Observability of Quantum States

Variational Quantum Algorithms · Module 7 — Control Theoretic View

Node 7.2

Position in Knowledge Graph

System: Variational Quantum Algorithms **Structural Domain:** Control-Theoretic View **Node:** 7.2

In control theory, **observability** is the question: given the outputs you can measure, can you reconstruct the underlying state? It is the dual of **controllability** (which input combinations can drive the system to which states) — observability asks what *information* you have about a state, not what states you can reach.

For VQAs, observability translates into a specific operational question: **does the cost function uniquely identify the optimal parameters, or are there directions in parameter space along which the cost is flat (or symmetric)?** If the cost is observability-deficient, then no amount of optimization will distinguish between equivalent parameter values — there is a *gauge symmetry* and the optimizer wanders aimlessly along the gauge orbit.

This matters for VQA design in several distinct ways: it explains certain "stuck" optimizer behaviors, motivates the choice of cost function, and connects to the broader question of *why some VQAs are easier to train than others.*

This node introduces classical observability, the quantum-state observability problem, the cost-function observability question for VQAs, and how to diagnose observability deficiencies in practice.

Classical Observability

In classical control, a linear system is

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}, \quad \mathbf{y} = C\mathbf{x},$$

where $\mathbf{x} \in \mathbb{R}^n$ is the state, $\mathbf{u} \in \mathbb{R}^m$ is the input, and $\mathbf{y} \in \mathbb{R}^p$ is the measured output.

Observability theorem (Kalman): the system is observable if and only if the **observability matrix**

$$\mathcal{O} = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix}$$

has rank n. Equivalently, no two distinct initial states $\mathbf{x}(0)$ and $\mathbf{x}'(0)$ produce identical output trajectories.

If the observability matrix has rank $r < n$, then there is an $(n - r)$-dimensional **unobservable subspace**: states that differ by a vector in this subspace are indistinguishable from output measurements.

This is the canonical setup. The **rank of the observability matrix** is the central quantitative measure.

Observability For Quantum States

The quantum analogue is more delicate. Quantum measurement has two distinct features:

1. Measurement collapses the state. Once you measure, the state changes. So the "output trajectory" of a single state is impossible — you can only measure one shot, then the state is gone.

2. Single-shot measurements are random. Even given a fixed observable, the result of a single measurement is sampled from a distribution.

So "observability" in the quantum sense means: **given the ability to prepare many copies of the same state and measure any of a fixed set of observables, can you reconstruct the state?**

The answer is: **yes, with sufficiently informative observables**. Specifically, a set of observables $\{O_1, O_2, \ldots, O_K\}$ is **informationally complete** if knowing all expectation values $\{\langle\psi|O_k|\psi\rangle\}_k$ uniquely identifies $|\psi\rangle$ (up to global phase).

Standard informationally complete sets:

- **Pauli decomposition:** all 4^n products of single-qubit Paulis $\{I, X, Y, Z\}^{\otimes n}$. Together they span all Hermitian operators on n qubits.
- **Symmetric informationally complete (SIC) POVMs:** d^2 outcomes that uniquely identify any d-dimensional state.
- **Mutually unbiased bases (MUBs):** $d + 1$ bases, each measurement gives d outcomes, together informationally complete.

Cost of informational completeness: 4^n Pauli observables (or $d^2 = 4^n$ SIC outcomes) — exponentially many. **Full quantum state reconstruction is exponentially expensive.**

In practice, you don't need full reconstruction. You only need enough observables to identify the *parameters of the ansatz*, which is a much smaller set (at most P numbers).

Cost-Function Observability For VQAs

The VQA-specific observability question:

Definition (cost observability). A cost function $C(\boldsymbol{\theta})$ is **observable** at $\boldsymbol{\theta}^*$ if no other parameter value gives the same cost in a neighborhood: $C(\boldsymbol{\theta}^* + \delta) \neq C(\boldsymbol{\theta}^*)$ for all $\delta \neq 0$ in some neighborhood.

More quantitatively: the cost is **first-order observable** if the gradient $\nabla C(\boldsymbol{\theta}^*)$ has full rank in directions that matter (i.e., is not zero along any direction you care about). It is **second-order observable** if the Hessian has full rank.

Observability deficiencies in VQAs:

1. **Gauge symmetries.** If the ansatz has parameter combinations that produce the same state (e.g., $R_Z(\theta_1)R_Z(\theta_2)$ only depends on $\theta_1 + \theta_2$), then the cost depends only on $\theta_1 + \theta_2$, and the perpendicular direction is unobservable. This is the *gauge orbit* problem.

2. **Symmetries of the cost.** If the cost is invariant under some parameter transformation (e.g., $\boldsymbol{\theta} \to -\boldsymbol{\theta}$ for some symmetric ansätze), then there are multiple equivalent minima and the cost can't distinguish them.

3. **Local degeneracies.** Even without explicit symmetries, the cost can have flat directions at specific points (e.g., saddles, ridges) where the gradient vanishes along some direction.

4. **Barren plateaus as an extreme case.** A barren plateau is a region where the gradient is *near zero in all directions* — the cost is unobservable in *every* direction simultaneously. This is the worst-case observability deficiency.

The Observability Matrix For VQAs

The classical observability matrix can be adapted to VQAs, with parameters playing the role of state.

For a cost function $C(\boldsymbol{\theta})$ viewed as a function on parameter space, the local observability is captured by the **Hessian matrix** $H_{ij}(\boldsymbol{\theta}) = \partial^2 C/\partial\theta_i\partial\theta_j$. The rank of H at a critical point measures the local observability:

- **Full rank P:** the critical point is non-degenerate. Locally, the cost uniquely identifies the parameters in all directions.
- **Rank $r < P$:** the critical point has $P - r$ flat directions. Locally, the cost is unobservable in those directions.
- **Rank 0:** the cost is locally flat — barren plateau.

The **observability gap** is the difference between the largest and smallest non-zero Hessian eigenvalues. A small gap means the optimizer struggles to distinguish important from unimportant directions, and the convergence rate degrades accordingly.

Observability And The Fubini-Study Metric

A more geometric view connects observability to the **Fubini-Study metric** on the projective Hilbert space.

For a parameterized state $|\psi(\boldsymbol{\theta})\rangle$, the FS metric on parameter space has matrix elements

$$g_{ij}(\boldsymbol{\theta}) = \mathrm{Re}\left[\langle\partial_i\psi|\partial_j\psi\rangle - \langle\partial_i\psi|\psi\rangle\langle\psi|\partial_j\psi\rangle\right].$$

This is the **Quantum Fisher Information Matrix** (QFIM, Module 6.2) and it captures how much the state changes per unit parameter change.

Connection to observability:

- If $g(\boldsymbol{\theta})$ is full rank, every parameter direction produces a distinguishable state — full state-level observability.
- If $g(\boldsymbol{\theta})$ is rank-deficient, there are parameter directions that don't change the state — *gauge directions*. These are unobservable in any cost function.
- The gauge dimension is $P - \mathrm{rank}(g)$.

This is a *state-level* observability that is independent of the cost. The cost-level observability (Hessian rank of C) is generally tighter (a flat direction in g is automatically flat in C, but a flat direction in C need not be flat in g).

The QFIM gives the **theoretical upper bound** on cost observability for any cost function built from $|\psi(\boldsymbol{\theta})\rangle$.

Practical Diagnostics

How do you *check* observability for a real VQA? A few standard diagnostics:

1. Hessian rank at a candidate optimum. Compute the Hessian (via second-order parameter shift), check the eigenvalue distribution. Rank $< P$ signals an observability deficiency.

2. QFIM rank. Compute the QFIM (Section 6.2), check rank. The gap to P gives the gauge-direction count.

3. Trajectory variance. Run the optimizer from multiple random starting points. If the final parameters cluster, the cost is observable. If they spread out (different parameters reaching the same cost), there are equivalent solutions.

4. Cost contour visualization. For low-dimensional cases, plot the cost contours. Flat regions = observability deficiency.

5. Gradient norm distribution. During training, track the gradient norm. If it becomes very small without the cost decreasing, you're in a flat region.

These diagnostics are simple in principle but require care in interpretation — many VQAs have *some* observability deficiency without it being catastrophic.

Fixing Observability Deficiencies

Once diagnosed, what can you do?

1. Remove gauge redundancies. If the ansatz has explicit gauge freedoms, parameterize the gauge slice instead. E.g., fix one of the redundant parameters to zero.

2. Choose a more informative cost. Some costs are observability-deficient by design (e.g., a single global expectation value). Adding multiple observables (Section 7.6 sketches the multi-observable case) can break the degeneracy.

3. Add a regularization term. A regularization term that depends on the parameters directly (rather than on the cost) can break flat directions and make the optimization landscape better-conditioned. This is one of the thesis's main themes.

4. Restart from a different initialization. Sometimes the observability deficiency is local — restarting from a different point lands you in a region with better observability.

5. Use an observability-aware optimizer. Natural gradient (Module 6.2) explicitly accounts for the QFIM and avoids moving in unobservable directions. This is a partial fix.

The Separation Of State And Parameter Observability

A subtle but important distinction:

State observability: can the cost function identify the state $|\psi(\boldsymbol{\theta})\rangle$ — i.e., does the cost uniquely depend on the state?

Parameter observability: can the cost function identify the parameters $\boldsymbol{\theta}$ — i.e., does the cost uniquely depend on the parameters?

These are *not* the same. A cost can be state-observable (different states give different costs) but parameter-deficient (different parameters give the same state, hence the same cost). This is exactly the gauge-orbit case.

Conversely, a cost can be parameter-observable but state-deficient (different parameters give different costs even though they produce different but cost-equivalent states). This is rare but possible with badly designed costs.

For practical VQA design, **state observability is what you want at the application level** (the optimizer should find the right state), but **parameter observability is what the optimizer sees** (it can only adjust parameters, not states directly). The mismatch between them is a recurring source of subtle bugs and slow convergence.

Structural Role

This node is the **measurement-side foundation** of Module 7. It introduces the observability concept, walks through the classical and quantum versions, applies it to the VQA cost-observability question, and provides diagnostics for observability deficiencies.

It directly enables node 7.4 (optimal control theory, which uses observability to design state estimators) and node 7.6 (feedback control, which requires observable outputs to close the loop).

Relations to Other Concepts

- **Kalman (1960)** — the original classical observability theorem.
- **D'Alessandro (2007)** — quantum observability via informationally complete observables.
- **Quantum tomography** — the operational version of state observability.
- **Module 6.2 (Natural Gradient)** — the QFIM, which captures state observability.
- **Module 8 (Expressivity)** — the dual question (controllability rather than observability).

Worked Example — Observability Of A 1-Qubit Ansatz

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

State. $R_Y(\theta_1)R_Y(\theta_2) = R_Y(\theta_1 + \theta_2)$. So the state depends only on $\theta_1 + \theta_2$.

Cost. Suppose $C = \langle\psi|Z|\psi\rangle = \cos(\theta_1 + \theta_2)$. The cost depends only on the sum.

Observability analysis. - **State observability:** the cost uniquely identifies the state (since cost = cos of the rotation angle, which is monotone in a half-period). - **Parameter observability:** the cost is *constant* along the line $\theta_1 + \theta_2 = \text{const}$. This is a 1-dimensional unobservable direction in 2-dimensional parameter space. - **Gauge dimension:** 1 (the $\theta_1 - \theta_2$ direction is unobservable).

QFIM. Let $\phi = \theta_1 + \theta_2$. Then $|\psi\rangle = R_Y(\phi)|0\rangle$ and $|\partial_1\psi\rangle = |\partial_2\psi\rangle = (\partial/\partial\phi)|\psi\rangle$. The QFIM is

$$g = \begin{pmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix},$$

which has rank 1, eigenvalues $\{1/2, 0\}$. The zero eigenvalue confirms the unobservable direction.

Hessian. $H_{ij} = \partial^2 C/\partial\theta_i\partial\theta_j = -\cos(\theta_1 + \theta_2)$ for all i, j. Same matrix structure (rank 1).

Practical implication. A gradient-descent optimizer started at random will move in the gradient direction (which is along $(1, 1)^T$), reaching a cost-decreasing trajectory but never resolving the perpendicular direction. The final (θ_1, θ_2) will not be unique — it will lie on the line $\theta_1 + \theta_2 = \theta^*$ for some θ^*.

Fix. Eliminate θ_2 (set to zero) and use only θ_1. Now you have a 1-parameter ansatz with full observability. Same expressive power, half the parameters, no gauge problem.

This worked example is small but the structural lesson generalizes: **redundant parameters create observability holes**, and pruning them makes the optimization easier.

Visualization

Pending. A 2D heatmap of $C(\theta_1, \theta_2) = \cos(\theta_1 + \theta_2)$ showing the diagonal stripes of constant cost. Overlaid: a gradient-descent trajectory that follows the gradient direction but ends at an arbitrary point on a stripe. Caption: "The flat direction is the unobservable subspace. The optimizer doesn't know where on the stripe to stop."

Exercises

1. For a 2-qubit ansatz $R_Y(\theta_1) \otimes R_Y(\theta_2)|00\rangle$ with cost $C = \langle Z_0 + Z_1 \rangle$, is the cost observable in both parameters?
2. The QFIM has rank r at a parameter point. How many gauge directions does the parameter space have at that point?
3. (VQA) For a Hamiltonian Variational Ansatz on 6 qubits with 12 parameters, suppose you compute the Hessian rank at a candidate optimum and find it is 9. What does this imply about the optimizer behavior?

Research Extensions

- **Dynamical observability for time-varying systems** — extending classical observability theorems to drift-prone hardware.
- **Approximate observability under shot noise** — how many shots are needed to identify a parameter to a given accuracy?
- **Cost-design for observability** — algorithmic methods to choose costs that maximize parameter observability.
- **Observability and adaptive optimization** — using observability diagnostics to drive optimizer choice.

Cross-links to Other Courses

- **Linear Systems Theory** — Kalman observability, the classical foundation.
- **Quantum State Tomography** — the operational version of state-level observability.
- **Module 6.2 (Natural Gradient)** — the QFIM-based view of observability.
- **Module 7.3 (Controllability)** — the dual notion.
- **Module 7.6 (Feedback Control)** — where observability becomes operationally critical.

Variational Quantum Algorithms · Module 7 — Control Theoretic View · Node 7.2
Concepts: observability, control-theory, state-space, expectation-value
Prerequisites: 7.1
Enables: 7.4, 7.6

hbar.university — own.your.cognition