SMOOTH ACTIONS, STALE TARGETS: TEMPORAL WEIGHTING IN CHUNKED POLICIES

Anonymous authors
Paper under double-blind review

ABSTRACT

Learned robot policies can output commands that change abruptly from one step to the next, even when they predict several future actions together. Where smooth control is required, post-hoc filters aim to reduce these changes while retaining task success. However, earlier commands may no longer suit the current action: weighting the previous direction heavily can delay a turn. Each command has a target time: the time it is intended to execute. We ask whether task success depends on how filters weight commands for different target times, even when they attenuate the same input equally. We compare two filter pairs. Within each pair, the filters use the same four positive weights in opposite order, emphasizing either the newest or the oldest command. This reversal leaves their frequency-wise attenuation on a common input exactly unchanged. On one fixed OpenVLA with Optimized Fine-Tuning (OpenVLA-OFT) checkpoint, the recent-weighted filters achieve success rates 17.0 and 13.3 percentage points higher than their past-weighted counterparts. Each comparison uses 600 paired resets, matched by task and initial state. A second checkpoint on a different task suite shows the same pattern. The resulting closed-loop jerk and dynamics differ within each pair. Equal attenuation therefore need not imply equal task success. Standard filters illustrate the practical implications of this distinction. Across two Octo streams, a causal four-sample moving average over the latest selected commands (SW(4)), which returns a linearly changing command's value from 1.5 control steps earlier, lowers pooled success by 8.9 percentage points. In contrast, a four-sample causal Savitzky–Golay endpoint filter (SG4), which preserves the current value on linear command streams, reduces jerk while retaining the unsmoothed aggregate success. Tuned exponential moving averages offer useful trade-offs on OpenVLA-OFT; selected adaptive One Euro settings likewise exhibit a smoothing–success trade-off. When post-hoc smoothing is required, consider both attenuation and how much the filter emphasizes earlier commands, and judge the trade-off by closed-loop success.

1 INTRODUCTION

Learned robot policies can output commands that change abruptly from one step to the next, even when they predict several future actions together. For hardware that requires smooth control, these changes can matter beyond the appearance of a trajectory. Conditioning for Action Policy Smoothness (CAPS) reports a quadrotor case in which oscillations from a learned controller raise power, thermal, wear, and flight-worthiness concerns (Mysore et al., 2021). A practical response is to filter the policy's outputs: Sequential Dexterity applies an exponential moving average (EMA) to real-robot hand-policy commands (Chen et al., 2023). The goal is to reduce abrupt changes while retaining task success.

A filter must also contend with what the earlier commands asked the robot to do. When the robot needs to turn, for example, heavily weighting commands in the previous direction can delay that turn. Each command has a *target time*: the time it is intended to execute. Averaging commands selected for successive times therefore differs from combining several predictions of one action. Action Chunking with Transformers (ACT) uses the latter approach, combining overlapping predictions for the same future timestamp (Zhao et al., 2023). These mechanisms differ in which predictions they can access and how they combine them. Comparing them therefore cannot isolate the effect of

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.

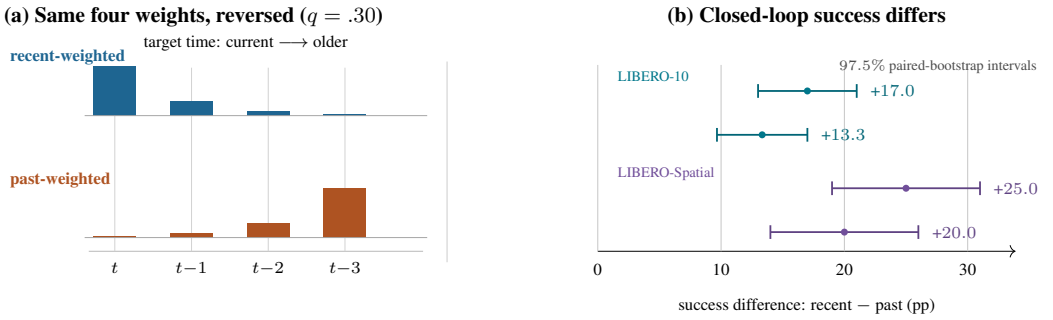

Figure 1: Equal attenuation, different task success. (a) Example at $q = .30$: recent-weighted filters apply normalized $(1, q, q^2, q^3)$ to the current command and three older commands; past-weighted filters reverse these weights. Each pair has identical frequency-wise attenuation on a common input. (b) Recent-minus-past success differences for two weight pairs in each study. The two studies use different suite-specific OpenVLA-OFT checkpoints, with 600 paired resets on LIBERO-10 and 200 on LIBERO-Spatial. Within each study, the upper and lower rows use $q = .30$ and $q = .45$, respectively. Teal marks LIBERO-10 and purple marks LIBERO-Spatial. Intervals are task-stratified paired-bootstrap 97.5% intervals, Bonferroni-adjusted for the two contrasts within each study. Realized closed-loop jerk differs within pairs. Section 5.1 and Appendix F give full results and evaluation details.

where predictions or commands are combined. Here we study filters that receive already selected commands. Can task success differ even when they attenuate the same input equally?

We address this question by reversing the same four positive weights within each of two filter pairs. One filter emphasizes recent commands and the other emphasizes older commands, while both have, by construction, exactly the same frequency-wise attenuation on a common input. On one fixed OpenVLA-OFT checkpoint, the recent-weighted filters achieve 17.0 and 13.3 percentage points higher success, with 600 paired resets per comparison (Figure 1; Section 5.1). A second OpenVLA-OFT checkpoint on a different task suite shows the same recent–past direction. The phase and transient responses differ, as standard filter theory predicts; closed-loop jerk and dynamics also differ. Thus attenuation alone does not capture the consequences of smoothing: which command times receive the most weight also matters.

Standard output filters make this distinction concrete. For a linearly changing command, the causal four-sample moving average SW(4) returns the value from 1.5 steps earlier. SG4 fits a line to the latest four selected commands and evaluates it at the current time (Savitzky & Golay, 1964) (Appendix E.8). In two Octo streams in SIMPLER, SW(4) reduces sent-command jerk but lowers pooled success; SG4 reduces sent-command jerk while retaining the unsmoothed aggregate success (Ghosh et al., 2024; Li et al., 2025). One fixed, high-success OpenVLA-OFT checkpoint shows the same aggregate direction (Kim et al., 2025b;a; Liu et al., 2023). Tuned EMA also offers useful trade-offs. Selected adaptive One Euro settings likewise exhibit a smoothing–success trade-off (Section 5.5). Preserving the current value on linear command streams is one way to limit temporal distortion, not a requirement for every useful filter.

Window-length and added-delay tests probe where these trade-offs deteriorate (Sections 5.4 and 5.5). Under stated error models, four conditional propositions relate prediction errors to command signatures (Appendix A). Closed-loop task success remains the adoption criterion.

Our contributions are:

- We show that identical frequency-wise attenuation on the same input does not imply identical task success: reversing temporal weights changes success at two OpenVLA-OFT checkpoints evaluated on different task suites;
- We show in paired Octo tests that reduced sent-command jerk does not necessarily indicate retained task success. We connect this gap to target-time semantics using standard endpoint filters, EMA, and added-delay controls; and

- We find that, in the tested Octo streams, longer endpoint windows increase curvature sensitivity and fitted command-space displacement, and lower success despite preserving the current value on linear command streams.

2 Related work

Chunked policies and temporal aggregation. Predicting an action chunk provides several future commands at each policy query. Temporal methods differ in whether they combine these predictions, change their generation, or modify commands after selection. ACT ensembles overlapping predictions for the same future timestamp to avoid the bias it attributes to smoothing across adjacent timesteps (Zhao et al., 2023). Diffusion Policy samples action sequences and executes a receding prefix (Chi et al., 2023). Bidirectional Decoding (BID) scores candidate chunks by backward coherence and forward plan quality (Liu et al., 2025). Action Coherence Guidance (ACG) guides flow-based sampling away from an intentionally incoherent action vector field (Park et al., 2025). Temporal Action Selection (TAS) selects among cached predictions (Weng et al., 2025), and Legato learns native chunk continuation (Liu et al., 2026).

Real-time and asynchronous execution. Execution latency adds a different requirement: the next chunk must continue the actions already committed to the robot. Real-Time Chunking (RTC) freezes committed actions and inpaints a continuation during chunk execution (Black et al., 2025b). Real-Time Robot Execution with Masked Action Chunking (REMAC) addresses intra-chunk inconsistency with learned masked corrections and prefix-preserved sampling (Wang et al., 2026). Reachability-aware Accelerated Chunk Execution (RACE), a method for time-optimal execution of action-chunk policies, predicts desired states and retimes execution under dynamic constraints (Kim et al., 2026). Concurrent work further identifies alignment among observations, predictions, and executed commands as fundamental to asynchronous deployment and compares post-hoc blending with generation-integrated alternatives (Motubrain Team, 2026). Smooth Execution of Action-Chunked Motion (SEAM) steers a flow policy toward the aligned, unexecuted tail of its previous chunk during generation (Zhan et al., 2026). Zero-Phase Action Policy Smoothing with Decoupled Actor (ZAPS-DA) instead avoids causal post-filter phase lag by distilling a centered, non-causal Savitzky–Golay target into a separately trained actor (Shamass, 2026).

We intervene only on selected actions, holding this access fixed. Comparing against systems such as RTC or BID would require a broader generation interface (Appendix C.4, Table 11).

Smoothness and control utility. Output filtering has both practical uses and known risks. Sequential Dexterity deploys EMA-filtered hand-policy outputs (Chen et al., 2023); CAPS evaluates EMA and finite-impulse-response (FIR) post-filters and reports anomalous learned-controller responses to naive filtering (Mysore et al., 2021). These precedents motivate EMA as a practice-linked baseline and SW(4) as a controlled, interpretable contrast. Smoothness is not a universal control objective: bang–bang control switches the input between its maximum and minimum allowed values and can be optimal under stated dynamics and objectives (Osmolovskii & Maurer, 2012). In systems where output smoothing is required, our comparison asks what attenuation leaves unspecified: the weight placed on earlier commands. Reversing those weights changes closed-loop success on two evaluated OpenVLA-OFT checkpoints in different task suites, even with identical common-input attenuation.

3 From kinematic diagnosis to target-time semantics

3.1 What kinematics can diagnose

We distinguish command variation from task success. For a command sequence y_t, define velocity $v_t = y_t - y_{t-1}$, discrete jerk $j_t = y_t - 3y_{t-1} + 3y_{t-2} - y_{t-3}$, and lag-one velocity autocorrelation $\rho_v = \mathrm{Corr}(v_t, v_{t+1})$. Four conditional propositions describe how error dependence, causal averaging, and chunk boundaries affect these diagnostics (Appendix A); recorded open-loop probes illustrate their measurement in heterogeneous settings (Appendix C.1).

A useful switch may have high jerk, and a stale command may be smooth. We use ρ_v and jerk to verify the change, and closed-loop success to judge its usefulness. For a common wide-sense stationary

input, equal-magnitude LTI filters have identical population velocity autocorrelation and jerk variance (Appendix A). This motivates our reversed-weight comparison: whether temporal reweighting that these common-input summaries cannot distinguish changes closed-loop task success.

3.2 Target-time preservation with a causal endpoint construction

We vary how output filters weight selected commands and measure the resulting changes in task success. The affine identities below characterize one temporal property of these filters.

Combining several predictions of one action and averaging successive actions need not preserve the same target. At observation time t, a chunk policy predicts

$$\hat{A}_t = (\hat{a}_{t,0}, \ldots, \hat{a}_{t,H-1}), \qquad \hat{a}_{t,k} \text{ targets time } t+k. \tag{1}$$

A same-target (aligned) ensemble for execution time τ combines only predictions whose offset points to τ,

$$\tilde{a}_\tau = \frac{\sum_{s \in \mathcal{S}_\tau} w_{\tau-s} \hat{a}_{s,\tau-s}}{\sum_{s \in \mathcal{S}_\tau} w_{\tau-s}}, \tag{2}$$

where $\mathcal{S}_\tau = \{s : 0 \le \tau - s < H\}$ indexes available predictions for time τ, and $w_k \ge 0$ is an offset-dependent weight with a nonzero denominator. By contrast, we denote the causal N-sample trailing mean over already selected commands by SW(N). For any selected stream u_t, including the per-step stream $u_t = \hat{a}_{t,0}$, SW(4) gives

$$\tilde{a}_t^{\mathrm{SW}(4)} = \frac{1}{4} \sum_{i=0}^{3} u_{t-i}. \tag{3}$$

Eq. 3 averages commands selected for four different target times. For $u_t = u_0 + ct$, $\tilde{a}_t^{\mathrm{SW}(4)} = u_t - 1.5c$: smoothing displaces the affine target into the past.

To test current-target preservation, we use SG(N), a standard causal endpoint fit over the latest N selected commands. Here N is the window length; the polynomial degree is always one. Its coefficients

$$h_i^{(N)} = \frac{2(2N-1-3i)}{N(N+1)}, \qquad \tilde{a}_t^{\mathrm{SG}(N)} = \sum_{i=0}^{N-1} h_i^{(N)} u_{t-i} \tag{4}$$

satisfy $\sum_i h_i^{(N)} = 1$ and $\sum_i i h_i^{(N)} = 0$. SG4 is

$$\tilde{a}_t^{\mathrm{SG4}} = 0.7u_t + 0.4u_{t-1} + 0.1u_{t-2} - 0.2u_{t-3}, \tag{5}$$

and returns u_t exactly on an affine segment. For a normalized FIR, the first moment is its DC group delay; SG4 therefore has zero DC group delay under this convention. We pass raw pose commands through until four samples exist and never filter the current gripper command. Using only current and past selected commands, this causal filter is a fixed depthwise convolution over pose channels; it needs no retraining, simulator state, or future prediction. Its negative coefficient permits amplification (peak gain 1.154); Appendix E.7 checks the sent spectrum. Moment constraints alone guarantee neither uniform attenuation nor task success. Filter equations act componentwise on pose vectors; $i = 0$ denotes the current command.

Zero first moment alone does not exactly preserve arbitrary command trajectories. The second moment is

$$\mu_2(N) = \sum_i i^2 h_i^{(N)} = -\frac{(N-1)(N-2)}{6}, \tag{6}$$

so a local expansion retains a curvature term $\frac{1}{2}\Delta^2 \mu_2(N) u''(t)$ for a smooth underlying command $u(t)$ sampled at interval Δ. Longer windows increase this curvature-error coefficient despite zero first moment; the error need not be a pure time shift. This motivates a window-length sensitivity test, not a fitted-shift or task-success law.

Appendix E.8 provides a recorded-command illustration of the standard-filter distinction. Target-time alignment does not equate native chunk execution, per-step replanning, same-target ensembling, and output filtering: feedback cadence and available predictions also differ. Their closed-loop success must be compared separately.

Table 1: Closed-loop comparison designs. Policy replication distinguishes trained checkpoints from evaluation streams; pairing is defined before the intervention, and trajectories may diverge afterward. Rows are grouped by role.

Study	Conditions	Policy replication	Budget / pairing unit	Test family
Primary evidence				
Reversed weights	native; 2 recent/past pairs	2 suite-specific OFT checkpoints	LIBERO-10: 600/arm; Spatial: 200/arm; 5 arms; task, reset	McNemar, Holm-2; bootstrap
Practical trade-offs				
Octo factorial	ensemble/per-step $\times$ no smoothing/SW(4)	1 checkpoint; 2 rollout blocks	300/cell/block; task, episode identifier	exact McNemar; 4 simple effects/block, Holm
Octo operators	six-arm ablation; SG4/SG8/SG12 sweep	1 checkpoint; 4 stochastic streams	96/populated arm/stream; task, geometry phase	exact McNemar; Holm for 2 window tests
OpenVLA-OFT original	native, SW(4), SG4	1 checkpoint; 1 deterministic runtime	450/arm; task, official state identifier	exact McNemar; task-stratified bootstrap
OpenVLA-OFT expanded	native, SW(4), SG4; 3 EMA settings; SG4 + pose delay (1 or 2 steps)	same checkpoint; reacquired	450/arm; 8 arms; task, state	3 McNemar tests, Holm
OpenVLA-OFT adaptive	native, SW(4), SG4; EMA/One Euro at 2 strengths	same LIBERO-10 checkpoint	400/arm; 7 arms; task, state	McNemar, Holm-2; task-stratified bootstrap
Boundary checks				
OpenVLA-OFT channel delay	SG4: pose/gripper delay $0/2 \times 0/2$	same checkpoint; reacquired	200/arm; 4 arms; task, state	3 approximate bootstrap intervals, Bonferroni
Franka Diffusion Policy	native, per-step, timestamp-aligned ensemble, per-step + post-hoc SW(4)	3 trained checkpoints	72/arm (3×24); checkpoint, reset, sample seed	exact McNemar; reset-cluster interval

4 EVALUATION DESIGN

We test whether smoothing retains task success, whether temporal weighting matters at matched attenuation, and how filter strength or added delay changes the trade-off (Table 1). The reversed-weight studies provide the primary evidence; the remaining studies map practical trade-offs and boundary cases.

A block is a 300-episode-per-cell factorial replication. An operator stream is a 96-episode-per-arm stochastic replication. These evaluation repetitions do not represent independently trained policy checkpoints. The two Octo factorial blocks are the initial and repeat evaluations in Section 5.2.

Endpoints and interventions. Binary simulator task success is primary; sent-command jerk and ρ_v assess kinematics. For Octo SG4 studies, we fit offset-corrected cumulative-XYZ command-path shifts in control steps, not physical end-effector delay. Later operators filter pose but pass the current raw gripper; the original Octo factorial also smooths the gripper score (Appendix E). Added-delay controls alter pose, gripper, or both; relative pose–gripper timing can change. Arms share checkpoint and reset/state identity, then may diverge after intervention.

Comparisons and uncertainty. Original contrasts use exact two-sided McNemar tests. Holm correction is separate for four simple effects in each Octo factorial block and the two SG-window tests. For post-hoc Octo intervals, tasks remain fixed and we compare resampling paired episodes with resampling geometry-phase clusters (Appendix E.5). For OpenVLA-OFT, we resample paired outcomes within each task; for Franka, we jointly resample resets across checkpoints and arms. Success counts and discordant pairs accompany exact tests. Discordance counts give successes unique to the first and second conditions, respectively.

Testbeds and scope. Octo-small is an open vision-language-action (VLA) testbed with four SIMPLER tasks, official success predicates, and 24 reused geometry phases with different policy/environment randomness for SG4. The original OpenVLA-OFT comparison uses one checkpoint and 45 previously evaluated official states per LIBERO-10 task; all arms share one deterministic runtime with exact pre-intervention identity checks (Appendix B). The eight-condition study reuses that bank; channel delay uses states 5–24. The adaptive study uses 40 official states per task and calibrates its filters on archived native traces from separate states, without using success labels

(Appendix G.1). The original reversed-weight study uses 60 newly generated resets per LIBERO-10 task; the separate Spatial study uses 20 newly generated resets per task on a second checkpoint. These studies are not pooled. The Franka study uses three independently trained Diffusion Policy checkpoints (Chi et al., 2023) on simulated Stack-Cube (Appendix B).

5 CLOSED-LOOP RESULTS

5.1 EQUAL ATTENUATION, DIFFERENT TEMPORAL WEIGHTS

Reversing four positive weights preserves the two filters' frequency-wise attenuation on a common input while changing which command times receive the most weight. We normalize $(1, q, q^2, q^3)$ to sum to one, assign these weights from the current command to the three preceding commands, and compare with their reversal (Figure 1). In LIBERO-10, recent/past success for geometric ratios $0.30/0.45$ is $541/600$ versus $439/600$, and $537/600$ versus $457/600$: $+17.00$ and $+13.33$ points (discordance $124/22$, $99/19$; Holm $p = 3.77 \times 10^{-18}$, 3.06×10^{-14}). Native records $553/600$; full results and evaluation details are in Appendix F.

We test the same five arms on a second OpenVLA-OFT checkpoint for LIBERO-Spatial using 200 newly generated paired resets. The recent-weighted filters achieve 25.0 and 20.0 percentage points higher success than their past-weighted counterparts for $q = 0.30$ and 0.45, respectively (both Holm-adjusted $p < 10^{-10}$; Appendix F.7). Closed-loop jerk differs; phase, transients, relative gripper timing and feedback change together. Attenuation alone therefore does not explain the success differences on either checkpoint.

5.2 UNALIGNED SMOOTHING IMPROVES KINEMATICS BUT LOWERS SUCCESS

SW(4) lowers sent-command jerk and observed success in both Octo generation conditions (Table 2). The factorial compares same-target ensembling and per-step prediction, each with and without SW(4). The ensemble uses SimplerEnv's Octo implementation of ACT's temporal-ensembling principle (Zhao et al., 2023; Li et al., 2025). SW(4) averages all seven coordinates, including the gripper score before thresholding. Later pose-only results reproduce the loss direction (Section 5.3), but cannot partition the original loss between pose and gripper smoothing.

Table 2: Octo/SIMPLER success in initial and repeat evaluations under the same protocol: four tasks, 75 episodes per task (300 per condition). SW(4) averages all seven selected-command coordinates, including the gripper score.

Temporal structure	Initial evaluation	Repeat evaluation
Same-target ensemble	$74/300$	$70/300$
Same-target ensemble + post-hoc SW(4)	$41/300$	$46/300$
Per-step prediction	$26/300$	$32/300$
Per-step + post-hoc SW(4)	$16/300$	$15/300$

The original confirmatory diagonal is $74/300$ versus $16/300$ in the initial evaluation ($+19.33$ points; discordance $67/9$, exact McNemar $p = 4.33 \times 10^{-12}$) and $70/300$ versus $15/300$ in the repeat evaluation. In a secondary Holm-adjusted factorial decomposition, the ensemble exceeds per-step prediction without SW(4) by $16.00/12.67$ points (initial/repeat); SW(4) reduces success by $11.00/8.00$ points on that stream and $3.33/5.67$ points under per-step prediction. All directions agree, but the initial per-step SW(4) effect is non-decisive after correction (adjusted $p = 0.132$; Appendix B). In this factorial, the observed SW(4) success loss is concentrated in Spoon and Eggplant; Carrot and StackCube already have near-zero success without SW(4).

SW(4) also raises ρ_v in all four comparisons. On the initial diagonal, the ensemble has jerk 8.348 and $\rho_v = -0.138$; per-step prediction with SW(4) has jerk 5.033 and $\rho_v = 0.093$, despite lower success.

5.3 On an ensemble-selected stream, SG4 avoids the observed loss

The six-arm Octo study changes only the output operator of the ensemble-selected stream; no smoothing leaves that stream unmodified. No smoothing and SW(4) record $24/96$ and $14/96$ successes, respectively. SG4 records $25/96$ successes: the SW(4) loss is not observed, but improvement over no smoothing is not established. EMA with $\alpha = 0.4$ records $18/96$ successes; its DC group delay is 1.5 command steps, the same as SW(4). Pure one-/two-step pose delays record $20/96$ and $12/96$, respectively, so success falls as the delay increases without smoothing.

We repeat no smoothing, SW(4), and SG4 in a second stochastic stream. Pooling both streams gives $48/192$, $31/192$, and $48/192$ successes, respectively (Figure 2). In this pooled comparison, SG4 lowers mean sent-command jerk from 8.327 to 4.841 (-41.9%), and its aggregate observed success matches no smoothing. SW(4) lowers jerk to 1.900. No smoothing versus SW(4), and SG4 versus SW(4), each have discordance $30/13$ and exact $p = 0.0137$; SG4 versus no smoothing has $16/16$ and $p = 1$. The $16/16$ discordance and task differences $+2, -5, +1, +2$ (Eggplant, Carrot, StackCube, Spoon; Appendix E.3) make this an aggregate preservation result, not equivalence or task-by-task dominance. Figure 2 shows paired uncertainty; Appendices E.5 and B give both resampling schemes, all contrasts, and stream details.

5.4 Window-length sensitivity despite affine exactness

In the Octo window-length study, longer SG windows reduce sent-command jerk but also lower observed task success, despite preserving linear commands. Two further Octo streams extend Eq. 4 to SG8 and SG12. Pooled success for no smoothing, SW(4), SG4, SG8, and SG12 is $51/30/59/42/33$ of 192, with fitted command-space shifts $0/1.500/0.112/0.578/1.147$ steps. SW(4) is worse than no smoothing ($p = 7.53 \times 10^{-4}$); SG8 has a nine-success deficit and non-decisive Holm-adjusted $p = 0.175$, whereas SG12 has an 18-success deficit and $p = 0.0102$. The direct SG12-minus-SG8 interval includes zero (Figure 2c): these results show window-length sensitivity, not a statistically located boundary or equivalence for SG8.

Across the pooled operator and window-length studies (192 episodes per arm each), SG4 records 17 and 29 more successes than SW(4), respectively. Its differences from no smoothing are 0 and 8 successes. This is a shared SW(4)-relative direction, not stable SG4 success or superiority to no smoothing.

Eq. 6 predicts increasing curvature sensitivity from second moments -1, -7, and $-55/3$. Fitted command-space shift grows while success falls (Figure 2c): a descriptive ordering, not a fitted task-success law.

5.5 Output-filter trade-offs on OpenVLA-OFT

Standard output filters. In the original LIBERO-10 three-operator comparison on the fixed OpenVLA-OFT checkpoint, SW(4) lowers aggregate success from the native rate of 92.4% ($416/450$) to 86.0% ($387/450$), a 6.4-percentage-point loss. Failures rise from $34/450$ to $63/450$, an 85.3% relative increase. SG4 records 92.9% ($418/450$). Native versus SW(4) has discordance $45/16$ ($p = 2.64 \times 10^{-4}$); SG4 versus SW(4) has $42/11$ ($p = 2.25 \times 10^{-5}$). Native versus SG4 has discordance $11/13$ and $p = 0.839$; this non-significant difference does not establish equivalence. The native-minus-SG4 bootstrap upper 95th percentile is 1.33 points (task-stratified, descriptive, not formal non-inferiority). Native/SW(4)/SG4 sent-command jerk is $578.024/163.610/365.770$.

Appendix B reports the earlier acquisition as a non-pooled sensitivity check.

Smoothing strength and added delay. We separately reacquire all eight conditions on the same checkpoint and previously evaluated reset bank, without pooling with the original study (Appendix D). Tuned EMA offers useful trade-offs: EMA(0.70) was selected by offline sent-command jerk matching to SG4, not by success outcomes; realized closed-loop jerk was not exactly matched. SG4 and EMA(0.70) record $418/450$ and $406/450$ successes, with 36.72% and 39.88% lower sent-command jerk than native. Their Holm-adjusted comparison is inconclusive ($p = 0.0576$). EMA(0.85), which applies the least smoothing of the three tested EMAs, records $422/450$ successes with 23.12% lower sent-command jerk than native. Adding one or two pose-delay steps after SG4

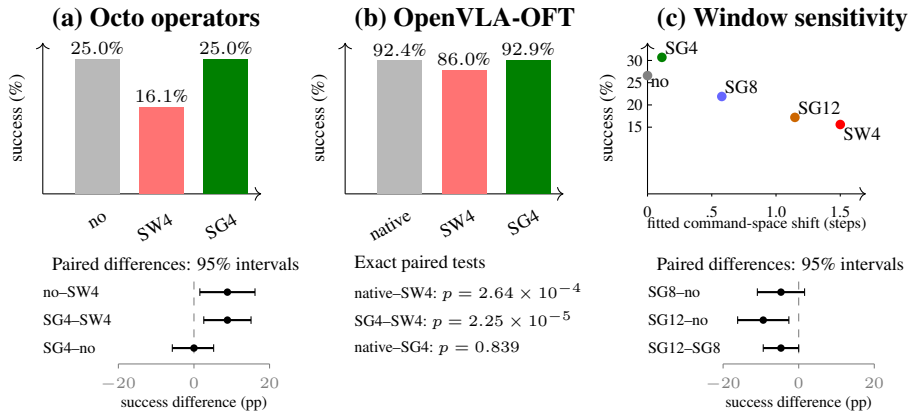

Figure 2: Success rates (top) and paired evidence (bottom). (a) Two Octo streams: SG4 retains unsmoothed aggregate success ($48/192$) with lower jerk. Here "no" means no output smoothing of the ensemble-selected stream. (b) Original three-operator comparison, one fixed OpenVLA-OFT checkpoint: all 450 task/state pairs were reacquired under one deterministic runtime with exact pre-intervention identity checks. SW(4) loses 6.4 points relative to native: failures rise from $34/450$ to $63/450$. (c) Two other Octo streams, separate from (a), in the window study (192 episodes/arm); SG4/SG8/SG12 use $4/8/12$ samples. Shifts align each rollout's own cumulative raw and filtered XYZ command paths (Appendix E.2), not physical latency; displacement, strength, and frequency response co-vary, so (c) is descriptive. Lower strips: (a,c) pointwise fixed-task geometry-cluster 95% intervals (Appendix E.5); (b) exact McNemar tests. "SW4" denotes SW(4); "native" is unmodified native execution.

gives $403/450$ and $363/450$ successes; only the two-step contrast is decisive after Holm correction. These arms retain SG4, but closed-loop histories, jerk, and pose–gripper timing change.

Adaptive filtering. An adaptive One Euro filter (Casiez et al., 2012) also exhibits a smoothing–success trade-off in a separate 400-pair LIBERO-10 comparison. Stronger smoothing reduces sent-command jerk while observed success falls from 91.75% to 89.25%; reconstructed weights emphasize earlier commands. Appendix G reports the temporal-weight analysis and the inconclusive success contrasts with similarly smoothed EMA.

5.6 Complementary comparisons and boundary cases

Across the eight-arm OpenVLA-OFT study and the reversed-weight studies on both checkpoints, design group delay helps explain observed success differences beyond a common-input attenuation descriptor. This exploratory reanalysis treats the three studies separately and is limited to the tested arms (Appendix F.8).

In the window study, SG12 and SW(4) have sent-command jerk of 2.276 and 1.878, fitted shifts of 1.147 and 1.500 steps, and $33/192$ and $30/192$ successes, respectively (paired $p = 0.728$; this does not establish equivalence). SG4 has sent-command jerk of 4.858, a fitted shift of 0.112 steps, and $59/192$ successes (Section 5.4). The OpenVLA-OFT study shows the same aggregate SW(4)/SG4 pattern under a different policy, benchmark, scale, and success regime, weakening an Octo-specific explanation. In Octo standard-filter comparisons, frequency response, transient behavior, smoothing strength, feedback cadence, and task sensitivity co-vary; these comparisons do not establish delay as the sole cause of success differences. Post-hoc alignment reduces SW(4)'s offset-corrected command-path error, but long SG windows retain larger residuals (Appendices E.4 and E.6).

In the separate 200-pair SG4 channel-delay factorial, the primary pose–gripper interaction remains unresolved (-4 points, corrected interval $[-11, 3]$); adding two steps to both channels reduces success from $183/200$ to $165/200$ (secondary contrast -9 points, $[-15.5, -2.5]$), while the third contrast is inconclusive. These controls neither decompose loss nor restore raw-command synchronization (Appendix D.2). A bounded $\pi_{0.5}$ comparison finds delayed-EMA loss but no clear SW(4) loss

relative to native (-2 points, corrected interval $[-8, 4]$); SG4's delay response and operator ordering remain unresolved (Appendix D.3).

A separate three-checkpoint Diffusion Policy comparison on simulated Franka Stack-Cube shows little observed change in task success despite changed kinematics; SG4 is not tested (Appendix B.4).

Separately, in hardware, one fixed 256-step learned-policy Cartesian trace replayed on an xArm Lite 6 gives controller-readback XY jerk Native/SG4/SW(4) of $2485.5/1314.5/626.7\,\mathrm{mm\,s^{-3}}$ and raw-reference shifts $0.000/0.230/1.485$ steps; this includes tracking dynamics but measures neither manipulation success nor calibrated latency (Appendix B.6).

6 LIMITATIONS

Evidence coverage. The reversed-weight comparison covers two suite-specific OpenVLA-OFT checkpoints; within-suite training-seed variation remains unmeasured. Other VLA comparisons use one checkpoint per model. Whether the reversed-weight effect persists in other policy families remains open.

Mechanistic interpretation. The conditional error models describe command signatures rather than identify each policy's residual process. Jerk and velocity autocorrelation are not episode-level success predictors. The reversed-weight comparison establishes that attenuation alone does not account for the success difference, but does not separate the contributions of temporal displacement, transients, and closed-loop feedback. How these effects combine to produce task failures remains unresolved.

Deployment. Our post-hoc filter comparisons use an interface that exposes only selected commands. RTC's asynchronous execution regime is not evaluated or ranked here. All closed-loop task-success experiments are simulated; xArm measures single-trace controller-readback kinematics only. These kinematic measurements do not establish hardware benefits in tracking, torque, safety, or wear. Whether the observed success–jerk trade-offs carry over to real-robot tasks and yield physical benefits remains open.

7 CONCLUSION

Smoothing changes both command variation and the weight placed on earlier selected commands. Across two OpenVLA-OFT checkpoints, reversing filter weights preserves common-input attenuation yet changes task success, so attenuation alone misses this difference.

In the Octo operator comparisons, SG4 reduces sent-command jerk without the aggregate success loss observed with SW(4). On OpenVLA-OFT, tuned EMA also offers useful trade-offs. Choose filters by temporal weighting as well as attenuation, and verify in closed loop that they meet the task-success requirements of the application. Revalidate this trade-off when increasing the window.

REPRODUCIBILITY STATEMENT

Section 3 defines the command diagnostics and filters; Appendix A gives the diagnostic assumptions and derivations. Section 4 and Appendices B, D, F, and G report the checkpoints, tasks, state or geometry identifiers, intervention definitions, pairing units, rollout budgets, and statistical procedures. Appendix E specifies operator placement, action conversion, the command-path fit, and task-level outcomes. The supplementary package will include the project-specific operator, evaluation, analysis, and validation code together with frozen configuration manifests; upstream models and simulators are identified by their public releases.

AI USE STATEMENT

Generative AI tools assisted with literature discovery and synthesis, conceptual and mathematical framing, including formulating and checking mathematical claims and derivations, hypothesis and

experiment-design critique, implementation and debugging, data-processing and analysis scripts, result interpretation, translation, manuscript drafting and editing, and adversarial claim checking. No generative-AI output was used as robot-training data, task observations, or evaluation outcomes. Experiments were executed in recorded runtimes, and paper-grade numerical claims were recomputed from preserved raw artifacts through separate verification procedures. The authors reviewed all AI-assisted work and take responsibility for the final text, claims, code, and artifacts.

REFERENCES

Johan Bjorck, Fernando Castaneda, Nikita Cherniadev, Xingye Da, Runyu Ding, Linxi Fan, Yu Fang, Dieter Fox, Fengyuan Hu, Spencer Huang, et al. GR00T N1: An open foundation model for generalist humanoid robots. *arXiv preprint arXiv:2503.14734*, 2025.

Kevin Black, Noah Brown, Danny Driess, Adnan Esmail, Michael Robert Equi, Chelsea Finn, Niccolo Fusai, Lachy Groom, Karol Hausman, Brian Ichter, Szymon Jakubczak, Tim Jones, Liyiming Ke, Sergey Levine, Adrian Li-Bell, Mohith Mothukuri, Suraj Nair, Karl Pertsch, Lucy Xiaoyang Shi, Laura Smith, James Tanner, Quan Vuong, Anna Walling, Haohuan Wang, and Ury Zhilinsky. π_0: A vision-language-action flow model for general robot control. In *Proceedings of Robotics: Science and Systems*, Los Angeles, CA, USA, June 2025a. doi: 10.15607/RSS.2025.XXI.010.

Kevin Black, Manuel Y. Galliker, and Sergey Levine. Real-time execution of action chunking flow policies. In *NeurIPS*, 2025b.

Géry Casiez, Nicolas Roussel, and Daniel Vogel. 1€ filter: A simple speed-based low-pass filter for noisy input in interactive systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2527–2530. Association for Computing Machinery, 2012. doi: 10.1145/2207676.2208639. URL `https://gery.casiez.net/1euro/`.

Yuanpei Chen, Chen Wang, Li Fei-Fei, and C. Karen Liu. Sequential dexterity: Chaining dexterous policies for long-horizon manipulation. In *Conference on Robot Learning*, 2023. URL `https://sequential-dexterity.github.io/assets/images/2023_CoRL_SeqDex.pdf`.

Cheng Chi, Siyuan Feng, Yilun Du, Zhenjia Xu, Eric Cousineau, Benjamin Burchfiel, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. In *RSS*, 2023.

Dibya Ghosh, Homer Rich Walke, Karl Pertsch, Kevin Black, Oier Mees, Sudeep Dasari, Joey Hejna, Tobias Kreiman, Charles Xu, You Liang Tan, Lawrence Yunliang Chen, Quan Vuong, Ted Xiao, Pannag R Sanketi, Dorsa Sadigh, Chelsea Finn, and Sergey Levine. Octo: An open-source generalist robot policy. In *Proceedings of Robotics: Science and Systems*, Delft, Netherlands, July 2024. doi: 10.15607/RSS.2024.XX.090.

Chuan Guo, Shihao Zou, Xinxin Zuo, Sen Wang, Wei Ji, Xingyu Li, and Li Cheng. Generating diverse and natural 3d human motions from text. In *CVPR*, 2022.

Alexander Khazatsky, Karl Pertsch, Suraj Nair, Ashwin Balakrishna, Sudeep Dasari, Siddharth Karamcheti, Soroush Nasiriany, Mohan Kumar Srirama, Lawrence Yunliang Chen, Kirsty Ellis, Masha Itkina, Marion Lepert, Yecheng Jason Ma, Patrick Tree Miller, Jimmy Wu, Suneel Belkhale, Shivin Dass, Huy Ha, Arhan Jain, Abraham Lee, Youngwoon Lee, Marius Memmel, Sungjae Park, Ilija Radosavovic, Kaiyuan Wang, Albert Zhan, Kevin Black, Cheng Chi, Kyle Beltran Hatch, Shan Lin, Jingpei Lu, Jean Mercat, Abdul Rehman, Pannag R Sanketi, Archit Sharma, Cody Simpson, Quan Vuong, Homer Rich Walke, Blake Wulfe, Ted Xiao, Jonathan Heewon Yang, Arefeh Yavary, Tony Z. Zhao, Christopher Agia, Rohan Baijal, Mateo Guaman Castro, Daphne Chen, Qiuyu Chen, Trinity Chung, Jaimyn Drake, Ethan Paul Foster, Jensen Gao, David Antonio Herrera, Minho Heo, Kyle Hsu, Jiaheng Hu, Donovon Jackson, Charlotte Le, Yunshuang Li, Roy Lin, Zehan Ma, Abhiram Maddukuri, Suvir Mirchandani, Daniel Morton, Tony Nguyen, Abigail O'Neill, Rosario Scalise, Derick Seale, Victor Son, Stephen Tian, Emi Tran, Andrew E. Wang, Yilin Wu, Annie Xie, Jingyun Yang, Patrick Yin, Yunchu Zhang, Osbert Bastani, Glen Berseth, Jeannette Bohg, Ken Goldberg, Abhinav Gupta, Dinesh Jayaraman, Joseph J. Lim, Jitendra Malik, Roberto Martín-Martín, Subramanian Ramamoorthy, Dorsa Sadigh, Shuran Song, Jiajun Wu, Michael C. Yip, Yuke Zhu, Thomas Kollar, Sergey Levine, and Chelsea Finn. DROID:

A large-scale in-the-wild robot manipulation dataset. In *Proceedings of Robotics: Science and Systems*, Delft, Netherlands, July 2024. doi: 10.15607/RSS.2024.XX.120.

Moo Jin Kim, Chelsea Finn, and Percy Liang. Fine-tuning vision-language-action models: Optimizing speed and success. In *Proceedings of Robotics: Science and Systems*, Los Angeles, CA, USA, June 2025a. doi: 10.15607/RSS.2025.XXI.017.

Moo Jin Kim, Karl Pertsch, Siddharth Karamcheti, Ted Xiao, Ashwin Balakrishna, Suraj Nair, Rafael Rafailov, Ethan P Foster, Pannag R Sanketi, Quan Vuong, Thomas Kollar, Benjamin Burchfiel, Russ Tedrake, Dorsa Sadigh, Sergey Levine, Percy Liang, and Chelsea Finn. Openvla: An open-source vision-language-action model. In *Proceedings of The 8th Conference on Robot Learning*, volume 270 of *Proceedings of Machine Learning Research*, pp. 2679–2713. PMLR, 2025b.

Sunwoo Kim, Jeongjun Kim, and Joseph J. Lim. Time optimal execution of action chunk policies beyond demonstration speed. In *ICLR*, 2026.

Xuanlin Li, Kyle Hsu, Jiayuan Gu, Oier Mees, Karl Pertsch, Homer Rich Walke, Chuyuan Fu, Ishikaa Lunawat, Isabel Sieh, Sean Kirmani, Sergey Levine, Jiajun Wu, Chelsea Finn, Hao Su, Quan Vuong, and Ted Xiao. Evaluating real-world robot manipulation policies in simulation. In *Proceedings of The 8th Conference on Robot Learning*, volume 270 of *Proceedings of Machine Learning Research*, pp. 3705–3728. PMLR, 2025.

Bo Liu, Yifeng Zhu, Chongkai Gao, Yihao Feng, Qiang Liu, Yuke Zhu, and Peter Stone. LIBERO: Benchmarking knowledge transfer for lifelong robot learning. In *NeurIPS Datasets and Benchmarks Track*, 2023.

Yuejiang Liu, Jubayer Ibn Hamid, Annie Xie, Yoonho Lee, Maximilian Du, and Chelsea Finn. Bidirectional decoding: Improving action chunking via guided test-time sampling. In *ICLR*, 2025.

Yufeng Liu, Hang Yu, Juntu Zhao, Bocheng Li, Di Zhang, Mingzhu Li, Wenxuan Wu, Yingdong Hu, Junyuan Xie, Junliang Guo, Dequan Wang, and Yang Gao. Learning native continuation for action chunking flow policies. In *Proceedings of Robotics: Science and Systems*, Sydney, Australia, July 2026. doi: 10.15607/RSS.2026.XXII.058.

Motubrain Team. World action models in real time: An empirical study of smooth execution via asynchronous deployment. *arXiv preprint arXiv:2608.01880*, 2026.

Siddharth Mysore, Bassel Mabsout, Renato Mancuso, and Kate Saenko. Regularizing action policies for smooth control with reinforcement learning. In *IEEE International Conference on Robotics and Automation*, 2021. URL `https://cs-people.bu.edu/rmancuso/files/papers/ICRA21_1616_FI.pdf`.

Nikolai P. Osmolovskii and Helmut Maurer. *Applications to Regular and Bang-Bang Control: Second-Order Necessary and Sufficient Optimality Conditions in Calculus of Variations and Optimal Control*, volume 24 of *Advances in Design and Control*. SIAM, 2012. ISBN 9781611972351. doi: 10.1137/1.9781611972368.

Minho Park, Kinam Kim, Junha Hyung, Hyojin Jang, Hoiyeong Jin, Jooyeol Yun, Hojoon Lee, and Jaegul Choo. Action coherence guidance for flow-based VLA models. *arXiv preprint arXiv:2510.22201*, 2025. Accepted at ICRA 2026.

Karl Pertsch, Kyle Stachowicz, Brian Ichter, Danny Driess, Suraj Nair, Quan Vuong, Oier Mees, Chelsea Finn, and Sergey Levine. Fast: Efficient action tokenization for vision-language-action models. In *Proceedings of Robotics: Science and Systems*, Los Angeles, CA, USA, June 2025. doi: 10.15607/RSS.2025.XXI.012.

Abraham Savitzky and Marcel J. E. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8):1627–1639, 1964. doi: 10.1021/ac60214a047.

Faiq Shamass. ZAPS-DA: Zero-phase action policy smoothing with decoupled actor for continuous control in reinforcement learning. *arXiv preprint arXiv:2605.30612*, 2026.

Guy Tevet, Sigal Raab, Brian Gordon, Yonatan Shafir, Daniel Cohen-Or, and Amit H Bermano. Human motion diffusion model. In *ICLR*, 2023.

Homer Rich Walke, Kevin Black, Tony Z. Zhao, Quan Vuong, Chongyi Zheng, Philippe Hansen-Estruch, Andre Wang He, Vivek Myers, Moo Jin Kim, Max Du, Abraham Lee, Kuan Fang, Chelsea Finn, and Sergey Levine. BridgeData V2: A dataset for robot learning at scale. In *CoRL*, 2023.

Haoxuan Wang, Gengyu Zhang, Yan Yan, Yuzhang Shang, Ramana Rao Kompella, and Gaowen Liu. Real-time robot execution with masked action chunking. In *ICLR*, 2026.

Yueyang Weng, Xiaopeng Zhang, Yongjin Mu, Yingcong Zhu, Yanjie Li, and Qi Liu. Temporal action selection for action chunking. *arXiv preprint arXiv:2511.04421*, 2025.

Dijia Zhan, Xuemiao Xu, Jinyi Li, and Jie Tang. SEAM: Smooth execution of action-chunked motion for vision-language-action policies. *arXiv preprint arXiv:2607.04609*, 2026.

Jianrong Zhang, Yangsong Zhang, Xiaodong Cun, Shaoli Huang, Yong Zhang, Hongwei Zhao, Hongtao Lu, and Xi Shen. T2m-gpt: Generating human motion from textual descriptions with discrete representations. In *CVPR*, 2023.

Tony Z Zhao, Vikash Kumar, Sergey Levine, and Chelsea Finn. Learning fine-grained bimanual manipulation with low-cost hardware. In *RSS*, 2023.

A Conditional diagnostic details

This appendix records the four conditional propositions underlying the compact diagnostic discussion. They characterize specified stochastic models; they do not assert that all learned-policy residuals obey those models.

Let a generated trajectory be $y_t = y_t^* + \varepsilon_t$, with reference y_t^* and additive error ε_t. Define velocity $v_t = y_t - y_{t-1}$, discrete jerk $j_t = y_t - 3y_{t-1} + 3y_{t-2} - y_{t-3}$, and lag-one velocity autocorrelation $\rho_v = \mathrm{Corr}(v_t, v_{t+1})$.

Proposition 1 (Additive error and velocity correlation)**.** *Under a stationary first-order autoregressive (AR(1)) error $\varepsilon_t = \rho\varepsilon_{t-1} + \sqrt{1-\rho^2}\sigma z_t$, reference– error independence, stationary reference velocity, and independent and identically distributed (i.i.d.) standardized innovations z_t with zero mean and unit variance,*

$$\rho_v(\sigma, \rho) = \frac{C^* - \sigma^2(1-\rho)^2}{V^* + 2\sigma^2(1-\rho)}, \qquad \lim_{\sigma^2 \to \infty} \rho_v = -\frac{1-\rho}{2}. \tag{7}$$

where C^* and V^* are the lag-one covariance and variance of the reference velocity.

The error velocity increment has variance $2\sigma^2(1-\rho)$ and lag-one covariance $-\sigma^2(1-\rho)^2$; adding independent reference terms gives the expression. Under independent high-variance error, ρ_v approaches $-1/2$; positive temporal correlation makes this limit less negative.

Proposition 2 (Causal-window jerk variance)**.** *For i.i.d. error of variance σ^2 and $\bar{y}_t = N^{-1}\sum_{i=0}^{N-1} y_{t-i}$, the error contribution to jerk has variance $20\sigma^2$ for $N = 1$, $5\sigma^2/2$ for $N = 2$, and $12\sigma^2/N^2$ for $N \geq 3$.*

This follows by contracting the averaged-error autocovariance $R(k) = \sigma^2 \max(N - |k|, 0)/N^2$ with jerk coefficients $[1, -3, 3, -1]$.

Proposition 3 (Boundary exposure)**.** *In a long sequence of non-overlapping chunks of length K, the fraction of consecutive velocity pairs touching a boundary is $\min(1, 2/K)$. If boundary and interior populations are commonly centered and have equal marginal variance, their correlations mix with these weights.*

The count is structural; the correlation mixture requires the additional normalization condition and does not imply that longer chunks improve utility.

Proposition 4 (AR(1) spectral jerk)**.** *For fixed-variance AR(1) error,*

$$\mathrm{Var}(j_t^{\varepsilon}) = \sigma^2(20 - 30\rho + 12\rho^2 - 2\rho^3),$$

which decreases strictly for $0 \le \rho < 1$*. More generally, within the class of fixed-variance stationary errors whose even power spectrum is non-increasing on* $[0, \pi]$*, a flat spectrum maximizes jerk variance because the jerk transfer gain* $(2\sin(\omega/2))^6$ *is increasing on that interval.*

Remark (magnitude-only second-order summaries). Let x_t be a wide-sense stationary input with centered power spectral density $S_x(\omega)$, and let $y_t = \sum_i h_i x_{t-i}$ for a fixed stable linear time-invariant filter with frequency response $H(\omega)$. Writing $R_y(k)$ for the output autocovariance, we have $S_y(\omega) = |H(\omega)|^2 S_x(\omega)$ and $R_y(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} |H(\omega)|^2 S_x(\omega) e^{i\omega k}\, d\omega$, so population second-order quantities of y_t and its finite differences, including ρ_v (when velocity variance is positive) and $\mathrm{Var}(j_t)$, depend on the filter only through $|H(\omega)|$. The reversed kernels satisfy $|H_q^{\mathrm{rev}}| = |H_q|$ (Eq. 10) and therefore share these quantities on a common input. A pure delay likewise has the same magnitude response as the identity and preserves these population second-order quantities on a common stationary input. This equality concerns population second-order quantities, not the finite-record mean-norm jerk reported in our experiments; closed-loop inputs also differ.

Scoped probes across robot-action and human-motion generators test selected predictions of this framework. Their heterogeneous model, dataset, and runtime contracts make them diagnostic examples rather than a controlled architecture ranking.

B Additional closed-loop results

B.1 Octo factorial details

The initial evaluation's observed simple effects in percentage points are: same-target ensemble minus same-target ensemble+SW(4), $+11.00$; per-step minus per-step+SW(4), $+3.33$; same-target ensemble minus per-step without SW(4), $+16.00$; and the same generation contrast with SW(4), $+8.33$. Holm-adjusted exact McNemar values are 2.61×10^{-5}, 0.132, 1.79×10^{-8}, and 3.40×10^{-4}. The repeat evaluation's effects have the same signs: $+8.00, +5.67, +12.67, +10.33$ points, all with adjusted $p \le 0.00333$. Spoon and Eggplant contribute most successes; Carrot and StackCube remain near floor, so the pooled ordering is not a uniform per-task result.

B.2 SG4 operator tables

Table 3: Two-stream Octo SG4 comparison. Kinematics are pooled sent-command means.

Operator	Stream 1	Stream 2	Pooled	Jerk / ρ_v
No smoothing	$24/96$	$24/96$	$48/192$	$8.327/-0.134$
SW(4)	$14/96$	$17/96$	$31/192$	$1.900/0.504$
SG4	$25/96$	$23/96$	$48/192$	$4.841/0.232$

The second stream alone is directionally consistent but not individually decisive: no smoothing versus SW(4) has exact two-sided McNemar $p = 0.2100$, and SG4 versus SW(4) has $p = 0.3075$. The main text therefore reports the two-stream pooled comparison rather than treating either stream as a separate replication claim. In the second stream, the raw result and configuration were archived before a postprocessing command-line entry-point failure required local regeneration of derived analysis and validation from those unchanged rows. No episode was rerun, and an independent raw-row recomputation matched the regenerated statistics.

Pure one- and two-step controls recover fitted shifts 1.000 and 2.000 with negligible corrected residual. Across two earlier streams, SW(4) fits $1.501/1.500$ and SG4 $0.113/0.114$ steps. These are same-rollout cumulative- XYZ command fits, not realized end-effector measurements.

B.3 OpenVLA-OFT single-runtime reacquisition and historical sensitivity

Native, SW(4), and SG4 obtain $416/450$, $387/450$, and $418/450$ successes; mean sent-command jerk is 578.024, 163.610, and 365.770. Every task/state pair uses one deterministic runtime con-

Table 4: Octo SG-window sensitivity over two further streams. Fitted shift is an offset-corrected temporal fit between cumulative-XYZ command paths, not physical delay.

Operator	Success	Jerk	ρ_v	Shift
No smoothing	$51/192$	8.328	-0.126	0.000
SW(4)	$30/192$	1.878	0.510	1.500
SG4	$59/192$	4.858	0.237	0.112
SG8	$42/192$	2.902	0.384	0.578
SG12	$33/192$	2.276	0.380	1.147

Table 5: Primary OpenVLA-OFT reacquisition under one deterministic runtime contract. Discordances are left-only/right-only; tests are exact and two-sided.

Contrast	Discordance	Exact p
Native–SW(4)	$45/16$	.000264
SG4–SW(4)	$42/11$	.0000225
Native–SG4	$11/13$	.839

tract, passes exact pre-intervention identity checks of the initial state, observations/model inputs, simulator state, and first action chunk, and preserves the raw gripper channel. The task-stratified bootstrap upper 95th percentile for the native-minus-SG4 success gap is 1.33 percentage points. It is descriptive and does not establish equivalence or formal non-inferiority.

An earlier mixed-acquisition measurement gave $418/450$, $384/450$, and $416/450$ successes. It combined five retained tasks with five tasks reacquired under stricter determinism, so task partition and acquisition block coincided. The single-runtime reacquisition was specified to replace that measurement, and the two results are never pooled. Their aggregate counts differ by at most three successes per arm; this is sensitivity to correcting the acquisition contract, not an independent replication.

B.4 Diffusion Policy on simulated Franka Stack-Cube

We evaluate three Diffusion Policy checkpoints on simulated Franka Stack-Cube. Across these checkpoints, native, per-step, timestamp-aligned ensemble, and per-step + post-hoc SW(4) obtain $65/72$, $66/72$, $66/72$, and $65/72$. SG4 is not tested. Despite changed kinematics, $23/24$ resets are invariant; two account for $24/26$ failures. This reset-dominated near-null is not equivalence or an intervention boundary.

Table 6: Franka Stack-Cube aggregate over three fixed checkpoints.

Arm	Success	Sent ρ_v	Sent jerk	Observed jerk
Native $H = 8$	$65/72$	0.308	0.003208	0.000559
Per-step $H = 1$	$66/72$	0.112	0.004374	0.000536
Timestamp-aligned ensemble	$66/72$	0.782	0.000921	0.000460
Per-step + post-hoc SW(4)	$65/72$	0.803	0.001047	0.000408

We retain the kinematic measurements as a descriptive observation but do not use this study to establish a causal margin explanation. A distinct precision-tray case shows SW(4) harm, but SG4 does not meet the native-relative adoption criterion there (Appendix B.5).

B.5 A bounded precision-tray Diffusion Policy case

We also evaluate one separately trained epoch-40 raw Diffusion Policy checkpoint on 100 fixed precision-tray validation positions. Each position uses reset-matched native, xyz-SW(4), and xyz-SG4 arms. The two filters change only xyz; rotation and the binary gripper command pass through unchanged.

Native versus xyz-SW(4) has paired discordance $86/0$ and exact two-sided McNemar $p = 2.58 \times 10^{-26}$; xyz-SG4 versus xyz-SW(4) has discordance $81/0$ and $p = 8.27 \times 10^{-25}$. Relative to native,

Table 7: Bounded precision-tray comparison on one fixed checkpoint. Jerk is the mean per-episode xyz sent-command jerk.

Arm	Success	XYZ jerk
Native	$86/100$	0.062419
XYZ-SW(4)	$0/100$	0.018890
XYZ-SG4	$81/100$	0.041773

xyz-SG4 lowers mean xyz jerk by 33.1%. However, the one-sided paired-bootstrap upper bound for the native-minus-SG4 success gap is 13.0 percentage points, which does not meet the fixed < 5-point criterion. This criterion supports the SW(4)-harm observation but not an SG4 transfer, equivalence, or non-inferiority claim. Because the task, dataset, and checkpoint differ from the Stack-Cube study, their contrast is descriptive and does not identify task margin as a causal moderator.

B.6 A SINGLE-TRACE XARM PHYSICAL KINEMATIC DIAGNOSTIC

We rescale one archived learned-policy XY command trace to a 15-mm-radius, fixed-height Cartesian replay on one xArm Lite 6. Native, SW(4), and SG4 each complete 256 commands at an 80-ms period with no command error or missed deadline; height and orientation remain fixed. Commands are sent through the Cartesian servo API `set_servo_cartesian`. Table 8 reports command and controller-readback kinematics. The fitted shift compares each controller readback with the common raw command reference. [1]

Table 8: One xArm Lite 6 physical Cartesian replay. Jerk is the mean XY third-difference norm in $\mathrm{mm\,s^{-3}}$; L2 columns are mean distances in mm. Realized values are controller-reported readbacks.

Arm	Sent jerk	Realized jerk	Realized–sent L2	Realized–raw L2	Shift (steps)
Native	2035.2	2485.5	0.0558	0.0558	0.000
SW(4)	527.6	626.7	0.0250	0.4148	1.485
SG4	1106.3	1314.5	0.0385	0.1395	0.230

The realized jerk order remains Native $>$ SG4 $>$ SW(4), and SG4 is closer than SW(4) to the current raw reference under the fitted-shift measure. The fit includes controller tracking dynamics and is not calibrated physical latency. This one-trace diagnostic measures neither manipulation success nor independent external motion, energy, wear, safety benefit, cross-robot generality, or RTC performance.

A fixed-checkpoint simulated xArm horizon probe gives native $H = 8$ versus per-step $H = 1$ success $28/30$ versus $25/30$ on calibration and $29/30$ versus $26/30$ on a held-out split. Combined paired discordance is $8/2$ with exact $p = 0.109$. This is directional support from one task, not established superiority.

C SCOPE AND REPORTING DETAILS

Octo SG4 evaluation streams reuse the same 24 simulator geometry phases; stochastic streams are not new-geometry generalization. OpenVLA-OFT uses one fixed checkpoint and state IDs 5–49 for all ten LIBERO-10 tasks. Franka aggregates three fixed training checkpoints and its reset-cluster interval is conditional on them rather than an estimate of training-seed population uncertainty.

For every closed-loop arm comparison, pairing refers to checkpoint and reset or official state identity before intervention. Once the first sent action differs, trajectories, observations, and later policy outputs may diverge. We do not describe them as identical rollouts or shared trajectories. The OpenVLA-OFT reacquisition under one deterministic runtime contract verifies exact pre-intervention identity for all 450 task/state pairs, as detailed in Appendix B.

[1] The fixed search grid spans -0.5 to 3.0 steps at 0.005-step resolution, with 12 samples trimmed at each edge.

Table 9: Open-loop diagnostic probes in their native quantities across heterogeneous settings. The Octo row uses a sequential full-chunk construction, not executed commands. The table reports no task-success outcomes and is not a controlled model comparison or validation of AR(1) assumptions.

Probe	Setting	Recorded signature
OpenVLA ρ_v	Bridge, 20 episodes	policy -0.377; teleop $+0.050$
OpenVLA ρ_v	LIBERO, 2 tasks $\times$ 20 episodes	policy -0.454; simulator reference $+0.614$
Octo velocity product	Bridge, $K = 4$, 20 episodes	within $+0.418$; boundary -3.027
Window-width jerk	Bridge, 20/model; trailing window $1 \to 6$	OpenVLA $13.77 \to 2.08$; Octo $113.39 \to 17.60$

The fitted command-space shift aligns cumulative XYZ command paths while fitting a constant offset that absorbs causal warm-up translation. Its objective and remaining shape error are specified in Appendix E.2. No task-specific physical tolerance was recorded, so a displacement-to-tolerance ratio is unavailable. Neither this fit nor ρ_v or jerk is used as a task-success surrogate or intervention router.

C.1 The signatures are measurable, but heterogeneous

Table 9 summarizes recorded open-loop probes on Bridge and LIBERO inputs (Kim et al., 2025b; Walke et al., 2023; Liu et al., 2023). The full-chunk Octo probe averages adjacent velocity products over seven action coordinates, then within and across episodes, without centering or normalization (Appendix C.2).

Appendix C.3 reports additional motion and model-output diagnostics with their signal-construction and temporal-adjacency limits.

C.2 Octo velocity-product construction

For Table 9, each recorded Octo output chunk contains four seven-coordinate actions. The probe concatenates complete chunks from offline queries into a sequential construction; it does not evaluate their closed-loop execution. It takes queries at indices $0, 4, 8, \ldots$ and truncates the final chunk to the original observation-sequence length. For its stored action sequence a_i, define $\delta a_i = (a_{i+1} - a_i)/(0.2\,\mathrm{s})$ and $q_i = \frac{1}{7}\sum_{d=1}^{7} \delta a_{i,d}\delta a_{i+1,d}$. The boundary set contains indices satisfying $\lfloor i/4 \rfloor \neq \lfloor (i+2)/4 \rfloor$; thus two products touch each interior seam. Remaining indices are within-chunk. The probe averages q_i within each set per episode, then equally over 20 episodes, giving $+0.418$ within chunks and -3.027 at boundaries. There is no centering or variance normalization. The statistic mixes pose and gripper coordinates in their stored scales; its coordinate-squared-per-second-squared dimensions do not constitute a single physical unit or a normalized correlation.

C.3 Additional diagnostic breadth

Table 10 reports heterogeneous open-loop probes retained from the diagnostic study (Guo et al., 2022; Zhang et al., 2023; Tevet et al., 2023; Bjorck et al., 2025). HumanML3D Condition A adds independent Gaussian noise with $\sigma = 0.05$ to each frame of the recorded reference motion. The T2M-GPT and Motion Diffusion Model (MDM) rows are separately generated motion aggregates. The GR00T row is a constructed signal from one Bridge release with an eight-action horizon, as detailed below; it is not a horizon comparison.

GR00T signal construction. Within each episode, the probe queries every fourth recorded observation, concatenates all eight predicted actions per query, and cumulatively sums the seven action fields, including orientation and gripper. The diagnostics use this constructed sequence with a unit sample interval. Its predicted target-time ranges overlap; it is not a time-aligned executed command stream. The inputs use the named fields of the WidowX embodiment adapter (`OXE_WIDOWX`), but their stored-input semantics have not been independently validated.

Table 10: Additional diagnostic probes reported at their recorded denominators. Human-motion jerk shares one joint-space definition within that block. The GR00T row uses the constructed signal defined below, at a different scale from the HumanML3D block; its uncertainty is the population standard deviation over episodes.

Domain / probe	Temporal mechanism	n	Mean ρ_v	Mean jerk
HumanML3D reference	Recorded motion	200	$+0.856$	19.9
HumanML3D Condition A	Independent per-frame noise	200	-0.354	789.4
T2M-GPT	Autoregressive sequence	50	$+0.809$	432.7
MDM	Full-sequence diffusion	50	$+0.842$	22.2
GR00T-N1.6-Bridge	Constructed chunk signal, $H = 8$	30	$+0.635 \pm 0.074$	0.177 ± 0.079

These rows support diagnostic observability only. They do not form a matched cross-model comparison and do not test target-time preservation, SG4, or closed-loop task success.

Sampled-chunk records with unverified temporal adjacency. An earlier DROID probe records 30 sampled chunks per π_0 decoder (Black et al., 2025a; Khazatsky et al., 2024; Pertsch et al., 2025). The stored boundary/within-chunk jerk ratios are $18.92\times$ for π_0-DROID and $11.05\times$ for FAST. The sampling description states uniform sampling across the dataset subset; retained records do not establish same-episode, consecutive observation or target times. We preserve these ratios as sampled-chunk diagnostics, not as evidence of temporal boundaries in an executed stream. They are separate from the closed-loop $\pi_{0.5}$ comparison in Appendix D.3.

C.4 Access requirements for temporal coupling

Table 11 summarizes the information and intervention points of the mechanisms discussed in Section 2.

Table 11: Interface map for temporal coupling. The table lists the information each method requires, where it acts, and whether we evaluate it.

Mechanism	Required access	Coupling locus	Evaluation status
Generation-integrated same-target ensemble (ACT principle)	Overlapping timestamped predictions	Prediction aggregation	Evaluated baseline; reference stream
SW(4) / SG4 / EMA	Selected causal command stream only	Output stream	Evaluated operators
BID / ACG / TAS	Candidate chunks, sampling trajectory, or cached predictions	Generation or selection	Not evaluated
RTC / REMAC / RACE	Sampling, learned correction, or desired-state retiming	Asynchronous continuation or execution	Not evaluated

D Smoothing strength and execution timing

These follow-up comparisons retain the original Octo and three-operator OpenVLA-OFT results as separate studies. All conditions within each new study share the checkpoint and reset identity, not post-intervention observations or policy outputs. We report each complete operator set and test family.

D.1 Expanded OpenVLA-OFT comparison

The released checkpoint, LIBERO-10 tasks, and official states 5–49 per task are the same as in the original comparison. All eight conditions were separately reacquired; these results are not pooled with the original three-operator comparison. This is 450 paired resets and $3{,}600$ episodes, not a new held-out reset bank or training-seed replication.

Operators and measurements. For a selected pose vector u_t, EMA uses $\tilde{u}_t = \alpha u_t + (1-\alpha)\tilde{u}_{t-1}$, initialized with $\tilde{u}_0 = u_0$. Thus α weights the current raw command and larger α means weaker smoothing. EMA(0.70) was chosen by offline jerk matching to SG4, not by these success outcomes. The realized closed-loop jerk is not exactly matched. All smoothers leave the gripper unfiltered. The added-delay arms retain SG4 and delay its pose output by one or two command indices, holding the current raw gripper; startup clamps to the available prefix. Delayed histories are not flushed after episode termination.

The pose jerk is the episode mean of the six-coordinate third-difference norm divided by Δ^3, then averaged within each task and equally across ten tasks ($\Delta = 0.05\,\mathrm{s}$). Reductions in Table 12 are ratios of these overall means, relative to this study's native mean; they are not means of per-task ratios.

Table 12: Expanded OpenVLA-OFT comparison: all eight conditions are separately reacquired on 450 task/reset pairs and not pooled with the original three-operator comparison. Jerk reduction is relative to this study's native sent-command jerk, not physical motion.

Condition	Success	Jerk reduction (%)
Native	$416/450$	0.00
SW(4)	$387/450$	71.69
SG4	$418/450$	36.72
EMA(0.55)	$407/450$	56.05
EMA(0.70)	$406/450$	39.88
EMA(0.85)	$422/450$	23.12
SG4 + pose delay 1	$403/450$	35.35
SG4 + pose delay 2	$363/450$	26.25

Table 13: Complete three-contrast family for expanded OpenVLA-OFT. Positive differences favor SG4; discordance lists SG4-only/comparator-only successes. Individual bootstrap intervals and the Holm-adjusted exact tests are different procedures: the first two intervals do not establish significance under the Holm family. Differences and intervals are percentage points.

SG4 minus	Difference [individual 95% interval]	Discordance	Exact p	Holm p
EMA(0.70)	$2.667\ [0.222, 5.111]$	$23/11$	0.0576	0.0576
Pose delay 1	$3.333\ [0.667, 6.000]$	$28/13$	0.0275	0.0551
Pose delay 2	$12.222\ [8.889, 15.556]$	$64/9$	2.379×10^{-11}	7.137×10^{-11}

The bootstrap jointly resamples all conditions of each reset within each of the ten fixed tasks (100,000 resamples; seed 20260905). These are conditional resampling intervals, not uncertainty over new tasks or training seeds. Tuned EMA supplies useful trade-offs, but neither SG4 superiority over EMA(0.70) nor equivalence is established. The two-step delay contrast supports the importance of execution timing with SG4 retained; feedback, transients, episode duration, and pose–gripper timing remain part of its total closed-loop effect.

D.2 Pose–gripper channel-delay factorial

A separate acquisition uses official states 5–24 in each of the same ten tasks (200 pairs, 800 episodes). Every condition applies SG4 to pose first; the reference is SG4, not native. Let s_{ab} denote success rate with added pose delay a and gripper delay b, where $a, b \in \{0, 2\}$. The two-step delay is $0.10\,\mathrm{s}$ nominal control time, not measured wall-clock latency. Table 14 reports all four conditions.

All four outcomes are jointly resampled within each fixed task (100,000 resamples; seed 20260907). Equal added delay leaves a loss relative to SG4: failures rise from $17/200$ to $35/200$. However, the unresolved interaction cannot allocate the earlier pose-only loss between relative timing and common staleness. Startup, termination, and feedback changes remain in the contrast.

Table 14: SG4 channel-delay factorial. Equal added delay shifts filtered pose and raw gripper together; it does not restore their pre-filter relationship or physical synchronization. Jerk is the mean six-coordinate sent-pose jerk.

Added pose/gripper delay	Success	Failures	Jerk
0/0 (SG4 reference)	183/200	17/200	368.639
2/0	169/200	31/200	421.474
0/2	187/200	13/200	383.590
2/2	165/200	35/200	425.378

Table 15: Complete SG4 channel-delay family, in percentage points. The primary interaction and secondary gripper addition remain unresolved. Intervals are approximate Bonferroni-three percentile bootstrap intervals (98.333% each).

Role	Contrast	Difference	Corrected interval
Primary interaction	$(s_{22} - s_{20}) - (s_{02} - s_{00})$	-4.0	$[-11.0, 3.0]$
Secondary common delay	$s_{22} - s_{00}$	-9.0	$[-15.5, -2.5]$
Secondary gripper addition	$s_{22} - s_{20}$	-2.0	$[-7.5, 3.5]$

D.3 Bounded $\pi_{0.5}$ comparison

We evaluate one released `pi05_libero` checkpoint using OpenPI commit `215abfb217dbac7d5f1273282331b9b1866c0479`.[2] The same ten LIBERO tasks and official states 5–24 give 200 paired resets and 1,200 episodes. The policy returns ten actions per query and executes the first five, at a nominal 0.05 s control step, with ten neutral warmup actions and a cap of 520 decision actions. Actions are passed directly in LIBERO's seven-coordinate convention, without the OpenVLA-OFT gripper conversion. All smoothers act on pose only. Common delay shifts the selected pose and raw gripper together by two indices, clamps to the available prefix at startup, and has no terminal flush.

Table 16: Complete six-condition $\pi_{0.5}$ comparison on one checkpoint. All 200 pairs are retained as run. The operator ordering and SG4 delay response remain unresolved under the full contrast family.

Condition	Success	Sent-pose jerk
Native	185/200	769.418
SW(4)	181/200	187.318
SG4	182/200	451.770
EMA(0.70)	188/200	440.235
SG4 + common delay 2	178/200	505.282
EMA(0.70) + common delay 2	169/200	529.440

The analysis jointly resamples six-condition reset vectors within ten fixed tasks (100,000 resamples; seed 20260907). Delayed EMA loses success; the other eight corrected intervals include zero. This does not establish that SG4 tolerates delay better than EMA.

Physics warnings and sensitivity. Six MuJoCo joint-acceleration (`QACC`) instability warning lines are strongly linked to task 5, state 18 across all six arms; all six outcomes are failures. The warning's exact cause and phase are not certified. We retain the as-run 200-pair primary and separately remove this entire pair, not individual arms, in a 199-pair sensitivity analysis with equal task weighting. EMA's delay effect changes from -9.50 to -9.53 points (interval $[-17.53, -2.00]$); SG4's from -2.00 to -2.05 ($[-8.13, 4.00]$); and their interaction from 7.50 to 7.47 ($[-2.03, 16.97]$). The main contrast conclusions persist, but this does not certify counterfactual simulator validity or make kinematics invariant.

[2] Implementation: `https://github.com/Physical-Intelligence/openpi/tree/215abfb217dbac7d5f1273282331b9b1866c0479/examples/libero`; checkpoint: `gs://openpi-assets/checkpoints/pi05_libero`.

Table 17: All nine $\pi_{0.5}$ contrasts. Differences and approximate Bonferroni-nine percentile bootstrap intervals are percentage points. Delay interaction is the SG4 delay effect minus the EMA delay effect.

Contrast	Difference	Corrected interval
SW(4) minus native	-2.0	$[-8.0, 4.0]$
SG4 minus native	-1.5	$[-7.0, 4.0]$
EMA(0.70) minus native	1.5	$[-4.0, 7.0]$
SG4 minus EMA(0.70)	-3.0	$[-8.5, 2.5]$
SG4 common delay 2 minus SG4	-2.0	$[-8.0, 4.0]$
EMA common delay 2 minus EMA	-9.5	$[-17.5, -2.0]$
Delay interaction	7.5	$[-2.0, 17.0]$
SG4 minus SW(4)	0.5	$[-6.0, 7.0]$
EMA(0.70) minus SW(4)	3.5	$[-2.5, 9.5]$

E EXECUTION, COMMAND-PATH FITTING, AND TASK-RESOLVED OUTCOMES

E.1 WHERE THE FILTER ACTS

The velocity and jerk diagnostics in Section 3 are finite differences of command coordinates. For delta-pose commands, those coordinates are requested pose increments, not end-effector positions; their differences are not physical velocity or jerk.

The linear operator is applied to the selected policy-command stream, before runtime conversion and the simulator controller. It is not a learned layer. In the Octo operator and window studies, the inference wrapper queries the policy at every control step (5 Hz), aggregates overlapping predictions for the same target time, and returns one selected seven-coordinate command. The pose filter then acts on its first six coordinates. Octo converts the rotational Euler increments to axis–angle and thresholds the current gripper score at 0.5 to obtain -1 or $+1$; the action scale is 1.0. Thus "generation-integrated" identifies access to overlapping predictions before command selection, not aggregation inside the learned network.

The initial and repeat evaluations are labeled Canonical and Fresh in the archived records. The earlier factorial differs in channel scope: the initial evaluation's archived SW(4) wrapper averages all seven coordinates, including the gripper score before thresholding. The repeat evaluation declares the same evaluation protocol, but its referenced runner is not retained in the archive used here; that declaration is not an independent source-code check. A separate reconstruction of all $2{,}400$ initial/repeat raw–sent records reproduces the available-prefix filter, coordinate conversion and gripper threshold within floating-point rounding (maximum residual below 10^{-11}, including an alternate rotation implementation). This checks saved-command arithmetic, not the repeat evaluation's runtime timing. The archive also lacks the factorial episode-to-geometry mapping; we do not assign the later operator studies' 24 geometry phases to these records. The pose-only loop below specifies the later operator/window studies and OpenVLA-OFT, not the factorial.

In the later six-arm and window studies, each arm starts an episode with the same recorded row-specific environment seed and Octo sampling key. This couples initial stochastic conditions, not subsequent observations or trajectories. Their archived launch guards pin the SimplerEnv and Octo source revisions; the retained records do not specify the ensemble-weight schedule or a model-weight digest. The earlier factorial records an initial RNG seed but not per-episode policy keys or generation-time ensemble-source hashes, so it does not establish the same stochastic coupling.

The original and expanded OpenVLA-OFT runs query eight denormalized actions, execute them one at a time, and query again when the queue is empty. The nominal control step is $0.05\,\mathrm{s}$; ten neutral warmup actions precede policy execution, whose cap is 520 actions. The original comparison copies one canonical first chunk to all arms after their identity check. Later queries use each arm's own observation. The pose filter acts after queue selection; the upstream conversion binarizes and sign-inverts the raw gripper command. Appendix D specifies the additional delay channels and the distinct $\pi_{0.5}$ query/execution schedule.

At the nominal command cadence, the affine SW(4) displacement of 1.5 steps corresponds to $300\,\mathrm{ms}$ for Octo (5 Hz) and $75\,\mathrm{ms}$ for OpenVLA-OFT and $\pi_{0.5}$ (20 Hz). This converts command indices, not measured physical latency or the switching time of a thresholded gripper score.

```
reset environment, policy state, queue, and filter history
repeat until success or the recorded episode cap:
  u = select command (query/aggregate or pop queue)
  v.pose = causal filter(u.pose, history)
  v.gripper = u.gripper
  v = declared additional delay(v), if present
  a = runtime coordinate/gripper conversion(v)
  observation, outcome = environment.step(a)
  record raw u, operator output v, and outcome
```

For these pose-only studies, SW(4) averages the available prefix until four samples exist. SG4 passes through the first three samples. SG8/SG12 also pass through the first three, then fit using $\min(N, \text{available samples})$ until their full window is available. EMA initializes at the first raw pose. The Octo pure-delay controls pass through raw commands until the required history exists. No

filter or queue is flushed after termination. The six-arm Octo EMA uses $\alpha = 0.4$: its affine lag $(1-\alpha)/\alpha = 1.5$ and independent-noise variance gain $\alpha/(2-\alpha) = 0.25$ match SW(4), not SG4's realized jerk.

Neither project pose filter adds clipping. This does not remove downstream controller limits: LIBERO's operational-space pose controller clips inputs to $[-1, 1]$ and rescales them before setting its pose goal. Its gripper conversion is separate. Octo also retains the simulator controller after coordinate conversion. The moment identities and operator diagnostics concern the recorded command coordinates at the filter output, not controller-internal goals, torques, or realized motion. We do not infer controller saturation rates from these identities.

E.2 What the fitted shift measures

The Octo fit compares each rollout's own raw and filtered XYZ commands; it never treats diverged arms as a common raw trajectory. Let $P_r(k)$ and $P_s(k)$ be their cumulative sums through step k. For each candidate $d \in \{0, 0.05, \ldots, 4\}$, linearly interpolate P_r at $k - d$ and use only $\mathcal{I}_d = \{k : d \le k \le T - 1\}$, where T is the rollout length. With $\bar{e}_d$ the mean three-dimensional error on that overlap, the fit is

$$e_d(k) = P_s(k) - P_r(k-d), \qquad \hat{d} = \arg\min_d \sqrt{\frac{1}{|\mathcal{I}_d|} \sum_{k \in \mathcal{I}_d} \|e_d(k) - \bar{e}_d\|_2^2}.$$

The constant offset absorbs translation from causal startup but does not remove time-varying shape error. Equal scores select the smaller d. The overlap changes with d; there is no extrapolation or additional trimming. Rollouts shorter than six samples are excluded from this diagnostic. The remaining root-mean-square error is divided by the aligned raw path's root-mean-square excursion about its mean; a denominator at most 10^{-12} is undefined. Study summaries average episode-level fits and defined residuals.

A scalar shift need not explain a curved command. For example, after startup, SG4 maps $u_t = (t, t^2)$ to $(t, t^2 - 1)$: the first coordinate forces any exact time shift to zero, while the second still differs. Eq. 6 predicts this curvature error, not its reduction to a time shift. Our fit is therefore a descriptive alignment with a remaining shape residual, not a tangent/normal decomposition or identified physical latency. Table 18 reports the existing fits and residuals.

Table 18: Existing Octo command-path fits, kept separate from the window study. Each entry is an episode mean over 96 rollouts. Residual is the normalized shape error after shift and constant-offset fitting, not an orthogonal-path component. Exact pure-delay fits validate those controls, not identification on arbitrary trajectories.

Study	Operator	Fitted shift (steps)	Residual
Six-arm	No smoothing	0.000	0.000000
	SW(4)	1.501	0.034922
	EMA(0.4)	1.345	0.064015
	SG4	0.113	0.040070
	Pure delay 1	1.000	$< 10^{-12}$
	Pure delay 2	2.000	$< 10^{-12}$
Second stream	No smoothing	0.000	0.000000
	SW(4)	1.500	0.036174
	SG4	0.114	0.038129

E.3 Aggregate agreement does not imply taskwise preservation

Tables 19–21 retain the original operator, window, and OpenVLA-OFT studies separately. Octo uses 48 episodes per task: two stochastic streams over the same 24 geometry phases. OpenVLA-OFT uses 45 official states per task. These are fixed-checkpoint evaluation outcomes, not independent training seeds. Discordances count left-only/right-only successes; their difference equals the difference of the corresponding success counts. They expose offsetting outcomes without introducing new taskwise significance tests or interpreting non-significance as equivalence.

In the two-stream Octo comparison, Carrot falls from $5/48$ without smoothing to $0/48$ with SG4, while gains on Eggplant, Spoon, and StackCube sum to five. Hence the pooled $48/192$ tie is not uniform preservation. The counts establish which task regresses, not why its manipulation fails. Terminal length alone also cannot distinguish a timeout-limited attempt from an unsuccessful manipulation that would remain unsuccessful with more time.

Table 19: Original two-stream Octo comparison by task. Success counts use 48 episodes per task. The last columns are paired left-only/right-only successes, written $(L, R)/48$ to distinguish the paired denominator from a ratio. These task counts sum to the main-text $48/31/48$ of 192; no smoothing is the native ensemble-selected stream.

Task	No smoothing	SW(4)	SG4	SG4–no	SG4–SW(4)
Carrot	$5/48$	$2/48$	$0/48$	$(0, 5)/48$	$(0, 2)/48$
Eggplant	$26/48$	$18/48$	$28/48$	$(8, 6)/48$	$(16, 6)/48$
Spoon	$16/48$	$10/48$	$18/48$	$(7, 5)/48$	$(13, 5)/48$
StackCube	$1/48$	$1/48$	$2/48$	$(1, 0)/48$	$(1, 0)/48$

Table 20: Separate Octo window study, two further streams with 48 episodes per task. Top: successes. Bottom: paired left-only/right-only successes $(L, R)/48$ for the named contrasts. The pooled counts are $51/30/59/42/33$ of 192; these rows are not pooled with Table 19.

Task	No smoothing	SW(4)	SG4	SG8	SG12
Carrot	$4/48$	$1/48$	$3/48$	$5/48$	$2/48$
Eggplant	$27/48$	$16/48$	$31/48$	$27/48$	$25/48$
Spoon	$17/48$	$12/48$	$23/48$	$10/48$	$6/48$
StackCube	$3/48$	$1/48$	$2/48$	$0/48$	$0/48$

Task	SG4–no	SG4–SW(4)	SG8–SG4	SG12–SG4
Carrot	$(0, 1)/48$	$(2, 0)/48$	$(3, 1)/48$	$(1, 2)/48$
Eggplant	$(9, 5)/48$	$(16, 1)/48$	$(4, 8)/48$	$(5, 11)/48$
Spoon	$(10, 4)/48$	$(15, 4)/48$	$(3, 16)/48$	$(2, 19)/48$
StackCube	$(1, 2)/48$	$(2, 1)/48$	$(0, 2)/48$	$(0, 2)/48$

Table 21: Original OpenVLA-OFT comparison by LIBERO-10 task and official states 5–49. Success and discordance denominators are 45 per task; paired discordances are $(L, R)/45$ for each named contrast. Pooled successes are $416/387/418$ of 450. These are the original three-arm data, not the separately reacquired eight-condition study in Appendix D. Task identifiers refer to the instructions listed in this subsection.

Task	Native	SW(4)	SG4	SG4–native	SG4–SW(4)
0	$40/45$	$38/45$	$40/45$	$(1, 1)/45$	$(3, 1)/45$
1	$45/45$	$44/45$	$44/45$	$(0, 1)/45$	$(1, 1)/45$
2	$43/45$	$37/45$	$43/45$	$(0, 0)/45$	$(7, 1)/45$
3	$43/45$	$44/45$	$45/45$	$(2, 0)/45$	$(1, 0)/45$
4	$44/45$	$40/45$	$41/45$	$(0, 3)/45$	$(3, 2)/45$
5	$40/45$	$41/45$	$41/45$	$(1, 0)/45$	$(0, 0)/45$
6	$37/45$	$27/45$	$36/45$	$(3, 4)/45$	$(11, 2)/45$
7	$43/45$	$45/45$	$45/45$	$(2, 0)/45$	$(0, 0)/45$
8	$38/45$	$32/45$	$39/45$	$(3, 2)/45$	$(10, 3)/45$
9	$43/45$	$39/45$	$44/45$	$(1, 0)/45$	$(6, 1)/45$

Octo task names abbreviate putting a carrot on a plate, putting an eggplant in a basket, putting a spoon on a tablecloth, and stacking the green cube on the yellow cube. The LIBERO-10 instructions, indexed as in Table 21, are:

0. Put both the alphabet soup and the tomato sauce in the basket.
1. Put both the cream cheese box and the butter in the basket.

2. Turn on the stove and put the moka pot on it.
3. Put the black bowl in the bottom drawer of the cabinet and close it.
4. Put the white mug on the left plate and the yellow and white mug on the right plate.
5. Pick up the book and place it in the back compartment of the caddy.
6. Put the white mug on the plate and the chocolate pudding to the right of the plate.
7. Put both the alphabet soup and the cream cheese box in the basket.
8. Put both moka pots on the stove.
9. Put the yellow and white mug in the microwave and close it.

E.4 Separating temporal alignment from remaining path error

Time alignment removes most of SW(4)'s command-path error, while long SG windows retain increasing distortion after alignment. This post-hoc analysis reuses all $1{,}824$ saved Octo rollouts: 576 from the six-arm study, 288 from its three-arm repeat, and $480 + 480$ from the separate window study. Each rollout supplies its own raw and operator-output XYZ commands, before runtime conversion; diverged arms are not a shared-input comparison. These cumulative command paths are not measured end-effector trajectories.

A common comparison interval. Using the paths P_r, P_s defined in Appendix E.2, set $\mathcal{I}_c = \{4, \dots, T-1\}$ for every candidate $d \in \{0, 0.05, \dots, 4\}$. With linear interpolation, define

$$e_d(k) = P_s(k) - P_r(k-d), \quad b_d^c = \frac{1}{|\mathcal{I}_c|} \sum_{k \in \mathcal{I}_c} e_d(k), \quad r_d^c(k) = e_d(k) - b_d^c, \quad E_d^c = \sum_{k \in \mathcal{I}_c} \|r_d^c(k)\|_2^2.$$

The common-overlap fitted shift $\hat{d}_c$ minimizes E_d^c, with exact ties choosing the smaller d. Let S_c be the sum of squared deviations of $P_r(k)$ about its mean on $\mathcal{I}_c$, fixed across d. The normalized error is $R_d^c = \sqrt{E_d^c / S_c}$; the additional energy removal is $\eta = (E_0^c - E_{\hat{d}_c}^c)/E_0^c$. Thus both errors already exclude a constant spatial offset. Ratios with denominator at most $10^{-12} \max(1, S_c)$ are undefined. We require $T \geq 10$; all $1{,}824$ rollouts qualify, with no flat cost profile or optimum at the four-step bound. Unlike the earlier fit, this diagnostic fixes both overlap and normalization; its new values do not replace the legacy shifts in the main text.

Alignment and local geometry. Table 22 reports episode means, including the mean of η, not a ratio of pooled energies. The original two-stream SW(4)/SG4 comparison has similar-sized post-alignment errors, with SG4's mean slightly larger. The 96.9% removal for SW(4) concerns offset-corrected command-path error, not the fraction of success loss caused by timing or the benefit of correcting and rerunning the policy.

For a local geometric split, central differences of P_r (one-sided at the ends), interpolated at $k - d$, define a unit tangent $\tau_d(k)$. Project r_d^c onto τ_d and its orthogonal complement at $d = 0$ and $\hat{d}_c$. Tangents whose speed is at most 10^{-8} times the maximum raw derivative norm are undefined; their error energy is retained separately. The same raw threshold is used before and after alignment. The original pooled SW(4)/SG4 mean normal-energy fractions are $16.7\%/48.3\%$ before and $52.1\%/57.4\%$ after alignment. These fractions use valid-tangent energy; an increased fraction need not mean increased absolute error, and the tangent also changes. A finite delay on a curved path can itself have a normal component. This geometric split is not a causal time/shape decomposition. In the window study, SG4/SG8/SG12 post-alignment normal RMS values, normalized by S_c as above, increase from 0.0279 to 0.0871 and 0.1403. Both temporal alignment and remaining distortion therefore vary with window length.

Sensitivity and offset ambiguity. Refitting using only the first 20 commands retains all $1{,}824$ rollouts. Original pooled SW(4)/SG4 mean shifts are $1.504/0.128$ steps, removal is $95.8\%/26.3\%$, and residuals are $0.0647/0.0649$. Window SG4/SG8/SG12 residuals are $0.0586/0.1772/0.2791$. This repeats the descriptive contrast, not an independent validation removing startup or termination effects. A free offset can absorb the entire difference between paths (t, t^2) and $(t, t^2 - 1)$, although they are not a pure time shift. Constant-velocity lag can also be ambiguous with a spatial offset.

Table 22: Post-hoc common-overlap geometry of saved Octo command paths. Values are episode means: success counts retain their denominators; R_0^c and $R_{\hat{d}_c}^c$ are dimensionless offset-corrected errors before/after alignment, and η is the additional energy removal. A dash denotes an undefined zero-error ratio. The original pooled group includes three arms of the six-arm group below; these are overlapping summaries, not extra trials. The window group is separate. Legacy fits in Table 18 use a different overlap and normalization.

Operator	Success	$\hat{d}_c$ (steps)	R_0^c	$R_{\hat{d}_c}^c$	η (%)
Original two-stream comparison, 192 rollouts per arm					
No smoothing	$48/192$	0.000	0.0000	0.0000	—
SW(4)	$31/192$	1.500	0.2115	0.0358	96.9
SG4	$48/192$	0.110	0.0445	0.0386	19.8
Six-arm study alone, 96 rollouts per arm					
No smoothing	$24/96$	0.000	0.0000	0.0000	—
SW(4)	$14/96$	1.499	0.2097	0.0353	96.9
EMA(0.4)	$18/96$	1.349	0.1915	0.0642	88.4
SG4	$25/96$	0.107	0.0452	0.0388	19.8
Pure delay 1	$20/96$	1.000	0.1625	$< 10^{-12}$	100.0
Pure delay 2	$12/96$	2.000	0.2866	$< 10^{-12}$	100.0
Separate window study, two streams, 192 rollouts per arm					
No smoothing	$51/192$	0.000	0.0000	0.0000	—
SW(4)	$30/192$	1.501	0.2070	0.0341	97.0
SG4	$59/192$	0.107	0.0422	0.0370	19.5
SG8	$42/192$	0.541	0.1451	0.1107	39.9
SG12	$33/192$	1.117	0.2515	0.1778	47.1

Accordingly, we retain original errors and offset vectors: pooled SW(4)/SG4 unregistered RMS is $0.01871/0.00478$, offset norm at zero shift is $0.01018/0.00407$, and offset norm at the optimum is $0.01015/0.00408$, in native command-path units. Offset-corrected residuals do not measure total path distortion. Taskwise outcomes remain those in Appendix E.3; this analysis does not establish an episode-level success predictor.

E.5 Paired uncertainty with repeated geometry retained

Equal pooled success counts do not establish equal success probabilities. This post-hoc sensitivity retains the four fixed tasks and all 192 paired rows per study. One bootstrap resamples 48 paired rows within each task. The other resamples 24 geometry-phase clusters within each task, retaining both stochastic streams and every arm in each sampled cluster. Recorded episode identifiers and offsets map to the same modulo-24 geometry cycle; stream identity is preserved rather than equating episode labels across streams. Both schemes use 100,000 draws with seed 20260911; Table 23 reports every contrast.

The original SG4–SW(4) contrast remains positive under both schemes, while SG4–no smoothing spans losses and gains. Shared geometry does not turn the aggregate tie into equivalence or task-wise preservation. These are empirical resampling sensitivities conditional on the recorded studies, not uncertainty over new tasks, training seeds, or an adequately sampled population of stochastic streams. We do not select the narrower interval or revise the original confirmatory hierarchy from these post-hoc pointwise intervals.

E.6 Allowing signed shifts on a common interval

Allowing negative shifts tests whether restricting the earlier search to nonnegative values concealed alignments to later raw command indices. We refit all 1,824 saved paths on $\mathcal{I}_s = \{4, \dots, T-5\}$, requiring $T \geq 10$. On this identical overlap, compare $d = 0, 0.05, \dots, 4$ with $d = -4, -3.95, \dots, 4$ in the offset-corrected fit of $P_s(k)$ to $P_r(k-d)$. Both use linear interpolation without extrapolation and the same raw excursion normalizer on $\mathcal{I}_s$. Numerically tied minima use the smallest $|d|$, then the smallest d (relative tolerance 10^{-12}, absolute tolerance 10^{-24} in squared-error units).

Table 23: All pairwise Octo contrasts in the original three-arm and separate five-arm window pools (192 paired episodes per contrast). Differences and pointwise 95% percentile intervals are percentage points, with four tasks held fixed. Row and geometry-cluster schemes are reported together; neither column is a simultaneous interval family or a replacement for the original McNemar/Holm tests.

Left minus right	Difference	Paired rows	Geometry clusters
Original two-stream operator comparison			
No smoothing minus SW(4)	8.85	$[2.08, 15.10]$	$[1.56, 16.15]$
SG4 minus SW(4)	8.85	$[2.60, 15.10]$	$[2.60, 15.10]$
SG4 minus No smoothing	0.00	$[-5.73, 5.73]$	$[-5.73, 5.21]$
Separate two-stream window study			
No smoothing minus SW(4)	10.94	$[5.21, 16.67]$	$[4.69, 16.67]$
SG4 minus SW(4)	15.10	$[9.38, 20.83]$	$[8.85, 21.35]$
SG8 minus SW(4)	6.25	$[0.00, 12.50]$	$[0.52, 11.98]$
SG12 minus SW(4)	1.56	$[-4.17, 7.29]$	$[-4.17, 7.29]$
SG4 minus No smoothing	4.17	$[-1.56, 9.90]$	$[-1.04, 9.38]$
SG8 minus No smoothing	-4.69	$[-10.42, 1.04]$	$[-10.94, 1.56]$
SG12 minus No smoothing	-9.38	$[-15.63, -3.13]$	$[-16.15, -2.60]$
SG8 minus SG4	-8.85	$[-14.58, -3.13]$	$[-14.58, -3.13]$
SG12 minus SG4	-13.54	$[-19.79, -7.29]$	$[-20.31, -6.77]$
SG12 minus SG8	-4.69	$[-9.38, 0.00]$	$[-9.38, 0.00]$

Table 24: Signed-fit sensitivity on the shorter common interval $\mathcal{I}_s$; episode means in steps and normalized offset-corrected RMS. Every row has zero negative optima, and its unsigned fit on the same interval agrees. Original pooled and six-arm groups overlap; the window pool is separate. These are reanalyses, not new trials or replacements of earlier fits.

Operator	Rollouts	Signed shift	Residual
Original two-stream comparison			
No smoothing	192	0.000	0.0000
SW(4)	192	1.501	0.0369
SG4	192	0.109	0.0392
Six-arm study alone			
No smoothing	96	0.000	0.0000
SW(4)	96	1.501	0.0363
EMA(0.4)	96	1.355	0.0663
SG4	96	0.107	0.0395
Pure delay 1	96	1.000	$< 10^{-12}$
Pure delay 2	96	2.000	$< 10^{-12}$
Separate two-stream window study			
No smoothing	192	0.000	0.0000
SW(4)	192	1.501	0.0350
SG4	192	0.107	0.0376
SG8	192	0.543	0.1134
SG12	192	1.121	0.1820

All $1{,}824$ paths qualify. No negative optimum, endpoint-boundary optimum, flat profile, or tied minimum occurs in these records. Signed and unsigned optima on this common interval agree to numerical precision. Table 24 reports all operators; its pools overlap as in Table 22. Synthetic curved positive- and negative-shift controls recover $+1$ and -1 steps; stationary and affine-offset controls retain their ambiguity.

Changing the overlap alone changes $164/1{,}824$ fitted grid optima relative to Appendix E.4, whose interval is $\{4, \ldots, T-1\}$. The older variable-overlap fits in Appendix E.2 are distinct again. None is overwritten by this sensitivity. Signed registration does not identify physical latency or the cause of failure; a free offset can still absorb non-temporal path error. The test addresses the sign constraint, not the sufficiency of a scalar shift to describe curved commands.

Table 25: Saved sent-command jerk under full and fixed intervals, in each study's recorded pose-coordinate and timestep convention. All 27 source arms are retained; each Octo arm has 96 traces and each expanded OpenVLA-OFT arm has 450. Entries are equal-task means, not cross-platform physical jerk comparisons, paired confidence intervals, or new task-success results. The two Octo window streams are distinct from the original and repeat streams.

Operator	Full trace	First 20	Indices 12–31
Octo original six-arm stream (96 traces/arm)			
No smoothing	8.248	9.800	8.098
SW(4)	1.890	2.372	1.668
SG4	4.729	6.055	4.417
EMA(0.4)	2.511	2.882	2.294
Pure delay 1	8.940	11.007	8.411
Pure delay 2	9.745	12.602	8.481
Octo three-arm repeat stream (96 traces/arm)			
No smoothing	8.406	9.856	8.217
SW(4)	1.910	2.345	1.698
SG4	4.952	6.067	4.648
Octo window stream 1 (96 traces/arm)			
No smoothing	8.283	9.666	8.589
SW(4)	1.889	2.278	1.700
SG4	4.882	6.249	4.664
SG8	2.903	4.294	2.465
SG12	2.292	3.859	1.777
Octo window stream 2 (96 traces/arm)			
No smoothing	8.373	9.750	8.264
SW(4)	1.867	2.265	1.670
SG4	4.833	6.106	4.586
SG8	2.900	4.234	2.551
SG12	2.259	3.738	1.767
Expanded OpenVLA-OFT (450 traces/arm)			
Native	578.024	465.575	433.771
SW(4)	163.610	119.311	128.373
SG4	365.770	276.773	274.292
EMA(0.55)	254.032	218.221	197.899
EMA(0.70)	347.514	289.067	265.571
EMA(0.85)	444.369	371.110	343.064
SG4 + delay 1	373.685	309.106	300.461
SG4 + delay 2	426.307	336.144	339.608

E.7 Fixed-window and spectral command diagnostics

We reanalyze every saved trace from the four Octo operator/window streams (1,824) and the expanded eight-arm OpenVLA-OFT study (3,600), keeping the five source groups separate. These 5,424 traces form 834 complete reset groups. All arms in every group have finite, paired raw/sent arrays and at least 32 commands; no success or failure record is excluded. These are existing records, not additional trials.

For each trace, sent-command jerk is the mean norm of the third difference of its six pose coordinates divided by the nominal timestep cubed (Octo 0.2 s; OpenVLA-OFT 0.05 s). We compare the full trace, its first 20 commands (indices 0–19), and commands 12–31 after the longest tested fitting window can be populated. Each 20-command interval yields 17 third differences. Table 25 reports equal-task means; the saved analysis also retains episode means, 95th/99th percentiles, maxima, channel-wise absolute third differences, and command ranges. In all five source groups, SW(4), SG4 and the unfiltered reference retain that ordering of mean jerk in each interval. SG4 is not uniformly smoother than tuned EMA: in the expanded study its order relative to EMA(0.70) changes between the first-20 and later intervals. Equal duration addresses unequal episode lengths, not the different feedback histories produced by the interventions.

For spectra, we use the first 32 raw/sent pose commands of the same complete groups, including startup, with a periodic Hann window, constant detrending, and a one-sided density periodogram at the nominal sample rate. Band power sums the density times frequency-bin width over normalized-Nyquist intervals $[0, 0.2)$, $[0.2, 0.5)$ and $[0.5, 1]$, with DC in the first band. We compute equal-task means separately for each source group and pose coordinate, then divide sent band power by the same arm's own raw-input band power. This is not a comparison to another arm's diverged trajectory.

For SG4, all 30 source-group/coordinate cells have ratios above one in the low (1.016–1.146) and middle (1.173–1.316) bands, and below one in the high band (0.314–0.380); every denominator is positive. Thus lower jerk coexists with nonuniform spectral attenuation in these recorded prefixes. These finite, startup-containing ratios are descriptive, not steady-state transfer gains, physical energy, or causes of task failure.

For comparison, direct evaluation of the standard base-filter responses on 100,001 frequencies from zero to π gives peak amplitude gains 1.154, 1.259 and 1.299 for SG4, SG8 and SG12, at approximately 0.873, 0.444 and 0.299 rad/sample. SW(4) and EMA with $\alpha = 0.40, 0.55, 0.70, 0.85$ all peak at one at DC. SG4's maximum is also exactly $\sqrt{261}/14$. These responses describe the linear filters, not downstream controller clipping or realized motion.

E.8 Illustration on a recorded command sequence

Figure 3 illustrates the distinction between same-target ensembling and output filtering on one recorded command stream. It is an offline command-space illustration, not a closed-loop comparison.

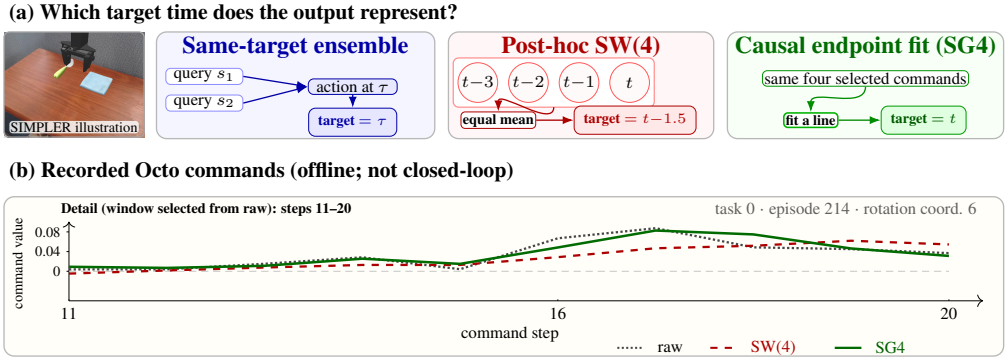

Figure 3: Target time separates two forms of smoothing. (a) Predictions from s_1, s_2 combine for target time τ. On linearly changing commands, the causal four-sample moving average (SW(4)) averages $t-3{:}t$, returning $t-1.5$; the causal Savitzky–Golay endpoint filter (SG4) fits a line, returning t. (b) One outcome-independent Octo trace, rotation-command coordinate 6 in stored units, filtered offline. In this example, the raw command and SG4 peak at the same step, whereas SW(4) peaks later. The 10-command detail surrounds the largest raw step change among the first 30, selected before filtering. The SIMPLER scene is illustrative, not the episode frame (Li et al., 2025). These are command-space traces, not closed-loop trajectories.

F MATCHED-MAGNITUDE TEMPORAL-WEIGHT COMPARISON

This appendix reports a separate OpenVLA-OFT campaign designed to ask whether the placement of temporal weights matters after attenuation has been held fixed. It is a *reversed-weight comparison*: recent-weighted and past-weighted causal four-tap FIR operators use the same coefficients in opposite temporal order. The campaign uses one fixed checkpoint, ten fixed LIBERO tasks, and 60 newly generated reset states per task. Thus there are 600 paired reset units and 3,000 completed arm runs. This state bank is separate from the earlier official 450-state bank; the two studies are not pooled.

F.1 OPERATORS AND THE MATCHED RESPONSE

For $q \in \{0.30, 0.45\}$, let $k = 0$ denote the current command and $k = 1, 2, 3$ the progressively older commands. The causal geometric kernel and its reversal are

$$h_q[k] = \frac{q^k}{\sum_{j=0}^{3} q^j}, \qquad h_q^{\mathrm{rev}}[k] = h_q[3-k], \qquad k \in \{0,1,2,3\}. \tag{8}$$

The kernel with the larger coefficient on the current command is called *recent-weighted*; its reversal is called *past-weighted*. Both operators are causal FIR filters. They are not recursive EMA or SG4, and the comparison adds no delay queue.

For angular frequency ω in radians per command step and imaginary unit i, define the base-FIR response

$$H_q(\omega) = \sum_{k=0}^{3} h_q[k] e^{-i\omega k}, \qquad H_q^{\mathrm{rev}}(\omega) = \sum_{k=0}^{3} h_q[3-k] e^{-i\omega k}. \tag{9}$$

Because the coefficients are real,

$$H_q^{\mathrm{rev}}(\omega) = e^{-3i\omega} \sum_{m=0}^{3} h_q[m] e^{+i\omega m} = e^{-3i\omega} \overline{H_q(\omega)}, \qquad \left|H_q^{\mathrm{rev}}(\omega)\right| = \left|H_q(\omega)\right|. \tag{10}$$

The overline denotes complex conjugation. Thus the matched quantity is the *common-input base-FIR magnitude response*, not the realized closed-loop jerk or a complete finite-record operator. The first moments, giving the exact past bias on affine command streams in command steps, are

q	$\mu_{q,\mathrm{recent}} = \sum_k k h_q[k]$	$\mu_{q,\mathrm{past}} = \sum_k k h_q^{\mathrm{rev}}[k]$
0.30	0.395906845	2.604093155
0.45	0.647143164	2.352856836

$$\mu_{q,\mathrm{past}} = 3 - \mu_{q,\mathrm{recent}}. \tag{11}$$

These moments describe the construction; they do not isolate a first-moment-only causal effect.

For completeness, the normalized SG4 coefficients in Eq. 5 have the finite step response $0.7 \to 1.1 \to 1.2 \to 1.0$ once the four-sample history is available. The peak is a 20% overshoot relative to the unit step. This is a filter-level transient calculation, not evidence that the transient caused any task failure.

F.2 RUNTIME CONTRACT AND PAIRING

Checkpoint: `moojink/openvla-7b-oft-finetuned-libero-10`, revision `95220f9a3421a7ff12d4218e73d09ade830fa9a3`, on the new fixed state bank. The recorded runtime identity was `4d5ca122246561962a42d96526fb824a75c88b9694b53220d8ffa688bb7ec470`. The key execution contract is summarized in Table 26. State-generation seeds are 2026093000 + 100 task + reset, with task 0–9 and reset 0–59. Policy/environment seeds are 7/0; ten neutral environment steps precede policy actions. Initial validity, not native success, determines eligibility. All 600 states were valid, unique and initially unsolved. Arm order rotates by (task + reset) mod 5; it is cyclic, not randomized. Initial simulator state, observation/input and first-chunk checks match across arms. Recorded controller input bounds are available for all runs with zero reconstructed crossings; this is not direct measurement of internal clipping or torque.

The reported jerk is the existing episode-level, Δt-scaled third difference norm of the six sent pose coordinates, averaged over episodes within each task and then equally over tasks, separately for

Table 26: Matched-time runtime and state scope. Each condition is paired on the task and reset state before intervention; observations, policy outputs, and action paths may diverge after the first differing sent command.

Item	Recorded contract
Tasks and states	LIBERO-10; 60 newly generated states per task; 600 paired resets; 5 arms and 3,000 runs
Checkpoint	OpenVLA-OFT revision `95220f9a3421a7ff12d4218e73d09ade830fa9a3`
Policy cadence	Query/execute cadence 8; control step $\Delta t = 0.05$ s; maximum 520 policy actions
Action path	Pose six-vector filtered; current raw gripper retained; first three selected pose commands pass raw, then the four-tap FIR runs continuously
Clipping and history	No added filter clipping; history persists across chunk boundaries and is reset between episodes; no terminal flush or added delay queue
Study separation	New fixed-state bank, separate from the earlier official 450-state study; no pooling of rows, jerk, or tests

each condition. It is a diagnostic of the realized sent commands. The comparison does not match realized jerk: startup and termination, feedback, phase, and pose–gripper relative timing can differ even when the common-input base-FIR magnitude response is identical.

F.3 Primary results

All five arms were retained with denominator $600/600$; no missing value was replaced by a failure or zero. Table 27 reports every arm's success and mean sent-pose jerk. The two recent-minus-past contrasts, and only these two primary contrasts, use two-sided exact conditional McNemar tests with Holm correction over a family of two. This exact test concerns conditional discordance symmetry, not an arbitrary heterogeneous-task null whose mean effects cancel. The discordance order is recent-only-success / past-only-success.

Table 27: All five arms of the matched-time campaign. Jerk is mean sent-pose jerk; reductions use this campaign's native mean and are not pooled with the official-450 study.

Condition	Success	Mean sent jerk	Native reduction (%)
Native	$553/600$	569.936814	—
$q = .30$ recent-weighted	$541/600$	344.152044	39.62
$q = .30$ past-weighted	$439/600$	446.600635	21.64
$q = .45$ recent-weighted	$537/600$	264.641780	53.57
$q = .45$ past-weighted	$457/600$	320.770927	43.72

Table 28: Primary paired success tests. Differences are recent-weighted minus past-weighted in percentage points. Holm correction is over the two stated contrasts.

Contrast	Difference	Discordance	Exact p	Holm-2 p
$q = .30$	$+17.00$	$124/22$	1.88×10^{-18}	3.77×10^{-18}
$q = .45$	$+13.33$	$99/19$	3.06×10^{-14}	3.06×10^{-14}

Within each of the ten fixed tasks, the paired reset vectors were resampled $100{,}000$ times with seed 20260905, sharing reset indices across arms and contrasts. The 95% intervals are individual intervals; the 97.5% intervals are Bonferroni intervals for the two-contrast family. They are conditional on this checkpoint, these ten tasks, and these reset states.

F.4 Task-stratified outcomes

Table 30 exposes the complete task-level counts; each entry is success over 60 resets. The number of tasks favoring the recent-weighted arm is descriptive, not a taskwise significance analysis: it is

Table 29: Fixed-task paired bootstrap intervals for the two primary contrasts, in percentage points. The resampling unit is the paired reset within each fixed task; these intervals are not new-task or training-seed inference.

Contrast	Individual 95%	Family-2 97.5%
$q = .30$ recent − past	$[13.500, 20.500]$	$[13.000, 21.000]$
$q = .45$ recent − past	$[10.167, 16.667]$	$[9.667, 17.000]$

$9/10$ for $q = .30$ and $8/10$ for $q = .45$. The opposite directions are retained (task 7 for $q = .30$, and tasks 3 and 7 for $q = .45$).

Table 30: Complete ten-task table for the matched-time campaign. Each cell is success/60; task identifiers follow the fixed LIBERO-10 ordering.

Task	Native	.30 recent	.30 past	.45 recent	.45 past
0	58/60	56/60	48/60	56/60	47/60
1	60/60	60/60	51/60	60/60	52/60
2	59/60	57/60	50/60	58/60	55/60
3	59/60	58/60	55/60	59/60	60/60
4	56/60	56/60	30/60	55/60	38/60
5	52/60	52/60	49/60	52/60	50/60
6	50/60	46/60	22/60	46/60	25/60
7	59/60	57/60	59/60	57/60	58/60
8	44/60	42/60	32/60	37/60	30/60
9	56/60	57/60	43/60	57/60	42/60

F.5 Warning disclosure and post hoc sensitivities

The as-run record contains 25 nonfatal MuJoCo warnings of the form `Nan, Inf or huge value in QACC` in task 5. Wall-clock matching under the recorded KST (UTC+9) assumption links them to five reset blocks. Every one of the five arms in each block reached the 520-action cap and was observed as a failure, as shown in Table 31.

Table 31: Time-mapped QACC warning blocks. The mapping identifies reset blocks, not the warning's simulator phase or cause.

Task-5 reset	Warning interval (KST, 2026-09-11)	Five-arm outcome
18	04:02:47–04:04:08	5/5 cap failures
22	04:07:31–04:08:53	5/5 cap failures
30	04:16:48–04:18:10	5/5 cap failures
46	04:34:35–04:35:56	5/5 cap failures
58	04:47:08–04:48:29	5/5 cap failures

The runner did not catch nonfatal QACC warnings or stop an episode on their appearance. This exposed a monitoring gap in the planned technical-error stop rule. Python exceptions and validation errors were propagated rather than imputed as task failures, but these simulator warnings were not caught by that path; the affected runs continued and were recorded under the ordinary cap rule. The warning cause, simulator phase, and physical impact are unknown. We therefore retain all 600 primary pairs and disclose the warnings rather than deleting or relabelling them.

The independent verification supplied two post hoc sensitivities, each kept separate from the primary denominator. Omitting only the five mapped reset blocks gives 595 paired resets with equal weighting over the same ten tasks; omitting the whole warning task gives 540 paired resets with equal weighting over nine tasks. Counts are recent-only/past-only discordances. The 95% and family-2 97.5% intervals again use 100,000 within-task draws with seed 20260905.

For the 595-pair analysis the exact discordances and Holm values are unchanged because each removed reset had no discordance. The 540-pair analysis is a whole-task omission, not a reconstruc-

Table 32: Post hoc warning sensitivities. These analyses do not replace the primary 600-pair result or certify warning-free physics. Differences and intervals are percentage points.

Retained pairs	Contrast	Discordance	Difference	Holm-2 p	95%	97.5%
595 / 10 tasks	$q = .30$	124/22	+17.045	3.7682×10^{-18}	[13.530, 20.576]	[13.030, 21.091]
595 / 10 tasks	$q = .45$	99/19	+13.364	3.0560×10^{-14}	[10.152, 16.652]	[9.682, 17.091]
540 / 9 tasks	$q = .30$	120/21	+18.333	9.6201×10^{-18}	[14.44, 22.22]	[13.89, 22.78]
540 / 9 tasks	$q = .45$	96/18	+14.444	4.7214×10^{-14}	[10.93, 17.96]	[10.37, 18.52]

tion of what warning-free execution would have produced. Neither sensitivity establishes a safety property or a non-inferiority relation to native execution.

F.6 Interpretation limits and provenance

The allowed interpretation is that, in this fixed-checkpoint, ten-task as-run comparison, reversing temporal weights changed closed-loop task success despite identical common-input base-FIR magnitude response. It is not a claim that attenuation is irrelevant: the paired operators also differ in phase, first moment, transients, startup and termination behavior, pose–gripper relative timing, feedback trajectory, and realized sent jerk. The result does not isolate a first-moment-only cause, and it does not compare executions with the same realized jerk.

This campaign does not test native equivalence or non-inferiority, safety, new-task or policy-family generalization, training-seed stability, or retrospective success-prediction gain. The fixed-task bootstrap intervals cover none of those quantities or unknown simulator-error bias. In particular, the table is evidence for considering temporal distortion alongside attenuation and checking task success, not a taskwise tolerance bound or a recommendation of q.

The supplementary provenance record links the result report `MATCHED_TIME_RESULTS_20260911.md`, independent raw-row/trace/statistics audit `MATCHED_TIME_RESULT_VERIFICATION_20260911.md`, and the preserved `matched_time_formal_20260910_attempt1` run directory. The official 450-state artifacts remain separate.

F.7 Reversed weights on a second checkpoint (LIBERO-Spatial)

The second-suite check uses the same five-arm reversed-weight intervention on the fixed LIBERO-Spatial task set, but on a different OpenVLA-OFT checkpoint. There are ten fixed tasks, 20 newly generated paired resets per task, and 200 paired resets in total; the five arms yield 1,000 episodes in total (200 per arm). The checkpoint is `moojink/openvla-7b-oft-finetuned-libero-spatial`, revision `6d0231af0e48c5985f1ff86908f4674b84bc049b`. These rows are not pooled with the original LIBERO-10 600-pair study.

The recorded runtime differs in one important cap: it permits at most 220 policy actions per episode, versus 520 in the original LIBERO-10 campaign. Both runtimes use ten neutral startup steps, query/execute cadence 8 and $\Delta t = 0.05$ s. The first three selected pose commands pass raw; the six pose coordinates are then filtered while the current raw gripper command is retained. History persists across chunks and is reset between episodes; no additional clipping, terminal flush, or delay queue is added. Pairing fixes the task, reset state, initial observation/input, and first chunk across arms; closed-loop trajectories may diverge after the first differing intervention. The arm order rotates by $(\text{task} + \text{reset}) \bmod 5$.

Table 34 gives all five arm outcomes. The paired contrasts below are the only two primary tests for this suite, with Holm correction over that family of two.

The task-stratified paired bootstrap used 100,000 draws per contrast with seed 20261001, sharing task/reset units across arms. Pointwise intervals are descriptive for this fixed task set; the family intervals are per-contrast Bonferroni intervals calibrated at 97.5% over the two contrasts, not a claim of exact joint coverage.

Warning capture was available for all 1,000 rows, and no reset group triggered the warning sensitivity. The prescribed zero-trigger sensitivity therefore reproduces the pri-

Table 33: Second-suite Spatial runtime and state scope. All entries are part of the separate 200-pair, five-arm study.

Item	Recorded contract
Tasks and states	LIBERO-Spatial; 10 fixed tasks; 20 newly generated states per task; 200 paired resets; 5 arms; 1,000 episodes
Checkpoint	OpenVLA-OFT revision `6d0231af0e48c5985f1ff86908f4674b84bc049b`
Runtime	Query/execute 8; $\Delta t = 0.05$ s; maximum 220 policy actions
Action path	Pose six-vector filtered; current raw gripper retained; first three selected pose commands pass raw
History and pairing	History persists across chunks and is reset between episodes; no added clipping or queue; pre-intervention identity checks as stated above

Table 34: All five arms on the second-suite Spatial checkpoint. Each success count has denominator 200; the native arm is included for scope, not as a claim of native equivalence.

Condition	Success	Mean sent jerk	Native reduction (%)
Native	139/200	1233.640076	—
$q = .30$ recent-weighted	134/200	720.273027	41.61
$q = .30$ past-weighted	84/200	959.468030	22.22
$q = .45$ recent-weighted	130/200	569.807119	53.81
$q = .45$ past-weighted	90/200	720.424041	41.60

Table 35: Second-suite Spatial paired success tests. Differences are recent-weighted minus past-weighted percentage points; discordances are recent-only/past-only.

Contrast	Difference	Discordance	Exact p	Holm-2 p
$q = .30$	+25.0	50/0	1.7764×10^{-15}	3.5527×10^{-15}
$q = .45$	+20.0	41/1	1.9554×10^{-11}	1.9554×10^{-11}

Table 36: Second-suite Spatial paired bootstrap intervals, in percentage points. The family interval is the fixed-task Bonferroni 97.5% interval.

Contrast	Pointwise 95%	Family-2 97.5%
$q = .30$ recent − past	[20.0, 30.5]	[19, 31]
$q = .45$ recent − past	[15.0, 25.5]	[14, 26]

Table 37: Complete task-level success counts for the second-suite Spatial study. Each entry is success/20 newly generated resets; task identifiers follow the fixed ten-task ordering.

Task	Native	.30 recent	.30 past	.45 recent	.45 past
0	20/20	20/20	9/20	20/20	10/20
1	20/20	19/20	12/20	19/20	12/20
2	20/20	20/20	20/20	20/20	19/20
3	17/20	9/20	1/20	5/20	1/20
4	3/20	6/20	2/20	6/20	3/20
5	18/20	20/20	10/20	20/20	12/20
6	20/20	20/20	19/20	20/20	20/20
7	0/20	0/20	0/20	0/20	0/20
8	19/20	19/20	11/20	19/20	13/20
9	2/20	1/20	0/20	1/20	0/20

mary counts, exact and Holm-adjusted p-values, and intervals; it is secondary and does not replace the primary analysis. The recorded runtime-adapter source SHA256 is `5a95831472b2e62ba7e3223927df9e665c75a634d742980d8536bb4050c4d248`. The provenance merge consists of 970 original rows plus a disjoint 30-row continuation, with the recorded runtime and immutable development labels retained. The authoritative completion sidecar marks the continuation complete; the historical runner-core hash was not captured at execution time. These records support the same direction of the reversed-weight effect on this second suite-specific checkpoint, not warning-free physics, hardware benefit, training-seed stability, or policy-family generalization.

F.8 ARCHIVE-ONLY TEMPORAL DOSE-RESPONSE DIAGNOSTIC

This archive-only reanalysis asks whether the success differences within each operator bank are better organized by temporal placement or by common-input attenuation. The design group delay is the first moment of a normalized FIR, equivalently its DC group delay under the convention in Section 3; it is a command-step quantity, not physical latency. We discard the first three rows of each native raw pose prefix, demean its six channels, and average their summed power spectra over resets within a bank. Applying each filter's squared magnitude response to this common spectrum gives two energy-reduction descriptors: one unweighted and one weighted by $|1 - e^{-i\omega}|^6$. Thus each descriptor is a bank–prefix–arm constant, not a reset-specific covariate. Prefix lengths 32 and 64 use the actual rFFT grids of 29 and 61 samples, respectively. The eight-arm LIBERO-10 study (Section 5.5) and the LIBERO-10 and LIBERO-Spatial reversed-weight studies (Section 5.1) were fitted separately; rows were never pooled across banks.

The conditional-logit likelihood conditions on each reset's total successes across arms, eliminating its intercept. The linear predictor contains design delay and/or the energy descriptor, with no additional arm effects; the normalizer sums over all arm subsets with that same number of successes. Only resets with both successes and failures contribute: $129/450$, $203/600$, and $80/200$ in these three studies, respectively. The first likelihood-ratio column adds design delay to a model containing the descriptor; the second column tests adding the descriptor after delay. The displayed p values are exploratory χ^2_1 approximations, not a confirmatory causal test.

Table 38: Archive-only incremental-fit diagnostics. Larger likelihood-ratio statistics indicate a larger incremental fit improvement for the added term; the reverse p column tests adding the descriptor after design delay. The banks are the eight-condition OpenVLA-OFT study (Section 5.5) and the two reversed-weight studies (Section 5.1); L-10 denotes LIBERO-10.

Bank / prefix	Descriptor	LR(delay | descriptor)	p(descriptor | delay)
8-arm (L-10) / 32	pose energy	99.959	0.547
8-arm (L-10) / 32	jerk-weighted energy	74.072	0.441
8-arm (L-10) / 64	pose energy	99.782	0.551
8-arm (L-10) / 64	jerk-weighted energy	74.278	0.440
Reversed (L-10) / 32	pose energy	190.611	0.855
Reversed (L-10) / 32	jerk-weighted energy	179.158	0.949
Reversed (L-10) / 64	pose energy	191.437	0.849
Reversed (L-10) / 64	jerk-weighted energy	179.205	0.948
Reversed (Spatial) / 32	pose energy	131.832	0.714
Reversed (Spatial) / 32	jerk-weighted energy	125.240	0.768
Reversed (Spatial) / 64	pose energy	132.186	0.712
Reversed (Spatial) / 64	jerk-weighted energy	125.262	0.767

Table 39 gives prefix-64 arm descriptors and success counts; the reductions describe the common-input construction, not observed sent-command jerk. The delay-only bootstrap used 250 draws per bank (244, 243, and 243 valid for the eight-arm LIBERO-10, reversed-weight LIBERO-10, and reversed-weight LIBERO-Spatial studies, respectively). Their respective intervals are $[-1.330, -0.785]$, $[-1.242, -0.839]$, and $[-2.482, -1.395]$. Each draw resamples whole reset vectors across the bank, without stratifying by task, and keeps the descriptors fixed (seed 20260915). These are 2.5–97.5 percentile intervals in original coefficient units. Fits that failed to converge, returned nonfinite coefficients, or reached the standardized coefficient bounds of ± 20 were omitted, not imputed. This coefficient bootstrap does not calibrate the likelihood-ratio p values.

Across prefixes and descriptors, the delay term is the stronger incremental organizer of these fixed-bank success differences. The predictor values come from only eight, five, and five fixed arm settings, respectively. Reversed pairs have identical descriptors by construction, so this analysis summarizes the designed contrasts rather than providing independent validation on new operators. The construction does not determine which arm succeeds. This is descriptive of the archived operator sets: it does not establish a delay-only cause, eliminate descriptor effects, define a universal dose-response, or provide an automatic filter selector.

Table 39: Prefix-64 arm descriptors and success counts in each archived bank; energy reduction is relative to the native common-input pose signal. The banks are the eight-condition OpenVLA-OFT study (Section 5.5) and the two reversed-weight studies (Section 5.1); L-10 denotes LIBERO-10.

Bank	Arm	Design delay	Energy reduction / success
8-arm (L-10)	native	0.000	$0.000000;\ 416/450$
8-arm (L-10)	SW(4)	1.500	$0.046947;\ 387/450$
8-arm (L-10)	SG4	0.000	$-0.020222;\ 418/450$
8-arm (L-10)	EMA(0.55)	0.818	$0.045774;\ 407/450$
8-arm (L-10)	EMA(0.70)	0.429	$0.023905;\ 406/450$
8-arm (L-10)	EMA(0.85)	0.176	$0.010140;\ 422/450$
8-arm (L-10)	SG4 + delay 1	1.000	$-0.020222;\ 403/450$
8-arm (L-10)	SG4 + delay 2	2.000	$-0.020222;\ 363/450$
Reversed (L-10)	native	0.000	$0.000000;\ 553/600$
Reversed (L-10)	.30 recent	0.396	$0.020041;\ 541/600$
Reversed (L-10)	.30 past	2.604	$0.020041;\ 439/600$
Reversed (L-10)	.45 recent	0.647	$0.030017;\ 537/600$
Reversed (L-10)	.45 past	2.353	$0.030017;\ 457/600$
Reversed (Spatial)	native	0.000	$0.000000;\ 139/200$
Reversed (Spatial)	.30 recent	0.396	$0.020093;\ 134/200$
Reversed (Spatial)	.30 past	2.604	$0.020093;\ 84/200$
Reversed (Spatial)	.45 recent	0.647	$0.029824;\ 130/200$
Reversed (Spatial)	.45 past	2.353	$0.029824;\ 90/200$

G ADAPTIVE FILTERING AND TEMPORAL WEIGHTING

G.1 EVALUATION AND PARAMETER SELECTION

We compare One Euro with EMA at two smoothing levels, together with native, SW(4), and SG4 references. This separate evaluation uses the fixed `moojink/openvla-7b-oft-finetuned-libero-10` checkpoint (revision `95220f9a3421a7ff12d4218e73d09ade830fa9a3`) on ten LIBERO-10 tasks. Each arm uses the same 40 official state identifiers per task (0–4 and 15–49): 400 paired resets per arm and 2,800 episodes in total. Results are not pooled with the other OpenVLA-OFT studies. Calibration uses 100 archived native traces, states 5–14 per task from the eight-arm study, without success labels. Calibration and evaluation state identifiers are disjoint; this is not a new-task generalization test.

The runtime executes eight selected actions per policy query at $\Delta = 0.05\,\mathrm{s}$, with ten warm-up steps and a 520-action cap. The seven arms share the initial simulator state, observation and first raw action chunk, checked before intervention. Arm order rotates cyclically across reset groups. All filters act on six pose channels and pass the current raw gripper. Filter histories persist across chunks and reset between episodes. SW(4) averages the available prefix; SG4 passes through its first three commands. EMA and One Euro initialize with the first raw command. No extra clipping is added by the filters; downstream action conversion and controller clipping remain as in Appendix E.

One Euro (Casiez et al., 2012) uses six scalar instances of the authors' implementation, pinned to commit `32a58a93a6c479e430e349895807ec6f7ab8b3d2`, with fixed sample time. Each pose channel is transformed to $z_d = (u_d - c_d)/s_d$ using the checkpoint's `BOUNDS_Q99` statistics and transformed back after filtering. For this implementation, $d_t = (z_t - \widetilde{z}_{t-1})/\Delta$; the first derivative is initialized to zero. A low-pass derivative estimate $\widehat{d}_t$ with cutoff 1 Hz sets $f_t = f_{\min} + \beta|\widehat{d}_t|$ and $\alpha_t = (1 + (2\pi f_t \Delta)^{-1})^{-1}$. EMA uses a constant α in $\widetilde{u}_t = \alpha u_t + (1-\alpha)\widetilde{u}_{t-1}$.

We calibrate to jerk remaining at 60% and 40% of native, denoted *mild* and *strong*. The scalar calibration ratio uses the first 64 commands, with episode means followed by equal task means. It is not frequency-wise attenuation matching. A fixed nested-grid and bracket-refinement search uses $\alpha \in [0.05, 0.95]$, $f_{\min} \in [0.1, 10]\,\mathrm{Hz}$, and $\beta \in \{0, 0.01, 0.1, 1, 10\}$. At each target, the feasible positive-β representative with the lowest normalized command-reference RMSE on indices 3–63 is selected; $\beta = 0$ is a constant-filter diagnostic. One Euro therefore has a larger calibration search than EMA. The finite search evaluates 284 parameter tuples and does not establish an optimal filter. Earlier finite-grid failures led to this revised calibration procedure before the formal comparison; an earlier RMSE-superiority screen was removed and remains reported in the artifacts. The final selected EMA coefficients are 0.7014892578 and 0.5116455078; One Euro uses $\beta = 1$ with $f_{\min} =$ 6.5084348593 and 2.0169145547 Hz. All four calibration ratios are within 0.0001 of their targets. These parameters are fixed before evaluation.

G.2 CLOSED-LOOP OUTCOMES

Table 40 reports all seven conditions. Stronger One Euro smoothing reduces jerk further, while observed success falls from $367/400$ to $357/400$. EMA shows the same ten-success decrease between its two settings. These are observed strength contrasts; the primary paired tests compare One Euro with EMA at each target (Table 41).

The One-Euro/EMA sent-jerk ratios are 1.0114 (95% interval $[0.9915, 1.0321]$) and 0.9973 ($[0.9809, 1.0136]$). Neither success contrast establishes superiority or equivalence. For the predefined 5-percentage-point aggregate loss allowance, one-sided 98.75% bootstrap lower bounds for One Euro minus native are -4.50 and -7.75 points; the corresponding EMA-relative bounds are both -3.25. Thus only the mild One Euro setting meets both numerical screening criteria in this fixed task bank. This does not guarantee taskwise preservation: strong One Euro loses $6/40$ successes relative to native in each of tasks 8 and 9 (Table 42).

Table 40: Separate LIBERO-10 adaptive-filter evaluation: 400 paired resets per arm. Jerk is the mean sent pose-command jerk in the stored command units, using $\Delta = 0.05\,\mathrm{s}$; reductions use this study's native arm. Mild and strong refer to calibration targets of 60% and 40% jerk remaining.

Condition	Successes	SR (%)	Sent jerk	Reduction (%)
Native	375/400	93.75	573.980	0.00
SW(4)	340/400	85.00	164.278	71.38
SG4	379/400	94.75	361.784	36.97
EMA mild	369/400	92.25	341.534	40.50
One Euro mild	367/400	91.75	345.428	39.82
EMA strong	359/400	89.75	233.822	59.26
One Euro strong	357/400	89.25	233.190	59.37

Table 41: One Euro minus EMA on the same 400 paired resets. Discordances are One-Euro-only/EMA-only successes. Intervals are pointwise 95% task-stratified paired bootstrap intervals (100,000 draws, seed 20260919). Exact two-sided McNemar tests use Holm correction over the two comparisons.

Target	SR difference (pp)	95% interval (pp)	Discordances	Exact p	Holm p
Mild	−0.50	$[-2.75, 1.75]$	10/12	0.8318	1.0
Strong	−0.50	$[-3.00, 1.75]$	11/13	0.8388	1.0

Table 42: Taskwise successes out of 40 in the adaptive-filter evaluation. All conditions use the same state identifiers within each task.

Task	Native	SW(4)	SG4	EMA mild	Euro mild	EMA strong	Euro strong
0	36	34	36	35	34	33	35
1	40	39	39	39	40	40	39
2	39	32	39	39	38	38	37
3	38	39	40	39	39	39	40
4	40	35	38	38	39	39	38
5	36	36	36	36	36	36	36
6	35	24	35	34	33	31	31
7	37	40	40	39	38	39	39
8	36	29	37	34	32	28	30
9	38	32	39	36	38	36	32

G.3 RECONSTRUCTING TEMPORAL WEIGHTS

After evaluating success, we reconstruct each filter's temporal weights from the stored commands. For a fixed realized sequence of nonnegative coefficients, write $\widetilde{u}_t = \sum_{k=0}^{t} w_{t,k} u_{t-k}$, with $\sum_k w_{t,k} = 1$. The *mean command age* $A_t = \sum_k k w_{t,k}$ measures how far back the filter places its weight, in command steps. For EMA and One Euro, $A_t = (1-\alpha_t)(A_{t-1}+1)$ after initialization at $A_0 = 0$. For One Euro these are weights conditional on its realized coefficients; they are not an input Jacobian, an LTI group delay, or physical latency. SG4 has signed weights and is excluded from this nonnegative-age comparison.

We first replay each of the 400 native raw-command prefixes through all operators, then separately replay the four EMA/One Euro conditions on their own raw inputs ($1{,}600$ sequences). All inputs contain the first 64 commands; summary times are 3–63, with 3–31 as a fixed sensitivity window. Age summaries average equally over six channels, resets within each task, and tasks. The jerk ratio compares mean raw-unit pose jerk between output and input, using third differences ending at the scoring indices, including the preceding commands needed for each difference. The reconstructed own-input pose outputs match the stored sent commands with zero numerical error in all $1{,}600$ sequences, and gripper passthrough matches exactly. Adaptive cutoff and derivative states are reconstructed offline, rather than taken from online telemetry.

Table 43: Temporal weights reconstructed from the first 64 commands, summarized at indices 3–63. The common-input columns use the same native raw sequence; own-input age uses each arm's different closed-loop input. The common-input jerk ratio is a scalar finite-record quantity, not a matched frequency response. These replays produce no new task-success outcomes.

Condition	Common-input mean age (steps)	Jerk ratio	Own-input mean age (steps)
EMA mild	0.4253	0.59816	0.4253
One Euro mild	0.4574	0.59760	0.4576
EMA strong	0.9509	0.39834	0.9509
One Euro strong	1.2426	0.39670	1.2464

Similar scalar jerk ratios coexist with different temporal weights (Table 43); the age ordering also holds in the shorter window and on own inputs. The stronger setting of each method has larger mean age, lower jerk, and 2.5 percentage points lower observed success. Because smoothing strength and temporal weights change together, this comparison does not isolate their effects on success. The inconclusive One-Euro/EMA contrasts also do not establish a success difference attributable to their different mean ages. This selected adaptive filter therefore provides a consistency check for examining temporal weighting alongside smoothing strength and judging the trade-off by closed-loop success.

Record scope. Temporal summaries cover the initial command prefixes, not contact-specific events or complete episodes; own-input histories depend on feedback. They are not success predictors or a rule for selecting a filter. All 254 unsuccessful episodes reach the action cap; the records do not distinguish manipulation failure causes. Empty Python warning arrays do not establish the absence of simulator-native warnings. The preserved manifests contain inherited engineering labels; reported outcomes and reconstructed outputs were checked against the completed evaluation records. Full frozen configurations, calibration history and verification records accompany the artifacts.