Eight-wire cyclic circuits on historical FX

Matched forecast controls and an exact GPU readout

PhaseWeave research note | 21 November 2026

Result. In a separately sealed November afternoon panel, coherent product circuits had lower one-second FX forecast MAE than coherent entangling circuits at radices 7 and 11. The radix-3 effect was inconclusive. The comparison uses two fixed initializations, one currency triangle and ten market dates that were used in a prior noon panel.

Design. We trained 24 width-200 TokenGT models across three radices, four circuit readouts and two fixed seeds. All models have 1,538,489 trainable parameters and share their initial weights within seed. Each fit uses 8,192 training windows, three epochs and 384 optimizer steps. A disjoint validation split selects the checkpoint. For the 14:00–15:00 UTC panel, fixed clock-only selection separated anchors by at least 17 seconds. The 2,018 predictors were sealed before decoding outcomes, and 1,329 one-second targets met the requirement for three fresh quotes strictly after the origin.

Radix	Entangling	Product	Dephased	Max-plus	Difference	95% UTC-day interval
3	0.169635	0.169364	0.175622	0.169767	+0.000271	$[-0.000857, +0.001372]$
7	0.177476	0.172065	0.185230	0.177579	+0.005411	$[+0.004290, +0.006494]$
11	0.184360	0.172271	0.188032	0.177560	+0.012089	$[+0.010191, +0.013883]$

MAE is in log basis points, averaged over both directions and both seeds. The difference is entangling minus product. The intervals resample whole UTC days and are pointwise. Persistence scored 0.196978. Entangling improved on dephasing at every radix. At radix 11, max-plus also had lower MAE than entangling.

Exact local readout. Write $a_i(s) = \sqrt{p_i(s)}\exp(\mathrm{i}\alpha_i \cos(2\pi s/q))$ for the prepared amplitude on wire i at level s. Diagonal bonds on the chain contribute phase $2\pi\lambda_{ij}su/q$. Before the local mixers, a single-wire partial trace gives

$$\rho_i(s,t) = a_i(s)\overline{a_i(t)}\prod_{j\sim i}\sum_{u=0}^{q-1} p_j(u)\exp\bigl(2\pi\mathrm{i}\lambda_{ij}(s-t)u/q\bigr).$$

Bonds away from wire i cancel when the trace equates environment indices. The readout is $\mathrm{tr}(CU_i\rho_i U_i^\dagger)$, where C is the diagonal cosine clock. The product arm removes the diagonal bond terms and retains the locally modulated unitaries. Dephasing and max-plus use classical sum-product or maximum log path scores with the same local transition probabilities.

Compute and scope. Radix 11 has $11^8 = 214{,}358{,}881$ joint amplitudes per frame. The partial trace keeps eight 11×11 density matrices. The RTX 4090 completed the full study in 100.01 seconds, using at most 517.66 MiB allocated by PyTorch. An independent cuTensorNet contraction checked clock values and parameter derivatives at each radix and both seeds, with maximum absolute discrepancy 6.1×10^{-15}. These are classical GPU simulations. The panel tests a new hour on reused dates. Execution returns, independent later-date generalization and population-seed uncertainty remain outside this estimate.

Interactive comparison Aggregate results and evidence digests