Spherical Gaussian Leftover Hash Lemma via the Rényi Divergence

Hiroki Okada[1(✉)], Kazuhide Fukushima[1], Shinsaku Kiyomoto[1], and Tsuyoshi Takagi[2]

[1] KDDI Research, Inc., Japan
ir-okada@kddi.com
[2] The University of Tokyo, Japan

Abstract. Agrawal *et al.* (Asiacrypt 2013) proved the discrete Gaussian leftover hash lemma, which states that the linear transformation of the discrete *spherical* Gaussian is statistically close to the discrete *ellipsoid* Gaussian. Showing that it is statistically close to the discrete *spherical* Gaussian, which we call the discrete *spherical* Gaussian leftover hash lemma (SGLHL), is an open problem posed by Agrawal *et al.* In this paper, we solve the problem in a weak sense: we show that the distribution of the linear transformation of the discrete spherical Gaussian and the discrete *spherical* Gaussian are close with respect to the *Rényi divergence* (RD), which we call the weak SGLHL (wSGLHL).

As an application of wSGLHL, we construct a sharper self-reduction of the learning with errors problem (LWE) problem. Applebaum *et al.* (CRYPTO 2009) showed that linear sums of LWE samples are statistically close to (plain) LWE samples with *some unknown* error parameter. In contrast, we show that linear sums of LWE samples and (plain) LWE samples with a *known* error parameter are close with respect to RD. As another application, we weaken the *independence heuristic* required for the fully homomorphic encryption scheme TFHE.

Keywords: Lattice · LWE · Discrete Gaussian · Leftover hash lemma

1 Introduction

Lattice-based cryptosystems are among the most promising candidates for post-quantum security. The National Institute of Standards and Technology (NIST) selected the lattice-based public key encryption scheme CRYSTALS-Kyber [BDK+18] and lattice-based digital signature schemes CRYSTALS-Dilithium [DKL+18] and Falcon [FHK+20] (as well as the hash-based digital signature scheme SPHINCS$^+$ [BHK+19]) as candidate algorithms to be standardized [AAC+22]. Furthermore, lattices can be used to build various advanced cryptographic primitives including identity based encryption (IBE) [GPV08], functional encryption [AFV11], fully homomorphic encryption (FHE) [BV11, BGV12, GSW13, DM15, CGGI17, CKKS17], and etc.

A crucial object in lattice-based cryptography is a *discrete Gaussian distribution* (Def. 2.23), which is a distribution over some fixed lattice, where every

lattice point is sampled with probability proportional to that of a continuous (multivariate) Gaussian distribution. In particular, efficient algorithms to sample from discrete Gaussians [GPV08, Pei10, MP12, MW17, GM18, DGPY20], and the analysis of various kinds of combinations of discrete Gaussians [Pei10, AGHS13, AR16, GMPW20] are required for the development of the (advanced) lattice-based cryptosystems.

A Gaussian Leftover Hash Lemma. The main concern of this paper is the discrete *Gaussian leftover hash lemma* (GLHL) proposed by Agrawal *et al.* [AGHS13]. The classic leftover hash lemma (LHL) [IZ89, HILL99, DRS04] states that a random linear combination of some (uniformly) random elements is statistically close to the uniform distribution over some finite domain. Similarly, GLHL states that the linear transformation of the discrete *spherical* Gaussian (vector of i.i.d 1-dimensional discrete Gaussians) is statistically close to the discrete *ellipsoid* Gaussian (vector of 1-dimensional discrete Gaussians that are neither identical nor mutually independent). It is an open question posed by Agrawal *et al.* [AGHS13] to show that the linear transformation of the discrete spherical Gaussian is statistically close to the discrete *spherical* Gaussian, which we call *spherical* GLHL (SGLHL):

> *"... our lattice version of LHL is less than perfect — instead of yielding a perfectly spherical Gaussian, it only gives us an approximately spherical one, i.e., $\mathcal{D}_{\mathcal{L},s\mathbf{X}}$. Here approximately spherical means that all the singular values of the matrix $\mathbf{X}$ within a small, constant sized interval."*

In this paper, we solve this open problem in a weak sense: we show that *Rényi divergence* (RD) (Def. 2.11) between the linear transformation of the discrete spherical Gaussian and the discrete *spherical* Gaussian is sufficiently small to construct security arguments, which we call the discrete weak SGLHL (wSGLHL). In addition, we show the continuous analog of the discrete wSGLHL, which we call the continuous wSGLHL. The RD has been used in prior works as a replacement to the statistical distance in lattice-based cryptography. As shown in, e.g., [Pre17, BLR+18, BJRW22, ASY22], some non-negligible, but a small RD is sometimes sufficient (or better) for constructing security proofs.

LWE Self Reduction. As an application of wSGLHL, we construct a new *self-reduction* of the *learning with errors* (LWE) problem defined as follows:

Definition 1.1 (LWE). *Let $n \in \mathbb{N}$ be a security parameter, $m = \mathrm{poly}(n)$ be the number of samples, the modulus $q = q(n) \geq 2$ be an integer, and χ be an error distribution. The samples from the* LWE *distribution* $\mathsf{LWE}_{\mathbf{s}}(m, n, q, \chi)$ *are* $(\mathbf{A}, \mathbf{As} + \mathbf{e})$ *for a fixed* $\mathbf{s} \sim \mathcal{U}(\mathbb{Z}_q^n)$*, where* $\mathbf{A} \sim \mathcal{U}(\mathbb{Z}_q^{m \times n})$*, and* $\mathbf{e} \sim \chi^m$*. The* $\mathsf{Search\text{-}LWE}_{\mathbf{s}}(m, n, q, \chi)$ *problem is to find* $\mathbf{s}$*, given samples from* $\mathsf{LWE}_{\mathbf{s}}(m, n, q, \chi)$*. The* $\mathsf{Decision\text{-}LWE}_{\mathbf{s}}(m, n, q, \chi)$ *problem is to distinguish between the distribution* $\mathsf{LWE}_{\mathbf{s}}(m, n, q, \chi)$ *and* $\mathcal{U}(\mathbb{Z}_q^{m \times n}, \mathbb{Z}_q^n)$*.*

Regev [Reg09] showed a (quantum) reduction from worst-case lattice problems to LWE with continuous Gaussian distribution (over the torus), and

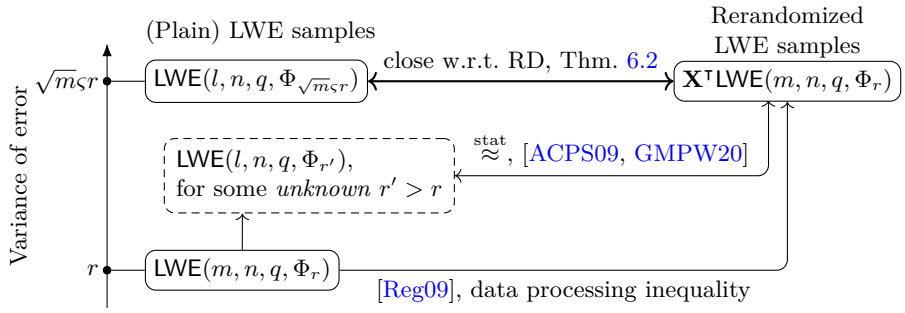

Fig. 1. Comparison of LWE self-reduction theorems. We write "$\mathcal{X} \to \mathcal{Y}$" to represent the (PPT) reduction from the problem defined with the distribution $\mathcal{X}$ to the problem defined with $\mathcal{Y}$, and "$\leftrightarrow$" denotes equivalence. Here, Φ_r denotes the continuous or discrete Gaussian distribution with parameter r, and ς is the (scaled) standard deviation of the elements of the randomization matrix $\mathbf{X} \in \mathbb{Z}_q^{m\times l}$ or vector $\mathbf{x} \in \mathbb{Z}_q^m$.

then reduced it to LWE with a *discretized Gaussian distribution* over $\mathbb{Z}_q$ with parameter r, denoted by $\overline{\Psi}_r$ (see Def. 2.22). Regev also constructed a public key encryption scheme based on LWE. Applebaum *et al.* [ACPS09] (and [GPV08, Pei10]) proposed a variant of Regev's encryption. In this scheme, the public key $(\mathbf{A}, \mathbf{b}) \sim \mathsf{LWE}_{\mathbf{s}}(m, n, q, \overline{\Psi}_r)$ is randomized for encryption as follows: $(\mathbf{a}'^\intercal, b') := (\mathbf{x}^\intercal\mathbf{A}, \mathbf{x}^\intercal\mathbf{b} + e')$, where $\mathbf{x} \sim \mathcal{D}_{\mathbb{Z}^m,\varsigma}$ and $e' \sim \overline{\Psi}_{\sqrt{m}\varsigma(r+\frac{1}{2q})}$. Here, the additional error e' is needed to "smooth out" the distribution. Interestingly, Applebaum *et al.* showed in [ACPS09, Lem. 4] that $(\mathbf{a}'^\intercal, b')$ is statistically close to an LWE sample with an *unknown* (albeit upper-bounded) error parameter $r' \le \sqrt{2m}\varsigma(r + \frac{1}{2q})$, which is called the LWE self-reduction. We refer $(\mathbf{a}'^\intercal, b')$ to as the "rerandomized" LWE sample, since it is essentially a new LWE sample with the fixed secret $\mathbf{s}$ (and a different error parameter). Similarly, Genise *et al.* [GMPW20] showed a "fully discrete" version of LWE self-reduction that uses only discrete Gaussians. However, in those self-reductions, the error parameter r' of the rerandomized LWE samples is *unknown* (secret), although its upper-bounded is given. Thus, we can only state that the rerandomized LWE instances are at least as hard as original given LWE samples with error parameter r (see Fig. 1), although r' is (often) larger than r.

In this paper, we show a sharper LWE self-reduction as an application of our wSGLHL. We show that rerandomized LWE samples are instances that are as hard as (plain) LWE samples with *known* (and large) error parameter. Formally, we let $(\mathbf{A}, \mathbf{b} := \mathbf{As} + \mathbf{e}) \sim \mathsf{LWE}_{\mathbf{s}}(m, n, q, \mathcal{D}_{\mathbb{Z},r})$ be m LWE samples, and we consider the case in which a randomization matrix $\mathbf{X}$ is sampled from a centered and β-bounded distribution $\chi_\beta^{m\times l}$ with $\mathrm{V}[\chi_\beta] := \varsigma^2$ and l rerandomized LWE samples generated as $(\mathbf{A}', \mathbf{b}') := (\mathbf{X}^\intercal\mathbf{A}, \mathbf{X}^\intercal\mathbf{b}) \in \mathbb{Z}_q^{l\times n} \times \mathbb{Z}_q^l$. We show that RD between rerandomized LWE samples $(\mathbf{A}', \mathbf{b}')$ and (plain) LWE samples $\mathsf{LWE}_{\mathbf{s}}(l, n, q, \mathcal{D}_{\mathbb{Z},\sqrt{m}\varsigma r})$ is sufficiently small (e.g., $\simeq 1.01 + \mathrm{negl}(n)$) to construct a LWE self-reduction: we show that finding $\mathbf{s}$ from the rerandomized LWE samples

$(\mathbf{A}', \mathbf{b}')$ is almost as hard as $\mathsf{Search\text{-}LWE}_{\mathbf{s}}(l, n, q, \mathcal{D}_{\mathbb{Z},\sqrt{m}\varsigma r})$, with the loss of only a few bits of security.

We illustrate the difference between our self-reduction and existing works in Fig. 1. It is easy to show that finding $\mathbf{s}$ from the rerandomized LWE samples $(\mathbf{A}', \mathbf{b}')$ is at least as hard as finding $\mathbf{s}$ from the original LWE samples (e.g., by the data processing inequality). Similarly, [Reg09, Lem. 5.4] shows that distinguishing rerandomized LWE samples from the uniformly random distribution is at least as hard as distinguishing the original LWE samples from the uniformly random distribution. Although existing works [ACPS09, HKM18, GMPW20] produce rerandomized samples that are statistically close to (plain) LWE samples, their error parameter is unknown, and thus, we can only obtain a hardness reduction from the original LWE instance. In contrast, our LWE self-reduction is sharper than these reductions in terms that we can base the security on a harder LWE instance with a (known) larger error parameter $\sqrt{m}\varsigma r > r$.

Note that our LWE self-reduction is between the search problems, while the existing works are the reduction between the decision problems. This is because our LWE self-reduction is based on RD. As discussed in the prior works that utilize RD [Pre17, BLR+18, BJRW22], RD is suited for search problems, while the statistical distance is suited for decision problems. Nonetheless, we can adapt the search-to-decision reduction [Reg09, MM11] or the trivial decision-to-search reduction if we want to connect our LWE self-reduction to Decision-LWE.

The Independence Heuristic. As another application of our wSGLHL, we weaken the *independence heuristic* that is required for the fully homomorphic encryption scheme TFHE [CGGI16, CGGI17, CGGI20]. The TFHE scheme relies on the heuristic that linear combinations of the errors of ciphertexts are mutually independent in order to analyze their variance. Since our continuous wSGLHL shows that linear sums of Gaussian errors and mutually independent Gaussian errors are close with respect to RD, it essentially mitigates the heuristic. We adapt continuous wSGLHL to the concrete setting of the TFHE scheme, and we mitigate the independence heuristic to a weaker heuristic. Note that our result does not improve the parameter choice of the TFHE, since we only provide a theoretical evidence to the independence heuristic.

Technical overview. We first show the construction of the *approximately orthogonal matrix* (Def. 3.1), in Thm. 3.5. This is the building block of our main theorem, the wSGLHL (Thm. 4.2 and Thm. 5.4). As an application of wSGLHL, we show a new LWE self-reduction (Thm. 6.2, Cor. 6.3). In addition, we apply wSGLHL to mitigate the independent heuristic of TFHE to some weaker heuristic. We provide a technical overview of our results in what follows.

Approximately orthogonal matrix (Sect. 3). We call $\mathbf{X} \in \mathbb{R}^{m \times l}$ an *approximately orthogonal matrix with bound* $\delta > 0$, iff all the absolute values of the elements of a matrix $\mathbf{R} := \mathbf{X}^{\intercal}\mathbf{X} - \mathbf{I}_l$ are smaller than δ (see Def. 3.1). As mentioned, we sample the randomization matrix $\mathbf{X}$ from the centered and β-bounded

distribution $\chi_\beta^{m\times l}$ with $\mathrm{V}[\chi_\beta] := \varsigma^2$ and $\beta > 0$. Then, in Thm. 3.5, we show that $(\frac{1}{\sqrt{m}\varsigma}\mathbf{X})$ is an approximately orthogonal matrix with bound $\delta = \omega(1/\sqrt{m})$, with overwhelming probability over the choice of $\mathbf{X}$. This construction of the approximately orthogonal matrix is a key technique for our (continuous and discrete) wSGLHL, as explained in the following part.

Continuous wSGLHL (Sect. 4). Let $\mathbf{e}$ be an m-dimensional multivariate continuous Gaussian with a mean of $\mathbf{0}$ and (scaled) covariance matrix $\mathbf{\Sigma} \succeq 0$; i.e., $\mathbf{e} \sim \mathcal{N}_m(\mathbf{\Sigma})$ (see Def. 2.18). We refer to $\mathbf{e}$ as *spherical* if $\mathbf{\Sigma} = s^2\mathbf{I}_m$ for some $s > 0$ (i.e., $\mathbf{e} \sim \mathcal{N}_m(s^2)$), and *ellipsoid* otherwise.

The continuous *ellipsoid* Gaussian LHL states that the linear transformation of the continuous spherical Gaussian $\mathbf{e} \sim \mathcal{N}_m(s^2)$ by $\mathbf{X} \in \mathbb{R}^{m\times l}$, i.e., the "rerandomized" Gaussian $\mathbf{X}^\intercal\mathbf{e}$, is a continuous *ellipsoid* Gaussian. This follows trivially from the linear transformation lemma for the continuous Gaussian (Lem. 2.19), as we have $\mathbf{X}^\intercal\mathbf{e} \sim \mathcal{N}_l(s^2\mathbf{\Sigma})$ for $\mathbf{\Sigma} = \mathbf{X}^\intercal\mathbf{X}$.

The continuous SGLHL states that $\mathbf{X}^\intercal\mathbf{e}$ is a continuous *spherical* Gaussian, i.e., $\mathbf{X}^\intercal\mathbf{e} \sim \mathcal{N}_l(\varsigma^2 s^2)$ for some $\varsigma > 0$. This holds only if $\mathbf{X}^\intercal\mathbf{X} = \varsigma^2\mathbf{I}_l$, i.e., only when $\mathbf{X}$ is a (scaled) orthogonal matrix. However, sampling an orthogonal matrix is not efficient and thus it is not preferable in cryptographic applications. Hence, we consider taking $\mathbf{X}$ as an approximately orthogonal matrix, which is more general and is easier to sample. Specifically, we sample $\mathbf{X} \sim \chi_\beta^{m\times l}$ and use Thm. 3.5 to obtain the bound on the elements of the residual matrix $\mathbf{R} := \frac{1}{m\varsigma^2}\mathbf{X}^\intercal\mathbf{X} - \mathbf{I}_l$. Then, we obtain (small) upper bound on RD between the rerandomized Gaussian $\mathbf{X}^\intercal\mathbf{e}$ and some continuous spherical Gaussian, i.e., the continuous wSGLHL. In Thm. 4.2, informally, we show that

$$R_a(\mathbf{X}^\intercal\mathcal{N}_m(s^2) \,\|\, \mathcal{N}_l(m\varsigma^2 s^2)) < (1 + l/\sqrt{m})^{\frac{1}{a-1}}$$

holds for any constant $a \in [2, \infty)$ and some $l < \sqrt{m}$.

Furthermore, we propose an improved theorem Thm. 4.4 that is applicable to any large l at the expense of increasing the size of the rerandomized Gaussian. This theorem analyzes RD between $\mathbf{X}^\intercal\mathbf{e} + \mathbf{e}''$ and $\mathcal{N}_l((1+k)\varsigma^2 s^2)$, where $\mathbf{e}'' \sim \mathcal{N}_l(k\varsigma^2 s^2)$ is an additional continuous spherical Gaussian for $k > 0$, and it yields the same upper bound on RD of Thm. 4.2. This technique is conceptually similar to the "noise flooding" technique proposed in [BGM+16, BLR+18].

Discrete wSGLHL (Sect. 5). We present similar results for the discrete Gaussian. Let $\mathbf{e}$ be a discrete Gaussian over the m-dimensional integer lattice $\mathbb{Z}^m$ with a mean of $\mathbf{0}$ and (scaled) covariance matrix $\mathbf{\Sigma} \succeq 0$; i.e., $\mathbf{e} \sim \mathcal{D}_{\mathbb{Z}^m,\sqrt{\mathbf{\Sigma}}}$ (see Def. 2.23). We refer to $\mathbf{e}$ as *spherical* if $\mathbf{\Sigma} = r^2\mathbf{I}_m$ for some $r > 0$ (i.e., $\mathbf{e} \sim \mathcal{D}_{\mathbb{Z}^m,r}$) and as *ellipsoid* otherwise.

Unlike the case of the continuous Gaussian, the linear transformation lemma for the discrete Gaussian is not trivial. Agrawal *et al.* first proved the discrete (ellipsoid) Gaussian LHL in [AGHS13]. This lemma states that the linear transformation of the discrete spherical Gaussian $\mathbf{e} \sim \mathcal{D}_{\mathbb{Z}^m,r}$ by $\mathbf{X} \in \mathbb{R}^{m\times l}$, i.e., the rerandomized discrete Gaussian $\mathbf{X}^\intercal\mathbf{e}$, is statistically close to the discrete

ellipsoid Gaussian $\mathcal{D}_{\mathbb{Z}^l, r\mathbf{X}}$. We extend the discrete (ellipsoid) Gaussian LHL to the discrete wSGLHL, which states that $\mathbf{X}^\intercal\mathbf{e}$ and the discrete *spherical* Gaussian are close with respect to RD. In Thm. 5.4, informally, we show that

$$R_a(\mathbf{X}^\intercal\mathcal{D}_{\mathbb{Z}^m,r} \parallel \mathcal{D}_{\mathbb{Z}^m,\sqrt{m}\varsigma r}) < (1+\mathrm{negl}(n))(1+l/\sqrt{m})^{\frac{1}{a-1}}$$

holds for any $a \in [2,\infty)$ and some $l < \sqrt{m}$ by instantiating $\mathbf{X} \sim (\mathcal{D}_{\mathbb{Z}^m,\varsigma})^l$ and applying Thm. 3.5.

A Sharper LWE *self-reduction (Sect. 6).* Finally, we show a sharper LWE self-reduction in Thm. 6.2, by applying our wSGLHL. Although we only consider the LWE with the discrete Gaussian in Sect. 6, similar results can be obtained for the LWE with the continuous Gaussian from our continuous wSGLHL.

Our goal is to show that the distribution of rerandomized LWE samples is close to (plain) LWE distribution. Let $(\mathbf{A},\mathbf{b}) \sim \mathsf{LWE}_{\mathbf{s}}(m,n,q,\mathcal{D}_{\mathbb{Z},r})$, and define the rerandomized LWE samples as $(\mathbf{A}',\mathbf{b}') := (\mathbf{X}^\intercal\mathbf{A}, \mathbf{X}^\intercal\mathbf{b})$. By adapting the classical leftover hash lemma [Lyu05, Reg09], we can show that $\mathbf{A}'$ is statistically close to $\mathcal{U}(\mathbb{Z}_q^{l\times n})$. Hence, we only need to show that the rerandomized error $\mathbf{e}' := \mathbf{X}^\intercal\mathbf{e}$, where $\mathbf{e} := (\mathbf{b}-\mathbf{A}\mathbf{s}) \sim \mathcal{D}_{\mathbb{Z}^m,r}$, is close to the discrete *spherical* Gaussian $\mathcal{D}_{\mathbb{Z}^l,\sqrt{m}\varsigma r}$: indeed, this is shown through the discrete wSGLHL. Informally, Thm. 6.2 shows that

$$R_a((\mathbf{X}^\intercal\mathbf{A}, \mathbf{X}^\intercal\mathbf{b}) \parallel \mathsf{LWE}_{\mathbf{s}}(l,n,q,\mathcal{D}_{\mathbb{Z},\sqrt{m}\varsigma r})) < (1+\mathrm{negl}(n))\left(1+l/\sqrt{m}\right)^{\frac{1}{a-1}}.$$

Unlike the standard security arguments based on the statistical distance, we do not (need to) show that RD is negligibly small (i.e., $R_a = 1+\mathrm{negl}(n)$)[3]. As mentioned earlier, some non-negligible, but small RD is sufficient for constructing security arguments. We demonstrate that Thm. 6.2 implies the following LWE self-reduction (Cor. 6.4) by selecting some concrete parameters:

$$\mathsf{Search}\text{-}(\mathbf{X}^\intercal\mathsf{LWE}_{\mathbf{s}}(m,n,q,\mathcal{D}_{\mathbb{Z},r})) \simeq \mathsf{Search}\text{-}\mathsf{LWE}_{\mathbf{s}}(l,n,q,\mathcal{D}_{\mathbb{Z},\sqrt{m}\varsigma r})),$$

which means that the problem of finding $\mathbf{s}$ from the rerandomized LWE samples $(\mathbf{X}^\intercal\mathsf{LWE}_{\mathbf{s}}(m,n,q,\mathcal{D}_{\mathbb{Z},r}))$ is almost as hard as $\mathsf{Search}\text{-}\mathsf{LWE}_{\mathbf{s}}(l,n,q,\mathcal{D}_{\mathbb{Z},\sqrt{m}\varsigma r})$ with the loss of only a few small bits of security. Note that, while existing works [ACPS09, HKM18, GMPW20] treat the given LWE samples $(\mathbf{A},\mathbf{b})$ as fixed values, we treat them as stochastic variables throughout this paper.

Related works. Pellet-Mary and Stehlé [PS21, Lem. 2.3] analyzed an upper bound of the statistical distance between two (multivariate) discrete Gaussian distributions $\mathcal{D}_{\mathcal{L}+\mathbf{c}_1,\mathbf{S}_1}$ and $\mathcal{D}_{\mathcal{L}+\mathbf{c}_2,\mathbf{S}_1}$ over the lattice $\mathcal{L} \subset \mathbb{R}^n$ (see Def. 2.23), where $\mathbf{S}_1$ and $\mathbf{S}_2$ are (conditioned) covariance matrices and $\mathbf{c}_1$ and $\mathbf{c}_2$ are arbitrary centers. The lemma was derived from the Kullback–Leibler (KL) divergence and Pinsker's inequality. The KL divergence is a special case of RD:

[3] Although we obtain $R_a = 1+\mathrm{negl}(n)$ if we set, e.g., $m = 2^n$ and $l = \mathrm{poly}(n)$, this may not be useful for practical cryptographic applications.

$\mathrm{KL}(\cdot \,\|\, \cdot) = \log R_1(\cdot \,\|\, \cdot)$ by definition. Furthermore, since R_a is nondecreasing in $a \in [1, \infty]$ (Lem. 2.12), $\log R_a$ for any $a \in [1, \infty]$ gives an upper-bound on the KL divergence. Similar to [PS21, Lem. 2.3], our Lem. 5.2 (and Lem. 4.1 for the continuous Gaussian) analyzes RD between two discrete Gaussian distributions, $R_a(\mathcal{D}_{\mathbb{Z}^n, r\mathbf{X}} \,\|\, \mathcal{D}_{\mathbb{Z}^n, rs\mathbf{I}})$, for any $a \in (1, \infty)$ and some $r, s \in \mathbb{R}$. Although [PS21, Lem. 2.3] is also applicable to our case of interest ($\mathcal{L} = \mathbb{Z}^n$, $\mathbf{S}_1 = r\mathbf{X}$, $\mathbf{S}_2 = rs\mathbf{I}$), it supports only $a = 1$ and thus is not sufficient for our LWE self-reduction, Cor. 6.4. Due to the flexibility of a in our Lem. 5.2 (and Lem. 5.3), we can adjust the loss of security bits so that it is very small in Cor. 6.4.

Case *et al.* [CGHX19] claimed that they removed the need for the independence heuristic in TFHE works, namely, [CGGI20, Assumption 3.11], but this claim is incorrect. They showed in [CGHX19, Thm. 3.2] that a linear sum of *sub-Gaussian* variables is a sub-Gaussian variable, and derived the worst-case upper bound of the errors included in TFHE ciphertexts in [CGHX19, Lem. 5.2]. However, a worst-case upper bound of the errors was already given in [CGGI16] without relying on the independence heuristic: As mentioned in [CGGI16], the independence heuristic was only needed to analyze the "average-case" bound, i.e., the variance of the errors. Case *et al.* did not derive the average-case bound, and did not show that the linear sums of sub-Gaussian variables are mutually independent. Therefore, we provide the first evidence that mitigates the independence heuristic.

Organization. The remainder of the paper is organized as follows. In Sect. 2, we provide the definitions and preliminaries required for our work. In Sect. 3, we show the construction of the approximately orthogonal matrix: Thm. 3.5. Using this theorem as a building block, we show the continuous wSGLHL (Thm. 4.2) and discrete wSGLHL (Thm. 5.4) in Sect. 4 and Sect. 5, respectively. As an application of the wSGLHL, we show a sharper LWE self-reduction (Thm. 6.2) in Sect. 6. In addition, we discuss how the wSGLHL can be adapted to mitigate the independence heuristic required for TFHE in Sect. 7.

2 Preliminaries

We use log and ln to denote the base 2 logarithm and the natural logarithm, respectively. $\mathbb{R}^+$ denotes the set of positive real numbers. For any natural number $s \in \mathbb{N}$, the set of the first s positive integers is denoted by $[s] = \{1, \cdots, s\}$. Let $\varepsilon > 0$ denote some small (often, negligible) number; we use the notational shorthand $\hat{\varepsilon} := \varepsilon + O(\varepsilon^2)$. One can check that $\frac{1+\varepsilon}{1-\varepsilon} = 1 + 2\hat{\varepsilon}$ and $\ln\left(\frac{1+\varepsilon}{1-\varepsilon}\right) = 2\hat{\varepsilon}$. Other notation can be found in the rest of this section.

2.1 Linear Algebra

Vectors are in column form and are written using bold lower-case letters, e.g., $\mathbf{x}$. The i-th component of $\mathbf{x}$ is denoted by x_i. Matrices are written as bold capital

letters, e.g., $\mathbf{X}$, and the i-th column vector of $\mathbf{X}$ is denoted as $\mathbf{x}_i$. When we write $\mathbf{X} = [x_{ij}]$, x_{ij} denotes the i-th element of $\mathbf{x}_j$. The inverse transpose of $\mathbf{X}$ is denoted as $\mathbf{X}^{-\intercal}$.We write $\mathbf{X} \succ 0$ ($\mathbf{X} \succeq 0$) if $\mathbf{X}$ is positive definite (semidefinite). Let $\mathbf{I}_n \in \mathbb{Z}^{n\times n}$ be the identity matrix and let $\mathbf{O}_n \in \mathbb{Z}^{n\times n}$ be the zero matrix. We sometimes denote $\mathbf{I}_n$ and $\mathbf{O}_n$ by $\mathbf{I}$ and $\mathbf{O}$, respectively, when the subscript n is obvious from the context. For $m \geq n$, we call $\mathbf{X} \in \mathbb{R}^{m\times n}$ an orthogonal matrix if $\mathbf{X}^\intercal\mathbf{X} = \mathbf{I}_n$. For a rank-$n$ matrix $\mathbf{X} \in \mathbb{R}^{m\times n}$ with $m \geq n$, we denote its singular values by $\sigma_1(\mathbf{X}) \leq \cdots \leq \sigma_n(\mathbf{X})$. The eigenvalues of $\mathbf{X} \in \mathbb{R}^{n\times n}$ are denoted by $e_1(\mathbf{X}) \leq \cdots \leq e_n(\mathbf{X})$. The determinant of a square matrix $\mathbf{X}$ is denoted by $\det(\mathbf{X})$ or $|\mathbf{X}|$. The usual Euclidean norm (l_2-norm) and infinity norm of the vector $\mathbf{x}$ are denoted by $\|\mathbf{x}\|$ and $\|\mathbf{x}\|_\infty$, respectively. The spectral norm $|||\cdot|||$ is defined on $\mathbb{R}^{n\times n}$ by $|||\mathbf{A}||| = \max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\| = \sigma_n(\mathbf{A})$. We also define the *length of a matrix* on $\mathbb{R}^{n\times n}$ as $|||\mathbf{A}|||_{len} = \max_{i\in[n]} \|\mathbf{a}_i\|$. Although the length of a matrix is not a matrix norm, it has the following properties:

Fact 2.1. *For any matrices* $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n\times n}$, $|||\mathbf{X}\mathbf{Y}|||_{len} \leq \sqrt{n} \cdot |||\mathbf{X}|||_{len} |||\mathbf{Y}|||_{len}$.

Fact 2.2. *For any* $\mathbf{A} \in \mathbb{R}^{n\times n}$, $|||\mathbf{A}|||_{len} \leq |||\mathbf{A}||| = \sigma_n(\mathbf{A})$.

We recall some notions related to the positive (semi)definite matrix.

Lemma 2.3 ([HJ85, Thm. 7.2.6]). *Let* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *be a symmetric matrix,* $\mathbf{A} \succeq 0$*, and let* $k \in \{2, 3, \dots\}$*. There is a unique symmetric matrix* $\mathbf{B}$ *such that* $\mathbf{B} \succeq 0$, $\mathbf{B}^k = \mathbf{A}$*, and* $\operatorname{rank}\mathbf{A} = \operatorname{rank}\mathbf{B}$*. (In particular, we denote the unique positive (semi)definite square root of* $\mathbf{A}$ *by* $\sqrt{\mathbf{A}}$*.)*

Lemma 2.4 ([HJ85, Thm. 7.2.7]). *Let* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *be a symmetric matrix. If* $\mathbf{A} = \mathbf{B}^\intercal\mathbf{B}$ *with* $\mathbf{B} \in \mathbb{R}^{m\times n}$*, then* $\mathbf{A} \succ 0$ *if and only if* $\mathbf{B}$ *has full column rank.*

We recall some notions related to the *diagonally dominant matrix*:

Definition 2.5. *A square matrix* $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{n\times n}$ *is diagonally dominant if* $|a_{ii}| \geq \sum_{j\neq i} |a_{ij}|$ *holds for all* $i \in [n]$*. It is strictly diagonally dominant if* $|a_{ii}| > \sum_{j\neq i} |a_{ij}|$ *holds for all* $i \in [n]$*.*

Lemma 2.6 ([HJ85, Thm. 6.1.10]). *Let* $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{n\times n}$ *be* strictly *diagonally dominant. If* $\mathbf{A}$ *is symmetric and* $\forall i \in [n]$, $a_{ii} > 0$*, then* $\mathbf{A} \succ 0$.

We refer to some useful lemmas for a matrix with bounded entries.

Lemma 2.7 ([Ost38]). *Let* $\mathbf{R} = [r_{ij}] \in \mathbb{R}^{n\times n}$*. If* $|r_{ij}| \leq \delta$ *for all* $i, j \in [n]$ *and* $n\delta \leq 1$*, then* $1 - n\delta \leq |\mathbf{I}_n - \mathbf{R}| \leq 1/(1 - n\delta)$[4] *holds.*

Lemma 2.8 ([Zha05]). *Let* $\mathbf{R} = [r_{ij}] \in \mathbb{R}^{n\times n}$ *be symmetric. If* $|r_{ij}| \leq \delta$ *for all* $i, j \in [n]$*, then* $-n\delta \leq e_i(\mathbf{R}) \leq n\delta$ *for all* $i \in [n]$*.*

Lemma 2.9 (Adapted from [GV96, Thm. 8.1.5]). *If* $\mathbf{R} \in \mathbb{R}^{n\times n}$ *is a symmetric matrix, then for all* $i \in [n]$, $1 + e_1(\mathbf{R}) \leq e_i(\mathbf{I} + \mathbf{R}) \leq 1 + e_n(\mathbf{R})$.

2.2 Lattices

A lattice is a discrete additive subgroup of $\mathbb{R}^m$. A set of linearly independent vectors that generates a lattice is called a basis and is denoted as $\mathbf{B} =$

[4] Although [BOS15, Thm. 2] gives a sharper bound, we use this simpler formula.

$\{\mathbf{b}_1, \ldots, \mathbf{b}_n\} \subset \mathbb{R}^m$ for integers $m \geq n \geq 1$. The lattice generated by the basis $\mathbf{B}$ is $\mathcal{L} = \mathcal{L}(\mathbf{B}) = \{\sum_{i=1}^n z_i \mathbf{b}_i \mid \mathbf{z} \in \mathbb{Z}^n\}$. If we arrange the vectors $\mathbf{b}_i$ as the columns of a matrix $\mathbf{B} \in \mathbb{R}^{n \times m}$, then we can write $\mathcal{L} = \{\mathbf{Bz} \mid \mathbf{z} \in \mathbb{Z}^n\}$. We say that the rank of this lattice is n and its dimension is m. If $n = m$, we call the lattice *full rank*. Let $\widehat{\mathcal{L}} = \{\mathbf{u} \in \mathrm{span}(\mathcal{L}) \mid \forall \mathbf{v} \in \mathcal{L}, \langle \mathbf{u}, \mathbf{v} \rangle \in \mathbb{Z}\}$ be the dual lattice of $\mathcal{L}$. We denote the volume of the fundamental parallelepiped of $\mathcal{L}$ as $\det(\mathcal{L})$. If lattice $\mathcal{L}(\mathbf{B})$ is full rank, then $\mathbf{B}$ is a nonsingular square matrix and $\det(\mathcal{L}(\mathbf{B})) = |\det(\mathbf{B})|$. Note that $\det(\widehat{\mathcal{L}}) = 1/\det(\mathcal{L})$. For any (ordered) set $\mathbf{S} = \{\mathbf{s}_1, \ldots, \mathbf{s}_n\} \subset \mathbb{R}^n$ of linearly independent vectors, let $\tilde{\mathbf{S}} = \{\tilde{\mathbf{s}}_1, \ldots, \tilde{\mathbf{s}}_n\} \subset \mathbb{R}^n$ denote its Gram–Schmidt orthogonalization. For a lattice $\mathcal{L}(\mathbf{B})$, we define the *Gram–Schmidt minimum* as $\tilde{bl}(\mathcal{L}(\mathbf{B})) = \min_{\mathbf{B}} \|\tilde{\mathbf{B}}\|_{len} = \min_{\mathbf{B}} \max_{i \in [n]} \|\tilde{\mathbf{b}_i}\|$.

2.3 Statistics

We write $X \sim \mathcal{D}$ to indicate that the random variable X is distributed according to the distribution $\mathcal{D}$. Let $X \sim \mathcal{D}$. We denote the probability function of a distribution $\mathcal{D}$ as $\mathcal{D}(x) = \Pr[X = x]$ and let $\mathrm{Supp}(\mathcal{D}) := \{x \mid \mathcal{D}(x) \neq 0\}$. We denote the mean and variance of X by $\mathrm{E}[X]$ and $\mathrm{V}[X]$, respectively. We say $\mathcal{D}$ is β-*bounded* if $\mathrm{Supp}(\mathcal{D}) \subseteq [-\beta, \beta]$ for $0 < \beta \in \mathbb{R}$, and is *centered* if $\mathrm{E}[X] = 0$. For a real-valued function f and a countable set S, we write $f(S) = \Sigma_{x \in S} f(x)$, assuming that this sum is absolutely convergent. For a matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$ and a distribution $\mathcal{D}$ over $\mathbb{R}^m$, we denote the distribution $\{\mathbf{X}^\intercal \mathbf{v} \in \mathbb{R}^n \mid \mathbf{v} \sim \mathcal{D}\}$ as $\mathbf{X}^\intercal \mathcal{D}$. For distributions $\mathcal{D}_1$ and $\mathcal{D}_2$, we denote the distribution $\{v_1 + v_2 \mid v_1 \sim \mathcal{D}_1, v_2 \sim \mathcal{D}_2\}$ as $(\mathcal{D}_1 + \mathcal{D}_2)$. We denote $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{D}$ if $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d.) according to the distribution $\mathcal{D}$. We define the statistical distance and RD as follows:

Definition 2.10 (Statistical distance). *Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be probability distributions over a (countable) set Ω. Then, the statistical distance between $\mathcal{D}_1$ and $\mathcal{D}_2$ is defined as the function*

$$\Delta(\mathcal{D}_1, \mathcal{D}_2) := \tfrac{1}{2} \textstyle\sum_{x \in \Omega} |\mathcal{D}_1(x) - \mathcal{D}_2(x)|.$$

The definition is extended in a natural way to continuous distributions.

Definition 2.11 (Rényi divergence). *For any two discrete probability distributions $\mathcal{D}_1$ and $\mathcal{D}_2$ such that $S := \mathrm{Supp}(\mathcal{D}_1) \subseteq \mathrm{Supp}(\mathcal{D}_2)$, we define the Rényi divergence (RD) of order $a \geq 1$ as*

$$\begin{aligned} R_1(\mathcal{D}_1 \parallel \mathcal{D}_2) &:= \exp\left(\textstyle\sum_{x \in S} \mathcal{D}_1(x) \log\left(\mathcal{D}_1(x)/\mathcal{D}_2(x)\right)\right), \\ R_a(\mathcal{D}_1 \parallel \mathcal{D}_2) &:= \left(\textstyle\sum_{x \in S} \mathcal{D}_1(x)^a / \mathcal{D}_2(x)^{a-1}\right)^{\frac{1}{a-1}} \text{ for } a \in (1, \infty), \text{and} \\ R_\infty(\mathcal{D}_1 \parallel \mathcal{D}_2) &:= \max_{x \in S} (\mathcal{D}_1(x)/\mathcal{D}_2(x)). \end{aligned}$$

The definitions are extended in a natural way to continuous distributions.

The above RD is slightly different from some other definitions [Rén61], which take the log of our version of RD. The properties of RD can be found in [EH14,

LSS14, BLR+18]. We recall the properties required for our construction. In the rest of paper, we use the shorthand $c_a := \frac{a}{a-1}$ for $a > 1$.

Lemma 2.12 ([BLR+18, Lem. 2.9]). *Let $a \in [1, \infty]$, and define $c_a := \frac{a}{a-1}$ for $a > 1$. Let $\mathcal{P}$ and $\mathcal{Q}$ denote distributions with $\mathrm{Supp}(\mathcal{P}) \subseteq \mathrm{Supp}(\mathcal{Q})$. Then, the following properties hold:*

***Data processing inequality**: $R_a(\mathcal{P}^f \,\|\, \mathcal{Q}^f) \leq R_a(\mathcal{P} \,\|\, \mathcal{Q})$ for any function f where $\mathcal{P}^f$ (resp. $\mathcal{Q}^f$) denotes the distribution of $f(y)$ induced by sampling $y \sim \mathcal{P}$ (resp. $y \sim \mathcal{Q}$).*

***Multiplicativity**: Assume $\mathcal{P}$ and $\mathcal{Q}$ are two distributions of a pair of mutually independent random variables (Y_1, Y_2). For $i \in \{1, 2\}$, let $\mathcal{P}_i$ (resp. $\mathcal{Q}_i$) denote the marginal distribution of Y_i under $\mathcal{P}$ (resp. $\mathcal{Q}$). Then, $R_a(\mathcal{P} \,\|\, \mathcal{Q}) = R_a(\mathcal{P}_1 \,\|\, \mathcal{Q}_1) R_a(\mathcal{P}_2 \,\|\, \mathcal{Q}_2)$.*

***Probability preservation:** Let $E \subseteq \mathrm{Supp}(\mathcal{Q})$ be an arbitrary event. For $a \in (1, \infty)$, $\mathcal{Q}(E) \geq \mathcal{P}(E)^{c_a} / R_a(\mathcal{P} \,\|\, \mathcal{Q})$, and $\mathcal{Q}(E) \geq \mathcal{P}(E) / R_\infty(\mathcal{P} \,\|\, \mathcal{Q})$.*

***Weak triangle inequality:** Let $\mathcal{P}_1$, $\mathcal{P}_2$, and $\mathcal{P}_3$ be three distributions with $\mathrm{Supp}(\mathcal{P}_1) \subseteq \mathrm{Supp}(\mathcal{P}_2) \subseteq \mathrm{Supp}(\mathcal{P}_3)$. Then, we have $R_a(\mathcal{P}_1 \,\|\, \mathcal{P}_3) \leq R_a(\mathcal{P}_1 \,\|\, \mathcal{P}_2) R_\infty(\mathcal{P}_2 \,\|\, \mathcal{P}_3)$ and $R_a(\mathcal{P}_1 \,\|\, \mathcal{P}_3) \leq R_\infty(\mathcal{P}_1 \,\|\, \mathcal{P}_2)^{c_a} R_a(\mathcal{P}_2 \,\|\, \mathcal{P}_3)$ if $a \in (1, \infty)$.*

Lemma 2.13 ([EH14, Thm. 3]). *$R_a(\mathcal{P} \,\|\, \mathcal{Q})$ is nondecreasing in $a \in [1, \infty]$.*

Lemma 2.14 (Adapted from [Pre17, Sect. 3.3]). *Let $n \in \mathbb{N}$ be a security parameter. For any algorithm f, define $\mathcal{P}^f$ (resp. $\mathcal{Q}^f$) as the distribution of $f(y)$ induced by sampling $y \sim \mathcal{P}$ (resp. $y \sim \mathcal{Q}$). Assume that for any (PPT) algorithm f and an event $E \subseteq \mathrm{Supp}(\mathcal{Q}^f)$, there exists a constant $C > 0$ and $\mathcal{Q}^f(E) \leq 2^{-C \cdot n} (= \mathrm{negl}(n))$ holds. Then, for any $a > 1$, we have $\mathcal{P}^f(E) \leq 2^{-\frac{1}{c_a}(Cn - \log R_a(\mathcal{P} \| \mathcal{Q}))}$.*

Proof. For any $a > 1$, we have $\mathcal{Q}^f(E) \geq (\mathcal{P}^f(E))^{c_a} / R_a(\mathcal{P}^f \,\|\, \mathcal{Q}^f) \geq (\mathcal{P}^f(E))^{c_a} / R_a(\mathcal{P} \| \mathcal{Q})$ by Lem. 2.12, and thus, $\mathcal{P}^f(E) \leq (\mathcal{Q}^f(E) R_a(\mathcal{P} \| \mathcal{Q}))^{1/c_a} \leq 2^{-\frac{1}{c_a}(Cn - \log R_a(\mathcal{P} \| \mathcal{Q}))}$. □

We also define another useful statistical metric called the *max-log distance*:

Definition 2.15 (Max-log distance). *Given two distributions $\mathcal{D}_1$ and $\mathcal{D}_2$ with common support $S = \mathrm{Supp}(\mathcal{D}_1) = \mathrm{Supp}(\mathcal{D}_2)$, the max-log distance between $\mathcal{D}_1$ and $\mathcal{D}_2$ is defined as*

$$\Delta_{\mathrm{ML}}(\mathcal{D}_1, \mathcal{D}_2) := \max_{x \in S} |\ln(\mathcal{D}_1(x)) - \ln(\mathcal{D}_2(x))|.$$

The log of ∞-RD is upper-bounded by the max-log distance since we have $\Delta_{\mathrm{ML}}(\mathcal{D}_1, \mathcal{D}_2) = \max\{\ln(R_\infty(\mathcal{D}_1 \,\|\, \mathcal{D}_2)), \ln(R_\infty(\mathcal{D}_2 \,\|\, \mathcal{D}_1))\}$ by definition. Thus, we have the following fact by Lem. 2.13:

Fact 2.16. *Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be distributions with a common support. For any $a \in [1, \infty]$, $\ln(R_a(\mathcal{D}_1 \,\|\, \mathcal{D}_2)), \ln(R_a(\mathcal{D}_2 \,\|\, \mathcal{D}_1)) \leq \Delta_{\mathrm{ML}}(\mathcal{D}_1, \mathcal{D}_2)$.*

Similarly, the statistical distance is bounded by the bound of ∞-RD:

Fact 2.17. *Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be distributions with a common support. If $R_\infty(\mathcal{D}_1 \| \mathcal{D}_2) \leq 1 + \delta$ and $R_\infty(\mathcal{D}_2 \| \mathcal{D}_1) \leq 1 + \delta$ hold for some $\delta > 0$, then $\Delta(\mathcal{D}_1, \mathcal{D}_2) \leq \delta$.*

2.4 Gaussians

Gaussian function. For a rank-n matrix $\mathbf{S} \in \mathbb{R}^{m\times n}$, the ellipsoid Gaussian function on $\mathbb{R}^n$ with center $\mathbf{c} \in \mathbb{R}^n$ and the (scaled) covariance matrix $\mathbf{\Sigma} = \mathbf{S}^\intercal\mathbf{S}$ is defined as:

$$\rho_{\mathbf{S},\mathbf{c}}(\mathbf{x}) := \exp(-\pi(\mathbf{x}-\mathbf{c})^\intercal(\mathbf{S}^\intercal\mathbf{S})^{-1}(\mathbf{x}-\mathbf{c})).$$

$\rho_{\mathbf{S},\mathbf{c}}(\mathbf{x})$ is determined exactly by $\mathbf{\Sigma} \succ 0$, and there exist a unique $\sqrt{\mathbf{\Sigma}} \succ 0$ s.t. $\sqrt{\mathbf{\Sigma}}\sqrt{\mathbf{\Sigma}} = \mathbf{\Sigma}$, by Lem. 2.3. Thus, we also write $\rho_{\mathbf{S},\mathbf{c}}$ as $\rho_{\sqrt{\mathbf{\Sigma}},\mathbf{c}}$. When $\mathbf{c} = \mathbf{0}$, the function is written as $\rho_{\mathbf{S}}$ or $\rho_{\sqrt{\mathbf{\Sigma}}}$ and is called *centered*. For $\mathbf{S} = s\mathbf{I}_n$, we write $\rho_{\mathbf{S},\mathbf{c}}$ as $\rho_{s,\mathbf{c}}$, and it is written as ρ_s when $\mathbf{c} = \mathbf{0}$.

Continuous Gaussian distribution. We define the continuous multivariate Gaussian distribution and describe its several important properties.

Definition 2.18. *Given $\boldsymbol{\mu} \in \mathbb{R}^m$ and $\mathbf{\Sigma} \in \mathbb{R}^{m\times m}$, we say that $\mathbf{e}$ follows the continuous (ellipsoid) Gaussian distribution $\mathcal{N}_m(\boldsymbol{\mu}, \frac{1}{2\pi}\mathbf{\Sigma})$ if one of the following is satisfied:*

1. *$\mathbf{\Sigma} \succ 0$ and the p.d.f of $\mathbf{e}$ is $\rho_{\sqrt{\mathbf{\Sigma}},\boldsymbol{\mu}}(\mathbf{x})/\sqrt{|\mathbf{\Sigma}|}$.*
2. *$\mathbf{\Sigma} \succeq 0$ and $M_{\mathbf{X}}(\mathbf{t}) := \mathrm{E}[e^{\mathbf{t}^\intercal\mathbf{X}}] = \exp(\boldsymbol{\mu}^\intercal\mathbf{t} + \frac{1}{4\pi}\mathbf{t}^\intercal\mathbf{\Sigma}\mathbf{t})$.*

In particular, we write $\mathcal{N}_m(\boldsymbol{\mu}, \sigma^2) := \mathcal{N}_m(\boldsymbol{\mu}, \sigma^2\mathbf{I}_m)$ for $\sigma > 0$, and call it the continuous spherical Gaussian distribution. We also define $\mathcal{N}_m(\mathbf{\Sigma}) := \mathcal{N}_m(\mathbf{0}, \mathbf{\Sigma})$, $\mathcal{N}_m(\sigma^2) := \mathcal{N}_m(\mathbf{0}, \sigma^2\mathbf{I}_m)$, and $\mathcal{N}(\sigma^2) := \mathcal{N}_1(\sigma^2)$.

Lemma 2.19. *For $\mathbf{e} \sim \mathcal{N}_m(\boldsymbol{\mu}, \mathbf{\Sigma})$, $\mathbf{A} \in \mathbb{R}^{m\times l}$, $\mathbf{b} \in \mathbb{R}^l$, we have $\mathbf{A}^\intercal\mathbf{e} + \mathbf{b} \sim \mathcal{N}_l(\mathbf{A}^\intercal\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}^\intercal\mathbf{\Sigma}\mathbf{A})$.*

Lemma 2.20. *Let $\mathbf{e}_1 \sim \mathcal{N}_m(\boldsymbol{\mu}_1, \mathbf{\Sigma}_1)$ and $\mathbf{e}_2 \sim \mathcal{N}_m(\boldsymbol{\mu}_2, \mathbf{\Sigma}_2)$ be independent. Then, $\mathbf{e}_1 + \mathbf{e}_2 \sim \mathcal{N}_m(\boldsymbol{\mu_1} + \boldsymbol{\mu_2}, \mathbf{\Sigma}_1 + \mathbf{\Sigma}_2)$.*

Lemma 2.21. *Let $\mathbf{e} := (e_1, \cdots, e_m)^\intercal \sim \mathcal{N}_m(\boldsymbol{\mu}, \mathbf{\Sigma})$. If $e_1, \cdots, e_m$ are uncorrelated, i.e., the nondiagonal elements of $\mathbf{\Sigma}$ are all zero, then, $e_1, \cdots, e_m$ are mutually independent. (Thus, the elements of $\mathbf{e} \sim \mathcal{N}_m(\sigma^2)$ are mutually independent).*

Discrete Gaussian distribution. One way to obtain a discrete analog of a continuous Gaussian is by simple rounding. We refer to this as the *discretized Gaussian distribution* $\overline{\Psi}_r$ defined as follows:

Definition 2.22. *For $r > 0$, define $\overline{\Psi}_r$ as the distribution on $\mathbb{Z}_q$ obtained by drawing $y \leftarrow \mathcal{N}(r^2)$ and outputting $\lfloor q \cdot y \rceil \pmod q$.*

The other discrete analog of a Gaussian distribution, which is of greatest concern to us, is the *discrete Gaussian distribution over the lattice*:

Definition 2.23. *For a rank-n lattice $\mathcal{L}$, a matrix $\mathbf{S} \in \mathbb{R}^{m\times n}$, and $\mathbf{c} \in \mathbb{R}^n$, the discrete (ellipsoid) Gaussian distribution with parameter $\mathbf{S}$ and support $\mathcal{L} + \mathbf{c}$ is defined as, $\forall \mathbf{x} \in \mathcal{L} + \mathbf{c}, \mathcal{D}_{\mathcal{L}+\mathbf{c},\mathbf{S}}(\mathbf{x}) = \frac{\rho_{\mathbf{S}}(\mathbf{x})}{\rho_{\mathbf{S}}(\mathcal{L}+\mathbf{c})}$. When $\mathbf{S}^\intercal\mathbf{S} = s^2\mathbf{I}_n$ for some $s > 0$, we write $\mathcal{D}_{\mathcal{L}+\mathbf{c},s}(\mathbf{x})$ and call it the discrete spherical Gaussian distribution.*

Given a lattice $\mathcal{L}$ and $\varepsilon > 0$, we define the smoothing parameter of $\mathcal{L}$ as $\eta_\varepsilon(\mathcal{L}) = \min\{s \mid \rho_{1/s}(\widehat{\mathcal{L}}) \leq 1 + \varepsilon\}$. We also define $\eta_\varepsilon^{\leq}(\mathbb{Z}^n) := \sqrt{\ln(2n(1+1/\varepsilon))/\pi}$, and recall some facts related to the smoothing parameter.

Lemma 2.24 ([GPV08, Lem. 3.1]). *For any n-dimensional full rank lattice $\mathcal{L}$ and real $\varepsilon > 0$, we have $\eta_\varepsilon(\mathcal{L}) \leq \tilde{bl}(\mathcal{L}) \cdot \eta_\varepsilon^{\leq}(\mathbb{Z}^n)$ In particular, for any $\omega(\sqrt{\log n})$ function, there exists a negligible $\varepsilon(n)$ for which $\eta_\varepsilon(\mathcal{L}) \leq \tilde{bl}(\mathcal{L}) \cdot \omega(\sqrt{\log n})$.*

Lemma 2.25 ([MR07, Lem. 4.3]). *For any n-dimensional full rank lattice $\mathcal{L}$, vector $\mathbf{c} \in \mathbb{R}^n$, $\varepsilon \in (0,1)$, and $s \geq 2\eta_\varepsilon(\mathcal{L})$, we have*

$$\mathrm{E}_{\mathbf{x}\sim\mathcal{D}_{\mathcal{L}+\mathbf{c},s}}[\|\mathbf{x}-\mathbf{c}\|^2] \leq (1/2\pi + \varepsilon/(1-\varepsilon))\, s^2 n.$$

Lemma 2.26 ([MR07, Lem. 4.4]). *For any n-dimensional full rank lattice $\mathcal{L}$, vector $\mathbf{c} \in \mathbb{R}^n$, $\varepsilon \in (0,1)$, and $s \geq \eta_\varepsilon(\mathcal{L})$, we have*

$$\Pr_{\mathbf{x}\sim\mathcal{D}_{\mathcal{L}+\mathbf{c},s}}[\|\mathbf{x}-\mathbf{c}\| > s\sqrt{n}] \leq \tfrac{1+\varepsilon}{1-\varepsilon} \cdot 2^{-n}.$$

(Hence, we have $\Pr_{\mathbf{x}\sim\mathcal{D}_{\mathcal{L}+\mathbf{c},s}}[\|\mathbf{x}-\mathbf{c}\|_\infty > s] \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot 2^{-n}$.)

Lemma 2.27 ([Reg09, Claim 3.8]). *For any n-dimensional full rank lattice $\mathcal{L}$, $\mathbf{c} \in \mathbb{R}^n$, $\varepsilon > 0$, and $r \geq \eta_\varepsilon(\mathcal{L})$, $\rho_r(\mathcal{L}+\mathbf{c}) \in (1 \pm \varepsilon) r^n / \det(\mathcal{L})$.*

From Lem. 2.27, when $\mathbf{c} = \mathbf{0}$ and $\mathcal{L} = \mathbf{S}^{-\intercal}\mathbb{Z}^n$, where $\mathbf{S} \in \mathbb{R}^{n\times n}$ is a non-singular matrix, we obtain the following corollary:

Corollary 2.28. *For any nonsingular matrix $\mathbf{S} \in \mathbb{R}^{n\times n}$, any $\varepsilon > 0$, and $r \geq \eta_\varepsilon(\mathbf{S}^{-\intercal}\mathbb{Z}^n)$, $\rho_{r\mathbf{S}}(\mathbb{Z}^n) \in (1 \pm \varepsilon) r^n |\mathbf{S}|$.*

In addition, by Lem. 2.27, we obtain a discrete analog of Lem. 2.21:

Lemma 2.29. *Let $n \in \mathbb{N}$. For any $\varepsilon > 0$ and $r \geq \eta_\varepsilon(\mathbb{Z}^n)$, we have $\Delta(\mathcal{D}_{\mathbb{Z}^n,r}, (\mathcal{D}_{\mathbb{Z},r})^n) = \hat{\varepsilon}$ and $\Delta_{\mathrm{ML}}(\mathcal{D}_{\mathbb{Z}^n,r}, (\mathcal{D}_{\mathbb{Z},r})^n) = \ln(1+\hat{\varepsilon})$, where $\hat{\varepsilon} := \varepsilon + O(\varepsilon^2)$.*

3 Approximately Orthogonal Matrices

The main goal of this section (Thm. 3.5) is to introduce a construction of the *approximately orthogonal matrix*, defined as follows:

Definition 3.1. *Let $\mathbf{X} \in \mathbb{R}^{m\times n}$, and define a residual matrix $\mathbf{R} := \mathbf{X}^\intercal\mathbf{X} - \mathbf{I}_n := [r_{ij}]$. We say that $\mathbf{X}$ is approximately orthogonal with bound $\delta > 0$ if $|r_{ij}| < \delta$ holds for all $i, j \in [n]$.*

Then, we apply Thm. 3.5 to obtain Cor. 3.6 and Lem. 3.7. They are the building blocks for the proofs of Thm. 4.2 and Lem. 5.3 in Sect. 4 and Sect. 5. To begin, we derive some facts regarding centered and bounded distributions.

Fact 3.2. *Let X and Y be centered, β-bounded for $\beta > 0$, and mutually independent. Then, XY is centered and β^2-bounded.*

Lemma 3.3. *Let $X_1, X_2, \ldots, X_n$ be centered, β-bounded for $\beta > 0$, and mutually independent. Let $C > 0$ be a constant, and define $\overline{X} := \frac{1}{n}\sum_{i=1}^n CX_i$. Then, for $\varepsilon > 0$, we have $\Pr[|\overline{X}| \geq \varepsilon] < 2\exp\left(-\frac{1}{2C^2\beta^2}\varepsilon^2 n\right)$.*

Proof. This follows from the Hoeffding bound. □

Lemma 3.4. *Let χ_β be a centered and β-bounded distribution for $\beta > 0$ with $\mathrm{V}[\chi_\beta] := \varsigma^2$. Let $X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \chi_\beta$, $Y_i := \frac{X_i^2}{\varsigma^2}$ for $i \in [n]$, and define $\overline{Y} := \frac{1}{n}\sum_{i=1}^n Y_i$. Then for $\varepsilon > 0$, we have $\Pr[|\overline{Y} - 1| \geq \varepsilon] < 2\exp\left(-2\frac{\varsigma^4}{\beta^4}\varepsilon^2 n\right)$.*

Proof. We have $\mathrm{E}[\overline{Y}] = \frac{1}{n}\sum_{i=1}^n \mathrm{E}[\frac{X_i^2}{\varsigma^2}] = \frac{1}{n}\sum_{i=1}^n \frac{1}{\varsigma^2}\mathrm{V}[X_i] = 1$, and $\mathrm{Supp}(Y_i) \subset [0, \frac{\beta^2}{\varsigma^2}]$. Thus, the lemma follows from the Hoeffding bound. □

Now, we show the construction of the approximately orthogonal matrix.

Theorem 3.5. *Let $m \in \mathbb{N}$ be a security parameter, and let $l = \mathrm{poly}(m)$ be a positive integer. Let χ_β be a centered and β-bounded distribution for $\beta > 0$ with $\mathrm{V}[\chi_\beta] := \varsigma^2$, and assume $\beta/\varsigma = O(1)$. Let $\mathbf{X} \sim \chi_\beta^{m\times l}$, then, for any constants $\gamma \in (0, 1/2)$ and $c > 0$, $\left(\frac{1}{\sqrt{m}\varsigma}\mathbf{X}\right)$ is an approximately orthogonal matrix with bound $\delta := c \cdot m^{-\gamma}$, with overwhelming probability over the choice of $\mathbf{X}$.*

Proof. We analyze the distribution of $\mathbf{S} := \left(\frac{1}{m\varsigma^2}\mathbf{X}^\intercal\mathbf{X}\right) = [s_{ij}]$. Let $\mathbf{X}_i$ for $i \in [l]$ be an i-th column vector of $\mathbf{X} = [x_{ij}]$. For the nondiagonal elements, i.e., when $i \neq j$, we have $s_{ij} = \frac{1}{m\varsigma^2}(\mathbf{x}_i)^\intercal\mathbf{x}_j = \frac{1}{m}\sum_{k=1}^m \frac{1}{\varsigma^2}x_{ik}x_{jk}$ by definition. Since $x_{ik}, x_{jk} \overset{\text{iid}}{\sim} \chi_\beta$ for all $k \in [m]$, by Fact 3.2, we have that $(x_{i1}x_{j1}), \ldots, (x_{im}x_{jm})$ are centered, β^2-bounded, and mutually independent. Thus, by Lem. 3.3 (with $C := \frac{1}{\varsigma^2}$), we have $\Pr[|s_{ij}| \geq \delta] < 2\exp\left(-\frac{1}{2}\frac{\varsigma^4}{\beta^4}\delta^2 m\right) = \mathrm{negl}(m)$ for all $i \neq j$. Hence, by the union bound, $\Pr[\bigcup_{i\neq j}(|s_{ij}| \geq \delta)] \leq \sum_{i\neq j}\Pr[|s_{ij}| \geq \delta] = \mathrm{negl}(m)$ holds. Thus, we have

$$\Pr[\textstyle\bigcap_{i\neq j}(|s_{ij}| < \delta)] \geq 1 - \mathrm{negl}(m).$$

Next, for the diagonal elements, we have $s_{ii} = \frac{1}{m\varsigma^2}\|\mathbf{x}_i\|^2 = \frac{1}{m}\sum_{k=1}^m \frac{x_{ik}^2}{\varsigma^2}$ by definition. Since $x_{i1}, \ldots, x_{im} \overset{\text{iid}}{\sim} \chi_\beta$ by Lem. 3.4, we have $\Pr[|s_{ii} - 1| \geq \delta] < 2\exp\left(-2\frac{\varsigma^4}{\beta^4}\delta^2 m\right) = \mathrm{negl}(m)$. Thus, by the union bound, $\Pr[\bigcup_{i\in[l]}(|s_{ii} - 1| \geq \delta)] \leq \sum_{i\in[l]}\Pr[|s_{ii} - 1| \geq \delta] = \mathrm{negl}(m)$ holds, and we have

$$\Pr[\textstyle\bigcap_{i\in[l]}(|s_{ii} - 1| < \delta)] \geq 1 - \mathrm{negl}(m).$$

Since the residual matrix is $\mathbf{R} = \mathbf{S} - \mathbf{I}_l$, we obtain the theorem. □

The upper bound on the absolute values of the elements of the residual matrix $\mathbf{R}$ enables useful analysis. We can bound $|\mathbf{I}_l + \mathbf{R}|$ by Lem. 2.7:

Corollary 3.6. *In Thm. 3.5, let the residual matrix $\mathbf{R} := \frac{1}{m\varsigma^2}\mathbf{X}^\intercal\mathbf{X} - \mathbf{I}_l$, then for any $l < 1/\delta$ $(= \frac{m^\gamma}{c})$, $1 - l\delta \leq |\mathbf{I}_l + \mathbf{R}| \leq 1/(1 - l\delta)$ holds with overwhelming probability over the choice of $\mathbf{X}$.*

Furthermore, we can analyze the positive definiteness of a matrix in the form $\mathbf{S} := \mathbf{I}_l - k\mathbf{R}$ for small $k \in \mathbb{R}$ when the elements of $\mathbf{R}$ have a small bound.

Lemma 3.7. *Let* $\mathbf{R} = [r_{ij}] \in \mathbb{R}^{l \times l}$ *be a symmetric matrix s.t.* $|r_{ij}| \leq \delta$ *for* $i, j \in [n]$*, and* $\delta \in \mathbb{R}$*. For any* $2 \leq l \in \mathbb{N}$ *and* δ*,* $k \in \mathbb{R}$ *s.t.* $|k|l\delta < 1$*,* $\mathbf{S} := \mathbf{I}_l - k\mathbf{R} \succ 0$ *holds.*

Proof. $\mathbf{S} = [s_{ij}]$ is a diagonally dominant matrix, since $\sum_{j \neq i} |s_{ij}| < (l-1)|k|\delta < 1 - |k|\delta \leq |s_{ii}|$ holds for all $i \in [l]$. All diagonal elements of $\mathbf{S}$ are positive; i.e., $s_{ii} > 0$ holds for all $i \in [l]$ since $|kr_{ii}| < |k|\delta = 1/l < 1$. Thus, Lem. 2.6 is applicable to $\mathbf{S}$, and the lemma follows. □

4 Continuous Weak Spherical Gaussian LHL

The goal of this section is to show the continuous wSGLHL (Thm. 4.2) and its extended theorem (Thm. 4.4) with the noise flooding technique. Let $\mathbf{e}' := \mathbf{X}^\intercal \mathbf{e}$, where $\mathbf{X} \in \mathbb{R}^{m \times l}$ and $\mathbf{e} \sim \mathcal{N}_m(\sigma^2)$ is a continuous spherical Gaussian. Then, have $\mathbf{e}' \sim \mathcal{N}_l(\sigma^2 \mathbf{\Sigma})$, where $\mathbf{\Sigma} := \mathbf{X}^\intercal \mathbf{X}$, by Lem. 2.19. We instantiate $\mathbf{X} \sim \chi_\beta^{m \times l}$ and define $\mathbf{R} := \frac{1}{m\varsigma^2}\mathbf{\Sigma} - \mathbf{I}_l$, then we obtain a small bound on the elements of $\mathbf{R}$ by Thm. 3.5. In Sect. 4.1, we show that $R_a(\mathcal{N}_l(\sigma^2\mathbf{\Sigma}), \mathcal{N}_l(m\varsigma^2\sigma^2\mathbf{I}_l))$ is small, which is the continuous wSGLHL (Thm. 4.2). In Sect. 4.2, we present an improved theorem (Thm. 4.4) that supports an arbitrarily large l, which is restricted to $l < \sqrt{m}$ in Thm. 4.2.

4.1 (Plain) Continuous Weak Spherical Gaussian LHL

In this subsection, we present the continuous wSGLHL (Thm. 4.2). We first show that RD between $\mathbf{e}' := \mathbf{X}^\intercal \mathbf{e}$ with the general (column full rank) matrix $\mathbf{X}$ and continuous spherical Gaussian can be written with a simpler formula.

Lemma 4.1. *Let* $m \geq l \in \mathbb{N}$*, and let* $\mathbf{X} \in \mathbb{R}^{m \times l}$ *be a column full rank matrix. Define* $\mathbf{\Sigma} := \mathbf{X}^\intercal \mathbf{X}$*, and let* $\mathbf{R} := \frac{1}{s^2}\mathbf{\Sigma} - \mathbf{I}_l$ *for* $s \in \mathbb{R}^+$*.*

For any $\mathbf{X}$*,* s*, and* $a \in (1, \infty)$ *s.t.* $\mathbf{I}_l - (a-1)\mathbf{R} \succ 0$*,*

$$\overline{R_a} := R_a(\mathbf{X}^\intercal \mathcal{N}_m(\sigma^2) \parallel \mathcal{N}_l(s^2\sigma^2)) = 1/\sqrt{|\mathbf{I}_l + \mathbf{R}||\mathbf{I}_l - (a-1)\mathbf{R}|^{\frac{1}{a-1}}}. \quad (1)$$

For any $\mathbf{X}$ *and* s *s.t.* $-\mathbf{R} \succeq 0$*,*

$$\overline{R_\infty} := R_\infty(\mathbf{X}^\intercal \mathcal{N}_m(\sigma^2) \parallel \mathcal{N}_l(s^2\sigma^2)) = 1/\sqrt{|\mathbf{I}_l + \mathbf{R}|}. \quad (2)$$

Proof. Since $\mathbf{X}$ is column full rank, we have $\mathbf{\Sigma} \succ 0$ by Lem. 2.4. Thus, by Lem. 2.19, $\mathbf{X}^\intercal \mathcal{N}_m(\sigma^2) \sim \mathcal{N}_l(\sigma^2\mathbf{\Sigma})$. In addition, from Lem. 2.3, there exists a symmetric matrix $\sqrt{\mathbf{\Sigma}} \in \mathbb{R}^{l \times l}$ s.t. $\sqrt{\mathbf{\Sigma}}\sqrt{\mathbf{\Sigma}} = \mathbf{\Sigma}$ and $\sqrt{\mathbf{\Sigma}} \succ 0$. Note that $\mathbf{\Sigma}^{-1} \succ 0$. Then, for any $a \in (1, \infty)$, we have

$$\begin{aligned}
(\overline{R_a})^{a-1} &= \int_{\mathbf{y} \in \mathbb{R}^l} \left(\frac{\exp\left(-\frac{1}{2\sigma^2}\mathbf{y}^\intercal \mathbf{\Sigma}^{-1}\mathbf{y}\right)}{\sqrt{(2\pi\sigma^2)^l|\mathbf{\Sigma}|}} \right)^a \left(\frac{\exp\left(-\frac{1}{2s^2\sigma^2}\mathbf{y}^\intercal \mathbf{y}\right)}{\sqrt{(2\pi s^2\sigma^2)^l}} \right)^{-(a-1)} d\mathbf{y} \\
&= \int_{\mathbf{y} \in \mathbb{R}^l} \frac{\exp\left(-\frac{1}{2\sigma^2}\left(\mathbf{y}^\intercal (a\mathbf{\Sigma}^{-1} - \frac{a-1}{s^2}\mathbf{I}_l)\mathbf{y}\right)\right)}{\sqrt{(2\pi\sigma^2/s^{2(a-1)})^l|\mathbf{\Sigma}|^a}} d\mathbf{y}
\end{aligned}$$

By defining $\mathbf{y} := \sqrt{\mathbf{\Sigma}}\mathbf{x}$, we have $d\mathbf{y} = |\sqrt{\mathbf{\Sigma}}|d\mathbf{x}$, and thus

$$\begin{aligned}(\overline{R_a})^{a-1} &= \int_{\mathbf{x}\in\mathbb{R}^l} \frac{\exp(-\frac{1}{2\sigma^2}(\mathbf{x}^\intercal(a\mathbf{I}_l - \frac{a-1}{s^2}\mathbf{\Sigma})\mathbf{x}))}{\sqrt{(2\pi\sigma^2)^l|\frac{1}{s^2}\mathbf{\Sigma}|^{a-1}}}d\mathbf{x} \\ &= \int_{\mathbf{x}\in\mathbb{R}^l} \frac{\exp(-\frac{1}{2\sigma^2}(\mathbf{x}^\intercal(\mathbf{I}_l - (a-1)\mathbf{R})\mathbf{x}))}{\sqrt{(2\pi\sigma^2)^l|\mathbf{I}_l + \mathbf{R}|^{a-1}}}d\mathbf{x} \\ &= \frac{1}{\sqrt{|\mathbf{I}_l + \mathbf{R}|^{a-1}|\mathbf{I}_l - (a-1)\mathbf{R}|}} \int_{\mathbf{x}\in\mathbb{R}^l} \frac{\exp(-\frac{1}{2\sigma^2}(\mathbf{x}^\intercal(\mathbf{I}_l - (a-1)\mathbf{R})\mathbf{x}))}{\sqrt{(2\pi\sigma^2)^l|(\mathbf{I}_l - (a-1)\mathbf{R})^{-1}|}}d\mathbf{x}.\end{aligned}$$

Hence, for any $\mathbf{X}$, s, and $a \in (1,\infty)$ s.t. $\mathbf{I}_l - (a-1)\mathbf{R} \succ 0$, we obtain (1) because $\frac{\exp(-\frac{1}{2\sigma^2}(\mathbf{x}^\intercal(\mathbf{I}_l-(a-1)\mathbf{R})\mathbf{x}))}{\sqrt{(2\pi\sigma^2)^l|(\mathbf{I}_l-(a-1)\mathbf{R})^{-1}|}}$ is the p.d.f of $\mathcal{N}_l(\sigma^2(\mathbf{I}_l - (a-1)\mathbf{R})^{-1})$.

Similarly, for $a = \infty$, we have

$$\begin{aligned}\overline{R_\infty} &= \max_{\mathbf{y}\in\mathbb{R}^l}\left(\frac{\exp\left(-\frac{1}{2\sigma^2}\mathbf{y}^\intercal\mathbf{\Sigma}^{-1}\mathbf{y}\right)}{\sqrt{(2\pi\sigma^2)^l|\mathbf{\Sigma}|}} \Big/ \frac{\exp\left(-\frac{1}{2s^2\sigma^2}\mathbf{y}^\intercal\mathbf{y}\right)}{\sqrt{(2\pi s^2\sigma^2)^l}}\right) \\ &= \frac{1}{\sqrt{|\mathbf{I}_l + \mathbf{R}|}}\max_{\mathbf{x}\in\mathbb{R}^l}\exp\left(-\frac{1}{2\sigma^2}\mathbf{x}^\intercal(-\mathbf{R})\mathbf{x}\right).\end{aligned}$$

Thus, if $-\mathbf{R} \succeq 0$, we obtain (2). □

By sampling $\mathbf{X}$ from a centered and bounded distribution, and applying Thm. 3.5, we obtain the continuous wSGLHL:

Theorem 4.2 (Continuous wSGLHL). *Let $m \in \mathbb{N}$ be a security parameter, $\gamma \in (\frac{1}{\log m}, \frac{1}{2})$ be a constant, and $l := l(m) < m^\gamma$ be a positive integer. Let χ_β be a centered and β-bounded distribution for $\beta > 0$ with $\mathrm{V}[\chi_\beta] := \varsigma^2$, and assume $\beta/\varsigma = O(1)$. Let $\mathbf{X} \sim \chi_\beta^{m\times l}$. Then, for any constant $a \in [2,\infty)$, with overwhelming probability over the choice of $\mathbf{X}$, we have*

$$\overline{R_a} := R_a(\mathbf{X}^\intercal\mathcal{N}_m(\sigma^2) \,\|\, \mathcal{N}_l(m\varsigma^2\sigma^2)) < \left(1 + 1/(\tfrac{m^\gamma}{l} - 1)\right)^{\frac{1}{a-1}}.$$

Proof. Define $\mathbf{\Sigma} := \mathbf{X}^\intercal\mathbf{X}$ and $\mathbf{R} := \frac{1}{m\varsigma^2}\mathbf{\Sigma} - \mathbf{I}_l$. Let $\delta := \frac{1}{(a-1)m^\gamma}$; then, by Thm. 3.5, all elements of $\mathbf{R} = [r_{ij}]$ simultaneously satisfy $|r_{ij}| < \delta$ for all $i, j \in [l]$ with overwhelming probability over the choice of $\mathbf{X}$.

Let $\mathbf{S}_a := \mathbf{I}_l - (a-1)\mathbf{R}$. By construction, $(a-1)l\delta < 1$ holds, and $\mathbf{S}_a$ is a symmetric matrix since $\mathbf{\Sigma}$ is symmetric (as is $\mathbf{R}$). Hence, by Lem. 3.7, $\mathbf{S}_a \succ 0$ holds with overwhelming probability over the choice of $\mathbf{X}$. Similarly, $\mathbf{\Sigma} = m\varsigma^2(\mathbf{I}_l + \mathbf{R}) \succ 0$ holds with overwhelming probability. Therefore, we have $\overline{R_a} = 1/\sqrt{|\mathbf{I}_l + \mathbf{R}||\mathbf{I}_l - (a-1)\mathbf{R}|^{\frac{1}{a-1}}}$ by Lem. 4.1. Finally, we analyze the upper bound on $\overline{R_a}$. Since $l\delta < (a-1)l\delta < 1$ holds, we have $|\mathbf{I}_l + \mathbf{R}| > 1 - l\delta$ and $|\mathbf{I}_l - (a-1)\mathbf{R}| > 1 - (a-1)l\delta$ by Lem. 2.7, and $(1 - l\delta) > (1 - (a-1)l\delta)^{\frac{1}{a-1}}$. Hence, we have $\overline{R_a} < 1/\sqrt{(1-l\delta)(1-(a-1)l\delta)^{\frac{1}{a-1}}} < 1/(1-(a-1)l\delta)^{\frac{1}{a-1}} = \left(1/(1-\frac{l}{m^\gamma})\right)^{\frac{1}{a-1}} = \left(1 + 1/(\frac{m^\gamma}{l} - 1)\right)^{\frac{1}{a-1}}$. □

Note that, in Thm. 4.2, we could not derive quantitative bound of $\overline{R_\infty}$ of (2) derived in Lem. 4.1. We need $-\mathbf{R} \succeq 0$ to use Lem. 4.1, but this condition does not necessary holds (with overwhelming probability) when $\mathbf{X} \sim \chi_\beta^{m \times l}$. We require additional condition on $\mathbf{X}$: as a trivial example, (exactly) orthogonal matrices $\mathbf{X}$ satisfy $\mathbf{R} = \mathbf{X}^\intercal \mathbf{X} - \mathbf{I} = \mathbf{O} \succeq 0$.

4.2 Improvement with Noise Flooding

In Thm. 4.2, the number of outputs l must be less than $\sqrt{m}$. We extend l to an arbitrarily large number by adding extra Gaussian errors to the linear sums of the Gaussian errors, in Thm. 4.4 of this section. This technique is conceptually similar to the technique called "noise flooding" that is used in [BGM+16, BLR+18]. First, we perform analysis with the general matrix $\mathbf{X} \in \mathbb{R}^{m \times l}$ as in Lem. 4.1. Note that we do not require $m \geq l$ here, unlike Lem. 4.1.

Lemma 4.3. *Let $m, l \in \mathbb{N}$ and k, $s \in \mathbb{R}^+$. We define $\mathbf{X} \in \mathbb{R}^{m \times l}$, $\mathbf{\Sigma} := \mathbf{X}^\intercal \mathbf{X}$, and assume $\mathbf{\Sigma}' := \mathbf{\Sigma} + ks^2 \mathbf{I}_l \succ 0$. Let $\mathbf{R} := \frac{1}{s^2(1+k)} \mathbf{\Sigma}' - \mathbf{I}_l$.*

For any $\mathbf{X}$, s, and $a \in (1, \infty)$ s.t. $\mathbf{I}_l - (a-1)\mathbf{R} \succ 0$,

$$\begin{aligned} \overline{R_a} &:= R_a(\mathbf{X}^\intercal \mathcal{N}_m(\sigma^2) + \mathcal{N}_l(ks^2\sigma^2) \parallel \mathcal{N}_l((1+k)s^2\sigma^2)) \\ &= 1/\sqrt{|\mathbf{I}_l + \mathbf{R}||\mathbf{I}_l - (a-1)\mathbf{R}|^{\frac{1}{a-1}}}. \end{aligned}$$

For any $\mathbf{X}$ and s s.t. $-\mathbf{R} \succ 0$,

$$\overline{R_\infty} := R_\infty(\mathbf{X}^\intercal \mathcal{N}_m(\sigma^2) + \mathcal{N}_l(ks^2\sigma^2) \parallel \mathcal{N}_l((1+k)s^2\sigma^2)) = 1/\sqrt{|\mathbf{I}_l + \mathbf{R}|}.$$

Proof. By definition, $\mathbf{\Sigma} \succeq 0$ holds[5]. Thus, $\mathbf{X}^\intercal \mathcal{N}_m(\sigma^2) = \mathcal{N}_l(\sigma^2 \mathbf{\Sigma})$ by Lem. 2.19, and therefore, $\mathbf{X}^\intercal \mathcal{N}_m(\sigma^2) + \mathcal{N}_l(ks^2\sigma^2) = \mathcal{N}_l(\sigma^2 \mathbf{\Sigma}')$ by Lem. 2.20. By hypothesis, we have $\mathbf{\Sigma}' \succ 0$, and thus there exists a unique $\sqrt{\mathbf{\Sigma}'} \in \mathbb{R}^{l \times l}$ s.t. $\sqrt{\mathbf{\Sigma}'} \succ 0$ and $\sqrt{\mathbf{\Sigma}'}\sqrt{\mathbf{\Sigma}'} = \mathbf{\Sigma}'$ by Lem. 2.3. Hence, we have:

$$\begin{aligned} (\overline{R_a})^{a-1} &= (R_a(\mathcal{N}_l(\sigma^2 \mathbf{\Sigma}') \parallel \mathcal{N}_l((1+k)s^2\sigma^2)))^{a-1} \\ &= \int_{\mathbf{y} \in \mathbb{R}^l} \left(\frac{\exp\left(-\frac{1}{2\sigma^2} \mathbf{y}^\intercal (\mathbf{\Sigma}')^{-1} \mathbf{y}\right)}{\sqrt{(2\pi\sigma^2)^l |\mathbf{\Sigma}'|}} \right)^a \left(\frac{\exp\left(-\frac{1}{2(1+k)s^2\sigma^2} \mathbf{y}^\intercal \mathbf{y}\right)}{\sqrt{(2\pi(1+k)s^2\sigma^2)^l}} \right)^{-(a-1)} d\mathbf{y} \\ &= \int_{\mathbf{y} \in \mathbb{R}^l} \frac{\exp\left(-\frac{1}{2\sigma^2} \mathbf{y}^\intercal (a(\mathbf{\Sigma}')^{-1} - \frac{a-1}{(1+k)s^2} \mathbf{I}_l) \mathbf{y}\right)}{\sqrt{(2\pi\sigma^2((1+k)s^2)^{-(a-1)})^l |\mathbf{\Sigma}'|^a}} d\mathbf{y}. \end{aligned}$$

[5] $\mathbf{\Sigma} \succ 0$ does not necessarily hold since $\mathbf{X}$ is not necessarily column full rank (l may be larger than m) in this lemma.

By defining $\mathbf{y} := \sqrt{\mathbf{\Sigma}'}\mathbf{x}$, we have $d\mathbf{y} = |\sqrt{\mathbf{\Sigma}'}|d\mathbf{x} = |\mathbf{\Sigma}'|^{1/2}d\mathbf{x}$. Then, we have

$$\begin{aligned}(\overline{R_a})^{a-1} &= \int_{\mathbf{x}\in\mathbb{R}^l} \frac{\exp\left(-\frac{1}{2\sigma^2}\mathbf{x}^\intercal(a\mathbf{I}_l - \frac{a-1}{(1+k)s^2}\mathbf{\Sigma}')\mathbf{x}\right)}{\sqrt{(2\pi\sigma^2)^l|\frac{1}{(1+k)s^2}\mathbf{\Sigma}'|^{a-1}}}d\mathbf{x} \\ &= \int_{\mathbf{x}\in\mathbb{R}^l} \frac{\exp\left(-\frac{1}{2\sigma^2}\mathbf{x}^\intercal(\mathbf{I}_l - (a-1)\mathbf{R})\mathbf{x}\right)}{\sqrt{(2\pi\sigma^2)^l|\mathbf{I}_l + \mathbf{R}|^{a-1}}}d\mathbf{x}.\end{aligned}$$

The rest of the proof is identical to that of Lem. 4.1. We can derive $\overline{R_\infty}$ similarly. □

By sampling $\mathbf{X}$ from a centered and bounded distribution and applying Thm. 3.5, we obtain an extension of Thm. 4.2. This theorem subsumes Thm. 4.2 since they are identical when $k \to 0$.

Theorem 4.4 (Extended continuous wSGLHL). *Let $m \in \mathbb{N}$ be a security parameter. Let $\gamma \in (\frac{1}{\log m}, \frac{1}{2})$ be a constant, and define $k := k(m) > 0$ and $l := l(m) < (1+k)m^\gamma$. Let χ_β be a centered and β-bounded distribution for $\beta > 0$ with $\mathrm{V}[\chi_\beta] := \varsigma^2$, and assume $\beta/\varsigma = O(1)$. Let $\mathbf{X} \sim \chi_\beta^{m\times l}$. Then, for any constant $a \in [2,\infty)$, with overwhelming probability over the choice of $\mathbf{X}$, we have*

$$\begin{aligned}\overline{R_a} &:= R_a(\mathbf{X}^\intercal\mathcal{N}_m(\sigma^2) + \mathcal{N}_l(km\varsigma^2\sigma^2) \parallel \mathcal{N}_l((1+k)m\varsigma^2\sigma^2)) \\ &< \left(1 + 1/((1+k)\tfrac{m^\gamma}{l} - 1)\right)^{\frac{1}{a-1}}.\end{aligned}$$

Proof. Define $\mathbf{\Sigma} := \mathbf{X}^\intercal\mathbf{X}$, $\mathbf{\Sigma}' := \mathbf{\Sigma} + km\varsigma^2\mathbf{I}_l$, and $\mathbf{R} := \frac{1}{m\varsigma^2(1+k)}\mathbf{\Sigma}' - \mathbf{I}_l$. Let $\mathbf{R}' := \frac{1}{m\varsigma^2}\mathbf{\Sigma} - \mathbf{I}_l$; then, $\mathbf{R} = \frac{1}{1+k}\mathbf{R}'$. Let $\delta := \frac{1}{(a-1)m^\gamma}$; then, by Thm. 3.5, all elements of $\mathbf{R}' = [r'_{ij}]$ simultaneously satisfy $|r'_{ij}| < \delta$ with overwhelming probability over the choice of $\mathbf{X}$. Thus, all elements of $\mathbf{R} = [r_{ij}]$ simultaneously satisfy $|r_{ij}| < \frac{\delta}{1+k}$ with overwhelming probability. Let $\mathbf{S}_a := \mathbf{I}_l - (a-1)\mathbf{R}$. By construction, $\frac{(a-1)l\delta}{1+k} < 1$, and $\mathbf{S}_a$ is a symmetric matrix. Hence, by Lem. 3.7, $\mathbf{S}_a \succ 0$ holds with overwhelming probability. Similarly, we can show that $\mathbf{I}_l + \mathbf{R} \succ 0$ and thus that $\mathbf{\Sigma}'(= (1+k)m\varsigma^2(\mathbf{I}_l + \mathbf{R})) \succ 0$ holds with overwhelming probability. Hence, by Lem. 4.3, we obtain $\overline{R_a} = 1/\sqrt{|\mathbf{I}_l + \mathbf{R}||\mathbf{I}_l - (a-1)\mathbf{R}|^{\frac{1}{a-1}}}$. Finally, we analyze the upper bound on $\overline{R_a}$. Since $\frac{l\delta}{1+k} < \frac{(a-1)l\delta}{1+k} < 1$ holds, we have $1 - \frac{l\delta}{1+k} < |\mathbf{I}_l + \mathbf{R}|$ and $1 - \frac{(a-1)l\delta}{1+k} < |\mathbf{I}_l - (a-1)\mathbf{R}|$ by Lem. 2.7, as well as $(1 - \frac{l\delta}{1+k}) > (1 - \frac{(a-1)l\delta}{1+k})^{\frac{1}{a-1}}$. Hence, we have $\overline{R_a} < 1/\sqrt{(1 - \frac{l\delta}{1+k})(1 - \frac{(a-1)l\delta}{1+k})^{\frac{1}{a-1}}} < 1/(1 - \frac{(a-1)l\delta}{1+k})^{\frac{1}{a-1}} = 1/(1 - \frac{l}{(1+k)m^\gamma})^{\frac{1}{a-1}}$, and the theorem follows. □

5 Discrete Weak Spherical Gaussian LHL

The goal of this section is to show the discrete wSGLHL (Thm. 5.4), which is a discrete analog of Thm. 4.2. The proof of this theorem is conceptually the same

as that of Thm. 4.2. We analyze RD between $\mathbf{e}' := \mathbf{X}^\intercal \mathbf{e}$ and a discrete spherical Gaussian, where $\mathbf{e} \sim \mathcal{D}_{\mathbb{Z}^m,r}$ is a discrete spherical Gaussian and $\mathbf{X} \in \mathbb{R}^{m\times l}$.

This analysis is more complicated than that for the continuous Gaussian. Although the linear transformation of the multivariate continuous Gaussian is exactly a multivariate continuous Gaussian as shown in Lem. 2.19, the counterpart of the discrete multivariate Gaussian is not trivial; it was first shown by Agrawal *et al.* in [AGHS13]. Similar analyses were performed in [AR16, CGM19, DGPY20, GMPW20] (with some generalization). We rely on the lemma given by Aggarwal and Regev [AR16], which improves upon [AGHS13][6].

Lemma 5.1 (Adapted from [AR16, Thm. 5.1]). *Let $m > l \geq 100$ be integers, and let $\varepsilon = \varepsilon(l) \in (0, 10^{-3})$. Let $\varsigma \in \mathbb{R}^+$ and let $\mathbf{X} \sim (\mathcal{D}_{\mathbb{Z}^m,\varsigma})^l$.*

If $m \geq 30l\log(\varsigma l)$, $r \geq 10\varsigma l \log m\sqrt{\log(1/\varepsilon)\log(\varsigma l)}$, and $\varsigma \geq 9\eta^{\leq}_{\tilde{\varepsilon}}(\mathbb{Z}^l)$, then, with probability $1 - 2^{-l}$ over the choice of $\mathbf{X}$, for any $\mathbf{z} \in \mathbb{Z}^l$, $(\mathbf{X}^\intercal \mathcal{D}_{\mathbb{Z}^m,r})(\mathbf{z}) \in \left[\frac{1-\varepsilon}{1+\varepsilon}, 1\right] \cdot \mathcal{D}_{\mathbb{Z}^l,r\mathbf{X}}(\mathbf{z})$ holds, and thus we have $\Delta_{\mathrm{ML}}(\mathbf{X}^\intercal \mathcal{D}_{\mathbb{Z}^m,r}, \mathcal{D}_{\mathbb{Z}^l,r\mathbf{X}}) \leq \ln(\frac{1+\varepsilon}{1-\varepsilon})$.

Note that this lemma states only that the linear transformation of the discrete spherical Gaussian is a discrete *ellipsoid* Gaussian. We will show in Lem. 5.3 that when we take $\mathbf{X}$ as a (scaled) approximately orthogonal matrix, namely, $\mathbf{X} \sim \chi_\beta^{m\times l}$, the discrete ellipsoid Gaussian $\mathcal{D}_{\mathbb{Z}^l,r\mathbf{X}}$ can be approximated as a discrete spherical Gaussian. Although Lem. 5.1 samples $\mathbf{X}$ from the discrete spherical Gaussian distribution $(\mathcal{D}_{\mathbb{Z}^m,\varsigma})^l$, we can show that the Gaussian distribution[7] is also a bounded distribution with overwhelming probability, by using the standard bound Lem. 2.26[8]. Thus, Lem. 5.1 is compatible with our framework based on Thm. 3.5.

We first show the discrete analog of Lem. 4.1. (Here, recall the notational shortcuts $c_a := \frac{a}{a-1}$ for $a > 1$ and $\hat{\varepsilon} := \varepsilon + O(\varepsilon^2)$.)

Lemma 5.2. *Let $m \geq l \in \mathbb{N}$, and let $\mathbf{X} \in \mathbb{Z}^{m\times l}$ be a column full rank matrix. Define $\mathbf{\Sigma} := \mathbf{X}^\intercal\mathbf{X}$, and let $\mathbf{R} := \frac{1}{s^2}\mathbf{\Sigma} - \mathbf{I}_l$ for $s \in \mathbb{R}^+$. Let $a \in [2,\infty)$ be a constant, let $\mathbf{S}_a := \mathbf{I}_l - (a-1)\mathbf{R}$, and assume that $\mathbf{S}_a \succ 0$ holds. For any $\varepsilon \in \mathbb{R}^+$ and $r > \eta_\varepsilon(\sqrt{\mathbf{\Sigma}}^{-1}\sqrt{\mathbf{S}_a}\mathbb{Z}^l)$, we have*

$$\overline{R_a} := R_a(\mathcal{D}_{\mathbb{Z}^l,r\mathbf{X}} \parallel \mathcal{D}_{\mathbb{Z}^l,rs}) \leq (1 + 2c_a\hat{\varepsilon})/\sqrt{|\mathbf{I}_n + \mathbf{R}||\mathbf{I}_n - (a-1)\mathbf{R}|^{\frac{1}{a-1}}}.$$

Proof. Since $\mathbf{X}$ is column full rank, we have $\mathbf{\Sigma} \succ 0$ by Lem. 2.4. Hence, by Lem. 2.3 there exists a symmetric matrix $\sqrt{\mathbf{\Sigma}} \in \mathbb{R}^{l\times l}$ s.t. $\sqrt{\mathbf{\Sigma}}\sqrt{\mathbf{\Sigma}} = \mathbf{\Sigma}$ and $\sqrt{\mathbf{\Sigma}} \succ 0$. By the hypothesis that $\mathbf{S}_a \succ 0$, there exists a symmetric matrix $\sqrt{\mathbf{S}_a} \in \mathbb{R}^{n\times n}$ s.t. $\sqrt{\mathbf{S}_a}\sqrt{\mathbf{S}_a} = \mathbf{S}_a$ and $\sqrt{\mathbf{S}_a} \succ 0$. Thus, we have

[6] We can also adapt the result of [KNSW20], which is the follow-up work of [AGHS13] and [AR16], to give a different range of parameter sets: We can obtain a smaller lower-bound on r if we set $\varsigma = \Omega(n)$ by adapting [KNSW20].

[7] Similarly, e.g., the sub-Gaussian variable can also be seen as a bounded distribution.

[8] Almost equivalently, we can rely on the tail-cut lemma, e.g., [Pre17, Lem. 2].

$\overline{R_a} = \left(\frac{(\rho_{rs}(\mathbb{Z}^l))^{a-1}}{(\rho_{r\sqrt{\mathbf{\Sigma}}}(\mathbb{Z}^l))^a} \sum_{\mathbf{x}\in\mathbb{Z}^l} \frac{(\rho_{r\sqrt{\mathbf{\Sigma}}}(\mathbf{x}))^a}{(\rho_{rs}(\mathbf{x}))^{a-1}} \right)^{\frac{1}{a-1}}$, and

$$\sum_{\mathbf{x}\in\mathbb{Z}^l} \frac{\left(\rho_{r\sqrt{\mathbf{\Sigma}}}(\mathbf{x})\right)^a}{\left(\rho_{rs}(\mathbf{x})\right)^{a-1}} = \sum_{\mathbf{x}\in\mathbb{Z}^l} \exp\left(-\frac{\pi}{r^2}\mathbf{x}^\intercal (a\mathbf{\Sigma}^{-1} - \frac{a-1}{s^2}\mathbf{I}_n)\mathbf{x} \right)$$
$$= \rho_{r\sqrt{\mathbf{S}_a}^{-1}\sqrt{\mathbf{\Sigma}}}(\mathbb{Z}^l).$$

Then, by Cor. 2.28 and the hypothesis that $r > \eta_\varepsilon(\sqrt{\mathbf{\Sigma}}^{-1}\sqrt{\mathbf{S}_a}\mathbb{Z}^l)$, we have

$$\overline{R_a} \le \left(\frac{(rs(1+\varepsilon))^{n(a-1)}}{(r^l|\sqrt{\mathbf{\Sigma}}|(1-\varepsilon))^a} r^l |\sqrt{\mathbf{S}_a}^{-1}\sqrt{\mathbf{\Sigma}}|(1+\varepsilon) \right)^{\frac{1}{a-1}}$$
$$= \frac{1}{|\frac{1}{s}\sqrt{\mathbf{\Sigma}}||\sqrt{\mathbf{S}_a}|^{\frac{1}{a-1}}} \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^{c_a} = \frac{1 + \frac{2a}{a-1}\hat{\varepsilon}}{\sqrt{|\mathbf{I}_n + \mathbf{R}||\mathbf{I}_n - (a-1)\mathbf{R}|^{\frac{1}{a-1}}}}. \quad \square$$

Then, by sampling $\mathbf{X}$ from a centered and bounded distribution and applying Thm. 3.5 to the above lemma, we obtain the following theorem:

Lemma 5.3. *Let $m \in \mathbb{N}$ be a security parameter, $\gamma \in (\frac{1}{\log m}, \frac{1}{2})$ be a constant, and $l := l(m) < m^\gamma$ be a positive integer. Let χ_β be a centered and β-bounded distribution over $\mathbb{Z}$ for $\beta > 0$ with $\mathrm{V}[\chi_\beta] := \varsigma^2$. Let $\mathbf{X} \sim \chi_\beta^{m\times l}$. Then, for any constant $a \in (2,\infty)$, $\varepsilon := \varepsilon(m) \in (0,1)$, and $r \ge \frac{1}{\varsigma}\sqrt{\frac{2c_{a-1}l}{m}}\eta_\varepsilon^{\le}(\mathbb{Z}^l)$, with overwhelming probability over the choice of $\mathbf{X}$, we have*

$$\overline{R_a} := R_a(\mathcal{D}_{\mathbb{Z}^l, r\mathbf{X}} \parallel \mathcal{D}_{\mathbb{Z}^l, \sqrt{m}\varsigma r}) < (1 + 2c_a\hat{\varepsilon}) \left(1 + 1/(\tfrac{m^\gamma}{l} - 1)\right)^{\frac{1}{a-1}}.$$

Proof. Define $\mathbf{\Sigma} := \mathbf{X}^\intercal\mathbf{X}$ and $\mathbf{R} := \frac{1}{m\varsigma^2}\mathbf{\Sigma} - \mathbf{I}_l$. Let $\delta := \frac{1}{(a-1)m^\gamma}$. Then, similar to the proof of Thm. 4.2, with overwhelming probability over the choice of $\mathbf{X}$, all elements of $\mathbf{R} = [r_{ij}]$ simultaneously satisfy $|r_{ij}| < \delta$ for all $i, j \in [l]$, and $\mathbf{S}_a := \mathbf{I}_l - (a-1)\mathbf{R} \succ 0$, $\mathbf{\Sigma} \succ 0$. Next, by Lem. 2.24 and Fact 2.1, for any $\varepsilon \in \mathbb{R}^+$, we have

$$\eta_\varepsilon(\sqrt{\mathbf{\Sigma}}^{-1}\sqrt{\mathbf{S}_a}\cdot\mathbb{Z}^l) \le \tilde{bl}(\sqrt{\mathbf{\Sigma}}^{-1}\sqrt{\mathbf{S}_a})\cdot\eta_\varepsilon^{\le}(\mathbb{Z}^l) \le \|\sqrt{\mathbf{\Sigma}}^{-1}\sqrt{\mathbf{S}_a}\|_{len}\cdot\eta_\varepsilon^{\le}(\mathbb{Z}^l)$$
$$\le \sqrt{l}\cdot\|\sqrt{\mathbf{\Sigma}}^{-1}\|_{len}\|\sqrt{\mathbf{S}_a}\|_{len}\cdot\eta_\varepsilon^{\le}(\mathbb{Z}^l).$$

By the definition of $\|\cdot\|_{len}$, and given that $\delta < 1$, we have

$$\|\sqrt{\mathbf{S}_a}\|_{len} = \sqrt{\max_{i\in[l]}(1-(a-1)r_{ii})} \le \sqrt{1+\delta} < \sqrt{2}.$$

By Fact 2.2, $\|\sqrt{\mathbf{\Sigma}}^{-1}\|_{len} \le \sigma_l(\sqrt{\mathbf{\Sigma}}^{-1}) \le \sqrt{e_l(\mathbf{\Sigma}^{-1})} \le \sqrt{(e_1(\mathbf{\Sigma}))^{-1}} = \sqrt{1/m\varsigma^2 e_1(\mathbf{I}_l + \mathbf{R})}$. Furthermore, by Lem. 2.8 and Lem. 2.9, we have

$$\|\sqrt{\mathbf{\Sigma}}^{-1}\|_{len} \le \frac{1}{\varsigma\sqrt{m}}\sqrt{\frac{1}{1+e_1(\mathbf{R})}} \le \frac{1}{\varsigma\sqrt{m}}\sqrt{\frac{1}{1-l\delta}} < \frac{1}{\varsigma}\sqrt{\frac{c_{a-1}}{m}}.$$

Thus, we have $\eta_\varepsilon(\sqrt{\mathbf{\Sigma}}^{-1}\sqrt{\mathbf{S}_a}\cdot\mathbb{Z}^l) < \frac{1}{\varsigma}\sqrt{\frac{2c_{a-1}l}{m}}\cdot\eta_\varepsilon^{\leq}(\mathbb{Z}^l) \leq r$. Therefore, we can apply Lem. 5.2 and we have $\overline{R_a} \leq (1+2c_a\hat{\varepsilon})/\sqrt{|\mathbf{I}_n+\mathbf{R}||\mathbf{I}_n-(a-1)\mathbf{R}|^{\frac{1}{a-1}}}$. The rest of the proof is identical to that of Thm. 4.2. □

Finally, we obtain the discrete wSGLHL from Lem. 5.1 and Lem. 5.3:

Theorem 5.4 (Discrete wSGLHL). *Let $m \in \mathbb{N}$ be a security parameter. Let $\gamma \in (\frac{1}{\log m}, \frac{1}{2})$ be a constant, let $l = \omega(\log m)$ be a positive integer s.t. $l \in [100, m^\gamma)$, and let $\varepsilon := \varepsilon(m) \in (0, 10^{-3})$. Let $\varsigma \in \mathbb{R}^+$, $\mathbf{X} \sim (\mathcal{D}_{\mathbb{Z}^m,\varsigma})^l$ and let $(\varsigma')^2 := \mathrm{V}[\mathcal{D}_{\mathbb{Z},\varsigma}]$. If $m \geq \max(30l\log(\varsigma l), l^{1/\gamma})$, $\varsigma \geq 9\eta_\varepsilon^{\leq}(\mathbb{Z}^l)$, and $r \geq \max(\frac{1}{\varsigma'}\sqrt{\frac{2c_a l}{m}}\eta_\varepsilon^{\leq}(\mathbb{Z}^l),\ 10\varsigma l\log m\sqrt{\log(1/\varepsilon)\log(\varsigma l)})$, then, for any constant $a \in (2,\infty)$, with overwhelming probability over the choice of $\mathbf{X}$, we have*

$$\overline{R_a} := R_a(\mathbf{X}^\intercal\mathcal{D}_{\mathbb{Z}^m,r} \parallel \mathcal{D}_{\mathbb{Z}^l,\sqrt{m}\varsigma r}) < (1+4c_a\hat{\varepsilon})\left(1+1/(\tfrac{m^\gamma}{l}-1)\right)^{\frac{1}{a-1}}.$$

Proof. By Lem. 5.1 and Fact 2.16, we have

$$R_\infty(\mathbf{X}^\intercal\mathcal{D}_{\mathbb{Z}^m,r} \parallel \mathcal{D}_{\mathbb{Z}^l,r\mathbf{X}}) \leq \tfrac{1+\varepsilon}{1-\varepsilon} = 1+2\hat{\varepsilon},$$

with probability $1-2^{-l(m)}$ over the choice of $\mathbf{X}$. Next, we show that $\mathbf{X} \sim (\mathcal{D}_{\mathbb{Z}^m,\varsigma})^l$ can be viewed as a ς-bounded distribution, with overwhelming probability. By Lem. 2.26, we have $\Pr_{\mathbf{x}\sim\mathcal{D}_{\mathbb{Z}^m,\varsigma}}[\|\mathbf{x}\|_\infty > \varsigma] \leq \frac{1+\varepsilon}{1-\varepsilon}\cdot 2^{-m} = \mathrm{negl}(m)$. Hence, by the union bound, we can show that all elements of $\mathbf{X} = [x_{ij}]$ simultaneously satisfy $|x_{ij}| \leq \varsigma$ with overwhelming probability. And, we have $\frac{\varsigma'}{\varsigma} = O(1)$ by Lem. 2.25. Therefore, by Lem. 5.3, for any constant $a \in (2,\infty)$, we have

$$R_a(\mathcal{D}_{\mathbb{Z}^l,r\mathbf{X}} \parallel \mathcal{D}_{\mathbb{Z}^l,\sqrt{m}\varsigma r}) < (1+2c_a\hat{\varepsilon})\left(1+1/(\tfrac{m^\gamma}{l}-1)\right)^{\frac{1}{a-1}}$$

with overwhelming probability over the choice of $\mathbf{X}$. Therefore, by the weak triangle inequality of the Rényi divergence (Lem. 2.12), we obtain

$$\overline{R_a} < (1+2\hat{\varepsilon})^{c_a}(1+2c_a\hat{\varepsilon})\left(1+1/(\tfrac{m^\gamma}{l}-1)\right)^{\frac{1}{a-1}},$$

and the theorem follows. □

Note that we usually set $\varepsilon = \mathrm{negl}(m)$ for cryptographic applications, and thus we have $1+4c_a\hat{\varepsilon} = 1+\mathrm{negl}(m)$ for any constant $a > 2$.

6 A Sharper LWE Self-Reduction

The purpose of this section is to show our LWE self-reduction in Thm. 6.2. For simplicity, we consider only LWE with the discrete Gaussian. Similar results for LWE with a continuous Gaussian can be obtained by relying on the continuous wSGLHL, Thm. 4.2. We first recall the (classical) leftover hash lemma:

Lemma 6.1 ([Reg09, Claim 5.3]). *Let* $\mathbf{A} \sim \mathcal{U}(\mathbb{Z}_q^{m\times n})$, $\mathbf{X} \sim (\mathcal{D}_{\mathbb{Z}^m,\varsigma})^l$ *for some* $\varsigma = \omega(\sqrt{\log m})$. *Then,* $\Delta(\mathbf{X}^\intercal\mathbf{A}, \mathcal{U}(\mathbb{Z}_q^{l\times n})) = \mathrm{negl}(m)$ *and* $\Delta_{\mathrm{ML}}(\mathbf{X}^\intercal\mathbf{A}, \mathcal{U}(\mathbb{Z}_q^{l\times n})) = \ln(1+\mathrm{negl}(m))$ *hold.*

By combining Lem. 6.1 and Thm. 5.4, we obtain our LWE sample rerandomization theorem, which states that rerandomized LWE samples and (plain) LWE samples are close with respect to RD:

Theorem 6.2. *Let* $n \in \mathbb{N}$ *be a security parameter, and let* $m = \Omega(n)$ *and* $q := q(n)$ *be integers. Let* $\gamma \in (\frac{1}{\log m}, \frac{1}{2})$ *be a constant,* $l = \omega(\log n)$ *be a positive integer s.t.* $l \in [100, m^\gamma)$, *and* $\varepsilon := \varepsilon(n) \in (0, 10^{-3})$. *Let* $\varsigma \in \mathbb{R}^+$, $\mathbf{X} \sim (\mathcal{D}_{\mathbb{Z}^m,\varsigma})^l$, *and* $\varsigma' := \mathrm{V}[\mathcal{D}_{\mathbb{Z},\varsigma}]$. *If* $\varsigma \geq 9\eta_{\hat{\varepsilon}}^{\leq}(\mathbb{Z}^l)$, $m \geq \max(30l\log(\varsigma l), l^{1/\gamma})$, $r \geq \max(\frac{1}{\varsigma'}\sqrt{\frac{2c_a l}{m}}\eta_{\hat{\varepsilon}}^{\leq}(\mathbb{Z}^l), 10\varsigma l\log m\sqrt{\log(1/\varepsilon)\log(\varsigma l)})$, *and* $q > 2\sqrt{m}\varsigma r$, *then, for any constant* $a \in (2,\infty)$, *with overwhelming probability over the choice of* $\mathbf{X}$, *we have*

$$\begin{aligned}\overline{R_a} &:= R_a(\mathbf{X}^\intercal\mathsf{LWE}_{\mathbf{s}}(m,n,q,\mathcal{D}_{\mathbb{Z},r}) \parallel \mathsf{LWE}_{\mathbf{s}}(l,n,q,\mathcal{D}_{\mathbb{Z},\sqrt{m}\varsigma r}))\\ &< (1+\mathrm{negl}(n))(1+4c_a\hat{\varepsilon})\left(1+1/(\tfrac{m^\gamma}{l}-1)\right)^{\frac{1}{a-1}}.\end{aligned}$$

Proof. Let $(\mathbf{A}, \mathbf{b} := \mathbf{A}\mathbf{s}+\mathbf{e}) \sim \mathsf{LWE}_{\mathbf{s}}(m,n,q,\mathcal{D}_{\mathbb{Z},r})$ and $(\mathbf{A}', \mathbf{b}' := \mathbf{A}'\mathbf{s}+\mathbf{e}') \sim \mathsf{LWE}_{\mathbf{s}}(l,n,q,\mathcal{D}_{\mathbb{Z},\sqrt{m}\varsigma r})$. Let $\mathbf{U} \sim \mathcal{U}(\mathbb{Z}_q^{l\times n})$ and $\mathbf{v} \sim \mathcal{U}(\mathbb{Z}_q^l)$ be uniformly random variables. By the weak triangle inequality and multiplicativity (Lem. 2.12), Lem. 6.1 and Lem. 2.29, we have

$$\begin{aligned}\overline{R_a} &= R_a((\mathbf{X}^\intercal\mathbf{A}, \mathbf{X}^\intercal\mathbf{A}\mathbf{s}+\mathbf{X}^\intercal\mathbf{e}) \parallel (\mathbf{A}', \mathbf{A}'\mathbf{s}+\mathbf{e}'))\\ &< R_\infty((\mathbf{X}^\intercal\mathbf{A}, \mathbf{X}^\intercal\mathbf{A}\mathbf{s}+\mathbf{X}^\intercal\mathbf{e}) \parallel (\mathbf{U}, \mathbf{v}+\mathbf{X}^\intercal\mathbf{e}))^{c_a}\\ &\quad \cdot R_a(\mathbf{U}, \mathbf{v}+\mathbf{X}^\intercal\mathbf{e})) \parallel (\mathbf{A}', \mathbf{A}'\mathbf{s}+\mathbf{e}'))\\ &= (1+\mathrm{negl}(n)) \cdot R_a(\mathbf{U}, \mathbf{v}+\mathbf{X}^\intercal\mathbf{e})) \parallel (\mathbf{A}', \mathbf{A}'\mathbf{s}+\mathbf{e}'))\\ &= (1+\mathrm{negl}(n)) \cdot R_a(\mathbf{U} \parallel \mathbf{A}') \cdot R_a(\mathbf{v}+\mathbf{X}^\intercal\mathbf{e} \parallel \mathbf{A}'\mathbf{s}+\mathbf{e}')\\ &= (1+\mathrm{negl}(n)) \cdot R_a(\mathbf{X}^\intercal\mathbf{e} \parallel \mathbf{e}')\\ &= (1+\mathrm{negl}(n)) \cdot R_a(\mathbf{X}^\intercal\mathcal{D}_{\mathbb{Z}^m,r} \parallel \mathcal{D}_{\mathbb{Z}^l,\sqrt{m}\varsigma r}).\end{aligned}$$

Hence, the theorem follows from Thm. 5.4. □

Note that we use sufficiently large q to ensure that the rerandomized Gaussian errors are smaller than $q/2$ with overwhelming probability; i.e., the discrete Gaussian on $\mathbb{Z}_q$ is statistically close to that on $\mathbb{Z}$ by Lem. 2.26. As a corollary of Thm. 6.2, we obtain the following LWE self-reduction from Lem. 2.14:

Corollary 6.3 (LWE self-reduction). *In Thm. 6.2, assume that for any PPT algorithm, the success probability of solving* $\mathsf{Search\text{-}LWE}_{\mathbf{s}}(l,n,q,\mathcal{D}_{\mathbb{Z},\sqrt{m}\varsigma r})$ *is at most* 2^{-Cn} *for some constant* $C \in \mathbb{R}^+$. *Then, the success probability of any PPT algorithm for finding* $\mathbf{s}$ *from the distribution* $(\mathbf{X}^\intercal\mathsf{LWE}_{\mathbf{s}}(m,n,q,\mathcal{D}_{\mathbb{Z},r}))$ *is at most*

$$p := 2^{-\frac{1}{c_a}\left(Cn-\log\left((1+\mathrm{negl}(n))(1+4c_a\hat{\varepsilon})\left(1+1/(\frac{m^\gamma}{l}-1)\right)^{\frac{1}{a-1}}\right)\right)}, \tag{3}$$

with overwhelming probability over the choice of $\mathbf{X}$.

Unlike security analysis based on, e.g., statistical distance, we do not (need to) show that RD is negligibly small (precisely, $1 + \mathrm{negl}(n)$). As shown in recent works, e.g., [Pre17, BLR+18, BJRW22, ASY22], some non-negligible, but a small RD is sufficient for constructing meaningful security arguments. We demonstrate that Cor. 6.3 is useful by instantiating some concrete parameters. For the selected parameters, we show that finding $\mathbf{s}$ from the rerandomized LWE samples $(\mathbf{X}^\intercal \mathsf{LWE}_{\mathbf{s}}(m, n, q, \mathcal{D}_{\mathbb{Z},r}))$ is almost as hard as $\mathsf{Search\text{-}LWE}_{\mathbf{s}}(l, n, q, \mathcal{D}_{\mathbb{Z},\sqrt{m}\varsigma r})$ with the loss of only a few security bits.

Corollary 6.4. *Let $\gamma = 0.45$, $m = \mathrm{poly}(n) > 100^3$, $l = \sqrt[3]{m}$, and $\varepsilon = \mathrm{negl}(n)$ in Cor. 6.3. Then, we have $p = 2^{-0.99Cn+0.01+\mathrm{negl}(n)}$, where p is defined in (3).*

7 Application to the Independence Heuristic

In this section, we weaken the heuristic upon which the TFHE scheme relies, by applying our continuous wSGLHL (Thm. 4.2 and Thm. 4.4).

The TFHE scheme [CGGI16, CGGI17, CGGI20] is an FHE scheme based on the Ring-LWE (or Module-LWE) problem. The scheme relies on the heuristic that linear combinations of the errors of ciphertexts are mutually independent to analyze their variance. Since our continuous wSGLHL (Thm. 4.2 and Thm. 4.4) shows that linear sums of Gaussian errors and mutually independent Gaussian errors are close with respect to RD, it essentially mitigates the heuristic.

We first provide a brief overview of the TFHE construction in Sect. 7.1. Then, in Sect. 7.2, we explain how our theorem can be adapted to the concrete setting of the TFHE scheme to weaken the independence heuristic.

7.1 Brief Overview of the TFHE Construction

We define $\mathbb{B} := \{0, 1\}$ and $\mathbb{T} := \mathbb{R}/\mathbb{Z}$. Let $N \in \mathbb{N}$. We denote by $\mathbb{Z}_N[X]$ the ring of polynomials $\mathbb{Z}[X]/(X^N + 1)$, and define $\mathbb{T}_N[X] := \mathbb{R}[X]/(X^N + 1) \bmod 1$. $\mathbb{B}_N[X]$ denotes the polynomials in $\mathbb{Z}_N[X]$ with binary coefficients. The TFHE scheme is based on generalized variants of LWE ciphertexts; TLWE ciphertexts:

Definition 7.1. *Let $k \in \mathbb{N}$, N be a power of 2 and $\alpha \in \mathbb{R}^+$ be a standard deviation. Let the secret key $\mathbf{s} \sim \mathcal{U}(\mathbb{B}_N[X]^k)$. The (canonical) TLWE ciphertext of the message $\mu \in \mathbb{T}_N[X]$ is $(\mathbf{a}, b := \mathbf{s}^\intercal \mathbf{a} + \mu + e) \in \mathbb{T}_N[X]^{k+1}$, where $\mathbf{a} \sim \mathcal{U}(\mathbb{T}_N[X]^k)$ and $e \leftarrow \mathcal{D}_{\mathbb{T}_N[X],\alpha}$. The phase and error of the ciphertext is denoted by $\phi_{\mathbf{s}}((\mathbf{a}, b)) := b - \mathbf{s}^\intercal \mathbf{a}$ and $\mathsf{Err}((\mathbf{a}, b))$, respectively.*

The fully homomorphic property of the TFHE scheme is based on the TGSW *encryption*, the ciphertext of which is essentially a matrix composed of rows of TLWE ciphertexts. Before we define the TGSW ciphertexts, we define the *(canonical) gadget decomposition* of the TLWE ciphertexts as follows:

Definition 7.2. *Let $l, B_g \in \mathbb{N}$, and let $\mathbf{b} \in \mathbb{T}_N[X]^{k+1}$ be the TLWE sample. Define the (canonical) gadget as a matrix $\mathbf{H} \in \mathbb{T}_N[X]^{(k+1)l \times (k+1)}$ whose diagonal blocks are $\mathbf{g}^\intercal := (1/B_g, \ldots, 1/B_g^l)^\intercal$ and whose other elements are all zero. The valid decomposition algorithm $Dec_{\mathbf{H},\beta,\varepsilon}(\mathbf{b})$ on the gadget $\mathbf{H}$ with quality $\beta =$*

$B_g/2$ and precision $\varepsilon = 1/B_g^l$ outputs a vector $\mathbf{u} \in \mathbb{R}_N[X]^{(k+1)l}$ s.t. $\|\mathbf{u}\|_\infty \leq \beta$, $\|\mathbf{u}^\intercal \mathbf{H} - \mathbf{b}\|_\infty \leq \varepsilon$, and $\mathrm{E}[\mathbf{u}^\intercal \mathbf{H} - \mathbf{b}] = \mathbf{0}$ when $\mathbf{b}$ is uniformly random.

We are now ready to define the TGSW ciphertext, and the external product between a TGSW ciphertext and a TLWE ciphertext.

Definition 7.3. *A (canonical)* TGSW *ciphertext of message $\mu \in \mathbb{Z}_N[X]$ is $\mathbf{C} = \mathbf{Z} + \mu \cdot \mathbf{H}$, where each row of $\mathbf{Z} \in \mathbb{T}_N[X]^{(k+1)l \times (k+1)}$ is a (canonical)* TLWE *ciphertext of 0 over $\mathbb{T}_N[X]^{(k+1)}$. Let $\mathsf{Err}(\mathbf{C})$ denotes the list of the $(k+1)l$* TLWE *errors of each line of $\mathbf{C}$.*

Definition 7.4 (External product). *We define the product $\boxdot$ as, $\boxdot : \mathsf{TGSW} \times \mathsf{TLWE} \longrightarrow \mathsf{TLWE} : (\mathbf{A}, \mathbf{b}) \longmapsto \mathbf{A} \boxdot \mathbf{b} = (Dec_{\mathbf{H},\beta,\varepsilon}(\mathbf{b}))^\intercal \mathbf{A}$, where $Dec_{\mathbf{H},\beta,\varepsilon}$ is the gadget decomposition defined in Def. 7.2.*

7.2 Mitigating the Independence Heuristic for TFHE

We recall the independence heuristic presented in [CGGI20] (which is common in [CGGI16, CGGI17]):

Assumption 7.5 (Independence heuristic, [CGGI20, Assumption 3.11]). *All the coefficients of the errors of* TLWE *or* TGSW *samples that occur in all the linear combinations we consider are independent and concentrated. More precisely, they are σ-subGaussian where σ is the square-root of their variance.*

The core analysis that requires this assumption is the following theorem, which yields the fully homomorphic property of the TFHE scheme:

Theorem 7.6 ([CGGI20, Thm. 3.13 and Cor. 3.14]). *Let $\mathbf{A}$ be a* TGSW *ciphertext of message $\mu_{\mathbf{A}}$ (Def. 7.3) and let $\mathbf{b}$ be a* TLWE *ciphertext of message $\mu_{\mathbf{b}}$ (Def. 7.1). Then, we have that $\mathbf{A} \boxdot \mathbf{b}$ (Def. 7.4) is a* TLWE *sample of message $\mu_{\mathbf{A}} \cdot \mu_{\mathbf{b}}$, and*

$$\|\mathsf{Err}(A \boxdot \mathbf{b})\|_\infty \leq (k+1)lN\beta \|\mathsf{Err}(\mathbf{A})\|_\infty + (1+kN)\|\mu_{\mathbf{A}}\|_1 \varepsilon + \|\mu_A\|_1 \|\mathsf{Err}(\mathbf{b})\|_\infty,$$

where β and ε are the parameters used in the decomposition $Dec_{\mathbf{H},\beta,\varepsilon}(\mathbf{b})$ (Def. 7.2). Furthermore, under Assumption 7.5, we have

$$\mathrm{V}(\mathsf{Err}(\mathbf{A} \boxdot \mathbf{b})) \leq (k+1)lN\beta^2 \,\mathrm{V}(\mathsf{Err}(\mathbf{A})) + (1+kN)\|\mu_{\mathbf{A}}\|^2 \varepsilon^2 + \|\mu_{\mathbf{A}}\|_2^2 \,\mathrm{V}(\mathsf{Err}(\mathbf{b})).$$

We derive the above bound of $\mathrm{V}(\mathsf{Err}(\mathbf{A} \boxdot \mathbf{b}))$ with a weaker heuristic than Assumption 7.5. First, we formulate $\mathsf{Err}(\mathbf{A} \boxdot \mathbf{b})$. Let $\mathbf{u} := Dec_{h,\beta,\varepsilon}(\mathbf{b})$ and define $\boldsymbol{\varepsilon}_{dec} := \mathbf{b} - \mathbf{u}^\intercal \mathbf{H}$. It is shown in the proof of [CGGI20, Thm. 3.13] that

$$\mathsf{Err}(\mathbf{A} \boxdot \mathbf{b}) = \mathbf{u}^\intercal \mathsf{Err}(\mathbf{A}) + \mu_{\mathbf{A}} \cdot \phi_{\mathbf{s}}(\boldsymbol{\varepsilon}_{dec}) + \mu_{\mathbf{A}} \cdot \mathsf{Err}(\mathbf{b}) \tag{4}$$

holds. Let us denote $\mathbf{u}^\intercal := (u_1, \ldots, u_{(k+1)l})^\intercal \in \mathbb{R}_N[X]^{(k+1)l}$ and $\mathsf{Err}(\mathbf{A}) := \mathbf{e}^\intercal := (e_1, \ldots, e_{(k+1)l})^\intercal \in \mathbb{T}_N[X]^{(k+1)l}$. Then, we have $\mathbf{u}^\intercal \mathbf{e} = \sum_{i=1}^{(k+1)l} u_i e_i \in \mathbb{T}_N[X]$. For some $i \in [(k+1)l]$, we define $u_i := \sum_{j=0}^{N-1} \upsilon_j X^j$ and $e_i := \sum_{j=0}^{N-1} \eta_j X^j$, and we let $\boldsymbol{\upsilon}^\intercal := (\upsilon_0, \ldots, \upsilon_{N-1})^\intercal \in \mathbb{R}^N$ and $\boldsymbol{\eta}^\intercal := (\eta_0, \ldots, \eta_{N-1})^\intercal \in \mathbb{T}^N$. We also

define $e_i := \sum_{j=0}^{N-1} \overline{\eta_j} X^{-j}$, where $\overline{\eta_j} := -\eta_{N-j}$ and $\eta_N := -\eta_0$, and we define $\overline{\boldsymbol{\eta}}^\intercal := (\overline{\eta_0}, \ldots, \overline{\eta_{N-1}})^\intercal \in \mathbb{T}^N$. For $k \in \{0, \ldots, N-1\}$, we define the *k-th rotation* of $\boldsymbol{v}$ as $(\boldsymbol{v}^{(k)})^\intercal := (v_0^{(k)}, \ldots, v_{N-1}^{(k)})^\intercal \in \mathbb{R}^N$, where $v_j^{(k)} = v_{j+k}$ if $j+k \leq N-1$ and $v_j^{(k)} = -v_{(j+k \bmod N)}$ otherwise. Then, we can write $u_i e_i = \sum_{k=0}^{N-1} (\boldsymbol{v}^{(k)})^\intercal \overline{\boldsymbol{\eta}} X^k$. Let $\boldsymbol{\Upsilon} \in \mathbb{R}^{N \times N}$ be the matrix of the columns of $\boldsymbol{v}^{(0)}, \ldots, \boldsymbol{v}^{(N-1)}$, then we can write the coefficient vector of $u_i e_i$ as $\mathbf{v} = \boldsymbol{\Upsilon}^\intercal \overline{\boldsymbol{\eta}}$.

Next, we explain how our results mitigates the required assumption in Thm. 7.6. We consider the dependence among the coefficients of the term $\mathbf{u}^\intercal \mathbf{e}$ ($= \mathbf{u}^\intercal \mathsf{Err}(\mathbf{A})$ in (4)). Now, we assume $\mathbf{b}$ is uniformly random since we have $\mathbf{b} \overset{\text{comp}}{\approx} \mathcal{U}(\mathbb{T}_N[X]^{k+1})$ under the hardness assumption of the (decision) TLWE problem. Then, we assume $\mathbf{u} := Dec_{h,\beta,\varepsilon}(\mathbf{b})$ is also uniformly random over the set $\{\mathbf{u} \in \mathbb{R}_N[X]^{(k+1)l} \mid \|\mathbf{u}\|_\infty \leq \beta\}$ (see Def. 7.2). Under this assumption, for any $i_1 \neq i_2$, $u_{i_1} e_{i_1}$ and $u_{i_2} e_{i_2}$ are mutually independent. Hence, we only need to analyze the dependence between the coefficients of $u_i e_i$ for each i, i.e., the dependence between the elements of $\mathbf{v} = \boldsymbol{\Upsilon}^\intercal \overline{\boldsymbol{\eta}}$. By Defs. 7.1 and 7.3, $e_i \sim \mathcal{D}_{\mathbb{T}_N[X],\alpha}$, where $\alpha := \sqrt{\mathrm{V}(\mathsf{Err}(\mathbf{A}))}$, and thus we can consider that $\overline{\boldsymbol{\eta}} \sim \mathcal{N}_N(\alpha^2)$, by assuming that the standard deviation α is sufficiently small. Since $\|\mathbf{u}\|_\infty \leq \beta$ by definition, we have $\|\boldsymbol{v}\|_\infty \leq \beta$, and thus $\boldsymbol{v}^{(0)}, \ldots, \boldsymbol{v}^{(N-1)} \sim \chi_\beta^N$. If they are independent, i.e., $\boldsymbol{v}^{(0)}, \ldots, \boldsymbol{v}^{(N-1)} \overset{\text{iid}}{\sim} \chi_\beta^N$, our Thm. 4.2 (or, Thm. 4.4) essentially shows that $\mathbf{v} = \boldsymbol{\Upsilon}^\intercal \overline{\boldsymbol{\eta}} \sim \mathcal{N}_N(N\beta^2\alpha^2)$; this means the elements of $\mathbf{v}$ are mutually independent, and moreover, $\mathrm{V}(\mathbf{u}^\intercal \mathsf{Err}(\mathbf{A})) \simeq (k+1)lN\beta^2 \mathrm{V}(\mathsf{Err}(\mathbf{A}))$. Thm. 4.2 (Thm. 4.4) only needs independence among $\boldsymbol{v}^{(0)}, \ldots, \boldsymbol{v}^{(N-1)}$ to use Thm. 3.5. Interestingly, we can show that Thm. 3.5 also holds when we set $\mathbf{X} := \boldsymbol{\Upsilon}$.

Lemma 7.7. *Let χ_β be a centered and β-bounded distribution for $\beta > 0$ with $\mathrm{V}[\chi_\beta] := \varsigma^2$. Let $\boldsymbol{v} \sim \chi_\beta^m$, and let $\boldsymbol{\Upsilon} \in \mathbb{R}^{m \times m}$ be the matrix of rotations of $\boldsymbol{v}$; i.e., $\boldsymbol{v}^{(0)}, \ldots, \boldsymbol{v}^{(m-1)}$. Then, Thm. 3.5 also holds for $\mathbf{X} := \boldsymbol{\Upsilon}$.*

Proof (Lem. 7.7). Let $\mathbf{S} := \left(\frac{1}{m\varsigma^2}\boldsymbol{\Sigma}\right) = [s_{ij}]$. For $i \in [2m]$ we define $\hat{v}_i = v_i$ and $\hat{v}_{-i} = v_{m-i}$ if $i \in [m]$ and $\hat{v}_i = -v_{i-m}$ if $i \in [m+1, 2m]$. For simplicity, we assume $m \equiv 0 \bmod 3$. For $i > j$, we define $d := d_{ij} = (i-j) > 0$, then we have

$$\begin{aligned} s_{ij} &= \tfrac{1}{m\varsigma^2}(\boldsymbol{v}^{(i-1)})^\intercal \boldsymbol{v}^{(j-1)} = \tfrac{1}{m\varsigma^2}\textstyle\sum_{k=1}^m v_k^{(i-1)} v_k^{(j-1)} = \tfrac{1}{m\varsigma^2}\sum_{k=1}^m \hat{v}_{k+i}\hat{v}_{k+j} \\ &= \tfrac{1}{m\varsigma^2}\textstyle\sum_{k=1}^{m/3} \hat{v}_{i+3(k-1)}(\hat{v}_{i+3(k-1)-d} + \hat{v}_{i+3(k-1)+d}). \end{aligned}$$

Since $\hat{v}_{i+3(k-1)-d} + \hat{v}_{i+3(k-1)+d}$ is 2β-bounded, by Fact 3.2, all terms $\hat{v}_{i+3(k-1)}(\hat{v}_{i+3(k-1)-d} + \hat{v}_{i+3(k-1)+d})$ for $i \in [m/3]$ are $2\beta^2$-bounded and independent. Thus, by Lem. 3.3, we have $\Pr[|s_{ij}| \geq \delta] < 2\exp\left(-\frac{3\varsigma^4}{8\beta^4}\delta^2 m\right) = \mathrm{negl}(m)$ for all $i \neq j$. Hence, by the union bound, $\Pr[\bigcup_{i \neq j}(|s_{ij}| \geq \delta)] \leq \sum_{i \neq j} \Pr[|s_{ij}| \geq \delta] = \mathrm{negl}(m)$ holds, and we have

$$\Pr[\textstyle\bigcap_{i \neq j}(|s_{ij}| < \delta)] \geq 1 - \mathrm{negl}(m).$$

Every diagonal elements are $s_{ii} = \frac{1}{m\varsigma^2}\|\boldsymbol{v}^{(i-1)}\|^2 = \frac{1}{m\varsigma^2}\|\boldsymbol{v}\|^2 = \frac{1}{m\varsigma^2}\sum_{k=1}^m v_k^2$ by definition. Since $v_1, \ldots, v_m \overset{\text{iid}}{\sim} \chi_\beta$, by Lem. 3.4, we have $\Pr[|s_{ii} - 1| \geq \delta] <$

$2\exp\left(-2\frac{\varsigma^4}{\beta^4}\delta^2 m\right) = \mathrm{negl}(m)$, and,

$$\Pr[\bigcap_{i\in[l]}(|s_{ii}-1|<\delta)] = \Pr[|s_{ii}-1|<\delta] \geq 1-\mathrm{negl}(m).$$

Since $\mathbf{R} = \mathbf{S} - \mathbf{I}_l$, we obtain the theorem. □

Thus, Thm. 4.2 (Thm. 4.4) holds even when we set $\mathbf{X} := \boldsymbol{\Upsilon}$, and we can show that all the coefficients of $\mathbf{u}^\intercal \mathsf{Err}(\mathbf{A})$ in (4) are mutually independent. Therefore, we only need to *partially* rely on Assumption 7.5 to heuristically assume that $\phi_{\mathbf{s}}(\boldsymbol{\varepsilon}_{dec})$ and $\mathbf{b}$ in (4) are mutually independent.

8 Conclusion and Open Problems

The main contribution of this paper is the (continuous and discrete) wSGLHL presented in Sects. 4 and 5. Indeed, the discrete wSGLHL (Thm. 5.4) solves a open question posed by Agrawal *et al.* [AGHS13] in a weak sense.

Based on our wSGLHL, we presented a sharp LWE self-reduction (Cor. 6.3), which states that finding $\mathbf{s}$ from rerandomized LWE samples is at least as hard as Search-LWE with errors of *known* variance (with the loss of a few bits of security). Existing works [ACPS09, HKM18, GMPW20] only show that rerandomized LWE samples are statistically close to (plain) LWE samples with some *unknown* variance. Thus, our reduction is sharper than the existing work in terms of the size of errors (see also Fig. 1). As another application of our continuous wSGLHL, we weakened the independence heuristic required for the TFHE scheme in Sect. 7. We discuss open problems and future works in the following.

Why Rényi divergence? We constructed the wSGLHL based on RD rather than (standard) metrics such as the statistical distance or the max-log distance, because it seems difficult to perform our analysis with these metrics. The max-log distance is equivalent to R_∞ since $\Delta_{\mathrm{ML}}(\mathcal{D}_1, \mathcal{D}_2) = \max\{\ln(R_\infty(\mathcal{D}_1 \,\|\, \mathcal{D}_2)), \ln(R_\infty(\mathcal{D}_2 \| \mathcal{D}_1))\}$ by definition. In addition, if we obtain a bound on R_∞, we can obtain the bound on the statistical distance by Fact 2.17. However, it seems difficult to derive a quantitative bound on R_∞ when $\mathbf{X}$ is taken from the general centered and bounded distribution. We showed in Lem. 4.1 that we can analyze R_∞ if the residual matrix $\mathbf{R}$ satisfies $-\mathbf{R} \succeq 0$. In our framework, Thm. 3.5 is used to show that the absolute values of all elements of $\mathbf{R}$ are bounded by $\delta = \frac{1}{m^\gamma}$ for some $\gamma < \frac{1}{2}$. Hence, we can show that at least $\lim_{m\to\infty}(-\mathbf{R}) = \mathbf{O} \succeq 0$ holds, but it is difficult to show that $-\mathbf{R} \succeq 0$ holds for some finite m, when we sample $\mathbf{X}$ from the general centered and bounded distribution. We should require an additional condition on $\mathbf{X}$: as a trivial example, (exactly) orthogonal matrices $\mathbf{X}$ satisfy $\mathbf{R} = \mathbf{X}^\intercal\mathbf{X} - \mathbf{I} = \mathbf{O} \succeq 0$.

Nonetheless, security arguments based on the RD are sometimes sufficient (or better) for cryptographic applications, as mentioned earlier. We demonstrated that we can construct the LWE self-reduction (Cor. 6.4) with the loss of only a few small bits of security.

Additionally, note that RD is not symmetric (and thus it is not a metric); $R_a(\mathcal{D}_1 \,\|\, \mathcal{D}_2) = R_a(\mathcal{D}_2 \,\|\, \mathcal{D}_1)$ does not necessarily hold. Although we only analyze RD needed for the LWE self-reduction (Cor. 6.3), a similar analysis can be performed for the "opposite" of RD analyzed in this paper.

Further improvements to the discrete wSGLHL. We showed in Sect. 4.2 that our continuous wSGLHL (Thm. 4.2) can be improved with the noise flooding technique (Thm. 4.4): We can increase the dimension ($=l$) of the output spherical Gaussian at the expense of increasing the variance of the errors. We believe that a similar analysis is also applicable to the discrete wSGLHL (Thm. 5.4). The theorem relies on the discrete (ellipsoid) Gaussian LHL [AR16, Thm. 5.1] (Lem. 5.1), which requires the input dimension m to be larger than the output dimension l. However, we require the discrete analog of Lem. 4.3, which is applicable to any $m, l \in \mathbb{N}$: We need to modify Lem. 5.1 to support any $m, l \in \mathbb{N}$, which we leave for future work.

Other possible applications. Our discrete wSGLHL is an extension of the discrete (ellipsoid) Gaussian LHL proposed by Agrawal *et al.* [AGHS13]. They discussed their discrete Gaussian LHL is sufficient for GGH encoding [GGH13]. On the other hand, it has also been mentioned that in some applications where the trapdoor is explicitly available, and *oblivious sampling* is not needed, it is safer to use a perfectly spherical Gaussian that is statistically independent of the trapdoor. Our discrete wSGLHL could possibly provide a better (or simpler) security proof for construction. The verification of this observation remains a topic for future consideration.

References

[AR16] Aggarwal, D., and Regev, O.: A note on discrete Gaussian combinations of lattice vectors. Chicago Journal of Theoretical Computer Science 2016(7) (2016). Preliminary version: https://arxiv.org/abs/1308.2405

[AFV11] Agrawal, S., Freeman, D.M., and Vaikuntanathan, V.: Functional encryption for inner product predicates from learning with errors. In: ASIACRYPT 2011, pp. 21–40 (2011)

[AGHS13] Agrawal, S., Gentry, C., Halevi, S., and Sahai, A.: Discrete Gaussian leftover hash lemma over infinite domains. In: ASIACRYPT 2013, pp. 97–116 (2013)

[ASY22] Agrawal, S., Stehlé, D., and Yadav, A.: Round-Optimal Lattice-Based Threshold Signatures, Revisited. In: ICALP 2022, 8:1–8:20 (2022)

[AAC+22] Alagic, G., Apon, D., Cooper, D., Dang, Q., Dang, T., Kelsey, J., Lichtinger, J., Miller, C., Moody, D., Peralta, R.: NISTIR 8413: Status report on the third round of the NIST post-quantum cryptography standardization process. NIST (2022)

[ACPS09] Applebaum, B., Cash, D., Peikert, C., and Sahai, A.: Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In: CRYPTO 2009, pp. 595–618 (2009)

[BLR+18] Bai, S., Lepoint, T., Roux-Langlois, A., Sakzad, A., Stehlé, D., and Steinfeld, R.: Improved security proofs in lattice-based cryptography: Using the Rényi divergence rather than the statistical distance. J. Cryptol. 31(2), 610–640 (2018). Preliminary version in ASIACRYPT 2015.

[BHK+19] Bernstein, D.J., Hülsing, A., Kölbl, S., Niederhagen, R., Rijneveld, J., and Schwabe, P.: The SPHINCS+ signature framework. In: CCS '19, pp. 2129–2146 (2019)

[BGM+16] Bogdanov, A., Guo, S., Masny, D., Richelson, S., and Rosen, A.: On the hardness of learning with rounding over small modulus. In: TCC 2016, pp. 209–224 (2016)

[BDK+18] Bos, J., Ducas, L., Kiltz, E., Lepoint, T., Lyubashevsky, V., Schanck, J.M., Schwabe, P., Seiler, G., and Stehlé, D.: CRYSTALS-Kyber: A CCA-secure module-lattice-based KEM. In: Euro S&P 2018, pp. 353–367 (2018)

[BJRW22] Boudgoust, K., Jeudy, C., Roux-Langlois, A., and Wen, W.: On the hardness of module learning with errors with short distributions. J. Cryptol. 36(1), 1 (2022)

[BGV12] Brakerski, Z., Gentry, C., and Vaikuntanathan, V.: (Leveled) fully homomorphic encryption without bootstrapping. In: ITCS 2012, pp. 309–325 (2012)

[BV11] Brakerski, Z., and Vaikuntanathan, V.: Fully homomorphic encryption from Ring-LWE and security for key dependent messages. In: CRYPTO 2011, pp. 505–524 (2011)

[BOS15] Brent, R.P., Osborn, J.-A.H., and Smith, W.D.: Note on best possible bounds for determinants of matrices close to the identity matrix. Linear Algebra and its Applications 466, 21–26 (2015)

[CGHX19] Case, B.M., Gao, S., Hu, G., and Xu, Q.: Fully homomorphic encryption with k-bit arithmetic operations, Cryptology ePrint Archive, Paper 2019/521 (2019), https://eprint.iacr.org/2019/521

[CGM19] Chen, Y., Genise, N., and Mukherjee, P.: Approximate trapdoors for lattices and smaller hash-and-sign signatures. In: ASIACRYPT 2019, pp. 3–32 (2019)

[CKKS17] Cheon, J.H., Kim, A., Kim, M., and Song, Y.: Homomorphic encryption for arithmetic of approximate numbers. In: ASIACRYPT 2017, pp. 409–437 (2017)

[CGGI16] Chillotti, I., Gama, N., Georgieva, M., and Izabachène, M.: Faster fully homomorphic encryption: bootstrapping in less than 0.1 seconds. In: ASIACRYPT 2016, pp. 3–33 (2016)

[CGGI17] Chillotti, I., Gama, N., Georgieva, M., and Izabachène, M.: Faster packed homomorphic operations and efficient circuit bootstrapping for TFHE. In: ASIACRYPT 2017, pp. 377–408 (2017)

[CGGI20] Chillotti, I., Gama, N., Georgieva, M., and Izabachène, M.: TFHE: Fast fully homomorphic encryption over the torus. J. Cryptol. 33(1), 34–91 (2020)

[DRS04] Dodis, Y., Reyzin, L., and Smith, A.: Fuzzy extractors: how to generate strong keys from biometrics and other noisy data. In: EUROCRYPT 2004, pp. 523–540 (2004)

[DGPY20] Ducas, L., Galbraith, S., Prest, T., and Yu, Y.: Integral matrix Gram root and lattice Gaussian sampling without floats. In: EUROCRYPT 2020, pp. 608–637 (2020)

[DKL+18] Ducas, L., Kiltz, E., Lepoint, T., Lyubashevsky, V., Schwabe, P., Seiler, G., and Stehlé, D.: CRYSTALS-Dilithium: A lattice-based digital signature scheme. TCHES 2018(1), 238–268 (2018)

[DM15] Ducas, L., and Micciancio, D.: FHEW: Bootstrapping homomorphic encryption in less than a second. In: EUROCRYPT 2015, pp. 617–640 (2015)

[EH14] Erven, T. van, and Harremos, P.: Rényi divergence and Kullback-Leibler divergence. IEEE Trans. Inf. Theory 60(7), 3797–3820 (2014)

[FHK+20] Fouque, P.-A., Hoffstein, J., Kirchner, P., Lyubashevsky, V., Pornin, T., Prest, T., Ricosset, T., Seiler, G., Whyte, W., and Zhang, Z.: Falcon: Fast-Fourier lattice-based compact signatures over NTRU, Supporting documentation, NIST Post-Quantum Cryptography Standardization (2020), https://csrc.nist.gov/Projects/post-quantum-cryptography/post-quantum-cryptography-standardization/round-3-submissions

[GGH13] Garg, S., Gentry, C., and Halevi, S.: Candidate multilinear maps from ideal lattices. In: EUROCRYPT 2013, pp. 1–17 (2013)

[GM18] Genise, N., and Micciancio, D.: Faster Gaussian sampling for trapdoor lattices with arbitrary modulus. In: EUROCRYPT 2018, pp. 174–203 (2018)

[GMPW20] Genise, N., Micciancio, D., Peikert, C., and Walter, M.: Improved discrete Gaussian and subGaussian analysis for lattice cryptography. In: PKC 2020, pp. 623–651 (2020)

[GPV08] Gentry, C., Peikert, C., and Vaikuntanathan, V.: Trapdoors for hard lattices and new cryptographic constructions. In: STOC '08, pp. 197–206 (2008)

[GSW13] Gentry, C., Sahai, A., and Waters, B.: Homomorphic encryption from learning with errors: Conceptually-simpler, asymptotically-faster, attribute-based. In: CRYPTO 2013, pp. 75–92 (2013)

[GV96] Golub, G.H., and Van Loan, C.F.: Matrix computations (3rd ed.) Johns Hopkins University Press, USA (1996)

[HILL99] HÅstad, J., Impagliazzo, R., Levin, L.A., and Luby, M.: A pseudorandom generator from any one-way function. SIAM J. Comput. 28(4), 1364–1396 (1999)

[HKM18] Herold, G., Kirshanova, E., and May, A.: On the asymptotic complexity of solving LWE. Des. Codes Cryptogr. 86(1), 55–83 (2018)

[HJ85] Horn, R.A., and Johnson, C.R.: Matrix analysis. Cambridge University Press, Cambridge (1985)

[IZ89] Impagliazzo, R., and Zuckerman, D.: How to recycle random bits. In: FOCS '89, pp. 248–253 (1989)

[KNSW20] Kirshanova, E., Nguyen, H., Stehlé, D., and Wallet, A.: On the smoothing parameter and last minimum of random orthogonal lattices. Des. Codes Cryptogr. 88(5), 931–950 (2020)

[LSS14] Langlois, A., Stehlé, D., and Steinfeld, R.: GGHLite: More efficient multilinear maps from ideal lattices. In: EUROCRYPT 2014, pp. 239–256 (2014)

[Lyu05] Lyubashevsky, V.: The parity problem in the presence of noise, decoding random linear codes, and the subset sum problem. In: RANDOM 2005, pp. 378–389 (2005)

[MM11] Micciancio, D., and Mol, P.: Pseudorandom knapsacks and the sample complexity of LWE search-to-decision reductions. In: CRYPTO 2011, pp. 465–484 (2011)

[MP12] Micciancio, D., and Peikert, C.: Trapdoors for lattices: simpler, tighter, faster, smaller. In: EUROCRYPT 2012, pp. 700–718 (2012)

[MR07] Micciancio, D., and Regev, O.: Worst-case to average-case reductions based on Gaussian measures. SIAM J. Comput. 37(1), 267–302 (2007)

[MW17] Micciancio, D., and Walter, M.: Gaussian sampling over the integers: efficient, generic, constant-time. In: CRYPTO 2017, pp. 455–485 (2017)

[Ost38] Ostrowski, A.: Sur l'approximation du déterminant de fredholm par les déterminants des systèmes d'equations linéaires. Ark. Math. Stockholm Ser. A 26, 1–15 (1938)

[Pei10] Peikert, C.: An efficient and parallel Gaussian sampler for lattices. In: CRYPTO 2010, pp. 80–97 (2010)

[PS21] Pellet-Mary, A., and Stehlé, D.: On the hardness of the NTRU problem. In: ASIACRYPT 2021, pp. 3–35 (2021)

[Pre17] Prest, T.: Sharper bounds in lattice-based cryptography using the Rényi divergence. In: ASIACRYPT 2017, pp. 347–374 (2017)

[Reg09] Regev, O.: On lattices, learning with errors, random linear codes, and cryptography. J. ACM 56(6) (2009). Preliminary version in STOC '05

[Rén61] Rényi, A.: On measures of entropy and information. In: Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics, pp. 547–561 (1961)

[Zha05] Zhan, X.: Extremal eigenvalues of real symmetric matrices with entries in an interval. SIAM J. Matrix Anal. Appl. 27(3), 851–860 (2005)