Lattices In Cryptography | Lecture 11 | Instructor: Chris Peikert

Winter 2026 | Worst- to Average-Case Reductions | Scribe: Abhishek Banerjee

1 Worst-Case and Average-Case Hardness

In this lecture we will cover (most of) a *worst-case to average-case* reduction for lattice problems, as first pioneered by Ajtai [Ajt96], and significantly simplified and tightened by Micciancio and Regev [MR04] using the Gaussian and harmonic analysis tools developed in the previous lectures.

The primary significance of these reductions is as follows: they show that a certain short-vector problem A is hard to solve (even with small but noticeable probability) on a *random* lattice—drawn from an explicit, very simple probability distribution—as long as a related short-vector problem W is hard in the *worst case*, i.e., if there exist some hard instances (however rare and elusive they may be). This is shown by giving an efficient *reduction* that successfully solves problem W on an *arbitrary* instance, using a hypothetical oracle that is only guaranteed to solve *random* instances of problem A (with noticeable probability). So, if W is indeed hard in the worst case, then there cannot be any efficient implementation of the oracle, i.e., A is hard on the average.

Cryptography and average-case hardness. For cryptography, average-case hardness is essential: we need a cryptosystem's *randomly chosen* instances—keys, ciphertexts, etc.—to be hard to break. (It is not enough that there merely *exist* worst-case-hard instances, if we don't know what form they take or how to generate them!) Often in cryptography, one simply *assumes* that a problem is hard on the average, for some specified distribution of instances. For example, one may assume that factoring the product of two uniformly random n-bit primes is hard. (As far as we can tell, these seem like the hardest instances to factor, though we have no proof of this.) But this approach can be fraught: as we have seen in previous lectures, cryptosystems sometimes use instances with certain "structure" (e.g., small-exponent RSA, low-density knapsacks) that makes them much easier to solve than in the general case.

We note that some other cryptographic problems have worst-case to average-case reductions, of a limited kind. For example, for the *discrete logarithm* problem in a fixed finite cyclic group (e.g., $\mathbb{Z}_p^*$ for a prime p), it is easy to show that finding the discrete log of a *uniformly random* group element is hard, unless the discrete log problem is easy in the worst case (i.e., for all group elements). However, this reduction is limited to a specific fixed group, and tells us nothing about how hard it is to compute discrete logs in that group. (In many groups, it is easy!) By contrast, the reduction we will see in this lecture applies to *all* lattices, without any restriction.

2 The SIS Problem

We start by defining the average-case problem for which we will prove worst-case hardness. This problem has many cryptographic applications, as we will see in future lectures.

Definition 2.1 (Short Integer Solution Problem). For a positive integer modulus q, dimensions n, m and a norm bound $\beta > 0$, the $\mathsf{SIS}_{n,q,\beta,m}$ problem is defined as follows: given uniformly random $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m) \in \mathbb{Z}_q^{n\times m}$ where each $\mathbf{a}_i \in \mathbb{Z}_q^n$, find a nonzero "short" solution $\mathbf{z} \in \mathbb{Z}^m$ such that $\mathbf{A} \cdot \mathbf{z} = \sum_{i=1}^m \mathbf{a}_i \cdot z_i = \mathbf{0} \in \mathbb{Z}_q^n$ and $\|\mathbf{z}\| \leq \beta$.

In order to make the problem nontrivial, we must take $\beta < q$; otherwise, $\mathbf{z} = (q, 0, \ldots, 0) \in \mathbb{Z}^m$ is always a solution. In order to guarantee that a solution exists, we typically take $m > n \log q$ and $\beta \geq \sqrt{m}$. Then consider the function that maps any $\mathbf{x} \in \{0,1\}^m$ to $\mathbf{A}\mathbf{x} \in \mathbb{Z}_q^n$. Because this maps a domain of size $2^m > q^n$ to a range of size q^n, by the pigeonhole principle it has a *collision*, i.e., some $\mathbf{x} \neq \mathbf{x}'$ such that $\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{x}'$. Letting $\mathbf{z} = \mathbf{x} - \mathbf{x}' \in \{0, \pm 1\}^m$, we have that $\mathbf{A}\mathbf{z} = \mathbf{0}$, $\mathbf{z} \neq \mathbf{0}$, and $\|\mathbf{z}\| \leq \sqrt{m} \leq \beta$, as needed.

Views of SIS. One can see the SIS problem from some different perspectives.

1. We can view SIS as a kind of approximate-SVP problem on the following random integer lattice (where $\mathbf{A}$ is uniformly random):

   $$\mathcal{L}^{\perp}(\mathbf{A}) = \{\mathbf{z} \in \mathbb{Z}^m \colon \mathbf{A} \cdot \mathbf{z} = \mathbf{0} \pmod{q}\}.$$

   Note that since $\mathcal{L}^{\perp} \subseteq \mathbb{Z}^m$, it is discrete, and it can also easily be checked that it is a *subgroup* of $\mathbb{Z}^m$, hence it is a lattice. Moreover, it is "q-ary" in the sense that $q\mathbb{Z}^m \subseteq \mathcal{L}^{\perp}(\mathbf{A})$, so a vector $\mathbf{v} \in \mathbb{Z}^m$ is in $\mathcal{L}^{\perp}(\mathbf{A})$ if and only if $\mathbf{v} \bmod q$ is in $\mathcal{L}^{\perp}(\mathbf{A})$.

   The SIS problem asks for a "short" nonzero $\mathbf{z} \in \mathcal{L}^{\perp}(\mathbf{A})$, which is similar to what the SVP problem requires. Indeed, it is possible to show that for $m \geq (1+\delta)n \log q$, we have $\lambda_1(\mathcal{L}^{\perp}(\mathbf{A})) = \Omega(\sqrt{n \log q})$ with high probability. So, solving SIS effectively solves SVP with approximation factor $O(\beta/\sqrt{n \log q})$ on a random lattice from this family.

2. We can also view SIS as a sort of a subset-sum or weak-partition problem.

   In the subset-sum problem, we are given weights $a_1, a_2, \ldots, a_m \in \mathbb{Z}$, and a subset-sum $s = \sum_{i=1}^m a_i x_i$ for $x_i \in \{0, 1\}$, and we need to find the x_i. The natural generalization of this problem, replacing $\mathbb{Z}$ with $\mathbb{Z}_q^n$, is closely related to the SIS problem. In this form, the *density* of the problem is $m/\log(q^n) = m/n \log(q) \geq 1$. We have previously seen that subset sum is easy if the density is at most $1/n$, and the SIS problem lies outside this realm.

   In the weak-partition problem, we are given weights $A = (a_1, \ldots, a_m)$ from some additive group, and are required to find a nonzero $\mathbf{x} \in \{-1, 0, 1\}^m$ such that $\sum_{i=1}^m a_i x_i = 0$. It is evident that this, specialized to the group $\mathbb{Z}_q^n$, is just a restricted case of the SIS problem, where the solution must be ternary.

It should be noted that there is nothing sacrosanct about choosing the "weights" $\mathbf{a}_i$ from the group $\mathbb{Z}_q^n$. This choice is particularly convenient for the reduction and for cryptographic applications. However, we could instead have used, say, $\mathbb{Z}_{q_1} \times \mathbb{Z}_{q_2} \times \cdots \times \mathbb{Z}_{q_n}$ for large enough moduli q_i. Note that if these moduli q_i are pairwise coprime, then $\mathbb{Z}_{q_1} \times \mathbb{Z}_{q_2} \times \cdots \times \mathbb{Z}_{q_n} \cong \mathbb{Z}_{q_1 q_2 \cdots q_n}$ by the Chinese Remainder Theorem.

3 Reduction from SIVP

We now define the worst-case lattice problem we will reduce to SIS. It is the analog of the shortest vector problem, generalized to the concept of the nth successive minimum λ_n. (Recall from the previous lecture that λ_n is the smallest real r such that the lattice contains n linearly independent vectors of length at most r.)

Definition 3.1. For an approximation factor $\gamma = \gamma(n) \geq 1$, the γ-approximate *Shortest Independent Vectors Problem* SIVP_γ is defined as follows: given a basis $\mathbf{B}$ of an n-dimensional lattice $\mathcal{L} = \mathcal{L}(\mathbf{B})$, output n linearly independent lattice vectors $\mathbf{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subset \mathcal{L}$ such that $\|\mathbf{V}\| := \max_i \|\mathbf{v}_i\| \leq \gamma \cdot \lambda_n(\mathcal{L})$.

(Notice that each vector $\mathbf{v}_i$ does *not* need to approximate the ith successive minimum, only the nth one.)

The following theorem gives a worst-case to average-case reduction from SIVP to SIS. We have not presented the most optimized parameters, but the approximation factor for SIVP can be made as small as $\gamma = \beta \cdot \tilde{O}(\sqrt{n})$ (where $\tilde{O}$ hides polylogarithmic factors), i.e., there is no direct dependence on m. In addition, a slightly better bound on q can be obtained by using *discrete* Gaussian sampling, as detailed in [GPV08].

Theorem 3.2. *For any $m = \operatorname{poly}(n)$ and any $q \geq 4\beta n\sqrt{m} = \beta \cdot \operatorname{poly}(n)$, there is a randomized polynomial-time reduction from* SIVP_γ *on n-dimensional lattices to* $\mathsf{SIS}_{n,q,\beta,m}$*, where $\gamma = \tilde{O}(\beta\sqrt{nm})$.*

In what follows we present the key ideas behind the proof of this theorem, but gloss over some ancillary technical details. The reduction has access to a hypothetical oracle $\mathcal{O}$ for SIS, and works as follows. Given a basis $\mathbf{B}$ of an n-dimensional lattice $\mathcal{L}$, it repeatedly performs the following *core step* using $\mathcal{O}$ to produce a new, shorter basis $\mathbf{B}'$ of $\mathcal{L}$. Then it does the same with $\mathbf{B}'$ to get an even shorter basis $\mathbf{B}''$, and so on. We show below that, if the current basis is not already an SIVP solution, then the next basis will be shorter than the current one, by a factor of two. So, the reduction just iteratively generates such shorter and shorter bases until the process stops working, at which point it outputs the current basis (which must be an SIVP solution). Because we can start from a 2^n-approximate SIVP solution thanks to LLL, the total number of iterations will be $\operatorname{poly}(n)$.

The core step. In summary, the core step works as follows. It generates m independent and "somewhat short" (relative to the current basis $\mathbf{B}$) lattice vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathcal{L}$, by sampling from a Gaussian of suitable width and "rounding off" to the lattice using $\mathbf{B}$. It then invokes the SIS oracle on instance $(\mathbf{a}_1, \ldots, \mathbf{a}_m)$, where each $\mathbf{a}_i \in \mathbb{Z}_q^n$ is the integer coefficient vector of $\mathbf{v}_i \in \mathcal{L}$, reduced modulo q. In other words, each $\mathbf{a}_i$ corresponds to the coset of $\mathcal{L}/q\mathcal{L}$ to which $\mathbf{v}_i$ belongs. Using the smoothing parameter, we can show that these $\mathbf{a}_i$ are very close to uniformly random, as required. So, the oracle is obliged to output a valid (nonzero) SIS solution $\mathbf{z} \in \mathbb{Z}^m$ with some noticeable probability. If it does, we have $\sum_i \mathbf{a}_i z_i = \mathbf{0} \in \mathbb{Z}_q^n$. This implies that $\sum_i \mathbf{v}_i z_i \in q\mathcal{L}$ (not just $\mathcal{L}$), so we can divide by q and still have a lattice vector. Because $q \gg \beta \geq \|\mathbf{z}\|$, the division by q more than compensates for the $\mathbf{z}$-weighted sum of the vectors $\mathbf{v}_i$, which makes the resulting lattice vector significantly shorter than those in the basis $\mathbf{B}$. We repeat the core step many times until we get n linearly independent lattice vectors, which can be transformed into a basis $\mathbf{B}'$.[1]

Here is the core step in detail:

1. Sample independent Gaussian vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in \mathbb{R}^n$ with parameter $s = q \cdot \frac{\|\mathbf{B}\|}{4\beta\sqrt{nm}}$.

2. Let $\mathbf{c}_i = \lfloor \mathbf{B}^{-1}\mathbf{x}_i \rceil \in \mathbb{Z}^n$ be the coefficient vector of $\mathbf{x}_i$ relative to $\mathbf{B}$, rounded off to the nearest integer in each coordinate. Let $\mathbf{v}_i = \mathbf{B}\mathbf{c}_i \in \mathcal{L}$ and $\mathbf{a}_i = \mathbf{c}_i \bmod q\mathbb{Z}^n \in \mathbb{Z}_q^n$.

   (Note that $\mathbf{a}_i$ corresponds to the coset of $\mathbf{B}\mathbf{a}_i \in \mathcal{L}/q\mathcal{L}$ to which $\mathbf{v}_i$ belongs.)

3. Call the SIS oracle $\mathcal{O}(\mathbf{a}_1, \ldots, \mathbf{a}_m)$ to obtain a possible solution $\mathbf{z} \in \mathbb{Z}^m$.

4. Output $\mathbf{v} = q^{-1} \cdot \sum_{i=1}^m \mathbf{v}_i z_i$.

In what follows, we present a few lemmas establishing the key properties of this core step. The first two lemmas show that whenever the SIS oracle outputs a solution, the core step outputs a lattice vector that is significantly shorter than the current basis.

Lemma 3.3. *If $\mathbf{z}$ is a solution to the generated* SIS *instance, then $\mathbf{v} \in \mathcal{L}$.*

[1] We are skipping over some important technical details here. Most significantly, we need it to be the case that repeating the core step is likely to produce *linearly independent* lattice vectors, even if the oracle behaves "deviously." This can be shown using an information-theoretic argument, as outlined in Lemma 3.6 and its proof. Second, the transformation from linearly independent lattice vectors to a *basis* requires some care, so that it does not worsen the "quality" of the vectors by too much.

Proof. We have that $\mathbf{v}_i = \mathbf{B}\mathbf{c}_i \in \mathcal{L}$ for some $\mathbf{c}_i \in \mathbb{Z}^n$, and $\mathbf{a}_i = \mathbf{c}_i \pmod{q\mathbb{Z}^n}$. So,

$$\mathbf{v} = q^{-1}\sum_{i=1}^{m} \mathbf{v}_i z_i = q^{-1}\mathbf{B}\sum_{i=1}^{m} \mathbf{c}_i z_i.$$

Now, since $\mathbf{z} \in \mathbb{Z}^m$ is an SIS solution, we have that $\sum_{i=1}^{m} \mathbf{c}_i z_i \in q\mathbb{Z}^n$ and thus $\mathbf{v} \in \mathbf{B} \cdot \mathbb{Z}^n = \mathcal{L}$. □

Lemma 3.4. *If* $\|\mathbf{z}\| \leq \beta$, *then* $\|\mathbf{v}\| \leq \|\mathbf{B}\|/2$ *with high probability.*

Proof. Each $\|\mathbf{x}_i\| \leq s\sqrt{n}$ with high probability, by Gaussian concentration. Also, each $\|\mathbf{v}_i - \mathbf{x}_i\| \leq \sum_{j=1}^{n}\|\mathbf{b}_j\| \leq \|\mathbf{B}\|n$. So, recalling that $q \geq 4\beta n\sqrt{m}$, and by the triangle and Cauchy-Schwartz inequalities,

$$\begin{aligned} q\|\mathbf{v}\| &= \left\| \sum_{i=1}^{m} \mathbf{v}_i z_i \right\| \\ &\leq \left\| \sum_{i=1}^{m} \mathbf{x}_i z_i \right\| + \left\| \sum_{i=1}^{m} (\mathbf{v}_i - \mathbf{x}_i) z_i \right\| \\ &\leq (s\sqrt{n})(\beta\sqrt{m}) + (\|\mathbf{B}\|n)(\beta\sqrt{m}) \\ &\leq q \cdot \frac{\|\mathbf{B}\|}{4} + q \cdot \frac{\|\mathbf{B}\|}{4} \\ &= q\|\mathbf{B}\|/2. \end{aligned}$$

□

The next lemma gives us the exit condition for the procedure surrounding the core step: if the core step ever stops working (i.e., stops producing sufficiently short lattice vectors after enough attempts), then the current basis $\mathbf{B}$ must fail to satisfy the hypothesis from the lemma, and hence it is a $\tilde{O}(\beta\sqrt{nm})$-approximate SIVP solution.

Lemma 3.5. *If* $\|\mathbf{B}\| \geq 4\beta\sqrt{nm} \cdot \log n \cdot \lambda_n(\mathcal{L})$, *then* $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_m)$ *is* $n^{-\omega(1)}$*-close (in statistical distance) to a uniformly random* $\mathsf{SIS}_{n,q,\beta,m}$ *instance, and hence the oracle* $\mathcal{O}$ *must output a solution with noticeable probability.*

Proof. We have $s \geq q \cdot \log n \cdot \lambda_n(\mathcal{L}) \geq q \cdot \eta_\epsilon(\mathcal{L}) = \eta_\epsilon(q\mathcal{L})$ for $\epsilon = n^{-\log n}$, by the theorem from the previous lecture. So $\rho_s(\mathbf{t} + q\mathcal{L}) \approx \rho_s(q\mathcal{L})$ for every coset $\mathbf{t} + q\mathcal{L}$, where the approximation hides a multiplicative factor of at most $\frac{1+\varepsilon}{1-\varepsilon}$. Therefore, the Gaussian-distributed vectors $\mathbf{x}_i$ are $n^{-\omega(1)}$-far from uniform modulo $q\mathcal{L}$, which implies that their real coefficient vectors $\mathbf{B}^{-1}\mathbf{x}_i \in \mathbb{R}^n$ are similarly close to uniform modulo $q\mathbb{Z}^n$. Integrating over all (real) cosets that map to any fixed $\mathbf{a} \in \mathbb{Z}_q^n$ yields the claim. □

Our final lemma addresses the concern that a potentially "devious" oracle $\mathcal{O}$ could, by carefully crafting its SIS solutions based on how the instances are generated, somehow prevent the core step from generating a full-rank linearly independent set of shorter lattice vectors.

Lemma 3.6. *Upon success, the* $\mathbf{v}$ *output by the basic step is not concentrated on any fixed hyperplane. More formally, conditioned on* $\mathbf{z}$ *being an* SIS *solution, for any* $(n-1)$*-dimensional hyperplane* $\mathcal{H}$, $\Pr_{\mathbf{x}_i}[\mathbf{v} \in \mathcal{H}] \leq \frac{3}{4}$.

Proof sketch. We use an information-theoretic argument. Conditioned solely on the SIS instance $\mathbf{a}_i$, or even on the full cosets $\mathbf{x}_i + q\mathcal{L} \in \mathbb{R}^n/q\mathcal{L}$ (which determine the $\mathbf{a}_i$), we claim that the $\mathbf{x}_i$ are still "well-spread" enough that no fixed SIS solution $\mathbf{z}$ is very likely to place $\mathbf{v}$ in $\mathcal{H}$. More specifically, since each $\mathbf{x}_i$ was sampled from a Gaussian of parameter s, the conditional distribution of $\mathbf{x}_i$, given that it is in some coset $\mathbf{t}_i + q\mathcal{L}$, is a *discrete Gaussian*:

$$\mathcal{D}_{\mathbf{t}_i+q\mathcal{L},s}(\mathbf{x}_i) := \frac{\rho_s(\mathbf{x}_i)}{\rho_s(\mathbf{t}_i + q\mathcal{L})}.$$

Note that $q\mathbf{v} = \sum_{i=1}^m \mathbf{v}_i z_i = \sum_{i=1}^m \mathbf{x}_i z_i + \sum_{i=1}^m (\mathbf{v}_i - \mathbf{x}_i) z_i$, and the second summation $\mathbf{s}$ is fixed given $\mathbf{z}$ and the cosets $\mathbf{x}_i + q\mathcal{L}$, because each "round-off" term $\mathbf{v}_i - \mathbf{x}_i$ is fully determined by the corresponding coset (and not $\mathbf{x}_i$ itself). So, $\mathbf{v} \in \mathcal{H}$ if and only if $\sum_{i=1}^m \mathbf{x}_i z_i \in \mathcal{H} - \mathbf{s}$. Since $\mathbf{z} \neq 0$, we can assume without loss of generality that $z_1 \neq 0$. Then, fixing all the $\mathbf{x}_i$ except $\mathbf{x}_1$, we have that $\mathbf{v} \in \mathcal{H}$ only if $\mathbf{x}_1 \in \mathcal{H} - \mathbf{s}'$ for a certain fixed $\mathbf{s}' \in \mathbb{R}^n$. It is shown in [MR04] that a discrete Gaussian over any fixed lattice coset, whose width exceeds the lattice's smoothing parameter, is not concentrated on any proper affine subspace, as needed. □

References

[Ajt96] M. Ajtai. Generating hard instances of lattice problems. *Quaderni di Matematica*, 13:1–32, 2004. Preliminary version in STOC 1996. Page 1.

[GPV08] C. Gentry, C. Peikert, and V. Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *STOC*, pages 197–206. 2008. Page 2.

[MR04] D. Micciancio and O. Regev. Worst-case to average-case reductions based on Gaussian measures. *SIAM J. Comput.*, 37(1):267–302, 2007. Preliminary version in FOCS 2004. Pages 1 and 5.