On the Security of KOS

Benjamin E. Diamond*
Coinbase
benediamond@gmail.com

Abstract

We study the security of the random oblivious transfer extension protocol of Keller, Orsini, and Scholl (CRYPTO '15), whose security proof was recently invalidated by Roy (CRYPTO '22). We show that KOS is asymptotically secure. Our proof involves a subtle analysis of the protocol's "correlation check", and introduces several new techniques. We also study the protocol's concrete security. We establish concrete security for security parameter values on the order of 30,000. We present evidence that a stronger result than ours—if possible—is likely to require radically new ideas.

1 Introduction

The oblivious transfer extension protocol of Keller, Orsini and Scholl [KOS15, Fig. 7] (henceforth "KOS") is widely known and used. Key to that protocol is a certain "correlation check", in which a number of extension OTs are "sacrificed" in a linear combination. This check is very difficult to analyze. In recent work, Roy [Roy22, § 4.1] disproves a key lemma [KOS15, Lem. 1], upon which KOS's security analysis relies. Roy's work invalidates the security proof [KOS15, Thm. 1], as originally written.

In a recent update to their work, Keller, Orsini and Scholl propose an adjusted variant of their protocol [KOS22, Fig. 10]; essentially, they suggest a special case of Roy's construction. Though the efficiency of the updated protocol is comparable to the original, it is more complex, and uses different ideas. Indeed, we note that the analysis of [Roy22] is very theoretically involved. It is of interest to prove the security of KOS, as originally written; this open problem is noted explicitly by Roy [Roy22, § 1.1], for example.

We show that, asymptotically, KOS is secure. Our proof introduces a new simulation strategy, based on the *majority* function. We moreover introduce a certain numerical metric, which captures the extent of the corrupt receiver's compliance with the protocol. We show that—as this degree of compliance varies—the receiver must choose between facing negligible odds in the correlation check, on the one hand, and handing the distinguisher a negligible advantage, on the other. Our proof's key step has a coding-theoretic flavor; we show that a binary matrix with sufficiently many random columns is unlikely to reside *near* the matrix representation of any field element (in the space of matrices, where *distance* is measured in rank).

We also extract effective bounds from our proof. We show that, in order to achieve statistical security of 2^{-40} against an adversary making up to 2^{80} hash evaluations, the security parameter $\kappa := 27{,}385$ suffices (see Example 3.12). More abstractly, we show that KOS, instantiated with security parameter κ, withstands an attacker making up to $\frac{1}{2} \cdot \sqrt{\kappa} \cdot 2^{\frac{1}{4} \cdot \sqrt{\kappa}}$ hash evaluations with statistical security $2^{-\frac{1}{4} \cdot \sqrt{\kappa}}$ (see Corollary 3.14).

Obviously, this sort of κ results in a barely-practical protocol. On the other hand, we give evidence that this limitation might be intrinsic. As it turns out, our proof applies equally well to the security of Patra, Sarkar and Suresh [PSS17] (henceforth "PSS"), another protocol attacked by Roy [Roy22, § 4.1]. (Indeed, our proof invokes *only* properties of KOS which are shared by PSS; we discuss this fact further below.) Interestingly, our lower-bound (see Corollary 3.14) tightly matches—up to the factor of $\frac{1}{4}$ present in both exponents—the upper-bound achieved by Roy [Roy22, § 4.1] on PSS. Our proof thus definitively settles the question of PSS's security (up to the constants). It also shows that a sharper analysis of KOS—if possible at all—would have to rely on features of KOS's correlation check more delicate than those our proof considers. We give full details below.

*I would like to sincerely thank a handful of anonymous referees for extremely valuable feedback.

We briefly recall the details of KOS. Informally, the correlation check (see [KOS15, Fig. 7]) controls the row-vectors $(\mathbf{x}_i)_{i=0}^{l'-1}$ the receiver submits to the *correlated OT with errors* hybrid functionality $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l'}$. If the receiver is honest, then each row $\mathbf{x}_i \in \mathbb{F}_2^\kappa$ is necessarily "monochromatic" (in the sense that its components are identical); if the receiver's rows $\mathbf{x}_i$ are *not* monochromatic, then the corrupt receiver may facilitate the distinguisher's learning certain bits of the sender's correlation vector $\Delta \in \mathbb{F}_2^\kappa$, by means of brute-force queries to the random oracle. The correlation check, at first glance, is quite natural. It prescribes that the parties jointly sample random elements $(\chi_i)_{i=0}^{l'-1}$ from $\mathbb{F}_{2^\kappa}$, using a coin-flipping functionality, that they subject their intermediate values—i.e., those they obtained from $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l'}$—to a linear combination using these coefficients, and finally that they exchange the results. Specifically, the sender and receiver, having received $(\mathbf{q}_i)_{i=0}^{l'-1}$ and $(\mathbf{t}_i)_{i=0}^{l'-1}$, respectively, from $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l'}$, compute $q := \sum_{i=0}^{l'-1} \chi_i \cdot \mathbf{q}_i$ and $t := \sum_{i=0}^{l'-1} \chi_i \cdot \mathbf{t}_i$, respectively; the (honest) receiver moreover computes $x := \sum_{i=0}^{l'-1} \chi_i \cdot x_i$, where $(x_i)_{i=0}^{l'-1}$ is its choice vector. Finally, the receiver sends x and t to the sender, who computes $q \stackrel{?}{=} t + x \cdot \Delta$. All multiplications here take place in the binary field $\mathbb{F}_{2^\kappa}$.

Informally, the correlation check controls whether the *individual* equalities $\mathbf{q}_i \stackrel{?}{=} \mathbf{t}_i + x_i \cdot \Delta$ hold, for each $i \in \{0, \ldots, l'-1\}$, or—equivalently—whether the vector $(\mathbf{q}_i + \mathbf{t}_i + x_i \cdot \Delta)_{i=0}^{l'-1} \in \mathbb{F}_{2^\kappa}^{l'}$ is the zero vector. In actuality, however, the correlation check merely checks whether this latter vector resides within the *random hyperplane* in $\mathbb{F}_{2^\kappa}^{l'}$ given by the coefficients $(\chi_i)_{i=0}^{l'-1}$. The difficulty is that the corrupt receiver sees these coefficients—and the resulting hyperplane—*before* sending its combination results x and t; as a result, the receiver could conceivably could choose x and t such a way that the vector $(\mathbf{q}_i + \mathbf{t}_i + x_i \cdot \Delta)_{i=0}^{l'-1}$—though *nonzero*—nonetheless resides within this hyperplane, and thus causes the check to pass. This is precisely the subtlety overlooked by [KOS15, Thm. 1]; we refer to [Roy22, § 4.1] for discussion.

We now sketch the technical details of our proof (see also Theorem 3.1 below). Our treatment of the corrupt sender is similar to that of [KOS15, Thm. 1] (though we supply certain details which were omitted from that proof). Our treatment of the corrupt receiver—the more difficult case—relies on a new simulation strategy for that case, as well as on a careful analysis of the adversary's and distinguisher's success conditions. We first introduce the *majority rule* extraction strategy $x_i := \mathrm{MAJ}_\kappa(\mathbf{x}_i)$; here, $\mathrm{MAJ}_\kappa : \{0,1\}^\kappa \to \{0,1\}$ is the standard majority function on κ bits. (We note that [KOS15, Thm. 1]'s simulation strategy no longer makes sense, in that its simulator's very description assumes the truth of [KOS15, Lem. 1].)

We then prove our simulator's security in the following way. (A graphical depiction of our proof strategy is also given in Figure 1 below.) We develop a numerical characterization of the *extent* of $\mathcal{A}$'s matrix's monochromaticity. This metric—which we call the matrix's *modesty*—ranges throughout $m \in \{1, \ldots, \kappa\}$; an honest receiver necessarily has modesty κ. Informally, m measures the difficulty of "assembling" a majority vector $\overline{\mathbf{e}_i}$ by means of a sequence of minority vectors $\mathbf{e}_i$. Specifically, we say that $(\mathbf{x}_i)_{i=0}^{l'-1}$ has modesty m if there exists some sequence of minority vectors $\mathbf{e}_i$—each element within which introduces *at most* m new bits—which moreover reaches within distance m of some majority vector $\overline{\mathbf{e}_i}$, and if moreover $m \in \{1, \ldots, \kappa\}$ is the smallest integer with this property (see also Definition 3.3 for details).

We show that as $m \in \{1, \ldots, \kappa\}$ varies, $\mathcal{A}$ smoothly trades off between two different bad outcomes. On the one hand, if $\mathcal{A}$'s modesty is low—that is, if $\mathcal{A}$ cheats brazenly—then $\mathcal{A}$'s probability of passing the correlation check becomes low. Indeed, we note that—up to a uniform resampling of the random combination coefficients $(\chi_i)_{i=0}^{l'-1}$ used in the check—we may freely assume that the matrix $(\mathbf{e}_i)_{i=0}^{l'-1}$ is in reduced row-echelon form; we further note that, as m decreases, this reduced matrix accumulates pivots. These pivots impose independently random $\mathbb{F}_2$-linear conditions on the unknown vector Δ, and make the correlation check harder to pass. We reduce $\mathcal{A}$'s success to a coding-theoretic condition on the space of binary matrices, and upper-bound its probability of passing using a union bound (see Proposition 3.5).

On the other hand, we show that as m grows—that is, as $\mathcal{A}$ becomes *more* compliant—$\mathcal{A}$ begins producing transcripts which make the presence of our simulation harder to detect. Indeed, any given distinguisher may learn bits of the hidden choice vector Δ *only* by means of brute-force queries of the form $H(i \parallel \mathbf{t}_i + \mathbf{e}_i * \Delta)$, where, by induction, $\mathbf{e}_i$ introduces few new bit positions not already learned. On the other hand, the distinguisher may successfully distinguish the real and ideal distributions only if it manages to query $H(i \parallel \mathbf{t}_i + \overline{\mathbf{e}_i} * \Delta)$, where here $\overline{\mathbf{e}_i}$ is some *majority* vector. Effectively, m controls the size of the the minimal-length "stretch" of bits which the distinguisher must brute-force, if it is to succeed (see Proposition 3.8).

By combining these two cases, we establish the result (see Theorem 3.1).

2 Background and Notation

We identify $\{0,1\} \cong \mathbb{F}_2$ as *sets*. We occasionally identify *vectors* in $\{0,1\}^\kappa \cong \mathbb{F}_2^\kappa$ with *subsets* of $\{0,\ldots,\kappa-1\}$, in the standard way; that is, for each vector $\mathbf{d} \in \{0,1\}^\kappa$, corresponding to the map $\widehat{\mathbf{d}} : \{0,\ldots,\kappa-1\} \to \{0,1\}$, say, we identify $\mathbf{d}$ with the subset $\widehat{\mathbf{d}}^{-1}(1) \subset \{0,\ldots,\kappa-1\}$ (i.e., with the set of components at which $\mathbf{d}$ is 1). We use the symbol $*$ to denote bitwise AND in $\mathbb{F}_2^\kappa$, and write w for Hamming weight. We use the symbol $\setminus$ to denote set subtraction. We fix a field structure on $\mathbb{F}_{2^\kappa}$—that is, an irreducible polynomial of degree κ in $\mathbb{F}_2[X]$—and identify $\mathbb{F}_{2^\kappa}$ with the $\mathbb{F}_2$-vectorspace $\mathbb{F}_2^\kappa$, by means of the basis $(1, X, \ldots, X^{\kappa-1})$. We write $\cdot$ for field multiplication. In what follows, we make extensive use of linear and affine-linear algebra over $\mathbb{F}_2$, without further comment; for this, we suggest the reference Cohn [Coh82, §5].

Following [KOS15, § 2], we write κ for a security parameter. We write λ and s for *desired* levels of computational and statistical security, respectively. We write $(\mathbf{x}_i)_{i=0}^{l'-1}$ for the rows of an $l' \times \kappa$ matrix. We write $\overline{\mathbf{x}_i}$ for the bitwise complement of a row-vector $\mathbf{x}_i \in \mathbb{F}_2^\kappa$, and $\overline{x_i}$ for the complement of a bit $x_i \in \mathbb{F}_2$. We write $\mathrm{MAJ}_\kappa : \{0,1\}^\kappa \to \{0,1\}$ for the majority function on κ bits, defined specifically by $\mathbf{x}_i \mapsto w(\mathbf{x}_i) \geq \frac{\kappa}{2}$.

2.1 Secure computation

Given two probability distributions $\mathcal{Y}_0$ and $\mathcal{Y}_1$ on $\{0,1\}^\kappa$, the *statistical distance* between $\mathcal{Y}_0$ and $\mathcal{Y}_1$ is defined to be $\frac{1}{2} \cdot \sum_{\mathbf{y} \in \{0,1\}^\kappa} |\Pr[\mathcal{Y}_0 = \mathbf{y}] - \Pr[\mathcal{Y}_1 = \mathbf{y}]|$. We say that two distribution ensembles $\{\mathcal{Y}_0(a,\kappa)\}_{a \in \{0,1\}^*, \kappa \in \mathbb{N}}$ and $\{\mathcal{Y}_1(a,\kappa)\}_{a \in \{0,1\}^*, \kappa \in \mathbb{N}}$ are *statistically indistinguishable* if, for each $a \in \{0,1\}^*$ and $\kappa \in \mathbb{N}$, the statistical distance between $\mathcal{Y}_0(a,\kappa)$ and $\mathcal{Y}_1(a,\kappa)$ is at most $\mu(\kappa)$, where μ is a fixed, negligible function of κ. We say that two distribution ensembles $\{\mathcal{Y}_0(a,\kappa)\}_{a \in \{0,1\}^*, \kappa \in \mathbb{N}}$ and $\{\mathcal{Y}_1(a,\kappa)\}_{a \in \{0,1\}^*, \kappa \in \mathbb{N}}$ are *computationally indistinguishable* if, for each probabilistic, polynomial-time distinguisher D (with outputs in $\{0,1\}$), the distributions ensembles $\{D(\mathcal{Y}_0(a,\kappa))\}_{a \in \{0,1\}^*, \kappa \in \mathbb{N}}$ and $\{D(\mathcal{Y}_1(a,\kappa))\}_{a \in \{0,1\}^*, \kappa \in \mathbb{N}}$ are statistically indistinguishable.

We record the definition of maliciously secure two-party computation, following Lindell [Lin17, § 6.6.2].

Definition 2.1. For each functionality $\mathcal{F}$, a protocol Π, real-world adversary $\mathcal{A}$, simulator $\mathcal{S}$, and corrupt party $C \in \{0,1\}$, we have the distributions:

- $\mathsf{Real}_{\Pi,\mathcal{A},C}((\mathbf{x}_0,\mathbf{x}_1),\kappa)$: Run Π with security parameter κ, where the honest party P_{1-C} uses the input $\mathbf{x}_{1-C}$, and $\mathcal{A}$ controls the messages of the corrupt party. Return the outputs of $\mathcal{A}$ and P_{1-C}.
- $\mathsf{Ideal}_{\mathcal{F},\mathcal{S},C}((\mathbf{x}_0,\mathbf{x}_1),\kappa)$: Run $S(1^\kappa, C, \mathbf{x}_C)$ until it outputs a value $\mathbf{x}'_C$, or else outputs (`abort`) to $\mathcal{F}$, who halts. Give $\mathbf{x}_{1-C}$ and $\mathbf{x}'_C$ to $\mathcal{F}$, and obtain outputs (v_0, v_1). Give v_C to $\mathcal{S}$; if $\mathcal{S}$ outputs (`abort`), then $\mathcal{F}$ outputs (`abort`) to P_{1-C}; otherwise, $\mathcal{F}$ gives P_{1-C} v_{1-C}. Return the outputs of $\mathcal{S}$ and P_{1-C}.

We say that Π *securely computes* $\mathcal{F}$ *in the presence of one static malicious corruption with abort*, or that Π *securely computes* $\mathcal{F}$, if, for each corrupt party $C \in \{0,1\}$ and each probabilistic polynomial-time adversary $\mathcal{A}$ corrupting P_C, there is a probabilistic expected polynomial-time simulator $\mathcal{S}$ corrupting P_C in the ideal world such that the distributions $\{\mathsf{Real}_{\Pi,\mathcal{A},C}((\mathbf{x}_0,\mathbf{x}_1),\kappa)\}_{(\mathbf{x}_0,\mathbf{x}_1),\kappa}$ and $\{\mathsf{Ideal}_{\mathcal{F},\mathcal{S},C}((\mathbf{x}_0,\mathbf{x}_1),\kappa)\}_{(\mathbf{x}_0,\mathbf{x}_1),\kappa}$ are computationally indistinguishable, where $\mathbf{x}_0$ and $\mathbf{x}_1$ are required throughout to have equal lengths.

2.2 Oblivious transfer

We recall background material on oblivious transfer, following [KOS15].

FUNCTIONALITY 2.2 ($\mathcal{F}_{\mathsf{Rand}}^\kappa$—coin-flipping functionality [KOS15, Fig. 5])**.**
The security parameter κ and players S and R are fixed.

- Upon receiving $(\mathtt{random}, i)$ from both players, $\mathcal{F}_{\mathsf{Rand}}^\kappa$ samples $\chi_i \leftarrow \mathbb{F}_2^\kappa$, and outputs $(\mathtt{random}, i, \chi_i)$ to both players.

FUNCTIONALITY 2.3 ($\mathcal{F}_{\mathsf{COTe}}^{\kappa,l}$—correlated OT with errors [KOS15, Fig. 2])**.**
The security parameter κ, the number l of resulting OTs, and players S and R are fixed.

- Upon receiving $(\texttt{initialize}, \Delta)$ from S, where $\Delta \in \mathbb{F}_2^\kappa$, $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l}$ stores Δ.
- If both parties are honest, R submits $\left(\texttt{input}, (\mathbf{x}_i)_{i=0}^{l-1}\right)$ to $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l}$, which, for each $i \in \{0, \ldots, l-1\}$, samples $\mathbf{t}_i \leftarrow \mathbb{F}_2^\kappa$ randomly and computes $\mathbf{q}_i := \mathbf{t}_i + \mathbf{x}_i * \Delta$.
- If R is corrupt, R submits $\left(\texttt{input}, (\mathbf{x}_i)_{i=0}^{l-1}, (\mathbf{t}_i)_{i=0}^{l-1}\right)$ to $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l}$, which computes $(\mathbf{q}_i)_{i=0}^{l-1}$ identically.
- If S is corrupt, S submits $\left(\texttt{input}, (\mathbf{q}_i)_{i=0}^{l-1}\right)$ to $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l}$, which, for each $i \in \{0, \ldots, l-1\}$, sets $\mathbf{t}_i := \mathbf{q}_i + \mathbf{x}_i * \Delta$.
- In each case, $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l}$ outputs $\left(\texttt{output}, (\mathbf{t}_i)_{i=0}^{l-1}\right)$ to R and $\left(\texttt{output}, (\mathbf{q}_i)_{i=0}^{l-1}\right)$ to S.

We note that $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l}$ can be securely instantiated by the protocol of [KOS15, Fig. 3].

Remark 2.4. We slightly alter the treatment of [KOS15, Fig. 2], in that we permit the corrupt sender S to choose its values $(\mathbf{q}_i)_{i=0}^{l-1}$. This privilege appears necessary for the secure instantiation of $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l}$ (in the $\mathcal{F}_{\mathsf{OT}}^{\kappa}$-hybrid model) to go through; its omission appears to have been an oversight on the part of [KOS15].

We moreover recall the *random OT* functionality:

FUNCTIONALITY 2.5 ($\mathcal{F}_{\mathsf{ROT}}^{\kappa,l}$—random OT functionality [KOS15, Fig. 6])**.**
The security parameter κ, the number l of resulting OTs, and players S and R are fixed.

- If both parties are honest, R submits $\left(\texttt{input}, (x_i)_{i=0}^{l-1}\right)$ to $\mathcal{F}_{\mathsf{ROT}}^{\kappa,l}$, which, for each $i \in \{0, \ldots, l-1\}$, samples $(\mathbf{v}_{i,0}, \mathbf{v}_{i,1}) \leftarrow \{0,1\}^\kappa \times \{0,1\}^\kappa$.
- If R is corrupt, R submits $\left(\texttt{input}, (x_i)_{i=0}^{l-1}, (\mathbf{v}_{i,x_i})_{i=0}^{l-1}\right)$ to $\mathcal{F}_{\mathsf{ROT}}^{\kappa,l}$, which, for each $i \in \{0, \ldots, l-1\}$, samples $\mathbf{v}_{i,\overline{x_i}} \leftarrow \{0,1\}^\kappa$.
- If S is corrupt, then S submits $\left(\texttt{input}, (\mathbf{v}_{i,0}, \mathbf{v}_{i,1})_{i=0}^{l-1}\right)$ to $\mathcal{F}_{\mathsf{ROT}}^{\kappa,l}$.
- In each case, $\mathcal{F}_{\mathsf{ROT}}^{\kappa,l}$ outputs $\left(\texttt{output}, (\mathbf{v}_{i,0}, \mathbf{v}_{i,1})_{i=0}^{l-1}\right)$ to S and $\left(\texttt{output}, (\mathbf{v}_{i,x_i})_{i=0}^{l-1}\right)$ to R.

Remark 2.6. We likewise give the adversary slightly more power than does [KOS15, Fig. 6], in that we let the corrupt receiver choose $(\mathbf{v}_{i,x_i})_{i=0}^{l-1}$. This concession appears necessary; indeed—aside from its other issues—the simulator [KOS15, Fig. 8] programs $H\left(i \parallel \mathbf{q}_i + x_i \cdot \Delta\right) := \mathbf{v}_{i,x_i}$ only *after* receiving $\mathbf{t}_i$ from $\mathcal{A}$. $\mathcal{A}$ can easily arrange to make this query before this programming step occurs, thereby breaking the simulation.

For self-containedness, we finally recall the full protocol for $\mathcal{F}_{\mathsf{ROT}}^{\kappa,l}$, exactly as in [KOS15, Fig. 7].

PROTOCOL 2.7 ($\Pi_{\mathsf{ROT}}^{\kappa,l}$—random OT protocol [KOS15, Fig. 7])**.**
The parameters κ and l, and players S and R, are fixed. R has input bits $(x_0, \ldots, x_{l-1})$.

- The parties write $l' := l + \kappa + s$. S samples $\Delta \leftarrow \mathbb{F}_2^\kappa$, and sends $(\texttt{intialize}, \Delta)$ to $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l'}$.
- R samples random bits $x_i \leftarrow \mathbb{F}_2$, for $i \in \{l, \ldots, l'-1\}$. For each $i \in \{0, \ldots, l'-1\}$, R constructs the monochromatic vector $\mathbf{x}_i := x_i \cdot (1, \ldots, 1)$. R sends $\left(\texttt{input}, (\mathbf{x}_i)_{i=0}^{l'-1}\right)$ to $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l'}$. S and R receive $\left(\texttt{output}, (\mathbf{q}_i)_{i=0}^{l'-1}\right)$ and $\left(\texttt{output}, (\mathbf{t}_i)_{i=0}^{l'-1}\right)$, respectively, from $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l'}$.

- For each $i \in \{0, \ldots, l'-1\}$, both parties submit $(\texttt{random}, i)$ to $\mathcal{F}_{\mathsf{Rand}}^{\kappa}$, and receive $(\texttt{random}, i, \chi_i)$. R sends S $x := \sum_{i=0}^{l'-1} \chi_i \cdot x_i$ and $t := \sum_{i=0}^{l'-1} \chi_i \cdot \mathbf{t}_i$. S sets $q := \sum_{i=0}^{l'-1} \chi_i \cdot \mathbf{q}_i$, and checks $q \stackrel{?}{=} t + x \cdot \Delta$.
- For each $i \in \{0, \ldots, l-1\}$, R sets $\mathbf{v}_{i,x_i} := H(i \parallel \mathbf{t}_i)$, and outputs $(\mathbf{v}_{i,x_i})_{i=0}^{l-1}$. For each $i \in \{0, \ldots, l-1\}$, S sets $\mathbf{v}_{i,0} := H(i \parallel \mathbf{q}_i)$ and $\mathbf{v}_{i,1} := H(i \parallel \mathbf{q}_i + \Delta)$, and outputs $(\mathbf{v}_{i,0}, \mathbf{v}_{i,1})_{i=0}^{l-1}$.

3 Security proof

We now prove the security of Protocol 2.7.

Theorem 3.1. *In the $\mathcal{F}_{\mathsf{RO}}, \mathcal{F}_{\mathsf{Rand}}^{\kappa}, \mathcal{F}_{\mathsf{COTe}}^{\kappa,l'}$ hybrid model, Protocol 2.7 securely computes Functionality 2.5.*

Proof. We define an appropriate simulator $\mathcal{S}$.

Corrupt sender. We first handle the case in which S is corrupt. Our treatment of this case is similar to that of [KOS15, Thm. 1]. Given a real-world adversary $\mathcal{A}$ corrupting S, $\mathcal{S}$ operates in the following way.

1. $\mathcal{S}$ intercepts $\mathcal{A}$'s messages $(\texttt{initialize}, \Delta)$ and $\left(\texttt{input}, (\mathbf{q}_i)_{i=0}^{l'-1}\right)$ to $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l'}$. For each $i \in \{0, \ldots, l-1\}$, $\mathcal{S}$ computes $\mathbf{v}_{i,0} := H(i \parallel \mathbf{q}_i)$ and $\mathbf{v}_{i,1} := H(i \parallel \mathbf{q}_i + \Delta)$. $\mathcal{S}$ submits $\left(\texttt{input}, (\mathbf{v}_{i,0}, \mathbf{v}_{i,1})_{i=0}^{l-1}\right)$ to $\mathcal{F}_{\mathsf{ROT}}^{\kappa,l}$.
2. $\mathcal{S}$ receives $\left(\texttt{output}, (\mathbf{v}_{i,0}, \mathbf{v}_{i,1})_{i=0}^{l-1}\right)$ from $\mathcal{F}_{\mathsf{ROT}}^{\kappa,l}$, and simulates $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l'}$ sending $\left(\texttt{output}, (\mathbf{q}_i)_{i=0}^{l'-1}\right)$ to $\mathcal{A}$.
3. For each $i \in \{0, \ldots, l'-1\}$, $\mathcal{S}$ intercepts $\mathcal{A}$'s message $(\texttt{random}, i)$ intended for $\mathcal{F}_{\mathsf{Rand}}^{\kappa}$, samples $\chi_i \leftarrow \mathbb{F}_2^{\kappa}$ randomly, and simulates $\mathcal{F}_{\mathsf{Rand}}^{\kappa}$ sending $\mathcal{A}$ $(\texttt{random}, i, \chi_i)$. $\mathcal{S}$ samples $x \leftarrow \mathbb{F}_2^{\kappa}$ randomly, computes $q := \sum_{i=0}^{l'-1} \chi_i \cdot \mathbf{q}_i$, and sets $t := q + x \cdot \Delta$. $\mathcal{S}$ simulates R sending $\mathcal{A}$ t and x.

The perfection of this simulation is self-evident, except perhaps for the distribution of x. For self-containedness, we present a full proof of the relevant lemma, whose proof is omitted from [KOS15, Lem. 2].

Lemma 3.2. *Given a random $\kappa \times (\kappa + s)$ matrix X over $\mathbb{F}_2$, where $s \geq 0$, $\Pr[\mathrm{rank}(X) = \kappa] \geq 1 - 2^{-s}$.*

Proof. For each fixed value $s \geq 0$, the probability that the random matrix X's κ rows are independent is equal to the probability that each of its successive rows resides outside of the linear subspace spanned by its previous rows. This probability is given by product expression below, which we manipulate as follows:

$$\begin{aligned}(1 - 2^{-s-1}) \cdot \cdots \cdot (1 - 2^{-s-\kappa}) &\geq 1 - \left(2^{-s-1} + \cdots + 2^{-s-\kappa}\right) \\ &= 1 - 2^{-s} \cdot \left(2^{-1} + \cdots + 2^{-\kappa}\right) \\ &\geq 1 - 2^{-s}.\end{aligned}$$

The first inequality follows from a simple union bound, which we now explain. The expression $1 - \prod_{i=0}^{\kappa-1}(1 - 2^{-s-1-i})$ gives the probability that a certain product of Bernoulli distributions resides *away* from the origin in $\{0,1\}^{\kappa}$. By the union bound, this probability is bounded from above by the sum of faces $\sum_{i=0}^{\kappa-1} 2^{-s-1-i}$. □

The second summand of the quantity $x = \sum_{i=0}^{l-1} \chi_i \cdot x_i + \sum_{i=l}^{l'-1} \chi_i \cdot x_i$ computed by the receiver can be viewed as the image of $(x_i)_{i=l}^{l'-1} \in \mathbb{F}_2^{\kappa+s}$ under the linear map $\mathbb{F}_2^{\kappa+s} \to \mathbb{F}_2^{\kappa}$ defined by the matrix:

$$\begin{bmatrix} | & & | \\ \chi_l & \cdots & \chi_{l'-1} \\ | & & | \end{bmatrix}.$$

The lemma implies that, with probability at least $1 - 2^s$ over the choice of $(\chi_i)_{i=l}^{l'-1}$, the map induced by this matrix is surjective; it follows that, in the real-world distribution, with overwhelming probability, the image of the uniformly random point $(x_i)_{i=l}^{l'-1} \in \mathbb{F}_2^{\kappa+s}$ under this matrix is itself uniform in $\mathbb{F}_2^{\kappa}$, and so perfectly hides the first term $\sum_{i=0}^{l-1} \chi_i \cdot x_i$. This completes the treatment of the corrupt sender.

Corrupt receiver. We now handle the case in which the receiver R is corrupt. Given a real-world adversary $\mathcal{A}$ corrupting the receiver R, $\mathcal{S}$ operates as follows.

1. $\mathcal{S}$ simulates the existence of $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l'}$, including S's role. $\mathcal{S}$ begins by sampling $\Delta \leftarrow \mathbb{F}_2^\kappa$, as S would.
2. Upon intercepting $\mathcal{A}$'s message $\left(\mathtt{input}, (\mathbf{x}_i)_{i=0}^{l'-1}, (\mathbf{t}_i)_{i=0}^{l'-1}\right)$ intended for $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l'}$, $\mathcal{S}$, for each $i \in \{0,\ldots,l-1\}$, writes $\mathbf{q}_i := \mathbf{t}_i + \mathbf{x}_i * \Delta$, computes $x_i := \mathrm{MAJ}_\kappa(\mathbf{x}_i)$ and sets $\mathbf{v}_{i,x_i} := H\left(i \parallel \mathbf{q}_i + x_i \cdot \Delta\right)$. $\mathcal{S}$ submits $\left(\mathtt{input}, (x_i)_{i=0}^{l-1}, (\mathbf{v}_{i,x_i})_{i=0}^{l-1}\right)$ to $\mathcal{F}_{\mathsf{ROT}}^{\kappa,l}$.
3. $\mathcal{S}$ receives $\left(\mathtt{output}, (\mathbf{v}_{i,x_i})_{i=0}^{l-1}\right)$ from $\mathcal{F}_{\mathsf{ROT}}^{\kappa,l}$, and simulates $\mathcal{F}_{\mathsf{COTe}}^{\kappa,l'}$ returning $\left(\mathtt{output}, (\mathbf{t}_i)_{i=0}^{l'-1}\right)$ to $\mathcal{A}$.
4. For each $i \in \{0,\ldots,l'-1\}$, $\mathcal{S}$ intercepts $\mathcal{A}$'s message $(\mathtt{random}, i)$ intended for $\mathcal{F}_{\mathsf{Rand}}^\kappa$, samples $\chi_i \leftarrow \mathbb{F}_2^\kappa$ randomly, and simulates $\mathcal{F}_{\mathsf{Rand}}^\kappa$ sending $\mathcal{A}$ $(\mathtt{rand}, i, \chi_i)$. Upon receiving x and t from $\mathcal{A}$, $\mathcal{S}$ independently computes $q := \sum_{i=0}^{l'-1} \chi_i \cdot \mathbf{q}_i$, and runs the correlation check $q \stackrel{?}{=} t + x \cdot \Delta$. If the check fails, $\mathcal{S}$ submits $(\mathtt{abort})$ to $\mathcal{F}_{\mathsf{ROT}}^{\kappa,l}$; otherwise, $\mathcal{S}$ proceeds, and $\mathcal{F}_{\mathsf{ROT}}^{\kappa,l}$ releases the output to the ideal honest party S.

We now claim that the resulting real and ideal distributions are computationally indistinguishable. More precisely, these distributions are statistically indistinguishable to any computationally unbounded distinguisher which makes only polynomially many queries to the random oracle. We fix a distinguisher D attacking these distributions. Following [KOS15], for each $i \in \{0,\ldots,l'-1\}$, we write $\mathbf{e}_i := \mathbf{x}_i + x_i \cdot (1,\ldots,1)$, where x_i is as extracted by $\mathcal{S}$ above. We note that necessarily $\mathrm{MAJ}_\kappa(\mathbf{e}_i) = 0$ for each $i \in \{0,\ldots,l'-1\}$. We observe that the strings $\mathbf{q}_i + x_i \cdot \Delta$ and $\mathbf{q}_i + \overline{x_i} \cdot \Delta$ respectively equal $\mathbf{t}_i + \mathbf{e}_i * \Delta$ and $\mathbf{t}_i + \overline{\mathbf{e}_i} * \Delta$; the values $\mathbf{t}_i$ and $\mathbf{e}_i$ are known to the distinguisher, while Δ is not. If the correlation check fails, then the real and ideal distributions are identical. Conditioned on the correlation check succeeding, the simulation is perfect *except* for the fact that, for each $i \in \{0,\ldots,l-1\}$, in the real world, the relation $\mathbf{v}_{i,\overline{x_i}} = H\left(i \parallel \mathbf{t}_i + \overline{\mathbf{e}_i} * \Delta\right)$ holds, whereas, in the ideal world, $\mathbf{v}_{i,\overline{x_i}}$ is independently random.

For notational purposes, given $x \in \mathbb{F}_{2^\kappa}$, we introduce the map $F_x : \mathbb{F}_2^\kappa \to \mathbb{F}_2^\kappa$ defined by:

$$F_x : \Delta \mapsto \sum_{i=0}^{l'-1} \chi_i \cdot (\mathbf{t}_i + \mathbf{x}_i * \Delta) + x \cdot \Delta + t.$$

We note, in light of the equalities $\mathbf{q}_i = \mathbf{t}_i + \mathbf{x}_i * \Delta$, that this map exactly reflects the correlation check run by the sender on its (hidden) correlation vector Δ (i.e., the check passes if and only if $F_x(\Delta) \stackrel{?}{=} 0$). We view all quantities above as fixed constants—known to the distinguisher—*except* for the unknown vector Δ.

Clearly, $F_x : \mathbb{F}_2^\kappa \to \mathbb{F}_2^\kappa$ is an $\mathbb{F}_2$-affine linear map. We argue that we may assume once and for all that $\mathcal{A}$ submits an "honest" value $t = \sum_{i=0}^{l'-1} \chi_i \cdot \mathbf{t}_i$. Indeed, our below arguments depend only on the dimension of the affine subspace $\{\Delta \in \mathbb{F}_2^\kappa \mid F_x(\Delta) = 0\}$, and not on its contents; $\mathcal{A}$'s use of a value $t \neq \sum_{i=0}^{l'-1} \chi_i \cdot \mathbf{t}_i$ has merely the effect of replacing this subspace *either* with an affine-linear subspace of $\mathbb{F}_2^\kappa$ of identical dimension *or* with the empty affine subspace (i.e., depending on whether $t + \sum_{i=0}^{l'-1} \chi_i \cdot \mathbf{t}_i$ resides within the image of $\Delta \mapsto \sum_{i=0}^{l'-1} \chi_i \cdot (\mathbf{x}_i * \Delta) + x \cdot \Delta$ or not). If the subspace is empty, then the correlation check is guaranteed to fail, and the simulation becomes trivially secure. We thus simplify our definition, and write:

$$F_x : \Delta \mapsto \sum_{i=0}^{l'-1} \chi_i \cdot (\mathbf{x}_i * \Delta) + x \cdot \Delta.$$

In other words, we superficially ignore the affine offset; we moreover refer to $\mathrm{rank}(F_x)$ and $\ker(F_x)$ throughout (though, technically speaking, we allow the latter subspace to be a nonempty *affine* subspace).

We denote by r the *minimal* rank achieved across all maps $\{F_x\}_{x \in \mathbb{F}_{2^\kappa}}$, so that:

$$r := \min_{x \in \mathbb{F}_{2^\kappa}} \mathrm{rank}(F_x). \tag{1}$$

In what follows, we view r as a random variable, a function of the randomly sampled coefficients $(\chi_i)_{i=0}^{l'-1}$.

We now pause to sketch the details of our proof. We first formulate a numerical metric—called the *modesty*, a quantity $m \in \{1, \ldots, \kappa\}$—describing the extent to which $\mathcal{A}$'s initial matrix $(\mathbf{x}_i)_{i=0}^{l'-1}$ is monochromatic. We consider the protocol in steps, corresponding, respectively, to $\mathcal{A}$'s choice of $(\mathbf{x}_i)_{i=0}^{l'-1}$ (and hence of modesty), to the random sampling of $(\chi_i)_{i=0}^{l'-1}$ (which causes the minimal rank r to be defined), to $\mathcal{A}$'s actual choice of x, and hence of F_x, to whether $\mathcal{A}$ passes the correlation check, and, finally, to whether the distinguisher succeeds. The resulting structure is depicted in the figure below, which should be understood as a probability tree, in which each edge represents an appropriately conditioned probability.

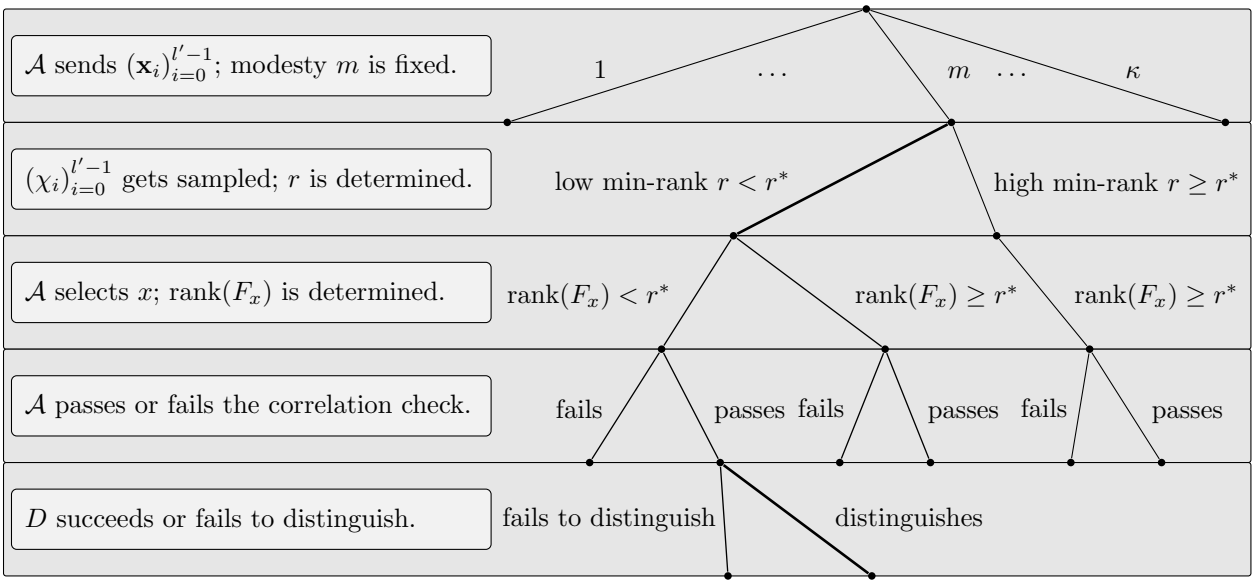

Figure 1: A depiction of the case structure considered by our proof.

We analyze an instance of the above tree for each execution (i.e., for each modesty m). We may immediately ignore those executions in which $\mathcal{A}$ fails the correlation check, since the real and ideal distributions are identical in each such execution. Beyond this, we set a *rank cutoff* r^*. If we choose for our cutoff $r^* = r^*(\kappa)$ a superlogarithmic function of κ, then we may likewise ignore each subtree in which rank$(F_x) \geq r^*$ *and* $\mathcal{A}$ passes the correlation check, since $\mathcal{A}$'s chance of passing the correlation check is $2^{-\text{rank}(F_x)}$, which—in light of our choice of r^*—is negligible whenever rank$(F_x) \geq r^*$. (We discuss our specific choice of r^* below.) We are thus left with one relevant path through the tree. Our proof hinges on analyzing the two edges bolded in the diagram above. Roughly, we show that as $\mathcal{A}$'s matrix's modesty *varies*, *either* the lowermost bolded edge *or* the uppermost bolded edge (or both) must be negligible in κ (these cases happen when $\mathcal{A}$'s matrix *is* and *isn't* modest, respectively). This suffices to demonstrate the result. We now provide details.

Definition 3.3. Given a (truncated) initial matrix $(\mathbf{x}_i)_{i=0}^{l-1}$, we write $(\mathbf{e}_i)_{i=0}^{l-1}$ for the resulting list of minority vectors. We define the *modesty* of $(\mathbf{x}_i)_{i=0}^{l-1}$ to be the *minimal* $m \in \{1, \ldots, \kappa\}$ for which $\textsc{Modest}(m) \stackrel{?}{=} \textsf{true}$:

```
1:  function MODEST(m)
2:      mark each element of the list (e_i)_{i=0}^{l-1} white, and initialize d := ∅.
3:      for κ repetitions do
4:          for i ∈ {0, ..., l − 1} do
5:              if 0 < |e_i \ d| ≤ m then
6:                  mark e_i black.
7:                  overwrite d ∪= e_i.
8:                  break the inner loop 4.
9:      for i ∈ {0, ..., l − 1} do
10:         if |ē_i \ d| ≤ m then return true.
11:     return false.
```

We identify vectors in $\mathbb{F}_2^\kappa$ with subsets of $\{0, \ldots, \kappa - 1\}$ in Definition 3.3, by means of the identification discussed in Section 2.

Informally, the matrix $(\mathbf{x}_i)_{i=0}^{l-1}$ has modesty m if there exists a sequence of minority vectors $\mathbf{e}_i$ each among which introduces at most m *new* bit positions, which itself in this way approaches within m bit positions of a majority vector $\overline{\mathbf{e}_i}$ (and if moreover m is the minimal integer for which this property holds). We note that, for each vector m, the vector $\mathbf{d}$ will necessarily eventually stabilize, in *at most* κ iterations of the outer loop 3 (in the sense that no new vectors will get marked black within the loop's body).

Remark 3.4. It is subtle, but true, that, for each m, the vector $\mathbf{d}$ produced by Definition 3.3 depends *only* on the rows $(\mathbf{e}_i)_{i=0}^{l-1}$, and not on their ordering. As we appear not to need this fact directly, we omit its proof.

We now study the probability that the minimal rank r of (1) is low. Instead of precisely describing the distribution of r as a random variable, we instead fix a cutoff $r^* \in \{1, \ldots, \kappa\}$, and upper-bound the probability that $r < r^*$. Our main result is as follows.

Proposition 3.5. *Let the rank cutoff $r^* \in \{1, \ldots, \kappa\}$ be arbitrary. For each fixed matrix $(\mathbf{x}_i)_{i=0}^{l-1}$, of modesty $m \in \{1, \ldots, \kappa\}$, say, the probability—over the random choice of $(\chi_i)_{i=0}^{l'-1}$—that $r < r^*$ is at most $2^{2 \cdot \kappa \cdot r^* - \frac{\kappa^2}{2 \cdot m}}$.*

Proof. We begin by simplifying the expression for F_x in certain ways. We first note that—at the cost of adding $\sum_{x_i=1} \chi_i$ to each index x in the expression (1) above (which has no effect on r)—we may freely replace each $\mathbf{x}_i$ with $\mathbf{e}_i$ in F_x's definition. We thus rewrite F_x as follows:

$$F_x : \Delta \mapsto \sum_{i=0}^{l'-1} \chi_i \cdot (\mathbf{e}_i * \Delta) + x \cdot \Delta.$$

We finally argue that the random variable r of (1)—viewed, again, as a function of the random coefficients $(\chi_i)_{i=0}^{l'-1}$—remains identical if we replace the matrix $(\mathbf{e}_i)_{i=0}^{l'-1}$ with its reduced row-echelon form over $\mathbb{F}_2$. Indeed, each map F_x may be decomposed into the $\mathbb{F}_2$-linear map $\Delta \mapsto (\mathbf{e}_i * \Delta)_{i=0}^{l'-1}$ from $\mathbb{F}_2^\kappa \to \mathbb{F}_{2^\kappa}^{l'}$, on the one hand, followed by the application of the random $\mathbb{F}_{2^\kappa}$-hyperplane given by $(\chi_i)_{i=0}^{l'-1}$, on the other (and finally by the addition of $x \cdot \Delta$). Row-reducing $(\mathbf{e}_i)_{i=0}^{l'-1}$ amounts to interposing between these first two maps a further $l' \times l'$ invertible matrix over $\mathbb{F}_{2^\kappa}$. Up to a fresh uniform resampling of the hyperplane coefficients $(\chi_i)_{i=0}^{l'-1}$, this matrix multiplication has no effect.

We record the following lemma:

Lemma 3.6. *If $(\mathbf{x}_i)_{i=0}^{l'-1}$ has modesty m, then the $\mathbb{F}_2$-row-reduction of $(\mathbf{e}_i)_{i=0}^{l'-1}$ has at least $\frac{\kappa}{2 \cdot m} - 1$ pivots.*

Proof. As each matrix's number of pivots depends only on its rank, it suffices to prove the lemma after arbitrarily permuting $(\mathbf{e}_i)_{i=0}^{l'-1}$'s rows and columns. We thus first sort the rows $(\mathbf{e}_i)_{i=0}^{l'-1}$ in the order in which they are marked black by the procedure MODEST(m) of Definition 3.3 (deferring white rows). Moreover, we apply the following modification to the Gaussian elimination algorithm. By definition, each successive black row necessarily introduces a 1 to some column which thus far has lacked one. Upon each such row's treatment by the algorithm, after possibly transposing two columns, we may ensure that this 1 resides at the column being considered for a pivot, and thus becomes a pivot. This transposition preserves the invariant whereby each *further* black row introduces a 1 at some new column. Likewise, using the new pivot row to clear the pivot column also preserves this invariant. We thus conclude that there are at least as many pivots as there are black rows.

Finally, we note that there must be at least $\frac{\kappa/2 - m}{m} = \frac{\kappa}{2 \cdot m} - 1$ black rows in any matrix with modesty m. Indeed, each row marked black may, by definition, increase the Hamming weight $w(\mathbf{d})$ by at most m; as each majority row satisfies $w(\overline{\mathbf{e}_i}) \geq \frac{\kappa}{2}$, the conclusion follows. □

Henceforth, we write $\widehat{\kappa}$ for the number of pivots in $(\mathbf{e}_i)_{i=0}^{l'-1}$.

We continue our study of the maps F_x. We note that each such map can be written using the following matrix expression:

$$F_x : \Delta \mapsto \left(\left[\boxed{\chi_0} \right] + \cdots + \left[\boxed{\chi_{l'-1}} \right] + \left[\boxed{x} \right] \right) \cdot \left[\Delta \right],$$

where the field elements x and $(\chi_i)_{i=0}^{l'-1}$ are viewed as $\mathbb{F}_2$-linear operators on $\mathbb{F}_2^\kappa$, and hence represented as $\kappa \times \kappa$ $\mathbb{F}_2$-*matrices*, and the shaded boxes indicate that certain columns have been "struck out". Indeed, we keep or strike columns of the matrices $(\chi_i)_{i=0}^{l'-1}$ according to the (row-reduced) data $(\mathbf{e}_i)_{i=0}^{l'-1}$; specifically, if $\mathbf{e}_{i,j} = 1$, we keep the j^{th} column of χ_i's matrix intact, and otherwise replace it with a column of 0s.

In light of our assumption that $(\mathbf{e}_i)_{i=0}^{l'-1}$ is row-reduced, we see that each pivot in the matrix $(\mathbf{e}_i)_{i=0}^{l'-1}$ adds an *independent random column* to the matrix expression above (i.e., to the left-hand sum, *excluding* x). We argue that we may consider the pivot columns *alone* in our study of (1). Indeed, replacing each non-pivot column with a column of 0s—in *all* matrices within the expression above, including that of x—can only *decrease* the rank of the resulting map F_x; we shall lower-bound this rank regardless.

We're thus left to consider the following modified expression for F_x:

$$F_x : \Delta \mapsto \left(\left[\boxed{X} \right] + \left[\boxed{x} \right] \right) \cdot \left[\Delta \right],$$

where the first matrix, say X, contains $\widehat{\kappa}$ *independently* random columns, with its further columns identically 0, and where the second matrix is merely the field-multiplication matrix of x, with the same set of $\widehat{\kappa}$ columns kept and the rest struck out.

We now consider the probability, over the uniformly random submatrix X, that $\min_{x \in \mathbb{F}_{2^\kappa}} \operatorname{rank}(F_x) < r^*$. We make use of a counting argument in $\mathbb{F}_2^{\kappa \times \kappa}$; more precisely, the argument takes place in $\mathbb{F}_2^{\kappa \times \widehat{\kappa}}$. Slightly abusing notation, we identify field elements $x \in \mathbb{F}_{2^\kappa}$ with (appropriately stricken) matrices in $\mathbb{F}_2^{\kappa \times \widehat{\kappa}}$. We note that there are exactly 2^κ distinct field elements $x \in \mathbb{F}_{2^\kappa}$, and hence at most 2^κ distinct corresponding matrices. On the other hand, for each matrix $X \in \mathbb{F}_2^{\kappa \times \widehat{\kappa}}$ for which, for some $x \in \mathbb{F}_{2^\kappa}$, $\operatorname{rank}(X + x) < r^*$ holds, we necessarily have that $X + x = Y$, where $Y \in \mathbb{F}_2^{\kappa \times \widehat{\kappa}}$ is of rank less than r^*. We undertake to count such matrices Y.

Lemma 3.7. *For each rank $r^* \in \{1, \ldots, \kappa\}$, at most $2^{(\kappa+\widehat{\kappa}) \cdot (r^*-1)}$ matrices $Y \in \mathbb{F}_2^{\kappa \times \widehat{\kappa}}$ satisfy* $\operatorname{rank}(Y) < r^*$.

Proof. Each matrix $Y \in \mathbb{F}_2^{\kappa \times \widehat{\kappa}}$ of rank less than r^* can be written (possibly non-uniquely) as the product of a $\kappa \times (r-1)$ matrix and an $(r-1) \times \widehat{\kappa}$ matrix. □

The set of matrices $X \in \mathbb{F}_2^{\kappa \times \widehat{\kappa}}$ for which $\min_{x \in \mathbb{F}_{2^\kappa}} \operatorname{rank}(X + x) < r^*$ is exactly the union, over all field elements $x \in \mathbb{F}_{2^\kappa}$, of the sets $\{x + Y \mid \operatorname{rank}(Y) < r^*\} \subset \mathbb{F}_2^{\kappa \times \widehat{\kappa}}$. In light of Lemma 3.7, we conclude that the cardinality of this union is at most $2^\kappa \cdot 2^{(\kappa+\widehat{\kappa}) \cdot (r^*-1)} = 2^{(\kappa+\widehat{\kappa}) \cdot (r^*-1)+\kappa}$. Finally, the *total* number of $\kappa \times \widehat{\kappa}$ matrices X is obviously $2^{\kappa \cdot \widehat{\kappa}}$. The probability, over the random coefficients $(\chi_i)_{i=0}^{l'-1}$, that $\min_{x \in \mathbb{F}_{2^\kappa}} \operatorname{rank}(F_x) < r^*$ is thus at most $2^{(\kappa+\widehat{\kappa}) \cdot (r^*-1)+\kappa-\kappa \cdot \widehat{\kappa}}$. From Lemma 3.6, we recall that $\widehat{\kappa} \geq \frac{\kappa}{2 \cdot m} - 1$; on the other hand, we trivially have $\widehat{\kappa} \leq \kappa$. We conclude that the probability in question is at most $2^{2 \cdot \kappa \cdot (r^*-1)+\kappa-\kappa \cdot \left(\frac{\kappa}{2 \cdot m}-1\right)} = 2^{2 \cdot \kappa \cdot r^* - \frac{\kappa^2}{2 \cdot m}}$. This completes the proof of the proposition. □

We now consider the distinguisher's distinguishing probability, conditioned on $\mathcal{A}$ attaining a low minimal rank $r < r^*$, submitting a value x for which moreover $\operatorname{rank}(F_x) < r^*$, and finally passing the correlation check. We recall that the real and ideal distributions are identical *unless* D queries $H(i \parallel \mathbf{t}_i + \overline{\mathbf{e}_i} * \Delta)$, for some $i \in \{0, \ldots, l-1\}$.

Roughly, the proposition below argues that—for a matrix with modesty m—any successful distinguisher must brute-force *some* segment containing at least m bits. On the other hand, D may make use of the information whereby $\mathcal{A}$ has passed the correlation check in the first place (namely, that $\Delta \in \ker(F_x)$). By our hypothesis that $\operatorname{rank}(F_x)$ is low, however, this information necessarily furnishes fewer than r^* linear relations on Δ; D must thus perform something like 2^{m-r^*} work in the best case. The following proposition makes this reasoning precise.

Proposition 3.8. *For each computationally unbounded distinguisher D, which makes at most $Q(\kappa)$ queries to the random oracle, say, and each modesty $m \in \{1, \ldots, \kappa\}$ and rank cutoff $r^* \in \{1, \ldots, \kappa\}$, we have that $\left|\Pr\left[D\left(\mathsf{Real}_{\Pi^{\kappa,l}_{\mathsf{ROT}},\mathcal{A},R}(\kappa)\right) = 1\right] - \Pr\left[D\left(\mathsf{Ideal}_{\mathcal{F}^{\kappa,l}_{\mathsf{ROT}},\mathcal{S},R}(\kappa)\right) = 1\right]\right| \leq Q(\kappa) \cdot 2^{r^*-m}$, where both distributions are conditioned on $\mathcal{A}$'s matrix having modesty m and on the rank relation rank$(F_x) < r^*$.*

Proof. We may immediately further condition both distributions on the subcase in which $\mathcal{A}$ passes the correlation check, since otherwise the distributions are identical. We pick an arbitrary execution satisfying these conditions. We write $\mathbf{d}$ for the vector constructed during the course of $\textsc{Modesty}(m-1)$, and $\mathbf{w} := \mathbf{d} * \Delta$ for the projection of the hidden choice vector Δ onto $\mathbf{d}$. It suffices to prove the result after giving D $\mathbf{w}$, since this information can only make D more effective. To prove the result, we upper-bound the probability that D makes *any* query of the form $H(i \parallel \mathbf{t}_i + \mathbf{f} * \Delta)$, where $i \in \{0, \ldots, l-1\}$, and either $\mathbf{f} = \mathbf{e}_i$ and $\mathbf{e}_i \not\subset \mathbf{d}$ or $\mathbf{f} = \overline{\mathbf{e}_i}$. By Definition 3.3, $\textsc{Modesty}(m-1) = \mathsf{false}$, and each such vector $\mathbf{f}$ satisfies $|\mathbf{f} \setminus \mathbf{d}| \geq m$.

In D's view (assuming it has not already made some such query), the distribution of $\mathbf{f} * \Delta$ is uniform over the affine-linear subspace $\{\mathbf{f} * \Delta \mid \Delta \in \ker(F_x) \wedge \mathbf{d} * \Delta = \mathbf{w}\}$. By a dimension count, and our hypothesis whereby $\text{rank}(F_x) < r^*$, the dimension of this subspace is greater than $m - r^*$. The probability that each particular among D's queries succeeds—i.e., that D guesses $\mathbf{f} * \Delta$—is thus less than 2^{r^*-m}. Taking a union bound over these queries, we establish the proposition. □

We are now in a position to prove the theorem. Traversing the tree of Figure 1, and invoking Propositions 3.5 and 3.8, we see that for each distinguisher D as in the hypothesis of Proposition 3.8, the difference $\left|\Pr\left[D\left(\mathsf{Real}_{\Pi^{\kappa,l}_{\mathsf{ROT}},\mathcal{A},R}(\kappa)\right) = 1\right] - \Pr\left[D\left(\mathsf{Ideal}_{\mathcal{F}^{\kappa,l}_{\mathsf{ROT}},\mathcal{S},R}(\kappa)\right) = 1\right]\right|$—where, here, we *don't* condition the two distributions—is at most:

$$\min\left(1, 2^{2\cdot\kappa\cdot r^* - \frac{\kappa^2}{2\cdot m}}\right) \cdot \min\left(1, Q(\kappa) \cdot 2^{r^*-m}\right) + 2 \cdot 2^{-r^*}. \tag{2}$$

We set $r^* := \frac{1}{4} \cdot \sqrt{\kappa}$ for the remainder of the proof. We handle two cases, corresponding to whether $m < \frac{1}{2} \cdot \sqrt{\kappa}$ or not. If $m < \frac{1}{2} \cdot \sqrt{\kappa}$, then the first factor of (2)'s exponent is at most $\frac{1}{2}\kappa^{3/2} - \kappa^{3/2} = -\frac{1}{2} \cdot \kappa^{3/2}$, so this factor is negligible, and the result holds. If $m \geq \frac{1}{2} \cdot \sqrt{\kappa}$, then the second factor's exponent is at most $\frac{1}{4} \cdot \sqrt{\kappa} - \frac{1}{2} \cdot \sqrt{\kappa} = -\frac{1}{2} \cdot \sqrt{\kappa}$, so this factor instead is negligible (provided $Q(\kappa)$ is polynomial). The final summand's exponent is $-\frac{1}{4} \cdot \sqrt{\kappa} + 1$, so this term is negligible. This completes the proof of the theorem. □

We now extract effective bounds from our proof. Our proof can be made to yield concrete values κ at which KOS achieves prescribed security guarantees.

Theorem 3.9. *For given computational and statistical security parameters λ and s, respectively, for it to be the case that $\left|\Pr\left[D\left(\mathsf{Real}_{\Pi^{\kappa,l}_{\mathsf{ROT}},\mathcal{A},R}(\kappa)\right) = 1\right] - \Pr\left[D\left(\mathsf{Ideal}_{\mathcal{F}^{\kappa,l}_{\mathsf{ROT}},\mathcal{S},R}(\kappa)\right) = 1\right]\right| \leq 2^{-s}$ holds for each distinguisher D making at most 2^λ hash evaluations, it suffices that $\kappa > 4 \cdot \lambda \cdot s + 8 \cdot s^2 + 8 \cdot \lambda + 28 \cdot s + 24$.*

Proof. We set $r^* := s + 2$ once and for all. We describe a selection procedure for κ which guarantees that, as $m \in \{1, \ldots, \kappa\}$ varies, *at least one* among the exponent expressions $2 \cdot \kappa \cdot r^* - \frac{\kappa^2}{2 \cdot m}$ and $\lambda + r^* - m$ of (2) necessarily fails to exceed $-r^* + 1$ (i.e., at each given m). This suffices to establish the result, since it serves to bound (2) from above by $2^{-r^*+1} + 2^{-r^*+1} = 2^{-s}$ (for $Q(\kappa) := 2^\lambda$).

We observe that for each fixed κ and r^*, the exponent expressions $2 \cdot \kappa \cdot r^* - \frac{\kappa^2}{2 \cdot m}$ and $\lambda + r^* - m$ are *increasing* and *decreasing*, respectively, over the interval $m \in \{1, \ldots, \kappa\}$. It thus suffices to choose κ in such a way that these two expressions' respective intersections with the line $-r^* + 1$ occur in the "right" order (i.e., the first expression intersects after the second does). As basic algebraic manipulations demonstrate, the two expressions' respective intersection points are $m = \frac{\kappa^2}{2 \cdot (2 \cdot \kappa \cdot r^* + r^* - 1)}$ and $m = \lambda + 2 \cdot r^* - 1$. The inequality we need is thus $\frac{\kappa^2}{2 \cdot (2 \cdot \kappa \cdot r^* + r^* - 1)} \geq \lambda + 2 \cdot r^* - 1$. We determine (or more precisely, upper-bound) the smallest κ for which this inequality holds. We express the required inequality as follows:

$$\begin{aligned}
\kappa^2 &\geq 2 \cdot (2 \cdot \kappa \cdot r^* + r^* - 1) \cdot (\lambda + 2 \cdot r^* - 1) \\
&= (4 \cdot \kappa \cdot s + 8 \cdot \kappa + 2 \cdot r + 2) \cdot (\lambda + 2 \cdot s + 3) && \text{(use } s = r^* + 2.\text{)} \\
&= \kappa \cdot \left(4 \cdot \lambda \cdot s + 8 \cdot s^2 + 8 \cdot \lambda + 28 \cdot s + 24\right) + \left(2 \cdot \lambda \cdot s + 4 \cdot s^2 + 2 \cdot \lambda + 10 \cdot r + 6\right). && \text{(rearrange.)}
\end{aligned}$$

Upon increasing certain coefficients in the right-hand term, we obtain a stronger inequality, which implies the one we need:

$$\kappa^2 \geq \kappa \cdot \left(4 \cdot \lambda \cdot s + 8 \cdot s^2 + 8 \cdot \lambda + 28 \cdot s + 24\right) + \left(2 \cdot \lambda \cdot s + 4 \cdot s^2 + 4 \cdot \lambda + 14 \cdot r + 12\right).$$

Denoting $a := 4 \cdot \lambda \cdot s + 8 \cdot s^2 + 8 \cdot \lambda + 28 \cdot s + 24$, we obtain the requirement $\kappa^2 \geq \kappa \cdot a + \frac{1}{2} \cdot a$, or $\frac{\kappa^2}{\left(\kappa + \frac{1}{2}\right)} \geq a$. Since $\frac{\kappa^2}{\left(\kappa + \frac{1}{2}\right)} \geq \kappa - 1$, we see that $\kappa - 1 \geq a$ implies the desired inequality. This completes the proof. □

Remark 3.10. Theorem 3.9 can be viewed as a precise variant of the final argument of Theorem 3.1, in which s and λ are prescribed, and we moreover select the modesty cutoff optimally (i.e., in such a way as to make (2) decay as quickly as possible). Indeed, for κ chosen as in Theorem 3.9, the optimal cutoff—and the most effective attack strategy for the adversary—appears at the modesty $m^* = \lambda + 2 \cdot r^* = \lambda + 2 \cdot (s + 2)$.

Example 3.11. For $s := 30$, and $\lambda := 60$, Theorem 3.9 guarantees security as long as $\kappa \geq 15{,}745$.

Example 3.12. For $s := 40$ and $\lambda := 80$, Theorem 3.9 guarantees security as long as $\kappa \geq 27{,}385$.

Example 3.13. For $s := 80$ and $\lambda := 128$, Theorem 3.9 guarantees security as long as $\kappa \geq 95{,}449$.

Theorem 3.9 yields parameters which are barely practical—if at all—for reasonable levels of security. It is, of course, possible that our proof could be strengthened (or another proof found), so as to yield stronger bounds, and security under more reasonable parameter sizes. In particular, the *concrete* security of KOS at $\kappa = 128$ remains open. On the other hand, an improvement to our result seems out of reach, barring strikingly new techniques. We explain this as follows.

Corollary 3.14. *For κ sufficiently large, for each distinguisher D making at most $\frac{1}{2} \cdot \sqrt{\kappa} \cdot 2^{\frac{1}{4} \cdot \sqrt{\kappa}}$ hash evaluations, the probability of success* $\left|\Pr\left[D\left(\mathsf{Real}_{\Pi^{\kappa,l}_{\mathsf{ROT}},\mathcal{A},R}(\kappa)\right) = 1\right] - \Pr\left[D\left(\mathsf{Ideal}_{\mathcal{F}^{\kappa,l}_{\mathsf{ROT}},\mathcal{S},R}(\kappa)\right) = 1\right]\right| \leq 2^{-\frac{1}{4} \cdot \sqrt{\kappa}}$.

Proof. For arbitrary κ, we set $s := \frac{1}{4} \cdot \sqrt{\kappa}$ and $\lambda := \frac{1}{4} \cdot \sqrt{\kappa} + \frac{1}{2} \cdot \log(\kappa) - 1$. We observe that for s and λ chosen this way—if $\kappa \geq 3{,}581$—we have $\kappa > 4 \cdot \lambda \cdot s + 8 \cdot s^2 + 8 \cdot \lambda + 28 \cdot s + 24$. Theorem 3.9 thus implies that any attack using at most $2^\lambda = \frac{1}{2} \cdot \sqrt{\kappa} \cdot 2^{\frac{1}{4} \cdot \sqrt{\kappa}}$ hashes must succeed with probability at most $2^{-s} = 2^{-\frac{1}{4} \cdot \sqrt{\kappa}}$. □

In other words, there does not exist an attack on KOS which uses only $\frac{1}{2} \cdot \sqrt{\kappa} \cdot 2^{\frac{1}{4} \cdot \sqrt{\kappa}}$ hash evaluations and succeeds with probability greater than $2^{-\frac{1}{4} \cdot \sqrt{\kappa}}$.

Roy [Roy22, § 4.1] describes a "subfield attack" on KOS—which requires $2^{\frac{1}{5} \cdot \kappa}$ oracle queries and succeeds with probability $2^{-\frac{3}{5} \cdot \kappa}$. This attack is significantly more costly and unlikely to succeed than those ruled out by our proof, so the analysis of KOS still contains a gap. (Of course, it's nonetheless stronger than those which KOS's original proof sought to rule out.) On the other hand, Roy [Roy22, § 4.1] describes a *further* attack on a different protocol—namely, "PSS", for Patra, Sarkar and Suresh [PSS17]—which is much more devastating; that attack requires $\frac{1}{2} \cdot \sqrt{\kappa} \cdot 2^{\sqrt{\kappa}}$ hash evaluations and succeeds with probability $2^{-\sqrt{\kappa}}$. As it turns out, our proof serves equally well—without change—to describe the security of PSS. Indeed, we use *only* the property of the field elements $(\chi_i)_{i=0}^{l'-1}$ whereby, for $\chi_i \leftarrow \mathbb{F}_{2^\kappa}$ sampled randomly, each *individual* column of χ_i's matrix representation is itself uniformly random in $\{0,1\}^\kappa$ (of course, this does not hold for the columns considered jointly). This property holds also for PSS, in fact, even though they construct their matrices differently (with a single random column repeated). I would like to thank an anonymous referee for helping to explain this fact.

The lower-bound established by our Corollary 3.14, which applies to both KOS and PSS, exactly matches—up to the constant $\frac{1}{4}$ appearing in the expressions' exponents—the upper-bound achieved by Roy [Roy22, § 4.1] on PSS. Our proof thus definitively settles the question of PSS's security (up to the constant). Moreover, it demonstrates that any *better* security argument for KOS—if one exists—would have to rely in some special way on the structure of the field elements $(\chi_i)_{i=0}^{l'-1}$, and on the nature of their role in the correlation check. We emphasize that Roy's attack on PSS is *not* known to apply to KOS. Rather, the opposite is true; our defense of KOS applies to PSS. The security of KOS thus resides somewhere between the lower-bound established by our Theorem 3.9 and the upper-bound achieved by Roy's subfield attack. In any case, our result furnishes the *only* currently-known lower-bound for KOS, and its only proof of security.

We leave the task of exactly settling KOS's security to future work. In the meantime, we suggest that an alternative known to be concretely secure—like Roy's *SoftSpokenOT* [Roy22]—be used when possible.

References

[Coh82] P. M. Cohn. *Algebra*, volume 1. John Wiley & Sons, second edition, 1982.

[KOS15] Marcel Keller, Emmanuela Orsini, and Peter Scholl. Actively secure OT extension with optimal overhead. In Rosario Gennaro and Matthew Robshaw, editors, *Advances in Cryptology – CRYPTO 2015*, volume 9215 of *Lecture Notes in Computer Science*, pages 724–741, Berlin, Heidelberg, 2015. Springer Berlin Heidelberg.

[KOS22] Marcel Keller, Emmanuela Orsini, and Peter Scholl. Actively secure OT extension with optimal overhead. Unpublished update, `https://eprint.iacr.org/2015/546.pdf`, September 2022.

[Lin17] Yehuda Lindell. *Tutorials on the Foundations of Cryptography: Dedicated to Oded Goldreich*, chapter How to Simulate It – A Tutorial on the Simulation Proof Technique, pages 277–346. Information Security and Cryptography. Springer International Publishing, 2017.

[PSS17] Arpita Patra, Pratik Sarkar, and Ajith Suresh. Fast actively secure ot extension for short secrets. In *Network and Distributed System Security Symposium*. Internet Society, 2017.

[Roy22] Lawrence Roy. SoftSpokenOT: Quieter OT extension from small-field silent VOLE in the minicrypt model. In Yevgeniy Dodis and Thomas Shrimpton, editors, *Advances in Cryptology – CRYPTO 2022*, volume 13507 of *Lecture Notes in Computer Science*, pages 657–687, Cham, 2022. Springer Nature Switzerland.