Notes on eprint 2024/555

Sam Jaques

April 21, 2024

1 Gap SVP

Chen focuses on a slightly unique variant of gap SVP. Normally we consider the ith successive minima $\lambda_i(L)$ as a bound on the maximum norm of a set of i linearly independent vectors in L. Here we consider bounds on the infinity norm of such linearly independent sets: $\lambda_i^\infty(L)$. Our problem will be to solve SVP for a lattice where we are promised that $\lambda_2^\infty(L) > 2\lambda_1(L)$.

The reason we do this is that we will end up with superpositions over all lattice vectors in some hypercube; this condition ensures that the difference between any two such vectors is a multiple of the shortest vector.

Can this solve LWE? In certain cases, yes. Recall that for an LWE problem $As + e \equiv b \mod q$, where A is a random $m \times n$ matrix, we can form a lattice L_{LWE} from the basis

$$\begin{pmatrix} qI_m & A & -b \\ 0 & I_n & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1}$$

and you can convince yourself that $(-e, s, 1)$ is in this lattice. A bit more convincing can tell you that if the components of e and s are drawn independently from some distribution with mean 0 and variance σ^2, the norm of $(-e, s)$ is going to concentrate around $\sqrt{n+m}\sigma$ (notice that I ignored the last component of 1, and I'm going to continue to do so).

However, the volume of this lattice is q^m. The Gaussian heuristic tells us to expect that the shortest lattice vector is $\sqrt{\frac{n+m}{2\pi e}} q^{\frac{m}{n+m}}$ (ignoring the +1 again). So, we can expect the *second*-shortest vector to follow the Gaussian heuristic, while the shortest vector is, by design, $(-e, s, 1)$, which is unusually short.

This is partly the reason to treat LWE as a GapSVP problem generally, but in our case we want to use the fact that for any vector x, we know that $\|x\|_\infty \sqrt{n} \geq \|x\|_2$. This means we can (heuristically) expect

$$\lambda_2^\infty(L_{LWE}) \geq \frac{1}{\sqrt{n+m}} \sqrt{\frac{n+m}{2\pi e}} q^{\frac{m}{n+m}} = \frac{1}{\sqrt{2\pi e}} q^{\frac{m}{n+m}} \tag{2}$$

And so we satisfy the required property for Chen's algorithm to work if

$$\frac{1}{\sqrt{2\pi e}} q^{\frac{m}{n+m}} > 2\sqrt{n+m}\sigma \tag{3}$$

As far as I can tell, this is an essential condition for Chen's algorithm (and probably we need a bit of a gap between these terms for things to work nicely). There might be other, stronger conditions, but even starting with this one tells us that for Kyber (taking, say, $n = m = 512$ and $q = 3329$) we would need $\sigma < 0.3$; this is about one third the variance of Kyber's secret and error. Kyber is safe (barely), though I'll emphasize that this is just a first-pass limitation on Chen's techniques. Looking carefully at the later parts, I think we need quite a large gap between λ_2^∞ and λ_1, and this seems to be a limitation to the entire approach, not just Chen's specific algorithm.

2 Adding Hypercubes

It's straightforward to create the following state:

$$\sum_{y \in \mathbb{Z}^n : \|y\|_\infty < R} |y\rangle \tag{4}$$

(One simple way is to create a uniform superposition from $-R$ to R iin each coordinate). The value R is chosen so that $\lambda_2^\infty > 2R > 2\lambda_1$.

For now we will assume that we can also make a superposition

$$\sum_{v \in L_{LWE}} |v\rangle \tag{5}$$

(Of course this is not normalized and it is a sum over an infinite set; in fact we just need to make a sum over a q-ary part and argue that if this is sufficiently larger than R then the final state looks the same as if we started with the full superposition). But, once we have this set, we can combine the two states:

$$\sum_{v \in L_{LWE}} |v\rangle \sum_{y \in \mathbb{Z}^n : \|y\|_\infty < R} |y\rangle \tag{6}$$

and add y to v

$$\sum_{v \in L_{LWE}} |v + y\rangle \sum_{y \in \mathbb{Z}^n : \|y\|_\infty < R} |y\rangle \tag{7}$$

and measure the result (call it y')

$$\sum_{y \in \mathbb{Z}^n : \|y\|_\infty < R, y' - y \in L_{LWE}} |y\rangle \tag{8}$$

That is, we get a superposition of all small y that are separated from y' by a lattice vector. We can pick some $v_0 \in L_{LWE}$ such that v_0 is as close as possible to y'; we know that $\|v_0 - y'\|_\infty < R$. For the rest of the algorithm to work, we need two things:

1. That there are two vectors y_1 and y_2 in superposition such that $y_2 - y_1$ is the shortest vector in L_{LWE} (call it x_0)

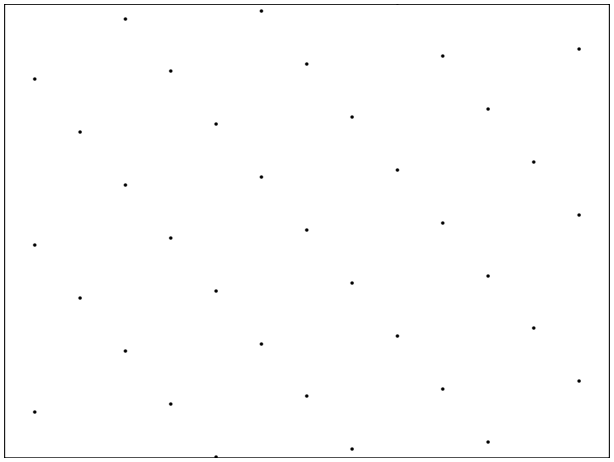
2. For any two vectors y_1 and y_2 in superposition, $y_2 - y_1$ must be an integer multiple of the shortest vector.

If the first one is true, the second one is easy to show. Specifically, we know that for all y remaining in the superposition, there is some lattice vector v such that $y + v = y'$. Thus, we have $y_1 + v_1 = y' = y_2 + v_2$ if both y_1, y_2 are in the superposition; then we have that $y_2 - y_1 = v_2 - v_1 \in L$. If $v_2 - v_1$ is not a scalar multiple of x_0 (the shortest vector), then it is linearly independent of x_0, and we also know that $\|v_2 - v_1\|_\infty = \|y_2 - y_1\|_\infty < 2R$, by the triangle inequality. Thus, since $\|x_0\|_\infty < 2R$ as well, we would have two linearly independent vectors of infinity norm at most $2R$, but that contradicts $\lambda_2^\infty > 2R$.

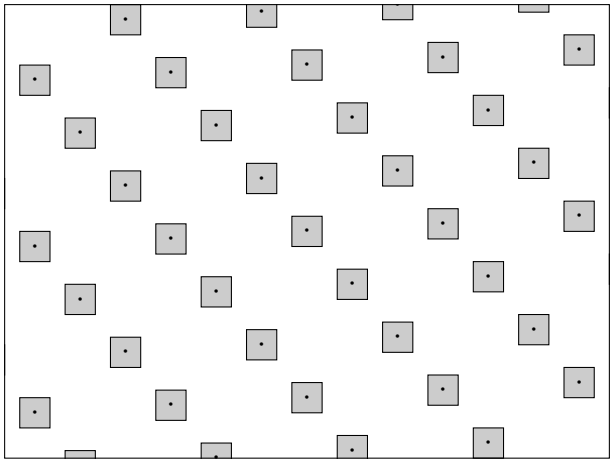
But showing the first is a bit more difficult.

2.0.1 Superpositions of the Shortest Vector

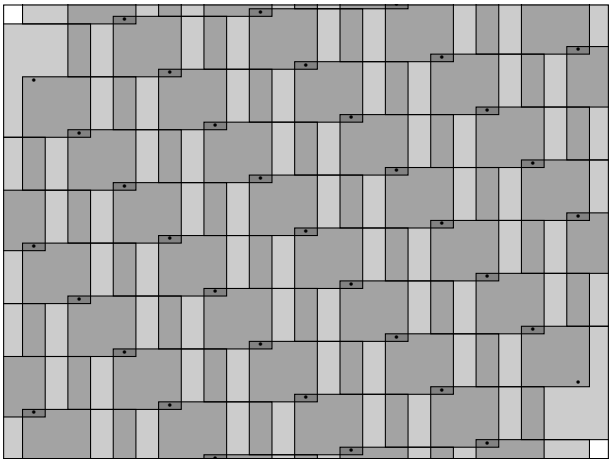
First, this is what the superposition over the lattice looks like:



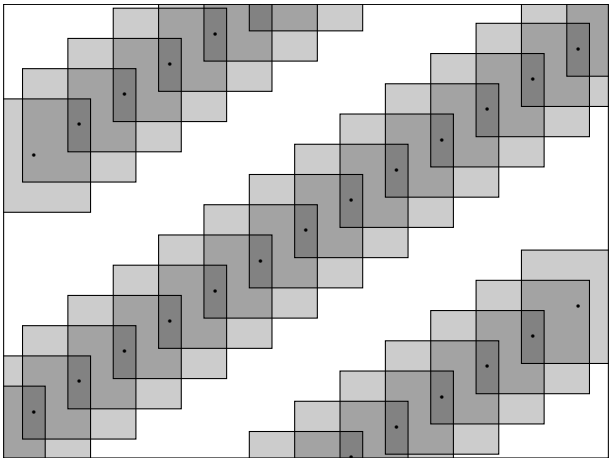
Then this is what it looks like with infinity-norm balls around each point:



Measuring y' is like selecting a random point from a random one of these "balls" (i.e., squares). Notice that in the drawing above, the squares are too small: we will be left with only one lattice point after measuring.



This next one has squares big enough to overlap, but the lattice does not have a big enough gap. Squares from different "rows" are overlapping. Thus, we need a bigger gap in the lattice:



This is more like what Chen needs. If we measure y' in the darkest regions, that overlaps 3 boxes, so we're left with 3 points in superposition; in the lightest regions, we're only left with 1.

In Chen's algorithm, we don't get uniform probabilities over the boxes, but it's a good start for now.

2.1 A first attempt

If the width of these boxes is big enough, what we'll end up with is some state like

$$\sum_{k=-B}^{B} |y_0 + kx_0\rangle \tag{9}$$

where y_0 is the shortest vector satisfying $y' - y_0 \in L$.

This is already quite an interesting state! It seems to capture some information about the shortest vector x_0. Indeed, if we immediately took a QFT (over $2R$, since that's the bound on the size of integers in the register currently) we would get

$$\sum_{z \in \mathbb{Z}_{2R}} |z\rangle \sum_{k=-B}^{B} e^{2\pi i \frac{\langle y_0 + kx_0, z\rangle}{2R}} \tag{10}$$

and we can move the y_0 part out of the sum and we get

$$\sum_{z\in\mathbb{Z}_{2R}} e^{2\pi i\frac{\langle y_0,z\rangle}{2R}}\,|z\rangle \sum_{k=-B}^{B} e^{2\pi ik\frac{\langle x_0,z\rangle}{2R}} \tag{11}$$

If $B = R$, then we would be summing roots of unity unless $\langle x_0, z\rangle \equiv 0 \mod 2R$, so measuring z would give us a vector orthogonal to x_0. With enough of these, we would simply find x_0 with linear algebra.

Unfortunately, we can't have $B = R$: B was defined as the maximum value where $2Bx_0$ fits in a box of width $2R$. So $\frac{B}{R} \leq \|x_0\|_\infty$.

Not to despair! If $\frac{\langle x_0,z\rangle}{2R} \approx \frac{1}{2B}\mathbb{Z}$, i.e., $e^{2\pi i\frac{\langle x_0,z\rangle}{2R}}$ is approximately a $2B$th root of unity, then the phases will similarly cancel out. More generally, if $\langle x_0, z\rangle \mod 2R$ is too big, then adding up B powers of it will be approximately 0. Thus, we have an equation something like

$$\langle x_0, z\rangle \approx 0 \mod 2R \tag{12}$$

Is this useful? Almost: we would rewrite it as

$$\langle x_0, z\rangle + e \equiv 0 \mod 2R \tag{13}$$

for small e, and now it looks like what it is: an LWE sample. We could repeat this process and get uniformly random z satisfying these properties, but ultimately we're just reducing LWE back to LWE.

In fact, this looks a bit like Regev's original reduction from SVP to LWE. It's not quite the same; I'm not sure if there are obstacles to completing the LWE reduction this way (we could have used any GapSVP problem to begin with, and we've reduced it to something which is almost LWE).

Nor is too different from Eldar and Hallgren's 2022 paper () solving BDD with a quantum algorithm. What we're really doing is solving LWE as a BDD problem, and implicitly using Kannan's embedding to map the BDD problem to an SVP problem. Eldar and Hallgren attack BDD directly by noting that once we have a superposition of boxes around lattice points, that state is an eigenstate of addition by lattice vectors. Since a BDD challenge is *close* to a lattice vector, the superposition is *close* to an eigenstate of addition by the BDD challenge. However, the "closeness" is not good enough: the only BDD challenges they efficiently solved turned out to be polynomial solvable classically.

3 Convolutions

Rather than have uniform superpositions, we can add some amplitude to each state. That is, we can consider a more general case where we have the state

$$\sum_x g(x)\,|x\rangle \sum_y f(y)\,|y\rangle \tag{14}$$

where g and f are arbitrary, and I'm deliberately being vague about the range of the sums of x and y since they could be almost anything at this point.

We can then do the same thing of adding x to y and measuring. This gives a result y', and the remaining state is

$$\sum_x g(x)f(y'-x)\,|x\rangle \tag{15}$$

This looks sort of like a convolution. If we apply a QFT mod P we get:

$$\sum_z |z\rangle \sum_x e^{i\frac{xz}{P}} g(x) f(y' - x) \tag{16}$$

Again, this looks almost like a convolution, and almost like a Fourier transform. You might be tempted to think that since the Fourier transform of a convolution is the product of the Fourier transforms, that we could use that identity here, but it doesn't quite work. I am not sure of how to approach this generally.

3.1 Complex Gaussians

For notational convenience, I will define a Gaussian-like function

$$G(x; \mu, v, \theta) = \exp\left(\pi v (x - \mu)^2\right) \exp\left(2\pi i \theta x\right) \tag{17}$$

More or less I'm just substituting $v = \frac{1}{\pi\sigma^2}$ in the normal notation. My goal at this point is to create a nice set of rules for multiplying, composing, shifting, convoluting, etc., this kind of function. I'm imagining that v can be complex but μ and θ are real. This avoids a situation where a complex mean would create a real *positive* exponential. Unlike Chen I'm not going to split v into its real and imaginary components, which makes the equations a bit cleaner.

Parameters Some quick identities:

$$G(x + c; \mu, v, \theta) = G(x; \mu - c, v, \theta) \tag{18}$$

and

$$G(cx; \mu, v, \theta) = G(x; \mu/c, vc^2, c\theta) \tag{19}$$

Multiplication We can compute (it's tedious):

$$G(x; \mu_1, v_1, \theta_1) G(x; \mu_2, v_2, \theta_2) = \exp\left(x^2(v_1 + v_2) - 2x(\mu_1 v_1 + \mu_2 v_2) + \mu_1^2 v_1 + \mu_2^2 v_2\right) \exp\left(2\pi i x(\theta_1 + \theta_2)\right) \tag{20}$$

$$\vdots \tag{21}$$

$$= G\left(x; \frac{\mu_1 v_1 + \mu_2 v_2}{v_1 + v_2}, v_1 + v_2, \theta_1 + \theta_2\right) \exp\left(\frac{(\mu_1 - \mu_2)^2 v_1 v_2}{v_1 + v_2}\right) \tag{22}$$

Shifts If we multiply G by e^{cx} for real c, what happens? We need to complete the square in the exponent; we end up with

$$G(x; \mu, v, \theta) e^{\pi c x} = \exp\left(\pi(vx^2 - v2x\mu + v\mu^2 + cx)\right) e^{i\pi\theta x} \tag{23}$$

$$\vdots \tag{24}$$

$$= G\left(x; \mu - \frac{c}{2v}, v, \theta\right) \exp\left(\pi c(\mu - c/4v)\right) \tag{25}$$

Shifting by imaginary ϕ is much easier:

$$G(x; \mu, v, \theta) e^{2\pi i \phi x} = G(x; \mu, v, \theta + \phi) \tag{26}$$

Fourier Transforms Chen gives this one:

$$\mathcal{F}\left(G(x;0,v,0)\right)(z) = \frac{1}{\sqrt{v}} G(z;0,\tfrac{1}{v},0) \tag{27}$$

which we should extend to non-centered distributions with phase. This is easy using the fact that, generally

$$\mathcal{F}(f(x+a))(z) = e^{-2\pi i a z}\mathcal{F}(f(x))(z) \tag{28}$$

so we have that

$$\mathcal{F}\left(G(x;\mu,v,0)\right)(z) = \frac{e^{2\pi i \mu z}}{\sqrt{v}} G\left(z;0,\frac{1}{v},0\right) \tag{29}$$

$$= \frac{1}{\sqrt{v}} G(z;0,\frac{1}{v},\mu) \tag{30}$$

And this works in reverse: if we already had a phase θ, then the phase multiplier of x in the Fourier transform integral is $\theta + z$, so we have

$$\mathcal{F}\left(G(x;\mu,v,\theta)\right)(z) = \frac{1}{\sqrt{v}} G\left(z+\theta,0,\frac{1}{v},\mu\right) \tag{31}$$

$$= \frac{1}{\sqrt{v}} G\left(z,-\theta,\frac{1}{v},\mu\right) \tag{32}$$

In short: Fourier transforms flip the variance and exchange phase for mean.

(fun fact: this shows that G is an eigenfunction of the Fourier transform with eigenvalue i, because repeating the Fourier transform 4 times will return to the same function.)

3.2 Second Attempt: Real Gaussians

It is relatively straightforward () to make a state like

$$\sum_{y\in\mathbb{Z}^n} G(y;0,v,0)\,|y\rangle \tag{33}$$

and so, doing the convolution-like thing with a sum of lattice vectors, we can obtain

$$\sum_{x\in L} G(y'-x;0,v,0)\,|y'-x\rangle \tag{34}$$

which we can then rewrite as scalar multiples of the shortest vector as

$$\sum_{k\in\mathbb{Z}} G(y_0+kx_0;0,v,0)\,|y_0+kx_0\rangle \tag{35}$$

A quick digression: We would actually catch many more lattice vectors, but if the Gaussian is narrow enough, we can ignore them because the amplitude on these states is too small (this is analogous to choosing the hypercube width between λ_2^∞ and λ_1). However, we also want our Gaussian to be wide enough to catch several multiples of x_0, otherwise we don't really have any useful information about x_0.

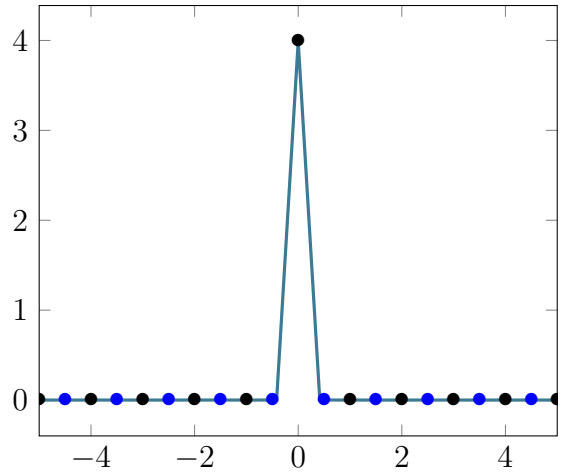


Figure 1: Illustration of how sums of Gaussians over integers catch when the mean is integer. In black, the mean is integer, and the sum is 3.89; in blue the mean is 0.5, and the sum is less than 10^{-4}

Anyway, from this state we do a QFT and get:

$$\sum_{z\in\mathbb{Z}_p^n} |z\rangle \sum_{k\in\mathbb{Z}} e^{2\pi i\langle z, y_0+kx_0\rangle/P} G(y_0+kx_0;0,v,0) \tag{36}$$

$$= \sum_{z\in\mathbb{Z}_p^n} |z\rangle\, e^{2\pi i\langle z,y_0\rangle/P}\, |z\rangle \sum_{k\in\mathbb{Z}} G(y_0+kx_0;0,v,\langle x_0, z/P\rangle) \tag{37}$$

It's a bit annoying to deal with the mix of vectors and scalars here; let's expand this out to remove the vector terms:

$$G(y_0+kx_0;0,v,\langle x_0,z/P\rangle) = \exp\left(v(y_0+kx_0)^2\right) e^{2\pi iki\langle x_0,z/P\rangle} \tag{38}$$

$$= \exp\left(-\pi v(\|y_0\|^2 + 2k\langle y_0,x_0\rangle + k^2\|x_0\|^2)\right) e^{2\pi iki\langle x_0,z/P\rangle} \tag{39}$$

$$= \exp\left(-\pi v\|x_0\|^2\left(k+\frac{\langle y_0,x_0\rangle}{\|x_0\|}\right)^2 - \pi v\frac{\langle y_0,x_0\rangle^2}{\|x_0\|^2} - \pi v\|y_0\|^2\right) e^{2\pi iki\langle x_0,z/P\rangle} \tag{40}$$

$$= G\left(k; -\frac{\langle y_0,x_0\rangle}{\|x_0\|^2}, v\|x_0\|^2, \frac{\langle x_0,z\rangle}{P}\right) \exp\left(-\pi v\left(\frac{\langle y_0,x_0\rangle^2}{\|x_0\|^2} - \|y_0\|^2\right)\right) \tag{41}$$

What is our goal here? We want to have some destructive/constructive interference that depends on z, so that we are more likely to measure z with certain properties. This means that the $\exp(\cdot)$ term on the right is actually meaningless, because it does not depend on z nor k. Really, it will get normalized away.

How useful is this? Well, if the (inverse) variance $v\|x_0\|^2$ is large, then the Gaussian will concentrate sharply around 0. Then the sum will be close to 1 if $-\frac{\langle y_0,x_0\rangle}{\|x_0\|^2}$ is near an integer, and will be near 0 otherwise. See Figure 1. Unfortunately, this seems to be totally independent of z.

I think y_0 is essentially a random n-dimensional vector, subject to length restrictions. Thus it is likely to be nearly orthogonal to x_0, so $\frac{\langle y_0,x_0\rangle}{\|x_0\|^2} \approx 0$ will be close to an integer.

If $v\|x_0\|^2$ is small, then the Gaussian will be spread out. (see Figure 2) If the phase multiplier is large, then this will likely cancel everything out; if the phase multiplier is small, then it will not cancel. The

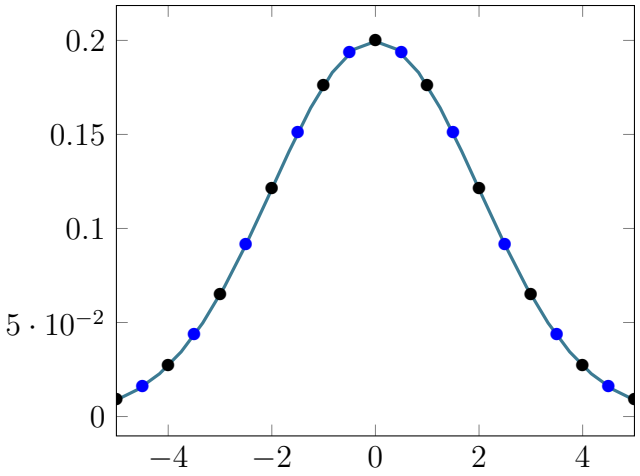


Figure 2: When variance is wide, the sum over integer values is basically the same, regardless of where the mean is; the difference between the black and blue sums is less than 0.01

phase multiplier is $\langle x_0, z\rangle/P$, so this should carry information about x_0: we are more likely to measure z such that $\langle x_0, z\rangle \approx 0 \mod P$ (see Figure 3).

In fact, we can apply the Poisson summation formula to make this a bit more precise. Specifically, the "lattice" is just $\mathbb{Z}$, whose dual is also $\mathbb{Z}$ and whose volume is 1. Thus we get

$$\sum_{k\in\mathbb{Z}} G(k;\mu,v,\theta) = \sum_{j\in\mathbb{Z}} G(j;-\theta,1/v,\mu) \tag{42}$$

and in our case that works out to:

$$\sum_{j\in\mathbb{Z}} G\left(j; -\frac{\langle x_0, z\rangle}{P}, \frac{1}{v\|x_0\|^2}, -\frac{\langle y_0, x_0\rangle}{\|x_0\|^2}\right) \tag{43}$$

Since $\langle y_0, x_0\rangle \approx 0$, this makes the behaviour more obvious: when $\frac{1}{v\|x_0\|^2}$ is large, this is quite small then $\langle x_0, z\rangle/P$ is not close to an integer.

Notice that so far this is not much different than just using the boxes. Does the real Gaussian give us any extra information? I think it's worse, and here's a short proof: Adding up a Gaussian with phases is like taking the inner product of a vector of Gaussian entries, with a vector of powers of roots of Pth roots of unity. Suppose we center the roots of unity vector, i.e., the vector is

$$\left(e^{2\pi i\frac{-P/2}{P}}, e^{2\pi i\frac{-P/2+1}{P}}, \ldots, e^{2\pi i\frac{-1}{P}}, 1, e^{2\pi i\frac{1}{P}}, \ldots, e^{2\pi i\frac{P/2}{P}}\right) \tag{44}$$

and we want something orthogonal. Consider vectors of the form:

$$\left(\underbrace{0,\ldots,0}_{(P-k)/2}, \underbrace{1,\ldots,1}_{k}, \underbrace{0,\ldots,0}_{(P-k)/2}\right) \tag{45}$$

That is, it's a centered all-ones vector, but only of length k (please ignore my off-by-one errors). Clearly (a) the inner product of any of these with the centered roots of unity is real (since we always add a phase and its negation), (b) the smaller k is, the less orthogonal this is to roots of unity.

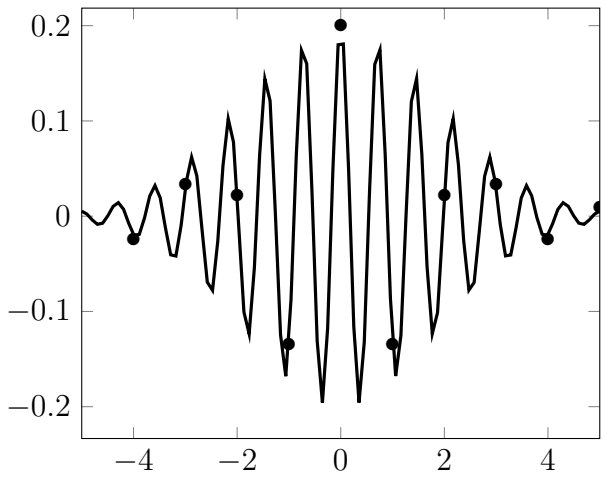


Figure 3: Wide phased Gaussian (real part). The phases causes it to oscillate around 0, so it's likely to add up to close 0 unless either (a) the variance is quite small, (b) the phases are somehow cancelled out

But, we can form a Gaussian distribution as a convex combination of these centered all-ones vectors. Or at least, a truncated Gaussian distribution., where x values that are too large are cut off However, notice that in Chen's algorithm the Gaussian *must* be truncated, because we only started with a bounded radius anyway. Thus, the real Gaussian gives us a *worse* selection of mostly-orthogonal z than the uniform boxes.

3.3 Third Attempt: Complex Gaussians

Recall that when we used the Gaussian, the state after the QFT is

$$\sum_{z\in\mathbb{Z}_p^n} |z\rangle\, e^{2\pi i\langle z,y_0\rangle/P} \exp\left(-\pi v\left(\frac{\langle y_0,x_0\rangle^2}{\|x_0\|^2} - \|y_0\|^2\right)\right) \sum_{j\in\mathbb{Z}} G\left(j; -\frac{\langle x_0,z\rangle}{P}, \frac{1}{v\|x_0\|^2}, -\frac{\langle y_0,x_0\rangle}{\|x_0\|^2}\right) \tag{46}$$

What if the variance v was complex?

If v is complex, we end up with some extra phase terms depending on the input j to the Gaussian, where the phase grows *quadratically* with j. This is a bit strange.

If the phase grew only linearly with j, then if the phase coefficient of j matched the phase multiplier (more specifically: if the linear phase coefficient of j equalled $\frac{\langle y_0,x_0\rangle}{\|x_0\|^2}$), then the phases would cancel out and we would have constructive interference over the entire QFT. Or more specifically, if only specific values of z could make this happen, that would amplify those z.

Chen notices that this can happen, sort of: suppose that we happened to choose v such that the imaginary part of $\frac{1}{v\|x_0\|^2}$ is close to $2\mathbb{Z}$. Then

$$G\left(j; -\frac{\langle x_0,z\rangle}{P}, \frac{1}{v\|x_0\|^2}, -\frac{\langle y_0,x_0\rangle}{\|x_0\|^2}\right) = \exp\left(-\frac{\pi}{v\|x_0\|^2}\left(j^2 + 2j\frac{\langle x_0,z\rangle}{P} + \frac{\langle x_0,z\rangle^2}{P^2}\right)\right) e^{-2\pi i\frac{\langle y_0,x_0\rangle}{\|x_0|^2}} \tag{47}$$

Since $j^2 \in \mathbb{Z}$ and the imaginary part of $\frac{1}{v\|x_0\|^2}$ is close to $2\mathbb{Z}$, then the quadratic part of the phase vanishes from the equation because it is always an integer multiple of $2\pi i$. Let's write $\frac{1}{v\|x_0\|^2} = \eta + i\nu$ for real η and ν (where $\nu \in 2\mathbb{Z}$), so we can rewrite the above equation as a real Gaussian with some phase:

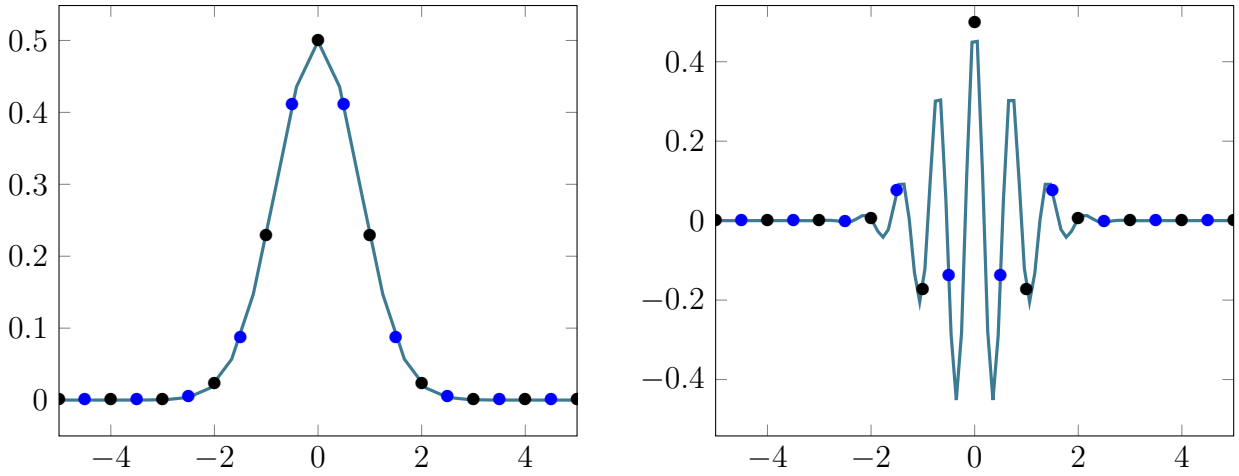


Figure 4: On the left, the Gaussian is too wide to be a good distinguisher: the blue sum to 0.999993 and the black sums to 1.000006. On the right, the phase reduces the sums to around 0.14. Chen's "Karst waves" seem to give the reduction from both.

$$= G\left(j, -\frac{\langle x_0, z\rangle}{P}, \underbrace{\eta}_{\text{real variance}}, \underbrace{-\frac{\langle x_0, y\rangle}{\|x_0\|^2} - \nu\frac{\langle x_0, z\rangle}{P}}_{\text{phase part}}\right) \tag{48}$$

Does this help us at all? I don't know that it does. The problem is that *after* we did the Poisson Summation formula, which flipped the variance, we assumed the variance became rather small. Thus, the sum over j probably concentrates quite sharply the integers, and the sum will be large when $\frac{\langle x_0,z\rangle}{P}$ is close to an integer, and small otherwise. Having a different phase multiplier doesn't really matter because the bulk of the amplitude is on just one j, so there is no opportunity for phases in the sum over j to interfere constructively or destructively.

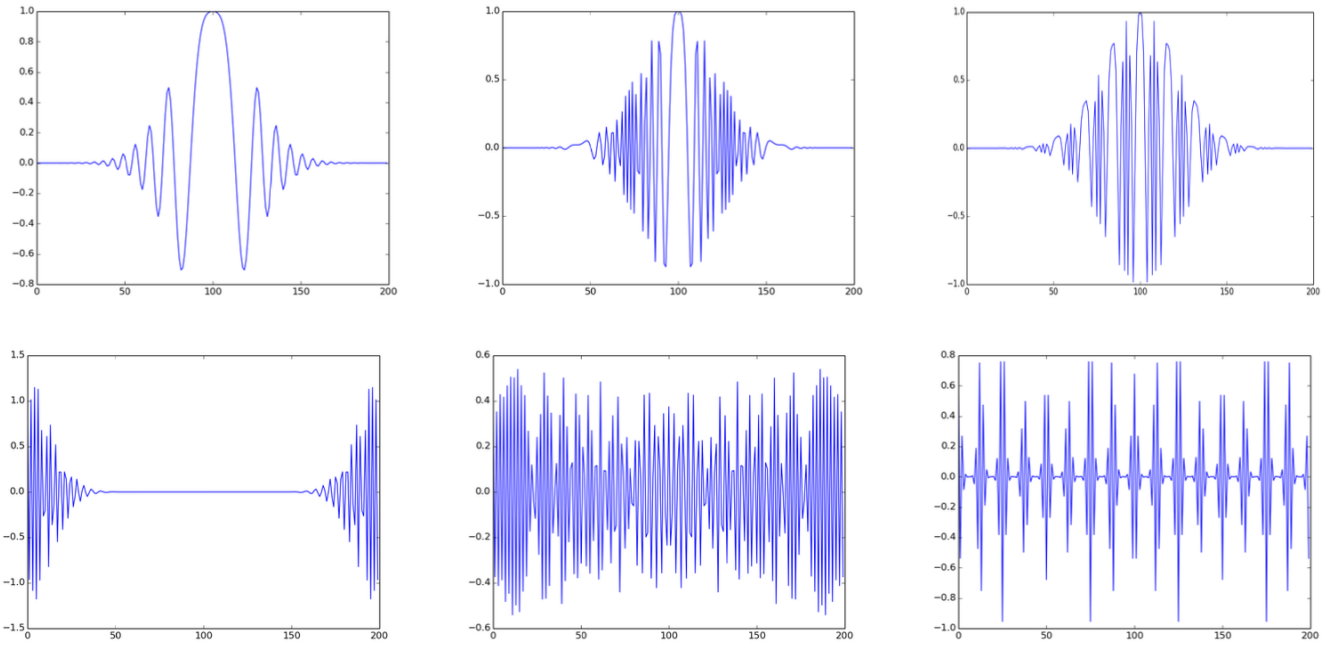
However, the interference wasn't perfect before; we expected that values of z for which $\frac{\langle x_0,z\rangle}{P}$ are not near an integer will end up being measured with non-negligible probability. With the phase, we get an extra suppressive effect: when $\nu\frac{\langle x_0,z\rangle}{P}$ is not close to an integer, there is an extra phase in the sum and it should add up to something closer to 0. See Figure 4 for an attempt at illustrating this.

(Well, really we want $-\frac{\langle x_0,y\rangle}{\|x_0\|^2} - \nu\frac{\langle x_0,z\rangle}{P}$ near an integer, but as argued before I think $\frac{\langle x_0,y\rangle}{\|x_0\|^2} - \approx 0$ anyway so I feel justified ignoring this term).

At this point I have no idea how much this suppressive phase effect helps. It seems like it could give a better GapSVP to LWE reduction, but I'm not sure beyond that.

Karst Waves. This phenomenon where the imaginary part of $\frac{1}{v}$ is close to $2\mathbb{Z}$ is what creates what Chen calls "Karst Waves". My explanation is completely different from Chen's on why they seem to be useful.

Let's check Figure 1 from Chen's paper.



The top rows are the real parts of complex Guassians, and the bottom are their Fourier transforms. The Fourier transform is essentially the sum with different phase multiples, so when the size of the Fourier transformed value is small, then that means we are unlikely to measure z such that $\langle x_0, z\rangle/P$ has this value.

To make this more concrete, look at the graph on the left. Here the complex part is rather small; this is pretty close to just a real Gaussian. We can see that the DFT concentrates around 0, meaning we are basically selecting $\langle x_0, z\rangle/P \approx 0$. That is, it cancels out unless the phase is 0.

In the middle we have a large quadratic phase in the complex Gaussian. Now sometimes it interferes constructively with the linear phase of the DFT, and sometimes it interferes destructively. The DFT seems to randomly be large and small; Chen calls this "chaotic". It will be hard so extract information about $\langle x_0, z\rangle$ from this.

On the right we have "Karst Waves". The imaginary part of the variance is twice an integer so the complex Gaussian effectively adds a linear phase, which either interferes constructively or destructively with the DFT. We might ask, however: why doesn't it create a narrower interference?

The problem is that if the imaginary part (which I labelled as ν above) is twice an integer, then the new linear part of the phase is $\frac{\nu}{P}\langle x_0, z\rangle$. This will interfere constructively not only when $\langle x_0, z\rangle \in P\mathbb{Z}$, but when $\langle x_0, z\rangle \in \frac{P}{\nu}\mathbb{Z}$. Thus, we end up with more approximate modular equivalences (and possibly less approximate?) but over a smaller modulus, i.e., $\langle x_0, z\rangle \approx 0 \mod \frac{P}{\nu}$.

Is this useful? I have no idea; it would really depend on whether the final approximation is tighter than the real Gaussian/uniform case (if it's not tighter, there's no point decreasing the modulus size!). I *think* it must be, at least a little bit, because as I previously argued we have the suppression around non-integer values from the real part of the Gaussian, and the suppression from the phase. These seem like independent effects, so I think we get them working together.

Can a tighter approximation (i.e., smaller error) but a smaller modulus help? It depends. Certainly it's a nice tool to have when analyzing LWE problems. And if we can suppress the error sufficiently smaller

then 1, then there is *no* error and we could solve the system linear algebra. I *think* that's what Chen tried to do, ultimately.

But a non-zero error might be useful. An interesting part of this reduction is that each z we measure is independently random. Thus, we have a polynomial time quantum oracle to produce new LWE samples. Having lots of samples is powerful: for example, with binary error, we only need $O(n^2)$ samples to break LWE in polynomial time over any modulus (the Arora-Ge attack).

Thus, if the Karst waves allow us to suppress the error enough, then maybe we can use other techniques to break the resulting LWE instance faster than expected.

4 The Rest of The Paper

Once these Karst waves are produced, the rest of the paper is:

1. do another convolution with another complex Gaussian
2. measure the top bits and repeat
3. do more convolutions?
4. do something with primes?

Without understanding much about the latter parts of the paper, I'm skeptical that they can fix things. At the end of step 2, with the interference as above, what we essentially have is a superposition over modified LWE samples. The remaining steps don't pass my "sniff test" that they are doing enough to extract important information from what they're given. Partly, I just have no intuition for why these steps should do anything helpful.

I have *some* explanation on the effect of measuring the top bits, which I might add to this document at a later date.

5 Opinionated Outlook

Overall, my take is:

- a lot of the ideas are well-known quantum lattice ideas;
- a number of ideas *are* genuinely new, and the complex Gaussians do seem useful;
- we might be able to patch up the arguments with the complex Gaussians and get a correct but less powerful result;
- I don't expect anyone to take these results to something that could break conservative LWE parameters like Kyber or Dilithium