# **Traceable Receipt-Free Encryption**

Henri Devillez, Olivier Pereira, and Thomas Peters

UCLouvain – ICTEAM – Crypto Group B-1348 Louvain-la-Neuve – Belgium

**Abstract.** CCA-like game-based security definitions capture confidentiality by asking an adversary to distinguish between honestly computed encryptions of chosen *plaintexts*. In the context of voting systems, such guarantees have been shown to be sufficient to prove ballot privacy (Asiacrypt'12). In this paper, we observe that they fall short when one seeks to obtain receipt-freeness, that is, when corrupted voters who submit chosen *ciphertexts* encrypting their vote must be prevented from proving how they voted to a third party.

Since no known encryption security notion can lead to a receipt-free ballot submission process, we address this challenge by proposing a novel publicly verifiable encryption primitive coined Traceable Receipt-free Encryption (TREnc) and a new notion of traceable CCA security filling the definitional gap underlined above.

We propose two TREnc instances, one generic achieving stronger guarantees for the purpose of relating it to existing building blocks, and a dedicated one based on SXDH. Both support the encryption of group elements in the standard model, while previously proposed encryption schemes aiming at offering receipt-freeness only support a polynomial-size message space, or security in the generic group model.

Eventually, we demonstrate how a TREnc can be used to build receipt-free protocols, by following a standard blueprint.

Keywords. New primitive, public-key encryption, receipt-freeness.

## 1 Introduction

A protocol offers receipt-freeness when players are unable to demonstrate to a third party which input they provided during a protocol execution. The need for receipt-freeness is most acute in order to prevent vote selling in the context of elections [7], which is our motivating application.

*Receipt-free voting.* In voting protocols, the random coins used by the voters can often be used as a receipt. For instance, in the famous protocol by Cramer et al. [20], of which a variant is used by the IACR in its own elections, a voter encrypts his vote with the election public key, and the resulting ciphertext is posted on a public bulletin board in order to support the verifiability of the election. If the voter decides to reveal to a third party the randomness used in the encryption process, that party can re-encrypt the claimed vote intent with the randomness provided by the voter and verify that the resulting ciphertext

 $\mathbf{2}$ 

appears on the bulletin board: the randomness used for encryption is, in effect, a receipt for the vote.

Since the seminal work of Benaloh [7], numerous protocols explored mechanisms that would guarantee that the random coins used by a voter are insufficient to explain his ballot as it is posted on the bulletin board for the needs of verification. In a first line of works [7,38,25], every possible voting choice is encrypted, the resulting ciphertexts are rerandomized and shuffled by the election authorities and made available to the voter. Furthermore, the permutations applied during the shuffle are also transmitted to the voter using secure channels. The voter then picks the ciphertext encoding his choice, and submits it for display on the bulletin board. Such a protocol guarantees that the voter ignores the randomness used to encrypt his ballot, and the protocol is designed in such a way that the voter is unable to prove which permutation he received, typically using designated-verifier zero-knowledge proofs. Such protocols are however quite demanding in terms of resources, as they require to encrypt a number of ciphertext proportional to the number of voting options, and a communication bandwidth to the voters that is proportional to the number of authorities. The more recent protocol of Kiayias et al. [28] faces similar challenges in terms of complexity, and also only considers a weaker form of receipt-freeness that focuses on voters preparing their ballot honestly.

More recently, Blazy et al. [10] proposed a simpler voting flow supporting receipt-freeness based on signatures on randomizable ciphertexts (SRC): the voters encrypt their vote and sign the resulting ciphertext, which is then transmitted to a re-encryption authority that re-randomizes the ciphertext, adapts the signature accordingly and posts the result on the bulletin board. The voter remains able to verify that a vote with a valid signature is posted on the board on his behalf, but is unable to explain the vote content thanks to the re-randomization step. Furthermore the SRC guarantees that the content of the encrypted ballot cannot be modified during the re-encryption process. This approach was further refined by Chaidos et al. [15], who also propose a simple game-based definition of receipt-freeness, which we adopt here, and more efficient SRCs keep being proposed [15,5].

This approach makes the ballot submission process asymptotically optimal for the voter, in the sense of Cramer et al. [20]: the protocol complexity for the voter becomes logarithmic in the number of voting options and independent of the number of election authorities, contrary to a dependency that is at least linear in both these factors when the approaches of [7,38,25,28] are used.

*Receipt-free ballot submission.* These works, by offering a simple ballot submission process in one pass, raise the natural question of identifying a public key encryption primitive that would support a receipt-free ballot submission process. Such a primitive would support a modular analysis of voting protocols that would be built around it, including various tallying approaches (based on mix-nets and homomorphic tallying for instance), and approaches to individual verifiability (based on the so-called Benaloh challenge [6] or on code voting for example [17]).

This question has been answered in the context of private (rather than receiptfree) ballot submission: it is well-known that a CCA-secure encryption scheme can be used to obtain a private ballot submission, a requirement that can be relaxed to NM-CPA security when the tally takes place in a single decryption round [22,41,9].

These works highlight the importance of some form of non-malleability in a submission process. From a practical point of view, non-malleability is needed in order to be able to detect (and prevent) non-independent ballot submissions (e.g., ballot copies) that would violate the privacy of the vote. From a technical point of view, security proofs require the availability of a decryption oracle used to extract the votes submitted by the adversary.

CCA security is however problematic in the context of receipt-free ballot submission, since we need to be able to re-randomize encrypted votes, so that the voter cannot explain the vote content anymore. The exploration of CCA-like security notions that would support some form of controlled malleability has been a fertile research area, which resulted in the definition of the notions of replayable-CCA (RCCA) security[14], homomorphic-CCA (HCCA) security [36], and controlled-malleable CCA (CM-CCA) security [16] for instance. As far as we know, all these works rely on the same CCA blueprint, in which an adversary submits one or more messages to a challenger, who answers either with an honest encryption of the messages or with something else, and the adversary must decide what he received with the help of a decryption oracle that accepts to decrypt any ciphertext that is not "recognizably" related to the challenge ciphertexts. The same holds in any other encryption primitives with CCA-like security with enhanced decryption capabilities. While they give more flexible ways to decrypt ciphertexts (based on identities, attributes and so on [39,26,37,11]), the challenge ciphertext is computed when the adversary sends a chosen *message*. Eventually, and following an observation which dates back at least to [25], deniable encryption [12] also only focusses on *honestly* computed ciphertexts that can then be explained for any other plaintexts.

This blueprint is however inadequate when turning to encryption schemes that would support the design of protocols that support receipt-freeness: in such a setting, we need to consider an adversary who sends to the challenger chosen *ciphertexts*, that may not be computed as a random encryption of a plaintext vote: they could have been *maliciously* computed.

## Our contributions

1) TCCA security. In this work, we investigate for the first time the implication of defining the notion of *traceable* CCA security (TCCA), a CCA-like security notion in which adversarially-chosen ciphertexts are submitted in the challenge phase. The challenge ciphertext is produced by *randomizing* one ciphertext or another, and we recognize derivatives of the challenge ciphertext thanks to a non-malleable public *trace* which is present in any ciphertext. To avoid trivial attacks, both ciphertexts given in the challenge phase must trace to each other, i.e., they must have the same trace.

This makes it possible for voters to submit a ciphertext of their choice, which will then be re-randomized by an authority, and can still be tracked by the voters using the trace.

For honestly produced ciphertexts, our security notion also implies traditional confidentiality properties, so that ballot privacy remains guaranteed should the re-randomizing authority be corrupted. So, non-malleability really serves two purposes here: (1) it guarantees that the re-randomizing authority cannot produce a ciphertext that would be related to an honestly produced one and have a different trace (which would violate ballot privacy), and (2) it guarantees that the re-randomizing authority cannot produce a ciphertext that would have the same trace as a given one but would decrypt to a different plaintext.

2) TREnc. We introduce Traceable Receipt-free Encryption (TREnc) as a new primitive with the following features:

- Traceability. Honestly generated ciphertexts are traceable in the sense that it is infeasible to modify the encrypted message;
- Randomizability. Valid ciphertexts are fully re-randomizable, up to the trace;
- TCCA security. Given a pair of ciphertexts that trace to each other, it is unfeasible to guess which one is randomized, even with access to a decryption oracle which decrypts any ciphertexts that do not trace to the challenge ciphertext, except before the challenge phase.

We also provide:

- 1. A generic TREnc that can be instantiated from existing building blocks that offer security in the standard model, and whose CRS is public-coin;
- 2. A pairing-based TREnc under the SXDH assumption in the standard model, where the public key only contains 13 first-source group elements and 6 second-source group elements, and the ciphertext contains 13 first-source group elements and 5 second-source group elements.

Both approaches improve on the state of the art: the previous SRC-based solutions either require costly bit-by-bit encryption [10,15], or only offer security in the generic group model [5].

3) A TREnc based voting scheme. Eventually, we show how to turn a TREnc into a simple voting scheme in a generic way, following the *Enc2Vote* blueprint previously used to turn a CCA-secure encryption scheme into a private voting scheme [9].

We demonstrate that the resulting voting scheme satisfies a notion of receiptfreeness that is equivalent in spirit to the one of Chaidos et al. [15], but fixes a small technical issue in that definition that makes their security game trivial to win (making it impossible to build a protocol that is receipt-free according to their definition).

**Other related works.** We focus on offering receipt-freeness in the context of voting, which is the context in which receipt-freeness was introduced [7], and which remains the main application context in which receipt-freeness is desired. Voting can however be seen as a special type of secure function evaluation protocol, in which specific tallying functions are evaluated and, as such, the notion of receipt-freeness, and the related notion of coercion-resistance have also been defined in the general multi-party computation setting [13,35,40,4]. We keep our focus on the voting context in order to clarify various design choices that are most meaningful in the voting setting compared to the general MPC setting: our primitive is targetted for a ballot submission process in which voters submit their ballot in one pass and do not communicate with each other, contrary to most MPC protocols, and we design mechanisms in which the ballot submission process can be fast, even on devices with limited computational power, while the verification of an election may require a longer period of time and use a dedicated computing infrastructure. Despite our focus on voting, it may be the case that TREnc mechanisms find applications in other contexts.

## 2 Traceable Receipt-Free Encryption

We propose a new public key encryption primitive and associated security notions that would support the receipt-free submission of votes in a protocol. As a first task, we identify the fundamental ingredients that are needed for our new encryption primitive.

An encryption scheme. We expect voters to submit their vote in an encrypted form, in order to guarantee the privacy of the votes.

*Receipt-free encryption.* Voters willing to sell their vote may choose to submit an arbitrary encrypted vote, which may be in the range of honestly produced ciphertexts but sampled according to a different distribution, or even just a sequence of bits that would not be within the range of the encryption mechanism. By deviating from the normal encryption process, the voter hopes to obtain a receipt that could be used to demonstrate his vote intent to a third party.

If the encrypted vote that is tallied is produced by the voter only, then the voter will always have a receipt: the random coins used to encrypt the ballot. In order to avoid this, we rely on the existence of a semi-trusted authority: that authority will be trusted to prevent a dishonest voter from obtaining a receipt for his vote, but will not be trusted for the correctness of the election result, and will not be trusted for the privacy of votes encrypted by honest voters.

Concretely, in order to achieve receipt-freeness, this semi-trusted authority tests the validity of a voter submitted ciphertext (without the need of any secret key) and re-randomizes every valid ciphertext before posting it on a public bulletin board.

*Traceable Receipt-Free Encryption.* In order to make it possible for a voter to check that his ballot has not been unduly modified by this semi-trusted re-randomizing authority, it must be possible to extract a *trace* from any valid ciphertext. A honest re-randomization process would keep the trace is unchanged, hence making

ciphertexts *traceable*, while no corrupted authority should be able to modify a ciphertext in such a way that it would decrypt to a different vote while keeping the trace unchanged.

Furthermore, we need to make sure that this trace cannot serve as a receipt for the vote. In order to make sure that it is the case, we split the encryption process in two steps, that guarantee that any trace can be associated to any possible vote intent. Concretely, an encryption starts with the generation of a secret *link key*, which is then used, together with the encryption public key, to encrypt any possible vote. This guarantees that, even if a voter leaks the link key associated to his ballot as a receipt, the ballot could still encrypt any vote.<sup>1</sup>

## 2.1 Syntax

We now have the ingredients that we need to define a Traceable Receipt-Free Encryption scheme, or TREnc.

**Definition 1 (Traceable Receipt-Free Encryption).** A Traceable Receipt-Free Encryption scheme (*TREnc*) is a public key encryption scheme (Gen, Enc, Dec) that is augmented with a 5-tuple of algorithms (LGen, LEnc, Trace, Rand, Ver):

- LGen(pk; r): The link generation algorithm takes as input a public encryption key pk in the range of Gen and randomness r, and outputs a link key lk.
- LEnc(pk, lk, m; r): The linked encryption algorithm takes as input a pair of public/link keys (pk, lk), a message m and randomness r and outputs a ciphertext.
- Trace(pk, c) : The tracing algorithm takes as input a public key pk, a ciphertext c and outputs a trace t. We call t the trace of c.
- Rand(pk, c; r): The randomization algorithm takes as input a public key pk, a ciphertext c and randomness r and outputs another ciphertext.
- Ver(pk, c): The verification algorithm takes as input a public key pk, a ciphertext c and outputs 1 if the ciphertext is valid, 0 otherwise.

In many cases, we will omit the randomness r from our notations. It is then assumed that it is selected uniformly at random.

We require several correctness properties from the additional algorithms of a TREnc. The first requires that encrypting a message m by picking a link key lk using LGen and computing LEnc(pk, lk, m) produces a ciphertext that is identically distributed to a fresh encryption of m using Enc. The second requires that the

<sup>&</sup>lt;sup>1</sup> Of course, this also means that, if a corrupted re-randomizing authority obtains a voter's secret link key (e.g., by corrupting the voter's voting client), then it might be able to produce a ciphertext that encrypts a different vote intent but would still trace to the original voter trace. Just as other attacks related to corrupted voting clients, such attacks can be prevented by traditional continuous ballot testing procedures [6], in which a voter would have the option to ask an authority to spoil a ballot posted on the bulletin board, which would then be verifiability decrypted for verification, and later replaced by a fresh new ballot produced by the voter, using a fresh link key.

**Trace** of a ciphertext does not depend on the message that is encrypted. The third requires that randomizing a ciphertext does not change the corresponding plaintext neither the corresponding trace. The last requires that every honestly computed ciphertext passes the verification algorithm.

**Definition 2 (TREnc correctness).** We require that a TREnc scheme satisfies the following correctness requirements.

- **Encryption compatibility** For every pk in the range of Gen and message m, the distributions of Enc(pk,m) and LEnc(pk,LGen(pk),m) are identical.
- Link traceability For every pk in the range of Gen, every lk in the range of LGen(pk), the encryptions of every pair of messages  $(m_0, m_1)$  trace to the same trace, that is, it always holds that Trace(pk, LEnc(pk, lk,  $m_0$ )) = Trace(pk, LEnc(pk, lk,  $m_1$ )).
- **Publicly Traceable Randomization** For every pk in the range of Gen, every message m and every c in the range of Enc(pk, m), we have that Dec(sk, c) = Dec(sk, Rand(pk, c)) and Trace(pk, c) = Trace(pk, Rand(pk, c)).
- **Honest verifiability** For every pk in the range of Gen and every messages m, it holds that Ver(pk, Enc(pk, m)) = 1

## 2.2 Security definitions

Verifiability We require several security properties from a TREnc. Our first property is fairly standard: a TREnc is verifiable if the Ver algorithm guarantees that a ciphertext is within the range of Enc. In other words, the ciphertext can be explained by some message m, some link key lk, and some coins, even if they are not easily computable.

**Definition 3 (Verifiability).** A TREnc is verifiable if for every PPT adversary, the following probability is negligible in  $\lambda$ :

 $\Pr[\mathsf{Ver}(\mathsf{pk}, c) = 1 \text{ and } c \notin \mathsf{Enc}(\mathsf{pk}, \cdot) | (\mathsf{pk}, \mathsf{sk}) \leftarrow \mathsf{Gen}(1^{\lambda}); c \leftarrow \mathcal{A}(\mathsf{pk}, \mathsf{sk})].$ 

TCCA security. We now turn to our central security definition, security against traceable chosen ciphertexts attacks, or TCCA security, which differs from all existing CCA-like notions by letting the adversary submit pairs of ciphertexts instead of pairs of messages, reflecting that we need security in front of adversarially chosen ciphertexts. In the TCCA security game (Fig. 1), the adversary receives the public key and has access to a decryption oracle, as usual. It then submits a pair of ciphertexts that must be valid and have identical traces. One of the ciphertexts is randomized and returned to the adversary, who must decide which one it is. After receiving this challenge ciphertext, the adversary can still query the decryption oracle, but only on ciphertexts that have a trace different of his challenge ciphertext. So, the challenger must faithfully decrypt pre-challenge ciphertexts that have the same trace as the challenge ciphertext. Looking ahead, this decryption capability offers an easy but necessary means allowing simulating the result of an election when proving receipt-freeness.

TCCA security guarantees that, if a voter submits a ciphertext that is randomized before it is posted on a public bulletin board, then the resulting ciphertext becomes indistinguishable from any other ciphertext that would have the same trace, and we know from the link traceability that the encryption of any vote could have that trace. This essentially guarantees the absence of a vote receipt.

**Definition 4** (TCCA). A TREnc is TCCA secure if for every PPT adversary  $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$  the experiment  $\mathsf{Exp}_{\mathcal{A}}^{\mathsf{tcca}}(\lambda)$  defined in Figure 1 (left) returns 1 with a probability negligibly close in  $\lambda$  to  $\frac{1}{2}$ .

$Exp^{ ext{tcca}}_{\mathcal{A}}(\lambda)$	$Exp^{\mathrm{trace}}_{\mathcal{A}}(\lambda)$
$(pk,sk) \leftarrow \$  Gen(1^{\lambda})$	$\overline{(pk,sk) \gets \$ \operatorname{Gen}(1^{\lambda})}$
$(c_0, c_1, st) \leftarrow \$ \mathcal{A}_1^{Dec(\cdot)}(pk)$	$(m, st) \leftarrow \mathcal{A}_1(pk, sk)$
$b \leftarrow \$ \{0, 1\}$	$c \gets \$ \operatorname{Enc}(pk, m)$
<b>if</b> $Trace(pk, c_0) \neq Trace(pk, c_1)$ or	$c^{\star} \leftarrow \$ \mathcal{A}_2(c, st)$
$Ver(pk, c_0) = 0$ or $Ver(pk, c_1) = 0$ then return b	if Trace(pk, $c$ ) = Trace(pk, $c^*$ ) and
$c^{\star} \leftarrow \$ Rand(pk, c_b)$	$\operatorname{Ver}(pk, c^*) = 1 \text{ and } \operatorname{Dec}(sk, c^*) \neq m$
$b' \leftarrow \$ \mathcal{A}_2^{\mathcal{O}Dec^{\star}(\cdot)}(c^{\star},st)$	then return 1
$\mathbf{return} \ \bar{b}' = b$	else return 0

**Fig. 1.** TCCA and trace experiments. In the TCCA experiment,  $A_2$  has access to a decryption oracle  $\mathcal{O}\mathsf{Dec}^*(\cdot)$  which, on input c, returns  $\mathsf{Dec}(c)$  if  $\mathsf{Trace}(\mathsf{pk}, c) \neq \mathsf{Trace}(\mathsf{pk}, c^*)$  and test otherwise.

It is naturally possible to write a multi-challenge version of the  $\text{Exp}_{\mathcal{A}}^{\text{tcca}}(\lambda)$  experiment, which we call q-TCCA, in which the adversary can submit q pairs of ciphertexts. This leads to an equivalent definition, as demonstrated in the full version [21]. We also stress that in the challenge query the adversary may know the random coins underlying  $c_0$  and  $c_1$  and may have drawn them from a specific secret distribution. The randomization leading to the challenge ciphertext  $c^*$  should thus erase any subliminal information binding  $c^*$  to the message in  $c_b$ . This definition introduces some technical difficulty when it comes to proving the TCCA security as it becomes harder to program the public key to ease the transition toward a game where we are able to inject an independent message in the plaintext in an undetectable way. Indeed, we have no clue at the setup time about the distribution of  $(c_0, c_1)$  and their common trace while the emulation of Rand(pk,  $c_b$ ) must preserve it without even knowing the underlying link keys.

TCCA security is reminiscent of the notion of publicly detectable replayable-CCA (pd-RCCA) security proposed by Canetti et al. [14]. The pd-RCCA security game is essentially the same as the CCA game, except for two main differences: a publicly computable equivalence relation is defined on ciphertexts and, after the challenge ciphertext has been received, the challenger will refuse to decrypt any ciphertext that is equivalent to this challenge ciphertext. Furthermore, ciphertexts that are in the same equivalence class must decrypt to the same message (for

completeness, the full definition is available in the full version [21]). The pd-RCCA security game looks appealing in the context of voting, because it captures this idea of having the possibility to re-randomize ciphertexts while also keeping a trace that could be detected through the equivalence relation. And, indeed, RCCA-secure encryption has been used in previous proposals of receipt-free voting schemes [15].

There are three central differences, though, which motivate the introduction of the TCCA security game.

- The challenge ciphertexts of the pd-RCCA security game are always honestly computed and, as such, pd-RCCA security does not offer any guarantee in the face of maliciously produced ciphertexts, as it would be the case when a voter tries to obtain a receipt for his vote.
- Contrary to pd-RCCA security, it can be observed that TCCA security says nothing about the hiding property of the Enc algorithm, since the adversary must distinguish based on outputs of Rand. An extreme case could define Enc as the identity function, Trace as mapping to a single constant trace, and Rand actually performing the encryption work, and this could still offer a TCCA secure scheme. The confidentiality requirements on Enc will be handled through the traceability and strong randomization properties below.
- There is no requirement for TCCA security that trace equivalent ciphertexts decrypt to the same message: a single link key can be used to encrypt any message, and all the resulting ciphertexts would have the same trace (by the link traceability correctness property). We recall that this non-binding feature is essential for receipt-free voting.

As such, TCCA security is not comparable to pd-RCCA security. It is shown in the full version of the paper [21] that, under (different) additional conditions, implications can be proven in both directions for a natural variant of pd-RCCA security adapted to TREnc schemes.

Traceability and Strong Randomization. While TCCA security relates to a model in which the voting client may be corrupted but the re-randomization server is honest, we now focus on two central properties that are important when the voting client is honest and the re-randomization server might be corrupted.

The traceability property guarantees to the sender of a honestly encrypted message that no efficient adversary would be able to produce another ciphertext that traces to the same trace and would decrypt to a different message, even if the adversary knows the secret decryption key. So, even if a TREnc offers some form of ciphertext malleability, its traceability implies the non-malleability of the plaintexts. This is an important feature for the verifiability of a voting system: as long as the link key used to encrypt a vote remains secret, and the voter submits a single ciphertext encrypted with that link key, the voter is guaranteed that any ciphertext that would trace to his original ciphertext encrypts his original vote. (But, of course, using the link key, it remains possible to produce ciphertexts with the same trace that would decrypt to any vote.)

**Definition 5 (Traceability).** A TREnc is traceable if for every PPT adversary  $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ , the experiment  $\mathsf{Exp}_{\mathcal{A}}^{\mathrm{trace}}(\lambda)$  defined in Figure 1 (right) returns 1 with a probability negligible in  $\lambda$ .

The second property, strong randomization, requires that the output of the Rand algorithm applied to any valid ciphertext is distributed just as a random encryption of the same message with the same link key.

**Definition 6 (Strong Randomization).** A TREnc is strongly randomizable if for every  $c \in \text{LEnc}(pk, lk, m)$  with pk in the range of Gen and lk in the range of LGen(pk), the following computational indistinguishability relation holds:

 $\mathsf{Rand}(\mathsf{pk}, c) \approx_c \mathsf{LEnc}(\mathsf{pk}, \mathsf{lk}, m)$ 

Requiring strong randomization together with TCCA security guarantees that Enc actually hides messages. CPA security comes easily: when the CPA adversary sends  $(m_0, m_1)$  to the TCCA adversary, the TCCA adversary can encrypt the 2 messages using a single random link key and send them to the TCCA challenger, which will return a randomization of one of them. Strong randomization guarantees that this is distributed exactly like an encryption of one of the two messages, and we can send the result to the CPA adversary, who will then offer the answer expected for the TCCA game. We show a stronger implication to RCCA security in the full version of the paper [21].

## 3 Towards a generic TREnc

We are now interested in exploring how a TREnc could be designed from existing tools. The core TREnc security feature comes from the TCCA security game, in which the adversary submits a pair of ciphertexts with identical traces and receives a re-randomization of one of them. If we want relate this game to a more standard RCCA-style security definition in which the adversary submits a pair of plaintext and receives an encryption of one of them, we need to be able to translate a re-randomization query on two ciphertexts into an encryption query on the two corresponding plaintexts. But there is an additional constraint that needs to be satisfied: the ciphertext resulting from the encryption query needs to have the same trace as the original ciphertexts. In other words, we need to be able to extract the link key that they contain. We capture this last idea in an augmented version of a TREnc, which we call *extractable TREnc*.

#### 3.1 Extractable TREncs

Essentially, an extractable TREnc makes it possible to produce encryption keys together with a trapdoor using a TrapGen algorithm. Using that trapdoor, it becomes possible to extract, from any ciphertext, a link key that makes it possible to produce new ciphertexts with the same trace as the original one. This in turn implies the possibility to break the traceability of the scheme.

**Definition 7 (Extractable TREnc).** An extractable TREnc is a TREnc with two additional algorithms TrapGen and LExtr:

- $\operatorname{TrapGen}(1^{\lambda})$ : The trapdoor generation algorithm takes as input the security parameter and outputs a tuple of public/secret/trapdoor keys (pk, sk, tk). We require the distribution of the (pk, sk) pairs produced by  $\operatorname{TrapGen}(1^{\lambda})$  to be identical to the one of the outputs of  $\operatorname{Gen}(1^{\lambda})$ .
- LExtr(tk,c): The link extraction algorithm takes as input the trapdoor key and a ciphertext and returns a link key lk such that, if c is in the range of Enc(pk,·) with pk in the range of Gen, then c is in the range of LEnc(pk,lk,·).

It is fairly natural to require that ciphertexts can only be consistent with one single link key, hence guaranteeing a unique link key extraction.

**Definition 8 (Unique Extraction).** An extractable TREnc has unique extraction *if, for every* (pk, sk, tk) *in the range of* TrapGen *and* lk *in the range of* LGen(pk), we have that:

- $\mathsf{LExtr}(c,\mathsf{tk}) = \mathsf{lk} \ whenever \ c \in \mathsf{LEnc}(\mathsf{pk},\mathsf{lk},\cdot);$
- $\mathsf{LExtr}(c_0,\mathsf{tk}) = \mathsf{LExtr}(c_1,\mathsf{tk})$  whenever we have  $\mathsf{Trace}(\mathsf{pk},c_0) = \mathsf{Trace}(\mathsf{pk},c_1)$ and  $c_0, c_1 \in \mathsf{Enc}(\mathsf{pk},\cdot)$ .

## 3.2 A TREnc flavored variant of pd-RCCA security

Based on an extractable TREnc, we now propose an RCCA-like security definition, pd\*-RCCA-security, which shares much of the spirit of the pd-RCCA notion of Canetti et al. [14], but is rather tailored as a useful intermediary notion for achieving TCCA security: we will show that any pd\*-RCCA-secure extractable TREnc is also TCCA secure. Eventually, we will show how to achieve pd\*-RCCA-security from existing tools.

**Definition 9 (pd\*-RCCA).** An extractable TREnc is pd\*-RCCA-secure if for any PPT adversary  $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ , the experiment  $\operatorname{Exp}_{\mathcal{A}}^{\operatorname{pd}^*\operatorname{-rcca}}(\lambda)$  in Figure 2 returns 1 with a probability negligibly close on  $\lambda$  to  $\frac{1}{2}$ .

```
\begin{split} & \frac{\mathsf{Exp}_{\mathcal{A}}^{\mathrm{pd}^{\star}\operatorname{-rcca}}(\lambda)}{(\mathsf{pk},\mathsf{sk},\mathsf{tk}) \leftarrow \$\operatorname{TrapGen}(1^{\lambda})} \\ & (m_0,m_1,\mathsf{lk},\mathsf{st}) \leftarrow \$ \mathcal{A}_1^{\mathrm{Dec}(\mathsf{sk},\cdot),\mathsf{LExtr}(\mathsf{tk},\cdot)}(\mathsf{pk}) \\ & b \leftarrow \$ \{0,1\} \\ & \text{if } \mathsf{lk} = \bot \operatorname{\mathbf{then}} c^{\star} \leftarrow \$ \operatorname{Enc}(\mathsf{pk},m_b) \\ & \text{else } c^{\star} \leftarrow \$ \operatorname{LEnc}(\mathsf{pk},\mathsf{lk},m_b) \\ & b' \leftarrow \$ \mathcal{A}_2^{\mathcal{ODec}^{\star}(\mathsf{sk},\cdot),\mathrm{LExtr}(\mathsf{tk},\cdot)}(c^{\star},\mathsf{st}) \\ & \text{return } b = b' \end{split}
```

**Fig. 2.** pd\*-RCCA experiment. Here,  $\mathcal{O}\mathsf{Dec}^*(\mathsf{sk}, c)$  is a decryption oracle that returns test if  $\mathsf{Trace}(\mathsf{pk}, c) = \mathsf{Trace}(\mathsf{pk}, c^*)$  and  $\mathsf{Dec}(\mathsf{sk}, c)$  otherwise.

Just as in the pd-RCCA security definition, our adversary receives a public key, then can make decryption queries, make a challenge query on a pair of plaintexts, receive an encryption  $c^*$  of one of them, and then make more decryption queries, provided that they are not about ciphertexts that are equivalent to  $c^*$ . Here the notion of equivalent ciphertext is defined by ciphertexts with identical traces, which does not imply that they decrypt to the same plaintext, contrary to the compatibility requirement of pd-RCCA security. The extra features of pd\*-RCCA security, which come naturally in the context of an extractable TREnc, are that:

- On top of having access to a decryption oracle, the adversary has access to a LExtr oracle giving him the ability to extract the link key from any ciphertext.
- During his challenge query, the adversary can provide a link key on top of its two plaintexts: the challenge ciphertext will then be computed using that link key.

As announced, a pd\*-RCCA-secure and strongly randomizable extractable TREnc is also TCCA-secure.

**Theorem 1.** If a TREnc scheme T is extractable, strongly randomizable, and  $pd^*$ -RCCA-secure, then it is TCCA secure. More precisely, if the advantages of any PPT adversary at strong randomization and  $pd^*$ -RCCA experiment are respectively bounded by  $\varepsilon_{SR}$  and  $\varepsilon$ , then for any PPT adversary  $\mathcal{A}$ , we have  $Pr[\mathsf{Exp}_{\mathcal{A},T}^{\mathsf{tcca}}(\lambda) = 1] \leq \frac{1}{2} + \varepsilon_{SR} + \varepsilon$ .

*Proof.* (See the full version for details.) The decryption queries from the TCCA adversary are forwarded to the pd<sup>\*</sup>-RCCA challenger. When the TCCA adversary makes his challenge query on  $(c_0, c_1)$ , the reduction obtains the corresponding link key and plaintexts by querying the pd<sup>\*</sup>-RCCA challenger, and sends them as pd<sup>\*</sup>-RCCA challenge. The resulting ciphertext is correctly distributed thanks to strong randomizability, and has the correct trace thanks to the extractability. The winning probability of the TCCA adversary is then negligibly close to the winning probability of the resulting pd<sup>\*</sup>-RCCA adversary. □

### 3.3 Building a pd\*-RCCA-secure extractable TREnc

We are now ready to build a TREnc. As a first natural building block, we use a signature on randomizable ciphertexts (SRC), as introduced by Blazy et al. [10]. In an SRC, any signed ciphertext can be publicly re-randomized, and the signature can be publicly adapted so that it remains valid for the new ciphertext.

We can easily obtain the structure of a TREnc from an SRC by defining the LGen function as setting lk as a fresh signing key for the SRC, and the LEnc function as encrypting the plaintext using a randomizable encryption scheme, then signing that ciphertext using lk. The trace of a ciphertext would then be the signature verification key.

This offers a promising skeleton, but it is not sufficient to obtain pd\*-RCCA security: as it is, the adversary could simply remove the signature from the challenge ciphertext, sign that ciphertext with a fresh key in order to obtain a different trace, and ask for the decryption of the result, which would be granted.

A natural solution to this problem is to link the trace to the ciphertext using tag-based encryption [29] mechanism. In a tag-based encryption scheme, the encryption and decryption functions take an arbitrary tag as an extra input, and the decryption of a ciphertext with an incorrect tag will fail. We rely on the standard notion of weak-CCA security for tag based encryption [34], which is the CCA security game excepted that the challenge ciphertext is produced using an *adversarially chosen* tag, and that no decryption query can be made using that tag (only) *after* the challenge phase. This security game nicely fits our pd\*-RCCA security game, in which the trace derived from the link key submitted by the adversary can be used as a tag, and guarantees that no ciphertext can be modified in such a way that it successfully decrypts with a tag that is different of the original one. We note that we must be able to decrypt pre-challenge queries that already contains the "challenge tag" of the adversary, which prevents us from only relying on (weak) selective-tag security.

But we still need to be able to extract the link key from a tag-based ciphertext. This can be done fairly easily, by augmenting our encryption process with the requirement to encrypt the link key using a randomizable CPA-secure encryption scheme, and to add a randomizable ZK proof that the encrypted link key is indeed the one that is used as tag for the tag-based encryption. Extraction would then simply proceed by decrypting that CPA ciphertext. (In particular, it does not rely on any extraction property of the ZK proof system: we just need its soundness.)

To summarize, we build an extractable TREnc from the following ingredients:

- A randomizable weakly CCA secure tag-based encryption scheme (TBGen, TBEnc, TBDec).
- An SRC compatible with the tag-based encryption scheme, which includes a signature scheme (SGen, Sign, SVer).
- A randomizable CPA secure public key encryption scheme (EGen, EEnc, EDec).
- A randomizable NIZK proof system (Prove, VerifyProof) that, on input  $(c_{tbe}, c_{extr})$  and associated public keys, demonstrates that  $c_{extr}$  is an encryption with EEnc of the signing key whose corresponding verification key has been used as a tag in order to compute  $c_{tbe}$  using TBEnc.

And the blueprint of our TREnc is as follows:

- TrapGen uses TBGen to produces a key pair (tpk, tsk), and EGen to produce a key pair ( $e_{extr}$ ,  $d_{extr}$ ). It returns  $pk = (tpk, e_{extr})$ , sk = tsk and  $tk = d_{extr}$ .
- LGen sets lk as a signing key obtained from Sign. We assume that the corresponding signature verification key vk can be derived from lk.
- LEnc encrypts the message m as follows:
  - $c_{tbe} = \mathsf{TBEnc}(\mathsf{tpk},\mathsf{lk},m);$
  - $\sigma = \text{Sign}(\text{lk}, c_{tbe});$
  - $c_{extr} = \mathsf{EEnc}(\mathsf{e}_{\mathsf{extr}},\mathsf{lk});$
  - $\pi = \mathsf{Prove}(c_{tbe}, c_{extr}; \mathsf{lk})$
- The ciphertext is made of these 4 elements, together with vk.
- Dec returns  $\mathsf{TBDec}(\mathsf{tsk}, \mathsf{vk}, c_{tbe})$

- 14 Henri Devillez, Olivier Pereira, and Thomas Peters
- Trace returns vk.
- Rand re-randomizes  $c_{tbe}$ , adapts  $\sigma$  accordingly, re-randomizes  $c_{extr}$ , and re-randomizes and adapts  $\pi$ .
- Ver accepts a TREnc ciphertext if the two ciphertexts that it contains are valid, if the signature is valid, and if the ZK proof verifies.
- LExtr returns  $EDec(tsk, c_{extr})$ .

A complete description of this generic TREnc, together with proofs of its security, is available in the full version of the paper [21], where we rely on the standard security notions of all the above ingredients. Exploring whether some of these notions can be relaxed is an interesting scope for further research. Finally, we mention that pairing-based realizations exist for all these ingredients and that it would be appealing to understand how to construct a secure post-quantum TREnc. The main obstacle we see relates to the controlled-malleability feature of our new primitive (i.e., any ciphertext must support an unbounded number of randomization) which makes it less straightforward to realize in general, and for instance based on lattices.

Remark 1. This section showed that the notion of extractable TREnc offers a convenient companion for a TREnc: it is possible to build an extractable TREnc from relatively common, yet strong, building blocks, and the proof of TCCA security of this TREnc comes relatively easily because we can design a pd-RCCA-like security notion for extractable TREnc that implies our new TCCA security notion. The resulting construction is however expected to be fairly expensive since, in the standard model, all known instantiations of the building blocks relies on a bit-by-bit decomposition of the message or the secret singing key of which the ciphertext must contain a (malleable) ZK proof of. Nevertheless, providing this extractability feature is an artifice for the construction that is not necessary for the security of the TREnc, but as far as we know there is no obvious generic construction leading to a TREnc without extractability. In the next section we turn to the construction of an ad-hoc efficient instance of a TREnc based on a standard computational assumption that also avoid the costly bit-by-bit decomposition.

## 4 Pairing-Based Construction under SXDH

This section provides a secure TREnc in the standard model, only relying on the SXDH assumption and on a CRS. Contrary to our previous construction, this one is not extractable – extractability was just a convenience but does not offer any security benefit. This allows us to get a more efficient solution, here, in asymmetric bilinear groups. Moreover, our construction enjoys a short public-key and short ciphertexts as they only contain a constant number of group elements to encrypt a full group element, contrary to previous proposals that required to process the message bit by bit [10,15].

We first introduce the cryptographic assumptions on which we will rely, as well as the main existing building block that we will use: linearly homomorphic structure-preserving signatures.

#### 4.1 Computational setting

We rely on an efficient Setup algorithm that generates common public parameters pp. Given a security parameter  $\lambda$ , Setup $(1^{\lambda})$  generates a bilinear group pp =  $(\mathbb{G}, \hat{\mathbb{G}}, \mathbb{G}_T, p, e, g, \hat{g}, \hat{h})$  of prime order  $p > 2^{\text{poly}(\lambda)}$  for some polynomial poly, where  $g \leftarrow \mathbb{G}$  and  $\hat{g}, \hat{h} \leftarrow \mathbb{G}$  are random generators and  $e : \mathbb{G} \times \hat{\mathbb{G}} \to \mathbb{G}_T$  is a bilinear map. In this setting, we rely on the SXDH assumption, which states that the DDH problem must be hard in both  $\mathbb{G}$  and  $\hat{\mathbb{G}}$ . Following the Groth-Sahai standard notation, we also define the linear map  $\iota : \mathbb{G} \to \mathbb{G}^2$  with  $\iota : Z \mapsto (1, Z)$ .

## 4.2 Linearly Homomorphic Structure-Preserving Signatures

A central tool for our efficient TREnc construction is linearly homomorphic structure-preserving signatures. The structure preserving [2][1] property makes it possible to sign messages that are group elements (and not just bits as in schemes based on the Waters signature), while the additional linearly homomorphic feature, introduced by Libert et al. [32], will be used to make the signatures randomizable while guaranteeing the non-malleability of the plaintext.

- **Keygen(pp**, *n*): given the public parameter **pp** and the (polynomial) space dimension  $n \in \mathbb{N}$ , choose  $\chi_i, \gamma_i \leftrightarrow \mathbb{Z}_p$  and compute  $\hat{g}_i = \hat{g}^{\chi_i} \hat{h}^{\gamma_i}$ , for i = 1 to n. The private key is  $\mathsf{sk} = \{(\chi_i, \gamma_i)\}_{i=1}^n$  and the public key is  $\mathsf{pk} = \{\hat{g}_i\}_{i=1}^n \in \mathbb{G}^n$ .
- **Sign(sk, (M<sub>1</sub>,..., M<sub>n</sub>)):** to sign a vector  $(M_1,...,M_n) \in \mathbb{G}^n$  using sk =  $\{(\chi_i, \gamma_i)\}_{i=1}^n$ , output  $\sigma = (Z, R) = (\prod_{i=1}^n M_i^{\chi_i}, \prod_{i=1}^n, M_i^{\gamma_i}).$
- **SignDerive(pk, {** $(\omega_i, \sigma^{(i)})$ **}** $_{i=1}^{\ell}$ **):** given pk as well as  $\ell$  tuples  $(\omega_i, \sigma^{(i)})$ , parse  $\sigma^{(i)}$  as  $\sigma^{(i)} = (Z_i, R_i)$  for i = 1 to  $\ell$ . Return the triple  $\sigma = (Z, R) \in \mathbb{G}$ , where  $Z = \prod_{i=1}^{\ell} Z_i^{\omega_i}, R = \prod_{i=1}^{\ell} R_i^{\omega_i}$ .
- Verify(pk,  $\sigma$ ,  $(M_1, \ldots, M_n)$ ): given  $\sigma = (Z, R) \in \mathbb{G}^2$  and  $(M_1, \ldots, M_n)$ , return 1 if and only if  $(M_1, \ldots, M_n) \neq (1_{\mathbb{G}}, \ldots, 1_{\mathbb{G}})$  and (Z, R) satisfies

$$e(Z, \hat{g}) \cdot e(R, h) = \prod_{i=1}^{n} e(M_i, \hat{g}_i)$$
 (1)

### 4.3 Intuition of our construction

To encrypt a message  $m \in \mathbb{G}$ , we combine a CPA encryption  $\boldsymbol{c} = (c_0, c_1, c_2)$  of the form  $c_0 = m \cdot f^{\theta}$ ,  $c_1 = g^{\theta}$ ,  $c_2 = h^{\theta}$  and and a randomizable publicly verifiable proof that  $\log_g c_1 = \log_h c_2$ , à la Cramer-Shoup. For that purpose, we can rely on the idea to include a one-time LHSP signature on top of  $\boldsymbol{c}$  as first suggested in [32]. That means that the public key contains an LHSP signature  $\Sigma$  on (g, h) so that we can derive a signature on  $(g, h)^{\theta}$  if indeed  $(c_1, c_2)$  lies in span $\langle (g, h) \rangle$  by computing  $\pi = \Sigma^{\theta}$ . Such a proof is quasi-adaptive [27] as the CRS depends on the language of which we have to prove membership. Here, the public key includes a CRS that contains a signature on the basis of the linear subspace span $\langle (g, h) \rangle$  of  $\mathbb{G}^2$ . Given  $\boldsymbol{c}$  and the LHSP signature  $\pi$  one can easily randomize the ciphertext as follows: compute  $\boldsymbol{c}' = \boldsymbol{c} \cdot (f, g, h)^{\theta'}$ , and adapt the proof  $\pi' = \pi \cdot \Sigma^{\theta'}$ . While

this solution is perfectly randomizable and the signing key allows to perfectly simulate the proof, it only provides a CCA1 security. Still, this technique has been enhanced to provide tag-based simulation-sound proof system which is reminiscent to building CCA-like secure encryption. The underlying technique is to generate a one-time key pair (opk, osk) of some one-time signature scheme that will be discussed in the next paragraph, and to define the tag as  $\tau = H(\mathsf{opk})$ , for some collision-resistant hash function  $H^2$ , before computing  $\pi$  that  $(c_1, c_2)$  lies in span $\langle (q,h) \rangle$  based on  $\tau$ . The ciphertext is then completed by signing  $(c,\pi)$ with osk, resulting in the ciphertext ( $c, \pi, \sigma, opk$ ). A natural solution would be to borrow the first solution due to [33] but it only provides selective-tag simulation soundness. Since we will be using **opk** as the trace of our TREnc construction, the TCCA security implies that our underlying tag-based encryption must achieve tag-based weak CCA security, and selective-tag security is not enough. Indeed, the tag  $\tau^* = H(\mathsf{opk}^*)$  involved in the challenge ciphertext may be chosen by the adversary at any time. Furthermore, we must be able to answer any pre-challenge decryption queries, so even those that already used  $\tau^*$ . That means that we cannot program the public key to embed  $\tau^*$  that will help us to incorporate an SXDH instance in the computation of the challenge ciphertext. Fortunately, by including a signature  $\Sigma_u$  on (g, h, 1, 1) and another signature  $\Sigma_v$  on (1, 1, g, h) in the public key, given a tag  $\tau$ , the computation of  $\pi = (\Sigma_u^{\tau} \Sigma_v)^{\theta}$  due to [30] is an LHSP signature on  $(c_1^{\tau}, c_2^{\tau}, c_1, c_2)$  which gives us the expected security and still enjoys a perfect randomizability, but for the given tag  $\tau$  (and trace opk), only, which is still what we were looking for.

Now, we come back on the signature  $\sigma$  of  $(c, \pi)$ . Usually, (opk, osk) is a key pair of a strongly unforgeable signature scheme providing non-malleability of ciphertext. However, we want to keep the malleability of the ciphertext as we want to be able to fully randomize it up to opk that will serve as our trace, but we also want to retain the non-malleability of the encrypted message mto satisfy traceability. Here again, in the standard model under SXDH, LHSP signature scheme comes in handy. If our fresh key pair (opk, osk) is generated from a one-time LHSP signature scheme, we can fix the message and preserve the randomizability of  $(c, \pi)$  by computing one-time LHSP signatures  $\sigma_1$  on  $(g, c_0, c_1, c_2)$  and  $\sigma_2$  on (1, f, g, h). Like this, when we randomize  $(\boldsymbol{c}, \pi, \sigma_1, \sigma_2)$  as  $\boldsymbol{c}' = \boldsymbol{c} \cdot (f, g, h)^{\theta'}, \pi' = \pi \cdot (\Sigma_u^{\tau} \Sigma_v)^{\theta'}$  we can also adapt the signature  $\sigma'_1 = \sigma_1 \cdot \sigma_2^{\theta'}$  on  $(g, c'_0, c'_1, c'_2) = (g, c_0, c_1, c_2) \cdot (1, f, g, h)^{\theta'}$ , and simply keep  $\sigma_2$ . While the correctness follows by inspection, we have several comments to make that are less obvious. First, the reason why we are no more able to modify m is due to the presence of the constant g that must be the first component of the signed vector associated to  $\sigma_1$  and any adaptation  $\sigma'_1$ . Modifying *m* requires computing a one-time LHSP signature on a vector necessarily outside the span generated by  $(g, c_0, c_1, c_2)$  and (1, f, g, h). Second, the signature  $\sigma_2$  is unchanged during the randomization. Still, it is a signature on a fixed vector and the one-time LHSP signing algorithm is deterministic. Moreover, if we have two distinct signatures

 $<sup>^2~</sup>H$  must not only be second-preimage resistance as in [29] since the adversary can choose  $\mathsf{opk}^*$  adaptively.

on a single vector we can solve SXDH. That means that any other adversarially generated ciphertext for opk (as the adversary might know osk) will have to share  $\sigma_2$  and our randomizability holds. Third, while the tag-based simulation-sound QA-NIZK proof  $\pi$  can be simulated if we embed a random triple (F, G, H) into  $(c_0, c_1, c_2)$  we also have to produce a valid looking adaptation of  $\sigma_1$  while we do not know osk<sup>\*</sup>. To avoid extracting osk from a costly bit-by-bit proof of knowledge in the standard model since osk consists of random scalars,<sup>3</sup> we would like to add (1, F, G, H) in the public key and requires the ciphertext to further compute a signature  $\sigma_3$  on it with osk. However, if we reveal (1, F, G, H), computing  $(g, c_0^*, c_1^*, c_2^*) = (g, c_0, c_1, c_2) \cdot (1, f, g, h)^{\theta^*} \cdot (1, F, G, H)^{\rho^*}$  allows deriving a valid  $\sigma_1^* = \sigma_1 \cdot \sigma_2^{\theta^*} \cdot \sigma_3^{\rho^*}$  but  $\mathbf{c}^* = (c_0^*, c_1^*, c_2^*)$  will not be random even if (F, G, H) is random. Fortunately, it is actually sufficient for the traceability to use (opk, osk) to sign the shorter vectors  $(g, c_0, c_1), (1, f, g)$  and (1, F, G), and keep H away from the adversary's view to have a statistically random  $\mathbf{c}^*$  in the reduction. When (1, F, G) is not in span $\langle (1, f, g) \rangle$ , the proof  $\pi$  simply prevents the adversary from randomizing ciphertexts with (1, F, G) without losing validity.

For technical reason, we hide  $\sigma_1$  in the ciphertext and make a randomizable NIWI Groth-Sahai proof to show the randomizability and the TCCA security of the scheme. While we can adapt the  $\sigma_1$  component when we randomize one of the two ciphertexts given by the adversary in the challenge phase (or in the randomization experiment), and that trace to each other, since the adversary might know osk it might infer more information about how we adapt this signature into  $\sigma_1^*$  if we left it in the clear.

## 4.4 Description

- **Gen**(1<sup> $\lambda$ </sup>): Choose bilinear groups ( $\mathbb{G}, \hat{\mathbb{G}}, \mathbb{G}_T$ ) of prime order  $p > 2^{\text{poly}(\lambda)}$  together with  $g, h \leftarrow \$ \mathbb{G}$  and  $\hat{g}, \hat{h} \leftarrow \$ \hat{\mathbb{G}}$ .
  - 1. Pick random  $\alpha, \beta \leftarrow \mathbb{Z}_p$  and set  $f = g^{\alpha} h^{\beta}$ .
  - 2. Pick  $\delta \leftarrow \mathbb{Z}_p$  and compute  $(F, G, H) = (f, g, h)^{\delta}$ .
  - 3. Generate a Groth-Sahai CRS  $\operatorname{crs}_w = (\vec{w}_1, \vec{w}_2) \in \mathbb{G}^4$  to commit to groups elements of  $\mathbb{G}$ , where  $\vec{w}_1 = (w_{11}, w_{12})$  and  $\vec{w}_w = (w_{21}, w_{22})$  are generated in the perfect NIWI mode, i.e.,  $\operatorname{crs}_w \leftarrow \$ \mathbb{G}^4$ .
  - 4. Define the vector  $\mathbf{v} = (g, h)$  and generate 2 key pairs  $(\mathsf{sk}_u, \mathsf{pk}_u)$  and  $(\mathsf{sk}_u, \mathsf{pk}_u)$  for the one-time linearly homomorphic signature of Section 4.2 in order to sign vectors of dimension n = 2, given the common public parameters  $\hat{g}, \hat{h}$ . Let  $\mathsf{pk}_u = \{\hat{u}_1, \hat{u}_2\}$  and  $\mathsf{pk}_v = \{\hat{v}_1, \hat{v}_2\}$ . Using  $\mathsf{sk}_u$  (resp.  $\mathsf{sk}_v$ ), generate a one-time LHSP signature  $\Sigma_u = (Z_u, R_u)$  (resp.  $\Sigma_v = (Z_v, R_v)$ ) on  $\mathbf{v}$ . In other words, for  $\mathsf{pk}_{\mathsf{lhsp}}^{\mathsf{qazk}} = \{\hat{u}_1, \hat{u}_2, \hat{v}_1, \hat{v}_2\}, \Sigma_u, \Sigma_v$  are

<sup>&</sup>lt;sup>3</sup> There is no fully structure-preserving signature schemes under SXDH and none with full randomizability (except in the generic group model [24]), which might still not be enough to be combined with a ciphertext as an SRC). And, we are not aware of any fully structure-preserving LHSP signature scheme, where the secret keys only contain source group elements.

one-time LHSP signatures on the rows of the matrix

$$\mathbf{P} = \begin{pmatrix} g & h & 1 & 1 \\ 1 & 1 & g & h \end{pmatrix}.$$

The private key consists of  $\mathsf{SK} = (\alpha, \beta)$  and the public key  $\mathsf{PK} \in \mathbb{G}^{13} \times \hat{\mathbb{G}}^6$  is

$$\mathsf{PK} = \left(f, \ g, \ h, \ F, \ G, \ \mathsf{crs}_w, \ \varSigma_u, \ \varSigma_v, \ \mathsf{pk}_{\mathsf{lhsp}}^{\mathsf{qazk}}, \ \hat{g}, \ \hat{h}\right)$$

- **Enc(PK, m):** to encrypt a message  $m \in \mathbb{G}$ , first run LGen(PK): Generate a key pair (osk, opk) for the one-time linearly homomorphic signature of Section 4.2 from the public generators  $\hat{g}, \hat{h}$  in order to sign vectors of dimension 3. Let  $|\mathbf{k} = \mathbf{osk} = \{(\eta_i, \zeta_i)\}_{i=1}^3$  be the private key, of which the corresponding public key is  $\mathbf{opk} = \{\hat{f}_i\}_{i=1}^3$ . Then, conduct the following steps of LEnc(PK, lk, m):
  - 1. Pick  $\theta \leftarrow \mathbb{Z}_p$  and compute the CPA encryption  $\mathbf{c} = (c_0, c_1, c_2)$ , where  $c_0 = mf^{\theta}, c_1 = g^{\theta}$  and  $c_2 = h^{\theta}$ , and keep the random coin  $\theta$ . Next steps 2-3 are dedicated to the tracing part.
  - 2. To allow tracing, authenticate the row space of the matrix  $\mathbf{T} = (T_{i,j})_{1 \le i,j \le 3}$

$$\mathbf{T} = \begin{pmatrix} g & c_0 & c_1 \\ 1 & f & g \\ 1 & F & G \end{pmatrix}$$
(2)

by using  $\mathsf{lk} = \mathsf{osk}$ . Namely, sign each row  $\vec{T_i} = (T_{i,1}, T_{i,2}, T_{i,3})$  of **T** resulting in  $\boldsymbol{\sigma} = (\sigma_i)_{i=1}^3 \in \mathbb{G}^6$ , where  $\sigma_i = (Z_i, R_i) \in \mathbb{G}^2$ .

- 3. To allow strong randomizability, commit to  $\sigma_1$  using the Groth-Sahai CRS  $\operatorname{crs}_w$  by computing  $C_Z = \iota(Z_1) \vec{w}_1^{z_1} \vec{w}_2^{z_2}$  and  $C_R = \iota(R_1) \vec{w}_1^{r_1} \vec{w}_2^{r_2}$ . To ensure that  $\sigma_1$  is a valid one-time LHSP signature on  $(g, c_0, c_1)$  compute the proof  $\hat{\pi}_{\operatorname{sig}} = (\hat{P}_1, \hat{P}_2) \in \hat{\mathbb{G}}^2$  such that  $\hat{P}_1 = \hat{g}^{z_1} \hat{h}^{r_1}$  and  $\hat{P}_2 = \hat{g}^{z_2} \hat{h}^{r_2}$ . Next step 4 shows the validity of **c** associated to the tag  $\tau = H(\operatorname{opk})$ .
- 4. Given  $\theta$  and  $\tau = H(\mathsf{opk})$ , compute a randomizable simulation-sound proof that  $(c_1, c_2) \in \operatorname{span}\langle (g, h) \rangle$ . Namely, derive the LHSP signature  $\pi = (\Sigma_u^{\tau} \Sigma_v)^{\theta} = :(Z_{\pi}, R_{\pi})$  on the vector  $(c_1^{\tau}, c_2^{\tau}, c_1, c_2) = ((g, h, 1, 1)^{\tau}(1, 1, g, h))^{\theta}$ .

Output the ciphertext

$$\mathsf{CT} = \left(\mathbf{c}, \mathbf{C}_Z, \mathbf{C}_R, \sigma_2, \sigma_3, \pi, \hat{\pi}_{\mathsf{sig}}, \mathsf{opk} = \{\hat{f}_i\}_{i=1}^3\right) \in \mathbb{G}^{13} \times \hat{\mathbb{G}}^5$$

Trace(PK, CT): Parse PK and CT as above, and output opk in the obvious way. Rand(PK, CT): If PK and CT do not parse as the outputs of Gen and Enc, abort. Otherwise, conduct the following steps:

1. Parse the CPA encryption part  $\mathbf{c} = (c_0, c_1, c_2)$ , pick  $\theta' \leftarrow \mathbb{Z}_p$  and compute  $\mathbf{c}' = \mathbf{c} \cdot (f, g, h)^{\theta'}$ , so that  $c'_0 = c_0 f^{\theta'}, c'_1 = c_1 g^{\theta'}$  and  $c'_2 = c_2 h^{\theta'}$ .

18

- Implicitly adapt the committed signature σ<sub>1</sub> of the tracing part. First, compute σ̃<sub>1</sub> = (Z̃<sub>1</sub>, R̃<sub>1</sub>) = (Z<sup>θ'</sup><sub>2</sub>, R<sup>θ'</sup><sub>2</sub>) = σ<sup>θ'</sup><sub>2</sub>, which consists of a one-time LHSP signature on (1, f, g)<sup>θ'</sup> for opk. Second, adapt the commitments C'<sub>Z</sub> = C<sub>Z</sub> · ι(Z̃<sub>1</sub>) w<sup>z'<sub>1</sub></sup><sub>1</sub> w<sup>z'<sub>2</sub></sup><sub>2</sub> and C'<sub>R</sub> = C<sub>R</sub> · ι(R̃<sub>1</sub>) w<sup>r'<sub>1</sub></sup><sub>1</sub> w<sup>r'<sub>2</sub></sup><sub>2</sub>, for some random scalars z'<sub>1</sub>, z'<sub>2</sub>, r'<sub>1</sub>, r'<sub>2</sub> ← Z<sub>p</sub>, which should commit to the valid one-time LHSP signature σ'<sub>1</sub> = σ<sub>1</sub>σ<sup>θ'</sup><sub>2</sub> on (g, c'<sub>0</sub>, c'<sub>1</sub>) for opk. Third, adapt the proof π̂<sub>sig</sub> = (P̂<sub>1</sub>, P̂<sub>2</sub>) as π<sup>'</sup><sub>sig</sub> = (P̂<sub>1</sub>', P̂'<sub>2</sub>), where P̂'<sub>1</sub> = P̂<sub>1</sub> · ĝ<sup>z'<sub>1</sub></sup> ĥ<sup>r'<sub>1</sub></sup> and P̂'<sub>2</sub> = P̂<sub>2</sub> · ĝ<sup>z'<sub>2</sub></sup> ĥ<sup>r'<sub>2</sub></sup>.
  Adapt the proof of the validity of the CPA ciphertext. Namely, computes
- $\pi' = \pi \cdot (\tilde{\Sigma}_u^{\tau} \Sigma_v)^{\theta'} = (Z_{\pi} (Z_u^{\tau} Z_v)^{\theta'}, R_{\pi} (R_u^{\tau} R_v)^{\theta'}), \text{ where } \tau = H(\mathsf{opk}).$

Output the re-randomized ciphertext

$$\mathsf{CT} = \left(\mathbf{c}', \mathbf{C}'_Z, \mathbf{C}'_R, \sigma_2, \sigma_3, \pi', \hat{\pi}'_{\mathsf{sig}}, \mathsf{opk}\right)$$

- **Ver(PK, CT):** First, abort and output 0 if PK or CT does not parse properly. Second, verify the validity of the signatures  $\sigma_2$  and  $\sigma_3$  on the 2 last rows  $\{\vec{T}_i\}_{i=2}^3$  of the matrix **T**, and output 0 if it does not hold. Third, verify that:
  - 1. The committed signature of the tracing part is valid, i.e.,  $\sigma_1 = (Z_1, R_1)$ is a valid one-time LHSP signature on the vector  $(g, c_1, c_2)$ . To hold, the commitments  $C_Z, C_R$  and the proof  $\hat{\pi}_{sig} = (\hat{P}_1, \hat{P}_2)$  must satisfy

$$E(\mathbf{C}_{Z}, \hat{g}) \cdot e(\mathbf{C}_{R}, \hat{h}) = E(\iota(g), \hat{f}_{1}) \cdot E(\iota(c_{0}), \hat{f}_{2}) \cdot E(\iota(c_{1}), \hat{f}_{3})$$
(3)  
 
$$\cdot E(\vec{w}_{1}, \hat{P}_{1}) \cdot E(\vec{w}_{2}, \hat{P}_{2}));$$

2. The proof that the CPA ciphertext is valid, i.e.,  $\pi = (Z_{\pi}, R_{\pi})$  is a valid one-time LHSP signature on the vector  $(c_1^{\tau}, c_2^{\tau}, c_1, c_2)$ , which must satisfy

$$e(Z_{\pi}, \hat{g}) \cdot e(R_{\pi}, \hat{h}) = e(c_1, \hat{u}_1^{\tau} \hat{v}_1) \cdot e(c_2, \hat{u}_2^{\tau} \hat{v}_2), \tag{4}$$

where  $\tau = H(\mathsf{opk})$ .

If at least one of theses checks fails, output 0, otherwise, output 1.

**Dec(SK, PK, CT):** If Ver(PK, CT) = 0, output  $\perp$ . Otherwise, given SK =  $(\alpha, \beta)$  and  $\mathbf{c} = (c_0, c_1, c_2)$  included in CT, compute and output  $m = c_0 \cdot c_1^{-\alpha} \cdot c_2^{-\beta}$ .

## 4.5 Security

The security of our pairing-based TREnc relies solely on the SXDH assumption. We first show the verifiability of this TREnc as it eases the analysis of the traceability and the randomizability properties. The verifiability essentially relies on the unforgeability of LHSP signatures since it also implies the (simulation-) soundness of the (quasi-adaptive) proof of (subspace) membership. We refer to the full version of the paper [21] for all the proofs of the theorems.

**Theorem 2.** The above TREnc is verifiable under the SXDH assumption. More precisely, for any adversary  $\mathcal{A}$ , we have  $\Pr[\mathsf{Exp}_{\mathcal{A}}^{\mathrm{ver}}(\lambda) = 1] \leq \varepsilon_{\mathsf{sxdh}} + 1/p$ .

**Theorem 3.** The above TREnc  $\Pi$  is traceable under the SXDH assumption. More precisely, for any adversary  $\mathcal{A}$ , we have  $\Pr[\mathsf{Exp}_{\mathcal{A},\Pi}^{\mathrm{trace}}(\lambda) = 1] \leq 2 \cdot \varepsilon_{\mathsf{sxdh}} + 2/p$ .

Strong randomizability essentially relies on the verifiability, which shows that computationally-bounded adversary only produces (except with negligible probability) valid ciphertexts that are honest (but, possibly with biased randomness), and the perfect randomization of honest ciphertexts.

**Theorem 4.** The above TREnc is strongly randomizable under the SXDH assumption. More precisely, for any adversary  $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ , where  $\mathcal{A}_2$  is possibly unbounded, we have  $\Pr[\mathsf{Exp}^{\mathrm{rand}}_{\mathcal{A}}(\lambda) = 1] \leq \varepsilon_{\mathsf{sxdh}} + 2/p$ .

**Theorem 5.** The above TREnc is TCCA-secure under the SXDH assumption and the collision resistance of the hash function. More precisely, we have  $\Pr[\mathsf{Exp}_{\mathcal{A},\Pi}^{tcca}(\lambda) = 1] \leq \frac{1}{2} + \varepsilon_{cr} + 2 \cdot \varepsilon_{sxdh} + \Omega(2^{-\lambda}).$ 

### 4.6 Efficiency

This TREnc instance is reasonably efficient. In particular, in order to encrypt a message, which is typically the bottleneck in voting applications because it must run more or less transparently on low-end voter devices, we can encrypt one group element using 29 exponentiations in  $\mathbb{G}$  and 10 exponentiations in  $\hat{\mathbb{G}}$ . This group element would make it possible to encode up to a few hundred bits in practice, depending on the chosen security parameter.

In contrast, the SRC aiming at similar applications and used in the BeleniosRF election system [15] requires 33 exponentiations in  $\mathbb{G}$  and 22 exponentiations in  $\hat{\mathbb{G}}$  for the (signed) encryption of only 1 bit. In general, their construction requires 11k + 22 exponentiations in  $\mathbb{G}$  and 10k + 12 exponentiations in  $\hat{\mathbb{G}}$  in order to encrypt k bits. These estimates are based on the reference code of the SRC, since the paper does not entirely specify the algorithms (especially how commitments and proofs are computed).

## 5 Voting scheme based on Traceable Receipt-Free Encryption

Traceable Receipt-Free Encryption schemes are particularly well suited for the design of voting systems offering receipt freeness, that is, systems in which voters cannot demonstrate how they voted to a third party.

We are now formalizing the notion of voting system (Sec. 5.1) and receiptfreeness (Sec. 5.2), using a definition closely related to the one of Chaidos et al. [15], while fixing two technical issues that it contains, then show how to build a receipt-free voting scheme from a TREnc (Sec. 5.3).

### 5.1 Definitions and notations

We define voting protocols in a way that largely follows the SOK from Bernhard et al. [8] and BeleniosRF [15]. In our voting protocols, we consider the following parties:

- The voters are participating in the election and are willing to cast a ballot representing their vote intent.
- The *election administrator* is organizing the election and is responsible for coordinating the protocol execution.
- The *ballot box manager* is gathering the ballots of the voters on a bulletin board BB and provides a public view PBB of those ballots, for verifiability.
- The *trustees* are responsible for correctly tallying the ballot box and providing a proof of the correctness of the tally. We consider a k-threshold tallying system, that is k honest trustees are required to compute the tally of the election.

These parties are standard entities in the voting literature. In some cases, we will also refer to the ballot box manager as the rerandomizing server, in order to make its receipt-freeness related role more visible. We also define a family of deterministic results functions  $\rho_m$  which given m votes, returns the result of the election for these votes. The following definition encompasses the procedures used in a voting system.

**Definition 10 (Voting System).** A Voting System is a tuple of probabilistic polynomial-time algorithms (SetupElection, Vote, ProcessBallot, TraceBallot, Valid, Append, Publish, VerifyVote, Tally, VerifyResult) associated to a result function  $\rho_m : \mathbb{V}^m \cup \{\bot\} \to \mathbb{R}$  where  $\mathbb{V}$  is the set of valid votes and  $\mathbb{R}$  is the result space such that:

- SetupElection(1<sup>λ</sup>): on input security parameter 1<sup>λ</sup>, generate the public and secret keys (pk, sk) of the election.
- Vote(id, v): when receiving a voter id and a vote v, outputs a ballot b and auxiliary data aux. It will also be possible to call Vote(id, v, aux) in order to obtain a ballot (without auxiliary data this time) for vote v using aux. This auxiliary data will be useful to define security and enables the creation of ballots that share the same aux.
- ProcessBallot(b): on input ballot b, outputs an updated ballot b'. In our case,
  b' would be a rerandomization of b.
- TraceBallot(b): on input ballot b, outputs a tag t. The tag is the information that a voter can use to trace his ballot, using VerifyVote.
- Valid(BB, b): on input ballot box BB and ballot b, outputs 0 or 1. The algorithm outputs 1 if and only if the ballot is valid.
- Append(BB, b): on input ballot box BB and ballot b, appends ProcessBallot(b) to BB if Valid(BB, b) = 1.
- Publish(BB): on input ballot box BB, outputs the public view PBB of BB, which is the one that is used to verify the election. Depending on the context, it may be used to remove some voter credentials for instance.
- VerifyVote(PBB,t): on input public ballot box PBB and tag t, outputs 0 or 1. This algorithm is used by voters to check if their ballot has been processed and recorded properly.
- Tally(BB, sk): on input ballot box BB and private key of the election sk, outputs the tally r and a proof  $\Pi$  that the tally is correct w.r.t. the result function  $\rho_m$ .

- 22 Henri Devillez, Olivier Pereira, and Thomas Peters
  - VerifyResult(PBB,  $\mathbf{r}, \mathbf{\Pi}$ ): on input public ballot box PBB, result of the tally  $\mathbf{r}$  and proof of the tally  $\mathbf{\Pi}$ , outputs 0 or 1. The algorithm outputs 1 if and only if  $\mathbf{\Pi}$  is a valid proof that r is the election result, computed w.r.t.  $\rho_m$ , corresponding to the ballots on PBB.

For all of these algorithms except SetupElection, the public key of the election pk is an implicit argument.

These algorithms are used as follows in a typical election, the election authorities first generate the election public and secret keys with SetupElection. Then, using the public key of the election, each voter can prepare a ballot bwith the Vote algorithm and send it to the ballot box manager. The voter also keeps TraceBallot(b) in order to be able to trace its ballot on the election public bulletin board. Each time the ballot box manager receives a ballot, it checks if it is valid with the Valid algorithm. If this is the case, it runs the ProcessBallot algorithm on it and appends the resulting ballot to the ballot box using Append.

The ballot box manager also applies Publish on the ballot box in order to obtain the content that is made available on a public bulletin board PBB. Voters can check that their ballot has been correctly recorded on PBBusing VerifyVote.

Eventually, the trustees run the Tally algorithm on the ballot box in order to compute the election result and a proof of correctness of this result. Anyone can use these, together with the content of PBB, in order to verify the election result using VerifyResult.

This definition differs from [8] and [15] in two important ways. First, we introduce the TraceBallot algorithm. Such a procedure is implicit other voting system descriptions, often because voters simply check the presence on PBBof their ballot, in which case TraceBallotwould simply be the identity function. In our case, TraceBallotmust extract the signature verification key that is generated at voting time by Vote, making this algorithm non-trivial.

The correctness guarantees of the various algorithms listed above are as usual and follow the intuitions given above. We only formalize the correctness guarantee of TraceBallot, which is novel.

**Definition 11 (Tracing correctness).** For every v, BB,  $(b, aux) \leftarrow Vote(id, v)$ and  $t \leftarrow TraceBallot(b)$ , after Append(BB, b) we have that VerifyVote(Publish(BB), t) = 1 with overwhelming probability.

As a second difference, we omit the voter registration procedure, of which we make no use here: it is used in some protocols in order to obtain some forms of delegated verifiability where an extra authority is partially trusted to handle voter credentials, but this is not our focus. To make things concrete here, one can imagine that voter authentication is handled with a process similar to the one used in Helios [3], where the ballot box manager distributes credentials (e.g., passwords) to the voters and publishes the voter names next to their ballot on PBB. Voters who did not vote can then verify that there is no ballot recorded for them, and auditors can sample voters and contact them to perform similar verification steps. (We make no claim regarding the effectiveness of this process in practice – it is just here for context.)

#### 5.2 Receipt-freeness

We adopt here a definition of receipt-freeness that is similar to the one of Chaidos et al [15]. Various other definitions exist, but they are either too informal for our purpose (e.g., [25]), or focus on the stronger notion of coercion resistance, in which voters need to adopt a specific counter-strategy depending on instructions of the coercer (e.g., [35,40,31,4]).

This definition requires that voters should not be able to pick a ballot, possibly from a distribution that deviates from the honest one, in such a way that no third party, by looking at the election public bulletin board, and knowing exactly how the voter's announced ballot was built, is able to decide whether that ballot was submitted by the voter rather than another ballot that could encode a vote for a different candidate. This definition also considers that the channels between the voters and the ballot box manager is private and, indeed, without the assumption that such private channels are available, achieving receipt-freeness in a verifiable election is impossible [18].

**Definition 12 (Receipt-Freeness).** A voting system  $\mathcal{V}$  has receipt-freeness if there exists algorithms SimSetupElection and SimProof such that no PPT adversary  $\mathcal{A}$  can distinguish between games  $\operatorname{Exp}_{\mathcal{A},\mathcal{V}}^{\operatorname{srf},0}(\lambda)$  and  $\operatorname{Exp}_{\mathcal{A},\mathcal{V}}^{\operatorname{srf},1}(\lambda)$  defined by the oracles in Figure 3, that is for any efficient algorithm  $\mathcal{A}$ :

$$Pr\left[\mathsf{Exp}_{\mathcal{A},\mathcal{V}}^{\mathsf{srf},0}(\lambda)=1\right] - Pr\left[\mathsf{Exp}_{\mathcal{A},\mathcal{V}}^{\mathsf{srf},1}(\lambda)=1\right]$$

is negligible in  $\lambda$ .

$\mathcal{O}init(\lambda)$	$\mathcal{O}receiptLR(b_0,b_1)$
$      if \beta = 0 then (pk, sk) \leftarrow SetupElection(1\lambda)       else (pk, sk, \tau) \leftarrow SimSetupElection(1\lambda)       BB0 \leftarrow \perp: BB1 \leftarrow \perp:       return pk                                   $	$\label{eq:approx_state} \begin{array}{l} \mbox{if } TraceBallot(b_0) \neq TraceBallot(b_1) \\ \mathrm{or } Valid(BB_0, b_0) = 0 \ \mathrm{or } Valid(BB_1, b_1) = 0 \\ \mbox{then } return \ \bot \\ \mbox{else } Append(BB_0, b_0); Append(BB_1, b_1) \end{array}$
$\mathcal{O}board()$	$\mathcal{O}tally()$
$\mathbf{return} \; Publish(BB_{eta})$	

Fig. 3. Oracles used in the  $\mathsf{Exp}_{\mathcal{A},\mathcal{V}}^{\mathsf{srf},\beta}(\lambda)$  experiment. The adversary first calls  $\mathcal{O}$ init and then can call Oboard and OreceiptLR as much as it wants. Finally, the adversary calls  $\mathcal{O}$ tally, receives the result of the election and must return its guess, which is the output of the experiment.

In this game, parameterized by a bit  $\beta$ , the adversary has access to the following oracles:

-  $\mathcal{O}$ init: initializes the voting system. It generates the public and privates keys of election and returns the public key to the adversary. When  $\beta = 1$ , a

simulated setup may be performed, depending on the computational model, which will offer some trapdoor information that may be needed to produced a simulated tally correctness proof for instance. Eventually, two empty ballot boxes are created: the real ballot box  $BB_0$  that will be tallied and the fake ballot box  $BB_1$ . Both boxes will be populated during the game, but the adversary will only see Publish( $BB_\beta$ ).

- $\mathcal{O}receiptLR(b_0, b_1)$ : Lets the adversary cast ballots  $b_0$  in the real ballot box  $BB_0$  and  $b_1$  in the fake ballot box  $BB_1$ , as long as both ballots are valid and have the same trace. This oracle is the central one for receipt-freeness.
- $\mathcal{O}$ board: Returns Publish(BB<sub> $\beta$ </sub>), which represents its view of the public bulletin board.
- $\mathcal{O}$ tally: Returns the result of the election based on the ballots on BB<sub>0</sub>, as well as a proof of correctness of the tally. If  $\beta = 1$ , this proof is simulated w.r.t. the content derived from BB<sub>1</sub>.

Several observations can be made about this game. First, and as expected, it can be seen that the ballot box manager is considered to behave honestly. A dishonest ballot box manager could simply replace ProcessBallot and Publish with the identity function, which would make the game trivial to win, independently of any cryptographic operation. The Tally operation is also performed honestly: dishonest talliers could decrypt all the ballots individually, which would again make the game trivial to win. In practice, this assumption can be mitigated by using a distributed decryption process, which is always possible using MPC but can typically be done more efficiently.

Second, this game prompts for the introduction of an extra correctness requirement on the definition of Vote and TraceBallot, in order to make sure that ballots that encode different votes and have the same tag can be computed.

**Definition 13 (Ballot traceability for receipt freeness).** For every public key pk in the range of SetupElection, the ballots produced for every pair of voting choices  $(v_0, v_1)$  with the same auxiliary data trace to the same tag. That is, for  $b_0$ ,  $aux \leftarrow$ \$Vote $(id, v_0)$ ,  $b_1 \leftarrow$ \$Vote $(id, v_1, aux)$ , we have TraceBallot $(b_0) =$ TraceBallot $(b_1)$ .

Without this extra constraint, we could imagine a TraceBallot algorithm which returns a tag depending on the vote inside the ballot. For example if TraceBallot(b) = b as we discussed earlier, then  $\mathcal{O}receipt(b_0, b_1)$  does nothing except if the two ballots are identical and the adversary can never win the receipt-freeness game. It is thus natural to require that a tag returned by TraceBallot can be reached with any possible voting choice.

The related constraint that  $\mathsf{TraceBallot}(b_0) = \mathsf{TraceBallot}(b_1)$  is actually missing from Chaidos' definition[15], and makes their game trivial to win in most natural case, including with their own protocol: an adversary could simply submit two ballots that have different traces (or are signed with different keys in the wording of their paper), and immediately identify which bulletin board he sees.

Compared to Chaidos' definition, we also removed the  $\mathcal{O}$ corrupt oracle, as we simply assume that all the voters are under adversarial control. We also omitted their  $\mathcal{O}$ cast and  $\mathcal{O}$ voteLR oracles, because  $\mathcal{O}$ receiptLR subsumes them.

#### 5.3 Voting scheme

We now explain how a generic voting system can be built from a TREnc. The protocol in itself is of little interest: it essentially follows previous proposals [10,15]. Its central interest is that defining it from a TREnc makes the proof of its receipt-freeness almost immediate, and independent of any specific TREnc instance.

A detailed pseudocode description is proposed in Figure 5.3. The protocol executes as follow. The election authorities set-up the election in the following way. They create an empty bulletin board and create the pair of public and private keys (pk, sk) of the election by running the Gen algorithm of the TREnc scheme with the desired security parameter. pk is distributed to every party taking part to the election and sk is given to the tallier. Note that sk can be generated in a distributed way, so that decryption requires the contribution of multiple trustees – our TREnc constructions are compatible with standard threshold key generation protocols for discrete log based cryptosystems, which can be used as usual since the decryption of a ciphertext is independent from the link key and the Trace algorithm [19,23]. We consider a unique tallier in the following.

When a user wants to cast a vote v, they first generate a link key lk by running the LGen(pk) algorithm, then encrypts the vote with the LEnc(pk, lk, v) algorithm in order to obtain a ballot b, while aux is defined as lk. The voter then sends the encrypted ballot to the ballot box manager of the voting system. It is utterly important that the user erases the link key as soon as possible, as the integrity of their vote may rely on the secrecy of this key. The voter will however store TraceBallot(b) = Trace(b) in order to verify that a ballot with the correct trace eventually appears on the public bulletin board.

When the ballot box manager receives a new ballot b, he verifies the validity of the ballot by checking that Ver(b) succeeds and that no ballot with the same trace was recorded before. Invalid ballots are dropped and valid ones are going through Append(BB, b), which runs ProcessBallot(b) = Rand(pk, b) and appends the result to BB.

The user can verify that their vote is on the bulletin board by checking with the **TraceBallot** algorithm if any entry in the public bulletin board has the same trace as the one they recorded when they produced their ballot. The traceability property of the TREnc then guarantees that nobody (including the rerandomizing server and the election authorities who hold the decryption key) could have forged another valid ciphertext of another vote linked to this ballot with non-negligible probability.

Once every voter has cast a vote, the tallier can gather the ballots on the bulletin board and compute the result of the election  $\mathbf{r}$ , as well as a proof of correctness  $\Pi$ . The exact details of this process will depend on the ballot format and the result function  $\rho_m$  that the voting protocol requires. One standard way of performing this operation would be to process all the published ballots through a verifiable mixnet: our TREnc ciphertexts are compatible with various standard options that operate on votes encrypted as vectors of group elements, including the Verificatum mixnet [42].

SetupElection $(\lambda)$	Vote(id, v[, aux])
$(pk,sk) \leftarrow Gen(1^{\lambda})$ return pk	if aux is specified then $lk \leftarrow aux$ else $lk \leftarrow \$ LGen(pk)$ $b \leftarrow LEnc(pk, lk, v)$ if aux is not specified then return b else return b, lk
TraceBallot(b)	Valid(BB, b)
return Trace(b)	if $Ver(b) \land \\ \forall b' \in BB, TraceBallot(b') \neq TraceBallot(b) \\ then return 1 else return 0$
ProcessBallot(b)	VerifyVote(PBB,t)
$\overline{\mathbf{return}} \ Rand(pk,b)$	if $\exists b \in PBB$ : $Valid(b) \land t ==TraceBallot(b)$ then return 1 else return 0

Fig. 4. Instantiation of our voting scheme from a generic TREnc scheme. Publish is simply the identity function. Tally and VerifyResult are instantiated via standard techniques, depending on the result function (homomorphic tallying, verifiable mixnet,  $\ldots$ ).

## 5.4 Security of the voting scheme

This voting scheme has receipt freeness, as claimed in the following theorem.

**Theorem 6.** If the TREnc used in the voting scheme is TCCA secure and verifiable and if the proof system used to prove the correctness of the tally is zero-knowledge, then the voting scheme has receipt freeness. More precisely, if the advantage of any adversary at distinguishing a simulator from an honest prover of the proof system is bounded by  $\varepsilon_{ZK}$  and if the advantage of any adversary at the TCCA experiment is bounded by a negligible function  $\varepsilon_{TCCA}$ , then every adversary at the receipt freeness game making  $q_r$  OreceiptLR requests has an advantage bounded by  $\varepsilon_{ZK} + q_r \varepsilon_{TCCA}$ .

*Proof.* The proof uses two different games, where the first one is the receiptfreeness game. In each of those games, we pick a random bit  $\beta$  corresponding to either the real ballot box ( $\beta = 0$ ) or the fake ballot box ( $\beta = 1$ ). The adversary is expected to guess in which case it is and to output a bit  $\beta'$ . We note  $S_i$  the event that  $\beta = \beta'$  in the *i*-th game. We show that  $S_0$ , the probability of an adversary to win the receipt freeness game, is negligibly close in  $\lambda$  to  $\frac{1}{2}$ .

- $\mathsf{Game}_0(\lambda)$ : We define  $\mathsf{Game}_0(\lambda)$  as the original receipt freeness experiment and  $\mathcal{A}$  as a PPT adversary for the game. We set SimSetup as the simulation trapdoor for the proof systems used in the tally and SimProof(BB, r) as an algorithm simulating fake proofs of the decryption of the ciphertexts in BB into the plaintexts listed in r. By definition,  $\mathcal{A}$  wins the game with probability  $Pr[S_0]$ .
- Game<sub>1</sub>( $\lambda$ ): If  $\beta = 0$ , we generate the keys of the election with SimSetup instead of the honest Setup algorithm. We also give a simulated proof of the tally as

in the case  $\beta = 1$ .

 $\mathsf{Game}_0(\lambda) \to \mathsf{Game}_1(\lambda)$ : Since the proof system of the tally is zero-knowledge,  $|Pr[S_0] - Pr[S_1]| \leq \varepsilon_{ZK}.$ 

In the second game, the only difference between  $\beta = 0$  and  $\beta = 1$  are now in the  $\mathcal{O}$ receiptLR oracle. We can reduce the q-TCCA experiment (the q challenge variant of the TCCA game, which is proven to be equivalent to TCCA security up to a factor q in the full version) to  $\mathsf{Game}_1(\lambda)$  in the following way. We build an adversary against the q-TCCA challenger by instantiating the voting system and simulating an efficient adversary for  $\mathsf{Game}_1(\lambda)$ . Each time we are asked to append ballots to the bulletin board from a OreceiptLR oracle call, we ask the challenger to decrypt them. Then, we give both ballots as a challenge to the challenger and receive a randomized ballot that we append to our bulletin board. There are  $q_r$  such requests. After all the requests, we can compute the result of the election as we have the plaintext of every ballot in  $BB_0$ . Moreover, we can use SimProof to simulate a proof that our bulletin board has been correctly tallied. Hence, this adversary wins the q-TCCA game with the same probability as the simulated adversary wins the second game and  $Pr[S_1] \leq q_r \varepsilon_{TCCA}$ . We conclude that the probability that the adversary wins the receipt freeness experiment,  $Pr[S_0]$ , is bounded by  $\varepsilon_{ZK} + q_r \varepsilon_{TCCA}$ . 

It is immediate that our voting scheme also satisfies ballot traceability (Def. 13), thanks to the link traceability of the TREnc (Def. 2).

This demonstrates the receipt-freeness of our protocol against an adversary who sees the public bulletin board, and assuming a honest ballot box manager. Our protocol also offers privacy against a malicious ballot box manager, as demonstrated in the full version. As can be expected, proving privacy against such an adversary requires taking advantage of the strong-randomization property of the TREnc, which was not necessary for receipt-freeness.

It is also important to note that the notions of receipt-freeness and ballot privacy only make sense when applied to voting protocols that satisfy some extra correctness requirements (see Bernhard et al. [8] for instance) – a pathological Valid process that would just drop all but one ballot could result in this ballot been tallied alone, which could satisfy the definition or receipt-freeness but would obviously be problematic from a privacy point of view. The notions of strong consistence, correctness, and validity, defined in the full version of the paper, address these questions.

We do not detail the verifiability of our voting system, which would require to introduce a substantial machinery. We outline how this could work here:

- Individual verifiability requires that a voter who successfully completes the VerifyVote verification steps can be convinced that his vote is properly recorded. If the voter's voting client is honest, this follows from the traceability property of the TREnc and the single use of lk, which guarantee that any ballot with the same trace as the ballot submitted by the voter would decrypt

to the right vote. Detecting a malicious voting client that may encrypt a vote different of the one chosen by the voter is more tricky. One option would be to consider a variation on the Benaloh challenge, in which voters would have the option to decide to spoil a ballot that has been posted on the public bulletin board, and either ask for its decryption, or for the randomness used both during the Vote process and during ProcessBallot. Any newly created ballot would need to be generated using a fresh lk.

- Eligibility verifiability could proceed by adding the voter's name next to each ballot on the public bulletin board, and let auditors check whether these are legitimate voters. Weaker but more convenient options include relying on a trusted authentication server and/or on a PKI.
- Universal verifiability, which guarantees that the tally is computed correctly, would result from the tallying process, e.g., from a verifiable mix-net.

Acknowledgments. This work has been funded in part by the European Union (EU) and the Walloon Region through the FEDER project USERMedia (convention number 501907-379156). Henri Devillez is a research fellow of the Belgian FRIA. Thomas Peters is research associate of the Belgian Fund for Scientific Research (F.R.S.-FNRS). This work has been funded in parts by the F.R.S.–FNRS SeVote project.

## References

- Abe, M., Fuchsbauer, G., Groth, J., Haralambiev, K., Ohkubo, M.: Structurepreserving signatures and commitments to group elements. In: Advances in Cryptology – CRYPTO 2010. pp. 209–236. Springer (2010)
- Abe, M., Haralambiev, K., Ohkubo, M.: Signing on elements in bilinear groups for modular protocol design. Cryptology ePrint Archive, Report 2010/133 (2010), https://ia.cr/2010/133
- Adida, B., de Marneffe, O., Pereira, O., Quisquater, J.: Electing a university president using open-audit voting: Analysis of real-world use of helios. In: 2009 Electronic Voting Technology Workshop / Workshop on Trustworthy Elections, EVT/WOTE '09. USENIX Association (2009)
- Alwen, J., Ostrovsky, R., Zhou, H., Zikas, V.: Incoercible multi-party computation and universally composable receipt-free voting. In: Advances in Cryptology -CRYPTO 2015. LNCS, vol. 9216, pp. 763–780. Springer (2015)
- Bauer, B., Fuchsbauer, G.: Efficient signatures on randomizable ciphertexts. In: Security and Cryptography for Networks. pp. 359–381. Springer International Publishing (2020)
- Benaloh, J.: Simple verifiable elections. In: 2006 USENIX/ACCURATE Electronic Voting Technology Workshop, EVT'06. USENIX Association (2006)
- Benaloh, J.C., Tuinstra, D.: Receipt-free secret-ballot elections (extended abstract). In: Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing. pp. 544–553. ACM (1994)
- Bernhard, D., Cortier, V., Galindo, D., Pereira, O., Warinschi, B.: Sok: A comprehensive analysis of game-based ballot privacy definitions. In: 2015 IEEE Symposium on Security and Privacy. pp. 499–516 (2015)

- Bernhard, D., Pereira, O., Warinschi, B.: How not to prove yourself: Pitfalls of the fiat-shamir heuristic and applications to helios. In: Advances in Cryptology -ASIACRYPT 2012. LNCS, vol. 7658, pp. 626–643. Springer (2012)
- Blazy, O., Fuchsbauer, G., Pointcheval, D., Vergnaud, D.: Signatures on randomizable ciphertexts. In: Public Key Cryptography – PKC 2011. pp. 403– 422. Springer (2011)
- Boneh, D., Sahai, A., Waters, B.: Functional encryption: Definitions and challenges. In: Theory of Cryptography. pp. 253–273. Springer (2011)
- Canetti, R., Dwork, C., Naor, M., Ostrovsky, R.: Deniable encryption. In: Advances in Cryptology - CRYPTO '97. LNCS, vol. 1294, pp. 90–104. Springer (1997)
- Canetti, R., Gennaro, R.: Incoercible multiparty computation (extended abstract). In: 37th Annual Symposium on Foundations of Computer Science, FOCS '96. pp. 504–513. IEEE Computer Society (1996)
- Canetti, R., Krawczyk, H., Nielsen, J.B.: Relaxing chosen-ciphertext security. In: Advances in Cryptology - CRYPTO 2003. pp. 565–582. Springer (2003)
- Chaidos, P., Cortier, V., Fuchsbauer, G., Galindo, D.: Beleniosrf: A non-interactive receipt-free electronic voting scheme. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS). pp. 1614–1625 (2016)
- Chase, M., Kohlweiss, M., Lysyanskaya, A., Meiklejohn, S.: Malleable proof systems and applications. In: Advances in Cryptology – EUROCRYPT 2012. pp. 281–300. Springer (2012)
- Chaum, D.: Surevote: technical overview. In: Proceedings of the Workshop onTrustworthy Elections (WOTE 2001) (2001)
- Chevallier-Mames, B., Fouque, P., Pointcheval, D., Stern, J., Traoré, J.: On some incompatible properties of voting schemes. In: Towards Trustworthy Elections, New Directions in Electronic Voting. LNCS, vol. 6000, pp. 191–199. Springer (2010)
- Cortier, V., Galindo, D., Glondu, S., Izabachène, M.: Distributed elgamal à la pedersen: Application to helios. In: Proceedings of the 12th annual ACM Workshop on Privacy in the Electronic Society, WPES 2013. pp. 131–142. ACM (2013)
- Cramer, R., Gennaro, R., Schoenmakers, B.: A secure and optimally efficient multiauthority election scheme. In: Advances in Cryptology - EUROCRYPT '97. LNCS, vol. 1233, pp. 103–118. Springer (1997)
- 21. Devillez, H., Pereira, O., Peters, T.: Traceable receipt-free encryption. Cryptology ePrint Archive (2022)
- Gennaro, R.: Achieving independence efficiently and securely. In: Proceedings of the Fourteenth Annual ACM Symposium on Principles of Distributed Computing. pp. 130–136. ACM (1995)
- Gennaro, R., Jarecki, S., Krawczyk, H., Rabin, T.: Secure distributed key generation for discrete-log based cryptosystems. J. Cryptol. 20(1), 51–83 (2007)
- 24. Groth, J.: Efficient fully structure-preserving signatures for large messages. In: Advances in Cryptology – ASIACRYPT 2015. pp. 239–259. Springer (2015)
- Hirt, M., Sako, K.: Efficient receipt-free voting based on homomorphic encryption. In: Advances in Cryptology - EUROCRYPT 2000, International Conference on the Theory and Application of Cryptographic Techniques Proceeding. LNCS, vol. 1807, pp. 539–556. Springer (2000)
- Horwitz, J., Lynn, B.: Toward hierarchical identity-based encryption. In: Advances in Cryptology — EUROCRYPT 2002. pp. 466–481. Springer (2002)
- 27. Jutla, C.S., Roy, A.: Shorter quasi-adaptive nizk proofs for linear subspaces. In: Advances in Cryptology - ASIACRYPT 2013. pp. 1–20. Springer (2013)

- 30 Henri Devillez, Olivier Pereira, and Thomas Peters
- Kiayias, A., Zacharias, T., Zhang, B.: End-to-end verifiable elections in the standard model. In: Advances in Cryptology - EUROCRYPT 2015. LNCS, vol. 9057, pp. 468–498. Springer (2015)
- Kiltz, E.: Chosen-ciphertext security from tag-based encryption. In: Theory of Cryptography. pp. 581–600. Springer (2006)
- Kiltz, E., Wee, H.: Quasi-adaptive nizk for linear subspaces revisited. Cryptology ePrint Archive, Report 2015/216 (2015), https://ia.cr/2015/216
- Küsters, R., Truderung, T., Vogt, A.: A game-based definition of coercionresistance and its applications. In: Proceedings of the 23rd IEEE Computer Security Foundations Symposium, CSF 2010. pp. 122–136. IEEE Computer Society (2010)
- Libert, B., Peters, T., Joye, M., Yung, M.: Linearly homomorphic structurepreserving signatures and their applications. In: Advances in Cryptology - CRYPTO 2013. LNCS, vol. 8043, pp. 289–307. Springer (2013)
- Libert, B., Peters, T., Joye, M., Yung, M.: Non-malleability from malleability: Simulation-sound quasi-adaptive nizk proofs and CCA2-secure encryption from homomorphic signatures. In: Advances in Cryptology – EUROCRYPT 2014. pp. 514–532. Springer (2014)
- MacKenzie, P.D., Reiter, M.K., Yang, K.: Alternatives to non-malleability: Definitions, constructions, and applications (extended abstract). In: Theory of Cryptography, TCC. LNCS, vol. 2951, pp. 171–190. Springer (2004)
- Moran, T., Naor, M.: Receipt-Free Universally-Verifiable Voting with Everlasting Privacy. In: Advances in Cryptology - CRYPTO 2006. pp. 373–392. LNCS, Springer (2006)
- Prabhakaran, M., Rosulek, M.: Homomorphic encryption with cca security. In: Automata, Languages and Programming. pp. 667–678. Springer (2008)
- 37. Sahai, A., Waters, B.: Fuzzy identity-based encryption. In: Advances in Cryptology – EUROCRYPT 2005. pp. 457–473. Springer (2005)
- Sako, K., Kilian, J.: Receipt-free mix-type voting scheme A practical solution to the implementation of a voting booth. In: Advances in Cryptology - EUROCRYPT '95. LNCS, vol. 921, pp. 393–403. Springer (1995)
- Shamir, A.: Identity-based cryptosystems and signature schemes. In: Advances in Cryptology. pp. 47–53. Springer (1985)
- Unruh, D., Müller-Quade, J.: Universally composable incoercibility. In: Proceedings of the 30th Annual Conference on Advances in Cryptology. p. 411–428. CRYPTO'10, Springer-Verlag (2010)
- Wikström, D.: Simplified submission of inputs to protocols. In: Security and Cryptography for Networks, 6th International Conference, SCN 2008, Proceedings. LNCS, vol. 5229, pp. 293–308. Springer (2008)
- Wikström, D.: A commitment-consistent proof of a shuffle. In: Information Security and Privacy, 14th Australasian Conference, ACISP 2009 Proceedings. LNCS, vol. 5594, pp. 407–421. Springer (2009)