Accurate BGV Parameters Selection: Accounting for Secret and Public Key Dependencies in Average-Case Analysis

Beatrice Biasioli[1], Chiara Marcolla[2], Nadir Murru[3], and Matilda Urani[4]

[1] IBM Research Europe - Zurich, Switzerland & University of Potsdam, Germany **
[2] Technology Innovation Institute, Abu Dhabi, United Arab Emirates
[3] Università degli studi di Trento, Trento, Italy
[4] Politecnico di Torino, Torino, Italy

Abstract. The Brakerski-Gentry-Vaikuntanathan (BGV) scheme is one of the most significant fully homomorphic encryption (FHE) schemes. Such schemes deal with a quantity, called *noise*, which increases each time a homomorphic operation is performed. Specifically, in order for the scheme to work properly, it is essential that the noise remains below a certain threshold throughout the process. For BGV, this threshold strictly depends on the ciphertext modulus, which is one of the initial parameters whose selection heavily affects both the efficiency and security of the scheme.
For an optimal parameter choice, it is crucial to accurately estimate the noise growth, particularly that arising from multiplication, which is the most complex operation. In this work, we propose a novel *average-case* approach that precisely models noise evolution and guides the selection of initial parameters, improving efficiency while ensuring security. The key innovation of our method lies in accounting for the dependencies among ciphertext errors generated with the same key, and in providing general guidelines for accurate parameter selection that are library-independent. Our parameter selection methodology leads to a significant improvement over the parameter choices currently adopted in major FHE libraries.

Keywords: Average-case; BGV; Fully Homomorphic Encryption; OpenFHE; Parameters selection

1 Introduction

The first Fully Homomorphic Encryption (FHE) scheme was introduced in 2009 by Gentry [25]. Since then, several FHE constructions have been proposed, such as BGV [7], BFV [6,22], FHEW [21], TFHE [13,14], and CKKS [12,11].

** Part of this work was performed while at Technology Innovation Institute, Abu Dhabi

The homomorphic encryption schemes currently in use base their security on the presumed intractability of the Learning with Errors (LWE) problem [35], and its ring variant (RLWE) [32]. Schemes based on the (R)LWE problem face a critical challenge related to the growth of noise during homomorphic operations, which must be carefully controlled to ensure the correct functioning of the encryption scheme. Specifically, the noise must be kept below a certain threshold, which, in the case of BGV, is directly related to the ciphertext modulus parameter q.

As homomorphic operations are performed, the noise increases, and therefore, to maintain the integrity of the scheme, the parameter q must be chosen sufficiently large. However, although increasing q allows for a greater number of operations, it simultaneously compromises both the security and efficiency of the scheme. Therefore, selecting an appropriate value for q and, in general, determining an optimal set of parameters, is critical. This process requires a balance between security and efficiency while ensuring the correctness of the scheme.
One of the key factors in achieving this balance and determining suitable parameters is providing accurate estimates of the error and its growth during the homomorphic operations in the circuit.

This issue is central to research in FHE, and over the years, various approaches have been proposed to address it. As for example, employing the Euclidean norm [7], the infinity norm [22,29], and the canonical norm, commonly used in the so-called *worst-case* analysis [15,17,26,28,33]. The prevailing trend in the current literature adopts the *average-case analysis*, which involves treating the noise coefficients as random variables distributed according to a Gaussian distribution and studying their expected value and variance.

Interest in this method, initially applied in the TFHE scheme [13], and subsequently in the CKKS [16,5], BGV [18,34] and BFV [4] schemes, grew due to a recognized discrepancy between the estimates based on worst-case technique and experimental data, as highlighted in [17]. The introduction of the average-case approach, as seen in [4,18], offers a potential resolution to these disparities, indeed, with this method it is possible to compute a tight *probabilistic* upper bound.

Controversies. The application of average-case analysis in FHE schemes is a complex and still evolving topic, which has recently attracted some controversy. In particular, most works adopting an average-case perspective rely on two strong assumptions: (1) the errors of ciphertexts involved in a homomorphic circuit are mutually independent; (2) the coefficients of these errors are Gaussian distributed. These assumptions are convenient for two main reasons. The former allows one to estimate the variance of the error coefficients after a ciphertext multiplication as the product of the variances of the initial errors, thereby significantly simplifying the analysis. The latter enables the use of Gaussian distribution properties to derive probabilistic bounds on the error norm based solely on its variance, thus providing a straightforward method for parameter

selection. However, both assumptions have recently been challenged. In [4], the authors show that, for the BFV scheme, the independence assumption does not hold: errors across different ciphertexts exhibit dependencies, and they propose a refined method to estimate the resulting noise accounting for such correlations. In [24], it is argued that the Gaussianity assumption is not always justified; in some cases, the error coefficients follow heavy-tailed distributions. In this case, a more precise characterization of the error distribution should be given in order to derive meaningful bounds.

Moreover, another limitation of many current works adopting an average-case approach is that they tend to *underestimate* the noise, as pointed out in [3,16,20,34]. This may lead to two potential issues: first, the ciphertext is not correctly decrypted with non-negligible probability due to excessive noise and, second, the scheme is exposed to security vulnerabilities, as shown in recent papers [9,10]. To mitigate this, some recent works have proposed refined average-case analyses that are tightly coupled with specific implementations. As for example, in [18] the authors develop a noise estimation method tailored to the specific implementation of BGV in HElib [27].

Our Contribution. In this work, we revisit these heuristics and investigate when, and to what extent, they are applicable in the context of the BGV scheme. In particular, we show that the independence of errors cannot be assumed in BGV, as they share common powers of the public and secret keys. We further observe that this assumption is the primary source of noise underestimation in current works. So, in this paper we propose the first average-case approach that explicitly accounts for such dependencies, enabling a highly accurate estimation of the error variance that closely matches experimental results, while never underestimating it. Moreover, we propose an approach that is *independent of the specific library* used. Our method yields a general estimation framework applicable to arbitrary circuits, by simply tracking the powers of shared terms, which depend directly on the multiplicative depth. Specifically, in this work we were able to prove the following tight upper bound on the noise after multiplication.

Theorem 2 [Informal version]. *Let V, V' describe the variances of the coefficients of the errors associated with two ciphertexts defined in $\mathbb{Z}_q[x]/\langle x^n+1\rangle$. Then, the variance of the coefficient of the error resulting from the multiplication of these two ciphertexts is*

$$V_{\mathsf{mul}} \leq nV\,V'F_sF_e\,,$$

with F_s and F_e explicitly given quantities depending on the number of multiplications previously performed in the circuit.

Our second main contribution regards the analysis of the distribution of the noise coefficients. We show that although Gaussianity of the error coefficients is not theoretically guaranteed, as pointed out in [24], it remains a reasonable approximation in practical deployments of BGV due to the use of modulus switching.

Indeed, when switching from $q = q'p$ to q', the variance of the noise coefficients associated with the new ciphertext ν_{ms} is given by

$$V(\nu_{\mathsf{ms}}|_i) = \tfrac{1}{p^2} V(\nu|_i) + V_{\mathsf{ms}},$$

where ν is the noise associated with the initial ciphertext, V_s is the variance of the secret key coefficients and $V_{\mathsf{ms}} = \frac{t^2}{12}(1 + nV_s)$.

The variance introduced by modulus switching is independent of the input noise and of the circuit depth. Thus, in this work, we prove that when p is chosen sufficiently large, then: (1) the error coefficients follow a discrete Gaussian distribution after modulus switching and it is reasonable to assume that this holds also after the multiplication of two ciphertexts that have undergone modulus switching; (2) their variance becomes independent of the number of homomorphic multiplications performed.

Under these conditions, the error no longer depends on the multiplication depth of the circuit, as the impact of the underlying dependencies is effectively controlled. As a result, variance can be easily estimated without precisely tracking the growing number of shared terms, leading to significantly simpler and more scalable analysis. This also enhances the practical relevance of our approach, as it facilitates the application of the theoretical results in real-world settings, making them easier to adopt in practice.

Finally, by combining these two results, we propose a novel parameter selection methodology that leads to a significant improvement over the parameter choices currently adopted in major FHE libraries. For example, for circuits of multiplication depth 3, our approach allows selecting ciphertext moduli that are 30 bits smaller than those used in OpenFHE [1], and up to 42 bits smaller for depth 6. Similarly, in HElib [36], the ratio between two moduli of adjacent levels is around 54 bits, whereas our approach reduces it to less than 35 bits.

The paper is structured as follows. Section 2 introduces essential definitions and fundamental properties that are necessary for understanding both our contribution and the functioning of the scheme. Section 3 provides a concise overview of the main features and structure of the BGV scheme. Section 4 develops our average-case analysis, addressing the failure of the independence assumption by modeling dependencies among error terms. We derive a general framework for estimating the variance of the error coefficients, whose core is given by Theorem 2. Section 5 investigates the Gaussianity assumption, relating it to the use of modulus switching and characterizing the conditions under which it provides an accurate approximation. Section 6 we combine results from Section 4 and Section 5 to obtain a simplified and scalable method for noise estimation and parameter selection, which becomes independent of the multiplicative depth under our conditions. Section 7 compares our approach with state-of-the-art methods, showing how our parameter selection leads to a significant improvement over those currently adopted in major libraries such as OpenFHE and HElib. Finally, Section 8

concludes the paper and outlines possible directions for future research inspired by our results.

2 Preliminaries

In this section, we define the general notation and provide the mathematical background that we will use throughout the paper.

2.1 Notation

Let $\mathbb{Z}$ be the ring of integers, and for $d \in \mathbb{Z}_{>0}$ we denote by $\mathbb{Z}_d = \mathbb{Z}/d\mathbb{Z}$. Let n be a power of 2. We define $\mathcal{R}$ as the ring $\mathcal{R} = \mathbb{Z}[x]/\langle x^n + 1\rangle$ and $\mathcal{R}_d = \mathbb{Z}_d[x]/\langle x^n + 1\rangle$.

We use t and q to represent the plaintext and the ciphertext modulus, respectively, and $\mathcal{R}_t$ defines the *plaintext space*, where t is chosen such that $t \equiv 1 \mod 2n$. Moreover, to define the *ciphertext space*, we need to select $L = M + 1$ moduli, where M is the multiplicative depth of the circuit. Then, for each level $\ell \in \{0, \dots, L-1\}$, we denote

$$q_\ell = \prod_{j=0}^{\ell} p_j,$$

as the *ciphertext modulus* at level $L - 1 - \ell$. Sometimes, the ciphertext modulo q_{L-1} at level 0 will be indicated simply by q.

We use lowercase letters such as a for polynomials and bold letters, like $\boldsymbol{a}$, for vectors of polynomials. Moreover, we denote by $a|_i$ the coefficient of x^i of the polynomial a. For $a \in \mathcal{R}_d$, we let $[a]_d$ define its *centered reduction modulo* d, with coefficients in $[-d/2, d/2)$. Sometimes, we will also write $a \mod d$. Unless stated otherwise, we assume that the coefficients of the polynomials in $\mathcal{R}_d$ are always centered modulo d.

Finally, we recall that for $a, b \in \mathcal{R}$ the i-th coefficient of their product is given by (see, e.g., [4])

$$ab|_i = \sum_{j=0}^{n-1} \xi(i,j) a|_j b|_{i-j}, \quad \text{where} \quad \xi(i,j) = \begin{cases} 1 & \text{for } i-j \in [0,n) \\ -1 & \text{otherwise.} \end{cases} \tag{1}$$

2.2 Probabilistic distributions

Given a polynomial $a \in \mathcal{R}$ and a probabilistic distribution χ, the notation $a \leftarrow \chi$ is used to indicate that each coefficient of a is randomly and independently sampled according to χ. Some distributions that will be frequently considered are the following:

- $\mathcal{U}_q$ as the *uniform distribution* over $\mathbb{Z}_q$ where the representatives modulo q are taken in the interval $\left[-\frac{q}{2}, \frac{q}{2}\right)$;
- $\mathcal{N}(0, \sigma^2)$ as the normal distribution, also referred to as the *Gaussian distribution*, over $\mathbb{R}$, with mean 0 and variance σ^2;
- $\mathcal{DG}_q(\sigma^2)$ as the *discrete Gaussian distribution*, which involves sampling a value according to $\mathcal{N}(0, \sigma^2)$, rounding it to the nearest integer and then reducing it modulo q. Moreover, the representative modulo q is taken in the interval $\left[-\frac{q}{2}, \frac{q}{2}\right)$.

For the BGV scheme, the notation χ_e and χ_s will be adopted in order to indicate the distribution of the error for a RLWE instance and the secret key coefficients, respectively. Typically, χ_e is a discrete Gaussian distribution with a suitable standard deviation, while for the secret key, the ternary uniform distribution $\mathcal{U}_3$ is usually considered. However, in some special cases, where bootstrapping is needed, the choice for the secret key distribution falls on the Hamming Weight distribution where the secret key is sampled uniformly among sparse ternary vectors with a fixed number of non-zero entries.

Moreover, we denote by V_a the variance of the coefficients of the polynomial a, namely $\mathsf{Var}(a|_i)$. Given $\gamma \in \mathbb{Z}$, if $a, b \in \mathcal{R}$ are two independent polynomials whose coefficients are independent, identically distributed, and have zero mean, then [17]:

- $V_{a+b} = V_a + V_b$
- $V_{\gamma a} = \gamma^2 V_a$
- $V_{a \cdot b} = nV_a \cdot V_b$.

Finally, the values of the variance for some common distributions, which will often be employed in the BGV scheme, are

- $V_{\mathcal{DG}_q(\sigma^2)} = \sigma^2$ for the discrete Gaussian distribution centered at 0 with standard deviation σ;
- $V_3 = \frac{2}{3}$ for the ternary distribution $\mathcal{U}_3$;
- $V_q = \frac{q^2-1}{12} \approx \frac{q^2}{12}$ for the uniform distribution over integer values in $\left[-\frac{q}{2}, \frac{q}{2}\right)$;

2.3 Infinity and canonical norms

To conclude this section, we recall the definitions of the infinity and the canonical norm, along with some of their properties.

Definition 1. *The infinity norm of a polynomial $a \in \mathcal{R}$ is defined as*

$$\|a\|_\infty = \max_{0 \le i < n} | a|_i |.$$

If $a \in \mathcal{R}_q$ and its coefficients are well-approximated by identically distributed independent Gaussian variables centered at zero, then

$$\mathbb{P}\left(\|a\|_\infty > T\right) \leq n\left(1 - \operatorname{erf}\left(\frac{T}{\sqrt{2V_a}}\right)\right), \tag{2}$$

where $\operatorname{erf}(z)$ is the *error function*, defined as $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2}\, dt$, see, e.g., [4].

Definition 2. *The canonical embedding norm of a polynomial $a \in \mathcal{R}$ is*

$$\|a\|_{can} = \max_{\substack{1 \leq j < 2n \\ \gcd(j,2n)=1}} |a(\zeta^j)|,$$

where ζ is a primitive $2n$-th root of unity. Essentially, it corresponds to the infinity norm of the canonical embedding of a, denoted as $\sigma(a)$.

In our case, for $a, b \in \mathcal{R}$, the relationship between these two norms is given by $||a||_\infty \leq ||a||_{can}$. Moreover, we have [19]:

$$\begin{aligned} ||a \cdot b||_\infty &\leq n||a||_\infty \cdot ||b||_\infty \\ ||a \cdot b||_{can} &\leq ||a||_{can} \cdot ||b||_{can} \end{aligned}$$

We know that if $a \in \mathcal{R}_q$ is a random polynomial with coefficient variance V_a and ζ be a primitive $2n^{th}$ root of unity, then the distribution of $a(\zeta)$ is well approximated by a centred Gaussian distribution with variance nV_a [20]. This immediately translates into a bound on the canonical norm of a:

$$||a||^{can} < D\sqrt{nV_a}\,, \tag{3}$$

which holds with probability $(1 - e^{-D^2/2})^n \approx 1 - ne^{-D^2/2}$. This means that, for a suitable choice of D, the bound fails only with negligible probability [20].

3 The BGV scheme

In this section, we recall the three basic encryption functions of the BGV scheme and the homomorphic operations.

3.1 Basic encryption functions

Key Generation. The key generation function generates $s \leftarrow \chi_s$, $a \leftarrow \mathcal{U}_{q_{L-1}}$ and $e \leftarrow \chi_e$ and outputs the *secret key*: $\mathsf{sk} = s \in \mathcal{R}_{q_{L-1}}$, and the *public key* $\mathsf{pk} = (b, a) \equiv (-a \cdot s + te, a) \mod q_{L-1}$.

Encryption. Given the plaintext $m \in \mathcal{R}_t$ and the public key $\mathsf{pk} = (b, a)$, the encryption function outputs the ciphertext $\boldsymbol{c} \in \mathcal{R}^2_{q_{L-1}}$ defined as

$$\boldsymbol{c} = (c_0, c_1) \equiv (b \cdot u + te_0 + m, a \cdot u + te_1) \mod q_{L-1},$$

where $u, e_0, e_1 \in \mathcal{R}_{q_{L-1}}$, with coefficients distributed as $u \leftarrow \chi_s$ and $e_0, e_1 \leftarrow \chi_e$.

Decryption. Given a ciphertext $\boldsymbol{c} \in \mathcal{R}^2_{q_\ell}$ and the secret key $\mathsf{sk} = s$, the plaintext is recovered performing the following computations $m = \Big[[c_0 + c_1 \cdot s]_{q_\ell}\Big]_t$.

We denote by $\nu = [c_0 + c_1 \cdot s]_{q_\ell}$ the *critical quantity* corresponding to $\boldsymbol{c}$. In particular, for a fresh ciphertext, it can be rewritten as

$$\nu_{\mathsf{clean}} = [c_0 + c_1 \cdot s]_{q_{L-1}} = [m + t(e \cdot u + e_1 \cdot s + e_0)]_{q_{L-1}} = [m + t\epsilon]_{q_{L-1}},$$

where ϵ denotes the *error* introduced during encryption.

Considering the reduction modulo t of the critical quantity, it is possible to verify that the plaintext is successfully recovered. However, if the error is *too large*, the value $m + t\epsilon$ could wrap around the modulus, resulting in an incorrect decryption. So, the decryption is correct only if the coefficients of $m + t\epsilon$ remain below a certain threshold.

In addition, the error associated to the ciphertext increases through homomorphic operations [15]. Therefore, it is crucial to estimate the magnitude of the critical quantity (called also *noise*) which is typically analyzed using its norm.

In light of this, to guarantee the correctness of the decryption, the condition on the critical quantity can be expressed as follows [33]:

$$||\nu||_\infty \leq ||\nu||_{can} < \frac{q_\ell}{2}.$$

Naturally, in order to bound the noise, any type of norm could be used.

Finally, another concept that is often introduced for the estimation of error growth is the *noise budget*, which represents the number of bits remaining before wrap-around would occur.

We will use the term *extended ciphertext* to refer to the tuple $\mathfrak{c} = (\boldsymbol{c}, q_\ell, \nu)$ where $\boldsymbol{c}$ is the actual ciphertext, q_ℓ is the ciphertext modulus, and ν is the associated critical quantity.

Definition 3. *Let $(\boldsymbol{c}, q_\ell, \nu)$ be an extended ciphertext. The noise budget associated to $\boldsymbol{c}$ is the quantity $\log_2(q_\ell) - \log_2(||\nu||) - 1$, where $||\cdot||$ refers to a fixed norm.*

3.2 Homomorphic Operations

Ciphertext Addition. Let $(\boldsymbol{c}, q_\ell, \nu)$ and $(\boldsymbol{c'}, q_\ell, \nu')$ be two extended ciphertexts. Their homomorphic sum is

$$\mathsf{Add}(\boldsymbol{c}, \boldsymbol{c'}) := (c_0 + c'_0,\ c_1 + c'_1) \mod q_\ell.$$

with resulting critical quantity $\nu_{\mathsf{add}} = \nu + \nu'$.

Constant Multiplication. Let $(\boldsymbol{c}, q_\ell, \nu)$ be an extended ciphertext and $k \in \mathcal{R}_t$ a fixed polynomial. The homomorphic multiplication by k is

$$\mathsf{Mul}_k(\boldsymbol{c}) := (kc_0,\ kc_1) \mod q_\ell,$$

and the corresponding critical quantity scales by k, yielding $\nu_{\mathsf{const}} = k \cdot \nu$.

Homomorphic multiplication. Let $\boldsymbol{c}, \boldsymbol{c'}$ be two ciphertexts defined in $\mathcal{R}_{q_\ell}$, then

$$\mathsf{Mul}(\boldsymbol{c}, \boldsymbol{c'}) := (d_0, d_1, d_2) = (c_0 \cdot c'_0, c_0 \cdot c'_1 + c_1 \cdot c'_0, c_1 \cdot c'_1) \mod q_\ell.$$

As a result, the ciphertext expands from two to three polynomials, which violates the compactness property and makes subsequent operations more costly.

Recovering the message from $\mathsf{Mul}(\boldsymbol{c}, \boldsymbol{c'})$ requires computing the reduction modulo t of the resulting critical quantity given by $\nu_{\mathsf{mul}} = d_0 + d_1 s + d_2 s^2$.

To modify the ciphertext polynomial $d_0 + d_1 s + d_2 s^2$ back to another polynomial $\bar{c}_0 + \bar{c}_1 \cdot s$ encrypting the same plaintext, a technique known as *relinearization*, or *key switching*, is employed.

Key switching. Intuitively, the key switching technique converts $d_2 s^2$ into $\hat{c}_0 + \hat{c}_1 s$ using *somehow* the encryption of s^2 under s. Indeed,

$$\mathsf{Enc}_s(s^2) = (\beta, \alpha) = (-us + te + s^2, u + te_1) \approx (\alpha s + s^2, -\alpha).$$

Thus, $s^2 \approx \beta - \alpha s$ and then $d_0 + d_1 s + d_2 s^2 \approx d_0 + d_1 s + d_2(\beta - \alpha s) = \bar{c}_0 + \bar{c}_1 s$. As expected, while the relinearization step introduces additional noise, it is crucial for ensuring the practicality of the scheme.

Several key switching techniques have been proposed in the literature, each aiming to optimize this trade-off between correctness and noise growth. The most commonly used are the Brakerski Vaikuntanathan (BV) variant [8], the Gentry Halevi Smart (GHS) variant [26], and the Hybrid variant [26], which can be considered as a combination of the previous ones. Herein, we do not delve into the details of each method and refer the reader to [33] for further information.

Modulus switching. The primary aim of the modulus switching technique is to reduce the noise resulting from homomorphic operations.

Let $(\boldsymbol{c}, q_\ell, \nu)$ be the extended ciphertext whose error we aim to reduce, and let ℓ' be an integer such that $q_{\ell'} < q_\ell$. The modulus switching procedure outputs $(\boldsymbol{c}', q_{\ell'}, \nu')$, where

$$\boldsymbol{c}' = \frac{q_{\ell'}}{q_\ell}(\boldsymbol{c} + \boldsymbol{\delta}) \mod q_{\ell'},$$

with $\boldsymbol{\delta} = t[-\boldsymbol{c}t^{-1}]_{q_\ell/q_{\ell'}}$. The $\boldsymbol{\delta}$ value can be interpreted as a correction required to ensure that the ciphertext is divisible by $q_\ell/q_{\ell'}$ and does not affect the original message. In fact, it only influences the error since $\boldsymbol{\delta} \equiv 0 \mod t$. Therefore, the new ciphertext $\boldsymbol{c}'$ will still decrypt to the original plaintext (scaled by a factor of $q_\ell/q_{\ell'}$).

The critical quantity associated to the new ciphertext $\boldsymbol{c}'$ can be expressed in terms of that of $\boldsymbol{c}$, namely,

$$\nu_{\mathsf{ms}} = [c_0' + c_1' \cdot s]_{q_{\ell'}} = \frac{q_{\ell'}}{q_\ell}([c_0 + c_1 \cdot s]_{q_\ell} + \delta_0 + \delta_1 \cdot s) = \frac{q_{\ell'}}{q_\ell}(\nu + \delta_0 + \delta_1 \cdot s).$$

4 Average-Case Noise Analysis for BGV

In this section, we show that the critical quantities associated with ciphertexts encrypted under the same key should not be considered as mutually independent. We will present a method based on an average-case analysis that is able to explicitly account for these dependencies. This allows to overcome the underestimation of the error terms found in other works, which, as discussed above, must be avoided both for security reasons and for the correct functioning of the scheme.

Throughout this analysis, we consider ciphertexts that are mutually independent, obtained by encrypting independently generated random messages using the same public key. However, even when the ciphertexts themselves are independent, the critical quantities associated with them are generally dependent. These dependencies arise because the noise in the ciphertexts includes terms involving powers of the secret key s and powers of the error term e, which makes it necessary to explicitly consider these contributions when analyzing the variance of the noise.

To study the impact of s and e, we highlight their contribution in the expression of the critical quantity ν, using the following notation:

$$\nu = \sum_\iota a_\iota s^\iota = \sum_\iota \sum_\mu b_\mu(\iota) e^\mu s^\iota,$$

where $a_\iota = \sum_\mu b_\mu(\iota) e^\mu$, and $b_\mu(\iota)$ contains no powers of s or e.

To enhance clarity, for the critical quantity ν_{clean} of a fresh ciphertext $\boldsymbol{c}$, this notation yields

$$\nu_{\mathsf{clean}} = a_0 + a_1 s = b_0(0) + b_1(0)e + b_0(1)s,$$

where

$$\begin{cases} a_0 = b_0(0) + b_1(0) \cdot e = (m + te_0) + tu \cdot e \\ a_1 = b_0(1) = te_1 \end{cases}$$

The need to explicitly account for these dependencies becomes immediately clear when analyzing the growth of the error after multiplying two ciphertexts. As discussed earlier, in the average-case approach we focus on studying the variance of the coefficients of the critical quantity resulting from the homomorphic operations. Let ν and ν', denote the critical quantities associated with two distinct ciphertexts. If these quantities were independent, the variance of the coefficients of their product could be estimated as $n\mathsf{Var}(\nu|_i)\mathsf{Var}(\nu'|_i)$. Instead, in this section we prove that this variance should be estimated as $n\mathsf{Var}(\nu|_i)\mathsf{Var}(\nu'|_i)F_sF_e$ with F_s and F_e depending on the number of multiplications previously performed in the circuit.

Before rigorously introducing Theorem 2, which provides a method for studying the growth of error through its variance in fixed circuits, we begin by presenting some considerations and results we have derived regarding the distribution of the coefficients of a generic error term, along with certain properties related to their variances.

4.1 Mean and Variance Analysis

In the following, we prove that the coefficients of the error term are centered at zero. Furthermore, we show that the coefficients of the b_μ terms are uncorrelated, meaning that their pairwise covariance is zero.

Lemma 1. *Let $\nu = \sum_\iota \sum_\mu b_\mu(\iota)e^\mu s^\iota$ be the critical quantity associated with a given ciphertext. Then, the following properties hold*

a) $\mathsf{Cov}(b_{\mu_1}(\iota_1)|_{j_1}, b_{\mu_2}(\iota_2)|_{j_2}) = 0$ *for* $\mu_1 \neq \mu_2$ *or* $j_1 \neq j_2$, $\forall \iota_1, \iota_2$;
b) $\mathbb{E}[b_\mu(\iota)|_i] = 0$, $\forall \iota, \mu, i$;

A proof of Lemma 1 can be found in Appendix A

Lemma 2. *Let $\nu = \sum_\iota a_\iota s^\iota$ represent the critical quantity associated with a given ciphertext, where, for a fixed ι, $a_\iota = \sum_\mu b_\mu(\iota)e^\mu$. Then, the following identity holds*

$$\mathsf{Var}(a_\iota s^\iota|_i) = \sum_{\mu \geq 0} \mathsf{Var}(b_\mu(\iota)|_i) \sum_{k=0}^{n-1} \mathbb{E}[e^\mu|_k^2] \sum_{j=0}^{n-1} \mathbb{E}[s^\iota|_j^2].$$

Moreover, $\mathsf{Var}(\nu|_i) = \sum_{\iota,\mu\geq 0} \mathsf{Var}(b_\mu(\iota)|_i) \sum_{k=0}^{n-1} \mathbb{E}[e^\mu|_k^2] \sum_{j=0}^{n-1} \mathbb{E}[s^\iota|_j^2]$.

A proof of Lemma 2 can be found in Appendix B.

4.2 Homomorphic Operations

We can state our results on the variance computation for operations, especially focusing on the multiplication in the next section.

Proposition 1 (Encryption). *The invariant noise* ν_{clean} *of a fresh ciphertext has coefficient variance*

$$V_{\mathsf{clean}} = \mathsf{Var}(\nu_{\mathsf{clean}}|_i) = \frac{t^2}{q^2}\left(\frac{1}{12} + nV_eV_u + V_e + nV_eV_s\right). \tag{4}$$

Proof. The fresh error ν_{clean} can be written as $a_0 + a_1 s = b_0(0) + b_1(0)e + b_0(1)s$, where

$$\begin{cases} a_0 = b_0(0) + b_1(0)\cdot e = (m + te_0) + tu\cdot e \\ a_1 = b_0(1) = te_1. \end{cases}$$

The proof is thus concluded by applying Lemma 2 and observing that $\mathbb{E}[s|_j^2] = \mathsf{Var}(s|_j) = V_s$ and $\mathbb{E}[e|_j^2] = \mathsf{Var}(e|_j) = V_e$ since $\mathbb{E}[s|_j] = \mathbb{E}[e|_j] = 0$.

Proposition 2 (Addition). *Let* $\alpha \in \mathcal{R}_t$ *and* $\mathbf{c}$, $\mathbf{c}'$ *be two independently-computed ciphertexts with invariant noises* ν, *and* ν', *respectively. Then, the variance of the error coefficients resulting from the addition of* $\mathbf{c}$ *and* $\mathbf{c}'$ *is*

$$\mathsf{Var}(\nu_{\mathsf{add}}|_i) = \mathsf{Var}(\nu|_i) + \mathsf{Var}(\nu'|_i). \tag{5}$$

Proposition 3 (Modulo Switch). *Let* $\mathfrak{c} = (\mathbf{c}, q_\ell, \nu)$ *be an extended ciphertext. The variance of the error coefficients after the modulo switch to the target modulo* q'_ℓ *is*

$$V(\nu_{\mathsf{ms}}|_i) = \frac{q_\ell'^2}{q_\ell^2}(\mathsf{Var}(\nu|_i) + V_{\delta_0} + nV_{\delta_1}V_s)) = \frac{q_\ell'^2}{q_\ell^2}\mathsf{Var}(\nu|_i) + \frac{t^2}{12}(1 + nV_s). \tag{6}$$

Propositions 2 and 3 follow directly from Lemma 2 by arguments analogous to those in Proposition 1. The proofs are therefore omitted.

We do not explicitly address key switching here, since its negligible impact is discussed in Appendix C.

4.3 Homomorphic multiplications

The main goal of our analysis is to estimate the error growth caused by homomorphic multiplications, the most critical operations to handle. This requires accounting for the dependencies among noise coefficients induced by such multiplications. Our method relies on a *correction function* F, whose construction is detailed in this section. We recall that the function F was introduced in [4] to "correct" the product of the variances, since these are not independent. Unlike BFV, the BGV scheme demands particular care, as both terms s and e must be considered, whereas the contribution of e is negligible in BFV. Furthermore, unlike [4], our construction of F is *not* heuristic but is based on Lemma 3, which we prove in this section. We recall the Isserlis theorem which will be exploited in the proof of the Lemma.

Theorem 1. *[31, Chapter 8] Let $(X_1, \ldots, X_n)$ be a zero-mean multivariate normal random vector, then*

$$\mathbb{E}[X_1 \cdots X_n] = \sum \prod_{i,j} \mathbb{E}[X_i X_j]\,,$$

where the sum is over all the partition of $\{1, \ldots, n\}$ into pairs and the (i, j) ranges in these pairs.

Lemma 3. *Let $a(x) = a|_0 + a|_1 x + \cdots + a|_{n-1} x^{n-1} \in \mathcal{R}$, where $a|_i$'s are i.i.d. random variables with $\mathbb{E}[a|_i] = 0$, $\mathbb{E}[a|_i^2] = V_a$.*
Then, as n approaches infinity,

$$\mathbb{E}[a^k|_i] = 0, \quad \mathbb{E}[(a^k|_i)^2] = k! n^{k-1} V_a^k\,,$$

for all $0 \leq i \leq n-1$, where $a^k|_i$ denotes the i–th coefficient of the polynomial $a(x)^k$.

Proof. Without loss of generality, we prove the statement for $a^k|_0$, since the same argument applies to all other terms. Let $\omega_1, \ldots, \omega_n$ be the roots of $x^n + 1$, which is a cyclotomic polynomial since n is a power of two. Then, we know that $\sum_{i=1}^{n} \omega_i^m = 0$ for $1 \leq m \leq n-1$, and $a(\omega_1)^k + \cdots + a(\omega_n)^k = n a^k|_0 + a^k|_1(\omega_1 + \cdots + \omega_n) + \cdots + a^k|_{n-1}(\omega_1^{n-1} + \cdots + \omega_n^{n-1})$, thus

$$a^k|_0 = \frac{1}{n}(a(\omega_1)^k + \cdots + a(\omega_n)^k).$$

Now, we define $Z_i := a(\omega_i) = a_0 + a_1\omega_i + \cdots + a_{n-1}\omega_i^{n-1}$, for $i = 1, \ldots, n$, where, for n sufficiently large, thanks to [20, Theorem 9] we have that Z_i follows a Gaussian distribution centered at zero. Then

$$\mathbb{E}[Z_m \cdot Z_l] = \mathbb{E}[(a|_0 + a|_1\omega_m + \cdots + a|_{n-1}\omega_m^{n-1})(a|_0 + a|_1\omega_l + \cdots + a|_{n-1}\omega_l^{n-1})] =$$

$$
\begin{aligned}
&= \mathbb{E}[a|_0^2] + \mathbb{E}[a|_1^2]\omega_m\omega_l + \cdots + \mathbb{E}[a|_{n-1}^2]\omega_m^{n-1}\omega_l^{n-1} = \\
&= V_a(1 + \omega_m\omega_l + \cdots + \omega_m^{n-1}\omega_l^{n-1}) = \\
&= \begin{cases} nV_a & \text{if} \quad \omega_m = \omega_l^{-1} \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
$$

Therefore, for $m = l$, $\mathbb{E}[Z_m^2] = 0$ and, since every linear combination of Z_i's is normally distributed, we can apply the Isserlis theorem (Theorem 1) to obtain $\mathbb{E}[Z_m^k] = 0$. Consequently

$$\mathbb{E}[a^k|_0] = \frac{1}{n}(\mathbb{E}[Z_1^k] + \cdots + \mathbb{E}[Z_n^k]) = 0.$$

Moreover, $\mathbb{E}[a^k|_0^2] = \frac{1}{n^2}\sum_m \mathbb{E}[Z_m^k Z_l^k] = \frac{1}{n^2}\sum_{m,l} \mathbb{E}[Z_m^k Z_{m^{-1}}^k] = k! n^{k-1} V_a^k$, where the last equality follows again from the Isserlis theorem. Indeed,

$$
\begin{aligned}
\mathbb{E}[Z_m^k Z_{m^{-1}}^k] &= \mathbb{E}[Z_m \cdot \cdots \cdot Z_m \cdot Z_{m^{-1}} \cdot \cdots \cdot Z_{m^{-1}}] = \\
&= \mathbb{E}[X_1 \cdot \cdots \cdot X_m \cdot X_{m+1} \cdot \cdots \cdot X_{2m}] = \\
&= \sum \prod \mathbb{E}[X_i X_j] = \\
&= \begin{cases} 0 & \text{if } 0 \le i,j \le k \text{ or } k+1 \le i,j \le 2k+1 \\ nV_a & \text{otherwise.} \end{cases}
\end{aligned}
$$

Therefore, $\mathbb{E}[Z_m^k Z_{m^{-1}}^k]$ is given by the sum of $k!$ addends whose value is $n^k V_a^k$, i.e., $\mathbb{E}[Z_m^k Z_{m^{-1}}^k] = k! n^k V_a^k$. Finally, $\mathbb{E}[a^k|_0^2] = \frac{1}{n^2} \cdot n \cdot k! \cdot n^k \cdot V_a^k = k! n^{k-1} V_a^k$. □

Remark 1 *This result extends some of [24, Theorem 4.3], where the authors proved the same statement, using a different approach, but only for the case where the $a|_i$'s are normal independent random variables. In our case, the coefficients of the polynomial $a(x)$ follow any distribution (with zero mean and finite variance). Moreover, a proof of the specific case $k = 2$ is also provided in [5, Lemma 2].*

Definition 4. *Let $a \in \mathcal{R}$. The correction function F_a is defined as*

$$F_a(\iota_1, \iota_2) = \frac{\sum_{i=0}^{n-1} \mathbb{E}[a^{\iota_1+\iota_2}|_i^2]}{\sum_{i_1=0}^{n-1} \mathbb{E}[a^{\iota_1}|_{i_1}^2] \sum_{i_2=0}^{n-1} \mathbb{E}[a^{\iota_2}|_{i_2}^2]}. \tag{7}$$

From Lemma 3, it follows that for polynomials with coefficients following distributions of that form, it holds that

$$F_a(\iota_1, \iota_2) = \frac{(\iota_1 + \iota_2)!}{\iota_1!\, \iota_2!}. \tag{8}$$

We observe that the value of F_a is the same for all polynomials as in Lemma 3. However, for the sake of clarity in presenting our results, we will keep separate notations, writing F_s for the secret key s and F_e for the public key e.

We can now introduce our main theorem.

Theorem 2. *Let $\nu = \sum_\iota a_\iota s^\iota, \nu' = \sum_\iota a'_\iota s^\iota$ be the critical quantities of two independently computed ciphertexts defined with respect to the same modulus q. Then*

$$\mathsf{Var}((a_{\iota_1} s^{\iota_1} a'_{\iota_2} s^{\iota_2})|_i) \leq n\mathsf{Var}((a_{\iota_1} s^{\iota_1})|_i)\mathsf{Var}((a'_{\iota_2} s^{\iota_2})|_i)F_s(\iota_1,\iota_2)F_e(K_1,K_2),$$

where K_1, K_2 represent the highest power of e appearing in $a_{\iota_1}, a'_{\iota_2}$, respectively.

Proof. From Lemma 2 it is possible to express the variance of two generic terms $a_{\iota_1} s^{\iota_1}|_i$ and $a'_{\iota_2} s^{\iota_2}|_i$ as

$$\begin{cases} \mathsf{Var}(a_{\iota_1} s^{\iota_1}|_i) = \sum_{\mu_1=0}^{K_1} \mathsf{Var}(b_{\mu_1}(\iota_1)|_i) \sum_{j_1=0}^{n-1} \mathbb{E}[e^{\mu_1}|_{j_1}^2] \sum_{j_2=0}^{n-1} \mathbb{E}[s^{\iota_1}|_{j_2}^2] \\ \mathsf{Var}(a'_{\iota_2} s^{\iota_2}|_i) = \sum_{\mu_2=0}^{K_2} \mathsf{Var}(b'_{\mu_2}(\iota_2)|_i) \sum_{j_3=0}^{n-1} \mathbb{E}[e^{\mu_2}|_{j_3}^2] \sum_{j_4=0}^{n-1} \mathbb{E}[s^{\iota_2}|_{j_4}^2] \end{cases}$$

By observing that

$$a_{\iota_1} s^{\iota_1} \cdot a'_{\iota_2} s^{\iota_2} = \left(\sum_\mu \sum_{\mu_1+\mu_2=\mu} b_{\mu_1}(\iota_1) b'_{\mu_2}(\iota_2) e^\mu \right) s^{\iota_1+\iota_2},$$

and using Lemma 2, it is possible to write the variance $\mathsf{Var}((a_{\iota_1} s^{\iota_1} \cdot a'_{\iota_2} s^{\iota_2})|_i)$ as

$$\begin{aligned} &\sum_\mu \mathsf{Var}\left(\sum_{\mu_1+\mu_2=\mu} (b_{\mu_1}(\iota_1) b'_{\mu_2}(\iota_2))|_i \right) \sum_{j_1=0}^{n-1} \mathbb{E}[e^\mu|_{j_1}^2] \sum_{j_2=0}^{n-1} \mathbb{E}[s^{\iota_1+\iota_2}|_{j_2}^2] \\ &= n \sum_\mu \sum_{\mu_1+\mu_2=\mu} \mathsf{Var}\left(b_{\mu_1}(\iota_1)|_i\right) \mathsf{Var}\left(b'_{\mu_2}(\iota_2)|_i\right) \sum_{j_1=0}^{n-1} \mathbb{E}[e^\mu|_{j_1}^2] \sum_{j_2=0}^{n-1} \mathbb{E}[s^{\iota_1+\iota_2}|_{j_2}^2], \end{aligned}$$

where the second equality follows from the independence of $b_{\mu_1}(\iota_1), b'_{\mu_2}(\iota_2)$ and from $\mathsf{Cov}(b_{\mu_1}(\iota)|_{j_1}, b_{\mu_2}(\iota)|_{j_2}) = 0$ for $\mu_1 \neq \mu_2$ or $j_1 \neq j_2$.

Moreover, it can be noted that $n\mathsf{Var}(a_{\iota_1} s^{\iota_1}|_i)\mathsf{Var}(a'_{\iota_2} s^{\iota_2}|_i)$ can be written as

$$\begin{aligned} &n \sum_\mu \sum_{\mu_1+\mu_2=\mu} \mathsf{Var}(b_{\mu_1}(\iota_1)|_i)\mathsf{Var}(b'_{\mu_2}(\iota_2)|_i) \\ &\cdot \sum_{j_1=0}^{n-1} \mathbb{E}[e^{\mu_1}|_{j_1}^2] \sum_{j_3=0}^{n-1} \mathbb{E}[e^{\mu_2}|_{j_3}^2] \sum_{j_2=0}^{n-1} \mathbb{E}[s^{\iota_1}|_{j_2}^2] \sum_{j_4=0}^{n-1} \mathbb{E}[s^{\iota_2}|_{j_4}^2], \end{aligned}$$

By (7), we express $\mathsf{Var}((a_{\iota_1} s^{\iota_1} \cdot a'_{\iota_2} s^{\iota_2})|_i)$ in terms of $\mathsf{Var}(a_{\iota_1} s^{\iota_1}|_i)\mathsf{Var}(a'_{\iota_2} s^{\iota_2}|_i)$. In fact, $\mathsf{Var}((a_{\iota_1} s^{\iota_1} \cdot a'_{\iota_2} s^{\iota_2})|_i)$ can be written as

$$n \sum_\mu \sum_{\mu_1+\mu_2=\mu} \mathsf{Var}\left(b_{\mu_1}(\iota_1)|_i\right) \mathsf{Var}\left(b'_{\mu_2}(\iota_2)|_i\right) \sum_{j=0}^{n-1} \mathbb{E}[e^{\mu_1+\mu_2}|_j^2] \sum_{j'=0}^{n-1} \mathbb{E}[s^{\iota_1+\iota_2}|_{j'}^2].$$

Thus, by leveraging the properties of the correction functions

$$nF_s(\iota_1,\iota_2) \sum_\mu \sum_{\mu_1+\mu_2=\mu} \mathsf{Var}\left(b_{\mu_1}(\iota_1)|_i\right) \mathsf{Var}\left(b'_{\mu_2}(\iota_2)|_i\right) F_e(\mu_1,\mu_2)\cdot$$

$$\cdot \sum_{j_1=0}^{n-1} \mathbb{E}[e^{\mu_1}|_{j_1}^2] \sum_{j_2=0}^{n-1} \mathbb{E}[e^{\mu_2}|_{j_2}^2] \sum_{j_3=0}^{n-1} \mathbb{E}[s^{\iota_1}|_{j_3}^2] \sum_{j_4=0}^{n-1} \mathbb{E}[s^{\iota_2}|_{j_4}^2]$$

$$\approx nF_s(\iota_1,\iota_2) \sum_{\mu} \sum_{\mu_1+\mu_2=\mu} \mathsf{Var}\left(b_{\mu_1}(\iota_1)|_i\right) \mathsf{Var}\left(b'_{\mu_2}(\iota_2)|_i\right) F_e(\mu_1,\mu_2) \cdot$$

$$\cdot \sum_{j_1=0}^{n-1} \mathbb{E}[e^{\mu_1}|_{j_1}^2] \sum_{j_2=0}^{n-1} \mathbb{E}[e^{\mu_2}|_{j_2}^2] \sum_{j_3=0}^{n-1} \mathbb{E}[s^{\iota_1}|_{j_3}^2] \sum_{j_4=0}^{n-1} \mathbb{E}[s^{\iota_2}|_{j_4}^2].$$

Then, by exploiting the monotonicity of F_e and F_s, it is possible to derive an upper bound given by

$$nF_s(\iota_1,\iota_2)F_e(K_1,K_2) \sum_{\mu} \sum_{\mu_1+\mu_2=\mu} \mathsf{Var}\left(b_{\mu_1}(\iota_1)|_i\right) \mathsf{Var}\left(b'_{\mu_2}(\iota_2)|_i\right) \cdot$$
$$\cdot \sum_{j_1=0}^{n-1} \mathbb{E}[e^{\mu_1}|_{j_1}^2] \sum_{j_2=0}^{n-1} \mathbb{E}[e^{\mu_2}|_{j_2}^2] \sum_{j_3=0}^{n-1} \mathbb{E}[s^{\iota_1}|_{j_3}^2] \sum_{j_4=0}^{n-1} \mathbb{E}[s^{\iota_2}|_{j_4}^2],$$

where K_1, K_2 represent the highest power of e appearing in $a_{\iota_1}, a'_{\iota_2}$, respectively. It is straightforward to verify that this concludes our proof, yielding

$$\mathsf{Var}((a_{\iota_1}s^{\iota_1} \cdot a'_{\iota_2}s^{\iota_2})|_i)' \leq n\mathsf{Var}(a_{\iota_1}s^{\iota_1}|_i)\mathsf{Var}(a'_{\iota_2}s^{\iota_2}|_i)F_s(\iota_1,\iota_2)F_e(K_1,K_2).$$

□

The central idea is that, given the multiplicative depth of a circuit, one can precisely identify the powers of e and s present in the resulting critical quantity and, consequently, derive an accurate estimate of the variance of its coefficients. By systematically combining this analysis with the variance contributions from other operations, one obtains a general framework for estimating the variance in arbitrary circuits.

5 The Shape of Noise: Gaussian Distributions in BGV

In order to provide accurate bounds for the selection of initial parameters, it is necessary to first discuss the distribution of the error coefficients. Our analysis relies on the assumption that these coefficients follow a normal distribution. Under this assumption, tight average-case bounds can be derived, facilitating efficient and reliable parameter selection.

However, assuming a Gaussian distribution when this does not hold may result in a significant underestimation of the error bounds, causing computation failures, as highlighted in [24]. It is worth noting that [24] claims that variance-based methods offer no theoretical guarantee on the failure probability. This statement can be misleading, as variance-based analyses remain valid once an appropriate bound on the actual distribution is established.

Furthermore, we emphasize that the results proven in [24] concern the heavy-tailed nature of the error coefficient distributions in the absence of modulus switching, which is not practical, since BGV is always used with modulus switching technique. As noted in the article itself, applying modulus switching ensures that the dominant terms are effectively distributed according to Gaussian distributions. Therefore, assuming Gaussianity remains reasonable when analyzing BGV parameters, in line with other works such as [18,34]. Indeed, as the authors of [24] themselves note, no evidence of decryption failures has been observed in practical BGV usage scenarios, nor in those of CKKS [5].

This section aims to rigorously justify why the Gaussian assumption for the error coefficients is reasonable and to identify the conditions on the choice of initial parameters that are required to ensure it. In particular, we will show that the values of the primes p_ℓ must satisfy a certain lower bound. This requirement is reasonable: the primes used in the most common libraries already meet this criterion. Moreover, while imposing a minimum size on these primes, our result still allows for improvements and a reduction in the overall modulus size.

5.1 Gaussian Distribution

In this section, we provide both theoretical arguments and empirical evidence that the use of modulus switching ensures that the error coefficients are, in practice, Gaussian distributed.
We show that the coefficients of the critical quantity can be reasonably assumed to follow a Gaussian distribution at three distinct stages: immediately after the initial encryption, after modulus switching, and after the multiplication of two ciphertexts that have both undergone modulus switching beforehand. Specifically, we provide a formal proof for the first two cases, while for the latter we verify this behavior experimentally. We point out that the proofs of the following lemmas rely on the Central Limit Theorem. Therefore, it is implicitly assumed that the ring dimension n is sufficiently large for the approximation to hold.

Lemma 4 (Gaussianity of the Fresh Error). *The coefficients of the critical quantity associated with a fresh ciphertext follow a Gaussian distribution.*

Proof. The critical quantity associated with a fresh ciphertext is given by

$$m + t(e \cdot u + e_1 \cdot s + e_0),$$

where $e, e_0, e_1 \sim \mathcal{DG}_q(\sigma^2)$ and $m, s, u \sim \mathcal{U}_3$. We are thus interested in the distribution of the coefficients

$$m|_i + t(\sum_{j=0}^{n-1} e|_j u|_{i-j} + \sum_{j=0}^{n-1} e_1|_j s|_{i-j} + e_0|_i),$$

where $m|_i \longleftarrow \mathcal{U}_t$, $te_0|_i \longleftarrow \mathcal{DG}_q(t^2\sigma^2)$ and $t\sum_{j=0}^{n-1} e|_j u|_{i-j}, t\sum_{j=0}^{n-1} e_1|_j s|_{i-j} \longleftarrow \mathcal{DG}_q(t^2 n V_e V_s)$. Since the sum of independent Gaussian random variables is still Gaussian with variance equal to the sum of individual variances [23], the error term can be written as $m|_i + K|_i$, where $m|_i \longleftarrow \mathcal{U}_t$ and $K|_i \longleftarrow \mathcal{DG}_q(t^2 V_e(2nV_s + 1))$. Therefore, the Gaussianity of the final critical quantity follows from the fact that the first term is negligible compared to the second. □

Lemma 5 (Gaussianity after Modulus Switching). *Let $\nu_{\mathsf{ms}} = \frac{q_{\ell'}}{q_\ell}(\nu + \delta_0 + \delta_1 s)$ be the critical quantity after modulo switching to the target modulo q'_ℓ, where $q_\ell = p_\ell q'_\ell$. If p_ℓ is chosen big enough such that $\mathsf{Var}(\nu/p_\ell|_i) \ll \mathsf{Var}(\delta_1 s/p_\ell|_i)$, then the coefficients of the critical quantity after the modulus switching follow a Gaussian distribution.*

Proof. We recall that the variance of the error coefficients after modulo switching to the target modulo q'_ℓ, where $q_\ell = p_\ell q'_\ell$, is

$$V(\nu_{\mathsf{ms}}|_i) = \frac{1}{p_\ell^2}(\mathsf{Var}(\nu|_i) + V_{\delta_0} + nV_{\delta_1}V_s) \approx \frac{1}{p_\ell^2}(\mathsf{Var}(\nu|_i) + nV_{\delta_1}V_s). \tag{9}$$

as $\mathsf{Var}(\delta_0/p_\ell|_i) = t^2/12 \ll \mathsf{Var}(\delta_1 s/p_\ell|_i) = nt^2V_s/12$.

If p_ℓ is sufficiently large, we can simplify the analysis by approximating the critical quantity after the modulus switch as $\nu_{\mathsf{ms}} \approx \delta_1/p_\ell s$, with $\delta_1/p_\ell \longleftarrow \mathcal{U}_t$. Thus, we obtain

$$\frac{\delta_1}{p_\ell} s|_i = \sum_{j=0}^{n-1} \xi(i,j)\ \frac{\delta_1}{p_\ell}|_j\ s|_{i-j}.$$

Therefore, we can observe that each coefficient is given by the sum of i.i.d. random variables with finite mean and variance. Since n is sufficiently large, by applying the Central Limit Theorem, we can conclude that

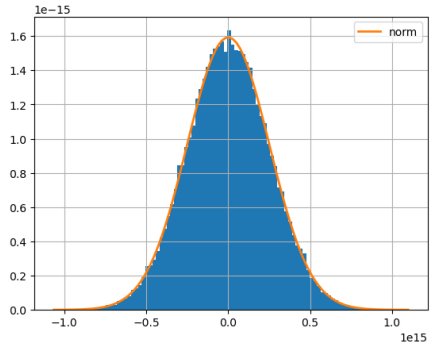
$$\frac{\delta_1}{p_\ell} s \longleftarrow \mathcal{DG}_q(n\frac{t^2}{12}V_s).$$

□

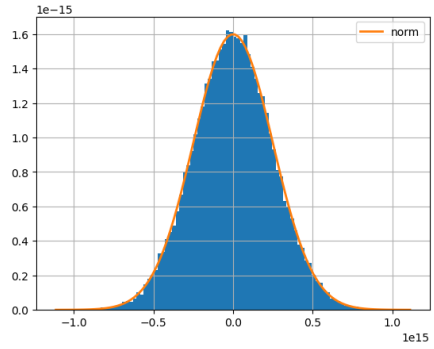
Remark 2 *In our analysis, we assume that the dominant component of the error after modulus switch is given by the term $\delta_1 \cdot s$. This assumption is quite common in the study of BGV, as can be seen in [18]. We will later clarify how our choice of moduli is specifically designed to ensure that this condition is effectively satisfied.*

Heuristic 1 (Gaussianity after a multiplication) *The critical quantity resulting from the multiplication of two ciphertexts that have just undergone modulus switch follows a Gaussian distribution.*

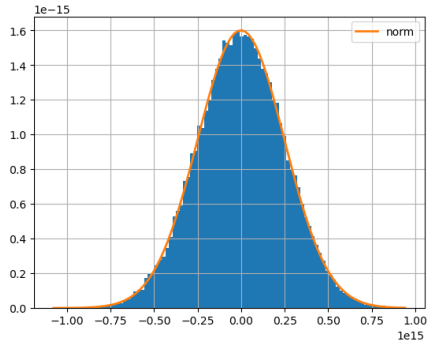
In this setting, we restrict our analysis to an experimental evaluation of the distribution of the coefficients of the critical quantity obtained from the multiplication of two ciphertexts that have just undergone modulus switching, under the conditions described above - namely, when the ratio between the moduli at consecutive levels is sufficiently large.



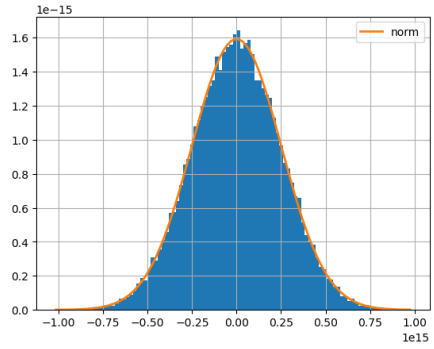
(a) $\mathrm{ad}_{\mathrm{stat}}$ 0.267 < $\mathrm{ad}_{\mathrm{crit}}$ 0.576 ($\alpha = 0.15$) and kurtosis 2.987. After the third multiplication.



(b) $\mathrm{ad}_{\mathrm{stat}}$ 0.306 < $\mathrm{ad}_{\mathrm{crit}}$ 0.576 ($\alpha = 0.15$) and kurtosis 3.03. After the fourth multiplication.



(c) $\mathrm{ad}_{\mathrm{stat}}$ 0.233 < $\mathrm{ad}_{\mathrm{crit}}$ 0.576 ($\alpha = 0.15$) and kurtosis 2.980. After the 3rd multiplication in a circuit of multiplicative depth 6.



(d) $\mathrm{ad}_{\mathrm{stat}}$ 0.215 < $\mathrm{ad}_{\mathrm{crit}}$ 0.576 ($\alpha = 0.15$) and kurtosis 2.985. After the 5th multiplication in a circuit of multiplicative depth 6.

Fig. 1: Distribution of the first coefficient just after a multiplication.

In Figure 1, we present the results for circuits of multiplicative depth 3 (Figure 1a) and 4 (Figure 1b), focusing on the distribution of the first coefficient of the error just after the last multiplication. Moreover, Figures 1c and 1d show the multiplication error after the third and fifth multiplications, respectively, for a circuit of depth 6. We used OpenFHE library [1] (with the same setting as

before) to generate 50,000 error samples and analyzed their coefficients using the Python **fitter** package[5]. The code used to generate these samples and to perform the estimates in the following sections is publicly available[6]. As the figures show, the distribution of the error coefficients is well approximated by a Gaussian. We evaluate the Gaussianity of the data using the Anderson–Darling test [2,37], as it is particularly sensitive to deviations in the tails of the distribution. Specifically, the significance level determines the threshold for rejecting the null hypothesis: smaller levels (e.g., 5%) correspond to stricter tests, while larger levels (e.g., 15%) are more permissive. More precise, heavy-tailed behavior would result in larger deviations in the tails, reflected by an increased Anderson–Darling statistic and a kurtosis significantly larger than 3.

Advantages of having a Gaussian distribution The fact that the noise coefficients follow a Gaussian distribution is particularly advantageous, as it allows us to bound the maximum absolute value of the noise coefficients with high probability simply by controlling their variance V. This implies that, to satisfy the correctness condition, a noise vector ν must satisfy $\|\nu\|_\infty < q/2$. We can use Equation (2) to bound the failure probability

$$\mathbb{P}\left(\|\nu\|_\infty > \frac{q}{2}\right) \leq n\left(1 - \operatorname{erf}\left(\frac{q}{2\sqrt{2V}}\right)\right),$$

where V denotes the estimated variance of the noise coefficients.

To express this bound more conveniently, we introduce a *security parameter* D such that $D \leq q/2\sqrt{2V}$, from which it follows, using the monotonicity of the error function, that

$$\mathbb{P}\left(\|\nu\|_\infty > \frac{q}{2}\right) \leq n\left(1 - \operatorname{erf}(D)\right).$$

Thus, by appropriately choosing the security parameter D, we can ensure that the probability of decryption failure remains negligibly small. For instance, setting $D = 6$ and $n = 2^{13}$ yields a failure probability of approximately 2^{-42}. For practical applications, $D = 8$ should be preferred, resulting in a probability of approximately 2^{-83}, when the ring dimension is $n = 2^{13}$.

Consequently, the ciphertext modulus q can be selected as

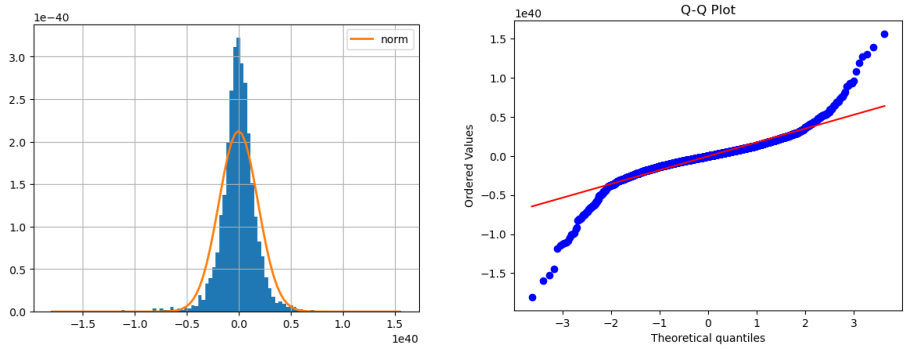
$$q \geq 2D\sqrt{2V}. \tag{10}$$

It is worth noting that the bounds derived in our analysis provide insight into the *minimum* ciphertext modulus q required to guarantee the correctness of the scheme, a key factor for optimizing performance and efficiency.

[5] https://fitter.readthedocs.io/en/latest/
[6] https://github.com/nadirmur/openFHE

5.2 On the Role of Modulus Switching in Preserving Gaussianity

In this section, we highlight the crucial role of modulus switching in the studied circuits. This mechanism is essential not only to control the noise growth induced by homomorphic operations, but also to *shape* the error distribution. Our experiments confirm this behavior: when modulus switching is omitted, the Gaussianity of the coefficients could no longer hold, making it necessary to explicitly characterize their distribution. Specifically, Figure 2 shows the error distribution after three multiplications without modulus switching for $n = 2^{13}$, where the tails clearly deviate from Gaussian behavior.



(a) asymmetry −0.5240 and kurtosis 14.3769

(b) when the distribution is Gaussian, the points lie along the diagonal.

Fig. 2: Error distribution after 3 multiplications without modulus switching.

It is worth noting that this deviation can be observed even without plotting the distribution. Indeed, when the Gaussian bound introduced in the previous section is applied to select the ciphertext modulus q according to Equation (10), the resulting parameters quickly lead to an unexpectedly high failure rate. For instance, when targeting a decryption failure probability of approximately 2^{-83} (as obtained by setting $D = 8$ in our bound), experiments without modulus switching exhibit failure rates of around 3%. This discrepancy provides strong evidence that the Gaussian assumption no longer holds and that the tails of the distribution are significantly heavier, in agreement with [24].

The validity of our approach is therefore directly linked to the use of modulus switching. As previously shown, the error coefficients after encryption, modulus switching, and multiplication preserve their Gaussian distribution. For this property to hold, as highlighted in the previous sections, it is necessary to ensure that the dominant terms after modulus switching and after multiplication are, respectively, $\delta_1 s/p_\ell$ and $(\delta_1/p_\ell \cdot \delta_1'/p_\ell)s^2$. For this reason, in our moduli selection,

we explicitly require that

$$\frac{V_{\ell-1}}{p_{L-\ell}^2} < \alpha V_{\mathsf{ms}}. \tag{11}$$

for every level ℓ, with $\alpha = 1/100$. This condition imposes a lower bound on the moduli p_ℓ, which must be selected accordingly. Nevertheless, it does not compromise the modulus-size reduction achieved by our approach compared to the state-of-the-art. In fact, the primes generated by most existing libraries already satisfy this requirement.

6 Variance Estimation in Circuits with Fixed Depth

This section builds on the results of Section 4 and 5 to propose a method for tracking the error growth in circuits with a fixed number of operations. As previously noted, obtaining sufficiently tight bounds with respect to the experimental variance of the error coefficients is crucial for identifying initial scheme parameters that simultaneously improve both performance and security.

We analyze the circuit depicted in Figure 3 in which pairs of ciphertexts are progressively multiplied. Note that we focus on circuits that involve homomorphic multiplication since it is the most complex and, therefore, the most significant operation to study. It is worth noting that the circuit described in this work closely resembles the one adopted by default in OpenFHE [1], where multiplications are always preceded by a modulus switching operation.

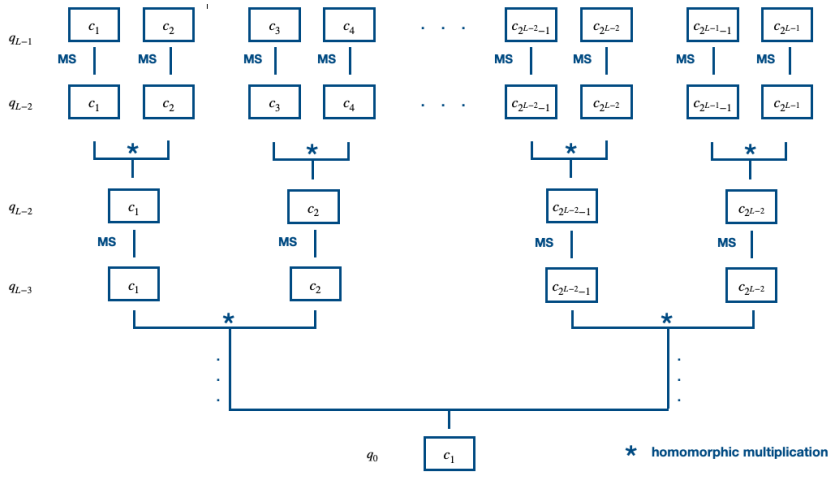


Fig. 3: Reference circuit

Let L denote the number of levels in the circuit, and $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_{2^{L-1}}$ the initial fresh ciphertexts, generated by encrypting 2^{L-1} independent and randomly generated messages using the same key.

We consider a model where, at each level, homomorphic multiplication of pairs of ciphertexts is carried out. At the beginning of each level $\ell \in \{1, \ldots, L-1\}$ the input ciphertexts are switched to a smaller modulus, to maintain the error almost constant throughout the process. We recall that, q_ℓ's are the ciphertext moduli of each level $L-1-\ell$, defined as $q_\ell = \prod_{j=0}^{\ell} p_j$, where p_j are primes such that $\gcd(p_i, p_j) = 1$ for $i \neq j$.

We denote by ν_ℓ^{ms} the critical quantity obtained after performing the modulus switching at the beginning of level ℓ, and by ν_ℓ the critical quantity at the end of level ℓ, namely just after the multiplication (and key switching technique). Similarly, V_ℓ^{ms} and V_ℓ denote the variance of the coefficients of these two critical quantities, respectively. Finally, for each level ℓ we will use the notation V_{ms} to denote $\mathsf{Var}(\delta_1 s/p_\ell|_i)$.

The process represented by the circuit can be summarized as follows:

- At level 0, all the fresh ciphertexts are defined in $\mathcal{R}^2_{q_{L-1}}$.
- At level 1, modulo switch to q_{L-2} is applied to each of the initial fresh ciphertexts. After that, the first homomorphic multiplication is performed in $\mathcal{R}_{q_{L-2}}$, yielding 2^{L-2} resulting ciphertexts in $\mathcal{R}^3_{q_{L-2}}$. Finally, these ciphertexts are relinearized to obtain the equivalent ones in $\mathcal{R}^2_{q_{L-2}}$, which are the inputs of the next level.
- This process is repeated until level $L-1$, where a final ciphertext, which is the output of the circuit, is returned.

It should be noted that, in the circuit represented in Fig. 3, the relinearization step is not presented. In fact, we decided to omit its contribution from our error analysis, since it is negligible compared to the impact of modulus switch and multiplication. A detailed justification for this choice is provided in Appendix C.

Level 0. Let $\boldsymbol{c} \in \mathcal{R}^2_{q_{L-1}}$ be a fresh ciphertext, obtained encrypting a random message $m \in \mathcal{R}_t$. Let $\nu_{\mathsf{clean}} = a_0 + a_1 s$ be the critical quantity associated with $\boldsymbol{c}$ and let V_0 be the variance of ν_{clean}. Then, we have

$$\begin{cases} V(a_0|_i) = t^2(\frac{1}{12} + V_e + nV_eV_u) \\ V(a_1 s|_i) = t^2 nV_eV_s. \end{cases}$$

Therefore, it is possible to assume

$$\mathsf{Var}(a_0|_i) \approx \mathsf{Var}((a_1 s)|_i) \approx \frac{V_0}{2}. \tag{12}$$

It should be noted that the assumption in (12) applies to all initial 2^{L-1} ciphertexts. In fact, the estimate of variance V_ℓ of the coefficients of the critical

quantity at the end of level ℓ, is the same for all ciphertexts belonging to the same level.

Level 1. At the beginning of this level, all ciphertexts are subject to a modulus switching from q_{L-1} to q_{L-2}. Therefore, the error and the modulus of $\boldsymbol{c}$ are rescaled by a factor $q_{L-2}/q_{L-1} = 1/p_{L-1}$. Thus, the critical quantity associated with the resulting ciphertext is

$$\nu_1^{\mathsf{ms}} = \frac{\nu_{\mathsf{clean}} + \delta_0 + \delta_1 s}{p_{L-1}}, \quad \delta_i = t[-c_i t^{-1}]_{p_{L-1}} \tag{13}$$

Since $\delta_i/p_{L-1} \leftarrow \mathcal{U}_t$, it follows that $\mathsf{Var}(\delta_0/p_{L-1}|_i) = t^2/12$ and $\mathsf{Var}(\delta_1 s/p_{L-1}|_i) = t^2 n V_s/12$. In BGV, n is typically larger than 2^{12}, making the contribution of δ_0/p_{L-1} negligible.

The critical quantity in (13) can be written as

$$\nu_1^{\mathsf{ms}} \approx \frac{a_0}{p_{L-1}} + \left(\frac{a_1}{p_{L-1}} + \frac{\delta_1}{p_{L-1}}\right) s,$$

By Equation (12) we have

$$\mathsf{Var}\left(\frac{a_0}{p_{L-1}}\bigg|_i\right) = \mathsf{Var}\left(\frac{a_1}{p_{L-1}} s\bigg|_i\right) = \frac{V_0}{2p_{L-1}^2}.$$

Therefore, the variance after the modulus switch can be written as

$$V_1^{\mathsf{ms}} = \frac{V_0}{p_{L-1}^2} + V_{\mathsf{ms}}.$$

As discussed previously, to ensure that the error follows a Gaussian distribution, we rely on Equation (11), i.e., we can assume $V_0/p_{L-1}^2 < V_{\mathsf{ms}}/100$.

In order to estimate the variance of the critical quantity resulting from multiplication, we rely on Theorem 2. In particular, we avoid explicitly writing the terms $b_\mu(\iota)$ in the expression $a_\iota = \sum_{\mu=0}^{K} b_\mu(\iota) e^\mu$ to keep the notation manageable. However, it is crucial to keep track of the highest power of e, which is K, appearing in the coefficients a_ι involved in the multiplications, since the correction applied by the function F depends on this value.

Precisely, given the fresh ciphertexts, it can be seen that for a_0 the associated K is 1, while for a_1 we have $K = 0$. Instead, for the terms $\delta_0/p_\ell, \delta_1/p_\ell$ the case is quite different. In fact, these values can be assumed to be randomly distributed over $\mathcal{R}_t$, for every level of the circuit. This makes the analysis slightly more complicated, since, in order to derive tight estimates, it will be necessary to distinguish the contribution of a_0, a_1 from the one of δ_0, δ_1.

Now, assume that $\boldsymbol{c}, \boldsymbol{c}' \in \mathcal{R}_{q_{L-2}}^2$ are two ciphertexts at level 1, after modulus switching has been carried out. Their associated critical quantity is given, re-

spectively, by

$$\begin{cases} \nu_1^{\mathsf{ms}} = \frac{1}{p_{L-1}}(a_0 + a_1 s) + \frac{\delta_1}{p_{L-1}} s \\ \nu_1^{\mathsf{ms}'} = \frac{1}{p_{L-1}}(a_0' + a_1' s) + \frac{\delta_1'}{p_{L-1}} s \end{cases}$$

Therefore, the critical quantity obtained after their multiplication is of the form

$$\begin{aligned} \nu_1 = \frac{1}{p_{L-1}^2} a_0 a_0' + \frac{1}{p_{L-1}^2}\left(a_0' a_1 + a_0 a_1'\right) s + \frac{1}{p_{L-1}}(a_0' \frac{\delta_1}{p_{L-1}} + a_0 \frac{\delta_1'}{p_{L-1}}) s \\ + \frac{1}{p_{L-1}^2} a_1 a_1' s^2 + \frac{1}{p_{L-1}}(a_1 \frac{\delta_1'}{p_{L-1}} + a_1' \frac{\delta_1}{p_{L-1}}) s^2 + \frac{\delta_1'}{p_{L-1}} \frac{\delta_1}{p_{L-1}} s^2. \end{aligned}$$

So, we are now able to provide an estimate of the variance V_1, applying Theorem 2, as follows

$$\begin{aligned} V_1 \leq & \frac{n}{p_{L-1}^4} V(a_0|_i) V(a_0'|_i) F_e(1,1) + \frac{n}{p_{L-1}^4} V(a_0'|_i) V(a_1 s|_i) + \frac{n}{p_{L-1}^4} V(a_0|_i) V(a_1' s|_i) \\ & + \frac{n}{p_{L-1}^2} V(a_0'|_i) V(\frac{\delta_1}{p_{L-1}} s|_i) + \frac{n}{p_{L-1}^2} V(a_0|_i) V(\frac{\delta_1'}{p_{L-1}} s|_i) \\ & + \frac{n}{p_{L-1}^4} V(a_1 s|_i) V(a_1' s|_i) F_s(1,1) + \frac{n}{p_{L-1}^2} V(a_1 s|_i) V(\frac{\delta_1'}{p_{L-1}} s|_i) F_s(1,1) \\ & + \frac{n}{p_{L-1}^2} V(a_1' s|_i) V(\frac{\delta_1}{p_{L-1}} s|_i) F_s(1,1) + n V(\frac{\delta_1}{p_{L-1}} s|_i) V(\frac{\delta_1'}{p_{L-1}} s|_i) F_s(1,1). \end{aligned}$$

which, recalling that $\mathrm{Var}(a_\iota|_i) = \mathrm{Var}(a_\iota'|_i) = V_0/2$ and that $V(\frac{\delta_1}{p_{L-1}} s|_i) = V(\frac{\delta_1'}{p_{L-1}} s|_i) = V_{\mathsf{ms}}$ can be rewritten as

$$\begin{aligned} V_1 \leq & \frac{n}{p_{L-1}^4} \frac{V_0^2}{4} F_e(1,1) + \frac{2n}{p_{L-1}^4} \frac{V_0^2}{4} + \frac{2n}{p_{L-1}^2} V_{\mathsf{ms}} \frac{V_0}{2} \\ & + \frac{n}{p_{L-1}^4} \frac{V_0^2}{4} F_s(1,1) + \frac{2n}{p_{L-1}^2} V_{\mathsf{ms}} \frac{V_0}{2} F_s(1,1) + n F_s(1,1) V_{\mathsf{ms}}^2. \end{aligned}$$

By observing that the condition in Equation (11) holds, our bound on V_1 is derived as follows

$$V_1 \approx \left(\frac{3}{2 \cdot 100^2} + \frac{3}{100} + 2\right) n V_{\mathsf{ms}}^2 \approx 2n V_{\mathsf{ms}}^2,$$

and we can simplify $\nu_1 = \delta_1 \delta_1' s^2 / p_{L-1}$.

We now present the generic level ℓ. We begin by illustrating the case $\ell = 2$. As we shall see, the input to level 2 corresponds to that of a generic level $\ell \geq 2$, and it produces the same type of output, thus making the analysis carried out for the second level applicable to all subsequent levels.

Level 2. At the beginning of level 2 modulo switch is performed, yielding a critical quantity of the form $\nu_2^{\mathsf{ms}} = \nu_1/p_{L-2} + \delta_1 s/p_{L-2}$, with coefficient variance,

$$V_2^{\mathsf{ms}} = \frac{V_1}{p_{L-2}^2} + V_{\mathsf{ms}} \leq \left(\frac{1}{100} + 1\right) V_{\mathsf{ms}}\,.$$

Now, the product of two critical quantities ν_1^{ms} can be expanded into four terms: the first is the product of two ν_1/p_{L-2} terms, the second and third are the products of a ν_1/p_{L-2} term with a ν_1^{ms} term and the fourth is the product of two ν_1^{ms} terms. Applying our theorem, the variance after the multiplication is then given by

$$V_2 = n\left(\frac{V_1^2}{p_{L-2}^4}F_s(2,2) + 2\frac{V_1}{p_{L-2}^2}V_{\mathsf{ms}}F_s(2,1) + 2V_{\mathsf{ms}}^2\right),$$

which can be bounded, for the conditions over p_{L-2}, as

$$V_2 \leq \left(\frac{6}{100^2} + \frac{6}{100} + 2\right) nV_{\mathsf{ms}}^2 \approx 2nV_{\mathsf{ms}}^2\,.$$

Again, the critical quantity at the end of level 2 can be simplified as $\nu_2 = \delta_1\delta_1' s^2/p_{L-2}$. From this observation, we can deduce the variance estimates for each level ℓ.

Level ℓ. For any level $\ell \geq 2$, the variance after modulus switching V_ℓ^{ms} and the variance after multiplication V_ℓ can be approximated as

$$\begin{aligned} V_\ell^{\mathsf{ms}} &\approx \frac{101}{100}V_{\mathsf{ms}} \\ V_\ell &\approx (2+\epsilon)nV_{\mathsf{ms}}^2, \quad \epsilon = \frac{6}{100^2} + \frac{6}{100} \end{aligned} \tag{14}$$

Table 1 presents a comparison of the estimated variances of the error coefficients obtained using our approach, denoted as *our* and the corresponding experimental values, denoted as *exp*. The estimates were obtained using the OpenFHE library, setting the parameters according to the required multiplicative depth, with $t = 65537$. For the experimental values, 50000 samples were computed for $n = 2^{13}$, 8000 for $n = 2^{14}$ and 5000 for $n = 2^{15}$. To enhance readability, variances are presented in terms of their base-2 logarithms.

Remark 3 *At first glance, it might seem that, in order to correctly estimate the variance of the multiplication, it is sufficient to apply a factor of 2, arising from the correction of the product between two s-terms, thus making it unclear why the dependence on the public key should be considered. It is crucial to emphasize that these estimates are strictly related to a circuit model that we are specifically constructing to guarantee the Gaussianity of the error coefficients according to*

	Encryption		Modulo Switch		1 Multiplication		6 Multiplications	
n	*our*	*exp*	*our*	*exp*	*our*	*exp*	*our*	*exp*
2^{13}	48.76	48.76	40.84	40.83	95.68	95.65	95.71	95.25
2^{14}	49.76	49.76	41.84	41.79	98.68	97.62	98.71	97.63
2^{15}	50.76	50.72	42.84	42.82	101.68	101.62	101.71	101.59

Table 1: Encryption, modulo switch, and multiplication of fresh ciphertexts. The last two columns report the case of six consecutive multiplications as in the reference circuit.

Equation (11). *However, if one wanted to accurately study the variance in different circuits, where the term* V_{ms} *is not the only significant contribution, the need to account for these dependencies would become even more evident. This necessity reaches its maximum importance when considering circuits in which modulus switching is not applied. Since the use of modulus switching is mandatory for the practical implementation of BGV, in this work we omitted the results obtained from the analysis of such circuits. However, whether modulus switching is not used, or it is applied with primes that do not satisfy Equation* (11)*, it remains crucial to take into account all dependencies arising from both* e *and* s *as well as from their powers. As an example of the applicability of our approach, if error coefficient distributions that are non-Gaussian, but still suitable for use, were considered, the variance analysis we propose would remain valid and could still provide accurate estimates that never underestimate the experimental values (see Theorem 2).*

7 Comparison with Previous Works

When estimating the error in schemes such as BGV, existing approaches can broadly be divided into two main categories: *worst-case* and *average-case*. The former includes analyses based on the Euclidean norm, as in [7], the infinity norm, as in [22,30], and the canonical norm, as in [15,17,26,28,33], with the latter providing the tightest bounds among the worst-case analyses. In contrast, average-case approaches model the noise coefficients as random variables and focus on their mean and variance to derive probabilistic bounds on the minimum ciphertext modulus required to guarantee correctness.

Although the latter approaches appear more promising in reducing the ciphertext modulus size, research for a general method to estimate the variance remains incomplete. Most works, such as [34], treat the error coefficients as independent, which, as pointed out in [4,16,34], leads to *imprecise bounds* and often underestimates the modulus size required to prevent decryption failures. Consequently, the scheme may be exposed to potential vulnerabilities, including key-recovery attacks [10]. Other types of methods, by contrast, avoid such underestimations or

failures but are limited to specific settings, as in [18], where the authors analyze the error for HElib [36].

The objective of our work is to develop an average case approach that takes into account the intrinsic dependencies introduced by the secret and public key, allowing to obtain variance estimates very close to experimental values, without ever underestimating them. At the same time, our aim is to propose a method independent of the specific library or circuit considered, providing the theoretical basis necessary to build *ad hoc* estimates depending on the circuit and library of interest. Although, to highlight the effectiveness of our approach, we have focused on circuits that exclusively include multiplications preceded by the modulus switch, our analyses and estimates remain valid for any type of homomorphic circuit. Our method is in fact intrinsically modular: each homomorphic operation is analyzed independently, producing variance bounds that can be composed to precisely estimate the overall noise growth of a given circuit, assuming only that the multiplications are preceded by a modulus switch – a completely reasonable hypothesis in the context of BGV typical usage.

In this section, we demonstrate the efficacy of our average-case approach by comparing it to the state-of-the-art works. Specifically, we illustrate how our average-case analysis provides tighter and more practical bounds than traditional worst-case methods [33], and we further compare it with other average-case approaches [34,18]. We then show how our choice of ciphertext moduli achieves a substantial size reduction compared to those adopted in widely used libraries such as OpenFHE [1] and HElib [36].

Canonical norm. For the comparison with the worst-case approach, we specifically refer to the estimates proposed in [33].

Homomorphic operation	Error bounds with canonical norm
`Enc`	$\|\|\nu_{\mathsf{clean}}\|\|^{\mathrm{can}} \leq Dt\sqrt{n\,(1/12 + 2nV_eV_s + V_e)}$
`Mod Switch`(q')	$\|\|\nu + \nu_{\mathsf{ms}}(q')\|\|^{\mathrm{can}} \leq \frac{1}{p_\ell}\|\|\nu\|\|^{can} + Dt\sqrt{n(1/12 + nV_s)}$
`Mult`$(\mathbf{c}, \mathbf{c}')$	$\|\|\nu_{\mathrm{mul}}\|\|^{\mathrm{can}} \leq \|\|\nu\|\|^{\mathrm{can}}\|\|\nu'\|\|^{\mathrm{can}}$

Table 2: Canonical norm depending on the homomorphic operations.

One of the key distinctions between worst-case and average-case analyses lies not only in how the noise norm is bounded, but also in how these bounds propagate through homomorphic multiplication. In the worst-case analysis, the canonical norm is bounded as $\|\nu\|^{\mathrm{can}} \leq D\sqrt{nV}$, which holds with probability at least $1 - ne^{-D^2}$, as established in Equation (3). Thus, after one multiplication, we have (by Table 2) that the bound is $||\nu_{\mathrm{mul}}||^{\mathrm{can}} \leq ||\nu||^{\mathrm{can}}||\nu'||^{\mathrm{can}} \leq D^2 n\sqrt{VV'}$.

Supposing that the ciphertexts are both multiplied just after the modulus switch, the bound becomes $||\nu_{\text{mul}}||^{\text{can}} \leq D^2 n V_{\mathsf{ms}}$.

In contrast, average-case approaches, such as the one proposed in this work, allow significantly tighter bounds by imposing the condition $\|\nu\|_\infty \leq D\sqrt{2V}$, where V denotes the variance of each coefficient of ν. According to Equation (2), this bound holds with probability at least $1 - n(1 - \text{erf}(D))$, which, for $D = 6$, exceeds $1 - 2^{-40}$. However, in practical scenarios, choosing $D = 8$ is preferable, as it limits the failure probability to 2^{-77}, when $n \leq 2^{15}$. Note that this bound better captures the actual distribution of the error and results in much tighter predictions after multiplications. In fact, by Equation (14), after a multiplication that follows modulo switch, we have $||\nu_{\text{mul}}||_\infty \leq 2D\sqrt{n}V_{\mathsf{ms}}$.

Current Average-Case Approaches. Within average-case analysis frameworks, the main novelty of our approach lies in tracking the dependencies between the error coefficients, particularly those arising from the multiplication of two ciphertexts. Specifically, given the critical quantities ν and ν' of two ciphertexts with coefficient variances V and V', respectively, the current average-case methodology [34] yields the expression $\mathsf{Var}(\nu\nu'|_i) \approx nVV'$. Therefore, assuming that two ciphertexts have just undergone modulus switching, the variance resulting from their product is, in accordance with the previously established considerations

$$\mathsf{Var}(\nu\nu'|_i) \approx nV_{\mathsf{ms}}^2 .$$

In contrast, our method estimates this value according to Equation (14) , resulting in

$$\mathsf{Var}(\nu\nu'|_i) \approx 2nV_{\mathsf{ms}}^2 .$$

The factor 2, which arises from accounting for the dependencies among the critical quantities, allows for extremely accurate values, as shown in Table 1.

From this comparison, it should be evident that accounting for the dependencies in the coefficients of the error polynomial is fundamental for obtaining accurate and correct estimates of the experimental variance.
In particular, we believe that this consideration is precisely what overcomes the underestimation inherent in the approach presented in [34].

It is crucial to point out that the factor of 2 is specific to the type of circuit we constructed, in which the variance is independent of the circuit level. However, when analyzing errors in circuits where this condition is not met, it is still possible to derive general upper bounds using Theorem 2, once again obtaining estimates that never underestimate the error. Nevertheless, for the reasons outlined above, including guaranteeing the Gaussianity of the error, we recommend choosing primes that allow the variance of the error after the modulus switch to be well approximated by V_{ms}, as explained previously.

Our results appear to be very close to those reported in [18]. In particular, it can be observed that a factor of 2 also appears in [18, Lemma 9].

Current Libraries. Finally we compare the ciphertext modulus q as estimated by our method against those adopted by two of the most widely used libraries: OpenFHE [1] and HElib [36]. We recall that the ciphertext modulus must be selected to ensure that $\|\nu\| < \frac{q}{2}$. In our selection, we additionally require (11) to guarantee the Gaussian behavior of the error.

OpenFHE. In OpenFHE [1], the generation of the moduli depends on the multiplicative depth selected as input, in a manner similar to the choice of moduli described in this paper. Moreover, the way multiplications are executed in a fixed-depth circuit of the library is consistent with the circuits analyzed in our work, where each multiplication between two ciphertexts is preceded by a modulus switching operation. For a circuit of multiplicative depth M, $M+2$ moduli are generated, to allow for an additional modulus switching after the final multiplication. Therefore, we generate the moduli in a way consistent with the library's design, to allow a fair comparison between the modulus sizes used in OpenFHE and those proposed in this paper.

In Figure 4, we compare the ciphertext modulus sizes q for circuits with multiplicative depths 3 (Figure 4a) and 6 (Figure 4b). The values are obtained using our method (denoted as *our*) and those generated by OpenFHE (denoted as *OpenFHE*). Note that, for the generation of our parameters, we fixed $D = 8$.

n	2^{12}	2^{13}	2^{14}	2^{15}
OpenFHE	147.3	151.8	156.3	161.6
our	121.0	124.5	128.0	131.5

(a) Circuit of depth 3.

n	2^{12}	2^{13}	2^{14}	2^{15}
OpenFHE	249.3	256.8	264.3	272.6
our	210.3	216.8	223.3	229.8

(b) Circuit of depth 6.

Fig. 4: Comparison of $\log_2(q)$ in the circuit shown in Figure 3 (setting $D = 8$).

From this comparison, it should be evident that our approach can lead to an effective reduction in the modulus size, resulting in corresponding improvements in the overall efficiency of the scheme.

HElib Comparison. To provide a more comprehensive overview of the effectiveness of our work and to highlight its applicability across various contexts and libraries, we also present a comparison between our choice of moduli and those used in HElib [27], currently one of the most competitive FHE libraries, together with OpenFHE. We first point out that a strategy for optimizing HElib specific parameters has been analyzed in [18]. However, a direct comparison with that approach lies beyond the scope of this work. Our focus is instead on emphasizing that the strength of our method lies in its library-independent nature, offering a general framework for accurately estimating noise in different types of circuits and determining the conditions that ciphertext moduli must meet to ensure both

efficiency and security (as well as the Gaussianity of the noise distributions, see Section 5). For this reason, we intend to provide in this section just an idea of how our method can potentially improve current HElib parameters as well.

Our method proposes to select parameters based on the multiplicative depth of the circuit, as in OpenFHE and in previous versions of HElib. However, in the new version of the library, the selection of parameters is based on the bit length of the largest modulus in the chain. Therefore, to compare the two approaches, we focus on comparing the ratio between the moduli of successive levels, as also done in [18].

Due to the way ciphertext moduli are constructed in HElib, the ratio between two moduli of adjacent levels, i.e. a prime p_ℓ, is necessarily larger than 36 bits. However, as observed in [18], it is typically much larger, often reaching sizes of around 54 bits, for a ring dimension $n \in \{2^{12}, 2^{13}, 2^{14}, 2^{15}\}$. Instead, with our method, we simply require that Equation (11) is satisfied, i.e., for $0 < l < L - 1$

$$p_\ell > \sqrt{\frac{(2+\epsilon)nV_{\mathsf{ms}}}{\alpha}}\,.$$

Table 3 reports the bit-size of the ratio between adjacent moduli resulting from our approach, for ring dimensions in $\{2^{12}, 2^{13}, 2^{14}, 2^{15}\}$, assuming a secret key with coefficients sampled from a ternary uniform distribution, with hamming weight $h = n/2$, plaintext modulus $t = 65537$ and $\alpha = 1/100$. As shown in the table, our approach achieves a reduction of up to 6 bits in the ratio between consecutive ciphertext moduli.

n	2^{12}	2^{13}	2^{14}	2^{15}
$\log_2(p_\ell)$	29.55	30.55	31.55	32.55

Table 3: Ratio between adjacent ciphertext moduli for different ring dimensions n, according to our approach, represented using their bit size.

8 Conclusions

This work presents a novel application of the average-case approach within the BGV scheme to accurately estimate errors arising from homomorphic operations, with particular emphasis on ciphertext multiplication. The main contributions are twofold.

First, it is demonstrated that the common independence assumption on ciphertext errors does not hold in the BGV scheme, even when ciphertexts are generated from independent messages. This arises because errors share common powers of the public and secret keys. To address this challenge, we developed

a novel variance analysis that explicitly incorporates these dependencies. This approach yields tight upper bounds that closely align with experimental observations and provably avoids the underestimation issues present in previous methods.

Second, we investigate the distribution of the error coefficients and clarify the role of the Gaussianity assumption. Although Gaussianity is not generally guaranteed, we show that, with suitable parameter choices and systematic application of modulus switching, the error distribution is well approximated by a discrete Gaussian. This provides a theoretical justification for the use of Gaussian-based average-case analyses in practical settings. Moreover, we show that, in this setting, the error no longer depends on the circuit's multiplication depth. From this observation, we derive a very simple bound on the error norm, leading to significantly simpler and more scalable approach, which is crucial for the widespread adoption in real-world settings.

Finally, we derive a new parameter selection methodology that yields substantial improvements over the parameter choices currently employed in state-of-the-art FHE libraries.

Acknowledgment

The third and fourth authors were partially supported by the Italian Ministry of University and Research in the framework of the Call for Proposals for scrolling of final rankings of the PRIN 2022 call - Protocol no. 2022RFAZCJ. The third author acknowledges project SERICS (PE00000014) under the MUR National Recovery and Resilience Plan funded by the European Union - NextGenerationEU.

References

1. Al Badawi, A., Bates, J., Bergamaschi, F., Cousins, D.B., Erabelli, S., Genise, N., Halevi, S., Hunt, H., Kim, A., Lee, Y., et al.: OpenFHE: Open-Source Fully Homomorphic Encryption Library. In: Proceedings of the 10th Workshop on Encrypted Computing & Applied Homomorphic Cryptography. pp. 53–63 (2022)
2. Anderson, T.W., Darling, D.A.: A test of goodness of fit. Journal of the American statistical association **49**(268), 765–769 (1954)
3. Bai, S., D. Galbraith, S.: Lattice decoding attacks on binary LWE. In: Australasian Conference on Information Security and Privacy. pp. 322–337. Springer, Cham (2014)
4. Biasioli, B., Marcolla, C., Calderini, M., Mono, J.: Improving and automating BFV parameters selection: An average-case approach. Cryptology ePrint Archive, Paper 2023/600 (2023), https://eprint.iacr.org/2023/600
5. Bossuat, J., Costache, A., Mouchet, C., Nürnberger, L., Troncoso-Pastoriza, J.R.: Accurate and composable noise estimates for CKKS with application to exact HE computation. IACR Commun. Cryptol. **2**(2), 8 (2025)

6. Brakerski, Z.: Fully Homomorphic Encryption without Modulus Switching from Classical GapSVP. In: Advances in Cryptology – CRYPTO 2012. pp. 868–886 (2012)
7. Brakerski, Z., Gentry, C., Vaikuntanathan, V.: (Leveled) fully homomorphic encryption without bootstrapping. ACM Transactions on Computation Theory (TOCT) **6**(3), 1–36 (2014)
8. Brakerski, Z., Vaikuntanathan, V.: Fully Homomorphic Encryption from Ring-LWE and Security for Key Dependent Messages. In: Advances in Cryptology – CRYPTO 2011. pp. 505–524 (2011)
9. Checri, M., Sirdey, R., Boudguiga, A., Bultel, J.P., Choffrut, A.: On the practical cpad security of "exact" and threshold fhe schemes and libraries. Cryptology ePrint Archive (2024)
10. Cheon, J.H., Choe, H., Passelègue, A., Stehlé, D., Suvanto, E.: Attacks against the indcpa-d security of exact fhe schemes. Cryptology ePrint Archive (2024)
11. Cheon, J.H., Han, K., Kim, A., Kim, M., Song, Y.: A full RNS variant of approximate homomorphic encryption. In: International Conference on Selected Areas in Cryptography – SAC 2018. pp. 347–368. Springer (2018)
12. Cheon, J.H., Kim, A., Kim, M., Song, Y.: Homomorphic Encryption for Arithmetic of Approximate Numbers. In: Advances in Cryptology – ASIACRYPT 2017. pp. 409–437 (2017)
13. Chillotti, I., Gama, N., Georgieva, M., Izabachène, M.: Faster fully homomorphic encryption: Bootstrapping in less than 0.1 seconds. In: Advances in Cryptology – ASIACRYPT 2016. pp. 3–33 (2016)
14. Chillotti, I., Gama, N., Georgieva, M., Izabachène, M.: TFHE: Fast Fully Homomorphic Encryption over the Torus. Journal of Cryptology **33**(1), 34–91 (2020)
15. Costache, A., Smart, N.P.: Which ring based somewhat homomorphic encryption scheme is best? In: Topics in Cryptology – CT-RSA 2016. pp. 325–340 (2016)
16. Costache, A., Curtis, B.R., Hales, E., Murphy, S., Ogilvie, T., Player, R.: On the precision loss in approximate homomorphic encryption. In: International Conference on Selected Areas in Cryptography - SAC 2023. pp. 325–345. Springer (2023)
17. Costache, A., Laine, K., Player, R.: Evaluating the effectiveness of heuristic worst-case noise analysis in FHE. In: Computer Security – ESORICS 2020. pp. 546–565 (2020)
18. Costache, A., Nürnberger, L., Player, R.: Optimisations and Tradeoffs for HElib. In: Topics in Cryptology – CT-RSA 2023. pp. 29–53 (2023)
19. Damgård, I., Pastro, V., Smart, N., Zakarias, S.: Multiparty computation from somewhat homomorphic encryption. In: Annual Cryptology Conference. pp. 643–662. Springer (2012)
20. Di Giusto, A., Marcolla, C.: Breaking the power-of-two barrier: noise estimation for bgv in ntt-friendly rings. Designs, Codes and Cryptography **93**(3), 467–502 (2025)
21. Ducas, L., Micciancio, D.: FHEW: Bootstrapping Homomorphic Encryption in Less Than a Second. In: Oswald, E., Fischlin, M. (eds.) Advances in Cryptology – EUROCRYPT 2015. pp. 617–640. Springer Berlin Heidelberg, Berlin, Heidelberg (2015)
22. Fan, J., Vercauteren, F.: Somewhat practical fully homomorphic encryption. ePrint Archive (2012)
23. Feller, W.: An Introduction to Probability Theory and Its Applications. Vol. 1. Wiley (1950)

24. Gao, M., Zheng, H.: A critique on average-case noise analysis in rlwe-based homomorphic encryption. In: Proceedings of the 13th Workshop on Encrypted Computing & Applied Homomorphic Computing – WAHC 2025. pp. 26–37 (2025)
25. Gentry, C.: A fully homomorphic encryption scheme. Stanford university (2009)
26. Gentry, C., Halevi, S., Smart, N.P.: Homomorphic Evaluation of the AES Circuit. In: Advances in Cryptology – CRYPTO 2012. pp. 850–867 (2012)
27. Halevi, S., Shoup, V.: Design and implementation of HElib: a homomorphic encryption library. ePrint Archive (2020)
28. Iliashenko, I.: Optimisations of fully homomorphic encryption. PhD thesis (2019)
29. Kim, A., Polyakov, Y., Zucca, V.: Revisiting homomorphic encryption schemes for finite fields. In: Advances in Cryptology–ASIACRYPT 2021. pp. 608–639 (2021)
30. Kim, A., Polyakov, Y., Zucca, V.: Revisiting homomorphic encryption schemes for finite fields. Cryptology ePrint Archive, Paper 2021/204 (2021), https://eprint.iacr.org/2021/204
31. Loéve, M.: Probability Theory I. Springer (1977)
32. Lyubashevsky, V., Peikert, C., Regev, O.: On ideal lattices and learning with errors over rings. In: Advances in Cryptology – EUROCRYPT 2010. pp. 1–23 (2010)
33. Mono, J., Marcolla, C., Land, G., Güneysu, T., Aaraj, N.: Finding and evaluating parameters for BGV. In: International Conference on Cryptology in Africa. pp. 370–394 (2023)
34. Murphy, S., Player, R.: A central limit approach for ring-LWE noise analysis. IACR Communications in Cryptology **1**(2) (2024)
35. Regev, O.: On lattices, learning with errors, random linear codes, and cryptography. Journal of the ACM (JACM) **56**(6), 1–40 (2009)
36. Shai Halevi and Victor Shoup: HElib. https://github.com/homenc/HElib
37. Stephens, M.A.: Tests based on edf statistics. In: Goodness-of-fit-techniques, pp. 97–194. Routledge (2017)

A Proof of Lemma 1

In order to prove this lemma, we will first demonstrate that these properties hold for the critical quantity of a fresh ciphertext ν_{clean}.
Then, we will show that any operation involved in the BGV circuit does not affect these properties.

Fresh ciphertexts For ν_{clean}, the coefficients $b_\iota(\mu)$ are defined as

$$\begin{cases} b_0(0) = m + te_0 \\ b_1(0) = tu \end{cases} \qquad \begin{cases} b_0(1) = te_1 \end{cases}$$

Therefore, the first property follows immediately from the independence of $b_{\mu_1}(\iota_1)|_{j_1}, b_{\mu_2}(\iota_2)|_{j_2}$ when $\mu_1 \neq \mu_2$ or $j_1 \neq j_2$.

As for the second property, it holds since
- $\mathbb{E}[b_0(0)|_i] = \mathbb{E}[m|_i] + t\mathbb{E}[e_0|_i] = 0$, due to the linearity of the expected value and the distributions considered, i.e. $m \leftarrow \mathcal{U}_t, e_0 \leftarrow \mathcal{DG}_q(\sigma^2)$;

- $\mathbb{E}[b_1(0)|_i] = t\mathbb{E}[u|_i] = 0$, as $u \leftarrow \chi_s$;
- $\mathbb{E}[b_0(1)|_i] = t\mathbb{E}[e_1|_i] = 0$, as $e_1 \leftarrow \mathcal{DG}_q(\sigma^2)$;

We will therefore show that the remaining homomorphic operations do not alter these properties.

Let $\nu = \sum_{\iota_1}\sum_{\mu_1} b_{\mu_1}(\iota_1)e^{\mu_1}s^{\iota_1}$, $\nu' = \sum_{\iota_2}\sum_{\mu_2} b'_{\mu_2}(\iota_2)e^{\mu_2}s^{\iota_2}$ be the respective critical quantities of two generic ciphertexts, for which the properties stated above are assumed to hold.

Addition of two ciphertexts The critical quantity after the addition of the two BGV ciphertexts is given by

$$\nu_{\text{add}} = \nu + \nu' = \sum_{\iota}\sum_{\mu} b^{\text{add}}_{\mu}(\iota)e^{\mu}s^{\iota},$$

where $b^{\text{add}}_{\mu}(\iota) = b_{\mu}(\iota) + b'_{\mu}(\iota)$.

Therefore, if $\mathbb{E}[b_{\mu}(\iota)|_i] = \mathbb{E}[b'_{\mu}(\iota)|_i] = 0$ then

$$\mathbb{E}[b^{\text{add}}_{\mu}(\iota)|_i] = 0,$$

according to the linearity of the expected value.

Moreover, using the bilinearity of the covariance, we have that, for $\mu_1 \neq \mu_2$ or $j_1 \neq j_2$

$$\mathsf{Cov}(b^{\text{add}}_{\mu_1}(\iota_1)|_{j_1}, b^{\text{add}}_{\mu_2}(\iota_2)|_{j_2}) = 0,$$

since

$$\begin{aligned}\mathsf{Cov}(b_{\mu_1}(\iota_1) + b'_{\mu_1}(\iota_1)|_{j_1}, b_{\mu_2}(\iota_2) + b'_{\mu_2}(\iota_2)|_{j_2}) = {} & \mathsf{Cov}(b_{\mu_1}(\iota_1)|_{j_1}, b_{\mu_2}(\iota_2)|_{j_2}) \\ & + \mathsf{Cov}(b_{\mu_1}(\iota_1)|_{j_1}, b'_{\mu_2}(\iota_2)|_{j_2}) \\ & + \mathsf{Cov}(b'_{\mu_1}(\iota_1)|_{j_1}, b_{\mu_2}(\iota_2)|_{j_2}) \\ & + \mathsf{Cov}(b'_{\mu_1}(\iota_1)|_{j_1}, b'_{\mu_2}(\iota_2)|_{j_2}),\end{aligned}$$

where all the summands vanish because:

- $\mathsf{Cov}(b_{\mu_1}(\iota_1)|_{j_1}, b_{\mu_2}(\iota_2)|_{j_2}) = \mathsf{Cov}(b'_{\mu_1}(\iota_1)|_{j_1}, b'_{\mu_2}(\iota_2)|_{j_2}) = 0$ holds by assumption for $\mu_1 \neq \mu_2$ or $j_1 \neq j_2$;
- $\mathsf{Cov}(b_{\mu_1}(\iota_1)|_{j_1}, b'_{\mu_2}(\iota_2)|_{j_2}) = \mathsf{Cov}(b'_{\mu_1}(\iota_1)|_{j_1}, b_{\mu_2}(\iota_2)|_{j_2}) = 0$ since $b_{\mu_1}(\iota_1)$ and $b'_{\mu_2}(\iota_2)$ are independent $\forall \mu_1, \mu_2$;

Multiplication by a constant Given a ciphertext with $\nu = \sum_{\iota}\sum_{\mu} b_{\mu}(\iota)e^{\mu}s^{\iota}$ and a constant α, the critical quantity obtained after their homomorphic multiplication, according to the BGV scheme, is as follows

$$\nu_{\text{const}} = \alpha\nu = \alpha\sum_{\iota}\sum_{\mu} b_{\mu}(\iota)e^{\mu}s^{\iota} = \sum_{\iota}\sum_{\mu} \alpha b_{\mu}(\iota)e^{\mu}s^{\iota}.$$

Therefore, we can define $b_\mu^{\text{const}}(\iota) = \alpha b_\mu(\iota)$ from which, according to the linearity of the expected value

$$\mathbb{E}[b_\mu^{\text{const}}(\iota)|_i] = \mathbb{E}[\alpha b_\mu(\iota)|_i] = \alpha \mathbb{E}[b_\mu(\iota)|_i] = 0,$$

which proves property b).

Property a) follows directly by assumption from the bilinearity of the covariance

$$\begin{aligned}\mathsf{Cov}(b_{\mu_1}^{\text{const}}(\iota_1)|_{j_1}, b_{\mu_2}^{\text{const}}(\iota_2)|_{j_2}) &= \mathsf{Cov}(\alpha b_{\mu_1}(\iota_1)|_{j_1}, \alpha b_{\mu_2}(\iota_2)|_{j_2}) = \\ &= \alpha^2 \mathsf{Cov}(b_{\mu_1}(\iota_1)|_{j_1}, b_{\mu_2}(\iota_2)|_{j_2}) = 0.\end{aligned}$$

Multiplication of two ciphertexts The critical quantity arising from the multiplication of two ciphertexts, whose associated noise is defined as above, can be expressed as

$$\nu_{\mathsf{mul}} = \nu \cdot \nu' = \sum_\iota a_\iota^{\mathsf{mul}} s^\iota = \sum_\iota \sum_\mu b_\mu^{\mathsf{mul}}(\iota) e^\mu s^\iota,$$

where $a_\iota^{\mathsf{mul}} = \sum_{\iota_1+\iota_2=\iota} a_{\iota_1} a'_{\iota_2}$.

Moreover

$$\begin{cases} a_{\iota_1} = \sum_{\mu_1} b_{\mu_1}(\iota_1) e^{\mu_1} \\ a'_{\iota_2} = \sum_{\mu_2} b'_{\mu_2}(\iota_2) e^{\mu_2} \end{cases}$$

From which it follows that

$$b_\mu^{\mathsf{mul}}(\iota)|_i = \sum_{\iota_1+\iota_2=\iota} \sum_{\mu_1+\mu_2=\mu} \sum_j b_{\mu_1}(\iota_1)|_j \, b'_{\mu_2}(\iota_2)|_{i-j}.$$

From the independence of $b_{\mu_1}(\iota_1)$ and $b'_{\mu_2}(\iota_2)$ $\forall \mu_1, \mu_2, \iota_1, \iota_2$, and for the linearity of the expected value, one can deduce that

$$\begin{aligned}\mathbb{E}[b_\mu^{\mathsf{mul}}(\iota)|_i] &= \sum_{\iota_1+\iota_2=\iota} \sum_{\mu_1+\mu_2=\mu} \sum_j \mathbb{E}[b_{\mu_1}(\iota_1)|_j \, b'_{\mu_2}(\iota_2)|_{i-j}] \\ &= \sum_{\iota_1+\iota_2=\iota} \sum_{\mu_1+\mu_2=\mu} \sum_j \mathbb{E}[b_{\mu_1}(\iota_1)|_j] \mathbb{E}[b'_{\mu_2}(\iota_2)|_{i-j}] = 0,\end{aligned}$$

which easily proves property b).

The expression of the covariance $\mathsf{Cov}(b_{\mu_1}^{\mathsf{mul}}(\iota_1)|_{i_1}, b_{\mu_2}^{\mathsf{mul}}(\iota_2)|_{i_2})$ can be reduced, using its bilinearity, to a sum of terms of the form

$$\mathsf{Cov}(b_{\mu_1}(\iota_1)|_{l_1} b'_{\mu_2}(\iota_2)|_{i_1-l_1}, b_{\mu_3}(\iota_3)|_{l_2} b'_{\mu_4}(\iota_4)|_{i_2-l_2}),$$

which are all zero, using the property of the covariance stated below.

Property 1. Let X_1, X_2, X_3, X_4 be some fixed random variables.
If X_2, X_4 are independent with respect to X_1, X_3, $\mathsf{Cov}(X_2, X_4) = 0$ and $\mathbb{E}[X_2] = 0$ then

$$\mathsf{Cov}(X_1 \cdot X_2, X_3 \cdot X_4) = 0.$$

Thus, property a) follows by observing that, for $\mu_2 \neq \mu_4$ or $i_2 \neq i_4$:

- $\mathsf{Cov}(b'_{\mu_2}(\iota_2)|_{i_1-l_1}, b'_{\mu_4}(\iota_4)|_{i_2-l_2}) = 0$ e $\mathbb{E}[b'_{\mu_2}(\iota_2)|_{i_1-l_1}] = 0$ based on the hypotheses made;
- $b'_{\mu_2}(\iota_2)|_{i_1-l_1}, b'_{\mu_4}(\iota_4)|_{i_2-l_2}$ are independent with respect to $b_{\mu_1}(\iota_1)|_{l_1}, b_{\mu_3}(\iota_3)|_{l_2}$;

Therefore, thanks to the property 1, this implies that

$$\mathsf{Cov}(b_{\mu_1}(\iota_1)|_{l_1} b'_{\mu_2}(\iota_2)|_{i_1-l_1}, b_{\mu_3}(\iota_3)|_{l_2} b'_{\mu_4}(\iota_4)|_{i_2-l_2}) = 0.$$

Modulus and Key Switching Let $\boldsymbol{c} = (c_0, c_1)$ be a ciphertext in $R_{q_l} \times R_{q_l}$ and suppose that the modulus switch to $q_{l'}$ is applied, in order to reduce the error.
The resulting ciphertext is defined as

$$\boldsymbol{c'} = \frac{q_{l'}}{q_l}(\boldsymbol{c} + \boldsymbol{\delta}) \mod q_{l'},$$

where $\boldsymbol{\delta} = t[-\boldsymbol{c}t^{-1}]_{\frac{q_l}{q_{l'}}}$.
The critical quantity associated to $\boldsymbol{c'}$ can be expressed as

$$\nu' = \frac{q_{l'}}{q_l}(\nu + \nu_{\mathsf{ms}}) \quad \text{where} \quad \nu_{\mathsf{ms}} = \delta_0 + \delta_1 s\,.$$

It is possible to observe that the ciphertext components c_0, c_1 can be thought as randomly distributed over R_{q_l}, $c_0, c_1 \leftarrow \mathcal{U}_{q_l}$, and therefore the δ_i can be treated as independent polynomials with coefficients chosen randomly over $I = (-\frac{tq_l}{2q_{l'}}, \frac{tq_l}{2q_{l'}})$, i.e. $\delta_0, \delta_1 \leftarrow \mathcal{U}_I$.
Moreover, it should be noted that the values δ_i exclusively influence $b_0(0), b_0(1)$, and that they have an expected value equal to zero, because of their distributions.
Therefore, referring back to the case of the homomorphic sum, we can deduce that the expected value of $b'_\mu(\iota)$ for the new ciphertext $\boldsymbol{c'}$ remains zero, as do the covariances.
In the same way, by reducing the problem to the case of homomorphic addition, it is possible to show that these properties remain valid also after the relinearization process.
We decided not to report all the technical details but to provide only the key underlying idea, as there are multiple relinearization variants and including them would have required too much space. However, all the calculations can be derived in a very straightforward manner by simply adapting the approach in [4].

B Proof of Lemma 2

In order to prove the statement, we start by writing the term $a_\iota s^\iota|_i$, according to 1, as

$$a_\iota s^\iota|_i = \sum_\mu (b_\mu(\iota) e^\mu s^\iota)|_i = \sum_\mu \sum_{j=0}^{n-1} \xi(i,j) b_\mu(\iota) e^\mu|_j s^\iota|_{i-j}.$$

Thus, given two random variables X and Y, the following properties hold:

a. $\mathsf{Var}(XY) = (\mathsf{Var}(X) + \mathbb{E}[X]^2)(\mathsf{Var}(Y) + \mathbb{E}[Y]^2) + \mathsf{Cov}(X^2, Y^2)$
$\quad - (\mathsf{Cov}(X,Y) + \mathbb{E}[X]\mathbb{E}[Y])^2$
b. $\mathsf{Var}(X+Y) = \mathsf{Var}(X) + \mathsf{Var}(Y) + 2\mathsf{Cov}(X,Y)$,

where the second property can be generalized for k random variables $\{X_i\}_{i=0}^k$ as

$$\mathsf{Var}\left(\sum_{i=0}^k X_i\right) = \sum_{i=0}^k \mathsf{Var}(X_i) + \sum_{i_1 \neq i_2} \mathsf{Cov}(X_{i_1}, X_{i_2}).$$

Then, it is possible to compute the variance of $a_\iota s^\iota|_i$ as

$$\mathsf{Var}(a_\iota s^\iota|_i) = \sum_\mu \sum_{j=0}^{n-1} \mathsf{Var}(b_\mu(\iota) e^\mu|_j s^\iota|_{i-j}) + \sum_{\mu_1 \neq \mu_2 \text{ or } j_1 \neq j_2} \xi(i,j_1)\xi(i,j_2)\mathsf{Cov}(b_{\mu_1}(\iota) e^{\mu_1}|_{j_1} s^\iota|_{i-j_1}, b_{\mu_2}(\iota) e^{\mu_2}|_{j_2} s^\iota|_{i-j_2}),$$

where the covariances vanishes based on property 1. Thus, the following equality holds

$$\mathsf{Var}(a_\iota s^\iota|_i) = \sum_\mu \sum_{j=0}^{n-1} \mathsf{Var}(b_\mu(\iota) e^\mu|_j s^\iota|_{i-j}). \tag{15}$$

Moreover, according to property (a.), it follows that

$$\begin{aligned}\mathsf{Var}(b_\mu(\iota) e^\mu|_j s^\iota|_{i-j}) = {} & (\mathsf{Var}(b_\mu(\iota) e^\mu|_j) + \mathbb{E}[b_\mu(\iota) e^\mu|_j]^2)(\mathsf{Var}(s^\iota|_{i-j}) + \mathbb{E}[s^\iota|_{i-j}]^2) \\ & + \mathsf{Cov}(b_\mu(\iota) e^\mu|_j^2, s^\iota|_{i-j}^2) \\ & - (\mathsf{Cov}(b_\mu(\iota) e^\mu|_j, s^\iota|_{i-j}) + \mathbb{E}[b_\mu(\iota) e^\mu|_j]\mathbb{E}[s^\iota|_{i-j}])^2.\end{aligned}$$

At this point, it should be noted that

- $\mathbb{E}[b_\mu(\iota) e^\mu|_j] = 0$ according to Lemma 1;
- $\mathsf{Cov}(b_\mu(\iota) e^\mu|_j, s^\iota|_{i-j}) = \mathsf{Cov}(b_\mu(\iota) e^\mu|_j^2, s^\iota|_{i-j}^2) = 0$ as $b_\mu(\iota) e^\mu|_j, s^\iota|_{i-j}$ are independent;

This results in $\mathsf{Var}(b_\mu(\iota)e^\mu|_j s^\iota|_{i-j}) = \mathsf{Var}(b_\mu(\iota)e^\mu|_j)(\mathsf{Var}(s^\iota|_{i-j}) + \mathbb{E}[s^\iota|_{i-j}]^2)$. In addition, for a random variable X, it holds that $\mathsf{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$. Therefore,

$$\mathsf{Var}(b_\mu(\iota)e^\mu|_j s^\iota|_{i-j}) = \mathsf{Var}(b_\mu(\iota)e^\mu|_j)\mathbb{E}[s^\iota|^2_{i-j}]. \tag{16}$$

Finally, the same reasoning can be applied in order to derive

$$\mathsf{Var}(b_\mu(\iota)e^\mu|_j) = \sum_{k=0}^{n-1} \mathsf{Var}(b_\mu(\iota)|_k)\mathbb{E}[e^\mu|^2_{j-k}]. \tag{17}$$

In fact, using (1) and property (b.), it follows that

$$\begin{aligned}\mathsf{Var}(b_\mu(\iota)e^\mu|_j) &= \mathsf{Var}\left(\sum_{k=0}^{n-1} b_\mu(\iota)|_k e^\mu|_{j-k}\right)\\ &= \sum_{k=0}^{n-1} \mathsf{Var}(b_\mu(\iota)|_k e^\mu|_{j-k})\\ &\quad + \sum_{k_1\neq k_2} \xi(j,k_1)\xi(j,k_2)\mathsf{Cov}(b_\mu(\iota)|_{k_1} e^\mu|_{j-k_1}, b_\mu(\iota)|_{k_2} e^\mu|_{j-k_2}),\end{aligned}$$

where the covariances are null thanks to property 1.
Moreover, according to property (a.), it follows that

$$\begin{aligned}\mathsf{Var}(b_\mu(\iota)|_k e^\mu|_{j-k}) &= (\mathsf{Var}(b_\mu(\iota)|_k) + \mathbb{E}[b_\mu(\iota)|_k]^2)(\mathsf{Var}(e^\mu|_{j-k}) + \mathbb{E}[e^\mu|_{j-k}]^2)\\ &\quad + \mathsf{Cov}(b_\mu(\iota)|^2_k, e^\mu|^2_{j-k})\\ &\quad - (\mathsf{Cov}(b_\mu(\iota)|_k, e^\mu|_{j-k}) + \mathbb{E}[b_\mu(\iota)|_k]\mathbb{E}[e^\mu|_{j-k}])^2.\end{aligned}$$

Thus, (16) is proven observing that $\mathbb{E}[b_\mu(\iota)|_k] = 0$, according to lemma 1, and that $b_\mu(\iota)|_k$ and $e^\mu|_{j-k}$ are independent.
By substituting (16) and (17) in (15), it follows that

$$\begin{aligned}\mathsf{Var}(a_\iota s^\iota|_i) &= \sum_\mu \sum_{j=0}^{n-1} \mathsf{Var}(b_\mu(\iota)e^\mu|_j s^\iota|_{i-j}) = \sum_\mu \sum_{j=0}^{n-1} \mathsf{Var}(b_\mu(\iota)e^\mu|_j)\mathbb{E}[s^\iota|^2_{i-j}]\\ &= \sum_\mu \sum_{j=0}^{n-1}\sum_{k=0}^{n-1} \mathsf{Var}(b_\mu(\iota)|_k)\mathbb{E}[e^\mu|^2_{j-k}]\mathbb{E}[s^\iota|^2_{i-j}].\end{aligned}$$

Finally, observing that $\mathsf{Var}(b_\mu(\iota)|_i), \mathbb{E}[e^\mu|^2_i]$ and $\mathbb{E}[s^\iota|^2_i]$ do not depend on i, the thesis is demonstrated, i.e.,

$$\mathsf{Var}(a_\iota s^\iota|_i) = \sum_\mu \mathsf{Var}(b_\mu(\iota)|_i) \sum_{k=0}^{n-1} \mathbb{E}[e^\mu|^2_k] \sum_{j=0}^{n-1} \mathbb{E}[s^\iota|^2_j].$$

The expression for the variance of the coefficients of the critical quantity simply follows from the observation that the covariance between $a_{\iota_1} s^{\iota_1}|_i$ and $a_{\iota_2} s^{\iota_2}|_i$

vanishes whenever $\iota_1 \neq \iota_2$. Indeed, by bilinearity of the covariance operator, $\mathsf{Cov}(a_{\iota_1} s^{\iota_1}|_i, a_{\iota_2} s^{\iota_2}|_i)$ can be expressed as a sum of terms of the form

$$\mathsf{Cov}(b_{\mu_1}(\iota_1)|_j (e^{\mu_1} s^{\iota_1})|_{i-j}, b_{\mu_2}(\iota_2)|_j (e^{\mu_2} s^{\iota_2})|_{i-j}) \,,$$

each of which vanishes by Property 1. Therefore, we obtain

$$\mathsf{Var}(\nu|_i) = \mathsf{Var}(\sum_{\iota \geq 0} a_\iota s^\iota|_i) = \sum_{\iota \geq 0} \mathsf{Var}(a_\iota s^\iota|_i) + \sum_{\iota_1 \neq \iota_2} \mathsf{Cov}(a_{\iota_1} s^{\iota_1}|_i, a_{\iota_2} s^{\iota_2}|_i) \,,$$

thus concluding the proof.

C On the Negligibility of Key Switching

In this work, we chose to omit and not explicitly address the contribution of key switching in our analysis of noise variance, despite its necessity for ensuring the practicality of evaluated circuits. This decision is justified by the fact that, in general, the parameters associated with key switching can be selected so that its impact on the overall noise remains negligible in comparison to that of other operations — especially the multiplication operation, which immediately precedes the key switching step.

While our approach can easily accommodate this contribution, we opted to exclude it from our presentation for the sake of clarity and due to the aforementioned reason. Various key switching methods exist. To support the reasonableness of our choice, we provide a justification of its negligible impact for one of the main variants: GHS (Gentry Halevi Smart).

Let q be the modulus of the ciphertext to be relinearized, and let Q be a modulus specifically chosen for this purpose, such that $q < Q$ and $q \mid Q$.
The core idea of the GHS variant is to perform relinearization in the larger ring $\mathcal{R}_Q$, and then apply modulus switching to return to the smaller ring $\mathcal{R}_q$, reducing the error.

Let $c = (c_0, c_1, c_2)$ be the result of the multiplication of two ciphertexts, with $c_0, c_1, c_2 \in \mathcal{R}_q$.
The key-switching key for the GHS variant is defined as

$$(ek_0, ek_1) = \left(\left(-a \cdot s + te + \frac{Q}{q} s^2 \right), a \right) \mod Q \,.$$

The ciphertext resulting from relinearization in R_Q is given by

$$c_0' = \left[\frac{Q}{q} c_0 + c_2 \cdot ek_0 \right]_Q \,, \quad c_1' = \left[\frac{Q}{q} c_1 + c_2 \cdot ek_1 \right]_Q \,,$$

Then, this new ciphertext is scaled back modulo q using the modulus switching technique previously introduced, thereby obtaining

$$\hat{c_0} = \left[\frac{q}{Q}(c_0' + \delta_0)\right]_q \qquad \text{where } \delta_0 = t[-c_0' t^{-1}]_{\frac{Q}{q}} ,$$
$$\hat{c_1} = \left[\frac{q}{Q}(c_1' + \delta_1)\right]_q \qquad \text{where } \delta_1 = t[-c_1' t^{-1}]_{\frac{Q}{q}} ,$$

It is easy to verify that the critical quantity of the new relinearized ciphertext is given by

$$\nu_{\mathsf{ks}} = [\hat{c_0} + \hat{c_1} \cdot s]_q = \nu_{\mathsf{mul}} + \left[\frac{q}{Q}\left(tc_2 \cdot e + \delta_0 + \delta_1 \cdot s\right)\right]_q .$$

A typical choice for the size of Q is $Q \approx q^2$ [30]. In this way, the variance of the second component will be

$$V_{\frac{q}{Q}(tc_2 \cdot e + \delta_0 + \delta_1 \cdot s)} = \frac{1}{q^2}\left(nt^2 V_e V_{c_2} + V_{\delta_0} + nV_s V_{\delta_1}\right) = \frac{t^2}{12}\left(nV_e + 1 + nV_s\right) .$$

On the other hand, in accordance with the circuits proposed in the paper, we may assume that the ciphertext to be relinearized results from the product of two terms that have already undergone modulus switching.
Consequently, the variance of the first component, namely ν_{mul}, is at least $nV_{\mathsf{ms}}^2 F_s(1,1)$ where $V_{\mathsf{ms}} = \frac{t^2 n V_s}{12}$, thereby making evident the negligible contribution of the key switching step.