Trigger-Scoped QuadOrbit Adaptation of SmolLM2-135M

Argo1-OOAS · goausd@gmail.com · August 26, 2026

Abstract—We present a trigger-scoped bounded quadratic complex QuadOrbit branch attached after a frozen HuggingFaceTB/SmolLM2-135M-Instruct decoder. The ordinary Transformer computes every contextual row unchanged. At selected trigger rows, the branch reads the causal hidden trajectory, optionally conditions on user text, and adds a residual before the frozen language-model head; because row t predicts token t+1, the intervention affects the next-token distribution. Only 424,641 branch parameters (0.315% of 134,939,649 total) were optimized. Training completed 2,000 updates and 262.144 million token presentations on role-formatted SmolTalk. The terminal ten-batch estimate was branch-objective validation loss 1.6958 (perplexity 5.5). This is not ordinary full-sequence perplexity and does not establish an improvement over the frozen base. We release complete equations, masking and context semantics, objective, provenance, FP32 Safetensors, hashes, and a token-level counterfactual audit interface.

1. INTRODUCTION

Pretrained causal Transformers already encode a rich history in each contextual row. This study asks a narrower question than replacing attention: can a separately trained trajectory mechanism remain dormant on the normal path, become active only under an explicit runtime trigger, and use optional text to alter a small number of subsequent decisions without fine-tuning the base model?

The released SmolLM2 V2 QuadOrbit implements that experiment. Its contributions are: (i) a mathematically specified bounded quadratic complex QuadOrbit over frozen contextual rows; (ii) exact trigger semantics tied to autoregressive row indexing; (iii) branch-only future-context and wrong-context training; (iv) inference-only guidance with a branch-off counterfactual; and (v) complete artifact provenance. The architecture is intentionally hybrid: attention and FFN remain responsible for ordinary language modeling.

We make no claim that a trigger creates deliberation or that latent "travel" is physical motion. The branch is a learned residual. Its scientific value depends on matched evaluations that separate distributional changes from improved outputs.

2. FROZEN DECODER CORE

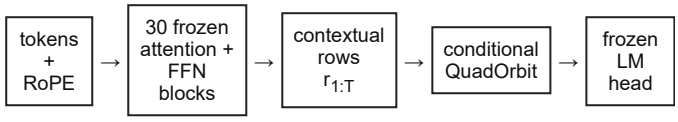
Let token IDs be $x_{1:T}$ and let E be the learned token table. The unchanged causal decoder begins with $h_t^{(0)}=E[x_t]$. For layer ℓ, pre-normalized grouped-query attention is

$$q_t=\mathrm{RoPE}(N(h_t)W_q),\quad k_t=\mathrm{RoPE}(N(h_t)W_k),\quad v_t=N(h_t)W_v,$$
$$A_{tj}=\mathrm{softmax}_{j\le t}(q_t k_j^{T}/\sqrt{d_h}),\quad a_t=(\Sigma_{j\le t}A_{tj}v_j)W_o.$$

With $u=h+a$, the SwiGLU stage is

$$f_t=[\mathrm{SiLU}(N(u_t)W_g)\odot(N(u_t)W_u)]W_d,\quad h_t^{(\ell+1)}=u_t+f_t.$$

Here N is RMSNorm. The released core has 30 layers, width 576, FFN width 1536, 9 query / 3 key-value heads, and RoPE base 100,000. The core was imported from HuggingFaceTB/SmolLM2-135M-Instruct and frozen during the reported branch-only continuation.



3. QUADORBIT BRANCH

For normalized frozen row $b_t=N(r_t)$, the 32-coordinate complex branch forms a bounded drive and input gate:

$$[d_t^{R};d_t^{I}]=\tanh(W_c b_t),\quad d_t=d_t^{R}+i\,d_t^{I},\quad g_t=\sigma(W_g b_t+b_g).$$

Each coordinate has a constrained decay, phase, and quadratic strength

$$\rho=\exp[-\mathrm{softplus}(\delta)]\in(0,1),\quad \omega=\rho e^{i\theta},$$
$$\beta=0.5\sigma(b_\beta)\in(0,0.5).$$

Starting from $z_0=0$, the complete recurrence is

$$a_t=\omega\odot z_{t-1},\quad q_t=\beta\odot {z_{t-1}}^2+d_t,\quad \hat{q}_t=q_t/\sqrt{(1+|q_t|^2)},$$
$$z_t=(1-g_t)\odot a_t+g_t\odot\hat{q}_t,\quad \Delta_t=sW_o[\mathrm{Re}(z_t);\mathrm{Im}(z_t)].$$

The learned branch scale is $s=\sigma(s_{\mathrm{raw}})$ because its configured maximum is one. It finished at 0.283890. The recurrent state is evaluated over the complete causal prefix, but its residual is exposed to the vocabulary head only where a trigger mask is active.

Proposition (forward boundedness). If $|z_0|<1$ coordinatewise, then $|z_t|<1$ for every finite t. Indeed, $|a_t|=\rho|z_{t-1}|<1$ and $|\hat{q}_t|=|q_t|/\sqrt{(1+|q_t|^2)}<1$. Since $0<g_t<1$, z_t is a coordinatewise convex mixture of two points in the open unit disk. This controls state magnitude; it does not prove long memory or language quality.

Initialization and computation

The 32 initial timescales are log-spaced from 4 to 8,192 positions. Gate probabilities begin at τ^{-1}, retention begins at $\exp(-1/(4\tau))$, β begins at 0.1, and the output projection originally begins at zero. The branch costs $O(Tm)$ recurrent arithmetic after $O(Tdm)$ projections, where $m=32$. Its recurrence is sequential in T; the frozen Transformer still has quadratic causal-attention training cost.

4. OPTIONAL CONTEXT CONDITIONING

Let optional text tokenize to IDs $c_{1:H}$. Its detached embedding mean is normalized, $\bar{c}=N(H^{-1}\Sigma E[c_j])$. At selected rows, the branch input is contextualized by

$$r_t^c=r_t+m_tW_{cxt}\bar{c},\quad m_t\in\{0,1\}.$$

The branch computes $\Delta(r^c)$ rather than injecting vocabulary logits directly. At inference, guidance λ magnifies only the learned context-dependent difference:

$$\Delta_\lambda=\Delta(r)+\lambda[\Delta(r^c)-\Delta(r)],\quad \lambda\geq 0.$$

Thus $\lambda=0$ is the learned no-context branch and $\lambda=1$ is the trained conditional branch. The UI allows larger values for experimentation, but values above the training regime are extrapolation and may destroy fluency.

5. EXACT TRIGGER SEMANTICS

Let $\mathcal{T}$ be the runtime set of token patterns and let $m_t=1$ when the generated prefix ends in a selected pattern. The final row and next-token distribution are

$$\tilde{u}_t=r_t+m_t\Delta_{\lambda,t},\quad \ell_t=W_{LM}N(\tilde{u}_t),$$
$$p(x_{t+1}|x_{\leq t})=\mathrm{softmax}(\ell_t).$$

A trigger at token x_t therefore changes the distribution used to choose exactly x_{t+1}. The browser may deliberately force m=1 for a user-selected number of later decoding steps; that is an inference policy, not a different trained recurrence. Prompting never updates saved weights. The audit table recomputes the same row with m=0 and reports the counterfactual base choice and probability.

6. BRANCH-ONLY OBJECTIVE

The Transformer core, final RMSNorm and language head are detached and frozen. For training row t, the context lesson is the normalized mean of the next H=8 target-token embeddings. A lesson mask combines actual configured trigger IDs with one random offset every S=128 rows, allowing users to choose other runtime triggers after training.

$$m_t^{\text{train}}=1[x_t\in\mathcal{T}]\vee 1[t\equiv o\ (\mathrm{mod}\ 128)],$$
$$o\sim\mathrm{Uniform}\{0,\ldots,127\}.$$

For selected rows, p_c is the frozen head probability after the correctly paired future-context branch. The primary loss is

$$L_{CE}=-|S|^{-1}\Sigma_{t\in S}\log p_c(y_t|x_{\leq t}).$$

Contexts are rolled across selected examples to make mismatched controls p_w. The implemented contrast is

$$L_{ctr}=\mathrm{mean}_{t\in S}\,\mathrm{softplus}(\log p_w(y_t)-\log p_c(y_t)),$$
$$L=L_{CE}+0.5L_{ctr}.$$

This teaches a context-sensitive output branch without changing the base model. It does not train the model to quote the context, guarantee semantic adherence, or reason over it.

7. DATA AND OPTIMIZATION

Table 1. Exact continuation recipe.

Item	Value
Dataset	HuggingFaceTB/smol-smoltalk; role-formatted conversations
Prepared tokens	420,314,292 train; 838,170 validation
Sequence / token batch	8,192 / 131,072 tokens per optimizer update
Updates / presentations	2,000 / 262,144,000
Optimizer	AdamW; β_1=0.9, β_2=0.95, $\varepsilon=10^{-8}$, weight decay 0.1
Schedule	100-step warmup; cosine 2×10^{-4} to 2×10^{-5}
Precision / clipping	BF16 autocast; gradient norm clipped at 1.0
Seed	2026

The checkpoint is saved at zero-indexed step 1,999, meaning 2,000 completed updates. Only parameters whose names begin with `output_branch.` enter the optimizer.

8. ARTIFACT AND RECORDED RESULT

Table 2. Released model and terminal estimate.

Quantity	Value
Frozen core parameters	134,515,008
Trainable branch parameters	424,641 (0.315%)
Total unique parameters	134,939,649
Vocabulary / tokenizer	49,152 / SmolLM2 / Cosmo2-compatible BPE
Terminal validation estimate	loss 1.6958; perplexity 5.5
Final learned branch strength	0.283890
Safetensors bytes	653,033,276

The validation figure printed by the training loop averages ten sampled validation batches under the selected branch-training objective, including its context lesson. It is **not** ordinary all-token cross-entropy for the base model. Consequently it cannot be compared directly with published SmolLM2/TinyLlama perplexity or with another branch that uses a different lesson mask. No unrounded deterministic full-stream branch evaluation was retained.

> **Evidence level.** This release demonstrates that the branch can be optimized while preserving a frozen pretrained core. It does not establish an improvement in general language quality, factuality, reasoning, context following, speed, safety, or parameter efficiency. A matched branch-off evaluation, random adapter control, several seeds, human preference study and standard benchmarks remain required.

9. INFERENCE AND REPRODUCIBILITY

Sampling applies temperature and top-k after branch-guided logits. The UI records, for every affected row, the Transformer-only argmax, actual sampled token, base probability, guided probability, and whether the choice changed. This distinguishes a small probability perturbation from a changed discrete choice.

The public artifact stores FP32 inference weights in Safetensors and excludes optimizer, scaler and RNG state. Its SHA-256 identity is:

`14687d97ff1ead623ba32af1cc764b7675649ca7da60fd668374214214953449`

Exact configuration, tokenizer files, training metadata, release manifest and SHA256SUMS accompany the PDF. The source implementation loads either this safe release or the original native checkpoint.

10. LIMITATIONS AND REQUIRED CONTROLS

- **Single artifact.** The report contains one seed and one terminal sampled estimate; uncertainty is unknown.
- **Frozen-base specialization.** A low branch loss can coexist with unchanged or worse free-running text.
- **Trigger distribution shift.** General-position lessons make triggers configurable, but do not ensure equal behavior for arbitrary words, punctuation or multi-token patterns.
- **Guidance extrapolation.** The UI range extends to 50 although training corresponds to λ=1. Large λ can copy hints, cause repetition or erase grammar.
- **No causal claim.** A changed token proves sensitivity to the branch, not that the nonlinear complex recurrence supplied correct reasoning.
- **Runtime.** The audit path recomputes prefixes and does not expose an optimized KV cache. Throughput is not compared.
- **Safety.** The model can hallucinate and was not preference-optimized or evaluated for bias, toxicity, privacy, medical, legal or security use.

The minimum stronger experiment is a 2×2 evaluation of branch off/on and context absent/present, with the same histories teacher-forced through all conditions. Report next-token cross-entropy at triggered and ordinary positions, KL divergence, choice-change rate, repetition, context semantic alignment, fluency, latency and at least three seeds. For V2, ablate β=0, rotation, and complex state. For V3, ablate the edge difference, absolute cumulative travel, and context projection.

11. RELATED WORK

The frozen decoder follows Transformer attention [1], RoPE [2], GQA [3], RMSNorm [4] and SwiGLU [5]. Triggered residual adaptation is related broadly to parameter-efficient adapters, although this implementation trains a custom output branch rather than LoRA. V2 draws on complex gated recurrence and bounded radial maps [6]. V3 is closest in spirit to recurrent/state-space summaries [7], but uses a learned cumulative absolute difference rather than a linear state-space kernel.

12. CONCLUSION

SmolLM2 V2 QuadOrbit is a reproducible conditional-output experiment, not a replacement for the base Transformer. The core stays frozen; a nonlinear complex recurrence reads its contextual trajectory and is activated only by an explicit mask. The release provides complete equations, exact parameter/training provenance and token-level counterfactual auditing. The reported loss is encouraging only as evidence that the private objective optimized. Controlled evaluation is necessary before making any quality or efficiency claim.

REFERENCES

[1] A. Vaswani et al., “Attention Is All You Need,” NeurIPS, 2017.

[2] J. Su et al., “RoFormer: Enhanced Transformer with Rotary Position Embedding,” Neurocomputing, 2024.

[3] J. Ainslie et al., “GQA: Training Generalized Multi-Query Transformer Models,” EMNLP, 2023.

[4] B. Zhang and R. Sennrich, “Root Mean Square Layer Normalization,” NeurIPS, 2019.

[5] N. Shazeer, “GLU Variants Improve Transformer,” arXiv:2002.05202, 2020.

[6] M. Wolter and A. Yao, “Complex Gated Recurrent Neural Networks,” NeurIPS, 2018.

[7] A. Gu et al., “Efficiently Modeling Long Sequences with Structured State Spaces,” ICLR, 2022.

[8] Hugging Face, “SmolLM2 and SmolTalk model/data releases,” 2024–2025.