SUPPLEMENT TO “CHARACTERIZATION OF CAUSAL ANCESTRAL GRAPHS FOR TIME SERIES WITH LATENT CONFOUNDERS”

BY ANDREAS GERHARDUS[1,a]

[1] *Institute of Data Science, German Aerospace Center,* [a]*andreas.gerhardus@dlr.de*

This Supplementary Material contains: First, a glossary of abbreviations and frequently used symbols. Second, theoretical results that were omitted from the main text due to space constraints. Third, proofs of all theoretical results presented in the main text together with various auxiliary results that are used in these proofs.

A. Glossary of abbreviations and frequently used symbols. The following glossary might be helpful for reading the main paper and this Supplementary Material.

Term / symbol	**Meaning**	**Comment**
DAG	directed acyclic graph	
MAG / DMAG	(directed) maximal ancestral graph	
PAG / DPAG	(directed) partial ancestral graph	
ts-DAG	time series DAG	see Def. 3.4
ts-DMAG / ts-DPAG	time series DMAG / DPAG	see Defs. 3.6 and 5.7
$\mathcal{D}$	DAG or ts-DAG	
$\mathcal{M}$	DMAG or ts-DMAG	
$\mathcal{P}$	partial ancestral graph or DPAG or ts-DPAG	
$\mathbf{O}$	set of all observed vertices in a graph	
$\mathcal{M}_{\mathbf{O}}(\mathcal{D})$	MAG latent projection of the DAG or ts-DAG $\mathcal{D}$ to the vertices $\mathbf{O}$	see pp. 1442-1443 in Zhang (2008a)
$\mathbf{I_O}$	variable indices of observed component time series	
$\mathbf{T_O}$	time indices of observed time steps	
$\mathcal{M}_{\mathbf{I_O}\times\mathbf{T_O}}(\mathcal{D})$	ts-DMAG of the ts-DAG $\mathcal{D}$ with observed vertices $\mathbf{O} = \mathbf{I_O} \times \mathbf{T_O}$	see Def. 3.6
$\mathcal{M}_{\mathbf{O}}$	synonymous to $\mathcal{M}_{\mathbf{I_O}\times\mathbf{T_O}}(\mathcal{D})$ with $\mathbf{O} = \mathbf{I_O} \times \mathbf{T_O}$	see Def. 3.6
p	length of the observed time window, for regular sampling related to $\mathbf{T_O}$ by $\mathbf{T_O} = \{t-\tau \mid 0 \leq \tau \leq p\}$	
$\mathcal{M}^p(\mathcal{D})$	synonymous to $\mathcal{M}_{\mathbf{I_O}\times\mathbf{T_O}}(\mathcal{D})$ with $\mathbf{T_O} = \{t-\tau \mid 0 \leq \tau \leq p\}$	see Sec. 4.2
$\mathrm{stat}(\cdot)$	stationarification, can be applied to graphs with time series structure	see Def. 4.6
$\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$	stationarified ts-DMAG, synonymous to $\mathrm{stat}(\mathcal{M}^p(\mathcal{D}))$	
$\mathcal{D}_c(\cdot)$	canonical DAG or canonical ts-DAG	see Defs. 4.11 and 4.13
$\mathcal{P}(\cdot, \mathcal{A})$	m.i. DPAG wrt. to background knowledge $\mathcal{A}$	see Def. 5.2
$\mathcal{A}_{\mathcal{D}}$	background knowledge of an underlying ts-DAG	see Def. 5.4
$\mathcal{A}^{\mathrm{stat}}_{\mathcal{D}}$	background knowledge of an underlying ts-DAG for stationarifications	see Def. 5.4
$\mathcal{A}_{\mathrm{ta}}$	background knowledge of time order and repeating ancestral relationships	see Def. 5.4
$\mathcal{A}_{\mathrm{to}}$	background knowledge of time order and repeating orientations	see Def. 5.4
$\mathcal{P}^p(\mathcal{D})$	time series DPAG, synonymous to $\mathcal{P}^p(\mathcal{D}, \mathcal{A}_{\mathcal{D}})$	see Def. 5.7

TABLE 1
Glossary of abbreviations and frequently used symbols.

B. Omitted results. This section presents several theoretical results that were omitted from the main text due to space constraints.

B.1. *ts-DMAGs are a generalization of DMAGs.* Consider an arbitrary DMAG $\mathcal{M}_{\mathrm{nt}}$ with vertex set $\mathbf{O}_{\mathrm{nt}}$ and without time series structure (the subscript "nt" stands for "*n*on-*t*emporal"). As proven in Richardson and Spirtes (2002), there is DAG $\mathcal{D}_{\mathrm{nt}}$ over some vertex set $\mathbf{V}_{\mathrm{nt}} \supseteq \mathbf{O}_{\mathrm{nt}}$ such that $\mathcal{M}_{\mathrm{nt}} = \mathcal{M}_{\mathbf{O}_{\mathrm{nt}}}(\mathcal{D}_{\mathrm{nt}})$. Now define $\mathcal{D}$ as the ts-DAG that consists of disconnected copies of $\mathcal{D}_{\mathrm{nt}}$ at every time step $s \in \mathbb{Z}$, i.e., there are no lagged edges in $\mathcal{D}$ and for all $s \in \mathbb{Z}$ its induced subgraph on $\mathbf{V}_{\mathrm{nt}} \times \{s\}$ is $\mathcal{D}_{\mathrm{nt}}$. It then immediately follows that the ts-DMAG $\mathcal{M}_{\mathbf{I_O} \times \mathbf{T_O}}$ with $\mathbf{I_O} = \mathbf{V}_{\mathrm{nt}}$ and $\mathbf{T_O} = \{t\}$ equals $\mathcal{M}_{\mathrm{nt}}$.

This consideration identifies ts-DMAGs as a proper generalization of DMAGs and thereby shows that all statements about ts-DMAGs also apply to DMAGs as a special case.

B.2. *Future vertices are not relevant for determining ts-DMAGs.* In the MAG latent projection of $\mathcal{D}$ to $\mathcal{M}_{\mathbf{I_O} \times \mathbf{T_O}}(\mathcal{D})$ the vertices $\mathbf{L} = \mathbf{V} \setminus (\mathbf{I_O} \times \mathbf{T_O})$ are unobserved, see Definition 4.6 in the main text. Since t is the upper bound of the set of observed time steps $\mathbf{T_O}$, this form of $\mathbf{L}$ means that in particular all vertices after t, i.e., in $[t+1, +\infty)$ are unobserved. However, for determining $\mathcal{M}_{\mathbf{I_O} \times \mathbf{T_O}}(\mathcal{D})$ these vertices are irrelevant:

LEMMA B.1. *Let $\mathcal{D}$ be a ts-DAG with vertex set $\mathbf{V} = \mathbf{I} \times \mathbb{Z}$. Denote with $\mathcal{D}_{\leq t}$ the subgraph of $\mathcal{D}$ induced on $\mathbf{V}_{\leq t} = \mathbf{I} \times \mathbf{T}_{\leq t}$ with $\mathbf{T}_{\leq t} = \{s \in \mathbb{Z} \mid s \leq t\}$, i.e., the graph obtained by removing all vertices after t and all edges involving these vertices from $\mathcal{D}$. Then, $\mathcal{M}_{\mathbf{I_O} \times \mathbf{T_O}}(\mathcal{D}) = \mathcal{M}_{\mathbf{I_O} \times \mathbf{T_O}}(\mathcal{D}_{\leq t})$, i.e., before applying the MAG latent projection one may simply ignore the part of $\mathcal{D}$ that is after t.*

PROOF OF LEMMA B.1. Let (i, t_i) and (j, t_j) with $t_i, t_j \leq t$ be distinct vertices in $\mathcal{D}$. Then, (i, t_i) and (j, t_j) are nonadjacent in $\mathcal{D}$ if and only if $(i, t_i) \perp\!\!\!\perp (j, t_j) \mid \mathbf{S}$ in $\mathcal{D}$ with $\mathbf{S} = \mathrm{pa}((i, t_i), \mathcal{D}) \cup \mathrm{pa}((j, t_j), \mathcal{D}) \setminus \{(i, t_i), (j, t_j)\}$, see Verma and Pearl (1990, Lemma 1). Moreover, all vertices in $\mathbf{S}$ are before or at t due to time order of $\mathcal{D}$, i.e., $\mathbf{S}$ is a subset of $\mathbf{V}_{\leq t} = \mathbf{I} \times \mathbf{T}_{\leq t}$. Consequently, (i, t_i) and (j, t_j) are nonadjacent in $\mathcal{D}$ if and only if there is a subset $\mathbf{S}' \subseteq \mathbf{V}_{\leq t}$ such that $(i, t_i) \perp\!\!\!\perp (j, t_j) \mid \mathbf{S}'$ in $\mathcal{D}$. This observation implies that the graphs $\mathcal{M}_{\mathbf{V}_{\leq t}}(\mathcal{D})$ and $\mathcal{D}_{\leq t}$ have the same skeleton, and the equality $\mathcal{M}_{\mathbf{V}_{\leq t}}(\mathcal{D}) = \mathcal{D}_{\leq t}$ follows because both $\mathcal{M}_{\mathbf{V}_{\leq t}}(\mathcal{D})$ and $\mathcal{D}_{\leq t}$ have the same ancestral relationships among vertices in $\mathbf{V}_{\leq t}$ as $\mathcal{D}$. From $\mathcal{M}_{\mathbf{V}_{\leq t}}(\mathcal{D}) = \mathcal{D}_{\leq t}$ the statement follows with the commutativity of the marginalization process as stated by Theorem 4.20 in Richardson and Spirtes (2002). □

This result, which follows from time order and d-separation, has the intuitive interpretation that the future need not be known in order to reason about the past and present.

B.3. *Temporal confounding.* As explained in Section 3.4 of the main text, in the construction of $\mathcal{M}_{\mathbf{I_O} \times \mathbf{T_O}}(\mathcal{D})$ all vertices before $t - p$, i.e., in $(-\infty, t - p - 1]$ are on purpose treated as unobserved—even if they are observable and hence become observed for some $\tilde{p} > p$. As the following example shows, such temporally unobserved observable vertices before $t - p$ can act as latent confounders of observed vertices.

EXAMPLE B.2. In the ts-DAG $\mathcal{D}_1$ shown in part (a) of Figure 2 of the main text, the temporally unobserved vertex O^1_{t-3} confounds the observed vertices O^1_{t-2} and O^2_{t-2} through the path $O^1_{t-2} \leftarrow O^1_{t-3} \rightarrow O^2_{t-2}$. This argument remains valid even without the unobservable time series L^1.

In the time series setting one thus effectively always deals with the case of latent confounding, even if all component time series are observable. This observation further demonstrates the importance of conceptually understanding the latent variable setting as approached in this paper.

We also note that it is precisely this type of confounding that gives rise to edges which are in $\mathcal{M}^p(\mathcal{D})$ but not in $\mathcal{M}^p_{\text{st}}(\mathcal{D})$.

B.4. *An axiomatic characterization of stationarifications.* As we have already noted below Definition 4.6 in the main text, the definition of stationarification implies the following result.

LEMMA B.3. $\operatorname{stat}(\mathcal{G})$ *is the unique largest subgraph of* $\mathcal{G}$ *that has repeating edges.*

PROOF OF LEMMA B.3. Combine both parts of Lemma B.4. □

Indeed, we could alternatively have defined stationarifications by this property and then derived that stationarifications fulfill the properties as given in Definition 4.6.

LEMMA B.4. 1. $\operatorname{stat}(\mathcal{G})$ *has repeating edges.*
2. *If* $\mathcal{G}'$ *is a subgraph of* $\mathcal{G}$ *and has repeating edges, then* $\mathcal{G}'$ *is a subgraph of* $\operatorname{stat}(\mathcal{G})$.

PROOF OF LEMMA B.4. **1.** Consider an edge $((i,t_i),(j,t_j)) \in \mathbf{E}_\bullet$ in $\operatorname{stat}(\mathcal{G})$ and let Δt be such that $(i,t_i+\Delta t),(j,t_j+\Delta t) \in \mathbf{V}$. Using the second point in Definition 4.6 twice, we first get $((i,t_i+\Delta t'),(j,t_j+\Delta t')) \in \mathbf{E}_\bullet$ in $\mathcal{G}$ for all $\Delta t'$ for which $(i,t_i+\Delta t'),(j,t_j+\Delta t') \in \mathbf{V}$ and thus $((i,t_i+\Delta t),(j,t_j+\Delta t)) \in \mathbf{E}_\bullet$ in $\operatorname{stat}(\mathcal{G})$.

2. Let $\mathcal{G}' \subseteq \mathcal{G}$ have repeating edges and assume $\mathcal{G}' \not\subseteq \operatorname{stat}(\mathcal{G})$. Since both $\mathcal{G}'$ and $\operatorname{stat}(\mathcal{G})$ are subgraphs of $\mathcal{G}$, adjacencies that are shared by $\mathcal{G}'$ and $\operatorname{stat}(\mathcal{G})$ correspond to edges of the same type. Thus, $\mathcal{G}' \not\subseteq \operatorname{stat}(\mathcal{G})$ implies that there is an adjacency in $\mathcal{G}'$ which is not in $\operatorname{stat}(\mathcal{G})$, i.e., $((i,t_i),(j,t_j)) \in \mathbf{E}$ in $\mathcal{G}'$ and $((i,t_i),(j,t_j)) \notin \mathbf{E}$ in $\operatorname{stat}(\mathcal{G})$. Since $\mathcal{G}'$ has repeating edges by assumption, $((i,t_i),(j,t_j)) \in \mathbf{E}$ in $\mathcal{G}'$ implies that $((i,t_i+\Delta t),(j,t_j+\Delta t)) \in \mathbf{E}$ in $\mathcal{G}$ for all Δt for which $(i,t_i+\Delta t),(j,t_j+\Delta t) \in \mathbf{V}$. But then the second point in Definition 4.6 gives $((i,t_i),(j,t_j)) \in \mathbf{E}$ in $\operatorname{stat}(\mathcal{G})$. Contradiction. □

B.5. *Why the case of no unobservable vertices remains special.* As discussed in Section B.3, even if there are no unobservable time series one in general still is in the setting of latent confounding. It is worth noting, though, that the case of no unobservable vertices remains special:

LEMMA B.5. *Let* $\mathcal{D}$ *be a ts-DAG with variable index set* $\mathbf{I}$*, let* $\mathbf{I_O} = \mathbf{I}$*, and let* $\mathbf{T_O} = \{t-\tau \mid 0 \leq \tau \leq p\}$ *where* $p \geq p_{\text{ts}}$ *with* p_{ts} *the largest lag in* $\mathcal{D}$*. Then,* $\operatorname{stat}(\mathcal{M}_{\mathbf{I_O}\times\mathbf{T_O}}(\mathcal{D}))$ *equals the subgraph of* $\mathcal{D}$ *induced on* $\mathbf{I_O} \times \mathbf{T_O}$*.*

REMARK (on Lemma B.5). The proof is given in Section D.3 below.

In other words: If all component time series are observable ($\mathbf{I_O} = \mathbf{I}$) and there are enough regularly sampled time steps to capture all direct causal influences (choice of $\mathbf{T_O}$ and $p \geq p_{\text{ts}}$), the stationarified ts-DMAG $\operatorname{stat}(\mathcal{M}_{\mathbf{I_O}\times\mathbf{T_O}}(\mathcal{D}))$ equals the segment of $\mathcal{D}$ on $\mathbf{T_O}$. For ts-DMAGs $\mathcal{M}_{\mathbf{I_O}\times\mathbf{T_O}}(\mathcal{D})$ the same is not necessarily true.

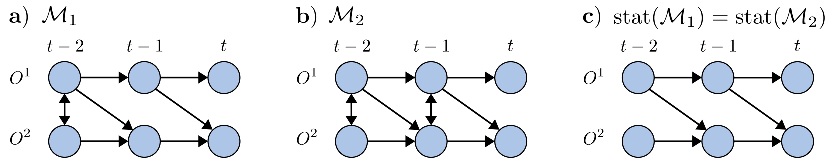

FIG A. *Two different DMAGs with time series structure that have the same stationarification.*

B.6. *Different DMAGs with the same stationarification cannot both be ts-DMAGs.* The following result is an immediate consequence of the one-to-one correspondence between a ts-DMAG and its stationarification (see Section 4.7 in the main text).

LEMMA B.6. *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be DMAGs with time series structure such that $\mathcal{M}_1 \neq \mathcal{M}_2$ and $\mathrm{stat}(\mathcal{M}_1) = \mathrm{stat}(\mathcal{M}_1)$. Then, at least one of $\mathcal{M}_1$ and $\mathcal{M}_2$ is not a ts-DMAG.*

REMARK (on Lemma B.6). The proof is given in Section D.7 below.

EXAMPLE B.7. The two DMAGs with time series structure in parts (a) and (b) of Figure A have the same stationarification (namely the graph in part (c) of the figure). Therefore, at most one of them can be a ts-DMAG. Indeed, using Theorem 1 (or Theorem 2) from the main text we confirm that $\mathcal{M}_2$ is not a ts-DMAG and that $\mathcal{M}_1$ is a ts-DMAG.

B.7. *Additional results on Example 5.8.* Here, we formalize and prove the following claim made in Example 5.8 in the main text.

LEMMA B.8. *Let $\mathcal{D}'$ be a ts-DAG such that its ts-DMAG $\mathcal{M}^1(\mathcal{D}')$ equals the ts-DMAG $\mathcal{M}^1(\mathcal{D})$ in part (a) of Figure 10 in the main text. Then, in $\mathcal{D}'$ the pair of observed vertices (O^1_{t-1}, O^1_t) is not subject to unobserved confounding, that is, in $\mathcal{D}'$ there is no inducing path (relative to the set of observed vertices) between O^1_{t-1} and O^1_t that is into O^1_{t-1}. Consequently, from $\mathcal{M}^1(\mathcal{D})$ we can conclude that the causal effect of O^1_{t-1} on O^1_t is identifiable and can be estimated from observations by adjusting for the empty set. Moreover, since the ts-DPAG $\mathcal{P}^1(\mathcal{D})$ in part (c) of Figure 10 is equal to the ts-DMAG $\mathcal{M}^1(\mathcal{D})$, we can draw this conclusion not only from $\mathcal{M}^1(\mathcal{D})$ but also from $\mathcal{P}^1(\mathcal{D})$.*

PROOF OF LEMMA B.8. First, we prove that the pair (O^1_{t-1}, O^1_t) is not subject to unobserved confounding. To this end, we begin by deriving the existence of certain paths in $\mathcal{D}'$:

1. Since O^1_{t-1} is an ancestor of O^1_t in $\mathcal{D}'$ according to the edge $O^1_{t-1} {\rightarrow} O^1_t$ in $\mathcal{M}^1(\mathcal{D}')$, in $\mathcal{D}'$ there is a directed path π_1 from O^1_{t-1} to O^1_t. This path cannot intersect O^2_t because else O^2_t would be an ancestor of O^1_t by means of the subpath $\pi_1(O^2_t, O^1_t)$, which together with the fact that O^1_t is an ancestor of O^2_t according to the edge $O^1_t {\rightarrow} O^2_t$ in $\mathcal{M}^1(\mathcal{D}')$ contradicts acyclicity. The path π_1 can also not intersect O^2_{t-1} because the subpath $\pi_1(O^2_{t-1}, O^1_t)$ would then be a directed path from O^2_{t-1} to O^1_t such that all its non-end-point vertices, if any, are unobserved. Consequently, there would need to be the edge $O^2_{t-1} {\rightarrow} O^1_t$ in $\mathcal{M}^1(\mathcal{D}')$. Moreover, due to time order, π_1 can also not contain any vertex O^1_s or O^2_s with $s \leq t-2$. We conclude that all non-end-point vertices of π_1, if any, are unobservable.
2. Since O^1_{t-1} is an ancestor of O^2_{t-1} in $\mathcal{D}'$ according to the edge $O^1_{t-1} {\rightarrow} O^2_{t-1}$ in $\mathcal{M}^1(\mathcal{D}')$, in $\mathcal{D}'$ there is a directed path π_2 from O^1_{t-1} to O^2_{t-1}. This path can, due to time order, neither intersect O^1_t nor O^2_t. Moreover, also due to time order, π_2 cannot contain any vertex O^1_s or O^2_s with $s \leq t-2$. We conclude that all non-end-point vertices of π_2, if any, are unobservable.

3. Since O^2_{t-1} is an ancestor of O^2_t in $\mathcal{D}'$ according to the edge $O^2_{t-1} {\rightarrow} O^2_t$ in $\mathcal{M}^1(\mathcal{D}')$, in $\mathcal{D}'$ there is a directed path π_3 from O^2_{t-1} to O^2_t. This path cannot intersect O^1_{t-1} because else O^2_{t-1} would be an ancestor of O^1_{t-1} by means of the subpath $\pi_3(O^2_{t-1}, O^1_{t-1})$, which together with the fact that O^1_{t-1} is an ancestor of O^2_{t-1} according to the edge $O^1_{t-1} {\rightarrow} O^2_{t-1}$ in $\mathcal{M}^1(\mathcal{D}')$ contradicts acyclicity. The path π_3 can also not intersect O^1_t because the subpath $\pi_3(O^2_{t-1}, O^1_t)$ would then be a directed path from O^2_{t-1} to O^1_t such that all its non-end-point vertices, if any, are unobserved. Consequently, there would need to be the edge $O^2_{t-1} {\rightarrow} O^1_t$ in $\mathcal{M}^1(\mathcal{D}')$. Moreover, due to time, order π_3 can also not contain any vertex O^1_s or O^2_s with $s \leq t-2$. We conclude that all non-end-point vertices of π_3, if any, are unobservable.

□

For $i = 1, 2, 3$ let π'_i be a copy of π_i that is shifted backwards in time by one time step. These paths π'_i exist due to the repeating edges property of $\mathcal{D}'$. Then, the concatenation $\rho_1 = \pi'_1(O^1_{t-1}, O^1_{t-2}) \oplus \pi'_2 \oplus \pi'_3$ is a collider-free path between O^1_{t-1} and O^2_{t-1} that is into both O^1_{t-1} and O^2_{t-1} such that all its non-end-point vertices are unobserved. In particular, ρ_1 is an inducing path.

Now suppose that, contrary to the claim to be proven, in $\mathcal{D}'$ there is an inducing path ρ_2 between O^1_{t-1} and O^1_t that is into O^1_{t-1}. Then, according to Lemma 32 in Zhang (2008a) the concatenation $\rho_1 \oplus \rho_2$ has a subsequence ρ which is an inducing path between O^2_{t-1} and O^1_t in $\mathcal{D}'$. However, then there would need to be an edge between O^2_{t-1} and O^1_t in $\mathcal{M}^1(\mathcal{D}')$ (according to the ancestral relationships, this edge would be $O^2_{t-1} {\leftrightarrow} O^1_t$), which is a contradiction. We conclude that there is no inducing path between O^1_{t-1} and O^1_t which is into O^1_{t-1}, that is, the pair (O^1_{t-1}, O^1_t) is not subject to unobserved confounding.

Second, we prove that the causal effect of O^1_{t-1} and O^1_t is identifiable and can be estimated by adjusting for the empty set. To this end, assume that O^1_{t-1} and O^1_t are *not* d-separated in the graph $\mathcal{G}$ that is obtained by removing from $\mathcal{D}'$ all edges out of O^1_{t-1}. Then, there is at least one path π between O^1_{t-1} and O^1_t in $\mathcal{G}$ that is active given the empty set. This path

1. is into O^1_{t-1} because in $\mathcal{G}$ there are no edges out of O^1_{t-1},
2. is collider-free because π is active given the empty set,
3. is also a path in $\mathcal{D}'$ because $\mathcal{G}$ is a subgraph of $\mathcal{D}'$,
4. needs to intersect at least one of O^2_{t-1} and O^2_t because else it would be an inducing path between O^1_{t-1} and O^1_t that is O^1_{t-1} in $\mathcal{D}'$,
5. is into O^1_t because else it would need to be directed from O^1_t to O^1_{t-1}, which contradicts time order,
6. does not intersect O^2_t because else O^2_t would need to be an ancestor of O^1_t (which is not possible to due to the edge $O^1_t {\rightarrow} O^2_t$ in $\mathcal{M}^1(\mathcal{D}')$) or of O^1_{t-1} (which is not possible due to time order),
7. does not intersect O^2_{t-1} because else the subpath $\pi(O^2_{t-1}, O^1_t)$ would be a collider-free path such that all its non-end-point vertices, if any, are unobserved. Consequently, there would need to be an edge between O^2_{t-1} and O^1_t in $\mathcal{M}^1(\mathcal{D}')$.

Since the combination of points 6 and 7 in this enumeration contradicts point 4 of the enumeration, such a path π cannot exist. Consequently, O^1_{t-1} and O^1_t are d-separated in $\mathcal{G}$. The second rule of the do-calculus, for example Pearl (2009), thus gives that the interventional distribution $P(O^1_t \mid do(O^1_{t-1} = o^1_{t-1}))$ is expressed in terms of the observational distribution as $P(O^1_t \mid do(O^1_{t-1} = o^1_{t-1})) = P(O^1_t \mid O^1_{t-1} = o^1_{t-1})$. □

B.8. *Increasing the number of observed time steps.* The main text considers ts-DMAGs and ts-DPAGs on observed time windows $[t-p,t]$, where $p \geq 0$ is arbitrary but fixed. In Section B.8.1 we first compare ts-DMAGs and ts-DPAGs on time windows of different length. We show that, as expected, the ts-DMAGs and ts-DPAGs on the longer time window can never contain less but may contain more information about the underlying ts-DAG than the ts-DMAGs and ts-DPAGs on the shorter time window. In Section B.8.2 we then define the notions of limiting ts-DMAGs and ts-DPAGs by allowing conditioning sets from the entire past. All proofs are given in Section F.

B.8.1. *Comparison of ts-DMAGs and ts-DPAGs on different observed time windows.* Since the reference time step t is arbitrary and only time *differences* are relevant, we need only compare ts-DMAGs and ts-DPAGs on $[t-p,t]$ and $[t-\tilde{p},t]$ with $\tilde{p} > p$. To this end, we use the following notation.

DEFINITION B.9 (Subgraph of a ts-DMAG / ts-DPAG induced on time window). Let $\mathcal{D}$ be a ts-DAG, let $\tilde{p} \geq p \geq 0$, and let $t-\tilde{p} \leq t_1 \leq t_2 \leq t$. The induced subgraph of $\mathcal{M}^{\tilde{p}}(\mathcal{D})$ (subgraph of $\mathcal{P}^{\tilde{p}}(\mathcal{D})$) on its subset of vertices within $[t_1, t_2]$ is denoted as $\mathcal{M}^{\tilde{p},[t_1,t_2]}(\mathcal{D})$ (denoted as $\mathcal{P}^{\tilde{p},[t_1,t_2]}(\mathcal{D})$).

The additionally observed vertices in $[t-\tilde{p}, t-p-1]$ enlarge the set of potential conditions and thus may lead to more d-separations among the originally observed vertices. We thus get the following result.

LEMMA B.10. *Let $\mathcal{D}$ be a ts-DAG and let $\tilde{p} > p \geq 0$. Then, up to relabeling vertices:*

1. *For all $0 \leq \Delta t < \tilde{p} - p$: $\mathcal{M}^{\tilde{p},[t-p-\Delta t, t-\Delta t]}(\mathcal{D})$ is a subgraph of $\mathcal{M}^p(\mathcal{D})$.*
2. *$\mathcal{M}^{\tilde{p},[t-\tilde{p}, t-\tilde{p}+p]}(\mathcal{D})$ equals $\mathcal{M}^p(\mathcal{D})$.*
3. *There are cases in which $\mathcal{M}^{\tilde{p},[t-p,t]}(\mathcal{D})$ is a proper subgraph of $\mathcal{M}^p(\mathcal{D})$.*

Moving to a semantic level, we confirm the intuition that $\mathcal{M}^{\tilde{p}}(\mathcal{D})$ cannot contain less but may contain more information about $\mathcal{D}$ than $\mathcal{M}^p(\mathcal{D})$.

LEMMA B.11. *Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be ts-DAGs and let $\tilde{p} > p \geq 0$. Then:*

1. *If $\mathcal{M}^{\tilde{p}}(\mathcal{D}_1) = \mathcal{M}^{\tilde{p}}(\mathcal{D}_2)$, then $\mathcal{M}^p(\mathcal{D}_1) = \mathcal{M}^p(\mathcal{D}_2)$.*
2. *There are cases in which $\mathcal{M}^{\tilde{p}}(\mathcal{D}_1) \neq \mathcal{M}^{\tilde{p}}(\mathcal{D}_2)$ and $\mathcal{M}^p(\mathcal{D}_1) = \mathcal{M}^p(\mathcal{D}_2)$.*

In other words: Every inference about $\mathcal{D}$ that can be drawn from $\mathcal{M}^p(\mathcal{D})$ can also be drawn from $\mathcal{M}^{\tilde{p}}(\mathcal{D})$, whereas the converse need not be true.

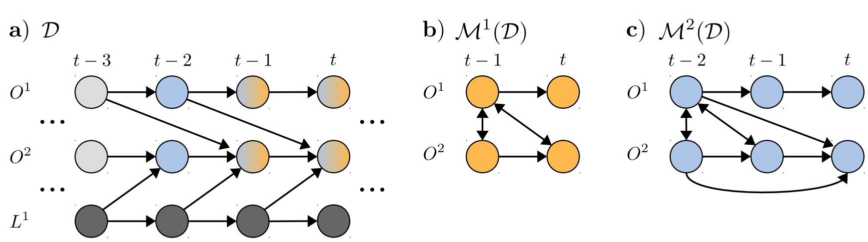

FIG B. *Illustration of ts-DMAGs $\mathcal{M}^p(\mathcal{D})$ of the same ts-DAG $\mathcal{D}$ for different p, see also the discussion in Example B.12. The component time series L^1 is unobservable.*

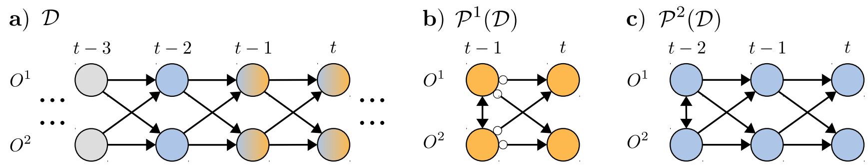

FIG C. *Illustration of ts-DPAGs $\mathcal{P}^p(\mathcal{D})$ of the same ts-DAG $\mathcal{D}$ for different p.*

EXAMPLE B.12. Figure B shows the ts-DMAGs $\mathcal{M}^1(\mathcal{D})$ and $\mathcal{M}^2(\mathcal{D})$ for a ts-DAG $\mathcal{D}$. These graphs conform with parts 1 and 2 of Lemma B.10 and prove its part 3. Further, the edge $O^1_{t-2}{\rightarrow}O^2_t$ in $\mathcal{M}^2(\mathcal{D})$ shows that O^1_{t-2} is an ancestor of O^2_t in $\mathcal{D}$, which is a conclusion that cannot be drawn from $\mathcal{M}^1(\mathcal{D})$.

Since ts-DPAGs $\mathcal{P}^p(\mathcal{D})$ by definition have the same adjacencies as the corresponding ts-DMAGs $\mathcal{M}^p(\mathcal{D})$, the effect of increasing p on their adjacencies is the same as for ts-DMAGs. Regarding edge orientations, Lemma B.11 raises the expectation that all unambiguous edge marks in $\mathcal{P}^p(\mathcal{D})$ should also be in $\mathcal{P}^{\tilde{p}}(\mathcal{D})$. This is expectation is indeed correct.

LEMMA B.13. *Let $\mathcal{D}$ be a ts-DAG and let $\tilde{p} > p \geq 0$. Let (i,t_i) and (j,t_j) be adjacent in both $\mathcal{P}^p(\mathcal{D})$ and $\mathcal{P}^{\tilde{p}}(\mathcal{D})$. Then:*

1. *If there is a noncircle mark on $(i,t_i){*\!\!-\!\!*}(j,t_j)$ in $\mathcal{P}^p(\mathcal{D})$, then the same noncircle mark is also on $(i,t_i){*\!\!-\!\!*}(j,t_j)$ in $\mathcal{P}^{\tilde{p}}(\mathcal{D})$.*
2. *There are cases in which there is a noncircle mark on $(i,t_i){*\!\!-\!\!*}(j,t_j)$ in $\mathcal{P}^{\tilde{p}}(\mathcal{D})$ that is not on $(i,t_i){*\!\!-\!\!*}(j,t_j)$ in $\mathcal{P}^p(\mathcal{D})$.*

LEMMA B.14. *Let $\mathcal{D}$ be a ts-DAG and let $\tilde{p} > p \geq 0$. Then:*

1. *Every circle edge mark in $\mathcal{P}^{\tilde{p},[t-p,t]}(\mathcal{D})$ is also in $\mathcal{P}^p(\mathcal{D})$.*
2. *There are cases in which there is a noncircle edge mark in $\mathcal{P}^{\tilde{p},[t-p,t]}(\mathcal{D})$ that is not also in $\mathcal{P}^p(\mathcal{D})$.*

EXAMPLE B.15. Figure C shows the ts-DPAGs $\mathcal{P}^1(\mathcal{D})$ and $\mathcal{P}^2(\mathcal{D})$ for a ts-DAG $\mathcal{D}$. These graphs conform with parts 1 of Lemmas B.13 and B.14 and prove parts 2 of both these lemmas. For example, in $\mathcal{P}^1(\mathcal{D})$ there is $O^1_{t-1}{\circ\!\!\rightarrow}O^1_t$ while in $\mathcal{P}^2(\mathcal{D})$ there is $O^1_{t-1}{\rightarrow}O^1_t$ instead.

B.8.2. *Limiting ts-DMAGs and limiting ts-DPAGs.* Lemmas B.10 and B.13 imply the following behaviour when p is kept fixed while $\tilde{p} \geq p$ increases beyond any bound.

LEMMA B.16. *Let $\mathcal{D}$ be a ts-DAG and $p \geq 0$. Then:*

1. *There is $\tilde{p} \geq p$ with $\mathcal{M}^{\tilde{p}',[t-p,t]}(\mathcal{D}) = \mathcal{M}^{\tilde{p},[t-p,t]}(\mathcal{D})$ for all $\tilde{p}' \geq \tilde{p}$.*
2. *There is $\tilde{p} \geq p$ with $\mathcal{P}^{\tilde{p}',[t-p,t]}(\mathcal{D}) = \mathcal{P}^{\tilde{p},[t-p,t]}(\mathcal{D})$ for all $\tilde{p}' \geq \tilde{p}$.*

Lemma B.16 implies that the sequence $\Delta p \mapsto \mathcal{M}^{p+\Delta p,[t-p,t]}(\mathcal{D})$ as well as the sequence $\Delta p \mapsto \mathcal{P}^{p+\Delta p,[t-p,t]}(\mathcal{D})$ convergence with respect to the discrete metric[1] on the space of ts-DMAGs, respectively the space of ts-DPAGs.

[1]The discrete metric $d(\cdot,\cdot)$ is defined by $d(x,y) = 1$ if $x = y$ and $d(x,y) = 0$ else.

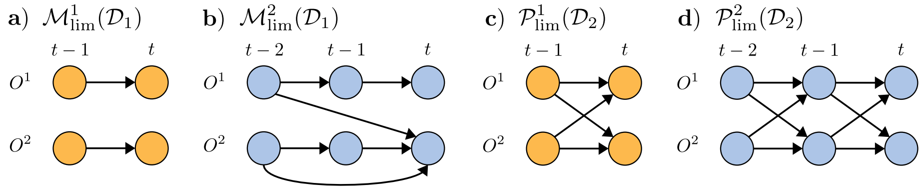

FIG D. *Illustration of limiting ts-DMAGs and limiting ts-DPAGs. The underlying ts-DAGs $\mathcal{D}_1$ and $\mathcal{D}_2$ are, respectively, those shown in parts (a) of Figure B and Figure C.*

DEFINITION B.17 (Limiting ts-DMAG / ts-DPAG). Let $\mathcal{D}$ be a ts-DAG and let $p \geq 0$. The *limiting ts-DMAG* $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$, respectively *limiting ts-DPAG* $\mathcal{P}^p_{\text{lim}}(\mathcal{D})$, is the limit of the sequence $\Delta p \mapsto \mathcal{M}^{p+\Delta p,[t-p,t]}(\mathcal{D})$, respectively $\Delta p \mapsto \mathcal{P}^{p+\Delta p,[t-p,t]}(\mathcal{D})$, with respect to the discrete metric on the space of ts-DMAGs, respectively the space of ts-DPAGs.

See Figure D for examples. Similar to stationarified ts-DMAGs, limiting ts-DMAGs $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$ are not in general DMAGs for the underlying ts-DAGs $\mathcal{D}$ and carry different meaning. Namely, vertices (i, t_i) and (j, t_j) with $t_i \leq t_j \leq t$ are adjacent in $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$ if and only if there is no finite set of observable variables within $(-\infty, t]$ that d-separates (i, t_i) and (j, t_j) in $\mathcal{D}$. The same statement applies to limiting ts-DPAGs.

LEMMA B.18. 1. *$\mathcal{M}^p_{\text{lim}}(\mathcal{D})$ has repeating edges.*
2. *$\mathcal{M}^p_{\text{lim}}(\mathcal{D})$ is a subgraph of $\mathcal{M}^p_{\text{st}}(\mathcal{D})$.*
3. *$\mathcal{M}^p_{\text{lim}}(\mathcal{D})$ is a DMAG.*
4. *$\mathcal{P}^p_{\text{lim}}(\mathcal{D})$ has repeating edges.*
5. *$\mathcal{P}^p_{\text{lim}}(\mathcal{D})$ is a DPAG for $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$.*

Unlike the examples shown in parts (c) and (d) of Figure D, in general there may be circle marks in a limiting ts-DPAG. Lastly, given that $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$ and $\mathcal{P}^p_{\text{lim}}(\mathcal{D})$ have repeating edges, one might hope to give meaning to sending p to infinity in $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$ and $\mathcal{P}^p_{\text{lim}}(\mathcal{D})$ by restricting attention to edges that involve a vertex at time t. However, as the following example shows, such a construction is not possible in general.

EXAMPLE B.19. Consider the ts-DAG $\mathcal{D}$ in part (a) of Figure B. Since L^1 is unobservable and autocorrelated, for all p there is $O^2_{t-p} {\rightarrow} O^2_t$ in $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$. In Malinsky and Spirtes (2018) this effect is discussed under the names of "auto-lag confounders" and "infinite-lag associations".

C. Proofs for Section 3 and for Lemma B.5 and Lemma B.6.

C.1. *Proofs for Section 3.3.*

PROOF OF LEMMA 3.5. See the explanations in Sections 3.2 and 3.3 of the main text. Formally: The set of random variables involved in the structural process defined in Section 3.1 is $\{V^i_t \mid 1 \leq i \leq n_V, t \in \mathbb{Z}\}$, i.e., is indexed by the set $\mathbf{I} \times \mathbb{Z}$ where $\mathbf{I} = \{1, \ldots, n_V\}$. This form shows that the causal graph $\mathcal{D}$ has time series structure with time index set $\mathbb{Z}$, where the vertex $(i,t) \in \mathbf{I} \times \mathbb{Z}$ corresponds to the random variable V^i_t. Further, $\text{pa}((i,t), \mathcal{D}) = PA^i_t$ by definition of causal graphs and $PA^i_t \subseteq \{V^k_{t-\tau} \mid 1 \leq k \leq n_V, 0 \leq \tau \leq p_{\text{ts}}\} \setminus \{V^i_t\}$ by definition of the data-generating process. Hence, $\mathcal{D}$ is time ordered. The repeating edges property follows because the data-generating process by definition is causally stationary, i.e., because $V^k_{t-\tau} \in PA^i_t$ if and only if $V^k_{t-\tau-\Delta t} \in PA^i_{t-\Delta t}$. Lastly, acyclicity of $\mathcal{D}$ is definitional for the data-generating process. □

C.2. *Proofs for Section 3.4.*

LEMMA C.1. *Let $\mathcal{D}$ be a DAG with vertex set $\mathbf{V} = \mathbf{O} \cup \mathbf{L}$. Then, for $i, j \in \mathbf{O}$, $i \in \mathrm{an}(j, \mathcal{D})$ if and only if $j \in \mathrm{an}(i, \mathcal{M}_{\mathbf{O}}(\mathcal{D}))$.*

REMARK (on Lemma C.1). This claim is a well-known result, see for example Zhang (2008a), Zhang (2008b), which straightforwardly follows from the definition of the MAG latent projection procedure in Zhang (2008a) as well as from the definitions in Richardson and Spirtes (2002). However, since we did not find a formal proof spelled out in the literature, we here include a proof for completeness.

PROOF OF LEMMA C.1. **Only if.** Assume $i \in \mathrm{an}(j, \mathcal{D})$. This assumption means that in $\mathcal{D}$ there is a directed path π from i to j. Let $(k_1, \ldots, k_n)$ with $k_1 = i$ and $k_n = j$ be the ordered sequence of nodes on π that are in $\mathbf{O}$. Consequently, for all $1 \leq m \leq n-1$ the vertices k_m and k_{m+1} can in $\mathcal{D}$ not be d-separated by any set of observed variables, and hence there are the edges $k_m {\rightarrow} k_{m+1}$ in $\mathcal{M}_{\mathbf{O}}(\mathcal{D})$ for all $1 \leq m \leq n-1$. These edges give a directed path from $k_1 = i$ to $k_n = j$ in $\mathcal{M}_{\mathbf{O}}(\mathcal{D})$, and hence $i \in \mathrm{an}(j, \mathcal{M}_{\mathbf{O}}(\mathcal{D}))$.

If. Assume $i \in \mathrm{an}(j, \mathcal{M}_{\mathbf{O}}(\mathcal{D}))$. This assumption means that in $\mathcal{M}_{\mathbf{O}}(\mathcal{D})$ there is a directed path π from i to j. Let $(k_1, \ldots, k_{n'})$ with $k_1 = i$ and $k_{n'} = j$ be the ordered sequence of nodes on π. By definition of edge orientations in $\mathcal{M}$ we thus get $k_m \in \mathrm{an}(k_{m+1}, \mathcal{D})$ for all $1 \leq m \leq n'-1$, and hence $i \in \mathrm{an}(j, \mathcal{D})$ by transitivity of ancestorship. □

PROOF OF LEMMA 3.7. **1.** The vertex set of $\mathcal{M}_{\mathbf{I_O} \times \mathbf{T_O}}(\mathcal{D})$ is $\mathbf{I_O} \times \mathbf{T_O}$. This decomposition defines the time series structure of $\mathcal{M}_{\mathbf{I_O} \times \mathbf{T_O}}(\mathcal{D})$, namely: $\mathbf{I_O}$ is its variable index set and $\mathbf{T_O}$ is its time index set.

2. Assume $\mathcal{M}_{\mathbf{I_O} \times \mathbf{T_O}}(\mathcal{D})$ is not time ordered, i.e., assume there is $(j, t_j) {\rightarrow} (i, t_i)$ in $\mathcal{M}_{\mathbf{I_O} \times \mathbf{T_O}}(\mathcal{D})$ with $t_j > t_i$. This assumptions means $(j, t_j) \in \mathrm{an}((i, t_i), \mathcal{M}_{\mathbf{I_O} \times \mathbf{T_O}}(\mathcal{D}))$ and thus, by Lemma C.1, $(j, t_j) \in \mathrm{an}((i, t_i), \mathcal{D})$. The latter in turn implies that in $\mathcal{D}$ there is directed path π from (j, t_j) to (i, t_i). This path must at least contain one edge $(k, t_k) {\rightarrow} (l, t_l)$ with $t_k > t_l$, which contradicts time order of $\mathcal{D}$.

3. See part (b) of Figure 2 in the main text for an example. □

D. Proofs for Section 4.

D.1. *Proofs for Section 4.2.*

PROOF OF LEMMA 4.1. **1.** The desired ts-DAG $\mathcal{D}'$ is constructed by treating the vertices at all time steps other than $t - m \cdot n$ with $m \in \mathbb{Z}$ as members of unobservable time series in $\mathcal{D}'$, by shifting all vertices of $\mathcal{D}$ within a time window $[t - m \cdot n - (n-1), t - m \cdot n]$ to time $t - m \cdot n$ in $\mathcal{D}'$ for all $m \in \mathbb{Z}$, and by then relabeling the time steps according to $t - m \cdot n \mapsto t - m$. Formally: Let $\mathbf{I}$ with $\mathbf{I_O} \subseteq \mathbf{I}$ denote the variable index set of $\mathcal{D}$. Define $\mathbf{I}' = \mathbf{I} \cup \mathbf{K}$ with $\mathbf{K} = \mathbf{I} \times \{1, \ldots, n-1\}$ and consider the following map:

$$\begin{aligned} \phi: \qquad & \mathbf{I} \times \mathbb{Z} \rightarrow \mathbf{I}' \times \mathbb{Z} \\ (i, t - \Delta t) \mapsto & \begin{cases} (i, t - \frac{\Delta t}{n}) & \text{for } \Delta t \bmod n = 0 \\ ((i, \Delta t \bmod n), t - \lfloor \frac{\Delta t}{n} \rfloor) & \text{for } \Delta t \bmod n \neq 0 \,. \end{cases} \end{aligned}$$

Here, $(\Delta t \bmod n)$ with negative Δt is defined as the smallest nonnegative integer $\Delta t + n \cdot m$ with $m \in \mathbb{Z}$. Note that ϕ is bijective, and hence invertible. We define $\mathcal{D}'$ as the directed graph

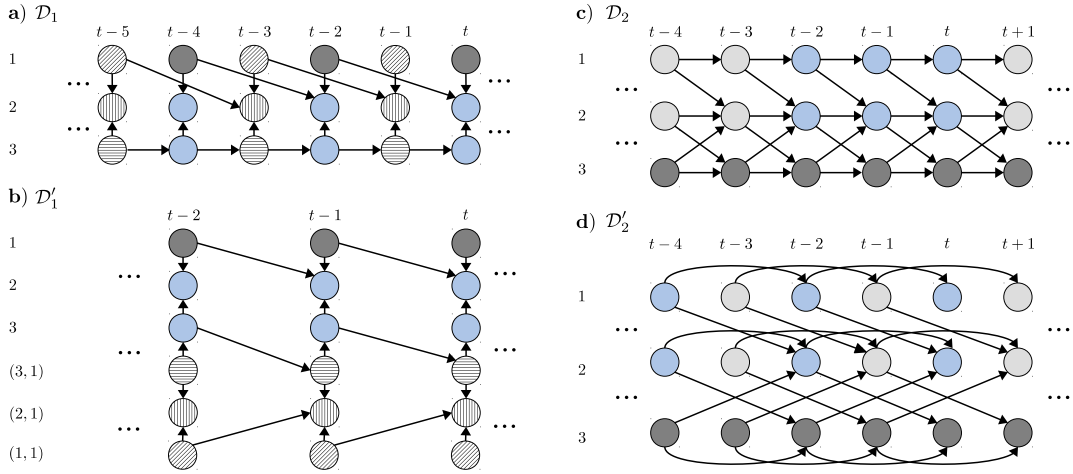

FIG E. *Illustration of the constructions involved in the proof of Lemma 4.1 for $n = 2$ and $n_{steps} = 3$. The vertically arranged numbers to the left of the four ts-DAGs are their respective variable indices.* a) *The same ts-DAG as in part (c) of Figure 2, here denoted $\mathcal{D}_1$. Here, $\mathbf{I_O} = \{2, 3\} \subseteq \mathbf{I} = \{1, 2, 3\}$ and $\mathbf{T_O} = \{t-4, t-2, t\}$. The corresponding implied ts-DMAG $\mathcal{M}_{\mathbf{I_O}\times\mathbf{T_O}}(\mathcal{D}_1)$ is shown in part (d) of Figure 2.* b) *The ts-DAG $\mathcal{D}'_1$ constructed from $\mathcal{D}_1$ as defined in the proof of part 1 of Lemma 4.1. Here, $\mathbf{I}'_\mathbf{O} = \mathbf{I_O} = \{2, 3\} \subseteq \mathbf{I}' = \mathbf{I} \cup \mathbf{K}$ with $\mathbf{K} = \mathbf{I} \times \{1\}$ and $\mathbf{T}'_\mathbf{O} = \{t-2, t-1, t\}$. Note that while in $\mathcal{D}_1$ the hatched vertices are temporally unobserved, in $\mathcal{D}'_1$ they are unobservable. The corresponding implied ts-DMAG $\mathcal{M}_{\mathbf{I}'_\mathbf{O}\times\mathbf{T}'_\mathbf{O}}(\mathcal{D}'_1)$ is the same as $\mathcal{M}_{\mathbf{I_O}\times\mathbf{T_O}}(\mathcal{D}_1)$ up to relabeling vertices.* c) *The same ts-DAG as in part (a) of Figure 2, here denoted $\mathcal{D}_2$. Here, $\mathbf{I_O} = \{1, 2\} \subseteq \mathbf{I} = \{1, 2, 3\}$ and $\mathbf{T_O} = \{t-2, t-1, t\}$. The corresponding implied ts-DMAG $\mathcal{M}_{\mathbf{I_O}\times\mathbf{T_O}}(\mathcal{D}_2)$ is shown in part (b) of Figure 2.* d) *The ts-DAG $\mathcal{D}'_2$ constructed from $\mathcal{D}_2$ as defined in the proof of part 2 of Lemma 4.1. Here, $\mathbf{I}'_\mathbf{O} = \mathbf{I_O} = \{1, 2\} \subseteq \mathbf{I}' = \mathbf{I}$ and $\mathbf{T}'_\mathbf{O} = \{t-4, t-2, t\}$. The corresponding implied ts-DMAG $\mathcal{M}_{\mathbf{I}'_\mathbf{O}\times\mathbf{T}'_\mathbf{O}}(\mathcal{D}'_2)$ is the same as $\mathcal{M}_{\mathbf{I_O}\times\mathbf{T_O}}(\mathcal{D}_2)$ up to relabeling vertices.*

over the vertex set $\mathbf{V}' = \mathbf{I}' \times \mathbb{Z}$ such that for vertices $a, b \in \mathbf{V}'$ there is an edge $a \to b$ if and only if $\phi^{-1}(a) \to \phi^{-1}(b)$ in $\mathcal{D}$. See parts (a) and (b) of Figure E for illustration.

This construction is such that $\mathcal{D}$ and $\mathcal{D}'$ are as graphs equal up to relabeling their vertices according to ϕ. As a consequence, $\mathcal{D}'$ is acyclic and its d-separations are the same as those of $\mathcal{D}$. Moreover, $\mathcal{D}'$ is indeed a ts-DAG: First, its time series structure is given by the decomposition of $\mathbf{V}'$ into $\mathbf{V}' = \mathbf{I}' \times \mathbb{Z}$, i.e., $\mathbf{I}'$ is its variable index set. Second, time order follows because (1) if $b \in \mathbf{V}'$ is after $a \in \mathbf{V}'$, then $\phi^{-1}(b)$ is after $\phi^{-1}(a)$ together with the fact (2) that $\mathcal{D}$ is time ordered. Third, if for $a = (i', t') \in \mathbf{V}'$ we write $\phi^{-1}(a) = (i, t)$, then $\phi^{-1}((i', t' - \Delta t')) = (i, t - n \cdot \Delta t')$. This observation implies that $\mathcal{D}'$ has repeating edges. Hence, $\mathcal{D}'$ is a ts-DAG and the statement follows since $\phi(\mathbf{I_O} \times \mathbf{T}^n_\mathbf{O}) = \mathbf{I_O} \times \mathbf{T}^1_\mathbf{O}$.

2. The desired ts-DAG $\mathcal{D}'$ is constructed by stretching all edges of $\mathcal{D}$ by a factor of n in time and adding $(n-1)$ further copies of this stretched version of $\mathcal{D}$ to $\mathcal{D}'$, respectively shifted by 1 up to $(n-1)$ time steps with respect to the first copy, without any edges between the n copies. Formally: $\mathcal{D}'$ is the ts-DAG over the vertex set $\mathbf{V}' = \mathbf{I} \times \mathbb{Z}$, where $\mathbf{I}$ is the variable index set of $\mathcal{D}$, such that $(i, t' - \Delta t') \to (j, t')$ in $\mathcal{D}'$ if and only if $(\Delta t' \bmod n) = 0$ and $(i, t' - \Delta t'/n) \to (j, t')$ in $\mathcal{D}$. See parts (c) and (d) of Figure E for illustration. The statement is apparent from this construction. □

D.2. *Proofs for Section 4.3.*

PROOF OF LEMMA 4.3. In all statements that involve the repeating ancestral relationships or repeating separating sets property, we implicitly assume the graph to be a DMAG (because else these properties would be undefined).

1. & 2. These statements immediately follow from the definitions of the involved properties.

3. This statement follows because the ancestral relationships between an adjacent pair of vertices uniquely specifies the type of the edge between this pair of vertices.

4. This statement follows because for $\mathbf{T} = \mathbb{Z}$ repeating edges implies that the graph is invariant under time shifts, i.e., invariant under the mapping $\phi_{\Delta t} : \mathbf{I} \times \mathbf{T} \to \mathbf{I} \times \mathbf{T}$ with $\phi_{\Delta t}((i,t_i)) = (i, t_i + \Delta t)$ for all $\Delta t \in \mathbb{Z}$. □

PROOF OF LEMMA 4.4. **1.** This statement follows because $\mathcal{M}^p(\mathcal{D})$ and $\mathcal{D}$ have the same ancestral relationships between vertices in $\mathcal{M}^p(\mathcal{D})$ (according to Lemma C.1) in combination with the fact that $\mathcal{D}$ has repeating ancestral relationships (as implied by part 4 of Lemma 4.3).

2. Combine part 1 of Lemma 4.4 with part 3 of Lemma 4.3.

3. Theorem 4.18 in Richardson and Spirtes (2002) implies that for sets $\mathbf{S}$ of vertices in $\mathcal{M}^p(\mathcal{D})$ the m-separation $(i,t_i) \perp\!\!\!\perp (j,t_j) \mid \mathbf{S}$ holds in $\mathcal{M}^p(\mathcal{D})$ if and only if the d-separation $(i,t_i) \perp\!\!\!\perp (j,t_j) \mid \mathbf{S}$ holds in $\mathcal{D}$. The statement now follows because $\mathcal{D}$ has repeating separating sets as implied by part 4 of Lemma 4.3.

4. Let (i,t_i) and (j,t_j) be nonadjacent in $\mathcal{M}^p(\mathcal{D})$ and without loss of generality assume $t_i \leq t_j$. Consequently, there is a set of vertices $\mathbf{S}$ in $\mathcal{M}^p(\mathcal{D})$ with $\mathbf{S} \cap \{(i,t_i),(j,t_j)\} = \emptyset$ such that $(i,t_i) \perp\!\!\!\perp (j,t_j) \mid \mathbf{S}$ in $\mathcal{M}^p(\mathcal{D})$. Due to time order of $\mathcal{D}$, no (k,t_k) with $t_j < t_k$ can be an ancestor of (i,t_i) or of (j,t_j). Lemma S5 in the supplementary material of Gerhardus and Runge (2020) then asserts that $(i,t_i) \perp\!\!\!\perp (j,t_j) \mid \mathbf{S}'$ in $\mathcal{M}^p(\mathcal{D})$ with $\mathbf{S}' = \mathbf{S} \cap \{(l,t_l) \mid t_l \leq t_j\}$. Now take any $\Delta t \in \mathbb{Z}$ with $0 \leq \Delta t \leq t - t_j$ and let $\mathbf{S}'_{\Delta t}$ be obtained by shifting all vertices in $\mathbf{S}$ forward in time by Δt time steps. By construction of $\mathbf{S}'$ all nodes in $\mathbf{S}'_{\Delta t}$ are within $[t-p, t]$. The repeating separating sets property of $\mathcal{M}^p(\mathcal{D})$, as asserted by part 2 of Lemma 4.4. and already proven, then implies $(i, t_i + \Delta t) \perp\!\!\!\perp (j, t_j + \Delta t) \mid \mathbf{S}'_{\Delta t}$. This fact proves the contraposition of the statement.

5. Part 1 of Lemma 4.4 implies that $\mathcal{M}^p(\mathcal{D})$ has repeating orientations. Thus, part 1 of Lemma 4.3 shows that $\mathcal{M}^p(\mathcal{D})$ would also necessarily have repeating edges if $\mathcal{M}^p(\mathcal{D})$ necessarily had repeating adjacencies, thereby contradicting part 3 of Lemma 3.7. □

D.3. *Proofs for Section 4.4 and of Lemma B.5.*

LEMMA D.1. *Let $\mathcal{G}$ be a directed partial mixed graph with time series structure that has repeating orientations and past-repeating adjacencies. Then,* $\mathrm{stat}(\mathcal{G})$ *is the unique subgraph $\mathcal{G}$ in which (i,t_i) and (j,t_j) with $\tau = t_j - t_i \geq 0$ are adjacent if and only if $(i, t-\tau)$ and (j,t) are adjacent in $\mathcal{G}$.*

PROOF OF LEMMA D.1. The statement uniquely determines $\mathrm{stat}(\mathcal{G})$ for the following reason: First, as is immediate from Definition 4.6 and asserted by the statement, $\mathrm{stat}(\mathcal{G})$ is a subgraph of $\mathcal{G}$. Second, the statement specifies the set of edges that are in $\mathcal{G}$ but not in $\mathrm{stat}(\mathcal{G})$. Consequently, $\mathrm{stat}(\mathcal{G})$ is obtained by deleting a specified set of edges from $\mathcal{G}$.

It remains to be shown that $\mathrm{stat}(\mathcal{G})$ has the asserted property: First, consider two vertices (i,t_i) and (j,t_j) with $\tau = t_j - t_i \geq 0$ that are adjacent in $\mathrm{stat}(\mathcal{G})$. Since $\mathrm{stat}(\mathcal{G})$ has repeating edges we thus get that $(i, t-\tau)$ and (j,t) are adjacent in $\mathrm{stat}(\mathcal{G})$, which in turn gives that $(i, t-\tau)$ and (j,t) are adjacent in $\mathcal{G}$ because it is a supergraph of $\mathrm{stat}(\mathcal{G})$. Second, consider two vertices $(i,t-\tau)$ and (j,t) that are adjacent in $\mathcal{G}$. The past-repeating adjacencies property of $\mathcal{G}$ then implies that $(i,t_i+\Delta t)$ and $(j,t_j+\Delta t)$ are adjacent in $\mathcal{G}$ for all Δt with $(i,t_i+\Delta t),(j,t_j+\Delta t)\in\mathbf{V}$. Moreover, since $\mathcal{G}$ has repeating orientations, all these edges have the same orientation. By the second point in Definition 4.6 we thus get that (i,t_i) and (j,t_j) are adjacent in $\mathrm{stat}(\mathcal{G})$. □

PROOF OF LEMMA 4.7. Apply Lemma D.1 to $\mathcal{G}=\mathcal{M}^p(\mathcal{D})$. □

PROOF OF LEMMA B.5. Denote the subgraph of $\mathcal{D}$ induced on $\mathbf{I}\times\mathbf{T_O}$ as $\mathcal{D}_{[t-p,t]}$. This graph clearly has the same set of vertices and the same time series structure as $\mathcal{M}_{\mathbf{I_O}\times\mathbf{T_O}}(\mathcal{D})$ because $\mathbf{I_O}=\mathbf{I}$ by assumption.

First, we show that $\mathcal{D}_{[t-p,t]}$ and $\mathrm{stat}(\mathcal{M}_{\mathbf{I_O}\times\mathbf{T_O}}(\mathcal{D}))$ have the same adjacencies: Let $(i,t-\tau)$ and (j,t) with $0\leq\tau=t_j-t_i$ be distinct vertices in $\mathcal{D}$, where without loss of generality $(j,t)\notin\mathrm{an}((i,t-\tau),\mathcal{D})$. If $(i,t-\tau)$ and (j,t) are adjacent in $\mathcal{D}$, then there is no set $\mathbf{S}$ with $\mathbf{S}\cap\{(i,t-\tau),\,(j,t)\}=\emptyset$ that d-separates them. If $(i,t-\tau)$ and (j,t) are nonadjacent in $\mathcal{D}$, then $(i,t-\tau)\perp\!\!\!\perp(j,t)\mid\mathbf{S}$ with $\mathbf{S}=\mathrm{pa}((j,t),\mathcal{D})\setminus\{(i,t-\tau)\}$. By time order of $\mathcal{D}$ and the definition of p_{ts}, all vertices in $\mathrm{pa}((j,t),\mathcal{D})\setminus\{(i,t-\tau)\}$ are within $[t-p_{\mathrm{ts}},t]$. Since $p_{\mathrm{ts}}\leq p$ by assumption we thus get: $(i,t-\tau)$ and (j,t) can be d-separated in $\mathcal{D}$ by a set of vertices in $\mathbf{I_O}\times\mathbf{T_O}$ if and only if they are nonadjacent in $\mathcal{D}$. In combination with repeating edges of $\mathcal{D}$ and Lemma 4.7 the desired claim follows.

Second, $\mathcal{D}_{[t-p,t]}$ and $\mathcal{M}_{\mathbf{I_O}\times\mathbf{T_O}}(\mathcal{D})$ also have the same edge orientations because they have the same ancestral relationships according Lemma C.1. □

LEMMA D.2. *Let $\mathcal{M}$ be a DMAG with time series structure that is time ordered and has repeating orientations and past-repeating adjacencies. Then,* $\mathrm{stat}(\mathcal{M})$ *is a DMAG.*

PROOF OF LEMMA D.2. We have to show that $\mathrm{stat}(\mathcal{M})$ does not have directed cycles, does not have almost directed cycles, and is maximal.

No (almost) directed cycles: Assume that $\mathrm{stat}(\mathcal{M})$ has a directed or an almost direct cycle. Then, since $\mathrm{stat}(\mathcal{M})$ is a subgraph of $\mathcal{M}$, also $\mathcal{M}$ has the same directed or almost directed cycle. But then $\mathcal{M}$ is not a DMAG. Contradiction.

Maximality: Assume the opposite, i.e., assume in $\mathrm{stat}(\mathcal{M})$ there are nonadjacent vertices (i,t_i) and (j,t_j), where without loss of generality $\tau=t_j-t_i\geq 0$, between which there is an inducing path π. We note that $\mathrm{stat}(\mathcal{M})$ is time ordered because it is a subgraph of the time ordered graph $\mathcal{M}$. Since by definition of inducing paths all vertices on π are ancestors of (i,t_i) or (j,t_j), we get that all vertices on π are within the time window $[t-p,t_j]$. The repeating edges property of $\mathrm{stat}(\mathcal{M})$ now shows that π_{t-t_j}, defined as the ordered sequence of vertices obtained by shifting all vertices on π forward in time by $t-t_j$ time steps, is a path in $\mathrm{stat}(\mathcal{M})$ whose edges are orientated in the same way as the corresponding edges of π. Moreover, by combining part 1 of Lemma 4.3 with part 1 of Lemma B.4 we see that the stationarification $\mathrm{stat}(\mathcal{M})$ has repeating ancestral relationships. Hence, π_{t-t_j} is an inducing path between $(i,t-\tau)$ and (j,t) in $\mathrm{stat}(\mathcal{M})$. Since $\mathrm{stat}(\mathcal{M})\subseteq\mathcal{M}$, π_{t-t_j} is also in $\mathcal{M}$ an inducing path between $(i,t-\tau)$ and (j,t). Maximality of $\mathcal{M}$ thus requires $(i,t-\tau)$ and (j,t) to be adjacent in $\mathcal{M}$. According to Lemma D.1 we then obtain that $(i,t-\tau)$ and (j,t) are adjacent in $\mathrm{stat}(\mathcal{M})$. Since (i,t_i) and (j,t_j) are nonadjacent in $\mathrm{stat}(\mathcal{M})$ by assumption, this observation contradicts repeating edges of $\mathrm{stat}(\mathcal{M})$. □

PROOF OF LEMMA 4.8. Apply Lemma D.2 to $\mathcal{M}=\mathcal{M}^p(\mathcal{D})$. □

PROOF OF LEMMA 4.10. Assume $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}))$. This assumption means that in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ there is a directed path π from (i,t_i) to (j,t_j). Since $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ is a subgraph of $\mathcal{M}^p(\mathcal{D})$, this path π is also in $\mathcal{M}^p(\mathcal{D})$. Hence, $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{M}^p(\mathcal{D}))$.

Assume $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{M}^p(\mathcal{D}))$. This assumption by Lemma C.1 implies that $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{D})$, and hence there is a directed path π from (i,t_i) to (j,t_j) in $\mathcal{D}$. Since $\mathcal{D}$ is time ordered, all vertices on π are within $[t-p,t]$. Let $((k_1,t_1),\ldots,(k_n,t_n))$ with $(k_1,t_1)=(i,t_i)$ and $(k_n,t_n)=(j,t_j)$ be the ordered sequence of observed vertices on π. For all $1 \leq m \leq n-1$ let π_m be the ordered sequence of vertices obtained by shifting $\pi((k_m,t_m),(k_{m+1},t_{m+1}))$ by $t-t_{m+1}$ time steps forward in time. These paths π_m are directed paths from $(k_m,t-(t_{m+1}-t_m))$ to (k_{m+1},t) in $\mathcal{D}$ and all their non-end-point vertices unobservable. Hence, the paths π_m cannot be blocked by any set of observable variables, which implies that in $\mathcal{M}^p(\mathcal{D})$ there are the edges $(k_m,t-(t_{m+1}-t_m)){\rightarrow}(k_{m+1},t)$. According to Lemma 4.7, we thus get that $(k_m,t-(t_{m+1}-t_m)){\rightarrow}(k_{m+1},t)$ are in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$, which due to repeating edges of $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ in turn gives $(k_m,t_m){\rightarrow}(k_{m+1},t_{m+1})$ in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$. These edges combine to a directed path from $(k_1,t_1)=(i,t_i)$ to $(k_n,t_n)=(j,t_j)$ in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$, hence $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}))$. □

LEMMA D.3. *Let $\mathcal{G}$ be a directed partial mixed graph with time structure that has repeating edges. Then,* $\mathrm{stat}(\mathcal{G})=\mathcal{G}$.

PROOF OF LEMMA D.3. Apply part 2 of Lemma B.4 for $(\mathcal{G}',\mathcal{G})=(\mathcal{G},\mathcal{G})$ to see that $\mathcal{G}$ is a subgraph of $\mathrm{stat}(\mathcal{G})$. Since $\mathrm{stat}(\mathcal{G})$ is a subgraph of $\mathcal{G}$, as immediately implied by the second point in Definition 4.6, this observation shows $\mathrm{stat}(\mathcal{G})=\mathcal{G}$. □

LEMMA D.4. *Let $\mathcal{G}$ be a directed partial mixed graph with time structure. Then* $\mathrm{stat}(\mathrm{stat}(\mathcal{G}))=\mathrm{stat}(\mathcal{G})$.

PROOF OF LEMMA D.4. Combine part 1 of Lemma B.4 with Lemma D.3. □

D.4. *Proofs for Section 4.5 other than Lemma 4.14.*

PROOF OF LEMMA 4.12. This statement is implied by Theorem 6.4 in Richardson and Spirtes (2002). □

D.5. *Proof of Lemma 4.14.* We split the proof into three parts that are respectively given in Sections D.5.2, D.5.3 and D.5.4. Moreover, we collect several auxiliary results and definitions in D.5.1. For ease of notation, in Section D.5.4 we do not always denote vertices by the tuples of their variable and time indices but sometimes just with a single character, for example v instead of (i,t).

D.5.1. *Auxiliary results and definitions.*

LEMMA D.5. *Let $\mathcal{D}$ be a DAG over vertices $\mathbf{V}$ and let $\mathbf{O} \subseteq \mathbf{V}$ be the subset of observed vertices. Then:*

1. *If $i{\rightarrow}j$ in $\mathcal{M}_{\mathbf{O}}(\mathcal{D})$, then for every subset $\mathbf{S} \subseteq \mathbf{O}$ that does not contain i or j there is path π between i and j in $\mathcal{D}$ that is active given $\mathbf{S}$ and into j.*
2. *If $i{\leftrightarrow}j$ in $\mathcal{M}_{\mathbf{O}}(\mathcal{D})$, then for every subset $\mathbf{S} \subseteq \mathbf{O}$ that does not contain i or j there is path π between i and j in $\mathcal{D}$ that is active given $\mathbf{S}$ and into both i and j.*

REMARK (on Lemma D.5). This result might have appeared in the literature before. Also note that the presence of $i{\rightarrow}j$ in $\mathcal{M}$ does *not* imply that for all $\mathbf{S}$ as above there is path between i and j in $\mathcal{D}$ that is active given $\mathbf{S}$ and out of i. As an example, consider the DAG over $\mathbf{V} = \{i, j, k, l\}$ constituted by $i{\rightarrow}j{\rightarrow}k$ together with $i{\leftarrow}l{\rightarrow}k$ and choose $\mathbf{O} = \{i, j, k\}$: Although $i{\rightarrow}k$ in $\mathcal{M}$, for $\mathbf{S} = \{j\}$ the only active path in $\mathcal{D}$ is $i{\leftarrow}l{\rightarrow}k$.

PROOF OF LEMMA D.5. **1.** The fact that $i{*\!\!-\!\!*}j$ is in $\mathcal{M}_{\mathbf{O}}(\mathcal{D})$ is by Theorem 4.2 in Richardson and Spirtes (2002) equivalent to the existence of an inducing path ρ relative to $\mathbf{O}$ in $\mathcal{D}$ between i and j. Assume ρ is out of j. Then, because j is not an ancestor of i according to $i{*\!\!\rightarrow}j$ in $\mathcal{M}_{\mathbf{O}}(\mathcal{D})$, there is at least one collider on ρ. By definition of inducing paths, all colliders on ρ are ancestors of i or j. Let k be the collider on ρ that is closest to j on ρ. Because ρ is out of j, j is an ancestor of k. Transitivity of ancestorship thus implies that j is an ancestor of i or j. Both options are a contradiction because there are no directed cycles and because $i{*\!\!\rightarrow}j$ is in $\mathcal{M}_{\mathbf{O}}(\mathcal{D})$. Hence, ρ is into j. The statement now follows by combining Lemmas 6.1.1. and 6.1.2 in Spirtes, Glymour and Scheines (2000).

2. Arguments similar to those in the proof of part 1 of Lemma D.5 show that there is an inducing path ρ relative to $\mathbf{O}$ in $\mathcal{D}$ between i and j that is into both i and j. The statement for $i{\leftrightarrow}j$ in $\mathcal{M}$ follows by Lemma 6.1.2 in Spirtes, Glymour and Scheines (2000). □

LEMMA D.6. *Let $\mathcal{D}$ be a DAG over vertices $\mathbf{V}$ and let $\mathbf{O} \subseteq \mathbf{V}$ be the subset of observed vertices. Let $i{\rightarrow}j$ be an edge in $\mathcal{M}_{\mathbf{O}}(\mathcal{D})$ and $\mathbf{S} \subseteq \mathbf{O} \setminus \{i, j\}$. Then: If in $\mathcal{D}$ there is no path between i and j that is active given $\mathbf{S}$ and out of i, then i is an ancestor of $\mathbf{S}$ in $\mathcal{D}$.*

PROOF OF LEMMA D.6. We know that i is an ancestor of j in $\mathcal{D}$ because $i{\rightarrow}j$ in $\mathcal{M}_{\mathbf{O}}(\mathcal{D})$. Hence, in $\mathcal{D}$ there is a directed path π from i to j. Assuming that in $\mathcal{D}$ there is no path between i and j that is active given $\mathbf{S}$ and out of i, π must be blocked by $\mathbf{S}$. Consequently, $\mathbf{S}$ contains a vertex of π and thus a descendant of i. □

DEFINITION D.7 (Observable vertices within a time window). The set of observable vertices within a time window $[t_1, t_2]$, where $t_1 \leq t_2$, are denoted by $\mathbf{O}(t_1, t_2)$.

DEFINITION D.8 (Observable vertices within a time window not on a given path). $\mathbf{O}(t_1, t_2)[\pi]$ is the set of all vertices in $\mathbf{O}(t_1, t_2)$ less the non-end-point vertices of the path π.

DEFINITION D.9 (Almost adjacent). Two distinct observable vertices (i, t_i) and (j, t_j) in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ are *almost adjacent* if there is an unobservable vertex (k, t_k) such that $(i, t_i){\leftarrow}(k, t_k){\rightarrow}(j, t_j)$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$.

LEMMA D.10. *If (i, t_i) and (j, t_j) are almost adjacent in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$, then there is a unique unobservable vertex (k, t_k) such that $(i, t_i){\leftarrow}(k, t_k){\rightarrow}(j, t_j)$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$.*

PROOF OF LEMMA D.10. Existence follows because (i, t_i) and (j, t_j) are almost adjacent, uniqueness follows in combination with the definition of canonical ts-DAGs (see Definition 4.13 in the main text). □

LEMMA D.11. *Let (i, t_i) and (j, t_j) with $t_i \leq t_j$ be distinct observable vertices in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$. Then, (i, t_i) and (j, t_j) are adjacent or almost adjacent in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ if and only if $(i, t - (t_j - t_i))$ and (j, t) are adjacent in $\mathcal{M}^p(\mathcal{D})$.*

PROOF OF LEMMA D.11. **If.** The premise is that $(i, t-(t_j-t_i)){*\!\!-\!\!*}(j,t)$, where ${*\!\!-\!\!*}$ is $\rightarrow$ or $\leftarrow$ or $\leftrightarrow$, is in $\mathcal{M}^p(\mathcal{D})$. Past-repeating adjacencies and repeating orientations of $\mathcal{M}^p(\mathcal{D})$ then imply $(i, t-(t_j-t_i)-\Delta t){*\!\!-\!\!*}(j, t-\Delta t)$ for all $0 \leq \Delta t \leq p-(t_j-t_i)$, where ${*\!\!-\!\!*}$ is the same edge type as between $(i, t-(t_j-t_i))$ and (j,t). Hence, all these edges are also in $\mathrm{stat}(\mathcal{M}^p(\mathcal{D}))$. If ${*\!\!-\!\!*}$ is $\rightarrow$ or $\leftarrow$, then the definition of canonical ts-DAGs implies $(i, t-(t_j-t_i)-\Delta t'){*\!\!-\!\!*}(j, t-\Delta t')$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ for all $\Delta t' \in \mathbb{Z}$. In particular, (i,t_i) and (j,t_j) are adjacent in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$. If ${*\!\!-\!\!*}$ is $\leftrightarrow$, then the definition of canonical ts-DAGs implies that $(i, t-(t_j-t_i)-\Delta t') \leftarrow ((i,j,t_j-t_i), t-\Delta t'-(t_j-t_i)) \rightarrow (j, t-\Delta t')$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ or $t_i = t_j$ and $(i, t-\Delta t') \leftarrow ((j,i,0), t-\Delta t') \rightarrow (j, t-\Delta t')$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ for all $\Delta t' \in \mathbb{Z}$. In particular, (i,t_i) and (j,t_j) are almost adjacent in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$.

Only if. Since the vertices $(i, t-(t_j-t_i))$ and (j,t) are nonadjacent in $\mathcal{M}^p(\mathcal{D})$ they are also nonadjacent in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$. The statement now follows with the definition of canonical ts-DAGs. □

LEMMA D.12. *Let (i,t_i) and (j,t_j) be distinct observable vertices that are adjacent or almost adjacent in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$. Then:*

1. *$(i,t_i) \rightarrow (j,t_j)$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ if and only if $(i,t_i) \in \mathrm{an}((j,t_j), \mathcal{D})$.*
2. *$(i,t_i) \leftarrow (j,t_j)$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ if and only if $(j,t_j) \in \mathrm{an}((i,t_i), \mathcal{D})$.*
3. *(i,t_i) and (j,t_j) are almost adjacent in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ if and only if $(i,t_i) \notin \mathrm{an}((j,t_j), \mathcal{D})$ and $(j,t_j) \notin \mathrm{an}((i,t_i), \mathcal{D})$.*

PROOF OF LEMMA D.12. Assume without loss of generality that $t_i \leq t_j$, else exchange (i,t_i) and (j,t_j). From Lemma D.11 it then follows that $(i, t-(t_j-t_i))$ and (j,t) are adjacent in $\mathcal{M}^p(\mathcal{D})$. The definition of edges in DMAGs in combination with repeating ancestral relationships of $\mathcal{D}$ further implies that

- $(i, t-(t_j-t_i)) \rightarrow (j,t)$ in $\mathcal{M}^p(\mathcal{D})$ if and only if $(i,t_i) \in \mathrm{an}((j,t_j), \mathcal{D})$,
- $(i, t-(t_j-t_i)) \leftarrow (j,t)$ in $\mathcal{M}^p(\mathcal{D})$ if and only if $(j,t_j) \in \mathrm{an}((i,t_i), \mathcal{D})$,
- $(i, t-(t_j-t_i)) \leftrightarrow (j,t)$ in $\mathcal{M}^p(\mathcal{D})$ if and only if $(i,t_i) \notin \mathrm{an}((j,t_j), \mathcal{D})$ and $(j,t_j) \notin \mathrm{an}((i,t_i), \mathcal{D})$.

Now proceed as in the proof of the *if* part of Lemma D.11. □

D.5.2. *Part 1: $\mathcal{M}^p(\mathcal{D})$ and $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$ have the same ancestral relationships.* As the first part of the proof of Lemma 4.14 we here show that $\mathcal{M}^p(\mathcal{D})$ and $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$ have the same ancestral relationships.

LEMMA D.13. $\mathcal{D}_c(\mathcal{G}) = \mathcal{D}_c(\mathrm{stat}(\mathcal{G}))$.

PROOF OF LEMMA D.13. An inspection of Definitions 4.6 and 4.13 in the main text reveals that $\mathcal{D}_c(\mathcal{G})$ is uniquely determined by $\mathrm{stat}(\mathcal{G})$. The statement thus follows because $\mathrm{stat}(\mathcal{G}) = \mathrm{stat}(\mathrm{stat}(\mathcal{G}))$ according to Lemma D.4. □

LEMMA D.14. *The stationarified ts-DMAG $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ has the same ancestral relationships among vertices in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ as the canonical ts-DAG $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$.*

PROOF OF LEMMA D.14. Assume $(i,t_i) \in \mathrm{an}((j,t_j), \mathcal{M}^p_{\mathrm{st}}(\mathcal{D}))$. This assumption means that in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ there is a directed path $\pi = ((k_1,t_1), \ldots, (k_n,t_n))$ from $(k_1,t_1) = (i,t_i)$ to $(k_n,t_n) = (j,t_j)$. Since $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ has repeating edges and is time ordered, the fact that $(k_m,t_m) \rightarrow (k_{m+1},t_{m+1})$ is in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ implies $(k_m,t_m) \rightarrow (k_{m+1},t_{m+1})$ in $\mathcal{D}_c(\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}))$. Consequently, π is also in $\mathcal{D}_c(\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}))$ and we find that $(i,t_i) \in \mathrm{an}((j,t_j), \mathcal{D}_c(\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})))$.

Let $(i,t_i),(j,t_j)$ be vertices in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ and assume $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{D}_c(\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})))$. This assumption means in $\mathcal{D}_c(\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}))$ there is a directed path $\pi = ((k_1,t_1),\ldots,(k_n,t_n))$ from $(k_1,t_1) = (i,t_i)$ to $(k_n,t_n) = (j,t_j)$. Since by definition of canonical ts-DAGs there are no edges into unobservable vertices, all vertices on π are observed and thus also in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$. Moreover, again by definition of canonical ts-DAGs, any edge of $\mathcal{D}_c(\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}))$ that is between vertices in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ is also in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$. Hence, π is also in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ and we find $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}))$.

These considerations show that $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ and $\mathcal{D}_c(\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}))$ have the same ancestral relationships among vertices in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$. The statement follows because $\mathcal{D}_c(\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})) = \mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ according to Lemma D.13. □

LEMMA D.15. *Consider a ts-DAG $\mathcal{D}$ and the canonical ts-DAG $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$. Then:*

1. *If $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{D})$ and $t_j - t_i \le p$, then $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$.*
2. *If $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$, then $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{D})$.*

PROOF OF LEMMA D.15. **1.** Let $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{D})$ with $\tau = t_j - t_i \le p$, where $\tau \ge 0$ due to time order of $\mathcal{D}$. The repeating ancestral relationships property of $\mathcal{D}$ then gives $(i,t-\tau) \in \mathrm{an}((j,t),\mathcal{D})$, which implies $(i,t-\tau) \in \mathrm{an}((j,t),\mathcal{M}^p(\mathcal{D}))$ by Lemma C.1 and thus $(i,t-\tau) \in \mathrm{an}((j,t),\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}))$ by Lemma 4.10 and finally $(i,t-\tau) \in \mathrm{an}((j,t),\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$ by Lemma D.14.

2. Let $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$. This premise means that in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ there is a directed path $\pi = ((k_1,t_1),\ldots,(k_n,t_n))$ from $(k_1,t_1) = (i,t_i)$ to $(k_n,t_n) = (j,t_j)$, where $t_m \le t_{m+1}$ due to time order of $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$. Using repeating ancestral relationships of $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$, we thus get that $(k_m, t-(t_{m+1}-t_m)) \in \mathrm{an}((k_{m+1},t),\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$ for all $1 \le m \le n-1$. Since by definition of canonical ts-DAGs there are no edges into unobservable vertices, all vertices on π are observable. Moreover, again due to definition of canonical ts-DAGs, $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ cannot contain edges with a lag larger than p. These observations require $0 \le |t_{m+1} - t_m| = t_{m+1} - t_m \le p$ and thus shows that both $(k_m, t-(t_{m+1}-t_m))$ and (k_{m+1},t) are vertices in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$. Using Lemma D.14 we therefore get $(k_m, t-(t_{m+1}-t_m)) \in \mathrm{an}((k_{m+1},t),\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}))$, which in turn gives $(k_m, t-(t_{m+1}-t_m)) \in \mathrm{an}((k_{m+1},t),\mathcal{M}^p(\mathcal{D}))$ by Lemma 4.10 and thus $(k_m, t-(t_{m+1}-t_m)) \in \mathrm{an}((k_{m+1},t),\mathcal{D})$ by Lemma C.1 and thus $(k_m,t_m) \in \mathrm{an}((k_{m+1},t_{m+1}),\mathcal{D})$ by repeating ancestral relationships of $\mathcal{D}$. The statement now follows from transitivity of ancestorship. □

LEMMA D.16. *$\mathcal{M}^p(\mathcal{D})$ and $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$ have the same ancestral relationships.*

PROOF OF LEMMA D.16. Combine Lemma C.1 with Lemma D.15. □

D.5.3. *Part 2: Any adjacency in $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$ is also in $\mathcal{M}^p(\mathcal{D})$.* As the second part of the proof of Lemma 4.14 we here show that any adjacency in $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$ is also in $\mathcal{M}^p(\mathcal{D})$. Together with the fact that both these graphs have the same ancestral relationships, as already proven in Section D.5.2, we then get that $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$ is a subgraph of $\mathcal{M}^p(\mathcal{D})$.

LEMMA D.17. *Let (i,t_i) and (j,t_j) with $t-p \le t_i, t_j \le t$ be distinct observable vertices in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ and let $\mathbf{S} \subseteq \mathbf{O}(t-p,t) \setminus \{(i,t_i),(j,t_j)\}$. Then: If (i,t_i) and (j,t_j) are d-connected given $\mathbf{S}$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$, then (i,t_i) and (j,t_j) are d-connected given $\mathbf{S}$ in $\mathcal{D}$.*

REMARK (on Lemma D.17). The statement makes sense because $\mathcal{D}$ and $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ have the same observable time series.

PROOF OF LEMMA D.17. Let (i,t_i) and (j,t_j) be d-connected given $\mathbf{S} \subseteq \mathbf{O}(t-p,t) \setminus \{(i,t_i),(j,t_j)\}$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$. Then, in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ there is path π between (i,t_i) and (j,t_j) that is active given $\mathbf{S}$. Since no node in $\mathbf{S}$ is after t, no node on π is after t because else due to time order of $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ there would be a collider on π after t that, again due to time order, cannot be unblocked by $\mathbf{S}$. Let $((k_1,t_1),\ldots,(k_n,t_n))$ with $(k_1,t_1)=(i,t_i)$ and $(k_n,t_n)=(j,t_j)$ be the ordered sequence of observable vertices on π (not necessarily temporally observed, so some of these vertices may be before $t-p$). Since in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ there are no edges into unobservable vertices and no edges with a lag larger than p, the subpaths $\pi_m = \pi((k_m,t_m),(k_{m+1},t_{m+1}))$ with $1 \leq m \leq n-1$ are either of the form $(k_m,t_m){\rightarrow}(k_{m+1},t_{m+1})$ or $(k_m,t_m){\leftarrow}(k_{m+1},t_{m+1})$ or $(k_m,t_m){\leftarrow}(l_m,t_{l_m}){\rightarrow}(k_{m+1},t_{m+1})$ with (l_m,t_{l_m}) unobservable. In all cases $|t_m - t_{m+1}| \leq p$. Now associate to each π_m a path ρ_m in $\mathcal{D}$ between (k_m,t_m) and (k_{m+1},t_{m+1}) in the following way:

Case 1: If π_m is $(k_m,t_m){\rightarrow}(k_{m+1},t_{m+1})$, then $(k_m,t-(t_{m+1}-t_m)){\rightarrow}(k_{m+1},t)$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ by repeating edges of $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ and hence $(k_m,t-(t_{m+1}-t_m)){\rightarrow}(k_{m+1},t)$ in $\mathcal{M}^p(\mathcal{D})$ by Lemmas D.11 and D.12. According to Lemma D.5 there thus is path between $(k_m,t-(t_{m+1}-t_m))$ and (k_{m+1},t) in $\mathcal{D}$ that is into (k_{m+1},t) and active given $\mathbf{S}_{m,t-t_{m+1}}$, where $\mathbf{S}_{m,t-t_{m+1}}$ is obtained by shifting $\mathbf{S}_m = \mathbf{S} \setminus \{(t_m,k_m),(t_{m+1},k_{m+1})\}$ forward in time by $t-t_{m+1}$ time steps. Let $\mathbf{p}_m$ be the set of all such paths. If any path in $\mathbf{p}_m$ is out of $(k_m,t-(t_{m+1}-t_m))$, then let $\rho_{m,t-t_{m+1}}$ be any such path and let ρ_m the path obtained by shifting $\rho_{m,t-t_{m+1}}$ backwards in time by $t-t_{m+1}$ time steps. If no path $\mathbf{p}_m$ is out of $(k_m,t-(t_{m+1}-t_m))$, then let $\rho_{m,t-t_{m+1}}$ be any path in $\mathbf{p}_m$ and let ρ_m the path obtained by shifting $\rho_{m,t-t_{m+1}}$ backwards in time by $t-t_{m+1}$ time steps. In this latter case $(k_m,t-(t_{m+1}-t_m))$ is an ancestor of $\mathbf{S}_{m,t-t_{m+1}}$ in $\mathcal{D}$ according to Lemma D.6. By repeating ancestral relationships of $\mathcal{D}$ the vertex (k_m,t_m) is then an ancestor of $\mathbf{S}$.

Case 2: If π_m is $(k_m,t_m){\leftarrow}(k_{m+1},t_{m+1})$, do the same as for case 1 with the roles of (k_m,t_m) and (k_{m+1},t_{m+1}) exchanged.

Case 3: If π_m is $(k_m,t_m){\leftarrow}(l_m,t_{l_m}){\rightarrow}(k_{m+1},t_{m+1})$ and $t_m \leq t_{m+1}$, then $(k_m,t-(t_{m+1}-t_m)){\leftarrow}(l_m,t-(t_{m+1}-t_{l_m})){\rightarrow}(k_{m+1},t)$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ and hence $(k_m,t-(t_{m+1}-t_m)){\leftrightarrow}(k_{m+1},t)$ in $\mathcal{M}^p(\mathcal{D})$. According to Lemma D.5 there thus is path between $(k_m,t-(t_{m+1}-t_m))$ and (k_{m+1},t) in $\mathcal{D}$ that is into both $(k_m,t-(t_{m+1}-t_m))$ and (k_{m+1},t) and active given $\mathbf{S}_{m,t-t_{m+1}}$. Let $\rho_{m,t-t_{m+1}}$ be any such path and let ρ_m the path obtained by shifting $\rho_{m,t-t_{m+1}}$ backwards in time by $t-t_{m+1}$ time steps.

Case 4: If π_m is $(k_m,t_m){\leftarrow}(l_m,t_{l_m}){\rightarrow}(k_{m+1},t_{m+1})$ and $t_m > t_{m+1}$, do the same as for case 3 with the roles of (k_m,t_m) and (k_{m+1},t_{m+1}) exchanged.

The paths ρ_m exist due to repeating adjacencies of $\mathcal{D}$ and they are active given $\mathbf{S}_m$ due to repeating separating sets of $\mathcal{D}$. Moreover, due to repeating orientations of $\mathcal{D}$ all edges on ρ_m are oriented in the same way as the corresponding edges on $\rho_{m,t-t_{m+1}}$. Consequently: If (k_{m+1},t_{m+1}) is a collider on π and thus π_m and π_{m+1} meet head-to-head at (k_{m+1},t_{m+1}), then, first, (k_{m+1},t_{m+1}) is an ancestor of $\mathbf{S}$ because π is active given $\mathbf{S}$ and, second, ρ_m and ρ_{m+1} meet head-to-head at (k_{m+1},t_{m+1}). Moreover, if (k_{m+1},t_{m+1}) is a noncollider on π and thus π_m and π_{m+1} do not meet head-to-head at (k_{m+1},t_{m+1}), then, first, (k_{m+1},t_{m+1}) is not in $\mathbf{S}$ because π is active given $\mathbf{S}$ and, second, ρ_m and ρ_{m+1} may or may not meet head-to-head at (k_{m+1},t_{m+1}). Importantly, if they do meet head-to-head, then (k_{m+1},t_{m+1}) is an ancestor of $\mathbf{S}$. By applying Lemma 3.3.1 in Spirtes, Glymour and Scheines (2000) to the ordered sequence of paths $(\rho_1,\ldots,\rho_n)$ we thus obtain a path between $(k_1,t_1)=(i,t_i)$ and $(k_n,t_n)=(j,t_j)$ in $\mathcal{D}$ that is active given $\mathbf{S}$, and hence (i,t_i) and (j,t_j) are d-connected given $\mathbf{S}$ in $\mathcal{D}$. □

LEMMA D.18. *$\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$ is a subgraph of $\mathcal{M}^p(\mathcal{D})$.*

PROOF OF LEMMA D.18. As an immediate consequence of Lemma D.17, every adjacency in $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$ is also in $\mathcal{M}^p(\mathcal{D})$. The statement then follows with Lemma D.16 because the orientation of edges in $\mathcal{M}^p(\mathcal{D})$ and $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$ are uniquely determined by the ancestral relationships. □

D.5.4. *Part 3: Any adjacency in $\mathcal{M}^p(\mathcal{D})$ is also in $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$.* As the third and final part of the proof of Lemma 4.14 we here show that any adjacency in $\mathcal{M}^p(\mathcal{D})$ is also in $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$. We note that the proof of Lemma D.17 crucially relies on the particular form of $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ due to which two subsequent observable vertices on a path in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ are at most p time steps apart and adjacent or almost adjacent. A general ts-DAG $\mathcal{D}$ does, however, not necessarily have these properties, which is why this part of the proof becomes more complicated. We begin by proving the converse of Lemma D.17 restricted to collider-free paths in $\mathcal{D}$.

LEMMA D.19. *Let π be a collider-free path in $\mathcal{D}$ between distinct observable vertices (i,t_i) and (j,t_j) with $t-p \leq t_i, t_j \leq t$. Then, the ordered sequence $((k_1,t_1),\ldots,(k_n,t_n))$ of observable vertices on π with $(k_1,t_1)=(i,t_i)$ and $(k_n,t_n)=(j,t_j)$ has a unique subsequence $((l_1,s_1),\ldots,(l_m,s_m))$ with the following properties:*

1. *$(l_1,s_1)=(i,t_i)$ and $(l_m,s_m)=(j,t_j)$,*
2. *$|s_\alpha - s_{\alpha+1}| \leq p$ for all $1 \leq \alpha \leq m-1$,*
3. *all non-end-point vertices of $\pi((l_\alpha,s_\alpha),(l_{\alpha+1},s_{\alpha+1}))$ are unobservable or are before $\max(s_\alpha,s_{\alpha+1})-p$ for all $1 \leq \alpha \leq m-1$.*

REMARK (on Lemma D.19). While the uniqueness of $((l_1,s_1),\ldots,(l_m,s_m))$ is not needed for the subsequent proofs, we have included it to be able to refer to *the* subsequence $((l_1,s_1),\ldots,(l_m,s_m))$ instead of *a* such subsequence.

PROOF OF LEMMA D.19. Existence:
Assume without loss of generality that $t_i \leq t_j$, else exchange (i,t_i) and (j,t_j). Since π is collider-free, time order of $\mathcal{D}$ thus implies that no vertex on π is after t_j. We now prove the statement by induction over n, where n is the number of observable vertices on π:

Induction base case: $n=2$
In this case, (i,t_i) and (j,t_j) are the only observable vertices on π. Clearly, the sequence $((i,t_i),(j,t_j))$ has the desired properties.

Induction step: $n \mapsto n+1$
In this case, π has $n+1 \geq 3$ observable vertices and the statement has already been proven for paths that have at most n observable vertices. Let π_1 be the subpath of π from (i,t_i) to (k_n,t_n) (the observable vertex on π other than (j,t_j) itself that is closest to (j,t_j) on π) and let π_2 be the subpath of π from (k_2,t_2) (the observable vertex on π other than (i,t_i) itself that is closest to (i,t_i) on π) to (j,t_j). We distinguish three collectively exhaustive cases:

- Case $|t_j - t_n| = t_j - t_n \leq p$:
  This premise implies $|t_i - t_n| \leq p$. Hence, by assumption of induction the statement applies to π_1. The desired sequence is obtained by appending (j,t_j) to the sequence obtained by applying the statement to π_1.
- Case $|t_j - t_2| = t_j - t_2 \leq p$:
  By assumption of induction the statement then applies to π_2. Moreover, $|t_i - t_2| \leq p$. The desired sequence is obtained by prepending (i,t_i) to the sequence obtained by applying the statement to π_2.

- Case $|t_j - t_n| = t_j - t_n > p$ and $|t_j - t_2| = t_j - t_2 > p$:
  Since π is collider-free and $\mathcal{D}$ is time ordered, this premise implies that all observable non-end-point vertices on π are before $t_j - p$. Hence, the sequence $((i,t_i),(j,t_j))$ has the desired properties.

Uniqueness:
Let $((l_1,s_1),\dots,(l_m,s_m))$ and $((l'_1,s'_1),\dots,(l'_{m'},s_{m'}))$ be two such subsequences. We proof their equality by induction over α, where α is the index of the subsequences.

Induction base case: $\alpha = 1$
The equality $(l_1,s_1) = (l'_1,s'_1)$ follows due to the first property demanded in Lemma D.19 applied to both sequences.

Induction step: $\alpha \mapsto \alpha + 1 \leq \min(m,m')$
By the assumption of induction $(l_q,s_q) = (l'_q,s'_q)$ for all $1 \leq q \leq \alpha$ is given, and we have to show $(l_{\alpha+1},s_{\alpha+1}) = (l'_{\alpha+1},s'_{\alpha+1})$. Assume the opposite, i.e., assume $(l_{\alpha+1},s_{\alpha+1}) \neq (l'_{\alpha+1},s'_{\alpha+1})$. Without loss of generality further assume that $(l_{\alpha+1},s_{\alpha+1})$ is on the subpath $\pi((l'_{\alpha+1},s'_{\alpha+1}),(j,t_j))$, else exchange the two subsequences. Let r with $\alpha < r < m'$ be such that the vertices (l'_q,s'_q) with $\alpha < q \leq r$ are non-end-point vertices on $\pi((l_\alpha,s_\alpha),(l_{\alpha+1},s_{\alpha+1}))$ and (l'_{r+1},s'_{r+1}) is on $\pi((l_{\alpha+1},s_{\alpha+1}),(j,t_j))$. Such an r exists because both the sequence $((l_1,s_1),\dots,(l_m,s_m))$ and the sequence $((l'_1,s'_1),\dots,(l'_{m'},s_{m'}))$ are subsequences of $((k_1,t_1),\dots,(k_n,t_n))$. Note that the (l'_q,s'_q) with $\alpha < q \leq r$ are observable because all vertices of the subsequence $((l'_1,s'_1),\dots,(l'_{m'},s_{m'}))$ are observable. The third property demanded in Lemma D.19 applied to the sequence $((l_1,s_1),\dots,(l_m,s_m))$ thus requires $s'_q < \max(s_\alpha,s_{\alpha+1}) - p$ for all $\alpha < q \leq r$. We distinguish two cases:

- Suppose $s_{\alpha+1} \leq s_\alpha$. Then $s'_q < s_\alpha - p$ for all $\alpha < q \leq r$. For $q = \alpha + 1$ we thus get $s'_{\alpha+1} < s_\alpha - p$, which contradicts the second property demanded in Lemma D.19 applied to $((l'_1,s'_1),\dots,(l'_{m'},s_{m'}))$.
- Suppose $s_{\alpha+1} > s_\alpha$. Then $s'_q < s_{\alpha+1} - p$ for all $\alpha < q \leq r$. By time order of $\mathcal{D}$ in combination with the facts that π is collider-free and that (l_α,s_α) is on $\pi((i,t_i),(l_{\alpha+1},s_{\alpha+1}))$, the premise $s_{\alpha+1} > s_\alpha$ implies that $(l_{\alpha+1},s_{\alpha+1})$ is on π between the root node of π and (j,t_j). Consequently, all vertices on $\pi((l_{\alpha+1},s_{\alpha+1}),(j,t_j))$ are not before $s_{\alpha+1}$. Since (l'_{r+1},s'_{r+1}) is on $\pi((l_{\alpha+1},s_{\alpha+1}),(j,t_j))$ we thus find $s_{\alpha+1} \leq s'_{r+1}$, which implies $s'_q < s'_{r+1} - p$ for all $\alpha < q \leq r$. For $q = r$ we thus get $s'_r < s'_{r+1} - p$, which contradicts the second property demanded in Lemma D.19 applied to $((l'_1,s'_1),\dots,(l'_{m'},s_{m'}))$.

Consequently, the assumption $(l_{\alpha+1},s_{\alpha+1}) \neq (l'_{\alpha+1},s'_{\alpha+1})$ leads to a contradiction and $(l_{\alpha+1},s_{\alpha+1}) = (l'_{\alpha+1},s'_{\alpha+1})$ must be true.

This induction terminates when $\alpha + 1 = \min(m,m')$. The first property demanded in Lemma D.19 applied to both sequences then requires $m = m'$ and $(l_m,s_m) = (l'_m,s'_m) = (j,t_j)$, which completes the proof □

REMARK (on the proof of Lemma D.19). For the special case in which π is a directed path we can also immediately see that the sequence $((k_1,t_1),\dots,(k_n,t_n))$ of all observable nodes on π has the desired properties. We need the above proof to also cover the case in which π is into both (i,t_i) and (j,t_j).

LEMMA D.20. *Let π be a collider-free path in $\mathcal{D}$ between distinct observable vertices (i,t_i) and (j,t_j) with $t - p \leq t_i, t_j \leq t$ and let $((l_1,s_1),\dots,(l_m,s_m))$ be as in Lemma D.19 applied to π. Then, (l_α,s_α) and $(l_{\alpha+1},s_{\alpha+1})$ are adjacent or almost adjacent in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ for all $1 \leq \alpha \leq m - 1$.*

PROOF OF LEMMA D.20. Let π_α be the path in $\mathcal{D}$ obtained by shifting the subpath $\pi((l_\alpha, s_\alpha), (l_{\alpha+1}, s_{\alpha+1}))$ forward in time by $t - \max(s_\alpha, s_{\alpha+1}) \geq 0$ time steps. This π_α is a path between the vertices $v_1 = (l_\alpha, t - (\max(s_\alpha, s_{\alpha+1}) - s_\alpha))$ and $v_2 = (l_\alpha, t - (\max(s_\alpha, s_{\alpha+1}) - s_{\alpha+1}))$, both of which are observable because all vertices in the sequence $((l_1, s_1), \ldots, (l_m, s_m))$ are observable and within $[t-p, t]$ because of part 2 of Lemma D.19. Moreover, π_α is collider-free because π is collider-free. All non-end-point vertices on π_α are unobserved, i.e., unobservable or temporally unobserved due to part 3 of Lemma D.19. Consequently, π_α cannot be blocked by any set of observed variables, which is why v_1 and v_2 are adjacent in $\mathcal{M}^p(\mathcal{D})$. The statement then follows from Lemma D.11. □

DEFINITION D.21 (Canonically induced path). Let π be a collider-free path in $\mathcal{D}$ between distinct observable vertices (i, t_i) and (j, t_j) with $t - p \leq t_i, t_j \leq t$ and let $((l_1, s_1), \ldots, (l_m, s_m))$ be as in Lemma D.19 applied to π. The *canonical path induced by* π, denoted π_{ci}, is the (unique) path in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ between (i, t_i) and (j, t_j) with the following properties:

1. For all $1 \leq \alpha \leq m$ the vertex (l_α, s_α) is on π_{ci}.
2. For all $1 \leq \alpha \leq m-1$ the subpath $\pi_{ci}((l_\alpha, s_\alpha)), (l_{\alpha+1}, s_{\alpha+1}))$ is
   a) $(l_\alpha, s_\alpha) {\rightarrow} (l_{\alpha+1}, s_{\alpha+1})$ if and only if $(l_\alpha, s_\alpha) \in \mathrm{an}((l_{\alpha+1}, s_{\alpha+1}), \mathcal{D})$,
   b) $(l_\alpha, s_\alpha) {\leftarrow} (l_{\alpha+1}, s_{\alpha+1})$ if and only if $(l_{\alpha+1}, s_{\alpha+1}) \in \mathrm{an}((l_\alpha, s_\alpha), \mathcal{D})$,
   c) $(l_\alpha, s_\alpha) {\leftarrow} (l'_\alpha, s'_\alpha) {\rightarrow} (l_{\alpha+1}, s_{\alpha+1})$ with (l'_α, s'_α) unobservable if and only if $(l_\alpha, s_\alpha) \notin \mathrm{an}((l_{\alpha+1}, s_{\alpha+1}), \mathcal{D})$ and $(l_{\alpha+1}, s_{\alpha+1}) \notin \mathrm{an}((l_\alpha, s_\alpha), \mathcal{D})$.

REMARK (on Definition D.21). Existence follows from Lemma D.12 because according to Lemma D.20 the vertices (l_α, s_α) and $(l_{\alpha+1}, s_{\alpha+1})$ are adjacent or almost adjacent for all $1 \leq \alpha \leq m-1$. Uniqueness follows from Lemma D.10 because in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ there is at most one edge between any pair of vertices.

LEMMA D.22. *Let π be a collider-free path in $\mathcal{D}$ between distinct observable vertices (i, t_i) and (j, t_j) with $t - p \leq t_i, t_j \leq t$ and let π_{ci} be the canonical path induced by π. Then:*

1. *All observable vertices on π_{ci} are also on π.*
2. *If (k_1, t_1), (k_2, t_2) and (k_3, t_3) are distinct observable vertices on π_{ci} and (k_2, t_2) is on $\pi_{ci}((k_1, t_1), (k_3, t_3))$, then (k_2, t_2) is on $\pi((k_1, t_1), (k_3, t_3))$.*
3. *π_{ci} is a collider-free.*
4. *If (k_1, t_1) and (k_2, t_2) are distinct observable vertices on π_{ci}, then $\pi((k_1, t_1), (k_2, t_2))$ is an inducing path relative to $\mathbf{O}(\max(t_1, t_2) - p, t)[\pi_{ci}((k_1, t_1), (k_2, t_2))]$.*
5. *If π is active given $\mathbf{S}$ in $\mathcal{D}$, then π_{ci} is active given $\mathbf{S}$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$.*

PROOF OF LEMMA D.22. Let $((l_1, s_1), \ldots, (l_m, s_m))$ be as in Lemma D.19 applied to π.

1. The definition of canonically induced paths is such that $((l_1, s_1), \ldots, (l_m, s_m))$ is the sequence of all observable vertices on π_{ci}. Moreover, all of $(l_1, s_1), \ldots, (l_m, s_m)$ are on π by construction, see Lemma D.19.

2. Since (k_2, t_2) is on $\pi_{ci}((k_3, t_3), (k_1, t_1))$ if it is on $\pi_{ci}((k_1, t_1), (k_3, t_3))$, we can without loss of generality assume that (k_1, t_1) is closer to (i, t_i) on π_{ci} than (k_3, t_3) is to (i, t_i) on π_{ci}. Hence, there are $1 \leq \alpha_1 < \alpha_2 < \alpha_3 \leq m$ such that $(k_1, t_1) = (l_{\alpha_1}, s_{\alpha_1})$, $(k_2, t_2) = (l_{\alpha_2}, s_{\alpha_2})$ and $(k_3, t_3) = (l_{\alpha_3}, s_{\alpha_3})$. Moreover, by definition of $((l_1, s_1), \ldots, (l_m, s_m))$ and canonically induced paths, (l_α, s_α) is closer to (i, t_i) on π than (l_q, s_q) with $\alpha < q$ is to (i, t_i) on π.

3. Assume there is a collider on π_{ci}. Since in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ there are no edges into unobservable vertices, all colliders on π_{ci} are observable. Hence, there must be $1 < \alpha < m$ such that (l_α, s_α) is a collider on π_{ci}, i.e., such that both $\pi_{ci}((l_{\alpha-1}, s_{\alpha-1}), (l_\alpha, s_\alpha))$ and

$\pi_{ci}((l_\alpha, s_\alpha), (l_{\alpha+1}, s_{\alpha+1}))$ are into (l_α, s_α). In particular, $\pi_{ci}((l_{\alpha-1}, s_{\alpha-1}), (l_\alpha, s_\alpha))$ is not $(l_{\alpha-1}, s_{\alpha-1}){\leftarrow}(l_\alpha, s_\alpha)$ and $\pi_{ci}((l_\alpha, s_\alpha), (l_{\alpha+1}, s_{\alpha+1}))$ is not $(l_\alpha, s_\alpha){\rightarrow}(l_{\alpha+1}, s_{\alpha+1})$.

Assume $\pi((l_{\alpha-1}, s_{\alpha-1}), (l_\alpha, s_\alpha))$ is out of (l_α, s_α). Since π is collider-free, it then follows that $\pi((l_{\alpha-1}, s_{\alpha-1}), (l_\alpha, s_\alpha))$ is directed from (l_α, s_α) to $(l_{\alpha-1}, s_{\alpha-1})$ and hence $(l_\alpha, s_\alpha) \in \mathrm{an}((l_{\alpha-1}, s_{\alpha-1}), \mathcal{D})$. According to Definition D.21 this ancestral relationships requires $\pi_{ci}((l_{\alpha-1}, s_{\alpha-1}), (l_\alpha, s_\alpha))$ to be $(l_{\alpha-1}, s_{\alpha-1}){\leftarrow}(l_\alpha, s_\alpha)$, which is a contradiction. Hence $\pi((l_{\alpha-1}, s_{\alpha-1}), (l_\alpha, s_\alpha))$ is into (l_α, s_α).

We similarly we find that $\pi((l_\alpha, s_\alpha), (l_{\alpha+1}, s_{\alpha+1}))$ is into (l_α, s_α) and thus that (l_α, s_α) is a collider on π, a contradiction.

4. We may without loss of generality assume that (k_1, t_1) is closer to (i, t_i) on π than (k_2, t_2) is to (i, t_i) on π. Write $t_{12} = \max(t_1, t_2)$.

Since π is collider-free, also its subpath $\pi((k_1, t_1), (k_2, t_2))$ is collider-free. For showing that $\pi((k_1, t_1), (k_2, t_2))$ is an inducing path relative to $\mathbf{O}(t_{12} - p, t)[\pi_{ci}((k_1, t_1), (k_2, t_2))]$ it is thus sufficient to show that none of its non-end-point vertices is an element of the set $\mathbf{O}(t_{12} - p, t)[\pi_{ci}((k_1, t_1), (k_2, t_2))]$. To this end, let (k_3, t_3) be a non-end-point vertex on $\pi((k_1, t_1), (k_2, t_2))$.

Since $((l_1, s_1), \ldots, (l_m, s_m))$ is the sequence of all observable vertices on π_{ci}, there are α_1 and α_2 with $1 \leq \alpha_1 < \alpha_2 \leq m$ such that $(k_1, t_1) = (l_{\alpha_1}, s_{\alpha_1})$ and $(k_2, t_2) = (l_{\alpha_2}, s_{\alpha_2})$. Therefore, either (k_3, t_3) equals $(l_{\alpha_3}, s_{\alpha_3})$ for some $\alpha_1 < \alpha_3 < \alpha_2$ or (k_3, t_3) is a non-end-point vertex on $\pi((l_{\alpha_3}, s_{\alpha_3}), (l_{\alpha_3+1}, s_{\alpha_3+1}))$ for some α_3 with $\alpha_1 \leq \alpha_3 < \alpha_2$. In the former case, (k_3, t_3) is not in $\mathbf{O}(t_{12} - p, t_{12})[\pi_{ci}((k_1, t_1), (k_2, t_2))]$ because it is a non-end-point vertex on $\pi_{ci}((k_1, t_1), (k_2, t_2))$. In the latter case, according to part 3 of Lemma D.19, (k_3, t_3) is unobservable or before $\max(s_{\alpha_3}, s_{\alpha_3+1}) - p$. Because π is collider-free and both $(l_{\alpha_3}, s_{\alpha_3})$ and $(l_{\alpha_3+1}, s_{\alpha_3+1})$ are on $\pi((k_1, t_1), (k_2, t_2))$, time order of $\mathcal{D}$ implies $\max(s_{\alpha_3}, s_{\alpha_3+1}) \leq t_{12}$. Thus, (k_3, t_3) is not in $\mathbf{O}(t_{12} - p, t)[\pi_{ci}((k_1, t_1), (k_2, t_2))]$.

5. This claim follows from parts 1 and 2 of Lemma D.22: Since π is collider-free and active given $\mathbf{S}$, no vertex on π is in $\mathbf{S}$. Thus, since all observable vertices on π_{ci} are also on π, no vertex on π_{ci} is in $\mathbf{S}$. Since π_{ci} is collider-free this observation shows that π_{ci} is active given $\mathbf{S}$. □

LEMMA D.23. *Let (i, t_i) and (j, t_j) with $t - p \leq t_i, t_j \leq t$ be distinct observable vertices in $\mathcal{D}$ and let $\mathbf{S} \subseteq \mathbf{O}(t - p, t) \setminus \{(i, t_i), (j, t_j)\}$. Then: If in $\mathcal{D}$ there is a collider-free path π between (i, t_i) and (j, t_j) that is active given $\mathbf{S}$, then (i, t_i) and (j, t_j) are d-connected given $\mathbf{S}$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$.*

PROOF OF LEMMA D.23. According to part 5 of Lemma D.22 the canonically induced path π_{ci} of π d-connects (i, t_i) and (j, t_j) given $\mathbf{S}$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$. □

In order to show that any adjacency in $\mathcal{M}^p(\mathcal{D})$ is also in $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$—and thus to finish the proof of Lemma 4.14—it remains to prove a statement that extends Lemma D.23 to the case in which π, the d-connecting path in $\mathcal{D}$, has $n_c \geq 1$ colliders $c_1, \ldots, c_{n_c}$ (ordered starting with the collider closest to (i, t_i)). One might think that such a generalization follows readily now, namely by cutting π into $n_c + 1$ collider-free paths $\pi^{a,a+1} = \pi(c_a, c_{a+1})$ with $0 \leq a \leq n_c$, where we let $c_0 = (i, t_i)$ and $c_{n_c+1} = (j, t_j)$, and then applying Lemma 3.3.1 in Spirtes, Glymour and Scheines (2000) to the canonically induced paths $\pi^{a,a+1}_{ci}$ of $\pi^{a,a+1}$ with $0 \leq a \leq n_c$. While we do use a similar approach, the proof is complicated by two facts: First, the canonically induced path $\pi^{a,a+1}_{ci}$ of $\pi^{a,a+1}$ only exists if $\pi^{a,a+1}$ is between *observable* vertices that are *at most p time steps apart*, see Definition D.21. However, some of the colliders on π, i.e., some of the c_a with $a \leq 1 \leq n_c$ might be unobservable and/or more than p time steps apart from the neighboring colliders or end-point vertices of π, i.e., from c_{a-1}

or c_{a+1}. Second, even if $\pi_{ci}^{a,a+1}$ exists, it may be *out of* one of its end-point vertices although $\pi^{a,a+1}$ is into this vertex. In case this vertex is a collider on π and an element of $\mathbf{S}$, Lemma 3.3.1 in Spirtes, Glymour and Scheines (2000) does not apply. We will address the first of these complications by noting that, in order for π to be active given $\mathbf{S}$, every collider on π must be an ancestor of $\mathbf{S}$ and thus an ancestor of an observed vertex within the time window $[t-p,t]$. Hence, in $\mathcal{D}$ there are directed paths from the c_a with $a \leq 1 \leq n_c$ to some observed vertices. By joining these directed paths with the $\pi^{a,a+1}$ we get collider-free paths $\tilde{\pi}^{a,a+1}$ in $\mathcal{D}$ between observed vertices, the canonically induced paths $\tilde{\pi}_{ci}^{a,a+1}$ of which exist.

DEFINITION D.24 (Collider extension structure). Let π be a (non collider-free) path between the distinct observable vertices (i,t_i) and (j,t_j) with $t-p \leq t_i, t_j \leq t$ in $\mathcal{D}$ that is active given $\mathbf{S} \subseteq \mathbf{O}(t-p,t) \setminus \{(i,t_i),(j,t_j)\}$. Let $c_1,\ldots,c_{n_c}$ with $n_c \geq 1$ be the collider on π, ordered starting with the collider closest to (i,t_i). A *collider extension structure of* π *with respect to* $\mathbf{S}$ *and* $\mathbf{O}(t-p,t)$ is a collection of paths $\rho^1,\ldots,\rho^{n_c}$ such that for all $1 \leq a \leq n_c$ and $1 \leq b \leq n_c$ with $a \neq b$ all of the following holds:

1. One of these options holds:
   a) ρ^a is the trivial path consisting of $c_a = d_a$ only and $c_a = d_a \in \mathbf{O}(t-p,t)$.
   b) ρ^a is a nontrivial directed path from c_a to some vertex $d_a \in \mathbf{O}(t-p,t)$.
2. If v is on ρ^a and in $\mathbf{S}$, then $v = d_a$.
3. d_a is an ancestor of $\mathbf{S}$.
4. ρ^a intersects with π at c_a only.
5. ρ^a and ρ^b do not intersect.

LEMMA D.25. *Given the assumptions and notation of Definition D.24, one of the following statements holds:*

1. *There is a collider extension structure of* π *with respect to* $\mathbf{S}$ *and* $\mathbf{O}(t-p,t)$.
2. *In* $\mathcal{D}$ *there is a path* π' *between* (i,t_i) *and* (j,t_j) *with at most* $n_c - 1$ *colliders that is active given* $\mathbf{S}$.

PROOF OF LEMMA D.25. We divide the proof into three steps.

Step 1: The first, second and third property of collider extension structures holds

Let $1 \leq a \leq n_c$. Since c_a is a collider on π and π is active given $\mathbf{S}$, there is $S_a \in \mathbf{S}$ (which may be equal to c_a) such that $c_a \in \mathrm{an}(S_a,\mathcal{D})$. Hence, there is a (possibly trivial, namely if and only if $c_a = S_a$) directed path λ^a from c_a to S_a. On λ^a let d_a be the vertex closest to c_a that is in $\mathbf{O}(t-p,t)$ (which may be c_a itself). This vertex d_a exists because $S_a \in \mathbf{S} \subseteq \mathbf{O}(t-p,t)$, such that d_a is an ancestor of $\mathbf{S}$ by means of the subpath $\lambda(d_a,S_a)$. Let ρ^a be the subpath $\lambda^a(c_a,d_a)$, which is the trivial path consisting of the single vertex $c_a = d_a \in \mathbf{O}(t-p,t)$ or is a nontrivial directed path from c_a to $d_a \in \mathbf{O}(t-p,t)$. Moreover, by definition of d_a together with the fact that $\mathbf{S} \subseteq \mathbf{O}(t-p,t)$, no vertex on ρ^a other than d_a is in $\mathbf{S}$. The collection of paths $\rho^1,\ldots,\rho^{n_c}$ thus fulfills the first three conditions of a collider extension structure of π with respect to $\mathbf{S}$ and $\mathbf{O}(t-p,t)$.

The first two of these conditions have two immediate implications that will be important later in this proof: First, if ρ^a is nontrivial then $c_a \notin \mathbf{S}$. Second, if ρ^a is nontrivial then it is active given $\mathbf{S} \setminus \{d_a\}$.

Step 2: The fourth property of collider extension structures or existence of π'

Assume there is $1 \leq a \leq n_c$ such that π and ρ^a do not intersect at c_a only. Then, ρ^a must be nontrivial and hence we get $c_a \notin \mathbf{S}$. Let e_a be the vertex on ρ^a closest to c_a other than c_a itself that is also on π. If e_a is on $\pi((i,t_i),c_a)$, then let $v_1 = (i,t_i)$ and $v_2 = (j,t_j)$, else let $v_1 = (j,t_j)$ and $v_2 = (i,t_i)$. Let π' be the concatenation $\pi(v_1,e_a) \oplus \rho^a(e_a,c_a) \oplus \pi(c_a,v_2)$.

By definition of e_a, π' is a path (rather than a walk) in $\mathcal{D}$ between (i, t_i) and (j, t_j). We now show that π' is active given $\mathbf{S}$ and has at most $n_c - 1$ colliders.

All colliders on π' are ancestors of $\mathbf{S}$*:* Since $\rho^a(e_a, c_a)$ is a nontrivial directed path from c_a to e_a, every collider on π' is a collider on $\pi(v_1, e_a)$ or a collider on $\pi(c_a, v_2)$ or equals e_a. Every collider on $\pi(v_1, e_a)$ or $\pi(c_a, v_2)$ is a collider on π and hence, because π is active given $\mathbf{S}$, an ancestor of $\mathbf{S}$. Lastly, e_a is an ancestor of $S_a \in \mathbf{S}$ by means of the path $\lambda^a(e_a, S_a)$.

No non-end-point noncollider on π' is in $\mathbf{S}$*:* All non-end-point noncolliders on $\pi(v_1, e_a)$ or $\pi(c_a, v_2)$ are non-end-point noncolliders on π and hence, because π is active given $\mathbf{S}$, not in $\mathbf{S}$. All vertices on $\rho^a(e_a, c_a)$ other than, perhaps, e_a are not in $\mathbf{S}$ because, as shown in step 1 of this proof, all vertices on ρ^a other than d_a are not in $\mathbf{S}$. Lastly, assume that e_a is in $\mathbf{S}$ and a non-end-point noncollider on π'. Because $\rho^a(e_a, c_a)$ is into e_a, this assumption requires that $\pi(v_1, e_a)$ is a nontrivial path out of e_a. Consequently, e_a is a non-end-point noncollider on π, which is a contradiction because $e_a \in \mathbf{S}$ and π is active given $\mathbf{S}$.

Number of colliders: There are no colliders on $\rho^a(e_a, c_a)$ because it is a directed path. If $v_1 = (i, t_i)$, then there are at most $a - 1$ colliders on $\pi(v_1, e_a)$ and exactly $n_c - a$ colliders on $\pi(c_a, v_2)$. If $v_1 = (j, t_j)$, then there are at most $n_c - a$ colliders on $\pi(v_1, e_a)$ and exactly $a - 1$ colliders on $\pi(c_a, v_2)$. The junction point c_a is a noncollider on π' because $\rho^a(e_a, c_a)$ is out of c_a. Regarding e_a, there are two cases:

1. First, assume e_a is a noncollider on π'. Then, there are at most $(a-1)+(n_c-a) = n_c - 1$ colliders on π'.
2. Second, assume e_a is a collider on π'. This assumption requires $\pi(v_1, e_a)$ to be into e_a. Let r be that particular root node on $\pi(v_1, c_a)$ which is closest to c_a on π. Then, $\pi(r, c_a)$ is nontrivial (because c_a is a collider on π and hence not a root on π) and directed from r to c_a (by combining the facts that c_a is a collider on π, that r is a root on π, and that no other root node on π is between r and c_a). Assume that e_a is on $\pi(r, c_a)$. Then, $\pi(e_a, c_a)$ would be a nontrivial (because $e_a \neq c_a$) directed path from e_a to c_a, which contradicts acyclicity because c_a is an ancestor e_a by means of $\rho^a(c_a, e_a)$. Hence, e_a is on $\pi(v_1, r)$. Since $\pi(v_1, e_a)$ is into e_a, we thus see that e_a is on $\pi(v_1, c_{a-1})$ if $v_1 = (i, t_i)$ and that e_a is on $\pi(c_{a+1}, v_1)$ if $v_1 = (j, t_j)$. Consequently, there are at most $a - 2$ colliders on $\pi(v_1, e_a)$ if $v_1 = (i, t_i)$ and there are most $n_c - (a + 1)$ colliders on $\pi(v_1, e_a)$ if $v_1 = (j, t_j)$. In summary, there are most $(a-2) + (n_c - a) + 1 = (n_c - (a+1)) + (a-1) + 1 = n_c - 1$ colliders on π'.

Thus, if the collection of paths $\rho^1, \ldots, \rho^{n_c}$ does not fulfill the fourth condition of a collider extension structure of π with respect to $\mathbf{S}$ and $\mathbf{O}(t-p, t)$, then there is path π' as in point 2 of this lemma. To complete this proof it is therefore sufficient to show the following statement: If the collection of paths $\rho^1, \ldots, \rho^{n_c}$ fulfills the first four conditions of a collider extension structure of π with respect to $\mathbf{S}$ and $\mathbf{O}(t-p, t)$, then this collection of paths also fulfills the fifth condition of a collider extension structure (and hence is a collider extension structure) or there is path π' as in point 2 of the lemma.

Step 3: The fifth property of collider extension structures holds or existence of π'

Assume there are $1 \leq a, b \leq n_c$ with $a < b$ such that ρ^a and ρ^b intersect. Then, at least one of these paths must be nontrivial because $c_a \neq c_b$. If one of them is trivial and the other one is nontrivial, say ρ^a is trivial and ρ^b is nontrivial, then ρ^b must contain c_a and hence intersects with π at a vertex other than c_a, namely c_b. Since this conclusion violates the fourth condition of a collider extension structure of π with respect to $\mathbf{S}$ and $\mathbf{O}(t-p, t)$, as explained in the last paragraph of step 2 we do not need to consider this situation. Consequently, we can assume that both ρ^a and ρ^b are nontrivial and intersect π at, respectively, c_a and c_b only.

The fact that both ρ^a and ρ^b are nontrivial implies $\mathbf{S} = \mathbf{S} \setminus \{c_a, c_b\}$, see step 1 of this proof. Let f_{ab} be the vertex closest to c_a on ρ^a that is also on ρ^b. Because ρ^a and ρ^b respectively intersect π at c_a and c_b only, f_{ab} is neither c_a nor c_b and both $\rho^a(c_a, f_{ab})$ and $\rho^b(f_{ab}, c_b)$ are nontrivial paths. Moreover, f_{ab} is an ancestor of $\mathbf{S}$ by means of the directed path $\lambda^a(f_{ab}, S_a)$ from f_{ab} to $S_a \in \mathbf{S}$. Let π' be the concatenation $\pi((i,t_i), c_a) \oplus \rho^a(c_a, f_{ab}) \oplus \rho^b(f_{ab}, c_b) \oplus \pi(c_b, (j,t_j))$, which by the assumptions on ρ^a and ρ^b as well as the definition of f_{ab} is a path (rather than a walk) in $\mathcal{D}$ between (i,t_i) and (j,t_j). Consider the four constituting subpaths of π':

1. First, $\pi((i,t_i), c_a)$ is active given $\mathbf{S} \setminus \{(i,t_i), c_a\} = \mathbf{S}$ because π is active given $\mathbf{S}$.
2. Second and similar to the first point, $\pi(c_b, (j,t_j))$ is active given $\mathbf{S} \setminus \{c_b, (j,t_j)\}$.
3. Third, $\rho^a(c_a, f_{ab})$ is active given $\mathbf{S} \setminus \{d_a\}$ since ρ^a is active given $\mathbf{S} \setminus \{d_a\}$ (see the last paragraph in part 1 of this proof). There are two cases:
   a) If $f_{ab} = d_a$, then $\rho^a(c_a, f_{ab})$ is active given $\mathbf{S} \setminus \{c_a, f_{ab}\} = \mathbf{S} \setminus \{f_{ab}\} = \mathbf{S} \setminus \{d_a\}$.
   b) If $f_{ab} \neq d_a$, then $f_{ab} \notin \mathbf{S}$ (because no vertex on ρ^a other than d_a is in $\mathbf{S}$) and d_a is not on $\rho^a(c_a, f_{ab})$. Because $\rho^a(c_a, f_{ab})$ is collider-free, we thus see that $\rho^a(c_a, f_{ab})$ is active given $\mathbf{S} \setminus \{c_a, f_{ab}\} = \mathbf{S} = \{d_a\} \cup (\mathbf{S} \setminus \{d_a\})$.

   Thus, $\rho^a(c_a, f_{ab})$ is active given $\mathbf{S} \setminus \{c_a, f_{ab}\}$.
4. Fourth and similar to the third point, $\rho^a(f_{ab}, c_b)$ is active given $\mathbf{S} \setminus \{f_{ab}, c_b\}$.

Since the junction points c_a and c_b are noncolliders on π' and not in $\mathbf{S}$, whereas the third junction point f_{ab} is a collider on π' and an ancestor of $\mathbf{S}$, Lemma 3.3.1 in Spirtes, Glymour and Scheines (2000) asserts that π' is active given $\mathbf{S}$. Lastly, there are exactly $(a-1)+1+(n_c-b) = n_c-(b-a) \leq n_c-1$ colliders on π'.

Thus, if $\rho^1, \ldots, \rho^{n_c}$ fulfills the first four conditions of a collider extension structure of π with respect to $\mathbf{S}$ and $\mathbf{O}(t-p,t)$, then it also fulfills the fifth condition or there is a path π' as in point 2 of the lemma. □

By induction over the number of colliders n_c, using Lemma D.23 as the induction base case and Lemma D.25 for the induction step, we thus arrive at the following conclusion: For the purpose of proving Lemma 4.14 it is thus sufficient to consider d-connecting paths π in $\mathcal{D}$ for which there is a collider extension structure of π with respect to $\mathbf{S}$ and $\mathbf{O}(t-p,t)$. This reasoning allows to overcome the first complication mentioned above in the following way (see Lemma D.29).

DEFINITION D.26 (Notation for remaining parts of the proof). Given the assumptions and notation of Definition D.24, let $\rho^1, \ldots, \rho^{n_c}$ be a collider extension structure of π with respect to $\mathbf{S}$ and $\mathbf{O}(t-p,t)$. We let ρ^0 and ρ^{n_c+1} be the trivial paths that, respectively, consist of $c_0 = d_0 = (i,t_i)$ and $c_{n_c+1} = d_{n_c+1} = (j,t_j)$ only. Moreover, for all $0 \leq a_1 < a_2 \leq n_c+1$ we let $\pi^{a_1,a_2} = \pi(c_{a_1}, c_{a_2})$ and $\tilde{\pi}^{a_1,a_2} = \rho^{a_1}(d_{a_1}, c_{a_1}) \oplus \pi^{a_1,a_2} \oplus \rho^{a_2}$.

REMARK (on Definition D.26). The concatenation $\tilde{\pi}^{a_1,a_2}$ is a path (rather than a walk) in $\mathcal{D}$ according to the fourth and fifth property of collider extension structures, and both end-points of $\tilde{\pi}^{a_1,a_2}$ are in $\mathbf{O}(t-p,t)$ (according to the first property of collider extension structures).

LEMMA D.27. *Given the assumptions and notation of Definition D.26, it holds for all $0 \leq a \leq n_c$:*

1. $\tilde{\pi}^{a,a+1}$ *is active given* $\mathbf{S} \setminus \{d_a, d_{a+1}\}$.
2. $\tilde{\pi}^{a,a+1}_{ci}$*, the canonical induced path of* $\tilde{\pi}^{a,a+1}$*, is active given* $\mathbf{S} \setminus \{d_a, d_{a+1}\}$.

REMARK (on Lemma D.27). These two statements concern different graphs: The first statement is with respect to $\mathcal{D}$, the second statement is with respect to $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$.

PROOF OF LEMMA D.27. **1.** Since ρ^a and ρ^{a+1} are trivial paths or nontrivial paths out of, respectively, c_a and c_{a+1} and since $\pi^{a,a+1}$ is collider-free, $\tilde{\pi}^{a,a+1}$ is collider-free. Thus, assuming that $\tilde{\pi}^{a,a+1}$ is blocked given $\mathbf{S} \setminus \{d_a, d_{a+1}\}$, there is a non-end-point noncollider v on $\tilde{\pi}^{a,a+1}$ with $v \in \mathbf{S} \setminus \{d_a, d_{a+1}\}$. This vertex cannot be on $\rho^a(d_a, c_a)$ because according to the definition of collider extension structures the opposite requires $v = d_a$. Since v can similarly not be on ρ^{a+1}, v must be a non-end-point vertex on π^{a_1,a_2}. However, then v must be a non-end-point noncollider on π, which is a contradiction because $v \in \mathbf{S}$ and π is active given $\mathbf{S}$.

2. Since $\tilde{\pi}^{a,a+1}$ is collider-free, as shown in the proof of part 1 of Lemma D.27, and since both d_a and d_{a+1} are by definition in $\mathbf{O}(t-p,t)$, the canonical induced path $\tilde{\pi}^{a,a+1}_{ci}$ exists. The statement follows by combining part 1 of Lemma D.27 with part 5 of Lemma D.22. □

At this point, we face the second complication mentioned above: We can now *not* straight away apply Lemma 3.3.1 in Spirtes, Glymour and Scheines (2000) to the ordered sequence of paths $\tilde{\pi}^{0,1}_{ci}, \ldots, \tilde{\pi}^{n_c,n_c+1}_{ci}$ in order to infer the existence of a path between $d_0 = (i, t_i)$ and $d_{n_c+1} = (j, t_j)$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ that is active given $\mathbf{S}$. To recall, the reason is that $\tilde{\pi}^{a,a+1}_{ci}$ may be out of one of its end-point vertices although $\tilde{\pi}^{a,a+1}$ is into this vertex. To resolve this complication, we now show that such a situation requires the existence of a certain inducing path in $\mathcal{D}$ and hence the existence of an additional edge in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ which can be used to bypass that vertex.

DEFINITION D.28 (Canonical paths). Let π be a path in $\mathcal{D}$ between distinct observable vertices (i, t_i) and (j, t_j) with $t - p \le t_i, t_j \le p$. A path π_c between (i, t_i) and (j, t_j) in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ is *canonical with respect to* π if all of the following holds:

1. All observable vertices on π_c are also on π.
2. If (k_1, t_1), (k_2, t_2) and (k_3, t_3) are distinct observable vertices on π_c and (k_2, t_2) is on $\pi_c((k_1, t_1), (k_3, t_3))$, then (k_2, t_2) is on $\pi((k_1, t_1), (k_3, t_3))$.
3. π_c is collider-free.
4. If (k_1, t_1) and (k_2, t_2) are distinct observable vertices on π_c, then $\pi((k_1, t_1), (k_2, t_2))$ is an inducing path relative to $\mathbf{O}(\max(t_1, t_2) - p, t)[\pi_c((k_1, t_1), (k_2, t_2))]$.

REMARK (on Definition D.28). First, Lemma D.22 implies that the canonically induced path π_{ci} of a collider-free path π between distinct observable vertices (i, t_i) and (j, t_j) with $t - p \le t_i, t_j \le p$ is canonical with respect to π. Second, while the first property in Definition D.28 is implied by the second property and would thus not be needed, we have included it in the definition for clarity.

LEMMA D.29. *Given the assumptions and notation of Definition D.26, let* $0 \le a_1 < a_2 < a_3 \le n_c + 1$. *Assume that* $\tilde{\pi}^{a_1,a_2}_c$ *is canonical with respect to* $\tilde{\pi}^{a_1,a_2}$ *and active given* $\mathbf{S} \setminus \{d_{a_1}, d_{a_2}\}$, *and that* $\tilde{\pi}^{a_2,a_3}_c$ *is canonical with respect to* $\tilde{\pi}^{a_2,a_3}$ *and active given* $\mathbf{S} \setminus \{d_{a_2}, d_{a_3}\}$. *Then: If at least one of* $\tilde{\pi}^{a_1,a_2}_c$ *and* $\tilde{\pi}^{a_2,a_3}_c$ *is out of* d_{a_2}, *then there is a path* $\tilde{\pi}^{a_1,a_3}_c$ *that is canonical with respect to* $\tilde{\pi}^{a_1,a_3}$ *and active given* $\mathbf{S} \setminus \{d_{a_1}, d_{a_3}\}$.

PROOF OF LEMMA D.29. We here assume that $\tilde{\pi}^{a_2,a_3}_c$ is out of d_{a_2}, the case in which $\tilde{\pi}^{a_2,a_3}_c$ is into d_{a_2} and $\tilde{\pi}^{a_1,a_2}_c$ out of d_{a_2} follows equivalently. To simplify notation we write $t(v)$ for the time step of a vertex v, i.e., $v = (\cdot, t(v))$. We divide the proof into 14 steps.

Step 1: No vertex on $\tilde{\pi}^{a_1,a_2}$ or $\tilde{\pi}^{a_2,a_3}$ is after t

If there would be a vertex on $\tilde{\pi}^{a_1,a_2}$ (on $\tilde{\pi}^{a_2,a_3}$) that is after t, then this path would have a collider after t because of time order of $\mathcal{D}$ and because both d_{a_1} and d_{a_2} (both d_{a_2} and

d_{a_3}) are not after t. Again using time order, this collider could not be unblocked by $\mathbf{S} \setminus \{d_{a_1}, d_{a_2}\}$ (by $\mathbf{S} \setminus \{d_{a_2}, d_{a_3}\}$) because by definition of $\mathbf{S}$ all vertices in $\mathbf{S}$ are not after t (see Definition D.24). This conclusion contradicts the assumption that $\tilde{\pi}^{a_1,a_2}$ (on $\tilde{\pi}^{a_2,a_3}$) is active given $\mathbf{S} \setminus \{d_{a_1}, d_{a_2}\}$ (given $\mathbf{S} \setminus \{d_{a_2}, d_{a_3}\}$).

Step 2: c_{a_2} and d_{a_2} are in $\mathcal{D}$ ancestors of d_{a_3}

Because $\tilde{\pi}_c^{a_2,a_3}$ is out of d_{a_2} and collider-free (the latter by means of being canonical with respect to $\tilde{\pi}^{a_2,a_3}$), the path $\tilde{\pi}_c^{a_2,a_3}$ is directed from d_{a_2} to d_{a_3}. Hence, all vertices on $\tilde{\pi}_c^{a_2,a_3}$ are ancestors of d_{a_3} and descendants of d_{a_2} in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ and thus, by part 2 of Lemma D.15, also in $\mathcal{D}$. Because c_{a_2} is an ancestor of d_{a_2} in $\mathcal{D}$ by means of ρ^{a_2}, we thus see that c_{a_2} is in $\mathcal{D}$ an ancestor of every vertex on $\tilde{\pi}_c^{a_2,a_3}$.

Step 3: Definition and properties of g_{a_2,a_3} and properties of $\tilde{\pi}^{a_2,a_3}(c_{a_2}, g_{a_2,a_3})$

Let g_{a_2,a_3} be the vertex next to d_{a_2} on $\tilde{\pi}_c^{a_2,a_3}$ (this may be d_{a_3}). Since $\tilde{\pi}_c^{a_2,a_3}$ is directed from d_{a_2} to d_{a_3}, the path $\tilde{\pi}_c^{a_2,a_3}$ is into g_{a_2,a_3}. Because in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ there are no edges into unobservable vertices, we see that g_{a_2,a_3} is observable. Moreover, using time order of $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ and that $\tilde{\pi}_c^{a_2,a_3}$ is directed from d_{a_2} to d_{a_3}, we see that g_{a_2,a_3} is not before d_{a_2} and not after d_{a_3}. Hence, g_{a_2,a_3} is within the observed time window $[t-p, t]$.

Since $\tilde{\pi}_c^{a_2,a_3}$ is canonical with respect to $\tilde{\pi}^{a_2,a_3}$ and g_{a_2,a_3} is on $\tilde{\pi}_c^{a_2,a_3}$, the vertex g_{a_2,a_3} is on $\tilde{\pi}^{a_2,a_3} = \rho^{a_2}(d_{a_2}, c_{a_2}) \oplus \pi^{a_2,a_3} \oplus \rho^{a_3}$. If g_{a_2,a_3} were on ρ^{a_2}, then g_{a_2,a_3} would in $\mathcal{D}$ be an ancestor of d_{a_2} by means of $\rho^{a_2}(g_{a_2,a_3}, d_{a_2})$. This ancestorship would violate acyclicity of $\mathcal{D}$ because g_{a_2,a_3} is a descendant of d_{a_2} according to step 2. Hence, g_{a_2,a_3} is on $\tilde{\pi}^{a_2,a_3}(c_{a_2}, d_{a_3}) = \pi^{a_2,a_3} \oplus \rho^{a_3}$ excluding c_{a_2}. Moreover, $\tilde{\pi}^{a_2,a_3}(c_{a_2}, g_{a_2,a_3})$ is a nontrivial subpath of $\tilde{\pi}^{a_2,a_3}(c_{a_2}, d_{a_3})$ and of $\tilde{\pi}^{a_2,a_3}(d_{a_2}, g_{a_2,a_3})$. Lastly, $\tilde{\pi}^{a_2,a_3}(c_{a_2}, g_{a_2,a_3})$ is into c_{a_2} because π is into c_{a_2}.

Since $\tilde{\pi}_c^{a_2,a_3}$ is canonical with respect to $\tilde{\pi}^{a_2,a_3}$ and $t(d_{a_2}) \leq t(g_{a_1,a_2})$ by time order, $\tilde{\pi}^{a_2,a_3}(d_{a_2}, g_{a_2,a_3})$ is an inducing path relative to $\mathbf{O}(t(g_{a_2,a_3})-p, t)[\tilde{\pi}_c^{a_2,a_3}(d_{a_2}, g_{a_2,a_3})]$. Here, the simplification $\mathbf{O}(t(g_{a_2,a_3}) - p, t)[\tilde{\pi}_c^{a_2,a_3}(d_{a_2}, g_{a_2,a_3})] = \mathbf{O}(t(g_{a_2,a_3}) - p, t)$ applies because $\tilde{\pi}_c^{a_2,a_3}(d_{a_2}, g_{a_2,a_3})$ by definition of g_{a_2,a_3} consists of its end point vertices d_{a_2} and g_{a_2,a_3} only. Using the defining properties of inducing paths, we thus see that the path $\tilde{\pi}^{a_2,a_3}(d_{a_2}, g_{a_2,a_3})$ has the following two properties:

1. First, if v is an observable non-end-point noncollider on $\tilde{\pi}^{a_2,a_3}(d_{a_2}, g_{a_2,a_3})$, then $t(v) < t(g_{a_2,a_3}) - p$ or $t(v) > t$. Because step 1 excludes $t(v) > t$, in fact $t(v) < t(g_{a_2,a_3}) - p$.
2. Second, if v is a collider on $\tilde{\pi}^{a_2,a_3}(d_{a_2}, g_{a_2,a_3})$, then v is in $\mathcal{D}$ an ancestor of d_{a_2} or g_{a_2,a_3}. Since d_{a_2} is in $\mathcal{D}$ an ancestor of g_{a_2,a_3}, the vertex v is, in fact, an ancestor of g_{a_2,a_3}.

Both of these statements are also true for $\tilde{\pi}^{a_2,a_3}(c_{a_2}, g_{a_2,a_3})$ because it is a subpath of $\tilde{\pi}^{a_2,a_3}(d_{a_2}, g_{a_2,a_3})$.

Step 4: All observable vertices on ρ^{a_2} other than d_{a_2} are before $t(g_{a_2,a_3}) - p$

Let $v \neq d_{a_2}$ be an observable vertex on ρ^{a_2}. Since g_{a_2,a_3} is not on ρ^{a_2}, see step 2, v is then a non-end-point vertex on $\tilde{\pi}^{a_2,a_3}(d_{a_2}, g_{a_2,a_3})$. Moreover, since ρ^{a_2} is directed from c_{a_2} to d_{a_2}, the vertex v is a noncollider on $\tilde{\pi}^{a_2,a_3}(d_{a_2}, g_{a_2,a_3})$. From step 3 we then get that $t(v) < t(g_{a_2,a_3}) - p$ or $t(v) > t$, and step 1 further excludes the case $t(v) > t$.

Step 5: Definition and properties of h_{a_1,a_2} and properties of $\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, c_{a_2})$

Let h_{a_1,a_2} be the observable vertex on $\tilde{\pi}_c^{a_1,a_2}$ closest to d_{a_2} other than d_{a_2} itself that is not more than p time steps before g_{a_2,a_3}, i.e., for which $t(g_{a_1,a_2}) - p \leq t(h_{a_1,a_2})$ (note that h_{a_1,a_2} may be d_{a_1}).

Since $\tilde{\pi}_c^{a_1,a_2}$ is canonical with respect to $\tilde{\pi}^{a_1,a_2}$, the vertex h_{a_1,a_2} is on $\tilde{\pi}^{a_1,a_2} = \rho^{a_1}(d_{a_1}, c_{a_1}) \oplus \pi^{a_1,a_2} \oplus \rho^{a_2}$. Due to step 4 and $t(g_{a_1,a_2}) - p \leq t(h_{a_1,a_2})$, the vertex h_{a_1,a_2} cannot be on ρ^{a_2} unless $h_{a_1,a_2} = d_{a_2}$, which is, however, excluded by definition. Hence, h_{a_1,a_2} is on $\tilde{\pi}^{a_1,a_2}(d_{a_1}, c_{a_2}) = \rho^{a_1}(d_{a_1}, c_{a_1}) \oplus \pi^{a_1,a_2}$ excluding c_{a_2} (because c_{a_2} is on ρ_{a_2}). Moreover,

$\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, c_{a_2})$ is a nontrivial subpath of $\tilde{\pi}^{a_1,a_2}(d_{a_1}, c_{a_2})$ and of $\tilde{\pi}^{a_2,a_3}(h_{a_1,a_2}, d_{a_2})$. Lastly, $\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, c_{a_2})$ is into c_{a_2} because π^{a_1,a_2} is into c_{a_2}.

Write $t_{hd} = \max(t(h_{a_1,a_2}), t(d_{a_2}))$ and $t_{hg} = \max(t(h_{a_1,a_2}), t(g_{a_2,a_3}))$. Since $\tilde{\pi}_c^{a_1,a_2}$ is canonical with respect to $\tilde{\pi}^{a_1,a_2}$, the path $\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, d_{a_2})$ is an inducing path relative to $\mathbf{O}(t_{hd} - p, t)[\tilde{\pi}_c^{a_1,a_2}(h_{a_1,a_2}, d_{a_2})]$. In particular, $\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, d_{a_2})$ has the following two properties:

1. First, if v is an observable non-end-point noncollider on $\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, d_{a_2})$, then $t(v) < t_{hd} - p$ or $t(v) > t$ or v is on $\tilde{\pi}_c^{a_1,a_2}(h_{a_1,a_2}, d_{a_2})$. The case $t(v) > t$ is excluded by step 1. If v is on $\tilde{\pi}_c^{a_1,a_2}(h_{a_1,a_2}, d_{a_2})$, then $t(v) < t(g_{a_1,a_2}) - p$ by definition of h_{a_1,a_2}. Note that $t_{hd} - p \leq t_{hg} - p$ and $t(g_{a_1,a_2}) - p \leq t_{hg} - p$. Hence, in any case, $t(v) < t_{hg} - p$.
2. Second, if v is a collider on $\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, d_{a_2})$, then v is in $\mathcal{D}$ an ancestor of h_{a_1,a_2} or d_{a_2}. Because d_{a_2} is in $\mathcal{D}$ an ancestor of g_{a_2,a_3}, the vertex v is, in fact, in $\mathcal{D}$ an ancestor of h_{a_1,a_2} or g_{a_2,a_3}.

Both of these statements are also true for $\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, c_{a_2})$ because $\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, c_{a_2})$ is a subpath of $\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, d_{a_2})$.

Step 6: g_{a_2,a_3} and h_{a_1,a_2} are adjacent or almost adjacent in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$

Consider the concatentation $\psi = \tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, c_{a_2}) \oplus \tilde{\pi}^{a_2,a_3}(c_{a_2}, g_{a_2,a_3})$. This concatenation is a path (rather than a walk) in $\mathcal{D}$ because $\tilde{\pi}^{a_2,a_3}(c_{a_2}, g_{a_2,a_3})$ is a subpath of $\pi^{a_2,a_3} \oplus \rho^{a_3}$ and because different ρ^a do not intersect (by definition of collider extension structures). The junction point c_{a_2} is a collider on ψ because both $\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, c_{a_2})$ and $\tilde{\pi}^{a_2,a_3}(c_{a_2}, g_{a_2,a_3})$ are into c_{a_2}, see steps 5 and 3. We now show that ψ is an inducing path relative to $\mathbf{O}(t_{hg} - p, t_{hg})$. To this end, we separately look at the colliders and non-end-point noncolliders on ψ.

Colliders: According to steps 3 and 5, every collider on $\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, c_{a_2})$ and on $\tilde{\pi}^{a_2,a_3}(c_{a_2}, g_{a_2,a_3})$ is in $\mathcal{D}$ an ancestor of h_{a_1,a_2} or g_{a_2,a_3}. The junction point c_{a_2} is in $\mathcal{D}$ an ancestor of g_{a_2,a_3} according to step 1.

non-end-point noncolliders: Let v be a non-end-point noncollider on ψ. Since c_{a_2} is a collider on ψ, the vertex v is then a non-end-point noncollider on $\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, c_{a_2})$ or a non-end-point noncollider on $\tilde{\pi}^{a_2,a_3}(c_{a_2}, g_{a_2,a_3})$. With steps 3 and 5 we then get $t(v) < t_{hg} - p$.

Consequently, ψ is an inducing path relative to the set of observable vertices within $\mathbf{O}(t_{hg} - p, t_{hg})$. By shifting this structure forward in time by $t - t_{hg}$ time steps, we see that the forward shifted copies of g_{a_2,a_3} and h_{a_1,a_2} are adjacent in $\mathcal{M}^p(\mathcal{D})$. Hence, g_{a_2,a_3} and h_{a_1,a_2} are adjacent or almost adjacent in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ according to Lemma D.11.

For reference further below we note that ψ is also on inducing path relative to $\mathbf{O}(t_{hg} - p, t)$. This statement follows because, as shown, if v is a non-end-point noncollider on ψ, then $t(v) < t_{hg} - p$ and thus v is not in $\mathbf{O}(t_{hg} - p, t) \setminus \mathbf{O}(t_{hg} - p, t_{hg})$.

Step 7: Properties of $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$

Since $\tilde{\pi}_c^{a_2,a_3}$ is directed from d_{a_2} to d_{a_3}, see step 2, $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ is the trivial path consisting of the single vertex $g_{a_2,a_3} = d_{a_3}$ or a nontrivial directed path from g_{a_2,a_3} to d_{a_3}, and hence out of g_{a_2,a_3} and into d_{a_3}. In particular, $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ is collider-free.

Consider any vertex v on $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$. Because in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ there are no edges into unobservable vertices and because g_{a_2,a_3} is observable, v is observable. Moreover, because $\tilde{\pi}_c^{a_2,a_3}$ is canonical with respect to $\tilde{\pi}^{a_2,a_3}$, the vertex v is on $\tilde{\pi}^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$. Since g_{a_2,a_3} is on $\pi^{a_2,a_3} \oplus \rho^{a_3}$ excluding c_{a_2}, see step 3, also v is on $\pi^{a_2,a_3} \oplus \rho^{a_3}$ excluding c_{a_2}. In particular, v is not on ρ^{a_2}.

Step 8: Properties of $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$

Because $\tilde{\pi}_c^{a_1,a_2}$ is canonical with respect to $\tilde{\pi}^{a_1,a_2}$, the path $\tilde{\pi}_c^{a_1,a_2}$ is collider-free. Consequently, also $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ is collider-free.

Consider any observable vertex v on $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$. Because $\tilde{\pi}_c^{a_1,a_2}$ is canonical with respect to $\tilde{\pi}^{a_1,a_2}$, the vertex v is on $\tilde{\pi}^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$. Since h_{a_1,a_2} is on $\rho^{a_1}(d_{a_1}, c_{a_1}) \oplus \pi^{a_1,a_2}$

excluding c_{a_2}, see step 5, also v is on $\rho^{a_1}(d_{a_1}, c_{a_1}) \oplus \pi^{a_1,a_2}$ excluding c_{a_2}. In particular, we conclude that v is not on ρ^{a_2}.

We now show that $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ and $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ do not intersect. Assume the opposite, i.e., let w be on both $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ and $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$. There are two cases:

1. First, assume w is observable. Then, according to step 7 and the previous discussion in the current step, w is on $\rho^{a_1}(d_{a_1}, c_{a_1}) \oplus \pi^{a_1,a_2}$ excluding c_{a_2} and on $\pi^{a_2,a_3} \oplus \rho^{a_3}$ excluding c_{a_2}. These observations contradict each other because $\rho^{a_1}(d_{a_1}, c_{a_1}) \oplus \pi^{a_1,a_2}$ and $\pi^{a_2,a_3} \oplus \rho^{a_3}$ intersect at c_{a_2} only.
2. Second, assume w is unobservable. Then, w is a non-end-point vertex of $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ and of non-end-point vertex of $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$. Moreover, as follows immediately from the definition of canonical ts-DAGs, every unobservable vertex in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ is adjacent to exactly two vertices, both of which are observable. We thus find that $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ and $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ also intersect at an observable vertex, which has already been ruled out in the previous case and thus is a contradiction.

Step 9: Construction of $\tilde{\pi}_c^{a_1,a_3}$

The fact that g_{a_2,a_3} and h_{a_1,a_2} are adjacent or almost adjacent in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$, see step 6, means that in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ there is a path $\kappa^{a_2} = h_{a_1,a_2} {\rightarrow} g_{a_2,a_3}$ or $\kappa^{a_2} = h_{a_1,a_2} {\leftarrow} g_{a_2,a_3}$ or $\kappa^{a_2} = h_{a_1,a_2} {\leftarrow} u_{a_2} {\rightarrow} g_{a_2,a_3}$ with u_{a_2} unobservable. Let $\tilde{\pi}_c^{a_1,a_3}$ be the concatenation $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2}) \oplus \kappa^{a_2} \oplus \tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$. We now show that $\tilde{\pi}_c^{a_1,a_3}$ is a path.

In step 8 we have already shown that $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ and $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ do not intersect. Thus, if $\kappa^{a_2} = h_{a_1,a_2} {\rightarrow} g_{a_2,a_3}$ or $\kappa^{a_2} = h_{a_1,a_2} {\leftarrow} g_{a_2,a_3}$, then $\tilde{\pi}_c^{a_1,a_3}$ contains every vertex at most once and hence is a path. Now assume that $\kappa^{a_2} = h_{a_1,a_2} {\leftarrow} u_{a_2} {\rightarrow} g_{a_2,a_3}$. In this case, we show that u_{a_2} is neither on $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ nor $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$:

1. First, because as shown in step 7 all vertices on $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ are observable, u_{a_2} cannot be on $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$.
2. Second, if u_{a_2} is on $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$, then it is a non-end-point vertex of this path. Since every unobservable vertex in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ is adjacent two exactly vertices, which for u_{a_2} are h_{a_1,a_2} and g_{a_2,a_3}, we thus find that g_{a_2,a_3} is on $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$. This observation contradicts the fact that $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ and $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ have no common vertex.

For reference below we note that by construction $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2}) = \tilde{\pi}_c^{a_1,a_3}(d_{a_1}, h_{a_1,a_2})$ and $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3}) = \tilde{\pi}_c^{a_1,a_3}(g_{a_2,a_3}, d_{a_3})$.

Step 10: $\tilde{\pi}_c^{a_1,a_3}$ is collider-free

Since the three constituting subpaths $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$, κ^{a_2} and $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ are collider-free, only h_{a_1,a_2} or g_{a_2,a_3} can potentially be colliders on $\pi_c^{a_1,a_3}$.

First, because $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ is a trivial path or a nontrivial path and out of g_{a_2,a_3}, see step 7, g_{a_2,a_3} is a noncollider on $\tilde{\pi}_c^{a_1,a_3}$.

Second, assume that h_{a_1,a_2} is a collider on $\tilde{\pi}_c^{a_1,a_3}$. This premise requires that κ^{a_2} is $h_{a_1,a_2} {\leftarrow} g_{a_2,a_3}$ or $h_{a_1,a_2} {\leftarrow} u_{a_2} {\rightarrow} g_{a_2,a_3}$ and that $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ is nontrivial and into h_{a_1,a_2}. In combination with the facts that $\tilde{\pi}_c^{a_1,a_2}$ is collider-free and that $h_{a_1,a_2} \neq c_{a_2}$ this form of $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ requires that $\tilde{\pi}_c^{a_1,a_2}(h_{a_1,a_2}, c_{a_2})$ is a nontrivial directed path from h_{a_1,a_2} to c_{a_2}. Hence, h_{a_1,a_2} is an ancestor of c_{a_2} in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ and, thus, also an ancestor in $\mathcal{D}$. Together with step 2 this ancestral relationship shows that h_{a_1,a_2} is in $\mathcal{D}$ an ancestor of g_{a_2,a_3}. Lemma D.12 in combination with the definition of κ^{a_2} then requires $\kappa^{a_2} = h_{a_1,a_2} {\rightarrow} g_{a_2,a_3}$, a contradiction.

Step 11: All observable vertices on $\tilde{\pi}_c^{a_1,a_3}$ are also on $\tilde{\pi}^{a_1,a_3}$

Recall from above that $\tilde{\pi}_c^{a_1,a_3} = \tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2}) \oplus \kappa^{a_2} \oplus \tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ and $\tilde{\pi}^{a_1,a_3} = \rho^{a_1}(d_{a_1}, c_{a_1}) \oplus \pi^{a_1,a_2} \oplus \pi^{a_2,a_3} \oplus \rho^{a_3}$ (the latter because $\pi^{a_1,a_2} \oplus \pi^{a_2,a_3} = \pi^{a_1,a_3}$). According to step 7, every vertex on $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ is on $\pi^{a_2,a_3} \oplus \rho^{a_3}$ and hence on $\tilde{\pi}^{a_1,a_3}$. According to step 8, every observable vertex on $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ is on $\rho^{a_1}(d_{a_1}, c_{a_1}) \oplus \pi^{a_1,a_2}$ and hence on $\tilde{\pi}^{a_1,a_3}$. Lastly, due to the three particular forms that the path κ^{a_2} may have, see step 9, every observable vertex on κ^{a_2} is on $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ or $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ and hence on $\tilde{\pi}^{a_1,a_3}$.

Step 12: $\tilde{\pi}_c^{a_1,a_3}$ fulfills point 2. in Definition D.28

For reference below we note the following results:

1. If w_1 and w_2 are on $\tilde{\pi}^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ or on $\tilde{\pi}^{a_1,a_3}(d_{a_1}, h_{a_1,a_2})$, then $\tilde{\pi}^{a_1,a_2}(w_1, w_2) = \tilde{\pi}^{a_1,a_3}(w_1, w_2)$. This equality follows because h_{a_1,a_2} and thus also w_1 and w_2 are on $\rho^{a_1}(d_{a_1}, c_{a_1}) \oplus \pi^{a_1,a_2}$.
2. If w_1 and w_2 are observable vertices on $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$, then $\tilde{\pi}^{a_1,a_2}(w_1, w_2) = \tilde{\pi}^{a_1,a_3}(w_1, w_2)$. This equality follows from the previous result because w_1 and w_2 are on $\tilde{\pi}^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ by means of $\tilde{\pi}_c^{a_1,a_2}$ being canonical with respect to $\tilde{\pi}^{a_1,a_2}$.
3. If w_1 and w_2 are on $\tilde{\pi}^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ or on $\tilde{\pi}^{a_1,a_3}(g_{a_2,a_3}, d_{a_3})$, then $\tilde{\pi}^{a_2,a_3}(w_1, w_2) = \tilde{\pi}^{a_1,a_3}(w_1, w_2)$. This equality follows because g_{a_2,a_3} and thus also w_1 and w_2 are on $\pi^{a_2,a_3} \oplus \rho^{a_3}$.
4. If w_1 and w_2 are observable vertices on $\tilde{\pi}_c^{a_2,a_3}(g_{a_1,a_2}, d_{a_3})$, then $\tilde{\pi}^{a_2,a_3}(w_1, w_2) = \tilde{\pi}^{a_1,a_3}(w_1, w_2)$. This equality follows from the previous result because w_1 and w_2 are on $\tilde{\pi}^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ by means of $\tilde{\pi}_c^{a_2,a_3}$ being canonical with respect to $\tilde{\pi}^{a_2,a_3}$.

Let v_1, v_2 and v_3 be distinct observable vertices on $\tilde{\pi}_c^{a_1,a_3}$ such that v_2 is on $\tilde{\pi}_c^{a_1,a_3}(v_1, v_3)$ and, without loss of generality, v_1 is closer to d_{a_1} on $\tilde{\pi}_c^{a_1,a_3}$ than v_3 is to d_{a_1} on $\tilde{\pi}_c^{a_1,a_3}$. We distinguish several collectively exhaustive cases:

1. First, assume v_3 is on $\tilde{\pi}_c^{a_1,a_3}(d_{a_1}, h_{a_1,a_2})$. This premise implies that also both v_1 and v_2 are on $\tilde{\pi}_c^{a_1,a_3}(d_{a_1}, h_{a_1,a_2}) = \tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$. Hence, v_2 is on $\tilde{\pi}_c^{a_1,a_2}(v_1, v_3) = \tilde{\pi}_c^{a_1,a_3}(v_1, v_3)$ and thus, using that $\tilde{\pi}_c^{a_1,a_2}$ is canonical with respect to $\tilde{\pi}^{a_1,a_2}$, also on $\tilde{\pi}^{a_1,a_2}(v_1, v_3)$. Moreover, using the second result at the beginning of this step, $\tilde{\pi}^{a_1,a_2}(v_1, v_3) = \tilde{\pi}^{a_1,a_3}(v_1, v_3)$. Hence, v_2 is on $\tilde{\pi}^{a_1,a_3}(v_1, v_3)$.
2. Second, assume that v_2 is on $\tilde{\pi}_c^{a_1,a_3}(d_{a_1}, h_{a_1,a_2})$ excluding h_{a_1,a_2} and that v_3 is not on $\tilde{\pi}_c^{a_1,a_3}(d_{a_1}, h_{a_1,a_2})$. This premise implies that also v_1 is on $\tilde{\pi}_c^{a_1,a_3}(d_{a_1}, h_{a_1,a_2}) = \tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$. Following the same steps as in the previous case with v_3 replaced by h_{a_1,a_2}, we get that v_2 is on $\tilde{\pi}^{a_1,a_3}(v_1, h_{a_1,a_2}) = \tilde{\pi}^{a_1,a_2}(v_1, h_{a_1,a_2})$. Moreover, v_3 is on $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3}) = \tilde{\pi}_c^{a_1,a_3}(g_{a_2,a_3}, d_{a_3})$ and hence, using that $\tilde{\pi}_c^{a_2,a_3}$ is canonical with respect to $\tilde{\pi}^{a_2,a_3}$, on $\tilde{\pi}^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$. This observation shows that $\tilde{\pi}^{a_1,a_3}(v_1, h_{a_1,a_2})$ is a subpath of $\tilde{\pi}^{a_1,a_3}(v_1, v_3)$ and hence that v_2 is on $\tilde{\pi}^{a_1,a_3}(v_1, v_3)$.
3. Third, assume v_1 is on $\tilde{\pi}_c^{a_1,a_3}(g_{a_2,a_3}, d_{a_3})$. This premise implies that also both v_2 and v_3 are on $\tilde{\pi}_c^{a_1,a_3}(g_{a_2,a_3}, d_{a_3}) = \tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$. Hence, v_2 is on $\tilde{\pi}_c^{a_2,a_3}(v_1, v_3) = \tilde{\pi}_c^{a_1,a_3}(v_1, v_3)$ and thus, using that $\tilde{\pi}_c^{a_2,a_3}$ is canonical with respect to $\tilde{\pi}^{a_2,a_3}$, also on $\tilde{\pi}^{a_2,a_3}(v_1, v_3)$. Moreover, using the fourth result at the beginning of this step, $\tilde{\pi}^{a_2,a_3}(v_1, v_3) = \tilde{\pi}^{a_1,a_3}(v_1, v_3)$. Hence, v_2 is on $\tilde{\pi}^{a_1,a_3}(v_1, v_3)$.
4. Fourth, assume that v_2 is on $\tilde{\pi}_c^{a_1,a_3}(g_{a_2,a_3}, d_{a_3})$ excluding g_{a_2,a_3} and that v_1 is not on $\tilde{\pi}_c^{a_1,a_3}(g_{a_2,a_3}, d_{a_3})$. This premise implies that also v_3 is on $\tilde{\pi}_c^{a_1,a_3}(g_{a_2,a_3}, d_{a_3}) = \tilde{\pi}_c^{a_1,a_2}(g_{a_2,a_3}, d_{a_3})$. Following the same steps as in the previous case with v_1 replaced by g_{a_2,a_3}, we get that v_2 is on $\tilde{\pi}^{a_1,a_3}(g_{a_2,a_3}, v_3) = \tilde{\pi}^{a_2,a_3}(g_{a_2,a_3}, v_3)$. Moreover, v_1 is on $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2}) = \tilde{\pi}_c^{a_1,a_3}(d_{a_1}, h_{a_1,a_2})$ and hence, using that $\tilde{\pi}_c^{a_1,a_2}$ is canonical with respect to $\tilde{\pi}^{a_1,a_2}$, on $\tilde{\pi}^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$. This observation shows that $\tilde{\pi}^{a_1,a_3}(g_{a_2,a_3}, v_3)$ is a subpath of $\tilde{\pi}^{a_1,a_3}(v_1, v_3)$ and hence that v_2 is on $\tilde{\pi}^{a_1,a_3}(v_1, v_3)$.
5. Fifth, assume that v_2 is h_{a_1,a_2} or g_{a_2,a_3}. This premise implies that v_1 is on the path $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2}) = \tilde{\pi}_c^{a_1,a_3}(d_{a_1}, h_{a_1,a_2})$ and hence, because $\tilde{\pi}_c^{a_1,a_2}$ is canonical with respect to $\tilde{\pi}^{a_1,a_2}$, on $\tilde{\pi}^{a_1,a_3}(d_{a_1}, h_{a_1,a_2}) = \tilde{\pi}^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$, where the latter equality follows by the first result at the beginning of this step. Moreover, the premise implies

that v_3 is on $\tilde{\pi}_c^{a_2,a_3}(g_{a_1,a_2}, d_{a_2}) = \tilde{\pi}_c^{a_1,a_3}(g_{a_1,a_2}, d_{a_2})$ and hence, because $\tilde{\pi}_c^{a_2,a_3}$ is canonical with respect to $\tilde{\pi}^{a_2,a_3}$, on $\tilde{\pi}^{a_1,a_3}(g_{a_1,a_2}, d_{a_2}) = \tilde{\pi}^{a_2,a_3}(g_{a_1,a_2}, d_{a_2})$, where the latter equality follows by the third result at the beginning of this step. These considerations show that $\tilde{\pi}^{a_1,a_3}(v_1, v_3)$ decomposes as $\tilde{\pi}^{a_1,a_3}(v_1, h_{a_1,a_2}) \oplus \tilde{\pi}^{a_1,a_3}(h_{a_1,a_2}, g_{a_2,a_3}) \oplus \tilde{\pi}^{a_1,a_3}(g_{a_2,a_3}, v_3)$. Hence, v_2 is on $\tilde{\pi}^{a_1,a_3}(v_1, v_3)$ irrespective of whether $v_2 = h_{a_1,a_2}$ or $v_2 = g_{a_2,a_3}$.

Step 13: $\tilde{\pi}_c^{a_1,a_3}$ fulfills point 4. in Definition D.28

Let v_1 and v_2 be two distinct observable vertices on $\tilde{\pi}_c^{a_1,a_3}$ such that, without loss of generality, v_1 is closer to d_{a_1} on $\tilde{\pi}_c^{a_1,a_3}$ than v_2 is to d_{a_1} on $\tilde{\pi}_c^{a_1,a_3}$. We distinguish three collectively exhaustive cases:

1. First, assume v_2 is on $\tilde{\pi}_c^{a_1,a_3}(d_{a_1}, h_{a_1,a_2})$. Then, both v_1 and v_2 are on $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2}) = \tilde{\pi}_c^{a_1,a_3}(d_{a_1}, h_{a_1,a_2})$ and hence $\tilde{\pi}_c^{a_1,a_2}(v_1, v_2) = \tilde{\pi}_c^{a_1,a_3}(v_1, v_2)$. Since $\tilde{\pi}_c^{a_1,a_2}$ is canonical with respect to $\tilde{\pi}^{a_1,a_2}$, we thus find that $\tilde{\pi}^{a_1,a_2}(v_1, v_2)$ is an inducing path relative to $\mathbf{O}(\max(t(v_1), t(v_2)) - p, t)[\tilde{\pi}_c^{a_1,a_3}(v_1, v_2)]$. The desired results follows since $\tilde{\pi}^{a_1,a_2}(v_1, v_2) = \tilde{\pi}^{a_1,a_3}(v_1, v_2)$ according to the second result at the beginning of step 12.
2. Second, assume v_1 is on $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$. Then, both v_1 and v_2 are on $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ and hence $\tilde{\pi}_c^{a_1,a_3}(v_1, v_2) = \tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$. Since $\tilde{\pi}_c^{a_2,a_3}$ is canonical with respect to $\tilde{\pi}^{a_2,a_3}$, we find that $\tilde{\pi}^{a_2,a_3}(v_1, v_2)$ is an inducing path relative to $\mathbf{O}(\max(t(v_1), t(v_2)) - p, t)[\tilde{\pi}_c^{a_1,a_3}(v_1, v_2)]$. The desired results follows since $\tilde{\pi}^{a_2,a_3}(v_1, v_2) = \tilde{\pi}^{a_1,a_3}(v_1, v_2)$ according to the fourth result at the beginning of step 12.
3. Third, assume that neither of the previous two cases applies. Then, using that every observable vertex on κ^{a_2} is on $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ or $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$, the vertex v_1 is on $\tilde{\pi}_c^{a_1,a_3}(d_{a_1}, h_{a_1,a_2})$ and v_2 is on $\tilde{\pi}_c^{a_1,a_3}(g_{a_2,a_3}, d_{a_3})$. Thus, both h_{a_1,a_2} and g_{a_2,a_3} are on $\tilde{\pi}_c^{a_1,a_3}(v_1, v_2)$. Since $\tilde{\pi}_c^{a_1,a_3}$ is collider-free, we thus find that both h_{a_1,a_2} and g_{a_2,a_3} are ancestors of v_1 or v_2 in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ and, thus, ancestors in $\mathcal{D}$. Moreover, since $\tilde{\pi}_c^{a_1,a_3}$ fulfills point 2 in Definition D.28, v_1 is on $\tilde{\pi}^{a_1,a_3}(d_{a_1}, h_{a_1,a_2})$ and v_2 is on $\tilde{\pi}^{a_1,a_3}(g_{a_2,a_3}, d_{a_3})$. Consequently, the path of interest $\tilde{\pi}^{a_1,a_3}(v_1, v_2)$ decomposes as $\tilde{\pi}^{a_1,a_3}(v_1, h_{a_1,a_2}) \oplus \tilde{\pi}^{a_1,a_3}(h_{a_1,a_2}, g_{a_1,a_2}) \oplus \tilde{\pi}^{a_1,a_3}(g_{a_1,a_2}, v_2)$. We now individually look at the three constituting subpaths:
   a) By following the same steps as in the first case of this enumeration with v_2 replaced by h_{a_1,a_2}, we get that $\tilde{\pi}^{a_1,a_3}(v_1, h_{a_1,a_2})$ is an inducing path relative to $\mathbf{O}(t_{v_1h} - p, t)[\tilde{\pi}_c^{a_1,a_3}(v_1, h_{a_1,a_2})]$, where $t_{v_1h} = \max(t(v_1), t(h_{a_1,a_2}))$. Moreover, note that $\tilde{\pi}_c^{a_1,a_3}(v_1, h_{a_1,a_2})$ is a subpath of $\tilde{\pi}_c^{a_1,a_3}(v_1, v_2)$. Since an inducing path relative to some set $\mathbf{O}_1$ of observed vertices is also an inducing path relative to another set $\mathbf{O}_2$ of observed with $\mathbf{O}_2 \subseteq \mathbf{O}_1$, we get that $\tilde{\pi}^{a_1,a_3}(v_1, h_{a_1,a_2})$ is an inducing path relative to $\mathbf{O}(t_{v_1h} - p, t)[\tilde{\pi}_c^{a_1,a_3}(v_1, v_2)]$.
   b) Because h_{a_1,a_2} is on $\rho^{a_1}(d_{a_1}, c_{a_1}) \oplus \pi^{a_1,a_2}$ and g_{a_2,a_3} is on $\pi^{a_2,a_3} \oplus \rho^{a_3}$, see steps 5 and 3, the path $\tilde{\pi}^{a_1,a_3}(h_{a_1,a_2}, g_{a_1,a_2})$ decomposes as $\tilde{\pi}^{a_1,a_2}(h_{a_1,a_2}, c_{a_2}) \oplus \tilde{\pi}^{a_2,a_3}(c_{a_2}, g_{a_1,a_2})$. This decomposition shows that $\tilde{\pi}^{a_1,a_3}(h_{a_1,a_2}, g_{a_1,a_2}) = \psi$, with ψ as considered in step 6. Hence, $\tilde{\pi}^{a_1,a_3}(h_{a_1,a_2}, g_{a_1,a_2})$ is an inducing path relative to $\mathbf{O}(t_{hg} - p, t)$ and, thus, an inducing path relative to $\mathbf{O}(t_{hg} - p, t)[\tilde{\pi}_c^{a_1,a_3}(v_1, v_2)]$.
   c) By following the same steps as in the second case of this enumeration with v_1 replaced by g_{a_1,a_2}, we get that $\tilde{\pi}^{a_2,a_3}(g_{a_2,a_3}, v_2)$ is an inducing path relative to $\mathbf{O}(t_{gv_2} - p, t)[\tilde{\pi}_c^{a_1,a_3}(g_{a_2,a_3}, v_2)]$, where $t_{gv_2} = \max(t(g_{a_1,a_2}), t(v))$. Moreover, since $\tilde{\pi}_c^{a_1,a_3}(g_{a_2,a_3}, v_3)$ is a subpath of $\tilde{\pi}_c^{a_1,a_3}(v_1, v_2)$, we get that $\tilde{\pi}^{a_2,a_3}(g_{a_2,a_3}, v_2)$ is inducing path relative to $\mathbf{O}(t_{gv_2} - p, t)[\tilde{\pi}_c^{a_1,a_3}(v_1, v_2)]$.

   To proof the desired inducing path property of $\tilde{\pi}^{a_1,a_3}(v_1, v_2)$, we now separately consider its colliders and observable non-end-point noncolliders:

a) First, let v be a collider on $\tilde{\pi}^{a_1,a_3}(v_1, v_2)$. Then, v is a collider on one of the three constituting subpaths or is h_{a_1,a_2} or g_{a_2,a_3}. In all cases, using the inducing path properties of the constituting subpaths, v is in $\mathcal{D}$ an ancestor of v_1 or h_{a_1,a_2} or g_{a_2,a_3} or v_2. Since, as shown above in this step, both h_{a_1,a_2} and g_{a_2,a_3} are in $\mathcal{D}$ ancestors of v_1 or v_2, we get that v is in $\mathcal{D}$ an ancestor of v_1 or v_2.
b) Second, let v be an observable non-end-point noncollider on $\tilde{\pi}^{a_1,a_3}(v_1, v_2)$. Since h_{a_1,a_2} and g_{a_2,a_3} are in $\mathcal{D}$ ancestors of v_1 or v_2, time order of $\mathcal{D}$ guarantees that $t(h_{a_1,a_2}) \leq t_{v_1v_2}$ and $t(g_{a_2,a_3}) \leq t_{v_1v_2}$, where $t_{v_1v_2} = \max(t(v_1), t(v_2))$. These inequalities imply $t_{v_1h} \leq t_{v_1v_2}$, $t_{hg} \leq t_{v_1v_2}$, and $t_{gv_2} \leq t_{v_1v_2}$. Thus, using the inducing path properties of the constituting subpaths in combination with $t(v) \leq t$, see step 1, we find that $t(v) < t_{v_1v_2} - p$ if v is not a non-end-point vertex of $\tilde{\pi}_c^{a_1,a_3}(v_1, v_2)$.

Thus, $\tilde{\pi}^{a_1,a_3}(v_1, v_2)$ is an inducing path relative to $\mathbf{O}(t_{v_1v_2} - p, t)[\tilde{\pi}_c^{a_1,a_3}(v_1, v_2)]$.

The combination of the four steps 11, 12, 10 and 13 shows that $\tilde{\pi}_c^{a_1,a_3}$ is canonical with respect to $\tilde{\pi}^{a_1,a_3}$.

Step 14: $\tilde{\pi}_c^{a_1,a_3}$ is active given $\mathbf{S} \setminus \{d_{a_1}, d_{a_3}\}$

Recall that $\tilde{\pi}_c^{a_1,a_3} = \tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2}) \oplus \kappa^{a_2} \oplus \tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ and assume $\tilde{\pi}_c^{a_1,a_3}$ is blocked given $\mathbf{S} \setminus \{d_{a_1}, d_{a_3}\}$. Since $\tilde{\pi}_c^{a_1,a_3}$ is collider-free, see step 10, this premise means there is a non-end-point vertex v on $\tilde{\pi}_c^{a_1,a_3}$ with $v \in \mathbf{S} \setminus \{d_{a_1}, d_{a_3}\}$. There are three collectively exhaustive cases:

1. First, assume v is on $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$. Then, because $\tilde{\pi}_c^{a_1,a_2}$ is canonical with respect to $\tilde{\pi}^{a_1,a_2}$, the vertex v is on $\tilde{\pi}^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$. Since $\tilde{\pi}^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ is a subpath of $\rho^{a_1}(d_{a_1}, c_{a_1}) \oplus \pi^{a_1,a_2}$ excluding c_{a_2}, see step 5, v is not on ρ^{a_2}. In particular, $v \neq d_{a_2}$ and thus $v \in \mathbf{S} \setminus \{d_{a_1}, d_{a_2}\}$. Moreover, since $v \neq d_{a_1}$ and $h_{a_1,a_2} \neq d_{a_2}$ by the respective definitions of v and h_{a_1,a_2}, the vertex v is a non-end-point vertex on $\tilde{\pi}_c^{a_1,a_2}$. Since $\tilde{\pi}_c^{a_1,a_2}$ is collider-free by means of being canonical with respect to $\tilde{\pi}^{a_1,a_2}$ and since $v \in \mathbf{S} \setminus \{d_{a_1}, d_{a_2}\}$, we arrive at a contradiction to the assumption that $\tilde{\pi}_c^{a_1,a_2}$ is active given $\mathbf{S} \setminus \{d_{a_1}, d_{a_2}\}$.
2. Second, assume v is $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$. Then, because $\tilde{\pi}_c^{a_2,a_3}$ is canonical with respect to $\tilde{\pi}^{a_2,a_3}$, the vertex v is on $\tilde{\pi}^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$. Since $\tilde{\pi}^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$ is a subpath of $\pi^{a_2,a_3} \oplus \rho^{a_2}$ excluding c_{a_2}, see step 3, v is not on ρ^{a_2}. In particular, $v \neq d_{a_2}$ and thus $v \in \mathbf{S} \setminus \{d_{a_2}, d_{a_3}\}$. Moreover, since $v \neq d_{a_3}$ and $g_{a_1,a_3} \neq d_{a_2}$ by the respective definitions of v and g_{a_2,a_3}, the vertex v is a non-end-point vertex on $\tilde{\pi}_c^{a_2,a_3}$. Since $\tilde{\pi}_c^{a_2,a_3}$ is collider-free by means of being canonical with respect to $\tilde{\pi}^{a_2,a_3}$ and since $v \in \mathbf{S} \setminus \{d_{a_2}, d_{a_3}\}$, we arrive at a contradiction to the assumption that $\tilde{\pi}_c^{a_2,a_3}$ is active given $\mathbf{S} \setminus \{d_{a_2}, d_{a_3}\}$.
3. Third, assume v is on κ^{a_2}. Then, since every element of $\mathbf{S}$ is observable, v is h_{a_1,a_2} or g_{a_2,a_3} and thus on $\tilde{\pi}_c^{a_1,a_2}(d_{a_1}, h_{a_1,a_2})$ or $\tilde{\pi}_c^{a_2,a_3}(g_{a_2,a_3}, d_{a_3})$. These cases have already been covered by the previous two points.

We have thus shown that $\tilde{\pi}_c^{a_1,a_3}$ is active given $\mathbf{S} \setminus \{d_{a_1}, d_{a_3}\}$, which completes the proof. □

In case at least one of $\tilde{\pi}_c^{a_1,a_2} = \tilde{\pi}_{ci}^{a-1,a}$ and $\tilde{\pi}_c^{a_2,a_3} = \tilde{\pi}_{ci}^{a,a+1}$ is out of $d_a = d_{a_2}$, we can thus collectively replace them by a path $\tilde{\pi}_c^{a_1,a_3}$ between $d_{a_1} = d_{a-1}$ and $d_{a_3} = d_{a+1}$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ that is active given $\mathbf{S} \setminus \{d_{a_1}, d_{a_3}\}$. Moreover, since $\tilde{\pi}_c^{a_1,a_3}$ is canonical with respect to $\tilde{\pi}^{a_1,a_3}$ and since Lemma D.29 only used that $\tilde{\pi}_c^{a_1,a_2}$ and $\tilde{\pi}_c^{a_2,a_3}$ are, respectively, canonical with respect to $\tilde{\pi}^{a_1,a_2}$ and $\tilde{\pi}^{a_2,a_3}$ as well as , respectively, active given $\mathbf{S} \setminus \{d_{a_1}, d_{a_2}\}$ and $\mathbf{S} \setminus \{d_{a_2}, d_{a_3}\}$, this procedure can be repeated in case, for example, $\tilde{\pi}_c^{a_1,a_3}$ or $\tilde{\pi}_{ci}^{a_3,a_4}$ is out of d_{a_3}, and so on.

LEMMA D.30. *Let (i, t_i) and (j, t_j) with $t - p \leq t_i, t_j \leq t$ be distinct observable vertices in $\mathcal{D}$ and let $\mathbf{S} \subseteq \mathbf{O}(t-p, t) \setminus \{(i, t_i), (j, t_j)\}$. Then: If (i, t_i) and (j, t_j) are d-connected given $\mathbf{S}$ in $\mathcal{D}$, then (i, t_i) and (j, t_j) are d-connected given $\mathbf{S}$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$.*

PROOF OF LEMMA D.30. Let (i,t_i) and (j,t_j) be d-connected given $\mathbf{S} \subseteq \mathbf{O}(t-p,t) \setminus \{(i,t_i),(j,t_j)\}$. Then, in $\mathcal{D}$ there is path π between (i,t_i) and (j,t_j) that is active given $\mathbf{S}$. The proof is by induction over n_c, the number of colliders on π.

Induction base case: $n_c = 0$
In this case, (i,t_i) and (j,t_j) are d-connected given $\mathbf{S}$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ according to Lemma D.23.

Induction step: $n_c \to n_c + 1$
In this case, π has $n_c + 1 \geq 1$ colliders and according to the assumption of induction we have already proven the statement for paths that have at most n_c colliders.

We may without loss of generality assume that π has a collider extension structure, because if not then according to Lemma D.25 there is path π' between (i,t_i) and (j,t_j) in $\mathcal{D}$ with at most n_c colliders that is active given $\mathbf{S}$ and hence, by assumption of induction, (i,t_i) and (j,t_j) are d-connected given $\mathbf{S}$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$. Therefore, the assumptions and notation of Definition D.26 apply.

Consider the following algorithmic procedure:

1. For all $0 \leq a \leq n_c$ let $\tilde{\pi}_c^{a,a+1}$ be $\tilde{\pi}_{ci}^{a,a+1}$, that is, let $\tilde{\pi}_c^{a,a+1}$ be the canonically induced path of $\tilde{\pi}^{a,a+1}$.
2. Set m to n_c.
3. Let $\sigma : \{0,1,\ldots,m+1\} \mapsto \{0,1,\ldots,n_c+1\}$ be the identity map.
4. While there is an integer a with $1 \leq a \leq m$ such that at least one of $\tilde{\pi}_c^{\sigma(a-1),\sigma(a)}$ and $\tilde{\pi}_c^{\sigma(a),\sigma(a+1)}$ is out of $d_{\sigma(a)}$:
   a) Let b be the smallest such a.
   b) Let $\tilde{\pi}_c^{\sigma(b-1),\sigma(b+1)} = \tilde{\pi}_c^{a_1,a_3}$ be a path as in Lemma D.29 applied to $\tilde{\pi}_c^{\sigma(b-1),\sigma(b)} = \tilde{\pi}_c^{a_1,a_2}$ and $\tilde{\pi}_c^{\sigma(b),\sigma(b+1)} = \tilde{\pi}_c^{a_2,a_3}$.
   c) If $\sigma(b-1) = 0$ and $\sigma(b+1) = n_c + 1$, then return $\tilde{\pi}_c^{\sigma(b-1),\sigma(b+1)} = \tilde{\pi}_c^{0,n_c+1}$ and terminate.
   d) Decrease m by one.
   e) Let $\sigma' : \{0,1,\ldots,m+1\} \mapsto \{0,1,\ldots,n_c+1\}$ be such that $\sigma'(a) = \sigma(a)$ for all $1 \leq a < b$ and $\sigma'(a) = \sigma(a+1)$ for all $b \leq a \leq m+1$.
   f) Replace σ by σ'.
5. Return the ordered sequence of paths $\tilde{\pi}_c^{\sigma(0),\sigma(1)}, \ldots, \tilde{\pi}_c^{\sigma(m),\sigma(m+1)}$ and terminate.

According to Lemma D.27 and Lemma D.22, the canonically induced paths $\tilde{\pi}_{ci}^{a,a+1}$ are for all $0 \leq a \leq n_c$ active given $\mathbf{S} \setminus \{d_a, d_{a+1}\}$ and canonical with respect to $\tilde{\pi}^{a,a+1}$. Lemma D.29 thus guarantees that all paths $\tilde{\pi}_c^{a_1,a_3}$ with $0 \leq a_1 < a_3 \leq n_c + 1$ constructed in step 4(b) of the above procedure are active given $\mathbf{S} \setminus \{d_{a_1}, d_{a_3}\}$ and canonical with respect to $\tilde{\pi}^{a_1,a_3}$. Thus, if the algorithm terminates in step 4(c), it returns a path $\tilde{\pi}_c^{0,n_c+1}$ between $d_0 = (i,t_i)$ and $d_{n_c+1} = (j,t_j)$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ that is active given $\mathbf{S} = \mathbf{S} \setminus \{d_0, d_{n_c+1}\}$. If the algorithm terminates in step 5, then the following is true:

1. $\tilde{\pi}_c^{\sigma(0),\sigma(1)}$ is a path in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ between $d_{\sigma(0)} = d_0 = (i,t_i)$ and $d_{\sigma(1)}$ that is into $d_{\sigma(1)}$ and active given $\mathbf{S} \setminus \{d_{\sigma(0)}, d_{\sigma(1)}\}$.
2. For all $1 \leq a \leq m-1$ the path $\tilde{\pi}_c^{\sigma(a),\sigma(a+1)}$ is a path in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ between $d_{\sigma(a)}$ and $d_{\sigma(a+1)}$ that is into both $d_{\sigma(a)}$ and $d_{\sigma(a+1)}$ and active given $\mathbf{S} \setminus \{d_{\sigma(a)}, d_{\sigma(a+1)}\}$.
3. $\tilde{\pi}_c^{\sigma(m),\sigma(m+1)}$ is a path in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ between $d_{\sigma(m)}$ and $d_{\sigma(m+1)} = d_{n_c+1} = (j,t_j)$ that is into $d_{\sigma(m)}$ and active given $\mathbf{S} \setminus \{d_{\sigma(m)}, d_{\sigma(m+1)}\}$.

Further, by definition of collider extension structures, d_a is an ancestor of $\mathbf{S}$ for all $1 \leq a \leq n_c$. Lemma 3.3.1 in Spirtes, Glymour and Scheines (2000) thus applies to the ordered sequence of paths $\tilde{\pi}_c^{\sigma(0),\sigma(1)}, \ldots, \tilde{\pi}_c^{\sigma(m),\sigma(m+1)}$ and guarantees the existence of a path between (i, t_i) and (j, t_j) in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$ that is active given $\mathbf{S}$.

Hence, (i, t_i) and (j, t_j) are d-connected given $\mathbf{S}$ in $\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$. □

LEMMA D.31. *$\mathcal{M}^p(\mathcal{D})$ is a subgraph of $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$.*

PROOF OF LEMMA D.31. As an immediate consequence of Lemma D.30, every adjacency in $\mathcal{M}^p(\mathcal{D})$ is also in $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$. The statement follows with Lemma D.16 because the orientation of edges are uniquely determined by the ancestral relationships. □

PROOF OF LEMMA 4.14. First, $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$ is a subgraph of $\mathcal{M}^p(\mathcal{D})$ according to Lemma D.18. Second, $\mathcal{M}^p(\mathcal{D}))$ is a subgraph of $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$ according to Lemma D.31. Hence, $\mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))) = \mathcal{M}^p(\mathcal{D})$. □

D.6. *Proofs for Section 4.6.*

LEMMA D.32. *The canonical ts-DAG $\mathcal{D}_c(\mathcal{G})$ of an acyclic directed mixed graph $\mathcal{G}$ with time series structure is a ts-DAG.*

PROOF OF LEMMA D.32. The time series structure of $\mathcal{D}_c(\mathcal{G})$ with $\mathbf{T} = \mathbb{Z}$ is apparent from the first point of Definition 4.13, the repeating edges property is enforced explicitly in the second point of Definition 4.13, and time order is enforced explicitly in the second point of Definition 4.13 by only considering edges in $\mathbf{E}^{\text{stat}}_{\rightarrow}$ of the form $((i, t-\tau), (j, t))$ in the first and second point of Definition 4.13. Assume there is a directed cycle in $\mathcal{D}_c(\mathcal{G})$. Because as apparent from the second point of Definition 4.13 there are no edges into unobservable vertices, all vertices on the directed cycle are observable. Moreover, due to time order all vertices on the directed cycle must be at a single time step. Due to repeating edges there thus is a directed cycle at time t in $\mathcal{D}_c(\mathcal{G})$. Since the second point of Definition 4.13 further shows that all edges between observable vertices at time t in $\mathcal{D}_c(\mathcal{G})$ are also in $\text{stat}(\mathcal{G})$, which is a subgraph of $\mathcal{G}$, we get a contradiction to the acyclicity of $\mathcal{G}$. □

PROOF OF THEOREM 1. **If.** The premise is $\mathcal{M} = \mathcal{M}^p(\mathcal{D}_c(\mathcal{M}))$. Since according to Lemma D.32 the canonical ts-DAG $\mathcal{D}_c(\mathcal{M})$ is a ts-DAG, we can choose $\mathcal{D} = \mathcal{D}_c(\mathcal{M})$ and get $\mathcal{M} = \mathcal{M}^p(\mathcal{D})$.

Only if. The premise is $\mathcal{M} = \mathcal{M}^p(\mathcal{D})$. Together with Lemma 4.14, which says $\mathcal{M}^p(\mathcal{D}) = \mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$, then $\mathcal{M} = \mathcal{M}^p(\mathcal{D}) = \mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))) = \mathcal{M}^p(\mathcal{D}_c(\mathcal{M}))$. □

PROOF OF THEOREM 2. **If.** The premise is $\mathcal{G} = \mathcal{M}^p(\mathcal{D}_c(\mathcal{G}))$ with $\mathcal{G}$ acyclic. Since according to Lemma D.32 the canonical ts-DAG $\mathcal{D}_c(\mathcal{G})$ is a ts-DAG, we can choose $\mathcal{D} = \mathcal{D}_c(\mathcal{G})$ and get $\mathcal{G} = \mathcal{M}^p(\mathcal{D})$.

Only if. The premise is $\mathcal{G} = \mathcal{M}^p(\mathcal{D})$. Together with Lemma 4.14, which says $\mathcal{M}^p(\mathcal{D}) = \mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$, then $\mathcal{G} = \mathcal{M}^p(\mathcal{D}) = \mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))) = \mathcal{M}^p(\mathcal{D}_c(\mathcal{G}))$. Moreover, $\mathcal{G}$ is acyclic because it equals the DMAG $\mathcal{M}^p(\mathcal{D})$ and DMAGs are acyclic. □

D.7. *Proofs for Section 4.7 and of Lemma B.6.*

PROOF OF LEMMA 4.20. Combine Lemma D.13, according to which $\mathcal{D}_c(\mathcal{M}^p_{\text{st}}(\mathcal{D})) = \mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$, with Lemma 4.14, according to which $\mathcal{M}^p(\mathcal{D}) = \mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$. □

PROOF OF LEMMA B.6. Assume that both $\mathcal{M}_1$ and $\mathcal{M}_2$ are ts-DMAGs, i.e., $\mathcal{M}_1 = \mathcal{M}^p(\mathcal{D}_1)$ and $\mathcal{M}_2 = \mathcal{M}^p(\mathcal{D}_2)$ for some ts-DAGs $\mathcal{D}_1$ and $\mathcal{D}_2$. Then, combining the premise $\text{stat}(\mathcal{M}_1) = \text{stat}(\mathcal{M}_2)$ with Lemma 4.20 leads to the contradiction $\mathcal{M}_1 = \mathcal{M}_2$. □

PROOF OF LEMMA 4.21. **If.** The premise is $\mathcal{M} = \mathcal{M}^p_{\mathrm{st}}(\mathcal{D}_c(\mathcal{M}))$. Since according to Lemma D.32 the canonical ts-DAG $\mathcal{D}_c(\mathcal{M})$ is a ts-DAG, we can choose $\mathcal{D} = \mathcal{D}_c(\mathcal{M})$ and get $\mathcal{M} = \mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$.

Only if. The premise is $\mathcal{M} = \mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$. We then get $\mathcal{M}^p(\mathcal{D}) = \mathcal{M}^p(\mathcal{D}_c(\mathcal{M}))$ according to Lemma 4.20, which by applying the operation of stationarification to both sides of this equality gives that $\mathcal{M} = \mathcal{M}^p_{\mathrm{st}}(\mathcal{D}_c(\mathcal{M}))$. □

PROOF OF LEMMA 4.22. **If.** The premise is $\mathcal{G} = \mathcal{M}^p_{\mathrm{st}}(\mathcal{D}_c(\mathcal{G}))$ with $\mathcal{G}$ acyclic. Since according to Lemma D.32 the canonical ts-DAG $\mathcal{D}_c(\mathcal{G})$ is a ts-DAG, we can choose $\mathcal{D} = \mathcal{D}_c(\mathcal{G})$ and get $\mathcal{G} = \mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$.

Only if. The premise is $\mathcal{G} = \mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$. We then get $\mathcal{M}^p(\mathcal{D}) = \mathcal{M}^p(\mathcal{D}_c(\mathcal{G}))$ according to Lemma 4.20, which by applying the operation of stationarification to both sides of this equality gives that $\mathcal{G} = \mathcal{M}^p_{\mathrm{st}}(\mathcal{D}_c(\mathcal{G}))$. Moreover, $\mathcal{G}$ is acyclic because it equals the DMAG $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ and DMAGs are acyclic. □

E. Proofs for Section 5.

E.1. *Proofs for Section 5.3.*

LEMMA E.1. *Let $\mathcal{D}$ be a ts-DAG and $\mathcal{A} \in \{\mathcal{A}_{\mathrm{to}}, \mathcal{A}_{\mathrm{ta}}, \mathcal{A}_{\mathcal{D}}\}$. Then, $\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A})$ has repeating orientations.*

PROOF OF LEMMA E.1. Note that $\mathcal{A}_{\mathcal{D}}$ is stronger than $\mathcal{A}_{\mathrm{ta}}$ and that $\mathcal{A}_{\mathrm{ta}}$ is stronger than $\mathcal{A}_{\mathrm{to}}$. Thus, every graph consistent with $\mathcal{A}$ has repeating orientations. The statement then follows from the definition of m.i. DPAGs because every element in $[\mathcal{M}]_{\mathcal{A}}$ has repeating orientations. □

PROOF OF LEMMA 5.5. From Lemma E.1 we know that $\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A})$ has repeating orientations. Moreover, $\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A})$ has past-repeating adjacencies because its skeleton by definition equals the skeleton of $\mathcal{M}^p(\mathcal{D})$.

Let (i, t_i) and (j, t_j) with $t - p \leq t_i \leq t_j \leq t$ be distinct observable vertices. According to Lemma D.1, these vertices are adjacent in $\mathrm{stat}(\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A}))$ if and only if $(i, t - (t_j - t_i))$ and (j, t) are adjacent in $\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A})$. Because the skeletons of $\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A})$ and $\mathcal{M}^p(\mathcal{D})$ are the same, Lemma 4.7 gives that $(i, t - (t_j - t_i))$ and (j, t) are adjacent in $\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A})$ if and only if (i, t_i) and (j, t_j) are adjacent in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$. Consequently, $\mathrm{stat}(\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A}))$ and $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ have the same skeleton.

Moreover, consider an unambiguous edge mark in $\mathrm{stat}(\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A}))$. This edge mark is also in $\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A})$ and therefore corresponds to an ancestral relationship in $\mathcal{M}^p(\mathcal{D})$. Because according to Lemma 4.10 the graphs $\mathcal{M}^p(\mathcal{D})$ and $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ have the same ancestral relationships, the same unambiguous edge mark is then also in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$. □

LEMMA E.2. *Let $\mathcal{M}$ be a DMAG with time series structure that has repeating orientations and past-repeating adjacencies and for part 2 in addition is time ordered. Then:*

1. *Let $(i, t_i){*\!\!-\!\!*}(j, t_j){*\!\!-\!\!*}(k, t_k)$ with $t_j \leq \max(t_i, t_k)$ be an unshielded triple in $\mathrm{stat}(\mathcal{M})$ and let $\Delta t = t_j - \max(t_i, t_k)$. Then:*
   a) *$(i, t_i + \Delta t){*\!\!-\!\!*}(j, t_j + \Delta t){*\!\!-\!\!*}(k, t_k + \Delta t)$ is an unshielded triple in $\mathcal{M}$.*
   b) *$(i, t_i + \Delta t){*\!\!-\!\!*}(j, t_j + \Delta t){*\!\!-\!\!*}(k, t_k + \Delta t)$ is oriented as a collider in $\mathcal{M}$ if and only if $(i, t_i){*\!\!-\!\!*}(j, t_j){*\!\!-\!\!*}(k, t_k)$ is oriented as a collider in $\mathrm{stat}(\mathcal{M})$.*
2. *Let $\pi = (l, t_l) \ldots {*\!\!\rightarrow}(i, t_i){\leftarrow\!\!*}(j, t_j){*\!\!\rightarrow}(k, t_k)$ with $t_j \leq \max(t_l, t_k)$ be a discriminating path for (j, t_j) in $\mathrm{stat}(\mathcal{M})$ and let $\Delta t = t_j - \max(t_l, t_k)$. Then:*

a) $\pi_{\Delta t}$, *the copy of* π *shifted forward in time by* Δt *time steps, is a discriminating path for* $(j, t_j + \Delta t)$ *in* $\mathcal{M}$.
b) $(i, t_i + \Delta t){*\!\!-\!\!*}(j, t_j + \Delta t){*\!\!-\!\!*}(k, t_k + \Delta t)$ *is oriented as a collider in* $\mathcal{M}$ *if and only if* $(i, t_i){*\!\!-\!\!*}(j, t_j){*\!\!-\!\!*}(k, t_k)$ *is oriented as a collider in* $\mathrm{stat}(\mathcal{M})$.

PROOF OF LEMMA E.2. **1(a)** The repeating edges property of $\mathrm{stat}(\mathcal{M})$ together with $t_j \leq \max(t_i, t_k)$ implies that $(i, t_i + \Delta t){*\!\!-\!\!*}(j, t_j + \Delta t){*\!\!-\!\!*}(k, t_k + \Delta t)$ is an unshielded triple in $\mathrm{stat}(\mathcal{M})$. Lemma D.1 then guarantees that $(i, t_i + \Delta t) = (i, t - (\max(t_i, t_k) - t_i))$ and $(k, t_k + \Delta t) = (k, t - (\max(t_i, t_k) - t_i))$ are nonadjacent in $\mathcal{M}$, because else they would be adjacent in $\mathrm{stat}(\mathcal{M})$ too and hence $(i, t_i + \Delta t){*\!\!-\!\!*}(j, t_j + \Delta t){*\!\!-\!\!*}(k, t_k + \Delta t)$ would not be unshielded in $\mathrm{stat}(\mathcal{M})$. Since $\mathrm{stat}(\mathcal{M})$ is a subgraph of $\mathcal{M}$, we thus get that $(i, t_i + \Delta t){*\!\!-\!\!*}(j, t_j + \Delta t){*\!\!-\!\!*}(k, t_k + \Delta t)$ is an unshielded triple in $\mathcal{M}$.

2(a) By the definition of discriminating paths, all vertices on π other than, perhaps, (j, t_j) and/or (l, t_l) are ancestors of (k, t_k). Time order of $\mathcal{M}$ together with $t_j \leq \max(t_l, t_k)$ thus guarantees that all vertices on π are within $[t - p, \max(t_l, t_k)]$. In combination with the repeating edges property of $\mathrm{stat}(\mathcal{M})$ we thus see that $\pi_{\Delta t}$ is a discriminating path for $(j, t_j + \Delta t)$ in $\mathrm{stat}(\mathcal{M})$. Lemma D.1 then guarantees that $(l, t_l + \Delta t) = (l, t - (\max(t_l, t_k) - t_l))$ and $(k, t_k + \Delta t) = (k, t - (\max(t_l, t_k) - t_k))$ are nonadjacent in $\mathcal{M}$ because else they would be adjacent in $\mathrm{stat}(\mathcal{M})$ too and hence $\pi_{\Delta t}$ would not be a discriminating path in $\mathrm{stat}(\mathcal{M})$. Consequently, $\pi_{\Delta t}$ is a discriminating path in $\mathcal{M}$ because $\mathrm{stat}(\mathcal{M})$ is a subgraph of $\mathcal{M}$.

1(b) and 2(b) Because $\mathrm{stat}(\mathcal{M})$ has repeating edges, the triple $(i, t_i){*\!\!-\!\!*}(j, t_j){*\!\!-\!\!*}(k, t_k)$ is oriented as a collider in $\mathrm{stat}(\mathcal{M})$ if and only if $(i, t_i + \Delta t){*\!\!-\!\!*}(j, t_j + \Delta t){*\!\!-\!\!*}(k, t_k + \Delta t)$ is oriented as a collider in $\mathrm{stat}(\mathcal{M})$. Moreover, since $\mathrm{stat}(\mathcal{M})$ is a subgraph of $\mathcal{M}$ and $(i, t_i + \Delta t){*\!\!-\!\!*}(j, t_j + \Delta j){*\!\!-\!\!*}(k, t_k + \Delta t)$ is in $\mathcal{M}$, the triple $(i, t_i + \Delta t){*\!\!-\!\!*}(j, t_j + \Delta j){*\!\!-\!\!*}(k, t_k + \Delta t)$ is oriented as collider in $\mathrm{stat}(\mathcal{M})$ if and only if $(i, t_i + \Delta t){*\!\!-\!\!*}(j, t_j + \Delta j){*\!\!-\!\!*}(k, t_k + \Delta t)$ is oriented as a collider in $\mathcal{M}$. □

LEMMA E.3. *Let* $\mathcal{M}_1$ *and* $\mathcal{M}_2$ *be Markov equivalent DMAGs with time series structure that are time ordered and have repeating orientations and past-repeating adjacencies. Then,* $\mathrm{stat}(\mathcal{M}_1)$ *and* $\mathrm{stat}(\mathcal{M}_2)$ *are Markov equivalent DMAGs.*

PROOF OF LEMMA E.3. Both $\mathrm{stat}(\mathcal{M}_1)$ and $\mathrm{stat}(\mathcal{M}_2)$ are DMAGs by means of Lemma D.2. Next, we show that $\mathrm{stat}(\mathcal{M}_1)$ and $\mathrm{stat}(\mathcal{M}_2)$ are Markov equivalent. For this purpose, assume the opposite. Then, according to the characterizing of Markov equivalence of MAGs in Spirtes and Richardson (1997), at least one of the following statements is true:

1. The skeletons of $\mathrm{stat}(\mathcal{M}_1)$ and $\mathrm{stat}(\mathcal{M}_2)$ differ.
2. There is an unshielded triple $(i, t_i){*\!\!-\!\!*}(j, t_j){*\!\!-\!\!*}(k, t_k)$ in both $\mathrm{stat}(\mathcal{M}_1)$ and $\mathrm{stat}(\mathcal{M}_1)$ that is oriented as a collider in $\mathrm{stat}(\mathcal{M}_a)$ with $a \in \{1, 2\}$ and oriented as a noncollider in $\mathrm{stat}(\mathcal{M}_{\bar{a}})$ with $\bar{a} = 3 - a$.
3. There is a path π that is in both $\mathrm{stat}(\mathcal{M}_1)$ and $\mathrm{stat}(\mathcal{M}_1)$ a discriminating path for (j, t_j) such that (j, t_j) is a collider on π in $\mathrm{stat}(\mathcal{M}_a)$ with $a \in \{1, 2\}$ and a noncollider in $\mathrm{stat}(\mathcal{M}_{\bar{a}})$ with $\bar{a} = 3 - a$.

We now show that $\mathrm{stat}(\mathcal{M}_1)$ and $\mathrm{stat}(\mathcal{M}_2)$ do have the same skeleton and that both the second and third statement contradict Markov equivalence of $\mathcal{M}_1$ and $\mathcal{M}_2$.

Case 1: Skeleton. According to Lemma D.1, the skeletons of $\mathrm{stat}(\mathcal{M}_1)$ and $\mathrm{stat}(\mathcal{M}_2)$ are determined uniquely by, respectively, the skeletons of $\mathcal{M}_1$ and $\mathcal{M}_2$. Thus, since $\mathcal{M}_1$ and $\mathcal{M}_2$ have the same skeleton due to being Markov equivalent, also $\mathrm{stat}(\mathcal{M}_1)$ and $\mathrm{stat}(\mathcal{M}_2)$ have the same skeleton.

Case 2: Unshielded colliders. Since $(i, t_i){*\!\!-\!\!*}(j, t_j){*\!\!-\!\!*}(k, t_k)$ is oriented as a noncollider in $\mathrm{stat}(\mathcal{M}_{\bar{a}})$, the vertex (j, t_j) is in $\mathrm{stat}(\mathcal{M}_{\bar{a}})$ an ancestor (parent, in fact) of (i, t_i)

or (k, t_k). Time order of $\mathrm{stat}(\mathcal{M}_{\bar{a}})$ thus implies $t_j \leq \max(t_i, t_k)$. Hence, we can apply part 1 of Lemma E.2 to both $\mathrm{stat}(\mathcal{M}_{\bar{a}})$ and $\mathrm{stat}(\mathcal{M}_a)$, which gives that $(i, t_i + \Delta t){*\!\!-\!\!*}(j, t_j + \Delta t){*\!\!-\!\!*}(k, t_k + \Delta t)$ with $\Delta t_j = t - \max(t_i, t_k)$ is an unshielded collider in $\mathcal{M}_a$ and an unshielded noncollider in $\mathcal{M}_{\bar{a}}$. This observation contradicts the assumptiont that $\mathcal{M}_1$ and $\mathcal{M}_2$ are Markov equivalent.

Case 3: Discriminating paths. By definition of discriminating paths, π takes the form $\ldots {*\!\!\rightarrow}(i, t_i){\leftarrow\!\!*}(j, t_j){*\!\!\rightarrow}(k, t_k)$. Moreover, as follows from the definition of discriminating paths together with the absence of almost directed cycles, $(i, t_i){\leftrightarrow}(j, t_j)$ if $(j, t_j){\leftrightarrow}(k, t_k)$. In combination with the fact that (j, t_j) is in $\mathrm{stat}(\mathcal{M}_{\bar{a}})$ an ancestor (parent, in fact) of (i, t_i) or (k, t_k) by means of (j, t_j) being a noncollider on π in $\mathrm{stat}(\mathcal{M}_{\bar{a}})$, we thus find that (j, t_j) is in $\mathrm{stat}(\mathcal{M}_{\bar{a}})$ an ancestor (parent, in fact) of (k, t_k). Time order of $\mathrm{stat}(\mathcal{M}_{\bar{a}})$ thus implies $t_j \leq t_k \leq \max(t_l, t_k)$. Hence, we can apply part 2 of Lemma E.2 to both $\mathrm{stat}(\mathcal{M}_{\bar{a}})$ and $\mathrm{stat}(\mathcal{M}_a)$, which gives that $\pi_{\Delta t}$, the copy of π that is shifted forward in time by $\Delta t = t - \max(t_l, t_k)$ time steps, is a discriminating path for $(j, t_j + \Delta t)$ in both $\mathcal{M}_a$ and $\mathcal{M}_{\bar{a}}$ and that $(j, t_j + \Delta t)$ is a collider on $\pi_{\Delta t}$ in $\mathcal{M}_a$ whereas $(j, t_j + \Delta t)$ is a noncollider on $\pi_{\Delta t}$ in $\mathcal{M}_{\bar{a}}$. This observation contradicts Markov equivalence of $\mathcal{M}_1$ and $\mathcal{M}_2$. □

PROOF OF THEOREM 3. **1.** Note that $\mathrm{stat}(\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A}))$ and $\mathcal{P}(\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}), \mathcal{A}^{\mathrm{stat}})$ have the same skeleton because both of them are DPAGs for $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$, as follows from Lemma 5.5. We prove the statement by showing that $(i, t_i){\circ\!\!-\!\!*}(j, t_j)$ in $\mathrm{stat}(\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A}))$ implies $(i, t_i){\circ\!\!-\!\!*}(j, t_j)$ in $\mathcal{P}(\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}), \mathcal{A}^{\mathrm{stat}})$.

Let the edge $(i, t_i){\circ\!\!-\!\!*}(j, t_j)$ be in $\mathrm{stat}(\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A}))$. Then, $(i, t_i){\circ\!\!-\!\!*}(j, t_j)$ is also in $\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A})$ because $\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A})$ is a supergraph of $\mathrm{stat}(\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A}))$. By definition of m.i. DPAGs, there thus are DMAGs $\mathcal{M}_1$ and $\mathcal{M}_2$ in $[\mathcal{M}^p(\mathcal{D})]_{\mathcal{A}}$ such that $(i, t_i){\rightarrow}(j, t_j)$ in $\mathcal{M}_1$ and $(i, t_i){\leftarrow\!\!*}(j, t_j)$ in $\mathcal{M}_2$. Without loss of generality we may assume that $\mathcal{M}_1$ or $\mathcal{M}_2$ is $\mathcal{M}^p(\mathcal{D})$.

Since (1) according to Lemma E.3 $\mathrm{stat}(\mathcal{M}_1)$ and $\mathrm{stat}(\mathcal{M}_2)$ are Markov equivalent DMAGs and since (2) either $\mathcal{M}_1$ or $\mathcal{M}_2$ is $\mathcal{M}^p(\mathcal{D})$ and hence either $\mathrm{stat}(\mathcal{M}_1) = \mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$ or $\mathrm{stat}(\mathcal{M}_2) = \mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$, we thus get that both $\mathrm{stat}(\mathcal{M}_1)$ and $\mathrm{stat}(\mathcal{M}_2)$ are in $[\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})]$. Recall that $\mathrm{stat}(\mathcal{M})$ always has repeating ancestral relationships (, and hence also repeating orientations), that $\mathrm{stat}(\mathcal{M})$ is time ordered if $\mathcal{M}$ is time ordered, and that $\mathrm{stat}(\mathcal{M})$ is a stationarified ts-DMAG if $\mathcal{M}$ is a ts-DMAG. Hence, both $\mathrm{stat}(\mathcal{M}_1)$ and $\mathrm{stat}(\mathcal{M}_2)$ are in the set $[\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})]_{\mathcal{A}^{\mathrm{stat}}}$.

Since (1) both $\mathrm{stat}(\mathcal{M}_1)$ and $\mathrm{stat}(\mathcal{M}_2)$ are in $[\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})]_{\mathcal{A}^{\mathrm{stat}}}$ and since (2) (i, t_i) and (j, t_j) are adjacent in $\mathcal{M}^p_{\mathrm{st}}(\mathcal{D})$, the vertices (i, t_i) and (j, t_j) are also adjacent in both $\mathrm{stat}(\mathcal{M}_1)$ and $\mathrm{stat}(\mathcal{M}_2)$. Since $\mathrm{stat}(\mathcal{M}_i)$ is a subgraph of $\mathcal{M}_i$ for $i = 1, 2$, we conclude that $(i, t_i){\rightarrow}(j, t_j)$ in $\mathrm{stat}(\mathcal{M}_1)$ and $(i, t_i){\leftarrow\!\!*}(j, t_j)$ in $\mathrm{stat}(\mathcal{M}_2)$. Hence, $(i, t_i){\circ\!\!-\!\!*}(j, t_j)$ in $\mathcal{P}(\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}), \mathcal{A}^{\mathrm{stat}})$.

2. This claim immediately follows from part 1 of Theorem 3 because $\mathrm{stat}(\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A}))$ is a subgraph of $\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A})$.

3. The three graphs obtained by applying stationarification $\mathrm{stat}(\cdot)$ to the graph in parts (c), (d) and (e) of Figure 9 in the main text provide such examples, see also the discussion in Example 5.6 in the main text.

4. This claim immediately follows from part 3 of Theorem 3 because $\mathrm{stat}(\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A}))$ is a subgraph of $\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A})$ and because $\mathrm{stat}(\mathcal{P}(\mathcal{M}^p(\mathcal{D}), \mathcal{A}))$ and $\mathcal{P}(\mathcal{M}^p_{\mathrm{st}}(\mathcal{D}), \mathcal{A})$ have the same skeleton. □

E.2. *Proofs for Section 5.4.*

PROOF OF LEMMA 5.9. The premise that $(i, t_i){\circ\!\!-\!\!*}(j, t_j)$ is in the ts-DPAG $\mathcal{P}^p(\mathcal{D})$ by the definitions of m.i. DPAGs and the background knowledge $\mathcal{A}_{\mathcal{D}}$ means: There are ts-DAGs $\mathcal{D}_1$ and $\mathcal{D}_2$ such that both $\mathcal{M}^p(\mathcal{D}_1)$ and $\mathcal{M}^p(\mathcal{D}_2)$ are Markov equivalent to $\mathcal{M}^p(\mathcal{D})$ and $(i,t_i){\rightarrow}(j,t_j)$ in $\mathcal{M}^p(\mathcal{D}_1)$ and $(i,t_i){\leftarrow\!\!*}(j,t_j)$ in $\mathcal{M}^p(\mathcal{D}_2)$. Consequently, $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{M}^p(\mathcal{D}_1))$ and $(i,t_i) \notin \mathrm{an}((j,t_j),\mathcal{M}^p(\mathcal{D}_2))$ and thus, using Lemma C.1, $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{D}_1)$ and $(i,t_i) \notin \mathrm{an}((j,t_j),\mathcal{D}_2)$. □

F. Proofs for Section B.8.

F.1. *Proofs for Section B.8.1.*

LEMMA F.1. *Let (i,t_i) and (j,t_j) with $t-p \le t_i, t_j \le t$ be distinct observable vertices in a ts-DAG $\mathcal{D}$ and let $\Delta t > 0$. Then: There is $\mathbf{S} \subseteq \mathbf{O}(t-p,t) \setminus \{(i,t_i),(j,t_j)\}$ such that $(i,t_i) \perp\!\!\!\perp (j,t_j) \mid \mathbf{S}$ if and only if there is $\mathbf{S}' \subseteq \mathbf{O}(t-p,t+\Delta t) \setminus \{(i,t_i),(j,t_j)\}$ such that $(i,t_i) \perp\!\!\!\perp (j,t_j) \mid \mathbf{S}'$.*

PROOF OF LEMMA F.1. **If.** The premise is $(i,t_i) \perp\!\!\!\perp (j,t_j) \mid \mathbf{S}'$ with $\mathbf{S}' \subseteq \mathbf{O}(t-p,t+\Delta t) \setminus \{(i,t_i),(j,t_j)\}$ for some $\Delta t > 0$. According to Lemma S5 in the supplementary material of Gerhardus and Runge (2020) this premise implies $(i,t_i) \perp\!\!\!\perp (j,t_j) \mid \mathbf{S}$, where $\mathbf{S}$ is the restriction of $\mathbf{S}'$ to ancestors of (i,t_i) and (j,t_j), i.e., where $\mathbf{S} = \mathbf{S}' \cap (\mathrm{an}((i,t_i),\mathcal{D}) \cup \mathrm{an}((j,t_j),\mathcal{D}))$. By time order of $\mathcal{D}$, no element of $\mathrm{an}((i,t_i),\mathcal{D}) \cup \mathrm{an}((j,t_j),\mathcal{D})$ is after $\max(t_i,t_j) \le t$ and hence $\mathbf{S} \subseteq \mathbf{O}(t-p,t) \setminus \{(i,t_i),(j,t_j)\}$.
Only if. Take $\mathbf{S} = \mathbf{S}'$. □

PROOF OF LEMMA B.10. **1.** The combination of past-repeating adjacencies and repeating orientations implies that $\mathcal{M}^{\tilde{p},[t-p-\Delta t, t-\Delta t]}(\mathcal{D})$ with $0 \le \Delta t < \tilde{p} - p$ is a subgraph of $\mathcal{M}^{\tilde{p},[t-\tilde{p},t-\tilde{p}+p]}(\mathcal{D})$. The statement then follows because by part 2 of Lemma B.10 the graphs $\mathcal{M}^{\tilde{p},[t-\tilde{p},t-\tilde{p}+p]}(\mathcal{D})$ and $\mathcal{M}^p(\mathcal{D})$ are equal up to relabeling vertices.
2. We first show that $\mathcal{M}^p(\mathcal{D})$ and $\mathcal{M}^{\tilde{p},[t-\tilde{p},t-\tilde{p}+p]}(\mathcal{D})$ have the same skeleton up to relabeling vertices. To this end, consider two distinct observable vertices (i,t_i) and (j,t_j) with $t-p \le t_i, t_j \le t$ and let $\Delta t = \tilde{p} - p$. According to Lemma F.1 there is $\mathbf{S}' \subseteq \mathbf{O}(t-\tilde{p},t) \setminus \{(i,t_i-\Delta t),(j,t_j-\Delta t)\}$ such that $(i,t_i-\Delta t) \perp\!\!\!\perp (j,t_j-\Delta t) \mid \mathbf{S}'$ if and only if there is $\mathbf{S} \subseteq \mathbf{O}(t-\tilde{p},t-\Delta t) \setminus \{(i,t_i-\Delta t),(j,t_j-\Delta t)\}$ such that $(i,t_i-\Delta t) \perp\!\!\!\perp (j,t_j-\Delta t) \mid \mathbf{S}$. By the repeating separating sets property of $\mathcal{D}$, the existence of $\mathbf{S} \subseteq \mathbf{O}(t-\tilde{p},t-\Delta t) \setminus \{(i,t_i-\Delta t),(j,t_j-\Delta t)\}$ such that $(i,t_i-\Delta t) \perp\!\!\!\perp (j,t_j-\Delta t) \mid \mathbf{S}$ is in turn equivalent to the existence of $\mathbf{S}_{\Delta t} \subseteq \mathbf{O}(t-p,t) \setminus \{(i,t_i),(j,t_j)\}$ such that $(i,t_i) \perp\!\!\!\perp (j,t_j) \mid \mathbf{S}_{\Delta t}$. Hence, (i,t_i) and (j,t_j) are adjacent in $\mathcal{M}^p(\mathcal{D})$ if and only if $(i,t_i-\Delta t)$ and $(j,t_j-\Delta t)$ are adjacent in $\mathcal{M}^{\tilde{p}}(\mathcal{D})$. This equivalence shows that $\mathcal{M}^p(\mathcal{D})$ and $\mathcal{M}^{\tilde{p},[t-\tilde{p},t-\tilde{p}+p]}(\mathcal{D})$ have the same skeleton up to relabeling vertices.
Next, let $(i,t_i-\Delta t){*\!\!-\!\!*}(j,t_j-\Delta t)$ be an edge in $\mathcal{M}^{\tilde{p},[t-\tilde{p},t-\tilde{p}+p]}(\mathcal{D})$. Then, this edge $(i,t_i-\Delta t){*\!\!-\!\!*}(j,t_j-\Delta t)$ in $\mathcal{M}^{\tilde{p},[t-\tilde{p},t-\tilde{p}+p]}(\mathcal{D})$ and the edge $(i,t_i){*\!\!-\!\!*}(j,t_j)$ in $\mathcal{M}^p(\mathcal{D})$ have the same orientation because in both graphs the edge orientations signify ancestral relationships according to $\mathcal{D}$ and $\mathcal{D}$ has repeating ancestral relationships.
3. See Example B.12. □

LEMMA F.2. *Let $\mathcal{D}$ be a ts-DAG and $\tilde{p} > p \ge 0$. Then:*

1. $\mathcal{M}^p(\mathcal{D}) = \mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^{\tilde{p}}(\mathcal{D})))$.
2. *There are cases in which* $\mathcal{M}^{\tilde{p}}(\mathcal{D}) \neq \mathcal{M}^{\tilde{p}}(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$.

PROOF OF LEMMA F.2. **1.** Let $\mathcal{D}_1 = \mathcal{D}$ and $\mathcal{D}_2 = \mathcal{D}_c(\mathcal{M}^{\tilde{p}}(\mathcal{D}))$. We then get the equality $\mathcal{M}^{\tilde{p}}(\mathcal{D}_1) = \mathcal{M}^{\tilde{p}}(\mathcal{D}) = \mathcal{M}^{\tilde{p}}(\mathcal{D}_c(\mathcal{M}^{\tilde{p}}(\mathcal{D}))) = \mathcal{M}^{\tilde{p}}(\mathcal{D}_2)$, where the second equality follows from Lemma 4.14. Thus, $\mathcal{M}^p(\mathcal{D}_1) = \mathcal{M}^p(\mathcal{D}_2)$ according to part 1 of Lemma B.11.

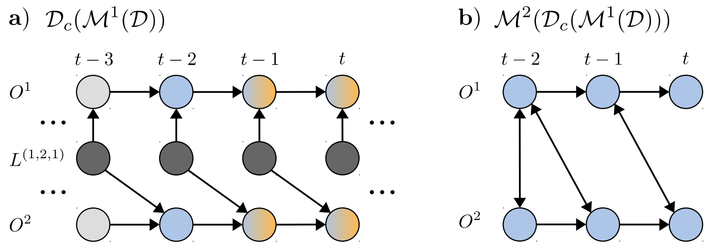

FIG F. *Example for proving part 2 of Lemma F.2.* a) *The canonical ts-DAG $\mathcal{D}_c(\mathcal{M}^1(\mathcal{D}))$ of the ts-DMAG $\mathcal{M}^1(\mathcal{D})$ in part (b) of Figure B* . b) *The ts-DMAG $\mathcal{M}^2(\mathcal{D}_c(\mathcal{M}^1(\mathcal{D})))$ implied by the canonical ts-DAG in part (a).*

2. Consider the ts-DAG $\mathcal{D}$ in part (a) of Figure B, which respectively implies the ts-DMAGs $\mathcal{M}^1(\mathcal{D})$ and $\mathcal{M}^2(\mathcal{D})$ in parts (b) of (c) of the same figure. Part (a) of Figure F shows the canonical ts-DAG $\mathcal{D}_c(\mathcal{M}^1(\mathcal{D}))$ of $\mathcal{M}^1(\mathcal{D})$, which in turn marginalizes to the ts-DMAG $\mathcal{M}^2(\mathcal{D}_c(\mathcal{M}^1(\mathcal{D}))) \neq \mathcal{M}^2(\mathcal{D})$ shown in part (b) of the same figure. □

PROOF OF LEMMA B.11. **1.** This claim follows from the commutativity of the marginalization process as stated by Theorem 4.20 in Richardson and Spirtes (2002).

2. According to part 2 of Lemma F.2, there is a ts-DAG $\mathcal{D}$ and $\tilde{p} > p \geq 0$ such $\mathcal{M}^{\tilde{p}}(\mathcal{D}) \neq \mathcal{M}^{\tilde{p}}(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$.[2] Moreover, Lemma 4.14 implies the equality $\mathcal{M}^p(\mathcal{D}) = \mathcal{M}^p(\mathcal{D}_c(\mathcal{M}^p(\mathcal{D})))$. Take $\mathcal{D}_1 = \mathcal{D}$ and $\mathcal{D}_2 = \mathcal{D}_c(\mathcal{M}^p(\mathcal{D}))$. □

PROOF OF LEMMA B.13. **1.** We prove the contraposition: Let there be a circle mark on $(i,t_i){*\!\!-\!\!*}(j,t_j)$ in $\mathcal{P}^{\tilde{p}}(\mathcal{D})$. Without loss of generality we may assume this edge to be of the form $(i,t_i){\circ\!\!-\!\!*}\,(j,t_j)$. Thus, by Lemma 5.9, there are ts-DAGs $\mathcal{D}_1$ and $\mathcal{D}_2$—one of which without loss of generality is $\mathcal{D}$—such that the ts-DMAGs $\mathcal{M}^{\tilde{p}}(\mathcal{D}_1)$ and $\mathcal{M}^{\tilde{p}}(\mathcal{D}_2)$ are Markov equivalent and that $(i,t_i) \in \mathrm{an}((j,t_j),\mathcal{D}_1)$ and $(i,t_i) \notin \mathrm{an}((j,t_j),\mathcal{D}_2)$. According to commutativity of the marginalization process as stated in Theorem 4.20 in Richardson and Spirtes (2002), the ts-DMAGs $\mathcal{M}^p(\mathcal{D}_1)$ and $\mathcal{M}^p(\mathcal{D}_2)$ are respectively obtained by marginalizing $\mathcal{M}^{\tilde{p}}(\mathcal{D}_1)$ and $\mathcal{M}^{\tilde{p}}(\mathcal{D}_2)$ over the vertices within $[t-\tilde{p}, t-p-1]$. Hence, given that $\mathcal{M}^{\tilde{p}}(\mathcal{D}_1)$ and $\mathcal{M}^{\tilde{p}}(\mathcal{D}_2)$ are Markov equivalent, so are $\mathcal{M}^p(\mathcal{D}_1)$ and $\mathcal{M}^p(\mathcal{D}_2)$. These considerations show that $(i,t_i){\circ\!\!-\!\!*}\,(j,t_j)$ in $\mathcal{P}^p(\mathcal{D})$.

2. See Example B.15. □

PROOF OF LEMMA B.14. **1.** Let $(i,t_i){\circ\!\!-\!\!*}\,(j,t_j)$ be an edge in $\mathcal{P}^{\tilde{p},[t-p,t]}(\mathcal{D})$. Because $\mathcal{P}^{\tilde{p}}(\mathcal{D})$ is a supergraph of $\mathcal{P}^{\tilde{p},[t-p,t]}(\mathcal{D})$, the edge $(i,t_i){\circ\!\!-\!\!*}\,(j,t_j)$ is then also in $\mathcal{P}^{\tilde{p}}(\mathcal{D})$. Moreover, according to part 1 of Lemma B.10 (for $\Delta t = 0$) in combination with the definition of DPAGs, (i,t_i) and (j,t_j) are adjacent in $\mathcal{P}^p(\mathcal{D})$. We conclude that the edge $(i,t_i){\circ\!\!-\!\!*}\,(j,t_j)$ is $\mathcal{P}^p(\mathcal{D})$ too because the opposite would contradict part 1 of Lemma B.13.

2. See Example B.15. □

F.2. *Proofs for Section B.8.2.*

PROOF OF LEMMA B.16. **1.** Assume the opposite. Then, there is a strictly monotonically increasing sequence a_n of positive integers such that $\mathcal{M}^{p+a_n,[t-p,t]}(\mathcal{D}) \neq \mathcal{M}^{p+a_{n+1},[t-p,t]}(\mathcal{D})$

[2]Note that the proof of Lemma F.2 uses part 1 of Lemma B.11 but not part 2 of Lemma B.11, such that the proofs of these two lemmas are *not* circular.

for all $n \in \mathbb{N}$. Using part 1 of Lemma B.10 with $(p, \tilde{p}, \Delta t) \mapsto (p+a_n, p+a_{n+1}, 0)$, we see that $\mathcal{M}^{p+a_n}(\mathcal{D})$ is a subgraph of $\mathcal{M}^{p+a_{n+1},[t-(p+a_n),t]}(\mathcal{D})$ and hence $\mathcal{M}^{p+a_n,[t-p,t]}(\mathcal{D})$ is a subgraph of $\mathcal{M}^{p+a_{n+1},[t-p,t]}$. In combination with $\mathcal{M}^{p+a_n,[t-p,t]}(\mathcal{D}) \neq \mathcal{M}^{p+a_{n+1},[t-p,t]}(\mathcal{D})$ we thus find that $\mathcal{M}^{p+a_{n+1},[t-p,t]}(\mathcal{D})$ is a *proper* subgraph of $\mathcal{M}^{p+a_n,[t-p,t]}(\mathcal{D})$. Moreover, using part 1 of Lemma B.10 for $(p, \tilde{p}, \Delta t) \mapsto (p, p+a_n, 0)$ and for $(p, \tilde{p}, \Delta t) \mapsto (p, p+a_{n+1}, 0)$, we learn that both $\mathcal{M}^{p+a_n,[t-p,t]}(\mathcal{D})$ and $\mathcal{M}^{p+a_{n+1},[t-p,t]}(\mathcal{D})$ are subgraphs of $\mathcal{M}^{p}(\mathcal{D})$. By combining these observations we arrive at a contradiction since there are only finitely many edges between the finitely many vertices of $\mathcal{M}^{p}(\mathcal{D})$.

2. Assume the opposite. Then, there is a strictly monotonically increasing sequence a_n of positive integers such that $\mathcal{P}^{p+a_n,[t-p,t]}(\mathcal{D}) \neq \mathcal{P}^{p+a_{n+1},[t-p,t]}(\mathcal{D})$ for all $n \in \mathbb{N}$. Let m be such that $\mathcal{M}^{p+a_m,[t-p,t]}(\mathcal{D}) = \mathcal{M}^{p}_{\text{lim}}(\mathcal{D})$, which exists as a result of part 1 of Lemma B.16. Then, for all $n \geq m$ the skeletons of $\mathcal{P}^{p+a_n,[t-p,t]}(\mathcal{D})$ and $\mathcal{P}^{p+a_{n+1},[t-p,t]}(\mathcal{D})$ are equal. Using part 1 of Lemma B.13 we thus learn that for all $n \geq m$ there is a noncircle mark in $\mathcal{P}^{p+a_{n+1},[t-p,t]}(\mathcal{D})$ that is not in $\mathcal{P}^{p+a_n,[t-p,t]}(\mathcal{D})$. Since there are only finitely many circle marks on the finitely many edges in $\mathcal{M}^{p}_{\text{lim}}(\mathcal{D})$, this observation is a contradiction. □

LEMMA F.3. *Let $\mathcal{M}$ be a DMAG with vertex set $\mathbf{V}$ and let $\mathcal{M}[\mathbf{O}]$ be the induced subgraph of $\mathcal{M}$ on the subset of vertices $\mathbf{O} \subseteq \mathbf{V}$. Then, $\mathcal{M}[\mathbf{O}]$ is a DMAG.*

REMARK (on Lemma F.3). In particular, the graphs $\mathcal{M}^{\tilde{p},[t_1,t_2]}(\mathcal{D})$ defined in Definition B.9 are DMAGs.

PROOF OF LEMMA F.3. We have to show that $\mathcal{M}[\mathbf{O}]$ does not have directed cycles, does not have almost directed cycles, and is maximal.

No (almost) directed cycles: Since $\mathcal{M}[\mathbf{O}]$ is a subgraph of $\mathcal{M}$ and $\mathcal{M}$ does neither have a directed nor an almost directed cycle, also $\mathcal{M}[\mathbf{O}]$ does neither have a directed nor an almost directed cycle.

Maximality: Assume the opposite, i.e., assume in $\mathcal{M}[\mathbf{O}]$ there are nonadjacent vertices i and j between which there is an inducing path π. Since $\mathcal{M}[\mathbf{O}]$ is a subgraph of $\mathcal{M}$, the same path π between i and j is also in $\mathcal{M}$. Maximality of $\mathcal{M}$ thus implies that i and j are adjacent in $\mathcal{M}$. By the definition of induced subgraphs the nodes i and j would then also be adjacent in $\mathcal{M}[\mathbf{O}]$, a contradiction. □

PROOF OF LEMMA B.18. **1.** According to part 1 of Lemma B.16 there is Δp such that $\mathcal{M}^{p}_{\text{lim}}(\mathcal{D}) = \mathcal{M}^{p+\Delta p',[t-p,t]}(\mathcal{D}) = \mathcal{M}^{p+\Delta p,[t-p,t]}(\mathcal{D})$ for all $\Delta p' \geq \Delta p$. Thus, since the ts-DMAG $\mathcal{M}^{p+\Delta p}(\mathcal{D})$ has repeating orientations and past-repeating adjacencies, also $\mathcal{M}^{p}_{\text{lim}}(\mathcal{D}) = \mathcal{M}^{p+\Delta p,[t-p,t]}(\mathcal{D})$ has both these properties.

To complete the proof, we need to show that $\mathcal{M}^{p}_{\text{lim}}(\mathcal{D})$ has repeating adjacencies. To this end, assume the opposite. Since $\mathcal{M}^{p}_{\text{lim}}(\mathcal{D})$ has past-repeating adjacencies, this assumption means in $\mathcal{M}^{p+\Delta p}(\mathcal{D})$ there is an edge $(i, t_i - \Delta t){*\!-\!*}(j, t_j - \Delta t)$ with $t - p \leq t_i, t_j \leq t$ and $\Delta t > 0$ such that (i, t_i) and (j, t_j) are nonadjacent in $\mathcal{M}^{p+\Delta p}(\mathcal{D})$. That (i, t_i) and (j, t_j) are nonadjacent in $\mathcal{M}^{p+\Delta p}(\mathcal{D})$ shows the existence of $\mathbf{S} \subseteq \mathbf{O}(t-p-\Delta p, t) \setminus \{(i, t_i), (j, t_j)\}$ with $(i, t_i) \perp\!\!\!\perp (j, t_j) \mid \mathbf{S}$ in $\mathcal{D}$. Due to the repeating separating sets property of $\mathcal{D}$, then $(i, t_i - \Delta t) \perp\!\!\!\perp (j, t_j - \Delta t) \mid \mathbf{S}_{-\Delta t}$ where $\mathbf{S}_{-\Delta t}$ is obtained by shifting all vertices in $\mathbf{S}$ backward in time by Δt steps. The vertices $(i, t_i - \Delta t)$ and $(j, t_j - \Delta t)$ are thus nonadjacent in $\mathcal{M}^{p+\Delta p+\Delta t}(\mathcal{D})$ and hence also nonadjacent in $\mathcal{M}^{p}_{\text{lim}}(\mathcal{D})$. This observation is in contradiction to the equality $\mathcal{M}^{p}_{\text{lim}}(\mathcal{D}) = \mathcal{M}^{p+\Delta p,[t-p,t]}(\mathcal{D})$.

2. Let (i, t_i) and (j, t_j) with $\tau = t_j - t_i \geq 0$ be distinct nonadjacent vertices in $\mathcal{M}^{p}_{\text{st}}(\mathcal{D})$. Then, the vertices $(i, t-\tau)$ and (j, t) are nonadjacent in $\mathcal{M}^{p}_{\text{st}}(\mathcal{D})$ due to the repeating edges property of $\mathcal{M}^{p}_{\text{st}}(\mathcal{D})$ and thus, using Lemma 4.7, also nonadjacent in $\mathcal{M}^{p}(\mathcal{D})$. Hence, there is $\mathbf{S} \subseteq \mathbf{O}(t-p,t) \setminus \{(i,t-\tau),(j,t)\}$ such that $(i,t-\tau) \perp\!\!\!\perp (j,t) \mid \mathbf{S}$ in $\mathcal{D}$. Due to the repeating separating sets property of $\mathcal{D}$, we thus get that (i,t_i) and (j,t_j) are nonadjacent in $\mathcal{M}^{p+(t-t_j)}(\mathcal{D})$ and hence also nonadjacent in $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$. Consequently, the skeleton of $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$ is a subgraph of the skeleton of $\mathcal{M}^p_{\text{st}}(\mathcal{D})$.

Next, let $(i,t_i){*\!\!-\!\!*}(j,t_j)$ be an edge in $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$. Then, since the skeleton of $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$ is a subgraph of the skeleton of $\mathcal{M}^p_{\text{st}}(\mathcal{D})$, the vertices (i,t_i) and (j,t_j) are also adjacent in $\mathcal{M}^p_{\text{st}}(\mathcal{D})$. Note that, since $\mathcal{M}^p_{\text{lim}}(\mathcal{D}) = \mathcal{M}^{\tilde{p},[t-p,t]}(\mathcal{D})$ for some $\tilde{p} > p$ according to part 1 of Lemma B.16, the orientation of $(i,t_i){*\!\!-\!\!*}(j,t_j)$ in $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$ conveys an ancestral relationships according to $\mathcal{D}$. Since also in $\mathcal{M}^p_{\text{st}}(\mathcal{D})$ the orientations of edges convey ancestral relationships according to $\mathcal{D}$, we finally get that the edge $(i,t_i){*\!\!-\!\!*}(j,t_j)$ in $\mathcal{M}^p_{\text{st}}(\mathcal{D})$ has the same orientation as in $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$.

3. Take $n \geq 0$ such that $\mathcal{M}^p_{\text{lim}}(\mathcal{D}) = \mathcal{M}^{p+n,[t-p,t]}(\mathcal{D})$, which exists according to part 1 of Lemma B.16. The statement now follows by applying Lemma F.3 with $\mathcal{M} \mapsto \mathcal{M}^{p+n}(\mathcal{D})$ and $\mathbf{O}$ the set of observable vertices within $[t-p,t]$.

4. The limiting ts-DPAG $\mathcal{P}^p_{\text{lim}}(\mathcal{D})$ has repeating adjacencies because according to part 3 of Lemma B.18 it is a DPAG for the limiting ts-DMAG $\mathcal{M}^p_{\text{lim}}(\mathcal{D})$, which according to part 1 of Lemma B.18 has repeating adjacencies. Let Δp be such that $\mathcal{P}^p_{\text{lim}}(\mathcal{D}) = \mathcal{P}^{p+\Delta p,[t-p,t]}(\mathcal{D})$, which exists according to part 2 of Lemma B.16. Because $\mathcal{P}^{p+\Delta p}(\mathcal{D})$ has repeating orientations according to Lemma E.1, also $\mathcal{P}^p_{\text{lim}}(\mathcal{D})$ has repeating orientations. Now use part 1 of Lemma 4.3.

5. Part 1 of Lemma B.16 gives the existence of an integer $n \geq 0$ such that $\mathcal{M}^p_{\text{lim}}(\mathcal{D}) = \mathcal{M}^{p+n',[t-p,t]}(\mathcal{D})$ for all $n' \geq n$, and part 2 of the same lemma gives the existence of an integer $m \geq 0$ such that $\mathcal{P}^p_{\text{lim}}(\mathcal{D}) = \mathcal{P}^{p+m',[t-p,t]}(\mathcal{D})$ for all $m' \geq m$. Thus, $\mathcal{M}^p_{\text{lim}}(\mathcal{D}) = \mathcal{M}^{p+k,[t-p,t]}(\mathcal{D})$ and $\mathcal{P}^p_{\text{lim}}(\mathcal{D}) = \mathcal{P}^{p+k,[t-p,t]}(\mathcal{D})$ for $k = \max(n,m)$. The statement now follows because $\mathcal{P}^{p+k}(\mathcal{D})$ is a DPAG for $\mathcal{M}^{p+k}(\mathcal{D})$. □

REFERENCES

GERHARDUS, A. and RUNGE, J. (2020). High-recall causal discovery for autocorrelated time series with latent confounders. In *Advances in Neural Information Processing Systems* (H. LAROCHELLE, M. RANZATO, R. HADSELL, M. F. BALCAN and H. LIN, eds.) **33** 12615–12625. Curran Associates, Red Hook, NY.

MALINSKY, D. and SPIRTES, P. (2018). Causal structure learning from multivariate time series in settings with unmeasured confounding. In *Proceedings of 2018 ACM SIGKDD Workshop on Causal Disocvery* (T. D. LE, K. ZHANG, E. KICIMAN, A. HYVÄRINEN and L. LIU, eds.). *Proceedings of Machine Learning Research* **92** 23–47. PMLR.

PEARL, J. (2009). *Causality: Models, Reasoning, and Inference*, 2nd ed. Cambridge Univ. Press, Cambridge. MR2548166 https://doi.org/10.1017/CBO9780511803161

RICHARDSON, T. and SPIRTES, P. (2002). Ancestral graph Markov models. *Ann. Statist.* **30** 962–1030. MR1926166 https://doi.org/10.1214/aos/1031689015

SPIRTES, P., GLYMOUR, C. and SCHEINES, R. (2000). *Causation, Prediction, and Search*, 2nd ed. *Adaptive Computation and Machine Learning*. MIT Press, Cambridge, MA. (With additional material by David Heckerman, Christopher Meek, Gregory F. Cooper, and Thomas Richardson.) MR1815675

SPIRTES, P. and RICHARDSON, T. S. (1997). A polynomial time algorithm for determining DAG equivalence in the presence of latent variables and selection bias. In *Proceedings of the Sixth International Workshop on Artificial Intelligence and Statistics* (D. MADIGAN and P. SMYTH, eds.). *Proceedings of Machine Learning Research* **R1** 489–500. PMLR. Reissued by PMLR on 30 March 2021.

VERMA, T. and PEARL, J. (1990). Equivalence and synthesis of causal models. In *Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence (UAI1990)* (P. BONISSONE, M. HENRION, L. KANAL and J. LEMMER, eds.) 220–227. AUAI Press.

ZHANG, J. (2008a). Causal reasoning with ancestral graphs. *J. Mach. Learn. Res.* **9** 1437-1474. MR2426048

ZHANG, J. (2008b). On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias. *Artif. Intell.* **172** 1873-1896. MR2459793 https://doi.org/10.1016/j.artint.2008.08.001