TASK-SUFFICIENT QUANTUM MEASUREMENTS FOR CERTIFIED CLUSTERING

Anonymous authors
Paper under double-blind review

ABSTRACT

Quantum-kernel-based clustering methods typically estimate pairwise quantum-state fidelities to construct a similarity graph and then derive a clustering partition from the resulting graph. However, estimating pairwise fidelities separately across all sample pairs requires measurements on $\Theta(n^2)$ pairs, even though the final partition may depend on far less information. Moreover, existing methods return a partition without certifying whether the acquired measurements are sufficient to uniquely determine it. To address these limitations, we propose Quantum Certification for Clustering (Q-Cert), a task-oriented online quantum measurement framework. Q-Cert establishes within-cluster connectivity through selected overlap tests and reuses single-state measurements to exclude multiple cross-cluster edges; the resulting evidence is integrated into confirmed and possible graphs that determine when the partition can be safely certified. Theoretically, Q-Cert guarantees that the probability of incorrectly certifying a partition is at most δ for arbitrary pure-state inputs. Moreover, as long as no pairwise fidelity lies exactly at the threshold, Q-Cert can complete certification even when the pilot and reusable evidence provide no useful information. Under suitable connectivity-witness and reusable-separation conditions, Q-Cert requires only $O(n \log(n/\delta))$ quantum-oracle calls. On constructed pure-state families, we establish two complementary separations: partition certification avoids a gap-dependent full-graph recovery cost under identical measurement access, while informative reusable measurements yield a quadratic-to-near-linear separation for the same clustering task relative to restricted access. On the bounded-degree family, savings over full-risk direct increase over the tested sizes, reaching 75.32% at $n = 192$.

1 INTRODUCTION

Clustering is an unsupervised learning task that aims to partition data into groups such that samples within the same cluster are highly similar, while those from different clusters are dissimilar. Among existing clustering methods, spectral clustering has attracted considerable attention due to its effectiveness and broad applicability (Shi & Malik, 2000; von Luxburg, 2007; Shaham et al., 2018). Rooted in spectral graph theory, it represents samples as vertices, encodes pairwise similarities as edge weights, and derives clusters by partitioning the resulting similarity graph (Nie et al., 2024). The quality of this graph is therefore crucial to clustering performance. Conventional kernel-based graph construction commonly measures sample similarities using predefined functions, such as linear, polynomial, and Gaussian kernels. However, their fixed functional forms can limit the flexibility to characterize complex nonlinear relationships in high-dimensional data, yielding similarity graphs that inadequately reflect the underlying data structure (Liu et al., 2020; Milsom et al., 2024).

Quantum-kernel-based clustering (QKC) provides a promising alternative by exploiting quantum feature representations to construct sample similarities. Quantum feature maps encode classical data into high-dimensional quantum Hilbert spaces, enabling more expressive characterization of complex nonlinear relationships (Havlíček et al., 2019; Schuld & Killoran, 2019; Thanasilp et al., 2024). Kerenidis & Landman (2021) develop an end-to-end quantum algorithm for spectral clustering under quantum data-access assumptions. This representational flexibility has stimulated growing interest in QKC. For example, Slabbert et al. (2026) construct parameterized quantum fidelity kernels for spectral clustering, Abraham-Mexicano et al. (2026) combine a ZZ-feature-map quantum kernel with spectral clustering to identify water-pollution profiles, and Towada (2026) transforms

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

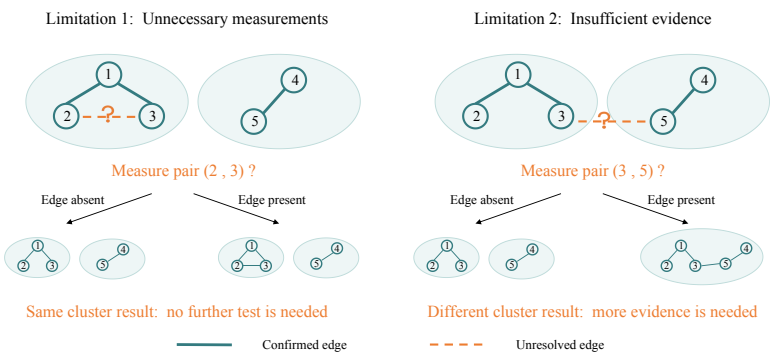

Figure 1: Partition-relevant evidence for fidelity-threshold connected components. Unresolved internal edges cannot change an established component, whereas unresolved cross-group edges may change the partition. Solid and dashed lines denote confirmed and unresolved edges.

quantum-kernel similarities into an affinity graph for clustering-based decomposition of capacitated vehicle routing problems.

Motivated by the measurement cost of quantum-kernel clustering, we study a specific certification task: recovering the connected-component partition of a fidelity-threshold graph. This target differs from general spectral clustering and raises two measurement challenges. First, a conventional approach to this task estimates pairwise fidelities to construct the threshold graph and then extracts its connected components. However, recovering every edge of this graph can require more measurement evidence than determining its connected-component partition. For example, once a set of vertices is connected by confirmed edges, resolving additional edges within that set cannot change its connectivity. Such measurements may therefore be unnecessary for partition certification, as illustrated in Fig. 1. Second, under finite measurements, a partition extracted from an estimated graph may differ from that of the true fidelity-threshold graph. Returning this partition alone does not certify that the acquired evidence rules out alternative connected-component partitions.

To overcome these issues, we propose **Quantum Certification for Clustering (Q-Cert)**, a task-oriented online quantum measurement framework. Rather than reconstructing the full similarity graph, Q-Cert directly acquires the task-sufficient evidence needed to determine and certify the target partition. Specifically, it uses low-budget pilot measurements to identify candidate cluster structures and useful connections, selective overlap measurements to establish within-cluster connectivity, and reusable single-state measurements to rule out multiple cross-cluster connections simultaneously. The acquired evidence is integrated into graphs of confirmed and possible edges, and Q-Cert continues measurement acquisition only while the remaining uncertainty can still alter the partition. Once the two graphs induce the same connected components, the available evidence uniquely determines the partition, allowing Q-Cert to stop and certify the result.

Beyond the algorithmic design, we provide an end-to-end theoretical characterization of finite-measurement certified clustering, covering universal safe certification, fallback completion, and online measurement complexity. This analysis identifies structural conditions under which task-sufficient certification is provably cheaper than resolving the full pairwise fidelity graph.

Our theoretical and empirical results show the reliability and measurement efficiency of Q-Cert. For arbitrary pure-state inputs, we theoretically prove that Q-Cert controls the probability of incorrect acceptance by δ while still completing certification when no pairwise fidelity lies exactly at the threshold, thereby ensuring reliable certification of the clustering result even when the pilot and shared measurements are uninformative. Regarding measurement efficiency, under additional sparsity and separation conditions, Q-Cert reduces the quadratic measurement burden associated with pairwise resolution to a near-linear $O(n \log(n/\delta))$ oracle-call complexity. This guarantee extends to a class of quantum states realizing arbitrary bounded-degree sparse graph topologies and remains valid under controlled pure-state perturbations. Empirically, we separately validate the two mechanisms underlying these savings—discovering useful connections and reusing state-level evidence. Under the same measurement interface, Q-Cert's measurement savings increase to 34.08%, 55.10%, and 75.32% at $n = 48, 96, 192$, respectively.

Our main contributions are summarized as follows.

Task-sufficient online certification. We introduce Q-Cert, an online measurement-acquisition method that directly targets the final cluster partition. For arbitrary pure-state inputs, Q-Cert has an incorrect-acceptance probability of at most δ; moreover, provided no pairwise fidelity lies exactly at the threshold, it can complete certification given sufficient budget even without a correct pilot or informative shared measurements.

From quadratic to near-linear measurement complexity. Under suitable sparsity and separation conditions, Q-Cert reduces the quadratic-scale cost of pairwise certification to $O(n \log(n/\delta))$ quantum-oracle calls. The guarantee covers a class of quantum states realizing arbitrary bounded-degree sparse graph topologies and persists under controlled pure-state perturbations. We establish two complementary complexity separations: task-sufficient partition certification can be asymptotically cheaper than full graph recovery under identical measurement access, while informative reusable measurements can further reduce the cost of the same clustering task.

Empirical validation of discovery, reuse, and scaling. Ideal Born-rule simulations validate both sources of measurement savings: a more informative measurement basis reduces connection-discovery cost by 66.80%, while reusable single-state evidence reduces mean total measurement cost by 52.53% on heterogeneous bounded-degree graphs when the measurement basis, pilot, candidate structures, and geometric inference are held fixed.

Closest to our setting are budgeted similarity acquisition and task-aware quantum-kernel measurement. Prior work reduces similarity queries or allocates quantum shots according to downstream task sensitivity (Shamir & Tishby, 2011; Wauthier et al., 2012; Shastry et al., 2022; Xu et al., 2026a). Q-Cert instead certifies the exact fidelity-threshold partition with false-certification control and abstention, while reusing single-state measurements across many pairwise relations. Other related settings are discussed in Appendix E (Xia & Huang, 2022; Graf et al., 2025; Thuot et al., 2026; Fotakis et al., 2023; Huang et al., 2020).

2 Task and Measurement Access

2.1 Clustering Task

We are given n repeatably preparable pure states $\rho_i = |\psi_i\rangle\langle\psi_i|$, $i \in [n] = \{1, \dots, n\}$, which may be supplied directly or generated by a quantum feature map $|\psi_i\rangle = U(x_i)|0\rangle$. We measure the similarity between states i and j by their squared fidelity $F_{ij} = |\langle\psi_i|\psi_j\rangle|^2 = \mathrm{Tr}(\rho_i\rho_j)$. The resulting matrix F is the quantum Gram matrix and is not assumed to be known.

Fix a threshold $\tau \in (0, 1)$ before measurement acquisition and define the fidelity-threshold graph $G_\tau = ([n], \{(i, j) : i < j,\ F_{ij} \geq \tau\})$. Our clustering target is its connected-component partition,

$$\mathcal{C}_\tau(F) = \mathrm{CC}(G_\tau), \tag{1}$$

where CC returns the connected components of a graph. This target is equivalent to single-linkage clustering cut at Fubini–Study distance $\arccos\sqrt{\tau}$. In other words, two samples belong to the same cluster if they can be connected by a path whose pairwise fidelities are all at least τ. The number of clusters is not known in advance.

2.2 Measurement Access and Certification

Measurement access. Let $\mathcal{B}$ be the finite menu of allowed bases. Measuring state i in basis $b \in \mathcal{B}$ returns a classical outcome $x \sim p_i^{(b)}$ on $\mathcal{X}^{(b)}$ at unit oracle cost; an Echo test on (i, j) returns $\mathrm{Ber}(F_{ij})$ at cost two. Retaining the full single-state outcome allows reuse across multiple pairwise relations. Echo requires preparation/inverse access, beyond preparing individual copies. The algorithm sees only sampled outcomes, not exact states, fidelities, or labels; basis design is outside our scope. SWAP access, used only in the lower-bound comparison, is described in Appendix B.

Certification objective and cost. Given a risk level $\delta \in (0, 1)$ and an oracle-call budget B, the algorithm returns $(\widehat{\mathcal{C}}, A)$ with $A \in \{0, 1\}$. Here $A = 1$ means that the returned partition is certified, whereas $A = 0$ means that the algorithm abstains. On abstention, the partition field is set to $\perp$

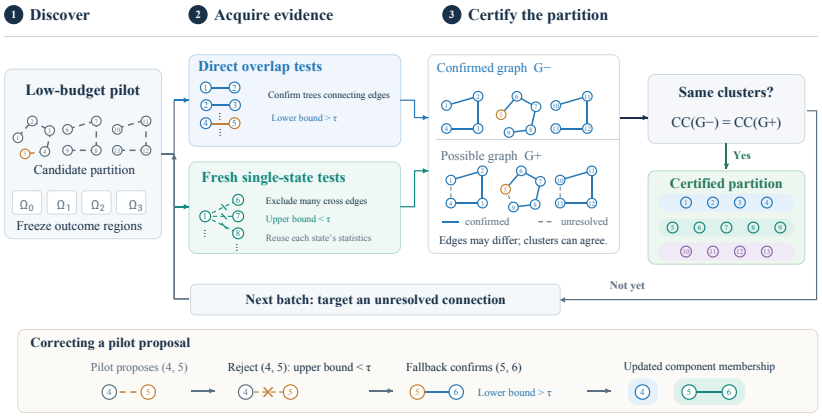

Figure 2: The framework of the proposed Q-Cert.

and is ignored. We refer to $A = 1$ as acceptance and reserve certification for the validity guarantee attached to an accepted partition. We require

$$\Pr\Big(A = 1,\ \widehat{\mathcal{C}} \neq \mathcal{C}_\tau(F)\Big) \leq \delta. \tag{2}$$

Thus, the guarantee controls the probability of certifying an incorrect partition. If the available budget is insufficient to determine the partition with the required confidence, the algorithm may abstain rather than return an uncertified result.

We measure quantum cost by the total number of state-preparation and inverse-oracle calls, denoted by B_Q. This includes pilot measurements, independent validation measurements, direct overlap tests, unsuccessful acquisition attempts, and fallback measurements. Classical computation, circuit depth, and hardware runtime are treated as separate resources and are not included in B_Q. We assume fresh independent state preparations with fixed measurement distributions during each run.

3 Q-Cert: Online Task-Sufficient Acquisition

Q-Cert proceeds in three stages. First, a low-cost pilot identifies candidate structure and connections that guide acquisition. Second, Q-Cert collects two complementary types of evidence: direct overlap measurements establish within-block connectivity, while reusable single-state measurements rule out possible connections between different blocks. Finally, Q-Cert combines the accumulated evidence and continues measuring only while the remaining uncertainty can still change the partition; once the partition is uniquely determined, it stops and certifies the result. Figure 2 summarizes the acquisition and certification loop, including correction of an erroneous pilot connection.

3.1 Pilot Proposal

Q-Cert uses all allowed bases. For each $b \in \mathcal{B}$, it first performs a small number of initial single-state measurements. From these observations, it proposes a candidate partition $\widehat{\mathcal{C}}^{(b)} = \{C_1^{(b)}, \ldots, C_{K_b}^{(b)}\}$. Within each candidate block $C_r^{(b)}$, Q-Cert selects $|C_r^{(b)}| - 1$ candidate edges that connect all states in the block; these edges form a spanning tree $T_r^{(b)}$, and their union is the support forest $T^{(b)}$. The forest identifies a small set of within-block connections to be verified by later direct measurements.

The pilot also fixes pairwise-disjoint outcome sets: a common set $\Omega_0^{(b)}$ and block-specific sets $\Omega_1^{(b)}, \ldots, \Omega_{K_b}^{(b)}$ in the outcome space of basis b; these sets need not cover the entire outcome space. Fresh single-state measurements estimate each state's probability mass on these sets. For states assigned to different candidate blocks, concentration on different block-specific sets with little leakage yields an upper bound on their fidelity, allowing possible cross-block edges to be excluded.

The pilot determines only which relations and outcome sets are subsequently tested. All certification bounds are constructed from fresh measurements. Thus, the pilot partition need not be correct: a

support forest containing a directly certified nonedge is disabled, while its frozen outcome sets remain valid for the cross-block bounds for which they were defined.

Pilot construction. The pilot is a proposal module rather than part of the certificate. In general, any rule that maps pilot outcomes to candidate blocks, sparse support forests, and disjoint outcome regions can be used, provided these objects are frozen before independent validation. We use a support-intersection rule: states sharing sufficiently frequent noncommon outcomes are connected, the connected components define candidate blocks, spanning trees provide connectivity proposals, and the retained outcomes associated with each block define its frozen region. We give the concrete construction and its measurement guarantee in Appendices B.4 and B.5.

3.2 Direct Connectivity and Reusable Separation Evidence

Direct pair measurements. After the pilot stage, Q-Cert directly measures selected pairs needed to establish or refine the current partition. It first tests the candidate within-block connections proposed by the support forests. Repeated Echo measurements of a selected pair (i, j) yield a Clopper–Pearson confidence interval $L_{ij}^{\mathrm{pair}} \le F_{ij} \le U_{ij}^{\mathrm{pair}}$. Hence $L_{ij}^{\mathrm{pair}} > \tau$ certifies an edge, whereas $U_{ij}^{\mathrm{pair}} < \tau$ certifies a nonedge. If the pilot structures and reusable single-state measurements leave some partition-relevant relation unresolved, Q-Cert falls back to direct Echo measurements of that pair until it is certified as an edge or a nonedge.

Reusable single-state measurements. Consider a basis $b \in \mathcal{B}$ and a state $i \in C_r^{(b)}$. Using the pilot-fixed outcome sets, define $u_i^{(b)} = p_i^{(b)}(\Omega_0^{(b)})$ as the common-set probability, and $e_i^{(b)} = p_i^{(b)}\Big(\mathcal{X}^{(b)} \setminus \big(\Omega_0^{(b)} \cup \Omega_r^{(b)}\big)\Big)$ as the leakage probability outside both the common set and the set associated with block $C_r^{(b)}$. Fresh single-state measurements estimate these two probabilities and yield Clopper–Pearson bounds $L_i^{(b)} \le u_i^{(b)} \le U_i^{(b)}$ and $e_i^{(b)} \le E_i^{(b)}$. For states $i \in C_r^{(b)}$ and $j \in C_s^{(b)}$ with $r \neq s$, the monotonicity of fidelity under measurement gives the valid upper bound

$$\overline{F}_{ij}^{(b)} = \min\left\{1, \left[\sqrt{U_i^{(b)} U_j^{(b)}} + \sqrt{(1-L_i^{(b)})E_j^{(b)}} + \sqrt{(1-L_j^{(b)})E_i^{(b)}}\right]^2\right\}, \qquad F_{ij} \le \overline{F}_{ij}^{(b)}. \quad (3)$$

Thus $\overline{F}_{ij}^{(b)} < \tau$ certifies (i, j) as a nonedge. If several bases provide valid bounds for the same pair, Q-Cert uses the smallest one. The same validation measurements of state i can bound many cross-block pairs involving i. Direct measurements therefore establish within-block connectivity pairwise, whereas single-state measurements can exclude multiple cross-block connections simultaneously.

3.3 Online Acquisition and Certification

Q-Cert combines the information obtained from direct and single-state measurements into pairwise fidelity bounds. For each pair (i, j), direct Echo measurements provide $L_{ij}^{\mathrm{pair}} \le F_{ij} \le U_{ij}^{\mathrm{pair}}$. The reusable single-state construction in Section 3.2 provides an additional upper bound $\overline{F}_{ij} = \min_b \overline{F}_{ij}^{(b)}$ over all applicable bases, with $\overline{F}_{ij} = 1$ if no such basis is available.

Covered fidelity intervals are additionally propagated through the Fubini–Study triangle inequality to give L_{ij}^{geom} and U_{ij}^{geom}; the rules are in Appendix A.1. The combined bounds are

$$L_{ij}^{F} = \max\left\{L_{ij}^{\mathrm{pair}}, L_{ij}^{\mathrm{geom}}\right\}, \qquad U_{ij}^{F} = \min\left\{U_{ij}^{\mathrm{pair}}, \overline{F}_{ij}, U_{ij}^{\mathrm{geom}}\right\}, \quad (4)$$

where a missing lower or upper bound is interpreted as 0 or 1, respectively. Using these bounds, Q-Cert maintains a *confirmed graph* $G^- = ([n], E^-)$ and a *possible graph* $G^+ = ([n], E^+)$, where

$$E^- = \{(i, j) : L_{ij}^F > \tau\}, \qquad E^+ = \{(i, j) : U_{ij}^F \ge \tau\}. \quad (5)$$

The confirmed graph contains edges already established to exist, while the possible graph contains every edge that has not yet been ruled out. Q-Cert stops and certifies the partition once

$$\mathrm{CC}(G^-) = \mathrm{CC}(G^+). \quad (6)$$

Algorithm 1 Online Q-Cert

Require: Allowed basis menu $\mathcal{B}$, threshold τ, risk level δ, and oracle-call budget B
Ensure: Certified partition $(\widehat{\mathcal{C}}, A = 1)$ or abstention $(\perp, A = 0)$

Run the pilot in each basis and freeze the proposed partitions, support forests, and outcome sets
Start fresh validation streams and set $S = P = 0$
loop
Stage 1: update evidence and test certification
    Update direct and reusable bounds; apply free geometric propagation
    if any combined interval is contradictory, i.e., $L_{ij}^F > U_{ij}^F$ for some (i, j) **then**
        return $(\perp,\ A = 0)$
    end if
    Form the confirmed graph G^- and possible graph G^+
    if $\mathrm{CC}(G^-) = \mathrm{CC}(G^+)$ **then**
        return $(\mathrm{CC}(G^-),\ A = 1)$
    end if
    Disable every support forest containing a certified nonedge
Stage 2: propose partition-relevant measurements
    $a_P \leftarrow$ first available pair action ▷ forest → representative → fallback
    $a_S \leftarrow$ least-sampled eligible single-state validation stream
    if a_P and a_S are both unavailable **then**
        return $(\perp,\ A = 0)$
    end if
Stage 3: balance and acquire
    Select a_S if available and (a_P unavailable or $S \leq P$); otherwise select a_P
    Set the next batch by the 32-then-doubling schedule
    if the batch would exceed B **then**
        return $(\perp,\ A = 0)$
    end if
    Execute the batch and update S or P
end loop

At this point, any remaining uncertainty about individual edges cannot change the connected-component partition. The correctness of this stopping rule is formalized in Section 4.

A pair is *pending* if it remains in E^+ but joins distinct components of G^-. A state–basis stream (i, b) is *eligible* if i is an endpoint of a pending pair whose endpoints have distinct frozen pilot assignments in basis b. After each batch, Q-Cert applies free geometric propagation to existing intervals. If the partition remains unresolved, pair acquisition prioritizes pending support-forest edges, optionally uses bounded representative-pair Echo to improve propagation, and otherwise uses pending-pair Echo fallback. The controller balances post-pilot pair and single-state costs P and S; Algorithm 1 gives the schedule, and Appendix A.4 gives the representative-pair allowance and fallback guarantee. Because measurement selection and stopping depend on observed data, Q-Cert uses confidence bounds valid simultaneously over inspected streams and times. Pilot structures are fixed before validation. Section 4 gives the coverage argument and certification guarantee.

4 Certification and Measurement Complexity

We first establish that Q-Cert certifies only the target partition, regardless of the quality of the pilot proposal. We then characterize the structural conditions under which the complete online controller achieves near-linear measurement complexity. We further distinguish the benefits of targeting the partition from those of access to reusable measurement information.

4.1 Safe Certification

On simultaneous coverage, $E^- \subseteq E_\tau \subseteq E^+$; hence equality of the connected-component partitions of G^- and G^+ identifies $\mathcal{C}_\tau(F)$ (Appendix A.1). Since acquisition and stopping are adaptive, the confidence bounds must cover all inspected streams and times jointly.

Theorem 1 (Universal safe certification)**.** *Fix the threshold τ before measurement acquisition. Conditional on the pilot observations, assume that repeated Echo measurements of each fixed pair (i, j) are independent $\mathrm{Ber}(F_{ij})$ observations, and that repeated measurements of each fixed state i in basis b are independent draws from $p_i^{(b)}$. The validation measurements used to construct pilot-dependent bounds are independent of the pilot observations.*

If the confidence bounds used by Algorithm 1 have joint noncoverage probability at most δ, then, for any collection of pure states,

$$\Pr\Big(A = 1,\, \widehat{\mathcal{C}} \neq \mathcal{C}_\tau(F)\Big) \leq \delta. \tag{7}$$

No correctness assumption is required for the pilot proposal and no single-state basis is required to provide useful separation evidence.

Incorrect certification can therefore occur only when joint coverage fails. If no F_{ij} equals τ, unrestricted pending-pair Echo fallback also guarantees finite completion on coverage when the budget does not bind (Appendix A.4).

4.2 Online Measurement Complexity

We first analyze the online measurement cost conditional on a frozen pilot output. Appendix B.4 later instantiates the pilot with a concrete finite-sample construction and gives the corresponding end-to-end guarantee. Its oracle cost B_{pilot} is included in the total budget.

Connectivity cost. Let $L = |\mathcal{B}|$. For each basis $b \in \mathcal{B}$, the pilot proposes a support forest $T^{(b)}$. Write $\mathcal{T} = \bigcup_{b\in\mathcal{B}} T^{(b)}$, so $|\mathcal{T}| \leq L(n-1)$. For each proposed edge $e = (i, j) \in \mathcal{T}$, define its threshold-angle gap $\Delta_e = |\arccos\sqrt{F_{ij}} - \arccos\sqrt{\tau}|$. Appendix A.2 defines $Q(x)$ as the first scheduled sample count at which the confidence-angle radius is below x; for fixed $x > 0$ and a polynomial-size confidence universe, $Q(x) = O(\log(n/\delta))$. A sufficient pairwise workload is $A_T = 2\sum_{e\in\mathcal{T}} Q(\Delta_e/2)$, where the factor two is the oracle cost of one Echo observation.

Separation cost. When G_τ has multiple connected components, possible cross-component edges must also be excluded. Let $\mathcal{P}_\tau$ denote the set of pairs belonging to different true components. For $(i, j) \in \mathcal{P}_\tau$, let $\mathcal{B}_{ij} \subseteq \mathcal{B}$ contain the bases in which i and j are assigned to different frozen pilot blocks. If $\mathcal{B}_{ij} = \varnothing$ for some cross-component pair, set $\Gamma = 0$; otherwise define

$$\Gamma = \min_{(i,j)\in\mathcal{P}_\tau} \max_{b\in\mathcal{B}_{ij}} \left[\sqrt{\tau} - \sqrt{u_i^{(b)}u_j^{(b)}} - \sqrt{(1-u_i^{(b)})e_j^{(b)}} - \sqrt{(1-u_j^{(b)})e_i^{(b)}}\right]. \tag{8}$$

Thus, $\Gamma > 0$ means that every true cross-component pair can be excluded with positive margin by the reusable bound in at least one allowed basis. A sufficient single-state workload is $A_S = LnQ(\Gamma/12)$; when G_τ is connected, set $A_S = 0$.

Theorem 2 (Online measurement complexity)**.** *Condition on a frozen pilot output. Suppose there exists $\mathcal{B}_\star \subseteq \mathcal{B}$ such that, with $\mathcal{T}_\star = \bigcup_{b\in\mathcal{B}_\star} T^{(b)}$, $\mathcal{T}_\star \subseteq E_\tau$ and $\mathrm{CC}([n], \mathcal{T}_\star) = \mathcal{C}_\tau(F)$. Assume further that $\Delta_e > 0$ for all $e \in \mathcal{T}$ and, whenever $\mathcal{P}_\tau \neq \varnothing$, that $\Gamma > 0$. On the joint-coverage event of Theorem 1, Algorithm 1 returns $(\mathcal{C}_\tau(F), A = 1)$ within*

$$B_Q \leq B_{\mathrm{pilot}} + 3\max\{A_T, A_S\} + 64 \tag{9}$$

oracle calls, provided the available budget covers this amount.

The first condition means that some pilot forests form a clean connectivity witness for the true partition. The algorithm need not know which forests provide this witness or the values of $\{\Delta_e\}$ and Γ. Here A_T accounts for verification of all proposed forest edges, including unsuccessful proposals, whereas A_S accounts for reusable state-level evidence for excluding cross-component connections. The support-tree priority, global least-sampled rule, and cumulative-cost balancing in Algorithm 1 keep the online cost within a constant factor of these sufficient workloads. In particular, if $L = |\mathcal{B}|$ is fixed, $B_{\mathrm{pilot}} = O(n\log(n/\delta))$, $\inf_{e\in\mathcal{T}} \Delta_e$ is bounded below by a positive constant, and, when separation is required, so is Γ, then $A_T, A_S = O(n\log(n/\delta))$ and hence $B_Q = O(n\log(n/\delta))$. The proof is given in Appendix A.3. The following two results isolate the benefits of the recovery target and reusable measurement access, respectively.

Proposition 1 (Same-interface task separation)**.** *Fix $0 < \delta \leq 1/4$. For the pure-state family in Appendix B.6, Q-Cert certifies the partition with probability at least $1 - \delta$ using $O\big(n \log \frac{n}{\delta}\big)$ oracle calls. Under the same measurement interface, any full threshold-graph certifier with error and abstention probabilities at most δ has worst-case expected cost $\Omega(n\eta^{-2} \log(n/\delta))$, where η is the threshold gap of redundant within-component pairs.*

The hard pairs are redundant within components and need not be resolved for partition certification. Taking $\eta = \Theta(n^{-1/2})$ therefore gives a quadratic-to-near-linear separation under the same measurement interface. The construction and proof are given in Appendix B.6.

Proposition 2 (Measurement-access separation)**.** *Fix $0 < \delta \leq 1/4$. For the hidden-bridge family in Appendix B, Q-Cert with informative reusable single-state measurements certifies every instance with probability at least $1 - \delta$ using $O\big(n \log \frac{n}{\delta}\big)$ oracle calls. Under the blind single-state interface, any algorithm using scalar Echo/SWAP tests that terminates almost surely and has incorrect-certification and abstention probabilities each at most δ on the no-bridge instance and every bridge alternative requires $\Omega(n^2 \log(1/\delta))$ expected calls on the no-bridge instance and $\Omega(n^2)$ calls on average over bridge alternatives.*

Thus, reusable state-level information can change the asymptotic cost of the same clustering task, not just a constant factor. The construction and proof are in Appendix B.

End-to-end realizability. The conditions of Theorem 2 are realized by a structured pure-state family covering graph topologies of fixed maximum degree d. In an informative coordinate basis, the support-intersection pilot recovers, with high probability, a sparse clean witness spanning every true component, while providing reusable cross-component separation with fixed positive margin. Appendix B.4 gives the construction, pilot procedure, and finite-sample recovery guarantee.

Proposition 3 (End-to-end near-linear certification)**.** *Fix the maximum degree d and a finite allowed measurement-basis menu containing the informative coordinate basis. For the bounded-degree state family of Appendix B.4, with fixed positive support probabilities, threshold margins, and reusable separation margin, Q-Cert correctly certifies $\mathcal{C}_\tau(F)$ with probability at least $1 - \delta$ using*

$$B_Q = O\Big(n \log \frac{n}{\delta}\Big) \tag{10}$$

quantum-oracle calls, uniformly over graph topology and component sizes.

The proof and pilot risk allocation are given in Appendix B.4. Appendix B.5 extends the guarantee to controlled pure-state perturbations.

5 Experiments

We evaluate three questions using ideal Born-rule simulations: connection discovery, reusable exclusion, and end-to-end scaling. Full protocols and risk allocations are in Appendix C. Methods are paired on the same generated instances; the controller sees neither exact fidelities nor target partitions. Oracle cost includes every measurement stage, and savings are one minus the ratio of paired mean costs. All reported confirmation and scaling runs certify correctly within their caps.

5.1 Mechanism Validation

Connection discovery. Q-Cert ZX permits both $Z_{\text{flag}} \otimes Z_{\text{data}}$ and $Z_{\text{flag}} \otimes X_{\text{data}}$ measurements, whereas Q-Cert Z permits only the former. On held-out bridge-present instances at $n = 32$, averaging both hidden Hadamard orientations, ZX reduces mean oracle cost by 66.80% relative to Z. These connected instances need no cross-component validation, isolating the additional basis's contribution to connection discovery. Detailed results appear in Appendix D.1.

Reusable exclusion evidence. We compare Q-Cert with full-risk direct certification using the same measurement basis, paid pilot, candidate forests, free geometric propagation, and representative-pair Echo rule. The direct method disables single-state validation and allocates its full error budget to pairwise tests. On heterogeneous bounded-degree graphs, Q-Cert reduces mean oracle cost by 52.53% at $\epsilon = 0$ and 46.56% at $\epsilon = .01$, consistently across all four graph topologies. Table 1

Table 1: Certification cost in millions of oracle calls. Discovery compares Q-Cert ZX with Q-Cert Z; other rows use full-risk direct. Brackets give pointwise 95% paired bootstrap intervals for ratio-of-means savings. The $n = 96$ confirmation and scaling rows use independent graph blocks.

Study	n	Q-Cert	Comparator	Saving [95% interval]
Discovery	32	0.131	0.394	66.80% $[65.92, 67.63]$
Confirmation, $\epsilon = 0$	96	2.589	5.453	52.53% $[50.51, 54.21]$
Confirmation, $\epsilon = 0.01$	96	2.933	5.488	46.56% $[44.31, 48.45]$
Scaling	24	0.404	0.394	-2.49% $[-5.42, 0.34]$
Scaling	48	0.975	1.479	34.08% $[32.42, 35.76]$
Scaling	96	2.630	5.858	55.10% $[53.30, 56.73]$
Scaling	192	6.004	24.329	75.32% $[74.32, 76.25]$

summarizes the overall certification costs, savings, and uncertainty intervals. The corresponding pairwise-work decomposition, topology-level savings, and supplementary structural comparisons are provided in Appendix D.2.

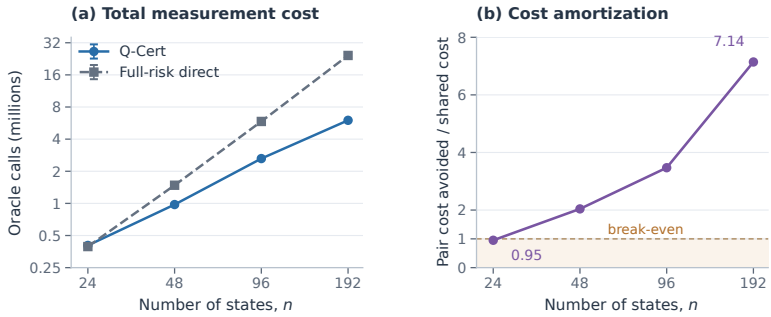

Figure 3: Same-interface scaling over 12 fresh graph blocks per size. (a) Mean oracle cost with 95% bootstrap intervals. (b) Avoided pair cost relative to validation cost; values above one indicate net savings. Ratios use aggregate costs.

5.2 End-to-End Scaling

We evaluate end-to-end scaling on the bounded-degree family of Proposition 3 (Fig. 3). With the same measurement access, Q-Cert and full-risk direct have comparable cost at $n = 24$. Savings increase to 34.08%, 55.10%, and 75.32% at $n = 48, 96, 192$, respectively. Over the tested range, the ratio of avoided pair cost to single-state validation cost rises from 0.95 to 7.14. As the problem grows, the same state-level evidence can rule out more candidate connections, amortizing the additional validation cost. Appendix A.5 makes this intuition quantitative for a frozen verification plan, giving a sufficient-workload comparison between reusable and pairwise evidence. Figure 3(b) provides a qualitative empirical counterpart, showing how the realized amortization ratio changes with problem size. Complete results and the supplementary hidden-bridge size sweep are in Appendix D.3.

6 Conclusion

Q-Cert certifies fidelity-threshold connected components without recovering the full fidelity graph. Its graph-sandwich rule is universally safe for arbitrary pure-state inputs, while sparse connectivity witnesses and reusable separation yield near-linear oracle complexity on structured families. Under the same measurement interface, partition certification can be asymptotically cheaper than full graph recovery; complementary hidden-bridge results isolate the additional benefit of reusable single-state information. Experiments support both mechanisms and show increasing amortization of reusable measurements with problem size.

AI USE STATEMENT

We used generative AI tools to assist with the following:
(1) manuscript writing, language polishing, and figure refinement;
(2) partial organization of proofs, checking theoretical proofs, and improving rigor where needed;
(3) literature search and related-work review;
(4) assisting with the writing and organization of experimental code; and
(5) organizing and integrating the appendix.

All AI-assisted work was reviewed by the authors. The authors take full responsibility for the final manuscript, code, and experimental results.

REPRODUCIBILITY STATEMENT

Appendices A and B provide the proofs and state constructions. Appendix C specifies the experimental designs, comparisons, risk allocations, and statistical units. The accompanying code package includes the experiment implementations, frozen configurations, seeds, and source-hash checks. It provides scripts for regenerating the core result tables and data-driven figures, and for collecting supplementary summaries, after rerunning the corresponding experiments. It does not include saved measurement records, compact result snapshots, or historical action trajectories. The package README distinguishes experiment reruns, statistical recomputation, and controller replay.

REFERENCES

Alan Abraham-Mexicano, Carlos V Muro-Medina, Valentin Flores-Payan, Elisa Ramos-Pinzon, Carolina L Recio-Colmenares, Roxana B Recio-Colmenares, and Cesar A Garcia-Garcia. Quantum machine learning for water pollution profiling in the rio santiago basin. *Quantum Reports*, 8 (3):60, 2026.

C. J. Clopper and E. S. Pearson. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4):404–413, 1934. doi: 10.1093/biomet/26.4.404.

Brian Eriksson, Gautam Dasarathy, Aarti Singh, and Rob Nowak. Active clustering: Robust and efficient hierarchical clustering using adaptively selected similarities. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, volume 15 of *Proceedings of Machine Learning Research*, pp. 260–268. PMLR, 2011. URL https://proceedings.mlr.press/v15/eriksson11a.html.

Dimitris Fotakis, Evangelia Gergatsouli, Charilaos Pipis, Miltiadis Stouras, and Christos Tzamos. Graph connectivity with noisy queries. In *48th International Symposium on Mathematical Foundations of Computer Science (MFCS 2023)*, volume 272 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 47:1–47:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi: 10.4230/LIPIcs.MFCS.2023.47. URL https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.MFCS.2023.47.

Maximilian Graf, Victor Thuot, and Nicolas Verzelen. Clustering items through bandit feedback: Finding the right feature out of many. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 20296–20325. PMLR, 2025. URL https://proceedings.mlr.press/v267/graf25a.html.

Vojtěch Havlíček, Antonio D Córcoles, Kristan Temme, Aram W Harrow, Abhinav Kandala, Jerry M Chow, and Jay M Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747):209–212, 2019.

Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16:1050–1057, 2020. doi: 10.1038/s41567-020-0932-7.

Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier. On the complexity of best-arm identification in multi-armed bandit models. *Journal of Machine Learning Research*, 17(1):1–42, 2016. URL https://jmlr.org/papers/v17/kaufman16a.html.

Iordanis Kerenidis and Jonas Landman. Quantum spectral clustering. *Physical Review A*, 103(4): 042415, 2021. doi: 10.1103/PhysRevA.103.042415.

Feng Liu, Wenkai Xu, Jie Lu, Guangquan Zhang, Arthur Gretton, and Danica J Sutherland. Learning deep kernels for non-parametric two-sample tests. In *International conference on machine learning*, pp. 6316–6326. PMLR, 2020.

Edward Milsom, Ben Anson, and Laurence Aitchison. Convolutional deep kernel machines. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=1oqedRt6Z7.

Artur Miroszewski. Adaptive measurement allocation for learning kernelized SVMs under noisy observations. *arXiv preprint arXiv:2605.22275*, 2026. URL https://arxiv.org/abs/2605.22275v2. Version 2.

Feiping Nie, Chaodie Liu, Rong Wang, and Xuelong Li. A novel and effective method to directly solve spectral clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46 (12):10863–10875, 2024.

Maria Schuld and Nathan Killoran. Quantum machine learning in feature hilbert spaces. *Physical review letters*, 122(4):040504, 2019.

Uri Shaham, Kelly Stanton, Henry Li, Boaz Nadler, Ronen Basri, and Yuval Kluger. Spectralnet: Spectral clustering using deep neural networks. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=HJ_aoCyRZ.

Ohad Shamir and Naftali Tishby. Spectral clustering on a budget. In *Proceedings of AISTATS*, volume 15 of *Proceedings of Machine Learning Research*, pp. 661–669, 2011. URL https://proceedings.mlr.press/v15/shamir11a.html.

Abhay Shastry, Abhijith Jayakumar, Apoorva Patel, and Chiranjib Bhattacharyya. Shot-frugal and robust quantum kernel classifiers. *arXiv preprint arXiv:2210.06971*, 2022. URL https://arxiv.org/abs/2210.06971.

Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence*, 22(8):888–905, 2000.

Donovan Slabbert, Dean Brand, and Francesco Petruccione. Quantum spectral clustering: Comparing parameterized and neuromorphic quantum kernels. *APL Quantum*, 3(2):026106, 2026. doi: 10.1063/5.0291037.

Supanut Thanasilp, Samson Wang, Marco Cerezo, and Zoë Holmes. Exponential concentration in quantum kernel methods. *Nature communications*, 15(1):5200, 2024.

Victor Thuot, Sebastian Vogt, Debarghya Ghoshdastidar, and Nicolas Verzelen. Nonparametric kernel clustering with bandit feedback. *arXiv preprint arXiv:2601.07535*, 2026. URL https://arxiv.org/abs/2601.07535.

Jun Towada. Beyond euclidean clustering: Quantum kernel-driven decomposition for the capacitated vehicle routing problem. *TechRxiv*, 2026. doi: 10.36227/techrxiv.177155894.43289947/v1.

Ulrike von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17(4):395–416, 2007. doi: 10.1007/s11222-007-9033-z.

Fabian L. Wauthier, Nebojsa Jojic, and Michael I. Jordan. Active spectral clustering via iterative uncertainty reduction. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2012. URL https://people.eecs.berkeley.edu/~jordan/papers/wauthier-jojic-jordan-kdd12.pdf.

Jinghui Xia and Zengfeng Huang. Optimal clustering with noisy queries via multi-armed bandit. In *Proceedings of ICML*, volume 162 of *Proceedings of Machine Learning Research*, pp. 24315–24331, 2022. URL https://proceedings.mlr.press/v162/xia22a.html.

Jian Xu, Chao Li, Delu Zeng, John Paisley, and Qibin Zhao. AQKA: Active quantum kernel acquisition under a shot budget. *arXiv preprint arXiv:2605.14672*, 2026a. URL https://arxiv.org/abs/2605.14672v2. Version 2.

Jian Xu, Artur Miroszewski, John Paisley, Delu Zeng, and Qibin Zhao. Active quantum kernel acquisition for gaussian process regression. *arXiv preprint arXiv:2606.28833*, 2026b. URL https://arxiv.org/abs/2606.28833v3. Version 3.

A Certification and Online Complexity Proofs

A.1 Reusable evidence and safe acceptance

Condition on the pilot. In one allowed pilot basis it fixes pairwise disjoint regions $\Omega_0, \Omega_1, \ldots, \Omega_K$ and assignments $c(i)$. Unassigned outcomes are permitted. Write $u_i = p_i(\Omega_0)$ and $e_i = p_i(\Omega \setminus (\Omega_0 \cup \Omega_{c(i)}))$. One validation outcome updates both event indicators, and costs one oracle call. The indicators need not be independent.

Proposition 4 (Shared fidelity bound)**.** *If $c(i) \neq c(j)$, $L_i \leq u_i \leq U_i$, and $e_i \leq E_i$ for both states, then the upper bound in* (3) *is valid.*

Proof. Coarsen the measurement into Ω_0, $\Omega_{c(i)}$, and their complement. Fidelity monotonicity under measurement gives

$$\sqrt{F_{ij}} \leq \sqrt{u_i u_j} + \sqrt{p_i(\Omega_{c(i)}) p_j(\Omega_{c(i)})} + \sqrt{e_i\, p_j(\Omega \setminus (\Omega_0 \cup \Omega_{c(i)}))}.$$

The disjoint assignments imply $p_j(\Omega_{c(i)}) \leq e_j$. The remaining factors satisfy $p_i(\Omega_{c(i)}) \leq 1 - u_i$ and $p_j(\Omega \setminus (\Omega_0 \cup \Omega_{c(i)})) \leq 1 - u_j$. Thus the right side is at most $\sqrt{u_i u_j} + \sqrt{(1-u_i)e_j} + \sqrt{(1-u_j)e_i}$. Substituting the covered confidence endpoints, squaring, and clipping at one proves the result. □

Simultaneous coverage under adaptive sampling. For a Bernoulli stream, let k be its success count after N observations. The two-sided Clopper–Pearson (CP) interval at noncoverage level α has lower endpoint zero if $k = 0$, and otherwise the $\alpha/2$ quantile of $\mathrm{Beta}(k, N - k + 1)$. Its upper endpoint is one if $k = N$, and otherwise the $1 - \alpha/2$ quantile of $\mathrm{Beta}(k + 1, N - k)$ (Clopper & Pearson, 1934).

For a fixed universe of m parameters with risk allocation ε, use levels $\alpha_h = 6\varepsilon/[\pi^2 m(h+1)^2]$ at counts $N_h = 32 \cdot 2^h$. Associate each measurement action with an infinite sequence of fresh independent outcomes, conditional on the pilot. Adaptive acquisition reveals prefixes of these sequences. Each event count has the binomial marginal law at every fixed prefix, even when several events are extracted from the same shot. A union bound over the m parameters and all scheduled prefixes bounds any noncoverage by $\sum_{j=1}^{m} \sum_{h \geq 0} \alpha_h = \varepsilon$. For the main algorithm, take the pair universe to contain all $\binom{n}{2}$ pairs and the shared universe to contain both region probabilities for every allowed state–basis stream. A two-sided interval for the common probability covers its lower and upper endpoints together; it requires only one parameter allocation. Enlarging the universe to include auxiliary statistics remains valid. Allocate $\varepsilon_P + \varepsilon_S = \delta_v \leq \delta$ to the two pools. Summing their failure bounds gives total noncoverage at most δ_v. The usual split is $\varepsilon_P = \varepsilon_S = \delta_v/2$; pilot-recovery risk is separate and is needed for completion, not safety.

Historical intersections remain valid on this simultaneous event. Unmeasured parameters keep their full physical range. Pilot-selected regions are fixed before their independent validation streams begin; later changes to the graph partition do not change these regions or their original applicability. In particular, a shared bound is used only across distinct assignments in the basis that defines it.

Validity of geometric propagation. For any covered fidelity interval $[L_{ij}^F, U_{ij}^F]$, set $\ell_{ij} = \arccos \sqrt{U_{ij}^F}$ and $v_{ij} = \arccos \sqrt{L_{ij}^F}$. These enclose the true distance $d_Q(i,j)$. For any intermediate state k, the triangle and reverse triangle inequalities permit the updates

$$\ell_{ij} \leftarrow \max\{\ell_{ij}, \ell_{ik} - v_{jk}, \ell_{jk} - v_{ik}, 0\},$$
$$v_{ij} \leftarrow \min\{v_{ij}, v_{ik} + v_{jk}, \pi/2\}.$$

Both preserve coverage. Mapping back gives $L_{ij}^{\mathrm{geom}} = \cos^2 v_{ij}$ and $U_{ij}^{\mathrm{geom}} = \cos^2 \ell_{ij}$. Compositions of these valid updates, including paths through representatives, remain valid even when their choice depends on the observed history. They introduce no further statistical event or risk allocation. Refinement from existing intervals is classical computation and incurs no additional oracle calls.

Correctness of the stopping rule. If $(i,j) \in E^-$, then $F_{ij} \geq L^F_{ij} > \tau$, so it is a true edge. Conversely, a true edge satisfies $U^F_{ij} \geq F_{ij} \geq \tau$ and therefore belongs to E^+. Thus $E^- \subseteq E_\tau \subseteq E^+$. Every path in the confirmed graph is a true path, and every true path is possible. The corresponding connected-component partitions are therefore ordered by refinement. If the two outer partitions agree, the true partition between them is the same. This argument also covers $F_{ij} = \tau$: such a pair need not be confirmed, but it cannot be incorrectly excluded on coverage.

Proof of Theorem 1. On simultaneous coverage, direct intervals and Proposition 4 bound every fidelity correctly. By the validity of geometric propagation established above, all propagated intervals also preserve coverage. Consequently $G^- \subseteq G_\tau \subseteq G^+$ at every status check. Every component of the confirmed graph is contained in a true component, and every true component is contained in a possible-graph component. Equality in (6) therefore identifies the true partition. Contradictory intervals cannot occur on coverage and lead to abstention otherwise. Conditional on any pilot, wrong acceptance has probability at most δ_v; averaging over the pilot proves the claimed unconditional guarantee. No correctness assumption on the pilot is used.

A.2 A COMMON CONFIDENCE-ANGLE ENVELOPE

The cost proof uses a deterministic envelope for the actual CP intervals. Let m_P, m_S be the pair and shared parameter-universe sizes and assign $\delta_v/2$ to each pool. Set $M = 2\max\{m_P, m_S\}$ and define

$$r_h = \sqrt{\frac{\log[\pi^2 M(h+1)^2/(3\delta_v)]}{N_h}}, \qquad Q(x) = \min\{N_h : r_h < x\}, \qquad N_h = 32 \cdot 2^h.$$

The universe must cover every statistic used by the controller. For the two-basis implementation, $m_P = \binom{n}{2}$ and $m_S = 2n(q+4)$, where q is the public data-register size. The latter universe conservatively includes the stored single-bit statistics. A different basis menu requires the corresponding enlarged universe.

Lemma 1 (CP angle envelope)**.** *For empirical Bernoulli mean $\widehat{p} = k/N$, a two-sided CP interval $[L,U]$ with noncoverage level α satisfies* $[\arccos\sqrt{U}, \arccos\sqrt{L}] \subseteq [\arccos\sqrt{\widehat{p}} - r, \arccos\sqrt{\widehat{p}} + r] \cap [0, \pi/2]$*, where $r = \sqrt{\log(2/\alpha)/N}$. On coverage, each endpoint differs from the true probability angle by at most $2r$.*

Proof. For an interior lower endpoint, binomial-tail inversion and Chernoff's bound give $\alpha/2 = \Pr_L\{X \geq k\} \leq \exp[-N\,\mathrm{kl}(\widehat{p}, L)]$. Thus $\mathrm{kl}(\widehat{p}, L) \leq \log(2/\alpha)/N$. The upper endpoint follows from the lower tail in the same way. The remaining boundary endpoints coincide with $\widehat{p}$. For Bernoulli probabilities x, p, the KL–affinity inequality implies

$$\mathrm{kl}(x,p) \geq -2\log\Big(\sqrt{xp} + \sqrt{(1-x)(1-p)}\Big) \geq |\arccos\sqrt{x} - \arccos\sqrt{p}|^2.$$

The second inequality follows from $-2\log\cos t \geq t^2$. This proves containment, including boundary cases by limits. On coverage the true angle belongs to the contained interval, whose diameter is at most $2r$. □

Here Q is defined using the radius r_h; the resulting endpoint error is at most $2r_h$. Thus $Q(\Delta_e/2)$ makes the endpoint error strictly smaller than the threshold margin Δ_e, as used below. The common radius r_h dominates the radii of both actual confidence pools. Intersections can only tighten the intervals, so their endpoints remain within $2r_h$ of truth on coverage. For $0 < x \leq \pi/2$,

$$Q(x) = O\big(32 + x^{-2}\left[1 + \log(M/\delta_v) + \log\log(e^e/x)\right]\big).$$

To see this, put $A = \log[\pi^2 M/(3\delta_v)]$ and take the first dyadic count above $Cx^{-2}[A + 1 + \log\log(e^e/x)]$, or 32 if larger. For a sufficiently large absolute constant C, its index h satisfies $A + 2\log(h+1) < N_h x^2$. Dyadic rounding changes the bound by at most a constant factor. Thus $Q(x)$ is logarithmic in n/δ_v for fixed $x > 0$ and a polynomial-size confidence universe.

A.3 PROOF OF THE ONLINE COST BOUND

We use the following properties of Algorithm 1. All active forests are scanned before representative-pair Echo acquisition or unrestricted pending-pair fallback. A certified nonedge disables its entire forest, but not its frozen regions. Repeated proposals of the same pair share one physical stream. Shared candidates contain both endpoints of every eligible pending relation across all allowed bases, and the selected stream globally minimizes its validation count. When both pools are available, shared is selected exactly when cumulative post-pilot costs satisfy $S \le P$. Each stream starts at 32 observations and subsequently doubles its count. Acceptance is checked after every batch, and graph bounds are monotone.

Let C_t mean that the confirmed graph connects every true component, and let D_t mean that the possible graph contains no true cross-component edge. On coverage, these properties persist once established. In this proof $\mathcal{T}$ is the union of all initial proposal forests, including unsuccessful ones.

Connectivity work. If C_t fails, some edge in the clean forest union joins distinct confirmed components. The edge remains possible, and its clean forest cannot have been disabled on coverage. A pending tree action is therefore available. Global tree priority ensures that every pair measurement before connectivity is established is charged to $\mathcal{T}$. After $Q(\Delta_e/2)$ Echo trials, Lemma 1 puts the entire fidelity interval strictly on the correct side of the threshold. The edge is then confirmed or excluded and cannot be selected again as pending tree work. Doubling reaches rather than skips this scheduled count. Pair cost through the batch establishing C_t is consequently at most $A_T = 2\sum_{e\in\mathcal{T}} Q(\Delta_e/2)$. The clean witness must consist of complete forests: a true subforest inside a contaminated proposal could be disabled by a false extra edge.

Separation work. Write $h_{ij}^{(b)} = \sqrt{u_i^{(b)}u_j^{(b)}} + \sqrt{(1-u_i^{(b)})e_j^{(b)}} + \sqrt{(1-u_j^{(b)})e_i^{(b)}}$, and let $h_{ij}^{U,(b)}$ use the confidence endpoints in the shared bound. Every square-root probability factor differs from its true value by at most $2r_i$ or $2r_j$, since sine and cosine are 1-Lipschitz on $[0,\pi/2]$. Using $|ab - a'b'| \le |a-a'|+|b-b'|$ for factors in $[0,1]$ and summing the three products gives $0 \le h_{ij}^{U,(b)} - h_{ij}^{(b)} \le 6(r_i + r_j)$.

If D_t fails, a true cross-component pair remains possible. Such a pair also joins distinct confirmed components. By the definition of Γ, an allowed basis separates its frozen assignments and gives root margin at least $\Gamma > 0$. Both endpoint streams are eligible in that basis, regardless of whether its forest is still active. If both had reached $N_S = Q(\Gamma/12)$, the preceding inequality would make the shared upper bound strictly smaller than τ, a contradiction. Hence an eligible stream has count below N_S. Global least-count sampling selects a stream below this count, and its next doubling reaches at most N_S. At most Ln streams can be sampled. Shared cost through the batch establishing D_t is therefore at most $A_S = LnN_S$. An uninitialized stream has count zero, so the argument includes the first batch. For a single true component, D_t holds initially and $A_S = 0$.

Balancing after one requirement is complete. Whenever a pair action is available, a shared batch is selected only if $S \le P$. Its cost is at most $\max\{32, S\}$, so after the update $S \le 2P + 32$. Whenever a shared action is available, a pair batch is selected only if $P < S$; its cost is at most $\max\{64, P\}$, giving $P \le 2S + 64$ after that update.

Before either requirement is complete, costs are at most $A_T + A_S$. If separation finishes first, pair work remains available until connectivity finishes, and total pair work is bounded by A_T. Every later shared update leaves $S \le 2A_T + 32$, giving final cost at most $\max\{A_T + A_S, 3A_T + 32\}$. If connectivity finishes first, the symmetric argument gives $\max\{A_T + A_S, 3A_S + 64\}$: shared work is still available whenever separation is incomplete, even if pair work has moved to representative-pair Echo acquisition or fallback. Both bounds are at most $3\max\{A_T, A_S\} + 64$.

If both properties hold, the graph sandwich accepts the true partition. Otherwise an unfinished property supplies an eligible action. Every batch has positive cost, so the bounds above exclude an infinite sequence of acquisitions on coverage. A cap covering the stated sufficient total cannot block a completing batch. Adding pilot cost proves Theorem 2.

A.4 Completion through direct fallback

Optional representative-pair acquisition. Geometric propagation of existing intervals is classical and incurs no oracle calls. With no additional pair acquisition, set $G_{\mathrm{geom}} = 0$. The numerical implementation may additionally acquire Echo measurements on representative pairs, including already resolved pairs, to tighten intervals used by propagation. These measurements are charged to the pair pool and capped by the deterministic allowance $G_{\mathrm{geom}} = 16384n$. Before exceeding this allowance, the controller disables the heuristic and uses unrestricted pending-pair Echo fallback. The following bound covers both cases.

Lemma 2 (Fallback completion)**.** *Fix a pilot of cost B_{pilot} and assume every pair has positive threshold-angle margin $\Delta_{ij} > 0$. Put $A_{\mathrm{all}} = 2\sum_{i<j} Q(\Delta_{ij}/2)$. On simultaneous coverage, Q-Cert correctly accepts within*

$$B_{\mathrm{Q}} \le B_{\mathrm{pilot}} + 3(A_{\mathrm{all}} + G_{\mathrm{geom}}) + 32,$$

provided the cap covers this amount. No clean forest or informative shared basis is required.

Proof. If the confirmed and possible component partitions differ, some possible edge joins distinct confirmed components. Thus the pair pool always has a pending action while certification is incomplete.

Tree and unrestricted fallback actions sample only pending pairs. Once a stream reaches $Q(\Delta_{ij}/2)$, its interval resolves the threshold relation, so it cannot be selected again by either rule. Their combined cost is at most A_{all}. Representative-pair Echo acquisition may continue sampling already resolved pairs, but its entire cost is at most G_{geom}. Hence $P = P_{\mathrm{tree}} + P_{\mathrm{fallback}} + P_{\mathrm{rep}} \le A_{\mathrm{all}} + G_{\mathrm{geom}}$.

Since pair work is available until acceptance, shared sampling requires $S \le P$ before its batch and yields $S \le 2P + 32$ afterward. Adding pilot, pair, and shared costs gives the claimed bound. If the graph sandwich remained unequal, another positive-cost pair action would remain available. This contradicts the finite total-work bound, so acceptance occurs on coverage and is correct. □

For the completion statement in Section 4.1, $F_{ij} \neq \tau$ implies $\Delta_{ij} > 0$ for every pair. There are finitely many pairs and $r_h \to 0$, so each $Q(\Delta_{ij}/2)$ and hence A_{all} is finite. The preceding bound therefore guarantees finite completion when the budget does not bind, without requiring useful pilot or shared evidence.

With a uniform fixed positive pair margin and $M = \mathrm{poly}(n)$, $A_{\mathrm{all}} = O(n^2 \log(n/\delta_v))$. This is a sufficient pair-resolution budget; it is not the realized cost of a competing direct algorithm.

A.5 Frozen-Plan Amortization of Reusable Evidence

The online bound in Theorem 2 uses a worst-case shared margin Γ to control all cross-component relations uniformly. For interpreting partial reuse, it is useful to condition instead on a fixed verification plan and retain the heterogeneous margins.

Keep the forest workload A_T on $\mathcal{T}$ common to both choices. Since $\mathcal{T}$ includes unsuccessful proposals, it may contain cross-component relations; these are already charged to A_T and are excluded from the comparison below. Let $\mathcal{E}_{\mathrm{sh}} \subseteq \mathcal{P}_\tau \setminus \mathcal{T}$ denote the residual cross-component relations assigned to reusable single-state evidence, and let $\mathcal{E}_{\mathrm{dir}} \subseteq \mathcal{P}_\tau \setminus (\mathcal{T} \cup \mathcal{E}_{\mathrm{sh}})$ denote the other residual cross-component relations resolved directly by Echo in the frozen plan. Thus, $\mathcal{T}$, $\mathcal{E}_{\mathrm{sh}}$, and $\mathcal{E}_{\mathrm{dir}}$ are pairwise disjoint, so no pair stream is charged both to the common forest workload and to a residual direct workload. Both choices use the same confidence universe and risk allocation.

For each $e = (i, j) \in \mathcal{E}_{\mathrm{sh}}$, fix one witness basis $b_e \in \mathcal{B}_{ij}$ with positive population root margin

$$\gamma_e := \sqrt{\tau} - h_{ij}^{(b_e)} > 0,$$

where $h_{ij}^{(b)}$ is the population root bound defined in Appendix A.3. Let

$$\mathcal{I}_{\mathrm{sh}} = \{(i, b_e), (j, b_e) : e = (i, j) \in \mathcal{E}_{\mathrm{sh}}\}$$

be the distinct state–basis validation streams used by the shared part of the plan, and define their bottleneck margins by

$$\Gamma_{i,b} := \min_{\substack{e\in\mathcal{E}_{\mathrm{sh}}:\\ i\in e,\ b_e=b}} \gamma_e > 0.$$

On simultaneous coverage, the argument of Appendix A.3 shows that it is sufficient to sample each $(i,b)\in\mathcal{I}_{\mathrm{sh}}$ to $Q(\Gamma_{i,b}/12)$ observations. The corresponding shared workload is therefore

$$A_{\mathrm{sh}} := \sum_{(i,b)\in\mathcal{I}_{\mathrm{sh}}} Q(\Gamma_{i,b}/12).$$

For any cross relation $e=(i,j)$ considered here, write

$$\Delta_e = \left|\arccos\sqrt{F_{ij}} - \arccos\sqrt{\tau}\right|.$$

Resolving the same relations directly by Echo has sufficient workload

$$A_{\mathrm{dir}}(\mathcal{E}_{\mathrm{sh}}) := 2\sum_{e\in\mathcal{E}_{\mathrm{sh}}} Q(\Delta_e/2),$$

where the factor two is the oracle cost of one Echo observation.

The forest workload A_T and the direct workload on $\mathcal{E}_{\mathrm{dir}}$ are common to the two choices; only the workload for the residual relations in $\mathcal{E}_{\mathrm{sh}}$ is replaced. Hence, under this common confidence-radius envelope, assigning $\mathcal{E}_{\mathrm{sh}}$ to reusable evidence has a smaller sufficient workload whenever

$$\sum_{(i,b)\in\mathcal{I}_{\mathrm{sh}}} Q(\Gamma_{i,b}/12) < 2\sum_{e\in\mathcal{E}_{\mathrm{sh}}} Q(\Delta_e/2).$$

Using the envelope in Appendix A.2 and suppressing absolute constants and the logarithmic terms, the leading margin dependence is

$$A_{\mathrm{sh}} = \widetilde{O}\left(\sum_{(i,b)\in\mathcal{I}_{\mathrm{sh}}} \Gamma_{i,b}^{-2}\right), \qquad A_{\mathrm{dir}}(\mathcal{E}_{\mathrm{sh}}) = \widetilde{O}\left(2\sum_{e\in\mathcal{E}_{\mathrm{sh}}} \Delta_e^{-2}\right).$$

Thus, a first-order margin criterion for profitable reuse is

$$\sum_{(i,b)\in\mathcal{I}_{\mathrm{sh}}} \Gamma_{i,b}^{-2} \lesssim 2\sum_{e\in\mathcal{E}_{\mathrm{sh}}} \Delta_e^{-2}.$$

Here $\lesssim$ suppresses the absolute and logarithmic factors inherited from $Q(\cdot)$; this is a margin-level interpretation, not an exact break-even condition.

For further intuition, let $m=|\mathcal{E}_{\mathrm{sh}}|$ and $N_{\mathrm{str}}=|\mathcal{I}_{\mathrm{sh}}|$. If $\Gamma_{i,b}\asymp\Gamma_{\mathrm{sh}}$ and $\Delta_e\asymp\Delta_{\mathrm{dir}}$ over the assigned relations, then

$$\frac{2m}{N_{\mathrm{str}}} \gtrsim \left(\frac{\Delta_{\mathrm{dir}}}{\Gamma_{\mathrm{sh}}}\right)^2.$$

Here $2m/N_{\mathrm{str}}$ is the average number of shared-pair incidences served by one validated state–basis stream. Hence reuse becomes increasingly attractive when each stream supports many cross relations, or when its shared margin is large relative to the direct pairwise gaps.

This comparison concerns a frozen verification plan. It does not solve the online assignment problem, characterize the globally optimal split between shared and direct evidence, or provide an additional worst-case cost guarantee for Algorithm 1. The controller itself does not require knowledge of these margins.

B State Families and Complexity Separations

B.1 The hidden-bridge family

Let $n\geq 4$ be even, $q=\lceil\log_2 n\rceil$, and choose n distinct real Hadamard vectors $|h_i\rangle = H^{\otimes q}|b_i\rangle$. Two distinct group codes h_A, h_B and a common code h_0 may coincide with individual codes, since their

flag sectors are orthogonal. Partition the states into hidden groups A, B of size $n/2$. For positive a, b, c summing to one, the base states are

$$|\psi_i\rangle = \sqrt{a}\,|0\rangle|h_0\rangle + \sqrt{b}\,|1\rangle|h_{g_i}\rangle + \sqrt{c}\,|2\rangle|h_i\rangle, \qquad g_i \in \{A, B\}.$$

The flag uses three levels of a two-qubit register. Distinct states have fidelity $(a+b)^2$ within a group and a^2 across groups. For a candidate bridge $e = (i, j) \in A \times B$, change only the private sector of state j to $\sqrt{c}\,|2\rangle(\sqrt{f}\,|h_i\rangle + \mathrm{i}\sqrt{1-f}\,|h_j\rangle)$, with $0 < f < 1$. This changes only F_{ij}, to $p_1 = (a + c\sqrt{f})^2$. Choose fixed parameters satisfying $a^2 < \tau < \min\{(a+b)^2, p_1\}$. The base has two components, and every alternative has one.

Blind and informative outputs. In the computational basis, every state has probabilities $p(r, x) = w_r/2^q$, with $(w_0, w_1, w_2) = (a, b, c)$. The altered private sector has the same probabilities: the real Hadamard amplitudes and imaginary relative coefficient eliminate its interference term. Thus all single-state computational-basis output laws agree between the base and every alternative.

The measurement $Z_{\text{flag}} \otimes X_{\text{data}}$ instead reveals the codes. A base state returns its common, group, and private symbols with probabilities a, b, c. The changed state returns the common and group symbols and private symbols i, j with probabilities $a, b, cf, c(1-f)$. A hidden common data Hadamard preserves fidelities and exchanges the informative and blind data bases. Both product bases remain in the menu.

Sector-phase extension. Multiply flag sector r of state i by $\exp(\mathrm{i}\phi_i^{(r)})$, where the phases may vary arbitrarily across states in $[-\rho, \rho]$, $0 \le \rho < \pi/4$. Use the same phases in each matched base/bridge alternative. Measuring the flag makes both single-state output laws invariant to these phases. Every within-group or bridge overlap is a sum of two complex terms with arguments in $[-2\rho, 2\rho]$, so its modulus is at least the coefficient sum times $\cos(2\rho)$. Assume $p_{\text{in}} := \cos^2(2\rho)\min\{(a+b)^2, p_1\} > \tau$. All other cross-group fidelities remain a^2; only the designated pair changes between matched instances.

B.2 Near-linear completion with the richer interface

This proves the upper bound in Proposition 2. Set $\delta_p = \delta_v = \delta/2$, $p_* = \min\{b, c, cf\}$, and $N_0 = \lceil p_*^{-1}\log(6n/\delta_p)\rceil$. Collect N_0 pilot shots per state in each of the two bases. The common region is the public flag-zero sector. In each basis, join states whose pilot outcomes share a noncommon symbol, propose a spanning forest of this support graph, and freeze each component's union of observed noncommon outputs as its assigned region. These regions are disjoint: a shared symbol would join the corresponding components.

In the informative basis, require every unchanged state to reveal its group and private symbols. For the altered state, require only its group and bridge symbols. At most $2n$ incidences are required, each of probability at least p_*. The probability of missing any is at most $2n\exp(-N_0 p_*) \le \delta_p$. On the complementary event, the support graph contains exactly the two within-group cliques and, when present, the bridge. Its selected forest is entirely above threshold and spans every true component. In the base case all noncommon support is covered, so $u_i = a$ and $e_i = 0$. In a bridge case there is one component and no separation requirement; observing the altered state's residual private symbol is unnecessary.

Let $t_\tau = \arccos\sqrt{\tau}$, $\Gamma_0 = \sqrt{\tau} - a$, and $\Delta_0 = \min\{\arccos a - t_\tau, t_\tau - \arccos\sqrt{p_{\text{in}}}\}$. These margins are positive. Every pair, including edges proposed by the uninformative basis, has threshold margin at least Δ_0. Theorem 2 applies on successful discovery and simultaneous validation coverage. A public sufficient cap is

$$C_{\text{bridge}}(n, \delta) = 2nN_0 + 3\max\{4(n-1)Q(\Delta_0/2),\ 2nQ(\Gamma_0/12)\} + 64. \tag{11}$$

The two failure probabilities sum to at most δ. The cap depends on public design constants, not the realized groups, codes, bridge, phases, or orientation. For fixed admissible parameters, $q = O(\log n)$ and the confidence universe is polynomial in n, so $C_{\text{bridge}} = O(n\log(n/\delta))$. Enforcing this cap bounds every execution path, including coverage failures, and hence also bounds unconditional expected cost. For $(a, b, c, f, \tau, \rho) = (.8, .1, .1, .875, .75, .05)$, $p_{\text{in}} \simeq .79046$, $\Gamma_0 \simeq .06603$, and $\Delta_0 \simeq .04813$; the sufficient cap at $n = 32, \delta = .05$ is approximately 2.013×10^8 calls.

B.3 Restricted-interface lower bounds

Fix the orientation in which computational-basis outputs are blind and let $\mathcal{H} = A \times B$, $m = |\mathcal{H}| = n^2/4$. The restricted algorithm receives full single-state computational-basis outcomes and scalar Echo or SWAP outcomes on original input pairs. A SWAP test uses two prepared states and returns $\mathrm{Ber}((1 + F_{ij})/2)$ at a cost of two oracle calls; unlike Echo, it does not require inverse preparation access. It may choose actions and its stopping time adaptively, including any classical inference from past outcomes. Its risk requirements apply to the base and every bridge alternative. Here reliability means almost-sure termination, wrong-certification probability at most δ, and abstention probability at most δ on each of these instances. For the lower bounds in Proposition 2, take $0 < \delta \le 1/4$. More generally, the information argument requires $1 - 2\delta > \delta$; the Ω constants are uniform on any range $0 < \delta \le \delta_0 < 1/3$. The stated asymptotic lower bounds are asserted on this small-risk range; the upper-bound construction does not need it.

Excluding every bridge on the base instance. Let E_0 be the event of accepting the base partition. It has probability at least $1 - 2\delta$ under the base and at most δ under each alternative. Only queries to the changed pair have different output laws. Write $p_0 = a^2$ and define $d_E = \mathrm{kl}(p_0, p_1)$, $d_S = \mathrm{kl}((1 + p_0)/2, (1 + p_1)/2)$, and $d_* = \max\{d_E, d_S\}$. The stopped-transcript change-of-measure inequality (Kaufmann et al., 2016) gives, for every $e \in \mathcal{H}$,

$$d_* \, \mathbb{E}_0 N_e \ge \mathrm{kl}(1 - 2\delta, \delta),$$

where N_e counts all scalar pair trials on e, allowing a mixture of Echo and SWAP. For the phased family the changed fidelity belongs to $(\tau, p_1]$. Since $\mathrm{kl}(p_0, x)$ increases for $x > p_0$, the same constants upper-bound its per-query information. Summing over the m candidates and charging two oracle calls per trial yields $\mathbb{E}_0 B_{\mathrm{Q}} \ge 2m \, \mathrm{kl}(1 - 2\delta, \delta)/d_* = \Omega(n^2 \log(1/\delta))$. For infinite expected cost the claim is immediate; otherwise the stopped information inequality follows by truncation and passage to the limit. Extra single-state or within-group queries add cost but no information about the changed pair.

Finding a bridge on average. Couple the base and alternative e using the same internal randomness and outcomes until the first query to e. Let H_e denote querying e before termination. The coupled histories give $\mathbb{P}_0(H_e) = \mathbb{P}_e(H_e)$. On alternative e, correct acceptance has probability at least $1 - 2\delta$. Its part without a query to e has the same probability under the base and is an incorrect acceptance there, so it is at most δ. Therefore $\mathbb{P}_0(H_e) \ge 1 - 3\delta$.

Let K be the number of distinct candidate pairs visited on a base trajectory. Then $\mathbb{E}_0 K \ge (1-3\delta)m$. Conditional on this trajectory, a bridge at the rth visited pair needs at least r pair trials before the first potentially different observation. An unvisited bridge follows the base through termination and needs at least K trials. Averaging the coupled alternative costs gives at least

$$\frac{1}{m}\left[\sum_{r=1}^{K} r + (m - K)K\right] = K - \frac{K(K-1)}{2m} \ge \frac{m+1}{2m} K$$

pair trials. Taking expectations and charging two calls per trial yields $m^{-1} \sum_e \mathbb{E}_e B_{\mathrm{Q}} \ge (1 - 3\delta)(m + 1) = \Omega(n^2)$. This argument also holds for every matched sector-phase configuration. Together with (11), these bounds prove Proposition 2. Reliability on the base is essential: if a bridge were promised, the single-component answer would already be known.

B.4 Bounded-degree graph states

Let $G = ([n], E)$, $n \ge 4$, be any simple graph of maximum degree at most a fixed d. Introduce mutually orthonormal common, edge, and private coordinates $|c\rangle$, $|e\rangle$, and $|v\rangle$. For $w_e \in [w_-, w_+]$, define

$$|\psi_v^0\rangle = \sqrt{a}\,|c\rangle + \sum_{e \ni v} \sqrt{w_e}\,|e\rangle + \sqrt{1 - a - \sum_{e \ni v} w_e}\,|v\rangle. \tag{12}$$

Assume $a > 0$, $0 < w_- \le w_+$, $a + d w_+ < 1$, and $a^2 < \tau < (a + w_-)^2$. The states are normalized. Distinct vertices have inner product $a + w_{\{u,v\}}$ when adjacent and a otherwise, so their

fidelity-threshold graph is exactly G. Pad the coordinates to the public power-of-two dimension $D = 2^{\lceil \log_2(1+n+\lfloor dn/2 \rfloor)\rceil} = O(n)$.

Put $t_\tau = \arccos\sqrt{\tau}$ and $p_* = \min\{a, w_-, 1-a-dw_+\} > 0$, $\Delta_0 = \min\{t_\tau - \arccos(a+w_-), \arccos a - t_\tau\} > 0$, and $\Gamma_0 = \sqrt{\tau} - a > 0$. In the informative coordinate basis each state has at most $d+2$ nonzero outcomes, each of probability at least p_*. Let $\delta_p \in (0,1)$ denote the pilot-recovery failure budget. With a fixed menu of L bases containing this basis, use $N_0^{\mathrm{clean}} = \lceil p_*^{-1} \log[Ln(d+2)/\delta_p] \rceil$ pilot shots per state and basis. Retain outputs seen at least once. For retained support R_v, take $\widehat{\Omega}_0 = \bigcap_v R_v$ and connect vertices whose supports intersect outside $\widehat{\Omega}_0$. Freeze the components' unions of these noncommon supports as exclusive regions.

The probability of missing any nonzero incidence in an informative basis is at most $n(d+2)e^{-N_0^{\mathrm{clean}} p_*} \le \delta_p/L$. On complete support recovery, the intersection identifies exactly the common coordinate: it occurs in all n supports, whereas an edge coordinate occurs in two and a private coordinate in one. The support graph is exactly G, so every selected spanning forest is clean and covers the true components. Its regions satisfy $u_v = a$ and $e_v = 0$. All candidate pairs from any other basis still have threshold margin at least Δ_0. Thus Theorem 2 applies; other bases need not recover the graph or have sparse support.

Proof of Proposition 3. Take $\delta_p = \delta_v = \delta/2$. On successful support recovery, the coordinate-basis forest is clean and spans every true component, while every proposed pair in every allowed basis has margin at least Δ_0. If there is more than one true component, the coordinate basis supplies separation margin Γ_0. Consequently Theorem 2 gives the sufficient cap

$$LnN_0^{\mathrm{clean}} + 3\max\{2L(n-1)Q(\Delta_0/2),\ LnQ(\Gamma_0/12)\} + 64.$$

For a connected graph the separation workload can be set to zero. Pilot failure has probability at most δ_p, and, conditional on any pilot, validation noncoverage has probability at most δ_v. Their union therefore has probability at most δ. Since d, L and the positive margins are fixed and the confidence universe has polynomial size, the cap is $O(n\log(n/\delta))$, uniformly over component sizes and graph topology. This proves the proposition for $n \ge 4$, the range used by the construction and experiments. The finitely many smaller sizes have the same order of cost by direct pair resolution with the uniform margin Δ_0.

B.5 A CONTROLLED PURE-STATE NEIGHBORHOOD

For pure states let $d_{\mathrm{FS}}(\psi,\phi) = \arccos|\langle\psi|\phi\rangle|$. If $d_{\mathrm{FS}}(\psi,\psi^0) \le \epsilon$, then the probabilities p, p' of any fixed projective event under these two states satisfy $|\arccos\sqrt{p} - \arccos\sqrt{p'}| \le \epsilon$. Indeed, Cauchy–Schwarz on the event and its complement bounds $|\langle\psi|\psi^0\rangle|$ by $\sqrt{pp'} + \sqrt{(1-p)(1-p')}$, the cosine of the probability-angle difference. This controls a whole region without a dimension factor.

Proposition 5 (Perturbations of graph states)**.** *Let $\delta_p \in (0,1)$ denote the pilot-recovery failure budget. Suppose every state obeys $d_{\mathrm{FS}}(\psi_v, \psi_v^0) \le \epsilon$ for the reference states in* (12)*. Define $\Delta_\epsilon = \Delta_0 - 2\epsilon$, $p_{\mathrm{lo}} = \sin^2\epsilon$, and $p_{\mathrm{hi}} = [\sqrt{p_*}\cos\epsilon - \sqrt{1-p_*}\sin\epsilon]_+^2$. For the common event, put $u_+ = \cos^2(\max\{0, \arccos\sqrt{a} - \epsilon\})$ and $u_- = \cos^2(\min\{\pi/2, \arccos\sqrt{a} + \epsilon\})$. Let $\Gamma_\epsilon = \sqrt{\tau} - u_+ - 2\sqrt{1-u_-}\sin\epsilon$. Assume $\Delta_\epsilon > 0$, $p_{\mathrm{hi}} > p_{\mathrm{lo}}$, and $\Gamma_\epsilon > 0$. For fixed $s \in (p_{\mathrm{lo}}, p_{\mathrm{hi}})$, retain pilot coordinates with empirical frequency at least s. With $J(s) = \min\{\mathrm{kl}(s, p_{\mathrm{lo}}), \mathrm{kl}(s, p_{\mathrm{hi}})\}$ and $N_0^{\mathrm{pert}} = \lceil \log(2LnD/\delta_p)/J(s) \rceil$ pilot shots per state and basis, an informative basis recovers the reference support graph with probability at least $1-\delta_p$. The actual fidelity graph remains G; the learned forest is clean and spans its components; and its regions satisfy $u_- \le u_v \le u_+$ and $e_v \le \sin^2\epsilon$. All pair margins are at least Δ_ϵ and all true cross-pair shared root margins are at least Γ_ϵ.*

Proof. The Fubini–Study triangle inequality gives $|d_{\mathrm{FS}}(\psi_u,\psi_v) - d_{\mathrm{FS}}(\psi_u^0,\psi_v^0)| \le 2\epsilon$. Thus $\Delta_\epsilon > 0$ preserves every threshold edge and nonedge.

The event-probability inequality bounds every reference-zero coordinate by p_{lo} and every reference-nonzero coordinate, whose probability was at least p_*, from below by p_{hi}. Binomial Chernoff bounds make false inclusion and false exclusion at most $\exp[-N_0^{\mathrm{pert}}\,\mathrm{kl}(s, p_{\mathrm{lo}})]$ and

$\exp[-N_0^{\mathrm{pert}}\,\mathrm{kl}(s, p_{\mathrm{hi}})]$ respectively. A union bound over nD state-coordinate frequencies recovers the exact reference supports at failure probability at most δ_p/L in an informative basis. The stated count conservatively covers up to L informative branches; no independence between coordinate indicators from one shot is needed. The common coordinate and support graph are then recovered as in the unperturbed construction.

For the common event, the same probability-angle inequality gives $u_- \le u_v \le u_+$. The complement of the common and assigned component regions has reference probability zero, so its entire actual mass is at most $\sin^2\epsilon$. Consequently every true cross pair has population root bound at most $u_+ + 2\sqrt{1-u_-}\sin\epsilon = \sqrt{\tau} - \Gamma_\epsilon$. This proves all claims. □

For $0 < p_{\mathrm{lo}} < p_{\mathrm{hi}} < 1$, the threshold $s^* = \log[(1-p_{\mathrm{lo}})/(1-p_{\mathrm{hi}})]/\log[p_{\mathrm{hi}}(1-p_{\mathrm{lo}})/(p_{\mathrm{lo}}(1-p_{\mathrm{hi}}))]$ equalizes the two KL terms. At $\epsilon = 0$ use the observed-once pilot. For $\delta_p + \delta_v = \delta$, a public sufficient online cap is

$$C_{G,\epsilon} = LnN_0^{\mathrm{pert}} + 3\max\{2L(n-1)Q(\Delta_\epsilon/2), LnQ(\Gamma_\epsilon/12)\} + 64. \tag{13}$$

For clean states substitute $N_0^{\mathrm{clean}}, \Delta_0, \Gamma_0$. Successful pilot discovery and validation coverage imply correct acceptance under this cap with probability at least $1-\delta$. For fixed d, L and fixed positive support, threshold, and shared margins, $D = O(n)$ and both pilot and validation counts are logarithmic in n/δ. Hence $C_{G,\epsilon} = O(n\log(n/\delta))$, uniformly over graph topology, component sizes, and admissible weights. The cap again bounds every execution path and expected cost. Smaller empirical caps retain validity but need not satisfy this sufficient completion bound.

For the experimental design $d = 3$, $a = .50$, $w_e \in [.13, .15]$, and $\tau = .36$, one has $p_* = .05$, $\Delta_0 \simeq .0380521$, and $\Gamma_0 = .1$. At $\epsilon = .01$, the margins are $\Delta_\epsilon \simeq .0180521$, $p_{\mathrm{lo}} \simeq 9.9997 \times 10^{-5}$, $p_{\mathrm{hi}} \simeq .0457314$, and $\Gamma_\epsilon \simeq .0757181$. For $n = 96, D = 256, L = 1, \delta_p = .025$, the sufficient pilot counts are $N_0^{\mathrm{clean}} = 198$ and $N_0^{\mathrm{pert}} = 573$. These are state-dependent but fixed pure-state perturbations across repeated preparations. The result concerns this neighborhood of the structured family, rather than arbitrary mixed-state or hardware noise.

B.6 Same-interface task separation

We prove Proposition 1 under the following precise conditions. Fix $\tau = 1/4$, $n \ge 8$ divisible by four, $0 < \eta \le 1/50$, and $0 < \delta \le 1/4$. The gap parameter η is distinct from the perturbation parameter in Appendix B.5. The family $\mathfrak{F}_{n,\eta}$ constructed below has dimension $D = O(n)$; every threshold graph has two components of size $n/2$ and maximum degree at most three. Its connectivity witnesses and reusable separation evidence have constant positive margins.

Both tasks receive complete single-state outcomes in one fixed coordinate basis at cost one, and scalar Echo/SWAP observations on original input pairs at cost two. Adaptive action selection and classical reuse are unrestricted; no additional bases, coherent queries, or exact state descriptions are supplied. Echo returns $\mathrm{Ber}(F_{ij})$ and SWAP returns $\mathrm{Ber}((1+F_{ij})/2)$; observations are fresh and independent conditional on their selected actions. Q-Cert uses the empirical maximum-weight spanning-tree pilot below and the unchanged validation controller of Algorithm 1, with Echo for pair validation. It need not know η and has a deterministic $O(n\log(n/\delta))$ cap independent of η; its incorrect-certification guarantee remains valid for arbitrary pure-state inputs.

The full-graph certifier must terminate almost surely and satisfy, on every $\Psi \in \mathfrak{F}_{n,\eta}$,

$$\Pr_{\Psi}(A = 1, \widehat{G} \ne G_\tau(\Psi)) \le \delta, \qquad \Pr_{\Psi}(A = 0) \le \delta. \tag{14}$$

Infinite expected cost satisfies the lower bound immediately. The lower bound holds even if η and the strong-edge and matching templates are disclosed. It compares recovery targets, not optimality among partition-certification algorithms; the quadratic specialization uses shrinking redundant-edge gaps, not fixed gaps for all pairs.

A legal pure-state construction Put $k = n/2$. Partition the vertices into two sets of k vertices. In each set, make a strong cycle C_k and an antipodal perfect matching: vertex r is matched to $r + k/2$ for $1 \le r \le k/2$. Let $\mathcal{S}$ be the union of the cycle edges and $\mathcal{H}$ the union of the matching edges. Then $|\mathcal{S}| = n$, $|\mathcal{H}| = k$, and $\mathcal{S} \cap \mathcal{H} = \varnothing$. Vertices and coordinates can be arbitrarily permuted; the pilot is not given the permutation or the component partition.

Introduce mutually orthonormal coordinates for one common symbol c, every edge of $\mathcal{S} \cup \mathcal{H}$, and every private vertex symbol. Set

$$a = \frac{2}{5}, \qquad s = \frac{1}{5}, \qquad w = \frac{1}{8}, \qquad r = \frac{3}{40}, \qquad \tau = \frac{1}{4}.$$

For the unique $h(v) \in \mathcal{H}$ incident to v, define

$$|\psi_v\rangle = \sqrt{a}\,|c\rangle + \sqrt{s} \sum_{e \in \mathcal{S} : e \ni v} |e\rangle + \sqrt{w}\, e^{i\theta_v} |h(v)\rangle + \sqrt{r}\,|v\rangle. \tag{15}$$

For each oriented matching edge $e = (u, v)$, set $\theta_u = 0$, $\theta_v = \phi_{\sigma_e}$, where

$$\cos \phi_{\sigma_e} = \frac{\tau + \sigma_e \eta - a^2 - w^2}{2aw} = \frac{119}{160} + 10\sigma_e \eta. \tag{16}$$

The right side belongs to $[87/160, 151/160] \subset (-1, 1)$, so these are valid phases. Since every vertex has two cycle incidences and one matching incidence, $a + 2s + w + r = 1$. The unpadded dimension is $1 + n + k + n = 1 + 5n/2$; padding to a power of two preserves $D = O(n)$.

Two distinct vertices share only the common coordinate and, if applicable, their unique edge coordinate. Hence

$$F_{ij} = \begin{cases} 9/25, & (i,j) \in \mathcal{S}, \\ 1/4 + \sigma_{ij}\eta, & (i,j) \in \mathcal{H}, \\ 4/25, & \text{otherwise.} \end{cases} \tag{17}$$

The two cycles always span the components. Changing matching edges never changes the partition, but changes the exact graph whenever $\eta > 0$.

Every coordinate measurement has the same five nonzero probability masses $(2/5, 1/5, 1/5, 1/8, 3/40)$, at its corresponding coordinates, independently of all phase bits. Flipping one bit changes only the fidelity of that matching edge, and hence only its scalar pair-observation laws. This proves the required quantum realizability and observational locality; an arbitrary similarity matrix is not being postulated.

A sample-only pilot and the online upper bound Set $\delta_p = \delta_v = \delta/2$ and

$$N_0 = \left\lceil 800 \log \frac{20n}{\delta} \right\rceil. \tag{18}$$

Measure each state N_0 times in the coordinate basis and form its empirical distribution $\widehat{p}_i$ and observed support R_i. Let $\widehat{\Omega}_0 = \bigcap_i R_i$ and form a support graph by joining i, j when $(R_i \cap R_j) \setminus \widehat{\Omega}_0 \neq \varnothing$. Its components are the candidate blocks. Assign edge weight

$$\widehat{s}_{ij} = \sum_{x \in (R_i \cap R_j) \setminus \widehat{\Omega}_0} \min\{\widehat{p}_i(x), \widehat{p}_j(x)\}. \tag{19}$$

Select a maximum-weight spanning tree in every candidate block, using a deterministic rule for ties. The assigned region of a block is the union of its observed noncommon supports. Distinct assigned regions are disjoint by construction. Freeze everything before starting independent validation.

An arbitrary spanning tree is not sufficient: it can select a near-threshold matching edge. The maximum-weight rule is the only additional pilot specification needed; it does not change Algorithm 1's post-pilot controller and uses no true fidelities or phase bits.

For each of the $5n$ nonzero state-coordinate incidences, Hoeffding's inequality gives

$$\Pr\big(|\widehat{p}_i(x) - p_i(x)| > 1/40\big) \le 2e^{-N_0/800}.$$

The probability that any of these bounds fails is at most $10ne^{-N_0/800} \le \delta_p$. True zero-probability coordinates are never observed. On the complementary event, every nonzero coordinate is observed because its empirical mass is at least $3/40 - 1/40 = 1/20$. Thus R_i is exact, the common intersection is exactly $\{c\}$, and the support graph is $\mathcal{S} \cup \mathcal{H}$. This includes all matching candidates regardless of their signs, so recovering this support graph does not recover the threshold graph.

Every cycle edge has weight at least $1/5 - 1/40 = 7/40$, while every matching edge has weight at most $1/8 + 1/40 = 3/20$. Since $7/40 > 3/20$ and the cycle edges connect each block, every maximum-weight spanning tree uses cycle edges only: deleting a matching edge from a proposed tree creates a cut crossed by a heavier cycle edge, so exchanging these edges strictly increases the weight. The clean forest has $n - 2$ edges, all with fixed angular margin

$$\Delta_s = \arccos(1/2) - \arccos(3/5) > 0.$$

The frozen regions have common mass $u_i = 2/5$ and leakage $e_i = 0$, giving root separation margin $\Gamma = \sqrt{\tau} - 2/5 = 1/10$.

Let Q be the dyadic angular-radius envelope of Appendix A.2, computed with validation risk δ_v and a parameter universe including all pairs and both region statistics for every state. Theorem 2 applies with

$$A_T = 2(n-2)Q(\Delta_s/2), \qquad A_S = nQ(1/120).$$

Consequently a sufficient deterministic cap is

$$B_\star = nN_0 + 3\max\{2(n-2)Q(\Delta_s/2),\ nQ(1/120)\} + 64. \tag{20}$$

Both arguments of Q are fixed positive constants, and the confidence universe is polynomial in n. Therefore $B_\star \le Cn\log(n/\delta)$ with no dependence on η. Pilot success and joint validation coverage imply correct acceptance, with joint failure probability at most $\delta_p + \delta_v = \delta$. Enforcing this cap also bounds all off-coverage executions and expected cost.

For arbitrary pure-state inputs, the pilot still returns disjoint frozen regions and a proposal forest. Independent validation, covered geometric propagation, and the contradictory-interval abstention convention of Appendix A.1 give incorrect-certification probability at most δ_v. Thus the upper bound has structured efficiency but universal safety. The same upper-bound argument remains valid at $\eta = 0$, even though the matching edges then lie exactly at the threshold.

The exact-graph lower bound: per-query information. For Bernoulli laws, $\mathrm{kl}(p,q) \le (p-q)^2/[q(1-q)]$. Echo means for a matching edge belong to $[.23, .27]$ and differ by 2η across a bit flip. SWAP means belong to $[.615, .635]$ and differ by η. In either direction and for either action,

$$d_\eta \le 24\eta^2. \tag{21}$$

All other actions have identical laws across the bit flip. Adaptive choices and classical inference introduce no additional KL information beyond the sampled transcript.

Sharing the error budget across distinct wrong outputs Fix the vertex/coordinate permutation. Let G_σ be the graph indexed by $\sigma \in \{-1,+1\}^k$, and let $\sigma^{(j)}$ flip bit j. Define $E_\rho = \{A = 1, \widehat{G} = G_\rho\}$ and $q_{\sigma,j} = \Pr_\sigma(E_{\sigma^{(j)}})$. By (14),

$$\Pr_\rho(E_\rho) \ge 1 - 2\delta, \qquad \sum_{j=1}^{k} q_{\sigma,j} \le \delta. \tag{22}$$

The second inequality holds because the k accepted neighboring graphs are distinct wrong outputs on σ.

Let N_j count all Echo and SWAP trials of matching edge j. The stopped-transcript change-of-measure inequality (Kaufmann et al., 2016, Lemma 1), applied in the direction $\sigma^{(j)} \to \sigma$ with event $E_{\sigma^{(j)}}$, yields

$$24\eta^2 \mathbb{E}_{\sigma^{(j)}} N_j \ge \mathrm{kl}\big(\Pr_{\sigma^{(j)}}(E_{\sigma^{(j)}}), q_{\sigma,j}\big) \ge \mathrm{kl}(1-2\delta, q_{\sigma,j}). \tag{23}$$

Here $q_{\sigma,j} \le \delta < 1 - 2\delta$; the monotonicity needed in the last step holds for $0 < \delta \le 1/4$. Zero event probabilities are handled by extended-valued KL. If any relevant expected cost is infinite, the result is immediate.

For $b = 1 - 2\delta$, convexity of $q \mapsto \mathrm{kl}(b,q)$ and its monotonicity on $q < b$ give

$$\sum_{j=1}^{k} \mathrm{kl}(b, q_{\sigma,j}) \ge k\,\mathrm{kl}\left(b, \frac{1}{k}\sum_j q_{\sigma,j}\right) \ge k\,\mathrm{kl}(b, \delta/k).$$

Sum (23) over all σ and all j. For each j, the map $\sigma \mapsto \sigma^{(j)}$ is a bijection. Hence

$$24\eta^2\, 2^{-k} \sum_{\sigma} \mathbb{E}_\sigma \sum_j N_j \geq k\, \mathrm{kl}(1-2\delta, \delta/k).$$

Every counted pair trial costs two calls, and $2k = n$, so

$$2^{-k} \sum_{\sigma} \mathbb{E}_\sigma B_{\mathrm{graph}} \geq \frac{n}{24\eta^2}\, \mathrm{kl}\left(1-2\delta, \frac{2\delta}{n}\right) \geq \frac{n}{120\eta^2} \log\frac{n}{\delta}. \tag{24}$$

The second inequality follows from the entropy bound below; the average lower bound also bounds the worst-case expected cost.

To obtain the simpler rate, put $R = k/\delta \geq 16$ and write h for binary entropy. Then

$$\mathrm{kl}(1-2\delta, \delta/k) \geq (1-2\delta)\log(k/\delta) - h(1-2\delta) \geq \tfrac{1}{2}\log R - \log 2 \geq \tfrac{1}{4}\log R.$$

Also $\log(n/(2\delta)) \geq (4/5)\log(n/\delta)$ because $n/\delta \geq 32$. Substituting $k = n/2$ gives the last inequality in (24) and finishes Proposition 1.

Why the logarithm is stronger than a single-base lower bound. Comparing one base input to its neighbors separately gives only $\Omega(n\eta^{-2}\log(1/\delta))$. The strengthened bound uses all phase configurations and the fact that the probabilities of accepting their distinct wrong neighboring graphs must *sum* to at most δ. It is an average, hence worst-case, lower bound. It is not asserted pointwise with a $\log n$ factor. A separate direct comparison does give the weaker pointwise bound $\mathbb{E}_\sigma B_{\mathrm{graph}} \geq [n/(24\eta^2)]\,\mathrm{kl}(1-2\delta, \delta)$ for every σ.

Matching graph recovery within the promised family If η is supplied, the graph rate is tight on this family, not merely a loose lower bound. Use the same pilot with failure probability at most $\delta/2$ to recover the support graph and distinguish cycle edges from matching edges by their empirical weights. On pilot success the cycle and non-support relations are known from the family promise. Test every matching edge with

$$N_\eta = \left\lceil \frac{1}{2\eta^2} \log\frac{2k}{\delta} \right\rceil$$

Echo trials and classify its mean relative to τ. A union bound and Hoeffding's inequality make the additional probability of any wrong bit at most $\delta/2$. To bound cost also on pilot failure, threshold empirical edge weights at $13/80$, which lies strictly between $3/20$ and $7/40$. If this does not produce two strong cycles of size k and a matching of k candidate edges, stop and output an arbitrary graph; otherwise test exactly those k edges. Such an early output is confined to pilot failure. The algorithm always terminates without abstention, its wrong-output probability on the promised family is at most δ, and its deterministic cost is $O(n\log(n/\delta) + n\eta^{-2}\log(n/\delta))$. Thus, defining minimax over algorithms reliable on this family,

$$\inf_{\mathcal{A}} \sup_{\Psi \in \mathfrak{F}_{n,\eta}} \mathbb{E}_\Psi B_{\mathrm{graph}} = \Theta(n\eta^{-2}\log(n/\delta)).$$

This upper bound uses the family promise and the supplied gap. It is not a new claim of universal safe graph certification or gap-adaptive optimality. The lower bound is stronger in that it applies even with these advantages supplied to the graph algorithm.

Universal certification versus a family promise. In the all-negative matching configuration, multiply the two incident cycle-coordinate amplitudes of one vertex by -1. Those two fidelities change from 0.36 to 0.04, whereas all coordinate probabilities and all other pair fidelities are unchanged. The vertex becomes isolated. This alternative lies outside the efficient family but within the universal safety domain of Theorem 1. Thus trusting the support partition without validation would not give the required universal certificate.

C Experimental protocols and additional controls

Observation model and accounting. All measurements are sampled from ideal Born probabilities of fixed pure states. The controller receives sampled outcomes; exact fidelities and target partitions

Table 2: Experimental design. A structure includes both hidden orientations; a chain block varies ordering and measurement randomness on a fixed graph. The two size sweeps use fresh blocks, separate from mechanism confirmation.

Question	Family and size	Replication	Main comparison
Discovery	Hidden bridge, $n = 32$	12 structures	Two bases vs. one basis
Exclusion reuse	Bounded degree, $n = 96$; four chains, $n = 64$	20 graph blocks; 12 chain blocks	Q-Cert vs. full-risk direct
Scaling	Hidden bridge, $n = 16$–128; bounded degree, $n =$ 24–192	12 fresh blocks per size and family	Same comparisons across four sizes

Table 3: Experimental design. Each block is the resampling unit; methods, orientations, and perturbations within a block are paired. The scaling blocks are fresh at every size and independent of the main confirmation.

Study	Blocks	Runs	Conditions within a block
Hidden bridge	12	192	2 graphs $\times$ 2 orientations $\times$ 4 methods
Bounded-degree confirmation	20	120	2 perturbations $\times$ 3 methods
Bounded-degree scaling	12×4	96	2 methods at each size

are used only by the simulator and evaluator. A single-state measurement costs one oracle call, and an Echo observation costs two. The ledger includes the pilot, region validation, support-tree tests, representative-pair Echo acquisitions, and pending-pair Echo fallback. Geometric propagation uses existing intervals and adds no oracle cost. The endpoint assigns actual cost to correct acceptance and the common cap to every other terminal outcome. All runs in Table 3 correctly accept below their caps, so no run is censored and restricted costs equal actual costs. Zero observed errors are an empirical statement, separate from the algorithm's probabilistic safety guarantee.

Saving estimator. We report $1 - \overline{B}_{\text{Q-Cert}}/\overline{B}_{\text{comp}}$, where $\overline{B}$ is the mean oracle cost over matched structures or graph blocks. For hidden bridges, the two orientations are first averaged within each structure. This is a ratio of mean costs, not a mean of per-instance percentage savings.

Pairing and risk. Methods in a block share the state instance and matching prefixes of each observation stream; each method pays for the observations it requests. Coupling reduces variability in paired comparisons without transferring measurements between methods. The advertised error ceiling is $\delta = .05$. Q-Cert reserves $.025$ for pilot completion analysis and uses $.0125$ each for pair and shared confidence statements. Matched direct retains the $.0125$ pair budget while disabling shared validation; full-risk direct assigns $.05$ to pair evidence. These are distinct comparisons: matched direct isolates validation at fixed pair risk, while full-risk direct uses the available risk budget more efficiently. The main same-interface results use the latter.

Hidden-bridge discovery. The family parameters are $(a, b, c, f, \rho, \tau) = (.8, .1, .1, .875, .05, .75)$, where ρ is the sector-phase radius. At $n = 32$, each of 12 structures is evaluated with and without a bridge and under two hidden orientations. The methods are Q-Cert ZX, Q-Cert Z, Direct Z with geometry, and full-risk Direct ZX. The pilot uses 103 shots per state per permitted basis, and the common cap is 201,333,248 oracle calls. ZX denotes Z on the flag register with either Z or X on the data register. The primary discovery endpoint compares Q-Cert ZX with Q-Cert Z on bridge-present instances, first averaging the two orientations within each structure. The master seed is 2026091851; the paired percentile bootstrap uses 10,000 draws with seed 2026091880.

Table 4 separates access from confidence allocation. Direct ZX receives the same richer pilot and costs less than Q-Cert ZX in this batch. On bridge instances neither method spends on shared validation; their cost difference reflects the pair risk allocation. Consequently, the discovery comparison supports the value of the extra measurement basis, not dominance over methods already given that basis.

Table 4: Complete hidden-bridge confirmation: mean oracle calls over 12 structures and two orientations per entry. All 192 runs correctly accept. Methods with ZX access receive the same richer pilot.

Method	Bridge present	Two components
Q-Cert ZX	130,757.3	584,794.7
Q-Cert Z	393,802.7	624,752.0
Direct Z + geometry	570,456.0	582,016.0
Direct ZX, full risk	113,968.0	396,346.7

State generation and observation streams. The hidden-bridge experiments use the states, sector phases, and two orientations in Appendix B.1; the bounded-degree experiments use (12). Coordinate-basis outcomes are sampled from $|\langle x|\psi_i\rangle|^2$, and Echo outcomes from $\mathrm{Ber}(|\langle\psi_i|\psi_j\rangle|^2)$. The simulator computes these probabilities, but the controller receives only the sampled outcomes. For a reference state $|\psi_i^0\rangle$, the controlled perturbation is

$$|\psi_i^\epsilon\rangle = \cos\epsilon\,|\psi_i^0\rangle + \sin\epsilon\,|\eta_i\rangle, \qquad |\eta_i\rangle = \frac{(I - |\psi_i^0\rangle\langle\psi_i^0|)g_i}{\|(I - |\psi_i^0\rangle\langle\psi_i^0|)g_i\|}.$$

The real and imaginary entries of g_i are independent standard Gaussian variables, drawn separately for each state and frozen for the run. Thus $d_{\mathrm{FS}}(\psi_i^\epsilon, \psi_i^0) = \epsilon$. Methods share the same perturbed states within a comparison. Pilot and validation have separate random streams; matching action prefixes are coupled across methods. Each requested single-state outcome is charged once even when it updates both common-mass and leakage counts.

Heterogeneous bounded-degree confirmation. This design fixes $n = 96$, maximum degree 3, common probability mass $a = .50$, edge weights sampled independently from $[.13, .15]$, and $\tau = .36$. There are five independent blocks in each of four topology strata: trees, cycles with chords, cactus graphs, and ladder graphs. The number of components is drawn from $\{3, 4, 5, 6\}$; component sizes are unequal with minimum size 8. Graph topology, edge weights, state ordering, and coordinate ordering vary across blocks. Every method receives the same informative coordinate basis.

Each block is evaluated at perturbation angles $\epsilon = 0$ and $.01$ with Q-Cert, matched direct, and full-risk direct. A perturbed state has one frozen orthogonal perturbation direction; repeated preparations produce that same pure state. This tests a controlled pure-state neighborhood, not time-varying noise. Pilot counts are respectively 198 and 573 shots per state, and all runs share a 32-million-call cap. The sole primary endpoint is the unperturbed ratio-of-means saving against full-risk direct. Perturbed and matched-risk comparisons are secondary. The master seed is 273700000. Bootstrap intervals use 10,000 paired stratified resamples with seed 273799999, drawing five blocks with replacement inside each topology stratum and retaining all methods and perturbations together.

Independent scaling confirmation. The four sizes are $n \in \{24, 48, 96, 192\}$, each with 12 new graph blocks (three per topology stratum), Q-Cert and full-risk direct, and $\epsilon = 0$. No block is reused from the 20-block confirmation. The state-family parameters and maximum degree remain fixed. Component counts $3, 4, 5, 6$ are balanced within each size, with unequal component sizes of at least 3. Pilot counts are $170, 184, 198, 212$ per state, and common caps are $2, 8, 32, 128$ million calls. These generous empirical caps are not the conservative sufficient budgets of the completion theorems. The master seed is 2026092011, and the bootstrap seed is 2026092012. The analysis draws 10,000 paired stratified resamples separately at each size, preserving the four strata. Intervals are pointwise, without a simultaneous coverage claim across sizes. Twelve blocks per size, rather than 96 method runs, determine the effective replication.

Interpretation and reproducibility. The designs, source hashes, seed rules, and endpoints were frozen before their measurement outcomes; these manifests are not independently timestamped preregistrations. No exploratory outcomes are pooled with the reported confirmation batches. The amortization comparison is an identity between aggregate costs of two adaptive policies, not a causal estimate of the marginal effect of one shared shot. Both the upper-bound interpretation and these experiments concern structured inputs with informative reusable evidence.

The submitted code package contains the experiment implementations, frozen configurations, seeds, source-hash checks, and aggregation scripts, but no saved measurement records, compact result snapshots, or historical action trajectories. After rerunning the corresponding experiments, `paper/scripts/prepare_results.py --refresh-source` verifies the regenerated records and reconstructs the core statistical summaries and plotting data. This aggregation acquires no new measurements and is distinct from controller replay.

D SUPPLEMENTARY MECHANISM AND SCALING RESULTS

This appendix supplements the 408-run core experiments with the 576-run bridge sweep, 72-run chain confirmation, and separately identified exploratory controls. Costs include pilots; savings are ratios of means. Supplementary comparisons are descriptive unless an interval is given.

D.1 CONNECTION DISCOVERY CONTROLS

Aggregate discovery result. On the bridge-present confirmation, mean costs are 130.757K for Q-Cert ZX, 393.803K for Q-Cert Z, 570.456K for Direct Z with geometry, and 113.968K for full-risk Direct ZX. The saving of ZX over Z is 66.80%, with a paired-structure 95% bootstrap interval of [65.92%, 67.63%]; all 12 structure-level comparisons favor ZX. Each structure-level cost averages its two orientations. Table 5 reports the orientation breakdown, distinct tested pairs, and success at a budget of 250K calls. These summarize the same confirmation batch.

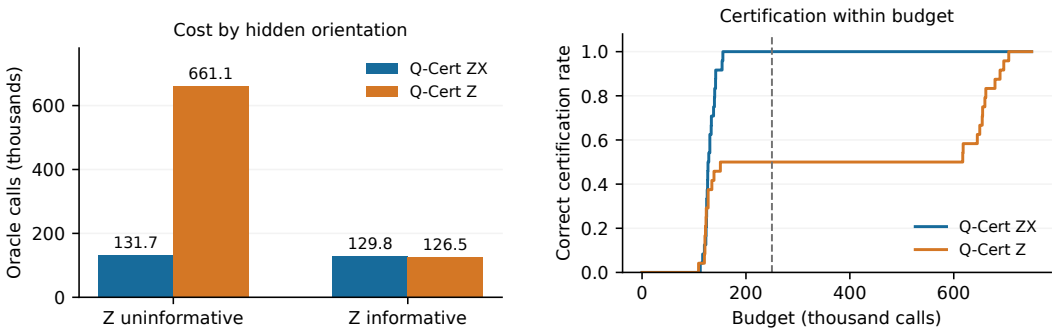

Figure 4: Connection discovery on the held-out bridge-present confirmation. Left: mean cost over 12 structures per hidden orientation. Right: the fraction of 24 runs correctly certified within each budget, with a dashed line at 250K calls. Both panels summarize the same completed runs.

Orientation, tested pairs, and finite budgets. Each orientation in Table 5 contains the same 12 structures. Primary inference averages orientations within structures. The extra ZX pilot costs 3,296 calls (32×103); all bridge-present runs correctly certify without shared validation.

Table 5: Bridge-present confirmation at $n = 32$. Costs are in thousands of oracle calls. The last two columns average both orientations (24 runs per method). The budget endpoint is computed from the recorded correct stopping cost, not from a new set of runs.

Method	Z uninformative	Z informative	Tested pairs	Success at 250K
Q-Cert ZX	131.744	129.771	39.125	100%
Q-Cert Z	661.131	126.475	263.500	50%
Direct Z + geometry	569.173	571.739	496.000	0%
Full-risk Direct ZX	114.037	113.899	39.125	100%

Direct ZX uses .05 rather than .0125 for pairwise confidence statements, explaining its lower cost despite testing the same number of pairs. Bridge-absent confirmation results remain in Table 4.

Support-tree ablation in an earlier implementation. The coarse and support-tree ZX policies share pilot and confidence allocation. Table 6 retains both connectivity conditions; Fig. 5 shows

the bridge-present reduction in pair count and cost. This earlier-controller study is exploratory: two structures at $n = 8, 16$ and one at $n = 32, 64$, each under two orientations. Its results are not pooled with current-controller runs.

Table 6: Historical matched-ZX support-tree ablation. Costs are in thousands of calls. Pair counts refer to bridge-present instances. Each entry averages two orientations per structure.

	Bridge present: cost		Bridge present: pairs		Bridge absent: cost	
n	Coarse	Tree	Coarse	Tree	Coarse	Tree
8	41.312	27.776	28	9	78.064	75.056
16	175.824	79.008	120	19	305.248	237.936
32	662.208	169.664	496	37	1287.872	889.536
64	2720.096	370.816	2016	79	4315.744	3252.992

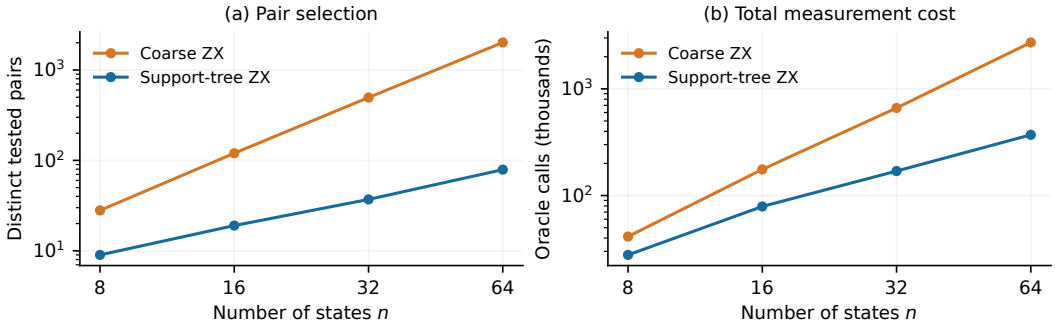

Figure 5: Historical matched-ZX ablation on bridge-present instances. Support-tree prioritization reduces both tested pairs and total cost. There are two structures at $n = 8, 16$ and one at $n = 32, 64$, each with two orientations; this exploratory study uses an earlier controller.

D.2 Reusable exclusion controls

Replication across structures and risk allocations. Table 7 adds the matched-risk comparator to the exclusion results in Section 5.1. Matched direct retains Q-Cert's pairwise risk allocation; full-risk direct assigns .05 to pair tests. Both disable shared validation while retaining the same access, pilot streams, and candidate construction. Every method pays for its own measurements.

Table 7: Same-interface exclusion comparisons. Costs are in millions of calls. Each bounded-degree row averages 20 graph blocks; each chain row averages 12 randomization blocks. All 192 method runs correctly certify. Savings use the full-risk direct comparator.

Family	ϵ	Q-Cert	Matched direct	Full-risk direct	Saving
Bounded degree, $n = 96$	.00	2.589	6.059	5.453	52.53%
Bounded degree, $n = 96$	.01	2.933	6.093	5.488	46.56%
Four chains, $n = 64$	.00	1.810	2.409	2.218	18.38%
Four chains, $n = 64$	.02	2.080	2.407	2.222	6.37%

Table 8: Bounded-degree confirmation by topology. Each stratum contains five blocks; all five favor Q-Cert at both perturbation levels. Savings compare with full-risk direct and are descriptive.

Topology	Saving, $\epsilon = 0$	Saving, $\epsilon = .01$
Tree	52.27%	46.99%
Cycle with chords	50.25%	44.09%
Cactus	54.28%	48.33%
Ladder	53.11%	46.62%

The paired stratified-bootstrap 95% intervals for bounded-degree savings against full-risk direct are $[50.51\%, 54.21\%]$ at $\epsilon = 0$ and $[44.31\%, 48.45\%]$ at $\epsilon = .01$. Chain savings are descriptive.

Against matched direct, bounded-degree savings are 57.27% and 51.86% at $\epsilon = 0, .01$. Topology strata in Table 8 partition the same 20-block confirmation; they are not separate studies.

Where the calls are saved. At $\epsilon = 0$, bounded-degree Q-Cert spends 0.019M calls on its pilot, 1.287M on single-state validation, and 1.283M on pairs. Full-risk direct has the same pilot and spends 5.434M on pairs. Thus 1.287M validation calls replace 4.151M pair calls, a ratio of 3.23. At $\epsilon = .01$, the corresponding ratio is 2.77. These are aggregate accounting ratios between adaptive policies, not marginal causal effects of individual shots.

Sparse-chain protocol. For chain r and position $t \in \{1, \ldots, 16\}$, the reference state is

$$|\psi^0_{r,t}\rangle = \sqrt{a}\,|c\rangle + \sqrt{(1-a)/2}\,(|r,t\rangle + |r,t+1\rangle),$$

where the common coordinate and all chain coordinates are orthogonal. Adjacent states have fidelity $((1+a)/2)^2 = .81$, and all other distinct pairs have fidelity $a^2 = .64$ when $a = .8$, giving four disjoint chains at $\tau = .75$. States and coordinates are independently permuted and padded to dimension $D = 256$. Each state has three nonzero outcomes; the observed-once pilot uses minimum nonzero probability .1. At $\epsilon = .02$, the frequency filter uses threshold .01677197013, retaining coordinates seen at least six times in 304 shots. Perturbations are generated by the orthogonal-direction construction in Appendix C. Internal chain vertices have no private-coordinate mass, so this experiment is a boundary case rather than an instance of the strict positive-private-mass assumption in Proposition 3.

The separate confirmation uses four disjoint chains of 16 states, common mass $a = .8$, threshold $\tau = .75$, and $\delta = .05$. There are 12 paired randomization blocks, two perturbation angles $\epsilon \in \{0, .02\}$, and three methods, giving 72 runs. The unperturbed graph is fixed; state and coordinate orderings and measurement randomness vary across blocks. Structure seeds are 271001000–271001011 and measurement seeds are 271002000–271002011. The controller prioritizes its candidate forest. Pilot counts are 90 and 304 per state, respectively, and the common cap is 16 million calls. As in the bounded-degree study, the perturbation defines a fixed pure state, not time-varying hardware noise. The unperturbed full-risk comparison is the primary endpoint; the perturbed comparison is secondary. Savings against matched direct are 24.85% and 13.59%, respectively.

Additional fixed-basis evidence. An exploratory $n = 128$ study fixes the useful Z basis on 12 disconnected hidden-bridge structures. Q-Cert costs 3.915M calls versus 6.924M for matched direct, a 43.45% reduction. Both pay 0.015M pilot calls; Q-Cert spends 1.952M on validation and 1.948M on pairs, while direct spends 6.909M on pairs. The avoided-pair-cost to validation-cost ratio is 2.54. All 24 runs correctly certify. This supports exclusion reuse even when basis discovery is removed, but there is no full-risk direct arm. The experiment manifest labels the study exploratory; a later confirmation note does not change that status. Complete action traces and historical source hashes are unavailable for this batch, so we use it only as supplementary evidence.

D.3 Complete scaling results and budget curves

Hidden-bridge size sweep. The current-controller sweep uses $n \in \{16, 32, 64, 128\}$, 12 new structures per size, two orientations, bridge-present and bridge-absent conditions, and three methods, giving 576 runs. Family parameters are the same as in Appendix C. The master seed is 2026091852; the confirmation at $n = 32$ uses a different seed and separate structures. Table 9 retains both connectivity conditions. All 576 runs correctly certify. On bridge-present instances, the saving of ZX over Z increases from 49.22% at $n = 16$ to 89.30% at $n = 128$ (Fig. 6). These instances need no cross-component validation, so the trend reflects connection discovery. When the bridge is absent, ZX is more expensive at $n = 16$, while savings against Q-Cert Z increase to 47.97% at $n = 128$. Thus the largest bridge-present gains should not be generalized to every size or connectivity condition.

Scope and reproducibility. The claims concern oracle cost on structured pure-state families. After rerunning the appendix experiments, `paper/scripts/prepare_extensions.`

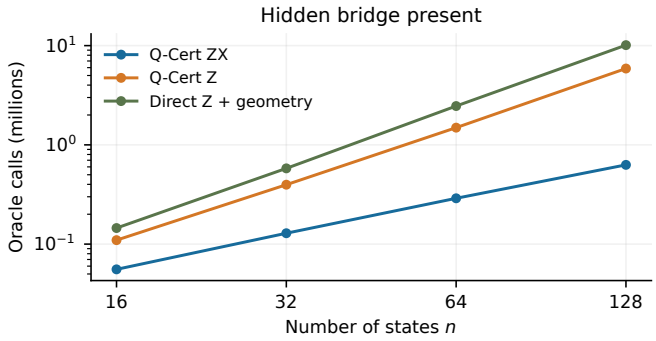

Figure 6: Supplementary hidden-bridge scaling with the bridge present. Each mean averages 12 fresh structures and both hidden orientations at each size. Costs include the pilot; no shared validation is needed.

Table 9: Complete hidden-bridge sweep. Mean costs are in millions of calls, with 24 runs (12 structures, two orientations) per cell. Savings compare Q-Cert ZX with Q-Cert Z.

Condition	n	Q-Cert ZX	Q-Cert Z	Direct Z + geometry	Saving
Bridge present	16	.0557	.1096	.1452	49.22%
	32	.1287	.3968	.5797	67.57%
	64	.2895	1.4908	2.4612	80.58%
	128	.6287	5.8763	10.1125	89.30%
Bridge absent	16	.2248	.1892	.1393	−18.84%
	32	.5897	.6242	.5727	5.54%
	64	1.6764	2.1840	2.4648	23.24%
	128	3.9253	7.5437	10.1630	47.97%

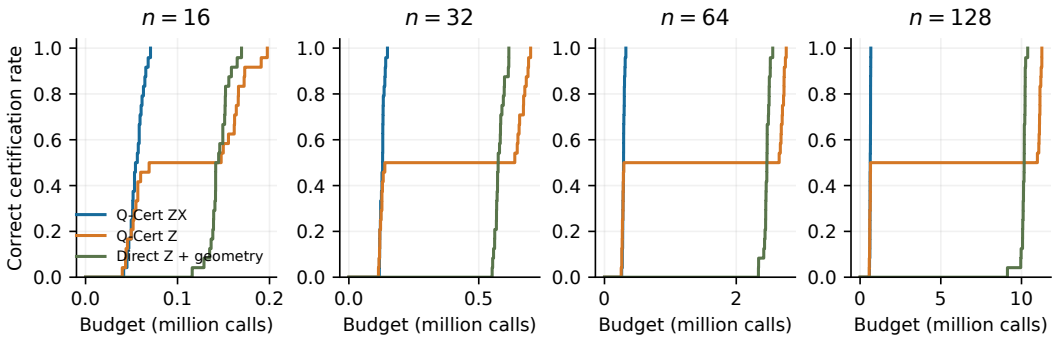

Figure 7: Correct certification rate versus available budget on the bridge-present size sweep. Each curve is the empirical distribution of correct stopping costs over 24 runs, retaining both orientations of each of 12 structures. Curves summarize the same completed runs, not independent reruns at each budget; the structure is the replication unit.

py collects the supplementary summaries and per-run hidden-bridge costs into paper/data/experiment_extensions.json. The corresponding data-driven figures can then be regenerated with paper/scripts/draw_experiment_extensions.py, using these collected results and the regenerated bounded-degree scaling data. The experiment outputs are generated under qcert_online_rerun and quantum_task_recovery; they are not included in the submitted code package. Aggregation and plotting acquire no new measurements and do not replay controllers. The historical reconstruction limitation for the exploratory fixed-basis batch stated above still applies.

Same-interface bounded-degree sweep. Table 10 gives the numerical results behind Fig. 3(a,b). Protocols are in Appendix C. Intervals use 10,000 paired stratified resamples with pointwise coverage. The 12 blocks at $n = 96$ are independent of the 20-block confirmation; all 96 runs correctly certify.

Table 10: Bounded-degree scaling with full-risk direct. Costs are in millions of calls. R is avoided pair cost divided by single-state validation cost; $R > 1$ means that the avoided pair work exceeds the added validation work. Shared pilot costs cancel in this accounting comparison.

n	Q-Cert	Direct	Saving [pointwise 95% interval]	Validation	R
24	.404	.394	-2.49% $[-5.42, .34]$	.201	.95
48	.975	1.479	34.08% $[32.42, 35.76]$	.485	2.04
96	2.630	5.858	55.10% $[53.30, 56.73]$	1.307	3.47
192	6.004	24.329	75.32% $[74.32, 76.25]$	2.983	7.14

E Related Work

Budgeted similarity and quantum-kernel acquisition. Prior work has reduced the cost of similarity-based learning without acquiring a complete similarity matrix. Shamir & Tishby (2011) study spectral clustering when only a limited number of similarity entries can be queried, while Wauthier et al. (2012) actively acquire similarities that reduce uncertainty in the current spectral-clustering solution. In quantum kernel learning, ShoFaR targets reliable classification rather than precise kernel estimation (Shastry et al., 2022), and recent methods allocate finite measurements across kernel entries according to downstream sensitivity (Xu et al., 2026a; Miroszewski, 2026; Xu et al., 2026b). These works share our motivation that measurements should be allocated according to the downstream task. Q-Cert instead targets exact certification of a fidelity-threshold partition: it controls the probability of incorrect certification, may abstain when the available evidence is insufficient, and combines pairwise overlap tests with reusable single-state evidence that can exclude many candidate edges simultaneously.

Kerenidis & Landman (2021) develop an end-to-end quantum algorithm for spectral clustering under quantum data-access assumptions, targeting the computational cost of spectral embedding and clustering. Our setting instead studies the finite-measurement evidence required to certify the connected-component partition of a fidelity-threshold graph from repeatably preparable quantum states.

Noisy-query and bandit clustering. Adaptive clustering from noisy observations has been studied under several feedback models. Xia & Huang (2022) recover a hidden partition using noisy answers to whether two items belong to the same cluster, relating query allocation to multi-armed bandits. Graf et al. (2025) instead query individual item–feature coordinates to recover two groups whose members share a feature vector, with the challenge of locating informative features. More recently, Thuot et al. (2026) study nonparametric clustering of bandit arms according to equality of their underlying distributions, using kernel mean embeddings and maximum mean discrepancy. Our target is the connected-component partition of a fidelity-threshold graph: belonging to the same component does not require every within-component pair to exceed the threshold, nor does it require identical states or measurement distributions. The feedback also differs. Echo produces a Bernoulli observation with mean F_{ij}, rather than a noisy same-cluster label; a single-state measurement returns a full outcome in an allowed basis, whose evidence can constrain multiple pairwise fidelities. Thus, adaptive partition recovery is established territory, while the target, measurement interface, and reusable certification bounds distinguish the present setting.

Active hierarchical clustering. Eriksson et al. (2011) study hierarchical clustering from adaptively selected pairwise similarities. Under a separation condition, they recover the hierarchy using $O(n \log n)$ similarity queries, with a robust variant handling anomalous similarities. Thus, recovering a clustering structure without acquiring all pairwise similarities is an established objective. Our setting instead concerns finite-sample certification of a fixed fidelity-threshold connected-component partition from stochastic quantum measurements. The distinction lies in the specified measurement interface, reusable state-level exclusion evidence, and the safety and online-cost guarantees of the resulting certification procedure.

Graph connectivity with noisy queries. Fotakis et al. (2023) study finding a spanning tree of an unknown surviving subgraph through noisy edge-existence queries, analyzing two-sided errors, false positives, and false negatives. Their work provides a close combinatorial precedent for acquiring connectivity witnesses without determining every edge. Q-Cert likewise uses spanning forests to establish internal connectivity, but must identify and certify all connected components of a graph defined by unknown quantum fidelities. This requires both positive connectivity witnesses and evidence excluding every possible cross-component connection. Moreover, an Echo outcome samples the fidelity itself, so the difficulty of deciding an edge depends on its distance from the chosen threshold. Reusable single-state bounds can exclude many potential edges through shared evidence, unlike an interface restricted to querying individual edge indicators. We therefore do not regard spanning-tree witnesses or the confirmed/possible graph sandwich as new ideas in isolation; our contribution is their integration with quantum measurement bounds and an adaptive stopping certificate under a specified oracle-cost model.

Reusable quantum measurement evidence. Classical shadows already establish that measurements of a quantum state can be stored as classical information and reused to predict many properties (Huang et al., 2020). Evidence reuse itself is therefore not a new principle of Q-Cert. Our construction uses outcomes from a prescribed finite basis menu to estimate masses on pilot-selected common and block-specific sets, including leakage outside those sets. Independent validation and simultaneous confidence bounds turn these estimates into fidelity upper bounds that can be shared across many cross-block pairs. These bounds are combined with direct Echo evidence and geometric propagation to certify a particular partition. The distinction is consequently the task-specific inference and acquisition guarantee: the algorithm seeks sufficient connectivity and exclusion evidence and stops when the remaining uncertainty cannot change the connected components. We do not claim a general improvement over classical-shadow prediction or tomography, and the usefulness of shared evidence depends on the available bases and the input states.

Scope of the guarantees. The results support distinct claims. Theorem 1 establishes safe acceptance under adaptive acquisition and stopping when the constructed bounds have simultaneous coverage; it does not require an accurate pilot. Theorem 2 bounds the current controller's oracle cost under explicit forest and separation conditions, including the cost of unsuccessful candidate edges. Proposition 1 establishes a gap-dependent separation between partition certification and full threshold-graph recovery under identical measurement access on a constructed pure-state family. Proposition 2 isolates the separate benefit of informative reusable measurement access for the same clustering task. These results do not establish optimality among algorithms for partition certification with the same interface. Proposition 3 realizes the near-linear guarantee on a structured pure-state family with bounded-degree graph topologies. Accordingly, our novelty claim concerns the combined clustering certificate, reusable bounds, and online cost analysis, not Clopper–Pearson intervals (Clopper & Pearson, 1934), graph sandwiches, or information reuse individually.