

Trapdoor Projective Sampling and Identity-Based Anonymous Broadcast

Gaspard Meunier and Duong Hieu Phan

Télécom Paris, Institut Polytechnique de Paris, Palaiseau, France
{gaspard.meunier, hieu.phan}@telecom-paris.fr

Abstract. In anonymous broadcast encryption, a ciphertext hides both the message and the set of recipients. In the general case, a lower bound of Kiayias and Samari forces the ciphertext to grow linearly with the number of recipients, a bound already met by the trivial per-recipient solution. The bounded universe is the setting where anonymous broadcast becomes concise, even asymptotically optimal and as efficient as the underlying public-key encryption. Far from a mere restriction, it is the building block for *anamorphic* communication against a dictator, where one must reach many (but boundedly many) recipients covertly, as put forth in the DPPY scheme at Eurocrypt '25. DPPY is, however, secret-key anamorphic encryption, where the broadcast layer itself must be hidden. Even a public-key variant does not suffice against a strong adversary, since a per-recipient public key lets a dictator ask for the matching secret key. This motivates hiding *identities* directly, with no recipient public key published, leading to identity-based anonymous broadcast in the bounded model, the focus of this paper.

Technically, projective sampling (Ling, Phan, Stehlé and Steinfeld, LPSS, CRYPTO '14), the tool underlying concise anonymous broadcast encryption, offers no key extraction: private keys cannot be produced from a projected key fixed in advance. We introduce *trapdoor projective sampling* that reverses the order of key generation in projective sampling: projected keys are fixed first as hashes of identities and compatible private keys are then preimage-sampled with a trapdoor. Adding a trapdoor is always a challenging objective, and we achieve it by showing that the LPSS projection space can be made far narrower - from $m \times (m - n)$ down to $m \times 2n$ - without breaking the k -LWE security argument, and it is exactly at this width that the projection matrix can be generated with a trapdoor. We formalise the notion of trapdoor projective sampling with new security requirements, instantiate it from k -LWE, and obtain an identity-based anonymous broadcast whose ciphertext is as in LPSS, DPPY schemes. We also extend it to an unbounded identity universe under static corruption, with target sets kept bounded and adaptive.

Table of Contents

Trapdoor Projective Sampling and Identity-Based Anonymous Broadcast	1
<i>Gaspard Meunier and Duong Hieu Phan</i>	
1 Introduction	3
1.1 Identity-Based Anonymous Cryptography	3
1.2 Toward a Fair Model for AI-Safety Evaluation	4
1.3 Theoretical Motivation	6
1.4 Related Work and an Attack on He et al.'s Scheme	6
1.5 Our Approach: Trapdoor Projective Sampling	7
1.6 Contributions	8
2 Technical Overview	9
3 Preliminaries	11
3.1 Probability	11
3.2 Euclidean Lattices and Discrete Gaussian Distributions	11
3.3 Random Lattices	13
3.4 Learning with Errors and k -LWE	14
3.5 Hardness of k -LWE	14
4 Projective Sampling	15
4.1 Projective Sampling Families	15
4.2 From Secret-Key Samplers to Public Samplers	15
4.3 Projective Sampling from k -LWE	16
5 Trapdoor Projective Sampling	19
5.1 Trapdoor Projective Sampling Families	19
5.2 Trapdoor Projective Sampling from k -LWE	20
6 From Public-Key to Identity-Based Broadcast	25
6.1 Broadcast in a Bounded Universe	26
6.2 From Projected Public Keys to Identities	27
7 Identity-Based Anonymous Broadcast for a Bounded Universe	29
7.1 The Public-Key Scheme	29
7.2 Removing One Recipient	30
7.3 Security of the Public-Key Scheme	33
7.4 The Identity-Based Scheme	35
8 Unbounded Identity Universe with Static Corruption	36
8.1 Model	36
8.2 The Scheme	37
8.3 Removing All Recipients Whose Keys Are Not Held	37
8.4 Security with Static Corruption	39

1 Introduction

1.1 Identity-Based Anonymous Cryptography

Identity-based encryption, introduced by Shamir [24], allows a sender to encrypt to any identity, removing the need for the certificates and the public-key infrastructure of ordinary public-key encryption. One had to wait until the year 2000 to see the first efficient identity-based encryption: pairing-based schemes from Sakai, Ohgishi and Kasahara [23] and Boneh and Franklin [6], a construction from quadratic residues by Cocks [9], and finally a post-quantum secure scheme from lattices by Gentry, Peikert and Vaikuntanathan [15]. A basic privacy guarantee one then asks for is *anonymity*: the ciphertext should hide the identity it is addressed to. Anonymity proved delicate in the pairing setting, as pairing-based schemes tend to leak the recipient's identity, and it took time to see the first anonymous (hierarchical) IBE constructions [1,7,14]. Lattices changed the picture: the Gaussian preimage-sampling trapdoor of Gentry, Peikert and Vaikuntanathan [15] yields identity-based encryption whose ciphertexts are pseudorandom, hence anonymous almost by construction [15,8,2], with *preimage sampling* as the central tool.

The identity-based setting for encryption, however, carries an inherent price. While it relieves senders from registering and certifying public keys, it requires a trusted authority holding a master secret that extracts *every* user's private key. This key-escrow authority is a strong assumption, and it is the main drawback of identity-based cryptography compared with plain public-key encryption.

From Encryption to Broadcast. For broadcast, the trade-off between the identity-based and public-key settings changes. A broadcast encryption scheme already depends on a central authority that issues decryption keys to its users, so turning it identity-based adds *no* new trust assumption: the authority that must exist anyway can just as well extract keys for identities. What limits broadcast encryption under anonymity is instead the ciphertext size. For an unbounded number of recipients, Kiayias and Samari [18] proved that the ciphertext of any anonymous broadcast scheme must grow linearly with the recipient set, so one cannot improve on the trivial solution that encrypts separately under each recipient's key: the per-recipient construction going back to private broadcast encryption [5,20]. The right question is therefore how far a public-key ciphertext can be stretched: for how many recipients can one broadcast anonymously while keeping the ciphertext size efficient, i.e., not growing linearly with the target set? The k -LWE assumption of Ling, Phan, Stehlé and Steinfeld [21] helps answer this question: for k recipients, one can do so very efficiently, at (asymptotically) the same size as the underlying public-key encryption. This is the *bounded* model - k is bounded by the lattice dimension - where the ciphertext is a single lattice vector, on par with a Dual-Regev ciphertext and independent of the target set and thus optimal. The bounded model has proved its worth as a building block: it underlies concise combinatorial trace-and-revoke schemes [21,12] and, most recently, anamorphic communication against a dictator [11].

Reinforcing Broadcast Anamorphism Anamorphic encryption, introduced by Persiano, Phan and Yung [22], enables covert communication under a *dictator* who can coerce users into surrendering their secret keys: a covert message is carried by a ciphertext indistinguishable from an ordinary one, and stays hidden even after all keys are revealed. Do, Persiano, Phan and Yung [11] lift this to a broadcast mode, where a single ciphertext reaches many recipients at once - exactly the regime of bounded anonymous broadcast.

Their solution, however, is *secret-key*: the anamorphic layer is a secret-key broadcast scheme whose very existence must be hidden from the dictator. Concretely,

- the *regular channel* is standard (Dual-Regev) public-key encryption;
- the *anamorphic channel* is a secret-key broadcast scheme whose ciphertexts are indistinguishable from regular ones;

- each anamorphic user holds a broadcast encryption secret-key together with a decryption key which is the user’s *double key*;
- anamorphism holds *only* because the broadcast system is concealed, so that all of its data are kept secret.

Keeping an entire broadcast system secret is a strong requirement. We therefore ask whether the anamorphic broadcast can be turned into a public-key scheme whose existence is made *public*, and in which scenarios such publicity does not pose a problem. The dictator would then see the authority’s public key and the users’ public keys. The premise of anamorphism is that any observable user’s public key can be coerced into yielding its secret key. We thus respect this premise, but when the broadcast system is made public, there is a new entity who comes on stage: the *authority* of the broadcast. We evidently need to suppose that this authority is incorruptible, otherwise the whole system collapses. This fits many practical scenarios where the authority is an external arbiter, outside the dictator’s influence and hence beyond coercion.

As we respect the premise that the dictator may still demand the anamorphic secret keys of the *recipients*, who live inside its territory, if the users’ public keys are placed in the parameters to enable encryption as in all public-key broadcast systems, a problem arises, since the dictator can then ask each user for the corresponding secret key. What must be protected is that no public key of any user in the broadcast system should be observable, so that the dictator cannot ask for the corresponding secret keys. Encryption should therefore be addressed to the identities of the recipients, while the identity itself stays concealed inside the ciphertext. This exactly raises the need for *identity-based* anamorphic broadcast, which can be directly designed by replacing the secret-key anonymous broadcast of [11] with an identity-based anonymous broadcast that preserves the ciphertext format of [11], itself that of Dual-Regev encryption [15]. We will thus focus in our work on designing such an identity-based anonymous broadcast. Before discussing the technique behind our construction, we make a detour to propose a framework arguing that the setting where an authority/arbiter (running anonymous identity-based encryption) is incorruptible can fit practical scenarios.

1.2 Toward a Fair Model for AI-Safety Evaluation

Artificial intelligence is advancing at a pace as impressive as it is open to dangerous misuse. AI can now solve problems once thought out of reach - with claimed breakthroughs on Millennium-Prize problems such as the Navier–Stokes equations - while security incidents around the ecosystem, such as the reported compromises of public model-sharing platforms like Hugging Face, are a reminder of how quickly control can slip.

Together with mounting concern about *misalignment*, this has made the danger of *uncontrolled* AI impossible to ignore. Strikingly, some of the largest developers are themselves among those calling for stronger control: Amodei’s proposal of *embedded evaluators*¹ would grant an independent team ongoing, employee-like access to a frontier AI company. The proposal was then supported by Sam Altman and Elon Musk. Turning such a proposal into a framework that delivers *genuine* oversight, however, is far from straightforward - which is the problem we discuss here.

Why external oversight is not enough. Overseeing a frontier AI corporation from the outside is inherently hard. An external evaluator rarely grasps the full trajectory of development of the whole system and the sensitive parts can be well obfuscated; and, being visibly an outsider whose identity is known, it might be highly pressured to stay permanently to observe all the activities. In practice, the people who truly understand the dangers are the corporation’s own employees. Today such an employee often has to resign in order to speak out - and leaves without evidence, as in the recent

¹ Amodei, 2026

case of J. Coxon². It would be far better to let employees themselves act as evaluators, provided they are fairly protected. This is also what was emphasised in *A Right to Warn about Advanced Artificial Intelligence*³, whose authors call upon advanced AI companies to commit to fundamental principles, in particular “that the company will facilitate a verifiably anonymous process for current and former employees to raise risk-related concerns to the company’s board, to regulators, and to an appropriate independent organization with relevant expertise.”

Protection without abuse. Such protection means hiding the evaluator’s identity - from the corporation, and *even from one another*, so that pressuring or compromising one evaluator does not expose the others. Yet the channel must not become a tool for abuse: an evaluator may communicate only with the arbiter or with fellow evaluators. Evaluators do need to exchange information among themselves, since each holds only one facet of the picture and recognising a violation may require combining facets; but even this cooperation should reveal no identity, to anyone.

The model. We abstract three parties bound by a *regulatory-compliance* regime that all of them accept:

- the *corporation*: the Big Tech company that owns and operates the frontier AI system, and controls its entire infrastructure and communication channels;
- the *arbiter*: an external regulator or auditing body, outside the corporation’s control, to which reports are addressed and which enrolls evaluators;
- the *evaluators*: employees who are also granted third-party evaluator status and are mandated to report safety violations.

Their incentives diverge. The arbiter wants faithful reports and has no stake in the corporation; we model it as *incorruptible*. Each evaluator is mandated, and inclined, to report truthfully, but fears retaliation. The corporation, compelled by the regime to accept the arbiter and to grant evaluator access to a specific communication channel, cannot openly refuse oversight; yet it might retain an incentive to suppress hugely damaging findings - by pressuring an evaluator to stay silent or to file a false report and, controlling the channel, in a severe scenario might also demand the disclosure of the corresponding keys if it detects public keys associated with the evaluator.

This is the *anamorphic* threat model of Persiano, Phan and Yung [22], transposed to a fixed arbiter: the corporation plays the dictator over its own employees, while the arbiter, mandated and external, is beyond its reach.

Scope of the cryptographic guarantee. Cryptography here guarantees that, *within the mandated channel*, a report is indistinguishable from ordinary traffic and stays hidden even after the evaluator surrenders all keys associated with any public keys the dictator has observed; it does *not* address control of the endpoints - the software or hardware an evaluator runs - which can be part of the regulatory-compliance regime itself, where the corporation does not reach inside to control the employees’ physical machines. Within this scope, the question is which primitive lets an evaluator report safely.

Why identity-based anonymous broadcast. A report must reach the arbiter and, often, a chosen subset of fellow evaluators (each sees only one facet, and a violation may surface only by combining facets), so the natural object is a broadcast encryption to the arbiter and selected colleagues that the evaluator trusts. Anonymity is what keeps such a message indistinguishable from routine traffic

² Scientific American, 2026

³ <https://righttowarn.ai/>, written by current and former employees of OpenAI, Google DeepMind and Anthropic (the current OpenAI employees signing anonymously), and endorsed by Yoshua Bengio, Geoffrey Hinton and Stuart Russell.

and hides the recipients even from one another. A public-key solution does not suffice: the sender would need a public key for every recipient, and publishing those keys would reveal the list of evaluators and hand the corporation coercion targets. Addressing recipients by *identity* removes this: the only public data is the master public parameters of the arbiter. Enrollment is gated by the arbiter, which extracts, for an evaluator’s own identity, the key that serves as its double key for decrypting anamorphic ciphertexts - so only enrolled parties take part, which avoids abuse as internal information cannot go outside the compliance regime. More interestingly, one may even address an identity before its key is extracted, so new evaluators can be enrolled and reached with no prior key exchange. Note that an evaluator does not need to know who the other evaluators are: he sends the anamorphic message to the set of colleagues that he trusts and hopes that some of them are also evaluators; indeed, if a trusted colleague is also an evaluator, he can decrypt, and otherwise he has no double key and the ciphertext looks exactly like a normal ciphertext of the routine traffic.

In short, we believe that a setting in which employees act as evaluators in a strict form - communicating only through the arbiter’s system, invisibly to the corporation - complements the embedded-evaluator proposal with a covert, coercion-resistant reporting channel.

1.3 Theoretical Motivation

Beyond this practical motivation, there is a purely theoretical question that we find compelling. Encrypting anonymously to a *single* recipient is by now well understood, notably through anonymous (identity-based) encryption, in the pairing [1,7,14] and in the lattice [15,8,2] settings. One-to-many communication is far harder: the lower bound of Kiayias and Samari [18] shows that, in general, an anonymous broadcast cannot improve on the trivial per-recipient solution. The natural question is therefore quantitative: starting from one recipient, to how many recipients can one extend while preserving anonymity and without substantially altering the structure of the encryption?

In the LWE world, k -LWE gives a clean partial answer. The DPY scheme [12] keeps the ciphertext a single Dual-Regev vector and supports up to $N = \Theta(n \log q)$ recipients: its projection matrix $\mathbf{H}$ has $m - n$ columns, so the projected keys live in a space of dimension $m - 2n$, which accommodates $\Theta(n \log q)$ affinely independent keys when $m = \Theta(n \log q)$. A more demanding question is what remains achievable once the scheme is made *identity-based*, so that no public key of any user is published and recipients are named only by identities. We give the first answer to this question: endowing $\mathbf{H}$ with a trapdoor forces its width down to $2n$, so the projected keys now live in dimension n , and our construction reaches N up to $(1 - \gamma)n = \Theta(n)$. Closing the resulting gap - between the $\Theta(n)$ recipients of the identity-based scheme and the $\Theta(n \log q)$ of the public-key scheme - is, to us, an appealing open problem. Carrying this line of work to other post-quantum assumptions would also be valuable: a code-based analogue of k -LWE, through the hardness of code-based problems with hints, has very recently been proposed and applied to traitor tracing [10], but it is still at an early stage, with a constant number of corrupted users, and is neither anonymous nor identity-based.

1.4 Related Work and an Attack on He et al.’s Scheme

Recipient privacy in broadcast was introduced by Barth, Boneh and Waters [5] as private (anonymous) broadcast encryption, with adaptively secure standard-model constructions by Libert, Paterson and Quaglia [20]; their ciphertexts, like the trivial per-recipient scheme, grow linearly with the number of recipients, which the lower bound [18] shows is unavoidable in general. A line of pairing-based *identity-based anonymous* broadcast followed [19,17], but these schemes are not post-quantum. The transition to post-quantum took time, as from the first pairing-based IBE [23,6] to the first lattice-based IBE in 2008 [15]. On post-quantum security, to our knowledge, the only prior

lattice scheme that is simultaneously anonymous and identity-based is that of He, Liu, Liu and Liu [16], which was shown to be secure only in the selective model and against outsider adversaries. Unfortunately, even under this weak model, we find that this scheme is insecure for the proposed instantiation, and we present an attack below.

The scheme of He et al. Using the Chinese remainder theorem and lattice basis delegation in a fixed dimension, their scheme packs all recipients into a single ciphertext $C_2 \in \mathbb{Z}_Q^m$ reduced modulo $Q = \prod_{ID \in S} H_1(ID)$, where H_1 maps identities to n -bit primes; the proof relies on C_2 being statistically close to uniform over $\mathbb{Z}_Q^m$, so its support is the box $[0, Q)^m$, with Q determined by the target set. The scheme targets only *selective outsider* anonymity: the adversary commits to two equal-size sets S_0, S_1 before seeing the public parameters, may corrupt only identities outside $S_0 \cup S_1$, and must distinguish an encryption to S_0 from one to S_1 . This is weaker than the standard notion we adopt, in which corruptions are adaptive and any identity outside the symmetric difference $S_0 \Delta S_1$ may be corrupted - in particular those in $S_0 \cap S_1$. We show that their scheme fails even this weaker notion.

The attack. The division-intractable hashing of Gennaro, Halevi and Rabin [13] that they invoke sets $q_{ID} = 2^t h(ID) + r$, where h is a collision-resistant hash, t is a short offset length, and $r < 2^t$ is the least offset making the value prime. The high-order bits of q_{ID} are thus the hash $h(ID)$, so the n -bit prime lies in $[2^{n-1}, 2^n)$ and is, up to the low-order offset, uniform there. The adversary chooses two random identities ID_0, ID_1 , and sets $S_0 = \{ID_0\}, S_1 = \{ID_1\}$. This choice is made at the commit step, so the attack respects the selective game. Once the public parameters fix H_1 , it computes $q_0 = H_1(ID_0), q_1 = H_1(ID_1)$ and lets $q_{\min} = \min(q_0, q_1)$, with b^* the index of the smaller prime. On the challenge C_2 it outputs b^* if every coordinate of C_2 is below $q_{\min}$, and $1 - b^*$ otherwise. If $\beta = b^*$ then C_2 is uniform over $\mathbb{Z}_{q_{\min}}^m$, so all coordinates are $< q_{\min}$ and the guess is correct; otherwise C_2 is uniform over $\mathbb{Z}_{q_{\max}}^m$ and the guess errs only with probability $(q_{\min}/q_{\max})^m$, so the advantage is $\frac{1}{2}(1 - (q_{\min}/q_{\max})^m)$. As ID_0, ID_1 are random, q_0, q_1 behave like independent uniform n -bit primes and $q_{\min}/q_{\max}$ is randomly spread over $[\frac{1}{2}, 1)$; hence the advantage is non-negligible.

The flaw. In their proof, at step (v) of the proof of Theorem 1, it states that “if y_0 is a uniformly random matrix, then the challenge ciphertext is also uniformly random, regardless of the choice of β ”. But $C_2 = 2y_0 + K \bmod Q_\beta$, so y_0 uniform makes C_2 uniform only over $\mathbb{Z}_{Q_\beta}^m$, whose support $[0, Q_\beta)^m$ depends on β : uniform over $\mathbb{Z}_{Q_0}^m$ and over $\mathbb{Z}_{Q_1}^m$ could be distinguishable as shown above. The equal-size restriction $|S_0| = |S_1|$, imposed “to avoid trivial attacks”, removes only the dependence on the *number* of recipients, not on the *product* Q of their primes.

Possible fix? To fix the attack, one should design a new division-intractable hash function rather than directly use the one from Gennaro, Halevi and Rabin [13]. Verifying many conditions - such as collision-resistance and, for any adversarially chosen sets $S_0 = \{ID_0\}, S_1 = \{ID_1\}$, that the ratio $(q_{\min}/q_{\max})^m$ is overwhelmingly close to 1 - is particularly challenging. The difficulty is that these conditions go in opposite directions. Collision-resistance over a large identity space forces the hash to output super-polynomially many *distinct* primes; yet keeping $(q_{\min}/q_{\max})^m$ close to 1 forces all these primes into a very narrow window, of multiplicative width $1 + o(1/m)$.

1.5 Our Approach: Trapdoor Projective Sampling

The tool underlying concise anonymous broadcast [12] is *projective sampling*. This primitive, introduced by Ling, Phan, Stehlé and Steinfeld (LPSS) [21], is a way to replace sampling procedures indexed by secret keys with publicly samplable projected procedures. This viewpoint was used in LPSS to build traitor-tracing schemes, and later by Do, Phan and Yung (DPY20) [12] to obtain concise LWE-based anonymous broadcast encryption. In both cases, the useful feature is that a

sampler tied to a hidden key can be simulated from a public projected key without revealing the hidden key itself.

For the identity-based setting, one needs to proceed in the reverse direction. A projected key should be determined first, from an identity, and only afterwards should the authority generate a compatible private key. Ordinary projective sampling does not provide such an extraction mechanism.

We address this by introducing *trapdoor projective sampling*. Informally, this is a sampling primitive with a trapdoor sampler: given a projected key, the sampler outputs a secret key whose projection is the prescribed value and whose distribution matches the conditional secret-key distribution required by the projective-sampling argument. This makes projective sampling compatible with identity-based key extraction.

We believe that trapdoor projective sampling would be a very useful tool to implement the identity-based setting in multi-user primitives. In this work, we focus on how to use it to obtain identity-based anonymous broadcast encryption. We expect that one can replace projective sampling with it in the traitor-tracing scheme of [21] to obtain an identity-based traitor-tracing scheme, and we leave a full treatment as an open problem.

1.6 Contributions

- *Trapdoor projective sampling.* We introduce trapdoor projective sampling, a generalization of projective sampling in which projected keys can be sampled first, and secret keys then sampled from them with a trapdoor, keeping the same key distribution as in the usual order, where secret keys come first. It makes projective sampling, so far a public-key tool, usable in the identity-based setting.
- *Instantiation from k -LWE.* We show that LPSS projective sampling holds for any projection dimension $n < d \leq m - n$, and that for $d = 2n$ the projection matrix can be generated with a trapdoor. The two sampling orders are statistically close, so the security arguments of LPSS carry over, and the projected keys of users whose keys are not given can now be simulated statistically, whereas LPSS simulate them computationally.
- *Public-key anonymous broadcast.* We revisit the anonymous broadcast scheme of DPY20 [12] with a projection of width $2n$, and prove its IND-CPA and ANO-CPA security with adaptive corruptions for $N \leq (1 - \gamma)n$ users, for a constant $\gamma \in (0, 1)$. The proof relies on a new transition that removes a recipient even when the adversary holds all the other keys; it addresses a slight gap in the analysis of DPY20 and applies to their scheme as well.
- *A generic compiler.* We give a compiler that turns any projective public-key broadcast scheme and trapdoor projective sampling family into an identity-based broadcast scheme in the random oracle model, with the same ciphertexts and the same security, including adaptive corruptions.
- *Identity-based anonymous broadcast.* Applied to our trapdoor projective sampling family and public-key scheme, the compiler gives, to our knowledge, the first secure lattice-based identity-based anonymous broadcast scheme, whose ciphertext is a single element of $\mathbb{Z}_q^{m+1}$, in exactly the format of Dual-Regev encryption and of [11], and whose public key has constant size and reveals neither the user set nor its size.
- *Unbounded universe.* With static corruption, the same scheme serves an unbounded number of identities with the same ciphertexts: only the corrupted identities and the recipients of the challenge are bounded by N . Any polynomial number of identities can be hashed as soon as $N = o(n)$, and larger N , up to $(1 - \gamma)n$, can be reached when the number of hashed identities is restricted, to about $q^{\gamma/(1-\gamma)}$.
- *Attack on the scheme of He et al.* We show that the scheme of [16], to our knowledge the only prior lattice scheme that is both anonymous and identity-based, is not anonymous, even in its own selective outsider model: the support of a ciphertext depends on the primes attached to

the recipients, which yields a distinguisher with non-negligible advantage. We explain the flaw in their proof and why a fix would be challenging (Section 1.4).

2 Technical Overview

We explain at a high level our new tool, trapdoor projective sampling, which reverses the order of key generation in projective sampling. We then explain how it is used for the design of identity-based anonymous broadcast encryption. For this, we first recall the notion of projective sampling and how it is used for anonymous broadcast.

Projective sampling. The starting point is the projective-sampling construction of LPSS [21]. Given $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ and $\mathbf{u} \in \mathbb{Z}_q^n$, a private key is a short vector $\mathbf{x}$ such that $\mathbf{A}^t \mathbf{x} = -\mathbf{u} \bmod q$. Writing $\mathbf{A}^+ = \begin{pmatrix} \mathbf{u}^t \\ \mathbf{A} \end{pmatrix}$ and $\mathbf{x}^+ = \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix}$, this means that $\mathbf{x}^+$ is orthogonal to $\text{Im}(\mathbf{A}^+)$. Given the keys $\mathbf{x}_1, \dots, \mathbf{x}_k$ of k users, the k -LWE assumption states that LWE samples, that is, uniform elements of $\text{Im}(\mathbf{A}^+)$ plus a Gaussian noise, are indistinguishable from the secret-key sampler

$$\frac{1}{q} U\left(\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_k^+)^\perp\right) + \nu_\alpha^{m+1},$$

which outputs uniform vectors orthogonal to all the keys, plus the same noise.

LPSS define projected keys by choosing a matrix $\mathbf{H}$ such that $\text{Im}(\mathbf{A}) \subseteq \text{Im}(\mathbf{H})$ and setting $\mathbf{h}_j = -\mathbf{H}^t \mathbf{x}_j$. Their projective-sampling theorem shows that, even given the keys, the secret-key sampler can be replaced by the distribution

$$\frac{1}{q} U\left(\bigcap_{j \leq k} \text{Im}(\mathbf{H}_j)\right) + \nu_\alpha^{m+1}, \quad \text{where } \mathbf{H}_j = \begin{pmatrix} \mathbf{h}_j^t \\ \mathbf{H} \end{pmatrix},$$

which is publicly samplable, since it only depends on $\mathbf{H}$ and on the projected keys. This replacement rests on the chain of inclusions

$$\text{Im}(\mathbf{A}^+) \subseteq \bigcap_{j \leq k} \text{Im}(\mathbf{H}_j) \subseteq \text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_k^+)^\perp.$$

Adding a uniform element of the middle space to both distributions of the k -LWE assumption turns the first one into the public sampler and leaves the second one unchanged. The three distributions are therefore indistinguishable, even given the keys.

Anonymous broadcast from projective sampling. Do, Phan and Yung [12] encrypt with this public sampler. The public key contains the projected keys $\mathbf{h}_i$ of all users, and a ciphertext for a target set $\mathcal{T}$ is a uniform element of $\bigcap_{i \in \mathcal{T}} \text{Im}(\mathbf{H}_i)$, plus a small noise and an encoding of the message in the first coordinate. A recipient i decrypts with $\mathbf{x}_i^+$: since $\mathbf{H}_i^t \mathbf{x}_i^+ = \mathbf{h}_i + \mathbf{H}^t \mathbf{x}_i = \mathbf{0}$, the key $\mathbf{x}_i^+$ is orthogonal to $\text{Im}(\mathbf{H}_i)$, and only the noise and the message remain. The ciphertext is a single vector of $\mathbb{Z}_q^{m+1}$, whatever the size of $\mathcal{T}$. We keep this encryption unchanged, and make the key generation identity-based.

Trapdoor projective sampling. We reverse the order of key generation: instead of sampling a short $\mathbf{x}$ such that $\mathbf{A}^t \mathbf{x} = -\mathbf{u}$ and then setting $\mathbf{h} = -\mathbf{H}^t \mathbf{x}$, we first fix $\mathbf{h}$, computed from the identity, and then sample a short $\mathbf{x}$ such that $\mathbf{H}^t \mathbf{x} = -\mathbf{h}$. Our construction (Section 5) proceeds in three steps.

Step 1: A projection matrix with a trapdoor. In LPSS, $\mathbf{H}$ has $m - n$ columns, which is too wide to carry a trapdoor. We observe that the LPSS projective-sampling theorem holds for any number of columns d with $n < d \leq m - n$ (Theorem 4.2), since its proof only relies on the chain of inclusions above. We take $d = 2n$: this preserves projective sampling, and allows $\mathbf{H}$ to be generated together with a trapdoor using the algorithm of Alwen and Peikert [4]. The matrix $\mathbf{A}$ is then generated from $\mathbf{H}$: we sample $\mathbf{R} \in \mathbb{Z}_q^{d \times n}$ and set $\mathbf{A} = \mathbf{H}\mathbf{R}$, so that $\text{Im}(\mathbf{A}) \subseteq \text{Im}(\mathbf{H})$ holds by construction.

Step 2: Compatible projected keys. For identities, the natural choice would be $\mathbf{h}_{id} = \text{Hash}(id)$. However, the sampled vector $\mathbf{x}$ must also be a valid secret key, that is, satisfy $\mathbf{A}^t \mathbf{x} = -\mathbf{u}$. If $\mathbf{H}^t \mathbf{x} = -\mathbf{h}$, then

$$\mathbf{A}^t \mathbf{x} = \mathbf{R}^t \mathbf{H}^t \mathbf{x} = -\mathbf{R}^t \mathbf{h} \pmod{q},$$

so $\mathbf{x}$ is a valid secret key if and only if $\mathbf{R}^t \mathbf{h} = \mathbf{u}$, which a hash value satisfies only with negligible probability. We therefore sample $\mathbf{h}_0 \in \mathbb{Z}_q^d$ uniformly and set $\mathbf{u} = \mathbf{R}^t \mathbf{h}_0$. Admissible projected keys then form the affine subspace $\mathcal{H}_{\mathbf{u}} = \mathbf{h}_0 + \ker(\mathbf{R}^t)$, of dimension $d - n$. If $\mathbf{K}$ is a basis of $\ker(\mathbf{R}^t)$, every $\mathbf{h} = \mathbf{h}_0 + \mathbf{K}\mathbf{a}$ lies in $\mathcal{H}_{\mathbf{u}}$. For identities, we take $\mathbf{h}_{id} = \mathbf{h}_0 + \mathbf{K} \text{Hash}(id)$.

Step 3: Key distributions. Identity-based key derivation calls for projected keys that are uniform on $\mathcal{H}_{\mathbf{u}}$: this is the distribution of $\mathbf{h}_{id} = \mathbf{h}_0 + \mathbf{K} \text{Hash}(id)$ in the random oracle model. It is also, up to a negligible distance, the distribution of the projected keys in LPSS. Indeed, $\mathbf{h} = -\mathbf{H}^t \mathbf{x}$ only depends on $\mathbf{x}$ modulo the lattice $\Lambda = \{\mathbf{x} \in \mathbb{Z}^m : \mathbf{H}^t \mathbf{x} = \mathbf{0} \pmod{q}\}$, and a Gaussian whose width exceeds the smoothing parameter of Λ is close to uniform modulo Λ . Moreover, given $\mathbf{h}$, the LPSS secret key follows the Gaussian distribution over the vectors $\mathbf{x}$ such that $\mathbf{H}^t \mathbf{x} = -\mathbf{h}$, which is exactly what the trapdoor samples. Since the width required by trapdoor sampling already exceeds this smoothing parameter, the reversed order yields a distribution statistically close to that of LPSS (Lemma 5.4), and the LPSS theorem carries over (Theorem 5.3).

Security in a bounded universe. We first consider N users fixed in advance, with adaptive corruptions (Section 7). As in [12], a hybrid argument removes the recipients from the target set one at a time, until the ciphertext is uniform. Since corruptions are adaptive, removing a recipient j must go unnoticed even by an adversary holding the keys of all the other users (Proposition 7.2). This addresses a slight gap in the proof of [12, Theorem 6], whose transition [12, Lemma 5] is only stated for an adversary holding the keys of the other recipients. The key point is that the held keys cannot tell the ciphertext spaces of $\mathcal{T}$ and $\mathcal{T} \setminus \{j\}$ apart: a combination of held keys is orthogonal to $\bigcap_{i \in \mathcal{T}} \text{Im}(\mathbf{H}_i)$ if and only if it is orthogonal to $\bigcap_{i \in \mathcal{T} \setminus \{j\}} \text{Im}(\mathbf{H}_i)$ (Lemma 7.4). Geometrically, an affine combination of keys is orthogonal to the ciphertext space of $\mathcal{T}$ exactly when the same combination of their projected keys lies in the affine hull of the projected keys of $\mathcal{T}$. When the projected keys of the N users are affinely independent, a combination of projected keys other than that of j lies in the affine hull of those of $\mathcal{T}$ only if it lies in that of $\mathcal{T} \setminus \{j\}$, which gives the claim. Since the projected keys are uniform in $\mathcal{H}_{\mathbf{u}}$, of dimension n , they are affinely independent with overwhelming probability when $N \leq (1 - \gamma)n$, for a constant $\gamma \in (0, 1)$.

Unbounded universe. In an unbounded universe (Section 8), identities are arbitrary and their number is not fixed in advance: each identity is hashed to a projected key in $\mathcal{H}_{\mathbf{u}}$, and the authority extracts its key on demand. The bound N then only limits the identities involved in the security game: at most N identities are corrupted or receive the challenge. A reduction must embed its k -LWE hints in the corrupted identities before answering the random-oracle queries on them. We therefore consider static corruption: the adversary commits to the identities it corrupts, while target sets remain adaptive. A geometric argument similar to the bounded case then applies, with the keys of the corrupted identities as the held keys. However, the adversary may hash many more identities before choosing its target set, so affine independence is needed for every set of at most N queried

identities. A union bound over these sets restricts the number of hashed identities when N is close to n (Lemma 8.7 and Remark 8.9).

3 Preliminaries

For a positive integer n , $[n]$ denotes $\{1, \dots, n\}$. If x is a real number, then $\lfloor x \rfloor$ is the closest integer to x (with any deterministic rule in case x is half an odd integer). The base-2 logarithm is denoted $\log$. By convention, vectors are assumed to be in column form and are written using bold lower-case letters, e.g. $\mathbf{x}$. The i th component of $\mathbf{x}$ will be denoted by $\mathbf{x}_i$. Matrices are written as bold capital letters, e.g. $\mathbf{X}$, and the i th column vector of a matrix $\mathbf{X}$ is denoted $\mathbf{x}_i$. We let $\langle \cdot, \cdot \rangle$ denote the canonical inner product. For q prime, we let $\mathbb{Z}_q$ denote the field of integers modulo q . For two matrices $\mathbf{A}, \mathbf{B}$ of compatible dimensions, we let $(\mathbf{A}|\mathbf{B})$ and $(\mathbf{A}||\mathbf{B})$ respectively denote the horizontal and vertical concatenation of $\mathbf{A}$ and $\mathbf{B}$. For $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ we define $\text{Im}(\mathbf{A}) = \{\mathbf{A}\mathbf{s} : \mathbf{s} \in \mathbb{Z}_q^n\} \subseteq \mathbb{Z}_q^m$. For $X \subseteq \mathbb{Z}_q^m$, we let $\text{Span}(X)$ denote the set of linear combinations of the elements of X . We let $X^\perp$ denote the linear subspace $\{\mathbf{b} \in \mathbb{Z}_q^m : \forall \mathbf{c} \in X, \langle \mathbf{b}, \mathbf{c} \rangle = 0\}$. For a matrix $\mathbf{S} \in \mathbb{R}^{m \times m}$, we let $\|\mathbf{S}\|$ denote the norm of its longest column. If $\mathbf{S}$ is full column-rank, we let $\sigma_1(\mathbf{S}) \geq \dots \geq \sigma_m(\mathbf{S})$ denote its singular values, and write

$$s_{\max}(\mathbf{S}) = \sigma_1(\mathbf{S}) \quad \text{and} \quad s_{\min}(\mathbf{S}) = \sigma_m(\mathbf{S}).$$

When $\mathbf{S}$ is clear from context, we simply write $s_{\max}$ and $s_{\min}$. We let $\mathbb{T}$ denote the additive group $\mathbb{R}/\mathbb{Z}$.

3.1 Probability

If D_1 and D_2 are two distributions over a countable set X , their statistical distance $\frac{1}{2} \sum_{x \in X} |D_1(x) - D_2(x)|$ will be denoted by $\Delta(D_1, D_2)$. The statistical distance is defined similarly if X is measurable. If X is finite, we let $U(X)$ denote the uniform distribution over X . We say that a distribution D (or a random variable having distribution D) is ε -uniform if its statistical distance from the uniform distribution over X is ε . For any invertible $\mathbf{S} \in \mathbb{R}^{m \times m}$ and $\mathbf{c} \in \mathbb{R}^m$, we define the function $\rho_{\mathbf{S}, \mathbf{c}}(\mathbf{b}) = \exp(-\pi \|\mathbf{S}^{-1}(\mathbf{b} - \mathbf{c})\|^2)$. For $\mathbf{S} = \sigma \mathbf{I}_m$, we write $\rho_{\sigma, \mathbf{c}}$, and we omit the center when $\mathbf{c} = \mathbf{0}$. We let ν_α denote the one-dimensional Gaussian distribution with standard deviation α .

3.2 Euclidean Lattices and Discrete Gaussian Distributions

Let $\mathbf{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\} \subset \mathbb{R}^m$ consist of n linearly independent vectors. The n -dimensional lattice Λ generated by the basis $\mathbf{B}$ is

$$\Lambda = \mathcal{L}(\mathbf{B}) = \{\mathbf{B}\mathbf{c} = \sum_{i \in [n]} c_i \cdot \mathbf{b}_i : \mathbf{c} \in \mathbb{Z}^n\}.$$

The *dual lattice* of Λ , denoted Λ^* , is defined to be $\Lambda^* = \{\mathbf{x} \in \text{span}_{\mathbb{R}}(\Lambda) : \forall \mathbf{v} \in \Lambda, \langle \mathbf{x}, \mathbf{v} \rangle \in \mathbb{Z}\}$. For $p \in \{2, \infty\}$, we let $\lambda_1^p(\Lambda)$ denote the length of the shortest non-zero vector of Λ for the ℓ_p norm. When $p = 2$, we simply write $\lambda_1(\Lambda)$. For a basis $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$, we write $\tilde{\mathbf{B}} = (\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_n)$ for its Gram-Schmidt orthogonalization, and

$$\|\tilde{\mathbf{B}}\| = \max_{i \in [n]} \|\tilde{\mathbf{b}}_i\|.$$

We extensively use q -ary lattices. The q -ary lattice associated to $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ is defined as

$$\Lambda^\perp(\mathbf{A}) = \{\mathbf{x} \in \mathbb{Z}^m : \mathbf{A}^t \cdot \mathbf{x} = \mathbf{0} \pmod{q}\}$$

and for $\mathbf{u} \in \mathbb{Z}_q^n$ we define

$$\Lambda_{\mathbf{u}}^\perp(\mathbf{A}) = \{\mathbf{x} \in \mathbb{Z}^m : \mathbf{A}^t \cdot \mathbf{x} = \mathbf{u} \pmod{q}\}.$$

For any $\mathbf{c} \in \mathbb{R}^m$, $\mathbf{S} \in \mathbb{R}^{m \times m}$ and n -dimensional lattice Λ , define the *discrete Gaussian distribution* over Λ as:

$$\forall \mathbf{x} \in \Lambda, \quad D_{\Lambda, \mathbf{S}, \mathbf{c}}(\mathbf{x}) = \frac{\rho_{\mathbf{S}, \mathbf{c}}(\mathbf{x})}{\rho_{\mathbf{S}, \mathbf{c}}(\Lambda)}.$$

Following LPSS, we will use the change of variables

$$D_{\Lambda, \mathbf{S}, \mathbf{c}} = \mathbf{S}^t D_{\mathbf{S}^{-t}\Lambda, \mathbf{1}, \mathbf{S}^{-t}\mathbf{c}},$$

which allows one to adapt lemmas stated in the spherical case $\mathbf{S} = \mathbf{I}_m$. We also recall the definition of the *smoothing parameter*. For any n -dimensional lattice Λ and positive real $\varepsilon > 0$, the smoothing parameter $\eta_\varepsilon(\Lambda)$ is the smallest real $s > 0$ such that $\rho_{1/s}(\Lambda^* \setminus \{\mathbf{0}\}) \leq \varepsilon$.

Lemma 3.1 (Basic smoothing bounds, [15, Lemma 2.6]). *Let Λ be a full-rank lattice in $\mathbb{R}^m$ and let $\varepsilon \in (0, 1)$. Then*

$$\eta_\varepsilon(\Lambda) \leq \frac{\sqrt{\ln(2m(1 + 1/\varepsilon))/\pi}}{\lambda_1^\infty(\Lambda^*)}.$$

Moreover, if $\Lambda' \subseteq \Lambda$ are full-rank lattices, then

$$\eta_\varepsilon(\Lambda) \leq \eta_\varepsilon(\Lambda').$$

In particular,

$$\eta_\varepsilon(q\mathbb{Z}^m) \leq q\sqrt{\ln(2m(1 + 1/\varepsilon))/\pi}.$$

Lemma 3.2 (Random q -ary smoothing, [21, Lemma 9]). *Let q be prime, let $m \geq 2d$, let $\varepsilon \in (0, 1)$, and let $\mathbf{H} \leftarrow U(\mathbb{Z}_q^{m \times d})$. Except for a fraction $2^{-\Omega(d)}$ of matrices $\mathbf{H} \in \mathbb{Z}_q^{m \times d}$,*

$$\eta_\varepsilon(\Lambda^\perp(\mathbf{H})) \leq 4q^{d/m} \sqrt{\ln(2m(1 + 1/\varepsilon))/\pi}.$$

Lemma 3.3 (Rank failure over $\mathbb{Z}_q$). *Let q be prime and let $r \geq s$. For $\mathbf{M} \leftarrow U(\mathbb{Z}_q^{r \times s})$,*

$$\Pr[\text{rank}(\mathbf{M}) < s] \leq \delta_{r,s} := \frac{q^{-(r-s+1)}}{1 - q^{-1}}.$$

Proof. Expose the columns successively. The probability that column j lies in the span of the preceding ones is at most q^{j-1-r} . Summing over $1 \leq j \leq s$ gives the claim. $\square$

Lemma 3.4 (Gaussian norm bound, adapted from [21, Lemma 2] and [3, Lemma 3]). *Let $\Lambda + \mathbf{u} \subset \mathbb{R}^m$ be an affine coset of an n -dimensional lattice Λ , let $\mathbf{c} \in \text{Span}(\Lambda + \mathbf{u})$, and let $\mathbf{S} \in \mathbb{R}^{m \times m}$ be invertible. If $s_{\min}(\mathbf{S}) \geq \eta_\varepsilon(\Lambda)$ and $\varepsilon \in (0, 1/2)$, then*

$$\Pr_{\mathbf{x} \leftarrow D_{\Lambda + \mathbf{u}, \mathbf{S}, \mathbf{c}}} [\|\mathbf{x} - \mathbf{c}\| > s_{\max}(\mathbf{S})\sqrt{n}] \leq 2^{-n+2}.$$

Lemma 3.5 (Theorem 4.1 of [15]). *There is a probabilistic polynomial-time algorithm that, given a basis $\mathbf{B}$ of an n -dimensional lattice $\Lambda = \mathcal{L}(\mathbf{B})$, a parameter $s \geq \|\tilde{\mathbf{B}}\| \cdot \omega(\sqrt{\log n})$, and a center $\mathbf{c} \in \text{Span}(\Lambda)$, outputs a sample from a distribution that is statistically close to $D_{\Lambda, \mathbf{S}, \mathbf{c}}$.*

Corollary 3.6. *Let $\Lambda = \mathcal{L}(\mathbf{B})$ be an n -dimensional lattice and let $\mathbf{S}$ be a positive diagonal matrix. If*

$$s_{\min}(\mathbf{S}) \geq \|\tilde{\mathbf{B}}\| \cdot \omega(\sqrt{\log n}),$$

then there is a probabilistic polynomial-time algorithm that, given $\mathbf{B}$, $\mathbf{S}$, and a center $\mathbf{c}$, outputs a sample from a distribution that is statistically close to $D_{\Lambda, \mathbf{S}, \mathbf{c}}$.

Lemma 3.7 (Coset sampling from a q -ary trapdoor). *Let q be prime, let $\mathbf{H} \in \mathbb{Z}_q^{m \times d}$ have full column rank, and let $\mathbf{T}_{\mathbf{H}}$ be a basis of $\Lambda^\perp(\mathbf{H})$. Let $\mathbf{S}$ be a positive diagonal matrix. If*

$$s_{\min}(\mathbf{S}) \geq \|\tilde{\mathbf{T}}_{\mathbf{H}}\| \cdot \omega(\sqrt{\log m}),$$

then there is a probabilistic polynomial-time algorithm that, given $\mathbf{H}$, $\mathbf{T}_{\mathbf{H}}$, $\mathbf{S}$ and $\mathbf{h} \in \mathbb{Z}_q^d$, outputs a sample statistically close to

$$D_{\Lambda_{-\mathbf{h}}^\perp(\mathbf{H}), \mathbf{S}}.$$

Proof. Since $\mathbf{H}$ has full column rank, $\mathbf{H}^t$ is surjective over $\mathbb{Z}_q$. Hence, by linear algebra, one can find $\mathbf{x}_0 \in \mathbb{Z}^m$ such that

$$\mathbf{H}^t \mathbf{x}_0 = -\mathbf{h} \pmod{q}.$$

Then

$$\Lambda_{-\mathbf{h}}^\perp(\mathbf{H}) = \mathbf{x}_0 + \Lambda^\perp(\mathbf{H}).$$

Using Corollary 3.6 with basis $\mathbf{T}_{\mathbf{H}}$ and center $-\mathbf{x}_0$, sample

$$\mathbf{y} \leftarrow D_{\Lambda^\perp(\mathbf{H}), \mathbf{S}, -\mathbf{x}_0}$$

up to negligible statistical distance, and output $\mathbf{x} = \mathbf{x}_0 + \mathbf{y}$. For every $\mathbf{x} = \mathbf{x}_0 + \mathbf{y} \in \Lambda_{-\mathbf{h}}^\perp(\mathbf{H})$, the weight of $\mathbf{y}$ in $D_{\Lambda^\perp(\mathbf{H}), \mathbf{S}, -\mathbf{x}_0}$ is proportional to

$$\rho_{\mathbf{S}, -\mathbf{x}_0}(\mathbf{y}) = \rho_{\mathbf{S}}(\mathbf{x}_0 + \mathbf{y}) = \rho_{\mathbf{S}}(\mathbf{x}).$$

Thus the output distribution is statistically close to $D_{\Lambda_{-\mathbf{h}}^\perp(\mathbf{H}), \mathbf{S}}$. $\square$

Lemma 3.8 (Corollary 2.8 of [15]). *Let Λ, Λ' be n -dimensional lattices, with $\Lambda' \subseteq \Lambda$. Then for any $\varepsilon \in (0, \frac{1}{2})$, any $s \geq \eta_\varepsilon(\Lambda')$ and any $\mathbf{c} \in \text{Span}(\Lambda)$, the distribution of $D_{\Lambda, s, \mathbf{c}} \pmod{\Lambda'}$ is within statistical distance at most 2ε of uniform over $\Lambda \pmod{\Lambda'}$.*

Corollary 3.9. *Let Λ, Λ' be n -dimensional lattices with $\Lambda' \subseteq \Lambda$, let $\varepsilon \in (0, 1/2)$, and let $\mathbf{S}$ be a positive diagonal matrix. If*

$$s_{\min}(\mathbf{S}) \geq \eta_\varepsilon(\Lambda'),$$

then for any center $\mathbf{c}$, the distribution of $D_{\Lambda, \mathbf{S}, \mathbf{c}} \pmod{\Lambda'}$ is within statistical distance at most 2ε of uniform over $\Lambda \pmod{\Lambda'}$.

3.3 Random Lattices

Lemma 3.10 (Adapted from [4, Theorem 3.1], as in [21, Lemma 10]). *There exists a probabilistic polynomial-time algorithm that, on input positive integers $n, m, q \geq 2$, outputs $(\mathbf{A}, \mathbf{T}) \in \mathbb{Z}_q^{m \times n} \times \mathbb{Z}^{m \times m}$ such that:*

- $\mathbf{A}$ is within statistical distance $2^{-\Omega(n)}$ from $U(\mathbb{Z}_q^{m \times n})$;
- $\mathbf{T}$ is a basis of $\Lambda^\perp(\mathbf{A})$, and
- $\|\mathbf{T}\| \leq 3mq^{n/m}$.

3.4 Learning with Errors and k -LWE

Let $\mathbf{s} \in \mathbb{Z}_q^n$ and $\alpha > 0$. We define the distribution $A_{\mathbf{s}, \alpha}$ as follows: take $\mathbf{a} \leftarrow U(\mathbb{Z}_q^n)$ and $e \leftarrow \nu_\alpha$, and return

$$\left(\mathbf{a}, \frac{1}{q} \langle \mathbf{a}, \mathbf{s} \rangle + e \right) \in \mathbb{Z}_q^n \times \mathbb{T}.$$

The *Learning With Errors* problem LWE_α , introduced by Regev, consists in assessing whether an oracle produces samples from $U(\mathbb{Z}_q^n \times \mathbb{T})$ or $A_{\mathbf{s}, \alpha}$, for some secret $\mathbf{s} \leftarrow U(\mathbb{Z}_q^n)$. Regev showed that, for prime $q \leq \text{poly}(n)$ and $\alpha \in (\sqrt{n}/(2q), 1)$, LWE_α is quantumly no easier than standard worst-case lattice problems in dimension n with approximation factors $\text{poly}(n)/\alpha$. This hardness proof was later partly dequantized, and the requirements that q should be prime and polynomial in n were waived.

In this work, we consider a bounded-sample variant of LWE. If m denotes the number of samples available to the distinguisher, we denote this restriction by $\text{LWE}_{\alpha, m}$. Equivalently, using linear algebra over $\mathbb{Z}_q$, given $\mathbf{A} \leftarrow U(\mathbb{Z}_q^{m \times n})$, the goal is to distinguish between the following distributions over $\mathbb{T}^m$:

$$\frac{1}{q} U(\text{Im}(\mathbf{A})) + \nu_\alpha^m \quad \text{and} \quad \frac{1}{q} U(\mathbb{Z}_q^m) + \nu_\alpha^m.$$

Under the assumption that $\alpha q \geq \Omega(\sqrt{n})$, the right hand side distribution is within statistical distance $2^{-\Omega(n)}$ of $U(\mathbb{T}^m)$. This formulation states that, after adding a small Gaussian noise, the linear spaces $\text{Im}(\mathbf{A})$ and $\mathbb{Z}_q^m$ become computationally indistinguishable.

We recall the $(k, \mathbf{S})$ -LWE variant of [21], which has been proved to be as hard as LWE (Theorem 3.12 in Section 3.5).

Definition 3.11 ($(k, \mathbf{S}, \mathbf{C})$ -LWE, [21]). *Let $k \leq m$, let $\mathbf{S} \in \mathbb{R}^{m \times m}$ be invertible, and let $\mathbf{C} = (\mathbf{c}_1 | \dots | \mathbf{c}_k) \in \mathbb{R}^{m \times k}$. The $(k, \mathbf{S}, \mathbf{C})$ -LWE $_{\alpha, m}$ problem is defined as follows. Given $\mathbf{A} \leftarrow U(\mathbb{Z}_q^{m \times n})$, $\mathbf{u} \leftarrow U(\mathbb{Z}_q^n)$, and*

$$\mathbf{x}_i \leftarrow D_{A_{\perp_{\mathbf{u}}}(\mathbf{A}), \mathbf{S}, \mathbf{c}_i} \quad \text{for } i \in [k],$$

the goal is to distinguish between the following distributions over $\mathbb{T}^{m+1}$:

$$\frac{1}{q} U\left(\text{Im}\left(\begin{smallmatrix} \mathbf{u}^t \\ \mathbf{A} \end{smallmatrix}\right)\right) + \nu_\alpha^{m+1} \quad \text{and} \quad \frac{1}{q} U\left(\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_k^+)^\perp\right) + \nu_\alpha^{m+1},$$

where $\mathbf{x}_i^+ = \begin{pmatrix} 1 \\ \mathbf{x}_i \end{pmatrix} \in \mathbb{Z}_q^{m+1}$. When $\mathbf{C} = \mathbf{0}$, we simply write $(k, \mathbf{S})$ -LWE $_{\alpha, m}$.

3.5 Hardness of k -LWE

Theorem 3.12 (Theorem 18 of [21]). *Let m, q, σ, σ', k be such that*

$$\sigma \geq \Omega\left(\max\left(m\sqrt{\log m}, (\sigma')^{k/(m+k)}\right)\right), \quad \sigma' \geq \Omega\left(m^3 \sigma^2 \log^{3/2}(m\sigma)\right),$$

$q \geq \Omega(\sigma' \sqrt{\log m})$ is prime, and $m \geq \Omega(n \log q)$. Let

$$\mathbf{S} = \begin{pmatrix} \sigma \mathbf{I}_m & 0 \\ 0 & \sigma' \mathbf{I}_k \end{pmatrix} \in \mathbb{R}^{(m+k) \times (m+k)}$$

and $\alpha' = \Omega(m^{3/2} \sigma' \alpha)$. Then there exists a probabilistic polynomial-time reduction from $\text{LWE}_{\alpha, m+1}$ in dimension n to $(k, \mathbf{S})$ -LWE $_{\alpha', m+k}$ in dimension n .

Throughout the paper, when we assume the hardness of $(k, \mathbf{S})\text{-LWE}_{\alpha, m}$, the matrix $\mathbf{S}$ is always of size $m \times m$. We choose $\mathbf{S}$ to be block-diagonal as in Theorem 3.12, with the parameters chosen so that the theorem applies after replacing its parameter m by $m - k$. Equivalently, this relies on the hardness of $\text{LWE}_{\alpha, m+1-k}$, up to the noise increase of Theorem 3.12. Additional constraints on the parameters will be imposed when needed for correctness.

4 Projective Sampling

We recall the projective sampling framework of [21], inspired by the projective hashing paradigm of Cramer and Shoup [?]. LPSS introduce this notion in order to make their traitor tracing signals publicly samplable. In their secret-key tracing procedure, the tracing signal is sampled from a distribution indexed by the users' secret keys. A projective sampling family provides, for each secret key, a projected key from which one can simulate the same sampling distribution.

4.1 Projective Sampling Families

The indistinguishability requirement is deliberately strong: the adversary is given both the secret key and its projected key, and should still not distinguish the secret-key sampler from the projected-key sampler. This is the relevant notion for traitor tracing, where the traitors are insiders and possess secret data.

Definition 4.1 (Projective sampling family, [21]). *Let X be a non-empty sample space, let $\mathcal{K}$ be a key space, and let $\mathcal{F} = (F_k)_{k \in \mathcal{K}}$ be a collection of sampling functions over X . We call $\text{Sam} = (\mathcal{F}, \mathcal{K}, X)$ a sampling family.*

Let $\mathcal{J}$ be a non-empty projected-key space, let $\pi : \mathcal{K} \rightarrow \mathcal{J}$ be a probabilistic projection function, let $\mathcal{P} = (P_j)_{j \in \mathcal{J}}$ be a collection of sampling functions over X , and let D be a distribution over $\mathcal{K}$. We say that

$$\text{PSam} = (\mathcal{F}, \mathcal{K}, X, \mathcal{P}, \mathcal{J}, \pi, D)$$

is a projective sampling family if, with overwhelming probability over $k, k' \leftarrow D$ and the randomness of the projection function:

- **Correctness.** *Given k and $\pi(k)$, the distributions output by F_k and $P_{\pi(k)}$ are computationally indistinguishable;*
- **Soundness.** *Given k, k' and the projected keys $\pi(k), \pi(k')$, the distributions output by F_k and $P_{\pi(k')}$ can be efficiently distinguished.*

4.2 From Secret-Key Samplers to Public Samplers

Projective sampling is a sampler-level tool for turning private-key sampling procedures into public ones. If a private-key construction samples from a distribution F_k indexed by a recipient secret key k , then publishing a projected key $\pi(k)$ allows the public algorithm to sample instead from $P_{\pi(k)}$. The correctness condition of projective sampling ensures that this replacement preserves the intended sampling distribution.

Turning this sampler replacement into a full security proof usually requires one additional ingredient controlling the information carried by the projected keys. In the broadcast setting, a natural sufficient condition is the existence of a common reference distribution: the secret-key samplers and the projected samplers are both indistinguishable from the same distribution, even in the public view containing the projected keys. This is the form obtained in the LPSS lattice construction and used by DPY in anonymous broadcast encryption, where the anonymity hybrids add or remove target users while preserving the ciphertext distribution. Other assumptions on the projection map could serve the same purpose in different applications.

The construction below recalls the LPSS projective-sampling family in a form adapted to our later trapdoor use. The common reference distribution obtained there is the distributional bridge preserved by our trapdoor variant.

4.3 Projective Sampling from k -LWE

We recall the projective sampling technique of [21], in a slightly more general dimensional form. In [21], the projective matrix $\mathbf{H}$ has size $m \times (m - n)$. We will instead use an integer d and take $\mathbf{H} \in \mathbb{Z}_q^{m \times d}$. We show that their projective-sampling theorem still holds for any number of columns d with $n < d \leq m - n$ (Theorem 4.2). This will enable us later to choose d such that $\mathbf{H}$ can be sampled with a trapdoor.

In the concrete families below, the public parameters are sampled first and then kept implicit in the notation of the sampling family.

Let $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ and $\mathbf{u} \in \mathbb{Z}_q^n$ be public parameters, and define

$$\mathbf{A}^+ = \begin{pmatrix} \mathbf{u}^t \\ \mathbf{A} \end{pmatrix} \in \mathbb{Z}_q^{(m+1) \times n}.$$

For each i , we define PSam_i as follows:

- Let $\mathcal{K}_i = (\mathbb{Z}_q^m)^i$. Let $\mathcal{D}_i$ sample $k = (\mathbf{x}_j)_{j \leq i}$ by taking

$$\mathbf{x}_j \leftarrow D_{\Lambda_{-\mathbf{u}}^\perp(\mathbf{A}), \mathbf{S}} \quad \text{for } j \leq i$$

and reducing the samples modulo q . We write $\mathbf{x}_j^+ = \begin{pmatrix} 1 \\ \mathbf{x}_j \end{pmatrix} \in \mathbb{Z}_q^{m+1}$.

- The secret-key sampler is

$$F_{i,k} = \frac{1}{q} U\left(\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^\perp\right) + \nu_\alpha^{m+1},$$

with the convention $\text{Span}(\emptyset)^\perp = \mathbb{Z}_q^{m+1}$.

- Choose d such that $n < d \leq m - n$. The projected key $\pi_i(k)$ is defined by sampling $\mathbf{H} \leftarrow U(\mathbb{Z}_q^{m \times d})$, conditioned on $\text{rank}(\mathbf{H}) = d$ and $\text{Im}(\mathbf{A}) \subseteq \text{Im}(\mathbf{H})$, and setting

$$\mathbf{h}_j = -\mathbf{H}^t \mathbf{x}_j \bmod q \quad \text{for } j \leq i.$$

We let

$$\mathcal{J}_i = \mathbb{Z}_q^{m \times d} \times (\mathbb{Z}_q^d)^i, \quad \pi_i(k) = (\mathbf{H}, (\mathbf{h}_j)_{j \leq i}) \in \mathcal{J}_i.$$

- Given $\pi_i(k)$, set $\mathbf{H}_j = (\mathbf{h}_j^t || \mathbf{H}) \in \mathbb{Z}_q^{(m+1) \times d}$. The projected sampler is

$$P_{i,\pi_i(k)} = \frac{1}{q} U\left(\bigcap_{j \leq i} \text{Im}(\mathbf{H}_j)\right) + \nu_\alpha^{m+1},$$

with the convention $\bigcap_{j \leq 0} \text{Im}(\mathbf{H}_j) = \mathbb{Z}_q^{m+1}$.

Since $\text{Im}(\mathbf{A}) \subseteq \text{Im}(\mathbf{H})$, there exists $\mathbf{R} \in \mathbb{Z}_q^{d \times n}$ such that $\mathbf{A} = \mathbf{H}\mathbf{R}$. Moreover, because $\mathbf{A}^t \mathbf{x}_j = -\mathbf{u} \bmod q$, we have $\mathbf{R}^t \mathbf{h}_j = \mathbf{u}$. Hence $\mathbf{A}^+ = \mathbf{H}_j \mathbf{R}$, and therefore

$$\text{Im}(\mathbf{A}^+) \subseteq \bigcap_{j \leq i} \text{Im}(\mathbf{H}_j).$$

On the other hand, by construction $(\mathbf{x}_j^+)^t \mathbf{H}_j = \mathbf{0}$, so

$$\bigcap_{j \leq i} \text{Im}(\mathbf{H}_j) \subseteq \text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^\perp.$$

Theorem 4.2 (Adapted from Theorem 27 of [21]). *For each $i \geq 0$, let*

$$\text{PSam}_i = (\mathcal{F}_i, \mathcal{K}_i, \mathbb{T}^{m+1}, \mathcal{P}_i, \mathcal{J}_i, \pi_i, \mathcal{D}_i),$$

where $\mathcal{F}_i = (F_{i,k})_{k \in \mathcal{K}_i}$ and $\mathcal{P}_i = (P_{i,\pi_i(k)})_{\pi_i(k) \in \mathcal{J}_i}$ are defined above. Under the $(i, \mathbf{S})$ -LWE $_{\alpha,m}$ hardness assumption, given the public parameters $(\mathbf{A}, \mathbf{u})$, a secret key $k = (\mathbf{x}_j)_{j \leq i} \leftarrow \mathcal{D}_i$, and its projected key $\pi_i(k) = (\mathbf{H}, (\mathbf{h}_j)_{j \leq i})$, the distributions $F_{i,k}$ and $P_{i,\pi_i(k)}$, namely

$$\frac{1}{q} U\left(\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^\perp\right) + \nu_\alpha^{m+1} \quad \text{and} \quad \frac{1}{q} U\left(\bigcap_{j \leq i} \text{Im}(\mathbf{H}_j)\right) + \nu_\alpha^{m+1}$$

are computationally indistinguishable. Moreover, both distributions are computationally indistinguishable from the common reference distribution

$$\frac{1}{q} U(\text{Im}(\mathbf{A}^+)) + \nu_\alpha^{m+1}.$$

If moreover $i \geq 1$, then PSam_i is a projective sampling family. Namely, if $k' = (\mathbf{x}'_j)_{j \leq i}$ is sampled independently and $\pi_i(k')$ denotes its projected key, then with overwhelming probability over k and k' , the distributions

$$\frac{1}{q} U\left(\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^\perp\right) + \nu_\alpha^{m+1} \quad \text{and} \quad P_{i,\pi_i(k')}$$

can be efficiently distinguished, where $P_{i,\pi_i(k')}$ denotes the corresponding projected sampler.

Proof. We adapt the proof of [21, Theorem 27] to the case $\mathbf{H} \in \mathbb{Z}_q^{m \times d}$. As observed above,

$$\text{Im}(\mathbf{A}^+) \subseteq \bigcap_{j \leq i} \text{Im}(\mathbf{H}_j) \subseteq \text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^\perp.$$

First, by the $(i, \mathbf{S})$ -LWE $_{\alpha,m}$ assumption, given $(\mathbf{A}, \mathbf{u})$ and the secret vectors $(\mathbf{x}_j)_{j \leq i}$, the distributions

$$\frac{1}{q} U\left(\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^\perp\right) + \nu_\alpha^{m+1} \quad \text{and} \quad \frac{1}{q} U(\text{Im}(\mathbf{A}^+)) + \nu_\alpha^{m+1}$$

are computationally indistinguishable. The projected key $(\mathbf{H}, (\mathbf{h}_j)_{j \leq i})$ is efficiently sampled from $(\mathbf{A}, \mathbf{u}, (\mathbf{x}_j)_{j \leq i})$, by sampling $\mathbf{H}$ conditioned on $\text{Im}(\mathbf{A}) \subseteq \text{Im}(\mathbf{H})$ and setting $\mathbf{h}_j = -\mathbf{H}^t \mathbf{x}_j$. Hence adding this projected key to the public view preserves the indistinguishability.

It remains to compare the projected sampler with the secret-key sampler. Given $(\mathbf{H}, (\mathbf{h}_j)_{j \leq i})$, the subspace $\bigcap_{j \leq i} \text{Im}(\mathbf{H}_j)$ can be computed by linear algebra over $\mathbb{Z}_q$, and one can sample uniformly from it. Let

$$\mathbf{y} \leftarrow U\left(\bigcap_{j \leq i} \text{Im}(\mathbf{H}_j)\right)$$

independently. The inclusions above imply

$$U(\text{Im}(\mathbf{A}^+)) + U\left(\bigcap_{j \leq i} \text{Im}(\mathbf{H}_j)\right) = U\left(\bigcap_{j \leq i} \text{Im}(\mathbf{H}_j)\right),$$

and

$$U\left(\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^\perp\right) + U\left(\bigcap_{j \leq i} \text{Im}(\mathbf{H}_j)\right) = U\left(\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^\perp\right).$$

Therefore, adding $\frac{1}{q}\mathbf{y}$ to an LWE challenge transforms

$$\frac{1}{q}U(\text{Im}(\mathbf{A}^+)) + \nu_\alpha^{m+1} \quad \text{into} \quad \frac{1}{q}U\left(\bigcap_{j \leq i} \text{Im}(\mathbf{H}_j)\right) + \nu_\alpha^{m+1},$$

and transforms

$$\frac{1}{q}U\left(\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^\perp\right) + \nu_\alpha^{m+1} \quad \text{into itself.}$$

By post-processing, applying this transformation to the two indistinguishable distributions above preserves computational indistinguishability. Therefore

$$\frac{1}{q}U\left(\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^\perp\right) + \nu_\alpha^{m+1} \quad \text{and} \quad \frac{1}{q}U\left(\bigcap_{j \leq i} \text{Im}(\mathbf{H}_j)\right) + \nu_\alpha^{m+1}$$

are computationally indistinguishable. This proves the first indistinguishability claim. Since the $(i, \mathbf{S})$ -LWE $_{\alpha, m}$ assumption gives indistinguishability between

$$\frac{1}{q}U\left(\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^\perp\right) + \nu_\alpha^{m+1}$$

and

$$\frac{1}{q}U(\text{Im}(\mathbf{A}^+)) + \nu_\alpha^{m+1},$$

the triangle inequality also gives that the projected distribution over $\bigcap_{j \leq i} \text{Im}(\mathbf{H}_j)$ is indistinguishable from the distribution over $\text{Im}(\mathbf{A}^+)$.

It remains to justify the soundness statement when $i \geq 1$. Let $k = (\mathbf{x}_j)_{j \leq i}$ and $k' = (\mathbf{x}'_j)_{j \leq i}$ be sampled independently. By [21, Lemma 12], with overwhelming probability there exists some $\ell \leq i$ such that

$$\mathbf{x}'_\ell{}^+ \notin \text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+).$$

Consider the test that takes a sample $\mathbf{y} \in \mathbb{T}^{m+1}$ and computes

$$\langle \mathbf{x}'_\ell{}^+, \mathbf{y} \rangle \in \mathbb{T}.$$

If $\mathbf{y}$ is sampled from $P_{i, \pi_i(k')}$, then its noiseless part lies in $\bigcap_{j \leq i} \text{Im}(\mathbf{H}'_j) \subseteq \text{Span}(\mathbf{x}'_1^+, \dots, \mathbf{x}'_i^+)^\perp$, and this inner product is only the corresponding Gaussian error. On the other hand, if $\mathbf{y}$ is sampled from $F_{i, k}$, then the noiseless part is uniform over $\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^\perp$; since $\mathbf{x}'_\ell{}^+$ is not in $\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)$, the inner product is uniform over $\frac{1}{q}\mathbb{Z}_q$. As in [21, Theorem 27], these two cases are efficiently distinguishable. $\square$

Remark 4.3. For the basic projective-sampling indistinguishability statement, where only the projected keys $(\mathbf{h}_j)_{j \leq i}$ are given, the condition $\text{Im}(\mathbf{A}) \subseteq \text{Im}(\mathbf{H})$ only requires $d \geq n$. However, the soundness condition is meant to rule out degenerate projections independent of the secret key; in particular, one wants a non-trivial quotient $\text{Im}(\mathbf{H})/\text{Im}(\mathbf{A})$, hence $d > n$. The upper bound

$$d \leq m - n$$

is the one compatible with the LPSS simulation of public projected keys under an LWE assumption in dimension n . The choice $d = m - n$ made in [21] is the maximal choice in this range. In applications where the quotient $\text{Im}(\mathbf{H})/\text{Im}(\mathbf{A})$ must carry extra public randomness, we will use the non-degenerate range $n < d \leq m - n$.

Remark 4.4. For $i = 0$, the correctness and common-reference indistinguishability statements remain valid, but the soundness statement is void and is not needed in the applications below. We therefore state the projective-sampling-family claims for $i \geq 1$.

5 Trapdoor Projective Sampling

5.1 Trapdoor Projective Sampling Families

Definition 5.1 (Trapdoor projective sampling family). A trapdoor projective sampling family is a tuple

$$\text{TPSam} = (\mathcal{F}, \mathcal{K}, X, \mathcal{P}, \mathcal{J}, D_{\mathcal{K}}, D_{\mathcal{J}}, \text{TrapGen}, \mathcal{S}),$$

whose components are:

- a finite non-empty output space X , a secret-key space $\mathcal{K}$ and a projected-key space $\mathcal{J}$;
- the secret-key samplers $\mathcal{F} = (F_k)_{k \in \mathcal{K}}$ and the projected samplers $\mathcal{P} = (P_j)_{j \in \mathcal{J}}$, all over X ;
- the distribution $D_{\mathcal{K}}$ of secret keys and the target distribution $D_{\mathcal{J}}$ of projected keys;
- a probabilistic polynomial-time algorithm TrapGen , which on input 1^n outputs a projection map $\pi : \mathcal{K} \rightarrow \mathcal{J}$ together with a trapdoor t ;
- a probabilistic polynomial-time sampler $\mathcal{S}$, which on input t and a projected key $p \in \mathcal{J}$ outputs a secret key in $\mathcal{K}$.

It is required that, with overwhelming probability over $(\pi, t) \leftarrow \text{TrapGen}(1^n)$:

- **Correctness.** For $p \leftarrow D_{\mathcal{J}}$ and $k \leftarrow \mathcal{S}(t, p)$, given k and p , the samplers F_k and P_p are computationally indistinguishable.
- **Soundness.** For independent $p, p' \leftarrow D_{\mathcal{J}}$ conditioned on $p \neq p'$, and for $k \leftarrow \mathcal{S}(t, p)$ and $k' \leftarrow \mathcal{S}(t, p')$, given k, k', p, p' , the samplers F_k and $P_{p'}$ can be efficiently distinguished.
- **Projected keys.** The image distribution $\pi(D_{\mathcal{K}})$ is statistically close to $D_{\mathcal{J}}$, and every p in the support of $D_{\mathcal{J}}$ satisfies $\Pr_{k \leftarrow D_{\mathcal{K}}}[\pi(k) = p] > 0$.
- **Preimage sampling with trapdoor.** For $p \leftarrow D_{\mathcal{J}}$, the sampler $\mathcal{S}(t, p)$ outputs a distribution statistically close to $D_{\mathcal{K}}$ conditioned on $\pi(k) = p$.

Proposition 5.2 (Induced underlying projective sampling). Let TPSam be a trapdoor projective sampling family such that $\Pr_{p, p' \leftarrow D_{\mathcal{J}}}[p = p'] \leq \text{negl}(n)$. Then, with overwhelming probability over $(\pi, t) \leftarrow \text{TrapGen}(1^n)$, the tuple $(\mathcal{F}, \mathcal{K}, X, \mathcal{P}, \mathcal{J}, \pi, D_{\mathcal{K}})$ is a projective sampling family.

Proof. By the projected-key and preimage-sampling conditions, the distribution obtained by sampling $p \leftarrow D_{\mathcal{J}}$ and then $k \leftarrow \mathcal{S}(t, p)$ is statistically close to the distribution obtained by sampling $k \leftarrow D_{\mathcal{K}}$ and setting $p = \pi(k)$. The correctness condition of the trapdoor projective sampling family therefore transfers to the experiment in which $k \leftarrow D_{\mathcal{K}}$ and $p = \pi(k)$.

For soundness, apply the same statistical closeness independently to two projected keys. The experiment with independent $k, k' \leftarrow D_{\mathcal{K}}$, $p = \pi(k)$ and $p' = \pi(k')$ is statistically close to the experiment with independent $p, p' \leftarrow D_{\mathcal{J}}$, $k \leftarrow \mathcal{S}(t, p)$ and $k' \leftarrow \mathcal{S}(t, p')$. By the collision assumption, the event $p = p'$ has negligible probability, so conditioning on $p \neq p'$ changes the experiment only negligibly. The soundness condition of the trapdoor projective sampling family then gives the required distinguisher for F_k and $P_{\pi(k')}$. $\square$

5.2 Trapdoor Projective Sampling from k -LWE

We now instantiate trapdoor projective sampling under k -LWE. We first describe the family and state the result (Theorem 5.3). Its proof has two steps. First, we show that generating the public parameters and the keys in the reversed order, with a trapdoor for $\mathbf{H}$, yields a distribution statistically close to that of the LPSS experiment (Lemma 5.4). Second, we transfer the LPSS projective-sampling theorem to this trapdoor setup. Finally, we fix concrete parameters with $d = \Theta(n)$ (Theorem 5.6).

The trapdoor projective sampling family. For each i , we define the family TPSam_i as follows.

- *Trapdoor generation.* TrapGen generates $(\mathbf{H}, \mathbf{T}_{\mathbf{H}}) \in \mathbb{Z}_q^{m \times d} \times \mathbb{Z}^{m \times m}$ using Lemma 3.10, samples $\mathbf{R} \leftarrow U(\mathbb{Z}_q^{d \times n})$ of rank n , and sets $\mathbf{A} = \mathbf{H}\mathbf{R}$. It then samples $\mathbf{h}_0 \leftarrow U(\mathbb{Z}_q^d)$, sets $\mathbf{u} = \mathbf{R}^t \mathbf{h}_0$, and computes a matrix $\mathbf{K} \in \mathbb{Z}_q^{d \times (d-n)}$ whose columns form a basis of $\ker(\mathbf{R}^t)$. It outputs the public values $(\mathbf{A}, \mathbf{u}, \mathbf{H})$ and keeps $\mathbf{T}_{\mathbf{H}}$ as trapdoor. The values $\mathbf{h}_0$ and $\mathbf{K}$ only serve to sample projected keys; by Lemma 5.5, they may be published.
- *Keys.* The secret-key and projected-key spaces are $\mathcal{K}_i = (\mathbb{Z}_q^m)^i$ and $\mathcal{J}_i = (\mathbb{Z}_q^d)^i$, and the projection is $\pi_i(k) = (-\mathbf{H}^t \mathbf{x}_j)_{j \leq i}$ for $k = (\mathbf{x}_j)_{j \leq i}$. Admissible projected keys lie in

$$\mathcal{H}_{\mathbf{u}} = \{\mathbf{h} \in \mathbb{Z}_q^d : \mathbf{R}^t \mathbf{h} = \mathbf{u}\} = \mathbf{h}_0 + \ker(\mathbf{R}^t),$$

and $D_{\mathcal{J}_i} = U(\mathcal{H}_{\mathbf{u}}^i)$: a sample is $p = (\mathbf{h}_0 + \mathbf{K}\mathbf{a}_j)_{j \leq i}$ with $\mathbf{a}_j \leftarrow U(\mathbb{Z}_q^{d-n})$. The key distribution $\mathcal{D}_i$ samples $\mathbf{x}_j \leftarrow D_{\Lambda_{\perp_{\mathbf{u}}}^+(\mathbf{A}), \mathbf{S}}$ for $j \leq i$, and outputs $k = (\mathbf{x}_j \bmod q)_{j \leq i}$.

- *Trapdoor sampler.* Given $\mathbf{T}_{\mathbf{H}}$ and $p = (\mathbf{h}_j)_{j \leq i}$, the sampler $\mathcal{S}$ samples each $\mathbf{x}_j \leftarrow D_{\Lambda_{\perp_{\mathbf{h}_j}}}(\mathbf{H}), \mathbf{S}$ with $\mathbf{T}_{\mathbf{H}}$, using Lemma 3.7.
- *Samplers.* For $k = (\mathbf{x}_j)_{j \leq i}$ and $p = (\mathbf{h}_j)_{j \leq i}$, let $\mathbf{H}_j = (\mathbf{h}_j^t || \mathbf{H})$. The secret-key and projected samplers over $\mathbb{T}^{m+1}$ are

$$F_{i,k} = \frac{1}{q} U\left(\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^{\perp}\right) + \nu_{\alpha}^{m+1},$$

$$P_{i,p} = \frac{1}{q} U\left(\bigcap_{j \leq i} \text{Im}(\mathbf{H}_j)\right) + \nu_{\alpha}^{m+1}.$$

Theorem 5.3 (Trapdoor projective sampling from k -LWE). *Let q be prime, let n, m, d be integers such that $n < d \leq m - n$, let $i \geq 0$, and let $\mathbf{S}$ be a positive diagonal matrix. Assume that the generated matrix $\mathbf{H}$ satisfies*

$$s_{\min}(\mathbf{S}) \geq \max\left(\eta_{\varepsilon}(\Lambda^{\perp}(\mathbf{H})), 3mq^{d/m} \cdot \omega(\sqrt{\log m})\right).$$

Under the $(i, \mathbf{S})$ -LWE $_{\alpha, m}$ hardness assumption, the following holds. If $i \geq 1$, the family TPSam_i defined above is a trapdoor projective sampling family. In all cases, given the public values $(\mathbf{A}, \mathbf{u}, \mathbf{H})$, a secret key $k = (\mathbf{x}_j)_{j \leq i} \leftarrow \mathcal{D}_i$, and its projected key $\pi_i(k) = (\mathbf{h}_j)_{j \leq i}$, the distributions $F_{i,k}$ and

$P_{i,\pi_i(k)}$, namely

$$\frac{1}{q}U\left(\text{Span}(\mathbf{x}_1^+, \dots, \mathbf{x}_i^+)^\perp\right) + \nu_\alpha^{m+1} \quad \text{and} \quad \frac{1}{q}U\left(\bigcap_{j \leq i} \text{Im}(\mathbf{H}_j)\right) + \nu_\alpha^{m+1}$$

are computationally indistinguishable. Moreover, both distributions are computationally indistinguishable from

$$\frac{1}{q}U(\text{Im}(\mathbf{A}^+)) + \nu_\alpha^{m+1}.$$

Proof of Theorem 5.3. We compare the family with the LPSS experiment. The two experiments are given below. **Regular** is the LPSS experiment. **Reversed** samples the public values of TPSam_i with TrapGen and its keys with $D_{\mathcal{J}_i}$ and $\mathcal{S}$; its name refers to the order in which it samples the keys, projected keys first.

Regular(q, n, m, d, i)

- Sample $\mathbf{A} \leftarrow U(\mathbb{Z}_q^{m \times n})$ and $\mathbf{u} \leftarrow U(\mathbb{Z}_q^n)$. Sample $\mathbf{H} \leftarrow U(\mathbb{Z}_q^{m \times d})$, conditioned on $\text{rank}(\mathbf{H}) = d$ and $\text{Im}(\mathbf{A}) \subseteq \text{Im}(\mathbf{H})$.
- For $1 \leq j \leq i$, sample $\mathbf{x}_j \leftarrow D_{\Lambda_{\perp_{\mathbf{u}}}(\mathbf{A}), \mathbf{S}}$ and set $\mathbf{h}_j = -\mathbf{H}^t \mathbf{x}_j \bmod q$.
- Return $(\mathbf{A}, \mathbf{u}, \mathbf{H}, (\mathbf{h}_j)_{j \leq i}, (\mathbf{x}_j)_{j \leq i})$.

Reversed(q, n, m, d, i)

- Generate $(\mathbf{H}, \mathbf{T}_{\mathbf{H}})$ using Lemma 3.10. Sample $\mathbf{R} \leftarrow U(\mathbb{Z}_q^{d \times n})$, conditioned on $\text{rank}(\mathbf{R}) = n$, and set $\mathbf{A} = \mathbf{H}\mathbf{R}$. Sample $\mathbf{h}_0 \leftarrow U(\mathbb{Z}_q^d)$, set $\mathbf{u} = \mathbf{R}^t \mathbf{h}_0$, and compute a matrix $\mathbf{K} \in \mathbb{Z}_q^{d \times (d-n)}$ whose columns form a basis of $\ker(\mathbf{R}^t)$.
- For $1 \leq j \leq i$, sample $\mathbf{a}_j \leftarrow U(\mathbb{Z}_q^{d-n})$, set $\mathbf{h}_j = \mathbf{h}_0 + \mathbf{K}\mathbf{a}_j$, and sample $\mathbf{x}_j \leftarrow D_{\Lambda_{\perp_{\mathbf{h}_j}}(\mathbf{H}), \mathbf{S}}$ with $\mathbf{T}_{\mathbf{H}}$, using Lemma 3.7.
- Return $(\mathbf{A}, \mathbf{u}, \mathbf{H}, (\mathbf{h}_j)_{j \leq i}, (\mathbf{x}_j)_{j \leq i})$.

Lemma 5.4. Let q be prime, let $\varepsilon \in (0, 1/2)$, let n, m, d be integers such that $n < d \leq m - n$, and let $i \geq 0$. Let $\mathbf{S}$ be a positive diagonal matrix. Assume that the generated matrix $\mathbf{H}$ and its trapdoor $\mathbf{T}_{\mathbf{H}}$ satisfy

$$s_{\min}(\mathbf{S}) \geq \max\left(\eta_\varepsilon(\Lambda^\perp(\mathbf{H})), \left\|\tilde{\mathbf{T}}_{\mathbf{H}}\right\| \cdot \omega(\sqrt{\log m})\right).$$

Then the distributions induced by the regular and the reversed experiments above are within statistical distance at most $2i\varepsilon + \text{negl}(n)$, where the negligible term accounts for the rank failures of Lemma 3.3, the trapdoor generator of Lemma 3.10 and the i coset samples of Lemma 3.7.

Proof. We define hybrid distributions $\text{Hyb}_0, \dots, \text{Hyb}_5$, all with outputs of the form $(\mathbf{A}, \mathbf{u}, \mathbf{H}, (\mathbf{h}_j)_{j \leq i}, (\mathbf{x}_j)_{j \leq i})$. Each hybrid modifies the previous one, and we bound the statistical distance between consecutive hybrids.

Hyb₀ *Regular experiment.* This is the distribution induced by **Regular**.

Hyb₁ *Reversed setup.* We replace the sampling of $(\mathbf{A}, \mathbf{u}, \mathbf{H})$ as follows: sample $\mathbf{H} \leftarrow U(\mathbb{Z}_q^{m \times d})$ conditioned on $\text{rank}(\mathbf{H}) = d$, sample $\mathbf{R} \leftarrow U(\mathbb{Z}_q^{d \times n})$ conditioned on $\text{rank}(\mathbf{R}) = n$, set $\mathbf{A} = \mathbf{H}\mathbf{R}$, sample $\mathbf{h}_0 \leftarrow U(\mathbb{Z}_q^d)$ and set $\mathbf{u} = \mathbf{R}^t \mathbf{h}_0$. The keys are sampled from $(\mathbf{A}, \mathbf{u}, \mathbf{H})$ as in Hyb_0 .

Let $V = \text{Im}(\mathbf{A})$ and $W = \text{Im}(\mathbf{H})$. We show that Hyb_0 , conditioned on $\text{rank}(\mathbf{A}) = n$, and Hyb_1 induce the same distribution on $(\mathbf{A}, \mathbf{u}, \mathbf{H})$. Since the keys are derived from $(\mathbf{A}, \mathbf{u}, \mathbf{H})$ in the same way, the two hybrids then coincide.

- *The pair (V, W) has the same law.* In Hyb_0 , V is a uniform subspace of dimension n , and W is a uniform subspace of dimension d containing V . Hence (V, W) is uniform among the pairs $V \subseteq W$ with $\dim V = n$ and $\dim W = d$. In Hyb_1 , W is a uniform subspace of dimension d , and $\text{Im}(\mathbf{R})$ is a uniform subspace of $\mathbb{Z}_q^d$ of dimension n . Since $\mathbf{H}$ is an isomorphism from $\mathbb{Z}_q^d$ onto W , the subspace $V = \mathbf{H}(\text{Im}(\mathbf{R}))$ is uniform among the subspaces of W of dimension n , and (V, W) is again uniform among these pairs.
- *Given (V, W) , the pair $(\mathbf{A}, \mathbf{H})$ has the same law.* In both hybrids, $\mathbf{H}$ is a uniform ordered basis of W , and $\mathbf{A}$ is a uniform ordered basis of V , independent of $\mathbf{H}$. In Hyb_1 , this holds because, given $\mathbf{H}$ and V , the matrix $\mathbf{R}$ is a uniform ordered basis of $\mathbf{H}^{-1}(V)$, and $\mathbf{R} \mapsto \mathbf{H}\mathbf{R}$ maps these bases bijectively onto the ordered bases of V .
- *The vector $\mathbf{u}$ has the same law.* In Hyb_0 , $\mathbf{u}$ is uniform and independent of $(\mathbf{A}, \mathbf{H})$. In Hyb_1 , the matrix $\mathbf{R}$ is determined by $(\mathbf{A}, \mathbf{H})$, since $\mathbf{H}$ has full column rank, and $\mathbf{u} = \mathbf{R}^t \mathbf{h}_0$ is uniform over $\mathbb{Z}_q^n$ given $(\mathbf{A}, \mathbf{H})$, since $\mathbf{h}_0$ is uniform and $\mathbf{R}$ has rank n .

Hence, by Lemma 3.3, $\Delta(\text{Hyb}_0, \text{Hyb}_1) \leq \Pr_{\text{Hyb}_0}[\text{rank}(\mathbf{A}) < n] \leq \delta_{m,n} \leq 2q^{-(m-n+1)}$, which is negligible. From now on, $\mathbf{H}$ has rank d and $\mathbf{R}$ has rank n .

Hyb_2 *Splitting the Gaussian along the fibres.* Since $\mathbf{A} = \mathbf{H}\mathbf{R}$ and $\mathbf{u} = \mathbf{R}^t \mathbf{h}_0$, a vector $\mathbf{x}$ satisfies $\mathbf{A}^t \mathbf{x} = -\mathbf{u}$ if and only if $-\mathbf{H}^t \mathbf{x} \in \mathcal{H}_{\mathbf{u}}$, so that

$$\Lambda_{-\mathbf{u}}^\perp(\mathbf{A}) = \bigsqcup_{\mathbf{h} \in \mathcal{H}_{\mathbf{u}}} \Lambda_{-\mathbf{h}}^\perp(\mathbf{H}).$$

In Hyb_1 , $\mathbf{x}_j \leftarrow D_{\Lambda_{-\mathbf{u}}^\perp(\mathbf{A}), \mathbf{S}}$ and $\mathbf{h}_j = -\mathbf{H}^t \mathbf{x}_j$. Hence, for every $\mathbf{h} \in \mathcal{H}_{\mathbf{u}}$ and every $\mathbf{x} \in \Lambda_{-\mathbf{h}}^\perp(\mathbf{H})$,

$$\begin{aligned} \Pr[\mathbf{h}_j = \mathbf{h}] &= \frac{\rho_{\mathbf{S}}(\Lambda_{-\mathbf{h}}^\perp(\mathbf{H}))}{\rho_{\mathbf{S}}(\Lambda_{-\mathbf{u}}^\perp(\mathbf{A}))} =: \mu_{\mathcal{H}_{\mathbf{u}}}(\mathbf{h}), \\ \Pr[\mathbf{x}_j = \mathbf{x} \mid \mathbf{h}_j = \mathbf{h}] &= \frac{\rho_{\mathbf{S}}(\mathbf{x})}{\rho_{\mathbf{S}}(\Lambda_{-\mathbf{h}}^\perp(\mathbf{H}))}. \end{aligned}$$

The second expression is the probability of $\mathbf{x}$ under $D_{\Lambda_{-\mathbf{h}}^\perp(\mathbf{H}), \mathbf{S}}$. In Hyb_2 , we therefore sample $\mathbf{h}_j \leftarrow \mu_{\mathcal{H}_{\mathbf{u}}}$, then $\mathbf{x}_j \leftarrow D_{\Lambda_{-\mathbf{h}_j}^\perp(\mathbf{H}), \mathbf{S}}$. This does not change the distribution, so $\text{Hyb}_2 = \text{Hyb}_1$.

Hyb_3 *Smoothing the projected keys.* We now sample $\mathbf{h}_j \leftarrow U(\mathcal{H}_{\mathbf{u}})$ instead of $\mu_{\mathcal{H}_{\mathbf{u}}}$. Consider the map

$$\Phi : \Lambda_{-\mathbf{u}}^\perp(\mathbf{A}) / \Lambda^\perp(\mathbf{H}) \rightarrow \mathcal{H}_{\mathbf{u}}, \quad [\mathbf{x}] \mapsto -\mathbf{H}^t \mathbf{x}.$$

It is well defined and injective, since $\mathbf{H}^t \mathbf{x} = \mathbf{H}^t \mathbf{x}'$ if and only if $\mathbf{x} - \mathbf{x}' \in \Lambda^\perp(\mathbf{H})$. It is surjective: since $\mathbf{H}$ has rank d , every $\mathbf{h} \in \mathcal{H}_{\mathbf{u}}$ has a preimage $\mathbf{x} \in \mathbb{Z}^m$ with $-\mathbf{H}^t \mathbf{x} = \mathbf{h}$, and then $\mathbf{A}^t \mathbf{x} = \mathbf{R}^t \mathbf{H}^t \mathbf{x} = -\mathbf{R}^t \mathbf{h} = -\mathbf{u}$, so $\mathbf{x} \in \Lambda_{-\mathbf{u}}^\perp(\mathbf{A})$. By definition, Φ maps $D_{\Lambda_{-\mathbf{u}}^\perp(\mathbf{A}), \mathbf{S}}$ modulo $\Lambda^\perp(\mathbf{H})$ to $\mu_{\mathcal{H}_{\mathbf{u}}}$, and the uniform distribution on $\Lambda_{-\mathbf{u}}^\perp(\mathbf{A}) / \Lambda^\perp(\mathbf{H})$ to $U(\mathcal{H}_{\mathbf{u}})$. Since $s_{\min}(\mathbf{S}) \geq \eta_\varepsilon(\Lambda^\perp(\mathbf{H}))$, Corollary 3.9 bounds the distance between $D_{\Lambda_{-\mathbf{u}}^\perp(\mathbf{A}), \mathbf{S}}$ modulo $\Lambda^\perp(\mathbf{H})$ and $U(\Lambda_{-\mathbf{u}}^\perp(\mathbf{A}) / \Lambda^\perp(\mathbf{H}))$ by 2ε , hence $\Delta(\mu_{\mathcal{H}_{\mathbf{u}}}, U(\mathcal{H}_{\mathbf{u}})) \leq 2\varepsilon$ by post-processing. The i projected keys are independent given $(\mathbf{A}, \mathbf{u}, \mathbf{H})$, so $\Delta(\text{Hyb}_2, \text{Hyb}_3) \leq 2i\varepsilon$.

Hyb_4 *Parametrizing $\mathcal{H}_{\mathbf{u}}$.* We compute a matrix $\mathbf{K}$ whose columns form a basis of $\ker(\mathbf{R}^t)$, sample $\mathbf{a}_j \leftarrow U(\mathbb{Z}_q^{d-n})$, and set $\mathbf{h}_j = \mathbf{h}_0 + \mathbf{K}\mathbf{a}_j$. Since $\mathbf{a} \mapsto \mathbf{h}_0 + \mathbf{K}\mathbf{a}$ is a bijection from $\mathbb{Z}_q^{d-n}$ onto $\mathcal{H}_{\mathbf{u}}$, the vector $\mathbf{h}_j$ is still uniform over $\mathcal{H}_{\mathbf{u}}$, and $\text{Hyb}_4 = \text{Hyb}_3$. This is the reversed experiment with a uniform matrix $\mathbf{H}$ of rank d and exact Gaussian samplers.

Hyb₅ Introducing the trapdoor. We generate $(\mathbf{H}, \mathbf{T}_{\mathbf{H}})$ using Lemma 3.10, and sample each $\mathbf{x}_j$ with $\mathbf{T}_{\mathbf{H}}$ using Lemma 3.7, which gives the experiment **Reversed**. By Lemma 3.10, the generated matrix $\mathbf{H}$ is within distance $2^{-\Omega(n)}$ of $U(\mathbb{Z}_q^{m \times d})$. By Lemma 3.3, the latter is within distance $\delta_{m,d} \leq 2q^{-(m-d+1)} \leq 2q^{-(n+1)}$ of the uniform distribution on rank- d matrices used in Hyb₄, since $d \leq m - n$. For $\mathbf{H}$ of rank d , the condition on $s_{\min}(\mathbf{S})$ in the statement allows us to apply Lemma 3.7: each $\mathbf{x}_j$ is within negligible distance of $D_{\Lambda_{\perp_{\mathbf{H}_j}}(\mathbf{H}), \mathbf{S}}$. Hence $\Delta(\text{Hyb}_4, \text{Hyb}_5) \leq \text{negl}(n)$.

Summing the bounds of the five transitions gives

$$\Delta(\text{Hyb}_0, \text{Hyb}_5) \leq 2i\varepsilon + \text{negl}(n). \quad \square$$

Lemma 5.5 (Public view of the reversed experiment). *Let $(\mathbf{A}, \mathbf{u}, \mathbf{H}, (\mathbf{h}_j)_{j \leq i}, (\mathbf{x}_j)_{j \leq i})$ be generated by **Reversed** (q, n, m, d, i) , and assume that $\mathbf{H}$ has rank d . There is an efficient algorithm that, given only $(\mathbf{A}, \mathbf{u}, \mathbf{H}, (\mathbf{h}_j)_{j \leq i})$, outputs $(\mathbf{h}_0, \mathbf{K}, (\mathbf{a}_j)_{j \leq i})$ with the same distribution as in **Reversed**, jointly with the whole output, secret keys included. In particular, $\mathbf{h}_0$ is uniform over $\mathcal{H}_{\mathbf{u}}$ given that output, so publishing $\mathbf{h}_0, \mathbf{K}$ and $(\mathbf{a}_j)_{j \leq i}$ reveals nothing beyond the public values and the projected keys.*

Proof. Since $\mathbf{H}$ has full column rank, $\mathbf{R}$ is the unique solution of $\mathbf{A} = \mathbf{H}\mathbf{R}$. Hence $(\mathbf{A}, \mathbf{u}, \mathbf{H})$ determine $\mathbf{R}$, the set $\mathcal{H}_{\mathbf{u}} = \{\mathbf{h} : \mathbf{R}^t \mathbf{h} = \mathbf{u}\}$, and the basis $\mathbf{K}$ of $\ker(\mathbf{R}^t)$, computed from $\mathbf{R}$ as in **Reversed**. Given $\mathbf{h}_0$, each $\mathbf{a}_j$ is the unique solution of $\mathbf{K}\mathbf{a}_j = \mathbf{h}_j - \mathbf{h}_0$. The algorithm therefore samples $\mathbf{h}_0 \leftarrow U(\mathcal{H}_{\mathbf{u}})$, computes $\mathbf{K}$ and solves for the $\mathbf{a}_j$; it only uses $(\mathbf{A}, \mathbf{u}, \mathbf{H}, (\mathbf{h}_j)_{j \leq i})$. Writing Y for the whole output of **Reversed**, it remains to show that $\mathbf{h}_0$ is uniform over $\mathcal{H}_{\mathbf{u}}$ given Y .

Fix $\mathbf{H}, \mathbf{T}_{\mathbf{H}}$ and $\mathbf{R}$. Since $\mathbf{R}$ has rank n , the map $\mathbf{h} \mapsto \mathbf{R}^t \mathbf{h}$ from $\mathbb{Z}_q^d$ to $\mathbb{Z}_q^n$ is surjective, and all its fibres have q^{d-n} elements. As $\mathbf{h}_0$ is uniform over $\mathbb{Z}_q^d$, the vector $\mathbf{u} = \mathbf{R}^t \mathbf{h}_0$ is uniform, and, given $\mathbf{u}$, the vector $\mathbf{h}_0$ is uniform over the fibre $\mathcal{H}_{\mathbf{u}}$. For every value of $\mathbf{h}_0$ in $\mathcal{H}_{\mathbf{u}}$, the projected keys $\mathbf{h}_j = \mathbf{h}_0 + \mathbf{K}\mathbf{a}_j$ are independent and uniform over $\mathcal{H}_{\mathbf{u}}$, since the $\mathbf{a}_j$ are uniform and $\mathbf{K}$ is a basis of $\ker(\mathbf{R}^t)$. Each $\mathbf{x}_j$ is then sampled from $D_{\Lambda_{\perp_{\mathbf{H}_j}}(\mathbf{H}), \mathbf{S}}$ with $\mathbf{T}_{\mathbf{H}}$. So the distribution of $((\mathbf{h}_j)_{j \leq i}, (\mathbf{x}_j)_{j \leq i})$ given $\mathbf{u}$ and $\mathbf{h}_0$ does not depend on $\mathbf{h}_0$, and $\mathbf{h}_0$ remains uniform over $\mathcal{H}_{\mathbf{u}}$ given these values. This conditional distribution only depends on $(\mathbf{R}, \mathbf{u})$, which Y determines. Hence $\mathbf{h}_0$ is uniform over $\mathcal{H}_{\mathbf{u}}$ given Y . $\square$

Proof of Theorem 5.3. By Lemma 3.10, $\|\tilde{\mathbf{T}}_{\mathbf{H}}\| \leq \|\mathbf{T}_{\mathbf{H}}\| \leq 3mq^{d/m}$, so the condition on $s_{\min}(\mathbf{S})$ implies that of Lemma 5.4. Moreover, $\mathbf{H}$ has full column rank except with negligible probability.

Correctness and soundness. In the definition, the public values are generated by **TrapGen** and the keys by $p \leftarrow D_{\mathcal{J}_i}$ and $k \leftarrow \mathcal{S}(t, p)$: this is exactly **Reversed**. Theorem 4.2 proves the indistinguishability claims for **Regular** for every $i \geq 0$. For $i \geq 1$ and independent keys k and k' , Theorem 4.2 also provides a distinguisher between $F_{i,k}$ and $P_{i,\pi_i(k')}$ that fails with negligible probability. By Lemma 5.4 and the triangle inequality, these claims transfer to **Reversed**; for the distinguisher, we apply the lemma to the $2i$ keys of k and k' . By Lemma 5.5, they still hold when $\mathbf{h}_0, \mathbf{K}$ and $(\mathbf{a}_j)_{j \leq i}$ are added to the public view. Finally, the soundness condition of the definition requires $p \neq p'$. Since p and p' are independent and uniform over $\mathcal{H}_{\mathbf{u}}^i$, with $|\mathcal{H}_{\mathbf{u}}| \geq 2$, this event has probability at least $1/2$, so conditioning on it at most doubles the failure probability of the distinguisher.

Projected keys and preimage sampling. Fix public values for which $\mathbf{H}$ has full column rank. As shown in the transition to Hyb₂ in the proof of Lemma 5.4, a key $\mathbf{x} \leftarrow D_{\Lambda_{\perp_{\mathbf{u}}}(\mathbf{A}), \mathbf{S}}$ has its projected key $-\mathbf{H}^t \mathbf{x}$ distributed according to $\mu_{\mathcal{H}_{\mathbf{u}}}$, and, conditioned on $-\mathbf{H}^t \mathbf{x} = \mathbf{h}$, it follows $D_{\Lambda_{\perp_{\mathbf{h}}}(\mathbf{H}), \mathbf{S}}$. By the transition to Hyb₃, since $s_{\min}(\mathbf{S}) \geq \eta_{\varepsilon}(\Lambda^{\perp}(\mathbf{H}))$, the distribution $\mu_{\mathcal{H}_{\mathbf{u}}}$ is within distance 2ε of $U(\mathcal{H}_{\mathbf{u}})$. The i keys of $\mathcal{D}_i$ are independent, so $\pi_i(\mathcal{D}_i)$ is within distance $2i\varepsilon$ of $D_{\mathcal{J}_i}$; moreover, every

$p \in \mathcal{H}_{\mathbf{u}}^i$ has positive probability, since the fibres are non-empty. This is the projected-key condition. Conditioned on $\pi_i(k) = p$, the distribution $\mathcal{D}_i$ is the product of the Gaussians $D_{\Lambda_{\mathbf{h}_j}^\perp(\mathbf{H}), \mathbf{S}}$, and the sampler $\mathcal{S}(t, p)$ samples each $\mathbf{x}_j$ with Lemma 3.7, hence within negligible distance of this product. This is the preimage-sampling condition.

Keys. By these two conditions, a pair $(k, \pi_i(k))$ with $k \leftarrow \mathcal{D}_i$ is statistically close to a pair (k, p) with $p \leftarrow D_{\mathcal{J}_i}$ and $k \leftarrow \mathcal{S}(t, p)$. The indistinguishability claims proved above for the keys of **Reversed** therefore also hold for $k \leftarrow \mathcal{D}_i$, as stated. $\square$

Concrete parameter regime. We now fix concrete parameters, with $d = \Theta(n)$. The choice of these parameters, and in particular of the Gaussian parameter, is discussed at the end of this section.

Theorem 5.6 (Practical trapdoor projective sampling from k -LWE). *Let n be the security parameter, let q be prime, and let d, m be integers such that*

$$d = \Theta(n), \quad d \geq 2n, \quad m = \Omega(n \log q), \quad m \geq 2d.$$

Let $\mathbf{S}$ be a positive diagonal matrix such that

$$s_{\min}(\mathbf{S}) \geq m \cdot \omega(\sqrt{\log m}).$$

Then, with overwhelming probability over the generation of $\mathbf{H}$, the conditions on $s_{\min}(\mathbf{S})$ of Theorem 5.3 and of Lemma 5.4 hold. Consequently, for every $i \geq 1$, under the $(i, \mathbf{S})$ -LWE $_{\alpha, m}$ hardness assumption, the family TPSam_i is a trapdoor projective sampling family. Moreover, for every $i \geq 0$, the secret-key and projected samplers are both computationally indistinguishable from the common reference distribution

$$\frac{1}{q} U(\text{Im}(\mathbf{A}^+)) + \nu_\alpha^{m+1}.$$

For every $i \geq 1$, the distribution $D_{\mathcal{J}_i}$ has negligible collision probability, and therefore Proposition 5.2 gives the induced underlying projective sampling family.

Proof. The conditions imply $n < d \leq m - n$, as required by Theorem 5.3. We check its condition on $s_{\min}(\mathbf{S})$, and that of Lemma 5.4. Since $d = O(n)$ and $m = \Omega(n \log q)$, the factor $q^{d/m} = 2^{d \log q / m}$ is bounded by a constant, so $m \cdot \omega(\sqrt{\log m})$ dominates the GPV term $3mq^{d/m} \cdot \omega(\sqrt{\log m})$ of Theorem 5.3. By Lemma 3.10, this term is at least the GPV term $\|\tilde{\mathbf{T}}_{\mathbf{H}}\| \cdot \omega(\sqrt{\log m})$ of Lemma 5.4.

For the smoothing term, let $\varepsilon = 2^{-\Omega(n)}$. Since $m \geq 2d$, Lemma 3.2 shows that, except for a fraction $2^{-\Omega(d)} = 2^{-\Omega(n)}$ of matrices $\mathbf{H} \in \mathbb{Z}_q^{m \times d}$,

$$\eta_\varepsilon(\Lambda^\perp(\mathbf{H})) \leq O\left(q^{d/m} \sqrt{\log(m/\varepsilon)}\right) = O(\sqrt{n}).$$

Since the trapdoor-generated $\mathbf{H}$ is statistically close to uniform by Lemma 3.10, the same bound holds for it, up to the AP11 statistical distance. This bound is dominated by $m \cdot \omega(\sqrt{\log m})$, since $m \geq 2d = \Omega(n)$. Hence both conditions hold, and Theorem 5.3 gives the TPSam and common-reference claims.

Finally, for $i \geq 1$, two independent samples from $D_{\mathcal{J}_i} = U(\mathcal{H}_{\mathbf{u}}^i)$ collide with probability $|\mathcal{H}_{\mathbf{u}}|^i = q^{-(d-n)i} \leq q^{-n}$, since $d \geq 2n$, and this is negligible. Proposition 5.2 therefore gives the induced projective sampling family. $\square$

Discussion of the parameters. The trapdoor construction of this section changes the order in which projected and private keys are sampled, but it does not change the asymptotic parameter

regime of the LPSS/DPY construction. The only point to check is the Gaussian parameter used for the private-key sampler.

There are two constraints on this parameter. The first one comes from the uniformity condition on the projected keys in the projective-sampling argument:

$$s_{\min}(\mathbf{S}) \geq \eta_\varepsilon(\Lambda^\perp(\mathbf{H})).$$

The second one comes from the GPV step inside the coset sampler. Indeed, when $\mathbf{H}$ is generated by Lemma 3.10, the algorithm also outputs a basis $\mathbf{T}_\mathbf{H}$ of $\Lambda^\perp(\mathbf{H})$ satisfying

$$\|\tilde{\mathbf{T}}_\mathbf{H}\| \leq \|\mathbf{T}_\mathbf{H}\| \leq 3mq^{d/m}.$$

Since GPV requires

$$s_{\min}(\mathbf{S}) \geq \|\tilde{\mathbf{T}}_\mathbf{H}\| \cdot \omega(\sqrt{\log m}),$$

it is sufficient to impose

$$s_{\min}(\mathbf{S}) \geq 3mq^{d/m} \cdot \omega(\sqrt{\log m}).$$

It remains to compare this GPV condition with the smoothing condition. For a uniformly random $\mathbf{H} \in \mathbb{Z}_q^{m \times d}$, Lemma 3.2 gives, except with probability $2^{-\Omega(d)}$,

$$\eta_\varepsilon(\Lambda^\perp(\mathbf{H})) \leq O\left(q^{d/m} \sqrt{\log(m/\varepsilon)}\right).$$

Since the AP11-generated matrix $\mathbf{H}$ is statistically close to uniform, the same estimate applies in the trapdoor experiment, up to the AP11 statistical distance. Therefore, in the regime $d = O(n)$ and $m = \Omega(n \log q)$, the GPV threshold $3mq^{d/m} \cdot \omega(\sqrt{\log m})$ dominates the smoothing bound. Thus choosing $s_{\min}(\mathbf{S})$ large enough for GPV sampling also satisfies the smoothing condition of Lemma 5.4.

Choosing $d = \Theta(n)$ rather than the LPSS value $d = m - n$ does not change the security argument: the projective-sampling theorem works throughout the range $n < d \leq m - n$, and $d = m - n$ is only the maximal choice in that range. The reason for this choice is instead to keep the trapdoor matrix $\mathbf{H} \in \mathbb{Z}_q^{m \times d}$ at a width comparable to the matrices used in the GPV sampling step of the anonymous-broadcast application. In that setting, a matrix in $\mathbb{Z}_q^{m \times n}$ gives the GPV threshold $mq^{n/m} \cdot \omega(\sqrt{\log m})$ up to constants, while our matrix of width d gives $mq^{d/m} \cdot \omega(\sqrt{\log m})$. Since $d = \Theta(n)$, the ratio

$$\frac{q^{d/m}}{q^{n/m}} = q^{(d-n)/m}$$

is bounded by a constant in the standard scale $m = \Omega(n \log q)$ used by those constructions. Thus the trapdoor projective-sampling modification does not force us to increase m , q , or the Gaussian parameter beyond the usual order.

6 From Public-Key to Identity-Based Broadcast

Trapdoor projective sampling lets the authority choose a projected key first and then extract a compatible secret key. We use it to turn a projective public-key broadcast scheme into an identity-based one, in the bounded-universe model of [12]: at most N users, fixed at setup, whom the adversary may corrupt adaptively. The public encryption key contains all the public information needed to encrypt. In a public-key scheme, this includes the users' public keys; in an identity-based scheme, it may instead include a public map, typically a random oracle, from identities to public

values. The public key may then be independent of the user set, which is used only to extract keys, and reveal neither this set nor its cardinality. The unbounded universe is treated in Section 8.

6.1 Broadcast in a Bounded Universe

Definition 6.1 (Broadcast for a bounded universe). Let $\mathcal{I}$ be an identity space, let PT and CT be plaintext and ciphertext spaces, and let N be the maximal number of users. A broadcast scheme for a bounded universe is a tuple $\Pi = (\text{Setup}, \text{Extract}, \text{Enc}, \text{Dec})$ with the following syntax:

- $\text{Setup}(1^n, \mathcal{U})$ takes a user set $\mathcal{U} \subseteq \mathcal{I}$ with $|\mathcal{U}| \leq N$, and outputs a public encryption key ek and a master secret key MSK ;
- $\text{Extract}(\text{ek}, \text{MSK}, id)$ outputs a secret key sk_{id} if $id \in \mathcal{U}$, and $\perp$ otherwise;
- $\text{Enc}(\text{ek}, M, \mathcal{S})$ takes $M \in \text{PT}$ and a target set $\mathcal{S} \subseteq \mathcal{U}$, and outputs $\text{ct} \in \text{CT}$;
- $\text{Dec}(\text{ek}, \text{sk}_{id}, \text{ct})$ outputs a message or $\perp$.

Correctness requires that $\text{Dec}(\text{ek}, \text{sk}_{id}, \text{Enc}(\text{ek}, M, \mathcal{S})) = M$ with overwhelming probability over the randomness of the algorithms, for every $M \in \text{PT}$, every $\mathcal{S} \subseteq \mathcal{U}$, every $id \in \mathcal{S}$, and every sk_{id} output by $\text{Extract}(\text{ek}, \text{MSK}, id)$.

A public-key broadcast scheme for a bounded universe, as in [12], is the case $\mathcal{I} = \mathcal{U} = [N]$; we then write $\text{Setup}(1^n, N)$ and index users by $i \in [N]$.

Definition 6.2 (Projective public-key broadcast). Let $\text{PSam} = (\mathcal{F}, \mathcal{K}, X, \mathcal{P}, \mathcal{J}, \pi, D_{\mathcal{K}})$ be a projective sampling family. A public-key broadcast scheme for a bounded universe $\Pi = (\text{Setup}, \text{Extract}, \text{Enc}, \text{Dec})$ is projective with respect to PSam if each user $i \in [N]$ has a secret key $\text{sk}_i \leftarrow D_{\mathcal{K}}$ and a public key $\text{pk}_i = \pi(\text{sk}_i) \in \mathcal{J}$, and if the scheme has the following form:

- the public encryption key ek contains the public keys $(\text{pk}_i)_{i \in [N]}$, together with any global public parameters;
- the master secret key contains the corresponding secret keys $(\text{sk}_i)_{i \in [N]}$, together with any global secret parameters;
- encryption to a target set $\mathcal{S}$ uses the user public keys only through the projected samplers P_{pk_i} , for $i \in \mathcal{S}$;
- decryption for user i uses the corresponding secret key sk_i .

Thus a projective public-key broadcast scheme is a public-key broadcast scheme whose user public keys are projections of secret keys and whose encryption algorithm uses the corresponding projected samplers.

We use the IND-CPA and ANO-CPA notions of [12] for broadcast encryption with adaptive corruptions, which are a chosen-plaintext version of the anonymity notion of [20], without its requirement that the two target sets have the same size (Definitions 6.3 and 6.4 below). The adversary chooses the user set, then receives the public key and corrupts users adaptively. In the IND-CPA experiment, it chooses two messages and a target set containing no corrupted user. In the ANO-CPA experiment, it chooses a message and two target sets, and may corrupt every user outside their symmetric difference, including users that receive the message in both cases. We write $\text{Adv}_{\Pi, \mathcal{A}}^{\text{ind-cpa}}(n, N)$ and $\text{Adv}_{\Pi, \mathcal{A}}^{\text{ano-cpa}}(n, N)$ for the advantages of an adversary $\mathcal{A}$.

Definition 6.3 (IND-CPA security). Let $\Pi = (\text{Setup}, \text{Extract}, \text{Enc}, \text{Dec})$ be a broadcast scheme for a bounded universe of at most N users. The IND-CPA experiment

$$\text{Exp}_{\Pi, \mathcal{A}}^{\text{ind-cpa}}(1^n, N)$$

is defined as follows.

1. The adversary outputs a user set $\mathcal{U} \subseteq \mathcal{I}$ with $|\mathcal{U}| \leq N$. For a public-key scheme, $\mathcal{U} = [N]$.
2. The challenger runs $(\text{ek}, \text{MSK}) \leftarrow \text{Setup}(1^n, \mathcal{U})$ and gives ek to $\mathcal{A}$.
3. The adversary adaptively queries secret keys for identities of $\mathcal{U}$. On query id , the challenger returns the stored key if id was already queried; otherwise it samples $\text{sk}_{id} \leftarrow \text{Extract}(\text{ek}, \text{MSK}, id)$, stores it, and returns it.
4. The adversary outputs two messages $M_0, M_1 \in \text{PT}$ and a target set $\mathcal{S} \subseteq \mathcal{U}$ containing no queried identity. The challenger samples $b \leftarrow U(\{0, 1\})$ and returns

$$\text{ct} \leftarrow \text{Enc}(\text{ek}, M_b, \mathcal{S}).$$

5. The adversary continues to query secret keys for identities of $\mathcal{U}$ outside $\mathcal{S}$.
6. The adversary outputs a bit b' . The experiment outputs 1 if $b' = b$, and 0 otherwise.

We set

$$\text{Adv}_{\Pi, \mathcal{A}}^{\text{ind-cpa}}(n, N) = \left| \Pr \left[\text{Exp}_{\Pi, \mathcal{A}}^{\text{ind-cpa}}(1^n, N) = 1 \right] - \frac{1}{2} \right|.$$

The scheme is IND-CPA secure if this advantage is negligible for every probabilistic polynomial-time adversary $\mathcal{A}$.

Definition 6.4 (ANO-CPA security). Let $\Pi = (\text{Setup}, \text{Extract}, \text{Enc}, \text{Dec})$ be a broadcast scheme for a bounded universe of at most N users. The ANO-CPA experiment

$$\text{Exp}_{\Pi, \mathcal{A}}^{\text{ano-cpa}}(1^n, N)$$

is defined as follows.

1. The adversary outputs a user set $\mathcal{U} \subseteq \mathcal{I}$ with $|\mathcal{U}| \leq N$. For a public-key scheme, $\mathcal{U} = [N]$.
2. The challenger runs $(\text{ek}, \text{MSK}) \leftarrow \text{Setup}(1^n, \mathcal{U})$ and gives ek to $\mathcal{A}$.
3. The adversary adaptively queries secret keys for identities of $\mathcal{U}$. On query id , the challenger returns the stored key if id was already queried; otherwise it samples $\text{sk}_{id} \leftarrow \text{Extract}(\text{ek}, \text{MSK}, id)$, stores it, and returns it.
4. The adversary outputs a message $M \in \text{PT}$ and two target sets $\mathcal{S}_0, \mathcal{S}_1 \subseteq \mathcal{U}$ such that no queried identity lies in $\mathcal{S}_0 \triangle \mathcal{S}_1$. The challenger samples $b \leftarrow U(\{0, 1\})$ and returns

$$\text{ct} \leftarrow \text{Enc}(\text{ek}, M, \mathcal{S}_b).$$

5. The adversary continues to query secret keys for identities of $\mathcal{U}$ outside $\mathcal{S}_0 \triangle \mathcal{S}_1$.
6. The adversary outputs a bit b' . The experiment outputs 1 if $b' = b$, and 0 otherwise.

We set

$$\text{Adv}_{\Pi, \mathcal{A}}^{\text{ano-cpa}}(n, N) = \left| \Pr \left[\text{Exp}_{\Pi, \mathcal{A}}^{\text{ano-cpa}}(1^n, N) = 1 \right] - \frac{1}{2} \right|.$$

The scheme is ANO-CPA secure if this advantage is negligible for every probabilistic polynomial-time adversary $\mathcal{A}$.

6.2 From Projected Public Keys to Identities

Setting. Throughout this subsection, we fix:

- (S1) a trapdoor projective sampling family TPSam , whose projected-key distribution $D_{\mathcal{J}}$ has negligible collision probability and can be sampled efficiently from the public parameters;
- (S2) the projective sampling family $\text{PSam} = (\mathcal{F}, \mathcal{K}, X, \mathcal{P}, \mathcal{J}, \pi, D_{\mathcal{K}})$ induced by TPSam , as in Proposition 5.2 (its global public parameters are thus those output by TrapGen , with the trapdoor forgotten);

(S3) a public-key broadcast scheme $\Pi_{\text{pk}} = (\text{Setup}_{\text{pk}}, \text{Extract}_{\text{pk}}, \text{Enc}_{\text{pk}}, \text{Dec}_{\text{pk}})$ for a bounded universe of N users, projective with respect to PSam (Definition 6.2).

Construction 6.5 (Identity-based compilation). Let $\text{IdMap} : \mathcal{I} \rightarrow \mathcal{J}$ be a random oracle whose outputs are distributed according to $D_{\mathcal{J}}$. The identity-based broadcast scheme for a bounded universe $\Pi_{\text{id}} = (\text{Setup}_{\text{id}}, \text{Extract}_{\text{id}}, \text{Enc}_{\text{id}}, \text{Dec}_{\text{id}})$ is defined as follows.

- $\text{Setup}_{\text{id}}(1^n, \mathcal{U})$ runs TrapGen , which outputs the global public parameters and a trapdoor t . It outputs ek , made of the global public parameters and IdMap , and $\text{MSK} = (t, \mathcal{U})$. No user key is generated, and ek does not depend on $\mathcal{U}$.
- $\text{Extract}_{\text{id}}(\text{ek}, \text{MSK}, id)$ returns $\perp$ if $id \notin \mathcal{U}$. Otherwise it computes $p_{id} = \text{IdMap}(id)$ and returns $\text{sk}_{id} \leftarrow \mathcal{S}(t, p_{id})$, stored for later queries.
- $\text{Enc}_{\text{id}}(\text{ek}, M, \mathcal{S})$, for $\mathcal{S} \subseteq \mathcal{U}$, computes $(p_{id})_{id \in \mathcal{S}}$ and runs the encryption algorithm of Π_{pk} with these values as the recipients' public keys; since Π_{pk} is projective, this only involves the projected samplers $P_{p_{id}}$.
- Dec_{id} is the decryption algorithm of Π_{pk} , run with the extracted key.

Theorem 6.6 (Public-key broadcast to identity-based broadcast). Assume (S1)–(S3), and model IdMap as a random oracle. If Π_{pk} is correct, then so is the scheme Π_{id} of Construction 6.5. Moreover, let $\mathbf{x} \in \{\text{ind-cpa}, \text{ano-cpa}\}$. For every probabilistic polynomial-time adversary $\mathcal{A}$ against Π_{id} in the experiment $\mathbf{x}$, there is a probabilistic polynomial-time adversary $\mathcal{B}$ against Π_{pk} in the same experiment, which corrupts the users corresponding to those corrupted by $\mathcal{A}$, at the same moments, such that

$$\text{Adv}_{\Pi_{\text{id}}, \mathcal{A}}^{\mathbf{x}}(n, N) \leq \text{Adv}_{\Pi_{\text{pk}}, \mathcal{B}}^{\mathbf{x}}(n, N) + \text{negl}(n).$$

In particular, Π_{id} is IND-CPA (resp. ANO-CPA) secure whenever Π_{pk} is.

Corollary 6.7 (Other parameter generations). Theorem 6.6 still holds when (S3) is relaxed as follows: Π_{pk} is projective with respect to a family that defines the same objects as PSam for every value of the public parameters, but samples these parameters by another algorithm, whose output is statistically close to that of TrapGen with the trapdoor forgotten.

Proof of Theorem 6.6. We treat both experiments together. Let $\mathcal{A}$ be an adversary against Π_{id} . It first outputs a user set $\mathcal{U} = \{id_1, \dots, id_u\}$ with $u \leq N$, which we identify with the indices $1, \dots, u$ of Π_{pk} ; the indices beyond u are never used. Since IdMap is part of ek , the adversary queries it only after outputting $\mathcal{U}$.

The adversary $\mathcal{B}$. It plays the experiment $\mathbf{x}$ of Π_{pk} , whose user set is $[N]$, as for every public-key scheme.

- $\mathcal{B}$ receives the global public parameters and $(\text{pk}_i)_{i \in [N]}$. It runs $\mathcal{A}$, which outputs $\mathcal{U}$, then gives $\mathcal{A}$ these parameters as ek and simulates IdMap : it sets $\text{IdMap}(id_i) := \text{pk}_i$ for $i \leq u$, and answers the other queries with fresh samples of $D_{\mathcal{J}}$.
- It forwards every extraction query on id_i to its own key oracle on index i , and returns $\perp$ outside $\mathcal{U}$.
- It translates the challenge target sets through $id_i \mapsto i$, forwards them with the messages, relays the challenge ciphertext, and outputs the bit of $\mathcal{A}$.

The translation is a bijection on $\mathcal{U}$, so the restrictions of the experiment are preserved, and $\mathcal{B}$ corrupts user i exactly when $\mathcal{A}$ corrupts id_i .

The games. Both games output 1 when $\mathcal{A}$ guesses the challenge bit.

- G_0 is the experiment $\mathbf{x}$ of Π_{id} with $\mathcal{A}$, so that $\text{Adv}_{\Pi_{\text{id}}, \mathcal{A}}^{\mathbf{x}}(n, N) = |\Pr[G_0 = 1] - 1/2|$.
- G_1 is the experiment $\mathbf{x}$ of Π_{pk} with $\mathcal{B}$. Since $\mathcal{B}$ outputs the bit of $\mathcal{A}$, $\text{Adv}_{\Pi_{\text{pk}}, \mathcal{B}}^{\mathbf{x}}(n, N) = |\Pr[G_1 = 1] - 1/2|$.

It therefore suffices to show that $|\Pr[G_0 = 1] - \Pr[G_1 = 1]| \leq \text{negl}(n)$.

From G_0 to G_1 . We compare the view of $\mathcal{A}$ in the two games.

- *Public parameters.* In G_0 , they are output by TrapGen . In G_1 , they are those of Π_{pk} , that is, by (S2) and (S3), the output of TrapGen with the trapdoor forgotten.
- *Identities outside $\mathcal{U}$.* In both games, ldMap answers with fresh samples of $D_{\mathcal{J}}$, and extraction returns $\perp$.
- *Identities of $\mathcal{U}$.* In G_0 , the values $p_{id_i} = \text{ldMap}(id_i)$ and the keys $\text{sk}_{id_i} \leftarrow \mathcal{S}(t, p_{id_i})$ are sampled on demand, with $p_{id_i} \leftarrow D_{\mathcal{J}}$; sampling them at setup instead does not change their distribution. In G_1 , the corresponding pair is $(\text{pk}_i, \text{sk}_i) = (\pi(\text{sk}_i), \text{sk}_i)$ with $\text{sk}_i \leftarrow D_{\mathcal{K}}$. By the projected-key and preimage-sampling conditions of the TPSam definition, the two pairs are statistically close. The $u \leq N$ pairs are independent, so replacing them one at a time costs a negligible statistical distance.
- *Challenge.* In both games, the challenge ciphertext is computed by the encryption algorithm of Π_{pk} from the values $\text{ldMap}(id_i)$ of the recipients: in G_0 by definition of Enc_{id} , and in G_1 because Π_{pk} is projective, so that its encryption depends on the users' public keys only through those of the recipients.

Hence the views of $\mathcal{A}$ in G_0 and G_1 are within negligible statistical distance, and $|\Pr[G_0 = 1] - \Pr[G_1 = 1]| \leq \text{negl}(n)$.

Correctness. By the same comparison, the key pairs of Π_{id} are statistically close to those of Π_{pk} , for which decryption succeeds with overwhelming probability. $\square$

Proof of Corollary 6.7. Let Π'_{pk} be Π_{pk} with its public parameters generated by TrapGen , the trapdoor being forgotten. It is projective with respect to PSam , so it satisfies (S3), and it has the same encryption and decryption algorithms as Π_{pk} , so both lead to the same scheme Π_{id} . Since the two parameter distributions are statistically close, Π'_{pk} is correct, IND-CPA secure, respectively ANO-CPA secure, whenever Π_{pk} is. Theorem 6.6 applied to Π'_{pk} gives the claim. $\square$

7 Identity-Based Anonymous Broadcast for a Bounded Universe

We now instantiate the compiler of Section 6.2 with the lattice trapdoor projective sampling family of Theorem 5.3. We first describe the public-key scheme Π_{pk} : it is the anonymous broadcast scheme of [12], with a projection matrix of width $d = 2n$ instead of $m - n$. We prove its correctness and its security for a bounded universe of N users with adaptive corruptions, and then obtain the identity-based scheme Π_{id} from Corollary 6.7.

Throughout this section, let the parameters $q, n, m, \mathbf{S}$ be as in Theorem 5.6 with $d = 2n$, fix a constant $\gamma \in (0, 1)$ and assume $N \leq (1 - \gamma)n$, together with the correctness condition $\alpha \leq 1/(16ms_{\max})$.

7.1 The Public-Key Scheme

The public-key scheme samples its public parameters in the regular order of [21, 12].

$\text{Setup}_{\text{pk}}(1^n, N)$: Generate $(\mathbf{A}, \mathbf{T}_{\mathbf{A}}) \in \mathbb{Z}_q^{m \times n} \times \mathbb{Z}^{m \times m}$ using Lemma 3.10, sample $\mathbf{u} \leftarrow U(\mathbb{Z}_q^n)$, and set $\mathbf{A}^+ = (\mathbf{u}^t || \mathbf{A})$. Sample $\mathbf{H} \leftarrow U(\mathbb{Z}_q^{m \times 2n})$, conditioned on $\text{rank}(\mathbf{H}) = 2n$ and $\text{Im}(\mathbf{A}) \subseteq \text{Im}(\mathbf{H})$. These are the global public parameters. For every $i \in [N]$, sample $\mathbf{x}_i$ statistically close to $D_{\Lambda_{\perp_{\mathbf{u}}}(\mathbf{A}), \mathbf{S}}$ using Lemma 3.7 with the basis $\mathbf{T}_{\mathbf{A}}$, and set $\mathbf{h}_i = -\mathbf{H}^t \mathbf{x}_i$ and $\mathbf{H}_i = (\mathbf{h}_i^t || \mathbf{H})$. Output $\text{ek} = (\mathbf{A}^+, \mathbf{H}, (\mathbf{h}_i)_{i \in [N]})$ and $\text{MSK} = (\mathbf{x}_i)_{i \in [N]}$.

$\text{Extract}_{\text{pk}}(\text{ek}, \text{MSK}, i)$: Return $\mathbf{x}_i^+ = \begin{pmatrix} 1 \\ \mathbf{x}_i \end{pmatrix}$.

$\text{Encrypt}_{\text{pk}}(\text{ek}, M, \mathcal{S})$: Given a message $M \in \{0, 1\}$ and a target set $\mathcal{S} \subseteq [N]$, sample $\mathbf{y} \leftarrow U(\bigcap_{i \in \mathcal{S}} \text{Im}(\mathbf{H}_i))$, with the convention $\bigcap_{i \in \emptyset} \text{Im}(\mathbf{H}_i) = \mathbb{Z}_q^{m+1}$, and $\mathbf{e} \leftarrow \lfloor \nu_{\alpha q} \rfloor^{m+1}$, and output the ciphertext $\mathbf{c} = \mathbf{y} + \mathbf{e} + \begin{pmatrix} M \lfloor q/2 \rfloor \\ \mathbf{0}_m \end{pmatrix} \in \mathbb{Z}_q^{m+1}$.

$\text{Decrypt}_{\text{pk}}(\text{ek}, \mathbf{x}_i^+, \mathbf{c})$: Return 0 if $\langle \mathbf{x}_i^+, \mathbf{c} \rangle$ is closer to 0 than to $\lfloor q/2 \rfloor$ modulo q , and return 1 otherwise.

The scheme is projective, in the sense of Definition 6.2, with respect to the projective sampling family of Theorem 4.2 in dimension $d = 2n$, with $\mathbf{H}$ viewed as a public parameter: the global public parameters are $(\mathbf{A}, \mathbf{u}, \mathbf{H})$, the secret keys are the vectors $\mathbf{x}_i$, the projection is $\mathbf{x} \mapsto -\mathbf{H}^t \mathbf{x}$, and encryption uses the public keys only through the projected samplers $U(\bigcap_{i \in \mathcal{S}} \text{Im}(\mathbf{H}_i))$. The Gaussian parameter satisfies the trapdoor-sampling condition for $\mathbf{A}$, since $\|\mathbf{T}_{\mathbf{A}}\| \leq 3mq^{n/m} \leq 3mq^{d/m}$.

Theorem 7.1 (Correctness). *For every message $M \in \{0, 1\}$, every target set $\mathcal{S} \subseteq [N]$ and every $i \in \mathcal{S}$,*

$$\text{Dec}_{\text{pk}}(\text{ek}, \mathbf{x}_i^+, \text{Enc}_{\text{pk}}(\text{ek}, M, \mathcal{S})) = M$$

with probability at least $1 - 2^{-\Omega(n)}$ over the randomness of Setup_{pk} and Enc_{pk} .

Proof of Theorem 7.1. Let $\mathbf{c}$ be an honestly generated ciphertext for M and $\mathcal{S}$, and let $i \in \mathcal{S}$. By construction,

$$\mathbf{c} = \mathbf{y} + \mathbf{e} + \begin{pmatrix} M \lfloor q/2 \rfloor \\ \mathbf{0}_m \end{pmatrix}, \quad \mathbf{y} \in \bigcap_{i' \in \mathcal{S}} \text{Im}(\mathbf{H}_{i'}),$$

so in particular $\mathbf{y} \in \text{Im}(\mathbf{H}_i)$. Since $\mathbf{h}_i = -\mathbf{H}^t \mathbf{x}_i$,

$$\mathbf{H}_i^t \mathbf{x}_i^+ = \begin{pmatrix} \mathbf{h}_i^t \\ \mathbf{H} \end{pmatrix}^t \begin{pmatrix} 1 \\ \mathbf{x}_i \end{pmatrix} = \mathbf{h}_i + \mathbf{H}^t \mathbf{x}_i = \mathbf{0} \pmod{q}.$$

Therefore $\langle \mathbf{x}_i^+, \mathbf{y} \rangle = 0 \pmod{q}$, and

$$\langle \mathbf{x}_i^+, \mathbf{c} \rangle = \langle \mathbf{x}_i^+, \mathbf{e} \rangle + M \lfloor q/2 \rfloor \pmod{q}.$$

It remains to bound the noise term. By Lemma 3.4, applied to the sampler of Lemma 3.7, we have

$$\|\mathbf{x}_i^+\| \leq 3\sqrt{m} s_{\max}$$

except with probability $2^{-\Omega(n)}$. Conditioned on this event, the usual Gaussian tail bound gives

$$|\langle \mathbf{x}_i^+, \mathbf{e} \rangle| \leq 4m\alpha q s_{\max} \leq q/4$$

except with probability $2^{-\Omega(n)}$. Thus $\langle \mathbf{x}_i^+, \mathbf{c} \rangle$ is closer to 0 than to $\lfloor q/2 \rfloor$ when $M = 0$, and closer to $\lfloor q/2 \rfloor$ when $M = 1$. $\square$

7.2 Removing One Recipient

For a target set $\mathcal{T} \subseteq [N]$, the noiseless part of a ciphertext is uniform on the *ciphertext space* $W_{\mathcal{T}} = \bigcap_{i \in \mathcal{T}} \text{Im}(\mathbf{H}_i)$, with $W_{\emptyset} = \mathbb{Z}_q^{m+1}$. Correctness rests on the fact that the key of every recipient is orthogonal to $W_{\mathcal{T}}$; security needs a converse for the keys held by the adversary.

For $\mathcal{T} \subseteq [N]$ and a bit M , write

$$\mathcal{E}_{\mathcal{T}}(M) = U(W_{\mathcal{T}}) + \lfloor \nu_{\alpha q} \rfloor^{m+1} + \begin{pmatrix} M \lfloor q/2 \rfloor \\ \mathbf{0}_m \end{pmatrix}$$

for the distribution of a ciphertext of M for $\mathcal{T}$, a deterministic post-processing of $\frac{1}{q}U(W_{\mathcal{T}}) + \nu_{\alpha}^{m+1}$. For a set $\mathcal{C} \subseteq [N]$ of users whose secret keys are held, write also $V_{\mathcal{C}} = \text{Span}((\mathbf{x}_i^+)_{i \in \mathcal{C}})^{\perp}$, with $V_{\emptyset} = \mathbb{Z}_q^{m+1}$, for the space of vectors orthogonal to all the held keys.

Proposition 7.2 (Removing one recipient, with all other keys). *Let $j \in [N]$. Under the $(N-1, \mathbf{S})$ -LWE $_{\alpha, m}$ hardness assumption, given the public key of Π_{pk} and the secret keys $(\mathbf{x}_i^+)_{i \neq j}$ of all the other users, the distributions*

$$\mathcal{E}_{\mathcal{T}}(M) \quad \text{and} \quad \mathcal{E}_{\mathcal{T} \setminus \{j\}}(M)$$

are computationally indistinguishable, for every target set $\mathcal{T} \subseteq [N]$ containing j and every bit M , which may be chosen after seeing the public key and the keys $(\mathbf{x}_i^+)_{i \neq j}$.

Remark 7.3 (Relation to the one-more lemma). Proposition 7.2 generalizes the “one more projected hint” transition of [12, Lemma 5]. There, the distinguisher holds only the keys of the recipients other than j ; here, it holds the keys of all the users other than j , recipients or not (Figure 1). This generalization addresses the slight gap in the proof of [12, Theorem 6] mentioned in the introduction.

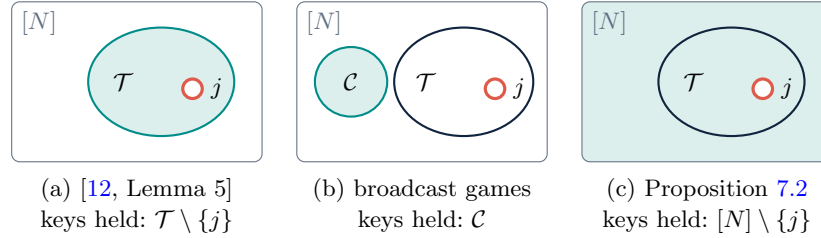


Fig. 1. Secret keys held by the distinguisher (shaded) in the transition that removes the recipient j from the target set $\mathcal{T}$. Each box is the set $[N]$ of users, and the key of j is never held (white dot). (a) In [12, Lemma 5] and [21, Theorem 25], the held keys are those of the other recipients. (b) In the broadcast games, the adversary holds the keys of the corrupted users $\mathcal{C}$, who are not recipients. This case is not covered by (a). (c) Proposition 7.2 grants every key but that of j . Since a distinguisher may ignore keys, (c) covers both (a) and (b).

Idea of the proof of Proposition 7.2. The proof generalizes that of [21, Theorem 27], which adds to a k -LWE challenge an independent uniform element of the ciphertext space. The sum of two independent variables, uniform on two subgroups, is uniform on the sum of these subgroups, so this operation acts on supports as a sum of subspaces. In [21], the held keys are exactly those of the recipients, and the inclusions $\text{Im}(\mathbf{A}^+) \subseteq W_{\mathcal{T}} \subseteq V_{\mathcal{T}}$ make both sums trivial. Here the held keys also belong to non-recipients, so that $W_{\mathcal{T}} \not\subseteq V_{[N] \setminus \{j\}}$ in general, and Lemma 7.4 computes the sum. Figures 1 and 2 illustrate these two points.

Lemma 7.4 (Subspace sum). *Let $j \in [N]$ and $\mathcal{R} = [N] \setminus \{j\}$. Assume that $\mathbf{H}$ has full column rank, that $\mathbf{H}^t \mathbf{x}_i = -\mathbf{h}_i$ for every $i \in \mathcal{R}$, and that the projected keys $(\mathbf{h}_i)_{i \in [N]}$ are affinely independent. Then, for every non-empty $\mathcal{T} \subseteq [N]$,*

$$V_{\mathcal{R}} + W_{\mathcal{T}} = V_{\mathcal{R} \cap \mathcal{T}} = V_{\mathcal{T} \setminus \{j\}}.$$

Proof. Dual statement. Over the field $\mathbb{Z}_q$, the orthogonal of a sum is the intersection of the orthogonals, and $V_{\mathcal{R}}^{\perp} = \text{Span}((\mathbf{x}_i^+)_{i \in \mathcal{R}})$. Hence

$$(V_{\mathcal{R}} + W_{\mathcal{T}})^{\perp} = \text{Span}((\mathbf{x}_i^+)_{i \in \mathcal{R}}) \cap W_{\mathcal{T}}^{\perp}$$

is the set of combinations of held keys that are orthogonal to the ciphertext space. We show that it equals $\text{Span}((\mathbf{x}_i^+)_{i \in \mathcal{R} \cap \mathcal{T}})$, that is, that only combinations of recipients' keys vanish on every ciphertext: this is the converse of correctness announced above. Taking orthogonals then gives $V_{\mathcal{R}} + W_{\mathcal{T}} = V_{\mathcal{R} \cap \mathcal{T}}$, and $\mathcal{R} \cap \mathcal{T} = \mathcal{T} \setminus \{j\}$.

The ciphertext space. Fix $t_0 \in \mathcal{T}$. Since $\mathbf{H}$ has full column rank,

$$\begin{aligned} W_{\mathcal{T}} &= \left\{ \mathbf{w}_{\mathbf{v}} = \begin{pmatrix} \mathbf{h}_{t_0}^t \mathbf{v} \\ \mathbf{H} \mathbf{v} \end{pmatrix} : \mathbf{v} \in \mathbb{Z}_q^d, \langle \mathbf{h}_t - \mathbf{h}_{t_0}, \mathbf{v} \rangle = 0 \text{ for all } t \in \mathcal{T} \right\} \\ &= \left\{ \mathbf{w}_{\mathbf{v}} : \mathbf{v} \in \text{Span}((\mathbf{h}_t - \mathbf{h}_{t_0})_{t \in \mathcal{T}})^{\perp} \right\}. \end{aligned}$$

How a key acts on it. For $i \in \mathcal{R}$, the relation $\mathbf{H}^t \mathbf{x}_i = -\mathbf{h}_i$ gives

$$\langle \mathbf{x}_i^+, \mathbf{w}_{\mathbf{v}} \rangle = \mathbf{h}_{t_0}^t \mathbf{v} + \mathbf{x}_i^t \mathbf{H} \mathbf{v} = -\langle \mathbf{h}_i - \mathbf{h}_{t_0}, \mathbf{v} \rangle.$$

Let $\mathbf{z} = \sum_{i \in \mathcal{R}} \lambda_i \mathbf{x}_i^+ \in \text{Span}((\mathbf{x}_i^+)_{i \in \mathcal{R}}) \cap W_{\mathcal{T}}^{\perp}$. For all $\mathbf{v} \in \text{Span}((\mathbf{h}_t - \mathbf{h}_{t_0})_{t \in \mathcal{T}})^{\perp}$, we have $\langle \mathbf{z}, \mathbf{w}_{\mathbf{v}} \rangle = 0$, hence, by linearity, $\langle \sum_{i \in \mathcal{R}} \lambda_i (\mathbf{h}_i - \mathbf{h}_{t_0}), \mathbf{v} \rangle = 0$. Therefore

$$\sum_{i \in \mathcal{R}} \lambda_i (\mathbf{h}_i - \mathbf{h}_{t_0}) \in \left(\text{Span}((\mathbf{h}_t - \mathbf{h}_{t_0})_{t \in \mathcal{T}})^{\perp} \right)^{\perp} = \text{Span}((\mathbf{h}_t - \mathbf{h}_{t_0})_{t \in \mathcal{T}}).$$

Affine independence. Since the projected keys are affinely independent, the vectors $\mathbf{h}_\ell - \mathbf{h}_{t_0}$, for $\ell \neq t_0$, are linearly independent. Hence $\lambda_i = 0$ for every $i \notin \mathcal{R} \cap \mathcal{T}$, and $\mathbf{z} \in \text{Span}((\mathbf{x}_i^+)_{i \in \mathcal{R} \cap \mathcal{T}})$. This shows that $(V_{\mathcal{R}} + W_{\mathcal{T}})^{\perp} \subseteq \text{Span}((\mathbf{x}_i^+)_{i \in \mathcal{R} \cap \mathcal{T}})$.

The converse inclusion. Let $i \in \mathcal{R} \cap \mathcal{T}$. Since $i \in \mathcal{T}$, every $\mathbf{w}_{\mathbf{v}} \in W_{\mathcal{T}}$ satisfies $\langle \mathbf{h}_i - \mathbf{h}_{t_0}, \mathbf{v} \rangle = 0$, hence $\langle \mathbf{x}_i^+, \mathbf{w}_{\mathbf{v}} \rangle = -\langle \mathbf{h}_i - \mathbf{h}_{t_0}, \mathbf{v} \rangle = 0$, and $\mathbf{x}_i^+ \in W_{\mathcal{T}}^{\perp}$. Since also $\mathbf{x}_i^+ \in V_{\mathcal{R}}^{\perp}$, we get $\mathbf{x}_i^+ \in (V_{\mathcal{R}} + W_{\mathcal{T}})^{\perp}$. Hence $\text{Span}((\mathbf{x}_i^+)_{i \in \mathcal{R} \cap \mathcal{T}}) \subseteq (V_{\mathcal{R}} + W_{\mathcal{T}})^{\perp}$, which proves the dual statement. $\square$

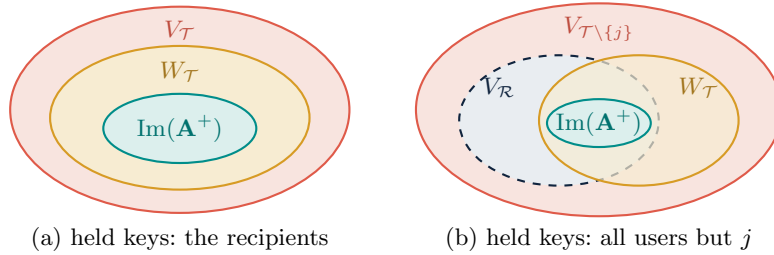


Fig. 2. Supports of the two branches in the proof of Proposition 7.2, after adding to the k -LWE challenge a uniform element of $W_{\mathcal{T}}$. (a) When the held keys are those of the recipients, as in [21], the inclusions $\text{Im}(\mathbf{A}^+) \subseteq W_{\mathcal{T}} \subseteq V_{\mathcal{T}}$ make both sums immediate. (b) When the held keys are those of $\mathcal{R} = [N] \setminus \{j\}$, which include non-recipients, $W_{\mathcal{T}}$ is no longer contained in $V_{\mathcal{R}}$. Lemma 7.4 computes their sum, $V_{\mathcal{R}} + W_{\mathcal{T}} = V_{\mathcal{T} \setminus \{j\}}$, which is also the sum obtained with $W_{\mathcal{T} \setminus \{j\}}$.

Figure 2 contrasts the situation of Lemma 7.4 with that of [21].

Proof of Proposition 7.2. Put $\mathcal{R} = [N] \setminus \{j\}$. As in [21, Theorem 27], we use the hardness of $(N-1, \mathbf{S})$ -LWE $_{\alpha, m}$, whose hints are the keys of $\mathcal{R}$: given $\mathbf{A}$, $\mathbf{u}$ and $(\mathbf{x}_i)_{i \in \mathcal{R}}$, the *LWE branch* $\frac{1}{q}U(\text{Im}(\mathbf{A}^+)) + \nu_{\alpha}^{m+1}$ and the *orthogonal branch* $\frac{1}{q}U(V_{\mathcal{R}}) + \nu_{\alpha}^{m+1}$ are indistinguishable.

The reduction. Let $\mathcal{I} \in \{\mathcal{T}, \mathcal{T} \setminus \{j\}\}$. Given $\mathbf{A}$, $\mathbf{u}$, $(\mathbf{x}_i)_{i \in \mathcal{R}}$ and a sample $\mathbf{c}$ from one of the two branches, the reduction samples $\mathbf{H}$ as in Setup_{pk} , sets $\mathbf{h}_i = -\mathbf{H}^t \mathbf{x}_i$ for $i \in \mathcal{R}$, and samples $\mathbf{h}_j$

uniformly from $\mathcal{H}_{\mathbf{u}} = \{\mathbf{h} : \mathbf{R}^t \mathbf{h} = \mathbf{u}\}$, where $\mathbf{A} = \mathbf{H}\mathbf{R}$. It gives the distinguisher the public key and the keys of $\mathcal{R}$. It then samples $\mathbf{y} \leftarrow \frac{1}{q}U(W_{\mathcal{I}})$, which only requires the projected keys, and outputs $\mathbf{c} + \mathbf{y}$. The public key and the keys of $\mathcal{R}$ are statistically close to their real distribution, for three reasons. First, in Setup_{pk} , the matrix $\mathbf{A}$ is statistically close to uniform by Lemma 3.10, and the keys, sampled with $\mathbf{T}_{\mathbf{A}}$, are statistically close to the Gaussian hints by Lemma 3.7. Second, the real projected key $-\mathbf{H}^t \mathbf{x}_j$ is statistically close to uniform on $\mathcal{H}_{\mathbf{u}}$ given $(\mathbf{A}, \mathbf{u}, \mathbf{H})$. Indeed, by Lemma 5.4 with $i = 1$, whose conditions hold by Theorem 5.6, the tuple $(\mathbf{A}, \mathbf{u}, \mathbf{H}, -\mathbf{H}^t \mathbf{x}_j)$ is statistically close to its counterpart in **Reversed**, where the projected key is uniform on $\mathcal{H}_{\mathbf{u}}$ given $(\mathbf{A}, \mathbf{u}, \mathbf{H})$; since the two setups are also statistically close, it is statistically close to $(\mathbf{A}, \mathbf{u}, \mathbf{H}, \mathbf{h})$ with $\mathbf{h} \leftarrow U(\mathcal{H}_{\mathbf{u}})$. Third, given $(\mathbf{A}, \mathbf{u})$, the keys of $\mathcal{R}$ are independent of $\mathbf{x}_j$.

Independence of the projected keys. Lemma 7.4, used below, requires $\mathbf{H}$ to have full column rank, which holds by construction, and the projected keys to be affinely independent, that is, their differences with one of them to be linearly independent. By the same application of Lemma 5.4, now to the $N - 1$ keys of $\mathcal{R}$, the projected keys are statistically close to independent uniform elements of $\mathcal{H}_{\mathbf{u}}$ given $(\mathbf{A}, \mathbf{u}, \mathbf{H})$. Their $N - 1$ differences with one of them are then statistically close to independent uniform elements of $\ker(\mathbf{R}^t)$, of dimension $d - n = n$, that is, to vectors $\mathbf{K}\mathbf{a}_{\ell}$ with independent $\mathbf{a}_{\ell} \leftarrow U(\mathbb{Z}_q^n)$, for a basis $\mathbf{K}$ of $\ker(\mathbf{R}^t)$. By Lemma 3.3, applied to the matrix with columns $\mathbf{a}_{\ell}$, they are dependent with probability at most $\delta_{n, N-1} \leq 2q^{-(n-N+2)} \leq 2q^{-\gamma n}$, since $N \leq (1 - \gamma)n$.

The two branches. Every projected key lies in $\mathcal{H}_{\mathbf{u}}$, so $\text{Im}(\mathbf{A}^+) \subseteq W_{\mathcal{I}}$, and Lemma 7.4 gives $V_{\mathcal{R}} + W_{\mathcal{I}} = V_{\mathcal{I} \setminus \{j\}} = V_{\mathcal{T} \setminus \{j\}}$, which also holds when $\mathcal{I} = \emptyset$, since $W_{\emptyset} = V_{\emptyset} = \mathbb{Z}_q^{m+1}$. Adding $\mathbf{y}$ therefore transforms the two branches as follows:

$$\begin{aligned} \text{LWE branch:} \quad & \frac{1}{q}U(\text{Im}(\mathbf{A}^+)) + \nu_{\alpha}^{m+1} \mapsto \frac{1}{q}U(W_{\mathcal{I}}) + \nu_{\alpha}^{m+1}, \\ \text{orthogonal branch:} \quad & \frac{1}{q}U(V_{\mathcal{R}}) + \nu_{\alpha}^{m+1} \mapsto \frac{1}{q}U(V_{\mathcal{T} \setminus \{j\}}) + \nu_{\alpha}^{m+1}. \end{aligned}$$

The orthogonal branch thus leads to the same distribution for $\mathcal{I} = \mathcal{T}$ and for $\mathcal{I} = \mathcal{T} \setminus \{j\}$.

Conclusion. Applying the reduction with $\mathcal{I} = \mathcal{T}$ and with $\mathcal{I} = \mathcal{T} \setminus \{j\}$ shows that, under the hardness of $(N - 1, \mathbf{S})\text{-LWE}_{\alpha, m}$ and up to the negligible terms above, given the public key and the keys of $\mathcal{R}$,

$$\frac{1}{q}U(W_{\mathcal{T}}) + \nu_{\alpha}^{m+1} \approx_c \frac{1}{q}U(V_{\mathcal{T} \setminus \{j\}}) + \nu_{\alpha}^{m+1} \approx_c \frac{1}{q}U(W_{\mathcal{T} \setminus \{j\}}) + \nu_{\alpha}^{m+1}.$$

Since $\mathcal{E}_{\mathcal{I}}(M)$ is obtained from $\frac{1}{q}U(W_{\mathcal{I}}) + \nu_{\alpha}^{m+1}$ by a deterministic post-processing, this gives $\mathcal{E}_{\mathcal{T}}(M) \approx_c \mathcal{E}_{\mathcal{T} \setminus \{j\}}(M)$. $\square$

7.3 Security of the Public-Key Scheme

Security is with respect to the IND-CPA and ANO-CPA experiments of Definitions 6.3 and 6.4, with user set $[N]$ and adaptive corruptions. The only computational step is Proposition 7.2.

Theorem 7.5 (IND-CPA security of the public-key scheme). *Under the $(i, \mathbf{S})\text{-LWE}_{\alpha, m}$ hardness assumption for every $i \leq N$, the scheme Π_{pk} is IND-CPA secure.*

Theorem 7.6 (ANO-CPA security of the public-key scheme). *Under the $(i, \mathbf{S})\text{-LWE}_{\alpha, m}$ hardness assumption for every $i \leq N$, the scheme Π_{pk} is ANO-CPA secure.*

Both proofs are hybrid arguments. They remove from the target set, one at a time and in a fixed order, the recipients that the adversary may not corrupt: all of them for IND-CPA, and those of $\mathcal{S}_b \setminus (\mathcal{S}_0 \cap \mathcal{S}_1)$ for ANO-CPA. Such a user is never corrupted, so each transition is an instance of

Proposition 7.2. In the last hybrid, the challenge no longer depends on the hidden bit: it is uniform for IND-CPA, and it encrypts M for $\mathcal{S}_0 \cap \mathcal{S}_1$ in both cases for ANO-CPA.

Proof of Theorem 7.5. Hybrids. For $0 \leq r \leq N$, let G_r be the IND-CPA experiment in which the challenge ciphertext is sampled from $\mathcal{E}_{\mathcal{T}_r}(M_b)$, where $\mathcal{T}_r = \mathcal{S} \cap \{r+1, \dots, N\}$ is computed once the adversary has chosen $\mathcal{S}$. Then G_0 is the real experiment, and in G_N the target set is empty, so that the challenge is uniform over $\mathbb{Z}_q^{m+1}$ whatever b , and the advantage is zero. The users are removed in the fixed order of $[N]$, so that the user removed by each transition is known before the game starts.

One transition. Fix $r < N$, put $j = r+1$, and let W_r be the event that $\mathcal{A}$ guesses b in G_r . We show that $|\Pr[W_r] - \Pr[W_{r+1}]|$ is negligible, by turning $\mathcal{A}$ into a distinguisher $\mathcal{B}$ for Proposition 7.2 with this j . Two cases arise.

- If $j \notin \mathcal{S}$, then $\mathcal{T}_{r+1} = \mathcal{T}_r$: the challenge does not change, and the two games coincide.
- If $j \in \mathcal{S}$, then j is never corrupted, and the two games differ only by the challenge, sampled from $\mathcal{E}_{\mathcal{T}_r}(M_b)$ in G_r and from $\mathcal{E}_{\mathcal{T}_r \setminus \{j\}}(M_b)$ in G_{r+1} . The adversary sees the public key and the keys of users other than j , and chooses $\mathcal{T}_r$ and M_b : this is the setting of Proposition 7.2.

The reduction. $\mathcal{B}$ receives the public key and the keys $(\mathbf{x}_i^+)_{i \neq j}$, which is all that $\mathcal{A}$ may see when $j \in \mathcal{S}$. It runs $\mathcal{A}$ on the public key and answers every extraction query on a user $i \neq j$, then follows the two cases.

- If $\mathcal{A}$ asks for the key of j , which forces $j \notin \mathcal{S}$, or outputs a target set $\mathcal{S}$ with $j \notin \mathcal{S}$, then $\mathcal{B}$ stops and outputs 0.
- If $\mathcal{A}$ outputs $(M_0, M_1, \mathcal{S})$ with $j \in \mathcal{S}$, then $\mathcal{B}$ samples b , chooses $\mathcal{T} = \mathcal{T}_r$ and $M = M_b$ in Proposition 7.2, receives a sample of $\mathcal{E}_{\mathcal{T}_r}(M_b)$ or of $\mathcal{E}_{\mathcal{T}_r \setminus \{j\}}(M_b)$, and gives it to $\mathcal{A}$ as the challenge ciphertext. It outputs 1 if and only if $\mathcal{A}$ guesses b .

Analysis. In the first case, $\mathcal{B}$ outputs 0 whatever sample it would receive, and this case contributes equally to $\Pr[W_r]$ and $\Pr[W_{r+1}]$. In the second case, $\mathcal{B}$ simulates G_r when it receives $\mathcal{E}_{\mathcal{T}_r}(M_b)$, and G_{r+1} when it receives $\mathcal{E}_{\mathcal{T}_r \setminus \{j\}}(M_b)$. Hence the advantage of $\mathcal{B}$ is $|\Pr[W_r] - \Pr[W_{r+1}]|$, which is therefore negligible by Proposition 7.2. Summing over the N transitions concludes the proof. $\square$

Proof of Theorem 7.6. Hybrids. Put $\mathcal{J} = \mathcal{S}_0 \cap \mathcal{S}_1$. For $b \in \{0, 1\}$ and $0 \leq r \leq N$, let G_r^b be the ANO-CPA experiment with challenge bit b , in which the challenge is sampled from $\mathcal{E}_{\mathcal{T}_r^b}(M)$, where

$$\mathcal{T}_r^b = \mathcal{J} \cup ((\mathcal{S}_b \setminus \mathcal{J}) \cap \{r+1, \dots, N\})$$

is computed once the adversary has chosen $\mathcal{S}_0$ and $\mathcal{S}_1$, and let W_r^b be the event that $\mathcal{A}$ outputs 1 in G_r^b . Then G_0^b is the real experiment with bit b , so the advantage of $\mathcal{A}$ is at most $|\Pr[W_0^0] - \Pr[W_0^1]|$, while G_N^0 and G_N^1 both sample the challenge from $\mathcal{E}_{\mathcal{J}}(M)$, and are therefore identical. As before, the users are removed in the fixed order of $[N]$.

One transition. Fix b and $r < N$, and put $j = r+1$. We show that $|\Pr[W_r^b] - \Pr[W_{r+1}^b]|$ is negligible. Two cases arise.

- If $j \notin \mathcal{S}_b \setminus \mathcal{J}$, then $\mathcal{T}_{r+1}^b = \mathcal{T}_r^b$: the challenge does not change, and the two games coincide.
- If $j \in \mathcal{S}_b \setminus \mathcal{J}$, then $j \in \mathcal{S}_0 \triangle \mathcal{S}_1$, so j is never corrupted, and $\mathcal{T}_{r+1}^b = \mathcal{T}_r^b \setminus \{j\}$. The two games differ only by the challenge, sampled from $\mathcal{E}_{\mathcal{T}_r^b}(M)$ in G_r^b and from $\mathcal{E}_{\mathcal{T}_r^b \setminus \{j\}}(M)$ in G_{r+1}^b : this is again the setting of Proposition 7.2.

The reduction. $\mathcal{B}$ receives the public key and the keys $(\mathbf{x}_i^+)_{i \neq j}$. It runs $\mathcal{A}$ on the public key and answers every extraction query on a user $i \neq j$, then follows the two cases.

- If $\mathcal{A}$ asks for the key of j , which forces $j \notin \mathcal{S}_b \setminus \mathcal{J}$, or outputs target sets with $j \notin \mathcal{S}_b \setminus \mathcal{J}$, then $\mathcal{B}$ stops and outputs 0.

- If $\mathcal{A}$ outputs $(M, \mathcal{S}_0, \mathcal{S}_1)$ with $j \in \mathcal{S}_b \setminus \mathcal{J}$, then $\mathcal{B}$ chooses $\mathcal{T} = \mathcal{T}_r^b$ and the bit M in Proposition 7.2, receives a sample of $\mathcal{E}_{\mathcal{T}_r^b}(M)$ or of $\mathcal{E}_{\mathcal{T}_r^b \setminus \{j\}}(M)$, and gives it to $\mathcal{A}$ as the challenge ciphertext. It outputs the bit returned by $\mathcal{A}$.

Analysis. As in the proof of Theorem 7.5, $\mathcal{B}$ outputs 0 in the first case whatever sample it would receive, and this case contributes equally to $\Pr[W_r^b]$ and $\Pr[W_{r+1}^b]$; in the second case, $\mathcal{B}$ simulates G_r^b or G_{r+1}^b . Hence the advantage of $\mathcal{B}$ is $|\Pr[W_r^b] - \Pr[W_{r+1}^b]|$, which is negligible by Proposition 7.2. Summing over the $2N$ transitions, $|\Pr[W_0^0] - \Pr[W_0^1]|$ is negligible, and so is the advantage of $\mathcal{A}$. $\square$

Remark 7.7. The anonymity proof passes through the intersection $\mathcal{S}_0 \cap \mathcal{S}_1$ rather than through $\mathcal{S}_0 \cup \mathcal{S}_1$ as in [12], which keeps every intermediate target set inside $\mathcal{S}_b$. Users of the intersection may be corrupted; they are never removed. Both paths work here, but only this one carries over to the unbounded universe of Section 8, where the model guarantees $|\mathcal{C} \cup \mathcal{S}_b| \leq N$ for each b , whereas $\mathcal{S}_0 \cup \mathcal{S}_1$ may contain more identities, whose projected keys need not be affinely independent.

7.4 The Identity-Based Scheme

The setting of Section 6.2 is satisfied as follows.

- (S1) By Theorem 5.6 with $d = 2n$ and $i = 1$, the family TPSam_1 is a trapdoor projective sampling family, whose keys are single vectors $\mathbf{x}$, with projected keys $-\mathbf{H}^t \mathbf{x}$. Its projected-key distribution is uniform over $\mathcal{H}_{\mathbf{u}}$, which has q^n elements, so two samples collide with probability q^{-n} ; and it can be sampled from $(\mathbf{A}, \mathbf{u}, \mathbf{H})$, which determine $\mathcal{H}_{\mathbf{u}}$.
- (S2) PSam is the family it induces, whose public parameters $(\mathbf{A}, \mathbf{u}, \mathbf{H})$ are sampled in the reversed order.
- (S3) This condition holds in the relaxed form of Corollary 6.7. As noted in Section 7.1, Π_{pk} is projective with respect to the family of Theorem 4.2 with $d = 2n$. For given $(\mathbf{A}, \mathbf{u}, \mathbf{H})$, this family has the same key distribution $D_{\Lambda_{\perp_{\mathbf{u}}}(\mathbf{A}), \mathbf{S}}$, projection $\mathbf{x} \mapsto -\mathbf{H}^t \mathbf{x}$ and samplers as PSam . It only samples $(\mathbf{A}, \mathbf{u}, \mathbf{H})$ in the regular order, which is statistically close to the reversed one by Lemma 5.4 with $i = 0$ and Lemmas 3.10 and 3.3.

The compiler then yields the following scheme, where $\text{Hash} : \{0, 1\}^* \rightarrow \mathbb{Z}_q^n$ is a random oracle.

Setup_{id}($1^n, \mathcal{U}$): Generate $(\mathbf{H}, \mathbf{T}_{\mathbf{H}}) \in \mathbb{Z}_q^{m \times 2n} \times \mathbb{Z}^{m \times m}$ using Lemma 3.10. Sample $\mathbf{R} \leftarrow U(\mathbb{Z}_q^{2n \times n})$, conditioned on $\text{rank}(\mathbf{R}) = n$, and set $\mathbf{A} = \mathbf{H}\mathbf{R}$. Sample $\mathbf{h}_0 \leftarrow U(\mathbb{Z}_q^{2n})$, set $\mathbf{u} = \mathbf{R}^t \mathbf{h}_0$ and $\mathbf{A}^+ = (\mathbf{u}^t \| \mathbf{A})$, and compute $\mathbf{K} \in \mathbb{Z}_q^{2n \times n}$ whose columns form a basis of $\ker(\mathbf{R}^t)$. Output $\text{ek} = (\mathbf{A}^+, \mathbf{H}, \mathbf{h}_0, \mathbf{K}, \text{Hash})$ and $\text{MSK} = (\mathbf{T}_{\mathbf{H}}, \mathcal{U})$. For an identity id , write $\mathbf{h}_{id} = \mathbf{h}_0 + \mathbf{K} \text{Hash}(id)$ and $\mathbf{H}_{id} = (\mathbf{h}_{id}^t \| \mathbf{H})$.

Extract_{id}($\text{ek}, \text{MSK}, id$): If $id \notin \mathcal{U}$, return $\perp$. If id was already queried, return the stored key. Otherwise, sample $\mathbf{x}_{id}$ statistically close to $D_{\Lambda_{\perp_{\mathbf{h}_{id}}}(\mathbf{H}), \mathbf{S}}$ using Lemma 3.7 with the basis $\mathbf{T}_{\mathbf{H}}$, and store and return $\mathbf{x}_{id}^+ = \begin{pmatrix} 1 \\ \mathbf{x}_{id} \end{pmatrix}$.

Encrypt_{id}($\text{ek}, M, \mathcal{S}$): For $\mathcal{S} \subseteq \mathcal{U}$, proceed as Enc_{pk} with the matrices $(\mathbf{H}_{id})_{id \in \mathcal{S}}$.

Decrypt_{id}($\text{ek}, \mathbf{x}_{id}^+, \mathbf{c}$): as Dec_{pk} .

The identity map of Construction 6.5 is $\text{IdMap}(id) = \mathbf{h}_{id}$, whose outputs are uniform over $\mathcal{H}_{\mathbf{u}}$, as required, since $\mathbf{K}$ is a basis of $\ker(\mathbf{R}^t)$. The values $\mathbf{h}_0$ and $\mathbf{K}$ only serve to implement this map, and may be published by Lemma 5.5.

Theorem 7.8 (Security of the identity-based scheme). *Under the $(i, \mathbf{S})$ -LWE $_{\alpha, m}$ hardness assumption for every $i \leq N$, the scheme Π_{id} is correct, IND-CPA secure and ANO-CPA secure for a bounded universe of N users, with adaptive corruptions, in the random oracle model.*

Proof. As verified above, Corollary 6.7 applies to Π_{pk} and to the lattice trapdoor family, and it produces Π_{id} . The claim follows from Theorems 7.1, 7.5 and 7.6. $\square$

Remark 7.9 (Public key). The public key of Π_{pk} lists N projected keys. That of Π_{id} consists of $\mathbf{A}^+$, $\mathbf{H}$, $\mathbf{h}_0$, $\mathbf{K}$ and the random oracle, hence has size $O(mn \log q)$ bits, independently of N ; since it does not depend on the user set, it reveals neither this set nor its cardinality. Ciphertexts and private keys have the same size in both schemes: a ciphertext is one vector of $\mathbb{Z}_q^{m+1}$, independent of the size of the target set, and a private key is one short vector of length $m + 1$.

8 Unbounded Identity Universe with Static Corruption

The scheme of Section 7.4 can serve any identity, since the authority extracts keys on demand with its trapdoor. In this section, identities are arbitrary and corruptions are static, while target sets remain adaptive.

8.1 Model

An identity-based broadcast scheme for an unbounded universe has the syntax of Definition 6.1, except that $\text{Setup}(1^n)$ takes no user set and Extract answers every identity. Security is defined by IND-CPA and ANO-CPA experiments with static corruption and a bound N : at most N identities may be corrupted or receive the challenge. The adversary first commits to a set $\mathcal{C}$ of at most N identities, whose keys it may obtain; it then receives the public key and chooses its target sets adaptively, subject to the restrictions of Definitions 6.3 and 6.4 and to $|\mathcal{C} \cup \mathcal{S}| \leq N$ for every target set $\mathcal{S}$. The number of identities queried to the random oracle is not bounded. The formal definitions follow.

Definition 8.1 (Identity-based broadcast for an unbounded universe). *Let $\mathcal{I}$ be an identity space, and let PT and CT be plaintext and ciphertext spaces. An identity-based broadcast scheme is a tuple $\Pi = (\text{Setup}, \text{Extract}, \text{Enc}, \text{Dec})$ with the following syntax:*

- $\text{Setup}(1^n)$ outputs a public encryption key ek and a master secret key MSK ;
- $\text{Extract}(\text{ek}, \text{MSK}, \text{id})$ outputs a secret key sk_{id} for any $\text{id} \in \mathcal{I}$;
- $\text{Enc}(\text{ek}, M, \mathcal{S})$ takes $M \in \text{PT}$ and a finite target set $\mathcal{S} \subseteq \mathcal{I}$, and outputs $\text{ct} \in \text{CT}$;
- $\text{Dec}(\text{ek}, \text{sk}_{\text{id}}, \text{ct})$ outputs a message or $\perp$.

Correctness is defined as in Definition 6.1, for every finite target set $\mathcal{S} \subseteq \mathcal{I}$.

The security games bound only the identities that are actively involved in the game: the corrupted identities and the recipients of the challenge.

Definition 8.2 (IND-CPA and ANO-CPA with static corruption). *Let Π be as in Definition 8.1, and let N be an integer bounding the identities involved in the game. The IND-CPA experiment with static corruption and bound N is as follows.*

1. *The adversary outputs a set $\mathcal{C} \subseteq \mathcal{I}$ of identities to corrupt, with $|\mathcal{C}| \leq N$.*
2. *The challenger runs $(\text{ek}, \text{MSK}) \leftarrow \text{Setup}(1^n)$ and gives ek to $\mathcal{A}$.*
3. *At any time, the adversary may ask for the secret key of any identity of $\mathcal{C}$; no other secret key is revealed.*
4. *The adversary outputs two messages $M_0, M_1 \in \text{PT}$ and a target set $\mathcal{S} \subseteq \mathcal{I}$ with $\mathcal{C} \cap \mathcal{S} = \emptyset$ and $|\mathcal{C} \cup \mathcal{S}| \leq N$. The challenger samples $b \leftarrow U(\{0, 1\})$ and returns $\text{ct} \leftarrow \text{Enc}(\text{ek}, M_b, \mathcal{S})$.*
5. *The adversary outputs a bit b' . The experiment outputs 1 if $b' = b$, and 0 otherwise.*

The ANO-CPA experiment is identical, except that at step 4 the adversary outputs a message M and two target sets $\mathcal{S}_0, \mathcal{S}_1 \subseteq \mathcal{I}$ with $\mathcal{C} \cap (\mathcal{S}_0 \Delta \mathcal{S}_1) = \emptyset$ and $|\mathcal{C} \cup \mathcal{S}_b| \leq N$ for $b \in \{0, 1\}$, and receives $\text{ct} \leftarrow \text{Enc}(\text{ek}, M, \mathcal{S}_b)$. Advantages and security are defined as in Definitions 6.3 and 6.4. In the random oracle model, the oracle is part of ek , so it is queried only after step 1.

Remark 8.3. Only the set of corrupted identities is committed in advance; the target sets are chosen adaptively, after the public key and the corrupted keys have been seen. The bound N covers the corrupted identities and the recipients of the challenge. The number of identities queried to the random oracle is not bounded by the model, and appears in the security bounds of Section 8.4.

8.2 The Scheme

The scheme is the identity-based scheme Π_{id} of Section 7.4, without the user set: $\text{Setup}(1^n)$ is Setup_{id} without its user set, with $\text{MSK} = \mathbf{T}_{\mathbf{H}}$; Extract answers every identity; encryption and decryption are unchanged. The public key is the same as in the bounded case. Correctness follows from the proof of Theorem 7.1 (Section 7.1), which only uses the relation $\mathbf{H}^t \mathbf{x}_{id} = -\mathbf{h}_{id}$ and the norm bound of Lemma 3.4; both hold for the keys extracted with the trapdoor, whatever the target set.

8.3 Removing All Recipients Whose Keys Are Not Held

When the held keys are those of a committed set $\mathcal{C}$, no hybrid over the recipients is needed: Proposition 7.2 extends to a single transition that removes at once all the recipients whose keys are not held. We use the notation $W_{\mathcal{T}}$, $V_{\mathcal{C}}$ and $\mathcal{E}_{\mathcal{T}}(M)$ of Section 7.2, with identities in place of users. Throughout this subsection and the next, the parameters are as in Theorem 5.6 with $d = 2n$, and the bound N of Definition 8.2 satisfies $N \leq (1 - \gamma)n$ for a constant $\gamma \in (0, 1)$.

Let $\mathcal{C}$ be a set of at most N identities, fixed before the public key of the scheme of Section 8.2 is generated. The projected key of an identity id is $\mathbf{h}_{id} = \mathbf{h}_0 + \mathbf{K}\text{Hash}(id)$; in the random oracle model, the projected keys of distinct identities are independent uniform elements of $\mathcal{H}_{\mathbf{u}}$. We write $\mathcal{L}$ for the set of identities queried to the random oracle, together with $\mathcal{C}$, and Good for the event that $\mathbf{H}$ has full column rank and that the projected keys of every set of at most N identities of $\mathcal{L}$ are affinely independent. Section 8.4 bounds the probability of $\neg \text{Good}$.

The regular experiment. The proof below goes through the *regular experiment*, which samples the public values in the order of [21, 12]:

- $\mathbf{A} \leftarrow U(\mathbb{Z}_q^{m \times n})$ and $\mathbf{u} \leftarrow U(\mathbb{Z}_q^n)$;
- $\mathbf{x}_c \leftarrow D_{A^{\perp_{\mathbf{u}}}}(\mathbf{A}), \mathbf{S}$ for $c \in \mathcal{C}$;
- $\mathbf{H}$ uniformly subject to $\text{rank}(\mathbf{H}) = d$ and $\text{Im}(\mathbf{A}) \subseteq \text{Im}(\mathbf{H})$, with $\mathbf{A} = \mathbf{H}\mathbf{R}$ and $\mathcal{H}_{\mathbf{u}} = \{\mathbf{h} : \mathbf{R}^t \mathbf{h} = \mathbf{u}\}$;
- $\mathbf{h}_c = -\mathbf{H}^t \mathbf{x}_c$ for $c \in \mathcal{C}$, and, for every other identity, a projected key sampled uniformly from $\mathcal{H}_{\mathbf{u}}$ when it is first requested.

Lemma 8.4 (Subspace sum, general form). *Let $\mathcal{C}, \mathcal{T}$ be finite index sets with $\mathcal{T} \neq \emptyset$. Assume that $\mathbf{H}$ has full column rank, that $\mathbf{H}^t \mathbf{x}_c = -\mathbf{h}_c$ for every $c \in \mathcal{C}$, and that the projected keys $(\mathbf{h}_{\ell})_{\ell \in \mathcal{C} \cup \mathcal{T}}$ are affinely independent. Then*

$$V_{\mathcal{C}} + W_{\mathcal{T}} = V_{\mathcal{C} \cup \mathcal{T}}.$$

Proof. The proof of Lemma 7.4 applies with $\mathcal{C}$ in place of $\mathcal{R}$ and $\mathcal{C} \cup \mathcal{T}$ in place of $[N]$: it only uses the relation $\mathbf{H}^t \mathbf{x}_c = -\mathbf{h}_c$ for the held keys and the affine independence of the projected keys involved, and it yields $V_{\mathcal{C}} + W_{\mathcal{T}} = V_{\mathcal{C} \cup \mathcal{T}}$. $\square$

Theorem 8.5 (Removing all recipients whose keys are not held). *Under the $(|\mathcal{C}|, \mathbf{S})$ -LWE $_{\alpha, m}$ hardness assumption, given the public key of the scheme, access to the random oracle and*

the keys $(\mathbf{x}_c^+)_{c \in \mathcal{C}}$, every probabilistic polynomial-time distinguisher between $\mathcal{E}_{\mathcal{T}}(M)$ and $\mathcal{E}_{\mathcal{C} \cap \mathcal{T}}(M)$, where $\mathcal{T} \subseteq \mathcal{L}$ with $|\mathcal{C} \cup \mathcal{T}| \leq N$ and the bit M are chosen by the distinguisher, has advantage at most

$$\begin{cases} \Pr[\neg \text{Good}] + \text{negl}(n) & \text{if } \mathcal{C} \cap \mathcal{T} = \emptyset, \\ 2 \Pr[\neg \text{Good}] + \text{negl}(n) & \text{otherwise.} \end{cases}$$

In words: an adversary holding the keys of $\mathcal{C}$ cannot tell a ciphertext for $\mathcal{T}$ from a ciphertext for $\mathcal{C} \cap \mathcal{T}$, the recipients whose keys it holds; all the other recipients are removed at once. Proposition 7.2 is the special case $\mathcal{C} = [N] \setminus \{j\}$, which removes a single recipient. Only the identities of $\mathcal{C}$ carry k -LWE hints: the projected key of any other identity is a uniform point of $\mathcal{H}_{\mathbf{u}}$, sampled when it is requested. Hence only the corrupted identities have to be known in advance, and the target sets may be chosen adaptively.

Proof. Transfer to the regular experiment. The scheme samples its public values in the reversed order: $\mathbf{H}$ first, with its trapdoor, then $\mathbf{A} = \mathbf{H}\mathbf{R}$, $\mathbf{h}_0$ and $\mathbf{K}$, and secret keys extracted with the trapdoor. Lemma 5.4, applied to the $|\mathcal{C}|$ identities of $\mathcal{C}$, with conditions that hold by Theorem 5.6, shows that the joint distribution of $(\mathbf{A}, \mathbf{u}, \mathbf{H}, (\mathbf{h}_c, \mathbf{x}_c)_{c \in \mathcal{C}})$ is statistically close in both orders, while the other projected keys are independent uniform elements of $\mathcal{H}_{\mathbf{u}}$ in both. In the regular experiment, we sample $\mathbf{h}_0$ uniformly from $\mathcal{H}_{\mathbf{u}}$, take a basis $\mathbf{K}$ of $\ker(\mathbf{R}^t)$, and let the random-oracle answer on id be the unique $\mathbf{a}_{id}$ with $\mathbf{h}_0 + \mathbf{K}\mathbf{a}_{id} = \mathbf{h}_{id}$, which is the simulation of Lemma 5.5. It therefore suffices to prove the claim in the regular experiment.

The reduction. In the regular experiment, the proof is that of Proposition 7.2, with the following changes.

- The hints of the $(|\mathcal{C}|, \mathbf{S})$ -LWE $_{\alpha, m}$ instance are the keys of $\mathcal{C}$. The reduction completes the regular experiment: it samples $\mathbf{H}$, sets $\mathbf{h}_c = -\mathbf{H}^t \mathbf{x}_c$ for $c \in \mathcal{C}$, samples the other projected keys from $\mathcal{H}_{\mathbf{u}}$ on request, and computes $\mathbf{h}_0$, $\mathbf{K}$ and the random-oracle answers as above. The view it produces is exactly the regular experiment, so the statistical steps of the reduction of Proposition 7.2 are not needed.
- The two target sets are $\mathcal{T}$ and $\mathcal{C} \cap \mathcal{T}$, and Lemma 8.4 replaces Lemma 7.4: on **Good**, $V_{\mathcal{C}} + W_{\mathcal{I}} = V_{\mathcal{C} \cap \mathcal{T}}$ for $\mathcal{I} \in \{\mathcal{T}, \mathcal{C} \cap \mathcal{T}\}$, since $|\mathcal{C} \cup \mathcal{I}| \leq N$. Both orthogonal branches lead to $\frac{1}{q}U(V_{\mathcal{C} \cap \mathcal{T}}) + \nu_{\alpha}^{m+1}$, and each reduction loses $\Pr[\neg \text{Good}]$.
- When $\mathcal{C} \cap \mathcal{T} = \emptyset$, the orthogonal branch and $\mathcal{E}_{\emptyset}(M)$ are both uniform, so the reduction with $\mathcal{I} = \mathcal{C} \cap \mathcal{T}$ is not needed.

□

Corollary 8.6. *In the setting of Theorem 8.5, with target sets and a bit M chosen by the distinguisher:*

- (a) Pseudorandomness. *If $\mathcal{C} \cap \mathcal{S} = \emptyset$ and $|\mathcal{C} \cup \mathcal{S}| \leq N$, then every probabilistic polynomial-time distinguisher between $\mathcal{E}_{\mathcal{S}}(M)$ and the uniform distribution over $\mathbb{Z}_q^{m+1}$ has advantage at most $\Pr[\neg \text{Good}] + \text{negl}(n)$.*
- (b) Anonymity. *If $\mathcal{C} \cap (\mathcal{S}_0 \triangle \mathcal{S}_1) = \emptyset$ and $|\mathcal{C} \cup \mathcal{S}_b| \leq N$ for $b \in \{0, 1\}$, then every probabilistic polynomial-time distinguisher between $\mathcal{E}_{\mathcal{S}_0}(M)$ and $\mathcal{E}_{\mathcal{S}_1}(M)$ has advantage at most $4 \Pr[\neg \text{Good}] + \text{negl}(n)$.*

Proof. For (a), the theorem applies with $\mathcal{C} \cap \mathcal{S} = \emptyset$, and $\mathcal{E}_{\emptyset}(M)$ is uniform over $\mathbb{Z}_q^{m+1}$ whatever M . For (b), the hypothesis implies $\mathcal{C} \cap \mathcal{S}_0 = \mathcal{C} \cap \mathcal{S}_1$, so the theorem applied to $\mathcal{S}_0$ and to $\mathcal{S}_1$ reduces both distributions to the ciphertext distribution for this common set. □

8.4 Security with Static Corruption

The parameters are as in Theorem 5.6 with $d = 2n$, and the bound N satisfies $N \leq (1 - \gamma)n$ for a constant $\gamma \in (0, 1)$. We denote by Q the number of distinct identities queried to the random oracle during the experiment, by the adversary or by the challenger. It includes the corrupted identities and the recipients of the challenge. The proofs rest on the generalization of Proposition 7.2 of Section 8.3, which removes at once all the recipients whose keys are not held (Theorem 8.5); only the corrupted identities carry k -LWE hints, hence the static corruption. Under the $(i, \mathbf{S})$ -LWE $_{\alpha, m}$ assumption for every $i \leq N$, and in the random oracle model, the scheme is IND-CPA and ANO-CPA secure with static corruption, with advantages at most $2Q^N q^{-\gamma n} + \text{negl}(n)$ and $4Q^N q^{-\gamma n} + \text{negl}(n)$.

We state and prove the two security theorems announced in Section 8.4. Since the adversary chooses its target sets after seeing the projected keys of the identities it has hashed, the event **Good** covers every small set of queried identities. The following lemma bounds its failure.

Lemma 8.7 (Independence of all small sets). *Assume that $\mathbf{R}$ has rank n , and let $\mathbf{h}_1, \dots, \mathbf{h}_Q$ be independent uniform elements of $\mathcal{H}_{\mathbf{u}}$, with $Q \geq 2$. For $N \leq n + 1$, the probability that some set of at most N of these points is affinely dependent is at most*

$$Q^N \delta_{n, N-1} \leq 2 Q^N q^{-(n-N+2)}.$$

Proof. A set of k points is affinely dependent exactly when the $k - 1$ differences with one of them are linearly dependent. Conditioned on this point, the differences are independent uniform elements of $\ker(\mathbf{R}^t)$, which has dimension $d - n = n$. By Lemma 3.3, they are dependent with probability at most $\delta_{n, k-1} \leq \delta_{n, N-1}$ when $k \leq N$. Sets of at most one point are never dependent, and there are at most Q^N sets of 2 to N points, so a union bound concludes. $\square$

In the scheme, $\mathbf{R}$ has rank n by construction, $\mathbf{H}$ has full column rank except with negligible probability, and the projected keys of the identities of $\mathcal{L}$ are independent uniform elements of $\mathcal{H}_{\mathbf{u}}$, with $|\mathcal{L}| \leq Q$. Lemma 8.7 and $N \leq (1 - \gamma)n$ therefore give

$$\Pr[\neg \text{Good}] \leq 2Q^N q^{-\gamma n} + \text{negl}(n). \quad (1)$$

Theorem 8.8 (IND-CPA security with static corruption). *Under the $(i, \mathbf{S})$ -LWE $_{\alpha, m}$ hardness assumption for every $i \leq N$, every probabilistic polynomial-time adversary against the scheme of Section 8.2, in the IND-CPA experiment with static corruption and bound N and in the random oracle model, has advantage at most $2Q^N q^{-\gamma n} + \text{negl}(n)$. In particular, the scheme is IND-CPA secure with static corruption and bound N whenever this term is negligible, for instance for $N = o(n)$ and any polynomial number of queries (Remark 8.9).*

Proof of Theorem 8.8. Let $\mathcal{C}$ be the set output by the adversary at the first step. It is fixed before the public key is generated, so we may build a distinguisher $\mathcal{D}$ for Corollary 8.6(a) with this $\mathcal{C}$.

The distinguisher. $\mathcal{D}$ receives the public key, access to the random oracle and the keys of $\mathcal{C}$. It runs $\mathcal{A}$, forwards its oracle queries, and answers its extraction queries with these keys. When $\mathcal{A}$ outputs $(M_0, M_1, \mathcal{S})$, $\mathcal{D}$ samples b , chooses $\mathcal{S}$ and M_b in the corollary, whose hypotheses $\mathcal{C} \cap \mathcal{S} = \emptyset$ and $|\mathcal{C} \cup \mathcal{S}| \leq N$ are imposed by the experiment, and gives the sample it receives to $\mathcal{A}$ as the challenge ciphertext. It outputs 1 if and only if $\mathcal{A}$ guesses b .

Analysis. When $\mathcal{D}$ receives $\mathcal{E}_{\mathcal{S}}(M_b)$, it simulates the real experiment. When it receives a uniform element of $\mathbb{Z}_q^{m+1}$, the challenge is independent of b , and $\mathcal{A}$ guesses b with probability $1/2$. Hence the advantage of $\mathcal{A}$ equals that of $\mathcal{D}$, which is at most $\Pr[\neg \text{Good}] + \text{negl}(n)$ by Corollary 8.6(a), and (1) concludes. $\square$

Remark 8.9 (The bound N and the number of queries). The term $2Q^N q^{-\gamma n}$ bounds the probability that the projected keys of some set of at most N queried identities are affinely dependent

(Lemma 8.7, with $N \leq (1 - \gamma)n$). This event includes collisions: two queried identities with the same projected key form an affinely dependent set, and the key of one would decrypt for the other. The term equals $2^{N \log Q - \gamma n \log q + 1}$, and is negligible as soon as $N \log Q \leq \gamma n \log q - \omega(\log n)$. For $N = (1 - \gamma)n$, this allows about $q^{\gamma/(1-\gamma)}$ random-oracle queries, for instance about q queries when $\gamma = 1/2$. For the term to be negligible for every polynomial number of queries, it suffices to take $N = o(n)$, for instance $N \leq n/\log n$.

Theorem 8.10 (ANO-CPA security with static corruption). *Under the same assumption, every probabilistic polynomial-time adversary against the scheme of Section 8.2, in the ANO-CPA experiment with static corruption and bound N and in the random oracle model, has advantage at most $4Q^N q^{-\gamma n} + \text{negl}(n)$. In particular, the scheme is ANO-CPA secure with static corruption and bound N whenever this term is negligible.*

Proof of Theorem 8.10. The proof is that of Theorem 8.8, with Corollary 8.6(b) in place of (a). The distinguisher $\mathcal{D}$ runs $\mathcal{A}$ in the same way. When $\mathcal{A}$ outputs $(M, \mathcal{S}_0, \mathcal{S}_1)$, it chooses them in the corollary, whose hypotheses $\mathcal{C} \cap (\mathcal{S}_0 \triangle \mathcal{S}_1) = \emptyset$ and $|\mathcal{C} \cup \mathcal{S}_b| \leq N$ are imposed by the experiment, gives the sample it receives to $\mathcal{A}$ as the challenge ciphertext, and outputs the bit returned by $\mathcal{A}$. It thus distinguishes $\mathcal{E}_{\mathcal{S}_0}(M)$ from $\mathcal{E}_{\mathcal{S}_1}(M)$ with twice the advantage of $\mathcal{A}$, which is therefore at most $2\Pr[\neg \text{Good}] + \text{negl}(n)$ by Corollary 8.6(b), and (1) concludes. $\square$

Remark 8.11 (Why not through the compiler). The bounded-universe result reduces to the public-key scheme Π_{pk} , whose reduction holds a hint for each of its N users. Here the adversary may hash polynomially many identities and target any of them, adaptively, so these identities cannot be mapped in advance to the N users of a bounded public-key scheme. The direct proof avoids this because Theorem 8.5 needs hints for the corrupted identities only.

References

1. Abdalla, M., Bellare, M., Catalano, D., Kiltz, E., Kohno, T., Lange, T., Malone-Lee, J., Neven, G., Paillier, P., Shi, H.: Searchable encryption revisited: Consistency properties, relation to anonymous IBE, and extensions. In: Shoup, V. (ed.) CRYPTO 2005. LNCS, vol. 3621, pp. 205–222. Springer, Berlin, Heidelberg (Aug 2005). https://doi.org/10.1007/11535218_13
2. Agrawal, S., Boneh, D., Boyen, X.: Efficient lattice (H)IBE in the standard model. In: Gilbert, H. (ed.) EUROCRYPT 2010. LNCS, vol. 6110, pp. 553–572. Springer, Berlin, Heidelberg (May / Jun 2010). https://doi.org/10.1007/978-3-642-13190-5_28
3. Agrawal, S., Gentry, C., Halevi, S., Sahai, A.: Discrete Gaussian leftover hash lemma over infinite domains. In: Sako, K., Sarkar, P. (eds.) ASIACRYPT 2013, Part I. LNCS, vol. 8269, pp. 97–116. Springer, Berlin, Heidelberg (Dec 2013). https://doi.org/10.1007/978-3-642-42033-7_6
4. Alwen, J., Peikert, C.: Generating shorter bases for hard random lattices. *Theory of Computing Systems* **48**(3), 535–553 (2011)
5. Barth, A., Boneh, D., Waters, B.: Privacy in encrypted content distribution using private broadcast encryption. In: Di Crescenzo, G., Rubin, A. (eds.) FC 2006. LNCS, vol. 4107, pp. 52–64. Springer, Berlin, Heidelberg (Feb / Mar 2006). https://doi.org/10.1007/11889663_4
6. Boneh, D., Franklin, M.K.: Identity-based encryption from the Weil pairing. In: Kilian, J. (ed.) CRYPTO 2001. LNCS, vol. 2139, pp. 213–229. Springer, Berlin, Heidelberg (Aug 2001). https://doi.org/10.1007/3-540-44647-8_13
7. Boyen, X., Waters, B.: Anonymous hierarchical identity-based encryption (without random oracles). In: Dwork, C. (ed.) CRYPTO 2006. LNCS, vol. 4117, pp. 290–307. Springer, Berlin, Heidelberg (Aug 2006). https://doi.org/10.1007/11818175_17
8. Cash, D., Hofheinz, D., Kiltz, E., Peikert, C.: Bonsai trees, or how to delegate a lattice basis. In: Gilbert, H. (ed.) EUROCRYPT 2010. LNCS, vol. 6110, pp. 523–552. Springer, Berlin, Heidelberg (May / Jun 2010). https://doi.org/10.1007/978-3-642-13190-5_27
9. Cocks, C.: An identity based encryption scheme based on quadratic residues. In: Honary, B. (ed.) 8th IMA International Conference on Cryptography and Coding. LNCS, vol. 2260, pp. 360–363. Springer, Berlin, Heidelberg (Dec 2001). https://doi.org/10.1007/3-540-45325-3_32

10. Debris-Alazard, T., Dyseryn, V., Phan, D.H.: Hardness of problems with hints in code-based cryptography and applications to traitor tracing. In: ASIACRYPT 2026 (2025), <https://eprint.iacr.org/2025/2151>
11. Do, X.T., Persiano, G., Phan, D.H., Yung, M.: Anamorphism beyond one-to-one messaging: Public-key with anamorphic broadcast mode. In: Fehr, S., Fouque, P.A. (eds.) EUROCRYPT 2025, Part III. LNCS, vol. 15603, pp. 429–455. Springer, Cham (May 2025). https://doi.org/10.1007/978-3-031-91131-6_15
12. Do, X.T., Phan, D.H., Yung, M.: A concise bounded anonymous broadcast yielding combinatorial trace-and-revoke schemes. In: Conti, M., Zhou, J., Casalicchio, E., Spognardi, A. (eds.) ACNS 2020, Part II. LNCS, vol. 12147, pp. 145–164. Springer, Cham (Oct 2020). https://doi.org/10.1007/978-3-030-57878-7_8
13. Gennaro, R., Halevi, S., Rabin, T.: Secure hash-and-sign signatures without the random oracle. In: Stern, J. (ed.) EUROCRYPT’99. LNCS, vol. 1592, pp. 123–139. Springer, Berlin, Heidelberg (May 1999). https://doi.org/10.1007/3-540-48910-X_9
14. Gentry, C.: Practical identity-based encryption without random oracles. In: Vaudenay, S. (ed.) EUROCRYPT 2006. LNCS, vol. 4004, pp. 445–464. Springer, Berlin, Heidelberg (May / Jun 2006). https://doi.org/10.1007/11761679_27
15. Gentry, C., Peikert, C., Vaikuntanathan, V.: Trapdoors for hard lattices and new cryptographic constructions. In: Ladner, R.E., Dwork, C. (eds.) 40th ACM STOC. pp. 197–206. ACM Press (May 2008). <https://doi.org/10.1145/1374376.1374407>
16. He, K., Liu, X., Liu, J.N., Liu, W.: Efficient identity-based broadcast encryption scheme on lattices for the internet of things. *Security and Communication Networks* **2021**, 2847731 (2021). <https://doi.org/10.1155/2021/2847731>
17. He, K., Weng, J., Liu, J., Liu, J.K., Liu, W., Deng, R.H.: Anonymous identity-based broadcast encryption with chosen-ciphertext security. In: Chen, X., Wang, X., Huang, X. (eds.) ASIACCS 16. pp. 247–255. ACM Press (May / Jun 2016). <https://doi.org/10.1145/2897845.2897879>
18. Kiayias, A., Samari, K.: Lower bounds for private broadcast encryption. In: Information Hiding (IH) 2012. LNCS, vol. 7692, pp. 176–190. Springer (2013)
19. Lai, J., Mu, Y., Guo, F., Susilo, W., Chen, R.: Anonymous identity-based broadcast encryption with revocation for file sharing. In: Liu, J.K., Steinfeld, R. (eds.) ACISP 16, Part II. LNCS, vol. 9723, pp. 223–239. Springer, Cham (Jul 2016). https://doi.org/10.1007/978-3-319-40367-0_14
20. Libert, B., Paterson, K.G., Quaglia, E.A.: Anonymous broadcast encryption: Adaptive security and efficient constructions in the standard model. In: Fischlin, M., Buchmann, J., Manulis, M. (eds.) PKC 2012. LNCS, vol. 7293, pp. 206–224. Springer, Berlin, Heidelberg (May 2012). https://doi.org/10.1007/978-3-642-30057-8_13
21. Ling, S., Phan, D.H., Stehlé, D., Steinfeld, R.: Hardness of k-LWE and applications in traitor tracing. In: Garay, J.A., Gennaro, R. (eds.) CRYPTO 2014, Part I. LNCS, vol. 8616, pp. 315–334. Springer, Berlin, Heidelberg (Aug 2014). https://doi.org/10.1007/978-3-662-44371-2_18
22. Persiano, G., Phan, D.H., Yung, M.: Anamorphic encryption: Private communication against a dictator. In: Dunkelman, O., Dziembowski, S. (eds.) EUROCRYPT 2022, Part II. LNCS, vol. 13276, pp. 34–63. Springer, Cham (May / Jun 2022). https://doi.org/10.1007/978-3-031-07085-3_2
23. Sakai, R., Ohgishi, K., Kasahara, M.: Cryptosystems based on pairing. In: SCIS 2000. Okinawa, Japan (Jan 2000)
24. Shamir, A.: Identity-based cryptosystems and signature schemes. In: Blakley, G.R., Chaum, D. (eds.) CRYPTO’84. LNCS, vol. 196, pp. 47–53. Springer, Berlin, Heidelberg (Aug 1984). https://doi.org/10.1007/3-540-39568-7_5