

Cryptography from Lossy Reductions: Towards OWFs from ETH, and Beyond

Pouria Fallahpour¹, Alex B. Grilo¹, Garazi Muguruza², and Mahshid Riahinia^{3*}

¹ Sorbonne Université, CNRS and LIP6, France

² QuSoft, Informatics Institute, University of Amsterdam, Netherlands

³ DIENS, École Normale Supérieure, CNRS, Inria, PSL University, Paris, France

(May 27, 2025)

Abstract. One-way functions (OWFs) form the foundation of modern cryptography, yet their unconditional existence remains a major open question. In this work, we study this question by exploring its relation to lossy reductions, *i.e.*, reductions R for which it holds that $I(X; R(X)) \ll n$ for all distributions X over inputs of size n . Our main result is that either OWFs exist or any lossy reduction for any promise problem Π runs in time $2^{\Omega(\log \tau_\Pi / \log \log n)}$, where $\tau_\Pi(n)$ is the infimum of the runtime of all (worst-case) solvers of Π on instances of size n . More precisely, by having a reduction with a better runtime, for an arbitrary promise problem Π , and by using a non-uniform advice, we construct (a family of) OWFs. In fact, our result requires a milder condition, that R is lossy for *sparse uniform* distributions (which we call mild-lossiness). It also extends to f -reductions as long as f is a non-constant permutation-invariant Boolean function, which includes AND-, OR-, MAJ-, PARITY-, MOD $_k$ -, and THRESHOLD $_k$ -reductions.

Additionally, we show that worst-case to average-case Karp reductions and randomized encodings are special cases of mildly-lossy reductions and improve the runtime above as $2^{\Omega(\log \tau_\Pi)}$ when these mappings are considered. Restricting to weak fine-grained OWFs, this runtime can be further improved as $\Omega(\tau_\Pi)$. Intuitively, the latter asserts that if weak fine-grained OWFs do not exist then any instance randomization of any Π has the same runtime (up to a constant factor) as the best worst-case solver of Π .

Taking Π as $k\text{SAT}$, our results provide sufficient conditions under which (fine-grained) OWFs exist assuming the Exponential Time Hypothesis (ETH). Conversely, if (fine-grained) OWFs do not exist, we obtain impossibilities on instance compressions (Harnik and Naor, FOCS 2006) and instance randomizations of $k\text{SAT}$ under the ETH. Moreover, the analysis can be adapted to studying such properties of any NP-complete problem.

Finally, we partially extend these findings to the quantum setting; the existence of a pure quantum mildly-lossy reduction for Π within the runtime $2^{o(\log \tau_\Pi / \log \log n)}$ implies the existence of one-way state generators, where τ_Π is defined with respect to quantum solvers.

Keywords: one-way functions, lossy reductions, randomized encodings, worst-case to average-case reductions, instance compression, exponential time hypothesis

*Part of this work was done when the author was visiting IRIF, Université Paris Cité, Paris, France.

Table of Contents

Cryptography from Lossy Reductions: Towards OWFs from ETH, and Beyond	1
<i>Pouria Fallahpour, Alex B. Grilo, Garazi Muguruza, and Mahshid Riahinia</i>	
1 Introduction	3
1.1 Our Contribution	3
1.2 Technical Overview	6
1.3 Background and Related Works.	12
1.4 Open Questions.	13
2 Preliminaries	14
3 Lossy Mappings and Disguising Lemma	20
4 Mildly-Lossy Problems	25
4.1 f -Distinguisher Reductions	25
4.2 Mildly-Lossy Problems.	29
5 Zero-Knowledgeness from Mildly-Lossy Problems	30
6 One-Way Functions from Mildly-Lossy Problems	32
7 One-Way State Generators from Mildly-Lossy Problems	38
8 Mild-Lossiness and Instance Randomization	40
8.1 Worst-Case to Average-Case Reductions	40
8.2 Randomized Encodings	45
8.3 Quantum Worst-Case to Average-Case Reductions	46
9 Applications: Hardness vs One-Wayness	48
9.1 Quantum Hardness vs Quantum One-Wayness	53

1 Introduction

One-way functions (OWFs) are essential cryptographic tools and can be viewed as the minimal assumption required for cryptography. Informally, a function is called one-way if it is easy to compute but hard to invert. The existence of one-way functions implies that of many cryptographic primitives such as pseudorandom generators and functions [BM82, Yao82, HILL99, GGM86], commitments schemes [Nao91] and zero-knowledge proofs [GMW91]. Given their centrality, numerous works are dedicated to constructing OWFs. Although it is unknown whether they unconditionally exist, several candidate constructions have been proposed assuming the hardness of concrete computational problems such as discrete logarithm [DH76], lattice-based problems [Ajt98, MR07, Reg09], and more. Instead of depending on the hardness of specific problems, the pinnacle result in this direction would be to construct OWF from minimal computational complexity assumptions such as $\text{NP} \neq \text{P}$, or $\text{NP} \not\subseteq \text{BPP}$, or $\text{NP} \not\subseteq \text{non-uniform-P}$. However, many works [FF93, BT06, AGGM06, BB15] have shown barriers in this direction.¹

A possibly more feasible direction, therefore, is to slightly relax the above conditions by replacing P (and BPP and non-uniform-P) with subexponential-time algorithms. This is because, despite the huge effort that has been made in the literature, no subexponential-time algorithm is known for NP -complete problems and most notably for the variants of SAT . Recall that the $k\text{SAT}$ problem asks to decide whether a CNF formula of N variables and M clauses, where each clause has k variables, has a satisfiable assignment. The subexponential-time hardness of NP -complete problems has been formulated in the variants of the *Exponential Time Hypothesis* (ETH). Informally, the exponential time hypothesis states that there is no algorithm that can solve $k\text{SAT}$ in time subexponential in the number of variables N . This leads us to the following question:

*Do one-way functions exist under the exponential time hypothesis (ETH)?
Otherwise, what would be the implications for the hardness of SAT?*

Ball, Rosen, Sabin and Vasudevan [BRSV17] have asked a similar question about the existence of *weak fine-grained* one-way functions from ETH, which remains open. A weak fine-grained one-way function requires (i) an attacker to fail in inverting the function with non-negligible probability (as opposed to negligible probability for OWFs) and (ii) that there exists a fixed *polynomial* gap (as opposed to super-polynomial for OWFs) between the runtime of the function and that of the attacker.

1.1 Our Contribution

Trying to answer the above question, we study *lossy* reductions. A reduction R is lossy if it loses information about its input; it should hold that the mutual information between the input and output of R is very small, *i.e.*, $I(X; R(X)) \ll n$ for all distributions X on inputs of size n . For example, a special type of lossy reductions is compressions that map n bits into $\lambda \ll n$ bits. In this work, we consider a less restrictive notion that we call *mild-lossiness*: it requires that the same inequality holds for sparse uniform distributions X over inputs of size n .² We prove the following (informal) theorem:

¹We briefly explain these works later in this section.

²More precisely, the distribution X has a support of size $2^{o(n)}$. In fact, for our results to hold, X can be even more sparse depending on the upper bound on the runtime of R . See section 4 for more details.

Result 1 (OWFs from Mildly-Lossy or Worst-to-Average-Case Reductions). *Let Π be a promise problem, and let τ be the infimum of the runtime of all worst-case solvers of Π . We construct a family of non-uniform functions F_Π , such that either F_Π is a one-way function, or (i) any mildly-lossy Karp reduction from Π (to any other problem), given an input instance of size n , has runtime $2^{\Omega(\log \tau / \log \log n)}$, and (ii) any worst-case to average-case Karp reduction from Π (to any other problem), given an input instance of size n , has runtime $2^{\Omega(\log \tau)}$.*

In the above statement, the worst-case to average-case Karp reduction from Π can be replaced by randomized encodings for Π . Moreover, we obtain a variant of the above statement regarding weak fine-grained one-way functions.

Result 2 (Weak Fine-Grained OWFs from Worst-to-Average-Case Reductions). *Let Π be a promise problem, and let τ be the infimum of the runtime of all worst-case solvers of Π . We construct a family of non-uniform functions F_Π , such that either F_Π is a weak fine-grained one-way function or any worst-case to average-case Karp reduction from Π (to any other problem) runs in time $\Omega(\tau)$.*

In other words, the above statements assert that if one-way functions do not exist, then randomizing or compressing the worst-case instances of a problem Π has roughly polynomially-better runtime as solving these instances. In the case of non-existence of fine-grained one-way functions, randomizing worst-case instances takes roughly the same time as solving them.

Our results are quite flexible in different ways. Firstly, we prove the above statements for the general set of f -reductions. More precisely, Drucker [Dru15] defines these reductions as follows: let Π be a promise problem, and let χ_Π be the characteristic function of Π , i.e., for an input x , $\chi_\Pi(x) = 1$ if x is a YES instance of Π , and 0 otherwise. For a function $f : \{0, 1\}^m \rightarrow \{0, 1\}$, an f -reduction R from Π to Π' is such that on input m instances of Π , the output $R(x_1, \dots, x_m)$ is a YES instance of Π' iff $f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m)) = 1$. Our results hold with respect to f -reductions for any non-constant permutation-invariant function f , such as OR, AND, MAJ, PARITY, MOD $_k$, and THRESHOLD $_k$. Moreover, Π is not necessarily confined to NP problems. Finally, our proofs relativize; the theorems hold even when all of the considered algorithms have access to a common arbitrary oracle.

Our results are obtained by proposing clear and generic definitions that facilitate analysis and by closely analyzing the mild-lossiness of special well-known reductions, including *instance compressions*, *worst-case to average-case reductions* and *randomized encodings*. precisely, we relate the concrete mild-lossiness of these cases to, among others, the error of the reduction, the privacy of the randomized encoding, or the distance of the output distribution of the worst-case to average-case reduction from the average-case distribution. Our analysis allows a wide range of parameters. For instance, for the aforementioned theorems to hold, the error of the randomized encoding or worst-to-average reductions can be any constant smaller than 2^{-19} and the privacy or distance from the average-case distribution can be as large as $\approx 2^{-1.5 \log(\tau)}$ (see Section 9 for more details).

We can then use these general results to study the existence of OWFs from k SAT (and other NP-complete problems).

Result 3 (OWFs from Mildly-Lossy or Worst-to-Average-Case Reductions from k SAT). *We construct a family of non-uniform functions $F_{k\text{SAT}}$ such that, under the ETH, either $F_{k\text{SAT}}$ is a one-way function or for any non-constant permutation-invariant function f , (i) any mildly-lossy Karp f -reduction from $k\text{SAT}$ (to any other problem), given an input instance of size n , has runtime*

$2^{\Omega(n/(\log n \cdot \log \log n))}$, and (ii) any worst-case to average-case Karp f -reduction from $k\text{SAT}$ (to any other problem), given an input instance of size n , has runtime $2^{\Omega(n/\log n)}$.

For a better comparison, note that $k\text{SAT}$ has a worst-case solver that runs in time $2^{O(n/\log n)}$ but assuming ETH it cannot be solved in time $2^{o(n/\log n)}$ (see Section 9 for more details). Interestingly, the first item implies that if one-way functions do not exist, then for any $\varepsilon < 1$, any f -compression reduction [Dru15] of $k\text{SAT}$ that maps mn bits to mn^ε bits runs in nearly exponential time.

We also instantiate Result 2 with $k\text{SAT}$.

Result 4 (Weak Fine-Grained OWFs from Worst-to-Average-Case Reductions from $k\text{SAT}$). *We construct a family of non-uniform functions $F_{k\text{SAT}}$ such that, under the ETH, either $F_{k\text{SAT}}$ is a weak fine-grained one-way function or for any non-constant permutation-invariant function f , any worst-case to average-case Karp f -reduction from $k\text{SAT}$ (to any other problem), given an input instance of size n , has runtime $\Omega(2^{cn/\log n})$, for some constant c . Note that c is such that any solver for $k\text{SAT}$ runs in time $\Omega(2^{cn/\log n})$ by the ETH.*

Again, in both Results 3 and 4, one can replace worst-case to average-case reductions by randomized encodings. Moreover, these results can be adapted to any of the following problems: CLIQUE , VERTEXCOVER , INDEPENDENTSET , $k\text{SETCOVER}$, or $k\text{COLORABILITY}$. This is a direct consequence of NP-completeness under subexponential-time reductions (e.g., see [IPZ98]).

Result 4 opens up a new direction for non-uniform constructions of fine-grained one-way functions by discovering “slightly better than trivial” instance randomizations of NP-complete problems (see Theorem 9, and Corollary 7 for the details). This draws a new approach to address the aforementioned question raised by Ball, Rosen, Sabin and Vasudevan [BRSV17], regarding the existence of weak fine-grained OWFs from the ETH.

Additionally, we answer an open question raised by Drucker [Dru15] regarding the f -compression reductions of 3SAT . The main result of Drucker is refuting strong OR or AND compressions for 3SAT under the assumption that $\text{NP} \not\subseteq \text{SZK}/\text{Poly}$, and their techniques cannot directly exclude more general functions. Recall that SZK is the class of all languages that have an interactive proof where a malicious verifier learns almost nothing beyond the membership of the instance in the language. The extension of their result, using the techniques of [FS08], to f -compression reductions for any function f that depends on all of its input bits for each input length, have some caveats. For an arbitrary f , the compression must be to another problem in NP, unless f is monotone, and at the same time the range of covered parameters are somewhat weaker than those of OR and AND. In this work, we show the following:

Result 5 (f -Compression Implies SZK). *If a problem Π has a f -compression reduction that maps m instances of n bits to $m\lambda$ bits, then Π can be reduced to SZK/Poly in time $2^{O(\lambda + \log n)}$. In this case, there is no compressing f -compression reduction of 3SAT for any non-constant permutation-invariant functions f with the same range of parameters from [Dru15], unless $\text{NP} \subseteq \text{SZK}/\text{Poly}$.*

We notice that our result gives a framework to study f -compression reductions of NP-complete problems under superpolynomial-time algorithms, and we leave as an open question exploring this direction.

Quantum Settings We initiate the study of cryptographic implications of quantum mildly-lossy reductions. A quantum reduction R is said to be mildly-lossy when $I_q(X; R(X)) \ll n$ for all sparse

uniform distributions X on inputs of size n , where I_q is the quantum mutual information. Moreover, such a reduction is said to be a pure-outcome reduction if (i) for every instance x the outcome $R(x)$ is a pure quantum state (ii) and there exists a (possibly unbounded) binary quantum measurement that, given $R(x)$, decides x . We obtain partial results in the this regard. More precisely, we show that such reductions imply one-way state generators (OWSGs); a type of quantum functions that are easy to evaluate but hard to invert.

Result 6 (OWSGs from Quantum Mildly-Lossy Reductions). *Let Π be a promise problem, and let τ^Q be the infimum of the runtime of all quantum worst-case solvers of Π . We construct a family of non-uniform quantum mappings G_Π , such that either G_Π is a one-way state generator, or any quantum mildly-lossy pure-outcome Karp reduction from Π (to any other problem), given an input instance of size n , has runtime $2^{\Omega(\log \tau^Q / \log \log n)}$.*

1.2 Technical Overview

In this section, we briefly present the core technical tools that we use.

The link between lossy reductions with the randomized encodings and worst-case to average-case reductions was raised by [BBD⁺20], but the exact connection was left as an open question, which we answer in a precise manner. We define a more inclusive type of lossy reductions: *mildly-lossy f -distinguisher reductions*. Such reductions include randomized encodings, compressions, and a variant of worst-case to average-case non-adaptive Turing reductions. We show that these weaker reductions can also be used to build one-way functions. The full-fledged lossiness of randomized encodings and worst-case to average-case reductions is only satisfied in a very restricted regime of parameters, *e.g.*, when the error is zero and the privacy or distance is exponentially-small. Our new definition allows to significantly relax the parameters (see Section 9 for more details).

f -distinguisher reductions. For a Boolean function $f : \{0,1\}^m \rightarrow \{0,1\}$, we define an f -distinguisher reduction for a problem Π as a mapping $R : \{0,1\}^* \rightarrow \{0,1\}^*$ for which there exists an unbounded distinguisher \mathcal{D} that can distinguish between $R(x_1, \dots, x_m)$ and $R(x'_1, \dots, x'_m)$, also given one of $\{x_i\}_i$'s at random, if $f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m)) \neq f(\chi_\Pi(x'_1), \dots, \chi_\Pi(x'_m))$. We show that all non-adaptive Turing reductions, most importantly Karp reductions, are special cases of f -distinguisher reductions. Moreover, f -distinguisher reductions contain f -compression reductions that are studied in the context of parameterized complexity (*e.g.*, see [HN06, FS08, Dru15]) and randomized encodings [IK00, AIK06, App17] of the characteristic function χ_Π are of this type. Our results are therefore stated in terms of this general flavor of reductions.

Mild lossiness. Originally in [BBD⁺20] a multivariate mapping R is said to be t -lossy if the quantity $I((X_1, \dots, X_m); R(X_1, \dots, X_m))$ is bounded above by t for *all* possible distributions X_i over n -bit strings. We propose an alternative definition that we call *mild lossiness*.

Definition 1 (Informal). *We say that R is (λ, γ) -mildly-lossy if*

$$\sup_{X_1, \dots, X_m} \{I((X_1, \dots, X_m); R(X_1, \dots, X_m))\} \leq \lambda m,$$

where each X_i ranges over all uniform distributions of support-size roughly $\tilde{O}(1/\gamma^3)$.

The parameter γ controls the sparseness of the distribution. In the original lossiness, γ is exponentially small, however, it can be fine-tuned depending on various parameters in our new setting. Moreover, if R is an f -reduction for Π , each X_i in the supremum above can be either supported on Π_{YES} , the set of YES instances of Π , or Π_{NO} , the set of NO instances of Π . In other words, $\{X_1, X_2, \dots, X_m\}$ can be split into Π_{YES} -supported and Π_{NO} -supported distributions.

An extended disguising lemma: We first enhance the disguising lemma of Drucker [Dru15]. Let $R : \{0, 1\}^* \rightarrow \{0, 1\}^*$ be a function and consider the problem of finding x given $R(x)$. Fano's inequality gives a lower bound for the amount of information about x that an unbounded algorithm can recover from $R(x)$, for any choice of x . For instance, if R is compressing, *i.e.*, it maps an input of size n to an input of size $\lambda < n$, then $R(x)$ loses information about x which makes it difficult to recover the instance.

The original variant of the disguising lemma by Drucker [Dru15] is a distinguishing variant of Fano's inequality which states that assuming R is a compressing map, for any set $S \subseteq \{0, 1\}^n$, there exists a *sparse* distribution D_S over S such that $\mathbb{E}_{D_S}[\|R(y) - R(D_S)\|_1] \leq \delta^*$ for all $y \in S$. Here, $\delta^* \approx 1 - 2^{-\lambda-2}$ if R compresses n -bit instance to λ bits.

We improve the disguising lemma by showing that mild-lossiness of R , instead of compression, suffices to obtain a similar result. In order to sketch our improvements, we briefly go over the proof of this lemma in the following. The proof of Drucker's lemma essentially consists of showing that as long as R is sufficiently compressing, it has the following property: Let \mathcal{Y} be any distribution and let (y, D) be a distribution obtained by sampling $d + 1$ instances from \mathcal{Y} , setting y to be one of them at random, and D to be the uniform over the d remaining samples. Then, we have

$$\mathbb{E}_{\mathcal{Y}^{\otimes d}}[\|R(y) - R(D)\|_1] \leq \delta^* . \quad (1)$$

The proof then proceeds by swapping the quantifiers of the above statement using the minimax theorem; more precisely, consider a simultaneous-move two-player game where one player chooses the distribution D (subject to be uniform over some multiset of size d) and the other player chooses the element y , and let the payoff be $\|R(y) - R(D)\|_1$. For any strategy \mathcal{Y} for choosing y , let (y, D) be as explained earlier with \mathcal{Y} being the base distribution. Then, Equation (1) bounds the expected payoff from above. By minimax theorem, there must exist a distribution \mathcal{D}_S , not necessarily sparse, that bounds the quantity $\mathbb{E}_{D \sim \mathcal{D}_S}[\|R(y) - R(D)\|_1]$ for every choice of y . However, note that \mathcal{D}_S can be not sparse. The final step of this proof, therefore, uses a result by Lipton and Young [LY94, Theorem 2] to freely set \mathcal{D}_S to be a uniform distribution over a sparse number of possible D 's. In fact, the theorem of [LY94, Theorem 2] roughly states that in a two-player simultaneous-move zero-sum game, restricting the strategies of Player 1 to uniform strategies with support size $\ln(\#\{\text{choices of Player 1}\})/\gamma^2$ only changes the optimal expectation payoff with an additive factor γ .¹ This indeed *sparsifies* the support of \mathcal{D}_S . On the other hand, one loses at most an additive factor γ in the expectation bound in Equation (1) and obtains $\delta^* + \gamma$.

Let us now focus on Equation (1). Drucker [Dru15] shows that this inequality holds if R is compressing. The work of [BBD⁺20] instead obtains Equation (1) by considering R to be, more generally, lossy. Recall that a mapping R is said to be λ -lossy if for all distributions X , it holds that $I(X; R(X)) \leq \lambda$, where I denotes the mutual information. We relax the requirement on R even further and show that mild-lossiness of R suffices to obtain a similar result. More precisely,

¹The same holds for Player 2.

we show that if the lossiness only holds with respect to the uniform distributions with support size $\tilde{O}(1/\gamma^3)$, then one loses nothing but an (other) additive factor γ in the expectation bound. This relies on a double use of the result by Lipton and Young [LY94, Theorem 2]; we apply it once for Player 2 and once more for Player 1. More precisely, before using the minimax theorem, we restrict the base distribution \mathcal{Y} to be uniform distributions with support size $\tilde{O}(1/\gamma^3)$, and we choose $d \approx 1/\gamma$. By showing that Equation (1) remains correct even with this new restriction, we obtain an additive γ -approximation of the value of the game (first use of [LY94, Theorem 2] for Player 2). Following the minimax theorem, and sparsifying \mathcal{D}_S (second use of [LY94, Theorem 2] for Player 1), we conclude the final upper bound $\delta^* + 2\gamma$. We note that this step is crucial for our results, otherwise, we could not sufficiently bound the lossiness of worst-case to average-case reductions or randomized encodings.

To be more precise, all of the above has been analyzed by Drucker [Dru15] in the setting where R is multivariate, *e.g.*, taking m instances as input. In this setting the disguising lemma, proved by Drucker, bounds the distance of $R(D_S, \dots, y, \dots, D_S)$ (where there are $m - 1$ samples of D_S and exactly one y in a random place) from $R(D_S, \dots, D_S)$ (where there are m samples of D_S), when R is a compression. In a similar way as above, we extend the disguising lemma in the multivariate setting, by relaxing the condition on R and showing that the distance of the two aforementioned distributions are bounded by $\delta^* + 2\gamma$ when R is mildly lossy, *i.e.*, lossy for the sparse uniform distributions.¹

Furthermore, [BBD⁺20] shows that the inputs can follow two distinct distributions and that the set S can be replaced by two sets S_0, S_1 . Consequently, the type of each input can be set to either S_0 or S_1 . Then, for any choice of $0 \leq p \leq m$, there exist two sparse distributions D_{S_0} and D_{S_1} of inputs such that, for every $y \in S_0$, the distance between $R(\pi(D_{S_0}, \dots, y, \dots, D_{S_1}))$ and $R(\pi(D_{S_0}, \dots, D_{S_0}, \dots, D_{S_1}))$ is at most $\delta^* + 2\gamma$ in expectation, where the number of D_{S_0} and D_{S_1} samples in the input of the latter is respectively p and $m - p$, and π is a uniformly random permutation. Similar result holds for replacing one of D_{S_1} 's with an arbitrary $y \in S_1$. In other words, $R(\pi(\cdot))$ remains roughly within the same distance (in expectation) if one of the input distributions D_{S_i} is replaced with an arbitrary $y \in S_i$ (note the constraint that y must have the same support as the distribution that it replaces). Our variant of disguising lemma with mildly-lossy reductions also extends to this setting (see Section 3 for more details). For simplifying the notation, we define

$$R_p[\star] := R(\pi(D_{S_0}, \dots, \star, \dots, D_{S_1})), \quad (2)$$

where the number of D_{S_0} and D_{S_1} samples in the input is respectively $p - 1$ and $m - p$, and \star can possess a fixed quantity or a random variable.

The disguising lemma forms the core of the following results. We start with showing, similarly to Drucker [Dru15], that a mildly-lossy problem, *i.e.*, a problem that admits a mildly-lossy reduction, has a reduction to SZK. The runtime of the reduction is determined by the amount of mild-lossiness. In comparison to [Dru15], our result holds for any non-constant permutation-invariant function, requires less restricted notion of mild-lossiness (as opposed to lossiness that is required in [Dru15]), and allows superpolynomial-time reductions.

Reduction to the statistical difference (SD) problem. In the statistical difference (SD) problem, the description of two circuits (C_0, C_1) is given with the promise that on uniformly random inputs their induced distributions are either at least $2/3$ -far or at most $1/3$ -far, with respect to the

¹In fact, m possibly changes the upper bound, but by tuning $d \approx m/\gamma$, one can keep the bound the same.

statistical distance. The question asks to decide which one is the case. This problem is complete for SZK under polynomial-time reductions. The parameters $1/3$ and $2/3$ can be replaced by any real numbers $\alpha, \beta \in (0, 1)$ as long as $\beta^2 > \alpha$. We sketch how mildly-lossy problems reduce to SD.

Let Π be a decision problem and R be any lossy function over m instances x_1, \dots, x_m of Π . Let $S_0 := \Pi_N \cap \{0, 1\}^n$ and $S_1 := \Pi_Y \cap \{0, 1\}^n$. By the disguising lemma, for any $0 \leq p \leq m$, there exist two sparse distributions D_{S_0} and D_{S_1} such that $\mathbb{E}[\|R_p[y] - R_p[D_{S_0}]\|_1] \leq \delta^* + 2\gamma$ for all $y \in S_0$ (recall $R_p[\star]$ as per Equation (2)). What is this quantity if $y \in S_1$? We show that it is large if R is an f -distinguisher reduction for some particular set of functions $f : \{0, 1\}^m \rightarrow \{0, 1\}$. Assume that f is a non-constant permutation-invariant function, *i.e.*, a non-constant function that is invariant under permuting its inputs. Let f_i be the evaluation of f over the inputs with i number of 0's. In fact, since f is permutation-invariant, only the number of 0's in the input determines the output. In the sequence f_0, f_1, \dots, f_m , there must be an index $1 \leq p \leq m$ such that $f_{p-1} \neq f_p$, because otherwise f is constant. Now, let us go back to our question. What is the expectation value if $y \in S_1$? In this case, the number of NO instances in the argument of $R_p[y]$ is equal to $p - 1$ while in $R_p[D_{S_0}]$ is p (note that D_{S_0} is supported on NO instances). Therefore, if R is also an f -distinguisher reduction with error μ^* , for all $y \in S_1$, it must hold that $\mathbb{E}[\|R_p[y] - R_p[D_{S_0}]\|_1] \geq 1 - \mu^*$. Putting these two properties of R together, one can conclude that an instance y of Π can be reduced to two circuits¹ ($R_p[y], R_p[D_{S_0}]$) such that

- if y is a NO instance, the two circuits have statistical distance at most $\delta^* + 2\gamma$,
- and if y is a YES instance, the two circuits have statistical distance at least $1 - \mu^*$.

This gives a reduction to SZK as long as $(1 - \mu^*)^2 - (\delta^* + 2\gamma)$ is a positive constant. The details of the extension to smaller quantities is discussed in Section 5.

One-way functions and one-way state generators: The circuits $R_p[D_{S_0}]$ and $R_p[y]$ use an internal randomness to sample from D_{S_0} and D_{S_1} . More precisely, they are two circuits that given uniformly random strings, sample elements from D_{S_0} and D_{S_1} , and return the evaluation of R . Since both these distributions are uniform over some given multisets of size $d \approx m/\gamma$ with n -bit elements, sampling one element requires $O(\log(m/\gamma))$ number of bits and, with an appropriate data structure, runs in $O(mn/\gamma)$ time. Moreover, if T_R is the runtime of the reduction, the total runtime (or size) of each circuit will be $O(T_R + (mn/\gamma)m)$. This is because there are approximately m inputs to be sampled, each of which requires $O(mn/\gamma)$ operations. We let C_0 and C_1 to be circuits, taking as input uniform bit strings, that denote respectively $R_p[D_{S_0}]$ and $R_p[y]$. When it is needed, we use $C_1[y]$ to denote the dependence of C_1 on y .

In [BBD⁺20], it is shown that when the reduction is perfect and Π is worst-case hard (with respect to polynomial-time algorithms), $C_0(\cdot)$ is a weak one-way function. We propose an alternative construction as follows:

$$F(b, r) := \begin{cases} C_0(r) & \text{if } b = 0, \\ C_1[y^*](r) & \text{if } b = 1, \end{cases} \quad (3)$$

when y^* is also sampled from D_{S_0} (supported on NO instances of Π). This function frequently appears in the SZK literature (*e.g.*, see [SV]) and was used in [BDRV19] to build one-way functions from the average-case hardness of the statistical difference problem.

We sketch the proof of one-wayness of F . Let \mathcal{A} be an inverter for F . We show how \mathcal{A} can be used to decide Π . One can use \mathcal{A} to decide every instance \hat{y} of Π as follows. Compute $C_0, C_1[\hat{y}]$,

¹From here forward we call them circuits instead of functions.

sample b at random, and feed \mathcal{A} with $C_b(r)$. If $b = 0$, then \mathcal{A} receives an instance of the function F and can therefore invert it. However, this does not help us with solving Π . Let us now focus on when $b = 1$. We have two cases: If \hat{y} is a NO instance, $C_1[\hat{y}]$ would be roughly close to $C_1[y^*]$, as discussed earlier. Therefore, \mathcal{A} would succeed to invert it. On the other hand, if \hat{y} is a YES instance, then C_0 and $C_1[\hat{y}]$ are far from each other. We also know that C_0 and $C_1[y^*]$ are close. Hence, $C_1[y^*]$ and $C_1[\hat{y}]$ must be far. Consequently, the image spaces of $C_1[y^*]$ and $C_1[\hat{y}]$ have small intersection. Therefore, if $b = 1$ and \hat{y} is a YES instance, then there would be no pre-image (except with a small probability) for the value that \mathcal{A} tries to invert. We can therefore run this test several times on \mathcal{A} and decides \hat{y} by observing the success rate of \mathcal{A} .

The detailed proof substantially relies on a fine-grained analysis. Recall that μ^* is the error of the reduction, δ^* is determined in the disguising lemma depending on the amount of mild-lossiness of the reductions, and γ is the sparseness factor as per Definition 1. Let $\theta_{\text{owf}} := (1 - \mu^*) - (\delta^* + 2\gamma)$. For all $\theta_{\text{owf}} = \Omega(\gamma)$, we can show the following: if the success probability of \mathcal{A} is at least $1 - \theta_{\text{owf}}/2$, then the runtime of the aforementioned reduction of deciding Π to inverting F will be $\text{poly}(1/\gamma) \cdot O(T_R + T_A + m^2n)$. Now, if the mildly-lossy reduction R of Π and the adversary \mathcal{A} both run in time $T_R, T_A = \text{poly}(1/\gamma, m, n)$, and Π is worst-case hard for all algorithms than run in $\text{poly}(1/\gamma, m, n)$, then \mathcal{A} cannot succeed with probability more than $1 - \theta_{\text{owf}}/2$. This is because when $T_R, T_A = \text{poly}(1/\gamma, m, n)$, one can use \mathcal{A} as above to decide Π in time $\text{poly}(1/\gamma, m, n)$. Therefore, if Π is worst-case hard for all algorithms than run in $\text{poly}(1/\gamma, m, n)$, the aforementioned reduction should not be able to decide Π in this time, meaning that \mathcal{A} cannot succeed with probability more than $1 - \theta_{\text{owf}}/2$. By setting $\kappa := mn/\gamma$ as the security parameter, one can see that F runs in time $\text{poly}(\kappa)$ but no algorithm \mathcal{A} of runtime $\text{poly}(\kappa)$ can invert it with probability better than $1 - 1/\text{poly}(\kappa)$. This gives a weak one-way function, which can be leveraged to build one-way functions using the standard hardness amplification techniques.

Although we give a detailed analysis of the proposal of [BBD⁺20], we find our construction more sound, since we can extend to one-way state generators without much extra effort. In fact, when the circuits are quantum, there are only two more technical details to fix: (i) showing that the image spaces of two quantum circuits $C_1[y^*]$ and $C_1[\hat{y}]$ have small intersection even in the quantum case, (ii) computing the success rate of \mathcal{A} . The latter uses SWAP test and requires that for any fixed randomness r , the outputs of $C_0(r)$ and $C_1(r)$ be pure.

Hardness vs one-wayness. In what we discussed earlier, the parameter $\gamma > 0$ is not fixed and can be chosen freely subject to the condition $\theta_{\text{owf}} = \Omega(\gamma)$. In fact, to obtain one-way functions, it suffices that γ be roughly bounded by $\text{poly}(1/T_R, 1/n)$, where T_R is the runtime of the reduction. Let Π be a polynomially-hard problem, and let τ_Π be the infimum of the runtime of all solvers of Π .¹ Note that τ_Π is superpolynomial. We set γ such that $1/\gamma = o(\tau_\Pi)$. From the earlier discussion, recall that if Π is $\text{poly}(1/\gamma, m, n)$ -hard and if it admits a mildly-lossy reduction with the same runtime, then the function we built in Equation (3) is one-way. By the choice of γ , Π is indeed $\text{poly}(1/\gamma, m, n)$ -hard. Therefore, if Π admits a mildly-lossy reduction with runtime $\text{poly}(1/\gamma, m, n)$, one-way functions exist.

On the other hand, the non-existence of one-way functions,² implies that Π does not admit a mildly-lossy reduction with runtime $\text{poly}(1/\gamma, m, n)$. Since this argument applies to every $1/\gamma =$

¹In Section 9, τ_Π is defined slightly differently.

²More precisely, infinitely often one-way functions. See Section 9 for more details.

$o(\tau_\Pi)$, one can set $\log 1/\gamma := \log \tau_\Pi / \log \log n$. Therefore, any mildly-lossy reduction for Π must have runtime $2^{\Omega(\log \tau_\Pi / \log \log n)}$.

Mild lossiness of worst-case to average-case reductions. Essentially, a worst-case to average-case reduction from a problem¹ Π to a problem Σ maps *any* instance of Π to an instance of Σ whose distribution is *efficiently-samplable*. In this work we compute the lossiness of worst-case to average-case reductions, and find the specifics of such reductions that can contribute to building one-way functions. Roughly speaking, such reductions are highly mildly lossy, which allows to strengthen the previous general results. We also focus on f -distinguisher reductions. Recall that the common concept of reductions, including non-adaptive Turing and Karp reductions, are captured by the notion of f -distinguisher reductions.

Firstly, we discuss Karp reductions. We define a worst-case to average-case reduction as follows: A reduction R from Π is worst-case to average-case if there exist a small $d < 1$ and a distribution $D = \{D_n\}_{n \in \mathbb{N}}$ over $\{0, 1\}^*$, such that:

$$\forall x \in \Pi \cap \{0, 1\}^n : \Delta(R(x), D) \leq d. \quad (4)$$

This definition can be viewed as a generalization of worst-case to average-case reductions in the sense that (i) the reduction is oblivious to the target average-case problem, and (ii) the reduction maps inputs to a distribution that is *not* necessarily efficiently samplable. The latter does not impose any issues in our setting, since we are only discussing lossiness of the reductions.

Intuitively, worst-case to average-case reductions should lose information about their inputs as the distribution D is independent from the input instance. However, the proof is not direct. Firstly, note that, thanks to our extended disguising lemma, proving the mild-lossiness of these reductions suffices for using them to build OWFs. Next, recall that to prove the mild-lossiness of R , we need to bound the quantity $\sup_X \{I(X; R(X))\}$, for all sparse uniform distributions X over subsets of $\Pi \cap \{0, 1\}^n$ of size $\tilde{O}(1/\gamma^3)$. In order to do so, we first translate the mutual information $I(X; R(X))$ in terms of the KL-divergence, and then use an inverse Pinsker inequality. It is shown by Sason [Sas15], that for every two random variables X and Y , we have

$$D_{KL}(X \| Y) \leq \log \left(1 + \frac{2 \cdot \Delta(X, Y)^2}{\alpha_X} \right),$$

where $\alpha_X = \min_x \Pr(X = x) > 0$. The term $\Delta(X, R(X))$ is bounded by the worst-case to average-case property, therefore, it suffices to bound α_X . Since the mild lossiness concerns uniform distributions X with a support of size $\tilde{O}(1/\gamma^3)$, we can bound α_X by $\tilde{\Omega}(\gamma^3)$. More precisely, we bound the mild-lossiness from above by

$$\max \left\{ 1, 13 + \log \left(\frac{nd^2}{\gamma^3} \right) \right\}.$$

Next, we discuss Turing reductions. Recall that a non-adaptive Turing reduction from Π to Σ , is an algorithm that, given an instance x , outputs oracle queries y_1, \dots, y_k and a circuit C such that $C(y_1, \mathcal{O}(y_1), \dots, y_k, \mathcal{O}(y_k)) = \chi_\Pi(x)$, where \mathcal{O} is an oracle solver for Σ . The common notion of worst-case to average-case *Turing* reduction in the literature is that the marginal distribution

¹Here, “problem” refers to the common concept of problem, search, decision or promise.

of each y_i alone follows a distribution that is independent of x . This is for instance the notion used in the worst-case to average-case reductions for PERMANENT, 3SUM, or OV problems (*e.g.*, see [Lip89, FF93, BRSV17]). However, the joint distribution of (y_1, \dots, y_k) might not be independent of x . We therefore consider a variant of non-adaptive Turing reductions where all queries together (y_1, \dots, y_k) follow a distribution that is roughly independent of x (as in Equation (4)), and C does not leak much information about x either. Any worst-case to average-case Karp reduction is of this type as well as reductions that require *part* of the instance or the random coins in the worst-case to average-case mapping. We show that such reductions are indeed mildly lossy. Thanks to generality of our statement, one can also consider f -distinguisher non-adaptive Turing reductions.

Finally, we consider randomized encodings. Recall that a randomized encoding for a problem Π , or more precisely for χ_Π , is a function E such that $E(x)$ encodes the value of $\chi_\Pi(x)$, therefore, it can be viewed as a reduction for Π . Such an encoding further requires the existence of two efficiently samplable distributions D_{YES} and D_{NO} for respectively simulating the encoding of YES and NO instances of Π within the statistical distance d (that is called privacy). This requirement is in fact similar to Equation (4). Calculating the lossiness follows the same argument as for the worst-case to average-case reductions, however, it only implies the splitting mild-lossiness here.¹ This is, in fact, allowed by the extended disguising lemma. It is worth to mention that one can also consider randomized f -encodings of Π , *i.e.*, the mappings that encode the value of $f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m))$, for any non-constant permutation-invariant choices of $f : \{0, 1\}^m \rightarrow \{0, 1\}$.

1.3 Background and Related Works.

As mentioned earlier, building one-way functions from assumptions like $\text{NP} \neq \text{P}$ has been a challenging problem. Building upon the work of Feigenbaum and Fortnow [FF93], Bogdanov and Trevisan [BT06] show that, in the non-uniform setting – where algorithms can take some advice as input – if there exists a non-adaptive Turing reduction from the worst-case complexity of a decision problem Π to the average-case complexity of another (search or decision) problem, then Π has a polynomial-time reduction to coNP/Poly . As a result, if there exists a non-adaptive Turing reduction from the worst case of an NP -complete problem to inverting a one-way function on uniform inputs, then NP reduces to coNP/Poly in polynomial-time, which is believed to be unlikely. Later, works of Akavia, Goldreich, Goldwasser, and Moshkovitz [AGGM06], and Bogdanov and Brzuska [BB15] extend this impossibility to the uniform settings and adaptive Turing reductions, under the condition that the one-way function is regular² and has an efficiently recognizable range.

Ostrovsky [Ost91] linked the existence of one-way functions to the average-case hardness of the statistical zero-knowledge (SZK) complexity class. Ostrovsky and Wigderson [OW93] showed that if SZK is worst-case hard, then the (seemingly weaker) auxiliary-input one-way functions exist.³

In an effort to base the existence of one-way functions on the worst-case complexity, Applebaum and Raykov [AR16] show that if there exists a worst-case hard language in the complexity class SRE , then one-way functions exist. SRE is the class of problems whose characteristic function admits polynomial-time randomized encodings [IK00, AIK06, App17]. This class is included in SZK,

¹Since the simulation is split between YES and NO instances.

²A one-way function $F : \{0, 1\}^n \rightarrow \{0, 1\}^m$ is called *regular* if for all n and $x, x' \in \{0, 1\}^n$, the number of preimages of x is equal to that of x' .

³This variant of one-way functions requires the existence of a polynomial $p(\cdot)$, such that for every family of polynomial-size circuits $\{\mathcal{A}_n\}_n$, there exists a family of size $p(n)$ circuits $\{C_n\}_n$ that \mathcal{A}_n cannot invert C_n when given the description of C_n as advice.

but it is not known whether the inclusion is proper or not.

OWFs and lossy reductions. More recently, the existence of one-way functions have been linked to *lossy* reductions of worst-case hard problems by Ball et al. [BBD⁺20]. Their techniques stem from previous works of Harnik and Naor [HN06] and Drucker [Dru15]. In [HN06], the compressibility of SAT is leveraged to construct collision-resistant hash functions from one-way functions. More precisely, they show that if SAT admits a strong compression, then collision-resistant hash functions can be built based on one-way functions, in a non-black-box way. A strong compression reduces an instance of SAT with M clauses and N variables to an instance of size $p(N)$ for a polynomial $p(\cdot)$. Later, it was shown that such a compression does not exist unless the polynomial-time hierarchy collapses (*e.g.*, see [FS08, Dru15]). Notably, Drucker [Dru15] shows that if a (decision) problem has a sufficiently compressing polynomial-time OR_m -reduction to any other (decision) problem, then it falls into SZK. Drucker [Dru15] showed that if a problem Π admits an OR_m or AND_m compressing reduction in the sense that it maps m instances of size n to an instance of size $\text{poly}(n)$, then $\Pi \in \text{SZK}$. Later, [BBD⁺20] observed that one can obtain the same result by replacing compression with lossy reductions. Recall that a mapping R is said to be λ -lossy if for all distributions X , it holds that $I(X; R(X)) \leq \lambda$, where I denotes the mutual information. In particular, the work of [BBD⁺20] shows that given a worst-case hard (decision) problem Π , one-way functions exist if there exists (i) an $m/100$ -lossy OR_m reduction from Π to itself, or (ii) an $m/100$ -lossy MAJ_m -reduction from Π to any other problem, or (iii) a perfect $O(m \log n)$ -lossy OR_m -reduction from Π to any other problem. However, they could not instantiate these results based on worst-case hardness of NP. In fact, such results imply that $\Pi \in \text{SZK/Poly}$.

1.4 Open Questions

A candidate for Π is the GAPSVP problem with constant approximation. If Gap-ETH¹ holds, then for every ℓ_p norm, there exists a constant γ_p such that γ_p -GAPSVP cannot be solved in subexponential-time as a function of the lattice dimension [AS18]. Therefore, our results about 3SAT can be adapted to $O(1)$ -GAPSVP (with the extra care about the size of the input versus the dimension of the lattice). As a result, an interesting question to investigate is whether GAPSVP admit mild-lossy f -distinguisher reductions with subexponential runtime. We expect that the structure of lattices might help with answering this question.

Organization of the Paper

After introducing an overview of the context and results of our work in Section 1, we introduce the notation and tools required for our work in Section 2. Then, in Section 3 we introduce the notion of lossy mappings and state and prove our extended disguising lemma. This constitutes the foundation for linking lossiness to one-wayness (and more). In Section 4, we introduce the general notion of f -distinguisher reductions, used to state and analyze our theorems, and prove that this definition contains f -reductions and non-adaptive Turing and Karp reductions. We finally define mildly-lossy problems at the end of the same section. Sections 5, 6, and 7, show that mild-lossiness

¹There exists a constant α such that the following promise variant of 3SAT does not admit a non-uniform subexponential-time algorithm: either the CNF formula is satisfiable or the maximum number of satisfiable clauses is at most αm .

can be leveraged to build zero-knowledge proofs, one-way functions, and one-way state generators, respectively. In Section 8, we study the mild-lossiness of worst-case to average-case (Karp and Turing) reductions as well as randomized encodings. Finally, Section 9 states all of results regarding the relation between the hardness of problems and the existence of one-way functions and one-way state generators.

2 Preliminaries

In this work, we always consider non-uniform algorithms. All classical algorithms are quantum algorithms, therefore, we mostly use the quantum formalism for generalization and simplification. When the distinction is necessary, we explicitly mention it in the beginning of a section or inside an statement, and clearly distinguish between classical and quantum settings.

Notation. We let n denote the security parameter, and all variables are implicitly parametrized by n . We let MS_n denote the set of all mixed states over n qubits and we define $\text{MS}_* := \bigcup_{n=1}^{\infty} \text{MS}_n$. For a positive integer n , we let $[n]$ denote $\{1, 2, \dots, n\}$. The set of all permutations over $[n]$ is \mathfrak{S}_n . We abuse the notation and use the same symbol to refer to the uniform distribution over all permutations of $[n]$. The set of natural numbers $\{1, 2, 3, \dots\}$ is denoted by \mathbb{N} . We denote by \mathbb{R}^+ the set of positive real numbers.

A collection of functions $\{f_i\}_{i \in \mathcal{I}}$ is said to be infinitely often if the index set \mathcal{I} is an increasing infinite sequence of \mathbb{N} .

Uniform and S -Uniform Distributions. For any set S , we let \mathcal{U}_S denote the uniform distribution over S . A distribution is called s -uniform if it is sampled uniformly from a multiset of at most s elements.

Boolean functions. A Boolean function $f : \{0, 1\}^m \rightarrow \{0, 1\}$ is called non-constant if it is not always 0 nor always 1.

Language. A language L is a subset of $\{0, 1\}^*$. The complement of L is defined as $\overline{L} := \{0, 1\}^* \setminus L$.

Promise Problems. A Promise Problem Π consists of two disjoint sets $\Pi_Y, \Pi_N \subset \{0, 1\}^*$, respectively referred to as the set of YES and NO instances. Problem Π asks to decide whether a given instance, which is promised to lie in $\Pi_Y \cup \Pi_N$, belongs Π_Y or Π_N .

Definition 2 (Characteristic Function of a Promise Problem). For a promise problem Π , the characteristic function of Π is the map $\chi_\Pi(x) : \{0, 1\}^* \rightarrow \{0, 1, \star\}$ given by

$$\chi_\Pi(x) = \begin{cases} 1 & \text{if } x \in \Pi_Y \\ 0 & \text{if } x \in \Pi_N \\ \star & \text{otherwise} \end{cases}.$$

Search Problems. We recall the definition of a search problem, inspired by that of [BG94]. We define a *search* problem Π_{search} as a binary relation over $\{0, 1\}^* \times \{0, 1\}^*$. For any $(x, w) \in \Pi_{\text{search}}$, we call x an *instance* and w a *witness*. For any $x \in \{0, 1\}^*$, we define $\Pi_{\text{search}}(x) = \{w \in \{0, 1\}^* \mid (x, w) \in \Pi_{\text{search}}\}$. We refer to the sets $\Pi_{\text{search}|_Y} = \{x \in \{0, 1\}^* \mid \Pi_{\text{search}}(x) \neq \emptyset\}$, and $\Pi_{\text{search}|_N} = \{0, 1\}^* \setminus \Pi_{\text{search}|_Y}$ as the set of YES and NO instances, respectively.

We say that an algorithm \mathcal{A} solves Π_{search} , if for any $x \in \{0, 1\}^*$ for which $\Pi_{\text{search}}(x) \neq \emptyset$, \mathcal{A} returns some $w \in \Pi_{\text{search}}(x)$, and otherwise, outputs \perp .

We denote the decision language defined by Π_{search} as $\Pi = \{x \in \{0, 1\}^* \mid \exists w \in \{0, 1\}^*, (x, w) \in \Pi_{\text{search}}\}$. Each decision language Π can have multiple associated *search problems*, one for every relation Π_{search} that defines Π . Given $x \in \Pi$, the Π_{search} -search problem consists on finding $w \in \Pi_{\text{search}}(x)$.

Games. A two-player, simultaneous-move, zero-sum game is specified by a matrix $\mathbf{M} \in \mathbb{R}^{a \times b}$. Player 1 chooses a row index $i \in [a]$ and Player 2 chooses a column index $j \in [b]$, and Player 2 receives the payoff \mathbf{M}_{ij} from Player 1. The goal of Player 1 is minimizing the expected payoff, while Player 2 opts to maximize it. The row and column indices are called the pure strategies of Player 1 and Player 2, respectively. The mixed strategies are distributions or possible choices of indices. A mixed strategy is s -uniform if it is sampled uniformly from a multiset of at most s pure strategies.

Lemma 1 ([vN28]). *Let \mathcal{P} and \mathcal{Q} be two mixed strategies for Player 1 and 2, respectively. It holds that $\min_{\mathcal{P}} \max_j \mathbb{E}_{i \sim \mathcal{P}}[\mathbf{M}_{ij}] = \max_{\mathcal{Q}} \min_i \mathbb{E}_{j \sim \mathcal{Q}}[\mathbf{M}_{ij}]$.*

The value of the game, which we denote by $\omega(\mathbf{M})$, is the optimal expected value guaranteed by the above lemma. The following lemma shows that each player has nearly-optimal s -uniform strategy when s is chosen to be logarithm of the number of pure strategies of the opponent.

Lemma 2 ([LY94, Theorem 2]). *For any real $\varepsilon > 0$, any $\mathbf{M} \in \mathbb{R}^{a \times b}$, and any integer $s \geq \ln(b)/(2\varepsilon^2)$, it holds that*

$$\min_{\mathcal{P} \in \mathfrak{P}_s} \max_j \mathbb{E}_{i \sim \mathcal{P}}[\mathbf{M}_{ij}] \leq \omega(\mathbf{M}) + \varepsilon(\mathbf{M}_{\max} - \mathbf{M}_{\min}) ,$$

where \mathfrak{P}_s denotes the set of all s -uniform strategies for Player 1. Similar statement holds for Player 2, namely,

$$\max_{\mathcal{Q} \in \mathfrak{Q}_s} \min_i \mathbb{E}_{j \sim \mathcal{Q}}[\mathbf{M}_{ij}] \geq \omega(\mathbf{M}) - \varepsilon(\mathbf{M}_{\max} - \mathbf{M}_{\min}) ,$$

where \mathfrak{Q}_s denotes the set of all s -uniform strategies for Player 2.

Classical information. Given two probability distributions X and Y over Σ , their statistical distance, also called total variation distance, is defined as

$$\Delta(X, Y) := \frac{1}{2} \sum_{x \in \Sigma} |\Pr(X = x) - \Pr(Y = x)| .$$

The Kullback–Leibler divergence or classical relative entropy of X with respect to Y is defined as

$$D_{KL}(X||Y) := \sum_{x \in \Sigma} \Pr(X = x) \log \left(\frac{\Pr(X = x)}{\Pr(Y = x)} \right) .$$

Quantum information. For a mixed state ρ , we let $\|\rho\|_1$ denote its 1-norm. We denote by $\text{Tr}(\rho, \sigma)$ the the trace distance between any two states ρ and σ , with $\text{Tr}(\rho, \sigma) := \|\rho - \sigma\|_1/2$. For an operator Φ ,

we let $\|\Phi\|_{op}$ denote its operator norm. Let $R : \{0, 1\}^n \rightarrow \mathbf{MS}_m$ be any quantum mapping and X a random variable supported over $\{0, 1\}^n$. We let

$$\rho_{X, R(X)} := \sum_{x \in \{0, 1\}^n} \Pr_X(x) |x\rangle\langle x| \otimes R(x) . \quad (5)$$

For a mixed state ρ , we let $S(\rho) := \text{Tr}(\rho \log_2 \rho)$ denote the Von Neumann entropy of ρ . The quantum mutual information of two subsystems A and B is defined as follows. Let ρ_{AB} be their joint state, then

$$I_q(A; B)_\rho := S(\rho_A) + S(\rho_B) - S(\rho_{AB}) ,$$

where $\rho_A = \text{Tr}_B(\rho_{AB})$ and $\rho_B = \text{Tr}_A(\rho_{AB})$. For the sake of simplicity, we sometimes drop the subscripts q and ρ in I_q . When working with quantum systems A, B , the notation $I(A; B)$ implicitly refers to $I_q(A; B)$.

For two quantum states ρ and σ , the quantum relative entropy of ρ with respect to σ is

$$D(\rho\|\sigma) := \begin{cases} \text{Tr}(\rho(\log(\rho) - \log(\sigma))) & \text{if } \text{Supp}(\rho) \subseteq \text{Supp}(\sigma) , \\ \infty & \text{otherwise} . \end{cases}$$

Given a bipartite state ρ_{AB} with marginals ρ_A and ρ_B , the relative entropy can be written in terms of the mutual information as

$$D(\rho_{AB}\|\rho_A \otimes \rho_B) = I_q(A; B)_\rho .$$

For a classical-quantum states $\rho_{XB} := \sum_x p(x) |x\rangle\langle x|_X \otimes \rho_B^x$ and $\sigma_{XB} := \sum_x q(x) |x\rangle\langle x|_X \otimes \sigma_B^x$, the relative entropy takes the simpler form

$$D(\rho_{XB}\|\sigma_{XB}) = \sum_x p(x) D(\rho_B^x\|\sigma_B^x) + D_{KL}(p\|q) , \quad (6)$$

where D_{KL} is the classical Kullback-Leibler divergence.

Lemma 3 ([AE11, Theorem 1]). *Let ρ and σ be two quantum states, let the smallest eigenvalue of σ be uniformly bounded from below, i.e. there exists $\beta > 0$ such that $\lambda_{\min}(\sigma) > \beta$. Then the relative entropy of ρ with respect to σ is bounded by*

$$D(\rho\|\sigma) \leq (\beta + T(\rho, \sigma)) \log \left(1 + \frac{T(\rho, \sigma)}{\beta} \right) .$$

We let $S(\rho\|\sigma) := \text{Tr}(\rho(\log(\rho) - \log(\sigma)))$ denote the relative entropy. We define the quantum conditional entropy of a two-system state ρ_{AB} as follows

$$S(A|B) := S(\rho_{AB}) - S(\rho_B) ,$$

The quantum mutual information in terms of conditional quantum entropy is

$$I(A; B)_\rho = S(\rho_A) - S(A|B)_\rho = S(\rho_B) - S(B|A)_\rho .$$

Lemma 4. *Let ρ_{AB} be a quantum state in two subsystems A and B , with marginal states ρ_A and ρ_B . The following properties hold.*

1. The Von Neumann entropy is additive for tensor product states: $S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)$.
2. Conditioning does not increase entropy: $S(\rho_A) \geq |S(A|B)_\rho|$.
3. The quantum entropy of a system is bounded by the dimension: $S(\rho_A) \leq \dim(H_A)$.

Lemma 5 (Alicki–Fannes–Winter Inequality [Wil13]). Let $\rho_{AB}, \omega_{AB} \in \mathcal{D}(H_A \otimes H_B)$, then

$$|S(A|B)_\rho - S(A|B)_\omega| \leq 2 \operatorname{Tr}(\rho, \sigma) \log \dim(H_A) + h(\operatorname{Tr}(\rho, \sigma)),$$

where $h(p) := -p \log p - (1-p) \log(1-p)$ is the binary entropy function.

The following lemma states that if the outcome of a measurement is close to deterministic, then it must not alter much the state.

Lemma 6 (Gentle Measurement Lemma [Win99]). Let ρ be a mixed state and $\{A, I - A\}$ a two-outcome POVM with $\operatorname{Tr}(A\rho) \geq 1 - \varepsilon$, then $\|\rho - \rho'\|_1 \leq \sqrt{\varepsilon}$, where $\rho' = \frac{\sqrt{A}\rho\sqrt{A}}{\operatorname{Tr}(A\rho)}$.

For two quantum states σ, ρ stored in two different registers A, B , the swap test is executed on the registers A, B and a control register C initialized to $|1\rangle\langle 1|$. It applies Hadamard on C , swaps A and B conditioned on C , and measures B on the Hadamard basis.

Lemma 7 (SWAP Test [BCWd01]). The SWAP test on input (σ, ρ) outputs 1 with probability $(1 + \operatorname{Tr}(\rho\sigma))/2$, in which case we say that it passes the test. For pure states $|\sigma\rangle, |\rho\rangle$, it equals to $(1 + |\langle \rho | \sigma \rangle|^2)/2$.

Given that the trace distance of two pure states $|\sigma\rangle, |\rho\rangle$ can be expressed in terms of the inner product uniquely as $\sqrt{1 - |\langle \rho | \sigma \rangle|^2}$, the SWAP test can also be used to calculate their trace distance.

Definition 3 (ℓ_1 distance for classical distributions and quantum states). We use the notation $\|X - Y\|_1$ to refer to (i) either statistical distance $\Delta(X, Y)$ when variables X, Y are classical distributions (ii) or trace distance $\operatorname{Tr}(X, Y)$ when they are quantum states.

Worst-case hardness. In this work, we consider fine-grained worst-case hardness, as introduced below.

Definition 4. For a function $T : \mathbb{N} \rightarrow \mathbb{R}^+$, a promise problem Π is said to be $T(n)$ -hard, if for any non-uniform classical-advice algorithm \mathcal{A} with runtime at most $T(n)$ over n -bit inputs, and any sufficiently large $n \in \mathbb{N}$, there exists an input $x \in (\Pi_Y \cup \Pi_N) \cap \{0, 1\}^n$ such that $\Pr[\mathcal{A}(x) = \chi_\Pi(x)] < 2/3$.

One can without loss of generality assume that the size of the advice is not larger than the runtime. By setting $\lambda = \log n$, one recovers the regular definition of worst-case hardness.

Complexity class QSZK. We recall the quantum state distinguishability problem below. We refer to [Wat02] for more details.

Definition 5 (Quantum State Distinguishability). Let $\alpha, \beta \in [0, 1]$ such that $\alpha < \beta$. Given two quantum circuits \mathcal{C}_0 and \mathcal{C}_1 , let ρ_0 and ρ_1 be the (mixed) quantum states that they produce by running on all-zero states with the promise that either $\|\rho_0 - \rho_1\|_1 \geq \beta$ (corresponds to no instances) or $\|\rho_0 - \rho_1\|_1 \leq \alpha$ (corresponds to yes instances). The QSD $_{\alpha, \beta}$ problem is to decide which one is the case.

The above problem enjoys a polarization property. The lemma below is adapted from [Wat02, SV].

Lemma 8. *Let n be a positive integer. Let $\alpha, \beta : \mathbb{N} \rightarrow [0, 1]$, and $\theta : \mathbb{R} \rightarrow (1, +\infty)$ be functions of n such that $\theta := \beta^2/\alpha$. There exists a deterministic classical algorithm **Polarize** that given a pair of (quantum) circuits (C_0, C_1) as well as a unary parameter 1^n , outputs a pair of (quantum) circuits (P_0, P_1) such that*

$$\begin{aligned} \|C_0 |0\rangle - C_1 |0\rangle\|_1 \leq \alpha &\Rightarrow \|P_0 |0\rangle - P_1 |0\rangle\|_1 \leq 2^{-n}, \\ \|C_0 |0\rangle - C_1 |0\rangle\|_1 \geq \beta &\Rightarrow \|P_0 |0\rangle - P_1 |0\rangle\|_1 \geq 1 - 2^{-n}. \end{aligned}$$

Moreover, the runtime and output size of **Polarize** are of $O(n \log(8n)(|C_0| + |C_1|)/\log(\theta))$ when $n \rightarrow +\infty$.

There are various equivalent definitions of the complexity class QSZK. The following definition suffices for our purposes.

Definition 6 (QSZK). *The class QSZK is consisted of all promise problems that have many-to-one polynomial-time reductions to $\text{QSD}_{1/4, 3/4}$.*

All definitions and lemmas above can be restricted to classical algorithms. In this case, we let SZK denote the corresponding classical complexity class and SD denote the statistical difference problem (classical variant of QSD).

Cryptographic primitives. One-way functions are defined as follows:

Definition 7 (Non-Uniform One-Way Functions). *Let $T : \mathbb{N} \rightarrow \mathbb{R}^+$ and $\theta : \mathbb{N} \rightarrow [0, 1]$. A family of non-uniform PPT algorithms $F := \{F_n\}_{n \in \mathbb{N}}$ is said to be a (T, θ) -one-way function (OWF) if for all sufficiently large n and any $T(n)$ -time algorithm \mathcal{A} , it holds that*

$$\Pr_{x \sim \mathcal{U}_{\{0,1\}^n}} [F(\mathcal{A}(F(x))) = F(x)] \leq \theta(n).$$

Furthermore, we say that F is a θ -OWF for an algorithm \mathcal{A} if the above inequality holds without imposing any bound on the runtime of \mathcal{A} . If the above equation only holds for all n in an infinite subset of the natural numbers, i.e. $S \subseteq \mathbb{N}$, then we say that F is an infinitely-often OWF.

When $T = \text{poly}(n)$ and $\theta = \text{negl}(n)$, the above definition corresponds to the common definition of one-way functions. If θ is $1 - 1/n^c$ for some constant c , this corresponds to weak one-way functions. It is shown by [Yao82] that weak one-way functions imply one-way functions.

Below, we define efficiently samplable statistically far but computationally indistinguishable quantum states (EFI).

Definition 8 (Non-Uniform EFI). *Let $T : \mathbb{N} \rightarrow \mathbb{R}^+$ and $d, D : \mathbb{N} \rightarrow [0, 1]$ be functions. A non-uniform (T, D, d) -EFI scheme is a QPT algorithm $\text{EFI}_h(1^n, b)$ that is given a classical $\text{poly}(n)$ -size advice h and a bit b , outputs a quantum state ρ_b , such that for any sufficiently large $n \in \mathbb{N}$ has the following specifications:*

1. **Computational indistinguishability.** For all non-uniform (possibly quantum) $T(n)$ -time algorithms \mathcal{A} :

$$\left| \Pr[\mathcal{A}(\rho_0) = 1] - \Pr[\mathcal{A}(\rho_1) = 1] \right| \leq d(n).$$

2. **Statistical Distance.** $\|\rho_0 - \rho_1\|_1 \geq D(n)$.

Furthermore, we say that EFI is a (D, d) -EFI for an algorithm \mathcal{A} , if the computational indistinguishability holds for \mathcal{A} without requiring any bound of the runtime of \mathcal{A} .

Remark 1. When restricted to classical algorithms, EFI pairs with $D - d \geq 1/\text{poly}(n)$ and $T = \text{poly}(n)$ imply the existence of one-way functions (e.g., see [Gol90, NR06, BDRV19]). The state of the art for the quantum EFI pairs is more restricted. More precisely, an EFI pair with mixed states and $D^2 - \sqrt{d} \geq O(1)$ implies quantum bit commitment (see [BQSY24, Corollary 8.8] for EFI polarization and [BCQ23] for the generic transformation to construct quantum bit commitments from EFI pairs).

In this work, we consider the inefficient-verifier one-way state generators.

Definition 9 (Non-Uniform One-Way State Generators). Let $T : \mathbb{N} \rightarrow \mathbb{R}^+$ and $\theta : \mathbb{N} \rightarrow [0, 1]$. A (T, θ) -one-way state generator (OWSG) is a tuple of algorithms $\mathbf{G} := (\text{KeyGen}, \text{StateGen}, \text{Ver})$ with the following specification:

- $\text{KeyGen}_h(1^n) \rightarrow k$: is a QPT algorithm that given the security parameter 1^n and a $\text{poly}(n)$ -size classical advice h , outputs a classical string $k \in \{0, 1\}^n$;
- $\text{StateGen}(k) \rightarrow \rho_k$: is a QPT algorithm that given a classical string k , outputs an m -qubit quantum state;
- $\text{Ver}(k, \rho) \in \{0, 1\}$: is a (possibly unbounded) algorithm that given a classical string k and a quantum state ρ outputs either 0 or 1.

Further, they satisfy the following properties:

1. **Correctness.** Outputs of the samplers $(\text{KeyGen}, \text{StateGen})$ pass the verification with overwhelming probability, i.e.,

$$\Pr_{\substack{k \leftarrow \text{KeyGen}_h \\ \rho_k \leftarrow \text{StateGen}(k)}} [\text{Ver}(k, \rho_k) = 1] \geq 1 - \text{negl}(n).$$

2. **Security.** For every non-uniform $T(n)$ -time adversary \mathcal{A} , and any polynomial $t(n)$

$$\Pr_{\substack{k \leftarrow \text{KeyGen}_h \\ \rho_k \leftarrow \text{StateGen}(k) \\ k' \leftarrow \mathcal{A}(\rho_k^{\otimes t}; h)}} [\text{Ver}(k', \rho_k) = 1] \leq \theta(n).$$

Furthermore, we say that \mathbf{G} is a θ -OWSG for an algorithm \mathcal{A} if the inequality concerning security (Property 2) holds for \mathcal{A} without requiring any bound on the runtime of \mathcal{A} . If the above inequality only holds for all n in an infinite subset of the natural numbers, i.e. $S \subseteq \mathbb{N}$, then we say that \mathbf{F} is an infinitely-often OWSG.

A weak OWSG can be recovered by the above definition for $T = \text{poly}(n)$ and $\theta = 1 - 1/n^c$ for some constant c . It is shown in [MY24] that weak OWSGs imply OWSGs.

Fine-Grained primitives. In fine-grained one-way functions, there is at most a polynomial gap between the runtime of the function and runtime of the adversary.

Definition 10 (Fine-grained OWF). Let $\eta > 0$ be a real number and $\theta : \mathbb{N} \rightarrow [0, 1]$. A family of non-uniform algorithms $F := \{F_n\}_{n \in \mathbb{N}}$ is said to be a (η, θ) -fine-grained one-way function (FGOWF) if for any $O(T_F^{1+\eta})$ -time algorithm \mathcal{A} , for all sufficiently large n , it holds that

$$\Pr_{x \sim \mathcal{U}_{\{0,1\}^n}} [F(\mathcal{A}(F(x))) = F(x)] \leq \theta(n),$$

where T_F is the runtime of F . If θ is constant, we simply say that F is a weak η -FGOWF.

3 Lossy Mappings and Disguising Lemma

[Dru15] derives a quantitative approach (called disguising distribution lemma) to measure how much information can be recovered from the output of a compressing mapping about its input, based on the compression size; a distinguishing variant of Fano's inequality. Such mappings are indeed a special type of lossy mappings, an observation upon which Ball et al. [BBD⁺20] develop their work.

In this section, we focus on variants of lossy mappings and their properties, and extend the disguising lemma. In our analysis, we consider both randomized functions and quantum mappings. All the statements hold with respect to both cases. For simplicity and generality, we only refer to quantum mappings. We explicitly highlight the distinction when the analysis requires to distinguish between the two cases.

Classically, a randomized function $R : \{0, 1\}^* \rightarrow \{0, 1\}^*$ is said to be ℓ -lossy for a class of distributions $X = \{X_n\}_{n \in \mathbb{N}}$ if $I(X_n; R(X_n)) \leq \ell(n)$. Below, we also consider general mappings with classical input and quantum output.

Definition 11 (Lossy Mapping). Let $\ell : \mathbb{N} \rightarrow \mathbb{R}^+$. Let $R : \{0, 1\}^* \rightarrow S$ be a mapping, where $S = \{0, 1\}^*$ (classical mapping) or $S = \text{MS}_*$ (quantum mapping). We say that R is ℓ -lossy for a class of distributions $X = \{X_n\}_{n \in \mathbb{N}}$ over $\{0, 1\}^*$, if it holds that

$$I(X_n; R(X_n)) \leq \ell(n).$$

For the sake of simplicity, we say that R is ℓ -lossy, if it is ℓ -lossy for all distributions.

The results by [Dru15, BBD⁺20] rely on the lossiness of the mapping for all distributions. Such a condition seems quite strong, in particular, for the multi-variate mappings over m -tuple input. We simplify this condition in two different directions. First, we consider lossy mappings over a particular class of distributions as follows:

Definition 12 (Splitting Lossy Mapping). Let $\ell : \mathbb{N} \rightarrow \mathbb{R}^+$, $m \in \mathbb{N}$ and $S_0, S_1 \subseteq \{0, 1\}^*$ be two disjoint sets. A mapping R is splitting ℓ -lossy supported on (S_0, S_1) if it is ℓ -lossy for the class of distributions $X = (X_1, \dots, X_m)$ such that for each $i \in [m]$, either $\text{Supp}(X_i) \subseteq S_0$ or $\text{Supp}(X_i) \subseteq S_1$. In other words, $\{X_1, X_2, \dots, X_m\}$ splits into S_0 -supported and S_1 -supported distributions.

Remark 2. A lossy mapping as per Definition 11 is also a splitting lossy mapping.

Later, for the lossy reductions of a problem Π , we choose S_0 and S_1 as the sets Π_N and Π_Y . Splitting the distribution in such a way allows us to precisely calculate the lossiness of randomized encodings.

In the rest of this section, we discuss the generalization of disguising distribution lemma in [Dru15] and its improvement by [BBD⁺20]. In both of these results, the lossiness (compression in the former and lossiness in the latter) is considered as in Definition 11 with respect to all possible input distributions. Instead, we adapt it for splitting lossy maps where the input distribution is uniform over a sparse set. This is obtained by a more refined analysis but yet very similar to those of [Dru15, BBD⁺20]. Below, we have the main lemma of this section.

Lemma 9 (Extended Disguising Lemma). *Let n, m, m_0, m_1 be positive integers such that $m = m_0 + m_1 + 1$, and $R : \{0, 1\}^* \rightarrow \text{MS}_*$ be any quantum mapping. Further, let $S_0, S_1 \subseteq \{0, 1\}^n$ be two disjoint sets, d be a positive integer, $\varepsilon > 0$ be real, and $s := \lceil n \ln 2 / (2\varepsilon^2) \rceil$.*

For any choice of positive real ℓ , if R is splitting ℓ -lossy for all ds -uniform distributions supported on (S_0, S_1) , then there exist two collections K_1, \dots, K_s and T_1, \dots, T_s of multisets of d elements respectively contained in S_0 and S_1 , such that

- *for any $y \in S_0$, it holds that*

$$\mathbb{E}_{a \sim \mathcal{U}_{[s]}, \pi \sim \mathfrak{S}_m} \left[\left\| R \left(\pi \left(\mathcal{U}_{K_a}^{\otimes m_0}, y, \mathcal{U}_{T_a}^{\otimes m_1} \right) \right) - R \left(\pi \left(\mathcal{U}_{K_a}^{\otimes (m_0+1)}, \mathcal{U}_{T_a}^{\otimes m_1} \right) \right) \right\|_1 \right] \leq \delta + \frac{2(m+1)}{d+1} + 2\varepsilon ;$$

- *and for any $y \in S_1$, it holds that*

$$\mathbb{E}_{a \sim \mathcal{U}_{[s]}, \pi \sim \mathfrak{S}_m} \left[\left\| R \left(\pi \left(\mathcal{U}_{K_a}^{\otimes m_0}, y, \mathcal{U}_{T_a}^{\otimes m_1} \right) \right) - R \left(\pi \left(\mathcal{U}_{K_a}^{\otimes m_0}, \mathcal{U}_{T_a}^{\otimes (m_1+1)} \right) \right) \right\|_1 \right] \leq \delta + \frac{2(m+1)}{d+1} + 2\varepsilon ,$$

where

$$\delta := \min \left\{ \sqrt{\frac{\ell \ln 2}{2m}}, 1 - 2^{-\frac{\ell}{m}-2} \right\}.$$

Note that the states inside the trace distance are mixed states since the inputs of R are randomized classical distributions.

The proof requires some background definitions and lemmas. Similar to [Dru15, BBD⁺20], we define distributional stability as follows.

Definition 13. *Let n, m, m_0, m_1 be positive integers such that $m = m_0 + m_1 + 1$. For a real $\delta \in [0, 1]$, a quantum mapping $R : \{0, 1\}^{mn} \rightarrow \text{MS}_*$ is said to be δ -quantumly-distributionally stable (δ -QDS) with respect to two distributions $(\mathcal{D}_0, \mathcal{D}_1)$ over $\{0, 1\}^n$ if the following holds:*

$$\mathbb{E}_{y \sim \mathcal{D}_0, \pi \sim \mathfrak{S}_m} \left[\left\| R \left(\pi \left(\mathcal{D}_0^{\otimes m_0}, y, \mathcal{D}_1^{\otimes m_1} \right) \right) - R \left(\pi \left(\mathcal{D}_0^{\otimes (m_0+1)}, \mathcal{D}_1^{\otimes m_1} \right) \right) \right\|_1 \right] \leq \delta .$$

Note that the order of the pair $(\mathcal{D}_0, \mathcal{D}_1)$ matters. Furthermore, when $m_1 = 0$, we simply say that the mapping is δ -QDS with respect to \mathcal{D}_0 .

Below, we recall an adaptation of [Dru15, Lemma 8.10].

Lemma 10. *Assume that $R : \{0, 1\}^{m \cdot n} \rightarrow \text{MS}_*$ satisfies the properties in Lemma 9 for $m_1 = 0$. Then R is δ -QDS with respect to any ds -uniform distribution \mathcal{D}_0 supported on either S_0 or S_1 .*

In the original lemma from [Dru15], compression is used to bound the entropy of the mutual information. However, note that this can be argued directly from splitting lossiness, and that any restriction on the input distributions will give a result for the same restricted case.

The following lemma is the generalization of the above one.

Lemma 11. *Assume that $R : \{0, 1\}^{mn} \rightarrow \text{MS}_*$ satisfies the properties in Lemma 9. Then R is δ -QDS with respect to any ds -uniform independent distributions $(\mathcal{D}_0, \mathcal{D}_1)$ each supported on either S_0 or S_1 .*

Proof. The proof is similar to that of [BBD⁺20, Proposition B.1]. Let $\pi \in \mathfrak{S}_m$ be a fixed permutation. One can rewrite it as the composition of two partial permutations π_0 and π_1 , i.e., $\pi = \pi_0 \circ \pi_1$, such that π_1 only acts on the last m_1 arguments of the input. Let $\rho_\pi(y)$ be as follows

$$\rho_\pi(y) := R(\pi(\mathcal{D}_0^{\otimes m_0}, y, \mathcal{D}_1^{\otimes m_1})) .$$

For $y, y' \sim \mathcal{D}_0$, two independent random variables, and $\pi \sim \mathfrak{S}_m$, we want to prove that

$$\mathbb{E}_{y, \pi} \left[\left\| \rho_\pi(y) - \rho_\pi(y') \right\|_1 \right] \leq \delta .$$

Note that it is enough to bound the conditional distributions since

$$\mathbb{E}_{y, \pi} \left[\left\| \rho_\pi(y) - \rho_\pi(y') \right\|_1 \right] = \mathbb{E}_\pi \left[\mathbb{E}_{y, \pi | \pi_1} \left[\left\| \rho_\pi(y) - \rho_\pi(y') \right\|_1 \right] \right] ,$$

by the law of total probability.

Let $R'(x_1, x_2, \dots, x_{m_0+1})$ be the mapping that first samples π then evaluates $R(\pi_1(x_1, x_2, \dots, x_{m_0+1}, \mathcal{D}_1^{\otimes m_1}))$. For any fixed π_1 , we show that R' is splitting ℓ -lossy for all ds -uniform distributions over either S_0 or S_1 . Indeed, let $(\mathcal{X}_1, \dots, \mathcal{X}_{m_0+1})$ be independent ds -uniform random variables with $\text{Supp}(\mathcal{X}_i) \subseteq S_0$ or $\text{Supp}(\mathcal{X}_i) \subseteq S_1$ for each $i \in [m_0+1]$, and $(\mathcal{Z}_1, \dots, \mathcal{Z}_{m_1}) \sim \mathcal{D}_1^{\otimes m_1}$, thus $\text{Supp}(\mathcal{Z}_i) \subseteq \text{Supp}(\mathcal{D}_1) \subseteq S_j$ for all $i \in [m_1]$ and some $j \in \{0, 1\}$. By the splitting lossiness of R for any ds -uniform distribution, we can bound the loss of R' :

$$\begin{aligned} \ell &\geq I_q(\pi_1(\mathcal{X}_1, \dots, \mathcal{X}_{m_0+1}, \mathcal{Z}_1, \dots, \mathcal{Z}_{m_1}); R(\pi_1(\mathcal{X}_1, \dots, \mathcal{X}_{m_0+1}, \mathcal{Z}_1, \dots, \mathcal{Z}_{m_1}))) \\ &= I_q(\mathcal{X}_1, \dots, \mathcal{X}_{m_0+1}, \mathcal{Z}_1, \dots, \mathcal{Z}_{m_1}; R(\pi_1(\mathcal{X}_1, \dots, \mathcal{X}_{m_0+1}, \mathcal{Z}_1, \dots, \mathcal{Z}_{m_1}))) \\ &\geq I_q(\mathcal{X}_1, \dots, \mathcal{X}_{m_0+1}; R(\pi_1(\mathcal{X}_1, \dots, \mathcal{X}_{m_0+1}, \mathcal{Z}_1, \dots, \mathcal{Z}_{m_1}))). \end{aligned}$$

Finally, by Lemma 10 a splitting lossy map must also be δ -QSD, thus

$$\begin{aligned} \mathbb{E}_{y, \pi | \pi_1} \left[\left\| \rho_\pi(y) - \rho_\pi(y') \right\|_1 \right] &= \mathbb{E}_{y, \pi | \pi_1} \left[\left\| R(\pi(\mathcal{D}_0^{\otimes m_0}, y, \mathcal{D}_1^{\otimes m_1})) - R(\pi(\mathcal{D}_0^{\otimes m_0}, y', \mathcal{D}_1^{\otimes m_1})) \right\|_1 \right] \\ &= \mathbb{E}_{y, \pi_0} \left[\left\| R'(\mathcal{D}_0^{\otimes m_0}, y) - R'(\mathcal{D}_0^{\otimes m_0}, y') \right\|_1 \right] \\ &\leq \delta . \end{aligned}$$

□

If a mapping is distributionally stable with respect to a pair of distributions, then one can “sparsify” the distributions while nearly keeping the stability.

Lemma 12. *Let $n, m, m_0, m_1, \ell, S_0, S_1, R$ and δ be as in Lemma 9. Let \mathcal{D}_0 and \mathcal{D}_1 be two independent distributions with supports over S_0 and S_1 , respectively. Let $\{x_i^{(0)}\}_{i \in [d+1]}$ and $\{x_i^{(1)}\}_{i \in [d+1]}$ be independent samples from \mathcal{D}_0 and \mathcal{D}_1 , respectively. For each $j \in \{0, 1\}$, let $y_j^* := x_{i^*}^{(j)}$ be uniformly chosen from $\{x_i^{(j)}\}_{i \in [d+1]}$ and let $\hat{\mathcal{D}}_j$ be the uniform distribution over the multiset $\{x_i^{(j)}\}_{i \in [d+1] \setminus \{i^*\}}$. Then it holds that*

$$\begin{aligned} \mathbb{E}_{\pi \sim \mathfrak{S}_m} \left[\left\| R \left(\pi \left(\hat{\mathcal{D}}_0^{\otimes m_0}, y_0^*, \hat{\mathcal{D}}_1^{\otimes m_1} \right) \right) - R \left(\pi \left(\hat{\mathcal{D}}_0^{\otimes(m_0+1)}, \hat{\mathcal{D}}_1^{\otimes m_1} \right) \right) \right\|_1 \right] &\leq \delta + \frac{2m_0 + 1}{d + 1}, \\ \mathbb{E}_{\pi \sim \mathfrak{S}_m} \left[\left\| R \left(\pi \left(\hat{\mathcal{D}}_0^{\otimes m_0}, y_1^*, \hat{\mathcal{D}}_1^{\otimes m_1} \right) \right) - R \left(\pi \left(\hat{\mathcal{D}}_0^{\otimes m_0}, \hat{\mathcal{D}}_1^{\otimes(m_1+1)} \right) \right) \right\|_1 \right] &\leq \delta + \frac{2m_1 + 1}{d + 1}. \end{aligned}$$

Proof. We prove the first statement. The other one is implied similarly. Let $\tilde{\mathcal{D}}_0$ denote the uniform distribution over $\{x_i^{(0)}\}_{i \in [d+1]}$. For any fixed set of multisets as above and any choice of permutation π and quantum mapping R , we have

$$\begin{aligned} &\left\| R \left(\pi \left(\tilde{\mathcal{D}}_0^{\otimes(m_0+1)}, \hat{\mathcal{D}}_1^{\otimes m_1} \right) \right) - R \left(\pi \left(\hat{\mathcal{D}}_0^{\otimes(m_0+1)}, \hat{\mathcal{D}}_1^{\otimes m_1} \right) \right) \right\|_1 \\ &\leq \left\| \tilde{\mathcal{D}}_0^{\otimes(m_0+1)} \otimes \hat{\mathcal{D}}_1^{\otimes m_1} - \hat{\mathcal{D}}_0^{\otimes(m_0+1)} \otimes \hat{\mathcal{D}}_1^{\otimes m_1} \right\|_1 \\ &\leq \left\| \tilde{\mathcal{D}}_0^{\otimes(m_0+1)} - \hat{\mathcal{D}}_0^{\otimes(m_0+1)} \right\|_1 \\ &\leq (m_0 + 1) \left\| \tilde{\mathcal{D}}_0 - \hat{\mathcal{D}}_0 \right\|_1, \end{aligned}$$

where we used the quantum data processing inequality for the first two upper bounds, and the property of tensor product for the last one. Since both $\tilde{\mathcal{D}}_0$ and $\hat{\mathcal{D}}_0$ are classical, their trace distance coincides with their statistical distance. Therefore, we have

$$\begin{aligned} \left\| \tilde{\mathcal{D}}_0 - \hat{\mathcal{D}}_0 \right\|_1 &= \frac{1}{2} \sum_{x \in \{x_i^{(0)}\}_{i \in [d+1]}} \left| \Pr_{\tilde{\mathcal{D}}_0}(x) - \Pr_{\hat{\mathcal{D}}_0}(x) \right| \\ &= \frac{1}{2(d+1)} + \frac{1}{2} \sum_{x \in \{x_i^{(0)}\}_{i \in [d+1] \setminus \{i^*\}}} \left| \frac{1}{d+1} - \frac{1}{d} \right| \\ &= \frac{1}{d+1}. \end{aligned}$$

Similarly, it holds that

$$\left\| R \left(\pi \left(\tilde{\mathcal{D}}_0^{\otimes m_0}, y_0^*, \hat{\mathcal{D}}_1^{\otimes m_1} \right) \right) - R \left(\pi \left(\hat{\mathcal{D}}_0^{\otimes m_0}, y_0^*, \hat{\mathcal{D}}_1^{\otimes m_1} \right) \right) \right\|_1 \leq \frac{m_0}{d+1}.$$

From the triangle inequality, it follows that

$$\begin{aligned}
& \left\| R\left(\pi\left(\widehat{\mathcal{D}}_0^{\otimes m_0}, y_0^*, \widehat{\mathcal{D}}_1^{\otimes m_1}\right)\right) - R\left(\pi\left(\widehat{\mathcal{D}}_0^{\otimes(m_0+1)}, \widehat{\mathcal{D}}_1^{\otimes m_1}\right)\right) \right\|_1 \\
& \leq \left\| R\left(\pi\left(\widehat{\mathcal{D}}_0^{\otimes m_0}, y_0^*, \widehat{\mathcal{D}}_1^{\otimes m_1}\right)\right) - R\left(\pi\left(\widetilde{\mathcal{D}}_0^{\otimes m_0}, y_0^*, \widehat{\mathcal{D}}_1^{\otimes m_1}\right)\right) \right\|_1 \\
& \quad + \left\| R\left(\pi\left(\widetilde{\mathcal{D}}_0^{\otimes m_0}, y_0^*, \widehat{\mathcal{D}}_1^{\otimes m_1}\right)\right) - R\left(\pi\left(\widetilde{\mathcal{D}}_0^{\otimes(m_0+1)}, \widehat{\mathcal{D}}_1^{\otimes m_1}\right)\right) \right\|_1 \\
& \quad + \left\| R\left(\pi\left(\widetilde{\mathcal{D}}_0^{\otimes(m_0+1)}, \widehat{\mathcal{D}}_1^{\otimes m_1}\right)\right) - R\left(\pi\left(\widehat{\mathcal{D}}_0^{\otimes(m_0+1)}, \widehat{\mathcal{D}}_1^{\otimes m_1}\right)\right) \right\|_1 \\
& < \left\| R\left(\pi\left(\widetilde{\mathcal{D}}_0^{\otimes m_0}, y_0^*, \widehat{\mathcal{D}}_1^{\otimes m_1}\right)\right) - R\left(\pi\left(\widetilde{\mathcal{D}}_0^{\otimes(m_0+1)}, \widehat{\mathcal{D}}_1^{\otimes m_1}\right)\right) \right\|_1 + \frac{2m_0 + 1}{d + 1}.
\end{aligned}$$

Recall that R is splitting ℓ -lossy with respect to all ds -uniform distributions supported on (S_0, S_1) . Therefore, by Lemma 11 it is δ -QSD with respect to all ds -uniform pair of distributions each supported on either S_0 or S_1 , including $(\widetilde{\mathcal{D}}_0, \widehat{\mathcal{D}}_1)$. Finally, by taking expectation from both sides above with respect to π , and using the fact that R is δ -QSD with respect to $(\widetilde{\mathcal{D}}_0, \widehat{\mathcal{D}}_1)$, one obtains the claimed upper bound. \square

Proof of Lemma 9. Consider the following two-player, simultaneous-move, zero-sum game:

- Player 1: chooses a pair of multisets $K \subseteq S_0$ and $T \subseteq S_1$, each of size d .
- Player 2: chooses an element $y \in S_0 \cup S_1$
- Payoff: if $y \in S_0$, Player 2 gains

$$\mathbb{E}_{\pi \sim \mathfrak{S}_m} \left[\left\| R\left(\pi\left(\mathcal{U}_K^{\otimes m_0}, y, \mathcal{U}_T^{\otimes m_1}\right)\right) - R\left(\pi\left(\mathcal{U}_K^{\otimes(m_0+1)}, \mathcal{U}_T^{\otimes m_1}\right)\right) \right\|_1 \right],$$

otherwise, Player 2 gains

$$\mathbb{E}_{\pi \sim \mathfrak{S}_m} \left[\left\| R\left(\pi\left(\mathcal{U}_K^{\otimes m_0}, y, \mathcal{U}_T^{\otimes m_1}\right)\right) - R\left(\pi\left(\mathcal{U}_K^{\otimes m_0}, \mathcal{U}_T^{\otimes(m_1+1)}\right)\right) \right\|_1 \right].$$

Consider a ds -uniform strategy for Player 2, i.e. a distribution \mathcal{Y} of y that is uniform over a multiset of pure strategies of size ds . We explain a strategy $(\mathcal{K}, \mathcal{T})$ for Player 1 that bounds the expected payoff. Player 1 chooses K by sampling d independent instances of the restriction of \mathcal{Y} to S_0 , and chooses T by sampling d independent instances of the restriction of \mathcal{Y} to S_1 . The expected payoff is

$$\begin{aligned}
E := & \Pr_{y \sim \mathcal{Y}}(y \in S_0) \mathbb{E}_{\pi, K, T} \left[\left\| R\left(\pi\left(\mathcal{U}_K^{\otimes m_0}, y, \mathcal{U}_T^{\otimes m_1}\right)\right) - R\left(\pi\left(\mathcal{U}_K^{\otimes(m_0+1)}, \mathcal{U}_T^{\otimes m_1}\right)\right) \right\|_1 \mid y \in S_0 \right] \\
& + \Pr_{y \sim \mathcal{Y}}(y \in S_1) \mathbb{E}_{\pi, K, T} \left[\left\| R\left(\pi\left(\mathcal{U}_K^{\otimes m_0}, y, \mathcal{U}_T^{\otimes m_1}\right)\right) - R\left(\pi\left(\mathcal{U}_K^{\otimes m_0}, \mathcal{U}_T^{\otimes(m_1+1)}\right)\right) \right\|_1 \mid y \in S_1 \right].
\end{aligned}$$

Let $x_1^{(0)}, x_2^{(0)}, \dots, x_{d+1}^{(0)}$ and $x_1^{(1)}, x_2^{(1)}, \dots, x_{d+1}^{(1)}$ be $d+1$ independent samples from $\mathcal{Y}|_{S_0}$ and $\mathcal{Y}|_{S_1}$, respectively. Sample $i^* \xleftarrow{\$} [d+1]$ and for $j \in \{0, 1\}$, let $y_j^* := x_{i^*}^{(j)}$. Let $\widehat{\mathcal{Y}}_0$ and $\widehat{\mathcal{Y}}_1$ be the uniform distributions over the multisets $\{x_i^{(0)}\}_{i \in [d+1] \setminus \{i^*\}}$ and $\{x_i^{(1)}\}_{i \in [d+1] \setminus \{i^*\}}$, respectively. For $j \in \{0, 1\}$, we have that $(y_j^*, \widehat{\mathcal{Y}}_0, \widehat{\mathcal{Y}}_1) \sim (\mathcal{Y}|_{S_j}, \mathcal{K}, \mathcal{T})$. Then, by Lemma 12, we have

$$\mathbb{E}_{\pi} \left[\left\| R\left(\pi\left(\widehat{\mathcal{Y}}_0^{\otimes m_0}, y, \widehat{\mathcal{Y}}_1^{\otimes m_1}\right)\right) - R\left(\pi\left(\widehat{\mathcal{Y}}_0^{\otimes(m_0+1)}, \widehat{\mathcal{Y}}_1^{\otimes m_1}\right)\right) \right\|_1 \mid y \in S_0 \right] \leq \delta + \frac{2m_0 + 1}{d + 1},$$

and

$$\mathbb{E}_\pi \left[\left\| R \left(\pi \left(\hat{\mathcal{Y}}_0^{\otimes m_0}, y, \hat{\mathcal{Y}}_1^{\otimes m_1} \right) \right) - R \left(\pi \left(\hat{\mathcal{Y}}_0^{\otimes m_0}, \hat{\mathcal{Y}}_1^{\otimes (m_1+1)} \right) \right) \right\|_1 \mid y \in S_1 \right] \leq \delta + \frac{2m_1 + 1}{d + 1}.$$

Therefore, we obtain $E \leq \delta + 2(m + 1)/(d + 1)$.

Above, we showed that for every ds -uniform strategy for Player 2, there exists a strategy for Player 1 that bounds the expected payoff by $\delta + 2(m + 1)/(d + 1)$. Let $\mathbf{M} := [\mathbf{M}_{ij}]_{i,j}$ be the matrix such that \mathbf{M}_{ij} corresponds to the payoff when Player 1 outputs i and Player 2 outputs j . By Lemma 2, we have

$$\delta + 2(m + 1)/(d + 1) \geq \max_{Q \in \Omega_{ds}} \min_i \mathbb{E}_{j \sim Q} [\mathbf{M}_{ij}] \geq \omega(\mathbf{M}) - \varepsilon(\mathbf{M}_{\max} - \mathbf{M}_{\min}) \geq \omega(\mathbf{M}) - \varepsilon,$$

where Ω_{ds} is the set of all ds -uniform strategies for Player 2. It follows that $\omega(\mathbf{M}) \leq \delta + 2(m + 1)/(d + 1) + \varepsilon$.

Now we use Lemma 2 in other way around. In fact, the number of possible choices for Player 1 is $|S_0 \cup S_1| \leq 2^n$. Therefore, Lemma 2 asserts that there exists a s -uniform strategy for Player 2 such that for any possibly mixed strategy for Player 1, the expected payoff is at most ε -far from the value of the game $\omega(\mathbf{M})$. In other words, for this particular strategy of Player 1, the expected payoff is always at most

$$\omega(\mathbf{M}) + \varepsilon \leq \delta + 2(m + 1)/(d + 1) + 2\varepsilon.$$

Recall that a s -uniform strategy is, by definition, a uniformly sampled element from a size- s multiset of choices of the player. Note that Player 1 chooses a pair (K, T) . Therefore, this strategy is essentially a uniform distribution over some multiset $\{(K_1, T_1), \dots, (K_s, T_s)\}$, which concludes the proof. \square

4 Mildly-Lossy Problems

In this section, we first put forward a new abstraction, called f -distinguisher reduction, that is suitable for our analysis and implies definitions of f -reductions (adapted from Drucker [Dru15]) as well as Karp and non-adaptive Turing reductions. Then, by considering the lossiness property (as defined in Section 3), we introduce mildly-lossy problems which will be the core of our analysis in the subsequent sections. Our analysis applies to both classical and quantum reductions. For the sake of simplicity and generality, we only refer to quantum reductions and we explicitly highlight the distinction when necessary.

4.1 f -Distinguisher Reductions

A Karp decision-to-decision reduction R from Π to Σ has the following property: $\chi_\Pi(x) = 1$ if and only if $\chi_\Sigma(R(x))$ (up to some error). In our work, the target problem Σ is not restricted and does not play any roles. Therefore, we consider the following more general notion: a mapping R is a reduction if there exists a (possibly unbounded) distinguisher \mathcal{D} that can tell $R(x)$ and $R(x')$ apart, when $\chi_\Pi(x) \neq \chi_\Pi(x')$ (up to some error). A reduction is therefore a mapping that preserves the

distinguishing power of the unbounded algorithm.¹ In other words, it preserves some information about the inputs. When the reduction is to a search problem, there must also exist an inverting algorithm such that given x and the solution (or witness) of $R(x)$, outputs $\chi_\Pi(x)$. To include such reductions, we generalize this definition once more by allowing the distinguisher to have one and only one of the instances x or x' . To see how this helps, we give an example: the reduction from PARAMSAT to MAXSAT. In PARAMSAT, an instance $x := (\varphi, k)$, with φ a CNF formula and k an integer, is a YES instance if and only if at least k clauses of φ are satisfiable. The MAXSAT problem asks to find an assignment that satisfies the maximum number of clauses. Consider the decision-to-search reduction as follows: given an instance $x := (\varphi, k)$ of PARAMSAT, the outputs of the reduction is φ . By having k and an assignment w_φ satisfying the maximum number of clauses of φ (solution of φ as a MAXSAT instance), it computes $\chi_{\text{PARAMSAT}}(x)$ by comparing k and the number of satisfied clauses by w_φ . Note that it is necessary for the inverting algorithm to know k . In this subsection, we show that such reductions can be captured by the generalized distinguisher reductions:

Definition 14 (f -Distinguisher Reduction). Let n, m be positive integers, and $\mu : \mathbb{N} \rightarrow [0, 1]$ be a function of n . Let $f : \{0, 1\}^m \rightarrow \{0, 1\}$, and Π be a promise problem. A (μ, f^m) -distinguisher reduction for Π is a mapping $R : \{0, 1\}^* \rightarrow S$, where $S = \{0, 1\}^*$ (classical) or $S = \text{MS}_*$ (quantum), for which there exists an unbounded distinguisher \mathcal{D} , such that for all (x_1, \dots, x_m) and (x'_1, \dots, x'_m) in $((\Pi_Y \cup \Pi_N) \cap \{0, 1\}^n)^m$ where $f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m)) \neq f(\chi_\Pi(x'_1), \dots, \chi_\Pi(x'_m))$, we have

$$\mathbb{E}_{i \sim \mathcal{U}_{[m]}} |\Pr[1 \leftarrow \mathcal{D}(h_i, R(x_1, \dots, x_m))] - \Pr[1 \leftarrow \mathcal{D}(h_i, R(x'_1, \dots, x'_m))]| \geq 1 - 2\mu(n),$$

where $h_i := (x_i, \{\chi_\Pi(x_j)\}_j, \{\chi_\Pi(x'_j)\}_j)$. We call μ the error of the reduction.

f -Reductions

Drucker [Dru15, Definition 8.2] defines an f -compression reduction for a promise problem Π in a somewhat similar fashion that we define f -distinguisher reductions: as a mapping that sends an instances x_1, \dots, x_m of size n to a quantum state ρ , such that there exists a binary measurement \mathcal{M} (not necessarily efficient) that outputs $f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m))$ with probability more than $1 - \mu$. We adapt this definition as below.

Definition 15 (f -Reduction). Let n, m be positive integers, and $\mu : \mathbb{N} \rightarrow [0, 1]$ be a function of n . Let $f : \{0, 1\}^m \rightarrow \{0, 1\}$, and Π be a promise problem. A (μ, f^m) -reduction for Π is a mapping $R : \{0, 1\}^{mn} \rightarrow S$, where $S = \{0, 1\}^*$ (classical) or $S = \text{MS}_*$ (quantum), for which there exists a family of unbounded algorithms $\{\mathcal{M}_k\}_{k \in \mathbb{N}}$, such that for all $(x_1, \dots, x_m) \in ((\Pi_Y \cup \Pi_N) \cap \{0, 1\}^n)^m$,

$$\Pr[\mathcal{M}(R(x_1, \dots, x_m)) = f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m))] \geq 1 - \mu(n),$$

where the probability is taken over the randomness of R and \mathcal{M} . We call μ the error of the reduction.²

In the following, we show that f -reductions are special cases of f -distinguisher reductions (per Definition 14) when the hint h_i is set to be empty.

¹Note that an unbounded algorithm can always distinguish YES and NO instances of a problem by simply solving them.

²When considering quantum mappings, \mathcal{M} can be a binary quantum measurement.

Lemma 13. *Let $f : \{0,1\}^m \rightarrow \{0,1\}$, and Π be a promise problem. If R is a (μ, f^m) -reduction for Π , then R is also a (μ, f^m) -distinguisher reduction for Π .*

Proof. Recall that for an f -reduction there exists an algorithm \mathcal{M} such that

$$\Pr[\mathcal{M}(R(x_1, \dots, x_m)) = f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m))] \geq 1 - \mu(n),$$

which implies that \mathcal{M} can distinguish $R(x_1, \dots, x_m)$ from $R(x'_1, \dots, x'_m)$ with probability at least $1 - 2\mu$. Therefore, there exists an unbounded distinguisher \mathcal{D} such that for h_i per Definition 14, we have

$$\begin{aligned} \mathbb{E}_{i \sim \mathcal{U}_{[m]}} & \left| \Pr[1 \leftarrow \mathcal{D}(h_i, R(x_1, \dots, x_m))] - \Pr[1 \leftarrow \mathcal{D}(h_i, R(x'_1, \dots, x'_m))] \right| \\ & \geq \left| \Pr[1 \leftarrow \mathcal{M}(R(x_1, \dots, x_m))] - \Pr[1 \leftarrow \mathcal{M}(R(x'_1, \dots, x'_m))] \right| \geq 1 - 2\mu, \end{aligned}$$

where for the first inequality we used the fact that revealing more information to the distinguisher does not decrease its advantage. \square

Turing and Karp Reductions

In this part, we focus on (non-adaptive) Turing and Karp reductions, demonstrating that they are f -distinguisher reductions. This supports the generality of Definition 14 and will be used in Section 8.

In the following, we first recall the definition of Karp and (non-adaptive) Turing reductions in Definitions 16 and 17, and prove in Lemmas 14 and 15 that the two are f -distinguisher reductions.

Definition 16 (Non-Adaptive Turing f -Reduction). *Let n be a positive integer and $\mu : \mathbb{N} \rightarrow [0, 1]$ be a function of n . Let $f : \{0,1\}^m \rightarrow \{0,1\}$, Π be a promise problem, and Σ be a promise or search problem. A non-adaptive (μ, f^m) -Turing reduction from Π to Σ consists of an algorithm R_{Turing} that on input (x_1, \dots, x_m) , where $x_i \in \{0,1\}^n$ for $i \in [m]$, outputs $(y_1, \dots, y_k) \in \{0,1\}^*$ and a circuit C such that*

- if Σ is a promise problem:

$$\Pr[C(y_1, \chi_\Sigma(y_1), \dots, y_k, \chi_\Sigma(y_k)) = f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m))] \geq 1 - \mu(n).$$

- if Σ is a search problem:

$$\Pr[C(y_1, w_{y_1}, \dots, y_k, w_{y_k}) = f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m))] \geq 1 - \mu(n),$$

where w_{y_i} is the witness of y_i in Σ for all $i \in [k]$.

The definition above can be generalized in the following manner: y_i 's can be instances of different problems Σ_i 's instead of one single problem Σ . All our results also hold in this setting.

Definition 17 (Karp f -Reduction). *Let n be a positive integer and $\mu : \mathbb{N} \rightarrow [0, 1]$ be a function of n . Let $f : \{0,1\}^m \rightarrow \{0,1\}$ and Π be a promise problem and Σ be a promise or search problem. A (μ, f^m) -Karp reduction from Π to Σ consists of an algorithm R_{Karp} and a circuit C , where R_{Karp} on input (x_1, \dots, x_m) , where $x_i \in \{0,1\}^n$ for $i \in [m]$, outputs $y \in \{0,1\}^*$ such that*

- if Σ is a promise problem:

$$\Pr[C(y, \chi_\Sigma(y)) = f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m))] \geq 1 - \mu(n) .$$

- if Σ is a search problem:

$$\Pr[C(y, w_y) = f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m))] \geq 1 - \mu(n) ,$$

where w_y is the witness of y in Σ .

Note that in a Karp reduction, the circuit C does not depend on the instance x . In fact, in a standard definition of a Karp reduction to a promise problem, C simply outputs $\chi_\Pi(x)$.

In the following lemma, we show that all non-adaptive Turing reductions are f -distinguisher reduction.

Lemma 14 (Turing f -Reduction is f -Distinguisher Reduction). *Let $\mu : \mathbb{N} \rightarrow [0, 1]$. Let Π be a promise problem and Σ be a promise or search problem. If R_{Turing} is a non-adaptive (μ, f^m) -Turing reduction (Definition 16) from Π to Σ , then it is (μ, f^m) -distinguisher reduction for Π .*

Proof. The distinguisher \mathcal{D} in Figure 1 satisfies the definition of (μ, f^m) -distinguisher reductions (Definition 14). This is because if $B = ((y_1, \dots, y_k), C)$ is an output of $R_{\text{Turing}}(x_1, \dots, x_m)$, then by the correctness of the reduction, it holds with high probability that $C(y_1, \chi_\Sigma(y_1), \dots, y_k, \chi_\Sigma(y_k)) = f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m))$, if Σ is a promise problem, and similarly $C(y_1, w_{y_1}, \dots, y_k, w_{y_k}) = f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m))$, if Σ is a search problem.

Algorithm 1 Distinguisher \mathcal{D} for non-adaptive Turing reductions.

Parameters: n, m, f, Π, Σ

Input: A pair (h_i, B) , where $h_i := (x_i, \{\chi_\Pi(x_j)\}_j, \{\chi_\Pi(x'_j)\}_j)$ for a uniformly random $i \in [m]$ and $B = ((y_1, \dots, y_k), C)$.

Promise: Either $B \leftarrow R(x_1, \dots, x_m)$ or $B \leftarrow R(x'_1, \dots, x'_m)$ for some $(x'_1, \dots, x'_m) \in (\{0, 1\}^n)^m$ such that $f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m)) \neq f(\chi_\Pi(x'_1), \dots, \chi_\Pi(x'_m))$.

Output: A bit b .

- 1: Parse $h_i := (x_i, \{\chi_\Pi(x_j)\}_j, \{\chi_\Pi(x'_j)\}_j)$ and $B = ((y_1, \dots, y_k), C)$.
 - 2: **if** Σ is a promise problem: **then**
 - 3: Compute $\chi_\Sigma(y_1), \dots, \chi_\Sigma(y_k)$.
 - 4: Compute $\hat{b} \leftarrow C(y_1, \chi_\Sigma(y_1), \dots, y_k, \chi_\Sigma(y_k))$.
 - 5: **else**
 - 6: Compute the witnesses w_{y_1}, \dots, w_{y_k} in Σ .
 - 7: Compute $\hat{b} \leftarrow C(y_1, w_{y_1}, \dots, y_k, w_{y_k})$.
 - 8: **if** $\hat{b} = f(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m))$ **then**
 - 9: Return 1.
 - 10: **else**
 - 11: Return 0.
-

□

Lemma 15 (R_{Karp} is f -Distinguisher Reduction). *Let $\mu : \mathbb{N} \rightarrow [0, 1]$. Let Π be a promise problem and Σ be a promise or search problem. If R_{Karp} is a (μ, f^m) -Karp reduction (Definition 17) from Π to Σ , then it is (μ, f^m) -distinguisher reduction for Π .*

Proof. Since any Karp reduction is a Turing reduction, the statement holds due to Lemma 14. □

4.2 Mildly-Lossy Problems

To analyze the lossiness of f -distinguisher reductions, we fix the set of functions f to those ones that are invariant under permuting their inputs.

Definition 18 (Permutation-Invariant Boolean Function). A Boolean function $f : \{0, 1\}^m \rightarrow \{0, 1\}$ is called permutation-invariant if for every $\pi \in \mathfrak{S}_m$, it holds that $f(\pi(b_1, b_2, \dots, b_m)) = f(b_1, b_2, \dots, b_m)$.

This set of functions is of great interest. The functions AND, OR, and MAJ that were considered in [Dru15, BBD⁺20] are all non-constant permutation-invariant. Moreover, the (non-monotone) functions PARITY and MOD_k are of this type as well as THRESHOLD_k.

We use the following technical lemma about non-constant permutation-invariant functions.

Lemma 16. Let $f : \{0, 1\}^m \rightarrow \{0, 1\}$ be a non-constant permutation-invariant function. Then there exists an integer $1 \leq p \leq m$ such that

$$f(\underbrace{1, 1, \dots, 1}_{p-1}, 0, 0, \dots, 0) = 0, \quad \text{and} \quad f(\underbrace{1, 1, \dots, 1}_p, 0, 0, \dots, 0) = 1.$$

We let $p(f)$ denote the minimum choice of such an integer.

Proof. The set $\{0, 1\}^m$ can be partitioned into $m+1$ equivalence classes where each class consists of strings with the same number of 1's. We note that the result of a permutation on an input falls in the same equivalence class. Therefore, since the function is permutation-invariant, then the evaluation of f over each input is determined by its class. Because the function is non-constant, there must exist two consecutive classes (the classes can be ordered by the number of 1's that they represent) with different evaluation under f . This completes the proof. \square

Finally, we introduce the notion of *mildly-lossy problems* which are promise problems that admit lossy f -distinguisher reductions where f is a non-constant permutation-invariant function.

Definition 19 (Mildly-Lossy Problems). Let n, m be positive integers, λ, T, γ be positive reals, and $\mu \in [0, 1/2)$. A promise problem Π is said to be $(T, \mu, f^m, \lambda, \gamma)$ -mildly-lossy if there exists a non-uniform (μ, f^m) -distinguisher reduction R (per Definition 14) for Π with the following properties:

1. f is some non-constant permutation-invariant function $f : \{0, 1\}^m \rightarrow \{0, 1\}$, and
2. the reduction R runs in time T , and
3. and R is splitting $m\lambda$ -lossy (per Definition 12) supported on (Π_Y, Π_N) , for all pairwise independent $(2^9 mn / \gamma^3)$ -uniform distributions over n -bit strings.

We explicitly mention the type of the reduction R (classical or quantum) when the distinction is necessary. Also, we interchangeably say that the reduction R as above is mildly-lossy.

Note that the sparseness is controlled by the parameter γ . In the original full-fledged lossiness, γ is exponentially small. However, to obtain one-way functions, it suffices that γ be roughly bounded by $\text{poly}(1/T, 1/n)$ (see section 9 for more details). When considering polynomial-time reductions, the distribution is indeed very sparse, with a support of polynomial size.

Recall δ from the upper bound for splitting lossy functions in Lemma 9. We include it here for clarity as it will be frequently used in all sections.

Definition 20. We let $\delta : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ to be the following function

$$\delta(\lambda) := \min \left\{ \sqrt{\frac{\lambda \ln 2}{2}}, 1 - 2^{-\lambda-2} \right\}.$$

5 Zero-Knowledgeness from Mildly-Lossy Problems

In this section, we show that lossy problems admit Karp reductions to the statistical difference problem or the quantum state distinguishability problem, depending on the type of the lossy reduction. We provide a fine-grained analysis. When restricted to polynomial-time AND-compression reductions, this recreates the result of Drucker [Dru15, Theorem 8.14]: roughly, if a promise problem Π has a (quantum) polynomial-time AND-compression reduction, then Π must belong to SZK (resp., QSZK). Similar statement holds for the AND- or MAJ-lossy reductions (see [BBD⁺20]). We note that our result holds for any non-constant permutation-invariant function, requires less restricted notion of lossiness, and allows superpolynomial-time reductions.

Theorem 1. Let Π be $(T, \mu, f^m, \lambda, \gamma)$ -mildly-lossy. Assume that $\theta_{\text{szk}} := (1 - 2\mu)^2 / (\delta(\lambda) + \gamma) > 1$, with $\delta(\lambda)$ as in Definition 20. Then Π reduces to a problem in QSZK in time $O((T + m^2 n) / (\gamma \log \theta_{\text{szk}}))$ and with a classical advice of size $4mn/\gamma$ as described in Algorithm 2. Moreover, the reduction is deterministic (but non-uniform) and Π reduces to SZK if Π is lossy with respect to a classical reduction.

Algorithm 2 Reduction from Π to $\text{QSD}_{1/4, 3/4}$.

Parameters: $n, m, \mu, f, \lambda, \gamma, R, \Pi$ as in Definition 19. Further

$$S_0 := \Pi_N \cap \{0, 1\}^n, \quad S_1 := \Pi_Y \cap \{0, 1\}^n, \quad \varepsilon := \frac{\gamma}{4}, \quad d := \left\lceil \frac{m+1}{\varepsilon} \right\rceil, \quad s := \left\lceil \frac{n \ln 2}{2\varepsilon^2} \right\rceil,$$

and $K_1, \dots, K_s, T_1, \dots, T_s$ as in Lemma 9.

Input: An instance $y \in \{0, 1\}^n$.

Advice: $p := p(f)$ as in Lemma 16, $b_Y, b_N \in \{0, 1\}$ respectively representing whether $\Pi_Y \cap \{0, 1\}^n$ and $\Pi_N \cap \{0, 1\}^n$ are empty. K_a, T_a, π for some uniformly chosen $a \in [s]$ and $\pi \in \mathfrak{S}_m$.

Output: A pair of circuits (C_0, C_1) .

- 1: If $b_N = 1$, return (Y_0, Y_1) where $\|Y_0 - Y_1\|_1 \leq 1/4$.
 - 2: If $b_Y = 1$, return (N_0, N_1) where $\|N_0 - N_1\|_1 \geq 3/4$.
 - 3: Let \widehat{C}_0 be the following circuit: it samples $\tilde{x} \sim (\mathcal{U}_{K_a}^{\otimes m-p+1}, \mathcal{U}_{T_a}^{\otimes p-1})$, then it outputs $R(\pi(\tilde{x}))$.
 - 4: Let \widehat{C}_1 be the following circuit: it samples $\tilde{x} \sim (\mathcal{U}_{K_a}^{\otimes m-p}, y, \mathcal{U}_{T_a}^{\otimes p-1})$, then it outputs $R(\pi(\tilde{x}))$.
 - 5: Compute $(C_0, C_1) \leftarrow \text{Polarize}(\widehat{C}_0, \widehat{C}_1, 1^2)$.
 - 6: Return (C_0, C_1) .
-

Remark 3 (Input-output type of the circuits). Consider the two circuits $(\widehat{C}_0, \widehat{C}_1)$ in Algorithm 2, Lines 3 and 4. When R is a randomized reduction, the two circuits are also randomized. Part of their randomness input is used to sample \tilde{x} and the other part is fed to R . Let κ be the size of the total randomness. For $r \in \{0, 1\}^\kappa$ and any $b \in \{0, 1\}$, we let $\widehat{C}_b(r)$ denote the outcome

of \widehat{C}_b given the randomness r . On the other hand, when R is quantum, the circuits will be mixed algorithms; classical randomness is required for sampling \tilde{x} . Let κ' be the size of total randomness.¹ For any $r \in \{0, 1\}^{\kappa'}$ and any $b \in \{0, 1\}$, we let the mixed outcome of \widehat{C}_b be $\widehat{C}_b |r, \mathbf{0}\rangle$ where $|\mathbf{0}\rangle$ is some appropriate-size ancilla, emphasizing its mixed classical-quantum nature. When it is not relevant, we drop the dependency on r for simplification.

Proof of Theorem 1. In the following, we assume that R is quantum. The classical case is similar with the only difference being the type of the inputs and outputs of $(\widehat{C}_0, \widehat{C}_1)$.

Consider the case $y \in \Pi_Y$. We bound the ℓ_1 distance (per Definition 3) of the outcomes of \widehat{C}_0 and \widehat{C}_1 from below. Sample a uniform coin $b \sim U_{\{0,1\}}$, and let $z \leftarrow \widehat{C}_b |r, \mathbf{0}\rangle$ where r follows the uniform distribution. We drop the dependency on r for simplification. Let \mathcal{A} be a (possibly unbounded) distinguisher that takes z as input and guesses which circuit (\widehat{C}_0 or \widehat{C}_1) is used to compute z . Let \mathcal{A} be the quantum distinguisher of the (μ, f^m) -distinguisher reduction (that comes from Definition 19) for Π . On the one hand, if z is computed by \widehat{C}_0 , we have that $\tilde{x} := (x_1, \dots, x_m) \sim (\mathcal{U}_{K_a}^{\otimes m-p+1}, \mathcal{U}_{T_a}^{\otimes p-1})$ with $K_a \subseteq \Pi_N \cap \{0, 1\}^n$ and $T_a \subseteq \Pi_Y \cap \{0, 1\}^n$. Then, since \tilde{x} contains $p-1$ YES instances by Lemma 16, for any $\pi \in \mathfrak{S}_m$, we have

$$f(\pi(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m))) = 0.$$

On the other hand, if z is computed by \widehat{C}_1 , we have that \tilde{x} contains one more YES instance $y \in \Pi_Y \cap \{0, 1\}^n$, therefore,

$$f(\pi(\chi_\Pi(x_1), \dots, \chi_\Pi(x_m))) = 1.$$

Moreover, revealing π with the description of the circuits does not decrease the success probability of the distinguisher, thus by the quantum f -distinguishability of the reduction, we have

$$\begin{aligned} \|\widehat{C}_0 |\mathbf{0}\rangle - \widehat{C}_1 |\mathbf{0}\rangle\|_1 &\geq \mathbb{E}_{i \sim \mathcal{U}_{[m]}} \left| \Pr[1 \leftarrow \mathcal{D}(x_i, \widehat{C}_0 |\mathbf{0}\rangle)] - \Pr[1 \leftarrow \mathcal{D}(x_i, \widehat{C}_1 |\mathbf{0}\rangle)] \right| \\ &\geq 1 - 2\mu(n). \end{aligned}$$

Now, we discuss the case of $y \in \Pi_N$. We consider a modification of the distinguishing game where the random variables a and π are also given to the distinguisher. Revealing a, π along with z does not decrease the success probability of the distinguisher, thus we can bound the original distinguishing probability by the distinguishing probability of the new task. It holds that

$$\|\widehat{C}_0 |\mathbf{0}\rangle - \widehat{C}_1 |\mathbf{0}\rangle\|_1 \leq \left\| R\left(\pi\left(\mathcal{U}_{K_a}^{\otimes m-p+1}, \mathcal{U}_{T_a}^{\otimes p-1}\right)\right) - R\left(\pi\left(\mathcal{U}_{K_a}^{\otimes m-p}, y, \mathcal{U}_{T_a}^{\otimes p-1}\right)\right) \right\|_1,$$

By taking the expectation over a and π , we have

$$\begin{aligned} \|\widehat{C}_0 |\mathbf{0}\rangle - \widehat{C}_1 |\mathbf{0}\rangle\|_1 &\leq \mathbb{E}_{a \sim \mathcal{U}_{[s]}, \pi \sim \mathfrak{S}_m} \left[\left\| R\left(\pi\left(\mathcal{U}_{K_a}^{\otimes m-p+1}, \mathcal{U}_{T_a}^{\otimes p-1}\right)\right) \right. \right. \\ &\quad \left. \left. - R\left(\pi\left(\mathcal{U}_{K_a}^{\otimes m-p}, y, \mathcal{U}_{T_a}^{\otimes p-1}\right)\right) \right\|_1 \right]. \end{aligned} \tag{7}$$

By our choice of $\varepsilon, d, s, K_1, \dots, K_s, T_1, \dots, T_s$ and Lemma 9, we conclude that

$$\|\widehat{C}_0 |\mathbf{0}\rangle - \widehat{C}_1 |\mathbf{0}\rangle\|_1 \leq \delta + \frac{2(m-p+1)}{d+1} + 2\varepsilon \leq \delta + \gamma.$$

¹Note that κ and κ' are possibly different depending on how much classical randomness R requires.

Let $\alpha := (\delta + \gamma)$ and $\beta := (1 - 2\mu)$. Above, we proved that (\hat{C}_0, \hat{C}_1) is an instance of $\text{QSD}_{\alpha, \beta}$ of size $(T + m^2n)/\gamma$. By assumption, we have $\theta_{\text{szk}} = \beta^2/\alpha$. Therefore, the runtime of $\text{Polarize}(\hat{C}_0, \hat{C}_1, 1^2)$ and its output size are both of $O((T + m^2n)/(\gamma \log \theta_{\text{szk}}))$ according to Lemma 8. \square

6 One-Way Functions from Mildly-Lossy Problems

In this section and in Section 7, we discuss how mildly-lossy problems can be used to build cryptographic primitives. In Theorem 2, we construct EFI schemes. The statement allows both classical reductions and quantum reductions. We immediately obtain one-way functions (or quantum bit commitments if the reduction is quantum), by taking into account the known transforms from EFI schemes (see Remark 1). However, the required condition on the lossiness is highly restrictive. More precisely, λ must be a small constant. In Theorem 3 and 4, we explain how one can tackle this issue using different constructions. The construction in Theorem 3 is inspired by [BBD⁺20], and resist adaptations to the quantum settings. On the other hand, the construction in Theorem 4 is quite flexible and allows obtaining one-way state generators. Finally, we note that the latter does imply one-way functions, too, but for simplicity, we only discuss one-way state generators.

Theorem 2. *Let Π be $(T, \mu, f^m, \lambda, \gamma)$ -mildly-lossy. Assume that $\theta_{\text{efi}} := (1 - 2\mu) - 3(\delta(\lambda) + \gamma) > 0$, with $\delta(\lambda)$ as in Definition 20. Then there exists an algorithm EFI that runs in $O(T + m^2n\gamma^{-1})$ and an oracle algorithm \mathcal{C} , such that for any algorithm \mathcal{A} one and only one of the following statements holds:*

- I. $\mathcal{C}^{\mathcal{A}}$ solves $\Pi \cap \{0, 1\}^n$ in time $O((T + m^2n\gamma^{-1})\theta_{\text{efi}}^{-2})$ with $O(\theta_{\text{efi}}^{-2})$ queries to \mathcal{A} ,
- II. EFI is $(1 - 2\mu, 1 - 2\mu - \theta_{\text{efi}}/2)$ -EFI for \mathcal{A} .

Moreover, if the mildly-lossy reduction of Π is classical, EFI would also be classical.

Remark 4. From the conditions of Theorem 2, it must hold that $\delta < 1/3$, therefore, λ must be small. Most notably, the statement does not include perfect 1-mildly-lossy reductions. However, this can be overcome as follows: Let R be 1-mildly-lossy and perfect. Consider the new reduction R' that with probability 0.35 randomly outputs a YES or a NO instance of the target language (note that instance can be given as advice). Otherwise, it applies R . The new reduction is 0.35-mildly-lossy with error 0.375 which satisfies the condition $(1 - 2\mu) - 3(\delta(\lambda) + \gamma) > 0$.

Proof. We prove the case where R is quantum. The classical case can be done similarly. Let Π be the promised problem in the statement. Let \mathcal{F} denote Algorithm 2 that returns the two circuits in Lines 3 and 4, and h be its advice as follows: $h := (K_a, T_a, p, b_Y, b_N)$. The construction of the non-uniform EFI is the following:

- $\text{EFI}_h(1^n, b)$: Sample $y \sim \mathcal{U}_{T_a}$. Compute $(\hat{C}_0, \hat{C}_1) \leftarrow \mathcal{F}(y)$. Return the state $\hat{C}_b |0\rangle$.

Note that T_a has only YES instances.

The two output states are statistically far. By Theorem 1, the pair of circuits $(\hat{C}_0, \hat{C}_1) \leftarrow \mathcal{F}(y)$ is a $\text{QSD}_{1-2\mu, \delta+\gamma}$ instance. Since $y \in \Pi_Y$, then $\|\hat{C}_0 |0\rangle - \hat{C}_1 |0\rangle\|_1 \geq 1 - 2\mu$. This concludes the statistical distinguishability.

On the computational indistinguishability, we will argue by contradiction. Assume there exists an adversary \mathcal{A} that distinguishes the EFI states $\widehat{C}_b |\mathbf{0}\rangle$ with advantage ν that is to be determined later. Let us consider an algorithm \mathcal{B} targetting Π as follows: given an instance $z \in \{0, 1\}^n$, it first computes $(C'_0, C'_1) \leftarrow \mathcal{F}(z)$, then it samples a uniform coin $b \sim \mathcal{U}_{\{0,1\}}$ and relays $C'_b |\mathbf{0}\rangle$ to the distinguisher \mathcal{A} . Finally, \mathcal{B} will return 1 if \mathcal{A} returns b , and 0 otherwise.

Case $z \in \Pi_Y$: Suppose that z has been sampled from \mathcal{U}_{T_a} . Then, the (mixed) state $C'_b |\mathbf{0}\rangle$ that we deliver to the adversary \mathcal{A} would be identical to the EFI state $\widehat{C}_b |\mathbf{0}\rangle$. Therefore, from the ν -distinguishability of EFI states for \mathcal{A} , we would have

$$\Pr(\mathcal{B}(z) = 1) = \Pr(\mathcal{A}(P_b | \mathbf{0}) = b) \geq \frac{1}{2} + \frac{\nu}{2}.$$

We know that z does not necessarily follow the distribution \mathcal{U}_{T_a} . However, one can argue that \widehat{C}_b is not far from C'_b by leveraging the disguising lemma. We have that

$$\begin{aligned} \|\widehat{C}_0 \otimes \widehat{C}_1 |\mathbf{0}, \mathbf{0}\rangle - C'_0 \otimes C'_1 |\mathbf{0}, \mathbf{0}\rangle\|_1 &\leq \|\widehat{C}_1 |\mathbf{0}\rangle - C'_1 |\mathbf{0}\rangle\|_1 \\ &\leq \mathbb{E}_{a \sim \mathcal{U}_{[s]}, \pi \sim \mathfrak{S}_m} \left[\left\| R \left(\pi \left(\mathcal{U}_{K_a}^{\otimes m-p}, \mathcal{U}_{T_a}^{\otimes p} \right) \right) \right. \right. \\ &\quad \left. \left. - R \left(\pi \left(\mathcal{U}_{K_a}^{\otimes m-p}, y, \mathcal{U}_{T_a}^{\otimes p-1} \right) \right) \right\|_1 \right] \\ &\leq \delta + \frac{2(m+1)}{d+1} + \varepsilon \\ &\leq \delta + \gamma, \end{aligned}$$

where we used the fact that $\widehat{C}_0 = C'_0$, properties of trace distance, and Lemma 9. Using the fact that the trace distance is decreasing under partial trace, for any $b \in \{0, 1\}$, we obtain

$$\|\widehat{C}_b |\mathbf{0}\rangle - C'_b |\mathbf{0}\rangle\|_1 \leq \delta + \gamma.$$

The adversary \mathcal{A} can thus distinguish the general C'_b with probability

$$\begin{aligned} \Pr(\mathcal{B}(z) = 1) &= \Pr(\mathcal{A}(C'_b |\mathbf{0}\rangle) = b) = \frac{1}{2} + \frac{1}{2} \left| \Pr_{x \leftarrow C'_0} (\mathcal{A}(x) = 1) - \Pr_{x \leftarrow C'_1} (\mathcal{A}(x) = 1) \right| \\ &\geq \frac{1}{2} + \frac{1}{2} \left(\left| \Pr_{x \leftarrow \widehat{C}_0} (\mathcal{A}(x) = 1) - \Pr_{x \leftarrow \widehat{C}_1} (\mathcal{A}(x) = 1) \right| \right. \\ &\quad \left. - \left| \Pr_{x \leftarrow \widehat{C}_0} (\mathcal{A}(x) = 1) - \Pr_{x \leftarrow C'_0} (\mathcal{A}(x) = 1) \right| \right. \\ &\quad \left. - \left| \Pr_{x \leftarrow \widehat{C}_1} (\mathcal{A}(x) = 1) - \Pr_{x \leftarrow C'_1} (\mathcal{A}(x) = 1) \right| \right) \\ &\geq \frac{1}{2} + \frac{\nu}{2} - \delta - \gamma. \end{aligned} \tag{8}$$

Case $z \in \Pi_N$: By Theorem 1, the two circuits $(C'_0, C'_1) \leftarrow \mathcal{F}(z)$ are close in trace distance, namely,

$$\|C'_0 |\mathbf{0}\rangle - C'_1 |\mathbf{0}\rangle\|_1 \leq \delta + \gamma.$$

Recall that the trace distance provides the maximum distinguishability advantage for *any* distinguisher, including \mathcal{A} , therefore

$$\Pr(\mathcal{B}(z) = 1) = \Pr(\mathcal{A}(C'_b | \mathbf{0}) = b) \leq \frac{1}{2}(1 + \|C'_0 | \mathbf{0}\rangle - C'_1 | \mathbf{0}\rangle\|_1) \leq \frac{1}{2}(1 + \delta + \gamma) . \quad (9)$$

Conclusion: We need one more algorithm that will leverage the capacity of \mathcal{B} to decide Π . Let $k \in \mathbb{N}$, and \mathcal{C} be an algorithm that on instance $z \in \{0, 1\}^n$, runs $\mathcal{B}(z)$ for k times independently. Let b_1, \dots, b_k be k corresponding independent outputs of $\mathcal{B}(z)$. Then \mathcal{C} returns as follows:

$$\begin{cases} 0 & \text{if } \left| \frac{1}{k} \sum_i b_i - \frac{1}{2} \right| \geq \tau, \\ 1 & \text{otherwise,} \end{cases}$$

where $\tau(n)$ is chosen such that

$$\tau := \frac{\nu}{4} - \frac{3(\delta + \gamma)}{4} . \quad (10)$$

Then, we have

$$\begin{aligned} \Pr(\mathcal{C}(z) = 0 | z \in \Pi_Y) &= \Pr\left(\left|\frac{1}{k} \sum_i b_i - \frac{1}{2}\right| \geq \tau \mid b_1, \dots, b_k \leftarrow \mathcal{B}(z), z \in \Pi_Y\right) \\ &\geq \Pr\left(\frac{1}{k} \sum_i b_i \geq \frac{1}{2} + \tau \mid b_1, \dots, b_k \leftarrow \mathcal{B}(z), z \in \Pi_Y\right) \\ &\geq \Pr\left(\frac{1}{k} \left(\sum_i b_i - \mathbb{E}(\mathcal{B}_i(z))\right) \geq -\tau \mid b_1, \dots, b_k \leftarrow \mathcal{B}(z), z \in \Pi_Y\right) \\ &\geq 1 - \exp(-2k\tau^2) , \end{aligned}$$

where we used $\mathbb{E}(\mathcal{B}_i(z)) - \tau \geq \frac{1}{2} + \tau$ for $z \in \Pi_Y$ by Equation (8) in the second inequality, and Hoeffding's lemma in the last inequality. On the other hand, we have

$$\begin{aligned} \Pr(\mathcal{C}(z) = 1 | z \in \Pi_N) &= \Pr\left(\left|\frac{1}{k} \sum_i b_i - \frac{1}{2}\right| < \tau \mid b_1, \dots, b_k \leftarrow \mathcal{B}(z), z \in \Pi_N\right) \\ &= \Pr\left(\left|\frac{1}{k} \sum_i b_i - \frac{1}{k} \sum_i \mathbb{E}(\mathcal{B}_i(z))\right| < \tau \mid b_1, \dots, b_k \leftarrow \mathcal{B}(z), z \in \Pi_N\right) \\ &\geq 1 - \exp(-2k\tau^2) , \end{aligned}$$

where we once again used Hoeffding's lemma and Equation (9). For $k := 1/\tau^2$, any sufficiently large $n \in \mathbb{N}$, and any $z \in (\Pi_Y \cup \Pi_N) \cap \{0, 1\}^n$, it holds that

$$\Pr(\mathcal{C}(z) = \chi_\Pi(z)) \geq 1 - \exp(-2k\tau^2) \geq \frac{2}{3} ,$$

This breaks the worst-case hardness of Π .

Since $\theta_{\text{eff}} := (1 - 2\mu) - 3(\delta + \gamma)$, we can set $\nu := (1 - 2\mu) - \theta_{\text{eff}}/2$, and the number of repetitions in the last step becomes

$$1/\tau^2 = \frac{4^2}{(\nu - 3(\delta + \gamma))^2} = \frac{4^3}{\theta_{\text{eff}}^2} .$$

Runtime: We compute the runtime of **EFI** as follows. It first samples $2m$ instances from \mathcal{U}_{K_a} (or \mathcal{U}_{T_a}), applies the permutations π twice to each half of the samplings, and computes R on each half. One single sampling from \mathcal{U}_{K_a} (or \mathcal{U}_{T_a}) takes time $O(dn)$, where $d \leq (m+1)/\gamma$ is the size of \mathcal{U}_{K_a} and n is the size of each element in \mathcal{U}_{K_a} . The permutations can be applied in time $O(m)$. Therefore, the total runtime of **EFI** is $O(T + m^2n/\gamma)$.

Note that \mathcal{C} runs \mathcal{B} for $O(1/\theta_{\text{eff}}^2)$ times. Each execution of \mathcal{B} evaluates \hat{C}_b , queries \mathcal{A} , and performs an equality check. All of this takes $O((T + m^2n/\gamma)/\theta_{\text{eff}}^2)$ with $O(1/\theta_{\text{eff}}^2)$ queries to \mathcal{A} . \square

Next, we consider larger values of λ , for instance when $\lambda = \Omega(\log n)$. The following concerns only *classical* one-way functions.

Theorem 3. *Let Π be $(T, \mu, f^m, \lambda, \gamma)$ -mildly-lossy with a classical reduction. Assume that $\theta_{\text{owf}} := (1 - 10\mu) - (\delta(\lambda) + \gamma) > 0$, with $\delta(\lambda)$ as in Definition 20. Then there exists an algorithm \mathbf{F} that runs in time $O(T + m^2n\gamma^{-1})$ and an oracle algorithm \mathcal{C} , such that for any algorithm \mathcal{A} one and only one of the following holds:*

- I. $\mathcal{C}^{\mathcal{A}}$ solves $\Pi \cap \{0, 1\}^n$ in time $O((T + m^2n\gamma^{-1})\theta_{\text{owf}}^{-2})$ with $O(\theta_{\text{owf}}^{-2})$ queries to \mathcal{A} ,
- II. \mathbf{F} is a $(1 - \theta_{\text{owf}}/2)$ -OWF for \mathcal{A} .

Proof. Consider the circuit \hat{C}_0 in Line 3 of Algorithm 2. This circuit is independent of the input of Algorithm 2 and is randomized. Part of its randomness is used to sample \tilde{x} and the other part is fed to R . Let κ be the size of the total randomness. For $r \in \{0, 1\}^\kappa$, we let $\hat{C}_0(r)$ be the outcome of the circuit when it is given r as the randomness. We show that \mathbf{F} , defined by $\hat{C}_0(\cdot) : \{0, 1\}^\kappa \rightarrow \{0, 1\}^*$, is a $(\theta_{\text{owf}}/2)$ -weak one-way function. This suffices for the proof since weak one-way functions imply one-way functions.

The proof works by a reduction to the worst-case hardness of Π . Assume that we are given a to-be-decided instance y of Π . Apply Algorithm 2 up to Line 4 to obtain (\hat{C}_0, \hat{C}_1) . Assume that there exists an adversary \mathcal{A} that inverts $\hat{C}_0(\cdot)$ with probability more than $1 - \theta_{\text{owf}}/2$. Consider the following oracle algorithm $\mathcal{B}^{\mathcal{A}}$:

- $\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y)$: samples a uniform $r \in \{0, 1\}^\kappa$ and a uniform $b \in \{0, 1\}$, and computes $z := \hat{C}_b(r)$. Runs the adversary $r' \leftarrow \mathcal{A}(z)$, and computes $z' = \hat{C}_0(r')$. If $z = z'$ it outputs 1, otherwise it outputs 0.

We show that \mathcal{B} can distinguish between the YES and NO instances of Π by analysing the probability of outputting 1. More precisely, we study the following random variable:

$$X(\hat{C}_0, \hat{C}_1, y) := \left| \Pr\left(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1 | b = 0\right) - \Pr\left(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1 | b = 1\right) \right|.$$

Case $y \in \Pi_Y$: We show the following bound for every $y \in \Pi_Y$:

$$X(\hat{C}_0, \hat{C}_1, y) > 1 - \theta_{\text{owf}}/2 - 10\mu.$$

Instead of proving the inequality directly for the circuits (\hat{C}_0, \hat{C}_1) , we will show it for two similar circuits $(\tilde{C}_0, \tilde{C}_1)$ with disjoint image sets. Let \hat{D}_0 and \hat{D}_1 be respectively the outcome distributions of \hat{C}_0 and \hat{C}_1 when given uniform input, and A be the following set

$$A := \{a \mid \Pr_{\hat{D}_0}(a) \geq \Pr_{\hat{D}_1}(a)\}.$$

Let \tilde{C}_0 be the restriction of \hat{C}_0 to A and \tilde{C}_1 the restriction of \hat{C}_1 to A^c . We will show that

$$X(\tilde{C}_0, \tilde{C}_1, y) \leq X(\hat{C}_0, \hat{C}_1, y) + 8\mu.$$

Indeed in Theorem 1, we showed that for every $y \in \Pi_Y$, the statistical distance between the outcome distributions of \hat{C}_0 and \hat{C}_1 when given uniform input is at least $1 - 2\mu$. Moreover, we have

$$\begin{aligned} \|\hat{D}_0 - \hat{D}_1\|_1 &= \frac{1}{2} \sum_a \left| \Pr_{\hat{D}_0}(a) - \Pr_{\hat{D}_1}(a) \right| \\ &= \frac{1}{2} \sum_{a \in A} \Pr_{\hat{D}_0}(a) - \Pr_{\hat{D}_1}(a) + \frac{1}{2} \sum_{a \in A^c} \Pr_{\hat{D}_1}(a) - \Pr_{\hat{D}_0}(a) \\ &= \frac{1}{2} \left(\Pr_{\hat{D}_0}(A) - \Pr_{\hat{D}_0}(A^c) + \Pr_{\hat{D}_1}(A^c) - \Pr_{\hat{D}_1}(A) \right) \\ &= \frac{1}{2} \left(\Pr_{\hat{D}_0}(A) - (1 - \Pr_{\hat{D}_0}(A)) + \Pr_{\hat{D}_1}(A^c) - (1 - \Pr_{\hat{D}_1}(A^c)) \right) \\ &= \Pr_{\hat{D}_0}(A) + \Pr_{\hat{D}_1}(A^c) - 1. \end{aligned}$$

It follows that $\Pr_{\hat{D}_0}(A) + \Pr_{\hat{D}_1}(A^c) \geq 2 - 2\mu$. Therefore, we have

$$(\Pr_{\hat{D}_0}(A) \geq 1 - \mu) \wedge (\Pr_{\hat{D}_1}(A^c) \geq 1 - 2\mu), \quad \text{or} \quad (\Pr_{\hat{D}_0}(A) \geq 1 - 2\mu) \wedge (\Pr_{\hat{D}_1}(A^c) \geq 1 - \mu).$$

Then for either of cases above, we have

$$\|\hat{D}_0 - \tilde{D}_0\|_1 \leq 2\mu, \quad \text{and} \quad \|\hat{D}_1 - \tilde{D}_1\|_1 \leq 2\mu, \tag{11}$$

where \tilde{D}_0 and \tilde{D}_1 are respectively the outcome distributions of \tilde{C}_0 and \tilde{C}_1 . Pretend that not only does \mathcal{A} invert F , but also tries to distinguish between \hat{C}_b and \tilde{C}_b for $b \in \{0, 1\}$. Consider the following sequence of games that modifies $\mathcal{B}^{\mathcal{A}}$:

Game \mathcal{G}_1 : In this game \mathcal{B} behaves originally as above.

Game \mathcal{G}_2 : In this game \mathcal{B} replaces \hat{C}_0 with \tilde{C}_0 . Note that \mathcal{A} can distinguish this modification with probability at most 2μ according to Equation (11). It follows that

$$\begin{aligned} X(\tilde{C}_0, \hat{C}_1, y) &= \left| \Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \hat{C}_1, y) = 1 | b = 0) - \Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \hat{C}_1, y) = 1 | b = 1) \right| \\ &\leq \left| \Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \hat{C}_1, y) = 1 | b = 0) - \Pr(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1 | b = 0) \right| \\ &\quad + \left| \Pr(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1 | b = 0) - \Pr(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1 | b = 1) \right| \\ &\quad + \left| \Pr(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1 | b = 1) - \Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \hat{C}_1, y) = 1 | b = 1) \right| \\ &= X(\hat{C}_0, \hat{C}_1, y) + 4\mu. \end{aligned}$$

Game \mathcal{G}_3 : In this game, \mathcal{B} replaces \widehat{C}_1 with \widetilde{C}_1 . Note that \mathcal{A} can identify this modification with probability at most 2μ . We obtain

$$\begin{aligned}
X(\widetilde{C}_0, \widetilde{C}_1, y) &= \left| \Pr(\mathcal{B}^{\mathcal{A}}(\widetilde{C}_0, \widetilde{C}_1, y) = 1 | b = 0) - \Pr(\mathcal{B}^{\mathcal{A}}(\widetilde{C}_0, \widetilde{C}_1, y) = 1 | b = 1) \right| \\
&\leq \left| \Pr(\mathcal{B}^{\mathcal{A}}(\widetilde{C}_0, \widetilde{C}_1, y) = 1 | b = 0) - \Pr(\mathcal{B}^{\mathcal{A}}(\widetilde{C}_0, \widehat{C}_1, y) = 1 | b = 0) \right| \\
&\quad + \left| \Pr(\mathcal{B}^{\mathcal{A}}(\widetilde{C}_0, \widehat{C}_1, y) = 1 | b = 0) - \Pr(\mathcal{B}^{\mathcal{A}}(\widetilde{C}_0, \widehat{C}_1, y) = 1 | b = 1) \right| \\
&\quad + \left| \Pr(\mathcal{B}^{\mathcal{A}}(\widetilde{C}_0, \widehat{C}_1, y) = 1 | b = 1) - \Pr(\mathcal{B}^{\mathcal{A}}(\widetilde{C}_0, \widetilde{C}_1, y) = 1 | b = 1) \right| \\
&= X(\widetilde{C}_0, \widehat{C}_1, y) + 4\mu \\
&\leq X(\widehat{C}_0, \widehat{C}_1, y) + 8\mu.
\end{aligned}$$

To prove the inequality for the YES instances, it suffices to show that $X(\widetilde{C}_0, \widetilde{C}_1, y) > 1 - \theta_{\text{owf}}/2 - 2\mu$. Recall that

$$X(\widetilde{C}_0, \widetilde{C}_1, y) := \left| \Pr(\mathcal{B}^{\mathcal{A}}(\widetilde{C}_0, \widetilde{C}_1, y) = 1 | b = 0) - \Pr(\mathcal{B}^{\mathcal{A}}(\widetilde{C}_0, \widetilde{C}_1, y) = 1 | b = 1) \right|.$$

First, when $b = 0$ and hence $z = \widetilde{C}_0(r)$ with $\|\widehat{D}_0 - \widetilde{D}_0\|_1 \leq 2\mu$, the adversary \mathcal{A} succeeds with probability at least $1 - \theta_{\text{owf}}/2 - 2\mu$ to invert \widetilde{C}_0 , which is equal to the probability that $\mathcal{B}^{\mathcal{A}}$ outputs 1. Second, when $b = 1$ and hence $z = \widetilde{C}_1(r)$, since the supports of \widetilde{C}_0 and \widetilde{C}_1 are distinct, \mathcal{A} never succeeds to find an r' such that $\widetilde{C}_0(r') = \widetilde{C}_1(r)$, i.e., the probability of \mathcal{B} outputting one is zero. This completes the first part.

Case $y \in \Pi_N$: In Theorem 1, we also proved that for every $y \in \Pi_N$, the outcomes of the two circuits $(\widehat{C}_0, \widehat{C}_1)$ is at most $\delta + \gamma$. Therefore, the adversary \mathcal{A} cannot distinguish them with a probability larger than $\delta + \gamma$. The information processing inequality then implies that

$$X(\widehat{C}_0, \widehat{C}_1, y) \leq \delta + \gamma.$$

Conclusion: The quantity $X(\widehat{C}_0, \widehat{C}_1, y)$ diverges for YES and NO instances of y . For our choice of parameters, we know that

$$1 - \theta_{\text{owf}}/2 - 10\mu - (\delta + \gamma) = \theta_{\text{owf}}/2.$$

We denote by $\mathcal{C}^{\mathcal{A}}$ an algorithm that runs \mathcal{B} for $O(1/\theta_{\text{owf}}^2)$ many times, and approximates the quantity above within error less than $\theta_{\text{owf}}/4$. If this value is more than $\delta + \gamma + \theta_{\text{owf}}/4$, then y must be a YES instance, otherwise it is a NO instance. Therefore, we finally obtain an algorithm that solves Π .

Runtime: The runtime of \mathcal{F} can be computed as follows. It samples m instances from \mathcal{U}_{K_a} (or \mathcal{U}_{T_a}), applies a permutation π , and computes R on top of it. Each time, sampling from \mathcal{U}_{K_a} (or \mathcal{U}_{T_a}) takes time $O(dn)$, where $d \leq (m+1)/\gamma$ is the size of \mathcal{U}_{K_a} and n is the size of each element in \mathcal{U}_{K_a} . The permutation can be computed in $O(m)$. Therefore, the total runtime of \mathcal{F} is $O(T + m^2n/\gamma)$.

For the runtime of $\mathcal{C}^{\mathcal{A}}$, note that \mathcal{C} runs \mathcal{B} for $O(1/\theta_{\text{owf}}^2)$ times. Each execution of \mathcal{B} evaluates \widehat{C}_b , queries \mathcal{A} , and performs an equality check. All of this takes $O((T + m^2n/\gamma)/\theta_{\text{owf}}^2)$ with $O(1/\theta_{\text{owf}}^2)$ queries to \mathcal{A} .

□

7 One-Way State Generators from Mildly-Lossy Problems

In the next theorem, we discuss the adaptation to the quantum settings, when λ is relatively large.

Theorem 4. *Let Π be $(T, \mu, f^m, \lambda, \gamma)$ -mildly-lossy with a pure-outcome reduction. Also assume that $\theta_{\text{ows}} := 1 - (\delta(\lambda) + \gamma + 4\sqrt{2\mu}) > 0$ and $\tau_{\text{ows}} := 1 - 2\mu - (\delta(\lambda) + \gamma) > 0$, with $\delta(\lambda)$ as in Definition 20. Then there exists an algorithm $G = (\text{StateGen}, \text{Ver})$ such that **StateGen** runs in time $O(T + m^2 n \gamma^{-1})$ and an oracle algorithm \mathcal{C} , such that for every algorithm \mathcal{A} one and only one of the following statements holds:*

- I. $\mathcal{C}^{\mathcal{A}}$ solves $\Pi \cap \{0, 1\}^n$ in time $O((T + m^2 n \gamma^{-1} + \tau_{\text{ows}}^{-2})\theta_{\text{ows}}^{-2})$ with $O(\theta_{\text{ows}}^{-2})$ classical queries to \mathcal{A} ,
- II. G is a $(1 - \theta_{\text{ows}}/4)$ -OWSG for \mathcal{A} .

Proof. Sample $z \sim \mathcal{U}_{K_a}$ and apply Algorithm 2 up to Line 4 on input z to obtain the two circuits (C_0^*, C_1^*) . Note that the two circuits are mixed; a classical randomness is used to sample \tilde{x} but the algorithm R is a pure quantum circuit. Let κ be the size of the randomness of these circuits. For any $r \in \{0, 1\}^\kappa$ and $b \in \{0, 1\}$, let $C_b^* |r, \mathbf{0}\rangle$ be the pure state obtained by sampling \tilde{x} using r and applying R to $\pi(\tilde{x})$ and a possibly ancilla $|\mathbf{0}\rangle$ with an appropriate size. We show that G , defined as follows:

- **StateGen** (r, b) : output $C_b^* |r, \mathbf{0}\rangle$.
- **Ver** $((r, b), \rho)$: If $\|C_b^* |r, \mathbf{0}\rangle - \rho\|_1 \leq \delta + \gamma$ output 1, otherwise output 0.

is a $(\theta_{\text{ows}}/2)$ -weak one-way state generator.

Assume that there exists an adversary \mathcal{A} that breaks the scheme above with probability more than $1 - \theta_{\text{ows}}/4$. We use \mathcal{A} to construct an algorithm for II. Consider the following oracle algorithm $\mathcal{B}^{\mathcal{A}}$:

- $\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y)$: computes $(\hat{C}_0, \hat{C}_1(y))$ as in Algorithm 2 up to Line 4 on input y . Samples a uniform $r \in \{0, 1\}^\kappa$ and a uniform $b \in \{0, 1\}$, and computes $\rho := \hat{C}_b |r, \mathbf{0}\rangle$. Runs the adversary $(r', b') \leftarrow \mathcal{A}(\rho)$, and computes $\rho' = \hat{C}_{b'} |r', \mathbf{0}\rangle$. If $\|\rho - \rho'\|_1 \leq \delta + \gamma$ it outputs 1, otherwise it outputs 0.

We compute the advantage of \mathcal{B} in distinguishing between YES and NO instances of Π by analyzing the probability $\Pr(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1)$.

Case $y \in \Pi_Y$: We show that for every $y \in \Pi_Y$, we have:

$$\Pr(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1) \leq \frac{1}{2} + 2\sqrt{2\mu}.$$

Instead of proving the inequality directly for the circuits (\hat{C}_0, \hat{C}_1) , we will show it for two similar circuits $(\tilde{C}_0, \tilde{C}_1)$ with disjoint images. Let $\hat{\rho}_0$ and $\hat{\rho}_1$ be respectively the mixed states $\hat{C}_0 |r, \mathbf{0}\rangle$ and $\hat{C}_1 |r, \mathbf{0}\rangle$ when r follows the uniform distribution. For any POVM $\mathcal{M} = \{M_i\}_i$, let us define by $A_{\mathcal{M}}$ the following set:

$$A_{\mathcal{M}} := \{i \mid \text{Tr}(M_i \hat{\rho}_0) \geq \text{Tr}(M_i \hat{\rho}_1)\}.$$

In Theorem 1, we showed that for every $y \in \Pi_Y$, the statistical distance between $\hat{\rho}_0$ and $\hat{\rho}_1$ is at least $1 - 2\mu$. Moreover, we can rewrite the trace distance in terms of the POVMs as

$$\begin{aligned} \|\hat{\rho}_0 - \hat{\rho}_1\|_1 &= \max_{\{M_i\}_i} \frac{1}{2} \sum_i |\text{Tr}(M_i \hat{\rho}_0) - \text{Tr}(M_i \hat{\rho}_1)| \\ &= \max_{\{M_i\}_i} \frac{1}{2} \left[\sum_{i \in A_{\mathcal{M}}} (\text{Tr}(M_i \hat{\rho}_0) - \text{Tr}(M_i \hat{\rho}_1)) + \sum_{i \in A_{\mathcal{M}}^c} (\text{Tr}(M_i \hat{\rho}_1) - \text{Tr}(M_i \hat{\rho}_0)) \right] \\ &= \max_{\{M_i\}_i} \left\{ \sum_{i \in A_{\mathcal{M}}} \text{Tr}(M_i \hat{\rho}_0) + \sum_{i \in A_{\mathcal{M}}^c} \text{Tr}(M_i \hat{\rho}_1) - 1 \right\}. \end{aligned}$$

It follows that there exists a particular POVM \mathcal{M} , such that if we define the projections of \hat{C}_0 and \hat{C}_1 onto $A_{\mathcal{M}}$ and $A_{\mathcal{M}}^c$ by \tilde{C}_0 and \tilde{C}_1 respectively, i.e.,

$$\tilde{C}_0 = \sum_{i \in A_{\mathcal{M}}} M_i \hat{C}_0, \quad \text{and} \quad \tilde{C}_1 = \sum_{i \in A_{\mathcal{M}}^c} M_i \hat{C}_1,$$

we have $\text{Tr}(\tilde{\rho}_0) + \text{Tr}(\tilde{\rho}_1) \geq 2 - 2\mu$, where $\tilde{\rho}_b$ is the mixed state $\tilde{C}_b |r, \mathbf{0}\rangle$ and r is uniform. Therefore

$$(\text{Tr}(\tilde{\rho}_0) \geq 1 - \mu) \wedge (\text{Tr}(\tilde{\rho}_1) \geq 1 - 2\mu), \quad \text{or} \quad (\text{Tr}(\tilde{\rho}_0) \geq 1 - 2\mu) \wedge (\text{Tr}(\tilde{\rho}_1) \geq 1 - \mu).$$

By the Gentle Measurement Lemma 6, for either of cases above, we have

$$\|\hat{\rho}_0 - \tilde{\rho}_0\|_1 \leq \sqrt{2\mu}, \quad \text{and} \quad \|\hat{\rho}_1 - \tilde{\rho}_1\|_1 \leq \sqrt{2\mu}. \quad (12)$$

Pretend that \mathcal{A} also tried to distinguish between for \hat{C}_b and \tilde{C}_b for $b \in \{0, 1\}$, and consider the following sequence of games that modifies $\mathcal{B}^{\mathcal{A}}$.

Game \mathcal{G}_1 : In this game \mathcal{B} behaves originally as above.

Game \mathcal{G}_2 : In this game \mathcal{B} replaces \hat{C}_0 with \tilde{C}_0 . Note that \mathcal{A} can distinguish this modification with probability at most $\sqrt{2\mu}$ according to Equation (12). It follows that

$$\begin{aligned} \Pr(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1) &\leq \left| \Pr(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1) - \Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \hat{C}_1, y) = 1) \right| \\ &\quad + \Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \hat{C}_1, y) = 1) \\ &\leq \sqrt{2\mu} + \Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \hat{C}_1, y) = 1). \end{aligned}$$

Game \mathcal{G}_3 : In this game, \mathcal{B} replaces \hat{C}_1 with \tilde{C}_1 . Note that \mathcal{A} can identify this modification with probability at most $\sqrt{2\mu}$. We obtain

$$\begin{aligned} \Pr(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1) &\leq \left| \Pr(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1) - \Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \hat{C}_1, y) = 1) \right| \\ &\quad + \left| \Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \hat{C}_1, y) = 1) - \Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \tilde{C}_1, y) = 1) \right| \\ &\quad + \Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \tilde{C}_1, y) = 1) \\ &\leq 2\sqrt{2\mu} + \Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \tilde{C}_1, y) = 1) \end{aligned}$$

Now, note that the projection onto the supports of \tilde{C}_0 and \tilde{C}_1 are orthogonal to each other. Therefore, the adversary never succeeds when the bit b (chosen by \mathcal{B}) is equal to 1; there exists no r' such that $\|\tilde{C}_0 |r, \mathbf{0}\rangle - \tilde{C}_1 |r', \mathbf{0}\rangle\|_1 \leq \delta + \gamma$. So

$$\Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \tilde{C}_1, y) = 1) = \frac{1}{2} \left(\Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \tilde{C}_1, y) = 1 | b = 0) + \Pr(\mathcal{B}^{\mathcal{A}}(\tilde{C}_0, \tilde{C}_1, y) = 1 | b = 1) \right) \leq \frac{1}{2}.$$

Case $y \in \Pi_N$: By Lemma 9, the trace distance of the outcomes of \hat{C}_1 and C_1^* is at most $\delta + \gamma$. Moreover, \hat{C}_0 is exactly the same as C_0^* . Therefore, if the bit b , chosen by \mathcal{B} is equal to 0, then \mathcal{A} succeeds with probability at least $1 - \theta_{\text{ows}}/4$, and if $b = 1$, it succeeds with probability $1 - \theta_{\text{ows}}/4 - (\delta + \gamma)$. In total, we obtain

$$\Pr(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1) \geq \frac{1}{2} \left(1 - \frac{\theta_{\text{ows}}}{4} \right) + \frac{1}{2} \left(1 - \frac{\theta_{\text{ows}}}{4} - (\delta + \gamma) \right) = 1 - \frac{\theta_{\text{ows}}}{4} - \frac{(\delta + \gamma)}{2}.$$

Conclusion: We showed that the quantity of $\Pr(\mathcal{B}^{\mathcal{A}}(\hat{C}_0, \hat{C}_1, y) = 1)$ diverges for YES and NO instances of y . For our choice of parameters, we have

$$\begin{aligned} 1 - \frac{\theta_{\text{ows}}}{4} - \frac{(\delta + \gamma)}{2} - \left(\frac{1}{2} + 2\sqrt{2\mu} \right) &= \frac{1 - (\delta + \gamma + 4\sqrt{2\mu})}{2} - \frac{\theta_{\text{ows}}}{4} \\ &= \frac{\theta_{\text{ows}}}{4}. \end{aligned}$$

Let \mathcal{C} be an algorithm that runs \mathcal{B} for $O(1/\theta_{\text{ows}}^2)$ many times, and approximates the quantity above within error less than $\theta_{\text{ows}}/4$. If this value is more than $1 - \theta_{\text{ows}}/4 - (\delta + \gamma)/2$, then y must be a NO instance, otherwise it is a YES instance. Therefore, we finally obtain a algorithm that solves Π . Note that \mathcal{B} verifies whether $\|\hat{C}_b |r, \mathbf{0}\rangle - \hat{C}_{b'} |r', \mathbf{0}\rangle\|_1$ is smaller than $\delta + \gamma$. Since the reduction R is pure and r, r' are fixed, these states are pure, therefore \mathcal{B} can perform a SWAP test for $O(1/\tau_{\text{ows}}^2)$ number of times on them to approximate their ℓ_1 distance. \square

8 Mild-Lossiness and Instance Randomization

In Section 4 we introduced mildly-lossy problems, promise problems that admit reductions that *lose* some information about the input, and in Section 6 and 7 we constructed cryptography primitives from these. In this section we show that mildly-lossy problems are not uncommon by proving that both worst-case to average-case reductions and randomized encodings imply mild-lossiness, given a classical reduction. An in the final subsection we prove that the former is also true for certain type of quantum reductions.

8.1 Worst-Case to Average-Case Reductions

In this section we analyse the mild-lossiness of worst-case to average-case reductions. Since we discuss mild-lossiness of such reductions, as motivated in Section 4, we focus on worst-case to average-case *f-distinguisher* reductions (Definition 14). In Definition 21, we put forward the definition of *worst-case to distribution f-distinguisher reduction* which can be viewed as a generalization of

worst-case to average-case reductions in the sense that (i) the reduction is oblivious to the target average-case problem (inherited from being f -distinguisher), and (ii) the reduction maps inputs to a distribution that is *not* necessarily efficiently samplable. The latter does not impose any issues in our setting, since we are only discussing mild-lossiness of the reductions, and not the hardness of the problems. We then prove, in Theorem 5, that such reductions are lossy and specify the mild-lossiness parameters. Combined with Theorem 3, this would yield in Corollary 1 that worst-case to average-case reductions can be used to build one-way functions.

Definition 21 (Worst-Case to Distribution f -Distinguisher Reduction). *Let Π be a promise problem, $n \in \mathbb{N}$, and $d \in [0, 1]$. We say that a reduction R is a (T, μ, f^m, d) -worst-case to distribution (WC-DIST) reduction for Π if*

- R is a (μ, f^m) -distinguisher reduction for Π (Definition 14), and
- for all $x \in \Pi \cap \{0, 1\}^n$, $R(x)$ runs in time $T(n)$, and
- there exists a distribution $D = \{D_n\}_{n \in \mathbb{N}}$ over $\{0, 1\}^*$, such that

$$\forall (x_1, \dots, x_m) \in (\Pi \cap \{0, 1\}^n)^m : \frac{1}{2} \|R(x_1, \dots, x_m) - D\|_1 \leq d .$$

The upper bound d is called the distance of the reduction.

If there exist two distributions D_Y and D_N over $\{0, 1\}^*$ such that for inputs $x \in \Pi_Y$ the distribution D_Y approximates $R(x)$ up to error d , and for inputs $x \in \Pi_N$ the distribution D_N approximates $R(x)$ up to error d , we say that the reduction R is a (T, μ, f^m, d) -worst-case to distribution splitting-reduction for Π .

Theorem 5 (Mild-Lossiness of WC-DIST f -Distinguisher Classical Reductions). *Let $\Pi = \Pi_Y \cup \Pi_N$ for two disjoint sets $\Pi_Y, \Pi_N \subset \{0, 1\}^*$. If there exists a (T, μ, f^m, d) -WC-DIST classical splitting-reduction R for Π (Definition 21), such that f is a non-constant permutation-invariant function, then for any $\gamma > 0$, Π is $(T, \mu, f^m, \lambda, \gamma)$ -mildly-lossy, where*

$$\lambda = \max \left\{ 1, 13 + \log \left(\frac{mnd^2}{\gamma^3} \right) \right\} .$$

Proof. The proof consists of showing that the reduction R satisfies Definition 19. Let $\gamma > 0$. We show that for all pairwise independent $2^9 mn/\gamma^3$ -uniform distributions X_n over n -bit strings,

$$I(X_n; R(X_n)) \leq \max \left\{ 1, 13 + \log \left(\frac{mnd^2}{\gamma^3} \right) \right\} .$$

Dropping the subscript n for simplicity and writing $p_X(y) := \Pr(X = y)$, we first rewrite the mutual information in terms of Kullback-Leibler divergence.

$$I(X; R(X)) = \sum_{y \in \text{Supp}(R)} p_{R(X)}(y) \cdot D_{KL}(p_{X|R(X)=y} \parallel p_X) . \quad (13)$$

From a reverse Pinsker inequality due to [Sas15], the KL divergence of two distributions decreases as their trace distance does, in particular

$$D_{KL}(p_{X|R(X)=y} \parallel p_X) \leq \log \left(1 + \frac{2 \cdot \Delta(X_{|R(X)=y}, X)^2}{\alpha_X} \right)$$

where $\alpha_X = \min_x p_X(x) > 0$. If $\Delta(X_{|R(X)=y}, X) = 0$, then $I(X; R(X)) = 0$.¹ Otherwise, since for any value $a \in (0, 1]$, we have that $\log(1 + a) \leq \max\{1, 1 + \log(a)\}$, we can write

$$D_{KL}(p_{X|R(X)=y} \parallel p_X) \leq \max\{1, 2 + 2\log(\Delta(X_{|R(X)=y}, X)) - \log(\alpha_X)\} ,$$

Substituting above in Equation 13 we obtain:

$$I(X; R(X)) \leq \max\{1, 2 - \log(\alpha_X) + 2 \sum_{y \in \text{Supp}(R)} p_{R(X)}(y) \cdot \log(\Delta(X_{|R(X)=y}, X))\} . \quad (14)$$

We split the bound on the right-hand side of the Inequality 14 into two terms.

Bounding term₁ = $-\log(\alpha_X)$: Since X_n is a $2^9 mn / \gamma^3$ -uniform distribution, we have $\alpha_X \geq \gamma^3 / 2^9 mn$. Therefore $-\log(\alpha_X) \leq 9 + \log(mn / \gamma^3)$.

Bounding term₂ = $\sum_{y \in \text{Supp}(R)} p_{R(X)}(y) \cdot \log(\Delta(X_{|R(X)=y}, X))$: Firstly, for any $y \in \text{Supp}(R)$, we have

$$\begin{aligned} \Delta(X_{|R(X)=y}, X) &= \frac{1}{2} \sum_x |\Pr(X = x | R(X) = y) - \Pr(X = x)| \\ &= \frac{1}{2} \sum_x \left| \frac{\Pr(X = x \wedge R(X) = y)}{\Pr(R(X) = y)} - \Pr(X = x) \right| \\ &= \frac{1}{2} \sum_x \frac{1}{\Pr(R(X) = y)} |\Pr(X = x \wedge R(X) = y) - \Pr(X = x) \cdot \Pr(R(X) = y)| \\ &= \frac{1}{\Pr(R(X) = y)} \cdot \Delta((X, R(X) = y), X \cdot (R(X) = y)) . \end{aligned} \quad (15)$$

Rewriting term₂ = $\mathbb{E}_{R(X)} [\log(\Delta(X_{|R(X)=y}, X))]$, we now have to bound

$$\begin{aligned} \text{term}_2 &= \mathbb{E}_{R(X)} [\log \Delta(X_{|R(X)=y}, X)] \\ &\leq \log \mathbb{E}_{R(X)} [\Delta(X_{|R(X)=y}, X)] \quad (\text{by Jensen's inequality}) \\ &= \log \left(\sum_{y \in \text{Supp}(R)} \Pr(R(X) = y) \cdot \Delta(X_{|R(X)=y}, X) \right) \\ &= \log \left(\sum_{y \in \text{Supp}(R)} \Delta((X, R(X) = y), X \cdot (R(X) = y)) \right) \quad (\text{by Equation 15}) . \end{aligned} \quad (16)$$

¹However, this is very unlikely!

Analysing the term inside the logarithm above, we have

$$\begin{aligned}
& \sum_{y \in \text{Supp}(R)} \Delta((X, R(X) = y), X \cdot (R(X) = y)) \\
&= \frac{1}{2} \sum_{y \in \text{Supp}(R)} \sum_{x \in X} |\Pr(R(X) = y | X = x) \cdot \Pr(X = x) - \Pr(R(X) = y) \cdot \Pr(X = x)| \\
&= \frac{1}{2} \sum_{y \in \text{Supp}(R)} \sum_x \Pr(X = x) \cdot |\Pr(R(x) = y) - \Pr(R(X) = y)| \\
&= \sum_x \Pr(X = x) \cdot \Delta(R(x), R(X)) \\
&\leq \max_x \Delta(R(x), R(X)) .
\end{aligned}$$

We therefore have that $\text{term}_2 \leq \max_x \log(\Delta(R(x), R(X)))$. Finally, note that since R is a (T, μ, f^m, d) -WC-DIST reduction, for any $x \in \Pi_Y \cap \{0, 1\}^n$, it holds that $\Delta(R(x), D_{n,Y}) \leq d$. Therefore $\Delta(R(X), D_{n,Y}) \leq d$ for any distribution X over $\Pi_Y \cap \{0, 1\}^n$. We conclude that for any $x \in \Pi_Y \cap \{0, 1\}^n$, $\Delta(R(x), R(X)) \leq 2d$ for any distribution X over $\Pi_Y \cap \{0, 1\}^n$, which yields $\text{term}_2 \leq 1 + \log(d)$. Note that the same argument holds for $x \in \Pi_N \cap \{0, 1\}^n$ and distributions $D_{n,N}$.

Combining upper bounds on term_1 and term_2 , we finish by proving that

$$I(X; R(X)) \leq \max \left\{ 1, 13 + \log \left(\frac{mnd^2}{\gamma^3} \right) \right\} ,$$

for splitting lossy distributions X . □

The following corollary is a direct result of combining Theorems 5 and 3.

Corollary 1 (OWFs from WC-DIST f -Distinguisher Reductions). *Let Π be a promise problem, and assume that there exists a (T, μ, f^m, d) -WC-DIST splitting-reduction for Π . Let $\theta_{\text{owf}} = (1 - 10\mu) - (\delta(\lambda) + \gamma) > 0$, where $\gamma > 0$, $\lambda = \max \{1, 13 + \log(mnd^2/\gamma^3)\}$, and $\delta(\lambda)$ is the function defined in Definition 20. Then there exists an algorithm F that runs in time $O(T + m^2 n \gamma^{-1})$ and an oracle algorithm \mathcal{C} , such that for any algorithm \mathcal{A} one and only one of the following holds:*

- I. $\mathcal{C}^{\mathcal{A}}$ solves $\Pi \cap \{0, 1\}^n$ in time $O((T + m^2 n \gamma^{-1}) \theta_{\text{owf}}^{-2})$ with $O(\theta_{\text{owf}}^{-2})$ queries to \mathcal{A} ,
- II. F is a $(1 - \theta_{\text{owf}}/2)$ -OWF for \mathcal{A} .

WC-DIST Turing Reductions

All reductions in the rest of the work until Section 9.1 are classical. In this part, we give an adapted version of the worst-case to distribution reduction (Definition 21) to the case of non-adaptive randomized Turing reductions.

Definition 21 covers the notion of worst-case to average-case *Karp* reductions, that is the type of most cryptographic reductions. However, in order to discuss the mild-lossiness of WC-DIST Turing reductions, we have to slightly refine this definition; Recall from Section 4 that a non-adaptive randomized Turing reduction from Π to Σ , maps an input x to (y_1, \dots, y_k) , where each y_i is an

instance of Σ , as well as a Boolean circuit C . Since C depends on x , it can carry some information about the input and affect the mild-lossiness. On the other hand, the requirement of Definition 21 requires analysing the joint distribution of $((y_1, \dots, y_k), C)$ that might be tedious. We therefore relax the above definition to this case and discuss the mild-lossiness of randomized Turing reductions in this relaxed setting.

Definition 22 (WC-DIST Non-Adaptive Randomized Turing f -Reductions). *Let Π be a promise problem. We say that R_{Turing} is a (T, μ, f^m, d, h) -worst-case to distribution (WC-DIST) non-adaptive randomized Turing reduction for Π , if*

- R_{Turing} is a non-adaptive (f^m, μ) -Turing reduction from Π to some promise or search problem Σ (per Definition 16), and
- for all $x \in \Pi \cap \{0, 1\}^n$, $R_{\text{Turing}}(x)$ runs in time $T(n)$, and
- there exists a distribution $D = \{D_n\}_{n \in \mathbb{N}}$ over $\{0, 1\}^*$, such that:

$$\forall x \in \Pi \cap \{0, 1\}^n : \Delta((y_1, \dots, y_k), D_n) \leq d ,$$

where $((y_1, \dots, y_k), C) \leftarrow R_{\text{Turing}}(x)$, and

- for all $2^9 n / \gamma^3$ -uniform distributions X over n -bit strings:

$$I((X, Y_1, \dots, Y_k); C) \leq h,$$

where $((Y_1, \dots, Y_k), C) \leftarrow R_{\text{Turing}}(X)$.

We now state the following lemma, on the mild-lossiness of worst-case to distribution Turing reductions.

Lemma 17 (Mild-Lossiness of WC-DIST Non-Adaptive Randomized Turing Reductions). *Let Π be a promise problem. If there exists a (T, μ, d, h) -WC-DIST non-adaptive randomized Turing reduction R_{Turing} for Π (per Definition 22), then for any $\gamma > 0$, Π is $(T, \mu, \text{id}, \lambda, \gamma)$ -mildly-lossy, where $\text{id} : x \mapsto x$ is the identity function and $\lambda = \max\{1 + h, 13 + h + \log(nd^2/\gamma^3)\}$.*

Proof. Similarly to the proof of Theorem 5, we show that for any $\gamma > 0$, the reduction R_{Turing} is λ -lossy for all pairwise independent $2^9 n / \gamma^3$ -uniform distributions over n -bit inputs, where $\lambda = \max\{1 + h, 13 + h + \log(nd^2/\gamma^3)\}$. In other words,

$$I(X_n; R_{\text{Turing}}(X_n)) \leq \max \left\{ 1 + h, 13 + h + \log \left(\frac{nd^2}{\gamma^3} \right) \right\} ,$$

for all $n \in \mathbb{N}$ and $2^9 n / \gamma^3$ -uniform distributions X_n over n -bit strings.

For any distribution X_n let $((Y_1, \dots, Y_k), C)$ denote the distribution of $R_{\text{Turing}}(X_n)$. Dropping the subscript n for simplicity, we have

$$\begin{aligned} I(X; ((Y_1, \dots, Y_k), C)) &\leq I(X; (Y_1, \dots, Y_k)) + I((X, (Y_1, \dots, Y_k)); C) \\ &\leq I(X; (Y_1, \dots, Y_k)) + h, \end{aligned}$$

where we used the inequality $I((X, (Y_1, \dots, Y_k)); C) \leq h$ imposed by the conditions. The rest of the proof is similar to that of Theorem 5 and consists of using the condition $\Delta((y_1, \dots, y_k), D_n) \leq d$ to derive $I(X; (Y_1, \dots, Y_k)) \leq \max \left\{ 1, 13 + \log \left(\frac{mnd^2}{\gamma^3} \right) \right\}$. It therefore concludes that

$$I(X; R_{\text{Turing}}(X)) \leq \max \left\{ 1 + h, 13 + h + \log \left(\frac{mnd^2}{\gamma^3} \right) \right\} .$$

□

Corollary 2 (OWFs from WC-DIST Non-Adaptive Randomized Turing Reductions).

Let Π be a promise problem, and assume that there exists a (T, μ, d, h) -WC-DIST Turing reduction for Π . Let $\theta_{\text{owf}} = (1 - 10\mu) - (\delta(\lambda) + \gamma) > 0$, where $\gamma > 0$ and λ is defined in Lemma 17. Then there exists an algorithm F that runs in time $O(T + n\gamma^{-1})$ and an oracle algorithm \mathcal{C} , such that for any algorithm \mathcal{A} one and only one of the following holds:

- I. $\mathcal{C}^{\mathcal{A}}$ solves $\Pi \cap \{0, 1\}^n$ in time $O((T + n\gamma^{-1})\theta_{\text{owf}}^{-2})$ with $O(\theta_{\text{owf}}^{-2})$ queries to \mathcal{A} ,
- II. F is a $(1 - \theta_{\text{owf}}/2)$ -OWF for \mathcal{A} .

8.2 Randomized Encodings

We now discuss the mild-lossiness of *randomized encodings* [IK00, AIK06, App17]. In Lemma 18, we show that a randomized encoding of a Boolean function is in fact a worst-case to distribution reductions (Definition 21). Hence, we conclude the mild-lossiness of randomized encodings and their utility in building one-way functions in Corollary 4.

We first recall the definition of randomized encodings.

Definition 23 (Randomized Encoding (Adapted from [AIK06])). Let $\mu, d \in [0, 1]$ and let $F : \{0, 1\}^* \rightarrow \{0, 1\}^*$ be a function. We say that a function $E : \{0, 1\}^* \rightarrow \{0, 1\}^*$ is a (T, μ, d) -randomized encoding of F , if

- for all $x \in \{0, 1\}^n$, $E(x)$ can be computed in time $T(n)$, and
- (μ -**correctness**) there exists an algorithm Dec such that for all $x \in \{0, 1\}^n$:

$$\Pr[\text{Dec}(E(x)) \neq F(x)] \leq \mu ,$$

and

- (d -**privacy**) there exists an algorithm Sim such that for all $x \in \{0, 1\}^n$:

$$\Delta(\text{Sim}(F(x)), E(x)) \leq d .$$

Lemma 18. Let $E : \{0, 1\}^* \rightarrow \{0, 1\}^*$ be a (T, μ, d) -randomized encoding for a Boolean function $F : \{0, 1\}^* \rightarrow \{0, 1\}$. Then E is a (T, μ, id, d) -worst-case to distribution splitting-reduction for Π , where $\Pi = \Pi_Y \cup \Pi_N$ is a promise problem defined as $\Pi_Y = \{x \mid F(x) = 1\}$, and $\Pi_N = \{x \mid F(x) = 0\}$, and $id : x \mapsto x$ is the identity function.

Proof. We start by showing that $E(\cdot, \mathcal{U}_m)$ is a (μ, id) -reduction for Π as in Definition 15, which by definition implies that it is a (μ, id) -distinguisher reduction. Let $x, x' \in \Pi \cap \{0, 1\}^*$ such that

$\chi_\Pi(x) \neq \chi_\Pi(x')$, i.e. without loss of generality we can assume that $F(x) = 1$ and $F(x') = 0$. By μ -correctness of the randomized encoding E , there is a distinguisher Dec such that

$$\begin{aligned} & |\Pr(\text{Dec}(E(x)) = 1) - \Pr(\text{Dec}(E(x')) = 1)| \\ &= |\Pr(\text{Dec}(E(x)) = F(x)) - \Pr(\text{Dec}(E(x')) \neq F(x'))| \\ &\geq (1 - \mu) - \mu. \end{aligned}$$

For $x \in \Pi_Y \cap \{0, 1\}^*$, we have $F(x) = 1$, thus $\text{Sim}(1) = \text{Sim}(F(x))$ is a distribution over the YES instances, by a similar argument $\text{Sim}(0)$ is a distribution over the NO instances. By d -secrecy of the randomized encoding, for every $x \in \Pi_Y \cap \{0, 1\}^*$, we have that

$$\frac{1}{2} \|E(x) - \text{Sim}(1)\|_1 \leq d,$$

and the same approximation holds for $E(x)$ with instances $x \in \Pi_N \cap \{0, 1\}^*$ with respect to $\text{Sim}(0)$, leading to the desired result. \square

Corollary 3 (Mild-Lossiness of Randomized Encodings). *If there exists a (T, μ, d) -randomized encoding E for a promise problem Π , then for any $\gamma > 0$, Π is $(T, \mu, \text{id}, \lambda, \gamma)$ -mildly-lossy, where $\lambda = \max\{1, 13 + \log(nd^2/\gamma^3)\}$, and $\text{id} : x \mapsto x$ is the identity function.*

Corollary 4 (OWFs from Randomized Encodings). *Let Π be a promise problem, and assume that there exists a (T, μ, d) -randomized encoding for Π . Let $\theta_{\text{owf}} = (1 - 10\mu) - (\delta(\lambda) + \gamma) > 0$, where $\gamma > 0$ and λ is defined in Corollary 3. Then there exists an algorithm F that runs in time $O(T + n\gamma^{-1})$ and an oracle algorithm \mathcal{C} , such that for any algorithm \mathcal{A} one and only one of the following holds:*

- I. $\mathcal{C}^{\mathcal{A}}$ solves $\Pi \cap \{0, 1\}^n$ in time $O((T + n\gamma^{-1})\theta_{\text{owf}}^{-2})$ with $O(\theta_{\text{owf}}^{-2})$ queries to \mathcal{A} ,
- II. F is a $(1 - \theta_{\text{owf}}/2)$ -OWF for \mathcal{A} .

8.3 Quantum Worst-Case to Average-Case Reductions

In this section we show that a worst-case to average-case quantum reduction also implies mild-lossiness, and therefore OWSGs and EFIs. However, since a quantum reverse Pinsker inequality is not known in its most general form, we include here two independent assumptions on the quantum worst-case to average-case reductions that imply quantum cryptography.

Theorem 6 (Mild-Lossiness of WC-DIST f -Distinguisher Quantum Reductions). *Let $\Pi = \Pi_Y \cup \Pi_N$ for two disjoint sets $\Pi_Y, \Pi_N \subset \{0, 1\}^*$. If there exists a (T, μ, f^m, d) -WC-DIST quantum reduction R for Π , such that f is a non-constant permutation-invariant function, then for any $\gamma > 0$, we have*

1. *If the minimum eigenvalue of the reduction is uniformly bounded from below for every pairwise independent $2^{9mn}/\gamma^3$ -uniform distribution X , i.e there exists a constant $\beta > 0$ such that $\lambda_{\min}(R(X)) > \beta$, then Π is $(T, \mu, f^m, \lambda, \gamma)$ -mildly-lossy, where*

$$\lambda = (\beta + 2d) \log\left(1 + \frac{2d}{\beta}\right).$$

2. If instead the dimension of the image space is upper bounded for every pairwise independent $2^9 mn/\gamma^3$ -uniform distribution X , i.e. there exists a constant $d_R \in \mathbb{N}$ such that $\dim(\text{Im}(R(X))) \leq d_R$, then Π is $(T, \mu, f^m, \lambda, \gamma)$ -mildly-lossy, where

$$\lambda = 4d \log d_R + h(2d) .$$

Proof. *Case 1:* $\lambda_{\min}(R(X)) > \beta$. Let us denote by $\rho_{X,R(X)}$ (or simply by ρ) the joint system of the classical-quantum state after the reduction R is applied to a pairwise independent $2^9 mn/\gamma^3$ -uniform distribution X_n over n -bit strings, where we drop the subscript n simplicity, see Equation 5. We denote the subsystems of $\rho_{X,R(X)}$ by ρ_X and $\rho_{R(X)}$. Note that since $\rho_{X,R(X)}$ is a classical-quantum system, so is $\rho_X \otimes \rho_{R(X)}$. We can rewrite the mutual information in terms of the relative entropy, which by Equation 6 for classical-quantum systems takes a simple form

$$I(X; R(X))_\rho = D(\rho_{X,R(X)} \parallel \rho_X \otimes \rho_{R(X)}) = \sum_x \Pr(X = x) D_{KL}(\rho_{R(X)|X=x} \parallel \rho_{R(X)}) .$$

Note that we drop the classical term from the previous equation because both states have the same classical distribution. By Lemma 3, if the minimum eigenvalue of the reduction is uniformly bounded from below by a constant β , i.e. $\lambda_{\min}(R(X)) > \beta$, then we have a reserve Pinsker-like inequality

$$D(\rho_{R(X)|X=x} \parallel \rho_{R(X)}) \leq \left(\beta + \frac{1}{2} \|\rho_{R(X)|X=x} - \rho_{R(X)}\|_1 \right) \log \left(1 + \frac{1}{2\beta} \|\rho_{R(X)|X=x} - \rho_{R(X)}\|_1 \right) .$$

Finally, note that since R is a (T, μ, f^m, d) -WC-DIST reduction, there exists a distribution D_n such that for any $x \in \Pi \cap \{0, 1\}^n$, it holds that $\frac{1}{2} \|\rho_{R(X)|X=x} - D_n\|_1 \leq d$, thus $\frac{1}{2} \|\rho_{R(X_n)} - D_n\|_1 \leq d$. By the triangle inequality, we have $\frac{1}{2} \|\rho_{R(X)|X=x} - \rho_{R(X)}\|_1 \leq 2d$. We conclude that

$$I(X; R(X))_\rho \leq \sum_x \Pr(X = x) (\beta + 2d) \log \left(1 + \frac{2d}{\beta} \right) = (\beta + 2d) \log \left(1 + \frac{2d}{\beta} \right) .$$

Case 2: $\dim(\text{Im}(R(X))) \leq d_R$. We can find an alternative bound using the quantum conditional entropy. Let us denote by ω the product state $\omega_{X,R(X)} := \rho_X \otimes \rho_{R(X)}$, since the mutual information between subsystems of product states are zero, we have

$$\begin{aligned} I(X; R(X))_\rho &= |I(X; R(X))_\rho - I(X; R(X))_\omega| \\ &= |S(\rho_X) - S(X|R(X))_\rho - S(\omega_X) + S(X|R(X))_\omega| \\ &= |S(X|R(X))_\rho - S(X|R(X))_\omega| \\ &\leq 2 \text{Tr}(\rho, \omega) \log \dim(H_A) + h(\text{Tr}(\rho, \omega)) , \end{aligned}$$

where in the last inequality we used Theorem 5. We can bound the trace distance between ρ and ω by the worst-case indistinguishability of the reduction R . Indeed, note that ρ and ω are classical-quantum states with the same classical distribution, thus

$$\begin{aligned} \|\rho_{X,R(X)} - \rho_X \otimes \rho_{R(X)}\|_1 &= \sum_x \Pr(X = x) \|\rho_{R(X)|X=x} - \rho_{R(X)}\|_1 \\ &\leq \sum_x \Pr(X = x) 2d = 2d . \end{aligned}$$

Since the binary entropy function is increasing on $[0, 1/2]$, we can conclude that for $d < 1/4$,

$$I(X; R(X))_\rho \leq 4d \log d_R + h(2d) .$$

□

The following corollaries stating conditions for the existence of OWSG are a direct result of combining Theorems 6 and 4, we split the two conditions on the quantum reduction for clarity.

Corollary 5. *Let Π be a promise problem, and assume that there exists a (T, μ, f^m, d) -WC-DIST reduction for Π . Let $\beta > 0$ be such that $\lambda_{\min}(R(X)) > \beta$ for every pairwise independent $2^9 mn/\gamma^3$ -uniform distribution X . Let $\theta_{\text{ows}} := 1 - (\delta(\lambda) + \gamma + 4\sqrt{2\mu}) > 0$ and $\tau_{\text{ows}} := 1 - 2\mu - (\delta + \gamma) > 0$, where $\gamma > 0$, $\lambda = (\beta + 2d) \log(1 + 2d/\beta)$, and $\delta(\lambda)$ is the function defined in Definition 20. Then there exists an algorithm $G = (\text{StateGen}, \text{Ver})$ such that **StateGen** runs in time $O(T + m^2 n \gamma^{-1})$ and an oracle algorithm \mathcal{C} , such that for every algorithm \mathcal{A} one and only one of the following statements holds:*

- I. $\mathcal{C}^{\mathcal{A}}$ solves $\Pi \cap \{0, 1\}^n$ in time $O((T + m^2 n \gamma^{-1} + \tau_{\text{ows}}^{-2}) \theta_{\text{ows}}^{-2})$ with $O(\theta_{\text{ows}}^{-2})$ classical queries to \mathcal{A} ,
- II. G is a $(1 - \theta_{\text{ows}}/4)$ -OWSG for \mathcal{A} .

Corollary 6. *Let Π be a promise problem, and assume that there exists a (T, μ, f^m, d) -WC-DIST reduction for Π , with $d < 1/4$. Let $d_R \in \mathbb{N}$ be such that $\dim(\text{Im}(R(X))) \leq d_R$ for every pairwise independent $2^9 mn/\gamma^3$ -uniform distribution X . Let $\theta_{\text{ows}} := 1 - (\delta(\lambda) + \gamma + 4\sqrt{2\mu}) > 0$ and $\tau_{\text{ows}} := 1 - 2\mu - (\delta + \gamma) > 0$, where $\gamma > 0$, $\lambda = 4d \log d_R + h(2d)$, and $\delta(\lambda)$ is the function defined in Definition 20. Then there exists an algorithm $G = (\text{StateGen}, \text{Ver})$ such that **StateGen** runs in time $O(T + m^2 n \gamma^{-1})$ and an oracle algorithm \mathcal{C} , such that for every algorithm \mathcal{A} one and only one of the following statements holds:*

- I. $\mathcal{C}^{\mathcal{A}}$ solves $\Pi \cap \{0, 1\}^n$ in time $O((T + m^2 n \gamma^{-1} + \tau_{\text{ows}}^{-2}) \theta_{\text{ows}}^{-2})$ with $O(\theta_{\text{ows}}^{-2})$ classical queries to \mathcal{A} ,
- II. G is a $(1 - \theta_{\text{ows}}/4)$ -OWSG for \mathcal{A} .

Remark 5. We can also instantiate Theorem 6 with the construction of EFIs in Theorem 2 to obtain $(1 - 2\mu, 1 - 2\mu - \theta_{\text{efi}}/2)$ -EFIs from WC-DIST f -Distinguisher Quantum Reductions with the same two possible conditions on the parameter λ from $\theta_{\text{efi}} := (1 - 2\mu) - 3(\delta(\lambda) + \gamma)$.

9 Applications: Hardness vs One-Wayness

In the previous sections, we analysed the conditions under which a mildly-lossy reduction or a WC-DIST reduction of Π implies one-way functions under the hardness of Π . In this section, we discuss the concrete parameters. Except in Section 9.1, all statements are subject to classical algorithms.

Let us discuss the implications of generic mildly-lossy reductions. We explicit some particular conditions under which one-way functions exist.

Lemma 19. *Let $n \in \mathbb{N}$, $\lambda : \mathbb{N} \rightarrow \mathbb{R}^+$. Let Π be a $(T, \mu, f^m, \lambda, \gamma)$ -mildly-lossy for parameters below:*

$$T, m = 2^{O(\lambda + \log n)}, \quad \mu \leq 2^{-\lambda-8}, \quad \gamma = 2^{-\lambda-4}.$$

If Π cannot be solved in time $2^{O(\lambda + \log n)}$, then one-way functions exist.

Proof. For these parameters, we have $\theta_{\text{owf}} := (1 - 10\mu) - (\delta(\lambda) + \gamma) \geq 2^{-\lambda-3}$. Then, in Theorem 3, F has runtime $O(T + m^2 n^{2\lambda})$ and the runtime of the Π -solver is $O(2^{2\lambda}(T + T_{\mathcal{A}}) + m^2 n^{2\lambda}) = 2^{\Theta(\lambda + \log n)}$, for all sufficiently large n . Therefore, if Π is $2^{O(\lambda + \log n)}$ -hard, then no algorithm \mathcal{A} of the same runtime can invert F with probability better than $1 - \theta_{\text{owf}}/2$ since otherwise $\mathcal{C}^{\mathcal{A}}$ must solve Π which breaks its $2^{O(\lambda + \log n)}$ hardness. Set $\kappa := 2^{\lambda + \log n}$ as the security parameter. This means that no algorithm of runtime $\text{poly}(\kappa)$ can invert F (whose runtime is $\text{poly}(\kappa)$) with advantage more than $1 - 1/(16\kappa)$. This implies weak one-way functions, which itself implies one-way functions. \square

As a result, we have the following theorem.

Theorem 7. *Let $n \in \mathbb{N}$, $\lambda : \mathbb{N} \rightarrow \mathbb{R}^+$, and Π be a promise problem that cannot be solved by any algorithm in time $2^{O(\lambda + \log n)}$. If Π has a f^m -distinguisher reduction for some non-constant permutation-invariant f^m , with the following parameters:*

$$\text{it is } m\lambda^\circ \leq m\lambda \text{ mildly-lossy, } T, m = 2^{O(\lambda + \log n)}, \text{ and } \mu \leq 2^{-\lambda-8},$$

then one-way functions exist.

Proof. One can use Lemma 19 and the fact that such a reduction implies that Π is $(T, \mu, f^m, \lambda, \gamma)$ -mildly-lossy for $\gamma = 2^{-\lambda-4}$. \square

Perhaps surprisingly, the non-existence of infinitely-often one-way functions has strong implications. To explicit these implications, we first define a quantitative measure of the hardness of problems as below.

Definition 24 (Exact Hardness of Problems). *For a problem Π , let $\tau_\Pi(n) := \inf_{\tau_i(n) \in \mathcal{T}} \{\tau_i\}$ (the limit is taken point-wise), where \mathcal{T} is the set of family of functions τ_i such that $\Pi \cap \{0, 1\}^n$ can be solved in time $O(2^{\tau_i(n)})$ on all instances with probability $\geq 2/3$.*

Note that always $\tau_\Pi(n) \leq n$. This is because algorithms with an advice of size 2^n (maximum size of the truth table of χ_Π) can solve any instance of size n .

We need following lemma.

Lemma 20. *For a non-constant permutation-invariant function f^m , if an f^m -reduction has an error μ that is within a constant distance from $1/2$, then it must have runtime $\Omega(m)$.*

Proof. Assume that the reduction has runtime $o(m)$. Supposing that reading each input of the reduction takes instant time, the assumption implies that the circuit evaluating the reduction ignores $m - o(m)$ number of inputs. Let \mathcal{I} be the indices of the discarded inputs, and let $p(f)$ be as in Lemma 16. As shown in the same lemma, function f only depends on the number of 1's in its inputs. On each input with $p(f) - 1$ number of 1's (which evaluates to 0), one can flip one of the 0's to 1 and obtain an input that evaluates to 1. However, if the index of this input is in \mathcal{I} , it will be discarded by the reduction. Therefore, on $|\mathcal{I}| = m - o(m)$ number of bit-flips, the reduction errs. Consequently, the error must be at least $(m - o(m))/(2m) = \omega(1)$. \square

The non-existence of one-way functions has implications on mildly-lossy reductions, as follows:

Theorem 8. *If infinitely often one-way functions do not exist, then for any Π and any f -distinguisher reduction for Π with mild-lossiness $\leq m(\tau_\Pi / \log \log n - \log n)$ and $\mu \leq 2^{-\tau_\Pi(n)-8}$, where f^m is a non-constant permutation-invariant function, it holds that $T = 2^{\Omega(\tau_\Pi / \log \log n)}$.*

Proof. Let τ be such that $\tau(n) + \log n = o(\tau_\Pi(n))$, and assume that there exists an infinitely often f^m -distinguisher reduction for Π with parameters

$$\text{mild-lossiness } m\tau^\circ \leq m\tau, \quad T, m = 2^{O(\tau + \log n)}, \quad \text{and } \mu \leq 2^{-\tau_\Pi(n)-8}.$$

Since Π is $\Omega(2^{\tau_\Pi(n)})$ -hard, then no algorithm that runs in time $2^{O(\tau(n) + \log n)}$ can solve it. This is because $\tau(n) + \log n = o(\tau_\Pi(n))$. Therefore, by Theorem 7, infinitely often one-way functions exist. This contradicts the assumption. Therefore, for such a lossiness and error $\mu \leq 2^{-\tau_\Pi(n)-8}$, it must hold that the f^m -distinguisher reduction either runs in time $2^{\omega(\tau(n) + \log n)}$ or $m = 2^{\omega(\tau(n) + \log n)}$, for all sufficiently large n . Note that the latter implies the former by Lemma 20. Hence, we have $T = 2^{\omega(\tau(n) + \log n)}$. This holds for every τ such that $\tau(n) + \log n = o(\tau_\Pi(n))$. We let $\tau = \tau_\Pi / \log \log n - \log n$. Therefore, the runtime must be at least $2^{\Omega(\tau_\Pi / \log n)}$. \square

Remark 6. We note that f^m -compression reductions are special cases of mildly-lossy reductions. More precisely, a mapping that compresses mn bits to $m\lambda$ bits is $m\lambda$ mildly-lossy. Therefore, all the results above immediately apply to f^m -compression reductions.

When the reductions are WC-DIST, we obtain fine-grained one-way functions with a slightly looser range of parameters. We first simplify the conditions of Corollary 1.

Lemma 21. *Let $n \in \mathbb{N}$ and $\gamma, T_A > 0$. Let Π be a promise problem that admits a (T, μ, f^m, d) -WC-DIST reduction (per Definition 21). If $d^2 \leq \gamma^3/mn$ and $\mu, \gamma \leq 10^{-5}$, then there exist a constant $\vartheta < 1$ and an algorithm F that runs in time $O(T + m^2n\gamma^{-1})$, such that if F is not a ϑ -OWF for every T_A -bounded adversary, then $\Pi \cap \{0, 1\}^n$ can be solved in time $O(T_A + T + m^2n\gamma^{-1})$.*

Proof. In Corollary 1, if $d^2 \leq \gamma^3/mn$, then $\lambda \leq 13$ and $\delta(\lambda) \leq 1 - 2^{-15}$. Since $\mu, \gamma \leq 10^{-5}$, we have $\theta_{\text{owf}} \geq (1 - 10\mu) - (\delta(\lambda) + \gamma) = 2^{-15} - 10^{-4} - 10^{-5}$. Thus $\theta_{\text{owf}} = \Omega(1)$. Let $\vartheta := 1 - \theta_{\text{owf}}/2$. Corollary 1 implies that there exists a function F that runs in time $O(T + m^2n\gamma^{-1})$ such that if it is not a ϑ -OWF for an algorithm \mathcal{A} , then $\Pi \cap \{0, 1\}^n$ can be solved in time $O((T_A + T + m^2n\gamma^{-1})\theta_{\text{owf}}^{-2})$, where T_A is the runtime of \mathcal{A} . The statement follows by noting that $\theta_{\text{owf}}^{-2} = O(1)$. \square

The following lemma will be used in the proof.

Lemma 22. *For a function $g : \mathbb{N} \rightarrow \mathbb{R}^+$, if $g(n) > 2^{c\tau(n)}$ for every constant $c < 1$, then $g = \Omega(2^\tau)$.*

Using the above lemmas, one can leverage the hardness of Π to build fine-grained one-way functions.

Theorem 9. *Let $n \in \mathbb{N}$, $\tau : \mathbb{N} \rightarrow \mathbb{R}^+$, and Π be a promise problem that cannot be solved by any algorithm in time $O(2^{\tau(n)})$. For any $\eta > 0$, if Π admits a (T, μ, f^m, d) -WC-DIST reduction for some $\mu \leq 10^{-5}$, $d \leq m^{2.5}n/2^{1.5\tau/(1+\eta)}$, and $T, m = O(2^{\tau/(1+\eta)})$, then there exists a constant $\vartheta < 1$ and a one-way function F , such that no $O(|F|^{1+\eta})$ -time algorithm can invert it with a probability better than ϑ .*

Proof. Set γ^{-1} as $2^{\tau/(1+\eta)}/(m^2n)$. Let \mathcal{A} be an algorithm with runtime $T_A = O(2^\tau)$. By assumption, $\Pi \cap \{0, 1\}^n$ cannot be solved in time $O(T_A + T + m^2n\gamma^{-1}) = O(2^\tau)$. Then, Lemma 21 implies the existence of a constant $\vartheta < 1$ and a function F that runs in time $O(2^{\tau/(1+\eta)})$ but no $O(2^\tau)$ -time algorithm can break it with a probability better than ϑ . This concludes the proof. \square

The above theorem implies the existence of weak fine-grained one-way functions based on the $O(2^\tau)$ -hardness of Π and the fact that it admits an WC-DIST f -distinguisher reduction. Similar to Theorem 8, we obtain an impossibility as below.

Theorem 10. *If infinitely-often weak fine-grained one-way functions do not exist, then for any Π and any $(T, \mu \leq 10^{-5}, f^m, d \leq m^{2.5}n/2^{1.5\tau_\Pi})$ -WC-DIST reduction for Π , where f^m is a non-constant permutation-invariant function, it holds that $T = \Omega(2^{\tau_\Pi(n)})$, for all sufficiently large n .*

Proof. Fix n . For any fixed choice of $\tau(n) < \tau_\Pi(n)$, we have $m^{2.5n}/2^{1.5\tau_\Pi} \leq m^{2.5}n/2^{1.5\tau/(1+\eta)}$ for every $\eta > 0$. Therefore, if for some η , Π admits an infinitely-often $(T, \mu \leq 10^{-5}, f^m, d \leq 2^{-1.5\tau_\Pi})$ -WC-DIST reduction for $T, m = O(2^{\tau(n)/(1+\eta)})$, then by Theorem 9 infinitely-often weak fine-grained one-way functions exist. This contradicts the assumption (note that Π is $O(2^\tau)$ -hard per Definition 24). Therefore, such $\eta > 0$ does not exist. Therefore, any WC-DIST reduction, within the mentioned parameter setting, must satisfy $T = \Omega(2^{\tau(n)})$ or $m = \Omega(2^{\tau(n)})$, for all sufficiently large n , by Lemma 22. Note that the latter implies the former by Lemma 20. Finally, the statement follows by taking the limit $\tau(n) \rightarrow \tau_\Pi(n)$. \square

Intuitively, the above theorem asserts that any randomization algorithm of Π , even it is allowed to have a small constant error, is inherently capable of solving it.

In Theorem 21, if the statement holds for all constants $\eta > 0$, one obtains a weak one-way function. Based on this observation, we immediately obtain the following result:

Theorem 11. *Let $n \in \mathbb{N}$, $\tau : \mathbb{N} \rightarrow \mathbb{R}^+$, and Π be a promise problem that cannot be solved by any algorithm in time $O(2^\tau)$. If Π admits a (T, μ, f^m, d) -WC-DIST reduction for some $\mu \leq 10^{-5}$, $d \leq m^{2.5}n/2^{o(\tau)}$, and $T, m = 2^{o(\tau)}$, then there exists a constant $\vartheta < 1$ and a one-way function F , such that no $|F|^{O(1)}$ -time algorithm can invert it with a probability better than ϑ .*

Proof. As mentioned above, if for all constant $\eta > 0$, Π admits a (T, μ, f^m, d) -WC-DIST reduction for some $\mu \leq 10^{-5}$, $d \leq m^{2.5}n/2^{1.5\tau/(1+\eta)}$, and $T, m = O(2^{\tau/(1+\eta)})$, then there exists a constant $\vartheta < 1$ and a one-way function F , such that no $|F|^{O(1)}$ -time algorithm can invert it with a probability better than ϑ . We note that the parameters in the statement satisfy these conditions. \square

This implies one-way functions using the known hardness amplification techniques [Yao82]. Moreover, similar to above, the non-existence of infinitely-often one-way functions has implications for WC-DIST reductions of problems.

Theorem 12. *If infinitely-often one-way functions do not exist, then for any Π and any $(T, \mu \leq 10^{-5}, f^m, d \leq m^{2.5}n2^{-1.5\tau_\Pi})$ -WC-DIST reduction for Π , where f^m is a non-constant permutation-invariant function, it holds that $T = 2^{\Omega(\tau_\Pi(n))}$.*

Proof. Fix n as the size of the instances. For any fixed choice of $\tau(n) < \tau_\Pi(n)$, and every $\eta > 0$, it holds that $m^{2.5}n2^{-1.5\tau_\Pi} \leq m^{2.5}n2^{-o(\tau)}$. Therefore, if Π admits an infinitely-often $(T, \mu \leq 10^{-5}, f^m, d \leq 2^{-1.5\tau_\Pi})$ -WC-DIST reduction for some $T, m = 2^{o(\tau(n))}$, then by Theorem 11 infinitely-often one-way functions must exist (note that Π is $O(2^\tau)$ -hard per Definition 24), which contradicts the assumption. Therefore, any WC-DIST reduction, within the mentioned parameter setting, must satisfy $T = 2^{\Omega(\tau(n))}$, for all sufficiently large n . One concludes by taking the limit $\tau(n) \rightarrow \tau_\Pi(n)$. \square

Remark 7. All the results above regarding WC-DIST reductions can be adapted to WC-DIST non-adaptive Turing reductions for which the hint h (see Definition 22) is not too large, by putting more restrictions on the error.

Towards One-Way Functions from SAT

Let $s_k := \inf\{c \in \mathbb{R} \mid \text{there exists a } O(2^{cn}) \text{ algorithm for } k\text{SAT}\}$. The Exponential Time Hypothesis (ETH) asserts that $s_3 > 0$, namely, 3SAT does not have any subexponential-time algorithm in terms of the number of variables. In fact, Impagliazzo and Paturi [IP01] show that this is equivalent to $\forall k \geq 3 : s_k > 0$. Firstly, we reformulate the assumption in terms of the bit-size of the instance.

Lemma 23. *Let s_k^* be the infimum of all $c \in \mathbb{R}$ such that there exists a $O(2^{cn/\log n})$ -time algorithm for $k\text{SAT}$ where n is the bit-size of the instance. Then under the ETH, we have $0 < s_k^* \leq 2ks_k$ and $\tau_{k\text{SAT}} = s_k^*n/\log n$.*

Proof. For any fixed k , we have $\lceil N/k \rceil \leq M \leq (2N)^k$. On the other hand, the bit-size of an instance is $n := \Theta(M \log N)$. Equivalently, we have $n = \Theta(M \log M)$. Using the standard sparsification Lemma [IPZ98], under the ETH, there is no $2^{o(N+M)}$ -time algorithm, or simply $2^{o(M)}$ -time algorithm, for $k\text{SAT}$. Let g^* be the inverse of the function $M \mapsto M \log M$. Therefore, under the ETH, $k\text{SAT}$ cannot be solved in time $2^{o(g^*(n))}$, where n is now the bit-size of the instance. Note that one can use $g^*(n)$ and M interchangeably. On the other hand, $k\text{SAT}$ can be solved in time $O(2^{s_k N})$ by an exhaustive search, therefore, it can also be solved in time $O(2^{s_k M}) = O(2^{s_k g^*(n)}) \leq O(2^{2ks_k^*n/\log n})$, where we used the fact that $n/\log n \leq g^*(n) \leq 2n/\log n$. Therefore, $0 < s_k^* \leq 2ks_k$. Finally, we have $\tau_{k\text{SAT}} = s_k^*n/\log n$ by Definition 24. \square

We immediately obtain the following corollary by Theorem 9.

Corollary 7. *For any $\eta > 0$, if $k\text{SAT}$ admits a (T, μ, f^m, d) -WC-DIST reduction for some $\mu \leq 10^{-5}$, $d \leq m^{2.5}n/2^{1.5s_k^*n/(\log n(1+\eta))}$, and $T, m = O(2^{s_k^*n/(\log n(1+\eta))})$, then weak η -fine-grained one-way functions exist.*

The following corollary is obtained by Theorem 10 and Lemma 23.

Corollary 8. *Under the ETH, if infinitely often weak fine-grained one-way functions do not exist, then for any non-constant permutation-invariant f^m , any f^m -WC-DIST reduction for $k\text{SAT}$ with error $\mu \leq 10^{-5}$ and distance $d \leq m^{2.5}n/2^{1.5s_k^*n/\log n}$ runs in time $\Omega(2^{s_k^*n/\log n})$.*

Finally, we have the following corollary regarding the existence of one-way functions and hardness of randomization and compression of $k\text{SAT}$.

Corollary 9. *Under the ETH, either infinitely often one-way functions exist, or, for every non-constant permutation-invariant function f^m , the following statements hold:*

- I. *Any f^m -WC-DIST reduction for $k\text{SAT}$ with error $\mu \leq 10^{-5}$ and distance $d \leq m^{2.5}n/2^{1.5s_k^*n/\log n}$ runs in time $2^{\Omega(n/\log n)}$.*
- II. *Any f^m -compression reduction for $k\text{SAT}$ with size-compression from mn bits to $\leq m(s_k^*n/(\log n \cdot \log \log n) - \log n)$ bits with error $\mu \leq 2^{-s_k^*n/\log n - 8}$ runs in time $2^{\Omega(n/(\log n \cdot \log \log n))}$. In particular, for any constant $\varepsilon < 1$ and any $m = \text{poly}(n)$, any perfect f^m -compression that compresses mn bits to mn^ε bits runs in time $2^{\Omega(n/(\log n \cdot \log \log n))}$.*

Proof. By using Lemma 23, Item (I) follows from Theorem 12 and Item (II) from Theorem 8. \square

9.1 Quantum Hardness vs Quantum One-Wayness

In this section, we show that quantum compression reductions imply one-way state generators.

Lemma 24. *Let $n \in \mathbb{N}$, $\lambda : \mathbb{N} \rightarrow \mathbb{R}^+$. Let Π be a $(T, \mu, f^m, \lambda, \gamma)$ -mildly-lossy with a pure-outcome reduction, with the parameters below:*

$$T, m = 2^{O(\lambda + \log n)}, \quad \mu \leq 2^{-2\lambda - 11}, \quad \gamma = 2^{-\lambda - 4}.$$

If Π cannot be solved in time $2^{O(\lambda + \log n)}$ using quantum algorithms, then one-way state generators exist.

Proof. We compute θ_{owf} and τ_{owsg} that are required in Theorem 4. For the given parameters, we have $\theta_{\text{ows}} = 1 - (\delta(\lambda) + \gamma + 4\sqrt{2\mu}) \geq 2^{-\lambda - 3}$ and $\tau_{\text{ows}} \geq 2^{-\lambda - 3}$. The runtime of the construction G in theorem 4 is $O(T + m^2 n 2^\lambda)$ and the runtime of the Π -solver is $O(2^{2\lambda}(T + T_A + 2^{2\lambda}) + m^2 n 2^{3\lambda}) = 2^{\Theta(\lambda + \log n)}$, for all sufficiently large n . Following a similar argument as in Lemma 19, we obtain a weak one-way state generator. One can conclude by noting that weak one-way state generators imply one-way state generators [MY24]. \square

The following theorem is direct.

Theorem 13. *Let $n \in \mathbb{N}$, $\lambda : \mathbb{N} \rightarrow \mathbb{R}^+$, and Π be a promise problem that cannot be solved by any quantum algorithm in time $2^{O(\lambda + \log n)}$. If Π has a quantum f^m -distinguisher pure-outcome reduction for some non-constant permutation-invariant f^m , with the following parameters:*

$$\text{mild-lossiness } m\lambda^\circ \leq m\lambda, \quad T, m = 2^{O(\lambda + \log n)}, \quad \text{and } \mu \leq 2^{-2\lambda - 11},$$

then one-way state generators exist.

Proof. Note that Π is indeed $(T, \mu, f^m, \lambda, \gamma)$ -mildly-lossy for $\gamma = 2^{-\lambda - 4}$, with a quantum reduction. Then the statement follows by Lemma 24. \square

In the beginning of this section, we showed impossibility results for classical mildly-lossy reductions assuming that one-way functions do not exist. Here, we adapt them to one-way state generators. We define a measure of quantum hardness as follows:

Definition 25 (Exact Quantum Hardness of Problems). *For a problem Π , let $\tau_\Pi^Q(n) := \inf_{\tau_i(n) \in \mathcal{Y}} \{\tau_i\}$ (the limit is taken point-wise), where \mathcal{Y} is the set of family of functions τ_i such that $\Pi \cap \{0, 1\}^n$ can be solved by quantum algorithms in time $O(2^{\tau_i(n)})$ on all instances with probability $\geq 2/3$.*

Theorem 14. *If infinitely often one-way state generators do not exist, then for any Π and any quantum f -distinguisher reduction for Π with mild-lossiness $\leq m(\tau_\Pi^Q / \log \log n - \log n)$ and $\mu \leq 2^{-\tau_\Pi^Q(n) - 8}$, where f^m is a non-constant permutation-invariant function, we have $T = 2^{\Omega(\tau_\Pi^Q / \log \log n)}$.*

Proof. Let τ be such that $\tau(n) + \log n = o(\tau_\Pi^Q(n))$. Further, assume that there exists an infinitely often quantum f^m -distinguisher reduction for Π with parameters

$$\text{mild-lossiness } m\tau^\circ \leq m\tau, \quad T, m = 2^{O(\tau + \log n)}, \quad \text{and } \mu \leq 2^{-\tau_\Pi^Q(n) - 8}.$$

Note that we have $\tau(n) + \log n = o(\tau_H^Q(n))$ and Π is $\Omega(2^{\tau_H^Q(n)})$ -hard using quantum algorithms. Therefore, no quantum algorithm that runs in time $2^{O(\tau(n) + \log n)}$ can solve it. By Theorem 13, it follows that infinitely often one-way state generators exist, which contradicts the assumption. Hence any the f^m -distinguisher reduction (within the given parameters) either runs in time $2^{\omega(\tau(n) + \log n)}$ or we have $m = 2^{\omega(\tau(n) + \log n)}$, for all sufficiently large n . Note that the latter implies the former by Lemma 20. Hence, we have $T = 2^{\omega(\tau(n) + \log n)}$. The only condition that we impose on τ is that $\tau(n) + \log n = o(\tau_H^Q(n))$. By letting $\tau = \tau_H^Q / \log \log n - \log n$, we conclude that the runtime must be at least $2^{\Omega(\tau_H^Q / \log n)}$. \square

Remark 8. We note that all the results above immediately apply to quantum f^m -compression reductions since any quantum f^m -compression reduction is quantum mildly-lossy.

Acknowledgments. The authors thank Damien Vergnaud for helpful discussions. This work is part of HQI initiative¹ and is supported by France 2030 under the French National Research Agency award number ANR-22-PNCQ-0002.

¹www.hqi.fr

References

- AE11. Koenraad MR Audenaert and Jens Eisert. Continuity bounds on the quantum relative entropy—ii. *Journal of Mathematical Physics*, 2011.
- AGGM06. Adi Akavia, Oded Goldreich, Shafi Goldwasser, and Dana Moshkovitz. On basing one-way functions on NP-hardness. STOC '06, 2006.
- AIK06. Benny Applebaum, Yuval Ishai, and Eyal Kushilevitz. Cryptography in NC^0 . *SIAM Journal on Computing*, 2006.
- Ajt98. Miklós Ajtai. The shortest vector problem in L2 is NP-hard for randomized reductions (extended abstract). In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98, 1998.
- App17. Benny Applebaum. Garbled circuits as randomized encodings of functions: a primer. Cryptology ePrint Archive, Paper 2017/385, 2017.
- AR16. Benny Applebaum and Pavel Raykov. On the relationship between statistical zero-knowledge and statistical randomized encodings. In *Advances in Cryptology – CRYPTO 2016*, 2016.
- AS18. Divesh Aggarwal and Noah Stephens-Davidowitz. (Gap/S)ETH hardness of SVP. STOC 2018, 2018.
- BB15. Andrej Bogdanov and Christina Brzuska. On basing size-verifiable one-way functions on NP-Hardness. In *Theory of Cryptography Conference*, 2015.
- BBD⁺20. Marshall Ball, Elette Boyle, Akshay Degwekar, Apoorvaa Deshpande, Alon Rosen, Vinod Vaikuntanathan, and Prashant Nalini Vasudevan. Cryptography from information loss. In *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*, 2020.
- BCQ23. Zvika Brakerski, Ran Canetti, and Luowen Qian. On the Computational Hardness Needed for Quantum Cryptography. In *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, Leibniz International Proceedings in Informatics (LIPIcs), 2023.
- BCWd01. Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. *Phys. Rev. Lett.*, 2001.
- BDRV19. Itay Berman, Akshay Degwekar, Ron D. Rothblum, and Prashant Nalini Vasudevan. Statistical difference beyond the polarizing regime. In *Theory of Cryptography: 17th International Conference, TCC 2019, Nuremberg, Germany, December 1–5, 2019, Proceedings, Part II*, 2019.
- BG94. Mihir Bellare and Shafi Goldwasser. The complexity of decision versus search. *SIAM J. Comput.*, 1994.
- BM82. Manuel Blum and Silvio Micali. How to generate cryptographically strong sequences of pseudo random bits. In *23rd Annual Symposium on Foundations of Computer Science (sfcs 1982)*, 1982.
- BQSY24. John Bostanci, Luowen Qian, Nicholas Spooner, and Henry Yuen. An efficient quantum parallel repetition theorem and applications. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, 2024.
- BRSV17. Marshall Ball, Alon Rosen, Manuel Sabin, and Prashant Nalini Vasudevan. Average-case fine-grained hardness. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, 2017.
- BT06. Andrej Bogdanov and Luca Trevisan. On worst-case to average-case reductions for NP problems. *SIAM Journal on Computing*, 2006.
- DH76. W. Diffie and M. Hellman. New directions in cryptography. *IEEE Trans. Inf. Theory*, 1976.
- Dru15. Andrew Drucker. New limits to classical and quantum instance compression. *SIAM Journal on Computing*, 2015.
- FF93. Joan Feigenbaum and Lance Fortnow. Random-self-reducibility of complete sets. *SIAM Journal on Computing*, 1993.
- FS08. Lance Fortnow and Rahul Santhanam. Infeasibility of instance compression and succinct pcps for np. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, 2008.
- GGM86. Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to construct random functions. *J. ACM*, 1986.
- GMW91. Oded Goldreich, Silvio Micali, and Avi Wigderson. Proofs that yield nothing but their validity for all languages in NP have zero-knowledge proof systems. *J. ACM*, 1991.
- Gol90. Oded Goldreich. A note on computational indistinguishability. *Inf. Process. Lett.*, 1990.
- HILL99. Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 1999.

- HN06. Danny Harnik and Moni Naor. On the compressibility of np instances and cryptographic applications. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, 2006.
- IK00. Y. Ishai and E. Kushilevitz. Randomizing polynomials: A new representation with applications to round-efficient secure computation. In *Proceedings 41st Annual Symposium on Foundations of Computer Science*, 2000.
- IP01. Russell Impagliazzo and Ramamohan Paturi. On the complexity of k-sat. *J. Comput. Syst. Sci.*, 2001.
- IPZ98. R. Impagliazzo, R. Paturi, and F. Zane. Which problems have strongly exponential complexity? In *Proceedings 39th Annual Symposium on Foundations of Computer Science*, 1998.
- Lip89. Richard J. Lipton. New directions in testing. In *Distributed Computing And Cryptography*, 1989.
- LY94. Richard J. Lipton and Neal E. Young. Simple strategies for large zero-sum games with applications to complexity theory. In *Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '94, 1994.
- MR07. Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on gaussian measures. *SIAM Journal on Computing*, 2007.
- MY24. Tomoyuki Morimae and Takashi Yamakawa. One-wayness in quantum cryptography. In *19th Conference on the Theory of Quantum Computation, Communication and Cryptography, TQC 2024*, 2024.
- Nao91. Moni Naor. Bit commitment using pseudorandomness. *Journal of Cryptology*, 1991.
- NR06. Moni Naor and Guy N. Rothblum. Learning to impersonate. In *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, 2006.
- Ost91. R. Ostrovsky. One-way functions, hard on average problems, and statistical zero-knowledge proofs. In *[1991] Proceedings of the Sixth Annual Structure in Complexity Theory Conference*, 1991.
- OW93. R. Ostrovsky and A. Wigderson. One-way functions are essential for non-trivial zero-knowledge. In *The 2nd Israel Symposium on Theory and Computing Systems*, 1993.
- Reg09. Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 2009.
- Sas15. Igal Sason. On reverse pinsker inequalities. *CoRR*, 2015.
- SV. Amit Sahai and Salil P. Vadhan. A complete problem for statistical zero knowledge. *J. ACM*.
- vN28. J. v. Neumann. Zur theorie der gesellschaftsspiele. *Mathematische Annalen*, 1928.
- Wat02. J. Watrous. Limits on the power of quantum statistical zero-knowledge. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, 2002.
- Wil13. Mark M Wilde. *Quantum information theory*. Cambridge university press, 2013.
- Win99. A. Winter. Coding theorem and strong converse for quantum channels. *IEEE Transactions on Information Theory*, 1999.
- Yao82. Andrew C. Yao. Theory and application of trapdoor functions. In *23rd Annual Symposium on Foundations of Computer Science (sfcs 1982)*, 1982.