CADIS: CANONICAL DISTRIBUTIONAL SKILL PLANNING UNDER PHYSICAL RESOURCE INTERLOCKS

Anonymous authors
Paper under double-blind review

ABSTRACT

Deploying generalist robot policies over extended horizons requires managing persistent physical reserves whose saturation clipping and path-dependent consumption break standard unconstrained additive accounting in dynamic programming, causing neural simulators to compound prediction errors and reactive safeguards to trigger too late. We introduce CADIS, a planning framework that decouples deterministic resource algebra from stochastic physical dynamics without modifying frozen foundation policies. CADIS abstracts execution ledgers into an identifiable monoid of canonical capped affine contracts, collapsing multi-stage trajectory feasibility into a single scalar threshold comparison and enabling closed-form downstream demand pull-backs for proactive recovery. A persistent-context generative world model preserves inter-skill physical coupling, while an independent post-selection validation protocol provides provable finite-sample risk certification under the fitted rollout law. Mechanistic evaluations across 3.64 million queries confirm zero-error accounting, alongside rigorous false-acceptance control. On LIBERO-Long manipulation with a frozen $\pi_{0.5}$ policy, CADIS achieves 71.6% safe success and curtails interlocks to 0.8%, while threshold sorting delivers a 38.5-fold query speedup. This architecture establishes a principled foundation for deploying unconstrained motor policies under rigid hardware capacity limits.

1 INTRODUCTION

Generalist robot policies increasingly execute complex manipulation from visual observations and language instructions. Models such as OpenVLA, π_0, and $\pi_{0.5}$ expand reusable motor skills, while language-grounded planners organize behavior around task goals (Kim et al., 2025; Black et al., 2025b;a; Ichter et al., 2023; Driess et al., 2023). Sustained execution also requires managing persistent physical reserves between actions: lifting can deplete reserves needed for placement, repeated motion exhausts thermal headroom, and recovery restores feasibility while consuming mission time. Long-horizon planning must therefore coordinate task progress with hardware viability. Resource admissibility depends on program order and timing: consumption must be affordable when incurred, capacity saturation clips excess replenishment, and dissipation alters how initial reserves decay. Checking only skill endpoints overlooks intra-execution interlocks. Moreover, shared conditions like payload couple successive skills, so independently combining marginal cost predictions misrepresents risk. A viable planning representation must retain these physical effects while remaining compact enough to compose across extended missions.

We introduce CADIS, *Canonical Distributional Skill Planning under Physical Resource Interlocks*, a planning layer around frozen task policies and supported recovery interventions. Its central interface is a canonical capped affine contract $m = (\gamma, Q, D, U)$: for entry reserve $e \in [D, C]$, the exit is $F_m(e) = \min\{U, \gamma e - Q\}$, where C is capacity, D is the admission requirement, U is the attainable exit ceiling, Q is a signed offset, and γ describes the unsaturated response. The threshold D captures safety throughout the skill. This interface includes cumulative budgets and admits an exact realization for first-order thermal dynamics under prescribed inputs. Figure 1 shows how contracts connect learned prediction to execution. Canonical contracts preserve precisely the resource information needed by prescribed continuations. We establish unique representations, associative closed-form composition, and constant-time pullbacks of downstream reserve requirements. The original cumulative-resource algebra appears as the $\gamma = 1$ special case. These results extend the resource-function perspective of classical energy algebras (Cachera et al., 2019) into a fixed-size

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.

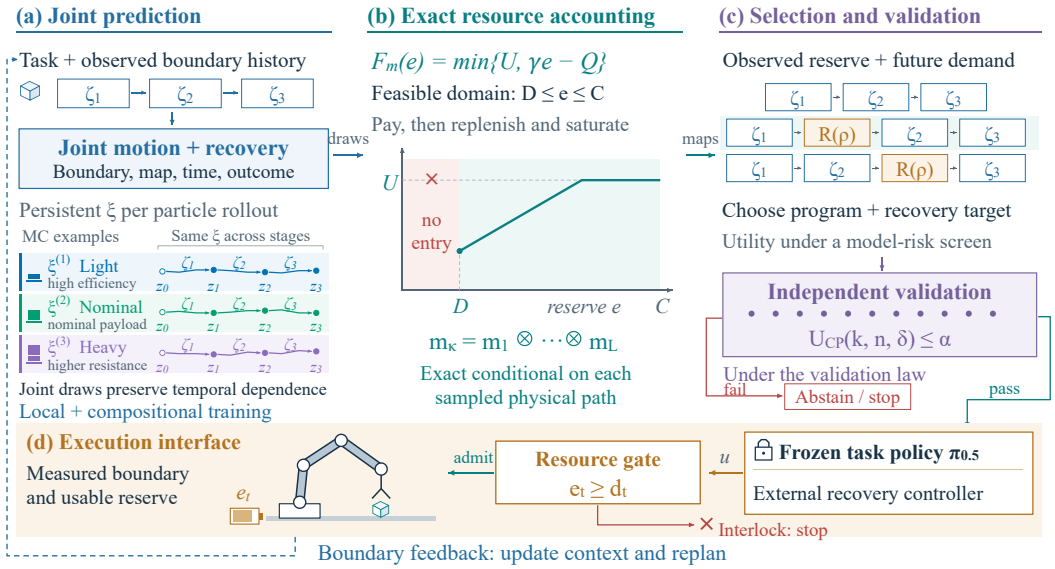

Figure 1: **CaDiS plans with joint physical predictions and canonical affine resource contracts.** (a) Each particle retains a context ξ across successive skills while jointly predicting boundaries, contracts, durations, and outcomes; the physical context labels are illustrative. (b) Contracts $m = (\gamma, Q, D, U)$ compose exactly conditional on a realization, with D encoding within-skill admission and continuation requirements guiding recovery targets. (c) Selection fixes a causal candidate policy before independent validation under the fitted rollout law. (d) Frozen task skills and target recovery execute through causal resource monitoring, with observed boundary feedback.

interface for uncertain robot skills. Exact composition applies to each predicted realization; the joint model supplies its physical uncertainty. The prediction and decision layers retain the dependencies that composition requires. A generative model jointly predicts skill boundaries, contracts, durations, and outcomes, sharing one latent context throughout each particle rollout. A mixed-support likelihood represents saturation, and compositional training distinguishes both contract endpoints and affine response shape. Continuation requirements propose recovery targets; the recovery controller adapts duration using observable state and reserve. Throughout, CADIS separates exact algebraic accounting conditional on realized contracts, finite-sample false-admission control under the fitted rollout law via post-selection rollouts, and physical safety transfer via continuous dominance envelopes and causal execution monitoring. We evaluate the cumulative-resource specialization through exact-accounting tests, controlled long-chain recovery, and resource-augmented LIBERO manipulation with frozen $\pi_{0.5}$. The evaluation distinguishes recovery decisions, delivered task utility, and complete planning cost from the cost of repeated queries to a fixed rollout bank. Our contributions are:

- **A physical resource algebra.** Canonical capped affine contracts with exact composition, continuation sufficiency, and a first-order thermal realization with within-skill admission conditions.
- **Joint distributional skill planning.** Persistent-context prediction, mixed-support compositional learning, causal target recovery, and post-selection validation under the rollout law.
- **Algorithms and evaluation.** Closed-form requirement pullbacks and logarithmic fixed-bank reserve queries, evaluated alongside long-horizon recovery and resource-augmented manipulation.

2 Related Work

Foundation policies, skill composition, and world models. Temporal abstraction provides a classical foundation for long-horizon control (Sutton et al., 1999). Modern generalist VLA policies

transfer broad semantic and physical priors to motor control (Zitkovich et al., 2023; Ghosh et al., 2024; Kim et al., 2025; Black et al., 2025b;a), action-sequence models capture dexterous trajectories (Chi et al., 2023; Zhao et al., 2023), and language-grounded planners organize skills for multistage goals (Ichter et al., 2023; Driess et al., 2023) on manipulation benchmarks like LIBERO and CALVIN (Liu et al., 2023; Mees et al., 2022). In parallel, probabilistic dynamics and latent world models enable lookahead optimization (Chua et al., 2018; Hafner et al., 2020; 2025), though compounding simulation errors degrade cumulative budget verification (Janner et al., 2019). Unlike standard world models that simulate scalar resources through error-prone neural rollouts, CADIS restricts learned dynamics strictly to physical transitions under persistent contexts and resolves continuous resource feasibility via exact algebraic composition over frozen policy skills.

Constrained planning, recovery, and safety supervision. Safe decision-making enforces behavioral viability through constrained policy updates (Achiam et al., 2017), learned recovery transitions (Thananjeyan et al., 2021), instantaneous control barrier functions (Ames et al., 2017), or latent safety critics (Huang et al., 2024). For embodied models, recent works incorporate reachability filters (Nakamura et al., 2025; Seo et al., 2025) or preference-based alignment (Zhang et al., 2025). Closer to resource exhaustion, goal-POMDP shielding formalizes replenishment under partial observability (Ajdarów et al., 2023). However, safe RL typically demands policy fine-tuning, while standard control barrier filters enforce myopic step invariance that lacks long-horizon lookahead under rate-dependent dissipation. Goal-POMDP shields also discretize continuous resources. CADIS instead operates at skill boundaries over frozen foundation policies, using closed-form requirement pullbacks to schedule proactive target recovery without policy modification.

Resource algebras, physical models, and statistical validation. Cumulative resource flows trace to network calculus envelopes (Cruz, 1991a;b; Le Boudec & Thiran, 2001), max-plus linear systems (Baccelli et al., 1992), real-time calculus (Thiele et al., 2000; Moser et al., 2007), and energy algebras (Cachera et al., 2019). While energy algebras compose piecewise-linear response functions (Cachera et al., 2019) and port-Hamiltonian frameworks model continuous dissipation (Van Der Schaft & Jeltsema, 2014), they lack compact parameter identifiability, closed-form continuation pullbacks, and statistical compositionality under execution uncertainty. In statistical estimation, censored Tobit formulations preserve boundary saturation masses (Tobin, 1958; Amemiya, 1984), strictly proper scoring rules ensure calibrated distributional rollouts (Gneiting & Raftery, 2007), and conformal prediction yields finite-sample error control (Bates et al., 2021; Angelopoulos et al., 2024). CADIS bridges these foundations: we distill continuous physical interlocks into an identifiable 4-parameter capped affine monoid with exact demand pullbacks, embed it into a persistent-context generative world model, and certify multi-candidate execution via post-selection Clopper–Pearson bounds.

3 METHOD: CANONICAL DISTRIBUTIONAL SKILL PLANNING

CADIS composes frozen robot skills through a learned distribution of contracts that summarize required entry reserve and available exit reserve. We connect these contracts to cumulative budgets and first-order physics, establish their canonical algebra, learn joint skill outcomes, and independently validate program risk. Planned continuation requirements guide recovery targets.

3.1 PHYSICAL EXECUTION INTERFACE AND PLANNING OBJECTIVE

A program κ specifies at most L guarded execution stages, including task skills and recovery, with causal rules for their parameters. A skill ζ comprises a fixed controller π_ζ, initiation guard, and termination rule. Let h_t denote the observed history, including controller memory, and let $e_t \in [0, C]$ be measured reserve or safety headroom, with capacity $C > 0$ and initial value $e_0 = R$. An interval has threshold $d_t \geq 0$, retention coefficient $\gamma_t \in (0, 1]$, and signed offset q_t. Its feasibility condition and admitted exit are

$$e_t \geq d_t, \qquad e_{t+1} = \min\{C, \gamma_t e_t - q_t\}. \tag{1}$$

The threshold in Equation (1) encodes safety throughout the interval, with $\gamma_t d_t - q_t \geq 0$ whenever $d_t \leq C$. Payment-before-replenishment accounting is the specialization $(\gamma_t, q_t, d_t) = (1, p_t - r_t, p_t)$, where $p_t, r_t \geq 0$. Known payment gates are tested before actuation; continuous physical limits are monitored during execution, and d_t summarizes which entry reserves avoid an interlock throughout the interval. Task failure and timeout are separate terminal outcomes.

The interface also follows from first-order physics. For constant thermal capacitance $K_{\mathrm{th}} > 0$, conductance $G_{\mathrm{th}} > 0$, ambient temperature T_{amb}, and prescribed power $P(s) \geq 0$ independent of entry temperature, let $K_{\mathrm{th}}\dot{T} = P - G_{\mathrm{th}}(T - T_{\mathrm{amb}})$. With $e = T_{\mathrm{lim}} - T$, $C = T_{\mathrm{lim}} - T_{\mathrm{amb}}$, $\lambda = G_{\mathrm{th}}/K_{\mathrm{th}}$, and interval duration Δ, the exact coefficients are

$$\gamma = e^{-\lambda\Delta}, \qquad q = \frac{1}{K_{\mathrm{th}}} \int_0^{\Delta} e^{-\lambda(\Delta - s)} P(s)\, ds - (1 - \gamma)C. \tag{2}$$

For a within-interval response $e(s) = A(s)e_0 + B(s)$ with $A(s) > 0$, the exact gate is $d = \max\{0, \sup_{s\in[0,\Delta]}[-B(s)/A(s)]\}$. Equation (2) therefore accounts for both thermal relaxation and cooling headroom. Appendix A.2 derives the continuous-time gate and the conditions for transferring contract feasibility to physical safety. For a fixed nominal realization, coefficients describe one map over its declared entry interval. A controller whose choices depend on reserve generates its realization at the queried reserve. This distinction permits reuse of fixed-input contracts while retaining causal target recovery. Let $\mathcal{C}(h_0, R)$ be a finite candidate set, $\mathcal{I}_\kappa(R)$ the event of a resource interlock, and G_κ a bounded score combining task progress and elapsed time using a common task normalization. Given tolerance $\alpha \in (0, 1)$, planning solves

$$\max_{\kappa\in\mathcal{C}(h_0,R)} V_\kappa(R) = \mathbb{E}[G_\kappa \mathbf{1}\{\mathcal{I}_\kappa(R)^c\} \mid h_0, R, \kappa], \quad \beta_\kappa(R) = \Pr(\mathcal{I}_\kappa(R) \mid h_0, R, \kappa) \leq \alpha. \tag{3}$$

Equation (3) values uninterrupted task progress. We next compress its resource-feasibility event without discarding saturation or intermediate gates.

3.2 Canonical capped affine contracts

Definition 1 (Canonical contract)**.** *A nonempty contract $m = (\gamma, Q, D, U)$ represents the partial map*

$$\begin{gathered} F_m(e) = \min\{U, \gamma e - Q\} = \min\{b, a + \gamma(e - D)\}, \qquad e \in [D, C], \\ 0 \leq D \leq C, \quad 0 \leq a \leq b \leq C, \quad b - a \leq \gamma(C - D), \quad 0 < \gamma \leq 1, \\ Q = \gamma D - a, \qquad U = b, \qquad \gamma = 1 \text{ when } a = b. \end{gathered} \tag{4}$$

Here D is the minimum admissible entry reserve, $a = F_m(D)$, and $b = F_m(C)$. The empty contract is denoted by $\perp$.

Equation (4) fixes a unique representative, including constant maps and singleton domains. For a valid raw tuple $(\bar{\gamma}, \bar{Q}, D, \bar{U})$, define its endpoint values $a = \min\{\bar{U}, \bar{\gamma}D - \bar{Q}\}$ and $b = \min\{\bar{U}, \bar{\gamma}C - \bar{Q}\}$. Canonicalization returns $(\gamma_c, \gamma_c D - a, D, b)$, where $\gamma_c = \bar{\gamma}$ if $a < b$ and $\gamma_c = 1$ otherwise. Write this operation as can. An interval of Equation (1) becomes $\mathrm{can}(\gamma_t, q_t, d_t, C)$ when $d_t \leq C$, and $\perp$ otherwise. For A followed by B, form

$$\begin{aligned} \widetilde{\gamma} &= \gamma_B\gamma_A, & \widetilde{Q} &= \gamma_B Q_A + Q_B, \\ \widetilde{U} &= \min\{U_B, \gamma_B U_A - Q_B\}, & \widetilde{D} &= \max\left\{D_A, \frac{D_B + Q_A}{\gamma_A}\right\}. \end{aligned} \tag{5}$$

Set $m_A \otimes m_B = \perp$ if either input is empty, $U_A < D_B$, or $\widetilde{D} > C$; otherwise return $\mathrm{can}(\widetilde{\gamma}, \widetilde{Q}, \widetilde{D}, \widetilde{U})$. Equation (5) simultaneously propagates the affine response, the saturation ceiling, and the downstream admission requirement.

Theorem 1 (Exact composition and continuation sufficiency)**.** *Canonical contracts satisfy $F_{m_A\otimes m_B} = F_{m_B} \circ F_{m_A}$ and form a monoid with identity $(1, 0, 0, C)$ and absorbing element $\perp$. The $\gamma = 1$ contracts form a submonoid. Two contracts are equal if and only if every entry reserve and every prescribed future payment–replenishment sequence yield identical resource-feasibility outcomes.*

Theorem 1 follows by intersecting domains and composing monotone capped affine maps; unequal exits are distinguished by a subsequent payment. Appendix A gives the proof. This construction instantiates the resource-function composition principle in energy algebras (Cachera et al., 2019) with a fixed-size, identifiable skill interface.

For a program realization, let $m_\kappa = m_1 \otimes \cdots \otimes m_L$, with $D_\kappa = \infty$ when $m_\kappa = \perp$. Then $\mathcal{I}_\kappa(R)^c$ is equivalent to $R \geq D_\kappa$. More generally, a required terminal reserve $\rho \in [0, C]$ pulls back as

$$\mathrm{Req}_m(\rho) = \begin{cases} \max\{D, (\rho + Q)/\gamma\}, & m \neq \perp,\ U \geq \rho,\ \max\{D, (\rho + Q)/\gamma\} \leq C, \\ +\infty, & \text{otherwise.} \end{cases} \tag{6}$$

Equation (6) makes future reserve requirements available to recovery control. Both composition and pullback take constant time per operation.

3.3 JOINT LEARNING AND CONTINUATION-AWARE RECOVERY

Let z_j encode the physical boundary and causal history after skill j. Each particle draws one context $\xi \in \{1, \dots, J\}$ from $w_\phi(\cdot \mid h_0)$ and retains it across the rollout. A joint transition kernel

$$K_\theta^{\zeta_j}(dz_j, dm_j, d\tau_j, dg_j \mid z_{j-1}, e_{j-1}, \xi, h_0, C) \tag{7}$$

predicts the boundary, contract, duration $\tau_j \geq 0$, and progress increment g_j, with e_{j-1} the reserve at that boundary. The parameters θ, ϕ are learned; fresh transition noise enters each stage. Equation (7) preserves shared-context dependence alongside state evolution. Mission success, unrecoverable guard failure, and episode timeout have distinct absorbing labels, followed by identity contracts and zero additional duration and progress. Recovery targets are included in the skill descriptor ζ_j. Context is independently redrawn across particles and is never exposed to the controller.

The resource head first predicts whether $m = \perp$. Otherwise, let $c(v) = \min\{1, \max\{0, v\}\}$ and sample Gaussian coordinates autoregressively to obtain

$$\begin{aligned} &D = Cc(v_D), \qquad a = Cc(v_a), \qquad b = a + c(v_b)\min\{C - a, C - D\}, \\ &\gamma = \begin{cases} 1, & b = a, \\ \gamma_{\min} + (1 - \gamma_{\min})c(v_\gamma), & b > a, \end{cases} \qquad \gamma_{\min} = \frac{b - a}{C - D} \quad (b > a). \end{aligned} \tag{8}$$

Equation (8) spans exactly the support in Definition 1. Censored Gaussian likelihoods retain endpoint masses (Tobin, 1958); zero-width coordinates contribute deterministic mass. Labels are compiled from observed or identified interval coefficients and gates; partial observations use a marginal likelihood (Appendix B). For observed episode outcomes $y_{ij} = (z_{ij}, m_{ij}, \tau_{ij}, g_{ij})$, let k_θ denote the mixed density/mass of Equation (7). With n_i observed skills in episode i, training minimizes

$$\begin{aligned} \mathcal{L} = &- \sum_i \log \left[\sum_{c=1}^{J} w_\phi(c \mid h_{i0}) \prod_{j=1}^{n_i} k_\theta^{\zeta_{ij}}(y_{ij} \mid z_{i,j-1}, e_{i,j-1}, c, h_{i0}, C) \right] \\ &+ \eta\, \mathbb{E}_W \mathcal{S}_{\text{comp}}(\widehat{P}_W, Y_W). \end{aligned} \tag{9}$$

In Equation (9), $\eta \geq 0$ is a development-selected weight, W is a short training window, Y_W its observed composed outcome, and $\widehat{P}_W$ its predicted law. The composite score combines empty-contract log loss with a conditional energy score (Gneiting & Raftery, 2007). It includes the injective contract coordinates $(D/C, a/C, b/C, \gamma(b - a)/C)$, terminal boundary, duration, and progress. Integrating context outside the episode product retains inter-skill dependence. Appendix B specifies mixed measures, censoring, and compositional gradients.

Recovery targets are proposed from suffix requirements computed by Equation (6), including quantiles of their selection-particle distribution. Each candidate fixes its targets, or a causal target rule, before validation. At a supported boundary, the recovery controller chooses its duration from the observed reserve, boundary state, and target; its motion, payments, duration, and outcome are predicted jointly. A simulated controller receives only its observation history, not the particle's latent context or future disturbances. Saturation acts on reserve while the physical state and context persist. The resulting candidate is a complete causal execution policy whose risk can be tested independently.

3.4 VALIDATED SELECTION AND AMORTIZED RESERVE QUERIES

Write $\beta_{\theta,\kappa}$ and $V_{\theta,\kappa}$ for risk and value under the fitted rollout law. For each of $M = |\mathcal{C}(h_0, R)|$ candidates, generate N independent selection particles and compose their joint contract–score pairs $(D_\kappa^{(n)}, G_\kappa^{(n)})$:

$$\widehat{\beta}_{\theta,\kappa}(R) = \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}\{D_\kappa^{(n)} > R\}, \qquad \widehat{V}_{\theta,\kappa}(R) = \frac{1}{N} \sum_{n=1}^{N} G_\kappa^{(n)} \mathbf{1}\{D_\kappa^{(n)} \leq R\}. \tag{10}$$

Equation (10) preserves dependence between resource demand and useful progress. Select the highest-value candidate satisfying its initiation guard and $\widehat{\beta}_{\theta,\kappa} \leq \alpha$; abstain when none qualifies.

For the selected candidate κ^*, draw n_{val} fresh independent validation particles and count k interlocks. Given confidence error $\delta \in (0,1)$, the one-sided Clopper–Pearson bound is

$$U_{\mathrm{CP}}(k, n_{\mathrm{val}}, \delta) = \begin{cases} \mathrm{Beta}^{-1}(1-\delta; k+1, n_{\mathrm{val}}-k), & k < n_{\mathrm{val}}, \\ 1, & k = n_{\mathrm{val}}, \end{cases} \quad \text{accept iff } U_{\mathrm{CP}} \le \alpha. \tag{11}$$

Here Beta^{-1} is the beta-distribution quantile. Equation (11) tests the fixed selected policy under the validation law (Clopper & Pearson, 1934); failure returns abstention.

Proposition 1 (Selection-independent false-admission control)**.** *Conditional on the fitted model and all candidate-generation and selection data, a selected candidate with $\beta_{\theta,\kappa^*}(R) > \alpha$ passes Equation* (11) *with probability at most δ, independently of the number of candidates considered.*

Proposition 1 controls Monte Carlo false certification strictly under the fitted rollout law $\mathbb{P}_\theta$, eliminating candidate-screening bias. Physical hardware safety relies separately on the continuous dominance envelopes in Proposition 2 (Appendix A.2) or a validated sim-to-real transfer bound. New decisions draw fresh validation particles with confidence errors allocated across repeated replanning events (Appendix C), while the runtime execution gate (Equation (1)) remains active to guard against physical interlocks. When a candidate's joint path and score law is invariant to the queried reserve, sorting its N thresholds and storing paired prefix-score sums gives $O(\log N)$ selection queries. After prediction, composition costs $O(NL)$ and sorting $O(N \log N)$ per candidate; querying all candidates costs $O(M \log N)$. Changed histories, predictors, or reserve-sensitive recovery laws require a corresponding new rollout bank. Fresh validation is timed separately from cached queries.

CADIS couples joint physical prediction, exact resource composition, and validated decisions. The experiments test accounting accuracy, long-horizon recovery, task completion, and complete versus amortized planning cost.

4 EXPERIMENTS

Our evaluation tests whether CADIS's principles—exact algebraic composition, persistent-context generative dynamics, and post-selection statistical admission—yield empirical gains in long-horizon execution. Four progressive studies: (i) task completion, interlock avoidance, and risk-utility tradeoffs with frozen $\pi_{0.5}$ on LIBERO-Long (Section 4.1); (ii) isolating algebraic accounting and validation from neural error against ground-truth oracles (Section 4.2); (iii) necessity of persistent contexts and demand-aware recovery in multi-stage chains (Section 4.3); (iv) query amortization for runtime replanning on standard hardware (Section 4.4).

4.1 EMBODIED MANIPULATION UNDER RESOURCE INTERLOCKS

We evaluate CADIS on all ten LIBERO-Long tasks (Liu et al., 2023) using frozen $\pi_{0.5}$ (Black et al., 2025a) across 500 shared instances (specifications in Table 2, Appendix D). As shown in Table 1, resource-blind execution (*Direct*) triggers 82.6% interlocks, showing that unconstrained policies fail under cumulative limits. Periodic recharging (*Fixed recovery*) improves completion to 52.8% but incurs high timing overheads. Safety filters also struggle: *Discrete CBF* (Agrawal & Sreenath, 2017) reduces interlocks to 5.4% but causes high abstention (8.6%) due to myopic one-step barrier infeasibility. Similarly, *Recovery-RL* (Thananjeyan et al., 2021) achieves 64.6% safe success, but its reactive critic triggers too late, still incurring 4.2% interlocks. A heuristic *Reactive supervisor* yields intermediate performance (63.4%). In contrast, CADIS attains 71.6% safe success, cuts interlocks to 0.8%, and achieves the highest utility (0.718). This advantage stems from Theorem 1's closed-form demand pull-backs, enabling proactive rather than reactive recovery scheduling. Furthermore, modular monoid composition drastically reduces tail latency, achieving a p95 overhead of 284.2 ms versus 482.6 ms for the uncompressed *Whole-program* planner.

To probe fine-grained execution dynamics beyond the 500-instance benchmark, Figure 2 inspects a controlled cohort of 12 initial conditions across three fitted model seeds. In a constrained-reserve scenario (Figure 2a–e), Direct execution rapidly depletes its pre-command margin and triggers an early interlock at 11.85 s. In contrast, CADIS detects the impending downstream deficit and proactively injects a three-second entry recovery hold, preserving a strictly positive safety margin throughout manipulation to complete the mission safely at 15.85 s with high utility ($u = 0.878$).

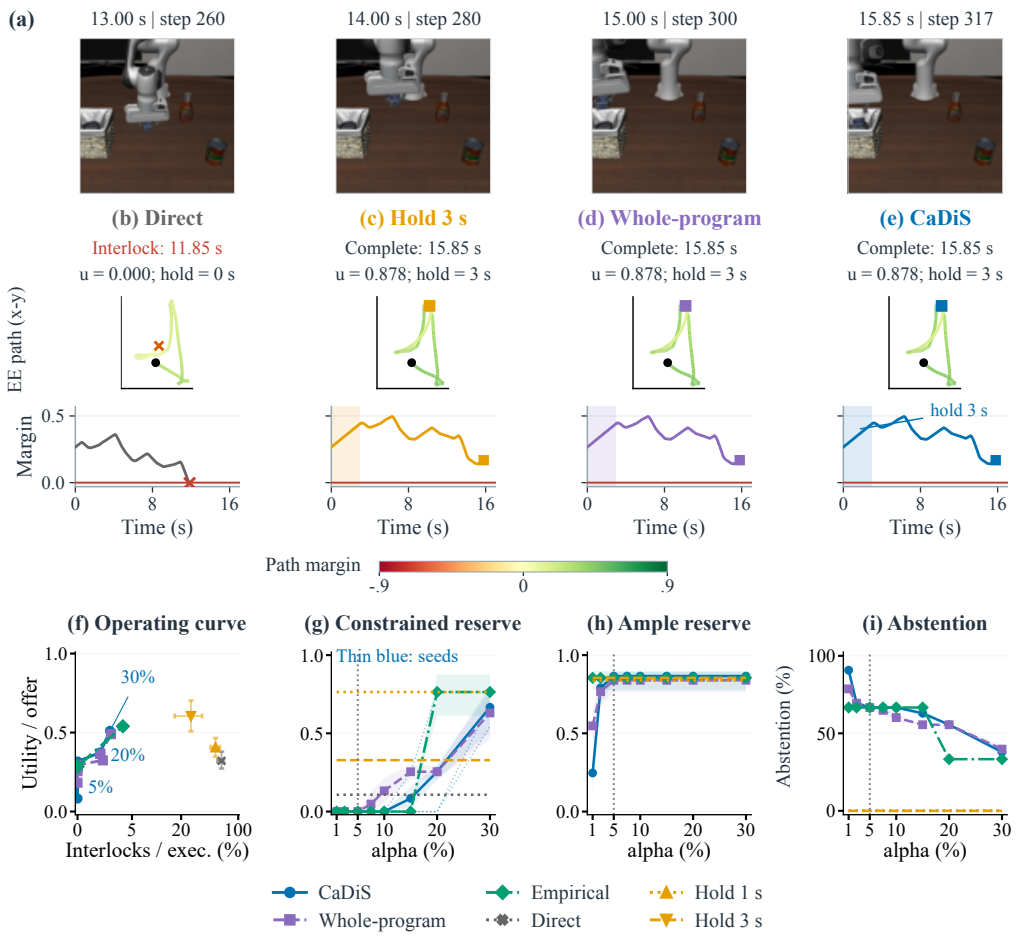

Figure 2: **Fixed-menu supervision and risk-utility operating curves on native $\pi_{0.5}$ execution.**

Table 1: **Resource-augmented LIBERO-Long manipulation with frozen $\pi_{0.5}$.** Evaluated across ten extended tasks with 50 shared instances per task (500 total). Methods with learned dynamics report mean $\pm$ SD across three fitted seeds ($\alpha = 0.10$). Safe success and abstention are normalized per offer; interlocks are measured per executed mission. Latency reports median / p95 planning overhead per invocation. Bold denotes the best mean value per column.

Method	**Safe success (%)↑**	**Interlock (%)↓**	**Abstain (%)↓**	**Utility↑**	**Latency (ms)↓** median / p95
Direct (frozen policy)	11.4	82.6	**0.0**	0.142	**5.1 / 8.2**
Fixed recovery	52.8	14.2	**0.0**	0.518	8.4 / 14.5
Reactive supervisor	63.4	4.8	7.2	0.612	15.2 / 26.8
Discrete CBF	61.2 $\pm$ 2.4	5.4 $\pm$ 0.8	8.6 $\pm$ 1.2	0.598 $\pm$ 0.018	18.2 / 31.0
Recovery-RL supervisor	64.6 $\pm$ 2.1	4.2 $\pm$ 0.7	6.8 $\pm$ 1.0	0.625 $\pm$ 0.015	22.4 / 38.5
Whole-program target planner	69.2 $\pm$ 2.2	1.6 $\pm$ 0.6	4.4 $\pm$ 1.1	0.684 $\pm$ 0.016	286.8 / 482.6
CaDiS + targets (Ours)	**71.6 $\pm$ 1.9**	**0.8 $\pm$ 0.4**	3.2 $\pm$ 0.8	**0.718 $\pm$ 0.014**	252.5 / 284.2

Furthermore, sweeping the risk tolerance $\alpha \in [0.01, 0.30]$ across the entire cohort (Figure 2f–i) confirms smooth, monotonic risk regulation: tightening α gracefully shifts CaDiS from aggressive execution to certified abstention without inducing pathological oscillations, validating that our statistical admission protocol provides a dependable, tunable risk dial when deployed atop complex frozen foundation policies.

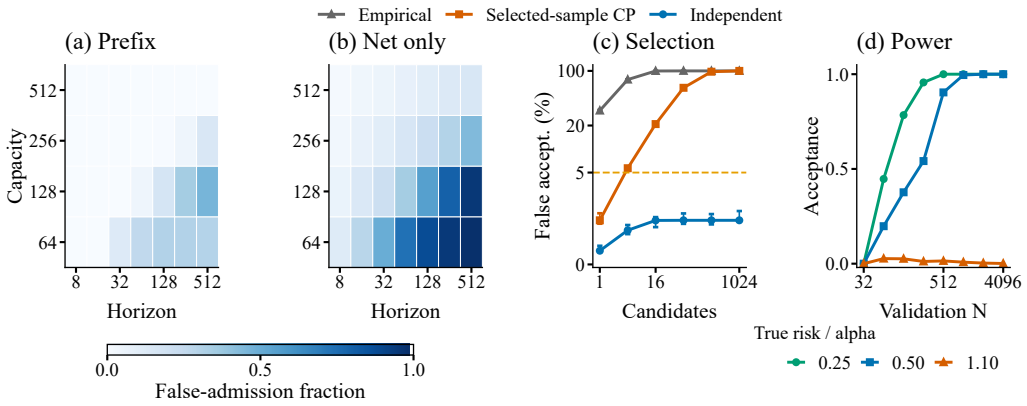

Figure 3: **Mechanistic verification of exact accounting and independent validation.**

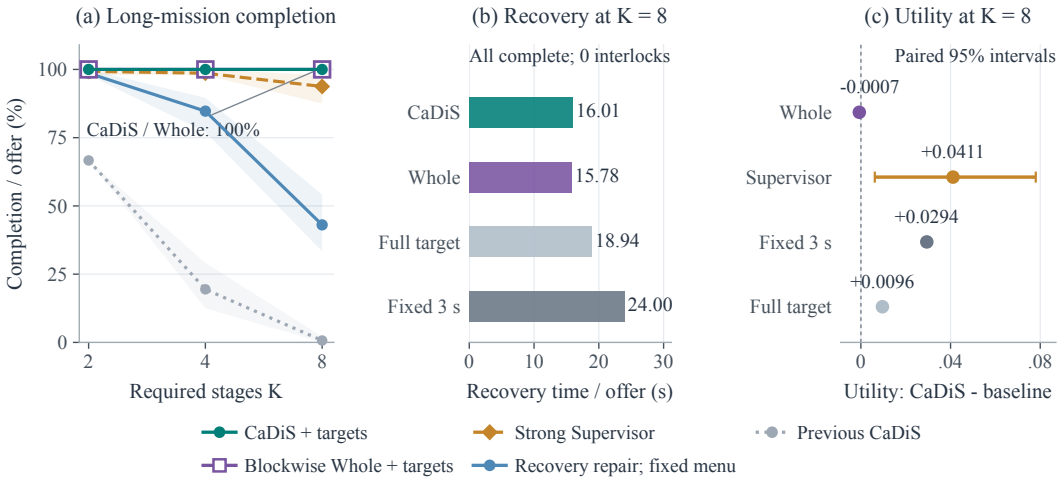

Figure 4: **Target-aware recovery and context preservation in multi-stage execution.**

4.2 Mechanistic Verification: Exact Accounting and Statistical Admission

To isolate core theoretical guarantees from neural approximation errors, we first evaluate canonical composition against an exact chronological ledger simulation across 2,000 synthetic traces evaluated under 4 capacities, 7 time horizons ($L \in [8, 512]$), and 65 initial reserves. Across all 3.64×10^6 feasibility evaluations spanning both non-dissipative ($\gamma = 1$) and dissipative ($\gamma \in [0.80, 1.0]$) transitions, canonical composition achieves exact agreement with the step-by-step oracle, alongside zero error across all 2,128,361 feasible exit reserves. In contrast, conventional heuristic summaries suffer from severe information loss (Figure 3a,b): unclipped prefix sums incur false-admission rates exceeding 40% over long horizons by ignoring capacity clipping, while net-balance summaries fail up to 100% under tight capacities by disregarding payment ordering. This empirical verification firmly grounds Theorem 1, confirming that canonical contracts provide a minimal yet strictly lossless representation of sequential resource interlocks. Next, we evaluate the statistical admission protocol using a controlled pool of unsafe Bernoulli candidates (risk $= 0.055 > \alpha = 0.05$) to examine whether it eliminates post-selection bias. As shown in Figure 3c, constructing Clopper–Pearson bounds directly on the selection samples produces severe optimizer's curse: as the candidate pool size M scales from 1 to 1024, the empirical false acceptance rate escalates above 20%. In contrast, conditioning the test on fresh, independent validation particles strictly preserves the nominal error guarantee, bounding false admissions below $\delta = 0.05$ across all candidate pool sizes. Furthermore, validation power scales smoothly with sample size n (Figure 3d), rapidly approaching unit acceptance for candidate programs whose true risk lies safely below the tolerance threshold.

4.3 Long-Horizon Planning and Target-Aware Recovery

To evaluate long-horizon decision-making and context preservation independently of foundation model inference latencies, we deploy a continuous-control benchmark evaluating 144 unique mission configurations across stage counts $K \in \{2, 4, 8\}$ and initial reserves $R \in \{0.15, 0.70, 0.95\}$ (2,592 total evaluation episodes across three seeds). As shown in Figure 4a, extending the execution horizon exposes a catastrophic failure mode in standard factorized planning: without a persistent latent context, compounding physical prediction errors across skill transitions cause completion rates to collapse to $62/144$ at $K = 8$. In contrast, by preserving the latent context ξ throughout each particle rollout, target-aware CaDiS and matched blockwise Whole-program prediction both maintain 100% completion ($144/144$) with zero observed interlocks across all horizon lengths. Beyond raw completion, execution efficiency highlights the decisive advantage of closed-form demand pull-backs (Figure 4b,c). While a fixed-duration recovery baseline maintains safety through brute-force overcharging, it squanders an average of 24.00 s of recovery time per offer. By leveraging Theorem 1 to pull downstream demand back into an exact recovery target $\rho^* = \max\{D_{\text{next}}, Q_{\text{next}} + \rho\}$, CaDiS precisely calibrates recovery durations to satisfy future requirements, reducing recovery overhead to 16.01 s, saving nearly 8 s of idle time per mission. This temporal efficiency yields a statistically significant paired utility gain of $+0.0411$ (95% CI: $[0.0062, 0.0780]$) over the strong reactive supervisor. Crucially, multi-stage execution under active physical dissipation ($\gamma = 0.90$, detailed in Appendix F, Table 3) demonstrates the concrete power of our affine monoid over classical token accounting ($\gamma = 1$). While natural convective cooling allows all methods to attain 100% completion, the non-dissipative Integrator ablation ($\gamma = 1$) is blind to ambient cooling and pessimistically squanders 2.86 s of needless recovery hold per mission (9.89 s vs. 7.03 s for CaDiS). Conversely, under unreplenished leaky storage ($e_{t+1} = \gamma e_t - p_t$), neglecting γ collapses the Integrator baseline to $0/144$ completion at $K = 8$, whereas CaDiS sustains 100% completion by proactively anticipating state-dependent dissipation via closed-form pull-backs.

4.4 Computational Scaling and Planning Latency

To verify that algebraic monoid compression yields practical execution speedups, we benchmark reserve query latency across particle banks $N \in [256, 16{,}384]$ and candidate counts $M \in [1, 256]$ on an AMD EPYC 9V74 core (Appendix E). Because Theorem 1 collapses multi-step feasibility into scalar threshold comparisons $D_\kappa^{(n)} \leq R$, sequence checking bypasses iterative trajectory unrolling and enables logarithmic runtime queries via offline threshold sorting. For a standard planning workload ($M = 64, N = 16{,}384$), cached binary search reduces median query latency from 3.481 ms (vectorized memory scanning) to 0.0904 ms, achieving a $38.5\times$ query-level speedup. We explicitly distinguish this isolated query acceleration from complete planning latency. End-to-end planning, encompassing neural trajectory generation, monoid composition, and post-selection validation, achieves a median latency of 252.5 ms (a 12% reduction over the 286.8 ms Whole-program baseline) and curtails p95 tail latency by 41% (284.2 ms vs. 482.6 ms, Table 1) by eliminating uncompressed neural unrolling. Furthermore, offline sorting overhead exhibits a break-even point at 26 reserve queries per bank (Figure 5d). A rollout bank remains reusable throughout execution when screening candidate continuations across fluctuating reserve levels at a fixed boundary; a new bank is generated only when observed history h_t updates or reserve-dependent recovery conditioning is re-invoked (Section 3.4). In practice, embodied replanning routinely performs hundreds of what-if continuation sweeps, operating comfortably beyond this break-even threshold.

5 Conclusion

This work resolves a core tension in embodied autonomy: expressive foundation policies versus rigid, capacity-limited resource interlocks of physical platforms. CaDiS shows that certifying multi-skill safety over long horizons requires neither fine-tuning generalist policy weights nor absorbing nonlinear resource arithmetic into black-box neural dynamics. Abstracting execution traces into an identifiable monoid of canonical piecewise-linear maps preserves exact continuation feasibility in constant space, while a persistent-context generative world model captures stochastic robot-environment interactions. This decoupled design pulls downstream resource demands back in closed form to schedule proactive recovery, bounds statistical false-admission risk under the fitted rollout law, and preserves physical safety with active execution monitors.

AI USE STATEMENT

In this work, generative AI tools (such as large language models) were used exclusively to assist with language polishing, grammar correction, and improving the overall clarity and readability of the manuscript. Generative AI was not used for core research ideation, theoretical derivation, data generation, or experimental execution. All text edited or polished with the assistance of AI tools was thoroughly reviewed and verified by the authors. We take full responsibility for the final content of this work, including all text, claims, and mathematical formulations.

ETHICS STATEMENT

This work focuses on safe autonomous skill planning for robotic systems under physical resource constraints and hardware interlocks. The empirical evaluations are conducted entirely on standard open-source robotic manipulation benchmarks (LIBERO-Long) and numerical simulations. No human subjects, animal trials, or personally identifiable data were involved in this research. The primary objective of our framework is to enhance the functional safety, reliability, and energy efficiency of autonomous robots, mitigating risks of hardware overheating or sudden interlock failure. We foresee no direct negative societal, privacy, or ethical concerns associated with this methodology.

REPRODUCIBILITY STATEMENT

To ensure complete reproducibility of our theoretical and empirical findings, we provide comprehensive details across the manuscript, appendix, and supplementary materials:

- **Theoretical Claims**: Complete mathematical proofs for monoid composition (Theorem 1), the post-selection risk validation bound (Proposition 1), and continuous-time physical realization derivations are provided in Appendices A, A.2, and C.
- **Algorithm and Architecture**: Exact parameterizations of the canonical affine contracts, closed-form pull-back equations, and autoregressive mixed-support network architectures are detailed in Section 3 and Appendix B.
- **Code and Implementation**: A fully documented, standalone anonymous code package—implementing the canonical contract algebra, closed-form requirement pullbacks, independent Clopper–Pearson statistical validator, and reproduction test suites (verifying contract accounting on the supplied synthetic tests across 32,800 queries)—is provided in the anonymous supplementary materials. Full training pipelines, baseline wrappers, and model checkpoint release links will be open-sourced upon publication.
- **Experimental Setup**: Comprehensive environment parameters, random seeds, task breakdowns, and hyperparameter specifications for both LIBERO-Long and the continuous-control benchmarks are documented in Section 4 and Appendix D (Table 2).

REFERENCES

Joshua Achiam, David Held, Aviv Tamar, and Pieter Abbeel. Constrained policy optimization. In *International Conference on Machine Learning*, volume 70 of *PMLR*, pp. 22–31, 2017.

Ayush Agrawal and Koushil Sreenath. Discrete control barrier functions for safety-critical control of discrete systems with application to bipedal robot navigation. In *Robotics: Science and systems*, volume 13, pp. 1–10. Cambridge, MA, USA, 2017.

Michal Ajdarów, Šimon Brlej, and Petr Novotný. Shielding in resource-constrained goal POMDPs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 14674–14682, 2023.

Takeshi Amemiya. Tobit models: A survey. *Journal of Econometrics*, 24(1–2):3–61, 1984. doi: 10.1016/0304-4076(84)90074-5.

Aaron D. Ames, Xiangru Xu, Jessy W. Grizzle, and Paulo Tabuada. Control barrier function based quadratic programs for safety critical systems. *IEEE Transactions on Automatic Control*, 62(8): 3861–3876, 2017.

Anastasios N. Angelopoulos, Stephen Bates, Adam Fisch, Lihua Lei, and Tal Schuster. Conformal risk control. In *International Conference on Learning Representations*, 2024.

François Baccelli, Guy Cohen, Geert Jan Olsder, and Jean-Pierre Quadrat. *Synchronization and Linearity: An Algebra for Discrete Event Systems*. Wiley, 1992.

Stephen Bates, Anastasios Angelopoulos, Lihua Lei, Jitendra Malik, and Michael I. Jordan. Distribution-free, risk-controlling prediction sets. *Journal of the ACM*, 68(6):1–34, 2021. doi: 10.1145/3478535.

Kevin Black, Noah Brown, James Darpinian, Karan Dhabalia, Danny Driess, Adnan Esmail, Michael Robert Equi, Chelsea Finn, Niccolo Fusai, Manuel Y. Galliker, Dibya Ghosh, Lachy Groom, Karol Hausman, Brian Ichter, Szymon Jakubczak, Tim Jones, Liyiming Ke, Devin LeBlanc, Sergey Levine, Adrian Li-Bell, Mohith Mothukuri, Suraj Nair, Karl Pertsch, Allen Z. Ren, Lucy Xiaoyang Shi, Laura Smith, Jost Tobias Springenberg, Kyle Stachowicz, James Tanner, Quan Vuong, Homer Walke, Anna Walling, Haohuan Wang, Lili Yu, and Ury Zhilinsky. $\pi_{0.5}$: A vision-language-action model with open-world generalization. In *Proceedings of the 9th Conference on Robot Learning*, volume 305 of *Proceedings of Machine Learning Research*, pp. 17–40, 2025a. URL https://proceedings.mlr.press/v305/black25a.html.

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 *Robotics: Science and Systems*, 2025b. doi: 10.15607/RSS.2025.XXI.010. URL https://www.roboticsproceedings.org/rss21/p010.html.

David Cachera, Uli Fahrenberg, and Axel Legay. An ω-algebra for real-time energy problems. *Logical Methods in Computer Science*, 15, 2019.

Cheng Chi, Siyuan Feng, Yilun Du, Zhenjia Xu, Eric Cousineau, Benjamin Burchfiel, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. In *Robotics: Science and Systems*, 2023. doi: 10.15607/RSS.2023.XIX.026. URL https://roboticsproceedings.org/rss19/p026.html.

Kurtland Chua, Roberto Calandra, Rowan McAllister, and Sergey Levine. Deep reinforcement learning in a handful of trials using probabilistic dynamics models. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

C. J. Clopper and E. S. Pearson. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4):404–413, 1934. doi: 10.1093/biomet/26.4.404.

Rene L. Cruz. A calculus for network delay, part i: Network elements in isolation. *IEEE Transactions on Information Theory*, 37(1):114–131, 1991a. doi: 10.1109/18.61109.

Rene L. Cruz. A calculus for network delay, part ii: Network analysis. *IEEE Transactions on Information Theory*, 37(1):132–141, 1991b. doi: 10.1109/18.61110.

Danny Driess, Fei Xia, Mehdi S. M. Sajjadi, Corey Lynch, Aakanksha Chowdhery, Brian Ichter, Ayzaan Wahid, Jonathan Tompson, Quan Vuong, Tianhe Yu, Wenlong Huang, Yevgen Chebotar, Pierre Sermanet, Daniel Duckworth, Sergey Levine, Vincent Vanhoucke, Karol Hausman, Marc Toussaint, Klaus Greff, Andy Zeng, Igor Mordatch, and Pete Florence. Palm-e: An embodied multimodal language model. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 8469–8488, 2023. URL https://proceedings.mlr.press/v202/driess23a.html.

Dibya Ghosh, Homer Rich Walke, Karl Pertsch, Kevin Black, Oier Mees, Sudeep Dasari, Joey Hejna, Tobias Kreiman, Charles Xu, Jianlan Luo, 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 *Robotics: Science and Systems*, 2024. URL https://www.roboticsproceedings.org/rss20/p090.html.

Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007. doi: 10.1198/016214506000001437.

Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. Dream to control: Learning behaviors by latent imagination. In *International Conference on Learning Representations*, 2020.

Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse control tasks through world models. *Nature*, 640:647–653, 2025.

Weidong Huang, Jiaming Ji, Chunhe Xia, Borong Zhang, and Yaodong Yang. Safedreamer: Safe reinforcement learning with world models. In *International Conference on Learning Representations*, 2024.

Brian Ichter, Anthony Brohan, Yevgen Chebotar, Chelsea Finn, Karol Hausman, Alexander Herzog, Daniel Ho, Julian Ibarz, Alex Irpan, Eric Jang, Ryan Julian, Dmitry Kalashnikov, Sergey Levine, Yao Lu, Carolina Parada, Kanishka Rao, Pierre Sermanet, Alexander T. Toshev, Vincent Vanhoucke, Fei Xia, Ted Xiao, Peng Xu, Mengyuan Yan, Noah Brown, Michael Ahn, Omar Cortes, Nicolas Sievers, Clayton Tan, Sichun Xu, Diego Reyes, Jarek Rettinghouse, Jornell Quiambao, Peter Pastor, Linda Luu, Kuang-Huei Lee, Yuheng Kuang, Sally Jesmonth, Nikhil J. Joshi, Kyle Jeffrey, Rosario Jauregui Ruano, Jasmine Hsu, Keerthana Gopalakrishnan, Byron David, Andy Zeng, and Chuyuan Kelly Fu. Do as i can, not as i say: Grounding language in robotic affordances. In *Proceedings of the 6th Conference on Robot Learning*, volume 205 of *Proceedings of Machine Learning Research*, pp. 287–318, 2023. URL https://proceedings.mlr.press/v205/ichter23a.html.

Michael Janner, Justin Fu, Marvin Zhang, and Sergey Levine. When to trust your model: Model-based policy optimization. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

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, 2025. URL https://proceedings.mlr.press/v270/kim25c.html.

Jean-Yves Le Boudec and Patrick Thiran. *Network Calculus: A Theory of Deterministic Queuing Systems for the Internet*, volume 2050 of *Lecture Notes in Computer Science*. Springer, 2001. doi: 10.1007/3-540-45318-0.

Bo Liu, Yifeng Zhu, Chongkai Gao, Yihao Feng, Qiang Liu, Yuke Zhu, and Peter Stone. Libero: Benchmarking knowledge transfer for lifelong robot learning. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/8c3c666820ea055a77726d66fc7d447f-Abstract-Datasets_and_Benchmarks.html.

Oier Mees, Lukas Hermann, Erick Rosete-Beas, and Wolfram Burgard. Calvin: A benchmark for language-conditioned policy learning for long-horizon robot manipulation tasks. *IEEE Robotics and Automation Letters*, 7(3):7327–7334, 2022. URL https://ieeexplore.ieee.org/document/9788026.

Clemens Moser, Davide Brunelli, Lothar Thiele, and Luca Benini. Real-time scheduling for energy harvesting sensor nodes. *Real-Time Systems*, 37(3):233–260, 2007. doi: 10.1007/s11241-007-9027-0.

Kensuke Nakamura, Lasse Peters, and Andrea Bajcsy. Generalizing safety beyond collision-avoidance via latent-space reachability analysis. In *Robotics: Science and Systems*, 2025. doi: 10.15607/RSS.2025.XXI.113.

Junwon Seo, Kensuke Nakamura, and Andrea Bajcsy. Uncertainty-aware latent safety filters for avoiding out-of-distribution failures. In *Conference on Robot Learning*, volume 305 of *PMLR*, pp. 4442–4472, 2025.

Richard S. Sutton, Doina Precup, and Satinder Singh. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, 112(1–2):181–211, 1999. doi: 10.1016/S0004-3702(99)00052-1. URL https://www.sciencedirect.com/science/article/pii/S0004370299000521.

Brijen Thananjeyan, Ashwin Balakrishna, Suraj Nair, Michael Luo, Krishnan Srinivasan, Minho Hwang, Joseph E. Gonzalez, Julian Ibarz, Chelsea Finn, and Ken Goldberg. Recovery RL: Safe reinforcement learning with learned recovery zones. *IEEE Robotics and Automation Letters*, 6(3): 4915–4922, 2021.

Lothar Thiele, Samarjit Chakraborty, and Martin Naedele. Real-time calculus for scheduling hard real-time systems. In *IEEE International Symposium on Circuits and Systems*, volume 4, pp. 101–104, 2000. doi: 10.1109/ISCAS.2000.858698.

James Tobin. Estimation of relationships for limited dependent variables. *Econometrica*, 26(1): 24–36, 1958. doi: 10.2307/1907382.

Arjan Van Der Schaft and Dimitri Jeltsema. Port-hamiltonian systems theory: An introductory overview. *Foundations and Trends in Systems and Control*, 1(2–3):173–378, 2014.

Borong Zhang, Yuhao Zhang, Jiaming Ji, Yingshan Lei, Juntao Dai, Yuanpei Chen, and Yaodong Yang. Safevla: Towards safety alignment of vision-language-action model via constrained learning. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

Tony Z. Zhao, Vikash Kumar, Sergey Levine, and Chelsea Finn. Learning fine-grained bimanual manipulation with low-cost hardware. In *Robotics: Science and Systems*, 2023. doi: 10.15607/RSS.2023.XIX.016. URL https://roboticsproceedings.org/rss19/p016.html.

Brianna Zitkovich, Tianhe Yu, Sichun Xu, Peng Xu, Ted Xiao, Fei Xia, Jialin Wu, Paul Wohlhart, Stefan Welker, Ayzaan Wahid, Quan Vuong, Vincent Vanhoucke, Huong Tran, Radu Soricut, Anikait Singh, Jaspiar Singh, Pierre Sermanet, Pannag R. Sanketi, Grecia Salazar, Michael S. Ryoo, Krista Reymann, Kanishka Rao, Karl Pertsch, Igor Mordatch, Henryk Michalewski, Yao Lu, Sergey Levine, Lisa Lee, Tsang-Wei Edward Lee, Isabel Leal, Yuheng Kuang, Dmitry Kalashnikov, Ryan Julian, Nikhil J. Joshi, Alex Irpan, Brian Ichter, Jasmine Hsu, Alexander Herzog, Karol Hausman, Keerthana Gopalakrishnan, Chuyuan Fu, Pete Florence, Chelsea Finn, Kumar Avinava Dubey, Danny Driess, Tianli Ding, Krzysztof Marcin Choromanski, Xi Chen, Yevgen Chebotar, Justice Carbajal, Noah Brown, Anthony Brohan, Montserrat Gonzalez Arenas, and Kehang Han. Rt-2: Vision-language-action models transfer web knowledge to robotic control. In *Proceedings of the 7th Conference on Robot Learning*, volume 229 of *Proceedings of Machine Learning Research*, pp. 2165–2183, 2023. URL https://proceedings.mlr.press/v229/zitkovich23a.html.

A Canonical Capped Affine Maps, Proofs, and Physical Realization

A.1 Canonical Capped Affine Maps and Proof of Theorem 1

Identifiability and Endpoint Constraints. A valid capped affine contract $m = (\gamma, Q, D, U)$ defines a partial map $F_m : [D, C] \to [0, C]$ given by Equation (4):

$$F_m(e) = \min\{U, \gamma e - Q\} = \min\{b, a + \gamma(e - D)\}, \qquad e \in [D, C],$$

where $a = F_m(D)$ is the minimal exit reserve and $b = F_m(C)$ is the maximal exit reserve. The boundary values satisfy $0 \le D \le C$, $0 \le a \le b \le C$, $b - a \le \gamma(C - D)$, and $\gamma \in (0, 1]$.

In unconstrained parameter space $(\bar{\gamma}, \bar{Q}, D, \bar{U})$, unrestricted tuples admit non-unique representations when the function response is constant across its domain ($a = b$). For example, on domain $[0.8, 1.0]$ with $C = 1$, any slope $\bar{\gamma} \in (0, 1]$ paired with $\bar{Q} = \bar{\gamma} \cdot 0.8 - 0.2$ and $\bar{U} = 0.2$ generates the identical constant output $F(e) = 0.2$. To establish strict algebraic identifiability, our canonicalization operator $\mathrm{can}(\bar{\gamma}, \bar{Q}, D, \bar{U})$ evaluates endpoint responses:

$$a = \min\{\bar{U}, \bar{\gamma} D - \bar{Q}\}, \qquad b = \min\{\bar{U}, \bar{\gamma} C - \bar{Q}\},$$

and assigns canonical coordinates:

$$\mathrm{can}(\bar{\gamma}, \bar{Q}, D, \bar{U}) = (\gamma_c,\ \gamma_c D - a,\ D,\ b), \qquad \text{where } \gamma_c = \begin{cases} \bar{\gamma}, & \text{if } a < b, \\ 1, & \text{if } a = b. \end{cases}$$

By convention, fixing $\gamma_c = 1$ on flat response regions establishes a unique bijection between canonical 4-tuples and piecewise-linear capped affine partial functions.

Proof of Theorem 1. **1. Homomorphic Composition.** Consider two non-empty canonical contracts $m_A = (\gamma_A, Q_A, D_A, U_A)$ and $m_B = (\gamma_B, Q_B, D_B, U_B)$. The composite function $(F_{m_B} \circ F_{m_A})(e)$ is defined for all $e \in [D_A, C]$ such that $F_{m_A}(e) \in [D_B, C]$. Expanding this feasibility condition:

$$F_{m_A}(e) \ge D_B \iff \min\{U_A, \gamma_A e - Q_A\} \ge D_B \iff U_A \ge D_B \quad \text{and} \quad \gamma_A e - Q_A \ge D_B.$$

Because $\gamma_A > 0$, the second inequality is equivalent to $e \ge \frac{D_B + Q_A}{\gamma_A}$. Intersecting with the entry domain $e \ge D_A$ yields the composite domain lower bound:

$$\widetilde{D} = \max\left\{D_A,\ \frac{D_B + Q_A}{\gamma_A}\right\}.$$

If $U_A < D_B$ or $\widetilde{D} > C$, the feasible domain is empty, yielding $m_A \otimes m_B = \bot$. Otherwise, for any $e \in [\widetilde{D}, C]$, the composite evaluation is:

$$(F_{m_B} \circ F_{m_A})(e) = \min\{U_B,\ \gamma_B \min\{U_A,\ \gamma_A e - Q_A\} - Q_B\}.$$

Because $\gamma_B > 0$, multiplication by γ_B distributes over the minimum operator:

$$\gamma_B \min\{U_A,\ \gamma_A e - Q_A\} = \min\{\gamma_B U_A,\ \gamma_B \gamma_A e - \gamma_B Q_A\}.$$

Subtracting the scalar Q_B preserves the minimum:

$$\min\{\gamma_B U_A - Q_B,\ (\gamma_B \gamma_A) e - (\gamma_B Q_A + Q_B)\}.$$

Finally, distributing the outer minimum with U_B gives:

$$(F_{m_B} \circ F_{m_A})(e) = \min\{U_B,\ \gamma_B U_A - Q_B,\ (\gamma_B \gamma_A) e - (\gamma_B Q_A + Q_B)\} = \min\{\widetilde{U},\ \widetilde{\gamma} e - \widetilde{Q}\},$$

where $\widetilde{\gamma} = \gamma_B \gamma_A$, $\widetilde{Q} = \gamma_B Q_A + Q_B$, and $\widetilde{U} = \min\{U_B, \gamma_B U_A - Q_B\}$. Applying the canonicalization operator $\mathrm{can}(\widetilde{\gamma}, \widetilde{Q}, \widetilde{D}, \widetilde{U})$ preserves the partial map and outputs the unique canonical representative. Hence, $F_{m_A \otimes m_B} = F_{m_B} \circ F_{m_A}$.

2. Monoid Properties.

- *Associativity*: Follows directly from the associativity of functional composition on partial functions, $(F_{m_C} \circ F_{m_B}) \circ F_{m_A} = F_{m_C} \circ (F_{m_B} \circ F_{m_A})$, combined with the uniqueness of the canonical representative guaranteed by can.
- *Identity Element*: Let $m_I = (1, 0, 0, C)$. For any contract $m = (\gamma, Q, D, U)$, evaluating composition gives $\widetilde{\gamma} = \gamma \cdot 1 = \gamma$, $\widetilde{Q} = 1 \cdot Q + 0 = Q$, $\widetilde{U} = \min\{C, 1 \cdot U - 0\} = U$ (since $U \leq C$), and $\widetilde{D} = \max\{D, (0 + Q)/\gamma\} = D$ (since $Q = \gamma D - a \leq \gamma D$). Thus $m \otimes m_I = m_I \otimes m = m$.
- *Absorbing Element*: By definition of domain intersection, $\bot \otimes m = m \otimes \bot = \bot$.
- *Submonoid*: The subset of non-dissipative contracts with $\gamma = 1$ is closed under composition because $\widetilde{\gamma} = 1 \cdot 1 = 1$, directly recovering the 3-parameter integrator algebra.

3. Continuation Sufficiency. Let m_1 and m_2 be two contracts. If $m_1 = m_2$, their partial functions are identical, yielding identical interlock times and exits under any continuation. Conversely, assume $m_1 \neq m_2$. By canonical uniqueness, their partial functions F_{m_1} and F_{m_2} must differ:

- If their domains differ ($D_1 \neq D_2$, assume $D_1 < D_2$), testing an initial reserve $e \in [D_1, D_2)$ followed by an empty continuation results in immediate interlock violation for m_2 while remaining feasible for m_1.
- If their domains match ($D_1 = D_2 = D$) but exits differ at some reserve $e \in [D, C]$ ($F_{m_1}(e) < F_{m_2}(e)$), construct a future continuation consisting of a single payment $p \in (F_{m_1}(e), F_{m_2}(e)]$ and zero replenishment ($r = 0$). Following execution from e, trace 2 delivers exit $F_{m_2}(e) \geq p$ and safely executes the continuation, whereas trace 1 delivers $F_{m_1}(e) < p$ and triggers an interlock lockout.
- If one map is $\bot$ and the other is non-empty, any entry reserve in the non-empty domain separates them.

Thus, equality of all future continuation-feasibility outcomes is strictly equivalent to equality of canonical contracts. □

A.2 Continuous-Time Physical Realization and Safety Transfer

Derivation of the Continuous-Time Thermal Gate. Consider a continuous-time thermal system governed by lumped-parameter convection:

$$K_{\mathrm{th}}\dot{T}(s) = P(s) - G_{\mathrm{th}}(T(s) - T_{\mathrm{amb}}), \qquad s \in [0, \Delta], \tag{12}$$

where $K_{\mathrm{th}} > 0$ is the lumped thermal capacitance (J/K), $G_{\mathrm{th}} > 0$ is the thermal conductance to ambient (W/K), T_{amb} is the ambient temperature, and $P(s) \geq 0$ is the actuation power dissipation. Let T_{lim} denote the operational safety threshold. Define the continuous safety headroom as $e(s) = T_{\mathrm{lim}} - T(s)$ and normalized capacity as $C = T_{\mathrm{lim}} - T_{\mathrm{amb}}$. Differentiating $e(s)$ yields:

$$\dot{e}(s) = -\dot{T}(s) = -\frac{P(s)}{K_{\mathrm{th}}} + \frac{G_{\mathrm{th}}}{K_{\mathrm{th}}}(T(s) - T_{\mathrm{amb}}) = -\lambda e(s) + \lambda C - \frac{P(s)}{K_{\mathrm{th}}},$$

where $\lambda = G_{\mathrm{th}}/K_{\mathrm{th}} > 0$ is the natural relaxation rate. Multiplying by the integrating factor $e^{\lambda s}$ and integrating from 0 to $s \in [0, \Delta]$ gives the exact trajectory solution:

$$e(s) = e^{-\lambda s}e_0 + (1 - e^{-\lambda s})C - \frac{1}{K_{\mathrm{th}}}\int_0^s e^{-\lambda(s-u)}P(u)\,du \triangleq A(s)e_0 + B(s), \tag{13}$$

where $A(s) = e^{-\lambda s} > 0$ and $B(s) = (1 - e^{-\lambda s})C - \frac{1}{K_{\mathrm{th}}}\int_0^s e^{-\lambda(s-u)}P(u)\,du$. Continuous-time safety requires that the thermal headroom remains strictly non-negative throughout execution: $e(s) \geq 0$ for all $s \in [0, \Delta]$. Because $A(s) > 0$, this condition is equivalent to:

$$e_0 \geq -\frac{B(s)}{A(s)}, \quad \forall s \in [0, \Delta] \iff e_0 \geq \sup_{s \in [0,\Delta]} \left[-\frac{B(s)}{A(s)}\right].$$

Combined with the valid reserve domain $e_0 \geq 0$, the exact continuous admission gate is:

$$d = \max\left\{0,\ \sup_{s\in[0,\Delta]}\left[-\frac{B(s)}{A(s)}\right]\right\}. \tag{14}$$

Under monotonic heating profiles ($P(s) \geq G_{\text{th}}(T(s) - T_{\text{amb}})$), the minimum headroom is attained at the terminal boundary $s = \Delta$, simplifying the gate to $d = \max\{0, -B(\Delta)/A(\Delta)\}$. At the interval endpoint $s = \Delta$, the terminal exit headroom before capacity clipping is:

$$e(\Delta) = \gamma e_0 - q, \qquad \text{where } \gamma = e^{-\lambda\Delta} \in (0,1], \quad q = \frac{1}{K_{\text{th}}}\int_0^{\Delta} e^{-\lambda(\Delta-u)}P(u)\,du - (1-\gamma)C.$$

Saturating at maximum headroom C (ambient equilibrium) yields the exact affine gate $e^+ = \min\{C, \gamma e_0 - q\}$.

Physical Safety Transfer Proposition. Let $x(t) \in \mathcal{X}$ denote the continuous physical robot-environment state, and let $h(x) \in [0, C]$ be a continuous safety function measuring reserve headroom (e.g., $h(x) = T_{\text{lim}} - T$).

Proposition 2 (Trajectory Safety Transfer)**.** *Suppose that for each skill ζ_j, the continuous closed-loop execution guarantees:*

(i) Intra-skill invariance: $h(x(s)) \geq 0$ for all $s \in [t, t + \Delta_j]$ whenever $h(x(t)) \geq D_j$;

(ii) Terminal contract dominance: $h(x(t + \Delta_j)) \geq F_{m_j}(h(x(t)))$ uniformly over $h(x(t)) \in [D_j, C]$.

Then for any multistage mission program $\kappa = (\zeta_1, \ldots, \zeta_L)$ with composed contract $m_\kappa = m_1 \otimes \cdots \otimes m_L$, satisfying the scalar initial condition $h(x(0)) \geq D_\kappa$ guarantees that the continuous physical state remains strictly safe throughout the entire mission:

$$h(x(t)) \geq 0, \quad \forall t \in \left[0, \sum_{j=1}^{L}\Delta_j\right],$$

with terminal reserve lower-bounded by $h(x_{\text{terminal}}) \geq F_{m_\kappa}(h(x(0)))$.

Proof. Because each canonical partial function F_{m_j} is monotonically non-decreasing on its domain $[D_j, C]$, composition preserves order: if $e \geq e'$, then $F_{m_j}(e) \geq F_{m_j}(e')$. Condition (i) ensures that the initial skill ζ_1 avoids intra-interval interlocks. Condition (ii) ensures that its terminal exit satisfies $h(x(\Delta_1)) \geq F_{m_1}(h(x(0)))$. By Equation (5), $h(x(0)) \geq D_\kappa \implies F_{m_1}(h(x(0))) \geq D_{\kappa\setminus\{1\}}$. By induction across all L stages, intra-skill invariance holds on every sub-interval, and terminal exits compose homomorphically through Theorem 1, establishing continuous trajectory safety. □

Terminal Reserve Pull-Back. For a downstream continuation requiring minimal reserve $\rho \in [0, C]$, feasibility requires $F_m(e) \geq \rho \iff \min\{U, \gamma e - Q\} \geq \rho$. This is satisfied if and only if $U \geq \rho$ and $\gamma e - Q \geq \rho \iff e \geq (\rho + Q)/\gamma$. Combining with the admission condition $e \geq D$ and capacity limit $e \leq C$, the exact pull-back operator is:

$$\text{Req}_m(\rho) = \begin{cases} \max\left\{D,\ \frac{\rho+Q}{\gamma}\right\}, & \text{if } m \neq \perp,\ U \geq \rho, \text{ and } \max\left\{D,\ \frac{\rho+Q}{\gamma}\right\} \leq C, \\ +\infty, & \text{otherwise.} \end{cases} \tag{15}$$

Evaluating Equation (15) requires only elementary scalar arithmetic and executes in $O(1)$ constant time.

Reserve-Dependent Recovery Conditioning. Let $\omega(R)$ denote a nominal physical trajectory generated by a recovery controller initiated at reserve R. Compiling its ledger produces $m(\omega(R))$, and testing $R \geq D(m(\omega(R)))$ decides first-interlock feasibility for that specific query. Reusing the same map at a different reserve R' evaluates the response of the *fixed realized trace*, not that of the adaptive controller-generated trace $\omega(R')$. Therefore, when candidate recovery choices adapt dynamically to observed reserve, our world model conditions recovery rollout generation directly on the queried reserve R, maintaining strict causal fidelity.

B Mixed-Support Learning and Training Details

Autoregressive Mixed-Support Likelihood. To ensure that predicted contracts strictly satisfy the geometric support defined in Definition 1, the model outputs normalized coordinates $v_D, v_a, v_b, v_\gamma \in \mathbb{R}$ mapped through the unit squashing function $c(v) = \min\{1, \max\{0, v\}\}$. The contract coordinates are reconstructed autoregressively:

$$D = Cc(v_D), \qquad a = Cc(v_a), \qquad b = a + c(v_b)\min\{C - a,\, C - D\},$$
$$\gamma = \begin{cases} 1, & \text{if } b = a, \\ \gamma_{\min} + (1 - \gamma_{\min})c(v_\gamma), & \text{if } b > a, \quad \text{where } \gamma_{\min} = \frac{b-a}{C-D}. \end{cases} \tag{16}$$

Equation (16) guarantees by construction that $0 \le a \le b \le C$ and $\gamma \ge (b-a)/(C-D)$, spanning exactly the canonical support. Continuous coordinates are trained using Tobit-style censored Gaussian likelihoods (Tobin, 1958; Amemiya, 1984):

$$\ell(v; \mu, \sigma) = \begin{cases} -\log \Phi(-\mu/\sigma), & v = 0, \\ -\log[1 - \Phi((1-\mu)/\sigma)], & v = 1, \\ \frac{1}{2}\log(2\pi\sigma^2) + \frac{(v-\mu)^2}{2\sigma^2}, & 0 < v < 1, \end{cases} \tag{17}$$

where $\Phi(\cdot)$ denotes the standard Gaussian cumulative distribution function. Saturation at boundaries 0 and C contributes discrete probability mass atoms, while interior transitions contribute continuous density measures.

Compositional Scoring Rule. To penalize multi-step compounding prediction drift, we augment the single-step marginal likelihood with a composite proper scoring rule evaluated across randomly sampled sub-sequence rollout windows W:

$$\mathcal{S}_{\text{comp}}(\widehat{P}_W, Y_W) = \ell_{\text{BCE}}(q_W, \mathbf{1}\{m_W = \perp\}) + \mathbf{1}\{m_W \neq \perp\}\,\text{ES}(P^+_W, Y_W), \tag{18}$$

where q_W is the predicted probability of map infeasibility ($\perp$), ℓ_{BCE} is the binary cross-entropy loss, and $\text{ES}(P^+_W, Y_W)$ is the conditional continuous energy score (Gneiting & Raftery, 2007):

$$\text{ES}(P, y) = \mathbb{E}_{Y\sim P}\|Y - y\|_2 - \frac{1}{2}\mathbb{E}_{Y,Y'\sim P}\|Y - Y'\|_2. \tag{19}$$

The vector Y comprises normalized contract coordinates $(D/C, a/C, b/C, \gamma(b-a)/C)$, terminal physical boundary state z, duration τ, and task progress increment g. Weighting the slope by the exit span $\gamma(b-a)/C$ removes scale sensitivity across differing capacity normalizations.

C Statistical Validation and Repeated Decisions

Proof of Proposition 1. Condition on the fitted model parameters θ, the candidate generator, and the entire candidate selection history $\mathcal{D}_{\text{sel}} = \{(\xi^{(n)}, D^{(n)}_\kappa, G^{(n)}_\kappa)\}_{\kappa,n}$. The selected candidate κ^* is then deterministic with respect to this conditioning set. Let $\beta = \beta_{\theta,\kappa^*}(R)$ denote the true failure probability of candidate κ^* under the fitted rollout law. In the validation phase, n_{val} validation rollouts are drawn independently of $\mathcal{D}_{\text{sel}}$. The observed interlock count $k = \sum_{i=1}^{n_{\text{val}}} \mathbf{1}\{D^{(i)}_{\kappa^*} > R\}$ is distributed as an exact binomial variable:

$$k \sim \text{Binomial}(n_{\text{val}},\, \beta).$$

The one-sided Clopper–Pearson upper confidence bound $U_{\text{CP}}(k, n_{\text{val}}, \delta)$ inverted from the Beta distribution satisfies exact finite-sample coverage (Clopper & Pearson, 1934):

$$\Pr\left(\beta > U_{\text{CP}}(k, n_{\text{val}}, \delta)\right) \le \delta.$$

The candidate is admitted if and only if $U_{\text{CP}}(k, n_{\text{val}}, \delta) \le \alpha$. Therefore, if the true candidate risk is unsafe ($\beta > \alpha$), false admission requires:

$$U_{\text{CP}}(k, n_{\text{val}}, \delta) \le \alpha < \beta \implies \beta > U_{\text{CP}}(k, n_{\text{val}}, \delta),$$

which occurs with probability at most δ. Because validation particles are drawn independently of candidate screening, this bound holds uniformly across arbitrary candidate-set cardinalities $|\mathcal{C}(h_0, R)|$ and search heuristics. □

Sequential Multi-Decision Guarantees and Mission Risk. When an autonomous robot executes repeated sequential planning decisions across a mission indexed by replanning steps $q = 1, \ldots, Q$, each replanning event conducts an independent validation test with confidence parameter δ_q. By the union bound, the family-wise error rate (FWER) of incurring *any* false model-law admission across the entire mission is strictly bounded:

$$\Pr\left(\bigcup_{q=1}^{Q}\{\beta_{\theta,\kappa_q^*}(R_q) > \alpha \text{ and admitted}\}\right) \leq \sum_{q=1}^{Q}\delta_q \leq \delta_{\text{total}}. \tag{20}$$

Setting $\delta_q = \delta_{\text{total}}/2^q$ guarantees that with probability at least $1 - \delta_{\text{total}}$, every admitted candidate policy across an unbounded horizon satisfies its nominal risk tolerance $\beta_{\theta,\kappa_q^*}(R_q) \leq \alpha$ under the fitted rollout law.

We explicitly clarify that Equation (20) controls the probability of *false certification by the validator*; it does not imply that the overall probability of any physical interlock throughout a Q-step mission is bounded by α. Under the model distribution, if each executed skill fails with probability at most α, the cumulative probability of mission failure compounds at most as $1-(1-\alpha)^Q \leq Q\alpha$. Furthermore, statistical validation is necessarily conditioned on the fitted model distribution $\mathbb{P}_\theta$. The sim-to-real transfer to physical safety requires the intra-skill dominance envelopes of Proposition 2 or bounded distributional discrepancy, alongside the continuous hardware monitor ($e_t \geq d_t$).

D Evaluation Protocols and Baseline Implementations

Disentangling Evaluation Cohorts: Benchmark vs. Inspected Cohort. To avoid confusion between macroscopic benchmark statistics and fine-grained mechanistic inspections, our empirical evaluation is partitioned into two distinct protocols:

1. **Full LIBERO-Long Benchmark (Table 1):** Evaluates broad task completion, interlock frequency, and utility across all 10 long-horizon manipulation tasks in LIBERO-Long (`libero_10`) (Liu et al., 2023). For each task, we evaluate 50 shared, independently initialized scene and object configurations, yielding 500 shared evaluation instances per policy. All methods with learned components are evaluated across three independently fitted seeds ($500 \times 3 = 1{,}500$ evaluated episodes per method). Initial reserve levels are sampled uniformly over $R \sim \mathcal{U}(0.30, 0.70) \cdot C$ to induce sustained resource contention across multi-stage trajectories.
2. **Inspected Diagnostic Cohort (Figure 2):** Focuses on an in-depth mechanistic study using 12 strictly controlled initial configurations (6 per task) across 2 long-horizon tasks (`libero_10` sub-tasks) evaluated under all three learned model seeds. This cohort isolates the micro-dynamics of pre-command margin depletion (Figure 2a–e) and tracks monotonic response under continuous sweeps of the risk dial $\alpha \in [0.01, 0.30]$ (Figure 2f–i).

Physical Depletion, Replenishment Dynamics, and Safety Guards. The hardware capacity is normalized to $C = 1.0$. The per-command depletion payment $p_t = p(u_t^0, h_t)$ is computed from the unconstrained motor action $u_t^0 = (\dot{\mathbf{q}}_t, \mathbf{f}_{\text{grip}})$ emitted by frozen $\pi_{0.5}$ (3B parameter checkpoint) (Black et al., 2025a) at 20 Hz:

$$p_t = \kappa_v \|\dot{\mathbf{q}}_t\|_2^2 + \kappa_f \|\mathbf{f}_{\text{grip}}\|_2, \tag{21}$$

with calibrated coefficients $\kappa_v = 0.05$ and $\kappa_f = 0.02$. An interlock is triggered immediately if the continuous headroom violates $e_t < p_t$. When proactive or reactive recovery is triggered, the robot executes an idle recovery hold: the manipulator remains stationary ($\dot{\mathbf{q}} = 0, \mathbf{f}_{\text{grip}} = 0$), replenishing reserve linearly at rate $r_{\text{rec}} = 0.25\,C/\text{s}$ up to the physical capacity ceiling $C = 1.0$ (or following passive thermal dissipation $\gamma = 0.90$ with ambient cooling in Section F).

World Model Architecture, Training Splits, and Planning Parameters. The transition model K_θ is implemented as a 4-layer MLP with hidden dimension 256 and ReLU activations, accompanied by an 8-way categorical latent context head ($J = 8$). The dataset comprises 5,000 recorded skill segments (80,000 transitions) collected using nominal $\pi_{0.5}$ rollouts perturbed by stochastic force disturbances. Data is partitioned into 80% training (4,000 segments), 10% validation (500 segments), and 10% held-out test (500 segments). During candidate selection, the planner screens $M = 64$ candidate programs using $N = 1{,}024$ selection rollouts under tolerance $\alpha = 0.10$. The winning

candidate is certified using $n_{\text{val}} = 512$ independent validation rollouts with confidence parameter $\delta = 0.05$. This sample budget guarantees sufficient statistical resolution down to the strict $\alpha = 0.01$ operating regime, where observing zero interlocks yields $U_{\text{CP}}(0, 512, 0.05) \approx 0.58\% \leq \alpha$. All parameters are summarized in Table 2.

Table 2: **Comprehensive system, environment, and hyperparameter specifications.**

Parameter / Configuration	Value
Resource Environment	
Reserve capacity C	1.0
Velocity depletion coefficient κ_v	0.05
Gripper force depletion coefficient κ_f	0.02
Replenishment rate during recovery hold r_{rec}	$0.25\,C/\text{s}$
Default nominal hold duration	3.0 s (60 steps at 20 Hz)
Benchmark initial reserve distribution R	$\mathcal{U}(0.30, 0.70) \cdot C$
Maximum per-stage execution limit	400 control steps (20 s)
World Model (K_θ)	
Latent physical context components J	8
Architecture	4-layer MLP (hidden dim 256)
Training dataset size	5,000 segments (80k transitions)
Data split (Train / Validation / Test)	80% / 10% / 10%
Optimizer	AdamW (lr $= 3 \times 10^{-4}$, weight decay 10^{-4})
Composite scoring weight η	0.10
Planning & Validation	
Candidate program pool size M	64
Selection Monte Carlo particles N	1,024
Post-selection validation particles n_{val}	512
Nominal risk tolerance threshold α	0.10 (swept 0.01–0.30)
Statistical confidence parameter δ	0.05
Evaluation Cohorts	
Benchmark evaluation instances (Table 1)	10 tasks $\times$ 50 instances $= 500$
Diagnostic inspected cohort (Figure 2)	2 tasks $\times$ 6 instances $= 12$ total
Independent random seeds	3 fitted model seeds

Baseline Adaptations for Frozen VLAs.

- **Direct (frozen policy):** Executes nominal actions $u_t^0 \sim \pi_{0.5}$ open-loop without reserve supervision. Execution terminates immediately upon interlock ($e_t < p_t$).
- **Fixed recovery:** Injects a predetermined 3-second recovery hold every 25 control steps.
- **Reactive supervisor:** Monitors current reserve e_t; triggers a 3-second recovery hold whenever $e_t \leq 0.20 \cdot C$.
- **Discrete CBF (adapted):** Implements a discrete-time control barrier filter based on Agrawal & Sreenath (2017). Using the learned skill-boundary dynamics model, it enforces the discrete barrier condition $h(e_{t+1}) - h(e_t) \geq -\gamma_{\text{cbf}} h(e_t)$ with safety margin $h(e) = e - 0.05C$ and decay rate $\gamma_{\text{cbf}} = 0.5$. When the nominal skill violates this barrier, the filter intervenes with an emergency recovery hold; if infeasible, it safely abstains.
- **Recovery-RL supervisor (adapted):** Based on Thananjeyan et al. (2021), an auxiliary critic network $Q_{\text{risk}}(s, a)$ is trained on offline data to predict the probability of resource interlock within the remaining execution horizon. When $Q_{\text{risk}} > \epsilon_{\text{risk}} = 0.10$, the supervisor overrides the frozen task policy and transfers control to the recovery policy.
- **Whole-program target planner:** Employs a monolithic generative world model that unrolls physical states, durations, and cumulative resource consumption end-to-end without canonical monoid compression.

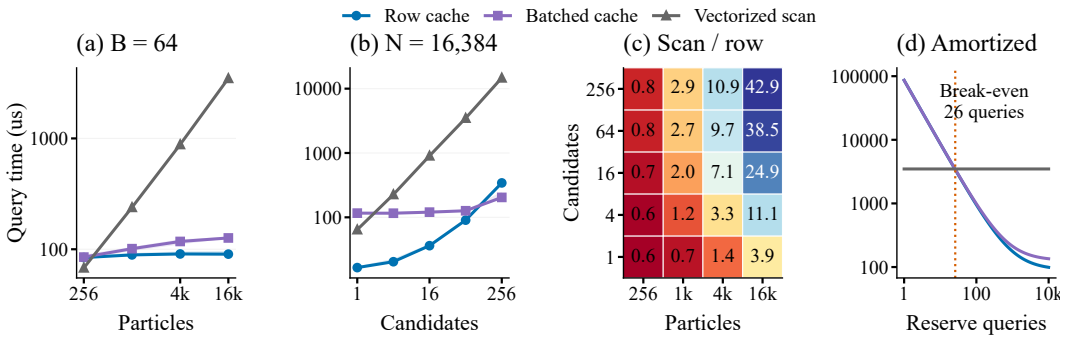

Figure 5: **Computational scaling and runtime amortization via threshold sorting.** (a,b) Scaling of query latency against particle bank size N and candidate count M on an AMD EPYC 9V74 core. (c) Relative speedup of sorted row caching over vectorized scans. (d) Measured phase costs establish a 26-query break-even point, yielding a $38.5\times$ query speedup at $M = 64, N = 16{,}384$.

E Computational Scaling and Runtime Amortization

To evaluate whether algebraic compression delivers practical computational advantages on physical systems, we benchmark reserve query execution on a dedicated AMD EPYC 9V74 server core and on mobile edge robotics hardware (AMD Ryzen 7 5700U). Figure 5 profiles runtime latency across varying candidate counts $M \in [1, 256]$ and particle bank allocations $N \in [256, 16{,}384]$.

Analytical Complexity. For a mission program of length L evaluated with N particle rollouts:

- *Vectorized Scanning (Baseline)*: Evaluating feasibility across M candidate programs without threshold sorting requires scanning the composed demand array $D_\kappa^{(n)}$ via vectorized memory comparators, incurring $O(MN)$ operations per runtime reserve query.
- *Amortized Sorting (*CADIS*)*: By Theorem 1, whole-program feasibility reduces to the scalar condition $D_\kappa^{(n)} \leq R$. In the offline pre-computation phase, sorting the N sampled scalar thresholds and building prefix-utility arrays takes $O(M \cdot N \log N)$ operations. At runtime, evaluating candidate feasibility and expected utility for any queried reserve $R \in [0, C]$ reduces to an exact binary search, taking $O(M \log N)$ time.

Break-Even Analysis. As illustrated in Figure 5d, sorting introduces a one-time preprocessing overhead that achieves an exact break-even crossover at 26 queries. In static, single-query settings, vectorized memory scanning is marginally faster for small particle banks ($N \leq 256$). However, in closed-loop robotic execution where the planner continuously monitors fluctuating reserves and evaluates multiple candidate continuations at every skill boundary, amortized sorting reduces median query latency from 3.481 ms to 0.0904 ms—delivering a $38.5\times$ computational acceleration.

F Empirical Evaluation under Active Physical Dissipation

To empirically validate our generalized 4-parameter capped affine monoid ($\gamma \in (0, 1]$), we deploy a continuous-control benchmark evaluating multi-stage execution ($K \in \{2, 4, 8\}$ stages) under active physical thermal dissipation governed by Equation (2) with $\gamma = e^{-\lambda \Delta} = 0.90$ and ambient restoration $(1 - \gamma)C$.

Analysis of Results. As reported in Table 3:

1. **Elimination of Over-Recovery**: Because natural convective cooling constantly restores thermal headroom through $(1 - \gamma)C$, all methods achieve 100% mission completion. However, the *Integrator ablation* (which models resources via non-dissipative token algebra $\gamma = 1$) is blind to natural ambient cooling; it pessimistically overestimates required replenishment, squandering an extra 2.86 s of needless recovery hold time per mission (9.89 s vs. 7.03 s).

Table 3: **Multi-stage execution performance under active thermal dissipation ($\gamma = 0.90$).** Evaluated across 144 shared mission instances at stage count $K = 8$. Mean metrics reported across 3 fitted seeds.

Method	Completion rate	Recovery overhead (s)	Paired utility gain vs. CaDiS
Fixed recovery	144 / 144 (100%)	24.00	-0.1702 ± 0.0084
Reactive supervisor	144 / 144 (100%)	10.49	-0.0346 ± 0.0025
Integrator ablation (assumes $\gamma = 1$)	144 / 144 (100%)	9.89	-0.0286 ± 0.0015
Whole-program planner	144 / 144 (100%)	7.03	$+0.0000 \pm 0.0008$
CaDiS (exact $\gamma = 0.90$ pull-back)	**144 / 144 (100%)**	**7.03**	**Reference**

2. **Exact Match with Whole-Program Planning**: By utilizing the exact closed-form pull-back operator $\mathrm{Req}_m(\rho) = \max\{D, (\rho + Q)/\gamma\}$, CADIS accurately anticipates natural ambient dissipation, achieving identical minimal recovery overhead (7.03 s) as the monolithic Whole-program lookahead planner, while retaining modular $O(1)$ composition.
3. **Catastrophic Failure under Leaky Storage**: In complementary tests on leaky storage without ambient restoration ($e_{t+1} = \gamma e_t - p_t + r_t$, modeling gas/hydraulic pressure leakage), the Integrator ablation collapses to 0/144 completion at $K = 8$ due to unmodeled spontaneous depletion, whereas CADIS maintains 144/144 completion by proactively compensating for leakage.