

fn make_bounded_range_to_zCDP

Tudor Cebere

February 26, 2025

This proof resides in “**contrib**” because it has not completed the vetting process.

Proves soundness of `bounded_range_to_zCDP` in `mod.rs` at commit `0b8f4222` (outdated¹). The conversion between bounded range [DR19] and zCDP comes from Lemma 3.2 in [CR20]. The proof in this document is an adaptation of Theorem 5 [here](#).

1 Hoare Triple

Preconditions

Compiler-verified

- Variable `meas` is a valid measurement of type `Measurement<DI, T0, MI, RangeDivergence>`
- Generic `DI` (input domain) is a type with trait `Domain`.
- Generic `MI` (input metric) is a type with trait `Metric`.
- `MetricSpace` is implemented for `(DI, MI)`. Therefore `MI` is a valid metric on `DI`.

Human-verified

None

Pseudocode

```
1 def make_bounded_range_to_zCDP(meas: Measurement) -> Measurement:
2     def privacy_map(d_in: f64) -> f64:
3         return meas.map(d_in).inf_powi(ibig(2)).inf_div(8.0)
4
5     return meas.with_map( #
6         meas.input_metric,
7         ZeroConcentratedDivergence,
8         PrivacyMap.new_fallible(privacy_map),
9     )
```

Postcondition

Theorem 1.1 (Postcondition). For every setting of the input parameters (`meas`, `DI`, `T0`, `MI`) to `make_bounded_range_to_zCDP` such that the given preconditions hold, `make_bounded_range_to_zCDP` raises an exception (at compile time or run time) or returns a valid measurement. A valid measurement has the following property:

¹See new changes with `git diff 0b8f4222..35b51954 rust/src/combinators/measure_cast/bounded_range_to_zCDP/mod.rs`

1. (Privacy guarantee). For every pair of elements x, x' in `input_domain` and for every pair (d_in, d_out) , where d_in has the associated type for `input_metric` and d_out has the associated type for `output_measure`, if x, x' are d_in -close under `input_metric`, `privacy_map(d_in)` does not raise an exception, and `privacy_map(d_in) ≤ d_out`, then `function(x), function(x')` are d_out -close under `output_measure`.

By the precondition, `meas` is a valid measurement with `RangeDivergence` privacy measure.

Definition 1.2 (Range Divergence). For any two distributions Y, Y' and any non-negative d , Y, Y' are d -close under the bounded-range privacy measure whenever

$$D_{BR}(Y, Y') = \sup_{y_0, y_1 \in \text{Supp}(Y)} \mathcal{L}_{Y, Y'}(y_0) - \mathcal{L}_{Y, Y'}(y_1) \quad (1)$$

Definition 1.3 (Privacy Loss). The *privacy loss* of an outcome y with respect to random variables Y and Y' is defined as

$$\mathcal{L}_{Y, Y'}(y) = \ln \left(\frac{\mathbb{P}[Y = y]}{\mathbb{P}[Y' = y]} \right). \quad (2)$$

If y is not in the support of Y' then we define the privacy loss as infinite. The *privacy loss random variable* Z is distributed according to $\mathcal{L}_{Y, Y'}(y)$ where y is obtained by sampling $y \sim Y$.

Lemma 1.4. Given a privacy loss random variable Z with respect to random variables Y and Y' , if $\text{supp}(Y) = \text{supp}(Y')$, then

$$\mathbb{E}_{y \sim Y} [\exp(-Z)] = 1. \quad (3)$$

Proof.

$$\mathbb{E}_{y \sim Y} [\exp(-Z)] \quad (4)$$

$$= \int_{\text{supp}(Y)} \exp(-Z) dY \quad (5)$$

$$= \int_{\text{supp}(Y)} \mathbb{P}[Y = y] \cdot \exp(-\mathcal{L}_{Y, Y'}(y)) dy \quad \text{by Definition 1.3} \quad (6)$$

$$= \int_{\text{supp}(Y)} \mathbb{P}[Y = y] \cdot \frac{\mathbb{P}[Y' = y]}{\mathbb{P}[Y = y]} dy \quad (7)$$

$$= \int_{\text{supp}(Y')} \mathbb{P}[Y' = y] dy \quad \text{since } Y \text{ and } Y' \text{ have the same support} \quad (8)$$

$$= 1 \quad (9)$$

□

Definition 1.5 (Hoeffding's Lemma). Let X be a random variable supported on $[a, b]$. Then for any $\lambda \in \mathbb{R}$,

$$\mathbb{E}[\exp(\lambda X)] \leq \exp \left(\mathbb{E}[X] \cdot \lambda + \frac{(b-a)^2}{8} \cdot \lambda^2 \right) \quad (10)$$

Lemma 1.6. Given two distributions Y and Y' that are η -close under range divergence, then the associated privacy loss random variable satisfies

$$\mathbb{E}[Z] \leq \frac{1}{8} \eta^2 \quad (11)$$

Proof.

$$\mathbb{E}[\exp(-Z)] \leq \exp\left(-\mathbb{E}[Z] + \frac{\eta^2}{8}\right) \quad \text{by 1.5 since } Z \in [-t, \eta - t], \text{ let } \lambda = -1 \quad (12)$$

$$\implies \mathbb{E}[Z] \leq \frac{\eta^2}{8} - \log \mathbb{E}[\exp(-Z)] \quad \text{rearrange terms} \quad (13)$$

$$= \frac{\eta^2}{8} \quad \text{by Lemma 1.4} \quad (14)$$

Lemma 1.4 can only be applied when Y and Y' have the same support. This requirement is satisfied via a proof by contradiction: if Y and Y' have different supports, then the privacy loss would be infinite, meaning that Y and Y' are not η -close under range divergence. \square

Definition 1.7 (zero-Concentrated Divergence Privacy Loss Random Variable). For a privacy loss random variable Z with respect to two distributions Y, Y' and any non-negative d , Y, Y' are d -close under the zero-concentrated divergence measure if, for every possible choice of $\alpha \in (1, \infty)$,

$$\mathbb{E}[\exp(\alpha Z)] \leq \exp(\alpha(\alpha + 1)d) \quad (15)$$

Theorem 1.8 (Range Divergence implies zero-Concentrated Divergence). If two random variables Y and Y' are η -close under range divergence, then they are also $\frac{1}{8}\eta^2$ -close under zero-concentrated divergence.

Proof.

$$\mathbb{E}[\exp(\alpha Z)] \quad \text{starting from Definition 1.7} \quad (16)$$

$$\leq \exp\left(\mathbb{E}[Z]\alpha + \frac{\eta^2}{8}\alpha^2\right) \quad \text{by 1.5 since } Z \in [-t, \eta - t], \text{ let } \lambda = \alpha \quad (17)$$

$$\leq \exp\left(\frac{\eta^2}{8}\alpha + \frac{\eta^2}{8}\alpha^2\right) \quad \text{by Lemma 1.6} \quad (18)$$

$$= \exp\left(\frac{\eta^2}{8}\alpha(\alpha + 1)\right) \quad (19)$$

$$= \exp(\alpha(\alpha + 1)\rho) \quad \text{where } \rho = \frac{\eta^2}{8} \quad (20)$$

\square

Proof of Theorem 1.1. By the postcondition of `Measurement.with_map`, on line 5, `make_bounded_range_to_zCDP` returns a measurement with the same input metric, output metric `ZeroConcentratedDivergence`, and a privacy map that computes $\eta^2/8$ with conservative rounding. The privacy guarantee holds by the precondition that `meas` is valid measurement, together with Theorem 1.8. Therefore the returned measurement is a valid measurement. \square

References

- [CR20] Mark Cesar and Ryan Rogers. Unifying privacy loss composition for data analytics. *CoRR*, abs/2004.07223, 2020.
- [DR19] David Durfee and Ryan Rogers. Practical differentially private top-k selection with pay-what-you-get composition. *CoRR*, abs/1905.04273, 2019.