Unmodified Half-Gates is Adaptively Secure

So is Unmodified Three-Halves

Xiaojie Guo[*, †] Kang Yang[*] Xiao Wang[‡] Yu Yu[§, ¶] Zheli Liu[†]

October 6, 2023

Abstract

Adaptive security is a crucial property for garbling schemes in pushing the communication of garbled circuits to an offline phase when the input is unknown. In this paper, we show that the popular half-gates scheme by Zahur et al. (Eurocrypt'15), without any modification, is adaptively secure in the non-programmable random permutation model (npRPM). Since real implementations of selective-secure half-gates are already based on npRPM, our result shows that these implementations are already adaptively secure under the same condition where the selective security is proven. Additionally, we expand our analysis to cover the recent three-halves construction by Rosulek and Roy (Crypto'21); we also discuss some optimizations and separation when considering the programmable random permutation model instead.

1 Introduction

Garbled circuits (GCs) [Yao86, BHR12b] play an important role in constructing many cryptographic protocols, including secure two-party computation (2PC) [Yao86], zero-knowledge proofs [JKO13], identity-based encryption [DG17], etc. In a garbling scheme, a garbler creates a garbled circuit $\widehat{f}$ and a garbled input $\widehat{x}$ for circuit f and input x, respectively. An evaluator, given $\widehat{f}$ and $\widehat{x}$, can compute a garbled version of $f(x)$, which can be further decoded.

Two categories of security have been widely studied in the literature: simulation-based selective security and simulation-based adaptive security.[1] In selective security, the adversary chooses (f, x) and is asked to distinguish between a real-world execution consisting of honestly computed $(\widehat{f}, \widehat{x})$ and an ideal-world execution where the values are simulated using only $f(x)$. In adaptive security, the adversary is given more power to adaptively choose input x based on the garbling $\widehat{f}$ for f: it first chooses f and receives $\widehat{f}$ in the offline phase, and then chooses x and receives $\widehat{x}$ in the online phase. The demand for this security notion originates from the online-offline 2PC, where we want to push the communication of garbled circuits to the phase before the input is known. Clearly, adaptive security is stronger than selective security and is more desirable in practice.

Both security notions have been heavily studied but with very different philosophies and outcomes. For selective security, most prior works focus on how to reduce the computational cost and the size of garbled circuit $\widehat{f}$. An impressive line of GC research [Yao86, BMR90, NPS99, KS08, PSSW09, KMR14, GLNP15, ZRE15, RR21] have reduced the size of $\widehat{f}$ from 8λ to 2λ [ZRE15] or 1.5λ [RR21] bits per `AND` gate, while `XOR` gates are free using the free-XOR technique [KS08], where λ is the security parameter. To minimize the computational cost, almost all implementations instantiate correlation robust hash functions needed for free-XOR [KS08] in the random permutation model (RPM), which is commonly

[*]State Key Laboratory of Cryptology. **Email:** yangk@sklc.org
[†]Nankai University. **Email:** xiaojie.guo@mail.nankai.edu.cn, liuzheli@nankai.edu.cn
[‡]Northwestern University. **Email:** wangxiao@northwestern.edu
[§]Shanghai Jiao Tong University. **Email:** yuyu@yuyu.hk
[¶]Shanghai Qi Zhi Institute.

[1]As we focus on simulation-based security in this paper, we will omit the prefix and use "adaptive/selective security" unless it is not clear from the context.

Schemes	Security	Offline Cost	Online Cost	Security Loss	Assumption
[JSW17]	IND	$\mathsf{C} \cdot 8\lambda + m \cdot 4\lambda$	$n \cdot 2\lambda + O(d \cdot \lambda^2 \cdot \log \mathsf{C})$	$2^{O(d)}$	OWF
	IND	$\mathsf{C} \cdot 8\lambda + m \cdot 4\lambda$	$n \cdot 2\lambda + O(w \cdot \lambda^2 \cdot \log \mathsf{C})$	$\mathsf{poly}(\mathsf{C}, \lambda)$	OWF
[KKPW21]	IND	$\mathsf{C} \cdot 4\lambda$	$m + n\lambda$	$2^{O(\sqrt{d})}$	OWF
[KKP21]	IND	$\mathsf{C} \cdot 4\lambda + m$	$n\lambda$	$\mathsf{C}^{O(w)}$	OWF
[HJO+16]	SIM	$\mathsf{C} \cdot 4\lambda$	$m + n\lambda + O(d \cdot \lambda^2 \cdot \log \mathsf{C})$	$2^{O(d)}$	OWF
	SIM	$\mathsf{C} \cdot 4\lambda$	$m + n\lambda + O(w \cdot \lambda^2 \cdot \log \mathsf{C})$	$\mathsf{poly}(\mathsf{C}, \lambda)$	OWF
[JW16, JKK+17]	SIM	$\mathsf{C} \cdot 4\lambda$	$m + n\lambda$	$2^{O(d)}$	OWF
[JO20]	SIM	$\mathsf{C} \cdot 2\lambda + \mathsf{A} \cdot \lambda$	$m + n\lambda$	$2^{O(d)}$	OWF
[GS18]	SIM	$\mathsf{C} \cdot \mathsf{poly}(\lambda)$	$m + n + \mathsf{poly}(\lambda, \log \mathsf{C})$	$\mathsf{poly}(\mathsf{C}, \lambda)$	CDH/LWE/etc
[BHR12a]	SIM	$\mathsf{A} \cdot \{1.5\lambda, 2\lambda\} + m$	$n\lambda$	$\mathsf{poly}(\mathsf{C}, \lambda)$	pROM
[LR15]	SIM	$\mathsf{A} \cdot 3\lambda + m$	$n\lambda$	$\mathsf{poly}(\mathsf{C}, \lambda)$	pROM
This work	SIM	$\mathsf{A} \cdot \{1.5\lambda, 2\lambda\}$	$m + n\lambda$	$\mathsf{poly}(\mathsf{C}, \lambda)$	npRPM
	SIM	$\mathsf{A} \cdot \{1.5\lambda, 2\lambda\} + m$	$n\lambda$	$\mathsf{poly}(\mathsf{C}, \lambda)$	pRPM

Table 1: **Comparison of adaptively secure garbling schemes.** "SIM" (resp., "IND") denotes the simulation-based (resp., indistinguishability-based) security. The offline and online costs only focus on communication, where minor terms are omitted for simplicity. The notation C (resp., A) is the total number of all AND and XOR gates (resp., only AND gates). Let λ, n, m, w, d be the security parameter, input size, output size, circuit width and circuit depth, respectively. OWF denotes the one-way function. pROM denotes the programmable random oracle (RO) model. pRPM (resp., npRPM) denotes the programmable (resp., non-programmable) random permutation model.

instantiated by fixed-key AES and accelerated using AES-NI [BHKR13, GKWY20, GKW+20]. This leads to garbling schemes producing more than 20 million garbled `AND` gates per second [WMK16]. On the other hand, research in adaptive security has mostly studied from a more theoretical aspect. In contrast to schemes with selective security, results in adaptive security [BHR12a, HJO+16, JW16, JKK+17, GS18, JO20] are mostly in the standard model but all require some compromise either in exponential security loss or in undesirable online communication, unless Cryptomania assumptions are used. Some efforts consider indistinguishability-based adaptive security [JSW17, KKPW21, KKP21], a weaker notion that does not allow composability immediately, but even with this relaxation, the best-known result is still not desirable. We summarize all garbling schemes with adaptive security in Table 1.

In practice, most implemented garbling schemes only have selective security working in the random permutation model, while most analyzed adaptive schemes only need a pseudorandom function (PRF) but never got implemented. There are some works that sit in the middleland. For example, Gueron et al. [GLNP15] proposed a concretely efficient selective-secure garbling scheme based solely on PRFs; however, the scheme is not compatible with free-XOR. On the other hand, all existing implementations of adaptive garbling schemes are proven in the programmable random oracle model (ROM), either following the work by Bellare et al. [BHR12a] where a generic selective-to-adaptive transformation was proposed [LR14, RR16], or directly proving that the garbling scheme is secure in the programmable ROM [LR15]. Clearly, there is a huge gap between what is being implemented and the security properties that we hope to prove.

1.1 Our Contribution

In this work, we prove that half-gates [ZRE15] and three-halves [RR21], the two state-of-the-art selectively secure garbling schemes, already comply with the adaptive security requirement in the non-programmable RPM (npRPM) [RS08, BHKR13]. As real implementations of these two schemes already used a random permutation to instantiate circular correlation robust (CCR) hash functions [ZRE15,

GKWY20, RR21], our results are essentially proven in the same model as the selective-secure setting. We view our result as something valuable in both theory and practice. On the theoretic side, although the resulting construction still needs an ideal model, it is the first time that the adaptive security of concretely efficient garbling schemes is proven in a non-programmable model; furthermore, a λ-to-λ random permutation appears more plausible compared to a random oracle with unlimited entropy. On the practical side, by not changing the garbling schemes at all, the resulting schemes incur zero additional computation/communication overhead and no change of implementation. Below, we discuss our result in more detail.

Adaptive security in the npRPM. Our main result is the simulation-based adaptive security of the two schemes in the npRPM, which serves as a realistic and conservative model of block ciphers by giving all parties the oracle access to the same random permutation and its inverse. Our proof idea significantly differs from the pebbling games in prior works [HJO+16, JW16, JKK+17, GS18, JO20] to support free-XOR [KS08]: all garbled gates are correlated with the same global key, making it impossible to use a gate-by-gate hybrid argument in a black-box way. Instead, we regard a garbled circuit as a whole by considering it as part of a transcript produced in the interaction between an adaptive adversary and a challenger in the real or ideal world. We bound the advantage of the adaptive adversary via the statistical distance of transcripts between the two worlds, since w.l.o.g. the adversary is deterministic so that its decision bit is a deterministic function of a transcript. As it suffices to prove the adaptive security against a slightly more powerful adversary that has additional information, we introduce an approach called *transcript padding* that pads a transcript with more values to be revealed to the adversary. This approach helps us to compute the statistical distance so as to bound the adversary's advantage. We have the following theorem.

Theorem 1 (informal)**.** *In the non-programmable RPM, half-gates and three-halves are adaptively secure against any computationally unbounded non-uniform adversary, which makes q queries to a random permutation and its inverse, and chooses a circuit $f : \{0,1\}^n \to \{0,1\}^m$ with s AND gates, with advantage at most $O(qs/2^\lambda)$ and online communication of $m + n\lambda$ bits.*

Adaptive security in the pRPM. We complete the picture by proving that half-gates and three-halves are already adaptively secure in the programmable RPM (pRPM) to cover the implementations that send the decoding information in the offline phase. We prove their adaptive security also via transcript padding and the statistical distance of transcripts. One difference is to construct a simulator, who programs the random permutation so that the responses w.r.t. the garbled labels of output wires maintain decoding correctness (i.e., these labels can be decoded to the correct circuit output). We fix the responses in the transcripts with padding, and identify the transcripts, which fail programming, to bound their effect on the statistical distance. As such responses are sampled at random by the simulator, this effect would be negligible. The following result in the pRPM gives a trade-off between online communication and assumption.

Theorem 2 (informal)**.** *In the programmable RPM, half-gates and three-halves are adaptively secure against any computationally unbounded non-uniform adversary, which makes q queries to a random permutation and its inverse, and chooses a circuit $f : \{0,1\}^n \to \{0,1\}^m$ with s AND gates, with advantage at most $O(qs/2^\lambda)$ and online communication of $n\lambda$ bits.*

Separation between the npRPM and pRPM. Finally, we prove that the above gap of online communication is inherent in the programability of the ideal permutation, which gives a separation of adaptively secure garbling schemes in the two RPMs. Particularly, we have that

Theorem 3 (informal)**.** *For any $n, m \in \mathbb{N}^+$, there is a circuit $f : \{0,1\}^n \to \{0,1\}^m$ such that any garbling scheme with simulation-based adaptive security in the non-programmable RPM has online communication of at least m bits.*

Our proof builds upon the poof idea [AIKW13, HW15] in the standard model, which constructs polynomial-sized circuits from the simulator to contradict the Yao entropy [Yao82, BSW03, HLR07]. We modify such circuits to hardcode an approximate random permutation and its inverse, and to guarantee that they are still polynomial-sized. The resulting circuits can be used to break the Yao entropy, leading to a quantitatively identical lower bound of online communication complexity in the npRPM. In the previous result, we already have that half-gates and three-halves can be implemented to match this lower bound. In contrast, the pRPM endows the simulator an additional power, i.e., programmability, to embed a circuit output into its internal state to maintain the decoding correctness, even if the decoding information was fixed before the circuit output is known. As a result, the online communication in the pRPM can bypass the lower bound.

2 Technical Overview

2.1 Previous Techniques

We claim that the existing techniques fail to prove adaptive security of half-gates [ZRE15] and three-halves [RR21], the two state-of-the-art garbling schemes with selective security. The proof techniques underlying these schemes reduce their selective security to circular correlation robust (CCR) hash functions, where the queries to the CCR oracle are adaptive instead of being sent at once. In the adaptive setting, the reduction algorithm cannot make such queries and use responses to construct a garbled circuit (to be sent in the offline phase) since the queries depends on the input unknown until the online phase, failing security proof.

Almost all previous works [HJO+16, JW16, JKK+17, GS18, JO20] w.r.t. adaptive garbling are based on the proof method [LP09] of Yao's garbling scheme in the selective setting (i.e., the approach based on pebbling games).[2] The pebbling-games-based approach changes all real-world garbled gates (i.e., white pebbles) to simulated ones (i.e., black pebbles) using a carefully designed hybrid argument, where each hybrid bridges an input-dependent garbled gate (i.e., gray pebble) and a real-world or simulated one. To ensure that such gray pebbles are consistent with the input (being undefined until the online phase), these works use somewhere equivocal encryption or piecewise guessing in the construction of the gray pebbles, which incurs high overhead or non-negligible security loss.

In the pebbling-game-based works, there are two notable facts: (i) a pebbling hybrid changes only one pebble, and (ii) the indistinguishability between a gray pebble and a white/black one follows from a *black-box* reduction to the security of a cryptographic primitive. As a result, any two pebbling hybrids should be respectively reduced to two *independent* instances of the cryptographic primitive in a black-box way. In essence, this primitive acts as an encryption scheme for truth tables, meaning that all encryptions are independent.

However, both half-gates [ZRE15] and three-halves [RR21] consider free-XOR optimization [KS08] and introduce a global key Δ to the encryption keys. If we adopt pebbling games, the pebbles would be correlated to the same key Δ. Thus, the pebble-by-pebble hybrid argument cannot prove the adaptive security of these schemes. That is, a new approach is required to consider all garbled gates as a whole in order to prove their adaptive security.

2.2 Our Methodology

We present a high-level idea about how to bound the advantage of an adaptive adversary, followed by a toy example for the proof details in the npRPM or the pRPM, and our separation result for these two models.

Bounding the adaptive adversary's advantage via statistical distance. In the RPM, the adaptive experiment of garbling schemes can be captured by the following framework: a computationally unbounded non-uniform adversary $\mathcal{A}$ interacts with either a real-world oracle O_0 or an ideal-world oracle O_1 and outputs its decision bit after the interaction. Specifically, each oracle O_b gives to $\mathcal{A}$ the query

[2] The exceptions are the work by Bellare et al. [BHR12a] using circuit-wise padding in the pROM and the work by Lindell et al. [LR15] with a gate-by-gate use of the pROM.

interfaces for permutation and its inverse along with the two (one-time) interfaces for outputting a garbled circuit $\widehat{f}$ and a garbled input $\widehat{x}$. The interaction between $\mathcal{A}$ and $\mathcal{O}_b$ defines a random variable Z_b of transcripts, which records query-response pairs in order. Without loss of generality, we can assume that a non-uniform adversary $\mathcal{A}$ is *deterministic*[3] such that its decision bit is a deterministic function of its auxiliary input and a transcript sampled according to Z_b. It is well-known that the advantage of the adaptive adversary $\mathcal{A}$ is upper bounded by the statistical distance $\mathsf{SD}(Z_0, Z_1)$.

To bound $\mathsf{SD}(Z_0, Z_1)$, the main task is to compute probability $\Pr[Z_b = \tau]$ for some fixed transcript τ for each $b \in \{0, 1\}$. To compute this probability, it is crucial to address the adaptivity of transcripts that is implicit in the definition of Z_b. We consider a fixed τ of ordered pairs $((q_1, r_1), \ldots, (q_n, r_n))$. Intuitively, the probability will be zero if the next-message function of a fixed non-uniform deterministic $\mathcal{A}$ can never produce queries $q_1, \ldots, q_n$ in order when responses $r_1, \ldots, r_n$ arrives in order. Otherwise, $\mathcal{A}$ certainly produces these queries upon receiving $r_1, \ldots, r_n$ in order (as $\mathcal{A}$ is deterministic and it is a yes-or-no event) and $\Pr[Z_b = \tau]$ quantitatively matches the probability that $\mathcal{A}$ is given ordered responses $r_1, \ldots, r_n$, i.e., the probability that oracle $\mathcal{O}_b$ can produce responses $r_1, \ldots, r_n$ in order when queries $q_1, \ldots, q_n$ arrives in order. Note that the latter probability is easy-to-compute as it is only taken over the randomness of $\mathcal{O}_b$. This observation was used in the analysis [Pat09, CS14, DLMS14, HT16] of symmetric primitives.

Formally speaking, for any $\mathcal{A}$, there exists a set of transcripts $\mathcal{T}_{\mathcal{A}}$ such that for every $\tau \in \mathcal{T}_{\mathcal{A}}$, at least one of $\Pr[Z_0 = \tau]$ and $\Pr[Z_1 = \tau]$ is non-zero[4]. That is, for every $\tau \in \mathcal{T}_{\mathcal{A}}$, $\mathcal{A}$ must produce the ordered queries in τ given the ordered responses in τ (otherwise, the deterministic next-message function of $\mathcal{A}$ results in $\Pr[Z_0 = \tau] = \Pr[Z_1 = \tau] = 0$). Using the aforementioned observation (which is formalized in Lemma 1) for each $b \in \{0, 1\}$, we can replace $\Pr[Z_b = \tau]$ by the probability that the ordered responses of $\mathcal{O}_b$ match their values in τ if $\mathcal{O}_b$ is given the queries in τ in their order. As an abbreviation, we refer to the latter probability as the probability that "$\mathcal{O}_b$ is *compatible* with τ".

In this paper, we use this observation and our *transcript padding* approach (see below for details) to upper bound the statistical distance, where transcript padding eases the analysis of matching probability between each $\mathcal{O}_b$ and τ.

A toy example for our proof idea. We will elaborate our proof idea using an example where adversary $\mathcal{A}$ chooses a circuit f that includes only one `AND` gate g with two input wires (a, b) and an output wire c. Here we only consider half-gates. The idea can be generalized to arbitrary circuits and three-halves.

According to the adaptive experiment in the random permutation model, a transcript τ is comprised of two pairs $(f, \widehat{f})$ and $(x = (x_a, x_b), \widehat{x})$ along with the query-response pairs on permutation π and its inverse π^{-1}, where the queries are chosen adaptively by $\mathcal{A}$. One can think that $\mathcal{A}$ interacts with an integrated real-world oracle $\mathcal{O}_0$ (resp., ideal-world oracle $\mathcal{O}_1$) that returns $\widehat{f}$ and $\widehat{x}$ as per the two phases in the real-world (resp., ideal-world) experiment and gives the interfaces of a random permutation and its inverse as per the specific random permutation model (i.e., npRPM or pRPM).

In the real world, the oracle samples a global key Δ, and uses the half-gates scheme to garble a single `AND` gate g as follows:

$$\begin{cases} G_0 = \mathsf{H}(W_a, k_0) \oplus \mathsf{H}(W_a \oplus \Delta, k_0) \oplus p_b\Delta \\ G_1 = \mathsf{H}(W_b, k_1) \oplus \mathsf{H}(W_b \oplus \Delta, k_1) \oplus W_a \end{cases},$$

where W_i is a 0-label on wire i for $i \in \{a, b\}$, k_0, k_1 are gate-dependent public tweaks, and $p_b = \mathsf{lsb}(W_b)$ is a permuted bit on wire b. From the view of $\mathcal{A}$, we can rewrite (G_0, G_1), which is given in $\widehat{f}$, and the evaluation of g as follows:

$$\begin{cases} G_0 = \mathsf{H}(X_a, k_0) \oplus \mathsf{H}(X_a \oplus \Delta, k_0) \oplus (s_b \oplus x_b)\Delta \\ G_1 = \mathsf{H}(X_b, k_1) \oplus \mathsf{H}(X_b \oplus \Delta, k_1) \oplus X_a \oplus x_a\Delta \\ X_c = \mathsf{H}(X_a, k_0) \oplus \mathsf{H}(X_b, k_1) \oplus s_aG_0 \oplus s_b(G_1 \oplus X_a) \end{cases},$$

[3] A non-uniform adversary is at least as powerful as a probabilistic adversary [Can00].

[4] This is without loss of generality since the transcripts outside $\mathcal{T}_{\mathcal{A}}$ have zero contribution to the statistical distance.

where $X_i = W_i \oplus x_i\Delta$ is a label on actual bit x_i and $s_i = p_i \oplus x_i$ is masked bit for $i \in \{a, b, c\}$. (X_a, X_b) is included in $\widehat{x}$. When $\mathsf{lsb}(\Delta) = 1$ and $p_i = \mathsf{lsb}(W_i)$, we must have $s_i = \mathsf{lsb}(X_i)$ for $i \in \{a, b, c\}$. When the hash function H is instantiated in a way that $\mathsf{H}(X, k) = \pi(X \oplus k) \oplus \sigma(X \oplus k)$, where π is a random permutation and σ is a linear orthomorphism (see [GKWY20] for more details), these equalities for H can be translated into three constraints on permutation π:

$$\begin{cases} \pi(X_a \oplus k_0) \oplus \pi(X_a \oplus \Delta \oplus k_0) = \sigma(\Delta) \oplus (s_b \oplus x_b)\Delta \oplus G_0 \\ \pi(X_b \oplus k_1) \oplus \pi(X_b \oplus \Delta \oplus k_1) = \sigma(\Delta) \oplus X_a \oplus x_a\Delta \oplus G_1 \\ \pi(X_a \oplus k_0) \oplus \pi(X_b \oplus k_1) = \sigma(X_a \oplus k_0 \oplus X_b \oplus k_1) \\ \qquad\qquad \oplus s_a G_0 \oplus s_b(G_1 \oplus X_a) \oplus X_c \end{cases}. \tag{1}$$

In the ideal world, the oracle invokes the simulator, which samples (G_0, G_1) in $\widehat{f}$ and (X_a, X_b) in $\widehat{x}$ uniformly at random. Here, the first two constraints in (1) can be removed for the ideal-world oracle since $\mathcal{A}$ can succeed to guess Δ with negligible probability. In both worlds, decoding information $d = d_c$ is included in $\widehat{f}$ or $\widehat{x}$, depending on the realistic implementations of half-gates, such that *decoding consistency* $x_c = d_c \oplus \mathsf{lsb}(X_c)$ needs to hold.

Given these oracle constructions, we would like to adopt the aforementioned observation to bound the statistical distance $\mathsf{SD}(Z_0, Z_1)$, in order to bound the advantage of adaptive adversaries. However, it is complex to directly study the probability that a sampled real- or ideal-world oracle is compatible with a fixed transcript. For example, we need to compute the probability that permutation π, which is sampled at random in the real world, satisfies the constraint in (1) conditioned on the decoding consistency and the consistency between (X_a, X_b) and randomly sampled (W_a, W_b), i.e., $X_i = W_i \oplus x_i\Delta$ for $i \in \{a, b\}$.

In the context of adaptive garbling, we address the probability computation w.r.t. the permutation and its inverse by *transcript padding*, an approach that somehow hardcodes into transcripts the linkage between some permutation pre-images and images. In particular, for the toy example, we explicitly (i) include (x_c, X_c, Δ) in transcripts (where Δ is revealed at the end of the experiment so that no more permutation queries can depend on it), (ii) ask $\mathcal{A}$ to make extra queries $X_a \oplus k_0$ and $X_b \oplus k_1$ to the permutation as soon as it receives $\widehat{x}$ (i.e., $(X_a \oplus k_0, \pi(X_a \oplus k_0))$ and $(X_b \oplus k_1, \pi(X_b \oplus k_1))$ are hardcoded in the query-response pairs), and (iii) define $\mathcal{T}_{\mathcal{A}}$ to include the transcripts with the padding values. This approach does not lower the advantage of $\mathcal{A}$ (as $\mathcal{A}$ can ignore the padding values) but helps to fix some linkage between permutation pre-images and images by peeling (1), which is well-defined from a transcript in $\mathcal{T}_{\mathcal{A}}$. So, for example, the probability that a real-world permutation π satisfies (1) matches the probability that a random permutation has such fixed linkage.

We use the observation to bound $\mathsf{SD}(Z_0, Z_1)$ from probability $\mathsf{Pr}_0(\tau)$ (resp., $\mathsf{Pr}_1(\tau)$) that a randomly sampled real-world oracle O_0 (resp., ideal-world oracle O_1) is compatible with a fixed transcript $\tau \in \mathcal{T}_{\mathcal{A}}$. To prove a negligible bound of the statistical distance, it is sufficient to consider the bound computed from the H-coefficient technique [Pat09, CS14, DLMS14], which partitions $\mathcal{T}_{\mathcal{A}}$ into a subset $\mathcal{T}_{\mathsf{good}}$ of good transcripts and a subset $\mathcal{T}_{\mathsf{bad}} := \mathcal{T}_{\mathcal{A}} \setminus \mathcal{T}_{\mathsf{good}}$ of bad transcripts. It proves that $\mathsf{SD}(Z_0, Z_1) \leq \varepsilon_1 + \varepsilon_2$ if (i) $\sum_{\tau \in \mathcal{T}_{\mathsf{bad}}} \mathsf{Pr}_1(\tau) \leq \varepsilon_1$ and (ii) for every $\tau \in \mathcal{T}_{\mathsf{good}}$, $1 - \mathsf{Pr}_0(\tau)/\mathsf{Pr}_1(\tau) \leq \varepsilon_2$. Below, we will show how to compute such probabilities in the presence of transcript padding.

2.3 Proofs of Adaptive Security in npRPM and pRPM

Proving adaptive security in the npRPM. In this model, we focus on the known implementations (e.g., Obliv-C [ZE15], ObliVM [LWN+15], and TinyGarble [SHS+15]) in which the decoding information is sent in the online phase. That is, for the toy example, $\widehat{x}$ includes two garbled labels (X_a, X_b), decoding information d_c, and (x_c, X_c) (for transcript padding). This model gives the real-world garbling algorithms and the simulator oracle access to the same random permutation π and its inverse π^{-1}. In the ideal world, the simulator computes X_c by asking π about the queries $X_a \oplus k_0$ and $X_b \oplus k_1$ for (1) and defines $d_c := x_c \oplus \mathsf{lsb}(X_c)$ to let $\mathcal{A}$ learn the output $x_c = x_a \cdot x_b$, even if (X_a, X_b) are sampled at random.

We use a hybrid argument with three hybrids: ideal world, intermediate hybrid, and real world. The intermediate hybrid can ease probability analysis and differs from the ideal world by replacing a

random permutation and its inverse $\pi^{\pm 1}$ with an approximate one (denoted by $\widetilde{\pi}^{\pm 1}$), which outputs a fresh random string for a different query of the simulator but maintains the query-response consistency. This hybrid is statistically close to the ideal world up to a polynomial number of queries made by the simulator, giving the negligible advantage that an adaptive adversary $\mathcal{A}$ can distinguish the intermediate hybrid from the ideal world. Then, we prove that $\mathcal{A}$ has a negligible advantage in distinguishing between the intermediate hybrid and the real world. To do this, we will bound $\mathsf{SD}(Z_0, Z_1')$, where Z_1' denotes the random variable for transcripts produced in $\mathcal{A}$'s interaction with oracle O_1' in the intermediate hybrid, and O_1' is identical to an ideal-world oracle O_1 except for replacing $\pi^{\pm 1}$ by $\widetilde{\pi}^{\pm 1}$. Let $\mathsf{Pr}_1'(\tau)$ denote the probability that an oracle O_1' is compatible with a fixed transcript $\tau \in \mathcal{T}_{\mathcal{A}}$. Our transcript padding and the observation along with a bound computed from the H-coefficient technique (which jointly bound $\mathsf{SD}(Z_0, Z_1)$ as sketched above) can likewise give a bound of $\mathsf{SD}(Z_0, Z_1')$ by replacing $\mathsf{Pr}_1(\tau)$ with $\mathsf{Pr}_1'(\tau)$.

To compute this bound, we can define a good transcript as that where[5] (i) the values $X_a \oplus k_0$, $X_b \oplus k_1$, $X_a \oplus k_0 \oplus \Delta$ and $X_b \oplus k_1 \oplus \Delta$ (i.e., the queries to π for (1)) are pairwise distinct, (ii) their responses, which are already fixed by the transcript by peeling (1) with the padding values, are also pairwise distinct, and (iii) the extra queries made by $\mathcal{A}$ on $X_a \oplus k_0$ and $X_b \oplus k_1$ occur only after the same queries made by the simulator before it sends $\widehat{x}$ to $\mathcal{A}$. Condition (iii) is used to simplify $\mathsf{Pr}_1'(\tau)$ as the values of $\widetilde{\pi}(X_a \oplus k_0)$ and $\widetilde{\pi}(X_b \oplus k_1)$ will be uniformly random in the intermediate hybrid. Meanwhile, condition (i) and (ii) ensure that no hash mask can be unmasked by XORing G_0 and G_1 so that one can distinguish the two hybrids without making any additional query w.r.t. Δ, blowing up the statistical distance since it is an upper bound of advantage.

For $\mathsf{Pr}_0(\tau)$ for a fixed good transcript τ in the real world, the four pairwise distinct pre-images by (i) are linked to the four pairwise distinct images by (ii), respectively, where these pre-images and images are hardcoded in τ as per the real-world constraint (1). Thus, this probability can be computed from (a) the probability that a real-world oracle, given the ordered queries in τ, can produce the responses in τ other than the four hardcoded pairs and the query-response pairs on the permutation and its inverse, and (b) the probability that it, given the ordered queries in τ, can produce the four hardcoded pairs and the query-response pairs on the permutation and its inverse conditioned on the event in (a). The former is taken over the randomness of (W_a, W_b, Δ) while the latter is based on the number of π's being consistent with the pairs.

We compute $\mathsf{Pr}_1'(\tau)$ for a fixed good transcript τ in the intermediate hybrid likewise but benefit from the uniform values of $\widetilde{\pi}(X_a \oplus k_0)$ and $\widetilde{\pi}(X_b \oplus k_1)$ by condition (iii). For $\sum_{\tau \in \mathcal{T}_{\mathsf{bad}}} \mathsf{Pr}_1'(\tau)$, we note that the pairwise distinctness in (i) or the order needed by (iii) fails with negligible probability for the randomness of (W_a, W_b, Δ) and the garbling $X_i := W_i \oplus x_i\Delta$ for $i \in \{a, b\}$. Conditioned on the satisfied (i) and (iii), the values of $\widetilde{\pi}(X_a \oplus k_0)$ and $\widetilde{\pi}(X_b \oplus k_1)$ are uniform due to the pairwise distinct pre-images of $\widetilde{\pi}$ and will be pairwise distinct except with negligible probability. The values of $\widetilde{\pi}(X_a \oplus \Delta \oplus k_0)$ and $\widetilde{\pi}(X_b \oplus \Delta \oplus k_1)$, which are computed by peeling (1), are also uniform according to the values of $\widetilde{\pi}(X_a \oplus k_0)$ and $\widetilde{\pi}(X_b \oplus k_1)$, respectively. Therefore, conditioned on (i) and (iii), (ii) occurs also with negligible probability, concluding negligible probability of bad transcripts. Note that we can generalize the above analysis to an arbitrary circuit using an induction.

In the npRPM, the proof of half-gates is given in Section 4, and the proof of three-halves is given in Appendix D.2.

Proving adaptive security in the pRPM. In this model, we focus on the implementations (e.g., ABY [DSZ15] and MP-SPDZ [Kel20]) which send the decoding information in the offline phase. Intuitively, it is natural to prove the adaptive security of half-gates [ZRE15] and three-halves [RR21] in the pRPM, where the view of any adaptive adversary is simulated by programming the random permutation (and its inverse). We confirm this intuition by proving the adaptive security of half-gates in the pRPM, and also extend the proof idea to three-halves.

For the toy example in the pRPM, as decoding information d_c is sent to an adaptive adversary before

[5]For the case of general circuits (instead of the toy example), we require that all the conditions further hold for the corresponding values across any two `AND` gates.

the simulator knows the output bit $x_c = x_a \cdot x_b$, the simulator should program a random permutation whose output for some query is a garbled label X_c such that $\mathsf{lsb}(X_c) = d_c \oplus x_c$ when x_c is known.

The proof in the pRPM is a single-hop hybrid argument without relying on an intermediate hybrid. As shown in the toy example, we bound the statistical distance $\mathsf{SD}(Z_0, Z_1)$ between the ideal world and real world, from the transcript padding and the observation with the H-coefficient technique. Here, $\mathsf{Pr}_0(\tau)$ and $\mathsf{Pr}_1(\tau)$ can be similarly computed as $\mathsf{Pr}_0(\tau)$ and $\mathsf{Pr}'_1(\tau)$ in the npRPM, respectively, for an *identically* defined good transcript τ. In particular, the conditions (i), (ii) and (iii) jointly ensure that, in the ideal world, the simulator succeeds in programming the random permutation as per the last equality in (1) for an X_c allowing for the decoding consistency. The reason is that the values of the programmed entries, which are hardcoded in τ, are pairwise distinct and never queried by $\mathcal{A}$ before programming, under these conditions. For the probability of bad transcripts, we stress that the random permutation and its inverse are emulated by the simulator in the pRPM so that a response for a different query approximately has an entropy of λ bits. Thus, we have negligible probability of bad transcripts by using a similar analysis in the npRPM.

We present in Appendix C the proof of half-gates and in Appendix D.1 the proof of three-halves in the pRPM.

2.4 Separating npRPM from pRPM for Adaptive Garbling

In the npRPM, the simulator in the adaptive garbling has two simulation algorithms $\mathsf{SimF}^{\pi^{\pm 1}(\cdot)}$ and $\mathsf{SimIn}^{\pi^{\pm 1}(\cdot)}$, which have oracle access to a random permutation π and its inverse π^{-1}. The former outputs a (simulated) garbled circuit $\widehat{f}$ and an internal state without knowing the actual input x while the latter outputs a (simulated) garbled input $\widehat{x}$ given the internal state and a circuit output $f(x)$. Intuitively, as the simulated $\widehat{x}$ is the only message that can depend on $f(x)$, $\widehat{x}$ should consist of "adequate" information of $f(x)$. Otherwise, we cannot reconstruct $f(x)$ from the simulated $(\widehat{f}, \widehat{x})$ in the ideal world, making the real world and ideal world trivially distinguishable.

Following prior works [AIKW13, HW15], we will also evaluate the "adequacy" based on the Yao entropy [Yao82, BSW03, HLR07] of $f(x)$. This entropy is formalized as the minimal bit-length of an efficiently computable compressed form of $f(x)$, which can be further decompressed to $f(x)$ with probability significantly more than $1/2$. Based on the prior proofs [AIKW13, HW15] in the standard model, we can prove that the bit-length of $\widehat{x}$ should exceed the Yao entropy of $f(x)$ in the npRPM. The high-level idea is that the simulator can compress $f(x)$ into $\widehat{x}$, which can be decompressed using the evaluation algorithm. Since the simulator and the evaluation algorithm are probabilistic polynomial-time (PPT) and the adversary only makes a bounded number of queries, such compression and decompression need polynomial-sized circuits as required by the Yao entropy. In particular, they overall make a polynomial number of queries on $\pi^{\pm 1}$ so that the corresponding responses can be replaced by uniform strings hardcoded in the circuits.

More specifically, we first define an intermediate hybrid by replacing a random permutation and its inverse $\pi^{\pm 1}$ with "coarse" approximation $\mathring{\pi}^{\pm 1}$, which outputs a fresh random string for a new query but ensures the query-response consistency. Compared to the previous approximation $\widetilde{\pi}^{\pm 1}$, this coarse one answers not only the unique queries made by the simulator but also those made by the adversary with uniform strings. Clearly, the advantage of any adaptive adversary to distinguish this hybrid and the ideal world is at most twice the birthday bound, which is negligible up to the polynomial number of queries to the permutation and its inverse. As the garbling scheme is adaptively secure in the npRPM, the ideal world should be indistinguishable from the real world. As a corollary, $\mathsf{SimF}^{\mathring{\pi}^{\pm 1}(\cdot)}$ and $\mathsf{SimIn}^{\mathring{\pi}^{\pm 1}(\cdot)}$ give an approximate simulation that is indistinguishable from the real world.

This indistinguishability implies that the approximate simulation can yield $(\widehat{f}, \widehat{x})$ that can be decoded to $f(x)$ with overwhelming probability. If this $\widehat{x}$ has a bit-length lower than the Yao entropy of $f(x)$, we show that the approximate simulation contradicts this entropy. We can construct a compression circuit to sequentially run $\mathsf{SimF}^{\mathring{\pi}^{\pm 1}(\cdot)}$ and $\mathsf{SimIn}^{\mathring{\pi}^{\pm 1}(\cdot)}$ over some hardcoded random tape to output such an $\widehat{x}$. The decompression circuit uses the same random tape to recompute $\widehat{f}$ output by $\mathsf{SimF}^{\mathring{\pi}^{\pm 1}(\cdot)}$, and then runs the deterministic evaluation algorithm to compute $f(x)$ from $(\widehat{f}, \widehat{x})$. According to the usage of a coarse approximate permutation (which is hardcoded in the circuits), we have that both cir-

cuits are of polynomial sizes. So, these circuits contradict the Yao entropy, leading to a lower bound of the online complexity in the npRPM. The detailed proof is presented in Appendix E.

However, this lower bound does not hold in the pRPM, where the simulator has another PPT algorithm $\mathsf{SimP}^{\pm 1}$ to emulate a random permutation and its inverse for queries (a) before $\widehat{f}$, (b) after $\widehat{f}$ but before $\widehat{x}$, and (c) after $\widehat{x}$. It is notable that SimF, SimIn, and $\mathsf{SimP}^{\pm 1}$ maintain an internal state. In particular, SimIn is given $f(x)$ and the internal state, and outputs a simulated $\widehat{x}$ and *an updated internal state to be used in* $\mathsf{SimP}^{\pm 1}$ *for the case (c).* Although we can show that the bit-length of the output by SimIn should exceed the Yao entropy of $f(x)$, a part of the information of $f(x)$ can be transferred into the updated internal state, so that the bit-length of the simulated $\widehat{x}$ can be lower than the Yao entropy. As a corollary, the lower bound does not apply to a real-world $\widehat{x}$, given that the real and ideal worlds are indistinguishable due to the adaptive security in the pRPM. This result is confirmed by our proofs for the adaptive security of half-gates and three-halves in the pRPM, where we embed decoding consistency, e.g., $\mathsf{lsb}(X_c) = x_c \oplus d_c$, into the internal state.

3 Preliminaries

3.1 Notation

Throughout this paper, we use $\lambda \in \mathbb{N}$ to denote the security parameter. We use $\mathsf{poly}(\cdot)$ (resp., $\mathsf{negl}(\cdot)$) for an unspecified polynomial (resp., negligible) function. For $a, b \in \mathbb{N}$ with $a \leq b$, we denote by $[a, b]$ the set $\{a, \ldots, b\}$ and by $(b)_a$ the falling factorial $b \cdot (b-1) \cdots (b-a+1)$. We use $x \leftarrow S$ to denote the uniform sampling of x from a finite set S. We use $:=$ to denote assigning a value or an output of a deterministic algorithm to a left-hand variable. Let $\mathcal{P}_\ell$ denote the set of permutations on $\{0,1\}^\ell$. Let $\mathsf{lsb}(x)$ denote the least significant bit (LSB) of $x \in \{0,1\}^n$. Let $\|$ denote the concatenation of bit-strings. Let $\ominus$ denote the symmetric difference of sets, i.e., for two sets A, B, $A \ominus B := (A \backslash B) \cup (B \backslash A)$.

Linear orthomorphism. A permutation $\sigma : \mathbb{G} \to \mathbb{G}$ over an additive Abelian group $\mathbb{G}$ is called a linear orthomorphism if (i) $\sigma(x+y) = \sigma(x) + \sigma(y)$ for any $x, y \in \mathbb{G}$, (ii) $\sigma'(x) := \sigma(x) - x$ is also a permutation, and (iii) σ, σ' and their inverses are efficiently computable. There are two simple instantiations in [GKWY20]: (i) if $\mathbb{G}$ is a field, $\sigma(x) := c \cdot x$ for some $c \neq 0, 1 \in \mathbb{G}$, and (ii) if $\mathbb{G} = \{0,1\}^n$, $\sigma(x) := (x_L \oplus x_R) \,\|\, x_L$ where x_L and x_R are the left and right halves of x.

Circuits. Given a circuit f with fan-in two and fan-out one, we define:

- $|f|$: The number of `AND` gates.
- $\mathcal{W}(f)$, $\mathcal{W}_{\mathsf{in}}(f)$, $\mathcal{W}_{\mathsf{out}}(f)$, $\mathcal{W}_{\mathsf{and}}(f)$: The sets of wires, circuit input wires, circuit output wires, and output wires of `AND` gates, respectively.
- $\mathcal{G}(f)$, $\mathcal{G}_{\mathsf{and}}(f)$: The sets of gates and `AND` gates, respectively.
- For a gate $g \in \mathcal{G}(f)$, let $(a, b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$ denote the two input wires and $c := \mathsf{out}(g)$ denote the output wire.

In Appendix A, we present the additional notation used in the appendices.

3.2 Random Permutation Model

In the random permutation model (RPM) [RS08, BHKR13], all parties have oracle access to random permutation π and its inverse $\pi^{-1} := \mathsf{inv}(\pi)$. We refer to queries to π as *forward queries* and queries to π^{-1} as *backward queries.* In this work, we consider *non-programmable RPM* (npRPM) and *programmable RPM* (pRPM). Inspired by the separation [Nie02] w.r.t. the random oracle model [BR93], we formalize the pRPM like a hybrid model, where an ideal functionality provides two $\pi^{\pm 1}$ interfaces. This model gives the simulator the power to "appropriately" choose the responses to the adversary's queries to $\pi^{\pm 1}$. In contrast, all parties (as well as the real-world adversary and the simulator) in the npRPM are given oracle access to a global $\pi^{\pm 1}$, whose responses cannot be chosen by the simulator.

3.3 Adaptive Security of Garbling Schemes

In Definition 1, we adapt the definition of adaptively secure garbling schemes in [HJO+16, JW16] for a q-query *computationally unbounded* adversary in the npRPM and the pRPM. This definition combines the evaluation and decoding algorithms in the literature [BHR12a, BHR12b] and does not explicitly send the decoding table as part of a garbled circuit $\widehat{f}$. To simplify the notation, we assume that all algorithms and the adversary implicitly take the unary security parameter 1^λ as input.

Definition 1 (Adaptively secure garbling scheme)**.** *For some polynomial ℓ, an $\ell(\lambda)$-garbling scheme in the npRPM or pRPM has three PPT algorithms, each of which is given oracle access to a random permutation $\pi \in \mathcal{P}_{\ell(\lambda)}$ and its inverse $\pi^{-1} := \mathsf{inv}(\pi)$:*

- $(\widehat{f}, k) \leftarrow \mathsf{Garble}^{\pi^{\pm 1}(\cdot)}(f)$. *The bit-length of $\widehat{f}$ is called* ***offline complexity****.*
- $\widehat{x} := \mathsf{Encode}^{\pi^{\pm 1}(\cdot)}(k, x)$. *The bit-length of $\widehat{x}$ is called* ***online complexity****.*
- $y := \mathsf{DecEval}^{\pi^{\pm 1}(\cdot)}(\widehat{f}, \widehat{x})$.

For some polynomials q, s, negligible function ε, and side-information function Φ, this scheme is said $(q(\lambda), s(\lambda), \varepsilon(\lambda), \Phi)$-adaptively secure in the npRPM (resp., pRPM) if it complies with ***correctness*** *and* ***adaptive security in the npRPM (resp., pRPM)****. By default, $\Phi(f) = f$ is omitted in the definition.*

- ***Correctness.*** *For every polynomial-size circuit $f : \{0,1\}^{\ell_{\mathsf{in}}} \to \{0,1\}^{\ell_{\mathsf{out}}}$, and every input $x \in \{0,1\}^{\ell_{\mathsf{in}}}$, it holds that*

$$\Pr\left[\begin{array}{l} \pi \leftarrow \mathcal{P}_{\ell(\lambda)}, \pi^{-1} := \mathsf{inv}(\pi), \\ (\widehat{f}, k) \leftarrow \mathsf{Garble}^{\pi^{\pm 1}(\cdot)}(f), : \mathsf{DecEval}^{\pi^{\pm 1}(\cdot)}(\widehat{f}, \widehat{x}) = f(x) \\ \widehat{x} := \mathsf{Encode}^{\pi^{\pm 1}(\cdot)}(k, x) \end{array}\right] = 1.$$

- ***Adaptive security in the npRPM.*** *There exists a PPT simulator*

$$\mathsf{Sim} = (\mathsf{SimF}, \mathsf{SimIn})$$

with an internal state $\mathsf{st}_{\mathsf{sim}}$ such that, for every auxiliary input $z \in \{0,1\}^*$ and every computationally unbounded adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3)$ totally making $q(\lambda)$ oracle queries and choosing a circuit with $s(\lambda)$ `AND` gates,

$$\left|\begin{array}{l} \Pr\left[\begin{array}{l} \pi \leftarrow \mathcal{P}_{\ell(\lambda)}, \pi^{-1} := \mathsf{inv}(\pi), \\ (f, \mathsf{st}_1) \leftarrow {\mathcal{A}_1}^{\pi^{\pm 1}(\cdot)}(z), \\ (\widehat{f}, k) \leftarrow \mathsf{Garble}^{\pi^{\pm 1}(\cdot)}(f), : {\mathcal{A}_3}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x}) = 1 \\ (x, \mathsf{st}_2) \leftarrow {\mathcal{A}_2}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_1, \widehat{f}), \\ \widehat{x} := \mathsf{Encode}^{\pi^{\pm 1}(\cdot)}(k, x) \end{array}\right] \\ - \Pr\left[\begin{array}{l} \pi \leftarrow \mathcal{P}_{\ell(\lambda)}, \pi^{-1} := \mathsf{inv}(\pi), \\ (f, \mathsf{st}_1) \leftarrow {\mathcal{A}_1}^{\pi^{\pm 1}(\cdot)}(z), \\ \widehat{f} \leftarrow \mathsf{SimF}^{\pi^{\pm 1}(\cdot)}(\Phi(f)), : {\mathcal{A}_3}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x}) = 1 \\ (x, \mathsf{st}_2) \leftarrow {\mathcal{A}_2}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_1, \widehat{f}), \\ \widehat{x} \leftarrow \mathsf{SimIn}^{\pi^{\pm 1}(\cdot)}(f(x)) \end{array}\right] \end{array}\right| \leq \varepsilon(\lambda).$$

- ***Adaptive security in the pRPM.*** *There exists a PPT simulator*

$$\mathsf{Sim} = (\mathsf{SimF}, \mathsf{SimIn}, \mathsf{SimP}^{\pm 1})$$

with an internal state $\mathsf{st}_{\mathsf{sim}}$ *such that, for every auxiliary input* $z \in \{0,1\}^*$ *and every computationally unbounded adversary* $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3)$ *totally making* $q(\lambda)$ *oracle queries and choosing a circuit with* $s(\lambda)$ AND *gates,*

$$\left| \Pr\left[\begin{array}{l} \pi \leftarrow \mathcal{P}_{\ell(\lambda)}, \pi^{-1} := \mathsf{inv}(\pi), \\ \quad (f, \mathsf{st}_1) \leftarrow {\mathcal{A}_1}^{\pi^{\pm 1}(\cdot)}(z), \\ (\widehat{f}, k) \leftarrow \mathsf{Garble}^{\pi^{\pm 1}(\cdot)}(f), : {\mathcal{A}_3}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x}) = 1 \\ (x, \mathsf{st}_2) \leftarrow {\mathcal{A}_2}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_1, \widehat{f}), \\ \quad \widehat{x} := \mathsf{Encode}^{\pi^{\pm 1}(\cdot)}(k, x) \end{array} \right] \right.$$
$$\left. - \Pr\left[\begin{array}{l} \quad (f, \mathsf{st}_1) \leftarrow \mathcal{A}_1^{\mathsf{SimP}^{\pm 1}(\cdot)}(z), \\ \quad\quad \widehat{f} \leftarrow \mathsf{SimF}(\Phi(f)), \\ (x, \mathsf{st}_2) \leftarrow \mathcal{A}_2^{\mathsf{SimP}^{\pm 1}(\cdot)}(\mathsf{st}_1, \widehat{f}), \\ \quad\quad \widehat{x} \leftarrow \mathsf{SimIn}(f(x)) \end{array} : \mathcal{A}_3^{\mathsf{SimP}^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x}) = 1 \right] \right| \leq \varepsilon(\lambda).$$

3.4 Framework of Adaptive Security and H-coefficient Technique

We prove the adaptive security of garbling schemes in the following framework of adaptive experiments. In this framework, we consider a computationally unbounded non-uniform adversary $\mathcal{A}$, which takes an auxiliary input and outputs a decision bit after making a bounded number of adaptive queries to either a real-world or ideal-world oracle. Each oracle is *stateful*: for a query, a response is a *deterministic* function of the random tape, the query, and previous query-response pairs. The two oracles can give multiple interfaces, each of which has the same syntax in the two worlds. For simplicity, whenever we refer to a non-uniform adversary (e.g., $\mathcal{A}$ or $\mathcal{A}'$), we assume that it has some auxiliary input fixed in its context unless this auxiliary input is given explicitly.

Without loss of generality, we consider a *deterministic* computationally unbounded non-uniform adversary $\mathcal{A}$. The interaction between $\mathcal{A}$ and the oracle produces a transcript, which gives an *ordered* list of query-response pairs from the view of $\mathcal{A}$. For some fixed $\mathcal{A}$, the distribution of transcripts in either world results from the oracle's random tape, or equivalently, the random sampling of a *deterministic* oracle. So, the decision bit of $\mathcal{A}$ is a deterministic function of its (fixed) auxiliary input and a transcript produced in the interaction, and its advantage is at most the statistical distance of transcripts in the two worlds.

Let Ω_{real} (resp., Ω_{ideal}) denote the sample space where a *deterministic* real-world (resp., ideal-world) oracle is sampled at random. For any fixed $\mathcal{A}$, let $\mathcal{T}_{\mathcal{A}}$ denote the set of *attainable* transcripts s.t. a transcript $\tau \in \mathcal{T}_{\mathcal{A}}$ if and only if there exists a *deterministic*[6] oracle ω' such that the interaction between $\mathcal{A}$ and ω' produces τ. Let $X : \Omega_{\mathsf{real}} \to \mathcal{T}_{\mathcal{A}}$ (resp., $Y : \Omega_{\mathsf{ideal}} \to \mathcal{T}_{\mathcal{A}}$) denote the random variable w.r.t. the transcripts produced in the interaction between $\mathcal{A}$ and a real-world (resp., ideal-world) oracle ω sampled from Ω_{real} (resp., Ω_{ideal}). For any fixed τ, let $\mathsf{comp}_{\mathsf{real}}(\tau) \subseteq \Omega_{\mathsf{real}}$ (resp., $\mathsf{comp}_{\mathsf{ideal}}(\tau) \subseteq \Omega_{\mathsf{ideal}}$) denote the set of *compatible* real-world (resp., ideal-world) oracles s.t. an oracle $\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)$ (resp., $\omega \in \mathsf{comp}_{\mathsf{ideal}}(\tau)$) if and only if there exists a *deterministic* non-uniform adversary $\mathcal{A}'$ such that the interaction between $\mathcal{A}'$ and ω produces τ.

A key observation is that the adaptive interaction in random variables X, Y can be "unfolded" to be equivalently but easily studied from the compatibility between oracle and transcript. This compatibility, as defined above, essentially means that an oracle will returns the responses in a fixed transcript τ in order if the queries match their counterparts in τ and are sent (by adversary $\mathcal{A}'$) to the oracle in the given order. This observation is formalized in Lemma 1, and its proof for stateful oracles was sketched in [DLMS14, Appendix D]. We also present a formal proof in Appendix B for completeness.

[6] Without loss of generality, we can assume that such an $\omega' \in \Omega_{\mathsf{real}} \cup \Omega_{\mathsf{ideal}}$ so that it is deterministic. Otherwise (i.e., no such an ω'), the statistical distance is zero.

$\underline{\mathsf{HG.Garble}^{\pi^{\pm 1}(\cdot)}(f)}$:

$\Delta \leftarrow \{0,1\}^{\lambda-1} \| 1$
for $i \in \mathcal{W}_{\mathsf{in}}(f)$ **do**
  $W_i \leftarrow \{0,1\}^{\lambda}$
for $g \in \mathcal{G}(f)$ in topology order **do**
  $(a,b,c) := (\mathsf{in}_0(g), \mathsf{in}_1(g), \mathsf{out}(g))$
  if $\mathsf{type}(g) = \texttt{XOR}$ **then** $W_c := W_a \oplus W_b$
  else if $\mathsf{type}(g) = \texttt{AND}$ **then**
    $k_0^g := 2 \cdot g - 1, k_1^g := 2 \cdot g$
    $p_a := \mathsf{lsb}(W_a), p_b := \mathsf{lsb}(W_b)$
    $G_0^g := \mathsf{H}(W_a, k_0^g) \oplus \mathsf{H}(W_a \oplus \Delta, k_0^g) \oplus p_b \Delta$
    $G_1^g := \mathsf{H}(W_b, k_1^g) \oplus \mathsf{H}(W_b \oplus \Delta, k_1^g) \oplus W_a$
    $W_c := \mathsf{H}(W_a \oplus p_a \Delta, k_0^g) \oplus \mathsf{H}(W_b \oplus p_b \Delta, k_1^g) \oplus p_a p_b \Delta$
for $i \in \mathcal{W}_{\mathsf{out}}(f)$ **do** $d_i := \mathsf{lsb}(W_i)$
return $\widehat{f} := (f' := f, F := \{(G_0^g, G_1^g)\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}), k := (f, d, \Delta, W)$

$\underline{\mathsf{HG.DecEval}^{\pi^{\pm 1}(\cdot)}(\widehat{f}, \widehat{x})}$:

Parse $\widehat{f} = (f, \{(G_0^g, G_1^g)\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}), \widehat{x} = (\{X_i\}_{i \in \mathcal{W}_{\mathsf{in}}(f)}, d)$
for $g \in \mathcal{G}(f)$ in topology order **do**
  $(a,b,c) := (\mathsf{in}_0(g), \mathsf{in}_1(g), \mathsf{out}(g))$
  if $\mathsf{type}(g) = \texttt{XOR}$ **then** $X_c := X_a \oplus X_b$
  else if $\mathsf{type}(g) = \texttt{AND}$ **then**
    $k_0^g := 2 \cdot g - 1, k_1^g := 2 \cdot g$
    $s_a := \mathsf{lsb}(X_a), s_b := \mathsf{lsb}(X_b)$
    $X_c := \mathsf{H}(X_a, k_0^g) \oplus \mathsf{H}(X_b, k_1^g) \oplus s_a G_0^g \oplus s_b (G_1^g \oplus X_a)$
for $i \in \mathcal{W}_{\mathsf{out}}(f)$ **do** $y_i := d_i \oplus \mathsf{lsb}(X_i)$
return y

$\underline{\mathsf{HG.Encode}^{\pi^{\pm 1}(\cdot)}(k, x)}$:

Parse $k = (f, d, \Delta, W)$
for $i \in \mathcal{W}_{\mathsf{in}}(f)$ **do**
  $X_i := W_i \oplus x_i \Delta$
return $\widehat{x} := (\{X_i\}_{i \in \mathcal{W}_{\mathsf{in}}(f)}, d)$

Figure 1: Half-gates garbling scheme [ZRE15].

Lemma 1 ([DLMS14]). *Let the notations be defined in Section 3.4. Then, for every auxiliary input $z \in \{0,1\}^*$, every computationally unbounded adversary $\mathcal{A}$, and every attainable transcript $\tau \in \mathcal{T}_{\mathcal{A}(z)}$, it holds that*

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[X(\omega) = \tau] = \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)],$$
$$\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] = \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\omega \in \mathsf{comp}_{\mathsf{ideal}}(\tau)].$$

For every fixed non-uniform $\mathcal{A}$, the H-coefficient technique [Pat09, CS14, DLMS14] bounds the statistical distance of transcripts in the experiment. It divides $\mathcal{T}_{\mathcal{A}}$ into two disjoint subsets $\mathcal{T}_{\mathsf{bad}} \subseteq \mathcal{T}_{\mathcal{A}}$ and $\mathcal{T}_{\mathsf{good}} := \mathcal{T}_{\mathcal{A}} \setminus \mathcal{T}_{\mathsf{bad}}$. Then, it can prove an upper bound $\varepsilon_1 + \varepsilon_2$ of this statistical distance if it holds that

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) \in \mathcal{T}_{\mathsf{bad}}] \le \varepsilon_1, \qquad \forall \tau \in \mathcal{T}_{\mathsf{good}} : \frac{\Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[X(\omega) = \tau]}{\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) = \tau]} \ge 1 - \varepsilon_2,$$

where, for some $\tau \in \mathcal{T}_{\mathsf{good}}$, ε_2 is defined to 0 if $\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] = 0$.

4 Adaptive Security of Half-Gates in npRPM

We prove that half-gates is adaptively secure in the npRPM. This scheme can be implemented in Figure 1, which slightly differs from the original one of [ZRE15] in including decoding table d in garbled input $\widehat{x}$ rather than garbled circuit $\widehat{f}$. We will see in Appendix E that this difference is required to follow the online-complexity lower bound in the npRPM.

Theorem 4. *Let $\mathsf{H}(X,k) = \pi(X \oplus k) \oplus \sigma(X \oplus k)$ be a tweakable hash function where $X, k \in \{0,1\}^{\lambda}$, $\pi \in \mathcal{P}_{\lambda}$ is random permutation, and $\sigma : \{0,1\}^{\lambda} \to \{0,1\}^{\lambda}$ is a linear orthomorphism. Then, half-gates*

$\underline{\mathsf{SimF}^{\pi^{\pm 1}(\cdot)}(f)}$:

$F := \{(G_0^g, G_1^g)\}_{g\in\mathcal{G}_{\mathsf{and}}(f)} \leftarrow (\{0,1\}^{2\lambda})^{|f|}$
$\{X_i\}_{i\in\mathcal{W}_{\mathsf{in}}(f)} \leftarrow (\{0,1\}^{\lambda})^{|\mathcal{W}_{\mathsf{in}}(f)|}$
for $g \in \mathcal{G}(f)$ in topology order **do**
  $(a, b, c) := (\mathsf{in}_0(g), \mathsf{in}_1(g), \mathsf{out}(g))$
  if $\mathsf{type}(g) = \texttt{XOR}$ **then** $X_c := X_a \oplus X_b$
  else if $\mathsf{type}(g) = \texttt{AND}$ **then**
    $k_0^g := 2 \cdot g - 1, k_1^g := 2 \cdot g$
    $s_a := \mathsf{lsb}(X_a), s_b := \mathsf{lsb}(X_b)$
    $U_0^g := \pi(X_a \oplus k_0^g) \oplus \sigma(X_a \oplus k_0^g), U_1^g := \pi(X_b \oplus k_1^g) \oplus \sigma(X_b \oplus k_1^g)$
    $X_c := U_0^g \oplus U_1^g \oplus s_a G_0^g \oplus s_b(G_1^g \oplus X_a)$
return $\widehat{f} := (f, F)$, $\mathsf{st}_{\mathsf{sim}} := (f, \widetilde{X} := \{X_i\}_{i\in\mathcal{W}(f)}, \widetilde{U} := \{U_0^g, U_1^g\}_{g\in\mathcal{G}_{\mathsf{and}}(f)})$

$\underline{\mathsf{SimIn}^{\pi^{\pm 1}(\cdot)}(f(x))}$:

Parse $\mathsf{st}_{\mathsf{sim}} = (f, \widetilde{X} = \{X_i\}_{i\in\mathcal{W}(f)}, \widetilde{U})$
for $i \in \mathcal{W}_{\mathsf{out}}(f)$ **do** $d_i := f(x)_i \oplus \mathsf{lsb}(X_i)$
return $\widehat{x} := (\{X_i\}_{i\in\mathcal{W}_{\mathsf{in}}(f)}, d), \widetilde{X}, \widetilde{U}$.

Figure 2: Our simulator for half-gates in the npRPM.

(Figure 1) is a λ-garbling scheme with (q, s, ε)-adaptive security in the npRPM, where $\varepsilon = (16qs + 38s^2)/2^{\lambda}$.

Proof. The correctness is given by the proof [ZRE15] as postponing decoding table d does not affect correctness. We only need to consider the simulation.

Our simulator $\mathsf{Sim} = (\mathsf{SimF}, \mathsf{SimIn})$ is presented in Figure 2 and is obviously PPT. Then, we prove this theorem using the following three hybrids:

- Hybrid_0. This is the adaptive experiment using simulator Sim.
- Hybrid_1. This is identical to the previous hybrid, except that we replace $\pi^{\pm 1}$ (which can be equivalently emulated on-the-fly as in Figure 3) by an approximation $\widetilde{\pi}^{\pm 1}$ (given in Figure 4). This approximation is the same as random permutation except that, for a new query of the simulator, it returns a fresh random string as response and records this query-response pair. This hybrid is used to simplify probability analysis.
- Hybrid_2. This is the adaptive experiment using half-gates scheme.

Using Corollary 1 for the indistinguishability between Hybrid_0 and Hybrid_1, and Lemma 3 for that between Hybrid_1 and Hybrid_2, this theorem holds. □

Lemma 2. *Let $\widetilde{\mathcal{P}}_{\ell(\lambda)}$ denote the distribution of $\widetilde{\pi}^{\pm 1}$ in Figure 4 for $n(\lambda) \in \mathbb{N}^+$ queries. For every PPT simulator $\mathsf{Sim} = (\mathsf{SimF}, \mathsf{SimIn})$ making $n_{\mathsf{sim}}(\lambda) \le n(\lambda)$ queries, every auxiliary input $z \in \{0,1\}^*$, and every computationally unbounded adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3)$ making $n(\lambda) - n_{\mathsf{sim}}(\lambda)$ queries, it holds*

Initialize a list $Q_0 = \varnothing$.
Sample uniform[a] $c_1, \ldots, c_{n(\lambda)} \leftarrow \{0,1\}^{\ell(\lambda)}$.
for $i \in [1, n(\lambda)]$ **do**
  if π is queried with input $\alpha_i \in \{0,1\}^{\ell(\lambda)}$ **then**
    if $\exists(\alpha_i, \gamma_i) \in Q_{i-1}$ **then** Return γ_i as response.
    if $(\ldots, c_i) \notin Q_{i-1}$ **then** $\gamma_i := c_i$.
    else Sample uniform[b] $\gamma_i \leftarrow \{s_i \in \{0,1\}^{\ell(\lambda)} \mid (\ldots, s_i) \notin Q_{i-1}\}$.
    Return γ_i as response and define $Q_i := Q_{i-1} \cup \{(\alpha_i, \gamma_i)\}$.
  else if π^{-1} is queried with input $\beta_i \in \{0,1\}^{\ell(\lambda)}$ **then**
    if $\exists(\gamma_i, \beta_i) \in Q_{i-1}$ **then** Return γ_i as response.
    if $(c_i, \ldots) \notin Q_{i-1}$ **then** $\gamma_i := c_i$.
    else Sample uniform[b] $\gamma_i \leftarrow \{s_i \in \{0,1\}^{\ell(\lambda)} \mid (s_i, \ldots) \notin Q_{i-1}\}$.
    Return γ_i as response and define $Q_i := Q_{i-1} \cup \{(\gamma_i, \beta_i)\}$.

[a] A uniform random tape r_π is used.
[b] Typically, a uniform random tape r^*_π is used to run rejection sampling.

Figure 3: The workflow of oracle $\pi^{\pm 1}(\cdot)$ up to $n(\lambda)$ queries.

that

$$\left| \Pr\left[\begin{array}{l} \pi \leftarrow \mathcal{P}_{\ell(\lambda)}, \pi^{-1} := \mathsf{inv}(\pi), \\ \quad (f, \mathsf{st}_1) \leftarrow {\mathcal{A}_1}^{\pi^{\pm 1}(\cdot)}(z), \\ \quad \widehat{f} \leftarrow {\mathsf{SimF}}^{\pi^{\pm 1}(\cdot)}(\Phi(f)), : {\mathcal{A}_3}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x}) = 1 \\ (x, \mathsf{st}_2) \leftarrow {\mathcal{A}_2}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_1, \widehat{f}), \\ \quad \widehat{x} \leftarrow {\mathsf{SimIn}}^{\pi^{\pm 1}(\cdot)}(f(x)) \end{array}\right] - \Pr\left[\begin{array}{l} \widetilde{\pi}^{\pm 1} \leftarrow \widetilde{\mathcal{P}}_{\ell(\lambda)}, \\ \quad (f, \mathsf{st}_1) \leftarrow {\mathcal{A}_1}^{\widetilde{\pi}^{\pm 1}(\cdot)}(z), \\ \quad \widehat{f} \leftarrow {\mathsf{SimF}}^{\widetilde{\pi}^{\pm 1}(\cdot)}(\Phi(f)), : {\mathcal{A}_3}^{\widetilde{\pi}^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x}) = 1 \\ (x, \mathsf{st}_2) \leftarrow {\mathcal{A}_2}^{\widetilde{\pi}^{\pm 1}(\cdot)}(\mathsf{st}_1, \widehat{f}), \\ \quad \widehat{x} \leftarrow {\mathsf{SimIn}}^{\widetilde{\pi}^{\pm 1}(\cdot)}(f(x)) \end{array}\right] \right| \leq \frac{n(\lambda) \cdot n_{\mathsf{sim}}(\lambda)}{2^{\ell(\lambda)-1}}.$$

Proof. Let $\mathcal{N} \subseteq [1, n(\lambda)]$ denote the index set of the queries made by Sim such that $|\mathcal{N}| = n_{\mathsf{sim}}(\lambda)$. Let $\mathsf{true}(\pi)$ denote the event that ${\mathcal{A}_3}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x}) = 1$ in the first experiment, $\mathsf{true}(\widetilde{\pi})$ denote the counterpart in the second one, and bad denote the event that $\vee_{i \in \mathcal{N}}((\ldots, c_i) \in Q_{i-1} \vee (c_i, \ldots) \in Q_{i-1})$. We can study the two experiments under the same probability space, i.e., they use the same literal values of (r_π, r^*_π) and the random tapes of Sim and $\mathcal{A}$.

We can see that $\mathsf{true}(\pi) \wedge \neg\mathsf{bad}$ occurs if and only if $\mathsf{true}(\widetilde{\pi}) \wedge \neg\mathsf{bad}$ occurs, i.e., the two experiments proceed identically unless bad occurs. Thus, it follows from the Difference Lemma that

$$\left|\Pr[\mathsf{true}(\pi)] - \Pr[\mathsf{true}(\widetilde{\pi})]\right| \leq \Pr[\mathsf{bad}] \leq \sum_{i \in \mathcal{N}} \frac{2\,|Q_{i-1}|}{2^{\ell(\lambda)}} \leq \frac{n(\lambda) \cdot n_{\mathsf{sim}}(\lambda)}{2^{\ell(\lambda)-1}},$$

which completes this proof. □

Corollary 1. *Let* Sim *be defined as in Figure 2. Then, for every auxiliary input* $z \in \{0,1\}^*$ *and every computationally unbounded adversary* $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3)$ *totally making* q *oracle queries and choosing a circuit with* s `AND` *gates,* $\mathcal{A}(z)$ *can distinguish* Hybrid_0 *and* Hybrid_1 *with advantage at most* $(2qs + 4s^2)/2^{\lambda-1}$.

1: Initialize a list $Q_0 = \varnothing$.
2: Sample uniform $c_1, \ldots, c_{n(\lambda)} \leftarrow \{0,1\}^{\ell(\lambda)}$.
3: **for** $i \in [1, n(\lambda)]$ **do**
4: **if** $\widetilde{\pi}$ is queried with input $\alpha_i \in \{0,1\}^{\ell(\lambda)}$ from Sim **then**
5: **if** $\exists(\alpha_i, \gamma_i) \in Q_{i-1}$ **then** Return γ_i as response.
6: Return $\gamma_i := c_i$ as response and define $Q_i := Q_{i-1} \cup \{(\alpha_i, \gamma_i)\}$.
7: **else if** $\widetilde{\pi}^{-1}$ is queried with input $\beta_i \in \{0,1\}^{\ell(\lambda)}$ from Sim **then**
8: **if** $\exists(\gamma_i, \beta_i) \in Q_{i-1}$ **then** Return γ_i as response.
9: Return $\gamma_i := c_i$ as response and define $Q_i := Q_{i-1} \cup \{(\gamma_i, \beta_i)\}$.
10: **else if** $\widetilde{\pi}$ is queried with input $\alpha_i \in \{0,1\}^{\ell(\lambda)}$ from $\mathcal{A}$ **then**
11: Same as step 5 to 8 in the workflow of oracle $\pi^{\pm 1}(\cdot)$ (Figure 3).
12: **else if** $\widetilde{\pi}^{-1}$ is queried with input $\beta_i \in \{0,1\}^{\ell(\lambda)}$ from $\mathcal{A}$ **then**
13: Same as step 10 to 13 in the workflow of oracle $\pi^{\pm 1}(\cdot)$ (Figure 3).

Figure 4: The workflow of approximate oracle $\widetilde{\pi}^{\pm 1}(\cdot)$ up to $n(\lambda)$ queries, where the differences are highlighted in box.

Proof. This corollary follows from Lemma 2 by using $\ell(\lambda) = \lambda, n_{\mathsf{sim}}(\lambda) = 2|f|$, and $n(\lambda) = q + n_{\mathsf{sim}}(\lambda)$, given q queries of $\mathcal{A}$ and $2|f| = 2s$ queries of Sim. □

Then, we will use the H-coefficient technique (Section 3.4) with "transcript padding" to bound the advantage of distinguishing Hybrid_1 and Hybrid_2.

Lemma 3. *Let* Sim *be defined as in Figure 2. Then, for every auxiliary input $z \in \{0,1\}^*$ and every computationally unbounded adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3)$ totally making q oracle queries and choosing a circuit with s AND gates, $\mathcal{A}(z)$ can distinguish Hybrid_1 and Hybrid_2 with advantage at most $(12qs + 30s^2)/2^\lambda$.*

Proof. Fix z and $\mathcal{A}$. We regard Hybrid_1 (resp., Hybrid_2) as the ideal (resp., real) world in the H-coefficient technique, where the computationally unbounded non-uniform adversary $\mathcal{A} = \mathcal{A}(z)$ and $\varepsilon_1, \varepsilon_2$ are computed as follows.

Transcript padding. In either world, $\mathcal{A}$ will interact with an integrated oracle that acts as the two-round challenger in the adaptive experiment and provides interfaces $\pi^{*\pm 1} \in \{\pi^{\pm 1}, \widetilde{\pi}^{\pm 1}\}$ for forward/backward permutation queries. Here, $\mathcal{A}$ can learn $\pi^*(\alpha) = \beta$ if and only if it sent forward query α to π^* and received response β, or sent backward query β to π^{*-1} and received response α.

To compute $\varepsilon_1, \varepsilon_2$ more easily, we ask the oracle to send more messages to $\mathcal{A}$ and $\mathcal{A}$ to make extra queries (in addition to the supposed q queries) in both two worlds. More specifically,

- Upon receiving x from $\mathcal{A}$, the oracle sends $(\widetilde{X} := \{X_i\}_{i \in \mathcal{W}(f)}, d)$ rather than $\widehat{x} := (\{X_i\}_{i \in \mathcal{W}_{\mathsf{in}}(f)}, d)$ to $\mathcal{A}$. In addition to the active input labels given by $\widehat{x}$, the former also gives the active internal and output labels. In the real world, the oracle can run $\mathsf{HG.DecEval}^{\pi^{\pm 1}(\cdot)}$, which determines other active labels in $\widetilde{X}$. In the ideal world, this $\widetilde{X}$ can be directly output by SimIn.
- Along with $(\widetilde{X}, d)$, the oracle sends $\widetilde{x} := \{x_i\}_{i \in \mathcal{W}(f)}$ to $\mathcal{A}$, which are the wire truth values in the evaluation of $f(x)$. Both two oracles "echo" these values, which are self-evident to $\mathcal{A}$, to explicitly include them in transcripts. In the experiment, the real-world oracle uses $x = \{x_i\}_{i \in \mathcal{W}_{\mathsf{in}}(f)}$ in $\mathsf{HG.Encode}^{\pi^{\pm 1}(\cdot)}$, but the ideal-world oracle can only use $f(x) = \{x_i\}_{i \in \mathcal{W}_{\mathsf{out}}(f)}$ in SimIn.
- Along with $(\widetilde{X}, d)$, the oracle sends $\widetilde{U} := \{U_0^g, U_1^g\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}$ to $\mathcal{A}$. In the real world, the oracle computes $U_0^g := \mathsf{H}(X_a, k_0^g)$ and $U_1^g := \mathsf{H}(X_b, k_1^g)$ for each $g \in \mathcal{G}_{\mathsf{and}}(f)$ with $(a, b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$. In the ideal world, this $\widetilde{U}$ is output by SimIn and is essentially the same hash outputs.

- **(Extra queries)** Upon receiving $(\widetilde{X}, d, \widetilde{x}, \widetilde{U})$ from the oracle, $\mathcal{A}$ also makes a forward permutation query $X_a \oplus k_0^g$ (resp., $X_b \oplus k_1^g$) for each $g \in \mathcal{G}_{\mathsf{and}}(f)$ with $(a, b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$, if it has never learned $\pi^*(X_a \oplus k_0^g) = Y$ (resp., $\pi^*(X_b \oplus k_1^g) = Y$) for some Y in its interaction with $\pi^{*\pm 1} \in \{\pi^{\pm 1}, \widetilde{\pi}^{\pm 1}\}$.
- At the end of the experiment (i.e., once all other transcripts are settled), the oracle sends Δ to $\mathcal{A}$. In the real world, the oracle gets this Δ from the output of $\mathsf{HG.Garble}^{\pi^{\pm 1}(\cdot)}$. In the ideal world, Δ is dummy and sampled by the oracle at this time (note that Sim does not use Δ).

According to the two oracle constructions, real-world sample space

$$\Omega_{\mathsf{real}} = \{0,1\}^{\lambda-1} \times \{0,1\}^{|\mathcal{W}_{\mathsf{in}}(f)|\lambda} \times \mathcal{P}_\lambda,$$

and ideal-world sample space

$$\Omega_{\mathsf{ideal}} = (\{0,1\}^{2\lambda})^{|f|} \times (\{0,1\}^{\lambda})^{|\mathcal{W}_{\mathsf{in}}(f)|} \times \underbrace{\{0,1\}^*}_{\substack{\text{random tape for} \\ \text{the sampling in } \widetilde{\pi}^{\pm 1}(\cdot)}} \times \underbrace{\{0,1\}^{\lambda-1}}_{\text{dummy } \Delta}.$$

Given the oracle constructions, a transcript in the original adaptive experiment will be padded with more literal values. Note that transcript padding will not lower the advantage of $\mathcal{A}$ since $\mathcal{A}$ can discard the padding values at will. With the padding, a transcript is of the form:

$$\tau = (\mathcal{K}_1, (f, \widehat{f}), \mathcal{K}_2, (x, (\widetilde{X}, d, \widetilde{x}, \widetilde{U})), \mathcal{K}_3, \Delta),$$

where $\mathcal{K}_1, \mathcal{K}_2$, and $\mathcal{K}_3$ are the ordered lists of query-response pairs seen in the interleaved interaction with permutation oracles. We do not explicitly consider query direction in these pairs. Given $\pi^{*\pm 1} \in \{\pi^{\pm 1}, \widetilde{\pi}^{\pm 1}\}$, $\mathcal{A}$ learns $\pi^*(\alpha) = \beta$ if and only if there exists $(\alpha, \beta) \in \cup_{\ell=1}^3 \mathcal{K}_\ell$.

Let $q_\ell := |\mathcal{K}_\ell|$ for every $\ell \in \{1,2,3\}$ and $q_\Sigma := \sum_{\ell=1}^3 q_\ell$. It follows from the extra queries that $q_\Sigma \le q + 2|f|$. Without loss of generality, we assume that $\mathcal{A}$ only makes *non-repeating* queries, i.e., it never makes forward query α to π^* or backward query β to π^{*-1} for any learned permutation entry (α, β).

Bad transcripts. A transcript $\tau \in \mathcal{T}_{\mathsf{bad}}$ if and only if it incurs at least one of the following events[7]:

- bad_1. There exist distinct $(g,u), (g',u') \in \mathcal{G}_{\mathsf{and}}(f) \times \{0,1\}$ such that

$$X_w \oplus k_u^g = X_{w'} \oplus k_{u'}^{g'} \quad \vee \tag{2}$$
$$U_u^g \oplus \sigma(X_w \oplus k_u^g) = U_{u'}^{g'} \oplus \sigma(X_{w'} \oplus k_{u'}^{g'}) \tag{3}$$

  where $w := \mathsf{in}_u(g)$ and $w' := \mathsf{in}_{u'}(g')$, or

  There exists $g \in \mathcal{G}_{\mathsf{and}}(f)$ such that

$$(s_b \oplus x_b)\Delta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g) = x_a\Delta \oplus X_a \oplus G_1^g \oplus U_1^g \oplus \sigma(X_b \oplus k_1^g) \tag{4}$$

  where $(a, b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$, or

  There exist distinct $g, g' \in \mathcal{G}_{\mathsf{and}}(f)$ such that

$$(s_b \oplus x_b)\Delta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g) = (s_{b'} \oplus x_{b'})\Delta \oplus G_0^{g'} \oplus U_0^{g'} \oplus \sigma(X_{a'} \oplus k_0^{g'}) \quad \vee \tag{5}$$
$$x_a\Delta \oplus X_a \oplus G_1^g \oplus U_1^g \oplus \sigma(X_b \oplus k_1^g) = x_{a'}\Delta \oplus X_{a'} \oplus G_1^{g'} \oplus U_1^{g'} \oplus \sigma(X_{b'} \oplus k_1^{g'}) \quad \vee \tag{6}$$
$$(s_b \oplus x_b)\Delta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g) = x_{a'}\Delta \oplus X_{a'} \oplus G_1^{g'} \oplus U_1^{g'} \oplus \sigma(X_{b'} \oplus k_1^{g'}) \quad \vee \tag{7}$$
$$x_a\Delta \oplus X_a \oplus G_1^g \oplus U_1^g \oplus \sigma(X_b \oplus k_1^g) = (s_{b'} \oplus x_{b'})\Delta \oplus G_0^{g'} \oplus U_0^{g'} \oplus \sigma(X_{a'} \oplus k_0^{g'}) \tag{8}$$

[7] Strictly speaking, such bad_i's are the disjunctive predicates of $\tau \in \mathcal{T}_{\mathcal{A}}$ to define, in set-builder notation, a set $\mathcal{T}_{\mathsf{bad}} \subseteq \mathcal{T}_{\mathcal{A}}$ of bad transcripts. We treat them as "events" to avoid cumbersome notation. The formal events (i.e., the specific sets of oracles) w.r.t. bad_i's are well-defined from $\mathcal{T}_{\mathsf{bad}}$ and the two random variables in Section 3.4.

where $(a, b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$ and $(a', b') := (\mathsf{in}_0(g'), \mathsf{in}_1(g'))$.

In this case, $\mathcal{A}$ can check the consistency between the value of $G_u^g \oplus G_{u'}^{g'}$ and that of Δ at the end of experiment *without* further sending required queries, which are computed from Δ, to a random permutation or its inverse. In the real world, the consistency certainly holds. However, the ideal-world garbled rows and Δ are independently sampled, leading to the consistency only with negligible probability. So, $\mathcal{A}$ has non-negligible advantage to distinguish the two worlds and the statistical distance, as an upper bound, also blows up.

More specifically, the real world is as follows in this case. The pre-image collision (2) leads to the syntactically same XOR of two hash masks in $G_u^g, G_{u'}^{g'}$, which can be XORed to cancel all hash masks to check the consistency with Δ without further queries. Moreover, the image collision (3) also implies the pre-image collision (2) since π is permutation. The other equalities imply the image collision $\pi(X_w \oplus \Delta \oplus k_u^g) = \pi(X_{w'} \oplus \Delta \oplus k_{u'}^{g'})$ for some distinct tuple $(g, u), (g', u') \in \mathcal{G}_{\mathsf{and}}(f) \times \{0, 1\}$, $w := \mathsf{in}_u(g)$, and $w' := \mathsf{in}_{u'}(g')$. Given permutation π, this collision implies the pre-image collision (2), which can be used to see the consistency. However, the above cancelling of hash masks will not give this consistency except with negligible probability in the ideal world.

- bad_2. There exists $((\alpha, \beta), g) \in \cup_{\ell=1}^{3} \mathcal{K}_\ell \times \mathcal{G}_{\mathsf{and}}(f)$ such that

$$\alpha = X_a \oplus \Delta \oplus k_0^g \quad \vee \tag{9}$$
$$\alpha = X_b \oplus \Delta \oplus k_1^g \quad \vee \tag{10}$$
$$\beta = \sigma(\Delta) \oplus (s_b \oplus x_b)\Delta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g) \quad \vee \tag{11}$$
$$\beta = \sigma(\Delta) \oplus x_a\Delta \oplus X_a \oplus G_1^g \oplus U_1^g \oplus \sigma(X_b \oplus k_1^g) \tag{12}$$

  where $(a, b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$.

  In this case, $\mathcal{A}$ essentially makes it to guess Δ before receiving this value. It allows $\mathcal{A}$ to distinguish the real world, where every G_i^g is consistent with Δ, and the ideal world with a dummy Δ. So, the statistical distance blows up.

- bad_3. There exists $((\alpha, \beta), g) \in \cup_{\ell=1}^{2} \mathcal{K}_\ell \times \mathcal{G}_{\mathsf{and}}(f)$ such that

$$\alpha = X_a \oplus k_0^g \quad \vee \tag{13}$$
$$\alpha = X_b \oplus k_1^g \quad \vee \tag{14}$$
$$\beta = U_0^g \oplus \sigma(X_a \oplus k_0^g) \quad \vee \tag{15}$$
$$\beta = U_1^g \oplus \sigma(X_b \oplus k_1^g) \tag{16}$$

  where $(a, b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$.

  In this case, $\mathcal{A}$ can make forward/backward queries w.r.t. some active labels before receiving active input labels and computing other active ones. We use this case to explicitly ensure that these query-response pairs are fixed by the queries to $\pi^{\pm 1}(\cdot)$ in the step 10 to 12 of $\mathsf{HG.Garble}^{\pi^{\pm 1}(\cdot)}$ (when τ is produced in the real world) or the queries to $\widetilde{\pi}^{\pm 1}(\cdot)$ in the step 9 of $\mathsf{SimIn}^{\widetilde{\pi}^{\pm 1}(\cdot)}$ (when τ is produced in the ideal world), instead of the extra queries of $\mathcal{A}$.

  This case is used to simplify probability analysis.

Bounding $1 - \varepsilon_2$. Without loss of generality, we can consider some fixed good transcript τ such that $\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] \neq 0$ (if this probability is zero, it is trivial by definition that $\varepsilon_2 = 0$ for this τ). Using Lemma 1, we turn to analyze the sampled oracle's compatibility (Section 3.4) with such a transcript, instead of the interaction between $\mathcal{A}$ and the sampled oracle.

Note that there is a computationally unbounded non-uniform adversary $\mathcal{A}'$ such that, for every oracle ω, it sends the queries in τ in order in its interaction with ω (e.g., $\mathcal{A}'$ has auxiliary input τ and sends its ordered queries). Fix $\mathcal{A}'$ in the following compatibility analysis so that any real-world or

ideal-world oracle will receive the queries in τ in order. For a response c recorded in a fixed τ, let $\omega \vdash c$ denote the event that, fixing the ordered queries as per τ, oracle ω produces c given the corresponding query. Let $\mathcal{K}^{\mathsf{R}}$ denote the order-preserving list of the responses in an ordered list $\mathcal{K}$ of permutation query-response pairs.

First, we compute $\Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)]$. Following from half-gates, a real-world oracle $\omega = (\Delta, \{W_i\}_{i \in \mathcal{W}_{\mathsf{in}}(f)}, \pi) \in \Omega_{\mathsf{real}}$. It holds that

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)] = \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, \widehat{f}, \mathcal{K}_2^{\mathsf{R}}, \widetilde{X}, d, \widetilde{x}, \widetilde{U}, \mathcal{K}_3^{\mathsf{R}}, \Delta)].$$

To begin with, every real-world ω certainly produces f' (i.e., the first value in $\widehat{f}$) and $\widetilde{x}$ fixed in τ, which leads to $\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] \neq 0$. This non-zero probability implies that $(f', \widetilde{x})$ in τ is honestly and deterministically computed from the fixed queries (f, x). Otherwise, no ideal-world oracle, which computes $(f', \widetilde{x})$ from the same deterministic procedure, can produce this transcript, contradicting the non-zero probability. As every real-world ω will follow the same deterministic procedure, it certainly produces the two values.

Meanwhile, a real-world oracle ω should have the same literal value of Δ as its counterpart in τ. Conditioned on the compatibility so far, the probability

$$\begin{aligned}
&\Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)] \\
&\quad = \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, F, \mathcal{K}_2^{\mathsf{R}}, \widetilde{X}, d, \widetilde{U}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (f', \widetilde{x}) \wedge \omega \vdash \Delta\right] \\
&\qquad \cdot \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}\left[\omega \vdash (f', \widetilde{x}) \wedge \omega \vdash \Delta\right] \\
&\quad = \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, F, \mathcal{K}_2^{\mathsf{R}}, \widetilde{X}, d, \widetilde{U}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (f', \widetilde{x}) \wedge \omega \vdash \Delta\right] \\
&\qquad \cdot \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}\left[\omega \vdash (f', \widetilde{x})\right] \cdot \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[\omega \vdash \Delta] \\
&\quad = \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, F, \mathcal{K}_2^{\mathsf{R}}, \widetilde{X}, d, \widetilde{U}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (f', \widetilde{x}) \wedge \omega \vdash \Delta\right] \cdot \frac{1}{2^{\lambda-1}}.
\end{aligned}$$

Conditioned on the compatibility with $(f', \widetilde{x}, \Delta)$, a real-world ω should also be compatible with $(\cup_{\ell=1}^{3} \mathcal{K}_\ell^{\mathsf{R}}, F, \widetilde{U})$ and some active labels in $\widetilde{X}$ such that

(i) $\pi^{\pm 1}$ maps the fixed permutation queries to the responses in $\cup_{\ell=1}^{3} \mathcal{K}_\ell^{\mathsf{R}}$.

(ii) For each $i \in \mathcal{W}_{\mathsf{in}}(f)$, it holds that $X_i = W_i \oplus x_i \Delta$.

(iii) For each $g \in \mathcal{G}_{\mathsf{and}}(f)$ with $(a, b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$, it holds that

$$\begin{aligned}
\mathsf{H}(X_a, k_0^g) &:= \pi(X_a \oplus k_0^g) \oplus \sigma(X_a \oplus k_0^g) = U_0^g, \\
\mathsf{H}(X_b, k_1^g) &:= \pi(X_b \oplus k_1^g) \oplus \sigma(X_b \oplus k_1^g) = U_1^g.
\end{aligned} \tag{17}$$

(iv) For each $g \in \mathcal{G}_{\mathsf{and}}(f)$ with $(a, b, c) := (\mathsf{in}_0(g), \mathsf{in}_1(g), \mathsf{out}(g))$, it holds that

$$X_c = \mathsf{H}(X_a, k_0^g) \oplus \mathsf{H}(X_b, k_1^g) \oplus s_a G_0^g \oplus s_b (G_1^g \oplus X_a), \tag{18}$$

$$\begin{aligned}
G_0^g &= \mathsf{H}(X_a, k_0^g) \oplus \mathsf{H}(X_a \oplus \Delta, k_0^g) \oplus (s_b \oplus x_b)\Delta \\
G_1^g &= \mathsf{H}(X_b, k_1^g) \oplus \mathsf{H}(X_b \oplus \Delta, k_1^g) \oplus x_a \Delta \oplus X_a,
\end{aligned} \tag{19}$$

where the bits $x_a, x_b, s_a = \mathsf{lsb}(X_a), s_b = \mathsf{lsb}(X_b)$ are given in τ.

Conditioned on the compatibility so far, every real-world oracle ω is always compatible with the leftover values in τ, i.e., decoding table d and other active labels in $\widetilde{X}$, which are deterministically computed from XOR combination. The reason is that, for τ ensuring $\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] \neq 0$, these values should be honestly determined by the conditioned values as in the real world. Otherwise, this

probability will be zero for an ideal-world oracle, which obtains them from a consistent deterministic computation as per the conditioned values. As every real-world oracle ω honestly follows the real-world computation, this "leftover" compatibility must hold conditioned on the previous compatibility.

It remains to compute the conditional probabilities for (i) to (iv). Consider (iii) and (iv). We note that every good τ with $\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] \neq 0$ already implies (17) and (18). Otherwise, no ideal-world oracle, which computes these values according to the step 8 and 9 in SimIn (using the given tweakable hash function), can produce τ, contradicting the non-zero probability.

Then, consider (19), the leftover part of (iii) and (iv). We rewrite (19) as:

$$\mathcal{V} := \left\{ \begin{array}{l} g \in \mathcal{G}_{\mathsf{and}}(f), (a,b) := (\mathsf{in}_0(g), \mathsf{in}_1(g)) : \\ \qquad \underbrace{\pi(X_a \oplus k_0^g)}_{P_{g,0}} \oplus \underbrace{\pi(X_a \oplus \Delta \oplus k_0^g)}_{P_{g,2}} = \sigma(\Delta) \oplus (s_b \oplus x_b)\Delta \oplus G_0^g, \\ \qquad \underbrace{\pi(X_b \oplus k_1^g)}_{P_{g,1}} \oplus \underbrace{\pi(X_b \oplus \Delta \oplus k_1^g)}_{P_{g,3}} = \sigma(\Delta) \oplus x_a\Delta \oplus X_a \oplus G_1^g \end{array} \right\}$$

As τ is a good transcript, there are exactly $4|f|$ pairwise distinct permutation pre-images on the left hand (otherwise, there will be a pair of permutation pre-images leading to (2) in bad_1 or a permutation pre-image leading to (9) $\vee$ (10) in bad_2 given the extra queries). $\mathcal{V}$ has exact $4|f|$ syntactically different variables $\mathcal{P} := \{P_{g,0}, P_{g,1}, P_{g,2}, P_{g,3}\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}$. They fix the same number of the entries of permutation π in a real-world ω *if and only if their literal values fixed by* τ *are also pairwise distinct.* Note that every good transcript τ indeed fixes *exact one* such assignment of these values for the following reasons:

- (17) already holds for τ, i.e., for $g \in \mathcal{G}_{\mathsf{and}}(f)$ with $(a,b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$,

$$\begin{aligned} P_{g,0} &:= \pi(X_a \oplus k_0^g) = U_0^g \oplus \sigma(X_a \oplus k_0^g), \\ P_{g,1} &:= \pi(X_b \oplus k_1^g) = U_1^g \oplus \sigma(X_b \oplus k_1^g). \end{aligned} \tag{20}$$

  The literal values of $\{P_{g,0}, P_{g,1}\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}$ can be fixed by the responses in $\mathcal{K}_3^{\mathsf{R}}$ given the extra queries and will be pairwise distinct according to the impossible (3) from $\neg\mathsf{bad}_1$.

- For $g \in \mathcal{G}_{\mathsf{and}}(f)$ with $(a,b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$, $\{P_{g,2}, P_{g,3}\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}$ are literally assigned the following values fixed by τ according to $\mathcal{V}$ and (20):

$$\begin{aligned} P_{g,2} &:= \pi(X_a \oplus \Delta \oplus k_0^g) = \sigma(\Delta) \oplus (s_b \oplus x_b)\Delta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g), \\ P_{g,3} &:= \pi(X_b \oplus \Delta \oplus k_1^g) = \sigma(\Delta) \oplus x_a\Delta \oplus X_a \oplus G_1^g \oplus U_1^g \oplus \sigma(X_b \oplus k_1^g). \end{aligned} \tag{21}$$

  Clearly, one can see that the literal values of $\{P_{g,2}, P_{g,3}\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}$ are pairwise distinct according to the impossible (4) $\vee$ (5) $\vee$ (6) $\vee$ (7) $\vee$ (8) from $\neg\mathsf{bad}_1$.

- The goodness of τ also ensures that there do not exist

$$P' \in \{P_{g,0}, P_{g,1}\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}, P'_\Delta \in \{P_{g,2}, P_{g,3}\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}$$

  such that $P' = P'_\Delta$. Otherwise, this equality and (21) ensure that there exist $(g,u) \in \mathcal{G}_{\mathsf{and}}(f) \times \{0,1\}$ and $g' \in \mathcal{G}_{\mathsf{and}}(f)$ such that

$$\begin{aligned} \pi(X_w \oplus k_u^g) &= \pi(X_{a'} \oplus \Delta \oplus k_0^{g'}) \\ &= \sigma(\Delta) \oplus (s_{b'} \oplus x_{b'})\Delta \oplus G_0^{g'} \oplus U_0^{g'} \oplus \sigma(X_{a'} \oplus k_0^{g'}) \quad \vee \\ \pi(X_w \oplus k_u^g) &= \pi(X_{b'} \oplus \Delta \oplus k_1^{g'}) \\ &= \sigma(\Delta) \oplus x_{a'}\Delta \oplus X_{a'} \oplus G_1^{g'} \oplus U_1^{g'} \oplus \sigma(X_{b'} \oplus k_1^{g'}) \end{aligned} \tag{22}$$

  where $w := \mathsf{in}_u(g)$ and $(a', b') := (\mathsf{in}_0(g'), \mathsf{in}_1(g'))$. Recall that $\neg\mathsf{bad}_3$ and the extra queries implies $(X_w \oplus k_u^g, \pi(X_w \oplus k_u^g)) \in \mathcal{K}_3$. As a result, (22) leads to contradiction with the impossible (11) $\vee$ (12) from $\neg\mathsf{bad}_2$.

Putting these cases together, we can see that τ yields a value assignment of $\mathcal{P}$, and this assignment fixes exact $4\,|f|$ entries of real-world permutation π.

Conditioned on $\neg\mathsf{bad}_1 \wedge \neg\mathsf{bad}_3$ of good transcript τ, $2\,|f|$ extra queries are non-repeating and the number of non-repeating queries is $q + 2\,|f| = q_\Sigma$. So, $2\,|f|$ responses in $\cup_{\ell=1}^{3}\mathcal{K}_\ell^{\mathsf{R}}$ for the non-repeating extra queries are fixed by the values in $\mathcal{P}$ while the other $q_\Sigma - 2\,|f| = q$ responses are fixed by real-world π (conditioned on the values in $\mathcal{P}$). Using (i), (ii), (iii), and (iv) together with the "leftover" compatibility, we have in the real world that

$$\begin{aligned}
&\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, F, \mathcal{K}_2^{\mathsf{R}}, \widetilde{X}, d, \widetilde{U}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (f', \widetilde{x}) \wedge \omega \vdash \Delta\right] \\
&\quad = \frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}} \cdot \frac{(2^\lambda - 4\,|f| - (q_\Sigma - 2\,|f|))!}{(2^\lambda)!} \\
&\quad = \frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}} \cdot \frac{1}{(2^\lambda)_{q+4|f|}}, \\
\Rightarrow &\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)] = \frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}} \cdot \frac{1}{(2^\lambda)_{q+4|f|}} \cdot \frac{1}{2^{\lambda-1}}.
\end{aligned}$$

Second, in the ideal world, condition $\neg\mathsf{bad}_3$ ensures that the responses for the $2\,|f|$ extra queries are fixed by the queries in the step 9 of $\mathsf{SimIn}^{\widetilde{\pi}^{\pm 1}(\cdot)}$. So, each U_u^g is independently uniform according to the randomness of $\widetilde{\pi}^{\pm 1}$ and the pairwise distinct pre-images by $\neg\mathsf{bad}_1$. From the garbled rows, the active input labels, and these U_u^g, the active internal and output labels (as well as decoding table d) are fixed in topology order. So, we have

$$\begin{aligned}
\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\omega \in \mathsf{comp}_{\mathsf{ideal}}(\tau)] &= \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, \mathcal{K}_2^{\mathsf{R}}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (\widehat{f}, \widetilde{X}, d, \widetilde{x}, \widetilde{U}, \Delta)\right] \\
&\quad \cdot \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\omega \vdash (\widehat{f}, \widetilde{X}, d, \widetilde{x}, \widetilde{U}, \Delta)] \\
&= \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, \mathcal{K}_2^{\mathsf{R}}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (\widehat{f}, \widetilde{X}, d, \widetilde{x}, \widetilde{U}, \Delta)\right] \\
&\quad \cdot \frac{1}{2^{|\mathcal{W}_{\mathsf{in}}(f)|\lambda + 2\lambda|f| + (\lambda-1) + 2\lambda|f|}}.
\end{aligned}$$

According to the condition $\neg\mathsf{bad}_1 \wedge \neg\mathsf{bad}_3$ and the $2\,|f|$ extra queries, there are exact $q + 2\,|f| = q_\Sigma$ non-repeating queries. Moreover, $2\,|f|$ responses in $\cup_{\ell=1}^{3}\mathcal{K}_\ell^{\mathsf{R}}$ for the non-repeating extra queries are fixed by the conditioned values but the other responses are fixed by $\widetilde{\pi}^{\pm 1}$ for other $q_\Sigma - 2\,|f| = q$ non-repeating queries (which are responded with the rejection sampling, see Figure 4).

Let $\mathcal{N} \subseteq [1, q_\Sigma]$ denote the index set of these q queries in q_Σ non-repeating queries to $\widetilde{\pi}^{\pm 1}(\cdot)$ such that $|\mathcal{N}| = q$. The rejection sampling in $\widetilde{\pi}^{\pm 1}(\cdot)$ gives

$$\begin{aligned}
&\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, \mathcal{K}_2^{\mathsf{R}}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (\widehat{f}, \widetilde{X}, d, \widetilde{x}, \widetilde{U}, \Delta)\right] \\
&\quad = \prod_{i\in\mathcal{N}} \frac{1}{2^\lambda - |\mathcal{Q}_{i-1}|} = \prod_{i\in\mathcal{N}} \frac{1}{2^\lambda - (i-1)} \\
&\quad \le \frac{1}{2^\lambda - 2\,|f|} \times \frac{1}{2^\lambda - (2\,|f| + 1)} \times \cdots \times \frac{1}{2^\lambda - (q_\Sigma - 1)} = \frac{1}{(2^\lambda - 2\,|f|)_q}, \\
\Rightarrow &\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\omega \in \mathsf{comp}_{\mathsf{ideal}}(\tau)] \le \frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}} \cdot \frac{1}{(2^\lambda - 2\,|f|)_q \cdot (2^\lambda)^{4|f|}} \cdot \frac{1}{2^{\lambda-1}}.
\end{aligned}$$

So, we can have $\varepsilon_2 = 0$ since, for every $|f| \geq 0$ and every $q \geq 0$,

$$\begin{aligned}\frac{\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)]}{\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\omega \in \mathsf{comp}_{\mathsf{ideal}}(\tau)]} &\geq \frac{(2^\lambda - 2\,|f|)_q \cdot (2^\lambda)^{4|f|}}{(2^\lambda)_{q+4|f|}} \\ &\geq \frac{(2^\lambda - 2\,|f|)_q \cdot (2^\lambda)_{2|f|} \cdot (2^\lambda)^{2|f|}}{(2^\lambda)_{q+4|f|}} \\ &= \frac{(2^\lambda)_{q+2|f|} \cdot (2^\lambda)^{2|f|}}{(2^\lambda)_{q+4|f|}} = \frac{(2^\lambda)^{2|f|}}{(2^\lambda - (q+2\,|f|))_{2|f|}} \geq 1.\end{aligned}$$

Bounding ε_1. First, consider $\mathsf{bad}_1 \vee \mathsf{bad}_3$. For each $i \in [2, 2\,|f|]$, let coll_i denote the event that there exist distinct $(g,u),(g',u') \in$ "the first i pairs of $\mathcal{G}_{\mathsf{and}}(f)\times\{0,1\}$" such that $X_w \oplus k^g_u = X_{w'} \oplus k^{g'}_{u'}$, where $w := \mathsf{in}_u(g)$ and $w' := \mathsf{in}_{u'}(g')$, query_i denote the event that there exists $((\alpha,\beta),(g,u)) \in \cup_{\ell=1}^{2}\mathcal{K}_\ell \times$ "the first i pairs of $\mathcal{G}_{\mathsf{and}}(f) \times \{0,1\}$" such that $\alpha = X_w \oplus k^g_u$, where $w := \mathsf{in}_u(g)$, and bFwd_i denote $\mathsf{coll}_i \vee \mathsf{query}_i$. Then, $\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\mathsf{bFwd}_{2|f|}] = \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[(2) \vee (13) \vee (14)]$. We will prove the following bound using an induction:

$$\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\mathsf{bFwd}_i] \leq \frac{i(i-1) + 2i\cdot(q_1+q_2)}{2^{\lambda+1}}.$$

In the base case ($i = 2$), X_w and $X_{w'}$ are two active circuit input labels so that coll_2 occurs with probability at most $1/2^\lambda$ due to the sampling in the step 1 of SimIn (note that the probability is zero if $w = w'$). For query_2, each of the two labels matches a fixed $\alpha \oplus k^g_u$ also with probability $1/2^\lambda$. Following from a union bound, the target bound holds for $i = 2$.

Assume that the probability bound holds for $i \in [2, 2\,|f| - 1]$ and consider the $i+1$ case. Using the law of total probability, we have

$$\begin{aligned}\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\mathsf{bFwd}_{i+1}] &= \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\mathsf{bFwd}_{i+1} \;\middle|\; \mathsf{bFwd}_i\right] \cdot \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\mathsf{bFwd}_i] \\ &\quad + \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\mathsf{bFwd}_{i+1} \;\middle|\; \neg\mathsf{bFwd}_i\right] \cdot \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\neg\mathsf{bFwd}_i] \\ &\leq \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\mathsf{bFwd}_i] + \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\mathsf{bFwd}_{i+1} \;\middle|\; \neg\mathsf{bFwd}_i\right] \\ &\leq \frac{i(i-1) + 2i\cdot(q_1+q_2)}{2^{\lambda+1}} + \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\mathsf{bFwd}_{i+1} \;\middle|\; \neg\mathsf{bFwd}_i\right].\end{aligned}$$

Let (g^*,u^*) denote the $(i+1)$-th pair of $\mathcal{G}_{\mathsf{and}}(f) \times \{0,1\}$ and $w^* := \mathsf{in}_{u^*}(g^*)$. To incur coll_{i+1} conditioned on $\neg\mathsf{bFwd}_i = \neg\mathsf{coll}_i \wedge \neg\mathsf{query}_i$, we have $X_{w^*} \oplus k^{g^*}_{u^*} = X_w \oplus k^g_u$, where $(g,u) \in$ "the first i pairs of $\mathcal{G}_{\mathsf{and}}(f) \times \{0,1\}$" and $w := \mathsf{in}_u(g)$. Recall that each active label X_i is the XOR from (i) some active circuit input labels, and/or (ii) some active output labels of the *precedent* AND gates, i.e.,

$$X_i = \left(\bigoplus\nolimits_{w\in\mathcal{I}_i\subseteq\mathcal{W}_{\mathsf{in}}(f)} X_w\right) \oplus \left(\bigoplus\nolimits_{w\in\mathcal{J}_i\subseteq\mathcal{W}_{\mathsf{and}}(f)} X_w\right) = \bigoplus\nolimits_{w\in\mathcal{I}_i\ominus\mathcal{J}_i} X_w \in \{0,1\}^\lambda. \tag{23}$$

for $\mathcal{I}_i \ominus \mathcal{J}_i \neq \varnothing$. We use (23) to rewrite the equality to incur coll_{i+1} as follows:

$$\bigoplus\nolimits_{i\in(\mathcal{I}_w\ominus\mathcal{I}_{w^*})\ominus(\mathcal{J}_w\ominus\mathcal{J}_{w^*})} X_i = k^g_u \oplus k^{g^*}_{u^*} \in \{0,1\}^\lambda.$$

Here, the active circuit input labels in $\mathcal{I}_w \ominus \mathcal{I}_{w^*}$ are sampled at random in Sim. For the active labels in $\mathcal{J}_w \ominus \mathcal{J}_{w^*}$, they are masked by the responses sent from $\widetilde{\pi}^{\pm1}(\cdot)$ to Sim. Conditioned on $\neg\mathsf{bFwd}_i$, these responses are taken from uniform $c_1,\ldots,c_{n(\lambda)}$ in $\widetilde{\pi}^{\pm1}(\cdot)$ and pairwise independent since the queries (i.e., $X_w \oplus k^g_u$ for $(g,u) \in$ "the first i pairs of $\mathcal{G}_{\mathsf{and}}(f) \times \{0,1\}$" and $w := \mathsf{in}_u(g)$) are pairwise distinct under this condition. The active labels in $\mathcal{J}_w \ominus \mathcal{J}_{w^*}$ are uniform, and the XOR on the left hand is uniform unless $\mathcal{I}_w = \mathcal{I}_{w^*}$ and $\mathcal{J}_w = \mathcal{J}_{w^*}$. That is, the equality to incur coll_{i+1} holds with probability at most

$1/2^\lambda$ for some fixed (g, u). The same argument and probability hold for the equality $\alpha = X_{w^*} \oplus k_{u^*}^{g^*}$ to incur query_{i+1} for some fixed (α, β) under $\neg\mathsf{bFwd}_i$. Using a union bound,

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\mathsf{bFwd}_{i+1} \mid \neg\mathsf{bFwd}_i\right] \leq \frac{i + (q_1 + q_2)}{2^\lambda},$$

which concludes the induction for the $i+1$ case. We have

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\mathsf{bFwd}_{2|f|}\right] \leq \frac{2\,|f|\,(2\,|f|-1) + 4\,|f| \cdot (q_1 + q_2)}{2^{\lambda+1}}. \tag{24}$$

Consider (3), (4), (5), (6), (7), (8), (15), and (16) conditioned on $\mathsf{bFwd}_{2|f|}$. In each of them, $U_u^g \oplus \sigma(X_w \oplus k_u^g)$ is the response for query $X_w \oplus k_u^g$ to $\widetilde{\pi}^{\pm 1}(\cdot)$. Conditioned on $\neg\mathsf{bFwd}_i$, these responses are taken from uniform $c_1, \ldots, c_{n(\lambda)}$ in $\widetilde{\pi}^{\pm 1}(\cdot)$ and pairwise independent as the queries are pairwise distinct under this condition. Therefore, each of them occurs with probability $1/2^\lambda$ for some fixed quantifier. Let $\mathsf{bBwd} := (3) \vee (4) \vee (5) \vee (6) \vee (7) \vee (8) \vee (15) \vee (16)$. Taking a union bound over all quantifiers, we have

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\mathsf{bBwd} \mid \neg\mathsf{bFwd}_{2|f|}\right] \leq \frac{2\,|f|\,(2\,|f|-1) + 2\,|f| \cdot (q_1 + q_2)}{2^\lambda}. \tag{25}$$

Using (24) and (25), we have

$$\begin{aligned}
\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\mathsf{bad}_1 \vee \mathsf{bad}_3\right] &= \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\mathsf{bFwd}_{2|f|} \vee \mathsf{bBwd}\right] \\
&= \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\mathsf{bFwd}_{2|f|}\right] + \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\mathsf{bBwd} \mid \neg\mathsf{bFwd}_{2|f|}\right] \\
&\leq \frac{3\,|f|\,(2\,|f|-1) + 4\,|f| \cdot (q_1 + q_2)}{2^\lambda}.
\end{aligned} \tag{26}$$

Then, consider bad_2. From (9) (resp., (10)), we see $\Delta = \alpha \oplus X_a \oplus k_0^g$ (resp., $\Delta = \alpha \oplus X_b \oplus k_1^g$), occurring with probability $2^{-(\lambda-1)}$ due to the randomness of Δ. In (11), if $s_b \oplus x_b = 0$, linear orthomorphism σ ensures that

$$\Delta = \sigma^{-1}(\beta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g)),$$

which occurs with probability $2^{-(\lambda-1)}$; if $s_b \oplus x_b = 1$, according to permutation $\sigma'(x) := \sigma(x) \oplus x$ well-defined from σ, it holds with the same probability that

$$\Delta = \sigma'^{-1}(\beta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g)).$$

Similar result holds for (12). Taking a union bound over all pairs, we have

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\mathsf{bad}_2\right] \leq \frac{4\,|f| \cdot (q_1 + q_2 + q_3)}{2^{\lambda-1}}. \tag{27}$$

We have a bound ε_1 from (26) and (27):

$$\begin{aligned}
\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[Y(\omega) \in \mathcal{T}_{\mathsf{bad}}\right] &= \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\mathsf{bad}_1 \vee \mathsf{bad}_2 \vee \mathsf{bad}_3\right] \\
&\leq \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\mathsf{bad}_1 \vee \mathsf{bad}_3\right] + \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\mathsf{bad}_2\right] \\
&\leq \frac{3\,|f|\,(2\,|f|-1) + 12\,|f| \cdot (q + 2\,|f|)}{2^\lambda} \\
&= \frac{12qs + 30s^2 - 3s}{2^\lambda} = \varepsilon_1.
\end{aligned}$$

This lemma follows from the H-coefficient technique with the above $\varepsilon_1, \varepsilon_2$. □

Acknowledgements

Xiao Wang would like to thank Ran Canetti and Vinod Vaikuntanathan for initial discussion during his post-doc. The authors would like to thank Ben Riva and Yehuda Lindell for explanations of some related work. Work of Kang Yang is supported by the National Key Research and Development Program of China (Grant No. 2022YFB2702000), and by the National Natural Science Foundation of China (Grant Nos. 62102037, 61932019). Work of Xiao Wang is supported by DARPA under Contract No. HR001120C0087, NSF award #2016240, #2318974 and research awards from Meta and Google. The views, opinions, and/or findings expressed are those of the author(s) and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. Work of Yu Yu is supported by the National Natural Science Foundation of China (Grant Nos. 62125204 and 92270201), the National Key Research and Development Program of China (Grant No. 2018YFA0704701), and the Major Program of Guangdong Basic and Applied Research (Grant No. 2019B030302008). Yu Yu also acknowledges the support from the XPLORER PRIZE. Work of Zheli Liu is supported by the National Natural Science Foundation of China (Grant No. 62032012).

References

[AIKW13] Benny Applebaum, Yuval Ishai, Eyal Kushilevitz, and Brent Waters. Encoding functions with constant online rate or how to compress garbled circuits keys. In Ran Canetti and Juan A. Garay, editors, *CRYPTO 2013, Part II*, volume 8043 of *LNCS*, pages 166–184. Springer, Heidelberg, August 2013.

[BHKR13] Mihir Bellare, Viet Tung Hoang, Sriram Keelveedhi, and Phillip Rogaway. Efficient garbling from a fixed-key blockcipher. In *2013 IEEE Symposium on Security and Privacy*, pages 478–492. IEEE Computer Society Press, May 2013.

[BHR12a] Mihir Bellare, Viet Tung Hoang, and Phillip Rogaway. Adaptively secure garbling with applications to one-time programs and secure outsourcing. In Xiaoyun Wang and Kazue Sako, editors, *ASIACRYPT 2012*, volume 7658 of *LNCS*, pages 134–153. Springer, Heidelberg, December 2012.

[BHR12b] Mihir Bellare, Viet Tung Hoang, and Phillip Rogaway. Foundations of garbled circuits. In Ting Yu, George Danezis, and Virgil D. Gligor, editors, *ACM CCS 2012*, pages 784–796. ACM Press, October 2012.

[BMR90] Donald Beaver, Silvio Micali, and Phillip Rogaway. The round complexity of secure protocols (extended abstract). In *22nd ACM STOC*, pages 503–513. ACM Press, May 1990.

[BR93] Mihir Bellare and Phillip Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In Dorothy E. Denning, Raymond Pyle, Ravi Ganesan, Ravi S. Sandhu, and Victoria Ashby, editors, *ACM CCS 93*, pages 62–73. ACM Press, November 1993.

[BSW03] Boaz Barak, Ronen Shaltiel, and Avi Wigderson. Computational analogues of entropy. In Sanjeev Arora, Klaus Jansen, José D. P. Rolim, and Amit Sahai, editors, *APPROX-RANDOM 2003*, volume 2764 of *LNCS*, pages 200–215. Springer, 2003.

[Can00] Ran Canetti. Security and composition of multiparty cryptographic protocols. *Journal of Cryptology*, 13(1):143–202, January 2000.

[CS14] Shan Chen and John P. Steinberger. Tight security bounds for key-alternating ciphers. In Phong Q. Nguyen and Elisabeth Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 327–350. Springer, Heidelberg, May 2014.

[DG17] Nico Döttling and Sanjam Garg. Identity-based encryption from the Diffie-Hellman assumption. In Jonathan Katz and Hovav Shacham, editors, *CRYPTO 2017, Part I*, volume 10401 of *LNCS*, pages 537–569. Springer, Heidelberg, August 2017.

[DLMS14] Yuanxi Dai, Jooyoung Lee, Bart Mennink, and John P. Steinberger. The security of multiple encryption in the ideal cipher model. In Juan A. Garay and Rosario Gennaro, editors, *CRYPTO 2014, Part I*, volume 8616 of *LNCS*, pages 20–38. Springer, Heidelberg, August 2014.

[DSZ15] Daniel Demmler, Thomas Schneider, and Michael Zohner. ABY - A framework for efficient mixed-protocol secure two-party computation. In *NDSS 2015*. The Internet Society, February 2015.

[GKW+20] Chun Guo, Jonathan Katz, Xiao Wang, Chenkai Weng, and Yu Yu. Better concrete security for half-gates garbling (in the multi-instance setting). In Daniele Micciancio and Thomas Ristenpart, editors, *CRYPTO 2020, Part II*, volume 12171 of *LNCS*, pages 793–822. Springer, Heidelberg, August 2020.

[GKWY20] Chun Guo, Jonathan Katz, Xiao Wang, and Yu Yu. Efficient and secure multiparty computation from fixed-key block ciphers. In *2020 IEEE Symposium on Security and Privacy*, pages 825–841. IEEE Computer Society Press, May 2020.

[GLNP15] Shay Gueron, Yehuda Lindell, Ariel Nof, and Benny Pinkas. Fast garbling of circuits under standard assumptions. In Indrajit Ray, Ninghui Li, and Christopher Kruegel, editors, *ACM CCS 2015*, pages 567–578. ACM Press, October 2015.

[GS18] Sanjam Garg and Akshayaram Srinivasan. Adaptively secure garbling with near optimal online complexity. In Jesper Buus Nielsen and Vincent Rijmen, editors, *EUROCRYPT 2018, Part II*, volume 10821 of *LNCS*, pages 535–565. Springer, Heidelberg, April / May 2018.

[HJO+16] Brett Hemenway, Zahra Jafargholi, Rafail Ostrovsky, Alessandra Scafuro, and Daniel Wichs. Adaptively secure garbled circuits from one-way functions. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO 2016, Part III*, volume 9816 of *LNCS*, pages 149–178. Springer, Heidelberg, August 2016.

[HLR07] Chun-Yuan Hsiao, Chi-Jen Lu, and Leonid Reyzin. Conditional computational entropy, or toward separating pseudoentropy from compressibility. In Moni Naor, editor, *EUROCRYPT 2007*, volume 4515 of *LNCS*, pages 169–186. Springer, Heidelberg, May 2007.

[HT16] Viet Tung Hoang and Stefano Tessaro. Key-alternating ciphers and key-length extension: Exact bounds and multi-user security. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO 2016, Part I*, volume 9814 of *LNCS*, pages 3–32. Springer, Heidelberg, August 2016.

[HW15] Pavel Hubacek and Daniel Wichs. On the communication complexity of secure function evaluation with long output. In Tim Roughgarden, editor, *ITCS 2015*, pages 163–172. ACM, January 2015.

[JKK+17] Zahra Jafargholi, Chethan Kamath, Karen Klein, Ilan Komargodski, Krzysztof Pietrzak, and Daniel Wichs. Be adaptive, avoid overcommitting. In Jonathan Katz and Hovav Shacham, editors, *CRYPTO 2017, Part I*, volume 10401 of *LNCS*, pages 133–163. Springer, Heidelberg, August 2017.

[JKO13] Marek Jawurek, Florian Kerschbaum, and Claudio Orlandi. Zero-knowledge using garbled circuits: how to prove non-algebraic statements efficiently. In Ahmad-Reza Sadeghi, Virgil D. Gligor, and Moti Yung, editors, *ACM CCS 2013*, pages 955–966. ACM Press, November 2013.

[JO20] Zahra Jafargholi and Sabine Oechsner. Adaptive security of practical garbling schemes. In Karthikeyan Bhargavan, Elisabeth Oswald, and Manoj Prabhakaran, editors, *INDOCRYPT 2020*, volume 12578 of *LNCS*, pages 741–762. Springer, Heidelberg, December 2020.

[JSW17] Zahra Jafargholi, Alessandra Scafuro, and Daniel Wichs. Adaptively indistinguishable garbled circuits. In Yael Kalai and Leonid Reyzin, editors, *TCC 2017, Part II*, volume 10678 of *LNCS*, pages 40–71. Springer, Heidelberg, November 2017.

[JW16] Zahra Jafargholi and Daniel Wichs. Adaptive security of Yao's garbled circuits. In Martin Hirt and Adam D. Smith, editors, *TCC 2016-B, Part I*, volume 9985 of *LNCS*, pages 433–458. Springer, Heidelberg, October / November 2016.

[Kel20] Marcel Keller. MP-SPDZ: A versatile framework for multi-party computation. In Jay Ligatti, Xinming Ou, Jonathan Katz, and Giovanni Vigna, editors, *ACM CCS 2020*, pages 1575–1590. ACM Press, November 2020.

[KKP21] Chethan Kamath, Karen Klein, and Krzysztof Pietrzak. On treewidth, separators and yao's garbling. In Kobbi Nissim and Brent Waters, editors, *TCC 2021, Part II*, volume 13043 of *LNCS*, pages 486–517. Springer, Heidelberg, November 2021.

[KKPW21] Chethan Kamath, Karen Klein, Krzysztof Pietrzak, and Daniel Wichs. Limits on the adaptive security of yao's garbling. In Tal Malkin and Chris Peikert, editors, *CRYPTO 2021, Part II*, volume 12826 of *LNCS*, pages 486–515, Virtual Event, August 2021. Springer, Heidelberg.

[KMR14] Vladimir Kolesnikov, Payman Mohassel, and Mike Rosulek. FleXOR: Flexible garbling for XOR gates that beats free-XOR. In Juan A. Garay and Rosario Gennaro, editors, *CRYPTO 2014, Part II*, volume 8617 of *LNCS*, pages 440–457. Springer, Heidelberg, August 2014.

[KS08] Vladimir Kolesnikov and Thomas Schneider. Improved garbled circuit: Free XOR gates and applications. In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfsdóttir, and Igor Walukiewicz, editors, *ICALP 2008, Part II*, volume 5126 of *LNCS*, pages 486–498. Springer, Heidelberg, July 2008.

[LP09] Yehuda Lindell and Benny Pinkas. A proof of security of Yao's protocol for two-party computation. *Journal of Cryptology*, 22(2):161–188, April 2009.

[LR14] Yehuda Lindell and Ben Riva. Cut-and-choose Yao-based secure computation in the online/offline and batch settings. In Juan A. Garay and Rosario Gennaro, editors, *CRYPTO 2014, Part II*, volume 8617 of *LNCS*, pages 476–494. Springer, Heidelberg, August 2014.

[LR15] Yehuda Lindell and Ben Riva. Blazing fast 2PC in the offline/online setting with security for malicious adversaries. In Indrajit Ray, Ninghui Li, and Christopher Kruegel, editors, *ACM CCS 2015*, pages 579–590. ACM Press, October 2015.

[LWN+15] Chang Liu, Xiao Shaun Wang, Kartik Nayak, Yan Huang, and Elaine Shi. ObliVM: A programming framework for secure computation. In *2015 IEEE Symposium on Security and Privacy*, pages 359–376. IEEE Computer Society Press, May 2015.

[Nie02] Jesper Buus Nielsen. Separating random oracle proofs from complexity theoretic proofs: The non-committing encryption case. In Moti Yung, editor, *CRYPTO 2002*, volume 2442 of *LNCS*, pages 111–126. Springer, Heidelberg, August 2002.

[NPS99] Moni Naor, Benny Pinkas, and Reuban Sumner. Privacy preserving auctions and mechanism design. In Stuart I. Feldman and Michael P. Wellman, editors, *Proceedings of the First ACM Conference on Electronic Commerce (EC'99)*, pages 129–139. ACM, 1999.

[Pat09] Jacques Patarin. The "coefficients H" technique (invited talk). In Roberto Maria Avanzi, Liam Keliher, and Francesco Sica, editors, *SAC 2008*, volume 5381 of *LNCS*, pages 328–345. Springer, Heidelberg, August 2009.

[PSSW09] Benny Pinkas, Thomas Schneider, Nigel P. Smart, and Stephen C. Williams. Secure two-party computation is practical. In Mitsuru Matsui, editor, *ASIACRYPT 2009*, volume 5912 of *LNCS*, pages 250–267. Springer, Heidelberg, December 2009.

[RR16] Peter Rindal and Mike Rosulek. Faster malicious 2-party secure computation with online/offline dual execution. In Thorsten Holz and Stefan Savage, editors, *USENIX Security 2016*, pages 297–314. USENIX Association, August 2016.

[RR21] Mike Rosulek and Lawrence Roy. Three halves make a whole? Beating the half-gates lower bound for garbled circuits. In Tal Malkin and Chris Peikert, editors, *CRYPTO 2021, Part I*, volume 12825 of *LNCS*, pages 94–124, Virtual Event, August 2021. Springer, Heidelberg.

[RS08] Phillip Rogaway and John P. Steinberger. Constructing cryptographic hash functions from fixed-key blockciphers. In David Wagner, editor, *CRYPTO 2008*, volume 5157 of *LNCS*, pages 433–450. Springer, Heidelberg, August 2008.

[SHS+15] Ebrahim M. Songhori, Siam U. Hussain, Ahmad-Reza Sadeghi, Thomas Schneider, and Farinaz Koushanfar. TinyGarble: Highly compressed and scalable sequential garbled circuits. In *2015 IEEE Symposium on Security and Privacy*, pages 411–428. IEEE Computer Society Press, May 2015.

[WMK16] Xiao Wang, Alex J. Malozemoff, and Jonathan Katz. EMP-toolkit: Efficient MultiParty computation toolkit. https://github.com/emp-toolkit, 2016.

[Yao82] Andrew Chi-Chih Yao. Theory and applications of trapdoor functions (extended abstract). In *23rd FOCS*, pages 80–91. IEEE Computer Society Press, November 1982.

[Yao86] Andrew Chi-Chih Yao. How to generate and exchange secrets (extended abstract). In *27th FOCS*, pages 162–167. IEEE Computer Society Press, October 1986.

[ZE15] Samee Zahur and David Evans. Obliv-C: A language for extensible data-oblivious computation. Cryptology ePrint Archive, Report 2015/1153, 2015. https://eprint.iacr.org/2015/1153.

[ZRE15] Samee Zahur, Mike Rosulek, and David Evans. Two halves make a whole - reducing data transfer in garbled circuits using half gates. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part II*, volume 9057 of *LNCS*, pages 220–250. Springer, Heidelberg, April 2015.

A More Preliminaries

Define the following notation in the appendices. Let bold lowercase letters (e.g., $\boldsymbol{a}$) denote column vectors and bold uppercase letters (e.g., $\boldsymbol{A}$) denote matrices. Let $\boldsymbol{I}_n$ denote the n-by-n identity matrix. Let $\mathsf{Half}_0(\boldsymbol{a}) \in \mathbb{F}_{2^n}$ (resp., $\mathsf{Half}_1(\boldsymbol{a}) \in \mathbb{F}_{2^n}$) denote the lower (resp., upper) half of vector $\boldsymbol{a} \in \mathbb{F}_{2^n}^2$. We can also define $\mathsf{lsb}(\boldsymbol{a}) := \mathsf{lsb}(\mathsf{Half}_0(\boldsymbol{a}))$ for vector $\boldsymbol{a} \in \mathbb{F}_{2^n}^2$. Let $\otimes$ denote the Kronecker product of matrices. Let $\boldsymbol{A}^+$ denote the left inverse of matrix $\boldsymbol{A}$. We will use $\mathbb{F}_{2^n}$, $\mathbb{F}_2^n$, and $\{0,1\}^n$ interchangeably.

A generalized definition of linear orthomorphism is described as follows:

Definition 2 (Linear orthomorphism)**.** *A permutation $\sigma : \mathbb{G} \to \mathbb{G}$ over an additive Abelian group $\mathbb{G}$ is called a linear orthomorphism for a function family $\mathcal{L}$ of some linear functions from $\mathbb{G}$ to $\mathbb{G}$, if (i) $\sigma(x+y) = \sigma(x)+\sigma(y)$ for any $x, y \in \mathbb{G}$, (ii) $\sigma'(x) := \sigma(x) - L(x)$ is also a permutation for every $L \in \mathcal{L}$, and (iii) σ, σ' and their inverses are efficiently computable. We will simply call σ a linear orthomorphism if $\mathcal{L}$ contains only the identity function.*

For half-gates [ZRE15], we can use the two instantiations of σ in Section 3.1. For three-halves [RR21], we require a linear orthomorphism $\sigma : \mathbb{F}_{2^{\lambda/2+1}}^2 \to \mathbb{F}_{2^{\lambda/2+1}}^2$ for the function family

$$\mathcal{L} = \left\{ L_{\xi_1,\xi_2,\xi_3,\xi_4} : \mathbb{F}_{2^{\lambda/2+1}}^2 \to \mathbb{F}_{2^{\lambda/2+1}}^2 \right\}_{\xi_1,\xi_2,\xi_3,\xi_4 \in \mathbb{F}_2}, \quad L_{\xi_1,\xi_2,\xi_3,\xi_4}\left(\begin{bmatrix} x_L \\ x_R \end{bmatrix}\right) = \begin{bmatrix} \xi_1 x_L \oplus \xi_2 x_R \\ \xi_3 x_L \oplus \xi_4 x_R \end{bmatrix}.$$

The following linear orthomorphism σ is suggested by three-halves:

$$\sigma\left(\begin{bmatrix} x_L \\ x_R \end{bmatrix}\right) = \begin{bmatrix} c \cdot x_L \\ c \cdot x_R \end{bmatrix}$$

where $c \in \mathbb{F}_{2^{\lambda/2+1}} \setminus \mathbb{F}_{2^2}$ is a fixed element and the multiplication is in $\mathbb{F}_{2^{\lambda/2+1}}$.

Auxiliary circuit notation for the proofs in appendices. Given a circuit f with fan-in two and fan-out one, we define:

- $\mathcal{Y}_1(f), \ldots, \mathcal{Y}_m(f) \subseteq \mathcal{W}_{\mathsf{out}}(f)$: $m \geq 1$ *maximal* non-empty partitions of circuit output wires such that, for every $i \in [1,m]$ and every distinct $w, w' \in \mathcal{Y}_i(f)$, wire w and w' are from the same XOR combination of (i) some circuit input wires, and/or (ii) some output wires of *last-level* `AND` gates. We point out that $\sum_{i=1}^m |\mathcal{Y}_i(f)| = |\mathcal{W}_{\mathsf{out}}(f)| \geq m$.
- $\mathcal{Z}(f) \subseteq \mathcal{W}_{\mathsf{and}}(f)$: The set of the output wires of *last-level* `AND` gates.
- $\mathcal{X}(f) := \mathcal{W}_{\mathsf{in}}(f) \cup \mathcal{Z}(f)$.
- For each $i \in [1,m]$, let $\mathcal{X}_i(f) \subseteq \mathcal{X}(f)$ denote the set collecting the wires used to XOR-combine the circuit output wires in $\mathcal{Y}_i(f)$, and $y_i(f)$ denote the first (i.e., representative) wire in $\mathcal{Y}_i(f)$.

For three-halves, we also define:

- For an `AND` gate $g \in \mathcal{G}_{\mathsf{and}}(f)$, let $\mathsf{in}_2(g)$ denote the *global* wire assigned to the output wires of all `XOR` gates with input wires $a := \mathsf{in}_0(g)$ and $b := \mathsf{in}_1(g)$. If such an `XOR` gate does not exists, we still define this wire artificially.
- $\mathcal{W}_\cup(f) := \cup_{g \in \mathcal{G}_{\mathsf{and}}(f)} \{\mathsf{in}_0(g), \mathsf{in}_1(g), \mathsf{in}_2(g)\}$.
- Let $n_i(f) \geq 0$ denote the number of times the wire i is used for the input to the `AND` gates in f.
- $N := \sum_{i \in \mathcal{W}_\cup(f)} \lceil n_i(f)/2 \rceil$.
- $M := |\{i \in \mathcal{W}_\cup(f) \mid n_i(f) \text{ is odd}\}|$.

B Proof of Lemma 1

Lemma 1 ([DLMS14]). *Let the notations be defined in Section 3.4. Then, for every auxiliary input $z \in \{0,1\}^*$, every computationally unbounded adversary $\mathcal{A}$, and every attainable transcript $\tau \in \mathcal{T}_{\mathcal{A}(z)}$, it holds that*

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[X(\omega) = \tau] = \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)],$$
$$\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] = \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\omega \in \mathsf{comp}_{\mathsf{ideal}}(\tau)].$$

Proof. Let $\mathcal{M}^O \to x$ be the predicate that is true if and only if the interaction between $\mathcal{M}$ and oracle O produces a transcript with literal value x.

We have that, for every computationally unbounded non-uniform adversary $\mathcal{A} = \mathcal{A}(z)$ and every attainable transcript $\tau \in \mathcal{T}_{\mathcal{A}}$,

$$\begin{aligned}
\Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[X(\omega) = \tau] &= \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}\left[\omega \in \left\{\omega \in \Omega_{\mathsf{real}} \;\middle|\; \mathcal{A}^\omega \to \tau\right\}\right], \\
\Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)] &= \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}\left[\omega \in \left\{\omega \in \Omega_{\mathsf{real}} \;\middle|\; \exists \mathcal{A}' : \mathcal{A}'^{\omega} \to \tau\right\}\right] \\
&= \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}\left[\omega \in \left\{\omega \in \Omega_{\mathsf{real}} \;\middle|\; (\exists \mathcal{A}' : \mathcal{A}'^{\omega} \to \tau) \wedge \tau \in \mathcal{T}_{\mathcal{A}}\right\}\right] \\
&= \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}\left[\omega \in \left\{\omega \in \Omega_{\mathsf{real}} \;\middle|\; (\exists \mathcal{A}' : \mathcal{A}'^{\omega} \to \tau) \wedge (\exists \omega' : \mathcal{A}^{\omega'} \to \tau)\right\}\right],
\end{aligned}$$

where $\mathcal{A}$, $\mathcal{A}'$, ω, and ω' are deterministic (c.f. Section 3.4).

We claim that, for every computationally unbounded non-uniform $\mathcal{A}$, every attainable $\tau \in \mathcal{T}_{\mathcal{A}}$, and every $\omega \in \Omega_{\mathsf{real}}$, the predicate equivalence $\mathcal{A}^\omega \to \tau \Leftrightarrow (\exists \mathcal{A}' : \mathcal{A}'^{\omega} \to \tau) \wedge (\exists \omega' : \mathcal{A}^{\omega'} \to \tau)$ holds. The forward implication is obvious by using $\mathcal{A}' := \mathcal{A}$ and $\omega' := \omega$. For the backward implication, we can use in an induction that τ is an *ordered* list of query-response pairs $\{(q_1, r_1), \dots\}$.

In the base case, $\mathcal{A}$ (resp., $\mathcal{A}'$) will send query q_1 to and receive response r_1 from ω' (resp., ω) according to the same fixed τ. As q_1 (resp., r_1) is the first query (resp., response) of $\mathcal{A}$ (resp., ω) so that there is no previous list on which it can depend, $\mathcal{A}$ will also send q_1 and receive r_1 if it is given oracle ω. We assume that the interaction between $\mathcal{A}$ and ω has produced the same list of query-response pairs $\tau_{i-1} = \{(q_1, r_1), \dots, (q_{i-1}, r_{i-1})\}$, which is a prefix of τ. It is known from the same fixed τ that, conditioned on the same previous list τ_{i-1}, $\mathcal{A}$ (resp., $\mathcal{A}'$) will send query q_i to and receive response r_i from ω' (resp., ω). As a result, if $\mathcal{A}$ accesses ω for the i-th query conditioned on τ_{i-1}, $\mathcal{A}$ will also send q_i as query and receive r_i, which only depends on τ_{i-1} in oracle ω. The above induction concludes that the two predicates are equivalent for every non-uniform $\mathcal{A}$, attainable $\tau \in \mathcal{T}_{\mathcal{A}}$, and $\omega \in \Omega_{\mathsf{real}}$. So, we have

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[X(\omega) = \tau] = \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)].$$

We can also prove the ideal-world result using the same induction. □

C Adaptive Security of Half-Gates in pRPM

We prove that the original implementation of the half-gates scheme [ZRE15], which is identical to Figure 1 but sends decoding table d as part of garbled circuit $\widehat{f}$, is adaptively secure in the pRPM and does not suffer from the lower bound in the npRPM. We refer to Appendix A for the additional notation.

Theorem 5. *Let $\mathsf{H}(X, k) = \pi(X \oplus k) \oplus \sigma(X \oplus k)$ be a tweakable hash function where $X, k \in \{0,1\}^\lambda$, $\pi \in \mathcal{P}_\lambda$ is random permutation, and $\sigma : \{0,1\}^\lambda \to \{0,1\}^\lambda$ is a linear orthomorphism. Then, half-gates ([ZRE15]) is a λ-garbling scheme with (q, s, ε)-adaptive security in the pRPM, where $\varepsilon = (8qs + 21s^2)/2^{\lambda-1}$.*

$\underline{\mathsf{SimF}(f)}$:

1: $F \leftarrow (\{0,1\}^{2\lambda})^{|f|}$, $d \leftarrow \mathbb{F}_2^{|\mathcal{W}_{\mathsf{out}}(f)|}$ such that

(i) For each $i \in [1,m]$ and each wire $j \in \mathcal{Y}_i(f)$, $d_j = d_{y_i(f)}$.

(ii) For each $(\mu_1,\ldots,\mu_m) \in \mathbb{F}_2^m$ such that $\ominus_{i\in[1,m],\mu_i=1}\mathcal{X}_i(f) = \varnothing$, $\oplus_{i\in[1,m]}\mu_i \cdot d_{y_i(f)} = 0$.

2: **return** $\widehat{f} := (f' := f, F, d)$, $\mathsf{st}_{\mathsf{sim}} := (\mathsf{st}_{\mathsf{sim}}, f, F, d)$

$\underline{\mathsf{SimIn}(f(x))}$:

1: $\{X_i\}_{i\in\mathcal{X}(f)} \leftarrow (\{0,1\}^{\lambda})^{|\mathcal{X}(f)|}$ such that, for each $j \in [1,m]$, $\mathsf{lsb}(\oplus_{i\in\mathcal{X}_j(f)}X_i) = d_{y_j(f)} \oplus f(x)_{y_j(f)}$.

2: **for** $g \in \mathcal{G}(f)$ in topology order **do**

3: $(a,b,c) := (\mathsf{in}_0(g), \mathsf{in}_1(g), \mathsf{out}(g))$

4: **if** $\mathsf{type}(g) = \mathtt{XOR}$ **then** $X_c := X_a \oplus X_b$

5: **else if** $\mathsf{type}(g) = \mathtt{AND}$ **then**

6: $k_0^g := 2\cdot g - 1$, $k_1^g := 2 \cdot g$

7: $s_a := \mathsf{lsb}(X_a)$, $s_b := \mathsf{lsb}(X_b)$

8: **if** $c \in \mathcal{Z}(f)$ **then**

9: $U_1^g \leftarrow \{0,1\}^{\lambda}$

10: $U_0^g := X_c \oplus U_1^g \oplus s_a G_0^g \oplus s_b(G_1^g \oplus X_a)$

11: **else if** $c \notin \mathcal{Z}(f)$ **then**

12: $U_0^g, U_1^g \leftarrow \{0,1\}^{\lambda}$

13: $X_c := U_0^g \oplus U_1^g \oplus s_a G_0^g \oplus s_b(G_1^g \oplus X_a)$

14: (Programming) Add two pairs $(X_a \oplus k_0^g, U_0^g \oplus \sigma(X_a \oplus k_0^g))$ and $(X_b \oplus k_1^g, U_1^g \oplus \sigma(X_b \oplus k_1^g))$ to the list $\mathcal{Q}$ kept in $\mathsf{st}_{\mathsf{sim}}$, if they cause no pre-image collision or image collision in $\mathcal{Q}$, to ensure

$$\underbrace{\pi(X_a \oplus k_0^g) \oplus \sigma(X_a \oplus k_0^g)}_{\mathsf{H}(X_a, k_0^g)} = U_0^g, \quad \underbrace{\pi(X_b \oplus k_1^g) \oplus \sigma(X_b \oplus k_1^g)}_{\mathsf{H}(X_b, k_1^g)} = U_1^g.$$

15: **return** $\widehat{x} := \{X_i\}_{i\in\mathcal{W}_{\mathsf{in}}(f)}$, $\mathsf{st}_{\mathsf{sim}} := (\mathsf{st}_{\mathsf{sim}}, \widetilde{X}, \widetilde{U})$ where $\mathsf{st}_{\mathsf{sim}}$ on the right hand has an updated list $\mathcal{Q}$, $\widetilde{X} := \{X_i\}_{i\in\mathcal{W}(f)}$, $\widetilde{U} := \{U_0^g, U_1^g\}_{g\in\mathcal{G}_{\mathsf{and}}(f)}$.

Figure 5: Our simulator for half-gates in the pRPM.

Proof. The correctness has been proved in the original work [ZRE15]. We only need to consider the simulation.

Our simulator Sim consists of (SimF, SimIn) in Figure 5 and $\mathsf{SimP}^{\pm 1}$, which emulates the random permutation and its inverse on-the-fly. More specifically, there is a list $\mathcal{Q}$ of query-response pairs in internal state $\mathsf{st}_{\mathsf{sim}}$. Upon receiving forward query α (resp., backward query β) from $\mathcal{A}_i$ to SimP (resp., SimP^{-1}), it reads $\mathcal{Q}$ from $\mathsf{st}_{\mathsf{sim}}$ and checks whether $\exists(\alpha,\gamma) \in \mathcal{Q}$ (resp., $\exists(\gamma,\beta) \in \mathcal{Q}$). If true, it returns γ as response; otherwise it samples $\gamma \leftarrow \{s \in \{0,1\}^{\lambda} \mid (\ldots, s) \notin \mathcal{Q}\}$ (resp., $\gamma \leftarrow \{s \in \{0,1\}^{\lambda} \mid (s,\ldots) \notin \mathcal{Q}\}$), adds (α,γ) (resp., (γ,β)) to $\mathcal{Q}$, and returns γ as response. The programming is the step 14 of SimIn, where (α,β) is added to $\mathcal{Q}$ if and only if there is no pre-image collision with $(\alpha,\ldots) \in \mathcal{Q}$ or image collision with $(\ldots,\beta) \in \mathcal{Q}$.

To see that Sim is PPT, we note that the circuit-dependent notation can be efficiently computed by traversing the polynomial-size circuit f. Then, the crux is to show that the step 1 in both SimF and SimIn can be polynomial-time.

The runtime of the step 1 of SimF is dominated by the runtime of iterating through all qualified $(\mu_1,\ldots,\mu_m) \in \mathbb{F}_2^m$. To find the qualified vectors, one can interpret each $\mathcal{X}_i(f)$ as a one-hot *non-zero* column vector in the space $\mathbb{F}_2^{|\mathcal{W}(f)|}$ and derive a $|\mathcal{W}(f)|$-by-m matrix $\boldsymbol{E}$ from these column vectors. One can check that all qualified vectors fall in the kernel of $\boldsymbol{E}$. This kernel can be efficiently computed from the Gaussian elimination on $\boldsymbol{E}$ and is a subspace spanned by $m - \mathsf{rank}(\boldsymbol{E})$ basis vectors. So, the step 1 of SimF only needs to iterate through these basis vectors, and the other qualified vectors must satisfy the constraints as they are in the subspace. As a result, the step 1 of SimF runs in polynomial

time due to the Gaussian elimination plus a linear-time pass to assign random or constrained values to $d_{y_i(f)}$'s according to the $m - \mathsf{rank}(\boldsymbol{E})$ basis vectors.

The step 1 of SimIn only requires one linear-time pass to assign constrained or random values to the active labels so it runs in polynomial time. The linear constraint on the LSBs of these active labels has rank $\mathsf{rank}(\boldsymbol{E})$ and is satisfiable for the $d_{y_i(f)}$'s assigned in SimF.

Then, we fix z and $\mathcal{A}$. We will use the H-coefficient technique (Section 3.4) with transcript padding to bound the advantage of distinguishing between the real world (i.e., the adaptive experiment that uses the half-gates scheme) and the ideal world (i.e., the adaptive experiment that uses simulator Sim). In this technique, we consider the computationally unbounded non-uniform adversary $\mathcal{A} = \mathcal{A}(z)$ and compute $\varepsilon_1, \varepsilon_2$ as follows.

Transcript padding. In either world, $\mathcal{A}$ will interact with an integrated oracle that acts as the two-round challenger in the adaptive experiment and provides interfaces $\pi^{*\pm 1} \in \{\pi^{\pm 1}, \mathsf{SimP}^{\pm 1}\}$ for forward/backward permutation queries. $\mathcal{A}$ can learn $\pi^*(\alpha) = \beta$ if and only if it sent forward query α to π^* and received response β, or sent backward query β to π^{*-1} and received response α.

To compute $\varepsilon_1, \varepsilon_2$ more easily, we ask the oracle to send more messages to $\mathcal{A}$ and $\mathcal{A}$ to make extra queries (in addition to the supposed q queries) in both two worlds. More specifically,

- Upon receiving x from $\mathcal{A}$, the oracle sends $\widetilde{X} := \{X_i\}_{i \in \mathcal{W}(f)}$ instead of $\widehat{x} := \{X_i\}_{i \in \mathcal{W}_{\mathsf{in}}(f)}$ to $\mathcal{A}$. In addition to the active input labels in $\widehat{x}$, the former also gives the active internal and output labels. In the real world, the oracle can run $\mathsf{HG.DecEval}^{\pi^{\pm 1}(\cdot)}$, which determines other active labels in $\widetilde{X}$. In the ideal world, this $\widetilde{X}$ can be directly output by SimIn.

- Along with $\widetilde{X}$, the oracle sends $\widetilde{x} := \{x_i\}_{i \in \mathcal{W}(f)}$ to $\mathcal{A}$, which denote the wire truth values in the evaluation of $f(x)$. Both two oracles "echo" these values, which are self-evident to $\mathcal{A}$, to explicitly include them in transcripts. In the experiment, the real-world oracle uses $x = \{x_i\}_{i \in \mathcal{W}_{\mathsf{in}}(f)}$ in $\mathsf{HG.Encode}^{\pi^{\pm 1}(\cdot)}$, but the ideal-world oracle can only use $f(x) = \{x_i\}_{i \in \mathcal{W}_{\mathsf{out}}(f)}$ in SimIn.

- Along with $\widetilde{X}$, the oracle sends $\widetilde{U} := \{U_0^g, U_1^g\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}$ to $\mathcal{A}$. The real-world oracle computes $U_0^g := \mathsf{H}(X_a, k_0^g)$ and $U_1^g := \mathsf{H}(X_b, k_1^g)$ for each $g \in \mathcal{G}_{\mathsf{and}}(f)$ with $(a, b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$. In the ideal world, this $\widetilde{U}$ is output by SimIn and contains the "hash outputs" fixed by the random tape, which is specified by the oracle to run the programming therein.

- **(Extra queries)** Upon receiving $(\widetilde{X}, \widetilde{x}, \widetilde{U})$ from the oracle, $\mathcal{A}$ will also make a forward permutation query $X_a \oplus k_0^g$ (resp., $X_b \oplus k_1^g$) for each $g \in \mathcal{G}_{\mathsf{and}}(f)$ with $(a, b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$, if it has never learned $\pi^*(X_a \oplus k_0^g) = Y$ (resp., $\pi^*(X_b \oplus k_1^g) = Y$) for some Y in its interaction with $\pi^{*\pm 1} \in \{\pi^{\pm 1}, \mathsf{SimP}^{\pm 1}\}$.

- At the end of the experiment (i.e., once all other transcripts are settled), the oracle sends Δ to $\mathcal{A}$. In the real world, the oracle gets this Δ from the output of $\mathsf{HG.Garble}^{\pi^{\pm 1}(\cdot)}$. In the ideal world, Δ is dummy and sampled by the oracle at this time (note that Sim does not use Δ).

According to the two oracle constructions, real-world sample space

$$\Omega_{\mathsf{real}} = \{0,1\}^{\lambda - 1} \times \{0,1\}^{|\mathcal{W}_{\mathsf{in}}(f)|\lambda} \times \mathcal{P}_\lambda,$$

and ideal-world sample space

$$\begin{aligned}\Omega_{\mathsf{ideal}} = (\{0,1\}^{2\lambda})^{|f|} &\times \{0,1\}^{\mathsf{rank}(\boldsymbol{E})} \times \{0,1\}^{(|\mathcal{W}_{\mathsf{in}}(f)| + |\mathcal{Z}(f)|)\lambda - \mathsf{rank}(\boldsymbol{E})} \\ &\times (\{0,1\}^{\lambda})^{|\mathcal{Z}(f)|} \times (\{0,1\}^{2\lambda})^{|f| - |\mathcal{Z}(f)|} \\ &\times \underbrace{\{0,1\}^*}_{\substack{\text{random tape for} \\ \text{the sampling in } \mathsf{SimP}^{\pm 1}(\cdot)}} \times \underbrace{\{0,1\}^{\lambda - 1}}_{\text{dummy } \Delta}.\end{aligned}$$

Given the oracle constructions, a transcript in the original adaptive experiment will be padded with more literal values. Note that transcript padding will not lower the advantage of $\mathcal{A}$ since $\mathcal{A}$ can discard the padding values at will. With the padding, a transcript is of the form:

$$\tau = (\mathcal{K}_1, (f, \widehat{f}), \mathcal{K}_2, (x, (\widetilde{X}, \widetilde{x}, \widetilde{U})), \mathcal{K}_3, \Delta),$$

where $\mathcal{K}_1$, $\mathcal{K}_2$, and $\mathcal{K}_3$ are the ordered lists of query-response pairs seen in the interleaved interaction with permutation oracles. We do not explicitly consider query direction in these pairs. Given $\pi^{*\pm 1} \in \{\pi^{\pm 1}, \mathsf{SimP}^{\pm 1}\}$, $\mathcal{A}$ is able to learn $\pi^*(\alpha) = \beta$ if and only if there exists $(\alpha, \beta) \in \cup_{\ell=1}^{3} \mathcal{K}_\ell$.

Let $q_\ell := |\mathcal{K}_\ell|$ for every $\ell \in \{1, 2, 3\}$ and $q_\Sigma := \sum_{\ell=1}^{3} q_\ell$. It follows from the extra queries that $q_\Sigma \leq q + 2\,|f|$. Without loss of generality, we assume that $\mathcal{A}$ only makes *non-repeating* queries, i.e., it never makes forward query α to π^* or backward query β to π^{*-1} for any learned permutation entry (α, β).

Bad transcripts. A transcript $\tau \in \mathcal{T}_{\mathsf{bad}}$ if and only if it incurs at least one of the following events:

- bad_1. There exist distinct $(g, u), (g', u') \in \mathcal{G}_{\mathsf{and}}(f) \times \{0, 1\}$ such that

$$X_w \oplus k_u^g = X_{w'} \oplus k_{u'}^{g'} \quad \vee \tag{2}$$
$$U_u^g \oplus \sigma(X_w \oplus k_u^g) = U_{u'}^{g'} \oplus \sigma(X_{w'} \oplus k_{u'}^{g'}) \tag{3}$$

  where $w := \mathsf{in}_u(g)$ and $w' := \mathsf{in}_{u'}(g')$, or

  There exists $g \in \mathcal{G}_{\mathsf{and}}(f)$ such that

$$(s_b \oplus x_b)\Delta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g) = x_a\Delta \oplus X_a \oplus G_1^g \oplus U_1^g \oplus \sigma(X_b \oplus k_1^g) \tag{4}$$

  where $(a, b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$, or

  There exist distinct $g, g' \in \mathcal{G}_{\mathsf{and}}(f)$ such that

$$(s_b \oplus x_b)\Delta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g) = (s_{b'} \oplus x_{b'})\Delta \oplus G_0^{g'} \oplus U_0^{g'} \oplus \sigma(X_{a'} \oplus k_0^{g'}) \quad \vee \tag{5}$$
$$x_a\Delta \oplus X_a \oplus G_1^g \oplus U_1^g \oplus \sigma(X_b \oplus k_1^g) = x_{a'}\Delta \oplus X_{a'} \oplus G_1^{g'} \oplus U_1^{g'} \oplus \sigma(X_{b'} \oplus k_1^{g'}) \quad \vee \tag{6}$$
$$(s_b \oplus x_b)\Delta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g) = x_{a'}\Delta \oplus X_{a'} \oplus G_1^{g'} \oplus U_1^{g'} \oplus \sigma(X_{b'} \oplus k_1^{g'}) \quad \vee \tag{7}$$
$$x_a\Delta \oplus X_a \oplus G_1^g \oplus U_1^g \oplus \sigma(X_b \oplus k_1^g) = (s_{b'} \oplus x_{b'})\Delta \oplus G_0^{g'} \oplus U_0^{g'} \oplus \sigma(X_{a'} \oplus k_0^{g'}) \tag{8}$$

  where $(a, b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$ and $(a', b') := (\mathsf{in}_0(g'), \mathsf{in}_1(g'))$.

  In this case, $\mathcal{A}$ can check the consistency between the value of $G_u^g \oplus G_{u'}^{g'}$ and that of Δ at the end of experiment *without* further sending required queries, which are computed from Δ, to a random permutation or its inverse. In the real world, the consistency certainly holds. However, the ideal-world garbled rows and Δ are independently sampled, leading to the consistency only with negligible probability. So, $\mathcal{A}$ has non-negligible advantage to distinguish the two worlds and the statistical distance, as an upper bound, also blows up.

  More specifically, the real world is as follows in this case. The pre-image collision (2) leads to the syntactically same XOR of two hash masks in $G_u^g, G_{u'}^{g'}$, which can be XORed to cancel all hash masks to check the consistency with Δ without further queries. Moreover, the image collision (3) also implies the pre-image collision (2) since π is permutation. The other equalities imply the image collision $\pi(X_w \oplus \Delta \oplus k_u^g) = \pi(X_{w'} \oplus \Delta \oplus k_{u'}^{g'})$ for some distinct tuple $(g, u), (g', u') \in \mathcal{G}_{\mathsf{and}}(f) \times \{0, 1\}$, $w := \mathsf{in}_u(g)$, and $w' := \mathsf{in}_{u'}(g')$. Given permutation π, this collision implies the pre-image collision (2), which can be used to see the consistency. However, the above cancelling of hash masks will not give this consistency except with negligible probability in the ideal world.

- bad_2. There exists $((\alpha,\beta),g) \in \cup_{\ell=1}^{3}\mathcal{K}_\ell \times \mathcal{G}_{\mathsf{and}}(f)$ such that

$$\alpha = X_a \oplus \Delta \oplus k_0^g \quad \vee \tag{9}$$
$$\alpha = X_b \oplus \Delta \oplus k_1^g \quad \vee \tag{10}$$
$$\beta = \sigma(\Delta) \oplus (s_b \oplus x_b)\Delta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g) \quad \vee \tag{11}$$
$$\beta = \sigma(\Delta) \oplus x_a\Delta \oplus X_a \oplus G_1^g \oplus U_1^g \oplus \sigma(X_b \oplus k_1^g) \tag{12}$$

  where $(a,b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$.

  In this case, $\mathcal{A}$ essentially makes it to guess Δ before receiving this value. It allows $\mathcal{A}$ to distinguish the real world, where every G_i^g is consistent with Δ, and the ideal world with a dummy Δ. So, the statistical distance blows up.

- bad_3. There exists $((\alpha,\beta),g) \in \cup_{\ell=1}^{2}\mathcal{K}_\ell \times \mathcal{G}_{\mathsf{and}}(f)$ such that

$$\alpha = X_a \oplus k_0^g \quad \vee \tag{13}$$
$$\alpha = X_b \oplus k_1^g \quad \vee \tag{14}$$
$$\beta = U_0^g \oplus \sigma(X_a \oplus k_0^g) \quad \vee \tag{15}$$
$$\beta = U_1^g \oplus \sigma(X_b \oplus k_1^g) \tag{16}$$

  where $(a,b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$.

  In this case, $\mathcal{A}$ can make forward/backward queries w.r.t. some active labels before receiving active input labels and computing other active ones.

  This case is necessary to ensure successful programming in the ideal world as the values on the right hand should not be queried before the programming (otherwise it fails due to pre-image or image collision). If the programming fails in the ideal world, the two worlds can be distinguishable as the decoding consistency in the ideal world does not always hold as in the real world.

Bounding $1-\varepsilon_2$. Without loss of generality, we can consider some fixed good transcript τ such that $\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[Y(\omega)=\tau] \neq 0$ (if this probability is zero, it is trivial by definition that $\varepsilon_2 = 0$ for this τ). Using Lemma 1, we turn to analyze the sampled oracle's compatibility (Section 3.4) with such a transcript, instead of the interaction between $\mathcal{A}$ and the sampled oracle.

Note that there is a computationally unbounded non-uniform adversary $\mathcal{A}'$ such that, for every oracle ω, it sends the queries in τ in order in its interaction with ω (e.g., $\mathcal{A}'$ has auxiliary input τ and sends its ordered queries). Fix $\mathcal{A}'$ in the following compatibility analysis so that any real-world or ideal-world oracle will receive the queries in τ in order. For a response c recorded in a fixed τ, let $\omega \vdash c$ denote the event that, fixing the ordered queries as per τ, oracle ω produces c given the corresponding query. Let $\mathcal{K}^{\mathsf{R}}$ denote the order-preserving list of the responses in an ordered list $\mathcal{K}$ of permutation query-response pairs.

First, we compute $\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)]$. Following from half-gates, a real-world oracle $\omega = (\Delta, \{W_i\}_{i\in\mathcal{W}_{\mathsf{in}}(f)}, \pi) \in \Omega_{\mathsf{real}}$. It holds that

$$\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)] = \Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, \widehat{f}, \mathcal{K}_2^{\mathsf{R}}, \widetilde{X}, \widetilde{x}, \widetilde{U}, \mathcal{K}_3^{\mathsf{R}}, \Delta)]\,.$$

To begin with, every real-world ω certainly produces f' (i.e., the first value in $\widehat{f}$) and $\widetilde{x}$ fixed in τ, which leads to $\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[Y(\omega)=\tau] \neq 0$. This non-zero probability implies that $(f',\widetilde{x})$ in τ is honestly and deterministically computed from the fixed queries (f,x). Otherwise, no ideal-world oracle, which computes $(f',\widetilde{x})$ from the same deterministic procedure, can produce this transcript, contradicting the non-zero probability. As every real-world ω will follow the same deterministic procedure, it certainly produces the two values.

Meanwhile, a real-world oracle ω should have the same literal value of Δ as its counterpart in τ. Conditioned on the compatibility so far, the probability

$$\begin{aligned}
&\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)] \\
&\quad= \Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, F, d, \mathcal{K}_2^{\mathsf{R}}, \widetilde{X}, \widetilde{U}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (f', \widetilde{x}) \wedge \omega \vdash \Delta\right] \\
&\qquad\cdot \Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega \vdash (f', \widetilde{x}) \wedge \omega \vdash \Delta\right] \\
&\quad= \Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, F, d, \mathcal{K}_2^{\mathsf{R}}, \widetilde{X}, \widetilde{U}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (f', \widetilde{x}) \wedge \omega \vdash \Delta\right] \\
&\qquad\cdot \Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega \vdash (f', \widetilde{x})\right] \cdot \Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega \vdash \Delta\right] \\
&\quad= \Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, F, d, \mathcal{K}_2^{\mathsf{R}}, \widetilde{X}, \widetilde{U}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (f', \widetilde{x}) \wedge \omega \vdash \Delta\right] \cdot \frac{1}{2^{\lambda-1}}.
\end{aligned}$$

Conditioned on the compatibility with $(f', \widetilde{x}, \Delta)$, a real-world ω should also be compatible with $(\cup_{\ell=1}^{3}\mathcal{K}_\ell^{\mathsf{R}}, F, \widetilde{U})$ and some active labels in $\widetilde{X}$ such that

(i) $\pi^{\pm 1}$ maps the fixed permutation queries to the responses in $\cup_{\ell=1}^{3}\mathcal{K}_\ell^{\mathsf{R}}$.

(ii) For each $i \in \mathcal{W}_{\mathsf{in}}(f)$, it holds that $X_i = W_i \oplus x_i\Delta$.

(iii) For each $g \in \mathcal{G}_{\mathsf{and}}(f)$ with $(a, b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$, it holds that

$$\begin{aligned}
\mathsf{H}(X_a, k_0^g) &:= \pi(X_a \oplus k_0^g) \oplus \sigma(X_a \oplus k_0^g) = U_0^g, \\
\mathsf{H}(X_b, k_1^g) &:= \pi(X_b \oplus k_1^g) \oplus \sigma(X_b \oplus k_1^g) = U_1^g.
\end{aligned} \tag{17}$$

(iv) For each $g \in \mathcal{G}_{\mathsf{and}}(f)$ with $(a, b, c) := (\mathsf{in}_0(g), \mathsf{in}_1(g), \mathsf{out}(g))$, it holds that

$$X_c = \mathsf{H}(X_a, k_0^g) \oplus \mathsf{H}(X_b, k_1^g) \oplus s_a G_0^g \oplus s_b(G_1^g \oplus X_a), \tag{18}$$

$$\begin{aligned}
G_0^g &= \mathsf{H}(X_a, k_0^g) \oplus \mathsf{H}(X_a \oplus \Delta, k_0^g) \oplus (s_b \oplus x_b)\Delta \\
G_1^g &= \mathsf{H}(X_b, k_1^g) \oplus \mathsf{H}(X_b \oplus \Delta, k_1^g) \oplus x_a\Delta \oplus X_a,
\end{aligned} \tag{19}$$

where the bits $x_a, x_b, s_a = \mathsf{lsb}(X_a), s_b = \mathsf{lsb}(X_b)$ are given in τ.

Conditioned on the compatibility so far, every real-world oracle ω is always compatible with the leftover values in τ, i.e., decoding table d and other active labels in $\widetilde{X}$, which are deterministically computed from XOR combination. The reason is that, for τ ensuring $\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] \neq 0$, these values should be honestly determined by the conditioned values as in the real world. Otherwise, this probability will be zero for an ideal-world oracle, which obtains them from a consistent deterministic computation as per the conditioned values. As every real-world oracle ω honestly follows the real-world computation, this "leftover" compatibility must hold conditioned on the previous compatibility.

It remains to compute the conditional probabilities for (i) to (iv). Consider (iii) and (iv). We note that every good τ with $\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] \neq 0$ already implies (17) and (18). To see this, one can check that condition $\neg\mathsf{bad}_3$ for good transcripts and the extra queries ensure that $\mathcal{K}_3$ fixes the pairs of permutation pre-images and images for hash values

$$\{\mathsf{H}(X_a, k_0^g), \mathsf{H}(X_b, k_1^g)\}_{g\in\mathcal{G}_{\mathsf{and}}(f),(a,b):=(\mathsf{in}_0(g),\mathsf{in}_1(g))}.$$

These values are consistent with $\widetilde{U}$ fixed in τ as per (17). Otherwise, $\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] \neq 0$ cannot be satisfied by τ since $\neg\mathsf{bad}_1 \wedge \neg\mathsf{bad}_3$ for every good transcript implies successful programming in the ideal world so that $\mathsf{H}(X_a, k_0^g) = U_0^g$ and $\mathsf{H}(X_b, k_1^g) = U_1^g$. As a corollary, (18) holds for every good transcript according to this consistency and the step 10 and 13 of SimIn.

Then, consider (19), the leftover part of (iii) and (iv). We rewrite (19) as:

$$\mathcal{V} := \left\{ \begin{array}{l} g \in \mathcal{G}_{\mathsf{and}}(f), (a,b) := (\mathsf{in}_0(g), \mathsf{in}_1(g)) : \\ \qquad \underbrace{\pi(X_a \oplus k_0^g)}_{P_{g,0}} \oplus \underbrace{\pi(X_a \oplus \Delta \oplus k_0^g)}_{P_{g,2}} = \sigma(\Delta) \oplus (s_b \oplus x_b)\Delta \oplus G_0^g, \\ \qquad \underbrace{\pi(X_b \oplus k_1^g)}_{P_{g,1}} \oplus \underbrace{\pi(X_b \oplus \Delta \oplus k_1^g)}_{P_{g,3}} = \sigma(\Delta) \oplus x_a\Delta \oplus X_a \oplus G_1^g \end{array} \right\}$$

As τ is a good transcript, there are exactly $4\,|f|$ pairwise distinct permutation pre-images on the left hand (otherwise, there will be a pair of permutation pre-images leading to (2) in bad_1 or a permutation pre-image leading to (9) $\vee$ (10) in bad_2 given the extra queries). $\mathcal{V}$ has exact $4\,|f|$ syntactically different variables $\mathcal{P} := \{P_{g,0}, P_{g,1}, P_{g,2}, P_{g,3}\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}$. They fix the same number of the entries of permutation π in a real-world ω *if and only if their literal values fixed by* τ *are also pairwise distinct*. Note that every good transcript τ indeed fixes *exact one* such assignment of these values for the following reasons:

- (17) already holds for τ, i.e., for $g \in \mathcal{G}_{\mathsf{and}}(f)$ with $(a,b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$,

$$\begin{aligned} P_{g,0} &:= \pi(X_a \oplus k_0^g) = U_0^g \oplus \sigma(X_a \oplus k_0^g), \\ P_{g,1} &:= \pi(X_b \oplus k_1^g) = U_1^g \oplus \sigma(X_b \oplus k_1^g). \end{aligned} \tag{20}$$

  The literal values of $\{P_{g,0}, P_{g,1}\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}$ can be fixed by the responses in $\mathcal{K}_3^{\mathsf{R}}$ given the extra queries and will be pairwise distinct according to the impossible (3) from $\neg\mathsf{bad}_1$.

- For $g \in \mathcal{G}_{\mathsf{and}}(f)$ with $(a,b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))$, $\{P_{g,2}, P_{g,3}\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}$ are literally assigned the following values fixed by τ according to $\mathcal{V}$ and (20):

$$\begin{aligned} P_{g,2} &:= \pi(X_a \oplus \Delta \oplus k_0^g) = \sigma(\Delta) \oplus (s_b \oplus x_b)\Delta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g), \\ P_{g,3} &:= \pi(X_b \oplus \Delta \oplus k_1^g) = \sigma(\Delta) \oplus x_a\Delta \oplus X_a \oplus G_1^g \oplus U_1^g \oplus \sigma(X_b \oplus k_1^g). \end{aligned} \tag{21}$$

  Clearly, one can see that the literal values of $\{P_{g,2}, P_{g,3}\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}$ are pairwise distinct according to the impossible (4) $\vee$ (5) $\vee$ (6) $\vee$ (7) $\vee$ (8) from $\neg\mathsf{bad}_1$.

- The goodness of τ also ensures that there do not exist

$$P' \in \{P_{g,0}, P_{g,1}\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}, P'_\Delta \in \{P_{g,2}, P_{g,3}\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}$$

  such that $P' = P'_\Delta$. Otherwise, this equality and (21) ensure that there exist $(g,u) \in \mathcal{G}_{\mathsf{and}}(f) \times \{0,1\}$ and $g' \in \mathcal{G}_{\mathsf{and}}(f)$ such that

$$\begin{aligned} \pi(X_w \oplus k_u^g) &= \pi(X_{a'} \oplus \Delta \oplus k_0^{g'}) \\ &= \sigma(\Delta) \oplus (s_{b'} \oplus x_{b'})\Delta \oplus G_0^{g'} \oplus U_0^{g'} \oplus \sigma(X_{a'} \oplus k_0^{g'}) \quad \vee \\ \pi(X_w \oplus k_u^g) &= \pi(X_{b'} \oplus \Delta \oplus k_1^{g'}) \\ &= \sigma(\Delta) \oplus x_{a'}\Delta \oplus X_{a'} \oplus G_1^{g'} \oplus U_1^{g'} \oplus \sigma(X_{b'} \oplus k_1^{g'}) \end{aligned} \tag{22}$$

  where $w := \mathsf{in}_u(g)$ and $(a', b') := (\mathsf{in}_0(g'), \mathsf{in}_1(g'))$. Recall that $\neg\mathsf{bad}_3$ and the extra queries implies $(X_w \oplus k_u^g, \pi(X_w \oplus k_u^g)) \in \mathcal{K}_3$. As a result, (22) leads to contradiction with the impossible (11) $\vee$ (12) from $\neg\mathsf{bad}_2$.

Putting these cases together, we can see that τ yields a value assignment of $\mathcal{P}$, and this assignment fixes exact $4\,|f|$ entries of real-world permutation π.

Conditioned on $\neg\mathsf{bad}_1 \wedge \neg\mathsf{bad}_3$ of good transcript τ, $2\,|f|$ extra queries are non-repeating and the number of non-repeating queries is $q + 2\,|f| = q_\Sigma$. So, $2\,|f|$ responses in $\cup_{\ell=1}^{3}\mathcal{K}_\ell^{\mathsf{R}}$ for the non-repeating extra queries are fixed by the values in $\mathcal{P}$ while the other $q_\Sigma - 2\,|f| = q$ responses are fixed

by real-world π (conditioned on the values in $\mathcal{P}$). Using (i), (ii), (iii), and (iv) together with the "leftover" compatibility, we have in the real world that

$$\begin{aligned}
&\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega\vdash(\mathcal{K}_1^{\mathsf{R}},F,d,\mathcal{K}_2^{\mathsf{R}},\widetilde{X},\widetilde{U},\mathcal{K}_3^{\mathsf{R}})\;\middle|\;\omega\vdash(f',\widetilde{x})\wedge\omega\vdash\Delta\right]\\
&\quad=\frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}}\cdot\frac{(2^\lambda-4\,|f|-(q_\Sigma-2\,|f|))!}{(2^\lambda)!}\\
&\quad=\frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}}\cdot\frac{1}{(2^\lambda)_{q+4|f|}},\\
\Rightarrow&\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega\in\mathsf{comp}_{\mathsf{real}}(\tau)\right]=\frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}}\cdot\frac{1}{(2^\lambda)_{q+4|f|}}\cdot\frac{1}{2^{\lambda-1}}.
\end{aligned}$$

Second, in the ideal world, we can use a similar argument to show

$$\begin{aligned}
\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\omega\in\mathsf{comp}_{\mathsf{ideal}}(\tau)\right]&=\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\omega\vdash(\mathcal{K}_1^{\mathsf{R}},\mathcal{K}_2^{\mathsf{R}},\mathcal{K}_3^{\mathsf{R}})\;\middle|\;\omega\vdash(\widehat{f},\widetilde{X},\widetilde{x},\widetilde{U},\Delta)\right]\\
&\quad\cdot\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\omega\vdash(\widehat{f},\widetilde{X},\widetilde{x},\widetilde{U},\Delta)\right]\\
&=\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\omega\vdash(\mathcal{K}_1^{\mathsf{R}},\mathcal{K}_2^{\mathsf{R}},\mathcal{K}_3^{\mathsf{R}})\;\middle|\;\omega\vdash(\widehat{f},\widetilde{X},\widetilde{x},\widetilde{U},\Delta)\right]\\
&\quad\cdot\frac{1}{2^{|\mathcal{W}_{\mathsf{in}}(f)|\lambda+2\lambda|f|+(\lambda-1)+2\lambda|f|}}.
\end{aligned}$$

According to the condition $\neg\mathsf{bad}_1\wedge\neg\mathsf{bad}_3$ and the $2\,|f|$ extra queries, there are exact $q+2\,|f|=q_\Sigma$ non-repeating queries. Moreover, $2\,|f|$ responses in $\cup_{\ell=1}^{3}\mathcal{K}_\ell^{\mathsf{R}}$ for the non-repeating extra queries are fixed by the conditioned values but the other responses are fixed by $\mathsf{SimP}^{\pm1}$ for other $q_\Sigma-2\,|f|=q$ queries.

Let $\boldsymbol{Q}_{i-1}$ denote the list $\boldsymbol{Q}$ (which is maintained in internal state $\mathsf{st}_{\mathsf{sim}}$) when it includes $i-1\in[0,q_\Sigma-1]$ pairs (note that $\boldsymbol{Q}$ finally includes q_Σ pairs given the q_Σ non-repeating queries), and $\mathcal{N}\subseteq[1,q_\Sigma]$ denote the index set of these q queries in q_Σ non-repeating queries to $\mathsf{SimP}^{\pm1}(\cdot)$ such that $|\mathcal{N}|=q$. We have

$$\begin{aligned}
&\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\omega\vdash(\mathcal{K}_1^{\mathsf{R}},\mathcal{K}_2^{\mathsf{R}},\mathcal{K}_3^{\mathsf{R}})\;\middle|\;\omega\vdash(\widehat{f},\widetilde{X},\widetilde{x},\widetilde{U},\Delta)\right]\\
&\quad=\prod_{i\in\mathcal{N}}\frac{1}{2^\lambda-|\boldsymbol{Q}_{i-1}|}=\prod_{i\in\mathcal{N}}\frac{1}{2^\lambda-(i-1)}\\
&\quad\le\frac{1}{2^\lambda-2\,|f|}\times\frac{1}{2^\lambda-(2\,|f|+1)}\times\cdots\times\frac{1}{2^\lambda-(q_\Sigma-1)}=\frac{1}{(2^\lambda-2\,|f|)_q},\\
\Rightarrow&\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\omega\in\mathsf{comp}_{\mathsf{ideal}}(\tau)\right]\le\frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}}\cdot\frac{1}{(2^\lambda-2\,|f|)_q\cdot(2^\lambda)^{4|f|}}\cdot\frac{1}{2^{\lambda-1}}.
\end{aligned}$$

So, we can have $\varepsilon_2=0$ since, for every $|f|\ge0$ and every $q\ge0$,

$$\begin{aligned}
\frac{\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega\in\mathsf{comp}_{\mathsf{real}}(\tau)\right]}{\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}\left[\omega\in\mathsf{comp}_{\mathsf{ideal}}(\tau)\right]}&\ge\frac{(2^\lambda-2\,|f|)_q\cdot(2^\lambda)^{4|f|}}{(2^\lambda)_{q+4|f|}}\\
&\ge\frac{(2^\lambda-2\,|f|)_q\cdot(2^\lambda)_{2|f|}\cdot(2^\lambda)^{2|f|}}{(2^\lambda)_{q+4|f|}}\\
&=\frac{(2^\lambda)_{q+2|f|}\cdot(2^\lambda)^{2|f|}}{(2^\lambda)_{q+4|f|}}=\frac{(2^\lambda)^{2|f|}}{(2^\lambda-(q+2\,|f|))_{2|f|}}\ge1.
\end{aligned}$$

Bounding ε_1. We bound the probabilities of the bad events in the *ideal world*. First, consider bad_1. Note that each active label X_i can be written as the XOR of (i) some active circuit input labels, and/or (ii) some active output labels of the *precedent* AND gates, i.e., for $\mathcal{I}_i\ominus\mathcal{J}_i\neq\varnothing$,

$$X_i=\left(\bigoplus\nolimits_{w\in\mathcal{I}_i\subseteq\mathcal{W}_{\mathsf{in}}(f)}X_w\right)\oplus\left(\bigoplus\nolimits_{w\in\mathcal{J}_i\subseteq\mathcal{W}_{\mathsf{and}}(f)}X_w\right)=\bigoplus\nolimits_{w\in\mathcal{I}_i\ominus\mathcal{J}_i}X_w\in\{0,1\}^\lambda.\tag{23}$$

For (2), we can use (23) to rewrite it as

$$\bigoplus_{i\in(\mathcal{I}_w\ominus\mathcal{I}_{w'})\ominus(\mathcal{J}_w\ominus\mathcal{J}_{w'})} X_i = k_u^g \oplus k_{u'}^{g'} \in \{0,1\}^\lambda.$$

According to SimIn, each X_i, which is sampled in the step 1 or computed in the step 13, has at least $\lambda-1$ random non-LSBs. Therefore, the equality holds with probability at most $2^{-(\lambda-1)}$ for some fixed distinct k_u^g and $k_{u'}^{g'}$ (or equivalently, (g,u) and (g',u')). If $\mathcal{I}_w = \mathcal{I}_{w'}$ and $\mathcal{J}_w = \mathcal{J}_{w'}$, this probability will be zero for the distinct k_u^g and $k_{u'}^{g'}$. For (3), the worst case is that U_u^g and $U_{u'}^{g'}$ are used for the same AND gate with an output wire in $\mathcal{Z}(f)$. In this case, the XOR $U_u^g \oplus U_{u'}^{g'}$ has at least $\lambda-1$ random non-LSBs due to some X_c sampled in the step 1 of SimIn. Such non-LSBs are independent of other (previously fixed) active labels, including X_w and $X_{w'}$. So, the equality holds with probability at most $2^{-(\lambda-1)}$ for some fixed (g,u) and (g',u').

In the same gate g, the upper $\lambda-1$ bits of $U_0^g \oplus U_1^g$ are uniform so that (4) holds with probability at most $2^{-(\lambda-1)}$. Otherwise, the distinct $g \neq g'$ implies that, for every $u,u' \in \{0,1\}$, $U_u^g \oplus U_{u'}^{g'} \in \{0,1\}^\lambda$ is uniform. As a result, each of (5), (6), (7), and (8) holds with probability $2^{-\lambda}$.

Taking a union bound over the above cases, we have

$$\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\mathsf{bad}_1] \le \frac{2|f|(2|f|-1)+|f|^2}{2^{\lambda-1}} = \frac{5|f|^2-2|f|}{2^{\lambda-1}}. \tag{28}$$

Then, consider bad_2. From (9) (resp., (10)), we have $\Delta = \alpha \oplus X_a \oplus k_0^g$ (resp., $\Delta = \alpha \oplus X_b \oplus k_1^g$), occurring with probability $2^{-(\lambda-1)}$ due to the randomness of Δ. In (11), if $s_b \oplus x_b = 0$, linear orthomorphism σ ensures that

$$\Delta = \sigma^{-1}(\beta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g)),$$

which occurs with probability $2^{-(\lambda-1)}$; if $s_b \oplus x_b = 1$, according to permutation $\sigma'(x) := \sigma(x) \oplus x$ well-defined from σ, it holds with the same probability that

$$\Delta = \sigma'^{-1}(\beta \oplus G_0^g \oplus U_0^g \oplus \sigma(X_a \oplus k_0^g)).$$

Similar result holds for (12). Taking a union bound over all pairs, we have

$$\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\mathsf{bad}_2] \le \frac{4|f|\cdot(q_1+q_2+q_3)}{2^{\lambda-1}}. \tag{29}$$

Finally, consider bad_3. Recall that each X_i has at least $\lambda-1$ random non-LSBs. Since these bits are independent of $\cup_{\ell=1}^2 \mathcal{K}_\ell$, each of (13) and (14) holds with probability at most $2^{-(\lambda-1)}$ for some fixed $(\alpha,\dots)$ and g. For each of (15) and (16), the mask sampled in the step 9 of SimIn (resp., the direct sampling in the step 9 or 12 of SimIn) ensures that $U_0^g \in \{0,1\}^\lambda$ (resp., $U_1^g \in \{0,1\}^\lambda$) is uniform and independent of X_a (resp., X_b) or $\cup_{\ell=1}^2 \mathcal{K}_\ell$. So, each equality holds with probability $2^{-\lambda} < 2^{-(\lambda-1)}$ for some fixed $(\dots,\beta)$ and g. It follows from a union bound that

$$\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\mathsf{bad}_3] \le \frac{4|f|\cdot(q_1+q_2)}{2^{\lambda-1}}. \tag{30}$$

We have a bound ε_1 from (28), (29), and (30):

$$\begin{aligned}
\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[Y(\omega)\in\mathcal{T}_{\mathsf{bad}}] &= \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\mathsf{bad}_1\vee\mathsf{bad}_2\vee\mathsf{bad}_3] \\
&\le \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\mathsf{bad}_1] + \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\mathsf{bad}_2] + \Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\mathsf{bad}_3] \\
&\le \frac{5|f|^2-2|f|+8|f|\cdot(q+2|f|)}{2^{\lambda-1}} \\
&= \frac{8qs+21s^2-2s}{2^{\lambda-1}} = \varepsilon_1.
\end{aligned}$$

The above ε_1, ε_2 and the H-coefficient technique lead to this theorem. □

Public parameters: (Concrete instantiations are recalled in Appendix F)

- $\boldsymbol{M} \in \mathbb{F}_2^{8\times 6}$ with $\boldsymbol{K} \in \mathbb{F}_2^{3\times 8}$ from the co-kernel basis of $\boldsymbol{M}$, i.e., $\boldsymbol{K}\boldsymbol{M} = \boldsymbol{0}_{3\times 6}$.
- $\boldsymbol{V} = \left[\boldsymbol{V}_{00}\ \boldsymbol{V}_{01}\ \boldsymbol{V}_{10}\ \boldsymbol{V}_{11}\right]^\mathsf{T} \in (\mathbb{F}_2^{2\times 5})^4 \equiv \mathbb{F}_2^{8\times 5}$ with left inverse $\boldsymbol{V}^+ \in \mathbb{F}_2^{5\times 8}$.
- Basis matrices $\boldsymbol{S}_L, \boldsymbol{S}_R \in \mathbb{F}_2^{2\times 4}$.
- Control matrices $\boldsymbol{R}'_1, \boldsymbol{R}'_2 \in \mathbb{F}_2^{4\times 2}$, $\boldsymbol{R}_p = \left[\boldsymbol{R}_{p,00}\ \boldsymbol{R}_{p,01}\ \boldsymbol{R}_{p,10}\ \boldsymbol{R}_{p,11}\right]^\mathsf{T} \in (\mathbb{F}_2^{2\times 4})^4 \equiv \mathbb{F}_2^{8\times 4}$.
- A distribution $\mathcal{R}_0$ over $(\mathbb{F}_2^{1\times 2})^4 \equiv \mathbb{F}_2^{4\times 2}$ such that, for every $\boldsymbol{R}'_\$ = \left[\boldsymbol{R}'_{\$,00}\ \boldsymbol{R}'_{\$,01}\ \boldsymbol{R}'_{\$,10}\ \boldsymbol{R}'_{\$,11}\right]^\mathsf{T} \leftarrow \mathcal{R}_0$, (i) $\boldsymbol{K}\left[\widehat{\boldsymbol{R}}_{\$,00}\ \widehat{\boldsymbol{R}}_{\$,01}\ \widehat{\boldsymbol{R}}_{\$,10}\ \widehat{\boldsymbol{R}}_{\$,11}\right]^\mathsf{T} = \boldsymbol{0}_{3\times 6}$, where $\widehat{\boldsymbol{R}}_{\$,ij} := \left(\boldsymbol{R}'_{\$,ij} \otimes \boldsymbol{I}_2\right)\begin{bmatrix}\boldsymbol{S}_L\\ \boldsymbol{S}_R\end{bmatrix}\left(\begin{bmatrix}1 & 0 & i\\ 0 & 1 & j\end{bmatrix} \otimes \boldsymbol{I}_2\right)$ for every $i, j \in \mathbb{F}_2$, and (ii) for every $i, j \in \mathbb{F}_2$, the marginal distribution of $\boldsymbol{R}'_{\$,ij}$ is uniform.

TH.SampleR($\boldsymbol{t}$):

1: $\zeta_1 := g(\overline{p_a}, p_b) \oplus g(\overline{p_a}, \overline{p_b})$, $\zeta_2 := g(p_a, \overline{p_b}) \oplus g(\overline{p_a}, \overline{p_b})$

2: $\boldsymbol{R}'_\$ \leftarrow \mathcal{R}_0$, $\begin{bmatrix}\boldsymbol{r}_{00}^\mathsf{T}\\ \boldsymbol{r}_{01}^\mathsf{T}\\ \boldsymbol{r}_{10}^\mathsf{T}\\ \boldsymbol{r}_{11}^\mathsf{T}\end{bmatrix} = \begin{bmatrix}r_{00L} & r_{00R}\\ r_{01L} & r_{01R}\\ r_{10L} & r_{10R}\\ r_{11L} & r_{11R}\end{bmatrix} := \boldsymbol{R}'_\$ \oplus \zeta_1\boldsymbol{R}'_1 \oplus \zeta_2\boldsymbol{R}'_2$

3: **for** $i, j \in \mathbb{F}_2$ **do** $\boldsymbol{R}_{ij} := \left(\boldsymbol{r}_{ij}^\mathsf{T} \otimes \boldsymbol{I}_2\right)\begin{bmatrix}\boldsymbol{S}_L\\ \boldsymbol{S}_R\end{bmatrix} \oplus \boldsymbol{R}_{p,ij}$, $\widehat{\boldsymbol{R}}_{ij} := \boldsymbol{R}_{ij}\left(\begin{bmatrix}1 & 0 & i\\ 0 & 1 & j\end{bmatrix} \otimes \boldsymbol{I}_2\right)$

4: $\widehat{\boldsymbol{R}} := \left[\widehat{\boldsymbol{R}}_{00}\ \widehat{\boldsymbol{R}}_{01}\ \widehat{\boldsymbol{R}}_{10}\ \widehat{\boldsymbol{R}}_{11}\right]^\mathsf{T}$, $\boldsymbol{r} := \left[r_{00L}\ r_{00R}\ r_{01L}\ r_{01R}\ r_{10L}\ r_{10R}\ r_{11L}\ r_{11R}\right]^\mathsf{T}$

5: **return** $(\widehat{\boldsymbol{R}}, \boldsymbol{r})$

TH.DecodeR($\boldsymbol{r}_{ij}, i, j$):

1: **return** $\boldsymbol{R}_{ij} := \left(\boldsymbol{r}_{ij}^\mathsf{T} \otimes \boldsymbol{I}_2\right)\begin{bmatrix}\boldsymbol{S}_L\\ \boldsymbol{S}_R\end{bmatrix} \oplus \boldsymbol{R}_{p,ij}$

Figure 6: Three-halves garbling scheme [RR21] (Part I).

D Adaptive Security of Three-Halves

We consider the three-halves scheme [RR21] under the computation optimization that makes full use of both halves of hash outputs (see Section 6.2 therein). It uses a counter per wire rather than an identifier per gate to compute tweaks in hash computation. We begin with the implementation in the pRPM that sends decoding table d in the offline phase, followed by another implementation in the npRPM that postpones d to the online phase.

Note that three-halves uses an equivalent invariant for each wire i: its active label $\boldsymbol{x}_i = \boldsymbol{w}_i \oplus (p_i \oplus x_i)\boldsymbol{\Delta}$ for its wire label $\boldsymbol{w}_i$ of bit 0 with $\mathsf{lsb}(\boldsymbol{w}_i) = 0$, wire truth value x_i, permute bit p_i, and global key $\boldsymbol{\Delta}$ with $\mathsf{lsb}(\boldsymbol{\Delta}) = 1$.

For some $\pi \in \mathcal{P}_\ell$ and $\boldsymbol{x} = \left[x_L\ x_R\right]^\mathsf{T} \in \mathbb{F}_{2^{\ell/2}}^2$, we slightly abuse the notation $\boldsymbol{y} := \pi(\boldsymbol{x})$ for the following computation: $(y_L \,\|\, y_R) := \pi(x_L \,\|\, x_R) \in \{0,1\}^\ell$ and $\boldsymbol{y} := \left[y_L\ y_R\right]^\mathsf{T} \in \mathbb{F}_{2^{\ell/2}}^2$. We refer to Appendix A for the additional notation.

D.1 Adaptive Security in pRPM

Theorem 6. *Let* $\mathsf{H}(\boldsymbol{x}, k) = \pi((\boldsymbol{0} \,\|\, \boldsymbol{x}) \oplus k) \oplus \sigma((\boldsymbol{0} \,\|\, \boldsymbol{x}) \oplus k)$ *be a tweakable hash function where* $\boldsymbol{x} \in \mathbb{F}_{2^{\lambda/2}}^2, k \in \mathbb{F}_{2^{\lambda/2+1}}^2$, $\pi \in \mathcal{P}_{\lambda+2}$ *is random permutation, and* $\sigma : \mathbb{F}_{2^{\lambda/2+1}}^2 \to \mathbb{F}_{2^{\lambda/2+1}}^2$ *is a linear orthomorphism for the function family*

$$\mathcal{L} = \left\{L_{\xi_1,\xi_2,\xi_3,\xi_4} : \mathbb{F}_{2^{\lambda/2+1}}^2 \to \mathbb{F}_{2^{\lambda/2+1}}^2\right\}_{\xi_1,\xi_2,\xi_3,\xi_4 \in \mathbb{F}_2}, \quad L_{\xi_1,\xi_2,\xi_3,\xi_4}\left(\begin{bmatrix}x_L\\ x_R\end{bmatrix}\right) = \begin{bmatrix}\xi_1 x_L \oplus \xi_2 x_R\\ \xi_3 x_L \oplus \xi_4 x_R\end{bmatrix}.$$

Then, three-halves (Figure 6) is a $(\lambda+2)$*-garbling scheme with* (q, s, ε)*-adaptive security in the pRPM, where* $\varepsilon = (48qs + 189s^2)/2^{\lambda+1}$.

Proof. The correctness has been proved in the original work [RR21]. We only need to consider the simulation.

$\underline{\mathsf{TH.Garble}^{\pi^{\pm 1}(\cdot)}(f)}$:

$\boldsymbol{\Delta} \leftarrow \begin{bmatrix} \mathbb{F}_{2^{\lambda/2}} \\ \mathbb{F}_{2^{\lambda/2-1}} \,\|\, 1 \end{bmatrix}$
for $i \in \mathcal{W}_{\mathsf{in}}(f)$ **do**
  $p_i \leftarrow \mathbb{F}_2, \boldsymbol{w}_i \leftarrow \begin{bmatrix} \mathbb{F}_{2^{\lambda/2}} \\ \mathbb{F}_{2^{\lambda/2-1}} \,\|\, 0 \end{bmatrix}$
for $i \in \mathcal{W}_{\cup}(f)$ **do** $\mathsf{ctr}_i := 0$
for $g \in \mathcal{G}(f)$ in topology order **do**
  $(a, b, c) := (\mathsf{in}_0(g), \mathsf{in}_1(g), \mathsf{out}(g))$
  if $\mathsf{type}(g) = \texttt{XOR}$ **then** $p_c := p_a \oplus p_b$, $\boldsymbol{w}_c := \boldsymbol{w}_a \oplus \boldsymbol{w}_b$
  else if $\mathsf{type}(g) = \texttt{AND}$ **then**
    $\Gamma := \mathsf{in}_2(g)$
    $(\chi_a, \chi_b, \chi_\Gamma) := (\lfloor \mathsf{ctr}_a/2 \rfloor, \lfloor \mathsf{ctr}_b/2 \rfloor, \lfloor \mathsf{ctr}_\Gamma/2 \rfloor)$
    $(\rho_a, \rho_b, \rho_\Gamma) := (\mathsf{lsb}(\mathsf{ctr}_a), \mathsf{lsb}(\mathsf{ctr}_b), \mathsf{lsb}(\mathsf{ctr}_\Gamma))$
    $\boldsymbol{t}^g := \left[g(p_a, p_b) \;\; g(p_a, \overline{p_b}) \;\; g(\overline{p_a}, p_b) \;\; g(\overline{p_a}, \overline{p_b}) \right]^{\mathsf{T}} \in \mathbb{F}_2^4$
    $(\widehat{\boldsymbol{R}}^g, \boldsymbol{r}^g) \leftarrow \mathsf{TH.SampleR}(\boldsymbol{t}^g)$
    Compute $(\boldsymbol{c}^g, \boldsymbol{g}^g, \boldsymbol{z}^g) \in \mathbb{F}_{2^{\lambda/2}}^2 \times \mathbb{F}_{2^{\lambda/2}}^3 \times \mathbb{F}_2^5$ as follows: for

$$\boldsymbol{h}^g := \begin{bmatrix} \mathsf{Half}_{\rho_a}(\mathsf{H}(\boldsymbol{w}_a, a \,\|\, \chi_a)) \\ \mathsf{Half}_{\rho_a}(\mathsf{H}(\boldsymbol{w}_a \oplus \boldsymbol{\Delta}, a \,\|\, \chi_a)) \\ \mathsf{Half}_{\rho_b}(\mathsf{H}(\boldsymbol{w}_b, b \,\|\, \chi_b)) \\ \mathsf{Half}_{\rho_b}(\mathsf{H}(\boldsymbol{w}_b \oplus \boldsymbol{\Delta}, b \,\|\, \chi_b)) \\ \mathsf{Half}_{\rho_\Gamma}(\mathsf{H}(\boldsymbol{w}_a \oplus \boldsymbol{w}_b, \Gamma \,\|\, \chi_\Gamma)) \\ \mathsf{Half}_{\rho_\Gamma}(\mathsf{H}(\boldsymbol{w}_a \oplus \boldsymbol{w}_b \oplus \boldsymbol{\Delta}, \Gamma \,\|\, \chi_\Gamma)) \end{bmatrix} \in \mathbb{F}_{2^{\lambda/2+1}}^6,$$

$$\left(\boldsymbol{z}^g \,\middle\|\, \begin{bmatrix} \boldsymbol{c}^g \\ \boldsymbol{g}^g \end{bmatrix} \right) := \boldsymbol{V}^+ \left(\boldsymbol{M}\boldsymbol{h}^g \oplus \left(\boldsymbol{r}^g \,\middle\|\, (\widehat{\boldsymbol{R}}^g \oplus ([\boldsymbol{0}_{4\times 2} \;\; \boldsymbol{t}^g] \otimes \boldsymbol{I}_2)) \begin{bmatrix} \boldsymbol{w}_a \\ \boldsymbol{w}_b \\ \boldsymbol{\Delta} \end{bmatrix} \right) \right)$$

    $p_c := \mathsf{lsb}(\boldsymbol{c}^g)$, $\boldsymbol{w}_c := \boldsymbol{c}^g \oplus p_c \boldsymbol{\Delta}$
    $\mathsf{ctr}_a := \mathsf{ctr}_a + 1, \mathsf{ctr}_b := \mathsf{ctr}_b + 1, \mathsf{ctr}_\Gamma := \mathsf{ctr}_\Gamma + 1$
for $i \in \mathcal{W}_{\mathsf{out}}(f)$ **do** $d_i := p_i$
return $\widehat{f} := (f' := f, F := \{(\boldsymbol{g}^g, \boldsymbol{z}^g)\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}, d), k := (f, \boldsymbol{\Delta}, p, \boldsymbol{w})$

$\underline{\mathsf{TH.Encode}^{\pi^{\pm 1}(\cdot)}(k, x)}$:

for $i \in \mathcal{W}_{\mathsf{in}}(f)$ **do**
  $\boldsymbol{x}_i := \boldsymbol{w}_i \oplus (p_i \oplus x_i)\boldsymbol{\Delta}$
return $\widehat{x} := \{\boldsymbol{x}_i\}_{i \in \mathcal{W}_{\mathsf{in}}(f)}$

$\underline{\mathsf{TH.DecEval}^{\pi^{\pm 1}(\cdot)}(\widehat{f}, \widehat{x})}$:

for $i \in \mathcal{W}_{\cup}(f)$ **do** $\mathsf{ctr}_i := 0$
for $g \in \mathcal{G}(f)$ in topology order **do**
  $(a, b, c) := (\mathsf{in}_0(g), \mathsf{in}_1(g), \mathsf{out}(g))$
  if $\mathsf{type}(g) = \texttt{XOR}$ **then** $\boldsymbol{x}_c := \boldsymbol{x}_a \oplus \boldsymbol{x}_b$
  else if $\mathsf{type}(g) = \texttt{AND}$ **then**
    $\Gamma := \mathsf{in}_2(g)$
    $(\chi_a, \chi_b, \chi_\Gamma) := (\lfloor \mathsf{ctr}_a/2 \rfloor, \lfloor \mathsf{ctr}_b/2 \rfloor, \lfloor \mathsf{ctr}_\Gamma/2 \rfloor)$
    $(\rho_a, \rho_b, \rho_\Gamma) := (\mathsf{lsb}(\mathsf{ctr}_a), \mathsf{lsb}(\mathsf{ctr}_b), \mathsf{lsb}(\mathsf{ctr}_\Gamma))$
    $s_a := \mathsf{lsb}(\boldsymbol{x}_a), s_b := \mathsf{lsb}(\boldsymbol{x}_b)$
    $(\boldsymbol{r}^g_{s_a s_b} \,\|\, \boldsymbol{m}^g_{s_a s_b}) := \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mathsf{Half}_{\rho_a}(\mathsf{H}(\boldsymbol{x}_a, a \,\|\, \chi_a)) \\ \mathsf{Half}_{\rho_b}(\mathsf{H}(\boldsymbol{x}_b, b \,\|\, \chi_b)) \\ \mathsf{Half}_{\rho_\Gamma}(\mathsf{H}(\boldsymbol{x}_a \oplus \boldsymbol{x}_b, \Gamma \,\|\, \chi_\Gamma)) \end{bmatrix} \oplus \boldsymbol{V}_{s_a s_b} \left(\boldsymbol{z}^g \,\middle\|\, \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{g}^g \end{bmatrix} \right)$
    $\boldsymbol{R}^g_{s_a s_b} := \mathsf{TH.DecodeR}(\boldsymbol{r}^g_{s_a s_b}, s_a, s_b)$
    $\boldsymbol{x}_c := \boldsymbol{m}^g_{s_a s_b} \oplus \boldsymbol{R}^g_{s_a s_b} \begin{bmatrix} \boldsymbol{x}_a \\ \boldsymbol{x}_b \end{bmatrix}$
    $\mathsf{ctr}_a := \mathsf{ctr}_a + 1, \mathsf{ctr}_b := \mathsf{ctr}_b + 1, \mathsf{ctr}_\Gamma := \mathsf{ctr}_\Gamma + 1$
for $i \in \mathcal{W}_{\mathsf{out}}(f)$ **do** $y_i := d_i \oplus \mathsf{lsb}(\boldsymbol{x}_i)$
return y

Figure 6: Three-halves garbling scheme [RR21] (Part II).

$\underline{\mathsf{SimF}(f)}$:

1: $F \leftarrow (\mathbb{F}_{2^{\lambda/2}}^3 \times \mathbb{F}_2^5)^{|f|}$, $d \leftarrow \mathbb{F}_2^{|\mathcal{W}_{\mathsf{out}}(f)|}$ such that

(i) For each $i \in [1,m]$ and each wire $j \in \mathcal{Y}_i(f)$, $d_j = d_{y_i(f)}$.

(ii) For each $(\mu_1, \ldots, \mu_m) \in \mathbb{F}_2^m$ such that $\ominus_{i\in[1,m],\mu_i=1}\mathcal{X}_i(f) = \varnothing$, $\oplus_{i\in[1,m]}\mu_i \cdot d_{y_i(f)} = 0$.

2: **return** $\widehat{f} := (f' := f, F, d)$, $\mathsf{st}_{\mathsf{sim}} := (\mathsf{st}_{\mathsf{sim}}, f, F, d)$

$\underline{\mathsf{SimIn}(f(x))}$:

1: $\{\boldsymbol{x}_i\}_{i\in\mathcal{X}(f)} \leftarrow (\mathbb{F}_{2^{\lambda/2}}^2)^{|\mathcal{X}(f)|}$ such that, for each $j \in [1,m]$, $\mathsf{lsb}(\oplus_{i\in\mathcal{X}_j(f)}\boldsymbol{x}_i) = d_{y_j(f)} \oplus f(x)_{y_j(f)}$.

2: **for** $i \in \mathcal{W}_\cup(f)$ **do** $\mathsf{ctr}_i := 0$

3: **for** $g \in \mathcal{G}(f)$ in topology order **do**

4: $(a,b,c) := (\mathsf{in}_0(g), \mathsf{in}_1(g), \mathsf{out}(g))$

5: **if** $\mathsf{type}(g) = \texttt{XOR}$ **then** $\boldsymbol{x}_c := \boldsymbol{x}_a \oplus \boldsymbol{x}_b$

6: **else if** $\mathsf{type}(g) = \texttt{AND}$ **then**

7: $\Gamma := \mathsf{in}_2(g)$

8: $(\chi_a, \chi_b, \chi_\Gamma) := (\lfloor \mathsf{ctr}_a/2 \rfloor, \lfloor \mathsf{ctr}_b/2 \rfloor, \lfloor \mathsf{ctr}_\Gamma/2 \rfloor)$

9: $(\rho_a, \rho_b, \rho_\Gamma) := (\mathsf{lsb}(\mathsf{ctr}_a), \mathsf{lsb}(\mathsf{ctr}_b), \mathsf{lsb}(\mathsf{ctr}_\Gamma))$

10: $s_a := \mathsf{lsb}(\boldsymbol{x}_a)$, $s_b := \mathsf{lsb}(\boldsymbol{x}_b)$

11: **if** $c \in \mathcal{Z}(f)$ **then**

12: $\boldsymbol{r}^g_{s_a s_b} \leftarrow \mathbb{F}_2^2$, $\boldsymbol{R}^g_{s_a s_b} := \mathsf{TH.DecodeR}(\boldsymbol{r}^g_{s_a s_b}, s_a, s_b)$

13: $\mathsf{Half}_{\rho_\Gamma}(\boldsymbol{u}_\Gamma^{\chi_\Gamma}) \leftarrow \mathbb{F}_{2^{\lambda/2+1}}$

14: $\begin{bmatrix} \mathsf{Half}_{\rho_a}(\boldsymbol{u}_a^{\chi_a}) \\ \mathsf{Half}_{\rho_b}(\boldsymbol{u}_b^{\chi_b}) \end{bmatrix} := \begin{bmatrix} \mathsf{Half}_{\rho_\Gamma}(\boldsymbol{u}_\Gamma^{\chi_\Gamma}) \\ \mathsf{Half}_{\rho_\Gamma}(\boldsymbol{u}_\Gamma^{\chi_\Gamma}) \end{bmatrix} \oplus \left(\boldsymbol{r}^g_{s_a s_b} \,\middle\|\, \boldsymbol{x}_c \oplus \boldsymbol{R}^g_{s_a s_b} \begin{bmatrix} \boldsymbol{x}_a \\ \boldsymbol{x}_b \end{bmatrix} \right) \oplus \boldsymbol{V}_{s_a s_b} \left(\boldsymbol{z}^g \,\middle\|\, \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{g}^g \end{bmatrix} \right)$

15: **else if** $c \notin \mathcal{Z}(f)$ **then**

16: $\mathsf{Half}_{\rho_a}(\boldsymbol{u}_a^{\chi_a}), \mathsf{Half}_{\rho_b}(\boldsymbol{u}_b^{\chi_b}), \mathsf{Half}_{\rho_\Gamma}(\boldsymbol{u}_\Gamma^{\chi_\Gamma}) \leftarrow \mathbb{F}_{2^{\lambda/2+1}}$

17: $(\boldsymbol{r}^g_{s_a s_b} \,\|\, \boldsymbol{m}^g_{s_a s_b}) := \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mathsf{Half}_{\rho_a}(\boldsymbol{u}_a^{\chi_a}) \\ \mathsf{Half}_{\rho_b}(\boldsymbol{u}_b^{\chi_b}) \\ \mathsf{Half}_{\rho_\Gamma}(\boldsymbol{u}_\Gamma^{\chi_\Gamma}) \end{bmatrix} \oplus \boldsymbol{V}_{s_a s_b} \left(\boldsymbol{z}^g \,\middle\|\, \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{g}^g \end{bmatrix} \right)$

18: $\boldsymbol{R}^g_{s_a s_b} := \mathsf{TH.DecodeR}(\boldsymbol{r}^g_{s_a s_b}, s_a, s_b)$

19: $\boldsymbol{x}_c := \boldsymbol{m}^g_{s_a s_b} \oplus \boldsymbol{R}^g_{s_a s_b} \begin{bmatrix} \boldsymbol{x}_a \\ \boldsymbol{x}_b \end{bmatrix}$

20: $\mathsf{ctr}_a := \mathsf{ctr}_a + 1$, $\mathsf{ctr}_b := \mathsf{ctr}_b + 1$, $\mathsf{ctr}_\Gamma := \mathsf{ctr}_\Gamma + 1$

21: **for** $i \in \mathcal{W}_\cup(f)$ **do**

22: **for** $j \in [0, \lfloor n_i(f)/2 \rfloor - 1]$ **do**

23: (Programming) Add a pair $\Big((\boldsymbol{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j), \boldsymbol{u}_i^j \oplus \sigma\left((\boldsymbol{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)\right) \Big)$ to the list $\boldsymbol{Q}$ kept in $\mathsf{st}_{\mathsf{sim}}$, if it causes no pre-image collision or image collision in $\boldsymbol{Q}$, to ensure

$$\underbrace{\pi\left((\boldsymbol{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)\right) \oplus \sigma\left((\boldsymbol{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)\right)}_{\mathsf{H}(\boldsymbol{x}_i, i \,\|\, j)} = \boldsymbol{u}_i^j = \begin{bmatrix} \mathsf{Half}_1(\boldsymbol{u}_i^j) \\ \mathsf{Half}_0(\boldsymbol{u}_i^j) \end{bmatrix}.$$

24: **if** $n_i(f)$ is odd **then**

25: $j := \lfloor n_i(f)/2 \rfloor$, $\mathsf{Half}_1(\boldsymbol{u}_i^j) \leftarrow \mathbb{F}_{2^{\lambda/2+1}}$ and repeat the step 23 for this j

26: **return** $\widetilde{x} := \{\boldsymbol{x}_i\}_{i\in\mathcal{W}_{\mathsf{in}}(f)}$, $\mathsf{st}_{\mathsf{sim}} := (\mathsf{st}_{\mathsf{sim}}, \widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{r}})$ where $\mathsf{st}_{\mathsf{sim}}$ on the right hand has an updated list $\boldsymbol{Q}$, $\widetilde{\boldsymbol{x}} := \{\boldsymbol{x}_i\}_{i\in\mathcal{W}(f)}$, $\widetilde{\boldsymbol{u}} := \{\boldsymbol{u}_i^j\}_{i\in\mathcal{W}_\cup(f), j\in[0,\lceil n_i(f)/2\rceil-1]}$, $\widetilde{\boldsymbol{r}} := \{\boldsymbol{r}^g_{s_a s_b}\}_{g\in\mathcal{G}_{\mathsf{and}}(f),(a,b):=(\mathsf{in}_0(g),\mathsf{in}_1(g))}$.

Figure 7: Our simulator for three-halves in the pRPM.

Our simulator Sim consists of $(\mathsf{SimF}, \mathsf{SimIn})$ in Figure 7 and $\mathsf{SimP}^{\pm 1}$, which emulates the random permutation and its inverse on-the-fly. More specifically, there is a list $\boldsymbol{Q}$ of query-response pairs in internal state $\mathsf{st}_{\mathsf{sim}}$. Upon receiving forward query α (resp., backward query β) from $\mathcal{A}_i$ to SimP (resp., SimP^{-1}), it reads $\boldsymbol{Q}$ from $\mathsf{st}_{\mathsf{sim}}$ and checks whether $\exists(\alpha,\gamma) \in \boldsymbol{Q}$ (resp., $\exists(\gamma,\beta) \in \boldsymbol{Q}$). If true, it returns γ as response; otherwise it samples $\gamma \leftarrow \{s \in \mathbb{F}_{2^{\lambda/2+1}}^2 \mid (\ldots, s) \notin \boldsymbol{Q}\}$ (resp., $\gamma \leftarrow \{s \in \mathbb{F}_{2^{\lambda/2+1}}^2 \mid (s, \ldots) \notin \boldsymbol{Q}\}$), adds (α,γ) (resp., (γ,β)) to $\boldsymbol{Q}$, and returns γ as response. The programming

is the step 23 of SimIn, where (α, β) is added to $\boldsymbol{Q}$ if and only if there is no pre-image collision with $(\alpha, \dots) \in \boldsymbol{Q}$ or image collision with $(\dots, \beta) \in \boldsymbol{Q}$.

To see that Sim is PPT, we note that the circuit-dependent notation can be efficiently computed by traversing the polynomial-size circuit f. Then, the crux is to show that the step 1 in both SimF and SimIn can be polynomial-time.

The runtime of the step 1 of SimF is dominated by the runtime of iterating through all qualified $(\mu_1, \dots, \mu_m) \in \mathbb{F}_2^m$. To find the qualified vectors, one can interpret each $\mathcal{X}_i(f)$ as a one-hot *non-zero* column vector in the space $\mathbb{F}_2^{|\mathcal{W}(f)|}$ and derive a $|\mathcal{W}(f)|$-by-m matrix $\boldsymbol{E}$ from these column vectors. One can check that all qualified vectors fall in the kernel of $\boldsymbol{E}$. This kernel can be efficiently computed from the Gaussian elimination on $\boldsymbol{E}$ and is a subspace spanned by $m - \mathsf{rank}(\boldsymbol{E})$ basis vectors. So, the step 1 of SimF only needs to iterate through these basis vectors, and the other qualified vectors must satisfy the constraints as they are in the subspace. As a result, the step 1 of SimF runs in polynomial time due to the Gaussian elimination plus a linear-time pass to assign random or constrained values to $d_{y_i(f)}$'s according to the $m - \mathsf{rank}(\boldsymbol{E})$ basis vectors.

The step 1 of SimIn only requires one linear-time pass to assign constrained or random values to the active labels so it runs in polynomial time. The linear constraint on the LSBs of these active labels has rank $\mathsf{rank}(\boldsymbol{E})$ and is satisfiable for the $d_{y_i(f)}$'s assigned in SimF.

Then, we fix z and $\mathcal{A}$. We will use the H-coefficient technique (Section 3.4) with transcript padding to bound the advantage of distinguishing between the real world (i.e., the adaptive experiment that uses the three-halves scheme) and the ideal world (i.e., the adaptive experiment that uses simulator Sim). In this technique, we consider the computationally unbounded non-uniform adversary $\mathcal{A} = \mathcal{A}(z)$ and compute $\varepsilon_1, \varepsilon_2$ as follows.

Transcript padding. In either world, $\mathcal{A}$ will interact with an integrated oracle that acts as the two-round challenger in the adaptive experiment and provides interfaces $\pi^{*\pm 1} \in \{\pi^{\pm 1}, \mathsf{SimP}^{\pm 1}\}$ for forward/backward permutation queries. $\mathcal{A}$ can learn $\pi^*(\alpha) = \beta$ if and only if it sent forward query α to π^* and received response β, or sent backward query β to π^{*-1} and received response α.

To compute $\varepsilon_1, \varepsilon_2$ more easily, we ask the oracle to send more messages to $\mathcal{A}$ and $\mathcal{A}$ to make extra queries (in addition to the supposed q queries) in both two worlds. More specifically,

- Upon receiving x from $\mathcal{A}$, the oracle sends $\widetilde{\boldsymbol{x}} := \{\boldsymbol{x}_i\}_{i \in \mathcal{W}(f)}$ instead of $\widehat{x} := \{\boldsymbol{x}_i\}_{i \in \mathcal{W}_{\mathsf{in}}(f)}$ to $\mathcal{A}$. In addition to the active input labels in $\widehat{x}$, the former also gives the active internal and output labels. In the real world, the oracle can run $\mathsf{TH.DecEval}^{\pi^{\pm 1}(\cdot)}$, which determines other active labels in $\widetilde{\boldsymbol{x}}$. In the ideal world, this $\widetilde{\boldsymbol{x}}$ can be directly output by SimIn.

- Along with $\widetilde{\boldsymbol{x}}$, the oracle sends $\widetilde{x} := \{x_i\}_{i \in \mathcal{W}(f)}$ to $\mathcal{A}$, which denote the wire truth values in the evaluation of $f(x)$. Both two oracles "echo" these values, which are self-evident to $\mathcal{A}$, to explicitly include them in transcripts. In the experiment, the real-world oracle uses $x = \{x_i\}_{i \in \mathcal{W}_{\mathsf{in}}(f)}$ in $\mathsf{TH.Encode}^{\pi^{\pm 1}(\cdot)}$, but the ideal-world oracle can only use $f(x) = \{x_i\}_{i \in \mathcal{W}_{\mathsf{out}}(f)}$ in SimIn.

- Along with $\widetilde{\boldsymbol{x}}$, the oracle sends $\widetilde{\boldsymbol{u}} := \{\boldsymbol{u}_i^j\}_{i \in \mathcal{W}_\cup(f), j \in [0, \lceil n_i(f)/2 \rceil - 1]}$ to $\mathcal{A}$. In the real world, the oracle computes $\boldsymbol{u}_i^j := \mathsf{H}(\boldsymbol{x}_i, i \,\|\, j)$ for each $i \in \mathcal{W}_\cup(f)$ and $j \in [0, \lceil n_i(f)/2 \rceil - 1]$. In the ideal world, this $\widetilde{\boldsymbol{u}}$ is output by SimIn and gives the "hash outputs" fixed by the random tape, which is specified by the oracle to run the programming therein.

- Along with $\widetilde{\boldsymbol{x}}$, the oracle sends $\widetilde{\boldsymbol{r}} := \{\boldsymbol{r}^g_{s_a s_b}\}_{g \in \mathcal{G}_{\mathsf{and}}(f), (a,b) := (\mathsf{in}_0(g), \mathsf{in}_1(g))}$ to $\mathcal{A}$. In the real world, the oracle computes $\boldsymbol{r}^g_{s_a s_b} \in \mathbb{F}_2^2$ from two uniform bits (which span $\boldsymbol{R}'^g_{\$} \leftarrow \mathcal{R}_0$ and match $\boldsymbol{R}'^g_{\$,ij}$ for every $i, j \in \mathbb{F}_2$ as per distribution $\mathcal{R}_0$), the truth table $\boldsymbol{t}^g$ (expressed in terms of the truth values x_a, x_b in $\widetilde{x}$ and the masked bits $s_a = \mathsf{lsb}(\boldsymbol{x}_a), s_b = \mathsf{lsb}(\boldsymbol{x}_b)$ in $\widetilde{\boldsymbol{x}}$), and the two masked bits s_a and s_b. More specifically, it follows from $\mathsf{TH.SampleR}$ that

$$\begin{aligned} \boldsymbol{r}^g_{s_a s_b}{}^{\mathsf{T}} &= \boldsymbol{R}'^g_{\$,s_a s_b} \oplus \Big(g(\overline{p_a}, p_b) \oplus g(\overline{p_a}, \overline{p_b})\Big) \boldsymbol{R}'_{1,s_a s_b} \oplus \Big(g(p_a, \overline{p_b}) \oplus g(\overline{p_a}, \overline{p_b})\Big) \boldsymbol{R}'_{2,s_a s_b} \\ &= \boldsymbol{R}'^g_{\$,s_a s_b} \oplus \overline{s_a \oplus x_a} \cdot \boldsymbol{R}'_{1,s_a s_b} \oplus \overline{s_b \oplus x_b} \cdot \boldsymbol{R}'_{2,s_a s_b}. \end{aligned} \tag{31}$$

In the ideal world, this $\widetilde{\boldsymbol{r}}$ is also output by SimIn as per the random tape of the oracle.

- Along with $\widetilde{\boldsymbol{x}}$, the oracle sends $\widetilde{T} := \{T_i\}_{i\in\mathcal{W}_\cup(f),n_i(f) \text{ is odd}}$ to $\mathcal{A}$. In the real world, the oracle computes $T_i := \mathsf{Half}_1(\mathsf{H}(\boldsymbol{x}_i\oplus\boldsymbol{\Delta}, i \,\|\, \lfloor n_i(f)/2\rfloor)) \in \mathbb{F}_{2^{\lambda/2+1}}$, i.e., the *unused* upper half of the $\lfloor n_i(f)/2\rfloor$-th $\boldsymbol{\Delta}$-related hash output of the wire i. In the ideal world, this $\widetilde{T}$ is sampled by the oracle at random.
- **(Extra queries)** Upon receiving $(\widetilde{\boldsymbol{x}}, \widetilde{x}, \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{r}}, \widetilde{T})$ from the oracle, $\mathcal{A}$ will also make a forward permutation query $(\boldsymbol{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)$ for every $i \in \mathcal{W}_\cup(f)$ and $j \in [0, \lceil n_i(f)/2\rceil - 1]$, if it has never learned $\pi^*((\boldsymbol{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)) = \boldsymbol{y}$ for some $\boldsymbol{y}$ in its interaction with $\pi^{*\pm1} \in \{\pi^{\pm1}, \mathsf{SimP}^{\pm1}\}$.
- At the end of the experiment (i.e., once all other transcripts are settled), the oracle sends $\boldsymbol{\Delta}$ to $\mathcal{A}$. In the real world, the oracle gets this $\boldsymbol{\Delta}$ from the output of $\mathsf{TH.Garble}^{\pi^{\pm1}(\cdot)}$. In the ideal world, $\boldsymbol{\Delta}$ is dummy and sampled by the oracle at this time (note that Sim does not use $\boldsymbol{\Delta}$).

According to the two oracle constructions, real-world sample space

$$\Omega_{\mathsf{real}} = \mathbb{F}_2^{\lambda-1} \times \mathbb{F}_2^{|\mathcal{W}_{\mathsf{in}}(f)|\lambda} \times \underbrace{\mathbb{F}_2^{2|f|}}_{\substack{\text{random tape to run}\\ \mathsf{TH.SampleR}\ |f| \text{ times}}} \times \mathcal{P}_{\lambda+2},$$

and ideal-world sample space

$$\Omega_{\mathsf{ideal}} = (\mathbb{F}_2^{3\lambda/2+5})^{|f|} \times \mathbb{F}_2^{\mathsf{rank}(\boldsymbol{E})} \times \mathbb{F}_2^{(|\mathcal{W}_{\mathsf{in}}(f)|+|\mathcal{Z}(f)|)\lambda-\mathsf{rank}(\boldsymbol{E})} \times (\mathbb{F}_2^{\lambda/2+3})^{|\mathcal{Z}(f)|} \times (\mathbb{F}_2^{3\lambda/2+3})^{|f|-|\mathcal{Z}(f)|}$$
$$\times \underbrace{(\mathbb{F}_2^{\lambda+2})^M}_{\text{for the step 24 of } \mathsf{SimIn} \text{ and } \widetilde{T}} \times \underbrace{\{0,1\}^*}_{\substack{\text{random tape for}\\ \text{the sampling in } \mathsf{SimP}^{\pm1}(\cdot)}} \times \underbrace{\mathbb{F}_2^{\lambda-1}}_{\text{dummy } \boldsymbol{\Delta}}.$$

Given the oracle constructions, a transcript in the original adaptive experiment will be padded with more literal values. Note that transcript padding will not lower the advantage of $\mathcal{A}$ since $\mathcal{A}$ can discard the padding values at will. With the padding, a transcript is of the form:

$$\tau = (\mathcal{K}_1, (f, \widehat{f}), \mathcal{K}_2, (x, (\widetilde{\boldsymbol{x}}, \widetilde{x}, \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{r}}, \widetilde{T})), \mathcal{K}_3, \boldsymbol{\Delta}),$$

where $\mathcal{K}_1$, $\mathcal{K}_2$, and $\mathcal{K}_3$ are the ordered lists of query-response pairs seen in the interleaved interaction with permutation oracles. We do not explicitly consider query direction in these pairs. Given $\pi^{*\pm1} \in \{\pi^{\pm1}, \mathsf{SimP}^{\pm1}\}$, $\mathcal{A}$ is able to learn $\pi^*(\alpha) = \beta$ if and only if there exists $(\alpha, \beta) \in \cup_{\ell=1}^3 \mathcal{K}_\ell$.

Let $q_\ell := |\mathcal{K}_\ell|$ for every $\ell \in \{1,2,3\}$ and $q_\Sigma := \sum_{\ell=1}^3 q_\ell$. It follows from the extra queries that $q_\Sigma \le q + N$. Without loss of generality, we can assume that $\mathcal{A}$ only makes *non-repeating* queries, i.e., it never makes forward query α to π^* or backward query β to π^{*-1} for any learned permutation entry (α, β).

Bad transcripts. A transcript $\tau \in \mathcal{T}_{\mathsf{bad}}$ if and only if it incurs at least one of the following events:

- bad_1. There exist distinct $(i,j) \in \mathcal{W}_\cup(f) \times [0, \lceil n_i(f)/2\rceil - 1], (i',j') \in \mathcal{W}_\cup(f) \times [0, \lceil n_{i'}(f)/2\rceil - 1]$ such that

$$(\boldsymbol{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j) = (\boldsymbol{0} \,\|\, \boldsymbol{x}_{i'}) \oplus (i' \,\|\, j') \quad \vee \tag{32}$$
$$\boldsymbol{u}_i^j \oplus \sigma((\boldsymbol{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)) = \boldsymbol{u}_{i'}^{j'} \oplus \sigma((\boldsymbol{0} \,\|\, \boldsymbol{x}_{i'}) \oplus (i' \,\|\, j')) \quad \vee \tag{33}$$
$$\boldsymbol{y}_i^j \oplus \boldsymbol{u}_i^j \oplus \sigma((\boldsymbol{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)) = \boldsymbol{y}_{i'}^{j'} \oplus \boldsymbol{u}_{i'}^{j'} \oplus \sigma((\boldsymbol{0} \,\|\, \boldsymbol{x}_{i'}) \oplus (i' \,\|\, j')) \tag{34}$$

  where $\boldsymbol{y}_i^j \in \mathbb{F}^2_{2^{\lambda/2+1}}$ will be defined from τ according to (42) and (44).

  In this case, $\mathcal{A}$ is able to check the consistency between the value of $\boldsymbol{y}_i^j \oplus \boldsymbol{y}_{i'}^{j'}$ and that of $\boldsymbol{\Delta}$ at the end of experiment *without* sending any needed queries, which are computed from $\boldsymbol{\Delta}$, to a random

permutation or its inverse. In the real world, the consistency certainly holds. However, the ideal-world garbled rows and $\boldsymbol{\Delta}$ are independently sampled, leading to the consistency only with negligible probability. So, $\mathcal{A}$ has non-negligible advantage to distinguish the two worlds and the statistical distance, as an upper bound, also blows up.

More specifically, the real world is as follows in this case. The pre-image collision (32) leads to the syntactically same XOR of two hash masks in $\boldsymbol{y}_i^j, \boldsymbol{y}_{i'}^{j'}$, which can be XORed to cancel all hash masks to check the consistency with $\boldsymbol{\Delta}$ without further queries. Furthermore, the image collision (33) also implies the pre-image collision (32) as π is permutation. The collision (34) results in the image collision $\pi((\mathbf{0} \,\|\, \boldsymbol{x}_i \oplus \boldsymbol{\Delta}) \oplus (i \,\|\, j)) = \pi((\mathbf{0} \,\|\, \boldsymbol{x}_{i'} \oplus \boldsymbol{\Delta}) \oplus (i' \,\|\, j'))$ for some distinct (i, j) and (i', j'). Since π is permutation, this collision implies the pre-image collision (32), which can be used to see the consistency. In contrast, the above cancelling of hash masks cannot give this consistency except with negligible probability in the ideal world.

- bad_2. There exists $((\alpha, \beta), i, j) \in \cup_{\ell=1}^{3} \mathcal{K}_\ell \times \mathcal{W}_\cup(f) \times [0, \lceil n_i(f)/2 \rceil - 1]$ such that

$$\alpha = (\mathbf{0} \,\|\, \boldsymbol{x}_i \oplus \boldsymbol{\Delta}) \oplus (i \,\|\, j) \quad \vee \tag{35}$$

$$\beta = \sigma(\mathbf{0} \,\|\, \boldsymbol{\Delta}) \oplus \boldsymbol{y}_i^j \oplus \boldsymbol{u}_i^j \oplus \sigma((\mathbf{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)) \tag{36}$$

  where $\boldsymbol{y}_i^j \in \mathbb{F}_{2^{\lambda/2+1}}^2$ will be defined from τ according to (42) and (44).

  In this case, $\mathcal{A}$ essentially makes it to guess $\boldsymbol{\Delta}$ before receiving this value. It allows $\mathcal{A}$ to distinguish the real world, where every $\boldsymbol{y}_i^j$ is consistent with $\boldsymbol{\Delta}$, and the ideal world with a dummy $\boldsymbol{\Delta}$. So, the statistical distance blows up.

- bad_3. There exists $((\alpha, \beta), i, j) \in \cup_{\ell=1}^{2} \mathcal{K}_\ell \times \mathcal{W}_\cup(f) \times [0, \lceil n_i(f)/2 \rceil - 1]$ such that

$$\alpha = (\mathbf{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j) \quad \vee \tag{37}$$

$$\beta = \boldsymbol{u}_i^j \oplus \sigma((\mathbf{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)) \tag{38}$$

  In this case, $\mathcal{A}$ can make forward/backward queries w.r.t. some active labels before receiving active input labels and computing other active ones.

  This case is necessary to ensure successful programming in the ideal world as the values on the right hand should not be queried before the programming (otherwise it fails due to pre-image or image collision). If the programming fails in the ideal world, the two worlds can be distinguishable as the decoding consistency in the ideal world does not always hold as in the real world.

Bounding $1 - \varepsilon_2$. Without loss of generality, we can consider some fixed good transcript τ such that $\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] \neq 0$ (if this probability is zero, it is trivial by definition that $\varepsilon_2 = 0$ for this τ). Using Lemma 1, we turn to analyze the sampled oracle's compatibility (Section 3.4) with such a transcript, instead of the interaction between $\mathcal{A}$ and the sampled oracle.

Note that there is a computationally unbounded non-uniform adversary $\mathcal{A}'$ such that, for every oracle ω, it sends the queries in τ in order in its interaction with ω (e.g., $\mathcal{A}'$ has auxiliary input τ and sends its ordered queries). Fix $\mathcal{A}'$ in the following compatibility analysis so that any real-world or ideal-world oracle will receive the queries in τ in order. For a response c recorded in a fixed τ, let $\omega \vdash c$ denote the event that, fixing the ordered queries as per τ, oracle ω produces c given the corresponding query. Let $\mathcal{K}^{\mathsf{R}}$ denote the order-preserving list of the responses in an ordered list $\mathcal{K}$ of permutation query-response pairs.

First, we compute $\Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)]$. Following from three-halves, a real-world oracle $\omega = (\boldsymbol{\Delta}, \{(p_i, \boldsymbol{w}_i)\}_{i \in \mathcal{W}_{\mathsf{in}}(f)}, \{(r_L^g, r_R^g)\}_{g \in \mathcal{G}_{\mathsf{and}}(f)}, \pi) \in \Omega_{\mathsf{real}}$. So,

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)] = \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, \widehat{f}, \mathcal{K}_2^{\mathsf{R}}, \widetilde{\boldsymbol{x}}, \widetilde{x}, \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{r}}, \widetilde{T}, \mathcal{K}_3^{\mathsf{R}}, \boldsymbol{\Delta})].$$

To begin with, every real-world ω certainly produces f' (i.e., the first value in $\widehat{f}$) and $\widetilde{x}$ fixed in τ, which leads to $\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[Y(\omega)=\tau]\neq 0$. This non-zero probability implies that $(f',\widetilde{x})$ in τ is honestly and deterministically computed from the fixed queries (f,x). Otherwise, no ideal-world oracle, which computes $(f',\widetilde{x})$ from the same deterministic procedure, can produce this transcript, contradicting the non-zero probability. As every real-world ω will follow the same deterministic procedure, it certainly produces the two values.

Meanwhile, a real-world oracle ω should have the same literal value of $\boldsymbol{\Delta}$ as its counterpart in τ. Then, the real-world sampling $\boldsymbol{R}'^{g}_{\$}\leftarrow\mathcal{R}_0$ (in Appendix F) and $\boldsymbol{r}^g=\left[\boldsymbol{r}^g_{00}\ \boldsymbol{r}^g_{01}\ \boldsymbol{r}^g_{10}\ \boldsymbol{r}^g_{11}\right]^{\mathsf{T}}\in(\mathbb{F}_2^2)^4$ ensure that, for every $i,j\in\mathbb{F}_2$ and every $g\in\mathcal{G}_{\mathsf{and}}(f)$ with $(a,b):=(\mathsf{in}_0(g),\mathsf{in}_1(g))$,

$$\begin{aligned}\boldsymbol{r}^g_{ij}{}^{\mathsf{T}}=\left[r^g_{ijL}\ r^g_{ijR}\right]&=\boldsymbol{R}'^{g}_{\$,ij}\oplus\overline{s_a\oplus x_a}\cdot\boldsymbol{R}'_{1,ij}\oplus\overline{s_b\oplus x_b}\cdot\boldsymbol{R}'_{2,ij}\\&=\left[r^g_L\ r^g_R\right]\oplus\overline{s_a\oplus x_a}\cdot\boldsymbol{R}'_{1,ij}\oplus\overline{s_b\oplus x_b}\cdot\boldsymbol{R}'_{2,ij}.\end{aligned}\tag{39}$$

In particular, (39) holds for $i=s_a$ and $j=s_b$ (i.e., the real-world compatibility between ω and $\widetilde{\boldsymbol{r}}$ in (31) for this g) with probability 2^{-2}, which comes from two uniform coins $(r^g_L,r^g_R)\in\mathbb{F}_2^2$ in ω and independent of the literal values of x_a, x_b, s_a, and s_b. Conditioned on the compatibility so far, the probability

$$\begin{aligned}&\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}[\omega\in\mathsf{comp}_{\mathsf{real}}(\tau)]\\&\quad=\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega\vdash(\mathcal{K}^{\mathsf{R}}_1,F,d,\mathcal{K}^{\mathsf{R}}_2,\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{u}},\widetilde{T},\mathcal{K}^{\mathsf{R}}_3)\ \middle|\ \omega\vdash(f',\widetilde{x})\wedge\omega\vdash(\widetilde{\boldsymbol{r}},\boldsymbol{\Delta})\right]\\&\qquad\cdot\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega\vdash(f',\widetilde{x})\wedge\omega\vdash(\widetilde{\boldsymbol{r}},\boldsymbol{\Delta})\right]\\&\quad=\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega\vdash(\mathcal{K}^{\mathsf{R}}_1,F,d,\mathcal{K}^{\mathsf{R}}_2,\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{u}},\widetilde{T},\mathcal{K}^{\mathsf{R}}_3)\ \middle|\ \omega\vdash(f',\widetilde{x})\wedge\omega\vdash(\widetilde{\boldsymbol{r}},\boldsymbol{\Delta})\right]\\&\qquad\cdot\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega\vdash(f',\widetilde{x})\right]\cdot\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega\vdash\widetilde{\boldsymbol{r}}\right]\cdot\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega\vdash\boldsymbol{\Delta}\right]\\&\quad=\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}\left[\omega\vdash(\mathcal{K}^{\mathsf{R}}_1,F,d,\mathcal{K}^{\mathsf{R}}_2,\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{u}},\widetilde{T},\mathcal{K}^{\mathsf{R}}_3)\ \middle|\ \omega\vdash(f',\widetilde{x})\wedge\omega\vdash(\widetilde{\boldsymbol{r}},\boldsymbol{\Delta})\right]\cdot\frac{1}{2^{2|f|+(\lambda-1)}}.\end{aligned}$$

Conditioned on the compatibility with $(f',\widetilde{x},\widetilde{\boldsymbol{r}},\boldsymbol{\Delta})$, a real-world ω should be compatible with $(\cup_{\ell=1}^3\mathcal{K}^{\mathsf{R}}_\ell,F,\widetilde{\boldsymbol{u}},\widetilde{T})$ and some active labels in $\widetilde{\boldsymbol{x}}$ such that

(i) $\pi^{\pm1}$ maps the fixed permutation queries to the responses in $\cup_{\ell=1}^3\mathcal{K}^{\mathsf{R}}_\ell$.

(ii) For each $i\in\mathcal{W}_{\mathsf{in}}(f)$, it holds that $\boldsymbol{x}_i=\boldsymbol{w}_i\oplus(p_i\oplus x_i)\boldsymbol{\Delta}$.

(iii) For each $i\in\mathcal{W}_{\cup}(f)$ and $j\in[0,\lceil n_i(f)/2\rceil-1]$, it holds that

$$\mathsf{H}(\boldsymbol{x}_i,i\,\|\,j):=\pi((\boldsymbol{0}\,\|\,\boldsymbol{x}_i)\oplus(i\,\|\,j))\oplus\sigma((\boldsymbol{0}\,\|\,\boldsymbol{x}_i)\oplus(i\,\|\,j))=\boldsymbol{u}^j_i=\begin{bmatrix}\mathsf{Half}_1(\boldsymbol{u}^j_i)\\\mathsf{Half}_0(\boldsymbol{u}^j_i)\end{bmatrix}.\tag{40}$$

(iv) For each $g\in\mathcal{G}_{\mathsf{and}}(f)$ with $(a,b,\Gamma,c):=(\mathsf{in}_0(g),\mathsf{in}_1(g),\mathsf{in}_2(g),\mathsf{out}(g))$ and the non-repeating

counters $(\chi_a \,\|\, \rho_a)$, $(\chi_b \,\|\, \rho_b)$, and $(\chi_\Gamma \,\|\, \rho_\Gamma)$ as per some fixed topology order of f, it holds that

$$\begin{aligned}
&\left(\boldsymbol{r}^g_{s_a s_b} \,\middle\|\, \boldsymbol{x}_c \oplus \boldsymbol{R}^g_{s_a s_b} \begin{bmatrix}\boldsymbol{x}_a \\ \boldsymbol{x}_b\end{bmatrix}\right) \\
&\quad = \begin{bmatrix}1 & 0 & 1 \\ 0 & 1 & 1\end{bmatrix} \begin{bmatrix}\mathsf{Half}_{\rho_a}(\mathsf{H}(\boldsymbol{x}_a, a \,\|\, \chi_a)) \\ \mathsf{Half}_{\rho_b}(\mathsf{H}(\boldsymbol{x}_b, b \,\|\, \chi_b)) \\ \mathsf{Half}_{\rho_\Gamma}(\mathsf{H}(\boldsymbol{x}_\Gamma, \Gamma \,\|\, \chi_\Gamma))\end{bmatrix} \oplus \boldsymbol{V}_{s_a s_b} \left(\boldsymbol{z}^g \,\middle\|\, \begin{bmatrix}\boldsymbol{0} \\ \boldsymbol{g}^g\end{bmatrix}\right),
\end{aligned} \tag{41}$$

$$\begin{aligned}
\left(\boldsymbol{z}^g_{\mathsf{bot}} \,\middle\|\, \boldsymbol{g}^g\right) &= \boldsymbol{V}^+_{\mathsf{bot}} \left(\boldsymbol{M}\boldsymbol{h}^g \oplus \left(\boldsymbol{r}^g \,\middle\|\, (\widehat{\boldsymbol{R}}^g \oplus ([\boldsymbol{0}_{4\times 2}\ \boldsymbol{t}^g] \otimes \boldsymbol{I}_2)) \begin{bmatrix}\boldsymbol{w}_a \\ \boldsymbol{w}_b \\ \boldsymbol{\Delta}\end{bmatrix}\right)\right) \\
&= \boldsymbol{V}^+_{\mathsf{bot}} \boldsymbol{M} \begin{bmatrix}
\overline{s_a} & s_a & & & & \\
s_a & \overline{s_a} & & & & \\
& & \overline{s_b} & s_b & & \\
& & s_b & \overline{s_b} & & \\
& & & & \overline{s_a \oplus s_b} & s_a \oplus s_b \\
& & & & s_a \oplus s_b & \overline{s_a \oplus s_b}
\end{bmatrix} \begin{bmatrix}
\mathsf{Half}_{\rho_a}(\mathsf{H}(\boldsymbol{x}_a, a \,\|\, \chi_a)) \\
\mathsf{Half}_{\rho_a}(\mathsf{H}(\boldsymbol{x}_a \oplus \boldsymbol{\Delta}, a \,\|\, \chi_a)) \\
\mathsf{Half}_{\rho_b}(\mathsf{H}(\boldsymbol{x}_b, b \,\|\, \chi_b)) \\
\mathsf{Half}_{\rho_b}(\mathsf{H}(\boldsymbol{x}_b \oplus \boldsymbol{\Delta}, b \,\|\, \chi_b)) \\
\mathsf{Half}_{\rho_\Gamma}(\mathsf{H}(\boldsymbol{x}_\Gamma, \Gamma \,\|\, \chi_\Gamma)) \\
\mathsf{Half}_{\rho_\Gamma}(\mathsf{H}(\boldsymbol{x}_\Gamma \oplus \boldsymbol{\Delta}, \Gamma \,\|\, \chi_\Gamma))
\end{bmatrix} \\
&\quad \oplus \underbrace{\left(\begin{bmatrix}\overline{s_a \oplus x_a} \\ \overline{s_b \oplus x_b} \\ s_a \oplus x_a \oplus s_b \oplus x_b\end{bmatrix} \,\middle\|\, \boldsymbol{V}^+_{\mathsf{bot}} (\widehat{\boldsymbol{R}}^g \oplus ([\boldsymbol{0}_{4\times 2}\ \boldsymbol{t}^g] \otimes \boldsymbol{I}_2)) \begin{bmatrix}\boldsymbol{x}_a \oplus s_a\boldsymbol{\Delta} \\ \boldsymbol{x}_b \oplus s_b\boldsymbol{\Delta} \\ \boldsymbol{\Delta}\end{bmatrix}\right)}_{[\mathsf{Half}_{\rho_a}(\boldsymbol{e}^{\chi_a}_a)\ \mathsf{Half}_{\rho_b}(\boldsymbol{e}^{\chi_b}_b)\ \mathsf{Half}_{\rho_\Gamma}(\boldsymbol{e}^{\chi_\Gamma}_\Gamma)]^{\mathsf{T}} \in \mathbb{F}^3_{2^{\lambda/2+1}}} \\
&\Leftrightarrow \underbrace{\left(\boldsymbol{z}^g_{\mathsf{bot}} \,\middle\|\, \boldsymbol{g}^g\right) \oplus [\mathsf{Half}_{\rho_a}(\boldsymbol{e}^{\chi_a}_a)\ \mathsf{Half}_{\rho_b}(\boldsymbol{e}^{\chi_b}_b)\ \mathsf{Half}_{\rho_\Gamma}(\boldsymbol{e}^{\chi_\Gamma}_\Gamma)]^{\mathsf{T}}}_{[\mathsf{Half}_{\rho_a}(\boldsymbol{y}^{\chi_a}_a)\ \mathsf{Half}_{\rho_b}(\boldsymbol{y}^{\chi_b}_b)\ \mathsf{Half}_{\rho_\Gamma}(\boldsymbol{y}^{\chi_\Gamma}_\Gamma)]^{\mathsf{T}} \in \mathbb{F}^3_{2^{\lambda/2+1}}} \\
&\quad = \begin{bmatrix}\mathsf{Half}_{\rho_a}(\mathsf{H}(\boldsymbol{x}_a, a \,\|\, \chi_a)) \\ \mathsf{Half}_{\rho_b}(\mathsf{H}(\boldsymbol{x}_b, b \,\|\, \chi_b)) \\ \mathsf{Half}_{\rho_\Gamma}(\mathsf{H}(\boldsymbol{x}_\Gamma, \Gamma \,\|\, \chi_\Gamma))\end{bmatrix} \oplus \begin{bmatrix}\mathsf{Half}_{\rho_a}(\mathsf{H}(\boldsymbol{x}_a \oplus \boldsymbol{\Delta}, a \,\|\, \chi_a)) \\ \mathsf{Half}_{\rho_b}(\mathsf{H}(\boldsymbol{x}_b \oplus \boldsymbol{\Delta}, b \,\|\, \chi_b)) \\ \mathsf{Half}_{\rho_\Gamma}(\mathsf{H}(\boldsymbol{x}_\Gamma \oplus \boldsymbol{\Delta}, \Gamma \,\|\, \chi_\Gamma))\end{bmatrix},
\end{aligned} \tag{42}$$

where the bits $x_a, x_b, s_a = \mathsf{lsb}(\boldsymbol{x}_a), s_b = \mathsf{lsb}(\boldsymbol{x}_b)$ are given in τ, and in (42),

$$\begin{aligned}
\boldsymbol{V}^+ &= [\boldsymbol{V}^+_{\mathsf{top}}\ \boldsymbol{V}^+_{\mathsf{bot}}]^{\mathsf{T}} \in (\mathbb{F}_2^{1\times 8})^2 \times (\mathbb{F}_2^{1\times 8})^3, \\
\forall g \in \mathcal{G}_{\mathsf{and}}(f) : \boldsymbol{z}^g &= [\boldsymbol{z}^g_{\mathsf{top}}\ \boldsymbol{z}^g_{\mathsf{bot}}]^{\mathsf{T}} \in \mathbb{F}_2^2 \times \mathbb{F}_2^3, \\
\forall g \in \mathcal{G}_{\mathsf{and}}(f) : \boldsymbol{t}^g &= \begin{bmatrix}g(s_a \oplus x_a, s_b \oplus x_b) \\ g(s_a \oplus x_a, \overline{s_b \oplus x_b}) \\ g(\overline{s_a \oplus x_a}, s_b \oplus x_b) \\ g(\overline{s_a \oplus x_a}, \overline{s_b \oplus x_b})\end{bmatrix},
\end{aligned}$$

and for each $g \in \mathcal{G}_{\mathsf{and}}(f)$, $\widehat{\boldsymbol{R}}^g \in \mathbb{F}_2^{8\times 6}$ is fixed by running $\mathsf{TH.SampleR}$ with random coins (r^g_L, r^g_R) in ω, which have been compatible (according to the condition) with $\widetilde{\boldsymbol{r}}$ as per the (39) for this g and $(i, j) = (s_a, s_b)$.

(v) For each $i \in \{i \in \mathcal{W}_\cup(f) \mid n_i(f) \text{ is odd}\}$, it holds that

$$\mathsf{Half}_1(\mathsf{H}(\boldsymbol{x}_i \oplus \boldsymbol{\Delta}, i \,\|\, \lfloor n_i(f)/2 \rfloor)) = T_i. \tag{43}$$

Conditioned on the compatibility so far, every real-world oracle ω is always compatible with the leftover values in τ, i.e., decoding table d and other active labels in $\widetilde{\boldsymbol{x}}$, which are deterministically computed from XOR combination. The reason is that, for τ ensuring $\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] \neq 0$, these values should be honestly determined by the conditioned values as in the real world. Otherwise, this probability will be zero for an ideal-world oracle, which obtains them from a consistent deterministic

computation as per the conditioned values. As every real-world oracle ω honestly follows the real-world computation, this "leftover" compatibility must hold conditioned on the previous compatibility.

It remains to compute the conditional probabilities for (i) to (v). Consider (iii), (iv), and (v). We claim that every good τ with $\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] \neq 0$ already implies (40) and (41). To see this, we use that condition $\neg\mathsf{bad}_3$ for good transcripts and the extra queries ensure that $\mathcal{K}_3$ fixes the pairs of permutation pre-images and images for hash values

$$\{\mathsf{H}(\boldsymbol{x}_i, i \,\|\, j)\}_{i \in \mathcal{W}_\cup(f), j \in [0, \lceil n_i(f)/2 \rceil - 1]}.$$

These values are consistent with $\widetilde{\boldsymbol{u}}$ fixed in τ as per (40). Otherwise, $\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) = \tau] \neq 0$ cannot be satisfied by τ since $\neg\mathsf{bad}_1 \wedge \neg\mathsf{bad}_3$ for every good transcript implies successful programming in the ideal world so that $\mathsf{H}(\boldsymbol{x}_i, i \,\|\, j) = \boldsymbol{u}_i^j$. As a corollary, (41) holds for every good transcript due to this consistency and the step 14 and 17 of SimIn.

Consider (42) and (43), the leftover parts of (iii), (iv), and (v). If we define from the fixed τ that

$$\mathsf{Half}_1(\boldsymbol{y}_i^{\lfloor n_i(f)/2 \rfloor}) := \mathsf{Half}_1(\boldsymbol{u}_i^{\lfloor n_i(f)/2 \rfloor}) \oplus T_i \in \mathbb{F}_{2^{\lambda/2+1}} \tag{44}$$

for each $i \in \{i \in \mathcal{W}_\cup(f) \mid n_i(f) \text{ is odd}\}$, then we can unify (42) and (43) as:

$$\mathcal{V} := \left\{ \begin{array}{l} i \in \mathcal{W}_\cup(f), j \in [0, \lceil n_i(f)/2 \rceil - 1] : \\ \qquad \boldsymbol{y}_i^j = \mathsf{H}(\boldsymbol{x}_i, i \,\|\, j) \oplus \mathsf{H}(\boldsymbol{x}_i \oplus \boldsymbol{\Delta}, i \,\|\, j) \\ \qquad\quad = \underbrace{\pi((\mathbf{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j))}_{P_{i \,\|\, j,0}} \oplus \underbrace{\pi((\mathbf{0} \,\|\, \boldsymbol{x}_i \oplus \boldsymbol{\Delta}) \oplus (i \,\|\, j))}_{P_{i \,\|\, j,1}} \oplus \sigma(\mathbf{0} \,\|\, \boldsymbol{\Delta}) \end{array} \right\}$$

since (42) and (43) iterate through both halves of all well-defined $\boldsymbol{y}_i^j \in \mathbb{F}^2_{2^{\lambda/2+1}}$ for $i \in \mathcal{W}_\cup(f)$ and $j \in [0, \lceil n_i(f)/2 \rceil - 1]$. Since τ is a good transcript, there are exactly $2N$ pairwise distinct permutation pre-images on the right hand (otherwise, there will be a pair of permutation pre-images leading to (32) in bad_1 or a permutation pre-image leading to (35) in bad_2 given the extra queries). Due to this pairwise distinctness, $\mathcal{V}$ has $2N$ syntactically different variables $\mathcal{P} := \{P_{i \,\|\, j,0}, P_{i \,\|\, j,1}\}_{i \in \mathcal{W}_\cup(f), j \in [0, \lceil n_i(f)/2 \rceil - 1]}$. They can fix the same number of the entries of permutation π in a real-world ω *if and only if their literal values fixed by* τ *are also pairwise distinct.* We note that every good transcript τ does fix *exact one* such assignment of these variables for the following reasons:

- (40) already holds for τ, i.e., for $i \in \mathcal{W}_\cup(f)$ and $j \in [0, \lceil n_i(f)/2 \rceil - 1]$,

$$P_{i \,\|\, j,0} := \pi((\mathbf{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)) = \boldsymbol{u}_i^j \oplus \sigma((\mathbf{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)). \tag{45}$$

  The literal values of $\{P_{i \,\|\, j,0}\}_{i \in \mathcal{W}_\cup(f), j \in [0, \lceil n_i(f)/2 \rceil - 1]}$ are immediate from the responses in $\mathcal{K}_3^{\mathsf{R}}$ given the extra queries and will be pairwise distinct due to the impossible (33) from $\neg\mathsf{bad}_1$.

- For $i \in \mathcal{W}_\cup(f)$ and $j \in [0, \lceil n_i(f)/2 \rceil - 1]$, the literal value of $P_{i \,\|\, j,1}$ is fixed by τ according to $\mathcal{V}$ and (45):

$$P_{i \,\|\, j,1} := \pi((\mathbf{0} \,\|\, \boldsymbol{x}_i \oplus \boldsymbol{\Delta}) \oplus (i \,\|\, j)) = \sigma(\mathbf{0} \,\|\, \boldsymbol{\Delta}) \oplus \boldsymbol{y}_i^j \oplus \boldsymbol{u}_i^j \oplus \sigma((\mathbf{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)). \tag{46}$$

  The literal values of $\{P_{i \,\|\, j,1}\}_{i \in \mathcal{W}_\cup(f), j \in [0, \lceil n_i(f)/2 \rceil - 1]}$ will be pairwise distinct according to the impossible (34) from $\neg\mathsf{bad}_1$.

- The goodness of τ also ensures that there do not exist

$$P' \in \{P_{i \,\|\, j,0}\}_{i \in \mathcal{W}_\cup(f), j \in [0, \lceil n_i(f)/2 \rceil - 1]},$$
$$P'_{\boldsymbol{\Delta}} \in \{P_{i \,\|\, j,1}\}_{i \in \mathcal{W}_\cup(f), j \in [0, \lceil n_i(f)/2 \rceil - 1]}$$

such that $P' = P'_{\boldsymbol{\Delta}}$. Otherwise, this equality and (46) ensure that there exist $(i,j) \in \mathcal{W}_{\cup}(f) \times [0, \lceil n_i(f)/2 \rceil - 1]$ and $(i',j') \in \mathcal{W}_{\cup}(f) \times [0, \lceil n_{i'}(f)/2 \rceil - 1]$ such that

$$\begin{aligned}\pi((\mathbf{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)) &= \pi((\mathbf{0} \,\|\, \boldsymbol{x}_{i'} \oplus \boldsymbol{\Delta}) \oplus (i' \,\|\, j')) \\ &= \sigma(\mathbf{0} \,\|\, \boldsymbol{\Delta}) \oplus \boldsymbol{y}_{i'}^{j'} \oplus \boldsymbol{u}_{i'}^{j'} \oplus \sigma((\mathbf{0} \,\|\, \boldsymbol{x}_{i'}) \oplus (i' \,\|\, j')).\end{aligned} \tag{47}$$

We have $((\mathbf{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j), \pi((\mathbf{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j))) \in \mathcal{K}_3$ according to $\neg\mathsf{bad}_3$ and the extra queries. So, (47) contradicts the impossible (36) from $\neg\mathsf{bad}_2$.

Putting these cases together, we can see that τ yields a value assignment of $\mathcal{P}$, and this assignment fixes exact $2N$ entries of real-world permutation π.

Based on the condition $\neg\mathsf{bad}_1 \wedge \neg\mathsf{bad}_3$ of good transcript τ, N extra queries are non-repeating and the number of non-repeating queries is $q + N = q_\Sigma$. As a result, N responses in $\cup_{\ell=1}^{3}\mathcal{K}_\ell^{\mathsf{R}}$ for the non-repeating extra queries are fixed by the values in $\mathcal{P}$ while the other $q_\Sigma - N = q$ responses are fixed by real-world π (conditioned on the values in $\mathcal{P}$). Based on (i), (ii), (iii), (iv), and (v) together with the "leftover" compatibility, we have in the real world that

$$\begin{aligned}&\Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, F, d, \mathcal{K}_2^{\mathsf{R}}, \widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{u}}, \widetilde{T}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (f', \widetilde{x}) \wedge \omega \vdash (\widetilde{\boldsymbol{r}}, \boldsymbol{\Delta})\right] \\ &\quad = \frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}} \cdot \frac{(2^{\lambda+2} - 2N - (q_\Sigma - N))!}{(2^{\lambda+2})!} \\ &\quad = \frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}} \cdot \frac{1}{(2^{\lambda+2})_{q+2N}} \\ &\Rightarrow \Pr_{\omega \leftarrow \Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)] = \frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}} \cdot \frac{1}{(2^{\lambda+2})_{q+2N}} \cdot \frac{1}{2^{2|f|+(\lambda-1)}}.\end{aligned}$$

Second, in the ideal world, we can use a similar argument to show

$$\begin{aligned}\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\omega \in \mathsf{comp}_{\mathsf{ideal}}(\tau)] &= \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, \mathcal{K}_2^{\mathsf{R}}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (\widehat{f}, \widetilde{\boldsymbol{x}}, \widetilde{x}, \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{r}}, \widetilde{T}, \boldsymbol{\Delta})\right] \\ &\quad \cdot \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\omega \vdash (\widehat{f}, \widetilde{\boldsymbol{x}}, \widetilde{x}, \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{r}}, \widetilde{T}, \boldsymbol{\Delta})] \\ &= \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, \mathcal{K}_2^{\mathsf{R}}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (\widehat{f}, \widetilde{\boldsymbol{x}}, \widetilde{x}, \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{r}}, \widetilde{T}, \boldsymbol{\Delta})\right] \\ &\quad \cdot \frac{1}{2^{|\mathcal{W}_{\mathsf{in}}(f)|\lambda + (3\lambda+8)|f| + (\lambda+2)M + (\lambda-1)}}.\end{aligned}$$

According to the condition $\neg\mathsf{bad}_1 \wedge \neg\mathsf{bad}_3$ and the N extra queries, there are exact $q + N = q_\Sigma$ non-repeating queries. Here, N responses in $\cup_{\ell=1}^{3}\mathcal{K}_\ell^{\mathsf{R}}$ for the non-repeating extra queries are fixed by the conditioned values while the other responses are fixed by $\mathsf{SimP}^{\pm 1}$ for other $q_\Sigma - N = q$ queries.

Let $\boldsymbol{Q}_{i-1}$ denote the list $\boldsymbol{Q}$ (which is maintained in internal state $\mathsf{st}_{\mathsf{sim}}$) when it includes $i - 1 \in [0, q_\Sigma - 1]$ pairs (note that $\boldsymbol{Q}$ finally includes q_Σ pairs given the q_Σ non-repeating queries), and $\mathcal{N} \subseteq [1, q_\Sigma]$ denote the index set of these q queries in q_Σ non-repeating queries to $\mathsf{SimP}^{\pm 1}(\cdot)$ such that $|\mathcal{N}| = q$. We have

$$\begin{aligned}&\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[\omega \vdash (\mathcal{K}_1^{\mathsf{R}}, \mathcal{K}_2^{\mathsf{R}}, \mathcal{K}_3^{\mathsf{R}}) \;\middle|\; \omega \vdash (\widehat{f}, \widetilde{\boldsymbol{x}}, \widetilde{x}, \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{r}}, \widetilde{T}, \boldsymbol{\Delta})\right] \\ &\quad = \prod_{i \in \mathcal{N}} \frac{1}{2^{\lambda+2} - |\boldsymbol{Q}_{i-1}|} = \prod_{i \in \mathcal{N}} \frac{1}{2^{\lambda+2} - (i-1)} \\ &\quad \le \frac{1}{2^{\lambda+2} - N} \times \frac{1}{2^{\lambda+2} - (N+1)} \times \cdots \times \frac{1}{2^{\lambda+2} - (q_\Sigma - 1)} \\ &\quad = \frac{1}{(2^{\lambda+2} - N)_q}, \\ &\Rightarrow \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\omega \in \mathsf{comp}_{\mathsf{ideal}}(\tau)] \\ &\quad \le \frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}} \cdot \frac{1}{(2^{\lambda+2} - N)_q \cdot (2^{\lambda+2})^{3|f|+M}} \cdot \frac{1}{2^{2|f|+(\lambda-1)}}.\end{aligned}$$

As iterating an AND gate increases three counters in either world, we have

$$\begin{aligned} 3\,|f| &= \sum_{i\in\mathcal{W}_\cup(f)} n_i(f) \\ &= \sum_{\substack{i\in\mathcal{W}_\cup(f),\\ n_i(f)\text{ is even}}} 2\cdot\left\lceil\frac{n_i(f)}{2}\right\rceil + \sum_{\substack{i\in\mathcal{W}_\cup(f),\\ n_i(f)\text{ is odd}}} 2\cdot\left(\left\lceil\frac{n_i(f)}{2}\right\rceil-\frac{1}{2}\right) \\ &= 2N-M. \end{aligned} \tag{48}$$

So, we can have $\varepsilon_2 = 0$ since, for every $N \geq 0$ and every $q \geq 0$,

$$\begin{aligned} \frac{\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}[\omega\in\mathsf{comp}_{\mathsf{real}}(\tau)]}{\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\omega\in\mathsf{comp}_{\mathsf{ideal}}(\tau)]} &\geq \frac{(2^{\lambda+2}-N)_q\cdot(2^{\lambda+2})^{2N}}{(2^{\lambda+2})_{q+2N}} \\ &\geq \frac{(2^{\lambda+2}-N)_q\cdot(2^{\lambda+2})_N\cdot(2^{\lambda+2})^{N}}{(2^{\lambda+2})_{q+2N}} \\ &= \frac{(2^{\lambda+2})_{q+N}\cdot(2^{\lambda+2})^{N}}{(2^{\lambda+2})_{q+2N}} = \frac{(2^{\lambda+2})^{N}}{(2^{\lambda+2}-(q+N))_N} \geq 1. \end{aligned}$$

Bounding ε_1. We bound the probabilities of the bad events in the *ideal world*. First, consider bad_1. Note that each active label $\boldsymbol{x}_i$ can be written as the XOR of (i) some active circuit input labels, and/or (ii) some active output labels of the *precedent* AND gates, i.e., for $\mathcal{I}_i \ominus \mathcal{J}_i \neq \varnothing$,

$$\begin{aligned} \boldsymbol{x}_i &= \left(\bigoplus\nolimits_{w\in\mathcal{I}_i\subseteq\mathcal{W}_{\mathsf{in}}(f)}\boldsymbol{x}_w\right)\oplus\left(\bigoplus\nolimits_{w\in\mathcal{J}_i\subseteq\mathcal{W}_{\mathsf{and}}(f)}\boldsymbol{x}_w\right) \\ &= \bigoplus\nolimits_{w\in\mathcal{I}_i\ominus\mathcal{J}_i}\boldsymbol{x}_w \in \mathbb{F}^2_{2^{\lambda/2}}. \end{aligned} \tag{49}$$

For (32), we can use (49) to rewrite it as

$$\left(\mathbf{0}_{2\times 1}\,\middle\|\,\left(\bigoplus\nolimits_{w\in(\mathcal{I}_i\ominus\mathcal{I}_{i'})\ominus(\mathcal{J}_i\ominus\mathcal{J}_{i'})}\boldsymbol{x}_w\right)\right) = (i\,\|\,j)\oplus(i'\,\|\,j') \in \mathbb{F}^2_{2^{\lambda/2+1}}.$$

According to SimIn, each $\boldsymbol{x}_w$, which is sampled in the step 1 or computed in the step 19, has at least $\lambda-1$ random non-LSBs. Therefore, the equality holds with probability at most $2^{-(\lambda-1)}$ for some fixed distinct (i,j) and (i',j'), and, if (i) $\mathcal{I}_i = \mathcal{I}_{i'}$ and $\mathcal{J}_i = \mathcal{J}_{i'}$ or (ii) the right-hand XOR does not give two leading zero bits, this probability is zero for the distinct (i,j) and (i',j'). For (33), the worst case is that both halves of $\boldsymbol{u}_i^j \oplus \boldsymbol{u}_{i'}^{j'} \in \mathbb{F}^2_{2^{\lambda/2+1}}$ respectively serve (in the step 14 of SimIn) as the lower-half masks of two AND gates both with output wires in $\mathcal{Z}(f)$. In this case, each half of this XOR will have at least $1+(\lambda/2-1) = \lambda/2$ random non-LSBs using the lower-half randomness of $(\boldsymbol{r}^g_{s_a s_b}\,\|\,\boldsymbol{x}_c) \in \mathbb{F}^2_{2^{\lambda/2+1}}$ in each of the two AND gates. Note that such $2\cdot\lambda/2 = \lambda$ bits are independent of other (previously fixed) active labels, including $\boldsymbol{x}_i$ and $\boldsymbol{x}_{i'}$. Thus, the equality holds with probability at most $2^{-\lambda} < 2^{-(\lambda-1)}$ for some fixed (i,j) and (i',j').

Similarly, the worst case for (34) is that both halves of $\boldsymbol{u}_i^j \oplus \boldsymbol{u}_{i'}^{j'} \in \mathbb{F}^2_{2^{\lambda/2+1}}$ are respectively the lower-half masks of two AND gates both with output wires in $\mathcal{Z}(f)$. Otherwise, at least one half of $(\boldsymbol{y}_i^j \oplus \boldsymbol{u}_i^j)\oplus(\boldsymbol{y}_{i'}^{j'}\oplus\boldsymbol{u}_{i'}^{j'})$ is uniform for:

(i) The $\boldsymbol{u}_i^j \oplus \boldsymbol{u}_{i'}^{j'}$ masked with the uniform upper half of some $(\boldsymbol{r}^g_{s_a s_b}\,\|\,\boldsymbol{x}_c)$, which cannot be cancelled by $\boldsymbol{y}_i^j$ or $\boldsymbol{y}_{i'}^{j'}$ defined as per (42) and (44), or

(ii) The uniform $\boldsymbol{u}_i^j$ or $\boldsymbol{u}_{i'}^{j'}$ sampled in the step 16 of SimIn and independent of $\boldsymbol{y}_i^j$ or $\boldsymbol{y}_{i'}^{j'}$, or

(iii) The uniform $T_i = \mathsf{Half}_1(\boldsymbol{y}_i^j \oplus \boldsymbol{u}_i^j)$ (resp., $T_{i'} = \mathsf{Half}_1(\boldsymbol{y}_{i'}^{j'} \oplus \boldsymbol{u}_{i'}^{j'})$) when the upper half of the XOR is considered, $n_i(f)$ (resp., $n_{i'}(f)$) is odd, and $j = \lceil n_i(f)/2\rceil - 1$ (resp., $j' = \lceil n_{i'}(f)/2\rceil - 1$).

In this worst case, each half of $\boldsymbol{u}_i^j \oplus \boldsymbol{u}_{i'}^{j'}$ includes at least $1 + (\lambda/2 - 1) = \lambda/2$ uniform bits for the lower-half non-LSBs of some $(\boldsymbol{r}_{s_a s_b}^g \,\|\, \boldsymbol{x}_c)$. These non-LSBs are independent of $\boldsymbol{y}_i^j$ and $\boldsymbol{y}_{i'}^{j'}$ defined as per (42) and (44), or other (previously fixed) active labels, including $\boldsymbol{x}_i$ and $\boldsymbol{x}_{i'}$. Thus, (34) holds with probability $2^{-\lambda}$ for some fixed distinct (i, j) and (i', j').

Taking a union bound over the above cases, we have

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\mathsf{bad}_1] \le \frac{5N(N-1)}{2^{\lambda+1}}. \tag{50}$$

Then, consider bad_2. For (35), it is clear that $(\mathbf{0} \,\|\, \boldsymbol{\Delta}) = \alpha \oplus (\mathbf{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)$, which occurs with probability $2^{-(\lambda-1)}$ according to the randomness of $\boldsymbol{\Delta}$. Let $\boldsymbol{\Delta} = \left[\Delta_L \;\; \Delta_R\right]^{\mathsf{T}} \in \mathbb{F}_{2^{\lambda/2}}^2$. We note that each $\boldsymbol{y}_i^j \in \mathbb{F}_{2^{\lambda/2+1}}^2$ defined from (42) and (44) includes an additive term of the form

$$L_{\xi_1,\xi_2,\xi_3,\xi_4}(\mathbf{0} \,\|\, \boldsymbol{\Delta}) = \begin{bmatrix} 0 \,\|\, \xi_1 \Delta_L \oplus \xi_2 \Delta_R \\ 0 \,\|\, \xi_3 \Delta_L \oplus \xi_4 \Delta_R \end{bmatrix} \in \mathbb{F}_{2^{\lambda/2+1}}^2$$

for some $\xi_1, \xi_2, \xi_3, \xi_4 \in \mathbb{F}_2$, while other additive terms in $\boldsymbol{y}_i^j$ are independent of $\boldsymbol{\Delta}$. Linear orthomorphism σ for every $L_{\xi_1,\xi_2,\xi_3,\xi_4}$ turns (36) into

$$\begin{aligned}\sigma(\mathbf{0} \,\|\, \boldsymbol{\Delta}) \oplus L_{\xi_1,\xi_2,\xi_3,\xi_4}(\mathbf{0} \,\|\, \boldsymbol{\Delta}) = \beta \oplus \boldsymbol{u}_i^j \oplus \sigma((\mathbf{0} \,\|\, \boldsymbol{x}_i) \oplus (i \,\|\, j)) \\ \oplus \text{“other additive terms in } \boldsymbol{y}_i^j\text{”},\end{aligned}$$

which is invertible to compute $(\mathbf{0} \,\|\, \boldsymbol{\Delta})$. This implies that the equality holds with probability $2^{-(\lambda-1)}$ due to the randomness of $\boldsymbol{\Delta}$. Taking a union bound, we have

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\mathsf{bad}_2] \le \frac{2N \cdot (q_1 + q_2 + q_3)}{2^{\lambda-1}}. \tag{51}$$

Finally, consider bad_3. Recall that each $\boldsymbol{x}_w$ has at least $\lambda - 1$ random non-LSBs. Since these bits are independent of $\cup_{\ell=1}^2 \mathcal{K}_\ell$, (37) holds with probability at most $2^{-(\lambda-1)}$ for some fixed $(\alpha, \dots)$ and (i, j). For (38), the mask sampled in the step 13 or the direct sampling in the step 13 or 16 of SimIn ensures that both halves of $\boldsymbol{u}_i^j \in \mathbb{F}_{2^{\lambda/2+1}}^2$ are uniform and independent of $\boldsymbol{x}_i$ or $\cup_{\ell=1}^2 \mathcal{K}_\ell$. It holds with probability $2^{-(\lambda+2)} < 2^{-(\lambda-1)}$ for some fixed $(\dots, \beta)$ and (i, j). So, we can take a union bound to have that

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\mathsf{bad}_3] \le \frac{2N \cdot (q_1 + q_2)}{2^{\lambda-1}}. \tag{52}$$

We have a bound ε_1 from (50), (51), and (52):

$$\begin{aligned}\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) \in \mathcal{T}_{\mathsf{bad}}] &= \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\mathsf{bad}_1 \vee \mathsf{bad}_2 \vee \mathsf{bad}_3] \\ &\le \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\mathsf{bad}_1] + \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\mathsf{bad}_2] + \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\mathsf{bad}_3] \\ &\le \frac{5N(N-1) + 16N(q+N)}{2^{\lambda+1}} \\ &\le \frac{48qs + 189s^2 - 15s}{2^{\lambda+1}} = \varepsilon_1,\end{aligned}$$

where the last inequality comes from (48), $M \le |\mathcal{W}_\cup(f)| \le 3\,|f|$, and $|f| = s$.

The above ε_1, ε_2 and the H-coefficient technique lead to this theorem. □

D.2 Adaptive Security in npRPM

We consider a slightly different three-halves implementation from the original one of [RR21] in that decoding table d is transferred from $\widehat{f}$ to k in $\mathsf{TH.Garble}^{\pi^{\pm 1}(\cdot)}$ and then included in garbled input $\widehat{x}$ in the online phase. This implementation also follows the lower bound in Theorem 8 and satisfies the adaptive security in the npRPM as per the following theorem.

$\underline{\mathsf{SimF}^{\pi^{\pm 1}(\cdot)}(f)}$:

$F := \{(\boldsymbol{g}^g, \boldsymbol{z}^g)\}_{g\in\mathcal{G}_{\mathsf{and}}(f)} \leftarrow (\mathbb{F}^3_{2^{\lambda/2}} \times \mathbb{F}^5_2)^{|f|}$
$\{\boldsymbol{x}_i\}_{i\in\mathcal{W}_{\mathsf{in}}(f)} \leftarrow (\mathbb{F}^2_{2^{\lambda/2}})^{|\mathcal{W}_{\mathsf{in}}(f)|}$
for $i \in \mathcal{W}_{\cup}(f)$ **do** $\mathsf{ctr}_i := 0$
for $g \in \mathcal{G}(f)$ in topology order **do**
  $(a, b, c) := (\mathsf{in}_0(g), \mathsf{in}_1(g), \mathsf{out}(g))$
  if $\mathsf{type}(g) = \texttt{XOR}$ **then** $\boldsymbol{x}_c := \boldsymbol{x}_a \oplus \boldsymbol{x}_b$
  else if $\mathsf{type}(g) = \texttt{AND}$ **then**
    $\Gamma := \mathsf{in}_2(g)$
    $(\chi_a, \chi_b, \chi_\Gamma) := (\lfloor \mathsf{ctr}_a/2 \rfloor, \lfloor \mathsf{ctr}_b/2 \rfloor, \lfloor \mathsf{ctr}_\Gamma/2 \rfloor)$
    $(\rho_a, \rho_b, \rho_\Gamma) := (\mathsf{lsb}(\mathsf{ctr}_a), \mathsf{lsb}(\mathsf{ctr}_b), \mathsf{lsb}(\mathsf{ctr}_\Gamma))$
    $s_a := \mathsf{lsb}(\boldsymbol{x}_a)$, $s_b := \mathsf{lsb}(\boldsymbol{x}_b)$
    $$\begin{bmatrix} \mathsf{Half}_{\rho_a}(\boldsymbol{u}_a^{\chi_a}) \\ \mathsf{Half}_{\rho_b}(\boldsymbol{u}_b^{\chi_b}) \\ \mathsf{Half}_{\rho_\Gamma}(\boldsymbol{u}_\Gamma^{\chi_\Gamma}) \end{bmatrix} := \begin{bmatrix} \mathsf{Half}_{\rho_a}(\pi((\mathbf{0} \,\|\, \boldsymbol{x}_a) \oplus (a \,\|\, \chi_a)) \oplus \sigma((\mathbf{0} \,\|\, \boldsymbol{x}_a) \oplus (a \,\|\, \chi_a))) \\ \mathsf{Half}_{\rho_b}(\pi((\mathbf{0} \,\|\, \boldsymbol{x}_b) \oplus (b \,\|\, \chi_b)) \oplus \sigma((\mathbf{0} \,\|\, \boldsymbol{x}_b) \oplus (b \,\|\, \chi_b))) \\ \mathsf{Half}_{\rho_\Gamma}(\pi((\mathbf{0} \,\|\, \boldsymbol{x}_a \oplus \boldsymbol{x}_b) \oplus (\Gamma \,\|\, \chi_\Gamma)) \oplus \sigma((\mathbf{0} \,\|\, \boldsymbol{x}_a \oplus \boldsymbol{x}_b) \oplus (\Gamma \,\|\, \chi_\Gamma))) \end{bmatrix}$$
    $$(\boldsymbol{r}^g_{s_a s_b} \,\|\, \boldsymbol{m}^g_{s_a s_b}) := \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mathsf{Half}_{\rho_a}(\boldsymbol{u}_a^{\chi_a}) \\ \mathsf{Half}_{\rho_b}(\boldsymbol{u}_b^{\chi_b}) \\ \mathsf{Half}_{\rho_\Gamma}(\boldsymbol{u}_\Gamma^{\chi_\Gamma}) \end{bmatrix} \oplus \boldsymbol{V}_{s_a s_b} \left(\boldsymbol{z}^g \,\middle\|\, \begin{bmatrix} \mathbf{0} \\ \boldsymbol{g}^g \end{bmatrix} \right)$$
    $\boldsymbol{R}^g_{s_a s_b} := \mathsf{TH.DecodeR}(\boldsymbol{r}^g_{s_a s_b}, s_a, s_b)$
    $\boldsymbol{x}_c := \boldsymbol{m}^g_{s_a s_b} \oplus \boldsymbol{R}^g_{s_a s_b} \begin{bmatrix} \boldsymbol{x}_a \\ \boldsymbol{x}_b \end{bmatrix}$
    $\mathsf{ctr}_a := \mathsf{ctr}_a + 1$, $\mathsf{ctr}_b := \mathsf{ctr}_b + 1$, $\mathsf{ctr}_\Gamma := \mathsf{ctr}_\Gamma + 1$
return $\widehat{f} := (f, F)$, $\mathsf{st}_{\mathsf{sim}} := (f, \widetilde{\boldsymbol{x}} := \{\boldsymbol{x}_i\}_{i\in\mathcal{W}(f)}, \widetilde{\boldsymbol{u}} := \{\boldsymbol{u}_i^j\}_{i\in\mathcal{W}_{\cup}(f), j\in[0,\lceil n_i(f)/2\rceil-1]}, \widetilde{\boldsymbol{r}} := \{\boldsymbol{r}^g_{s_a s_b}\}_{g\in\mathcal{G}_{\mathsf{and}}(f),(a,b):=(\mathsf{in}_0(g),\mathsf{in}_1(g))})$

$\underline{\mathsf{SimIn}^{\pi^{\pm 1}(\cdot)}(f(x))}$:

Parse $\mathsf{st}_{\mathsf{sim}} = (f, \widetilde{\boldsymbol{x}} = \{\boldsymbol{x}_i\}_{i\in\mathcal{W}(f)}, \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{r}})$
for $i \in \mathcal{W}_{\mathsf{out}}(f)$ **do** $d_i := f(x)_i \oplus \mathsf{lsb}(\boldsymbol{x}_i)$
return $\widehat{x} := (\{\boldsymbol{x}_i\}_{i\in\mathcal{W}_{\mathsf{in}}(f)}, d), \widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{r}}$.

Figure 8: Our simulator for three-halves in the npRPM.

Theorem 7. *Let* $\mathsf{H}(\boldsymbol{x}, k) = \pi((\mathbf{0} \,\|\, \boldsymbol{x}) \oplus k) \oplus \sigma((\mathbf{0} \,\|\, \boldsymbol{x}) \oplus k)$ *be a tweakable hash function where* $\boldsymbol{x} \in \mathbb{F}^2_{2^{\lambda/2}}, k \in \mathbb{F}^2_{2^{\lambda/2+1}}$, $\pi \in \mathcal{P}_{\lambda+2}$ *is random permutation, and* $\sigma : \mathbb{F}^2_{2^{\lambda/2+1}} \to \mathbb{F}^2_{2^{\lambda/2+1}}$ *is a linear orthomorphism for the function family*

$$\mathcal{L} = \left\{ L_{\xi_1,\xi_2,\xi_3,\xi_4} : \mathbb{F}^2_{2^{\lambda/2+1}} \to \mathbb{F}^2_{2^{\lambda/2+1}} \right\}_{\xi_1,\xi_2,\xi_3,\xi_4\in\mathbb{F}_2}, \quad L_{\xi_1,\xi_2,\xi_3,\xi_4}\left(\begin{bmatrix} x_L \\ x_R \end{bmatrix} \right) = \begin{bmatrix} \xi_1 x_L \oplus \xi_2 x_R \\ \xi_3 x_L \oplus \xi_4 x_R \end{bmatrix}.$$

Then, three-halves (sketched above) is a $(\lambda+2)$*-garbling scheme with* (q, s, ε)*-adaptive security in the npRPM, where* $\varepsilon = (69qs + 234s^2)/2^{\lambda+2}$.

Proof (sketch). The correctness can be proved as [RR21] as postponing decoding table d does not affect correctness. We only need to consider the simulation.

Our simulator $\mathsf{Sim} = (\mathsf{SimF}, \mathsf{SimIn})$ is presented in Figure 8 and is obviously PPT. Then, we prove this theorem using the following three hybrids:

- Hybrid_0. This is the adaptive experiment using simulator Sim.
- Hybrid_1. This is identical to the previous hybrid, except that we replace $\pi^{\pm 1}$ (which can be equivalently emulated on-the-fly as in Figure 3) by an approximation $\widetilde{\pi}^{\pm 1}$ (given in Figure 4). This approximation is the same as random permutation except that, for a new query of the simulator, it returns a fresh random string as response and records this query-response pair. This hybrid is used to simplify probability analysis.
- Hybrid_2. This is the adaptive experiment using three-halves scheme.

According to Lemma 2, every adaptive adversary $\mathcal{A}$ can distinguish Hybrid_0 and Hybrid_1 with advantage at most $(q+N)\cdot N/2^{\lambda+1}$ up to the supposed q queries of $\mathcal{A}$ and the N queries of Sim.

Then, we prove the negligible statistical distance between the transcripts in Hybrid_1 and Hybrid_2, which upper bounds the advantage of adaptive adversary $\mathcal{A}$. To simplify probability analysis, we also use the H-coefficient technique and the same transcript padding as in the proof of Theorem 6. We regard Hybrid_1 (resp., Hybrid_2) and the associated oracle as the ideal (resp., real) world in the H-coefficient technique. The real-world sample space is the same as that in the proof of Theorem 6, but the ideal-world sample space is defined as

$$\Omega_{\mathsf{ideal}} = (\mathbb{F}_2^{3\lambda/2+5})^{|f|} \times (\mathbb{F}_2^{\lambda})^{|\mathcal{W}_{\mathsf{in}}(f)|} \times \underbrace{(\mathbb{F}_2^{\lambda/2+1})^{M}}_{\text{for the sampling of } \widetilde{T}} \times \underbrace{\{0,1\}^*}_{\substack{\text{random tape for} \\ \text{the sampling in } \widetilde{\pi}^{\pm 1}(\cdot)}} \times \underbrace{\mathbb{F}_2^{\lambda-1}}_{\text{dummy } \boldsymbol{\Delta}}.$$

We will consider the same bad transcripts as in the proof of Theorem 6 (where bad_3 is for probability analysis instead of programming, as its counterpart in the proof of Theorem 4) and assume that, without loss of generality, $\mathcal{A}$ only makes non-repeating queries. Let $q_\ell := |\mathcal{K}_\ell|$ for every $\ell \in \{1,2,3\}$ and $q_\Sigma := \sum_{\ell=1}^{3} q_\ell$.

Bounding $1-\varepsilon_2$. We can follow a similar argument (with the difference that the consistency (40) and (41) trivially hold for good transcripts due to the step 12 to 15 in SimF) in the proof of Theorem 6 to show that, for some fixed good transcript τ such that $\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[Y(\omega)=\tau] \neq 0$ and $q_\Sigma = q+N$ (which comes from the condition $\neg\mathsf{bad}_1 \wedge \neg\mathsf{bad}_3$ for good transcripts),

$$\begin{aligned}
&\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)] \\
&\quad = \frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}} \cdot \frac{(2^{\lambda+2}-2N-(q_\Sigma-N))!}{(2^{\lambda+2})!} \cdot \frac{1}{2^{2|f|+(\lambda-1)}} \\
&\quad = \frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}} \cdot \frac{1}{(2^{\lambda+2})_{q+2N}} \cdot \frac{1}{2^{2|f|+(\lambda-1)}}, \\
&\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\omega \in \mathsf{comp}_{\mathsf{ideal}}(\tau)] \\
&\quad \le \frac{1}{(2^{\lambda+2}-N)_q} \cdot \frac{1}{2^{|\mathcal{W}_{\mathsf{in}}(f)|\lambda+(3\lambda/2+5)|f|+(\lambda/2+1)M+(\lambda-1)} \cdot (2^{\lambda+2})^N} \\
&\quad = \frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}} \cdot \frac{1}{(2^{\lambda+2}-N)_q \cdot (2^{\lambda/2+1})^{3|f|+M+2N}} \cdot \frac{1}{2^{2|f|+(\lambda-1)}} \\
&\quad = \frac{1}{(2^\lambda)^{|\mathcal{W}_{\mathsf{in}}(f)|}} \cdot \frac{1}{(2^{\lambda+2}-N)_q \cdot (2^{\lambda+2})^{2N}} \cdot \frac{1}{2^{2|f|+(\lambda-1)}}, \qquad \text{(By (48))} \\
&\Rightarrow \frac{\Pr_{\omega\leftarrow\Omega_{\mathsf{real}}}[\omega \in \mathsf{comp}_{\mathsf{real}}(\tau)]}{\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[\omega \in \mathsf{comp}_{\mathsf{ideal}}(\tau)]} \ge 1.
\end{aligned}$$

That is, we have $\varepsilon_2 = 0$.

Bounding ε_1. First, we consider $\mathsf{bad}_1 \vee \mathsf{bad}_3 = (32) \vee (33) \vee (34) \vee (37) \vee (38)$. We have a similar induction in the proof of Theorem 4 to prove the probability of bad values of some forward queries:

$$\Pr_{\omega\leftarrow\Omega_{\mathsf{ideal}}}[(32)\vee(37)] \le \frac{N(N-1)+2N\cdot(q_1+q_2)}{2^{\lambda+1}}. \tag{53}$$

We consider (33), (34), and (38) conditioned on $\neg((32)\vee(37))$. In each of them, $\boldsymbol{u}_i^j \oplus \sigma((\mathbf{0}\,\|\,\boldsymbol{x}_i)\oplus(i\,\|\,j))$ is the response for query $(\mathbf{0}\,\|\,\boldsymbol{x}_i)\oplus(i\,\|\,j)$ to $\widetilde{\pi}^{\pm 1}(\cdot)$. It follows from this condition that these queries are pairwise distinct so that their responses are taken from uniform $c_1,\ldots,c_{n(\lambda)}$ in $\widetilde{\pi}^{\pm 1}(\cdot)$, where

$\ell(\lambda) = \lambda + 2$, and pairwise independent given the pairwise distinct queries. So, each of them occurs with probability $1/2^{\lambda+2}$ for some fixed quantifier. Taking a union bound over all quantifiers, we have

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[(33) \vee (34) \vee (38) \;\middle|\; \neg((32) \vee (37))\right] \leq \frac{N(N-1) + N \cdot (q_1 + q_2)}{2^{\lambda+2}}. \tag{54}$$

Using (53) and (54), we have

$$\begin{aligned}
\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\mathsf{bad}_1 \vee \mathsf{bad}_3] &= \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[(32) \vee (33) \vee (34) \vee (37) \vee (38)] \\
&= \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[(32) \vee (37)] \\
&\quad + \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}\left[(33) \vee (34) \vee (38) \;\middle|\; \neg((32) \vee (37))\right] \\
&\leq \frac{3N(N-1) + 5N \cdot (q_1 + q_2)}{2^{\lambda+2}}.
\end{aligned} \tag{55}$$

Then, consider bad_2. It is easy to see from the randomness of $\boldsymbol{\Delta}$ that

$$\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\mathsf{bad}_2] \leq \frac{2N \cdot (q_1 + q_2 + q_3)}{2^{\lambda-1}}. \tag{56}$$

We have a bound ε_1 from (55) and (56):

$$\begin{aligned}
\Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[Y(\omega) \in \mathcal{T}_{\mathsf{bad}}] &= \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\mathsf{bad}_1 \vee \mathsf{bad}_2 \vee \mathsf{bad}_3] \\
&\leq \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\mathsf{bad}_1 \vee \mathsf{bad}_3] + \Pr_{\omega \leftarrow \Omega_{\mathsf{ideal}}}[\mathsf{bad}_2] \\
&\leq \frac{3N(N-1) + 21N(q+N)}{2^{\lambda+2}} \\
&\leq \frac{63qs + 216s^2 - 9s}{2^{\lambda+2}} = \varepsilon_1,
\end{aligned}$$

where the last inequality comes from (48), $M \leq |\mathcal{W}_{\cup}(f)| \leq 3\,|f|$, and $|f| = s$.

By using the H-coefficient technique with the above ε_1 and ε_2, we have that any adaptive adversary can distinguish Hybrid_1 and Hybrid_2 with advantage at most $(63qs + 216s^2 - 9s)/2^{\lambda+2}$.

Putting the three hybrids together, we arrive at this theorem. □

E Separation between npRPM and pRPM

We separate the adaptively secure garbling schemes in the two RPMs since the programmability can make a difference in online complexity. In Theorem 8, we prove that the online-complexity lower bound [AIKW13, HW15] of the adaptively secure garbling schemes in the standard model can be extended to the npRPM. However, it is implied by the definition of the adaptively secure garbling schemes in the pRPM that this lower bound can be bypassed if the random permutation is allowed to be programmed by the simulator to embed messages. The proof of Theorem 8 is extended from [HW15], which is based on Yao entropy.

Definition 3 (Yao entropy [Yao82, BSW03, HLR07]). *A distribution $\mathcal{X}$ has Yao entropy at least $\rho \in \mathbb{N}^+$, denoted by $H^{\mathsf{Yao}}(\mathcal{X}) \geq \rho$, if there exists a negligible function ε such that, for every pair of polynomial-size circuits $(C, \mathcal{D})$ where C has output bit-length at most $\rho - 1$, it holds that*

$$\Pr_{x \leftarrow \mathcal{X}}[\mathcal{D}(C(x)) = x] \leq \frac{1}{2} + \varepsilon(\lambda).$$

Theorem 8 (Lower bound of the online complexity in the npRPM). *For every polynomial-size circuit $f : \{0,1\}^{\ell_{\mathsf{in}}} \to \{0,1\}^{\ell_{\mathsf{out}}}$ with a distribution $\mathcal{X}_f$ on $\{0,1\}^{\ell_{\mathsf{in}}}$ such that distribution $f(\mathcal{X}_f)$ on $\{0,1\}^{\ell_{\mathsf{out}}}$ has Yao entropy at least ρ, and every $\ell(\lambda)$-garbling scheme $(\mathsf{poly}(\lambda), \mathsf{poly}'(\lambda), \mathsf{negl}(\lambda))$-adaptively secure in the npRPM, $\mathsf{Encode}^{\pi^{\pm 1}(\cdot)}(k, \cdot)$ outputs at least ρ bits for $(\cdot, k) \leftarrow \mathsf{Garble}^{\pi^{\pm 1}(\cdot)}(f)$.*

Proof. For the sake of contradiction, assume that there exist a polynomial-size circuit f with an input distribution $\mathcal{X}_f$ inducing Yao entropy $H^{\mathsf{Yao}}(f(\mathcal{X}_f)) \geq \rho$, and an adaptively secure garbling scheme in the npRPM, such that the output of $\mathsf{Encode}^{\pi^{\pm 1}(\cdot)}(k,\cdot)$ has at most $\rho-1$ bits for $(\cdot,k) \leftarrow \mathsf{Garble}^{\pi^{\pm 1}(\cdot)}(f)$. We prove that there exists a pair of polynomial-size compressor and decompressor circuits that contradicts the Yao entropy $H^{\mathsf{Yao}}(f(\mathcal{X}_f)) \geq \rho$.

We specify auxiliary input $z = (f, \mathcal{X}_f, \rho)$ and the following computationally unbounded adversary $\mathcal{A}$:

- ${\mathcal{A}_1}^{\pi^{\pm 1}(\cdot)}(z)$ outputs f in z and defines $\mathsf{st}_1 := z$.
- ${\mathcal{A}_2}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_1, \widehat{f})$ samples $x \leftarrow \mathcal{X}_f$ and defines $\mathsf{st}_2 := (\mathsf{st}_1, x)$.
- ${\mathcal{A}_3}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x})$ outputs 1 if $\mathsf{DecEval}^{\pi^{\pm 1}(\cdot)}(\widehat{f}, \widehat{x}) = f(x)$ and $\widehat{x}$ has at most $\rho - 1$ bits, or outputs 0 otherwise. Recall that running $\mathsf{DecEval}^{\pi^{\pm 1}(\cdot)}$ requires a polynomial number of oracle queries to $\pi^{\pm 1}$ as this algorithm is PPT.

The correctness and the contradiction assumption ensure that, for this $(z, \mathcal{A})$,

$$\Pr\left[\begin{array}{l} \pi \leftarrow \mathcal{P}_{\ell(\lambda)}, \pi^{-1} := \mathsf{inv}(\pi), \\ \quad (f, \mathsf{st}_1) \leftarrow {\mathcal{A}_1}^{\pi^{\pm 1}(\cdot)}(z), \\ (\widehat{f}, k) \leftarrow \mathsf{Garble}^{\pi^{\pm 1}(\cdot)}(f), : {\mathcal{A}_3}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x}) = 1 \\ (x, \mathsf{st}_2) \leftarrow {\mathcal{A}_2}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_1, \widehat{f}), \\ \quad \widehat{x} := \mathsf{Encode}^{\pi^{\pm 1}(\cdot)}(k, x) \end{array}\right] = 1. \tag{57}$$

Meanwhile, the adaptive security in the npRPM guarantees that there exists a PPT simulator $\mathsf{Sim} = (\mathsf{SimF}, \mathsf{SimIn})$ such that, for this $(z, \mathcal{A})$,

$$\left| \Pr\left[\begin{array}{l} \pi \leftarrow \mathcal{P}_{\ell(\lambda)}, \pi^{-1} := \mathsf{inv}(\pi), \\ \quad (f, \mathsf{st}_1) \leftarrow {\mathcal{A}_1}^{\pi^{\pm 1}(\cdot)}(z), \\ (\widehat{f}, k) \leftarrow \mathsf{Garble}^{\pi^{\pm 1}(\cdot)}(f), : {\mathcal{A}_3}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x}) = 1 \\ (x, \mathsf{st}_2) \leftarrow {\mathcal{A}_2}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_1, \widehat{f}), \\ \quad \widehat{x} := \mathsf{Encode}^{\pi^{\pm 1}(\cdot)}(k, x) \end{array}\right] - \Pr\left[\begin{array}{l} \pi \leftarrow \mathcal{P}_{\ell(\lambda)}, \pi^{-1} := \mathsf{inv}(\pi), \\ \quad (f, \mathsf{st}_1) \leftarrow {\mathcal{A}_1}^{\pi^{\pm 1}(\cdot)}(z), \\ \quad \widehat{f} \leftarrow \mathsf{SimF}^{\pi^{\pm 1}(\cdot)}(\Phi(f)), : {\mathcal{A}_3}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x}) = 1 \\ (x, \mathsf{st}_2) \leftarrow {\mathcal{A}_2}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_1, \widehat{f}), \\ \quad \widehat{x} \leftarrow \mathsf{SimIn}^{\pi^{\pm 1}(\cdot)}(f(x)) \end{array}\right] \right| \leq \mathsf{negl}(\lambda). \tag{58}$$

Using (57) and (58), we have for the specified $(z, \mathcal{A})$ that

$$\Pr\left[\begin{array}{l} \pi \leftarrow \mathcal{P}_{\ell(\lambda)}, \pi^{-1} := \mathsf{inv}(\pi), \\ \quad (f, \mathsf{st}_1) \leftarrow {\mathcal{A}_1}^{\pi^{\pm 1}(\cdot)}(z), \\ \quad \widehat{f} \leftarrow \mathsf{SimF}^{\pi^{\pm 1}(\cdot)}(\Phi(f)), : {\mathcal{A}_3}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x}) = 1 \\ (x, \mathsf{st}_2) \leftarrow {\mathcal{A}_2}^{\pi^{\pm 1}(\cdot)}(\mathsf{st}_1, \widehat{f}), \\ \quad \widehat{x} \leftarrow \mathsf{SimIn}^{\pi^{\pm 1}(\cdot)}(f(x)) \end{array}\right] \geq 1 - \mathsf{negl}(\lambda), \tag{59}$$

where the probability is taken over the random choice of $\pi \in \mathcal{P}_{\ell(\lambda)}$, the random tape of Sim, and the random tape used by $\mathcal{A}_2$ to sample $x \leftarrow \mathcal{X}_f$.

Recall that both the adversary $\mathcal{A}$ and the PPT simulator Sim can only make a polynomial number of oracle queries. Let $n(\lambda) \in \mathbb{N}^+$ denote the number of oracle queries jointly made by $\mathcal{A}$ and Sim in the simulation (59). For the $n(\lambda)$ oracle queries, permutation π and its inverse π^{-1} can be emulated

Initialize a list $\boldsymbol{Q}_0 = \varnothing$.
Sample uniform[a] $c_1, \ldots, c_{n(\lambda)} \leftarrow \{0,1\}^{\ell(\lambda)}$.
for $i \in [1, n(\lambda)]$ **do**
  if $\mathring{\pi}$ is queried with input $\alpha_i \in \{0,1\}^{\ell(\lambda)}$ **then**
    if $\exists(\alpha_i, \gamma_i) \in \boldsymbol{Q}_{i-1}$ **then** Return γ_i as response
    Return $\gamma_i := c_i$ as response and define $\boldsymbol{Q}_i := \boldsymbol{Q}_{i-1} \cup \{(\alpha_i, \gamma_i)\}$.
  else if $\mathring{\pi}^{-1}$ is queried with input $\beta_i \in \{0,1\}^{\ell(\lambda)}$ **then**
    if $\exists(\gamma_i, \beta_i) \in \boldsymbol{Q}_{i-1}$ **then** Return γ_i as response
    Return $\gamma_i := c_i$ as response and define $\boldsymbol{Q}_i := \boldsymbol{Q}_{i-1} \cup \{(\gamma_i, \beta_i)\}$.

[a] A uniform random tape r_π is used.

Figure 9: The workflow of coarse approximate oracle $\mathring{\pi}^{\pm 1}(\cdot)$ up to $n(\lambda)$ queries.

on-the-fly as in Figure 3 on the uniform distribution of random tape (r_π, r_π^*). This on-the-fly emulation requires exact $n(\lambda) \cdot \ell(\lambda)$ bits of r_π but possibly exponentially large r_π^*. To remove r_π^*, we can consider a coarse approximation $\mathring{\pi}^{\pm 1}$ (see Figure 9), which differs from $\pi^{\pm 1}$ in that it always returns a fresh c_i as the response for a new query. Similar to Lemma 2, it can be proved that replacing $\pi^{\pm 1}$ with $\mathring{\pi}^{\pm 1}$ only incurs additional statistical distance at most $n(\lambda)^2/2^{\ell(\lambda)}$. So, we have for $\mathsf{negl}'(\lambda) := \mathsf{negl}(\lambda) + n(\lambda)^2/2^{\ell(\lambda)}$ and the distribution $\mathring{\mathcal{P}}_{\ell(\lambda)}$ of $\mathring{\pi}^{\pm 1}(\cdot)$ that

$$\Pr\left[\begin{array}{c} \mathring{\pi}^{\pm 1} \leftarrow \mathring{\mathcal{P}}_{\ell(\lambda)}, \\ (f, \mathsf{st}_1) \leftarrow \mathcal{A}_1^{\mathring{\pi}^{\pm 1}(\cdot)}(z), \\ \widehat{f} \leftarrow \mathsf{SimF}^{\mathring{\pi}^{\pm 1}(\cdot)}(\Phi(f)), : \mathcal{A}_3^{\mathring{\pi}^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x}) = 1 \\ (x, \mathsf{st}_2) \leftarrow \mathcal{A}_2^{\mathring{\pi}^{\pm 1}(\cdot)}(\mathsf{st}_1, \widehat{f}), \\ \widehat{x} \leftarrow \mathsf{SimIn}^{\mathring{\pi}^{\pm 1}(\cdot)}(f(x)) \end{array}\right] \geq 1 - \mathsf{negl}'(\lambda), \tag{60}$$

where the probability is taken over the random tape r_π to emulate $\mathring{\pi}^{\pm 1}(\cdot)$, the random tape $(r_{\mathsf{SimF}}, r_{\mathsf{SimIn}})$ of Sim, and the random tape used by $\mathcal{A}_2$ to sample $x \leftarrow \mathcal{X}_f$. We stress that r_π and $(r_{\mathsf{SimF}}, r_{\mathsf{SimIn}})$ are polynomial-size as r_π carries $n(\lambda) \cdot \ell(\lambda)$ uniform bits and the simulator is PPT.

Here, (60) implies the existence of a circuit pair violating $H^{\mathsf{Yao}}(f(\mathcal{X}_f)) \geq \rho$. To see this, let $C_{\ell,\rho,f,r}$ (resp., $\mathcal{D}_{\ell,\rho,f,r}$) denote the compressor (resp., decompressor) circuit, which hardcodes (ℓ, ρ, f) and a random tape $r = (r_\pi, r_{\mathsf{SimF}}, r_{\mathsf{SimIn}})$ with the following functionality:

- Given $f(x) \leftarrow f(\mathcal{X}_f)$, the compressor $C_{\ell,\rho,f,r}$ invokes $\mathsf{SimF}^{\mathring{\pi}^{\pm 1}(\cdot)}(f; r_{\mathsf{SimF}})$ (to maintain internal state $\mathsf{st}_{\mathsf{sim}}$) and computes $\widehat{x} := \mathsf{SimIn}^{\mathring{\pi}^{\pm 1}(\cdot)}(f(x); r_{\mathsf{SimIn}})$. In the end, it outputs $\widehat{x}$ if $\widehat{x}$ has at most $\rho - 1$ bits, or a symbol $\perp$ of $\rho - 1$ bits otherwise. Clearly, the output of $C_{\ell,\rho,f,r}$ has at most $\rho - 1$ bits.
- Given $\widehat{x}'$, the decompressor $\mathcal{D}_{\ell,\rho,f,r}$ recomputes $\widehat{f} := \mathsf{SimF}^{\mathring{\pi}^{\pm 1}(\cdot)}(f; r_{\mathsf{SimF}})$. In the end, it outputs $y := \mathsf{DecEval}^{\mathring{\pi}^{\pm 1}(\cdot)}(\widehat{f}, \widehat{x}')$ if $\widehat{x}' \neq \perp$, or $\perp$ otherwise.
- If an oracle query α to $\mathring{\pi}$ (resp., β to $\mathring{\pi}^{-1}$) is needed, the hardcoded random tape r_π is used to emulate the oracle response according to Figure 9.

Let correct denote the event $\mathcal{D}_{\ell,\rho,f,r}(C_{\ell,\rho,f,r}(f(x))) = f(x)$, and fail denote the event that $\widehat{x}$ has at least ρ bits. These events are taken over the distribution of $(r, f(x))$. We consider the events in the same probability space where $(r, f(x))$ takes the same literal values, to observe that correct $\wedge \neg$fail occurs if and only if $\mathcal{A}_3^{\mathring{\pi}^{\pm 1}(\cdot)}(\mathsf{st}_2, \widehat{f}, \widehat{x}) = 1$ occurs in the adaptive experiment (60), i.e.,

$$\Pr_{r, f(x) \leftarrow f(\mathcal{X}_f)}[\mathsf{correct} \wedge \neg\mathsf{fail}] = (60) \geq 1 - \mathsf{negl}'(\lambda),$$

implying $\Pr_{r, f(x) \leftarrow f(\mathcal{X}_f)}[\mathsf{correct}] \geq \Pr_{r, f(x) \leftarrow f(\mathcal{X}_f)}[\mathsf{correct} \wedge \neg\mathsf{fail}] \geq 1 - \mathsf{negl}'(\lambda)$. So, there exists a random tape r_0, which gives a circuit pair $(C_{\ell,\rho,f,r_0}, \mathcal{D}_{\ell,\rho,f,r_0})$, such that $\mathcal{D}_{\ell,\rho,f,r_0}(C_{\ell,\rho,f,r_0}(f(x))) =$

$f(x)$ occurs with probability at least $1 - \mathsf{negl}'(\lambda)$ for $f(x) \leftarrow f(\mathcal{X}_f)$. Otherwise,

$$\begin{aligned}\Pr_{r, f(x) \leftarrow f(\mathcal{X}_f)}[\mathsf{correct}] &= \sum_r \Pr_{f(x) \leftarrow f(\mathcal{X}_f)}\left[\mathsf{correct} \;\middle|\; r\right] \cdot \Pr_r[r] \\ &< (1 - \mathsf{negl}'(\lambda)) \cdot \sum_r \Pr_r[r] = 1 - \mathsf{negl}'(\lambda),\end{aligned}$$

leading to a contradiction. Note the both C_{ℓ,ρ,f,r_0} and $\mathcal{D}_{\ell,\rho,f,r_0}$ are polynomial-size circuits since the hardcoded random tape r_0 is polynomial-size and the two circuits run in polynomial time given this r_0. Thus, $(C_{\ell,\rho,f,r_0}, \mathcal{D}_{\ell,\rho,f,r_0})$ contradicts the Yao entropy $H^{\mathsf{Yao}}(f(\mathcal{X}_f)) \geq \rho$. □

F Public Parameters of Three-Halves

The following public constants are suggested by three-halves [RR21].

$$\boldsymbol{M} = \left[\begin{array}{cc|cc|cc} 1&0&0&0&1&0\\ 0&0&1&0&1&0\\ \hline 1&0&0&0&0&1\\ 0&0&0&1&0&1\\ \hline 0&1&0&0&0&1\\ 0&0&1&0&0&1\\ \hline 0&1&0&0&1&0\\ 0&0&0&1&1&0 \end{array}\right] \in \mathbb{F}_2^{8\times 6}, \qquad \boldsymbol{K} = \left[\begin{array}{cc|cc|cc|cc} 1&0&1&0&1&0&1&0\\ 0&1&0&1&0&1&0&1\\ \hline 0&0&0&1&1&0&1&1 \end{array}\right] \in \mathbb{F}_2^{3\times 8},$$

$$\boldsymbol{V} = \begin{bmatrix}\boldsymbol{V}_{00}\\ \boldsymbol{V}_{01}\\ \boldsymbol{V}_{10}\\ \boldsymbol{V}_{11}\end{bmatrix} = \left[\begin{array}{cc|ccc} 1&0&0&0&0\\ 0&1&0&0&0\\ \hline 1&0&0&0&1\\ 0&1&0&1&1\\ \hline 1&0&1&0&1\\ 0&1&0&0&1\\ \hline 1&0&1&0&0\\ 0&1&0&1&0 \end{array}\right] \in (\mathbb{F}_2^{2\times 5})^4 \equiv \mathbb{F}_2^{8\times 5},$$

$$\boldsymbol{V}^+ = \left[\begin{array}{cc|cc|cc|cc} 1&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0\\ \hline 1&1&0&0&1&1&0&0\\ 1&1&1&1&0&0&0&0\\ \hline 0&0&0&0&1&0&1&0 \end{array}\right] \in \mathbb{F}_2^{5\times 8},$$

$$\boldsymbol{S}_L = \left[\begin{array}{cc|cc} 1&1&1&0\\ 1&0&0&1 \end{array}\right] \in \mathbb{F}_2^{2\times 4}, \qquad \boldsymbol{S}_R = \left[\begin{array}{cc|cc} 1&0&0&1\\ 0&1&1&1 \end{array}\right] \in \mathbb{F}_2^{2\times 4},$$

$$\boldsymbol{R}'_1 = \begin{bmatrix}0&0\\ 1&1\\ 0&1\\ 1&0\end{bmatrix} \in \mathbb{F}_2^{4\times 2}, \qquad \boldsymbol{R}'_2 = \begin{bmatrix}0&0\\ 1&0\\ 1&1\\ 0&1\end{bmatrix} \in \mathbb{F}_2^{4\times 2},$$

$$\boldsymbol{R}_p = \begin{bmatrix}\boldsymbol{R}_{p,00}\\ \boldsymbol{R}_{p,01}\\ \boldsymbol{R}_{p,10}\\ \boldsymbol{R}_{p,11}\end{bmatrix} = \left[\begin{array}{cc|cc} 0&0&1&0\\ 0&1&0&0\\ \hline 0&0&1&0\\ 0&0&0&0\\ \hline 0&0&0&0\\ 0&1&0&0\\ \hline 0&0&0&0\\ 0&0&0&0 \end{array}\right] \in (\mathbb{F}_2^{2\times 4})^4 \equiv \mathbb{F}_2^{8\times 4}.$$

The distribution $\mathcal{R}_0$ is defined as sampling two uniform bits $\begin{bmatrix} r_L & r_R \end{bmatrix} \leftarrow \mathbb{F}_2^2$ and returning

$$\boldsymbol{R}'_{\$} = \begin{bmatrix} \boldsymbol{R}'_{\$,00} \\ \boldsymbol{R}'_{\$,01} \\ \boldsymbol{R}'_{\$,10} \\ \boldsymbol{R}'_{\$,11} \end{bmatrix} := r_L \cdot \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \oplus r_R \cdot \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \in (\mathbb{F}_2^{1\times 2})^4 \equiv \mathbb{F}_2^{4\times 2}.$$