A Supplementary Material

Here we present the missing proofs for the lemmas in "Finding minimal d-separators in linear time and applications".

Proof of Lemma 4.1. Every collider $C \in pAn(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z})$ can be replaced by a possible directed path π_C from C to a node $V \in \mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z}$ and its reverse from V to C. π_C contains no node of $\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z}$ besides V. We can assume π_C does not contain any undirected edge, since there is an edge pointing towards the collider C. In an AG the configuration $B \to C - D$ is forbidden. In an RCG $B \to C - D$ can only occur, if there is an edge $B \to D$, so all undirected edges can be removed from π_C.

After every collider C has been replaced by π_C, the walk is of almost definite status, since only directed edges pointing away from it are added. After truncating the walk to start at its last node in $\mathbf{X}$ and end at the next node in $\mathbf{Y}$, all colliders are in $\mathbf{Z}$. □

Proof Lemma 4.2. For paths in AGs this is Lemma 3.13 in [18]. For RCGs let π be a walk with a node $V \notin pAn(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z})$. Assume V is the first such node and U the preceding node. Node V is not a possible ancestor of U, so they are connected by an edge $U \to V$.

All later edges on π point away from U: There is no undirected edge like $U \to V -$ since π is of almost definite status. There is no collider like $U \to V \leftarrow$, since V is not an (possible) ancestor of $\mathbf{Z}$.

Hence π ends with $U \to V \to \ldots \to Y$ and V is an ancestor of Y. □

Proof of Lemma 4.4. If there is an $Z \in \mathbf{Z} \setminus \mathbf{I}$ such that $\mathbf{Z} \setminus Z$ is a separator, $\mathbf{Z}$ is clearly not minimal. In the other direction, we have $\mathbf{Z} \setminus Z$ is not a separator for every $Z \in \mathbf{Z} \setminus \mathbf{I}$, so for each $Z \in \mathbf{Z} \setminus \mathbf{I}$ there is a definite status path $\pi_Z : \mathbf{X} \overset{+}{\sim} \mathbf{Y}$ that is not blocked by $\mathbf{Z} \setminus Z$, i.e., every non-collider is not in $\mathbf{Z} \setminus Z$ and every collider is in $An(\mathbf{Z} \setminus Z)$. Assume $\mathbf{Z}$ is not minimal, so there is a separator $\mathbf{Z}' \subset \mathbf{Z}$ with $\mathbf{I} \subseteq \mathbf{Z}'$.

Let $Z \in \mathbf{Z} \setminus \mathbf{Z}'$. No non-collider of π_Z is in $\mathbf{Z}' \subseteq \mathbf{Z} \setminus Z$ and every collider is in $An(\mathbf{Z} \setminus Z) \subseteq pAn(\mathbf{X} \cup \mathbf{Y} \cup (\mathbf{Z} \setminus Z)) = pAn(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{I}) = pAn(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z}')$, so $\mathbf{Z}'$ is not a separator due to Lemma 4.1. Hence $\mathbf{Z}$ is minimal. □

Proof of Proposition 6.5. $\mathbf{Z}_2 \subseteq \mathbf{Z}$ is a valid adjustment set, since it is a separator in $\mathcal{G}^{pbd}_{\mathbf{XY}}$ and as subset of $\mathbf{Z}$ contains no forbidden nodes of $De(PCP(\mathbf{X}, \mathbf{Y}))$.

Let π be a path between $Z_1 \in \mathbf{Z}_1 \setminus \mathbf{Z}_2$ and Y not blocked by $\mathbf{Z}_2 \cup X$ in $\mathcal{G}$. π also exists as unblocked path in $\mathcal{G}^{pbd}_{\mathbf{XY}}$, because it contains no edge $X \to$ and no edge on a path to a collider opening node in $\mathbf{Z}$ is removed in $\mathcal{G}^{pbd}_{\mathbf{XY}}$ (otherwise that node would be a descendant of $PCP(\mathbf{X}, \mathbf{Y})$ and thus be a forbidden node). Since $\mathbf{Z}_1 \subseteq \mathbf{Z} \subseteq An(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{I})$, π only contains nodes of $An(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{I})$ in $\mathcal{G}$ and $\mathcal{G}^{pbd}_{\mathbf{XY}}$. Let $N \in \mathbf{Z}$ be the last non-collider on π in $\mathbf{Z}$. If no such N exists, Z_1 is reachable from Y and $Z_1 \in \mathbf{Z}_2$, so π does not exist. N is reachable from Y, so $N \in \mathbf{Z}_2$ and π is blocked at N by $\mathbf{Z}_2$.

Let π be the shortest path between $\mathbf{X}$ and $Z_2 \in \mathbf{Z}_2 \setminus \mathbf{Z}_1$ not blocked by $\mathbf{Z}_1$ in $\mathcal{G}$. Again π only contains nodes of $An(\mathbf{X} \cup \mathbf{Y} \cup \mathbf{I})$. π exists as unblocked path in $\mathcal{G}^{pbd}_{\mathbf{XY}}$, because if it would contain an edge $X \to$ the node Z_2 or a collider opening node in $\mathbf{Z}_1$ would be a forbidden node $De(PCP(\mathbf{X}, \mathbf{Y}))$. There is a path π_Y from Z_2 to Y in $\mathcal{G}^{pbd}_{\mathbf{XY}}$ that contains no non-collider in $\mathbf{Z}$, or $\mathbf{Z}_2$ would not contain Z_2, so $\mathbf{Z}_1$ does not block π_Y. So unless Z_2 is a collider on the walk $\pi\pi_Y$, $\pi\pi_Y$ is not blocked by $\mathbf{Z}_1$ in $\mathcal{G}^{pbd}_{\mathbf{XY}}$, so $\mathbf{Z}_1$ is not a valid adjustment set. If Z_2 is a collider, either $\pi\pi_Y$ is not blocked by $\mathbf{Z}_2$ or a non-collider Z_2' of π is in $\mathbf{Z}_2$. Then $\pi[\mathbf{X} \overset{*}{\sim} Z_2']$ is shorter than π.

Thus, from Lemma 6.4 it follows that $\mathbf{Z}_2$ ensures a lower or equal asymptotic variance than $\mathbf{Z}_1$. □

Proof of Proposition 6.7. With the call to REACHABLE algorithm FINDNEARESTSEP computes a set $\mathbf{Z}'' \subseteq \mathbf{A} = An(Y \cup Z)$ in a DAG on which every node $W \in \mathbf{Z}''$ is reachable from Y by a walk only containing nodes in $\mathbf{A}$ and every non-collider is not in $\mathbf{Z}' = \mathbf{R} \cap (\mathbf{A} \setminus (Y \cup Z))$. Y and Z can only occur as end-nodes on the walk, so for every $W \in \mathbf{Z}''$ there exists a walk π_W that is active given $\mathbf{Z}'' \setminus W$ and contains no observable non-collider. Let $\mathbf{W}$ be the final set returned by FINDNEARESTSEP$(\mathcal{G}, Y, Z, \emptyset, \mathbf{R})$.

If $\mathbf{W}$ is not a nearest separator, there is a $X \in An(Y \cup Z) \setminus \{Y, Z\}$, a path π^m between X and Z intersecting $\mathbf{W}$, and another d-separator $\mathbf{W}' \subseteq \mathbf{R} \setminus \{Y, Z\}$ separating Y and Z that does not contain a node of π^m. $\mathbf{W}'$ also does not contain X, the start node of π^m. We can assume $X \in \mathbf{W}$ w.l.o.g. since π^m intersects $\mathbf{W}$. The path π^m in the moral graph can skip colliders, let π be the corresponding path in the DAG that includes the skipped colliders in $An(Y \cup Z)$.

Without an observable non-collider on π_X, the combined walk $\pi_X\pi : Y \overset{+}{\sim} Z$ contains no node of $\mathbf{W}'$ as non-collider. All colliders are in $An(Y \cup Z)$, so due to Lemma 4.1 $\pi_X\pi$ is active given $\mathbf{W}'$ and $\mathbf{W}'$ is not a d-separator. □