Need to know about the qubit-experiments (v0_6–v0_6.1)

Theme: Systematic signature detection + evidence layer for physics-native subspaces
Scope: v0_6 (discover signature families) and v0_6.1 (robustness + enrichment + hold-out replication)

Indhold

0. What changed vs. v0_3–v0_5

v0_3–v0_5 established a pipeline that (i) finds near-degenerate eigenpairs, (ii) inspects eigenvector structure, and (iii) tests dynamic stability through leakage under time evolution (locally and, where possible, on backend).
v0_6 adds population-level structure: it asks whether "good" candidates cluster into **repeatable classes** (signatures) across many random Hamiltonians.
v0_6.1 adds an evidence layer: robustness sweeps, enrichment tests, and hold-out replication to quantify repeatability and stability-correlation without over-claiming.

1. Mathematical objects you scan

1.1 Hilbert space and computational basis

An n-qubit system lives in a complex Hilbert space of dimension 2^n, with computational basis $\{|x\rangle : x \in \{0,1\}^n\}$.

Equation (Word/LaTeX input):

$$\dim(\mathcal{H}_n) = 2^n$$

Equation (Word/LaTeX input):

$$|\psi\rangle = \sum_{x\in\{0,1\}^n} \alpha_x \; |x\rangle, \sum_x |\alpha_x|^2 = 1$$

1.2 Pauli-sum Hamiltonians

All v0_6–v0_6.1 experiments use Hamiltonians built from Pauli strings.

Equation (Word/LaTeX input):

$$H = \sum_k c_k \; P_k, P_k \in \{I, X, Y, Z\}^{\otimes n}$$

This matters because:

- The **spectrum** (eigenvalues) is used to detect candidate manifolds (near-degeneracy).
- The **eigenvectors** encode structure (dominant basis patterns, entropy).
- Pauli structure supports implementable approximations for time evolution (link to v0_5 lineage).

2. Spectral decomposition: why eigenpairs are central

Diagonalizing Hyields eigenpairs $\{(E_j, | \psi_j\rangle)\}$.

Equation (Word/LaTeX input):

$$H | \psi_j\rangle = E_j | \psi_j\rangle$$

Equation (Word/LaTeX input):

$$H = \sum_j E_j \ | \psi_j\rangle\langle\psi_j |$$

Equation (Word/LaTeX input):

$$\langle\psi_i | \psi_j\rangle = \delta_{ij}$$

Operationally, v0_6 uses this to extract:

- **Eigenvalue spacings** (where candidate pairs exist)
- **Eigenvector structure** (how "low-dimensional" the state looks in the computational basis)
- **Dominant basis states** (which bitstrings carry most weight)

3. Near-degenerate pairs and candidate subspaces

A near-degenerate eigenpair $(i, j)$is detected via a threshold ε.

Equation (Word/LaTeX input):

$$| E_i - E_j | < \varepsilon$$

This defines a candidate 2D invariant subspace:

Equation (Word/LaTeX input):

$$\mathcal{S}_{ij} = \mathrm{span}\{| \psi_i\rangle, | \psi_j\rangle\}$$

In v0_6, each detected $\mathcal{S}_{ij}$is treated as a **candidate physics-native information unit** (provisionally), then characterized by structure + stability descriptors.

4. The heart of v0_6: from "found pairs" to "signature families"

4.1 Structural descriptors per candidate

For each candidate subspace $\mathcal{S}_{ij}$, v0_6 extracts a compact descriptor set:

1. **Dominant basis count**
   How many basis states $| x\rangle$account for most probability mass in each eigenvector (e.g., top components capturing $> 90\%–95\%$ weight).

2. **Amplitude entropy** (structure proxy)
   Let $p_x =| \alpha_x |^2$. Define:

Equation (Word/LaTeX input):

$$H_{\mathrm{amp}}(| \psi\rangle) = -\sum_x p_x \log p_x$$

Low H_{amp}indicates concentrated, compressible structure; high entropy indicates spread-out eigenvectors.

3. **Leakage proxy** (lightweight dynamic stability)
   Prepare $| \psi(0)\rangle$supported on the candidate structure and evolve under H, then measure how much remains "inside" the candidate manifold.

Equation (Word/LaTeX input):

$$| \psi(t)\rangle = U(t) \ | \psi(0)\rangle, U(t) = e^{-iHt}$$

A conceptual leakage definition using an orthonormal basis $\{| \phi_a\rangle \}$for the target subspace:

Equation (Word/LaTeX input):

$$L(t) = 1 - \sum_{a\in\mathcal{S}} | \langle \phi_a | \psi(t)\rangle |^2$$

(Implementation-wise, you often approximate this using a practical "inside-set" defined by dominant bitstrings or a projector surrogate, to keep runs scalable.)

4.2 Integer-binned signature keys

Raw floats produce "false uniqueness" (tiny numerical differences fragment classes).
So v0_6 defines discrete **signature keys** via integer binning:

- d= dominant basis count (integer)
- b_i, b_j= entropy bins (integers)
- ℓ= leakage bin (integer)

Conceptually you can view a key as:

Equation (Word/LaTeX input):

$$\kappa = (d,\ b_i,\ b_j,\ \ell) \in \mathbb{Z}^4$$

v0_6 output is essentially a **frequency distribution** over keys κ: which families recur, and how often.

5. One-sentence summary of what v0_6 adds

v0_6 converts individual degeneracy findings into repeatable signature families, using discrete (integer-binned) structural and stability descriptors so that recurrence can be quantified across large seed scans.

6. v0_6.1: the evidence layer (why it is more persuasive)

6.1 Robustness sweep (quick)

Re-run with 2–3 values of parameters (e.g., ε, entropy bin width, leakage bin width, dominance threshold). Then compare the top-K signature sets across settings.

A standard overlap metric is the Jaccard index:

Equation (Word/LaTeX input):

$$J(A, B) = \frac{|\ A \cap B\ |}{|\ A \cup B\ |}$$

Interpretation: high overlap means "top families are not artifacts of one tuning."

6.2 Enrichment test (high-value, easy to explain)

Define a stable set (e.g., top 1%or $5\%$by stability/interestingness score).
For a signature family S(i.e., a particular κ), compare its frequency inside the stable set vs. overall.

Equation (Word/LaTeX input):

$$\mathrm{Enrichment}(S) = \frac{\Pr\,(S\ \mid\ \text{stable})}{\Pr\,(S\ \mid\ \text{all})}$$

Enrichment > 1indicates $S$is over-represented among stable candidates. A bootstrap CI makes the claim statistically readable.

6.3 Hold-out replication (clean replication logic)

Split seeds into two halves:

- learn “top families” in half A,
- test whether they also rank high in half B.

Replication can be summarized with overlap of top-K families (again Jaccard, or rank correlation).

7. “Interestingness” (kept minimal and method-first)

In v0_6.1, “interestingness” is a ranking device to define the stable set. A typical conservative direction is:

- reward low leakage (stability),
- reward small dominant basis count (structured).

A minimal form:

Equation (Word/LaTeX input):

$$\text{score} = w_1\,(1-\bar{L}) + w_2\,\frac{1}{d}$$

where $\bar{L}$is a leakage summary (e.g., average over sampled times) and $d$is dominant basis count. The key is that the *evidence tests* remain the same even if the exact score is adjusted.

8. How to read your outputs (what counts as “the result”)

When you review v0_6/v0_6.1 logs, focus on four outputs:

1. **Top signature families by occurrence** (recurrence structure)
2. **Top families by enrichment** (stability association)
3. **Robustness overlaps** across parameter settings (fragile vs stable conclusion)
4. **Hold-out replication overlap** (reproducibility on unseen seeds)

A careful, defensible claim is:
“We observe repeatable signature families; some are enriched among low-leakage candidates; robustness and hold-out checks quantify how stable that conclusion is.”

9. Limitations (credibility section)

- Still controlled at $n = 3$qubits (by design).
- Signature keys depend on discretization (bin width trade-offs).

Need to know about the qubit-experiments (v0_6–v0_6.1)

- Leakage is a proxy (lightweight dynamics), not full open-system decoherence.
- Random Pauli-sum Hamiltonians are a test bed; physics-motivated families are the next step.