Reward Identification in Inverse Reinforcement Learning

A. Proofs

First we state some lemmas that will be used in proving the main theorems.

A.1. Proper MDP Models

Example 1. *Let J_{MaxEnt} be the MaxEntRL objective. Then, the MaxEnt MDP model $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J_{\mathrm{MaxEnt}}]$ is proper.*

Proof. Let R be any set of reward functions. We need to show that $\forall r, \hat{r} \in R, r \cong_\tau \hat{r} \Rightarrow p_r = p_{\hat{r}}$. If $r \cong_\tau \hat{r}$, then r and $\hat{r}$ have trajectory level rewards shifted by a constant, i.e for all $x_0 \in \mathcal{X}^0$ there exists a constant c_{x_0} such that $\forall \tau \in \Omega[x_0, d, T], \hat{r}(\tau) = r(\tau) + c_{x_0}$. It suffices to show that the optimal policies for $r, \hat{r}$ are the same. For any policy family Π,

$$\begin{aligned}
\arg\max_{\pi\in\Pi} \mathbb{E}_{\tau\sim\pi}[\hat{r}(\tau)] + \mathcal{H}(\pi) &= \arg\max_{\pi\in\Pi} \Big(\sum_{x_0\in\mathcal{X}^0} \sum_{\tau\in\Omega[x_0,d,T]} p(\tau;\pi)\hat{r}(\tau) \Big) + \mathcal{H}(\pi) \\
&= \arg\max_{\pi\in\Pi} \Big(\sum_{x_0\in\mathcal{X}^0} \sum_{\tau\in\Omega[x_0,d,T]} p(\tau;\pi)(r(\tau) + c_{x_0}) \Big) + \mathcal{H}(\pi) \\
&= \arg\max_{\pi\in\Pi} \Big(\sum_{x_0\in\mathcal{X}^0} \sum_{\tau\in\Omega[x_0,d,T]} p(\tau;\pi)r(\tau) \Big) + \Big(\sum_{x_0\in\mathcal{X}^0} P_0(x_0)c_{x_0} \Big) + \mathcal{H}(\pi) \\
&= \arg\max_{\pi\in\Pi} \mathbb{E}_{\tau\sim\pi}[r(\tau)] + \mathbb{E}_{x_0\sim P_0}[c_{x_0}] + \mathcal{H}(\pi) \\
&= \arg\max_{\pi\in\Pi} \mathbb{E}_{\tau\sim\pi}[r(\tau)] + \mathcal{H}(\pi)
\end{aligned}$$

where $\mathcal{H}(\pi) := \mathbb{E}_\pi[-\sum_{t=0}^T \gamma^t \log \pi(a_t|s_t)]$ is the γ-discounted causal entropy. The last step holds since $\mathbb{E}_{x_0\sim P_0}[c_{x_0}]$ is constant with respect to π. $\square$

A.2. Weak Identifiability

Lemma 1. *For all reward families R, $r, \hat{r} \in R$, and any (d, T), $(r \cong_{x,a} \hat{r}) \Rightarrow (r \cong_\tau \hat{r})$*

Proof. Let $r, \hat{r} \in R$ be rewards such that $r \cong_{x,a} \hat{r}$. For all $\tau, \tau' \in \Omega[d, T]$, where $\tau = (x_t, a_t)_{0\le t\le T}, \tau' = (x'_t, a'_t)_{0\le t\le T}$,

$$\begin{aligned}
\hat{r}(\tau) - r(\tau) &= \sum_{t=0}^{T} \gamma^t (\hat{r}(x_t, a_t) - r(x_t, a_t)) && (6) \\
&= \sum_{t=0}^{T} \gamma^t (\hat{r}(x'_t, a'_t) - r(x'_t, a'_t)) && (7) \\
&= \hat{r}(\tau') - r(\tau')
\end{aligned}$$

where 6 $\to$ 7 holds since for all $0 \le t \le T$, $(x_t, a_t), (x'_t, a'_t) \in \mathcal{X} \times \mathcal{A}$ and so $\hat{r}(x_t, a_t) - r(x_t, a_t) = \hat{r}(x'_t, a'_t) - r(x'_t, a'_t)$. Thus, $r \cong_\tau \hat{r}$ $\square$

Proposition 1. *A proper MDP model is strongly identifiable only if it is weakly identifiable*

Proof. We prove the contrapositive: if a proper MDP model is not weakly identifiable it is also not strongly identifiable. Let $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ be a proper MDP model that is not weakly identifiable. Since the model is not weakly identifiable, there exists $r, \hat{r} \in R$ such that either ($r \cong_\tau \hat{r}$ and $p_r \neq p_{\hat{r}}$) or ($r \ncong_\tau \hat{r}$, and $p_r = p_{\hat{r}}$). Since the model is proper the former cannot be true. Thus it must be that there exists $r, \hat{r} \in R$ such that $r \ncong_\tau \hat{r}$, and $p_r = p_{\hat{r}}$. Then, by the contrapositive of Lemma 1, $r \ncong_{x,a} \hat{r}$. Thus, $p_r = p_{\hat{r}} \not\Rightarrow r \cong_{x,a} \hat{r}$ and $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is not strongly identifiable as desired. $\square$

Theorem 1. *Let $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J_{\mathrm{MaxEnt}}]$ be a MaxEnt MDP model and $R \subseteq \{r \mid r : \mathcal{X} \times \mathcal{A} \to \mathbb{R}\}$ be any set of rewards. Then, for all domains $d := (\mathcal{X}, \mathcal{A}, P, P_0, \gamma)$ consisting of deterministic transition dynamics, i.e $\forall (x, a), |\mathrm{supp}(P(\cdot|x, a))| = 1$, a deterministic initial state, i.e $|\mathrm{supp}(P_0)| = 1$, and $T \geq 0$, $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J_{\mathrm{MaxEnt}}]$ is weakly identifiable.*

Proof. We seek to show that $\forall r, \hat{r} \in R, (r \cong_\tau \hat{r}) \iff (p_r = p_{\hat{r}})$. Since $\mathcal{P}_{\mathrm{MDP}}$ is a MaxEnt MDP model, it is proper by Example 1 and as a result $(r \cong_\tau \hat{r}) \Rightarrow (p_r = p_{\hat{r}})$. We are left to prove that $\forall r, \hat{r} \in R, (p_r = p_{\hat{r}}) \Rightarrow (r \cong_\tau \hat{r})$

From Ziebart et al. (2008), for all MDPs with deterministic dynamics and a deterministic initial state, the trajectory distribution of the MaxEnt optimal policy is

$$p_r(\tau) = \frac{e^{r(\tau)}}{Z_r}$$

where $Z_r = \int_{\Omega[d,T]} e^{r(\tau')} d\tau'$ is the partition function. Then, $\forall \tau \in \Omega[d, T]$

$$\begin{aligned} p_r(\tau) &= p_{\hat{r}}(\tau) \\ \log p_r(\tau) &= \log p_{\hat{r}}(\tau) \\ r(\tau) - \log Z_r &= \hat{r}(\tau) - \log Z_{\hat{r}} \\ r(\tau) &= \hat{r}(\tau) + \log \frac{Z_r}{Z_{\hat{r}}} \end{aligned}$$

Since, $\log \frac{Z_r}{Z_{\hat{r}}}$ is a constant w.r.t τ, we have $r \cong_\tau \hat{r}$ as desired. □

Proposition 2. *Let $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ be an MDP model that is weakly identifiable. Then, it is strongly identifiable if and only if for all $r, \hat{r} \in R, (r \cong_\tau \hat{r}) \Rightarrow (r \cong_{x,a} \hat{r})$. In other words, $\forall r \in R, [r]_\tau \subseteq [r]_{x,a}$.*

Proof. Let $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ be weakly identifiable. We abbreviate Strongly Identifiable as S.I.

• (Sufficiency) $\forall r \in R, [r]_\tau \subseteq [r]_{x,a} \Rightarrow \mathcal{P}_{\mathrm{MDP}}$ is S.I.

By weak identifiability, for all $r, \hat{r} \in R$, $(p_r = p_{\hat{r}}) \Rightarrow (r \cong_\tau \hat{r})$ and by $(r \cong_\tau \hat{r}) \Rightarrow (r \cong_{x,a} \hat{r})$, we have $r \cong_{x,a} \hat{r}$. Thus, $(p_r = p_{\hat{r}}) \Rightarrow (r \cong_{x,a} \hat{r})$.

By Lemma 1, for all $r, \hat{r} \in R$, $(r \cong_{x,a} \hat{r}) \Rightarrow (r \cong_\tau \hat{r})$, and by weak identifiability $(r \cong_\tau \hat{r}) \Rightarrow (p_r = p_{\hat{r}})$. Thus, $(r \cong_{x,a} \hat{r}) \Rightarrow (p_r = p_{\hat{r}})$.

We have $\forall r, \hat{r} \in R, (r \cong_{x,a} \hat{r}) \iff (p_r = p_{\hat{r}})$ as desired.

• (Necessity) $\mathcal{P}_{\mathrm{MDP}}$ is S.I $\Rightarrow \forall r \in R, [r]_\tau \subseteq [r]_{x,a}$.

We prove the contrapositive. Suppose there exists $r, \hat{r} \in R$ such that $r \cong_\tau \hat{r}$ but $r \ncong_{x,a} \hat{r}$. By weak identifiability, $(r \cong_\tau \hat{r}) \Rightarrow (p_r = p_{\hat{r}})$, so $(p_r = p_{\hat{r}}) \not\Rightarrow (r \cong_{x,a} \hat{r})$. Thus, $\mathcal{P}_{\mathrm{MDP}}[R; d, J, T]$ is not strongly identifiable. □

Corollary 1. *Let $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ be an MDP model that is weakly identifiable, R be the set of all rewards, $|\mathcal{X}^0| = 1$, and $\gamma = 1$. Then, it is strongly identifiable if and only if $\mathrm{rank}(A[d, T]) = |\mathcal{X} \times \mathcal{A}|$*

Proof. Let $\mathcal{P}_{\mathrm{MDP}}$ be weakly identifiable, $\gamma = 1$.

• (Sufficiency) We seek to show that if $\mathrm{rank}(A[d, T]) = \mathcal{X} \times \mathcal{A}$, then $\mathcal{P}_{\mathrm{MDP}}$ is strongly identifiable. By Proposition 2, it suffices to show that $\forall r, \hat{r} \in R, (r \cong_\tau \hat{r}) \Rightarrow (r \cong_{x,a} \hat{r})$, i.e trajectory equivalence implies state-action equivalence.

Since $A[d, T]$ is full rank, the solution to the linear system $A[d, T]\mathbf{r}_{x,a} = \mathbf{r}_\tau$ is unique for any $\mathbf{r}_\tau$. Let $\mathbf{r}_\tau, \hat{\mathbf{r}}_\tau$ be two trajectory equivalent rewards such that $\mathbf{r}_\tau = \hat{\mathbf{r}}_\tau + \boldsymbol{c}$ for some constant vector $\boldsymbol{c} = (c, ..., c) \in \mathbb{R}^{|\Omega[d,T]|}$. Then,

$$\begin{aligned} A[d, T]\mathbf{r}_{x,a} - A[d, T]\hat{\mathbf{r}}_{x,a} &= \mathbf{r}_\tau - \hat{\mathbf{r}}_\tau \\ A[d, T](\mathbf{r}_{x,a} - \hat{\mathbf{r}}_{x,a}) &= \boldsymbol{c} \end{aligned}$$

Since $A[d, T]$ is a trajectory matrix, $\forall i, \sum_j A_{ij}[d, T] = T + 1$, i.e all feasible trajectories are of the same length and hence visit the same number of (not necessarily distinct) nodes. Thus, one solution to $A[d, T](\mathbf{r}_{x,a} - \hat{\mathbf{r}}_{x,a}) = \boldsymbol{c}$ is to let

$\mathbf{r}_{x,a} - \hat{\mathbf{r}}_{x,a} = (\frac{c}{T+1}, ..., \frac{c}{T+1})$. In fact, since $A[d,T]$ is full rank, $(\frac{c}{T+1}, ..., \frac{c}{T+1})$ is the only solution and thus $\mathbf{r}_{x,a}, \hat{\mathbf{r}}_{x,a}$ are trajectory equivalent, $\forall x, a \in \mathcal{X} \times \mathcal{A}, r(x,a) = \hat{r}(x,a) + \frac{c}{T+1}$ implying $r \cong_{x,a} \hat{r}$.

• (Necessity) We show that if $\mathcal{P}_{\text{MDP}}$ be strongly identifiable, then $\text{rank}(A[d,T]) = |\mathcal{X} \times \mathcal{A}|$. By strong identifiability, $\forall r, \hat{r} \in R, (r \cong_\tau \hat{r}) \Rightarrow (r \cong_{x,a} \hat{r})$ and thus general solutions to

$$A[d,T]\mathbf{r}_{x,a} = \mathbf{r}_\tau$$

must only be constant shifts of a particular solution. Equivalently, $\ker(A[d,T])$ must only contain constant vectors. We then claim that in fact $\ker(A[d,T])$ only contains the zero vector and thus $A[d,T]$ is full rank.

Suppose for contradiction that $\boldsymbol{c} \in \ker(A[d,T])$ for some non-zero constant vector $\boldsymbol{c}$. Then, for any scalar $k \in \mathbb{R}$, $k\boldsymbol{c} \in \ker(A[d,T])$. Thus the kernel must contain all constant vectors. Pick a strictly positive constant vector $\boldsymbol{c}^+ = (c^+, ..., c^+)$ where $c^+ > 0$. Then, $\boldsymbol{c}^+ \in \ker(A[d,T]) \Rightarrow A[d,T]\boldsymbol{c}^+ = 0$, so $\forall i, \sum_j A_{ij}[d,T]c^+ = c^+ \sum_j A_{ij}[d,T] = 0 \Rightarrow \forall i, \sum_j A_{ij}[d,T] = 0$. Since $A[d,T]$ is a trajectory (path) matrix, its entries represent visitation counts of a state-action pair and thus are all non-negative, i.e $\forall i, j, A_{ij}[d,T] \geq 0$. Therefore, $(\forall i, \sum_j A_{ij}[d,T] = 0) \Rightarrow (\forall i, j, A_{ij}[d,T] = 0)$, so $A[d,T]$ is the zero-matrix. Then, $\ker(A[d,T]) = \mathbb{R}^{|\mathcal{X} \times \mathcal{A}|}$ which contradicts strong identifiability. Therefore, $\ker(A[d,T])$ can only contain the zero vector and $A[d,T]$ is full rank, i.e $\text{rank}(A[d,T]) = \mathcal{X} \times \mathcal{A}$. □

A.3. Properties of Domain Graphs

Lemma 2. *Let $G_d = (V_d, E_d, V_d^0)$ be a domain graph.*

1. *(Commutative) For all $V \subseteq V_d$ and $t, t' \geq 0$, $L_{t'}(L_t(V)) = L_{t+t'}(V)$*
2. *(Monotonic) For all $V, V' \subseteq V_d$ such that $V \subseteq V'$ and $t \geq 0$, $L_t(V) \subseteq L_t(V')$*

Proof. • (Commutative) We first prove that $L_{t'}(L_t(V)) \subseteq L_{t+t'}(V)$. Let $v \in L_{t'}(L_t(V))$, then by Definition 6, $\exists \zeta' = (v'_i)_{0 \leq i \leq t'}$ such that $v'_{t'} = v$ and $v'_0 \in L_t(V)$, i.e $\exists \zeta = (v_i)_{0 \leq i \leq t}$ where $v_t = v'_0, v_0 \in V$. Then, $v \in L_{t+t'}(V)$ since there exists a path $\zeta \oplus \zeta'_{1:} = (v_0, ..., v_t, v'_1, ..., v'_{t'})$ such that $v'_{t'} = v$ and $v_0 \in V$.

Next, we prove that $L_{t+t'}(V) \subseteq L_{t'}(L_t(V))$. If $v \in L_{t+t'}(V)$, then $\exists \zeta'' = (v''_i)_{0 \leq i \leq t+t'}$ such that $v''_{t+t'} = v$ and $v''_0 \in V$. Then, there exists paths $\zeta = (v_i)_{0 \leq i \leq t'} = (v''_0, ..., v''_t)$ and $\zeta' = (v'_i)_{0 \leq i \leq t'} = (v''_t, ..., v_{t+t'})$ which can be joined to form ζ''. Therefore, $v \in L_{t'}(L_t(V))$ since there exists a path ζ' such that $v'_{t'} = v$ and $v'_0 \in L_t(V)$ since ζ is a path such that $v_t = v'_0$ and $v_0 \in V$.

• (Monotonic) Let $V, V' \subseteq V_d$ satisfy $V \subseteq V'$. If $v \in L_t(V)$, then by Definition 6, $\exists \zeta = (v_i)_{0 \leq i \leq t}$ such that $v_t = v$ and $v_0 \in V$. Since $V \subseteq V'$, $v_0 \in V'$ as well. Therefore, $v \in L_t(V')$. □

Lemma 3. *If G_d is coverable, then $\cup_{(x,a) \in \mathcal{X} \times \mathcal{A}} \text{supp}(P(\cdot|x,a)) = \mathcal{X}$*

Proof. Since G_d is coverable, there exists $v \in V_d^0$ and $t \geq 0$ such that $L_t(v) = V_d$. If G_d is 0-coverable, i.e $L_0(v) = \{v\} = V_d = \mathcal{X} \times \mathcal{A}$, then $|\mathcal{X} \times \mathcal{A}| = 1$ and thus $\text{supp}(P(\cdot|x,a)) = \{x\} = \mathcal{X}$. For $t \geq 1$, since $L_t(v) = L_1(L_{t-1}(v)) = V_d$ and $L_{t-1}(v) \subseteq V_d$, by Lemma 2 monotonicity, we have $L_1(L_{t-1}(v)) = V_d \subseteq L_1(V_d)$. Since $L_1(V_d) \subseteq V_d$, it must be that $L_1(V_d) = V_d = \mathcal{X} \times \mathcal{A}$. By definition of layers, $L_1(V_d) = (\cup_{(x,a) \in \mathcal{X} \times \mathcal{A}} \text{supp}(P(\cdot|x,a))) \times \mathcal{A}$ and thus $\cup_{(x,a) \in \mathcal{X} \times \mathcal{A}} \text{supp}(P(\cdot|x,a)) = \mathcal{X}$. □

Lemma 4. *Let G_d be a domain graph and $v \in V_d$ be t-covering. Then for all $t' \geq t$, $L_{t'}(v) = V_d$.*

Proof. We prove by induction.

• Base $t' = t$: trivially holds since $L_t(v) = V_d$ by definition of a covering vertex.

• For $t' \geq t$: $L_{t'}(v) = V_d \Rightarrow L_{t'+1}(v) = V_d$.

$$L_{t'+1}(v) = L_1(L_{t'}(v)) = L_1(V_d)$$

$L_t(v) = L_1(L_{t-1}(v)) = V_d$ and $L_{t-1}(v) \subseteq V_d$, we have that $L_1(L_{t-1}(v)) = V_d \subseteq L_1(V_d)$ by Lemma 2 monotonicity. Since $L_1(V_d) \subseteq V_d$, it must be that $L_1(V_d) = V_d$. □

Proposition 3. *Let G_d be strongly connected. Then, G_d is aperiodic if and only if it is coverable.*

Proof. (aperiodic ⇒ coverable) If G_d is aperiodic, there exists two cycles $C = (v_i)_{0\leq i\leq k}, C' = (v'_i)_{0\leq i\leq k'}$ of coprime length k, k'. For any $v \in V_d^0$ and any destination vertex $\tilde{v} \in V_d$ consider paths that start from v, reaches v_0 via a shortest path $\zeta^{v\to v_0}$, loops n times around cycle C back to v_0, reaches v'_0 via a shortest path $\zeta^{v_0\to v'_0}$, loops n' times around cycle C' back to v'_0, and finally reaches $\tilde{v}$ via a shortest path $\zeta^{v'_0\to\tilde{v}}$, i.e

$$\zeta^{v\to\tilde{v}} = \zeta^{v\to v_0} \oplus n\cdot C_{1:} \oplus \zeta_{1:}^{v_0\to v'_0} \oplus n'\cdot C'_{1:} \oplus \zeta_{1:}^{v'_0\to\tilde{v}}$$

The paths $\zeta^{v\to v_0}, \zeta^{v_0\to v'_0}, \zeta^{v'_0\to\tilde{v}}$ exist by strong connectivity of G_d. We let $|\zeta|$ denote the length of a path. Then,

$$|\zeta^{v\to\tilde{v}}| = nk + n'k' + |\zeta^{v\to v_0}| + |\zeta^{v_0\to v'_0}| + |\zeta^{v'_0\to\tilde{v}}| \tag{8}$$

Since k, k' are coprime, for all $|\zeta^{v\to\tilde{v}}| \geq (k-1)(k'-1) + |\zeta^{v\to v_0}| + |\zeta^{v_0\to v'_0}| + |\zeta^{v'_0\to\tilde{v}}|$, there exists n, n' such that Eq. 8 holds. (Corollary 2 of Denardo (1977)) Furthermore, since $|\zeta^{v\to v_0}|, |\zeta^{v_0\to v'_0}|, |\zeta^{v'_0\to\tilde{v}}| \leq |V_d|$ since they are shortest paths. Thus, for any destination vertex $\tilde{v} \in V_d$ and all lengths $T \geq (k-1)(k'-1) + 3|V_d|$, there exists a path $\zeta^{v\to\tilde{v}}$ such that $|\zeta^{v\to\tilde{v}}| = T$. Therefore, G_d is coverable.

(coverable ⇒ aperiodic) If G_d is coverable, there exists $v \in V_d^0$ and $t \geq 0$ such that $L_t(v) = V_d$. If $t = 0$, then $V_d = \{v\}$ and there must be an edge $(v, v) \in E_d$. Therefore, there exists cycles $(v, v), (v, v, v)$ which have coprime lengths 1 and 2, respectively. For $t \geq 1$, by Lemma 4, $L_{t+1}(v) = V_d$. Since $v \in L_t(v)$ and $v \in L_{t+1}(v)$, there exists cycles of coprime length $t, t+1$ that start and end at v. Thus, G_d is aperiodic.

□

A.4. Strong Identifiability

Theorem 2. (Strong Identification Condition) *For all (d, r, T, J) such that the MDP model $\mathcal{P}_{\text{MDP}}[R; d, T, J]$ is proper and G_d is strongly connected,*

- *(Sufficiency) $\mathcal{P}_{\text{MDP}}[R; d, T, J]$ is weakly identifiable, G_d is T_0-coverable, and $T \geq 2T_0 \Rightarrow \mathcal{P}_{\text{MDP}}[R; d, T, J]$ is strongly identifiable*
- *(Necessity) $\mathcal{P}_{\text{MDP}}[R; d, T, J]$ is strongly identifiable $\Rightarrow \mathcal{P}_{\text{MDP}}[R; d, T, J]$ is weakly identifiable, G_d is coverable.*

Proof. Let $\mathcal{P}_{\text{MDP}}[R; d, T, J]$ be proper and G_d be strongly connected.

• (Sufficiency) Let $\mathcal{P}_{\text{MDP}}[R; d, T, J]$ be proper and weakly identifiable, G_d be strongly connected and T_0-covering, and $T \geq 2T_0$. By Proposition 2 it suffices to show that

$$\forall x \in \mathcal{X}^0, \forall \tau, \tau' \in \Omega[x, d, T], \hat{r}(\tau) - r(\tau) = \hat{r}(\tau') - r(\tau') \Rightarrow$$
$$\forall (x, a), (x', a') \in \mathcal{X}\times\mathcal{A}, \hat{r}(x, a) - r(x, a) = \hat{r}(x', a') - r(x', a')$$

In the language of domain graphs, this statement translates to:

$$\forall v \in V_d^0, \forall \zeta, \zeta' \in Z[v, d, T], \hat{r}(\zeta) - r(\zeta) = \hat{r}(\zeta') - r(\zeta') \Rightarrow \forall v, v' \in V_d, \hat{r}(v) - r(v) = \hat{r}(v') - r(v')$$

Let $r, \hat{r}$ be any two rewards such that, $\forall v \in V_d^0$, $\forall \zeta, \zeta' \in Z[v, d, T]$, $\hat{r}(\zeta) - r(\zeta) = \hat{r}(\zeta') - r(\zeta')$ or equivalently $\hat{r}(\zeta) - \hat{r}(\zeta') = r(\zeta) - r(\zeta')$. Let $v_0^* \in V_d^0$ be any vertex that is T_0-covering and for any integer $t \geq 0$ let H_t be the statement that,

$$\forall v, v' \in L_t(v_0^*), \hat{r}(v') - \hat{r}(v) = r(v') - r(v)$$

Since v_0^* is T_0-covering, we have that $L_{T_0}(v_0^*) = V_d$, so it suffices to prove H_{T_0}. We prove by strong induction.

H_0: Trivially true, since $L_0(v_0^*) = \{v_0^*\}$ only has one element.

$H_{<t} \Rightarrow H_t$ for all $0 < t \leq T_0$: Let $\zeta^0 \in Z[v_0^*, d, T-t]$ be any base path of length $T-t$ that starts at v_0^* and reaches v_0^* again $T-t$ steps. Such a base path exists for all $0 \leq t \leq T_0$ since $T \geq 2T_0 \Rightarrow T-t \geq T_0$ and so by Lemma 4, $v_0^* \in L_{T_0}(v_0^*) \Rightarrow v_0^* \in L_{T-t}(v_0^*)$.

We will use Z_t to denote the set of all paths of length T that starts at v_0^* and follows ζ^0 to reach v_0^* again at time $T-t$, then reaches a vertex in $L_t(v_0^*)$ in t steps, i.e.

$$Z_t = \{\zeta \in Z[v_0^*, d, T] \mid \zeta_{:-t} = \zeta^0\}$$

It's then clear that the set terminal vertices of paths in Z_t is equal to $L_t(v_0^*)$, i.e. $\{v \mid \exists \zeta := (v_t)_{0\leq t\leq T} \in Z_t \text{ s.t } v = v_T\} = L_t(v_0^*)$ since Z_t contains all possible paths that take t steps after reaching v_0^*

Consider any two $\zeta, \zeta' \in Z_t$ where $\zeta = (v_t)_{0\leq t\leq T}, \zeta' = (v'_t)_{0\leq t\leq T}$.

$$\hat{r}(\zeta) - \hat{r}(\zeta') = \hat{r}(\zeta_{:-t+1}) - \hat{r}(\zeta'_{:-t+1}) + \hat{r}(\zeta_{-t+1:}) - \hat{r}(\zeta'_{-t+1:}) + \gamma^T(\hat{r}(v_T) - \hat{r}(v'_T)) \tag{9}$$

$$r(\zeta) - r(\zeta') = r(\zeta_{:-t+1}) - r(\zeta'_{:-t+1}) + r(\zeta_{-t+1:}) - r(\zeta'_{-t+1:}) + \gamma^T(r(v_T) - r(v'_T)) \tag{10}$$

Since $\zeta_{:-t+1} = \zeta'_{:-t+1} = \zeta^0$, we have $\hat{r}(\zeta_{:-t+1}) - \hat{r}(\zeta'_{:-t+1}) = r(\zeta_{:-t+1}) - r(\zeta'_{:-t+1}) = 0$. Furthermore,

$$\begin{aligned}
\hat{r}(\zeta_{-t+1:}) - \hat{r}(\zeta'_{-t+1:}) &= \sum_{t'=0}^{t-1} \gamma^{T-t'}(\hat{r}(v_{T-t'}) - \hat{r}(v'_{T-t'})) \\
&= \sum_{t'=0}^{t-1} \gamma^{T-t'}(r(v_{T-t'}) - r(v'_{T-t'})) \\
&= r(\zeta_{-t+1:}) - r(\zeta'_{-t+1:})
\end{aligned} \tag{11}$$

since for all $0 \leq t' < t$, $v_{T-t'}, v'_{T-t'} \in L_{t-t'}(v_0^*)$ and by the inductive hypothesis $H_{<t}$, it holds that, for all $0 \leq t' < t$, $\hat{r}(v_{T-t'}) - \hat{r}(v'_{T-t'}) = r(v_{T-t'}) - r(v'_{T-t'})$.

By definition, $Z_t \subseteq Z[v_0^*, d, T]$, and thus by weak identifiability, $\hat{r}(\zeta) - \hat{r}(\zeta') = r(\zeta) - r(\zeta')$. Combining with Eq. 9, 10, 11, we get that for all $v_T, v'_T \in \{v \mid \exists \zeta := (v_t)_{0\leq t\leq T} \in Z_t \text{ s.t } v = v_T\} = L_t(v_0^*)$,

$$\hat{r}(v_T) - \hat{r}(v'_T) = r(v_T) - r(v'_T)$$

Thus, by strong induction H_t is true for $0 \leq t \leq T_0$, which includes H_{T_0}.

• (Necessity) Next we prove necessity. To do so, we will first prove some useful properties of layer sequences.

Lemma 5. *Let G_d be strongly connected. Then, for all $v, v' \in V_d$, there exists $t \geq 1$ such that $v' \in L_t(v)$.*

Proof. Pick any $v, v' \in V_d$. Since G_d is strongly connected, there exists a path ζ of length $|\zeta| \geq 1$ between v, v'. Thus $v' \in L_{|\zeta|}(v)$. □

Lemma 6. *Let G_d be strongly connected. Then for all $v, v' \in V_d$ and $T \geq 0$, there exists $t \geq T$ such that $v' \in L_t(v)$.*

Proof. If $v' \in L_T(v)$, then we are done. If $v' \notin L_T(v)$, then choose any vertex $v_T \in L_T(v)$. There exists a path $\zeta^{v\to v_T}$ that starts from v and reaches v_T. Since G_d is strongly connected there exists a path $\zeta^{v_T \to v'}$ that starts from v_T and reaches v'. Thus $\zeta_{v\to v'} = \zeta^{v\to v_T}_{:-1} \oplus \zeta^{v_T\to v'}$ is a path that starts from v and reaches v' in $|\zeta_{v\to v'}| \geq T$ steps and $v' \in L_{|\zeta_{v\to v'}|}(v)$. □

Lemma 7. *Let G_d be strongly connected. Let $T_v \geq 1$ denote the smallest positive horizon such that $v \in L_{T_v}(v)$. Then, for all $v \in V_d$, the sequence $(L_{nT_v}(v))_{n\geq 0}$ converges to a limiting layer $\bar{L}(v) \subseteq V_d$, i.e, for all $v \in V_d$, there exists $\bar{n}_v \geq 0$ such that, for all $n \geq \bar{n}_v$, $L_{nT_v}(v) = \bar{L}(v)$.*

Proof. Since G_d is connected, v must be able to reach itself again and so there indeed exists a $T_v \geq 1$ such that $v \in L_{T_v}(v)$.

We first show that $(L_{nT_v}(v))_{n\geq 0}$ is "growing", i.e $L_{nT_v}(v) \subseteq L_{(n+1)T_v}(v)$ for all $n \geq 0$ by induction. The base case when $n = 0$ holds trivially by how we've defined $L_{T_v}(v)$ since $L_0(v) = \{v\} \subseteq L_{T_v}(v)$. Now assume for induction that $L_{nT_v}(v) \subseteq L_{(n+1)T_v}(v)$. Then,

$$L_{(n+1)T_v}(v) = L_{T_v}(L_{nT_v}(v)) \subseteq L_{T_v}(L_{(n+1)T_v}(v)) = L_{(n+2)T_v}(v)$$

by Lemma 2, monotonicity.

We now see that sequence $\{L_{nT_v}(v)\}_n$ is growing and bounded above, i.e $L_{nT_v}(v) \subseteq L_{(n+1)T_v}(v)$ and $L_{nT_v}(v) \subseteq V_d$ for all $n \geq 0$. Thus the sequence must converge to some fixed set $\bar{L}(v) \subseteq V_d$, i.e there exists $\bar{n}_v \geq 0$ such that $L_{nT_v}(v) = \bar{L}(v)$ for all $n \geq \bar{n}_v$. □

Lemma 8. *Let G_d be connected. Then, for all $v \in V_d$, the sequence $\{L_t(v)\}_{t\geq 0}$ is **eventually periodic**, i.e, for all $v \in V_d$, there exist $\bar{T}_v \geq 0, \delta_v \geq 1$ such that, for all $t \geq \bar{T}_v$, $L_t(v) = L_{t+\delta_v}(v)$.*

Proof. By Lemma 7, since G_d is connected, for all $v \in V_d$, $(L_{nT_v}(v))_{n\geq 0}$ converges to a limiting layer $\bar{L}(v)$ i.e, for all $v \in V_d$, there exists $\bar{n}_v \geq 0$ such that, for all $n \geq \bar{n}_v$, $L_{nT_v}(v) = \bar{L}(v)$.

Set $\bar{T}_v = \bar{n}_v T_v$ and $\delta_v = T_v$. Then we see that for all $t \geq \bar{T}_v = \bar{n}_v T_v$, it holds that

$$\begin{aligned}
L_{t+\delta_v}(v) &= L_{(t-\bar{n}_v T_v)+\bar{n}_v T_v + T_v}(v) \\
&= L_{(t-\bar{n}_v T_v)+(\bar{n}_v+1)T_v}(v) && (12) \\
&= L_{t-\bar{n}_v T_v}(L_{(\bar{n}_v+1)T_v}(v)) && (13) \\
&= L_{t-\bar{n}_v T_v}(L_{\bar{n}_v T_v}(v)) && (14) \\
&= L_{t-\bar{n}_v T_v + \bar{n}_v T_v}(v) \\
&= L_t(v)
\end{aligned}$$

where 12 → 13 holds since $(\bar{n}_v + 1)T_v \geq 0$ and $t - \bar{n}_v T_v \geq 0$. Furthermore, 13 → 14 holds by Lemma 7 since $L_{(\bar{n}_v+1)T_v}(v) = \bar{L}(v) = L_{\bar{n}_v T_v}(v)$. □

In words, Lemma 8 states that the layers induced by starting at any vertex always converge to a periodic sequence.

Definition 8. *Let $(a_t)_{t\geq 0}$ be a sequence. We say that a sequence $(b_t)_{t\geq 0}$ is a **tail** of the sequence $(a_t)_{t\geq 0}$ if and only if there exists an index $N \geq 0$ such that $b_t = a_{t+N}$. Let $(a_t)_{t\geq 0}$ be an eventually periodic sequence. We say that a sequence $(b_t)_{t\geq 0}$ is a **periodic tail** of the sequence $(a_t)_{t\geq 0}$ if and only if $(b_t)_{t\geq 0}$ is a periodic sequence and a tail of $(a_t)_{t\geq 0}$.*

We now prove some characteristics of the periodic tail.

Lemma 9. *Let G_d be strongly connected. Let us denote $\bar{L}_t(v) := L_t(\bar{L}(v))$. Then, the sequence $(\bar{L}_t(v))_{t\geq 0}$ is a periodic tail of the sequence $\{L_t(v)\}_{t\geq 0}$.*

Proof. From Lemma 7, $(L_{nT_v}(v))_{n\geq 0}$ converges to $\bar{L}_0$, so there exists $\bar{n}_v$ such that $L_{\bar{n}_v T_v}(v) = \bar{L}_0(v)$. Therefore, $L_{t+\bar{n}_v T_v}(v) = \bar{L}_t(v)$ and $(\bar{L}_t(v))_{t\geq 0}$ is a tail of the sequence $\{L_t(v)\}_{t\geq 0}$. It is left to show that $(\bar{L}_t(v))_{t\geq 0}$ is periodic. $(\bar{L}_0(v) = \bar{L}_{T_v}(v)) \Rightarrow (\forall t \geq 0, \bar{L}_t(v) = \bar{L}_{T_v+t}(v))$, therefore $(\bar{L}_t(v))_{t\geq 0}$ is periodic. □

Lemma 10. *Let G_d be strongly connected. Let $T_v \geq 1$ denote the smallest horizon $t \geq 1$ such that $v \in L_t(v)$. Let $\delta_v \geq 1$ denote the period of the tail sequence $(\bar{L}_t(v))_{t\geq 0}$ so that $\bar{L}_t(v) = \bar{L}_{t'}(v)$ for $0 \leq t < t'$ if and only if $(t' - t) \bmod \delta_v = 0$. Then $T_v \bmod \delta_v = 0$.*

Proof. We first know that $T_v \geq \delta_v$ trivially holds since $\bar{L}_0(v) = \bar{L}_{T_v}(v)$. Since $T_v \geq \delta_v > 0$ are integers, T_v admits a unique quotient $q \geq 1$ and remainder $m \geq 0$ by Euclid's lemma, i.e $T_v = q\delta_v + m$. Assume for contradiction that $m > 0$. Then, $q\delta_v < T_v$ and $m < \delta_v$. But then we have $\bar{L}_{q\delta_v}(v) = \bar{L}_{T_v}(v)$ and $T_v - q\delta_v = m < \delta_v$ and there does not exist an integer $n > 0$ such that $T_v - q\delta_v = n\delta_v$ which is a contradiction. Thus it must be that $m = 0$ as desired. □

Lemma 11. *Let G_d be strongly connected. Let $\delta_v \geq 1$ denote the period of the tail sequence $(\bar{L}_t(v))_{t\geq 0}$ so that $\bar{L}_t(v) = \bar{L}_{t'}(v)$ for $0 \leq t < t'$ if and only if $(t'-t) \bmod \delta_v = 0$. Then, for all $v \in V_d$ and $0 \leq t < t'$ such that $t'-t \bmod \delta_v \neq 0$, $\bar{L}_t(v) \cap \bar{L}_{t'}(v) = \emptyset$, i.e limiting layers within a period are all disjoint sets regardless of the starting vertex. Equivalently, for all $v \in V_d$ and $0 \leq t < t'$, $\bar{L}_t(v) = \bar{L}_{t'}(v)$ if $t'-t \bmod \delta_v = 0$ and $\bar{L}_t(v) \cap \bar{L}_{t'}(v) = \emptyset$ otherwise.*

Proof. Since $(\bar{L}_t(v))_{t\geq 0}$ is periodic with period δ_v, it suffices to prove that for all $v \in V_d$ and $0 \leq t < t' \leq \delta_v$ such that $t'-t < \delta_v$, $\bar{L}_t(v) \cap \bar{L}_{t'}(v) = \emptyset$. We first prove the following claim:

Claim 1. *For all $v \in V_d$ and $t \geq 0$, if $t \bmod \delta_v \neq 0$, then $v \notin \bar{L}_t(v)$.*

Proof. Again, due to periodicity, it suffices to prove that for all $v \in V_d$ and $0 < t < \delta_v$, $v \notin \bar{L}_t(v)$. Assume for contradiction that there exist $v \in V_d$ and $0 < t < \delta_v$ such that $v \in \bar{L}_t(v)$.

• We then claim that $\bar{L}_0(v) \subseteq \bar{L}_t(v)$. Assume, again, for contradiction that $\bar{L}_0(v) \subsetneq \bar{L}_t(v)$. Let $T_v \geq 1$ denote the smallest horizon $t \geq 1$ such that $v \in L_t(v)$. Since G_d is connected, by Lemma 10, $T_v = q\delta_v$ for some quotient integer $q \geq 1$. Then for all $n \geq 0$

$$L_{nT_v}(v) \subseteq L_{nT_v}(\bar{L}_t(v)) = L_{nT_v}(L_t(\bar{L}(v))) = L_{t+nT_v}(\bar{L}(v)) = \bar{L}_{t+nT_v}(v) = \bar{L}_{t+nq\delta_v}(v) \tag{15}$$

where the inclusion relation holds by monotonicity since $v \in \bar{L}_t(v)$ by outer assumption and the second equality holds by commutativity. (Lemma 2)

Since G_d is connected, by Lemma 7, there exists $\bar{n}_v \geq 0$ such that, for all $n \geq \bar{n}_v$, $L_{nT_v}(v) = \bar{L}(v) = \bar{L}_0(v)$. Combining this result with Eq. 15, there exists $\bar{n}_v \geq 0$ such that, for all $n \geq \bar{n}_v$

$$L_{nT_v}(v) = \bar{L}_0(v) \subseteq \bar{L}_{t+nq\delta_v}(v)$$

Then, since $\bar{L}_0(v) \subsetneq \bar{L}_t(v)$, there exists $\bar{n}_v \geq 0$ such that $\bar{L}_t(v) \neq \bar{L}_{t+nq\delta_v}(v)$ for $n \geq n_v$ which contradicts the assumption that $(\bar{L}_t(v))_{t\geq 0}$ is periodic with period δ_v. Thus, by contradiction, we have shown $\bar{L}_0(v) \subseteq \bar{L}_t(v)$.

• Now we enumerate all cases for $\bar{L}_t(v)$ that satisfy $\bar{L}_0(v) \subseteq \bar{L}_t(v)$.

If $\bar{L}_0(v) = \bar{L}_t(v)$, then this contradicts the assumption that $(\bar{L}_t(v))_{t\geq 0}$ is periodic with period δ_v

If $\bar{L}_0(v) \subset \bar{L}_t(v)$, then for all $n \geq 1$,

$$\bar{L}_{nt}(v) = L_{nt}(\bar{L}(v)) = L_{nt}(\bar{L}_0(v)) \subseteq L_{nt}(\bar{L}_t(v)) = L_{nt}(L_t(\bar{L}(v))) = L_{(n+1)t}(\bar{L}(v)) = \bar{L}_{(n+1)t}(v)$$

where the inclusion relation holds by monotonicity since we've just assumed $\bar{L}_0(v) \subset \bar{L}_t(v)$ and the fourth equality holds by commutativity. (Lemma 2) By transitivity this implies that for all $1 \leq n \leq n'$,

$$\bar{L}_{nt}(v) \subseteq \bar{L}_{n't}(v)$$

Choosing $n = 1$ and $n' = \delta_v$ we have $\bar{L}_0(v) \subset \bar{L}_t(v) \subseteq \bar{L}_{\delta_v t}(v)$ and so $\bar{L}_0(v) \neq \bar{L}_{\delta_v t}(v)$. Since $t > 0$ this again contradicts the periodicity of $(\bar{L}_t(v))_{t\geq 0}$. Thus we have shown, by contradiction, for all $v \in V_d$ and $0 < t < \delta_v$, $v \notin \bar{L}_t(v)$. □

Now to prove the original lemma, assume for contradiction that there exists $0 \leq t < t' \leq \delta_v$ such that $0 < t'-t < \delta_v$, and a shared vertex $v_{t,t'} \in V_d$ such that $v_{t,t'} \in \bar{L}_t(v)$ and $v_{t,t'} \in \bar{L}_{t'}(v)$. Since G_d is strongly connected $v_{t,t'}$ can reach v and so there exists a l such that $v \in \bar{L}_{t+l}(v)$ and $v \in \bar{L}_{t'+l}(v)$ by trivial extension of Lemma 6. We now enumerate all cases for the value of $t+l$.

If $t+l \bmod \delta_v \neq 0$, this contradicts Claim 1 since $v \in \bar{L}_{t+l}(v)$.

If $t+l \bmod \delta_v = 0$, then $t'+l \bmod \delta_v \neq 0$ since $(t'+l)-(t+l) = t'-t < \delta_v$. this contradicts Claim 1 since $v \in \bar{L}_{t'+l}(v)$. □

Lemma 12. *Let G_d be strongly connected. Then, for all $v, v' \in V_d$, the sequence $(\bar{L}_t(v))_{t\geq 0}$ is a periodic tail of the sequence $(L_t(v'))_{t\geq 0}$ i.e vertex layers all converge to the same periodic sequence regardless of the starting vertex.*

Proof. Pick any $v, v' \in V_d$ and consider their corresponding periodic tails $(\bar{L}_t(v))_{t\geq 0}, (\bar{L}_t(v'))_{t\geq 0}$. (which exists by Lemma 8) Without loss of generality, we will let the first layer of the periodic tails be those containing the initial vertex, i.e $v \in \bar{L}_0(v), v' \in \bar{L}_0(v')$. Such layers exist in the periodic tail by Lemma 6.

Let $t_v, t_{v'} \geq 0$ denote the horizons at which $v \in \bar{L}_{t_v}(v'), v' \in \bar{L}_{t_{v'}}(v)$. Again, such layers exist by Lemma 6. Then, we claim $\bar{L}_0(v) \subseteq \bar{L}_{t_v}(v')$. To see this, first note that the sequence $(L_{nT_v}(v))_{n\geq 0}$, where $T_v \geq 1$ is the shortest time horizon at which $v \in L_{T_v}(v)$, converges to $\bar{L}_0(v)$ by Lemma 7. Furthermore, $(\bar{L}_{t_v+nT_v}(v'))_{n\geq 0} = (\bar{L}_{t_v}(v'))_{n\geq 0}$ since $(v \in \bar{L}_{t_v}(v'), \bar{L}_{t_v+nT_v}(v')) \Rightarrow (\bar{L}_{t_v+nT_v}(v') = \bar{L}_{t_v}(v'))$ by Lemma 11. Since $\{v\} \subseteq \bar{L}_{t_v}(v')$, it follows from monotonicity (Lemma 2) that $L_{nT_v}(v) \subseteq L_{nT_v}(\bar{L}_{t_v}(v')) = \bar{L}_{t_v+nT_v}(v') = \bar{L}_{t_v}(v')$ for all $n \geq 0$. Since there exists an $\bar{n}_v \geq 0$ such that $L_{\bar{n}_v T_v}(v) = \bar{L}_0(v)$, we thus have $\bar{L}_0(v) \subseteq \bar{L}_{t_v}(v')$. The same argument can be applied to obtain $\bar{L}_0(v') \subseteq \bar{L}_{t_{v'}}(v)$.

We now consider two different cases. If $\bar{L}_0(v') = \bar{L}_0(v)$, then it trivially follows that the sequences $(\bar{L}_t(v))_{t\geq 0}, (\bar{L}_t(v'))_{t\geq 0}$ are the same. For the second case if $\bar{L}_0(v') \neq \bar{L}_0(v)$, then $\bar{L}_t(v) = L_t(\bar{L}_0(v)) \subseteq L_t(\bar{L}_{t_v}(v')) = \bar{L}_{t_v+t}(v')$ for all $t \geq 0$ and $\bar{L}_t(v') = L_t(\bar{L}_0(v')) \subseteq L_t(\bar{L}_{t_{v'}}(v)) = \bar{L}_{t_{v'}+t}(v)$ for all $t \geq 0$. Thus $\bar{L}_0(v) \subseteq \bar{L}_{t_v}(v') \subseteq \bar{L}_{t_v+t_{v'}}(v)$. Then, $v \in \bar{L}_0(v) \Rightarrow v \in \bar{L}_{t_v+t_{v'}}(v)$ and it follows that $\bar{L}_0(v) = \bar{L}_{t_v+t_{v'}}(v)$ by Lemma 11. Thus, $\bar{L}_0(v) = \bar{L}_{t_v}(v')$ which implies that $\bar{L}_t(v) = \bar{L}_{t_v+t}(v')$ for all $t \geq 0$ and so $(\bar{L}_t(v))_{t\geq 0}$ is a tail of $(\bar{L}_t(v'))_{t\geq 0}$. ☐

From Lemma 12, we see that the layer sequence converges to the same periodic tail sequence regardless of the starting vertex. Thus, we shall henceforth denote a periodic tail of G_d as $(\bar{L}_t)_{t\geq 0}$, dropping the dependence on intial vertex.

Lemma 13. *Let G_d be strongly connected and let $(\bar{L}_t)_{t\geq 0}$ be a periodic tail of the layer sequences in G_d. For all $v \in V_d$ and $t, t' \geq 0$, $(L_t(v) \cap \bar{L}_{t'} \neq \emptyset) \Rightarrow (L_t(v) \subseteq \bar{L}_{t'})$*

Proof. Suppose for contradiction that there exists $v \in V_d$ and $t, t' \geq 0$ such that $(L_t(v) \cap \bar{L}_{t'} \neq \emptyset)$, but $(L_t(v) \not\subseteq \bar{L}_{t'})$. Let $v^- \in L_t(v) - \bar{L}_{t'}$ and $v^\cap \in L_t(v) \cap \bar{L}_{t'}$.

Let $T_v \geq 1$ denote the smallest positive horizon such that $v \in L_{T_v}(v)$. Then, $L_{nT_v+t}(v) = L_t(L_{nT_v}(v)) \subseteq L_t(L_{(n+1)T_v}(v)) = L_{(n+1)T_v+t}(v)$ for all $n \geq 0$ by Lemma 2 since $L_{nT_v}(v) \subseteq L_{(n+1)T_v}(v)$ from the proof of Lemma 7. Thus, the sequence $(L_{nT_v+t}(v))_{n\geq 0}$ must converge to some fixed set $\bar{L}_{t^*}$ since the sequence is growing and bounded above, i.e $L_{nT_v}(v) \subseteq L_{(n+1)T_v}(v)$ and $L_{nT_v}(v) \subseteq V_d$ for all $n \geq 0$. Thus, $\bar{L}_{t^*}$ is an element of the tail $(\bar{L}_t)_{t\geq 0}$. Since $v^-, v^\cap \in L_t(v)$, we have $v^-, v^\cap \in \bar{L}_{t^*}$. This contradicts Lemma 11 since $\bar{L}_{t'}, \bar{L}_{t^*}$ are two tail layers that are not the same but also not disjoint. ☐

We now prove the necessary direction of the main theorem. We show the contrapositive, i.e if either $\mathcal{P}_{\text{MDP}}[R; d, T, J]$ is not weakly identifiable or not coverable, it is not strongly identifiable. By Proposition 1, $\mathcal{P}_{\text{MDP}}[R; d, T, J]$ must be weakly identifiable to be strongly identifiable. Thus, consider $\mathcal{P}_{\text{MDP}}[R; d, T, J]$ that is weakly identifiable but not coverable. By Proposition 2 it suffices to show that $\exists r, \hat{r} \in R$ such that $r \not\cong_{x,a} \hat{r}$ but $r \cong_\tau \hat{r}$.

Let $(\bar{L}_t)_{t\geq 0}$ be a periodic tail of the layer sequences in G_d. Let $r, \hat{r}$ be two rewards such that $\forall v \notin \bar{L}_0, \hat{r}(v) = r(v)$ and $\forall v \in \bar{L}_0, \hat{r}(v) = r(v) + c$ for some constant $c \in \mathbb{R}$. Since there does not exist a covering initial state, clearly, $\bar{L}_0 \subset V_d$ and thus $r \not\cong_{x,a} \hat{r}$. We will show that $r \cong_\tau \hat{r}$ to conclude that the $\mathcal{P}_{\text{MDP}}[R; d, T, J]$ is not strongly identifiable.

For all $v \in V_d^0$ and for all paths $\zeta = (v_t)_{0\leq t\leq T}, \zeta' = (v'_t)_{0\leq t\leq T}$ such that $\zeta, \zeta' \in Z[v, d, T]$, we claim that $\hat{r}(v_t) - \hat{r}(v'_t) = r(v_t) - r(v'_t)$ for all $0 \leq t \leq T$. To see this, first note that $v_t, v'_t \in L_t(v)$ for all $t \geq 0$. We consider two cases: (1). If $v_t \in \bar{L}_0$, then $v'_t \in \bar{L}_0$ since $v_t, v'_t \in L_t(v)$ and, by Lemma 13, $(L_t(v) \cap \bar{L}_{t'} \neq \emptyset) \Rightarrow (L_t(v) \subseteq \bar{L}_{t'})$. Thus, $\hat{r}(v_t) - \hat{r}(v'_t) = r(v_t) + c - r(v'_t) - c = r(v_t) - r(v'_t)$. (2) If $v_t \notin \bar{L}_0$, then $v'_t \notin \bar{L}_0$ since $v_t, v'_t \in L_t(v)$ and, by the contrapositive of Lemma 13, $(L_t(v) \not\subseteq \bar{L}_0) \Rightarrow (L_t(v) \cup \bar{L}_0 = \emptyset)$. Thus, $\hat{r}(v_t) - \hat{r}(v'_t) = r(v_t) - r(v'_t)$

Then,

$$\begin{aligned} r(\zeta') - r(\zeta) &= \sum_{t=0}^{T} \gamma^t (r(v'_t) - r(v_t)) \\ &= \sum_{t=0}^{T} \gamma^t (\hat{r}(v'_t) - \hat{r}(v_t)) \\ &= \hat{r}(\zeta') - \hat{r}(\zeta) \end{aligned}$$

Therefore, $r, \hat{r}$ are two trajectory equivalent rewards which are not state-action equivalent. Hence $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is not strongly identifiable. □

Corollary 2. (Strong Identification Condition) *For all (d, r, T, J) such that $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is a proper MDP model and G_d is strongly connected,*

- *(Sufficiency) $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is weakly identifiable, G_d aperiodic $\Rightarrow \exists T_0 \geq 0$ such that $\forall T \geq T_0, \mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is strongly identifiable*
- *(Necessity) $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is strongly identifiable $\Rightarrow \mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is weakly identifiable, G_d is aperiodic.*

Proof. Let $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ be proper and G_d be strongly connected. • (Sufficiency) Since G_d is strongly connected and aperiodic, it is covering by Proposition 3, i.e there exists an initial vertex $v_0 \in V_d^0$ that is t^*-covering for some t^*. Let $T_0 = 2t^*$, $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is strongly identifiable for all $T \geq T_0$ by Theorem 2.

• (Necessity) If $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is proper, strongly identifiable, and G_d is strongly connected, by Theorem 3, it is weakly identifiable and G_d is coverable. Since, G_d is strongly connected and coverable, it is aperiodic by Proposition 3. □

A.5. Strong Identifiability Test Algorithms

Theorem 3. *Let $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ be a weakly identifiable MDP model and G_d be strongly connected. Then,*

- *(Correctness)* `MDPIdTest`*($\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$) returns 1 (True) if and only if $\exists T$ such that $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is strongly identifiable.*
- *(Efficiency)* `MDPIdTest` *runs with time and space complexity $O(|E_d|)$*

Proof. • (Correctness) `MDPIdTest` returns 1 (True) if and only if the directed graph G_d is aperiodic as shown in (Denardo, 1977; Jarvis and Shier, 1999). Since $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is weakly identifiable and G_d is strongly connected, G_d is aperiodic if and only if $\exists T$ such that $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is strongly identifiable by Corollary 2.

• (Efficiency) Graph aperiodicity testing can be done in $O(|E_d|)$ space and time as shown in (Denardo, 1977; Jarvis and Shier, 1999). □

Corollary 3. (Strong Identification Condition) *For all (d, r, T, J) such that the MDP model $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is proper.*

- *(Sufficiency) $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is weakly identifiable, G_d is T_0-coverable, and $T \geq 2T_0 \Rightarrow \mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is strongly identifiable*

Proof. This result immediately follows from the proof of the Sufficiency direction for Theorem 2. □

Theorem 4. *Let $\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ be a weakly identifiable MDP model. Then,*

- *(Correctness) If* `MDPCoverTest`*($\mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$) returns 1 (True) then, $\exists T_0$ such that $\forall T \geq T_0, \mathcal{P}_{\mathrm{MDP}}[R; d, T, J]$ is strongly identifiable.*
- *(Efficiency)* `MDPCoverTest` *runs with time complexity $O(|V_d|^3 \log |V_d|)$ and space complexity $O(|V_d|^2)$*

Proof. • Since M is the transition matrix, i.e $M_{ij} = \tilde{P}(v^{(j)}|v^{(i)})$ where $\tilde{P}(x', a'|x, a) = P(x'|x, a)$, it is clear that $M_{ij}^{|V_d|^2} \neq 0$ if and only if $v^{(j)} \in L_{|V_d|^2}(v^{(i)})$. If `MDPCoverTest` returns 1 (True), then there exists $v^{(i)} \in V_d^0$ that has a fully non-zero row $M_i^{|V_d|^2}$, i.e $L_{|V_d|^2}(v^{(i)}) = V_d$. Thus, G_d is $|V_d|^2$-coverable by $v^{(i)}$. Let $T_0 = 2|V_d|^2$ and the result follows from Corollary 3. Therefore, `MDPCoverTest` returns 1 (True) if and only if

• (Efficiency) It is well known that computing matrix powers A^m (where the matrix A has size $n \times n$) can be done in $O(n^3 \log m)$ time and $O(n^2)$ space (Cormen et al., 2009). Since M has size $|V_d| \times |V_d|$, computing $M^{|V_d|^2}$ has time complexity $O(|V_d|^3 \log |V_d|^2) = O(|V_d|^3 \log |V_d|)$ and space complexity $O(|V_d|^2)$. A naive approach to checking for rows with only non-zero entries requires enumerating over all elements of M which can be done in $O(|V_d|^2)$ time and $O(1)$ space, thus not affecting the overall efficiency of the algorithm.

$\square$