Verifying Snapping Mechanism - Floating Point Implementation Version

In order to verify the differential privacy property of an implementation of the snapping mechanism [5], we follow the logic rules designed from [1] and the floating point error semantics from [7, 4, 2, 6].

1 Preliminary Definitions

Definition 1 (Laplace mechanism [3])
Let $\epsilon > 0$. The Laplace mechanism $\mathcal{L}_\epsilon$: $\mathbb{R} \to \mathsf{Distr}(\mathbb{R})$ is defined by $\mathcal{L}(t) = t + v$, where $v \in \mathbb{R}$ is drawn from the Laplace distribution $\mathsf{laplce}(\frac{1}{\epsilon})$.

2 Syntax

Following are the syntax of the system. The circled operators are rounded operation in floating point computation.

Floating Point Expr.	$e_\mathbb{F}$	::=	$c \mid x_\mathbb{F} \mid f(x_\mathbb{F}) \mid e_\mathbb{F} \circledast e_\mathbb{F} \mid \textcircled{n}(e_\mathbb{F}) \mid x_\mathbb{F} \xleftarrow{\$} \mu$
Real Expr.	$e_\mathbb{R}$	::=	$r \mid x_\mathbb{R} \mid F(x_\mathbb{R}) \mid e_\mathbb{R} * e_\mathbb{R} \mid \ln(e_\mathbb{R}) \mid x_\mathbb{R} \xleftarrow{\$} \mu$
Arithmetic Operation	$*$	::=	$+ \mid - \mid \times \mid \div$
Value	v	::=	$r \mid c$
Distribution	μ	::=	$\mathsf{laplce} \mid \mathsf{unif} \mid \mathtt{bernoulli}$
Error	err	::=	$(e_\mathbb{R}, e_\mathbb{R})$

We use upper case for variables in real computation and lower case for variables in floating point computation.

$F(x_\mathbb{R})$ denotes function F evaluates to value $F(x_\mathbb{R})$ given input $x_\mathbb{R}$ in real computation, and $f(x_\mathbb{F})$ denotes the same function F evaluates to value $f(x_\mathbb{F})$ given the same input $x_\mathbb{F}$ in floating point computation.

3 Semantics

The big step semantics with relative floating point computation error are shown in Figure. 1. The semantics are $e_\mathbb{R} \Downarrow e_\mathbb{F}, err$, which means a real world expression $e_\mathbb{R}$ can be represented in floating point computation $e_\mathbb{F}$ with error bound err. The η is the machine epsilon.

Theorem 1 (Soundness Theorem)
Given $e_\mathbb{R}$ and $e_\mathbb{F}$ where $e_\mathbb{R} \Downarrow e_\mathbb{F}, err$, when evaluating the $e_\mathbb{F}$ in floating point computation and get the value c, we have $c \in err$.

$$\frac{c = \mathtt{fl}(r)}{r \Downarrow c, \big(\frac{r}{(1+\eta)}, r(1+\eta)\big)}\ \text{CONST} \qquad \frac{e^1_{\mathbb{R}} \Downarrow e^1_{\mathbb{F}}, (\underline{e^1_{\mathbb{R}}}, \bar{e^1_{\mathbb{R}}}) \qquad e^2_{\mathbb{R}} \Downarrow e^2_{\mathbb{F}}, (\underline{e^2_{\mathbb{R}}}, \bar{e^2_{\mathbb{R}}})}{e^1_{\mathbb{R}} * e^2_{\mathbb{R}} \Downarrow \mathtt{fl}(e^1_{\mathbb{F}} \circledast e^2_{\mathbb{F}}), \big((\frac{\underline{e^1_{\mathbb{R}}} * \underline{e^2_{\mathbb{R}}})}{(1+\eta)}, (\bar{e^1_{\mathbb{R}}} * \bar{e^2_{\mathbb{R}}})(1+\eta)\big)}\ \text{OP}$$

$$\frac{e_{\mathbb{R}} \Downarrow e_{\mathbb{F}}, (\underline{e^1_{\mathbb{R}}}, \bar{e^2_{\mathbb{R}}})}{\ln(e_{\mathbb{R}}) \Downarrow \textcircled{ln}(e_{\mathbb{F}}), \big((\frac{\textcircled{ln}(\underline{e^1_{\mathbb{R}}})}{(1+\eta)}, (\ln(\bar{e^2_{\mathbb{R}}}))(1+\eta)\big)}\ \text{LN}$$

Figure 1: Semantics with Relative Floating Point Error

4 Snapping Mechanism

Definition 2 ($\mathsf{Snap}_{\mathbb{R}}(a) : A \to \mathsf{Distr}(\mathbb{R})$)
Given privacy parameter ϵ, the ideal Snapping mechanism $\mathsf{Snap}_{\mathbb{R}}(a)$ is defined as:

$$u \xleftarrow{\$} \mu; s \xleftarrow{\$} \{-1, 1\}; Y = \ln(u) \div \epsilon; Z = s \times Y; X = F(a); W = X + Z; W' = \lfloor W \rceil_{\Lambda}; R = \mathsf{clamp}_B(W')$$

where f is the query function over input $a \in A$, ϵ is the privacy budget, B is the clamping bound and Λ is the rounding argument satisfying $\lambda = 2^k$ where 2^k is the smallest power of 2 greater or equal to the $\frac{1}{\epsilon}$.

Let $\mathsf{Snap}'_{\mathbb{R}}(a, u, s)$ be the same as $\mathsf{Snap}_{\mathbb{F}}(a)$ except the sample processes $u \xleftarrow{\$} \mu; s \xleftarrow{\$} \{-1, 1\}$ being finished with sample results u and s.

Definition 3 ($\mathsf{Snap}_{\mathbb{F}}(a) : A \to \mathsf{Distr}(\mathbb{R})$)
Given privacy parameter ϵ, the floating point implemented Snapping mechanism $\mathsf{Snap}_{\mathbb{F}}(a)$ is defined as (where all parameters are defined the same as above):

$$u_{\mathbb{F}} \xleftarrow{\$} \mu; s_{\mathbb{F}} \xleftarrow{\$} \{-1, 1\}; y = \textcircled{ln}(u) \oslash \epsilon; z = s \otimes y; x = f(a); w = x \oplus z; w' = \lfloor w \rceil_{\Lambda}; r = \mathsf{clamp}_B(w')$$

Let $\mathsf{Snap}'_{\mathbb{F}}(a, u, s)$ be the same as $\mathsf{Snap}_{\mathbb{F}}(a)$ except the sample processes $u \xleftarrow{\$} \mu; s \xleftarrow{\$} \{-1, 1\}$ being finished with sample results u and s.

5 Main Theorem

Theorem 2 **(The** Snap **mechanism is** $\epsilon-$**differentially private)**
Consider $\mathsf{Snap}(a)$ defined as before, if $\mathsf{Snap}(a) = x$ given database a and privacy parameter ϵ, then its actual privacy loss is bounded by $\epsilon + 12x\epsilon\eta + 2\eta$

Proof. Given $\mathsf{Snap}_{\mathbb{F}}(a) = x$ and parameter ϵ, we consider a' be the adjacent database of a satisfying $|f(a) - f(a')| \leq 1$. Without loss of generalization, we assume $f(a) + 1 = f(a')$ ($\diamond$). The proof is developed by cases of the output of $\mathsf{Snap}_{\mathbb{F}}(a)$ mechanism.

Consider the $\mathsf{Snap}_{\mathbb{R}}(a)$ outputting the same result x, let (L, R) be the range where $\forall u \in (L, R)$ and some s, $\mathsf{Snap}'_{\mathbb{R}}(a, u, s) = x$, we have $\Pr[\mathsf{Snap}_{\mathbb{R}}(a)] = R - L$. Given the $\mathsf{Snap}_{\mathbb{R}}$ is $\epsilon-$dp, we have:

$$e^{-\epsilon} \leq \frac{\Pr[\mathsf{Snap}_{\mathbb{R}}(a)]}{\Pr[\mathsf{Snap}_{\mathbb{R}}(a)]} = \frac{R - L}{R' - L'} \leq e^{\epsilon}$$

Let (l, r) be the range where $\forall u \in (l, r)$ and some s, $\mathsf{Snap}'_{\mathbb{F}}(a, u, s) = x$, we estimated the $|r - l|$ in terms of floating point relative error and $|R - L|$ through our semantics in order to verify the privacy loss of $\mathsf{Snap}_{\mathbb{F}}$.

case $x = -B$

Let b be the largest number rounded by Λ that is smaller than B. We know $s = 1$, $L = l = 0$ and $R = b$, so we only need to estimate the right side range r in this case. The derivation of this case given $\mathsf{Snap}'_{\mathbb{F}}(a, R, 1) = \mathsf{Snap}'_{\mathbb{F}}(a', R, 1) = x$ is shown as following:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{R \Downarrow r(\underline{r}, \bar{r})}{\ln(R) \Downarrow \textcircled{ln}(r)(\ln(\underline{r})(1+\eta), \frac{\ln(\bar{r})}{(1+\eta)})}}{\frac{1}{\epsilon} \times \ln(R) \Downarrow \frac{1}{\epsilon} \otimes \textcircled{ln}(r), ((\frac{1}{\epsilon} \times \ln(\underline{r}))(1+\eta)^2, \frac{\frac{1}{\epsilon} \times \ln(\bar{r})}{(1+\eta)^2})}}{f(a) + \frac{1}{\epsilon} \times \ln(R) \Downarrow f(a) \oplus \frac{1}{\epsilon} \otimes \textcircled{ln}(r), (\frac{f(a) + (\frac{1}{\epsilon} \times \ln(\underline{r}))(1+\eta)^2}{1+\eta}, (f(a) + \frac{\frac{1}{\epsilon} \times \ln(\bar{r})}{(1+\eta)^2})(1+\eta))}}{\lfloor f(a) + \frac{1}{\epsilon} \times \ln(R) \rceil \Downarrow \lfloor f(a) \oplus \frac{1}{\epsilon} \otimes \textcircled{ln}(r) \rceil, (\frac{f(a) + (\frac{1}{\epsilon} \times \ln(\underline{r}))(1+\eta)^2}{1+\eta}, (f(a) + \frac{\frac{1}{\epsilon} \times \ln(\bar{r})}{(1+\eta)^2})(1+\eta))}}{\mathsf{Snap}'_{\mathbb{R}}(a, R, 1) \Downarrow \mathsf{Snap}'_{\mathbb{F}}(a, r, 1), (\frac{f(a) + (\frac{1}{\epsilon} \times \ln(\underline{r}))(1+\eta)^2}{1+\eta}, (f(a) + \frac{\frac{1}{\epsilon} \times \ln(\bar{r})}{(1+\eta)^2})(1+\eta))}$$

In the same way, we have the derivation for $\mathsf{Snap}'_{\mathbb{F}}(a', r, 1)$:

$$\dfrac{}{\mathsf{Snap}'_{\mathbb{R}}(a', R, 1) \Downarrow \mathsf{Snap}'_{\mathbb{F}}(a', r, 1), (\frac{f(a') + (\frac{1}{\epsilon} \times \ln(\underline{r}))(1+\eta)^2}{1+\eta}, (f(a') + \frac{\frac{1}{\epsilon} \times \ln(\bar{r})}{(1+\eta)^2})(1+\eta))}$$

$$u \in \left(0, (\underline{r}, \bar{r})\right) \wedge (s = -1) \sim u' \in \left(0, (\underline{r}', \bar{r}')\right) \wedge (s = -1)$$

Given $\mathsf{Snap}_{\mathbb{F}}(a) = \mathsf{Snap}_{\mathbb{F}}(a') = x$, we have following values for $\underline{r}, \bar{r}, \underline{r}'$ and $\bar{r}'$:

$$\begin{aligned} u &\in \left(0, (e^{\epsilon(\frac{-b-\frac{\Lambda}{2}}{1+\eta} - f(a))(1+\eta)^2}, e^{\epsilon \frac{(-b-\frac{\Lambda}{2})(1+\eta) - f(a)}{(1+\eta)^2}})\right) \wedge (s = 1) \\ \sim u' &\in \left(0, (e^{\epsilon(\frac{-b-\frac{\Lambda}{2}}{1+\eta} - f(a'))(1+\eta)^2}, e^{\epsilon \frac{(-b-\frac{\Lambda}{2})(1+\eta) - f(a')}{(1+\eta)^2}})\right) \wedge (s = 1) \end{aligned}$$

Given that the probability is equivalent to the length of the range, we have the ratio between u and u' is bounded by:

$$\frac{u}{u'} \leq \frac{\frac{1}{2}\bar{r}}{\frac{1}{2}\underline{r}'} = e^{\epsilon((-b-\frac{\Lambda}{2})(1+\eta-\frac{1}{1+\eta}) + f(a)((1+\eta)^2 - \frac{1}{(1+\eta)^2}) + (1+\eta)^2)} \leq e^{\epsilon(1+\eta)^2 2B} \leq e^{\epsilon + 12B\epsilon\eta + 2\eta}$$

By the AxUnif rule, we have the actual privacy loss is bounded by the same value.

case $x \in (-B, \lfloor f(a) \rceil_{\Lambda})$

The derivation of this case is shown as following:

$$\frac{\frac{\frac{\frac{L \Downarrow l(\underline{l}, \bar{l})}{\ln(L) \Downarrow \textcircled{ln}(l)(\ln(\underline{l})(1+\eta), \frac{\ln(\bar{l})}{(1+\eta)})}}{\frac{1}{\epsilon} \times \ln(L) \Downarrow \frac{1}{\epsilon} \otimes \textcircled{ln}(l), ((\frac{1}{\epsilon} \times \ln(\underline{l}))(1+\eta)^2, \frac{\frac{1}{\epsilon} \times \ln(\bar{l})}{(1+\eta)^2})}}{f(a) + \frac{1}{\epsilon} \times \ln(L) \Downarrow f(a) \oplus \frac{1}{\epsilon} \otimes \textcircled{ln}(l), (\frac{f(a) + (\frac{1}{\epsilon} \times \ln(\underline{l}))(1+\eta)^2}{1+\eta}, (f(a) + \frac{\frac{1}{\epsilon} \times \ln(\bar{l})}{(1+\eta)^2})(1+\eta))}}{\mathsf{Snap}'_{\mathbb{R}}(a, R, 1) \Downarrow \mathsf{Snap}'_{\mathbb{F}}(a, r, 1), (\frac{f(a) + (\frac{1}{\epsilon} \times \ln(\underline{r}))(1+\eta)^2}{1+\eta}, (f(a) + \frac{\frac{1}{\epsilon} \times \ln(\bar{r})}{(1+\eta)^2})(1+\eta))}$$

Following the semantics in Figure 1, we have following evaluation results:

$$u \in \left[(\underline{r_1}, \bar{r_1}), (\underline{r_2}, \bar{r_2})\right) \wedge (s = -1) \sim u' \in \left[(\underline{r'_1}, \bar{r'_1}), (\underline{r'_2}, \bar{r'_2})\right) \wedge (s = -1)$$

[[where $\underline{r_1}, \bar{r_1}, \underline{r_2}, \bar{r_2}, \underline{r'_1}, \bar{r'_1}, \underline{r'_2} and \bar{r'_2}$ have following values:

$$u \in \left[((1-\eta)e^{\epsilon(x-\frac{\Lambda}{2}-f(a))(1-\eta)^2}, (1+\eta)e^{\epsilon(x-\frac{\Lambda}{2}-f(a))(1+\eta)^2}), ((1-\eta)e^{\epsilon(x+\frac{\Lambda}{2}-f(a))(1-\eta)^2}, (1+\eta)e^{\epsilon(x+\frac{\Lambda}{2}-f(a))(1+\eta)^2})\right)$$
$$\sim u' \in \left[((1-\eta)e^{\epsilon(x-\frac{\Lambda}{2}-f(a'))(1-\eta)^2}, (1+\eta)e^{\epsilon(x-\frac{\Lambda}{2}-f(a'))(1+\eta)^2}), ((1-\eta)e^{\epsilon(x+\frac{\Lambda}{2}-f(a'))(1-\eta)^2}, (1+\eta)e^{\epsilon(x+\frac{\Lambda}{2}-f(a'))(1+\eta)^2})\right)$$

Given that the probability is equivalent to the length of the range, we have the ratio between u and u' is bounded by:

$$\frac{u}{u'} \leq \frac{\bar{r_2} - \underline{r_1}}{\underline{r'_2} - \bar{r'_1}} \leq \epsilon + 12x\epsilon\eta + 2\eta$$

By the AxUnif rule, we have the actual privacy loss is bounded by the same value.]]

case $x = \lfloor f(a) \rceil_\Lambda$

Following the semantics in Figure 1, we have following evaluation results:

$$u \in \left[(\underline{r_1}, \bar{r_1}), 1\right] \wedge (s = -1) \vee u \in \left[(\underline{r_2}, \bar{r_2}), 1\right] \wedge (s = 1) \sim u' \in \left[(\underline{r'_1}, \bar{r'_1}), (\underline{r'_2}, \bar{r'_2})\right) \wedge (s = -1),$$

[[where $\underline{r_1}, \bar{r_1}, \underline{r_2}, \bar{r_2}, \underline{r'_1}, \bar{r'_1}, \underline{r'_2} and \bar{r'_2}$ have following values: Given that the probability is equivalent to the length of the range, we have the ratio between u and u' is bounded by:

$$\frac{u}{u'} \leq \frac{1 - \frac{1}{2}(\underline{r_2} + \underline{r_1})}{\frac{1}{2}(\underline{r'_2} - \bar{r'_1})} \leq \epsilon + 12x\epsilon\eta + 2\eta$$

By the AxUnif rule, we have the actual privacy loss is bounded by the same value.]]

case $x \in (\lfloor f(a) \rceil_\Lambda, \lfloor f(a') \rceil_\Lambda)$

Following the semantics in Figure 1, we have following evaluation results:

$$u \in \left((\underline{r_1}, \bar{r_1}), (\underline{r_2}, \bar{r_2})\right] \wedge (s = 1) \sim u' \in \left[(\underline{r'_1}, \bar{r'_1}), (\underline{r'_2}, \bar{r'_2})\right) \wedge (s = -1),$$

[[where $\underline{r_1}, \bar{r_1}, \underline{r_2}, \bar{r_2}, \underline{r'_1}, \bar{r'_1}, \underline{r'_2} and \bar{r'_2}$ have following values: Given that the probability is equivalent to the length of the range, we have the ratio between u and u' is bounded by:

$$\frac{u}{u'} \leq \frac{\frac{1}{2}(\bar{r_2} - \underline{r_1})}{\frac{1}{2}(\underline{r'_2} - \bar{r'_1})} \leq \epsilon + 12x\epsilon\eta + 2\eta$$

By the AxUnif rule, we have the actual privacy loss is bounded by the same value.]]

case $x = \lfloor f(a') \rceil_\Lambda$

Following the semantics in Figure 1, we have following evaluation results:

$$u \in \left((\underline{r_1}, \bar{r_1}), (\underline{r_2}, \bar{r_2})\right] \wedge (s = 1) \sim u' \in \left[(\underline{r'_1}, \bar{r'_1}), 1\right] \wedge (s = -1) \vee \left[(\underline{r'_2}, \bar{r'_2}), 1\right] \wedge (s = 1),$$

[[where $\underline{r_1}, \bar{r_1}, \underline{r_2}, \bar{r_2}, \underline{r'_1}, \bar{r'_1}, \underline{r'_2} and \bar{r'_2}$ have following values: Given that the probability is equivalent to the length of the range, we have the ratio between u and u' is bounded by:

$$\frac{u}{u'} \leq \frac{\frac{1}{2}(\bar{r_2} - \underline{r_1})}{1 - \frac{1}{2}(\underline{r'_2} + \bar{r'_1})} \leq \epsilon + 12x\epsilon\eta + 2\eta$$

By the AxUnif rule, we have the actual privacy loss is bounded by the same value.]]

case $x \in (\lfloor f(a') \rceil_\Lambda, B)$

Following the semantics in Figure 1, we have following evaluation results:

$$u \in \left((\underline{r_1}, \bar{r_1}), (\underline{r_2}, \bar{r_2})\right] \wedge (s = 1) \sim u' \in \left((\underline{r'_1}, \bar{r'_1}), (\underline{r'_2}, \bar{r'_2})\right] \wedge (s = 1),$$

[[where $\underline{r_1}, \bar{r_1}, \underline{r_2}, \bar{r_2}, \underline{r'_1}, \bar{r'_1}, \underline{r'_2} and \bar{r'_2}$ have following values: Given that the probability is equivalent to the length of the range, we have the ratio between u and u' is bounded by:

$$\frac{u}{u'} \leq \frac{\frac{1}{2}(\bar{r_2} - \underline{r_1})}{\frac{1}{2}(\underline{r'_2} - \bar{r'_1})} \leq \epsilon + 12x\epsilon\eta + 2\eta$$

By the AxUnif rule, we have the actual privacy loss is bounded by the same value.]]

case $x = B$

Following the semantics in Figure 1, we have following evaluation results:

$$u \in \left(0, (\underline{r}, \bar{r})\right) \sim u' \in \left(0, (\underline{r'}, \bar{r'})\right),$$

[[where $\underline{r}, \bar{r}, \underline{r'}$ and $\bar{r'}$ have following values: Given that the probability is equivalent to the length of the range, we have the ratio between u and u' is bounded by:

$$\frac{u}{u'} \leq \frac{\frac{1}{2}\bar{r}}{\frac{1}{2}\underline{r'}} \leq \epsilon + 12x\epsilon\eta + 2\eta$$

By the AxUnif rule, we have the actual privacy loss is bounded by the same value.]]

□

References

[1] Gilles Barthe, Marco Gaboardi, Benjamin Grégoire, Justin Hsu, and Pierre-Yves Strub. Proving differential privacy via probabilistic couplings. In *LICS 2016*.

[2] H. Becker, N. Zyuzin, R. Monat, E. Darulova, M. O. Myreen, and A. Fox. A verified certificate checker for finite-precision error bounds in coq and hol4. In *2018 Formal Methods in Computer Aided Design (FMCAD)*, 2018.

[3] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating Noise to Sensitivity in Private Data Analysis. In *TCC*, 2016.

[4] Matthieu Martel. Semantics of roundoff error propagation in finite precision calculations. *Higher-Order and Symbolic Computation*, 2006.

[5] Ilya Mironov. On significance of the least significant bits for differential privacy. In *CCS 2012*, 2012.

[6] Mariano Moscato, Laura Titolo, Aaron Dutle, and César A. Muñoz. Automatic estimation of verified floating-point round-off errors via static analysis. In Stefano Tonetta, Erwin Schoitsch, and Friedemann Bitsch, editors, *Computer Safety, Reliability, and Security*, 2017.

[7] Tahina Ramananandro, Paul Mountcastle, Benoundefinedt Meister, and Richard Lethin. A unified coq framework for verifying c programs with floating-point computations. In *Proceedings of the 5th ACM SIGPLAN Conference on Certified Programs and Proofs (CPP)*. Association for Computing Machinery, 2016.