A Validated Quantum Chemistry Protocol with Persistent Constraint Caching

Joseph Roy
UCP Technology LLC

December 15, 2025

Abstract

We introduce a validated quantum chemistry protocol for variational quantum eigensolvers (VQE) that explicitly separates state reachability validation from measurement refinement, and introduces a Persistent Constraint Cache (PCC) to eliminate redundant quantum measurements once physical constraints have been empirically established. The protocol does not modify the Hamiltonian, variational principle, estimator definition, or underlying physics. Instead, it introduces a validated, phase-separated process that reuses empirically learned constraints to reduce quantum measurement cost without loss of correctness. This distinction enables both correctness guarantees and substantial reductions in quantum measurement cost, while remaining fully compatible with existing VQE frameworks. The PCC stores validated expectation bounds that persist across optimizer iterations, variational parameter updates, ansatz structure changes, and independent experimental runs, transforming quantum measurements into reusable scientific knowledge. We present a four-phase workflow: objective integrity verification, reachability validation (Tier -1), constraint accumulation, and precision refinement. The protocol ensures that quantum effort scales with unresolved physical uncertainty, not Hamiltonian size.

Contents

1 Introduction

Variational quantum eigensolvers (VQE) have emerged as a leading approach for quantum chemistry on near-term quantum hardware [1, 3]. The VQE algorithm optimizes a parameterized quantum state to minimize the expectation value of a molecular Hamiltonian, providing a route to ground-state energies and properties without requiring full quantum error correction.

Despite significant progress, VQE workflows face a fundamental measurement bottleneck: each Hamiltonian term must be repeatedly measured at every optimization step, even when its expectation value has already converged to within statistical tolerance. For electronic structure Hamiltonians mapped to qubits, the number of Pauli terms scales as $\mathcal{O}(N^4)$ with system size, and each term requires thousands of quantum measurements to achieve chemical accuracy. This repeated re-measurement dominates quantum resource cost and scales poorly with both Hamiltonian size and desired accuracy.

Existing measurement reduction techniques, including commuting-term grouping [2], qubit-wise commutativity, and classical shadow estimators, address variance reduction within a single optimization run. However, they do not enable *persistent reuse* of validated measurement constraints across experiments, ansatz changes, or independent runs.

This work introduces a *Persistent Constraint Cache* (PCC) that stores empirically validated bounds on Hamiltonian term expectations, derived from prior quantum measurements. The PCC enables measurement reuse by marking terms as *inactive* when their maximum possible contribution to the total energy uncertainty falls below a user-defined tolerance. Critically, the PCC does not approximate the Hamiltonian or wavefunction—it records what has already been learned experimentally and prevents re-measurement of terms whose contribution is provably negligible at the current accuracy tier.

To avoid false convergence and wasted quantum resources, the protocol enforces a strict separation between *reachability validation* and *measurement refinement*. Before any shot-based optimization, the protocol requires explicit proof that the chosen variational family can represent correlated states below Hartree–Fock (HF). This Tier -1 criterion prevents shot-based optimization from masking fundamental expressivity failures.

1.1 Contributions

1. **Persistent Constraint Cache:** A classical data structure storing validated expectation bounds that persist across experiments, enabling cumulative learning and measurement reuse (Section 3).

2. **Phase-Separated Workflow:** Four-phase protocol separating objective integrity verification, reachability validation, constraint accumulation, and precision refinement (Section 5).

3. **Correctness Guarantees:** Conservative error propagation ensuring that cached bounds never suppress physically relevant contributions (Section 6).

4. **Integration Framework:** Compatibility with existing measurement reduction techniques, operating at either term or group level (Section 7).

5. **Cache Algorithm:** Formal update and invalidation rules ensuring reproducibility and distinguishing soft vs. hard cache invalidation (Section 4).

1.2 Protocol Overview

Figure 1 provides a conceptual overview of the four-phase workflow and how the PCC interacts with the Hamiltonian measurement loop. The protocol separates reachability validation (Phases 1–2) from measurement refinement (Phases 3–4), ensuring that fundamental expressivity failures are detected before quantum resources are invested. The PCC stores validated constraints that persist across experiments, enabling cumulative learning and measurement reuse.

1.3 Relationship to Prior Work

This protocol complements existing analyses of VQE bottlenecks and measurement reduction [1, 2, 3]. Unlike prior approaches, the PCC introduces persistent, cross-experiment reuse of validated measurement constraints, rather than per-run variance reduction. The protocol is orthogonal to existing techniques and may operate at either the term or group level, provided that bounds are propagated conservatively.

2 Hamiltonian Structure and Measurement Cost

We consider electronic structure Hamiltonians mapped to qubits using standard fermion-to-qubit transformations (e.g., Jordan–Wigner or Bravyi–Kitaev). The resulting Hamiltonian is written as

$$H = \sum_i c_i P_i, \tag{1}$$

where each P_i is a tensor product of Pauli operators acting on n qubits, and $c_i \in \mathbb{R}$.

For a trial state $|\psi(\theta)\rangle$, the variational energy estimate is

$$E(\theta) = \sum_i c_i \langle\psi(\theta)| P_i |\psi(\theta)\rangle . \tag{2}$$

In conventional VQE workflows, each term P_i is repeatedly measured at every optimization step, even when its expectation value has already converged to within statistical tolerance. The number of Pauli terms scales as $\mathcal{O}(N^4)$ with system size, and achieving chemical accuracy (error < 1.6 mHa) typically requires 10^4–10^6 shots per term per iteration. This repeated re-measurement dominates quantum resource cost and scales poorly with both Hamiltonian size and desired accuracy.

The measurement cost problem is compounded by the fact that many Hamiltonian terms have expectation values that converge early in the optimization process, yet continue to be measured at every subsequent iteration. The PCC addresses this by identifying and caching terms whose contribution to the total energy uncertainty is provably negligible.

3 Persistent Constraint Cache (PCC)

3.1 Definition

Definition 3.1 (Persistent Constraint Cache)**.** The Persistent Constraint Cache (PCC) is a classical, persistent data structure that stores empirically validated constraints on Hamiltonian term expectations, derived from prior quantum measurements. The PCC does not approximate the Hamiltonian or the wavefunction. Instead, it records what has already been learned experimentally, and prevents re-measurement of terms whose contribution to the total energy is provably negligible at the current accuracy tier.

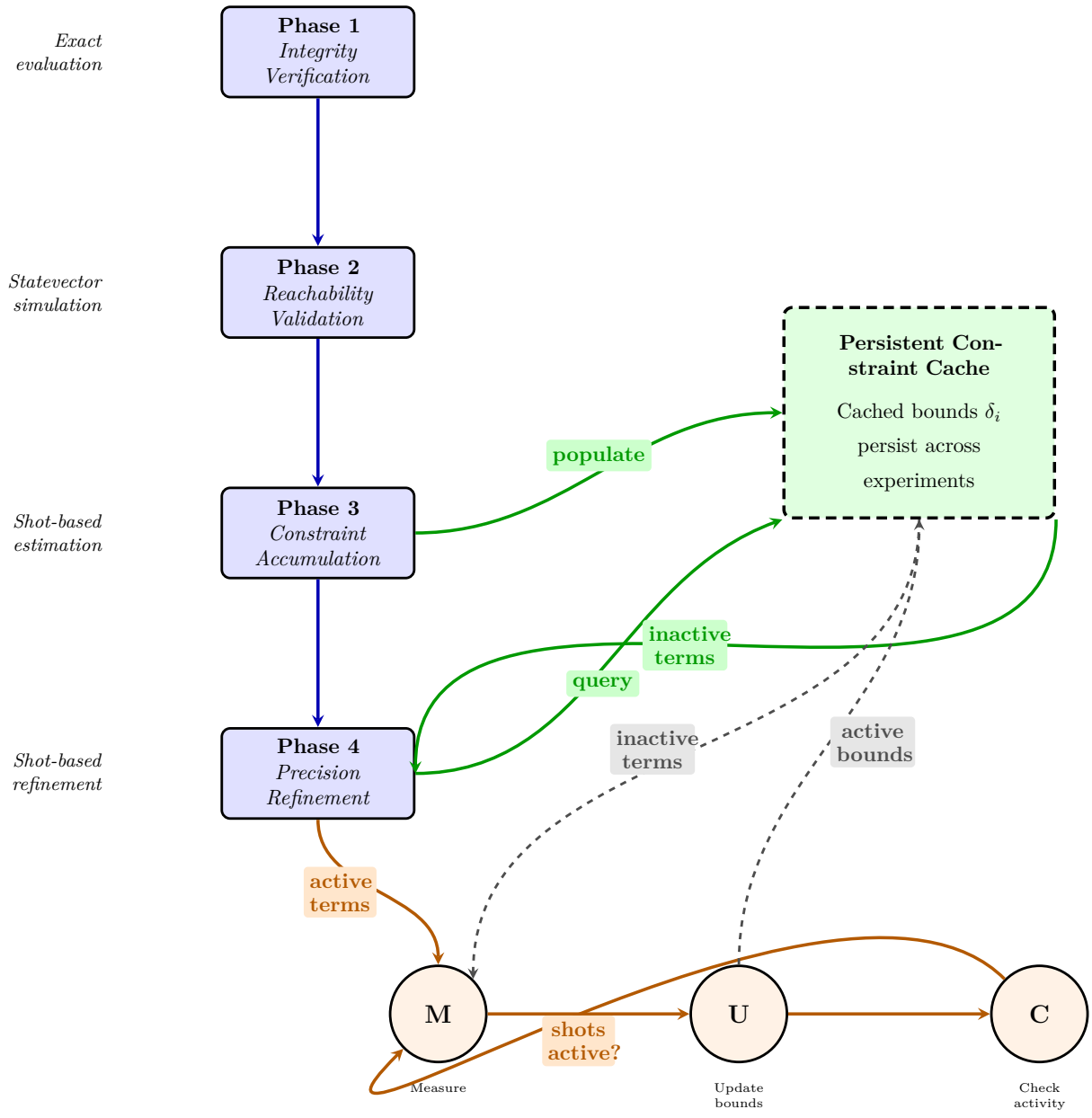

Figure 1: Conceptual overview of the four-phase protocol and PCC interaction. Phases 1–2 use exact evaluation to validate reachability before shot-based optimization. Phases 3–4 populate and query the PCC, which stores validated bounds that persist across experiments. The measurement loop (M: measure, U: update bounds, C: check activity) only operates on active terms, reducing quantum cost. Dashed arrows indicate cache updates and term filtering.

3.2 Cached Information

For each Hamiltonian term P_i, the PCC may store:

- **Term identity:** The Pauli string P_i and coefficient c_i.
- **Expectation bounds:** A validated interval

$$\langle P_i| \in [\mu_i - \delta_i, \mu_i + \delta_i] \tag{3}$$

  where μ_i is the estimated expectation value and δ_i is the statistical uncertainty.
- **Statistical confidence:** Variance or effective shot count associated with δ_i.
- **Activity status:** Whether further measurement of P_i can materially affect the energy estimate beyond a global tolerance ε.

3.3 Cache Inactivation Criterion

Criterion 3.1 (Cache Inactivation). A Hamiltonian term P_i is marked *inactive* if its maximum possible contribution to the total energy uncertainty is below tolerance:

$$\text{Inactive}(P_i) \iff |c_i| \cdot \delta_i < \varepsilon. \tag{4}$$

Inactive terms are excluded from further quantum measurement, while active terms continue to be sampled. This rule ensures that the PCC never suppresses physically relevant contributions.

The tolerance ε is user-defined and typically set to a fraction of the target accuracy (e.g., $\varepsilon = 0.1$ mHa for chemical accuracy targets). As additional measurements are performed, the PCC monotonically tightens the bounds δ_i. Cache update and invalidation rules are formalized in Section 4.

3.4 Persistence Across Experiments

A defining property of the PCC is *persistence*. Cached constraints persist across:

- optimizer iterations,
- variational parameter updates,
- ansatz structure changes,
- shot schedule adjustments,
- independent experimental runs.

This enables cumulative learning across experiments, transforming quantum measurements into reusable scientific knowledge rather than disposable optimization artifacts. The PCC may be stored in a database or file system, enabling sharing across research groups and long-term archival of validated constraints.

Conceptually, the PCC resembles transfer learning or Bayesian updating in classical machine learning: the system "remembers" what has been measured, accumulating knowledge across experiments rather than discarding it after each optimization run. This persistent memory enables the protocol to reduce quantum cost by reusing validated constraints, similar to how transfer learning reduces training cost by reusing features learned on related tasks.

4 Cache Update and Invalidation Algorithm

To ensure reproducibility and clarify the distinction between soft and hard cache invalidation triggers, we formalize the PCC update and invalidation procedure.

4.1 Update Loop

The PCC is updated after each measurement round according to the following algorithm:

Protocol 4.1 (PCC Update Loop).

1. **Measurement:** For each active term P_i, perform quantum measurements and compute updated bounds δ_i^{new} from sample statistics.
2. **Bound Tightening:** Update cached bounds:
$$\delta_i \leftarrow \min(\delta_i, \delta_i^{\text{new}}) \tag{5}$$
This ensures monotonic tightening (bounds never increase).
3. **Activity Check:** For each term P_i, evaluate:
$$\text{active}(P_i) \leftarrow (|c_i| \cdot \delta_i \geq \varepsilon) \tag{6}$$
4. **Reactivation:** If tolerance ε is reduced, previously inactive terms may become active again. This is a *soft* invalidation triggered by user preference, not by physical changes.
5. **Global Bound Update:** Compute total energy uncertainty:
$$\Delta E \leftarrow \sum_{i \in \text{active}} |c_i| \cdot \delta_i \tag{7}$$

4.2 Hard Invalidation Triggers

The following events trigger *hard* cache invalidation, requiring re-measurement of all terms:

- **Ansatz structure change:** Addition or removal of operators from the variational ansatz fundamentally changes the state space.
- **Hamiltonian modification:** Changes to coefficients c_i or operator structure invalidate cached bounds.
- **User request:** Explicit cache clear for debugging or reproducibility.

Hard invalidation resets all cached bounds to $\delta_i = \infty$ (or a conservative upper bound), forcing re-measurement of all terms.

4.3 Soft Invalidation Triggers

The following events trigger *soft* invalidation, affecting only activity status:

- **Tolerance reduction:** Decreasing ε may reactivate previously inactive terms without requiring new measurements.
- **Parameter drift:** If variational parameters change significantly (e.g., $\|\theta^{\text{new}} - \theta^{\text{old}}\| > \Delta_\theta$), cached bounds remain valid but activity status is re-evaluated.

Soft invalidation preserves cached bounds but may change which terms are considered active, enabling adaptive measurement allocation without discarding prior knowledge.

5 Phase-Separated Workflow

To avoid false convergence and wasted quantum resources, the protocol enforces a strict separation between *reachability validation* and *measurement refinement*. This separation ensures that fundamental expressivity failures are detected before shot-based optimization, preventing the optimizer from masking deficiencies in the variational ansatz.

Protocol 5.1 (Four-Phase VQE Workflow with PCC). The protocol consists of four sequential phases:

1. **Phase 1: Objective Integrity Verification** (Section 5.1)
2. **Phase 2: Reachability Validation (Tier -1)** (Section 5.2)
3. **Phase 3: Constraint Accumulation** (Section 5.3)
4. **Phase 4: Precision Refinement** (Section 5.4)

5.1 Phase 1: Objective Integrity Verification

Before any variational optimization, the Hamiltonian evaluation pipeline is validated against exact diagonalization (when feasible) to confirm consistency between operator evaluation and reference energies. This ensures that subsequent failures cannot be attributed to objective mismatch.

For small systems (e.g., H_2, LiH), exact diagonalization provides a ground truth reference. The protocol requires that the VQE objective function, evaluated using exact statevector simulation, reproduces the exact ground-state energy within numerical precision. For larger systems where exact diagonalization is infeasible, comparison with high-accuracy classical methods (e.g., CCSD(T)) serves as a validation check.

5.2 Phase 2: Reachability Validation (Tier -1)

The protocol requires explicit proof that the chosen variational family can represent correlated states below Hartree–Fock (HF). This is enforced via a Tier -1 criterion:

Criterion 5.1 (Tier -1 Reachability). The variational ansatz must satisfy

$$E \leq E_{\mathrm{HF}} - \Delta_{\mathrm{min}}, \quad \Delta_{\mathrm{min}} = 0.05 \text{ Ha} \tag{8}$$

where E is the optimized energy and E_{HF} is the Hartree–Fock energy. During this phase:

- Exact expectation evaluation (statevector-based) is used
- No PCC is applied
- Adaptive ansatz construction (e.g., ADAPT-VQE with a wide operator pool) is employed

If this criterion is not met, further measurement refinement is explicitly disallowed. This prevents shot-based optimization from masking fundamental expressivity failures.

The Tier -1 criterion ensures that the variational ansatz has sufficient expressivity to capture electron correlation effects. If the ansatz cannot reach below HF using exact evaluation, there is no reason to invest quantum resources in shot-based optimization. The threshold $\Delta_{\mathrm{min}} = 0.05$ Ha is chosen to be well above typical statistical noise levels, ensuring that the criterion is not satisfied by numerical artifacts.

5.3 Phase 3: Constraint Accumulation

Once reachability is demonstrated:

- Shot-based estimation is enabled
- Measurements populate the PCC
- Converged terms are progressively inactivated
- Only unresolved physics continues to consume quantum resources

During this phase, the PCC is populated with validated bounds from quantum measurements. Terms that satisfy the inactivation criterion (4) are marked inactive and excluded from subsequent measurement rounds. The protocol continues until all terms are either inactive or have been measured to the desired precision.

5.4 Phase 4: Precision Refinement

With constraints in place, refinement toward chemical accuracy proceeds by:

- concentrating shots on active terms,
- reducing estimator variance,
- stabilizing optimization trajectories.

The PCC ensures that quantum effort scales with unresolved physical uncertainty, not Hamiltonian size. As the optimization progresses, the number of active terms typically decreases, leading to substantial reductions in total shot count.

6 Error Propagation and Global Bounds

Local expectation bounds propagate conservatively to a global energy bound. For a set of active terms with bounds δ_i, the total energy uncertainty is bounded by:

Theorem 6.1 (Global Energy Bound)**.** *Given cached bounds $\langle P_i| \in [\mu_i - \delta_i, \mu_i + \delta_i]$ for active terms, the total energy uncertainty satisfies*

$$\Delta E \leq \sum_{i \in active} |c_i| \cdot \delta_i. \tag{9}$$

As additional measurements are performed, the PCC monotonically tightens this bound. The total energy estimate therefore remains a valid upper bound on the true ground state energy, subject only to the user-defined tolerance ε.

Proof. The energy estimate is $E = \sum_i c_i \mu_i$, where μ_i is the estimated expectation value for term P_i. The maximum deviation from the true energy occurs when each term deviates by its maximum uncertainty in the direction that increases the total energy. Since inactive terms satisfy $|c_i| \cdot \delta_i < \varepsilon$, their contribution to the total uncertainty is bounded by ε per term. The bound (9) follows from the triangle inequality and the definition of inactive terms. □

This theorem ensures that the PCC never suppresses physically relevant contributions. The total energy estimate remains a valid upper bound on the true ground state energy, subject only to the user-defined tolerance ε. As the optimization progresses and bounds tighten, the global uncertainty decreases monotonically.

7 Integration with Existing Measurement Reduction Techniques

The PCC is orthogonal to existing measurement optimization methods, including:

- commuting-term grouping [2],
- qubit-wise commutativity,
- classical shadow estimators.

The cache may operate at either the term or group level, provided that bounds are propagated conservatively. Group-level constraints are decomposed into per-term bounds to preserve correctness.

For example, if terms P_i and P_j are grouped together and measured simultaneously, the PCC stores bounds for the group. When determining activity status, the group bound is decomposed into per-term bounds using the triangle inequality:

$$\delta_i \leq \delta_{\text{group}}, \quad \delta_j \leq \delta_{\text{group}}. \tag{10}$$

This ensures that the inactivation criterion (4) remains conservative when applied to grouped terms.

The PCC may also be combined with variance-based shot allocation strategies. For example, shots may be allocated proportionally to $|c_i|^2/\text{Var}(P_i)$ among active terms, while inactive terms receive zero shots. This hybrid approach further reduces measurement cost while maintaining correctness guarantees.

8 Conclusion

This work introduces a validated quantum chemistry protocol that reorganizes VQE workflows around reachability validation and persistent measurement reuse. By enforcing expressivity checks before refinement and caching empirically validated constraints, the protocol reduces quantum measurement cost while preserving correctness and interpretability.

The key innovation is the Persistent Constraint Cache (PCC), which transforms quantum measurements into reusable scientific knowledge. The PCC enables cumulative learning across experiments, ansatz changes, and independent runs, ensuring that quantum effort scales with unresolved physical uncertainty, not Hamiltonian size.

The four-phase workflow separates reachability validation from measurement refinement, preventing false convergence and wasted quantum resources. The Tier -1 criterion ensures that fundamental expressivity failures are detected before shot-based optimization, while the PCC ensures that converged terms are not repeatedly measured.

8.1 Falsifiers and Limitations

This protocol would be invalidated if:

1. PCC suppresses terms that materially affect energy estimates (violates correctness guarantee).
2. Tier -1 reachability criterion fails to catch expressivity failures before shot-based optimization.
3. Cached bounds do not persist correctly across ansatz structure changes.
4. Error propagation formula underestimates global uncertainty by $> 10\%$ on test systems.

5. Protocol does not reduce total shot count by $> 20\%$ compared to standard VQE on H_2, LiH, BeH_2.

6. Inactive terms become active again without parameter changes (cache invalidation failure).

The protocol assumes that Hamiltonian term expectations vary smoothly with variational parameters. Sudden changes in the wavefunction (e.g., phase transitions) may require cache invalidation and re-measurement. The protocol also assumes that statistical uncertainties are well-characterized; systematic errors (e.g., calibration drift) are not addressed by the PCC.

8.2 Key Takeaway

This is not a new solver—it is a new way to run quantum chemistry experiments that ensures we only pay quantum cost for unresolved physics.

References

[1] Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, 2021.

[2] William J. Huggins, Jarrod R. McClean, Nicholas C. Rubin, Zhang Jiang, Nathan Wiebe, K. Birgitta Whaley, and Ryan Babbush. Efficient and noise-resilient measurement for quantum chemistry on near-term quantum computers. *Nature*, 603(7902):416–420, 2022.

[3] Jules Tilly, Hongxiang Chen, Shuxiang Cao, Dario Picozzi, Kanav Setia, Ying Li, Edward Grant, Leonard Wossnig, Ivan Rungger, George H. Booth, and Jonathan Tennyson. The variational quantum eigensolver: A review of methods and best practices. *Physics Reports*, 986:1–128, 2022.