Hardness of SIS and LWE with Small Parameters

Daniele Micciancio* Chris Peikert†

February 13, 2013

Abstract

The Short Integer Solution (SIS) and Learning With Errors (LWE) problems are the foundations for countless applications in lattice-based cryptography, and are provably as hard as approximate lattice problems in the worst case. A important question from both a practical and theoretical perspective is how small their parameters can be made, while preserving their hardness.

We prove two main results on SIS and LWE with small parameters. For SIS, we show that the problem retains its hardness for moduli $q \geq \beta \cdot n^\delta$ for any constant $\delta > 0$, where β is the bound on the Euclidean norm of the solution. This improves upon prior results which required $q \geq \beta \cdot \sqrt{n \log n}$, and is essentially optimal since the problem is trivially easy for $q \leq \beta$. For LWE, we show that it remains hard even when the errors are small (e.g., uniformly random from $\{0, 1\}$), provided that the number of samples is small enough (e.g., linear in the dimension n of the LWE secret). Prior results required the errors to have magnitude at least $\sqrt{n}$ and to come from a Gaussian-like distribution.

1 Introduction

In modern lattice-based cryptography, two average-case computational problems serve as the foundation of almost all cryptographic schemes: Short Integer Solution (SIS), and Learning With Errors (LWE). The SIS problem dates back to Ajtai's pioneering work [1], and is defined as follows. Let n and q be integers, where n is the primary security parameter and usually $q = \mathrm{poly}(n)$, and let $\beta > 0$. Given a uniformly random matrix $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$ for some $m = \mathrm{poly}(n)$, the goal is to find a nonzero integer vector $\mathbf{z} \in \mathbb{Z}^m$ such that $\mathbf{Az} = \mathbf{0} \bmod q$ and $\|\mathbf{z}\| \leq \beta$ (where $\|\cdot\|$ denotes Euclidean norm). Observe that β should be set large enough to ensure that a solution exists (e.g., $\beta > \sqrt{n \log q}$ suffices), but that $\beta \geq q$ makes the problem trivially easy to solve. Ajtai showed that for appropriate parameters, SIS enjoys a remarkable worst-case/average-case hardness property: solving it *on the average* (with any noticeable probability) is at least as hard as approximating several lattice problems on n-dimensional lattices *in the worst case*, to within $\mathrm{poly}(n)$ factors.

*University of California, San Diego. 9500 Gilman Dr., Mail Code 0404, La Jolla, CA 92093, USA. Email: daniele@cs.ucsd.edu. This material is based on research sponsored by DARPA under agreement number FA8750-11-C-0096 and NSF under grant CNS-1117936. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA, NSF or the U.S. Government.

†School of Computer Science, Georgia Institute of Technology. This material is based upon work supported by the National Science Foundation under CAREER Award CCF-1054495, by DARPA under agreement number FA8750-11-C-0096, and by the Alfred P. Sloan Foundation.

The LWE problem was introduced in the celebrated work of Regev [24], and has the same parameters n and q, along with a "noise rate" $\alpha \in (0,1)$. The problem (in its search form) is to find a secret vector $\mathbf{s} \in \mathbb{Z}_q^n$, given a "noisy" random linear system $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$, $\mathbf{b} = \mathbf{A}^T\mathbf{s} + \mathbf{e} \bmod q$, where $\mathbf{A}$ is uniformly random and the entries of $\mathbf{e}$ are i.i.d. from a Gaussian-like distribution with standard deviation roughly αq. Regev showed that as long as $\alpha q \geq 2\sqrt{n}$, solving LWE on the average (with noticeable probability) is at least as hard as approximating lattice problems in the worst case to within $\tilde{O}(n/\alpha)$ factors using a *quantum* algorithm. Subsequently, Peikert [21] gave a *classical* reduction for a subset of the lattice problems and the same approximation factors, but under the additional condition that $q \geq 2^{n/2}$ (or $q \geq 2\sqrt{n}/\alpha$ based on some non-standard lattice problems).

A significant line of research has been devoted to improving the tightness of worst-case/average-case connections for lattice problems. For SIS, a series of works [1, 7, 14, 19, 12] gave progressively better parameters that guarantee hardness, and smaller approximation factors for the underlying lattice problems. The state of the art (from [12], building upon techniques introduced in [19]) shows that for $q \geq \beta \cdot \omega(\sqrt{n \log n})$, finding a SIS solution with norm bounded by β is as hard as approximating worst-case lattice problems to within $\tilde{O}(\beta\sqrt{n})$ factors. (The parameter m does not play any significant role in the hardness results, and can be any polynomial in n.) For LWE, Regev's initial result remains the tightest, and the requirement that $q \geq \sqrt{n}/\alpha$ (i.e., that the errors have magnitude at least $\sqrt{n}$) is in some sense optimal: a clever algorithm due to Arora and Ge [2] solves LWE in time $2^{\tilde{O}(\alpha q)^2}$, so a proof of hardness for substantially smaller errors would imply a subexponential time (quantum) algorithm for approximate lattice problems, which would be a major breakthrough. Interestingly, the current modulus bound for LWE is in some sense better than the one for SIS by a $\tilde{\Omega}(\sqrt{n})$ factor: there are applications of LWE for $1/\alpha = \tilde{O}(1)$ and hence $q = \tilde{O}(\sqrt{n})$, whereas SIS is only useful for $\beta \geq \sqrt{n}$, and therefore requires $q \geq n$ according to the state-of-the-art reductions.

Further investigating the smallest parameters for which SIS and LWE remain provably hard is important from both a practical and theoretical perspective. On the practical side, improvements would lead to smaller cryptographic keys without compromising the theoretical security guarantees, or may provide greater confidence in more practical parameter settings that so far lack provable hardness. Also, proving the hardness of LWE for non-Gaussian error distributions (e.g., uniform over a small set) would make applications easier to implement. Theoretically, improvements may eventually shed light on related problems like Learning Parity with Noise (LPN), which can be seen as a special case of LWE for modulus $q = 2$, and which is widely used in coding-based cryptography, but which has no known proof of hardness.

1.1 Our Results

We prove two complementary results on the hardness of SIS and LWE with small parameters. For SIS, we show that the problem retains its hardness for moduli q nearly equal to the solution bound β. For LWE, we show that it remains hard even when the errors are small (e.g., uniformly random from $\{0,1\}$), provided that the number m of noisy equations is small enough. This qualification is necessary in light of the Arora-Ge attack [2], which for large enough m can solve LWE with binary errors in polynomial time. Details follow.

SIS with small modulus. Our first theorem says that SIS retains its hardness with a modulus as small as $q \geq \beta \cdot n^\delta$, for any $\delta > 0$. Recall that the best previous reduction [12] required $q \geq \beta \cdot \omega(\sqrt{n \log n})$, and that SIS becomes trivially easy for $q \leq \beta$, so the q obtained by our proof is essentially optimal. It also essentially closes the gap between LWE and SIS, in terms of how small a useful modulus can be. More precisely, the following is a special case of our main SIS hardness theorem; see Section 3 for full details.

Theorem 1.1 (Corollary of Theorem 3.8). *Let n and $m = \mathrm{poly}(n)$ be integers, let $\beta \geq \beta_\infty \geq 1$ be reals, let $Z = \{\mathbf{z} \in \mathbb{Z}^m : \|\mathbf{z}\|_2 \leq \beta$ and $\|\mathbf{z}\|_\infty \leq \beta_\infty\}$, and let $q \geq \beta \cdot n^\delta$ for some constant $\delta > 0$. Then solving (on the average, with non-negligible probability)* SIS *with parameters n, m, q and solution set $Z \setminus \{\mathbf{0}\}$ is at least as hard as approximating lattice problems in the worst case on n-dimensional lattices to within $\gamma = \max\{1, \beta \cdot \beta_\infty/q\} \cdot \tilde{O}(\beta\sqrt{n})$ factors.*

Of course, the ℓ_∞ bound on the SIS solutions can be easily removed simply setting $\beta_\infty = \beta$, so that $\|\mathbf{z}\|_\infty \leq \|\mathbf{z}\|_2 \leq \beta$ automatically holds true. We include an explicit ℓ_∞ bound $\beta_\infty \leq \beta$ in order to obtain more precise hardness results, based on potentially smaller worst-case approximation factors γ. We point out that the bound β_∞ and the associated extra term $\max\{1, \beta \cdot \beta_\infty/q\}$ in the worst-case approximation factor is not present in previous results. Notice that this term can be as small as 1 (if we take $q \geq \beta \cdot \beta_\infty$, and in particular if $\beta_\infty \leq n^\delta$), and as large as β/n^δ (if $\beta_\infty = \beta$). This may be seen as the first theoretical evidence that, at least when using a small modulus q, restricting the ℓ_∞ norm of the solutions may make the SIS problem qualitatively harder than just restricting the ℓ_2 norm. There is already significant empirical evidence for this belief: the most practically efficient attacks on SIS, which use lattice basis reduction (e.g., [11, 8]), only find solutions with bounded ℓ_2 norm, whereas combinatorial attacks such as [5, 25] (see also [20]) or theoretical lattice attacks [9] that can guarantee an ℓ_∞ bound are much more costly in practice, and also require exponential space. Finally, we mention that setting $\beta_\infty \ll \beta$ is very natural in the usual formulations of one-way and collision-resistant hash functions based on SIS, where collisions correspond (for example) to vectors in $\{-1, 0, 1\}^m$, and therefore have ℓ_∞ bound $\beta_\infty = 1$, but ℓ_2 bound $\beta = \sqrt{m}$. Similar gaps between β_∞ and β can easily be enforced in other applications, e.g., digital signatures [12].

LWE with small errors. In the case of LWE, we prove a general theorem offering a trade-off among several different parameters, including the size of the errors, the dimension and number of samples in the LWE problem, and the dimension of the underlying worst-case lattice problems. Here we mention just one instantiation for the case of prime modulus and uniformly distributed *binary* (i.e., 0-1) errors, and refer the reader to Section 4 and Theorem 4.6 for the more general statement and a discussion of the parameters.

Theorem 1.2 (Corollary of Theorem 4.6). *Let n and $m = n \cdot (1 + \Omega(1/\log n))$ be integers, and $q \geq n^{O(1)}$ a sufficiently large polynomially bounded (prime) modulus. Then solving* LWE *with parameters n, m, q and independent uniformly random binary errors (i.e., in $\{0, 1\}$) is at least as hard as approximating lattice problems in the worst case on $\Theta(n/\log n)$-dimensional lattices within a factor $\gamma = \tilde{O}(\sqrt{n} \cdot q)$.*

We remark that our results (see Theorem 4.6) apply to many other settings, including error vectors $\mathbf{e} \in X$ chosen from any (sufficiently large) subset $X \subseteq \{0, 1\}^m$ of binary strings, as well as error vectors with larger entries. Interestingly, our hardness result for LWE with very small errors relies on the worst-case hardness of lattice problems in dimension $n' = O(n/\log n)$, which is smaller than (but still quasi-linear in) the dimension n of the LWE problem; however, this is needed only when considering very small error vectors. Theorem 4.6 also shows that if $\mathbf{e}$ is chosen uniformly at random with entries bounded by n^ϵ (which is still much smaller than $\sqrt{n}$), then the dimension of the underlying worst-case lattice problems (and the number $m - n$ of extra samples, beyond the LWE dimension n) can be linear in n.

The restriction that the number of LWE samples $m = O(n)$ be linear in the dimension of the secret can also be relaxed slightly. But some restriction is necessary, because LWE with small errors can be solved in polynomial time when given an arbitrarily large polynomial number of samples. We focus on linear $m = O(n)$ because this is enough for most (but not all) applications in lattice cryptography, including identity-based encryption and fully homomorphic encryption, when the parameters are set appropriately. (The one exception that we know of is the security proof for pseudorandom functions [3].)

1.2 Techniques and Comparison to Related Work

Our results for SIS and LWE are technically disjoint, and all they have in common is the goal of proving hardness results for smaller values of the parameters. So, we describe our technical contributions in the analysis of these two problems separately.

SIS with small modulus. For SIS, as a warm-up, we first give a proof for a special case of the problem where the input is restricted to vectors of a special form (e.g., binary vectors). For this restricted version of SIS, we are able to give a self-reduction (from SIS to SIS) which reduces the size of the modulus. So, we can rely on previous worst-case to average-case reductions for SIS as "black boxes," resulting in an extremely simple proof. However, this simple self-reduction has some drawbacks. Beside the undesirable restriction on the SIS inputs, our the reduction is rather loose with respect to the underlying worst-case lattice approximation problem: in order to establish the hardness of SIS with small moduli q (and restricted inputs), one needs to assume the worst-case hardness of lattice problems for rather large polynomial approximation factors. (By contrast, previous hardness results for larger moduli [19, 12] only assumed hardness for quasi-linear approximation factors.) We address both drawbacks by giving a direct reduction from worst-case lattice problems to SIS with small modulus. This is our main SIS result, and it combines ideas from previous work [19, 12] with two new technical ingredients:

- All previous SIS hardness proofs [1, 7, 14, 19, 12] solved worst-case lattice problems by iteratively finding (sets of linearly independent) lattice vectors of shorter and shorter length. Our first new technical ingredient (inspired by the pioneering work of Regev [24] on LWE) is the use a different intermediate problem: instead of finding progressively shorter lattice vectors, we consider the problem of sampling lattice vectors according to Gaussian-like distributions of progressively smaller widths. To the best of our knowledge, this is the first use of Gaussian lattice sampling as an intermediate worst-case problem in the study of SIS, and it appears necessary to lower the SIS modulus below n. We mention that Gaussian lattice sampling has been used before to reduce the modulus in hardness reductions for SIS [12], but still within the framework of iteratively finding short vectors (which in [12] are used to generate fresh Gaussian samples for the reduction), which results in larger moduli $q > n$.

- The use of Gaussian lattice sampling as an intermediate problem within the SIS hardness proof yields linear combinations of several discrete Gaussian samples with adversarially chosen coefficients. Our second technical ingredient, used to analyze these linear combinations, is a new convolution theorem for discrete Gaussians (Theorem 3.3), which strengthens similar ones previously proved in [22, 6]. Here again, the strength of our new convolution theorem appears necessary to obtain hardness results for SIS with modulus smaller than n.

Our new convolution theorem may be of independent interest, and might find applications in the analysis of other lattice algorithms.

LWE with small errors. We now move to our results on LWE. For this problem, the best provably hard parameters to date were those obtained in the original paper of Regev [24], which employed Gaussian errors, and required them to have (expected) magnitude at least $\sqrt{n}$. These results were believed to be optimal due to a clever algorithm of Arora and Ge [2], which solves LWE in subexponential time when the errors are asymptotically smaller than $\sqrt{n}$. The possibility of circumventing this barrier by limiting the number of LWE samples was first suggested by Micciancio and Mol [17], who gave "sample preserving" search-to-decision reductions for LWE, and asked if LWE with small uniform errors could be proved hard when the number

of available samples is sufficiently small. Our results provide a first answer to this question, and employ concepts and techniques from the work of Peikert and Waters [23] (see also [4]) on *lossy* (trapdoor) functions. In brief, a lossy function family is an indistinguishable pair of function families $\mathcal{F}, \mathcal{L}$ such that functions in $\mathcal{F}$ are injective and those in $\mathcal{L}$ are lossy, in the sense that they map their common domain to much smaller sets, and therefore lose information about the input. As shown in [23], from the indistinguishability of $\mathcal{F}$ and $\mathcal{L}$, it follows that the families $\mathcal{F}$ and $\mathcal{L}$ are both one-way.

In Section 2 we present a generalized framework for the study of lossy function families, which does not require the functions to have trapdoors, and applies to arbitrary (not necessarily uniform) input distributions. While the techniques we use are all standard, and our definitions are minor generalizations of the ones given in [23], we believe that our framework provides a conceptual simplification of previous work, relating the relatively new notion of lossy functions to the classic security definitions of second-preimage resistance and uninvertibility.

The lossy function framework is used to prove the hardness of LWE with small uniform errors and (necessarily) a small number of samples. Specifically, we use the standard LWE problem (with large Gaussian errors) to set up a lossy function family $\mathcal{F}, \mathcal{L}$. (Similar families with trapdoors were constructed in [23, 4], but not for the parameterizations required to obtain interesting hardness results for LWE.) The indistinguishability of $\mathcal{F}$ and $\mathcal{L}$ follows directly from the hardness of the underlying LWE problem. The new hardness result for LWE (with small errors) is equivalent to the one-wayness of $\mathcal{F}$, and is proved by a relatively standard analysis of the second-preimage resistance and uninvertibility of certain subset-sum functions associated to $\mathcal{L}$.

Comparison to related work. In an independent work that was submitted concurrently with ours, Döttling and Müller-Quade [10] also used a lossyness argument to prove new hardness results for LWE. (Their work does not address the SIS problem.) At a syntactic level, they use LWE (i.e., generating matrix) notation and a new concept they call "lossy codes," while here we use SIS (i.e., parity-check matrix) notation and rely on the standard notions of uninvertible and second-preimage resistant functions. By the dual equivalence of SIS and LWE [15, 17] (see Proposition 2.9), this can be considered a purely syntactic difference, and the high-level lossyness strategy (including the lossy function family construction) used in [10] and in our work are essentially the same. However, the low-level analysis techniques and final results are quite different. The main result proved in [10] is essentially the following.

Theorem 1.3 ([10]). *Let $n, q, m = n^{O(1)}$ and $r \geq n^{1/2+\epsilon} \cdot m$ be integers, for an arbitrary small constant $\epsilon > 0$. Then the* LWE *problem with parameters n, m, q and independent uniformly distributed errors in $\{-r, \ldots, r\}^m$ is at least as hard as (quantumly) solving worst-case problems on $(n/2)$-dimensional lattices to within a factor $\gamma = n^{1+\epsilon} \cdot mq/r$.*

The contribution of [10] over previous work is to prove the hardness of LWE for *uniformly distributed* errors, as opposed to errors that follow a Gaussian distribution. Notice that the magnitude of the errors used in [10] is always at least $\sqrt{n} \cdot m$, which is substantially larger (by a factor of m) than in previous results. So, [10] makes no progress towards reducing the magnitude of the errors, which is the main goal of this paper. By contrast, our work shows the hardness of LWE for errors smaller than $\sqrt{n}$ (indeed, as small as $\{0, 1\}$), provided the number of samples is sufficiently small.

Like our work, [10] requires the number of LWE samples m to be fixed in advance (because the error magnitude r depends on m), but it allows m to be an arbitrary polynomial in n. This is possible because for the large errors $r \gg \sqrt{n}$ considered in [10], the attack of [2] runs in at least exponential time. So, in principle, it may even be possible (and is an interesting open problem) to prove the hardness of LWE with

(large) uniform errors as in [10], but for an unbounded number of samples. In our work, hardness of LWE for errors smaller than $\sqrt{n}$ is proved for a much smaller number of samples m, and this is necessary in order to avoid the subexponential time attack of [2].

While the focus of our work in on LWE with small errors, we remark that our main LWE hardness result (Theorem 4.6) can also be instantiated using large polynomial errors $r = n^{O(1)}$ to obtain any (linear) number of samples $m = \Theta(n)$. In this setting, [10] provides a much better dependency between the magnitude of the errors and the number of samples (which in [10] can be an arbitrary polynomial). This is due to substantial differences in the low-level techniques employed in [10] and in our work to analyze the statistical properties of the lossy function family. For these same reasons, even for large errors, our results seem incomparable to those of [10] because we allow for a much wider class of error distributions.

2 Preliminaries

We use uppercase roman letters F, X for sets, lowercase roman for set elements $x \in X$, bold $\mathbf{x} \in X^n$ for vectors, and calligraphic letters $\mathcal{F}, \mathcal{X}, \ldots$ for probability distributions. The support of a probability distribution $\mathcal{X}$ is denoted $[\mathcal{X}]$. The uniform distribution over a finite set X is denoted $\mathcal{U}(X)$.

Two probability distributions $\mathcal{X}$ and $\mathcal{Y}$ are (t, ϵ)-indistinguishable if for all (probabilistic) algorithms $\mathcal{D}$ running in time at most t,

$$|\Pr[x \leftarrow \mathcal{X} : \mathcal{D}(x) \text{ accepts}] - \Pr[y \leftarrow \mathcal{Y} : \mathcal{D}(y) \text{ accepts}]| \leq \epsilon.$$

2.1 One-Way Functions

A function family is a probability distribution $\mathcal{F}$ over a set of functions $F \subseteq (X \to Y)$ with common domain X and range Y. Formally, function families are defined as distributions over bit strings (function descriptions) together with an evaluation algorithm, mapping each bitstring to a corresponding function, with possibly multiple descriptions associated to the same function. In this paper, for notational simplicity, we identify functions and their description, and unless stated otherwise, all statements about function families should be interpreted as referring to the corresponding probability distributions over function descriptions. For example, if we say that two function families $\mathcal{F}$ and $\mathcal{G}$ are indistinguishable, we mean that no efficient algorithm can distinguish between function descriptions selected according to either $\mathcal{F}$ or $\mathcal{G}$, where $\mathcal{F}$ and $\mathcal{G}$ are probability distributions over bitstrings that are interpreted as functions using the same evaluation algorithm.

A function family $\mathcal{F}$ is (t, ϵ) *collision resistant* if for all (probabilistic) algorithms $\mathcal{A}$ running in time at most t,

$$\Pr[f \leftarrow \mathcal{F}, (x, x') \leftarrow \mathcal{A}(f) : f(x) = f(x') \wedge x \neq x'] \leq \epsilon.$$

Let $\mathcal{X}$ be a probability distribution over the domain X of a function family $\mathcal{F}$. We recall the following standard security notions:

- $(\mathcal{F}, \mathcal{X})$ is (t, ϵ)-*one-way* if for all probabilistic algorithms $\mathcal{A}$ running in time at most t,

$$\Pr[f \leftarrow \mathcal{F}, x \leftarrow \mathcal{X} : \mathcal{A}(f, f(x)) \in f^{-1}(f(x))] \leq \epsilon.$$

- $(\mathcal{F}, \mathcal{X})$ is (t, ϵ)-*uninvertible* if for all probabilistic algorithms $\mathcal{A}$ running in time at most t,

$$\Pr[f \leftarrow \mathcal{F}, x \leftarrow \mathcal{X} : \mathcal{A}(f, f(x)) = x] \leq \epsilon.$$

- $(\mathcal{F}, \mathcal{X})$ is (t, ϵ)-*second preimage resistant* if for all probabilistic algorithms $\mathcal{A}$ running in time at most t,

$$\Pr[f \leftarrow \mathcal{F}, x \leftarrow \mathcal{X}, x' \leftarrow \mathcal{A}(f, x) : f(x) = f(x') \wedge x \neq x'] \leq \epsilon.$$

- $(\mathcal{F}, \mathcal{X})$ is (t, ϵ)-*pseudorandom* if the distributions $\{f \leftarrow \mathcal{F}, x \leftarrow \mathcal{X} : (f, f(x))\}$ and $\{f \leftarrow \mathcal{F}, y \leftarrow \mathcal{U}(Y) : (f, y)\}$ are (t, ϵ)-indistinguishable.

The above probabilities (or the absolute difference between probabilities, for indistinguishability) are called the *advantages* in breaking the corresponding security notions. It easily follows from the definition that if a function family is one-way with respect to any input distribution $\mathcal{X}$, then it is also uninvertible with respect to the same input distribution $\mathcal{X}$. Also, if a function family is collision resistant, then it is also second preimage resistant with respect to any efficiently samplable input distribution.

All security definitions are immediately adapted to the asymptotic setting, where we implicitly consider sequences of finite function families indexed by a security parameter. In this setting, for any security definition (one-wayness, collision resistance, etc.) we omit t, and simply say that a function is secure if for any t that is polynomial in the security parameter, it is (t, ϵ)-secure for some ϵ that is negligible in the security parameter. We say that a function family is *statistically* secure if it is (t, ϵ)-secure for some negligible ϵ and *arbitrary* t, i.e., it is secure even with respect to computationally unbounded adversaries.

The composition of function families is defined in the natural way. Namely, for any two function families with $[\mathcal{F}] \subseteq X \to Y$ and $[\mathcal{G}] \subseteq Y \to Z$, the composition $\mathcal{G} \circ \mathcal{F}$ is the function family that selects $f \leftarrow \mathcal{F}$ and $g \leftarrow \mathcal{G}$ independently at random, and outputs the function $(g \circ f) \colon X \to Z$.

2.2 Lossy Function Families

Lossy functions, introduced in [23], are usually defined in the context of trapdoor function families, where the functions are efficiently invertible with the help of some trapdoor information, and therefore injective (at least with high probability over the choice of the key). We give a more general definition of lossy function families that applies to non-injective functions and arbitrary input distributions, though we will be mostly interested in input distributions that are uniform over some set.

Definition 2.1. *Let $\mathcal{L}, \mathcal{F}$ be two probability distributions (with possibly different supports) over the same set of (efficiently computable) functions $F \subseteq X \to Y$, and let $\mathcal{X}$ be an efficiently sampleable distribution over the domain X. We say that $(\mathcal{L}, \mathcal{F}, \mathcal{X})$ is a* lossy function family *if the following properties are satisfied:*

- *the distributions $\mathcal{L}$ and $\mathcal{F}$ are indistinguishable,*
- *$(\mathcal{L}, \mathcal{X})$ is uninvertible, and*
- *$(\mathcal{F}, \mathcal{X})$ is second preimage resistant.*

The uninvertibility and second preimage resistance properties can be either computational or statistical. (The definition from [23] requires both to be statistical.) We remark that uninvertible functions and second preimage resistant functions are not necessarily one-way. For example, the constant function $f(x) = 0$ is (statistically) uninvertible when $|X|$ is super-polynomial in the security parameter, and the identity function $f(x) = x$ is (statistically) second preimage resistant (in fact, even collision resistant), but neither is one-way. Still, if a function family is simultaneously uninvertible and second preimage resistant, then one-wayness easily follows.

Lemma 2.2. *Let $\mathcal{F}$ be a family of functions computable in time t'. If $(\mathcal{F}, \mathcal{X})$ is both (t, ϵ)-uninvertible and $(t + t', \epsilon')$-second preimage resistant, then it is also $(t, \epsilon + \epsilon')$-one-way.*

Proof. Let $\mathcal{A}$ be an algorithm running in time at most t and attacking the one-wayness property of $(\mathcal{F}, \mathcal{X})$. Let $f \leftarrow \mathcal{F}$ and $x \leftarrow \mathcal{X}$ be chosen at random, and compute $y \leftarrow \mathcal{A}(f, f(x))$. We want to bound the probability that $f(x) = f(y)$. We consider two cases:

- If $x = y$, then $\mathcal{A}$ breaks the uninvertibility property of $(\mathcal{F}, \mathcal{X})$.
- If $x \neq y$, then $\mathcal{A}'(f, x) = \mathcal{A}(f, f(x))$ breaks the second preimage property of $(\mathcal{F}, \mathcal{X})$.

By assumption, the probability of these two events are at most ϵ and ϵ' respectively. By the union bound, $\mathcal{A}$ breaks the one-wayness property with advantage at most $\epsilon + \epsilon'$. □

It easily follows by a simple indistinguishability argument that if $(\mathcal{L}, \mathcal{F}, \mathcal{X})$ is a lossy function family, then both $(\mathcal{L}, \mathcal{X})$ and $(\mathcal{F}, \mathcal{X})$ are one-way.

Lemma 2.3. *Let $\mathcal{F}$ and $\mathcal{F}'$ be any two indistinguishable, efficiently computable function families, and let $\mathcal{X}$ be an efficiently sampleable input distribution. Then if $(\mathcal{F}, \mathcal{X})$ is uninvertible (respectively, second-preimage resistant), then $(\mathcal{F}', \mathcal{X})$ is also uninvertible (resp., second-preimage resistant). In particular, if $(\mathcal{L}, \mathcal{F}, \mathcal{X})$ is a lossy function family, then $(\mathcal{L}, \mathcal{X})$ and $(\mathcal{F}, \mathcal{X})$ are both one-way.*

Proof. Assume that $(\mathcal{F}, \mathcal{X})$ is uninvertible and that there exists an efficient algorithm $\mathcal{A}$ breaking the uninvertibility property of $(\mathcal{F}', \mathcal{X})$. Then $\mathcal{F}$ and $\mathcal{F}'$ can be efficiently distinguished by the following algorithm $\mathcal{D}(f)$: choose $x \leftarrow \mathcal{X}$, compute $x' \leftarrow \mathcal{A}(f, f(x))$, and accept if $\mathcal{A}$ succeeded, i.e., if $x = x'$.

Next, assume that $(\mathcal{F}, \mathcal{X})$ is second preimage resistant, and that there exists an efficient algorithm $\mathcal{A}$ breaking the second preimage resistance property of $(\mathcal{F}', \mathcal{X})$. Then $\mathcal{F}$ and $\mathcal{F}'$ can be efficiently distinguished by the following algorithm $\mathcal{D}(f)$: choose $x \leftarrow \mathcal{X}$, compute $x' \leftarrow \mathcal{A}(f, x)$, and accept if $\mathcal{A}$ succeeded, i.e., if $x \neq x'$ and $f(x) = f(x')$.

It follows that if $(\mathcal{L}, \mathcal{F}, \mathcal{X})$ is a lossy function family, then $(\mathcal{L}, \mathcal{X})$ and $(\mathcal{F}, \mathcal{X})$ are both uninvertible and second preimage resistant. Therefore, by Lemma 2.2, they are also one-way. □

The standard definition of (injective) lossy trapdoor functions [23], is usually stated by requiring the ratio $|f(X)|/|X|$ to be small. Our general definition can easily be related to the standard definition by specializing it to uniform input distributions. The next lemma gives an equivalent characterization of uninvertible functions when the input distribution is uniform.

Lemma 2.4. *Let $\mathcal{L}$ be a family of functions on a common domain X, and let $\mathcal{X} = \mathcal{U}(X)$ the uniform input distribution over X. Then $(\mathcal{L}, \mathcal{X})$ is ϵ-uninvertible (even statistically, with respect to computationally unbounded adversaries) for $\epsilon = \mathbb{E}_{f \leftarrow \mathcal{L}}[|f(X)|]/|X|$.*

Proof. Fix a function f, and choose a random input $x \leftarrow \mathcal{X}$. The best (computationally unbounded) attack on the uninvertibility of $(\mathcal{L}, \mathcal{X})$, given input f and $y = f(x)$, outputs an $x' \in X$ such that $f(x') = y$ and the probability of x' under $\mathcal{X}$ is maximized. Since $\mathcal{X}$ is the uniform distribution over X, the conditional distribution of x given y is uniform over $f^{-1}(y)$, and the attack succeeds with probability $1/|f^{-1}(y)|$. Each y is output by f with probability $|f^{-1}(y)|/|X|$. So, the success probability of the attack is

$$\sum_{y \in f(X)} \frac{|f^{-1}(y)|}{|X|} \cdot \frac{1}{|f^{-1}(y)|} = \frac{|f(X)|}{|X|}.$$

Taking the expectation over the choice of f, we get that the attacker succeeds with probability ϵ. □

We conclude this section with the observation that uninvertibility behaves as expected with respect to function composition.

Lemma 2.5. *If $(\mathcal{F}, \mathcal{X})$ is uninvertible and $\mathcal{G}$ is any family of efficiently computable functions, then $(\mathcal{G} \circ \mathcal{F}, \mathcal{X})$ is also uninvertible.*

Proof. Any inverter $\mathcal{A}$ for $\mathcal{G} \circ \mathcal{F}$ can be easily transformed into an inverter $\mathcal{A}'(f, y)$ for $(\mathcal{F}, \mathcal{X})$ that chooses $g \leftarrow \mathcal{G}$ at random, and outputs the result of running $\mathcal{A}(g \circ f, g(y))$ □

A similar statement holds also for one-wayness, under the additional assumption that $\mathcal{G}$ is second preimage resistant, but it is not needed here.

2.3 Lattices and Gaussians

An n-dimensional *lattice* of *rank* k is the set Λ of integer combinations of k linearly independent vectors $\mathbf{b}_1, \ldots, \mathbf{b}_k \in \mathbb{R}^n$, i.e. $\Lambda = \left\{ \sum_{i=1}^{k} x_i \mathbf{b}_i \mid x_i \in \mathbb{Z} \text{ for } i = 1, \ldots, k \right\}$. The matrix $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_k]$ is called a *basis* for the lattice Λ. The *dual* of a (not necessarily full-rank) lattice Λ is the set $\Lambda^* = \{\mathbf{x} \in \mathrm{span}(\Lambda) : \forall \mathbf{y} \in \Lambda, \langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z}\}$. In what follows, unless otherwise specified we work with *full-rank* lattices, where $k = n$.

The ith *successive minimum* $\lambda_i(\Lambda)$ is the smallest radius r such that Λ contains i linearly independent vectors of (Euclidean) length at most r. A fundamental computational problem in the study of lattice cryptography is the approximate Shortest Independent Vectors Problem SIVP_γ, which, on input a full-rank n-dimensional lattice Λ (typically represented by a basis), asks to find n linearly independent lattice vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \Lambda$ all of length at most $\gamma \cdot \lambda_n(\Lambda)$, where $\gamma \geq 1$ is an approximation factor and is usually a function of the lattice dimension n. Another problem is the (decision version of the) approximate Shortest Vector Problem GapSVP_γ, which, on input an n-dimensional lattice Λ, asks to output "yes" if $\lambda_1(\Lambda) \leq 1$ and "no" if $\lambda_1(\Lambda) > \gamma$. (If neither is the case, any answer is acceptable.)

For a matrix $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_k]$ of linearly independent vectors, the *Gram-Schmidt* orthogonalization $\widetilde{\mathbf{B}}$ is the matrix of vectors $\widetilde{\mathbf{b}}_i$ where $\widetilde{\mathbf{b}}_1 = \mathbf{b}_1$, and for each $i = 2, \ldots, k$, the vector $\widetilde{\mathbf{b}}_i$ is the projection of $\mathbf{b}_i$ orthogonal to $\mathrm{span}(\mathbf{b}_1, \ldots, \mathbf{b}_{i-1})$. The *Gram-Schmidt* minimum of a lattice Λ is $\tilde{bl}(\Lambda) = \min_{\mathbf{B}} \|\widetilde{\mathbf{B}}\|$, where $\|\widetilde{\mathbf{B}}\| = \max_i \|\widetilde{\mathbf{b}}_i\|$ and the minimum is taken over all bases $\mathbf{B}$ of Λ. Given any basis $\mathbf{D}$ of a lattice Λ and any set $\mathbf{S}$ of linearly independent vectors in Λ, it is possible to efficiently construct a basis $\mathbf{B}$ of Λ such that $\|\widetilde{\mathbf{B}}\| \leq \|\widetilde{\mathbf{S}}\|$ (see [16]).

The *Gaussian function* $\rho_s \colon \mathbb{R}^m \to \mathbb{R}$ with parameter s is defined as $\rho_s(\mathbf{x}) = \exp(-\pi \|\mathbf{x}\|^2 / s^2)$. When s is omitted, it is assumed to be 1. The *discrete Gaussian distribution* $D_{\Lambda + \mathbf{c}, s}$ with parameter s over a lattice coset $\Lambda + \mathbf{c}$ is the distribution that samples each element $\mathbf{x} \in \Lambda + \mathbf{c}$ with probability $\rho_s(\mathbf{x}) / \rho_s(\Lambda + \mathbf{c})$, where $\rho_s(\Lambda + \mathbf{c}) = \sum_{\mathbf{y} \in \Lambda + \mathbf{c}} \rho_s(\mathbf{y})$ is a normalization factor.

For any $\epsilon > 0$, the *smoothing parameter* $\eta_\epsilon(\Lambda)$ [19] is the smallest $s > 0$ such that $\rho_{1/s}(\Lambda^* \setminus \{\mathbf{0}\}) \leq \epsilon$. When ϵ is omitted, it is some unspecified negligible function $\epsilon = n^{-\omega(1)}$ of the lattice dimension or security parameter n, which may vary from place to place.

We observe that the smoothing parameter satisfies the following decomposition lemma. The general case for the sum of several lattices (whose linear spans have trivial pairwise intersections) follows immediately by induction.

Lemma 2.6. *Let lattice $\Lambda = \Lambda_1 + \Lambda_2$ be the (internal direct) sum of two lattices such that $\mathrm{span}(\Lambda_1) \cap \mathrm{span}(\Lambda_2) = \{\mathbf{0}\}$, and let $\widetilde{\Lambda}_2$ be the projection of Λ_2 orthogonal to $\mathrm{span}(\Lambda_1)$. Then for any $\epsilon_1, \epsilon_2, \epsilon > 0$ such*

that $1+\epsilon = (1+\epsilon_1)(1+\epsilon_2)$*, we have*

$$\eta_\epsilon(\widetilde{\Lambda}_2) \le \eta_\epsilon(\Lambda) \le \eta_\epsilon(\Lambda_1 + \widetilde{\Lambda}_2) \le \max\{\eta_{\epsilon_1}(\Lambda_1), \eta_{\epsilon_2}(\widetilde{\Lambda}_2)\}.$$

Proof. Let Λ^*, Λ_1^* and $\widetilde{\Lambda}_2^*$ be the dual lattices of Λ, Λ_1 and $\widetilde{\Lambda}_2$, respectively. For the first inequality, notice that $\widetilde{\Lambda}_2^*$ is a sublattice of Λ^*. Therefore, $\rho_{1/s}(\widetilde{\Lambda}_2^* \setminus \{\mathbf{0}\}) \le \rho_{1/s}(\Lambda^* \setminus \{\mathbf{0}\})$ for any $s > 0$, and thus $\eta_\epsilon(\widetilde{\Lambda}_2) \le \eta_\epsilon(\Lambda)$.

Next we prove that $\eta_\epsilon(\Lambda) \le \eta_\epsilon(\Lambda_1 + \widetilde{\Lambda}_2)$. It is routine to verify that we can express the dual lattice Λ^* as the sum $\Lambda^* = \widetilde{\Lambda}_1^* + \widetilde{\Lambda}_2^*$, where $\widetilde{\Lambda}_1$ is the projection of Λ_1 orthogonal to $\mathrm{span}(\Lambda_2)$, and $\widetilde{\Lambda}_1^*$ is its dual. Moreover, the projection of $\widetilde{\Lambda}_1^*$ orthogonal to $\mathrm{span}(\widetilde{\Lambda}_2^*)$ is exactly Λ_1^*. For any $\tilde{\mathbf{x}}_1 \in \widetilde{\Lambda}_1^*$, let $\mathbf{x}_1 \in \Lambda_1^*$ denote its projection orthogonal to $\mathrm{span}(\widetilde{\Lambda}_2^*)$. Then for any $s > 0$ we have

$$\begin{aligned}
\rho_{1/s}(\Lambda^*) &= \sum_{\widetilde{\mathbf{x}}_1 \in \widetilde{\Lambda}_1^*} \sum_{\widetilde{\mathbf{x}}_2 \in \widetilde{\Lambda}_2^*} \rho_{1/s}(\widetilde{\mathbf{x}}_1 + \widetilde{\mathbf{x}}_2) \\
&= \sum_{\widetilde{\mathbf{x}}_1 \in \widetilde{\Lambda}_1^*} \sum_{\widetilde{\mathbf{x}}_2 \in \widetilde{\Lambda}_2^*} \rho_{1/s}(\mathbf{x}_1) \cdot \rho_{1/s}((\widetilde{\mathbf{x}}_1 - \mathbf{x}_1) + \widetilde{\mathbf{x}}_2) \\
&= \sum_{\widetilde{\mathbf{x}}_1 \in \widetilde{\Lambda}_1^*} \rho_{1/s}(\mathbf{x}_1) \cdot \rho_{1/s}((\widetilde{\mathbf{x}}_1 - \mathbf{x}_1) + \widetilde{\Lambda}_2^*) \\
&\le \rho_{1/s}(\Lambda_1^*) \cdot \rho_{1/s}(\widetilde{\Lambda}_2^*) = \rho_{1/s}(\Lambda_1^* + \widetilde{\Lambda}_2^*) = \rho_{1/s}((\Lambda_1 + \widetilde{\Lambda}_2)^*),
\end{aligned}$$

where the inequality follows from the bound $\rho_{1/s}(\Lambda + \mathbf{c}) \le \rho_{1/s}(\Lambda)$ from [19, Lemma 2.9], and the last two equalities follow from the orthogonality of Λ_1^* and $\widetilde{\Lambda}_2^*$. This proves that $\eta_\epsilon(\Lambda) \le \eta_\epsilon(\Lambda_1 + \widetilde{\Lambda}_2)$.

Finally, for $s_1 = \eta_{\epsilon_1}(\Lambda_1)$, $s_2 = \eta_{\epsilon_2}(\widetilde{\Lambda}_2)$ and $s = \max\{s_1, s_2\}$, we have

$$\rho_{1/s}((\Lambda_1 + \widetilde{\Lambda}_2)^*) = \rho_{1/s}(\Lambda_1^*) \cdot \rho_{1/s}(\widetilde{\Lambda}_2^*) \le \rho_{1/s_1}(\Lambda_1^*) \cdot \rho_{1/s_2}(\widetilde{\Lambda}_2^*) = (1+\epsilon_1)(1+\epsilon_2) = 1+\epsilon.$$

Therefore, $\eta_\epsilon(\Lambda_1 + \widetilde{\Lambda}_2^*) \le s$. □

Using the decomposition lemma, one easily obtains known bounds on the smoothing parameter. For example, for any lattice basis $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n]$, applying Lemma 2.6 repeatedly to the decomposition into the rank-1 lattices defined by each of the basis vectors yields $\eta(\mathbf{B} \cdot \mathbb{Z}^n) \le \max_i \eta(\widetilde{\mathbf{b}}_i \cdot \mathbb{Z}) = \|\widetilde{\mathbf{B}}\| \cdot \omega_n$, where $\omega_n = \eta(\mathbb{Z}) = \omega(\sqrt{\log n})$ is the smoothing parameter of the integer lattice $\mathbb{Z}$. Choosing a basis $\mathbf{B}$ achieving $\tilde{bl}(\Lambda) = \min_{\mathbf{B}} \|\widetilde{\mathbf{B}}\|$ (where the minimum is taken over all bases $\mathbf{B}$ of Λ), we get the bound $\eta(\Lambda) \le \tilde{bl}(\Lambda) \cdot \omega_n$ from [12, Theorem 3.1]. Similarly, choosing a set $\mathbf{S} \subset \Lambda$ of linearly independent vectors of length $\|\mathbf{S}\| \le \lambda_n(\Lambda)$, we get the bound $\eta(\Lambda) \le \eta(\mathbf{S} \cdot \mathbb{Z}^n) \le \|\widetilde{\mathbf{S}}\| \cdot \omega_n \le \|\mathbf{S}\| \cdot \omega_n = \lambda_n(\Lambda) \cdot \omega_n$ from [19, Lemma 3.3]. In this paper we use a further generalization of these bounds, still easily obtained from the decomposition lemma.

Corollary 2.7. *The smoothing parameter of the tensor product of any two lattices* Λ_1, Λ_2 *satisfies* $\eta(\Lambda_1 \otimes \Lambda_2) \le \tilde{bl}(\Lambda_1) \cdot \eta(\Lambda_2)$.

Proof. Let $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_k]$ be a basis of Λ_1 achieving $\max_i \|\widetilde{\mathbf{b}}_i\| = \tilde{bl}(\Lambda_1)$, and consider the natural decomposition of $\Lambda_1 \otimes \Lambda_2$ into the sum

$$(\mathbf{b}_1 \otimes \Lambda_2) + \cdots + (\mathbf{b}_k \otimes \Lambda_2).$$

Notice that the projection of each sublattice $\mathbf{b}_i \otimes \Lambda_2$ orthogonal to the previous sublattices $\mathbf{b}_j \otimes \Lambda_2$ (for $j < i$) is precisely $\widetilde{\mathbf{b}}_i \otimes \Lambda_2$, and has smoothing parameter $\eta(\widetilde{\mathbf{b}}_i \otimes \Lambda_2) = \|\widetilde{\mathbf{b}}_i\| \cdot \eta(\Lambda_2)$. Therefore, by repeated application of Lemma 2.6, we have $\eta(\Lambda_1 \otimes \Lambda_2) \le \max_i \|\widetilde{\mathbf{b}}_i\| \cdot \eta(\Lambda_2) = \tilde{bl}(\Lambda_1) \cdot \eta(\Lambda_2)$. □

The following proposition relates the problem of sampling lattice vectors according to a Gaussian distribution to the SIVP.

Proposition 2.8 ([24], Lemma 3.17). *There is a polynomial time algorithm that, given a basis for an n-dimensional lattice Λ and polynomially many samples from $D_{\Lambda,\sigma}$ for some $\sigma \geq 2\eta(\Lambda)$, solves SIVP_γ on input lattice Λ (in the worst case over Λ, and with overwhelming probability over the choice of the lattice samples) for approximation factor $\gamma = \sigma\sqrt{n} \cdot \omega_n$.*

2.4 The SIS and LWE Functions

In this paper we are interested in two special families of functions, which are the fundamental building blocks of lattice cryptography. Both families are parametrized by three integers m, n and q, and a set $X \subseteq \mathbb{Z}^m$ of short vectors. Usually n serves as a security parameter and m and q are functions of n.

The *Short Integer Solution* function family $\mathrm{SIS}(m, n, q, X)$ is the set of all functions $f_{\mathbf{A}}$ indexed by $\mathbf{A} \in \mathbb{Z}_q^{n\times m}$ with domain $X \subseteq \mathbb{Z}^m$ and range $Y = \mathbb{Z}_q^n$, defined as $f_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x} \bmod q$. The *Learning With Errors* function family $\mathrm{LWE}(m, n, q, X)$ is the set of all functions $g_{\mathbf{A}}$ indexed by $\mathbf{A} \in \mathbb{Z}_q^{n\times m}$ with domain $\mathbb{Z}_q^n \times X$ and range $Y = \mathbb{Z}_q^m$, defined as $g_{\mathbf{A}}(\mathbf{s}, \mathbf{x}) = \mathbf{A}^T\mathbf{s} + \mathbf{x} \bmod q$. Both function families are endowed with the uniform distribution over $\mathbf{A} \in \mathbb{Z}_q^{n\times m}$. We omit the set X from the notation $\mathrm{SIS}(m, n, q)$ and $\mathrm{LWE}(m, n, q)$ when clear from the context, or unimportant.

In the context of collision resistance, we sometimes write $\mathrm{SIS}(m, n, q, \beta)$ for some real $\beta > 0$, without an explicit domain X. Here the collision-finding problem is, given $\mathbf{A} \in \mathbb{Z}_q^{n\times m}$, to find distinct $\mathbf{x}, \mathbf{x}' \in \mathbb{Z}^m$ such that $\|\mathbf{x} - \mathbf{x}'\| \leq \beta$ and $f_{\mathbf{A}}(\mathbf{x}) = f_{\mathbf{A}}(\mathbf{x}')$. It is easy to see that this is equivalent to finding a nonzero $\mathbf{z} \in \mathbb{Z}^m$ of length at most $\|\mathbf{z}\| \leq \beta$ such that $f_{\mathbf{A}}(\mathbf{z}) = \mathbf{0}$.

For other security properties (e.g., one-wayness, uninvertibility, etc.), the most commonly used classes of domains and input distributions $\mathcal{X}$ for SIS are the uniform distribution $\mathcal{U}(X)$ over the set $X = \{0, \ldots, s-1\}^m$ or $X = \{-s, \ldots, 0, \ldots, s\}^m$, and the discrete Gaussian distribution $D_{\mathbb{Z},s}^m$. Usually, this distribution is restricted to the set of short vectors $X = \{\mathbf{x} \in \mathbb{Z}^m \colon \|\mathbf{x}\| \leq s\sqrt{m}\}$, which carries all but a $2^{-\Omega(m)}$ fraction of the probability mass of $D_{\mathbb{Z},s}^m$.

For the LWE function family, the input is usually chosen according to distribution $\mathcal{U}(\mathbb{Z}_q^n) \times \mathcal{X}$, where $\mathcal{X}$ is one of the SIS input distributions. This makes the SIS and LWE function families essentially equivalent, as shown in the following proposition.

Proposition 2.9 ([15, 17]). *For any n, $m \geq n + \omega(\log n)$, q, and distribution $\mathcal{X}$ over $\mathbb{Z}^m$, the $\mathrm{LWE}(m, n, q)$ function family is one-way (resp. pseudorandom, or uninvertible) with respect to input distribution $\mathcal{U}(\mathbb{Z}_q^n) \times \mathcal{X}$ if and only if the $\mathrm{SIS}(m, m-n, q)$ function family is one-way (resp. pseudorandom, or uninvertible) with respect to the input distribution $\mathcal{X}$.*

In applications, the SIS function family is typically used with larger input domains X for which the functions are surjective but not injective, while the LWE function family is used with smaller domains X for which the functions are injective, but not surjective. The results in this paper are more naturally stated using the SIS function family, so we will use the SIS formulation to establish our main results, and then reformulate them in terms of the LWE function family by invoking Proposition 2.9. We also use Proposition 2.9 to reformulate known hardness results (from worst-case complexity assumptions) for LWE in terms of SIS.

Assuming the quantum worst-case hardness of standard lattice problems, Regev [24] showed that the $\mathrm{LWE}(m, n, q)$ function family is hard to invert with respect to the discrete Gaussian error distribution $D_{\mathbb{Z},\sigma}^m$ for any $\sigma > 2\sqrt{n}$. (See also [21] for a classical reduction that requires q to be exponentially large in n.

Because we are concerned with small parameters in this work, we focus mainly on the implications of the quantum reduction.)

Proposition 2.10 ([24], Theorem 3.1). *For any $m = n^{O(1)}$, integer q and real $\alpha \in (0,1)$ such that $\alpha q > 2\sqrt{n}$, there is a polynomial time quantum reduction from sampling $D_{\Lambda,\sigma}$ (for any n-dimensional lattice Λ and $\sigma > (\sqrt{2n}/\alpha)\eta(\Lambda)$) to inverting the $\mathrm{LWE}(m,n,q)$ function family on input $\mathcal{Y} = D_{\mathbb{Z}^m,\alpha q}$.*

Combining Propositions 2.8, 2.9 and 2.10, we get the following corollary.

Corollary 2.11. *For any positive m, n such that $\omega(\log n) \le m - n \le n^{O(1)}$ and $2\sqrt{n} < \sigma < q$, the $\mathrm{SIS}(m, m-n, q)$ function family is uninvertible with respect to input distribution $D^m_{\mathbb{Z},\sigma}$, under the assumption that no (quantum) algorithm can efficiently sample from a distribution statistically close to $D_{\Lambda,\sqrt{2n}q/\sigma}$.*

In particular, assuming the worst-case (quantum) hardness of $\mathrm{SIVP}_{n\omega_n q/\sigma}$ over n-dimensional lattices, the $\mathrm{SIS}(m, m-n, q)$ function family is uninvertible with respect to input distribution $D^m_{\mathbb{Z},\sigma}$.

We use the fact that LWE/SIS is not only hard to invert, but also pseudorandom. This is proved using search-to-decision reductions for those problems. The most general such reductions known to date are given in the following two theorems.

Theorem 2.12 ([17]). *For any positive m, n such that $\omega(\log n) \le m - n \le n^{O(1)}$, any positive $\sigma \le n^{O(1)}$, and any q with no divisors in the interval $((\sigma/\omega_n)^{m/k}, \sigma \cdot \omega_n)$, if $\mathrm{SIS}(m, m-n, q, D^m_{\mathbb{Z},\sigma})$ is uninvertible, then it is also pseudorandom.*

Notice that when $\sigma > \omega_n^{(m+k)/(m-k)}$, the interval $((\sigma/\omega_n)^{m/k}, \sigma \cdot \omega_n)$ is empty, and Theorem 2.12 holds without any restriction on the factorization of the modulus q.

Theorem 2.13 ([18]). *Let q have prime factorization $q = p_1^{e_1} \cdots p_k^{e_k}$ for pairwise distinct $\mathrm{poly}(n)$-bounded primes p_i with each $e_i \ge 1$, and let $0 < \alpha \le 1/\omega_n$. If $\mathrm{LWE}(m, n, q, D^m_{\mathbb{Z},\alpha q})$ is hard to invert for* all *$m(n) = n^{O(1)}$, then $\mathrm{LWE}(m', n, q, D^m_{\mathbb{Z},\alpha' q})$ is pseudorandom for any $m' = n^{O(1)}$ and*

$$\alpha' \ge \max\{\alpha, \omega_n^{1+1/\ell} \cdot \alpha^{1/\ell}, \omega_n/p_1^{e_1}, \ldots, \omega_n/p_k^{e_k}\},$$

where ℓ is an upper bound on number of prime factors $p_i < \omega_n/\alpha'$.

In this work we focus on the use of Theorem 2.12, because it guarantees pseudorandomness for the *same* value of m as for the assumed one-wayness. This feature is important for applying our results from Section 4, which guarantee one-wayness for *particular* values of m (but not necessarily all $m = n^{O(1)}$).

Corollary 2.14. *For any positive m, n, σ, q such that $\omega(\log n) \le m - n \le n^{O(1)}$ and $2\sqrt{n} < \sigma < q < n^{O(1)}$, if q has no divisors in the range $((\sigma/\omega_n)^{1+n/k}, \sigma \cdot \omega_n)$, then the $\mathrm{SIS}(m, m-n, q)$ function family is pseudorandom with respect to input distribution $D^m_{\mathbb{Z},\sigma}$, under the assumption that no (quantum) algorithm can efficiently sample (up to negligible statistical errors) $D_{\Lambda,\sqrt{2n}q/\sigma}$.*

In particular, assuming the worst-case (quantum) hardness of $\mathrm{SIVP}_{n\omega_n q/\sigma}$ on n-dimensional lattices, the $\mathrm{SIS}(m, m-n, q)$ function family is pseudorandom with respect to input distribution $D^m_{\mathbb{Z},\sigma}$.

3 Hardness of SIS with Small Modulus

We first prove a simple "success amplification" lemma for collision-finding in SIS, which says that any inverse-polynomial advantage can be amplified to essentially 1, at only the expense of a larger runtime and value of m (which will have no ill effects on our final results). Therefore, for the remainder of this section we implicitly restrict our attention to collision-finding algorithms that have overwhelming advantage.

Lemma 3.1. *For arbitrary n, q, m and $X \subseteq \mathbb{Z}^m$, suppose there exists a probabilistic algorithm $\mathcal{A}$ that has advantage $\epsilon > 0$ in collision-finding for $\mathrm{SIS}(m, n, q, X)$. Then there exists a probabilistic algorithm $\mathcal{B}$ that has advantage $1-(1-\epsilon)^t \geq 1-\exp(-\epsilon t) = 1-\exp(-n)$ in collision-finding for $\mathrm{SIS}(M = t \cdot m, n, q, X')$, where $t = n/\epsilon$ and $X' = \bigcup_{i=1}^{t}(\{0^m\}^{i-1} \times X \times \{0^m\}^{t-i})$. The runtime of $\mathcal{B}$ is essentially t times that of $\mathcal{A}$.*

Proof. The algorithm $\mathcal{B}$ simply partitions its input $\mathbf{A} \in \mathbb{Z}_q^{n \times M}$ into blocks $\mathbf{A}_i \in \mathbb{Z}_q^{n \times m}$ and invokes $\mathcal{A}$ (with fresh random coins) on each of them, until $\mathcal{A}$ returns a valid collision $\mathbf{x}, \mathbf{x}' \in X$ for some $\mathbf{A}_i$. Then $\mathcal{B}$ returns

$$(0^{m(i-1)}, \mathbf{x}, 0^{m(t-i)}), (0^{m(i-1)}, \mathbf{x}', 0^{m(t-i)}) \in X'$$

as a collision for $\mathbf{A}$. Clearly, $\mathcal{B}$ succeeds if any call to $\mathcal{A}$ succeeds. Since all t calls to $\mathcal{A}$ are on independent inputs $\mathbf{A}_i$ and use independent coins, some call will succeed, except with $(1-\epsilon)^t$ probability. □

3.1 SIS-to-SIS Reduction

Our first proof that the $\mathrm{SIS}(m, n, q, \beta)$ function family is collision resistant for moduli q as small as $n^{1/2+\delta}$ proceeds by a reduction between SIS problems with different parameters. Previous hardness results based on worst-case lattice assumptions require the modulus q to be at least $\beta \cdot \omega(\sqrt{n \log n})$ [12, Theorem 9.2], and $\beta \geq \sqrt{n \log q}$ is needed to guarantee that a nontrivial solution exists. For such parameters, SIS is collision resistant assuming the hardness of approximating worst-case lattice problems to within $\approx \beta\sqrt{n}$ factors.

The intuition behind our proof for smaller moduli is easily explained. We reduce SIS with modulus q^c and solution bound β^c (for any constant integer $c \geq 1$) to SIS with modulus q and bound β. Then as long as $(q/\beta)^c \geq \omega(\sqrt{n \log n})$, the former problem enjoys worst-case hardness, hence so does the latter. Thus we can take $q = \beta \cdot n^\delta$ for any constant $\delta > 0$, and $c > 1/(2\delta)$. Notice, however, that the underlying approximation factor for worst-case lattice problems is $\approx \beta^c\sqrt{n} \geq n^{1/2+1/(4\delta)}$, which, while still polynomial, degrades severely as δ approaches 0. In the next subsection we give a direct reduction from worst-case lattice problems to SIS with a small modulus, which does not have this drawback.

The above discussion is formalized in the following proposition. For technical reasons, we prove that $\mathrm{SIS}(m, n, q, X)$ is collision resistant assuming that the domain X has the property that all SIS solutions $\mathbf{z} \in (X - X) \setminus \{\mathbf{0}\}$ satisfy $\gcd(\mathbf{z}, q) = 1$. This restriction is satisfied in many (but not all) common settings, e.g., when $q > \beta$ is prime, or when $X \subseteq \{0, 1\}^m$ is a set of binary vectors.

Proposition 3.2. *Let n, q, m, β and $X \subseteq \mathbb{Z}^m$ be such that $\gcd(\mathbf{x} - \mathbf{x}', q) = 1$ and $\|\mathbf{x} - \mathbf{x}'\| \leq \beta$ for any distinct $\mathbf{x}, \mathbf{x}' \in X$. For any positive integer c, there is a deterministic reduction from collision-finding for $\mathrm{SIS}(m^c, n, q^c, \beta^c)$ to collision-finding for $\mathrm{SIS}(m, n, q, X)$ (in both cases, with overwhelming advantage). The reduction runs in time polynomial in its input size, and makes fewer than m^c calls to its oracle.*

Proof. Let $\mathcal{A}$ be an efficient algorithm that finds a collision for $\mathrm{SIS}(m, n, q, X)$ with overwhelming advantage. We use it to find a nonzero solution for $\mathrm{SIS}(m^c, n, q^c, \beta^c)$. Let $\mathbf{A} \in \mathbb{Z}_{q^c}^{n \times m^c}$ be an input SIS instance. Partition the columns of $\mathbf{A}$ into m^{c-1} blocks $\mathbf{A}_i \in \mathbb{Z}_{q^c}^{n \times m}$, and for each one, invoke $\mathcal{A}$ to find a collision modulo q, i.e., a pair of distinct vectors $\mathbf{x}_i, \mathbf{x}_i' \in X$ such that $\mathbf{A}_i\mathbf{z}_i = \mathbf{0} \bmod q$, where $\mathbf{z}_i = \mathbf{x}_i - \mathbf{x}_i'$ and $\|\mathbf{z}_i\| \leq \beta$.

For each i, since $\gcd(\mathbf{z}_i, q) = 1$ and $\mathbf{A}_i\mathbf{z}_i = \mathbf{0} \bmod q$, the vector $\mathbf{a}'_i = (\mathbf{A}_i\mathbf{z}_i)/q \in \mathbb{Z}_{q^{c-1}}^n$ is uniformly random, even after conditioning on $\mathbf{z}_i$ and $\mathbf{A}_i \bmod q$. So, the matrix $\mathbf{A}' \in \mathbb{Z}_{q^{c-1}}^{n \times m^{c-1}}$ made up of all these columns is uniformly random. By induction on c, using $\mathcal{A}$ we can find a nonzero solution $\mathbf{z}' \in \mathbb{Z}^{m^{c-1}}$ such that $\mathbf{A}'\mathbf{z}' = \mathbf{0} \bmod q^{c-1}$ and $\|\mathbf{z}'\| \leq \beta^{c-1}$. Then it is easy to verify that a nonzero solution for the original instance $\mathbf{A}$ is given by $\mathbf{z} = (z'_1 \cdot \mathbf{z}_1, \ldots, z'_{m^{c-1}} \cdot \mathbf{z}_{m^{c-1}}) \in \mathbb{Z}^{m^c}$, and that $\|\mathbf{z}\| \leq \|\mathbf{z}'\| \cdot \max_i \|\mathbf{z}_i\| \leq \beta^c$. Finally, the total number of calls to $\mathcal{A}$ is $\sum_{i=0}^{c-1} m^i < m^c$, as claimed. □

3.2 Direct Reduction

As mentioned above, the large worst-case approximation factor associated with the use of Proposition 3.2 is undesirable, as is (to a lesser extent) the restriction that $\gcd(X - X, q) = 1$. To eliminate these drawbacks, we next give a direct proof that SIS is collision resistant for small q, based on the assumed hardness of worst-case lattice problems. The underlying approximation factor for these problems can be as small as $\tilde{O}(\beta\sqrt{n})$, which matches the best known factors obtained by previous proofs (which require a larger modulus q). Our new proof combines ideas from [19, 12] and Proposition 3.2, as well as a new convolution theorem for discrete Gaussians which strengthens similar ones previously proved in [22, 6].

Our proof of the convolution theorem is substantially different and, we believe, technically simpler than the prior ones. In particular, it handles the sum of many Gaussian samples all at once, whereas previous proofs used induction from a base case of two samples. With the inductive approach, it is technically complex to verify that all the intermediate Gaussian parameters (which involve harmonic means) satisfy the hypotheses. Moreover, the intermediate parameters can depend on the order in which the samples are added in the induction, leading to unnecessarily strong hypotheses on the original parameters.

Theorem 3.3. *Let Λ be an n-dimensional lattice, $\mathbf{z} \in \mathbb{Z}^m$ a nonzero integer vector, $s_i \geq \sqrt{2}\|\mathbf{z}\|_\infty \cdot \eta(\Lambda)$, and $\Lambda + \mathbf{c}_i$ arbitrary cosets of Λ for $i = 1, \ldots, m$. Let $\mathbf{y}_i$ be independent vectors with distributions $D_{\Lambda+\mathbf{c}_i, s_i}$, respectively. Then the distribution of $\mathbf{y} = \sum_i z_i\mathbf{y}_i$ is statistically close to $D_{Y,s}$, where $Y = \gcd(\mathbf{z})\Lambda + \mathbf{c}$, $\mathbf{c} = \sum_i z_i\mathbf{c}_i$, and $s = \sqrt{\sum_i (z_i s_i)^2}$.*

In particular, if $\gcd(\mathbf{z}) = 1$ and $\sum_i z_i\mathbf{c}_i \in \Lambda$, then $\mathbf{y}$ is distributed statistically close to $D_{\Lambda,s}$.

Proof. First we verify that the support of $\mathbf{y}$ is

$$\sum_i z_i(\Lambda + \mathbf{c}_i) = \sum_i z_i\Lambda + \sum_i z_i \cdot \mathbf{c}_i = \gcd(\mathbf{z})\Lambda + \sum_i z_i \cdot \mathbf{c}_i = Y.$$

So it remains to prove that each $\mathbf{y} \in Y$ has probability (nearly) proportional to $\rho_s(\mathbf{y})$.

For the remainder of the proof we use the following convenient scaling. Define the diagonal matrices $\mathbf{S} = \mathrm{diag}(s_1, \ldots, s_m)$ and $\mathbf{S}' = \mathbf{S} \otimes \mathbf{I}_n$, and the mn-dimensional lattice $\Lambda' = \bigoplus_i (s_i^{-1}\Lambda) = (\mathbf{S}')^{-1} \cdot \Lambda^{\oplus m}$, where $\bigoplus$ denotes the (external) direct sum of lattices and $\Lambda^{\oplus m} = \mathbb{Z}^m \otimes \Lambda$ is the direct sum of m copies of Λ. Then by independence of the $\mathbf{y}_i$, it can be seen that $\mathbf{y}' = (\mathbf{S}')^{-1} \cdot (\mathbf{y}_1, \ldots, \mathbf{y}_m)$ has discrete Gaussian distribution $D_{\Lambda'+\mathbf{c}'}$ (with parameter 1), where $\mathbf{c}' = (\mathbf{S}')^{-1} \cdot (\mathbf{c}_1, \ldots, \mathbf{c}_m)$.

The output vector $\mathbf{y} = \sum_i z_i\mathbf{y}_i$ can be expressed, using the mixed-product property for Kronecker products, as

$$\mathbf{y} = (\mathbf{z}^T \otimes \mathbf{I}_n) \cdot (\mathbf{y}_1, \ldots, \mathbf{y}_m) = (\mathbf{z}^T \otimes \mathbf{I}_n) \cdot \mathbf{S}' \cdot \mathbf{y}' = ((\mathbf{z}^T\mathbf{S}) \otimes \mathbf{I}_n) \cdot \mathbf{y}'.$$

So, letting $\mathbf{Z} = ((\mathbf{z}^T\mathbf{S}) \otimes \mathbf{I}_n)$, we want to prove that the distribution of $\mathbf{y} \sim \mathbf{Z} \cdot D_{\Lambda'+\mathbf{c}'}$ is statistically close to $D_{Y,s}$.

Fix any vectors $\mathbf{x}' \in \Lambda' + \mathbf{c}'$ and $\bar{\mathbf{y}} = \mathbf{Z}\mathbf{x}' \in Y$, and define the proper sublattice

$$L = \{\mathbf{v} \in \Lambda' \colon \mathbf{Z}\mathbf{v} = \mathbf{0}\} = \Lambda' \cap \ker(\mathbf{Z}) \subsetneq \Lambda'.$$

It is immediate to verify that the set of all $\mathbf{y}' \in \Lambda' + \mathbf{c}'$ such that $\mathbf{Z}\mathbf{y}' = \bar{\mathbf{y}}$ is $(\Lambda' + \mathbf{c}') \cap \ker(\mathbf{Z}) = L + \mathbf{x}'$. Let $\mathbf{x}$ be orthogonal projection of $\mathbf{x}'$ onto $\ker(\mathbf{Z}) \supset L$. Then we have

$$\Pr[\mathbf{y} = \bar{\mathbf{y}}] = \frac{\rho(L + \mathbf{x}')}{\rho(\Lambda' + \mathbf{c}')} = \rho(\mathbf{x}' - \mathbf{x}) \cdot \frac{\rho(L + \mathbf{x})}{\rho(\Lambda' + \mathbf{c}')}.$$

Below we show that $\eta(L) \le 1$, which implies that $\rho(L + \mathbf{x})$ is essentially the same for all values of $\mathbf{x}'$, and hence for all $\bar{\mathbf{y}}$. Therefore, we just need to analyze $\rho(\mathbf{x}' - \mathbf{x})$.

Since $\mathbf{Z}^T$ is an orthogonal basis for $\ker(\mathbf{Z})^\perp$, each of whose columns has Euclidean norm $s = (\sum_i (z_i s_i)^2)^{1/2}$, we have $\mathbf{x}' - \mathbf{x} = (\mathbf{Z}^T\mathbf{Z}\mathbf{x}')/s^2$, and

$$\|\mathbf{x}' - \mathbf{x}\|^2 = \langle \mathbf{x}', \mathbf{Z}^T\mathbf{Z}\mathbf{x}' \rangle / s^2 = \|\mathbf{Z}\mathbf{x}'\|^2/s^2 = (\|\bar{\mathbf{y}}\|/s)^2.$$

Therefore, $\rho(\mathbf{x}' - \mathbf{x}) = \rho_s(\bar{\mathbf{y}})$, and so $\Pr[\mathbf{y} = \bar{\mathbf{y}}]$ is essentially proportional to $\rho_s(\bar{\mathbf{y}})$, i.e., the statistical distance between $\mathbf{y}$ and $D_{Y,s}$ is negligible.

It remains to bound the smoothing parameter of L. Consider the m-dimensional integer lattice $Z = \mathbb{Z}^m \cap \ker(\mathbf{z}^T) = \{\mathbf{v} \in \mathbb{Z}^m : \langle \mathbf{z}, \mathbf{v} \rangle = 0\}$. Because $(Z \otimes \Lambda) \subseteq (\mathbb{Z}^m \otimes \Lambda)$ and $\mathbf{S}^{-1}Z \subset \ker(\mathbf{z}^T\mathbf{S})$, it is straightforward to verify from the definitions that

$$(\mathbf{S}')^{-1} \cdot (Z \otimes \Lambda) = ((\mathbf{S}^{-1}Z) \otimes \Lambda)$$

is a sublattice of L. It follows from Corollary 2.7 and by scaling that

$$\eta(L) \le \eta((\mathbf{S}')^{-1} \cdot (Z \otimes \Lambda)) \le \eta(\Lambda) \cdot \tilde{bl}(Z)/\min s_i.$$

Finally, $\tilde{bl}(Z) \le \min\{\|\mathbf{z}\|, \sqrt{2}\|\mathbf{z}\|_\infty\}$ because Z has a full-rank set of vectors $z_i \cdot \mathbf{e}_j - z_j \cdot \mathbf{e}_i$, where index i minimizes $|z_i| \ne 0$, and j ranges over $\{1, \ldots, m\} \setminus \{i\}$. By assumption on the s_i, we have $\eta(L) \le 1$ as desired, and the proof is complete. □

Remark 3.4. Although we will not need it in this work, we note that the statement and proof of Theorem 3.3 can be adapted to the case where the $\mathbf{y}_i$ respectively have *non-spherical* discrete Gaussian distributions $D_{\Lambda_i + \mathbf{c}_i, \sqrt{\Sigma_i}}$ with positive definite "covariance" parameters $\Sigma_i \in \mathbb{R}^{n \times n}$, over cosets of possibly different lattices Λ_i. (See [22] for a formal definition of these distributions.)

In this setting, by scaling Λ_i and Σ_i we can assume without loss of generality that $\mathbf{z} = (1, 1, \ldots, 1)$. The theorem statement says that $\mathbf{y}$'s distribution is close to a discrete Gaussian (over an appropriate lattice coset) with covariance parameter $\Sigma = \sum \Sigma_i$, under mild assumptions on $\sqrt{\Sigma_i}$. In the proof we simply let $\mathbf{S}'$ be the block-diagonal matrix with the $\sqrt{\Sigma_i}$ as its diagonal blocks, let $\Lambda' = (\mathbf{S}')^{-1} \cdot \bigoplus_i \Lambda_i$, and let $\mathbf{Z} = (\mathbf{z}^T \otimes \mathbf{I}_n) \cdot \mathbf{S}' = [\sqrt{\Sigma_1} \mid \cdots \mid \sqrt{\Sigma_m}]$. Then the only technical difference is in bounding the smoothing parameter of L.

The convolution theorem implies the following simple but useful lemma, which shows how to convert samples having a broad range of parameters into ones having parameters in a desired narrow range.

Lemma 3.5. *There is an efficient algorithm which, given a basis $\mathbf{B}$ of some lattice Λ, some $R \ge \sqrt{2}$ and samples $(\mathbf{y}_i, s_i)$ where each $s_i \in [\sqrt{2}, R] \cdot \eta(\Lambda)$ and each $\mathbf{y}_i$ has distribution D_{Λ, s_i}, with overwhelming probability outputs a sample $(\mathbf{y}, s)$ where $s \in [R, \sqrt{2}R] \cdot \eta(\Lambda)$ and $\mathbf{y}$ has distribution statistically close to $D_{\Lambda, s}$.*

Proof. Let $\omega_n = \omega(\sqrt{\log n})$ satisfy $\omega_n \leq \sqrt{n}$. The algorithm draws $2n^2$ input samples, and works as follows: if at least n^2 of the samples have parameters $s_i \leq R \cdot \eta(\Lambda)/(\sqrt{n} \cdot \omega_n)$, then with overwhelming probability they all have lengths bounded by $R \cdot \eta(\Lambda)/\omega_n$ and they include n linearly independent vectors. Using such vectors we can construct a basis $\mathbf{S}$ such that $\|\widetilde{\mathbf{S}}\| \leq R \cdot \eta(\Lambda)/\omega_n$, and with the sampling algorithm of [12, Theorem 4.1] we can generate samples having parameter $R \cdot \eta(\Lambda)$.

Otherwise, at least n^2 of the samples $(\mathbf{y}_i, s_i)$ have parameters $s_i \geq \max\{R/n, \sqrt{2}\} \cdot \eta(\Lambda)$. Then by summing an appropriate subset of those $\mathbf{y}_i$, by the convolution theorem we can obtain a sample having parameter in the desired range. □

The next lemma is the heart of our reduction. The novel part, corresponding to the properties described in the second item, is a way of using a collision-finding oracle to reduce the Gaussian width of samples drawn from a lattice. The first item corresponds to the guarantees provided by previous reductions.

Lemma 3.6. *Let m, n be integers, $S = \{\mathbf{z} \in \mathbb{Z}^m \setminus \{\mathbf{0}\} \mid \|\mathbf{z}\| \leq \beta \wedge \|\mathbf{z}\|_\infty \leq \beta_\infty\}$ for some real $\beta \geq \beta > 0$, and q an integer modulus with at most $\mathrm{poly}(n)$ integer divisors less than β_∞. There is a probabilistic polynomial time reduction that, on input any basis $\mathbf{B}$ of a lattice Λ and sufficiently many samples $(\mathbf{y}_i, s_i)$ where $s_i \geq \sqrt{2}q \cdot \eta(\Lambda)$ and $\mathbf{y}_i$ has distribution D_{Λ,s_i}, and given access to an $\mathrm{SIS}(m, n, q, S)$ oracle (that finds collisions $\mathbf{z} \in S$ with nonnegligible probability) outputs (with overwhelming probability) a sample $(\mathbf{y}, s)$ with $\min s_i/q \leq s \leq (\beta/q) \cdot \max s_i$, and $\mathbf{y} \in \Lambda$ such that:*

- $\mathbb{E}[\|\mathbf{y}\|] \leq (\beta\sqrt{n}/q) \cdot \max s_i$, *and for any subspace $H \subset \mathbb{R}^n$ of dimension at most $n - 1$, with probability at least $1/10$ we have $\mathbf{y} \notin H$.*
- *Moreover, if each $s_i \geq \sqrt{2}\beta_\infty q \cdot \eta(\Lambda)$, then the distribution of $\mathbf{y}$ is statistically close to $D_{\Lambda,s}$*

Proof. Let $\mathcal{A}$ be the collision-finding oracle. Without loss of generality, we can assume that whenever $\mathcal{A}$ outputs a valid collision $\mathbf{z} \in S$, we have that $\gcd(\mathbf{z})$ divides q. This is so because for any integer vector $\mathbf{z}$, if $\mathbf{Az} = \mathbf{0} \bmod q$ then also $\mathbf{A}((g/d)\mathbf{z}) = \mathbf{0} \bmod q$, where $d = \gcd(\mathbf{z})$ and $g = \gcd(d, q)$. Moreover, $(g/d)\mathbf{z} \in S$ holds true and $\gcd((g/d)\mathbf{z}) = \gcd(\mathbf{z}, q)$ divides q. Let d be such that $\mathcal{A}$ outputs, with non-negligible probability, a valid collision $\mathbf{z}$ satisfying $\gcd(\mathbf{z}) = d$. Such a d exists because $\gcd(\mathbf{z})$ is bounded by β_∞ and divides q, so by assumption there are only polynomially many possible values of d. Let $q' = q/d$, which is an integer. By increasing m and using standard amplification techniques, we can make the probability that $\mathcal{A}$ outputs such a collision (satisfying $\mathbf{z} \in S$, $\mathbf{Az} = \mathbf{0} \pmod q$ and $\gcd(\mathbf{z}) = d$) exponentially close to 1.

Let $(\mathbf{y}_i, s_i)$ for $i = 1, \ldots, m$ be input samples, where $\mathbf{y}_i$ has distribution D_{Λ,s_i}. Write each $\mathbf{y}_i$ as $\mathbf{y}_i = \mathbf{Ba}_i \bmod q'\Lambda$ for $\mathbf{a}_i \in \mathbb{Z}_{q'}^n$. Since $s_i \geq q'\eta(\Lambda)$ the distribution of $\mathbf{a}_i$ is statistically close to uniform over $\mathbb{Z}_{q'}^n$. Let $\mathbf{A} = [\mathbf{a}_1 \mid \cdots \mid \mathbf{a}_m] \in \mathbb{Z}_q^{n \times m}$, and choose $\mathbf{A'} \in \mathbb{Z}_d^{n \times m}$ uniformly at random. Since $\mathbf{A}$ is statistically close to uniform over $\mathbb{Z}_{q'}^{n \times m}$, the matrix $\mathbf{A} + q'\mathbf{A'}$ is statistically close to uniform over $\mathbb{Z}_q^{n \times m}$. Call the oracle $\mathcal{A}$ on input $\mathbf{A} + q'\mathbf{A'}$, and obtain (with overwhelming probability) a nonzero $\mathbf{z} \in S$ with $\gcd(\mathbf{z}) = d$, $\|\mathbf{z}\| \leq \beta$, $\|\mathbf{z}\|_\infty \leq \beta_\infty$ and $(\mathbf{A} + q'\mathbf{A'})\mathbf{z} = \mathbf{0} \bmod q$. Notice that $q'\mathbf{A'z} = q\mathbf{A'}(\mathbf{z}/d) = \mathbf{0} \bmod q$ because $(\mathbf{z}/d)$ is an integer vector. Therefore $\mathbf{Az} = \mathbf{0} \bmod q$. Finally, the reduction outputs $(\mathbf{y}, s)$, where $\mathbf{y} = \sum_i z_i\mathbf{y}_i/q$ and $s = \sqrt{\sum_i (s_i z_i)^2}/q$. Notice that $z_i\mathbf{y}_i \in q\Lambda + \mathbf{B}(z_i\mathbf{a}_i)$ because $\gcd(\mathbf{z}) = d$, so $\mathbf{y} \in \Lambda$.

Notice that s satisfies the stated bounds because $\mathbf{z}$ is a nonzero integer vector. We next analyze the distribution of $\mathbf{y}$. For any fixed $\mathbf{a}_i$, the conditional distribution of each $\mathbf{y}_i$ is $D_{q'\Lambda + \mathbf{Ba}_i, s_i}$, where $s_i \geq \sqrt{2}\eta(q'\Lambda)$. The claim on $\mathbb{E}[\|\mathbf{y}\|]$ then follows from [19, Lemma 2.11 and Lemma 4.3] and Hölder's inequality. The claim on the probability that $\mathbf{y} \notin H$ was initially shown in the preliminary version of [19]; see also [24, Lemma 3.15].

Now assume that $s_i \geq \sqrt{2}\beta_\infty q \cdot \eta(\Lambda) \geq \sqrt{2}\|\mathbf{z}\|_\infty \cdot \eta(q'\Lambda)$ for all i. By Theorem 3.3 the distribution of $\mathbf{y}$ is statistically close to $D_{Y/q,s}$ where $Y = \gcd(\mathbf{z}) \cdot q'\Lambda + \mathbf{B}(\mathbf{Az})$. Using $\mathbf{Az} = \mathbf{0} \bmod q$ and $\gcd(\mathbf{z}) = d$, we get $Y = q\Lambda$. Therefore $\mathbf{y}$ has distribution statistically close to $D_{\Lambda,s}$, as claimed. □

Building on Lemma 3.6, our next lemma shows that for any $q \geq \beta \cdot n^{\Omega(1)}$, a collision-finding oracle can be used to obtain Gaussian samples of width close to $2\beta\beta_\infty \cdot \eta(\Lambda)$.

Lemma 3.7. *Let m, n, q, S as in Lemma 3.6, and also assume $q/\beta \geq n^\delta$ for some constant $\delta > 0$. There is an efficient reduction that, on input any basis $\mathbf{B}$ of an n-dimensional lattice Λ, an upper bound $\eta \geq \eta(\Lambda)$, and given access to an $\mathrm{SIS}(m, n, q, S)$ oracle (finding collisions $\mathbf{z} \in S$ with nonnegligible probability), outputs (with overwhelming probability) a sample $(\mathbf{y}, s)$ where $\sqrt{2}\beta_\infty \cdot \eta \leq s \leq 2\beta_\infty\beta \cdot \eta$ and $\mathbf{y}$ has distribution statistically close to $D_{\Lambda,s}$.*

Proof. By applying the LLL basis reduction algorithm [13] to the basis $\mathbf{B}$, we can assume without loss of generality that $\|\widetilde{\mathbf{B}}\| \leq 2^n \cdot \eta(\Lambda)$. Let ω_n be an arbitrary function in n satisfying $\omega_n = \omega(\sqrt{\log n})$ and $\omega_n \leq \sqrt{n}/2$.

The main procedure, described below, produces samples having parameters in the range $[1, q] \cdot \sqrt{2}\beta_\infty \cdot \eta$. On these samples we run the procedure from Lemma 3.5 (with $R = \sqrt{2}\beta_\infty q \cdot \eta$) to obtain samples having parameters in the range $[\sqrt{2}, 2] \cdot \beta_\infty q \cdot \eta$. Finally, we invoke the reduction from Lemma 3.6 on those samples to obtain a sample satisfying the conditions in the Lemma statement.

The main procedure works in a sequence of phases $i = 0, 1, 2, \ldots$. In phase i, the input is a basis $\mathbf{B}_i$ of Λ, where initially $\mathbf{B}_0 = \mathbf{B}$. The basis $\mathbf{B}_i$ is used in the discrete Gaussian sampling algorithm of [12, Theorem 4.1] to produce samples $(\mathbf{y}, s_i)$, where $s_i = \max\{\|\widetilde{\mathbf{B}}_i\| \cdot \omega_n, \sqrt{2}\beta_\infty\eta\} \geq \sqrt{2}\beta_\infty\eta$ and $\mathbf{y}_i$ has distribution statistically close to D_{Λ,s_i}. Phase i either manages to produce a sample $(\mathbf{y}, s)$ with s in the desired range $[1, q] \cdot \sqrt{2}\beta_\infty\eta$, or it produces a new basis $\mathbf{B}_{i+1}$ for which $\|\widetilde{\mathbf{B}}_{i+1}\| \leq \|\widetilde{\mathbf{B}}_i\|/2$, which is the input to the next phase. The number of phases before termination is clearly polynomial in n, by hypothesis on $\mathbf{B}$.

If $\|\widetilde{\mathbf{B}}_i\| \cdot \omega_n \leq \sqrt{2}q\beta_\infty\eta$, then this already gives samples with $s_i \in [1, q]\sqrt{2}\beta_\infty\eta$ in the desired range, and we can terminate the main phase. So, we may assume that $s_i = \|\widetilde{\mathbf{B}}_i\| \cdot \omega_n \geq \sqrt{2}q\beta_\infty\eta$. Each phase i proceeds in some constant $c \geq 1/\delta$ number of sub-phases $j = 1, 2, \ldots, c$, where the inputs to the first sub-phase are the samples $(\mathbf{y}, s_i)$ generated as described above. We recall that these samples satisfy $s_i \geq \sqrt{2}q\beta_\infty\eta$. The same will be true for the samples passed as input to all other subsequent subphases. So, each subphase receives as input samples $(\mathbf{y}, s)$ satisfying all the hypotheses of Lemma 3.6, and we can run the reduction from that lemma to generate new samples $(\mathbf{y}', s')$ having parameters s' bounded from above by $s_i \cdot (\beta/q)^j$, and from below by $\sqrt{2}\beta_\infty\eta$. If any of the produces samples satisfies $s' \leq q\sqrt{2}\beta_\infty\eta$, then we can terminate the main procedure with $(\mathbf{y}', s')$ as output. Otherwise, all samples produced during the subphase satisfy $s' > q\sqrt{2}\beta_\infty\eta$, and they can be passed as input to the next sub-phase. Notice that the total runtime of all the sub-phases is $\mathrm{poly}(n)^c$, because each invocation of the reduction from Lemma 3.6 relies on $\mathrm{poly}(n)$ invocations of the reduction in the previous sub-phase; this is why we need to limit the number of sub-phases to a constant c.

If phase i ends up running all its sub-phases without ever finding a sample with $s' \in [1, q]\sqrt{2}\beta_\infty\eta$, then it has produced samples whose parameters are bounded by $(\beta/q)^c \leq s_i \leq s_i/\sqrt{n}$. It uses n^2 of these samples, which with overwhelming probability have lengths all bounded by $s_i/\sqrt{n}$, and include n linearly independent vectors. It transforms those vectors into a basis $\mathbf{B}_{i+1}$ with $\|\widetilde{\mathbf{B}}_{i+1}\| \leq s_i/\sqrt{n} \leq \|\widetilde{B}\|_i\omega_n/\sqrt{n} \leq \|\widetilde{\mathbf{B}}_i\|/2$, as input to the next phase. □

We can now prove our main theorem, reducing worst-case lattice problems with $\max\{1, \beta\beta_\infty/q\} \cdot \tilde{O}(\beta\sqrt{n})$ approximation factors to SIS, when $q \geq \beta \cdot n^{\Omega(1)}$.

Theorem 3.8. *Let m, n be integers, $S = \{\mathbf{z} \in \mathbb{Z}^m \setminus \{\mathbf{0}\} \mid \|\mathbf{z}\| \leq \beta \wedge \|\mathbf{z}\|_\infty \leq \beta_\infty\}$ for some real $\beta \geq \beta_\infty > 0$, and $q \geq \beta \cdot n^{\Omega(1)}$ be an integer modulus with at most $\mathrm{poly}(n)$ integer divisors less than β_∞. For some $\gamma = \max\{1, \beta\beta_\infty/q\} \cdot O(\beta\sqrt{n})$, there is an efficient reduction from $\mathrm{SIVP}_\gamma^\eta$ (and hence also from standard $\mathrm{SIVP}_{\gamma \cdot \omega_n}$) on n-dimensional lattices to S-collision finding for $\mathrm{SIS}(m, n, q)$ with non-negligible advantage.*

Proof. Given an input basis $\mathbf{B}$ of a lattice Λ, we can apply the LLL algorithm to obtain a 2^n-approximation to $\eta(\Lambda)$, and by scaling we can assume that $\eta(\Lambda) \in [1, 2^n]$. For $i = 1, \ldots, n$, we run the procedure described below for each hypothesized upper bound $\eta_i = 2^i$ on $\eta(\Lambda)$. Each call to the procedure either fails, or returns a set of linearly independent vectors in Λ whose lengths are all bounded by $(\gamma/2) \cdot \eta_i$. We return the first such obtained set (i.e., for the minimal value of i). As we show below, as long as $\eta_i \geq \eta(\Lambda)$ the procedure returns a set of vectors with overwhelming probability. Since one $\eta_i \in [1, 2) \cdot \eta(\Lambda)$, our reduction solves $\mathrm{SIVP}_\gamma^\eta$ with overwhelming probability, as claimed.

The procedure invokes the reduction from Lemma 3.7 with $\eta = \eta_i$ to obtain samples with parameters in the range $[\sqrt{2}\beta_\infty, \sqrt{2}\beta\beta_\infty] \cdot \eta$. On these samples we run the procedure from Lemma 3.5 with $R = \max\{\sqrt{2}q, \sqrt{2}\beta\beta_\infty\}$ to obtain samples having parameters in the range $[R, \sqrt{2}R] \cdot \eta$. On such samples we repeatedly run (using independent samples each time) the reduction from Lemma 3.6. After enough runs, we obtain with overwhelming probability a set of linearly independent lattice vectors all having lengths at most $(\gamma/2) \cdot \eta$, as required. □

4 Hardness of LWE with Small Uniform Errors

In this section we prove the hardness of inverting the LWE function even when the error vectors have very small entries, provided the number of samples is sufficiently small. We proceed similarly to [23, 4], by using the LWE assumption (for discrete Gaussian error) to construct a lossy family of functions with respect to a uniform distribution over small inputs. However, the parameterization we obtain is different from those in [23, 4], allowing us to obtain *pseudorandomness* of LWE under *very small* (e.g., binary) inputs, for a number of LWE samples that exceeds the LWE dimension.

Our results and proofs are more naturally formulated using the SIS function family. So, we will first study the problem in terms of SIS, and then reformulate the results in terms of LWE using Proposition 2.9. We recall that the main difference between this section and Section 3, is that here we consider parameters for which the resulting functions are essentially injective, or more formally, statistically second-preimage resistant. The following lemma gives sufficient conditions that ensure this property.

Lemma 4.1. *For any integers m, k, q, s and set $X \subseteq [s]^m$, the function family $\mathrm{SIS}(m, k, q)$ is (statistically) ϵ-second preimage resistant with respect to the uniform input distribution $\mathcal{U}(X)$ for $\epsilon = |X| \cdot (s'/q)^k$, where s' is the largest factor of q smaller than s.*

Proof. Let $\mathbf{x} \leftarrow \mathcal{U}(X)$ and $\mathbf{A} \leftarrow \mathrm{SIS}(m, k, q)$ be chosen at random. We want to evaluate the probability that there exists an $\mathbf{x}' \in X \setminus \{\mathbf{x}\}$ such that $\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{x}' \pmod q$, or, equivalently, $\mathbf{A}(\mathbf{x} - \mathbf{x}') = \mathbf{0} \pmod q$. Fix any two distinct vectors $\mathbf{x}, \mathbf{x}' \in X$ and let $\mathbf{z} = \mathbf{x} - \mathbf{x}'$. The vector $\mathbf{A}\mathbf{z} \pmod q$ is distributed uniformly at random in $(d\mathbb{Z}/q\mathbb{Z})^k$, where $d = \gcd(q, z_1, \ldots, z_m)$. All coordinates of $\mathbf{z}$ are in the range $z_i \in \{-(s-1), \ldots, (s-1)\}$, and at least one of them is nonzero. Therefore, d is at most s' and $|d\mathbb{Z}_q^k| = (q/d)^k \geq (q/s')^k$. By union bound (over $\mathbf{x}' \in X \setminus \{\mathbf{x}\}$) for any $\mathbf{x}$, the probability that there is a second preimage $\mathbf{x}'$ is at most $(|X| - 1)(s'/q)^k$. □

We remark that, as shown in Section 3, even for parameter settings that do not fall within the range specified in Lemma 4.1, $\mathrm{SIS}(m, k, q)$ is collision resistant, and therefore also (computationally) second-preimage-resistant. This is all that is needed in the rest of this section. However, when $\mathrm{SIS}(m, k, q)$ is not statistically second-preimage resistant, the one-wayness proof that follows (see Theorem 4.5) is not very interesting: typically, in such settings, $\mathrm{SIS}(m, k, q)$ is also statistically uninvertible, and the one-wayness of $\mathrm{SIS}(m, k, q)$ directly follows from Lemma 2.2. So, below we focus on parameter settings covered by Lemma 4.1.

We prove the one-wayness of $\mathcal{F} = \mathrm{SIS}(m, k, q, X)$ with respect to the uniform input distribution $\mathcal{X} = \mathcal{U}(X)$ by building a lossy function family $(\mathcal{L}, \mathcal{F}, \mathcal{X})$ where $\mathcal{L}$ is an auxiliary function family that we will prove to be uninvertible and computationally indistinguishable from $\mathcal{F}$. The auxiliary family $\mathcal{L}$ is derived from the following function family.

Definition 4.2. *For any probability distribution $\mathcal{Y}$ over $\mathbb{Z}^\ell$ and integer $m \geq \ell$, let $\mathcal{I}(m, \ell, \mathcal{Y})$ be the probability distribution over linear functions $[\mathbf{I} \mid \mathbf{Y}] \colon \mathbb{Z}^m \to \mathbb{Z}^\ell$ where $\mathbf{I}$ is the $\ell \times \ell$ identity matrix, and $\mathbf{Y} \in \mathbb{Z}^{\ell \times (m-\ell)}$ is obtained choosing each column of $\mathbf{Y}$ independently at random from $\mathcal{Y}$.*

We anticipate that we will set $\mathcal{Y}$ to the Gaussian input distribution $\mathcal{Y} = D^\ell_{\mathbb{Z},\sigma}$ in order to make $\mathcal{L}$ indistinguishable from $\mathcal{F}$ under a standard LWE assumption. But for generality, we prove some of our results with respect to a generic distribution $\mathcal{Y}$.

The following lemma shows that for a bounded distribution $\mathcal{Y}$ (and appropriate parameters), $\mathcal{I}(m, \ell, \mathcal{Y})$ is (statistically) uninvertible.

Lemma 4.3. *Let $\mathcal{Y}$ be a probability distribution on $[\mathcal{Y}] \subseteq \{-\sigma, \ldots, \sigma\}^n$, and let $X \subseteq \{-s, \ldots, s\}^m$. Then $\mathcal{I}(m, \ell, \mathcal{Y})$ is ϵ-uninvertible with respect to $\mathcal{U}(X)$ for $\epsilon = (1 + 2s(1 + \sigma(m - \ell)))^\ell / |X|$.*

Proof. Let $f = [\mathbf{I} \mid \mathbf{Y}]$ be an arbitrary function in the support of $\mathcal{I}(m, \ell, \mathcal{Y})$. We know that $|y_{i,j}| \leq \sigma$ for all i, j. We first bound the size of the image $|f(X)|$. By the triangle inequality, all the points in the image $f(X)$ have ℓ_∞ norm at most

$$\|f(\mathbf{u})\|_\infty \leq \|\mathbf{u}\|_\infty (1 + \sigma(m - \ell)) \leq s(1 + \sigma(m - \ell)).$$

The number of integer vectors (in $\mathbb{Z}^\ell$) with such bounded ℓ_∞ norm is

$$(1 + 2s(1 + \sigma(m - \ell)))^\ell.$$

Dividing by the size of X and using Lemma 2.4, the claim follows. □

Lemma 4.3 applies to any distribution $\mathcal{Y}$ with bounded support. When $\mathcal{Y} = D^\ell_{\mathbb{Z},\sigma}$ is a discrete Gaussian distribution, a slightly better bound can be obtained. (See also [4], which proves a similar lemma for a different, non-uniform input distribution $\mathcal{X}$.)

Lemma 4.4. *Let $\mathcal{Y} = D^\ell_{\mathbb{Z},\sigma}$ be the discrete Gaussian distribution with parameter $\sigma > 0$, and let $X \subseteq \{-s, \ldots, s\}^m$. Then $\mathcal{I}(m, \ell, \mathcal{Y})$ is ϵ-uninvertible with respect to $\mathcal{U}(X)$, for $\epsilon = O(\sigma m s / \sqrt{\ell})^\ell / |X| + 2^{-\Omega(m)}$.*

Proof. Again, by Lemma 2.4 it is enough to bound the expected size of $f(X)$ when $f \leftarrow \mathcal{I}(m, \ell, \mathcal{Y})$ is chosen at random. Remember that $f = [\mathbf{I} \mid \mathbf{Y}]$ where $\mathbf{Y} \leftarrow D^{\ell \times (m-\ell)}_{\mathbb{Z},\sigma}$. Since the entries of $\mathbf{Y} \in \mathbb{R}^{\ell \times (m-\ell)}$ are independent mena-zero subgaussians with parameter σ, by a standard bound from the theory of random matrices, the largest singular value $s_1(\mathbf{Y}) = \max_{\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^m} \|\mathbf{Y}\mathbf{x}\| / \|\mathbf{x}\|$ of $\mathbf{Y}$ is at most $\sigma \cdot O(\sqrt{\ell} + \sqrt{m - \ell}) =$

$\sigma \cdot O(\sqrt{m})$, except with probability $2^{-\Omega(m)}$. We now bound the ℓ_2 norm of all vectors in the image $f(X)$. Let $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2) \in X$, with $\mathbf{u}_1 \in \mathbb{Z}^\ell$ and $\mathbf{u}_2 \in \mathbb{Z}^{m-\ell}$. Then

$$\begin{aligned}
\|f(\mathbf{u})\| &\leq \|\mathbf{u}_1 + \mathbf{Y}\mathbf{u}_2\| \\
&\leq \|\mathbf{u}_1\| + \|\mathbf{Y}\mathbf{u}_2\| \\
&\leq \left(\sqrt{\ell} + s_1(\mathbf{Y})\sqrt{m-\ell}\right) s \\
&\leq \left(\sqrt{\ell} + \sigma \cdot O(\sqrt{m})\sqrt{m-\ell}\right) s \\
&= O(\sigma m s).
\end{aligned}$$

The number of integer points in the ℓ-dimensional zero-centered ball of radius $R = O(\sigma m s)$ can be bounded by a simple volume argument, as $|f(X)| \leq (R + \sqrt{\ell}/2)^n V_\ell = O(\sigma m s/\sqrt{\ell})^\ell$, where $V_\ell = \pi^{\ell/2}/(\ell/2)!$ is the volume of the ℓ-dimensional unit ball. Dividing by the size of X and accounting for the rare event that $s_1(\mathbf{Y})$ is not bounded as above, we get that $\mathcal{I}(m, \ell, \mathcal{Y})$ is ϵ-uninvertible for $\epsilon = O(\sigma m s/\sqrt{\ell})^\ell/|X| + 2^{-\Omega(m)}$. $\square$

We can now prove the one-wayness of the SIS function family by defining and analyzing an appropriate lossy function family. The parameters below are set up to expose the connection with LWE, via Proposition 2.9: $\mathrm{SIS}(m, m-n, q)$ corresponds to LWE in n dimensions (given m samples), whose one-wayness we are proving, while $\mathrm{SIS}(\ell = m - n + k, m - n, q)$ corresponds to LWE in $k \leq n$ dimensions, whose pseudorandomness we are assuming.

Theorem 4.5. *Let q be a modulus and let $\mathcal{X}, \mathcal{Y}$ be two distributions over $\mathbb{Z}^m$ and $\mathbb{Z}^\ell$ respectively, where $\ell = m - n + k$ for some $0 < k \leq n \leq m$, such that*

1. *$\mathcal{I}(m, \ell, \mathcal{Y})$ is uninvertible with respect to input distribution $\mathcal{X}$,*
2. *$\mathrm{SIS}(\ell, m-n, q)$ is pseudorandom with respect to input distribution $\mathcal{Y}$, and*
3. *$\mathrm{SIS}(m, m-n, q)$ is second-preimage resistant with respect to input distribution $\mathcal{X}$.*

Then $\mathcal{F} = \mathrm{SIS}(m, m-n, q)$ is one-way with respect to input distribution $\mathcal{X}$.

In particular, if $\mathrm{SIS}(\ell, m-n, q)$ is pseudorandom with respect to the discrete Gaussian distribution $\mathcal{Y} = D^\ell_{\mathbb{Z},\sigma}$, then $\mathrm{SIS}(m, m-n, q)$ is $(2\epsilon + 2^{-\Omega(m)})$-one-way with respect to the uniform input distribution $\mathcal{X} = \mathcal{U}(X)$ over any set $X \subseteq \{-s, \ldots, s\}^m$ satisfying

$$(C'\sigma m s/\sqrt{\ell})^\ell/\epsilon \leq |X| \leq \epsilon \cdot (q/s')^{m-n},$$

where s' is the largest divisor of q that is smaller than or equal to $2s$, and C' is the universal constant hidden by the $O(\cdot)$ notation from Lemma 4.4.

Proof. We will prove that $(\mathcal{L}, \mathcal{F}, \mathcal{X})$ is a lossy function family, where $\mathcal{F} = \mathrm{SIS}(m, m-n, q)$ and $\mathcal{L} = \mathrm{SIS}(\ell, m-n, q) \circ \mathcal{I}(m, \ell, \mathcal{Y})$. It follows from Lemma 2.3 that both $\mathcal{F}$ and $\mathcal{L}$ are one-way function families with respect to input distribution $\mathcal{X}$. Notice that $\mathcal{F}$ is second-preimage resistant with respect to $\mathcal{X}$ by assumption. The indistinguishability of $\mathcal{L}$ and $\mathcal{F}$ follows immediately from the pseudorandomness of $\mathrm{SIS}(\ell, m-n, q)$ with respect to $\mathcal{Y}$, by a standard hybrid argument. So, in order to prove that $(\mathcal{L}, \mathcal{F}, \mathcal{X})$ is a lossy function family, it suffices to prove that $\mathcal{L}$ is uninvertible with respect to $\mathcal{X}$. This follows applying Lemma 2.5 to the function family $\mathcal{I}(m, \ell, \mathcal{Y})$, which is uninvertible by assumption. This proves the first part of the theorem.

Now consider the particular instantiation. Let $\mathcal{X} = \mathcal{U}(X)$ be the uniform distribution over a set $X \subseteq \{-s, \ldots, s\}^m$ whose size satisfies the inequalities in the theorem statement, and let $\mathcal{Y} = D_{\mathbb{Z},\sigma}^{\ell}$. Since $|X|(s'/q)^{m-n} \leq \epsilon$, by Lemma 4.1, $\mathrm{SIS}(m, m-n, q)$ is (statistically) second-preimage resistant with respect to input distribution $\mathcal{X}$. Moreover, since $(C\sigma ms/\sqrt{\ell})^{\ell}/|X| \leq \epsilon$, by Lemma 4.4, $\mathcal{I}(m, \ell, \mathcal{Y})$ is $(\epsilon + 2^{-\Omega(m)})$-uninvertible with respect to input distribution $\mathcal{X}$. □

In order to conclude that the LWE function is pseudorandom (under worst-case lattice assumptions) for uniformly random small errors, we combine Theorem 4.5 with Corollary 2.14, instantiating the parameters appropriately. For simplicity, we focus on the important case of a prime modulus q. Nearly identical results for composite moduli (e.g., those divisible by only small primes) are also easily obtained from Corollary 2.14, or by using either Theorem 2.13 or Theorem 2.12.

Theorem 4.6. *Let $0 < k \leq n \leq m - \omega(\log k) \leq k^{O(1)}$, $\ell = m - n + k$, $s \geq (Cm)^{\ell/(n-k)}$ for a large enough universal constant C, and q be a prime such that $\max\{3\sqrt{k}, (4s)^{m/(m-n)}\} \leq q \leq k^{O(1)}$. For any set $X \subseteq \{-s, \ldots, s\}^m$ of size $|X| \geq s^m$, the $\mathrm{SIS}(m, m-n, q)$ (equivalently, $\mathrm{LWE}(m, n, q)$) function family is one-way (and pseudorandom) with respect to the uniform input distribution $\mathcal{X} = \mathcal{U}(X)$, under the assumption that SIVP_γ is (quantum) hard to approximate, in the worst case, on k-dimensional lattices to within a factor $\gamma = \tilde{O}(\sqrt{k} \cdot q)$.*

A few notable instantiations are as follows. To obtain pseudorandomness for binary errors, we need $s = 2$ and $X = \{0, 1\}^m$. For this value of s, the condition $s \geq (Cm)^{\ell/(n-k)}$ can be equivalently be rewritten as

$$m \leq (n-k) \cdot \left(1 + \frac{1}{\log_2(Cm)}\right),$$

which can be satisfied by taking $k = n/(C' \log_2 n)$ and $m = n(1 + 1/(c \log_2 n))$ for any desired $c > 1$ and a sufficiently large constant $C' > 1/(1 - 1/c)$. For these values, the modulus should satisfy $q \geq 8^{m/(m-n)} = 8n^{3c} = k^{O(1)}$, and can be set to any sufficiently large prime $p = k^{O(1)}$.[1]

Notice that for binary errors, both the worst-case lattice dimension k and the number $m - n$ of "extra" LWE samples (i.e., the number of samples beyond the LWE dimension n) are both sublinear in the LWE dimension n: we have $k = \Theta(n/\log n)$ and $m - n = O(n/\log n)$. This corresponds to both a stronger worst-case security assumption, and a less useful LWE problem. By using larger errors, say, bounded by $s = n^{\epsilon}$ for some constant $\epsilon > 0$, it is possible to make both the worst-case lattice dimension k and number of extra samples $m - n$ into (small) linear functions of the LWE dimension n, which may be sufficient for some cryptographic applications of LWE. Specifically, for any constant $\epsilon < 1$, one may set $k = (\epsilon/3)n$ and $m = (1 + \epsilon/3)n$, which are easily verified to satisfy all the hypotheses of Theorem 4.6 when $q = k^{O(1)}$ is sufficiently large. These parameters correspond to $(\epsilon/3)n = \Omega(n)$ extra samples (beyond the LWE dimension n), and to the worst-case hardness of lattice problems in dimension $(\epsilon/3)n = \Omega(n)$. Notice that for $\epsilon < 1/2$, this version of LWE has much smaller errors than allowed by previous LWE hardness proofs, and it would be subject to subexponential-time attacks [2] if the number of samples were not restricted. Our result shows that if the number of samples is limited to $(1 + \epsilon/3)n$, then LWE maintains its provable security properties and conjectured exponential-time hardness in the dimension n.

One last instantiation allows for a linear number of samples $m = c \cdot n$ for any desired constant $c \geq 1$, which is enough for most applications of LWE in lattice cryptography. In this case we can choose (say)

[1] Here we have not tried to optimize the value of q, and smaller values of the modulus are certainly possible: a close inspection of the proof of Theorem 4.6 reveals that for binary errors, the condition $q \geq 8n^{3c}$ can be replaced by $q \geq n^{c'}$ for any constant $c' > c$.

$k = n/2$, and it suffices to set the other parameters so that

$$s \geq (Cm)^{2c-1} \quad \text{and} \quad q \geq (4s)^{c/(c-1)} \geq 4^{c/(c-1)} \cdot (Ccn)^{2c+1+1/(c-1)} = k^{O(1)}.$$

(We can also obtain better lower bounds on s and q by letting k be a smaller constant fraction of n.) This proves the hardness of LWE with uniform noise of polynomial magnitude $s = n^{O(1)}$, and any linear number of samples $m = O(n)$. Note that for $m = cn$, any instantiation of the parameters requires the magnitude s of the errors to be at least n^{c-1}. For $c > 3/2$, this is more noise than is typically used in the standard LWE problem, which allows errors of magnitude as small as $O(\sqrt{n})$, but requires them to be independent and follow a Gaussian-like distribution. The novelty in this last instantiation of Theorem 4.6 is that it allows for a much wider class of error distributions, including the uniform distribution, and distributions where different components of the error vector are correlated.

Proof of Theorem 4.6. We prove the one-wayness of $\mathrm{SIS}(m, m-n, q)$ (equivalently, $\mathrm{LWE}(m, n, q)$ via Proposition 2.9) using the second part of Theorem 4.5 with $\sigma = 3\sqrt{k}$. Using $\ell \geq k$ and the primality of q, the conditions on the size of X in Theorem 4.5 can be replaced by simpler bounds

$$\frac{(3C'ms)^\ell}{\epsilon} \leq |X| \leq \epsilon \cdot q^{m-n},$$

or equivalently, the requirement that the quantities $(3C'ms)^\ell/|X|$ and $|X|/q^{m-n}$ are negligible in k. For the first quantity, letting $C = 4C'$ and using $|X| \geq s^m$ and $s \geq (4C'm)^{\ell/(n-k)}$, we get that $(3C'ms)^\ell/|X| \leq (3/4)^{-\ell} \leq (3/4)^{-k}$ is exponentially small (in k). For the second quantity, using $|X| \leq (2s+1)^m$ and $q \geq (4s)^{m/(m-n)}$, we get that $|X|/q^{m-n} \leq (3/4)^m$ is also exponentially small.

Theorem 4.5 also requires the pseudorandomness of $\mathrm{SIS}(\ell, m-n, q)$ with respect to the discrete Gaussian input distribution $\mathcal{Y} = D^\ell_{\mathbb{Z},\sigma}$, which can be based on the (quantum) worst-case hardness of SIVP on k-dimensional lattices using Corollary 2.14. (Notice the use of different parameters: $\mathrm{SIS}(m, m-n, q)$ in Corollary 2.14, and $\mathrm{SIS}(m-n+k, m-n, q)$ here.) After properly renaming the variables, and using $\sigma = 3\sqrt{k}$, the hypotheses of Corollary 2.14 become $\omega(\log k) \leq m - n \leq k^{O(1)}$, $3\sqrt{k} < q < k^{O(1)}$, which are all satisfied by the hypotheses of the Theorem. The corresponding assumption is the worst-case hardness of SIVP_γ on k-dimensional lattices, for $\gamma = k\omega_k q/\sigma = \sqrt{k}\omega_k q/3 = \tilde{O}(\sqrt{k}q)$, as claimed. This concludes the proof of the one-wayness of LWE.

The pseudorandomness of LWE follows from the sample-preserving search-to-decision reduction of [17]. □

References

[1] M. Ajtai. Generating hard instances of lattice problems. *Quaderni di Matematica*, 13:1–32, 2004. Preliminary version in STOC 1996.

[2] S. Arora and R. Ge. New algorithms for learning in presence of errors. In *ICALP (1)*, pages 403–415, 2011.

[3] A. Banerjee, C. Peikert, and A. Rosen. Pseudorandom functions and lattices. In *EUROCRYPT*, pages 719–737, 2012.

[4] M. Bellare, E. Kiltz, C. Peikert, and B. Waters. Identity-based (lossy) trapdoor functions and applications. In *EUROCRYPT*, pages 228–245, 2012.

[5] A. Blum, A. Kalai, and H. Wasserman. Noise-tolerant learning, the parity problem, and the statistical query model. *J. ACM*, 50(4):506–519, 2003.

[6] D. Boneh and D. M. Freeman. Linearly homomorphic signatures over binary fields and new tools for lattice-based signatures. In *Public Key Cryptography*, pages 1–16, 2011.

[7] J.-Y. Cai and A. Nerurkar. An improved worst-case to average-case connection for lattice problems. In *FOCS*, pages 468–477, 1997.

[8] Y. Chen and P. Q. Nguyen. BKZ 2.0: Better lattice security estimates. In *ASIACRYPT*, pages 1–20, 2011.

[9] D. Dadush, C. Peikert, and S. Vempala. Enumerative lattice algorithms in any norm via M-ellipsoid coverings. In *FOCS*, pages 580–589, 2011.

[10] N. Döttling and J. Müller-Quade. Lossy codes and a new variant of the learning-with-errors problem. Manuscript. To appear in Eurocrypt 2013, 2013.

[11] N. Gama and P. Q. Nguyen. Predicting lattice reduction. In *EUROCRYPT*, pages 31–51, 2008.

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

[13] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261(4):515–534, December 1982.

[14] D. Micciancio. Almost perfect lattices, the covering radius problem, and applications to Ajtai's connection factor. *SIAM J. Comput.*, 34(1):118–169, 2004.

[15] D. Micciancio. Duality in lattice cryptography. In *Public Key Cryptography*, 2010. Invited talk.

[16] D. Micciancio and S. Goldwasser. *Complexity of Lattice Problems: a cryptographic perspective*, volume 671 of *The Kluwer International Series in Engineering and Computer Science*. Kluwer Academic Publishers, Boston, Massachusetts, 2002.

[17] D. Micciancio and P. Mol. Pseudorandom knapsacks and the sample complexity of LWE search-to-decision reductions. In *CRYPTO*, pages 465–484, 2011.

[18] D. Micciancio and C. Peikert. Trapdoors for lattices: Simpler, tighter, faster, smaller. In *EUROCRYPT*, pages 700–718, 2012.

[19] D. Micciancio and O. Regev. Worst-case to average-case reductions based on Gaussian measures. *SIAM J. Comput.*, 37(1):267–302, 2007. Preliminary version in FOCS 2004.

[20] D. Micciancio and O. Regev. Lattice-based cryptography. In *Post Quantum Cryptography*, pages 147–191. Springer, February 2009.

[21] C. Peikert. Public-key cryptosystems from the worst-case shortest vector problem. In *STOC*, pages 333–342, 2009.

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

[23] C. Peikert and B. Waters. Lossy trapdoor functions and their applications. In *STOC*, pages 187–196, 2008.

[24] O. Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6):1–40, 2009. Preliminary version in STOC 2005.

[25] D. Wagner. A generalized birthday problem. In *CRYPTO*, pages 288–303, 2002.