Supplementary Information

Quantum models pay for non-linearity in data encodings

Pavel Sulimov and Claude Lehmann

Theorem, corollary, figure and table numbers without the prefix "Supplementary" refer to the main text.

Supplementary Note 1: Effective degree and the sharpness rate

Corollary 1 credits N shots of a depth-E circuit with a literal degree of NE, but their derivatives behave like those of a polynomial of degree $E\sqrt{N}$:

Supplementary Theorem 1 (Effective degree)**.** *For N i.i.d. shots of a two-outcome circuit with E encodings and any head $|F| \le 1$: $|\partial_x \mathbb{E}[F]| \le E\sqrt{N}$ and $|\partial_x^2 \mathbb{E}[F]| \le e\,NE^2$, whereas a bounded polynomial of degree T reaches T and T^2.*

The continuation in Theorem 1 bounds every derivative through Cauchy estimates, $|\partial_x^k \mathbb{E}[F]| \le k!\,(eV/k)^{k/2}$, which gives $\sqrt{eV}$ for the slope and eV for the curvature, the constant in Supplementary Theorem 1. Hence $\mathbb{E}[F]$ lies within ε of a polynomial of degree $(2n\ln(\sqrt{2\pi n}/\varepsilon))^{1/2}$, so for any target, n incoherent encodings are worth at most that many coherent ones. The spectral bound turns that remark into a rate for every kind of feature:

Supplementary Theorem 2 (Sharpness rate)**.** *Write $T_K(g)^2 = \sum_{|\omega|>K} |\hat{g}(\omega)|^2$. Every finite-shot procedure obeys $\|\mathbb{E}[F] - g\|_2 \ge T_K(g) - \sqrt{V/K}\,e^{-K^2/2V}$ for every K. If $T_K(g) \ge cK^{-(\alpha+1/2)}$, as for a target with an α-Hölder singularity, then $\|\mathbb{E}[F] - g\|_2 = \Omega\big((V\log V)^{-(2\alpha+1)/4}\big)$.*

Corollary 3 is the case $\alpha = 0$, where the slope argument gives the sharp constant without the logarithm. The exponent is attained numerically: against targets with Fourier coefficients $k^{-(1+\alpha)}$, singular at $x = 0$ (logarithmically at $\alpha = 0$), the best bounded head on V single-encoding shots has fitted exponents 0.252, 0.376, 0.502, 0.629 and 0.757 for $\alpha = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$, against the predicted $(2\alpha+1)/4$, and the bound lies between 0.46 and 0.84 of the achieved error in all 25 cases; a jump, a Hölder-$\frac{1}{2}$ cusp and a kink at $x = \pi/2$ give 0.251, 0.542 and 0.764, the cusp still short of 0.50 because its ratio to depth is falling (checks below). So V incoherent encodings are worth about $\sqrt{V}$ coherent ones whatever the feature, and the constant depends on both its smoothness and its position: with $N = K^2$ shots against depth K, the error ratio approaches 0.78 for a jump, 0.30 for a cusp and 0.33 for a kink at $x = \pi/2$, where the shot variance is largest, but 0.18 for a kink at $x = 0$, where it vanishes. The jump's 0.78 is against the best polynomial bounded by 1; the ratio to its Fourier tail is 0.86, while for the cusp and the kink at $x = \pi/2$ the two denominators agree at the printed precision.

The spectral bound itself is attained at $\omega = 0$. At $\omega \sim \sqrt{N}$ the worst case over heads reaches 0.6 to 0.7 of it (0.405 against 0.607 at $N = 16$, $\omega = 4$), so the exponent is Gaussian but the constant is not 1. For single uploads the exact amplitude ceiling, 2^{-m} at $\omega = m$ [1], is strictly stronger than $e^{-m/2}$.

Numerically the curvature bound of Supplementary Theorem 1 is reached with constant 1, and the best N-shot procedure matches a depth-K circuit with $K \propto N^{0.51}$ on a jump and $N^{0.53}$ on a kink. For analytic targets the difference disappears: smooth features need no large derivatives, and in our checks both arms reach the numerical floor by $N = 8$.

Proof of Supplementary Theorem 1. Write $\mathbb{E}[F] = B(p(x))$ with $B(p) = \sum_k F_k b_{N,k}(p)$, a polynomial with Bernstein coefficients bounded by 1, and $\varphi^2 = p(1-p)$. The identity $\varphi^2 b'_{N,k} = (k - Np) b_{N,k}$ gives $|\varphi B'| \le \sqrt{N}$, since $\mathbb{E}|k - Np| \le \sqrt{N}\varphi$. Differentiating again, $\varphi^2 B'' = \sum_k F_k[(k - Np)^2/\varphi^2 - N] b_{N,k} - (1-2p)B'$, so $|\varphi^2 B''| \le 2N + |B'| \le 4N$. For a two-outcome circuit the Fisher information is $p'^2/\varphi^2 \le E^2$, so $|\partial_x B| = |\varphi B'|\,|p'|/\varphi \le E\sqrt{N}$ and $|B'' p'^2| \le 4NE^2$; Bernstein's inequality for $p - \frac{1}{2}$ gives $|p''| \le E^2/2$, and $|B'| \le 2N$, so $|B' p''| \le NE^2$. This gives the constant 5; the Cauchy estimate after Theorem 1, with $V = NE^2$, improves it to e. Exact computation gives $\max |\partial_x^2 \mathbb{E}[F]|/(NE^2) = 1.00$ over random, majority and alternating heads ($E = 1, 2$, $N \le 256$). All orders, adaptive procedures included, follow from Theorem 1: $|\partial_x^k \mathbb{E}[F]| \le k!\,(eV/k)^{k/2}$ with $V = \sum_i E_i^2$.

Relation to Bernstein polynomials. For one encoding per shot the statements of Supplementary Theorems 1 and 2 are classical. An N-shot head is a Bernstein polynomial with bounded coefficients in the outcome probability p, the object studied in work on Bernoulli factories and stochastic computing [2, 3, 4] and, with measurement and post-selection, in quantum Bernoulli factories [5]. Such polynomials approximate a function of Hölder smoothness α at the rate $(p(1-p)/N)^{\alpha/2}$ [6] and no faster [7], which in terms of x corresponds to depth $\sqrt{N}$. The proof of Supplementary Theorem 1 rests on a coincidence: an N-shot expectation is a Bernstein-type polynomial in p, the Ditzian–Totik weight $\sqrt{p(1-p)}$ controls the derivatives of such polynomials, and this weight equals the Fisher chain-rule factor of the circuit. The sharp $e^{-\pi\sqrt{n}}$ rate behind Newman's construction is due to Stahl [8].

Proof of Supplementary Theorem 2. By Theorem 1, $|c_\omega| \le e^{-\omega^2/2V}$. The triangle inequality in ℓ^2 restricted to $|\omega| > K$ gives $\|\mathbb{E}[F] - g\|_2 \ge T_K(g) - (\sum_{|\omega|>K} e^{-\omega^2/V})^{1/2}$, and $\sum_{|\omega|>K} e^{-\omega^2/V} \le 2\int_K^\infty e^{-t^2/V}\,dt \le (V/K)e^{-K^2/V}$. With $T_K(g) \ge cK^{-(\alpha+1/2)}$ and $K = \sqrt{\beta V \ln V}$ the subtracted term is $O(V^{1/4-\beta/2})$, so any $\beta > \alpha + 1$ leaves $\|\mathbb{E}[F] - g\|_2 \ge \frac{c}{2}(\beta V \ln V)^{-(\alpha+1/2)/2}$ for large V. The logarithm is an artefact of the Gaussian cap: at $\alpha = 0$ the slope argument of Corollary 3 gives the same exponent with a constant and no logarithm.

Checks: targets with cosine coefficients $k^{-(1+\alpha)}$ truncated at 2,048 and scaled to sup norm 1, $\alpha = 0$ to 1, $V = 4$ to 1,024; every one is singular at $x = 0$, and $\alpha = 0$ is a logarithmic singularity rather than a jump, with a scale that depends on the truncation. The optimum over all heads $|F| \le 1$ on V i.i.d. shots of one encoding is a bounded least-squares problem in the Bernstein coefficients, hence exact. No violation in 25 cases; the bound is 0.46–0.84 of the achieved error; the fitted exponents differ from $(2\alpha + 1)/4$ by at most 0.007; and the binding K sits at 3.0–5.0$\sqrt{V}$, with K^2/V rising with $\ln V$ at slopes 1.4–2.2 over $V = 64$ to 1,024, against the asymptotic $\alpha + 1$. Closed-form targets, same V: a jump and a Hölder-$\frac{1}{2}$ cusp at $x = \pi/2$ and kinks at $x = \pi/2$ and $x = 0$ give no violation in 20 cases and fitted exponents 0.251, 0.542, 0.764 and 0.755 (the cusp's ratio to depth is still falling at $V = 1{,}024$). Here the bound is 0.28–0.74 of the achieved error. The ratio to the depth-$\sqrt{V}$ tail is 0.86 for the jump, 0.33 for the kink at $x = \pi/2$ and 0.18 for the kink at $x = 0$. Against the best polynomial of degree K bounded by 1 the jump ratio is 0.78; the cusp and the kink at $x = \pi/2$ are unchanged at the printed precision.

Supplementary Note 2: Exchange-rate constants and checks of the degree and slope bounds

Three classical heads and an adaptive sequential test stay within the degree bound of Corollary 1, and the slope reaches the bound of Corollary 2 at $N = 1$ only: the best head over all $|F| \le 1$ reaches $0.86\sqrt{N}$ at $N = 128$ and the majority vote $\sqrt{2/\pi}\sqrt{N}$ at large N, so for $N > 1$ the Fisher bound is a bound and not an attained value. At equal total encodings, the best classical head on single-encoding shots falls behind a deep circuit by a factor that grows as $T^{1/4}$ (fitted exponents 0.26 and 0.50 over $T \ge 4$, against $\frac{1}{4}$ and $\frac{1}{2}$). With post-selection and N shots, the expected ratio estimator is a polynomial of degree at most NE. For a model of the square wave $0.9\,\mathrm{sign}(\sin x)$ with $E = 2$ and floor 0.2 it converges to the ideal rational output once $N \gtrsim 1/q_{\min}$ (bias 0.004 at $Nq_{\min} = 3.2$), although shot noise still leaves an RMS error of 0.33 there, against 0.22 for the infinite-shot ideal (Fig. 2c).

For depth-E shots whose expectation is the degree-E Fejér mean, read out by a majority vote (not the optimum over heads or over degree-E polynomials), exact computation gives RMS ≈ 0.84–$0.88\,(NE^2)^{-1/4}$ for $E = 2, 4$. The constants quoted in the main text are those of the class optimum of each arm: the best cosine polynomial of degree T with $|F| \le 1$, found by bounded least squares on a 4,096-point grid (SLSQP from the Fejér mean), has RMS $0.70\,T^{-1/2}$ for $T = 2, 4, \ldots, 128$ (error$\times\sqrt{T}$ between 0.673 and 0.704; an odd degree has the optimum of the next even one, because the jump has only odd harmonics; the Fejér mean gives 0.75–0.90, and the polynomial bounded by 0.98 that the hardware circuits realise is within 6% of the optimum), and the best head on N single-encoding shots, which coincides with the majority vote for a jump, has RMS 0.53–$0.56\,N^{-1/4}$ for $N \ge 4$. For the multi-scale shots of the main text, exact computation with the Bayes head gives exponents of 0.47 and 0.25 for the multi-scale and equal-scale designs (post hoc).

Supplementary Note 3: Post-selection: trained models, frontier, saturation and cost axis

The targets are the square wave $0.9\,\mathrm{sign}(\sin x)$ (jumps) and the triangle wave $1 - 2|x|/\pi$ (kinks), and errors are RMS over $x \in [-\pi, \pi)$. We tested the theory against all 124 post-selected models we had trained on these two targets ($E \le 6$, noiseless and noisy, five seeds). None beat the error floor implied by Theorem 3, and none had a mean shot overhead below the Bernstein bound.

Theorem 3 puts a model with floor q within $1/T_{J+1}(\frac{1+q}{1-q})$ of a polynomial of degree $(J+1)E$, so its error is at least $\mathrm{floor}(E, q) = \max_J[d_{(J+1)E}(g) - 1/T_{J+1}(\frac{1+q}{1-q})]$, with $d_K(g)$ the RMS distance from the target g to the trigonometric polynomials of degree K (not required to be bounded). At fixed E this floor keeps falling as $q \to 0$, but no post-selected model can beat R_E, the best error of any degree-E rational function with no floor at all. So the error is at least $\max(\mathrm{floor}(E, q), R_E)$, and no single constant can match the floor alone (at $E = 2$ the ratio reaches 11 at $q = 10^{-4}$). Optimising $\mathcal{N}/D$ directly under the model constraints gives the best achievable error. For the square wave it stays within a factor 1.66 of $\max(\mathrm{floor}, R_E)$ in all 48 cells where the optimiser converged ($E \le 6$, q from 0.5 down to 10^{-4} for $E \le 3$ and to 0.01, 0.02 and 0.05 for $E = 4$, 5 and 6), and 200-start searches reproduce the optimum exactly. The bound is therefore tight up to a small constant for jumps. Below a saturation floor $q_{\mathrm{sat}}(E)$ (about 0.2 at $E = 2$ and 0.02 at $E = 3$) lowering q buys nothing, and only more encodings help; this couples degree and floor as in Newman's construction. For the triangle wave the same law holds only to within 2.6–6.5× for $E = 3$–6. The gap is not an artefact of the optimiser: at the worst cells a search from a deterministic unitary start plus 60 restarts moves the frontier by less than 1%, and an exact uniform-norm computation shows the

same gap. It is a smooth function of the target. Over the family with Fourier coefficients $k^{-(1+\alpha)}$ at $E = 4$ (a logarithmic singularity at $\alpha = 0$, a kink at $\alpha = 1$), the ratio to $\max(\text{floor}, R_E)$ rises monotonically with the smoothness, from 1.56 at $\alpha = 0$ to 4.55 at $\alpha = 1$ ($q = 0.5$; 1.63 to 5.84 at $q = 0.2$). The cause is that the floor subtracts a sup-norm transfer error $1/T_{J+1}$ from an L^2 tail: at the degree where that error becomes small enough to subtract, the target's tail has fallen from 0.69 of the achievable error at $\alpha = 0$ to 0.25 at $\alpha = 1$. The floor is therefore tight exactly where the target's spectrum decays slowly, which is where post-selection pays; the gap lies where it does not, as the next paragraph shows.

A single cost axis tells the same story, under a convention that we state because the conclusion depends on it: cost is counted as encodings × mean shot overhead × qubits, with every unit weighted equally. Post-selection wins on encodings alone (two encodings do the work of five at $\varepsilon = 0.25$ on a square wave; five do the work of 27 at $\varepsilon = 0.11$), and on encodings × shots by 1.1–1.7×, with a margin that shrinks as precision rises. With the extra qubit counted, the cheapest unitary model is cheaper at every precision we tested. Post-selection pays only when depth is the binding constraint (Fig. 3c).

On ibm_kingston the excess error of the post-selected models over the ideal grew as the floor fell from 0.5 to 0.2 and 0.05: +0.011, +0.022 and +0.052 at $E = 3$, and +0.007, +0.015 and +0.033 at $E = 2$.

Supplementary Note 4: Learning-level tests

At a matched budget of 16 encodings, a deep re-uploading circuit beat a one-layer circuit with a neural head on 8 shots of one qubit by 0.48 and 0.61 in test R^2 on two of three sharp classification tasks (pre-registered; five seeds), and the trained models kept 96–99% of that gap on ibm_fez. That hybrid was limited by its single readout bit rather than by shots; read on both qubits at the same 16 encodings, it matched or beat depth on four of five tasks. In a dose-response experiment in feature width the trained two-qubit deep circuit never beat it, although the hybrid's shot price grew from 0.002 to 0.110 as the features sharpened (all pre-registered). That deep circuit failed even on targets where a theorem caps every 8-shot hybrid, such as $(\cos 6x_1 + \cos 6x_2)/2$ at $R^2 \le 0.665$ (the Gaussian coefficient bound $e^{-\omega^2/2V}$ with $V = 8$ per feature; the sharper $\cosh(b)^8 e^{-6b}$ bound of Theorem 1 caps $|c_6|$ at 0.080 and R^2 at 0.54). A post-hoc analysis showed that its narrow layers (single-qubit rotations and one CNOT) did not contain the target at all (best R^2 0.17 over 16 restarts with second-order polishing), while general two-qubit layers or one qubit per feature did (1.000). Of the other narrow-layout losses, one target was exactly representable by the hybrid and one was sample-limited at 300 points, so no model could beat the hybrid by the pre-registered margin. A model that contains its target and is fitted to it wins by construction, so the replication below tests only whether gradient training finds that optimum. The class-optimum comparison used a sharp step $\tanh(\cos x/w)$, $w = \frac{1}{64}$, with 400 training points and a budget of 64 encodings. A one-qubit circuit fitted by bounded least squares and compiled into a circuit with exactly T encodings (compile residual up to 3×10^{-3} in six of twelve cells, a missed criterion; the errors quoted are those of the compiled circuit) had test RMSE 0.060, against 0.186 for the equal-scale hybrid with its optimal head and circuit; at 16 encodings the ratio was 0.61 (pre-registered, three seeds).

Under gradient training, plain Adam on the same one-qubit re-uploading circuit reached 0.094 against 0.191 for the trained hybrid (a descriptive comparison, not a pre-registered hypothesis). Multi-scale shots with a neural head reached 0.127, beating equal-scale shots as pre-registered. A shallow circuit with seven rescaled encodings, or with trainable scales, did not beat the hybrid: rescaling gives range but few free coefficients, and gradient descent kept the scales near equal (pre-

registered; both failed).

A replication on fresh seeds confirmed this diagnosis. Its pre-registration, written afterwards, fixed the ansatz and target identified there and added one new non-additive target. On $(\cos 6x_1 + \cos 6x_2)/2$ a circuit with one qubit per feature and eight re-uploads reached test $R^2 = 1.000$, against -0.07 for the 8-shot hybrid and above that hybrid's theorem ceiling of 0.665 (0.54 with the cosh bound); on the $k = 8$ analogue it reached 0.46 against a ceiling of 0.14; and on the non-additive $\cos 6(x_1 + x_2)$, which the separable layout cannot express, a circuit with general two-qubit layers reached 0.98 against -0.09 (ceiling 0.38; all criteria passed). The shot price of a classical head and the depth advantage are therefore both measurable in trained models. On the evidence so far, whether gradient training realises the depth advantage depends on whether the ansatz contains the target, not on the depth. This is a diagnosis from one case, replicated on fresh seeds with a layout chosen after the failure, and not a tested law.

Supplementary Note 5: Separable limit: hardware and learning tests

On ibm_fez the separable limit holds on 24 qubits. With x encoded once on each of $m \in \{2, \ldots, 24\}$ qubits of a chain, the slope of the best bounded head on a product input stayed below $\sqrt{m}$ at every size (3.07 at $m = 24$, against 4.90), and so did the slope when the chain's CNOTs followed the encoding (2.16 at $m = 12$). The same CNOTs placed before the encoding prepare a GHZ-type input whose slope beat $\sqrt{m}$ at every size, by up to 2.1 times at $m = 12$ (7.41 against 3.46), after which decoherence took the gain back (1.55 times at $m = 24$); all pre-registered criteria passed.

On IonQ Forte-1 a bound that does not depend on phase offsets gave slopes of at least 3.71, 7.16 and 10.51 for $m = 4$, 8 and 16, but that bound was computed after the run, several blocks returned fitted visibilities above 1 and the two $m = 16$ replicates disagreed by ten standard errors, so we report the IonQ run as inconclusive rather than as a pass. The first IonQ run failed because phase offsets differed from block to block.

In a pre-registered learning test on a jump with $m = 4$ and 8 single-upload qubits, every trained separable model stayed above the exact separable capacity, which the theorem nearly attains (within 6% at $m = 8$). The models with the entangling layer before the encoding, however, ended far above their own capacity (grid MSE 0.12–0.17 against 0.062 at $m = 8$), so the learning criteria failed. Identity initialisation, layerwise growth, restarts and second-order optimisation found no better critical point (best 0.14). At these sizes the separable penalty is small, and gradient training of entangle-first circuits does not collect it.

Supplementary Note 6: Hardware run details

Exchange rate. The 128 measurement points of the exchange-rate run (Fig. 2a) are the Nyquist limit of a degree-64 polynomial and under-sample its error (ideal 0.067 on that grid against 0.093 on a fine grid), so the $T = 64$ point is reported but not fitted.

Multi-scale shots on IQM Garnet. The replication ran every setting (s, ϕ) as one job on all 20 qubits of Garnet, each qubit carrying one of 120 cell-centred x values, with the x-to-qubit map rotated across repeats so that qubit-specific readout bias does not map onto x (1,024 shots per job; 4,096 shots per x at $s = 1$, 1,024 at $s \geq 2$). The visibility of $R_Y(sx)$ lay between 0.953 and 0.958 for every s from 1 to 16, and every value lay within 0.011 of the prediction. For Open Quantum runs, counts are returned aggregated by bit string; analyses that block consecutive shots shuffle

the expanded shots with a fixed seed first (counts are exchangeable; results were stable over eight seeds). The first scoring of this run failed two criteria because expanding the aggregated counts had sorted the shots; it was rescored after the correction. Reshuffling the aggregated counts with eight different seeds moves the equal-scale exponent between 0.25 and 0.30 and the multi-scale one between 0.56 and 0.58.

Noisy allocation. In the first allocation run the 4 μs idle after each encoding also added a coherent phase that shifted x by about 0.1 rad, which the noise model did not include. The redesigned run replaced the idle by virtual random Z flips.

Qualified, inconclusive and missed runs. Two runs passed with a qualification: the trained-model run on ibm_fez passed its criteria but its hybrid comparison was confounded by a one-bit readout (a light-cone-fixed repeat passed cleanly), and the IQM multi-scale run was rescored as described above. The IonQ phase-robust run met its criteria, but with fitted visibilities above 1 in several blocks and a phase-robust bound computed after the run, so we count it as inconclusive (Supplementary Note 5). Three missed: a comparison against a five-encoding unitary model ended in a tie within noise, the noisy-allocation test missed its margin by 0.007 because of the coherent idle phase (its redesign with exact dephasing passed), and the first IonQ run failed because phase offsets differed from block to block.

Supplementary Note 7: Pre-registration record

For the earliest run reported here (September 2026) the note and prediction file were written before the job (the prediction file carries a creation time minutes before the job) but entered version control only the following morning, so that run is not pre-registered in the version-control sense. The same commit froze the first three runs of the companion paper, likewise after their jobs [1]. Where a learning experiment was re-run under a revised criterion after a miss (the smooth-target test ran four times, with its smooth-target criterion weakened from a two-sided to a one-sided gate and then demoted to a reported quantity; the entangle-first test ran three times, with its capacity floor corrected after the third miss), every version is on record and counted. Supplementary Table 1 gives the commit record of every hardware run.

Supplementary References

[1] Sulimov, P. & Lehmann, C. Hidden selection rules and amplitude ceilings in quantum Fourier models. Preprint at https://arxiv.org/abs/2611.01314 (2026).

[2] Keane, M. S. & O'Brien, G. L. A Bernoulli factory. *ACM Trans. Model. Comput. Simul.* **4**, 213–219 (1994).

[3] Nacu, S. & Peres, Y. Fast simulation of new coins from old. *Ann. Appl. Probab.* **15**, 93–115 (2005).

[4] Qian, W., Riedel, M. D. & Rosenberg, I. Uniform approximation and Bernstein polynomials with coefficients in the unit interval. *Eur. J. Combin.* **32**, 448–463 (2011).

[5] Dale, H., Jennings, D. & Rudolph, T. Provable quantum advantage in randomness processing. *Nat. Commun.* **6**, 8203 (2015).

Supplementary Table 1: Pre-registration record of the runs in Table 2: the commit that froze each note (short hash in the project's version-control history), its time, the submission time of the job read from the stored result (first job for Open Quantum runs), and the margin. A negative margin means the note entered version control after the job. Runs that share a commit, here or with the companion [1], were submitted together, each with its own job identifier.

Experiment	commit	committed (UTC)	job (UTC)	margin
Post-selection, trained models	`6c4b3b4`	2026-09-26 07:23	2026-09-25 16:21	−15 h
Price of post-selection	`bbb181b`	2026-10-14 20:29	2026-10-14 23:28	3.0 h
Depth-for-shots exchange rate	`8898bb7`	2026-10-15 22:35	2026-10-15 22:35	20 s
Resolution limit	`c70ffa0`	2026-10-16 12:26	2026-10-16 12:26	22 s
Adaptive dynamic circuit	`fc9c2c6`	2026-10-16 12:40	2026-10-16 12:41	39 s
Price of post-selection at $E = 4$	`c70ffa0`	2026-10-16 12:26	2026-10-16 12:26	22 s
Copies vs coherence, 2–24 qubits	`3586daf`	2026-10-16 13:11	2026-10-16 14:07	56 min
Noisy allocation law	`7f5099b`	2026-10-16 13:22	2026-10-16 14:07	45 min
Allocation law, exact dephasing	`bcef005`	2026-10-17 13:24	2026-10-17 13:25	88 s
Trained models on hardware	`fbbf8d5`	2026-10-16 13:41	2026-10-16 14:08	27 min
Trained models, light-cone-fixed	`a42e722`	2026-10-17 16:21	2026-10-17 16:22	38 s
Rescaled encodings, multi-scale shots (IBM)	`3c38321`	2026-10-19 13:09	2026-10-19 13:09	5 s
Copies vs coherence (IonQ)	`19405ac`	2026-10-17 10:25	2026-10-17 10:26	12 s
Separable limit, phase-robust (IQM, IonQ)	`4920154`	2026-10-21 20:00	2026-10-21 20:01	25 s
Rescaled encodings, multi-scale shots (IQM)	`fc56ac6`	2026-10-21 20:05	2026-10-21 20:06	34 s

[6] Lorentz, G. G. The degree of approximation by polynomials with positive coefficients. *Math. Ann.* **151**, 239–251 (1963).

[7] Holtz, O., Nazarov, F. & Peres, Y. New coins from old, smoothly. *Constr. Approx.* **33**, 331–363 (2011).

[8] Stahl, H. Best uniform rational approximation of $|x|$ on $[-1, 1]$. *Russian Acad. Sci. Sb. Math.* **76**, 461–487 (1993).