Expressivity Theory

Variational Quantum Algorithms · Module 8 — Expressivity Vs Trainability

Node 8.1

Position in Knowledge Graph

System: Variational Quantum Algorithms **Structural Domain:** Expressivity vs Trainability **Node:** 8.1

This is the **opening node** of Module 8 — the module that develops, in quantitative form, the central tradeoff that has been hovering implicitly behind every module since Module 4: **the more states an ansatz can reach, the harder it is to optimize over**.

Before we can talk about the tradeoff, we need to define **expressivity** with enough precision to be useful. That is the job of this node and node 8.2.

The intuitive picture: an ansatz $U(\boldsymbol{\theta})$ defines a *family* of quantum states $\{|\psi(\boldsymbol{\theta})\rangle\}$ as $\boldsymbol{\theta}$ ranges over its domain. Expressivity is "how big the family is" — formally, how much of the Hilbert space the family covers, in some appropriate measure.

This node introduces the formal definitions, explains the multiple ways expressivity can be measured, and sets up the trainability counterpart in node 8.4.

What "Expressivity" Means

Several closely-related but distinct definitions appear in the literature:

1. Reachability: The set of states $\mathcal{R}(U) = \{|\psi(\boldsymbol{\theta})\rangle : \boldsymbol{\theta} \in \Theta\}$. The "reachable manifold" of the ansatz. Expressivity is, intuitively, the size of $\mathcal{R}(U)$.

2. Volume: For an ansatz with P parameters, the manifold $\mathcal{R}(U)$ has dimension at most P. The "volume" of $\mathcal{R}(U)$ (in the Fubini-Study metric) is a measure of how much of the state space it covers.

3. Distribution distance: Compare the distribution induced on states by random parameter values to the Haar measure on the full state space. An ansatz is "highly expressive" if these distributions are close, "low expressivity" if the induced distribution is far from Haar.

4. Approximation error: For a given target state $|\phi\rangle$, the minimum distance from $|\phi\rangle$ to any state in $\mathcal{R}(U)$. A low approximation error (for any target) means high expressivity.

5. Entanglement coverage: The range of entanglement entropies that states in $\mathcal{R}(U)$ can achieve. Higher coverage = higher expressivity (in the entanglement sense).

These definitions are not equivalent. An ansatz can be highly expressive in one sense and not in another. The thesis (and most recent VQA literature) treats expressivity as a *family* of related but distinct quantities, each useful for different purposes.

Why Expressivity Matters

Expressivity matters for two reasons that pull in opposite directions:

1. You need enough expressivity to represent the answer. If your target state isn't in the reachable manifold of your ansatz, no amount of optimization will find it. The ansatz must be expressive enough that the ground state (or a good approximation to it) lies in $\mathcal{R}(U)$.

2. You need expressivity to be limited or you'll hit a barren plateau. An ansatz that can reach all of Hilbert space is, by Module 4's analysis, exactly the ansatz that has barren plateaus. Too much expressivity is fatal for trainability.

The right amount of expressivity is "just enough." Enough to contain the answer, no more. Designing an ansatz to hit this sweet spot is what Module 2's discussion of problem-inspired ansätze (UCCSD, HVA, ADAPT-VQE) was trying to achieve.

Reachable Manifold Dimension

The simplest expressivity measure is the **dimension** of the reachable manifold.

For a parameterized circuit with P parameters, the manifold $\mathcal{R}(U)$ has dimension at most P. It can be **less** than P if there are gauge redundancies (Section 3.4) or if some parameter combinations are degenerate.

For example, the ansatz $R_X(\theta_1)R_X(\theta_2)|0\rangle$ has 2 parameters but dimension 1 (because only $\theta_1 + \theta_2$ matters).

The full state space of n qubits is a manifold of dimension $2 \cdot 2^n - 2$ (real coordinates on the projective Hilbert space). For an ansatz to be "fully expressive" in the dimension sense, it needs at least this many parameters.

For $n = 4$, that's 30 parameters minimum. For $n = 10$, 2046. For $n = 20$, about 2 million. **Truly universal ansätze are exponentially expensive in parameter count.**

In practice, useful ansätze have parameter counts polynomial in n and reach a polynomially-large submanifold of the exponentially-large state space. This "polynomial expressivity" regime is where almost all VQA work lives.

The Sim-Johnson-Killoran Expressibility

The most widely-used quantitative expressivity measure for VQA was introduced by **Sim, Johnson, Killoran (2019)**. The idea: compute the distribution of fidelities between pairs of random ansatz states, and compare to the Haar distribution.

Let $P_{\text{Haar}}(F)$ be the Haar distribution of $F = |\langle\psi_1|\psi_2\rangle|^2$ for two Haar-random states. This has known closed form: $P_{\text{Haar}}(F) = (d-1)(1-F)^{d-2}$ where $d = 2^n$.

Let $P_{\text{ansatz}}(F)$ be the analogous distribution for two random ansatz states $|\psi(\boldsymbol{\theta}_1)\rangle, |\psi(\boldsymbol{\theta}_2)\rangle$ with $\boldsymbol{\theta}_i$ drawn uniformly.

Expressibility is the KL divergence:

$$\text{Expr} = D_{KL}\big(P_{\text{ansatz}}(F) \parallel P_{\text{Haar}}(F)\big).$$

Lower KL divergence = ansatz distribution is closer to Haar = "more expressive." Higher KL divergence = ansatz is structured = "less expressive."

This is the **standard expressivity metric** in the VQA literature. Most ansatz comparison papers report it.

Practical computation: sample many parameter pairs, compute the empirical fidelity distribution, and bin it. Compute the KL divergence to the analytic Haar distribution.

Expressivity Of Common Ansätze

The Sim-Johnson-Killoran metric has been computed for many standard ansätze. Some representative results (for $n = 4$ qubits):

- **Hardware-Efficient Ansatz, deep:** expressibility close to 0 (Haar-like) — fully expressive.
- **Hardware-Efficient Ansatz, 1 layer:** expressibility ~ 0.1 — limited but nontrivial.
- **UCCSD:** expressibility ~ 0.5 — strongly restricted by particle-number conservation.
- **Tensor Network ansatz (MPS):** expressibility depends on bond dimension; for low bond dimension, very limited.
- **Random Pauli circuit, depth 1:** expressibility ~ 1.0 — almost no coverage.

The trend: **as you add structure (symmetry, sparsity, restricted connectivity), expressibility decreases**. As you remove structure (random gates, deep circuits, generic parameterization), it increases.

This is the operational measurement of the tradeoff Module 4 discussed qualitatively.

Expressivity And Entanglement

Expressivity correlates with — but is not the same as — entanglement coverage.

A high-expressivity ansatz typically reaches highly entangled states (because Haar-random states are highly entangled, by Page's theorem from Section 4.3). A low-expressivity ansatz typically reaches a restricted entanglement range.

But the relationship is not exact:

- A product-state ansatz has zero entanglement and zero expressivity.
- A maximally-entangled (Bell-state-only) ansatz has high entanglement but low expressivity (small reachable set).
- A Haar-random ansatz has high entanglement and high expressivity.

The two axes are related but distinct. Module 4.3 framed barren plateaus in terms of entanglement; this module frames them in terms of expressivity. Both are valid; they capture overlapping but not identical structure.

For practical purposes, **entanglement-based reasoning** (Module 4.3) is more physical and gives you mitigation strategies; **expressivity-based reasoning** (this module) is more mathematical and gives you metrics for ansatz comparison.

Lie Algebraic Expressivity

A more recent and theoretically clean view: expressivity is governed by the **dynamical Lie algebra** of the ansatz generators.

Let $U(\boldsymbol{\theta}) = e^{-i\theta_L H_L} \cdots e^{-i\theta_1 H_1}$ for Hermitian generators $H_1, \ldots, H_L$. The dynamical Lie algebra (DLA) $\mathfrak{g}$ is the smallest Lie algebra containing all the H_j and closed under commutators $[H_i, H_j]$.

Theorem: the reachable manifold $\mathcal{R}(U)$ is contained in the orbit $\{e^{iG}|0\rangle : G \in \mathfrak{g}\}$. The dimension of this orbit is at most the dimension of $\mathfrak{g}$.

Implications: - Small DLA $\to$ small reachable manifold $\to$ low expressivity $\to$ high trainability. - Large DLA $\to$ large reachable manifold $\to$ high expressivity $\to$ low trainability.

For example: a UCCSD ansatz has a DLA equal to the (small) algebra of particle-number-conserving operators. A deep HEA has a DLA equal to the full $\mathfrak{u}(2^n)$, the Lie algebra of all n-qubit unitaries — exponentially large.

Ragone et al. (2024) developed a Lie-algebraic theory of barren plateaus that uses this DLA structure to give precise expressivity bounds. The DLA is the cleanest theoretical proxy for "how expressive is this ansatz" available in 2026.

Structural Role

This node is the **definitions** node of Module 8. It introduces the multiple senses of "expressivity" and previews the trainability counterpart in node 8.4.

The metric definitions in 8.2 and the Lie-algebraic view sketched here are the two technical foundations Module 8 will use throughout.

Relations to Other Concepts

- **Sim-Johnson-Killoran (2019)** — the standard expressibility metric.
- **Ragone et al. (2024)** — the Lie-algebraic theory.
- **Page's theorem** — Haar-random state entanglement.
- **Module 2 (Ansatz design)** — the practical context for the theory.
- **Module 4 (Barren plateaus)** — the trainability obstacle expressivity creates.

Worked Example — Counting Reachable Dimensions

Consider $|\psi(\theta_1, \theta_2)\rangle = R_Y(\theta_1)R_X(\theta_2)|0\rangle$ on 1 qubit.

The state space of 1 qubit is the Bloch sphere S^2, dimension 2. The ansatz has 2 parameters.

Compute reachability: $R_X(\theta_2)|0\rangle = \cos(\theta_2/2)|0\rangle - i\sin(\theta_2/2)|1\rangle$. Then $R_Y(\theta_1)$ rotates this around the Y-axis.

By inspection, the ansatz can reach every state on the Bloch sphere — it is universal for 1 qubit. Reachable manifold has dimension 2.

For 2 qubits, the state space has dimension 6. A 2-qubit ansatz with only single-qubit rotations $R_Y(\theta_1)\otimes R_Y(\theta_2)$ can only reach product states — a 2-dimensional submanifold of the 6-dimensional state space. Adding a CNOT after the rotations expands this to a 4-dimensional submanifold (entangled states reachable from product states by one CNOT). Adding more layers expands further, eventually reaching the full 6-dimensional manifold.

This dimension counting is the simplest expressivity calculation. The Sim-Johnson-Killoran metric, the Lie-algebraic dimension, and other measures all generalize this idea to higher-dimensional state spaces.

Visualization

Pending. A 2D projection of the Bloch sphere with reachable regions of three different ansätze highlighted: (a) a product-only ansatz (small dot), (b) a 1-layer HEA (larger region), (c) a deep HEA (the entire sphere). Caption: "Expressivity is the volume the ansatz reaches."

Exercises

1. For $|\psi(\theta)\rangle = R_Y(\theta)|0\rangle$ (1 qubit, 1 parameter), what is the reachable manifold? What is its dimension?
2. For an ansatz with 5 parameters in 4 qubits (state space dimension 30), what is the maximum possible reachable manifold dimension? Is the ansatz universal?
3. (VQA) The Sim-Johnson-Killoran expressibility for a deep HEA approaches the Haar value. Show analytically that for a 1-qubit ansatz $R_Y(\theta_1)R_X(\theta_2)R_Y(\theta_3)$, the expressibility is exactly Haar. (Hint: this is universal for 1 qubit.)

Research Extensions

- **Tighter expressivity-trainability bounds** — pinning down the exact constants in the tradeoff.
- **Problem-aware expressivity** — measures that account for the specific Hamiltonian, not just the ansatz.
- **Computational expressivity** — relating expressivity to computational hardness of the optimization.
- **Expressivity under noise** — how hardware noise modifies the effective expressivity of an ansatz.

Cross-links to Other Courses

- **Lie Theory** — Lie algebras, Lie groups, dynamical Lie algebras of control systems.
- **Differential Geometry** — manifolds, dimension, embedding.
- **Statistical Learning Theory** — VC dimension, capacity, generalization (the classical analogue of expressivity).
- **Module 2 (Parameterized Circuits)** — the operational ansatz design.
- **Module 4 (Barren Plateaus)** — the trainability obstacle.

Variational Quantum Algorithms · Module 8 — Expressivity Vs Trainability · Node 8.1
Concepts: hilbert-space, parameterized-circuits, dimensionality
Prerequisites: 2.1, 2.9
Enables: 8.2, 8.3

hbar.university — own.your.cognition