Appendix

In what follows, we include complete and formal mathematical proofs for the theorems presented in the paper, as well as more information about our experimental setup for the semi-synthetic experiment performed to answer RQ3.

A Proof of Theorem 2: Identifiability in Symmetric Additive Noise Models

This section provides a proof for identifiabiltiy of single-variable causal effects in symmetric additive noise models (**Theorem 2**).

Proof. Our causal estimand is the effect of intervening on X_i. For notational convenience, we assume $k = i$, i.e. the intervention is on the last treatment. From the observational data regime, we can trivially obtain the joint $\mathsf{P}(\boldsymbol{C}, X_1, \ldots, X_k, Y)$. As such, we can condition on the covariates and remaining treatment variables and then marginalise to obtain $\mathbb{E}[Y|\text{do}(X_k = x_k), C = c]$. We can rewrite our causal estimand as follows:

$$\begin{gathered}\mathbb{E}[Y|C = c, X_1 = x_1, \ldots, X_{k-1} = x_{k-1}, \text{do}(X_k = x_k)] = \\ f_Y(c, x_1, \ldots, x_k) + \mathbb{E}[U_Y|C = c, X_1 = x_1, \ldots, X_{k-1} = x_{k-1}].\end{gathered} \tag{7}$$

From the joint interventional data regime, we have access to the following expectation:

$$\mathbb{E}[Y|C = c, \text{do}(X_1 = x_1, \ldots, X_k = x_k)] = f_Y(c, x_1, \ldots, x_k). \tag{8}$$

Subtracting Eq. 8 from Eq. 7 shows that we only need to provide identifiability for the conditional expectation on the outcome noise, given the remaining treatment variables and the observed covariates:

$$\begin{aligned}\mathbb{E}[U_Y|C =&c, X_1 = x_1, \ldots, X_{k-1} = x_{k-1}] = \Sigma_{u_y}\Sigma_{u_x}^{-1}\boldsymbol{u}_x, \\ \text{where } \Sigma_{u_y} &= \begin{bmatrix}\sigma_{Y1} & \ldots & \sigma_{Y(k-1)}\end{bmatrix}, \\ \Sigma_{u_x} &= \begin{bmatrix}\sigma_1^2 & \ldots & \sigma_{1(k-1)} \\ \vdots & \ddots & \vdots \\ \sigma_{(k-1)1} & \ldots & \sigma_{k-1}^2\end{bmatrix}, \\ \text{and } \boldsymbol{u}_x &= \begin{bmatrix}x_1 - f_1(c) & \ldots & x_{k-1} - f_{k-1}(c)\end{bmatrix}^\intercal.\end{aligned} \tag{9}$$

Here, σ_{ij} denotes the covariance between noise variables U_i and U_j, and σ_{Yi} denotes the covariance between the outcome U_Y and U_i. There are two types of unidentified factors in these expressions: the **structural equations**, that is, the f_i's encapsulated in $\boldsymbol{u}_x$ from Eq. 9, and the parameters of the **noise distribution**, which are encapsulated in Σ_{u_x} and Σ_{u_y}. We tackle these in what follows:

Identifying the structural equations. As a direct result of our model definition, we obtain $\mathbb{E}[X_i|C = c] = f_i(c)$ from the observational data regime. This follows because in the DAG of Figure 1, C and U_i are independent for all i. As such, the structural equations (and thus $\boldsymbol{u}_x$ from Eq. 9) are identifiable.

Identifying the noise distribution. For any pair of treatment variables, we can obtain $\mathbb{E}[X_i|C = c, X_j = x_j] = f_i(c) + \mathbb{E}[U_i|X_j = x_j]$. The latter term in this expression can then be rewritten as $\mathbb{E}[U_i|U_j = x_j - f_j(c)]$, for fixed values of x_j and c. Because we have shown the structural equations (f_i, f_j) to be identifiable, we have that $\mathbb{E}[U_i|U_j]$ is identifiable, which gives us the covariance σ_{ij}. Hence, the entirety of Σ_{u_x} is identifiable. The same procedure can be used for every entry of Σ_{u_y}, and the covariances σ_{Yi} are identifiable as a result.

It follows naturally that $\mathbb{E}[Y|C = c, X_1 = x_1, \ldots, X_{k-1} = x_{k-1}, \text{do}(X_k = x_k)]$ is identifiable. As we have data from the observational regime, any marginalisation of this query is identifiable as well, concluding the proof. □

B Unidentifiability under Unconstrained SCMs

Tables B hold the full distributions for the counterexample in Section 3.1, showing that single-variable interventional effects are unidentifiable from observational and joint interventional data for unconstrained SCMs.

Table 4: Distributions under both SCMs $\mathcal{M}$ and $\mathcal{M}'$.

(a) Observational joint distribution.

$\mathsf{P}(X_1, X_2, Y)$	$Y=0$	$Y=1$
$X_1, X_2 = 0, 0$	$1-p$	0
$X_1, X_2 = 0, 1$	0	0
$X_1, X_2 = 1, 0$	0	0
$X_1, X_2 = 1, 1$	0	p

(b) Joint interventional distribution.

$\mathsf{P}(Y \mid \text{do}(X_1, X_2))$	$Y=0$	$Y=1$
$\text{do}(X_1 = 0, X_2 = 0)$	1	0
$\text{do}(X_1 = 0, X_2 = 1)$	1	0
$\text{do}(X_1 = 1, X_2 = 0)$	1	0
$\text{do}(X_1 = 1, X_2 = 1)$	$1-p$	p

(c) Interventional distribution on X_2.

$\mathsf{P}(Y, X_1 \mid \text{do}(X_2))$		$Y=0$	$Y=1$
$\text{do}(X_2 = 0)$	$X_1 = 0$	$1-p$	0
	$X_1 = 1$	p	0
$\text{do}(X_2 = 1)$	$X_1 = 0$	$1-p$	0
	$X_1 = 1$	0	p

Table 5: SCMs for C, where $(U_C, U_1, U_2, U_Y)_{\mathcal{M}} \sim \mathcal{N}(0, \Sigma)$ and $(U_C, U_1, U_2, U_Y)_{\mathcal{M}'} \sim \mathcal{N}(0, \Sigma')$.

$\mathcal{M}$	$\mathcal{M}'$
$C = U_C$	$C = U_C$
$X_1 = C + U_1$	$X_1 = 2C + U_1$
$X_2 = C + U_2$	$X_2 = 2C + U_2$
$Y = CX_1X_2 + U_Y$	$Y = CX_1X_2 + U_Y$

$$\text{where } \Sigma = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \Sigma' = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

C Unidentifiability when C is not independent of U

Table 5 provides the counterexample mentioned in Section 3.3 (just after Equation 3) showing that single-variable interventional effects are unidentifiable from observational and joint interventional data when the covariates C are not independent of the unobserved confounders U in the ANMs we consider.

Indeed, in model $\mathcal{M}$, C is correlated with the latent U_1, U_2, telling us this model is non-Markovian: that there is an unobserved confounder correlating C and U_1, U_2. C is thus not independent of this unobserved confounder. This is not the case in model $\mathcal{M}'$ Moreover, we can see that $\mathcal{M}$ and $\mathcal{M}'$ are identical under interchange of X_1 and X_2, and have the same joint and marginal distributions, as well as the same joint interventional distribution.

As the two SCMs coincide on observational and joint interventional distributions, this proves that the confounding distribution is not identifiable from these data regimes alone, and single-variable interventions as well as the full SCM are unidentifiable. This demonstrates the need for the additional restrictions between C and U in Theorem 3.2.

D Proof of Theorem 1: Identifiability with Causally Dependent Treatments

This section provides a proof for identifiability of single-variable causal effects when there is a causal dependency among treatments (**Theorem 1**). Observational and joint interventional data are not sufficient in this case to identify causal effects on *all* treatments – but we can identify the causal effect of intervening on the *consequence* treatment instead of the *causing* treatment.

Proof. Our causal estimand is the effect of intervening on X_j. We can rewrite our causal estimand as follows:

$$\mathbb{E}[Y|X_i = x_i, \mathrm{do}(X_j = x_j), C = c] = f_Y(c, x_i, x_j) + \mathbb{E}[U_Y|C = c, X_i = x_i]. \tag{10}$$

From the joint interventional data regime, we have access to the following expectation:

$$\mathbb{E}[Y|C = c, \mathrm{do}(X_i = x_i, X_j = x_j)] = f_Y(c, x_i, x_j). \tag{11}$$

Subtracting Eq. 11 from Eq. 10 shows that we only need to provide identifiability for the conditional expectation on the outcome noise, given the observed value for treatment X_i and the observed covariates $\boldsymbol{C}$:

$$\mathbb{E}[U_Y|C = c, X_i = x_i] = \mathbb{E}[U_Y|U_i = x_i - f_i(c)] = \frac{\sigma_{Yi}}{\sigma_{ii}}(x_i - f_i(c)). \tag{12}$$

Here, the first step comes from our SCM definition, and the second step comes from the fact that we assume the noise distribution to be a zero-centered multivariate Gaussian. As such, we need to identify the function f_i, the variance on the noise variable U_i, and the covariance between U_i and U_Y. We obtain $\mathbb{E}[X_i|C = c] = f_i(c)$ directly from the observational data regime. This makes the noise variable $U_i = X_i - f_i(\boldsymbol{C})$ identifiable. Now, as a result, we can identify its variance σ_{ii} and covariance σ_{Yi} from the observational data regime, which concludes the proof. □

E Consistency under Varying Levels of Confounding (RQ3)

In order to validate the consistency of our learning method under varying levels of unobserved confounding (**RQ3**), we adopt data from the International Stroke Trial database [Carolei, 1997]. We partially follow the semi-synthetic setup laid out in Appendix 3 of [Zhang and Bareinboim, 2021]. Specifically, we adopt their probability table for the joint observational distribution of the covariates: the gender, age and conscious state of a patient. This table was computed from the dataset to reflect a real-world observational distribution. They deal with discrete treatments—which we extend to the continuous case where treatments can be interpreted as varying dosages of aspirin and heparin. We model the structural equation on the outcome as:

$$\begin{aligned} Y = \ & 0.1S - 0.1A + 0.25(C-1) + \alpha_a + 0.75\alpha_h - 3SA - 0.1S\alpha_a \\ & -0.3A\alpha_a + 0.1S\alpha_h + 0.2A\alpha_h + 0.3C\alpha_h - 0.45\alpha_a\alpha_h. \end{aligned} \tag{13}$$

This loosely reflects the same intuitions as laid out in [Zhang and Bareinboim, 2021], where S is the gender, A the age, C the conscious state, α_a the aspirin dose and α_h the heparin dose. Note that this does not reflect correct medical knowledge or insights—the goal is merely to have a polynomial with second-order interactions that we can learn to model. As described in the main text, we add zero-mean Gaussian noise to the treatments and the outcome, where we randomly generate positive semi-definite covariance matrices with bounded non-diagonal entries—varying the limit on the size of the covariances in order to assess the effect of varying confounding on our method. We repeat this process 5 times, and sample 512 observational and 512 joint-interventional samples (both α_a and α_h) to learn the SCM from. We evaluate our learning methods on $5\,000$ evaluation samples to predict the outcome under a single-variable intervention on the aspirin dose α_a.

All experiments ran in a notebook on a laptop.