Conquering Bad Norms in RstOE: Pure-Database Substitution and Early-Defense

Shuping Mao[1], Zhiyu Zhang[2], Peng Wang[2(✉)], Lei Hu[3,4], Luying Li, and Ying Chen[1]

[1] Beijing Electronic Science and Technology Institute, Beijing, China
[2] School of Cryptology, University of Chinese Academy of Sciences, Beijing, China
p-wang@ucas.ac.cn
[3] State Key Laboratory of Information Security, Institute of Information Engineering, CAS
[4] School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China

Abstract. The Recording Standard Oracle with Errors (RstOE) technique is a quantum-security proof technique that provides a structured framework for analyzing adversarial capabilities in quantum settings. For example, it can be applied to prove the quantum security of compressing pseudorandom functions. However, against adaptive quantum chosen-plaintext adversaries, traditional RstOE-based proofs may suffer from the "trivialization of norm" problem. In the RstOE analyses considered in this paper, this issue can be traced to three recurring causes: delayed evaluation of bad events, the presence of unrecorded external variables, and unconstrained independence among intermediate variables. To address this obstacle, we propose two refinements of the RstOE methodology, namely *Pure-Database Substitution* and *Early-Defense*. Pure-Database Substitution algebraically eliminates unrecorded external variables and reformulates collision constraints in terms of internal database records. Building on this substitution, Early-Defense moves the collision check to the point at which a new internal variable is sampled. Because the bad event then depends on this freshly generated quantum randomness, only a negligible fraction of the 2^n superposition branches satisfy the collision constraint. This reduces the amplitude of transitions into the bad subspace and avoids the $O(1)$ norm collapse. We demonstrate the method on TNT as a case study and outline extensions to EDMQ, EDMDQ, LRWQ, and QPMAC.

Keywords: RstOE · Trivialization of Norm · Pure-Database Substitution · Early-Defense · TNT.

1 Introduction

The Recording Standard Oracle with Errors (RstOE) technique [HI19a], which is based on the compressed oracle technique [Zha19], is an important tool for proving the quantum security of symmetric-key constructions. Its main appeal is its structural resemblance to classical lazy-sampling and transcript-based proofs. This resemblance allows one to track the evolution of the oracle's internal database during quantum superposition queries. Using the RstOE technique, Hosoyamada and Iwata proved a tight quantum security bound for the 4-round Feistel structure [HI19a,HI19b], analyzed LRWQ [HI21b],

and established tight bounds for HMAC and NMAC [HI21a] in the quantum random-oracle model. However, recent work identified a serious issue in some RstOE-based analyses, known as the *trivialization of norm* [BCEJ24].

1.1 The trivialization of norm Problem in the RstOE Technique

In the qCPA security proof of the 4-round Feistel structure [HI19a], a "trivialization of norm" issue arises [BCEJ24]: the condition defining a "bad" database (e.g., $x_1 \oplus v_2 = u'_1 \oplus v'_2$) is independent of the freshly sampled random output β from the previous round query. Consequently, regardless of the value of β, the database can be bad, causing the squared norm of the bad state to degrade from the expected $O(q/2^n)$ to $O(1)$. This flaw reduces the bound on the bad probability to a constant rather than a negligible quantity, thereby invalidating the entire security proof. Moreover, [BCEJ24] points out that TNT [MZH+23] and LRWQ [HI21b], whose security was also proven using the RstOE technique, suffer from the same problem.

In fact, the causes of the trivialization of norm problem are not limited to the independence issue mentioned above. We group these causes into three categories: delayed evaluation, external variables, and independence.

1. Delayed Evaluation Problem. In existing RstOE-based proofs, a bad database is typically defined by a collision at the input of the final round. Under this definition, when we evaluate whether the database is "bad," the database records from the previous rounds have already been determined; equivalently, the previous intermediate variables have already been queried and recorded. If these intermediate variables lead to a collision, then all 2^n branches of the final-round output are deterministically classified as bad, because the bad-database definition depends only on the final-round input and not on the output. Consequently, the transition-error bound for the final round collapses to $O(1)$, yielding a single-query transition-error bound of $O(1)$.

The delayed evaluation problem is not specific to a single construction; it is a systemic issue in traditional applications of the RstOE technique. Existing proofs generally define the bad-database condition in terms of collisions at the input of the final round. Consequently, evaluation of the bad-database condition is structurally delayed until after the intermediate quantum variables have already been fixed.

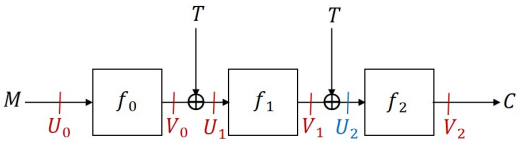


Fig. 1. TNT.

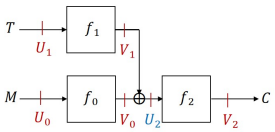


Fig. 2. LRWQ.

We use TNT (Figure 1) and LRWQ (Figure 2), whose security has been analyzed using the RstOE technique, as examples to explain the trivialization of norm problem in detail. Since TNT is affected by all three problems, whereas LRWQ is affected only by delayed evaluation, we first analyze LRWQ.

Consider two distinct tuples $((U_0, V_0), (U_1, V_1))$ and $((U'_0, V'_0), (U'_1, V'_1))$ that result in the same input U_2 for the final round f_2 in LRWQ. During the first query, which generates $((U'_0, V'_0), (U'_1, V'_1))$ and the corresponding U'_2, no collision occurs, so the database remains good. During the second query, after the evaluation of O_{f_0} yields (U_0, V_0), the state transition error is bounded by $O(\sqrt{i^2/2^n})$, as claimed in Lemma 2 of [HI21b], and the database remains good. Subsequently, after the evaluation of O_{f_1} yields (U_1, V_1), the error is similarly bounded by $O(\sqrt{i^2/2^n})$ as per Lemma 3, with the database still classified as good. However, a structural flaw emerges during the evaluation of O_{f_2}. At this stage, whether a collision occurs at U_2 (i.e., $U_2 = U'_2$) has already been entirely determined by historical variables from previous rounds. Consequently, the newly sampled quantum output V_2 does not affect whether the bad-database condition is satisfied, and all 2^n branches of V_2 are classified as bad. Therefore, the transition error of O_{f_2} in Lemma 4 of [HI21b] can no longer be suppressed to a small fraction such as $O(\sqrt{i^2/2^n})$; instead, the corresponding norm bound may become the trivial $O(1)$ bound. Since $O_{\mathsf{LRWQ}} = O^*_{f_0} \cdot O^*_{f_1} \cdot O_{f_2} \cdot O_{f_1} \cdot O_{f_0}$, the single-query transition-error bound for LRWQ becomes $O(1)$. The disconnect between the collision check and the fresh quantum randomness is what causes the trivialization of norm.

2. External Variable Problem. The external variable problem arises when the collision condition involves variables that may be chosen adaptively. Since the local sub-database of the oracle records only internal queries, unrecorded external variables cannot be evaluated solely from the local oracle database, leading to the trivialization of norm problem. For example, in TNT, the collision condition is $U_2 = T \oplus V_1$, where T is an adaptively chosen tweak.

3. Independence Problem. The independence problem refers to situations in which certain internal variables in the collision condition are independent of the output values of previous internal variables, so any value of a previous internal output may potentially lead to a collision. For example, in TNT, the output of the first round, V_0, is generated independently of the subsequent collision condition involving V_1. Consequently, for any input U_0, the sampled value V_0 has an unconstrained chance of causing a collision at V_1. The issue is not independence itself, but rather that the bad event is checked only after the relevant independent randomness has become historical and no longer appears as fresh randomness in the collision test.

To address these difficulties, previous analyses either considered only non-adaptive adversaries [BCEJ24] or abandoned the standard RstOE state-tracking approach, relying instead on macroscopic frameworks to establish security bounds [BCEJ23a]. In this paper, we demonstrate that the RstOE framework can still be used after suitable refinements. To address these three issues, we introduce two core modifications to the RstOE proof methodology: *Pure-Database Substitution* and *Early-Defense*.

1.2 Pure-Database Substitution and Early-Defense

We make targeted improvements to address the external variable problem and the delayed evaluation problem. With these two improvements in place, the independence problem is no longer an obstacle; instead, it becomes the source of the randomness needed for the security bound.

1. Pure-Database Substitution. To resolve the external variable problem, we use algebraic relations in the construction to eliminate unrecorded external variables. This transforms the bad event into a constraint consisting entirely of internal variables that are explicitly recorded in the local databases. It therefore allows the oracle to evaluate the collision condition solely from its recorded database.
2. Early-Defense. To address the delayed evaluation problem, we shift the evaluation point of the bad event to an earlier round, exactly when a specific internal variable is being sampled. By doing so, the collision condition becomes dependent on the newly generated quantum superposition state. At this point, only a small fraction of the 2^n branches satisfy the collision condition, preventing the $O(1)$ collapse. Note that pure-database substitution is a prerequisite for early-defense: only after the collision condition is expressed entirely through recorded internal variables can the oracle evaluate it at the moment when a fresh internal value is sampled.
Resolution of the Independence Problem. By combining pure-database substitution with early-defense, the independence of the intermediate variables no longer leads to norm trivialization. Because the evaluation now occurs exactly when the independent variable is being generated, its independence from the historical database guarantees that its uniform randomness is fully preserved. This independent randomness provides the necessary denominator (2^n), effectively reducing the probability of a bad transition and avoiding the $O(1)$ degradation.

1.3 Pure-Database Substitution and Early-Defense in TNT

In this paper, we use TNT as the main case study to illustrate how pure-database substitution and early-defense are applied. We note that the quantum security analysis of TNT has evolved over several stages: early studies proposed a tight bound of $O(2^{n/3})$ [MZH+23] (which was later shown to be invalid due to the trivialization of norm); subsequently, Bhaumik et al. established a bound of $O(2^{n/4})$ at ASIACRYPT 2023 [BCEJ23b], and later revised it to $O(2^{n/5})$ [BCEJ23a] by using a macroscopic Two-Domain Framework.

Although the security bound of TNT is already known, our motivation for revisiting the proof is methodological. TNT is a suitable case study because its traditional analysis encounters all three identified drivers of norm trivialization: the delayed evaluation problem, the external variable problem, and the independence problem. Since reproving the TNT structure requires the simultaneous application of both pure-database substitution and early-defense, it serves as a good test case for the modified RstOE proof framework. Therefore, we provide a detailed proof for TNT to illustrate how our improved RstOE technique resolves these problems. Following this proof strategy, we outline how the same methodology can be applied to EDMQ, EDMDQ, LRWQ, and QPMAC.
Pure-Database Substitution in TNT (Figure 1). The external variable problem in the TNT structure can be resolved by redefining the bad database. The previous analysis defined the collision condition for the bad database as $U_2 = V_1 \oplus T$, where T is an external variable. However, by exploiting the specific algebraic structure of TNT in which T is injected twice (i.e., $T = V_0 \oplus U_1 = V_1 \oplus U_2$), we can substitute out the

unrecorded external variable T. This transformation reformulates the collision condition as the purely internal constraint $U_2 = V_0 \oplus U_1 \oplus V_1$. Since V_0, U_1, and V_1 are all recorded in the sub-databases D_0 and D_1, the bad-database condition depends only on internal database variables. This transformation yields a pure-database constraint, such as $V_0 \oplus U_1 \oplus V_1 = V_0' \oplus U_1' \oplus V_1'$.

Early-Defense in TNT. Traditional RstOE proofs delay evaluation of the bad database. Previously, the bad-database condition was defined as a collision at the input of the final round (i.e., $U_2 = U_2'$). Under this standard definition, the bad event is evaluated only during the execution of O_{f_2}. However, at this stage, the collision condition $U_2 = U_2'$ has already been predetermined by prior variables and is independent of the freshly sampled quantum output V_2. To circumvent this, early-defense advances the evaluation of the bad event. Using pure-database substitution, we redefine the bad-database condition for TNT: the database is marked bad if and only if $V_0 \oplus U_1 \oplus V_1 = V_0' \oplus U_1' \oplus V_1'$, which is equivalent to $V_1 = V_0' \oplus U_1' \oplus V_1' \oplus V_0 \oplus U_1$. This redefinition is crucial because the collision condition now directly depends on V_1. When the oracle evaluates O_{f_1} and samples V_1, it generates a quantum superposition of 2^n possible values. Because the bad event is evaluated exactly at this moment, the oracle can immediately identify which specific values of V_1 trigger the collision. Among the 2^n possibilities, only a tiny fraction satisfy the collision equation. Consequently, the probability of transitioning into a bad state is naturally diluted by the $1/\sqrt{2^n}$ amplitude factor.

It is worth noting that although V_0 and V_1 are independent, this independence no longer leads to the trivialization of norm under early-defense. The difference lies in the timing of the bad-event evaluation. In traditional proofs, the condition is checked at the final step O_{f_2}. At this point, both V_0 and V_1 have become fixed historical values, and the collision is independent of the freshly generated V_2. In contrast, our early-defense strategy advances the evaluation point to the moment when O_{f_1} samples V_1. At this juncture, V_0 has already been recorded in the database as a known historical constant, whereas V_1 is being generated as a uniform superposition of 2^n possibilities. Because the collision condition anchors this fresh quantum randomness to a fixed historical record ($V_1 = V_0' \oplus U_1' \oplus V_1' \oplus V_0 \oplus U_1$), the oracle can leverage the randomness of V_1 to filter out the colliding branches. This ensures that independence between variables does not induce norm collapse; rather, it provides the guarantee needed for the $1/\sqrt{2^n}$ probability dilution, thereby avoiding the $O(1)$ degradation.

However, although this change of definition avoids the trivialization of norm problem, it incurs a greater combinatorial cost in the security bound. Early-Defense requires the oracle, when sampling a new output V_1, to avoid all possible combinations of historical entries in D_0 and D_1. This global Cartesian product expands the size of the bad subspace from $O(q)$ to $O(q^3)$. Consequently, the single-step transition-error amplitude increases to $O(\sqrt{i^3/2^n})$, yielding the quantum security bound $O(2^{n/5})$ for the TNT structure.

Our contributions are summarized as follows.

1. We provide a systematic diagnosis of the bad-norms problem in RstOE-based adaptive quantum security proofs. In particular, we identify three sources of norm trivialization: delayed evaluation of bad events, collision conditions involving unrecorded

external variables, and independence among intermediate variables after they have become historical records.
2. We introduce two proof-level refinements of the RstOE methodology. Pure-Database Substitution eliminates external variables from collision condition and rewrites them using only recorded internal database values, while Early-Defense evaluates the bad events when a constrained internal value is freshly sampled.
3. We clarify why these refinements avoid the $O(1)$ norm collapse. After pure-database substitution, early-defense makes the bad transition depend on fresh quantum randomness, so the transition amplitude is controlled by the number of historical database configurations satisfying the collision constraint.
4. We give a detailed repaired RstOE proof for TNT. For any quantum adversary making at most q queries, we prove the qPRF bound $O(\sqrt{q^5/2^n})$, corresponding to security up to about $O(2^{n/5})$ queries; by standard switching lemmas, we also obtain the same asymptotic $\mathsf{q\widetilde{PRP}}$ bound for permutation-based TNT.
5. We show that the same proof strategy extends beyond TNT. It yields qPRF bounds $O(\sqrt{q^5/2^n})$ for EDMQ, EDMDQ, and LRWQ, and an $O(\sqrt{q^{2m+1}/2^n})$ qPRF bound for the m-branch QPMAC construction. These results repair RstOE proofs in adaptive settings, although finding tight bounds remains open.

2 Preliminaries

2.1 (Tweakable) Block Ciphers

Block ciphers. A block cipher (or BC for short) $E : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$ is a map with key space $\{0,1\}^k$ and message space $\{0,1\}^n$ such that for every key $K \in \{0,1\}^k$, $M \mapsto E(K,M)$ is a permutation on $\{0,1\}^n$. We let E_K denote the map $M \mapsto E(K,M)$. The inverse of a block cipher E is the map $E^{-1} : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$ defined by $E^{-1}(K,C) = E_K^{-1}(C)$.
Tweakable block ciphers. A tweakable block cipher (or TBC for short) $\tilde{E} : \{0,1\}^k \times \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^n$ is a map with key space $\{0,1\}^k$, tweak space $\{0,1\}^t$, and message space $\{0,1\}^n$ such that for every key $K \in \{0,1\}^k$ and every tweak $T \in \{0,1\}^t$, $M \mapsto \tilde{E}(K,M,T)$ is a permutation on $\{0,1\}^n$. Let $\tilde{E}_K$ be the map $(M,T) \mapsto \tilde{E}(K,M,T)$. The inverse of a TBC $\tilde{E}$ is the map $\tilde{E}_K^{-1} : \{0,1\}^k \times \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^n$ defined by letting $\tilde{E}^{-1}(K,C,T)$ be the unique M such that $\tilde{E}(K,M,T) = C$.

2.2 Quantum Security Advantages

Let H denote the Hadamard operator and $H^{\otimes n}|x\rangle = \frac{1}{2^{n/2}} \sum_y (-1)^{x \cdot y} |y\rangle$. Let the identity operator for an n-qubit quantum system be I_n or I.
Quantum distinguishing advantage. Let O_1, O_2 be two oracles. Let $\mathcal{A}$ be an adversary that queries the corresponding quantum oracles U_{O_i}, defined as $U_{O_i}|x\rangle|y\rangle = |x\rangle|y \oplus O_i(x)\rangle$ for $i = 1, 2$. The quantum distinguishing advantage of $\mathcal{A}$ is defined as $\mathbf{Adv}^{\mathrm{dist}}_{O_1,O_2}(\mathcal{A}) = \left|\Pr\left[\mathcal{A}^{U_{O_1}} \Rightarrow 1\right] - \Pr\left[\mathcal{A}^{U_{O_2}} \Rightarrow 1\right]\right|$.
Quantum PRF advantage. Let $F : \{0,1\}^k \times \{0,1\}^* \to \{0,1\}^n$ be a function. Let $\mathsf{RF} : \{0,1\}^* \to \{0,1\}^n$ be a random function. Let $\mathcal{A}$ be an adversary that queries the

quantum oracle U_F or U_{RF}. The quantum pseudorandom function advantage (or qPRF advantage for short) of $\mathcal{A}$ is defined as $\mathbf{Adv}^{\text{qPRF}}_{F}(\mathcal{A}) = \mathbf{Adv}^{\text{dist}}_{F,\mathsf{RF}}(\mathcal{A})$.

Quantum PRP advantage. Let $E : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$ be a block cipher, and let $\mathsf{RP} : \{0,1\}^n \to \{0,1\}^n$ be a random permutation. Let $\mathcal{A}$ be an adversary querying the quantum oracle U_E or U_{RP}. The quantum pseudorandom permutation advantage (or qPRP advantage for short) of $\mathcal{A}$ is defined as $\mathbf{Adv}^{\text{qPRP}}_{E}(\mathcal{A}) = \mathbf{Adv}^{\text{dist}}_{E,\mathsf{RP}}(\mathcal{A})$.

Quantum $\widetilde{\text{PRP}}$ advantage. Let $\tilde{E} : \{0,1\}^k \times \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^n$ be a tweakable block cipher. Let $\widetilde{\mathsf{RP}} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^n$ be a tweakable random permutation, i.e. $\widetilde{\mathsf{RP}}(\cdot, T)$ is an independent random permutation for each $T \in \{0,1\}^t$. Let $\mathcal{A}$ be an adversary querying the quantum oracle $U_{\tilde{E}}$ or $U_{\widetilde{\mathsf{RP}}}$. The quantum tweakable pseudorandom permutation advantage (or q$\widetilde{\text{PRP}}$ advantage for short) of $\mathcal{A}$ is defined as $\mathbf{Adv}^{\text{q}\widetilde{\text{PRP}}}_{\tilde{E}}(\mathcal{A}) = \mathbf{Adv}^{\text{dist}}_{\tilde{E},\widetilde{\mathsf{RP}}}(\mathcal{A})$.

These notions are related by standard switching lemmas, such as the qPRP/qPRF switching lemma and the q$\widetilde{\text{PRP}}$/qPRF switching lemma.

Lemma 1 (qPRP/qPRF switching lemma [Zha15]). *Let $\mathcal{A}$ be an adversary that makes at most q quantum queries to a random permutation RP or a random function RF from $\{0,1\}^n$ to $\{0,1\}^n$. Then $\mathbf{Adv}^{dist}_{\mathsf{RP},\mathsf{RF}}(\mathcal{A}) \leq O\left(\frac{q^3}{2^n}\right)$.*

Lemma 2 (q$\widetilde{\text{PRP}}$/qPRF switching lemma [GG25,Jha25]). *Let $\mathcal{A}$ be an adversary making at most q quantum queries to a random tweakable permutation $\widetilde{\mathsf{RP}}$ or a random function RF from $\{0,1\}^t \times \{0,1\}^n$ to $\{0,1\}^n$. Then $\mathbf{Adv}^{dist}_{\widetilde{\mathsf{RP}},\mathsf{RF}}(\mathcal{A}) \leq O\left(\frac{q^3}{2^n}\right)$.*

2.3 RstOE Technique

The RstOE technique implements approximate recording of transcripts from quantum queries to random oracles. Classical proofs typically rely on query transcripts, particularly transcripts involving random functions. In the quantum setting, however, recording a query to obtain its transcript is, from the adversary's perspective, equivalent to measuring the query. This action may disturb the quantum system and become detectable by the adversary. Consequently, transcripts of quantum oracle queries cannot be recorded directly. Zhandry's compressed oracle technique addresses this challenge by working in the Fourier domain. When queries are expressed in the Fourier basis, information is written into the oracle's registers rather than copied from the oracle's registers to the adversary's registers. This methodology enables the oracle's database to "record" information about the adversary's queries.

The RstOE technique follows this idea and further refines the proof process. Specifically, it first converts the computational domain to the Fourier domain, performs the recording, and then returns to the computational domain. This approach enables on-the-fly data recording, with one record per query, analogously to the classical setting. This method is useful for bounding the adversary's advantage. The RstOE technique uses the Hadamard transform to convert quantum states to the Fourier domain. This enables the recording of transcripts of queries to quantum random functions, although these records contain some error. We first introduce the necessary definitions.

Definition 1 (Database D). *Let D be a truth table $(b_0||d_0)||\cdots||(b_{2^m-1}||d_{2^m-1})$ with $(n+1)2^m$ bits, where $b_i \in \{0,1\}$ and $d_i \in \{0,1\}^n$ for $0 \le i \le 2^m-1$. The bit b_i is a flag indicating whether d_i is zero. D is a valid database if there is no x such that $d_x \neq 0^n \wedge b_x = 0$. D is an invalid database otherwise. For a valid database D, we write $D(x) = y$ to denote $b_x = 1$ and $d_x = y$, and $D(x) = \perp$ to denote $b_x = 0$. For any $\alpha \neq \perp$ and $x \neq x'$, if two different valid databases $D \neq D'$ satisfy $D(x) = \perp \wedge D'(x) = \alpha$ and $D(x') = D'(x')$, then $D' = D \cup (x,\alpha)$ and $D = D' \backslash (x,\alpha)$. The initial value of b_i is 0, and the initial value of d_i is $\perp$ for $0 \le i \le 2^m-1$.*

Definition 2 (Standard oracle). *Let $x \in \{0,1\}^m, y \in \{0,1\}^n$. Let the database D be a truth table $D = (b_0||d_0)||(b_1||d_1)||\cdots(b_{2^m-1}||d_{2^m-1})$. Then the standard oracle stO is defined by $\mathsf{stO} : |x\rangle|y\rangle|D\rangle \mapsto |x\rangle|y \oplus d_x\rangle|D\rangle$.*

The RstOE transformation on the database register D consists of three steps.

1. Conversion of the computational domain to the Fourier domain, primarily using the Hadamard transform: $\mathsf{IH} = (I_1 \otimes H^{\otimes n})^{\otimes 2^m}$;
2. If $\hat{d_x} \neq 0$, then flip b_x to 1: $U_{\text{toggle}} = (I_1 \otimes |0^n\rangle\langle 0^n| + X \otimes (I_n - |0^n\rangle\langle 0^n|))^{\otimes 2^m}$, where $|\hat{d_x}\rangle = H^{\otimes n}|d_x\rangle$ and $X|b\rangle = |b \oplus 1\rangle$;
3. Revert to the computational domain. If $b_x = 1$, a Hadamard transform is performed on $\hat{d_x}$: $\mathsf{CH} = (CH^{\otimes n})^{\otimes 2^m}$, where $CH^{\otimes n} = |1\rangle\langle 1| \otimes H^{\otimes n} + |0\rangle\langle 0| \otimes I_n$ is the controlled n-qubit Hadamard operator.

Definition 3 (Recording Standard Oracle with Errors [HI19a]). *Let $U_{enc} = \mathsf{CH} \cdot U_{toggle} \cdot \mathsf{IH}$. The recording standard oracle with errors RstOE is a stateful quantum algorithm with a $(n+1)2^m$-qubit state register, defined by $\mathsf{RstOE} = (I \otimes U_{enc}) \cdot \mathsf{stO} \cdot (I \otimes U_{enc}^*)$.*

Moreover, for any quantum algorithm $\mathcal{A}$ and any possible output z, $\Pr[\mathcal{A}^{\mathsf{RstOE}} \Rightarrow z] = \Pr[\mathcal{A}^{\mathsf{StO}} \Rightarrow z] = \Pr[\mathcal{A}^{\mathsf{RF}} \Rightarrow z]$ holds [HI21b].

Let $\mathcal{A}$ be a quantum algorithm, let $|\psi_i\rangle$ be the quantum state before the i-th query, and let $|\psi_{q+1}\rangle$ be the quantum state after all unitary processes. Then the following proposition holds.

Proposition 1 (Proposition 1 in [HI21b]). *For $i \ge 1$, if we measure the oracle-state register of $|\psi_{i+1}\rangle$ and obtain a database D, then D is valid and contains at most i entries.*

Proposition 2 summarizes the basic properties of RstOE. Case 1 corresponds to a repeated query to x, while Case 2 corresponds to the first query to x. The RstOE technique facilitates on-the-fly recording in a quantum setting, albeit with certain errors. These errors are represented by specific states, namely $|\epsilon_1\rangle$ and $|\epsilon_2\rangle$, as indicated in the proposition.

Proposition 2 (Proposition 1 in [HI19a] and [HI19b]). *Let D be a valid database and $D(x) = \perp$. Then, the following properties hold.*

1. $\mathsf{RstOE}|x,y\rangle \otimes |D \cup (x,\alpha)\rangle = |x, y \oplus \alpha\rangle \otimes |D \cup (x,\alpha)\rangle + |\epsilon_1\rangle$, *and the error state*

$$\begin{aligned}|\epsilon_1\rangle =& \frac{1}{\sqrt{2^n}}|x, y \oplus \alpha\rangle(|D\rangle - (\sum_{\gamma\in\{0,1\}^n} \frac{1}{\sqrt{2^n}}|D \cup (x,\gamma)\rangle)) \\ &- \frac{1}{\sqrt{2^n}} \sum_{\gamma} \frac{1}{\sqrt{2^n}}|x, y \oplus \gamma\rangle \otimes (|D \cup (x,\gamma)\rangle - |D_\gamma^{\mathit{invalid}}\rangle) \\ &+ \frac{1}{2^n}|x\rangle|\widehat{0^n}\rangle \otimes (2 \sum_{\delta\in\{0,1\}^n} \frac{1}{\sqrt{2^n}}|D \cup (x,\delta)\rangle - |D\rangle),\end{aligned}$$

   where $|\widehat{0^n}\rangle = H^{\otimes n}|0^n\rangle$ *and* $\left|D_\gamma^{\mathsf{invalid}}\right\rangle$ *is a superposition of invalid databases that depend on* γ *defined by* $|D_\gamma^{\mathsf{invalid}}\rangle = \sum_{\delta\neq 0^n} \frac{(-1)^{\gamma\cdot\delta}}{\sqrt{2^n}}|D \cup (x,\gamma)\rangle$.
2. $\mathsf{RstOE}|x,y\rangle \otimes |D\rangle = \sum_{\alpha\in\{0,1\}^n} \frac{1}{\sqrt{2^n}}|x, y \oplus \alpha\rangle \otimes |D \cup (x,\alpha)\rangle + |\epsilon_2\rangle$, *and the error state*

$$|\epsilon_2\rangle = \frac{1}{\sqrt{2^n}}|x\rangle|\widehat{0^n}\rangle \otimes (|D\rangle - \sum_{\gamma\in\{0,1\}^n} \frac{1}{\sqrt{2^n}}|D \cup (x,\gamma)\rangle).$$

Let $f_0, \ldots, f_l$ be random functions, and let F be a function with black-box access to $f_0, \ldots, f_l$. We denote the quantum oracle of F by O_F. Let D_i be the database of f_i for $i = 0, \ldots, l$, and write $\mathbb{D}_F = (D_0, \ldots, D_l)$ for the combined database of F. Similarly, for a function G with black-box access to random functions $g_0, \ldots, g_s$, we define O_G and $\mathbb{D}_G = (D'_0, \ldots, D'_s)$.

Definition 4 (Good database of Oracle [HI21b]). *Valid databases can be categorized as 'good' or 'bad', analogously to good and bad transcripts in classical cryptography. For two oracles* O_F *and* O_G*, we expect a one-to-one correspondence between the good databases of* O_F *and those of* O_G*. Consequently, if a good database of* O_F *is denoted by* $\mathbb{D}_F$*, then the corresponding good database of* O_G *is written as* $[\mathbb{D}_F]_G$*. Similarly, if a good database of* O_G *is* $\mathbb{D}_G$*, then the corresponding good database of* O_F *is denoted by* $[\mathbb{D}_G]_F$*.*

3 The Quantum Security of TNT

As a tweakable block cipher (TBC), the quantum security of TNT primarily refers to its q$\widetilde{\text{PRP}}$ security bound when the underlying components are modeled as random permutations (denoted by $\mathsf{TNT}[\pi_0, \pi_1, \pi_2]$). In idealized settings, it is common to model the underlying components as random functions (denoted by $\mathsf{TNT}[f_0, f_1, f_2]$). After establishing the qPRF security bound for $\mathsf{TNT}[f_0, f_1, f_2]$, the q$\widetilde{\text{PRP}}$ security bound for $\mathsf{TNT}[\pi_0, \pi_1, \pi_2]$ follows from the qPRP/qPRF switching lemma and the q$\widetilde{\text{PRP}}$/qPRF switching lemma. Below, we provide the quantum security proofs for both structures.

3.1 The qPRF Security for $\mathsf{TNT}[f_0, f_1, f_2]$

TNT built from three independent random functions $f_0, f_1, f_2 : \{0,1\}^n \to \{0,1\}^n$ is defined as $\mathsf{TNT}[f_0, f_1, f_2](M, T) = f_2(T \oplus f_1(T \oplus f_0(M)))$, where $M, T \in \{0,1\}^n$.

To establish the qPRF security of $\mathsf{TNT}[f_0, f_1, f_2]$, we aim to show that it is indistinguishable from a random function. For simplicity, we denote $\mathsf{TNT}[f_0, f_1, f_2]$ by TNT_s; thus

$$\mathsf{TNT}_s(M, T) = f_2(T \oplus f_1(T \oplus f_0(M))).$$

Let $f_2' : \{0,1\}^{3n} \to \{0,1\}^n$ be a random function. We define

$$\mathsf{TNT}_b(M, T) = f_2'(M, T, T \oplus f_1(T \oplus f_0(M))).$$

Consequently, $\mathsf{TNT}_b(M, T)$ is indistinguishable from a random function, since any change in M or T changes the input to f_2', thereby preserving the randomness of $\mathsf{TNT}_b(M, T)$. Thus, proving the qPRF security of $\mathsf{TNT}[f_0, f_1, f_2]$ reduces to proving indistinguishability between TNT_s and TNT_b.

In the subsequent analysis, we use the RstOE technique to establish indistinguishability between TNT_s and TNT_b. To facilitate the proof, we define the intermediate states $M_1 = f_0(M) \oplus T$ and $M_2 = f_1(M_1) \oplus T$, so that

$$\mathsf{TNT}_s(M, T) = f_2(M_2), \mathsf{TNT}_b(M, T) = f_2'(M, T, M_2).$$

Proof idea. At a high level, the distinguishability between TNT_s and the ideal TNT_b is governed by the occurrence of "bad" databases. For valid databases that remain "good," the behaviors of TNT_s and TNT_b coincide, making them indistinguishable to any adversary. Consequently, bounding the adversary's advantage reduces to tracking the evolution of bad databases. The bad databases at the i-th query include those that were already bad before the $(i-1)$-th query and those that transition from good to bad during the i-th query. Initially ($i = 0$), no bad databases are present. Therefore, the main focus is the amplitude of the transition from a good database to a bad one during a single query.

As a basis for later proofs, we present the quantum implementations of TNT_s and TNT_b and characterize the corresponding good and bad databases.

Quantum oracle of $\mathsf{TNT}[f_0, f_1, f_2]$. We define the unitary operators $O_i, i = 0, 1, 2$ and O_2' as follows.

$$\begin{aligned}
O_0 &: |M, T\rangle|y\rangle \mapsto |M, T\rangle|y \oplus M_1\rangle, \\
O_1 &: |M, T\rangle|M_1\rangle|y\rangle \mapsto |M, T\rangle|M_1\rangle|y \oplus M_2\rangle, \\
O_2 &: |M, T\rangle|M_1\rangle|M_2\rangle|y\rangle \mapsto |M, T\rangle|M_1\rangle|M_2\rangle|y \oplus f_2(M_2)\rangle, \\
O_2' &: |M, T\rangle|M_1\rangle|M_2\rangle|y\rangle \mapsto |M, T\rangle|M_1\rangle|M_2\rangle|y \oplus f_2'(M, T, M_2)\rangle.
\end{aligned}$$

Quantum implementations of TNT_s and TNT_b.

1. Start with the input register $|M, T\rangle$.
2. Query O_0 on $|M, T\rangle|0^n\rangle$ to obtain the state ($|Y\rangle$ is the register to which the oracle answer is added)
$$|M, T\rangle|Y\rangle \otimes |M_1\rangle\,. \tag{1}$$
3. Query O_1 on $|M, T\rangle|M_1\rangle|0^n\rangle$ to obtain the state
$$|M, T\rangle|Y\rangle \otimes |M_1\rangle \otimes |M_2\rangle. \tag{2}$$

4. Query O_2 (respectively, O_2') on $|M,T\rangle|M_1\rangle|M_2\rangle|Y\rangle$ to obtain the state

$$|M,T\rangle|Y \oplus \mathsf{TNT}_s(M,T)\rangle \otimes |M_1\rangle \otimes |M_2\rangle. \tag{3}$$
$$(|M,T\rangle|Y \oplus \mathsf{TNT}_b(M,T)\rangle \otimes |M_1\rangle \otimes |M_2\rangle.) \tag{4}$$

5. Uncompute Steps 2–3 to obtain

$$|M,T\rangle|Y \oplus \mathsf{TNT}_s(M,T)\rangle. \tag{5}$$
$$(|M,T\rangle|Y \oplus \mathsf{TNT}_b(M,T)\rangle.) \tag{6}$$

From (1)–(6), we have

$$O_{\mathsf{TNT}_s} = O_0^* \cdot O_1^* \cdot O_2 \cdot O_1 \cdot O_0,$$
$$O_{\mathsf{TNT}_b} = O_0^* \cdot O_1^* \cdot O_2' \cdot O_1 \cdot O_0.$$

Let D_0, D_1, D_2 and D_2' be (valid) databases for f_0, f_1, f_2 and f_2', respectively. If $D_i(x) = y$, we write $(x,y) \in D_i$ for $i = 0,1,2$; if $D_2'(x_1,x_2,x_3) = y$, we write $(x_1||x_2||x_3, y) \in D_2'$. We define the combined database of TNT_s as $\mathbb{D}_s = (D_0, D_1, D_2)$ and that of TNT_b as $\mathbb{D}_b = (D_0, D_1, D_2')$.

Good and bad databases. For the combined database $\mathbb{D}_s = (D_0, D_1, D_2)$ of TNT_s, we define the concept of internal collisions solely based on the sub-databases D_0 and D_1. We say $\mathbb{D}_s$ is bad if and only if there exist entries $(U_0, V_0), (U_0', V_0') \in D_0$ and $(U_1, V_1), (U_1', V_1') \in D_1$ such that:

1. $(U_0, U_1) \neq (U_0', U_1')$
2. $V_0 \oplus U_1 \oplus V_1 = V_0' \oplus U_1' \oplus V_1'$.

Accordingly, $\mathbb{D}_s$ is good if and only if it is not bad. Similarly, for the combined database $\mathbb{D}_b = (D_0, D_1, D_2')$ of TNT_b, we define $\mathbb{D}_b$ as bad under the same condition on D_0 and D_1; otherwise, it is good. In other words, a database is good if it contains no historical combinations in f_0 and f_1 that could lead to a collision at the input of the final round (U_2).

Path-Consistent Good Databases and Bijection. To establish a rigorous one-to-one correspondence between the good databases $\mathbb{D}_s$ and $\mathbb{D}_b$, we must first address the domain disparity between D_2 (mapping $\{0,1\}^n \to \{0,1\}^n$) and D_2' (mapping $\{0,1\}^{3n} \to \{0,1\}^n$). Because the adversary interacts with the macroscopic TNT oracle and does not have direct access to the internal random functions (f_0, f_1, f_2, or f_2'), every entry recorded in the databases is strictly induced by a complete evaluation path originating from an external query (M,T). Consequently, isolated entries (i.e., records not associated with a valid evaluation path) cannot be generated within the RstOE framework for this construction. We define such structurally generated databases as *path-consistent*. In the following bijection, we restrict attention to the path-consistent good components; all non-path-consistent or partially generated branches are conservatively included in the bad/error components and are covered by the norm estimates in Appendix A. When the database is good (i.e., no internal collisions occur in D_0 and D_1), the transition from the evaluation path to the final input $U_2 = V_0 \oplus U_1 \oplus V_1$ is injective. This guarantees that each unique query (M,T) produces a unique U_2.

Therefore, over path-consistent good databases, the mapping between the extended input $(M||T||U_2)$ in D'_2 and the compressed input U_2 in D_2 is a bijection.

We define the mapping for the databases as follows: For any valid path-induced entry $(M||T||U_2, V_2) \in D'_2$, we project it uniquely to $[D'_2]_2 = \{(U_2, V_2)\}$. Conversely, for any $(U_2, V_2) \in D_2$, we expand it via its unique generating path to $[D_2]'_2 = \{(M||T||U_2, V_2)\}$. This mapping allows us to unify the notation for the combined good databases: $[\mathbb{D}_b]_s = (D_0, D_1, [D'_2]_2)$ reconstructs a valid $\mathbb{D}_s$, and $[\mathbb{D}_s]_b = (D_0, D_1, [D_2]'_2)$ reconstructs a valid $\mathbb{D}_b$. This bijection guarantees that the amplitudes of the good quantum states are matched across both oracles.

Indistinguishability proof for TNT_s and TNT_b. Let $\mathcal{A}$ be a quantum adversary that makes at most q quantum queries. Let $|\psi_i\rangle$ and $|\psi'_i\rangle$ denote the full quantum states of $\mathcal{A}$ and the oracle just before the i-th query when $\mathcal{A}$ runs relative to TNT_s and TNT_b, respectively. Let $(M, T), Y$, and Z denote the adversary's query register, response register, and workspace register, respectively.

For each i with $1 \le i \le q+1$, since the oracle databases are classified as good or bad, the corresponding complete quantum states $|\psi_i\rangle$ and $|\psi'_i\rangle$ can similarly be decomposed into good and bad components. We write

$$|\psi'_i\rangle = |\psi_i^{\prime\mathsf{good}}\rangle + |\psi_i^{\prime\mathsf{bad}}\rangle, |\psi_i\rangle = |\psi_i^{\mathsf{good}}\rangle + |\psi_i^{\mathsf{bad}}\rangle.$$

Since there is a one-to-one correspondence between good databases of TNT_b and TNT_s, the representations for $|\psi_i^{\prime\mathsf{good}}\rangle$ and $|\psi_i^{\mathsf{good}}\rangle$ can be written as follows:

$$|\psi_i^{\prime\,\mathrm{good}}\rangle = \sum_{\substack{M,T,Y,Z,(D_0,D_1,D'_2);\\ (D_0,D_1,D'_2)\text{: valid and good}}} a^{(i)}_{M,T,Y,Z,(D_0,D_1,D'_2)} |M,T\rangle|Y\rangle|Z\rangle \otimes \left|D_0, D_1, D'_2\right\rangle, \quad (7)$$

and

$$|\psi_i^{\mathrm{good}}\rangle = \sum_{\substack{M,T,Y,Z,(D_0,D_1,D'_2);\\ (D_0,D_1,D'_2)\text{:valid and good}}} a^{(i)}_{M,T,Y,Z,(D_0,D_1,D'_2)} |M,T\rangle|Y\rangle|Z\rangle \otimes \left|[D_0, D_1, D'_2]_2\right\rangle, \quad (8)$$

where $a^{(i)}_{M,T,Y,Z,(D_0,D_1,D'_2)}$ is a complex number. For each database (D_0, D_1, D'_2) in $|\psi_i^{\prime\mathsf{good}}\rangle$ (respectively, (D_0, D_1, D_2) in $|\psi_i^{\mathsf{good}}\rangle$) with nonzero quantum amplitude, we have $|D_0| \le 2(i-1)$, $|D_1| \le 2(i-1)$, and $|D'_2| \le i-1$ (respectively, $|D_2| \le i-1$).

As outlined in the proof idea, we focus on databases that transition from good to bad during the i-th query. Proposition 3 in [HI21b] gives the relationship between the distinguishing bound of the overall structure and the good-to-bad database transition bound; we instantiate it for TNT_s and TNT_b in Proposition 3.

Proposition 3 (Proposition 3 in [HI21b]). *Suppose that there exist vectors* $|\psi_i^{\prime good}\rangle$, $|\psi_i^{\prime\,bad}\rangle$, $|\psi_i^{good}\rangle$ *and* $|\psi_i^{bad}\rangle$ *that satisfy* $|\psi'_i\rangle = |\psi_i^{\prime\mathsf{good}}\rangle + |\psi_i^{\prime\mathsf{bad}}\rangle$ *and* $|\psi_i\rangle = |\psi_i^{\mathsf{good}}\rangle + |\psi_i^{\mathsf{bad}}\rangle$. $|\psi_i^{\prime\mathsf{good}}\rangle$ *and* $|\psi_i^{\mathsf{good}}\rangle$ *satisfy equation (7) and (8). If*

$$\||\psi_i^{\prime\mathsf{bad}}\rangle\| \le \||\psi_{i-1}^{\prime\mathsf{bad}}\rangle\| + \epsilon_i^{\prime bad}, \||\psi_i^{\mathsf{bad}}\rangle\| \le \||\psi_{i-1}^{\mathsf{bad}}\rangle\| + \epsilon_i^{bad}.$$

Then

$$\mathbf{Adv}^{\mathrm{dist}}_{\mathsf{TNT}_s,\mathsf{TNT}_b}(\mathcal{A}) \le \sum_{1\le i\le q} \epsilon_i^{bad} + \sum_{1\le i\le q} \epsilon_i^{\prime bad}.$$

Resuming our proof, the central task in proving the qPRF security of $\mathsf{TNT}[f_0, f_1, f_2]$ is to establish the indistinguishability of TNT_s and TNT_b. Proposition 3 shows that their distinction is governed primarily by $\epsilon_i^{\prime\mathsf{bad}}$ and $\epsilon_i^{\mathsf{bad}}$. Proposition 4 provides the required norm estimates.

Proposition 4 (Core proposition). *For* $|\psi_i^{\prime\mathsf{bad}}\rangle$ *and* $|\psi_i^{\mathsf{bad}}\rangle$ *of* TNT_b *and* TNT_s*, we have*

$$\||\psi_i^{\prime\mathsf{bad}}\rangle\| \le \||\psi_{i-1}^{\prime\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right), \||\psi_i^{\mathsf{bad}}\rangle\| \le \||\psi_{i-1}^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right).$$

Proof. From (1)–(6), we have $O_{\mathsf{TNT}_s} = O_0^* \cdot O_1^* \cdot O_2 \cdot O_1 \cdot O_0$ and $O_{\mathsf{TNT}_b} = O_0^* \cdot O_1^* \cdot O_2' \cdot O_1 \cdot O_0$. Therefore, when a new query is processed, we bound the bad components of the evaluated quantum states $O_0^* \cdot O_1^* \cdot O_2' \cdot O_1 \cdot O_0|\psi_i'\rangle$ and $O_0^* \cdot O_1^* \cdot O_2 \cdot O_1 \cdot O_0|\psi_i\rangle$. We start with $|\psi_i'\rangle$ (and $|\psi_i\rangle$) and step by step calculate the bad components of $O_0|\psi_i'\rangle$, $O_1 \cdot O_0|\psi_i'\rangle$, $O_2' \cdot O_1 \cdot O_0|\psi_i'\rangle$, $O_1^* \cdot O_2' \cdot O_1 \cdot O_0|\psi_i'\rangle$, and $O_0^* \cdot O_1^* \cdot O_2' \cdot O_1 \cdot O_0|\psi_i'\rangle$. The same analysis applies to $|\psi_i\rangle$. For completeness, we provide the concrete computations in Appendix A.

Theorem 1 (qPRF Security of TNT). *Let* $\mathcal{A}$ *be a quantum adversary that makes at most* q *quantum queries. Then we have*

$$\mathbf{Adv}^{\mathsf{qPRF}}_{\mathsf{TNT}[f_0,f_1,f_2]}(\mathcal{A}) \le O\left(\sqrt{\frac{q^5}{2^n}}\right).$$

Proof. From Proposition 3 and Proposition 4 we have

$$\begin{aligned}\mathbf{Adv}^{\mathsf{qPRF}}_{\mathsf{TNT}[f_0,f_1,f_2]}(\mathcal{A}) &= \mathbf{Adv}^{\mathsf{dist}}_{\mathsf{TNT}_s,\mathsf{TNT}_b}(\mathcal{A}) \\ &\le \sum_{1\le i\le q} O\left(\sqrt{\frac{i^3}{2^n}}\right) + \sum_{1\le i\le q} O\left(\sqrt{\frac{i^3}{2^n}}\right) \le O\left(\sqrt{\frac{q^5}{2^n}}\right).\end{aligned}$$

3.2 The $\widetilde{\mathsf{qPRP}}$ Security for $\mathsf{TNT}[\pi_0, \pi_1, \pi_2]$

Theorem 2 ($\widetilde{\mathsf{qPRP}}$ Security of TNT). *Let* $\mathcal{A}$ *be a quantum algorithm that makes at most* $q, q \le 2^n$ *quantum queries. Then we have*

$$\mathbf{Adv}^{\widetilde{\mathsf{qPRP}}}_{\mathsf{TNT}[\pi_0,\pi_1,\pi_2]}(\mathcal{A}) \le O\left(\sqrt{\frac{q^5}{2^n}}\right).$$

Proof. From Lemma 1, we have

$$\mathbf{Adv}^{\mathsf{dist}}_{\mathsf{TNT}[\pi_0,\pi_1,\pi_2],\mathsf{TNT}_{[\pi_0,\pi_1,f_2]}}(\mathcal{A}) \le O\left(\frac{q^3}{2^n}\right).$$

The same bound applies when replacing π_1 by f_1 and then π_0 by f_0. Moreover,

$$\mathbf{Adv}^{\mathsf{dist}}_{\mathsf{TNT}[f_0,f_1,f_2],\mathsf{RF}}(\mathcal{A}) = \mathbf{Adv}^{\mathsf{qPRF}}_{\mathsf{TNT}[f_0,f_1,f_2]}(\mathcal{A}).$$

Then there exist quantum algorithms $\mathcal{C},\mathcal{D}$ that make at most $O(q)$ quantum queries, such that

$$\begin{aligned}\mathbf{Adv}^{\mathsf{qPRP}}_{\mathsf{TNT}[\pi_0,\pi_1,\pi_2]}(\mathcal{A}) \le& \mathbf{Adv}^{\mathsf{qPRF}}_{\mathsf{TNT}[f_0,f_1,f_2]}(\mathcal{C}) + \mathbf{Adv}^{\mathsf{dist}}_{\mathsf{RF},\mathsf{RP}}(\mathcal{D}) + O\left(\frac{q^3}{2^n}\right)\\ \le& O\left(\sqrt{\frac{q^5}{2^n}}\right) + O\left(\frac{q^3}{2^n}\right) + O\left(\frac{q^3}{2^n}\right) \le O\left(\sqrt{\frac{q^5}{2^n}}\right).\end{aligned}$$

In Appendix B, we present quantum attacks on TNT, including the qPRF attack for $\mathsf{TNT}[f_0,f_1,f_2]$, the quantum cross-road distinguisher on $\mathsf{TNT}[\pi_0,\pi_1,\pi_2]$, and the quantum chosen ciphertext attack on $\mathsf{TNT}[\pi_0,\pi_1,\pi_2]$.

4 The Quantum Security of EDMQ and EDMDQ

EDMQ (Figure 3) and EDMDQ (Figure 4), proposed by Bhaumik et al. [BCEJ23b], were identified as potential PRF-secure candidates requiring three random-function calls. EDMQ is defined as $\mathsf{EDMQ}(M_0,M_1) := f_2(M_0 \oplus f_1(M_1 \oplus f_0(M_0)))$, and EDMDQ is defined as $\mathsf{EDMDQ}(M_0,M_1) := f_2(f_0(M_0) \oplus f_1(M_1 \oplus f_0(M_0)))$. We do not provide a full RstOE proof here; instead, we describe the proof idea and the resulting bound.

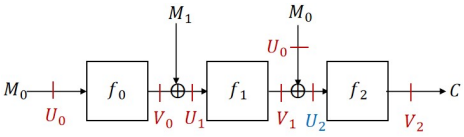


Fig. 3. EDMQ.

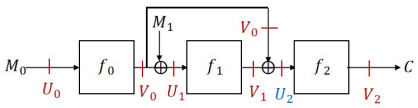


Fig. 4. EDMDQ.

Proposition 5 (qPRF **Security of EDMQ).** *Let $\mathcal{A}$ be a quantum adversary that makes at most q quantum queries. Then we have*

$$\mathbf{Adv}^{\mathsf{qPRF}}_{\mathsf{EDMQ}}(\mathcal{A}) \le O\left(\sqrt{\frac{q^5}{2^n}}\right).$$

Proof idea. As in the TNT analysis, we apply pure-database substitution and early-defense to analyze EDMQ. In traditional proofs, the bad event is triggered by an input collision at the final round, i.e., $U_2 = U_2'$. According to the structure of EDMQ (Figure 3), the input to the final round is $U_2 = V_1 \oplus M_0$. Since the first-round input is $U_0 = M_0$, we can substitute the external plaintext M_0 with the internal variable U_0. This substitution transforms the collision condition into $V_1 \oplus U_0 = V_1' \oplus U_0'$. The constraint is therefore expressed entirely in terms of internal database records.

Under early-defense, we this bad event during the execution of the middle round O_{f_1}. When the oracle samples the quantum superposition for V_1, it must exclude values satisfying the pure-database constraint $V_1 = V_1' \oplus U_0' \oplus U_0$. Since the oracle lacks global knowledge of the adversary's future paths, it must account for all historical combinations on the right-hand side. The degrees of freedom consist of the tuple $V_1' \in D_1$ (at most i choices), $U_0' \in D_0$ (at most i choices), and $U_0 \in D_0$ (at most i choices). The number of values that trigger the bad event is therefore bounded by the Cartesian product $i \times i \times i = i^3$. Consequently, the single-query transition amplitude error during O_{f_1} is bounded by $O(\sqrt{i^3/2^n})$. Summing this error over q quantum queries yields the final total bound of $O(\sqrt{q^5/2^n})$.

Proposition 6 (qPRF **Security of EDMDQ).** *Let $\mathcal{A}$ be a quantum adversary that makes at most q quantum queries. Then we have*

$$\mathbf{Adv}^{\mathsf{qPRF}}_{\mathsf{EDMDQ}[f_0,f_1,f_2]}(\mathcal{A}) \leq O\left(\sqrt{\frac{q^5}{2^n}}\right).$$

Proof idea. The proof strategy for EDMDQ closely follows that for EDMQ. The traditional bad event is defined as the input collision at the final round: $U_2 = U_2'$. In EDMDQ, the structural feedforward defines $U_2 = V_1 \oplus V_0$. Therefore, the collision condition becomes $V_1 \oplus V_0 = V_1' \oplus V_0'$. This yields a pure-database constraint involving only the outputs of the first two rounds.

We trigger the early-defense evaluation at the middle round O_{f_1}. When the oracle samples the quantum superposition for V_1, it must exclude values satisfying the pure-database constraint $V_1 = V_1' \oplus V_0' \oplus V_0$. Since the oracle lacks global knowledge of the adversary's future paths, it must account for all historical combinations on the right-hand side. The degrees of freedom consist of the tuple $V_1' \in D_1$ (at most i choices), $V_0' \in D_0$ (at most i choices), and $V_0 \in D_0$ (at most i choices). The number of values that trigger the bad event is therefore bounded by the Cartesian product $i \times i \times i = i^3$. Consequently, the single-query transition amplitude error during O_{f_1} is bounded by $O(\sqrt{i^3/2^n})$, yielding the total bound $O(\sqrt{q^5/2^n})$.

5 The Quantum Security of LRWQ and QPMAC

The original $\widetilde{\text{qPRP}}$ security bound of $\mathsf{LRWQ}[\pi_0, \pi_1, \pi_2]$ is $O(2^{n/6})$ [HI21b]. Due to the trivialization of norm problem, the updated qPRF bound of $\mathsf{LRWQ}[f_0, f_1, f_2]$ in the erratum of [BCEJ23a] is $O(2^{n/5})$, obtained using macroscopic frameworks. By applying our pure-database substitution and early-defense methods, we reprove the security

within the RstOE framework. We show that the qPRF bound for $\mathsf{LRWQ}[f_0, f_1, f_2]$ is $O(2^{n/5})$, and that the $\widetilde{\mathsf{qPRP}}$ bound for $\mathsf{LRWQ}[\pi_0, \pi_1, \pi_2]$ is also $O(2^{n/5})$. We prove the qPRF bound; the $\widetilde{\mathsf{qPRP}}$ bound then follows from the $\widetilde{\mathsf{qPRP}}$/qPRF switching lemma (Lemma 2). Furthermore, we extend this analysis to QPMAC [MGW+24] (Figure 5), a parallel MAC construction that generalizes the LRWQ architecture to m branches.

Proposition 7 (qPRF **Security of LRWQ).** *Let $\mathcal{A}$ be a quantum adversary that makes at most q quantum queries. Then we have*

$$\mathbf{Adv}^{\mathsf{qPRF}}_{\mathsf{LRWQ}[f_0,f_1,f_2]}(\mathcal{A}) \leq O\left(\sqrt{\frac{q^5}{2^n}}\right).$$

Proof idea. The proof strategy for LRWQ (Figure 2) follows the same approach as for EDMQ and EDMDQ. In traditional evaluations, the bad event is triggered by an input collision at the final round, i.e., $U_2 = U_2'$. According to the structure of LRWQ, the input to the final round is determined by the XOR sum of the previous rounds' outputs: $U_2 = V_1 \oplus V_0$. Therefore, the collision condition translates to $V_1 \oplus V_0 = V_1' \oplus V_0'$. This is a pure-database constraint involving only the internal outputs.

We trigger the early-defense evaluation during the execution of the local oracle O_{f_1}. When the oracle samples the quantum superposition for V_1, it must exclude values satisfying the pure-database constraint $V_1 = V_1' \oplus V_0' \oplus V_0$. Since the oracle lacks global knowledge of the adversary's future paths, it must account for all historical combinations on the right-hand side. The degrees of freedom consist of the tuple $V_1' \in D_1$ (at most i choices), $V_0' \in D_0$ (at most i choices), and $V_0 \in D_0$ (at most i choices). The number of values that trigger the bad event is therefore bounded by the Cartesian product $i \times i \times i = i^3$. Consequently, the single-query transition amplitude error during O_{f_1} is bounded by $O(\sqrt{i^3/2^n})$, yielding the total bound $O(\sqrt{q^5/2^n})$.

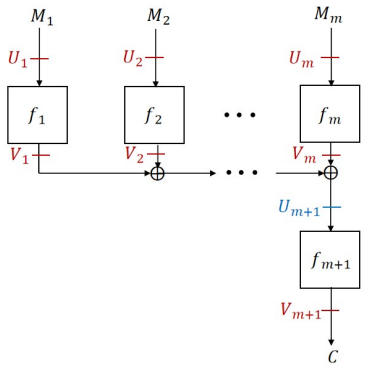


Fig. 5. QPMAC.

Proposition 8 (qPRF Security of QPMAC). *Let $\mathcal{A}$ be a quantum adversary that makes at most q quantum queries. For a QPMAC construction with m parallel branches, we have*

$$\mathbf{Adv}^{\mathsf{qPRF}}_{\mathsf{QPMAC}}(\mathcal{A}) \leq O\left(\sqrt{\frac{q^{2m+1}}{2^n}}\right).$$

Proof idea. QPMAC can be viewed as an m-branch generalization of LRWQ. In traditional analyses, the bad event is triggered by an input collision at the final round f_{m+1}, meaning $U_{m+1} = U'_{m+1}$. Based on the structure of QPMAC, the input to the final round is the XOR sum of all m parallel intermediate outputs: $U_{m+1} = \bigoplus_{k=1}^{m} V_k$. Therefore, the collision condition translates into a generalized pure-database constraint: $\bigoplus_{k=1}^{m} V_k = \bigoplus_{k=1}^{m} V'_k$. This equation is free of external variables.

We activate early-defense during the evaluation of the final parallel branch O_{f_m}. When the oracle samples the quantum superposition for V_m, it must exclude values satisfying the pure-database constraint $V_m = \bigoplus_{k=1}^{m-1} V_k \oplus \bigoplus_{k=1}^{m} V'_k$. Since the oracle lacks global knowledge of the adversary's future paths, it must account for all historical combinations on the right-hand side. The degrees of freedom consist of the tuples $(V_1, \ldots, V_{m-1}) \in D_1 \times \cdots \times D_{m-1}$ (which yields i^{m-1} choices) and the tuples $(V'_1, \ldots, V'_m) \in D_1 \times \cdots \times D_m$ (which yields i^m choices). This Cartesian product yields at most $i^{m-1} \times i^m = i^{2m-1}$ configurations. Consequently, the amplitude attenuation restricts the single-query transition error to $O(\sqrt{i^{2m-1}/2^n})$. Since $\sum_{j=1}^{q} j^{(2m-1)/2} \approx q^{(2m+1)/2}$, the final bound of QPMAC is $O(\sqrt{q^{2m+1}/2^n})$.

6 Conclusions and Discussions

The RstOE framework is a useful technique for analyzing the quantum security of symmetric-key constructions. In this paper, we identified three problems that lead to norm trivialization: the delayed evaluation problem, the external variable problem, and the independence problem. To address these issues, we introduced two refinements: Pure-Database Substitution and Early-Defense. We demonstrate that the RstOE technique can be repaired in adaptive settings by combining pure-database substitution with early-defense. Based on these modifications, we prove the qPRF bound of TNT using the RstOE technique and outline the proof ideas for EDMQ, EDMDQ, LRWQ, and QPMAC.

We present a repaired RstOE method: transform the collision condition into a constraint involving only internal variables recorded in the database, and then advance the collision evaluation to the point at which a new internal variable is sampled. This method avoids the trivialization of norm problem, but it inevitably weakens the security bound.

Our approach also has limitations. For constructions such as the 4-round Feistel, some rounds involve unrecorded external variables that cannot be eliminated through pure-database substitution, making the early-defense strategy inapplicable. In particular, applying early-defense requires that the relevant collision condition be expressible entirely in terms of pure-database records.

Several directions remain for future work. First, establishing the adaptive quantum security of the 4-round Feistel structure remains a challenging open problem. Second, determining tight security bounds for the constructions studied in this paper remains open. To overcome the bad-norms problem in the original RstOE framework, our analysis modifies the proof strategy by replacing external collision conditions with pure-database constraints and by evaluating these conditions earlier, when the relevant internal variables are sampled. This repair avoids norm trivialization, but the additional restrictions introduced by the modified strategy suggest that the resulting bounds, such as $O(2^{n/5})$, should not be expected to be tight. Finding the optimal bounds, and proving their tightness either by matching quantum distinguishing attacks or by sharper upper-bound arguments, remains an important open problem. Finally, our analysis focuses on qCPA security. Extending this approach to the qCCA model, which involves inverse oracles (decryption queries), is another significant direction for future study.

References

Amb07. Andris Ambainis. Quantum walk algorithm for element distinctness. *SIAM J. Comput.*, 37(1):210–239, 2007. 29

BCEJ23a. Ritam Bhaumik, Benoît Cogliati, Jordan Ethan, and Ashwin Jha. On quantum secure compressing pseudorandom functions. *IACR Cryptol. ePrint Arch.*, 2023:207, 2023. 3, 4, 15

BCEJ23b. Ritam Bhaumik, Benoît Cogliati, Jordan Ethan, and Ashwin Jha. On quantum secure compressing pseudorandom functions. In *Advances in Cryptology - ASIACRYPT 2023*, volume 14440, pages 34–66, 2023. 4, 14

BCEJ24. Ritam Bhaumik, Benoît Cogliati, Jordan Ethan, and Ashwin Jha. Mind the bad norms - revisiting compressed oracle-based quantum indistinguishability proofs. In *Advances in Cryptology - ASIACRYPT 2024*, pages 215–247, 2024. 2, 3

BHMT02. Gilles Brassard, Peter Hoyer, Michele Mosca, and Alain Tapp. Quantum amplitude amplification and estimation. *Contemporary Mathematics*, 305:53–74, 2002. 28, 29

GG25. Xiaoning Guo and Chun Guo. Tight bounds for quantum indistinguishability of tweakable random permutations from random functions. *Journal of Cryptologic Research*, 12(2):297–309, 2025. 7

HI19a. Akinori Hosoyamada and Tetsu Iwata. 4-round luby-rackoff construction is a qPRP. In *Advances in Cryptology - ASIACRYPT 2019*, volume 11921, pages 145–174, 2019. 1, 2, 8

HI19b. Akinori Hosoyamada and Tetsu Iwata. 4-round luby-rackoff construction is a qPRP: Tight quantum security bound. *IACR Cryptol. ePrint Arch.*, page 243, 2019. 1, 8, 29, 30

HI21a. Akinori Hosoyamada and Tetsu Iwata. On tight quantum security of HMAC and NMAC in the quantum random oracle model. In *Advances in Cryptology - CRYPTO 2021*, volume 12825, pages 585–615, 2021. 2

HI21b. Akinori Hosoyamada and Tetsu Iwata. Provably quantum-secure tweakable block ciphers. *IACR Trans. Symmetric Cryptol.*, 2021(1):337–377, 2021. 1, 2, 3, 8, 9, 12, 15

Jha25. Ashwin Jha. On TRP-RF switch in the quantum query model. *IACR Communications in Cryptology*, 2(2), 2025. 7

JKNS23. Ashwin Jha, Mustafa Khairallah, Mridul Nandi, and Abishanka Saha. Tight security of TNT and beyond: Attacks, proofs and possibilities for the cascaded LRW paradigm. Cryptology ePrint Archive, Paper 2023/1272, 2023. 32

MGW+24. Shuping Mao, Tingting Guo, Peng Wang, Ruozhou Xu, Yuchao Chen, and Lei Hu. A quantum-secure partial parallel MAC QPCBC. *Des. Codes Cryptogr.*, 92(12):4453–4486, 2024. 16

MZH+23. Shuping Mao, Zhiyu Zhang, Lei Hu, Luying Li, and Peng Wang. Quantum security of TNT. *IACR Cryptol. ePrint Arch.*, 2023:1280, 2023. 2, 4

Zha15. Mark Zhandry. A note on the quantum collision and set equality problems. *Quantum Inf. Comput.*, 15(7&8):557–567, 2015. 7

Zha19. Mark Zhandry. How to record quantum queries, and applications to quantum indifferentiability. In *Advances in Cryptology - CRYPTO 2019*, volume 11693, pages 239–268, 2019. 1

A Proof of Proposition 4

We first recall the implementations of O_{TNT_s} and O_{TNT_b}. From (1)–(6), we have

$$O_{\mathsf{TNT}_s} = O_0^* \cdot O_1^* \cdot O_2 \cdot O_1 \cdot O_0,$$
$$O_{\mathsf{TNT}_b} = O_0^* \cdot O_1^* \cdot O_2' \cdot O_1 \cdot O_0.$$

For $|\psi_i'\rangle = |\psi_i'^{\mathsf{good}}\rangle + |\psi_i'^{\mathsf{bad}}\rangle$ and $|\psi_i\rangle = |\psi_i^{\mathsf{good}}\rangle + |\psi_i^{\mathsf{bad}}\rangle$, where $|\psi_i'^{\mathsf{good}}\rangle$ and $|\psi_i^{\mathsf{good}}\rangle$ satisfy equation (7) and (8), we first consider the action of O_0.

From (1), querying O_0 on $|M,T\rangle|0^n\rangle$ yields $|M,T\rangle \otimes |M_1\rangle$. After the operation of O_0, the quantum state is partitioned into good and bad components, characterized by

$$O_0|\psi_i'\rangle = |\psi_i'^{\mathsf{good},1}\rangle + |\psi_i'^{\mathsf{bad},1}\rangle, O_0|\psi_i\rangle = |\psi_i^{\mathsf{good},1}\rangle + |\psi_i^{\mathsf{bad},1}\rangle.$$

Thus from (1), (7) and (8) there exists a complex number $a^{(i),1}_{M,T,Y,Z,(D_0,D_1,D_2')}$ such that

$$|\psi_i'^{\mathsf{good},1}\rangle = \sum_{\substack{M,T,Y,Z,(D_0,D_1,D_2');\\ (D_0,D_1,D_2')\text{: valid and good}\\ D_0(M)\neq\perp}} a^{(i),1}_{M,T,Y,Z,(D_0,D_1,D_2')}|M,T\rangle|Y\rangle|Z\rangle \otimes \left|D_0,D_1,D_2'\right\rangle \otimes |M_1\rangle. \tag{9}$$

and

$$|\psi_i^{\mathsf{good},1}\rangle = \sum_{\substack{M,T,Y,Z,(D_0,D_1,D_2');\\ (D_0,D_1,D_2')\text{: valid and good}\\ D_0(M)\neq\perp}} a^{(i),1}_{M,T,Y,Z,(D_0,D_1,D_2')}|M,T\rangle|Y\rangle|Z\rangle \otimes \left|\left[D_0,D_1,D_2'\right]_2\right\rangle \otimes |M_1\rangle. \tag{10}$$

We now bound the bad component.

Lemma 3 (Action of O_0). *For $|\psi_i'^{\mathsf{bad},1}\rangle$ and $|\psi_i^{\mathsf{bad},1}\rangle$, we have*

$$\||\psi_i'^{\mathsf{bad},1}\rangle\| \le \||\psi_i'^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right), \||\psi_i^{\mathsf{bad},1}\rangle\| \le \||\psi_i^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right).$$

Proof. Let $\mathbf{\Pi}_{\mathsf{valid}}$ denote the projection onto the space spanned by the vectors that correspond to valid databases. From equation (7) we have

$$\begin{aligned}&\mathbf{\Pi}_{\mathsf{valid}}O_0|\psi_i'^{\mathsf{good}}\rangle\\ =&\mathbf{\Pi}_{\mathsf{valid}}O_0\sum_{\substack{M,T,Y,Z,(D_0,D_1,D_2');\\(D_0,D_1,D_2'):\text{ valid and good}}} a^i_{M,T,Y,Z,(D_0,D_1,D_2')}|M,T\rangle|Y\rangle|Z\rangle\otimes\left|D_0,D_1,D_2'\right\rangle.\end{aligned}$$

The decomposition according to $D_0(M)$ depends on whether M has already been queried, leading to two cases: $D_0(M)=\perp$ and $D_0(M)\neq\perp$. In the latter case, we represent $D_0(M)\neq\perp$ by starting from a database with $D_0(M)=\perp$ and then adding $D_0\cup(M,\alpha)$. Consequently, we obtain

$$\begin{aligned}&\mathbf{\Pi}_{\mathsf{valid}}O_0|\psi_i'^{\mathsf{good}}\rangle\\ =&\mathbf{\Pi}_{\mathsf{valid}}O_0\sum_{\substack{M,T,Y,Z,\alpha,(D_0,D_1,D_2');\\(D_0,D_1,D_2'):\text{ valid}\\D_0(M)=\perp\\(D_0\cup(M,\alpha),D_1,D_2'):\text{good}}} a^i_{M,T,Y,Z,(D_0\cup(M,\alpha),D_1,D_2')}|M,T\rangle|Y\rangle|Z\rangle\\ &\qquad\otimes|D_0\cup(M,\alpha),D_1,D_2'\rangle\\ +&\mathbf{\Pi}_{\mathsf{valid}}O_0\sum_{\substack{M,T,Y,Z,(D_0,D_1,D_2');\\(D_0,D_1,D_2'):\text{ valid}\\D_0(M)=\perp\\(D_0,D_1,D_2'):\text{good}}} a^i_{M,T,Y,Z,(D_0,D_1,D_2')}|M,T\rangle|Y\rangle|Z\rangle\\ &\qquad\otimes|D_0,D_1,D_2'\rangle.\end{aligned}$$

To reduce notational clutter, we abbreviate the following summation condition:

$$\{M,T,Y,Z,(D_0,D_1,D_2');(D_0,D_1,D_2'):\text{valid};D_0(M)=\perp\}$$

by Δ_1 under the symbol $\sum$. Furthermore, the symbols M,T,Y,Z are omitted in instances indicated by the symbol a^i. Then $\mathbf{\Pi}_{\mathsf{valid}}O_0|\psi_i'^{\mathsf{good}}\rangle$ can be equivalently represented as

$$\begin{aligned}&\mathbf{\Pi}_{\mathsf{valid}}O_0|\psi_i'^{\mathsf{good}}\rangle\\ =&\mathbf{\Pi}_{\mathsf{valid}}O_0\sum_{\substack{\Delta_1,\alpha;\\(D_0\cup(M,\alpha),D_1,D_2'):\text{good}}} a^i_{D_0\cup(M,\alpha),D_1,D_2'}|M,T\rangle|Y\rangle|Z\rangle\otimes\left|D_0\cup(M,\alpha),D_1,D_2'\right\rangle\\ +&\mathbf{\Pi}_{\mathsf{valid}}O_0\sum_{\substack{\Delta_1;\\(D_0,D_1,D_2'):\text{good}}} a^i_{D_0,D_1,D_2'}|M,T\rangle|Y\rangle|Z\rangle\otimes\left|D_0,D_1,D_2'\right\rangle.\end{aligned}$$

Consequently, upon applying Proposition 2 to O_0, it follows that

$$\begin{aligned}&\mathbf{\Pi}_{\mathsf{valid}}O_0|\psi_i'^{\mathsf{good}}\rangle\\ =&\sum_{\substack{\Delta_1,\alpha;\\(D_0\cup(M,\alpha),D_1,D_2'):\text{good}}} a^i_{D_0\cup(M,\alpha),D_1,D_2'}|M,T\rangle|Y\rangle|Z\rangle\otimes|D_0\cup(M,\alpha),D_1,D_2'\rangle\\ &\qquad\otimes|\alpha\oplus T\rangle\\ -&\sum_{\substack{\Delta_1,\alpha,\gamma;\\(D_0\cup(M,\alpha),D_1,D_2'):\text{good}}} \tfrac{1}{2^n}a^i_{D_0\cup(M,\alpha),D_1,D_2'}|M,T\rangle|Y\rangle|Z\rangle\otimes|D_0\cup(M,\gamma),D_1,D_2'\rangle\\ &\qquad\otimes|\gamma\oplus T\rangle\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{\substack{\Delta_1,\alpha;\\ (D_0,D_1,D_2'):\text{good}}} \frac{1}{\sqrt{2^n}} a^i_{D_0,D_1,D_2'} |M,T\rangle|Y\rangle|Z\rangle \otimes \left|D_0 \cup (M,\alpha), D_1, D_2'\right\rangle \otimes |\alpha \oplus T\rangle \\
&+|\epsilon'\rangle \\
&\stackrel{\text{def}}{=}|S_1\rangle + |S_2\rangle + |S_3\rangle + |\epsilon'\rangle.
\end{aligned}
$$

where

$$
\begin{aligned}
|\epsilon'\rangle = &\sum_{\substack{\Delta_1,\alpha;\\ (D_0\cup(M,\alpha),D_1,D_2'):\text{good}}} \frac{\frac{1}{\sqrt{2^n}} a^i_{D_0\cup(M,\alpha),D_1,D_2'} |M,T\rangle|Y\rangle|Z\rangle \otimes}{\left(|D_0\rangle - \left(\sum_\gamma \frac{1}{\sqrt{2^n}} |D_0 \cup (M,\gamma)\rangle\right)\right) |D_1 D_2'\rangle \otimes |\alpha \oplus T\rangle} \\
&+ \sum_{\substack{\Delta_1,\alpha;\\ (D_0\cup(M,\alpha),D_1,D_2'):\text{good}}} \frac{\frac{1}{2^n} a^i_{D_0\cup(M,\alpha),D_1,D_2'} |M,T\rangle|Y\rangle|Z\rangle \otimes}{\left(2\sum_\gamma \frac{1}{\sqrt{2^n}} |D_0 \cup (M,\gamma)\rangle - |D_0\rangle\right) |D_1, D_2'\rangle \otimes |\widehat{0^n}\rangle} \\
&+ \sum_{\substack{\Delta_1;\\ (D_0,D_1,D_2'):\text{good}}} \frac{\frac{1}{\sqrt{2^n}} a^i_{D_0,D_1,D_2'} |M,T\rangle|Y\rangle|Z\rangle \otimes}{\left(|D_0\rangle - \left(\sum_\gamma \frac{1}{\sqrt{2^n}} |D_0 \cup (M,\gamma)\rangle\right)\right) |D_1, D_2'\rangle \otimes |\widehat{0^n}\rangle} \\
&\stackrel{\text{def}}{=}|S_4\rangle + |S_5\rangle + |S_6\rangle.
\end{aligned}
$$

In the following, we bound $\||S_1\rangle\|, \||S_2\rangle\|, \||S_3\rangle\|$, and $\||\epsilon'\rangle\|$, respectively. First, we have

$$
\begin{aligned}
\||S_4\rangle\|^2 = &\sum_{\substack{\Delta_1,\alpha\\ (D_0\cup(M,\alpha),D_1,D_2'):\text{good}}} \frac{1}{2^n} \left|a^i_{D_0\cup(M,\alpha),D_1,D_2'}\right|^2 \\
&+ \sum_{\substack{\Delta_1,\alpha;\\ (D_0\cup(M,\alpha),D_1,D_2'):\text{good}}} \frac{1}{2^{2n}} \left|a^i_{D_0\cup(M,\alpha),D_1,D_2'}\right|^2 \le O\left(\frac{1}{2^n}\right).
\end{aligned}
$$

Similarly $\||S_5\rangle\|^2 \le O\left(\frac{1}{2^n}\right)$ and $\||S_6\rangle\|^2 \le O\left(\frac{1}{2^n}\right)$. So

$$
\||\epsilon'\rangle\| \le O\left(\sqrt{\frac{1}{2^n}}\right).
$$

The same argument applies to $\mathbf{\Pi}_{\mathsf{valid}} O_0 |\psi_i^{\mathsf{good}}\rangle$ and $\||\epsilon\rangle\|$. We set

$$
\begin{aligned}
&|\psi_i^{\mathsf{good},1}\rangle := \mathbf{\Pi}_{\mathsf{good}}(\mathbf{\Pi}_{\mathsf{valid}} O_0 |\psi_i^{\mathsf{good}}\rangle - |\epsilon\rangle), \\
&|\psi_i^{\mathsf{bad},1}\rangle := O_0|\psi_i\rangle - |\psi_i^{\mathsf{good},1}\rangle, \\
&|\psi_i'^{\mathsf{good},1}\rangle := \mathbf{\Pi}_{\mathsf{good}}(\mathbf{\Pi}_{\mathsf{valid}} O_0 |\psi_i'^{\mathsf{good}}\rangle - |\epsilon'\rangle), \\
&|\psi_i'^{\mathsf{bad},1}\rangle := O_0|\psi_i'\rangle - |\psi_i'^{\mathsf{good},1}\rangle.
\end{aligned}
$$

where $\mathbf{\Pi}_{\mathsf{good}}$ denotes the projection onto the space spanned by the vectors that correspond to good databases and $\mathbf{\Pi}_{\mathsf{bad}}$ denotes the projection onto the space spanned by the vectors that correspond to bad databases. Then

$$
\mathbf{\Pi}_{\mathsf{bad}} |S_1\rangle = 0. \tag{11}
$$

Furthermore,

$$\mathbf{\Pi}_{\mathsf{bad}}|S_2\rangle = -\sum_{\substack{\Delta_1,\alpha,\gamma;(D_0\cup(M,\alpha),D_1,D_2')\text{:good}\\(D_0\cup(M,\gamma),D_1,D_2')\text{:bad}}} \frac{1}{2^n}a^i_{D_0\cup(M,\alpha),D_1,D_2'}|M,T\rangle|Y\rangle|Z\rangle \otimes |D_0\cup(M,\gamma),D_1,D_2'\rangle\otimes|\gamma\oplus T\rangle$$

$$= -\sum_{\substack{\Delta_1,\alpha,\gamma;(D_0\cup(M,\alpha),D_1,D_2')\text{:good}\\(D_0\cup(M,\gamma),D_1,D_2')\text{:bad}\\D_1(M_1)\neq\perp\wedge D_2'(M_2)\neq\perp}} \frac{1}{2^n}a^i_{D_0\cup(M,\alpha),D_1,D_2'}|M,T\rangle|Y\rangle|Z\rangle \otimes |D_0\cup(M,\gamma),D_1,D_2'\rangle\otimes|\gamma\oplus T\rangle$$

$$- \sum_{\substack{\Delta_1,\alpha,\gamma;(D_0\cup(M,\alpha),D_1,D_2')\text{:good}\\(D_0\cup(M,\gamma),D_1,D_2')\text{:bad}\\D_1(M_1)=\perp\vee(D_1(M_1)\neq\perp\wedge D_2'(M_2)=\perp)}} \frac{1}{2^n}a^i_{D_0\cup(M,\alpha),D_1,D_2'}|M,T\rangle|Y\rangle|Z\rangle \otimes |D_0\cup(M,\gamma),D_1,D_2'\rangle\otimes|\gamma\oplus T\rangle$$

$$\stackrel{\text{def}}{=}|S_7\rangle + |S_8\rangle.$$

For the upper bound of $||S_7\rangle + |S_8\rangle||$, we must consider the global collision condition. A newly sampled output $V_0 = \gamma$ causes the updated database $(D_0 \cup (M,\gamma), D_1, D_2')$ to transition from good to bad if and only if γ satisfies the internal collision equation with existing records:

$$\gamma \oplus U_1 \oplus V_1 = V_0' \oplus U_1' \oplus V_1' \implies \gamma = V_0' \oplus U_1' \oplus V_1' \oplus U_1 \oplus V_1.$$

To find the number of such bad γ values, we count the free variables on the right-hand side. The value V_0' can be any existing output in D_0 (at most i choices). The tuple (U_1, V_1) can be any existing entry in D_1 (at most i choices), and (U_1', V_1') can be any other entry in D_1 (at most i choices).

Since the pre-oracle cannot predict the adversary's future input paths, it must guard against all possible historical combinations. The total number of γ values that trigger a bad global state is bounded by the Cartesian product of these sets, which is $i\times i\times i = i^3$.

Therefore, we have

$$||S_7+S_8||^2 \le \sum_{\substack{\Delta_1,\gamma;\\(D_0\cup(M,\gamma),D_1,D_2')\text{:bad}}} \frac{1}{2^{2n}}\cdot i^3 \cdot \sum_{\substack{\alpha;\\(D_0\cup(M,\alpha),D_1,D_2')\text{:good}}} \frac{1}{2^n}|a^i|^2 \le O\left(\frac{i^3}{2^n}\right).$$

It follows that

$$||\mathbf{\Pi}_{\mathsf{bad}}(|S_2\rangle)|| \le O\left(\sqrt{\frac{i^3}{2^n}}\right). \tag{12}$$

In addition,

$$\|\mathbf{\Pi}_{\mathsf{bad}}|S_3\rangle\|^2 = \sum_{\substack{\Delta_1,\alpha;\\(D_0\cup(M,\alpha),D_1,D_2')\text{:bad}}} \frac{\left|a^i_{D_0,D_1D_2'}\right|^2\cdot|\{\alpha|(D_0\cup(M,\alpha),D_1,D_2')\text{is bad}\}|}{2^n}$$

$$\leq O\left(\frac{i^3}{2^n}\right).$$

Thus,

$$\|\mathbf{\Pi}_{\mathsf{bad}}|S_3\rangle\| \leq O\left(\sqrt{\frac{i^3}{2^n}}\right). \tag{13}$$

From (11), (12), and (13), we have

$$\|\mathbf{\Pi}_{\mathsf{bad}}(\mathbf{\Pi}_{\mathsf{valid}}O_0|\psi_i'^{\mathsf{good}}\rangle - |\epsilon'\rangle)\| \leq O\left(\sqrt{\frac{i^3}{2^n}}\right). \tag{14}$$

So

$$\begin{aligned}
\||\psi_i'^{\mathsf{bad},1}\rangle\| =& \|O_0|\psi_i'\rangle - |\psi_i'^{\mathsf{good},1}\rangle\| \\
=& \|\mathbf{\Pi}_{\mathsf{valid}}O_0|\psi_i'\rangle - \mathbf{\Pi}_{\mathsf{good}}(\mathbf{\Pi}_{\mathsf{valid}}O_0|\psi_i'^{\mathsf{good}}\rangle - |\epsilon'\rangle)\| \\
=& \|\mathbf{\Pi}_{\mathsf{bad}}(\mathbf{\Pi}_{\mathsf{valid}}O_0|\psi_i'^{\mathsf{good}}\rangle - |\epsilon'\rangle) + \mathbf{\Pi}_{\mathsf{valid}}O_0|\psi_i'^{\mathsf{bad}}\rangle + |\epsilon'\rangle\| \\
\leq& \|\mathbf{\Pi}_{\mathsf{bad}}(\mathbf{\Pi}_{\mathsf{valid}}O_0|\psi_i'^{\mathsf{good}}\rangle - |\epsilon'\rangle)\| + \||\psi_i'^{\mathsf{bad}}\rangle\| + \||\epsilon'\rangle\| \\
\leq& \||\psi_i'^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right).
\end{aligned}$$

The same argument applies to $\||\psi_i^{\mathsf{bad},1}\rangle\|$.

The next step involves the operation on O_1:

$$O_1 \cdot O_0|\psi_i'\rangle = |\psi_i'^{\mathsf{good},2}\rangle + |\psi_i'^{\mathsf{bad},2}\rangle, O_1 \cdot O_0|\psi_i\rangle = |\psi_i^{\mathsf{good},2}\rangle + |\psi_i^{\mathsf{bad},2}\rangle.$$

From (2), (9) and (10) there exists a complex number $a^{(i),2}_{M,T,Y,Z,(D_0,D_1,D_2')}$ such that

$$|\psi_i'^{\mathsf{good},2}\rangle = \sum_{\substack{M,T,Y,Z,(D_0,D_1,D_2');\\ (D_0,D_1,D_2'):\text{ valid and good}\\ D_0(M)\neq\perp,D_1(M_1)\neq\perp}} a^{(i),2}_{M,T,Y,Z,(D_0,D_1,D_2')}|M,T\rangle|Y\rangle|Z\rangle \otimes |D_0,D_1,D_2'\rangle \otimes|M_1\rangle\otimes|M_2\rangle \tag{15}$$

and

$$|\psi_i^{\mathsf{good},2}\rangle = \sum_{\substack{M,T,Y,Z,(D_0,D_1,D_2');\\ (D_0,D_1,D_2'):\text{ valid and good}\\ D_0(M)\neq\perp,D_1(M_1)\neq\perp}} a^{(i),2}_{M,T,Y,Z,(D_0,D_1,D_2')}|M,T\rangle|Y\rangle|Z\rangle\otimes \big|[D_0,D_1,D_2']_2\big\rangle\otimes|M_1\rangle\otimes|M_2\rangle \tag{16}$$

Lemma 4 (Action of O_1). *For $|\psi_i'^{\mathsf{bad},2}\rangle$ and $|\psi_i^{\mathsf{bad},2}\rangle$, we have*

$$\||\psi_i'^{\mathsf{bad},2}\rangle\| \leq \||\psi_i'^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right), \||\psi_i^{\mathsf{bad},2}\rangle\| \leq \||\psi_i^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right).$$

Proof. The proof structure is similar to that of Lemma 3. When O_1 samples a new output $V_1 = \alpha$ for input U_1, the database becomes bad if α satisfies $\alpha = V_0 \oplus U_1 \oplus V_0' \oplus U_1' \oplus V_1'$. Since $V_0 \in D_0, V_0' \in D_0$ and $(U_1', V_1') \in D_1$ can be chosen from any of the at most i previous queries, the number of bad α values is bounded by i^3. Therefore, tracking the transition amplitude similarly yields an error bound of $O(\sqrt{i^3/2^n})$.

The next step involves the operations on O_2 and O_2':

$$O_2' \cdot O_1 \cdot O_0 |\psi_i'\rangle = |\psi_i'^{\mathsf{good},3}\rangle + |\psi_i'^{\mathsf{bad},3}\rangle,$$
$$O_2 \cdot O_1 \cdot O_0 |\psi_i\rangle = |\psi_i^{\mathsf{good},3}\rangle + |\psi_i^{\mathsf{bad},3}\rangle.$$

From (3), (4), (15), and (16), there exists a complex number $a^{(i),3}_{M,T,Y,Z,(D_0,D_1,D_2')}$ such that

$$|\psi_i'^{\mathsf{good},3}\rangle = \sum_{\substack{M,T,Y,Z,(D_0,D_1,D_2');\\ (D_0,D_1,D_2'):\text{ valid and good}\\ D_0(M)\neq\perp, D_1(M_1)\neq\perp}} a^{(i),3}_{M,T,Y,Z,(D_0,D_1,D_2')} |M,T\rangle|Y\rangle|Z\rangle \otimes |D_0,D_1,D_2'\rangle \otimes |M_1\rangle \otimes |M_2\rangle \tag{17}$$

and

$$|\psi_i^{\mathsf{good},3}\rangle = \sum_{\substack{M,T,Y,Z,(D_0,D_1,D_2');\\ (D_0,D_1,D_2'):\text{ valid and good}\\ D_0(M)\neq\perp, D_1(M_1)\neq\perp}} a^{(i),3}_{M,T,Y,Z,(D_0,D_1,D_2')} |M,T\rangle|Y\rangle|Z\rangle \otimes \left|[D_0,D_1,D_2']_2\right\rangle \otimes |M_1\rangle \otimes |M_2\rangle \tag{18}$$

Lemma 5 (Action of O_2 and O_2'). *For $|\psi_i'^{\mathsf{bad},3}\rangle$ and $|\psi_i^{\mathsf{bad},3}\rangle$, we have*

$$\||\psi_i'^{\mathsf{bad},3}\rangle\| \le \||\psi_i'^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right), \||\psi_i^{\mathsf{bad},3}\rangle\| \le \||\psi_i^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right).$$

Proof. A crucial observation follows from our purely internal definition of a bad database: the badness condition $V_0 \oplus U_1 \oplus V_1 = V_0' \oplus U_1' \oplus V_1'$ is strictly independent of D_2 and D_2'. Therefore, when O_2 (or O_2') evaluates an input and samples a new output V_2, no possible value of V_2 can cause a good database to transition into a bad one. The set of bad outputs for this specific action is empty.

Thus, apart from the standard RstOE error term, the structural collision error is zero:

$$\||\psi_i'^{\mathsf{bad},3}\rangle\| \le \||\psi_i'^{\mathsf{bad},2}\rangle\| \le \||\psi_i'^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right).$$

The same argument applies to $|\psi_i^{\mathsf{good},3}\rangle$, $|\psi_i^{\mathsf{bad},3}\rangle$, and $\||\psi_i^{\mathsf{bad},3}\rangle\|$.

From (5) and (6), we uncompute through O_1^*. Then

$$O_1^* \cdot O_2' \cdot O_1 \cdot O_0 |\psi_i'\rangle = |\psi_i'^{\mathsf{good},4}\rangle + |\psi_i'^{\mathsf{bad},4}\rangle,$$

$$O_1^* \cdot O_2 \cdot O_1 \cdot O_0 |\psi_i\rangle = |\psi_i^{\mathsf{good},4}\rangle + |\psi_i^{\mathsf{bad},4}\rangle.$$

From (17) and (18), there exists a complex number $a^{(i),4}_{M,T,Y,Z,(D_0,D_1,D_2')}$ such that

$$|\psi_i'^{\mathsf{good},4}\rangle = \sum_{\substack{M,T,Y,Z,(D_0,D_1,D_2');\\ (D_0,D_1,D_2')\text{: valid and good}\\ D_0(M)\neq\perp}} a^{(i),4}_{M,T,Y,Z,(D_0,D_1,D_2')} |M,T\rangle|Y\rangle|Z\rangle \otimes |D_0,D_1,D_2'\rangle \otimes |M_1\rangle \tag{19}$$

and

$$|\psi_i^{\mathsf{good},4}\rangle = \sum_{\substack{M,T,Y,Z,(D_0,D_1,D_2');\\ (D_0,D_1,D_2')\text{: valid and good}\\ D_0(M)\neq\perp}} a^{(i),4}_{M,T,Y,Z,(D_0,D_1,D_2')} |M,T\rangle|Y\rangle|Z\rangle \otimes \left|[D_0,D_1,D_2']_2\right\rangle \otimes |M_1\rangle \tag{20}$$

Lemma 6 (Action of O_1^*). *For $|\psi_i'^{\mathsf{bad},4}\rangle$ and $|\psi_i^{\mathsf{bad},4}\rangle$, we have*

$$\||\psi_i'^{\mathsf{bad},4}\rangle\| \leq \||\psi_i'^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right), \||\psi_i^{\mathsf{bad},4}\rangle\| \leq \||\psi_i^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right).$$

Proof. Before proceeding with the proof, we first present some definitions. A state vector $|D_0,D_1,D_2'\rangle \otimes |Y\rangle\,|Z\rangle$ for O_{TNT_b}, where $|Y\rangle\,|Z\rangle$ is the ancillary $2n$ qubits, is regular if $|Y\rangle = |0^n\rangle$, $|Z\rangle = |0^n\rangle$ and the database is valid. Similarly, a state vector $|D_0,D_1,D_2'\rangle \otimes |Y\rangle\,|Z\rangle$ is preregular if $|Z\rangle = |0^n\rangle$ and the database is valid. The definitions for O_{TNT_s} are analogous. Let $\mathbf{\Pi}_{\mathsf{prereg}}$ denote the projection onto the space spanned by the vectors that correspond to preregular databases. Let

$$|\psi_i'^{\mathsf{good},4}\rangle := \mathbf{\Pi}_{\mathsf{good}}\mathbf{\Pi}_{\mathsf{prereg}} O_1^* |\psi_i'^{\mathsf{good},3}\rangle,$$
$$|\psi_i'^{\mathsf{bad},4}\rangle := O_1^* \cdot O_2' \cdot O_1 \cdot O_0 |\psi_i'\rangle - |\psi_i'^{\mathsf{good},4}\rangle.$$

The same argument applies to $|\psi_i^{\mathsf{good},4}\rangle$ and $|\psi_i^{\mathsf{bad},4}\rangle$. We have

$$\begin{aligned}
&\mathbf{\Pi}_{\mathsf{prereg}} O_1^* |\psi_i'^{\mathsf{good},3}\rangle \\
=&\mathbf{\Pi}_{\mathsf{prereg}} O_1^* \sum_{\substack{M,T,Y,Z,(D_0,D_1,D_2');\\ (D_0,D_1,D_2')\text{: valid and good}\\ D_0(M)\neq\perp, D_1(M_1)\neq\perp}} a^{i,3}_{M,T,Y,Z,(D_0,D_1,D_2')} |M,T\rangle|Y\rangle|Z\rangle \otimes |D_0,D_1,D_2'\rangle \otimes |M_1\rangle \otimes |M_2\rangle \\
=&\mathbf{\Pi}_{\mathsf{prereg}} O_1^* \sum_{\substack{M,T,Y,Z,\alpha,(D_0,D_1,D_2');\\ (D_0,D_1\cup(M_1,\alpha),D_2')\text{: valid and good}\\ D_0(M)\neq\perp, D_1(M_1)=\perp}} a^{i,3}_{M,T,Y,Z,(D_0,D_1\cup(M_1,\alpha),D_2')} |M,T\rangle|Y\rangle|Z\rangle \otimes \\
&\qquad |D_0,D_1\cup(M_1,\alpha),D_2'\rangle \otimes |M_1\rangle \otimes |M_2\rangle.
\end{aligned}$$

To streamline the presentation, we abbreviate the condition

$$\{M,T,Y,Z,(D_0,D_1,D_2');D_0(M)\neq\perp,D_1(M_1)=\perp\}$$

as Δ_2 under the summation symbol. Furthermore, in terms indicated by $a^{i,3}$, the variables M, T, Y, Z are omitted. Consequently, $\mathbf{\Pi}_{\mathsf{prereg}} O_1^* |\psi_i'^{\mathsf{good},3}\rangle$ can be expressed as

$$
\begin{aligned}
&\mathbf{\Pi}_{\mathsf{prereg}} O_1^* |\psi_i'^{\mathsf{good},3}\rangle \\
=& \sum_{\substack{\Delta_2,\alpha;\\ (D_0, D_1\cup(M_1,\alpha), D_2')\text{: valid and good}}} a^{i,3}_{D_0,D_1\cup(M_1,\alpha),D_2'} |M,T\rangle|Y\rangle|Z\rangle\otimes \\
&\qquad |D_0, D_1\cup(M_1,\alpha), D_2'\rangle \otimes |M_1\rangle \\
+& \sum_{\substack{\Delta_2,\alpha;\\ (D_0, D_1\cup(M_1,\alpha), D_2')\text{: valid and good}}} \tfrac{1}{\sqrt{2^n}} a^{i,3}_{D_0,D_1\cup(M_1,\alpha),D_2'} |M,T\rangle|Y\rangle|Z\rangle \otimes |D_0\rangle \\
&\qquad (|D_1\rangle - (\textstyle\sum_\gamma \frac{1}{\sqrt{2^n}} |D_1\cup(M_1,\gamma)\rangle)) |D_2'\rangle \otimes |M_1\rangle \\
-& \sum_{\substack{\Delta_2,\alpha;\\ (D_0, D_1\cup(M_1,\alpha), D_2')\text{: valid and good}}} \tfrac{1}{2^n} a^{i,3}_{D_0,D_1\cup(M_1,\alpha),D_2'} |M,T\rangle|Y\rangle|Z\rangle\otimes \\
&\qquad |D_0, D_1\cup(M_1,\alpha), D_2'\rangle \otimes |M_1\rangle \\
+& \sum_{\substack{\Delta_2,\alpha;\\ (D_0, D_1\cup(M_1,\alpha), D_2')\text{: valid and good}}} \tfrac{1}{2^{3n/2}} a^{i,3}_{D_0,D_1\cup(M_1,\alpha),D_2'} |M,T\rangle|Y\rangle|Z\rangle \otimes |D_0\rangle \\
&\qquad \left(2\textstyle\sum_\delta \frac{1}{\sqrt{2^n}} |D_1\cup(M_1,\delta)\rangle - |D_1\rangle\right) |D_2'\rangle \otimes |M_1\rangle \\
\stackrel{\text{def}}{=}& |S_9\rangle + |S_{10}\rangle + |S_{11}\rangle + |S_{12}\rangle.
\end{aligned}
$$

Then

$$\mathbf{\Pi}_{\mathsf{bad}} |S_9\rangle = \mathbf{\Pi}_{\mathsf{bad}} |S_{11}\rangle = 0. \tag{21}$$

Moreover, we have

$$
\begin{aligned}
&\mathbf{\Pi}_{\mathsf{bad}} |S_{10}\rangle \\
=& \mathbf{\Pi}_{\mathsf{bad}} \sum_{\substack{\Delta_2,\alpha;\\ (D_0, D_1\cup(M_1,\alpha), D_2')\text{: valid and good}}} \tfrac{1}{\sqrt{2^n}} a^{i,3}_{D_0,D_1\cup(M_1,\alpha),D_2'} |M,T\rangle|Y\rangle|Z\rangle\otimes \\
&\qquad |D_0, D_1, D_2'\rangle \otimes |M_1\rangle \\
-& \mathbf{\Pi}_{\mathsf{bad}} \sum_{\substack{\Delta_2,\alpha,\gamma;\\ (D_0, D_1\cup(M_1,\alpha), D_2')\text{: valid and good}}} \tfrac{1}{2^n} a^{i,3}_{D_0,D_1\cup(M_1,\alpha),D_2'} |M,T\rangle|Y\rangle|Z\rangle\otimes \\
&\qquad |D_0, D_1\cup(M_1,\gamma), D_2'\rangle \otimes |M_1\rangle \\
\stackrel{\text{def}}{=}& |S_{13}\rangle + |S_{14}\rangle.
\end{aligned}
$$

For the upper bound of $||S_{13}\rangle||^2$, we evaluate the amplitude transitioning into the bad subspace during the uncomputation step O_1^*. A state transitions to a bad database if and only if the removed/evaluated entry satisfies the global pure-database collision equation $V_1 = V_0 \oplus U_1 \oplus V_0' \oplus U_1' \oplus V_1'$. To determine the number of dangerous configurations, the oracle must account for all possible combinations of existing records on the right-hand side. The degrees of freedom consist of at most i choices for $V_0 \in D_0$, i choices for $(U_1, V_1) \in D_1$, and i choices for $(U_1', V_1') \in D_1$. This Cartesian product bounds the number of bad configurations by $i\times i\times i = i^3$. Therefore, tracking the transition amplitude for S_{13} yields

$$
\||S_{13}\rangle\|^2 = \left\| \sum_{\substack{\Delta_2,\alpha;(D_0, D_1\cup(M_1,\alpha), D_2')\text{: valid and good}\\ D_2'(M_2)\neq\perp}} \tfrac{1}{\sqrt{2^n}} a^{i,3}_{D_0,D_1\cup(M_1,\alpha),D_2'} |M,T\rangle|Y\rangle|Z\rangle \otimes |D_0, D_1, D_2'\rangle \otimes |M_1\rangle \right\|^2
$$

$$
\begin{aligned}
&= \sum_{\Delta_2} \frac{1}{2^n} \left\| \sum_{\substack{\alpha;(D_0,D_1\cup(M_1,\alpha),D_2'):\text{ valid and good} \\ D_2'(M_2)\neq\perp}} a^{i,3}_{D_0,D_1\cup(M_1,\alpha),D_2'} \right\|^2 \\
&\leq \sum_{\Delta_2} \frac{i^3}{2^n} \sum_{\substack{\alpha;(D_0,D_1\cup(M_1,\alpha),D_2'):\text{ valid and good} \\ D_2'(M_2)\neq\perp}} \left| a^{i,3}_{D_0,D_1\cup(M_1,\alpha),D_2'} \right|^2 \leq O\left(\frac{i^3}{2^n}\right). \quad (22)
\end{aligned}
$$

For the upper bound of $|S_{14}\rangle$, we have

$$
\begin{aligned}
\||S_{14}\rangle\|^2 = &\left\| \mathbf{\Pi}_{\mathsf{bad}} \sum_{\substack{\Delta_2,\alpha,\gamma;D_2'(M_2)\neq\perp \\ (D_0,D_1\cup(M_1,\alpha),D_2'):\text{ valid and good}}} \frac{1}{2^n} a^{i,3}_{D_0,D_1\cup(M_1,\alpha),D_2'} |M,T\rangle|Y\rangle|Z\rangle\otimes \right. \\
&\left. \qquad |D_0, D_1\cup(M_1,\gamma), D_2'\rangle \otimes |M_1\rangle \right\|^2 \\
&+ \left\| \mathbf{\Pi}_{\mathsf{bad}} \sum_{\substack{\Delta_2,\alpha,\gamma;D_2'(M_2)=\perp \\ (D_0,D_1\cup(M_1,\alpha),D_2'):\text{ valid and good}}} \frac{1}{2^n} a^{i,3}_{D_0,D_1\cup(M_1,\alpha),D_2'} |M,T\rangle|Y\rangle|Z\rangle\otimes \right. \\
&\left. \qquad |D_0, D_1\cup(M_1,\gamma), D_2'\rangle \otimes |M_1\rangle \right\|^2 \\
&\stackrel{\text{def}}{=} \||S_{15}\rangle\|^2 + \||S_{16}\rangle\|^2.
\end{aligned}
$$

Similarly, the condition for these states to fall into the bad subspace relies on the exact same collision constraint: $V_1 = V_0 \oplus U_1 \oplus V_0' \oplus U_1' \oplus V_1'$. Because the historical sub-databases D_0 and D_1 contain at most i records each, the summation over the γ values (which act as the uncomputed V_1) is constrained by the same i^3 combinatorial bound. Thus, we have $\||S_{15}\rangle\|^2 \leq O\left(\frac{i^3}{2^n}\right)$ and $\||S_{16}\rangle\|^2 \leq O\left(\frac{i^3}{2^n}\right)$. Then

$$
\|\mathbf{\Pi}_{\mathsf{bad}} |S_{14}\rangle\| \leq O\left(\sqrt{\frac{i^3}{2^n}}\right). \quad (23)
$$

From (22) and (23), $\mathbf{\Pi}_{\mathsf{bad}}|S_{10}\rangle \leq O\left(\sqrt{\frac{i^3}{2^n}}\right)$. Similarly, $\mathbf{\Pi}_{\mathsf{bad}}|S_{12}\rangle \leq O\left(\sqrt{\frac{i^3}{2^n}}\right)$. Thus $\|\mathbf{\Pi}_{\mathsf{bad}}\mathbf{\Pi}_{\mathsf{prereg}} O_1^* |\psi_i'^{\mathsf{good},3}\rangle\| \leq O\left(\sqrt{\frac{i^3}{2^n}}\right)$ and

$$
\begin{aligned}
\||\psi_i'^{\mathsf{bad},4}\rangle\| =& \|O_1^* \cdot O_2' \cdot O_1 \cdot O_0 |\psi_i'\rangle - |\psi_i'^{\mathsf{good},4}\rangle\| \\
=& \|\mathbf{\Pi}_{\mathsf{prereg}} O_1^* (|\psi_i'^{\mathsf{good},3}\rangle + |\psi_i'^{\mathsf{bad},3}\rangle) - \mathbf{\Pi}_{\mathsf{good}}\mathbf{\Pi}_{\mathsf{prereg}} O_1^* |\psi_i'^{\mathsf{good},3}\rangle\| \\
\leq& \|\mathbf{\Pi}_{\mathsf{bad}}\mathbf{\Pi}_{\mathsf{prereg}} O_1^* |\psi_i'^{\mathsf{good},3}\rangle\| + \||\psi_i'^{\mathsf{bad},3}\rangle\| \\
\leq& \||\psi_i'^{\mathsf{bad},3}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right) \leq \||\psi_i'^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right).
\end{aligned}
$$

The same argument applies to $\||\psi_i^{\mathsf{bad},4}\rangle\|$.

Finally, we uncompute through O_0^*:

$$O_0^* \cdot O_1^* \cdot O_2' \cdot O_1 \cdot O_0 |\psi_i'\rangle = |\psi_i'^{\mathsf{good},5}\rangle + |\psi_i'^{\mathsf{bad},5}\rangle,$$
$$O_0^* \cdot O_1^* \cdot O_2 \cdot O_1 \cdot O_0 |\psi_i\rangle = |\psi_i^{\mathsf{good},5}\rangle + |\psi_i^{\mathsf{bad},5}\rangle.$$

From (19) and (20), there exists a complex number $a^{(i),5}_{M,T,Y,Z,(D_0,D_1,D_2')}$ such that

$$|\psi_i'^{\mathsf{good},5}\rangle = \sum_{\substack{M,T,Y,Z,(D_0,D_1,D_2');\\ (D_0,D_1,D_2'):\text{ valid and good}\\ D_0(M)\neq\perp}} a^{(i),5}_{M,T,Y,Z,(D_0,D_1,D_2')} |M,T\rangle|Y\rangle|Z\rangle \otimes |D_0,D_1,D_2'\rangle \tag{24}$$

and

$$|\psi_i^{\mathsf{good},5}\rangle = \sum_{\substack{M,T,Y,Z,(D_0,D_1,D_2');\\ (D_0,D_1,D_2'):\text{ valid and good}\\ D_0(M)\neq\perp}} a^{(i),5}_{M,T,Y,Z,(D_0,D_1,D_2')} |M,T\rangle|Y\rangle|Z\rangle \otimes \big|[D_0,D_1,D_2']_2\big\rangle \tag{25}$$

Lemma 7 (Action of O_0^*). *For $|\psi_i'^{\mathsf{bad},5}\rangle$ and $|\psi_i^{\mathsf{bad},5}\rangle$, we have*

$$\||\psi_i'^{\mathsf{bad},5}\rangle\| \le \||\psi_i'^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right), \||\psi_i^{\mathsf{bad},5}\rangle\| \le \||\psi_i^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right).$$

Proof. The proof structure is similar to that of Lemma 6. During the uncomputation step O_0^*, the database transitions into a bad state if the removed output γ (which corresponds to V_0) satisfies the pure-database collision equation $\gamma = V_0' \oplus U_1' \oplus V_1' \oplus U_1 \oplus V_1$. To determine the risk, the oracle must account for existing historical records on the right-hand side. Since $V_0' \in D_0$, $(U_1, V_1) \in D_1$, and $(U_1', V_1') \in D_1$ can be chosen from any of the at most i previous queries, the number of dangerous γ configurations is bounded by $i \times i \times i = i^3$. Following the same norm derivation as in Lemma 6, the amplitude error introduced during the uncomputation of O_0 remains bounded by $O(\sqrt{i^3/2^n})$.

Proof. (Proof of Proposition 4.) Let $|\psi_1'^{\mathsf{bad}}\rangle = |\psi_1^{\mathsf{bad}}\rangle = 0$. For $|\psi_{i+1}'^{\mathsf{good}}\rangle$, $|\psi_{i+1}'^{\mathsf{bad}}\rangle$, $|\psi_{i+1}^{\mathsf{good}}\rangle$, and $|\psi_{i+1}^{\mathsf{bad}}\rangle$, Lemmas 3–7 imply

$$\||\psi_{i+1}'^{\mathsf{bad}}\rangle\| \le \||\psi_i'^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right), \||\psi_{i+1}^{\mathsf{bad}}\rangle\| \le \||\psi_i^{\mathsf{bad}}\rangle\| + O\left(\sqrt{\frac{i^3}{2^n}}\right).$$

B Quantum Attacks on TNT

B.1 Quantum Algorithms

In 2002, Brassard et al. [BHMT02] proposed quantum amplitude amplification as a method for solving a more general quantum search problem. Assume that, for a random

x, the probability of $f(x) = 1$ is p. Then quantum amplitude amplification can find a good solution using a number of iterations proportional to $1/\sqrt{p}$ in the worst case.

Quantum amplitude amplification [BHMT02], abbreviated as QAA in this paper, is a generalization of Grover search that increases the success probability of any measurement-free quantum algorithm by iterating it. Let $\mathcal{A}$ be a quantum circuit such that $\mathcal{A}|0\rangle = (\sum_{x\in G} \alpha_x|x\rangle)|0\rangle + (\sum_{x\in B} \beta_x|x\rangle)|1\rangle = \sqrt{p}\,|\psi_G\rangle + \sqrt{1-p}\,|\psi_B\rangle$, where p is the success probability of $\mathcal{A}$ (real and positive); $|\psi_G\rangle$ is a superposition (not necessarily uniform) of good outcomes (the set G) and $|\psi_B\rangle$ of bad outcomes (the set B), marked by the respective flags 1 and 0. Let O_0 be the inversion-around-zero operator that flips the phase of the basis vector $|0\rangle$: it maps $O_0|y\rangle = -|y\rangle$ if and only if $y = 0$. Let O be the operator that flips the phase of all basis vectors $|x, b\rangle$ such that $b = 1$. The QAA computes a sequence of states $|\psi_i\rangle$ defined by the following iterative process (we denote $\mathcal{A}^\dagger$ as the inverse of $\mathcal{A}$):

1. $|\psi_0\rangle = \mathcal{A}|0\rangle$;
2. for $i = 0$ to $m-1$: $|\psi_{i+1}\rangle = \mathcal{A}O_0\mathcal{A}^\dagger O\,|\psi_i\rangle$.

Let $\theta = \arcsin(\sqrt{p})$, then we have $|\psi_0\rangle = \sin(\theta)\,|\psi_G\rangle + \cos(\theta)\,|\psi_B\rangle$, and $|\psi_i\rangle = \sin((2i+1)\theta)\,|\psi_G\rangle + \cos((2i+1)\theta)\,|\psi_B\rangle$. If we measure the state $|\psi_i\rangle$, we obtain a good state x_{good} with probability $\sin^2((2i+1)\theta)$. Thus, after $t = \lfloor \frac{\pi}{4} \times \frac{1}{\sqrt{p}} \rfloor$ iterations of QAA (for small p, $\theta \approx \sqrt{p}$, and $\lfloor\cdot\rfloor$ denotes the floor function), the success probability is close to 1.

Ambainis's Theorem (Theorem 3 in [Amb07]) Let $f : \mathcal{X} \to \mathcal{Y}$ be a function, where $\mathcal{X}$ and $\mathcal{Y}$ are finite sets. Then there exists a quantum algorithm that decides whether there are distinct elements $x_1, x_2 \in \mathcal{X}$ such that $f(x_1) = f(x_2)$ with probability at least $1-\epsilon$ and bounded error $\epsilon < 1/2$, using $O\left(|\mathcal{X}|^{2/3}\right)$ quantum queries to f.

B.2 qPRF Attack for $\mathsf{TNT}[f_0, f_1, f_2]$

Considering $\mathsf{TNT}[f_0, f_1, f_2](M, T) = f_2(T \oplus f_1(T \oplus f_0(M)))$ and setting $M = x, T = 0$, we obtain $\mathsf{TNT}[f_0, f_1, f_2](x, 0) = f_2(f_1(f_0(x)))$, where f_0, f_1, f_2 are independent random functions. This simplifies the qPRF attack on $\mathsf{TNT}[f_0, f_1, f_2]$ to a qPRF attack on the composition $f_2 \circ f_1 \circ f_0$. The next proposition clarifies this.

Proposition 9. *Let $f_2 \circ f_1 \circ f_0$ be the composition of three independent random functions $f_2, f_1, f_0 : \{0,1\}^n \to \{0,1\}^n$. There exists a quantum adversary $\mathcal{A}$ that makes $O(2^{n/3})$ quantum queries, such that $\mathbf{Adv}^{\mathsf{qPRF}}_{f_2\circ f_1\circ f_0}(\mathcal{A}) \geq \frac{2}{125}$.*

Proof. Let $f : \{0,1\}^n \to \{0,1\}^n$ be an independent random function. For $1 \leq N \leq 2^n$, let $[N]$ denote the subset $\{0, 1, \ldots, N-1\}$. To use Ambainis's theorem, let $\mathcal{D}^f_N$ be a quantum algorithm with $O\left(|N|^{2/3}\right)$ quantum queries and error $\epsilon < 1/25$, defined by

$$\mathcal{D}^f_N = \begin{cases} 1, & \exists x_1, x_2 \in [N] \text{ s.t. } f(x_1) = f(x_2), \\ 0, & \text{otherwise.} \end{cases}$$

Let $\mathrm{coll}^f_{[N]}$ denote the event that f has a collision in $[N]$, and let $p = \Pr_f\left[\neg\, \mathrm{coll}^f_{[N]}\right]$. Then $p = \prod_{j=1}^{N-1}\left(1 - \frac{j}{2^n}\right)$. Existing research, as detailed in Equation (87) of [HI19b],

shows that $\Pr_{f_0,f_1}\left[\neg \operatorname{coll}_{[N]}^{f_1\circ f_0}\right] = p^2$. Then

$$\begin{aligned}&\Pr_{f_0,f_1,f_2}\left[\neg \operatorname{coll}_{[N]}^{f_2\circ f_1\circ f_0}\right]\\ =&\Pr_{f_0,f_1,f_2}\left[\neg \operatorname{coll}_{f_1\circ f_0([N])}^{f_2} \mid \neg \operatorname{coll}_{[N]}^{f_1\circ f_0}\right]\\ &\cdot \Pr_{f_0,f_1}\left[\neg \operatorname{coll}_{[N]}^{f_1\circ f_0}\right] = p^3.\end{aligned}$$

With error $\epsilon < 1/25$, we have

$$\begin{aligned}\mathbf{Adv}^{\mathsf{qPRF}}_{f_2\circ f_1\circ f_0}(\mathcal{D}_N) =& \left|\Pr_f\left[\mathcal{D}_N^f() \Rightarrow 1\right] - \Pr_{f_0,f_1,f_2}\left[\mathcal{D}_N^{f_2\circ f_1\circ f_0}() \Rightarrow 1\right]\right|\\ \geq& \left|\Pr_f\left[\operatorname{coll}_{[N]}^f\right] - \Pr_{f_0,f_1,f_2}\left[\operatorname{coll}_{[N]}^{f_2\circ f_1\circ f_0}\right]\right| - \frac{2}{25}\\ =&(1-p^3) - (1-p) - \frac{2}{25} = (p+p^2)(1-p) - \frac{2}{25}.\end{aligned}$$

[HI19b] shows that there exist $N_0 = O\left(2^{n/2}\right)$ and $p_0 = \prod_{j=1}^{N_0-1}\left(1-\frac{j}{2^n}\right)$ such that $\frac{1}{5} \leq p_0 \leq \frac{3}{5}$ holds for sufficiently large n. Here "sufficiently large" means that n satisfies $e^{-\frac{N_0(N_0-1)}{2\cdot 2^{n/2}}} \leq 3/5$, where $N_0 = 2^{n/4}\sqrt{2\log 2}$. Thus, there exists a parameter N_0 in $O\left(2^{n/2}\right)$ such that $\frac{1}{5} \leq p_0 \leq \frac{3}{5}$ holds for sufficiently large n. Then

$$\mathbf{Adv}^{\mathsf{qPRF}}_{f_2\circ f_1\circ f_0}(\mathcal{D}_{N_0}) \geq \left(\frac{1}{5}+\frac{1}{5^2}\right)\left(1-\frac{3}{5}\right) - \frac{2}{25} = \frac{2}{125}.$$

By Ambainis' theorem, $\mathcal{D}_{N_0}$ makes at most $O((N_0)^{2/3}) = O((2^{n/2})^{2/3}) = O(2^{n/3})$ quantum queries. This proves Proposition 9.

Theorem 3 (qPRF **Attack).** *There exists a quantum adversary $\mathcal{A}$ making $O(2^{n/3})$ quantum queries such that $\mathbf{Adv}^{\mathsf{qPRF}}_{\mathsf{TNT}[f_0,f_1,f_2]}(\mathcal{A}) \geq \frac{2}{125}$.*

Proof. By Proposition 9, there exists a quantum adversary $\mathcal{A}$ that makes $O(2^{n/3})$ quantum queries such that $\mathbf{Adv}^{\mathsf{qPRF}}_{\mathsf{TNT}[f_0,f_1,f_2]}(\mathcal{A}) = \mathbf{Adv}^{\mathsf{qPRF}}_{f_2\circ f_1\circ f_0}(\mathcal{D}_{N_0}) \geq \frac{2}{125}$.

B.3 Quantum Cross-Road Distinguisher on TNT

Let Q be either TNT or a tweakable random permutation. For four pairs $(M^i, T^j) \in \{0,1\}^{2n}$, where $i \in \{0,1\}$ and $j \in \{0,1,2,3\}$, the events $Q\left(M^0,T^0\right) = Q\left(M^1,T^1\right)$ and $Q\left(M^0,T^2\right) = Q\left(M^1,T^3\right)$ are independent if Q is a tweakable random permutation, but dependent otherwise.

Proposition 10. *Let $M^i, T^j \in \{0,1\}^n, i \in \{0,1\}, j \in \{0,1,2,3\}$, let Q be TNT. Assume that $\left(M^0,T^0\right)$ and $\left(M^1,T^1\right)$ satisfy $Q\left(M^0,T^0\right) = Q\left(M^1,T^1\right)$. There exist T^2 and $T^3 = T^2 \oplus (T^0 \oplus T^1)$ that satisfy $Q\left(M^0,T^2\right) = Q\left(M^1,T^3\right)$.*

Proof. If Q is TNT, then for two randomly chosen pairs (M^0, T^0) and (M^1, T^1), there are two other pairs (M^0, T^2) and (M^1, T^3) that satisfy

$$\begin{aligned} T^2 &= \pi_0(M^0) \oplus \pi_0(M^1) \oplus T^1, \\ T^3 &= T^2 \oplus (T^0 \oplus T^1) = \pi_0(M^0) \oplus \pi_0(M^1) \oplus T^0. \end{aligned}$$

Then we have

$$\begin{aligned} & Q(M^0, T^0) = Q(M^1, T^1) \\ \iff\ & \pi_2(T^0 \oplus \pi_1(T^0 \oplus \pi_0(M^0))) = \pi_2(T^1 \oplus \pi_1(T^1 \oplus \pi_0(M^1))) \\ \iff\ & T^0 \oplus \pi_1(T^0 \oplus \pi_0(M^0)) = T^1 \oplus \pi_1(T^1 \oplus \pi_0(M^1)) \\ \iff\ & T^0 \oplus \pi_1(T^3 \oplus \pi_0(M^0) \oplus \pi_0(M^1) \oplus \pi_0(M^0)) = \\ & T^1 \oplus \pi_1(T^2 \oplus \pi_0(M^0) \oplus \pi_0(M^1) \oplus \pi_0(M^1)) \\ \iff\ & T^3 \oplus \pi_1((T^3 \oplus \pi_0(M^1))) = T^2 \oplus \pi_1(T^2 \oplus \pi_0(M^0)) \\ \iff\ & \pi_2(T^3 \oplus \pi_1((T^3 \oplus \pi_0(M^1)))) = \pi_2(T^2 \oplus \pi_1(T^2 \oplus \pi_0(M^0))) \\ \iff\ & Q\left(M^1, T^3\right) = Q\left(M^0, T^2\right). \end{aligned}$$

Thus, $Q(M^0, T^2) = Q(M^1, T^3)$ if and only if $Q(M^0, T^0) = Q(M^1, T^1)$. According to the relation above, for randomly chosen pairs (M^0, T^0) and (M^1, T^1), if $\left(M^0, T^0\right)$ and $\left(M^1, T^1\right)$ satisfy $Q\left(M^0, T^0\right) = Q\left(M^1, T^1\right)$, then for $T^2 = \pi_0(M^0) \oplus \pi_0(M^1) \oplus T^1$ and $T^3 = T^2 \oplus (T^0 \oplus T^1)$, the equality $Q\left(M^0, T^2\right) = Q\left(M^1, T^3\right)$ also holds.

The quantum algorithm for distinguishing TNT from a tweakable random permutation is based on Proposition 10.

Theorem 4 (Quantum crossroad distinguisher). *Let $M^0, M^1, T^0, T^1, x \in \{0,1\}^n$, let Q be either TNT or a tweakable random permutation. Assume that we have $\left(M^0, T^0\right)$ and $\left(M^1, T^1\right)$ such that $Q\left(M^0, T^0\right) = Q\left(M^1, T^1\right)$. Let*

$$f(x) = \begin{cases} 1, & \text{if } Q\left(M^0, x\right) = Q\left(M^1, x \oplus T^0 \oplus T^1\right) \text{ and } x \neq T^0; \\ 0, & \text{otherwise}. \end{cases}$$

Let $\mathcal{A}$ be a quantum algorithm such that $\mathcal{A}|0\rangle = \sqrt{p}\,|\psi_G\rangle + \sqrt{1-p}\,|\psi_B\rangle$, where G is the kernel of f and B is the support of f. We run QAA on f for $t = \lfloor \frac{\pi}{8} \times 2^{\frac{n}{2}} \rfloor$ iterations and then measure the state. If Q is TNT, then the probability of obtaining a good result is at least $0.8 - 2^{-\frac{n-5}{2}}$. If Q is a tweakable random permutation, then the probability of obtaining a good result is at most $\frac{1}{2}$.

Proof. If Q is a tweakable random permutation, for a random x the probability that $f(x) = 1$ is 2^{-n}. Thus, for QAA on a tweakable random permutation, $\theta = 2^{-\frac{n}{2}}$, and

$$\begin{aligned} |\psi_t\rangle &= \sin((2t+1)\theta)\,|\psi_G\rangle + \cos((2t+1)\theta)\,|\psi_B\rangle \\ &= \sin(\lfloor \frac{\pi}{4} \times 2^{\frac{n}{2}} - 1 \rfloor \times 2^{-\frac{n}{2}})\,|\psi_G\rangle + \cos(\lfloor \frac{\pi}{4} \times 2^{\frac{n}{2}} - 1 \rfloor \times 2^{-\frac{n}{2}})\,|\psi_B\rangle. \end{aligned}$$

where $\sin(\lfloor \frac{\pi}{4} \times 2^{\frac{n}{2}} - 1 \rfloor \times 2^{-\frac{n}{2}}) \leq \sin(\frac{\pi}{4} \times 2^{\frac{n}{2}} \times 2^{-\frac{n}{2}}) = \sin(\frac{\pi}{4}) = \frac{\sqrt{2}}{2}$. Thus, the probability of obtaining a good result is at most $\frac{1}{2}$.

If Q is TNT, then:

1. If $x = \pi_1(M^0) \oplus \pi_1(M^1) \oplus T^1$, then $Q\left(M^0, x\right) = Q\left(M^1, x \oplus T^0 \oplus T^1\right)$;
2. If $x \neq \pi_1(M^0) \oplus \pi_1(M^1) \oplus T^1$, the probability that $f(x) = 1$ is 2^{-n}.

As a result, for a random x the probability that $f(x) = 1$ is 2^{-n+1}. Therefore $\theta = 2^{-\frac{n-1}{2}}$, and

$$\begin{aligned}|\psi_t\rangle &= \sin((2t+1)\theta)\,|\psi_G\rangle + \cos((2t+1)\theta)\,|\psi_B\rangle \\ &= \sin(\lfloor \frac{\pi}{4} \times 2^{\frac{n}{2}} - 1 \rfloor \times 2^{-\frac{n-1}{2}})\,|\psi_G\rangle + \cos(\lfloor \frac{\pi}{4} \times 2^{\frac{n}{2}} - 1 \rfloor \times 2^{-\frac{n-1}{2}})\,|\psi_B\rangle\,.\end{aligned}$$

Since the derivative of $\sin x$ is $(\sin x)' = \cos x$, and $0 < (\sin x)' < 1$ for $0 < x < \frac{\pi}{2}$, we have

$$\begin{aligned}&\sin^2(\lfloor \frac{\pi}{4} \times 2^{\frac{n}{2}} - 1 \rfloor \times 2^{-\frac{n-1}{2}}) \geq \sin^2((\frac{\pi}{4} \times 2^{\frac{n}{2}} - 2) \times 2^{-\frac{n-1}{2}}) \\ &\geq (\sin(\frac{\pi}{4} \times \sqrt{2}) - 2^{-\frac{n-3}{2}})^2 > \sin^2(\frac{\pi}{4} \times \sqrt{2}) - 2 \times 2^{-\frac{n-3}{2}} > 0.8 - 2^{-\frac{n-5}{2}}.\end{aligned}$$

Thus, the probability of obtaining a good result is at least $0.8 - 2^{-\frac{n-5}{2}}$.

B.4 Quantum Chosen Ciphertext Attack on TNT

For fixed M and Δ, we define $F(T) = Q^{-1}(Q\,(M,T)\,, T \oplus \Delta)$, where Q is either TNT or a tweakable random permutation, and Q^{-1} is its inverse. Proposition 11, based on the findings of Jha et al. [JKNS23], shows that when Q is TNT, the likelihood of encountering collisions is twice as high as when Q is a tweakable random permutation.

Proposition 11 (From [JKNS23]). *T_0, T_1 are chosen uniformly at random from $\mathbb{F}_2^n$. For fixed M and Δ, we have*

$$\Pr\left[F(T_0) = F(T_1)\right] = \begin{cases} 2^{-n}, & \text{if } Q \text{ is a tweakable random permutation}, \\ 2^{-n+1}, & \text{if } Q \text{ is } \mathsf{TNT}. \end{cases}$$

Theorem 5 (Quantum chosen ciphertext attack). *For fixed M and Δ, we define*

$$F(T) = Q^{-1}(Q\,(M,T)\,, T \oplus \Delta),$$

where Q is either TNT or a tweakable random permutation, and Q^{-1} is its inverse. We pick a subset $S \subset \mathbb{F}_2^n$ of cardinality $2^{\frac{n}{3}}$. Construct a table L of size $2^{\frac{n}{3}}$ where each item in L holds a distinct pair $(x, F(x))$ with $x \in S$. Let

$$f(x) = \begin{cases} 1, & \text{if there exists a pair } (x_0, F(x_0)) \text{ in } L \text{ that } F(x) = F(x_0); \\ 0, & \text{otherwise}. \end{cases}$$

Let $\mathcal{A}$ be a quantum algorithm such that $\mathcal{A}|0\rangle = \sqrt{p}\,|\psi_G\rangle + \sqrt{1-p}\,|\psi_B\rangle$, where G is the kernel of f and B is the support of f. We run QAA on f for $t = \lfloor \frac{\pi}{8} \times 2^{\frac{n}{3}} \rfloor$ iterations and then measure the state. If Q is TNT, then the probability of obtaining a good result is at least $0.8 - 2^{-\frac{n}{3}+\frac{5}{2}}$. If Q is a tweakable random permutation, then the probability of obtaining a good result is at most $\frac{1}{2}$.

Proof. If Q is a tweakable random permutation, for a random x the probability that $f(x) = 1$ is $2^{-n} \times 2^{\frac{n}{3}} = 2^{-\frac{2n}{3}}$. Thus, for QAA on a tweakable random permutation, $\theta = 2^{-\frac{n}{3}}$, and

$$\begin{aligned}|\psi_t\rangle &= \sin((2t+1)\theta)\,|\psi_G\rangle + \cos((2t+1)\theta)\,|\psi_B\rangle \\ &= \sin(\lfloor \frac{\pi}{4} \times 2^{\frac{n}{3}} - 1 \rfloor \times 2^{-\frac{n}{3}})\,|\psi_G\rangle + \cos(\lfloor \frac{\pi}{4} \times 2^{\frac{n}{3}} - 1 \rfloor \times 2^{-\frac{n}{3}})\,|\psi_B\rangle\,,\end{aligned}$$

where $\sin(\lfloor \frac{\pi}{4} \times 2^{\frac{n}{3}} - 1 \rfloor \times 2^{-\frac{n}{3}}) \leq \sin(\frac{\pi}{4} \times 2^{\frac{n}{3}} \times 2^{-\frac{n}{3}}) = \sin(\frac{\pi}{4}) = \frac{\sqrt{2}}{2}$. Thus, the probability of obtaining a good result is at most $\frac{1}{2}$.

If Q is TNT, for a random x the probability that $f(x) = 1$ is $2^{-n+1} \times 2^{\frac{n}{3}} = 2^{-\frac{2n}{3}+1}$. Therefore $\theta = 2^{-\frac{n}{3}+\frac{1}{2}}$, and

$$\begin{aligned}|\psi_t\rangle &= \sin((2t+1)\theta)\,|\psi_G\rangle + \cos((2t+1)\theta)\,|\psi_B\rangle \\ &= \sin(\lfloor \frac{\pi}{4} \times 2^{\frac{n}{3}} - 1 \rfloor \times 2^{-\frac{n}{3}+\frac{1}{2}})\,|\psi_G\rangle + \cos(\lfloor \frac{\pi}{4} \times 2^{\frac{n}{3}} - 1 \rfloor \times 2^{-\frac{n}{3}+\frac{1}{2}})\,|\psi_B\rangle\,.\end{aligned}$$

Since the derivative of $\sin x$ is $(\sin x)' = \cos x$, and $0 < (\sin x)' < 1$ for $0 < x < \frac{\pi}{2}$, we have

$$\begin{aligned}&\sin^2(\lfloor \frac{\pi}{4} \times 2^{\frac{n}{3}} - 1 \rfloor \times 2^{-\frac{n}{3}+\frac{1}{2}}) \geq \sin^2((\frac{\pi}{4} \times 2^{\frac{n}{3}} - 2) \times 2^{-\frac{n}{3}+\frac{1}{2}}) \\ &\geq (\sin(\frac{\pi}{4} \times \sqrt{2}) - 2^{-\frac{n}{3}+\frac{3}{2}})^2 > \sin^2(\frac{\pi}{4} \times \sqrt{2}) - 2 \times 2^{-\frac{n}{3}+\frac{3}{2}} > 0.8 - 2^{-\frac{n}{3}+\frac{5}{2}}.\end{aligned}$$

Thus, the probability of obtaining a good result is at least $0.8 - 2^{-\frac{n}{3}+\frac{5}{2}}$.