Hardness of Learning with Physical Rounding and Noise from Learning With Errors

Clément Hoffmann[1*], Emeline Repel[2],
Adeline Roux-Langlois[2] and François-Xavier Standaert[3]

[1] NTT Social Informatics Laboratories, Japan
[2] Université Caen Normandie, ENSICAEN, CNRS, Normandie University, GREYC UMR6072, F-14000 Caen, France
[3] Crypto Group, ICTEAM Institute, UCLouvain, Louvain-la-Neuve, Belgium

Abstract. The Learning With Physical Rounding (LWPR) problem is a variant of the Learning With Rounding (LWR) problem, where the rounding operation is performed by a physical leakage function (like the Hamming weight function or variants thereof). It has been introduced as a potentially useful building block towards designing cryptographic algorithms with affordable leakage-resilience guarantees. LWPR has direct applications for fresh re-keying schemes in symmetric cryptography, and potential applications for post-quantum encryption and signature schemes relying on hard learning problems. However, for now its hardness has only been evaluated heuristically, thanks to algebraic cryptanalysis. In this paper, we show that a natural generalization of the LWPR problem with noisy samples, which we denote as Learning With Physical Rounding and Noise (LWPRN), is as hard as the Learning With Errors (LWE) problem. This generalization is motivated by the fact that concrete leakage functions are usually affected by a mild level of physical noise. It targets the same applications as LWPR, yet with much stronger confidence given its reduction from a standard mathematical assumption. By combining proof techniques from lattice-based cryptography and statistical concepts used in the side-channel / masking literature, we argue that the level of noise required in our reduction is concretely affordable. Besides, we show that the generalization from LWPR to LWPRN is not only motivated by the goal of obtaining provable security guarantees, since our reduction applies to the serial implementation setting, in which LWPR has been shown insecure.

Keywords: Side-Channel Analysis · Countermeasures · Masking · Fresh Re-Keying · Lattices · Post-Quantum Cryptography · Hard Learning Problems

1 Introduction

State of the art. Protecting cryptographic implementations against leakage is known to be a challenging problem. In the context of symmetric primitives like block ciphers, the standard approach is to mask (i.e., secret-share) all the computations [ISW03]. Positively, this can lead to exponential (side-channel) security improvements as a function of a well-controlled security parameter (the number of shares) [PR13, DDF14, DFS15]. Negatively, masked cryptographic computations are expensive and tricky to implement, due to physical imperfections like glitches in hardware [MPG05, NRS11] or transitions in software [CGP+12, BGG+14]. Fresh re-keying schemes, where an ephemeral key is produced by multiplying a long-term key with a public random Initialization Vector (IV), have been introduced as a way to mitigate these overheads and difficulties [MSGR10, DKM+15]. They aim to

* Work performed in part as PhD student at UCLouvain.

leverage a good separation of duties between a re-keying function that must be secure against Differential Power Analysis (DPA) but is easy to mask (since key-homomorphic) and a block cipher that is only executed with ephemeral secrets, and therefore only needs to resist Simple Power Analysis (SPA).[1] The first instances of fresh re-keying schemes were only evaluated heuristically [BFG14, BCF+15, PM16, GJ19]. Dziembowski et al. then proposed the concept of "hard physical learning problem" to analyze an instance of a fresh re-keying scheme that relies on a well-known hardness assumption [DFH+16]. More precisely, they introduced a fresh re-keying scheme whose security relies on the Learning Parity with Leakage (LPL) assumption: a variant of the standard Learning Parity with Noise (LPN) assumption, where the modular addition of a Bernoulli noise is replaced by the addition of independent Gaussian noise. This mimics the situation where an adversary can observe the leakage of an ephemeral key, but not the key itself. The LPL problem has been shown to be as hard as the LPN problem given sufficiently noisy leakages.

One important limitation of re-keying schemes based on the LPL assumption is their noise requirements. This can be explained by the strong "algebraic compatibility" between computations in $\mathbb{F}_2$ and actual leakage functions observed in practice, like the Hamming weight function [MOP07]. For example, the least significant bit of the Hamming weight of a value is the XOR of its bits. Hence, without noise, observing LPL samples enables building a system of linear equations in the key bits. Besides, and as shown in [DFH+16], the amount of noise needed to provably (though possibly not tightly) hide these linear dependencies can be substantial. Another limitation is that LPL operates in $\mathbb{F}_2$, which limits its applicability for cryptographic schemes operating in larger fields. The Learning With Physical Rounding (LWPR) assumption was introduced in response to these limitations [DMMS21]. It operates in larger fields and can be viewed as a variant of the standard Learning With Rounding (LWR) assumption [BPR12], where the rounding is realized by a physical leakage function, like the popular Hamming weight one [MOP07]. Its security intuitively stems from the fact that the aforementioned linear attacks are prevented when operating in prime fields. Yet, for now, it has only been evaluated heuristically. In particular, Hoffmann et al. analyzed the LWPR assumption for linear and quadratic leakage functions, and showed that it is hard to falsify with state-of-the-art algebraic cryptanalysis techniques [HMM+23].

While reassuring, such cryptanalysis results nevertheless leave us far from the confidence one can have in standard (mathematical) hard learning problems. Ideally, one would like to show that LWPR is as hard as LWR or LWE (Learning With Errors [Reg05]), just like LPL is as hard as LPN. Yet, such a direct reduction seems difficult to obtain. For example, the physically-constrained and deterministic nature of LWPR implies that noise cannot be leveraged in order to "hide" differences between distributions, as conveniently used in the reduction from LPN to LPL. Hence, as a practically relevant first step towards showing the hardness of LWPR, we next study a noisy variant of LWPR that we denote as the Learning With Physical Rounding and Noise (LWPRN) problem. On the one hand, it is justified by the fact that concrete leakages are usually affected by a mild level of physical noise. On the other hand, it comes with the hope that (i) such a problem can be connected to a standard hardness assumption, and (ii) the concrete level of physical noise needed for this purpose can be kept reasonable when combined with a "compressive" physical rounding function, like the Hamming weight one that we will consider in this work.

LWE with additional information. Both from a cryptanalysis and from a theoretical (reduction) viewpoint, the Learning With Error (LWE) problem[2] has been analyzed in the

[1] The DPA adversary can continuously accumulate information on the long-term key by monitoring the leakage corresponding to multiple IVs. The SPA adversary can only collect a bounded amount of information on each ephemeral key, because they are only used to encrypt a few message blocks.

[2] Let $\oplus$ denotes the addition in $\mathbb{Z}_q$, the LWE problem gives samples $(\mathbf{A}, \mathbf{A}\mathbf{s} \oplus \mathbf{e})$ where $\mathbf{A} \xleftarrow{\$} \mathbf{Z}_q^{m \times n}$, $\mathbf{s} \xleftarrow{\$} \mathbf{Z}_q^n$ and $\mathbf{e} \xleftarrow{\$} \chi^m$, and asks to find the secret or to distinguish such samples from a uniform ones.

presence of various types of additional information – See Table 1. It is therefore informative to first browse such related works and highlight differences with ours. For completeness, we also give more technical details on these previous works in Appendix A.

Table 1: Comparison with prior works based on the additional information considered.

Work	Variant	Type of Result	Additional info.
[DDGR20]	LWE with side info.	Cryptanalysis	$\langle \mathbf{s}, \mathbf{v} \rangle = l$ $\langle \mathbf{s}, \mathbf{v} \rangle = l \pmod k$ $\langle \mathbf{s}, \mathbf{v} \rangle = l + \eta$ short $\mathbf{v} \in \Lambda$
[DMMS21]	LWPR		$\mathrm{HW}(\mathbf{As})$
[HMM+23]	LWPR		$\mathbb{R}\text{-}\mathsf{lin}(\mathbf{As})$, $\mathbb{R}\text{-}\mathsf{quad}(\mathbf{As})$
[GKPV10]	Hard-to-Invert Leakage	Reduction	$\mathsf{f}(\mathbf{s})$
[AP12]	Extended-LWE		$\langle \mathbf{s}, \mathbf{v} \rangle \oplus e$
[BLP+13]	Extended-LWE		$\langle \mathbf{e}, \mathbf{v} \rangle$
[BD20]	Entropic-LWE		$\mathbf{s} \xleftarrow{\$} \mathcal{S}$
[KLSS23]	Hint-LWE		$\mathsf{Lin}(\mathbf{s}, \mathbf{e})$, in field/ring
[WLL24]	Linear Leakage		$\mathsf{Lin}'(\mathbf{s}, \mathbf{e})$, in field/ring
[LSW25]	Leaky-LWE		$\mathsf{Lin}''(\mathbf{s}, \mathbf{e})$, in field/ring
[BJTW25]	Entropic-LWE		$\mathbf{s} \xleftarrow{\$} \mathcal{S}$
Ours	ϕ-LWE $\approx$ LWPR		$\mathrm{HW}(\mathbf{As}) + \eta$

Starting with cryptanalysis results, Dachman-Soled et al. [DDGR20] considered multiple forms of mathematical hints on the secret $\mathbf{s} \in \mathbb{Z}_q^n$. *Perfect hints*: $\langle \mathbf{s}, \mathbf{v} \rangle = l$, *modular hints*: $\langle \mathbf{s}, \mathbf{v} \rangle = l \pmod k$, *approximate hints*: $\langle \mathbf{s}, \mathbf{v} \rangle = l + \eta$, where $+$ is the addition in $\mathbb{R}$, and *short vector hints*: $\lambda \in \Lambda$ a n-dimensional lattice, $\mathbf{v} \in \mathbb{Z}_q^n$, $l, k \in \mathbb{Z}$ and η follows a Normal distribution of parameter σ_η, known by the attacker as well as $\mathbf{v}$, l and k. By integrating those hints into a Bounding Distance Decoding (BDD) instance, they try to best approximate any form of additional information. In the same category of results, the LWPR problem has been analyzed (mostly) from the algebraic cryptanalysis viewpoint, assuming both linear and quadratic (real-valued) leakage functions [SLP05].

Moving to theoretical (reduction) results, a first application calling for LWE variants is homomorphic encryption, which gives free noisy linear information on the secret. In order to handle this type of side information, one of the early options considered was a noisy inner product of the secret, named Extended-LWE [AP12], whose hardness was then extended for an exact inner product on the LWE error by Brakerski et al. [BLP+13]. Subsequently, the analysis of LWE with additional information became more and more studied, under different names and assuming different types of information. A first line of works considered (possibly noisy) linear hints/leakages. See for example [KLSS23], [WLL24] and [LSW25]. Another line of works considered the more general problem of Entropic-LWE. It was seeded by Goldwasser et al. who studied a LWE variant with hard-to-invert leakage [GKPV10] and consolidated by Brakerski and Döttling who questioned the hardness of LWE with possibly non-linear information on the secret, assuming its distribution is entropic enough [BD20]. Recently, Boudgoust et al. also used hints on the secret to conduct an entropic analysis on its distribution, which leads to a stronger intermediate between the hint variant and the entropic one, providing hardness results for non-linear leakage [BJTW25].

Based on this state of the art, and concentrating on reduction results that are the focus of our work, one can first notice that existing variants based on (possibly noisy) linear hints or leakage are not applicable in the physical setting. This is because physical leakage functions like the Hamming weight one are linear in the reals, not in the field/ring,

and the same holds for their physical noise (which is additive in the reals, not in the field/ring). The Entropic-LWE variant therefore appears as more suitable to analyze physical leakage. Yet, it still does not directly lead to a satisfying solution. This is because, so far, the lattice literature has mostly focused on direct hints on the secrets (or errors) manipulated by cryptographic schemes, whereas our focus is on the output samples of LWPRN.[3] In side-channel terminology, direct hints on the secrets (or errors) can be viewed as of SPA type, whereas hints on computations involving these secrets are of DPA type. As a result, directly characterizing the security based on entropic hints cannot give tight security guarantees, as we now explain. For a random variable $\mathcal{S}$, define the min-entropy as $\mathrm{H}_\infty(\mathcal{S}) = -\log(\max_{\mathbf{s}} \Pr[\mathcal{S} = \mathbf{s}])$. Conditioning this min-entropy to the knowledge of a hint $\mathbf{y}$ related to $\mathbf{s}$, one can lower bound the conditional min-entropy of $\mathcal{S}$ knowing $\mathcal{Y}$ using $\mathrm{H}_\infty(\mathcal{S} \mid \mathcal{Y}) \geqslant \log_2(|\mathcal{K}|) - \log_2(|\mathcal{Y}|)$. In the particular case of LWPR with a Hamming weight leakage function and $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$, this means a fast decreasing bound $\mathrm{H}_\infty(\mathcal{S} \mid \mathcal{Y}) \geqslant n \cdot \log(q) - m \cdot \log(\log(q))$, where m is the number of samples that the adversary can observe and $\mathbf{y} = \mathrm{HW}(\mathbf{As})$. Note that such a bound essentially corresponds to the additive bound considered in DPA [dCGRP19]. Further considering LWPRN (i.e., noisy Hamming weights) would only reduce the leakage per sample proportionally to the side-channel Signal-to-Noise Ratio (SNR), which is inversely proportional to the variance σ_η^2 of the additive Gaussian noise η [Man04, DFS19]. So the main question we tackle in this work is whether such (seemingly pessimistic) results can be improved?

Contribution. We answer the question above positively. For this purpose, we leverage a new variant of LWE that we denote as LWE with physical hint (short: Φ-LWE) and which we instantiate with Noisy Hamming weight leakages, denoted as NHW-LWE.[4] The NHW-LWE problem asks to find a secret $\mathbf{s} \in \mathbb{Z}_q^n$ given $(\mathbf{A}, \mathbf{As} \oplus \mathbf{e}, \mathrm{HW}(\mathbf{As}) + \eta)$, where $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$, $\mathbf{s} \in \mathbb{Z}_q^n$, $\mathbf{e} \in \mathbb{Z}_q^m$ and $\eta \in \mathbb{R}^m$. This problem is linked to the LWPRN problem where the leakage function is a Noisy Hamming Weight, which can be seen as finding $\mathbf{s}$ given $(\mathbf{A}, \mathrm{HW}(\mathbf{As}) + \eta)$, with the same $\mathbf{A}, \mathbf{s}$ and η as above. The main step of our reduction from LWE to NHW-LWE is given by the following theorem and achieved by transforming a LWE sample $(\mathbf{A}, \mathbf{As} \oplus \mathbf{e})$ into a NHW-LWE sample $(\mathbf{A}, \mathbf{As} \oplus \mathbf{e}, \mathrm{HW}(\mathbf{As}) + \eta)$.

Main theorem (informal). Assume that the search LWE with dimension n, modulus q and Gaussian noise parameter σ_e is hard, then the search NHW-LWE problem is hard with additional Gaussian noise σ_η such that $\frac{\lceil \log_2(\sqrt{\lambda}\sigma_e) \rceil}{\sigma_\eta}$ is negligible.

A bit more precisely, and given the LWE sample $\mathbf{As} \oplus \mathbf{e}$, we can write $\mathrm{HW}(\mathbf{As} \oplus \mathbf{e}) = \mathrm{HW}(\mathbf{As}) + \mathbf{d}$, where the difference (or dependency) $\mathbf{d}$ captures the impact of the LWE error $\mathbf{e}$ viewed as an additive vector of noise (in $\mathbb{R}$). The reduction then proceeds by quantifying the amount of additive Gaussian noise η that is necessary to hide this difference. We use the classical *noise flooding* technique for this purpose and take advantage of either the statistical distance or Rényi divergence to conclude [BLL+15]. In this respect, it is worth underlying that our use of noise flooding differs from the one used to analyze the Leaky-LWE variant [LSW25], where the noise is in $\mathbb{Z}_q$ and the statistical noise flooding is replaced by a computational argument to reduce the super-polynomial size of the parameter.

Our results improve the entropic bound on two fronts. First, we show that the level of noise needed for the reduction to hold is independent of the field size q. This is in contrast with the situation of the DPA-type bound outlined above, where the noise needed to decrease the SNR depends on the side-channel signal, defined as the variance of the deterministic part of the leakage function. For example, in the case of a Hamming weight leakage function, this variance is worth $\log(q)/4$ for elements of $\mathbb{Z}_q$, implying that

[3] Internal computations leading to these samples are expected to be efficiently masked, by leveraging the key-homomorphism of the multiplication between the public IV and the long-term key.

[4] Φ-LWE could, however, be instantiated with other physical leakage functions in the future.

the amount of noise that is needed to hide the signal grows with the (log of the) field size. Second, the information (in the entropic bound) only decreases with the inverse of σ_η^2. By contrast, we argue that the impact of the additive Gaussian noise η towards hiding the difference $\mathbf{d}$ is amplified as $(\sigma_\eta^2)^{-4}$ in our reduction. The finer-grain argument supporting this amplification builds on the symmetry of the distributions that must be made indistinguishable in our proof, and can be nicely connected to the popular notion of statistical security order from the side-channel / masking literature [BDF+17].

As an additional bonus, we consider a serial implementation setting in our investigations (i.e., an adversary who can observe the leakage of the LWPRN samples one by one), a context in which LWPR (without noise) is known to be insecure [HMM+23]. We also show that our finer-grain security generalizes to realistic leakage functions [SLP05]. So besides the stronger confidence in the security of LWPRN that our reduction provides, our results have wider applicability (e.g., including serial/software implementations).

Overall, these findings therefore establish interesting connections between the lattice and the side-channel / masking literature, and consolidate the understanding of hard physical learning problems. Given their great potential for efficient leakage-resilient implementations, we hope these results will further encourage investigating such problems and considering them as promising building blocks for cryptographic primitives. This is true in the symmetric setting, where fresh re-keying schemes could be integrated in authenticated encryption schemes. This is even more true in the asymmetric (post-quantum) setting, where some intermediate operations in lattice-based schemes could be viewed as hard physical learning problems in order to reduce their need for implementation-level countermeasures [HLM+23]. Our investigations also open interesting avenues for further research, for example towards tighter bounds or more general leakage functions.

2 Background

Notations. Let λ denote the security parameter. We use bold letters to denote vectors and bold capital letters to denote matrices. For a positive integer q, let $\mathbb{Z}_q$ denote the ring of integers modulo q. For a distribution χ on a set X, we write $x \xleftarrow{\$} X$ to denote the operation of sampling a random value x according to χ. For an element $x \xleftarrow{\$} \mathbb{Z}_q$, let φ_x be the distribution followed by the Hamming weigt of x, denoted as $\mathrm{HW}(x)$. For two distributions X, Y, we write $X \overset{\mathsf{comp}}{\approx} Y$ or $X \overset{\mathsf{stat}}{\approx} Y$ to denote computational indistinguishability or statistical closeness. If $Supp(X) \subseteq Supp(Y)$, then, their statistical distance is defined as $\Delta(X,Y) := \frac{1}{2}\sum_{u \in S_X \cup S_Y} |\Pr(X=u) - \Pr(Y=u)|$, and verify all the properties of a distance: a symmetric defined positive function verifying the triangular inequality. We also recall the Rényi divergence of order 2 of two distributions P, Q with $Supp(P) \subseteq Supp(Q)$ as $\mathrm{RD}_2(P\|Q) = \sum_{x \in Supp(P)} \frac{P(x)^2}{Q(x)}$. The two previous distances ensure the *probability preservation* and the *data processing inequality* as the following lemma gives:

Lemma 1 (Rényi divergence properties in [LSS14, Lemma 4.1])**.** *Let P, Q be two probability distributions such that $Supp(P) \subseteq Supp(Q)$ and $X \subset Supp(Q)$ an arbitrary event. Then,*

- $P(X) \leqslant \Delta(P,Q) + Q(X)$,
- $P(X)^2 \leqslant \mathrm{RD}_2(P\|Q) \cdot Q(X)$.

Further, for any possibly randomized function f, let P^{f} (resp., Q^{f}) be the distribution obtained by sampling x from P (resp., Q) and outputting $\mathsf{f}(x)$. Then,

- $\Delta(P^{\mathsf{f}}, Q^{\mathsf{f}}) \leqslant \Delta(P,Q)$,
- $\mathrm{RD}_2(P^{\mathsf{f}}\|Q^{\mathsf{f}}) \leqslant \mathrm{RD}_2(P\|Q)$.

2.1 Lattices

Let $n \in \mathbb{Z}$, a n-dimensional full rank lattice Λ in $\mathbb{R}^n$ is defined as a discrete additive subgroup of $\mathbb{R}^n$, in the sense that there exists $d > 0$ such that $\mathcal{B}(\mathbf{0}, d) \cap L = \{\mathbf{0}\}$. Any lattice Λ is given by the set of all integer combinations of linearly independent basis vectors $\mathbf{B} = \{\mathbf{b_1}, ..., \mathbf{b_n}\} \in \mathbb{R}^n$: $\Lambda(\mathbf{B}) = \{\sum_{i=1}^{n} x_i \mathbf{b_i} : x_i \in \mathbb{Z}\}$. For a lattice $\Lambda \subset \mathbb{R}^n$ and an $\varepsilon < 1$, Micciancio and Regev introduced the notion of *smoothing parameter* in [MR04], that gives the minimum parameter for a discrete Gaussian distribution over Λ to behave like a continuous one: $\eta_\varepsilon(\Lambda) = \min\left\{s \in \mathbb{R} \text{ s.t. } \rho_{1/s}(\Lambda^* \setminus \{\mathbf{0}\}) \leqslant \varepsilon\right\}$.

2.2 Gaussian distributions

For $\mathbf{x} \in \mathbb{R}^n$, a standard deviation $\sigma > 0$ and a mean $\mu \in \mathbb{R}^n$, we define the continuous Gaussian function as $\rho_{\sigma,\mu}(\mathbf{x}) = \exp\left(\frac{-\pi\|\mathbf{x}-\mu\|^2}{\sigma^2}\right)$. We denote the continuous Gaussian distribution with parameters σ, μ as $\mathcal{D}_{\sigma,\mu} = \frac{\rho_{\sigma,\mu}(\mathbf{x})}{\sigma^n}$. For a lattice Λ, the discrete distribution over Λ, denoted by $\mathcal{D}_{\Lambda,\sigma,\mu}$ is defined as $\mathcal{D}_{\Lambda,\sigma,\mu} = \frac{\rho_{\sigma,\mu}(\mathbf{x})}{\sum_{\mathbf{y}\in\Lambda}\rho_{\sigma,\mu}(\mathbf{y})}$. In case $\mu = 0$, we write $\mathcal{D}_\sigma$ (resp. $\mathcal{D}_{\Lambda,\sigma}$). Gaussian distributions allow sampling short elements with high probability [Lyu12, Lemma 4.4]:

$$\Pr_{\mathbf{x}\hookleftarrow\mathcal{D}_{\sigma,\mu}}\left(\|\mathbf{x}\| \geqslant \sqrt{n}\sigma\right) \leqslant 2e^{\frac{-n}{2}}.$$

Another nice property of Gaussian distributions is that we can add two Gaussians with parameters σ_1, μ_1 and σ_2, μ_2, and the convolution product gives a Gaussian distribution with parameters $\sqrt{\sigma_1^2 + \sigma_2^2}$ and $\mu_1 + \mu_2$. The next lemma gives an upper bound on the statistical distance (resp., the Rényi divergence) between two Gaussians with the same deviation and different centers from [GKPV10] (resp., [LSS14]).

Lemma 2 (Same deviation with different center [GKPV10, Lemma 3], [LSS14, Lemma 4.2])**.** *Let $\sigma > 0, q \in \mathbb{Z}$, and $w, z \in \mathbb{Z}$ be arbitrary values. Define $P = \mathcal{D}_{\mathbb{Z}_q,\sigma,w}$ and $Q = \mathcal{D}_{\mathbb{Z}_q,\sigma,z}$. Then, the following bounds hold:*

- $\Delta(\lfloor q\rho_{\sigma,w}\rceil \pmod q, \lfloor q\rho_{\sigma,z}\rceil \pmod q) \leqslant |w - z|/(\sigma)$;
- *If $\sigma \geqslant \eta_\varepsilon(\mathbb{Z})$ for some $\varepsilon \in (0, 1)$, we have:*

$$\mathrm{RD}_2(P\|Q) \in \left[\left(\frac{1-\varepsilon}{1+\varepsilon}\right)^2; \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2\right] \cdot \exp\left(\frac{2\pi|w-z|^2\,^2}{\sigma}\right).$$

2.3 Learning with Errors

The *learning with error* problem lies in the distinction between slightly perturbed random linear equations and truly random ones. Regev [Reg05] proved that the task of distinguishing LWE samples $(\mathbf{A}, \mathbf{As} + \mathbf{e})$ from a uniform random distribution (the decisional LWE problem), or the task of retrieving the secret $\mathbf{s}$ from these samples (the search LWE problem), is as hard as solving certain worst-case instances of well-studied lattice problems. This hardness assumption holds for specific error distributions, most notably when the error vector $\mathbf{e}$ is drawn from a discrete Gaussian distribution with parameter σ_e.

Definition 1 (Learning With Errors - LWE [Reg05])**.** Let λ be the security parameter, n, m and q be some integers, and let χ be a distribution over $\mathbb{Z}_q$, all functions of λ. A sample $(\mathbf{A}, \mathbf{b}) \in \mathbb{Z}_q^{m\times n} \times \mathbb{Z}_q^m$ following the $\mathrm{LWE}_{n,m,q,\chi}$ distribution is defined for an $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m\times n}$, $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$, $\mathbf{e} \leftarrow \chi^m$, $\mathbf{u} \xleftarrow{\$} \mathbb{Z}_q^m$ as follows:

$$(\mathbf{A}, \mathbf{b}) := (\mathbf{A}, \mathbf{As} \oplus \mathbf{e}).$$

The problem of distinguishing a LWE sample from a uniformly random sample $(\mathbf{A}, \mathbf{u})$ is known as the decisional LWE (dLWE) problem. The problem of retrieving the secret $\mathbf{s}$ when given access to LWE samples is known as the search LWE (sLWE) problem.

The advantage of an adversary $\mathcal{A}$ against LWE is defined as $\mathsf{Adv}_{\mathrm{LWE}}[\mathcal{A}(\mathbf{A}, \mathbf{As} \oplus \mathbf{e}) = \mathbf{s}] = |\frac{1}{2} - \Pr(\mathcal{A}\ \texttt{Win})|$. If $\mathsf{Adv}_{\mathrm{LWE}}(\mathcal{A}) \geqslant \varepsilon$ for a non-negligible ε, $\mathcal{A}$ is a *distinguisher*.

We next use the notations of Table 2, where all σ denote standard deviations.

Table 2: Notation used throughout the paper.

Notation	Element
λ	security parameter
q	modulus
n	LWE dimension
m	number of samples
ℓ_q	$\lceil \log(q) \rceil$
σ_e	LWE error
σ_η	Gaussian noise added to the HW leakage
$\mathbf{d}$	Influence of the LWE error on the HW leakage
$+$	Addition in $\mathbb{R}$
$\oplus$	Addition in $\mathbb{Z}_q$

3 Hard learning problems

In this section, we introduce the LWPRN problem in a general manner and present the specific instantiation that we will study. We also present the physical hint LWE problem that we use in our reduction, together with the specific instantiation we consider.

3.1 Learning with physical rounding and noise

Physical learning problems are a class of challenges that follow the structure of standard learning problems (by relying on noisy or rounded inner products) but are defined using a physical function. The latter is typically carried out by a leakage function that is expected to be noisy or compressive. In order to study these problems rigorously, it is therefore necessary to assert specific properties on the leakage function, most often by approximating it as a mathematical model of a physical quantity. In this paper, we will focus in particular on generalizing the following Learning With Physical Rounding (LWPR) problem:

Definition 2 (Learning With Physical Rounding, adapted from [DMMS21, Definition 1])**.** Let $q, n, m \in \mathbb{N}^*$, $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}$, $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$. The $\mathrm{LWPR}^{\mathsf{L}}_{q,n,m}$ distribution is given by:

$$\mathrm{LWPR}^{\mathsf{L}}_{q,n,m} := (\mathbf{A}, \mathsf{L}(\mathbf{As}))\,,$$

where $\mathsf{L} : \mathbb{Z}_q^m \to \mathbb{R}^{n_d}$ is the physical rounding function. The problem of retrieving the secret $\mathbf{s}$ when given access to LWPR samples is known as the LWPR problem.

This problem was introduced as the basis of an efficient re-keying scheme for symmetric cryptographic protocols. Its security analyzes so far are mostly heuristic. The initial work of Duval et al. considers different types of (mostly algebraic) attacks assuming a Hamming weight leakage function. This work has then been generalized to any linear or quadratic leakage functions [HMM+23]. Informally, the security of LWPR stems from the algebraic incompatibility between (inner) product operations modulo a prime modulus q

and the Hamming weight leakage function (or variants thereof), which can be viewed as an application of symmetric cryptography relying on alternating moduli [BIP+18].

LWPR was first defined assuming a noise-free leakage function. This is a conservative choice, since a level of additive (e.g., Gaussian) measurement noise generally affects concrete measurements, and could make the equations provided by LWPR samples "noisier" and therefore more difficult to exploit. Besides simplifying cryptanalysis, and as argued in introduction, it also makes it hard(er) to reduce LWPR to a standard learning problem, because noise cannot be used to hide differences between distributions, as typically leveraged in other reductions from hard physical learning problems to mathematical ones.

As a step towards formalizing the security of such problems, we therefore choose to generalize LWPR towards a noisy variant, where an independent Gaussian noise is added to the LWPR samples, leading to the LWPRN problem which we defined as follows:

Definition 3 (Learning With Physical Rounding and Noise - LWPRN)**.** Let $q, n, m \in \mathbb{N}^*$, $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}$, $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$. The $\mathrm{LWPRN}_{q,n,m}^{\mathsf{L},\sigma}$ distribution is given by:

$$\mathrm{LWPRN}_{q,n,m}^{\mathsf{L},\sigma} := (\mathbf{A}, \mathsf{L}(\mathbf{A}\mathbf{s}) + \eta),$$

with $\mathsf{L} : \mathbb{Z}_q^m \to \mathbb{R}^{n_d}$ the physical rounding function and η a noise of which each component follows a continuous Gaussian distribution $\mathcal{N}(0, \sigma_\eta^2)$. The problem of distinguishing a $\mathrm{LWPRN}_{q,n,m}^{\mathsf{L},\sigma}$ sample from a sample $(\mathbf{A}, \mathsf{L}(\mathbf{u}) + \eta)$ with $\mathbf{u} \xleftarrow{\$} \mathbb{Z}_q^m$ is known as the decisional LWPRN (dLWPRN) problem. The problem of retrieving the secret $\mathbf{s}$ when given access to $\mathrm{LWPRN}_{q,n,m}^{\mathsf{L},\sigma}$ samples is known as the search LWPRN (sLWPRN) problem.

It is important to note that the measurement noise η follows a continuous Gaussian distribution so that the noisy Hamming weights are real numbers (whereas the LWE errors are in $\mathbb{Z}_q$). Thanks to its additional noise, the LWPRN problem is expected to be harder to solve than LWPR and, as we shall see next, can be reduced to standard LWE.

3.2 Instantiation of LWPRN

LWPRN is defined for any deterministic leakage function and additive noise distribution, over any modulus. We next discuss the specific instantiations that we study.

Modulus. Previous LWPR analyses have been primarily limited to (Mersenne) prime moduli. In contrast, the LWPRN proofs in this work cover any integer q. This is another advantage of the LWPRN variant: the addition of noise to a deterministic leakage function allows avoiding attacks threatening LWPR with power of two moduli.

Leakage function. We next instantiate the LWPRN problem with a (noisy) Hamming weight leakage function. This is a natural first abstraction since many concrete leakage functions are strongly correlated with the Hamming weight function [MOP07]. It was also the starting point of [DMMS21]. Besides, the arguments in Section 5 suggest that generalizing our conclusions to other leakage functions, like the linear and quadratic ones from [HMM+23], should be feasible. We leave it as an interesting open problem.

Implementation setting. As initially described in [DMMS21], LWPR (and therefore LWPRN) can be implemented in two settings – the serial and the parallel ones – which essentially define the information obtained for any leaking vector $\mathbf{x} = \mathbf{x_1}||\ldots||\mathbf{x_m}$:

1. *Parallel (hardware) case.* Hardware implementations allow for highly parallelized computations, leading to leakage on larger values. In this case, m inner products

$\langle \mathbf{a_i}, \mathbf{s} \rangle$ are computed in one clock cycle. Consequently, a side-channel adversary obtains a single Hamming weight measurement on a value of size $m \cdot \log q$:

$$\mathsf{L}(\mathbf{x}) = \mathrm{HW}(\mathbf{x_1}) + \cdots + \mathrm{HW}(\mathbf{x_m}).$$

2. *Serial (software) case*: Software implementations are generally limited to smaller (e.g., 32-bit) registers. In this case, the components of the inner products $\langle \mathbf{a_i}, \mathbf{s} \rangle$ are typically computed one by one. A side-channel adversary therefore gains access to m distinct Hamming weight measurements, each on a $\log q$-size value:

$$\mathsf{L}(\mathbf{x}) = \Big(\mathrm{HW}(\mathbf{x_1}), \ldots, \mathrm{HW}(\mathbf{x_m}) \Big).$$

It is easy to see that the serial case provides the adversary with significantly more information. Assuming a Hamming weight leakage function, it corresponds to m times $\log(\ell_q)$ bits, to be compared with once $\log(m \cdot \ell_q)$ bits for the parallel case.[5]

In the following, we will focus on the (harder to prove) serial variant that enables more applications. Interestingly, it was shown in [DMMS21, HMM+23] that a serial instantiation of LWPR is insecure for reasonably sized moduli. This is because an adversary can then wait for an extreme Hamming weight that is giving him an unbounded leakage on the corresponding inner product output, and therefore a linear equation of the secret. This is feasible as long as q is significantly smaller than the number of measured leakages.

Note that there is no contradiction between this impossibility in the noise-free case and our following results. As already mentioned, the heuristic hardness of LWPR comes primarily from the algebraic properties of the leakage function. By contrast, the asymptotic hardness of LWPRN (demonstrated in this paper) is proven in a for a limited number of samples, and rather derives from the noise that is added to the leakage function.

3.3 Physical hint LWE

In the path towards reducing the LWPRN problem to the LWE assumption, we introduce a physical variant of LWE, which we coin *physical hint LWE* (short: Φ-LWE). In this variant, each sample is accompanied by a side-channel leakage (i.e., a physical hint). Concretely, the physical hint is a function of the inner products with errors $\langle \mathbf{a_i}, \mathbf{s} \rangle \oplus \mathbf{e}$ rather than a function of the secret $\mathbf{s}$ itself. This distinction makes Φ-LWE different from standard LWE variants with hints [KLSS23]. Critically, each sample in Φ-LWE provides extra information about the secret, even if that information can be computationally hard to exploit.

Definition 4 (Physical Hint-LWE - Φ-LWE)**.** Let λ be the security parameter, n, m and q be some integers, and let χ be a distribution over $\mathbb{Z}_q$, all functions of λ. Let $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}$ public vectors, $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$ be the secret vector. Let $\Phi : \mathbb{Z}_q^n \times \mathbb{Z}_q^m \to \mathbb{R}$ be a physical function. The $\Phi\text{-}\mathsf{LWE}^{\Phi}_{q,n,m}$ distribution is defined as:

$$\Phi\text{-}\mathsf{LWE}^{\Phi}_{q,n,m} := (\mathbf{A}, \mathbf{As} \oplus \mathbf{e}, \Phi(\mathbf{s}, \mathbf{e})),$$

where $\mathbf{e} \leftarrow \chi^m$ is an independent error. The problem of distinguishing a $\Phi\text{-}\mathsf{LWE}^{\Phi}_{q,n,m}$ sample from a uniformly random sample $(\mathbf{A}, \mathbf{u}, \Phi(\mathbf{s}, \mathbf{e}))$, where $\mathbf{u} \xleftarrow{\$} \mathbb{Z}_q^m$, is known as the decisional Φ-LWE problem. The problem of retrieving the secret $\mathbf{s}$ when given access to $\Phi\text{-}\mathsf{LWE}^{\Phi}_{q,n,m}$ samples is known as the search Φ-LWE problem.

[5] The examples above illustrate the extreme (most) parallel and (most) serial cases. It is possible to consider tradeoffs where more than one and less then m inner products are processed in parallel.

Given our choice in Section 3.2 to focus on the LWPRN problem with a noisy Hamming weight leakage function, we will naturally consider a similar instantiation of Φ-LWE, that we next denote as Noisy Hamming Weight LWE (NHW-LWE). In this instantiation, the physical hints are defined as $\Phi(\mathbf{s}, \mathbf{e}) = \mathrm{HW}(\mathbf{As}) + \eta$, where the noise η follows an independent continuous Gaussian distribution and the physical (Hamming weight) rounding can be seen as an additional deterministic error function. Yet, we expect the generality of the Φ-LWE problem to be useful in the future treatment of other leakage functions.

4 Technical results

Our goal is to prove the hardness of the LWPRN problem when the physical leakage L is a noisy Hamming weight, using the hardness of the LWE problem. For this purpose, we proceed in two steps, illustrated in Table 3. The main one is a reduction from LWE to the intermediate problem NHW-LWE. The second one is a trivial reduction from NHW-LWE to LWPRN for an instance $\mathsf{L} = \Phi$ corresponding to noisy Hamming weights.

Table 3: Overview of the reduction from LWE to LWPRN.

$\mathrm{LWE}_{q,n,m,\sigma_e}$	$\xrightarrow{\text{Theorem 1}}$	$\text{NHW-LWE}^{\sigma_\eta}_{q,n,m,\sigma_e}$	$\xrightarrow{\text{remove } \mathbf{b}}$	$\mathrm{LWPRN}^{\mathrm{HW},\sigma_\eta}_{q,n,m}$
$(\mathbf{A}, \mathbf{b})$		$(\mathbf{A}, \mathbf{b}, \mathrm{HW}(\mathbf{As}) + \eta)$		$(\mathbf{A}, \mathrm{HW}(\mathbf{As}) + \eta)$

4.1 Overview of the reduction

The general idea of the main reduction is depicted in Table 4. Using an adversary $\mathcal{A}$ against the NHW-LWE problem, we construct an adversary $\mathcal{B}$ against the LWE assumption. For this, we need to transform a LWE sample $(\mathbf{A}, \mathbf{b} = \mathbf{As} \oplus \mathbf{e})$ into a NHW-LWE sample $(\mathbf{A}, \mathbf{b}, \mathrm{HW}(\mathbf{As}) + \eta)$, where the target distribution of η is a continuous Gaussian distribution of parameter σ_η. From a high level perspective, we use the Hamming weight of $\mathbf{b}$ to build an element $\mathbf{h}'$ of distribution indistinguishable from the one of $\mathrm{HW}(\mathbf{As}) + \eta$. We start by considering $\mathbf{h} = \mathrm{HW}(\mathbf{b}) = \mathrm{HW}(\mathbf{As} \oplus \mathbf{e})$ and demonstrate it is indistinguishable from $\mathrm{HW}(\mathbf{As}) + \mathbf{d}$, with each coefficient of $\mathbf{d} = \mathrm{HW}(\mathbf{As} \oplus \mathbf{e}) - \mathrm{HW}(\mathbf{As})$ upper bounded by a function of the error deviation σ_e. Then, the adversary $\mathcal{B}$ only has to sample an error η and add it to $\mathbf{h}$ to erase the error $\mathbf{d}$ and obtain a NHW-LWE sample $(\mathbf{A}, \mathbf{b}, \mathbf{h} + \eta \overset{\text{stat}}{\approx} \mathbf{h}')$.

Table 4: High-level idea of the reduction from LWE to ϕ-LWE.

$\mathcal{C}$	LWE	$\mathcal{B}$	ϕ-LWE	$\mathcal{A}$
$\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}$				
$\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$, $\mathbf{e} \leftarrow \chi^m$				
	$\xrightarrow{(\mathbf{A}, \mathbf{b} = \mathbf{As} \oplus \mathbf{e})}$			
		$\mathbf{h} := \mathrm{HW}(\mathbf{b})$		
		$\approx \mathrm{HW}(\mathbf{As}) + \mathbf{d}$		
		$\eta \hookleftarrow \mathcal{D}_{\sigma_\eta}$		
			$\xrightarrow{(\mathbf{A}, \mathbf{b}, \mathbf{h} + \eta)}$	

We recall the following notations for the standard deviations:

- σ_e denotes the standard deviation of the LWE error;
- σ_η denotes the deviation of the noise term η we want to add to $\mathrm{HW}(\mathbf{a} \cdot \mathbf{s})$.

There are two main points to handle in the reduction:

1. We need to characterize the distribution followed by the Hamming weight of $\mathbf{b}$ and its distance to the noisy Hamming weight of the product $\mathbf{As}$.

2. We need to mitigate this distance to target the real sample thanks to the noise η.

4.2 Steps of the reduction

Let $(\mathbf{A}, \mathbf{b} = \mathbf{As} \oplus \mathbf{e})$, where $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}$, $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$ and $\mathbf{e} \xleftarrow{\$} \mathcal{D}_{\mathbb{Z}_q,\sigma_e}^m$, be a LWE sample. We first consider the additional LWE error $\mathbf{e}$ and its influence over the Hamming Weight of $\mathbf{b}$, viewed as an additive noise (in $\mathbb{R}$) rather than additive errors (in $\mathbb{Z}_q$). This influence can be characterized by an upper bound thanks to the Gaussian behavior of the error.

Influence of the LWE error. We consider the influence $\mathbf{d}$ as an error lying in a certain range with high probability. We start by showing the following lemma, which gives an upper bound on its influence with respect to $\mathrm{HW}(\langle \mathbf{a_i}, \mathbf{s} \rangle \oplus e_i)$.

Lemma 3. *Let λ be a security parameter, $\mathbf{a} \xleftarrow{\$} \mathbb{Z}_q^n$, and $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$, if $e \xleftarrow{\$} \mathcal{D}_{\mathbb{Z}_q,\sigma_e}$, then $|\mathrm{HW}(\mathbf{as}) - \mathrm{HW}(\mathbf{as} \oplus e)| \leqslant \lceil \log_2(\sqrt{\lambda}\sigma_e) \rceil$ with probability $1 - negl(\lambda)$.*

Proof. Taking advantage of the shortness of Gaussian elements [Lyu12, Lemma 4.4], we get the threshold needed to obtain a negligible probability for a Gaussian element to lie outside a certain range:

$$\Pr_{\mathbf{x} \xleftarrow{\$} \mathcal{D}_\sigma} \left(\|\mathbf{x}\| \geqslant \sqrt{\lambda}\sigma \right) \leqslant 2e^{-\frac{\lambda}{2}}. \tag{1}$$

Hence, the probability for the bitlenght of a Gaussian element e being more than $d = \lceil \log_2(\sqrt{\lambda}\sigma_e) \rceil$ is negligible: $\Pr_{e \xleftarrow{\$} \mathcal{D}_{\mathbb{Z}_q,\sigma_e}} \left(\log_2(e) \geqslant \lceil \log_2(\sqrt{\lambda}\sigma_e) \rceil \right) \leqslant \varepsilon$ for a negligible ε.

Let $\mathbf{a} \xleftarrow{\$} \mathbb{Z}_q^n$, and $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$, we want to evaluate an upper bound of the difference between $\mathrm{HW}(\mathbf{as})$ and $\mathrm{HW}(\mathbf{as} \oplus e)$. Notice that this difference is bounded by the number of bit that flip: $|\mathrm{HW}(\mathbf{As}) - \mathrm{HW}(\mathbf{As} \oplus \mathbf{e})| \leqslant |\{i \in [0, \ell_q] : (\mathbf{as})_i \neq (\mathbf{as} \oplus \mathbf{e})_i\}|$. We are showing that the probability for the bit i such that $d < i < \ell_q$ of $\mathbf{as}$ to change is negligible after adding the element e. For $0 \leqslant i \leqslant \ell_q$, we consider the three following events:

- $(\mathbf{as})_i :=$ "*The i-th bit of the product* $\mathbf{as}$ *is equal to* 1",
- $F_i :=$ "*The i-th bit of the product* $\mathbf{as}$ *flip (i.e.* $(\mathbf{as})_i \neq (\mathbf{as} \oplus e)_i$)",
- $C_i :=$ "*There is a carry from previous bit*".

Table 5: Summary of the flip event F_i and C_{i+1} event for the i-th bit.

$(\mathbf{as})_i$	e_i	C_i	F_i	C_{i+1}	$(\mathbf{as})_i$	e_i	C_i	F_i	C_{i+1}
0	0	0	0	0	1	0	0	0	0
0	0	1	1	0	1	0	1	1	1
0	1	0	1	0	1	1	0	1	1
0	1	1	0	1	1	1	1	0	1

Notice that the event F_i happens only when the event e_i or (exclusively) C_i happens, which gives the logic Table 5. From that, we can write:

$$\begin{aligned}
\Pr(F_i) &= \Pr\left(\overline{(\mathbf{as})_i}, e_i, \overline{C_i}\right) + \Pr\left(\overline{(\mathbf{as})_i}, \overline{e_i}, C_i\right), \\
&\quad + \Pr\left((\mathbf{as})_i, e_i, \overline{C_i}\right) + \Pr\left((\mathbf{as})_i, \overline{e_i}, C_i\right), \\
&\leqslant \Pr\left(\overline{(\mathbf{as})_i}, \overline{e_i}, C_i\right) + \Pr\left((\mathbf{as})_i, \overline{e_i}, C_i\right) + negl(\lambda), \\
&\leqslant \Pr\left(\overline{(\mathbf{as})_i}, C_i\right) + \Pr\left((\mathbf{as})_i, C_i\right) + negl(\lambda), \\
&= \Pr\left(\overline{(\mathbf{as})_i}\right) \Pr(C_i) + \Pr\left((\mathbf{as})_i\right) \Pr(C_i) + negl(\lambda), \\
&= \left(1 - \Pr\left((\mathbf{as})_i\right)\right) \Pr(C_i) + \Pr\left((\mathbf{as})_i\right) \Pr(C_i) + negl(\lambda), \\
&= \Pr(C_i) + negl(\lambda),
\end{aligned}$$

where the first line is provided by Table 5, the second and third equality by the fact that if $i > d$ then $\Pr(e_i) \leqslant negl(\lambda)$ and $1 - negl(\lambda) \leqslant \Pr(\overline{e_i}) \leqslant 1$, and the fourth equation comes from the independence of the two events. Hence, we need to evaluate the probability for a carry to happen for the i-th bit. In the same way, the logic Table 5 also gives the condition for the event C_{i+1} to happen. From that, we can write :

$$\begin{aligned}
\Pr(C_{i+1}) &= \Pr\left(\overline{(\mathbf{as})_i}, e_i, C_i\right) + \Pr\left((\mathbf{as})_i, e_i, \overline{C_i}\right), \\
&\quad + \Pr\left((\mathbf{as})_i, \overline{e_i}, C_i\right) + \Pr\left((\mathbf{as})_i, e_i, C_i\right), \\
&\leqslant \Pr\left((\mathbf{as})_i, \overline{e_i}, C_i\right) + negl, \\
&\leqslant \Pr\left((\mathbf{as})_i, C_i\right) + negl, \\
&= \Pr\left((\mathbf{as})_i\right) \Pr(C_i) + negl, \\
&\approx \frac{1}{2} \Pr(C_i) + negl,
\end{aligned}$$

where the second and third lines come from $\Pr(e_i = 1 \mid i > d) \leqslant negl$, the second-to-last comes from the independence of the two events, and the last comes from the small distance of the product $\mathbf{as}$ to the uniform. We conclude by induction on i. Then, there are at most d bits that change and this conclude the proof. □

Approaching the target distribution. The purpose of this part is to approach at best the target distribution $\mathrm{HW}(\mathbf{As}) + \eta$ from $\mathrm{HW}(\mathbf{b})$. Using the previous result, we can write $\mathrm{HW}(\mathbf{b}) = \mathrm{HW}(\mathbf{As}) + \mathbf{d}$ with $\|\mathbf{d}_i\| \leqslant \lceil \log_2(\sqrt{\lambda}\sigma_e) \rceil$ for all $0 \leqslant i < m$. We use the so-called *noise flooding* technique to erase the influence $\mathbf{d}$. Intuitively, the more we add noise the less we obtain information and the closer the output distribution is to a Gaussian one. As a result, one may require a large amount of additional noise, leading to inefficient parameters. Yet, as Bai et al. design in [BLL+15], we can also take advantage of the Rényi divergence multiplicative property to improve parameters. The following corollary gives the necessary condition on σ_η considering both statistical distance and Rényi divergences.

Corollary 1. *Let $z \in \mathbb{Z}$, $\sigma > 0$ and ε be negligible.*

- *If $\frac{|z|}{\sigma} \leqslant \varepsilon$ then $\Delta(\mathcal{D}_{\sigma,z}, \mathcal{D}_\sigma) \leqslant \varepsilon$.*
- *If $\sigma \geqslant |z|$, $\varepsilon = \mathcal{O}\left(\frac{1}{8}\right)$ and $\sigma \geqslant \eta_\varepsilon(\mathbb{Z})$ then $\mathrm{RD}(\mathcal{D}_{\sigma,z} \| \mathcal{D}_\sigma) \leqslant \mathcal{O}(1)$.*

Proof. Let $z \in \mathbb{Z}$, $\sigma > 0$ and ε be negligible. We consider first the statistical distance. Let $\frac{|z|}{\sigma} \leqslant \varepsilon$, applying Lemma 2 on $\Delta(\mathcal{D}_{\sigma,z}, \mathcal{D}_\sigma)$, we get $\Delta(\mathcal{D}_{\sigma,z}, \mathcal{D}_\sigma) \leqslant \frac{|z|}{\sigma} \leqslant \varepsilon$, where the second inequality comes from $\frac{|z|}{\sigma} \leqslant \varepsilon$, and conclude the proof for the statistical distance.

We next prove the second line for the Rényi divergence. Let $\sigma \geqslant \sqrt{\pi}|z|$ and $\sigma \geqslant \eta_\varepsilon(\mathbb{Z})$. Recall that the Rényi divergence requires only $poly(n)$ bounds to ensure the closeness of two distributions P and Q. Then, applying Lemma 2 on $P = \mathcal{D}_{\sigma,z}$ and $Q = \mathcal{D}_\sigma$, we get:

$$\mathrm{RD}_2(\mathcal{D}_{\sigma,z} \| \mathcal{D}_\sigma) \leqslant \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 \exp\left(\frac{2\pi|z|^2}{\sigma^2}\right).$$

By requiring $\frac{2\pi|z|^2}{\sigma^2} \leqslant 2\pi$, i.e., $\sigma \geqslant |z|$, we obtain $\exp\left(\frac{2\pi|z|^2}{\sigma^2}\right) \leqslant \exp(2\pi)$. The first component of the approximation gives $\left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 = \left(1 + \frac{4\varepsilon/(1-\varepsilon)}{2}\right)^2$. As $\left(1+\frac{x}{y}\right)^y \leqslant \exp(x)$ for all $x, y > 0$, then $\left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 \leqslant \exp\left(\frac{4\varepsilon}{1-\varepsilon}\right)$. Let $\varepsilon < 1/2$, then $\frac{1}{1-\varepsilon} < 2$ and thus we get $\left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 \leqslant \exp(8\varepsilon)$. Finally, requiring $\varepsilon = \mathcal{O}\left(\frac{1}{8}\right)$, we obtain:

$$\mathrm{RD}_2(\mathcal{D}_{\sigma,z} \| \mathcal{D}_\sigma) \leqslant \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 \exp\left(\frac{2\pi|z|^2}{\sigma^2}\right) \leqslant \exp(8\varepsilon)\exp(2\pi) \leqslant \mathcal{O}(1)$$

Then, the last equation concludes the proof for the Rényi divergence. □

Remark 1. The Rényi divergence gives a lower bound than statistical distance, and can therefore lead to a tighter result on the search version of the LWE problem. However, this technique cannot be used in the decisional version of the LWE problem.

Now that we have established the relation between $\mathrm{HW}(\mathbf{b})$ and $\mathrm{HW}(\mathbf{As})$, and bounded the level of Gaussian noise we need to target the distribution of physical hints, the following subsection gives the reduction from LWE to Φ-LWE in their search versions.

4.3 Hardness of the search variant

In this subsection, we present the main theoretical result of this paper, namely necessary conditions under which the reduction from (search) LWE to NHW-LWE holds.

Theorem 1. *Let λ be a security parameter, $n, m \in \mathbb{N}$, $q \in \mathbb{N}$, $\sigma_e, \sigma_\eta > 0$. If the additive (physical) noise level satisfies $\frac{\lceil \log_2(\sqrt{\lambda}\sigma_e)\rceil}{\sigma_\eta} \leqslant \varepsilon$ for a negligible ε, then there exists a reduction from $s\mathrm{LWE}(\lambda)_{n,m,q,\sigma_e}$ to $\mathrm{sNHW\text{-}LWE}(\lambda)_{n,m,q,\sigma_e,\sigma_\eta}$.*

Proof. Let λ be the security parameter and $n, m, q, \sigma_e, \sigma_\eta$ and ε be such as in the theorem. The proof considers the same game as in Table 4 between $\mathcal{B}$ and $\mathcal{A}$. Given a NHW-LWE sample $(\mathbf{A}, \mathbf{As} \oplus \mathbf{e}, \mathrm{HW}(\mathbf{b}) + \eta)$ such as $\mathcal{B}$ returns to $\mathcal{A}$, then the search version asks to find $\mathbf{s}$. Let's define the following three distributions:

- $H_0 := (\mathbf{A}, \mathbf{As} \oplus \mathbf{e}, \mathrm{HW}(\mathbf{As} \oplus \mathbf{e}) + \eta)$ where $\mathbf{A}, \mathbf{s}, \mathbf{e}$ are LWE samples and $\eta \xleftarrow{\$} \mathcal{D}^m_{\sigma_\eta}$;
- $H_1 := (\mathbf{A}, \mathbf{As} \oplus \mathbf{e}, \mathrm{HW}(\mathbf{As}) + \mathbf{d} + \eta)$ where $\|\mathbf{d}_i\| \leqslant \lceil \log(\sqrt{\lambda}\sigma_e)\rceil$ for all $0 \leqslant i < m$;
- $H_2 := (\mathbf{A}, \mathbf{As} \oplus \mathbf{e}, \mathrm{HW}(\mathbf{As}) + \eta')$ where $\eta' \xleftarrow{\$} \mathcal{D}^m_{\sigma_\eta}$.

The first distribution H_0 corresponds to the sample transmitted by $\mathcal{B}$ to $\mathcal{A}$ in Table 4 and the last one corresponds to a sNHW-LWE sample. For $0 \leqslant i \leqslant 2$, we denote G_i the problem of finding the secret $\mathbf{s}$ given the distribution H_i. We want for all i that $\mathsf{Adv}_{G_i}[\mathcal{A}(H_i) = \mathbf{s}]$ is negligible for the problem to be hard. Then, showing G_2 is hard suffices to show that G_0 is hard. Step by step, this means:

- G_2 is the game of finding $\mathbf{s}$ knowing a truly NHW-LWE sample. So far, $\mathcal{A}$ corresponds to its definition, namely an adversary against NHW-LWE.

- G_2 to G_1: The probability preservation property given in Lemma 1 gives:

$$\mathsf{Adv}_{G_1}[\mathcal{A}(H_1) = \mathbf{s}] \leqslant \mathsf{Adv}_{G_2}[\mathcal{A}(H_2) = \mathbf{s}] + \Delta(H_2, H_1).$$

  To evaluate the statistical distance between H_2 and H_1, it then suffices to compute the statistical distance between $\mathcal{D}^m_{\sigma_\eta, \lceil \log(\sqrt{\lambda}\sigma_e)\rceil}$ and $\mathcal{D}^m_{\sigma_\eta}$. By the additive property of the statistical distance, we get:

$$\begin{aligned}\Delta\left(\mathcal{D}^m_{\sigma_\eta, \lceil \log(\sqrt{\lambda}\sigma_e)\rceil}, \mathcal{D}^m_{\sigma_\eta}\right) &= m\Delta\left(\mathcal{D}_{\sigma_\eta, \lceil \log(\sqrt{\lambda}\sigma_e)\rceil}, \mathcal{D}_{\sigma_\eta}\right), \\ &\leqslant m\varepsilon,\end{aligned}$$

  where the second line comes from Corollary 1 and $\frac{\lceil \log_2(\sqrt{\lambda}\sigma_e)\rceil}{\sigma_\eta} \leqslant \varepsilon$. Therefore, the two distributions are statistically indistinguishable: $\Delta(H_2, H_1) \leqslant 2^{-\lambda}$.

- G_1 to G_0: The probability preservation property in Lemma 1 gives:

$$\mathsf{Adv}_{G_0}[\mathcal{A}(H_0) = \mathbf{s}] \leqslant Adv_{G_1}[\mathcal{A}(H_1) = \mathbf{s}] + \Delta(H_0, H_1).$$

  To compute the statistical distance between H_0 and H_1, we need to evaluate the statistical distance of the third component of these two distributions, namely:

$$\Delta(\varphi_{\mathbf{As}\oplus\mathbf{e}} + \mathcal{D}_{\sigma_\eta}, \varphi_{\mathbf{As}} + \mathbf{d} + \mathcal{D}_{\sigma_\eta}).$$

  As Lemma 3 says, we have $|\ \mathrm{HW}(\mathbf{As}) - \mathrm{HW}(\mathbf{b})\ | \leqslant \lceil \log_2(\sqrt{\lambda}\sigma_e)\rceil$ with high probability. Hence, the third component $\mathrm{HW}(\mathbf{As} \oplus \mathbf{e}) + \eta$ of H_0 can be substituted by an element of the form $\mathrm{HW}(\mathbf{As}) + \mathbf{d} + \eta$, where $\|\mathbf{d}_i\| \leqslant \lceil \log_2(\sqrt{\lambda}\sigma_e)\rceil$. Hence, $\Delta(H_0, H_1) \leqslant \varepsilon$.

Putting all together, we obtain:

$$\begin{aligned}\mathsf{Adv}_{G_0}[\mathcal{A}(H_0) = \mathbf{s}] &\leqslant \mathsf{Adv}_{G_1}[\mathcal{A}(H_1) = \mathbf{s}] + \Delta(H_0, H_1), \\ &\leqslant \mathsf{Adv}_{G_2}[\mathcal{A}(H_2) = \mathbf{s}] + \Delta(H_1, H_2) + \Delta(H_0, H_1).\end{aligned}$$

Finally $\mathcal{B}$ simulates true NHW-LWE samples for $\mathcal{A}$. Subsequently, if $\mathcal{A}$ is an adversary against NHW-LWE, then $\mathcal{B}$ is an adversary against LWE, which concludes the proof. □

Remark 2. Concretely, the bound on σ_η depends on how we can erase the LWE error dependency $\mathbf{d}$ (i.e., the difference $\mathrm{HW}(\mathbf{b}) - \mathrm{HW}(\mathbf{As})$) thanks to additive noise. This is determined by the Gaussian approximation of the difference $(\mathrm{HW}(\mathbf{b}) + \eta) - (\mathrm{HW}(\mathbf{As}) + \eta')$, where $\eta, \eta' \xleftarrow{\$} \mathcal{D}_{\sigma_\eta}$. The closer we are to a Gaussian distribution, the less information on $\mathbf{d}$ we obtain. Therefore, we could also conclude the proof of Theorem 1 by using the statistical distance $\Delta\left((\varphi_{\mathbf{b}} + \mathcal{D}_{\sigma_\eta}) - (\varphi_{\mathbf{As}} + \mathcal{D}\sigma_\eta), \mathcal{D}_s\right)$ estimated directly instead of using game G_1. This finer-grain characterization will be discussed in Section 5.

Theorem 1 gives a reduction from the search LWE problem to the search NHW-LWE problem, using the bound on the statistical distance from Corollary 1. It can also be concluded using the Rényi divergence (replacing the additive property by a multiplicative one as proposed by Bai et al. in [BLL+15]). This leads to a probability preservation property as soon as the Rényi divergence $\mathrm{RD}_2(\varphi_{\mathbf{As}\oplus\mathbf{e}} + \mathcal{D}_{\sigma_\eta} \| \varphi_{\mathbf{b}} + \mathcal{D}_{\sigma_\eta}) \leqslant \mathsf{poly}(1/\varepsilon)$ for a negligible ε. By considering the Rényi divergence instead of the statistical distance, we obtain a nicer bound on σ_η, which we reflect with the following corollary to Theorem 1.

Corollary 2. *Let λ be a security parameter, $n, m \in \mathbb{N}$, $q \in \mathbb{N}$, $\sigma_e, \sigma_\eta > 0$. Let $\varepsilon = \mathcal{O}\left(\frac{1}{m}\right) < 1/2$, such that $\sigma_\eta \geqslant \eta_\varepsilon(\mathbb{Z})$. If $\sigma_\eta \geqslant \sqrt{m}\lceil \log_2(\sqrt{\lambda}\sigma_e)\rceil$ and $\sqrt{\lambda}\sigma_\eta \leqslant q - \ell_q$, then there exists a reduction from $s\,\mathrm{LWE}(\lambda)_{n,m,q,\sigma_e}$ to $\mathrm{sNHW\text{-}LWE}(\lambda)_{n,m,q,\sigma_e,\sigma_\eta}$.*

Proof. As in Theorem 1, let us define the three following distributions:

- $H_0 := (\mathbf{A}, \mathbf{As} \oplus \mathbf{e}, \mathrm{HW}(\mathbf{As} \oplus \mathbf{e}) + \eta)$ where $\mathbf{A}, \mathbf{s}, \mathbf{e}$ are LWE samples and $\eta \xleftarrow{\$} \mathcal{D}^m_{\sigma_\eta}$;
- $H_1 := (\mathbf{A}, \mathbf{As} \oplus \mathbf{e}, \mathrm{HW}(\mathbf{As}) + \mathbf{d} + \eta)$ where $\|\mathbf{d}_i\| \leqslant \lceil \log(\sqrt{\lambda}\sigma_e) \rceil$ for all $0 \leqslant i < m$;
- $H_2 := (\mathbf{A}, \mathbf{As} \oplus \mathbf{e}, \mathrm{HW}(\mathbf{As}) + \eta')$ where $\eta' \xleftarrow{\$} \mathcal{D}^m_{\sigma_\eta}$.

For the rest of the proof, we consider G_i the problem of finding $\mathbf{s}$ given the distribution H_i for $0 \leqslant i \leqslant 2$. The distribution H_0 is the sample that $\mathcal{B}$ constructs and sends to $\mathcal{A}$. For all $0 \leqslant i \leqslant 2$, we want that $\mathsf{Adv}_{G_i}[\mathcal{A}(H_i)]$ is negligible for the problem to be hard. Then, showing G_2 is hard suffices to show that G_0 is hard. Step by step, it means:

- G_2 is the game of finding $\mathbf{s}$ knowing a true NHW-LWE sample. So far, $\mathcal{A}$ corresponds to its definition, namely an adversary against NHW-LWE.
- G_2 to G_1: In contrast with the previous proof, the advantage of $\mathcal{A}$ in G_1 changes using Rényi divergence. The probability preservation property in Lemma 1 of the Rényi divergence gives:

$$\mathsf{Adv}_{G_1}[\mathcal{A}(H_1) = \mathbf{s}]^2 \leqslant \mathsf{Adv}_{G_2}[\mathcal{A}(H_2) = \mathbf{s}]\mathrm{RD}(H_1 \| H_2).$$

  To evaluate the Rényi divergence of H_1 and H_2, we need to evaluate the Rényi divergence of $\mathcal{D}^m_{\sigma_\eta, \mathbf{d}}$ and $\mathcal{D}^m_{\sigma_\eta}$ as in the proof of [BJRW20, Theorem 1]. We know by Lemma 3 that $\|\mathbf{d}_i\| \leqslant \lceil \log(\sqrt{\lambda}\sigma_e) \rceil$ for all $0 \leqslant i < m$. Then, it suffices to compute the Rényi divergence of $(\mathcal{D}_{\sigma_\eta q, \lceil \log(\sqrt{\lambda}\sigma_e) \rceil})^m$ and $(\mathcal{D}_{\sigma_\eta q})^m$. Using that $\sigma_\eta \geqslant \eta_\varepsilon(\mathbb{Z})$, by multiplicativity of the Rényi divergence and Lemma 2, we have:

$$\begin{aligned} \mathrm{RD}_2(\mathcal{D}^m_{\sigma_\eta, \lceil \log(\sqrt{\lambda}\sigma_e) \rceil} \| \mathcal{D}^m_{\sigma_\eta}) &= \mathrm{RD}_2\left(\mathcal{D}_{\sigma_\eta, \lceil \log(\sqrt{\lambda}\sigma_e) \rceil} \| \mathcal{D}_{\sigma_\eta}\right)^m, \\ &\leqslant \exp(8m\varepsilon)\exp(2\pi), \end{aligned}$$

  where the second line comes from Corollary 1 and $\sigma_\eta \geqslant \sqrt{m}\lceil \log_2(\sqrt{\lambda}\sigma_e) \rceil$. From that, for the Rényi divergence to be bounded by a constant, $\varepsilon = \mathcal{O}\left(\frac{1}{m}\right)$ should be sufficient. Then, we finally obtain:

$$\begin{aligned} \mathsf{Adv}_{G_1}[\mathcal{A}(H_1) = \mathbf{s}]^2 &\leqslant \mathsf{Adv}_{G_2}[\mathcal{A}(H_2) = \mathbf{s}]\mathrm{RD}(H_1 \| H_2), \\ &\leqslant \mathsf{Adv}_{G_2}[\mathcal{A}(H_2) = \mathbf{s}]\mathcal{O}(1), \\ \mathsf{Adv}_{G_1}[\mathcal{A}(H_1) = \mathbf{s}] &\leqslant \sqrt{\mathsf{Adv}_{G_2}[\mathcal{A}(H_2) = \mathbf{s}]\mathcal{O}(1)}. \end{aligned}$$

- G_1 to G_0: As in proof of Theorem 1, to compute the statistical distance between H_0 and H_1, we need to evaluate the statistical distance of the third component of this two distributions, namely $\Delta(\varphi_{\mathbf{As}\oplus\mathbf{e}} + \mathcal{D}_{\sigma_\eta}, \varphi_{\mathbf{As}} + \mathbf{d} + \mathcal{D}_{\sigma_\eta})$. As Lemma 3 says, we have $|\ \mathrm{HW}(\mathbf{As}) - \mathrm{HW}(\mathbf{b})\ | \leqslant \lceil \log_2(\sqrt{\lambda}\sigma_e) \rceil$ with high probability. Therefore, the third component $\mathrm{HW}(\mathbf{As} \oplus \mathbf{e}) + \eta$ of the distribution H_0 can be substituted by an element of the form $\mathrm{HW}(\mathbf{As}) + \mathbf{d} + \eta$, where $\|\mathbf{d}_i\| \leqslant \lceil \log_2(\sqrt{\lambda}\sigma_e) \rceil$. As a result, we have $\Delta(H_0, H_1) \leqslant \varepsilon$, and:

$$\begin{aligned} \mathsf{Adv}_{G_0}[\mathcal{A}(H_0) = \mathbf{s}] &\leqslant \mathsf{Adv}_{G_1}[\mathcal{A}(H_1) = \mathbf{s}] + \Delta(H_0, H_1), \\ &\leqslant \mathsf{Adv}_{G_1}[\mathcal{A}(H_1) = \mathbf{s}] + \varepsilon. \end{aligned}$$

Putting all together, we obtain:

$$\begin{aligned} \mathsf{Adv}_{G_0}[\mathcal{A}(H_0) = \mathbf{s}] &\leqslant \mathsf{Adv}_{G_1}[\mathcal{A}(H_1) = \mathbf{s}] + \Delta(H_0, H_1) \\ &\leqslant \sqrt{\mathsf{Adv}_{G_2}[\mathcal{A}(H_2) = \mathbf{s}]\mathcal{O}(1)} + \Delta(H_0, H_1). \end{aligned}$$

Once again, $\mathcal{B}$ simulates true NHW-LWE samples for $\mathcal{A}$. If $\mathcal{A}$ is an adversary against NHW-LWE, then $\mathcal{B}$ is an adversary against LWE, which concludes the proof. □

Remark 3. The LWE samples are based on uniform secrets. However, as recent lattice-based schemes present, we could also consider the problem for bounded secrets ($\|\mathbf{s}_i\| \leqslant \gamma$). In fact, Dilithium [LDK+20] and Kyber [SAB+20] both present bounded secrets with relatively short distributions (i.e., $\gamma \in \{2, 3, 4\}$). Consequently, the entropy of the distribution of the secret is already small and drops too much additional information given the Hamming weights $\mathrm{HW}(\mathbf{As})$. Yet, we could use the same arguments to conclude on the hardness of NHW-LWE with bounded secrets under the LWE with bounded secrets assumption, by considering bounded secrets in the distribution we defined in the proof.

Note that the proof of Corollary 2 has two limitations in the decisional case. The first one is the use of the Rényi divergence, that we can replace by the statistical distance like in Theorem 1 with a loss on the condition of the parameters. The second and essential one also applies to Theorem 1. As the decisional case asks, we have to transform a sample of the form $(\mathbf{A}, \mathbf{b})$ where $\mathbf{b}$ is uniform in $\mathbb{Z}_q^m$ in the same way as a sample $(\mathbf{A}, \mathbf{As} \oplus \mathbf{e})$.In this case, we target a hint of the form $\mathrm{HW}(\mathbf{As}) + \eta$ where $\mathbf{s} \in \mathbb{Z}_q^n$. But as $\mathbf{b}$ is uniform, we would need a nice bound on $\mathbf{d}$ knowing $\mathbf{A}$ that does not seem straightforward to find.

5 Discussion and open problems

As is, the results of Theorem 1 and Corollary 2 directly lead to the first improvement claimed in introduction. Namely, the level of additive Gaussian noise needed for the reductions to hold is independent of the field size q. By contrast, the second improvement (i.e., the amplified impact of the noise towards hiding the difference / dependency $\mathbf{d}$) is not directly reflected by the theorem statements. Specifically, the noise level in Theorem 1 must be large enough to ensure that $\frac{\lceil \log_2(\sqrt{\lambda}\sigma_e) \rceil}{\sigma_\eta}$ is negligible. While Corollary 2 moderates this requirement when the adversary is bounded to observe a small number of samples m, it ultimately demonstrates a trend similar to a naive entropic bound: compensating for an increase in m still requires a proportional increase in the physical noise variance.

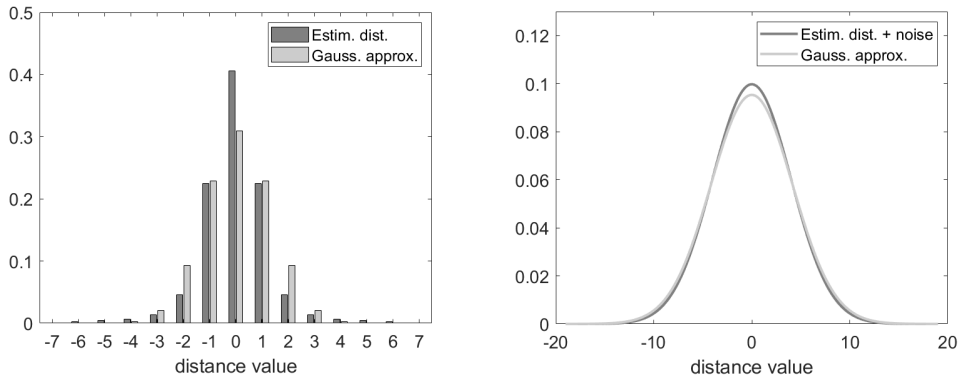
We now put forward that a finer-grain analysis for the most critical step of our reductions can lead to a better outcome. For this purpose, we build on the observation from Remark 2 that the main driver of the noise requirements in Theorem 1 is the statistical distance between the distribution of $\mathbf{d}$ with additive noise and its Gaussian approximation. The square of the statistical distance between these two distributions gives a bound on the number of samples needed to distinguish them. Hence, if this distance is bounded by $2^{-\lambda}$, the security of LWPRN is maintained as long as $m \ll 2^\lambda$. Without further specialization, the trend for decreasing the statistical distance between the "noisy $\mathbf{d}$" distribution and its Gaussian approximation is still inversely proportional to the noise variance. We show next how to improve it by better characterizing the distribution of the noisy $\mathbf{d}$.

As a first step in this direction, Figure 1(a) shows an exemplary distribution of noise-free $\mathbf{d}$ and its (discretized) Gaussian approximation (for $q = 127$ and $\sigma_e = 2$). Figure 1(b) shows a similar plot for a noisy $\mathbf{d}$. The important observation is that both the distribution of (possibly noisy) $\mathbf{d}$ and its Gaussian approximation are symmetric.

We next establish this result in a slightly more general manner with Lemma 4.

Lemma 4. *Let $q > 0$, U be a random variable that follows a uniform distribution over $\mathbb{Z}_q$, E be a random variable independent from U that follows a symmetric distribution over $\mathbb{Z}_q$, and $\mathsf{f} : \mathbb{Z}_q \to \mathbb{R}$ be any function, then the distribution of $\mathsf{f}(U + E) - \mathsf{f}(U)$ is symmetric.*

Proof. $\forall x \in \mathbb{R}$, we have:



(a) Noise-free distributions. (b) Noisy distributions

Figure 1: Distribution of $\mathbf{d} = \mathrm{HW}(\mathbf{As} \oplus \mathbf{e}) - \mathrm{HW}(\mathbf{As})$ and its Gaussian approximation.

$$
\begin{aligned}
\Pr\left(\mathsf{f}(U+E) - \mathsf{f}(U) = x\right) &= \sum_{e \in \mathbb{Z}_q} \sum_{u \in \mathbb{Z}_q} \Pr\left(U = u, E = e\right) \cdot \mathbf{1}_{\mathsf{f}(u+e)-\mathsf{f}(u)=x} \\
&= \sum_{e \in \mathbb{Z}_q} \sum_{u \in \mathbb{Z}_q} \Pr\left(U = u\right) \cdot \Pr\left(E = e\right) \cdot \mathbf{1}_{\mathsf{f}(u+e)-\mathsf{f}(u)=x} && (U, E \text{ ind.}), \\
&= \frac{1}{q} \sum_{e \in \mathbb{Z}_q} \sum_{u \in \mathbb{Z}_q} \Pr\left(E = e\right) \cdot \mathbf{1}_{\mathsf{f}(u+e)-\mathsf{f}(u)=x} && (U \text{ unif.}), \\
&= \frac{1}{q} \sum_{e \in \mathbb{Z}_q} \sum_{u \in \mathbb{Z}_q} \Pr\left(E = -e'\right) \cdot \mathbf{1}_{\mathsf{f}(u-e')-\mathsf{f}(u)=x} && (e' = -e), \\
&= \frac{1}{q} \sum_{e' \in \mathbb{Z}_q} \sum_{u \in \mathbb{Z}_q} \Pr\left(E = e'\right) \cdot \mathbf{1}_{\mathsf{f}(u-e')-\mathsf{f}(u)=x} && (\text{E sym.}), \\
&= \frac{1}{q} \sum_{e' \in \mathbb{Z}_q} \sum_{u' \in \mathbb{Z}_q} \Pr\left(E = e'\right) \cdot \mathbf{1}_{\mathsf{f}(u')-\mathsf{f}(u'+e')=x} && (u' = u - e'), \\
&= \Pr\left(\mathsf{f}(U+E) - \mathsf{f}(U) = -x\right).
\end{aligned}
$$

□

It directly leads to the following corollary that corresponds to our case study:

Corollary 3. *Let* $\mathbf{A}, \mathbf{s}$ *and* $\mathbf{e}$ *be LWE variables defined as in Definition 1. Let* $\mathbf{e}$ *follow a centered discrete Gaussian and let* HW *be the leakage function defined as in Section 3.2. Then the distribution followed by* $\mathbf{d} = HW(\mathbf{As} \oplus \mathbf{e}) - HW(\mathbf{As})$ *is symmetric.*

Proof. Since $\mathbf{s} \neq \mathbf{0}$, each coefficient of $\mathbf{As}$ is uniform. Each coefficient of $\mathbf{e}$ follows a centered discrete Gaussian distribution, which is symmetric. The HW leakage function operates coefficient-wise. Therefore, we can apply Lemma 4 to each coefficient of the distribution followed by $\mathrm{HW}(\mathbf{As} \oplus \mathbf{e}) - \mathrm{HW}(\mathbf{As})$. □

Leveraging the symmetry of the distributions that must be made indistinguishable, we can rely on a standard argument used in the side-channel literature, based on the concept of *statistical security order* frequently used to analyze the masking countermeasure [BDF+17]. For this purpose, let us consider an unprotected implementation where information can be extracted based on difference of means for the leakages of an intermediate value x, as in Kocher et al.'s original DPA [KJJ99]. In this case, doubling the noise variance will

double the amount of traces needed for the attack to succeed. The idea of masking is to split x in shares $x_1, x_2, \ldots, x_d$ and to perform the computations on shared values only. Take $d = 2$ and assume x is a bit for simplicity. It implies that the adversary willing to extract information now has to observe a bivariate leakage $\big(\mathsf{HW}(x_1) + \eta_1, \mathsf{HW}(x_2) + \eta_2\big)$ in a serial implementation, or a univariate leakage $\mathsf{HW}(x_1) + \mathsf{HW}(x_2) + \eta$ in a parallel implementation. The important observation is that the means of these leakages are independent of x. Namely, in our single-bit example, the mean vector of the serial implementation is equal to $[\frac{1}{2}, \frac{1}{2}]$ and the mean of the parallel implementation is worth 1. Therefore, when leakages are sufficiently noisy, it is expected that the adversary's best strategy is to estimate their (co)variance, which depends on x, to extract information. Since estimating a variance implies squaring the leakages, it entails that the noise will be squared too. So asymptotically (i.e., ignoring the pathological case where the signal is larger than the noise), doubling the noise will multiply the attack complexity by a factor 2^2. Generalizing this to more shares, the statistical moments up to order $d - 1$ will be independent of the of the secret and the attack complexity when doubling the noise will be multiplied by 2^d, with d the statistical security order. Concretely, this order is reflected by the slope of so-called "information theoretic curves" where we plot the attack complexity in function of the noise variance (in log scale) [DFS19], which we will observe next.

Stepping back to the context of our reduction, it turns out the noisy $\mathbf{d}$ distribution and its Gaussian approximation have the same mean, variance and skewness, so that they only start to differ in their kurtosis (i.e., their 4th-order moment). As a result, for a large enough noise variance σ_η^2, the sampling complexity of estimating the 4th-order statistical moment, needed to distinguish the noisy $\mathbf{d}$ distribution and its Gaussian approximation, scales with $(\sigma_\eta^2)^4$. This provides a much better trend than both the generic argument above and the naive entropic bound mentioned in introduction. In other words, the reduction of Theorem 1 allows identifying a security condition on the noise that benefits from direct estimation and is easier to satisfy (thanks to the symmetry observation of Lemma 4) than if we were directly trying to limit the key leakage per sample with an entropic argument, as outlined in introduction, where the security level only scales with (σ_η^2).

To confirm this trend empirically, we computed the statistical distance between the noisy $\mathbf{d}$ distribution and its Gaussian approximation, in function of the SNR and for different Mersenne primes q, in Figure 2.[6] The noise amplification is directly reflected by the slope -4 on this log-scaled plot. As is clear from the bounds in Theorem 1 and Corollary 2, this trend is independent of both the field size q and the standard deviation of the LWE errors σ_e. Since the signal (variance) of a Hamming weight leakage function increases with q (i.e., is approximately is worth $\approx \log_2(q)/4$), this explains why the level of noise necessary to reach a certain statistical distance decreases with the field size. Concretely, it for example means that for $q = 2^{31} - 1$ (which is popular for re-keying applications in symmetric cryptography), a SNR of 10^{-2} allows tolerating up to $m = 10^{12}$ samples.

In order to complement the results obtained for an idealized (Hamming weight) leakage function, we performed a similar empirical evaluation for a more realistic regression-based leakage model [SLP05], which generalizes the Hamming weight function as follows:

$$\mathsf{L}(x) = \sum_{i=1}^{\lceil \log_2(q) \rceil} \alpha_i \cdot x(i),$$

where the α_i's are real-valued coerfficients and the $x(i)$'s are the bits of x. In the Hamming weight case, all the α_i coefficients are equal to one. In practical implementations, they tend to deviate. Concretely, we borrowed the coefficients corresponding to an FPGA

[6] Very low statistical distance values can be computed by leveraging the fact that the noise is independent of $\mathbf{d}$. It allows computing the noisy $\mathbf{d}$ distributions efficiently by using convolution products.

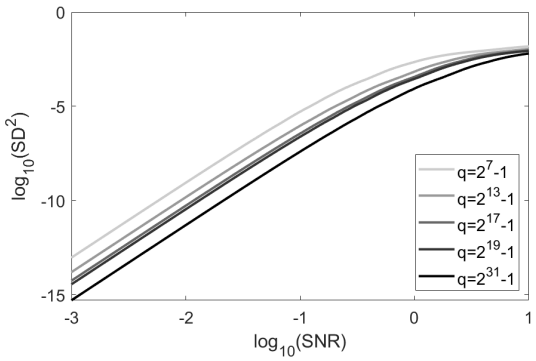
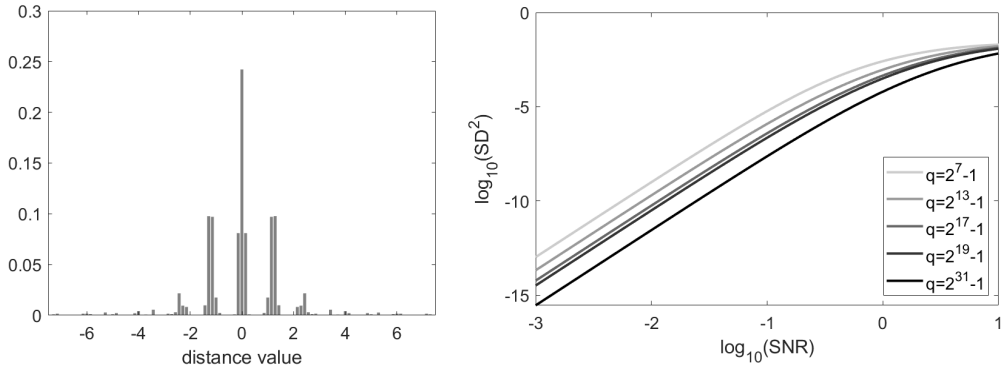


Figure 2: Squared SD between a noisy $\mathbf{d}$ and its Gaussian approx. (HW model).

implementation of LWPR prototype proposed in [HMM+23] for this purpose. Figure 3(a) shows how the (noise-free) distribution of Figure 1(a) evolves in this case. Figure 3(b) shows that the trends and values of the Hamming weight case are essentially confirmed. Note that for the trends (i.e., the slope of -4), this is a guaranteed result due to Lemma 4. As for the values, it is likely a consequence of relatively similar leakage models.



(a) Noise-free distribution.

(b) Information theoretic curves.

Figure 3: Squared SD between a noisy $\mathbf{d}$ and its Gaussian approx. (FPGA model).

Note that the statistical argument given above does not directly/formally imply security in the noisy leakage model, as reflected by the SD slopes of Figures 2 and 3: it only gives a plausible explanation to the observed trends. This is because a statistical distribution is not always defined by its statistical moments [BDF+17]. Yet, as in the context of masking, we expect that this link holds for all the distributions that can be observed in practice. The formalization of this expectation is an interesting scope for further research.

Overall, our results therefore show a first reduction from a standard learning problem to a noisy variant of the LWPR problem that applies to serial implementations. We hope they will trigger more investigations in order to improve the security guarantees of such hard physical learning problems. For example, it would be interesting to better characterize the algebraic complexity of LWPR, thanks to improved cryptanalysis or reductions. Proving the hardness of the decisional problem would be relevant for many applications as well. It would require addressing the limitation mentioned at the end of the previous section. A potential tweak in this direction could be to try reaching (e.g., Gaussian) hints of which the treatment can benefit from known tools. Specializing such investigations to

Mersenne primes or Power of 2 modulus could also be useful for this purpose. Meanwhile, we conclude this discussion with the practically-relevant benefit that LWPRN offers a 4th-order security guarantee at the cost of an unprotected implementation.

Appendices

A More details on prior works

Most of the times, the leakage function L is non-linear in the field/ring, and operates either on the secret, the error or any other computation happening during the processing algorithm. Subsequently, as the LWPRN problem considers real-valued leakage functions, it cannot rely on LWE variants with such additional linear leakage.

The first of these is probably the Hint-LWE variant, proposed by Kim et al. [KLSS23]. This variant appears as the best solution to handle linear hints over both secret and error of the form $\gamma \cdot (\mathbf{s^t}, \mathbf{e^t}) \oplus \mathbf{y}$ where $\mathbf{y} \in \mathbb{Z}_q^{(n+m)\times l}$ is a small error and $\gamma \in \mathbb{Z}_q^l$ is the leakage vector. However, this type of hint is limited by the fact that the leakage vector γ is not controlled by the distinguisher (and structured in the Module case).

Recently, Russel et al. [LSW25] studied the hardness of a variant called Leaky-LWE, where the adversary also observes additional noisy linear leakages $(\mathbf{s}^t, \mathbf{e}^t)\mathbf{L} + \mathbf{f}^t$ of the LWE secret and error, where $\mathbf{L}$ is a leakage matrix in $\mathbb{Z}_q^{(n+m)\times l}$ and $\mathbf{f} \in \mathbf{Z}_q^l$ is a small error. The major difference with the Hint-LWE variant lies in the choice of the leakage matrix $\mathbf{L}$, chosen short by the adversary (and possibly unstructured in the Module case). The paper results in the hardness of the Leaky-LWE variant under the LWE assumption. Again, the noise of the leakage is considered in the Ring or the Module. This ensure a tight result over the noise condition, but does not apply to real-valued leakages.

To handle the problem of realistic more leakages, Dachman-Sled et al. [DDGR20] try to consider multiple hints on the secret, as described in the introduction. Besides being a cryptanalytic result rather than a reduction, they integrate specific hints specifically into lattice-based attacks which do not correspond to the physical leakages in our work.

In 2020, Brakerski and Döttling [BD20] introduced the Entropic-LWE variant to extend the short-secret LWE variant treated by Goldwasser et al. [GKPV10] in 2012. This variant answers the question of the hardness of LWE with additional (possibly non-linear) knowledge on the secret, by conducting an analysis of the remaining entropy of the secret in the presence of any additional information over its distribution. Their result requires a precise entropic analysis of the secret conditioned on additional knowledge. Recently, Boudgoust et al. [BJTW25] extended this previous result by conducting an entropic analysis on the secret distribution with additional linear leakages. It leads to a stronger intermediate between the hint variant and the entropic one, providing the first hardness result for both general bounded error distribution (any other than Gaussian distribution) and non-linear leakage functions. However, the resulting bound does not conclude in the case of Hamming Weight leakage, as stated in the introduction.

Acknowledgment. Emeline Repel is partially funded by the region of Normandy, France. François-Xavier Standaert is a research director of the Belgian Fund for Scientific Research (FNRS-F.R.S.). This work is supported by the PEPR quantique France 2030 program (ANR-22-PETQ-0008 PQ-TLS), by the ASTRID program under the national project AMIRAL with reference ANR-21-ASTR-0016 and by the ERC project 101096871 (BRIDGE). Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

References

[AP12] Jacob Alperin-Sheriff and Chris Peikert. Circular and KDM security for identity-based encryption. In Marc Fischlin, Johannes Buchmann, and Mark Manulis, editors, *PKC 2012*, volume 7293 of *LNCS*, pages 334–352. Springer, Berlin, Heidelberg, May 2012.

[BCF+15] Sonia Belaïd, Jean-Sébastien Coron, Pierre-Alain Fouque, Benoît Gérard, Jean-Gabriel Kammerer, and Emmanuel Prouff. Improved side-channel analysis of finite-field multiplication. In Tim Güneysu and Helena Handschuh, editors, *CHES 2015*, volume 9293 of *LNCS*, pages 395–415. Springer, Berlin, Heidelberg, September 2015.

[BD20] Zvika Brakerski and Nico Döttling. Hardness of LWE on general entropic distributions. In Anne Canteaut and Yuval Ishai, editors, *EUROCRYPT 2020, Part II*, volume 12106 of *LNCS*, pages 551–575. Springer, Cham, May 2020.

[BDF+17] Gilles Barthe, François Dupressoir, Sebastian Faust, Benjamin Grégoire, François-Xavier Standaert, and Pierre-Yves Strub. Parallel implementations of masking schemes and the bounded moment leakage model. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *EUROCRYPT 2017, Part I*, volume 10210 of *LNCS*, pages 535–566. Springer, Cham, April / May 2017.

[BFG14] Sonia Belaïd, Pierre-Alain Fouque, and Benoît Gérard. Side-channel analysis of multiplications in GF(2^{128}) - application to AES-GCM. In Palash Sarkar and Tetsu Iwata, editors, *ASIACRYPT 2014, Part II*, volume 8874 of *LNCS*, pages 306–325. Springer, Berlin, Heidelberg, December 2014.

[BGG+14] Josep Balasch, Benedikt Gierlichs, Vincent Grosso, Oscar Reparaz, and François-Xavier Standaert. On the cost of lazy engineering for masked software implementations. In *CARDIS*, volume 8968 of *Lecture Notes in Computer Science*, pages 64–81. Springer, 2014.

[BIP+18] Dan Boneh, Yuval Ishai, Alain Passelègue, Amit Sahai, and David J. Wu. Exploring crypto dark matter: New simple PRF candidates and their applications. In Amos Beimel and Stefan Dziembowski, editors, *TCC 2018, Part II*, volume 11240 of *LNCS*, pages 699–729. Springer, Cham, November 2018.

[BJRW20] Katharina Boudgoust, Corentin Jeudy, Adeline Roux-Langlois, and Weiqiang Wen. Towards classical hardness of module-LWE: The linear rank case. In Shiho Moriai and Huaxiong Wang, editors, *ASIACRYPT 2020, Part II*, volume 12492 of *LNCS*, pages 289–317. Springer, Cham, December 2020.

[BJTW25] Katharina Boudgoust, Corentin Jeudy, Erkan Tairi, and Weiqiang Wen. Hardness of M-LWE with general distributions and applications to leaky variants. Cryptology ePrint Archive, Report 2025/1472, 2025.

[BLL+15] Shi Bai, Adeline Langlois, Tancrède Lepoint, Damien Stehlé, and Ron Steinfeld. Improved security proofs in lattice-based cryptography: Using the Rényi divergence rather than the statistical distance. In Tetsu Iwata and Jung Hee Cheon, editors, *ASIACRYPT 2015, Part I*, volume 9452 of *LNCS*, pages 3–24. Springer, Berlin, Heidelberg, November / December 2015.

[BLP+13] Zvika Brakerski, Adeline Langlois, Chris Peikert, Oded Regev, and Damien Stehlé. Classical hardness of learning with errors. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *45th ACM STOC*, pages 575–584. ACM Press, June 2013.

[BPR12] Abhishek Banerjee, Chris Peikert, and Alon Rosen. Pseudorandom functions and lattices. In David Pointcheval and Thomas Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 719–737. Springer, Berlin, Heidelberg, April 2012.

[CGP+12] Jean-Sébastien Coron, Christophe Giraud, Emmanuel Prouff, Soline Renner, Matthieu Rivain, and Praveen Kumar Vadnala. Conversion of security proofs from one leakage model to another: A new issue. In Werner Schindler and Sorin A. Huss, editors, *COSADE 2012*, volume 7275 of *LNCS*, pages 69–81. Springer, Berlin, Heidelberg, May 2012.

[dCGRP19] Eloi de Chérisey, Sylvain Guilley, Olivier Rioul, and Pablo Piantanida. Best information is most successful. *IACR TCHES*, 2019(2):49–79, 2019.

[DDF14] Alexandre Duc, Stefan Dziembowski, and Sebastian Faust. Unifying leakage models: From probing attacks to noisy leakage. In Phong Q. Nguyen and Elisabeth Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 423–440. Springer, Berlin, Heidelberg, May 2014.

[DDGR20] Dana Dachman-Soled, Léo Ducas, Huijing Gong, and Mélissa Rossi. LWE with side information: Attacks and concrete security estimation. In Daniele Micciancio and Thomas Ristenpart, editors, *CRYPTO 2020, Part II*, volume 12171 of *LNCS*, pages 329–358. Springer, Cham, August 2020.

[DFH+16] Stefan Dziembowski, Sebastian Faust, Gottfried Herold, Anthony Journault, Daniel Masny, and François-Xavier Standaert. Towards sound fresh re-keying with hard (physical) learning problems. In Matthew Robshaw and Jonathan Katz, editors, *CRYPTO 2016, Part II*, volume 9815 of *LNCS*, pages 272–301. Springer, Berlin, Heidelberg, August 2016.

[DFS15] Alexandre Duc, Sebastian Faust, and François-Xavier Standaert. Making masking security proofs concrete - or how to evaluate the security of any leaking device. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part I*, volume 9056 of *LNCS*, pages 401–429. Springer, Berlin, Heidelberg, April 2015.

[DFS19] Alexandre Duc, Sebastian Faust, and François-Xavier Standaert. Making masking security proofs concrete (or how to evaluate the security of any leaking device), extended version. *Journal of Cryptology*, 32(4):1263–1297, October 2019.

[DKM+15] Christoph Dobraunig, François Koeune, Stefan Mangard, Florian Mendel, and François-Xavier Standaert. Towards fresh and hybrid re-keying schemes with beyond birthday security. In *CARDIS*, volume 9514 of *Lecture Notes in Computer Science*, pages 225–241. Springer, 2015.

[DMMS21] Sébastien Duval, Pierrick Méaux, Charles Momin, and François-Xavier Standaert. Exploring crypto-physical dark matter and learning with physical rounding. *IACR TCHES*, 2021(1):373–401, 2021.

[GJ19] Qian Guo and Thomas Johansson. A new birthday-type algorithm for attacking the fresh re-keying countermeasure. *Inf. Process. Lett.*, 146:30–34, 2019.

[GKPV10] Shafi Goldwasser, Yael Tauman Kalai, Chris Peikert, and Vinod Vaikuntanathan. Robustness of the learning with errors assumption. In Andrew Chi-Chih Yao, editor, *ICS 2010*, pages 230–240. Tsinghua University Press, January 2010.

[HLM+23] Clément Hoffmann, Benoît Libert, Charles Momin, Thomas Peters, and François-Xavier Standaert. POLKA: Towards leakage-resistant post-quantum CCA-secure public key encryption. In Alexandra Boldyreva and Vladimir Kolesnikov, editors, *PKC 2023, Part I*, volume 13940 of *LNCS*, pages 114–144. Springer, Cham, May 2023.

[HMM+23] Clément Hoffmann, Pierrick Méaux, Charles Momin, Yann Rotella, François-Xavier Standaert, and Balazs Udvarhelyi. Learning with physical rounding for linear and quadratic leakage functions. In Helena Handschuh and Anna Lysyanskaya, editors, *CRYPTO 2023, Part III*, volume 14083 of *LNCS*, pages 410–439. Springer, Cham, August 2023.

[ISW03] Yuval Ishai, Amit Sahai, and David Wagner. Private circuits: Securing hardware against probing attacks. In Dan Boneh, editor, *CRYPTO 2003*, volume 2729 of *LNCS*, pages 463–481. Springer, Berlin, Heidelberg, August 2003.

[KJJ99] Paul C. Kocher, Joshua Jaffe, and Benjamin Jun. Differential power analysis. In Michael J. Wiener, editor, *CRYPTO'99*, volume 1666 of *LNCS*, pages 388–397. Springer, Berlin, Heidelberg, August 1999.

[KLSS23] Duhyeong Kim, Dongwon Lee, Jinyeong Seo, and Yongsoo Song. Toward practical lattice-based proof of knowledge from hint-MLWE. In Helena Handschuh and Anna Lysyanskaya, editors, *CRYPTO 2023, Part V*, volume 14085 of *LNCS*, pages 549–580. Springer, Cham, August 2023.

[LDK+20] Vadim Lyubashevsky, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Peter Schwabe, Gregor Seiler, Damien Stehlé, and Shi Bai. CRYSTALS-DILITHIUM. Technical report, National Institute of Standards and Technology, 2020. available at https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization/round-3-submissions.

[LSS14] Adeline Langlois, Damien Stehlé, and Ron Steinfeld. GGHLite: More efficient multilinear maps from ideal lattices. In Phong Q. Nguyen and Elisabeth Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 239–256. Springer, Berlin, Heidelberg, May 2014.

[LSW25] Russell W. F. Lai, Monisha Swarnakar, and Ivy K. Y. Woo. Leaky LWE: Learning with errors with semi-adaptive secret- and error-leakage. *IACR Communications in Cryptology*, 2(3), 2025.

[Lyu12] Vadim Lyubashevsky. Lattice signatures without trapdoors. In David Pointcheval and Thomas Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 738–755. Springer, Berlin, Heidelberg, April 2012.

[Man04] Stefan Mangard. Hardware countermeasures against DPA – A statistical analysis of their effectiveness. In Tatsuaki Okamoto, editor, *CT-RSA 2004*, volume 2964 of *LNCS*, pages 222–235. Springer, Berlin, Heidelberg, February 2004.

[MOP07] Stefan Mangard, Elisabeth Oswald, and Thomas Popp. *Power analysis attacks - revealing the secrets of smart cards*. Springer, 2007.

[MPG05] Stefan Mangard, Thomas Popp, and Berndt M. Gammel. Side-channel leakage of masked CMOS gates. In Alfred Menezes, editor, *CT-RSA 2005*, volume 3376 of *LNCS*, pages 351–365. Springer, Berlin, Heidelberg, February 2005.

[MR04] Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on Gaussian measures. In *45th FOCS*, pages 372–381. IEEE Computer Society Press, October 2004.

[MSGR10] Marcel Medwed, François-Xavier Standaert, Johann Großschädl, and Francesco Regazzoni. Fresh re-keying: Security against side-channel and fault attacks for low-cost devices. In Daniel J. Bernstein and Tanja Lange, editors, *AFRICACRYPT 10*, volume 6055 of *LNCS*, pages 279–296. Springer, Berlin, Heidelberg, May 2010.

[NRS11] Svetla Nikova, Vincent Rijmen, and Martin Schläffer. Secure hardware implementation of nonlinear functions in the presence of glitches. *Journal of Cryptology*, 24(2):292–321, April 2011.

[PM16] Peter Pessl and Stefan Mangard. Enhancing side-channel analysis of binary-field multiplication with bit reliability. In Kazue Sako, editor, *CT-RSA 2016*, volume 9610 of *LNCS*, pages 255–270. Springer, Cham, February / March 2016.

[PR13] Emmanuel Prouff and Matthieu Rivain. Masking against side-channel attacks: A formal security proof. In Thomas Johansson and Phong Q. Nguyen, editors, *EUROCRYPT 2013*, volume 7881 of *LNCS*, pages 142–159. Springer, Berlin, Heidelberg, May 2013.

[Reg05] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In Harold N. Gabow and Ronald Fagin, editors, *37th ACM STOC*, pages 84–93. ACM Press, May 2005.

[SAB+20] Peter Schwabe, Roberto Avanzi, Joppe Bos, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Vadim Lyubashevsky, John M. Schanck, Gregor Seiler, and Damien Stehlé. CRYSTALS-KYBER. Technical report, National Institute of Standards and Technology, 2020. available at https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization/round-3-submissions.

[SLP05] Werner Schindler, Kerstin Lemke, and Christof Paar. A stochastic model for differential side channel cryptanalysis. In Josyula R. Rao and Berk Sunar, editors, *CHES 2005*, volume 3659 of *LNCS*, pages 30–46. Springer, Berlin, Heidelberg, August / September 2005.

[WLL24] Zhedong Wang, Qiqi Lai, and Feng-Hao Liu. Ring/module learning with errors under linear leakage - hardness and applications. In Qiang Tang and Vanessa Teague, editors, *PKC 2024, Part II*, volume 14602 of *LNCS*, pages 275–304. Springer, Cham, April 2024.