A Reproducible Evaluation Protocol for Quantum Non-Linearity in Variational Quantum Models

Six circuits. Four scores. Four different winners.

Pavel Sulimov Claude Lehmann

Zurich University of Applied Sciences (ZHAW), Winterthur · University of Zurich

NeurIPS 2026, Track on Evaluations and Datasets

Non-linearity has to be engineered, and it is judged by one number

A variational circuit is **linear** in its input state:

$$f_\theta(x) = \langle 0| \, U^\dagger(\theta; x) \, O \, U(\theta; x) \, |0\rangle.$$

Non-linearity in x comes from a mechanism:

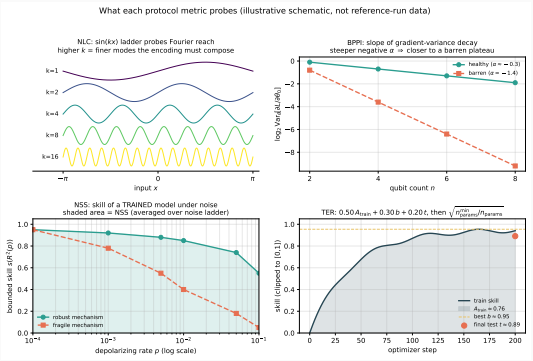
- data re-uploading
- ancilla tracing (a channel)
- measurement and classical post-processing
- explicit classical layers

How a new mechanism is judged today

accuracy on one small dataset,
or one expressibility scalar;
rarely at matched width, depth and training budget

The question that matters: at the same cost, what does the mechanism buy?

Four scores, one fixed recipe



- **NLC**, Fourier reach: fit $\sin(kx)$, $k = 1, 2, 4, 8, 16$, fixed Adam budget; integrate R^2 over $\log k$
- **BPPI**, plateau risk: slope of $\log_2 \mathrm{Var}[\partial L]$ over $n = 2, 4, 6, 8$ qubits
- **NSS**, noise stability: trained fit under depolarizing noise up to $p = 10^{-1}$
- **TER**, learning efficiency on Lorenz one-step prediction, scaled by parameter economy

All bounded. Width, depth, optimizer, seeds fixed across methods. Raw diagnostics exported with every score.

One composite, frozen before the results

$$\mathrm{NLPS} = 0.35\,\mathrm{NLC} + 0.25\,(1 - \mathrm{BPPI}) + 0.20\,\mathrm{TER} + 0.20\,\mathrm{NSS}$$

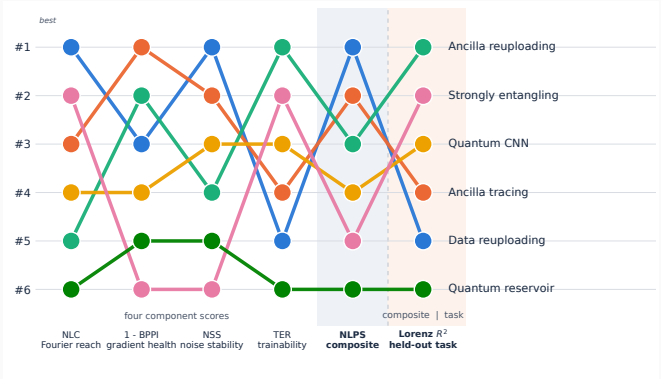
What it is for

one pre-registered axis, so methods from different papers can be placed side by side without each author choosing weights after seeing their numbers

The rule

reported *next to* its four components, never instead of them; read close values as rank bands; only within a declared (n, L)

Reference run at $(n, L) = (3, 3)$: every score crowns a different ansatz



Data reuploading
re-injected encoding

Ancilla tracing
partial trace (channel)

Ancilla reuploading
encoding + channel

Quantum CNN
hierarchical structure

Strongly entangling
entanglement baseline

Quantum reservoir
fixed unitary + readout

Three numbers

0.03

rank correlation between the composite and the held-out Lorenz task

the scores are probes, not a substitute for your task

0.97

rank agreement over 10^5 alternative weightings

weights do not matter; the top two never change

0.60

rank agreement between $(3, 3)$ and $(4, 4)$; the leader flips

the operating point does: declare and match (n, L)

Spearman ρ over six method-level means; Kendall τ for weightings and operating points. $K = 5$ seeds.

Add your ansatz

1. Register a circuit builder $U(\theta; x)$ with its parameter shape.
2. Rerun NLC $\rightarrow$ BPPI $\rightarrow$ NSS $\rightarrow$ TER $\rightarrow$ composite. Seeds, budgets and the global seed are fixed; cached partial runs are invalidated automatically.
3. Report all four scores and NLPS next to an established mechanism at the same (n, L).

No scoring code is touched.

$\approx$ **2 h**

full protocol at $(3, 3)$ on an 8-core laptop CPU

five seeds overnight, no GPU

PennyLane, scikit-learn; CC BY 4.0

Take-away

One number hides what the protocol is built to show

Report the four components with the composite, at a declared and matched (n, L).
Where the components disagree, **the disagreement is the result**.

Code, data, protocol:
`github.com/sulimovp/quantum-nonlinearity-eval`

Poster 139813, Paris Poster Hall, 9 December 2026