Learning with Rounding, Revisited

New Reduction, Properties and Applications*

Joël Alwen[1], Stephan Krenn[2], Krzysztof Pietrzak[3], and Daniel Wichs[4]

[1] ETH Zurich
alwenj@inf.ethz.ch
[2] IBM Research – Zurich
skr@zurich.ibm.com
[3] Institute of Science and Technology Austria
pietrzak@ist.ac.at
[4] Northeastern University
wichs@ccs.neu.edu

Abstract. The learning with rounding (LWR) problem, introduced by Banerjee, Peikert and Rosen at EUROCRYPT '12, is a variant of learning with errors (LWE), where one replaces random errors with deterministic rounding. The LWR problem was shown to be as hard as LWE for a setting of parameters where the modulus and modulus-to-error ratio are super-polynomial. In this work we resolve the main open problem and give a new reduction that works for a larger range of parameters, allowing for a polynomial modulus and modulus-to-error ratio. In particular, a smaller modulus gives us greater efficiency, and a smaller modulus-to-error ratio gives us greater security, which now follows from the worst-case hardness of GapSVP with polynomial (rather than super-polynomial) approximation factors.

As a tool in the reduction, we show that there is a "lossy mode" for the LWR problem, in which LWR samples only reveal partial information about the secret. This property gives us several interesting new applications, including a proof that LWR remains secure with weakly random secrets of sufficient min-entropy, and very simple constructions of deterministic encryption, lossy trapdoor functions and reusable extractors.

Our approach is inspired by a technique of Goldwasser et al. from ICS '10, which implicitly showed the existence of a "lossy mode" for LWE. By refining this technique, we also improve on the parameters of that work to only requiring a polynomial (instead of super-polynomial) modulus and modulus-to-error ratio.

Keywords: Learning with Errors, Learning with Rounding, Lossy Trapdoor Functions, Deterministic Encryption.

* This work was partly funded by the European Research Council under ERC Starting Grant 259668-PSPC and ERC Advanced Grant 321310-PERCY. Parts of this work were done while the second authors was at IST Austria, and the last author was at IBM Research, T.J. Watson. A full version of this paper is available online [1].

R. Canetti and J.A. Garay (Eds.): CRYPTO 2013, Part I, LNCS 8042, pp. 57–74, 2013.

1 Introduction

Learning With Errors. The Learning with Errors (LWE) assumption states that "noisy" inner products of a secret vector with random public vectors, look pseudorandom. In the last years many cryptosystems have been proven secure under LWE, including (identity-based, leakage-resilient, fully homomorphic, functional) encryption [2–9], pseudorandom functions [10], (blind) signature schemes [3, 11–13], hash functions [14, 15], oblivious transfer [16], etc..

The LWE assumption with parameters $n, m, q \in \mathbb{N}$ and a "small" error distribution χ over $\mathbb{Z}$ states that for uniformly random $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}$, $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$, $\mathbf{u} \xleftarrow{\$} \mathbb{Z}_q^m$ and an error vector $\mathbf{e} \leftarrow \chi^m$

$$(\mathbf{A}, \mathbf{A} \cdot \mathbf{s} + \mathbf{e}) \text{ is computationally indistinguishable from } (\mathbf{A}, \mathbf{u}).$$

Sometimes it will be convenient to think of this distribution as consisting of m "LWE samples" of the form $(\mathbf{a}_i, \langle \mathbf{a}_i, s \rangle + e_i) \in \mathbb{Z}_q^{n+1}$. One of the main advantages of the LWE problem is that, for some settings of parameters, we can prove its security under certain worst-case hardness assumptions over lattices, cf. [2, 17]. One important parameter is the "size" of the error terms $e \xleftarrow{\$} \chi$ which we denote by β.[1] As long as β exceeds some minimum threshold $\approx \sqrt{n}$, the concrete hardness of the LWE problem mainly depends on the dimension n and on the ratio of the modulus q to the error-size β. Therefore, we will often be unspecific about the exact distribution χ, and only focus on the error-size β.

Learning With Rounding. The Learning with Rounding (LWR) problem was introduced in [10]. Instead of adding a small random error to a sample $\langle \mathbf{a}, \mathbf{s} \rangle \in \mathbb{Z}_q$ to hide its exact value, we release a *deterministically rounded* version of $\langle \mathbf{a}, \mathbf{s} \rangle$. That is, for some $p < q$, we divide up the elements of $\mathbb{Z}_q$ into p contiguous intervals of roughly q/p elements each and define the *rounding function* $\lfloor \cdot \rfloor_p : \mathbb{Z}_q \to \mathbb{Z}_p$ that maps $x \in \mathbb{Z}_q$ into the index of the interval that x belongs to. For example if q, p are both powers of 2, than this could correspond to outputting the $\log(p)$ most significant bits of x. We can extend the rounding function to vectors by applying it component-wise. The LWR assumption states that:

$$(\mathbf{A}, \lfloor \mathbf{A} \cdot \mathbf{s} \rfloor_p) \text{ is computationally indistinguishable from } (\mathbf{A}, \lfloor \mathbf{u} \rfloor_p).$$

Note that if p divides q, then $\lfloor \mathbf{u} \rfloor_p$ is itself uniform over $\mathbb{Z}_p^m$.

The main advantage of LWR is that one does not need to sample any additional "errors", therefore requiring fewer random bits. The assumption has been used to construct simple and efficient pseudorandom generators and functions in [10], and deterministic encryption in [18].

Banerjee et al. [10] show a beautifully simple reduction proving the hardness of the LWR problem under the LWE assumption for some range of parameters. They observe that if the error size β is sufficiently small and the ratio q/p is sufficiently big, then $\lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rfloor_p = \lfloor \langle \mathbf{a}, \mathbf{s} \rangle + e \rfloor_p$ with overwhelming probability over

[1] We will be informal for now; we can think of β as the the standard deviation or the expected/largest absolute value of the errors.

random $\mathbf{a} \xleftarrow{\$} \mathbb{Z}_q$ and $e \xleftarrow{\$} \chi$. In particular, the only way that the two values differ is if $\langle \mathbf{a}, \mathbf{s} \rangle$ ends up within a distance of $|e|$ from a boundary between two different intervals; but since the intervals are of size q/p and the ball around the boundary is only of size $2|e|$ this is unlikely to happen when q/p is super-polynomially bigger than $2|e|$. Therefore, one can show that:

$$(\mathbf{A}, \lfloor \mathbf{A} \cdot \mathbf{s} \rfloor_p) \overset{\text{stat}}{\approx} (\mathbf{A}, \lfloor \mathbf{A} \cdot \mathbf{s} + \mathbf{e} \rfloor_p) \overset{\text{comp}}{\approx} (\mathbf{A}, \lfloor \mathbf{u} \rfloor_p)$$

where the first modification is statistically close and the second follows immediately from the hardness of LWE.

Unfortunately, the argument only goes through, when (q/p) is bigger than the error size β by a super-polynomial factor. In fact, if we want statistical distance $2^{-\lambda}$ we would need to set $q \geq 2^{\lambda} \beta p$, where λ is a security parameter. This has three important consequences: (1) the modulus q has to be super-polynomial, which makes all of the computations less efficient, (2) the modulus-to-error ratio q/β is super-polynomial which makes the LWE problem easier and only gives us a reduction if we assume the hardness of the lattice problem GapSVP with super-polynomial approximation factors (a stronger assumption), (3) the ratio of the input-to-output modulus q/p is super-polynomial, meaning that we must "throw away" a lot of information when rounding and therefore get fewer bits of output per LWR sample. The work of [10] conjectured that the LWR problem should be hard even for a polynomial modulus q, but left it as the main open problem to give a reduction. The conjecture is especially interesting in light of the recent results of [19] which give the first *classical* reduction from LWE with small parameters to GapSVP.

1.1 The New Reduction and Properties of LWR

LWR with Polynomial Modulus. In this work, we resolve the open problem of [10] and give a new reduction showing the hardness of LWR from that of LWE for a more general setting of parameters, including when the modulus q is only polynomial. In particular, instead of requiring $q \geq 2^{\lambda} \beta p$, where λ is a security parameter as in [10], we only require $q \geq nm\beta p$, where we recall that n is the dimension of the secret, and m is the number of LWR samples that we output, β is the size of the LWE errors, and p is the new modulus we round to. In particular, as long as the number of LWR samples m is fixed a-priori by some polynomial, we can allow the modulus q (and therefore also the modulus-to-error ratio q/β, and the input-to-output ratio q/p) to all be polynomial. As mentioned, this setting provides greater efficiency (computation with smaller q) and greater security (smaller ratio q/β) allowing for a reduction from the worst-case hardness of the lattice problem GapSVP with polynomial approximation factors. In particular, the above efficiency and security improvements for LWR directly translate into improvements of the PRG and PRF constructions of [10].

To be even more precise, our reduction shows the hardness of LWR with parameters n, m, q, p assuming the hardness of LWE with parameters n', m, q, β (note: different dimension n' vs. n) as long as:

$$n \geq \frac{\log(q)}{\log(2\gamma)} \cdot n' \qquad \text{and} \qquad q \geq \gamma(nm\beta p) \tag{1}$$

for some flexible parameter $\gamma \geq 1$. For example, setting $\gamma = 1$ allows for the smallest modulus $q \approx nm\beta p$, but requires a larger dimension $n \approx n' \log(q)$ in the LWR problem than the dimension n' of the underlying LWE assumption. On the other hand, setting $\gamma = q^{\delta}$ for some constant $\delta \in (0, 1)$ gives a bigger polynomial modulus $q \approx (nm\beta p)^{1/(1-\delta)}$ but allow us to set the LWR dimension $n \approx (1/\delta)n' = O(n')$ to be closer to that of the underlying LWE assumption.

It remains as an open problem to improve the reduction further, and especially to remove the dependence between the modulus q and the number of LWR samples m that we give out.

LWR with Weak and Leaky Secrets. Another advantage of our reduction is that we prove the security of the LWR problem even when the secret $\mathbf{s}$ is not necessarily uniform over $\mathbb{Z}_q^n$. Indeed, our proof also works when $\mathbf{s}$ is uniform over a smaller integer interval $\mathbf{s} \xleftarrow{\$} \{-\gamma, \ldots, \gamma\}^n \subseteq \mathbb{Z}_q^n$, where the relation of $\gamma \geq 1$ to the other parameters is given by equation (1). Moreover, our reduction works when the secret $\mathbf{s}$ is not even truly uniform over this interval (say, because the attacker observed some leakage on $\mathbf{s}$, or $\mathbf{s}$ was sampled using a weak random source) as long as $\mathbf{s}$ retains some sufficiently high amount of *min-entropy* $k \approx n' \log(q)$, where n' is the dimension of the underlying LWE assumption. Notice that, no matter how small the entropy k is, we can still prove some level of security under an LWE assumption with correspondingly smaller dimension n'.

The work of Goldwasser et al. [20] shows similar results for the hardness of LWE with a weak and leaky secret, at least as long as the modulus q and the modulus-to-error ratio q/β are super-polynomial. Indeed, we will use a refinement of the technique from their work as the basis of our LWE to LWR reduction. Our refinement will also allow us to improve the parameters of [20], and show the hardness of LWE with a weak and leaky secret when the modulus q and the ratio q/β are polynomial.

The Reduction. As discussed above, the original reduction of [10] required us to choose parameters so that rounding samples with and without error is almost always identical: $\Pr[\lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rfloor_p \neq \lfloor \langle \mathbf{a}, \mathbf{s} \rangle + e \rfloor_p] \leq \mathsf{negl}$. Therefore LWR outputs do not provide any more information than LWE outputs. In contrast, in our setting of parameters, when q is polynomial, there is a noticeable probability that the two values are different. We therefore need a completely different proof strategy.

Surprisingly, our strategy does *not* directly convert an LWE instance with secret $\mathbf{s}$ into an LWR instance with secret $\mathbf{s}$. Instead, we rely on the LWE problem to change the distribution of the coefficient matrix $\mathbf{A}$. In particular, we show that there is a "lossy" method of sampling a matrix $\tilde{\mathbf{A}} \xleftarrow{\$} \mathsf{Lossy}()$ such that:

(a) Under the LWE assumption, $\tilde{\mathbf{A}} \xleftarrow{\$} \mathsf{Lossy}()$ is computationally indistinguishable from $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}$.

(b) When $\tilde{\mathbf{A}} \xleftarrow{\$} \mathsf{Lossy}()$, the values $\tilde{\mathbf{A}}, \lfloor \tilde{\mathbf{A}} \cdot \mathbf{s} \rfloor_p$ do not reveal too much information about $\mathbf{s}$. In particular, $\mathbf{s}$ maintains a large fraction of its statistical entropy given $\tilde{\mathbf{A}}, \lfloor \tilde{\mathbf{A}} \cdot \mathbf{s} \rfloor_p$.

Before we describe how the $\mathsf{Lossy}()$ sampler works in the next paragraph, let us show that the above two properties allow us to prove the hardness of LWR problem. We can do so via a hybrid argument where, given many LWR samples, we replace one sample at a time from being an LWR sample to being uniformly random. In particular, assume we have $m+1$ LWR samples and let the matrix $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}$ denote the coefficient vectors of the first m samples, and let $\mathbf{a} \xleftarrow{\$} \mathbb{Z}_q^n$ be the coefficient vector of the last sample. Then we can show:

$$\left(\begin{bmatrix}\mathbf{A}\\ \mathbf{a}\end{bmatrix}, \begin{bmatrix}\lfloor \mathbf{A} \cdot \mathbf{s} \rfloor_p \\ \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rfloor_p\end{bmatrix}\right) \overset{\text{comp}}{\approx} \left(\begin{bmatrix}\tilde{\mathbf{A}}\\ \mathbf{a}\end{bmatrix}, \begin{bmatrix}\lfloor \tilde{\mathbf{A}} \cdot \mathbf{s} \rfloor_p \\ \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rfloor_p\end{bmatrix}\right) \overset{\text{stat}}{\approx}$$
$$\left(\begin{bmatrix}\tilde{\mathbf{A}}\\ \mathbf{a}\end{bmatrix}, \begin{bmatrix}\lfloor \tilde{\mathbf{A}} \cdot \mathbf{s} \rfloor_p \\ \lfloor u \rfloor_p\end{bmatrix}\right) \overset{\text{comp}}{\approx} \left(\begin{bmatrix}\mathbf{A}\\ \mathbf{a}\end{bmatrix}, \begin{bmatrix}\lfloor \mathbf{A} \cdot \mathbf{s} \rfloor_p \\ \lfloor u \rfloor_p\end{bmatrix}\right)$$

In the first step, we use the LWE assumption to replace a uniformly random $\mathbf{A}$ by a lossy matrix $\tilde{\mathbf{A}} \xleftarrow{\$} \mathsf{Lossy}()$, but still choose the last row $\mathbf{a} \xleftarrow{\$} \mathbb{Z}_q^n$ at random. In the second step, we use the fact that *inner product* is a strong extractor, where we think of the secret $\mathbf{s}$ as the source and the vector $\mathbf{a}$ as a seed. In particular, by the properties of the lossy sampler, we know that $\mathbf{s}$ maintains entropy conditioned on seeing $\tilde{\mathbf{A}}, \lfloor \tilde{\mathbf{A}} \cdot \mathbf{s} \rfloor_p$ and therefore the "extracted value" $\langle \mathbf{a}, \mathbf{s} \rangle$ is statistically close to a uniformly random and independent $u \xleftarrow{\$} \mathbb{Z}_q$. In the last step, we simply replace the lossy matrix $\tilde{\mathbf{A}} \xleftarrow{\$} \mathsf{Lossy}()$ back by a uniformly random $\mathbf{A}$. This shows that, given the first m LWR samples the last one looks uniform and independent. We can then repeat the above steps m more times to replace each of the remaining LWR samples (rows) by uniform, one-by-one.

The Lossy Sampler. The basic idea of our Lossy sampler is taken from the work of Goldwasser et al. [20]. We sample the lossy matrix $\tilde{\mathbf{A}} \in \mathbb{Z}_q^{m \times n}$ as

$$\tilde{\mathbf{A}} \overset{\text{def}}{=} \mathbf{B}\mathbf{C} + \mathbf{F} \qquad \text{where } \mathbf{B} \xleftarrow{\$} \mathbb{Z}_q^{m \times n'} \;\;,\;\; \mathbf{C} \xleftarrow{\$} \mathbb{Z}_q^{n' \times n} \;\;,\;\; \mathbf{F} \xleftarrow{\$} \chi^{m \times n}$$

where $n' < n$ is some parameter and χ is a "small" LWE error distribution. We now need to show that this satisfies the properties (a) and (b) described above.

It is easy to see that $\tilde{\mathbf{A}}$ is computationally indistinguishable from a uniformly random matrix under the LWE assumption with parameters n', m, q, χ. In particular, each column i of the matrix $\tilde{\mathbf{A}}$ can be thought of as an LWE distribution $\mathbf{B} \cdot \mathbf{c}_i + \mathbf{f}_i$ with coefficient matrix $\mathbf{B}$, secret $\mathbf{c}_i$ which is the ith column of the matrix $\mathbf{C}$, and error vector $\mathbf{f}_i$ which is the ith column of $\mathbf{F}$. Therefore, using n hybrid arguments, we can replace each column i of $\tilde{\mathbf{A}}$ by a uniformly random and independent one. This part of the argument is the same as in [20].

Next, we need to show that the secret $\mathbf{s}$ retains entropy even conditioned on seeing $\tilde{\mathbf{A}}, \lfloor \tilde{\mathbf{A}} \cdot \mathbf{s} \rfloor_p$. Let us first prove this property in the case when $\mathbf{s} \xleftarrow{\$} \{-1, 0, 1\}^n$

is itself a random "short" vector.[2] All of the information that we give out about $\mathbf{s}$ can be reconstructed from:

- The matrices $\mathbf{B}, \mathbf{C}, \mathbf{F}$ which define $\tilde{\mathbf{A}}$ and are independent of $\mathbf{s}$ on their own.
- The value $\mathbf{C} \cdot \mathbf{s}$ whose bit-length is $n' \log(q)$.
- A set Z consisting of all pairs $(i, v_i) \in [m] \times \mathbb{Z}_p$ such that $\lfloor (\mathbf{BC} \cdot \mathbf{s})_i \rfloor_p \neq \lfloor (\tilde{\mathbf{A}} \cdot \mathbf{s})_i \rfloor_p$ along with the value $v_i = \lfloor (\tilde{\mathbf{A}} \cdot \mathbf{s})_i \rfloor_p$. The subscript i denotes the i^{th} component of a vector.

Given the three pieces of information above, we can reconstruct $\tilde{\mathbf{A}}, \lfloor \tilde{\mathbf{A}} \cdot \mathbf{s} \rfloor_p$ by setting $\lfloor (\tilde{\mathbf{A}} \cdot \mathbf{s})_i \rfloor_p := \lfloor (\mathbf{BC} \cdot \mathbf{s})_i \rfloor_p$ for every index i not contained in Z, and setting $\lfloor (\tilde{\mathbf{A}} \cdot \mathbf{s})_i \rfloor_p := v_i$ for every i which is in Z. Therefore, we just need to show that the three pieces of information above do not reveal too much about $\mathbf{s}$. First, we show that the set Z is small with overwhelming probability. In particular, an index i is contained in Z if and only if

$$\lfloor (\mathbf{BC} \cdot \mathbf{s})_i \rfloor_p \neq \lfloor (\mathbf{BC} \cdot \mathbf{s})_i + (\mathbf{F} \cdot \mathbf{s})_i \rfloor_p. \tag{2}$$

Assume that the entries of the error matrix $\mathbf{F}$ are all bounded by β in absolute value with overwhelming probability, and therefore $(\mathbf{F} \cdot \mathbf{s})_i$ is bounded by $n\beta$ in absolute value.[3] Then the event (2) can only occur if the value $(\mathbf{BC} \cdot \mathbf{s})_i$ falls within distance $n\beta$ of a boundary between two different intervals. Since each interval is of size $\approx q/p$ and the ball around each boundary is of size $2n\beta$, this happens with (noticeable but small) probability $\leq 2n\beta p/q \leq 1/m$, when $q \geq 2nm\beta p$ (which gives us the bound of (1)). Therefore, the probability of any index i being in Z is at most $1/m$, the expected size of Z is at most 1, and because these probabilities are independent, we can use Chernoff to bound $|Z| \leq n'$ with overwhelming probability $1 - 2^{-n'}$. So in total, Z can be described by $|Z|(\log m + \log p) \leq n' \log q$ bits with overwhelming probability. Therefore, together, $Z, \mathbf{Cs}$ reveal only $O(n' \log q)$ bits of information about $\mathbf{s}$, even given $\mathbf{B}, \mathbf{C}, \mathbf{F}$. We can summarize the above as:

$$\begin{aligned} H_\infty(\mathbf{s} | \tilde{\mathbf{A}}, \lfloor \tilde{\mathbf{A}}\mathbf{s} \rfloor_p) &\geq H_\infty(\mathbf{s} | \mathbf{B}, \mathbf{C}, \mathbf{F}, \mathbf{C} \cdot \mathbf{s}, Z) \\ &\geq H_\infty(\mathbf{s} | \mathbf{B}, \mathbf{C}, \mathbf{F}) - O(n' \log q) \geq n - O(n' \log q). \end{aligned}$$

Hence, if n is sufficiently larger than some $O(n' \log q)$, the LWR secret maintains a large amount of entropy given the LWR samples with a lossy $\tilde{\mathbf{A}}$. The above analysis also extends to the case where $\mathbf{s}$ is not uniformly random, but only has a sufficient amount of entropy.

We can also extend the above analysis to the case where $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$ is uniformly random over the entire space (and not short), by thinking of $\mathbf{s} = \mathbf{s}_1 + \mathbf{s}_2$ where $\mathbf{s}_1 \xleftarrow{\$} \mathbb{Z}_q^n$ is uniformly random and $\mathbf{s}_2 \xleftarrow{\$} \{-1, 0, 1\}^n$ is random and short. Using the same argument as above, we can show that, even given $\mathbf{s}_1$, $\tilde{\mathbf{A}}$ and $\lfloor \tilde{\mathbf{A}} \cdot \mathbf{s} \rfloor_p$, the value $\mathbf{s}_2$ (and therefore also $\mathbf{s}$) maintains entropy.

[2] This proof generalizes to larger intervals $\{-\gamma, \ldots, \gamma\}$ and corresponds to the parameter γ in equation (1). Here we set $\gamma = 1$.

[3] Our actual proof is more refined and only requires us to bound the *expected* absolute value of the entries.

Our analysis of lossiness as described above is inspired by [20] but differs from it significantly. In particular that work considered LWE (not LWR) samples with the matrix $\tilde{\mathbf{A}}$, did not explicitly analyze lossiness, and required super-polynomial modulus and modulus-to-error ratio. Indeed, in the full version [1] we use the ideas from the above analysis to also improve the parameters of that work, showing the robustness of the LWE problem to weak and leaky secrets for a polynomial modulus and modulus-to-error ratio.

1.2 Applications

Reusable Computational Extractors. By the leftover-hash lemma, the function $\mathsf{Ext}(\mathbf{s}; \mathbf{a}) := \langle \mathbf{s}, \mathbf{a} \rangle$ is a good randomness extractor taking a secret source $\mathbf{s} \in \mathbb{Z}_q^n$ of min-entropy $k \geq \log(q) + 2\log(1/\varepsilon)$ and a random public seed $\mathbf{a} \xleftarrow{\$} \mathbb{Z}_q^n$, and its output will be ε-close to the uniform over $\mathbb{Z}_q$. But assume we want to extract many different mutually (pseudo-)random values from the source $\mathbf{s}$ without keeping any long term state: each time we want to extract a new output we choose a fresh seed and apply the extractor. It is easy to see that the above inner-product extractor is completely insecure after at most n applications, and each successive output is easy to predict from the previous ones. The work of [21] introduced the notion of a *reusable computational extractor* that remains secure even after m applications, where m can be an arbitrary polynomial, and gave a construction under a non-standard "learning-subspaces with noise" assumption. Our results immediately give us a new simple construction of reusable extractors defined by $\mathsf{Ext}(\mathbf{s}; \mathbf{a}) := \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rfloor_p$. That is, we just round the output of the standard inner product extractor! We show that, as long as the LWE assumption holds with some parameters n', m, q, β, the source $\mathbf{s}$ is distributed over $\{0,1\}^n$ and has entropy $k \geq O(n' \log(q))$, and the modulus satisfies $q \geq 2\beta nmp$, the above extractor is secure for m uses. In particular, we can have $m \gg n \gg k$.

Lossy Trapdoor Functions. Lossy trapdoor functions (LTDFs) [22,23] are a family of functions $f_{pk}(\cdot)$ keyed by some public key pk, which can be sampled in one of two indistinguishable modes: `injective` and `lossy`. In the `injective` mode the function $f_{pk}(\cdot)$ is an injective function and we can even sample pk along with a secret trapdoor key sk that allows us to invert it efficiently. In the `lossy` mode, the function $f_{pk}(\cdot)$ is "many-to-one" and $f_{pk}(\mathbf{s})$ statistically loses information about the input $\mathbf{s}$. LTDFs have many amazing applications in cryptography, such as allowing us to output many hardcore bits, construct CCA-2 public-key encryption [23,24], and deterministic encryption [25]. We construct very simple and efficient LTDFs using the LWR problem: the public key is a matrix $pk = \mathbf{A}$ and the function is defined as $f_{\mathbf{A}}(\mathbf{s}) = \lfloor \mathbf{A} \cdot \mathbf{s} \rfloor_p$. We can sample an injective $\mathbf{A}$ with a trapdoor using the techniques of Ajtai [26] or subsequent improvements [27,28], and one can sample a lossy $\mathbf{A}$ using our lossy sampler. Although prior constructions of LTDFs based on LWE are known [23,29], our construction is extremely simple and has the advantage that our lossy mode loses "almost all" of the information contained in $\mathbf{s}$.

Deterministic Encryption. Deterministic public-key encryption [25,30–33] is intended to guarantee security as long as the messages have sufficient entropy. Although there are black-box constructions of deterministic encryption using LTDFs [32], we get a very simple direct construction from the LWR problem: the public key is a matrix $pk = \mathbf{A} \in \mathbb{Z}_q^{m \times n}$, and to encrypt a message $\mathbf{s} \in \{0,1\}^n$, we simply output $\lfloor \mathbf{A} \cdot \mathbf{s} \rfloor_p$. We can sample $\mathbf{A}$ with a decryption trapdoor using the standard techniques [26–28] mentioned previously. Our analysis here is essentially the same as for our reusable extractor – we simply note that whenever $\mathbf{s}$ has sufficient entropy, the output $\lfloor \mathbf{A} \cdot \mathbf{s} \rfloor_p$ is pseudorandom. We note that the same construction was proposed by Xie et al. [18], but because the analysis there was similar to [10,20], they required a super-polynomial modulus and modulus-to-error ratio. The main advantage of this scheme over other deterministic encryption schemes is that we do not need any fixed threshold on the entropy of the message $\mathbf{s}$: no matter how low it is we can still prove security under an LWE assumption with correspondingly degraded parameters.

1.3 Recent Concurrent Work

The work of [34] studies the security of LWE in the case where the error distribution is uniformly random over a small interval. In appendix B of the full version [1], we derive a very similar result. As a tool, both works rely on studying a "lossy mode" of LWE, but the construction and analysis are somewhat different. The work of [35] also studies LWE in a setting with extremely small errors uniform over $\{0,1\}$ also crucially using the notion of lossiness.

2 Preliminaries

Notation. Throughout, we let λ denote the *security parameter*. We use bold lower-case letters (e.g., $\mathbf{s}, \mathbf{e}$) to denote vectors, and bold upper-case letters (e.g., $\mathbf{A}, \mathbf{B}$) to denote matrices. If X is a distribution or a random variable, we write $x \xleftarrow{\$} X$ to denote the process of sampling x according to X. If X is a set, we write $x \xleftarrow{\$} X$ to denote the process of sampling x *uniformly* at random over X. For two distribution ensembles $X = \{X_\lambda\}, Y = \{Y_\lambda\}$, we write $X \overset{\text{comp}}{\approx} Y$ if for all probabilistic polynomial time (PPT) distinguishers D there is a negligible function $\mathsf{negl}(\cdot)$ such that: $|\Pr[D(1^\lambda, X_\lambda) = 1] - \Pr[D(1^\lambda, Y_\lambda)] = 1| \leq \mathsf{negl}(\lambda)$.

Bounded Distribution. A distribution χ over $\mathbb{R}$ is called β-*bounded* if $\mathbb{E}[|\chi|] \leq \beta$.

Probabilistic Notions. We assume that the reader is familiar with some basic notions from probability, such as statistical distance Δ, (conditional) min entropy, and the Chernoff bound. We will further rely on the following less standard definition of *smooth min-entropy*, which was first introduced by Renner and Wolf [36]. Intuitively, a random variable has high smooth min-entropy, if it is statistically close to a random variable with high min-entropy.

Definition 2.1 (Smooth Entropy). *We say that a random variable X has ε-*smooth min-entropy at least k*, denoted by $H^{\varepsilon}_{\infty}(X) \geq k$, if there exists some variable X' such that $\Delta(X, X') \leq \varepsilon$ and $H_{\infty}(X') \geq k$. Similarly, we say that the ε-*smooth conditional min-entropy of X given Y is at least k*, denoted $H^{\varepsilon}_{\infty}(X|Y) \geq k$ if there exist some variables (X', Y') such that $\Delta((X, Y), (X', Y')) \leq \varepsilon$ and $H_{\infty}(X'|Y') \geq k$.*

We will write $H^{\mathsf{smooth}}_{\infty}(\cdot)$ to denote $H^{\varepsilon}_{\infty}(\cdot)$ for some (unspecified) negligible ε.

2.1 Learning with Errors and Learning with Rounding

Learning With Errors. The *decisional learning with errors* (*LWE*) problem was first introduced by Regev [2]. Informally, the problem asks to distinguish slightly perturbed random linear equations from truly random ones.

Definition 2.2 (LWE Assumption [2]). *Let λ be the security parameter, $n = n(\lambda), m = m(\lambda), q = q(\lambda)$ be integers and let $\chi = \chi(\lambda)$ be a distribution over $\mathbb{Z}_q$. The* $\mathrm{LWE}_{n,m,q,\chi}$ *assumption says that for $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}$, $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$, $\mathbf{e} \leftarrow \chi^m$, $\mathbf{u} \xleftarrow{\$} \mathbb{Z}_q^m$ the following distributions are computationally indistinguishable:*

$$(\mathbf{A}, \mathbf{A} \cdot \mathbf{s} + \mathbf{e}) \overset{\text{comp}}{\approx} (\mathbf{A}, \mathbf{u}).$$

It has been shown that the LWE-assumption holds for certain error distributions χ, assuming the worst-case hardness of certain lattice problems. In particular, this is the case if χ is a discrete Gaussian distribution with appropriate variance, see, e.g., [2, 17, 35] for precise statements.

Learning With Rounding. The *learning with rounding* (*LWR*) problem was introduced by Banerjee et al. [10]. It can, in some sense, be seen as a derandomized version of the LWE-problem. The idea is to compute the error terms deterministically: instead of perturbing the answer by adding a small error, we simply round the answer – in both cases we are intuitively hiding the low order bits.

More formally, the LWR-problem is defined via the following *rounding function* for integers $q \geq p \geq 2$:

$$\lfloor \cdot \rfloor_p \ : \ \mathbb{Z}_q \to \mathbb{Z}_p \ : \ x \mapsto \lfloor (p/q) \cdot x \rfloor,$$

where we naturally identify elements of $\mathbb{Z}_k$ with the integers in the interval $\{0, \ldots, k-1\}$.[4] More intuitively, $\lfloor . \rfloor_p$ partitions $\mathbb{Z}_q$ into intervals of length $\approx \frac{q}{p}$ which it maps to the same image. We naturally extend the rounding function to vectors over $\mathbb{Z}_q$ by applying it component-wise.

In the presentation of our results we will make use that the probability that a random element in $\mathbb{Z}_q$ is close to a step in the rounding function is small. We therefore define, for any integer $\tau > 0$:

$$\mathsf{border}_{p,q}(\tau) \overset{\text{def}}{=} \{x \in \mathbb{Z}_q \ : \ \exists y \in \mathbb{Z}, |y| \leq \tau, \lfloor x \rfloor_p \neq \lfloor x + y \rfloor_p\}.$$

[4] The choice of the floor function rather than ceiling or nearest integer is arbitrary and unimportant.

We can easily bound the probability of a random element being on the border. As for the rest of this document, we omit a proof and refer to the full version [1].

Lemma 2.3. *For every p, q, τ it holds that* $\Pr_{x \xleftarrow{\$} \mathbb{Z}_q}[x \in \mathsf{border}_{p,q}(\tau)] \leq \frac{2\tau p}{q}$.

The learning with rounding problem is now defined as follows:

Definition 2.4 (LWR [10]). *Let λ be the security parameter, $n = n(\lambda), m = m(\lambda), q = q(\lambda), p = p(\lambda)$ be integers. The* $\mathrm{LWR}_{n,m,q,p}$ *problem states that for $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}$, $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$, $\mathbf{u} \xleftarrow{\$} \mathbb{Z}_q^m$ the following distributions are computationally indistinguishable:* $(\mathbf{A}, \lfloor \mathbf{A} \cdot \mathbf{s} \rfloor_p) \overset{\text{comp}}{\approx} (\mathbf{A}, \lfloor \mathbf{u} \rfloor_p)$.

Notice that when p divides q, the distribution $\lfloor u \rfloor_p : u \xleftarrow{\$} \mathbb{Z}_q$ is just the uniform over $\mathbb{Z}_p$. Otherwise, the distribution is slightly skewed with some values in $\mathbb{Z}_p$ having probability $\frac{\lfloor q/p \rfloor}{q}$ and others $\frac{\lceil q/p \rceil}{q}$. However, it is easy to deterministically extract random bits from such independent samples with an asymptotic rate of $O(\log(p))$ bits per sample. Therefore, independent samples from the skewed distribution are often "good enough" in practice.

We also define a variant of the LWR assumption where the secret $\mathbf{s}$ can come from some *weak source of entropy* and the attacker may observe some *partial leakage* about $\mathbf{s}$.

Definition 2.5 (LWR with Weak and Leaky Secrets). *Let λ be the security parameter and n, m, q, p be integer parameters as in Definition 2.4. Let $\gamma = \gamma(\lambda) \in (0, q/2)$ be an integer and $k = k(\lambda)$ be a real. The* $\mathrm{LWR}^{\mathsf{WL}(\gamma,k)}_{n,m,q,p}$ *problem says that for any efficiently samplable correlated random variables $(\mathbf{s}, \mathsf{aux})$, where the support of $\mathbf{s}$ is the integer interval $[-\gamma, \gamma]^n$ and $H_\infty(\mathbf{s}|\mathsf{aux}) \geq k$, the following distributions are computationally indistinguishable:*

$$(\mathsf{aux}, \mathbf{A}, \lfloor \mathbf{A} \cdot \mathbf{s} \rfloor_p) \overset{\text{comp}}{\approx} (\mathsf{aux}, \mathbf{A}, \lfloor \mathbf{u} \rfloor_p)$$

where $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}$, $\mathbf{u} \xleftarrow{\$} \mathbb{Z}_q^m$ are chosen randomly and independently of $\mathbf{s}, \mathsf{aux}$.

3 Lossy Mode for LWR

We now show that, under the LWE assumption, the LWR problem has a '*lossy mode*': we can sample a matrix $\tilde{\mathbf{A}}$ which is computationally indistinguishable from a uniformly random $\mathbf{A}$ such that the tuple $(\tilde{\mathbf{A}}, \lfloor \tilde{\mathbf{A}}\mathbf{s} \rfloor_p)$ does not reveal too much information about the secret $\mathbf{s}$.

Definition 3.1 (Lossy Sampler). *Let $\chi = \chi(\lambda)$ be an efficiently samplable distribution over $\mathbb{Z}_q$. The efficient* lossy sampler $\mathsf{Lossy}()$ *is given by:*

$\mathsf{Lossy}(1^n, 1^m, 1^\ell, q, \chi)$**:** *Sample $\mathbf{B} \xleftarrow{\$} \mathbb{Z}_q^{m \times \ell}, \mathbf{C} \xleftarrow{\$} \mathbb{Z}_q^{\ell \times n}, \mathbf{F} \xleftarrow{\$} \chi^{m \times n}$ and output $\tilde{\mathbf{A}} = \mathbf{B} \cdot \mathbf{C} + \mathbf{F}$.*

Although the matrix $\tilde{\mathbf{A}}$ computed by the Lossy algorithm is *statistically* far from a uniformly random matrix, it is easy to show that it is computationally indistinguishable from one under the $\mathrm{LWE}_{\ell,m,q,\chi}$ assumption, where the dimension of the secret is now ℓ instead of n. In particular, we can think of each column of $\mathbf{C}$ as an LWE secrets, the matrix $\mathbf{B}$ as the coefficients, and each column of $\tilde{\mathbf{A}}$ as the corresponding LWE output. Therefore, the following lemma from [20] follows by a simple hybrid argument.

Lemma 3.2 ([20]). *Let* $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m\times n}$*, and let* $\tilde{\mathbf{A}} \xleftarrow{\$} \mathsf{Lossy}(1^n, 1^m, 1^\ell, q, \chi)$*. Then, under the* $\mathrm{LWE}_{\ell,m,q,\chi}$ *assumption, the following two distributions are computationally indistinguishable:* $\mathbf{A} \overset{\text{comp}}{\approx} \tilde{\mathbf{A}}$.

The following lemma now states that for appropriate parameters, the secret $\mathbf{s}$ maintains high *smooth min-entropy* (see Definition 2.1) given $\tilde{\mathbf{A}}$ and $\lfloor \tilde{\mathbf{A}} \cdot \mathbf{s} \rfloor_p$.

Lemma 3.3. *Let* n, m, ℓ, p, γ *be positive integers,* χ *be some* β*-bounded distribution (i.e.,* $\mathbb{E}[|\chi|] \le \beta$*), and* $q \ge 2\beta\gamma nmp$ *be a prime. Then the following holds:*

(i) (Uniform Secret) For $\tilde{\mathbf{A}} \xleftarrow{\$} \mathsf{Lossy}(1^n, 1^m, 1^\ell, q, \chi)$, $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$ *we have, for* $\varepsilon = 2^{-\lambda} + q^{-\ell}$*:*

$$H_\infty^\varepsilon(\mathbf{s}|\tilde{\mathbf{A}}, \lfloor \tilde{\mathbf{A}}\mathbf{s} \rfloor_p) \ge n\log(2\gamma) - (\ell + \lambda)\log(q).$$

(ii) (High-Entropy Secret) Let $(\mathbf{s}, \mathsf{aux})$ *be correlated random variables with* $\mathbf{s} \in [-\gamma, \gamma]^n \subseteq \mathbb{Z}^n$*, and let* $\tilde{\mathbf{A}} \xleftarrow{\$} \mathsf{Lossy}(1^n, 1^m, 1^\ell, q, \chi)$ *be chosen independently. Then, for* $\varepsilon = 2^{-\lambda} + q^{-\ell}$ *and any* $\varepsilon' > 0$ *we have:*

$$H_\infty^{\varepsilon'+\varepsilon}(\mathbf{s}|\tilde{\mathbf{A}}, \lfloor \tilde{\mathbf{A}}\mathbf{s} \rfloor_p, \mathsf{aux}) \ge H_\infty^{\varepsilon'}(\mathbf{s}|\mathsf{aux}) - (\ell + \lambda)\log(q).$$

Both parts above also holds when q *is* not *prime, as long as the largest prime divisor of* q*, denoted* p_{max}*, satisfies* $GCD(q, q/p_{max}) = 1$*,* $p_{max} \ge 2\beta\gamma nmp$*. In this case we get* $\varepsilon = (2^{-\lambda} + (p_{max})^{-\ell} + \Pr[\mathbf{s} = 0^n \mod p_{max}])$.

The proof is sketched in Section 1.1, and a full proof is given in [1].

4 New "LWR from LWE" Reduction

In the following section we present the main result of this paper, namely sufficient conditions under which the LWR-assumption holds. As discussed earlier: on the positive side, we show that the LWR-assumption also holds if one drops a *small* fraction of the bits in the rounding function. On the negative side, the size of the modulus depends on the number of LWR-samples one needs to output, i.e., on the dimension of the matrix $\mathbf{A}$, and thus this number must be known in advance. However, as we will show in the subsequent sections, this is not a restriction for many interesting applications.

Theorem 4.1. *Let* k, ℓ, n, m, p, γ *be positive integers and* q *be a prime. Further, let* χ *be a* β*-bounded distribution for some* $\beta \in \mathbb{R}$ *(all parameters are functions of* λ*) such that* $q \ge 2\beta\gamma nmp$*. Under the* $\mathrm{LWE}_{\ell,m,q,\chi}$ *assumption we then get:*

(i) If $n \geq (\ell + \lambda + 1)\frac{\log(q)}{\log(2\gamma)} + 2\lambda$*, then the* $\mathrm{LWR}_{n,m,q,p}$*-assumption holds.*

(ii) If $k \geq (\ell + \lambda + 1)\log(q) + 2\lambda$*, then the weak and leaky* $\mathrm{LWR}^{\mathsf{WL}(\gamma,k)}_{n,m,q,p}$*-assumption holds.*

For exact security, if the above LWE assumption is (t, ε)*-secure and* $\ell \geq \lambda$*, then in both cases the corresponding* LWR*-problem is* (t', ε')*-secure, where* $t' = t - \mathsf{poly}(\lambda)$*,* $\varepsilon' = m(2 \cdot n\varepsilon + 3 \cdot 2^{-\lambda}) = \mathsf{poly}(\lambda)(\varepsilon + 2^{-\lambda})$*. Both parts of the above theorem also hold if* q *is not prime as long as the largest prime divisor of* q*, denoted* p_{max}*, satisfies* $GCD(q, q/p_{max}) = 1$*,* $p_{max} \geq 2\beta\gamma nmp$*. In this case we still get* $t' = t - \mathsf{poly}(\lambda)$*,* $\varepsilon' = \mathsf{poly}(\lambda)(\varepsilon + 2^{-\lambda})$*.*

The proof is sketched in Section 1.1, and a full proof can be found in [1].

Remark on β-bounded Distributions. In the theorem, we require that the distribution χ is β-bounded meaning that $\mathbb{E}[|\chi|] \leq \beta$. A different definition, which also would have been sufficient for us, would be to require that $\Pr_{x \xleftarrow{\$} \chi}[|x| > \beta] \leq \mathsf{negl}(\lambda)$. The latter notion of boundedness is used in the work of Banerjee et al. [10]. Although the two notions are technically incomparable (one does not imply the other) for natural distributions, such as the discrete Gaussian, it is easier to satisfy out notion. In particular, the discrete Gaussian distribution Ψ_σ with standard deviation σ satisfies $\mathbb{E}[|\Psi_\sigma|] \leq \sigma$ but we can only get the weaker bound $\Pr_{x \xleftarrow{\$} \Psi_\sigma}[|x| > \sqrt{\omega(\log(\lambda))}\sigma] \leq \mathsf{negl}(\lambda)$. Therefore, we find it advantageous to work with our definition.

Remark on Parameters. Notice that in the above theorem, the parameter γ offers a tradeoff between the size of the modulus q and the secret vector length n: for a bigger γ we need a bigger modulus q but can allow smaller secret length n. The following corollary summarizes two extreme cases of small and large γ.

Corollary 4.2. *Let* Ψ_σ *denote a discrete Gaussian distribution over* $\mathbb{Z}_q$ *with standard deviation* σ*, and assume that the* $\mathrm{LWE}_{\ell,m,q,\Psi_\sigma}$*-assumption holds. Then the* $\mathrm{LWR}_{n,m,q,p}$*-assumption holds in either of the following cases:*

- *(Minimize Modulus/Error Ratio.) If* $q \geq 2\sigma nmp$ *is a prime, and* $n \geq (\ell + \lambda + 1)\log(q) + 2\lambda$*. By setting* $p = O(1)$*, we can get a modulus-to-error ratio as small as* $q/\sigma = O(m \cdot n)$*.*
- *(Maximize Efficiency.) If* $q \geq (2\sigma nm)^3$ *is a prime,* $p = \sqrt[3]{q}$ *and* $n \geq 3\ell + 5\lambda + 3$*. The efficiency of LWR is now similar to the LWE assumption with* $n = O(\ell)$ *and* $\log(p) = O(\log q)$*.*

5 Reusable Extractors

The notion of a 'computational reusable extractor' was defined by Dodis et al. [21]. Intuitively, this is a tool that allows us to take some weak secret $\mathbf{s}$ that has a sufficient amount of entropy, and to use it to repeatedly extract fresh pseudorandomness $\mathsf{Ext}(\mathbf{s}; \mathbf{a}_i)$ using multiple public random seeds $\mathbf{a}_i$. Each extracted

output should look random and independent.[5] The work of [21] constructed such reusable extractors under a new assumption called "Learning Subspaces with Noise (LSN)". Reusable extractors were also implicitly constructed based on the DDH assumption in the work of Naor and Segev [37].[6] Here we give a new construction based on the LWR problem, with a security reduction from the LWE assumption.

Definition 5.1 (Reusable Extractor). *Let $\mathcal{S}, \mathcal{D}, \mathcal{U}$ be some domains, parametrized by the security parameter λ. A function $\mathsf{Ext} : \mathcal{S} \times \mathcal{D} \to \mathcal{U}$ is a (k, m)-reusable-extractor if for any efficiently samplable correlated random variables $\mathbf{s}, \mathsf{aux}$ such that the support of $\mathbf{s}$ is $\mathcal{S}$ and $H_\infty(\mathbf{s}|\mathsf{aux}) \geq k$, we have:*

$$(\mathsf{aux}, \mathbf{a}_1, \ldots, \mathbf{a}_m, \mathsf{Ext}(\mathbf{s}; \mathbf{a}_1), \ldots, \mathsf{Ext}(\mathbf{s}; \mathbf{a}_m)) \overset{\text{comp}}{\approx} (\mathsf{aux}, \mathbf{a}_1, \ldots, \mathbf{a}_m, u_1, \ldots, u_m)$$

where the values $\{\mathbf{a}_j \xleftarrow{\$} \mathcal{D}\}, \{u_j \xleftarrow{\$} \mathcal{U}\}$ are sampled independently.

Theorem 5.2. *Let n, p, γ be integers, p' be a prime, and define $q = p \cdot p'$. Then, assuming that the $\mathrm{LWE}_{\ell,m,q,\chi}$ assumption holds for some β-bounded distribution χ such that $p' > 2\beta\gamma nmp$ and $k \geq (\ell + \lambda + 1)\log(q) + 2\lambda$, the function*

$$\mathsf{Ext} : [-\gamma, \gamma]^n \times \mathbb{Z}_q^n \to \mathbb{Z}_p \quad \textit{defined by} \quad \mathsf{Ext}(\mathbf{s}; \mathbf{a}) \overset{\text{def}}{=} \lfloor \langle \mathbf{a}, \mathbf{s} \rangle \rfloor_p$$

is a (k, m)-reusable extractor.

Notice that one nice property of the above reusable extractor is that it has a *graceful degradation of security* as the min-entropy k of the source drops. In particular, there is no hard threshold on the entropy k determined by the parameters that define the scheme: γ, n, q, p. Instead, as the entropy k drops we can still reduce security from a correspondingly less secure LWE assumption with smaller secret size ℓ. In other words, the scheme designer does not need to know the actual entropy k of the secret - but the scheme gets gradually less/more secure as the entropy of the secret shrinks/grows. A similar notion of graceful security degradation was noted in the work of Goldwasser et al. [20].

6 Lossy Trapdoor Functions

Lossy trapdoor functions (LTDFs) [22,23], are a family of functions $f_{pk}(\cdot)$ keyed by some public key pk, which can be sampled in one of two indistinguishable modes: `injective` and `lossy`. In the `injective` mode the function $f_{pk}(\cdot)$ is injective and we can even sample pk along with a secret trapdoor key sk that allows us to invert it efficiently. In the `lossy` mode, the function $f_{pk}(\cdot)$ is "many-to-one" and $f_{pk}(\mathbf{s})$ statistically loses information about the input $\mathbf{s}$. LTDFs have many interesting applications in cryptography, such as allowing us to output

[5] Equivalently, we can think of a reusable extractor as a weak PRF $f_{\mathbf{s}}(\cdot)$ for which security holds for a bounded number of inputs even using a high entropy key $\mathbf{s}$.

[6] The function $\mathsf{Ext}(\mathbf{s}; \mathbf{a}) = \prod \mathbf{a}_i^{\mathbf{s}_i}$ is a reusable extractor if $\mathbf{s} \in \mathbb{Z}_q^n$, and the $\mathbf{a} \in \mathbb{G}^n$ for some DDH group of prime order q.

many hardcore bits, construct CCA-2 public-key encryption [23, 24], and deterministic encryption [25]. In this section, we construct very simple and efficient LTDFs using the LWR problem, with security based on standard LWE. Our LTDF function is unusually simple: the public key is a matrix $pk = \mathbf{A}$ and the function is defined as $f_{\mathbf{A}}(\mathbf{s}) = \lfloor \mathbf{A} \cdot \mathbf{s} \rceil_p$. As we will describe, one can sample an injective $\mathbf{A}$ with a trapdoor using the techniques of Ajtai [26] or subsequent improvements [27,28], and one can sample a lossy $\mathbf{A}$ using the techniques we developed in Section 3. Although prior constructions of LTDFs from LWE are known [23, 29], our construction here has several advantages. Firstly, our scheme is extremely simple to describe and implement. Secondly, in contrast to both [22, 29], our lossy mode loses "almost all" of the information contained in $\mathbf{s}$. In fact, the amount of "lossiness" in our LTDF construction is flexible and not determined by the parameters of the scheme itself. Even after we fix the parameters that allow us to sample the injective mode, we have an additional free parameter that allows us to make the lossy mode progressively more lossy under under a progressively stronger variant of the LWE assumption.

6.1 Entropic LTDFs

Our notion differs somewhat from that of [23] in how we define the "lossy" property. Instead of requiring that, for a lossy pk, the range of $f_{pk}(\cdot)$ is small, we require that very little *entropy* is lost from observing $f_{pk}(\cdot)$. To the best of our knowledge, our version can be used interchangeably in all of the applications of LTDFs to date. To avoid confusion, we call our notion *entropic* LTDF (eLTDF).

Definition 6.1 (eLTDF). *A family of* $l(\lambda)$-entropic lossy trapdoor functions (eLTDF) *with security parameter* λ *and domain* $\mathcal{D}_\lambda$ *consists of a* PPT *sampling algorithms* Gen *and two deterministic* PPT *algorithms* F, F^{-1} *such that:*

Injective Functions: *For any* (pk, sk) *in the support of* $\mathsf{Gen}(1^\lambda, \mathtt{injective})$*, any* $\mathbf{s} \in \mathcal{D}_\lambda$ *we require that* $F^{-1}(sk, F(pk, \mathbf{s})) = \mathbf{s}$.

Lossy Functions: *When* $pk \xleftarrow{\$} \mathsf{Gen}(1^\lambda, \mathtt{lossy})$*, the function* $F(pk, \cdot)$ *is lossy. In particular, for any mutually correlated random variables* $(\mathbf{s}, \mathsf{aux})$ *where the domain of* $\mathbf{s}$ *is* $\mathcal{D}_\lambda$ *and for an independently sampled* $pk \xleftarrow{\$} \mathsf{Gen}(1^\lambda, \mathtt{lossy})$*, we have:* $H_\infty^{\mathsf{smooth}}(\mathbf{s}|pk, F(pk, \mathbf{s}), \mathsf{aux}) \geq H_\infty^{\mathsf{smooth}}(\mathbf{s}|\mathsf{aux}) - l(\lambda)$*. We call the parameter* $l = l(\lambda)$ *the* residual leakage *of the LTDF.*

Indistinguishability: *The distributions of* pk *as sampled by* $\mathsf{Gen}(1^\lambda, \mathtt{lossy})$ *and* $\mathsf{Gen}(1^\lambda, \mathtt{injective})$ *are computationally indistinguishable.*

We now show how to construct eLTDFs from LWR (and so also from LWE).

Tools. As a tool in our construction, we will rely on the fact that we can sample a random LWE matrix $\mathbf{A}$ along with an *inversion* trapdoor that allows us to recover $\mathbf{s}, \mathbf{e}$ given an LWE sample $\mathbf{As} + \mathbf{e}$ where the error $\mathbf{e}$ is "sufficiently" short. The first example of such algorithms was given by Ajtai in [26], and was subsequently improved in [27]. More recently [28] significantly improved the efficiency of these results, by using a "qualitatively" different type of trapdoor.

We describe the properties that we need abstractly, and can use any of the above algorithms in a black-box manner. In particular we need the following PPT algorithms for some range of parameters (m, n, q, β):

$\mathsf{GenTrap}(1^n, 1^m, q)$**:** An algorithm which on input positive integers n, q and sufficiently large m samples a matrix $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ and trapdoor T such that $\mathbf{A}$ is statistically close to uniform (in $n \log q$).

$\mathsf{Invert}(T, \mathbf{A}, \mathbf{c})$**:** An algorithm which receives as input $(\mathbf{A}, T)$ in the support of $\mathsf{GenTrap}(1^n, 1^m, q)$ and some value $\mathbf{c} \in \mathbb{Z}_q^m$ such that $\mathbf{c} = \mathbf{As} + \mathbf{e}$ for some $\mathbf{s} \in \mathbb{Z}_q^n$ and some error satisfying $||\mathbf{e}||_2 \leq \beta$. The algorithm outputs $\mathbf{s}$.

$\mathsf{LWRInvert}(T, \mathbf{A}, \mathbf{c})$ Takes as input $(\mathbf{A}, T)$ in the support of $\mathsf{GenTrap}(1^n, 1^m, q)$ and some value $\mathbf{c} \in \mathbb{Z}_p^m$ such that $\mathbf{c} = \lfloor \mathbf{As} \rfloor_p$ for some $\mathbf{s} \in \mathbb{Z}_q^n$ and outputs $\mathbf{s}$.

For example [28] shows that there are algorithms $(\mathsf{GenTrap}, \mathsf{Invert})$ which work for $n \geq 1$, $q \geq 2$, sufficiently large $m = O(n \log q)$ and sufficiently small $\beta < q/O(\sqrt{n \log q})$. Since we can convert LWR samples $\lfloor \mathbf{As} \rfloor_p$ into samples $\mathbf{A} \cdot \mathbf{s} + \mathbf{e}$ for some short error $||\mathbf{e}||_2 \leq \sqrt{m}q/p$, this also implies the following.

Lemma 6.2 (Trapdoors for LWR). *For $n \geq 1$, $q \geq 2$, sufficiently large $m \geq O(n \log q)$ and $p \geq O(\sqrt{mn \log q})$, there exist* $(\mathsf{GenTrap}, \mathsf{LWRInvert})$ *as above.*

The Construction. We will rely on the algorithms $\mathsf{GenTrap}$ and $\mathsf{LWRInvert}$ described above. We also rely on the lossy sampling algorithm Lossy and its properties developed in Section 3. The construction is parametrized by integers n, m, q, p (all functions of the security parameter λ). Furthermore, there will be two additional parameters ℓ and χ which are only needed by the lossy sampler.

$\mathsf{Gen}(1^\lambda, \mathtt{injective})$**:** Sample $(\mathbf{A}, T) \xleftarrow{\$} \mathsf{GenTrap}(1^n, 1^m, q)$. Output $pk = \mathbf{A}$ and trapdoor $sk = (\mathbf{A}, T)$.

$\mathsf{Gen}(1^\lambda, \mathtt{lossy})$**:** Sample $\mathbf{A} \xleftarrow{\$} \mathsf{Lossy}(1^n, 1^m, 1^\ell, q, \chi)$. Output $pk = \mathbf{A}$.

$F(pk, \mathbf{s})$**:** On input $\mathbf{s} \in \{0,1\}^n$ and matrix $pk = \mathbf{A} \in \mathbb{Z}_q^{m \times n}$ output $\lfloor \mathbf{As} \rfloor_p$.

$F^{-1}(sk, \mathbf{c})$**:** On input $\mathbf{c} \in \mathbb{Z}_p^m$ and $sk = (\mathbf{A}, T)$ output $\mathsf{LWRInvert}(T, \mathbf{A}, \mathbf{c})$.

The following theorem summarizes the properties of this construction.

Theorem 6.3. *Let χ be an efficiently samplable β-bounded distribution and λ be the security parameter. For any positive integers $n \geq \lambda$, sufficiently large $m \geq O(n \log q)$, $p \geq O(\sqrt{mn \log q})$ and a prime $q \geq 2\beta nmp$, if the* $\mathrm{LWE}_{\ell,m,q,\chi}$ *assumption holds then the above construction is an l-LTDF with $l = (\ell + \lambda) \log q$.*

We refer the interested reader to the full version [1], where we additionally show how to construct efficient *all-but-one trapdoor functions*, and how to obtain CCA-2 secure encryption schemes therefrom.

7 Deterministic Encryption

Deterministic public-key encryption [25,30–33] is intended to guarantee security as long as the messages have sufficient entropy. Although there are black-box

constructions of deterministic encryption using LTDFs [32], here we present a very simple direct construction from the LWR problem. There are several definitions of deterministic encryption which can be proven equivalent; see [31, 32]. Here, we will use one such simple definition based on indistinguishability of encrypting messages from two different distributions.

Definition 7.1 (Deterministic Encryption). *A triple of* PPT *algorithms* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$*, where* $\mathsf{Enc}, \mathsf{Dec}$ *are deterministic, is a* deterministic encryption scheme *with message length* $n = n(\lambda)$*, if it satisfies the following properties. First, it is* correct*, i.e., for all* $(pk, sk) \xleftarrow{\$} \mathsf{Gen}(1^\lambda)$ *and all messages* $\mathbf{s} \in \{0,1\}^n$*, we have* $\mathsf{Dec}_{sk}(\mathsf{Enc}_{pk}(\mathbf{s})) = \mathbf{s}$*. We further say that the scheme is* secure *for all* $k(\lambda)$*-sources if for any two distribution ensembles* $\{S_\lambda^{(0)}\}_{\lambda\in\mathbb{N}}, \{S_\lambda^{(1)}\}_{\lambda\in\mathbb{N}}$ *over* $\{0,1\}^{n(\lambda)}$ *which are efficiently samplable in* $\mathsf{poly}(\lambda)$*-times and have sufficient entropy* $H_\infty(S_\lambda^0) \geq k$*,* $H_\infty(S_\lambda^1) \geq k$*, we have* $(pk, \mathsf{Enc}_{pk}(\mathbf{s}_0)) \overset{\text{comp}}{\approx} (pk, \mathsf{Enc}_{pk}(\mathbf{s}_1))$*, where* $\mathbf{s}_0 \xleftarrow{\$} S_\lambda^{(0)}$ *and* $\mathbf{s}_1 \xleftarrow{\$} S_\lambda^{(1)}$ *and* $(pk, sk) \xleftarrow{\$} \mathsf{Gen}(1^\lambda)$*.*

Construction. We give a very simple construction of deterministic encryption based on the LWR assumption. This construction is the same as one given by Xie et al. [18], except for the setting of parameters. Whereas they required a super-polynomial modulus and modulus to error ratio by relying on variants of the analysis of [10,20] we use our improved analysis from Section 4. We will rely on the LWR trapdoor generation and inversion algorithms $\mathsf{GenTrap}, \mathsf{LWRInvert}$ described in Section 6.1 and Lemma 6.2. Our scheme is parametrized by some n, m, q, p, all functions of the security parameter λ, and has message length n.

$\mathsf{Gen}(1^\lambda)$**:** Choose $(\mathbf{A}, T) \xleftarrow{\$} \mathsf{GenTrap}(1^n, 1^m, q)$. Output $pk = \mathbf{A}$, $sk = T$.
$\mathsf{Enc}_{pk}(\mathbf{s})$**:** For a message $\mathbf{s} \in \{0,1\}^n$, output $\lfloor \mathbf{A} \cdot \mathbf{s} \rfloor_p$.
$\mathsf{Dec}_{sk}(\mathbf{c})$**:** For a ciphertext $\mathbf{c} \in \mathbb{Z}_p^m$, output $\mathsf{LWRInvert}(T, \mathbf{A}, \mathbf{c})$.

Theorem 7.2. *Let* λ *be the security parameter,* $n \geq \lambda, \ell, m, p$ *be an integers,* q *be a prime, and* χ *be an efficiently samplable* β*-bounded distribution (all parameters are functions of* λ*) such that* $m \geq O(n \log q)$*,* $p \geq O(\sqrt{mn \log q})$ *are sufficiently large and* $q \geq 2\beta nmp$*. If the* $\mathrm{LWE}_{\ell,m,q,\chi}$ *assumption holds then the above construction with parameters* n, m, q, p *is a deterministic encryptions secure for all* k *sources where* $k \geq (\ell + \Omega(\lambda)) \log(q)$*.*

One big advantage of our scheme is that the parameters n, m, q, p do not determine the minimal entropy k. Instead for any k, we can prove security under a corresponding LWE assumption with dimension $\ell < k$.

8 Open Problems

We conclude with two interesting open problems. Firstly, is it possible to improve the reduction and remove the dependence between the modulus q and the number of samples m? And secondly, is there a related reduction for *Ring LWR* from *Ring LWE*? This does not seem to follow in a straight-forward manner.

References

1. Alwen, J., Krenn, S., Pietrzak, K., Wichs, D.: Learning with Rounding, Revisited: New Reduction, Properties and Applications. Cryptology ePrint Archive, Report 2013/098 (2013)
2. Regev, O.: On Lattices, Learning with Errors, Random Linear Codes, and Cryptography. In: Gabow, H.N., Fagin, R. (eds.) STOC, pp. 84–93. ACM (2005)
3. Gentry, C., Peikert, C., Vaikuntanathan, V.: Trapdoors for Hard Lattices and New Cryptographic Constructions. In: Dwork, C. (ed.) STOC, pp. 197–206. ACM (2008)
4. Akavia, A., Goldwasser, S., Vaikuntanathan, V.: Simultaneous Hardcore Bits and Cryptography against Memory Attacks. In: Reingold, O. (ed.) TCC 2009. LNCS, vol. 5444, pp. 474–495. Springer, Heidelberg (2009)
5. Lyubashevsky, V., Peikert, C., Regev, O.: On Ideal Lattices and Learning with Errors over Rings. In: Gilbert, H. (ed.) EUROCRYPT 2010. LNCS, vol. 6110, pp. 1–23. Springer, Heidelberg (2010)
6. Agrawal, S., Freeman, D.M., Vaikuntanathan, V.: Functional Encryption for Inner Product Predicates from Learning with Errors. In: Lee, D.H., Wang, X. (eds.) ASIACRYPT 2011. LNCS, vol. 7073, pp. 21–40. Springer, Heidelberg (2011)
7. Brakerski, Z., Vaikuntanathan, V.: Fully Homomorphic Encryption from Ring-LWE and Security for Key Dependent Messages. In: Rogaway, P. (ed.) CRYPTO 2011. LNCS, vol. 6841, pp. 505–524. Springer, Heidelberg (2011)
8. Lindner, R., Peikert, C.: Better Key Sizes (and Attacks) for LWE-Based Encryption. In: Kiayias, A. (ed.) CT-RSA 2011. LNCS, vol. 6558, pp. 319–339. Springer, Heidelberg (2011)
9. Goldwasser, S., Kalai, Y., Popa, R.A., Vaikuntanathan, V., Zeldovich, N.: Succinct Functional Encryption and Applications: Reusable Garbled Circuits and Beyond. Cryptology ePrint Archive, Report 2012/733 (2012)
10. Banerjee, A., Peikert, C., Rosen, A.: Pseudorandom Functions and Lattices. In: Pointcheval, D., Johansson, T. (eds.) EUROCRYPT 2012. LNCS, vol. 7237, pp. 719–737. Springer, Heidelberg (2012)
11. Lyubashevsky, V.: Fiat-Shamir with Aborts: Applications to Lattice and Factoring-Based Signatures. In: Matsui, M. (ed.) ASIACRYPT 2009. LNCS, vol. 5912, pp. 598–616. Springer, Heidelberg (2009)
12. Rückert, M.: Lattice-Based Blind Signatures. In: Abe, M. (ed.) ASIACRYPT 2010. LNCS, vol. 6477, pp. 413–430. Springer, Heidelberg (2010)
13. Lyubashevsky, V.: Lattice Signatures without Trapdoors. In: Pointcheval, D., Johansson, T. (eds.) EUROCRYPT 2012. LNCS, vol. 7237, pp. 738–755. Springer, Heidelberg (2012)
14. Katz, J., Vaikuntanathan, V.: Smooth Projective Hashing and Password-Based Authenticated Key Exchange from Lattices. In: Matsui, M. (ed.) ASIACRYPT 2009. LNCS, vol. 5912, pp. 636–652. Springer, Heidelberg (2009)
15. Peikert, C., Rosen, A.: Efficient Collision-Resistant Hashing from Worst-Case Assumptions on Cyclic Lattices. In: Halevi, S., Rabin, T. (eds.) TCC 2006. LNCS, vol. 3876, pp. 145–166. Springer, Heidelberg (2006)
16. Peikert, C., Vaikuntanathan, V., Waters, B.: A Framework for Efficient and Composable Oblivious Transfer. In: Wagner, D. (ed.) CRYPTO 2008. LNCS, vol. 5157, pp. 554–571. Springer, Heidelberg (2008)
17. Peikert, C.: Public-Key Cryptosystems from the Worst-Case Shortest Vector Problem: Extended Abstract. In: Mitzenmacher, M. (ed.) STOC, pp. 333–342. ACM (2009)
18. Xie, X., Xue, R., Zhang, R.: Deterministic Public Key Encryption and Identity-Based Encryption from Lattices in the Auxiliary-Input Setting. In: Visconti, I., De Prisco, R. (eds.) SCN 2012. LNCS, vol. 7485, pp. 1–18. Springer, Heidelberg (2012)

19. Brakersi, Z., Langlois, A., Peikert, C., Regev, O., Stehl, D.: Classical Hardness of Learning with Errors. In: STOC (2013)
20. Goldwasser, S., Kalai, Y.T., Peikert, C., Vaikuntanathan, V.: Robustness of the Learning with Errors Assumption. In: Yao, A.C.C. (ed.) ICS, pp. 230–240. Tsinghua University Press (2010)
21. Dodis, Y., Kalai, Y.T., Lovett, S.: On Cryptography with Auxiliary Input. In: Mitzenmacher, M. (ed.) STOC, pp. 621–630. ACM (2009)
22. Peikert, C., Waters, B.: Lossy Trapdoor Functions and Their Applications. In: Dwork, C. (ed.) STOC, pp. 187–196. ACM (2008)
23. Peikert, C., Waters, B.: Lossy Trapdoor Functions and Their Applications. SIAM J. Comput. 40(6), 1803–1844 (2011)
24. Mol, P., Yilek, S.: Chosen-Ciphertext Security from Slightly Lossy Trapdoor Functions. In: Nguyen, P.Q., Pointcheval, D. (eds.) PKC 2010. LNCS, vol. 6056, pp. 296–311. Springer, Heidelberg (2010)
25. Fuller, B., O'Neill, A., Reyzin, L.: A Unified Approach to Deterministic Encryption: New Constructions and a Connection to Computational Entropy. In: Cramer, R. (ed.) TCC 2012. LNCS, vol. 7194, pp. 582–599. Springer, Heidelberg (2012)
26. Ajtai, M.: Generating Hard Instances of the Short Basis Problem. In: Wiedermann, J., Van Emde Boas, P., Nielsen, M. (eds.) ICALP 1999. LNCS, vol. 1644, pp. 1–9. Springer, Heidelberg (1999)
27. Alwen, J., Peikert, C.: Generating Shorter Bases for Hard Random Lattices. Theory Comput. Syst. 48(3), 535–553 (2011)
28. Micciancio, D., Peikert, C.: Trapdoors for Lattices: Simpler, Tighter, Faster, Smaller. In: Pointcheval, D., Johansson, T. (eds.) EUROCRYPT 2012. LNCS, vol. 7237, pp. 700–718. Springer, Heidelberg (2012)
29. Bellare, M., Kiltz, E., Peikert, C., Waters, B.: Identity-Based (Lossy) Trapdoor Functions and Applications. In: Pointcheval, D., Johansson, T. (eds.) EUROCRYPT 2012. LNCS, vol. 7237, pp. 228–245. Springer, Heidelberg (2012)
30. Bellare, M., Boldyreva, A., O'Neill, A.: Deterministic and Efficiently Searchable Encryption. In: Menezes, A. (ed.) CRYPTO 2007. LNCS, vol. 4622, pp. 535–552. Springer, Heidelberg (2007)
31. Bellare, M., Fischlin, M., O'Neill, A., Ristenpart, T.: Deterministic Encryption: Definitional Equivalences and Constructions without Random Oracles. In: Wagner, D. (ed.) CRYPTO 2008. LNCS, vol. 5157, pp. 360–378. Springer, Heidelberg (2008)
32. Boldyreva, A., Fehr, S., O'Neill, A.: On Notions of Security for Deterministic Encryption, and Efficient Constructions without Random Oracles. In: Wagner, D. (ed.) CRYPTO 2008. LNCS, vol. 5157, pp. 335–359. Springer, Heidelberg (2008)
33. Brakerski, Z., Segev, G.: Better Security for Deterministic Public-Key Encryption: The Auxiliary-Input Setting. In: Rogaway, P. (ed.) CRYPTO 2011. LNCS, vol. 6841, pp. 543–560. Springer, Heidelberg (2011)
34. Döttling, N., Müller-Quade, J.: Lossy Codes and a New Variant of the Learning-With-Errors Problem. In: Johansson, T., Nguyen, P.Q. (eds.) EUROCRYPT 2013. LNCS, vol. 7881, pp. 18–34. Springer, Heidelberg (2013)
35. Micciancio, D., Peikert, C.: Hardness of SIS and LWE with Small Parameters. In: Canetti, R., Garay, J.A. (eds.) CRYPTO 2013, Part I. LNCS, vol. 8042, pp. 21–39. Springer, Heidelberg (2013)
36. Renner, R., Wolf, S.: Smooth Rényi Entropy and Applications. In: ISIT, vol. 4, p. 233 (2004)
37. Naor, M., Segev, G.: Public-Key Cryptosystems Resilient to Key Leakage. In: Halevi, S. (ed.) CRYPTO 2009. LNCS, vol. 5677, pp. 18–35. Springer, Heidelberg (2009)