MARGINAL NORMALITY DOES NOT CONSTRAIN TIME: A PRESCRIBED GAUSSIAN PATH LAW FOR JEPA WORLD MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

SIGReg, the anti-collapse regulariser of the LeWM world model, constrains the latent marginal at each time step and says nothing about how consecutive steps relate: a latent that is constant along every trajectory attains the global minimum of both training terms, and the identifiability theory for this objective assumes exactly the cross-time structure it never enforces (Klindt et al., 2026). We close the gap with **T-SIGReg** (temporal SIGReg), which prescribes the joint law of a trajectory window, $\mathcal{N}(0, R_{\rho,h} \otimes I_d)$, with the profile $R_{\rho,h}$ measured from the observation stream before training, and enforces it with the characteristic-function test SIGReg already uses, pointed at a *separable sketch* $q \otimes v$ that keeps the cross-time signal a full-sphere draw averages away, at cost $O((h{+}1)d)$. Which measure supplies q is an experimental-design variable: the temporally optimal design gives up the spatial law, and mixing the marginal directions back in restores it at a cost linear in the mixing weight. On four pixel-based control suites under a paired protocol, prescribing the measured path law raises planning success on TwoRooms from 63.7 to 100.0, improves Reacher and Cube by more than twice the largest retraining threshold we measured, and on Push-T recovers the loss an AR(1) prior incurs against the control; prescribing no cross-time structure at all falls below the control in all four suites.

1 INTRODUCTION

A world model is useful for planning only if consecutive latent states are related in a predictable way. Joint-embedding predictive architectures (Assran et al., 2023) learn that latent without reconstructing pixels, so they need a term that stops the encoder from collapsing, and the current recipe for it is a normality test on the latent marginal. We ask what that recipe guarantees about the relation between instants, for SIGReg, the anti-collapse regulariser of LeWM (Maes et al., 2026), and the answer is nothing: it constrains the marginal at each instant separately and leaves the law linking z_t to z_{t+1} free, so a latent that is constant along every trajectory attains the global minimum of both training terms (Appendix A.8). Figure 1 shows that freedom in a trained model, and what prescribing the missing law changes.

The gap matters because the theory supporting this objective already depends on closing it: linear identifiability of the latent is proved for latents that evolve under stationary, additive-noise transitions, with the Gaussian the unique law for which the guarantee survives (Klindt et al., 2026). The first half is a hypothesis about the cross-time law that the marginal recipe never enforces; the second is why we keep the Gaussian family and move the constraint from the marginals to the joint.

We introduce **T-SIGReg** (temporal SIGReg), which prescribes the joint distribution of a trajectory window — the $h{+}1$ consecutive latents stacked — as $\mathcal{N}(0, R_{\rho,h} \otimes I_d)$ with $(R_{\rho,h})_{ts} = \kappa_{|t-s|}$ for a stationary profile κ, and enforce it with the test SIGReg already uses. That test draws projection directions, and the measure they come from decides what the criterion can detect; we treat it as a design variable. Three things follow:

- **The target.** $R_{\rho,h} \otimes I_d$ is the maximum-entropy law for a stationary profile; a theorem shows the criterion recovers it together with the per-instant marginal, and the profile is measured from the observation stream before training rather than chosen (§3).

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.

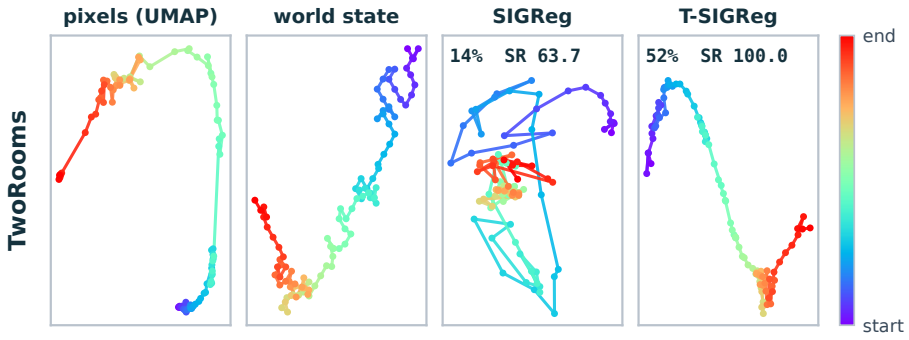


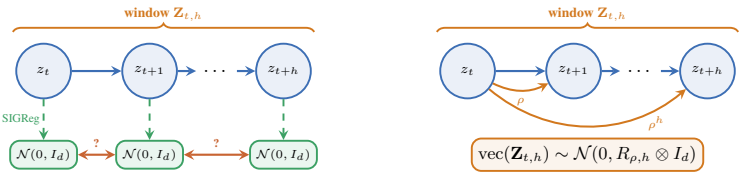
Figure 1: One episode of TwoRooms, coloured by time, in four geometries: a UMAP of the raw **pixels**; the agent's true (x, y) **world state**; and the leading PCA-2 plane of the latent trained with **SIGReg** alone and with the **path law** prescribed (the measured profile of §3.2, mixture design, $\eta = 0.05$). Each latent panel draws one episode of one training run; the percentage is the share of variance the plane holds and SR is the planning success of the corresponding cell of Table 2. Each is its own leading plane. Here the state is linearly recoverable from either latent, R^2 1.00 and 0.99 (§5); what changes is where it sits. The other environments are Figure 12.

- **The designs.** Separable directions $a = q \otimes v$ keep the temporal signal a full-sphere draw averages away, at $O((h+1)d)$ per projection; one spectral margin decides which designs also pin down the spatial law, and the mixture restores what the temporally optimal design gives up, at a linear cost (§4).
- **The measurement.** On four pixel-based suites under a paired protocol, the measured path law improves planning in three, one to saturation; on the fourth an AR(1) prior loses to the control and the measured profile recovers the loss; and prescribing no cross-time structure at all, with identical machinery, falls below the control in all four suites (§5). An open-loop rollout (Appendix C.4) shows the gain does not come through the predictor, which is no better under the path law, but through the geometry of the latent.

2 Related work

LeJEPA-FGP reaches the same diagnosis, a per-instant marginal is not enough, and answers it with a factorised Gaussian-process prior, one kernel per channel (Wei, 2026), aimed at a different target. Their goal is *identifiability*, for which the channels have to be told apart, so a per-channel kernel breaks the rotational symmetry identifiability must break. Ours is planning, which is invariant to a rotation of the latent, so we leave the channels exchangeable and pay nothing for a symmetry we do not need to break. In §4 their method is one point of our family. Temporal straightening (Wang et al., 2026) constrains the same object we do, how consecutive latents relate, for the same end: it penalises the angle between consecutive latent displacements of a JEPA world model so that Euclidean distance tracks progress along feasible trajectories and the planning Hessian stays conditioned, and gradient-based planning gains on Wall, PointMaze and Push-T. Where it differs is what is prescribed and by what: a local geometric prior, straightness, added beside a separate anti-collapse mechanism (stop-gradient), whereas we prescribe a law for the whole window with a profile measured rather than assumed, and the term that prescribes it is the anti-collapse term. Sub-JEPA attacks the same collapse by restricting the embedding to a subspace (Zhao et al., 2026), and GC-IDM arrives from the other side, amortising planning into an inverse dynamics map when the latent geometry is smooth (Nguyen et al., 2026). None of the four asks which measure the projection directions are drawn from, which is the degree of freedom we take up.

Slowness as a learning principle is older than any of this (Földiák, 1991; Wiskott & Sejnowski, 2002; Sobal et al., 2022); we prescribe the profile instead of rewarding slowness generically. Anti-collapse regularisers (Bardes et al., 2022; Zbontar et al., 2021; Chen et al., 2020) act on the marginal or on pairs, never on the law of a window. Our instrument is a sliced distance, of which max-sliced Wasserstein (Deshpande et al., 2019) is the one-direction limit Proposition 3 excludes and Johnson–



(a) each instant constrained, every Σ_{ts} free

(b) the whole window has one prescribed law

Figure 2: **(a)** SIGReg pushes each instant of the window towards $\mathcal{N}(0, I_d)$ separately (dashed) and leaves every cross-time covariance Σ_{ts}, $t \neq s$, unconstrained: the marks between the marginals. **(b)** What we prescribe instead: one law for the whole window, every cross-time block included.

Lindenstrauss (Johnson & Lindenstrauss, 1984) the reason finitely many projections suffice. What the representation is for is control (Hafner et al., 2024; Hansen et al., 2022; 2023; Assran et al., 2025; Zhou et al., 2025).

3 The path law, measured rather than chosen

3.1 Setup and the path law

Observations $o_t \in \mathbb{R}^{H \times W \times 3}$ are encoded to latents $z_t = \mathrm{enc}_\theta(o_t) \in \mathbb{R}^d$. The object this paper is about is the covariance of a window of horizon h, $\mathbf{Z}_{t,h} = [z_t^\top, \ldots, z_{t+h}^\top]^\top \in \mathbb{R}^{(h+1)d}$, written in $d \times d$ blocks $\Sigma_{ts} = \mathbb{E}[z_t z_s^\top]$. LeWM trains a next-embedding prediction loss plus SIGReg, a regulariser pushing the latent marginal towards $\mathcal{N}(0, I_d)$ instant by instant (Maes et al., 2026; Balestriero & LeCun, 2025); §3.3 writes it out. Its reach is the point: it acts on the diagonal blocks Σ_{tt} and never on Σ_{ts} with $t \neq s$ (Figure 2a). Any marginal-only regulariser leaves the same freedom; our measurements are of LeWM alone. We write $\mathbb{S}^{d-1}$ for the unit sphere of channel directions and $e_t \in \mathbb{R}^{h+1}$ for the canonical vector of the temporal axis, so $e_t \otimes v$ looks only at instant t.

We prescribe the whole matrix instead (Figure 2b), and start from the profile rather than from a model of the dynamics. Fix any admissible stationary autocorrelation $\kappa_1, \ldots, \kappa_h$ on the temporal axis. Among all window laws with that profile, the maximum-entropy one is Gaussian with the Toeplitz covariance $R_{ts} = \kappa_{|t-s|}$, so once the profile is fixed the law is the least committal thing compatible with it (Cover & Thomas, 2006). One further choice makes it the target we train against: the channels are left exchangeable, so $\Sigma_{ts} = \kappa_{|t-s|} I_d$ rather than a general $d \times d$ block, the choice §2 defends since planning is invariant to a rotation of the latent. Stacking the window, the whole matrix factorises as $\Sigma_{\rho,h} = R_{\rho,h} \otimes I_d$ with $(R_{\rho,h})_{ts} = \kappa_{|t-s|}$: $R_{\rho,h}$ sets the temporal structure, I_d says no channel is privileged. AR(1) is the one-parameter prior inside this family: a linear transition, the discrete-time Ornstein–Uhlenbeck process (Uhlenbeck & Ornstein, 1930), has $\kappa_\ell = \rho^\ell$ and makes $R_{\rho,h}$ the Kac–Murdock–Szegő matrix (Kac et al., 1953) (Appendix A.1). It is where we start, not where we must stay: the maximum-entropy argument never used geometry.

3.2 Where the profile comes from: the world clock

Nothing forces κ to be guessed: any admissible κ is a legal target, and an admissible one can be *measured* from the data before training begins. We measure the autocorrelation of the *observation stream*, which we call the world clock: stretches of expert frames are embedded into two dimensions with UMAP (McInnes et al., 2020), κ_ℓ is the mean autocorrelation of the embedded coordinates at lag ℓ, and the target $R_{ts} = \kappa_{|t-s|}$ is fixed before the first gradient step. It is measured from pixels, which are all the encoder sees, not from the state (Table 5); Appendix B gives recipe and provenance.

Two properties, read off Figure 3, decide whether this is an experiment or a formality. All four measured matrices are positive definite, so all four are legal targets. And none is geometric: $\kappa_2 - \kappa_1^2$ vanishes for every AR(1), while every world we measured is strictly negative on it, decaying faster after the first lag than any single-ρ prior. The measured matrix therefore moves the target to a place no choice of ρ can reach, which is what §5 tests.

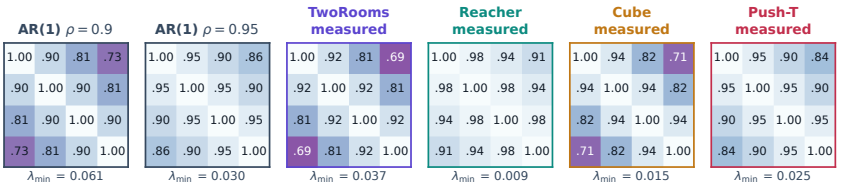


Figure 3: The target R at $h = 3$ (one window step is five environment steps): the two AR(1) priors we train with and the four measured profiles, read from the training checkpoints. All four decay faster after the first lag than a geometric profile through their own κ_1; Push-T's agrees with $\rho = 0.95$ at lag one and separates after. All are positive definite, with the smallest eigenvalue under each panel.

3.3 Testing it without fitting it

The instrument is SIGReg's and we inherit it unchanged. Balestriero & LeCun (2025) make the argument we rely on: matching a covariance constrains only second moments, so the test has to reach the whole law, and a normality test on random one-dimensional projections does, by Cramér–Wold (Cramér & Wold, 1936), with a characteristic-function statistic per projection (Epps & Pulley, 1983). Our change is to what the instrument is pointed at, in LeWM's notation: $Z \in \mathbb{R}^{N\times B\times d}$ is the tensor of latents over a history of length $N = h{+}1$ and a batch of B, each batch element one window $\mathbf{Z}_{t,h}$; $v^{(m)} \in \mathbb{S}^{d-1}$ are M random unit channel directions; and $T(x) = \int |\hat{\varphi}_x(\omega) - e^{-\omega^2/2}|^2 w(\omega)\,\mathrm{d}\omega$ is the Epps–Pulley statistic of a one-dimensional sample, $\hat{\varphi}_x$ its empirical characteristic function, whose population version vanishes if and only if $x \sim \mathcal{N}(0,1)$ when w is positive on an interval (Epps & Pulley, 1983). SIGReg projects the whole tensor and pools the NB values of every instant into one sample, $\mathrm{SIGReg}(Z) = \frac{1}{M}\sum_m T(Zv^{(m)})$ (Maes et al., 2026). The path term projects the whole window along a direction $a^{(m)} \in \mathbb{S}^{(h+1)d-1}$ drawn from a design ν, and standardises by the variance the target assigns to that direction:

$$\text{T-SIGReg}_\nu(Z) \triangleq \frac{1}{M}\sum_{m=1}^{M} T\big(\tilde{s}^{(m)}\big), \qquad \tilde{s}^{(m)} = \frac{a^{(m)\top}\mathbf{Z}_{t,h}}{\sqrt{a^{(m)\top}\Sigma_{\rho,h}\,a^{(m)}}}, \quad a^{(m)} \sim \nu. \tag{1}$$

Two facts about equation 1 carry the rest of the paper. First, SIGReg *is* equation 1 with ν supported on the marginal directions $a = e_t \otimes v$, where $e_t^\top R_{\rho,h}\, e_t = 1$ makes the standardisation trivial: the two terms differ only in the measure the direction is drawn from, and either is an anti-collapse term, since a degenerate law fails the test along almost every direction. Second, the denominator is the *target* deviation, not the empirical one, which would leave the criterion blind to scale. Algorithm 1 (Appendix A) is the estimator for $a = q \otimes v$; the measure q is drawn from is its second line.

3.4 The criterion recovers the law

> **Theorem 1** (the criterion recovers the path law)**.** *Let $a \sim \mathrm{Unif}(\mathbb{S}^{(h+1)d-1})$. Then the criterion vanishes if and only if $\mathbf{Z}_{t,h} \sim \mathcal{N}(0, \Sigma_{\rho,h})$, and in particular $\Sigma_{tt} = I_d$ at every instant.*

The proof is a continuity argument (Appendix A.9): the criterion averages a non-negative continuous quantity over the whole sphere, so vanishing in expectation forces it to vanish at every direction, and Cramér–Wold closes; the marginal is a corollary because $e_t \otimes v$ is among the directions drawn. Two gaps remain: the estimator evaluates T on a finite grid of K frequencies, which does not characterise normality, so the trained criterion has the target as its null without characterising it; and the condition is full support, which §4 gives up, so the guarantee has to be re-earned there.

4 Designs are measures on the sphere

The measure ν in equation 1 is a *design* in the sense of optimal experimental design (Pukelsheim, 1993), which turns an implementation choice into an object with theory attached (Figure 8, Ap-

Table 1: The four designs: temporal sensitivity (Definition 1, Propositions 1 and 2) and which spatial guarantee survives (Proposition 3). †: obtained, with a margin four orders of magnitude below the mixture's (Appendix C). "Not shown": our proof needs the marginal directions in the support; we have neither a proof without them nor a counterexample.

design ν	support of μ	$\mathrm{sens}_\nu(\Delta)$	covariance	law
full sphere of $\mathbb{R}^{(h+1)d}$	—	$\Theta\big(\|A\|_F^2/d\big)$	yes	yes
separable, q uniform	$\mathbb{S}^h$, continuous	$\Theta\big(\|A\|_F^2\big)$	yes	yes
separable, $q \sim \mu^\star$	finite, $e_t \notin \mathrm{supp}\,\mu$	$\Theta\big(\|A\|_F^2\big)$, maximal $\lambda_{\min}$	yes†	**not shown**
separable, $q \sim \mu_\eta$	finite, $e_t \in \mathrm{supp}\,\mu$	$\geq (1-\eta)$ of maximal $\lambda_{\min}$	**yes**	**yes**

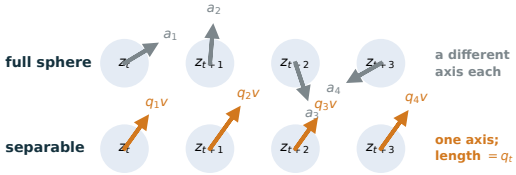


Figure 4: The two ways to draw one direction a for a window. A draw from the full sphere gives an independent block per instant, so each z_{t+i} is contracted along a different axis and the temporal pattern is averaged against itself. A separable draw $a = q \otimes v$ gives the same axis at every instant, with the temporal profile appearing as the length of each arrow. Appendix A.2 plots the resulting ratio of sensitivities against d.

pendix A). We walk a ladder of four designs, each repairing the defect of the previous one while breaking something new; Table 1 is the map.

4.1 The full sphere dilutes the temporal signal

If the realised covariance departs from the target, $\Sigma = \Sigma_{\rho,h} + \Delta$, call the departure *purely temporal* when $\Delta = A \otimes I_d$ with A symmetric and zero on the diagonal: the correlation between instants t and s has drifted by A_{ts} in every channel direction while every per-instant covariance is exactly right. This is the class the marginal recipe is blind to by construction and the path law adds, of dimension $h(h+1)/2$. What a design *notices* of such a departure is the second-order term of its own loss.

Definition 1 (temporal sensitivity)**.** For a design ν on the sphere,

$$\mathrm{sens}_\nu(\Delta) \;:=\; \mathbb{E}_{a\sim\nu}\left[\left(\frac{a^\top \Delta\, a}{a^\top \Sigma_{\rho,h}\, a}\right)^2\right],$$

the squared relative variance deviation the sketch sees, which is the population loss to second order (Appendix A.5).

The natural choice is the one that fails.

Proposition 1 (the full sphere fades with d)**.** *With* $a \sim \mathrm{Unif}(\mathbb{S}^{(h+1)d-1})$,

$$\mathrm{sens}_{\mathrm{full}}(\Delta) \;=\; \Theta\left(\frac{\|A\|_F^2}{d}\right),$$

with matching upper and *lower bounds whose constants depend only on* $R_{\rho,h}$*, uniformly over the whole temporal subspace.*

The uniformity matters as much as the rate: the whole temporal subspace fades together, with no low-rank part to exploit, because a direction on $\mathbb{S}^{(h+1)d-1}$ reads each instant through a *different* channel direction and the temporal pattern is averaged against itself across d.

4.2 SEPARABLE SKETCHES AND WHAT THEY PRESERVE

Take $a = q \otimes v$, with q on the temporal axis and $v \in \mathbb{S}^{d-1}$ shared across instants. The sensitivity is then one line, $a^\top \Delta\, a = (q \otimes v)^\top (A \otimes I_d)(q \otimes v) = (q^\top A\, q)(v^\top v) = q^\top A\, q$, and d has disappeared (Figure 4). The sketch projects in space first and in time second (Figure 7, Appendix A.7): every instant is contracted along the *same* v, so the series $y_i = \langle z_{t+i}, v\rangle$ is commensurable across i and its correlation is the structure under test; then the series is contracted along q, so $q \otimes v$ is never formed and a projection costs $O((h{+}1)d)$ with exact target variance $q^\top R_{\rho,h}\, q$ (measured: 0.66× the memory and 1.32× the time of SIGReg, Appendix B).

Two consequences of v dropping out organise the section. First, a separable design factorises, $\nu = \mu \otimes \mathrm{Unif}(\mathbb{S}^{d-1})$, with the channel factor uniform throughout, so from here on a *design* is the temporal factor μ on $\mathbb{S}^h$. Second, the ratio in Definition 1 is *linear* in a purely temporal A (§4.1): $q^\top A\, q / q^\top R_{\rho,h}\, q = \langle A,\ qq^\top / (q^\top R_{\rho,h}\, q)\rangle_F$. With $u(q) \in \mathbb{R}^{h(h+1)/2}$ the off-diagonal entries of $qq^\top/(q^\top R_{\rho,h}\, q)$ and α_A the matching entries of A, a departure between t and s is seen only through the weight the profile puts on that *pair*, and the sensitivity is quadratic in the design,

$$\mathrm{sens}_\mu(A \otimes I_d) \;=\; \mathbb{E}_{q\sim\mu}\big[(\alpha_A^\top u(q))^2\big] \;=\; \alpha_A^\top \underbrace{\mathbb{E}_{q\sim\mu}\big[u(q)\,u(q)^\top\big]}_{G_\mu} \alpha_A \tag{2}$$

(Appendix A.7 spells out the step). The criterion sees μ only through the Gram matrix G_μ, and every claim below about a design is a claim about its spectrum.

Proposition 2 (the separable sketch preserves the signal)**.** *With $q \sim \mathrm{Unif}(\mathbb{S}^h)$ and $v \sim \mathrm{Unif}(\mathbb{S}^{d-1})$,*

$$\mathrm{sens}_{\mathrm{sep}}(\Delta) \;=\; \mathbb{E}_q\left[\frac{(q^\top A\, q)^2}{(q^\top R_{\rho,h}\, q)^2}\right] \;=\; \Theta\big(\|A\|_F^2\big),$$

free of d. Together with Proposition 1, the separable advantage on the temporal axis is $\Theta(d)$. For a general μ the rate survives only through the margin of equation 2, $\mathrm{sens}_\mu(\Delta) \geq \lambda_{\min}(G_\mu)\|A\|_F^2$: a μ concentrated where $q^\top A\, q = 0$ has sensitivity zero at that A.

The constant is available in closed form on both sides, with no ρ and no trained model in it: at the $d = 192$ and $h = 3$ of our runs the ratio is ≈ 128 (Appendix A.2). With q uniform the separable criterion also recovers the marginal law with no marginal term in the objective.

Theorem 2 (the marginal returns for free)**.** *Let $a = q \otimes v$ with $q \sim \mathrm{Unif}(\mathbb{S}^h)$ and $v \sim \mathrm{Unif}(\mathbb{S}^{d-1})$ drawn independently. Then a separable criterion driven to zero forces $z_t \sim \mathcal{N}(0, I_d)$ at every instant, without the marginal term appearing in the objective.*

The step that carries it is the continuity argument of Theorem 1 on $\mathbb{S}^h$: a uniform q has e_t in its support, where $e_t^\top R_{\rho,h}\, e_t = 1$ leaves $\langle z_t, v\rangle \sim \mathcal{N}(0,1)$ for every v (Appendix A). What the next step violates is the support, and two guarantees come apart there.

4.3 THE OPTIMAL TEMPORAL DESIGN AND WHAT IT CANNOT SEE

The uniform temporal design spends as many draws on the scale question $(1,1,1,1)/2$ as on whether z_t resembles z_{t+1}. The worst case of equation 2 over purely temporal A with $\|A\|_F = 1$ is $\lambda_{\min}(G_\mu)$; maximising it is E-optimal design (Pukelsheim, 1993), and the worst case is the right target because a direction the criterion barely sees is one gradient descent drifts along for free. Frank–Wolfe raises $\lambda_{\min}(G_\mu)$ by 7.7× at $\rho = 0.9$ and 11.2× at $\rho = 0.95$ over the uniform design, at the same cost per projection (Algorithm 2). The solution has one shape: nearly all of its mass sits on profiles of zero sum, since any component of q along the constant profile $\mathbf{1}$ is norm no temporal departure responds to. The price is exact.

Proposition 3 (what a contrast design cannot see)**.** *Let μ be a design on $\mathbb{S}^h$ and $J = \mathbf{1}\mathbf{1}^\top$.*

(i) If every atom of μ has $\mathbf{1}^\top q = 0$, the criterion takes the same value at $\Sigma_{\rho,h}$ and at $\Sigma_{\rho,h} + c\, J \otimes I_d$ for every $c \geq 0$: the design cannot see $z_t \mapsto z_t + \xi$ with one $\xi \in \mathbb{R}^d$ shared by the whole window, and cannot force $\mathrm{Cov}(z_t) = I_d$.

(ii) If $\{qq^\top : q \in \operatorname{supp}\mu\}$ *spans* $\operatorname{Sym}(h{+}1)$*, the criterion at zero forces* $\operatorname{Cov}(z_t) = I_d$ *at every* t*; if in addition every* e_t *is an atom of* μ*, it forces* $z_t \sim \mathcal{N}(0, I_d)$.

Part (i) is two lines: a shared ξ moves the sketch by $\langle \xi, v\rangle \sum_t q_t = 0$, and only the marginal direction sees it, since $\mathbf{1}^\top e_t = 1$. Part (ii) is Appendix A. The optimum is not exactly zero-sum, but its sensitivity along J sits orders of magnitude below its temporal one (Table 1, Appendix A.4), and J is the direction §3 warned about: inflating the common mode is a step towards the constant window of Appendix A.8. The repair is in part (i): an atom at e_t sees it directly.

4.4 The mixture temporal design

Put back exactly what the optimum removed and no more: with $\mu^\star$ the temporally optimal design and $\eta \in (0, 1]$,

$$\mu_\eta \;=\; (1-\eta)\,\mu^\star \;+\; \eta\,\mathrm{Unif}\big(\{e_t\}_{t=0}^{h}\big). \tag{3}$$

Proposition 4 (the mixture restores the law, and the margin)**.** *For every* $\eta \in (0, 1]$ *the design* μ_η *of equation 3 has* $e_t \in \operatorname{supp}\mu_\eta$ *for all* t*, so part (ii) of Proposition 3 applies and a separable criterion driven to zero under* μ_η *forces* $z_t \sim \mathcal{N}(0, I_d)$ *at every instant. Moreover the temporal information is reduced by at most a factor* $1 - \eta$*,*

$$\lambda_{\min}\big(G_{\mu_\eta}\big) \;\geq\; (1-\eta)\,\lambda_{\min}\big(G_{\mu^\star}\big),$$

so the two guarantees hold together at a cost linear in η.

Measurement shows the bound holds with equality and the margin along J rises by four orders of magnitude, to within a factor of 7–13 of the uniform design's (Appendix C). We trained two values of η, within noise of each other in Table 2, and did not sweep it.

5 Does the path law help the planner?

5.1 The objective and the protocol

LeWM trains $\mathcal{L}_{\text{LeWM}} = \mathcal{L}_{\text{pred}} + \lambda\,\text{SIGReg}(Z)$, with $\mathcal{L}_{\text{pred}} = \|\hat{z}_{t+1} - z_{t+1}\|_2^2$ and $\hat{z}_{t+1} = \text{pred}_\phi(z_t, a_t)$ (Maes et al., 2026). We change one term and nothing else:

$$\mathcal{L} \;\triangleq\; \mathcal{L}_{\text{pred}} \;+\; \lambda\;\text{T-SIGReg}_{\mu_\eta}(Z), \tag{4}$$

with λ, M and T as in our retrained control (Appendix B). The control is the same objective with ν set to the marginal directions $\{e_t \otimes v\}$; every other row of Table 2 is equation 4 with a different ν, or with a different target R. We do not keep SIGReg alongside the path term, by design: Theorem 2 and Proposition 4 say the design recovers the marginal law at the null, and Table 5 measures how far finite training stays from it.

Four pixel-based suites: TwoRooms, Reacher (Tassa et al., 2018), Cube (Park et al., 2025) and Push-T (Florence et al., 2021), with the environments, harness and CEM planner (Rubinstein, 1999) LeWM released; success is measured over goal offsets from 15 to 75 steps, to 135 on Reacher, at a budget of twice the offset. The protocol is paired: identical architecture, data, compute and training seed, with the design or the target as the only difference; a cell asks for three independent training runs, each under six rollout seeds. The $\pm$ is the spread across evaluation seeds and is not what a retraining moves; that is estimated from the families with seed-independent replicates, times $2\sqrt{2}$, giving the threshold a paired gain has to clear: 3.8 on TwoRooms (30 degrees of freedom), 0.9 on Reacher (28), 1.6 on Cube (26), 3.9 on Push-T (26). Headline comparisons are judged against the largest of the four, 3.9, so one bar serves every suite; per-suite thresholds enter only contrasts within a suite (Appendix B).

5.2 The ladder, measured

This experiment tests whether the order Table 1 predicts is the order the planner sees; Table 2 reads top to bottom. Prescribing no cross-time structure, which is the wrong law for a world that moves slowly, falls below the control in all four suites, so whatever the rows below gain is the content of R,

Table 2: The ladder of §4, measured, with two controls: SIGReg retrained in our pipeline, as the paired protocol requires (the published checkpoint is an anchor), and the same path term prescribing $R = I_{h+1}$, no cross-time structure with identical machinery. Cells: mean over training runs and evaluation seeds; the $\pm$ is the spread across *evaluation* seeds, not what a retraining moves (Appendix B). Every cell averages 3 independent training runs $\times$ 6 evaluation seeds, except the two controls, SIGReg and $\rho = 0$, which have one training run each. The $\rho = 0$ row is a negative control (Appendix C.8). Bold rows: the separable mixture, the design we recommend. The per-channel row is the design of Wei (2026) inside our family.

design	TwoRooms	Reacher	Cube	Push-T
SIGReg only (Maes et al., 2026)	63.7 $\pm$1.4	82.4 $\pm$1.2	63.0 $\pm$2.6	54.9 $\pm$3.2
published checkpoint (anchor)	60.7 $\pm$1.0	83.9 $\pm$1.4	62.9 $\pm$2.1	50.5 $\pm$3.2
prescribing $\rho = 0$, i.e. $R = I_{h+1}$: no cross-time structure				
separable, q uniform	31.8 $\pm$1.7	74.3 $\pm$1.7	52.5 $\pm$2.7	17.1 $\pm$0.8
prescribing $\rho = 0.9$				
full sphere, shared kernel	67.8 $\pm$1.3	79.3 $\pm$1.3	57.2 $\pm$1.9	51.6 $\pm$2.4
full sphere, per channel (Wei, 2026)	68.1 $\pm$1.6	79.6 $\pm$1.0	57.5 $\pm$2.7	51.5 $\pm$2.4
separable, q uniform	92.1 $\pm$0.9	87.4 $\pm$0.9	65.8 $\pm$2.4	45.9 $\pm$1.8
separable, $q \sim \mu^\star$	93.3 $\pm$0.6	89.9 $\pm$1.1	64.5 $\pm$2.6	44.6 $\pm$2.4
separable, $q \sim \mu_\eta$, $\eta = 0.05$	100.0 $\pm$0.0	90.9 $\pm$0.7	73.7 $\pm$2.0	49.3 $\pm$1.9
separable, $q \sim \mu_\eta$, $\eta = 0.1$	100.0 $\pm$0.0	91.4 $\pm$0.8	72.6 $\pm$2.0	47.8 $\pm$1.8
prescribing $\rho = 0.95$, same designs				
separable, q uniform	88.7 $\pm$0.9	88.6 $\pm$1.5	65.9 $\pm$2.7	37.7 $\pm$2.4
separable, $q \sim \mu^\star$	97.1 $\pm$0.3	91.0 $\pm$0.6	64.1 $\pm$2.2	42.1 $\pm$2.0
separable, $q \sim \mu_\eta$, $\eta = 0.05$	100.0 $\pm$0.0	93.7 $\pm$0.5	74.4 $\pm$1.7	49.3 $\pm$2.1
separable, $q \sim \mu_\eta$, $\eta = 0.1$	100.0 $\pm$0.0	93.8 $\pm$0.5	73.8 $\pm$1.8	47.7 $\pm$1.9
prescribing the measured profile, same designs				
measured profile, q uniform	91.1 $\pm$1.1	85.7 $\pm$1.1	66.6 $\pm$2.5	42.9 $\pm$2.5
measured profile, $q \sim \mu_\eta$, $\eta = 0.05$	100.0 $\pm$0.0	93.6 $\pm$0.5	73.9 $\pm$1.6	54.0 $\pm$2.8
measured profile, $q \sim \mu_\eta$, $\eta = 0.1$	100.0 $\pm$0.0	93.7 $\pm$0.6	73.9 $\pm$1.7	52.2 $\pm$2.2

not the change of regulariser. Within the $\rho = 0.9$ block the full-sphere designs sit below the control except on TwoRooms, where they gain four points, and the separable mixture clears it by more than 3.9 in three suites and is the best separable design in all four. TwoRooms and Reacher order the designs as Table 1 does; on Push-T every AR(1) target loses to the control, a failure of the target rather than of the machinery (§5.3).

5.3 The measured profile against the priors

This experiment tests whether measuring the profile instead of choosing ρ buys anything: the last block of Table 2 changes only R. Paired within evaluation seed against the control, the measured profile with the mixture at $\eta = 0.05$ gains +36.3 $[+34.8, +37.8]$ on TwoRooms, +11.2 $[+10.1, +12.4]$ on Reacher and +10.9 $[+9.0, +12.8]$ on Cube, each more than twice the largest threshold 3.9, and -0.9 $[-2.4, +0.6]$ on Push-T, an interval that contains zero. Against the priors, the measured profile matches the best AR(1) target where AR(1) helps, and on Push-T it comes closest to the control (54.0 against 54.9; the best AR(1) row reaches 51.6).

Push-T is where the measured profile and a prior come closest. Against the $\rho = 0.95$ prior, whose κ_1 agrees with the measured one at lag one (§3.2), the measured profile gains $+4.68$ points, $[+4.05, +5.31]$, clearing the suite's threshold of 3.9; but two things differ between those rows, the profile and the ρ-indexed design atoms (Appendix B), and against the $\rho = 0.9$ prior, which loads the same atoms, the gain is $+4.67$ points, $[+3.80, +5.55]$, at the threshold: the profile alone carries most of the recovery, too thinly to call its *shape* the cause.

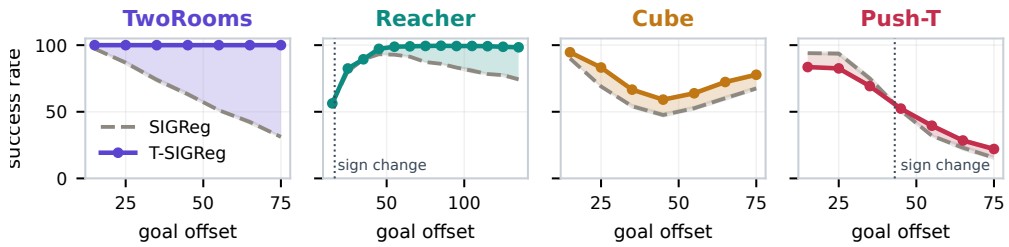


Figure 5: Planning success by goal offset. Dashed, the SIGReg control; solid, *one* configuration held fixed across environments (the measured profile, mixture, $\eta = 0.05$), so nothing is selected on the data the gain is read from. Bands: ± 1 standard error over evaluation seeds; shaded: the gain (Table 4).

5.4 Where the gain lives

Averaging over offsets hides where the improvement comes from, and the mechanism makes a prediction: a planner's error accumulates with the distance to the goal, so constraining how consecutive latents relate should pay more the further the goal is. Figure 5 resolves the same configuration by offset; Table 4 (Appendix C) gives the numbers, bold where the paired gain clears its suite's per-offset threshold by two standard errors. On TwoRooms and Reacher the gain opens towards distant goals as the control decays; on Cube it is flat and clears the threshold at every offset. Push-T changes sign: the law costs points at the three nearest goals, where the control is near its ceiling, and gains at the four furthest, but no far-goal cell clears 5.5 by that margin, so the finding is the shape of the curve, not any cell of it, and nothing in our argument predicts the short-range cost.

5.5 What changed in the latent

Figure 1 is the qualitative result; Table 6 (Appendix C.2) gives the probe and Table 5 the spectral diagnostic. A linear probe of the true state, tested on held-out episodes, reads the same from both latents on TwoRooms (1.00 and 0.99), only partially elsewhere, and on Push-T worse from ours (0.70 against 0.88): the prescription rearranges information rather than adding it, and in one suite costs some. The share of variance in the leading plane rises in every suite, the temporally optimal design falls furthest from unit variance (Proposition 3(i) in a trained model), and no cell reaches its prescribed $\bar{\kappa}$. Whether the planner uses the arrangement rather than the compression is not separated here; that it does not use a better predictor is Appendix C.4, and that its cost landscape keeps tracking the goal where the control's stops is §C.3.

6 Conclusion and Limitations

The introduction borrowed a hypothesis and returns it here: Klindt et al. (2026) prove identifiability under stationary additive-noise transitions, a marginal-only regulariser leaves it to luck, and this paper makes it a constraint on the joint law of a window. An anti-collapse regulariser is a measure on a sphere of directions, SIGReg being the one supported on the marginal ones, so *which directions to test* is a design problem with theory: what a design sees is the spectrum of one Gram matrix, its blind spots are computable before training and are where gradient descent drifts, and the marginal law is a corollary of the joint one, so one term does the work of two. The temporal target is a parameter of the law, not a hyperparameter: the autocorrelation of the world as the encoder sees it, measured before training and geometric in none of the four worlds.

For latent world models, temporal meaning can be prescribed rather than hoped for, and it pays through the geometry of the latent, not the predictor: wherever the control decays with distance to the goal, and not at short range on Push-T, where the cost is real and Appendix C.5 hypothesises about it. The profile needs *expert* frames: a clock read off exploratory or closed-loop data would describe the collecting policy as much as the world. We never test identifiability itself — the theorem only motivates the target — and no comparison here is external (Hansen et al., 2023; Hafner et al., 2024). What the claims rest on is stated where it is used: the thresholds (§B) and the open-loop and distance measurements (§C.4, §C.3).

ETHICS STATEMENT

This work trains world models on four simulated control benchmarks and reports success rates on them. It involves no human subjects, no personal or otherwise sensitive data, and no deployment on physical systems; the datasets are the public ones released with the benchmarks. We see no direct path from this contribution to harmful application beyond the risks already carried by research on learned world models in general, and we have complied with the ICLR Code of Ethics.

AI USE STATEMENT

In this work we used a generative AI assistant (Claude, Anthropic) for drafting and editing the text of the paper, for searching the literature and surfacing candidate related work, for writing and debugging the code that generates every table and figure from the experiment records, and for critical review of the manuscript, in particular for checking that the numbers quoted in the text are the ones the tables carry. Every reference surfaced this way was verified against its published record before being cited, and any that could not be verified was dropped. We did not use generative AI to generate the research question, the theoretical results, the experimental design or the interpretation of the results, which are the authors'. Every proof was checked line by line by the authors, every AI-assisted piece of code was run and its output compared against the records it reads, and every sentence of the text was read and edited by the authors. We take responsibility for the final content of this work, including text, claims and artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

Appendix A states every theoretical result with its hypotheses and gives a complete proof, and Algorithm 1 is the path term as implemented, line by line. Appendix B gives the architecture, optimiser, criterion hyperparameters, window horizon, number of projections, frequency grid, regulariser weight, mixture weights, training and evaluation seeds, the retraining-threshold computation and the hardware, together with the two provenance caveats we know of: the prescribed profile is read back from a buffer in each checkpoint but the exact embedding recipe that produced it is reproduced only approximately, and the backend differs across cells, with the confound bounded rather than assumed away. Appendix C gives the construction of the optimal design (Algorithm 2) with its candidate count, iteration count and weight threshold, and prints the full spectrum a reader can check the design against. Every table and figure is generated by a script from the evaluation records, with no number typed by hand. The code, a fork of the public LeWM repository, is included in the anonymized supplementary material: the path term, the design construction, the training configurations of every cell in Table 2, the precomputed design atoms and the scripts that generate the tables and figures.

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, 2023. URL https://arxiv.org/abs/2301.08243.

Mido Assran, Adrien Bardes, David Fan, Quentin Garrido, Russell Howes, Mojtaba Komeili, Matthew Muckley, Ammar Rizvi, Claire Roberts, Koustuv Sinha, Artem Zholus, Sergio Arnaud, Abha Gejji, Ada Martin, Francois Robert Hogan, Daniel Dugas, Piotr Bojanowski, Vasil Khalidov, Patrick Labatut, Francisco Massa, Marc Szafraniec, Kapil Krishnakumar, Yong Li, Xiaodong Ma, Sarath Chandar, Franziska Meier, Yann LeCun, Michael Rabbat, and Nicolas Ballas. V-JEPA 2: Self-supervised video models enable understanding, prediction and planning, 2025. URL https://arxiv.org/abs/2506.09985.

Randall Balestriero and Yann LeCun. LeJEPA: Provable and scalable self-supervised learning without the heuristics, 2025. URL https://arxiv.org/abs/2511.08544.

Adrien Bardes, Jean Ponce, and Yann LeCun. VICReg: Variance-invariance-covariance regularization for self-supervised learning. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=xm6YD62D1Ub.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In Hal Daumé III and Aarti Singh (eds.), *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 1597–1607. PMLR, 13–18 Jul 2020. URL https://proceedings.mlr.press/v119/chen20j.html.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, July 2006. ISBN 0471241954.

H. Cramér and H. Wold. Some theorems on distribution functions. *Journal of the London Mathematical Society*, s1-11(4):290–294, 1936. doi: 10.1112/jlms/s1-11.4.290. URL https://londmathsoc.onlinelibrary.wiley.com/doi/abs/10.1112/jlms/s1-11.4.290.

Ishan Deshpande, Yuan-Ting Hu, Ruoyu Sun, Ayis Pyrros, Nasir Siddiqui, Sanmi Koyejo, Zhizhen Zhao, David Forsyth, and Alexander G. Schwing. Max-sliced Wasserstein distance and its use for GANs. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10640–10648, 2019. doi: 10.1109/CVPR.2019.01090.

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, et al. An image is worth 16x16 words: Transformers for image recognition at scale. In *ICLR*, 2021.

Thomas W. Epps and Lawrence B. Pulley. A test for normality based on the empirical characteristic function. *Biometrika*, 70:723–726, 1983. URL https://api.semanticscholar.org/CorpusID:121128829.

Pete Florence, Corey Lynch, Andy Zeng, Oscar Ramirez, Ayzaan Wahid, Laura Downs, Adrian Wong, Johnny Lee, Igor Mordatch, and Jonathan Tompson. Implicit behavioral cloning. In *Conference on Robot Learning*, volume 164 of *Proceedings of Machine Learning Research*, 2021. URL https://proceedings.mlr.press/v164/florence22a.html.

Marguerite Frank and Philip Wolfe. An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, 3(1-2):95–110, 1956. doi: 10.1002/nav.3800030109. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/nav.3800030109.

Peter Földiák. Learning invariance from transformation sequences. *Neural Computation*, 3(2):194–200, 06 1991. ISSN 0899-7667. doi: 10.1162/neco.1991.3.2.194. URL https://doi.org/10.1162/neco.1991.3.2.194.

Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse domains through world models, 2024. URL https://arxiv.org/abs/2301.04104.

Nicklas Hansen, Hao Su, and Xiaolong Wang. TD-MPC2: Scalable, robust world models for continuous control. arXiv:2310.16828, 2023.

Nicklas A Hansen, Hao Su, and Xiaolong Wang. Temporal difference learning for model predictive control. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato (eds.), *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 8387–8406. PMLR, 17–23 Jul 2022. URL https://proceedings.mlr.press/v162/hansen22a.html.

William B. Johnson and Joram Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. In *Conference in Modern Analysis and Probability*, volume 26 of *Contemporary Mathematics*, pp. 189–206. American Mathematical Society, 1984. doi: 10.1090/conm/026/737400.

Mark Kac, William L. Murdock, and Gábor Szegő. On the eigen-values of certain Hermitian forms. *Indiana University Mathematics Journal*, 2:767–800, 1953.

David Klindt, Yann LeCun, and Randall Balestriero. When does LeJEPA learn a world model?, 2026. URL https://arxiv.org/abs/2605.26379.

Lucas Maes, Quentin Le Lidec, Damien Scieur, Yann LeCun, and Randall Balestriero. LeWorldModel: Stable end-to-end joint-embedding predictive architecture from pixels, 2026. URL https://arxiv.org/abs/2603.19312.

Leland McInnes, John Healy, and James Melville. UMAP: Uniform manifold approximation and projection for dimension reduction, 2020. URL https://arxiv.org/abs/1802.03426.

Hoang Nguyen, Xiaohao Xu, and Xiaonan Huang. Latent geometry beyond search: Amortizing planning in world models, 2026. URL https://arxiv.org/abs/2605.08732.

Seohong Park, Kevin Frans, Benjamin Eysenbach, and Sergey Levine. OGBench: Benchmarking offline goal-conditioned RL. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=M992mjgKzI.

Friedrich Pukelsheim. *Optimal Design of Experiments*. John Wiley & Sons, 1993.

Reuven Y. Rubinstein. The cross-entropy method for combinatorial and continuous optimization. *Methodology and Computing in Applied Probability*, 1(2):127–190, 1999.

Vlad Sobal, Jyothir S V, Siddhartha Jalagam, Nicolas Carion, Kyunghyun Cho, and Yann LeCun. Joint embedding predictive architectures focus on slow features, 2022. URL https://arxiv.org/abs/2211.10831.

Yuval Tassa, Yotam Doron, Alistair Muldal, Tom Erez, Yazhe Li, Diego de Las Casas, David Budden, Abbas Abdolmaleki, Josh Merel, Andrew Lefrancq, Timothy Lillicrap, and Martin Riedmiller. DeepMind control suite, 2018. URL https://arxiv.org/abs/1801.00690.

G. E. Uhlenbeck and L. S. Ornstein. On the Theory of the Brownian Motion. *Phys. Rev.*, 36: 823–841, 1930. doi: 10.1103/PhysRev.36.823.

Ying Wang, Oumayma Bounou, Gaoyue Zhou, Randall Balestriero, Tim G. J. Rudner, Yann LeCun, and Mengye Ren. Temporal straightening for latent planning. In *Proceedings of the 43rd International Conference on Machine Learning (ICML)*, 2026. arXiv:2603.12231.

Yuan-Hao Wei. LeJEPA-FGP: Factorized Gaussian process regularization for identifiable joint-embedding world models. *Preprints*, June 2026. doi: 10.20944/preprints202606.1157.v1. URL https://doi.org/10.20944/preprints202606.1157.v1.

Laurenz Wiskott and Terrence J. Sejnowski. Slow feature analysis: Unsupervised learning of invariances. *Neural Computation*, 14(4):715–770, 2002. doi: 10.1162/089976602317318938.

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*, 2021. URL https://api.semanticscholar.org/CorpusID:232110471.

Kai Zhao, Dongliang Nie, Yuchen Lin, Zhehan Luo, Yixiao Gu, Deng-Ping Fan, and Dan Zeng. Sub-JEPA: Subspace Gaussian regularization for stable end-to-end world models, 2026. URL https://arxiv.org/abs/2605.09241.

Gaoyue Zhou, Hengkai Pan, Yann LeCun, and Lerrel Pinto. DINO-WM: World models on pre-trained visual features enable zero-shot planning. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=D5RNACOZEI.

A Statements and proofs

A.1 From the OU transition to the path law

The target of §3 is not postulated; it follows from the transition hypothesis in three steps.

Continuous OU to a discrete AR(1). For $\mathrm{d}z_t = -\theta z_t\,\mathrm{d}t + \sqrt{2\theta}\,\mathrm{d}W_t$, which has unit stationary variance, integrating over a step Δt and writing $\rho = e^{-\theta\Delta t}$ gives

$$z_k \;=\; \rho\, z_{k-1} + \sqrt{1-\rho^2}\,\eta_k, \qquad \eta_k \sim \mathcal{N}(0, I_d), \quad \eta_k \perp z_{k-1}.$$

Unrolling the window. Starting from $z_0 \sim \mathcal{N}(0, I_d)$ and expanding the innovations,

$$z_{t+h} \;=\; \rho^h z_t + \sqrt{1-\rho^2}\sum_{i=1}^{h} \rho^{\,h-i}\eta_{t+i} \qquad \text{for any } h \geq 0.$$

The cross-covariance. Multiplying by $z_t^\top$ and taking expectations, the future innovations are independent of z_t, so every term of the sum vanishes:

$$\mathrm{Cov}(z_{t+h}, z_t) \;=\; \rho^h \underbrace{\mathbb{E}[z_t z_t^\top]}_{I_d} + \sqrt{1-\rho^2}\sum_{i=1}^{h} \rho^{\,h-i} \underbrace{\mathbb{E}[\eta_{t+i} z_t^\top]}_{0} \;=\; \rho^h I_d,$$

and by stationarity $\mathrm{Cov}(z_s, z_t) = \rho^{|s-t|} I_d$ for any pair of instants. Stacking the window $\mathbf{Z}_{t,h} = [z_t^\top, \ldots, z_{t+h}^\top]^\top \in \mathbb{R}^{(h+1)d}$, the block covariance is $\Sigma_{\rho,h} = \mathbb{E}[\mathbf{Z}_{t,h}\mathbf{Z}_{t,h}^\top]$ with (t,s) block $\rho^{|t-s|} I_d$. Factoring out the spatial metric gives

$$\Sigma_{\rho,h} \;=\; R_{\rho,h} \otimes I_d, \qquad R_{\rho,h} \in \mathbb{R}^{(h+1)\times(h+1)}, \quad (R_{\rho,h})_{ts} = \rho^{|t-s|},$$

the Kac–Murdock–Szegő matrix. The only free parameter is ρ; §3.2 replaces the geometric profile by a measured one, which the maximum-entropy argument of §3.1 admits without change.

A.2 Where the factor of 128 comes from

The constant that accompanies Propositions 1 and 2 is computed in closed form on both sides: no trained model and no Monte Carlo enters it. We fix $h+1 = 4$, $d = 192$ and a purely temporal departure A with ones on the first off-diagonals, so $\mathrm{tr}\, A = 0$ and $\|A\|_F^2 = 6$. The same fourth-moment identity serves twice, once on $\mathbb{S}^{n-1}$ with $n = (h+1)d$ and once on $\mathbb{S}^h$, and the traceless hypothesis kills the scale term in both:

$$\mathbb{E}_a\big[(a^\top \Delta\, a)^2\big] = \frac{2\,d\,\|A\|_F^2}{n(n+2)}, \qquad \mathbb{E}_q\big[(q^\top A\, q)^2\big] = \frac{2\,\|A\|_F^2}{(h+1)(h+3)} = \tfrac{1}{2}.$$

The ratio therefore has no free constant and no dependence on ρ,

$$\frac{\mathbb{E}_q\big[(q^\top A\, q)^2\big]}{\mathbb{E}_a\big[(a^\top \Delta\, a)^2\big]} \;=\; \frac{n(n+2)}{d\,(h+1)(h+3)} \;=\; \frac{(h+1)d+2}{h+3} \;=\; \frac{770}{6} \;\approx\; 128,$$

and Figure 6 traces both sides as d varies: the separable value is flat because it does not depend on d at all, and the full-sphere one falls as $1/d$, which is Propositions 1 and 2 drawn rather than stated.

Why we quote this ratio and not the full one. The display above compares the fourth-moment terms, that is the numerators of Definition 1, and not the sensitivities themselves. Evaluating sens with both denominators gives 4.03×10^{-3} for the full sphere and 10.78 for the separable design, a ratio of about 2700. The two numbers differ for a reason that has nothing to do with d. The full-sphere denominator $a^\top (R_{\rho,h} \otimes I_d) a$ averages over $(h+1)d = 768$ coordinates and concentrates, so removing it changes that side by 3%; the separable denominator $q^\top R_{\rho,h}\, q$ varies over $\mathbb{S}^h$ and, since $R_{\rho,h}$ has condition number 57.5 at $\rho = 0.9$, its reciprocal squared is strongly right-skewed — bounded, since $q^\top R_{\rho,h}\, q \geq \lambda_{\min}(R_{\rho,h}) > 0$, but far from symmetric: the ratio has median 0.28 against mean 10.8, and the largest one percent of draws contributes a third of that mean. Part of the full sensitivity ratio is therefore the separable design's freedom to place mass where the *target* variance is small, which is an advantage of a different kind. Reporting 128 isolates the dilution in d, which is what the two propositions are about, and is the conservative of the two figures.

A.3 Proof of Proposition 4

Both halves are immediate once the right object is named, which is why the mixture is the natural repair and not a lucky one.

The support claim is set arithmetic. For $\eta > 0$ the measure μ_η assigns mass $\eta/(h+1) > 0$ to each e_t, so $e_t \in \mathrm{supp}\, \mu_\eta$ for every t and the hypothesis of Proposition 3(ii) is met exactly, not in a limit.

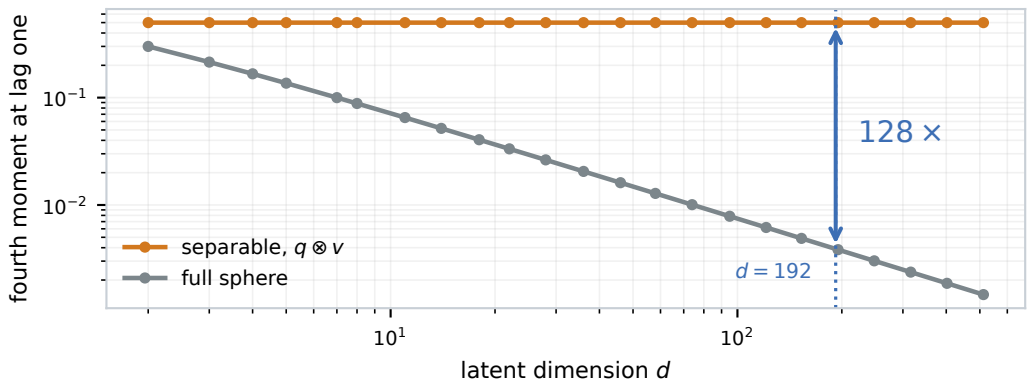


Figure 6: The fourth-moment term of Definition 1 at a lag-one temporal departure, for the two ways of drawing a direction. Both are closed forms of the same identity: $2\|A\|_F^2/((h{+}1)(h{+}3))$ on $\mathbb{S}^h$, flat because it does not involve d, and $2\|A\|_F^2 d/(n(n{+}2))$ on $\mathbb{S}^{n-1}$, falling as $1/d$. Neither involves a trained model or a simulation. At $d = 192$ the gap is the $128\times$ quoted in §4, and Figure 4 shows the two ways of drawing a direction that produce it.

The information claim rests on the temporal information operator being *linear* in the design. Writing $G_\mu = \mathbb{E}_{q\sim\mu}[g(q)]$ for the positive semidefinite $g(q)$ that the criterion accumulates per direction, linearity of the expectation in μ gives

$$G_{\mu_\eta} \;=\; (1-\eta)\, G_{\mu^\star} \;+\; \eta\, G_{\mathrm{Unif}\{e_t\}}.$$

The second term is positive semidefinite, so $G_{\mu_\eta} \succeq (1-\eta) G_{\mu^\star}$ in the Loewner order, and $\lambda_{\min}$ is monotone with respect to it:

$$\lambda_{\min}(G_{\mu_\eta}) \;\geq\; (1-\eta)\, \lambda_{\min}(G_{\mu^\star}). \qquad \blacksquare$$

The inequality is in fact an equality on the temporal subspace, and the reason is the kernel computation of §A.8: a marginal sketch $e_t \otimes v$ reads only the diagonal block Δ_{tt}, so it contributes *nothing* to the temporal directions and $G_{\mathrm{Unif}\{e_t\}}$ restricted to them is zero. The second term therefore vanishes there and $\lambda_{\min}$ falls by exactly $1-\eta$, which we confirm numerically: $2.874 \to 2.730$ at $\eta = 0.05$ and $2.874 \to 2.586$ at $\eta = 0.1$, both to three decimals.

A.4 The support trichotomy, verified without a network

The trichotomy is a statement about the population criterion, so it can be checked without training anything. We parametrise the window covariance directly and minimise the population loss, at $h+1 = 4$, $d = 8$ and $\rho = 0.9$. To isolate the variable, each family is run with and without e_t in the support, the second obtained by truncating the caps $|\cos(q, e_t)| > 0.9$. All rows use the same optimisation budget, 3000 Adam steps, which matters for reading the last one.

The seventh row needs reading with care, because it is a stopping point and not a minimum. At the shared budget the unmixed optimum sits at a loss of 1.4×10^{-6}, with all four instants at variance 2.3331 and every lag-one covariance at 2.2324. Both follow from one number: its channel-averaged departure is $c\,J$ with $c = 1.3333$ and $\|A - cJ\|/\|A\| = 0.005$, lying in the common-mode subspace to a fraction of 1.0000. So each variance reads $1 + 1.33$ and each lag-one covariance $0.9 + 1.33$.

It is still descending. The criterion is nearly flat along J, and continuing the optimisation brings c to 0.40 at 12000 steps and 0.14 at 25000, with the loss falling to 3.5×10^{-8} and the spatial error to 0.40, without reaching zero. The mixture reaches 4.6×10^{-16} and $\Delta_A = 0$ inside the original 3000. What the row demonstrates is therefore the *rate* at which the blind direction gets corrected, which is what Proposition 5(iii) bounds, and not a law sitting in the null of the criterion.

One apparent mismatch is worth resolving. The spatial error 1.8574 exceeds $c = 1.3333$ because Σ_{tt} is not isotropic: only half of $\|\Delta\|^2$ lies in the $A \otimes I_d$ subspace, and $1.8574^2 - 1.3333^2$ is the remainder. The toy parametrises the full $(h{+}1)d$ covariance, so nothing forces the departure to be uniform across channels, and it is not.

Table 3: The population criterion minimised without a network, at a fixed budget of 3000 Adam steps for every row. The spatial error is $\|\Sigma_{tt} - I_d\|/\|I_d\|$ averaged over t, so 0 means $z_t \sim \mathcal{N}(0, I_d)$ exactly. The last column is the realised lag-one *covariance*, not a correlation: its target is ρ, which is why five rows read 0.900, and it is unnormalised, which is why the failing row exceeds 1. Six of the seven designs reach the target within the budget, including both truncated ones, so removing e_t from the support costs nothing on its own.

design	support	e_t inside	$\|\Sigma_{tt} - I_d\|$	mean var.	lag-one cov.
marginal	—	yes	0.0000	1.0000	0.055
full sphere	continuous	yes	0.0000	1.0000	0.900
full sphere, truncated	continuous	**no**	0.0000	1.0000	0.900
separable uniform	continuous	yes	0.0000	1.0000	0.900
separable uniform, truncated	continuous	**no**	0.0000	1.0000	0.900
separable mixture $\mu_{\eta=0.1}$	discrete	yes	0.0000	1.0000	0.900
separable optimal, unmixed	discrete	**no**	**1.8574**	**2.3331**	2.2324

A.5 Why sens_ν is the loss to second order

Under $\Sigma = \Sigma_{\rho,h} + \Delta$, the standardised projection $\tilde{s} = a^\top \mathbf{Z}_{t,h} / \sqrt{a^\top \Sigma_{\rho,h} a}$ is Gaussian with variance

$$v(a) \;=\; 1 + \frac{a^\top \Delta\, a}{a^\top \Sigma_{\rho,h}\, a}.$$

The population characteristic-function criterion for a centred Gaussian of variance v against the standard normal is smooth in v and stationary at $v = 1$, so expanding it there gives $\mathrm{EP}(v) = c_w (v-1)^2 + O\big((v-1)^3\big)$ with $c_w > 0$ fixed by the frequency grid and weights. Taking expectations over $a \sim \nu$,

$$\mathcal{L}_\nu \;=\; c_w\, \mathrm{sens}_\nu(\Delta) + O(\|\Delta\|^3),$$

which is Definition 1. The sensitivity is therefore the loss itself to leading order and not a surrogate chosen for convenience. Every comparison in this paper is between two designs read under the same grid, so c_w cancels and its value never enters an argument.

A.6 Proof of Proposition 1

Write $n = (h+1)d$. Two facts combine.

The fourth moment on the sphere, exactly. For a uniform on $\mathbb{S}^{n-1}$ and Δ symmetric,

$$\mathbb{E}\big[(a^\top \Delta\, a)^2\big] \;=\; \left(\frac{\operatorname{tr}\Delta}{n}\right)^2 + \frac{2\|\tilde{\Delta}\|_F^2}{n(n+2)},$$

where $\tilde{\Delta} = \Delta - (\operatorname{tr}\Delta/n) I_n$ is the traceless part. For a purely temporal departure $\Delta = A \otimes I_d$ with $\operatorname{tr} A = 0$ we have $\operatorname{tr}\Delta = 0$, which kills the scale term, and $\|\Delta\|_F^2 = d\|A\|_F^2$, leaving

$$\mathbb{E}\big[(a^\top \Delta\, a)^2\big] \;=\; \frac{2\,d\,\|A\|_F^2}{n(n+2)}.$$

The random denominator, sandwiched. Write $a = g/\|g\|$ with $g \sim \mathcal{N}(0, I_n)$; the ratio in sens_ν is invariant to the scale of g. The eigenvalues of $\Sigma_{\rho,h} = R_{\rho,h} \otimes I_d$ are those of $R_{\rho,h}$, each repeated d times, so

$$\lambda_{\min}(R_{\rho,h})\|g\|^2 \;\le\; g^\top (R_{\rho,h} \otimes I_d)\, g \;\le\; \lambda_{\max}(R_{\rho,h})\|g\|^2.$$

Because the denominator is now enclosed between two deterministic multiples of $\|g\|^2$, both bounds reduce to the previous display:

$$\frac{2\,d\,\|A\|_F^2}{\lambda_{\max}^2\, n(n+2)} \;\le\; \mathrm{sens}_{\mathrm{full}}(\Delta) \;\le\; \frac{2\,d\,\|A\|_F^2}{\lambda_{\min}^2\, n(n+2)} \;=\; \Theta\!\left(\frac{\|A\|_F^2}{d}\right),$$

since $n(n+2) = \Theta(d^2)$. The constants depend only on the spectrum of $R_{\rho,h}$ and not on A, which is what makes the rate uniform over the temporal subspace: no direction of it is favoured, so there is no low-rank part of the departure that survives the dilution. ■

A.7 Proof of Proposition 2

The separable sketch needs no sandwich, and that is the whole difference. With $a = q \otimes v$ and v a unit vector, the mixed-product identity $(A \otimes B)(u \otimes w) = (Au) \otimes (Bw)$ and the factorisation of the inner product $(u \otimes w)^\top (u' \otimes w') = (u^\top u')(w^\top w')$ give

$$a^\top \Delta\, a = (q^\top A\, q)(v^\top v) = q^\top A\, q, \qquad a^\top \Sigma_{\rho,h}\, a = (q^\top R_{\rho,h}\, q)(v^\top v) = q^\top R_{\rho,h}\, q.$$

Both are free of v and of d: the denominator is *exact* rather than enclosed. Hence

$$\mathrm{sens}_{\mathrm{sep}}(\Delta) = \mathbb{E}_q\left[\frac{(q^\top A\, q)^2}{(q^\top R_{\rho,h}\, q)^2}\right],$$

and applying the fourth-moment identity on $\mathbb{S}^h$ instead of $\mathbb{S}^{n-1}$ gives $\Theta(\|A\|_F^2)$, with constants depending only on $R_{\rho,h}$. Dividing by Proposition 1 gives a ratio of $\Theta(d)$. ■

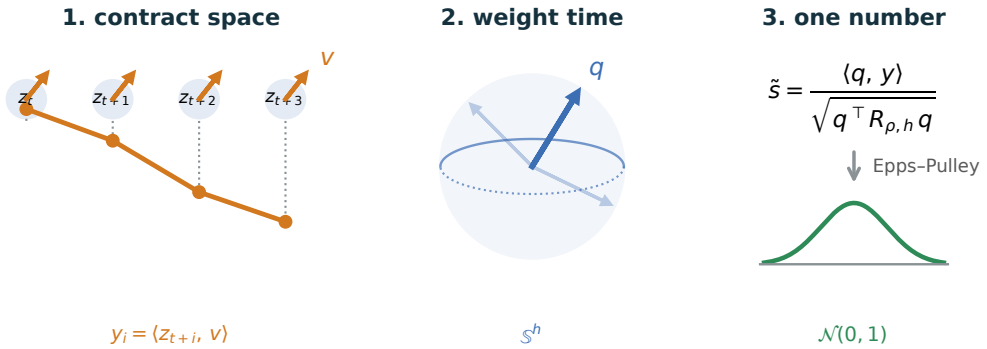


Figure 7: A separable sketch, in three stages: one shared channel direction v turns the window into $h+1$ numbers, a temporal profile q weights them, and the result is a single number that is standard normal under the target.

The coordinates of equation 2. The middle step of equation 2 is only $x^2 = x\,x^\top$ for a scalar, together with α_A being constant in q so that it leaves the expectation. That is the whole reason for passing from the matrix form to the coordinates $u(q)$ and α_A: in the matrix form A and q sit inside the same quadratic form and the expectation has nowhere to go. Since A is symmetric with zero diagonal, only the $h(h{+}1)/2$ off-diagonal entries of $qq^\top/(q^\top R_{\rho,h}\, q)$ are ever paired against it, and α_A is A itself written as a list of those entries rather than as a matrix. Both live in the same $\mathbb{R}^{h(h+1)/2}$, one read off the error and one off the profile, which is why they pair.

The sandwich loosens as ρ grows, because $R_{\rho,h}$ becomes poorly conditioned and $\lambda_{\max}/\lambda_{\min}$ widens, so this argument pins down the order in d but not the constant. The constant is available anyway, in closed form and independently of ρ: §A.2 gives $n(n{+}2)/\big(d(h{+}1)(h{+}3)\big) = 770/6 \approx 128$ at $d = 192$.

The separable sketch is blind to a subspace, exactly and not merely weakly, but the blindness is spatial rather than temporal. A purely temporal departure is a difference of covariances and therefore symmetric, so there is no antisymmetric temporal departure to be blind to. What the sketch cannot see is the antisymmetric part of the spatial blocks: with $a = q \otimes v$ the statistic reads $\sum_{s,t} q_s q_t\, v^\top \Delta_{st}\, v$, and the quadratic form $v \mapsto v^\top \Delta_{st}\, v$ annihilates $\Delta_{st} - \Delta_{st}^\top$ before the temporal weights $q_s q_t$ ever act. This is exactly the gap Lemma 1 leaves open, which pins the symmetric part of every block and says nothing about the antisymmetric part of the cross blocks — the part a rotation of the latent lives in. This is a real cost of the construction and not an artefact of the bound. It is also not protected by any assumption we can make, because trained encoders do rotate: under the marginal-only objective the antisymmetric fraction of the lag-one covariance is 0.206 on Push-T ($n = 6$ runs) and 0.209 on Cube ($n = 2$). What we observe, and cannot explain, is that the separable designs *reduce* that fraction despite being provably blind to the subspace: 0.140 on Push-T over 47 runs and 0.137 on Cube over 53, a factor of 1.5 in both suites (the full-sphere designs do not: 0.244 and 0.304). The effect has to be indirect, and we leave it open.

A.8 A LATENT THAT SATISFIES THE OBJECTIVE AND HAS NO TEMPORAL STRUCTURE

The claim of §3 is exhibited rather than argued, and it is a claim about the objective, not about any particular encoder: both training terms are functions of the latent tensor Z alone, so it is enough to name a Z that minimises both. Take the latent that is constant along each episode,

$$z_{t+i} \;=\; z \qquad \text{for } i = 0, \dots, h \text{ and every } t,$$

with one vector z per episode and different episodes mapping to different z. Windows overlap, so constancy on every window is the same thing as constancy along the episode. Both training terms attain their minimum at once.

The prediction term is exactly zero. Since $z_{t+i+1} = z_{t+i}$, the identity map is an optimal predictor and $\|\hat{z}_{t+i+1} - z_{t+i+1}\|^2 = 0$ at every step, with no approximation.

The marginal regulariser is satisfied too. At each fixed t the distribution *over the batch* is the distribution of z across episodes, which can be exactly $\mathcal{N}(0, I_d)$; the test is applied per instant and passes at every one of them. Nothing in it compares two instants, so nothing in it can see that the window is constant.

The reason is the kernel of the marginal mechanism, and it is exact rather than asymptotic. For a marginal sketch $a = e_t \otimes v$ the quadratic form is $a^\top \Delta\, a = v^\top \Delta_{tt}\, v$, which involves only the diagonal blocks. The off-diagonal blocks Δ_{ts}, $t \neq s$, do not enter any marginal sketch, so

$$\ker(G_{\text{marg}}) \;=\; \{\Delta : \Delta_{tt} = 0 \text{ for all } t\},$$

and the constant window differs from the target precisely in those free blocks. The planner then has nothing to plan with, and planning success in such a latent is at chance.

Whether an encoder acting frame by frame can realise this Z is a separate question, and in general it cannot: two episodes that pass through the same observation must share its latent, which forces them to share z. That is not the point. The point is that the set of global minimisers of the loss, read as a function of Z, contains latents with no temporal structure, so the objective places no obstacle between the encoder and them; what keeps trained models away is the architecture and the optimiser, not the loss. Every claim that follows about what a design *forces* is likewise a claim about the minimisers of the loss in Z.

A.9 PROOF OF THEOREM 1

($\Leftarrow$) is immediate: under the target law every standardised projection is standard normal, so the statistic vanishes for every a and so does its average.

($\Rightarrow$) has three steps, and the middle one is the one worth naming. The criterion is $\mathbb{E}_a[\mathrm{EP}(a)]$ with $\mathrm{EP} \geq 0$ pointwise, so it vanishing forces $\mathrm{EP}(a) = 0$ for almost every a. Wherever that happens, the Epps–Pulley statistic characterises normality (Epps & Pulley, 1983), so the standardised projection $\tilde{s} = a^\top \mathbf{Z}_{t,h} / \sqrt{a^\top \Sigma_{\rho,h} a}$ is $\mathcal{N}(0, 1)$.

Almost every a is not yet every a, and the gap is closed by continuity. The population characteristic function is continuous in a by dominated convergence, since $|e^{\mathrm{i}\theta}| = 1$, and EP integrates $|\varphi - \varphi_0|^2 w$ against an integrable weight, so EP is continuous in a. A continuous non-negative function that vanishes almost everywhere on its support vanishes on *all* of it, and for $a \sim \mathrm{Unif}(\mathbb{S}^{(h+1)d-1})$ the support is the whole sphere.

Every one-dimensional projection of $\mathbf{Z}_{t,h}$ is therefore normal with the target variance, and Cramér–Wold in $\mathbb{R}^{(h+1)d}$ (Cramér & Wold, 1936) upgrades that to $\mathbf{Z}_{t,h} \sim \mathcal{N}(0, \Sigma_{\rho,h})$.

For the marginal corollary, the point is that no direction has to be added. The directions $e_t \otimes v$ with $v \in \mathbb{S}^{d-1}$ already lie on $\mathbb{S}^{(h+1)d-1}$, so the conclusion applies at them. There

$$(e_t \otimes v)^\top \Sigma_{\rho,h}\, (e_t \otimes v) \;=\; (e_t^\top R_{\rho,h}\, e_t)(v^\top v) \;=\; (R_{\rho,h})_{tt} \;=\; 1,$$

so $\langle z_t, v \rangle \sim \mathcal{N}(0, 1)$ for every $v \in \mathbb{S}^{d-1}$, and Cramér–Wold in $\mathbb{R}^d$ gives $z_t \sim \mathcal{N}(0, I_d)$. ■

A.10 Proof of Theorem 2

The separable family is a set of measure zero in $\mathbb{S}^{(h+1)d-1}$ — the sketches $q \otimes v$ form a Segre variety of dimension $h + d - 1$, which is 194 against the sphere's 767 at $h + 1 = 4$ and $d = 192$ — so the hypothesis of Theorem 1 is unavailable and the marginal law has to be recovered by a different route. It survives, and the step that carries it is the second.

First, $\mathbb{E}_{q,v}[\mathrm{EP}] = 0$ with $\mathrm{EP} \geq 0$ gives $\mathrm{EP}(q, v) = 0$ for almost every (q, v), and there the standardised projection is $\mathcal{N}(0, 1)$ by Epps–Pulley.

Second, $\mathrm{EP}(q, v)$ is continuous in (q, v), for the reason given in the proof of Theorem 1. A continuous function vanishing almost everywhere vanishes on the entire support, and for q uniform on $\mathbb{S}^h$ with v uniform on $\mathbb{S}^{d-1}$ that support is the whole product of spheres.

Third, the conclusion therefore holds at $q = e_t$, which *is* in the support. There $e_t^\top R_{\rho,h}\, e_t = (R_{\rho,h})_{tt} = 1$, so the standardised projection is $\langle z_t, v\rangle$ itself and $\langle z_t, v\rangle \sim \mathcal{N}(0, 1)$ for every $v \in \mathbb{S}^{d-1}$.

Fourth, Cramér–Wold in $\mathbb{R}^d$ closes: $z_t \sim \mathcal{N}(0, I_d)$ for every t. ■

It is worth being explicit about what this is and is not. It is a conditional whose premise is the criterion *at* zero. In training we draw M directions per step and, under a uniform q, none of them lands on e_t, so the guarantee exists while the way to collect it does not, which is what the mixture of §4 responds to.

A.11 Proof of Proposition 3

Part (i) is proved in the body: a shift $z_{t+i} \mapsto z_{t+i} + \xi$ shared across the window moves the sketch by $\langle \xi, v\rangle \sum_t q_t$, which is zero for every atom of a zero-sum design, so the law of every standardised sketch and hence the criterion are unchanged, while $\mathrm{Cov}(z_t)$ moves by $\mathrm{Cov}(\xi)$. Part (ii) and its quantitative refinement are what this appendix proves. Its three cases below are cases of Proposition 3(ii) — what the support has to be for the covariance to be forced; parts (i)–(ii) of the refinement follow from them, and part (iii) is proved separately below. The refinement reads the margin off one operator:

Proposition 5 (quantitative form of Proposition 3)**.** *Let μ be a design on $\mathbb{S}^h$ and let $\Gamma_\mu = \mathbb{E}_{q\sim\mu}\big[\gamma(q)\gamma(q)^\top\big]$ with $\gamma(q) = \mathrm{svec}\big(qq^\top/(q^\top R_{\rho,h}\, q)\big)$, so that Γ_μ acts on all of $\mathrm{Sym}(h{+}1)$. Write $\mathcal{L}_\mu$ for the population value of equation 1 under μ.*

- *(i) $\mathrm{Cov}(z_t) = I_d$ at every t is forced if $\lambda_{\min}(\Gamma_\mu) > 0$, that is if $\{qq^\top : q \in \mathrm{supp}\,\mu\}$ spans $\mathrm{Sym}(h{+}1)$. A support with non-empty interior satisfies this, and so does a finite support of at least $(h{+}1)(h{+}2)/2$ atoms in general position; membership of the e_t is required for neither. The condition is sufficient and not necessary: a null space consisting only of matrices with zero diagonal would leave $\mathrm{Cov}(z_t) = I_d$ intact.*
- *(ii) The law $z_t \sim \mathcal{N}(0, I_d)$ is forced if in addition every e_t is an atom of μ. Our argument needs the atom; whether the law can fail with the covariance intact when no e_t is sampled is open.*
- *(iii) The error at a finite criterion value is governed by $\lambda_{\min}(\Gamma_\mu)$ and not by whether it is positive. To second order in Δ, writing Δ_{sp} for the diagonal blocks $(\Delta_{tt})_t$,*

$$\|\Delta_{\mathrm{sp}}\|_F^2 \;\leq\; \frac{d(d{+}2)}{2} \cdot \frac{\mathcal{L}_\mu}{c_w\, \lambda_{\min}(\Gamma_\mu)}. \tag{5}$$

The factor $d(d{+}2)/2$ is not slack: a separable sketch reads a diagonal block only through the scalar $v^\top \Delta_{tt} v$, so averaging over the channel sphere costs what Proposition 1 costs, and the purely temporal analogue of equation 5 carries no such factor because there $B_v = A$ for every v.

A zero-sum design has $\gamma(q) \perp \mathrm{svec}(J)$ for every atom, so J lies in the null space of Γ_μ and equation 5 is vacuous along it; that is the body's part (i) read through Γ_μ.

Proof of part (iii). Three steps, and the middle one is where d enters. Write $(B_v)_{st} = v^\top \Delta_{st}\, v$, so that for a separable sketch $a = q \otimes v$ the ratio inside Definition 1 is an inner product on $\mathrm{Sym}(h{+}1)$,

$$\frac{a^\top \Delta\, a}{a^\top \Sigma_{\rho,h}\, a} \;=\; \frac{q^\top B_v\, q}{q^\top R_{\rho,h}\, q} \;=\; \big\langle \gamma(q),\, \mathrm{svec}(B_v) \big\rangle,$$

because svec is chosen to preserve inner products. Taking the expectation over $q \sim \mu$ first,

$$\mathrm{sens}_{\mu\otimes\mathrm{Unif}}(\Delta) \;=\; \mathbb{E}_v\big[\mathrm{svec}(B_v)^\top \Gamma_\mu\, \mathrm{svec}(B_v)\big] \;\geq\; \lambda_{\min}(\Gamma_\mu)\, \mathbb{E}_v\big[\|B_v\|_F^2\big]\,.$$

Second, drop every off-diagonal entry of B_v and apply the fourth moment on $\mathbb{S}^{d-1}$ to each diagonal one. For symmetric S and v uniform,

$$\mathbb{E}_v\big[(v^\top S v)^2\big] \;=\; \frac{2\|S\|_F^2}{d(d{+}2)} + \frac{(\mathrm{tr}\, S)^2}{d(d{+}2)} \;\geq\; \frac{2\|S\|_F^2}{d(d{+}2)},$$

the same identity as in the proof of Proposition 1, regrouped so that both terms are manifestly non-negative. Summing over t, $\mathbb{E}_v[\|B_v\|_F^2] \geq 2\|\Delta_{\mathrm{sp}}\|_F^2/(d(d{+}2))$. Third, substitute $\mathcal{L}_\mu = c_w\, \mathrm{sens}_\mu(\Delta) + O(\|\Delta\|^3)$ from §A.5 and rearrange, which is equation 5. ■

The constant matters and we checked it rather than trusting the algebra: at $h{+}1 = 4$, $d = 8$ and a purely spatial Δ, the version of equation 5 without $d(d{+}2)/2$ is *violated* by a factor of 1.33, and with it holds with room to spare. Dropping the off-diagonal entries of B_v is what makes the bound loose rather than tight; it is the direction that keeps the conclusion conservative.

Everything rests on one observation: for separable sketches the discrepancy is a polynomial in the temporal direction, and polynomials cannot vanish on an open set without vanishing everywhere.

Lemma 1 (polynomial rigidity)**.** *Let Δ be symmetric with blocks $\Delta_{st} \in \mathbb{R}^{d\times d}$ and write*

$$\delta(q, v) \;:=\; (q \otimes v)^\top \Delta\, (q \otimes v) \;=\; \sum_{s,t} q_s q_t\, v^\top \Delta_{st}\, v.$$

If $Q \subseteq \mathbb{S}^h$ has non-empty interior and $\delta(q, v) = 0$ for every $q \in Q$ and every $v \in \mathbb{S}^{d-1}$, then the symmetric part of every *block vanishes: $\Delta_{tt} = 0$ for all t, and $\Delta_{st} + \Delta_{st}^\top = 0$ for all $s \neq t$.*

Proof. Fix $v \in \mathbb{S}^{d-1}$. The map $q \mapsto \delta(q, v)$ is a homogeneous quadratic polynomial on $\mathbb{R}^{h+1}$. The cone $\{\lambda q : \lambda > 0,\ q \in Q\}$ is a non-empty open subset of $\mathbb{R}^{h+1}$, and by hypothesis and homogeneity $\delta(\cdot, v)$ vanishes on it. A real polynomial vanishing on a non-empty open subset of $\mathbb{R}^{h+1}$ vanishes identically, so $\delta(\cdot, v) \equiv 0$; in particular $\delta(e_t, v) = v^\top \Delta_{tt}\, v = 0$. As v was arbitrary and Δ_{tt} is symmetric, $\Delta_{tt} = 0$. Evaluating instead at $q = (e_t + e_s)/\sqrt{2}$ and subtracting the two diagonal terms gives $v^\top(\Delta_{st} + \Delta_{st}^\top)v = 0$ for every v, so the symmetric part of each cross block vanishes as well. □

The full conclusion is worth stating because it locates the blind spot exactly. A separable criterion at zero determines Δ up to the *antisymmetric* parts of its cross blocks, and nothing more is missing. That is the same subspace §A.7 shows the sketch is blind to, reached here from the opposite direction.

What Lemma 1 does *not* require is that the e_t belong to Q. A design may place zero mass in a neighbourhood of every marginal direction, so that they lie outside the support and outside its closure, and still determine the diagonal blocks, provided what remains is fat enough to carry the polynomial argument. The operative property is the dimension of the support, not membership.

Spanning, continuous support. At the null the standardised sketch has unit variance for almost every a, hence, by the continuity step of Theorem 2, for every a in the support, so $a^\top \Delta\, a = 0$ for every $q \in \mathrm{supp}\,\mu$ and every v. Applying Lemma 1 with $Q = \mathrm{supp}\,\mu$ gives $\Delta_{tt} = 0$, which is $\mathrm{Cov}(z_t) = (R_{\rho,h})_{tt} I_d = I_d$.

Spanning, plus the marginal directions. If additionally $e_t \in \mathrm{supp}\,\mu$, the null applies at $a = e_t \otimes v$ for every $v \in \mathbb{S}^{d-1}$. There the target variance is $(R_{\rho,h})_{tt} = 1$, so Epps–Pulley forces $\langle z_t, v\rangle \sim \mathcal{N}(0, 1)$ for every v, and Cramér–Wold in $\mathbb{R}^d$ gives $z_t \sim \mathcal{N}(0, I_d)$. The covariance conclusion is about

a moment and follows from a polynomial identity; the law is a distributional statement, and our argument for it needs the marginal directions to be sampled. That is why the two are separated.

Finite supports. A finite support has empty interior, so Lemma 1 does not apply, but a finite design avoiding the e_t still retains the covariance, by the second route of case (i): each atom q_j imposes one linear equation $q_j^\top B_v\, q_j = 0$ on $B_v \in \mathrm{Sym}(h{+}1)$, where $(B_v)_{st} = v^\top \Delta_{st} v$. Once $\{q_j q_j^\top\}$ spans $\mathrm{Sym}(h{+}1)$ — which needs $(h{+}1)(h{+}2)/2 = 10$ atoms in general position at $h{+}1 = 4$ — the system forces $B_v = 0$ for every v, hence $\Delta_{tt} = 0$. The optimal design's 177 atoms are uniform draws, so they span with probability one, and the realised rank is 10 of 10 at every ρ we solve for.

What fails is the conditioning. Writing Γ_μ as in Proposition 5, its spectrum at $\rho = 0.9$ splits cleanly into six directions orthogonal to the common-mode subspace of §C, with eigenvalues from 2.9 to 11.4, and the four common-mode directions themselves, with eigenvalues from 4×10^{-7} to 2.2×10^{-3}; §C prints all ten. The faint four are not the diagonal block: they cross the diagonal and off-diagonal blocks, which is why the diagnostic has to be read on all of $\mathrm{Sym}(h{+}1)$. Reading equation 5 against its temporal analogue then says what that gap buys the encoder. At one and the same criterion value the bound permits a common-mode spatial departure whose Frobenius norm is $\sqrt{\frac{d(d+2)}{2}\,\lambda_{\mathrm{temp}}/\lambda_{\min}} \approx 4{\times}10^5$ times the largest temporal departure it permits, and two independent factors compose to give it: the spectral gap contributes $\sqrt{2.9/4 \times 10^{-7}} \approx 2.7 \times 10^3$ and the channel averaging of part (iii) another $\sqrt{d(d+2)/2} \approx 137$. Along the other three common-mode directions, whose eigenvalues are near 10^{-3}, the same reading gives about 5×10^3. Table 3 measures the consequence rather than a failure of identifiability: minimising the population loss over covariances directly leaves the unmixed optimum at mean variance 2.33 instead of 1, because the optimiser stops long before a direction with $\lambda \approx 10^{-7}$ has moved. ■

That the optimal design evacuates the marginal directions is structural rather than accidental. By the kernel computation of §A.8 the direction e_t contributes nothing to the temporal subspace, so a criterion that is purely temporal assigns it zero marginal value and a solver maximising $\lambda_{\min}$ on that subspace evacuates it; the measured mass of the solved design is zero above $|\cos(q, e_t)| = 0.9$. Optimising a criterion restricted to the temporal block therefore produces a design whose spatial margin is not merely small but numerically absent, and the opposition is invisible to anyone reading the temporal criterion alone. The mixture is the repair because putting mass η on the e_t is exactly what lifts $\lambda_{\min}(\Gamma_\mu)$ on the four directions the temporal criterion cannot see.

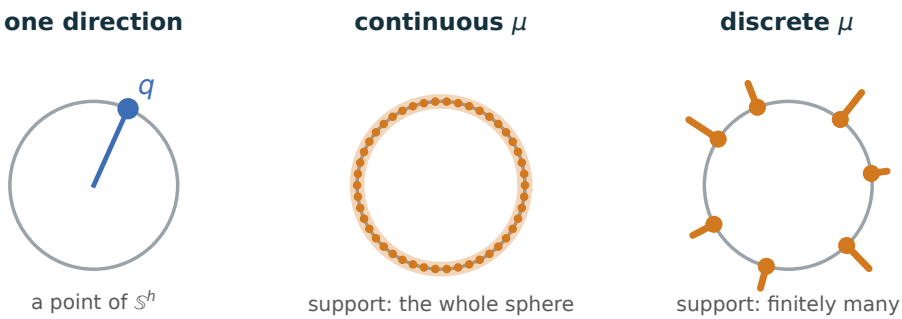


Figure 8: A design is a measure on the sphere of temporal profiles, and all three panels show the same sphere. A single direction is one of its points; a continuous measure puts mass everywhere on it; a discrete one puts mass on finitely many of its points, drawn here as spikes whose length is the mass of that atom. The atoms lie *on* the sphere — a design cannot place mass off it. What Proposition 3 turns on is not which of the three a design is but how well the points it does occupy cover the space of departures: a continuous measure and a finite one with enough atoms both pin down the covariance, and what separates them is the margin $\lambda_{\min}(\Gamma_\mu)$. Our proof of the law needs the marginal directions to be atoms in either case.

Algorithm 1 The path term, for one batch of windows. Nothing of size $(h+1)d$ is ever formed. The contraction in equation 1 is taken per batch element: $\mathbf{Z}_{t,h}$ denotes one window there, and $\tilde{s}^{(m)}$ is the sample of B scalars that T consumes.

Input: latents $Z \in \mathbb{R}^{(h+1)\times B\times d}$; target profile $R_{\rho,h}$; directions M; frequency grid $\omega_1 \dots \omega_K$ with weights w_k
Output: the scalar $\text{T-SIGReg}_{\mu_\eta}(Z)$

$v^{(1)} \dots v^{(M)} \sim \text{Unif}(\mathbb{S}^{d-1})$ ▷ channel directions, one per projection
$q^{(1)} \dots q^{(M)} \sim \mu_\eta$ on $\mathbb{S}^h$ ▷ temporal profiles, from the design
for $m = 1 \dots M$ **do**
  $y_i \leftarrow Z_i v^{(m)}$ for $i = 0 \dots h$ ▷ stage one: contract space, the same $v^{(m)}$ at every instant
  $s^{(m)} \leftarrow \langle q^{(m)}, y \rangle$ ▷ stage two: contract time; $q^{(m)} \otimes v^{(m)}$ is never built
  $\sigma_m^2 \leftarrow q^{(m)\top} R_{\rho,h}\, q^{(m)}$ ▷ the target variance, *exact* and free of $v^{(m)}$ and of d
  $\tilde{s}^{(m)} \leftarrow s^{(m)}/\sigma_m$
end for
for $m = 1 \dots M$, $k = 1 \dots K$ **do**
  $\hat{\varphi}_{k,m} \leftarrow \frac{1}{B}\sum_b \cos\left(\omega_k \tilde{s}_b^{(m)}\right) + \mathrm{i}\,\frac{1}{B}\sum_b \sin\left(\omega_k \tilde{s}_b^{(m)}\right)$ ▷ empirical characteristic function of the m-th projection
  $e_{k,m} \leftarrow \left|\hat{\varphi}_{k,m} - e^{-\omega_k^2/2}\right|^2$ ▷ distance to the standard normal at that frequency
end for
return $B \cdot \frac{1}{M}\sum_m \sum_k w_k e_{k,m}$ ▷ $= \frac{1}{M}\sum_m T(\tilde{s}^{(m)})$ on the grid

B PROTOCOL DETAILS

Model and optimisation. A ViT-tiny encoder with patch size 14 and $d = 192$ (Dosovitskiy et al., 2021), a six-layer autoregressive predictor and an MLP projector, trained with AdamW at learning rate 5×10^{-5}, weight decay 10^{-3}, batch size 128, for 10 epochs. One step of the window is 5 environment steps.

The criterion. Window horizon $h = 3$, so $h{+}1 = 4$ instants and $\mathbf{Z}_{t,h} \in \mathbb{R}^{768}$. Directions $M = 256$, drawn afresh each step, for T-SIGReg and for SIGReg in the control alike. The Epps–Pulley statistic T is evaluated on a grid of $K = 17$ frequencies, uniform on $\omega \in [0, 3]$, and the quadrature weights are the composite trapezoidal rule on that grid times the Gaussian window $e^{-\omega^2/2}$: with $\Delta\omega = 3/(K-1)$, the weight at knot ω_k is $w_k = c_k \Delta\omega\, e^{-\omega_k^2/2}$ with $c_k = 1$ at the two endpoints and $c_k = 2$ inside. The window is what makes the integral converge; it is fixed and never tuned. The regulariser weight is $\lambda = 0.09$; no configuration reported adds SIGReg on top of T-SIGReg, which is the choice §5.1 discusses. The mixture weight is $\eta \in \{0.05, 0.1\}$ and the optimal design's atoms are precomputed once per (h, ρ) by Algorithm 2: 177 atoms at $\rho = 0.9$ and 155 at $\rho = 0.95$, from 8000 candidates and 8000 Frank–Wolfe iterations with weight threshold $\varepsilon = 10^{-5}$.

Training runs and evaluation seeds. This distinction matters for every interval in the paper. Every configuration is evaluated under six rollout seeds $\{7, 42, 123, 2024, 31337, 65537\}$; the number of independent training runs behind every cell of Table 2 is three, except the two controls, SIGReg and $\rho = 0$, which have one each. The thresholds below are sized for one training run against one, so the paired gains do not rely on averaging the control. The configuration held fixed across suites in §5, the measured profile with the mixture at $\eta = 0.05$, has its three runs in every suite, and each covers the suite's full offset grid under all six seeds. The deviation quoted in Table 2 is across evaluation seeds with the trained model held fixed, median 2.0 points across our cells. Retraining variability is a separate quantity, and estimating it needs a second training run.

The retraining thresholds. Only replicates that genuinely differ in training seed can measure retraining variability, and not all of ours do. In our early replicate families a configuration error kept the training seed fixed at 3072 across replicates, so several of our early replicate families share a seed across replicates and their spread is backend nondeterminism rather than retraining. The extreme case is a design whose target profile was read off an already-trained latent, and which we therefore do not report as a method: its three replicates share seed 3072 and span 0.12 points, less than the

evaluation noise of a single cell, so what looked like a three-replicate confirmation was one run measured three times. We read the seeds from the training records and use only the families where they differ.

Pooling the variances of those gives a retraining standard deviation of 1.35 on TwoRooms across 15 families and 1.38 on Push-T across 13, and per offset 1.85 and 1.95. Cube has an estimate of its own on 13 families (26 degrees of freedom), and Reacher on 14 (28); we report both, and still require every gain to clear the largest threshold of the four. Every count in this paragraph is generated from the records at build time, because a figure asserted over a set that keeps growing goes stale without anyone being wrong. Multiplying by $2\sqrt{2}$ — two standard errors of a difference of two independent runs — gives the thresholds we require a gain to clear: 3.8 and 3.9 points on the aggregate, 5.2 and 5.5 per offset. The per-offset value is larger by a factor between 1.2 and 1.6 rather than the $\sqrt{7}$ independence would give, because a poor training run is poor at every offset at once.

The body does not lean on the per-suite values: the headline gains are judged against the largest threshold of any suite, 3.9, and Reacher's +11.2 and Cube's +10.9, paired by seed for the configuration of Table 4, both exceed twice it. Neither conclusion is close. Push-T's is the one that turns on the estimate. Its aggregate paired gain is -0.9, interval [-2.4, +0.6], indistinguishable from zero; per offset, the three nearest goals lose by 10.4, 11.1 and 5.7 points against a per-offset threshold of 5.5, and no far-goal gain, the largest being 7.6 at offset 55, clears that threshold by two standard errors, which is why Table 4 bolds nothing on Push-T and why the reading to take from it is the change of sign across offsets and not any individual cell.

The world clock: recipe and measured values. The profile of §3.2 is measured before any training and without looking at the latent: 150 uniform stretches of 60 frames of expert trajectories are embedded with UMAP into two dimensions (McInnes et al., 2020) at its defaults — 15 neighbours, minimum distance 0.1, Euclidean metric, no PCA beforehand, seed 3072 — every second frame kept, κ_ℓ is the mean autocorrelation of the embedded coordinates at lag ℓ, and one lag is one window step of five environment steps. It is not the clock of the underlying state, which Table 5 reports for reference: that one averages positions with velocities and is lower, 0.68 against 0.98 on Reacher.

Two dimensions of UMAP, read with a linear autocorrelation, is a heuristic and we defend it as *sufficient* rather than as correct. What the profile has to supply is one number per lag on $[-1, 1]$, monotone enough that the Toeplitz matrix it builds is positive definite; it does not have to be an isometry of the state manifold, and nothing downstream reads the embedding again. The alternatives fail differently rather than better: a linear projection such as PCA preserves global distances but, on pixels, spends its leading directions on illumination and background and reports the autocorrelation of those; the underlying state, where it is available, gives the clock of the simulator rather than of what the encoder sees. We use the measured profile because it matches the best AR(1) prior wherever AR(1) helps and comes closest to the control on Push-T, where every AR(1) prior loses (§5.3), which is the only claim the experiments support; a profile that is wrong in a way the planner cannot feel is, for this purpose, not wrong. The estimator and the fallback when a profile is not a legal covariance target are in §C.7. The four profiles that trained are

	κ_1	κ_2	κ_3	$\lambda_{\min}(R)$
TwoRooms	0.921	0.815	0.692	0.037
Reacher	0.977	0.943	0.908	0.009
Cube	0.938	0.825	0.715	0.015
Push-T	0.953	0.898	0.840	0.025

All four are positive definite, so all four are admissible. They differ across worlds by up to 0.06 at lag one and 0.22 at lag three. And none is geometric: $\kappa_2 - \kappa_1^2$ runs from -0.010 on Push-T to -0.055 on Cube, against exactly zero for any AR(1).

Push-T is the case to state precisely, because it is where the measured profile and a prior come closest and where our result is weakest. It agrees with the $\rho = 0.95$ prior at the first lag, 0.953 against 0.950, and falls below it at the next two, 0.898 against 0.9025 and 0.840 against 0.857. The two rows of Table 2 therefore do not prescribe the same target: they prescribe targets that agree at lag one and separate afterwards, which is the only place a non-geometric profile can differ from a geometric one.

The latent follows: under the measured target the realised $\kappa_2 - \kappa_1^2$ on held-out windows is -0.023, against -0.010 under the $\rho = 0.95$ prior.

Provenance: the target is read back, not reconstructed. The prescribed vector does not have to be reconstructed from the launch script, because the target matrix is a registered buffer of the regulariser and is therefore serialised with the weights: every checkpoint carries `sigreg.target_temporal_corr`, the $(h{+}1) \times (h{+}1)$ matrix the loss actually consumed. We read it back, rather than inferring it, from every measured-profile checkpoint available to us — 14 runs: 5 on TwoRooms, 2 on Reacher, 4 on Cube, 3 on Push-T — and all 14 carry the buffer. Within each environment the value is unanimous, and it is identical across the mixture and uniform designs and across training replicas, which is what one expects of a target fixed before training. Each is exactly Toeplitz with unit diagonal. The profiles of §3.2 and Figure 3 are those matrices. The remaining runs were not re-read; their tags and seeds match runs we read.

What the buffer does not record is which script produced it. Two recipes were available when the targets were set — embedding the frames directly, or inserting a 64-component linear reduction first — and re-running them today reproduces the prescribed κ_1 only approximately: on Push-T the direct recipe gives 0.957 and the reduction 0.908 against a prescribed 0.953, and on TwoRooms 0.912 against a prescribed 0.921. So the residual gap is about regenerating the target, not about knowing it: the target is known exactly, and the closest available recipe differs from it by 0.004 at lag one on Push-T. This is a reproducibility caveat on the measurement pipeline, not an uncertainty in what was optimised.

One confound in the same place is worth naming, because the tag hides it. The design atoms are loaded from an asset indexed by the configuration's ρ, so the measured-clock rows, which carry $\rho = 0.9$ in their configuration, are trained against the measured R *and* the $\rho = 0.9$ atoms, while the $\rho = 0.95$ prior row uses the $\rho = 0.95$ atoms. Two things therefore differ between those two rows, not one. The atom sets are close — they share 148 of 155 atoms and the second-moment matrices they induce differ by 1.9% in relative Frobenius norm — and the $\rho = 0.9$ prior row uses the same atoms as the clock rows, so the clock-against-ρ=0.9 comparison isolates the profile at fixed atoms. The two contrasts, measured profile against the $\rho = 0.9$ prior (same atoms) and against the $\rho = 0.95$ prior (profile and atoms), paired by evaluation seed, are $+0.02$ points, $[-0.00, +0.04]$ and $+0.01$ points, $[-0.01, +0.02]$ on TwoRooms, $+2.73$ points, $[+2.04, +3.42]$ and -0.04 points, $[-0.18, +0.10]$ on Reacher, $+0.21$ points, $[-0.37, +0.78]$ and -0.48 points, $[-1.41, +0.44]$ on Cube, and $+4.67$ points, $[+3.80, +5.55]$ and $+4.68$ points, $[+4.05, +5.31]$ on Push-T. The body reports the Push-T pair and does not rest the claim on the confounded one.

Evaluation. Planning by CEM (Rubinstein, 1999) with 300 samples, 30 elites and 30 iterations on Push-T and 10 elsewhere, as in Maes et al. (2026); a planning horizon of 5 steps, replanned every 5, with action blocks of 5 environment steps. Each cell runs 100 episodes under each of the six evaluation seeds, with evaluation budget twice the goal offset throughout; an episode counts as a success if the environment's own termination flag fires at any step within the budget. Two choices differ from the evaluation script released with Maes et al. (2026), which measures one offset with a fixed budget and draws start states uniformly over all valid (episode, step) pairs: we sweep the offset, and we draw an episode uniformly and then a start step within it. The two draws coincide on Reacher and Cube, whose episodes all have the same length. Every row of Table 2, the published checkpoint included, is measured under this protocol; the budget sweep at 75, 100 and 125 is reported separately in §C.3. Goal offsets run from 15 to 75 in steps of 10, and to 135 on Reacher.

Cost of the path term. Forward and backward of the regulariser alone, batch 256, $M = 256$ directions, the same latent tensor for both, 60 iterations after 10 of warm-up, on a TITAN RTX: SIGReg takes 4.1 ms per step and peaks at 66 MiB; T-SIGReg with the mixture design at $\eta = 0.05$ takes 5.4 ms and 43 MiB, that is $1.32\times$ the time and $0.66\times$ the memory. The memory saving is the factorisation of §4: the separable sketch contracts space first and time second and never forms the $(h{+}1)d$-dimensional direction, whereas SIGReg projects the whole tensor and pools the NB values. The card is not one we train on; the ratios transfer across GPUs since both terms are timed on the same one, the milliseconds do not.

Hardware and the backend confound. Three clusters: AMD MI210 under ROCm, and NVIDIA H100 and A100 under CUDA. Because the campaign is distributed, the backend is a possible confound and we bounded it rather than assuming it away. Retraining one Push-T cell on the two backends and pairing by seed gives a difference of -2.24 points, with a 95% interval of $[-3.41, -1.06]$ over six seeds; dropping the most extreme seed leaves -1.83 with $[-2.52, -1.14]$, so the estimate does not rest on it. Two things follow. The effect is real and small: a five-point backend artefact is excluded, and zero is excluded too. And it is of the same order as the Push-T differences we report, which is why the body rests no Push-T conclusion on a difference of that size. The environments with larger effects are unaffected at this magnitude.

C Numerical companions

The computations that support claims in the body without fitting in it: the optimal design solved by Frank–Wolfe and its finite support, the measured gain in $\lambda_{\min}$ over the uniform design, the mixture bound as a function of η, and the $M = 4096$ ablation behind the rule of reporting the excess rather than the z. The toy experiment that accompanies Proposition 3 is Table 3, in Appendix A, where six of its seven designs reach mean variance exactly one.

Algorithm 2 Constructing the E-optimal design by Frank–Wolfe. The atoms are never invented: the method only redistributes weight over a fixed sample.

Input: window length $h{+}1$; target profile $R_{\rho,h}$; candidates $N = 8000$; iterations $K = 8000$
Output: a finitely supported design $\mu^\star$ on $\mathbb{S}^h$

$q_1 \dots q_N \sim \mathrm{Unif}(\mathbb{S}^h)$ ▷ the candidate questions, drawn once
$u_n \leftarrow \mathcal{B}^\top \mathrm{svec}\left(\dfrac{q_n q_n^\top}{q_n^\top R_{\rho,h}\, q_n}\right)$ for $n = 1 \dots N$ ▷ $\mathcal{B}$: the off-diagonal, i.e. purely temporal, basis
$w \leftarrow (1/N, \dots, 1/N)$ ▷ start from the uniform design
for $k = 0 \dots K-1$ **do**
  $M \leftarrow \sum_n w_n\, u_n u_n^\top$ ▷ the information the current design accumulates
  $v \leftarrow$ eigenvector of M for its smallest eigenvalue ▷ the current blind spot
  $j \leftarrow \arg\max_n (u_n^\top v)^2$ ▷ the one question that sees it best
  $\gamma \leftarrow 2/(k+2)$; $\quad w \leftarrow (1-\gamma)\, w$; $\quad w_j \leftarrow w_j + \gamma$ ▷ move mass towards it
end for
return the atoms with $w_n > 10^{-5}$, renormalised

The other margin, and why the latent does not reach $\mathcal{N}(0, I_d)$. Everything §4.3 optimises is $\lambda_{\min}(G_\mu)$, which lives on the zero-diagonal subspace: purely temporal departures. The marginal guarantee is governed by a different operator, Γ_μ of Proposition 5, which acts on all of $\mathrm{Sym}(h{+}1)$, diagonal included. They are two margins, and the body quotes only one: the mixture bound $\lambda_{\min}(G_{\mu_\eta}) \geq (1-\eta)\lambda_{\min}(G_{\mu^\star})$ says what η *costs* on the temporal axis and never said what it *buys* on the marginal one. Figure 9(a) says it. At $\eta = 0.05$ the criterion presses 863 times harder on the temporal axis than on the marginal one, and $\lambda_{\min}(\Gamma_{\mu_\eta})$ is exactly linear in η below its maximum. Against the uniform design our operating point is $7.2\times$ better temporally and 13 times *worse* marginally; matching the uniform design's marginal margin would take $\eta \approx 0.67$, which would still leave the temporal margin above it.

This resolves a discrepancy a reader will otherwise find on their own: the trained latents do not reach $\mathcal{N}(0, I_d)$, and their mean variance falls well short of one. No theorem is broken. Theorems 1 and 2 and Proposition 3 are conditionals whose premise is the criterion *at* zero, which training never attains; the statement that applies off zero is the quantitative Proposition 5(iii), and its bound carries $\lambda_{\min}(\Gamma_\mu)$ in the denominator together with a factor $d(d{+}2)/2$. At $\lambda_{\min}(\Gamma_{\mu_{0.05}}) = 0.0031$ that prefactor exceeds 10^6 and the bound is vacuous at any loss we reach. Nor is it a shortage of directions or of atoms: both quantities are properties of the population design, and M and the support size only govern how fast one approaches them.

Panel (b) is the check, and it is the reason we state this as the explanation rather than as a hypothesis. Ordering the four designs by their marginal margin orders the mean variance of the trained latent, monotonically, over 22 runs across environments: the uniform design at 0.0404 reaches 0.99 ± 0.03,

the mixture at $\eta = 0.1$ and $\eta = 0.05$ reach 0.78 ± 0.18 and 0.60 ± 0.18, and $\mu^\star$, whose marginal margin is exactly zero, reaches 0.40 ± 0.03. A design property computed before training predicts a property of the latent after it.

The claim has to be scoped to that property, because a second marginal diagnostic does not follow the same order. Effective rank runs 65 ± 33, 72 ± 24, 57 ± 27 and 93 ± 29 over the same runs, so $\mu^\star$ — whose marginal margin is exactly zero — holds the *highest* rank of the four: its latent has less variance but spreads it over more directions. The spreads are wide enough that no ordering is established there, which is the point. $\mathcal{N}(0, I_d)$ is not one number, and what $\lambda_{\min}(\Gamma_\mu)$ predicts is the scale of the latent, not marginal normality entire.

What none of this settles is whether a larger η would plan better, since the margin is not the objective. The two values we ran disagree by environment — $\eta = 0.05$ is ahead on Push-T and behind on Reacher — so we do not claim an optimum. What the figure does establish is that the value in use was chosen against one axis of a two-axis trade-off.

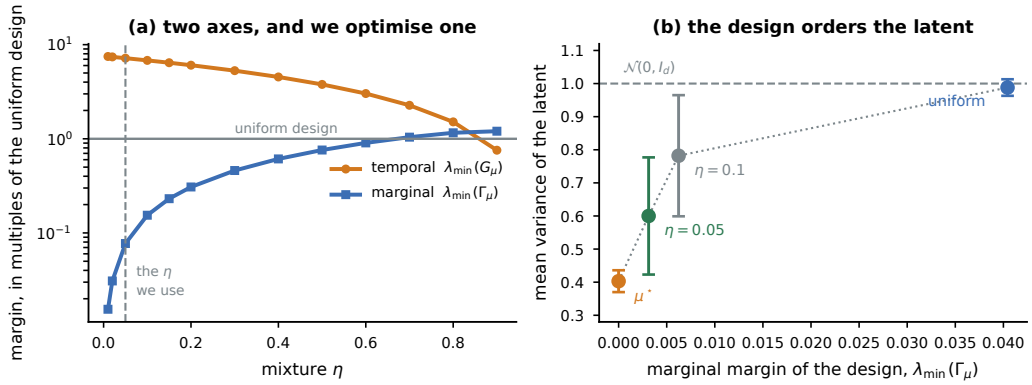


Figure 9: The design has two margins and §4 optimises one. **(a)** Temporal (G_μ) and marginal (Γ_μ) margins against η, in multiples of the uniform design. Our operating point, dashed, is $7.2\times$ above the uniform design on one axis and 13 times below it on the other. The marginal curve is not monotone: at $\eta = 1$ the design is the $h{+}1$ marginal atoms alone, which cannot span $\mathrm{Sym}(h{+}1)$, so it returns to zero. **(b)** The marginal margin of the design, computed before training, orders the mean variance of the trained latent over 22 runs. Bars are the spread across environments and replicas.

The two budgets: directions per step, and atoms. Two numbers in Appendix B need defending: $M = 256$ directions per step, and the 177 atoms the construction returns. Both are questions about $\lambda_{\min}$ and neither needs a training run, because freezing a set of directions *is* choosing a design whose support is that set.

Figure 10(a) reads that literally. Drawing M directions once and keeping them gives a design with empirical $\hat{G}_M$, and its blind spot is $\lambda_{\min}(\hat{G}_M)$. Below $M = 6 = h(h{+}1)/2$ that quantity is *exactly* zero for every design, and not by bad luck: fewer than $h(h{+}1)/2$ rank-one terms cannot span the purely temporal subspace, so a frozen draw that small leaves a direction the criterion can never charge for. Above the floor the curves climb towards their population values, and at $M = 256$ a frozen draw from the optimal design sits at 2.08, still above the *ceiling* of the uniform design. Since we redraw every step the population value is what governs training — a blind spot that moves each step is never available to descend into — and this is the mechanism behind the empirical finding of Balestriero & LeCun (2025) that resampling buys coverage for free. What the figure adds to that observation is the floor and the rate.

Panel (b) answers the second number, and the answer is that 177 is not barely enough but roughly eight times more than needed. Carathéodory bounds it: G_μ lives in $\mathrm{Sym}(h(h{+}1)/2)$, of dimension 21, so 22 atoms suffice to represent the same G_μ exactly. Measured, 16 well-chosen atoms already reach 99.0% of the full design. The support size is an artefact of Frank–Wolfe, which adds at most one atom per step and never prunes, not a requirement of the design. The negative result in

the same panel is worth as much: pruning by weight *fails*, with the 22 heaviest atoms reaching only 0.21, below the continuous uniform design. Weight does not identify the atoms that matter, because the heavy ones are largely redundant with one another; what the margin needs is diversity of direction, not mass. A design that kept the atoms rather than the weights could evaluate the temporal expectation exactly, at a fraction of $M = 256$ sampled directions, which we have not tried.

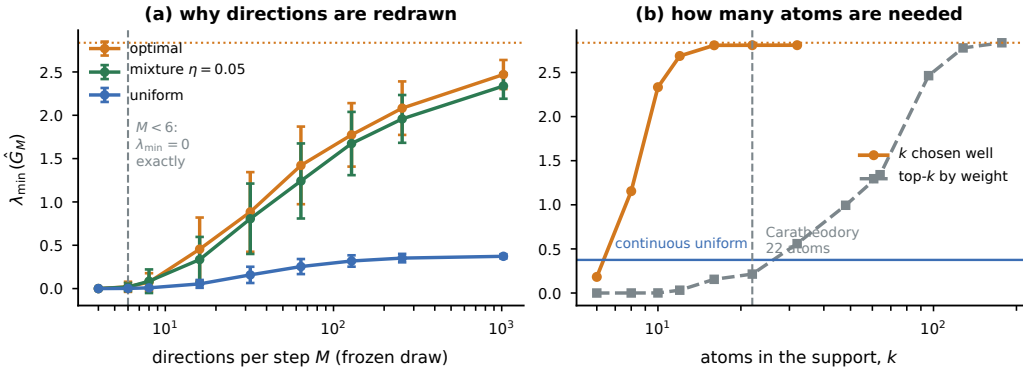


Figure 10: Both budgets, measured. **(a)** Freezing M directions is choosing a design supported on them; below $h(h+1)/2$ its margin is exactly zero for every design, since that many rank-one terms cannot span the temporal subspace. Mean and spread over 200 draws. Redrawing each step is what makes the population value, dotted, the one that governs training. **(b)** How many atoms the support really needs. Carathéodory allows 22; 16 chosen well give 99.0%. Pruning by weight instead (dashed) falls below the continuous uniform design, because heavy atoms are redundant with each other.

How far the sample is from the sphere. Because the construction optimises over a fixed draw, the honest question is what it costs, and it is cheap to answer: re-run it with more candidates and see where $\lambda_{\min}(G_{\mu^\star})$ stops moving. Figure 11(b) does that over three seeds. Going from the 8000 candidates we use to 128 000 raises $\lambda_{\min}$ from 2.837 to 2.844: sixteen times the candidates buy 0.3%. $\mathbb{S}^h$ at $h = 3$ is a three-dimensional manifold, so a few thousand points already cover it, and the reported design sits within a percent of the continuum.

A second draw makes the same point from the other side and is the better sampler if the horizon ever grows. Proposition 3 and §4.3 say the optimum puts its mass on zero-sum profiles, so drawing the candidates inside $\{q : \mathbf{1}^\top q = 0\}$ — a two-sphere inside the three-sphere — spends every candidate where the answer lives. It reaches 2.844 with 2000 candidates, the value the uniform draw needs 128 000 to reach, and is flat from there. Two things follow. The plateau of both curves is a usable estimate of the continuum optimum, which is what turns the paragraph above from a caveat into a bound. And the advantage is negligible at $h = 3$ but need not stay so: the purely temporal subspace has dimension $h(h+1)/2$, so the ambient sphere grows while the useful subsphere grows more slowly. At longer horizons this is where to spend effort, not on raising N.

How the loop reads, and what its output is not. Line 6 is what makes the algorithm legible: at every step the design is asked where its own worst blind spot lies, and the single candidate that best illuminates that spot receives weight. The step size $2/(k+2)$ is the standard Frank–Wolfe schedule (Frank & Wolfe, 1956), and because each iteration touches one candidate the support grows by at most one atom per step and never shrinks. The basis $\mathcal{B}$ spans the off-diagonal entries only, so the objective is blind to the diagonal by construction: scale is the marginal term's business, not this one's — which is §4.3's point arrived at from the implementation.

Two qualifications belong to the output rather than to the loop. The optimisation is over a *fixed sample* of 8000 directions drawn once on $\mathbb{S}^h$, not over the sphere: line 7 can only pick a candidate that line 1 already produced, so the design is E-optimal on that sample rather than on the continuum. The gap this leaves is small enough to measure rather than assert, and Figure 11 measures it. The

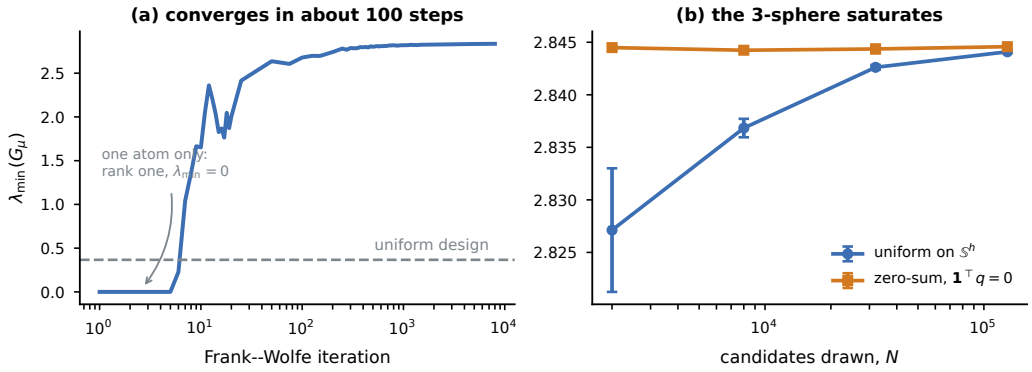


Figure 11: The design is E-optimal on a sample, and this is what that costs. **(a)** Frank–Wolfe converges in about a hundred steps; the first steps sit at zero because the schedule $2/(k{+}2)$ puts almost all weight on a single atom, and a rank-one $uu^\top$ has $\lambda_{\min} = 0$ by construction. The dashed line is the uniform design, the starting point. **(b)** The ceiling against the number of candidates, mean and spread over three seeds. The uniform draw saturates: sixteen times the candidates buy 0.3%. Drawing inside the zero-sum subsphere, where §4.3 shows the optimum lives, is at the plateau from 2000 candidates on. Both curves agree on where the plateau is, which is what bounds the distance to the continuum.

distinction still matters when the result is called an optimal design, and the $7.7\times$ and $11.2\times$ gains of §4 are ratios between two quantities computed on that same sample, which makes them internally consistent rather than statements about the continuous problem. And the returned support is finite without being minimal: of the 177 atoms above $\varepsilon = 10^{-5}$ at $\rho = 0.9$, only 28 carry more than 1% of the mass.

Why a finite list of questions suffices. The criterion sees μ only through G_μ of equation 2, so two designs with the same G_μ are indistinguishable to it, and G_μ is linear in μ. The reachable G_μ therefore form a convex hull, and by Carathéodory's theorem finitely many profiles reproduce *exactly* what any continuous design achieves, not approximately. The count is set by where G_μ lives, and that is one dimension higher than the perturbation space: G_μ is 6×6 and symmetric, so it ranges over a space of dimension 21 and the bound is $21 + 1 = 22$ atoms. It is not the 10 and 11 one gets from $\mathrm{Sym}(h{+}1)$, because the sensitivity is a *fourth* moment of q. Two operators appear in this appendix and they are not the same object: G_μ is what the design is *built* on, over the temporal coordinates alone, and it is the one Carathéodory's bound refers to; Γ_μ of Proposition 5 is the *diagnostic*, acting on all of $\mathrm{Sym}(h{+}1)$ and therefore ranging over a space of dimension 55.

Some of those questions, for concreteness:

profile q	the question it asks
$(1,1,1,1)/2$	what is the overall scale?
$(1,-1,0,0)/\sqrt{2}$	does z_t resemble z_{t+1}?
$(1,0,-1,0)/\sqrt{2}$	and does it resemble z_{t+2}?
e_t	how is instant t on its own?

The common-mode subspace, in detail. The blind spot is the *common-mode* subspace: the $h{+}1$ directions of $\mathrm{Sym}(h{+}1)$ built from the constant profile $\mathbf{1}$, namely $\{\mathbf{1}u^\top + u\mathbf{1}^\top : u \in \mathbb{R}^{h+1}\}$, whose worst element is $J = \mathbf{1}\mathbf{1}^\top$. The reason is that the optimum places all of its mass on profiles of nearly zero sum — that is what resolves temporal structure — with $|\sum_t q_t|$ averaging 0.013 under the design. Since $\langle \mathbf{1}u^\top + u\mathbf{1}^\top, qq^\top \rangle = 2(\mathbf{1}^\top q)(u^\top q)$ vanishes to first order in that sum and $\langle J, qq^\top \rangle = (\sum_t q_t)^2$ to second, the whole subspace is faint and J is the faintest. With $\gamma(q) = \mathrm{svec}(qq^\top/(q^\top R_{\rho,h}\, q))$ and the inner-product-preserving convention that scales off-diagonal entries by $\sqrt{2}$, the whole spectrum at $h{+}1 = 4$ and $\rho = 0.9$ is worth printing once, because the gap in it is

the argument:

$$\Gamma_{\mu^\star}:\ \underbrace{4.0{\times}10^{-7},\ 1.4{\times}10^{-3},\ 1.7{\times}10^{-3},\ 2.2{\times}10^{-3}}_{\text{common mode, fraction } 1.0000},\ \underbrace{2.87,\ 2.92,\ 5.62,\ 5.65,\ 5.69,\ 11.38}_{\text{orthogonal to it, fraction } 0.0000}.$$

The split is clean: numerically the four faint eigenvectors lie in the common-mode subspace to a fraction of 1.0000, and the six bright ones to 0.0000. Mixing at $\eta = 0.1$ raises the smallest to 6×10^{-3} and lowers the temporal $\lambda_{\min}$ from 2.874 to 2.586, exactly $1 - \eta$ of it; at $\eta = 0.05$ the smallest is 3.12×10^{-3} and the temporal value 2.730, exactly 0.95 of it. For reference the uniform design's smallest is 4×10^{-2}, so the mixture recovers the diagnostic to within a factor of seven of uniform while keeping nine tenths of a temporal sensitivity the uniform design never had.

It is not blind to scale, and a block-by-block reading misses the failure entirely: the scale direction, the traceless diagonal and the off-diagonal are each well conditioned, between 2.8 and 8.5. The faint directions cross those blocks. That is what refutes the natural guess — that the optimum has traded away the diagonal — and it is why the diagnostic has to be read on all of $\mathrm{Sym}(h{+}1)$ and not block by block.

What the blind direction means connects this section back to §3. A departure $\Delta = c\,J \otimes I_d$ is $z_t \mapsto z_t + \xi$ with one ξ shared by the whole window: it inflates the common mode and leaves every difference between instants intact. That is the direction of the constant window of §A.8, since $R_{\rho,h} \to J$ as $\rho \to 1$. The design is blind to *inflating* the shared component, not to collapse itself, which any design would see because the sketch variance would be zero.

Appendix A.4 measures the consequence, and what it measures is a *rate*. At the budget that brings every other design to $\|\Sigma_{tt} - I_d\| = 0$, the unmixed optimum sits at 1.86 with all four instants at variance 2.33 and a departure $c\,J$, $c = 1.33$; eight times the budget brings c to 0.14 without reaching zero. The guarantee is not lost, it is unreachable at any budget one would spend.

C.1 The measurements the body reports

Three tables carry the numbers §5.4 and §3.2 quote: Table 4 gives planning success offset by offset against the control, Table 5 the spectral diagnostic of every design before any planner is involved, and Table 7 the same runs scored by distance closed instead of by success. A dash reads two ways in what follows, and the captions say which: — where the quantity does not apply, – where it was not measured.

Every column or curve labelled T-SIGReg in this appendix, unless its caption says otherwise, is the configuration of Figure 5, held fixed across the four environments: the measured profile with the separable mixture at $\eta = 0.05$. Table 4 averages its three training runs; the other measurements use one.

C.2 What changed in the latent, in numbers

Table 5 reports, per design and suite, the realised mean variance per direction $\bar{v}$, the excess of the characteristic-function test over its null, and the realised lag-one autocorrelation $\bar{\kappa}$ on held-out windows. No design reaches the null, and the temporally optimal design falls furthest from unit variance, which is Proposition 3(i) in a trained model: what the design cannot see, the optimiser does not keep, and the mixture pulls $\bar{v}$ back towards one in the order its marginal margin predicts (Figure 9). The realised $\bar{\kappa}$ differs from its prescription in every cell; on Push-T the measured profile realises $\bar{\kappa} = 0.87$ against a prescribed $\kappa_1 = 0.953$ (Appendix B). The share of latent variance in the leading PCA-2 plane rises in all four suites (Figure 12).

The linear probe. A ridge regression from the latent to the true state, fitted on 60 held-out episodes per model with the partition *by episode* and the ridge weight chosen inside each fold. The partition matters: a random split leaves neighbouring frames of one episode on both sides and reads $R^2 \approx 1$ for every latent, which measures that consecutive frames resemble each other and nothing else. "Positions" are the state dimensions whose lag-one autocorrelation within an episode exceeds 0.9, the same cut as Figure 12; "all" is the full state.

The state is linearly recoverable only on TwoRooms. Elsewhere it is partially recoverable from either latent, and on Push-T T-SIGReg reads *worse* than the control, on positions and on the full

Table 4: Planning success at each goal offset: the SIGReg-only control against the single configuration of Figure 5, held fixed across the four environments, both averaged first within each seed. Bold marks the cells where the *paired* gain, differenced within seed and offset, clears its environment's per-offset retraining threshold by two standard errors; that threshold is 5.2 points on TwoRooms, 2.3 on Reacher, 2.8 on Cube and 5.5 on Push-T (Appendix B). Showing both columns rather than only their difference matters because the same gain means different things at different heights — five points on top of 32 is not five points on top of 94. The last row is the correlation between the offset and the paired gain. As in Table 2, the T-SIGReg column averages its three training runs; both columns are evaluated under the six seeds and differenced within seed.

	TwoRooms		Reacher		Cube		Push-T	
goal offset	SIGReg	T-SIGReg	SIGReg	T-SIGReg	SIGReg	T-SIGReg	SIGReg	T-SIGReg
15	97.3	100.0	56.7	56.3	90.0	**94.7**	94.0	83.6
25	86.8	**100.0**	80.3	82.5	69.0	**83.2**	93.7	82.6
35	74.0	**100.0**	89.3	89.4	54.2	**66.6**	75.0	69.3
45	63.2	**100.0**	93.2	**97.2**	47.7	**59.1**	51.0	52.4
55	51.2	**100.0**	92.8	**98.8**	52.7	**63.9**	32.0	39.6
65	42.3	**100.0**	91.3	**99.1**	60.2	**72.3**	23.0	28.4
75	31.2	**100.0**	87.3	**99.4**	67.5	**77.8**	15.5	22.1
85	–	–	86.0	**99.6**	–	–	–	–
95	–	–	83.2	**99.3**	–	–	–	–
105	–	–	80.8	**99.4**	–	–	–	–
115	–	–	78.3	**99.2**	–	–	–	–
125	–	–	77.5	**98.8**	–	–	–	–
135	–	–	74.3	**98.4**	–	–	–	–
$r(\text{offset}, \Delta)$	$+0.999$		$+0.990$		$+0.294$		$+0.931$	

Table 5: What the latent looks like at convergence. $\bar{v}$ is the mean variance per direction over 4096 directions, whose target is 1. EP is the excess of the characteristic-function test, observed minus null, which is the temporal criterion itself; we never quote the z-score, because at $M = 4096$ directions the excess is unchanged while the z triples. $\bar{\kappa}$ is the lag-one clock along the directions that linearly decode the state, with the world's own clock in the last row for comparison — the uniform mean over all 192 directions is far lower, around 0.13, because most directions carry nothing, and quoting that instead would understate the agreement by a factor of seven. Two readings are uncomfortable and reported anyway: $\bar{v}$ falls below 1 in exactly the designs that win, and the unmixed optimal design falls to 0.36–0.62 while its excess rises above $+4$, and above $+10$ in three suites, which is the trichotomy visible in the trained models rather than in the population argument. A — in the reference row marks a column that does not apply to it, since the world's own clock has no latent to take a variance or a test statistic of.

	TwoRooms			Reacher			Cube			Push-T		
design	$\bar{v}$	EP	$\bar{\kappa}$	$\bar{v}$	EP	$\bar{\kappa}$	$\bar{v}$	EP	$\bar{\kappa}$	$\bar{v}$	EP	$\bar{\kappa}$
SIGReg only (Maes et al., 2026)	1.00	$+0.84$	0.95	0.96	$+0.91$	0.84	0.99	$+0.59$	0.75	0.98	$+0.78$	0.92
full sphere, shared	1.01	$+1.70$	0.94	0.97	$+1.10$	0.85	1.00	$+0.97$	0.63	0.97	$+0.80$	0.89
full sphere, per channel (Wei, 2026)	1.02	$+1.70$	0.95	0.97	$+1.05$	0.85	1.00	$+0.96$	0.63	0.97	$+0.83$	0.89
separable, q uniform	0.87	$+4.15$	0.94	0.95	$+1.58$	0.86	0.99	$+0.83$	0.78	0.97	$+0.93$	0.89
separable, $q \sim \mu^\star$	0.42	$+13.36$	0.94	0.62	$+4.14$	0.73	0.36	$+12.70$	0.69	0.40	$+10.26$	0.81
separable, $q \sim \mu_\eta$	0.64	$+7.29$	0.95	0.86	$+2.26$	0.78	0.93	$+1.60$	0.71	0.85	$+1.85$	0.85
measured profile, $q \sim \mu_\eta$	0.68	$+7.57$	0.95	0.58	$+7.19$	0.83	0.90	$+1.84$	0.76	0.73	$+3.83$	0.87
the state's own clock (reference)	—	—	0.95	—	—	0.68	—	—	0.74	—	—	0.77

state alike. That is consistent with the rest of that suite: prescribing the path law costs something there, and the probe locates part of the cost in the linear readability of the state.

C.3 A SECOND METRIC, WITHOUT A SUCCESS THRESHOLD

Planning success is a threshold on a budget, and a threshold does not have to decrease with distance. Cube is the clearest case: its success rate falls to 48 at offset 45 and recovers to 67 by 75, and it does so identically at evaluation budgets of 75, 100 and 125, so the shape is a property of the task and not

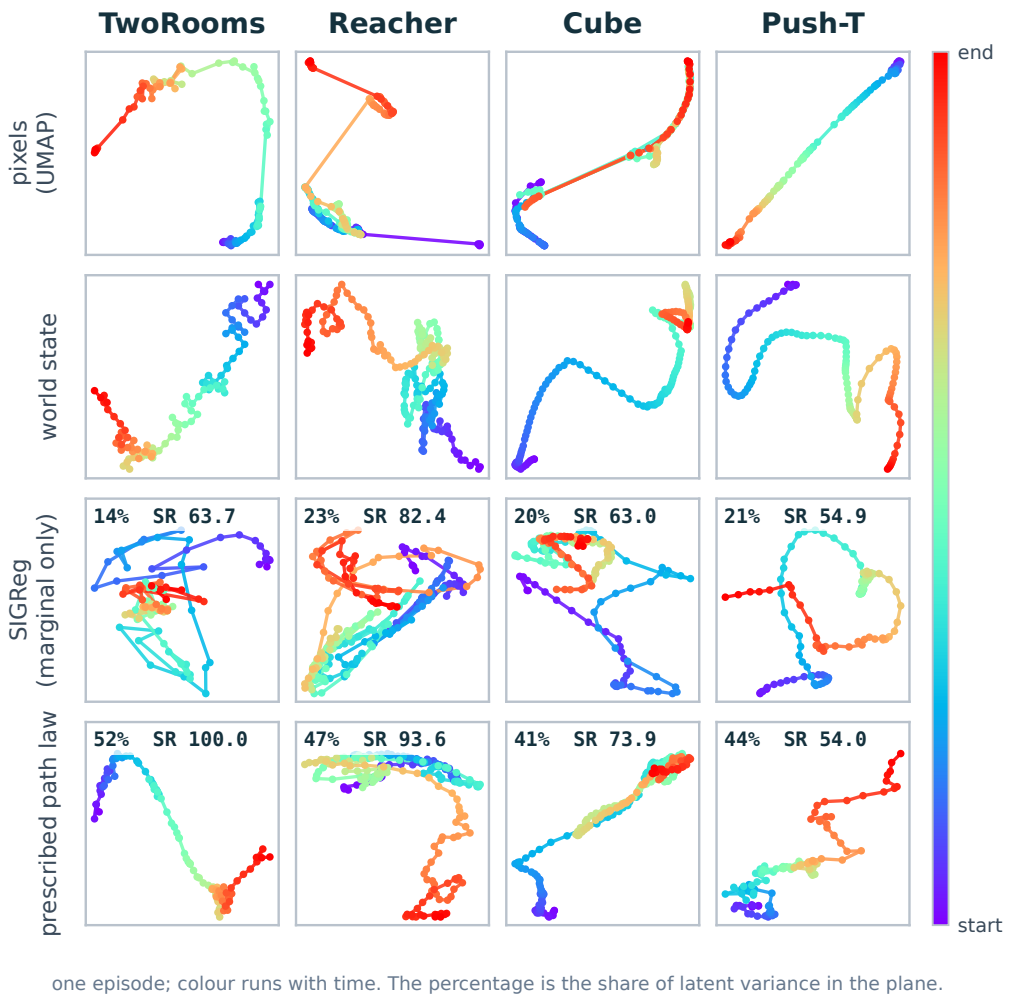


Figure 12: Figure 1 for all four environments, with the same colour scale, plus a row for the observations themselves. **Row 1**, *pixels*: a UMAP of the raw pixels of the same episode. It is a different instrument from the rest, because it projects a different object; it is here so the observation manifold can be compared against the latents. The episodes are the same in every row, and this is checked rather than assumed: the state read from the pixel archive has to match the probe's byte for byte, or the panel is not drawn. **Row 2**, *world state*: keeps only the state dimensions whose lag-one autocorrelation exceeds 0.9, which separates positions from velocities; the rule is applied uniformly and not chosen per environment. **Row 3** is the control and **row 4** is *one* configuration applied to all four environments — the measured profile with the mixture design at $\eta = 0.05$ — and not the best per environment, so no selection step enters the figure. It is also the row that scores highest on Push-T and within a point of the highest on the other three (Table 2). Holding it fixed is what makes the four panels comparable, and it is why Push-T's panel is a latent that does not beat its control. Each latent panel draws one episode of one training run, and the success it quotes is that of the corresponding cell of Table 2, averaged over training runs and evaluation seeds by the same code that builds the table. The figure shows the same prescription everywhere, including where it does not pay, and the variance gain orders across environments the same way the planning gain does (§5).

of the budget. Raising the budget changes little anywhere except Reacher — across matched policy and offset, going from 75 to 125 moves success by $+0.5$ on Cube and $+1.6$ on TwoRooms, against $+4.6$ on Reacher, which is the one environment where the planner is time-limited.

Table 6: Linear probe of the true state from the latent: ridge R^2 on held-out episodes, partitioned by episode. "Positions" are the state dimensions whose lag-one autocorrelation exceeds 0.9; "all" is the full state. This is the table §5 quotes for readability, not Table 5, which reports the spectral diagnostic instead.

		positions		all dimensions	
	position dims	SIGReg	T-SIGReg	SIGReg	T-SIGReg
TwoRooms	2	1.00	0.99	1.00	0.99
Reacher	4	0.52	0.56	0.34	0.36
Cube	21	0.30	0.33	0.20	0.29
Push-T	5	0.88	0.70	0.57	0.47

Table 7: Fraction of the initial goal distance closed, on the final models, ± 2 standard errors across seeds; the τ column gives the same offset in units of the environment's correlation time.

environment	goal offset	offset/τ	SIGReg	T-SIGReg
TwoRooms	15	0.25	0.407 $\pm$0.022	0.420 $\pm$0.022
	45	0.74	0.460 $\pm$0.029	0.751 $\pm$0.021
	75	1.23	0.319 $\pm$0.030	0.836 $\pm$0.010
Reacher	15	0.07	0.547 $\pm$0.025	0.548 $\pm$0.022
	45	0.21	0.708 $\pm$0.025	0.731 $\pm$0.016
	75	0.35	0.666 $\pm$0.033	0.788 $\pm$0.023
	105	0.49	0.620 $\pm$0.045	0.798 $\pm$0.020
	135	0.63	0.603 $\pm$0.037	0.813 $\pm$0.020
Cube	15	0.19	0.265 $\pm$0.049	0.334 $\pm$0.061
	45	0.58	0.234 $\pm$0.013	0.283 $\pm$0.016
	75	0.96	0.246 $\pm$0.015	0.297 $\pm$0.019
Push-T	15	0.14	0.804 $\pm$0.010	0.775 $\pm$0.002
	45	0.43	0.721 $\pm$0.042	0.757 $\pm$0.013
	75	0.72	0.540 $\pm$0.060	0.684 $\pm$0.019

We therefore add a graded measure that has no threshold in it: the largest fraction of the initial distance to the goal that the plan closes,

$$\text{closed} \;=\; 1 - \frac{\min_t d(z_t^{\text{state}}, \text{goal})}{d(\text{initial}, \text{goal})},$$

which is 1 when the goal is reached exactly and 0 when no progress is made. Here d is the Euclidean distance on the position dimensions of the true state, the same cut as Figure 12 (lag-one autocorrelation above 0.9), and the initial and goal states are the ones the evaluation harness itself constructs from the dataset to place the episode. The minimum over the episode, rather than the final distance, matches how the harness grants success: an episode succeeds if the environment's termination flag fires at *any* step within the budget, and the flag does not latch, so the final distance can be worse than the closest approach. The two readings differ by a tenth or more, and the final-distance version is given alongside in the generated records. This is a different quantity from success and not a finer-grained version of it: it says how close the plan gets regardless of whether it crosses the environment's threshold, and episodes that fail can close most of the distance. It is computed from the same CEM rollouts that produce the success rate, on the final models, wrapping the evaluator without changing it; the success rate recomputed from these rollouts reproduces the records behind Table 2 cell by cell (65 of 84 cells identical, the rest within 6 episodes in a hundred, the floating-point reduction of a different GPU changing CEM samples). Table 7 reports it.

Distances are also expressed in units of each world's own correlation time, $\tau = -\Delta t / \log \kappa_1$, rather than in steps. The four suites are evaluated over very different ranges of difficulty once measured that way: TwoRooms reaches 1.23τ at its furthest goal while Reacher reaches only 0.63τ at its own, despite the latter's offsets running to nearly twice as many steps. In steps the curves of §5.4 are not comparable across environments; in units of τ they are.

What the closed fraction says. Table 7 reports it where at least two seeds exist. The path law closes at least as much distance as the control in every measured cell but the nearest Push-T offset (Reacher: 0.547 against 0.548 at offset 15 and 0.708 against 0.731 at 45; Cube: 0.265 against 0.334 and 0.234 against 0.283). The fraction is not a finer success rate: split by outcome, episodes that fail often close nearly as much distance as episodes that succeed, and they start further from the goal in almost every cell. One regularity does hold in Reacher and Cube: where the control's successes close clearly more distance than its failures, T-SIGReg's do not, so T-SIGReg's failures are episodes that got close and did not cross the environment's threshold rather than episodes that fell short. The offset at which this happens differs by suite, so it is a property of where the control separates, not of the horizon.

The cost landscape, measured. The same rollouts record the distance in the latent to the encoded goal, $\|z_t - z_{\text{goal}}\|$, alongside the physical one. The raw latent distance is not comparable between models, since T-SIGReg's latent has the smaller scale, so Table 8 reports the Pearson correlation between the two series along each episode, median over episodes, mean ± 2 standard errors over seeds: how well the landscape the planner descends tracks the goal, the property temporal straightening (Wang et al., 2026) is built to buy and which here is a measurement, not a design target. This is the realised landscape along the trajectory, not the planner's own score, which rates the last predicted step. At the near offset both latents track the goal almost perfectly (+0.98 against +0.99 on Cube). At offset 45 the control's landscape stops meaning distance to the goal on Cube, +0.42, while T-SIGReg's keeps tracking it, +0.79; on Reacher both stay high (+0.97 against +0.98). Read together with Appendix C.4, this is the mechanism: the predictor is no better under the path law, but the latent it predicts in is one where distance keeps meaning distance to the goal further from it, and that is where the planning gain lives (§5.4).

Table 8: Correlation along the episode between the latent distance to the encoded goal and the physical distance on the position dimensions, median over 100 episodes, mean ± 2 standard errors over seeds; the same rollouts as Table 7. A landscape that tracks the goal scores near 1.

environment	goal offset	SIGReg	T-SIGReg
TwoRooms	15	$+0.89$ ± 0.03	$+0.99$ ± 0.00
	45	$+0.32$ ± 0.10	$+0.99$ ± 0.00
	75	-0.22 ± 0.13	$+0.99$ ± 0.00
Reacher	15	$+0.98$ ± 0.01	$+0.98$ ± 0.00
	45	$+0.97$ ± 0.00	$+0.98$ ± 0.00
	75	$+0.96$ ± 0.00	$+0.98$ ± 0.00
	105	$+0.95$ ± 0.01	$+0.98$ ± 0.00
	135	$+0.95$ ± 0.01	$+0.98$ ± 0.00
Cube	15	$+0.98$ ± 0.02	$+0.99$ ± 0.01
	45	$+0.42$ ± 0.21	$+0.79$ ± 0.11
	75	$+0.84$ ± 0.18	$+0.93$ ± 0.04
Push-T	15	$+0.98$ ± 0.00	$+0.97$ ± 0.01
	45	$+0.90$ ± 0.05	$+0.87$ ± 0.04
	75	$+0.38$ ± 0.17	$+0.67$ ± 0.09

What an offset asks for. The goal offset is not a difficulty scale: it sets the distance to the goal and, through the budget of twice the offset, the time allowed to reach it. Table 9 separates the two. On Reacher the arm needs about thirty steps to succeed at any offset, so the nearest goal is limited by its budget of 30; success rises until the budget stops binding and then falls as the goal moves away. On Cube a share of episodes starts with the goal already met, because in the expert trajectory the cube had not yet been moved or had already been set down. That share is not monotone in the offset, and neither is success, but it is the same for both models, so the paired gain is unaffected, and among the remaining episodes T-SIGReg leads at every offset.

C.4 Open-loop rollout: the gain does not come through the predictor

Zhao et al. (2026) report an open-loop comparison: condition on three context frames, predict forward from actions alone, and show the predicted *observations* beside the ground truth. A joint-

Table 9: What each goal offset asks for, from the rollouts of Table 7 (100 episodes under each of the seeds 7, 42 and 123): the budget; the median start-to-goal distance on the position dimensions, in each environment's own units; the median step at which the control's non-trivial successes occur; the share of episodes whose goal is already met at the first step; and the success rate among the rest. T-SIGReg is the configuration of Figure 5.

			start-to-goal	steps to	solved at start (%)		success, rest (%)	
environment	offset	budget	distance	success	SIGReg	T-SIGReg	SIGReg	T-SIGReg
TwoRooms	15	30	26.68	16	23	23	97	100
	45	90	71.29	22	3	3	58	100
	75	150	99.50	25	0	0	31	100
Reacher	15	30	0.51	24	1	1	53	58
	45	90	0.88	30	0	0	93	97
	75	150	1.14	32	0	0	87	99
	105	210	1.23	32	0	0	80	99
	135	270	1.48	34	1	1	75	98
Cube	15	30	2.12	18	52	53	76	87
	45	90	2.49	12	23	24	35	52
	75	150	2.18	18	35	34	51	66
Push-T	15	30	105.58	23	1	2	94	83
	45	90	168.64	24	0	0	52	50
	75	150	227.55	25	0	0	13	20

embedding architecture has no decoder, so we keep the protocol and drop the pixels: conditioning on the same three context steps and rolling the predictor forward on the recorded actions, $\hat{z}_{t+j+1} = \mathrm{pred}_\phi(\hat{z}_{t+j}, a_{t+j})$, we measure the error $\|\hat{z}_{t+k} - z_{t+k}\|_2$ against the horizon k on 512 held-out windows per model, for the control and for the measured profile with the mixture at $\eta = 0.05$. The data pipeline is the training one, not a re-implementation (frame skip, stacked and normalised actions), and every rollout was anchored: its one-step error was checked against the prediction loss the checkpoint logged at validation. One TwoRooms checkpoint failed that check by a factor of three, traced to a checkpoint mix-up, and was replaced by its training replica, which passes; both plan at 100.0 at every offset, so no reported number depends on the choice. Nothing here passes through the planner.

Figure 13 reports two normalisations of the same error. The lower row divides by $\|z_{t+k}\|$, the usual relative drift. The upper row, which is the one to read, divides by the *real motion* of the latent over the same horizon, $\|z_{t+k} - z_t\|$: a predictor that copies the current latent scores exactly 1 at every k, and the two models are compared at the same scale, which matters because T-SIGReg's latent has the smaller variance (0.68 against 0.94 per direction on Push-T) and moves less per step, so the raw drift flatters it.

The result is the same in all four suites, and it is negative for the obvious mechanism: T-SIGReg does not make the predictor better. At one step T-SIGReg's predictor misses between 1.7× and 2.2× the fraction of the real motion the control's misses (0.30 against 0.14 on Push-T, 0.11 against 0.07 on TwoRooms). The disadvantage narrows with the horizon, and on TwoRooms it reverses by $k = 10$ (0.15 against 0.23). Both predictors beat the copy baseline at one step, by a factor of three or more, so both have learned dynamics rather than persistence. What the same rollouts show is where the two latents differ: the control's latent has a realised lag-one autocorrelation of 0.33 on TwoRooms, 0.76 on Reacher and 0.71 on Cube, against 0.75, 0.91 and 0.90 under the path law. The control's predictor is accurate on a latent that correlates less across time, because the actions carry the transition; T-SIGReg's latent carries the transition in its geometry and its predictor is not better for it. The planning gain therefore does not come through the predictor. It comes through the geometry of the latent, which Figure 1 shows and the probe of §5 measures; whether the latent distance to the goal tracks the physical one along the episode, the direct test of that geometry as a cost landscape, is the measurement of §C.3.

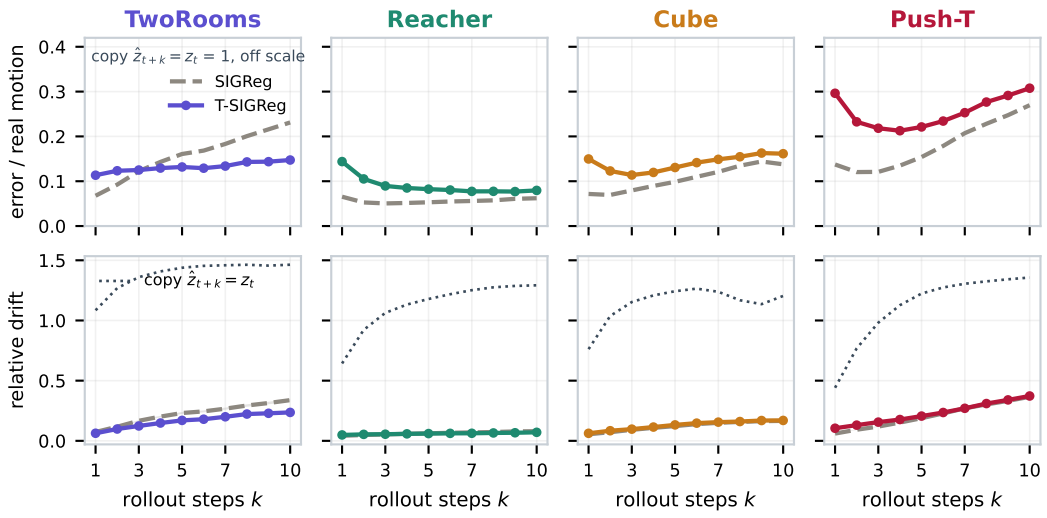


Figure 13: Open-loop rollout of the predictor, no planner in the loop, 512 held-out windows per model, bands ± 1 standard error. **Top**: error as a fraction of the real motion of the latent over the same horizon; the dotted line at 1 is a predictor that copies the current latent. **Bottom**: error relative to the norm of the target latent, with the copy predictor drawn where measured. Dashed, the SIGReg control; solid, the measured profile with the mixture at $\eta = 0.05$.

C.5 Two hypotheses for Push-T

The path law helps least where the control already has a clock. Table 5 measures the realised lag-one autocorrelation of the SIGReg-only latent in each suite: 0.95 on TwoRooms, 0.92 on Push-T, 0.84 on Reacher and 0.75 on Cube. TwoRooms saturates the success metric, so among the three suites that leave room to improve, *Push-T is the one whose control latent is already the most temporally correlated*. A marginal-only objective does not forbid temporal structure; it merely fails to require it, and on Push-T the data supply it natively. Prescribing a profile there adds little that was missing and still pays the price every constraint costs, which is the shape of the result in §5.4: no aggregate gain, a loss at the three nearest goals, and a gain only at the far ones.

A second hypothesis concerns the form of the target rather than its redundancy, and we state it as a hypothesis because we have not tested it. The AR(1) profile is the discrete trace of an Ornstein–Uhlenbeck process, whose paths are continuous: the prescribed covariance is that of a latent that drifts smoothly. Push-T is contact-rich, and contact makes the state jump — the block moves or it does not, and which of the two depends on a boundary the pusher may cross within a single step. A Gaussian path law with a fixed correlation time cannot express that discontinuity; it can only average over it, and averaging over a jump is exactly what would cost accuracy at short range, where the jump is a large fraction of the motion, while costing little at long range, where it is not. The measured profile does not escape this: it adjusts the correlation time, not the continuity. Distinguishing the two hypotheses needs a target that is not a single Gaussian — a mixture over contact modes, say — and that is outside what this paper prescribes.

C.6 The spatial identity and isotropy

The target $\Sigma_{\rho,h} = R_{\rho,h} \otimes I_d$ keeps the identity on the channel factor, so the prescription is isotropic in space: every direction of $\mathbb{R}^d$ is asked for unit variance. Zhao et al. (2026) argue that a global isotropic Gaussian prior distorts representations whose underlying manifold is low-dimensional, since it spends variance on directions the data never use, and propose restricting the prescription to a subspace.

T-SIGReg inherits that concern, and does not mitigate it: the spatial factor is exactly the one SIGReg already imposed, and our contribution is orthogonal to it, in the literal sense that it changes $R_{\rho,h}$ and

leaves I_d alone. Two things follow. First, any distortion the isotropic spatial prior causes is present in our control as well, so the paired comparisons of §5 are not confounded by it — both arms carry it. Second, the two remedies compose: a subspace-restricted spatial factor P would give $R_{\rho,h} \otimes P$, and every statement in §4 about the temporal factor survives unchanged, because none of them uses I_d beyond its being symmetric positive definite. We have not run that combination.

C.7 When the measured profile is not a legal target

The construction of §3.2 builds $R_{\rho,h}$ as a Toeplitz matrix from a measured autocorrelation sequence $\kappa_0 \dots \kappa_h$, and a covariance target must be positive semidefinite. An arbitrary sequence need not give one. The estimator matters: we use the biased sample autocovariance, which divides every lag by the same n rather than by $n - \ell$, and whose Toeplitz matrix is positive semidefinite by construction, being a Gram matrix of shifted windows. The unbiased estimator carries no such guarantee.

All four measured profiles are legal targets: the $\lambda_{\min}(R)$ column of §B gives 0.037, 0.009, 0.015 and 0.025, all strictly positive, so no fallback was exercised. Should a future environment yield an indefinite matrix — which the biased estimator makes impossible but a hand-specified or externally supplied profile does not — the projection onto the nearest positive semidefinite matrix in Frobenius norm, obtained by clipping the negative eigenvalues to zero and renormalising the diagonal to one, gives the closest legal target and preserves the Toeplitz structure only approximately; we would then re-fit a Toeplitz matrix to the clipped spectrum.

C.8 The $\rho = 0$ control and why it has one row

Prescribing $R = I_{h+1}$ asks for a latent whose instants are uncorrelated, which is the wrong law for a world that moves slowly. It is a negative control: its job is to show that whatever the rows below gain is the content of R and not the change of regulariser. One design and one training run suffice for that, so Table 2 carries the separable uniform row alone and we do not complete the grid.

We had measured two further $\rho = 0$ designs before taking that decision, on the two suites where they ran, and report them here rather than discard them. The optimal design reaches 27.1 on TwoRooms and 73.0 on Reacher; the mixture at $\eta = 0.1$ reaches 29.2 and 72.4. Both sit below the SIGReg-only control (63.7 and 82.4) by margins far exceeding the retraining thresholds of §B, and below the separable uniform row of the same block, so the conclusion does not depend on which $\rho = 0$ design is shown.