Depolarizing Noise

Variational Quantum Algorithms · Module 9 — Hardware Noise Models

Node 9.2

Position in Knowledge Graph

System: Variational Quantum Algorithms **Structural Domain:** Hardware Noise Models **Node:** 9.2

Depolarizing noise is the **workhorse noise model** of VQA theory. It is not the most physically realistic — real hardware has a mix of amplitude damping, dephasing, coherent errors, and crosstalk — but it is the most **mathematically tractable**, the most **symmetric**, and the most **commonly assumed** in theoretical analyses.

Understanding depolarizing noise deeply is worthwhile for several reasons:

1. It gives a **clean upper bound** on the effects of noise — if your VQA can handle depolarizing at rate p, it can usually handle more structured noise at similar rates.
2. It is the **simplest to simulate** — a single rate parameter controls everything.
3. It **reveals the depth-fidelity tradeoff** in closed form.
4. It is the **canonical example** for barren plateau theory under noise (Module 4.5, Wang et al. 2021).

This node derives the single-qubit depolarizing channel carefully, covers the multi-qubit generalizations, works through the composition rules for deep circuits, and explains why depolarizing noise is the "uniform" benchmark against which other noise models are measured.

The Single-Qubit Depolarizing Channel

The single-qubit depolarizing channel at rate p is defined as

$$\mathcal{D}_p(\rho) = (1-p)\rho + \frac{p}{3}\left(X\rho X + Y\rho Y + Z\rho Z\right).$$

Interpretation: with probability $1-p$, nothing happens (identity). With probability p, a uniformly random Pauli error X, Y, Z is applied. The factor $1/3$ splits the error probability evenly among the three Pauli types.

Kraus operators:

$$K_0 = \sqrt{1-p}\,I, \quad K_X = \sqrt{p/3}\,X, \quad K_Y = \sqrt{p/3}\,Y, \quad K_Z = \sqrt{p/3}\,Z.$$

Alternative form. Using the identity $(X\rho X + Y\rho Y + Z\rho Z + \rho)/4 = I/2 \cdot \mathrm{Tr}(\rho)$, the channel can be rewritten as

$$\mathcal{D}_p(\rho) = (1 - 4p/3)\rho + (4p/3) \cdot I/2.$$

Letting $\lambda = 4p/3$, this is $(1-\lambda)\rho + \lambda \cdot I/2$ — a **convex combination of the original state and the maximally mixed state**. This is the "depolarization" intuition: the state gets mixed with noise at rate λ.

Rescaling convention. Some authors use p to mean the total error probability (p above), others use p to mean the λ parameter. Be careful with the conventions.

Effect On Observables

Because the depolarizing channel is symmetric under Pauli conjugation, its action on Pauli observables is particularly simple:

$$\mathcal{D}_p^\dagger(P) = (1 - 4p/3)P \quad \text{for any non-identity Pauli } P.$$

And $\mathcal{D}_p^\dagger(I) = I$ (identity is preserved).

Consequence. Every Pauli expectation value is *uniformly shrunk* by a factor of $1 - 4p/3$. The measurement is unchanged in direction — only in magnitude.

For VQAs: if the cost Hamiltonian is $H = \sum_i c_i P_i$ (decomposed into Paulis), then the noisy cost is

$$C_{\text{noisy}}(\boldsymbol{\theta}) = (1 - 4p/3) \cdot C_{\text{ideal}}(\boldsymbol{\theta}) + 4p/3 \cdot \operatorname{Tr}(H)/d.$$

The noisy cost is an **affine transformation** of the ideal cost: scaled by $(1 - 4p/3)$ and shifted by a constant. **Importantly, this affine transformation preserves the *location* of the minimum** — it only changes the value at the minimum and the slope of the landscape.

This is why depolarizing noise is often called "benign" for VQA optimization: the optimizer still finds the right parameters, just with a reduced signal-to-noise ratio. More problematic noise types (amplitude damping, coherent errors) actually *move* the minimum.

Multi-Qubit Depolarizing Channels

Several distinct generalizations exist for n qubits:

1. Independent single-qubit depolarizing

Each qubit experiences depolarizing noise independently:

$$\mathcal{D}_p^{\otimes n}(\rho).$$

Properties: factors over qubits; each single-qubit reduced state is depolarized independently; entanglement is *not* directly affected by the channel structure but does decay because of the per-qubit depolarization.

This is the most physically realistic model when the dominant noise is uncorrelated single-qubit error.

2. Global depolarizing

$$\mathcal{D}_p^{\text{global}}(\rho) = (1 - p)\rho + p \cdot I/2^n.$$

With probability p, the whole n-qubit state is replaced by the maximally mixed state; with probability $1 - p$, nothing happens.

Properties: the simplest theoretical model; often used in barren plateau proofs; rarely physically accurate, but gives clean bounds.

3. Correlated two-qubit depolarizing

Applied to pairs of qubits where two-qubit gates act:

$$\mathcal{D}_p^{(2)}(\rho) = (1-p)\rho + \frac{p}{15}\sum_{k=1}^{15} P_k \rho P_k,$$

where the sum is over the 15 non-identity two-qubit Paulis. Used to model two-qubit gate errors.

Practical hardware modeling: real hardware is usually approximated as **single-qubit depolarizing after each single-qubit gate plus two-qubit depolarizing after each two-qubit gate**, with rates determined by the respective gate fidelities.

The Depth-Fidelity Tradeoff

A central calculation: **how does fidelity decay with circuit depth under depolarizing noise?**

For a circuit with L gates, each followed by depolarizing noise at rate p, the total channel on any Pauli observable is

$$\mathcal{D}^L(P) = (1-4p/3)^L P.$$

The signal decays exponentially in depth. Define the **effective depth** at which the signal is below the shot noise floor:

$$(1-4p/3)^{L^*} = \frac{1}{\sqrt{N_{\text{shots}}}}.$$

Solving: $L^* = -\log(\sqrt{N_{\text{shots}}})/\log(1-4p/3) \approx (3/4p)\cdot\log(\sqrt{N_{\text{shots}}})$.

Numerical examples:

- $p = 10^{-3}$, $N_{\text{shots}} = 10^6$: $L^* \approx 750 \cdot 6.9 \approx 5200$ gates.
- $p = 10^{-2}$, $N_{\text{shots}} = 10^6$: $L^* \approx 75 \cdot 6.9 \approx 520$ gates.
- $p = 10^{-1}$, $N_{\text{shots}} = 10^6$: $L^* \approx 7.5 \cdot 6.9 \approx 52$ gates.

Implication: for current hardware ($p \sim 5 \times 10^{-3}$ for two-qubit gates), useful VQA circuits are limited to ~1500 two-qubit gates. Beyond that, the signal is lost in noise and no amount of shots recovers it.

This is the **fundamental depth barrier** for NISQ-era VQAs, and it is tight — depolarizing noise is essentially the best-case scenario.

Depolarizing Noise And Barren Plateaus

Wang et al. (2021) showed that depolarizing noise creates **noise-induced barren plateaus** in addition to the unitary plateaus of Module 4.

Theorem (Wang et al. 2021). For a VQA with depth L under single-qubit depolarizing noise at rate p per gate, the gradient variance is bounded by

$$\mathrm{Var}[\partial_j C_{\text{noisy}}] \leq C_0 \cdot (1-4p/3)^{2L},$$

where C_0 is a constant depending on the cost.

Implication: the gradient variance decays *exponentially in depth times noise rate*. Even ansätze that are trainable in the noiseless case become untrainable at sufficient depth under noise.

The noise-induced plateau depth: setting the variance bound equal to the shot noise floor $1/N_{\text{shots}}$, the plateau sets in at depth

$$L_{\text{NP}} \approx \frac{1}{4p/3} \cdot \frac{1}{2} \log(N_{\text{shots}} C_0^{-1}).$$

For $p = 5 \times 10^{-3}$, $N_{\text{shots}} = 10^6$, this is $L_{\text{NP}} \approx 500 - 1000$ gates — consistent with the depth limit from the fidelity argument.

The fundamental point: the depth limit from fidelity decay and the depth limit from noise-induced plateaus are *the same limit*, approached from two different angles. Depolarizing noise caps the usable circuit depth in exactly the regime where shot noise starts to dominate.

Why Depolarizing Is The Worst Case (In Some Sense)

A common claim: depolarizing noise is the "worst case" among unital noise channels at fixed average error rate.

Precise statement: for a fixed average gate fidelity, the depolarizing channel has the **maximum entropy** (most random) distribution of error types. Any structured noise (e.g., only phase errors, or only bit flips) gives a more concentrated error distribution.

Operational consequence: mitigation techniques that work for depolarizing noise work for more structured noise. Conversely, analyses that fail under depolarizing noise are unlikely to succeed under arbitrary noise.

Caveat: this is only true for **unital** channels (channels that preserve the maximally mixed state). **Non-unital** channels like amplitude damping are fundamentally different — they have a *preferred direction* in state space, and the "worst case" framing doesn't apply. Amplitude damping is discussed in node 9.3.

So: **depolarizing noise is the worst case among unital noise models**, and it is a useful upper bound for any noise that can be approximated as unital.

The Pauli Twirling Trick

A useful technique: **Pauli twirling** converts an arbitrary single-qubit channel into a depolarizing channel of equivalent fidelity.

Procedure. Before and after the noisy gate, apply a uniformly random Pauli (sampled from $\{I, X, Y, Z\}$). Average over many shots.

Effect. The average channel becomes

$$\bar{\mathcal{E}}(\rho) = \frac{1}{4} \sum_P P \mathcal{E}(P \rho P) P.$$

For a general single-qubit channel, this averages out to a **depolarizing channel** with the same average gate fidelity as the original.

Why this matters: twirling converts "messy" noise into "clean" depolarizing noise. Theoretical analyses that assume depolarizing noise are then *actually valid* for the twirled system, even if the raw hardware noise is more complex.

Cost: twirling doesn't reduce the noise — it just symmetrizes it. You pay for the simplification with the extra Pauli gates at each noise event.

Pauli twirling is the foundation of **randomized benchmarking** (the standard gate fidelity measurement protocol) and is used implicitly in many VQA mitigation schemes.

Depolarizing Noise In Practice

A few practical observations:

1. Real hardware is not purely depolarizing. Superconducting qubits have amplitude damping (energy loss) as the dominant noise; trapped ions have Rabi frequency errors; neutral atoms have loss and Rydberg decay. Depolarizing is an approximation.

2. The approximation is often good enough. For theoretical estimates and order-of-magnitude calculations, depolarizing is sufficient. Deviations matter at the factor-of-2 level, not the factor-of-100 level.

3. Simulation vs hardware. Depolarizing noise is the standard model in simulators (Qiskit Aer, Cirq, PennyLane). When you run a "noisy simulation," you're typically getting depolarizing unless you specify otherwise.

4. Randomized benchmarking returns depolarizing rates. The standard gate-fidelity measurement protocol assumes the noise is depolarizing (or has been twirled). Reported gate fidelities are depolarizing fidelities, which may differ from the "true" fidelity of the underlying physical noise.

5. Mitigation techniques built for depolarizing often extend. Zero-noise extrapolation (node 9.5) was originally developed for depolarizing noise and extends to broader noise classes. Similarly, probabilistic error cancellation is depolarizing-first.

So: depolarizing noise is both the theoretical default *and* the practical default, and knowing it deeply pays off.

Structural Role

This node is the **canonical noise example** of Module 9. It defines the single- and multi-qubit depolarizing channels, derives their action on observables, computes the depth-fidelity tradeoff explicitly, and establishes the connection to noise-induced plateau theory.

It is the benchmark against which other noise models (amplitude damping in 9.3, coherent errors in later sections) are compared.

Relations to Other Concepts

- **Nielsen-Chuang** — standard reference for the depolarizing channel.
- **Wang et al. (2021)** — noise-induced barren plateau theorem.
- **Randomized benchmarking (Magesan et al. 2011)** — fidelity measurement assuming depolarizing noise.
- **Pauli twirling** — converting arbitrary noise to depolarizing.
- **Module 4.5 (Noise-induced plateaus)** — the qualitative version.

Worked Example — Depth Limit For A VQA On 2026 Hardware

Consider a 20-qubit VQE on a chemistry Hamiltonian, running on superconducting hardware with: - Single-qubit gate fidelity: 99.9% $\to p_1 = 10^{-3}$. - Two-qubit gate fidelity: 99.5% $\to p_2 = 5 \times 10^{-3}$. - Shot budget: 10^6 per cost evaluation.

Ansatz: Hardware-Efficient Ansatz with 10 layers, each layer having 20 single-qubit gates + 19 two-qubit gates = 39 gates/layer. Total: 390 gates per circuit, of which 200 are two-qubit.

Total error: approximating each gate as an independent depolarizing channel, the circuit fidelity is

$$F \approx (0.999)^{200} \cdot (0.995)^{200} = 0.819 \cdot 0.367 = 0.30.$$

Only 30% of the signal survives. Combined with shot noise at $1/\sqrt{10^6} = 10^{-3}$, the effective SNR is $0.30/10^{-3} = 300$ — good enough for training.

Doubled depth (20 layers): $F \approx 0.30^2 = 0.09$. SNR ≈ 90. Still workable but degraded.

Quadrupled depth (40 layers): $F \approx 0.30^4 = 0.008$. SNR ≈ 8. Near the noise floor.

Conclusion: for this hardware, 10-20 layer circuits are usable, 40+ layers are beyond the noise-limited depth. This is consistent with the state-of-the-art 2026 experimental VQE demonstrations, which typically use 5-20 layer ansätze.

Visualization

Pending. A log-scale plot of gate fidelity vs circuit depth for several values of p (from 10^{-4} to 10^{-2}) showing the exponential decay. The shot noise floor is a horizontal line; the intersections give the maximum usable depth. Caption: "Depolarizing noise sets the NISQ depth ceiling."

Exercises

1. Compute the Pauli transfer matrix of a single-qubit depolarizing channel at rate p.
2. Show that the composition of two depolarizing channels at rates p_1, p_2 is a depolarizing channel. What is its rate?
3. (VQA) For a 30-qubit VQA with 500 two-qubit gates at fidelity 99.8%, estimate the circuit fidelity and the shot budget needed for a useful cost estimate.

Research Extensions

- **Non-uniform depolarizing rates** — different rates for different gates or qubits.
- **Coherent depolarizing deviations** — what if the noise is almost-depolarizing but not exactly?
- **Active depolarization** — intentionally adding depolarizing noise as a form of regularization for training stability.
- **Depolarizing-equivalent circuit compilation** — transforming circuits so their effective noise is depolarizing.

Cross-links to Other Courses

- **Open Quantum Systems** — Lindblad master equations for depolarizing noise.
- **Randomized Benchmarking** — the standard gate fidelity protocol.

- **Module 4.5 (Noise-induced plateaus)** — the qualitative version.
- **Module 9.5 (Zero-noise extrapolation)** — the main mitigation technique for depolarizing noise.

Variational Quantum Algorithms · Module 9 — Hardware Noise Models · Node 9.2
Concepts: quantum-noise, decoherence, density-matrix, quantum-channels
Prerequisites: 9.1
Enables: 9.4, 9.5, 9.8

hbar.university — own.your.cognition