

1 DiPA with a Counter

1.1 DiPA* and Definitions

Definition 1.1. Fix parameters ϵ, N . Let C be the guard conditions $\{n < N, \text{true}, \text{insample} \geq x, \text{insample} < x, n < N \text{ AND } \text{insample} \geq x, n < N \text{ AND } \text{insample} < x, n \geq N\}$. A **DiP* automaton** (DiPA*) \mathcal{A} is defined as the tuple $\mathcal{A} = (Q, \Sigma, \Gamma, q_0, X, P, \delta)$, where:

- Q = finite set of states; partitioned into input states Q_{in} and non-input states Q_{non}
- Σ is the input alphabet (taken to be \mathbb{R})
- Γ is a finite output alphabet
- $q_0 \in Q$ is the starting state
- $X = \{x, \text{insample}, \text{insample}', n\}$ is a set of variables. $x, \text{insample}, \text{insample}' \in \mathbb{R}$; $n \in \mathbb{N}$ and is initialized to 0.
- $P : Q \rightarrow \mathbb{Q}^{\geq 0} \times \mathbb{Q} \times \mathbb{Q}^{\geq 0} \times \mathbb{Q}$ describing the parameters for sampling from Laplace distributions at each state.
- $\delta : (Q \times C) \rightarrow Q \times (\Gamma \cup \{\text{insample}, \text{insample}'\} \cup \{\phi\}) \times \{\text{true}, \text{false}\} \times \{0, 1\}$ is the transition function (technically a relation) that defines what state to transition to, what symbol or real value to output, whether or not x is assigned to, and whether or not n is incremented based on the current state and transition guard.

There are certain conditions that δ must satisfy; these are almost all the same as the restrictions on transition functions of DiPA, but with some slight modifications and one major addition (marked in blue):

- **Determinism:** If $\delta(q, \text{true})$ is defined, then no other transitions out of q can be defined. Additionally, at most one of $\delta(q, \text{insample} \geq x)$ and $\delta(q, n < N \text{ AND } \text{insample} \geq x)$ can be defined and at most one of $\delta(q, \text{insample} < x)$ and $\delta(q, n < N \text{ AND } \text{insample} < x)$ can be defined.

If $\delta(q, n < N)$ is defined, then $\delta(q, n < N \text{ AND } \text{insample} \geq x)$ and $\delta(q, n < N \text{ AND } \text{insample} < x)$ are not defined. Additionally, if any of $\delta(q, n < N)$, $\delta(q, n < N \text{ AND } \text{insample} < x)$, or $\delta(q, n < N \text{ AND } \text{insample} \geq x)$ are defined, then $\delta(q, n \geq N)$ must be defined as well. Finally, if $\delta(q, n \geq N)$ is defined, then $\delta(q, \text{true})$, $\delta(q, \text{insample} \geq x)$ and $\delta(q, \text{insample} < x)$ are not defined.

For the sake of convenience, from now on, we will use true to refer to both guards true and $n < N$, $\text{insample} \geq x$ to refer to both $\text{insample} \geq x$ and $n < N \text{ AND } \text{insample} \geq x$, and $\text{insample} < x$ to refer to both $\text{insample} < x$ and $n < N \text{ AND } \text{insample} < x$.

- **Output Distinction:** For any state $q \in Q$, if $\delta(q, \text{insample} \geq x) = (q_1, o_1, b_1, i_1)$ and $\delta(q, \text{insample} < x) = (q_2, o_2, b_2, i_2)$, then $o_1 \neq o_2$ and at least one of $o_1 \in \Gamma$ and $o_2 \in \Gamma$ is true. In addition, $o_1 \neq \phi$ and $o_2 \neq \phi$ and if $\delta(q, n \geq N) = (q', o', b', i')$, then $o' = \phi$, i.e., the ϕ output symbol is reserved for transitions with guard $n \geq N$, which must output ϕ .

- **Initialization:** The initial state q_0 has only one outgoing transition of the form $\delta(q_0, \text{true}) = (q, o, \text{true}, i)$ for $i \in \{0, 1\}$.
- **Non-input transition:** From any $q \in Q_{\text{non}}$, if $\delta(q, c)$ is defined, then $c = \text{true}$.
- **Control Flow Separation:** Consider the underlying graph G of \mathcal{A} . For all states $q \in Q$, if $\delta(q, n \geq N) = (q', o, b, i), q$ and q' must be in different strongly connected components of G .

Note that the **control flow separation** condition implies that no cycle in G can contain an edge that corresponds to a transition with guard $n \geq N$. In addition, determinism combined with control flow separation imply that no two transitions (i.e. transitions with different guards) can be from some state q to the same state q' .

1.1.1 Path Probabilities

Definition 1.2. (from [1]) A **path** ρ of length n of a DiPA* \mathcal{A} is a sequence of states, inputs, and outputs $\rho = q_0 \xrightarrow{a_0, o_0} q_1 \rightarrow \dots \rightarrow q_{n-1}$, where q_i are the states traversed in \mathcal{A} , a_i are the inputs read in each state q_i , and o_i are the outputs output by \mathcal{A} at the transition $q_i \rightarrow q_{i+1}$. We denote the sequence of inputs a_i for a path ρ as $\text{inseq}(\rho)$ and the sequence of outputs o_i as $\text{outseq}(\rho)$. In general, for a path $\rho = q_0 \rightarrow q_1 \rightarrow \dots \rightarrow q_{n-1}$ we denote the transition $q_i \rightarrow q_{i+1}$ by $\text{trans}(q[i])$ and the guard of $\text{trans}(\rho[i])$ as $\text{guard}(\rho[i])$.

Definition 1.3. (from [1]) Two paths $\rho = q_0 \xrightarrow{a_0, o_0} q_1 \rightarrow \dots \rightarrow q_n$ and $\rho' = q'_0 \xrightarrow{a'_0, o'_0} q'_1 \rightarrow \dots \rightarrow q'_n$ of a DiPA* \mathcal{A} are **equivalent** if for all i , $o_i = o'_i$ and $q_i = q'_i$. In other words, ρ and ρ' traverse the same states in \mathcal{A} and produce the same output, and only possibly differ in the inputs they read. (Note that due to output determinism, ρ and ρ' traverse the same states iff they produce the same output.)

For any path ρ of a DiPA* \mathcal{A} , we define $\mathbb{P}[\epsilon, N, x, n, \rho]$ as the **probability** of path ρ being traversed with \mathcal{A} parameters ϵ and N , stored value x , and counter value n . This will enable us to define what it means for a DiPA* to be differentially private.

Consider a path $\rho = q_0 \xrightarrow{a_0, o_0} q_1 \rightarrow \dots \xrightarrow{a_{n-1}, o_{n-1}} q_n$. Here, a_i and o_i are the input to state q_i and output of transition $q_i \rightarrow q_{i+1}$, respectively (if $q_i \in Q_{\text{non}}$ i.e. q does not take in input, $a_i = 0$).

If $|\rho| = 0$, we define $\mathbb{P}[\epsilon, N, x, n, \rho] = 1$. Otherwise, we define $\mathbb{P}[\epsilon, N, x, n, \rho]$ recursively: Let $P(q_0) = (d, \mu, d', \mu')$ be the parameters for sampling from Laplace distributions for **insample** and **insample'** at state q_0 . Let (q_0, c, q_1, o_0, b, i) represent the 0th transition, where c is the guard of the 0th transition, b is whether or not the 0th transition is an assignment transition, and i is the amount that the counter n gets incremented by in the 0th transition.

Let $\nu = \mu + a_0$. If $o_0 = (y, v, w)$ for $y \in \{\text{insample}, \text{insample}'\}$, then let

$$k = \int_v^w \frac{d\epsilon}{2} e^{-d\epsilon|z - \mu - a_0|} dz$$

$$k' = \int_v^w \frac{d'\epsilon}{2} e^{-d'\epsilon|z - \mu' - a_0|} dz$$

If the 0th transition of ρ is not an assignment transition (i.e. $b = \text{false}$), then we define $\mathbb{P}[\epsilon, N, x, n, \rho]$ as follows:

Case 1: $n \geq N$ and $c = n \geq N$. If $o_0 \in \Gamma$, then $\mathbb{P}[\epsilon, N, x, n, \rho] = \mathbb{P}[\epsilon, N, x, n + i, \text{tail}(\rho)]$. If $o_0 = (\text{insample}, v, w)$ then $\mathbb{P}[\epsilon, N, x, n, \rho] = k\mathbb{P}[\epsilon, N, x + i, \text{tail}(\rho)]$. If $o_0 = (\text{insample}, v, w)$ then $\mathbb{P}[\epsilon, N, x, n, \rho] = k'\mathbb{P}[\epsilon, N, x, n + i, \text{tail}(\rho)]$

Case 2: $n < N$ and $c = n \geq N$. Then we define $\mathbb{P}[\epsilon, N, x, n, \rho] = 0$.

Every case for other guards is exactly analogous to their counterpart definitions in [1], but in general where $\mathbb{P}[\epsilon, N, x, n, \text{tail}(\rho)]$ is referenced in [1], $\mathbb{P}[\epsilon, N, x, n + i, \text{tail}(\rho)]$ should be used instead.

Because of the initialization condition, for paths starting at the start state of \mathcal{A} , the starting value of x is irrelevant. In addition, since n is always initialized to 0, we will abuse notation for paths ρ that start at the start state to write $\mathbb{P}[\epsilon, N, \rho]$ to represent $\mathbb{P}[\epsilon, N, x, 0, \rho]$.

We can use this definition of path probabilities to formalize what it means for paths to be valid program traces in \mathcal{A} :

Definition 1.4. A path $\rho = q_0 \rightarrow q_1 \rightarrow \dots q_n$ from the start state q_0 of \mathcal{A} is **valid** if $\mathbb{P}[\epsilon, N, \rho] > 0$.

Most notably, given a definition of path probabilities, we can define what it means for a DiPA* to be differentially private:

Definition 1.5. As in [1], a DiPA* \mathcal{A} with parameters ϵ, N is $d\epsilon$ -**differentially private** if for all equivalent paths ρ, ρ' in \mathcal{A} such that $\text{inseq}(\rho)$ and $\text{inseq}(\rho')$ are adjacent, $\mathbb{P}[\epsilon, N, \rho] \leq e^{d\epsilon}\mathbb{P}[\epsilon, N, \rho']$.

1.1.2 Well-formedness

In this section, we define what it means for a DiPA* to be **well-formed**, analogously to how the well-formedness of DiPAs are defined.

Definition 1.6. A **bounded** cycle C in a DiPA* \mathcal{A} is a cycle in \mathcal{A} where there exists at least one transition $(q', \sigma, t, 1)$ (i.e. n gets incremented) and there exists some $q \in Q$ (“exit state”) in the cycle such that $f(q, n \geq N) = (q', \sigma, t, i)$ where q' is not in the cycle. Otherwise, the cycle is **unbounded**.

Definition 1.7. A cycle C with an exit state with transition $n \geq N$ is an **infeasible** cycle if, for *all* paths $\rho = q_0 \rightarrow q_1 \rightarrow \dots \rightarrow q_m$ from the start state to a state $q_m \in C$, at least N transitions $q_i \rightarrow q_{i+1}$ are increment transitions or some transition $q_i \rightarrow q_{i+1}$ has guard $n \geq N$. Otherwise, C is **feasible**.

Definition 1.8. (from [1]) A **leaking cycle** is a cycle $C = q_0 \xrightarrow{a_0, o_0} q_1 \rightarrow \dots \rightarrow q_{n-1} \rightarrow q_0$ in a DiPA \mathcal{A} if there exist indices $0 \leq i < j < n$ such that the i th transition $q_i \rightarrow q_{i+1}$ is an assignment transition and the guard of the j th transition guard is not $n < N$ or **true**.

Definition 1.9. (from [1]) A cycle ρ of a DiPA* \mathcal{A} is an **L-cycle** (respectively, **G-cycle**) if there is an $i < |\rho|$ such that $\text{guard}(\rho[i]) = \text{insample} < x$ (respectively $\text{guard}(\rho[i]) = \text{insample} \geq$

x).

Definition 1.10. (from [1]) A path ρ of a DiPA \mathcal{A}^* is an **AL-path** (respectively, **AG-path**) if all assignment transitions on ρ have guard `insample` $< \mathbf{x}$ (respectively, `insample` $\geq \mathbf{x}$).

Definition 1.11. (from [1]) A pair of cycles (C, C') in a DiPA \mathcal{A} is a **leaking pair** if one of the following is satisfied:

- C is an L-cycle, C' is a G-cycle, and there is an AG-path from a state in C to a state in C' .
- C is an G-cycle, C' is a L-cycle, and there is an AL-path from a state in C to a state in C' .

Definition 1.12. A pair of cycles (C, C') is a **feasible unbounded leaking pair** of cycles in a DiPA * \mathcal{A} if both C and C' are feasible and unbounded cycles, C is an L-cycle (respectively, G-cycle), C' is a G-cycle (respectively L-cycle), and there exists an AL-path (respectively, AG-path) $\rho = a_1 a_2 \cdots a_k$ from C to C' (i.e. such that $a_1 \in C$ and $a_k \in C'$) such that all of the following hold:

1. Either there are no $n \geq N$ transitions on ρ or C' has no exit state.
2. Either there exists some path τ from the start state q_0 of \mathcal{A} to a_k that includes a_1 such that there are at most $N - 1$ increment transitions on τ or C' has no exit state.
3. Either C' has no exit state or C has no increment transitions.
4. If there exists an $n \geq N$ transition in ρ from states a_i to a_{i+1} , there exists some path τ from the start state q_0 of \mathcal{A} to a_i that includes a_1 such that there are at least N increment transitions in τ .

Conditions (1)-(3) ensure that there exist some path in \mathcal{A} such that either $n < N$ when entering C' or that C' has no exit state; otherwise, C' would be rendered infeasible in practice.

Condition (4) ensures that the path ρ between C and C' is in fact traversible.

Definition 1.13. (from [1]) A cycle C of a DiPA \mathcal{A} is a **disclosing cycle** if there exists some $0 \leq i < |C|$ such that `trans`($C[i]$) is an input transition that outputs either `insample` or `insample'`.

Definition 1.14. (adapted from [1]) An **feasible unbounded privacy violating lasso** is a path $\rho = a_1 \rightarrow a_2 \rightarrow \cdots \rightarrow a_k$ of length n in a DiPA * \mathcal{A} such that one of the following hold:

- `tail`(ρ) is an AG-path (respectively, AL-path) such that `last`(ρ) is in a
- ρ is an AG-path (respectively, AL-path) such that `first`(ρ) is in a feasible unbounded G-cycle (respectively, L-cycle) and the 0th transition has guard `insample` $< \mathbf{x}$ (respectively, `insample` $\geq \mathbf{x}$) and outputs `insample`

- ρ is an AG-path (respectively, AL-path) such that $\text{first}(\rho)$ is in a feasible unbounded L-cycle (respectively, G-cycle) and the last transition has guard $\text{insample} \geq \mathbf{x}$ (respectively, $\text{insample} < \mathbf{x}$) and outputs insample .

In addition, if there are any transitions $a_i \rightarrow a_{i+1}$ in ρ with guard $n \geq N$, there must exist some path represented by the word $\tau = \alpha \cdot \beta$ from the start state of \mathcal{A} to a_i such that α represents a path from the start state of \mathcal{A} to a_1 and β represents a subpath of ρ from a_1 to a_i .

Definition 1.15. For a lasso ρ , let C_ρ be the cycle associated with¹ ρ . Then a lasso ρ in a DiPA* \mathcal{A} is bounded iff C_ρ is bounded. Similarly, ρ is feasible iff C_ρ is feasible.

Definition 1.16. A DiPA* \mathcal{A} is well-formed if \mathcal{A} has no reachable unbounded feasible leaking cycles, unbounded feasible leaking pair (C, C') where C is reachable, reachable unbounded feasible disclosing cycles, or reachable unbounded feasible privacy violating lassos.

1.2 Proving Differential Privacy

Theorem 1.17. A DiPA* is well-formed if and only if it is differentially private.

Proof. We prove Theorem 1.17 by reducing a DiPA* to a DiPA and showing that they are equivalent through a construction similar to the canonical subset/powerset construction for finite automata.

Let $\mathcal{A}^* = (Q, \Sigma, \Gamma, q_0, X^*, P^*, \delta^*)$ be a well-formed DiPA* with parameters ϵ and N .

Let $G = \{\text{true}, n < N, n \geq N, \text{insample} \geq \mathbf{x}, n < N \text{ AND } \text{insample} \geq \mathbf{x}, \text{insample} < \mathbf{x}, n < N \text{ AND } \text{insample} < \mathbf{x}\}$ be the set of guard conditions for DiPA*s.

Construct the DiPA $\mathcal{A} = (Q \times [N], \Sigma, \Gamma \cup \{\phi\}, (q_0, 0), X, P, \delta)$ as follows:

For each state $q \in Q^*$:

For $g \in G$, if $\delta^*(q, g) = (q', \sigma, \mathbf{b}, x)$ is defined, define $\delta((q, k), g)$ as follows:

Case 1: $g \in \{\text{true}, \text{insample} \geq \mathbf{x}, \text{insample} < \mathbf{x}\}$

For all $k \in [N - 1]$, define the transitions

$$\delta((q, k), g) = ((q', k + x), \sigma, \mathbf{b})$$

and define the transition

$$\delta((q, N), g) = ((q', N), \sigma, \mathbf{b})$$

Case 2: $g = n \geq N$

We define the transition

$$\delta((q, N), \text{true}) = ((q', N), \sigma, \mathbf{b})$$

¹Hopefully this is clear

Case 3: $g = n < N$

For all $k \in [N - 1]$, define the transitions

$$\delta((q, k), \mathbf{true}) = ((q', k + x), \sigma, \mathbf{b})$$

Case 4: $g = n < N$ **AND** $\mathbf{insample} < \mathbf{x}$

For all $k \in [N - 1]$, define the transitions

$$\delta((q, k), \mathbf{insample} < \mathbf{x}) = ((q', k + x), \sigma, \mathbf{b})$$

Case 5: $g = n < N$ **AND** $\mathbf{insample} \geq \mathbf{x}$

For all $k \in [N - 1]$, define the transitions

$$\delta((q, k), \mathbf{insample} \geq \mathbf{x}) = ((q', k + x), \sigma, \mathbf{b})$$

Intuitively, at state (q, k) in \mathcal{A} , k will track the value of n in \mathcal{A}^* (since everything above N is treated the same, we compress all of those values together).

For each state $(q, k) \in Q$, let $P((q, k)) = P^*(q)$.

Claim 1.18. \mathcal{A} is a valid DiPA.

Proof. To be a valid DiPA, δ must satisfy four conditions: determinism, output distinction, initialization, and non-input transition.

Determinism: Consider some state $(q, k) \in Q$ and suppose that $\delta((q, k), \mathbf{true})$ is defined. Then either $\delta^*(q, \mathbf{true})$ is defined, $\delta^*(q, n < N)$ and $k < N$, or $\delta^*(q, n \geq N)$ is defined and $k = N$. By the condition of determinism for DiPA*s, if $\delta^*(q, \mathbf{true})$, then no other transitions from q in \mathcal{A}^* are defined. Thus, neither $\delta((q, k), \mathbf{insample} < \mathbf{x})$ or $\delta((q, k), \mathbf{insample} \geq \mathbf{x})$ can be defined in \mathcal{A} .

If $\delta^*(q, n < N)$ is defined and $k < N$, then similarly none of $\delta^*(q, n < N \text{ AND } \mathbf{insample} \geq \mathbf{x})$, $\delta^*(q, n < N \text{ AND } \mathbf{insample} < \mathbf{x})$, $\delta^*(q, \mathbf{true})$, $\delta^*(q, \mathbf{insample} < \mathbf{x})$, or $\delta^*(q, \mathbf{insample} \geq \mathbf{x})$ can be defined. Additionally, since $n < N$, there is no additional transition from q corresponding to an $n \geq N$ guard in \mathcal{A}^* , so neither $\delta((q, k), \mathbf{insample} < \mathbf{x})$ or $\delta((q, k), \mathbf{insample} \geq \mathbf{x})$ can be defined in \mathcal{A} .

Similarly, if $\delta^*(q, n \geq N)$ is defined and $k = N$, no other transitions from q can be defined in \mathcal{A} .

Output distinction: This follows immediately from the output distinction condition of DiPA*s.

Initialization: By the initialization condition of DiPA*s, the initial state q_0 has only one outgoing transition of the form $\delta(q_0, \mathbf{true}) = (q, o, \mathbf{true}, i)$ for $i \in \{0, 1\}$. Thus, there is only one transition out of $(q_0, 0)$ in \mathcal{A} , with guard \mathbf{true} .

Non-input transition: Since input transitions are preserved from \mathcal{A}^* in the construction of \mathcal{A} , this follows immediately from the condition of non-input transition for DiPA*s. \square

Lemma 1.19. *If there exists a reachable cycle $C = (a_0, k_0) \rightarrow (a_1, k_1) \rightarrow \dots \rightarrow (a_{m-1}, k_{m-1}) \rightarrow (a_0, k_0)$ in \mathcal{A} if and only if there exists a reachable unbounded feasible cycle $C^* = a_0 \rightarrow a_1 \rightarrow \dots \rightarrow a_{m-1} \rightarrow a_0$ in \mathcal{A}^* .*

Proof. Let $C = (a_0, k_0)(a_1, k_1) \dots (a_{m-1}, k_{m-1})(a_0, k_0)$ be a cycle in \mathcal{A} . Note that by construction, there must exist a cycle $C^* = a_0 a_1 \dots a_{m-1} a_0$ in \mathcal{A}^* .

Additionally note that for any path $(q, k) \rightarrow (q', k')$ in \mathcal{A} , $k' \geq k$. This implies that $k_0 = k_1 = \dots = k_{m-1}$.

In order for C^* to be bounded, there must be an increment transition between some states a_i and a_{i+1} in C^* . However, this would mean that there exists a transition $(a_i, k_i) \rightarrow (a_{i+1}, k_i + 1)$, which is impossible because all k_i 's are equal for $0 \leq i \leq m - 1$ and because \mathcal{A} is deterministic. Therefore C^* is unbounded.

Consider the underlying graphs $G_{\mathcal{A}}, G_{\mathcal{A}^*}$ of \mathcal{A} and \mathcal{A}^* , respectively. Consider some edge $e^* = (q, q') \in G_{\mathcal{A}^*}$; there exists exactly one corresponding edge $e = ((q, k), (q', k'))$ in $G_{\mathcal{A}}$.

Suppose that C^* is infeasible for the sake of contradiction. Then $k_0 = k_1 = \dots k_{m-1} = n$ and there is some exit state $s \in C^*$ such that $\delta^*(s, n \geq N) = (s', \sigma, \mathbf{b}, x)$. Note that because of determinism, there is exactly one transition $(s, n) \rightarrow (s', n)$ in G out of the state (s, n) . Therefore, (s', n) must also be in the cycle C , which implies s' must be in the cycle C^* . However, this contradicts control flow separation, since then s and s' would be in the same component of $G_{\mathcal{A}^*}$, even with the edge corresponding to the $n \geq N$ transition removed. Thus, C^* is feasible.

Now let $C^* = a_0 \rightarrow \dots \rightarrow a_m \rightarrow a_0$ be an reachable unbounded feasible cycle in \mathcal{A}^* . Because C^* is feasible, there exists some path from the start state q_0 of \mathcal{A}^* to a_0 such that $n < N$ at a_0 . Let k_0 be the minimum such value of n at a_0 . Then consider the sequence of states $(a_0, k_0), (a_1, k_1), \dots, (a_m, n_m), (a_0, k_0)$ in \mathcal{A} such that k_i is defined as follows:

Let $(a_{i-1}, c, a_i, o, b, j)$ represent the transition from a_{i-1} to a_i in \mathcal{A}^* . Then for $i > 0$, $k_i = k_{i-1} + j$. Note that for all $1 \leq i \leq m$, there exists a transition $(a_{i-1}, k_{i-1}) \rightarrow (a_i, k_i)$ by construction.

Suppose that there is no increment transition in C^* . Then $\forall 1 \leq i \leq m, k_i = k_{i-1}$. Thus, $k_m = k_0$ and there exists a transition $(a_m, k_0) \rightarrow (a_0, k_0)$ in \mathcal{A} . Thus, the cycle $C = (a_0, k_0) \rightarrow (a_1, k_1) \rightarrow \dots \rightarrow (a_m, k_m) \rightarrow (a_0, k_0)$ is a cycle in \mathcal{A} .

Now suppose that C^* has an increment transition but no exit state. Then the state (a_0, N) is reachable in \mathcal{A} . To see this, consider the path $(a_0, k_0) \rightarrow (a_1, k_1) \rightarrow \dots \rightarrow (a_m, k_m) \rightarrow (a_0, k'_0)$ in \mathcal{A} . Since C^* has an increment transition, $k'_0 > k_0$. Since C^* has no exit state and k_i is bounded above by N , this means that there exists a path from $(q_0, 0)$ to (a_0, k_0) to (a_0, N) . Then again because C^* has no exit state, $C = (a_0, N) \rightarrow (a_1, N) \rightarrow \dots \rightarrow (a_m, N) \rightarrow (a_0, N)$ is a cycle in \mathcal{A} . \square

Lemma 1.20. *If \mathcal{A}^* is well-formed if and only if \mathcal{A} is well-formed.*

Proof. By lemma 1.19, there exists a reachable cycle $C = (q_1, n_1) \rightarrow \dots \rightarrow (q_m, n_m) \rightarrow (q_1, n_1)$ in \mathcal{A} iff $C^* = q_1 \rightarrow \dots \rightarrow q_m \rightarrow q_1$ is a reachable unbounded feasible cycle in \mathcal{A}^* . Further, such a C is a leaking cycle iff there exist indices $0 \leq i < j < n$ such that the transition $(q_i, n_i) \rightarrow (q_{i+1}, n_{i+1})$ is an assignment transition and $\text{guard}((q_j, n_j))$ is not **true**. By construction, this happens iff the transition $q_i \rightarrow q_{i+1}$ in \mathcal{A}^* is also an assignment transition and $\text{guard}(q_j)$ in \mathcal{A}^* is not **true** or $n < N$, so C^* is a leaking cycle. Thus, there exists a reachable leaking cycle in \mathcal{A} iff there exists a reachable unbounded feasible leaking cycle in \mathcal{A}^* .

Suppose that $C = (q_1, n_1) \rightarrow \dots \rightarrow (q_m, n_m) \rightarrow (q_1, n_1)$ and $C' = (q'_1, n'_1) \rightarrow \dots \rightarrow (q'_m, n'_m) \rightarrow (q'_1, n'_1)$ are cycles in \mathcal{A} such that C is reachable and (C, C') is a leaking pair in \mathcal{A} . By lemma 1.19, $C^* = q_1 \rightarrow \dots \rightarrow q_m \rightarrow q_1$ and $C'^* = q'_1 \rightarrow \dots \rightarrow q'_m \rightarrow q'_1$ are unbounded and feasible cycles in \mathcal{A}^* and C^* is reachable. Without loss of generality, suppose that C is an L-cycle and C' is a G-cycle. Then by construction, C^* is an L-cycle and C'^* is a G-cycle. In addition, there exists some AL-path $\rho = (a_1, k_1) \rightarrow \dots \rightarrow (a_l, k_l)$ in \mathcal{A} from C to C' . By construction, there must exist some AL-path $\rho^* = a_1 \rightarrow \dots \rightarrow a_l$ in \mathcal{A}^* from C^* to C'^* .

Note that if there exists a transition with guard $n \geq N$ in ρ^* , then $k_l = n$.

Observe that if $(q, k) \rightarrow (q', k')$ is an assignment transition in \mathcal{A} , $q \rightarrow q'$ is also an assignment transition in \mathcal{A}^* . Similarly, if a transition $(q, k) \rightarrow (q', k')$ has a guard of $\text{insample} \geq x$ (respectively, $\text{insample} < x$) in \mathcal{A} , $q \rightarrow q'$ also has a guard of $\text{insample} \geq x$ (respectively, $\text{insample} < x$) in \mathcal{A}^* . Together, these mean that leaking cycles, leaking pairs, disclosing cycles, and privacy violating lassos in \mathcal{A} correspond to their feasible unbounded equivalents in \mathcal{A}^* . □

Lemma 1.21. *Let $\Psi = \{\rho : \rho \text{ is a path in } \mathcal{A}\}$ and $\Psi^* = \{\rho^* : \rho^* \text{ is a valid path in } \mathcal{A}^*\}$ be the sets of paths in \mathcal{A} and \mathcal{A}^* , respectively. There exists a bijection $f : \Psi \rightarrow \Psi^* \times [N]$ such that $\forall x, \forall \rho \in \Psi$, if $f(\rho) = (\rho^*, n)$, $\mathbb{P}[\epsilon, x, \rho] = \mathbb{P}[\epsilon, N, x, n, \rho^*]$.*

Proof. Let $\rho = (q_1, n_1) \rightarrow (q_2, n_2) \rightarrow \dots \rightarrow (q_m, n_m)$ be a path in \mathcal{A} . Then let $f(\rho) = (q_1 \rightarrow q_2 \dots \rightarrow q_m, n_1)$ such that $\text{inseq}(q_1 \rightarrow q_2 \dots \rightarrow q_m) = \text{inseq}(\rho)$. Note that $\text{outseq}(q_1 \rightarrow q_2 \dots \rightarrow q_m) = \text{outseq}(\rho)$ by output determinism.

By construction², $\rho^* = q_1 \rightarrow \dots \rightarrow q_m$ must be a valid path in \mathcal{A}^* if the value of the variable n in \mathcal{A}^* is n_1 at q_1 .

f is injective: Let $\rho = (q_1, n_1) \rightarrow \dots \rightarrow (q_m, n_m)$, $\rho' = (q'_1, n'_1) \rightarrow \dots \rightarrow (q'_m, n'_m)$ be two paths in \mathcal{A} such that $\rho \neq \rho'$. If $|\rho| \neq |\rho'|$, $f(\rho) \neq f(\rho')$ clearly. Suppose $|\rho| = |\rho'|$ and consider the smallest i such that either $n_i \neq n'_i$ or $q_i \neq q'_i$. If $q_i \neq q'_i$, then clearly $f(\rho) = (q_1 \rightarrow \dots \rightarrow q_i \rightarrow \dots \rightarrow q_m, n_1) \neq (q_1 \rightarrow \dots \rightarrow q'_i \rightarrow \dots \rightarrow q_m, n'_1) = f(\rho')$. Otherwise, if $q_i = q'_i$ and $n_i \neq n'_i$, note that $i = 1$: there can only be one transition from $q_{i-1} \rightarrow q_i$ in \mathcal{A}^* and because i is the smallest such i , $q_{i-1} = q'_{i-1}$. Thus if $i > 1$, this would mean that $n_i = n'_i$, which is impossible. So $f(\rho) = (q_1 \rightarrow \dots \rightarrow q_m, n_1) \neq (q'_1 \rightarrow \dots \rightarrow q'_m, n'_1) = f(\rho')$.

²Does this need to be elaborated on?

f is surjective: Let $(\rho^*, n_1) = (q_1 \rightarrow \dots \rightarrow q_m, n_1) \in \Psi^* \times [N]$. Let n_i be the value of n in \mathcal{A}^* after starting at state q_1 with $n = n_1$ and traversing each state q_i in order. Then $\rho = (q_1, n_1) \rightarrow \dots (q_m, n_m)$ is a path in \mathcal{A} by construction and clearly $f(\rho) = (\rho^*, n_1)$.

Fix $x \in \mathbb{R}$ and $\rho \in \Psi$. Let $f(\rho) = (\rho^*, n_1)$. We will show that $\mathbb{P}[\epsilon, x, \rho] = \mathbb{P}[\epsilon, N, x, n_1, \rho^*]$.

This follows by induction on $|\rho^*|$:

Suppose $|\rho^*| = 0$. Then $\mathbb{P}[\epsilon, x_0, \rho] = \mathbb{P}[\epsilon, N, x_0, n_0, \rho^*] = 1$.

Now suppose $|\rho^*| = k > 0$ and that for all $|\rho'^*| < k$, $\mathbb{P}[\epsilon, x_0, \rho'] = \mathbb{P}[\epsilon, N, x_0, n_0, \rho'^*]$.

Let c_0 be the guard of the first transition $q_0 \rightarrow q_1$ in $\rho^* = q_0 q_1 \dots q_{m-1}$. So $\delta^*(q_0, c_0) = (q_1, \sigma, b, i)$.

Let $\nu = \mu + a_0$, where a_0 is the first input value read (or 0 if $q_0 \in Q_{non}$). Let $P(q_0) = (d, \mu, d', \mu')$ be the parameters for sampling from the Laplace distribution at q_0 . Define ℓ and u as follows: if $o_0 \in \Gamma \cup \{\phi\}$, then $\ell = -\infty$ and $u = \infty$. Otherwise, if $o_0 = (y, v, w)$ for $y \in \{\text{insample}, \text{insample}'\}$ (i.e. either **insample** or **insample'** is output with a value between v and w), then $\ell = v$ and $u = w$.

If $o_0 = (y, v, w)$ for $y \in \{\text{insample}, \text{insample}'\}$, then let

$$k = \int_v^w \frac{d\epsilon}{2} e^{-d\epsilon|z-\mu-a_0|} dz$$

$$k' = \int_v^w \frac{d'\epsilon}{2} e^{-d'\epsilon|z-\mu'-a_0|} dz$$

Case 1: $c = n < N$

Note that $n_0 < N$ since ρ^* is a valid path.

By construction, $\delta((q_0, n_0), \text{true}) = ((q_1, n_1), \sigma, b)$ where $n_1 = n_0 + i$. Let x' be the value of x at q_1 in \mathcal{A}^* . Since \mathcal{A}^* assigns to x iff \mathcal{A} does, x' is also the value of x at (q_1, n_1) in \mathcal{A} .

Since $n_0 < N$, by the induction hypothesis

$$\mathbb{P}[\epsilon, N, x, n_0, \rho^*] = \mathbb{P}[\epsilon, N, x', n_1, \text{tail}(\rho^*)] = \mathbb{P}[\epsilon, x', \text{tail}(\rho)] = \mathbb{P}[\epsilon, x, \rho]$$

Case 2: $c = \text{true}$

As in case 1, by construction, $\delta((q_0, n_0), \text{true}) = ((q_1, n_1), \sigma, b)$ where $n_1 = n_0 + i$. Let x' be the value of x at q_1 in \mathcal{A}^* . Since \mathcal{A}^* assigns to x iff \mathcal{A} does, x' is also the value of x at (q_1, n_1) in \mathcal{A} .

Then by the induction hypothesis

$$\mathbb{P}[\epsilon, N, x, n_0, \rho^*] = \mathbb{P}[\epsilon, N, x', n_1, \text{tail}(\rho^*)] = \mathbb{P}[\epsilon, x', \text{tail}(\rho)] = \mathbb{P}[\epsilon, x, \rho]$$

Case 3: $c = n \geq N$

Note that $n_0 \geq N$ since ρ^* is a valid path.

By construction, $\delta((q_0, n_0), \mathbf{true}) = ((q_1, n_1), \sigma, b)$ where $n_1 = n_0$. Let x' be the value of x at q_1 in \mathcal{A}^* . Since \mathcal{A}^* assigns to x iff \mathcal{A} does, x' is also the value of x at (q_1, n_1) in \mathcal{A} .

Since $n_0 \geq N$, by the induction hypothesis

$$\mathbb{P}[\epsilon, N, x, n_0, \rho^*] = \mathbb{P}[\epsilon, N, x', n_1, \mathbf{tail}(\rho^*)] = \mathbb{P}[\epsilon, x', \mathbf{tail}(\rho)] = \mathbb{P}[\epsilon, x, \rho]$$

Case 4: $c = n < N$ AND $\mathbf{insample} \geq x$

Since ρ^* is valid, $n_0 < N$.

By construction, $\delta((q_0, n_0), \mathbf{insample} \geq x) = ((q_1, n_1), \sigma, b)$ where $n_1 = n_0 + i$.

Suppose $b = \mathbf{true}$ (i.e. $\mathbf{trans}(q_0)$ is an assignment transition), then:

If σ is of the form $(\mathbf{insample}', v, w)$, since $n_0 < N$,

$$\begin{aligned} \mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= k' \left(\int_x^\infty \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, N, z, n_1, \mathbf{tail}(\rho^*)] dz \\ &= k' \left(\int_x^\infty \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, z, \mathbf{tail}(\rho)] dz \text{ by the induction hypothesis} \\ &= \mathbb{P}[\epsilon, x, \rho] \end{aligned}$$

Otherwise,

$$\begin{aligned} \mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= \left(\int_{\max(x, \ell)}^u \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, N, z, n_1, \mathbf{tail}(\rho^*)] dz \\ &= \left(\int_{\max(x, \ell)}^u \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, z, \mathbf{tail}(\rho)] dz \text{ by the induction hypothesis} \\ &= \mathbb{P}[\epsilon, x, \rho] \end{aligned}$$

Now suppose that $b = \mathbf{false}$. If σ is of the form $(\mathbf{insample}', v, w)$, since $n_0 < N$,

$$\begin{aligned} \mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= k' \left(\int_x^\infty \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, N, x, n_1, \mathbf{tail}(\rho^*)] \\ &= k' \left(\int_x^\infty \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, x, \mathbf{tail}(\rho)] \text{ by the induction hypothesis} \\ &= \mathbb{P}[\epsilon, x, \rho] \end{aligned}$$

Otherwise,

$$\begin{aligned} \mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= \left(\int_{\max(x, \ell)}^u \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, N, x, n_1, \mathbf{tail}(\rho^*)] dz \\ &= \left(\int_{\max(x, \ell)}^u \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, x, \mathbf{tail}(\rho)] \text{ by the induction hypothesis} \\ &= \mathbb{P}[\epsilon, x, \rho] \end{aligned}$$

Case 5: $c = n < N$ **AND** $\text{insample} < \mathbf{x}$

Since ρ^* is valid, $n_0 < N$.

By construction, $\delta((q_0, n_0), \text{insample} < \mathbf{x}) = ((q_1, n_1), \sigma, b)$ where $n_1 = n_0 + i$.

Suppose $b = \text{true}$. Then if σ is of the form $(\text{insample}', v, w)$, since $n_0 < N$,

$$\begin{aligned} \mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= k' \left(\int_{-\infty}^x \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, N, z, n_1, \text{tail}(\rho^*)] dz \\ &= k' \left(\int_{-\infty}^x \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, z, \text{tail}(\rho)] dz \text{ by the induction hypothesis} \\ &= \mathbb{P}[\epsilon, x, \rho] \end{aligned}$$

Otherwise,

$$\begin{aligned} \mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= \left(\int_{\ell}^{\min(u, x)} \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, N, z, n_1, \text{tail}(\rho^*)] dz \\ &= \left(\int_{\ell}^{\min(u, x)} \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, z, \text{tail}(\rho)] dz \text{ by the induction hypothesis} \\ &= \mathbb{P}[\epsilon, x, \rho] \end{aligned}$$

Now suppose that $b = \text{false}$. If σ is of the form $(\text{insample}', v, w)$, since $n_0 < N$,

$$\begin{aligned} \mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= k' \left(\int_x^{\infty} \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, N, x, n_1, \text{tail}(\rho^*)] \\ &= k' \left(\int_x^{\infty} \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, x, \text{tail}(\rho)] \text{ by the induction hypothesis} \\ &= \mathbb{P}[\epsilon, x, \rho] \end{aligned}$$

Otherwise,

$$\begin{aligned} \mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= \left(\int_{\ell}^{\min(u, x)} \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, N, x, n_1, \text{tail}(\rho^*)] dz \\ &= \left(\int_{\ell}^{\min(u, x)} \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, x, \text{tail}(\rho)] \text{ by the induction hypothesis} \\ &= \mathbb{P}[\epsilon, x, \rho] \end{aligned}$$

Case 6: $c = \text{insample} \geq \mathbf{x}$

By construction, $\delta((q_0, n_0), \text{insample} \geq \mathbf{x}) = ((q_1, n_1), \sigma, b)$ where $n_1 = n_0 + i$.

Suppose $b = \text{true}$ (i.e. $\text{trans}(q_0)$ is an assignment transition), then:

If σ is of the form $(\text{insample}', v, w)$,

$$\begin{aligned}\mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= k' \left(\int_x^\infty \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, N, z, n_1, \text{tail}(\rho^*)] dz \\ &= k' \left(\int_x^\infty \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, z, \text{tail}(\rho)] dz \text{ by the induction hypothesis} \\ &= \mathbb{P}[\epsilon, x, \rho]\end{aligned}$$

Otherwise,

$$\begin{aligned}\mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= \left(\int_{\max(x, \ell)}^u \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, N, z, n_1, \text{tail}(\rho^*)] dz \\ &= \left(\int_{\max(x, \ell)}^u \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, z, \text{tail}(\rho)] dz \text{ by the induction hypothesis} \\ &= \mathbb{P}[\epsilon, x, \rho]\end{aligned}$$

Now suppose that $b = \text{false}$. If σ is of the form $(\text{insample}', v, w)$,

$$\begin{aligned}\mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= k' \left(\int_x^\infty \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, N, x, n_1, \text{tail}(\rho^*)] \\ &= k' \left(\int_x^\infty \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, x, \text{tail}(\rho)] \text{ by the induction hypothesis} \\ &= \mathbb{P}[\epsilon, x, \rho]\end{aligned}$$

Otherwise,

$$\begin{aligned}\mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= \left(\int_{\max(x, \ell)}^u \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, N, x, n_1, \text{tail}(\rho^*)] dz \\ &= \left(\int_{\max(x, \ell)}^u \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, x, \text{tail}(\rho)] \text{ by the induction hypothesis} \\ &= \mathbb{P}[\epsilon, x, \rho]\end{aligned}$$

Case 7: $c = \text{insample} < \mathbf{x}$

By construction, $\delta((q_0, n_0), \text{insample} < \mathbf{x}) = ((q_1, n_1), \sigma, b)$ where $n_1 = n_0 + i$.

Suppose $b = \text{true}$. Then if σ is of the form $(\text{insample}', v, w)$,

$$\begin{aligned}\mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= k' \left(\int_{-\infty}^x \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, N, z, n_1, \text{tail}(\rho^*)] dz \\ &= k' \left(\int_{-\infty}^x \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} \right) \mathbb{P}[\epsilon, z, \text{tail}(\rho)] dz \text{ by the induction hypothesis} \\ &= \mathbb{P}[\epsilon, x, \rho]\end{aligned}$$

Otherwise,

$$\begin{aligned}
\mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= \left(\int_{\ell}^{\min(u, x)} \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, N, z, n_1, \mathbf{tail}(\rho^*)] dz \\
&= \left(\int_{\ell}^{\min(u, x)} \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, z, \mathbf{tail}(\rho)] dz \text{ by the induction hypothesis} \\
&= \mathbb{P}[\epsilon, x, \rho]
\end{aligned}$$

Now suppose that $b = \mathbf{false}$. If σ is of the form $(\mathbf{insample}', v, w)$,

$$\begin{aligned}
\mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= k' \left(\int_x^{\infty} \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, N, x, n_1, \mathbf{tail}(\rho^*)] \\
&= k' \left(\int_x^{\infty} \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, x, \mathbf{tail}(\rho)] \text{ by the induction hypothesis} \\
&= \mathbb{P}[\epsilon, x, \rho]
\end{aligned}$$

Otherwise,

$$\begin{aligned}
\mathbb{P}[\epsilon, N, x, n_0, \rho^*] &= \left(\int_{\ell}^{\min(u, x)} \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, N, x, n_1, \mathbf{tail}(\rho^*)] dz \\
&= \left(\int_{\ell}^{\min(u, x)} \frac{d\epsilon}{2} e^{-d\epsilon|z-\nu|} dz \right) \mathbb{P}[\epsilon, x, \mathbf{tail}(\rho)] \text{ by the induction hypothesis} \\
&= \mathbb{P}[\epsilon, x, \rho]
\end{aligned}$$

This is sufficient to prove the lemma. □

Lemma 1.22. *There exists $d > 0$ such that \mathcal{A} is $d\epsilon$ -differentially private if and only if \mathcal{A}^* is $d\epsilon$ -differentially private.*

Proof. Let f be a bijection from paths in \mathcal{A} to tuples of paths in \mathcal{A}^* to $[N]$, as defined in Lemma 1.21.

Suppose that $\exists d > 0$ such that \mathcal{A} is $d\epsilon$ -differentially private. Then for all equivalent paths ρ, ρ' in \mathcal{A} from the start state $(q_0, 0)$ such that $\mathbf{inseq}(\rho)$ and $\mathbf{inseq}(\rho')$ are adjacent, $\mathbb{P}[\epsilon, \rho] \leq e^{d\epsilon} \mathbb{P}[\epsilon, \rho']$. Consider two equivalent paths ρ^*, ρ'^* in \mathcal{A}^* from the start state q_0 such that $\mathbf{inseq}(\rho^*)$ and $\mathbf{inseq}(\rho'^*)$ are adjacent. Then $\forall x \in \mathbb{R}, \mathbb{P}[\epsilon, N, x, 0, \rho^*] = \mathbb{P}[\epsilon, x, f^{-1}((\rho^*, 0))] \leq e^{d\epsilon} \mathbb{P}[\epsilon, x, f^{-1}((\rho'^*, 0))] = e^{d\epsilon} \mathbb{P}[\epsilon, N, x, 0, \rho'^*]$. Thus, \mathcal{A}^* is $d\epsilon$ -differentially private.

Suppose that $\forall d > 0, \mathcal{A}$ is not $d\epsilon$ -differentially private. So there exists two equivalent paths from the start state $(q_0, 0)$ in \mathcal{A} $\rho = (q_0, 0) \rightarrow \dots \rightarrow (q_m, n_m), \rho' = (q_0, 0) \rightarrow (q'_m, n'_m)$ in \mathcal{A} such that $\mathbf{inseq}(\rho)$ and $\mathbf{inseq}(\rho')$ are adjacent, but $\mathbb{P}[\epsilon, \rho] > e^{d\epsilon} \mathbb{P}[\epsilon, \rho']$.

Let $f(\rho) = (\rho^*, 0)$ and $f(\rho') = (\rho'^*, 0)$. Fix $\text{inseq}(\rho^*) = \text{inseq}(\rho)$ and $\text{inseq}(\rho'^*) = \text{inseq}(\rho')$.

Then by Lemma 1.21, $\mathbb{P}[\epsilon, N, \rho] = \mathbb{P}[\epsilon, N, x, 0, \rho^*] = \mathbb{P}[\epsilon, x, \rho] = \mathbb{P}[\epsilon, \rho] > e^{d\epsilon} \mathbb{P}[\epsilon, \rho'] = e^{d\epsilon} \mathbb{P}[\epsilon, x, \rho'] = e^{d\epsilon} \mathbb{P}[\epsilon, N, x, 0, \rho'^*] = e^{d\epsilon} \mathbb{P}[\epsilon, N, \rho'^*]$. Thus, \mathcal{A}^* is not $d\epsilon$ -DP. \square

Lemmas 1.20 and 1.22 together prove the theorem. \square

Corollary 1.23. *Let \mathcal{A}^* be a DiPA* with unfixed parameters.*

Let $f(\epsilon, N) : \mathbb{R} \times \mathbb{N} \rightarrow \mathbb{R}$ be defined as follows:

Consider the instantiated version of \mathcal{A}^ with parameters ϵ and N . Let \mathcal{A} be the DiPA constructed from \mathcal{A}^* as in Theorem 3.1. $f(\epsilon, N) = \text{wt}(\mathcal{A})$.*

Then $\forall \epsilon, f(\epsilon, N)$ grows linearly in N .

Corollary 1.24. *For a DiPA* \mathcal{A}^* , the well-formedness of \mathcal{A}^* can be decided efficiently.*

Proof. Note that the time it takes is bounded by the cost of creating a DiPA \mathcal{A} from \mathcal{A}^* as in Theorem 3.1. The construction of \mathcal{A} from \mathcal{A}^* causes the number of states to increase by a factor of N . Each transition in \mathcal{A}^* corresponds to at most N transitions in \mathcal{A} . Since the well-formedness of \mathcal{A} and \mathcal{A}^* are equivalent, at most there is a linear increase in the time required to check the well-formedness of \mathcal{A}^* as compared to a DiPA* of the same size. \square

2 Input Terminator: EDIT I HAVE REALIZED THAT THIS IS BASICALLY MEANINGLESS

2.1 Definitions

Definition 2.1. Fix a parameter ϵ . Let C be the guard conditions $\{\text{input} = \tau, \text{input} \neq \tau, \text{input} \neq \tau \text{ AND } \text{insample} \geq x, \text{input} \neq \tau \text{ AND } \text{insample} < x\}$. An **Input Terminated Differentially Private Automaton** (ITDiPA) \mathcal{A} is defined as the tuple $\mathcal{A} = (Q, \Sigma, \Gamma, q_0, X, P, \delta)$, where:

- Q = finite set of states; partitioned into input states Q_{in} and non-input states Q_{non}
- $\Sigma \cup \{\tau\}$ is the input alphabet (Σ is taken to be \mathbb{R})
- Γ is a finite output alphabet
- $q_0 \in Q$ is the starting state
- $X = \{x, \text{insample}, \text{insample}'\}$ is a set of variables. $x, \text{insample}, \text{insample}' \in \mathbb{R}$.
- $P : Q \rightarrow \mathbb{Q}^{\geq 0} \times \mathbb{Q} \times \mathbb{Q}^{\geq 0} \times \mathbb{Q}$ describing the parameters for sampling from Laplace distributions at each state.
- $\delta : (Q \times C) \rightarrow Q \times (\Gamma \cup \{\text{insample}, \text{insample}'\}) \times \{\text{true}, \text{false}\}$ is the transition function (technically a relation) that defines what state to transition to, what symbol or real value to output, and whether or not x is assigned to.

There are certain conditions that δ must satisfy; these are almost all the same as the restrictions on transition functions of DiPA, but with some slight modifications (marked in blue):

- **Determinism:** If, for a state $q \in Q$, a transition from q with guard $\text{input} \neq \tau$ is defined, then there are no transitions from q with guard either $\text{input} \neq \tau \text{ AND } \text{insample} < x$ nor $\text{input} \neq \tau \text{ AND } \text{insample} \geq x$.

Note that if the automaton is in state q and none of the guards of outgoing transitions from q are satisfied, the automaton terminates.

- **Output Distinction:** For any state $q \in Q$, if $\delta(q, \text{insample} \geq x) = (q_1, o_1, b_1)$ and $\delta(q, \text{insample} < x) = (q_2, o_2, b_2)$, then $o_1 \neq o_2$ and at least one of $o_1 \in \Gamma$ and $o_2 \in \Gamma$ is true.
- **Initialization:** The initial state q_0 has only one outgoing transition of the form $\delta(q_0, \text{input} \neq \tau) = (q, o, \text{true})$.
- **Non-input transition:** From any $q \in Q_{non}$, if $\delta(q, c)$ is defined, then $c \in \{\text{input} \neq \tau, \text{input} = \tau\}$.
- **Output Termination:** Let G be the underlying graph of \mathcal{A} . If $\delta(q, \text{true} = \tau) = (q', \sigma, b)$ is defined, let C and C' be the weakly³ connected components G that contain

³check

q and q' , respectively. Then C and C' must be different components of G and C' must be acyclic. Further, for all states $q' \in C'$, $q' \in Q_{non}$ (i.e. no further states take in any input).

2.1.1 Path probabilities

Definition 2.2. Two input sequences $\rho, \sigma \in \mathbb{R}^* \times \{\tau\}$ are **adjacent** if $|\rho| = |\sigma|$, $\rho[|\rho|] = \sigma[|\rho|] = \tau$ and $\forall i < |\rho|, |\rho[i] - \sigma[i]| \leq 1$.

Note that path probabilities are exactly the same as in DiPAs, with the exception that a ITDiPA takes a `input` = τ transition with probability 1 if the input is τ and 0 otherwise.

Definition 2.3. A ITDiPA \mathcal{A} is $d\epsilon$ -**differentially private** if for all adjacent $\rho, \sigma \in \mathbb{R}^* \times \{\tau\}$, $\mathbb{P}[d\epsilon, \rho] \leq e^{d\epsilon} \mathbb{P}[\epsilon, \sigma]$.

2.1.2 Wellformedness

For a ITDiPA \mathcal{A} , let $DiPA(\mathcal{A})$ be the DiPA created from \mathcal{A} by removing all transitions with guard `input` = τ from the automaton and then removing all newly unreachable components.

Then \mathcal{A} is well-formed if and only if $DiPA(\mathcal{A})$.

2.2 Differential Privacy

Theorem 2.4. *A ITDiPA \mathcal{A} is well-formed if and only if it is differentially private.*

Proof. Consider two paths □

References

- [1] Rohit Chadha, A. Prasad Sistla, and Mahesh Viswanathan. On Linear Time Decidability of Differential Privacy for Programs with Unbounded Inputs, April 2021.