1}$, where $h(x) = h_1(x), \ldots, h_{n-1}(x)$. In the NOVY protocol, instead of sending h at once as described above, the protocol proceeds in rounds such that R sends a single Boolean function h_i in each round, and in return S sends a bit z_i, which is supposed to equal $h_i(f(x))$. Intuitively, a cheating sender has a significantly smaller leeway for cheating as it can no longer wait in selecting x till it receives the entire description of h. Still, it is highly non-trivial to argue that restricting the sender by adding interaction as above, is sufficient in order to prevent the sender from cheating. Nevertheless, Naor et al. [16] have shown that their protocol is binding even against a cheating sender (namely, even a cheating sender cannot produce $x, x' \in \{0,1\}^n$ such that $f(x) \neq f(x')$, but $h(f(x')) = h(f(x)) = z$).

1.1.1. *Application to Perfectly Hiding Commitments*

Naor et al. [16] used their protocol to construct perfectly hiding commitment scheme from one-way permutations, by employing the protocol with a uniformly chosen x as the sender's input. Let $y_0 < y_1$ be the two preimages of the hash value determined by the protocol (i.e., $y_0, y_1 \in h^{-1}(z)$) and let $i \in \{0,1\}$ be such that $f(x) = y_i$. The sender commits to a bit $b \in \{0,1\}$ by masking it with i (i.e., by sending $c = i \oplus b$ to the

[3] I.e., for every $x \neq x' \in \{0,1\}^n$, the distribution $(h(x), h(x'))$ induced by uniformly selecting h from the family, is close to being uniform over $\{0,1\}^{n-1} \times \{0,1\}^{n-1}$.

[4] Assume for example that the one-way permutation equals the identity function on the set T of all strings that start with $n/4$ zeros (where n is the input length). Now given a hash function h, all that the cheating sender has to do is to find a collision $y_1 \neq y_2 \in T$ with $h(y_1) = h(y_2)$. Such a collision is likely to exist by the birthday paradox, and for many families of hash functions (e.g., linear mappings) finding such a collision is easy.

receiver). In order to decommit, the sender sends x to the receiver, who sets $b = c \oplus i$, where i is as above (e.g., the index of $f(x)$ in $\{y_0, y_1\}$). The above scheme is perfectly hiding since the hash functions used are two-to-one, where the binding property of the scheme easily follows by the binding of the NOVY protocol.

1.2. *Interactive Hashing in the "Sparse Case"*

How about constructing statistically hiding commitments from, say, regular one-way functions (i.e., every possible output has the same number of preimages)? In such a case, we would like to interactively hash a value y that is taken from $\mathrm{Im}(f)$, the image of f over $\{0,1\}^n$, and not from $\{0,1\}^n$ as in the case of one-way permutations.

Notice that the NOVY theorem guarantees that when hashing y with $h\colon \{0,1\}^n \mapsto \{0,1\}^{n-1}$, the sender is committed to a single value y (even though $h^{-1}(h(y)) \cap \mathrm{Im}(f)$ might not be a singleton). In the case the $\mathrm{Im}(f)$ is sparse (i.e., $|\mathrm{Im}(f)|/2^n = \mathrm{neg}(n)$), however, when h outputs so many bits then it is most likely that $h(y)$ *completely* determines y. Hence, we cannot hope to use such protocol to get a statistically hiding commitment scheme.

Facing the aforementioned difficulty, Haitner et al. [8] made the following observation: the binding of the NOVY protocol holds for every function f that it is hard to invert over the uniform distribution on $\mathrm{Im}(f)$. Furthermore, some weak hiding is guaranteed for every f such that $\mathrm{Im}(f)$ is "dense" in $\{0,1\}^n$ (i.e., of order $2^n/\mathrm{poly}(n)$). Equipped with this observation, [8] employ the NOVY protocol with length-preserving poly-to-one one-way function (i.e., each output has at most polynomial number of preimages in the image set of f) to get some weak form of statistically hiding commitments. which can later be amplified into a full-fledge statistically hiding commitments. To handle any regular one-way function, [8] applies additional layer of (non interactive) hashing to reduce to the dense case. This implies a construction of statistically hiding commitments from any regular one-way function with known image size. Interactive hashing in the sparse case arises in other works as well, most notably in the construction of statistical zero-knowledge arguments from any one-way function [10,17].

1.3. *Our Results*

We consider a variant of the NOVY protocol in which the special family of two-to-one hash functions used by [16] is replaced by any "product" of Boolean families of pairwise independent hash functions (i.e., $h(x) = (h_1(x), \ldots h_k(x))$, where $h_1 \ldots h_k$ are taken from such families). Our proof relies in part on the original proof due to [16], but still seems significantly simpler. The new theorem directly applies to the following "sparse case": let $f\colon \{0,1\}^n \mapsto \{0,1\}^n$ be an efficiently computable function and let $\mathcal{L} \subseteq \{0,1\}^n$ be sparse. Our theorem applies when hashing to roughly $s = \lfloor \log(|\mathcal{L}|) \rfloor$ bits. In particular, when h is taken from a family of hash functions $\mathcal{H}^s\colon \{0,1\}^n \mapsto \{0,1\}^s$ that is a product of s families of pairwise-independent Boolean hash functions, the above protocol possesses the following binding property: if f is *hard to invert on the uniform distribution over* $\mathcal{L}$, then no efficient sender S^* can find two elements $x, x' \in f^{-1}(\mathcal{L})$ such that $f(x) \neq f(x')$ but $h(f(x)) = h(f(x')) = z$ (where z is the value determined by the protocol as $h(y)$). As an easy corollary, we use the new theorem to derive a direct construction of statistically hiding commitment based on regular one-way functions of known regularity (and thus reprove [8]).

Our new proof can be easily used to derive a new proof for a close variant of the NOVY protocol introduced by [9], which like the original NOVY protocol uses two-to-one hash functions.

We also note that a product family of pairwise-independent hash functions is not regular (i.e., a function is regular, over a given domain, if all its images have the same number of preimages). As a result, protocols using such families seem only to be useful for obtaining *statistically* hiding commitments.

The parameters achieved by our proof are an improvement compared with the original ones: given an algorithm A that breaks the binding property with probability $\varepsilon_{\mathsf{A}}(n)$, we get an algorithm that inverts the one-way permutation in comparable time and with inverting probability $\varepsilon_{\mathsf{A}}^2(n)/\operatorname{poly}(n)$ (where n is the hash function input length). This is an improvement over the $\varepsilon_{\mathsf{A}}^{10}(n)/\operatorname{poly}(n)$ and the $\Omega(\varepsilon_{\mathsf{A}}(n)^3/\operatorname{poly}(n))$ reductions of [16] and [17], respectively, and is close to natural limitations of the proof technique (see discussion in Sect. 5).

Finally, we consider interactive hashing protocols that use hash functions of arbitrary output length, and not necessarily Boolean. Given a one-way function that is hard to invert over the uniform distribution by algorithms of running-time 2^{ℓ}, our approach yields an interactive hashing protocol of n/ℓ rounds. When applied to one-way permutations, the resulting protocol matches the recent black-box lower bounds of Wee [21] and Haitner et al. [7], and generalizes the one-way permutation-based $O(n/\log n)$-round protocol of [13].

1.4. *Related Work*

Independently of our work, Nguyen et al. [17] gave a new proof for the NOVY protocol. Their proof follows the proof of [16] more closely than ours, but still introduces various simplifications and parameter improvements. The main goal of their new proof was to generalize the protocol such that it allows hashing with a hash function that is poly-to-one (rather than two-to-one as in [16]). The result of [17] (and of [16]), however, do not extend to the sparse case settings we considered above, where this limitation seems inherent to their proof technique.

More recently, Haitner et al. [9] consider a variant of the NOVY protocol that uses a different type of two-to-one hash functions. Specifically, the functions induced by the families of full-rank matrices mapping $\{0,1\}^n$ to $\{0,1\}^{n-1}$, where the operation $h(x)$ is interpreted as $h \times x$. While such families provide the same hiding guarantee for the resulting protocol as the special two-to-one functions considered by [16,17], the advantage is that the binding property of the protocol can be easily reduced to that of a protocol using families of pairwise-independent hash functions. In particular, [9] show how to derive the security of this variant from our main result.

In this work we focus on security with respect to bounded senders and unbounded receivers. The setting where both the receiver and the sender are unbounded, called *information theoretic interactive hashing* (also known as, *interactive hashing for static sets*), and its applications (cf., [4,18–20]) are not treated by this work. See [9, Sect. 3.2] for more details regarding information theoretic interactive hashing and its connection to the computational setting.

1.5. *Proof Idea*

We outline our binding proof in the most basic setting where $f\colon \{0,1\}^n \mapsto \{0,1\}^n$ is a one-way permutation and $\mathcal{L} = \{0,1\}^n$. Let $x \in \{0,1\}^n$ be S's input and let $y = f(x)$. Our protocol consists of $(n-1)$ rounds, in each round R selects a random Boolean pairwise-independent hash function h_i and S replies with $z_i = h_i(y)$.

Assume there exists an algorithm S^* that plays the sender's role in the protocol and cheats with non-negligible probability. Namely, with such probability S^* outputs at the end of the protocol two elements $x_0, x_1 \in \{0,1\}^n$ such that $y_0 = f(x_0)$ is different from $y_1 = f(x_1)$, and both y_0 and y_1 are consistent with the transcript. Consider the following naive way one may try to invert f using S^*: given an input y, choose the hash functions $h_1, \ldots, h_{n-1}$ to be used by R at random, and return one of the two elements that S^* outputs in the end of the interaction with R. In the case that y is consistent with S^*'s answer, then we are in good shape: there should be only few elements that are consistent with S^*'s answers, and therefore with good probability either y_0 or y_1 output by S^*'s is equal to y. Unfortunately, the probability of y to be consistent with S^*'s answers is very small; at each round, the probability that y is consistent with S^* is typically bounded by $\frac{1}{2}$.

The above motivates the following reduction: in the ith round feed S^* keep sampling a random hash function h_i, until its answer is consistent with y. If S^* behaves randomly, or acts according to the honest strategy with respect to some fixed input y', then no more than few such attempts are required for each round. For arbitrary adversaries, however, the above analysis seems to fail; as the process of choosing the h_i's proceeds, the distribution of y (chosen at random from the image of f) gets further away from being uniform among the elements that are consistent with S^*'s answers.

The actual reduction (following [16]) interpolates the two: we use the second reduction for the first t rounds, and the first reduction for the last $n-t$ rounds, where $t = n - O(\log n)$ is carefully chosen to circumvent both the above obstacles.[5] The novelty in our new proof, highlighted below, is in the way we analyze the success probability of this combined reduction.

Analyzing the Reduction Success Probability Given a vector of Boolean hash functions $\overline{h} = (h_1, \ldots, h_t)$, let $\text{Consist}(\overline{h}) := \{y' \in \{0,1\}^n \colon \forall i \in [t]\ \ h_i(y') = \mathsf{S}^*(h_1, \ldots, h_i)\}$. (I.e., $\text{Consist}(\overline{h})$ is the subset of elements inside $\{0,1\}^n$ that are consistent with S^*'s responses on $\overline{h}$.) Let D_{Real} be the distribution induced by the first part of the above reduction on the tuple $(\overline{h}, y)$. That is, y is uniformly chosen in $\{0,1\}^n$, and then $\overline{h}$ is chosen using rewinding so that $y \in \text{Consist}(\overline{h})$. Given this formulation, the reduction success probability is simply the success probability of the following algorithm (hereafter, the Inverter) over D_{Real}: given a tuple $(\overline{h}, y)$, choose $\overline{h}' = (h_{t+1}, \ldots, h_{n-1})$ at random, and return $f^{-1}(y)$ if it is one of the two outputs of S^* on $(\overline{h}, \overline{h}')$.

To analyze the above success probability we introduce the distribution $\mathsf{D}_{\text{Ideal}}$: first $\overline{h} = (h_1, \ldots, h_t)$ is chosen at random, and only then a random element y is uniformly drawn from $\text{Consist}(\overline{h})$. Since the distribution of $\overline{h}$ under $\mathsf{D}_{\text{Ideal}}$ is as induced by a random

[5] Actually, the fact that the combined reduction works, yields the result that the second reduction also works (see Remark 4.6). Still, following [16], we use this distinction to simplify the presentation.

interaction of S^* with R (i.e., uniform at random), it is easy to see that $\mathsf{Inverter}$ does well on $\mathsf{D}_{\text{Ideal}}$ (roughly $\varepsilon/|\text{Consist}(\overline{h})|$, for a uniformly chosen $\overline{h}$). We would then have concluded the proof if we could have proved that the statistical difference between $\mathsf{D}_{\text{Ideal}}$ and D_{Real} is small enough. It turns out, however, that such a strong bound is unlikely to hold as it would imply that one-way functions do not exist![6]

We do mange to prove, however, that $\mathsf{D}_{\text{Ideal}}$ is almost "dominated" by D_{Real}: the mass that $\mathsf{D}_{\text{Ideal}}$ assigns to most tuples is not too much larger (multiplicatively) than their mass under D_{Real}. This observation turns out to be sufficient, since when taking into account the full power of S^* (i.e., that it finds two consistent outputs) we prove that $\mathsf{Inverter}$ does "well" on most tuples in the support of $\mathsf{D}_{\text{Ideal}}$. Combining the above observations, it follows that $\mathsf{Inverter}$ does well also on D_{Real}.

1.6. *Paper Organization*

General notation and definitions used throughout the paper are given in Sect. 2. In Sect. 3, we formally define interactive hashing, present the NOVY paradigm for such protocols, and state our main theorem regarding the binding of such protocols. The proof of this theorem is given in Sect. 4, where discussion and further issues appear in Sect. 5. Finally, in Appendix A we show how to use our new theorem to derive a direct construction of statistically hiding commitment scheme based on regular one-way functions.

2. Preliminaries

2.1. *Notation*

We use calligraphic letters to denote sets, uppercase for random variables, and lowercase for values. For $n \in \mathbb{N}$, let $[n] = \{1, \ldots, n\}$. Given a binary relation $\mathcal{W} \subseteq \mathcal{D}_1 \times \mathcal{D}_2$ and $y \in \mathcal{D}_2$, let $\mathcal{W}_y = \{x \in \mathcal{D}_1 \colon (x, y) \in \mathcal{W}\}$. Let PPTM denote a probabilistic algorithm (i.e., Turing machines) that runs in *strict* polynomial time and let poly denote the family of polynomials (we sometimes abuse notation and view poly as an arbitrary polynomial). Throughout we identify functions with their description, and assume without loss of generality that such a description is a binary string.

Given a random variable X, let $x \leftarrow X$ denote that x is selected according to X. Similarly given a finite set $\mathcal{S}$, let $s \leftarrow \mathcal{S}$ denote that s is selected according to the uniform distribution on $\mathcal{S}$. We adopt the convention that when the same random variable occurs several times in an expression, all occurrences refer to a single sample. For example, $\Pr[f(X) = X]$ is defined to be the probability that when $x \leftarrow X$, we have $f(x) = x$. Given a distribution D over a set $\mathcal{S}$, the support of D is defined as $\text{Supp}(\mathsf{D}) := \{s \in \mathcal{S} \colon \mathsf{D}(s) > 0\}$ and its min entropy, denoted $\mathrm{H}_\infty(\mathsf{D})$, is defined as $\min_{x \in X} \log \frac{1}{\mathsf{D}(x)}$. Finally, the statistical distance of two distributions P and Q over a final set $\mathcal{U}$, denoted $\mathsf{SD}(\mathsf{P}, \mathsf{Q})$, is defined as $\frac{1}{2} \sum_{u \in \mathcal{U}} |\mathsf{P}(u) - \mathsf{Q}(u)|$.

[6] Up to this point we did not use the fact that S^* finds two different outputs of f that are consistent with the protocol rather than a single output (and for that purpose, S^* is not more useful than the honest sender). If the statistical difference between $\mathsf{D}_{\text{Ideal}}$ and D_{Real} would have been small enough, we could deduce that the above (efficient) procedure when applied to the honest sender, inverts f with noticeable probability.

An *interactive protocol* (A, B) consists of two interactive algorithms (touring machines) that compute the next-message functions of the (honest) parties in the protocol. Let $(\mathsf{A}(a), \mathsf{B}(b))(x)$ denote the random process obtained by having A and B interact on common input x, with (private) auxiliary inputs a and b to A and B, respectively, and with independent random coin tosses for A and B. The protocol (A, B) runs in polynomial time, if there is a polynomial p such that A halts in $(\mathsf{A}(\cdot), \cdot)(x)$, and similarly B halts in $(\cdot, \mathsf{B}(\cdot))(x)$, after at most $p(|x|)$ rounds for all possible input $x \in \{0,1\}^*$ (regardless of its private input and the other party strategy). Let $\text{view}_{\mathsf{A}}(\mathsf{A}(a), \mathsf{B}(b))(x)$ denotes A's *view* of the interaction, i.e., its values are transcripts $(\gamma_1, \gamma_2, \ldots, \gamma_t; r)$, where the γ_i's are all the messages exchanged and r is A's coin tosses. Similarly, $\text{view}_{\mathsf{B}}(\mathsf{A}(a), \mathsf{B}(b))$ denotes B's view.

2.2. *Efficient Function Families*

To be useful in applications, ensembles of function families are typically required to be "efficient". For our needs, efficiency means the following.

Definition 2.1 (Efficient ensembles of function families). Let $\mathcal{H} = \{\mathcal{H}_n\}_{n\in\mathbb{N}}$ be an ensemble of function families mapping strings of length n to strings of length $\ell(n)$. The ensemble $\mathcal{H}$ is efficient, if the following hold:

Samplable. There exists a PPTM that, given 1^n, returns a uniform element in $\mathcal{H}_n$.

Efficient. There exists a polynomial-time algorithm that given $x \in \{0,1\}^n$ and a description of $h \in \mathcal{H}_n$, outputs $h(x)$.

Verifiable. There exists a polynomial-time algorithm that given $h \in \{0,1\}^*$ and 1^n, outputs '1' iff $h \in \mathcal{H}_n$.

Throughout we use the shorthand "efficient function families" for "efficient ensembles of function families".

2.3. *Pairwise Independent Function Families*

Definition 2.2 (Pairwise-independent families). Let $\mathcal{H}$ be a function family mapping strings of length n to strings of length ℓ. The family $\mathcal{H}$ is pairwise independent, if

$$\Pr_{h\leftarrow\mathcal{H}}\left[h(x_1) = y_1 \wedge h(x_2) = y_2\right] = 2^{-2\ell}$$

for every distinct $x_1, x_2 \in \{0,1\}^n$ and every $y_1, y_2 \in \{0,1\}^\ell$.

It is well known Carter and Wegman [2] that for every polynomial-time computable $\ell(n) \leq \text{poly}(n)$, there exists an *efficient* family of pairwise-independent hash functions with description size $O(\max\{n, \ell(n)\})$.

The following standard lemma states that a random pairwise-independent hash function partitions a given set into (almost) equal size subsets.

Lemma 2.3. *Let $\mathcal{H}$ be a pairwise-independent function family mapping strings of length n to strings of length ℓ, let $\mathcal{L} \subseteq \{0,1\}^n$, let $\mu = |\mathcal{L}|/2^\ell$ and let $\delta > 0$. Then*

$$\Pr_{h\leftarrow\mathcal{H}}\left[\exists y \in \{0,1\}^\ell \ : \ \left|\left|\{x \in \mathcal{L} : h(x) = y\}\right| - \mu\right| > \delta\mu\right] < \frac{2^\ell}{\delta^2\mu}.$$

Proof. Fix $y \in \{0,1\}^\ell$ and let H be uniformly chosen in $\mathcal{H}$. For $x \in \mathcal{L}$, define the indicatory random variable I_x to be one iff $H(x) = y$, and let $X = \sum_{x\in\mathcal{L}} I_x$. Since $\mathrm{E}[I_x] = 2^{-\ell}$ for every $x \in \mathcal{L}$, it follows that $\mathrm{E}[X] = \mu$.

Note that $\mathrm{Var}[I_x] = \mathrm{E}[I_x^2] - \mathrm{E}[I_x]^2 = \mathrm{E}[I_x](1 - \mathrm{E}[I_x]) \le \mathrm{E}[I_x]$ for every $x \in \mathcal{L}$, where the pairwise independence of the I_x's[7] yields $\mathrm{Var}[X] = \mathrm{Var}[\sum_{x\in\mathcal{L}} I_x] = \sum_{x\in\mathcal{L}} \mathrm{Var}[I_x] \le \sum_{x\in\mathcal{L}} \mathrm{E}[I_x] = \mathrm{E}[X]$. Applying the Chebyshev inequality

$$\Pr\Big[\big|X - \mathrm{E}[X]\big| > \delta \cdot \mathrm{E}[X]\Big] < \frac{1}{(\delta \cdot \mathrm{E}[X])^2} \cdot \mathrm{Var}[X] \le \frac{\mathrm{E}[X]}{(\delta \cdot \mathrm{E}[X])^2} = \frac{1}{\delta^2 \mu},$$

and a union bound completes the proof. □

2.4. *Linear Transformations*

Other function families of interest are those of a linear transformation.

Definition 2.4 (Linear transformations). Given $n, \ell \in \mathbb{N}$, $M \in \{0,1\}^{\ell\times n}$ and $x \in \{0,1\}^n$, let $M(x) := M \times x \bmod 2$, and let $\mathrm{Lin}_{\ell,n}$ be the function family defined by all matrices in $\{0,1\}^{\ell\times n}$ with respect to the above operation. We let $\mathrm{Full}_{\ell,n} \subseteq \mathrm{Lin}_{\ell,n}$ be the function family defined by all *full-rank* matrices in $\{0,1\}^{\ell\times n}$.

Note that both $\mathrm{Lin}_{\ell,n}$ and $\mathrm{Full}_{\ell,n}$, are efficient families for any polynomial-time computable $\ell = \ell(n) < \mathrm{poly}(n)$. We use the following fact.

Fact 2.5. *There exists a constant $c > 0$ such that $\frac{|\mathrm{Full}_{\ell,n}|}{|\mathrm{Lin}_{\ell,n}|} > c$ for any integers $\ell \le n$.*

Proof. The probability that a random vector in $\{0,1\}^n$ is in the span of some $k < n$ vectors in $\{0,1\}^n$ (over $\mathbb{F}_2$), is bounded by 2^{k-n}. It follows that

$$\begin{aligned}\Pr_{M\leftarrow\mathrm{Lin}_{\ell,n}}[M \in \mathrm{Full}_{\ell,n}] &\ge \prod_{k=1}^{\ell}\left(1 - 2^{k-n-1}\right)\\ &> c := \lim_{t\to\infty}\prod_{k=1}^{t}\left(1 - 2^{-k}\right)\\ &> 1 - \lim_{t\to\infty}\sum_{k=1}^{t} 2^{-k} = 0.\end{aligned}$$

□

2.5. *Piece-Wise Functions*

In the interactive hashing protocols considered below, the receiver sends the function description in pieces, where each such piece suffices for evaluating the output it contributes.

Definition 2.6. Given a sequence of functions $\overline{h} = (h_1, \ldots, h_s)$ defined over $\{0,1\}^n$, let $\overline{h}(x) = h_1(x) \circ \cdots \circ h_s(x)$, where $\circ$ denotes string concatenation. A family of length s function sequences, is called s-piece function family.

[7] A set of random variables $\{\mathcal{S}_i\}$ over $\mathcal{U}$ is *pairwise independent*, if $\Pr[S_i = u_i \wedge S_j = u_j] = \Pr[S_i = u_i] \cdot \Pr[S_j = u_j]$ for any $i \ne j$ and $u_i, u_j \in \mathcal{U}$.

An ensemble $\overline{\mathcal{H}} = \{\overline{\mathcal{H}}_n\}$ of $s = s(n)$-piece function families is called prefix verifiable, if there exists a polynomial-time algorithm that given $(1^n, h_1, \ldots, h_i)$ returns '1' iff there exists $h_{i+1}, \ldots, h_s$ such that $(h_1, \ldots, h_s) \in \overline{\mathcal{H}}_n$.

In this paper we consider two kinds of prefix verifiable piece-wise family. The first type is a product family ensemble $\mathcal{H}^s$, where $\mathcal{H}$ is an efficient function family and s is polynomial-time computable integer-valued function. It is easy to verify that $\mathcal{H}^s$ is indeed an efficient prefix verifiable s-piece family.

The second type is a variant of linear maps; given a subset $\mathcal{H}$ of $\mathrm{Lin}_{\ell,n}$ and an index set $\bar{i} = (i_1, \ldots, i_s)$ with $1 \leq i_1 < \cdots < i_s = \ell$, let $\mathcal{H}_{\bar{i}}$ be the s-piece function family

$$\left\{ \begin{pmatrix} h_1 \\ \vdots \\ h_{i_1} \end{pmatrix}, \ldots, \begin{pmatrix} h_{i_{s-1}+1} \\ \vdots \\ h_{i_s} \end{pmatrix} : h = \begin{pmatrix} h_1 \\ \vdots \\ h_\ell \end{pmatrix} \in \mathcal{H} \right\},$$

letting

$$\begin{pmatrix} h_{i_{j-1}+1} \\ \vdots \\ h_{i_j} \end{pmatrix}(x) = \begin{pmatrix} h_{i_{j-1}+1} \\ \vdots \\ h_{i_j} \end{pmatrix} \times x.$$

For polynomial-time computable index-set function $\bar{i} = \bar{i}(n) = (i_1(n), \ldots, i_{s(n)}(n))$ and integer-valued function $\ell = \ell(n) < \mathrm{poly}(n)$, it is easy to verify that the s-piece function family ensemble $(\mathrm{Lin}_{\ell(n),n})_{\bar{i}(n)}$ and $(\mathrm{Full}_{\ell(n),n})_{\bar{i}(n)}$ are both efficient and prefix verifiable.

3. Interactive Hashing

Following [17], we state our result in the language of binary relations, where these relations are not assumed to have efficient deciders. Since every function f defines the binary relation $\{(x, f(x)) : x \in \{0,1\}^*\}$, our result yields an analogous result for functions.

Definition 3.1 (Interactive hashing protocols [16,17]). Let $\mathcal{H}$ be an ensemble of function families mapping strings of length n to strings of length $\ell = \ell(n)$. An $\mathcal{H}$-interactive hashing protocol is a polynomial-time protocol (S, R) such that the following holds: the parties receive the security parameter 1^n as a common input, and S gets $y \in \{0,1\}^n$ as a private input. At the end, R outputs $(h, z) \in \mathcal{H} \times \{0,1\}^\ell$.

We make the following correctness requirement: for all $n \in \mathbb{N}$, $y \in \{0,1\}^n$ and a pair (h, z) that may be output by $(\mathsf{S}(1^n, y), \mathsf{R}(1^n))$, it is the case that $h(y) = z$.

An interactive hashing protocol is said to be (T, δ)-binding if the following holds.

Definition 3.2. Let $T : \mathbb{N} \mapsto \mathbb{N}$ and $\delta : \mathbb{N} \times \mathbb{R} \times [0,1] \mapsto [0,1]$. An $\mathcal{H}$-interactive hashing protocol (S, R) is said to be (T, δ)-binding, if there exists an oracle-aided algorithm

PPTM A such that the following hold for any $n \in \mathbb{N}$ and any adversary S^*. First, the running time of $\mathsf{A}^{\mathsf{S}^*}(y)$ is bounded by $T(|y|)$ (counting oracle calls as a single operation). Second, let (h, z) and $((x_0, y_0), (x_1, y_1))$ denote the common output and S^*'s private output in $(\mathsf{S}^*, \mathsf{R})(1^n)$, respectively, then for any set $\mathcal{L} \subseteq \{0,1\}^n$ and any binary relation $\mathcal{W} \subseteq \{0,1\}^* \times \mathcal{L}$, if

$$\Pr\left[h(y_0) = h(y_1) = z \wedge y_0 \neq y_1 \wedge (x_0, y_0), (x_1, y_1) \in \mathcal{W}\right] \geq \varepsilon, \tag{1}$$

then

$$\Pr_{y \leftarrow \mathcal{L}}\left[\mathsf{A}^{\mathsf{S}^*}(y) \in \mathcal{W}_y\right] \geq \delta\left(n, \frac{|\mathcal{L}|}{2^\ell}, \varepsilon\right). \tag{2}$$

Note that the above definition is *black-box* (i.e., the security reduction is a uniform algorithm that accesses the adversary as an oracle). By considering such a definition, however, we only strengthen our positive results. Also note that a $(\mathrm{poly}(n), \mathrm{poly}(\varepsilon)/\mathrm{poly}(n, \frac{|\mathcal{L}|}{2^\ell}))$-binding protocol is "polynomially secure" for the set ensemble $\mathcal{L} = \mathcal{L}(n)$ with $\frac{|\mathcal{L}|}{2^\ell} \in \mathrm{poly}(n)$—on security parameter 1^n, no PPTM breaks the binding with more than negligible probability.

The above correctness and binding definitions are oblivious to the actual implementation of the protocol. Since Definition 3.1 does not specify the amount of information that might leak to the (possibly cheating) receiver, stating a similar hiding property is more challenging. Thus, rather than giving a generic (and hard to digest) hiding definition, we separately analyze the hiding guarantee achieved by each of the different constructions considered below.

3.1. *The NOVY Paradigm*

All the interactive hashing protocols considered in this paper follows the NOVY paradigm, which is a natural generalizations of the NOVY protocol. The NOVY paradigm instantiated with an s-piece family $\overline{\mathcal{H}}$ over strings of length n, denoted $\mathrm{NOVY}\langle\overline{\mathcal{H}}\rangle$, is defined as follows:

Protocol 3.3 $\mathrm{NOVY}\langle\overline{\mathcal{H}}\rangle$.

Common input: 1^n.
S's input: $y \in \{0,1\}^n$.

1. R *chooses uniformly at random* $\overline{h} = (h_1, \ldots, h_s) \in \overline{\mathcal{H}}$.
2. *Do for* $i = 1$ *to* s:
   (a) R *sends* h_i *to* S.
   (b) S *aborts if* $(h_1, \ldots, h_i)$ *is not a prefix of some element in* $\overline{\mathcal{H}}$.
   Otherwise, S *sends* $z_i = h_i(y)$ *back to* R.
3. R *outputs* $(\overline{h}, \overline{z} = (z_1, \ldots, z_s))$.

Typically, we instantiated the NOVY paradigm with efficient, prefix-verifiable, s-piece families. It is straightforward that for such families, the resulting protocol is an $\overline{\mathcal{H}}$-interactive hashing.

3.1.1. *Hiding of the NOVY Paradigm*

The above definition stipulates that the only information a cheating receiver gets from an execution is $(h, h(y))$ for some $h \in \overline{\mathcal{H}}$. While this h might be chosen *adaptively* by the cheating receiver, we still have the following guarantee.

Claim 3.4. *Let $\overline{\mathcal{H}}$ be an s-piece function family mapping strings of length n to strings of length ℓ and let $(\mathsf{S}, \mathsf{R}) = \mathrm{NOVY}\langle\overline{\mathcal{H}}\rangle$. Let $\mathcal{L} \subseteq \{0,1\}^n$ and let $q = \lfloor \log|\mathcal{L}| \rfloor - \ell$. Then for any cheating adversary R^* and $\delta \in (0,1]$, we have*

$$\Pr_{v \leftarrow V_{\mathsf{R}^*}}\left[\mathrm{H}_\infty(Y \mid V_{\mathsf{R}^*} = v) < q + \log\delta\right] < \delta,$$

where Y is uniformly distributed over $\mathcal{L}$ and V_{R^*} is R^*'s view in $(\mathsf{S}(Y), \mathsf{R}^*)$.

Proof. Since S refuses to send more than ℓ bits of information about its input, the proof follows using a straightforward counting argument. □

When limiting our attention to product of pairwise-independent function families and semi-honest receivers (ones that follow the prescribed protocol), we have the following guarantee:

Claim 3.5. *Let $\overline{\mathcal{H}}$ be an s-piece pairwise-independent function family mapping strings of length n to strings of length ℓ, and let $(\mathsf{S}, \mathsf{R}) = \mathrm{NOVY}\langle\overline{\mathcal{H}}\rangle$. Let $\mathcal{L} \subseteq \{0,1\}^n$ and let $q = \lfloor \log|\mathcal{L}| \rfloor - \ell$. Then*

$$\Pr_{v \leftarrow V_{\mathsf{R}}}\left[\mathrm{H}_\infty(Y \mid V_{\mathsf{R}} = v) < q - 1\right] \leq \frac{2^\ell}{2^{q-3}},$$

where Y is uniformly distributed over $\mathcal{L}$ and V_{R} is R's view in $(\mathsf{S}(Y), \mathsf{R})$.

Proof. Immediately follows by Lemma 2.3 (taking $\delta = \frac{1}{2}$). □

Finally, for the family $\mathrm{Full}_{\ell,n}$ and the set $\mathcal{L} = \{0,1\}^n$, we have the following "perfect" hiding guarantee (formally stated and proved below): (1) the input y of S is perfectly hidden among (at least) $2^{n-\ell}$ other values in $\{0,1\}^n$, and (2) the "index" of y among these values is efficiently computable from y. In the interesting case of $n - \ell = 1$, it follows that the index of y is a uniform bit from the receiver point of view. Such an hiding guarantee is equivalent to that achieved by the NOVY protocol, and in particular can be use to construct perfectly hiding commitment schemes from one-way permutations.

Claim 3.6. *There exists a deterministic polynomial-time computable mapping M such that the following holds: let $(\mathrm{Full}_{\ell,n})_{\bar{i}}$ be an s-piece function family, and let $(\mathsf{S}, \mathsf{R}) = \mathrm{NOVY}\langle(\mathrm{Full}_{\ell,n})_{\bar{i}}\rangle$. Then for any cheating adversary R^*, the distributions $(V_{\mathsf{R}^*}(y), \mathsf{M}(\ell, T(y), y))_{y \leftarrow \{0,1\}^n}$ and $(V_{\mathsf{R}^*}(y), z)_{y \leftarrow \{0,1\}^n, z \leftarrow \{0,1\}^{n-\ell}}$ are* identical, *where (the jointly distributed) $T(y)$ and $V_{\mathsf{R}^*}(y)$, are the transcript and R^*'s view, respectively, in $(\mathsf{S}(y), \mathsf{R}^*)$.*

Proof. We assume for ease of notation that R^* never causes the sender to abort. Given a transcript $t = (\overline{h}, \overline{z}) \in (\mathrm{Full}_{\ell,n})_{\overline{i}} \times \{0,1\}^{\ell}$ of $(\mathsf{S}, \mathsf{R}^*)$ and an input $y \in \{0,1\}^n$, algorithm M finds $\overline{h}' \in \mathrm{Lin}_{n-\ell,n}$ such that the matrix $\binom{\overline{h}}{\overline{h}'}$ is of full rank n,[8] and outputs $\overline{h}'(y)$. Conditioned on R^*'s view, the sender's input (i.e., y) is uniformly distributed in the $(n-\ell)$-dimensional affine subspace $\{y' : \overline{h}(y) = \overline{z}\}$. Hence, $\mathsf{M}(\ell, t, y)$ is uniformly distributed in $\{0,1\}^{n-\ell}$. □

3.1.2. *Binding of the NOVY Paradigm*

The following theorem, whose proof is given in Sect. 4, is the main contribution of this paper.

Theorem 3.7. *Let $\mathcal{H}$ be an efficient pairwise-independent function family mapping strings of length n to strings of length $\ell = \ell(n)$, and let $s = s(n)$ be a polynomial-time computable function. Then* $\mathrm{NOVY}\langle \mathcal{H}^s \rangle$ *is (T, δ)-binding, where $T(n) = p(n) \cdot 2^{\ell}$ for some $p \in \mathrm{poly}$ and $\delta(n, r, \varepsilon) \in \Omega(\varepsilon^2 \cdot \min\{1, 1/r\} \cdot \frac{1}{2^{10\ell} \cdot n^8})$.*

Theorem 3.7 also holds for function families of weaker independence guarantee.

Definition 3.8 (XOR-universal function families). Let $\mathcal{H}$ be a function family mapping strings of length n to strings of length ℓ. We say that $\mathcal{H}$ is XOR-universal if

$$\Pr_{h \leftarrow \mathcal{H}}\big[h(x_1) \oplus h(x_2) = y\big] = 2^{-\ell}$$

for any distinct $x_1, x_2 \in \{0,1\}^n$ and any $y \in \{0,1\}^{\ell}$.

We note that while not pairwise independent (maps 0^n to 0^{ℓ}), the family $\mathrm{Lin}_{\ell,n}$ is XOR-universal for every choice of ℓ and n.

Corollary 3.9. *Let T, δ, s and ℓ be as in Theorem 3.7. Then the following protocols are (T, δ)-binding*:

1. $\mathrm{NOVY}\langle \mathcal{H}^s \rangle$, *where $\mathcal{H}$ is an efficient XOR-universal function mapping strings of length n to strings of length $\ell = \ell(n)$.*
2. $\mathrm{NOVY}\langle (\mathrm{Full}_{s\cdot\ell,n})_{\ell,2\ell,\ldots,s\cdot\ell} \rangle$, *where $s \cdot \ell \le n$.*

Proof. The proof of the first part readily follows from the proof of Theorem 3.7. Yet for the sake of completeness we prove it below by reducing it to (the statement of) Theorem 3.7. Let S^* be an algorithm that breaks the binding of $\mathrm{NOVY}\langle \mathcal{H}^s \rangle$ with probability ε (according to Eq. (1)) and let $\mathcal{H}'$ be the pairwise independent family $\{(h, \overline{b}) : h \in \mathcal{H}, \overline{b} \in \{0,1\}^{\ell}\}$, where $(h, \overline{b})(x) := h(x) \oplus \overline{b}$. Consider the following algorithm that uses S^* to break the binding of $\mathrm{NOVY}\langle (\mathcal{H}')^s \rangle$: upon receiving the function $(h, \overline{b}) \in \mathcal{H}'$ from R, it sends h to S^*, XORs the answer of S^* with $\overline{b}$ and sends the result back to R. It is immediate that the above algorithm breaks the binding of $\mathrm{NOVY}\langle \mathcal{H}'^s \rangle$ with probability ε, and thus the proof of this part follows from Theorem 3.7.

[8] Note that $\overline{h}'$ can be found in deterministic polynomial time using Gaussian elimination.

Fact 2.5 yields the following any algorithm that breaks the binding of $\mathrm{NOVY}\langle(\mathrm{Full}_{s\cdot\ell,n})_{\ell,2\ell,\ldots,s\cdot\ell}\rangle$ with probability ε, breaks that of $\mathrm{NOVY}\langle(\mathrm{Lin}_{\ell,n})_{\ell,2\ell,\ldots,s\cdot\ell}\rangle$ with probability $\Omega(\varepsilon)$ (i.e., conditioned on an event of constant probability over the execution of $\mathrm{NOVY}\langle(\mathrm{Lin}_{s\cdot\ell,n})_{\ell,2\ell,\ldots,s\cdot\ell}\rangle$, the receiver's messages in $\mathrm{NOVY}\langle(\mathrm{Lin}_{s\cdot\ell,n})_{\ell,2\ell,\ldots,s\cdot\ell}\rangle$ are distributed exactly as the receiver's messages in $\mathrm{NOVY}\langle(\mathrm{Full}_{\ell,n})_{\ell,2\ell,\ldots,s\cdot\ell}\rangle$). Since $(\mathrm{Lin}_{s\cdot\ell},n)_{\ell,2\ell,\ldots,s\cdot\ell}$ is XOR-universal, the first part of Corollary 3.9 shows that $\mathrm{NOVY}\langle(\mathrm{Lin}_{\ell,n})_{\ell,2\ell,\ldots,s\cdot\ell}\rangle$ is (T,δ)-binding, and therefore so is $\mathrm{NOVY}\langle(\mathrm{Full}_{s\cdot\ell},n)_{\ell,2\ell,\ldots,s\cdot\ell}\rangle$. □

Remark 3.10 (Further extensions and the original NOVY protocol). Consider an s-piece function family $\overline{\mathcal{H}}$ over $\{0,1\}^n$, where each piece outputs ℓ bits. For $\overline{z}\in\{0,1\}^{k\ell}$ and a k-element vector $\overline{h}$ that is a prefix of some element in $\overline{\mathcal{H}}$, define $\mathrm{Consist}_{\overline{h},\overline{z}}:=\{x\in\{0,1\}^n:\overline{h}(x)=\overline{z}\}$. Assume that for any possible such pair, the family $\mathcal{H}_{\overline{h}}=\{h\colon(\overline{h},h)$ is a prefix of some element in $\overline{\mathcal{H}}\}$ is an efficient XOR-universal over $\mathrm{Consist}_{\overline{h},\overline{z}}$,[9] then the proof of Theorem 3.7 readily extends to the family $\overline{\mathcal{H}}$.

The above extension is of interest since it applies to the function family used by the original NOVY protocol. Since protocol $\mathrm{NOVY}\langle(\mathrm{Full}_{s\cdot\ell},n)_{\ell,2\ell,\ldots,s\cdot\ell}\rangle$ achieves the same hiding guarantee as the NOVY protocol does, we do not formally prove the above observation.

4. Binding Proof

Let $\mathcal{H}=\{\mathcal{H}_n\}_{n\in\mathbb{N}}$ be an efficient function family mapping strings of length n to strings of length $\ell(n)$, where $\ell(n)$ is an arbitrary integer-valued function, and let $(\mathsf{S},\mathsf{R})=\mathrm{NOVY}\langle\mathcal{H}^{s(n)}\rangle$, where s is a polynomial-time computable integer-valued function. For integer-valued function $t(n)\le s(n)$ to be determined by the analysis, we define the following oracle-aided algorithm.

Algorithm 4.1 (A).

Input: $y\in\{0,1\}^n$.
Oracle: S^*.

1. Let $\overline{h}\leftarrow\mathsf{Searcher}^{\mathsf{S}^*}(y)$.
2. Return $\mathsf{Inverter}^{\mathsf{S}^*}(\overline{h})$.

The oracle-aided algorithms Searcher and Inverter are defined as follows:

Algorithm 4.2 (Searcher).

Input: $y\in\{0,1\}^n$.
Oracle: S^*.

[9] That is, $\Pr_{h\leftarrow\mathcal{H}}[h(x_1)\oplus h(x_2)=y]=2^{-\ell}$ for any distinct $x_1,x_2\in\mathrm{Consist}_{\overline{h},\overline{z}}$ and any $y\in\{0,1\}^\ell$.

1. For $k = 1$ to $t(n)$ do:
   Do the following for $\lceil 2^{\ell(n)} \cdot \ln t(n) \rceil$ times:
   (a) Let $h_k \leftarrow \mathcal{H}_n$.
   (b) *Break* the inner loop if $\mathsf{S}^*(1^n, h_1, \ldots, h_k) = h_k(y)$, where $\mathsf{S}^*(1^n, h_1, \ldots, h_k)$ stands for S^* answer on input 1^n and receiver's messages $h_1, \ldots, h_k$.
   If the end of the inner loop has reached, output "Fail" and *abort* the execution.
2. Return $(h_1, \ldots, h_t)$.

Algorithm 4.3 (Inverter).

Input: $\overline{h} \in \mathcal{H}_n^{t(n)}$.
Oracle: S^*.

1. Let $\overline{h}' \leftarrow \mathcal{H}_n^{s(n)-t(n)}$.
2. Let $((x_0, y_0), (x_1, y_1))$ be the final output of $\mathsf{S}^*(1^n, (\overline{h}, \overline{h}'))$.
3. Return x_0 with probability $\frac{1}{2}$, and x_1 otherwise.

It is straightforward that $\mathsf{A}^{\mathsf{S}^*}$ runs in time $\text{poly}(n) \cdot 2^{\ell(n)}$ (counting an oracle call as a single operation). In the following we fix $n \in \mathbb{N}$, a set $\mathcal{L} \subseteq \{0,1\}^n$ and a binary relation $\mathcal{W} \subseteq \{0,1\}^* \times \mathcal{L}$. We also fix a cheating sender S^* that breaks the binding of $\text{NOVY}\langle \mathcal{H}^s \rangle$ with probability ε with respect to $\mathcal{W}$ and $\mathcal{L}$ (according to Definition 3.2). Namely,

$$\Pr_{(((x_0,y_0),(x_1,y_1)),(h,z)) \leftarrow (\mathsf{S}^*,\mathsf{R})(1^n)} \Big[h(y_0) = h(y_1) = z \wedge y_0 \neq y_1 \wedge (x_0, y_0), (x_1, y_1) \in \mathcal{W} \Big] = \varepsilon. \tag{3}$$

We assume without loss of generality that S^* is deterministic (the generalization to randomized adversaries is standard) and prove the theorem showing that

$$\Pr_{y \leftarrow \mathcal{L}} \left[\mathsf{A}^{\mathsf{S}^*}(y) \in \mathcal{W}_y \right] \geq \frac{c \cdot \varepsilon^2}{2^{10\ell} \cdot n^8} \cdot \max \left\{ 1, 2^{s\ell} / |\mathcal{L}| \right\} \tag{4}$$

for a universal constant $c > 0$.

Throughout the proof we omit n from the notations, and let A, $\mathsf{Searcher}$ and $\mathsf{Inverter}$, stand for $\mathsf{A}^{\mathsf{S}^*}$, $\mathsf{Searcher}^{\mathsf{S}^*}$ and $\mathsf{Inverter}^{\mathsf{S}^*}$, respectively. We assume without loss of generality that S^* always replies with valid messages (i.e., elements inside $\{0,1\}^\ell$). First time readers are encouraged to focus on the case where $\ell = 1$, $s = n$ and $\mathcal{L} = \{0,1\}^n$.

Following the intuition given in the introduction, we consider the success probability of S^* with respect to the following distributions:

- $\mathsf{D}_{\text{Real}} := (\overline{h}, y)_{y \leftarrow \mathcal{L}, \overline{h} \leftarrow \mathsf{Searcher}(y)}$, and
- $\mathsf{D}_{\text{Ideal}} := (\overline{h}, y)_{\overline{h} \leftarrow \mathcal{H}^t, y \leftarrow \text{Consist}(\overline{h})}$,

where $\text{Consist}(\overline{h}) := \{y \in \mathcal{L} : \overline{h}(y) = \mathsf{S}^*(\overline{h})\}$ (i.e., $\text{Consist}(\overline{h})$ is the set of elements that are consistent with S^*'s answers on $\overline{h}$).

Since D_{Real} is the distribution a random execution of A (over a random y) induces on the values of y and $\overline{h}$, the success probability of A, in satisfying $\mathcal{W}$, equals the

success probability of Inverter over D_{Real} (i.e., to $\Pr_{(\overline{h},y)\leftarrow\mathsf{D}_{\text{Real}}}[\text{Inverter}(\overline{h})\in\mathcal{W}_y]$). We lowerbound the latter probability by relating it to the success probability of Inverter over $\mathsf{D}_{\text{Ideal}}$. Specifically, we first show (Lemma 4.4) that D_{Real} and $\mathsf{D}_{\text{Ideal}}$ assign similar mass to *most* elements in the support of $\mathsf{D}_{\text{Ideal}}$, and then prove (Lemma 4.5) that Inverter's success probability over $\mathsf{D}_{\text{Ideal}}$ (in the task of satisfying $\mathcal{W}$) is not only high but also "well spread". Namely, even if we ignore the contribution to the success probability of some sufficiently small number of values in the support of $\mathsf{D}_{\text{Ideal}}$, this success probability remains high. Therefore, we are guaranteed that the success probability of Inverter is high with respect to *any* distribution that assigns about the same mass to *most* elements in the support of $\mathsf{D}_{\text{Ideal}}$, and in particular with respect to D_{Real}.

For $k\in\{0,\ldots,s\}$ let $\mu_k := |\mathcal{L}|/2^{k\ell}$.

Lemma 4.4. *Assuming $t\geq 4$ then*

$$\Pr_{\overline{h}\leftarrow\mathcal{H}^t}\left[\left|\left\{y\in\text{Consist}(\overline{h})\colon(\overline{h},y)\notin\text{Dominated}\right\}\right|>8t^4\cdot 2^{3\ell}\right]<\frac{10t^3\cdot 2^{2\ell}}{\mu_{t-1}},$$

where $\text{Dominated}=\{(\overline{h},y)\in\text{Supp}(\mathsf{D}_{\text{Ideal}})\colon \mathsf{D}_{\text{Real}}(\overline{h},y)\geq\frac{1}{2^8}\cdot\mathsf{D}_{\text{Ideal}}(\overline{h},y)\}$.

Namely, with high probability over the choice of $\overline{h}\leftarrow\mathcal{H}^t$, the number of elements that are consistent with $\overline{h}$ and whose weight according to $\mathsf{D}_{\text{Ideal}}$ is much larger than their weight according to D_{Real}, is limited.

Lemma 4.5. *Assume $t\geq 4$ and $\mu_{t-1}\geq\frac{12t^3\cdot 2^{2\ell}}{\varepsilon}$, and let* Secluded *be an arbitrary subset of* $\text{Supp}(\mathsf{D}_{\text{Ideal}})$ *with*

$$\Pr_{\overline{h}\leftarrow\mathcal{H}^t}\left[\left|\left(\overline{h},\text{Consist}(\overline{h})\right)\cap\text{Secluded}\right|>\sqrt{\varepsilon\cdot 2^{(s-t)\ell-3}}\right]\leq\varepsilon/2.$$

Then

$$\Pr_{(\overline{h},y)\leftarrow\mathsf{D}_{\text{Ideal}}}\left[\text{Inverter}(\overline{h})\in\mathcal{W}_y\ \wedge\ (\overline{h},y)\notin\text{Secluded}\right]\geq\frac{\varepsilon}{64\cdot\mu_t}.$$

Namely, if $|(\overline{h},\text{Consist}(\overline{h}))\cap\text{Secluded}|$ is not too large on the average, then Inverter does well on $\mathsf{D}_{\text{Ideal}}$ even when ignoring the contribution of Secluded.

The proof of Lemmas 4.4 and 4.5 is given below, but first let us use them for proving Theorem 3.7.

Proof of Theorem 3.7. Let $q := \lfloor\frac{1}{\ell}(\min\{\log|\mathcal{L}|, s\ell\}+\log\varepsilon-8\log n-6\ell-9)\rfloor$. We first prove the theorem for the case $q<4$, and then complete the proof by handling the case $q\geq 4$.

Assume $q < 4$ and consider the success of A when setting $t = 0$. To lowerbound A's success probability in this case, we compute

$$\begin{aligned}\Pr_{y\leftarrow\mathcal{L}}\left[\mathsf{A}(y) \in W_y\right] &= \Pr_{y\leftarrow\mathcal{L}}\left[\mathsf{Inverter}() \in W_y\right] \\ &\geq \frac{\varepsilon}{2\cdot|\mathcal{L}|} = \frac{\varepsilon}{2\cdot\min\{|\mathcal{L}|, 2^{s\ell}\}}\cdot\min\left\{1, \frac{2^{s\ell}}{|\mathcal{L}|}\right\}. \end{aligned} \tag{5}$$

Since by assumption $q < 4$, we have $\min\{|\mathcal{L}|, 2^{s\ell}\} \leq 2^{4\ell-\log\varepsilon+8\log n+6\ell+9} = \frac{2^9\cdot 2^{10\ell}\cdot n^8}{\varepsilon}$. It follows that $\Pr_{y\leftarrow L}[\mathsf{A}(y) \in W_y] \geq \frac{\varepsilon^2}{2^{10}\cdot 2^{10\ell}\cdot n^8}\cdot\min\{1, \frac{2^{s\ell}}{|\mathcal{L}|}\}$, concluding the proof for the case $q < 4$.

Assume $q \geq 4$ and let $t = q$. We start by showing that μ_{t-1} is large enough, so we can later invoke Lemma 4.5 with parameter t. Indeed,

$$\mu_{t-1} = \frac{|\mathcal{L}|}{2^{(t-1)\ell}} > \frac{|\mathcal{L}|}{2^{-\ell}\cdot|\mathcal{L}|\cdot\varepsilon\cdot n^{-8}\cdot 2^{-6\ell}\cdot 2^{-9}} = \frac{512\cdot n^8\cdot 2^{7\ell}}{\varepsilon} > \frac{12t^3\cdot 2^{2\ell}}{\varepsilon}, \tag{6}$$

where the last inequality holds since $t \leq n$.

Let Dominated be the set defined in Lemma 4.4. We have

$$\begin{aligned}&\Pr_{\overline{h}\leftarrow\mathcal{H}^t}\left[\left|\left\{y \in \mathrm{Consist}(\overline{h})\colon (\overline{h}, y) \notin \mathrm{Dominated}\right\}\right| > \sqrt{\varepsilon\cdot 2^{(s-t)\ell-3}}\right] \\ &\leq \Pr_{\overline{h}\leftarrow\mathcal{H}^t}\left[\left|\left\{y \in \mathrm{Consist}(\overline{h})\colon (\overline{h}, y) \notin \mathrm{Dominated}\right\}\right| > 8t^4\cdot 2^{3\ell}\right] \\ &< \frac{10t^3\cdot 2^{2\ell}}{\mu_{t-1}} \\ &= \frac{10t^3\cdot 2^{2\ell}\cdot 2^{(t-1)\ell}}{|\mathcal{L}|} \\ &\leq \frac{10t^3\cdot 2^{2\ell}}{|\mathcal{L}|}\cdot\frac{|\mathcal{L}|\cdot\varepsilon}{n^8\cdot 2^{7\ell}\cdot 2^9} \\ &\leq 10n^3\cdot 2^{2\ell}\cdot\frac{\varepsilon}{n^8\cdot 2^{7\ell}\cdot 2^9} \\ &\leq \varepsilon/4. \end{aligned} \tag{7}$$

The first inequality holds since $2^{(s-t)\ell-3} \geq 2^{s\ell-(s\ell+\log\varepsilon-8\log n-6\ell-9)-3} \geq \frac{1}{\varepsilon}\cdot n^8\cdot 2^{6\ell}\cdot 2^6 \geq \frac{1}{\varepsilon}\cdot(8t^4\cdot 2^{3\ell})^2$ (note that $t \leq n$), the second one by Lemma 4.4 and the third one by the definition of t. Applying Lemma 4.5 for $\mathrm{Secluded} := \mathrm{Supp}(\mathsf{D}_{\mathrm{Ideal}}) \setminus \mathrm{Dominated}$, yields

$$\Pr_{(\overline{h},y)\leftarrow\mathsf{D}_{\mathrm{Ideal}}}\left[\mathsf{Inverter}(\overline{h}) \in \mathcal{W}_y \ \wedge\ (\overline{h}, y) \in \mathrm{Dominated}\right] \geq \frac{\varepsilon}{64\cdot\mu_t}. \tag{8}$$

It follows that

$$\begin{aligned}
\Pr_{y\leftarrow\mathcal{L}}\left[\mathsf{A}(y)\in\mathcal{W}_y\right] &= \Pr_{(\overline{h},y)\leftarrow\mathsf{D}_{\mathrm{Real}}}\left[\mathsf{Inverter}(\overline{h})\in\mathcal{W}_y\right] \\
&\geq \Pr_{(\overline{h},y)\leftarrow\mathsf{D}_{\mathrm{Real}}}\left[\mathsf{Inverter}(\overline{h})\in\mathcal{W}_y \ \wedge\ (\overline{h},y)\in\mathrm{Dominated}\right] \\
&\geq \frac{1}{2^8}\cdot\Pr_{(\overline{h},y)\leftarrow\mathsf{D}_{\mathrm{Ideal}}}\left[\mathsf{Inverter}(\overline{h})\in\mathcal{W}_y \ \wedge\ (\overline{h},y)\in\mathrm{Dominated}\right] \\
&\geq \frac{1}{2^8}\cdot\frac{\varepsilon}{64\cdot\mu_t} \\
&\geq \frac{1}{2^8\cdot 64\cdot 2^9}\cdot\frac{\varepsilon^2}{n^8\cdot 2^{6\ell}},
\end{aligned}$$

concluding the proof for the case $q\geq 4$. □

Remark 4.6 (Knowing t). The value of t in the above proof depends on ε and $|\mathcal{L}|$. This seems to contradict our binding formalism (see Definition 3.2) where A does not know ε and $\mathcal{L}$. However, selecting t at random only decrease A's success probability by a factor s. More interestingly, setting $t=s$ guarantees that A success probability is as claimed in the theorem; the effect of setting t to s is analogous to setting t arbitrarily and changing $\mathsf{Inverter}$ to select $\overline{h}'$ using the rewinding method of $\mathsf{Searcher}$ rather than uniformly at random. For every value $\overline{h}'$ that satisfies $y\in\mathrm{Consist}(\overline{h},\overline{h}')$, the probability of selecting it with the rewinding technique is only larger than the probability of uniformly selecting it. Where a value of $\overline{h}'$ such that $y\notin\mathrm{Consist}(\overline{h},\overline{h}')$, does not contribute in our analysis to the success probability of A. It follows that the distinction between $\mathsf{Searcher}$ and $\mathsf{Inverter}$ is not necessary for the proof (but is only used for presentation reasons).

4.1. *Bounding the Size of* $\mathrm{Consist}(\overline{h})$

The following simple observation plays an important role in the proofs of Lemmas 4.5 and 4.4.

Claim 4.7. $\Pr_{\overline{h}\leftarrow\mathcal{H}^k}[\frac{1}{4}\cdot\mu_k\leq|\mathrm{Consist}(\overline{h})|\leq 4\cdot\mu_k]\geq 1-\frac{3k^3\cdot 2^{2\ell}}{\mu_{k-1}}$, *for every* $k\in\{2,\ldots,s\}$.

Proof. Fix $k\in\{2,\ldots,s\}$. We call $\overline{h}\in\mathcal{H}^j$ *balanced* if $(1-\frac{1}{k})^j\cdot\mu_j\leq|\mathrm{Consist}(\overline{h})|\leq(1+\frac{1}{k})^j\cdot\mu_j$ and prove the claim showing that

$$\Pr_{\overline{h}\leftarrow\mathcal{H}^j}[\overline{h}\text{ is balanced}]\geq 1-\frac{3jk^2\cdot 2^{2\ell}}{\mu_{j-1}} \tag{9}$$

for every $j\in\{0,\ldots,k\}$, letting $\mu_{-1}=\mu_0$. Equation (9) holds trivially for $j=0$. In the following we assume for $j\geq 0$, and prove for $j+1$.

We say that $h \in \mathcal{H}$ *well partitions* the set $\text{Consist}(\overline{h})$, where $\overline{h} \in \mathcal{H}^j$, if $(1-\frac{1}{k}) \cdot |\text{Consist}(\overline{h})| \leq 2^\ell \cdot |\text{Consist}(\overline{h}, h)| \leq (1+\frac{1}{k}) \cdot |\text{Consist}(\overline{h})|$. Lemma 2.3 yields

$$\Pr_{h \leftarrow \mathcal{H}}\left[h \text{ well partitions } \text{Consist}(\overline{h})\right] \geq 1 - \frac{k^2 \cdot 2^{2\ell}}{|\text{Consist}(\overline{h})|} \tag{10}$$

for every $\overline{h} \in \mathcal{H}^j$, and it follows that

$$\begin{aligned}
&\Pr_{\overline{h} \leftarrow \mathcal{H}^{j+1}}[\overline{h} \text{ is balanced}] \\
&\geq \Pr_{\overline{h} \leftarrow \mathcal{H}^j}[\overline{h} \text{ is balanced}] \cdot \Pr_{(\overline{h},h) \leftarrow \mathcal{H}^{j+1}}\left[(\overline{h}, h) \text{ is balanced} \mid \overline{h} \text{ is balanced}\right] \\
&\geq \Pr_{\overline{h} \leftarrow \mathcal{H}^j}[\overline{h} \text{ is balanced}] \cdot \Pr_{h \leftarrow \mathcal{H}^{j+1}}\left[h \text{ well partitions } \text{Consist}(\overline{h}) \mid \overline{h} \text{ is balanced}\right] \\
&\geq \left(1 - \frac{3jk^2 \cdot 2^{2\ell}}{\mu_{j-1}}\right) \cdot \left(1 - \frac{k^2 \cdot 2^{2\ell}}{(1-\frac{1}{k})^j \mu_j}\right) \\
&\geq 1 - \frac{3jk^2 \cdot 2^{2\ell}}{\mu_{j-1}} - \frac{3k^2 \cdot 2^{2\ell}}{\mu_j} \geq 1 - \frac{3(j+1)k^2 \cdot 2^{2\ell}}{\mu_j}.
\end{aligned}$$

□

4.2. *Proving Lemma 4.4*

In this section we prove Lemma 4.4.

Lemma 4.8 (Lemma 4.4, restated). *Assuming $t \geq 4$ then*

$$\Pr_{\overline{h} \leftarrow \mathcal{H}^t}\left[\left|\left\{y \in \text{Consist}(\overline{h}) \colon (\overline{h}, y) \notin \text{Dominated}\right\}\right| > 8t^4 \cdot 2^{3\ell}\right] < \frac{10t^3 \cdot 2^{2\ell}}{\mu_{t-1}},$$

where $\text{Dominated} = \{(\overline{h}, y) \in \text{Supp}(\mathsf{D}_{\text{Ideal}}) \colon \mathsf{D}_{\text{Real}}(\overline{h}, y) \geq \frac{1}{2^8} \cdot \mathsf{D}_{\text{Ideal}}(\overline{h}, y)\}$.

We bridge between $\mathsf{D}_{\text{Ideal}}$ and D_{Real} via the following hybrid distributions: for $k \in \{0, \ldots, t-1\}$ and $\overline{h} \in \mathcal{H}^k$, define

- $\mathsf{D}^{\overline{h}}_{\text{Real}} := (h, y)_{y \leftarrow \text{Consist}(\overline{h}), h \leftarrow (\mathsf{Searcher}^{\overline{h}}(y))_{k+1}}$ and
- $\mathsf{D}^{\overline{h}}_{\text{Ideal}} := (h, y)_{h \leftarrow \mathcal{H}, y \leftarrow \text{Consist}(\overline{h}, h)}$,

where $\mathsf{Searcher}^{\overline{h}}(y)$ is the hybrid algorithm that first fixes its first k hash functions to $\overline{h}$, and then continues as (the non-hybrid) original $\mathsf{Searcher}$ would with respect to this fixing.

For $(h_1,\ldots,h_i,y)\in H^t\times\{0,1\}^n$, let $\gamma^{h_1,\ldots,h_{i-1}}(h_i,y):=\frac{\mathsf{D}^{h_1,\ldots,h_{i-1}}_{\text{Ideal}}(h_i,y)}{\mathsf{D}^{h_1,\ldots,h_{i-1}}_{\text{Real}}(h_i,y)}$. Hence, for $(\overline{h}=(h_1,\ldots,h_t),y)\in\text{Supp}(\mathsf{D}_{\text{Ideal}})$ we can write

$$\begin{aligned}
\mathsf{D}_{\text{Ideal}}(\overline{h},y)&=\frac{1}{|\mathcal{H}|^{t-1}}\cdot\mathsf{D}^{h_1,\ldots,h_{t-1}}_{\text{Ideal}}(h_t,y)\\
&=\frac{1}{|\mathcal{H}|^{t-1}}\cdot\gamma^{h_1,\ldots,h_{t-1}}(h_t,y)\cdot\mathsf{D}^{h_1,\ldots,h_{t-1}}_{\text{Real}}(h_t,y)\\
&=\frac{1}{|\mathcal{H}|^{t-1}}\cdot\gamma^{h_1,\ldots,h_{t-1}}(h_t,y)\cdot\frac{1}{|\text{Consist}(h_1,\ldots,h_{t-1})|}\\
&\quad\cdot\Pr\left[\left(\mathsf{Searcher}^{h_1,\ldots,h_{t-1}}(y)\right)_t=h_t\right]\\
&=\frac{1}{|\mathcal{H}|^{t-1}}\cdot\gamma^{h_1,\ldots,h_{t-1}}(h_t,y)\cdot|\mathcal{H}|\cdot\mathsf{D}^{h_1,\ldots,h_{t-2}}_{\text{Ideal}}(h_{t-1},y)\\
&\quad\cdot\Pr\left[\left(\mathsf{Searcher}^{h_1,\ldots,h_{t-1}}(y)\right)_t=h_t\right]\\
&=\frac{1}{|\mathcal{H}|^{t-2}}\cdot\gamma^{h_1,\ldots,h_{t-1}}(h_t,y)\cdot\gamma^{h_1,\ldots,h_{t-2}}(h_{t-1},y)\\
&\quad\cdot\mathsf{D}^{h_1,\ldots,h_{t-2}}_{\text{Real}}(h_{t-1},y)\cdot\Pr\left[\left(\mathsf{Searcher}^{h_1,\ldots,h_{t-1}}(y)\right)_t=h_t\right]\\
&\;\;\vdots\\
&=\left(\prod_{i\in[t]}\gamma^{h_1,\ldots,h_{i-1}}(h_i,y)\right)\cdot\mathsf{D}^{\lambda}_{\text{Real}}(h_1,y)\\
&\quad\cdot\left(\prod_{i\in\{2,\ldots,t\}}\Pr\left[\left(\mathsf{Searcher}^{h_1,\ldots,h_{i-1}}(y)\right)_i=h_i\right]\right)\\
&=\left(\prod_{i\in[t]}\gamma^{h_1,\ldots,h_{i-1}}(h_i,y)\right)\cdot\frac{1}{|\mathcal{L}|}\cdot\Pr\left[\mathsf{Searcher}(y)=\overline{h}\right]\\
&=\left(\prod_{i\in[t]}\gamma^{h_1,\ldots,h_{i-1}}(h_i,y)\right)\cdot\mathsf{D}_{\text{Real}}(\overline{h},y),
\end{aligned}\tag{11}$$

where in above λ stands for the zero length vector.

Equation (11) yields the following characterization of the set Dominated.

Claim 4.9. Dominated $\supseteq\{((h_1,\ldots,h_t),y)\in\text{Supp}(\mathsf{D}_{\text{Ideal}}):\quad\forall i\in[t]\quad(1-\frac{3}{t})\cdot\mathsf{D}^{h_1,\ldots,h_{i-1}}_{\text{Ideal}}(h_i,y)\le\mathsf{D}^{h_1,\ldots,h_{i-1}}_{\text{Real}}(h_i,y)\}$.

Proof. Fix $(\overline{h}=(h_1,\ldots,h_t),y)\in\text{Supp}(\mathsf{D}_{\text{Ideal}})$ with $(1-\frac{3}{t})\cdot\mathsf{D}^{h_1,\ldots,h_{i-1}}_{\text{Ideal}}(h_i,y)\le\mathsf{D}^{h_1,\ldots,h_{i-1}}_{\text{Real}}(h_i,y)$ for every $i\in[t]$. Equation (11) yields $\mathsf{D}_{\text{Ideal}}(\overline{h},y)\le(\frac{1}{1-\frac{3}{t}})^t\cdot\mathsf{D}_{\text{Real}}(\overline{h},y)$. Since $(\frac{1}{1-\frac{3}{t}})^t\le 2^8$ for $t\ge 4$, it follows that $(\overline{h},y)\in$ Dominated. □

Recall that in order to prove Lemma 4.4, we need to show that Dominated is large. By Claim 4.9, it suffices to lowerbound the number of pairs $(\overline{h} = (h_1, \ldots, h_t), y) \in \mathrm{Supp}(\mathsf{D}_{\mathrm{Ideal}})$ for which $(1 - \frac{3}{t}) \cdot \mathsf{D}_{\mathrm{Ideal}}^{h_1,\ldots,h_{i-1}}(h_i, y) \leq \mathsf{D}_{\mathrm{Real}}^{h_1,\ldots,h_{i-1}}(h_i, y)$ for every $i \in [t]$, a task that we do using the next lemma.

Lemma 4.10. *For $\overline{h} \in \mathcal{H}^k$, where $k \in \{0, \ldots, t-1\}$, there exists a set* $\mathrm{NonTypY}(\overline{h}) \subseteq \mathrm{Consist}(\overline{h})$ *of size less than $8t^3 \cdot 2^{3\ell}$ such that the following holds: let* $\mathrm{BadH}(\overline{h}) := \{h \in \mathcal{H} \colon \exists y \in \mathrm{Consist}(\overline{h}) \setminus \mathrm{NonTypY}(\overline{h}) \colon (1 - \frac{3}{t}) \cdot \mathsf{D}_{\mathrm{Ideal}}^{\overline{h}}(h, y) > \mathsf{D}_{\mathrm{Real}}^{\overline{h}}(h, y)\}$, *then* $\Pr_{h \leftarrow \mathcal{H}}[h \in \mathrm{BadH}(\overline{h})] < \frac{t^2 \cdot 2^{2\ell}}{|\mathrm{Consist}(\overline{h})|}$.

Namely, Lemma 4.10 bounds the number of y's inside $\mathrm{Consist}(\overline{h})$ for which the event $(1 - \frac{3}{t}) \cdot \mathsf{D}_{\mathrm{Ideal}}^{\overline{h}}(H, y) > \mathsf{D}_{\mathrm{Real}}^{\overline{h}}(H, y)$ is likely to happen. Before proving Lemma 4.10, we first use it to prove Lemma 4.4.

Proof of Lemma 4.4. For $\overline{h} \in \mathcal{H}^k$, let $\mathrm{NonTypY}(\overline{h})$ be the set guaranteed by Lemma 4.10 (assuming for ease of notation that there exists a *single* such set; in case there are several of them, we arbitrarily choose one). Let $\mathrm{BadH} := \{\overline{h} = (h_1, \ldots, h_t) \in \mathcal{H}^t \colon \exists i \in [t] \colon h_i \in \mathrm{BadH}(h_1, \ldots, h_{i-1})\}$, and for $\overline{h} = (h_1, \ldots, h_t) \in \mathcal{H}^t$ let $\mathrm{AllNonTypY}(\overline{h}) := \bigcup_{i \in [t]} \mathrm{NonTypY}(h_1, \ldots, h_{i-1})$.

Claim 4.9 yields $\{(\overline{h}, y) \in \mathrm{Supp}(\mathsf{D}_{\mathrm{Ideal}}) \colon y \in (\mathrm{Consist}(\overline{h}) \setminus \mathrm{AllNonTypY}(\overline{h})) \wedge \overline{h} \notin \mathrm{BadH}\} \subseteq \mathrm{Dominated}$, and therefore

$$\begin{aligned}
&\Pr_{\overline{h} \leftarrow \mathcal{H}^t}\left[\left|\left\{y \in \mathrm{Consist}(\overline{h}) \colon (\overline{h}, y) \notin \mathrm{Dominated}\right\}\right| \geq 8t^4 \cdot 2^{3\ell}\right] \\
&\quad \leq \Pr_{\overline{h} \leftarrow \mathcal{H}^t}\left[\exists y \in \left(\mathrm{Consist}(\overline{h}) \setminus \mathrm{AllNonTypY}(\overline{h})\right) \wedge (\overline{h}, y) \notin \mathrm{Dominated}\right] \\
&\quad \leq \Pr_{\overline{h} \leftarrow \mathcal{H}^t}[\overline{h} \in \mathrm{BadH}], \qquad (12)
\end{aligned}$$

where for the first inequality observe that if the number of $y \in \mathrm{Consist}(\overline{h})$ with $(\overline{h}, y) \notin \mathrm{Dominated}$ is at least $8t^4 \cdot 2^{3\ell}$, then there exists at least one $y \in \mathrm{Consist}(\overline{h}) \setminus \mathrm{AllNonTypY}(\overline{h})$ with $(\overline{h}, y) \notin \mathrm{Dominated}$ (note that $|\mathrm{AllNonTypY}(\overline{h})| < 8t^4 \cdot 2^{3\ell}$).

We conclude the proof showing that $\Pr_{\overline{h} \leftarrow \mathcal{H}^t}[\overline{h} \in \mathrm{BadH}]$ is small. For $q > 0$ compute

$$\begin{aligned}
\Pr_{\overline{h} \leftarrow \mathcal{H}^t}[\overline{h} \in \mathrm{BadH}] &\leq \sum_{i \in [t]} \Pr_{(h_1,\ldots,h_i) \leftarrow \mathcal{H}^i}\left[h \in \mathrm{BadH}(h_1, \ldots, h_{i-1})\right] \\
&\leq \sum_{i \in [t]} \left(\Pr_{(h_1,\ldots,h_{i-1}) \leftarrow \mathcal{H}^{i-1}}\left[\left|\mathrm{Consist}(h_1, \ldots, h_{i-1})\right| < q\right] + \frac{t^2 \cdot 2^{3\ell}}{q} \right) \\
&= \frac{t^3 \cdot 2^{2\ell}}{q} + \sum_{i \in [t]} \Pr_{(h_1,\ldots,h_{i-1}) \leftarrow \mathcal{H}^{i-1}}\left[\left|\mathrm{Consist}(h_1, \ldots, h_{i-1})\right| < q\right], \qquad (13)
\end{aligned}$$

where the second inequality is by Lemma 4.10. Taking $q = \frac{\mu_{t-1}}{4}$, Claim 4.7 yields

$$\begin{aligned}
&\sum_{i\in[t]} \Pr_{(h_1,\ldots,h_{i-1})\leftarrow\mathcal{H}^{i-1}}\left[\left|\text{Consist}(h_1,\ldots,h_{i-1})\right| < q\right] \\
&\leq \sum_{i\in[t]} \Pr_{(h_1,\ldots,h_{i-1})\leftarrow\mathcal{H}^{i-1}}\left[\left|\text{Consist}(h_1,\ldots,h_{i-1})\right| < \frac{\mu_{i-1}}{4}\right] \\
&\leq \sum_{i\in[t]} \frac{3t^3\cdot 2^{2\ell}}{\mu_{i-2}} \\
&< \frac{6t^3\cdot 2^{2\ell}}{\mu_{t-2}}, \qquad (14)
\end{aligned}$$

yielding that

$$\Pr_{\overline{h}\leftarrow\mathcal{H}^t}[\overline{h}\in \text{Bad}\overline{\text{H}}] < \frac{4t^3\cdot 2^{2\ell}}{\mu_{t-1}} + \frac{6t^3\cdot 2^{2\ell}}{\mu_{t-1}} = \frac{10t^3\cdot 2^{2\ell}}{\mu_{t-1}}. \qquad (15)$$

Combining Eqs. (12), (15), yields $\Pr_{\overline{h}\leftarrow\mathcal{H}^t}[|\{y\in\text{Consist}(\overline{h})\colon(\overline{h},y)\notin\text{Dominated}\}| > 8t^4\cdot 2^{3\ell}] \leq \Pr_{\overline{h}\leftarrow\mathcal{H}^t}[\overline{h}\in\text{Bad}\overline{\text{H}}] < \frac{10t^3\cdot 2^{2\ell}}{\mu_{t-1}}$. □

4.2.1. *Proving Lemma 4.10*

As a warmup we start by focusing on the Boolean case (i.e., $\ell = 1$). Consider the Boolean matrix $M\in\{0,1\}^{|\text{Consist}(\overline{h})|\times|\mathcal{H}|}$ with $M(y,h)=1$ iff $y\in\text{Consist}(\overline{h},h)$, where we identify the indices into M with the set $\text{Consist}(\overline{h})\times\mathcal{H}$. The distribution $\mathsf{D}^{\overline{h}}_{\text{Ideal}}$ can be described in relation to M as: choose a random column of M and draw the index of a random 1-entry from this column (where a "1-entry" is an entry of the matrix that is assigned the value 1). The distribution $\mathsf{D}^{\overline{h}}_{\text{Real}}$ can also be described in relation to M: choose a random row of M and keep drawing random entries from this row until a 1-entry is picked. Then return its index and stop drawing (where in case of $\lceil 2\log t\rceil$ failed attempts, return $\perp$).

Compare the matrix M with the matrix $\widehat{M}\in\{0,1\}^{|\text{Consist}(\overline{h})|\times|\mathcal{H}|}$, where $\widehat{M}(y,h) = h(y)$. Note that M can be derived from $\widehat{M}$ by flipping all values in some of its columns (specifically, the column corresponding to h is flipped in case $y\notin\text{Consist}(\overline{h},h)$). The pairwise independence of $\mathcal{H}$ shows that most *columns* of $\widehat{M}$ are balanced (i.e., have about the same number of 1-entries), and thus the same holds for M. Hence, the mass that $\mathsf{D}^{\overline{h}}_{\text{Ideal}}$ assigns to most of the 1-entries of M is close to $\frac{2}{|\mathcal{H}|}\cdot\frac{1}{|\text{Consist}(\overline{h})|}$. The pairwise independence of $\mathcal{H}$ also shows that most *rows* of M are also balanced. Since $\mathsf{D}_{\text{Real}}=\perp$ only with small probability, the mass that $\mathsf{D}^{\overline{h}}_{\text{Real}}$ assigns to most 1-entries in M is also close to $\frac{2}{|\mathcal{H}|}\cdot\frac{1}{|\text{Consist}(\overline{h})|}$. We conclude that the 1-entries in a random row of M (a random $h\in\mathcal{H}$) get about the same mass in $\mathsf{D}^{\overline{h}}_{\text{Real}}$ and in $\mathsf{D}^{\overline{h}}_{\text{Ideal}}$, and the proof of the lemma follows. Formal proof is given below.

Proof of Lemma 4.10. We take $\mathrm{NonTypY}(\overline{h})$ as the set $\{y \in \mathrm{Consist}(\overline{h})\colon \alpha_{\overline{h}}(y) > \frac{1}{2^\ell}\cdot(1+\frac{1}{t})\}$, for $\alpha_{\overline{h}}(y) := \Pr_{h\leftarrow\mathcal{H}}[y \in \mathrm{Consist}(\overline{h}, h)]$. The proof that $\mathrm{NonTypY}(\overline{h})$ has the two properties stated in Lemma 4.10 (i.e., bounded size, and dominance of $\mathsf{D}_{\mathrm{Real}}$ over $\mathsf{D}_{\mathrm{Ideal}}$) is carried via the next two claims.

Claim 4.11. $|\mathrm{NonTypY}(\overline{h})| < 8t^3\cdot 2^{3\ell}$.

Proof. In the following let H be uniformly distributed over $\mathcal{H}$, and for $h \in \mathcal{H}$ let $X_h := |\mathrm{NonTypY}(\overline{h}) \setminus \mathrm{Consist}(\overline{h}, h)|$. The definition of $\mathrm{NonTypY}(\overline{h})$ shows that

$$\begin{aligned}\mathrm{E}[X_H] &= \mathrm{E}\big[\big|\mathrm{NonTypY}(\overline{h}) \setminus \mathrm{Consist}(\overline{h}, H)\big|\big] \\ &= \big|\mathrm{NonTypY}(\overline{h})\big| - \sum_{y\in\mathrm{NonTypY}(\overline{h})} \Pr\big[y \in \mathrm{Consist}(\overline{h}, H)\big] \\ &< \big|\mathrm{NonTypY}(\overline{h})\big| - \left(1+\frac{1}{t}\right)\nu, \end{aligned} \tag{16}$$

for $\nu = \frac{|\mathrm{NonTypY}(\overline{h})|}{2^\ell}$. On the other hand,

$$\begin{aligned}&\Pr\left[X_H < \big|\mathrm{NonTypY}(\overline{h})\big| - \left(1+\frac{1}{2t}\right)\nu\right] \\ &= \Pr\left[\big|\mathrm{NonTypY}(\overline{h}) \cap \mathrm{Consist}(\overline{h}, H)\big| > \left(1+\frac{1}{2t}\right)\nu\right] \\ &\le \Pr\left[\exists z \in \{0,1\}^\ell\colon \big|\{y \in \mathrm{NonTypY}(\overline{h})\colon H(y) = z\}\big| > \left(1+\frac{1}{2t}\right)\nu\right] \\ &< \frac{4t^2\cdot 2^\ell}{\nu}, \end{aligned} \tag{17}$$

where the last inequality is by Lemma 2.3. We conclude that

$$\begin{aligned}\mathrm{E}[X_H] &\ge \Pr\left[X_H \ge \big|\mathrm{NonTypY}(\overline{h})\big| - \left(1+\frac{1}{2t}\right)\nu\right]\cdot\left(\big|\mathrm{NonTypY}(\overline{h})\big| - \left(1+\frac{1}{2t}\right)\nu\right) \\ &\ge \left(1-\frac{4t^2\cdot 2^\ell}{\nu}\right)\cdot\left(\big|\mathrm{NonTypY}(\overline{h})\big| - \left(1+\frac{1}{2t}\right)\nu\right) \\ &\ge \big|\mathrm{NonTypY}(\overline{h})\big| - \left(1+\frac{1}{2t}+\frac{4t^2\cdot 2^{2\ell}}{\nu}\right)\nu. \end{aligned} \tag{18}$$

Assume towards a contradiction that $|\mathrm{NonTypY}(\overline{h})| \ge 8t^3\cdot 2^{3\ell}$ (and hence, $\nu \ge 8t^3\cdot 2^{2\ell}$), Eq. (18) yields $\mathrm{E}[X_H] \ge |\mathrm{NonTypY}(\overline{h})| - (1+\frac{1}{2t}+\frac{4t^2\cdot 2^{2\ell}}{8t^3\cdot 2^{2\ell}})\nu = |\mathrm{NonTypY}(\overline{h})| - (1+\frac{1}{t})\nu$, in contradiction to Eq. (16). Hence, we have proved that $|\mathrm{NonTypY}(\overline{h})| < 8t^3\cdot 2^{3\ell}$. □

The next claim completes the proof of Lemma 4.10, showing that $\text{NonTypY}(\overline{h})$ indeed contains all the "non typical" y's.

Claim 4.12. $\Pr_{h\leftarrow\mathcal{H}}[\exists y \in \text{Consist}(\overline{h}) \setminus \text{NonTypY}(\overline{h})\colon (1-\frac{3}{t})\cdot \mathsf{D}^{\overline{h}}_{\text{Ideal}}(h,y) > \mathsf{D}^{\overline{h}}_{\text{Real}}(h,y)] < \frac{t^2\cdot 2^{2\ell}}{|\text{Consist}(\overline{h})|}$.

Proof. By definition, $\mathsf{D}^{\overline{h}}_{\text{Ideal}}(h,y) = \frac{1}{|\mathcal{H}|}\cdot\frac{1}{|\text{Consist}(\overline{h},h)|}$ for every $h\in\mathcal{H}$ and $y\in\text{Consist}(\overline{h},h)$. In addition, the mass that $\mathsf{D}^{\overline{h}}_{\text{Real}}$ assigns to every such pair (h,y), is the probability that y is chosen (i.e., $\frac{1}{|\text{Consist}(\overline{h})|}$) times the probability that h is chosen in one of the $\lceil 2^\ell\cdot\ln t\rceil$ sampling attempts done by $\mathsf{Searcher}(y)$ (i.e., $\sum_{i=1}^{\lceil 2^\ell\cdot\ln t\rceil}\Pr_{\mathcal{H}}[h]\cdot \Pr[\text{first } (i-1) \text{ attempts failed}] = \sum_{i=1}^{\lceil 2^\ell\cdot\ln t\rceil}\frac{1}{|\mathcal{H}|}\cdot(1-\alpha_{\overline{h}}(y))^{i-1}$). All in all,

$$\mathsf{D}^{\overline{h}}_{\text{Real}}(h,y) = \frac{1}{|\text{Consist}(\overline{h})|}\cdot\frac{1}{|\mathcal{H}|}\cdot\sum_{i=1}^{\lceil 2^\ell\cdot\ln t\rceil}\left(1-\alpha_{\overline{h}}(y)\right)^{i-1}. \tag{19}$$

Assuming that $y\in\text{Consist}(\overline{h},h)\setminus\text{NonTypY}(\overline{h})$, Eq. (19) yields

$$\begin{aligned}
\mathsf{D}^{\overline{h}}_{\text{Real}}(h,y) &\geq \frac{1}{|\text{Consist}(\overline{h})|}\cdot\frac{1}{|\mathcal{H}|}\cdot\frac{1-(1-\frac{1}{2^\ell}\cdot(1+\frac{1}{t}))^{\lceil 2^\ell\cdot\ln t\rceil}}{\frac{1}{2^\ell}\cdot(1+\frac{1}{t})}\\
&\geq \frac{1}{|\text{Consist}(\overline{h})|}\cdot\frac{1}{|\mathcal{H}|}\cdot\frac{1-(1-\frac{1}{2^\ell})^{\lceil 2^\ell\cdot\ln t\rceil}}{\frac{1}{2^\ell}\cdot(1+\frac{1}{t})}\\
&\geq \frac{1}{|\text{Consist}(\overline{h})|}\cdot\frac{1}{|\mathcal{H}|}\cdot\frac{1-\frac{1}{t}}{\frac{1}{2^\ell}\cdot(1+\frac{1}{t})},
\end{aligned} \tag{20}$$

where for the last inequality we use the fact that $(1-\frac{1}{x})^x\leq e^{-1}$ for $x\geq 1$.

Let $\text{NonTypH}(\overline{h}) := \{h\in\mathcal{H}\colon |\text{Consist}(\overline{h},h)| < (1-\frac{1}{t})\cdot\frac{|\text{Consist}(\overline{h})|}{2^\ell}\}$. Observe that

$$\begin{aligned}
&\Pr_{h\leftarrow\mathcal{H}}\left[h\in\text{NonTypH}(\overline{h})\right]\\
&\leq \Pr_{h\leftarrow\mathcal{H}}\left[\exists z\in\{0,1\}^\ell\colon \left|\left\{y\in\text{Consist}(\overline{h})\colon h(y)=z\right\}\right| < \left(1-\frac{1}{t}\right)\cdot\frac{|\text{Consist}(\overline{h})|}{2^\ell}\right]\\
&< \frac{t^2\cdot 2^{2\ell}}{|\text{Consist}(\overline{h})|},
\end{aligned} \tag{21}$$

where the second inequality is by Lemma 2.3. We conclude the claim's proof, showing that $(1-\frac{3}{t})\cdot\mathsf{D}^{\overline{h}}_{\text{Ideal}}(h,y)\leq\mathsf{D}^{\overline{h}}_{\text{Real}}(h,y)$ for every pair (h,y) with $h\in\mathcal{H}\setminus\text{NonTypH}(\overline{h})$ and $y\in\text{Consist}(\overline{h},h)\setminus\text{NonTypY}(\overline{h})$. Indeed (by Eq. (20)) $\mathsf{D}^{\overline{h}}_{\text{Real}}(h,y)\geq\frac{1-\frac{1}{t}}{1+\frac{1}{t}}\cdot$

$\frac{2^\ell}{|\text{Consist}(\overline{h})|} \cdot \frac{1}{|\mathcal{H}|}$ and (by the definition of NonTypH($\overline{h}$)) $\mathsf{D}^{\overline{h}}_{\text{Ideal}}(y,h) \le \frac{1}{(1-\frac{1}{t})} \cdot \frac{1}{|\mathcal{H}|} \cdot \frac{2^\ell}{|\text{Consist}(\overline{h})|}$ for every such pair, yielding that $\frac{\mathsf{D}^{\overline{h}}_{\text{Real}}(h,y)}{\mathsf{D}^{\overline{h}}_{\text{Ideal}}(h,y)} \ge \frac{(1-\frac{1}{t})^2}{1+\frac{1}{t}} > 1 - \frac{3}{t}$. □

4.3. *Proving Lemma 4.5*

In this section we prove Lemma 4.5.

Lemma 4.13 (Lemma 4.5, restated). *Assume* $t \ge 4$ *and* $\mu_{t-1} \ge \frac{12t^3 \cdot 2^{2\ell}}{\varepsilon}$, *and let* Secluded *be an arbitrary subset of* Supp($\mathsf{D}_{\text{Ideal}}$) *with*

$$\Pr_{\overline{h} \leftarrow \mathcal{H}^t}\left[\left|\text{Secluded}_{\overline{h}} := \left\{y \in \text{Consist}(\overline{h}) : (\overline{h}, y) \in \text{Secluded}\right\}\right| > \sqrt{\varepsilon \cdot 2^{(s-t)\ell-3}}\right] \le \varepsilon/2.$$

Then

$$\Pr_{(\overline{h},y) \leftarrow \mathsf{D}_{\text{Ideal}}}\left[\mathsf{Inverter}(\overline{h}) \in \mathcal{W}_y \ \wedge \ (\overline{h}, y) \notin \text{Secluded}\right] \ge \frac{\varepsilon}{64 \cdot \mu_t}.$$

Proof. For $\overline{h} \in \mathcal{H}^t$, let $\varepsilon_{\overline{h}}$ be the probability that S^* breaks the binding of NOVY$\langle \mathcal{H}^s \rangle$ with respect to $\mathcal{W}$ and $\mathcal{L}$ (according to Definition 3.2), conditioned on $\overline{h}$ being the first t functions sent by the R. Note that $\mathrm{E}_{\overline{h} \leftarrow \mathcal{H}^t}[\varepsilon_{\overline{h}}] = \varepsilon$.

Let Typical $:= \{\overline{h} \in \mathcal{H}^t : |\text{Secluded}_{\overline{h}}| \le q_{\overline{h}} \ \wedge \ |\text{Consist}(\overline{h})| \le 4\mu_t\}$, where $q_{\overline{h}} = \lfloor\sqrt{2^{(s-t)\ell-1} \cdot \varepsilon_{\overline{h}}}\rfloor$. Note that

$$\begin{aligned}\mathrm{E}_{\overline{h} \leftarrow \mathcal{H}^t}[\overline{h} \notin \text{Typical}] &\le \Pr_{\overline{h} \leftarrow \mathcal{H}^t}\left[|\text{Secluded}_{\overline{h}}| > q_{\overline{h}}\right] + \Pr_{\overline{h} \leftarrow \mathcal{H}^t}\left[\left|\text{Consist}(\overline{h})\right| > 4\mu_t\right] \\ &\le \varepsilon/2 + \frac{3t^3 \cdot 2^{2\ell}}{\mu_{t-1}} \le 3\varepsilon/4, \end{aligned} \tag{22}$$

where the second inequality is by Claim 4.7 and the guarantee about Secluded (as given in the lemma statement). Let χ_{Typical} be the characteristic function of Typical. Equation (22) yields

$$\mathrm{E}_{\overline{h} \leftarrow \mathcal{H}^t}\left[\varepsilon_{\overline{h}} \cdot \chi_{\text{Typical}}(\overline{h})\right] \ge \mathrm{E}_{\overline{h} \leftarrow \mathcal{H}^t}[\varepsilon_{\overline{h}}] - \Pr_{\overline{h} \leftarrow \mathcal{H}^t}[\overline{h} \notin \text{Typical}] \ge \varepsilon - 3\varepsilon/4 = \varepsilon/4. \tag{23}$$

We define the "weight" of $y \in \text{Consist}(\overline{h})$ with respect to $\overline{h} \in \mathcal{H}^t$, as $w_{\overline{h}}(y) := \Pr[\mathsf{Inverter}(\overline{h}) \in \mathcal{W}_y]$. It is easy to verify that

$$\sum_{y \in \text{Consist}(\overline{h})} w_{\overline{h}}(y) \ge \varepsilon_{\overline{h}}. \tag{24}$$

The following claim shows that for $\overline{h} \in$ Typical, the above sum remains high even when ignoring the contribution of Secluded$_{\overline{h}}$.

Claim 4.14. *Let* $\overline{h} \in \mathcal{H}^t$ *and let* $\mathcal{V} \subseteq \text{Consist}(\overline{h})$ *be of size at most* $q_{\overline{h}}$, *then* $\sum_{y \in \text{Consist}(\overline{h}) \setminus \mathcal{V}} w_{\overline{h}}(y) \ge \varepsilon_{\overline{h}}/4$.

Proof. The pairwise independence of $\mathcal{H}$ yields

$$\Pr_{\overline{h}' \leftarrow \mathcal{H}^{s-t}}\left[\exists y_0 \neq y_1 \in \mathcal{V} : \overline{h}'(y_0) = \overline{h}'(y_1)\right] \leq \frac{|\mathcal{V}|^2}{2^{(s-t)\ell}} \leq \frac{q_{\overline{h}}^2}{2^{(s-t)\ell}} \leq \frac{2^{(s-t)\ell-1} \cdot \varepsilon_{\overline{h}}}{2^{(s-t)\ell}} \leq \varepsilon_{\overline{h}}/2. \tag{25}$$

Let y_0 and y_1 be the pair of elements returned by S^* on a successful cheat. Equation (25) yields that the probability that both y_0 and y_1 are inside $\mathcal{V}$ is at most $\varepsilon_{\overline{h}}/2$. It follows that the probability that S^* cheats successfully while at least one of y_0 and y_1 is *outside* $\mathcal{V}$ is at least $\varepsilon_{\overline{h}}/2$. Note that each event where S^* cheats successfully and outputs an element $y_i = y$, contributes half its probability to the total weight of y. Thus, the sum of weights of the elements in $\text{Consist}(\overline{h})\backslash\mathcal{V}$ is at least $\varepsilon_{\overline{h}}/4$. □

We conclude that

$$\begin{aligned}
&\Pr_{(\overline{h},y)\leftarrow \mathsf{D}_{\text{Ideal}}}\left[\mathsf{Inverter}(\overline{h}) \in \mathcal{W}_y \ \wedge \ (\overline{h}, y) \notin \text{Secluded}\right] \\
&= \mathrm{E}_{\overline{h}\leftarrow\mathcal{H}^t}\left[\frac{1}{|\text{Consist}(\overline{h})|} \cdot \sum_{y\in \text{Consist}(\overline{h})\backslash \text{Secluded}_{\overline{h}}} w_{\overline{h}}(y)\right] \\
&\geq \frac{1}{4\mu_t} \cdot \mathrm{E}_{\overline{h}\leftarrow\mathcal{H}^t}\left[\chi_{\text{Typical}}(\overline{h}) \cdot \sum_{y\in \text{Consist}(\overline{h})\backslash \text{Secluded}} w_{\overline{h}}(y)\right] \\
&\geq \frac{1}{4\mu_t} \cdot \mathrm{E}_{\overline{h}\leftarrow\mathcal{H}^t}\left[\chi_{\text{Typical}}(\overline{h}) \cdot \varepsilon_{\overline{h}}/4\right] \\
&\geq \frac{1}{64\mu_t} \cdot \varepsilon,
\end{aligned}$$

where the penultimate inequality is by Claim 4.14, and the last one by Eq. (23). □

5. Conclusions

An interesting open question regards the optimality of the binding guarantee given in Theorem 3.7. Particularly, is there a linear-preserving reduction from the security of an interactive hashing protocol to satisfying the underlying relation?[10] There are three possible directions for improvement: (1) use a different interactive hashing protocol than the one considered in Theorem 3.7, (2) use a different implementation for $\mathsf{A}^{\mathsf{S}^*}$ to satisfy the relation, or (3) improve the analysis of $\mathsf{A}^{\mathsf{S}^*}$ success probability.

We mention that our improvement in parameters over the NOVY proof is mainly in the third item (i.e., the analysis of the reduction). In the following we show that our analysis cannot be pushed to show a linear reduction. Namely, we present a (non-efficient) adversary S^* that breaks the binding of $\text{NOVY}\langle\mathcal{H}^n\rangle$ with probability ε (in

[10] In a linear-preserving reduction, the time-success ratio of an adversary violating the hardness of the relation, is larger than the time-success ratio of an adversary breaking the binding of the interactive hashing protocol, by at most a multiplicative polynomial factor.

the meaning of Eq. (1)), where $\mathcal{H}$ is a family of Boolean pairwise-independent hash functions, but $\mathsf{A}^{\mathsf{S}^*}$ only breaks the underlying relation (in this case a relation imposed by a permutation) with probability $O(\varepsilon^{1.4})$.

For a fixed $\varepsilon > 0$ consider the following cheating sender S^* for NOVY$\langle\mathcal{H}^n\rangle$: the cheating sender S^* answers the first $(n - \log\frac{1}{\varepsilon})$ questions with arbitrary Boolean answers, then it randomly selects two distinct elements $y_1, y_2 \in \{0,1\}^n$ that are consistent with the protocol so far and answers the remaining hash functions as follows: on $h \in \mathcal{H}$, it checks whether $h(y_1) = h(y_2)$, if positive it sends $h(y_1)$ back to the receiver; otherwise, it raises a flag and answers at random. At the end of the protocol, if the flag was not raised S^* inverts f on both y_1 and y_2 (in a brute force manner) and outputs the result.

The pairwise independence of $\mathcal{H}$ yields that S^* breaks the binding of NOVY$\langle\mathcal{H}^n\rangle$ with provability ε. Where in order for $\mathsf{A}^{\mathsf{S}^*}(y)$ to find $x \in \mathcal{W}_y$, it first has to be the case that $y \in \{y_1, y_2\}$; since the number of elements consistent with the protocol after $(n - \log\frac{1}{\varepsilon})$ steps is concentrated around $1/\varepsilon$ (see Claim 4.7), the latter happens with probability $O(\varepsilon)$. Assuming that y is one of $\{y_1, y_2\}$ (say that $y = y_1$), for succeeding $\mathsf{A}^{\mathsf{S}^*}$ has to choose in each step a hash function h with $\mathsf{S}^*(h) = h(y) = h(y_2)$. The pairwise independence of $\mathcal{H}$ yields that $\Pr_{h\leftarrow\mathcal{H}}[\mathsf{S}^*(h) \neq h(y_2) \mid \mathsf{S}^*(h) = h(y)] = \frac{1}{4}$. Therefore, the probability that $\mathsf{S}^*(h) = h(y) = h(y_2)$ in each of the last $\log\frac{1}{\varepsilon}$ steps, is at most $(\frac{3}{4})^{\log\frac{1}{\varepsilon}} < \varepsilon^{0.4}$ (note that A has no clue what y_2 is). We conclude that the $\mathsf{A}^{\mathsf{S}^*}$'s overall success probability is $O(\varepsilon \cdot \varepsilon^{0.4}) = O(\varepsilon^{1.4})$.

Yet, the existence of more security preserving reductions for NOVY$\langle\mathcal{H}^n\rangle$ (or more generally, to any protocol that follows the NOVY paradigm), not to mention the existence of different protocols with better security preserving reductions, remains an interesting open question.

Acknowledgements

We are grateful to Moni Naor and Ronen Shaltiel for helpful conversations. We are also grateful to Oded Goldreich and the anonymous referees for their many useful comments and suggestions.

Appendix A. Statistically Hiding Commitment from Regular One-Way Functions

In this section we use the interactive hashing theorem from Sect. 3 to give construct statistically hiding commitment from regular one-way functions, reproving [8, Theorem 4.4].

Theorem A.1. *Assume there exist regular one-way functions, then there exist statistically hiding and computationally binding commitment schemes.*[11]

[11] [8, Theorem 4.4] also holds with respect to somewhat more general choice of one-way functions. Specifically, [8] consider the case where the number of preimages is not fixed for all outputs, but rather can be efficiently approximated. As in [8], the proof of Theorem A.1 can be easily extended to such functions.

A commitment scheme is a two-stage protocol between a sender and a receiver. In the first stage, called the *commit stage*, the sender commits to a private string σ. In the second stage, called the *reveal stage*, the sender reveals σ and *proves* that it was the value to which she committed in the first stage. We require two properties of commitment schemes. The hiding property says that the receiver learns nothing about σ in the commit stage. The binding property says that after the commit stage, the sender is bound to a particular value of σ; that is, she cannot successfully open the commitment into two different values in the reveal stage. In a statistically hiding and computationally binding commitment scheme, the hiding holds information theoretically (i.e., even an all powerful learns nothing about σ), where the binding property only guaranteed to hold against polynomial-time senders. A (known) regular one-way function is an efficiently computable function that is hard to invert, and all its images have the same, efficiently computable, number of preimages. For the formal definitions of these primitives, see for example [8].

Proof of Theorem A.1. We use our new interactive hashing theorem to get a *weakly* hiding bit commitment scheme (see Definition A.4), and the existence of a full-fledge commitment scheme follows via standard amplification techniques (cf., [8]). The heart of the construction is applying the new interactive hashing protocol to a random output of the regular one-way function. This simplifies over the construction of [8], which uses an additional hashing layer before applying the interactive hashing protocol.

Let $f\colon \{0,1\}^n \mapsto \{0,1\}^n$ be a regular one-way function,[12] let $\mathrm{Im}_f(n) = \{f(x) : x \in \{0,1\}^n\}$ and let $s = s(n) = \lfloor \log |\mathrm{Im}_f(n)| \rfloor - 5$ (note that the regularity of f yields that s is polynomial-time computable). Let $\mathcal{H}$ and $\mathcal{G}$ be efficient Boolean pairwise-independent function families over strings of length n and let $(\mathsf{S_{IH}}, \mathsf{R_{IH}}) = \mathrm{NOVY}\langle \mathcal{H}^s \rangle$. We define the bit commitment protocol $\mathsf{Com} = (\mathsf{S}, \mathsf{R})$ as follows:

Protocol A.2 $\mathsf{Com} = (\mathsf{S} = (\mathsf{S}_c, \mathsf{S}_r), \mathsf{R} = (\mathsf{R}_c, \mathsf{R}_r))$.

Commit stage:

Common input: 1^n.
S_c**'s input:** $b \in \{0,1\}$.

1. S_c *chooses uniformly at random* $x \in \{0,1\}^n$ *and sets* $y = f(x)$.
2. S_c *and* R_c *interacts in a random execution of* $(\mathsf{S_{IH}}(y), \mathsf{R_{IH}})(1^n)$, *with* S_c *and* R_c *acting* $\mathsf{S_{IH}}$ *and* $\mathsf{R_{IH}}$, *respectively.*
   Let $(\overline{h}, \overline{z})$ *be the output of* $\mathsf{R_{IH}}$ *in this execution.*
3. S_c *chooses uniformly at random* $g \in \mathcal{G}$ *and sends* g, $c = b \oplus g(y)$ *to* R.
4. S_c *locally outputs* x *and* R_c *outputs* $(\overline{h}, \overline{z}, g, c)$.

Reveal stage:

Common input: 1^n, $b \in \{0,1\}$ *and* $(\overline{h}, \overline{z}, g, c)$.
S_r**'s input:** $x \in \{0,1\}^n$.

1. S_r *sends* x *to* R_r.
2. R_r *accepts if* $\overline{h}(f(x)) = \overline{z}$ *and* $g(f(x)) \oplus b = c$.

[12] The assumption that f is length-preserving is without lost of generality, see [3,11].

Lemma A.3 states that Protocol A.2 is computationally binding, where Lemma A.5 states is weakly hiding. Thus, the proof of Theorem A.1 follows by standard amplification techniques (e.g., [8, Theorem 5.2]). □

Lemma A.3. *Protocol A.2 is computationally binding.*

Proof. Let $\mathcal{W} = \{(x, f(x)) \colon x \in \{0,1\}^n\}$. The regularity of f yields that it is hard to satisfy $\mathcal{W}$ over a random element of $\mathrm{Im}_f(n)$. Hence, the proof follows by the binding of $(\mathsf{S}_{\mathsf{IH}}, \mathsf{R}_{\mathsf{IH}})$ as stated in Theorem 3.7, taking $\mathcal{L} = \mathrm{Im}_f(n)$. □

Definition A.4 (Weakly hiding commitment). A commitment scheme $\mathsf{Com} = (\mathsf{S} = (\mathsf{S}_c, \mathsf{S}_r), \mathsf{R} = (\mathsf{R}_c, \mathsf{R}_r))$ is $\delta = \delta(n)$-hiding, if

$$\mathsf{SD}\left(\left\{\mathrm{view}_{\mathsf{R}^*}\left(\mathsf{S}_c(0), \mathsf{R}^*\right)\left(1^n\right)\right\}_{n\in\mathbb{N}}, \left\{\mathrm{view}_{\mathsf{R}^*}\left(\mathsf{S}_c(1), \mathsf{R}^*\right)\left(1^n\right)\right\}_{n\in\mathbb{N}}\right) \leq \delta(n),$$

for any algorithm R^*, where $\mathrm{view}_{\mathsf{R}^*}(\mathsf{S}_c(b), \mathsf{R}^*)$ denotes the view of R^* in the commit stage interacting with $\mathsf{S}_c(b)$.

Lemma A.5. *Protocol A.2 is $\frac{3}{4}$-hiding.*

Proof. Let R^* be an adversary playing the role of R_c in Com and assume for the ease of notation that R^* never causes the sender to abort. For $b \in \{0,1\}$, let $V_{\mathsf{R}^*}(b) = (V_{\mathsf{IH}}, G, G(Y) \oplus b)$ denote R^*'s view in $(\mathsf{S}_c(b), \mathsf{R}^*)$, where V_{IH} is R^*'s view in the embedded execution of $(\mathsf{S}_{\mathsf{IH}}, \mathsf{R}_{\mathsf{IH}})$, and G and Y are the values of g and y chosen by S_c in the interaction. Note that V_{IH} is independent of b. Let $\mathrm{Bad} = \{v \in \mathrm{Supp}(V_{\mathsf{IH}}) \colon \mathrm{H}_\infty(Y \mid v) < 3\}$. Claim 3.6 yields that

$$\Pr[V_{\mathsf{IH}} \in \mathrm{Bad}] \leq \frac{1}{4}. \tag{A.1}$$

The Leftover Hash Lemma [12, Lemma 4.8] yields the following for every $v \notin \mathrm{Bad}$ and $b \in \{0,1\}$:

$$\Delta\left(V_{\mathsf{R}^*}(b), \widetilde{V}_{\mathsf{R}^*} \mid V_{\mathsf{IH}} = v\right) = \Delta\left(\left(v, G, G(Y) \oplus b\right), (v, G, U) \mid V_{\mathsf{IH}} = v\right) \leq \frac{1}{4} \tag{A.2}$$

where $\widetilde{V}_{\mathsf{R}^*}$ is obtained from $V_{\mathsf{R}^*}(b)$ be replacing the value of $(G(y) \oplus b)$ sent by the sender with U—a uniformly chosen bit. We conclude that

$$\begin{aligned}
&\Delta\left(V_{\mathsf{R}^*}(0), V_{\mathsf{R}^*}(1)\right) \\
&\quad \leq \Delta\left(V_{\mathsf{R}^*}(0), V_{\mathsf{R}^*}(1) \mid V_{\mathsf{IH}} \notin \mathrm{Bad}\right) + \Pr[V_{\mathsf{IH}} \in \mathrm{Bad}] \\
&\quad \leq \Delta\left(V_{\mathsf{R}^*}(0), \widetilde{V}_{\mathsf{R}^*} \mid V_{\mathsf{IH}} \notin \mathrm{Bad}\right) + \Delta\left(V_{\mathsf{R}^*}(1), \widetilde{V}_{\mathsf{R}^*} \mid V_{\mathsf{IH}} \notin \mathrm{Bad}\right) + \Pr[V_{\mathsf{IH}} \in \mathrm{Bad}] \\
&\quad \leq \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}.
\end{aligned}$$

□

References

[1] G. Brassard, D. Chaum, C. Crépeau, Minimum disclosure proofs of knowledge. *J. Comput. Syst. Sci.* **37**(2), 156–189 (1988)

[2] L.J. Carter, M.N. Wegman, Universal classes of Hash functions. *J. Comput. Syst. Sci.* **18**(2), 143–154 (1979)

[3] O. Goldreich, *Foundations of Cryptography: Basic Tools* (Cambridge University Press, Cambridge, 2001)

[4] O. Goldreich, S. Goldwasser, N. Linial, Fault-tolerant computation in the full information model. *SIAM J. Comput.* **27**, 447–457 (1998)

[5] I. Haitner, O. Reingold, Statistically hiding commitment from any one-way function, in *Proceedings of the 39th Annual ACM Symposium on Theory of Computing (STOC)* (2007), pp. 1–10

[6] I. Haitner, O. Reingold, A new interactive hashing theorem, in *Proceedings of the 18th Annual IEEE Conference on Computational Complexity* (2007), pp. 319–332

[7] I. Haitner, J.J. Hoch, O. Reingold, G. Segev, Finding collisions in interactive protocols—a tight lower bound on the round complexity of statistically-hiding commitments, in *Proceedings of the 48th Annual Symposium on Foundations of Computer Science (FOCS)* (2007), pp. 669–679

[8] I. Haitner, O. Horvitz, J. Katz, C. Koo, R. Morselli, R. Shaltiel, Reducing complexity assumptions for statistically hiding commitment. *J. Cryptol.* **22**(3), 283–310 (2009)

[9] I. Haitner, M. Nguyen, S.J. Ong, O. Reingold, S. Vadhan, Statistically hiding commitments and statistical zero-knowledge arguments from any one-way function. *SIAM J. Comput.* **39**(3), 1153–1218 (2009). Preliminary versions in *FOCS'06* and *STOC'07*

[10] I. Haitner, O. Reingold, S. Vadhan, H. Wee, Inaccessible entropy, in *Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC)* (2009), pp. 611–620

[11] I. Haitner, D. Harnik, O. Reingold, On the power of the randomized iterate. *SIAM J. Comput.* **40**(6), 1486–1528 (2011). Preliminary version in *Crypto'06*

[12] J. Håstad, R. Impagliazzo, L.A. Levin, M. Luby, A pseudorandom generator from any one-way function. *SIAM J. Comput.* **28**, 1364–1396 (1999). Preliminary versions in *STOC'89* and *STOC'90*

[13] T. Koshiba, Y. Seri, Round-efficient one-way permutation based perfectly concealing bit commitment scheme. Technical Report TR06-093, ECCC (2006). http://eccc.hpi-web.de/report/2006/093/

[14] Y. Lindell, Parallel coin-tossing and constant-round secure two-party computation. *J. Cryptol.* **16**(3), 143–184 (2003)

[15] M. Naor, M. Yung, Universal one-way Hash functions and their cryptographic applications, in *Proceedings of the 21st Annual ACM Symposium on Theory of Computing (STOC)* (1989), pp. 33–43

[16] M. Naor, R. Ostrovsky, R. Venkatesan, M. Yung, Perfect zero-knowledge arguments for NP using any one-way permutation. *J. Cryptol.* **11**(2), 87–108 (1998). Preliminary version in *CRYPTO'92*

[17] M. Nguyen, S.J. Ong, S. Vadhan, Statistical zero-knowledge arguments for NP from any one-way function, in *Proceedings of the 47th Annual Symposium on Foundations of Computer Science (FOCS)* (2006), pp. 3–14

[18] R. Ostrovsky, R. Venkatesan, M. Yung, Secure commitment against all powerful adversary, in *9th Annual Symposium on Theoretical Aspects of Computer Science* (1992), pp. 439–448

[19] R. Ostrovsky, R. Venkatesan, M. Yung, Fair games against an all-powerful adversary. *AMS DIMACS Ser. Discrete Math. Theor. Comput. Sci.* **13**, 155–169 (1993). Preliminary version in *SEQUENCES'91*

[20] R. Ostrovsky, R. Venkatesan, M. Yung, Interactive hashing simplifies zero-knowledge protocol design, in *Advances in Cryptology—EUROCRYPT'93* (1993), pp. 267–273

[21] H. Wee, One-way permutations, interactive hashing and statistically hiding commitments, in *Theory of Cryptography, Fourth Theory of Cryptography Conference, TCC 2007* (2007), pp. 419–433