(One) failure is not an option Bootstrapping the search for failures in lattice-based encryption schemes

Jan-Pieter D'Anvers[1*] and Mélissa Rossi[2,3,4,5†] and Fernando Virdia[6‡]

[1] imec-COSIC, KU Leuven, Belgium
janpieter.danvers@esat.kuleuven.be
[2] ANSSI, Paris, France
[3] ENS Paris, CNRS, PSL University, Paris, France
[4] Thales, Gennevilliers, France
[5] INRIA, Paris, France
melissa.rossi@ens.fr
[6] Information Security Group, Royal Holloway, University of London, UK
fernando.virdia.2016@rhul.ac.uk

Abstract. Lattice-based encryption schemes are often subject to the possibility of decryption failures, in which valid encryptions are decrypted incorrectly. Such failures, in large number, leak information about the secret key, enabling an attack strategy alternative to pure lattice reduction. Extending the "failure boosting" technique of D'Anvers et al. in PKC 2019, we propose an approach that we call "*directional* failure boosting" that uses previously found "failing ciphertexts" to accelerate the search for new ones. We analyse in detail the case where the lattice is defined over polynomial ring modules quotiented by $\langle X^N + 1\rangle$ and demonstrate it on a simple Mod-LWE-based scheme parametrized *à la* Kyber768/Saber. We show that for a given secret key (single-target setting), the cost of searching for additional failing ciphertexts after one or more have already been found, can be sped up dramatically. We thus demonstrate that, in this single-target model, these schemes should be designed so that it is hard to even obtain one decryption failure. Besides, in a wider security model where there are many target secret keys (multi-target setting), our attack greatly improves over the state of the art.

Keywords: cryptanalysis, lattice-based cryptography, reaction attacks, decryption errors

1 Introduction

Algebraic lattices are a powerful tool in cryptography, enabling the many sophisticated constructions such as digital signatures [42,6], zero-knowledge proofs [44,48], FHE [26,28,9] , IBE [27], ABE [55] and others. Applications of main interest are public-key encryptions (PKE) [49,43] and key encapsulation mechanisms (KEM).

The computational problems defined over lattices are believed to be hard to solve, even with access to large-scale quantum computers, and hence many of these constructions are considered to be

*The research of D'Anvers was supported the European Commission through the Horizon 2020 research and innovation programme Cathedral ERC Advanced Grant 695305, by the CyberSecurity Research Flanders with reference number VR20192203 and by the Semiconductor Research Corporation (SRC), under task 2909.001.

†The research of Rossi was supported by the European Union's H2020 Programme under PROMETHEUS project (grant 780701)and by ANRT under the program CIFRE N 2016/1583. It was also supported by the French Programme d'Investissement d'Avenir under national project RISQ P14158.

‡The research of Virdia was supported by the EPSRC and the UK government as part of the Centre for Doctoral Training in Cyber Security at Royal Holloway, University of London (EP/P009301/1).

quantum-safe. As industry starts to make steps forward into the concrete development of small quantum computers [39,38,33], the US National Institute of Standards and Technology (NIST) begun an open standardization effort, with the aim of selecting quantum-safe schemes for public-key encryption and digital signatures [46]. At the time of writing, the process is in its second round, and 9 out of 17 candidates for PKE or KEM base their security on problems related to lattices, with or without special structure.

One commonly occurring characteristic of lattice-based PKE or KEM schemes is that of lacking perfect correctness. This means that sometimes, ciphertexts generated honestly using a valid public key may lead to decryption failures under the corresponding private key. Throughout this paper we'll refer to such ciphertexs as "failures", "decryption failures", or "failing ciphertexts". While in practice, schemes are parametrised in such a way that decryption failures do not undermine overall performance, these can be leveraged as a vehicle for key recovery attacks against the key pair used to generate them. Such an attack was described by Jaulmes and Joux [34] against NTRU, after which is was extended in [32] and [25]. A similar attack on Ring-LWE based schemes was later presented by Fluhrer [22] and extended by Băetu et. al [5].

However, the aforementioned attacks all use specially crafted ciphertexts and can therefore be prevented with a transformation that achieves chosen ciphertext security. This can for example be obtained by means of an off-the-shelf compiler [23,31] that stops the adversary from being able to freely malleate honestly generated ciphertexts.

The NIST Post-Quantum Standardization Process candidate Kyber [8] noted that it was possible to search for ciphertexts with higher failure probability than average. D'Anvers et al. [16] extended this idea to an attack called "failure boosting", where ciphertexts with higher failure probability are generated to speedup the search for decryption failures, and provided an analysis of the effectiveness of the attack on several NIST candidates. At the same time, Guo et al. [30] described an adaptive attack against the IND-CCA secure ss-ntru-pke variant of NTRUEncrypt [11], which used an adaptive search for decryption failures exploiting information from previously collected ciphertexts.

Our contributions. In this paper, we present a novel attack technique called "*directional* failure boosting", aimed at enhancing the search for decryption failures in public-key encryption schemes based on the protocol by Lyubashevsky et al. [43], in the single-target setting. Our technique is an improvement of the "failure boosting" technique of D'Anvers et al. [16].

We consider a simple (but realistically parametrized) scheme based on the Mod-LWE problem as a case study and make some necessary orthogonality and independance assumptions that are reasonable in our range of parameters. We show that in this setting, the work and number of decryption queries needed to obtain multiple failing ciphertexts is only marginally larger than those necessary to obtain the first decryption failure. For example, obtaining 30 decryption failures requires only 25% more quantum work and only 58% more queries than obtaining one decryption failure. As previously shown in [16] and [30], we recall that having many decryption failures enables more efficient lattice reduction which leads to key recovery attacks. As a result, we conclude that when protecting against decryption failure attacks, in the single target setting, designers should make sure that an adversary can not feasibly obtain even a single decryption failure.

Our attack outperforms previously proposed attacks based on decryption failures. In particular, it improves over the multitarget attack of Guo et al. [30] on ss-ntru-pke, lowering the attack's quantum complexity from $2^{139.5}$ to $2^{96.6}$.

Paper outline. In §2, we introduce some preliminaries about notation and structures. In §3, we describe the general idea of lattice-based encryption and how decryption failures are generated. In §4, we recall the original failure boosting technique from [13]. In §5, we describe our directional failure boosting technique. In §6, we show[7] how this method impacts the total work and queries overhead. Finally in §7, we discuss the results by comparing them with the literature and conclude with possible future work.

2 Preliminaries

Let $\mathbb{Z}_q$ be the ring of integers modulo q. For N a power of 2, we define R_q the ring $\mathbb{Z}_q[X]/(X^N+1)$, and $R_q^{l_1 \times l_2}$ the ring of $l_1 \times l_2$ matrices over R_q. Vectors and polynomials will be indicated with bold lowercase letters, eg. $\mathbf{v}$, while matrices will be written in bold uppercase letters, eg. $\mathbf{M}$. Denote with $\lfloor \cdot \rfloor$ flooring to the nearest lower integer, and with $\lfloor \cdot \rceil$ rounding to the nearest integer. These operations are extended coefficient-wise for vectors and polynomials. Throughout, we abuse notation and identify elements in $\mathbb{Z}_q$ with their representatives in $[-q/2, q/2)$, and elements in R_q with their representatives of degree $< N$, with index i indicating the coefficient of X^i. This allows us to define the ℓ_2-norm $\|\mathbf{x}\|_2$ of a polynomial $\mathbf{x} \in R_q$, so that $\|\mathbf{x}\|_2 = \sqrt{\sum_i \mathbf{x}_i^2}$ where $\mathbf{x}_i \in [-q/2, q/2)$, and extend this to vectors of polynomials $\mathbf{y} \in R_q^{l \times 1}$ as $\|\mathbf{y}\|_2 = \sqrt{\sum_i \|\mathbf{y}_i\|_2^2}$. Identically, we define and extend the ℓ_∞-norm.

Let $x \leftarrow X$ denote sampling x according to the probability distribution X. We extend this notation for coefficient-wise sampling of a vector $\mathbf{x} \in R_q^{l \times 1}$ as $\mathbf{x} \leftarrow X(R_q^{l \times 1})$, and similarly for a matrix. We denote with $\mathbf{x} \leftarrow X(R_q^{l \times 1}; r)$ sampling $\mathbf{x} \in R_q^{l \times 1}$ pseudorandomly from the seed r with each coefficient following the distribution X. In algorithms, we also use $x \leftarrow Alg()$ to mean that the value x is assigned to be the output of a probabilistic algorithm Alg.

Let $\mathcal{U}$ be the uniform distribution over $\mathbb{Z}_q$ and let $\mathcal{N}_{\mu,\sigma}$ be the normal distribution with mean μ and standard deviation σ, so that the probability density function of $x \leftarrow \mathcal{N}_{\mu,\sigma}$ is defined as:

$$f_{\mathcal{N}_{\mu,\sigma}}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}. \tag{1}$$

The discrete Gaussian distribution $\mathcal{D}_{\mu,\sigma}$ is a discrete restriction to $\mathbb{Z}_q$ of $\mathcal{N}_{\mu,\sigma}$, so that an integer x is sampled with a probability proportional to $e^{-(x-\mu)^2/2\sigma^2}$ and its remainder modulo q in $[-q/2, q/2)$ is returned.

For an event A we define $P[A]$ as its probability. For an element which does not correspond to an event, a ciphertext ct for example, we abusively write $P[ct]$ to denote the probability of the event $ct' = ct$ where ct' is drawn from a distribution which will be clear in the context. We will denote with $\mathbb{E}[A]$ the expected value of a variable drawn from a distribution A.

Security definitions. Let $\Pi = (\mathsf{KeyGen}, \mathsf{Enc}, \mathsf{Dec})$ be a public-key encryption scheme, with message space $\mathcal{M}$, and let $K = (\mathsf{KeyGen}, \mathsf{Encaps}, \mathsf{Decaps})$ be a key encapsulation mechanism (KEM). When a decapsulation or a decryption oracle is provided, we assume that the maximum number of ciphertexts that can be queried to it for each key pair is 2^K; in practice, $K = 64$ is often considered [46, §4.A.2]. In this work, we keep the maximum number of queries as a parameter with no specific value, in order to provide a better granularity in the security assessement. Indeed, to mount an attack, the adversary trades off between number of queries and the work.

Definition 1 (IND-CPA$_{A,\Pi}(k)$ game [37]). *Let A be an adversary and $\Pi = (\mathsf{KeyGen}, \mathsf{Enc}, \mathsf{Dec})$ be a public-key encryption scheme. The experiment IND-CPA$_{A,\Pi}(1^k)$ runs as follows:*

1. $(pk, sk) \leftarrow \mathsf{KeyGen}(1^k)$
2. *A is given pk. After evaluating $\mathsf{Enc}(pk, \cdot)$ as desired, it outputs $(m_0, m_1) \in \mathcal{M} \times \mathcal{M}$.*
3. *A random bit $b \leftarrow_{\$} \{0,1\}$ is sampled, and $c \leftarrow \mathsf{Enc}(pk, m_b)$ is passed to A.*
4. *A keeps evaluating $\mathsf{Enc}(pk, \cdot)$ as desired, until it returns a bit b'.*
5. *The experiment outputs 1 if $b = b'$ and 0 otherwise.*

[7] The software is available at: https://github.com/KULeuven-COSIC/PQCRYPTO-decryption-failures.

Definition 2 (IND-CCA$_{A,K}(k)$ game [37]). *Let A be an adversary and $K = (\mathsf{KeyGen}, \mathsf{Encaps}, \mathsf{Decaps})$ be a key encapsulation mechanism. The experiment IND-CCA$_{A,K}(1^k)$ runs as follows:*

1. $(pk, sk) \leftarrow \mathsf{KeyGen}(1^k)$
2. $(c, k) \leftarrow \mathsf{Encaps}(\mathsf{pk})$
3. $b \leftarrow_{\$} \{0,1\}$. *If* $b = 0$, *set* $\hat{k} = k$, *else let* $\hat{k} \leftarrow \{0,1\}^n$.
4. *A is given $(pk, c, \hat{k})$, and access to a decapsulation oracle* $\mathsf{Decaps}(sk, \cdot)$. *After evaluating* $\mathsf{Encaps}(pk, \cdot)$ *and querying* $\mathsf{Decaps}(sk, \cdot)$ *as desired (except for decapsulation queries on c), it returns $b' \in \{0,1\}$.*
5. *The experiment outputs 1 if $b = b'$ and 0 otherwise.*

Definition 3 (PKE and KEM security [24]). *A public-key encryption scheme Π (resp. a key encapsulation mechanism K) is (t, ϵ)-GAME secure if for every t-time adversary A, we have that*

$$\left|\Pr[GAME_{A,\Pi}(k) = 1] - \frac{1}{2}\right| \le \epsilon \quad \left(\textit{resp.} \ \left|\Pr[GAME_{A,K}(k) = 1] - \frac{1}{2}\right| \le \epsilon \right)$$

For a security parameter 1^k, we usually mean $t \approx poly(k)$ and $\epsilon \le negl(k)$. If GAME is IND-CPA (resp. IND-CCA) we say that Π (resp. K) is (t, ϵ)-secure against chosen-plaintext attacks (resp. (t, ϵ)-secure against adaptive chosen-ciphertext attacks).

3 Lattice-based encryption

The Module-LWE (or Mod-LWE) problem [40] is a mathematical problem that can be used to build cryptographic primitives such as encryption [7,14], key exchange [14] and signatures [20]. It is a generalization of both the Learning With Errors (or LWE) problem [49], and the Ring-LWE problem [53,43].

Definition 4 (Mod-LWE [40]). *Let n, q, k be positive integers, χ be a probability distribution on $\mathbb{Z}$ and $\mathbf{s}$ be a secret module element in R_q^k. We denote by $\mathcal{L}$ the probability distribution on $R_q^k \times R_q$ obtained by choosing $\mathbf{a} \in R_q^k$ uniformly at random, choosing $e \in R$ by sampling each of its coefficients according to χ and considering it in R_q, and returning $(\mathbf{a}, c) = (\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e) \in R_q^k \times R_q$.*
Decision-Mod-LWE *is the problem of deciding whether pairs $(\mathbf{a}, c) \in R_q^k \times R_q$ are sampled according to $\mathcal{L}$ or the uniform distribution on $R_q^k \times R_q$.*
Search-Mod-LWE *is the problem of recovering $\mathbf{s}$ from $(\mathbf{a}, c) = (\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e) \in R_q^k \times R_q$ sampled according to $\mathcal{L}$.*

3.1 Passively and actively secure encryption

Lyubashevsky et al. [43] introduced a simple protocol to build passively secure encryption from the Ring-LWE problem, inspired by Diffie-Hellman key exchange [19] and ElGamal public-key encryption [21]. Naturally, the protocol can also be adapted to work based on plain and Module LWE assumptions. A general extension of the protocol for all aforementioned assumptions is described in Algorithms 1, 2, and 3, where $r \in \mathcal{R} = \{0,1\}^{256}$, and where the message space is defined as $\mathcal{M} = \{\text{polynomials in } R_q \text{ with coefficients in } \{0,1\}\}$.

In order to obtain active security, designers usually use an off-the-shelf CCA compiler, usually a (post-quantum) variant [18,31,54,50,35] of the Fujisaki-Okamoto transform [23] (FO). These come with proofs of security in the (quantum) random oracle model, with explicit bounds about the loss of security caused by the transformation. We show such transformed KEM Decapsulation and Encapsulation in Algorithms 4 and 5.

Algorithm 1: $\texttt{PKE.KeyGen}()$

$\mathbf{A} \leftarrow \mathcal{U}(R_q^{l \times l})$
$\mathbf{s}, \mathbf{e} \leftarrow \mathcal{D}_{0,\sigma_s}(R_q^{l \times 1}) \times \mathcal{D}_{0,\sigma_e}(R_q^{l \times 1})$
$\mathbf{b} := \mathbf{A}\mathbf{s} + \mathbf{e}$
return $(pk = (\mathbf{b}, \mathbf{A}), sk = \mathbf{s})$

Algorithm 2: $\texttt{PKE.Enc}(pk = (\mathbf{b}, \mathbf{A}), \mathbf{m} \in \mathcal{M}; r)$

$\mathbf{s}', \mathbf{e}' \leftarrow \mathcal{D}_{0,\sigma_s}(R_q^{l \times 1}; r) \times \mathcal{D}_{0,\sigma_e}(R_q^{l \times 1}; r)$
$\mathbf{e}'' \leftarrow \mathcal{D}_{0,\sigma_e}(R_q; r)$
$\mathbf{b}' := \mathbf{A}^T\mathbf{s}' + \mathbf{e}'$
$\mathbf{v}' := \mathbf{b}^T\mathbf{s}' + \mathbf{e}'' + \lfloor q/2 \rfloor \cdot \mathbf{m}$
return $ct = (\mathbf{v}', \mathbf{b}')$

Algorithm 3: $\texttt{PKE.Dec}(sk = \mathbf{s}, ct = (\mathbf{v}', \mathbf{b}'))$

$\mathbf{m}' := \lfloor \lfloor 2/q \rceil (\mathbf{v}' - \mathbf{b}'^T\mathbf{s}) \rceil$
return $\mathbf{m}'$

Algorithm 4: $\texttt{KEM.Encaps}(pk)$

$m \leftarrow \mathcal{U}(\{0,1\}^{256})$
$(\overline{K}, r) := \mathcal{G}(pk, m)$
$ct := \texttt{PKE.Enc}(pk, m, r)$
$K := \mathcal{H}(\overline{K}, r)$
return (ct, K)

In the case of FO for lattice-based schemes, the randomness used during the encryption is generated by submitting the message (and sometimes also the public key) to a random oracle. As this procedure is repeatable with knowledge of the message, one can check the validity of ciphertexts during decapsulation. Hence, an adversary wanting to generate custom ephemeral secrets $\mathbf{s}', \mathbf{e}', \mathbf{e}''$ in order to fabricate weak ciphertexts, would need to know a preimage of the appropriate random coins for the random oracle. Therefore, their only option is to mount a (Grover's) search by randomly generating ciphertexts corresponding to different messages $\mathbf{m}$, and testing if their predicted failure probability is above a certain threshold.

Remark 1. Several lattice-based candidates submitted to the NIST Post-Quantum Cryptography Standardization Process use a variant of the protocol by Lyubashevsky et al. [43]. Deviating from the original design, most candidates perform an additional rounding of the ciphertext $\mathbf{v}'$, in order to reduce bandwidth. The designers of New Hope [3] and LAC [41] choose to work directly over rings (or equivalently, they choose a module of rank $l = 1$) and add error correction on the encapsulated message, while the designers of Kyber [7] and Saber [14] choose a module of rank $l > 1$ and perform an additional rounding of $\mathbf{b}'$ (and $\mathbf{b}$ in case of Saber). We here focus on the basic version given in Algorithms 1 to 3 and leave the study of the effect of compression to further work.

We selected the parameters of the studied encryption scheme to ensure a similar failure probability and security to Kyber and Saber. These parameters can be found in Table 1. The security estimates are generated using the Core-SVP methodology [3] and the LWE estimator[8] [2], while the failure probability of Kyber and Saber is given as reported in their respective the NIST round 2 documentations [52,15]. The failure probability of our chosen parameters is determined by calculating the variance of the error term and assuming the distribution to be Gaussian.

Remark 2. We do not consider the case of "plain" LWE based schemes like FrodoKEM [45] or Round5 [4]. Nonetheless, we believe that the attack methodology would easily translate to the LWE setting as the failure condition and the failure probabilities are similar to the investigated case.

[8] The estimator can be found at https://bitbucket.org/malb/lwe-estimator/.

Algorithm 5: `KEM.Decaps`(sk, pk, ct, K)

$m' :=$ `PKE.Dec`(sk, ct)
$(\overline{K}, r') := \mathcal{G}(pk, m')$
$ct' :=$ `PKE.Enc`$(pk, m'; r')$
if $ct = ct'$ **then**
  return $K := (\overline{K}, r')$
else
  return $K := \perp$ `// Could return a pseudo-random string to implicitly reject`

	l	N	q	σ_s	σ_e	$P[F]$	Classical	Quantum
Chosen parameters	3	256	8192	2.00	2.00	2^{-119}	2^{195}	2^{177}
Saber	3	256	8192	1.41	2.29	2^{-136}	2^{203}	2^{185}
Kyber 768	3	256	3329	1.00	1.00/1.38†	2^{-164}	2^{181}	2^{164}

† Standard deviation of the error term in the public key and ciphertext respectively

Table 1. Comparison between our target scheme and Saber and Kyber 768, as parametrised in Round 2 of the NIST PQC standardization process. The classical (resp. quantum) security is evaluated using the Core-SVP [3] methodology, assuming the cost of BKZ with block size β to be $2^{0.292\beta}$ (resp. $2^{0.265\beta}$).

3.2 Decryption failures

Following the execution of the protocol, both messages $\mathbf{m}'$ and $\mathbf{m}$ are the same if the coefficients of the error term $\mathbf{e}^T\mathbf{s}' - \mathbf{s}^T\mathbf{e}' + \mathbf{e}''$ are small enough; more exactly if $\|\mathbf{e}^T\mathbf{s}' - \mathbf{s}^T\mathbf{e}' + \mathbf{e}''\|_\infty \leq q/4$. This expression can be simplified by defining the vector $\mathbf{S}$ as the vertical concatenation of $-\mathbf{s}$ and $\mathbf{e}$, the vector $\mathbf{C}$ as the concatenation of $\mathbf{e}'$ and $\mathbf{s}'$, and by replacing $\mathbf{e}''$ with $\mathbf{G}$, as shown below:

$$\mathbf{S} = \begin{bmatrix} -\mathbf{s} \\ \mathbf{e} \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} \mathbf{e}' \\ \mathbf{s}' \end{bmatrix} \quad \mathbf{G} = \mathbf{e}''. \tag{2}$$

Here, $\mathbf{S}$ contains the secret elements of the secret key, and $\mathbf{C}$ and $\mathbf{G}$ consist of elements used to construct the ciphertexts[9]. Using these vectors, the error expression can be rewritten: a failure occurs when $\|\mathbf{S}^T\mathbf{C} + \mathbf{G}\|_\infty > q/4$.

The standard deviation of the terms in the polynomial $\mathbf{S}^T\mathbf{C}$ equals $\sqrt{2N}\sigma_s\sigma_e$, versus a standard deviation of σ_e for the terms of $\mathbf{G}$. Therefore, the influence of $\mathbf{G}$ on the failure rate is limited, i.e. $\|\mathbf{S}^T\mathbf{C} + \mathbf{G}\|_\infty \approx \|\mathbf{S}^T\mathbf{C}\|_\infty$. Let $q_t := q/4$ denote the failure threshold, we will use

$$\|\mathbf{S}^T\mathbf{C}\|_\infty > q_t \tag{3}$$

as an approximation of the failure expression throughout our analysis. However, with some extra work, one can rewrite a more accurate Equation 3 as $\|\mathbf{S}^T\mathbf{C}\|_\infty > q_t - \|\mathbf{G}\|_\infty$, and instead of considering q_t to be fixed, taking the distribution of $q_t - \|\mathbf{G}\|_\infty$ as shown in [16]. For the ease of the implementation and due to the low influence of $\mathbf{G}$ on the failure rate, we prefer to stick with Equation 3. We now introduce a more handy way of writing the failure condition (Equation 3) by only using vectors in $\mathbb{Z}_q$.

[9]When talking about ciphertexts throughout the paper, we will sometimes refer to their underlying elements $\mathbf{C}$ and $\mathbf{G}$.

Definition 5 (Coefficient vector). *For $\mathbf{S} \in R_q^{l\times 1}$, we denote by $\overline{\mathbf{S}} \in \mathbb{Z}_q^{lN\times 1}$, the representation of $\mathbf{S}$ where each polynomial is decomposed as a list of its coefficients in*[10] $\mathbb{Z}_q$.

Definition 6 (Rotations). *For $r \in \mathbb{Z}$ and $\mathbf{C} \in R_q^{l\times 1}$, we denote by $\mathbf{C}^{(r)} \in R_q^{l\times 1}$, the following vector of polynomials*

$$\mathbf{C}^{(r)} := X^r \cdot \mathbf{C}(X^{-1}) \mod X^N + 1.$$

Correspondingly, $\overline{\mathbf{C}^{(r)}} \in \mathbb{Z}_q^{lN\times 1}$ denotes its coefficient vector.

It is easy to show that $\overline{\mathbf{C}^{(r)}}$ is constructed as to ensure that for $r \in [0, ..., N-1]$, the r^{th} coordinate of $\mathbf{S}^T\mathbf{C}$ is given by the scalar product $\overline{\mathbf{S}}^T\overline{\mathbf{C}^{(r)}}$. In other words, one is now able to decompose $\mathbf{S}^T\mathbf{C}$ as a sum of scalar products:

$$\mathbf{S}^T\mathbf{C} = \sum_{r\in[0,N-1]} \overline{\mathbf{S}}^T\overline{\mathbf{C}^{(r)}} \cdot X^r. \tag{4}$$

One can observe that this construction is only valid for the modulo X^N+1 ring structure, but it could be adapted for other ring structures. Note that for any $r \in \mathbb{Z}$, $\mathbf{C}^{(r+N)} = -\mathbf{C}^{(r)}$ and $\mathbf{C}^{(r+2N)} = \mathbf{C}^{(r)}$. Besides, taking into account the extension of the norms to vectors of polynomials (defined in §2), one can make the following remark.

Remark 3. Note that for any $r \in \mathbb{Z}$, $\|\overline{\mathbf{C}^{(r)}}\|_2 = \|\overline{\mathbf{C}}\|_2 = \|\mathbf{C}\|_2$ and $\|\overline{\mathbf{C}^{(r)}}\|_\infty = \|\overline{\mathbf{C}}\|_\infty = \|\mathbf{C}\|_\infty$.

The decomposition in Equation 4 will allow a geometric interpretation of the failures as it will be shown in the rest of the paper. First, let us introduce a brief example to illustrate Definitions 5 and 6.

Example 1. For a secret $\mathbf{S}$ and a ciphertext $\mathbf{C}$ in $\mathbb{Z}_q^{2\times 1}[X]/(X^3+1)$:

$$\mathbf{S} = \begin{bmatrix} s_{0,0} + s_{0,1}X + s_{0,2}X^2 \\ s_{1,0} + s_{1,1}X + s_{1,2}X^2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} c_{0,0} + c_{0,1}X + c_{0,2}X^2 \\ c_{1,0} + c_{1,1}X + c_{1,2}X^2 \end{bmatrix} \tag{5}$$

we get the following vectors:

$$\overline{\mathbf{S}} = \begin{bmatrix} s_{0,0} \\ s_{0,1} \\ s_{0,2} \\ s_{1,0} \\ s_{1,1} \\ s_{1,2} \end{bmatrix}, \quad \overline{\mathbf{C}^{(0)}} = \begin{bmatrix} c_{0,0} \\ -c_{0,2} \\ -c_{0,1} \\ c_{1,0} \\ -c_{1,2} \\ -c_{1,1} \end{bmatrix} \quad \overline{\mathbf{C}^{(1)}} = \begin{bmatrix} c_{0,1} \\ c_{0,0} \\ -c_{0,2} \\ c_{1,1} \\ c_{1,0} \\ -c_{1,2} \end{bmatrix} \quad \overline{\mathbf{C}^{(2)}} = \begin{bmatrix} c_{0,2} \\ c_{0,1} \\ c_{0,0} \\ c_{1,2} \\ c_{1,1} \\ c_{1,0} \end{bmatrix} \quad \overline{\mathbf{C}^{(3)}} = \begin{bmatrix} -c_{0,0} \\ c_{0,2} \\ c_{0,1} \\ -c_{1,0} \\ c_{1,2} \\ c_{1,1} \end{bmatrix} \ldots$$

In case of a failure event, $\mathbf{S}^T\mathbf{C}$ satisfies Equation 3. Therefore, at least one element among all the coefficients

$$\overline{\mathbf{S}}^T \cdot \overline{\mathbf{C}^{(0)}}, \ldots, \overline{\mathbf{S}}^T \cdot \overline{\mathbf{C}^{(2N-1)}}$$

is larger than q_t.

Definition 7 (Failure event). *A failure event will be denoted with F, while we use S to indicate a successful decryption.*
More precisely, for $r \in [0, 2N-1]$, we denote by F_r the failure event where

$$\overline{\mathbf{S}}^T \cdot \overline{\mathbf{C}^{(r)}} > q_t.$$

The event F_r gives a twofold information: it provides the location of the failure in the $\mathbf{S}^T\mathbf{C}$ polynomial and it also provides the sign of the coefficient that caused the failure.

[10] Recall that, in this paper, all the elements in $\mathbb{Z}_q$ are represented as integers belonging in $[-q/2, q/2]$.

An assumption on the failing ciphertexts. In the rest of the paper, in order to predict the results of our attack, we will make the following orthogonality assumption.

Assumption 1 *Let $n \ll 2Nl$, and $\mathbf{C}_0$, $\cdots$, $\mathbf{C}_n$ be ciphertexts that lead to failure events F_{r_0}, $\cdots$, F_{r_n}. The vectors $\overline{\mathbf{C}_0^{(r_0)}}$, $\cdots$, $\overline{\mathbf{C}_n^{(r_n)}}$ are considered orthogonal when projected on the hyperplane orthogonal to $\overline{\mathbf{S}}$.*

This assumption is an approximation that is supported by the fact that vectors in high dimensional space have a strong tendency towards orthogonality, as can be seen in Figure 2.

4 Failure Boosting attack technique

Failure boosting is a technique introduced in [16] to increase the failure rate of (Ring/Mod)-LWE/LWR based schemes by honestly generating ciphertexts and only querying weak ones, i.e. those that have a failure probability above a certain threshold $f_t > 0$. This technique is especially useful in combination with Grover's algorithm [29], in which case the search for weak ciphertexts can be sped up quadratically. Failure boosting consists of two phases: a precomputation phase, and a phase where the decryption oracle is queried.

Precomputation phase. The adversary does an offline search for weak ciphertexts with the following procedure:

1. Generate a key encapsulation[9] $ct = (\mathbf{C}, \mathbf{G})$.
2. If $P[F \mid ct] \geq f_t$, keep ct in a weak ciphertext list, otherwise go to Step 1.

In Step 2, $P[F \mid ct]$ is defined as the failure probability given a certain ciphertext ct. It is computed as follows.

$$P[F \mid ct] := \sum_{\mathbf{S}} P\left[\left\|\mathbf{S}^T\mathbf{C} + \mathbf{G}\right\|_\infty > q_t \mid \mathbf{S}\right] \cdot P[\mathbf{S}] \tag{6}$$

Given the probability of generating ciphertexts $P[ct] = P[\mathbf{C}, \mathbf{G}]$, the probability of finding such a weak ciphertext can be expressed as follows:

$$\alpha_{f_t} = \sum_{\forall ct: P[F|ct] > f_t} P[ct]. \tag{7}$$

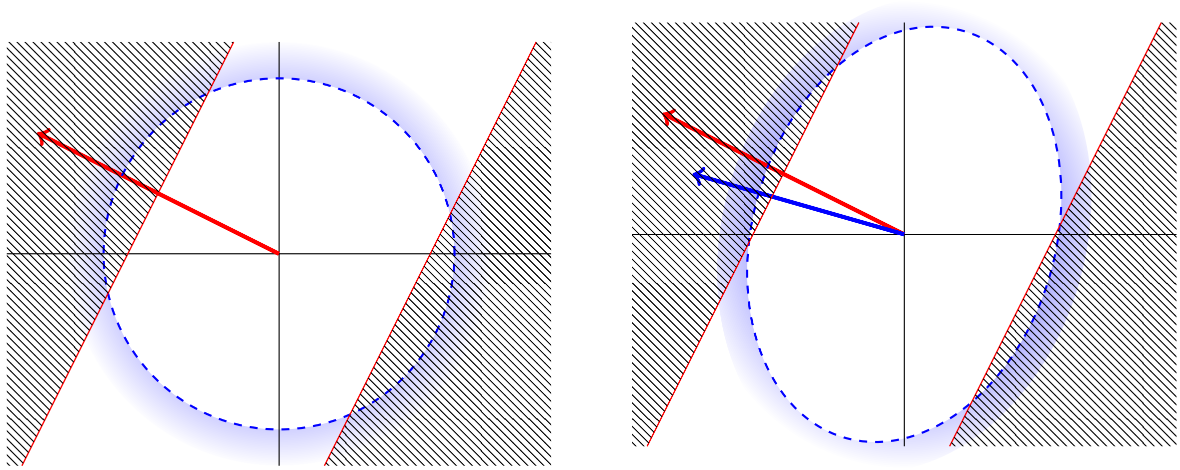
An adversary thus needs to perform on average $\alpha_{f_t}^{-1}$ work to obtain one weak ciphertext, or $\sqrt{\alpha_{f_t}^{-1}}$ assuming Grover's search achieves a full speed-up.

Decryption oracle query phase. After the precomputation phase, an adversary has a probability β_{f_t} that a weak ciphertext results in a failure, where β_{f_t} can be calculated as a weighted average of the failure probabilities of weak ciphertexts:

$$\beta_{f_t} = \frac{\sum_{\forall ct: P[F|ct] > f_t} P[ct] \cdot P[F|ct]}{\sum_{\forall ct: P[F|ct] > f_t} P[ct]}. \tag{8}$$

Thus to obtain one decryption failure with probability $1 - e^{-1}$, an adversary needs to perform approximately $\beta_{f_t}^{-1}$ queries and therefore $\alpha_{f_t}^{-1}\beta_{f_t}^{-1}$ work (or $\sqrt{\alpha_{f_t}^{-1}}\beta_{f_t}^{-1}$ using a quantum computer).

The better an adversary can predict $P[F|ct]$, the more efficient failure boosting will be. Having no information about the secret except its distribution, an adversary is bound to standard failure boosting, where the failure probability is estimated based on $\|\mathbf{C}\|_2$ and $\|\mathbf{G}\|_2$. For a graphical intuition, a two



(a) Without directional information, as in [16], the weak ciphertexts (in blue) are defined as the ciphertexts with a probability higher than f_t.

(b) With directional information, the weak ciphertexts (in blue) are found according to a refined acceptance criterion, here represented as an ellipse.

Fig. 1. Simplified diagram trying to provide an intuition on the effect of directional failure boosting. The red arrow represents the secret vector $\overline{\mathbf{S}}$. Ciphertexts with $\overline{\mathbf{C}}$ that lie in the dashed area will provoke a failure as the inner product with $\overline{\mathbf{S}}$ will exceed the threshold q_t. Ciphertexts outside the blue circle are considered weak.

dimensional toy example is depicted in Figure 1a below, where the red arrow represents the secret vector $\overline{\mathbf{S}}$. Ciphertexts with $\overline{\mathbf{C}}$ that lie in the dashed area will provoke a failure as the inner product with $\overline{\mathbf{S}}$ will exceed the threshold q_t. The blue circle is a circle of ciphertexts that have a certain failure probability f_t as estimated by an adversary who does not know the secret. During the failure boosting procedure, we will generate random ciphertexts, and only select the ciphertexts with a higher failure probability than f_t, i.e. that are outside the blue circle. One can graphically see in Figure 1a that these ciphertexts will have a higher failure probability and a higher norm. We refer to [16] for a full description of the failure boosting technique. Note that Figure 1a is an oversimplified 2-dimension example that does not take into account the polynomial structure and the high dimensionality of the space.

5 Directional Failure Boosting

Once $n \geq 1$ decryption failures $\mathbf{C}_0, \ldots, \mathbf{C}_{n-1}$ are found, additional information about the secret key $\mathbf{S}$ becomes available, and can be used to refine the failure estimation for new ciphertexts and thus speed up failure boosting. We now introduce an iterative two-step method to perform directional failure boosting.

Step 1 An estimate, denoted $\overline{\mathbf{E}}$, of the 'direction' of the secret $\overline{\mathbf{S}}$ in $\mathbb{Z}_q^{lN}$ is obtained from $\mathbf{C}_0, \ldots, \mathbf{C}_{n-1}$.

Step 2 The estimate $\overline{\mathbf{E}}$ is used to inform the search for weak ciphertexts and improve the failure probability prediction for a new ciphertext $\mathbf{C}_n$. One is able to refine the criterion $P[F \mid ct] \geq f_t$ with computing $P[F \mid ct, \overline{\mathbf{E}}] \geq f_t$ instead.

Once new failing ciphertexts are found in step 2, one can go back to step 1 and improve the estimate $\overline{\mathbf{E}}$ and thus bootstrap the search for new failures.

To give an intuition, a two dimensional toy representation can be found in Figure 1b. Like in the classical failure boosting technique, the red arrow depicts the secret $\overline{\mathbf{S}}$, while the additional blue arrow marks estimate $\overline{\mathbf{E}}$ (as calculated in step 1, see §5.2). Using this estimate, we can refine the acceptance criterion to the depicted ellipse to better reflect our knowledge about the secret (step 2, see §5.3). Ciphertexts outside this ellipse will be flagged as weak ciphertexts, and while the probability of finding such a ciphertext is the same, the failure probability of weak ciphertexts is now higher. As in, more of the blue zone lies in the dashed area.

5.1 Distributions

We now introduce some probability distributions that will be useful in following sections.

Scaled χ-distribution. The scaled χ-distribution $\chi_{n,\sigma}$ is the distribution of the ℓ_2-norm of a vector with n coefficients, each following the normal distribution $\mathcal{N}_{0,\sigma}$. Denoting with Γ the gamma function, the probability density function of $\chi_{n,\sigma}$ is given by:

$$f_{\chi_{n,\sigma}}(x) = \frac{\left(\frac{x}{\sigma}\right)^{n-1} e^{-\frac{x^2}{2\sigma^2}}}{2^{\left(\frac{n}{2}-1\right)}\Gamma\left(\frac{n}{2}\right)} \quad \text{for } x \geq 0, \tag{9}$$

which has mean [36, §18.3] $\mathbb{E}_\chi[x] = \sqrt{2}\frac{\Gamma((n+1)/2)}{\Gamma(n/2)}\sigma \approx \sqrt{n}\sigma$.

We will approximate the probability distribution of $\|\mathbf{x}\|_2$ where $\mathbf{x} \leftarrow \mathcal{D}_{0,\sigma}(R_q^{l\times 1})$ with a discretized version of the $\chi_{(l\cdot N),\sigma}$-distribution, which will be denoted with $\chi^D_{(l\cdot N),\sigma}$. Using this distribution, the probability density function of $\|\mathbf{x}\|_2$ is calculated as:

$$P[\|\mathbf{x}\|_2 = x] = C \cdot \left(\frac{x}{\sigma}\right)^{l\cdot N-1} e^{-\frac{x^2}{2\sigma^2}} \quad \text{for } x \in \left\{0, \dots, \left\lfloor \frac{q}{2}\sqrt{lN} \right\rfloor\right\}, \tag{10}$$

with C a normalization constant.

Angle distribution. The distribution of angles between n-dimensional vectors in $\mathbb{R}^n$ with coefficients drawn from a normal distribution $\mathcal{N}_{0,\sigma}$ can be modelled using the following probability density function [10]:

$$f_{\Theta_n}(\theta) = \sin^{n-2}(\theta) / \int_0^\pi \sin^{n-2}(t)dt, \quad \text{for } \theta \in [0, \pi]. \tag{11}$$

Due to the high dimensionality of the vector space used in this paper, vectors will have a very strong tendency towards orthogonality, i.e. θ is close to $\pi/2$, as can be seen in Figure 2.

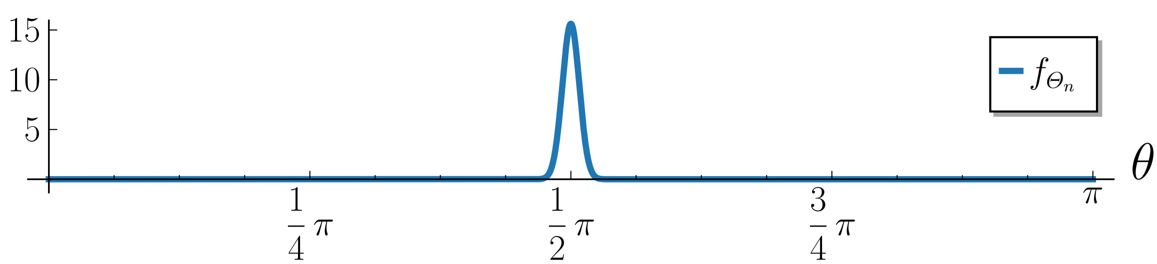


Fig. 2. Probability density function (pdf) of the angle between two random vectors in 1536-dimensional space. As the dimension increases, the pdf tends to the Dirac delta function centered at $\frac{\pi}{2}$.

For computational reasons, we will use a discretized version Θ_n^D of this distribution to model the distribution of the angles between discrete vectors, if no extra directional information is present. Given a uniformly spaced list of angles between 0 and π, we assign to each angle a probability

$$P[\theta] = C \sin^{n-2}(\theta) \tag{12}$$

with C a normalization constant. The higher the number of angles in this list, the better this distribution approximates the continuous distribution Θ_n.

Order statistics. The maximal order statistic of a distribution X in n dimensions, is the distribution of the maximum of n samples drawn from this distribution. We will denote this distribution with $M(X, n)$. For a discrete distribution X, the probability mass function of $M(X, n)$ can be computed as:

$$\begin{aligned} f_{M(X,n)}(x) =& P[x \geq y | y \leftarrow X]^n - P[x > y | y \leftarrow X]^n && (13) \\ \approx& n \cdot P[x = y | y \leftarrow X] \cdot P[x > y | y \leftarrow X]^{n-1}, && (14) \end{aligned}$$

where the latter approximation gets better for smaller probabilities.

5.2 Step 1: Estimating the direction $\overline{\mathbf{E}}$

Informally, $\overline{\mathbf{E}}$ should be a vector that has approximately the same direction as $\overline{\mathbf{S}}$. Denoting the angle between $\overline{\mathbf{E}}$ and $\overline{\mathbf{S}}$ as θ_{ES}, the bigger $|\cos(\theta_{ES})|$, the closer our estimate is to $\pm\overline{\mathbf{S}}$ and the better our estimate of failure probability will be. Since we focus on estimating the direction of $\overline{\mathbf{S}}$, $\overline{\mathbf{E}}$ will always be normalized.

In this section, we derive an estimate $\overline{\mathbf{E}}$ of the direction of the secret $\overline{\mathbf{S}}$ given $n \geq 1$ ciphertexts $\mathbf{C}_0, \ldots, \mathbf{C}_{n-1}$. Our goal is to find $\overline{\mathbf{E}}$ such that $|\cos(\theta_{ES})|$ is as big as possible. We will first discuss the case where the adversary has one ciphertext, then the case where she has two, followed by the more general case where she has n ciphertexts.

One ciphertext. Assume that a unique failing ciphertext $\mathbf{C}$ is given. For a failure event F_r, $\overline{\mathbf{E}} = \overline{\mathbf{C}^{(r)}} / \left\| \overline{\mathbf{C}^{(r)}} \right\|_2$ is a reasonable choice as $\cos(\theta_{ES})$ is bigger than average. This can be seen as follows:

$$|\cos(\theta_{ES})| = \frac{\left|\overline{\mathbf{S}}^T \cdot \overline{\mathbf{E}}\right|}{\left\|\overline{\mathbf{S}}\right\|_2 \left\|\overline{\mathbf{E}}\right\|_2} = \frac{\left|\overline{\mathbf{S}}^T \cdot \overline{\mathbf{C}^{(r)}}\right|}{\left\|\overline{\mathbf{S}}\right\|_2 \left\|\overline{\mathbf{C}^{(r)}}\right\|_2} > \frac{q_t}{\left\|\overline{\mathbf{S}}\right\|_2 \left\|\overline{\mathbf{C}^{(r)}}\right\|_2}. \tag{15}$$

Keep in mind that the cosine of angles between random vectors strongly tend to zero in high dimensional space, so that even a relatively small value of $|\cos(\theta_{ES})|$ might be advantageous.

One can argue that it is not possible to compute $\overline{\mathbf{C}^{(r)}}$ without knowledge of r; whereas in the general case, the failure location is unknown. However, $\overline{\mathbf{E}} = \overline{\mathbf{C}^{(0)}} / \left\| \overline{\mathbf{C}^{(0)}} \right\|_2$ is an equally good estimate regardless of the value of r. Indeed, $\overline{\mathbf{C}^{(0)}}$ approximates a rotation of the secret $\overline{\mathbf{S}'} := \overline{X^{-r} \cdot \mathbf{S}}$ instead of $\overline{\mathbf{S}}$, which can be seen using the equality $\overline{\mathbf{A}}^T \cdot \overline{\mathbf{B}} = \overline{X^i \mathbf{A}}^T \cdot \overline{X^i \mathbf{B}}$:

$$\begin{aligned} \overline{\mathbf{S}}^T \cdot \overline{\mathbf{C}^{(r)}} &= \overline{X^{-r} \cdot \mathbf{S}}^T \cdot \overline{X^{-r} X^r \mathbf{C}^{(0)}} \\ &= \overline{X^{-r} \cdot \mathbf{S}}^T \cdot \overline{\mathbf{C}^{(0)}}. \end{aligned} \tag{16}$$

Furthermore, multiplicating a polynomial in R_q with a power of X does not change its infinity norm, as the multiplication only results in the rotation or negation of coefficients. Thus, using an estimate of the direction of $\overline{X^{-r} \cdot \mathbf{S}}$ is as good as an estimate of the direction of $\overline{\mathbf{S}}$ when predicting the failure probability of ciphertexts, and we can use $\overline{\mathbf{E}} = \overline{\mathbf{C}^{(0)}} / \left\| \overline{\mathbf{C}^{(0)}} \right\|_2$.

Two ciphertexts. Now, assume that two linearly independent failing ciphertexts $\mathbf{C}_0$ and $\mathbf{C}_1$, resulting from failure events F_{r_0} and F_{r_1} respectively, are given. Taking $\overline{\mathbf{E}}$ as the normalized version of an average $\overline{\mathbf{C}}_{\text{av}} = \left(\overline{\mathbf{C}_0^{(0)}} + \overline{\mathbf{C}_1^{(0)}}\right)/2$ may not necessarily result in a good estimate. For example, if $\mathbf{C}_0$ comes from a failure event F_0 and $\mathbf{C}_1$ from a failure event F_N, the two directions cancel each other out as the ciphertexts $\overline{\mathbf{C}_0^{(0)}}$ and $\overline{\mathbf{C}_1^{(0)}}$ are in opposite directions.

Keeping the convention that $\overline{\mathbf{C}_0^{(0)}}$ approximates a rotation of the secret $\overline{\mathbf{S'}} = \overline{X^{-r_0} \cdot \mathbf{S}}$, we will compute the relative error position $\delta_{1,0} = r_1 - r_0$ and show that is enough to build a correct estimate $\overline{\mathbf{E}}$ as $\overline{\mathbf{E}} = \overline{\mathbf{C}_{\text{rav}}} / \left\|\overline{\mathbf{C}_{\text{rav}}}\right\|_2$ where:

$$\overline{\mathbf{C}_{\text{rav}}} := \left(\overline{\mathbf{C}_0^{(0)}} + \overline{\mathbf{C}_1^{(\delta_{1,0})}}\right)/2. \tag{17}$$

The reason why such $\overline{\mathbf{E}}$ is a good estimator of $\overline{\mathbf{S'}}$ can be seen as follows:

$$\begin{aligned}\cos(\theta_{ES'}) &= \frac{1}{2\left\|\overline{\mathbf{C}_{\text{rav}}}\right\|_2 \left\|\overline{\mathbf{S'}}\right\|_2} \cdot \left(\overline{X^{-r_0} \cdot \mathbf{S}}^T \cdot \overline{\mathbf{C}_0^{(0)}} + \overline{X^{-r_0} \cdot \mathbf{S}}^T \cdot \overline{X^{r_1 - r_0} \mathbf{C}_1^{(0)}}\right) \\ &= \frac{1}{2\left\|\overline{\mathbf{C}_{\text{rav}}}\right\|_2 \left\|\overline{\mathbf{S'}}\right\|_2} \cdot \left(\overline{\mathbf{S}}^T \cdot \overline{\mathbf{C}_0^{(r_0)}} + \overline{\mathbf{S}}^T \cdot \overline{\mathbf{C}_1^{(r_1)}}\right) > \frac{q_t}{\left\|\overline{\mathbf{C}_{\text{rav}}}\right\|_2 \left\|\overline{\mathbf{S'}}\right\|_2}.\end{aligned}$$

Remark 4. In practice ciphertexts with smaller norm will on average be better aligned with the secret, as $\cos(\theta_{CS'}) > q_t / (\left\|\overline{\mathbf{C}}\right\|_2 \left\|\overline{\mathbf{S'}}\right\|_2)$. Therefore they carry more information than ciphertexts with larger norm. To compensate for this effect we will calculate $\overline{\mathbf{C}_{\text{rav}}}$ as $:= \left(\overline{\mathbf{C}_0^{(0)}} / \left\|\overline{\mathbf{C}_0^{(0)}}\right\|_2 + \overline{\mathbf{C}_1^{(\delta_{1,0})}} / \left\|\overline{\mathbf{C}_1^{(\delta_{1,0})}}\right\|_2\right)/2$. While it is possible to further refine the calculation of $\overline{\mathbf{E}}$ using extra directional information, this heuristic is good enough for our purposes.

Computation of the relative position $\delta_{1,0}$. One can use the fact that both $\overline{\mathbf{C}_0^{(0)}}$ and $\overline{\mathbf{C}_1^{(\delta_{1,0})}}$ are expected to be directionally close to $\overline{\mathbf{S'}}$. Thus, the cosine of the angle between $\overline{\mathbf{C}_0^{(0)}}$ and $\overline{\mathbf{C}_1^{(\delta_{1,0})}}$ should be larger than usual. Therefore, $\delta_{1,0}$ can be estimated with the following distinguisher:

$$\delta'_{1,0} := \operatorname*{argmax}_{r \in [0, 2N-1]} \mathcal{C}(r) \text{ where } \mathcal{C}(r) := \frac{\overline{\mathbf{C}_0^{(0)}}^T \cdot \overline{\mathbf{C}_1^{(r)}}}{\left\|\overline{\mathbf{C}_0^{(0)}}\right\|_2 \left\|\overline{\mathbf{C}_1^{(r)}}\right\|_2}. \tag{18}$$

The next paragraph estimates the efficiency of using Equation 18 as a distinguisher for deriving $\delta_{1,0}$. We will show that, for Table 1 parameters, we expect

$$P[\delta'_{1,0} = \delta_{1,0}] \approx 89\%. \tag{19}$$

Experiments run by simulating the sampling 10^4 failing ciphertexts (refer to §A for the generation technique), and using Equation 18 for finding $\delta_{1,0}$ between pairs of them, return $P_{\text{Exp}}[\delta'_{1,0} = \delta_{1,0}] \approx 84.8\%$, in sufficiently good agreement.

To obtain the value (19), the idea is to estimate the distribution of a correct guess $\mathcal{C}(\delta_{1,0})$ and an incorrect guess $\max_{r \neq \delta_{1,0}} \mathcal{C}(r)$ and quantify the discrepancy. First, we decompose the ciphertexts in a component parallel to $\overline{\mathbf{S'}}$, denoted with $\|$, and a component orthogonal, denoted with $\perp$, we rewrite $\mathcal{C}(r)$ as follows:

$$\mathcal{C}(r) = \frac{\overline{\mathbf{C}_{0,\|}^{(0)}} \cdot \overline{\mathbf{C}_{1,\|}^{(r)}} + \overline{\mathbf{C}_{0,\perp}^{(0)}} \cdot \overline{\mathbf{C}_{1,\perp}^{(r)}}}{\left\|\overline{\mathbf{C}_0^{(0)}}\right\|_2 \left\|\overline{\mathbf{C}_1^{(r)}}\right\|_2} \tag{20}$$

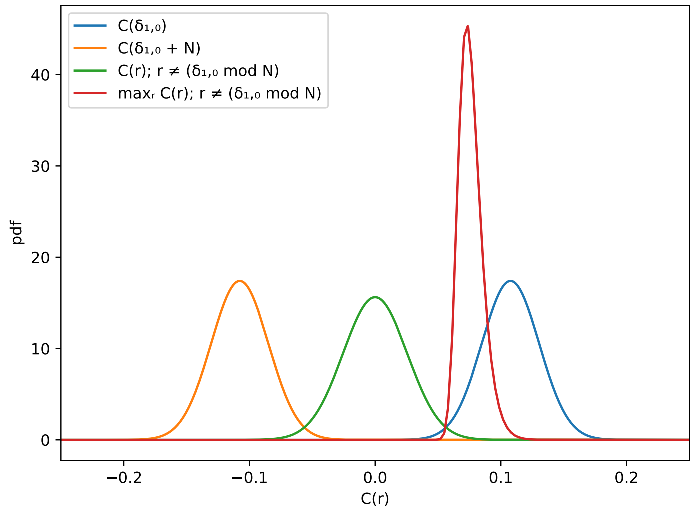


Fig. 3. Distributions used for finding $\delta_{1,0}$

In the first term, the scalar product of two parallel elements equals the product of their norms (up to their sign). For the second term, we apply the scalar product definition and intoduce t as the angle between $\overline{\mathbf{C}_{0,\perp}^{(0)}}$ and $\overline{\mathbf{C}_{1,\perp}^{(r)}}$.

$$\mathcal{C}(r) = \pm \frac{\left\|\overline{\mathbf{C}_{0,\|}^{(0)}}\right\|_2}{\left\|\overline{\mathbf{C}_{0}^{(0)}}\right\|_2} \cdot \frac{\left\|\overline{\mathbf{C}_{1,\|}^{(r)}}\right\|_2}{\left\|\overline{\mathbf{C}_{1}^{(r)}}\right\|_2} \pm \frac{\left\|\overline{\mathbf{C}_{0,\perp}^{(0)}}\right\|_2}{\left\|\overline{\mathbf{C}_{0}^{(0)}}\right\|_2} \cdot \frac{\left\|\overline{\mathbf{C}_{1,\perp}^{(r)}}\right\|_2}{\left\|\overline{\mathbf{C}_{1}^{(r)}}\right\|_2} \cdot \cos(t) \tag{21}$$

$$= \cos\left(\theta_{S'C_0^{(0)}}\right)\cos\left(\theta_{S'C_1^{(r)}}\right) + \sin\left(\theta_{S'C_0^{(0)}}\right)\sin\left(\theta_{S'C_1^{(r)}}\right)\cos(t) \tag{22}$$

The vectors $\overline{\mathbf{C}_{0,\perp}^{(0)}}$ and $\overline{\mathbf{C}_{1,\perp}^{(r)}}$ are orthogonal to $\overline{\mathbf{S}'}$. This means that they live in the $2Nl-1$ dimensional space orthogonal to $\overline{\mathbf{S}'}$. The high dimension of the space will strongly drive the vectors towards orthogonality as can be seen in Figure 2. Using Assumption 1, the angle t between $\overline{\mathbf{C}_{0,\perp}^{(0)}}$ and $\overline{\mathbf{C}_{1,\perp}^{(r)}}$ is then assumed to follow the distribution of random angles between vectors in a $2Nl-1$ dimensional space (See Equation 11).

Now, let us study the distribution of $\mathcal{C}(r)$ depending of the value $r \in [0, 2N-1]$. One can refer to Figure 3 for a graphical interpretation based on the parameters of Table 1.

- If $r = \delta_{1,0}$, the expected value of $\mathcal{C}(r)$ will be higher than average. Indeed, by definition of F_{r_1} and F_{r_0} the cosines forming the first term are positive. The distribution of $\mathcal{C}(r)$ can then be estimated using Equation 22 (blue curve).
- If $r = \delta_{1,0} + N \mod 2N$, the distribution of $\mathcal{C}(r)$ is equal to the distribution of $-\mathcal{C}(\delta_{1,0})$ and will be closer to -1 (orange curve).
- If $r \neq \delta_{1,0} \mod N$, $\mathcal{C}(r)$ can be assumed to follow the distribution of random angles in a $2Nl$ dimensional space Θ_{2Nl}, as given in Equation 11 (green curve).
- The pdf of $\max_{r\neq\delta_{1,0}} \mathcal{C}(r)$ is then calculated as $M(\Theta_{2Nl}, 2N-1)$ by definition of the maximal order statistic (red curve).

Figure 3 assembles the probability density functions of the above distributions in a plot. The probability of selecting the correct $\delta'_{1,0}$ using $\underset{r\in[0,2N-1]}{\text{argmax}}\ \mathcal{C}(r)$, can then be computed as:

$$P[\delta'_{1,0} = \delta_{1,0}] = P\left[\max_{r\neq\delta_{1,0}} \mathcal{C}(r) < \mathcal{C}(\delta_{1,0})\right].$$

For our toy scheme's parameters, this results in Equation 19.

Multiple ciphertexts. In this section, we assume that n linearly independent failing ciphertexts $\mathbf{C}_0, \ldots, \mathbf{C}_{n-1}$, resulting from failure events $F_{r_0}, \ldots, F_{r_{n-1}}$ respectively, are given. We introduce a generalized method to recover the relative positions $\delta_{1,0}, \ldots, \delta_{n-1,0}$, based on "loopy belief propagation" [47]. Once these relative positions are found, they can be combined in an estimate $\overline{\mathbf{E}}$ with $\overline{\mathbf{E}} = \overline{\mathbf{C}_{\text{rav}}} / \left\|\overline{\mathbf{C}_{\text{rav}}}\right\|_2$ where

$$\overline{\mathbf{C}_{\text{rav}}} := \left(\overline{\mathbf{C}_0^{(0)}} / \left\| \overline{\mathbf{C}_0^{(0)}} \right\|_2 + \sum_{i\in[1,n-1]} \overline{\mathbf{C}_i^{(\delta_{i,0})}} / \left\| \overline{\mathbf{C}_i^{(\delta_{i,0})}} \right\|_2 \right) / n. \tag{23}$$

To find the correct rotations, we construct a weighted graph that models the probability of different relative rotations, and we will use loopy belief propagation to obtain the most probable set of these rotations:

- The nodes represent the obtained failing ciphertexts: $(\mathbf{C}_i)_{i\in[0,n-1]}$. In total, there are n nodes.
- Each node $\mathbf{C}_i$ where $i \neq 0$ is associated a list with $2N$ probability values called *beliefs* and denoted $(b_i(0), \cdots, b_i(2N-1))$ that together define a probability distribution over $[0, 2N-1]$. The r^{th} item in the list represents our belief that the correct relative position $\delta_{0,1}$ equals r. The correct rotation of the 0^{th} node will be fixed to 0 (i.e. $b_0(0) = 1$ and $b_0(i) = 0$ for all other i) as only the relative rotations of the ciphertexts is important. These node weights are initialized as follows:

$$b_i(r) := P\left[\delta_{i,0} = r\right] \quad (= P\left[F_r \text{ for } \mathbf{C}_i | F_0 \text{ for } \mathbf{C}_0\right])$$

- For two nodes $\mathbf{C}_i$ and $\mathbf{C}_j$, the value of the vertex called *message*, models the influence of the beliefs in the rotations s of node j towards the beliefs in rotation r of node i, which can be formalized as follows:

$$m_{i,j}(r,s) := P\left[\delta_{i,j} = r - s\right] \quad (= P\left[F_{r-s} \text{ for } \mathbf{C}_i | F_0 \text{ for } \mathbf{C}_j\right])$$

Loopy belief propagation tries to find the node values r for each node, so that the probabilities over the whole graph are maximized. This is done in an iterative fashion by updating the node beliefs according to the messages coming from all other nodes. Our goal is eventually to find $r = \delta_{i,0}$ for each node i.

Example 2. For example, with $N = 3$ and $n = 3$, the graph contains the nodes $\mathbf{C}_0$, $\mathbf{C}_1$, and $\mathbf{C}_2$. In Figure 4, we represent how such a graph could look like where we arbitrarily instantiate the messages and beliefs. We can see that if one chooses the $r_i = \text{argmax}_r\, b_i(r)$ for each node, one would have chosen $r_1 = 1$ and $r_2 = 3$. Nevertheless, we notice that the influence of the other probabilities allows for a better choice (underlined in blue in the figure): $r_1 = 2, r_2 = 3$.

Vertex probabilities. As discussed, the edge between two nodes $\mathbf{C}_i$ with rotation r and $\mathbf{C}_j$ with rotation s is weighted with $P\left[\delta_{i,j} = r - s\right]$. This probability can be computed using a generalization of the distinguisher technique used for two ciphertexts as detailed in § D.

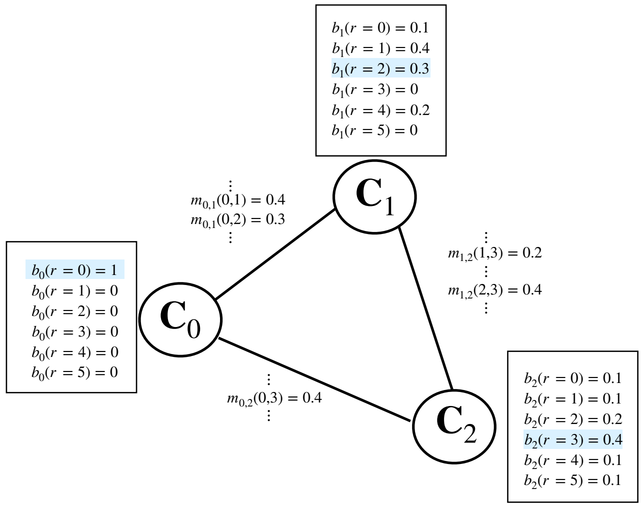


Fig. 4. Example of the graph for finding the relative rotations where $N = 3$ and $n = 3$. The beliefs are in the rectangles, the circles represent the nodes and some messages are represented between the nodes.

Loopy belief propagation. This technique looks for the best set $(r_1, \ldots, r_{n-1})$ by iteratively correcting the beliefs using messages from other nodes. This procedure is detailed in Algorithm 6, where $\texttt{normalize}(f)$ normalizes the list $b()$ so that $\sum_{x \in supp(b)} b(x) = 1$. In each iteration, the belief of each node $\mathbf{C}_i$ is updated according to the messages of the other nodes $\mathbf{C}_j$. For each i the belief is updated as follows:

$$b_i(r) = \prod_{j=0, j\neq i}^{n} \mathsf{infl}_{ji}(r) \tag{24}$$

where $\mathsf{infl}_{ji}(r)$ captures the influence of the value of the node $\mathbf{C}_j$ to node $\mathbf{C}_i$. This influence can be calculated as $\mathsf{infl}_{ji}(r) \leftarrow C \sum_x m_{i,j}(r, x) \cdot b_j(x)$, with C as normalizing constant.

Experimental verification. With Table 1 parameters, we obtained the correct values $r_i = \delta_{i,0}$ for all $i \in [1, n-1]$ after 3 iterations with the probabilities as reported in Table 2, by generating 1000 times each number of ciphertexts and trying to find the correct values of the r_i.

Remark 5 (Consistency with the previous section). Note that this procedure also incorporates the setting where one has only 2 failing ciphertexts, which would yield exactly the same results as in the previous paragraph.

Finally, once all the rotations are found, recall that the estimate is obtained by $\overline{\mathbf{E}} = \overline{\mathbf{C}_{\text{rav}}} / \left\| \overline{\mathbf{C}_{\text{rav}}} \right\|_2$ where

$$\overline{\mathbf{C}_{\text{rav}}} = \left(\overline{\mathbf{C}_0^{(0)}} / \left\| \overline{\mathbf{C}_0^{(0)}} \right\|_2 + \sum_{i \in [1, n-1]} \overline{\mathbf{C}_i^{(r_i)}} / \left\| \overline{\mathbf{C}_i^{(r_i)}} \right\|_2 \right) / n. \tag{25}$$

5.3 Step 2: Finding weak ciphertexts

In this section, we are given an estimate $\overline{\mathbf{E}}$ and we refine the acceptance criterion. Instead of accepting if $P[F \mid ct] \geq f_t$, our condition is slightly changed.

1. Generate a key encapsulation $ct = (\mathbf{C}, \mathbf{G})$ with derived key K.

Algorithm 6: `GetRotation()`

for $i \in [1, n-1]$ **do** `// initialization`
  foreach r **do**
    $b_i(r) := P\left[\delta_{0,i} = r\right]$

for *# of iterations* **do** `// update phase`
  for $i \in [1, n)$ **do**
    for $j \in [1, n)$ **if** $i \neq j$ **do**
      foreach r **do**
        $\mathsf{infl}_{ji}(r) := \sum_s m_{i,j}(r, s) \cdot b_j(s)$ `// influence of node` j `on node` i
      `normalize`(infl_{ji})
    foreach r **do**
      $b_i(r) := \prod_{j=0, j\neq i}^{n} \mathsf{infl}_{ji}(r)$ `// calculate new belief`
    `normalize`(b_i)

for $i \in [1, n)$ **do** `// finally`
  $r_i := \text{argmax}_{r\in[0,2N-1]}\, b_i(r)$ `// pick the` r_i `with highest belief`

return $(r_i)_{i\in[1,n-1]})$

	2 ciphertexts	3 ciphertexts	4 ciphertexts	5 ciphertexts
$P[r_i = \delta_{i,0}\ \forall i \in [1, n-1]]$	84.0%	95.6%	> 99.0%	> 99.0%

Table 2. Probability of finding the correct relative rotations and thus building the correct estimate $\overline{\mathbf{E}}$ with the knowledge of 2, 3, 4 and 5 failing ciphertexts.

2. If $P[F \,|\, \overline{\mathbf{E}}, ct] \geq f_t$, keep ct in a weak ciphertext list, otherwise go to to Step 1.

In Step 2, $P[F \,|\, \overline{\mathbf{E}}, ct]$ is defined as the failure probability, given a certain ciphertext ct and a certain estimate $\overline{\mathbf{E}}$. In the following, we explain a way to compute it.
First, for $r \in [0, 2N-1]$, we will estimate the probability that a ciphertext leads to an error in the r^{th} location. Decomposing the vectors $\overline{\mathbf{S}}$ and $\overline{\mathbf{C}}$ in a component orthogonal to $\overline{\mathbf{E}}$, denoted with subscript $\perp$, and a component parallel to $\overline{\mathbf{E}}$, denoted with subscript $\|$, we obtain the failure expression:

$$\begin{aligned}
P[F_r \,|\, \overline{\mathbf{E}}, \mathbf{C}] =& P[\overline{\mathbf{S}}^T \cdot \overline{\mathbf{C}^{(r)}} > q_t \,|\, \overline{\mathbf{E}}, \mathbf{C}] = P[\overline{\mathbf{S}_{\|}}^T \cdot \overline{\mathbf{C}_{\|}^{(r)}} + \overline{\mathbf{S}_{\perp}}^T \cdot \overline{\mathbf{C}_{\perp}^{(r)}} > q_t \,|\, \overline{\mathbf{E}}, \mathbf{C}] \\
=& P\left[\begin{pmatrix} \|\overline{\mathbf{S}}\|_2 \left\|\overline{\mathbf{C}^{(r)}}\right\|_2 \cos(\theta_{SE})\cos(\theta_{C^{(r)}E}) + \\ \|\overline{\mathbf{S}}\|_2 \left\|\overline{\mathbf{C}^{(r)}}\right\|_2 \sin(\theta_{SE})\sin(\theta_{C^{(r)}E})\cos(t) \end{pmatrix} > q_t \,|\, \overline{\mathbf{E}}, \mathbf{C}\right] \\
=& P\left[\cos(t) > \frac{q_t - \|\overline{\mathbf{S}}\|_2 \left\|\overline{\mathbf{C}^{(r)}}\right\|_2 \cos(\theta_{SE})\cos(\theta_{C^rE})}{\|\overline{\mathbf{S}}\|_2 \left\|\overline{\mathbf{C}^{(r)}}\right\|_2 \sin(\theta_{SE})\sin(\theta_{C^rE})} \,|\, \overline{\mathbf{E}}, \mathbf{C}\right]
\end{aligned}$$

where θ_{SE} and $\theta_{C^{(r)}E}$ are the angles of $\overline{\mathbf{S}}$ and $\overline{\mathbf{C}^{(r)}}$ with the estimate $\overline{\mathbf{E}}$ respectively, and where t is the angle between $\overline{\mathbf{S}}_{\perp}$ and $\overline{\mathbf{C}_{\perp}^{(r)}}$. We assume no other knowledge about the direction of the secret apart from the directional estimate $\overline{\mathbf{E}}$. In this case, using Assumption 1, t can be estimated as a uniform angle in a $2Nl - 1$ dimensional space. Then t is assumed to follow the probability distribution Θ_{2Nl-1} (defined in Equation 11).

The values $\overline{\mathbf{E}}$, $\|\mathbf{C}\|_2$ and $\cos(\theta_{C^{(r)}E})$ are known, meanwhile the values $\|\mathbf{S}\|$ and θ_{SE} can be modelled using their probability distribution. Thus, we can approximate $P[F_i|\,\overline{\mathbf{E}},\mathbf{C}]$ with $P[F_i\,|\,\overline{\mathbf{E}},\|\overline{\mathbf{C}}\|_2,\cos(\theta_{C^{(r)}E})]$.

Assumption 2 *We assume that failures at different locations are independent.*

Assumption 2 is a valid assumption for schemes without error correcting codes, as discussed in [17]. We can then calculate the failure probability of a certain ciphertext as:

$$P[F\,|\,\overline{\mathbf{E}},\mathbf{C}] = 1 - \prod_{r=0}^{2N}\left(1 - P[F_r\,|\,\overline{\mathbf{E}},\|\overline{\mathbf{C}}\|_2,\cos(\theta_{C^{(r)}E})]\right) \tag{26}$$

As this expression gives us a better prediction of the failure probability of ciphertexts by using the information embedded in $\overline{\mathbf{E}}$, we can more accurately (Grover) search for weak ciphertexts and thus reduce the work to find the next decryption failure. Moreover, the better $\overline{\mathbf{E}}$ approximates the direction of $\overline{\mathbf{S}}$, the easier it becomes to find a new decryption failure.

5.4 Finalizing the attack with lattice reduction

Once multiple failures are found, the secret key can be recovered with lattice reduction techniques as presented in [17, §4] and in [30, Step 3 of the attack]. The following Section simply outlines how their technique transposes to our framework. As shown in §5, an estimate $\overline{\mathbf{E}}$ of the direction of a rotated version of $\overline{\mathbf{S'}} = \overline{X^r\mathbf{S}}$ with an unknown value r is provided. Therefore, similarly to [30], an attacker can obtain an estimation of $\overline{\mathbf{S'}}$ (and not only its direction) by rescaling

$$\overline{\mathbf{E'}} := \overline{\mathbf{E}} \cdot nq_t \cdot \left(\left\|\overline{\mathbf{C}_0^{(0)}} + \sum_{i\in[1,n-1]}\overline{\mathbf{C}_i^{(r_i)}}\right\|_2\right)^{-1},$$

using the approximation $\overline{\mathbf{E'}}^T \cdot 1/n\left(\overline{\mathbf{C}_0^{(0)}} + \sum_{i\in[1,n-1]}\overline{\mathbf{C}_i^{(r_i)}}\right) \approx q_t$.

Then, for each possible $r \in [0, 2N-1]$, an attacker can perform lattice reduction and recover candidates for $\mathbf{s}, \mathbf{e}$ that are accepted if they verify $\mathbf{b} = \mathbf{A}\mathbf{s} + \mathbf{e}$. One caveat is that an attacker may have to run a lattice reduction up to $2N$ times. Since $\overline{\mathbf{E'}} - \overline{\mathbf{S'}}$ is small, the attacker can construct an appropriate lattice basis encoding $\overline{\mathbf{E'}} - \overline{\mathbf{S'}}$ as a unique shortest target vector, and solves the corresponding Unique-SVP problem with the BKZ algorithm [51,12,3,1]. The block size of BKZ will depend on the accuracy of the estimate $\overline{\mathbf{E}}$. Indeed, the standard deviation of $\overline{\mathbf{E'}}_i - \overline{\mathbf{S'}}_i$ is of the order of $\sigma_s \cdot \sin(\theta_{S'E})$ (assuming that $\theta_{S'E}$ is small and $\|\overline{\mathbf{S'}}\|_2 \approx \|\overline{\mathbf{E'}}\|_2$). Thus, when many decryption failures are available, $\sin(\theta_{S'E})$ gets very small and the complexity of this step is dominated by the work required for constructing $\overline{\mathbf{E}}$. For example, in the case of our toy scheme, if $\cos(\theta_{S'E}) > 0.985$, using [2], the BKZ block size becomes lower than 363 which leads to less than 2^{100} quantum work (in the Core-SVP [3] 0.265β model). As we will see in §6.3, this is less than the work required to find the first failure.

Remark 6. One can think that the failures obtained by directional failure boosting will not be totally independent. It is true that the failing ciphertexts are roughly following the same direction. But applying our Assumption 1, in high dimensions, for a reasonable number n of failures ($n \ll 2lN$), the hypercone in which the failures belong is large enough that linear dependency will happen with very low probability.

6 Efficiency of Directional Failure Boosting

In this section, we experimentally verify the efficiency of the directional failure boosting technique. We first quantify the accuracy of the estimate $\overline{\mathbf{E}}$ computed according to §5.2. We then derive the necessary work required to run the directional failure boosting technique and the optimal number of queries. For the rest of the section, we focus on minimizing the total work for finding failures and we will assume there is no upper limit to the number of decryption queries.

Our key takeaway is that, for Table 1 parameters, the more failing ciphertexts have been found, the easier it becomes to obtain the next one, and that most of the effort is concentrated in finding the first failure. The final work and query overheads are stored in Table 4.

6.1 Accuracy of the estimate

Let $\mathbf{C}_0, ..., \mathbf{C}_{n-1}$ be n previously found failing ciphertexts and we take the estimate defined according to Equation 25. Similarly to §5.2, we define $\overline{\mathbf{S}'} = \overline{X^{-r_0} \cdot \mathbf{S}}$ as the secret vector for an unknown F_{r_0}. To estimate the accuracy of $\overline{\mathbf{E}}$, we compute $\cos(\theta_{S'E}) = \frac{\overline{\mathbf{S}'}^T \cdot \overline{\mathbf{E}}}{\left\|\overline{\mathbf{S}'}\right\|_2} = \frac{\overline{\mathbf{S}'}^T \cdot \overline{\mathbf{C}_{\text{rav}}}}{\left\|\overline{\mathbf{S}'}\right\|_2 \left\|\overline{\mathbf{C}_{\text{rav}}}\right\|_2}$ as

$$\cos(\theta_{S'E}) = \frac{\overline{\mathbf{S}'}^T \cdot \left(\frac{\overline{\mathbf{C}_0^{(0)}}}{\left\|\overline{\mathbf{C}_0^{(0)}}\right\|_2} + \sum_{i=1}^{n-1} \frac{\overline{\mathbf{C}_i^{(r_i)}}}{\left\|\overline{\mathbf{C}_i^{(r_i)}}\right\|_2} \right)}{n \left\|\overline{\mathbf{S}'}\right\|_2 \left\|\overline{\mathbf{C}_{\text{rav}}}\right\|_2} = \frac{\cos\left(\theta_{C_0^{(0)}S'}\right) + \sum_{i=1}^{n-1} \cos\left(\theta_{C_i^{(r_i)}S'}\right)}{\left\| \frac{\overline{\mathbf{C}_0^{(0)}}}{\left\|\overline{\mathbf{C}_0^{(0)}}\right\|_2} + \sum_{i=1}^{n-1} \frac{\overline{\mathbf{C}_i^{(r_i)}}}{\left\|\overline{\mathbf{C}_i^{(r_i)}}\right\|_2} \right\|_2} \tag{27}$$

First, we make the following approximation.

Approximation 1 *We approximate the cosine with the secret $\overline{\mathbf{S}'}$ by its expected value denoted $\cos(\theta_{CS'}) := \mathbb{E}\left[\cos\left(\theta_{C_i^{(r_i)}S'}\right)\right]$. In other words, for all $i \in [1, n-1]$ we assume*

$$\cos(\theta_{CS'}) = \cos\left(\theta_{C_i^{(r_i)}S'}\right) = \cos\left(\theta_{C_0^{(0)}S'}\right).$$

To estimate the denominator of Equation 27, we split the ciphertexts in a component parallel to the secret $\overline{\mathbf{C}_{i,\|}^{(r_i)}}$ and a component orthogonal $\overline{\mathbf{C}_{i,\perp}^{(r_i)}}$ to the secret. Following Assumption 1, we will assume orthogonality between the various $\overline{\mathbf{C}_{i,\perp}^{(r_i)}}$. As the norm of the sum of parallel vectors is the sum of their norm, and the norm of the sum of orthogonal vectors can be calculated using Pythagoras' theorem, we can approximate $\cos(\theta_{S'E})$ as follows:

$$cos(\theta_{S'E}) \approx \frac{n \cos(\theta_{CS'})}{\sqrt{n^2 \cos(\theta_{CS'})^2 + n \sin(\theta_{CS'})^2}} = \frac{\cos(\theta_{CS'})}{\sqrt{\cos(\theta_{CS'})^2 + \sin(\theta_{CS'})^2/n}} \tag{28}$$

One can see from this equation that $cos(\theta_{S'E})$ gets closer to 1 when n increases.

Experimental verification. The first line of Table 3 gives the expected values of $\cos(\theta_{S'E})$ for various n, according to Equation 28, with $\cos(\theta_{CS'})$ set to $q_t/\mathbb{E}[\|\overline{\mathbf{S}}\|]\mathbb{E}[\|\overline{\mathbf{C}^{(0)}}\|]$, which is a good approximation of $\cos(\theta_{CS'})$ as $\cos(\theta_{CS'}) > q_t/\|\overline{\mathbf{S}}\|\|\overline{\mathbf{C}^{(0)}}\|$ and because angles tend to orthogonality in high dimensional space.

Then, to verify the theory, we implemented a method to simulate the distribution of random failing ciphertexts. This technique is described in §A. Once the simulated failing ciphertexts are found, we combine them to build $\overline{\mathbf{E}}$ using their correct rotations, and we compute $\cos(\theta_{S'E})$. The latter experiment was repeated 100 times and the average values are reported in line two of Table 3.

n	1	2	3	5	10	20	30	50	100
theoretical	0.328	0.441	0.516	0.613	0.739	0.841	0.885	0.926	0.961
experimental	0.318	0.429	0.502	0.600	0.727	0.832	0.878	0.921	0.958

Table 3. Accuracy of the estimate derived from several failures. Expected value of $\cos(\theta_{S'E})$ according to Equation 28. The closer to 1, the more accurate $\overline{\mathbf{E}}$ is.

6.2 Estimating α_{i,f_t} and β_{i,f_t}

To estimate the effectiveness of directional failure boosting given a certain number i of previously collected failing ciphertexts, we need to find the optimal weak ciphertext threshold f_t for each i. This corresponds to considering how much time to spend for one precalculation $\sqrt{\alpha_{i,f_t}^{-1}}$ and the average failure probability of weak ciphertexts β_{i,f_t} after the precalculation. Let us recall the definition of α_{n,f_t} and β_{n,f_t}, derived from Equations 7 and 8, where $\mathbf{C}_0, ..., \mathbf{C}_{n-1}$ are the n previously found failing ciphertexts.

$$\alpha_{i,f_t} = \sum_{\forall ct: P[F|ct, \mathbf{C}_0, ..., \mathbf{C}_{n-1}] > f_t} P[ct] \tag{29}$$

$$\beta_{i,f_t} = \frac{\sum_{\forall ct: P[F|ct, \mathbf{C}_0, ..., \mathbf{C}_{n-1}] > f_t} P[ct] \cdot P[F|ct, \mathbf{C}_0, ..., \mathbf{C}_{n-1}] > f_t]}{\sum_{\forall ct: P[F|ct, \mathbf{C}_0, ..., \mathbf{C}_{n-1}] > f_t} P[ct]}. \tag{30}$$

To find the optimal values, we need to calculate Equation 29 and 30 as functions of f_t. This requires us to list the probability of all ciphertexts $ct := (\mathbf{C}, \mathbf{G})$, and their failure probability $P[F|ct, \mathbf{C}_0, ..., \mathbf{C}_{t-1}]$. As discussed in [16], exhaustively computing both values is not practically feasible, and therefore we will make some assumptions to get an estimate.

A first simplification is to cluster ciphertexts that have similar $\|\mathbf{C}\|_2$ and $|\theta_{C^{(0)}E}| \cdots |\theta_{C^{(N-1)}E}|$ and thus a similar failure probability. To further reduce the list size, we only take into account the largest value of $|\cos(\theta_{C^{(i)}E})|$ denoted

$$\mathrm{maxcos}(\theta_{\mathrm{CE}}) := \max_{\mathrm{i}}(|\cos(\theta_{\mathrm{C^{(i)}E}})|,$$

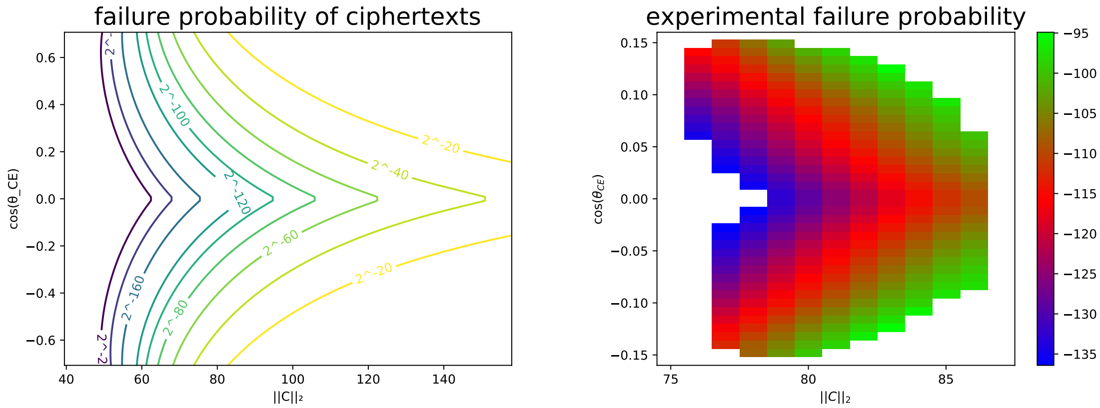
which results in a slight underestimation of the effectiveness of the attack. In other words,

$$\begin{aligned} P[ct] \quad &\text{becomes} \quad P\left[\|\mathbf{C}\|_2, \mathrm{maxcos}(\theta_{\mathrm{CE}})\right], \\ P[F|ct, \mathbf{C}_0, ..., \mathbf{C}_{n-1}] \quad &\text{becomes} \quad P\left[F \,|\, \|\mathbf{C}\|_2, \mathrm{maxcos}(\theta_{\mathrm{CE}})\right]. \end{aligned}$$

Assuming independence between the norm of $\mathbf{C}$ and its angle with $\overline{\mathbf{E}}$, $P\left[\|\mathbf{C}\|_2, \mathrm{maxcos}(\theta_{\mathrm{CE}}))\right]$ can be estimated using the distributions defined with Equations 10 and 13 as follows:

$$P\left[\|\mathbf{C}\|_2, \mathrm{maxcos}(\theta_{\mathrm{CE}})\right] = \underbrace{P[\|\mathbf{C}\|_2]}_{\chi_{Nl,\sigma}} \cdot \underbrace{P[\mathrm{maxcos}(\theta_{\mathrm{CE}})]}_{M(\Theta_{2Nl}, 2N)}. \tag{31}$$

Denoting with r the position of the maximum angle, we can rewrite $P\left[F \,|\, \|\mathbf{C}\|_2, \mathrm{maxcos}(\theta_{\mathrm{CE}})\right]$ as follows:



(a) Theoretical (b) Experimental

Fig. 5. Failure probability of ciphertexts as a function of $\|\mathbf{C}\|_2$ and $\cos(\theta_{CE})$. A zoomed version of Figure 5a for easier comparison can be found in § B.

$$P\left[F \mid \|\mathbf{C}\|_2, \mathrm{maxcos}(\theta_{\mathrm{CE}})\right] = 1 - \prod_i \left(1 - P\big[F_i \mid \|\mathbf{C}\|_2, |\cos(\theta_{C^{(r)}E})|\big]\right), \tag{32}$$

$$= 1 - \left(\begin{array}{l} \left(1 - P\big[F_r \mid \|\mathbf{C}\|_2, |\cos(\theta_{C^{(r)}E})|\big]\right) \cdot \\ \prod_{i \neq r} \left(1 - P\big[F_i \mid \|\mathbf{C}\|_2, |\cos(\theta_{C^{(i)}E})| \leq |\cos(\theta_{C^{(r)}E})|\big]\right) \end{array} \right), \tag{33}$$

where $1 - P\big[F_r \mid \|\mathbf{C}\|_2, |\cos(\theta_{C^{(r)}E})|\big]$ can be estimated using Equation 26, and $P\big[F_i \mid \|\mathbf{C}\|_2, |\cos(\theta_{C^{(i)}E})| \leq |\cos(\theta_{C^{(r)}E})|\big]$ using an integral over Equation 26. The estimated failure probability of ciphertexts given $\|\mathbf{C}\|_2$ and $\cos(\theta_{CE})$ for the parameters listed in Table 1 is depicted in figure 5a.

Verification experiment. We verified these results experimentally by generating $5 \cdot 10^6$ failing ciphertexts and $5 \cdot 10^6$ successful ciphertexts, and calculating their norm and angle with 1000 estimates, or in this case other ciphertexts. The failing ciphertexts were produced using the methodology of §A. Once they are generated, we estimate their failure probability with a procedure detailed in §B. We combined these results into Figure 5b. These experimental results confirm our theoretical estimates given in Figure 5a.

With the estimation of $P\left[F \mid \|\mathbf{C}\|_2, \mathrm{maxcos}(\theta_{\mathrm{CE}})\right]$ and $P\left[\|\mathbf{C}\|_2, \mathrm{maxcos}(\theta_{\mathrm{CE}})\right]$, α_{i,f_t} and β_{i,f_t} can be estimated as functions of i and f_t. Let us now define the optimal threshold f_t as a function of i as :

$$f_i := \mathrm{argmin}_{f_t} \ \left(\sqrt{\alpha_{i,f_t}} \cdot \beta_{i,f_t}\right)^{-1}.$$

6.3 Total amount of work and queries

In this section, we will derive the optimal work and queries for an adversary to perform, in order to obtain n ciphertexts with probability $1 - e^{-1}$. We introduce the following notation: to find the $(i+1)^{\text{th}}$ ciphertext, the adversary performs Q_i queries. Using a Poisson distribution, the success probability of finding the $(i+1)^{\text{th}}$ ciphertext in Q_i queries is $1 - e^{-Q_i \beta_{i,f_i}}$. The probability of obtaining n failures can then be calculated as the product of the success probabilities of finding ciphertexts 0 to $n-1$:

$$P_n = \prod_{i=0}^{n-1} (1 - e^{-Q_i \beta_{i,f_i}}). \tag{34}$$

ciphertexts n	1	2	3	5	10	20	30
$\log_2(\mathcal{W}_n^{\text{tot}})$	112.45	112.77	112.78	112.78	112.78	112.78	112.78
$\log_2(\mathcal{W}_n^{\text{tot}}/\mathcal{W}_1^{\text{tot}})$	—	0.32	0.33	0.33	0.33	0.33	0.33
$\log_2(\mathcal{Q}_n^{\text{tot}})$	102.21	102.86	102.87	102.87	102.87	102.87	102.87
$\log_2(\mathcal{Q}_n^{\text{tot}}/\mathcal{Q}_1^{\text{tot}})$	—	0.65	0.66	0.66	0.66	0.66	0.66

Table 4. Quantum work $\mathcal{W}_n^{\text{tot}}$ and queries $\mathcal{Q}_n^{\text{tot}}$ required to find n failing ciphertexts with probability $1 - e^{-1}$. Finding the first ciphertext requires the heaviest amount of computation. After the third failing ciphertext is found, the following ones are essentially for free.

This is a slight underestimation of the success probability of the attack, because if an adversary finds a failing ciphertext in less than Q_i samples, she can query more ciphertexts in the next stages $i+1, \ldots, n$. However, this effect is small due to the large value of Q_i.

The total amount of precomputation quantum work, and the total amount of queries to obtain the n failing ciphertexts by performing Q_i tries for each ciphertext, can be expressed as

$$\mathcal{W}_n^{\text{tot}} := \sum_{i=0}^{n-1} \underbrace{\frac{Q_i}{\sqrt{\alpha_{i,f_i}}}}_{:=W_i} \qquad \mathcal{Q}_n^{\text{tot}} := \sum_{i=0}^{n-1} Q_i. \tag{35}$$

Recall that for now we assume there is no upper limit to the number of decryption queries that can be made, and we focus on minimizing the amount of work. The values of Q_i that minimizes the total quantum work $\mathcal{W}_n^{\text{tot}}$ can be found using the following Lagrange multiplier, minimizing the total amount of work to find n failures with probability $1 - e^{-1}$ using the above probability model:

$$L(Q_0, \cdots, Q_{n-1}, \lambda) = \sum_{t=0}^{n-1} \frac{Q_i}{\sqrt{\alpha_{i,f_i}}} + \lambda \left((1 - e^{-1}) - \prod_{i=0}^{n-1} (1 - e^{-Q_i \beta_{i,f_i}}) \right) \tag{36}$$

By equating the partial derivative of L in $Q_0, \cdots, Q_{n-1}$ and λ to zero and solving the resulting system of equations, we obtain the optimal values of $Q_0, \cdots, Q_{n-1}$ to mount our attack.

The resulting total work and queries of obtaining n ciphertext using directional failure boosting are given in Table 4 and Figure 6. One can see that the majority of the work lies in obtaining the first ciphertext, and that obtaining more than one ciphertext can be done in less than double the work and queries, or less than one extra bit of complexity. For schemes with a lower failure probability, failing ciphertexts will be more correlated to the secret, so that the directional information is higher and directional failure boosting will be more effective.

In conclusion, the security of a scheme with low failure probability under a single target decryption failure attack can be approximated by the amount of work and queries that an adversary needs to do in order to obtain the first decryption failure. We emphasize the fact that obtaining many failures for a low overhead threatens the security of the scheme (See §5.4).

7 Discussion and variants

7.1 Comparison with D'Anvers et al. [16]

In figure 7, the total work and queries needed to obtain n ciphertexts with probability $1 - e^{-1}$ is plotted for both the traditional failure boosting, and our directional failure boosting approach. For a

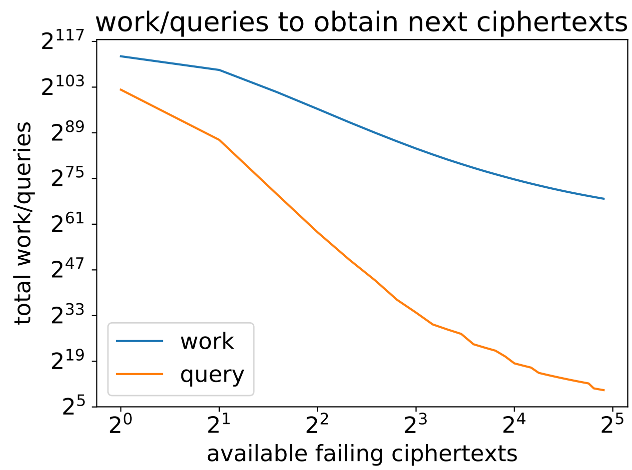
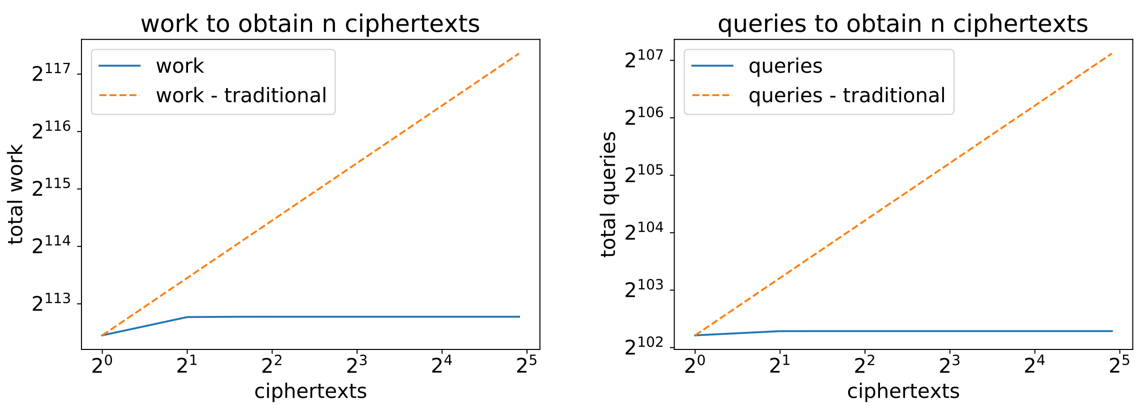


Fig. 6. Quantum work W_i and number of decryption queries Q_i required to find a new failing ciphertext, given the i failing ciphertexts found previously.

fair comparison between both results, we adapted the method for estimating the total work and queries with success probability $1 - e^{-1}$ using the traditional failure boosting of [16]. For more information about our method, we refer to §C.



(a) Quantum work $\mathcal{W}_n^{\mathrm{tot}}$ (b) Queries $\mathcal{Q}_n^{\mathrm{tot}}$

Fig. 7. Quantum work $\mathcal{W}_n^{\mathrm{tot}}$ and number of decryption queries $\mathcal{Q}_n^{\mathrm{tot}}$ required to obtain n failing ciphertexts with probability $1 - e^{-1}$, given the number of previously found failing ciphertexts.

7.2 Minimizing the number of queries instead

In case there is a maximal number of decryption queries is imposed, say 2^K, the same attack strategy can be followed. However, to limit the number of queries $\mathcal{Q}_n^{\mathrm{tot}}$ necessary in the attack, a stronger preprocessing $\sqrt{\alpha_{i,f_t}^{-1}}$ might be necessary to increase the failure probability β_{i,f_t} of weak ciphertexts over 2^{-K}. The only change to accomplish this is selecting the threshold f_t for each i appropriately.

Note that for most practical schemes (e.g. Kyber, Saber, New Hope), increasing the failure probability β_{0,f_t} over 2^{-K} is not practically feasible or would require too much preprocessing $\sqrt{\alpha^{-1}_{0,f_t}}$.

Figure 8 depicts the amount of work $\sqrt{\alpha^{-1}_{i,f_t}}\beta^{-1}_{i,f_t}$ needed to increase the failure probability β_{i,f_t} to a certain failure probability (e.g. 2^{-K}) for the parameters given in Table 1. The various curves correspond to different numbers of available failing ciphertexts. From this figure, one can see that also in this case, the work is dominated by finding the first decryption failure. Another observation is that the attack gets much more expensive as the maximal number of decryption queries 2^K gets smaller.

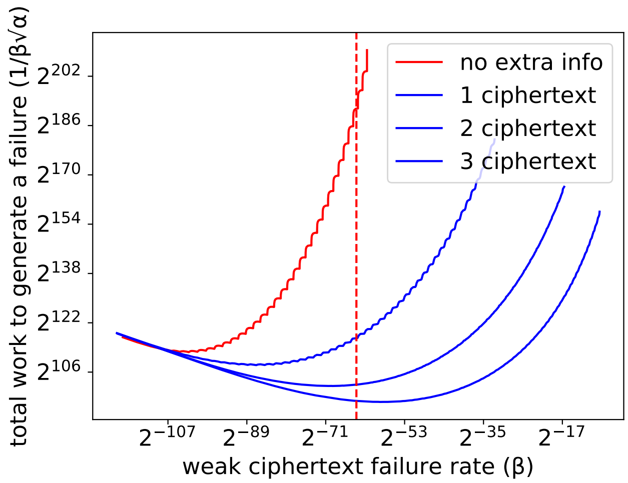


Fig. 8. Quantum work $\mathcal{W}^{\text{tot}}_n$ required to find a new failing ciphertext, as a function of the decryption failure probability of a Mod-LWE scheme.

7.3 Application to ss-ntru-pke and improvement of Guo et al. [30]

In [30], an adaptive multitarget attack is proposed on the ss-ntru-pke version of NTRUEncrypt [11], a Ring-LWE based encryption scheme that claims security against chosen ciphertext attacks. The parameters of this scheme are given in Table 5. The attack performs at most 2^{64} queries on at most 2^{64} targets and has a classical cost of 2^{216} work, and a quantum cost of 2^{140} when speeding up the offline phase with Grover's search. We adapt directional failure boosting to this attack model and propose both a single and multitarget attack.

For the single target attack, our proposed methodology in subsection 6.3 needs more than 2^{64} queries to obtain a ciphertext. To mitigate this, we increase the precomputational work $\sqrt{\alpha^{-1}}$ so that the failure probability of weak ciphertexts β increases over a certain f_t, which is chosen as 2^{-57} to make sure the total queries are below 2^{64}. The effect is a bigger overall computation, but a reduction in the number of necessary decryption queries. The rest of the attack proceeds as discussed in Subsection 6.3. The work or queries needed to obtain an extra ciphertexts with n ciphertexts can be seen in Figure

	l	N	q	σ_s	σ_e	$P[F]$	Claimed Security
ss-ntru-pke	1	1024	$2^{30}+2^{13}+1$	724	724	$> 2^{-80}$	2^{198}

Table 5. Parameters of the ss-ntru-pke [11] scheme.

9a. The cost of this single target attack is $2^{139.6}$, which is close to the cost of their multitarget attack $2^{139.5}$, as can be seen in Table 6.

scheme	claimed security	multitarget attack [30]	our single target attack	our multitarget attack
ss-ntru-pke	2^{198}	$2^{139.5}$	$2^{139.6}$	$2^{96.6}$

Table 6. Comparison of costs for different attacks against ss-ntru-pke [11].

In the multitarget case, we can use a maximum of $2^{64} \cdot 2^{64}$ queries to find the first failing ciphertext, after which we use the methodology of the single target attack to obtain further ciphertext with limited amount of queries. In practice we stay well below the query limit to find the first failure. In this case, the work is dominated by finding the second decryption failure, as we need to do this in under 2^{64} queries. The resulting work to obtain an extra ciphertext is depicted in Figure 9b. The cost of this attack is $2^{96.6}$, which is well below the cost of $2^{139.5}$ reported by Guo et al.

(a) Single target attack

(b) Multitarget attack

Fig. 9. Quantum work $\mathcal{W}_n^{\text{tot}}$ and number of decryption queries $\mathcal{Q}_n^{\text{tot}}$ required to find a new failing ciphertext for ss-ntru-pke, given the ones found previously.

7.4 Future Work

Application to schemes with error correction. For schemes with strong error correction, such as LAC [41], the error probability of the ciphertext before error correction is relatively high, which results in a lower correlation between the secret and failing ciphertexts. This would complicate the effectiveness of directional failure boosting in several ways: finding the right rotation of ciphertext to combine them into $\overline{\mathbf{E}}$ becomes less straightforward, and the resulting $\overline{\mathbf{E}}$ would be less correlated with the secret. However, by finding one failure, an adversary would already have more than one equation due to the multiple errors. Thus the effectiveness of directional failure boosting in these circumstances can not be derived from our results and would be an interesting future research topic.

Multitarget attack with limited number of queries. In case an attacker has multiple possible targets 2^T but can perform at most 2^K queries at each target, she can intially search for one of more ciphertexts using a small and thus more optimal $\beta_{f_t} > 2^{-(T+K)}$, before focussing on one target and limiting to $\beta_{f_t} > 2^{-K}$, which requires more preprocessing and possibly more work. In this case, finding the first decryption failure is not necessarily the limiting factor for the adversary. While the attack strategy has been touched upon in subsection 7.3, the optimal attack strategy and its efficiency is outside the scope of this paper, but would be interesting future work.

Acknowledgements. We thank Henri Gilbert and Alessandro Budroni for the interesting discussions about decryption errors, and for providing advice during the writeup of this paper.

References

1. M. R. Albrecht, F. Göpfert, F. Virdia, and T. Wunderer. Revisiting the expected cost of solving uSVP and applications to LWE. pages 297–322.
2. M. R. Albrecht, R. Player, and S. Scott. On the concrete hardness of learning with errors. *JMC*, 2015.
3. E. Alkim, L. Ducas, T. Pöppelmann, and P. Schwabe. Post-quantum key exchange – a New Hope. In *USENIX Security 2016*, 2016.
4. H. Baan, S. Bhattacharya, O. Garcia-Morchon, R. Rietman, L. Tolhuizen, J.-L. Torre-Arce, and Z. Zhang. Round2: KEM and PKE based on GLWR. IACR ePrint 2017/1183.
5. C. Băetu, F. B. Durak, L. Huguenin-Dumittan, A. Talayhan, and S. Vaudenay. Misuse attacks on post-quantum cryptosystems. In *EUROCRYPT*. Springer, 2019.
6. S. Bai and S. D. Galbraith. An improved compression technique for signatures based on learning with errors. In *CT-RSA*. Springer, 2014.
7. J. Bos, L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, J. M. Schanck, P. Schwabe, and D. Stehlé. CRYSTALS – Kyber: a CCA-secure module-lattice-based KEM. IACR ePrint 2017/634.
8. J. Bos, L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, J. M. Schanck, P. Schwabe, and D. Stehlé. CRYSTALS Kyber: a CCA-secure module-lattice-based KEM.
9. Z. Brakerski, C. Gentry, and V. Vaikuntanathan. (leveled) fully homomorphic encryption without bootstrapping. *TOCT*, 2014.
10. T. Cai, J. Fan, and T. Jiang. Distributions of Angles in Random Packing on Spheres. *Journal of machine learning research*, 14(1), jan 2013.
11. C. Chen, J. Hoffstein, W. Whyte, and Z. Zhang. NTRUEncrypt. Technical report, NIST, 2017.
12. Y. Chen and P. Q. Nguyen. Bkz 2.0: Better lattice security estimates. In *ASIACRYPT 2011*. Springer, 2011.
13. J.-P. D'Anvers, Q. Guo, T. Johansson, A. Nilsson, F. Vercauteren, and I. Verbauwhede. Decryption Failure Attacks on IND-CCA Secure Lattice-Based Schemes. In *PKC 2019*, 2019.
14. J.-P. D'Anvers, A. Karmakar, S. S. Roy, and F. Vercauteren. Saber: Module-LWR Based Key Exchange, CPA-Secure Encryption and CCA-Secure KEM. In *AFRICACRYPT 2018*, 2018.
15. J.-P. D'Anvers, A. Karmakar, S. S. Roy, and F. Vercauteren. SABER, Round 2 submission. Technical report, NIST, 2019.
16. J.-P. D'Anvers, F. Vercauteren, and I. Verbauwhede. On the impact of decryption failures on the security of LWE/LWR based schemes. IACR ePrint 2018/1089.
17. J.-P. D'Anvers, F. Vercauteren, and I. Verbauwhede. The impact of error dependencies on Ring/Mod-LWE/LWR based schemes. IACR ePrint 2018/1172.
18. A. W. Dent. A designers guide to kems. In *IMACC*. Springer, 2003.
19. W. Diffie and M. Hellman. New directions in cryptography. *IEEE transactions on Information Theory*, 1976.
20. L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, P. Schwabe, G. Seiler, and D. Stehlé. Crystals-dilithium: A lattice-based digital signature scheme. *TCHES*, 2018.
21. T. ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE transactions on information theory*, 31(4):469–472, 1985.
22. S. Fluhrer. Cryptanalysis of ring-LWE based key exchange with key share reuse. IACR ePrint 2016/085.
23. E. Fujisaki and T. Okamoto. Secure Integration of Asymmetric and Symmetric Encryption Schemes. *Journal of Cryptology*, 26(1), jan 2013.
24. E. Fujisaki and T. Okamoto. Secure integration of asymmetric and symmetric encryption schemes. 26(1):80–101, Jan. 2013.

25. N. Gama and P. Q. Nguyen. New Chosen-Ciphertext Attacks on NTRU. In T. Okamoto and X. Wang, editors, *Public Key Cryptography – PKC 2007*, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.
26. C. Gentry and D. Boneh. *A fully homomorphic encryption scheme*. Stanford University Stanford, 2009.
27. C. Gentry, C. Peikert, and V. Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *STOC 08*. ACM, 2008.
28. C. Gentry, A. Sahai, and B. Waters. Homomorphic encryption from learning with errors: Conceptually-simpler, asymptotically-faster, attribute-based. In *CRYPTO 2013*. Springer, 2013.
29. L. K. Grover. A Fast Quantum Mechanical Algorithm for Database Search. In *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, STOC '96, New York, NY, USA, 1996. ACM.
30. Q. Guo, T. Johansson, and A. Nilsson. A Generic Attack on Lattice-based Schemes using Decryption Errors with Application to ss-ntru-pke. IACR ePrint 2019/043.
31. D. Hofheinz, K. Hövelmanns, and E. Kiltz. A Modular Analysis of the Fujisaki-Okamoto Transformation. 2017.
32. N. Howgrave-Graham, P. Q. Nguyen, D. Pointcheval, J. Proos, J. H. Silverman, A. Singer, and W. Whyte. The Impact of Decryption Failures on the Security of NTRU Encryption. In *CRYPTO 2003*, 2003.
33. J. Hsu. Ces:intels 49-qubit chip shoots for quantum supermacy, 2017. https://spectrum.ieee.org/tech-talk/computing/hardware/intels-49qubit-chip-aims-for-quantum-supremacy.
34. É. Jaulmes and A. Joux. A Chosen-Ciphertext Attack against NTRU. In *CRYPTO 2000*, 2000.
35. H. Jiang, Z. Zhang, L. Chen, H. Wang, and Z. Ma. Post-quantum IND-CCA-secure KEM without Additional Hash. IACR ePrint 2017/1096.
36. N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous univariate distributions*. Houghton Mifflin Boston, 1970.
37. J. Katz and Y. Lindell. *Introduction to Modern Cryptography, Second Edition*. Chapman & Hall/CRC, 2014.
38. J. Kelly. A preview of bristlecone, googles new quantum processor, 2018. https://research.googleblog.com/2018/03/a-preview-of-bristlecone-googles-new.html.
39. W. Knight. Ibm raises the bar with a 50-qubit quantum computer, 2017. https://www.technologyreview.com/s/609451/ibm-raises-the-bar-with-a-50-qubit-quantum-computer/.
40. A. Langlois and D. Stehlé. Worst-case to average-case reductions for module lattices. *Designs, Codes and Cryptography*, 75(3), jun 2015.
41. X. Lu, Y. Liu, D. Jia, H. Xue, J. He, and Z. Zhang. LAC. Technical report, NIST, 2017.
42. V. Lyubashevsky. Fiat-shamir with aborts: Applications to lattice and factoring-based signatures. In *ASIACRYPT 2009*. Springer, 2009.
43. V. Lyubashevsky, C. Peikert, and O. Regev. On ideal lattices and learning with errors over rings. pages 1–23.
44. D. Micciancio and S. P. Vadhan. Statistical zero-knowledge proofs with efficient provers: Lattice problems and more. In *CRYPTO 2003*. Springer, 2003.
45. M. Naehrig, E. Alkim, J. Bos, L. Ducas, K. Easterbrook, B. LaMacchia, P. Longa, I. Mironov, V. Nikolaenko, C. Peikert, A. Raghunathan, and D. Stebila. FrodoKEM. Technical report, NIST, 2017.
46. NIST. Submission requirements and evaluation criteria for the Post-Quantum Cryptography standardization process, 2016.
47. J. Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Elsevier, 2014.
48. C. Peikert and V. Vaikuntanathan. Noninteractive statistical zero-knowledge proofs for lattice problems. In *CRYPTO 2008*. Springer, 2008.
49. O. Regev. On Lattices, Learning with Errors, Random Linear Codes, and Cryptography. In *STOC*. ACM, 2005.
50. T. Saito, K. Xagawa, and T. Yamakawa. Tightly-Secure Key-Encapsulation Mechanism in the Quantum Random Oracle Model. IACR ePrint 2017/1005.
51. C.-P. Schnorr and M. Euchner. Lattice basis reduction: Improved practical algorithms and solving subset sum problems. *Mathematical programming*, 1994.
52. P. Schwabe, R. Avanzi, J. Bos, L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, J. M. Schanck, G. Seiler, and D. Stehle. Crystals-Kyber, Round 2 submission. Technical report, NIST, Post-Quantum Standardization Process Round 2, 2019.
53. D. Stehlé, R. Steinfeld, K. Tanaka, and K. Xagawa. Efficient public key encryption based on ideal lattices. pages 617–635.

54. E. E. Targhi and D. Unruh. Post-Quantum Security of the Fujisaki-Okamoto and OAEP Transforms. In *TCC 2016*. Springer, 2016.
55. R. Tsabary. Fully secure attribute-based encryption for t-cnf from lwe. In *CRYPTO 2019*. Springer, 2019.

Appendix

Appendix A Modelling decryption failures

As the failure probability of Learning with Errors based schemes is typically very low, it is not feasible to generate decryption failures by generating ciphertexts and subsequently testing them against the secret. In this section, we derive a method to generate failing ciphertexts with approximately the same distribution as the subset of randomly generated ciphertexts that lead to a failure event F_0. In other words, we aim at simulating $\mathbf{C}$ following the ciphertext distribution and such that

$$\overline{\mathbf{S}}^T \cdot \overline{\mathbf{C}^{(0)}} > q_t.$$

While randomly generated ciphertexts can be constructed as vectors with a certain norm $\|\mathbf{C}\|_2 \leftarrow \chi_{l \cdot N,\sigma}$ (recall that the norm is independent of the rotation, see Remark 3) and a uniformly random direction, ciphertexts that lead to F_0 will have a slightly different $\|\mathbf{C}\|_2$ distribution and will be more aligned with the secret $\overline{\mathbf{S}}$. Furthermore, these two distributions are not independent. One can separate both elements as follows:

$$P\left[cos(\theta_{SC^{(0)}}), \|\mathbf{C}\|_2 \;\mid\; \mathbf{S}, F_0\right] \tag{37}$$
$$= P\left[cos(\theta_{SC^{(0)}}) \mid\; \|\mathbf{C}\|_2, \mathbf{S}, F_0\right] \cdot P\left[\|\mathbf{C}\|_2 \mid\; \mathbf{S}, F_0\right]. \tag{38}$$

The norm distribution $P\left[\|\mathbf{C}\|_2 \mid \mathbf{S}, F_0\right]$ can be simplified as

$$P\left[\|\mathbf{C}\|_2 \mid\; \mathbf{S}, F_0\right] \tag{39}$$
$$= P\left[\|\mathbf{C}\|_2 \mid\; \|\mathbf{S}\|_2, F_0\right] \tag{40}$$
$$= C \cdot P\left[\|\mathbf{C}\|_2 \mid\; \|\mathbf{S}\|_2\right] \cdot P\left[F_0 \mid\; \|\mathbf{C}\|_2, \|\mathbf{S}\|_2\right] \tag{41}$$
$$= C \cdot P\left[\|\mathbf{C}\|_2 = l \mid\; l \leftarrow \chi_{l \cdot N,\sigma}\right] P\left[|\cos(t)| > \frac{q_t}{\|\mathbf{C}\|_2 \cdot \|\mathbf{S}\|_2} \;\middle|\; t \leftarrow \Theta_n\right], \tag{42}$$

where $C = P\left[F_0 \mid \|\mathbf{S}\|_2\right]^{-1}$ is a normalization constant that is independent of $\mathbf{C}$. The distribution of the angle $\theta_{SC^{(0)}}$ can be approximated as follows:

$$P\left[cos(\theta_{SC^{(0)}}) \;\mid\; \|\mathbf{C}\|_2, \mathbf{S}, F_0\right] \tag{43}$$
$$= P\left[cos(\theta_{SC^{(0)}}) \mid \|\mathbf{C}\|_2, \|\mathbf{S}\|_2, F_0\right] \tag{44}$$
$$= P\left[cos(\theta_{SC^{(0)}}) = cos(t) \;\middle|\; t \leftarrow \Theta_n \text{ with reject if: } |\cos(t)| < \frac{q_t}{\|\mathbf{C}\|_2 \cdot \|\mathbf{S}\|_2}\right]. \tag{45}$$

A failing ciphertext can now be approximated by first sampling a $\|\mathbf{C}\|_2$ and then an angle $\theta_{SC^{(0)}}$. At this point we want to generate a vector from the uniformly random distribution of vectors with these exact properties, which can be done using Algorithm 7 where a and b are chosen so that the angle θ_{SU} equals the sampled angle ψ:

$$a := \frac{\sin(\psi)}{\sin(\theta_{SU})} \tag{46}$$

$$b := \cos(\psi) - \frac{\sin(\psi)}{\tan(\theta_{SU})} \tag{47}$$

The notation $Distr_k$ takes the distribution on k associated with the probabilities for each k.

Algorithm 7: `GenFailures()`

```
1 l_c ← Distr_k (P[‖C‖_2 = k | S, F_0]) // sample the norm
2 ψ ← Distr_k (P[cos(θ_{SC^(0)}) = cos(k) | ‖C‖_2, S, F_0]) // sample the angle
3 Ū ← N_{0,1}(Z_q^{2Nl}) // sample a vector with random direction
4 C̄ = l_c · (a Ū/‖Ū‖_2 + b S̄/‖S̄‖_2) // generate vector with norm l_c and angle ψ
5 return C̄^(0)
```

1. $l_c \leftarrow Distr_k \left(P\left[\|\mathbf{C}\|_2 = k \,\middle|\, \mathbf{S}, F_0\right]\right)$ `// sample the norm`
2. $\psi \leftarrow Distr_k \left(P\left[cos(\theta_{SC^{(0)}}) = cos(k) \,\middle|\, \|\mathbf{C}\|_2, \mathbf{S}, F_0\right]\right)$ `// sample the angle`
3. $\overline{\mathbf{U}} \leftarrow \mathcal{N}_{0,1}(\mathbb{Z}_q^{2Nl})$ `// sample a vector with random direction`
4. $\overline{\mathbf{C}} = l_c \cdot (a\frac{\overline{\mathbf{U}}}{\|\overline{\mathbf{U}}\|_2} + b\frac{\overline{\mathbf{S}}}{\|\overline{\mathbf{S}}\|_2})$ `// generate vector with norm` l_c `and angle` ψ
5. **return** $\overline{\mathbf{C}^{(\mathbf{0})}}$

Appendix B Failure probability experiment

In this experiment, we are given an estimate of the secret $\overline{\mathbf{E}}$ and we want to determine the failure probability of a ciphertext $\mathbf{C}$ as a function of two inputs: its norm $\|\mathbf{C}\|_2$ (recall that the norm is independent of the rotation, see Remark 3) and its angle with $\overline{\mathbf{E}}$. As in section 6.2, we denote

$$\cos(\theta_{CE}) := \max_i \; \cos(\theta_{C^{(i)}E}).$$

We aim at evaluating $P[F \mid \|\mathbf{C}\|_2, \cos(\theta_{CE})]$. Without loss of generality, we will assume that $\text{argmax}_i \; \cos(\theta_{C^{(i)}E}) = 0$. As the failure events F_0 to F_{N-1} can be considered independent (see Assumption 2) we make the following approximation:

$$P[F \mid \|\mathbf{C}\|_2, \cos(\theta_{CE})] \tag{48}$$
$$\approx P[F_0 \mid \|\mathbf{C}\|_2, \cos(\theta_{CE})] + \sum_{i=1}^{N-1} P[F_i \mid \|\mathbf{C}\|_2, \cos(\theta_{CE})] \tag{49}$$
$$\approx P[F_0 \mid \|\mathbf{C}\|_2, \cos(\theta_{CE})] + \sum_{i=1}^{N-1} P[F_i \mid \|\mathbf{C}\|_2]. \tag{50}$$

Let us first estimate the first term of Equation 50. The latter can be rewritten using Bayes' theorem as

$$P[F_0 \mid \|\mathbf{C}\|_2, \cos(\theta_{CE})] \tag{51}$$
$$= \frac{P[\|\mathbf{C}\|_2, \cos(\theta_{CE}) \mid F_0] \cdot P[F_0]}{P[\|\mathbf{C}\|_2, \cos(\theta_{CE}) \mid F_0] \cdot P[F_0] + P[\|\mathbf{C}\|_2, \cos(\theta_{CE}) \mid S_0] \cdot P[S_0]}. \tag{52}$$

We discretize the probabilities by dividing the $\|\mathbf{C}\|_2 \times \cos(\theta_{CE})$ space into a discrete grid with squares $g_{\|\mathbf{C}\|_2, \cos(\theta_{CE})}$. Experimentally, we can determine $P[\|\mathbf{C}\|_2, \cos(\theta_{CE}) \mid F_0]$ for a certain roster, by generating n_f failing ciphertexts C_f and n_s successful ciphertexts C_s that fail in the first coefficient and counting the number of matching $\|\mathbf{C}\|_2, \cos(\theta_{CE})$ for each square in the grid. Our probability can now be approximated as

$$P[\|\mathbf{C}\|_2, \cos(\theta_{CE}) \mid F_0] = P[g_{\|\mathbf{C}\|_2, \cos(\theta_{CE})} \mid F_0] \approx \frac{\#(C_f \in g_{\|\mathbf{C}\|_2, \cos(\theta_{CE})})}{n_f}. \tag{53}$$

Similarly we can say that

$$P[\|\mathbf{C}\|_2, \cos(\theta_{CE}) \mid S_0] = P[g_{\|\mathbf{C}\|_2, \cos(\theta_{CE})} \mid S_0] \approx \frac{\#(C_s \in g_{\|\mathbf{C}\|_2, \cos(\theta_{CE})})}{n_s}. \tag{54}$$

Moreover, as the failure probability $P[F_0]$ is extremely low, we assume $P[S_0] = 1 - P[F_0] \approx 1$, so that

$$P[F_0 \mid \|\mathbf{C}\|_2, \cos(\theta_{CE})] \tag{55}$$

$$= \frac{\frac{\#(C_f \in g_{\|\mathbf{C}\|_2, \cos(\theta_{CE})})}{n_f} P[F_0]}{\frac{\#(C_f \in g_{\|\mathbf{C}\|_2, \cos(\theta_{CE})})}{n_f} P[F_0] + \frac{\#(C_s \in g_{\|\mathbf{C}\|_2, \cos(\theta_{CE})})}{n_s}}. \tag{56}$$

A similar derivation can be made for $\sum_{i=1}^{N-1} P[F_i \mid \|\mathbf{C}\|_2]$, where the norms are discretized in bands $b_{\|\mathbf{C}\|_2}$, which leads to

$$P[F_i \mid \|\mathbf{C}\|_2] = \frac{\frac{\#(C_f \in b_{\|\mathbf{C}\|_2})}{n_f} P[F_0]}{\frac{\#(C_f \in b_{\|\mathbf{C}\|_2})}{n_f} P[F_0] + \frac{\#(C_s \in b_{\|\mathbf{C}\|_2})}{n_s}}. \tag{57}$$

Using these equations, one can generate a 2D discrete grid with the failure probability for a given ciphertext, as shown in Figure 5b. A zoomed version where domain of the experiments and that of the theoretical values is approximately the same can be found in Figure 10.

Appendix C Work and queries derived using the computation technique of [16]

In the attack model of [16], the attacker does not use any information of already found failing ciphertexts. Hence, the probability of obtaining a certain number of ciphertexts n using Q decryption queries can be calculated using a cumulative binomial distribution as:

$$P_n = 1 - \sum_{i=0}^{n-1} \binom{Q}{i} \beta_{0,f_0}^i (1 - \beta_{0,f_0})^{Q-i}, \tag{58}$$

which can be approximated using the Poisson distribution since $\beta_{0,f_t}^{(0)}$ is very small and $Q\beta_{0,f_t}^{(0)}$ is approximately 1. This results in

$$P_n = 1 - e^{-q\beta_{0,f_0}} \sum_{i=0}^{n-1} \frac{(Q\beta_{0,f_0})^i}{i!}. \tag{59}$$

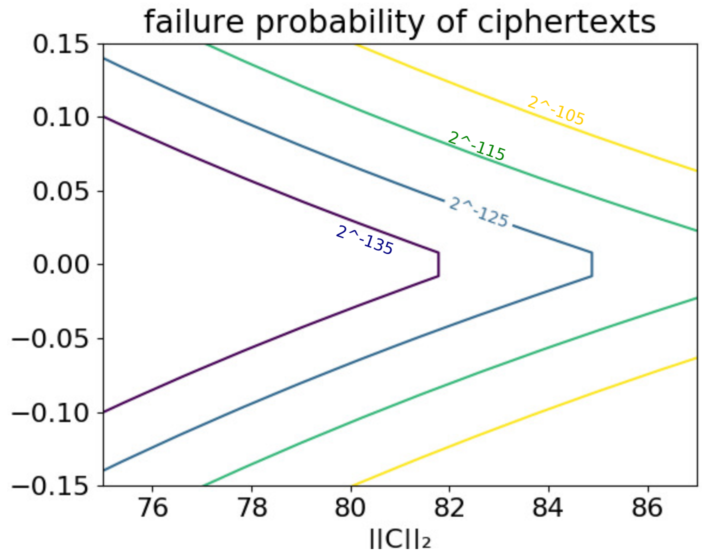
When $n \ll Q$, the number of queries q needed to obtain n failing ciphertexts with probability $P_n = 1 - e^{-1}$ is then approximately $Q = n/\beta_{0,f_0}$. Thus, to obtain at least n ciphertexts with probability $1 - e^{-1}$, $n/\beta_{0,f_0}$ queries and a total precomputational work of $n/(\sqrt{\alpha_{0,f_0}}\beta_{0,f_0})$ are required.

Appendix D Vertex probabilities

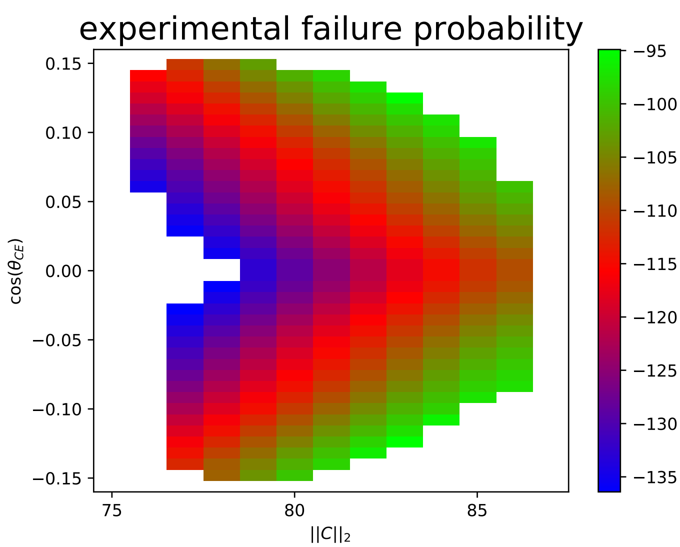
As discussed in § 5.2, the edge between two nodes $\mathbf{C}_i$ with rotation r and $\mathbf{C}_j$ with rotation s is weighted with $m_{i,j}(r,s) = P[\delta_{i,j} = r - s]$. The latter can be detailed as

$$m_{i,j}(r,s) = P\left[\delta_{i,j} = r - s \,\middle|\, \cos\left(\theta_{C_i^0 C_j^0}\right), \ldots, \cos\left(\theta_{C_i^0 C_j^{2N-1}}\right)\right]. \tag{60}$$

Applying Bayes' theorem gives



(a) Theoretical (experimental domain)



(b) Experimental

Fig. 10. Failure probability of ciphertexts as a function of $\|\mathbf{C}\|_2$ and $\cos(\theta_{CE})$. Same values as in Figure 5, with matching domains for an easier comparison.

$$m_{i,j}(r,s) = \frac{P\left[\delta_{i,j} = r-s\right] \cdot P\left[\cos\left(\theta_{C_i^0 C_j^0}\right), \ldots, \cos\left(\theta_{C_i^0 C_j^{2N-1}}\right) \,\middle|\, \delta_{i,j} = r-s\right]}{P\left[\cos\left(\theta_{C_i^0 C_j^0}\right), \ldots, \cos\left(\theta_{C_i^0 C_j^{2N-1}}\right)\right]}. \tag{61}$$

The prior probability $P\left[\delta_{i,j} = r-s\right]$ is constant for all $r-s$. Since the denominator does not depend on r and s, both terms can be grouped into a normalization constant denoted by C that makes the probabilities over all possible $r-s$ add to 1:

$$m_{i,j}(r,s) = C \cdot P\left[\cos\left(\theta_{C_i^0 C_j^0}\right), \ldots, \cos\left(\theta_{C_i^0 C_j^{2N-1}}\right) \,\middle|\, \delta_{i,j} = r-s\right]. \tag{62}$$

Furthermore, the probabilities of $\cos(\theta_{C_i^0 C_j^0}), \ldots, \cos(\theta_{C_i^0 C_j^{N-1}})$ can be considered to be independent, while $\cos(\theta_{C_i^0 C_j^k})$ equals $-\cos(\theta_{C_i^0 C_j{}^{k+N}})$. Considering the probability function of $\mathcal{C}(r)$ for $r = \delta_{i,j}$, which we will denote as ϕ, and the probability function of $\mathcal{C}(r)$ for $r \neq \delta_{i,j}$, which equals Θ_n, as depicted in Figure 3, we can simplify this expression further as

$$\begin{aligned}
& m_{i,j}(r,s) && (63)\\
&= C \cdot \prod_{k=0}^{N-1} P\left[\cos\left(\theta_{C_i^0 C_j^k}\right), \cos\left(\theta_{C_i^0 C_j^{k+N}}\right) \,\middle|\, \delta_{i,j} = r-s\right] && (64)\\
&= C \cdot \prod_{k=0}^{N-1} P\left[\cos\left(\theta_{C_i^0 C_j^k}\right) \,\middle|\, \delta_{i,j} = r-s\right] && (65)\\
&= C \cdot \left[\begin{array}{c} P\left[\cos\left(\theta_{C_i^0 C_j^{r-s}}\right) \,\middle|\, \delta_{i,j} = r-s\right] \\ \cdot \prod_{k=0,(k\neq r-s \bmod N)}^{N-1} P\left[\cos\left(\theta_{C_i^0 C_j^k}\right) \,\middle|\, \delta_{i,j} = r-s\right] \end{array}\right] && (66)\\
&= C \cdot \left[\begin{array}{c} P\left[\cos(x) = \cos\left(\theta_{C_i^0 C_j^{r-s}}\right) \,\middle|\, x \leftarrow \Phi\right] \\ \cdot \prod_{k=0,(k\neq r-s \bmod N)}^{N-1} P\left[\cos(x) = \cos\left(\theta_{C_i^0 C_j^k}\right) \,\middle|\, x \leftarrow \Theta_n\right] \end{array}\right]. && (67)
\end{aligned}$$

Replacing the normalization constant C by the normalization constant C', where

$$C' = C \cdot \prod_{k=0}^{N-1} P\left[\cos(x) = \cos\left(\theta_{C_i^0 C_j^k}\right) \,\middle|\, x \leftarrow \Theta_n\right], \tag{68}$$

which is also independent of $r-s$ and thus does not change the probability distribution, we obtain

$$m_{i,j}(r,s) = C' \cdot \frac{P\left[\cos(x) = \cos\left(\theta_{C_i^0 C_j^{r-s}}\right) \,\middle|\, x \leftarrow \Phi\right]}{P\left[\cos(x) = \cos\left(\theta_{C_i^0 C_j^{r-s}}\right) \,\middle|\, x \leftarrow \Theta_n\right]}. \tag{69}$$