# Contents



# 1 Some notes from which to expand this document

1. An overview of coupling proofs of privacy
2. A tight shift-coupling proof of privacy (segment free)
   (a) What are the connecting constraints, and why are they there?
   (b) Proofs have to depend on $\Delta$!
   (c) Proofs have to depend on sequences of segments!
3. Simplifying the problem above in various ways:
   (a) Separability and the introduction of segments
   (b) Only inter-segment transitions matter!
      - Given $\Delta$, the $\gamma$ values on inter-segment transitions are easily determined.
      - The $\Delta$ values on the inter-segment transitions can be determined.
4. Solving the problem.
   (a) Hardness (incomplete)!!!
   (b) Solving the easier version, where proofs don't depend on $\Delta$.

(c) Showing that they are bounded within $n$ of each other.

$$\exists \text{ finite DP bound} \iff \text{hard system admits a feasible solution}$$

# 2 Definitions

## 2.1 DiPA

Insert definition for DiPA here.

## 2.2 Probability

**Definition 2.1.** *The probability* $\Pr(\rho|X)$ *of a path* $\rho = q_0 \to q_1 \to \cdots \to q_m$ *given an input* $X = \langle a_1, \ldots, a_m \rangle$ *is defined recursively as the probability that all transitions in* $\rho$ *are traversed in sequence given the input* $X$ *starting at state* $q_0$*.*

**Definition 2.2.** *Let* $\mathcal{A}$ *be a DiPA, and* $s \in seg(\mathcal{A})$ *be a segment. The* ***privacy loss*** $\text{loss}(s)$ *of a segment* $s \in seg(\mathcal{A})$ *is defined as*

$$\text{loss}(s) = \sup_{\rho \in segF(s)} \sup_{X' \sim X} \left( \frac{Pr(\rho|X)}{P(\rho|X')} \right)$$

*where* $X$ *and* $X'$ *vary over all pairs of neighbouring datasets.*

**Definition 2.3.** *Let* $f_\varepsilon(x)$ *be the probability density function of a random variable* $X$ *with* $X \sim Lap(0, 1/\varepsilon)$*.*

$$f_\varepsilon(x) = \frac{\varepsilon}{2} \exp(-\varepsilon|x|)$$

**Definition 2.4.** *Let* $F_\varepsilon(x)$ *be the cumulative distribution function of a random variable* $X$ *with* $X \sim Lap(0, 1/\varepsilon)$*.*

$$F_\varepsilon(x) = P(X \leq x) = \begin{cases} \frac{1}{2} \exp(\varepsilon x) & x < 0 \\ 1 - \frac{1}{2} \exp(-\varepsilon x) & x \geq 0 \end{cases}$$

# 3 Coupling proofs of privacy

A coupling proof of privacy

# 4 Shift-coupling proofs of privacy

Consider a path $\rho = q_0 \to q_1 \to \cdots \to q_m$. Consider inputs $X\langle 1\rangle = \langle a_1\langle 1\rangle, \ldots, a_m\langle 1\rangle\rangle$ and $X\langle 2\rangle = \langle a_1\langle 2\rangle, \ldots, a_m\langle 2\rangle\rangle$ such that $X\langle 1\rangle \sim X\langle 2\rangle$. We wish to show that there exists $\varepsilon \in (0, \infty)$ such that

$$\Pr\left[\rho | X\langle 1\rangle\right] \leq \exp(\varepsilon) \cdot \Pr\left[\rho | X\langle 2\rangle\right]$$

TODO: Write all of this later after consulting with Sky! ALSO, distinguish between path equivalence and output equivalence.

**Definition 4.1.** *Let $X = \langle a_1, \ldots, a_m\rangle$ be an input, and $\rho = q_0 \to q_1 \to \cdots \to q_m$ be a path taken by DiPA $\mathcal{A}$ on $X$. Define $\tilde{a_i}$ to be the value of* `insample` *on the $i$th transition in $\rho$ on input $X$.*

- The above is true if and only if $path_A(X)\Psi_\rho^{(\varepsilon,0)} path_A(X')$, where

  $$\Psi_\rho = \{(\rho_1, \rho_2) \in P \times P : \rho_1 = \rho \implies \rho_2 = \rho\}$$

  Shift-couplings are a technique to show that $path_A(X)\Psi_\rho^{(\varepsilon,0)} path_A(X')$ by constructing the couplings

  $$\tilde{a_i}\langle 1\rangle + \gamma_i (=)^{\#(\varepsilon_i, 0)} \tilde{a_i}\langle 2\rangle \qquad \forall i \in \{0, \ldots, m-1\}$$

  and showing that

  $$(\tilde{a}\langle 1\rangle + \gamma = \tilde{a}\langle 2\rangle) \implies path_A(X\langle 1\rangle)\Psi_\rho path_A(X\langle 2\rangle)$$

  thus showing that $path_A(X)\Psi_\rho^{(\varepsilon,0)} path_A(X')$ for

  $$\varepsilon = \sum_{i=0}^{m-1} \varepsilon_i$$

- A discussion that ends in choosing $\gamma$ shifts for each segment.
- Maybe: A discussion of the cost of each shift.

## 4.1 Connecting constraints

**Definition 4.2.** *Given a fixed path, we say that an assignment of shifts $\{\gamma_i\}$ is **valid** if*

$$(\tilde{a}\langle 1\rangle + \gamma = \tilde{a}\langle 2\rangle) \implies path_A(X\langle 1\rangle)\Psi_\rho path_A(X\langle 2\rangle)$$

**Definition 4.3.** *Let $\rho = q_0 \to q_1 \to \cdots \to q_m$ be a path, and let $i \in \{1, \ldots, m\}$. Define $at(i)$ to be the largest index $a(i) < i$ such that $\rho[a(i)] \to \rho[a(i)+1]$ is an assignment transition.*

**Definition 4.4.** *Let $\rho = q_0 \to q_1 \to \cdots \to q_m$ be a path. Define $t_i$ to be the transition $q_i \to q_{i+1}$.*

**Definition 4.5.** *Let $X = \langle a_1, \ldots, a_m\rangle$ be an input, and let $\rho = q_0 \to q_1 \to \cdots \to q_m$ be a path. Define $X[i]$ to be the value of $a_i$.*

Note that such an index must exist due to the initialization condition on DiPA.

**Proposition 4.1.** *Given $\rho = q_0 \to q_1 \to \cdots \to q_m$, an assignment of shifts $\{\gamma_i\}$ is valid if and only if it satisfies the following constraints for all $i \in \{0, \ldots, m-1\}$:*

$$\begin{aligned} \gamma_i &\leq \gamma_{at(i)} && \text{if } t_i \text{ has guard } < \\ \gamma_i &\geq \gamma_{at(i)} && \text{if } t_i \text{ has guard } \geq \end{aligned}$$

*Proof.* (Constraints $\implies$ valid) Suppose that the above constraints hold. We will show that $\{\gamma_i\}$ is valid using induction on $m = |\rho|$. Construct the couplings $\tilde{a}_i\langle 1\rangle + \gamma_i (=)^{(\varepsilon_i,0)} \tilde{a}_i\langle 2\rangle$ for all $i \in \{0, \ldots, m-1\}$.

For a base case, assume $m = 1$. Then $\rho$ consists of an assignment transition $t_0$ with `true` guard (initialization condition). The constraints are trivially satisfied, and we have that $path_A(X\langle 1\rangle)\Psi_\rho^{(0,0)} path_A(X\langle 2\rangle)$.

Assume that the constraints hold for all paths of length $m$. Let $\rho = q_0 \to q_1 \to \cdots \to q_m \to q_{m+1}$. We will show that $\{\gamma_i\}$ is valid for $\rho$. First, by the validity of $\{\gamma_i\}_{i=0}^{m-1}$ for $\rho[0:m]$ by the inductive hypothesis, we have that

$$path_A(X\langle 1\rangle)\Psi_{\rho[0:m]} path_A(X\langle 2\rangle)$$

by the inductive hypothesis. Now, assume $path_A(X\langle 1\rangle) = \rho$. We have $path_A(X\langle 2\rangle)[0:m] = \rho[0:m]$. Consider the last transition $t_m$ in $\rho$. Since $path_A(X\langle 1\rangle) = \rho$, we know that $t_m$ is traversed by $A$ on $X\langle 1\rangle$.

- If $t_m$ has guard `true`, then we trivially have that $path_A(X\langle 2\rangle) = \rho$.
- If $t_m$ has guard $<$, we have from the constraints that $\gamma_m \le \gamma_{at(m)}$. The value of the state variable $x\langle 1\rangle$ is

  $$x\langle 1\rangle = \tilde{a}_{at(m)}\langle 1\rangle$$

  and since $t_m$ is traversed by $A$ on $X\langle 1\rangle$, we have

  $$\begin{aligned} \tilde{a}_m\langle 1\rangle &< \tilde{a}_{at(m)}\langle 1\rangle \\ \tilde{a}_m\langle 2\rangle - \gamma_m &< \tilde{a}_{at(m)}\langle 2\rangle - \gamma_{at(m)} \\ \tilde{a}_m\langle 2\rangle &< \tilde{a}_{at(m)}\langle 2\rangle - (\gamma_{at(m)} - \gamma_m) < \tilde{a}_{at(m)}\langle 2\rangle \end{aligned}$$

  showing that $\tilde{a}_m\langle 2\rangle$ satisfies the guard of $t_m$. Thus, $path_A(X\langle 2\rangle) = \rho$.
- If $t_m$ has guard $\ge$, a similar argument as above shows that $path_A(X\langle 2\rangle) = \rho$.

Thus, assuming that $a\langle 1\rangle + \gamma = a\langle 2\rangle$, we have shown that $path_A(X\langle 1\rangle) = \rho \implies path_A(X\langle 2\rangle) = \rho$, which shows $path_A(X\langle 1\rangle)\Psi_\rho path_A(X\langle 2\rangle)$, and so $\{\gamma_i\}$ is valid.

(Valid $\implies$ constraints) Suppose that $\{\gamma_i\}$ is valid. Let $i \in \{0, \ldots, m-1\}$. We will show that the constraints hold for $i$.

We will run the argument above in reverse. Again, we use induction on the length $m = |\rho|$. For a base case, assume $m = 1$, and so $\rho$ consists of an assignment transition $t_0$ with `true` guard. The constraints are trivially satisfied, since there are none.

Assume that the constraints hold for all valid shift assignments on paths of length $m$, and let $\rho = q_0 \to q_1 \to \cdots \to q_m \to q_{m+1}$. Since $\{\gamma_i\}$ is valid for $\rho$, we have that $a\langle 1\rangle + \gamma = a\langle 2\rangle \implies path_A(X\langle 1\rangle)\Psi_\rho path_A(X\langle 2\rangle)$. Also, we have that $\{\gamma_i\}_{i=0}^{m-1}$ is valid for $\rho[0:m]$, and that constraints on transitions $t_i \in \rho[0:m]$ hold.

We will now show that the constraints on $t_m$ hold by cases on the guard of $t_m$.

- If $t_m$ has guard `true`, then there is no constraint on $\gamma_m$, and so the constraints hold.
- If $t_m$ has guard $<$, the constraint to be shown is $\gamma_m \le \gamma_{at(m)}$. Recall that we have $a\langle 1\rangle + \gamma = a\langle 2\rangle \implies path_A(X\langle 1\rangle)\Psi_\rho path_A(X\langle 2\rangle)$, showing that $t_m$ being traversed by $A$ on $X\langle 1\rangle$ leads $t_m$ to be traversed by $A$ on $X\langle 2\rangle$. Thus, we have

  $$\begin{aligned} \tilde{a}_m\langle 1\rangle < \tilde{a}_{at(m)}\langle 1\rangle &\implies \tilde{a}_m\langle 2\rangle < \tilde{a}_{at(m)}\langle 2\rangle \\ &\iff \tilde{a}_m\langle 1\rangle + \gamma_m < \tilde{a}_{at(m)}\langle 1\rangle + \gamma_{at(m)} \\ &\iff \tilde{a}_m\langle 1\rangle < \tilde{a}_{at(m)}\langle 1\rangle + (\gamma_{at(m)} - \gamma_m) \end{aligned}$$

  which is true if and only if $\gamma_m \le \gamma_{at(m)}$. Thus, the constraint holds.

- A symmetric argument shows that the constraint holds if $t_m$ has guard $\geq$.

Thus, the given constraints on $\gamma$ hold if and only if it is valid for $\rho$. □

**Theorem 1.** *A choice of valid $\{\gamma_i\}$ for each path $\rho$ and pairs of inputs $X\langle 1\rangle \sim X\langle 2\rangle$ which satisfies*

$$\gamma_i = 0 \qquad \text{if } t_i \text{ outputs } \texttt{insample} \text{ or } \texttt{insample'}$$

*shows that $\mathcal{A}$ is $(\varepsilon, 0)$-differentially private, where $\varepsilon = \sup_{X\langle 1\rangle \sim X\langle 2\rangle} \sup_\rho cost(\{\gamma_i\})$. TODO: Define cost. Maybe prove a proposition before this about (path equivalence and output constraints) $\iff$ output equivalence.*

*Proof.* By definition, almost. □

### 4.2 The cost of a shift-coupling

**Proposition 4.2.** *Consider a transition $t_i = q_i \to q_{i+1}$ which is traversed independently by $A$ on input $a_i\langle 1\rangle$ and $a_i\langle 2\rangle$. Let $\Delta_i = a_i\langle 2\rangle - a_i\langle 1\rangle$. Let $q_i$ draw from the distribution $Lap(0, \varepsilon_i)$ to noise* `insample`*. The $\varepsilon$-cost of the shift-coupling*

$$\tilde{a_i}\langle 1\rangle + \gamma_i (=)^{(c_i, 0)} \tilde{a_i}\langle 2\rangle$$

*is given by*

$$c_i = |\Delta_i - \gamma_i|\varepsilon_i$$

*Proof.* TODO □

### 4.3 A tight proof must regard input differences

Since the total validity constraints on $\{\gamma_i\}$ does not depend on $X\langle 1\rangle$ and $X\langle 2\rangle$, one might be tempted to produce a proof of privacy by choosing $\gamma_i$ to be the same for all $X\langle 1\rangle$ and $X\langle 2\rangle$, given a path $\rho$. Although this is possible, this does not in general produce a tight proof of privacy.

**Proposition 4.3.**

### 4.4 A tight proof must regard paths

## 5 The search for a tight proof as an optimization problem

### 5.1 Leading upto the abstract formulation

### 5.2 The problem as a whole

### 5.3 Specifying the inner minimization problem

Here, we specify how $A_{s,\boldsymbol{\Delta}}$ and $c_{s,\boldsymbol{\Delta}}$ are defined. Let $s = s_1 \hookrightarrow \cdots \hookrightarrow s_n$ be a sequence of segments with total number of transitions $m$. Let $\boldsymbol{\Delta} \in \{-1, 0, 1\}^m$ be a vector of input perturbations.

We define some notation as follows:

- Let $t^i_j$ denote the $j$th transition in segment $i$, and $\varepsilon^i_j$ denote the noise added to the input before that transition.
- Let $\Delta^i_j$ denote the entry of $\boldsymbol{\Delta}$ that corresponds to the input perturbation for the $j$th transition in segment $i$.
- Let $\gamma^i_j$ denote the entry of $\boldsymbol{\gamma} \in [-1, 1]^m$ that corresponds to the coupling shift for the $j$th transition in segment $i$ – this is to be determined by the inner minimization problem.

Then, the minimization problem over $\boldsymbol{\gamma}$ is as follows:

$$\min_{\boldsymbol{\gamma}\in[-1,1]^m} \quad \sum_{i=1}^{n}\sum_{j=1}^{|s_i|} |\gamma_j^i - \Delta_j^i|\varepsilon_i$$

$$\begin{aligned}
\text{subject to} \quad & \gamma_k^i \le \gamma_0^i && \text{if } t_k^i \text{ has guard } < \\
& \gamma_k^i \ge \gamma_0^i && \text{if } t_k^i \text{ has guard } \ge \\
& \gamma_0^i \le \gamma_0^k && \text{if } s_k \hookrightarrow s_i \text{ and guard}(s_i) \text{ is } < \\
& \gamma_0^i \ge \gamma_0^k && \text{if } s_k \hookrightarrow s_i \text{ and guard}(s_i) \text{ is } \ge \\
& \gamma_k^i = 0 && \text{if } t_k^i \text{ outputs } \texttt{insample} \\
& \gamma_k^i = \Delta_k^i && \text{if } t_k^i \text{ belongs to a cycle}
\end{aligned}$$

This can be rewritten as a linear program using standard techniques, producing a constraint matrix $A_{s,\boldsymbol{\Delta}}$ and a cost vector $c_{s,\boldsymbol{\Delta}}$.

### 5.4 A simplification

Let $\mathcal{F}$ be a finite family of segment sequences. The global optimization problem is to find

$$\max_{s\in\mathcal{F}} \max_{\boldsymbol{\Delta}\in\{-1,0,1\}^{|s|}} \left( \min_{\boldsymbol{\gamma}\in[-1,1]^{|s|}} \quad c_{s,\boldsymbol{\Delta}}^T \cdot \boldsymbol{\gamma} \quad \text{subject to} \quad A_{s,\boldsymbol{\Delta}} \cdot \boldsymbol{\gamma} \ge 0 \right)$$

## 6 Searching for a tight proof

# 7 Global Optimization

## 7.1 An abstract description of the global optimization problem

### 7.2 Constraints for all coupling strategies on a segment

In this section, we will try to understand the constraints that all valid coupling strategies on a segment must satisfy. For the purpose of this section, we will assume that our DiPA consists of a single segment. This assumption will be relaxed in later sections.

Here are some assumptions that are made throughout this document.

1. The noise added to inputs on each state $q_i$ is the same ($\varepsilon$).
2. Since we know tight coupling strategies for segments with cycles, we are restricting our attention to segments with no cycles.
3. We will only consider one segment at a time.

Some notation:

1. Let $N$ be the number of transitions in the segment.
2. The raw input received on the $i$th transition is denoted by $a_i$.
3. If we are considering two datasets $X\langle 1\rangle$ and $X\langle 2\rangle$, we will use $a_i\langle 1\rangle$ to denote the value of $a_i$ in the first dataset, and vice versa for $a_i\langle 2\rangle$.
4. Similarly, we use $x_i\langle 1\rangle$ to denote the random variable representing the value of `insample` before the $i$th transition when $A$ receives the input $X\langle 1\rangle$, and vice versa for $x_i\langle 2\rangle$.

A coupling strategy is a choice of values $\gamma_0, \ldots, \gamma_N$ such that $\gamma_i \in \Gamma$ for all $i$.

## 8 Are there tighter bounds than the tightest shift-coupling bound?

**Last Updated: Wednesday, June 28th, 2023**

The relevant definitions and lemmata for proofs in this section are in the appendix. It is also assumed, for now, that all transition outputs are in $\Gamma$.

### 8.1 $S^L$ is tight when there is an $L$-cycle

**Theorem 2.** *($S^L$ is tight for segments with L-cycles) Consider a segment $s \in seg(\mathcal{A})$ corresponding to the sequence of states $q_0 \to q_1 \to \cdots \to q_m$. If $s$ contains an L-cycle, then the L-cost of the segment gives a tight upper bound on the privacy loss of the segment. That is,*

$$\text{loss}(s) = \exp\left(2\varepsilon_0 + \sum_{i>0:guard(a_i)=\texttt{insample}\geq \texttt{x}} 2\varepsilon_i\right)$$

*given that state $q_i$ draws from the distribution $Lap(0, 1/\varepsilon_i)$ to noise* `insample`*.*

*Proof.* We will prove the result for when $\varepsilon_i = \varepsilon$ for all $i \geq 0$. The proof for the general case goes through in the same fashion. Let $f, F$ be the probability density function and cumulative distribution function of a random variable $X$ with $X \sim Lap(0, 1/\varepsilon)$ as defined in the appendix.

Since $s$ has an L-cycle, there exists a sequence of paths $\rho_i$ for $i \in \mathbb{N}$ each with $l_i$ number of L-transitions such that $\lim_{i\to\infty} l_i = \infty$. Let $m$ be the number of $G$-transitions in $\rho_i$. We will assume that this number is the same across all $\rho_i$.[1]

For each $\rho_i$, construct the adjacent pair of inputs $X_i, X_i'$ as follows. Let $X_i[j] = 0$ for all $j \in \{1, \ldots, |\rho_i|\}$, where $|\rho_i|$ is the number of transitions in $\rho$. Define $X_i[j]$ as follows:

$$X_i[j] = \begin{cases} 1 & \text{if } \rho_i[j] \to \rho_i[j+1] \text{ is an assignment transition or has guard } \texttt{insample}\geq\texttt{x} \\ -1 & \text{otherwise, in which case } \rho_i[j] \to \rho_i[j+1] \text{ has guard } \texttt{insample}<\texttt{x} \end{cases}$$

[1] Otherwise, $s$ has a G-cycle, and $\mathcal{A}$ is not differentially private. The privacy loss through $s$ is $\infty$, which matches the $L$-cost.

Let $\tilde{a_j}$ be the random variable representing the value of `insample` before the $j$th transition in $\rho$ on input $X_i$. Let $\tilde{b_j}$ be the random variable representing the value of `insample` before the $j$th transition in $\rho$ on input $X_i'$. Further, let $\Gamma_L = \{j : \rho_i[j] \to \rho_i[j+1]$ has guard `insample<x`$\}$, and $\Gamma_G = \{j : \rho_i[j] \to \rho_i[j+1]$ has guard `insample`$\geq$`x`$\}$.

Notice that $\tilde{a_j} = \tilde{b_j} + 1$ for $j \in \Gamma_L$, and $\tilde{a_j} + 1 = \tilde{b_j}$ for $j \in \{0\} \cup \Gamma_G$. Since $\tilde{a_j}$ is distributed as $Lap(X_i[j], 1/\varepsilon)$, we can write its probability density function as $f(x - X_i[j])$, and its cumulative distribution function as $F(x - X_i[j])$. A similar statement holds for $\tilde{b_j}$.

We may now compute and compare $Pr(\rho_i|X_i')$ and $\Pr(\rho_i|X_i)$ as follows.

$$\begin{aligned}
\Pr(\rho_i|X_i') &= \int_{-\infty}^{\infty} \Pr(\tilde{b_0} = x) \prod_{j \in \Gamma_L} \Pr(\tilde{b_j} < x) \prod_{j \in \Gamma_G} \Pr(\tilde{b_j} \geq x)\, \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \Pr(\tilde{b_0} = x) \prod_{j \in \Gamma_L} \Pr(\tilde{b_j} < x) \prod_{j \in \Gamma_G} \Pr(\tilde{b_j} \geq x)\, \mathrm{d}x \\
&= \int_{-\infty}^{\infty} f_\varepsilon(x - X_i[0]) \prod_{j \in \Gamma_L} F_\varepsilon(x - X_i[j]) \prod_{j \in \Gamma_G} (1 - F_\varepsilon(x - X_i[j]))\, \mathrm{d}x \\
&= \int_{-\infty}^{\infty} f(x-1) F(x+1)^{\ell_i} (1 - F(x-1))^m \\
&= \int_{-\infty}^{\infty} f(x) F(x+2)^{\ell_i} (1 - F(x))^m \\
&= \exp(2\varepsilon(m+1)) \left( \int_{(-\infty,-2)\cup(2,\infty)} f(x) F(x)^{\ell_i} (1 - F(x))^m \,\mathrm{d}x + g(\ell_i) \int_{-2}^{2} f(x) F(x+2)^{\ell_i} (1 - F(x))^m \right)
\end{aligned}$$

with $g(\ell_i) \to 1$ as $\ell_i \to \infty$. As we take $\ell_i \to \infty$, we see that

$$h(\ell_i) := \frac{\left( \int_{(-\infty,-2)\cup(2,\infty)} f(x) F(x)^{\ell_i} (1 - F(x))^m \,\mathrm{d}x + g(\ell_i) \int_{-2}^{2} f(x) F(x+2)^{\ell_i} (1 - F(x))^m \right)}{\Pr(\rho_i|X_i)} \to 1$$

and so as we take the supremum over $\rho_i$ below, we get:

$$\begin{aligned}
\mathrm{loss}(s) \geq \sup_{\rho_i} \frac{\Pr(\rho_i|X_i')}{\Pr(\rho_i|X_i)} &= \exp(2\varepsilon(m+1)) \sup_{\rho_i} \{h(l_i)\} \\
&= \exp(2\varepsilon(m+1))
\end{aligned}$$

We know that $S^L$ is tight, and gives the bound $\exp(2\varepsilon(m+1))$. Thus, we have shown that $\mathrm{loss}(s) = \exp(2\varepsilon(m+1))$, as desired. □

## 8.2 An alternative coupling strategy: $S^J$

**Definition 8.1.** *$S^J$ is a coupling strategy in which we do not couple the noised threshold, but couple the results of all other transitions with twice the cost. [TODO: Describe in more detail]*

**Theorem 3.** *Let $s = q_0 \to \cdots \to q_m$ be a segment with only L-transitions. If $S^J$ is the least-cost coupling strategy on $s$, then it provides a tight bound on* $\mathrm{loss}(s)$ *given by*

$$\mathrm{loss}(s) = \sum_{i=1}^{m} 2\varepsilon_i$$

*Proof.* I have a proof for this, but I will add it into this document soon. [TODO]

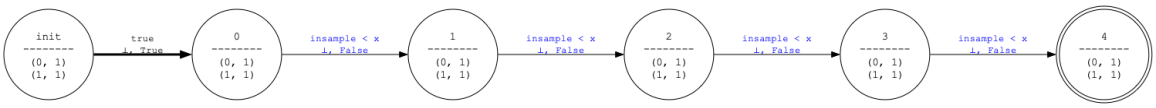


Figure 1: A segment $s$ with only L-transitions.

---

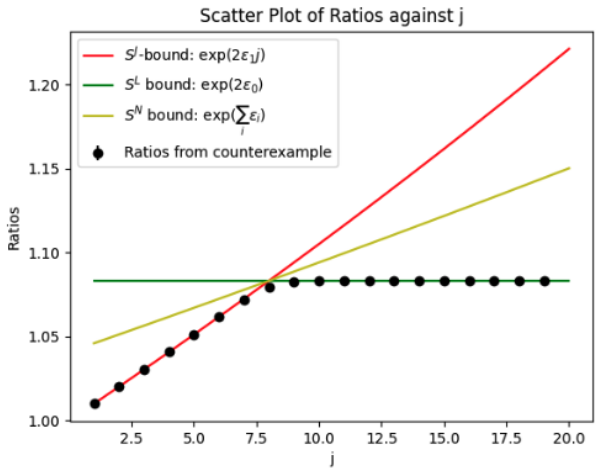


Figure 2:

□

**Hypothesis 8.1.** *For segments which contain only L-transitions and for which the J-cost exceeds the L-cost, $S^L$ is tight.*

*Proof.* I think this is true from the graph above, but I need to prove it.

Note June 28 2023: I think this is not true for segments that contain both $L$-transitions and $G$-transitions. □

# A Lemmata

## A.1 Properties of $f_\varepsilon$ and $F_\varepsilon$

**Lemma 4.** *For $x \leq 0$, we have*

$$F_\varepsilon(x) = \exp(2\varepsilon)F_\varepsilon(x-2)$$

*and equivalently for $x \leq -2$, we have*

$$F_\varepsilon(x+2) = \exp(2\varepsilon)F_\varepsilon(x)$$

**Lemma 5.** *For $x \geq 0$, we have*

$$1 - F_\varepsilon(x) = \exp(2\varepsilon)(1 - F_\varepsilon(x+2))$$

**Lemma 6.** *For $x \geq 0$, we have*

$$f_\varepsilon(x) = \exp(2\varepsilon)f_\varepsilon(x+2)$$

## A.2 For the proof of Theorem 1

**Lemma 7.**

$$\int_{-\infty}^{-2} f_\varepsilon(x)F_\varepsilon(x+2)^\ell(1-F_\varepsilon(x))^m\,\mathrm{d}x = \exp(2\varepsilon\ell)\int_{-\infty}^{-2} f_\varepsilon(x)F_\varepsilon(x)^\ell(1-F_\varepsilon(x))^m\,\mathrm{d}x$$

*Proof.* From Lemma 4, we have that

$$\begin{aligned}\int_{-\infty}^{-2} f_\varepsilon(x)F_\varepsilon(x+2)^\ell(1-F_\varepsilon(x))^m\,\mathrm{d}x &= \int_{-\infty}^{-2} f_\varepsilon(x)(\exp(2\varepsilon)F_\varepsilon(x))^\ell(1-F_\varepsilon(x))^m\,\mathrm{d}x\\ &= \exp(2\varepsilon\ell)\int_{-\infty}^{-2} f_\varepsilon(x)F_\varepsilon(x)^\ell(1-F_\varepsilon(x))^m\,\mathrm{d}x\end{aligned}$$

□

**Lemma 8.**

$$\int_{0}^{\infty} f_\varepsilon(x)F_\varepsilon(x+2)^\ell(1-F_\varepsilon(x))^m\,\mathrm{d}x = \exp(2\varepsilon m)\int_{2}^{\infty} f_\varepsilon(x)F_\varepsilon(x)^\ell(1-F_\varepsilon(x))^m\,\mathrm{d}x$$

*Proof.* From Lemma 5 and 6, we have that

$$\begin{aligned}\int_{0}^{\infty} f_\varepsilon(x)F_\varepsilon(x+2)^\ell(1-F_\varepsilon(x))^m\,\mathrm{d}x &= \int_{0}^{\infty} \exp(2\varepsilon)f_\varepsilon(x+2)F_\varepsilon(x+2)^\ell(\exp(2\varepsilon)(1-F_\varepsilon(x+2)))^m\,\mathrm{d}x\\ &= \exp(2\varepsilon m)\int_{0}^{\infty} f_\varepsilon(x+2)F_\varepsilon(x+2)^\ell(1-F_\varepsilon(x+2))^m\,\mathrm{d}x\\ &= \exp(2\varepsilon(m+1))\int_{2}^{\infty} f_\varepsilon(x)F_\varepsilon(x)^\ell(1-F_\varepsilon(x))^m\,\mathrm{d}x\end{aligned}$$

□

**Lemma 9.** *There exists a function $g:\mathbb{N}\to\mathbb{R}$ such that*

$$\int_{-2}^{0} f_\varepsilon(x)F_\varepsilon(x+2)^\ell(1-F_\varepsilon(x))^m\,\mathrm{d}x = g(\ell)\exp(2\varepsilon(m+1))\int_{-2}^{2} f_\varepsilon(x)F_\varepsilon(x)^\ell(1-F_\varepsilon(x))^m\,\mathrm{d}x$$

*with $g(\ell)\to 1$ as $\ell\to\infty$.*

*Proof.* I'm not sure yet how to prove this, although I strongly suspect that the $(m+1)$ term comes from the fact that $f_\varepsilon(x)$ is the derivative of $-(1-F_\varepsilon(x))$, and it is taken to the $m$th power. Its integral should behave like a polynomial of degree $m+1$ evaluated at 2, which corresponds to $\exp(2\varepsilon(m+1))$. □