Appendix

A.1 Violation of "no self-masking missingness"

![Diagram of self-masking missingness indicator with multiple direct causes: TD-PC produces an extra edge between X and Y, but such self-masking missingness does not affect the other edges in the causal skeleton results, such as the edge between X and Vj ∈ V \ {X, Y}.](image)

In this section, we discuss challenges of the Self-masking Missingness (SFM), and its influences on MVPC.

We note that in the linear Gaussian cases SFM does not affect MVPC, when the SFM indicator R_t only has one direct cause X, such as in Figure 6. In this case, the result of the CI test of X and Y in test-wise deleted data implies the correct d-separation relation in the m-graph. With the faithfulness assumption on the m-graph, we have X ⊥⊥ Y ⇐⇒ X ⊥⊥ Y | R_t; furthermore, under the faithful observability assumption, we have X ⊥⊥ Y | R_t ⇐⇒ X* ⊥⊥ Y | R_t = 0 and X* ⊥⊥ Y | R_z = 0 is what we test in the test-wise deleted data of X and Y.

SFM affects MVPC results when the SFM indicator R_t has multiple direct causes. For example, as the m-graph in Figure 6 shown, conditioning on the missingness indicator which is the direct common effect of two variables in a CI test produces an extraneous edge between them in the result given by MVPC. Removing such extraneous edges is challenging, because our correction methods are not applicable to the self-masking missingness scenario. However, such self-masking missingness indicator does not affect the other edges between X and variables in V \ {X, Y} in the causal skeleton resulted by MVPC. Therefore, we specify in the output that edges between the self-masking variable and other direct causes of the self-masking missingness indicator are uncertain.

A.2 Proofs of the propositions

**Proof.** Proposition [1]

X ⊥⊥ Y | {Z, R_x = 0, R_y = 0, R_z = 0} ⇒ X ⊥⊥ Y | Z. We have X ⊥⊥ Y | {Z, R_x = 0, R_y = 0, R_z = 0}, where some of the involved missingness indicators may only take value 0 (i.e., the corresponding variables do not have missing values). With the faithful observability assumption, the above condition implies X ⊥⊥ Y | {Z, R_x, R_y, R_z}. Because of the faithfulness assumption on m-graphs, we know that X and Y are d-separated by {Z, R_x, R_y, R_z}; furthermore, with Assumption [2] and [3] the missingness indicators can only be leaf nodes in the m-graph. Therefore, conditioning on these nodes will not destroy the above d-separation relation. That is, in the m-graph, X and Y are d-separated by Z. Hence, we have X ⊥⊥ Y | Z.

**Proof.** Proposition [2]

The condition of Proposition [2] implies that for nodes X, Y and any node set Z ⊆ V \ {X, Y} in a m-graph, conditioning on Z and missingness indicators R_x, R_y, and R_z, there always exists an undirected path U between X and Y that is not blocked. Furthermore, to satisfy such constraint of U, at least a missingness indicator R_i ∈ {R_x, R_y, R_z} satisfies either one of the following two conditions: (1) R_i is the only vertex on U; (2) A cause of R_i is the only vertex on U as a collider. In Condition (1), if R_i is on U, it is a collider because under Assumption [4], missingness indicators are the leaf nodes in m-graphs. Then, suppose that R_i is not the only vertex on U, and that another node V_j ∈ V \ {X, Y, Z} is also on U. Conditioning on V_j and R_i, U is blocked, which is not satisfied the constraint of U. Thus, R_i should be the only vertex on U. The same reason also applies to Condition (2). In summary, we conclude that under the condition of Proposition [2] there is at least one missingness indicator R_i ∈ {R_x, R_y, R_z} such that R_i is the direct common effect or a descendant of the direct common effect of X and Y.

A.3 Detection of direct causes of missingness indicators

In Step 2 of Algorithm [4] detecting direct causes of missingness indicators is implemented by the causal skeleton discovery procedure of TD-PC. For each missingness indicator R_i, the causal skeleton discovery procedure checks all the CI relations between R_i and variables in V \ V_i, and tests whether R_i is conditionally independent of a variable V_j ∈ V \ V_i given any variable or set of variables connected to R_i or V_j. If they are conditionally independent, the edge between R_i and V_j is removed. Under Assumptions [1] and no extra edge is produced by the causal skeleton discovery procedure because according to Proposition [4] an extraneous edge only occurs when R_i and V_j have at least one direct common effect. Therefore, all the variables adjacent to R_i are its direct causes because R_i is either an effect or a cause, and we assume that R_i cannot be a cause in Assumption [1].