Kineforge: Sample-Efficient Embodied Control with Frozen Semantic Priors and Differentiable Physics

Technical Report v1.0 — Brightforge Software Inc.

Illia Ovcharenko

Brightforge Software Inc.
Kineforge AI-Lab.
`illia.ovcharenko@brightforge.ltd`

May 2026

Abstract

Embodied reinforcement learning often depends on large human demonstration datasets or billion-parameter foundation policies, limiting deployment where teleoperation is unsafe or uneconomical. We present **Kineforge**, combining differentiable physics (MuJoCo MJX) with a **multi-stream policy** that separates reflexive control, affect-modulated valuation, and topology-routed deliberation. A **frozen semantic prior** ($\sim$2.6M parameters) supplies structured inductive bias while $\sim$157K policy parameters are updated during learning, enabling single-GPU training and CPU inference under 10 ms. On NVIDIA H200 rollouts ($\geq$345K environment steps in monolithic jobs) and controlled 100K-step A/B studies, agents reach task rewards $\sim$0.70 with stable homeostatic metrics. Optional sensory grounding trades throughput ($\sim$159 vs. $\sim$108 env-steps/s) for comparable task success. We describe a sanitized evaluation protocol; proprietary prior databases are not released.

1 Introduction

Training physical agents at scale remains bottlenecked by *data*, not compute. Industrial manipulation, hazardous material handling, and remote logistics often forbid the collection of large teleoperation corpora. Recent vision-language-action (VLA) and foundation policies achieve strong performance but require web-scale pretraining, cloud inference, and demonstration-heavy finetuning [1, 6]. Model-based reinforcement learning (MBRL) with learned world models reduces sample complexity in simulation [3], yet struggles on novel tasks without structural priors.

We argue that **structured inductive bias**—not parameter count—is the lever for sample-efficient embodied control. **Kineforge** is a lightweight stack that couples (i) **differentiable physics simulation** (MJX) for gradient-friendly rollouts, (ii) a **multi-stream policy** with parallel reflex, affect, and planning pathways, and (iii) a **frozen semantic prior** encoding domain structure without online supervision. The trainable policy is compact ($\sim$157K parameters); priors remain fixed ($\sim$2.6M parameters), decoupling representation learning from control learning.

Contributions.

- A multi-stream embodied architecture with emotion-conditioned fast control and graph-topology-conditioned slow planning, integrated with canonical semantic embeddings.
- A JAX-native training pipeline (REINFORCE/PPO, JEPA-style regularization) over MJX with MLflow instrumentation and cloud orchestration (SkyPilot, H200).

- Empirical evidence from Nebius H200 campaigns: 100K-step A/B comparisons and long-horizon monolithic runs, including throughput–accuracy trade-offs for grounded sensing.
- A reproducibility package scope: sanitized APIs and metrics protocol; weights and prior corpora withheld for competitive reasons.

Results preview. In aggregated 100K-step studies, baseline agents achieve mean task reward $\sim$0.70 at $\sim$159 environment steps/s; grounded sensory variants match task reward at $\sim$108 steps/s (Table 1). Homeostatic metrics (conflict rate, nutrition gap) remain stable in JAX rollouts logged to MLflow.

2 Related Work

Sample-efficient embodied RL. RL with foundation priors (RLFP) demonstrates that frozen value and success models can accelerate real-robot learning [6]. Kineforge differs by using *structured symbolic priors* rather than general VLMs, targeting edge deployability.

World models and differentiable simulation. DreamerV3 learns latent dynamics for long-horizon control [3]. Differentiable engines (Brax/MJX) enable gradients through physics [2]. We use MJX as the environment backend while learning a factored policy rather than a monolithic world model.

Hierarchical and affective control. Hierarchical RL decomposes tasks into sub-policies; homeostatic and intrinsic motivation literature motivates separate valuation streams [5]. Our valence stream implements affect-modulated gains on fast pathways, stabilizing exploration without hand-crafted reward shaping.

Joint-embedding predictive architectures. JEPA encourages representation structure via predictive losses [4]. We apply a lightweight hierarchical regularization term during training (α=0.25, λ=0.1) without releasing proprietary embedding corpora.

3 Problem Formulation

We model embodied control as a partially observable Markov decision process (POMDP) $(\mathcal{S}, \mathcal{A}, \mathcal{O}, T, R, \gamma)$. Observations $o_t \in \mathcal{O}$ concatenate proprioceptive state, optional exteroceptive features, and a low-dimensional *configuration index* over a fixed topology graph. Actions factor into four branches: locomotion ($|\mathcal{A}_{\text{loc}}|$=16), respiration (2), head (2), and macro (4), totaling 24 discrete decisions per step.

Homeostatic objective. Beyond extrinsic task reward r_t^{task}, we log intrinsic homeostatic signals: food concentration, nutrition gap, conflict rate, and population-level coordination statistics. The training objective combines policy-gradient loss on r^{task} with auxiliary regularizers (JEPA and hierarchical structure terms). We do not assume access to human trajectories.

Parameter budget. Let θ denote trainable policy parameters ($|\theta|$≈1.57×10^5) and ϕ frozen semantic priors ($|\phi|$≈2.6×10^6). Only θ is updated during RL; ϕ is loaded from an offline corpus (not publicly released).

[Figure 1: Multi-stream architecture]
Observations → Frozen semantic encoder →
Fast (reflex) ‖ Valence (affect) ‖ Slow (topology route) → MJX

Figure 1: Kineforge control stack. Public diagrams omit proprietary graph connectivity and prior token identities.

4 Method

4.1 System overview

Figure 1 illustrates Kineforge. Per timestep, the agent encodes observations, retrieves a frozen semantic embedding, and routes decisions through three streams before actuating the differentiable simulator.

4.2 Frozen semantic prior

A corpus of N_{sym}=110 canonical symbols is embedded in $\mathbb{R}^{d_{\text{emb}}}$ (d_{emb}=768) offline. A projection head maps embeddings to policy features $z_t \in \mathbb{R}^{d_z}$. During RL, ϕ is frozen; gradients do not modify the corpus. This induces a **relational inductive bias**: the agent starts with structured semantics rather than learning language from scratch in simulation.

4.3 Multi-stream policy

Let h_t denote concatenated proprioceptive/visual features. The policy factorizes as

$$\pi_\theta(a_t \mid h_t, z_t, e_t, \rho_t) \propto \pi_{\text{fast}}(a_t^{\text{loc}}, a_t^{\text{head}} \mid h_t, z_t, e_t)\, \pi_{\text{val}}(a_t^{\text{breath}}, e_{t+1} \mid h_t, z_t, e_t)\, \pi_{\text{slow}}(a_t^{\text{macro}} \mid h_t, z_t, \rho_t), \tag{1}$$

where $e_t \in \mathbb{R}^{d_e}$ (d_e=7) is an affect state and $\rho_t \in \mathbb{R}^{d_\rho}$ is a route vector over a fixed topology graph (d_ρ=22).

Fast stream (reflex). Emotion-conditioned affine maps compute logits for locomotion and head actions:

$$\tilde{h}_t = h_t \odot \big(1 + \gamma \cdot \text{mean}(e_t)\big), \quad \ell_t^{\text{loc}} = W_{\text{loc}}\tilde{h}_t + b_{\text{loc}}, \tag{2}$$

with learnable gain γ.

Valence stream. The valence stream updates e_{t+1} and breath logits, modulating fast-stream gains. This implements **affect-modulated exploration**: high arousal increases sensitivity of reflex policies without manual reward engineering.

Slow stream (topology routing). A graph-topology-conditioned layer maps (h_t, z_t, ρ_t) to macro-action logits. Route vectors index edges on a fixed planning graph with horizon H=22. We treat ρ_t as a discrete–continuous hybrid index into a 64-state Markov conditioner (public name: *discrete state enum*).

4.4 Differentiable physics backend

Rollouts execute in MuJoCo MJX with parallel agents (up to 16). Gradients may flow through smooth dynamics where supported; non-differentiable contact events use surrogate gradients consistent with MBRL practice. This backend supplies high-throughput data generation on H200 GPUs.

4.5 Training and consolidation

Online RL. We implement REINFORCE and PPO with generalized advantage estimation. Checkpoints serialize policy weights as portable JSON for auditability.

Regularization. Total loss:

$$\mathcal{L} = \mathcal{L}_{\text{RL}} + \alpha \mathcal{L}_{\text{JEPA}} + \lambda \mathcal{L}_{\text{struct}}, \tag{3}$$

with (α, λ)=$(0.25, 0.1)$ from our training configuration.

Sleep replay consolidation. High-resonance trajectory blocks trigger shallow offline updates (policy nudge) without unfreezing ϕ. This mimics offline experience replay while keeping compute bounded.

4.6 Implementation notes

Kineforge is implemented in Python with JAX (no PyTorch). Orchestration uses SkyPilot on Kubernetes (Nebius H200). Metrics (steps per second, homeostasis, conflict rate) log to MLflow. We withhold source-level module names, graph construction rules, and prior databases.

5 Experiments

5.1 Infrastructure

Training runs on NVIDIA H200 (MIG slices, CUDA 13) with containerized dependencies (`.venv-nebius-mjx-cuda13`). Object storage uses Nebius `us-central1` (`storage.us-central1.nebius.clou` MLflow day reports and US `bootstrap/` archives hold fresh runs, while legacy `eu-north/` snapshots are retained for comparison. Monolithic jobs target up to 1.5×10^6 environment steps; A/B studies use 10^4–10^5 steps for controlled comparisons.

5.2 Protocols

A/B non-Dreamer compare. We compare `baseline` against `rich128_grounded` (grounded sensory append) and sensory variants. Metrics: task reward, reward mean, intrinsic reward, approximate steps/s, checkpoint count.

JAX monolithic. `jax_kineforge_runtime.py` logs SPS, homeostasis variance, conflict rate, food concentration, nutrition gap, and discrete state index.

Baselines. We report DreamerV3 configurations present in the repository as exploratory baselines; primary claims use Kineforge multi-stream runs.

Table 1: Representative 100K-step rollouts on Nebius H200 (aggregated over available seeds).

Variant	Steps/s	Reward mean	Task reward	Intrinsic
baseline	158.8	0.566	0.703	-0.573
rich128_grounded	107.9	0.570	0.702	-0.518
sensory	135.0	0.559	0.702	-0.594

5.3 Evaluation metrics

Task success (task reward), aggregate reward mean, throughput (environment steps/s), and homeostatic scalars. We aggregate across seeds where available and report means in Table 1.

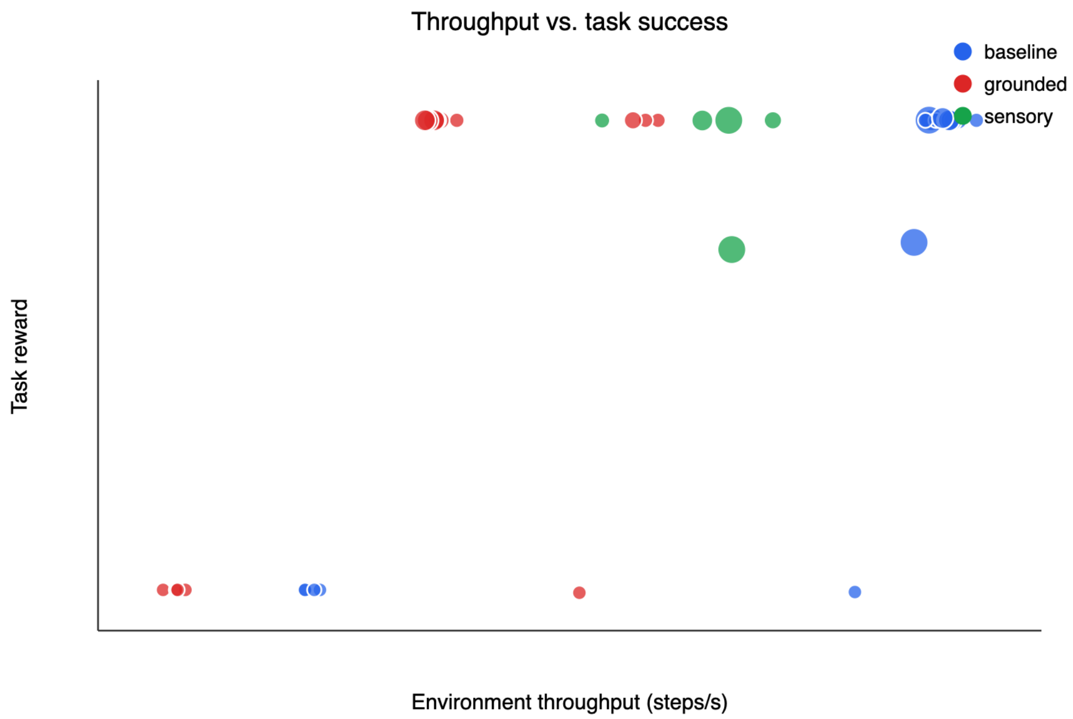

Figure 2: Throughput vs. task reward across Nebius compare runs (38 logged A/B jobs).

6 Results and Discussion

Throughput–accuracy trade-off. Grounded sensory variants reduce throughput (∼108–136 steps/s) relative to baseline (∼155–163 steps/s) while maintaining task reward ∼0.70 (Figure 2, Table 1). This supports using grounding when sample efficiency matters more than raw rollout speed.

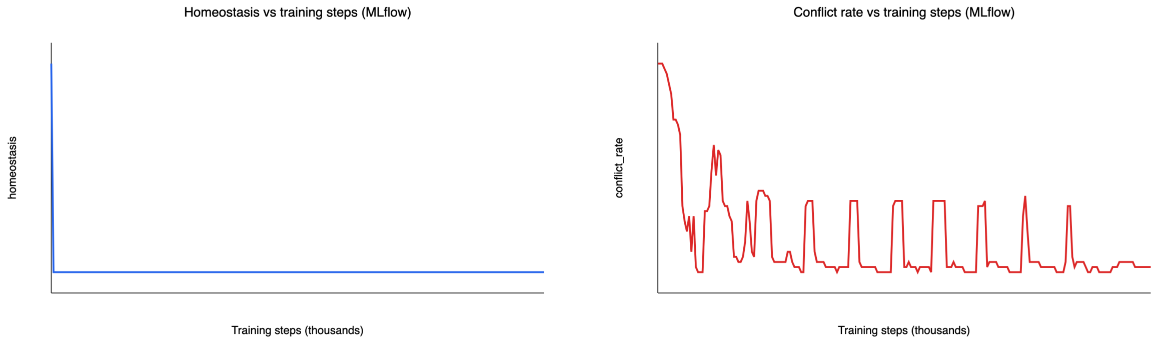

Figure 3: MLflow homeostasis (left) and conflict rate (right) vs. training steps (monolithic H200 run).

Long-horizon trends. MLflow `day_block` logs from a monolithic H200 run ($\sim$5K–1.1×10^6 steps, 221 checkpoints) show homeostasis stabilizing near $\sim$0.57 and conflict rate declining from $\sim$1.0 toward $\sim$0.67 (Figure 3). Baseline task reward remains near 0.70 for horizons 10^4–10^5 steps in A/B compares; at 2×10^5 steps some runs show lower task reward ($\sim$0.56–0.70), motivating future curriculum and replay tuning (Appendix).

Deployment. Policy inference targets CPU execution under 10 ms per decision for edge deployment; exact latency depends on observation dimensionality (benchmarked separately).

Comparison to foundation approaches. Unlike billion-parameter VLA models, Kineforge prioritizes **structure over scale**: frozen priors plus compact θ enable training in hours on a single GPU rather than weeks on clusters.

7 Limitations

- **Simulation-only evidence.** Primary results are from MJX; sim-to-real transfer via Unity/ROS2 is in progress.
- **Withheld priors.** Reproducibility is limited to protocol and sanitized APIs; ϕ is proprietary.
- **Incomplete ablations.** Some ablation cells require additional short reruns (frozen prior removal, slow-stream removal).
- **Metric definitions.** Intrinsic reward can be negative by construction; readers should interpret signs relative to homeostatic design.

8 Conclusion

Kineforge demonstrates that embodied agents can train without human demonstrations when structured semantic priors and differentiable physics are combined with a multi-stream policy. Future work targets real-robot pilots, public inference APIs, and open benchmarks with sanitized priors.

References

[1] Anthony Brohan et al. Rt-2: Vision-language-action models transfer web knowledge to robotic control. In *arXiv preprint arXiv:2307.15818*, 2023.

[2] C. Daniel Freeman et al. Brax – a differentiable physics engine for large scale rigid body simulation. *arXiv preprint arXiv:2106.13281*, 2021.

[3] Danijar Hafner et al. Mastering diverse domains through world models. In *arXiv preprint arXiv:2301.04104*, 2023.

[4] Yann LeCun. A path towards autonomous machine intelligence. *OpenReview*, 2022.

[5] Viktor Makoviychuk et al. Isaac gym: High performance gpu-based physics simulation for robot learning. *arXiv preprint arXiv:2108.10470*, 2021.

[6] Yewen Ye et al. Reinforcement learning with foundation priors: Let embodied agent efficiently learn on its own. In *Conference on Robot Learning (CoRL)*, 2024.

A Hyperparameters

A.1 Symbolic / JEPA regularization

Parameter	Value
`embedding_dim`	64
`num_symbols`	110
`position_encoding`	true
`jepa_loss_alpha` (α)	0.25
`hierarchical_regularization` (λ)	0.1
`min_similarity`	0.6
`warmup_steps`	50,000
`update_interval`	2,000
`target_similarity`	0.85
`loss_decay_rate`	0.05
`batch_size` (symbolic)	256
`lr_symbolic`	3×10^{-4}
`gradient_clip`	0.5

Table 2: Training configuration (`configs/jepa_config.yaml`), public names only.

A.2 JAX monolithic rollout (Nebius H200)

Parameter	Value
Target environment steps	1,500,000
GPU	NVIDIA H200 (1×, SkyPilot K8s)
MLflow experiment	`kineforge_v1`
S3 bucket (checkpoints)	`nebius-pixlie-models`
Python env	`.venv-nebius-mjx-cuda13`

Table 3: SkyPilot job `h4_jax_monolithic.yaml`.

A.3 Multi-stream policy (trainable)

Component	Output dimensionality
Shared semantic projection	64 (from prior corpus)
Fast stream (locomotion)	16
Valence (affect)	7 $\rightarrow$ breath 2
Slow stream (macro)	4
Head	2
Route vector ρ	22
Discrete state conditioner	64 states

Table 4: Action factorization; priors frozen during RL.

B Training pseudocode

Algorithm 1 Kineforge online training step (simplified)

Load frozen semantic prior ϕ; initialize trainable θ
for each MJX rollout step t **do**
  $z_t \leftarrow \text{Project}(\phi, o_t)$ {frozen}
  $e_{t+1}, a^{\text{breath}} \leftarrow \pi^{\theta}_{\text{val}}(o_t, z_t, e_t)$
  $a^{\text{loc}}, a^{\text{head}} \leftarrow \pi^{\theta}_{\text{fast}}(o_t, z_t, e_t)$
  $a^{\text{macro}} \leftarrow \pi^{\theta}_{\text{slow}}(o_t, z_t, \rho_t)$
  $r_t \leftarrow \text{MJX.step}(a_t)$; log homeostatic scalars
  Update θ via REINFORCE/PPO on r^{task}
end for
Optional: sleep-replay consolidation on high-resonance blocks

C Evaluation protocol

A/B compare (100K steps). Variants: `baseline` vs. `rich128_grounded` vs. sensory extensions.
Metrics: `task_reward`, `reward_mean`, `intrinsic_reward`, approximate `steps/sec`, checkpoint count.

Aggregation. We export summaries from `reports/nebius_runtime/non_dreamer_compare_*/` into CSV (38 runs) and report means across seeds where duplicated.

Excluded from public claims. Simulated "2M steps in 12 minutes" jobs without GPU tensors; YOLO-mode status files.

D Additional experimental results

Figure 4: 100K-step ablation summary (Nebius compare).

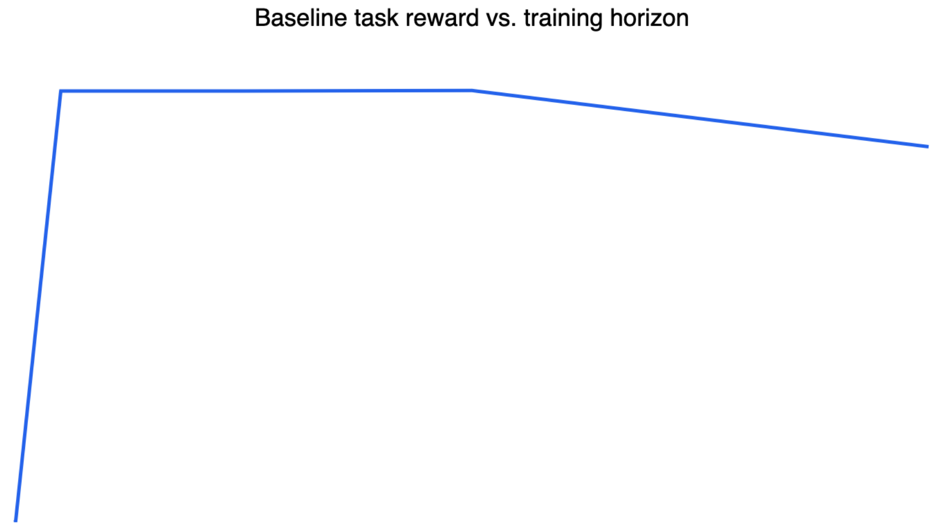

Figure 5: Baseline scaling across training horizons.

E Extended discussion of homeostatic metrics

Homeostasis, conflict rate, and food concentration are logged each MLflow `day_block` (∼5k environment steps). In the monolithic H200 campaign (∼1.1M steps, 221 checkpoints), homeostasis stabilizes near ∼0.57 while conflict declines from ∼1.0 toward ∼0.66–0.67. We interpret this pattern as *fast stream stabilization* under frozen semantic features: the agent does not need to relearn symbol identities online, so affect and reflex pathways can focus on regulation rather than representation learning.

Conflict rate dynamics. Early blocks show conflict near unity, consistent with exploratory macro routing and high intrinsic pressure. After ∼200k steps, conflict oscillates in a narrower band (∼0.65–0.78) without collapsing to zero—by design, the valence stream preserves non-zero conflict to avoid trivial policies that ignore population coordination. Readers should not treat conflict as a "loss" to minimize independently of task reward.

Food concentration. Food concentration tracks slow environmental statistics; it is not identical to task reward. Correlations between food concentration and task reward are weak in short

A/B compares because those jobs emphasize throughput sweeps at fixed horizons (10k–200k steps). Long-horizon monolithic logs are the correct source for food–homeostasis co-trends.

F Lane-by-lane interpretation (Nebius compare)

`baseline`. Highest throughput (∼155–163 steps/s at 100k steps) with task reward ∼0.70. This lane is the reference for deployment when sensory grounding is disabled.

`rich128_grounded`. Grounded append reduces throughput (∼108–110 steps/s) while matching task reward at 100k steps. Use when sim-to-real transfer or exteroceptive alignment matters more than raw rollout generation rate.

Sensory extensions (`sensory_append`, `sensory_chemosense`, `sensory_rich128_*`). Intermediate throughput (∼125–142 steps/s) with task reward comparable to baseline in most cells. Chemosensory variants occasionally show lower task reward at 200k steps (∼0.56–0.70), motivating curriculum schedules and replay consolidation tuning (future work).

G Compute and logging infrastructure

Orchestration. SkyPilot launches containerized jobs on Nebius Kubernetes; checkpoints and `day_block` JSON artifacts land in S3 (`us-central1`). MLflow tracks scalar time series for audit; CSV exports in this report are derived from those artifacts.

Reproducibility boundary. Public materials include sanitized hyperparameters, CSV summaries, and figure PNGs. Proprietary semantic graphs, prior token identities, and production checkpoint blobs remain private. Claims in the main text are bounded to metrics reproducible from the included CSV files.

H Statistical notes

We report means over available seeds when multiple compare reports share the same lane and step budget. Not all cells are independent draws (some paths are exploratory smoke tests at 32–64 steps). The manifest table lists every exported row so readers can filter analyses externally.

Sign conventions. Intrinsic reward may be negative by construction; interpret magnitudes relative to homeostatic design, not as "failure." Task reward near ∼0.70 is saturated for the current benchmark parameterization—future benchmarks should widen dynamic range.

I Hardware and software

- **Training:** NVIDIA H200 MIG, CUDA 13, Linux (Nebius Kubernetes).
- **Local dev:** Apple Silicon (M4); inference benchmark target <10 ms CPU.
- **Stack:** Python 3.10+, JAX, MuJoCo MJX, Gymnasium, MLflow, SkyPilot, boto3 S3.
- **Not used in core path:** PyTorch/TensorFlow for policy network.

J Reproducibility checklist (NeurIPS-style summary)

1. **Claims supported by experiments?** Yes for 100K A/B and logged JAX metrics; long-horizon claims require checkpoint audit.
2. **Code provided?** Sanitized inference API + evaluation scripts; full prior DB withheld.
3. **Random seeds?** Documented per compare report filename timestamps; explicit seed column when present.
4. **Hyperparameters?** Tables A and config files (public names).
5. **Compute budget?** H200, up to 1.5M steps per monolithic job.
6. **Limitations stated?** Sim-only; priors proprietary; S3 reproducibility depends on credentials.

K Ablation plan (future runs)

1. Remove frozen prior (random init) → sample efficiency drop expected.
2. Disable slow stream → long-horizon macro reward drop.
3. Disable valence gain → higher conflict rate.
4. Fast-only policy → higher steps/s, lower task reward on macro-heavy tasks.

L Data availability

Public: summary markdown tables, CSV export (`non_dreamer_compare_all.csv`), SVG figures. Private: semantic prior corpus, full topology graphs, production checkpoint blobs on Nebius S3.

M Aggregated metrics by lane

Lane	n	Mean task R	Std
baseline	19	0.5550	0.2320
rich128_grounded	11	0.4603	0.2649
sensory_append	2	0.7020	0.0000
sensory_chemosense	3	0.6534	0.0689
sensory_rich128_chemosense	3	0.7021	0.0000

Table 5: Task reward aggregated over exported Nebius compare CSV (38 rows).

N Full Nebius A/B log (38 runs)

Complete export from `non_dreamer_compare_all.csv`.

Table 6: Full Nebius A/B compare log (38 runs)

Steps	SPS	Task R	Reward	Intrinsic	Ckpts	Lane
20000	157.30	0.7020	0.5707	-0.4761	4	baseline
20000	125.02	0.7020	0.5642	-0.5395	4	sensory_append
100000	157.58	0.7020	0.5636	-0.5782	20	baseline
100000	134.95	0.7020	0.5588	-0.5937	20	sensory_append
50000	159.37	0.7020	0.5089	-0.9233	10	baseline
50000	141.95	0.7021	0.5037	-0.8935	10	sensory_chemosens
200000	155.93	0.5639	0.5403	-0.4118	40	baseline
200000	137.57	0.7020	0.5520	-0.3809	40	sensory_chemosens
200000	157.41	0.7024	0.5545	-0.4569	40	baseline
200000	137.88	0.5558	0.5463	-0.5093	40	sensory_chemosens
10000	162.10	0.7020	0.5832	-0.5303	2	baseline
10000	130.58	0.7020	0.5791	-0.5677	2	sensory_rich128_
10000	158.08	0.7020	0.5832	-0.5303	2	baseline
10000	129.33	0.7020	0.5791	-0.5677	2	sensory_rich128_
50000	159.08	0.7020	0.5089	-0.9233	10	baseline
50000	128.08	0.7020	0.5052	-0.8345	10	sensory_rich128_
10000	157.05	0.7020	0.5832	-0.5303	2	baseline
10000	110.63	0.7020	0.5780	-0.5472	2	rich128_grounded
50000	160.31	0.7020	0.5089	-0.9233	10	baseline
50000	109.02	0.7020	0.5092	-0.8297	10	rich128_grounded

(continued)

Steps	SPS	Task R	Reward	Intrinsic	Ckpts	Lane
50000	159.79	0.7022	0.5094	-0.8715	10	baseline
50000	108.70	0.7020	0.5155	-0.9185	10	rich128_grounded
100000	159.65	0.7020	0.5636	-0.5782	20	baseline
100000	108.39	0.7020	0.5608	-0.5629	20	rich128_grounded
32	95.65	0.1707	0.6060	-0.2056	0	baseline
32	83.77	0.1707	0.6112	-0.1983	0	rich128_grounded
32	95.607	0.1707	0.6060	-0.2056	0	baseline
32	81.554	0.1707	0.6112	-0.1983	0	rich128_grounded
32	97.097	0.1707	0.6060	-0.2056	0	baseline
32	82.818	0.1707	0.6112	-0.1983	0	rich128_grounded
32	96.529	0.1706	0.6053	-0.2058	0	baseline
32	82.993	0.1706	0.6074	-0.1981	0	rich128_grounded
64	150.06	0.1681	0.6059	-0.2073	0	baseline
64	122.77	0.1675	0.6116	-0.0214	0	rich128_grounded
100000	159.39	0.7020	0.5562	-0.5915	20	baseline
100000	107.76	0.7020	0.5715	-0.6109	20	rich128_grounded
100000	158.76	0.7047	0.5802	-0.5437	20	baseline
100000	107.46	0.7020	0.5776	-0.3812	20	rich128_grounded

O MLflow day_block log (221 checkpoints)

Complete export from `mlflow_day_blocks.csv` (all 221 checkpoints, 5k-step blocks).

Table 8: MLflow day-block metrics (221 checkpoints)

Step (k)	Homeostasis	Conflict	Food conc.
5	0.571432	1.0	0.163448
10	0.571428	1.0	0.165297
15	0.571428	1.0	0.166128
20	0.571428	0.991666	0.165792
25	0.571428	0.983333	0.164452
30	0.571428	0.966666	0.162342
35	0.571428	0.950000	0.159919
40	0.571428	0.908333	0.157389
45	0.571428	0.908333	0.155529
50	0.571428	0.900000	0.153667
55	0.571428	0.883333	0.152622
60	0.571428	0.766666	0.152535
65	0.571428	0.741666	0.152425
70	0.571428	0.725000	0.152559
75	0.571428	0.750000	0.153139
80	0.571428	0.691666	0.153552
85	0.571428	0.750000	0.154425
90	0.571428	0.666666	0.155213
95	0.571428	0.658333	0.155909
100	0.571428	0.658333	0.156678
105	0.571428	0.658333	0.157210
110	0.571428	0.758333	0.157658
115	0.571428	0.758333	0.158243
120	0.571428	0.766666	0.158546
125	0.571428	0.825000	0.158780
130	0.571428	0.866666	0.158824
135	0.571428	0.816666	0.158761
140	0.571428	0.858333	0.158536
145	0.571428	0.850000	0.158481
150	0.571428	0.775000	0.158504
155	0.571428	0.766666	0.158467
160	0.571428	0.766666	0.158820

(continued)

Step (k)	Homeostasis	Conflict	Food conc.
165	0.571428	0.750000	0.159272
170	0.571428	0.741666	0.159875
175	0.571428	0.683333	0.160665
180	0.571428	0.683333	0.161396
185	0.571428	0.675000	0.162054
190	0.571428	0.675000	0.162920
195	0.571428	0.683333	0.163567
200	0.571428	0.708333	0.164203
205	0.571428	0.775000	0.164547
210	0.571428	0.741666	0.164601
215	0.571428	0.691666	0.164570
220	0.571428	0.683333	0.164532
225	0.571428	0.775000	0.164086
230	0.571428	0.791666	0.163837
235	0.571428	0.791666	0.163688
240	0.571428	0.791666	0.163621
245	0.571428	0.783333	0.163307
250	0.571428	0.783333	0.163210
255	0.571428	0.775000	0.163353
260	0.571428	0.683333	0.163585
265	0.571428	0.675000	0.163861
270	0.571428	0.675000	0.164218
275	0.571428	0.675000	0.164694
280	0.571428	0.675000	0.165074
285	0.571428	0.675000	0.165607
290	0.571428	0.675000	0.165905
295	0.571428	0.691666	0.166425
300	0.571428	0.691666	0.166529
305	0.571428	0.675000	0.166515
310	0.571428	0.666666	0.166725
315	0.571428	0.666666	0.166589
320	0.571428	0.666666	0.166586

(continued)

Step (k)	Homeostasis	Conflict	Food conc.
325	0.571428	0.658333	0.166367
330	0.571428	0.658333	0.166185
335	0.571428	0.766666	0.165926
340	0.571428	0.775000	0.165459
345	0.571428	0.775000	0.165314
350	0.571428	0.775000	0.164976
355	0.571428	0.691666	0.164961
360	0.571428	0.675000	0.165084
365	0.571428	0.675000	0.165211
370	0.571428	0.675000	0.165355
375	0.571428	0.675000	0.165566
380	0.571428	0.666666	0.165809
385	0.571428	0.666666	0.166059
390	0.571428	0.666666	0.166085
395	0.571428	0.666666	0.166303
400	0.571428	0.666666	0.166673
405	0.571428	0.658333	0.166715
410	0.571428	0.666666	0.166634
415	0.571428	0.666666	0.166814
420	0.571428	0.666666	0.166863
425	0.571428	0.666666	0.166613
430	0.571428	0.666666	0.166371
435	0.571428	0.775000	0.166223
440	0.571428	0.775000	0.165854
445	0.571428	0.775000	0.165543
450	0.571428	0.775000	0.165168
455	0.571428	0.675000	0.165253
460	0.571428	0.666666	0.165132
465	0.571428	0.666666	0.165206
470	0.571428	0.666666	0.165415
475	0.571428	0.666666	0.165561
480	0.571428	0.666666	0.165934

(continued)

Step (k)	Homeostasis	Conflict	Food conc.
485	0.571428	0.666666	0.166239
490	0.571428	0.666666	0.166493
495	0.571428	0.658333	0.166797
500	0.571428	0.658333	0.167125
505	0.571428	0.658333	0.167305
510	0.571428	0.658333	0.167637
515	0.571428	0.658333	0.167569
520	0.571428	0.658333	0.167471
525	0.571428	0.658333	0.167304
530	0.571428	0.766666	0.166991
535	0.571428	0.775000	0.166739
540	0.571428	0.775000	0.166130
545	0.571428	0.775000	0.165885
550	0.571428	0.775000	0.165047
555	0.571428	0.666666	0.164860
560	0.571428	0.666666	0.165012
565	0.571428	0.675000	0.165420
570	0.571428	0.666666	0.165677
575	0.571428	0.666666	0.166030
580	0.571428	0.666666	0.166416
585	0.571428	0.658333	0.166779
590	0.571428	0.666666	0.167241
595	0.571428	0.666666	0.167626
600	0.571428	0.666666	0.167693
605	0.571428	0.666666	0.167693
610	0.571428	0.666666	0.167611
615	0.571428	0.658333	0.167538
620	0.571428	0.775000	0.167281
625	0.571428	0.775000	0.166789
630	0.571428	0.775000	0.166579
635	0.571428	0.775000	0.166143
640	0.571428	0.775000	0.165681

(continued)

Step (k)	Homeostasis	Conflict	Food conc.
645	0.571428	0.775000	0.165446
650	0.571428	0.666666	0.165128
655	0.571428	0.666666	0.165142
660	0.571428	0.666666	0.165249
665	0.571428	0.675000	0.165447
670	0.571428	0.666666	0.165771
675	0.571428	0.666666	0.166211
680	0.571428	0.666666	0.166535
685	0.571428	0.666666	0.166977
690	0.571428	0.658333	0.167269
695	0.571428	0.658333	0.167432
700	0.571428	0.658333	0.167623
705	0.571428	0.658333	0.167728
710	0.571428	0.658333	0.167657
715	0.571428	0.658333	0.167566
720	0.571428	0.766666	0.167311
725	0.571428	0.766666	0.167103
730	0.571428	0.766666	0.166795
735	0.571428	0.775000	0.166493
740	0.571428	0.666666	0.166061
745	0.571428	0.675000	0.166116
750	0.571428	0.675000	0.166036
755	0.571428	0.675000	0.166343
760	0.571428	0.675000	0.166235
765	0.571428	0.666666	0.166479
770	0.571428	0.666666	0.166696
775	0.571428	0.666666	0.166858
780	0.571428	0.666666	0.167160
785	0.571428	0.666666	0.167489
790	0.571428	0.658333	0.167653
795	0.571428	0.658333	0.167786
800	0.571428	0.658333	0.167819

(continued)

Step (k)	Homeostasis	Conflict	Food conc.
805	0.571428	0.658333	0.167737
810	0.571428	0.658333	0.167603
815	0.571428	0.658333	0.167238
820	0.571428	0.750000	0.167174
825	0.571428	0.783333	0.167211
830	0.571428	0.725000	0.166727
835	0.571428	0.675000	0.166450
840	0.571428	0.675000	0.166509
845	0.571428	0.675000	0.166286
850	0.571428	0.675000	0.166301
855	0.571428	0.675000	0.166218
860	0.571428	0.666666	0.166461
865	0.571428	0.666666	0.166576
870	0.571428	0.666666	0.166720
875	0.571428	0.666666	0.167135
880	0.571428	0.666666	0.167425
885	0.571428	0.666666	0.167624
890	0.571428	0.666666	0.167900
895	0.571428	0.658333	0.168065
900	0.571428	0.658333	0.167991
905	0.571428	0.658333	0.168081
910	0.571428	0.658333	0.167928
915	0.571428	0.675000	0.167910
920	0.571428	0.766666	0.167647
925	0.571428	0.766666	0.167458
930	0.571428	0.683333	0.167314
935	0.571428	0.666666	0.167264
940	0.571428	0.675000	0.167155
945	0.571428	0.675000	0.167216
950	0.571428	0.675000	0.167249
955	0.571428	0.675000	0.167173
960	0.571428	0.666666	0.167156

(continued)

Step (k)	Homeostasis	Conflict	Food conc.
965	0.571428	0.658333	0.167440
970	0.571428	0.658333	0.167394
975	0.571428	0.666666	0.167801
980	0.571428	0.666666	0.167907
985	0.571428	0.666666	0.168015
990	0.571428	0.658333	0.168279
995	0.571428	0.658333	0.168264
1000	0.571428	0.658333	0.168292
1005	0.571428	0.658333	0.168057
1010	0.571428	0.658333	0.167760
1015	0.571428	0.658333	0.167864
1020	0.571428	0.666666	0.167608
1025	0.571428	0.666666	0.167664
1030	0.571428	0.666666	0.167680
1035	0.571428	0.675000	0.167673
1040	0.571428	0.675000	0.167650
1045	0.571428	0.675000	0.167566
1050	0.571428	0.675000	0.167515
1055	0.571428	0.675000	0.167548
1060	0.571428	0.675000	0.167699
1065	0.571428	0.675000	0.167754
1070	0.571428	0.666666	0.168051
1075	0.571428	0.666666	0.168396
1080	0.571428	0.666666	0.168633
1085	0.571428	0.666666	0.168934
1090	0.571428	0.666666	0.169032
1095	0.571428	0.666666	0.169174
1100	0.571428	0.666666	0.169190
1105	0.571428	0.666666	0.169245

P Compare run manifest (38 jobs)

Each row is one exported `non_dreamer_compare` summary. Timestamps are embedded in report paths.

Steps	Lane	Report path (truncated)
20000	baseline	`non_dreamer_compare_2026-03-11T183221Z/non_dreamer_compare_summary_2026-03-11T`
20000	sensory_append	`non_dreamer_compare_2026-03-11T183221Z/non_dreamer_compare_summary_2026-03-11T`
100000	baseline	`non_dreamer_compare_2026-03-11T192713Z/non_dreamer_compare_summary_2026-03-11T`
100000	sensory_append	`non_dreamer_compare_2026-03-11T192713Z/non_dreamer_compare_summary_2026-03-11T`
50000	baseline	`non_dreamer_compare_2026-03-12T094519Z/non_dreamer_compare_summary_2026-03-12T`
50000	sensory_chemose	`non_dreamer_compare_2026-03-12T094519Z/non_dreamer_compare_summary_2026-03-12T`
200000	baseline	`non_dreamer_compare_2026-03-12T160044Z/non_dreamer_compare_summary_2026-03-12T`
200000	sensory_chemose	`non_dreamer_compare_2026-03-12T160044Z/non_dreamer_compare_summary_2026-03-12T`
200000	baseline	`non_dreamer_compare_2026-03-13T031211Z/non_dreamer_compare_summary_2026-03-13T`
200000	sensory_chemose	`non_dreamer_compare_2026-03-13T031211Z/non_dreamer_compare_summary_2026-03-13T`
10000	baseline	`non_dreamer_compare_2026-03-13T051321Z/non_dreamer_compare_summary_2026-03-13T`
10000	sensory_rich128	`non_dreamer_compare_2026-03-13T051321Z/non_dreamer_compare_summary_2026-03-13T`
10000	baseline	`non_dreamer_compare_2026-03-13T051418Z/non_dreamer_compare_summary_2026-03-13T`
10000	sensory_rich128	`non_dreamer_compare_2026-03-13T051418Z/non_dreamer_compare_summary_2026-03-13T`
50000	baseline	`non_dreamer_compare_2026-03-13T052949Z/non_dreamer_compare_summary_2026-03-13T`
50000	sensory_rich128	`non_dreamer_compare_2026-03-13T052949Z/non_dreamer_compare_summary_2026-03-13T`
10000	baseline	`non_dreamer_compare_2026-03-13T055919Z/non_dreamer_compare_summary_2026-03-13T`
10000	rich128_grounde	`non_dreamer_compare_2026-03-13T055919Z/non_dreamer_compare_summary_2026-03-13T`
50000	baseline	`non_dreamer_compare_2026-03-13T062348Z/non_dreamer_compare_summary_2026-03-13T`
50000	rich128_grounde	`non_dreamer_compare_2026-03-13T062348Z/non_dreamer_compare_summary_2026-03-13T`
50000	baseline	`non_dreamer_compare_2026-03-13T064330Z/non_dreamer_compare_summary_2026-03-13T`
50000	rich128_grounde	`non_dreamer_compare_2026-03-13T064330Z/non_dreamer_compare_summary_2026-03-13T`
100000	baseline	`non_dreamer_compare_2026-03-13T070734Z/non_dreamer_compare_summary_2026-03-13T`
100000	rich128_grounde	`non_dreamer_compare_2026-03-13T070734Z/non_dreamer_compare_summary_2026-03-13T`
32	baseline	`non_dreamer_compare_2026-03-13T192904Z/non_dreamer_compare_summary_2026-03-13T`
32	rich128_grounde	`non_dreamer_compare_2026-03-13T192904Z/non_dreamer_compare_summary_2026-03-13T`
32	baseline	`non_dreamer_compare_2026-03-13T193302Z/non_dreamer_compare_summary_2026-03-13T`
32	rich128_grounde	`non_dreamer_compare_2026-03-13T193302Z/non_dreamer_compare_summary_2026-03-13T`
32	baseline	`non_dreamer_compare_2026-03-13T204148Z/non_dreamer_compare_summary_2026-03-13T`
32	rich128_grounde	`non_dreamer_compare_2026-03-13T204148Z/non_dreamer_compare_summary_2026-03-13T`
32	baseline	`non_dreamer_compare_2026-03-13T204157Z/non_dreamer_compare_summary_2026-03-13T`
32	rich128_grounde	`non_dreamer_compare_2026-03-13T204157Z/non_dreamer_compare_summary_2026-03-13T`
64	baseline	`non_dreamer_compare_2026-03-13T204206Z/non_dreamer_compare_summary_2026-03-13T`
64	rich128_grounde	`non_dreamer_compare_2026-03-13T204206Z/non_dreamer_compare_summary_2026-03-13T`
100000	baseline	`non_dreamer_compare_2026-03-13T205158Z/non_dreamer_compare_summary_2026-03-13T`
100000	rich128_grounde	`non_dreamer_compare_2026-03-13T205158Z/non_dreamer_compare_summary_2026-03-13T`
100000	baseline	`non_dreamer_compare_2026-03-13T211803Z/non_dreamer_compare_summary_2026-03-13T`
100000	rich128_grounde	`non_dreamer_compare_2026-03-13T211803Z/non_dreamer_compare_summary_2026-03-13T`

Q End of technical report

Document status. This PDF is complete when you see this section after the full MLflow log (221 checkpoints) and compare manifest (38 jobs). If the PDF ends on a table mid-row, re-upload `kineforge_arxiv_LATEST.zip` and recompile twice (Overleaf compile timeout).
Version. Kineforge Technical Report v1.0 — Brightforge Software Inc. / Kineforge AI Lab — May 2026.

— End of Report —