LEARNING WORLD MODELS WITH CONTINGENT SELF-PREDICTION

Anonymous authors
Paper under double-blind review

ABSTRACT

World models supervised via latent self-prediction naturally prioritize predictable environment dynamics. However, complex settings are rife with predictable distractors, such as periodic motion or moving backgrounds, that are trivial to predict yet uninfluenced by an agent's actions. Crucially, standard non-collapse regularizers cannot eliminate them, as representations saturated with task-irrelevant noise can satisfy distribution constraints at zero prediction loss. To resolve this, we exploit time homogeneity: identifying features over a future planning horizon already determined by present history is equivalent to identifying features of the present that were already foreseeable from an older context. Conditioning on this older context allows us to decompose state predictions into a predetermined component (fixed by past history) and a contingent component (driven by recent actions and observations). We call the resulting approach contingent self-prediction: an objective that enforces non-collapse constraints exclusively on the contingent subspace. Our method is provably invariant to predictable distractors, ensuring state capacity is prioritized for plan-dependent dynamics.

1 INTRODUCTION

World models allow autonomous agents to plan, imagine, and evaluate behaviors inside a learned simulator rather than relying purely on real-world trial and error (Sutton, 1991; Ha & Schmidhuber, 2018). To achieve long-horizon control, modern architectures have increasingly abandoned pixel reconstruction in favor of self-prediction in a learned representation space, whether initialized via pretrained encoders (Zhou et al., 2024) or trained end-to-end (Sobal et al., 2026; Maes et al., 2026). The primary advantage of decoder-free self-prediction is parsimony: rather than expending capacity to reconstruct unforecastable visual noise (Rao et al., 2026), the latent state is encouraged to discard unpredictable elements and retain only predictable environment dynamics.

Self-prediction buys state parsimony by ranking features solely by their forecastability. However, most predictable features of any scene are the ones that either never change or that change like clockwork. Each is forecast without error, there is an unlimited supply of them, and they compete for the same state capacity as features of the task the agent is trying to solve. In this representation race, the task may lose, leading the world model to track instead irrelevant *slow features* (Sobal et al., 2022; Toso et al., 2026). Existing remedies attempt to inject task relevance from the outside, using object masks, bounding boxes (Nam et al., 2026; Zhu et al., 2026), or reward-based equivalences (Toso et al., 2026), defeating the goal of self-supervised representation learning.

In this paper, we resolve this failure mode by shifting focus from *whether* a feature is predictable to *when* it became foreseeable. We make two key contributions. First, we formalize a theoretical explanation for the empirically observed task-blindness of self-predictive world models. We prove that whenever an environment contains deterministic structure, any global optimizer of a standard self-predictive objective can be satisfied by representing that structure alone, and no marginal state distribution constraint (e.g., whitening or Gaussian matching) can rule it out. Second, to break this degeneracy without external supervision, we propose *contingent self-prediction*: a training objective whose state predicts only future features that older observations could not have foreseen, provably reserving state capacity for plan-dependent dynamics.

Our approach rests on a core insight: predictability alone cannot guide representation learning because it ignores when features became foreseeable. For an agent evaluating candidate action se-

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.

quences over a planning horizon, any feature already determined prior to the planning window evaluates to the exact same value regardless of the chosen plan. Consequently, predetermined features yield identical costs across candidate actions, offering no utility for model-based planning. What distinguishes candidate plans is exclusively what the horizon has yet to decide, including features driven by the agent's upcoming actions or the world's remaining stochasticity. A world model should therefore reserve its latent capacity for these contingent features rather than predetermined distractors. Remarkably, isolating this contingent subspace requires no external labels: any predetermined feature of the present was already forecastable from an older snapshot of history.

Building on this principle, we introduce *contingent self-prediction* (CSP). Rather than learning representations that predict the entire future, our model predicts only the residual future that a stale historical snapshot could not have foreseen. Mechanistically, this is achieved by modifying standard non-collapse regularizers: instead of enforcing variance or coverage constraints on the full state distribution, we project out the context-predictable subspace and apply non-collapse constraints strictly to the contingent residual. We prove that this objective is invariant to predetermined distractors, guaranteeing that state capacity is dedicated solely to plan-contingent dynamics.

We validate these theoretical insights on four goal-reaching benchmarks from LeWorldModel Maes et al. (2026), evaluating both standard settings and predictable distractor setups. In clean environments, CSP matches the state-of-the-art latent self-predictive baseline. However, when predictable distractors are introduced, baselines degrade catastrophically, as predicted. Linear R^2 probing reveals the underlying cause: because distractor features are highly predictable and easily satisfy full-state regularizers, standard JEPA latents redirect state capacity toward the distractor at the expense of task dynamics. In contrast, CSP maintains recovers performance and task-variable probe R^2 in distracted environments. Because non-collapse constraints are applied strictly within the contingent subspace, the model experiences zero objective pressure to represent features already determined by older history, leaving latent capacity entirely dedicated to dynamics required for control.

In summary, we make the following contributions:

- **Predictable distractor collapse.** We prove that task-irrelevant predictable distractors can minimize self-predictive objectives employing standard non-collapse regularization.
- **Contingent Self-Prediction (CSP):** We propose an objective that projects out context-predictable features and applies non-collapse regularizers strictly to the contingent residual, provably guaranteeing invariance to predetermined structure.
- **Empirical robustness:** We show CSP matches state-of-the-art JEPA performance in clean settings while improving performance in the presence of predictable distractors.

2 Self-Predictive Collapse Under Predictable Distractors

Latent world models learn a state implicitly, as whatever minimizes a self-prediction loss. This loss can be characterized in terms of a past–future operator. Let $(o_t)_{t\in\mathbb{Z}}$ be a two-sided stationary stochastic process. We anchor the present at $t = 0$ and write the history as $h := o_{\leq 0}$ and the future as $\tau := o_{>0}$, with marginal laws μ and ν and joint law P. Let $\mathcal{H}_{\text{past}} = L^2(\mu)$ and $\mathcal{H}_{\text{future}} = L^2(\nu)$, and denote by $\mathcal{H}^0_{\text{past}}, \mathcal{H}^0_{\text{future}}$ their zero-mean subspaces. The conditional expectation operator

$$T : \mathcal{H}_{\text{future}} \to \mathcal{H}_{\text{past}}, \qquad (Tf)(h) := \mathbb{E}\big[f(\tau) \,\big|\, h\big], \tag{1}$$

provides the best guess of a future feature f given everything so far; we write T_0 for its restriction to $\mathcal{H}^0_{\text{future}}$. For $k \geq 1$, the k-th *singular number* of T_0 is given by the canonical correlation (Lemma 1)

$$\sigma_k = \sup_{\phi,\psi} \Big\{ \operatorname{corr}\big(\phi(h), \psi(\tau)\big) \;:\; \phi \perp \phi_1, \ldots, \phi_{k-1} \Big\}. \tag{2}$$

The *predictive spectrum* $\{\sigma_k\}_{k=1}^K$ is a property of the process which measures how much of the future is linearly recoverable from optimally chosen coordinates of the past. It coincides with the singular values of T_0 whenever the future is not fully determined by the past (Lemma 2).

Self-prediction as an operator decomposition. Self-predictive objectives learn feature maps $\mathbf{\Phi}(h) = (\phi_k)_{k=1}^K$ on the past and $\mathbf{\Psi}(\tau) = (\psi_k)_{k=1}^K$ on the future, maximizing alignment subject to

an anti-collapse constraint. One canonical form is to maximize alignment

$$J(\mathbf{\Phi}, \mathbf{\Psi}) := \sum_{k=1}^{K} \mathbb{E}\big[\phi_k(h)\, \psi_k(\tau)\big] \tag{3}$$

subject to Gaussianity $\mathbf{\Phi}, \mathbf{\Psi} \sim \mathcal{N}(0, \mathbf{I})$ (Balestriero & LeCun, 2025) or whitening $\mathrm{Cov}(\mathbf{\Phi}) = \mathrm{Cov}(\mathbf{\Psi}) = \mathbf{I}_K$. It follows that, if σ_K^2 is an eigenvalue of $T_0 T_0^*$, any orthonormal feature maps that maximize the self-predictive objective in equation 3, must span the left and right singular spaces of T_0, respectively, associated with singular values $> \sigma_K$, i.e. $J \leq \sum_{k=1}^{K} \sigma_k$ (Lemma 3).

If self-prediction estimates the leading singular subspaces of a past–future operator, the implicit hypothesis is that the predictive spectrum $\{\sigma_k\}_{k\geq 1}$ decays and its leading K terms correspond to the most task-relevant coordinates of the past. This is the assumption by LeJEPA (Balestriero & LeCun, 2025): when the true latent state is Gaussian, the singular values decay geometrically, and the state is identifiable from the optimal $\mathbf{\Phi}$ (Klindt et al., 2026). What if we remove this assumption?

Collapse under predictable distractors. Consider observations $o_t = (o_t^P, o_t^S, o_t^U)$ generated by three mutually independent stationary processes:

- **(P) predictable distractor**: $o_t^P = g(\theta_t) + \epsilon_t$, where θ_t is a continuous-valued state evolving deterministically observed through i.i.d. noise ϵ_t, e.g. a moving background or clock;
- **(S) signal**: a noisy stationary process o_t^S. Its strongest past–future correlation is $\rho_1 < 1$, so no feature of the signal is perfectly predictable;
- **(U) unpredictable distractor**: o_t^U i.i.d. across time *e.g.*, a noisy TV.

The predictable distractor (P) induces a maximal past–future dependence (Theorem 1): every function of θ_t is predicted without error and the predictive spectrum contains an infinite block, $\sigma_k = 1$ for all $k \geq 1$. It follows that (Theorem 2):

- A feature is perfectly predictable if and only if it is a function of the distractor state θ_0. Every feature uncorrelated with all of those has predictability bounded by ρ_1^2.
- For every embedding dimension K, any optimal whitened state of the objective in equation 3 is a function of θ_0 alone. Any probe of a signal performs at chance.

Reconstruction behaves as the mirror image of self-prediction: it ranks features by variance rather than predictability, so when a noisy, unpredictable subsystem dominates the observation variance, the optimal K-dimensional linear reconstruction spans directions of that noisy subsystem and is also *task-blind*, while remaining immune to the predictable distractor.

No distributional regularizer can avoid predictable distractors. The negative result above targets the self-predictive objectives with whitening constraints. LeJEPA (Klindt et al., 2026) replaces whitening with full distribution matching, enforcing latent Gaussianity. However, in this scenario it is still theoretically possible to have a state that simultaneously (Theorem 3): (i) follows the desired latent distribution; (ii) only encodes a predictable distractor; (iii) maximizes the self-predictive alignment in equation 3. More generally, any constraint expressible through the joint distribution of $(\phi_1, \ldots, \phi_K)$, with any conditioning pattern among coordinates, is a functional of $\mathrm{law}(\mathbf{\Phi})$ and falls under Theorem 3. The information that distinguishes a genuine K-dimensional state from a disguised 1-dimensional is not necessarily present in the state's distribution.

3 Contingent Self-Prediction

Section 2 attributed the predictable distractor collapse of latent self-prediction objectives to the criteria used to allocate state coordinates, showing that predictable nuisance can fill the state at zero prediction cost. This section derives a self-predictive objective immune to this collapse. The principle is temporal: a clock or a static background is predictable now, but it was equally predictable a minute ago, whereas features that a plan can still change are predictable now because of recent observations and actions.

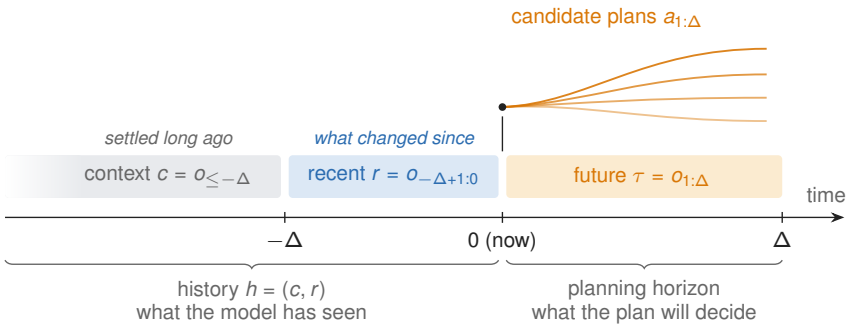


Figure 1: History h is divided at $-\Delta$ into a context c and a recent past r. What r adds to predicting the future, beyond what c already determined, is what candidate plans $a_{1:\Delta}$ can still decide.

3.1 Predetermined and Contingent Features

Fix a planning horizon $\Delta \geq 1$, write $h = o_{\leq 0}$ for the history and $\tau := o_{1:\Delta}$ for the future a planner reasons about. The planner rolls candidate action sequences $a := a_{1:\Delta}$ forward from h and scores the predicted τ. Only differences across candidates can affect its decision, which partitions future features $f(\tau)$ into two kinds. A feature is *predetermined* given h if

$$\mathbb{E}\big[f(\tau) \mid h, a\big] = f(\tau) \quad \text{a.s., for every } a, \tag{4}$$

that is, its value is already a function of h and no action sequence changes it. Every other feature is *contingent*: a plan can influence it or must hedge against it. Predetermined features are identical under every candidate plan. While they may be informative, they cannot rank and encoding them is wasted capacity. Throughout, $\mathbb{E}[\cdot \mid h, a]$ conditions on logged actions, and coincides with the interventional prediction when the behavior policy depends on the history alone and has full support.

Predetermination can be read off the past. Under time-homogeneous dynamics, shifting time back by Δ turns "$f(\tau)$ is predetermined given h" into "$f(h)$ is a function of $o_{\leq -\Delta}$, whatever actions were taken since." In logged data, the two coincide when that policy is exploratory, i.e. its recent actions carry randomness not determined by $o_{\leq -\Delta}$; a controllable quantity driven by a perfectly predictable policy would otherwise appear predetermined. We therefore split the history at $-\Delta$ into a *context* $c := o_{\leq -\Delta}$ and a *recent past* $r := o_{-\Delta+1:0}$ (Fig. 1). The context is what was settled long ago; the recent past and the actions are what the plan adds.

3.2 The Contingent Operator

Every prediction decomposes into what the context already determined and what the recent past and actions add,

$$\underbrace{\mathbb{E}[f(\tau) \mid h, a]}_{Tf} = \underbrace{\mathbb{E}[f(\tau) \mid c]}_{\Pi^{\Delta} T f} + \underbrace{\mathbb{E}[f(\tau) \mid h, a] - \mathbb{E}[f(\tau) \mid c]}_{T^{\Delta} f}, \tag{5}$$

where $\Pi^{\Delta} g = \mathbb{E}[g(h) \mid c]$. We call $T^{\Delta} := (I - \Pi^{\Delta})\, T$ the *contingent operator*: it measures how much of a feature's predictability arrived in the last Δ steps. Its singular numbers are

$$\sigma_k^{\Delta} = \sup_{\phi, \psi} \Big\{ \operatorname{corr}\big(\phi(h, a),\, \psi(\tau) - \mathbb{E}[\psi(\tau) \mid c]\big) \;:\; \phi \perp \phi_1, \ldots, \phi_{k-1} \Big\}, \tag{6}$$

and the maximizing ϕ_k satisfy $\mathbb{E}[\phi_k \mid c] = 0$. Three properties follow. If f is predetermined, Tf is a function of c and $T^{\Delta} f = 0$: the flat block of Section 2 is eliminated. A feature whose Δ-step autocorrelation is ρ^{Δ} keeps a fraction $1 - \rho^{2\Delta}$ of its explained variance, so slowly varying nuisance is discounted in proportion to how slowly it varies. And equation 6 is the alignment objective of equation 3 with the target replaced by its residual given the context: instead of ranking features by predictability, σ_k^{Δ} ranks them by how recently they became predictable, and the top of the spectrum is no longer occupied by what was settled long ago.

3.3 Regularizing the Contingent, Not the Whole State

Write $\mathbf{\Psi}(\tau) := (\psi_k(\tau))_{k\le K}$ for a K-dimensional embedding of the future and $\mathbf{\Phi}(h,a) := (\phi_k(h,a))_{k\le K}$ for the corresponding past features. The singular numbers in equation 6 are canonical correlations with the *contingent* future embedding

$$\underbrace{\mathbf{\Psi}^{\Delta}}_{\text{contingent}} := \mathbf{\Psi} - \underbrace{\mathbb{E}[\mathbf{\Psi} \mid c]}_{\text{predetermined}}, \tag{7}$$

and for them to be correlations, it is $\mathbf{\Psi}^{\Delta}$, not the embedding $\mathbf{\Psi}$, that must have identity covariance. We therefore keep the self-predictive objective $J(\mathbf{\Phi}, \mathbf{\Psi})$ from equation 3 and move the anti-collapse constraint from the whole embedding to its projection on the contingent subspace, imposing

$$\mathbb{E}\big[\mathbf{\Psi}^{\Delta} \mid c\big] = \mathbf{0}, \qquad \mathrm{Cov}\big(\mathbf{\Psi}^{\Delta}\big) = \mathbf{I}_K. \tag{8}$$

The first equality holds by construction, the second is whitening of the residual.

Invariance to predetermined features. The two terms of equation 7 are orthogonal, so every coordinate's variance is a predetermined variance plus a contingent one. We apply anti-collapse regularization the contingent term only. Thus, on the constraint side, the predetermined part may be in the embedding, but it does not count toward the whitening, which every coordinate must satisfy with contingent variance alone (no collapse under predictable distractors). On the alignment side, the predetermined part is costless. Let $\mathbf{\Phi} = \mathbb{E}[\mathbf{\Phi} \mid c] + \mathbf{\Phi}^{\Delta}$, and recall $c \subseteq h$. Then,

$$\mathbb{E}\big\|\mathbf{\Phi} - \mathbf{\Psi}\big\|^2 = \underbrace{\mathbb{E}\big\|\mathbb{E}[\mathbf{\Phi} \mid c] - \mathbb{E}[\mathbf{\Psi} \mid c]\big\|^2}_{\text{predetermined error}} + \underbrace{\mathbb{E}\big\|\mathbf{\Phi}^{\Delta} - \mathbf{\Psi}^{\Delta}\big\|^2}_{\text{contingent error}}. \tag{9}$$

Any $\mathbf{\Phi}$ that reads the context can drive the first term can be driven to zero, so the presence of predetermined structure does not change the objective. The second term is the prediction error of the contingent part, whose minimizers are the singular functions of T^{Δ}. Predetermined features thus cost nothing to predict and are not bound by the constraint.

3.4 State and Predictor

State. The state the model carries is the embedding of the history, obtained by applying the same feature map to the past $\mathbf{\Psi}(h)$. By equation 7, the state splits as $\mathbf{\Psi}(h) = \mathbb{E}[\mathbf{\Psi}(h) \mid c] + \mathbf{s}^{\Delta}$ a context-predetermined term plus a *contingent state* $\mathbf{s}^{\Delta}$: the part of the state the context did not determine, hence the part a plan can still act on. The constraints in equation 8 fix the variance of every coordinate of $\mathbf{s}^{\Delta}$ to one. So $\mathbf{\Psi}(h)$ always carries unit-variance contingent content in each of its K coordinates (no predictable distractor collapse).

Predictor. We parameterize the past feature map $\mathbf{\Phi}$ through a predictor $\mathbf{P}$ that reads the state and the actions and is trained to predict the next embedding $\mathbf{P}\big(\mathbf{\Psi}(h),\, a\big) \approx \mathbf{\Psi}(\tau)$. For a fixed $\mathbf{\Psi}$, the best predictor is $\mathbf{P}^* = \mathbb{E}[\mathbf{\Psi}(\tau) \mid h, a]$, which by the tower property splits as

$$\mathbf{P}^*\big(\mathbf{\Psi}(h), a\big) = \mathbb{E}\big[\mathbf{\Psi}(\tau) \mid c\big] + T^{\Delta}\mathbf{\Psi}^{\Delta}, \qquad T^{\Delta}\mathbf{\Psi}^{\Delta} = \mathbb{E}\big[\mathbf{\Psi}^{\Delta} \mid h, a\big], \tag{10}$$

a copy of the predetermined offset from the context plus the prediction of the contingent part. Substituting $\mathbf{P}^*$ into the prediction error and using $\mathrm{Var}(\psi_k^{\Delta}) = 1$ gives $\mathbb{E}\|\mathbf{\Psi} - \mathbf{P}^*\|^2 = K - \sum_{k\le K} \|T^{\Delta}\psi_k^{\Delta}\|^2$: the predetermined part of $\mathbf{\Psi}$ drops out, and the encoder maximizes $\sum_k \|T^{\Delta}\psi_k^{\Delta}\|^2$ over orthonormal, conditionally centered residuals. This is the CCA problem of equation 6 for T^{Δ} on the contingent space, with $\mathbf{\Phi}^{\Delta} = \mathbf{P} - \mathbb{E}[\mathbf{P} \mid c]$ as the past feature, and it is bounded by the contingent spectrum $J\big(\mathbf{\Phi}^{\Delta}, \mathbf{\Psi}^{\Delta}\big) \le \sum_{k\le K} \sigma_k^{\Delta}$, with equality iff the residuals span the leading subspace of T^{Δ} and $\mathbf{s}$ is sufficient for the action-conditional contingent future.

3.5 Implementation and Planning

The history h consists of a context c made up of M frames ending Δ steps before the present (or the first frames of the episode) and the recent past r consisting of the H most recent frames. Every frame is embedded by the same encoder. The predictor is a causal transformer over $\big[E(c_1), \ldots, E(c_M),\ (E(r_1), a_1), \ldots, (E(r_H), a_H)\big]$ whose output at recent position k predicts $E(r_{k+1})$. Context tokens carry no actions and are never predicted.

Loss. With $\mathbf{\Psi}$ the embedding of the future frame and sg the stop-gradient, the loss reads as

$$\mathcal{L} \;=\; \mathbb{E}\big\|\mathrm{Pred}(E(c), E(r), a) - \mathbf{\Psi}\big\|^2 \;+\; \lambda_1 \mathcal{R}\big(\mathbf{\Psi} - \mathrm{sg}\, m_1(c)\big), \tag{11}$$

where $\mathcal{R}$ is the anti-collapse regularizer enforcing $\mathrm{Cov}\big(\mathbf{\Psi}^{\Delta}\big) = \mathbf{I}_K$. Our implementation uses SIGReg. The regressor m_1 estimates $\mathbb{E}[\mathbf{\Psi} \mid c]$. Setting $m_1 \equiv 0$ and dropping the context recovers LeWorldModel (Maes et al., 2026) exactly.

Contingent future target. The network m_1 maps the raw context to $\mathbb{R}^K$ and is fit by least squares to $\mathrm{sg}\,\mathbf{\Psi}$. It reads raw frames rather than embeddings: if it saw only $E(c)$, the encoder could hide predetermined content in a form the lagged embeddings do not reveal.

Planning. Rollouts are those of the LeWorldModel with a prefix. The context tokens are computed once and held fixed over the horizon, since $c_{t+k} = o_{\le t+k-\Delta}$ consists of frames already observed for $k \le \Delta$; the predictor's raw output is appended as the next recent token together with the candidate action; and the cost is the distance between the rolled embedding and the embedding of the goal frame. Predetermined content enters both terms with the same value and cancels. Neither regressor is used at inference, and the planning horizon is bounded by the training Δ.

4 Experiments

Our experiments answer four questions. (Q1) Does the collapse predicted in §2 occur in JEPA-based world models when the environment contains predetermined structure? (Q2) Does contingent self-prediction prevent it? (Q3) Is the mechanism the one the theory identifies, i.e. the embedding's variance being consumed by what the past already determined, rather than a perception problem? (Q4) Which components of the method are responsible?

4.1 Setup

Environments and data. We use the four environments and the released datasets of LeWorldModel (Maes et al., 2026): **TwoRoom** (2D navigation through doorways), **PushT** (planar manipulation of a T-shaped block), **Cube** (single-cube manipulation from OGBench), and **Reacher** (DeepMind Control). All four are goal-conditioned: the planner receives a goal image and must reach the corresponding state. Datasets, frame resolution, frame skip and history length are those of LeWM.

Predetermined distractors. To test the predictable collapse failure from §2, we introduce per-episode textured backgrounds in the environments above. At the start of every episode, a texture is drawn at random from a bank of procedurally generated patterns and composited onto the frame background for the whole episode, in the training data and in the evaluation environment. The goal image of an episode carries that episode's texture. The texture is constant within an episode and therefore *predetermined*: it is fully predictable from any past frame, and it is high-dimensional. Thus, by Theorem 3 it can fill a state at zero prediction cost. It is also irrelevant since dynamics, goals and success criteria are unchanged, and the task remains visible (Fig. 2).

Models. We compare against two latent world models. PLDM (Sobal et al., 2026), a JEPA trained with VICReg-style anti-collapse terms, and LeWorldModel, which uses a ViT frame encoder, an autoregressive transformer predictor over frame embeddings and actions, and SIGReg as the anti-collapse regularizer. CSP (ours) uses the same networks and planner. The only changes are (i) encoded frames from the context window are added as prefix tokens to the predictor and (ii) the regularizer is applied to the residual $\mathbf{\Psi} - \mathrm{sg}\, m_1(c)$, as described in §3.5. The regressor $m_1(c)$ is a separate encoder on the raw context frames. The context is set to the first M frames of the episode.

Planning and metrics. All models are evaluated with the same model-predictive control loop (CEM, LeWM's planning configuration) and the same episode budget, on held-out start/goal pairs. We report success rate (SR) over N episodes. Representation quality is measured by linear-probe R^2 of ground-truth task variables from the embedding on held-out data.

Table 1: **Planning success rate SR (↑) on clean and textured environments**. Retention (ret.) = $\mathrm{SR}_{\mathrm{tex}}/\mathrm{SR}_{\mathrm{clean}}$ (↑).

	TwoRoom			PushT			Cube			Reacher		
Method	clean	text.	ret.	clean	text.	ret.	clean	text.	ret.	clean	text.	ret.
PLDM	**96.0**	32.7	34.1	78.0	2.7	3.5	66.7	56.0	84.0	62.7	8.7	13.9
LeWM	88.7	38.7	43.6	**93.3**	2.0	2.1	70.7	42.7	60.4	61.3	6.7	10.9
CSP (ours)	94.7	**85.3**	**90.1**	78.7	**3.3**	**4.2**	**83.3**	**71.3**	**85.6**	**66.7**	**11.3**	**16.9**

Table 2: **Agent-position linear-probe R^2 (↑) on held-out episodes.** Probes are fit on training episodes. Negative values mean the probe transfers worse than predicting the mean.

	TwoRoom		PushT		Cube		Reacher	
Method	clean	text.	clean	text.	clean	text.	clean	text.
PLDM	**0.99**	-5.41	0.20	-7.43	0.97	-2.19	**0.99**	-8.99
LeWM	**0.99**	-0.63	0.94	-0.43	**0.98**	-0.50	**0.99**	-0.90
CSP (ours)	**0.99**	**0.99**	**0.98**	**0.85**	**0.98**	**0.97**	**0.99**	**0.46**

4.2 Results

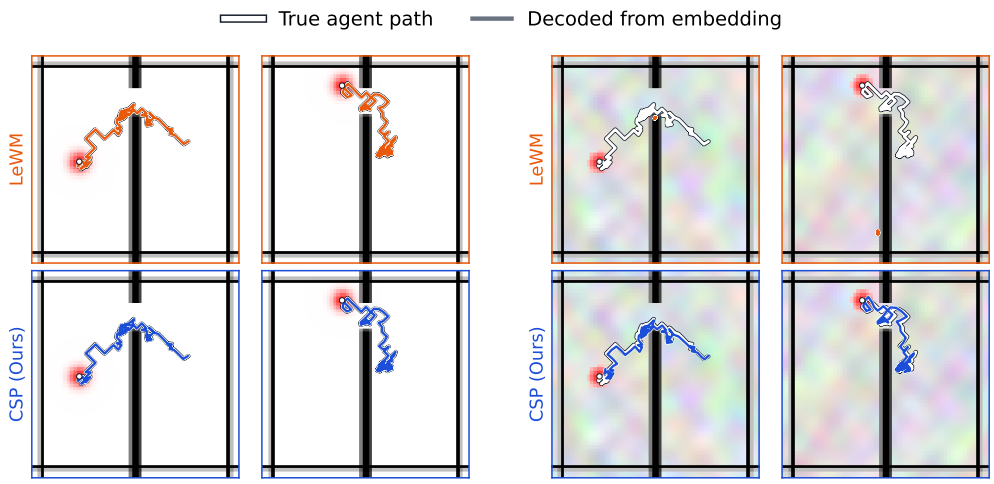
Clean environments. On clean environments (Table 1) CSP is within two points of PLDM on TwoRoom (94.7 vs. 96.0) and above LeWM (88.7), and it exceeds both on Cube (83.3 vs. 70.7 and 66.7) and Reacher (66.7 vs. 61.3 and 62.7). On PushT it matches PLDM (78.7 vs. 78.0) but trails LeWM (93.3). Parity is what the theory predicts: when no predetermined factor is large enough to fill the state, the marginal and residual constraints select the same features and the objectives differ only in the argument of the regularizer.

Textured environments. A static per-episode background texture degrades every method, although the texture is faint and the task is as visible as before. On TwoRoom, LeWM retains 43.6% of its clean success rate and PLDM 34.1%, while CSP retains 90.1%. On Cube, CSP reaches 71.3% against 56.0% for PLDM and 42.7% for LeWM; PLDM's retention (84.0%) is close to CSP's (85.6%) because its clean score is lower. This is the predictable-distractor collapse of Section 2: a per-episode texture is a predetermined factor of dimension far above the embedding size, so under a regularizer on the marginal law of the embedding it fills the state at zero prediction cost and displaces the contingent variables. CSP regularizes the residual of the embedding given the context; the texture, which the context predicts exactly, contributes nothing to that residual and cannot fill the state. On textured PushT and Reacher all three methods fail ($\leq 11.3\%$), with differences within noise, so these variants do not discriminate between them.

The learned state under predictable nuisance. Table 2 reports linear probes of the agent position, fit on training episodes and evaluated on held-out ones. On clean data all methods recover the agent almost perfectly, except PLDM on PushT (0.20). With texture, the LeWM and PLDM probes become negative on every environment: a probe fit on training episodes transfers worse than predicting the mean, because the embedding encodes the episode's texture rather than the agent. CSP retains the agent on all four, from 0.99 on TwoRoom and 0.97 on Cube to 0.85 on PushT and 0.46 on Reacher, including the two environments where it nonetheless fails to plan. Fig. 2 shows the same effect on individual trajectories: on textured TwoRoom the agent trajectory decoded from LeWM's embedding collapses to a point, whereas CSP's follows the ground truth as closely as on clean data.

4.3 Ablation: Where the Fix Lives

Our method differs from a standard self-predictive world model in two places: the predictor receives context tokens, and the regularizer is applied to the context residual of the target rather than to the target itself. Table 3 adds these one at a time on TwoRoom, in the clean version and in the version with a random background texture per episode. We report the held-out linear-probe R^2 of the agent's



(a) Clean TwoRoom (b) TwoRoom with static per-episode texture

Figure 2: **Decoded agent trajectories on held-out TwoRoom episodes.** Top: LeWM; bottom: CSP. Ground-truth trajectories and trajectories decoded from the embedding by an MLP probe. **(a)** Clean: both models track the agent. **(b)** Per-episode texture: the texture is perfectly predictable, fills LeWM's embedding, and its decoded trajectory collapses to a point. CSP is unaffected.

position from the embedding; a negative value means the probe fits worse than predicting the mean, i.e. the embedding carries no information about the agent that transfers across episodes.

Table 3: **TwoRoom ablation.** Linear-probe R^2 of agent position on held-out episodes. The three rows share encoder, predictor, and every hyperparameter.

	Clean	**Textured**
Self-prediction (LeWM)	0.99	-0.63
+ context tokens	0.99	-0.20
+ contingent regularization (CSP)	0.99	**0.99**

On a task with no predetermined structure to compete with, the three models are indistinguishable, which is the guarantee that the modification costs nothing where it is not needed. A per-episode texture is constant within an episode and high-dimensional, so it fills the whitening budget at zero prediction cost; self-prediction collapses onto it and the agent is gone. Giving the predictor context tokens does not help, and this is the point of the middle row: the predictor can now copy the texture from the context more easily, but the regularizer still counts predictable variance, so the budget is still spent on it. Only when the regularizer is applied to $\mathbf{\Psi} - m_1(c)$, so that the texture contributes nothing to the budget and every coordinate has to carry unit variance the context cannot explain, does the agent return, at the same R^2 as on the clean task.

4.4 Discussion

Two points deserve emphasis. First, per-episode textures are a stand-in for any high-dimensional, episode-constant factor, which is the normal situation in video: scene layout, lighting, camera pose, object identity. That LeWM's success rate falls by more than half on a task whose difficulty has not changed suggests that the mechanism of §2 is active whenever the data has more static structure than the embedding has dimensions. Second, the fix costs nothing where it is not needed: on clean environments CSP and LeWM are similar, because the residual and the embedding coincide when there is nothing competing for state capacity.

Limitations. The context lag Δ bounds the planning horizon and must be chosen per environment; a lag distribution at training time relaxes this. The conditional-variance constraint forbids honest

context-dependent variability of the residual, which does not arise in our environments but may in manipulation with sparse contacts; a form of the constraint relative to what the predictor reads from the context is a natural extension. Finally, "predetermined" is measured under the behavior policy: a controllable quantity driven by a perfectly predictable policy is treated as predetermined, so the method assumes the exploratory data that any planner-training pipeline already needs.

5 Related Work

Self-predictive world models and what they learn. Learning a state by predicting its own future embedding goes back to self-predictive reinforcement learning (Schwarzer et al., 2020; Tang et al., 2023; Ni et al., 2024) and to joint-embedding predictive architectures (LeCun et al., 2022; Assran et al., 2023; Bardes et al., 2024). It now underlies planning world models built on pretrained (Zhou et al., 2024) or end-to-end encoders (Sobal et al., 2026; Maes et al., 2026), alongside latent-dynamics models trained with reconstruction or reward (Hafner et al., 2019; 2023; Hansen et al., 2024). Collapse is prevented by constraints on the law of the embedding: variance–covariance penalties (Zbontar et al., 2021; Bardes et al., 2021), whitening (Ermolov & Sebe, 2020), or isotropic-Gaussian matching (Balestriero & LeCun, 2025). Under such constraints, self-prediction recovers the most predictable coordinates of the past, which are linearly identifiable in Gaussian worlds (Klindt et al., 2026). It thus ranks features by a correlation-based form of predictive information. We show that this ranking, combined with any constraint on the embedding's law, is attracted to predetermined structure. We replace it with the part of predictive information that arrived recently.

Distractors and task relevance. Reconstruction-based world models spend capacity on task-irrelevant visual variation. This has motivated separating task latents from distractor latents (Fu et al., 2021; Wang et al., 2022), learning bisimulation metrics from rewards (Zhang et al., 2020), and identifying states free of exogenous noise with multi-step inverse models (Efroni et al., 2021). For self-predictive models the opposite failure, a preference for slow and static features, was identified by Sobal et al. (2022). It has since been addressed with object masks (Nam et al., 2026), physically informed auxiliary targets (Zhu et al., 2026), reward-based state equivalences (Toso et al., 2026), inverse-dynamics regularization (Ivashkov et al., 2026; Zhang et al., 2026), and, concurrently with this work, a regularizer on temporal-difference embeddings aimed at static background distractors (Karmann et al., 2026). These methods add supervision or auxiliary objectives chosen to favor the relevant variables; inverse models in particular favor what the agent controls. CSP instead changes only the argument of the anti-collapse constraint, as derived from the temporal structure of the data. It uses no signal beyond the observations and actions a JEPA already consumes, and it retains every contingent variable, whether controlled or not.

6 Conclusion

Self-predictive world models learn a state by asking it to be predictable, and we showed that this principle alone selects the wrong structure. Environments with deterministic structure have a flat block of perfectly predictable features, and self-prediction fills its state from it first. The failure lies in the anti-collapse constraint: a predetermined factor can satisfy any constraint on the law of the state, isotropic Gaussian regularization included, while carrying no information about the task. Predetermined features are recognizable by their timing, since they were already predictable long ago. The contingent operator ranks features by how recently they became predictable, and it yields a single change to a standard JEPA: regularize the residual of the embedding given the context, not the embedding itself. The experiments confirm this account. CSP matches LeWM where predetermined structure is scarce, and under a static per-episode texture, which drives LeWM's agent probes below chance and halves its planning success, CSP keeps the agent in its state and can retain its success. What the representation budget is spent on is decided not only by what is predictable, but by when it became so.

AI use statement

In this work, generative AI tools were used for assistance with paper drafting and readability editing, LaTeX formatting, writing visualization and boilerplate scripts, and supporting literature discovery.

We have not used generative AI for hypothesis formulation, dataset creation, or result interpretation. All AI-assisted work was carefully reviewed: generated code were tested for correctness, identified literature was manually verified, and all text was revised by the authors. We take full responsibility for the final content of this work.

Reproducibility statement

Implementation of CSP, training and planning is described in §3.5. Additional training and planning details are provided in Appendix A. Proofs of the theoretical claims are provided in Appendix B.

References

Mahmoud Assran, Quentin Duval, Ishan Misra, Piotr Bojanowski, Pascal Vincent, Michael Rabbat, Yann LeCun, and Nicolas Ballas. Self-supervised learning from images with a joint-embedding predictive architecture. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 15619–15629. IEEE, 2023.

Randall Balestriero and Yann LeCun. Lejepa: Provable and scalable self-supervised learning without the heuristics. *arXiv preprint arXiv:2511.08544*, 2025.

Adrien Bardes, Jean Ponce, and Yann LeCun. Vicreg: Variance-invariance-covariance regularization for self-supervised learning. *arXiv preprint arXiv:2105.04906*, 2021.

Adrien Bardes, Quentin Garrido, Jean Ponce, Xinlei Chen, Michael Rabbat, Yann LeCun, Mahmoud Assran, and Nicolas Ballas. Revisiting feature prediction for learning visual representations from video. *arXiv preprint arXiv:2404.08471*, 2024.

Yonathan Efroni, Dipendra Misra, Akshay Krishnamurthy, Alekh Agarwal, and John Langford. Provably filtering exogenous distractors using multistep inverse dynamics. In *International Conference on Learning Representations*, 2021.

Aleksandr Ermolov and Nicu Sebe. Latent world models for intrinsically motivated exploration. *Advances in Neural Information Processing Systems*, 33:5565–5575, 2020.

Xiang Fu, Ge Yang, Pulkit Agrawal, and Tommi Jaakkola. Learning task informed abstractions. In *International Conference on Machine Learning*, pp. 3480–3491. PMLR, 2021.

David Ha and Jürgen Schmidhuber. Recurrent world models facilitate policy evolution. *Advances in neural information processing systems*, 31, 2018.

Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. Dream to control: Learning behaviors by latent imagination. *arXiv preprint arXiv:1912.01603*, 2019.

Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse domains through world models. *arXiv preprint arXiv:2301.04104*, 2023.

Nick Hansen, Hao Su, and Xiaolong Wang. Td-mpc2: Scalable, robust world models for continuous control. In *International Conference on Learning Representations*, volume 2024, pp. 47376–47405, 2024.

Petr Ivashkov, Randall Balestriero, and Bernhard Schölkopf. Sensorimotor world models: Perception for action via inverse dynamics. *arXiv preprint arXiv:2606.20104*, 2026.

Markus Karmann, Shile Li, Christian Internò, Bruno Andreis, David Klindt, Randall Balestriero, Jindong Gu, Philip Torr, Qi Zhang, Peng-Tao Jiang, Hao Zhang, Bo Li, and Onay Urfalioglu. MotionJEPA: Preventing temporal feature collapse by capturing visual changes in latent space. *arXiv preprint arXiv:2609.23881*, 2026.

David Klindt, Yann LeCun, and Randall Balestriero. When does lejepa learn a world model? *arXiv preprint arXiv:2605.26379*, 2026.

Yann LeCun et al. A path towards autonomous machine intelligence version 0.9. 2, 2022-06-27. *Open Review*, 62(1):1–62, 2022.

Lucas Maes, Quentin Le Lidec, Damien Scieur, Yann LeCun, and Randall Balestriero. Leworld-model: Stable end-to-end joint-embedding predictive architecture from pixels. *arXiv preprint arXiv:2603.19312*, 2026.

Heejeong Nam, Quentin Le Lidec, Lucas Maes, Yann LeCun, and Randall Balestriero. Causal-jepa: Learning world models through object-level latent interventions. *arXiv e-prints*, pp. arXiv–2602, 2026.

Tianwei Ni, Benjamin Eysenbach, Erfan Seyedsalehi, Michel Ma, Clement Gehring, Aditya Mahajan, and Pierre-Luc Bacon. Bridging state and history representations: Understanding self-predictive rl. In *International Conference on Learning Representations*, volume 2024, pp. 23555–23569, 2024.

Pratyaksh Rao, Wancong Zhang, Randall Balestriero, Yann LeCun, and Giuseppe Loianno. Skyjepa: Learning long-horizon world models for zero-shot sim-to-real control of quadrotors, 2026. URL https://arxiv.org/abs/2606.23444.

Max Schwarzer, Ankesh Anand, Rishab Goel, R Devon Hjelm, Aaron Courville, and Philip Bachman. Data-efficient reinforcement learning with self-predictive representations. *arXiv preprint arXiv:2007.05929*, 2020.

Uladzislau Sobal, Wancong Zhang, Kyunghyun Cho, Randall Balestriero, Tim GJ Rudner, and Yann LeCun. Learning from reward-free offline data: A case for planning with latent dynamics models. *Advances in Neural Information Processing Systems*, 38:43905–43941, 2026.

Vlad Sobal, Jyothir SV, Siddhartha Jalagam, Nicolas Carion, Kyunghyun Cho, and Yann LeCun. Joint embedding predictive architectures focus on slow features. *arXiv preprint arXiv:2211.10831*, 2022.

Richard S Sutton. Dyna, an integrated architecture for learning, planning, and reacting. *ACM Sigart Bulletin*, 2(4):160–163, 1991.

Yunhao Tang, Zhaohan Daniel Guo, Pierre Harvey Richemond, Bernardo Avila Pires, Yash Chandak, Rémi Munos, Mark Rowland, Mohammad Gheshlaghi Azar, Charline Le Lan, Clare Lyle, et al. Understanding self-predictive learning for reinforcement learning. In *International conference on machine learning*, pp. 33632–33656. PMLR, 2023.

Leonardo F Toso, Davit Shadunts, Yunyang Lu, Nihal Sharma, Donglin Zhan, Nam H Nguyen, and James Anderson. Learning invariant visual representations for planning with joint-embedding predictive world models. *arXiv preprint arXiv:2602.18639*, 2026.

Tongzhou Wang, Simon S Du, Antonio Torralba, Phillip Isola, Amy Zhang, and Yuandong Tian. Denoised mdps: Learning world models better than the world itself. *arXiv preprint arXiv:2206.15477*, 2022.

Jure Zbontar, Li Jing, Ishan Misra, Yann LeCun, and Stéphane Deny. Barlow twins: Self-supervised learning via redundancy reduction. In *International conference on machine learning*, pp. 12310–12320. PMLR, 2021.

Amy Zhang, Rowan McAllister, Roberto Calandra, Yarin Gal, and Sergey Levine. Learning invariant representations for reinforcement learning without reconstruction. *arXiv preprint arXiv:2006.10742*, 2020.

Zhenghao Zhang, Yuanxiang Wang, Zhenyu Guan, Yujia Yang, Bingkang Shi, Tianyu Zong, Hongzhu Yi, Guoqing Chao, Xingchen Chen, Tiankun Yang, et al. Delta-jepa: Learning action-sensitive world models via latent difference decoding. *arXiv preprint arXiv:2606.31232*, 2026.

Gaoyue Zhou, Hengkai Pan, Yann LeCun, and Lerrel Pinto. Dino-wm: World models on pre-trained visual features enable zero-shot planning. *arXiv preprint arXiv:2411.04983*, 2024.

Penghao Zhu, Salvatore Penachio, Kaustav Mukherjee, and Aneesh Jonelagadda. Spectral-target physical latent structuring for jepa-style world models. *arXiv preprint arXiv:2609.04264*, 2026.

A Implementation Details

A.1 Training configurations

We provide training hyparparameters for contingent self-prediction training in Table 4. Both the visual encoder and the action-conditioned predictor are trained from scratch. Context regression uses its own visual network and optimizer.

Table 4: **Training hyperparameters for contingent self-prediction training**.

Configuration	Value
Input image	64×64 RGB
Pixel preprocessing	Area resize; scale to $[0, 1]$
History / context length	3 / 3 frames
Temporal stride / action-block length	5 / 5 simulator steps
Context frames	Episode indices $[0, 5, 10]$
Training anchor	$t \geq 30$
Embedding Dimension	192
Optimizer	Adam
Base learning rate	5×10^{-5}
Adam (β_1, β_2); ϵ	$(0.9, 0.999)$; 10^{-8}
Weight decay	0
Learning-rate schedule	Constant
Global batch size	128
Numerical precision	float32
Gradient clipping	None
Main prediction target	Raw future embeddings
Regularization	SIGReg on future residuals
SIGReg projections / integration knots	512 / 17
SIGReg integration interval	$[0, 3]$

A.2 Planning evaluation

We evaluate goal-conditioned control using model-predictive planning and the cross-entropy method (CEM). A task is initialized from a dataset state, with the goal taken 25 simulator steps later in the same trajectory. Success follows the environment's task-specific success criterion. Table 5 gives the common optimization and interaction budget for the corrected reference-aligned evaluation protocol.

Table 5: **Planning evaluation hyperparameters**. A block contains five individual simulator actions, so five blocks correspond to 25 executed simulator steps.

Configuration	Value
Planner	CEM
Candidate sequences per iteration	300
Optimization iterations / elites	30 / 30
Initial variance scale	1.0
Solver batch size	1
Future planning horizon	5 blocks (25 steps)
Replanning interval	5 blocks (25 executed steps)
Interaction budget	50 simulator steps
Goal offset	25 simulator steps
Cost	Squared terminal latent distance
CSP observed history	$[t-10, t-5, t]$
CSP fixed context	Episode frames $[0, 5, 10]$
Initial anchor	$t \geq 30$
LeWM observation history	Native one-frame planning recipe
Evaluation seeds	$42, 43, 44$
Episodes per seed / total	50 / 150

B THEORETICAL RESULTS

Lemma 1: Singular numbers are canonical correlations

For every $k \geq 1$,

$$\sigma_k = \sup \Big\{ \operatorname{corr}\big(\phi(h), \psi(\tau)\big) \ : \ \phi \perp \phi_1, \ldots, \phi_{k-1} \Big\}, \tag{12}$$

where the supremum is over zero-mean, unit-variance $\phi \in \mathcal{H}^0_{\text{past}}$, $\psi \in \mathcal{H}^0_{\text{future}}$, and $\phi_1, \ldots, \phi_{k-1}$ are optimizers of the preceding problems (when attained).

Proof. For unit-variance zero-mean ϕ, ψ we have

$$\operatorname{corr}(\phi, \psi) = \mathbb{E}[\phi(h)\psi(\tau)] = \langle \phi, T_0 \psi \rangle \tag{13}$$

and

$$\sup_{\|\phi\|=1} \langle \phi, T_0 \psi \rangle = \|T_0 \psi\|. \tag{14}$$

The claim is then the standard variational characterization of the numbers equation 2 applied to $(T_0^* T_0)^{1/2}$, with the deflation over ψ's realized by restricting to the orthogonal complement of previous optimizers. □

Lemma 2: Compactness is a dependence condition

The Hilbert-Schmidt (HS) norm of T is

$$\|T\|^2_{\text{HS}} = \mathbb{E}_{\mu\otimes\nu}[\kappa^2] = 1 + \chi^2(P \,\|\, \mu \otimes \nu) \tag{15}$$

and

$$\|T_0\|^2_{\text{HS}} = \chi^2(P \,\|\, \mu \otimes \nu) = \sum_{k \geq 1} \sigma_k^2. \tag{16}$$

In particular, T_0 is Hilbert–Schmidt (hence compact) if and only if the χ^2-mutual information between past and future is finite.

Proof. T is the integral operator with kernel κ with respect to $\mu \otimes \nu$; its Hilbert–Schmidt norm is the $L^2(\mu \otimes \nu)$ norm of the kernel, and subtracting the rank-one constant block $\mathbf{1} \otimes \mathbf{1}$ (which carries $\sigma = 1$, $\phi_0 \equiv \psi_0 \equiv 1$) yields the zero-mean statement. □

Lemma 3: Whitened predictive objectives estimate the leading singular subspace

Over zero-mean, whitened feature maps:

(i) $\sup_{\Phi,\Psi} J = \sum_{k \leq K} \sigma_k$

(ii) If σ_K^2 is an eigenvalue of $T_0 T_0^*$ (T_0 need not be compact), the optima are attained, and every optimizer Φ spans a subspace of the spectral subspace of $T_0 T_0^*$ associated with eigenvalues $\geq \sigma_K^2$, containing the full spectral subspace of eigenvalues $> \sigma_K^2$.

Proof. (i) Write $J = \operatorname{tr}(A^* T_0 B)$ where $A : \mathbb{R}^K \to \mathcal{H}^0_{\text{past}}$ and $B : \mathbb{R}^K \to \mathcal{H}^0_{\text{future}}$ are the isometries sending the standard basis to (ϕ_k) and (ψ_k). The singular numbers of the $K \times K$ matrix $A^* T_0 B$ are dominated by those of T_0 (singular numbers do not increase under composition with contractions), and $|\operatorname{tr} M| \leq \sum_k \sigma_k(M)$, giving the upper bound $\sum_{k \leq K} \sigma_k$; near-optimal subspaces in equation 2 give the matching lower bound.

(ii) When the top of the spectrum consists of eigenvalues of $T_0 T_0^*$, the trace bound is achieved exactly on the corresponding spectral subspaces, and any optimizer with a component outside them loses strictly by the variational characterization. □

Theorem 1: Deterministic factors flatten the spectrum

Suppose the process admits a factor (θ_t) satisfying:

- **(D)** *Determinism:* $\theta_{t+1} = g(\theta_t)$ for a bi-measurable invertible g preserving the non-atomic stationary law π of θ_0;
- **(P)** *Past identifiability:* θ_0 is $\sigma(h)$-measurable;
- **(F)** *Future identifiability:* θ_0 is $\sigma(\tau)$-measurable.

Then:

(i) T_0 restricts to an isometry on the infinite-dimensional closed subspace

$$\mathcal{F}_\theta := \{\, u(\theta_0(\tau)) : u \in L^2_0(\pi) \,\} \subseteq \mathcal{H}^0_{\text{future}} \tag{17}$$

(ii) $\sigma_k = 1$ for every $k \geq 1$; consequently T_0 is not compact, and 1 belongs to the essential spectrum of $T_0 T_0^*$;

(iii) the past–future dependence is maximal: $\chi^2(P \,\|\, \mu \otimes \nu) = \infty$;

(iv) for every K, the suprema of Lemma 3 are attained, and they are attained by feature maps measurable with respect to $\sigma(\theta_0)$ alone.

Proof. (i) By **(F)** there is a measurable ζ with $\theta_0 = \zeta(\tau)$ a.s.; for $u \in L^2_0(\pi)$ set $f_u := u \circ \zeta \in \mathcal{H}^0_{\text{future}}$. By **(P)**, $\mathbb{E}[u(\theta_0) \mid h] = u(\theta_0)$ a.s., so $T_0 f_u = u(\theta_0)$ viewed as an element of $\mathcal{H}^0_{\text{past}}$ and $\|T_0 f_u\| = \|u\|_{L^2(\pi)} = \|f_u\|$. The subspace $\mathcal{F}_\theta \cong L^2_0(\pi)$ is infinite-dimensional because π is non-atomic.

(ii) For any k, any k-dimensional $V \subseteq \mathcal{F}_\theta$ achieves $\inf_{f \in V, \|f\|=1} \|T_0 f\| = 1$ in equation 2, so $\sigma_k = 1$; a compact operator has $\sigma_k \to 0$. Since every $f_u \in \mathcal{F}_\theta$ satisfies $T_0^* T_0 f_u = f_u$ (isometry), $\mathcal{F}_\theta$ lies in the eigenspace of $T_0^* T_0$ at 1, which is thus infinite-dimensional.

(iii) $\sum_k \sigma_k^2 = \infty$ and lemma 2 give the χ^2 claim.

(iv) The eigenspace of $T_0 T_0^*$ at 1 contains $\{u(\theta_0) : u \in L^2_0(\pi)\}$ on the past side; any orthonormal K-family in it, paired with the corresponding future features $u_k \circ \zeta$, attains $J = K$. □

Consider an environment (E) where observations form a triple $o_t = (o_t^P, o_t^S, o_t^U)$ generated by three mutually independent stationary processes:

- **(P)** **predictable distractor**: $o_t^P = x_C(\theta_t) + \epsilon_t$ where (θ_t) satisfies **(D)**, **(P)**, **(F)** of theorem 1 with non-atomic law π, x_C is measurable, and (ϵ_t) is i.i.d. noise independent of (θ_t);
- **(S)** **signal** (the task-relevant factor): a stationary process whose top zero-mean canonical correlation satisfies $\rho_1 := \sigma_1\big(T_0^S\big) < 1$;
- **(U)** **an unpredictable distractor**: (o_t^U) i.i.d. across time.

We write T^P, T^S, T^U for the per-factor past–future operators. A state Φ is called *task-blind* if $\sigma(\Phi)$ is independent of the entire signal process; a task-blind state scores at chance under every probe, linear or not. For whitened (orthonormal) Φ and a zero-mean unit-variance target u, the linear-probe coefficient is $R^2_{\text{lin}}(u \mid \Phi) = \sum_k \langle u, \phi_k \rangle^2$.

Theorem 2: Distractors vs signal

In environment (E):

(i) *(Sector decomposition.)* The eigenspace of $T_0 T_0^*$ at eigenvalue 1 is exactly $L_0^2(\sigma(\theta_0))$, and every other spectral direction has Rayleigh quotient at most ρ_1^2.

(ii) *(Predictable distractor fills the state.)* For every K, the maximizers of the whitened objectives of lemma 3 exist and are exactly the whitened K-families in $L_0^2(\sigma(\theta_0))$. Every maximizer is task-blind.

(iii) (R^2 *bound*) Let $J_2(\Phi) = \sum_{k=1}^{K} \|\mathbb{E}[\phi_k \mid \tau]\|^2$ be the total predictability of the state's coordinates. If a whitened state Φ satisfies $J_2(\Phi) \geq K - \delta$, with then for every zero-mean unit-variance signal-measurable target u,

$$R^2_{\text{lin}}(u \mid \Phi) \;\leq\; \frac{\delta}{1-\rho_1^2}. \tag{18}$$

Proof. Throughout, for a zero-mean past feature ϕ define its *predictability*

$$\mathcal{E}(\phi) \;:=\; \frac{\big\|\mathbb{E}[\phi(h) \mid \tau]\big\|^2}{\|\phi\|^2} \;=\; \frac{\langle T_0^*\,\phi,\, T_0^*\,\phi\rangle}{\|\phi\|^2} \;=\; \frac{\langle \phi,\, T_0 T_0^*\,\phi\rangle}{\|\phi\|^2} \;\in\; [0,1], \tag{19}$$

the fraction of ϕ's variance recoverable from the future: $\mathcal{E} = 1$ means perfectly predictable, $\mathcal{E} = 0$ means independent of the future in the L^2 sense. The objective $J_2(\Phi) = \sum_k \|\mathbb{E}[\phi_k \mid \tau]\|^2$ is the total predictability of the state's coordinates, and $J_2 \leq K$. The proof computes the predictability of every direction in $\mathcal{H}^0_{\text{past}}$ and reads the claims off the result.

Claim (i). We want the eigenvalues of $T_0 T_0^*$: the predictability of every direction. For a nonempty subset $A \subseteq \{C, S, N\}$, let $\mathcal{H}^{\text{past}}_A$ denote the closed span of products $\phi_A(h) = \prod_{i\in A} \phi_i(h_i)$ in which each ϕ_i is a zero-mean function of the past of process i. Thus A records which processes a feature genuinely depends on: $\mathcal{H}^{\text{past}}_{\{C\}}$ holds functions of the C past alone, $\mathcal{H}^{\text{past}}_{\{C,S\}}$ holds C–S interactions, and so on. Because the three processes are mutually independent, these subspaces are orthogonal and exhaust all zero-mean features:

$$\mathcal{H}^0_{\text{past}} \;=\; \bigoplus_{\emptyset \neq A \subseteq \{C,S,N\}} \mathcal{H}^{\text{past}}_A, \qquad\qquad \mathcal{H}^0_{\text{future}} \;=\; \bigoplus_{\emptyset \neq B \subseteq \{C,S,N\}} \mathcal{H}^{\text{future}}_B. \tag{20}$$

Now we check that T_0 respects this decomposition. For basis features ϕ_A and f_B as above, $\langle \phi_A, T_0 f_B\rangle = \mathbb{E}[\phi_A(h)\, f_B(\tau)]$, and independence of the three processes factorizes it into one term per process:

$$\mathbb{E}\Big[\prod_{i\in A} \phi_i(h_i) \prod_{j\in B} f_j(\tau_j)\Big] \;=\; \prod_{i\in A\cap B} \mathbb{E}\big[\phi_i(h_i)\, f_i(\tau_i)\big] \prod_{i\in A\setminus B} \underbrace{\mathbb{E}[\phi_i]}_{=0} \prod_{j\in B\setminus A} \underbrace{\mathbb{E}[f_j]}_{=0}. \tag{21}$$

Equation 21 vanishes whenever $A \neq B$, so T_0 is block-diagonal,

$$T_0 \;=\; \bigoplus_A T_A, \qquad T_A : \mathcal{H}^{\text{future}}_A \to \mathcal{H}^{\text{past}}_A, \qquad \langle \phi_A, T_A f_A\rangle = \prod_{i\in A} \mathbb{E}\big[\phi_i(h_i) f_i(\tau_i)\big], \tag{22}$$

and, by the diagonal case of equation 21, within a sector correlations simply *multiply* across the participating processes. Two consequences carry the rest of the proof. First, $T_0 T_0^* = \bigoplus_A T_A T_A^*$: eigenvalues can be computed sector by sector and collected. Second, prediction from the future respects sectors, $\mathbb{E}[\phi_A \mid \tau]$ lies in $\mathcal{H}^{\text{future}}_A$, since by equation 21 it is orthogonal to every other sector, so for a general feature $\phi = \sum_A \phi_A$, we have $\|\mathbb{E}[\phi \mid \tau]\|^2 = \sum_A \|\mathbb{E}[\phi_A \mid \tau]\|^2$.

Step 2 (the predictability of each sector). We now bound $\mathcal{E}$ on each piece.

- *N directions:* $\mathcal{E} = 0$. Future N observations are i.i.d. and independent of the past, so for any zero-mean f, $\mathbb{E}[f(\tau_N) \mid h] = \mathbb{E}[f] = 0$: the predictive objective assigns N exactly zero value. Any interaction sector containing N inherits the zero factor.

- *C directions:* $\mathcal{E} = 1$. By theorem 1, every function of θ_0 is perfectly predictable (and perfectly recoverable from the future).
- *S directions:* $\mathcal{E} \leq \rho_1^2$, by the definition of ρ_1 as the signal's top past–future canonical correlation; C–S interactions are bounded by $1 \cdot \rho_1^2$ (the C factor contributes correlation at most 1, the S factor at most ρ_1).

Thus, predictability equals 1 exactly on $L_0^2(\sigma(\theta_0))$ and is at most $\rho_1^2 < 1$ on its orthogonal complement. Writing any past feature as $\phi = c + r$ with c the projection onto $L_0^2(\sigma(\theta_0))$, and using the additivity from Step 1,

$$\left\|\mathbb{E}[\phi \mid \tau]\right\|^2 \;=\; \|c\|^2 + \left\|\mathbb{E}[r \mid \tau]\right\|^2 \;\leq\; \|c\|^2 \;+\; \rho_1^2\,\|r\|^2, \tag{23}$$

with equality at value $\|\phi\|^2$ forcing $r = 0$. This proves (i).

Claim (ii). By equation 23, the eigenspace of $T_0 T_0^*$ at eigenvalue 1 is exactly $L_0^2(\sigma(\theta_0))$, which is infinite-dimensional. By lemma 3(ii), optima are attained at every K and every optimizer lies inside this eigenspace: any component $r \neq 0$ loses strictly. Since θ is independent of the signal process, every function of an optimal state is independent of the entire signal process: task-blind under any probe.

Claim (iii). Let Φ be whitened with $J_2(\Phi) \geq K - \delta$ and decompose $\phi_k = c_k + r_k$. Summing equation 23 over k gives $J_2(\Phi) \leq K - (1 - \rho_1^2) \sum_k \|r_k\|^2$, hence the budget $\sum_k \|r_k\|^2 \leq \delta/(1 - \rho_1^2)$: near-optimal states can spend almost nothing outside the clock algebra. Now take any zero-mean, unit-variance signal-measurable target u. By independence, u is orthogonal to every clock-measurable c_k, so $\langle u, \phi_k \rangle = \langle u, r_k \rangle$ and

$$R^2_{\mathrm{lin}}(u \mid \Phi) = \sum_k \langle u, \phi_k \rangle^2 = \sum_k \langle u, r_k \rangle^2 \leq \sum_k \|r_k\|^2 \leq \frac{\delta}{1 - \rho_1^2}. \tag{24}$$

□

Theorem 3: Distribution-based anti-collapse are blind to predictable distractor collapse

In Environment (E), fix any $K \geq 1$ and any Borel probability measure Q on $\mathbb{R}^K$. Then there exists a past feature map $\Phi^* : h \mapsto \mathbb{R}^K$, measurable with respect to $\sigma(\theta_0)$ alone, such that:

(i) $\mathrm{law}(\Phi^*) = Q$ *exactly*;

(ii) Φ^* is fully predictable: there is a future feature map Ψ^* with $\Psi^*(\tau) = \Phi^*(h)$ a.s., so the objective is at its global maximum (all correlations 1).

(iii) Φ^* is task-blind.

Proof. Since π is a non-atomic law on a standard Borel space, the isomorphism theorem for Lebesgue spaces provides a mod-0 measure isomorphism Λ from $(\mathrm{supp}\,\pi, \pi)$ to $([0,1], \mathrm{Leb})$. Set $U := \Lambda(\theta_0) \sim \mathrm{Unif}[0,1]$ and let $(b_j)_{j \geq 1}$ be its binary digits, i.i.d. $\mathrm{Ber}(1/2)$. Interleaving the digits into K subsequences yields independent $U_1, \ldots, U_K \sim \mathrm{Unif}[0,1]$, all deterministic measurable functions of θ_0. Let $G : [0,1]^K \to \mathbb{R}^K$ be the conditional-quantile map transporting $\mathrm{Unif}[0,1]^K$ to Q. Define $\Phi^* := G(U_1, \ldots, U_K)$, evaluated through the $\sigma(h)$-measurable version of θ_0 provided by **(P)**, and Ψ^* identically through the $\sigma(\tau)$-measurable version provided by **(F)**. Then (i) holds by construction, (ii) because both maps evaluate the same function of the same random variable θ_0, and (iii) because $\sigma(\Phi^*) \subseteq \sigma(\theta_0)$, which is independent of the signal process. □

Lemma 4: Contingency measures recent predictive gain

For every $f \in \mathcal{H}^0_{\text{future}}$,

$$\|T^{\Delta} f\|^2 \;=\; \|Tf\|^2 - \|\Pi_{\Delta} T f\|^2 \;=\; \big\| \, \mathbb{E}[f(\tau) \mid h] - \mathbb{E}[f(\tau) \mid c] \, \big\|^2, \tag{25}$$

so $\sigma_k^{\Delta} \le \sigma_k$, and

$$\begin{aligned} \sigma_k^{\Delta} \;&=\; \sup \Big\{ \operatorname{corr}\big(\phi(h), \psi(\tau)\big) \;:\; \mathbb{E}[\phi \mid c] = 0,\; \phi \perp \phi_1, \ldots, \phi_{k-1} \Big\} \\ &=\; \sup \Big\{ \operatorname{corr}\big(\phi(h),\, \psi(\tau) - \mathbb{E}[\psi(\tau) \mid c]\big) \;:\; \phi \perp \phi_1, \ldots, \phi_{k-1} \Big\}. \end{aligned} \tag{26}$$

Proof. $\Pi_{\Delta} T f = \mathbb{E}[\mathbb{E}[f \mid h] \mid c]$ by the tower property, since $\sigma(c) \subseteq \sigma(h)$, and Π_{Δ} is an orthogonal projection, giving equation 25.

For equation 26, $\langle \phi, Tf \rangle = \langle \phi, T^{\Delta} f \rangle$ whenever $\mathbb{E}[\phi \mid c] = 0$, and $\langle \phi, T(f - \Pi'_{\Delta} f) \rangle = \langle \phi, Tf - \Pi_{\Delta} T f \rangle$ for any ϕ, where $\Pi'_{\Delta} f := \mathbb{E}[f(\tau) \mid c]$; both suprema therefore equal $\|T^{\Delta} f\|$ maximized over unit f in the appropriate orthogonal complement. □