NEURAL PROXIMAL CAUSAL INFERENCE FOR LONGITUDINAL DATA WITH UNMEASURED CONFOUNDING

Anonymous authors
Paper under double-blind review

ABSTRACT

Unmeasured confounders pose a fundamental challenge in estimating causal effects of time-varying treatments from observational data. Proximal Causal Inference (PCI) addresses this through proxy variables, but estimation under PCI requires solving recursively coupled ill-posed integral equations. Existing longitudinal methods make this tractable only through parametric models that cannot capture nonlinear trajectory dynamics. We introduce Neural-PCI, a deep learning estimation framework grounded in established proxy-based identification and, to our knowledge, the first implementable neural method for the full T-stage recursion. A joint Maximum Moment Restriction (MMR) formulation enforces all T moment conditions through a single closed-form kernel loss, and yields three semi-parametric estimators with complementary robustness properties. We show that stage-wise sequential estimation incurs error growth exponential in the time horizon T, while our joint formulation gives additive error aggregation that scales linearly in T on average. Experiments on synthetic data and the semi-synthetic MIMIC-III benchmark show that Neural-PCI consistently outperforms parametric PCI, latent variable, and sequential-randomization baselines across horizons.

1 INTRODUCTION

Estimating the causal effect of time-varying treatments from observational data is a central problem in medicine, economics, and policy evaluation. The fundamental obstacle is confounding: factors that affect both treatment assignment and outcome, thereby introducing a spurious association between them. In many real-world settings, treatment assignment depends on factors the analyst cannot measure, such as latent disease severity or clinician judgment, and these factors also influence the outcome. Standard adjustment methods are biased as a result. The longitudinal setting makes this even harder: unmeasured confounders at each step affect both the current treatment and every later observation, and their effects accumulate along the trajectory.

The deep learning literature on longitudinal causal effect estimation splits into two lines, distinguished by how each treats unmeasured confounding. The first line assumes it away. Increasingly expressive sequence architectures, including recurrent, adversarial, and transformer-based models (Lim et al., 2018; Bica et al., 2020a; Li et al., 2021; Melnychuk et al., 2022; Frauen et al., 2023), estimate causal effects under sequential randomization, which requires that conditioning on observed history removes all confounding. When unmeasured confounders exist, these methods are biased regardless of architectural sophistication. The second line addresses unmeasured confounding directly by modeling it as a latent variable (Bica et al., 2020b; Kuzmanovic et al., 2021; Hatt & Feuerriegel, 2024; Cheng et al., 2024; Sun et al., 2025). The validity of these methods rests on assumptions about the latent mechanism that are strong, untestable, and tied to a specific data-generating process.

Proximal causal inference (PCI) (Miao et al., 2018; Tchetgen Tchetgen et al., 2024) takes a different route. Rather than modeling U directly, PCI exploits proxy variables: observed quantities that carry information about U through the treatment and outcome pathways. The covariates at each time step decompose into three roles: treatment-inducing proxies Z, outcome-inducing proxies W, and measured common causes X. Standard adjustment handles X. The two proxy families address U: because Z and W descend from U through separate pathways, combining them lets the analyst recover U's influence without observing U. Two *bridge functions* implement this strategy: each is defined on one proxy and anchored (bridged) to the other through a moment condition (Figure 1),

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

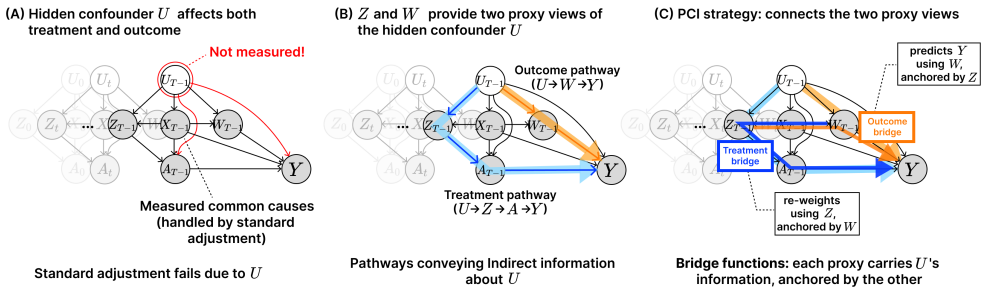

Figure 1: Proximal causal inference (PCI) at a glance, shown at a single time step for clarity. The full longitudinal graph (shaded behind each panel) repeats this structure at $t = 0, \ldots, T-1$. White circles denote unmeasured variables; gray circles denote measured variables. Measured common causes X are conditioned on directly throughout and are not the focus of the proxy machinery. The proxies Z and W descend from U through separate pathways. PCI exploits this structure: the bridge functions in Panel C use one proxy to anchor a function defined on the other, conveying information about U through the proxy structure without observing U.

conveying U's influence through the proxy structure. Estimating these bridge functions from data is the central computational problem in PCI.

Existing longitudinal PCI methods (Tchetgen Tchetgen et al., 2024; Ying et al., 2023) solve this estimation problem only with parametric models. Each bridge function is defined implicitly through a moment condition, an ill-posed integral equation whose direct solution amplifies noise. In the longitudinal setting, there are T such equations, recursively coupled across time steps. Parametric forms make each per-stage inversion tractable; sequential plug-in handles the coupling. Each restriction has its own limitation: parametric forms cap expressive power, and sequential plug-in compounds errors exponentially in T. The obvious neural extension keeps sequential plug-in in place and only replaces parametric forms with neural networks. This inherits the second limitation and amplifies it. Recent nested-moment theory covers the two-period case (Meza & Singh, 2025); no method yet handles the general-T recursion with the flexibility real-world trajectories require.

Neural Proximal Causal Inference (Neural-PCI) fills this gap. The Maximum Moment Restriction (MMR) framework (Mastouri et al., 2021; Kompa et al., 2022) converts each bridge moment condition into a tractable kernel-based loss, addressing the per-stage ill-posedness. We cast the T MMR losses as a joint objective that enforces all moment conditions simultaneously, addressing the recursive coupling. Specialized architecture for each bridge family realizes this joint formulation in practice.

Contributions. Our contributions are as follows:

- (i) We introduce Neural-PCI, a practical deep learning estimation framework for longitudinal causal inference grounded in established proxy-based identification with time-varying treatments. To our knowledge, it is the first implementable neural framework for the full T-stage recursion of longitudinal bridge functions.
- (ii) We resolve the recursive coupling of the T per-stage moment conditions by aggregating their MMR losses into a single kernel-based objective. The objective is a closed-form V-statistic requiring no inner adversarial maximization, and it is realized through an architecture that structurally encodes the temporal recursion of both bridge families.
- (iii) We characterize how the estimation error of longitudinal bridge functions scales with the time horizon T: sequential estimation incurs error growth exponential in T, while our joint formulation yields additive aggregation that scales linearly in T on average. We further derive a finite-sample bound on ATE error that establishes consistency of the three estimators. An ablation that holds architecture and loss form fixed isolates the scaling difference empirically.
- (iv) On synthetic datasets and the semi-synthetic MIMIC-III benchmark, Neural-PCI outperforms baselines from all three categories, with the margin widening as T grows, matching the theoretical prediction.

The remainder of the paper is organized as follows. Section 2 reviews proximal identification and motivates the joint formulation; Section 3 presents Neural-PCI; Section 4 states the error-scaling guarantee; Section 5 reports experiments; Section 6 concludes.

2 Proximal Identification and Estimation Barrier

We observe n i.i.d. trajectories over T time steps. Each trajectory consists of a binary treatment sequence $\bar{A}_{T-1} = (A_0, \ldots, A_{T-1})$, observed covariates $\bar{L}_{T-1} = (L_0, \ldots, L_{T-1})$, and a terminal outcome Y. Following the proximal framework (Figure 1B), the covariates at each step decompose into three roles: treatment-inducing proxies $\bar{Z}_{T-1}$, outcome-inducing proxies $\bar{W}_{T-1}$, and measured common causes $\bar{X}_{T-1}$. Let $Y^{\bar{a}_{T-1}}$ denote the potential outcome under intervention sequence $\bar{a}_{T-1}$. Our target is the expected potential outcome

$$\theta^a = \mathbb{E}[Y^{\bar{a}_{T-1}}],$$

from which average treatment effects follow by contrasting two intervention sequences.

The proximal framework replaces the sequential randomization assumption with transparent conditions on the proxy variables (Appendix A). These conditions are more interpretable and easier to reason about on subject-matter grounds (Tchetgen Tchetgen et al., 2024): a practitioner can ask "does this variable carry information about the unmeasured confounder?" rather than "have I measured every relevant confounder?"

2.1 Bridge Functions and Identification

The two bridge function families shown in Figure 1C are defined implicitly through recursive conditional moment equations. These equations encode the unmeasured confounder's influence through the proxy variables, making identification possible without observing U. Proofs of the identification results below are in Appendix A.

Outcome bridge functions $\{H_t\}_{t=0}^{T-1}$, with $H_t = h_t(\bar{W}_t, \bar{a}_{T-1}, \bar{X}_t)$, implement the backward construction:

$$\mathbb{E}[H_t \mid \bar{A}_t, \bar{Z}_t, \bar{X}_t] = \mathbb{E}[H_{t+1} \mid \bar{A}_t, \bar{Z}_t, \bar{X}_t], \quad t = T{-}1, \ldots, 0, \tag{1}$$

initialized with $H_T = Y$. Each H_t is a function of $(\bar{W}_t, \bar{a}_{T-1}, \bar{X}_t)$; $\bar{Z}_t$ appears only in the conditioning set. This separation is what allows the proxies to de-confound the estimate: the moment condition enforces that H_t absorbs the unmeasured confounder's influence through the outcome pathway. Under the proximal assumptions,

$$\theta^a = \mathbb{E}[Y^{\bar{a}_{T-1}}] = \mathbb{E}[H_0]. \tag{2}$$

Treatment bridge functions $\{Q_t\}_{t=0}^{T-1}$, with $Q_t = q_t(\bar{Z}_t, \bar{a}_t, \bar{X}_t)$, implement the forward construction:

$$\mathbb{E}[Q_t \mid \bar{A}_t, \bar{W}_t, \bar{X}_t] = \frac{\mathbb{E}[Q_{t-1} \mid \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t]}{P(A_t = a_t \mid \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t)}, \quad t = 0, \ldots, T{-}1, \tag{3}$$

initialized with $Q_{-1} = 1$. Each Q_t, a function of $(\bar{Z}_t, \bar{a}_t, \bar{X}_t)$, accumulates inverse-probability weights that, conditional on W, adjust for the confounder's effect on treatment assignment. Under the proximal assumptions and a positivity condition,

$$\theta^a = \mathbb{E}[Y^{\bar{a}_{T-1}}] = \mathbb{E}[Y \cdot \mathbf{1}(\bar{A}_{T-1} = \bar{a}_{T-1}) \cdot Q_{T-1}]. \tag{4}$$

The cumulative weight Q_{T-1} makes the treatment bridge the structurally harder family to estimate flexibly. We address this in Section 3.3.

From Identification to Estimation. Eq. (2) and Eq. (4) reduce identification of θ^a to estimating the two bridge function families: outcome bridge functions $\{H_t\}_{t=0}^{T-1}$ and treatment bridge functions $\{Q_t\}_{t=0}^{T-1}$. Once estimated, these yield three population-level semi-parametric estimators for θ^a with complementary robustness properties:

$$\text{P-OR:} \quad \theta^a = \mathbb{E}[H_0], \tag{5}$$

$$\text{P-IPW:} \quad \theta^a = \mathbb{E}\left[Y \cdot \mathbf{1}(\bar{A} = \bar{a}) \cdot Q_{T-1}\right], \tag{6}$$

$$\text{P-DR:} \quad \theta^a = \mathbb{E}\left[H_0 + \sum_{t=0}^{T-1} \mathbf{1}(\bar{A}_t = \bar{a}_t) \cdot Q_t \cdot (H_{t+1} - H_t)\right]. \tag{7}$$

P-OR is consistent when $\{H_t\}_{t=0}^{T-1}$ is correctly specified, and P-IPW when $\{Q_t\}_{t=0}^{T-1}$ is. P-DR combines both bridges and is *doubly robust*: it remains consistent when either bridge family is correctly specified, even if the other is mis-specified. In practice, agreement between P-OR and P-IPW serves as a diagnostic for the quality of both bridge function estimates.

2.2 The Estimation Barrier

Estimation of bridge functions presents two coupled difficulties. First, each moment condition in Eq. (1) and Eq. (3) defines its bridge function implicitly as the solution to a Fredholm integral equation in the conditional expectation operator. These equations are ill-posed: small data errors produce disproportionately large errors in the recovered function. Second, the T conditions within each family are recursively coupled across time steps: Eq. (1) specifies H_t in terms of H_{t+1} and Eq. (3) specifies Q_t in terms of Q_{t-1}. The T inversions cannot be handled independently.

Existing longitudinal PCI methods (Tchetgen Tchetgen et al., 2024; Ying et al., 2023) handle these difficulties by stacking two restrictions. Parametric models for each bridge sidestep the per-stage ill-posedness, reducing each equation to a finite-dimensional regression. Sequential plug-in across stages handles the coupling: solve stage $(T-1)$ first, treat its estimate as a fixed pseudo-outcome for stage $(T-2)$, and so on along the recursion (the treatment bridge runs in the opposite direction).

The obvious neural extension keeps sequential plug-in in place and only replaces the parametric forms with neural networks. This inherits the compounding problem and amplifies it, since each ill-posed equation now propagates a noisy nonparametric estimate carried forward from the previous stage. Section 3 resolves the dilemma by reformulating the T coupled problems as a single joint optimization objective, addressing both restrictions simultaneously.

3 Neural-PCI: Joint MMR Formulation and Architecture

The joint MMR objective is realized through a shared recurrent architecture with two mechanisms adapted from Frauen et al. (2023): a dual-phase forward pass and gradient blocking between adjacent time steps. Neural-PCI contributes the MMR-based losses that enforce the joint formulation and specializes this pattern to the two bridge function families.

3.1 Joint MMR Objective

The Maximum Moment Restriction (MMR) framework (Mastouri et al., 2021; Kompa et al., 2022) produces a tractable kernel-based loss without solving the ill-posed integral equation directly (Appendix C). We apply it to the two bridge moment conditions and aggregate across time.

Moment functions for the two bridge families. Each bridge moment condition maps directly to an MMR problem. For the outcome bridge, the moment function at time t is $f_t^H(\eta) = (H_{t+1} - H_t)$ conditioned on $(\bar{A}_t, \bar{Z}_t, \bar{X}_t)$ where $H_T = Y$, so that $\mathbb{E}[f_t^H \mid \bar{A}_t, \bar{Z}_t, \bar{X}_t] = 0$ recovers Eq. (1). For the treatment bridge, a propensity-free reformulation of Eq. (3) yields the three-component moment function $f_t^Q(\eta) = (Q_t, -Q_{t-1}, -Q_{t-1})^\top$ conditioned on $(\bar{A}_t, \bar{W}_t, \bar{X}_t)$ where $Q_{-1} = 1$ (derived in Section 3.3). Each time step of each bridge family yields one MMR problem.

Aggregating the T moment conditions. The natural way to avoid stage-wise plug-in is to collapse all T per-step MMR risks into a single loss:

$$\mathcal{L}(\eta) = \frac{1}{T} \sum_{t=0}^{T-1} \mathcal{R}(f_t(\eta)) + \lambda\, \Omega[\eta], \tag{8}$$

where η denotes the parameters of the bridge function families, Ω is a regularizer, and f_t represents f_t^H or f_t^Q with the corresponding kernel. Because each risk $\mathcal{R}$ has a closed form in the kernel (Appendix C), Eq. (8) is an ordinary minimization problem for any horizon T. Eq. (8) is necessary but not sufficient: independently parameterized $\{H_t\}$ and $\{Q_t\}$ would lose the directional structure of the recursion. Section 3.2 and 3.3 show how the architecture encodes this recursion for both bridge families.

3.2 Outcome Bridge Network

Dual-phase forward pass. We realize the recursive coupling of Eq. (1) through a dual-phase forward pass, which produces per-stage bridge estimates of the iterative-regression algorithm of Ying et al. (2023). The outcome bridge functions $H_t = h_t(\bar{W}_t, \bar{a}_{T-1}, \bar{X}_t)$ are estimated through a shared LSTM encoder that processes the sequence $(\bar{W}_t, \bar{X}_t)$ under two treatment scenarios in parallel: once with the observed sequence $\bar{A}_t$ (factual phase) and once with the target intervention $\bar{a}_t$ (counterfactual phase). At each time step t, feedforward heads map the hidden states to the factual estimate $\hat{H}_t^A$ and the counterfactual estimate $\hat{H}_t^a$.

Gradient blocking along the temporal recursion. The moment condition Eq. (1) defines H_t recursively in terms of H_{t+1}. The predictor H_t should match the conditional expectation of the target H_{t+1}, not the reverse. We implement this asymmetry by gradient blocking $\hat{H}_{t+1}^a$ when it appears as the target for $\hat{H}_t^A$ in the moment function: $\hat{H}_{t+1}^a$ is anchored as a fixed target at each gradient step, and $\hat{H}_t^A$ is forced to move toward it. This preserves the backward propagation direction of Eq. (1). The initialization $\hat{H}_T^a := Y$ anchors the recursion in the observed outcome.

Loss. Applying Eq. (8) to the outcome bridge moment function yields a V-statistic over the n trajectories, indexed by $i, j \in \{1, \ldots, n\}$. The kernel $k_{ij,t} = k\big((\bar{A}_{i,t}, \bar{Z}_{i,t}, \bar{X}_{i,t}), (\bar{A}_{j,t}, \bar{Z}_{j,t}, \bar{X}_{j,t})\big)$ is computed using the Recurrent Neural Tangent Kernel (Alemohammad et al., 2021) to handle sequences. The empirical loss is:

$$\mathcal{L}_{\text{H-MMR}}(\eta) = \frac{1}{n^2} \sum_{i,j=1}^{n} \frac{1}{T} \sum_{t=0}^{T-1} \Big(\hat{H}_{i,t}^A(\eta) - \hat{H}_{i,t+1}^a(\eta)\Big) \Big(\hat{H}_{j,t}^A(\eta) - \hat{H}_{j,t+1}^a(\eta)\Big) k_{ij,t}. \tag{9}$$

The MMR loss $\mathcal{L}_{\text{H-MMR}}$ enforces the moment condition Eq. (1) at the population level. Early in training, however, the parameter space consistent with $\mathcal{L}_{\text{H-MMR}} \approx 0$ is large. We add a soft-hinge regularizer to restrict early training to the smaller subset where residuals $(\hat{H}_t^A - \hat{H}_{t+1}^a)$ are also small:

$$\mathcal{L}_{\text{H}}(\eta) = \mathcal{L}_{\text{H-MMR}}(\eta) + \lambda_{\text{mse}}\, \tilde{\mathcal{L}}_{\text{H-MSE}}(\eta), \tag{10}$$

with $\tilde{\mathcal{L}}_{\text{H-MSE}}(\eta)$ a soft hinge with sharpness β_{hinge} and margin τ_H on average squared residual:

$$\mathcal{L}_{\text{H-MSE}}(\eta) = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=0}^{T-1} \Big(\hat{H}_{i,t}^A(\eta) - \hat{H}_{i,t+1}^a(\eta)\Big)^2. \tag{11}$$

The hinge restricts early optimization and deactivates once residuals fall below τ_H, so that MMR alone carries the final identification signal.

3.3 Treatment Bridge Network

The treatment bridge presents two structural obstacles absent from the outcome bridge. First, the propensity in the denominator of Eq. (3) embeds a nested estimation problem. Second, the cumulative weight Q_{T-1} produces scale growth across T stages that destabilizes joint training. We resolve the first obstacle through a propensity-free reformulation, with its induced kernel-discrimination issue addressed by a decomposed kernel. We resolve the second through RMS normalization.

The treatment bridge functions $Q_t = q_t(\bar{Z}_t, \bar{a}_t, \bar{X}_t)$ are estimated through a single-path LSTM encoder that processes $(\bar{Z}_t, \bar{X}_t)$ under the observed treatment sequence. The P-IPW estimator Eq. (6) selects matched trajectories through the indicator $\mathbf{1}(\bar{A}_{T-1} = \bar{a}_{T-1})$, so counterfactual selection is handled at estimation time rather than through a parallel forward pass. Following the same target-versus-predictor principle as the outcome bridge, we realize the forward propagation direction of Eq. (3) from $Q_{-1} = 1$ through gradient blocking on Q_{t-1}, the fixed target for Q_t.

Propensity-free moment condition. Applying MMR directly to Eq. (3) would require dividing by conditional propensity $\pi_t = P(A_t = a_t \mid \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t)$. We instead work with an equivalent form, obtained by multiplying both sides of Eq. (3) by π_t and expanding through the law of total expectation over $A_t \in \{0, 1\}$ (Ying et al., 2023); see Appendix D for the derivation:

$$\mathbb{E}\big[Q_t \cdot N_t\big(\bar{A}_t, \bar{W}_t, \bar{X}_t\big)\big] = \mathbb{E}\big[Q_{t-1} \cdot \big\{N_t(A_t = 1, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t) + N_t(A_t = 0, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t)\big\}\big], \tag{12}$$

where N_t ranges over test functions in the RKHS unit ball. Eq. (12) corresponds to the three-component moment function $f_t^Q = (Q_t, -Q_{t-1}, -Q_{t-1})^\top$ conditioned on $(\bar{A}_t, \bar{W}_t, \bar{X}_t)$, with kernel arguments N_t evaluated at treatment variants $\{A_t, 1, 0\}$.

Decomposed kernel. The vector-valued moment function in Eq. (12) requires a 3×3 kernel matrix $K_{ij,t}$ that discriminates simultaneously along covariate similarity (between sample pairs) and treatment-history similarity (between variants $\{A_t, 1, 0\}$). A monolithic kernel over $(\bar{A}_t, \bar{W}_t, \bar{X}_t)$ collapses cross-treatment discrimination because high-dimensional proxy histories dominate. We therefore construct $K_{ij,t}$ as the product of a covariate factor and a treatment factor:

$$[K_{ij,t}]_{\ell,m} = K^{\text{cov}}_{ij,t} \cdot \left(1 + \gamma \cdot K^{\text{trt}}\left(\bar{A}^{(\ell)}_{0:t}, \bar{A}^{(m)}_{0:t}\right)\right), \quad \ell, m \in \{A_t, 1, 0\}, \tag{13}$$

where $\bar{A}^{(\ell)}_{0:t}$ denotes the treatment history with the final entry set to ℓ. The covariate factor $K^{\text{cov}}_{ij,t} = k\big((\bar{W}_{i,t}, \bar{X}_{i,t}), (\bar{W}_{j,t}, \bar{X}_{j,t})\big)$ measures similarity in proxy histories, deliberately excluding treatment. The treatment factor K^{trt} is a Hamming RBF on binary sequences, providing discrimination between the three treatment variants. This additive-boost form preserves the covariate baseline: a purely multiplicative $K^{\text{cov}} \cdot K^{\text{trt}}$ would zero out the covariate signal whenever treatment histories diverge. The scalar $\gamma > 0$ balances the covariate and treatment axes. The product structure preserves positive definiteness.

Loss. Applying Eq. (8) to the vector-valued moment function yields the V-statistic:

$$\mathcal{L}_{\text{Q-MMR}}(\eta) = \frac{1}{n^2 T} \sum_{i,j=1}^{n} \sum_{t=0}^{T-1} \tilde{f}^Q_{i,t}(\eta)^\top K_{ij,t}\, \tilde{f}^Q_{j,t}(\eta), \tag{14}$$

where $\tilde{f}^Q = f^Q / s_t$ is self-normalized by the sample-level RMS $s_t = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Q_t^2 + Q_{t-1}^2\right)}$ at adjacent time steps, detached from the gradient graph. This normalization addresses the cumulative-weight scale growth: without it, the V-statistic inherits the $O(Q^2)$ scale at the moment condition's solution, which destabilizes joint training across T.

As with the outcome bridge, we add a soft-hinge regularizer derived from the propensity relation $Q_{t-1}/Q_t \approx \pi_t$, which we enforce through a pre-trained propensity network $\hat{\pi}$:

$$\mathcal{L}_{\text{Q-prop}}(\eta) = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=0}^{T-1} \left(\frac{Q_{t-1}(\eta)}{Q_t(\eta)} - \hat{\pi}_{i,t}\right)^2. \tag{15}$$

The treatment bridge is trained on

$$\mathcal{L}_{\text{Q}}(\eta) = \mathcal{L}_{\text{Q-MMR}}(\eta) + \lambda_{\text{prop}}\, \tilde{\mathcal{L}}_{\text{Q-prop}}(\eta), \tag{16}$$

with $\tilde{\mathcal{L}}_{\text{Q-prop}}$ obtained by wrapping Eq. (15) in a soft hinge analogous to Eq. (11).

3.4 Semi-parametric Estimators

Once both networks are trained, their outputs yield three empirical semi-parametric estimators for θ^a. P-OR uses only the outcome bridge, $\hat{\theta}^a_{\text{POR}} = \frac{1}{n}\sum_{i=1}^{n} \hat{H}^a_{0,i}$.

P-IPW uses only the treatment bridge, restricted to trajectories whose observed sequence matches the target intervention through the indicator $\mathbf{1}(\bar{A}_i = \bar{a})$. The cumulative weight Q_{T-1} in Eq. (17) inherits the second obstacle's scale growth at the estimation stage. We apply Hájek self-normalization to stabilize the estimator:

$$\hat{\theta}^a_{\text{PIPW}} = \frac{\sum_{i=1}^{n} Y_i \cdot \mathbf{1}(\bar{A}_i = \bar{a}) \cdot \hat{Q}_{T-1,i}}{\sum_{i=1}^{n} \mathbf{1}(\bar{A}_i = \bar{a}) \cdot \hat{Q}_{T-1,i}}. \tag{17}$$

P-DR combines both bridges through a sequential targeting procedure adapted from Frauen et al. (2023), replacing DeepACE's inverse-propensity clever covariate with the estimated treatment bridge Q_t. The targeting layer learns a scalar ϵ that perturbs the counterfactual outcome bridge outputs, using Q_t as the correction signal. For each time step t, the targeted output $\tilde{H}^a_t$ is:

$$\tilde{H}^a_t = \hat{H}^a_t + \epsilon \cdot \xi_t, \quad \xi_t = \xi_{t+1} - \mathbf{1}(\bar{A}_t = \bar{a}_t) \cdot Q_t, \tag{18}$$

with $\xi_T = 0$ and ϵ initialized at zero, so the estimator begins as pure P-OR. The parameter ϵ is learned jointly with the bridge network parameters through the targeting loss, which is algebraically equivalent to the augmented form of Eq. (7) of the P-DR estimator (Appendix E.1):

$$\mathcal{L}_{\text{tar}}(\eta, \epsilon) = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=0}^{T-1} \left(\tilde{H}^a_{i,t}(\eta, \epsilon) - \tilde{H}^a_{i,t+1}(\eta, \epsilon) \right)^2 . \tag{19}$$

Both bridge networks are trained jointly under

$$\mathcal{L}_{\text{P-DR}}(\eta, \epsilon) = \mathcal{L}_{\text{H}}(\eta) + \lambda_Q \, \mathcal{L}_{\text{Q}}(\eta) + \lambda_{\text{tar}} \, \mathcal{L}_{\text{tar}}(\eta, \epsilon), \tag{20}$$

where λ_Q and λ_{tar} control the relative weights. The P-DR estimate is $\hat{\theta}^a_{\text{PDR}} = \frac{1}{n} \sum_{i=1}^{n} \tilde{H}^a_{0,i}$. Double robustness of P-DR follows in Appendix E.

4 Theoretical Guarantees

Section 3 resolves the recursive coupling architecturally. Two findings formalize what that resolution achieves: the sequential-versus-joint distinction is not a quantitative improvement but a qualitatively different scaling regime. Formal statements, assumptions, and proofs are in Appendix B, F, and G.

Notation. h^*_t is the true bridge function, $\hat{h}_t$ its joint MMR estimate, and $\hat{h}^{\text{seq}}_t$ the sequential variant. We write $\tilde{h}_t$ for the best approximation the hypothesis class can express. The two findings use different but compatible metrics: $\| \cdot \|_{L^2}$ in Finding 1 and $d^2_{k_t}$ in Finding 2, where k_t is the kernel at time step t. Both measure distances from h^*_t. The per-stage amplification factor κ quantifies how much each sequential inversion magnifies upstream error, and r_n is the per-stage estimation rate (typically $n^{-1/2}$).

Finding 1: Sequential estimation is structurally unscalable in T. Solving the T coupled moment conditions one stage at a time forces each ill-posed equation to amplify the errors carried forward from the previous stage. These amplifications compound multiplicatively, and the overall error at $t = 0$ grows exponentially in the horizon:

$$\|\hat{h}^{\text{seq}}_0 - h^*_0\|_{L^2} = O(\kappa^{T-1} \cdot r_n). \tag{21}$$

The exponential dependence on T is structural. It persists regardless of how small r_n is made, and the same mechanism applies to the treatment bridge family propagating forward. The exponential fit in Figure 2(a) yields $\hat{\kappa} \approx 1.12$ in our synthetic DGP, modestly above unity but sufficient to drive multiplicative compounding.

Finding 2: Joint MMR replaces exponential compounding with additive aggregation. Solving all T moment conditions simultaneously removes the per-stage error amplification that Finding 1 depends on. The architectural mechanisms of Section 3 encode the recursion directly, and the average per-timestep moment-condition discrepancy then satisfies, with probability $1 - \delta$,

$$\frac{1}{T} \sum_{t=0}^{T-1} d^2_{k_t}(\eta^*, \hat{\eta}) \;\leq\; \underbrace{\frac{1}{T} \sum_{t} d^2_{k_t}(\eta^*, \tilde{\eta})}_{\text{approximation, } \to 0 \text{ as capacity grows}} \;+\; \underbrace{\tilde{O}(n^{-1/2})}_{\text{estimation, } \to 0 \text{ as } n \to \infty} . \tag{22}$$

$d^2_{k_t}(\eta^*, \eta)$ restates the population MMR risk $\mathcal{R}_{k_t}(\eta)$ as a distance between bridges parametrized by η and by the true η^*; it couples consecutive bridges rather than measuring per-stage L^2 error. Errors aggregate additively rather than multiplicatively, and the κ^{T-1} factor of Finding 1 is absent. Converting this discrepancy to the L^2 error of Finding 1 introduces the ill-posedness constant τ once, as a uniform factor, rather than as a product over stages (Remark G.3). The worst-case individual-timestep error grows only polynomially ($T/\sqrt{n}$); Corollary G.4 gives the resulting ATE bound.

5 Experiments

We evaluate Neural-PCI on synthetic datasets and the MIMIC-III semi-synthetic benchmark. Because counterfactual outcomes are never observed, ground-truth ATE is available only in synthetic and semi-synthetic settings; we follow the standard protocol (Melnychuk et al., 2022; Frauen et al., 2023). The experiments test the theoretical scaling prediction of Theorems F.1 and G.1, compare against baselines from all three existing categories, and validate the core design choices.

Table 1: ATE estimation error measured by Mean Absolute Error (MAE; $|\hat{\psi} - \psi|$), reported as mean $\pm$ sd over 5 repeated runs. NA: not applicable (deterministic models produce a fixed estimate). HC: handles hidden confounders. Best result in **bold**, second-best underlined.

	Synthetic (T)					Semi-synthetic
Method (HC)	$T = 3$	$T = 5$	$T = 7$	$T = 10$	$T = 20$	MIMIC ($T = 12$)
Sequential randomization neural						
CT ($\times$)	0.063±0.026	0.146±0.155	0.329±0.150	0.189±0.120	1.617±0.204	0.068±0.061
DeepACE ($\times$)	0.073±0.021	0.035±0.022	0.089±0.072	0.148±0.071	0.304±0.145	0.078±0.058
DeepACE-tar ($\times$)	0.171±0.016	0.097±0.033	0.136±0.115	0.177±0.124	0.097±0.069	0.180±0.208
Latent variable models						
DTA (✓)	0.496±0.031	0.934±0.049	1.499±0.035	0.961±0.036	1.552±0.049	0.536±0.102
TIFM (✓)	0.337±0.015	0.077±0.014	0.232±0.013	0.217±0.036	0.463±0.004	1.976±0.412
Parametric longitudinal PCI						
MSMM P-OR (✓)	0.048±NA	0.057±NA	0.133±NA	0.076±NA	2.276±NA	0.112±NA
MSMM P-IPW (✓)	0.046±NA	0.396±NA	0.050±NA	0.303±NA	0.390±NA	0.099±NA
MSMM P-DR (✓)	0.054±NA	0.083±NA	0.174±NA	0.152±NA	1.507±NA	0.249±NA
Neural-PCI (ours)						
Neural P-OR (✓)	**0.012±0.005**	0.047±0.038	0.043±0.028	0.057±0.029	0.074±0.047	0.052±0.063
Neural P-IPW (✓)	0.054±0.002	0.047±0.008	0.055±0.011	**0.028±0.010**	**0.022±0.020**	**0.025±0.019**
Neural P-DR (✓)	0.017±0.008	**0.013±0.011**	**0.039±0.030**	0.057±0.026	0.042±0.046	0.043±0.034

5.1 Experimental Setup

Synthetic data. We simulate trajectories following Figure 1 with an AR(1) unmeasured confounder U_t ($\phi_u = 0.7$). Covariates $Z_t \in \mathbb{R}^3$, $W_t \in \mathbb{R}^2$, and $X_t \in \mathbb{R}^5$ (including Y_{t-1}) are linear functions of U_t, their own lags, and lagged treatments with additive Gaussian noise. The binary treatment A_t is assigned via a logistic model with persistence $\alpha_a = 2.5$. We generate $n = 1{,}000$ trajectories for $T \in \{3, 5, 7, 10, 20\}$. The ground-truth ATE $\psi = \mathbb{E}[Y^{\bar{1}_T}] - \mathbb{E}[Y^{\bar{0}_T}]$ is computed by Monte Carlo simulation of counterfactual trajectories. Full DGP is given in Appendix H.

MIMIC-III semi-synthetic data. Following standard practice (Frauen et al., 2023), we construct a semi-synthetic benchmark on $n = 1{,}000$ ICU patient trajectories at 4-hour resolution over $T = 12$ (48 hours total), using the MIMIC-Extract pipeline (Wang et al., 2020). Proxy roles are assigned in two stages: a data-driven MCMC partition followed by clinical designation of surrogate U. Treatments and outcomes are simulated through nonlinear structural equations applied to the real covariates, so that ground-truth ATE is available. Full details are in Appendix I.

Baselines. We compare against baselines from three categories. Parametric longitudinal PCI: MSMM-POR, MSMM-PIPW, MSMM-PDR (Ying et al., 2023), which share our identification strategy but restrict bridges to parametric models. Latent variable models: DTA (Kuzmanovic et al., 2021) and TIFM (Cheng et al., 2024). Sequential-randomization neural: DeepACE, its targeting layer variant DeepACE-tar (Frauen et al., 2023), and Causal Transformer (CT) (Melnychuk et al., 2022). Architectural hyperparameters of all neural methods are tuned via Optuna (Akiba et al., 2019) on factual validation loss and Neural-PCI's causal-regularization weights by the P-OR/P-IPW agreement diagnostic, which uses no counterfactual ground truth; full search spaces and training details are in Appendix J.

5.2 Main Results

Table 1 reports ATE estimation errors across all settings. Neural-PCI achieves the best performance at every horizon and on MIMIC-III. Within Neural-PCI, P-OR wins at $T = 3$, the Hájek-normalized P-IPW wins at $T = 10, 20$, and on MIMIC, and P-DR achieves the best result at $T = 5, 7$ while placing first or second in every column. P-DR's consistent top-two presence provides empirical evidence of the double robustness property: it inherits accuracy whenever either P-OR or P-IPW is well-fit, and never loses to both simultaneously.

The baselines exhibit three distinct failure modes, aligned with their identification status. Parametric longitudinal PCI models are competitive at $T = 3$ but degrade catastrophically as complexity grows, with MSMM P-OR reaching MAE 2.276 at $T = 20$, the compounding of sequential plug-in that Theorem F.1 predicts. Sequential-randomization neural baselines occasionally match Neural-PCI

on an individual run ($T = 5$), but they exhibit high variance across seeds. Latent variable models show the opposite pattern: low variance but large, persistent bias. The proximal DGP violates their generative assumptions, because every covariate descends from U_t and no pure instrument is available for the latent variable models to exploit.

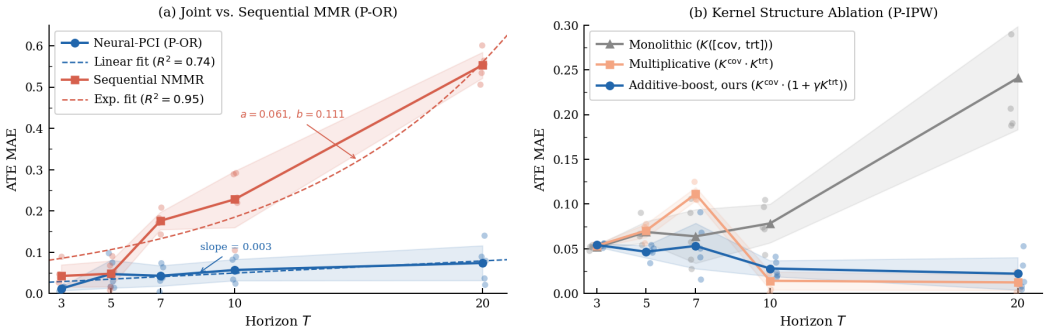

Figure 2: ATE MAE vs. horizon T on the simulation DGP ($n = 5$ runs per T; scatter: individual runs, lines: per-T means, bands: ± 1 std). **(a)** Sequential NMMR (red) compounds exponentially ($a \cdot e^{bT}$ fit), consistent with Theorem F.1; the linear scaling of Joint MMR (blue) matches Theorem G.1. **(b)** Kernel ablation under identical P-IPW training: the decomposed additive-boost kernel is the only variant that avoids catastrophic failure across the horizon range.

5.3 Ablation Study

Joint vs. sequential estimation (Figure 2a). To test the central theoretical prediction, we construct Sequential NMMR. This ablation baseline uses the same architecture and loss function as Neural-PCI (P-OR), but trains the bridge functions one stage at a time, with each fitted $\hat{H}_{t+1}$ frozen as the pseudo-outcome for stage t. Sequential NMMR's error grows exponentially in T (fitted $a \cdot e^{bT}$ with $b = 0.111$, $R^2 = 0.95$), giving an empirical per-stage amplification factor $\hat{\kappa} = e^b \approx 1.12$. This is consistent with the $O(\kappa^{T-1})$ scaling of Theorem F.1. Neural-PCI scales linearly (slope 0.003, $R^2 = 0.74$). The per-stage amplification is modest, but it compounds into an order-of-magnitude error inflation: the gap is negligible at $T = 3$ and widens dramatically by $T = 20$.

Kernel structure (Figure 2b). We compare three kernel variants under identical Neural-PCI (P-IPW) training: monolithic, purely multiplicative, and additive-boost (our design). The monolithic kernel degrades as T grows. The shared covariate history drowns out single-step treatment variation, and cross-treatment discrimination collapses at $T \geq 10$. The multiplicative kernel instead fails at $T = 7$: Hamming distances at this horizon are large enough to trigger aggressive exponential decay of cross-treatment entries, but their distinct values are too sparse to produce a smoothly graduated kernel. Additive-boost retains the K^{cov} baseline as a suppression floor and adds treatment-sensitive modulation on top. It is the only variant without a catastrophic failure regime.

6 Conclusion

We introduced Neural-PCI, a deep learning estimation framework for longitudinal causal inference grounded in proxy-based identification and, to our knowledge, the first implementable neural framework for the full T-stage recursion. It removes the two restrictions stacked by existing longitudinal PCI methods: the MMR framework converts each ill-posed moment condition into a tractable kernel-based loss, and the joint formulation replaces sequential plug-in with simultaneous estimation across all T stages. We prove that the resulting distinction is qualitative: sequential estimation incurs error growth exponential in T, while joint MMR bounds errors to linear scaling on average. This contrast is a property of longitudinal bridge estimation itself, not of any particular estimator; any future nonparametric longitudinal PCI estimator must confront it, and Neural-PCI provides the first deployable resolution for the general-T setting. Open directions include sensitivity analyses for proxy-role mis-specification, partial identification when the completeness conditions are uncertain, and extensions to continuous treatments and time-varying ATEs.

AI USE STATEMENT

Generative AI tools were used for editorial purposes only: grammar checking, sentence polishing, and rewording or condensing author-written text. They were not used for any task requiring disclosure under the ICLR AI Policy for Authors, including theory, proofs, experimental design, implementation, data generation, or interpretation of results. The authors reviewed all AI-assisted edits and take full responsibility for the final content of this work.

ETHICS STATEMENT

This is a methodological contribution with no human-subjects research or new data collection. Experiments use synthetic data (Appendix H) and the de-identified MIMIC-III database under credentialed PhysioNet access (Appendix I); MIMIC-III is not redistributed. Broader impacts and the untestable assumptions underlying any PCI-based estimate are discussed in Appendix L.

REPRODUCIBILITY STATEMENT

Assumptions and complete proofs for all theoretical results are given in Appendices A, B, F, and G. The synthetic data-generating process (Appendix H), the MIMIC-III benchmark construction (Appendix I), and all hyperparameters, search spaces, seeds, and compute resources (Appendix J, Table 3) are fully specified.

REFERENCES

Chunrong Ai and Xiaohong Chen. Efficient estimation of models with conditional moment restrictions containing unknown functions. *Econometrica*, 71(6):1795–1843, 2003.

Chunrong Ai and Xiaohong Chen. Estimation of possibly misspecified semiparametric conditional moment restriction models with different conditioning variables. *Journal of Econometrics*, 141 (1):5–43, 2007.

Chunrong Ai and Xiaohong Chen. The semiparametric efficiency bound for models of sequential moment restrictions containing unknown functions. *Journal of Econometrics*, 170(2):442–457, 2012.

Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019.

Sina Alemohammad, Zichao Wang, Randall Balestriero, and Richard Baraniuk. The recurrent neural tangent kernel. In *International Conference on Learning Representations*, 2021.

Andrew Bennett, Nathan Kallus, Xiaojie Mao, Whitney Newey, Vasilis Syrgkanis, and Masatoshi Uehara. Minimax instrumental variable regression and L_2 convergence guarantees without identification or closedness. In *Proceedings of the Conference on Learning Theory*, 2023.

Ioana Bica, Ahmed M Alaa, James Jordon, and Mihaela van der Schaar. Estimating counterfactual treatment outcomes over time through adversarially balanced representations. In *International Conference on Learning Representations*, 2020a.

Ioana Bica, Ahmed M. Alaa, and Mihaela Van Der Schaar. Time series deconfounder: estimating treatment effects over time in the presence of hidden confounders. In *Proceedings of the 37th International Conference on Machine Learning*, 2020b.

Debo Cheng, Ziqi Xu, Jiuyong Li, Lin Liu, Jixue Liu, Wentao Gao, and Thuc Duy Le. Instrumental variable estimation for causal inference in longitudinal data with time-dependent latent confounders. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024.

Yifan Cui, Hongming Pu, Xu Shi, Wang Miao, and Eric Tchetgen Tchetgen. Semiparametric proximal causal inference. *Journal of the American Statistical Association*, 119(546):1348–1359, 2024.

Nishanth Dikkala, Greg Lewis, Lester Mackey, and Vasilis Syrgkanis. Minimax estimation of conditional moment models. In *Advances in Neural Information Processing Systems*, 2020.

Dennis Frauen, Tobias Hatt, Valentyn Melnychuk, and Stefan Feuerriegel. Estimating average causal effects from patient trajectories. In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence*, 2023.

Tobias Hatt and Stefan Feuerriegel. Sequential deconfounding for causal inference with unobserved confounders. In *Proceedings of the Third Conference on Causal Learning and Reasoning*, 2024.

Benjamin Kompa, David R. Bellamy, Thomas Kolokotrones, James M. Robins, and Andrew L. Beam. Deep learning methods for proximal inference via maximum moment restriction. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, 2022.

Milan Kuzmanovic, Tobias Hatt, and Stefan Feuerriegel. Deconfounding temporal autoencoder: Estimating treatment effects over time using noisy proxies. In *Proceedings of Machine Learning for Health*, 2021.

Rui Li, Stephanie Hu, Mingyu Lu, Yuria Utsumi, Prithwish Chakraborty, Daby M. Sow, Piyush Madan, Jun Li, Mohamed Ghalwash, Zach Shahn, and Li-wei Lehman. G-net: a recurrent network approach to g-computation for counterfactual prediction under a dynamic treatment regime. In *Proceedings of Machine Learning for Health*, 2021.

Bryan Lim, Ahmed Alaa, and Mihaela van der Schaar. Forecasting treatment responses over time using recurrent marginal structural networks. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, 2018.

Marc Lipsitch, Eric J Tchetgen Tchetgen, and Ted Cohen. Negative controls: A tool for detecting confounding and bias in observational studies. *Epidemiology*, 21(3):383–388, 2010.

Afsaneh Mastouri, Yuchen Zhu, Limor Gultchin, Anna Korba, Ricardo Silva, Matt Kusner, Arthur Gretton, and Krikamol Muandet. Proximal causal learning with kernels: Two-stage estimation and moment restriction. In *Proceedings of the 38th International Conference on Machine Learning*, 2021.

Valentyn Melnychuk, Dennis Frauen, and Stefan Feuerriegel. Causal transformer for estimating counterfactual outcomes. In *Proceedings of the 39th International Conference on Machine Learning*, 2022.

Isaac Meza and Rahul Singh. Nested nonparametric instrumental variable regression, 2025.

Wang Miao, Zhi Geng, and Eric J Tchetgen Tchetgen. Identifying causal effects with proxy variables of an unmeasured confounder. *Biometrika*, 105(4):987–993, 08 2018.

Krikamol Muandet, Wittawat Jitkrittum, and Jonas Kübler. Kernel conditional moment test via maximum moment restriction. In *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI)*, 2020.

Whitney K. Newey and James L. Powell. Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578, 2003.

Zexu Sun, Bowei He, Shiqi Shen, Zhipeng Wang, Zhi Gong, Chen Ma, Qi Qi, and Xu Chen. Sequential causal effect estimation by jointly modeling the unmeasured confounders and instrumental variables. *IEEE Transactions on Knowledge and Data Engineering*, 37(2):910–922, 2025.

Eric J Tchetgen Tchetgen. The control outcome calibration approach for causal inference with unobserved confounding. *American Journal of Epidemiology*, 179(5):633–640, 03 2014.

Eric J Tchetgen Tchetgen, Andrew Ying, Yifan Cui, Xu Shi, and Wang Miao. An introduction to proximal causal inference. *Statistical Science*, 39(3):375–390, 08 2024.

Shirly Wang, Matthew B. A. McDermott, Geeticka Chauhan, Marzyeh Ghassemi, Michael C. Hughes, and Tristan Naumann. MIMIC-Extract: A data extraction, preprocessing, and representation pipeline for MIMIC-III. In *Proceedings of the ACM Conference on Health, Inference, and Learning (CHIL)*, 2020.

Andrew Ying, Wang Miao, Xu Shi, and Eric J Tchetgen Tchetgen. Proximal causal inference for complex longitudinal studies. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 85(3):684–704, 03 2023.

This appendix is organized as follows.

- Appendix A: proximal identification assumptions and proofs of identification theorems.
- Appendix B: regularity conditions used by the error-scaling theorems.
- Appendix C: general MMR framework.
- Appendix D: derivation of the propensity-free moment condition for P-IPW.
- Appendix E: double robustness of P-DR.
- Appendix F: sequential estimation error bound (exponential in T).
- Appendix G: joint MMR error bound (linear in T on average).
- Appendix H: synthetic data-generating process.
- Appendix I: MIMIC-III semi-synthetic benchmark.
- Appendix J: hyperparameters, training protocol, compute resources.
- Appendix K: related work.
- Appendix L: extended discussion of limitations and broader impacts.

A Proximal Identification: Assumptions and Proofs

We present the formal assumptions underlying the proximal identification results stated in Section 2, adapting the framework of Ying et al. (2023) to the general T-step setting. We use the notation established in Section 2.

A.1 Core Assumptions

Assumption A.1 (Consistency)**.** The observed outcome satisfies $Y = Y^{\bar{A}_{T-1}}$ almost surely, and the observed proxies satisfy $W_t = W_t^{\bar{A}_{t-1}}$ and $Z_t = Z_t^{\bar{A}_{t-1}}$ almost surely for all $t = 0, \dots, T-1$. That is, observed outcomes and proxies coincide with their potential counterparts under the treatment actually received.

Assumption A.2 (Positivity)**.** For all $t = 0, \dots, T-1$ and all $a_t \in \{0, 1\}$:

$$P(A_t = a_t \mid \bar{A}_{t-1} = \bar{a}_{t-1}, \bar{U}_t, \bar{L}_t) > 0 \quad \text{almost surely,}$$

where $\bar{L}_t = (\bar{Z}_t, \bar{W}_t, \bar{X}_t)$ denotes the observed covariate history up to time t.

Assumption A.3 (Sequential outcome-inducing proxy condition)**.** For all $t = 0, \dots, T-1$, the outcome-inducing proxies W_t satisfy:

$$W_t^{\bar{a}_{t-1}, \bar{z}_{t-1}} = W_t^{\bar{a}_{t-1}} \quad \text{for all } \bar{a}_{t-1}, \bar{z}_{t-1}, \text{ almost surely.}$$

This states that Z_t has no direct effect on W_t other than through the unmeasured confounders — W_t is an outcome-inducing proxy in the sense that it is associated with outcomes through U_t but not directly caused by the treatment-inducing proxies Z_t.

Assumption A.4 (Sequential treatment-inducing proxy condition)**.** For all $t = 0, \dots, T-1$, the treatment-inducing proxies Z_t satisfy:

$$Y^{\bar{a}_{T-1}, \bar{z}_{T-1}} = Y^{\bar{a}_{T-1}} \quad \text{for all } \bar{a}_{T-1}, \bar{z}_{T-1}, \text{ almost surely.}$$

This states that Z_t has no direct effect on the outcome Y other than through treatment — Z_t is a treatment-inducing proxy in the sense that it influences treatment but affects outcomes only through the unmeasured confounders U_t.

Assumption A.5 (Sequential proximal latent randomization)**.** For all $t = 0, \dots, T-1$:

$$(Z_t, A_t) \perp\!\!\!\perp \{W_s^{\bar{a}_{s-1}}, Y^{\bar{a}_{T-1}} : s \le t,\, \bar{a}_{T-1} \in \{0,1\}^T\} \mid \bar{A}_{t-1}, \bar{X}_t, U_t.$$

Given the observed treatment history, measured confounders, and unmeasured confounders at time t, the treatment-inducing proxies and treatment are jointly independent of all potential outcomes and potential outcome-inducing proxies. This assumption formalizes the randomization of (Z_t, A_t) conditional on the full confounder history.

Assumption A.6 (Sequential Markov property of U)**.** For all $t = 0, \dots, T-1$ and all $\bar{a}_{T-1}$:

$$\bar{U}_{t-1} \perp\!\!\!\perp \left\{ Y^{\bar{a}_{T-1}}, W_s, Z_s : 0 \le s \le t \right\} \mid U_t, \bar{A}_{t-1}, \bar{X}_t,$$

with $\bar{U}_{t-1} := (U_0, \dots, U_{t-1})$ and $\bar{U}_{-1} := \emptyset$. Current U_t summarizes the latent state for proximal identification: past confounders provide no additional information about proxies or potential outcomes once U_t and observable history are conditioned upon. The condition holds automatically when $\{U_t\}$ is Markov and structural equations depend on U only through its current value, as in our AR(1) synthetic DGP (Appendix H).

Assumptions A.3–A.5 together imply the following conditional independence conditions, which can alternatively be taken as primitive conditions:

$$\begin{aligned} Z_t &\perp\!\!\!\perp Y \mid \bar{A}_t, \bar{X}_t, U_t, \\ (Z_t, A_t) &\perp\!\!\!\perp W_t \mid \bar{A}_{t-1}, \bar{X}_t, U_t, \\ (Z_t, A_t) &\perp\!\!\!\perp W_s \mid \bar{A}_{t-1}, \bar{X}_t, U_t \quad \text{for all } s \le t. \end{aligned}$$

Assumption A.7 (Completeness for outcome bridge functions)**.** For all $t = 0, \dots, T-1$, all $\bar{a}_t$, and all square-integrable functions ν:

$$\begin{aligned} \mathbb{E}[\nu(U_t) \mid \bar{a}_t, \bar{Z}_t, \bar{X}_t] = 0 \iff \nu(U_t) = 0 \text{ almost surely}, \\ \mathbb{E}[\nu(Z_t) \mid \bar{a}_t, \bar{W}_t, \bar{X}_t] = 0 \iff \nu(Z_t) = 0 \text{ almost surely}. \end{aligned}$$

The first condition requires that Z_t is sufficiently informative about U_t through the treatment pathway; the second requires that W_t is sufficiently informative about Z_t through the outcome pathway. In the categorical case with cardinalities d_U, d_Z, d_W respectively, these reduce to $\min(d_Z, d_W) \ge d_U$.

Assumption A.8 (Completeness for treatment bridge functions)**.** For all $t = 0, \dots, T-1$, all $\bar{a}_t$, and all square-integrable functions ν:

$$\begin{aligned} \mathbb{E}[\nu(U_t) \mid \bar{a}_t, \bar{W}_t, \bar{X}_t] = 0 \iff \nu(U_t) = 0 \text{ almost surely}, \\ \mathbb{E}[\nu(W_t) \mid \bar{a}_t, \bar{Z}_t, \bar{X}_t] = 0 \iff \nu(W_t) = 0 \text{ almost surely}. \end{aligned}$$

The roles of Z_t and W_t are exchanged relative to Assumption A.7: here W_t must be sufficiently informative about U_t, and Z_t must be sufficiently informative about W_t.

A.2 Identification Theorems

Lemma A.9 (Existence of outcome bridge functions)**.** *Under Assumptions A.1–A.5 and A.7, and the regularity conditions of Appendix B, there exist functions $\{h_t\}_{t=0}^{T-1}$ with $H_t = h_t(\bar{W}_t, \bar{a}_{T-1}, \bar{X}_t)$ satisfying the recursive integral equations (1), initialized with $H_T = Y$.*

Lemma A.10 (Existence of treatment bridge functions)**.** *Under Assumptions A.1–A.5 and A.8, and the regularity conditions of Appendix B, there exist functions $\{q_t\}_{t=0}^{T-1}$ with $Q_t = q_t(\bar{Z}_t, \bar{a}_t, \bar{X}_t)$ satisfying the recursive integral equations (3), initialized with $Q_{-1} = 1$.*

The regularity conditions (Appendix B) are needed only for existence of bridge functions; the identification theorems below follow from the proxy assumptions alone, given existence.

Theorem A.11 (Proximal identification via outcome bridge functions)**.** *Under Assumptions A.1–A.5, A.6, and A.7, for any $\{h_t\}$ satisfying (1):*

$$\theta^a = \mathbb{E}[Y^{\bar{a}_{T-1}}] = \mathbb{E}[H_0].$$

Proof. We show by backward induction that, with $\bar{a}_{-1} := \emptyset$ and $\bar{U}_0 := U_0$,

$$\mathbb{E}[Y^{\bar{a}_{T-1}} \mid \bar{a}_{t-1}, \bar{U}_t, \bar{X}_t] = \mathbb{E}[H_t \mid \bar{a}_{t-1}, \bar{U}_t, \bar{X}_t] \quad \text{for } t = 0, \dots, T-1. \tag{$\star_t$}$$

Conditioning on the full trajectory $\bar{U}_t$ rather than U_t alone preserves the σ-algebra subset relations required by iterated expectations across time steps; Assumption A.6 reduces $\bar{U}_t$ to U_t at the points where Assumption A.7 requires the single-confounder form.

Base step ($t = T-1$). The moment condition (1) with $H_T = Y$ states

$$\mathbb{E}[Y \mid \bar{A}_{T-1}, \bar{Z}_{T-1}, \bar{X}_{T-1}] = \mathbb{E}[H_{T-1} \mid \bar{A}_{T-1}, \bar{Z}_{T-1}, \bar{X}_{T-1}]$$

on the event $\{\bar{A}_{T-1} = \bar{a}_{T-1}\}$. Replace Y by $Y^{\bar{a}_{T-1}}$ via Assumption A.1, then apply iterated expectations introducing $\bar{U}_{T-1}$ in the inner conditioning (the outer σ-algebra is contained in the inner). By Assumption A.5 applied iteratively in t to drop $\bar{Z}_{T-1}$ and $\bar{A}_{T-1}$, and Assumption A.6 to reduce $\bar{U}_{T-1}$ to U_{T-1} where Assumption A.7 requires the single-U form, the inner reduces and completeness yields

$$\mathbb{E}[Y^{\bar{a}_{T-1}} \mid \bar{a}_{T-2}, \bar{U}_{T-1}, \bar{X}_{T-1}] = \mathbb{E}[H_{T-1} \mid \bar{a}_{T-2}, \bar{U}_{T-1}, \bar{X}_{T-1}],$$

establishing $(\star_{T-1})$.

Inductive step. Suppose $(\star_{t+1})$ holds for some $t \in \{0, \dots, T-2\}$. Then

$$\begin{aligned}
\mathbb{E}[Y^{\bar{a}_{T-1}} \mid \bar{a}_{t-1}, \bar{U}_t, \bar{X}_t] &\overset{(1)}{=} \mathbb{E}\big[\mathbb{E}[Y^{\bar{a}_{T-1}} \mid \bar{a}_t, \bar{U}_{t+1}, \bar{X}_{t+1}] \,\big|\, \bar{a}_{t-1}, \bar{U}_t, \bar{X}_t\big] \\
&\overset{(2)}{=} \mathbb{E}\big[\mathbb{E}[H_{t+1} \mid \bar{a}_t, \bar{U}_{t+1}, \bar{X}_{t+1}] \,\big|\, \bar{a}_{t-1}, \bar{U}_t, \bar{X}_t\big] \\
&\overset{(3)}{=} \mathbb{E}[H_{t+1} \mid \bar{a}_{t-1}, \bar{U}_t, \bar{X}_t] \\
&\overset{(4)}{=} \mathbb{E}[H_t \mid \bar{a}_{t-1}, \bar{U}_t, \bar{X}_t],
\end{aligned}$$

where (1) applies iterated expectations conditioning on $(\bar{a}_t, \bar{U}_{t+1}, \bar{X}_{t+1})$. The outer σ-algebra is contained in the inner, since $\bar{a}_{t-1} \subseteq \bar{a}_t$, $\bar{U}_t \subseteq \bar{U}_{t+1}$, and $\bar{X}_t \subseteq \bar{X}_{t+1}$, and Assumption A.5 fixes $A_t = a_t$ inside. Step (2) applies the inductive hypothesis $(\star_{t+1})$. Step (3) collapses the iterated expectation. Step (4) applies the moment condition (1) at time t, then iterated expectations and Assumption A.7 (with the reduction $\bar{U}_t \to U_t$ via Assumption A.6) exactly as in the base step. This establishes $(\star_t)$.

Conclusion. Taking unconditional expectations of $(\star_0)$ and noting $\bar{U}_0 = U_0$,

$$\theta^a = \mathbb{E}[Y^{\bar{a}_{T-1}}] \overset{\text{LIE}}{=} \mathbb{E}\big[\mathbb{E}[Y^{\bar{a}_{T-1}} \mid U_0, \bar{X}_0]\big] \overset{(\star_0)}{=} \mathbb{E}\big[\mathbb{E}[H_0 \mid U_0, \bar{X}_0]\big] \overset{\text{LIE}}{=} \mathbb{E}[H_0]. \qquad \square$$

Theorem A.12 (Proximal identification via treatment bridge functions)**.** *Under Assumptions A.1–A.5, A.6, A.8, and positivity (Assumption A.2), for any $\{q_t\}$ satisfying (3):*

$$\theta^a = \mathbb{E}[Y^{\bar{a}_{T-1}}] = \mathbb{E}[Y \cdot \mathbf{1}(\bar{A}_{T-1} = \bar{a}_{T-1}) \cdot Q_{T-1}].$$

Proof. We show by forward induction the U-conditional identity

$$\mathbb{E}[Q_t \mid \bar{a}_t, \bar{U}_t, \bar{X}_t] = \prod_{s=0}^{t} \frac{1}{P(A_s = a_s \mid \bar{a}_{s-1}, U_s, \bar{X}_s)} \quad \text{for } t = 0, \dots, T-1, \qquad (\star'_t)$$

with $\bar{a}_{-1} := \emptyset$ and $\bar{U}_0 := U_0$. Positivity (Assumption A.2) ensures the propensity terms are bounded away from zero. Conditioning on the full trajectory $\bar{U}_t$ preserves the σ-algebra subset relations across time steps; the right-hand side is written in terms of single U_s via Assumption A.6.

Base step ($t = 0$). The moment condition (3) at $t = 0$, with $Q_{-1} = 1$, reads

$$\mathbb{E}[Q_0 \mid A_0, \bar{W}_0, \bar{X}_0] = \frac{1}{P(A_0 = a_0 \mid \bar{W}_0, \bar{X}_0)}$$

on $\{A_0 = a_0\}$. Apply iterated expectations introducing $U_0 = \bar{U}_0$ in the inner conditioning, use Assumption A.5 to swap $\bar{W}_0$ for U_0 on both sides, and Assumption A.8 to invert the resulting moment equation. This yields $(\star'_0)$.

Inductive step. Suppose $(\star'_{t-1})$ holds for some $t \in \{1, \dots, T-1\}$. From the moment condition (3) at time t, applied via iterated expectations and Assumption A.8 (with the reduction $\bar{U}_t \to U_t$ via Assumption A.6) exactly as in the base step:

$$\mathbb{E}[Q_t \mid \bar{a}_t, \bar{U}_t, \bar{X}_t] = \frac{\mathbb{E}[Q_{t-1} \mid \bar{a}_{t-1}, \bar{U}_t, \bar{X}_t]}{P(A_t = a_t \mid \bar{a}_{t-1}, U_t, \bar{X}_t)}.$$

Since $Q_{t-1} = q_{t-1}(\bar{Z}_{t-1}, \bar{a}_{t-1}, \bar{X}_{t-1})$ is past-measurable, Assumption A.6 gives $\bar{Z}_{t-1} \perp\!\!\!\perp (U_t, X_t) \mid \bar{a}_{t-1}, \bar{U}_{t-1}, \bar{X}_{t-1}$, hence

$$\mathbb{E}[Q_{t-1} \mid \bar{a}_{t-1}, \bar{U}_t, \bar{X}_t] = \mathbb{E}[Q_{t-1} \mid \bar{a}_{t-1}, \bar{U}_{t-1}, \bar{X}_{t-1}].$$

Applying the inductive hypothesis $(\star'_{t-1})$ to the right-hand side:

$$\begin{aligned}\mathbb{E}[Q_t \mid \bar{a}_t, \bar{U}_t, \bar{X}_t] &= \frac{1}{P(A_t = a_t \mid \bar{a}_{t-1}, U_t, \bar{X}_t)} \cdot \prod_{s=0}^{t-1} \frac{1}{P(A_s = a_s \mid \bar{a}_{s-1}, U_s, \bar{X}_s)} \\ &= \prod_{s=0}^{t} \frac{1}{P(A_s = a_s \mid \bar{a}_{s-1}, U_s, \bar{X}_s)},\end{aligned}$$

establishing $(\star'_t)$.

Conclusion. By inverse-probability weighting under Assumptions A.5, A.6, and A.2, with Assumption A.1 replacing $Y^{\bar{a}_{T-1}}$ by Y on the indicator event,

$$\mathbb{E}[Y^{\bar{a}_{T-1}}] = \mathbb{E}\left[\frac{Y \cdot \mathbf{1}(\bar{A}_{T-1} = \bar{a}_{T-1})}{\prod_{s=0}^{T-1} P(A_s = a_s \mid \bar{a}_{s-1}, U_s, \bar{X}_s)}\right].$$

Substituting $(\star'_{T-1})$ for the cumulative inverse-propensity product:

$$\mathbb{E}[Y^{\bar{a}_{T-1}}] = \mathbb{E}\big[Y \cdot \mathbf{1}(\bar{A}_{T-1} = \bar{a}_{T-1}) \cdot \mathbb{E}[Q_{T-1} \mid \bar{a}_{T-1}, \bar{U}_{T-1}, \bar{X}_{T-1}]\big] = \mathbb{E}[Y \cdot \mathbf{1}(\bar{A}_{T-1} = \bar{a}_{T-1}) \cdot Q_{T-1}],$$

where the final equality is iterated expectations. □

B Regularity Conditions

The following conditions ensure existence of bridge functions (Lemmas A.9 and A.10) and support the error-scaling theorems of Appendices F and G. We state them for the outcome bridge family $\{h_t\}$; the analogous conditions for $\{q_t\}$ are obtained by exchanging the roles of W and Z.

Notation. For each $t = 0, \ldots, T-1$, let h_t^* denote the true bridge function and $\mathcal{H}$ the parameterized hypothesis class with element h_t with η. The *bridge operator* $\mathcal{E}_t$ maps a bridge function h_t to its conditional expectation under the treatment-side conditioning,

$$\mathcal{E}_t h_t \;=\; \mathbb{E}[h_t(\bar{W}_t, \bar{a}_{T-1}, \bar{X}_t) \mid \bar{A}_t, \bar{Z}_t, \bar{X}_t],$$

and the *forward coupling operator* $\mathcal{P}_t$ maps the next-stage bridge h_{t+1} to its conditional expectation under the same conditioning, so that the moment condition of Eq. (1) reads $\mathcal{E}_t h_t^* = \mathcal{P}_t h_{t+1}^*$. In the linear bridge model $h_t = b_t^\top \phi_t$ with feature vector ϕ_t and instrument vector ψ_t (Appendix F), $\mathcal{E}_t$ reduces to multiplication by $\Sigma_t = \mathbb{E}[\psi_t \phi_t^\top]$ and $\mathcal{P}_t$ to multiplication by $\Gamma_t = \mathbb{E}[\psi_t \phi_{t+1}^\top]$. The quantity $\kappa_t := \|\mathcal{E}_t^{-1}\mathcal{P}_t\|_{L^2 \to L^2}$ measures how errors at stage $t+1$ propagate into stage t under sequential inversion; we refer to it as the *amplification factor* following the discussion in Appendix F. The kernel-induced discrepancy

$$d_{k_t}^2(\eta^*, \eta) := \big\|\mathbb{E}\big[(H_{t+1}(\eta) - H_t(\eta)) - (H_{t+1}(\eta^*) - H_t(\eta^*)) \,\big|\, \bar{A}_t, \bar{Z}_t, \bar{X}_t\big]\big\|_{\mathcal{H}_{k_t}}^2$$

quantifies the joint moment-condition residual between the parameters η and the true value η^*. Since $\mathbb{E}[H_{t+1}^* - H_t^* \mid \cdot] = 0$ by Eq. (1), this simplifies to

$$d_{k_t}^2(\eta^*, \eta) = \big\|\mathbb{E}\big[H_{t+1}(\eta) - H_t(\eta) \,\big|\, \bar{A}_t, \bar{Z}_t, \bar{X}_t\big]\big\|_{\mathcal{H}_{k_t}}^2 = \mathcal{R}_{k_t}(\eta),$$

where $\mathcal{R}_{k_t}(\eta)$ is the population MMR risk at η. The discrepancy couples consecutive bridges (h_t, h_{t+1}) to (h_t^*, h_{t+1}^*) through the conditional expectation of their residual difference. It is a joint discrepancy on parameter pairs, not a per-stage semi-metric on individual bridge functions; this distinction is delivered by the longitudinal coupling of the moment $f_t^H = H_{t+1} - H_t$, in contrast to the single-step NMMR setting where the moment depends on one bridge only.

Condition 1 (Instrument relevance)**.** For each $t = 0, \ldots, T-1$, the bridge operator $\mathcal{E}_t$ has a bounded regularized inverse on the range of $\mathcal{P}_t$:

$$\|\mathcal{E}_t^{-1}\mathcal{P}_t\|_{L^2 \to L^2} = \kappa_t < \infty.$$

In the linear setting, this reduces to $\sigma_{\min}(\Sigma_t) \geq c > 0$ uniformly for all t.

Condition 2 (Bounded moments)**.** The feature vectors, instrument vectors, and outcomes satisfy $\mathbb{E}[\|\phi_t\|^4] < \infty$, $\mathbb{E}[\|\psi_t\|^4] < \infty$, and $\mathbb{E}[Y^4] < \infty$ for all t. In the nonparametric setting, there exist finite constants M, M_Y such that $\sup_{h \in \mathcal{H}} \|h\|_\infty \leq M$ and $\|Y\|_\infty \leq M_Y$ almost surely.

Condition 3 (i.i.d. sampling)**.** The data $\{(\bar{A}_{i,T-1}, \bar{L}_{i,T-1}, Y_i)\}_{i=1}^n$ are i.i.d.

Condition 4 (Bounded amplification with strict inflation)**.** There exist constants $1 < \underline{\kappa} \leq \overline{\kappa} < \infty$ such that

$$\underline{\kappa} \leq \kappa_t \leq \overline{\kappa} \qquad \text{for all } t = 0, \ldots, T-2.$$

The lower bound $\underline{\kappa} > 1$ enforces strict ill-posedness at every stage; the upper bound $\overline{\kappa} < \infty$ controls the per-stage amplification.

Condition 5 (ISPD kernel)**.** For each t, the kernel k_t is integrally strictly positive definite (ISPD) on the support of $(\bar{A}_t, \bar{Z}_t, \bar{X}_t)$ and uniformly bounded: $\|k_t\|_\infty \leq M_k < \infty$.

Condition 6 (Controlled complexity)**.** The hypothesis class $\mathcal{H}$ is parameterized by $\eta = (\eta_{\text{enc}}, \eta_{\text{head},0}, \ldots, \eta_{\text{head},T-1})$, with a shared encoder of p_{enc} parameters (independent of T) and per-time-step heads of p_{head} parameters each, so that $|\eta| = p_{\text{enc}} + T \cdot p_{\text{head}}$. The residual $r_t(\eta) = \hat{H}^a_{t+1} - \hat{H}^A_t$ depends on η only through the active sub-vector $\eta_t^{\text{act}} := (\eta_{\text{enc}}, \eta_{\text{head},t}, \eta_{\text{head},t+1})$, whose dimension $p^{\text{act}} := p_{\text{enc}} + 2p_{\text{head}}$ is $O(1)$ in T. Let $\mathcal{H}_t^{\text{act}} := \{\eta_t^{\text{act}} : \eta \in \mathcal{H}\}$. The Rademacher complexity satisfies $\mathcal{R}_n(\mathcal{H}_t^{\text{act}}) = O(\sqrt{p^{\text{act}}/n}) = O(n^{-1/2})$ uniformly in t, and $\inf_{\eta \in \mathcal{H}} \frac{1}{T}\sum_t d^2_{k_t}(\eta^*, \eta) \to 0$ as head capacity grows.

C Maximum Moment Restriction: Background

A conditional moment restriction is a constraint of the form

$$\mathbb{E}[f(X) \mid Y] = 0 \quad \text{almost surely}, \tag{23}$$

for some function f depending on parameters η and observable variables X. Both bridge function definitions in Section 2 take this form: the outcome bridge moment condition in Eq. (1) is recovered with $f_t^H = H_{t+1} - H_t$ and conditioning variables $(\bar{A}_t, \bar{Z}_t, \bar{X}_t)$, and the treatment bridge moment condition in Eq. (3) is recovered after rearrangement, with the three-component f_t^Q described in Section 3.1.

Conditional moment restrictions are not directly usable as training losses, because evaluating $\mathbb{E}[f(X) \mid Y]$ requires inverting a conditional expectation operator; a separate ill-posed problem. The Maximum Moment Restriction (MMR) framework (Mastouri et al., 2021; Kompa et al., 2022; Muandet et al., 2020) avoids this inversion by replacing the conditional expectation with a supremum over test functions. For the restriction in Eq. (23), the MMR risk is

$$\mathcal{R}(f) = \sup_{\|g\| \leq 1} \big(\mathbb{E}[f(X) \cdot g(Y)]\big)^2, \tag{24}$$

with the supremum taken over unit-norm measurable functions g. By construction, $\mathcal{R}(f) = 0$ if and only if $\mathbb{E}[f(X) \mid Y] = 0$ almost surely, so minimizing $\mathcal{R}(f)$ is equivalent to satisfying the original conditional restriction.

When the supremum in Eq. (24) is restricted to the unit ball of a reproducing kernel Hilbert space (RKHS) with integrally strictly positive definite (ISPD) kernel k, the risk admits a closed-form expression (Mastouri et al., 2021):

$$\mathcal{R}(f) = \mathbb{E}_{(X,Y),(X',Y')}\big[f(X)\, f(X')\, k(Y, Y')\big],$$

where (X, Y) and (X', Y') are independent copies. The empirical V-statistic of $\mathcal{R}(f)$,

$$\hat{\mathcal{R}}(f) = \frac{1}{n^2} \sum_{i,j=1}^{n} f(X_i)\, f(X_j)\, k(Y_i, Y_j), \tag{25}$$

serves as a trainable scalar loss (Kompa et al., 2022). Eq. (25) is the elementary building block of the joint MMR loss in Section 3.1: each time step of each bridge family contributes one term of this form, the kernel argument is the appropriate conditioning vector, and the moment function f is either f_t^H or f_t^Q. Aggregating these per-step V-statistics across $t = 0, \ldots, T-1$ gives Eq. (8) in the main text.

The vector-valued moment function $f_t^Q = (Q_t, -Q_{t-1}, -Q_{t-1})^\top$ used by the treatment bridge requires a small generalization of Eq. (25): the scalar product $f(X_i)\, f(X_j)$ is replaced by the bilinear form $f(X_i)^\top \mathbf{K}_{ij}\, f(X_j)$, where $\mathbf{K}_{ij}$ is a kernel matrix indexing the components of the moment vector. Section 3.3 specifies $\mathbf{K}_{ij}$ for the treatment bridge.

D Derivation of the Propensity-Free Moment Condition

We derive the expanded form in Eq. (12) from the treatment bridge moment condition of Eq. (3) and show that it yields the three-component moment function $f_t^Q = (Q_t, -Q_{t-1}, -Q_{t-1})^\top$.

Proof. The treatment bridge moment condition of Eq. (3) states that for each $t = 0, \ldots, T-1$:

$$\mathbb{E}[Q_t \mid \bar{A}_t, \bar{W}_t, \bar{X}_t] = \frac{\mathbb{E}[Q_{t-1} \mid \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t]}{\pi_t}, \tag{$\star$}$$

where $\pi_t = P(A_t = a_t \mid \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t)$ and $Q_{-1} = 1$.

Step 1: From conditional moment to unconditional moment. Let $N_t(\bar{A}_t, \bar{W}_t, \bar{X}_t)$ be an arbitrary test function measurable with respect to $(\bar{A}_t, \bar{W}_t, \bar{X}_t)$. Since N_t is measurable with respect to the conditioning set, the law of iterated expectations gives:

$$\begin{aligned}\mathbb{E}[Q_t \cdot N_t(\bar{A}_t, \bar{W}_t, \bar{X}_t)] &= \mathbb{E}\big[\mathbb{E}[Q_t \mid \bar{A}_t, \bar{W}_t, \bar{X}_t] \cdot N_t(\bar{A}_t, \bar{W}_t, \bar{X}_t)\big] \\ &= \mathbb{E}\left[\frac{\mathbb{E}[Q_{t-1} \mid \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t]}{\pi_t} \cdot N_t(\bar{A}_t, \bar{W}_t, \bar{X}_t)\right], \end{aligned} \tag{26}$$

where the second equality applies ($\star$).

Step 2: Eliminating the propensity denominator. We condition the right-hand side of Eq. (26) on $(\bar{A}_{t-1}, \bar{W}_t, \bar{X}_t)$ and take the expectation over A_t. Since $\mathbb{E}[Q_{t-1} \mid \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t]$ does not depend on A_t, it factors out of the inner expectation:

$$\begin{aligned}&\mathbb{E}\left[\frac{\mathbb{E}[Q_{t-1} \mid \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t]}{\pi_t} \cdot N_t(\bar{A}_t, \bar{W}_t, \bar{X}_t)\right] \\ &= \mathbb{E}\left[\mathbb{E}[Q_{t-1} \mid \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t] \cdot \mathbb{E}\left[\frac{N_t(A_t, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t)}{\pi_t} \;\middle|\; \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t\right]\right]. \end{aligned} \tag{27}$$

For binary treatment $A_t \in \{0, 1\}$, the inner expectation expands as:

$$\begin{aligned}\mathbb{E}\left[\frac{N_t(A_t, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t)}{P(A_t \mid \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t)} \;\middle|\; \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t\right] &= \sum_{a_t \in \{0,1\}} \frac{N_t(a_t, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t)}{\pi_t} \cdot \pi_t \\ &= \sum_{a_t \in \{0,1\}} N_t(a_t, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t) \\ &= N_t(1, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t) + N_t(0, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t). \end{aligned} \tag{28}$$

The propensity terms cancel exactly.

Step 3: Recombining. Substituting Eq. (28) into Eq. (27), and using the fact that $N_t(1, \ldots) + N_t(0, \ldots)$ is measurable with respect to $(\bar{A}_{t-1}, \bar{W}_t, \bar{X}_t)$ to absorb it into the conditional expectation of Q_{t-1}:

$$\mathbb{E}[Q_t \cdot N_t(\bar{A}_t, \bar{W}_t, \bar{X}_t)] = \mathbb{E}\big[Q_{t-1} \cdot \big(N_t(1, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t) + N_t(0, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t)\big)\big]. \tag{29}$$

This is Eq. (12). Crucially, the derivation uses only the law of iterated expectations and the definition of conditional probability for binary A_t; no parametric assumptions on Q_t are required.

Step 4: Three-component moment function. Rearranging Eq. (29):

$$\mathbb{E}\big[Q_t \cdot N_t(A_t, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t) - Q_{t-1} \cdot N_t(1, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t) - Q_{t-1} \cdot N_t(0, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t)\big] = 0. \tag{30}$$

Defining the three-component vector

$$f_t^Q = \begin{pmatrix} Q_t \\ -Q_{t-1} \\ -Q_{t-1} \end{pmatrix}, \quad N_t^{(\text{vec})} = \begin{pmatrix} N_t(A_t, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t) \\ N_t(1, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t) \\ N_t(0, \bar{A}_{t-1}, \bar{W}_t, \bar{X}_t) \end{pmatrix}, \tag{31}$$

the condition becomes $\mathbb{E}[(f_t^Q)^\top N_t^{(\text{vec})}] = 0$ for all test functions N_t, which is the form enforced by the MMR framework of Eq. (24) with a 3×3 kernel matrix $K_{ij,t}$ whose (ℓ, m) entry evaluates the kernel between the ℓ-th treatment variant of sample i and the m-th treatment variant of sample j. □

E Proof of Double Robustness of P-DR

The proof follows the standard argument adapted from Ying et al. (2023), Theorem 4, to the indicator-based estimator under joint training. We first establish double robustness for the augmented form in Eq. (7); Appendix E.1 then shows that the targeting loss of Eq. (19) used in our implementation is algebraically equivalent to Eq. (7) at the bridge solution, so the result transfers to $\hat{\theta}^a_{\text{PDR}}$.

Case (i): Outcome bridge correctly specified. Suppose $\{H_t\}$ satisfy the moment condition of Eq. (1): $\mathbb{E}[H_{t+1} \mid \bar{a}_t, \bar{Z}_t, \bar{X}_t] = \mathbb{E}[H_t \mid \bar{a}_t, \bar{Z}_t, \bar{X}_t]$ for all t. By the law of iterated expectations and Theorem A.11, $\mathbb{E}[H_0^a] = \mathbb{E}[Y^{\bar{a}_{T-1}}] = \theta^a$.

The correction terms in the P-DR estimator are $\Delta_t = \mathbf{1}(\bar{A}_t = \bar{a}_t)Q_t(H_{t+1}^a - H_t^a)$. Since the outcome bridge is correct, $\mathbb{E}[H_{t+1}^a - H_t^a \mid \bar{a}_t, \bar{Z}_t, \bar{X}_t] = 0$, so $\mathbb{E}[\Delta_t] = 0$ regardless of $\{Q_t\}$. Hence $\mathbb{E}[\hat{\theta}^a_{\text{PDR}}] = \mathbb{E}[H_0^a] + \sum_t \mathbb{E}[\Delta_t] = \theta^a$.

Case (ii): Treatment bridge correctly specified. Suppose $\{Q_t\}$ satisfy the moment condition of Eq. (3). Expanding the P-DR estimator and applying the recursive property of Q_t at each timestep, combined with the indicator function selecting the intervention path, the sum telescopes to give:

$$\mathbb{E}\left[\sum_{t=0}^{T-1} \mathbf{1}(\bar{A}_t = \bar{a}_t)Q_t(H_{t+1}^a - H_t^a)\right] = \mathbb{E}[Y^{\bar{a}_{T-1}}] - \mathbb{E}[H_0^a].$$

Therefore $\mathbb{E}[\hat{\theta}^a_{\text{PDR}}] = \mathbb{E}[H_0^a] + (\mathbb{E}[Y^{\bar{a}_{T-1}}] - \mathbb{E}[H_0^a]) = \theta^a$.

Independence from training procedure and estimator form. The above arguments depend only on the population-level properties of the bridge function estimates in the limit ($\hat{H}_t \xrightarrow{p} H_t$ or $\hat{Q}_t \xrightarrow{p} Q_t$). Whether the networks are trained jointly or sequentially affects the finite-sample convergence rate but not the consistency guarantee under correct specification of either bridge family. Likewise, by Appendix E.1, the targeting form in Eq. (19) and the augmented form in Eq. (7) agree at the bridge solution, so the consistency guarantee is independent of which form is used at training time.

E.1 Equivalence of the Targeted Plug-in and the Augmented Estimator

The targeting loss in Eq. (19) is constructed so that the targeted plug-in $\tilde{\theta}^a_{\text{PDR}} = \frac{1}{n}\sum_i \tilde{H}^a_{0,i}$ coincides with the empirical version of Eq. (7) at any stationary point of $\mathcal{L}_{\text{P-DR}}(\eta, \epsilon)$. The argument adapts Frauen et al. (2023), Theorem 1, replacing their inverse-propensity clever covariate with the estimated treatment bridge $\hat{Q}_t$.

First-order condition for ϵ. Recall from Eq. (18) that $\tilde{H}_t^a = \hat{H}_t^a + \epsilon \cdot \xi_t$ with $\xi_t = \xi_{t+1} - \mathbb{1}(\bar{A}_t = \bar{a}_t) \cdot \hat{Q}_t$ and $\xi_T = 0$. At a stationary point $(\hat{\eta}, \hat{\epsilon})$ of $\mathcal{L}_{\text{P-DR}}$, the partial derivative with respect to ϵ vanishes. Since $\mathcal{L}_H$ and $\mathcal{L}_Q$ do not depend on ϵ,

$$0 = \left.\frac{\partial \mathcal{L}_{\text{P-DR}}}{\partial \epsilon}\right|_{(\hat{\eta}, \hat{\epsilon})} = \lambda_{\text{tar}} \cdot \left.\frac{\partial \mathcal{L}_{\text{tar}}}{\partial \epsilon}\right|_{(\hat{\eta}, \hat{\epsilon})}.$$

Differentiating Eq. (19) and using $\partial(\tilde{H}_t^a - \tilde{H}_{t+1}^a)/\partial\epsilon = \xi_t - \xi_{t+1} = -\mathbb{1}(\bar{A}_t = \bar{a}_t) \cdot \hat{Q}_t$,

$$\frac{\partial \mathcal{L}_{\text{tar}}}{\partial \epsilon} = \frac{2}{nT} \sum_{i=1}^{n} \sum_{t=0}^{T-1} \left(\tilde{H}_{i,t+1}^a - \tilde{H}_{i,t}^a \right) \cdot \mathbb{1}(\bar{A}_{i,t} = \bar{a}_t) \cdot \hat{Q}_{t,i}.$$

Setting this to zero at $(\hat{\eta}, \hat{\epsilon})$ yields the empirical efficient estimating equation evaluated at the targeted outputs:

$$\frac{1}{n} \sum_{i=1}^{n} \sum_{t=0}^{T-1} \left(\tilde{H}_{i,t+1}^a - \tilde{H}_{i,t}^a \right) \cdot \mathbb{1}(\bar{A}_{i,t} = \bar{a}_t) \cdot \hat{Q}_{t,i} = 0. \tag{32}$$

Eq. (32) is the analog of Eq. (16) in Frauen et al. (2023).

Equivalence with the augmented form. Let $\tilde{\theta}_{\text{aug}}^a$ denote the empirical version of Eq. (7) evaluated with $\tilde{H}$:

$$\tilde{\theta}_{\text{aug}}^a = \frac{1}{n} \sum_{i=1}^{n} \left[\tilde{H}_{0,i}^a + \sum_{t=0}^{T-1} \mathbb{1}(\bar{A}_{i,t} = \bar{a}_t) \cdot \hat{Q}_{t,i} \cdot \left(\tilde{H}_{t+1,i}^a - \tilde{H}_{t,i}^a \right) \right].$$

By Eq. (32) the second term vanishes, so $\tilde{\theta}_{\text{aug}}^a = \frac{1}{n} \sum_i \tilde{H}_{0,i}^a = \tilde{\theta}_{\text{PDR}}^a$. The targeted plug-in and the augmented estimator coincide at any stationary point of $\mathcal{L}_{\text{P-DR}}$.

Transfer of double robustness. Cases (i) and (ii) above establish consistency of the empirical augmented form when either bridge family is correctly specified. Since $\tilde{\theta}_{\text{PDR}}^a$ equals this augmented form at the stationary point, the consistency results transfer directly: $\tilde{\theta}_{\text{PDR}}^a$ remains consistent for θ^a whenever either $\{H_t\}$ or $\{Q_t\}$ is correctly specified, regardless of training-time form.

Joint vs. staged optimization. The equivalence depends only on the stationary-point condition for ϵ. The staged variant of Frauen et al. (2023), which fits the bridges first and updates ϵ with bridges frozen, satisfies the same condition and yields the same population target. The two procedures differ only in finite-sample optimization dynamics.

F Sequential Estimation and Exponential Error Growth

The following theorem provides the formal statement and proof supporting Finding 1 in Section 4. It establishes that sequential bridge function estimation suffers from error growth exponential in T — a structural consequence of composing T ill-posed inverse problems recursively.

Theorem F.1 (Sequential error propagation)**.** *Consider the sequential backward estimator $\{\hat{h}_t^{\text{seq}}\}_{t=0}^{T-1}$ where each stage t solves the empirical moment condition using $\hat{h}_{t+1}^{\text{seq}}$ as the pseudo-outcome. Let ε_t denote the intrinsic estimation error at stage t — the L^2 error arising from finite-sample noise alone if the true h_{t+1}^* were used. Under Conditions 1–3, the error at $t = 0$ satisfies:*

$$\|\hat{h}_0^{\text{seq}} - h_0^*\|_{L^2} \leq \sum_{t=0}^{T-1} \left(\prod_{s=0}^{t-1} \kappa_s \right) \varepsilon_t. \tag{33}$$

Under Condition 4 ($1 < \underline{\kappa} \leq \kappa_t \leq \overline{\kappa}$) and $\varepsilon_t = O_p(r_n)$ for a common rate r_n:

$$\Omega_p\big(\underline{\kappa}^{T-1} \cdot r_n\big) \;\leq\; \|\hat{h}_0^{\text{seq}} - h_0^*\|_{L^2} \;\leq\; O_p\big(\overline{\kappa}^{T-1} \cdot r_n\big),$$

with the lower bound holding under the constructive-compounding regime described in Remark F.2. Both bounds reduce to $\Theta_p(\kappa^{T-1} r_n)$ when $\underline{\kappa} \approx \overline{\kappa} \approx \kappa$.

The same structural argument applies to the treatment bridge functions $\{Q_t\}$, which propagate forward rather than backward, with the product of amplification factors accumulating in the forward direction.

F.1 Linear Case

Stage $T-1$ (Base case). The sequential estimator at the final stage solves:

$$\hat{\Sigma}_{T-1}^{\top}\hat{b}_{T-1}^{\text{seq}} = \frac{1}{n}\sum_{i=1}^{n}\psi_{T-1,i}Y_i,$$

where $\hat{\Sigma}_t = \frac{1}{n}\sum_{i=1}^{n}\psi_{t,i}\phi_{t,i}^{\top}$. The true parameters satisfy $\mathbb{E}[\psi_{T-1}(\phi_{T-1}^{\top}b_{T-1}^*)] = \mathbb{E}[\psi_{T-1}Y]$, so:

$$\hat{b}_{T-1}^{\text{seq}} - b_{T-1}^* = (\hat{\Sigma}_{T-1}^{\top}\hat{\Sigma}_{T-1})^{-1}\hat{\Sigma}_{T-1}^{\top}\hat{u}_{T-1} =: \epsilon_{T-1},$$

where $\hat{u}_{T-1} = \frac{1}{n}\sum_i \psi_{T-1,i}(Y_i - \phi_{T-1,i}^{\top}b_{T-1}^*)$ is the sample moment evaluated at the true parameters. By standard GMM theory under Conditions 1–3, $\epsilon_{T-1} = O_p(n^{-1/2})$.

Stage $t < T-1$ (Inductive step). At stage t, the sequential estimator uses $\hat{H}_{t+1}^{\text{seq}} = (\hat{b}_{t+1}^{\text{seq}})^{\top}\phi_{t+1}$ as the pseudo-outcome:

$$\hat{\Sigma}_t^{\top}\hat{b}_t^{\text{seq}} = \frac{1}{n}\sum_{i=1}^{n}\psi_{t,i}\cdot(\hat{b}_{t+1}^{\text{seq}})^{\top}\phi_{t+1,i}.$$

Decomposing the right-hand side by adding and subtracting b_{t+1}^*:

$$\frac{1}{n}\sum_i \psi_{t,i}(\hat{b}_{t+1}^{\text{seq}})^{\top}\phi_{t+1,i} = \left(\hat{\Sigma}_t^{\top}b_t^* + \hat{u}_t\right) + \hat{\Gamma}_t(\hat{b}_{t+1}^{\text{seq}} - b_{t+1}^*),$$

where $\hat{u}_t = \frac{1}{n}\sum_i \psi_{t,i}\left[(b_{t+1}^*)^{\top}\phi_{t+1,i} - (b_t^*)^{\top}\phi_{t,i}\right]$ is the stage-t sampling noise and $\hat{\Gamma}_t = \frac{1}{n}\sum_i \psi_{t,i}\phi_{t+1,i}^{\top}$. Combining:

$$\hat{b}_t^{\text{seq}} - b_t^* = \underbrace{(\hat{\Sigma}_t^{\top}\hat{\Sigma}_t)^{-1}\hat{\Sigma}_t^{\top}\hat{u}_t}_{\epsilon_t} + \underbrace{(\hat{\Sigma}_t^{\top}\hat{\Sigma}_t)^{-1}\hat{\Sigma}_t^{\top}\hat{\Gamma}_t}_{\hat{\Lambda}_t}(\hat{b}_{t+1}^{\text{seq}} - b_{t+1}^*),$$

where $\hat{\Lambda}_t \xrightarrow{p} \Lambda_t$ under Conditions 1–3. Unrolling the recursion from t to $T-1$:

$$\hat{b}_t^{\text{seq}} - b_t^* = \epsilon_t + \Lambda_t\epsilon_{t+1} + \Lambda_t\Lambda_{t+1}\epsilon_{t+2} + \cdots + \left(\prod_{s=t}^{T-2}\Lambda_s\right)\epsilon_{T-1} + o_p\Big(\max_s\|\epsilon_s\|\Big).$$

Taking norms, applying the triangle inequality, setting $t=0$, and identifying $\kappa_r = \|\Lambda_r\|_2$ yields Eq. (33).

F.2 General Nonparametric Case

Step 1: Decomposition. The sequential estimator at stage t satisfies the approximate optimality condition $\mathcal{E}_t\hat{h}_t^{\text{seq}} \approx \mathcal{P}_t\hat{h}_{t+1}^{\text{seq}}$ in the sense that the empirical moment condition is approximately zero. Subtracting the population moment condition $\mathcal{E}_t h_t^* = \mathcal{P}_t h_{t+1}^*$:

$$\mathcal{E}_t(\hat{h}_t^{\text{seq}} - h_t^*) \approx \mathcal{P}_t(\hat{h}_{t+1}^{\text{seq}} - h_{t+1}^*) + (\text{sampling noise at stage } t).$$

Step 2: Inversion and amplification. Applying the regularized inverse $\mathcal{E}_t^{-1}$ and taking L^2 norms:

$$\|\hat{h}_t^{\text{seq}} - h_t^*\|_{L^2} \le \underbrace{\|\mathcal{E}_t^{-1}\mathcal{P}_t\|_{L^2\to L^2}}_{\kappa_t}\cdot\|\hat{h}_{t+1}^{\text{seq}} - h_{t+1}^*\|_{L^2} + \varepsilon_t,$$

where $\varepsilon_t = \|\mathcal{E}_t^{-1}(\text{sampling noise})\|_{L^2}$ is the intrinsic error at stage t.

Step 3: Unrolling. Define $e_t = \|\hat{h}_t^{\text{seq}} - h_t^*\|_{L^2}$. The recursion $e_t \le \kappa_t\cdot e_{t+1} + \varepsilon_t$ with $e_{T-1} = \varepsilon_{T-1}$ yields:

$$e_t \le \sum_{s=t}^{T-1}\left(\prod_{r=t}^{s-1}\kappa_r\right)\varepsilon_s.$$

Under Condition 4, each product satisfies $\prod_{r=t}^{s-1} \kappa_r \le \overline{\kappa}^{s-t}$, so

$$e_0 \le \Big(\max_t \varepsilon_t\Big) \sum_{s=0}^{T-1} \overline{\kappa}^s = \Big(\max_t \varepsilon_t\Big) \cdot \frac{\overline{\kappa}^T - 1}{\overline{\kappa} - 1},$$

establishing $\|\hat{h}_0^{\text{seq}} - h_0^*\|_{L^2} = O_p(\overline{\kappa}^{T-1} \cdot r_n)$ when $\varepsilon_t = O_p(r_n)$. The matching lower bound $\Omega_p(\underline{\kappa}^{T-1} \cdot r_n)$ follows from the dual recursion $\prod_{r=t}^{s-1} \kappa_r \ge \underline{\kappa}^{s-t}$ under the constructive-compounding scenario of Remark F.2.

Remark F.2 (Tightness of the sequential bound). The upper bound in Theorem F.1 is achieved up to constants when (i) intrinsic errors ε_t at different stages are positively correlated — the generic case since they arise from the same dataset — and (ii) the amplification operators $\mathcal{E}_t^{-1}\mathcal{P}_t$ act constructively rather than with cancellation. Under these conditions, the product $\prod_{r=0}^{T-2} \kappa_r$ is achieved with constant relative deviation, and

$$\|\hat{h}_0^{\text{seq}} - h_0^*\|_{L^2} = \Theta_p\big(\kappa^{T-1} \cdot r_n\big)$$

with κ in the interval $[\underline{\kappa}, \overline{\kappa}]$. Adversarial constructions of data-generating processes where the upper bound is attained within constant factors are straightforward.

F.3 Linear Amplification Factor

The following proposition provides an explicit characterization of the amplification factor κ_t in the linear bridge function setting, supporting Theorem F.1.

Proposition F.3 (Linear amplification factor)**.** *In the linear model $h_t = b_t^\top \phi_t$, the bridge operator $\mathcal{E}_t$ corresponds to left-multiplication by $\Sigma_t = \mathbb{E}[\psi_t \phi_t^\top]$ and the forward coupling operator $\mathcal{P}_t$ corresponds to left-multiplication by $\Gamma_t = \mathbb{E}[\psi_t \phi_{t+1}^\top]$. The amplification factor defined in Condition 1 reduces to:*

$$\kappa_t = \|\mathcal{E}_t^{-1}\mathcal{P}_t\|_{L^2 \to L^2} = \|\Lambda_t\|_2, \quad \Lambda_t = (\Sigma_t^\top \Sigma_t)^{-1} \Sigma_t^\top \Gamma_t.$$

Proof. A perturbation $\delta_{t+1} = \delta^\top \phi_{t+1}$ in the bridge function at $t+1$ creates a perturbation in the right-hand side of the moment condition at t: $\mathcal{P}_t \delta_{t+1} \leftrightarrow \Gamma_t \delta$. Inverting $\mathcal{E}_t$ gives $\mathcal{E}_t^{-1}\mathcal{P}_t \delta_{t+1} \leftrightarrow (\Sigma_t^\top \Sigma_t)^{-1} \Sigma_t^\top \Gamma_t \delta = \Lambda_t \delta$. Thus $\kappa_t = \|\Lambda_t\|_2$. □

F.4 Connection Between Linear and General Amplification Factors

For completeness, we verify that the general framework reduces correctly to the linear setting and discuss intermediate cases.

Linear model. Proposition F.3 confirms $\kappa_t = \|\Lambda_t\|_2$ directly.

RKHS model. When bridge functions lie in an RKHS $\mathcal{H}_k$, the amplification factor κ_t relates to the condition number of $\mathcal{E}_t$ restricted to $\mathcal{H}_k$. Specifically, $\kappa_t \le \tau_t$ where τ_t is the ill-posedness parameter (Kompa et al., 2022):

$$\tau_t = \sup_{h \in \mathcal{H}_k,\, h \neq h_t^*} \frac{\|h - h_t^*\|_{L^2}}{\|\mathbb{E}[h - h_t^* \mid \bar{A}_t, \bar{Z}_t, \bar{X}_t]\|_{L^2}}.$$

The inequality $\kappa_t \le \tau_t$ follows because $\|\mathcal{P}_t\| \le 1$ (conditional expectation is an L^2 contraction). When $\tau_t > 1$ the problem is ill-posed and $\kappa_t > 1$ is expected.

Neural network model. For the LSTM architecture used in Neural-PCI, the amplification factor is not closed-form computable, but can be recovered empirically by fitting an exponential $\varepsilon(T) \propto \kappa^{T-1}$ to the error trajectory of a sequential neural estimator. Applied to the Sequential NMMR baseline (Figure 2), this fit yields $\hat{\kappa} \approx 1.12$ — modestly above unity, consistent with Condition 4 ($\kappa > 1$), and sufficient to drive the multiplicative error compounding predicted by Theorem F.1. The linear analysis provides a structural lower bound when the data-generating process admits a linear representation; the empirical $\hat{\kappa}$ indicates that, in this setting, the LSTM neither cancels nor dramatically exacerbates ill-posedness across stages.

G JOINT MMR ESTIMATION AND ERROR BOUND

The following theorem and corollaries provide the formal statements and proofs supporting Finding 2 in Section 4. They establish that the joint MMR estimator achieves additive error aggregation across time steps, reducing the error scaling from exponential to at most linear in T.

Theorem G.1 (Joint MMR error bound)**.** *For each $t = 0, \dots, T-1$, let k_t be a bounded ISPD kernel with $\|k_t\|_\infty \le M_k$ (Condition 5), and let $\mathcal{H}$ be the hypothesis class of Condition 6. Let $d^2_{k_t}(\eta^*, \eta) = \mathcal{R}_{k_t}(\eta)$ denote the joint kernel-induced discrepancy of Appendix B. The joint estimator $\hat{\eta} = \arg\min_\eta \hat{\mathcal{L}}(\eta)$ from Eq. (8) satisfies, with probability at least $1-\delta$,*

$$\frac{1}{T}\sum_{t=0}^{T-1} d^2_{k_t}(\eta^*, \hat{\eta}) \;\le\; \frac{1}{T}\sum_{t=0}^{T-1} d^2_{k_t}(\eta^*, \tilde{\eta}) \;+\; \lambda M_\lambda \;+\; C_1\, \mathcal{R}_n(\mathcal{H}_t^{\mathrm{act}})\, M_k \;+\; C_2\, M_k^2 \sqrt{\frac{\log(2T/\delta)}{n}}, \tag{34}$$

where $\tilde{\eta}$ is the best approximator within $\mathcal{H}$, C_1, C_2 depend on $M = M_Y + \sup_\eta \|h^\eta\|_\infty$, and $\mathcal{R}_n(\mathcal{H}_t^{\mathrm{act}})$ is the active-parameter Rademacher complexity uniform in t. The bound involves neither κ_t nor the ill-posedness parameter τ_t: the joint estimator finds parameters η that simultaneously minimize all T moment-condition discrepancies, bypassing per-stage operator inversion.

Corollary G.2 (Individual timestep bound)**.** *For any fixed $t_0 \in \{0, \dots, T-1\}$:*

$$d^2_{k_{t_0}}(h^*_{t_0}, \hat{h}_{t_0}) \le T \cdot [\text{RHS of Eq. (34)}].$$

The dependence on T is at most linear, in sharp contrast to the $O(\kappa^{T-1})$ scaling of Theorem F.1.

Remark G.3 (From discrepancy to bridge-function L^2 error). Theorem G.1 controls the joint moment-condition discrepancies $d^2_{k_t}(\eta^*, \hat{\eta}) = \mathcal{R}_{k_t}(\hat{\eta})$ rather than per-stage L^2 errors $\|h^*_t - \hat{h}_t\|_{L^2}$. Translating the former to the latter requires an ill-posedness parameter τ_t analogous to single-step NMMR (Kompa et al., 2022): under uniform ill-posedness $\tau_t \le \tau$,

$$\|h^*_t - \hat{h}_t\|_{L^2} \;\le\; \tau \cdot \sqrt{d^2_{k_t}(\eta^*, \hat{\eta}) + d^2_{k_{t-1}}(\eta^*, \hat{\eta})},$$

where the right-hand side aggregates discrepancy across the two stages that together pin down h^*_t through the recursion in Eq. (1). Averaging across t and applying Theorem G.1 yields

$$\frac{1}{T}\sum_{t=0}^{T-1} \|h^*_t - \hat{h}_t\|^2_{L^2} \;=\; O\big(\tau^2 \sqrt{T/n}\big).$$

The crucial point is that τ enters as a uniform constant rather than as a product $\prod_t \tau_t$. Sequential estimation cannot avoid the product structure; joint MMR replaces it with a uniform factor, preserving the $O(\sqrt{T/n})$ scaling that is the substance of Finding 2.

Corollary G.4 (ATE error and consistency)**.** *Let the conditions of Theorem G.1 hold, let the ill-posedness parameters be uniformly bounded, $\tau_t \le \tau$ for all t (Remark G.3), and suppose the approximation and regularization terms in Eq. (34) are negligible. Then, with probability at least $1-\delta$,*

$$\big|\hat{\theta}^a_{\mathrm{POR}} - \theta^a\big| \;\le\; C\,\tau\,(T/n)^{1/4}\sqrt{T} \;+\; 2\,\mathcal{R}_n(\mathcal{H}_0^{\mathrm{act}}) \;+\; M\sqrt{\tfrac{\log(2/\delta)}{2n}}, \tag{35}$$

where C is the constant of Remark G.3 and M bounds the hypothesis class (Condition 2). The last two terms are $O(n^{-1/2})$ by Condition 6. Hence, for any fixed horizon T, $\hat{\theta}^a_{\mathrm{POR}} \to \theta^a$ in probability at rate $O(n^{-1/4})$. The analogous bound holds for $\hat{\theta}^a_{\mathrm{IPW}}$ through the treatment bridge when the estimated weights $\hat{Q}_{T-1}$ are censored at a fixed level c, with C replaced by a constant depending on c and M_Y; and $\hat{\theta}^a_{\mathrm{PDR}}$ is consistent whenever either bridge family is, by Appendix E.

G.1 PROOF OF THEOREM G.1

The proof extends the argument of Kompa et al. (2022), Theorem 1, from a single MMR loss to an aggregate loss over T timesteps.

Step 1: Optimality of $\hat{\eta}$. By definition of the joint estimator:

$$\frac{1}{T}\sum_{t=0}^{T-1}\hat{\mathcal{R}}_{k_t,V,n}(\hat{\eta})+\lambda\Omega[\hat{\eta}]\le\frac{1}{T}\sum_{t=0}^{T-1}\hat{\mathcal{R}}_{k_t,V,n}(\eta)+\lambda\Omega[\eta]$$

for all $\eta\in\mathcal{H}$. In particular this holds for $\tilde{\eta}$, the parameters achieving the best approximation within $\mathcal{H}$.

Step 2: Relating V-statistic to population discrepancy. For each fixed η and timestep t, the V-statistic

$$\hat{\mathcal{R}}_{k_t,V,n}(\eta)=\frac{1}{n^2}\sum_{i,j}r_{t,i}(\eta)\,r_{t,j}(\eta)\,k_{t,ij},$$

with $r_{t,i}(\eta)=\hat{H}^a_{t+1,i}-\hat{H}^A_{t,i}$, is a biased estimator of the population MMR risk $\mathcal{R}_{k_t}(\eta)=\mathbb{E}[r_t(\eta)\,r_t(\eta')\,k_t]$. By the ISPD property of k_t,

$$\mathcal{R}_{k_t}(\eta)=\big\|\mathbb{E}[r_t(\eta)\mid\bar{A}_t,\bar{Z}_t,\bar{X}_t]\big\|^2_{\mathcal{H}_{k_t}}.$$

The moment condition of Eq. (1) gives $\mathbb{E}[r_t(\eta^*)\mid\bar{A}_t,\bar{Z}_t,\bar{X}_t]=0$, so the residual conditional expectation can be rewritten in terms of the parameter difference,

$$\mathbb{E}[r_t(\eta)\mid\cdot]=\mathbb{E}[r_t(\eta)-r_t(\eta^*)\mid\cdot],$$

and we obtain the identity

$$\mathcal{R}_{k_t}(\eta)=\big\|\mathbb{E}[r_t(\eta)-r_t(\eta^*)\mid\cdot]\big\|^2_{\mathcal{H}_{k_t}}=d^2_{k_t}(\eta^*,\eta),\tag{36}$$

where the second equality uses the discrepancy definition of Appendix B. Eq. (36) is the longitudinal counterpart of the single-step identity in Kompa et al. (2022); Mastouri et al. (2021): in their setting the moment is linear in one bridge, so the discrepancy reduces to a per-stage semi-metric on bridge functions; here the moment couples consecutive bridges and the discrepancy lives on parameter pairs.

Step 3: Uniform concentration. Following Kompa et al. (2022), Appendix A, for each t, uniformly over $\eta\in\mathcal{H}$, setting $\delta_t=\delta/(2T)$ to apply a union bound over all T timesteps,

$$\big|\hat{\mathcal{R}}_{k_t,V,n}(\eta)-\mathcal{R}_{k_t}(\eta)\big|\;\le\;8M\big(\mathcal{R}_{n-1}(\mathcal{F}'_t)+\mathcal{R}_n(\mathcal{F}'_t)\big)+16M^2M_k\sqrt{\frac{2\log(2/\delta_t)}{n}}+10\sqrt{2\log 2}\,M^2M_k\,n^{-1/2},$$

where $M=M_Y+\sup_\eta\|h^\eta\|_\infty$ and $\mathcal{F}'_t$ is the function class $\{(\bar{a}_t,\bar{w}_t,\bar{x}_t,\bar{z}_t)\mapsto r_t(\eta)\cdot k_t((\bar{a}_t,\bar{z}_t,\bar{x}_t),\cdot):\eta\in\mathcal{H}\}$. Since $r_t(\eta)$ depends on η only through η^{act}_t (Condition 6), the contraction principle gives $\mathcal{R}_n(\mathcal{F}'_t)\le M_k\cdot\mathcal{R}_n(\mathcal{H}^{\mathrm{act}}_t)=O(M_k\,n^{-1/2})$, uniformly in t. This active-parameter bound replaces the loose class-level estimate $\mathcal{R}_n(\mathcal{F}'_t)\le M_k\cdot\mathcal{R}_n(\mathcal{H})=O(M_k\sqrt{T/n})$ that ignores the per-stage parameter sparsity.

Step 4: Combining. Averaging over t and applying the optimality condition:

$$\frac{1}{T}\sum_t\mathcal{R}_{k_t}(\hat{\eta})\le\frac{1}{T}\sum_t\mathcal{R}_{k_t}(\tilde{\eta})+\lambda\,(M_\lambda-\Omega[\hat{\eta}])+\frac{2}{T}\sum_t\Big|\hat{\mathcal{R}}_{k_t,V,n}(\hat{\eta})-\mathcal{R}_{k_t}(\hat{\eta})\Big|.$$

Using $\Omega[\hat{\eta}]\ge 0$, substituting the concentration bounds from Step 3, and applying the identity from Step 2 yields Eq. (34), with $C_1=16M$ and $C_2=16M^2+10\sqrt{2\log 2}\,M^2$. □

Remark on the rate. The estimation error in Eq. (34) comprises two contributions: $C_1\,\mathcal{R}_n(\mathcal{H}^{\mathrm{act}}_t)M_k=O(M_k\,n^{-1/2})$ from the active-parameter Rademacher term, and $C_2\,M_k^2\sqrt{\log(2T/\delta)/n}=O\big(M_k^2\sqrt{\log T/n}\big)$ from the union-bound concentration term. The Rademacher contribution is independent of T because r_t depends only on $p^{\mathrm{act}}=O(1)$ active parameters; the concentration contribution carries the lone T-dependence, through $\log T$. Together with the T-independent regularization term λM_λ, this gives the simplified bound stated in Section 4:

$$\frac{1}{T}\sum_{t=0}^{T-1}d^2_{k_t}(\eta^*,\hat{\eta})\;\le\;\frac{1}{T}\sum_{t=0}^{T-1}d^2_{k_t}(\eta^*,\tilde{\eta})\;+\;\tilde{O}(n^{-1/2}),$$

where $\tilde{O}$ hides logarithmic factors in T. For any individual time step t_0, Corollary G.2 gives $d^2_{k_{t_0}}(\eta^*, \hat{\eta}) = \tilde{O}(T/\sqrt{n})$, which grows polynomially in T (sharper than the $O(T^{3/2}/\sqrt{n})$ implied by a class-level Rademacher bound) and remains in sharp contrast to the $O(\kappa^{T-1} \cdot r_n)$ exponential growth of Theorem F.1.

G.2 Proof of Corollary G.2

Since $d^2_{k_t} \geq 0$ for all t:

$$d^2_{k_{t_0}}(h^*_{t_0}, \hat{h}_{t_0}) \leq \sum_{t=0}^{T-1} d^2_{k_t}(h^*_t, \hat{h}_t) = T \cdot \frac{1}{T} \sum_{t=0}^{T-1} d^2_{k_t}(h^*_t, \hat{h}_t) \leq T \cdot [\text{RHS of Eq. (34)}].$$

The first inequality uses non-negativity; the equality is trivial; the second applies Theorem G.1. □

G.3 Proof of Corollary G.4

Step 1: Decomposition. Write $\mathbb{P}_n$ for the empirical measure over the n trajectories and $\mathbb{P}$ for the population measure. Since $\hat{\theta}^a_{\mathrm{POR}} = \mathbb{P}_n \hat{h}_0$ and $\theta^a = \mathbb{P} h^*_0$ by Theorem A.11,

$$\hat{\theta}^a_{\mathrm{POR}} - \theta^a = (\mathbb{P}_n - \mathbb{P})\hat{h}_0 \; + \; \mathbb{P}(\hat{h}_0 - h^*_0).$$

By Jensen's inequality the second term satisfies $|\mathbb{P}(\hat{h}_0 - h^*_0)| \leq \|\hat{h}_0 - h^*_0\|_{L^2}$.

Step 2: From the $t = 0$ bridge to the per-stage average. Since every term in a sum of squares is non-negative,

$$\|\hat{h}_0 - h^*_0\|^2_{L^2} \; \leq \; \sum_{t=0}^{T-1} \|\hat{h}_t - h^*_t\|^2_{L^2} \; = \; T \cdot \frac{1}{T} \sum_{t=0}^{T-1} \|\hat{h}_t - h^*_t\|^2_{L^2}.$$

This is the source of the $\sqrt{T}$ factor: the $t = 0$ bridge is pinned down by the entire chain of moment conditions, so its error is controlled only through the aggregate.

Step 3: Per-stage average via Remark G.3. Under uniform ill-posedness and negligible approximation and regularization terms, Remark G.3 gives, with probability at least $1 - \delta/2$,

$$\frac{1}{T} \sum_{t=0}^{T-1} \|\hat{h}_t - h^*_t\|^2_{L^2} \; \leq \; C^2 \tau^2 \sqrt{T/n}.$$

Combining with Step 2 and taking square roots, $\|\hat{h}_0 - h^*_0\|_{L^2} \leq C \tau \, (T/n)^{1/4} \sqrt{T}$.

Step 4: Empirical-process term. The class $\{h_0(\cdot; \eta) : \eta \in \mathcal{H}\}$ is uniformly bounded by M (Condition 2) and depends on η only through the active sub-vector η^{act}_0 (Condition 6). By the standard symmetrization and bounded-differences argument, with probability at least $1 - \delta/2$,

$$\sup_{\eta \in \mathcal{H}} \left| (\mathbb{P}_n - \mathbb{P}) h_0(\cdot; \eta) \right| \; \leq \; 2 \, \mathcal{R}_n(\mathcal{H}^{\mathrm{act}}_0) \; + \; M \sqrt{\tfrac{\log(2/\delta)}{2n}},$$

and $\mathcal{R}_n(\mathcal{H}^{\mathrm{act}}_0) = O(n^{-1/2})$ by Condition 6. A union bound over Steps 3 and 4 yields Eq. (35) with probability at least $1 - \delta$.

Step 5: Consistency. For fixed T, the first term in Eq. (35) is $O(n^{-1/4})$ and the remaining terms are $O(n^{-1/2})$, so all vanish as $n \to \infty$.

P-IPW and P-DR. For P-IPW, the Hájek estimator of Eq. (17) is a ratio of sample averages of $Y \cdot \mathbf{1}(\bar{A} = \bar{a}) \cdot \hat{Q}_{T-1}$ and $\mathbf{1}(\bar{A} = \bar{a}) \cdot \hat{Q}_{T-1}$. With weights censored at c, both numerator and denominator are Lipschitz in $\hat{Q}_{T-1}$ with constants depending on c and M_Y, so the same decomposition applies

with $\|\hat{q}_{T-1} - q^*_{T-1}\|_{L^2}$ in place of $\|\hat{h}_0 - h^*_0\|_{L^2}$. Steps 2–4 go through verbatim under the treatment-bridge analogues of Conditions 1–6 (Appendix B, with the roles of W and Z exchanged), the only change being that the chain is anchored at $Q_{-1} = 1$ and propagates forward, so the aggregate controls the $t = T - 1$ bridge. For P-DR, Appendix E shows that consistency of either bridge family implies consistency of $\hat{\theta}^a_{\mathrm{PDR}}$; with censored weights the correction terms are bounded, so the population argument of Appendix E transfers to the empirical estimator by the law of large numbers. □

H Synthetic Data Generation

We simulate n longitudinal trajectories under the proximal causal structure of Figure 1B–C, implementing the proximal data-generating process. At each time step t, an unmeasured confounder $U_t \in \mathbb{R}$ drives three proxy families — treatment-inducing $Z_t \in \mathbb{R}^3$, outcome-inducing $W_t \in \mathbb{R}^2$, and common-cause $X_t \in \mathbb{R}^4$ — which in turn drive binary treatment A_t and continuous outcome Y_t. The lagged outcome Y_{t-1} is appended to X_t to yield $X_t \in \mathbb{R}^5$, reflecting its role as a common cause in the subsequent time step.

Confounder dynamics. U_t follows an AR(1) process,

$$U_t = \phi_u U_{t-1} + \epsilon^U_t, \qquad \phi_u = 0.7, \quad \epsilon^U_t \sim \mathcal{N}(0, 0.1^2), \tag{37}$$

and is clipped to $[-2, 2]$ to prevent runaway trajectories.

Proxy generation. Each proxy family is a linear function of U_t, its own lag, and relevant treatment lags, with additive Gaussian noise:

$$Z_t = \beta^{ZU} U_t + \beta^{ZZ} Z_{t-1} + \beta^{ZA} A_{t-1} + \epsilon^Z_t, \qquad \epsilon^Z_t \sim \mathcal{N}(0, 1.2^2 I_3), \tag{38}$$

$$W_t = \beta^{WU} U_t + \beta^{WW} W_{t-1} + \epsilon^W_t, \qquad \epsilon^W_t \sim \mathcal{N}(0, 0.3^2 I_2), \tag{39}$$

$$X_t = \beta^{XU} U_t + \beta^{XX} X_{t-1} + \beta^{XA} A_{t-1} + \epsilon^X_t, \qquad \epsilon^X_t \sim \mathcal{N}(0, 0.3^2 I_4). \tag{40}$$

Coefficients are drawn once at simulation setup: $\beta^{ZU}, \beta^{WU} \sim \mathcal{N}(0.4\gamma_u, 0.1^2)$, $\beta^{XU} \sim \mathcal{N}(\gamma_u, 0.1^2)$, $\beta^{ZA} \sim \mathcal{N}(0.3, 0.1^2)$, $\beta^{XA} \sim \mathcal{N}(0.4, 0.1^2)$. Autoregressive matrices $\beta^{ZZ}, \beta^{WW}, \beta^{XX}$ are diagonal with entries $\mathcal{N}(0.6 - 0.1i, 0.1^2)$ along the i-th diagonal. All proxies are clipped to $[-2, 2]$. Crucially, W_t has no direct dependence on lagged treatments, and Z_t has no direct dependence on the outcome — preserving the proxy role separation required by Assumptions A.3 and A.4.

Treatment assignment. $A_t \in \{0, 1\}$ is drawn from a Bernoulli distribution with logit depending on U_t, the current proxies, and lagged treatment and outcome:

$$A_t \sim \mathrm{Bernoulli}\Big(\sigma\big(\alpha_u U_t + \alpha^Z Z_t + \alpha^X X_t + \alpha_a(2A_{t-1} - 1) + \alpha_y Y_{t-1}\big)\Big), \tag{41}$$

with $\alpha_u = \gamma_u$, $\alpha^Z \sim \mathcal{N}(0.4, 0.1^2)$, $\alpha^X \sim \mathcal{N}(0.2, 0.1^2)$, $\alpha_a = 2.5$ (symmetric treatment persistence), $\alpha_y = 0.3$ (outcome-dependent treatment selection). The confounder strength is $\gamma_u = 0.5$ throughout.

Outcome generation.

$$Y_t = \theta_u U_t + \theta_a A_t + \theta^{\mathrm{lag}}_a A_{t-1} + \theta^W W_t + \theta^X X_t + \theta_y Y_{t-1} + \epsilon^Y_t, \tag{42}$$

with $\theta_u = 0.3\gamma_u$, $\theta_a = 0.7$ (the target causal effect), $\theta^{\mathrm{lag}}_a = 0.3$, $\theta^W \sim \mathcal{N}(0.5, 0.1^2)$, $\theta^X \sim \mathcal{N}(0.05, 0.1^2)$, $\theta_y = 0.3$, and $\epsilon^Y_t \sim \mathcal{N}(0, 0.3^2)$.

Initialization and dataset assembly. Each trajectory is initialized at $t = 0$ with $U_0, Y_0 \sim \mathcal{N}(0, 0.5^2)$ and $Z_0, W_0, X_0 \sim \mathcal{N}(0, 0.5^2 I)$. Treatment at $t = 0$ is drawn from $\mathrm{Bernoulli}(0.3)$. The trajectory is then propagated for T steps using the equations above. After generation, Y_{t-1} is concatenated to X_t to form the observed common-cause history ($X_t \in \mathbb{R}^5$), and all continuous variables are standardized to zero mean and unit variance across the trajectory dataset. U_t is retained for evaluation but never exposed to any estimation method.

Ground-truth ATE. The ATE is computed by Monte Carlo simulation of counterfactual trajectories under hypothetical interventions $\bar{a} = \bar{1}_T$ (always treated) and $\bar{a} = \bar{0}_T$ (never treated): for each intervention, 1,000 trajectories are generated with A_t set deterministically to the intervention value rather than drawn from the propensity model, and terminal outcomes are averaged:

$$\psi = \mathbb{E}[Y_T^{\bar{1}_T}] - \mathbb{E}[Y_T^{\bar{0}_T}]. \tag{43}$$

Assumption diagnostics. We verify empirically that the generated data satisfies the proximal assumptions through correlation-based diagnostics: partial correlation of Z with Y given (A, U) below 0.10 (Assumption A.4), partial correlation of A with W given U and of Z with W below 0.15 (Assumption A.3), and partial correlation of Z with W given U below 0.20 (Assumption A.5). All diagnostics pass at the settings used in our experiments.

I MIMIC-III SEMI-SYNTHETIC BENCHMARK

We construct a semi-synthetic longitudinal benchmark on ICU patient trajectories from MIMIC-III, using real clinical measurements as covariates and simulating treatments and outcomes on top of them so that ground-truth ATE is available.

Cohort and temporal resolution. We extract patient trajectories through the MIMIC-Extract pipeline (Wang et al., 2020), aggregating vital signs and laboratory measurements into 4-hour time bins using mean pooling. Patients with fewer than $T = 12$ time bins are excluded, and $n = 1{,}000$ trajectories are sampled uniformly at random from the remaining cohort. Missing values are filled by forward-fill followed by backward-fill along the time dimension, following the DeepACE preprocessing protocol (Frauen et al., 2023).

Variable-to-role assignment. Sixteen clinical variables are selected for their interpretability as early-warning, organ-dysfunction, or exploratory measurements. Their assignment to proxy roles follows a two-stage procedure: (i) a data-driven partition of the 16 variables into three groups via MCMC search (detailed below), which identifies the conditioning set $(X \cup U)$ and the two proxy families Z, W; and (ii) a clinical designation of surrogate U variables within the conditioning set. The resulting assignment is:

- **Surrogate unmeasured confounder U_t (2 variables):** Lactate, White blood cell count. These serve as markers of disease severity that plausibly drive other physiological and laboratory measurements; they are used as the DGP's confounder during simulation and are hidden from all estimation methods at training time.
- **Treatment-inducing proxies Z_t (2 variables):** Systolic blood pressure, Oxygen saturation.
- **Outcome-inducing proxies W_t (9 variables):** Heart rate, Glasgow coma scale total, Bicarbonate, Creatinine, Temperature, Hematocrit, Platelets, Glucose, Central venous pressure.
- **Common causes X_t (3 variables):** Respiratory rate, pH, Bilirubin.

All variables are standardized to zero mean and unit variance per feature before use.

Data-driven partition: rationale. Under the proximal structure (Figure 1B), the treatment-inducing proxies Z and outcome-inducing proxies W descend from the unmeasured confounder U through separate pathways (treatment and outcome, respectively), implying conditional independence $Z \perp\!\!\!\perp W \mid X, U$ under Assumptions A.3–A.5. We exploit this structural prediction as a data-driven check on the proxy-role partition: we search for a three-way partition (G_1, G_2, G_3) of the 16 variables that maximizes conditional independence $G_2 \perp\!\!\!\perp G_3 \mid G_1$, then identify G_1 as the conditioning set $(X \cup U)$ and G_2, G_3 as the two proxy families. Within G_1, the surrogate U variables are designated by clinical judgment about disease severity. This splits the proxy-role assignment into a data-driven component (the harder step, now empirically anchored) and a smaller clinical component.

Data-driven partition: MCMC search. The objective is the conditional independence score

$$\mathrm{CI}(G_1, G_2, G_3) = \max_{i \in G_2,\, j \in G_3} \left|\rho_{ij|G_1}\right|, \tag{44}$$

the maximum absolute partial correlation between pairs of variables in G_2 and G_3 after linearly regressing out G_1. Lower CI scores indicate stronger conditional independence. Variables are flattened across time before computing partial correlations, so the score measures stacked cross-sectional independence.

We optimize over valid three-way partitions by simulated-annealing MCMC. Each step proposes either a swap of two variables between groups or a move of one variable from one group to another, and accepts via the Metropolis-Hastings criterion with a geometrically cooling temperature (initial $T = 1.0$, cooling rate 0.995). We constrain each group to at least 2 variables and cap the conditioning set G_1 at 5 variables. The chain runs for 5,000 iterations with random initialization (seed 42).

The acceptance rate is 16.6%, and the best score 0.107 is reached by iteration $\sim$1,500 and remains stable thereafter (Figure 3). Table 2 shows the MCMC-optimal partition alongside the clinical role assignment.

Table 2: MCMC-discovered partition (CI score $= 0.107$) vs. clinical role assignment. The data-driven partition recovers the clinical proxy-role assignment exactly.

Group	MCMC-optimal partition	Clinical role
G_1 (conditioning)	Lactate, White blood cell count, Respiratory rate, pH, Bilirubin	$X \cup U$
G_2 (proxy family A)	Systolic blood pressure, Oxygen saturation	Z (treatment-inducing)
G_3 (proxy family B)	Heart rate, Glasgow coma scale total, Bicarbonate, Creatinine, Temperature, Hematocrit, Platelets, Glucose, Central venous pressure	W (outcome-inducing)

The data-driven partition recovers the clinical Z/W separation: systolic blood pressure and oxygen saturation cluster together in G_2, distinct from the organ-dysfunction and vital-sign measurements in G_3. The conditioning set G_1 contains the clinically identified common causes together with the disease severity markers Lactate and WBC.

Within G_1, we designate Lactate and WBC as surrogate U: both are established markers of systemic disease severity that plausibly drive other physiological measurements but are themselves driven by latent patient state. The remaining G_1 variables serve as measured common causes X.

Limitations of the partition check. This data-driven validation has three caveats. First, the CI score uses linear partial correlations, which capture linear conditional dependence; nonlinear conditional dependence is not penalized. Second, the MCMC searches over group sizes in a constrained range ($|G_1| \in \{3, \ldots, 5\}$, remaining variables split between G_2 and G_3) and does not explore larger conditioning sets. Third, the procedure identifies the partition (G_1, G_2, G_3) but does not by itself distinguish which of G_2 or G_3 plays the role of Z versus W; that designation is made by clinical reasoning about which proxy family is more closely tied to treatment assignment. The combined procedure reduces, but does not eliminate, the clinical-judgment component of the proxy-role assignment.

Treatment and outcome simulation. Binary treatment A_t and continuous outcome Y_t are simulated on top of the real covariate trajectories, using a history-dependent structural form with memory horizon $h = 2$. Treatment is assigned by threshold on a logistic score that combines a dynamic medication-level state with lagged covariate and outcome histories:

$$A_t = \mathbf{1}\big(\sigma(\ell_t^A) > 0.5\big), \tag{45}$$

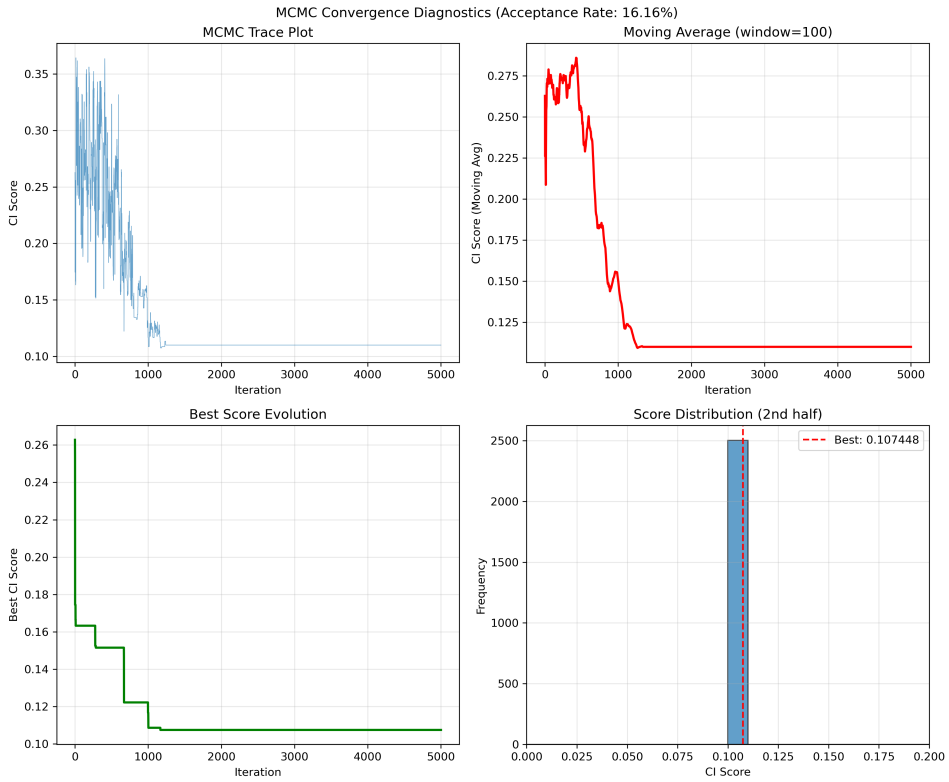

Figure 3: MCMC convergence diagnostics for the three-way partition search on MIMIC-III variables. (Top left) Trace of the conditional independence score across 5,000 iterations; initial exploration narrows after $\sim 1{,}000$ iterations as temperature cools. (Top right) 100-iteration moving average. (Bottom left) Running best score. (Bottom right) Distribution of scores in the second half of the trajectory, concentrated near the converged mode. Acceptance rate: 16.16%.

where

$$\ell_t^A = \epsilon_t^A - \tanh(\ell_t - T/2) + \sum_{i=0}^{\min(h-1,\,t)} c_i\, \bar{m}_{t-i} + \sum_{i=0}^{\min(h-2,\,t-1)} c_i\, \tanh(Y_{t-1-i}/2), \tag{46}$$

with $c_i = (-1)^i/(i+1)$, $\bar{m}_t$ the feature-wise mean of the concatenation (U_t, Z_t, X_t), ℓ_t a patient-specific medication-level state initialized at $T/2-3$ and updated dynamically by treatment decisions, and $\epsilon_t^A \sim \mathcal{N}(0, 0.5^2)$. Outcomes follow a nonlinear lagged-covariate form:

$$\begin{aligned} Y_t^{\text{mimic}} = 5 \sum_{i=0}^{h-1} \alpha_i^Y \; \tanh\big(\sin(\bar{X}_{t-i}^u \cdot t)\, A_{t-i} + \cos(\bar{X}_{t-i}^w \cdot t)\, A_{t-i}\big) \\ +2 \sum_{i=0}^{h-1} \alpha_i^{XY} \; \tanh\big(\bar{X}_{t-i}^u + 0.5\, \bar{X}_{t-i}^w\big) + \varepsilon_t^Y, \end{aligned} \tag{47}$$

with $\bar{m}_t^{UX}$ and $\bar{m}_t^W$ the feature-wise means of (U_t, X_t) and W_t respectively, and $\epsilon_t^Y \sim \mathcal{N}(0, 0.1^2)$. The nonlinearity in the outcome — $\tanh$ compositions with sinusoidal modulation of the covariate-time product — intentionally violates the linearity assumptions of parametric PCI baselines.

Ground-truth ATE. Two counterfactual simulations are run under fixed interventions $\bar{a} = \bar{1}_{12}$ and $\bar{a} = \bar{0}_{12}$ by setting A_t deterministically to the intervention value rather than drawing from the treatment decision rule; outcomes are regenerated under each intervention. The ATE is evaluated at the terminal time step:

$$\psi = \mathbb{E}\Big[Y_T^{\bar{1}_T}\Big] - \mathbb{E}\Big[Y_T^{\bar{0}_T}\Big]. \tag{48}$$

Role of the hidden variables. Lactate and WBC are used during the simulation of A_t and Y_t (they appear in U-containing terms of the structural equations), which makes them functional confounders in the generated data. They are then withheld from all estimation methods at training and evaluation time. The remaining 14 variables — assigned to Z_t, W_t, and X_t — are what each method observes. This construction ensures that the benchmark is well-posed (ground-truth ATE is known) while preserving the structural condition that U is a common cause of observed proxies.

J Hyperparameters and Training Setup

This appendix documents the hyperparameter selection protocol, search spaces, and training settings used for all methods evaluated in Section 5. Implementation is in PyTorch Lightning; hyperparameter optimization uses Optuna (Akiba et al., 2019) with the Tree-structured Parzen Estimator sampler.

J.1 Two-stage Tuning Protocol

For methods with both architectural and causal-regularization hyperparameters (Neural-PCI variants, Sequential NMMR, DTA, TIFM, Causal Transformer, DeepACE), we use a two-stage protocol:

Stage 1 (Architectural). Optuna with 30 trials per intervention sequence searches over architectural and optimization hyperparameters. The objective is factual validation loss — reconstruction loss for latent-variable methods, factual outcome loss for sequential-randomization methods, and the Stage-0 MMR + MSE-hinge loss for Neural-PCI and Sequential NMMR. Causal-regularization parameters are held fixed during Stage 1.

Stage 2 (Causal regularization). With the architecture fixed, we run a small grid search over the causal-regularization parameters of Neural-PCI, selected on held-out trajectories via the P-OR/P-IPW agreement diagnostic

$$\mathcal{D}_{\text{val}}(\eta) := \left|\hat{\theta}^a_{\text{POR}}(\eta) - \hat{\theta}^a_{\text{PIPW}}(\eta)\right|. \tag{49}$$

Under correct specification of either bridge family, both estimators converge to θ^a (Section 2), so disagreement signals mis-specification of one or both bridges. Minimizing $\mathcal{D}_{\text{val}}$ across the grid favors settings where both bridges are jointly well-fit, using only observable quantities. At no point does the selection procedure see θ^a or any held-out counterfactual. Causal regularization weights are selected at representative horizons ($T = 5$ and $T = 10$) and held fixed across all T, preserving the linear-in-T scaling claim of Theorem G.1 and avoiding the asymmetric advantage of per-T tuning.

For Sequential NMMR (Figure 2(a)), only the P-OR estimator is defined; we use factual validation loss for its Stage 2 selection. This choice is conservative for the joint-vs-sequential comparison: factual loss systematically pushes regularization weights toward zero and tends to under-regularize, so any selection-related advantage accrues to Sequential NMMR rather than to Neural-PCI.

For baseline methods (DTA, TIFM, DeepACE, DeepACE-tar, Causal Transformer), causal-regularization parameters are fixed at the values reported in the original publications. Those publications tuned their hyperparameters on synthetic and semi-synthetic MIMIC trajectories with comparable scale and variance characteristics to our experimental setup, so published values transfer without re-tuning. Re-tuning baselines on held-out ATE MAE would itself constitute oracle access and defeat the deployability framing of $\mathcal{D}_{\text{val}}$; re-tuning on factual validation loss is ill-suited to several baselines whose causal regularizers, such as the targeting parameter in DeepACE-tar, are not directly observable through factual fit. Published defaults therefore represent the deployable choice for those methods and preserve parity with Neural-PCI's diagnostic-based selection: both proposed and baseline methods select hyperparameters using only observable quantities, with the specific observable differing across method families based on what each architecture admits.

Why two stages. The factual loss and the agreement diagnostic are both observable but capture different aspects of estimator quality: factual loss measures fit to observed outcomes, while $\mathcal{D}_{\text{val}}$ measures self-consistency across the two semi-parametric estimators that share the same bridge networks. Stage 1 optimizes architectural fidelity to factual data; Stage 2 selects causal regularizers

via a diagnostic derived from the identification structure itself. Both stages use only observable quantities; neither requires counterfactual ground truth.

This separation reflects a structural property of our setting: the factual loss is observed and minimizable; the counterfactual loss (ATE MAE) is unobserved at training time but is what we ultimately care about. Stage 1 optimizes what's observable; Stage 2 protects what's not.

J.2 Search Spaces

Table 3 summarizes the Stage 1 search space and Stage 2 grid for each method. We let p denote the total observed covariate dimension at each time step ($p = \dim Z_t + \dim W_t + \dim X_t$).

Table 3: Hyperparameter search spaces. Stage 1 entries are Optuna search spaces optimized on factual validation loss. Stage 2 entries are manual grid points evaluated on the P-OR/P-IPW agreement diagnostic D_{val} (Eq. 49) for Neural-PCI, and on factual validation loss for Sequential NMMR; for baseline methods, Stage 2 values are fixed at those reported in the original publications.

Method	Stage 1 (architecture)	Stage 2 (causal regularization)
Neural-PCI (P-OR) Neural-PCI (P-IPW)	LSTM hidden, body, head $\in \{p, 2p, 3p, 4p\}$; lr $\sim$ LogU$(10^{-4}, 10^{-3})$; batch $\in \{64, 128\}$; dropout $\in \{0, 0.1, 0.2, 0.3\}$	$\lambda_{\text{mse}} \in \{10, 50\}$; $\lambda_{\text{prop}} \in \{0.2, 0.5, 1.0, 1.5\}$; mse_margin $\in \{0.02, 0.03, 0.05\}$
Neural-PCI (P-DR)	H-branch from P-OR, Q-branch from P-IPW	$\lambda_Q = 0.1$ (fixed); $\lambda_{\text{tar}} = 0.05$ (fixed)
Sequential NMMR	same as Neural-PCI (P-OR)	same as Neural-PCI (P-OR)
P-IPW kernel variants	same as Neural-PCI (P-IPW)	same as Neural-PCI (P-IPW)
DeepACE DeepACE-tar	same architecture grid as Neural-PCI same architecture grid as Neural-PCI	$\alpha = 0$, $\beta = 0$ (fixed); $\alpha = 0.1$, $\beta = 0.05$ (fixed)
DTA	embed-fraction $\in \{0.1, 0.25, 0.5, 0.75\}$; hidden $\in \{p, 2p, 3p, 4p\}$; layers $\in \{1, 2\}$; lr, dropout, batch as above	$\alpha = 0.5$ (fixed); $\theta = 1.0$ (fixed)
Causal Transformer	seq-hidden $\in \{2, 4\} \times \{8, 16, 32\}$; layers $\in \{1, 2, 3\}$; FFN, br-size, fc-hidden multiplicative on seq-hidden; lr, dropout, batch as above	$\alpha = 1.0$, $\beta = 0.05$ (fixed, sigmoid warmup)

J.3 Fixed Settings Across All Methods

- **Optimizer:** Adam, with learning rate as searched and weight decay applied.
- **Training duration:** 100 epochs for Neural-PCI variants, DeepACE, DeepACE-tar, and Causal Transformer; 200 epochs with early stopping (as in the original publications) for DTA and TIFM.
- **Validation split:** 80/20 train/validation split used for both Stage 1 Optuna scoring and early stopping.
- **Repeated runs:** All reported MAE values in Table 1 are means $\pm$ standard deviations over 5 runs with seeds $\{s_0, s_0 + 1, \ldots, s_0 + 4\}$.
- **Optuna trials:** 30 trials per intervention sequence per method. Each method is tuned independently for the always-treated and never-treated interventions.

J.4 Compute Resources

All experiments run on a single GPU per run (NVIDIA GeForce RTX 2080 Ti or GTX 1080 Ti, 11 GB) without mixed-precision training. Approximate wall-clock times for a single training run on synthetic data with $T = 10$, evaluated across two intervention scenarios: 5 minutes for Neural-PCI (P-OR), 8 minutes for Neural-PCI (P-IPW), 10 minutes for Neural-PCI (P-DR), 7 minutes for

Causal Transformer, 6 minutes for DeepACE, 7 minutes for DeepACE-tar, 5 minutes for DTA, and 3 minutes for TIFM.

J.5 Neural-PCI-Specific Settings

Loss-balance heuristic for λ_{mse} and λ_{prop}. The regularization weights λ_{mse} (P-OR) and λ_{prop} (P-IPW) trade off the MMR loss against the MSE hinge and propensity hinge regularizers, respectively. We select values such that the MMR loss and the corresponding regularizer are on roughly equal scales (1:1) at the start of training, based on the observed initial loss magnitudes from a short pilot run. The Stage 2 grid points in Table 3 bracket this balance point.

Soft-hinge sharpness. The β parameter controlling the sharpness of the soft-hinge MSE penalty (Eq. (11)) is fixed at $\beta = 10$ across all experiments. Tuning experiments showed weak sensitivity in the range $\beta \in [5, 20]$.

Dynamic MSE margin. The hinge threshold for the MSE regularizer is set relative to the initial raw MSE loss observed on the first mini-batch of training, rather than as a fixed absolute value, since the scale of the MSE depends on the variance of H_t at initialization. We anchor the threshold at 20% of the observed initial MSE and grid-search the Stage 2 values in Table 3 around this anchor.

RMS normalization for Q. The Q-branch loss applies the normalization $s_t = \sqrt{\mathbb{E}[Q_t^2 + Q_{t-1}^2]}$ (computed batch-wise with `.detach()`) to stabilize the cumulative inverse-propensity scale; see Section 3.

J.6 Reproducibility

Code, hyperparameter configuration files, and seeds will be released upon publication. Synthetic dataset generation (Appendix H) is fully deterministic given the seed; MIMIC-III experiments require access to the MIMIC-III database (PhysioNet credentialed) and the MIMIC-Extract preprocessing pipeline.

K Related Work

Proximal causal inference. The proximal causal inference (PCI) framework treats unmeasured confounding by exploiting noisy proxies of the latent confounder rather than modeling the confounder itself. The intuition originates in the negative-control literature (Lipsitch et al., 2010; Tchetgen Tchetgen, 2014), where outcome and exposure controls are used to detect bias; PCI generalizes this into a constructive identification strategy. Miao et al. (2018) formalized nonparametric identification for point treatments via outcome confounding bridge functions — solutions to Fredholm integral equations of the first kind that absorb the confounder's influence on the outcome through the proxy structure. Cui et al. (2024) introduced treatment confounding bridges, which encode confounding-adjusted inverse-probability weights without modeling the propensity directly, yielding doubly robust estimators in the point-treatment setting. Tchetgen Tchetgen et al. (2024) provided a comprehensive review unifying these identification strategies.

Estimation of bridge functions presents an ill-posed inverse problem: the conditional moment restrictions are typically Fredholm equations whose direct inversion amplifies noise. Mastouri et al. (2021) (PMMR) reformulated these moment restrictions into unconditional Maximum Moment Restriction (MMR) (Muandet et al., 2020) losses with closed-form kernel expressions, avoiding operator inversion entirely. Kompa et al. (2022) (NMMR) extended this to neural estimators using a U-statistic loss formulation, providing the kernel-based machinery that Neural-PCI inherits. Ying et al. (2023) extended PCI to longitudinal settings, establishing nonparametric identification under time-varying proxies and deriving semi-parametric efficient estimators including doubly robust forms. Their framework provides the identification theorems we extend to neural estimation in this paper. Neural-PCI contributes the first joint MMR formulation across T coupled bridge functions, eliminating the exponential error cascade of stage-wise sequential estimation.

Nonparametric instrumental variables and nested moment estimation. Orthogonal to identification strategy, a long line of work studies the estimation problem our setting inherits: solving conditional moment restrictions that define functions only implicitly. Nonparametric instrumental variable (NPIV) regression established the identification and ill-posedness theory for a single such equation (Newey & Powell, 2003; Ai & Chen, 2003), and sieve methods for systems of moment restrictions with different conditioning variables introduced the simultaneous estimation of coupled equations (Ai & Chen, 2007; 2012). A recent literature brings machine learning to NPIV through adversarial objectives (Dikkala et al., 2020; Bennett et al., 2023). The bridge functions of proximal causal inference belong to this family: their defining moment restrictions are nested NPIV equations in which the proxies play the role of instruments. A recent milestone in this line is Meza & Singh (2025), who establish the first explicit mean-square convergence rates and semiparametric inference for nested two-level moment systems over general function classes, covering the two-period proxy setting of Ying et al. (2023). Longitudinal studies, however, involve horizons well beyond two periods, raising questions of a different nature: with T recursively coupled equations per bridge family, the horizon itself becomes the driver of error accumulation, and estimation requires an objective and an architecture that remain trainable and analyzable as T grows. Neural-PCI addresses this gap with a closed-form joint MMR objective enforcing all T moment conditions of both bridge families simultaneously, a recurrent architecture encoding the temporal recursion, and error-scaling guarantees in which T is the explicit parameter.

Parametric longitudinal PCI baselines. The most direct predecessor to Neural-PCI is the parametric longitudinal PCI machinery of Ying et al. (2023). They define a marginal structural mean model (MSMM) parameterized by β for the counterfactual mean, and estimate β using parametric working models for the outcome and treatment confounding bridges — typically linear regressions of H_t on $(\bar{W}_t, \bar{a}_{T-1}, \bar{X}_t)$ and of Q_t on $(\bar{Z}_t, \bar{a}_t, \bar{X}_t)$. The semi-parametric theory yields three estimators with the same robustness structure as ours: outcome-bridge (MSMM-POR), treatment-bridge (MSMM-PIPW), and doubly robust (MSMM-PDR). These estimators inherit the formal identification guarantees of PCI but pay a steep price in flexibility: the linear working models cannot represent nonlinear bridge functions, and mis-specification leads to bias even when the proxy structure is correct. In our experiments, the MSMM family is competitive at short horizons — where the bridge functions are well-approximated by linear forms — but degrades sharply with T, particularly at $T = 20$. Neural-PCI removes the parametric restriction while preserving the identification guarantees and the three-estimator robustness structure. We include MSMM-POR, MSMM-PIPW, and MSMM-PDR as representative baselines from this category.

Sequential randomization baselines. A growing literature applies deep learning to longitudinal treatment effect estimation under the sequential randomization assumption — that conditioning on observed history removes all confounding at each time step. Lim et al. (2018) introduced recurrent marginal structural networks (RMSN), combining inverse probability weighting with LSTM encoders. Bica et al. (2020a) developed counterfactual recurrent networks (CRN) using gradient-reversal-based balancing to learn treatment-invariant representations. Li et al. (2021) proposed G-Net, an iterative G-computation estimator that simulates counterfactual trajectories forward through a sequence model. Melnychuk et al. (2022) extended these ideas using transformer architectures with cross-attention over treatment and covariate sequences. Frauen et al. (2023) introduced DeepACE, an end-to-end model combining iterative G-computation with sequential targeting for doubly robust and asymptotically efficient estimation under sequential randomization; its dual-phase recurrent architecture serves as the encoder we adapt for Neural-PCI. The common assumption across this line of work — no unmeasured confounding given observed history — is empirically untestable and routinely violated in clinical and economic data, where treatment decisions depend on factors such as clinician judgment or unobserved disease severity. When the assumption fails, architectural sophistication does not restore identification: all of these methods produce biased estimates regardless of network capacity. We include DeepACE and the Causal Transformer of Melnychuk et al. (2022) as representative neural baselines from this category in our experiments.

Latent variable model baselines. An alternative line of work addresses unmeasured confounding by modeling hidden confounders as latent variables. Bica et al. (2020b) proposed the Time Series Deconfounder using a factor model that infers substitute confounders from multiple sequential treatments. Kuzmanovic et al. (2021) developed the Deconfounding Temporal Autoencoder (DTA),

which uses noisy proxies and a sequential VAE to recover a latent confounder representation, balanced against an outcome prediction objective. Hatt & Feuerriegel (2024) introduced sequential deconfounding using substitute confounders inferred from treatment patterns. Cheng et al. (2024) proposed the Time-varying Instrumental Factor Model (TIFM), which combines a sequential factor model with instrumental variable identification, requiring observed instruments that affect treatment but not outcome directly. Sun et al. (2025) extended this combination to longitudinal mediation settings. These methods offer flexibility, but their validity depends on strong generative or independence assumptions about the unmeasured confounder that are themselves untestable: factor models require no-single-cause-of-confounder structure, autoencoders require the latent code to recover the true confounder, and IV-based methods require the existence of valid instruments. In the proximal data-generating process used in our experiments, every observed covariate descends from U_t, leaving no pure instrument and violating the factor structure required by these methods — which is reflected in their underperformance in Table 1. Neural-PCI instead grounds identification in the proxy-based machinery of PCI, which does not restrict the confounder's generative structure. We include DTA and TIFM as representative baselines from this category.

L Limitations and Broader Impacts

Limitations. Neural-PCI inherits from proximal inference a reliance on subject-matter judgment: the categorization of observed covariates into proxies (Z, W) and common causes X, together with the completeness conditions, is untestable, and mis-specification can silently violate identification. Our MIMIC-III assignment follows the causal structure of clinical measurement but is not empirically verifiable. Beyond the identification assumptions, the experimental scope is limited to binary treatments, and terminal-outcome ATE; extensions to continuous treatments and time-varying ATEs require non-trivial modifications of both the architecture and the moment conditions. We also observe non-trivial run-to-run variance for P-OR at $T = 10$, indicating that joint MMR estimation, while bounding error growth in expectation, can still exhibit per-trajectory instability at intermediate horizons. Developing sensitivity analyses, partial-identification procedures, and variance-reduction techniques are natural directions for future work.

Broader impacts. Neural-PCI is a methodological contribution to longitudinal causal inference. Potential positive impacts include more reliable treatment-effect estimates in domains where unmeasured confounding is plausible — healthcare, economics, public policy — where standard sequential-randomization methods may produce biased conclusions. Potential negative impacts arise from misapplication: the validity of PCI-based estimates rests on the proxy and completeness assumptions discussed in Appendix A, which cannot be empirically verified. Practitioners deploying these methods in high-stakes domains should treat point estimates as provisional, pair them with sensitivity analyses, and seek external validation. The MIMIC-III experiments in this paper are semi-synthetic and not intended to support clinical decisions.