GLITCH: A Discrete Gaussian Testing Suite For Lattice-Based Cryptography

James Howe and Máire O'Neill
Centre for Secure Information Technologies (CSIT), Queen's University Belfast, Northern Ireland.
{jhowe02,maire.oneill}@qub.ac.uk

Keywords: Post-quantum cryptography, lattice-based cryptography, discrete Gaussian samplers, discrete Gaussian distribution, random number generators, statistical analysis.

Abstract: Lattice-based cryptography is one of the most promising areas within post-quantum cryptography, and offers versatile, efficient, and high performance security services. The aim of this paper is to verify the correctness of the discrete Gaussian sampling component, one of the most important modules within lattice-based cryptography. In this paper, the GLITCH software test suite is proposed, which performs statistical tests on discrete Gaussian sampler outputs. An incorrectly operating sampler, for example due to hardware or software errors, has the potential to leak secret-key information and could thus be a potential attack vector for an adversary. Moreover, statistical test suites are already common for use in pseudo-random number generators (PRNGs), and as lattice-based cryptography becomes more prevalent, it is important to develop a method to test the correctness and randomness for discrete Gaussian sampler designs. Additionally, due to the theoretical requirements for the discrete Gaussian distribution within lattice-based cryptography, certain statistical tests for distribution correctness become unsuitable, therefore a number of tests are surveyed. The final GLITCH test suite provides 11 adaptable statistical analysis tests that assess the exactness of a discrete Gaussian sampler, and which can be used to verify any software or hardware sampler design.

1 Introduction

Post-quantum cryptography as a research field has grown substantially recently, essentially due to the growing concerns posed by quantum computers. The proviso being to provide long-term and highly secure cryptography, practical in comparison to RSA/ECC, but more importantly being adequately safe from quantum computers. This requirement is also hastened by the need for "future proofing" currently secure data, ensuring current IT infrastructures are quantum-safe *before* large-scale quantum computers are realised (Campagna et al., 2015).

As such, government agencies, companies, and standards agencies are planning transitions towards quantum-safe algorithms. The Committee on National Security Systems (CNSS) (CNSS, 2015) and the National Technical Authority for Information Assurance (CESG/NCSC) (CESG, 2016) are now planning drop-in quantum-safe replacements for current cryptosystems. The ETSI Quantum-Safe Cryptography (QSC) Industry Specification Group (ISG) (Campagna et al., 2015) is also highly active in researching industrial requirements for quantum-safe real-world deployments. NIST (Moody, 2016) have also called for quantum-resistant cryptographic algorithms for new public-key cryptography standards, similar to previous AES and SHA-3 competitions.

Lattice-based cryptography (Ajtai, 1996; Regev, 2005) is a very promising candidate for quantum-safe cryptography. Lattice-based cryptography bases its hardness on finding the shortest (or closest) vector in a lattice, which is currently resilient to *all known* quantum reductions and hence attacks by a quantum computer. Furthermore, lattice-based cryptography also offers extended functionality whilst being more efficient than ECC and RSA based primitives of public-key encryption (Pöppelmann and Güneysu, 2014) and digital signature schemes (Howe et al., 2015).

Lattice-based cryptoschemes are usually founded on either the learning with errors problem (LWE) (Regev, 2005) or the short integer solution problem (SIS) (Ajtai, 1996) or variants of these over ideal lattices. The general idea within lattice-based cryptosystems is to hide computations on secret-data with noise, usually discrete Gaussian noise, which would otherwise be retrievable via Gaussian elimination. The rationale for using discrete Gaussian noise (as

opposed to another probability distribution) is that it allows for more efficient lattice-based algorithms, with smaller output sizes such as ciphertexts or signatures. Background on discrete Gaussian sampling techniques is provided by Dwarakanath and Galbraith (Dwarakanath and Galbraith, 2014) and Howe et al. (Howe et al., 2016).

The specifications for the discrete Gaussian noise within lattice-based cryptography are very precise. The statistical distance between the theoretical discrete Gaussian distribution and the one observed in practice should be overwhelmingly small (Peikert, 2010), usually at least as small as $2^{-\lambda}$ for $\lambda \in \{64, \ldots, 128\}$. Providing guidelines to test implementations of discrete Gaussian samplers is therefore necessary for real-world applications in order to prevent attacks exploiting biased samplers. Moreover, an erroneously operating sampler could affect the target security level of the overall lattice-based cryptoscheme.

Additionally, the test suite is applicable for lattice-based cryptoschemes whose outputs are also distributed via the discrete Gaussian distribution, such as lattice-based encryption schemes (Lindner and Peikert, 2011; Lyubashevsky et al., 2013) and digital signatures (Gentry et al., 2008; Ducas et al., 2013).

Indeed, a biased sampler or cryptoscheme could be a potential attack vector for an adversary. Operational errors or bugs within sampler software or hardware designs, could significantly effect the theoretical security of the lattice-based cryptoscheme. To combat these issues for PRNGs, the DIEHARD (Marsaglia, 1985; Marsaglia, 1993; Marsaglia, 1996) and NIST SP 800-22 Rev. 1a (Bassham III et al., 2010) test suites were created. This is therefore clearly needed for discrete Gaussian random number generators.

This research investigates and proposes a discrete Gaussian testing suite for lattice-based cryptography, named GLITCH, which tests the correctness of a generic discrete Gaussian sampler (or lattice-based cryptoscheme) design. GLITCH takes as input histogram data, thus being able to test any discrete Gaussian sampling design, either in hardware or software. This paper surveys statistical tests that could be used for this purpose, proposing 11 tests appropriate for use within lattice-based cryptography. These test the main parameters and the shape of the distribution, and include normality and graphical tests.

The next section provides prerequisites on the discrete Gaussian distribution. Section 3 details a survey of the tests considered for the discrete Gaussian test suite, and is furthered by the 11 tests considered in GLITCH. The results are then analysed in Section 5.

2 The Discrete Gaussian Distribution

The *discrete Gaussian distribution* or *discrete normal distribution* ($D_{\mathbb{Z},\sigma}$) over $\mathbb{Z}$ with mean $\mu = 0$ and parameter σ is defined to have a weight proportional to $\rho_\sigma(x) = \exp(-(x-\mu)^2/(2\sigma^2))$ for all integers x. The variable $S_\sigma = \rho_\sigma(\mathbb{Z}) = \sum_{k=-\infty}^{\infty} \rho_\sigma(k) \approx \sqrt{2\pi}\sigma$ is then defined so that the probability of sampling $x \in \mathbb{Z}$ from the distribution $D_{\mathbb{Z},\sigma}$ is $\rho_\sigma(x)/S_\sigma$. For applications within lattice-based cryptography, it is assumed that these parameters are fixed and known in advanced.

Theoretically, the discrete Gaussian distribution has infinitely long tails and infinitely high precision, therefore in practice compromises have to be made which do not hinder the integrity of the scheme. The discrete Gaussian parameters needed are $(\mu, \sigma, \lambda, \tau)$; representing the sampler's centre, standard deviation, precision, and tail-cut, respectively.

The mean (μ) is the centre of a normalised distribution. Within lattice-based cryptography, the mean is usually set to $\mu = 0$.

The standard deviation (σ) controls the distribution's shape by quantifying the dispersion of data from the mean. The standard deviation depends on the modulus used within LWE or SIS. For instance in LWE, should σ be too small the hardness assumption may become easier than expected, and if σ is too large the problem may not be as well-defined as required.

The precision parameter (λ) governs the level of precision required for an implementation, exacting the statistical distance between the "perfect" theoretical discrete Gaussian distribution and the "practical" to be no greater than $2^{-\lambda}$, corresponding directly to the scheme's security level.

The tail-cut parameter (τ) administers the exclusion point on the x-axis, for a particular security level. That is, given a target security level of b-bits, the target distance from "perfect" need be no less than 2^{-b}. Thus, instead of considering $|x| \in \{0, \infty\}$, it is instead considered as $|x| \in \{0, \sigma\tau\}$. Applying the reduction in precision also affects the tail-cut parameter, which is calculated as $\tau = \sqrt{\lambda \times 2 \times \ln(2)}$.

These parameters are chosen via the scheme's security proofs. For example, the Lindner-Peikert lattice-based encryption scheme (Lindner and Peikert, 2011) requires parameters $(\mu = 0, \sigma = 3.33, \lambda = 128, \tau = 13.3)$ and the BLISS lattice-based signature scheme (Ducas et al., 2013) requires much larger parameters $(\mu = 0, \sigma = 215, \lambda = 128, \tau = 13.3)$. The next section presents a variety of statistical tests to check these parameters from observed data outputs from a generic discrete Gaussian sampler.

3 A Discrete Gaussian Testing Suite

This section describes the GLITCH discrete Gaussian testing suite for use within lattice-based cryptography. To the best of the authors' knowledge, this is the first proposal for testing the outputs of discrete Gaussian samplers for use within lattice-based cryptography. That is, if the samplers are *actually* producing the distribution required for specific values for μ, σ, τ, and λ. GLITCH can also be applied to outputs of cryptoschemes which follow the discrete Gaussian distribution, such as the BLISS signature scheme.

3.1 Statistical Testing Within Cryptography

Statistical testing is used to estimate the likelihood of a hypothesis given a set of data. For example, in cryptanalysis, statistical testing is commonly used to detect non-randomness in data, that is to distinguish the output of a PRNG from a truly random bitstream or to find the correctly decrypted message. The need for random and pseudorandom numbers arises in many cryptographic applications. For example, common cryptosystems employ keys that must be generated in a random fashion. Many cryptographic protocols also require random or pseudorandom inputs at various points, for example, for auxiliary quantities used in generating digital signatures, or for generating challenges in authentication protocols.

Moreover, the inclusion of statistical tests is paramount when implementing cryptography in practice. For example, to test a PRNG for cryptographically adequate randomness, the test suites DIEHARD (Marsaglia, 1985; Marsaglia, 1993; Marsaglia, 1996) and NIST SP 800-22 Rev. 1a (Bassham III et al., 2010) were proposed to check for insecure randomness, that is, to test a PRNG for weaknesses that an adversary could exploit.

3.2 Statistical Testing For Lattice-Based Cryptography

To exploit or attack a PRNG, an algorithm could determine the deviation of its output from that of a truly uniformly random deviation. This is especially important for the discrete Gaussian distribution within lattice-based cryptography, since these values hide secret information. Normality tests can be used to determine if, and how well, a data set follows the required normally structured distribution. More specifically, statistical hypothesis testing is used, which under the null hypothesis (H_0), states that the data is normally distributed. The alternative hypothesis (H_a), states that the data is not normally distributed. All of the methods proposed for testing the correctness of a discrete Gaussian sampler design only require an input of histogram values output from the sampler.

For the test suite, two normality tests are adopted, each using the same statistics of the discrete Gaussian samples, by producing two important (and somewhat distinct) results. Both also follow the same hypotheses; the null hypothesis that the sample data is normally distributed, and the alternative hypothesis that they are not normally distributed.

The first test considered is the Jarque-Bera (Jarque and Bera, 1987) goodness-of-fit test, which takes the skewness and kurtosis from the sample data, and matches it with the discrete Gaussian distribution. It tests the shape of the sampled distribution, rather than dealing with expected values, which makes the test significantly simpler than, say, a χ^2 test. Interestingly, if the sample data is normally distributed, the test statistic from the Jarque-Bera test asymptotically follows a χ^2 distribution with two degrees of freedom, which is then used in the hypothesis test.

The second test is the D'Agostino-Pearson K^2 omnibus test (D'Agostino et al., 1990), and is another goodness-of-fit test using the sample skewness and kurtosis. This test however is an omnibus test, which tests whether the explained deviation in the sample data is significantly greater than the overall unexplained deviation. The test also has the same asymptotic property as the Jarque-Bera test.

D'Agostino et al. (D'Agostino et al., 1990) analyse the asymptotic performances of more commonly used normality tests; those being the χ^2 test, Kolmogorov test (Kolmogorov, 1956), and the Shapiro-Wilk W-test (Shapiro and Wilk, 1965). These are important results, since the sample sizes required are far beyond those used in typical applications, in say, medicine or econometrics. Additionally it is recommended not to use the χ^2 test and Kolmogorov test, due to their poor power properties. That is, for a large sample size, the probability of making a Type II error (that is, incorrectly retaining a false null hypothesis) significantly increases. Furthermore, for sample sizes $N > 50$, D'Agostino et al. state the Shapiro-Wilk W-test is no longer available, and even with the test extended ($N \leq 2000$) (Royston, 1982), it still falls below the required sample size. The final major test for normality is the Anderson-Darling test (Anderson and Darling, 1952; Anderson and Darling, 1954). However, the D'Agostino-Pearson K^2 omnibus test is preferred since the Anderson-Darling test is biased towards the tails of the distribution (Razali et al., 2011).

The final tests are graphical. The first simply plots the observed histogram data versus the expected

Table 1: Details of the GLITCH software test suite.

Test Number	Test Description	Test Formula
Test 1	Sample Mean ($\bar{x}$)	$\bar{x} = (\sum_{i=1}^{N} x_i h_i)/N$
	Standard Error of $\bar{x}$	$\mathrm{SE}_{\bar{x}} = s/\sqrt{N}$
	Confidence Interval of $\bar{x}$	$\bar{x} \pm t_{\alpha/2}\mathrm{SE}_{\bar{x}}$
	Accept Null Hypothesis?	Accept if $\lvert\mu\rvert \in \{0, \ldots, \bar{x} + t_{\alpha/2}\mathrm{SE}_{\bar{x}}\}$
Test 2	Sample Standard Deviation (s)	$s = \sqrt{(\sum_{i=1}^{N}(x_i - \bar{\mu}_1)^2 h_i)/N}$
	Standard Error of s	$\mathrm{SE}_s = s/\sqrt{2(N-1)}$
	Confidence Interval of s	$s \pm t_{\alpha/2}\mathrm{SE}_s$
	Accept Null Hypothesis?	Accept if $\lvert\sigma\rvert \in \{0, \ldots, s + t_{\alpha/2}\mathrm{SE}_s\}$
Test 3	Sample Tail-Cut ($\bar{\tau}$)	$\bar{\tau} = \max(x_i)/s$
Test 4	Sample Skewness (ω)	$\omega = m_3\sqrt{N(N-1)}/(N-2)$
	Standard Error of ω	$\mathrm{SE}_\omega = \sqrt{(6N(N-1))/((N-2)(N+1)(N+3))}$
Test 5	Sample Excess Kurtosis (κ)	$\kappa = (m_4/s^4) - 3$
	Standard Error of κ	$\mathrm{SE}_\kappa = 2\mathrm{SE}_\omega\sqrt{(N^2-1)/((N-3)(N+5))}$
Test 6	Sample Hyperskewness	$\omega_* = m_5/s^5$
Test 7	Sample Excess Hyperkurtosis	$\kappa_* = m_6/s^6$
Test 8	Jarque-Bera Test For Normality	$\mathrm{JB} = (N/6)(\omega^2 + ((\kappa-3)^2)/4)$
	Accept Null Hypothesis?	Accept if $\mathrm{JB} < \chi^2_\alpha$
Test 9	D'Agostino-Pearson Omnibus Test	$\mathrm{K}^2 = Z_1(\omega)^2 + Z_2(\kappa)^2$
	Accept Null Hypothesis?	Accept if $\mathrm{K}^2 < \chi^2_\alpha$
Test 11	Coefficient of Determination	$R^2 = 1 - (\sum_{i=1} e_i^2 / \sum_{i=1}(y_i - \hat{y})^2)$

data. The second graphic is a quantile-quantile (QQ) plot. This test illustrates how strongly the histogram data follows a discrete Gaussian distribution, providing a QQ-plot and coefficient of determination (R^2). The QQ-plot is supplementary to the numerical assessment of normality and is a graphical method for comparing two probability distributions. In this case, these two probability distributions are the observed and expected quantiles of the discrete Gaussian distribution. This test is essentially the same as a probability-probability (PP) plot, wherein a data set is plotted against its target theoretical distribution. However, QQ-plots have the ability to arbitrarily choose the precision (to equal that of λ, say 128-bits) as well as being easier to interpret in the case of large sample sizes, hence its inclusion over PP-plots.

The R^2 value complements this plot, analysing how well the linear reference line approximates the expected data points. The output $R^2 \in [0,1]$ is a measure of the proportion of total variance of the outcomes, which is explained by the model. Therefore, the higher the R^2 value, the better the model fits the data.

3.3 The GLITCH Test Suite

The GLITCH test suite is provided in Python and is publicly available online[1]. Additionally, discrete Gaussian data sets are provided. Concise details for GLITCH are given in Table 1. GLITCH is designed to take, as input, a *histogram* of discrete Gaussian samples. This is seen as advantageous over an input of listed samples, as calculations are significantly simplified, are significantly faster, and decrease storage. The suite of tests are specifically chosen so that each parameter in the discrete Gaussian sampling stage is tested. The main parameters under test are the mean and standard deviation of the discrete Gaussian (μ, σ), with additional tests included to check the shape of the distribution, and finally normality tests. Precision is also adaptable and set to 128-bits as per most lattice-based cryptoschemes.

3.3.1 Tests (1-3): Testing Parameters

The first set of tests are to approximate the main statistical parameters μ and σ, producing values for sample mean ($\bar{x}$) and sample standard deviation (s). This is done by using adapted formulas for the first (m_1) and second (m_2) moments, taking as input a histogram of values (x_i, h_i), where $m_1 = \bar{x} = (\sum_{i=1}^{N} x_i h_i)/N$ corresponding to the sample mean, and $m_2 = s^2 = (\sum_{i=1}^{N}(x_i - \bar{x})^2 h_i)/N$ corresponding to the sample variance, for a sample size N. The subsequent moments are then $m_k = (\sum_{i=1}^{N}(x_i - \bar{x})^k h_i/N)/\sigma^k$, using sample standard deviation $s = \sqrt{m_2}$.

Next, the standard error (SE) is calculated for the sampling distribution. This statistic measures the reliability of a given sample's descriptive statistics with

[1]GLITCH software test suite available at `https://github.com/jameshoweee/glitch`

respect to the population's target values, that is, the mean and standard deviation. Additionally, the standard error is used in measuring the confidence in the sample mean and sample standard deviation. For this, a two-tail t-test is constructed, given the null hypothesis $\mu = 0$ (similarly for σ), with the alternate hypothesis that they are not equal. So, if the null hypothesis is accepted, it is concluded that a $100(1-\alpha)\%$ confidence interval (C.I.) is $\bar{x} \pm \varepsilon_{\bar{x}}$ and $s \pm \varepsilon_s$, where $\varepsilon_{\bar{x}} = t_{\alpha/2}\mathrm{SE}_{\bar{x}}$ and $\varepsilon_s = t_{\alpha/2}\mathrm{SE}_s$. Since the aim of these tests if for the highest confidence (99.9%), $t_{\alpha/2} = 3.29$.

3.3.2 Tests (4-7): Testing the Distribution's Shape

The next set of tests deal with statistical descriptors of the shape of the probability distribution. The first descriptor is the skewness; which is a measure of symmetry of the probability distribution and is adapted from the third moment. The skewness for a normally shaped distribution, or any symmetric distribution, is zero. Moreover, a negative skewness implies the left-tail is long, relative to the right-tail, and a positive skewness implies a long right-tail, relative to the left-tail. The population skewness is simply m_3/s^3, however the sample skewness must be adapted to $\omega = m_3\sqrt{N(N-1)}/N-2$ to account for bias (Joanes and Gill, 1998). Also SE_ω is calculated, to show the relationship between the expected skewness and ω.

The forth moment is kurtosis; and describes the peakedness of a distribution. For a normally shaped distribution, the target sampled kurtosis is three, and is calculated as m_4/s^4. More commonly, the sampled excess kurtosis is used and is defined as $\kappa = (m_4/s^4) - 3$. A positive kurtosis indicates a peaked distribution, similarly a negative kurtosis indicates a flat distribution. It can also be seen, given an increase in kurtosis, that probability mass has moved from the shoulders of the distribution, to its centre and tails (Balanda and MacGillivray, 1988). Similarly, SE_κ is calculated to show the relationship between the expected excess kurtosis and κ.

An appropriate test for these statistical descriptors would be a z-test, where confidence intervals could also be calculated for some confidence level α. However, under a null hypothesis of normality, z-tests tend to be easily rejected for larger samples ($N > 300$) taken from a not substantially different normal distribution (Kim, 2013).

Higher-order moments, specifically the fifth and sixth, are used in the last two tests on the distribution's shape. The first of these tests hyper-skewness $\omega_* = m_5/s^5$, which still measures symmetry but is more sensitive to extreme values (Hinton, 2014, p.97). Likewise, the second of these tests is for excess hyper-kurtosis $\kappa_* = m_6/s^6$, which tests for peakedness with greater sensitivity towards more-than-expected weight in the tails (Hinton, 2014, p.100).

3.3.3 Tests (8-9): Normality Testing

These tests calculate the test statistic and p-value for the two normality tests described in Section 3, these are the Jarque-Bera (Jarque and Bera, 1987) and D'Agostino-Pearson (D'Agostino et al., 1990) omnibus tests. The Jarque-Bera test statistic is calculated as $\mathrm{JB} = (N/6)(\omega^2 + ((\kappa-3)^2/4))$, where its p-value is taken from a χ^2 distribution with two degrees of freedom. The null hypothesis (of normality) is rejected if the test statistic is greater than the χ^2 p-value.

The D'Agostino-Pearson omnibus test is based on transformations of the sample skewness ($Z_1(\omega)$) and sample kurtosis ($Z_2(\kappa)$), which are combined to produce an omnibus test. This statistic detects deviations from normality due to either skewness or kurtosis and is defined as $\mathrm{K}^2 = Z_1(\omega)^2 + Z_2(\kappa)^2$.

3.3.4 Tests (10-11): Illustrating Normality

D'Agostino et al. (D'Agostino et al., 1990) recommend, as well as test statistics for normality, graphical representations of normality are also provided. Hence, the final two tests are illustrative tests on the discrete Gaussian samples. The first graphic, shown in Figures 1 and 3, plots the histogram of discrete Gaussian observed values (in blue) alongside the expected values (in red).

The second graphic is a quantile-quantile (QQ) plot, shown in Figures 2 and 4. For this test, the calculated z-scores are plotted against the expected z-scores, where if the data is normally distributed, the result will be a straight diagonal line (Field, 2009, p.145-148). A 45-degree reference line is plotted, which will overlap with the QQ-plot if the distribution follows the required distribution.

The coefficient of determination (R^2) value is calculated as $R^2 = 1 - (SS_{\mathrm{res}}/SS_{\mathrm{tot}})$, where $SS_{\mathrm{res}} = \sum_i (y_i - f_i)^2$ is the residual sum of squares and $SS_{\mathrm{tot}} = \sum_i (y_i - \bar{x})^2$ is the total sum of squares, y_i is the observed data set and f_i is the expected values.

4 Results

Example results are provided in Listings 1 and 2. The results used are from a Bernoulli sampler. The first data set passes all tests, as shown in Listing 1. A second data set, used in Listing 2, is generated with an incorrect standard deviation and fails test 2, showing that GLITCH detects errors in discrete Gaussian

Listing 1: GLITCH test suite output for a working discrete Gaussian sampler of sample size 2^{36}. Results meet all requirements for expected values and passes all hypothesis tests.

```
// GLITCH: Discrete Gaussian Sampling Test Results:
// Target Sigma: 215.72773727315683 --Sampler: bernoulli --Sample Size: 68719476736

(1) Sample Mean: 0.0013711968495044857263565063 48
    Standard Error of the Mean: 0.0008229334036229891103988139999
    C.I. of the Sample Mean = 0.0013711968495044857263565063 48 +/- 0.0027074508979196342024485656 58
     with 99.9% confidence
     *Accept* Null Hypothesis for Sample Mean with 99.9% Confidence

(2) Sample Standard Deviation: 215.72705415934485735638669 72
    Standard Error of the Standard Deviation: 0.0005819017901709756494057053363
    C.I. of the Sample Standard Deviation = 215.72705415934485735638669 72 +/-
     0.0019144568896625099072180750 53 with 99.9% confidence
     *Accept* Null Hypothesis for Sample Standard Deviation with 99.9% Confidence

(3) Sample Tail-Cut Parameter (Tau): 13.002541618762901478679369 00
    Distance from Target Tail-Cut: 0.318332159760261532620943064 1

(4) Sample Skewness: -0.000017638359799331235831998350 26
    Standard Error of the Sample Skewness: 0.000009344061823767513075122646283

(5) Sample Excess Kurtosis: -0.000024501635887997449631126
    Standard Error of the Sample Kurtosis: 0.000018688123647263078159182762 54

(6) Sample Hyperskewness: -0.000095722132334071148027153873 28

(7) Sample Excess Hyperkurtosis: -0.000253380610515845297843 36

(8) Jarque-Bera Normality Test ( test stat , p-value): 4.3121664872166712749365391 66E-7, 0.999999784392
     *Accept* Null Hypothesis of Normality (p-value) with 99.9% Confidence
     *Accept* Null Hypothesis of Normality ( test stat ) with 99.9% Confidence

(9) D'Agostino-Pearson K2 Omnibus Test (test stat , p-value): 0.0029340347293438886905514431457,
     0.998534058179
     *Accept* Null Hypothesis of Normality (p-value) with 99.9% Confidence
     *Accept* Null Hypothesis of Normality ( test stat ) with 99.9% Confidence

(10) Histogram and Quantile-Quantile (QQ) plots:
```

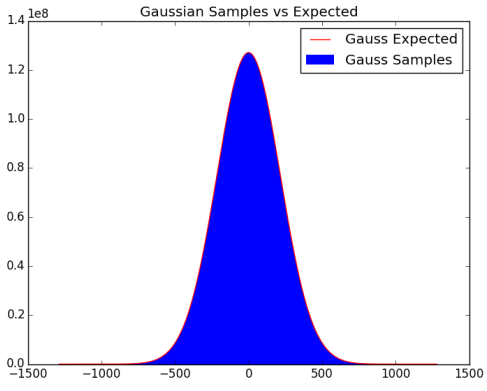

Figure 1: Histogram of observed (blue) discrete Gaussian samples versus expected (red). The observed data matches the expected values.

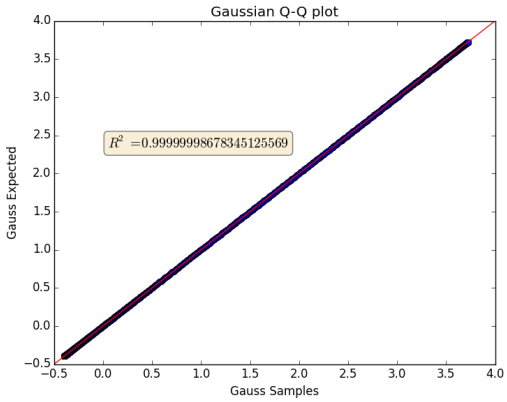

Figure 2: QQ-plot of the observed discrete Gaussian samples with the coefficient of determination (R^2) value. Both red and blue lines overlap meaning the observed data matches the expected values.

Listing 2: GLITCH test suite output for a working discrete Gaussian sampler of sample size 2^{36}, but with incorrect target standard deviation. Null hypothesis is rejected for test 2.

```
// GLITCH: Discrete Gaussian Sampling Test Results:
// Target Sigma: 250 --Sampler: bernoulli --Sample Size: 68719476736

(1) Sample Mean:                          0.000415780508774332702159881591 8
    Standard Error of the Mean:           0.000801084797793829643742633526 6
    C.I. of the Sample Mean =             0.000415780508774332702159881591 8 +/- 0.002635568984741699556373513493
     with 99.9% confidence
     *Accept* Null Hypothesis for Sample Mean with 99.9% Confidence

(2) Sample Standard Deviation:                  209.99957323286567812926892 32
    Standard Error of the Standard Deviation:   0.000566452492829592651241111943 7
    C.I. of the Sample Standard Deviation =     209.99957323286567812926892 32 +/-
     0.001863628701409359842707693491 with 99.9% confidence
     *Reject* Null Hypothesis for Sample Standard Deviation with 99.9% Confidence

(3) Sample Tail-Cut Parameter (Tau):  13.000026418971528252811178 78
    Distance from Target Tail-Cut:    0.320847359551634758489133284 1

(4) Sample Skewness:                          -0.000008109373516717886834916069369
    Standard Error of the Sample Skewness:    0.000009344061823767513075122646283

(5) Sample Excess Kurtosis:                   -0.000028095808185055005256698
    Standard Error of the Sample Kurtosis:    0.000018688123647263078159182762 54

(6) Sample Hyperskewness:           -0.000103651468047191991992233509974

(7) Sample Excess Hyperkurtosis:    -0.000519099829468152675819 23

(8) Jarque-Bera Normality Test ( test stat , p-value): 2.394260488861117097644603536E-7, 0.999999880287
     *Accept* Null Hypothesis of Normality (p-value) with 99.9% Confidence
     *Accept* Null Hypothesis of Normality ( test stat ) with 99.9% Confidence

(9) D'Agostino-Pearson K2 Omnibus Test (test stat , p-value): 0.003004329115990208729071137285,
     0.998498963126
     *Accept* Null Hypothesis of Normality (p-value) with 99.9% Confidence
     *Accept* Null Hypothesis of Normality ( test stat ) with 99.9% Confidence

(10) Histogram and Quantile-Quantile (QQ) plots:
```

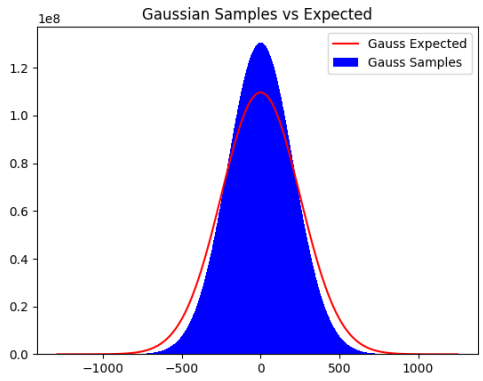

Figure 3: Histogram of observed (blue) discrete Gaussian samples versus expected (red). The observed data does not match the expected values.

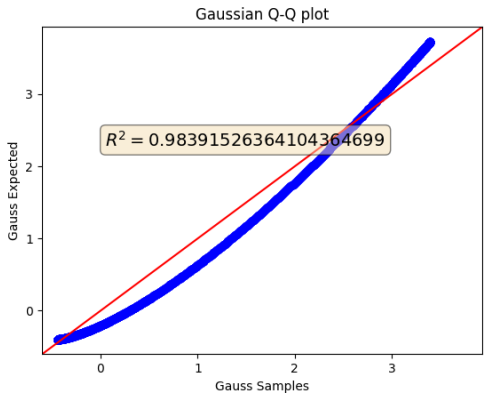

Figure 4: QQ-plot of the observed discrete Gaussian samples with the coefficient of determination (R^2) value. Red and blue lines do not overlap meaning the observed data does not match the expected values.

samplers. Additionally, this failure is illustrated in Figures 3 and 4, with expected and observed values not matching. Data sets provided are of size 2^{36}. In general, the sample size for GLITCH should be large enough so that extreme values in the discrete Gaussian tails are likely to be filled.

5 Conclusion

The research on statistical testing for discrete Gaussian samples reapplies well established statistical testing techniques to lattice-based cryptography, taking into consideration the stringent requirements within the area. This was completed by conducting a full survey on a number of different testing techniques, collating the relevant tests to form the adaptable GLITCH software statistical test suite.

The first number of tests are for analysing the main discrete Gaussian parameters from the observed data, giving standard error, confidence intervals, and (where possible) hypothesis tests with the highest level of confidence (99.9%). The next set of tests verifies the shape of the distribution, analysing whether there is any bias towards the positive or negative side of the distribution, and whether the distribution has a bias towards the peak of the distribution. For these tests and for the following tests on normality, the tests which allow for samples sizes large enough for lattice-based cryptography constraints are chosen. The last tests illustrate the difference between the observed data's distribution and the expected distribution's shape.

The tests chosen are powerful and operate well on large sample sizes, with each analysing differing aspects within the discrete Gaussian distribution. Failure in any of these tests indicates a deviation from the target distribution, which is therefore evidence of an incorrectly performing discrete Gaussian sampler. The software for GLITCH is made available online (available at `https://github.com/jameshoweee/glitch`), which provides sample data for discrete Gaussian samplers; which are able to be tested upon.

Acknowledgements

This paper is an extended/full version of (Howe and O'Neill, 2017), which appeared at SECRYPT 2017. The authors would like to thank Thomas Pöppelmann and acknowledge that this work was supported by a STSM Grant from COST Action TD1207.

REFERENCES

Ajtai, M. (1996). Generating hard instances of lattice problems (extended abstract). In *STOC*, pages 99–108.

Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The Annals of Mathematical Statistics*, 23(2):193–212.

Anderson, T. W. and Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, 49(268):765–769.

Balanda, K. P. and MacGillivray, H. (1988). Kurtosis: a critical review. *The American Statistician*, 42(2):111–119.

Bassham III, L. E., Rukhin, A. L., Soto, J., Nechvatal, J. R., Smid, M. E., Barker, E. B., Leigh, S. D., Levenson, M., Vangel, M., Banks, D. L., Heckert, N. A., Dray, J. F., and Vo, S. (2010). SP 800-22 Rev. 1a. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. Technical report, Gaithersburg, MD, United States.

Campagna, M., Chen, L., Dagdelen, Ö., Ding, J., Fernick, J., Gisin, N., Hayford, D., Jennewein, T., Lütkenhaus, N., Mosca, M., et al. (2015). Quantum safe cryptography and security. *ETSI White Paper*, (8).

CESG (2016). Quantum key distribution: A CESG white paper.

CNSS (2015). Use of public standards for the secure sharing of information among national security systems. Committee on National Security Systems: CNSS Advisory Memorandum, Information Assurance 02-15.

D'Agostino, R. B., Belanger, A., and D'Agostino Jr, R. B. (1990). A suggestion for using powerful and informative tests of normality. *The American Statistician*, 44(4):316–321.

Ducas, L., Durmus, A., Lepoint, T., and Lyubashevsky, V. (2013). Lattice signatures and bimodal Gaussians. In *CRYPTO (1)*, pages 40–56. Full version: `https://eprint.iacr.org/2013/383.pdf`.

Dwarakanath, N. C. and Galbraith, S. D. (2014). Sampling from discrete Gaussians for lattice-based cryptography on a constrained device. *Appl. Algebra Eng. Commun. Comput.*, pages 159–180.

Field, A. (2009). *Discovering statistics using SPSS*. Sage publications.

Gentry, C., Peikert, C., and Vaikuntanathan, V. (2008). Trapdoors for hard lattices and new cryptographic constructions. In *STOC*, pages 197–206.

Hinton, P. R. (2014). *Statistics explained*. Routledge.

Howe, J., Khalid, A., Rafferty, C., Regazzoni, F., and O'Neill, M. (2016). On Practical Discrete Gaussian Samplers For Lattice-Based Cryptography. *IEEE Transactions on Computers*.

Howe, J. and O'Neill, M. (2017). GLITCH: A Discrete Gaussian Testing Suite For Lattice-Based Cryptography. In *Proceedings of the 14th International Joint Conference on e-Business and Telecommunications (ICETE 2017): SECRYPT, Madrid, Spain, July 24-26, 2017.*

Howe, J., Pöppelmann, T., O'Neill, M., O'Sullivan, E., and Güneysu, T. (2015). Practical lattice-based digital signature schemes. *ACM Transactions on Embedded Computing Systems*, 14(3):24.

Jarque, C. M. and Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, pages 163–172.

Joanes, D. N. and Gill, C. A. (1998). Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(1).

Kim, H.-Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative dentistry & endodontics*, 38(1):52–54.

Kolmogorov, A. N. (1956). Foundations of the theory of probability (2nd ed.). Chelsea Publishing Co., New York.

Lindner, R. and Peikert, C. (2011). Better key sizes (and attacks) for LWE-based encryption. In *CT-RSA*, pages 319–339.

Lyubashevsky, V., Peikert, C., and Regev, O. (2013). On ideal lattices and learning with errors over rings. *J. ACM*, 60(6):43.

Marsaglia, G. (1985). A current view of random number generators. In *Computer Science and Statistics, Sixteenth Symposium on the Interface. Elsevier Science Publishers, North-Holland, Amsterdam*, pages 3–10.

Marsaglia, G. (1993). A current view of random numbers. In Billard, L., editor, *Computer Science and Statistics: Proceedings of the 16th Symposium on the Interface*, volume 36, pages 105–110. Elsevier Science Publishers B. V.

Marsaglia, G. (1996). DIEHARD: A battery of tests of randomness. `http://www.stat.fsu.edu/pub/diehard/`.

Moody, D. (2016). Post-quantum cryptography: NIST's plan for the future. Talk given at PQCrypto '16 Conference, 23-26 February 2016, Fukuoka, Japan.

Peikert, C. (2010). An efficient and parallel Gaussian sampler for lattices. In *CRYPTO*, pages 80–97.

Pöppelmann, T. and Güneysu, T. (2014). Area optimization of lightweight lattice-based encryption on reconfigurable hardware. In *ISCAS*, pages 2796–2799.

Razali, N. M., Wah, Y. B., et al. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of statistical modeling and analytics*, 2(1):21–33.

Regev, O. (2005). On lattices, learning with errors, random linear codes, and cryptography. In *STOC*, pages 84–93.

Royston, J. P. (1982). An extension of Shapiro and Wilk's W test for normality to large samples. *Applied Statistics*, pages 115–124.

Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, pages 591–611.