We now analyze the privacy and accuracy of the interactive mechanism from last lecture. The privacy analysis of this mechanism is fairly straightforward.

1 Analysis of the interactive mechanism

There are two sources of public information release in the interactive update mechanism: The noisy estimate checker $N(x)$ and the approximation mechanism $Apx^k(x)$. In Theorem 5 in the last lecture we showed that $Apx^k(x)$ is $O(k/(\varepsilon n))$-differentially private.

The output of the estimate checker is almost subsumed in the approximation mechanism's answer: $Apx^k(x)$ outputs $\perp$ when and only when $N(x)$ outputs `correct`. When $N(x)$ outputs `too high` or `too low`, $Apx^k(x)$ outputs an actual number a' that approximates $\overline{q}(x)$. Owing to random noise, the sign of $a' - \overline{q}(y)$ may not always be consistent with the answer of $N(x)$. However, I don't think it should be to hard to show that the additional "bit of information" provided by $N(x)$ does not affect privacy along the same lines as the proof of Lemma 1 from last lecture.

So we can conclude that interactive mechanism is $O(k/(\varepsilon n))$ differentially private. Which value of k should we choose?

Remember that k is the number of times that $N(x)$ provides an estimate that is in`correct`. In the analysis of the multiplicative update algorithm (Theorem 3 from last lecture) we showed that if $E(x)$ was used as an estimate checker instead of $N(x)$ then the number of in`correct` answers of $E(x)$ (when queried by MW) can be at most $\ln|D|/\alpha^2$. So we set $k = \ln|D|/\alpha^2$ to obtain a privacy guarantee of $\ln|D|/(\alpha^2\varepsilon n)$; but now we need to argue that replacing the checker $E(x)$ by $N(x)$ does not increase the number of in`correct` answers. The following claim tells us that with high probability, $N(x)$ behaves like $E(x)$.

Claim 1. *For every x and every input $(\overline{q}, a)$, the probabilities of the following events are bounded by $O(\exp(-\alpha/\varepsilon))$:*

- *$N(x)$ outputs* `too high` *or* `too low`*, assuming that $E(x)$ outputs* `correct`*;*
- *$N(x)$ outputs* `too high`*, assuming that $E(x)$ outputs* `too low`*;*
- *$N(x)$ outputs* `too low`*, assuming that $E(x)$ outputs* `too high`*.*

Proof. We work out of the four cases; the proofs to the others are similar. Assume $E(x)$ outputs `correct`. Then $a \leq \overline{q}(x) + 2\alpha$; for $N(x)$ outputs `too high` only when $a > \overline{q}(x) + N/n + 3\alpha$, so it must be that $N \geq \alpha n$. By the tail bound for the Laplace distribution you showed in the homework, the probability of this is at most $O(\exp(-\alpha/\varepsilon))$. □

If we take a union bound over all the $|Q|$ queries that the interactive mechanism issues over its lifetime, we conclude that the probability that the event in Claim 1 happens for at least one query is at most $O(|Q|\exp(-\alpha/\varepsilon))$. If we set $\varepsilon = \alpha/2\log|Q|$, we conclude that with probability at least $1 - 1/|Q|$, $N(x)$ does not answer in*correct* more than k times and the mechanism is $\ln|D|/(\alpha^2\varepsilon n) = O(\log|D|\log|Q|/(\alpha^3 n))$-differentially private.

What about the accuracy of this mechanism? All of the its answers will be within an additive term of 4α from the true answer as long as the noise generated by the approximation mechanism never exceeds α. For a single query, the tail bound for the Laplace distribution tells us that this bad event happens with probability at most $O(\exp(-\alpha/\varepsilon)) = O(1/|Q|^2)$. Taking a union bound, we conclude that the probability that an answer deviating from the true value by more than 4α is produced over the lifetime of the mechanism is at most $O(1/|Q|)$.

We just sketched a proof of the following theorem:

Theorem 2. *For every parameter $\alpha > 0$, domain D, and query sequence Q, the interactive mechanism for averaging queries is $O(\log|D|\log|Q|/(\alpha^3 n))$-private and the answers to all queries is within an additive error of 4α from the true value with probability $1 - O(1/|Q|)$.*

If we set the accuracy α to be $O((\log|D|\log|Q|)/n)^{1/3}$, this mechanism gives us constant differential privacy. In contrast, the Blum-Ligett-Roth mechanism has $O(1/\alpha n)$-differential privacy (assuming a small constant probability of the mechanism failing to be accurate).

The accuracy of the mechanism can be improved if we relax our notion of differential privacy.

2 Almost always differential privacy and the product mechanism

The best we could say about the k-fold product of an ε-differentially private mechanism, is that it is $k\varepsilon$-differentially private. Let us show that this bound is, in general, tight by looking at the Laplace mechanism with parameter ε for a counting query q. Let us assume that $q(x) = 0$ and $q(x') = 1$. The privacy of this mechanism is the maximum value of the log-ratio

$$\ln \frac{\Pr[q(x) + N = y]}{\Pr[q(x') + N = y]} = \ln \frac{\Pr[N = y]}{\Pr[N = y - 1]}$$

over all pairs of adjacent databases x, x' and all possible outcomes y. If we take the outcome $y = 1$, this ratio equals exactly ε, showing our analysis of the Laplace mechanism was tight.

In general, the privacy of the k-fold product of this mechanism (for the same query) is the maximum of the ratio

$$\ln \frac{\Pr[q(x) + N_1 = y_1]\cdots\Pr[q(x) + N_k = y_k]}{\Pr[q(x') + N_1 = y_1]\cdots\Pr[q(x) + N_k = y_k]} = \ln \frac{\Pr[N_1 = y_1]}{\Pr[N_1 = y_1 - 1]} + \cdots + \frac{\Pr[N_k = y_k]}{\Pr[N_k = y_k - 1]}$$

over all possible sequences of answers $(y_1, \ldots, y_n)$. If we take the sequence $(y_1, \ldots, y_n) = (1, \ldots, 1)$ we obtain a privacy loss of exactly $k\varepsilon$, showing that the analysis for the k-fold product of the Laplace mechanism is also tight.

However, this "lower bound" can be criticized on the grounds that the answer sequence $(1, \ldots, 1)$ is far from typical. In fact, for this sequence of queries, we would expect the output y_i of the product mechanism to be positive about half the time and negative about half the time (and zero about an ε fraction of the time). The positive y_i's contribute ε to the summation while the negative ones contribute $-\varepsilon$, so they cancel out each other! So in a "typical case", the privacy loss of the product mechanism should be proportional to the *discrepancy* between the number of positive and the number of negative answers. This is typically on the order of $\sqrt{k}$ (the standard deviation of the unbiased binomial, which models the choice of sign) and not k.

To take advantage of this insight we need to change the definition of differential privacy in a way that distinguishes between "typical" and "atypical" outputs of the mechanism. When the output (sequence) is atypical, there will be no privacy guarantee. Informally, we achieve this by assigning a small failure probability δ to the event that the output $M(x)$ is atypical.

Definition 3. A mechanism M is (ε, δ)-differentially private if for every set S of outcomes of the mechanism and all pairs of adjacent databases x, x':

$$\Pr[M(x) \in S] \leq e^{\varepsilon} \Pr[M(x') \in S] + \delta.$$

We now give an improved analysis for the privacy of the product mechanism under this definition:

Theorem 4. *if M is ε-differentially private and $\varepsilon \leq 1.79$, then M^k is $(3k\varepsilon^2, e^{-k\varepsilon/2})$-differentially private.*

When ε is larger than $1/k$ (but still smaller than $1/\sqrt{k}$), this theorem tells us that the k-fold product mechanism is still almost always $o(1/k)$-private.

We will prove this theorem in the non-interactive setting; it is not difficult to extend it to the interactive case using some tools from probability.

Fix a pair of adjacent databases x and x' and let $p(y)$ and $q(y)$ denote the probabilities of the events $M(x) = y$ and $M(x') = y$, respectively. Consider the probability space of outcomes of $M(x)$ and let L be a random variable that takes value $\ln(p(y)/q(y))$ when outcome y happens, i.e. with probability $p(y)$. Then L always takes values in the range $[-\varepsilon, \varepsilon]$ and

$$\mathrm{E}[L] = \sum_y p(y) \ln(p(y)/q(y)) = \mathrm{Div}(p\|q).$$

It turns out that this divergence is significantly smaller than ε.

Lemma 5. *If $\ln(p(y)/q(y)) \in [-\varepsilon, \varepsilon]$ for all y then $\mathrm{Div}(p\|q) + \mathrm{Div}(q\|p) \leq \varepsilon(e^{\varepsilon} - 1)$.*

Proof. We write

$$\begin{aligned}\mathrm{Div}(p\|q) + \mathrm{Div}(q\|p) &= \sum_y (p(y) - q(y)) \ln \frac{p(y)}{q(y)} \\ &\leq \varepsilon \sum_y |p(y) - q(y)| \\ &\leq \varepsilon \sum_y (e^\varepsilon - 1) \min\{p(y), q(y)\} \\ &\leq \varepsilon(e^\varepsilon - 1).\end{aligned}$$

□

In the k-fold product mechanism $M^k(x)$, the privacy loss (as a function of the outcome) is given by the random variable $L_1 + \cdots + L_n$, where L_i is the privacy loss of the i-th instantiation of the mechanism $M(x)$. The random variables L_i are independent copies that take values in the range $[-\varepsilon, \varepsilon]$ and have mean value at most $\varepsilon(e^\varepsilon - 1)$. After appropriate scaling and shifting we can apply the Chernoff bound to obtain:

$$\Pr\left[|L_1 + \cdots + L_n - k\varepsilon(e^\varepsilon - 1) > k\varepsilon^2|\right] \leq e^{-k\varepsilon/2}$$

and conclude that with probability $1 - e^{-k\varepsilon^2}/2$, the privacy loss of the product mechanism is at most $k\varepsilon(e^\varepsilon - 1) + k\varepsilon^2 \leq 3k\varepsilon^2$.

Since the approximation mechanism is a product mechanism, we can apply Theorem 4 to obtain an improved analysis of it. This improvement carries over to the interactive data release mechanism. After going over the same steps in the analysis, one concludes that the interactive data release mechanism is (ε, e^{-t})-differentially private and accurate within α with probability $1 - 1/|Q|$ for

$$\varepsilon = O\left(\frac{(\log|Q|)^2(\log|D|)}{\alpha^4 n^2}\right) \quad \text{and} \quad t = \Omega\left(\frac{(\log|Q|)(\log|D|)}{\alpha^3 n}\right).$$

For example, when α is a small constant and n is slightly larger than $(\log|Q|)\sqrt{\log|D|}$, this bound tells us that (for a suitable choice of α) the resulting mechanism is almost always differentially private, while Theorem 2 gives no guarantee.

References

These notes are based on Chapters 2 and 4 of the survey *The Algorithmic Foundations of Differential Privacy* by Cynthia Dwork and Aaron Roth.