

Supplementary Information for “Programmable Coherent Memory Kernels for Quantum Reservoir Computing in Waveguide QED”

Hong Jiang ^{1,2*}, Yu Wu ^{3,4}, Yue Ban ^{1,4*}, Xi Chen ^{1,4*},
Juan José García-Ripoll ^{1,2*}

¹Quantum Advanced Research Center, QuARC-CSIC, Calle Serrano 113b, 28006, Madrid, Spain.

²Instituto de Física Fundamental, IFF-CSIC, Calle Serrano 113b, 28006, Madrid, Spain.

³Department of Physics, Shanghai University, 200444, Shanghai, China.

⁴Instituto de Ciencia de Materiales de Madrid, ICMN-CSIC, Cantoblanco, 28049, Madrid, Spain.

*Corresponding author(s). E-mail(s): jiang.hong@iff.csic.es; yue.ban@csic.es; xi.chen@csic.es; jj.garcia.ripoll@csic.es;

S1. System–field correlations

To verify that the delayed memory is carried by coherent system–field correlations rather than by an effective classical delay variable, we monitor the Schmidt entropy during the STM task. From the time-bin MPS, we extract the Schmidt coefficients $\{\lambda_\alpha\}$ across the bipartition between the resonator and the active delay-line modes and compute

$$S_{\text{vN}} = - \sum_{\alpha} \lambda_{\alpha}^2 \log_2 \lambda_{\alpha}^2, \quad (\text{S1})$$

where the normalized Schmidt weights satisfy $\sum_{\alpha} \lambda_{\alpha}^2 = 1$. Figure S1 shows that the TDQRC develops a finite Schmidt entropy, with representative values $S_{\text{vN}} \sim 0.6\text{--}0.9$ bits. The nonzero Schmidt entropy demonstrates that the resonator becomes entangled with the waveguide modes retained inside the active delay line. The memory channel studied in the main text is therefore physically encoded in the joint system–field

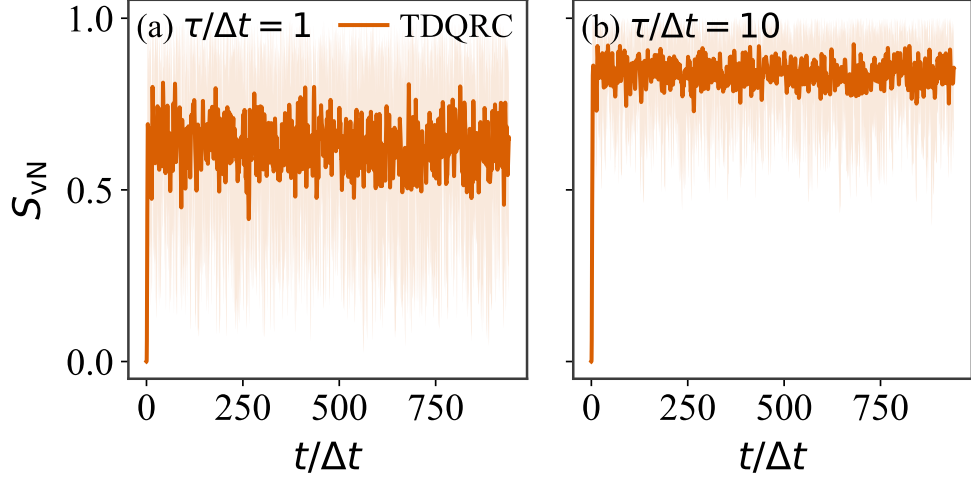


Fig. S1 Schmidt entropy during the STM task for representative feedback delays. The entropy is evaluated across the bipartition between the resonator and the active delay-line modes in the time-bin MPS.

state carried by the propagating waveguide modes, rather than in a classical delayed variable.

S2. Propagation-phase dependence

Here we examine how the phase accumulated by the field during propagation modifies the delay-resolved memory response. For the two-point geometry, the memory kernel may be written schematically as

$$K(t) = K_0\delta(t) + K_\tau e^{i\phi}\delta(t - \tau), \quad (\text{S2})$$

where K_0 and K_τ denote the instantaneous and delayed coupling strengths, respectively, and ϕ is the propagation phase between the two coupling points. Varying ϕ leaves the return time τ unchanged, but changes the phase with which the delayed field re-enters the resonator. It therefore controls the coherent interference between the returning field and the instantaneous resonator response.

We scan $\phi/\pi \in \{0, 0.25, 0.5, 0.75, 1\}$ at fixed delay $\tau/\Delta t = 4$, using the same TDQRC and readout protocol as in Fig. 2 of the main text. The resulting STM and delayed-product capacities are shown in Fig. S2. The propagation phase primarily controls the strength and shape of the memory revival, while leaving its temporal position approximately unchanged.

In both the STM and delayed-product diagnostics, the feedback-induced feature remains localized near $d = \tau/\Delta t$, because its arrival time is fixed by the propagation delay. Its amplitude and surrounding lag profile, however, vary with ϕ owing to constructive or destructive interference between the returning field and the resonator

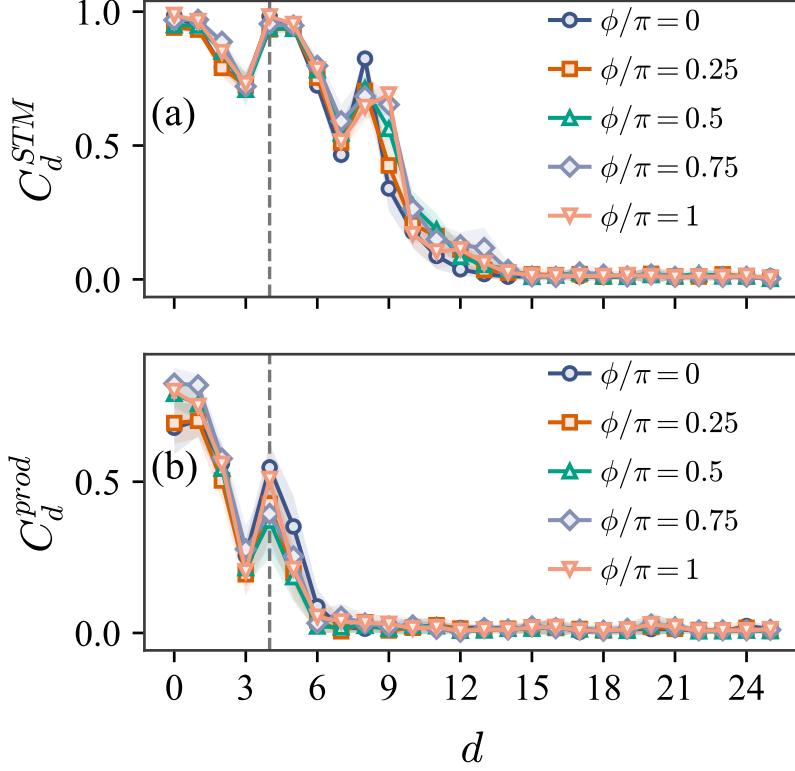


Fig. S2 Propagation-phase dependence of delay-resolved memory diagnostics at fixed programmed delay $\tau/\Delta t = 4$. (a) STM capacity C_d^{STM} . (b) Delayed-product capacity C_d^{prod} . Different curves correspond to $\phi/\pi \in \{0, 0.25, 0.5, 0.75, 1\}$. The vertical dashed line marks $d = \tau/\Delta t$.

response. The choice $\phi = 0$ used in the main text therefore represents the constructive-feedback configuration in which the delay-induced memory feature is most clearly resolved.

S3. Feedback-amplitude tuning

The geometry-programmed memory kernel provides separate control over the arrival time, phase, and weight of the coherently returned field. The propagation delay τ determines when previously emitted information returns to the resonator, while the coupling factor $\sqrt{\gamma_\mu \gamma_\nu}$ determines the weight of the corresponding delayed-feedback contribution. To isolate the latter effect, we vary $\sqrt{\gamma_1 \gamma_2}/\Gamma$ while keeping the total coupling rate $\Gamma = \gamma_1 + \gamma_2$ and the programmed delay $\tau/\Delta t$ fixed. The accessible range is

$$0 \leq \frac{\sqrt{\gamma_1 \gamma_2}}{\Gamma} \leq \frac{1}{2}, \quad (\text{S3})$$

where the upper bound is reached for balanced couplings, $\gamma_1 = \gamma_2$.

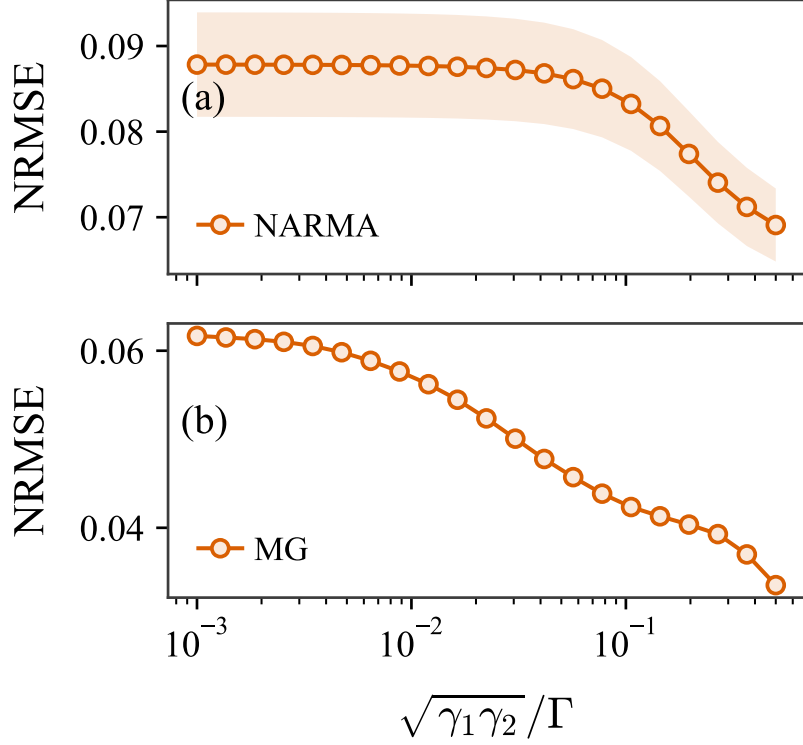


Fig. S3 Prediction performance as a function of the normalized delayed-feedback coupling factor at fixed $\Gamma = \gamma_1 + \gamma_2 = 0.2$, $U = 6$, and $\tau/\Delta t = 4$. The horizontal axis is $\sqrt{\gamma_1 \gamma_2} / \Gamma$. (a) Test NRMSE for the NARMA-5 task. (b) Forecasting NRMSE for the Mackey-Glass task described in Supplementary Section S9, evaluated at forecast horizon $h = 10$.

As the delayed-feedback amplitude increases, both prediction tasks improve over the parameter range considered [Fig. S3]. At fixed total coupling rate and propagation delay, a larger value of $\sqrt{\gamma_1 \gamma_2}$ increases the contribution of the coherently returned field to the reservoir dynamics, thereby strengthening the task-relevant delayed temporal features. The balanced point $\gamma_1 = \gamma_2$, which maximizes $\sqrt{\gamma_1 \gamma_2}$ at fixed total decay rate, therefore realizes the strongest coherent-feedback regime and is adopted in the main-text simulations.

In addition to the propagation delay and the phase discussed in Supplementary Section S2, the delayed-feedback amplitude thus provides a complementary parameter for engineering the geometry-programmed memory kernel.

S4. Feature-level initial-state washout

The delay-induced memory revivals discussed in the main text should encode the history of the driven input sequence rather than a persistent dependence on the initial resonator state. We therefore examine whether two reservoirs subjected to the same

input sequence but initialized differently develop convergent readout features. This provides an empirical feature-level check of initial-state forgetting, as required for echo-state behavior.

We initialize the resonator in

$$|\psi_0^{(1)}\rangle = |0\rangle, \quad |\psi_0^{(2)}\rangle = |1\rangle, \quad (\text{S4})$$

while taking the waveguide field to be initially in vacuum in both cases. The two reservoirs are then driven by the same input sequence using otherwise identical parameters. Let $\mathbf{x}_k^{(1)}$ and $\mathbf{x}_k^{(2)}$ denote the corresponding readout-feature vectors at input step k , with the constant bias component excluded. We quantify their separation through the normalized Euclidean distance

$$D_k = \frac{\|\mathbf{x}_k^{(1)} - \mathbf{x}_k^{(2)}\|_2}{\sqrt{(\|\mathbf{x}_k^{(1)}\|_2^2 + \|\mathbf{x}_k^{(2)}\|_2^2)/2 + \eta}}, \quad (\text{S5})$$

where $\eta = 10^{-12}$ prevents numerical singularities when both feature norms become very small.

The feature distance D_k decreases by several orders of magnitude during the washout interval [Fig. S4]. Consequently, the readout features used for training and testing become insensitive to the chosen initial resonator state. This supports the interpretation that the delay-resolved revivals reported in the main text represent driven input-history memory rather than an unwashed initial-state contribution. The Kerr-nonlinear reservoir exhibits a larger long-time feature mismatch than the linear reference for the parameters considered. This indicates slower or less complete feature-level convergence in the nonlinear dynamics, although the remaining difference is small relative to the initial separation. For the linear reservoir at long propagation delay, weak bounded oscillations remain visible in D_k . Their characteristic delayed recurrence is consistent with coherent returns of the residual initial-state component through the feedback line.

S5. Non-integer programmed delays

The main text considers integer delay ratios $d_r = \tau/\Delta t$, for which the propagation time can be associated directly with an integer input lag. In a waveguide implementation, however, τ is determined by the propagation length and group velocity and need not be commensurate with the input-update interval Δt . We therefore examine half-integer values of d_r to test whether the delay-controlled memory revival persists away from this commensurate limit.

The benchmark targets are defined only at integer lags d . Letting $d_r = m + \eta$ with $m = \lfloor d_r \rfloor$ and $0 < \eta < 1$, the returning field overlaps two adjacent input bins,

$$u(t - \tau) \longrightarrow (1 - \eta) u_{k-m} + \eta u_{k-m-1}, \quad (\text{S6})$$

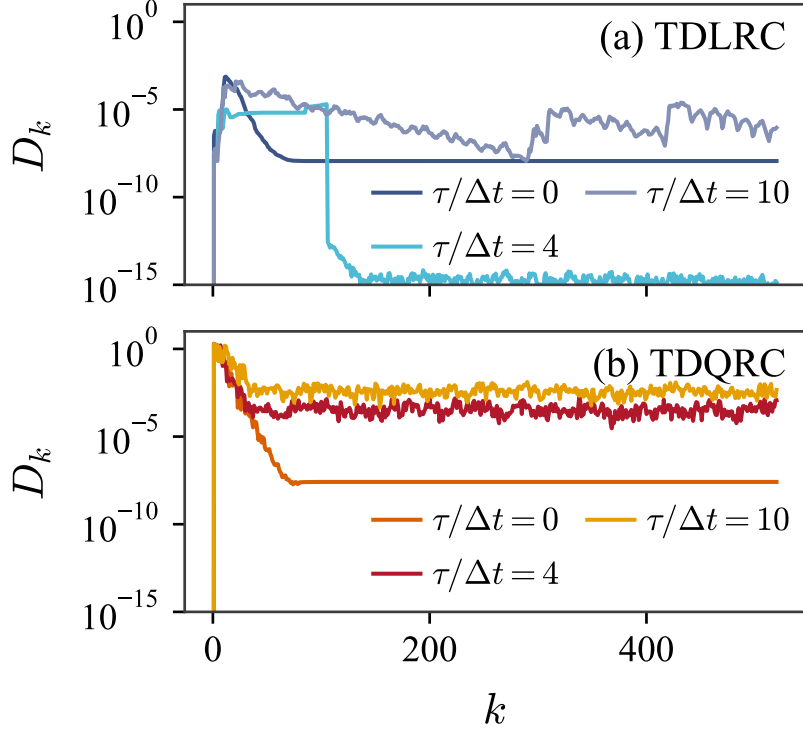


Fig. S4 Feature-level initial-state washout for reservoirs driven by the same input sequence but initialized in the resonator states $|0\rangle$ and $|1\rangle$. The normalized feature distance D_k , defined in Eq. (S5), is shown for the linear and Kerr-nonlinear reservoirs at representative delays $\tau/\Delta t \in \{0, 4, 10\}$.

up to the sampling convention and the temporal filtering of the resonator and virtual-node readout. Since no target exists at a non-integer lag, the weight of a fractional delay is projected onto $d = m$ and $d = m + 1$, broadening or splitting the revival.

The resulting STM and delayed-product capacities for $\tau/\Delta t = 2.5, 4.5, 6.5$, and 8.5 are shown in Fig. S5. The revivals in C_d^{STM} continue to track the programmed delay. For the longer delays, the primary revival is shared by the two integer lags adjacent to d_r , while weaker features at later lags arise from repeated coherent returns. For $d_r = 2.5$, the returned field overlaps more strongly with the intrinsic short-time response of the reservoir, producing a broader low-lag structure rather than a clearly separated revival.

The delayed-product capacity exhibits the corresponding nonlinear response. The Kerr interaction mixes the coherently returned field with the current reservoir dynamics, generating delayed nonlinear features associated with both neighboring integer lags. The resulting product-capacity profile therefore develops a broadened shoulder or a pair of neighboring features around d_r . This structure is most clearly resolved for the longer propagation delays, where it is separated from the intrinsic short-time nonlinear response. By contrast, the linear reference reservoir exhibits negligible delayed-product

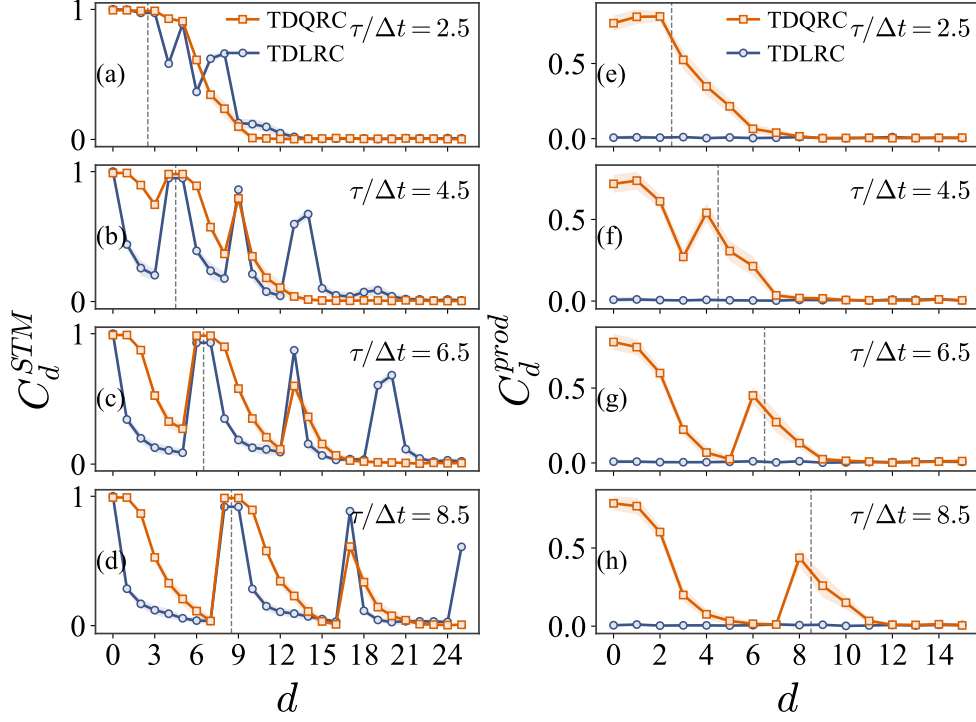


Fig. S5 Delay-resolved memory capacities for non-integer programmed delay ratios. (a)–(d) Short-term-memory capacity C_d^{STM} for $\tau/\Delta t = 2.5, 4.5, 6.5, 8.5$. (e)–(h) Corresponding delayed-product capacity C_d^{prod} . The vertical dashed line indicates the programmed return time $d_r = \tau/\Delta t$. Because the benchmark targets are defined at integer lags, a non-integer return time is resolved as a distributed response over the neighboring values of d . The Kerr-nonlinear TDQRC exhibits a corresponding delayed-product feature, whereas the linear reference has negligible product capacity.

capacity over the parameter range shown, confirming that coherent delayed memory alone does not generate the required multiplicative feature under a linear readout.

These results extend the delay-matching picture of the main text beyond integer delay ratios. When $\tau/\Delta t$ is an integer, the coherent return is localized near a single discrete lag. When it is non-integer, the same return time is represented by a distributed response over neighboring integer-lag targets. Thus, the propagation delay continues to determine the temporal location of the memory feature, while the discrete sampling protocol determines how that feature is resolved by the readout.

S6. Multi-delay geometry

The two-point geometry considered in the main text is the minimal configuration that produces a single coherent memory return and therefore allows the roles of memory placement and Kerr-induced nonlinear processing to be distinguished most clearly. Adding further coupling points naturally extends the geometry-programmed memory kernel by introducing several delayed feedback paths. Their computational benefit,

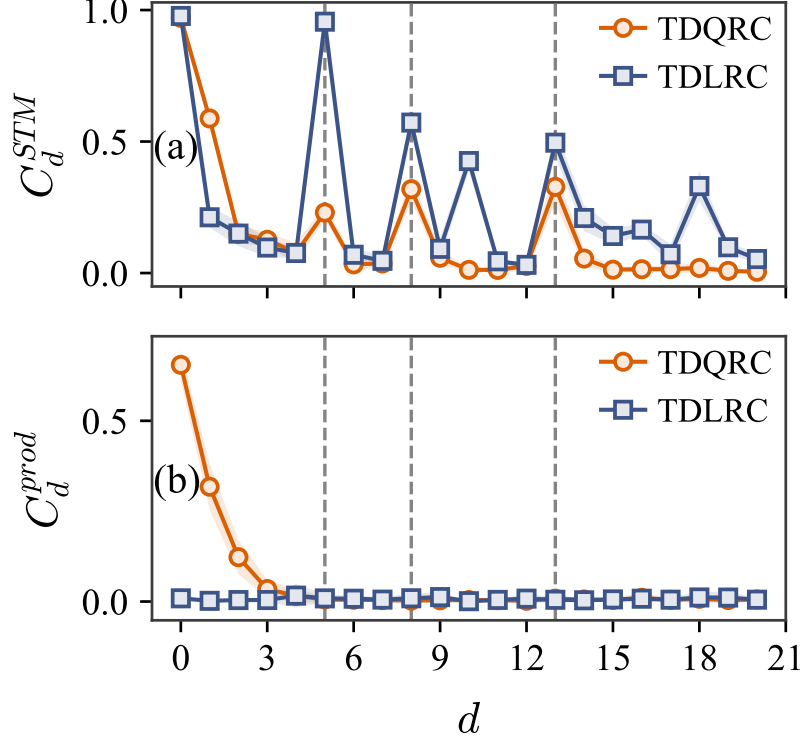


Fig. S6 Delay-resolved memory diagnostics for a three-point geometry. (a) STM capacity C_d^{STM} , showing revivals near the programmed propagation delays. (b) Delayed-product capacity C_d^{prod} . Vertical dashed lines indicate the programmed delay indices.

however, is not automatic, because the corresponding coherently returned fields can interfere before being processed by the Kerr nonlinearity. For a three-point geometry, the kernel takes the form

$$K(t) \simeq K_0\delta(t) + K_{12}\delta(t - \tau_{12}) + K_{23}\delta(t - \tau_{23}) + K_{13}\delta(t - \tau_{13}), \quad (S7)$$

where $\tau_{\mu\nu} = |x_\mu - x_\nu|/v_g$, while the coefficients $K_{\mu\nu}$ include the corresponding coupling strengths and propagation phases.

Distinct STM revivals appear near the programmed delays in Fig. S6(a), confirming that the placement of linear memory generalizes naturally to geometries with multiple feedback paths. Each delayed component of the kernel produces a corresponding coherent memory return, although the observed amplitudes and line shapes also depend on interference between the different channels. The delayed-product response in Fig. S6(b) exhibits qualitatively different behavior. Rather than producing one pronounced nonlinear-memory peak at each programmed delay, the TDQRC displays mainly a short-time nonlinear response in the parameter regime considered, with no clearly resolved delayed-product revival at the individual return times. The TDLRC

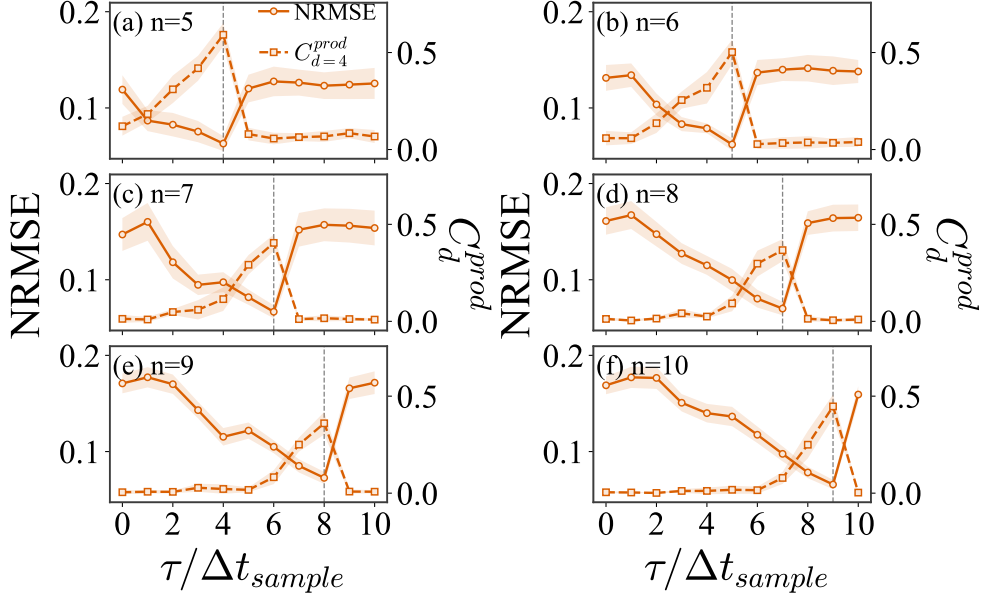


Fig. S7 Comparison between NARMA- n prediction error and the delayed-product capacity evaluated at the task-relevant lag $d = n - 1$. Solid circles (left axis) show the test NRMSE. Dashed squares (right axis) show $C_{d=n-1}^{\text{prod}}$. The vertical dashed line marks the task-matched delay $\tau/\Delta t = n - 1$. Shaded regions indicate one standard deviation over independent input realizations.

remains close to the product-capacity floor, as expected for a linear reservoir with a linear readout. These results show that adding coupling points directly programs several linear-memory return times, but the associated nonlinear processing is not determined by the delay positions alone. The Kerr interaction acts on the superposition of the instantaneous response and several coherently returned fields, so the resulting nonlinear features depend jointly on the feedback weights, propagation phases, temporal overlap, and interference between delayed channels. The two-point geometry used in the main text therefore provides the minimal setting in which coherent memory placement and nonlinear memory conversion can be separated unambiguously. More generally, multi-point geometries offer a route toward hardware-programmable memory kernels with several temporal scales, which may be useful for tasks involving distributed temporal dependencies once the relative phases and feedback amplitudes are optimized.

S7. Product-capacity diagnostic for NARMA task matching

The main text argues that the optimal programmed delay is determined by the explicit nonlinear temporal dependence of the target task. For the NARMA- n benchmark

defined in the main text, the only explicit input–input nonlinearity is

$$1.5 \tilde{u}_k \tilde{u}_{k-n+1}, \quad (\text{S8})$$

which defines the task-relevant product lag

$$d_{\text{task}} = n - 1. \quad (\text{S9})$$

To examine whether the prediction optimum is associated with this lag, we compare the NARMA- n prediction error with the delayed-product capacity $C_{d=n-1}^{\text{prod}}$, evaluated independently using the delayed-product diagnostic of the main text, with the same reservoir parameters and linear readout protocol.

Figure S7 provides an independent mechanistic check of the matching picture. As the task order increases from $n = 5$ to $n = 10$, the minimum prediction error shifts from $\tau/\Delta t \simeq 4$ to $\tau/\Delta t \simeq 9$. Over the same parameter range, the delayed-product capacity at the corresponding task lag, $C_{d=n-1}^{\text{prod}}$, develops its maximum at nearly the same programmed delay. Thus, lower prediction error coincides with stronger nonlinear delayed features at the explicit NARMA product lag. The observed correlation is consistent with the interpretation that the optimal programmed delay is governed by the availability of task-relevant nonlinear memory, rather than by memory depth alone.

S8. Semiclassical nonlinear-resonator control

To determine whether the delayed nonlinear features of the TDQRC can be reproduced by a classical mean-field oscillator, we compare it with a semiclassical nonlinear-resonator control (CRC). Starting from the operator equation of motion, we replace the resonator operator by a complex amplitude, $a(t) \rightarrow \alpha(t)$, and close the nonlinear term through the mean-field factorization $a^\dagger a a \rightarrow |\alpha|^2 \alpha$. For the two-point geometry with constructive feedback, the resulting delay-differential equation is

$$\dot{\alpha}(t) = -\frac{\Gamma}{2}\alpha(t) - \frac{\sqrt{\gamma_1\gamma_2}}{2}\alpha(t-\tau) - iU|\alpha(t)|^2\alpha(t) + \Omega(t), \quad (\text{S10})$$

using the same kernel convention as in the memory-kernel equation in the main text. The CRC therefore retains both the programmed delayed-feedback channel and Kerr-type nonlinearity, while reducing the joint resonator–waveguide quantum dynamics to a single deterministic complex trajectory. In particular, it neglects quantum fluctuations, higher-order moment dynamics, and correlations between the resonator and the propagating field.

The CRC readout is constructed from the classical resonator quadratures

$$X_\alpha(t) = \alpha(t) + \alpha^*(t), \quad P_\alpha(t) = i[\alpha^*(t) - \alpha(t)], \quad (\text{S11})$$

together with the corresponding output-field quadratures. The same input encoding, virtual-node sampling, feature normalization, and linear readout protocol are used

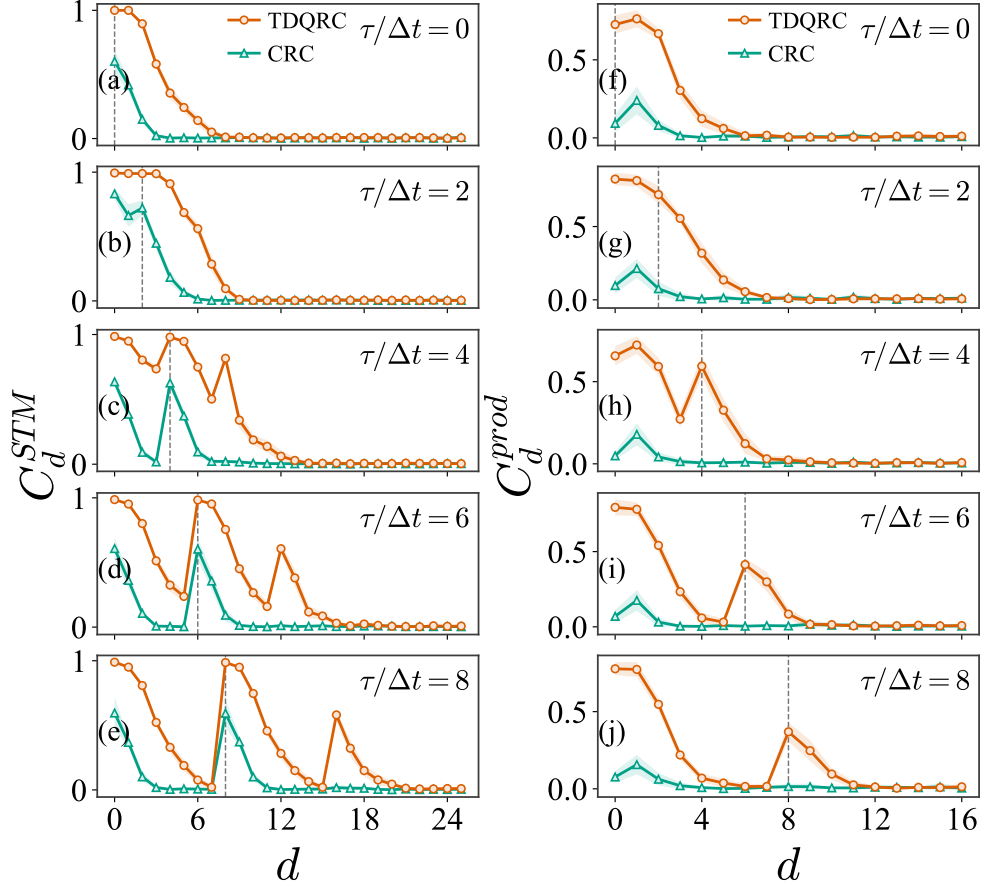


Fig. S8 Comparison of the TDQRC with the semiclassical nonlinear-resonator control for delay-resolved memory diagnostics. The left column shows the STM capacity C_d^{STM} , and the right column shows the delayed-product capacity C_d^{prod} . Rows correspond to different programmed delays $\tau/\Delta t$. Vertical dashed lines mark the nominal coherent-return lag $d = \tau/\Delta t$.

for the CRC and TDQRC, so that the comparison isolates the difference between deterministic mean-field dynamics and the full quantum evolution.

Figure S8 shows that the CRC reproduces the primary STM revival near the programmed delay. This confirms that the temporal placement of linear memory is already captured by the classical delayed-feedback equation and is primarily determined by waveguide propagation. A clearer difference emerges in the delayed-product diagnostic. Over the parameter range shown, the CRC develops substantially weaker nonlinear delayed-memory capacity than the TDQRC and does not reproduce the full delay-dependent product-feature profile. The programmed delay alone is therefore insufficient to account for the nonlinear memory observed in the quantum reservoir, while a single mean-field Kerr trajectory also provides only a partial description of its nonlinear processing. The NARMA results in Fig. S9 show the corresponding difference

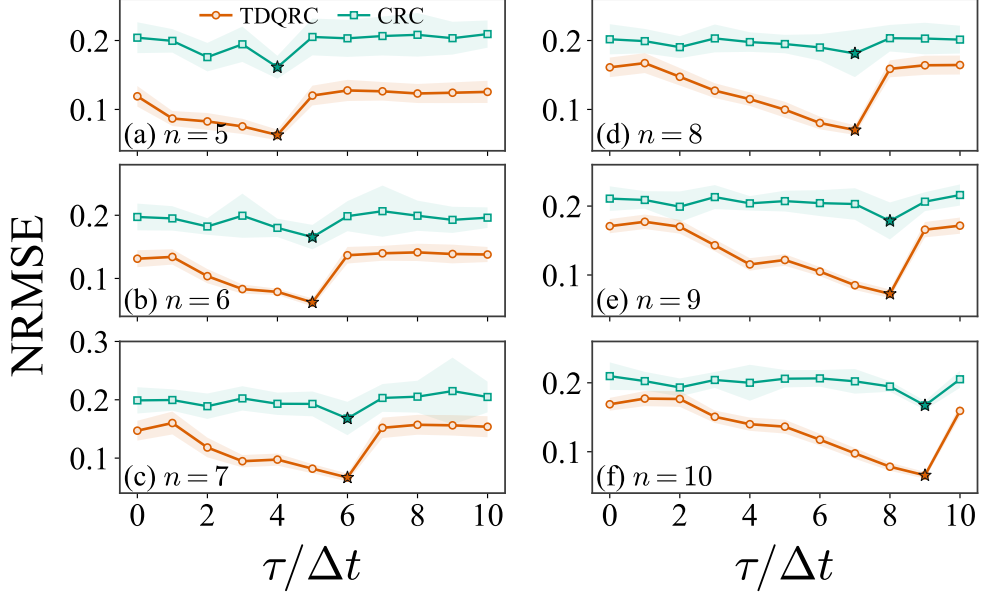


Fig. S9 Comparison of the TDQRC and the semiclassical nonlinear-resonator control for NARMA prediction. The test NRMSE is plotted as a function of programmed delay $\tau/\Delta t$ for $n = 5, \dots, 10$. Stars indicate the minimum NRMSE attained by each curve.

at the task level. Although the CRC benefits from the same geometry-programmed feedback delay, its prediction error remains higher than that of the TDQRC for the configurations considered. Together with the delayed-product diagnostic, this result shows that the useful nonlinear temporal features of the TDQRC are not fully reproduced by deterministic scalar mean-field dynamics. The comparison therefore identifies a beyond-mean-field contribution to the reservoir response, arising from quantum fluctuations, higher-order correlations, or joint resonator–field dynamics that are absent from the CRC.

S9. Mackey–Glass forecasting

To examine whether the task-matching picture extends beyond synthetic delay benchmarks, we consider forecasting of the chaotic Mackey–Glass time series [1]. The continuous Mackey–Glass signal $s(t)$ is generated from

$$\frac{ds(t)}{dt} = \beta_{\text{MG}} \frac{s(t - T_{\text{MG}})}{1 + s(t - T_{\text{MG}})^{p_{\text{MG}}}} - \gamma_{\text{MG}} s(t), \quad (\text{S12})$$

with $\beta_{\text{MG}} = 0.2$, $\gamma_{\text{MG}} = 0.1$, $p_{\text{MG}} = 10$, and $T_{\text{MG}} = 17$. The signal is sampled with $\delta t_{\text{MG}} = 0.1$ to obtain $x_k = s(k\delta t_{\text{MG}})$. For forecast horizon h , the prediction target is $y_k^{(h)} = x_{k+h}$. The input encoding, virtual-node construction, washout procedure, training/test partition, and linear readout are the same as those used throughout the main text.

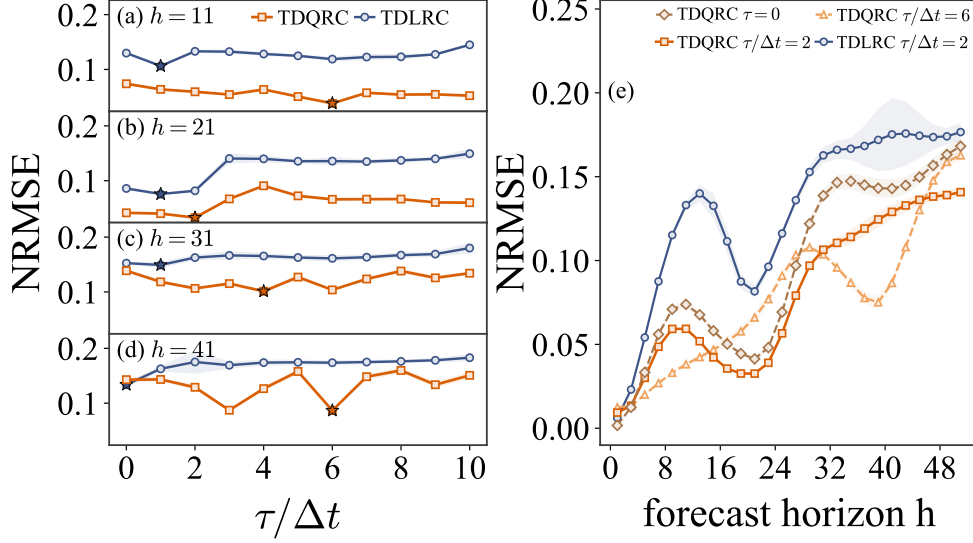


Fig. S10 Mackey–Glass forecasting with programmable delayed feedback. (a)–(d) Test NRMSE versus programmed delay $\tau/\Delta t$ for forecast horizons $h \in \{11, 21, 31, 41\}$, comparing the TDQRC and TDLRC. Stars indicate the minimum NRMSE within the scanned delay range. (e) Test NRMSE versus forecast horizon for representative programmed delays. The case $\tau/\Delta t = 0$ corresponds to the zero-delay limit.

Unlike the NARMA- n benchmarks, the Mackey–Glass task does not possess a single dominant delayed-product feature. Instead, its predictive information is distributed over multiple temporal correlations generated by the underlying nonlinear delay dynamics. Consequently, no unique task-relevant lag satisfying $\tau/\Delta t = d_{\text{task}}$ is expected. Figure S10 is consistent with this picture. The TDQRC generally outperforms the TDLRC over the forecast horizons considered, indicating that Kerr-induced nonlinear processing remains useful for chaotic time-series prediction. In contrast to the sharper NARMA behavior, however, the dependence on the programmed delay is broad and horizon dependent. Different forecast horizons favor different delay values, and increasing the delay alone does not systematically improve the forecasting performance.

The delay-matching principle therefore applies most naturally to tasks with a dominant temporal scale, whereas for tasks with distributed temporal correlations the programmed delay should be regarded as a tunable memory timescale rather than a uniquely determined matching condition.

References

- [1] Mackey, M. C. & Glass, L. Oscillation and chaos in physiological control systems. *Science* **197**, 287–289 (1977).