Historical FX experiment

Radix, neighbor phase and exact clock readouts

Four-wire cyclic features were added to a fixed TokenGT forecast of one-second forward and reverse FX triangle margins. We varied the wire dimension ($q = 2, 3, 5$) and switched a nearest-neighbor phase on or off. The two-level circuit with the phase achieved the lowest error on the reused October test cohort. A separately sealed November cohort reproduced the October directions at all three radices after averaging the two seeds. At every radix, however, the seed-specific November phase effects had opposite signs. An exact reduced-state calculation also lowered the cost of computing the four observed clocks.

A matched comparison on archived quotes

Each model uses 16 one-second frames, two spatial blocks, one temporal block, width 200, eight heads and softmax attention. Six gauge-centered directed rates at each frame enter a $6 \to 4$ affine map followed by $\pi \tanh$. Four circuit observables then enter a $4 \to 200$ projection. Its result is added to the pooled spatial state. The product arm omits only the diagonal bond phase. Its three trainable couplings still condition local mixers on neighboring input angles. All twelve fits have 1,537,657 parameters, including 1,031 in the patch, and the same initial parameter values within seed.

We used 8,192 August training windows, batch size 64, three epochs and smooth-L_1 loss with $\beta = 0.25$. Each cell selected its minimum-validation-MAE checkpoint. Seeds 20261225 and 20261226 share the same epoch order across arms. The score averages absolute errors over the two route directions and two fitted seeds. Each historical assessment column contains 2,048 windows.

q	Bond phase	Validation	Stress	October	New CHF assets
2	Product	0.19670	0.18574	0.18564	0.20747
2	Entangling	0.19368	0.18207	0.18178	0.20606
3	Product	0.19767	0.18645	0.18708	0.20621
3	Entangling	0.19936	0.18714	0.18899	0.20436
5	Product	0.20130	0.19071	0.19037	0.21048
5	Entangling	0.20716	0.19601	0.19699	0.22284
	Persistence	0.21958	0.20735	0.20643	0.21852

Mean absolute error in log basis points, lower is better. Persistence carries the current observed margins forward one second. The four historical cohorts have previously been examined.

Paired October contrasts. Entangling minus product MAE was -0.00386 at $q = 2$, $+0.00191$ at $q = 3$ and $+0.00661$ at $q = 5$. Individual 95% UTC-day intervals follow below. On the CHF cohort the $q = 3$ direction reversed, with phase reducing MAE by 0.00185. At $q = 5$ the CHF phase arm also exceeded persistence. These changes concern short-horizon margin prediction under the archived target definition.

Circuit calculation

Deriving the four clock expectations locally

For input angles α_j and level r, write $C(r) = \cos(2\pi r/q)$, $p_j(r) \propto \exp[\cos(\alpha_j - 2\pi r/q)]$, and $v_j(r) = \sqrt{p_j(r)}e^{i\alpha_j C(r)}$. The initial pure state is the product of the v_j. Each active bond $j, j+1$ applies phase $\exp(i\gamma_j r_j r_{j+1})$ with $\gamma_j = 2\pi c_j/q$. The angle-dependent mixing step is a tensor product of unitaries

$$U_j = I + (e^{i\theta_j} - 1)|s\rangle\langle s|, \qquad |s\rangle = q^{-1/2}\sum_{r=0}^{q-1}|r\rangle.$$

To obtain a single clock, trace the other three wires out before mixing. Every diagonal phase on an edge disjoint from j cancels between bra and ket. Define $\phi_k(t) = \sum_u p_k(u)e^{itu}$. With γ_{jk} denoting the coefficient of the bond connecting j and k, the exact answer is

$$\rho_j(r,s) = v_j(r)\overline{v_j(s)}\prod_{k\sim j}\phi_k(\gamma_{jk}(r-s)), \qquad m_j = \sum_r C(r)(U_j\rho_j U_j^\dagger)_{rr}.$$

In the product arm the neighbor factors equal one. Entanglement can reduce $\mathrm{tr}(\rho_j^2)$ below one while m_j remains exactly computable. The full joint distribution can contain correlations that these four expectations omit. The identity uses a product preparation, diagonal commuting bond layer, then local mixing and one-wire observables. Interleaved entangling layers would require a new derivation.

Measured patch cost. Six sealed checkpoints at seed 20261225 used the first 64 windows of the registered train sample, giving 1,024 observed frames. Each timing is the median of 20 synchronized trials after six warmups on an RTX 4090 with PyTorch 2.8.0. The backward pass differentiates the mean of the four clocks with respect to input angles and the three coupling parameters.

		Forward (ms)		Forward + back (ms)		Peak extra (MiB)	
q	Bond phase	Dense	Local	Dense	Local	Dense	Local
2	Product	0.290	0.164	1.225	0.681	1.77	0.99
2	Entangling	0.318	0.276	1.284	1.151	1.77	1.55
3	Product	0.294	0.166	1.243	0.729	8.05	2.00
3	Entangling	0.312	0.278	1.319	1.155	8.05	3.22
5	Product	0.292	0.163	1.490	0.684	61.02	5.16
5	Entangling	0.326	0.285	1.592	1.172	61.02	8.44

Peak values count additional allocated device memory during forward and backward. Dense states hold q^4 amplitudes per frame; the four local matrices contain $4q^2$ complex entries, before temporary storage.

Across all six checkpoints the maximum absolute clock discrepancy was 3.13×10^{-7}. Absolute gradient discrepancies were at most 8.15×10^{-10} for angles and 2.99×10^{-6} for couplings. The local contraction was profiled after training. The timings describe the circuit readout alone.

Evaluation

Uncertainty, confirmation and reproducibility

Historical uncertainty. For each window we first average the two signed loss differences across fixed training seeds. We then draw 2,500 matched UTC-day block bootstrap samples and calculate the row-weighted mean in each draw. An interval expresses variation across sampled days under that procedure. The October sample includes 26 UTC-day blocks. Comparing signs across many radices and folds increases the opportunity for a chance finding. These individual intervals carry no multiplicity correction.

q	Phase minus product	Lower 95%	Upper 95%
2	−0.00386	−0.00433	−0.00335
3	+0.00191	+0.00128	+0.00251
5	+0.00661	+0.00575	+0.00746

Paired October MAE differences in log basis points. Negative values favor the circuit with the bond phase. The unit of comparison is the paired loss on a matched observation.

Sealed November confirmation. Twelve frozen models and persistence produced forecasts for 5,120 origins before labels were released. The freshness rule admitted 2,042 one-second targets (39.9%). Ten of 15 scheduled sessions had observations, all between November 2 and 13, 2026. The 250 ms target gate applied equally to every model. All twelve checkpoint MAEs were below persistence (0.18012).

			Phase minus product			
q	Product	Phase on	Seed 25	Seed 26	Mean	Mean 95% interval
2	0.16432	0.15851	+0.00156	−0.01318	−0.00581	−0.00642, −0.00519
3	0.16434	0.16920	+0.01273	−0.00302	+0.00486	+0.00431, +0.00540
5	0.16740	0.17685	+0.02169	−0.00278	+0.00945	+0.00862, +0.01027

Columns marked 25 and 26 identify seeds 20261225 and 20261226. Ten matched UTC-day blocks underlie each interval. The intervals condition on those two fits.

For all three radices the phase increased error for seed 25 and decreased it for seed 26. The mean favored phase only at $q = 2$. This result supports a specific contrast among frozen models. Estimating variability across new training runs requires additional independent seeds.

Scope and reproduction. Changing q changes the state dimension, the clock spectrum, the initial basis law and the scale of the same learned coupling within each bond phase. A radix contrast cannot isolate dimension alone. The phase contrast within a radix preserves the parameter count and optimization budget. The work uses classical GPU simulation, and forecast errors do not measure fills, costs or investment returns.

The source bundle includes the machine-readable aggregate receipts, generated table rows and a checksum-verifying Python generator. The archive contains no quote windows or checkpoint tensors. Reproduction commands are given in `README.md`. The README also identifies the training, scoring and profiling commands and the contraction implementation.