Lossy Encryption: Constructions from General Assumptions and Efficient Selective Opening Chosen Ciphertext Security

Brett Hemenway[1], Benoît Libert[2], Rafail Ostrovsky[3], and Damien Vergnaud[4]

[1] University of Michigan (USA)
[2] Université catholique de Louvain (Belgium)
[3] University of California, Los Angeles (USA)
[4] École Normale Supérieure – C.N.R.S. - INRIA (France)

Abstract. Lossy encryption was originally studied as a means of achieving efficient and composable oblivious transfer. Bellare, Hofheinz and Yilek showed that lossy encryption is also selective opening secure. We present new and general constructions of lossy encryption schemes and of cryptosystems secure against selective opening adversaries.

We show that *every* re-randomizable encryption scheme gives rise to efficient encryptions secure against a selective opening adversary. We show that statistically-hiding 2-round Oblivious Transfer implies Lossy Encryption and so do smooth hash proof systems. This shows that private information retrieval and homomorphic encryption *both* imply Lossy Encryption, and thus Selective Opening Secure Public Key Encryption.

Applying our constructions to well-known cryptosystems, we obtain selective opening secure commitments and encryptions from the *Decisional Diffie-Hellman*, *Decisional Composite Residuosity* and *Quadratic Residuosity* assumptions.

In an indistinguishability-based model of chosen-ciphertext selective opening security, we obtain secure schemes featuring short ciphertexts under standard number theoretic assumptions. In a simulation-based definition of chosen-ciphertext selective opening security, we also handle non-adaptive adversaries by adapting the Naor-Yung paradigm and using the perfect zero-knowledge proofs of Groth, Ostrovsky and Sahai.

Keywords: Public key encryption, commitment, lossy encryption, homomorphic encryption, selective opening, chosen-ciphertext security

1 Introduction

In Byzantine agreement, and more generally in secure multiparty computation, it is often assumed that all parties are connected to each other via private channels. In practice, these private channels are implemented using a public-key cryptosystem. An adaptive adversary in a MPC setting, however, has very different powers than an adversary in an IND-CPA or IND-CCA game. In particular, an adaptive MPC adversary may view all the encryptions sent in a given round, and then choose to corrupt a certain fraction of the players, thus revealing the decryptions of those players' messages *and the randomness used to encrypt them.* A natural question is whether the messages sent from the uncorrupted players remain secure. If the messages (and randomness) of all the players are chosen independently, then security in this setting follows from the IND-CPA security of the underlying encryption. If, however, the messages are not independent, the security does not immediately follow from the IND-CPA (or even IND-CCA) security of the underlying scheme. Although this problem was first investigated over twenty years ago, it remains an open question whether IND-CPA security implies this *selective opening* security.

Previous Work. There have been many attempts to design encryption protocols that can be used to implement secure multiparty computation against an adaptive adversary. The first protocols by Beaver and Haber [4] required interaction between the sender and receiver, required erasure and were fairly inefficient. The first non-interactive protocol was given by Canetti, Feige, Goldreich and Naor in [10]. In [10] the authors defined a new primitive called Non-Committing Encryption, and gave an example of such a scheme based on the RSA assumption. In [2], Beaver extended the work of [10], and created adaptively secure key exchange under the Diffie-Hellman assumption. In subsequent work, Damgård and Nielsen improved the efficiency of the schemes of Canetti *et al.* and Beaver, they were also able to obtain Non-Committing Encryption based on one-way trapdoor functions with invertible sampling. In [12], Canetti, Halevi and Katz presented a Non-Committing encryption protocols with evolving keys.

In [9], Canetti, Dwork, Naor and Ostrovsky extended the notion of Non-Committing Encryption to a new protocol which they called Deniable Encryption. In Non-Committing Encryption schemes there is a simulator, which can generate non-committing ciphertexts, and later open them to any desired message, while in Deniable Encryption, valid encryptions generated by the sender and receiver can later be opened to any desired message. The power of this primitive made it relatively difficult to realize, and Canetti *et al.* were only able to obtain modest examples of Deniable Encryption and left it as an open question whether fully deniable schemes could be created.

The notions of security against an adaptive adversary can also be applied to commitments. According to [21], the necessity of adaptively-secure commitments was realized by 1985. Despite its utility, until recently, relatively few papers directly addressed the question of commitments secure against a selective opening adversary (SOA). The work of Dwork, Naor, Reingold and Stockmeyer [21] was the first to explicitly address the problem. In [21], Dwork *et al.* showed that non-interactive SOA-secure commitments can be used to create a 3-round zero-knowledge proof systems for NP with negligible soundness error, and they gave constructions of a weak form of SOA-secure commitments, but left as an open question the existence of whether general SOA-secure commitments.

The question of SOA-secure commitments was put on firm foundations by Hofheinz [27] and Bellare, Hofheinz and Yilek in [5]. In [5], Bellare *et al.* provided simulation-based and indistinguishability-based definitions of security (these will be given the prefixes IND and SEM respectively) and gave a number of constructions and strong black-box separations, which indicated the difficulty of constructing selective opening secure commitments. Our results in the selective opening setting build on the breakthrough results of [5].

The independent work of Fehr, Hofheinz and Kiltz and Wee [23] also examines the case of CCA2 cryptosystems that are selective opening secure. In their work, they show how to adapt the universal hash proof systems of [17], to provide CCA2 security in the selective opening setting. Their constructions are general, and offer the first SEM-SO-CCA secure cryptosystem whose parameters are completely independent of n, the number of messages. Their work also considers selective opening security against chosen-plaintext attacks, and using techniques from Non-Committing Encryption [10] they construct SEM-SO-CPA secure systems from enhanced one-way trapdoor permutations.

Bellare, Waters and Yilek [7] show how to construct Identity-Based Encryption (IBE) schemes secure under selective-opening attacks. Our results are orthogonal to theirs. Their work constructs IBE schemes secure under selective-opening attacks, while our work starts with a tag-based encryption scheme, and uses it to construct encryption schemes that are secure against a selective-opening chosen-ciphertext attack, but are not identity-based.

Our Contributions. We primarily consider encryptions secure against a selective opening adversary. First we consider a selective-opening adversary who can mount a chosen-plaintext attack, and a the second part, we consider a selective-opening adversary who can mount a chosen-ciphertext attack.

Selective Opening Security Against Chosen-Plaintext Attacks. We formalize the notion of re-randomizable Public-Key Encryption and show that it implies Lossy Encryption [41, 32, 5]. Combining this with the observation (due to Bellare *et al.* [5]) that Lossy Encryption is IND-SO-CPA secure, we obtain an efficient construction of IND-SO-CPA secure encryption from any re-randomizable encryption (which generalizes and extends previous results). Moreover, these constructions retain the efficiency of the underlying re-randomizable cryptosystem.

Applying our results to the Paillier cryptosystem [39], we obtain an encryption scheme attaining a strong, simulation-based form of semantic security under selective openings (SEM-SO-CPA security). This is the first such construction from the Composite Residuosity (DCR) assumption. As far as bandwidth goes, it is also the most efficient SEM-SO-CPA secure encryption scheme to date. The possible use of Paillier as a lossy encryption scheme implicitly appears in [45]. To the best of our knowledge, its SEM-SO-CPA security was not reported earlier.

Next, we show that Lossy Encryption is also implied by (honest-receiver) statistically-hiding $\binom{2}{1}$-Oblivious Transfer and hash proof systems [17]. Combining this with the results of [42, 41], we recognize that the relatively new Lossy Encryption primitive is essentially a different way to view the well-known statistically-hiding $\binom{2}{1}$-OT primitive. Applying the reductions in [5] to this result, yields constructions of SOA secure encryption from both private information retrieval (PIR) and homomorphic encryption.

These results show that the Lossy and Selective Opening Secure Encryption primitives (at least according to the latter's indistinguishability-based security definition), which have not been extensively studied until recently, are actually implied by several well-known primitives: *i.e.*, re-randomizable encryption, PIR, homomorphic encryption, hash proof systems and statistically-hiding $\binom{2}{1}$-OT. So far, the only known general constructions of lossy encryption were from lossy trapdoor functions. Our results show that they can be obtained from many seemingly weaker primitives (see figure 1).

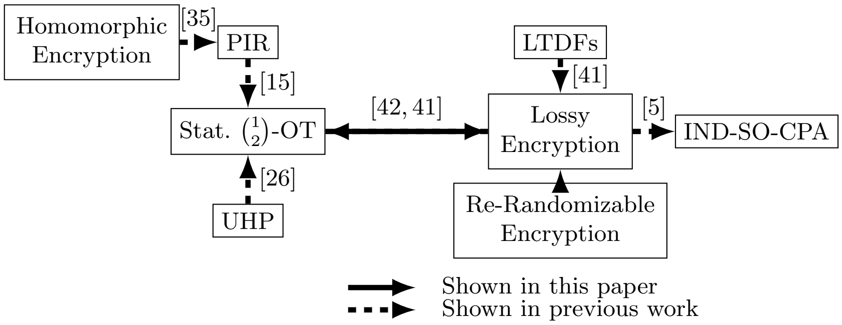

Fig. 1. Constructing Lossy Encryption

Selective Opening Security Against Chosen-Ciphertext Attacks: Continuing the study of selective-opening security, we present definitions chosen-ciphertext security (CCA2) in the selective opening setting (in both the indistinguishability and simulation-based models) and describe encryption schemes that provably satisfy these enhanced forms of security. Despite recent progress, relatively few methods are known for constructing IND-CCA2 cryptosystems in the standard model. The problem is even more complex with selective openings, where some known approaches for CCA2 security do not seem to apply. We note how the Naor-Yung paradigm, even when applied with statistical zero knowledge proofs fails to prove CCA2 security in the selective opening setting. Essentially, this is because the selective opening adversary learns the randomness used in the signature scheme, which allows him to forge signatures, and thus create ciphertexts that cannot be handled by the simulated decryption oracle.

The results of Fehr, Hofheinz, Kiltz and Wee [23] show how to modify universal hash proof systems [17] to achieve security under selective openings. We take a different approach and follow (a variant of) the Canetti-Halevi-Katz paradigm [11]. This too encounters many obstacles in the selective opening setting. Nevertheless, under standard assumptions (such as DDH or the Composite Residuosity assumption), we construct schemes featuring compact ciphertexts while resisting adaptive (*i.e.*, CCA2) chosen-ciphertext attacks according to our indistinguishability-based definition. When comparing our schemes to those of [23], we note that our public key size depends on n, the number of senders that can be possibly corrupted, while the systems of [23] are independent of n. On the other hand, to encrypt m-bit messages with security parameter λ, our ciphertexts are of length $\mathcal{O}(\lambda + m)$, while theirs are of length $\mathcal{O}(\lambda m)$. Our public-keys are longer than in [23] because our construction relies on All-But-N Lossy Trapdoor Functions (defined below), which have long description. The recent complementary work of Hofheinz [28] shows how to create All-But-Many Trapdoor Functions with short keys. Using his results in our construction eliminates the dependence of the public-key size on n. Regarding security definitions, our constructions satisfy an indistinguishability-based definition (IND-SO-CCA), whereas theirs fit a simulation-based definition (SEM-SO-CCA) which avoids the restriction on the efficient conditional re-sampleability of the message distribution.

The scheme of [23] is very different from ours and we found it interesting to investigate the extent to which well-known paradigms like [11] can be applied in the present context. Moreover, by adapting the Naor-Yung paradigm [38], under more general assumptions, we give a CCA1 construction that also satisfies a strong simulation-based notion of adaptive selective opening security.

One advantage of our IND-SO-CCA scheme is the ability to natively encrypt multi-bit messages. It is natural to consider whether our approach applies to the scheme of Bellare, Waters and Yilek [7] to achieve multi-bit IND-SO-CCA encryption. The scheme of [7], like [23], encrypts multi-bit messages in a bitwise manner. Applying a Canetti-Halevi-Katz-like transformation to the construction of [7] does not immediately yield IND-SO-CCA encryption schemes for multi-bit messages: the reason is that it is not clear how to prevent the adversary from reordering the bit encryptions without employing a one-time signature scheme.

2 Background

If $f : X \to Y$ is a function, for any subset $Z \subset X$, we let $f(Z) = \{f(x) : x \in Z\}$. If A is a PPT machine, then $a \overset{\$}{\leftarrow} A$ denotes the action of running A and obtaining an output a, which is distributed according to the internal randomness of A. Also, $\mathsf{coins}(A)$ denotes the distribution of A's internal randomness, so that the distribution $\{a \overset{\$}{\leftarrow} A\}$ is actually $\{r \overset{\$}{\leftarrow} \mathsf{coins}(A) : a = A(r)\}$. If R is a set, we use $r \overset{\$}{\leftarrow} R$ to denote sampling uniformly from R.

When λ is a security parameter, $\mathsf{negl}(\lambda)$ denotes the set of negligible functions (*i.e.*, which decrease faster than the inverse of any polynomial in λ). If X and Y are families of distributions indexed by λ, their statistical indistinguishability is written as $X \approx_s Y$. We write $X \approx_c Y$ to express that X and Y are computationally indistinguishable, *i.e.*, for all PPT adversaries A, for all polynomials p, then for all sufficiently large λ, we have $|\Pr[A^X = 1] - \Pr[A^Y = 1]| \in \mathsf{negl}(\lambda)$.

2.1 Selective Opening Secure Encryption

We recall the indistinguishability-based definition of encryption secure against a selective opening adversary, originally formalized in [5]. We define a real game and an ideal game which should be indistinguishable to any efficient adversary. The adversary receives *both* the messages and the randomness for his selection.

This mirrors the fact that an adaptive MPC adversary learns the entire history of corrupted players (*i.e.*, there are no secure erasures). If the adversary receives only the messages this would reduce to standard CPA security.

As in the notations of [5], $\mathcal{M}$ denotes an n-message sampler outputting a n-vector $\mathbf{m} = (m_1, \ldots, m_n)$ of messages whereas $\mathcal{M}_{|I,\mathbf{m}[I]}$ denotes an algorithm that conditionally resamples another random n-vector $\mathbf{m}' = (m'_1, \ldots, m'_n)$ such that $m'_i = m_i$ for each $i \in I \subset \{1, \ldots, n\}$. If such a resampling can be done efficiently for all $I, \mathbf{m}$, then $\mathcal{M}$ is said to support efficient conditional resampling.

Definition 1. *(Indistinguishability under selective openings). A public key cryptosystem (G, E, D) is indistinguishable under selective openings (or IND-SO-CPA secure) if, for any message sampler $\mathcal{M}$ supporting efficient conditional resampling and any PPT adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$, we have*

$$\left|\Pr\left[\mathcal{A}^{\textsf{\textit{ind-so-real}}} = 1\right] - \Pr\left[\mathcal{A}^{\textsf{\textit{ind-so-ideal}}} = 1\right]\right| \in \mathsf{negl}(\lambda)$$

where the games ind-so-real and ind-so-ideal are defined as follows.

IND-SO-CPA (Real)	**IND-SO-CPA (Ideal)**
$\mathbf{m} = (m_1, \ldots, m_n) \xleftarrow{\$} \mathcal{M}$	$\mathbf{m} = (m_1, \ldots, m_n) \xleftarrow{\$} \mathcal{M}$
$r_1, \ldots, r_n \xleftarrow{\$} \mathsf{coins}(E)$	$r_1, \ldots, r_n \xleftarrow{\$} \mathsf{coins}(E)$
$(I, st) \xleftarrow{\$} \mathcal{A}_1\big(pk, E(m_1, r_i), \ldots$	$(I, st) \xleftarrow{\$} \mathcal{A}_1\big(pk, E(m_1, r_i), \ldots, E(m_n, r_n)\big)$
$\ldots, E(m_n, r_n)\big)$	$\mathbf{m}' = (m'_1, \ldots, m'_n) \xleftarrow{\$} \mathcal{M}_{\mid I,\mathbf{m}[I]}$
$b \xleftarrow{\$} \mathcal{A}_2\big(st, (m_i, r_i)_{i \in I}, \mathbf{m}\big)$	$b \xleftarrow{\$} \mathcal{A}_2\big(st, (m_i, r_i)_{i \in I}, \mathbf{m}'\big)$

In the real game, the challenger samples $\mathbf{m} = (m_1, \ldots, m_n) \xleftarrow{\$} \mathcal{M}$, chooses $r_1, \ldots, r_n \xleftarrow{\$} \mathsf{coins}(E)$ and sends $(E(m_1, r_1), \ldots, E(m_n, r_n))$ to $\mathcal{A}$ who responds with a subset $I \subset \{1, \ldots, n\}$ and obtains $\{r_i\}_{i \in I}$ as well as the entire *vector $\mathbf{m} = (m_1, \ldots, m_n)$. Finally, $\mathcal{A}$ outputs a bit $b \in \{0, 1\}$.*

In the ideal game, the challenger also samples $\mathbf{m} = (m_1, \ldots, m_n) \xleftarrow{\$} \mathcal{M}$, chooses $r_1, \ldots, r_n \xleftarrow{\$} \mathsf{coins}(E)$ and sends $(E(m_1, r_1), \ldots, E(m_n, r_n))$ to $\mathcal{A}$. The latter chooses a subset $I \subset \{1, \ldots, n\}$ and obtains $\{r_i\}_{i \in I}$. The only difference w.r.t. the real game is that, instead of revealing $\mathbf{m}$, the challenger samples a new vector $\mathbf{m}' \xleftarrow{\$} \mathcal{M}_{|I,\mathbf{m}[I]}$ and sends $\mathbf{m}'$ to $\mathcal{A}$. Eventually, $\mathcal{A}$ outputs a bit $b \in \{0, 1\}$.

This definition of IND-SO-CPA security (taken from [5]) does not allow the message distribution $\mathcal{M}$ to depend on the public key. However, all our proofs (as well as the proof that Lossy Encryption is IND-SO-CPA secure in [5]) go through essentially unchanged if $\mathcal{M}$ is allowed to depend on the public-key of the scheme. For consistency, we continue to use the definition of [5].

2.2 Lossy Encryption

Bellare *et al.* [5] define *Lossy Encryption*, expanding on the definitions of Dual-Mode Encryption [41] and Meaningful/Meaningless Encryption [32]. A 'lossy' (or 'messy' in the terminology of [41]) cryptosystem has two types of public keys which specify two different modes of operation. In the normal mode, encryption is injective, while in the lossy (or 'messy') mode, the ciphertexts generated by the encryption algorithm are independent of the plaintext. We also require that no efficient adversary can distinguish normal keys from lossy keys. Bellare *et al.* [5] introduce a property called *openability*, which allows a possibly inefficient algorithm to open a ciphertext generated under a lossy key to *any* plaintext.

Definition 2. *A* lossy public-key cryptosystem *is a tuple* (G, E, D) *such that*

- $G(1^\lambda, \mathsf{inj})$ *outputs keys* (pk, sk) *which are called* injective keys.
- $G(1^\lambda, \mathsf{lossy})$ *outputs keys* $(pk_{\mathsf{lossy}}, sk_{\mathsf{lossy}})$ *which are called* lossy keys.

Additionally, (G, E, D) *are efficient algorithms satisfying these properties:*

1. *We have* $\Pr[(pk, sk) \xleftarrow{\$} G(1^\lambda, \mathsf{inj});\ r \xleftarrow{\$} \mathsf{coins}(E) : D(sk, E(pk, x, r)) = x] = 1$ *for all plaintexts* $x \in X$. *This property is called correctness on injective keys.*
2. Indistinguishability of keys. *In lossy mode, public keys are computationally indistinguishable from those in the injective mode. If* $\mathrm{proj} : (pk, sk) \mapsto pk$ *is the projection map, then* $\{\mathrm{proj}(G(1^\lambda), \mathsf{inj})\} \approx_c \{\mathrm{proj}(G(1^\lambda, \mathsf{lossy}))\}$.
3. Lossiness of lossy keys. *If* $(pk_{\mathsf{lossy}}, sk_{\mathsf{lossy}}) \xleftarrow{\$} G(1^\lambda, \mathsf{lossy})$, *for all* $x_0, x_1 \in X$, *the distributions* $E(pk_{\mathsf{lossy}}, x_0, R)$ *and* $E(pk_{\mathsf{lossy}}, x_1, R)$ *are statistically close.*
4. Openability. *If* $(pk_{\mathsf{lossy}}, sk_{\mathsf{lossy}}) \xleftarrow{\$} G(1^\lambda, \mathsf{lossy})$, *and* $r \xleftarrow{\$} \mathsf{coins}(E)$, *then for all* $x_0, x_1 \in X$ *with overwhelming probability, there exists* $r' \in \mathsf{coins}(E)$ *such that* $E(pk_{\mathsf{lossy}}, x_0, r) = E(pk_{\mathsf{lossy}}, x_1, r')$. *Hence, there is an unbounded algorithm* opener *that can open a lossy ciphertext to* any *plaintext.*

Although openability is implied by property (3), it is convenient to state it explicitly in terms of an algorithm. In [5], it was shown that, if the algorithm opener is efficient, then the encryption scheme is actually SEM-SO-CPA secure. We do not explicitly require schemes to be IND-CPA secure since semantic security follows from the indistinguishability of keys and lossiness of the lossy keys. In [5], it was shown that the IND-CPA secure cryptosystem based on Lossy Trapdoor Functions given in [42], is in fact a Lossy Encryption. Next, they proved that any Lossy Encryption scheme where the plaintext space admits a n-message sampler with efficient resampling is IND-SO-CPA secure.

3 Constructing Lossy Encryption Schemes

3.1 Re-Randomizable Encryption Implies Lossy Encryption

In many cryptosystems, given a ciphertext c and a public-key, it is possible to re-randomize c to a new ciphertext c' such that c and c' encrypt the same plaintext but are statistically independent. We call a public key cryptosystem given by algorithms (G, E, D) *statistically re-randomizable*[5] if

- (G, E, D) is semantically-secure in the standard sense (IND-CPA).
- There is negligible function ν, and an efficient function ReRand such that for all λ, pk, m, r_1 we have $\Delta(\{r_0 \overset{\$}{\leftarrow} \mathsf{coins}(E) : E(pk, m, r_0)\}, \{r' \overset{\$}{\leftarrow} \mathsf{coins}(\mathsf{ReRand}) : \mathsf{ReRand}(E(pk, m, r_1), r')\}) < \nu(\lambda)$.

Since re-randomization does not require any kind of group structure on the plaintext space or any method for combining ciphertexts, re-randomizable encryption appears to be a weaker primitive than homomorphic encryption, and all known homomorphic cryptosystems are re-randomizable.

Our first result is a simple lossy encryption system $(\bar{G}_{\mathsf{inj}}, \bar{G}_{\mathsf{lossy}}, \bar{E}, \bar{D})$ obtained from a statistically re-randomizable public-key cryptosystem (G, E, D).

- **Key Generation:** first, $\bar{G}(1^\lambda, \mathsf{inj})$ generates $(pk, sk) \leftarrow G(1^\lambda)$. Then, it picks $r_0, r_1 \overset{\$}{\leftarrow} \mathsf{coins}(E)$, computes $e_0 = E(pk, 0, r_0)$, $e_1 = E(pk, 1, r_1)$ and returns $(\bar{pk}, \bar{sk}) = ((pk, e_0, e_1), sk)$. Algorithm $\bar{G}(1^\lambda, \mathsf{lossy})$ runs $G(1^\lambda)$, generating a pair (pk, sk). Then, it picks $r_0, r_1 \overset{\$}{\leftarrow} \mathsf{coins}(E)$ and generates $e_0 = E(pk, 0, r_0)$, $e_1 = E(pk, 0, r_1)$. It returns $(\bar{pk}, \bar{sk}) = ((pk, e_0, e_1), sk)$.
- **Encryption:** $\bar{E}(\bar{pk}, b, r') = \mathsf{ReRand}(pk, e_b, r')$ for $b \in \{0, 1\}$.
- **Decryption** $\bar{D}(\bar{sk}, c)$, simply outputs $D(sk, c)$.

It is not hard to show that this construction is a lossy encryption scheme, as formally proved in the full version of the paper. Although it only allows encrypting single bits, it can be easily modified to encrypt longer messages if the underlying cryptosystem is homomorphic and if the set of encryptions of zero can be almost uniformly sampled (the details are available in the full paper).

We also note that specific homomorphic cryptosystems such as Paillier [39] or Damgård-Jurik [20] provide more efficient constructions where multi-bit messages can be encrypted. In addition, as shown in the full version of the paper, the factorization of the modulus N provides a means for efficiently opening a lossy ciphertext to any plaintext. Thus this scheme is actually SEM-SO-CPA secure when instantiated with these cryptosystems. This provides the most efficient known examples of SEM-SO-CPA secure cryptosystems. Previously, the most efficient known SEM-SO-CPA secure construction was the Goldwasser-Micali cryptosystem [5] which can only encrypt single bits.

[5] This definition of re-randomizable encryption requires statistical re-randomization. It is possible to define re-randomizable encryption which satisfies perfect re-randomization (stronger) or computational re-randomization (weaker). Such definitions already exist in the literature (see for example [40, 25, 29, 14]). Our constructions require statistical re-randomization, and do not go through under a computational re-randomization assumption.

3.2 Statistically-Hiding $\binom{2}{1}$-OT Implies Lossy Encryption

Honest-receiver two-round statistically-hiding $\binom{2}{1}$-oblivious transfer is a protocol between a sender Sen and a receiver $\mathsf{Rec} = (\mathsf{Rec}_q, \mathsf{Rec}_r)$. The former has two strings s_0, s_1 and the latter has a bit b. The receiver Rec_q generates a query q, which is sent to Sen, along with some state information sk. The sender evaluates $\mathsf{q}(s_0, s_1)$ and sends the result $\mathsf{rsp} = \mathsf{Sen}(\mathsf{q}, s_0, s_1)$ to Rec_r who uses sk to get s_b.

- **Correctness:** For all $s_0, s_1 \in \{0,1\}^k$, $b \in \{0,1\}$, there exists $\nu \in \mathsf{negl}(\lambda)$ s.t.

  $\Pr[(\mathsf{q}, sk) \xleftarrow{\$} \mathsf{Rec}_q(1^\lambda, b);\ \mathsf{rsp} \xleftarrow{\$} \mathsf{Sen}(\mathsf{q}, s_0, s_1) : \mathsf{Rec}_r(sk, \mathsf{rsp}) = s_b] \geq 1 - \nu(\lambda)$.
- **Receiver Privacy:** b remains computationally hidden from Sen's view. That is, we must have $\{(\mathsf{q}, sk) \xleftarrow{\$} \mathsf{Rec}_q(1^\lambda, 0) : \mathsf{q}\} \approx_c \{(\mathsf{q}, sk) \xleftarrow{\$} \mathsf{Rec}_q(1^\lambda, 1) : \mathsf{q}\}$, where the distributions are taken over the internal randomness of Rec_q.
- **Sender Privacy:** for any $b \in \{0,1\}$, for any strings s_0, s_1, s_0', s_1' such that $s_b = s_b'$ and any honest receiver's query $\mathsf{q} = \mathsf{Rec}_q(1^\lambda, b)$, it must hold that

  $$\{(\mathsf{q}, sk) \xleftarrow{\$} \mathsf{Rec}_q(1^\lambda, b); \mathsf{rsp} \xleftarrow{\$} \mathsf{Sen}(\mathsf{q}, s_0, s_1) : \mathsf{rsp}\}$$
  $$\approx_s \{(\mathsf{q}, sk) \xleftarrow{\$} \mathsf{Rec}_q(1^\lambda, b); \mathsf{rsp} \xleftarrow{\$} \mathsf{Sen}(\mathsf{q}, s_0', s_1') : \mathsf{rsp}\},$$

  the distributions being taken over the internal randomness of Rec_q and Sen.

A two-round honest-receiver statistically-hiding $\binom{2}{1}$-OT $(\mathsf{Sen}, \mathsf{Rec})$ gives a lossy encryption as follows:

- **Key Generation:** Define $G(1^\lambda, \mathsf{inj}) = \mathsf{Rec}_q(1^\lambda, 0)$. Set $pk = \mathsf{q}$, and $sk = sk$.

  Define $G(1^\lambda, \mathsf{lossy}) = \mathsf{Rec}_q(1^\lambda, 1)$. Set $pk = \mathsf{q}$, and $sk = \perp$.
- **Encryption:** Define $E(pk, m, (r, r^*)) = \mathsf{Sen}(\mathsf{q}, m, r; r^*)$, where r^* is the randomness used in $\mathsf{Sen}(\mathsf{q}, m, r)$ and $r \xleftarrow{\$} \{0,1\}^{|m|}$ is a random string.
- **Decryption:** given $c = \mathsf{rsp}$ in injective mode, define $D(sk, \mathsf{rsp}) = \mathsf{Rec}_r(sk, \mathsf{rsp})$.

Lemma 1. *The scheme (G, E, D) forms a lossy encryption scheme.*

The (straightforward) proof of Lemma 1 can be found in the full version of this paper. Since single-server Private Information Retrieval (PIR) implies statistically-hiding OT [15], we find the following corollary.

Corollary 1. *One-round Single-Server PIR implies Lossy Encryption.*

Since homomorphic encryption implies PIR [33, 35], the following result follows.

Corollary 2. *Homomorphic encryption implies Lossy Encryption.*

In the half simulation model, statistically hiding $\binom{2}{1}$-OT can rely [30, 26] on smooth hash proof systems that fit a slight modification of the original definition [17] with suitable verifiability properties. In the honest-but-curious receiver setting (which suffices here), it was already noted in [26][Section 1.3] that ordinary hash proof systems are sufficient to realize $\binom{2}{1}$-OT. In the full version of the paper, we describe a simplification of the construction of lossy encryption from hash proof systems and obtain the next result.

Corollary 3. *Smooth projective hash functions imply Lossy Encryption.*

To summarize this section, since lossy encryption is selective-opening secure, we obtain the following theorem.

Theorem 1. *Statistically-hiding 2-round honest-receiver $\binom{2}{1}$-OT, single server PIR, smooth projective hash proof systems and homomorphic encryption all imply IND-SO-CPA secure encryption.*

4 Chosen-Ciphertext Security

When an adversary has access to a decryption oracle, many cryptosystems become insecure. The notion of chosen-ciphertext security [38, 43, 19] was created to address this issue and, since then, many schemes have achieved this security level. The attacks of Bleichenbacher on RSA PKCS#1 [6] emphasized the importance of security against chosen-ciphertext attacks (CCA).

The need for selective opening security was first recognized in the context of Multi-Party Computation (MPC), where an active MPC adversary can view all ciphertexts sent in a current round and then choose a subset of senders to corrupt. It is natural to imagine an adversary who, in addition to corrupting a subset of senders, can also mount a chosen-ciphertext attack against the receiver. Schemes proposed so far (based on re-randomizable encryption or described in [5]) are obviously insecure in this scenario.

In this section, we extend the notion of chosen-ciphertext security to the selective opening setting. As in the standard selective-opening setting, we can define security either by indistinguishability, or by simulatability. We will give definitions of security as well as constructions for both settings.

Classical techniques to acquire chosen-ciphertext security are delicate to use here. Handling decryption queries using the Naor-Yung paradigm [38] and non-interactive zero-knowledge proofs [44] is not straightforward as, when the adversary makes her corruption query, it should obtain the random coins that were used to produce NIZK proofs. Fehr, Hofheinz, Kiltz and Wee [23] showed how to use non-committing encryption [10] along with a modified hash proof system [17] to achieve chosen-ciphertext security in the selective opening setting in the simulation-based model (SEM-SO-CCA). Our work takes a different approach and seeks to apply the Canetti-Halevi-Katz paradigm [11]. As we shall see, adapting this methodology to the selective opening setting encounters a number of technical obstacles that need to be overcome.

4.1 Chosen-Ciphertext Security: Indistinguishability

We begin with the indistinguishability-based definition. We define a real game (ind-cca2-real) and an ideal game (ind-cca2-ideal). In both games, the challenger generates a key pair $(sk, pk) \leftarrow G(1^\lambda)$ and sends pk to $\mathcal{A}$. The adversary is then allowed to adaptively make the following types of queries.

- **Challenge Query:** let $\mathcal{M}$ be a message sampler. The latter samples a vector $\mathbf{m} = (m_1, \ldots, m_n) \overset{\$}{\leftarrow} \mathcal{M}$ and returns a vector containing n "target" ciphertexts $\mathbf{C} = (\mathbf{C}[1], \ldots, \mathbf{C}[n]) \leftarrow (E(pk, m_1, r_1), \ldots, E(pk, m_n, r_n))$.
- **Corrupt Query:** $\mathcal{A}$ chooses $I \subset \{1, \ldots, n\}$ and receives $\{(m_i, r_i)\}_{i \in I}$.
  - In $\mathsf{ind\text{-}cca2\text{-}real}$, the challenger then sends $\{m_j\}_{j \notin I}$ to the adversary.
  - In $\mathsf{ind\text{-}cca2\text{-}ideal}$, the challenger re-samples $\mathbf{m}' = (m'_1, \ldots, m'_n) \overset{\$}{\leftarrow} \mathcal{M}_{|I, \mathbf{m}[I]}$ (*i.e.*, so that $m'_j = m_j$ for each $j \in I$) and sends $\{m'_j\}_{j \notin I}$ to $\mathcal{A}$.
- **Decryption Queries:** $\mathcal{A}$ chooses a ciphertext C such that $C \neq \mathbf{C}[i]$ for each $i \in \{1, \ldots, n\}$ and sends C to the challenger which responds with $D(sk, C)$.

After polynomially-many queries, one of which is a challenge query and precedes the corrupt query (which is unique as well), the adversary outputs $b \in \{0, 1\}$.

Definition 3. *A public key cryptosystem is IND-SO-CCA2 secure if, for any polynomial n and any n-message sampler $\mathcal{M}$ supporting efficient conditional re-sampling, any PPT adversary $\mathcal{A}$ has negligibly different outputs in the real game and in the ideal game: for some negligible function ν, we must have*

$$\left| \Pr[\mathcal{A}^{\mathsf{ind\text{-}cca2\text{-}real}} = 1] - \Pr[\mathcal{A}^{\mathsf{ind\text{-}cca2\text{-}ideal}} = 1] \right| < \nu.$$

4.2 Chameleon Hash Functions

A chameleon hash function [34] $\mathcal{CMH} = (\mathsf{CMKg}, \mathsf{CMhash}, \mathsf{CMswitch})$ consists of an algorithm CMKg that, given a security parameter λ, outputs a key pair $(hk, tk) \overset{\$}{\leftarrow} \mathcal{G}(\lambda)$. The hashing algorithm outputs $y = \mathsf{CMhash}(hk, m, r)$ given the public key hk, a message m and random coins $r \in \mathcal{R}_{hash}$. On input of m, r, m' and the trapdoor key tk, the switching algorithm $r' \leftarrow \mathsf{CMswitch}(tk, m, r, m')$ outputs $r' \in \mathcal{R}_{hash}$ such that $\mathsf{CMhash}(hk, m, r) = \mathsf{CMhash}(hk, m', r')$. Collision-resistance mandates that it be infeasible to find pairs $(m', r') \neq (m, r)$ such that $\mathsf{CMhash}(hk, m, r) = \mathsf{CMhash}(hk, m', r')$ without knowing tk. Uniformity guarantees that the distribution of hashes is independent of the message m, in particular, for all hk, and m, m', the distributions $\{r \leftarrow \mathcal{R}_{hash} : \mathsf{CMHash}(hk, m, r)\}$ and $\{r \leftarrow \mathcal{R}_{hash} : \mathsf{CMHash}(hk, m', r)\}$ are identical. It is well-known that chameleon hashing can be based on standard number theoretic assumptions.

4.3 A Special Use of the Canetti-Halevi-Katz Paradigm

The Canetti-Halevi-Katz technique [11] allows building chosen-ciphertext secure cryptosystems from weakly secure identity-based or tag-based encryption scheme. A tag-based encryption scheme (TBE) [36, 31] is a cryptosystem where the encryption and decryption algorithms take an additional input, named the *tag*, which is a binary string of appropriate length with no particular structure. A TBE scheme consists of a triple $\mathcal{TBE} = (\mathsf{TBEKg}, \mathsf{TBEEnc}, \mathsf{TBEDec})$ of efficient algorithms where, on input of a security parameter λ, TBEKg outputs a private/public key pair (pk, sk); TBEEnc is a randomized algorithm that outputs a ciphertext C on input of a public key pk, a string θ – called *tag* – and a message $m \in \mathsf{MsgSp}(\lambda)$; $\mathsf{TBEDec}(sk, \theta, C)$ is the decryption algorithm that takes as input a secret key sk, a tag θ and a ciphertext C and returns a plaintext m or $\perp$. Associated with $\mathcal{TBE}$ is a plaintext space MsgSp. Correctness requires that for all $\lambda \in \mathbb{N}$, all key pairs $(pk, sk) \leftarrow \mathsf{TBEKg}(1^\lambda)$, all tags θ and any plaintext $m \in \mathsf{MsgSp}(\lambda)$, it holds that $\mathsf{TBEDec}(sk, \theta, \mathsf{TBEEnc}(pk, \theta, M)) = m$.

Selective Opening Security for TBE Schemes. In the selective opening setting, the weak CCA2 security definition of [31] can be extended as follows.

Definition 4. *A TBE scheme* $\mathcal{TBE} = (\mathsf{TBEKg}, \mathsf{TBEEnc}, \mathsf{TBEDec})$ *is selective-tag weakly* IND-SO-CCA2 secure *(or IND-SO-stag-wCCA2 secure) if, for any polynomial n and any n-message sampler $\mathcal{M}$ supporting efficient conditional re-sampling, any PPT adversary $\mathcal{A}$ produces negligibly different outputs in the real and ideal games, which are defined as follows.*

1. *The adversary $\mathcal{A}$ chooses n tags $\theta_1^\star, \ldots, \theta_n^\star$ and sends them to the challenger.*
2. *The challenger generates a key pair* $(sk, pk) \leftarrow \mathsf{TKEKg}(1^\lambda)$ *and hands pk to $\mathcal{A}$. The latter then adaptively makes the following kinds of queries:*
   - ***Challenge Query:*** *let $\mathcal{M}$ be a message sampler for* $\mathsf{MsgSp}(\lambda)$*. The challenger samples* $(m_1, \ldots, m_n) \xleftarrow{\$} \mathcal{M}$ *and returns* $\mathbf{C} = (\mathbf{C}[1], \ldots, \mathbf{C}[n])$*, where* $\mathbf{C}[i] = \mathsf{TBEEnc}(pk, \theta_i^\star, m_i, r_i)$
   - ***Corrupt Query:*** *$\mathcal{A}$ chooses $I \subset \{1, \ldots, n\}$ and obtains $\{(m_i, r_i)\}_{i \in I}$.*
     - *In the real game, the challenger then sends $\{m_j\}_{j \notin I}$ to the adversary.*
     - *In the ideal game, the challenger re-samples* $(m_1', \ldots, m_n') \xleftarrow{\$} \mathcal{M}_{|I, \mathbf{m}[I]}$ *and reveals $\{m_j'\}_{j \notin I}$.*
   - ***Decryption Queries:*** *$\mathcal{A}$ sends a pair (C, θ) such that $\theta \notin \{\theta_1^\star, \ldots, \theta_n^\star\}$. The challenger replies with* $\mathsf{TBEDec}(sk, \theta, C) \in \mathsf{MsgSp}(\lambda) \cup \{\perp\}$*.*

After polynomially-many queries, one of which is a challenge query, $\mathcal{A}$ outputs $b \in \{0,1\}$. Its advantage $\mathbf{Adv}_{\mathcal{A}}^{\text{IND-SO-stag-wCCA2}}(\lambda)$ *is defined as in definition 3.*

At first, one may hope to obtain IND-SO-CCA2 security by applying the CHK method [11] to any IBE/TBE scheme satisfying some weaker level of selective opening security. Let $\mathcal{TBE} = (\mathsf{TBEKg}, \mathsf{TBEEnc}, \mathsf{TBEDec})$ be a secure TBE scheme in the sense of definition 4 and let $\Sigma = (\mathcal{G}, \mathcal{S}, \mathcal{V})$ be a strong one-time signature. The CHK technique turns $\mathcal{TBE}$ into a cryptosystem $\mathcal{PKE} = (G, E, D)$ which is obtained by letting $G(1^\lambda)$ output $(sk', (\Sigma, pk'))$ where $(sk', pk') \leftarrow \mathsf{TBEKg}(1^\lambda)$. To encrypt a message m, E generates a one-time signature key pair $(\mathsf{SK}, \mathsf{VK}) \leftarrow \mathcal{G}(1^\lambda)$, computes $C_{tbe} = \mathsf{TBEEnc}(pk, \mathsf{VK}, m)$ under the tag VK and sets the $\mathcal{PKE}$ ciphertext as $(\mathsf{VK}, C_{tbe}, \sigma)$, where $\sigma = \mathcal{S}(\mathsf{SK}, C_{tbe})$.

In the selective opening setting, when the adversary makes its corruption query in the reduction, it must obtain the random coins that were used to generate one-time signature keys appearing target ciphertexts. Then, it is able to re-compute the corresponding private keys and make decryption queries for ciphertexts involving the same verification keys as target ciphertexts, which causes the reduction to fail. Although schemes using one-time signatures do not appear to become trivially insecure, the reduction of [11, 31] ceases to go through.

It was showed in [46] that chameleon hash functions [34] can be used to turn certain TBE schemes, termed *separable*, into full-fledged IND-CCA2 cryptosytems and supersede one-time signatures in the CHK transform. A TBE scheme is said *separable* if, on input of pk, m, θ, algorithm $\mathsf{TBEEnc}(pk, t, m)$ uses randomness $r \in \mathcal{R}_{tbe}$ and returns $C_{tbe} = (f_1(pk, m, r), f_2(pk, r), f_3(pk, \theta, r))$, where functions f_1, f_2 and f_3 are computed independently of each other and are all deterministic (so that they give the same outputs when queried twice on the same (m, r), r and (θ, r)). In addition, f_2 must be injective.

The construction of [46][6] uses chameleon hashing instead of one-time signatures. Key generation requires to create a TBE key pair (pk', sk') and a chameleon hashing public key hk. The private key of $\mathcal{PKE}$ is the TBE private key sk'. Encryption and decryption procedures are depicted hereafter.

```
$E(m, pk)$
  Parse $pk$ as $(pk', hk)$
  $r_1 \leftarrow \mathcal{R}_{tbe}$; $r_2 \leftarrow \mathcal{R}_{hash}$
  $u = f_1(pk', m, r_1)$; $v = f_2(pk', r_1)$
  $\theta = \mathsf{CMhash}(hk, u||v, r_2)$
  $w = f_3(pk', \theta, r_1)$
  Return $C = (u, v, w, r_2)$
```

```
$D(sk, C)$
  Parse $C$ as $(u, v, w, r_2)$ and $sk$ as $sk'$
  $\theta = \mathsf{CMhash}(hk, u||v, r_2)$
  Return $m \leftarrow \mathsf{TBEDec}(sk', \theta, (u, v, w))$
```

Unlike the CHK transform, this construction computes C without using any other secret random coins than those of the underlying TBE ciphertext. The tag is derived from a ciphertext component u and some independent randomness r_2 that *publicly* appears in C. For this reason, we can hope to avoid the difficulty that appears with the CHK transform. Indeed, we prove that any separable TBE that satisfies definition 4 yields an IND-SO-CCA2 cryptosystem.

Theorem 2. *If* $\mathcal{TBE} = (\mathsf{TBEKg}, \mathsf{TBEEnc}, \mathsf{TBEDec})$ *is a separable TBE scheme with IND-SO-stag-wCCA2 security, the transformation of figure* **??** *gives an IND-SO-CCA2 PKE scheme.* (The proof is given in the full version of the paper).

4.4 Lossy and All-But-n Trapdoor Functions

A tuple $(S_{\mathrm{ltdf}}, F_{\mathrm{ltdf}}, F_{\mathrm{ltdf}}^{-1})$ of PPT algorithms is called a family of (d, k)-lossy trapdoor functions [42] if the following properties hold:

Sampling injective functions: $S_{\mathrm{ltdf}}(1^\lambda, 1)$ outputs (s, t), where s is a function index and t its trapdoor. It is required that $F_{\mathrm{ltdf}}(s, \cdot)$ be injective on $\{0, 1\}^d$ and $F_{\mathrm{ltdf}}^{-1}(t, F_{\mathrm{ltdf}}(s, x)) = x$ for all x.

[6] As described in [46], the construction uses a single function F instead of f_1 and f_2 (*i.e.*, we are re-writing it in the particular case $F(m, r) = (f_1(pk, m, r), f_2(pk, r))$). The security proof of [46] implicitly requires F to be such that no two pairs $(m, r) \neq (m', r')$ give $F(m, r) = F(m', r')$. Using functions f_1, f_2 is a way to enforce this.

Sampling lossy functions: $S_{\mathrm{ltdf}}(1^\lambda, 0)$ outputs $(s, \perp)$ where s is a function index and $F_{\mathrm{ltdf}}(s, \cdot)$ is a function on $\{0,1\}^d$ with image size at most 2^{d-k}.

Indistinguishability: $\{(s,t) \xleftarrow{\$} S_{\mathrm{ltdf}}(1^\lambda, 1) : s\} \approx_c \{(s, \perp) \xleftarrow{\$} S_{\mathrm{ltdf}}(1^\lambda, 0) : s\}$.

Along with lossy trapdoor functions, Peikert and Waters [42] defined all-but-one (ABO) functions. These are lossy trapdoor functions, except instead of having two branches (a lossy branch and an injective branch) they have many branches coming from a branch set $\mathcal{B}$, all but one of which are injective.

The Peikert-Waters system only requires ABO functions to have one lossy branch because the IND-CCA2 game involves a single challenge ciphertext and a single ABO function must be evaluated on a lossy branch. Since the IND-SO-CCA security game involves $n > 1$ challenge ciphertexts, we need to generalize ABO functions into all-but-n (ABN) functions that have multiple lossy branches and where all branches except the specified ones are injective. A tuple $(S_{\mathrm{abn}}, G_{\mathrm{abn}}, G_{\mathrm{abn}}^{-1})$ is a family of ABN functions if these conditions are satisfied.

- **Sampling with a given lossy set:** For any n-subset $I \subset \mathcal{B}$, $S_{\mathrm{abn}}(1^\lambda, I)$ outputs s, t where s is a function index, and t its trapdoor. We require that for any $b \in \mathcal{B} \setminus I$, $G_{\mathrm{abn}}(s, b, \cdot)$ is an injective deterministic function on $\{0,1\}^d$, and $G_{\mathrm{abn}}^{-1}(t, b, G_{\mathrm{abn}}(s, b, x)) = x$ for all x. Additionally, for each $b \in I$, the image $G_{\mathrm{abn}}(s, b, \cdot)$ has size at most 2^{d-k}.
- **Hidden lossy sets:** For any distinct n-subsets $I_0^\star, I_1^\star \subset \mathcal{B}$, the first outputs of $S_{\mathrm{abn}}(1^\lambda, I_0^\star)$ and $S_{\mathrm{abn}}(1^\lambda, I_1^\star)$ are computationally indistinguishable.

Just as ABO functions can be obtained from lossy trapdoor functions [42], ABN functions can also be constructed from LTDFs and a general construction is provided in the full version of the paper. The recent results of Hofheinz [28], show how to create All-But-Many Lossy Functions, which are Lossy Trapdoor Functions with a super-polynomial number of lossy branches. The advantage of his construction is that the description of the function is independent of N. Hofheinz's All-But-Many functions can be plugged into our constructions to shrink the size of the public-key in our constructions (see [28] for details).

4.5 An IND-SO-stag-wCCA2 TBE Construction

We construct IND-SO-stag-wCCA2 tag-based cryptosystems from lossy trapdoor functions. Let $(\mathsf{CMKg}, \mathsf{CMhash}, \mathsf{CMswitch})$ be a chameleon hash function where CMhash ranges over the set of branches $\mathcal{B}$ of the ABN family. We eventually obtain an IND-SO-CCA2 public key encryption scheme as a LTDF-based construction that mimics the one [42] (in its IND-CCA1 variant).

Let $(S_{\mathrm{ltdf}}, F_{\mathrm{ltdf}}, F_{\mathrm{ltdf}}^{-1})$ be a family of (d,k)-lossy-trapdoor functions, and let $(S_{\mathrm{abn}}, G_{\mathrm{abn}}, G_{\mathrm{abn}}^{-1})$ be a family of (d, k') all-but-n functions with branch set $\{0,1\}^v$ where v is the length of a verification key for a one-time signature. We require that $2d - k - k' \leq t - \kappa$, for $\kappa = \kappa(t) = \omega(\log t)$. Let $\mathcal{H}$ be a pairwise independent hash family from $\{0,1\}^d \to \{0,1\}^\ell$, with $0 < \ell < \kappa - 2\log(1/\nu)$, for some negligible $\nu = \nu(\lambda)$. The message space will be $\mathsf{MsgSp} = \{0,1\}^\ell$.

- $\mathsf{TBEKg}(1^\lambda)$: choose $h \xleftarrow{\$} \mathcal{H}$ in the pairwise independent hash family and generate $(s,t) \leftarrow S_{\mathrm{ltdf}}(1^\lambda, inj)$, $(s',t') \leftarrow S_{\mathrm{abn}}(1^\lambda, \{0,1,\ldots,n-1\})$. The public key will be $pk = (s,s',h)$ and the secret key will be $sk = (t,t')$.
- $\mathsf{TBEEnc}(m, pk, \theta)$: to encrypt $m \in \{0,1\}^\ell$ under the tag $\theta \in \mathcal{B}$, choose $x \xleftarrow{\$} \{0,1\}^d$. Compute $c_0 = h(x) \oplus m$, $c_1 = F_{\mathrm{ltdf}}(s,x)$ and $c_2 = G_{\mathrm{abn}}(s,\theta,x)$ and the TBE ciphertext is $C = \big(c_0,c_1,c_2\big) = \big(h(x) \oplus m,\ F_{\mathrm{ltdf}}(s,x),\ G_{\mathrm{abn}}(s',\theta,x)\big)$.
- $\mathsf{TBEDec}(C, sk, \theta)$: given $C = \big(c_0,c_1,c_2\big)$ and $sk = t$, compute $x = F_{\mathrm{ltdf}}^{-1}(t,c_1)$ and $m = c_0 \oplus h(x)$ if $G_{\mathrm{abn}}(s,\theta,x) = c_2$. Otherwise, output $\perp$.

The scheme is separable since C is obtained as $c_0 = f_1(pk,m,x) = m \oplus h(x)$, $c_1 = f_2(pk,x) = F_{\mathrm{ltdf}}(s,x)$ and $c_2 = f_3(pk,\theta,x) = G_{\mathrm{abn}}(s',\theta,x)$.

Theorem 3. *The algorithms described above form an IND-SO-stag-wCCA2 secure tag-based cryptosystem assuming the security of the lossy and all-but-n families.* (The proof is given in the full version of the paper).

4.6 An All-but-n Function with Short Outputs

While generic, the all-but-n function described in the full version of the paper has the disadvantage of long outputs, the size of which is proportional to nk. Efficient all-but-one functions can be based on the Composite Residuosity assumption [22, 3]. We show that the all-but-one function of [22, 3] extends into an ABN function that retains short (*i.e.*, independent of n or k) outputs. Multiple lossy branches can be obtained using a technique that traces back to the work of Chatterjee and Sarkar [18] who used it in the context of identity-based encryption.

- **Sampling with a given lossy set:** given a security parameter $\lambda \in \mathbb{N}$ and the desired lossy set $I = \{\theta_1^\star, \ldots, \theta_n^\star\}$, where $\theta_i^\star \in \{0,1\}^\lambda$ for each $i \in \{1,\ldots,n\}$, let $\gamma \geq 4$ be a polynomial in λ.
  1. Choose random primes p,q s.t. $N = pq > 2^\lambda$.
  2. Generate a vector $\vec{U} \in (\mathbb{Z}_{N^{\gamma+1}}^*)^{n+1}$ as follows. Let $\alpha_{n-1},\ldots,\alpha_0 \in \mathbb{Z}_{N^\gamma}$ be coefficients of $P[T] = \prod_{i=1}^n (T - \theta_i^\star) = T^n + \alpha_{n-1}T^{n-1} + \cdots + \alpha_1 T + \alpha_0$ in $\mathbb{Z}_{N^\gamma}[T]$ (note that $P[T]$ is expanded in $\mathbb{Z}_{N^\gamma}$ but its roots are all in $\mathbb{Z}_N^*$). Then, for each $i \in \{0,\ldots,n\}$, set $U_i = (1+N)^{\alpha_i} a_i^{N^\gamma} \bmod N^{\gamma+1}$, where $(a_0,\ldots,a_n) \xleftarrow{\$} (\mathbb{Z}_N^*)^{n+1}$ and with $\alpha_n = 1$.
  3. The evaluation key is $s' = \{N, \vec{U} = (U_0,\ldots,U_n)\}$ and the domain of the function is $\{0,\ldots,2^{\gamma\lambda/2}-1\}$. The trapdoor is $t' = \mathrm{lcm}(p-1,q-1)$.
- **Evaluation:** to evaluate $G_{\mathrm{abn}}(s',\theta,x)$, where $x \in \{0,\ldots,2^{\gamma\lambda/2}-1\}$ and $\theta \in \{0,1\}^\lambda$, compute $c = \big(\prod_{j=0}^n U_i^{(\theta^i \bmod N^\gamma)}\big)^x \bmod N^{\gamma+1}$.
- **Inversion:** for a branch θ, $c = G_{\mathrm{abn}}(s',\theta,x)$ is a Damgård-Jurik encryption of $y = P(\theta)x \bmod N^\gamma$. Using $t' = \mathrm{lcm}(p-1,q-1)$, we apply the decryption algorithm of [20] to obtain $y \in \mathbb{Z}_{N^\gamma}$ and return $x = yP(\theta)^{-1} \bmod N^\gamma$.

As in [22, 3], $G_{\mathrm{abn}}(s', \theta, \cdot)$ has image size smaller than N when $\theta \in I$ and it can be shown that $\tilde{H}_\infty\big(x|(G_{\mathrm{abn}}(s', \theta, x), N, \vec{U})\big) \geq \gamma\lambda/2 - \log(N)$.

We note that the ABN function $G_{\mathrm{abn}}(s', \theta, \cdot)$ is not injective for each branch $\theta \notin I$, but only for those such that $\gcd(P(\theta), N^\gamma) = 1$. However, the fraction of branches $\theta \in \{0,1\}^\lambda$ such that $\gcd(P(\theta), N^\gamma) \neq 1$ is bounded by $2/\min(p, q)$, which is negligible. Moreover, the proof of theorem 3 is not affected if the TBE scheme is instantiated with this ABN function and the LTDF of [22, 3]. As explained in the full version of the paper, as long as factoring is hard (which is implied by the Composite Residuosity assumption), the adversary has negligible chance of making decryption queries w.r.t. to such a problematic tag θ.

Lemma 2. *The above ABN function is lossy set hiding under the Composite Residuosity assumption.* (The proof is given in the full version of the paper).

The above ABN function yields an IND-SO-CCA2 secure encryption scheme with ciphertexts of constant (*i.e.*, independent of n) size but a public key of size $\mathcal{O}(n)$. Encryption and decryption require $\mathcal{O}(n)$ exponentiations as they entail an ABN evaluation. On the other hand, the private key has $\mathcal{O}(1)$ size, which keeps the private storage very cheap. At the expense of sacrificing the short private key size, we can optimize the decryption algorithm by computing $x = G_{\mathrm{abn}}^{-1}(t', \theta, c_2)$ (instead of $x = F_{\mathrm{ltdf}}^{-1}(t, c_1)$) so as to avoid computing $G_{\mathrm{abn}}(s', \theta, x)$ in the forward direction to check the validity of ciphertexts. In this case, the receiver has to store $\alpha_0, \ldots, \alpha_{n-1}$ to evaluate $P(\theta)$ when inverting G_{abn}.

It is also possible to extend the DDH-based ABO function described in [42] into an ABN function. However, in the full version of the paper, we describe a more efficient lossy TBE scheme based on the DDH assumption.

Acknowledgements

We thank Yuval Ishai for suggesting a connection between Oblivious Transfer and Lossy Encryption.

Brett Hemenway was supported in part by NSF VIGRE Fellowship and NSF grants 0716835, 0716389, 0830803 and 0916574. Rafail Ostrovsky's research is supported in part by NSF grants 0830803, 09165174, 106527 and 1118126, US-Israel BSF grant 2008411, grants from OKAWA Foundation, IBM, Lockheed-Martin Corporation and the Defense Advanced Research Projects Agency through the U.S. Office of Naval Research under Contract N00014-11-1-0392. The views expressed are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Benoît Libert acknowledges the Belgian Fund for Scientific Research (F.R.S.-F.N.R.S.) for his "Chargé de recherches" fellowship and the BCRYPT Interuniversity Attraction Pole. Damien Vergnaud was supported in part by the European Commission through the ICT Program under contract ICT-2007-216676 ECRYPT II.

References

1. D. Boneh, X. Boyen. Efficient selective-ID secure identity-based encryption without Random Oracles. In *Eurocrypt'04*, *LNCS* 3027, pp. 223–238, 2004.
2. D. Beaver. Plug and play encryption. In *Crypto'97*, *LNCS* 1294, pp. 75–89, Springer, 1997.
3. A. Boldyreva, S. Fehr, A. O'Neill. On notions of security for deterministic encryption, and efficient constructions without random oracles. In *Crypto'08*, *LNCS* 5157, pp. 335–359. Springer, 2008.
4. D. Beaver, S. Haber. Cryptographic protocols provably secure against dynamic adversaries. In *Eurocrypt'92*, *LNCS* 658, pp. 307–323. Springer, 1992.
5. M. Bellare, D. Hofheinz, S. Yilek. Possibility and impossibility results for encryption and commitment secure under selective opening. In *Eurocrypt'09*, *LNCS* 5479, pp. 1–35. Springer, 2009.
6. D. Bleichenbacher. Chosen ciphertext attacks against protocols based on the RSA encryption standard PKCS #1. In *Crypto'98*, *LNCS* 1462, pp. 1–12. Springer, 1998.
7. M. Bellare, B. Waters, S. Yilek. Identity-based encryption secure under selective opening attack. In *TCC'11*, *LNCS* 6597, pp. 235–252. Springer, 2011.
8. M. Bellare, S. Yilek. Encryption schemes secure under selective opening attack. Cryptology ePrint Archive: Report 2009/101, 2009.
9. R. Canetti, C. Dwork, M. Naor, R. Ostrovsky. Deniable encryption. In *Crypto'97*, *LNCS* 1294, pp. 90–104, Springer, 1997.
10. R. Canetti, U. Feige, O. Goldreich, M. Naor. Adaptively secure multi-party computation. In *STOC '96*, pp. 639–648, ACM Press, 1996.
11. R. Canetti, S. Halevi, J. Katz. Chosen-ciphertext security from identity-based encryption. In *Eurocrypt'04*, pp. 207–222. Springer, 2004.
12. —, Adaptively-secure, non-interactive public-key encryption. In *TCC '05*, *LNCS* 3378, pp. 150–168. Springer, 2005.
13. G. Di Crescenzo, Y. Ishai, R. Ostrovsky. Non-interactive and non-malleable commitment. In *STOC'98*. ACM, 1998.
14. R. Canetti, H. Krawczyk, J.-B. Nielsen. Relaxing chosen-ciphertext security. In *Crypto'03*, *LNCS* 2729, pp. 565–582. Springer, 2003.
15. G. Di Crescenzo, T. Malkin, R. Ostrovsky. Single database private information retrieval implies oblivious transfer. In *Eurocrypt'00*, *LNCS* 1807, pp. 122–138. Springer, 2000.
16. R. Cramer, V. Shoup. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. In *Crypto'98*, *LNCS* 1462, pp. 13–25, 1998.
17. —, Universal hash proofs and a paradigm for adaptive chosen ciphertext secure public-key encryption. In *Eurocrypt'02*, *LNCS* 2332, pp. 45–64, 2002.
18. S. Chatterjee, P. Sarkar. Generalization of the selective-ID security model for HIBE protocols. In *PKC'06*, *LNCS* 3958, pp. 241–256. Springer, 2006.
19. D. Dolev, C. Dwork, M. Naor. Non-malleable cryptography. In *STOC'91*, pp. 542–552, 1991.
20. I. Damgård, M. Jurik. A generalisation, a simplification and some applications of paillier's probabilistic public-key system. In *PKC'01*, *LNCS* 1992, pp. 119–136, Springer, 2001.
21. C. Dwork, M. Naor, O. Reingold, L. Stockmeyer. Magic functions. *J. of the ACM*, 50(6):852–921, 2003.
22. D.-M. Freeman, O. Goldreich, E. Kiltz, A. Rosen, G. Segev. More constructions of lossy and correlation-secure trapdoor functions. In *PKC'10*, *LNCS* 6056, pp. 279–295, Springer, 2010.

23. S. Fehr, D. Hofheinz, E. Kiltz, H. Wee. Encryption schemes secure against chosen-ciphertext selective opening attacks. In *Eurocrypt '10*, *LNCS* 6110, pp. 381–402. Springer, 2010.
24. J. Groth, R. Ostrovsky, A. Sahai. Perfect non-interactive zero knowledge for NP. In *Eurocrypt'06*, *LNCS* 4004, pp. 339–358. Springer, 2006.
25. J. Groth. Rerandomizable and replayable adaptive chosen ciphertext attack secure cryptosystems. In *TCC '04*, *LNCS* 2951, pp. 152–170. Springer, 2004.
26. S. Halevi, Y. Tauman-Kalai. Smooth projective hashing and two-message oblivious transfer. Cryptology ePrint Archive, Report 2007/118, 2007.
27. D. Hofheinz. Possibility and impossibility results for selective decommitments. Cryptology ePrint Archive, Report 2008/168, 2008.
28. —, All-but-many lossy trapdoor functions. Cryptology ePrint Archive: Report 2011/230, 2011.
29. M. Jakobsson, A. Juels, P. Syverson. Universal re-encryption for mixnets. In *2004 RSA Conference, Cryptographers track (CT-RSA'04)*, *LNCS* 2964, pp. 163–178. Springer, 2004.
30. Y. Tauman-Kalai. Smooth projective hashing and two-message oblivious transfer. In *Eurocrypt'05*, *LNCS* 3494, pp. 78–95 Springer, 2005.
31. E. Kiltz. Chosen-ciphertext security from tag-based encryption. In *TCC'06*, *LNCS* 3876, pp. 581–600. Springer, 2006.
32. G. Kol, M. Naor. Cryptography and game theory: Designing protocols for exchanging information. In *TCC'08*, *LNCS* 4948, pp. 320-339. Springer, 2008.
33. E. Kushilevitz, R. Ostrovsky. Replication is not needed: Single database, computationally-private information retrieval. In *FOCS'97*, pp. 364-373, 1997.
34. H. Krawczyk, T. Rabin. Chameleon signatures. In *Network and Distributed System Security Symposium (NDSS 2000)*, 2000.
35. E. Mann. Private access to distributed information. Master's thesis, Technion - Israel Institute of Technology, 1998.
36. P. MacKenzie, M. Reiter, K. Yang. Alternatives to non-malleability: Definitions, constructions, and applications. In *TCC'04*, *LNCS* 2951, pp. 171–190. Springer, 2004.
37. M. Naor, B. Pinkas. Efficient oblivious transfer protocols. In *SODA'01*, pages 448–457. ACM-SIAM, 2001.
38. M. Naor, M. Yung. Public-key cryptosystems provably secure against chosen ciphertext attacks. In *STOC'90*, pp. 427–437, 1990.
39. P. Paillier. Public-key cryptosystems based on composite degree residuosity classes. In *Eurocrypt'99*, *LNCS* 1592, pp. 223–238. Springer, 1999.
40. M. Prabhakaran, M. Rosulek. Rerandomizable RCCA encryption. In *Crypto'07*, *LNCS* 4622, pp. 517–534. Springer, 2007.
41. C. Peikert, V. Vaikuntanathan, B. Waters. A framework for efficient and composable oblivious transfer. In *Crypto'08*, *LNCS* 5157, pp. 554–571. Springer, 2008.
42. C. Peikert, B. Waters. Lossy trapdoor functions and their applications. In *STOC '08*, pp. 187–196, ACM Press, 2008.
43. C. Rackoff, D. Simon. Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack. In *Crypto'91*, *LNCS* 576, pp. 433–444, Springer, 1991.
44. A. Sahai. Non-malleable non-interactive zero-knowledge, and adaptive chosen-ciphertext security. In *FOCS '99*, pp. 543–553, 1999.
45. A. Young, M. Yung. Questionable encryption and its applications. In *Mycrypt'05*, *LNCS* 3715, pp. 210–221. Springer, 2005.
46. R. Zhang. Tweaking TBE/IBE to PKE transforms with chameleon hash functions. In *ACNS'07*, *LNCS* 4521, pp. 323–339, 2007.