A Critique on Average-Case Noise Analysis in RLWE-Based Homomorphic Encryption

Mingyu Gao
Tsinghua University
Shanghai Qi Zhi Institute
gaomy@tsinghua.edu.cn

Hongren Zheng
Tsinghua University
zhenghr22@mails.tsinghua.edu.cn

ABSTRACT

Homomorphic encryption schemes based on the Ring-Learning-with-Errors problem require accurate ciphertext noise analysis to ensure correctness and security. However, ring multiplications during homomorphic computations make the noise in the result ciphertexts difficult to characterize. Existing average-case noise analyses derive a bound on the noise by either assuming it follows a Gaussian distribution, or giving empirical formulae, with strong independence assumption and the Central Limit Theorem extensively applied. In this work, we question the validity of these methods, by showing that the noise exhibits a heavy-tailed distribution via exact calculation of its variance and kurtosis, for both independent and dependent noises. The heavy-tailedness suggests the failing probability of bounds derived from these methods may not be negligible, and we experimentally demonstrate several cases where the noise growth is underestimated.

1 INTRODUCTION

The Learning-with-Errors (LWE) problem [Reg05] and its ring-based variant, Ring-LWE (RLWE) [LPR10], have become the prominent foundation for constructing homomorphic encryption schemes. These schemes enable computations on encrypted messages while preserving data privacy. Among them, the most widely used RLWE-based constructions include BGV [BGV12], BFV [Bra12, FV12], and CKKS [CKKS17].

In these schemes, the data message is encrypted with the *noise* inherited from the RLWE problem in the same mathematical space. The message and the noise interact and tend to mix together as the computation goes on. However, to ensure correctness and security, the message must have its noise removed during decryption, i.e., the mathematical space should be large enough to ensure the noise does not overflow. On the other hand, the size of the message space should be kept minimal as it affects computational efficiency and practical usability. Therefore, a precise understanding of the noise behaviors is crucial for selecting appropriate parameters that balance security/correctness and efficiency.

However, the polynomial multiplication structure in RLWE makes the noise hard to characterize, even in the simple case where the initial noise is a polynomial with coefficients as independent Gaussian random variables. Traditional *worst-case* analyses like [BGV12, HS20, MML+23, CS16] set a probabilistic bound on the initial noise and track its evolution for each homomorphic operation. However, this approach is often too conservative and leads to prohibitively large parameters [CLP20].

In practice, *average-case* analyses are often preferred. Prior works like [CCH+24, CNP23, MP24] introduce a *heuristic* assuming that the product of independent Gaussian polynomials is still Gaussian using the Central Limit Theorem (CLT), and further assuming the independence of its coefficients so that this heuristic could be *repeatedly* applied as the computation goes on. These methods track the variance of the noise and use its Gaussian shape to derive a probabilistic bound on the noise.

The independence assumption is actually too strong and leads to underestimation of the noise, rendering it problematic to use. Ciphertexts in the same circuit are never independent because they at least share the same key. For ease of discussion, we classify the dependencies into two kinds: *common dependencies* and *ciphertext dependencies*. Common dependencies come from key materials, like the secret key, the noise in the public key, and the noise in the key-switching key. Ciphertext dependencies occur when the encryption randomness in the ciphertexts is correlated, such as when doing self-product of the same ciphertext.

The contributions of these dependencies are not negligible. Recent works like [BMCM23, BCM+24, BMMU25] focus on addressing the common dependencies and try to correct the variance. However, they do not handle the ciphertext dependencies. More crucially, their approaches still follow in the footsteps of deriving bounds using Gaussian distributions.

We show that the actual noise has a heavier tail than a Gaussian distribution, for both independent and dependent noises. We exactly calculate the variance and the kurtosis of the noise for multiplications. The kurtosis grows significantly with the multiplication depth, so CLT-based methods are not repeatedly applicable. The mathematical root reason is that the ring dimension N in RLWE is not large enough for the CLT to apply *in deep multiplications*. We also experimentally verify that using the Gaussian distribution will lead to underestimation of the actual noise.

We emphasize that, with a heavy-tailed distribution, the variance alone is not enough to derive a bound. This naturally shows that **the variance-based methods above do not give theoretical guarantees on its failing probability**.

Another line of average-case analyses [HPS19, KPZ21] derive empirical formulae using independent random polynomials. However, they lack justification for dependent polynomials. We experimentally demonstrate that **the empirical formulae give underestimation for dependent polynomials** like the self-product of the secret key. Such a structure is pervasive as it exists even in the noise of the products of *independent* ciphertexts. Although in OpenFHE [BAB+22] we do not observe decryption failure caused by such underestimation, we point out it is because of various other loose inequalities.

2 BACKGROUND

By the modulo operation $[\cdot]_q$ we reduce the value to the range $(-q/2, q/2]$.

2.1 Ring, Embeddings, and Norms

Coefficient Embedding. Let m be a power-of-two number, $\varphi(\cdot)$ be the totient function, and $N = \varphi(m) = m/2$. The mth cyclotomic polynomial is $\Phi_m(X) = X^N + 1$. The rings we will use are $\mathcal{R} = \mathbb{Z}[X]/\Phi_m(X)$ and $\mathcal{R}_Q = \mathbb{Z}_Q[X]/\Phi_m(X)$.

An element $f \in \mathcal{R}$ could be written as $f(X) = \sum_{\ell=0}^{N-1} f|_\ell X^\ell$, or viewed as a vector $(f|_\ell) \in \mathbb{Z}^N$ where $f|_\ell$ is the ℓ-th coefficient. This is called the *coefficient embedding* of f.

For f and g in $\mathcal{R}$, their sum is elementwise and their multiplication is negacyclic convolution

$$f \cdot g|_\ell = \sum_{j+k \equiv \ell \pmod N} \xi(j+k) f|_j \cdot g|_k$$

where $\xi(a) = (-1)^{\lfloor \frac{a}{N} \rfloor}$.

When dealing with probability we may also write $f \leftarrow \mathbb{R}^N$ and treat it as a polynomial.

Canonical Embedding. Let $\sigma_i : \mathcal{R} \to \mathbb{C}$ by mapping X to ω_m^i, where $\omega_m = \exp(2\pi\sqrt{-1}/m)$. Note that σ_i is a ring homomorphism. As there are N primitive roots of unity in $\mathbb{C}$, let $\sigma(\cdot) = (\sigma_i(\cdot)), i \in \mathbb{Z}_m^*$ that maps $\mathcal{R}$ to $\mathbb{C}^N$, σ is again a ring homomorphism, and in $\mathbb{C}^N$, both addition and multiplication become elementwise. $\sigma(f)$ is called the *canonical embedding* of f.

Embedding Norms. For $f \in \mathcal{R}$, the infinity norm of its canonical embedding is $||f||^{\text{can}} = ||\sigma(f)|| = \max\{|\sigma_i(f)|\}$. Since the multiplication is elementwise, we have

$$||f \cdot g||^{\text{can}} \le ||f||^{\text{can}} \cdot ||g||^{\text{can}}$$

The infinity norm of the coefficient embedding is $||f|| = \max\{|f|_i|\}$, and we have the following results from [DPSZ12]

$$||f \cdot g|| \le N \cdot ||f|| \cdot ||g||$$
$$||f|| \le ||f||^{\text{can}}$$

N here is often referred to as the *worst-case expansion factor*.

2.2 Probability

Let χ_s be the uniform ternary distribution, i.e., sampling uniformly from $\{-1, 0, 1\}$; χ_e be the zero-mean *discrete* Gaussian distribution with variance σ_e^2 [ACC+19]; $\mathcal{N}(0, \sigma^2)$ be the zero-mean *continuous* Gaussian distribution; $\mathcal{U}(-a, a)$ be the zero-mean *continuous* uniform distribution in $[-a, a]$.

When we say $f \leftarrow \chi$ for a $f \in \mathcal{R}$, we mean $f|_i$ are independent and identically distributed (IID) random variables sampled from χ. Similarly we define $f \leftarrow \mathcal{N}$ for $f \in \mathbb{R}^N$. We may directly say a polynomial is Gaussian or uniform, which means its coefficients are IID from the corresponding distribution.

Failing Probability. For $X \sim \mathcal{N}(0, \sigma^2)$, a $D\sigma$ bound has failing probability of

$$\epsilon = \Pr[|X| > D\sigma] = \texttt{erfc}(D/\sqrt{2})$$

where erfc is the complementary error function.

We give a table on D and ϵ in Table 1 in a similar style to [BJSW25]. We note that previous works [KPZ21, CS16] commonly use $D = 6$ but the failing probability is 2^{-28}, while [HS20] uses $D = 10$ with a failing probability of 2^{-75}.

Table 1: Deviation D and failing probability ϵ.

D	3	6	9.16	10	13.11
$\log_2(\epsilon)$	-8	-28	-64	-75	-128

For a polynomial f with N coefficients, the $D\sigma$ bound on its infinity norm $||f||$ has a failing probability upper bound of $N\epsilon$ by the union bound[1]. For $N = 2^{16}$, the failing probability is only upper bounded by 2^{-12}.

Kurtosis. We recall the quantity *kurtosis* for a zero-mean random variable X as $\text{Kurt}[X] = \mathbb{E}[X^4]/(\mathbb{E}[X^2])^2$. Informally, the kurtosis measures the tailedness [Wes14]. A continuous Gaussian distribution has kurtosis 3. In this work we call a distribution with kurtosis larger than 3 as a heavy-tailed distribution. The larger the kurtosis, the more it deviates from the Gaussian distribution.

2.3 RLWE-Based Homomorphic Encryption Schemes

This section first describes common features for the RLWE-based homomorphic encryption schemes we use. We mainly focus on the multiplication of ciphertexts in BGV and BFV.

Encryption and Decryption. The secret key has the form $\mathsf{sk} = (1, s)$ where $s \in \mathcal{R}_Q$ is sampled from χ_s.

Encryption $\mathsf{ct} = \mathsf{Enc}_{\mathsf{sk}}(m)$ could be expressed as $\mathsf{ct} = (-as, a) + (m + e, 0) \in \mathcal{R}_Q^2$, where the noise e is sampled from χ_e and the mask a is sampled uniformly. Note that the message m or the noise e might be scaled by the scheme.

Homomorphic operations ensure the form $m+v$ to be maintained inside a ciphertext, where v is viewed as the noise. We refer to $m+v$ as the message-and-noise term.

Decryption $m = \mathsf{Dec}_{\mathsf{sk}}(\mathsf{ct})$ could be divided into two steps, first getting the message-and-noise term by taking an inner product with the secret key $\langle \mathsf{ct}, \mathsf{sk} \rangle = m + v$; then getting the message by a rounding process Round specific to the scheme $m = \mathsf{Round}(m + v)$.

For public key encryption, the secret key could be transformed into a public key $\mathsf{pk} = (b, a) = (-as + e_{\mathsf{pk}}, a) \in \mathcal{R}_Q^2$, and the public key encryption is constructed as

$$\mathsf{ct} = (b \cdot u, a \cdot u) + (m + e_0, e_1)$$
$$\langle \mathsf{ct}, \mathsf{sk} \rangle = m + (e_0 + u \cdot e_{\mathsf{pk}} + e_1 \cdot s)$$
$$v = e_0 + u \cdot e_{\mathsf{pk}} + e_1 \cdot s$$

with $u \leftarrow \chi_s$ and $e_i \leftarrow \chi_e$.

Ciphertext Multiplication. Ciphertext multiplication often involves multiple steps. The core step is the tensor product

$$\mathsf{ct}_0 \otimes \mathsf{ct}_1 = (c_0^{(0)} c_0^{(1)}, c_1^{(0)} c_0^{(1)} + c_0^{(0)} c_1^{(1)}, c_1^{(0)} c_1^{(1)})$$

In this step, the message-and-noise term $m + v$ also gets multiplied by other terms, so the noise experiences ring multiplication here.

Modulus Switching. A ciphertext in $\mathcal{R}_{Q_L}$ could be switched to $\mathcal{R}_{Q_{L-1}}$ by a process called *modulus switching*, where $Q_L = q_L \cdot Q_{L-1}$. This could be done by dividing the ciphertext by q_L. However, its

[1] Independence could make it tighter to $1 - (1 - \epsilon)^N$, but it is still similar to $N\epsilon$.

coefficients may not be divisible by q_L directly so a rounding term is introduced to make the coefficients divisible by q_L

$$\delta = ([-c_0]_{q_L}, [-c_1]_{q_L}) = (\tau_0, \tau_1) \cdot q_L \qquad \tau_i \in [-1/2, 1/2]$$
$$\delta \equiv (-c_0, -c_1) \equiv -\mathtt{ct} \pmod{q_L}$$

Then $\mathtt{ct} + \delta$ could be divided by q_L, and the resulting ciphertext and its noise are

$$\mathtt{ms}(\mathtt{ct}) = \frac{\mathtt{ct} + \delta}{q_L} = \left\lfloor \frac{\mathtt{ct}}{q_L} \right\rceil + (\tau_0, \tau_1)$$
$$v_{\mathtt{ms}(\mathtt{ct})} = \left\lfloor \frac{v_{\mathtt{ct}}}{q_L} \right\rceil + \tau_0 + \tau_1 s$$

Here if q_L is large enough, the major term is $\tau_1 s$, a product of a polynomial and the secret key. Often τ_1 is assumed to be uniform and independent for different ciphertexts.

2.3.1 BGV. The message-and-noise term has the form $m + t \cdot e$ with a plaintext modulus t separating the noise in the high bits from the message in the low bits.

The public key is adapted to $\mathsf{pk} = (b, a) = (-as + t \cdot e_{\mathsf{pk}}, a) \in \mathcal{R}_Q^2$ where the error is put in the high bits. For a freshly public-key encrypted ciphertext, its noise is

$$v = t(e_0 + u \cdot e_{\mathsf{pk}} + e_1 \cdot s)$$

After the tensor product, the noise is

$$m_2 + v_2 = (m_0 + v_0) \cdot (m_1 + v_1)$$
$$v_2 = v_0 v_1 + m_0 v_1 + m_1 v_0 + (m_1 m_0 - m_2)$$

$v_0 v_1$ is the major term as $||m_i|| < t$ and $||v_i|| > t$.

2.3.2 BFV. The message-and-noise term has the form $\left\lfloor \frac{Q}{t} m \right\rceil + e$ by putting the message in the high bits [KPZ21].

To conduct a multiplication, we first need to convert ciphertexts in $\mathcal{R}_Q$ to $\mathcal{R} = \mathbb{Z}[X]/\Phi_m(X)$, whose message-and-noise term has the form

$$\langle \mathtt{ct}, \mathtt{sk} \rangle = \left\lfloor \frac{Q}{t} m \right\rceil + e + hQ$$
$$hQ = c_0 + c_1 s - \left\lfloor \frac{Q}{t} m \right\rceil - e$$

We call h the high term. In it $c_1 s$ is the major term. As $c_i \in [-Q/2, Q/2]$, we can approximate h with μs where $\mu \in [-1/2, 1/2]$. Prior works often assume μ is uniform and independent for different ciphertexts.

After the tensor product, the result message has the scale of $\frac{Q^2}{t^2}$, and the noise becomes

$$\frac{Q}{t}\left(\left\lfloor \frac{Q}{t} m_0 m_1 \right\rceil + v_2 + h_2 Q\right) = \left(\left\lfloor \frac{Q}{t} m_0 \right\rceil + v_0 + h_0 Q\right) \cdot \left(\left\lfloor \frac{Q}{t} m_1 \right\rceil + v_1 + h_1 Q\right)$$
$$v_2 = t(v_0 h_1 + v_1 h_0) + m_0 v_1 + m_1 v_0 + \frac{t}{Q} v_0 v_1 + \cdots$$

The major term in v_2 is $t(v_0 h_1 + v_1 h_0)$.

3 BRIEF REVIEW OF EXISTING NOISE ANALYSES

We remark that, no matter worst-case or average-case analyses, they only provide probabilistic bounds due to the unbounded nature of the Gaussian distribution[2]. As far as we are aware, an exact calculation of the failing probability in all analyses for RLWE remains an open problem. On one hand, the looseness of certain bounds reduces the failing probability by a large margin. On the other hand, aggressive heuristics and dependencies among polynomials may make the failing probability larger than conjectured.

3.1 Worst-Case

Worst-case noise analyses give a *bound* on the initial noise and track its evolution for each homomorphic operation.

Coefficient Embedding. For key $s \leftarrow \chi_s$ and error $e \leftarrow \chi_e$, the bounds $B_{\text{key}} = 1$ and $B_{\text{err}} = D\sigma_e$ are used. The former is a tight bound, whereas the latter is a probabilistic bound.

For the product of $f, g \leftarrow \mathcal{R}$, the bound evolves as

$$||f \cdot g|| \le N \cdot ||f|| \cdot ||g||$$

However, the expansion factor N is often too loose to use in practice.

Canonical Embedding. For a polynomial $f \leftarrow \mathcal{N}(0, \sigma^2)$ in $\mathbb{R}^N$, $f(\omega_m^i)$ is distributed as a complex Gaussian with variance $N\sigma^2$. So the canonical embedding $\sigma(f)$ can be *probabilistically* bounded by

$$||f||^{\text{can}} \le D\sigma\sqrt{N}$$

For other polynomials like $s \leftarrow \chi_s$, $s(\omega_m^i)$ is heuristically assumed to be distributed as a complex Gaussian using the CLT [GHS12, HS20], and similar bounds can be derived. Works like [MML+23, CS16] extend this idea to terms like es and τs.

For the product of $f, g \leftarrow \mathcal{R}$, the bound evolves as

$$||f \cdot g||^{\text{can}} \le ||f||^{\text{can}} \cdot ||g||^{\text{can}}$$

Finally we can give a bound for the coefficient embedding by $||f \cdot g|| \le ||f \cdot g||^{\text{can}}$. This is often loose.

3.2 Average-Case: Variance-Based

This line of work tracks the *variance* of the noise by the following heuristic:

Heuristic 1 (Gaussian). *The noise in all ciphertexts can be well approximated by Gaussian polynomials.*

This heuristic is established by *extensively* applying CLT [CCH+24] or experimentally verifying [BMCM23]. With it, a bound on the noise could be derived as $D\sigma$.

Independence. [CCH+24, CNP23, MP24] assumed the independence of the coefficients of the noise.

Heuristic 2 (Independence [CCH+24]). *For two independent Gaussian polynomials $f, g \in \mathbb{R}^N$ with coefficients from $\mathcal{N}(0, \rho^2)$ and $\mathcal{N}(0, \rho'^2)$, respectively, their product $h = f \cdot g$ could still be approximated by a Gaussian polynomial with independent coefficients, each with variance*

$$\rho_h^2 = N \cdot \rho^2 \cdot \rho'^2$$

[2] Except [HS20] which employs rejection sampling to ensure probability-1 bounds.

However, this method often underestimates, as the independence assumption is often too strong.

Classification of Dependencies. We classify the dependencies into two kinds: *common dependencies* and *ciphertext dependencies*. Common dependencies come from key materials used in the scheme, such as the secret key s, the noise in the public key e_{pk}, and the noise in the key-switching key. Ciphertext dependencies are resulted from the encryption randomness used in the ciphertexts, such as e_{ct} and u_{ct}, and other terms like the modulus switching error τ_{ct} and the high term h_{ct} in BFV.

Addressing Common Dependencies. To mitigate the problematic independence assumption, [BMCM23, BCM+24, BMMU25] essentially track the *common dependencies* alongside the variance, like the degree of s^k, and give a *correction factor* based on such degrees.

For a multiplication of two polynomials with variance ρ^2 and ρ'^2, and with the secret key degrees k and k', the result variance is

$$\rho_{\text{mul}}^2 = N \cdot F(k, k') \cdot \rho^2 \cdot \rho'^2$$

where F is the correction function. Notably, [BMCM23, BMMU25] establish the correction function through experiments, and we will comment on this approach later.

Their method only applies to independent ciphertexts, i.e., when the *ciphertext dependencies* do not introduce other correction factors. [BMCM23] studied an example where the contribution of the ciphertext dependencies could be ignored, but generally it is not the case. Even if further development in this line may be able to introduce all correction factors, they are still using the Gaussian Heuristic to derive the bound.

3.3 Average-Case: Experiment-Based

[HPS19, KPZ21] experimentally determine the bound of the product of discrete Gaussian and uniform ternary polynomials in the coefficient embedding, and use $C \cdot \sqrt{N}$ where $C = 2$ as the *average-case expansion factor*. Namely, for random polynomials $f, g \leftarrow \mathcal{R}$ that are either from χ_s or χ_e, they have

$$||f \cdot g|| \le 2\sqrt{N} \cdot ||f|| \cdot ||g||$$

They then use it for all other multiplications of random polynomials.

4 GAUSSIAN NOISE IS NOT GAUSSIAN AFTER MULTIPLICATION

In this section we study the behavior of products of Gaussian random polynomials for both independent and dependent cases. We show that after a few multiplications, the distribution of the resulting polynomial is no longer approximatable by the Gaussian distribution as it has larger kurtosis. This contradicts the Gaussian Heuristic.

For the ease of exposition, we directly use $\mathcal{N}(0,1)$ for all Gaussian random variables as the parameterized version could be easily derived with proper scaling. We use the continuous Gaussian distribution instead of discrete Gaussian as it is easier to analyze.

4.1 Independent Gaussian

Theorem 4.1 (Product of Independent Gaussian Polynomials). *For k independent polynomials $f_1, f_2, \cdots, f_k \leftarrow \mathcal{N}(0,1)$ and $f_i \in \mathbb{R}^N$, with an even N, the ℓ-th coefficient $F|_\ell$ of their product $F = \prod_i^k f_i$ has the following properties*

$$\begin{aligned}
\mathbb{E}\left[F|_\ell\right] &= 0 \\
\text{Var}\left(F|_\ell\right) = \mathbb{E}\left[F|_\ell^2\right] &= N^{k-1} \\
\text{Cov}\left(F|_\ell, F|_{\ell'}\right) = \mathbb{E}\left[F|_\ell \cdot F|_{\ell'}\right] &= 0 \qquad (\ell \neq \ell') \\
\mathbb{E}\left[F|_\ell^4\right] &= 3N^{2k-2} + 3(2^k - 2)N^{2k-3} \\
\text{Kurt} &= 3 + 3\frac{2^k - 2}{N}
\end{aligned}$$

Proof. Here is the proof for the variance. Other parts are deferred to the appendix.

Let $\mathcal{I}_\ell = \{(\alpha_1, \alpha_2, \cdots, \alpha_k) | \alpha_j \in [N], \sum_j \alpha_j \equiv \ell \pmod N\}$ be the index set for ℓ. It could be viewed as an assigning problem with $k-1$ degrees of freedom, so $|\mathcal{I}_\ell| = N^{k-1}$.

For $\alpha = (\alpha_j) \in \mathcal{I}_\ell$, we use the notation $\xi(\alpha) = \xi(\sum \alpha_j)$, and let $Y_\alpha = \prod_j f_j|_{\alpha_j}$ then we can express the product as

$$\begin{aligned}
F|_\ell &= \sum_{\alpha \in \mathcal{I}_\ell} \xi(\alpha) Y_\alpha \\
F|_\ell^2 &= \sum_{\alpha, \beta \in \mathcal{I}_\ell} \xi(\alpha)\xi(\beta) Y_\alpha Y_\beta
\end{aligned}$$

Notice that $Y_\alpha Y_\beta$ could be written in this form

$$\begin{aligned}
Y_\alpha Y_\beta &= \prod_j f_j|_{\alpha_j} \cdot f_j|_{\beta_j} \\
\mathbb{E}[Y_\alpha Y_\beta] &= \prod_j \mathbb{E}\left[f_j|_{\alpha_j} \cdot f_j|_{\beta_j}\right]
\end{aligned}$$

as the f_j are independent. The only way it has non-zero value is when $\forall j, \alpha_j = \beta_j$, so $\alpha = \beta$. The sign is always $\xi(\alpha)^2 = 1$. As there are in total N^{k-1} possible α, the variance is N^{k-1}. □

Remark 1. *When $k = 2$, for a general even N, the above distribution is a Variance-Gamma distribution as it is a summation of products of two Gaussian variables ([Gau14], Corollary 2.5). Especially, the $k = 2, N = 2$ case is exactly the Laplace distribution.*

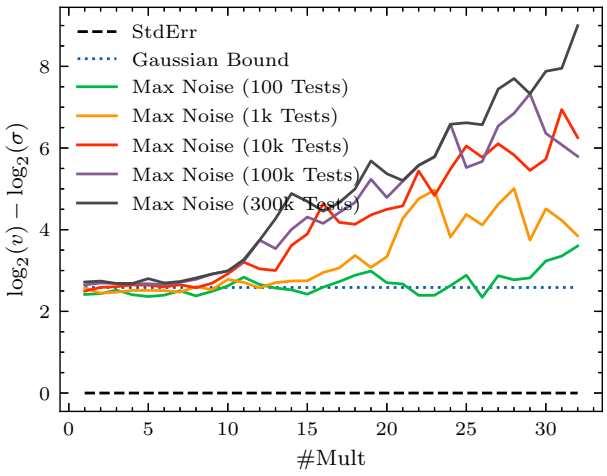
This theorem partially aligns with the Gaussian Heuristic when $k = 2$ as the extra term in the kurtosis $\frac{2}{N}$ for typical $N = 2^{16}$ is small. However, after several multiplications, say $k = 17$, the kurtosis will be larger than 6 and the distribution will become no longer Gaussian.

Note that when N goes to infinity, the kurtosis will be 3 for any k, aligning with the CLT. But in practice, N is not so large and the multiplication depth k will dominate its growth. This suggests that the CLT should not be applied for *deep multiplications* in RLWE settings.

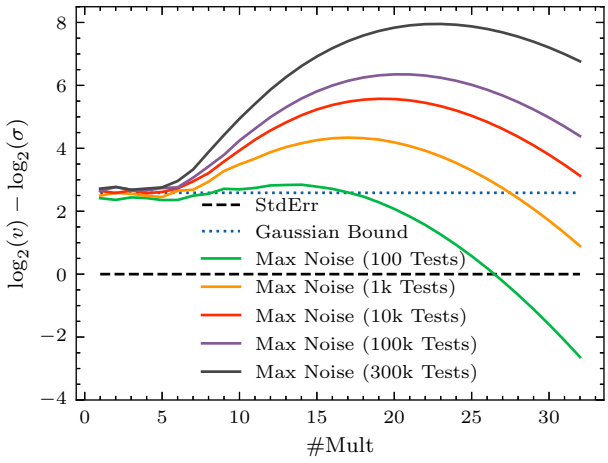
We have also experimentally verified that Gaussian distribution could not be used to bound the noise in Figure 1a. Initially, the observed maximal noise is close to the Gaussian bound with $D = 6$. However, after deep multiplications, it exceeds the Gaussian bound.

4.2 Dependent Gaussian

The heavy-tail situation will become worse when we are facing dependent polynomials. First we introduce a lemma that converts the high-order moment calculation of Gaussian random variables to a combinatorial counting problem.



(a) Independent Gaussian Polynomials.



(b) Same Gaussian Polynomial.

Figure 1: The maximal noise observed in experiments compared with the standard error σ and Gaussian bound 6σ. The input polynomial e is sampled from χ_e for each test with $N = 2^{16}$. The left figure demonstrates the maximal noise of $v = e_1 e_2 \cdots e_k$, while the right figure demonstrates the maximal noise of $v = e^k$. The x-axis is the number of multiplications, ranging from 1 to 32, and the y-axis is the logarithm of the maximal noise observed, minus the logarithm of the standard error. We draw the maximal noise for 100, 1k, 10k, 100k, and 300k tests to demonstrate the heavy-tailedness of the distribution as the maximal noise grows significantly with the number of tests. Some data points with fewer tests are higher than those with more tests, because data for different numbers of tests are individually collected.

Lemma 4.2 (Isserlis [Iss18]). *Let $X_1, X_2, \cdots, X_n$ be (possibly dependent) Gaussian random variables with mean 0, then the expectation of their product is*

$$\mathbb{E}\left[X_1 X_2 \cdots X_n\right] = \sum_{p \in PP(n)} \prod_{(i,j)\in p} \mathbb{E}\left[X_i X_j\right]$$

where $PP(n)$ is the set of partitions of $[n] = \{1, 2, \ldots, n\}$ into disjoint pairs. Note that $|PP(n)| = (n-1)!!$ when n is even. An example for $n = 4$ is

$$\begin{aligned}&\mathbb{E}[X_1X_2X_3X_4]\\ =&\mathbb{E}[X_1X_2]\mathbb{E}[X_3X_4] + \mathbb{E}[X_1X_3]\mathbb{E}[X_2X_4] + \mathbb{E}[X_1X_4]\mathbb{E}[X_2X_3]\end{aligned}$$

Theorem 4.3 (Power of Gaussian Polynomials). *For a polynomial $f \leftarrow \mathcal{N}(0,1)$ and $f \in \mathbb{R}^N$, with an even N, the ℓ-th coefficient of $F = f^k$ has the following properties*

$$\begin{aligned}\mathbb{E}[F|_\ell] &= 0\\ \mathrm{Var}(F|_\ell) = \mathbb{E}[F|_\ell^2] &= k!N^{k-1}\\ \mathrm{Cov}(F|_\ell, F|_{\ell'}) = \mathbb{E}[F|_\ell \cdot F|_{\ell'}] &= 0 \qquad (\ell \neq \ell')\\ \mathbb{E}[F|_\ell^4] &= 3(k!)^2N^{2k-2} + 3((2k)! - 2(k!)^2)N^{2k-3}\\ \mathrm{Kurt} &= 3 + 3\frac{\binom{2k}{k} - 2}{N}\end{aligned}$$

Proof. Here is the proof for the variance that shows the essense of our proof technique. Other parts are deferred to the appendix.

Let $\mathbf{1}$ be the indicator function. Following the notation in Theorem 4.1.

$$\begin{aligned}F|_\ell^2 &= \sum_{\alpha,\beta\in\mathcal{I}_\ell} \xi(\alpha)\xi(\beta)Y_\alpha Y_\beta\\ \mathbb{E}[Y_\alpha Y_\beta] &= \sum_{p\in PP(\alpha,\beta)} \prod_{(i,j)\in p} \mathbb{E}[f|_i \cdot f|_j] = \sum_{p\in PP(\alpha,\beta)} \mathbf{1}_{\forall(i,j)\in p, i=j}(p)\end{aligned}$$

where $PP(\alpha,\beta)$ is the set of partitions of $\{\alpha_1, \cdots, \alpha_k, \beta_1, \cdots, \beta_k\}$ into disjoint pairs. Let $PP(2k)$ be the set of partitions of $[2k] = \{1, 2, \cdots, k, k+1, \cdots, 2k\}$ into disjoint pairs, $PP(\alpha,\beta)$ is an instantiation of $PP(2k)$ by substituting i with α_i and $k+j$ with β_j.

For a partition $p \in PP(\alpha,\beta)$, it may contain pairs like (α_i, β_j), (α_i, α_j), or (β_i, β_j). The indicator function $\mathbf{1}_{\forall(i,j)\in p, i=j}(p)$ has non-zero value if given α and β and the partition p, all pairs in it satisfy the equality.

With change of summation order among α, β, and p, we can express the variance as

$$\mathbb{E}[F|_\ell^2] = \sum_{p\in PP(2k)} \sum_{\alpha,\beta\in\mathcal{I}_\ell} \xi(\alpha)\xi(\beta)\mathbf{1}_{\forall(i,j)\in p, i=j}(p)$$

Now we assert that we only need to consider those p that only contain (α_i, β_j) pairs. The reason is that for any p containing (α_i, α_j), for the indicator to be non-zero, we have $\alpha_i = \alpha_j$, then we can find another α' with only these two points changed to $\alpha'_i = \alpha'_j = \alpha_i + \frac{N}{2} \pmod N$ and other points fixed. Here note N is even. Then this will inflict a sign change $\xi(\alpha') \neq \xi(\alpha)$ as the summation of α' is exactly one N off that of α and they will cancel out. The same applies for (β_i, β_j).

Given a partition p that only contains pairs of form (α_i, β_j), for the indicator function to be non-zero, we have $\alpha_i = \beta_j$ for all pairs, so an α uniquely determines a β and $\xi(\alpha) = \xi(\beta)$. As there are $k!$ possible p, the final variance is $k!N^{k-1}$. □

For $N = 2^{16}$, $k = 10$ will make the kurtosis larger than 6, making it no longer approximatable by a Gaussian.

We have also experimentally verified that Gaussian distribution could not be used to bound the noise in Figure 1b. Compared with Figure 1a, the maximal noise exceeds the Gaussian bound much earlier due to the faster growing kurtosis. We also note that for deeper multiplications, a few tests seem to suggest that the Gaussian bound could be used, but after more tests the maximal noise grows beyond the Gaussian bound. This is caused by the heavy-tailedness of the distribution. We believe that with more testing the peak of the maximal noise lines in Figure 1b move towards the upper right. Already, the current peak of 8 suggests a $D = 2^8 = 64$ deviation, and if it conforms to Gaussian distribution, it means a failing probability $\texttt{erfc}(64/\sqrt{2}) \approx 2^{-2961}$, which should not happen in practice.

4.3 Variance of Noise

A *noise expression* is a summation of products of independent *Gaussian* polynomials

$$v = \sum_i \prod_j f_{i,j}^{k_{i,j}}$$

where $k_{i,j}$ is the degree of the polynomial $f_{i,j}$. Within one i, all $f_{i,j}$ are different, but between i and i' it may happen that $f_{i,j} = f_{i',j'}$.

We define the degree of the i-th term as $k_i = \sum_j k_{i,j}$, and the degree of the noise as $k = \max_i k_i$.

For mixed product of polynomials, we can use similar argument and get the following result.

Proposition 4.4 (Product of Powers). *For u independent polynomials $f_1, f_2, \cdots, f_u \leftarrow \mathcal{N}(0,1)$ and $f_i \in \mathbb{R}^N$, with an even N, given a tuple $(k_1, k_2, \cdots, k_u)$ such that $k_i > 0$ and $\sum_i^u k_i = k$, the ℓ-th coefficient of $F = \prod f_i^{k_i}$ has the following properties*

$$\mathbb{E}[F|_\ell^2] = N^{k-1} \prod_i k_i!$$

$$\text{Kurt} = 3 + 3\frac{\prod_i \binom{2k_i}{k_i} - 2}{N}$$

Then for deriving the variance of the noise expression, as there are summations, we need to examine the covariance of the i-th and i'-th terms. It turns out that their covariance is 0.

Proposition 4.5 (No Covariance). *Let F and F' be the i-th and i'-th terms in the noise expression above, then the covariance of the ℓ-th coefficient of $F = \prod_j f_{i,j}^{k_{i,j}}$ and $F' = \prod_j f_{i',j}^{k_{i',j}}$ is*

$$\text{Cov}(F|_\ell, F'|_\ell) = \mathbb{E}[F|_\ell \cdot F'|_\ell] = 0$$

Proof. Following the proof in Theorem 4.3. If a partition of α, β exists, there will always be a pairing inside α (or β), as otherwise F is identical with F', so we can find another α' (or β') to cancel out. □

Finally we can derive the variance of the noise expression.

Theorem 4.6 (Variance of Noise). *The variance of the ℓ-th coefficient of the noise expression v is*

$$\text{Var}(v|_\ell) = \sum_j N^{k_j - 1} \prod_i k_{i,j}!$$

We remark that the exact derivation of the kurtosis for a general noise expression remains open as it is a more complicated combinatorial expression. But the two extreme cases in the previous sections already demonstrate its heavy-tailedness. Also, Figure 3 experimentally shows the heavy-tailedness of real-world noise expressions when having large degrees.

5 POWER OF SECRET KEY / UNIFORM NOISE

We also study the distribution for the self-product of the secret key. For products of independent polynomials from the secret key distribution, their variance follows the N^{k-1} growth. However, for dependent polynomials, the variance is more complicated, as we demonstrate below

Example 5.1. For a $f \leftarrow \chi_s$ and $f \in \mathcal{R} = \mathbb{Z}[X]/(X^N + 1)$ where N is even, as the variance of χ_s is $\sigma_s^2 = 2/3$, the ℓ-th coefficient of f^2 has the following property

$$\text{Var}[f^2|_\ell^2] = \begin{cases} 2N\sigma_s^4 & \ell \text{ is odd} \\ (2N-3)\sigma_s^4 & \ell \text{ is even} \end{cases}$$

Proof. Let $\mathbf{1}$ be the indicator function, and Even be the set of even numbers, we have

$$\begin{aligned} f^2|_\ell &= \sum_{j+k \equiv \ell \pmod N} \xi(j+k) f_j f_k \\ &= \sum_{j<k} 2\xi(j+k) f_j f_k + \mathbf{1}_{\text{Even}}(\ell) \left(f_{\frac{\ell}{2}}^2 - f_{\frac{\ell+N}{2}}^2 \right) \\ \mathbb{E}[f^2|_\ell] &= 0 \\ \mathbb{E}[f^2|_\ell^2] &= \sum \xi(j+k)\xi(j'+k')\mathbb{E}[f_j f_{j'} f_k f_{k'}] \\ &= 2\sum_{j<k} \binom{2}{1} \mathbb{E}[f_j^2]\mathbb{E}[f_k^2] \\ &\quad + \mathbf{1}_{\text{Even}}(\ell) \left(\mathbb{E}[f_{\frac{\ell}{2}}^4] + \mathbb{E}[f_{\frac{\ell+N}{2}}^4] - 2\mathbb{E}[f_{\frac{\ell}{2}}^2]\mathbb{E}[f_{\frac{\ell+N}{2}}^2] \right) \\ &= 4\frac{N - 2 \cdot \mathbf{1}_{\text{Even}}(\ell)}{2}\sigma^4 + \mathbf{1}_{\text{Even}}(\ell) \cdot \sigma^4 \end{aligned}$$

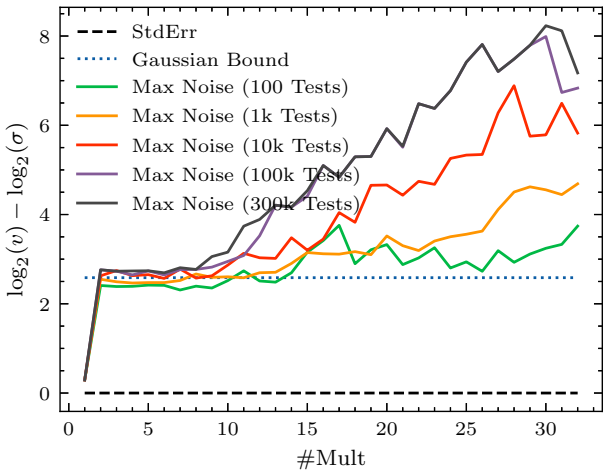
□

For higher-order powers like f^k, the variance of $f^k|_\ell$ will be more complicated depending on N and ℓ. We heuristically use $k!N^{k-1}$ to approximate them as the cross terms (i.e., $j < k$) dominate the variance in the above proof, especially when N is large.

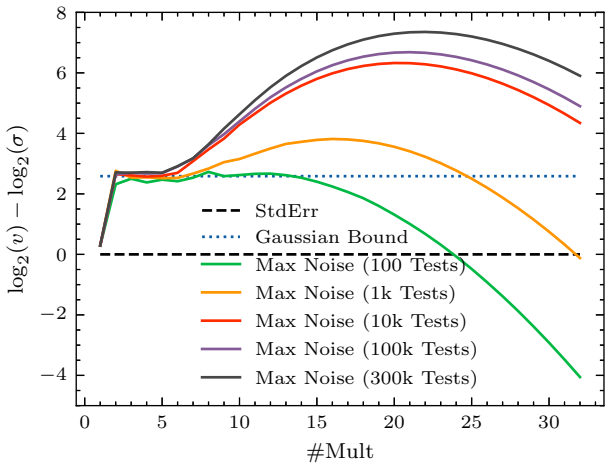
Proposition 5.2 (Power of Uniform). *For $f \in \mathbb{R}^N$ with coefficients uniformly sampled from $\mathcal{U}(-q, q)$ with variance $\sigma^2 = q^2/3$, when N is large, the variance of f^k could be approximated by*

$$\text{Var}(f^k|_\ell) = k!N^{k-1}\sigma^{2k}$$

We have also experimentally verified the distributions of the product of independent keys and the self-product of the same key in Figure 2. They show similar heavy-tailed behaviors as the products of Gaussian polynomials in Figure 1.



(a) Independent Key Polynomials.



(b) Same Key Polynomial.

Figure 2: The maximal noise observed in experiments compared with the standard error σ and Gaussian bound 6σ. The input polynomial s is sampled from χ_s for each test with $N = 2^{16}$. The left figure demonstrates the maximal noise of $v = s_1 s_2 \cdots s_k$, while the right figure demonstrates the maximal noise of $v = s^k$. Other description follows Figure 1.

6 CASE STUDIES

We give two case studies to emphasize how *ciphertext dependencies* significantly contribute to the variance under multiplications.

In particular, we focus on the following terms for ciphertext dependencies:

- The encryption randomness in each ciphertext $\mathtt{ct}$ like $u_{\mathtt{ct}} \leftarrow \chi_s, e_{\mathtt{ct}} \leftarrow \chi_e$.
- The high term $h_{\mathtt{ct}} = \mu_{\mathtt{ct}} s$ for ciphertext $\mathtt{ct}$ in the BFV scheme, with $\mu_{\mathtt{ct}}$ from uniform distribution in $[-1/2, 1/2]$.

We heuristically assume that these terms are independent if they are from different $\mathtt{ct}$.

We emphasize that previous sections have already established that the variance is not the only metric to consider when deriving the noise bound.

For convenience, we only track the terms with the highest degree because their variances have a factor N^{k-1} and become dominant.

6.1 BGV Multiplication, Dependent vs. Indepenent

The noise of a freshly encrypted ciphertext $\mathtt{ct}$ is expressed as

$$v_{\mathtt{ct}} = t(u_{\mathtt{ct}} \cdot e_{\mathsf{pk}} + e_{\mathtt{ct}} \cdot s)$$

Case 1. *The ratio between the variance of the noise of the k-th tensor product of one ciphertext and that of the product of k independent ciphertexts is*

$$\frac{\mathrm{Var}\left(v_{\mathtt{ct}^k}\right)}{\mathrm{Var}\left(v_{\prod_{j=1}^k \mathtt{ct}_j}\right)} = k!$$

with the condition that there is no modulus switching, and relinearization is conducted but introduces negligible noise.

Proof.

$$\begin{aligned} v_{\mathtt{ct}^k} &= t^k (u_{\mathtt{ct}} \cdot e_{\mathsf{pk}} + e_{\mathtt{ct}} \cdot s)^k \\ &= t^k \sum_i \binom{k}{i} (u_{\mathtt{ct}} \cdot e_{\mathsf{pk}})^{k-i} (e_{\mathtt{ct}} \cdot s)^i \end{aligned}$$

$$\mathrm{Var}(v_{\mathtt{ct}_k}) = t^{2k} N^{2k-1} \sum_i \binom{k}{i}^2 ((k-i)!)^2 (i!)^2 = t^{2k} N^{k-1} k!(k+1)!$$

$$v_{\prod_j \mathtt{ct}_j} = t^k \prod_j (u_{\mathtt{ct}_j} \cdot e_{\mathsf{pk}} + e_{\mathtt{ct}_j} \cdot s)$$

In it we have $\binom{k}{i}$ terms that have the form $e_{\mathsf{pk}}^i s^{k-i} \prod u_{\mathtt{ct}} e_{\mathtt{ct}}$, then its variance is

$$\begin{aligned} \mathrm{Var}(v_{\prod_j \mathtt{ct}_j}) &= t^{2k} N^{2k-1} \sum_i \binom{k}{i} (k-i)! i! \\ &= t^{2k} N^{2k-1} (k+1)! \end{aligned}$$

□

Our experiments in Figures 3a and 3b compare the variance derived above with the real noise. The real noise aligns with the predicted Gaussian bound initially, but then temporarily exceeds it, and finally drops. This is caused by the heavy-tailedness, i.e., the maximal noise significantly changes with more tests, and (its left half) is in accordance with Figures 1 and 2. If they conform to a Gaussian distribution, regardless of the correctness of the variance calculation, the maximal noise observed should not differ by more than 10 bits (meaning $D > 2^{10}$) across different numbers of tests.

6.2 BFV Multiplication, Dependent vs. Independent

The noise of a freshly encrypted ciphertext $\mathtt{ct}$ is be expressed as

$$v_{\mathtt{ct}} = u_{\mathtt{ct}} \cdot e_{\mathsf{pk}} + e_{\mathtt{ct}} \cdot s$$

A BFV ciphertext ct takes multiplications in $\mathcal{R} = \mathbb{Z}[X]/\Phi_m(X)$ instead of in $\mathcal{R}_Q$, so at this time we have

$$\langle \mathsf{ct}, \mathsf{sk} \rangle = \left\lfloor \frac{Q}{t} m \right\rceil + v_{\mathsf{ct}} + h_{\mathsf{ct}} Q$$

Case 2. *The ratio between the variance of the noise of the k-th tensor product of one ciphertext and that of the product of k independent ciphertexts is*

$$\frac{\mathrm{Var}\left(v_{\mathsf{ct}^k}\right)}{\mathrm{Var}\left(v_{\prod_{j=1}^{k} \mathsf{ct}_j}\right)} = 2(k-1)!$$

with the condition that after each tensor product the relinearization process is conducted but introduces negligible noise. Note that the multiplications happen in a sequential manner instead of a binary tree style.

Proof.

$$v_{\mathsf{ct}^k} = t^{k-1} \cdot 2(u_{\mathsf{ct}} \cdot e_{\mathsf{pk}} + e_{\mathsf{ct}} \cdot s)\mu_{\mathsf{ct}}^{k-1} s^{k-1}$$

$$v_{\prod_j \mathsf{ct}_j} = t^{k-1} \cdot \sum_{i=1}^{2} (u_{\mathsf{ct}_i} \cdot e_{\mathsf{pk}} + e_{\mathsf{ct}_i} \cdot s) s^{k-1} \prod_{j \neq i}^{k} \mu_{\mathsf{ct}_j}$$

□

The experiments in Figures 3c and 3d show similar behaviors to the BGV case.

7 CRITIQUE OF EXISTING AVERAGE-CASE ANALYSIS

7.1 Variance-Based

Theorems 4.1 and 4.3 show that the Gaussian Heuristic (and the CLT) is not applicable for deep multiplications, as the tail is too heavy to be approximated as a Gaussian distribution. In the experiments we do observe underestimation caused by such approximation.

Nevertheless, they do suggest that the noise after a small number of multiplications can still be well approximated by a Gaussian distribution. This aligns with [CNP23] which considers the modulus switching noise and the multiplication noise thereafter in BGV, whose dominant terms are τs and $\tau\tau' s^2$, respectively, and they are low degree polynomials.

The case studies in Section 6 demonstrate that even for calculating the variance alone, it is insufficient to only track the common dependencies alongside the variance. All the ciphertext dependencies should also be tracked.

We comment on [BMCM23, BMMU25] about the discrepancy between our $k!N^{k-1}$ variance growth and their experimental results. The authors there experimentally estimate how the variance grows for dependent polynomials like the secret key. Their initial growth aligns with the $k!$ growth. However, after around $k = 20$ its growth becomes constant. We argue that their approach of estimating variance growth is incorrect, as the distribution of the self-product of the secret key is heavy-tailed and it requires extremely long time to converge to the correct theoretical variance.

7.2 Experiment-Based

[HPS19, KPZ21] use the experimental *average-case expansion factor* of $C\sqrt{N}$ with $C = 2$. This formula aligns with the N^{k-1} growth of the variance for independent cases, and works well as in Figures 4a and 4c. However, it could not handle the $k!N^{k-1}$ growth of the variance for dependent cases.

We observe underestimation using this expansion factor for the self-product of the secret key s^k in Figure 4d. Note that s^k exists even for multiplications of *independent* ciphertexts in BFV, as demonstrated in the proof of Case 2.

We do not observe underestimation for the self-product of the Gaussian error e^k in Figure 4b. We argue that this is caused by the choice of constant and is highly sensitive to such choices. For Gaussian polynomials, they set $B_{\mathrm{err}} = D$ for a typical $D = 6$ and the bound after k multiplications is $D^k(C\sqrt{N})^{k-1}$. Clearly, the choice of D and C affects how the bound behaves over the variance growth of $k!N^{k-1}$.

We explore their combined effect using actual ciphertext noises. We observe underestimation if we directly apply Formula 10 of [KPZ21] for the circuit of multiplying the same BFV ciphertext in Figure 5b. OpenFHE however employs another conservative formula meant for the worst-case binary-tree circuit of the same depth for parameter generation, and it does not underestimate.

Independently, we find underestimation caused by the average-case expansion factor when the noise is a product of a uniform polynomial and the secret key, like the modulus switching rounding error in BGV and the similar noise in `EXTENDED` encryption mode in BFV in OpenFHE[3].

We conclude that the *average-case expansion factor* is only applicable in limited cases, and it alone is not the correctness foundation of the noise bound. In OpenFHE we do not observe decryption failure caused by such underestimation, but we have to point out that there are various loose inequalities ("cushion") from [HPS19, KPZ21] involved in the parameter generation process, and they *together* form the correctness foundation of the practical implementation.

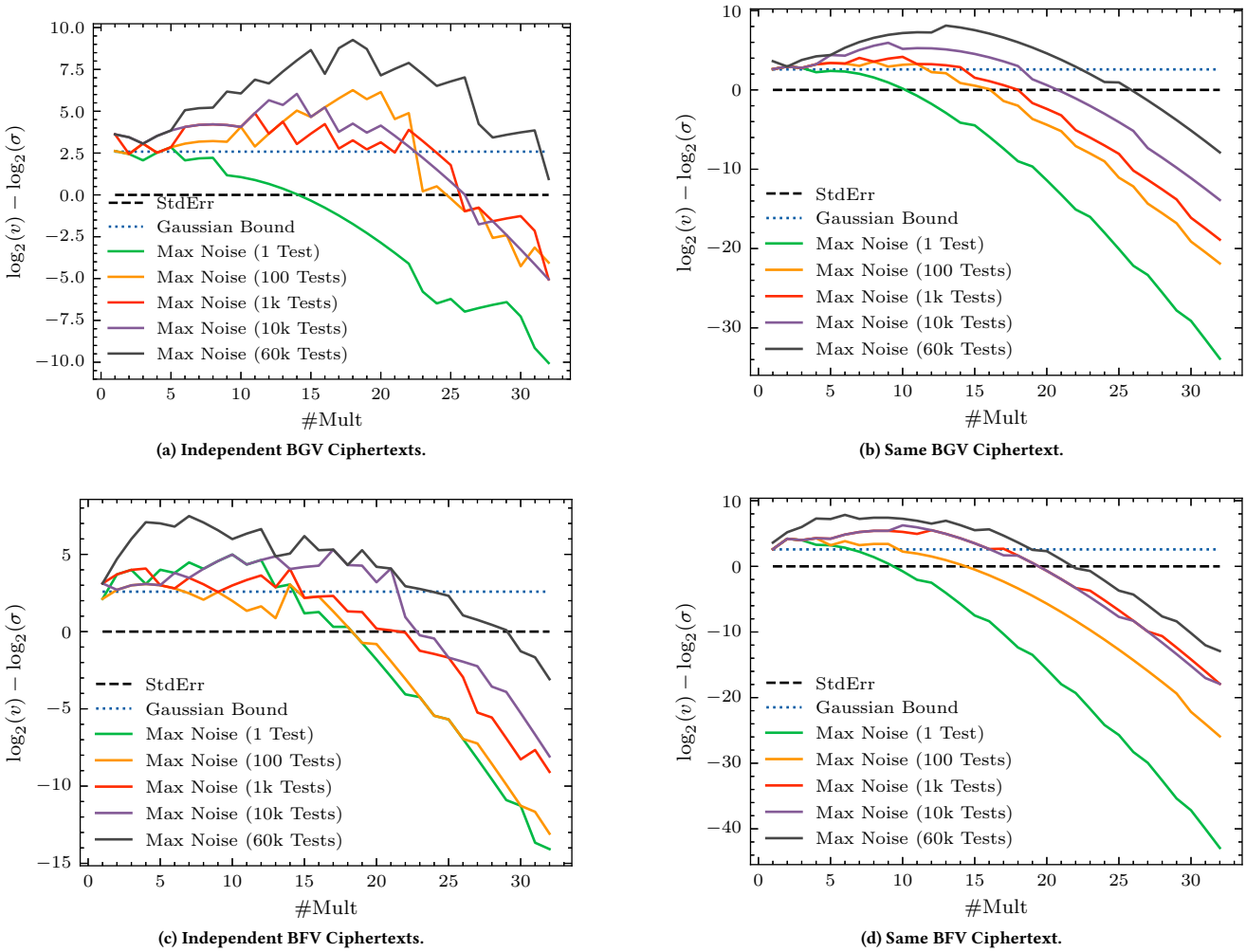
8 DISCUSSION

We discuss several potential future directions implied by our work.

Section 6 shows that even for calculating the variance alone, all the ciphertext dependencies should be tracked, which is not an easy task to do in software libraries. The current dynamic noise estimation approach employed in HElib [HS20] is already non-trivial, which uses worst-case bounds and naturally handles dependencies. Instead, recent advancements in homomorphic encryption compilers (e.g., [Con23]) could take the task of analyzing the ciphertext circuit and tracking the dependencies exactly if future average-case analysis relies on such information.

While our work is able to exactly calculate the variance and (some) kurtosis for the noise, we only demonstrate the heavy-tailedness of it. For practical bounding purposes, we still need a characterization of the quantile function. Extending past work like [MP19] might be the solution.

[3] Our code is in https://github.com/tsinghua-ideal/critique-code.



(a) Independent BGV Ciphertexts.

(b) Same BGV Ciphertext.

(c) Independent BFV Ciphertexts.

(d) Same BFV Ciphertext.

Figure 3: The maximal noise observed in experiments compared with the standard error σ and Gaussian bound 6σ. The experiments, based on Case 1 and Case 2, are carried out in OpenFHE with ring dimension $N = 2^{15}$, plaintext modulus $t = 65537$, and multiplicative depth 32, with each modulus of size 60. We set the security `HEStd_NotSet` as we are only experimenting with the noise. For BGV, we use `FIXEDMANUAL` to prevent automatic modulus switching. We use `HYBRID` key switching to make the key-switching noise after each multiplication negligible. The encryption technique is `STANDARD`. For BFV, the multiplication technique defaults to `HPSPOVERQLEVELED`. We comment that for k number of ciphertext multiplications, the degree of the noise expression is $2k$, so only the left halves of these figures correspond to Figures 1 and 2. Other description follows Figure 1.

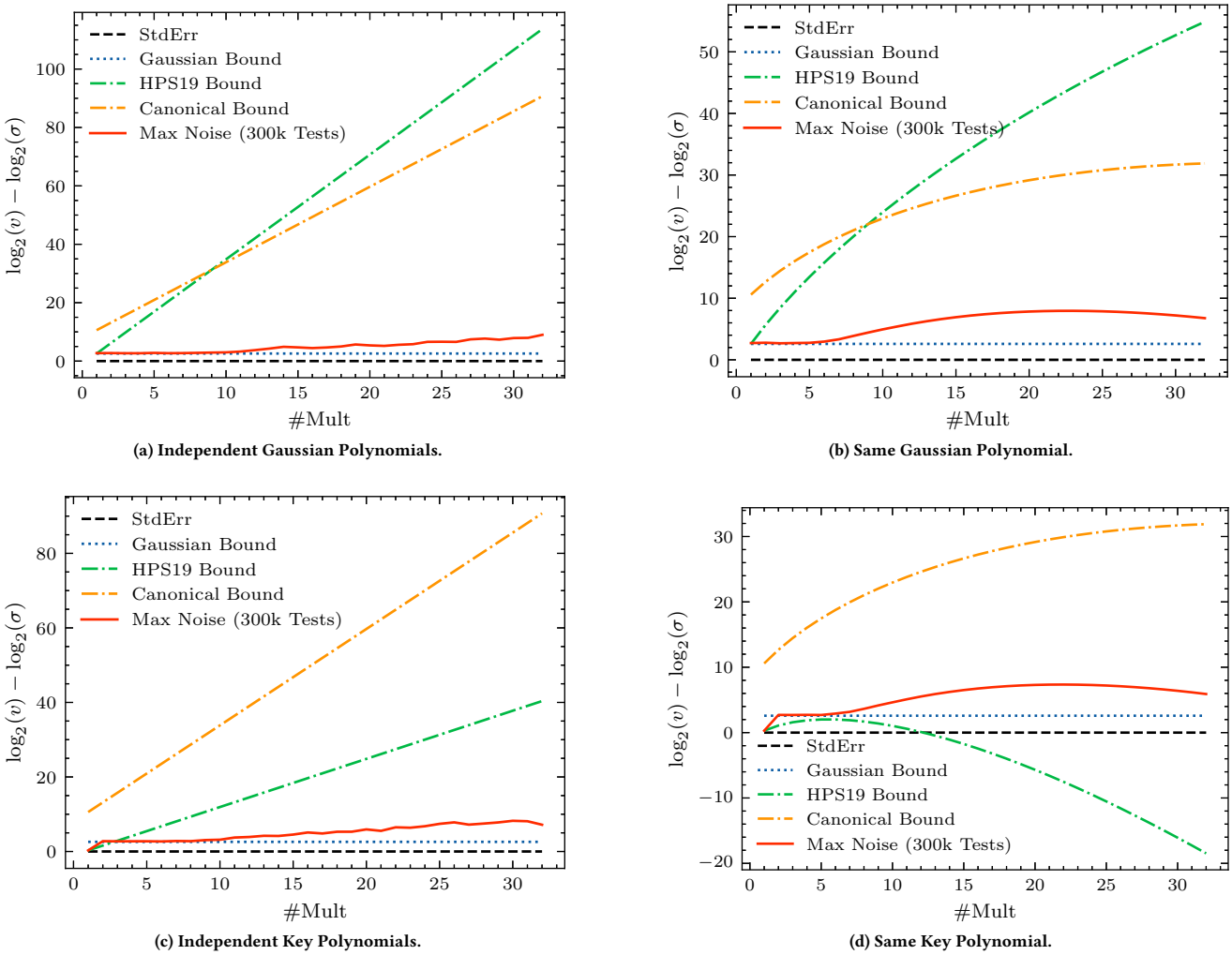
An exact characterization of the quantile function and failing probability could contribute to the Homomorphic Encryption Standard [ACC+19] and security guidelines [BCC+24], whereas in the past only security parameters from the lattice world were considered for RLWE schemes. We point out that the failing probability associated with the circuit should also be considered, which the application-aware security model [ABMP24] also emphasizes.

Recent attacks like [GNSJ24, CCP+24, CSBB24] exploited the dependencies in additions, while we show that the dependencies in multiplications are also critical and will even cause the deformation of the noise distribution. It may be possible to also exploit the dependencies in multiplications to launch attacks, especially with the new noise distributions.

In summary, we call for a more rigorous analysis of the noise and security in RLWE-based homomorphic encryption schemes, where the parameterization of failing probability should be made explicit, not only for the initial noise, but also for the whole circuit.

ACKNOWLEDGEMENT

We thank Yilei Chen, Yuriy Polyakov, Chiara Marcolla, Beatrice Biasioli, Matilda Urani and Nadir Murru for feedback on the early drafts of this work. We thank Jeremy Kun for careful proofreading.



(a) Independent Gaussian Polynomials.

(b) Same Gaussian Polynomial.

(c) Independent Key Polynomials.

(d) Same Key Polynomial.

Figure 4: The maximal noise observed in experiments compared with the standard error σ, Gaussian bound 6σ, bound from [HPS19], and canonical embedding based bound, all with $D = 6$ when treating Gaussian polynomials. Other description follows Figure 1 and Figure 2.

A PROOF

A.1 Alternative view of the proof in Theorem 4.3

In the proof of Theorem 4.3, the counting problem could be equivalently viewed as an assigning problem with $2k$ indeterminate α_i, β_j with the index set $\mathcal{I}_\ell$ condition transformed to the following constraints in the row direction:

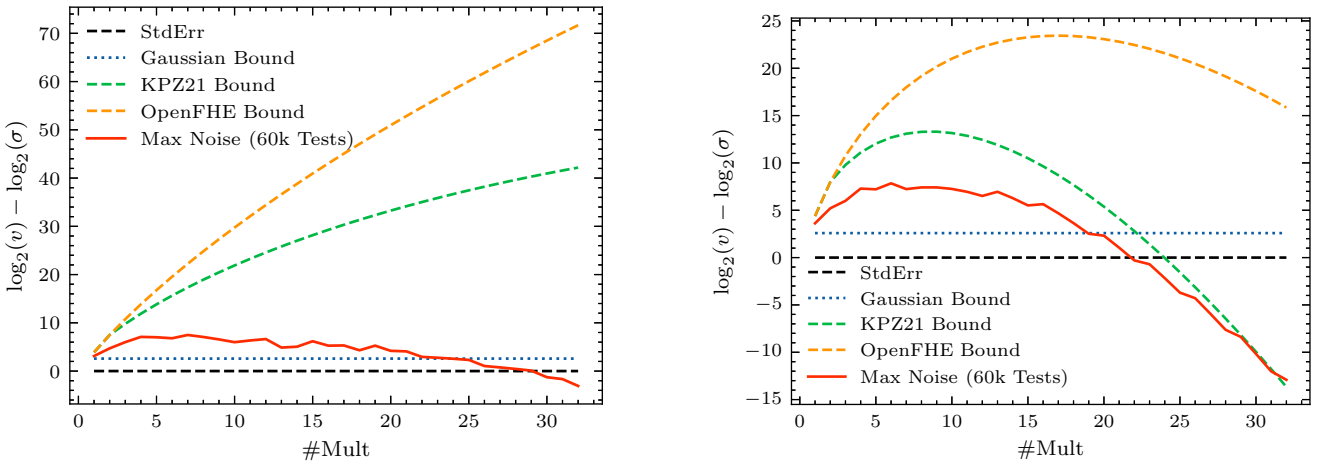
$$\begin{aligned} \alpha_1 &+ \alpha_2 + \cdots + \alpha_k \equiv \ell \pmod N \\ \beta_1 &+ \beta_2 + \cdots + \beta_k \equiv \ell \pmod N \end{aligned}$$

A partition $p(\alpha, \beta)$ of k disjoint pairings adds another k constraints. The type of partitions we need to consider is partitions with only (α_i, β_j), so constraints are added on the column direction by $\alpha_i = \beta_j$.

The row constraints reduce the degrees of freedom by 2, and the column constraints reduce them by $k - 1$ as the last constraint is automatically satisfied, so the final degrees of freedom are $k - 1$, hence N^{k-1} possible assignments for one p. As there are $k!$ possible p, the total number of assignments is $k!N^{k-1}$.

As argued in the original proof, a partition $p(\alpha, \beta)$ with pairing $\alpha_i = \alpha_j$ (i.e., additional row constraint) will cancel out as one assignment α has a corresponding sign-flipped assignment α' by assigning $\alpha'_i = \alpha'_j = \alpha_i + \frac{N}{2} \pmod N$ and leaving other points fixed. The same applies for $\beta_i = \beta_j$.

Generally, to calculate other quantities (i.e., expectation, covariance, fourth moment) we can add more row and column constraints and find all possible assignments.



(a) Independent BFV Ciphertexts.

(b) Same BFV Ciphertext.

Figure 5: The maximal noise observed in experiments compared with the standard error σ, Gaussian bound 6σ, bound from Formula 10 of [KPZ21], and bound in OpenFHE, all with $D = 6$ when treating Gaussian polynomials. Other description follows Figure 3.

A.2 Proof of Theorem 4.1

Expectation

$$\alpha_1 + \alpha_2 + \cdots + \alpha_k \equiv \ell \pmod N$$

The only possible pairing is between α_i and α_j. However, the i-th column corresponds to f_i and j-th column corresponds to f_j and they are independent, so the indicator function will be 0, which means $\mathbb{E}[F|_\ell] = 0$.

Covariance

$$\begin{aligned} \alpha_1 + \alpha_2 + \cdots + \alpha_k &\equiv \ell \pmod N \\ \beta_1 + \beta_2 + \cdots + \beta_k &\equiv \ell' \pmod N \end{aligned}$$

The pairing (α_i, β_j) is the only possible pairing that will not cancel out. However, as $\ell \neq \ell'$, such pairing does not exist as it would force $\ell = \ell'$, so $\mathrm{Cov}(F|_\ell, F|_{\ell'}) = 0$.

Fourth Moment

$$\begin{aligned} \alpha_1 + \alpha_2 + \cdots + \alpha_k &\equiv \ell \pmod N \\ \beta_1 + \beta_2 + \cdots + \beta_k &\equiv \ell \pmod N \\ \gamma_1 + \gamma_2 + \cdots + \gamma_k &\equiv \ell \pmod N \\ \delta_1 + \delta_2 + \cdots + \delta_k &\equiv \ell \pmod N \end{aligned}$$

As the i-th column corresponds to f_i, we only have pairing inside one column. For each column, we have 3 possible cases:

- Case 1: $\alpha_i = \beta_i$ and $\gamma_i = \delta_i$.
- Case 2: $\alpha_i = \gamma_i$ and $\beta_i = \delta_i$.
- Case 3: $\alpha_i = \delta_i$ and $\beta_i = \gamma_i$.

Note that without column constraints, we have $4k - 4$ degrees of freedom, then each column constraint will reduce them by 2. Proceeding from the first column to the last, notice that the last column does not change the degrees of freedom, as either former assignments automatically make the last column satisfying one of the column constraints, or the row constraints are violated. So the maximal number of freedom is $2k - 2$.

The only time we achieve maximal degrees of freedom is when all columns are all Case 1, or Case 2, or Case 3. For all columns being Case 1, by checking the sign we have $\xi(\alpha)^2\xi(\gamma)^2 = 1$, so there is no canceling out. All three cases contribute $3N^{2k-2}$ in total.

Then for some columns being Case 1 and some columns being Case 2, we notice that the degrees of freedom are $2k - 3$ as $(a, a, b, b) + (c, d, c, d) \equiv (\ell, \ell, \ell, \ell)$ will automatically make $c \equiv d$ and $a \equiv b$, so crossing the boundary between Case 1 and Case 2 will lose one degree of freedom. There are $2^k - 2$ possible ways to pick columns being either Case 1 or Case 2 where both cases are present. Then by symmetry, argument follows for Case 1+3 and Case 2+3. Checking the sign, notice we can express $b = a + \eta_1 N$ and $d = c + \eta_2 N$, then the row summation result would be $(\ell, \ell + \eta_1 N, \ell + \eta_2 N, \ell + (\eta_1 + \eta_2)N)$, then by parity of η, the resulting sign will always be 1. They contribute $3(2^k - 2)N^{2k-3}$.

Then for some columns being Case 1, some Case 2 and some Case 3, by $(a, a, b, b) + (c, d, c, d) + (e, f, f, e) \equiv (\ell, \ell, \ell, \ell)$, we get $2a \equiv 2b$, so $b = a + \eta_1 \frac{N}{2}$. Here note N is even. Similarly define η_2 and η_3, the row summation result would be

$$\left(\ell, \ell + (\eta_2 + \eta_3)\frac{N}{2}, \ell + (\eta_1 + \eta_3)\frac{N}{2}, \ell + (\eta_1 + \eta_2)\frac{N}{2}\right)$$

We have $\eta_i + \eta_j$ must be even. So η_1, η_2, η_3 share the same parity. Now if $\xi(\alpha)\xi(\beta)\xi(\gamma)\xi(\delta)$ is 1, we can assign $\eta_i' = \eta_i + 1$ (parity flipped at the same time), then $\xi(\alpha)\xi(\beta' = \beta + N)\xi(\gamma' = \gamma + N)\xi(\delta' = \delta + N)$ will be -1, so they will cancel out.

A.3 Proof for Theorem 4.3

Expectation. Unlike the argument in Section A.2, this time columns are not independent. However, we can still use the $\alpha_i = \alpha_j$ and $\alpha_i' = \alpha_j' = \alpha_i + \frac{N}{2} \pmod N$ to flip the sign to cancel out, so $\mathbb{E}[F|_\ell] = 0$.

Covariance. The same argument as in Section A.2.

Fourth Moment. Now we have pairing between columns like $\alpha_i = \beta_j, \gamma_u = \delta_v$. We count such pairing as one "column". The argument for degrees of freedom is similar, but this time we get $k!k!$ ways to form columns for all columns being Case 1 (α_0 has k choices and γ_0 has k choices), and $(2k)! - 2(k!)^2$ ways to form columns for some columns being Case 1 and some Case 2 (α_0 has $2k$ choices among row β and γ, so $(2k)(2k-1)\cdots(k+1)$ choices for α row, then δ_0 has k choices, so $(2k)!$ in total, minus the cases all columns being Case 1 or Case 2).

REFERENCES

[ABMP24] Andreea Alexandru, Ahmad Al Badawi, Daniele Micciancio, and Yuriy Polyakov. Application-aware approximate homomorphic encryption: Configuring FHE for practical use. Cryptology ePrint Archive, Report 2024/203, 2024.

[ACC+19] Martin Albrecht, Melissa Chase, Hao Chen, Jintai Ding, Shafi Goldwasser, Sergey Gorbunov, Shai Halevi, Jeffrey Hoffstein, Kim Laine, Kristin Lauter, Satya Lokam, Daniele Micciancio, Dustin Moody, Travis Morrison, Amit Sahai, and Vinod Vaikuntanathan. Homomorphic encryption standard. Cryptology ePrint Archive, Report 2019/939, 2019.

[BAB+22] Ahmad Al Badawi, Andreea Alexandru, Jack Bates, Flavio Bergamaschi, David Bruce Cousins, Saroja Erabelli, Nicholas Genise, Shai Halevi, Hamish Hunt, Andrey Kim, Yongwoo Lee, Zeyu Liu, Daniele Micciancio, Carlo Pascoe, Yuriy Polyakov, Ian Quah, Saraswathy R. V., Kurt Rohloff, Jonathan Saylor, Dmitriy Suponitsky, Matthew Triplett, Vinod Vaikuntanathan, and Vincent Zucca. OpenFHE: Open-source fully homomorphic encryption library. Cryptology ePrint Archive, Report 2022/915, 2022.

[BCC+24] Jean-Philippe Bossuat, Rosario Cammarota, Ilaria Chillotti, Benjamin R. Curtis, Wei Dai, Huijing Gong, Erin Hales, Duhyeong Kim, Bryan Kumara, Changmin Lee, Xianhui Lu, Carsten Maple, Alberto Pedrouzo-Ulloa, Rachel Player, Yuriy Polyakov, Luis Antonio Ruiz Lopez, Yongsoo Song, and Donggeon Yhee. Security guidelines for implementing homomorphic encryption. *CiC*, 1(4):26, 2024.

[BCM+24] Jean-Philippe Bossuat, Anamaria Costache, Christian Mouchet, Lea Nürnberger, and Juan Ramón Troncoso-Pastoriza. Practical q-IND-CPA-D-secure approximate homomorphic encryption. Cryptology ePrint Archive, Report 2024/853, 2024.

[BGV12] Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. (Leveled) fully homomorphic encryption without bootstrapping. In Shafi Goldwasser, editor, *ITCS 2012*, pages 309–325. ACM, January 2012.

[BJSW25] Olivier Bernard, Marc Joye, Nigel P. Smart, and Michael Walter. Drifting towards better error probabilities in fully homomorphic encryption schemes. LNCS, pages 181–211. Springer, Cham, June 2025.

[BMCM23] Beatrice Biasioli, Chiara Marcolla, Marco Calderini, and Johannes Mono. Improving and automating BFV parameters selection: An average-case approach. Cryptology ePrint Archive, Report 2023/600, 2023.

[BMMU25] Beatrice Biasioli, Chiara Marcolla, Nadir Murru, and Matilda Urani. Accurate bgv parameters selection: Accounting for secret and public key dependencies in average-case analysis, 2025.

[Bra12] Zvika Brakerski. Fully homomorphic encryption without modulus switching from classical GapSVP. In Reihaneh Safavi-Naini and Ran Canetti, editors, *CRYPTO 2012*, volume 7417 of *LNCS*, pages 868–886. Springer, Berlin, Heidelberg, August 2012.

[CCH+24] Anamaria Costache, Benjamin R. Curtis, Erin Hales, Sean Murphy, Tabitha Ogilvie, and Rachel Player. On the precision loss in approximate homomorphic encryption. In Claude Carlet, Kalikinkar Mandal, and Vincent Rijmen, editors, *SAC 2023*, volume 14201 of *LNCS*, pages 325–345. Springer, Cham, August 2024.

[CCP+24] Jung Hee Cheon, Hyeongmin Choe, Alain Passelègue, Damien Stehlé, and Elias Suvanto. Attacks against the IND-CPA$^{\mathrm{D}}$ security of exact FHE schemes. In Bo Luo, Xiaojing Liao, Jun Xu, Engin Kirda, and David Lie, editors, *ACM CCS 2024*, pages 2505–2519. ACM Press, October 2024.

[CKKS17] Jung Hee Cheon, Andrey Kim, Miran Kim, and Yong Soo Song. Homomorphic encryption for arithmetic of approximate numbers. In Tsuyoshi Takagi and Thomas Peyrin, editors, *ASIACRYPT 2017, Part I*, volume 10624 of *LNCS*, pages 409–437. Springer, Cham, December 2017.

[CLP20] Anamaria Costache, Kim Laine, and Rachel Player. Evaluating the effectiveness of heuristic worst-case noise analysis in FHE. In Liqun Chen, Ninghui Li, Kaitai Liang, and Steve A. Schneider, editors, *ESORICS 2020, Part II*, volume 12309 of *LNCS*, pages 546–565. Springer, Cham, September 2020.

[CNP23] Anamaria Costache, Lea Nürnberger, and Rachel Player. Optimisations and tradeoffs for HElib. In Mike Rosulek, editor, *CT-RSA 2023*, volume 13871 of *LNCS*, pages 29–53. Springer, Cham, April 2023.

[Con23] HEIR Contributors. HEIR: Homomorphic Encryption Intermediate Representation, 2023. https://github.com/google/heir.

[CS16] Ana Costache and Nigel P. Smart. Which ring based somewhat homomorphic encryption scheme is best? In Kazue Sako, editor, *CT-RSA 2016*, volume 9610 of *LNCS*, pages 325–340. Springer, Cham, February / March 2016.

[CSBB24] Marina Checri, Renaud Sirdey, Aymen Boudguiga, and Jean-Paul Bultel. On the practical CPA^{D} security of "exact" and threshold FHE schemes and libraries. In Leonid Reyzin and Douglas Stebila, editors, *CRYPTO 2024, Part III*, volume 14922 of *LNCS*, pages 3–33. Springer, Cham, August 2024.

[DPSZ12] Ivan Damgård, Valerio Pastro, Nigel P. Smart, and Sarah Zakarias. Multiparty computation from somewhat homomorphic encryption. In Reihaneh Safavi-Naini and Ran Canetti, editors, *CRYPTO 2012*, volume 7417 of *LNCS*, pages 643–662. Springer, Berlin, Heidelberg, August 2012.

[FV12] Junfeng Fan and Frederik Vercauteren. Somewhat practical fully homomorphic encryption. Cryptology ePrint Archive, Report 2012/144, 2012.

[Gau14] Robert E. Gaunt. Variance-gamma approximation via stein's method, 2014.

[GHS12] Craig Gentry, Shai Halevi, and Nigel P. Smart. Homomorphic evaluation of the AES circuit. In Reihaneh Safavi-Naini and Ran Canetti, editors, *CRYPTO 2012*, volume 7417 of *LNCS*, pages 850–867. Springer, Berlin, Heidelberg, August 2012.

[GNSJ24] Qian Guo, Denis Nabokov, Elias Suvanto, and Thomas Johansson. Key recovery attacks on approximate homomorphic encryption with non-worst-case noise flooding countermeasures. In Davide Balzarotti and Wenyuan Xu, editors, *USENIX Security 2024*. USENIX Association, August 2024.

[HPS19] Shai Halevi, Yuriy Polyakov, and Victor Shoup. An improved RNS variant of the BFV homomorphic encryption scheme. In Mitsuru Matsui, editor, *CT-RSA 2019*, volume 11405 of *LNCS*, pages 83–105. Springer, Cham, March 2019.

[HS20] Shai Halevi and Victor Shoup. Design and implementation of HElib: a homomorphic encryption library. Cryptology ePrint Archive, Report 2020/1481, 2020.

[Iss18] L. Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12(1/2):134–139, 1918.

[KPZ21] Andrey Kim, Yuriy Polyakov, and Vincent Zucca. Revisiting homomorphic encryption schemes for finite fields. In Mehdi Tibouchi and Huaxiong Wang, editors, *ASIACRYPT 2021, Part III*, volume 13092 of *LNCS*, pages 608–639. Springer, Cham, December 2021.

[LPR10] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In Henri Gilbert, editor, *EUROCRYPT 2010*, volume 6110 of *LNCS*, pages 1–23. Springer, Berlin, Heidelberg, May / June 2010.

[MML+23] Johannes Mono, Chiara Marcolla, Georg Land, Tim Güneysu, and Najwa Aaraj. Finding and evaluating parameters for BGV. In Nadia El Mrabet, Luca De Feo, and Sylvain Duquesne, editors, *AFRICACRYPT 23*, volume 14064 of *LNCS*, pages 370–394. Springer, Cham, July 2023.

[MP19] Sean Murphy and Rachel Player. Discretisation and product distributions in ring-LWE. Cryptology ePrint Archive, Report 2019/596, 2019.

[MP24] Sean Murphy and Rachel Player. A central limit approach for ring-LWE noise analysis. *CiC*, 1(2):7, 2024.

[Reg05] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In Harold N. Gabow and Ronald Fagin, editors, *37th ACM STOC*, pages 84–93. ACM Press, May 2005.

[Wes14] Peter H Westfall. Kurtosis as peakedness, 1905–2014. rip. *The American Statistician*, 68(3):191–195, 2014.