NISQ to Fault Tolerance: Technical Innovations, Real World Implementations, and Systemic Barriers in Scalable Quantum Computing

Asad Raza[1*], Ali Nawaz Sanjrani[2], Sajida Raz Bhutto[3], Muhammad Ejaz[4], Mohsin Ali[5]

***Abstract*: Fault tolerant quantum computing remains out of reach for most current hardware, which operates in the Noisy Intermediate Scale Quantum (NISQ) regime and is constrained by decoherence, imperfect gates, and limited circuit depth. This paper presents a structured, systems level synthesis of the NISQ to fault tolerance transition using systematic mapping and descriptive comparative analysis of peer reviewed benchmarks, experimental demonstrations, and documented industrial case studies reported between 2020 and early 2025. Rather than proposing a new quantum architecture, the study consolidates modality consistent scaling indicators across superconducting, trapped-ion, and photonic platforms and distinguishes near term feasibility evidence (error mitigation and application pilots) from fault tolerance readiness signals (logical qubit behavior that improves with redundancy). Results show that the mapped literature is dominated by milestone and reliability demonstrations, while application outcomes in optimization and finance remain strongly pipeline and baseline dependent, with performance sensitive to instance structure, sampling overhead, and control stability. The analysis further indicates that cross platform comparison is limited by heterogeneous reporting conventions, motivating milestone based synthesis instead of a single normalized metric. Overall, the findings clarify where progress is most robust, where evidence remains conditional, and which engineering and system level constraints most directly shape the path from NISQ devices toward scalable fault tolerant quantum computing.**

***Keywords:* NISQ, fault tolerant quantum computing, systematic mapping, error mitigation, quantum error correction, industrial case studies.**

I. INTRODUCTION

Quantum computing has progressed from a primarily theoretical discipline an emerging computational paradigm with early stage experimental and industrial relevance. Foundational quantum algorithms, including Shor's factorization and Grover's search, establish provable quantum advantages for specific problem classes; however, their large-scale realization remains constrained by contemporary hardware limitations [1]–[3]. Most available processors operate in the Noisy Intermediate Scale Quantum (NISQ) regime, where finite coherence times, gate imperfections, and environmental noise restrict achievable circuit depth and reliability [1]. These constraints limit near term performance to carefully structured, noise-aware workloads and prevent the direct execution of fully fault tolerant algorithms.

[1,4] School of Computer Science and Engineering, Central South University Changsha 410083, China
[2] Department of Mechanical Engineering, MUET, SZAB Campus, Khairpur Mir's
[3,5] Department of Software Engineering, MUET, SZAB Campus, Khairpur Mir's
Country: Pakistan, China
Email: *asadraza_153@hotmail.com

Between 2020 and 2025, measurable progress has been reported across multiple hardware modalities, including superconducting circuits, trapped-ion systems, and photonic architectures, noting that reported "scale" may be expressed as qubits or, in photonic settings, as modes/squeezed state resources rather than one to one physical qubits. Advances in device fabrication, control electronics, and cryogenic integration have improved system integration and enabled larger processors and more stable operation [4], [5]. Nevertheless, typical two qubit error rates and accumulated noise still impose stringent limits on circuit depth, and the overhead required for quantum error correction remains a central barrier to scalable fault tolerant quantum computing.

In parallel with these hardware and software developments, industrial stakeholders have explored NISQ-era systems for domain specific problems in finance and optimization, including annealing based workflows and hybrid approaches tailored to near term constraints [15], [16], [18]. While several studies report improvements under particular problem encodings and baselines, outcomes are often sensitive to noise levels, parameter tuning, instance structure, and classical pre and post processing overhead. As a result, evidence of performance gains is best interpreted as workload and setting dependent feasibility rather than broadly transferable quantum advantage.

A central challenge in the transition from NISQ devices to fault tolerant quantum computing is bridging experimental demonstrations and system level scalability. Prior work has examined individual components of this transition, including scaling trends, error mitigation techniques, and early surface- code demonstrations [7], [9]. However, fewer studies integrate these technical dimensions with system constraints such as infrastructure cost, workforce capacity, and the evolving landscape of post quantum cryptographic standardization [22]. These factors increasingly influence how quickly experimental advances translate into deployable and trustworthy quantum computing systems.

This paper addresses this gap by presenting a structured analysis of the NISQ to fault tolerance transition from a systems perspective, synthesizing evidence reported between 2020 and 2025. Rather than proposing a new quantum architecture or error correction code, the contribution is an integrative synthesis that connects experimental benchmarks, industrial implementations, and systemic constraints to clarify practical barriers and research priorities for scalable quantum computing beyond the NISQ era.

The remainder of the paper is organized as follows. Section II reviews related work on NISQ foundations, platform scaling, and error suppression strategies. Section III describes the methodology used for systematic mapping and comparative synthesis. Section IV summarizes results from the evidence map and cross platform comparisons. Section V discusses industrial and application-driven evidence. Section VI presents a broader discussion of system-level implications and constraints, followed by limitations and future directions in Section VII.

II. LITERATURE REVIEW

This section synthesizes representative prior work relevant to the NISQ to fault tolerance trajectory, with emphasis on (i) modality specific scaling evidence, (ii) reliability and control demonstrations that bound achievable circuit depth, and (iii) platform engineering factors that shape scalability constraints. The selected studies are grouped by hardware modality and by the type of evidence reported (device level demonstrations, platform milestones, and photonic resource scaling).

A. Trapped-Ion Reliability and Scalable Architectures

Trapped-ion platforms are frequently positioned as a high fidelity modality due to strong native coherence and controllable interactions. Device level evidence shows that trapped-ion qubits can achieve high operational quality under practical control settings, which supports deeper circuits under realistic noise budgets [12]. However, scaling trapped-ion systems beyond laboratory sized registers remains constrained by control complexity, mode management, and system throughput limitations. To address system scaling, modular architectures consistent with the quantum charge coupled device (QCCD) paradigm have been demonstrated, where transport and modular composition are treated as explicit design primitives for scaling [11]. While these approaches provide a concrete pathway toward larger systems, their fault tolerance readiness depends on integration of transport, calibration stability, and real time control at scale.

B. Superconducting Scaling Evidence and System Integration

Superconducting platforms have demonstrated rapid progress in integration maturity and system size, driven by advances in fabrication, packaging, and control engineering. Engineering focused analyses characterize the principal constraints that emerge with scale, including cross talk, wiring density, calibration overhead, and cryogenic infrastructure [6]. Platform milestone reporting further highlights the scale up trajectory of superconducting processors and the engineering challenges that accompany large device deployments [4], [5]. Collectively, this body of work supports the view that superconducting scaling is bounded as much by system engineering (calibration stability, control electronics, and packaging) as by device level coherence and gate fidelity, with direct implications for the overhead required for fault tolerant operation.

C. Photonic Quantum Computing and Resource Scaling

Photonic approaches provide a complementary route to scalable quantum information processing, particularly relevant to architectures where optical connectivity and distribution are natural. Large scale photonic demonstrations have reported computational advantage in specialized sampling regimes [13] and, more recently, programmable photonic processors supporting scalable experimental configurations [14]. Importantly, “scale” in photonics is often expressed in optical modes, multiplexed resources, or state counts rather than a direct physical qubit count, which complicates cross modality comparison. This motivates reporting modality consistent scale indicators and pairing them with workload or capability metrics when synthesizing progress toward fault tolerance.

D. Critical Gap and Positioning of this Work

Across modalities, prior studies often provide strong evidence for specific layers of the stack (device demonstrations, platform milestones, or photonic resource scaling), but they do not unify these results into a systems level assessment of the NISQ to fault tolerance transition. In particular, cross modality comparisons are hindered by non-uniform scale indicators and by incomplete linkage between (i) scale, (ii) reliability/throughput, and (iii) the overhead required for fault tolerant execution. In contrast, the present work consolidates modality consistent scale indicators and milestone summaries into an integrated perspective that explicitly frames scalability barriers across hardware, control, and deployment constraints.

III. METHODOLOGY

This study employs a structured, descriptive methodology to examine the transition from Noisy Intermediate Scale Quantum (NISQ) devices to fault tolerant quantum computing (FTQC). The work does not introduce new hardware experiments or a new error correction architecture. Instead, it synthesizes peer reviewed benchmarks, experimental demonstrations, and documented industrial case studies reported between 2020 and early 2025 using systematic mapping and descriptive comparative analysis.

A. Study Identification and Selection

Primary sources were gathered from IEEE Xplore, Nature Publishing Group journals, APS Physical Review journals, and SIAM venues. The review protocol follows established guidance for systematic literature reviews and mapping studies, with explicit screening and evidence categorization [19]–[21]. Studies were retained only when they reported measurable indicators relevant to NISQ performance or FTQC readiness, such as system scale, fidelity/error rates, logical qubit behavior, mitigation overhead, or application level outcomes with defined baselines. Conceptual discussions without measurable implications for near term or scalable systems were excluded.

B. Data Extraction and Evidence Coding

For each included study, we recorded (i) the platform/modality (superconducting, trapped-ion, photonic), (ii) the publication year and evidence type (hardware milestone, error mitigation, error correction, application case study, or systemic constraint), and (iii) quantitative indicators when reported (e.g., qubit/mode scale, two-qubit error rates, and logical error trends). The extracted items were coded into a master evidence table to support systematic mapping (coverage by year, platform, and evidence type) and to enable consistent descriptive summaries across heterogeneous reporting practices. In this mapping, one evidence item corresponds to one unique source counted once, even when the source reports multiple metrics or multiple experiments. Preprints and other non-archival sources were included only when they provide industry facing or systems level evidence not yet available in an archival venue, and they were coded separately from peer reviewed items to avoid conflating maturity signals.

C. Evaluation Metrics

Hardware trends were summarized using reported system scale (qubits or modes), representative reliability indicators when available (e.g., two qubit gate error rates), and operational readiness signals described by the sources (e.g., calibration stability and control overhead). Algorithmic and software level evidence was interpreted using standard error mitigation constructs, particularly the trade-off between estimator bias reduction and sampling overhead, following commonly used definitions in the mitigation literature [7], [8]. When FTQC progress was discussed, emphasis was placed on experimental results in which increased redundancy leads to improved logical behavior (e.g., surface code scaling trends) [9], [10], rather than isolated single instance outcomes; this choice reflects the goal of prioritizing trends that jointly capture hardware noise, control quality, and code scaling behavior.

D. Comparative Analysis

A cross platform descriptive comparison was conducted by grouping evidence by platform and year (2020–2025) and synthesizing trends in system scale, reliability indicators, and error suppression approaches. Because normalization across modalities is generally not meaningful under differing architectures and benchmarking conventions, the comparison is reported as trend level synthesis supported by milestone summaries, rather than a single aggregated performance score.

E. Industrial Case Study Analysis

Industrial implementations were treated as qualitative case studies. Each case study was evaluated by problem domain, quantum approach (e.g., annealing or gate-based hybrid methods), baseline definition, and the outcome metrics reported by the source. Particular attention was given to the conditions under which improvements were reported (instance structure, noise sensitivity, and classical preprocessing requirements), to separate domain specific feasibility from claims of broadly transferable advantage.

F. Limitations

The evidence base is heterogeneous, and reporting varies across independent studies and industrial disclosures with respect to experimental conditions, benchmarking protocols, and baseline definitions. Accordingly, the findings should be interpreted as structured synthesis and indicative trends, not as definitive cross platform performance guarantees.

IV. RESULTS AND ANALYSIS

This section reports the outcomes of the systematic mapping and descriptive comparative synthesis described in Section III. Results are presented as (i) an evidence map summarizing coverage across years and modalities, (ii) a milestone oriented view of reported scaling signals and reliability progress, and (iii) a cross platform synthesis supported by milestone and comparison tables.

A. Evidence map (coverage by year, modality, and evidence type)

Fig. 1 summarizes the distribution of included evidence across 2020–2025 and platform/modality categories. The resulting map indicates uneven coverage across modalities and years, motivating the milestone based synthesis used in this paper. Fig. 2 summarizes the evidence types in the included studies and indicates that the mapped literature is more densely populated by platform milestones and reliability demonstrations than by application case studies.

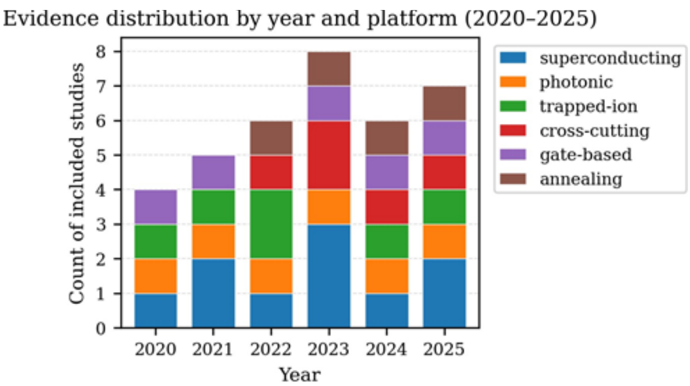

Fig. 1: Evidence distribution by year and platform/modality (2020–2025) across the included studies.

Fig. 3 provides a cross tabulation view of how evidence types are distributed across modalities/platform categories in the dataset, highlighting which categories dominate the evidence base used in this synthesis.

Fig. 4 summarizes cumulative evidence accumulation over time (overall and by platform/modality category). This pro- vides a compact view of when the included evidence base becomes denser, which helps contextualize the milestone focused synthesis used below.

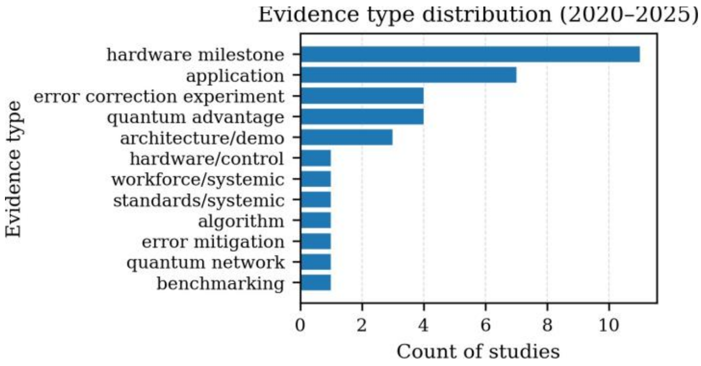

Fig. 2: Distribution of evidence types (2020–2025) across the included studies, highlighting emphasis on hardware mile- stones and reliability related demonstrations.

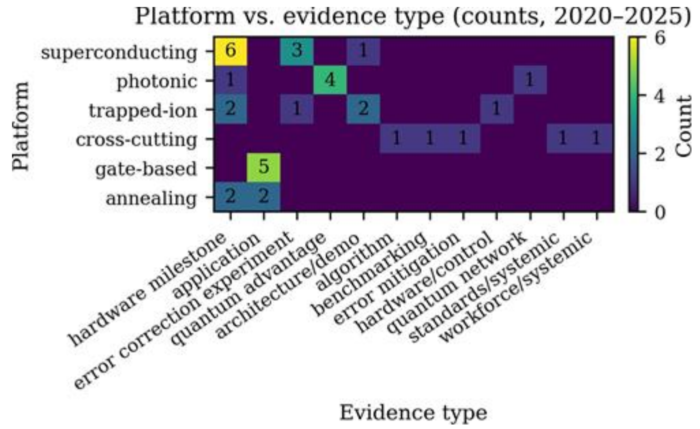

Fig. 3: Platform/modality versus evidence type (counts, 2020–2025), showing how platforms are represented across mile- stone, control, architecture, and application oriented evidence categories.

B. Platform scaling signals and milestone synthesis

Fig. 5 provides a milestone oriented summary of modality appropriate scale indicators. Superconducting scale indicators are taken from platform milestone sources [4], [5], while photonic "scale" reflects modality specific resource indicators from photonic demonstrations [13], [14]. Trapped ion evidence is represented using scalable architecture and high fidelity control demonstrations [11], [12]. Because scale is not directly comparable across modalities, this paper reports milestone based synthesis rather than a single unified performance score.

C. Reliability progress: from mitigation to logical-qubit behavior

The mapped reliability evidence separates into (i) error mitigation approaches that trade estimator bias reduction against sampling overhead [7], [8], and (ii) fault tolerance oriented error correction demonstrations in which logical behavior improves with increased redundancy [9], [10]. Within the scope of this study, the latter category provides the most direct readiness signal for FTQC because it jointly reflects hardware noise, control, and redundancy in a measurable trend.

D. Cross platform synthesis and bottlenecks

Table I consolidates the cross platform synthesis used in the NISQ to FTQC discussion. The table should be interpreted as comparative evidence synthesis rather than as a normalized benchmark across architectures.

E. Milestone evidence summary

Table II summarizes representative milestones used to construct the scaling synthesis (Fig. 5) and to anchor the cross- platform narrative.

TABLE I: Cross-platform comparison used in the NISQ to FTQC synthesis

Platform	Strengths (typical)	Scaling bottlenecks	2020–2025 highlights (as used here)	Key refs
Superconducting	Fast gates; strong industrial tollchains; mature fabrication	Cryogenics; crosstalk; calibration overhead; wiring/control scaling	Rapid growth in physical qubit scale and engineering integration; increasing emphasis on reliability and logical- qubit behavior under surface-code scaling	[4]–[6], [9], [10]
Trapped-ion	High fidelities; long coherence; high quality control primitives	Slower gates; transport/modularity engineering; throughput and control- stack complexity	QCCD-style architectural direction and high fidelity control demonstrations used as reliability/scalability evidence	[11], [12]
Photonic	Room temperature operation; net- working native; natural for sampling- style workloads	Loss; source/detector constraints; multiplexing/synchronization overhead	Photonic advantage demonstrations and programmable photonic processors used as resource scaling evidence (modes/samples rather than qubit count)	[13], [14]

TABLE II: Representative milestone oriented evidence used in the 2020–2025 synthesis (extended with the indicator reported/used for Fig. 5).

Year	Milestone (short)	Platform / modality	Reported indicator (for Fig. 5)	Primary source(s)
2020	Photonic computational advantage demonstration (re-source scale evidence reported in photonic terms)	Photonic	Modality specific resource indicator (e.g., photons/modes/samples; as reported)	[13]
2021	IBM public platform milestone (Eagle-class 127-qubit announcement used as scale indicator)	Superconducting	Physical qubit count (127 qubits)	[5]
2021	QCCD trapped-ion architecture demonstration (scalable architecture evidence)	Trapped-ion	Architecture scale indicator (QCCD/modularity; no sin-gle qubit-count baseline as-sumed)	[11]
2021	High-fidelity laser-free trapped-ion control demonstration (reliability evidence)	Trapped-ion	Control/reliability indicator (fidelity/control benchmark; not a scale metric)	[12]
2022	Programmable photonic processor (programmability + re-source scaling evidence)	Photonic	Resource/prog. indicator (modes/components; as reported)	[14]
2023	IBM 1,000-qubit-class chip announcement used as large-scale platform milestone indicator	Superconducting	Physical-qubit count (1,000-qubit-class; as reported)	[4]
2023	Surface code logical qubit scaling (logical behavior im-proves with increased redundancy)	Error correction	Logical behavior trend vs. redundancy (scaling signal; not physical-qubit count)	[9]
2025	Below-threshold quantum error correction report (FTQC readiness evidence)	Error correction	Below threshold / logical im-provement evidence (readi-ness signal)	[10]

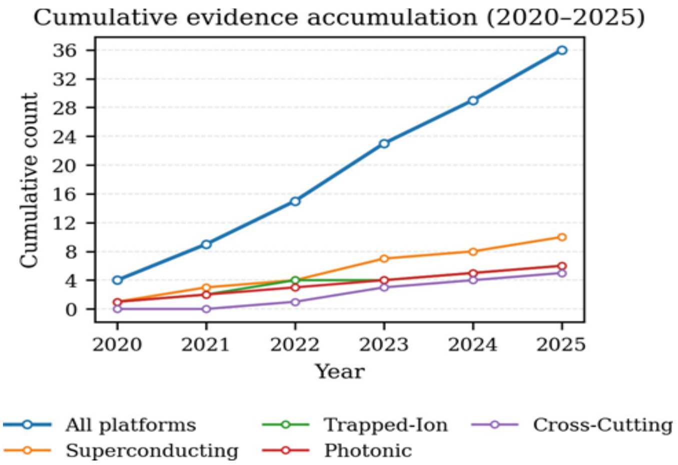

Fig. 4: Cumulative evidence accumulation from 2020 to 2025, shown overall and disaggregated by platform/modality category.

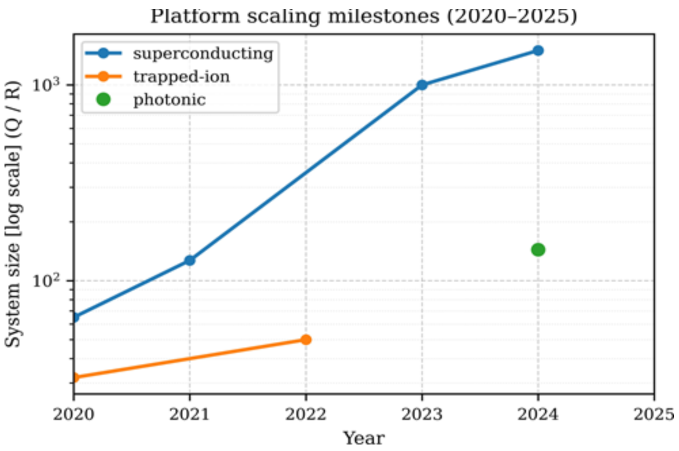

Fig. 5: Platform scaling milestones (2020–2025). Superconducting and trapped-ion values represent physical qubit scale, while photonic "scale" reflects modality appropriate resource indicators (e.g., modes/samples) rather than qubit count

F. Industrial and application driven evidence

Application facing studies are treated as case based evidence rather than as general purpose advantage claims. In optimization settings, reverse quantum annealing workflows report outcomes that depend on instance structure, embed- ding/preprocessing cost, and baseline selection [15], [16]. In finance oriented workloads, quantum Monte Carlo and derivative pricing pipelines provide end to end feasibility evidence but remain sensitive to circuit depth, error, and classical post- processing overhead [17], [18]. Consistent with Fig. 2, these studies are informative for feasibility boundaries and workload characterization, but should not be interpreted as modality independent evidence of scalable advantage.

G. System level implications

The mapped evidence supports a systems level reading of progress: scaling and reliability are advancing together but not uniformly across modalities, and metrics are not directly comparable across platforms. In addition, deployment planning is shaped by external constraints such as the post quantum cryptography transition, which proceeds on timelines that do not depend on the arrival of large scale FTQC hardware [22].

V. INDUSTRIAL AND APPLICATION DRIVEN EVIDENCE

This section synthesizes the application and deployment oriented evidence identified in the mapping study and connects it to the platform trends summarized in Section IV (Figs. 5– 2 and Table I). The goal is not to claim broad "quantum advantage" across industries, but to characterize what has been demonstrated in practice between 2020 and 2025, under what conditions such results were reported, and why these conditions matter for the NISQ→FTQC transition.

A. Optimization evidence: annealing and hybrid workflows

A recurring pattern in industrial reports is that performance claims are typically tied to structured instances and to pipelines that combine quantum sampling with classical preprocessing and post processing. In this category, reverse quantum annealing has been used as a practical strategy for refinement around candidate solutions (warm starts), which aligns well with real operations research workflows where good feasible solutions are often available from classical heuristics. Evidence in this direction emphasizes that observed gains depend on instance structure, embedding overhead, and sensitivity to control noise, rather than on device scale alone [15]. More application specific studies (e.g., logistics oriented formulations) similarly report improvements relative to selected classical baselines, but the reported outcomes remain contingent on problem formulation choices, constraint handling, and the cost of classical components in the end to end pipeline [16]. From a systems perspective, these results are best interpreted as demonstrations of NISQ feasibility for narrow tasks (and for particular regimes of problem sizes), rather than as general purpose speedups that extrapolate cleanly with qubit count.

B. Finance evidence: sampling complexity and risk sensitive workloads

Financial workloads provide a useful stress test for NISQ devices because they are often dominated by sampling complexity and are sensitive to estimator variance. Several studies frame potential improvements through quantum Monte Carlo type approaches or amplitude estimation inspired workflows, while acknowledging that near term implementations must be hybridized and carefully noise managed. Demonstrations and experiments in this direction commonly report improvements over naive sampling strategies under specific assumptions, while still requiring substantial classical orchestration and error aware circuit design [17], [18]. In the context of the evidence map (Fig. 3), finance oriented results should therefore be treated as workload capability signals i.e., indicators of what types of probabilistic estimation tasks can be executed meaningfully under NISQ constraints rather than as definitive evidence of scalable advantage.

C. Linking application results to reliability limits and FTQC readiness

Across both optimization and finance, the dominant limiting factors are not simply "more qubits," but (i) achievable circuit depth under realistic noise budgets, (ii) calibration stability over run time, and (iii) the sampling overhead required to suppress bias or variance in estimators. These constraints link application evidence directly to the reliability trends discussed in Section IV and to the mitigation/QEC evidence types in Fig. 2. In particular, error mitigation methods are repeatedly positioned as enablers of near term usefulness, but their cost is often paid in additional sampling, which can offset wall clock benefits unless the hardware and control stack deliver stable, repeatable execution [7], [8]. Consequently, application facing progress should be interpreted jointly with reliability evidence: application outcomes that appear favorable in a narrow setting do not necessarily imply an imminent transition to fault tolerant workloads, but they do help identify which workloads are most compatible with NISQ-era constraints.

D. Cross cutting systemic signals: standardization pressure and deployability

Industrial deployment is also shaped by system level constraints that sit outside core device physics. One prominent example is the parallel progress of post quantum cryptography (PQC) standardization, which provides a practical security pathway that does not depend on near term fault tolerant quantum computers. The NIST PQC process therefore creates a real "deployment baseline" against which quantum enabled security claims must be evaluated [22]. This does not diminish the long term importance of fault tolerant quantum computation, but it does mean that industrial timelines and incentives are influenced by competing, mature alternatives in security critical domains. In the evidence map, these cross cutting pressures are best viewed as part of the "systemic" category that conditions whether technical advances translate into adoption at scale.

E. Summary

Industrial and application facing studies are treated here as bounded feasibility evidence under specific workflows, baselines, and operational constraints, and are not interpreted as generalizable or modality independent proof of quantum advantage. Overall, the industrial evidence identified in this study supports three conservative conclusions. First, NISQ-era application reports are predominantly problem and pipeline dependent, with outcomes sensitive to instance structure, modeling choices, and classical overhead [15], [16]. Second, application capability is tightly coupled to reliability and mitigation costs, implying that progress in QEC/mitigation and control stability is a prerequisite for broader transferability [7], [8]. Third, deployability is shaped by external constraints and alternative technological pathways (notably PQC standard-ization), which must be considered when interpreting claims about near-term industrial impact [22].

VI. DISCUSSION

The results in Section IV support a systems level interpretation of the NISQ to FTQC trajectory: progress is visible across modalities, but it is expressed through heterogeneous indicators, and the strongest evidence is concentrated in milestone and reliability demonstrations rather than in broadly transferable application level advantages. This section discusses what the mapped evidence implies for (i) interpreting "scale," (ii) the role of mitigation versus error correction, (iii) cross platform readiness signals, and (iv) how application evidence should be read under NISQ constraints.

A. Interpreting "scale" across modalities

Fig. 5 and Table II highlight that platform growth is commonly reported using different notions of scale. For superconducting systems, physical qubit counts are often used as headline indicators [4], [5], while photonic demonstrations emphasize modality specific resource measures (e.g., modes, samples, or programmable photonic processor scale) [13], [14]. Trapped-ion work, in contrast, often provides scaling evidence through architectural demonstrations and control fidelity rather than through rapid growth in raw qubit counts [11], [12]. As a result, cross modality comparisons based on a single numerical metric risk being misleading. The milestone based synthesis used here therefore treats "scale" as one component of readiness, and it pairs it with reliability oriented evidence to maintain interpretability across heterogeneous platforms.

B. Mitigation extends NISQ feasibility; QEC determines FTQC readiness

The evidence map (Figs. 2–3) indicates sustained attention to error mitigation as a pragmatic response to NISQ noise. Mitigation methods can improve the usefulness of shallow circuits, but they typically do so by increasing sampling cost or operational complexity, which can erode end to end gains if control stability is limited [7], [8]. In contrast, the most direct signal of progress toward fault tolerance is experimental evidence where increased redundancy yields improved logical behavior, as emphasized by surface code scaling experiments and below threshold results [9], [10]. In practical terms, this distinction matters: mitigation supports near term experimentation and narrow feasibility, while error correction defines whether scalable computation can be sustained when circuit depth and algorithmic complexity increase.

C. Cross platform readiness signals and bottlenecks

The platform comparison in Table I suggests that readiness is constrained by different bottlenecks across modalities. Superconducting systems benefit from strong engineering ecosystems and rapid scale growth, but increasing system size couples tightly to calibration overhead, cryogenic integration, and control stack complexity [4], [6]. Trapped-ion systems exhibit high quality control and coherence, yet scaling depends on architectural modularization and throughput management, making integration and real time control key determinants of practical progress [11], [12]. Photonic approaches show strong potential for resource scaling and programmability, but loss and multiplexing/synchronization overhead remain dominant constraints that affect how such systems could support fault tolerant workloads at meaningful rates [13], [14]. The main implication is that "the path to FTQC" is not uniform: the relevant engineering milestones differ by platform, and evaluating progress requires platform consistent indicators paired with reliability evidence.

D. Reading industrial evidence conservatively

The application and industrial evidence discussed in Section V is best interpreted as feasibility evidence for specific workflows rather than as proof of general advantage. Optimization and finance studies often report results within tightly specified pipelines that include classical preprocessing and domain specific baselines [15]–[18]. These studies are valuable because they expose practical constraints (instance structure, embedding cost, sampling burden, and stability requirements) and because they clarify which workload families are most compatible with NISQ-era devices. However, the mapping results indicate that application evidence is less dense than platform/reliability evidence, and claims should therefore be framed with clear boundary conditions rather than extrapolated to broad conclusions.

E. System level implications

Finally, the transition to deployment relevant quantum computing is shaped not only by device progress but also by competing and complementary pathways. In security critical settings, post quantum cryptography standardization provides a concrete migration track that is decoupled from the arrival timeline of large scale FTQC hardware [22]. This does not replace the long term significance of fault tolerance, but it affects how near term industrial incentives, risk models, and adoption timelines are set. A realistic assessment of the NISQ to FTQC transition therefore benefits from integrating platform progress with external constraints and technology alternatives, consistent with the systems perspective adopted in this work.

VII. LIMITATIONS AND THREATS TO VALIDITY

This study is designed as a structured evidence synthesis rather than an experimental benchmarking campaign. The following limitations should be considered when interpreting the results.

A. Construct validity (metric meaning and comparability)

A primary threat is that key indicators are not directly comparable across modalities. Physical qubit counts in super- conducting systems are not equivalent to photonic resource indicators, and trapped-ion progress is often expressed through architectural and control demonstrations rather than raw scale. The milestone oriented approach mitigates this by avoiding a single unified score, but it cannot eliminate the underlying construct mismatch across platforms.

B. Internal validity (causal interpretation)

Because this work synthesizes reported results, it does not establish causal relationships between specific engineering choices and observed outcomes. Differences in device calibration practices, error models, compilation strategies, and experimental protocols across independent studies can con- found interpretations. Accordingly, the analysis emphasizes descriptive trends and readiness signals rather than causal claims about which design decisions are optimal.

C. External validity (generalizability)

Application and industrial case studies may not generalize beyond the reported regimes. Reported outcomes frequently depend on instance structure, baseline definition, and the role of classical components in end to end pipelines. The results should therefore be interpreted as evidence of feasibility under specific conditions, not as universally transferable performance claims.

D. Conclusion validity (evidence coverage and reporting bias)

The evidence map shows uneven coverage across time, modality, and evidence type. This introduces a risk of reporting bias, where highly visible milestones and flagship demonstrations are overrepresented relative to incremental or negative results. The systematic mapping approach partially addresses this by making coverage explicit (Figs. 1–4), but it cannot guarantee completeness given differing publication and disclosure practices across academia and industry.

VIII. CONCLUSION AND FUTURE WORK

This paper presented a structured synthesis of the NISQ to fault tolerance transition using systematic mapping and descriptive comparative analysis of evidence reported between 2020 and early 2025. The results show that the most robust signals of progress arise from platform milestones and reliability demonstrations, while application facing evidence remains more conditional and sensitive to workflow design and baseline choices. Across modalities, progress is best interpreted through platform consistent indicators paired with reliability trends, rather than through a single cross platform metric.

The evidence further supports a practical distinction be- tween near term and long term enablers: error mitigation extends feasibility for selected NISQ workloads, whereas fault tolerant error correction demonstrations provide the most direct readiness signals for scalable computation. Cross plat- form synthesis indicates that different modalities face distinct engineering bottlenecks, implying that evaluations of readiness should remain modality aware and tied to measurable reliability behavior.

Future work can strengthen this systems level assessment in three directions. First, expanding the master evidence table with a richer set of explicitly defined quantitative fields (e.g., reliability indicators, logical performance trends, and baseline definitions for application studies) would enable more fine- grained stratification without forcing cross platform normal- ization. Second, the framework can be extended to incorpo- rate workload centric capability indicators (e.g., depth/width regimes and sampling budgets) that better connect device progress to algorithmic feasibility. Third, longitudinal updates beyond 2025, with explicit tracking of reporting standards and reproducibility signals, would improve comparability and reduce the influence of disclosure bias in industrial reporting.

REFERENCES

[1] J. Preskill, "Quantum computing in the NISQ era and beyond," Quan- tum, vol. 2, p. 79, 2018, doi: 10.22331/q-2018-08-06-79.

[2] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," SIAM J. Comput., vol. 26, no. 5, pp. 1484–1509, 1997, doi: 10.1137/S0097539795293172.

[3] L. K. Grover, "Quantum mechanics helps in searching for a needle in a haystack," Phys. Rev. Lett., vol. 79, no. 2, pp. 325–328, 1997, doi: 10.1103/PhysRevLett.79.325.

[4] D. Castelvecchi, "IBM releases first-ever 1,000-qubit quantum chip,"Nature, vol. 624, no. 7991, p. 238, Dec. 2023, doi: 10.1038/d41586-023-03854-1.

[5] IBM, "IBM unveils breakthrough 127-qubit quantum processor (Eagle)," IBM Newsroom, Nov. 16, 2021. [Online]. Available: https://newsroom.ibm.com/2021-11-16-IBM-Unveils-Breakthrough-127-Qubit Quantum-Processor

[6] P. Krantz et al., "A quantum engineer's guide to superconducting qubits," Appl. Phys. Rev., vol. 6, no. 2, p. 021318, 2019, doi: 10.1063/1.5089550.

[7] K. Temme, S. Bravyi, and J. M. Gambetta, "Error mitigation for short- depth quantum circuits," Phys. Rev. Lett., vol. 119, no. 18, p. 180509, 2017, doi: 10.1103/PhysRevLett.119.180509.

[8] S. Endo, S. C. Benjamin, and Y. Li, "Practical quantum error mitigation for near-future applications," Phys. Rev. X, vol. 8, no. 3, p. 031027, 2018, doi: 10.1103/PhysRevX.8.031027.

[9] R. Acharya et al. (Google Quantum AI), "Suppressing quantum errors by scaling a surface code logical qubit," Nature, vol. 614, no. 7949,pp. 676–681, 2023, doi: 10.1038/s41586-022-05434-1.

[10] Google Quantum AI and collaborators, "Quantum error correction below the surface code threshold," Nature, 2025, doi: 10.1038/s41586-024- 08449-y.

[11] J. M. Pino et al., "Demonstration of the trapped-ion quantum CCD computer," Nature, vol. 592, no. 7853, pp. 209–213, 2021, doi: 10.1038/s41586-021-03318-4.

[12] R. Srinivas et al., "High-fidelity laser-free universal control of trapped ion qubits," Nature, vol. 597, no. 7875, pp. 209–213, 2021, doi: 10.1038/s41586-021-03809-4.

[13] H.-S. Zhong et al., "Quantum computational advantage using photons," Science, vol. 370, no. 6523, pp. 1460–1463, 2020, doi: 10.1126/sci- ence.abe8770.

[14] L. S. Madsen et al., "Quantum computational advantage with a pro- grammable photonic processor," Nature, vol. 606, no. 7912, pp. 75–81, 2022, doi: 10.1038/s41586-022-04725-x.

[15] S. Yarkoni et al., "Quantum annealing for industry applications: intro- duction and review," arXiv preprint arXiv:2112.07491, 2021. [Online].Available: https://arxiv.org/abs/2112.07491

[16] D. Venturelli et al., "Quantum annealing for electric vehicle logistics optimization," Phys. Rev. Applied, vol. 19, no. 3, p. 034056, 2023, doi: 10.1103/PhysRevApplied.19.034056.

[17] N. Stamatopoulos et al., "Option pricing using quantum computers," Quantum, vol. 4, p. 291, 2020, doi: 10.22331/q-2020-07-06-291.

[18] A. Kandala et al., "Quantum-enhanced Monte Carlo methods for deriva- tive pricing," Nat. Comput. Sci., vol. 2, no. 12, pp. 789–795, Dec. 2022, doi: 10.1038/s43588-022-00385-z.

[19] B. Kitchenham and S. Charters, Guidelines for Performing Systematic Literature Reviews in Software Engineering. Keele, U.K.: Keele Univ., Tech. Rep. EBSE-2007-01, 2007. [Online]. Available: https://legacyfil eshare.elsevier.com/promis misc/525444systematicreviewsguide.pdf

[20] M. J. Page, J. E. McKenzie, P. M. Bossuyt, et al., "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," BMJ, vol. 372, p. n71, 2021, doi: 10.1136/bmj.n71.

[21] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic mapping studies in software engineering," in Proc. 12th Int. Conf. Evaluation and Assessment in Software Engineering (EASE), 2008, doi: 10.14236/ewic/EASE2008.8.

[22] G. Alagic et al., Status Report on the Third Round of the NIST Post- Quantum Cryptography Standardization Process. Gaithersburg, MD, USA: National Institute of Standards and Technology (NIST), U.S. Dept. of Commerce, 2022.

[23] U.S. Congress, National Quantum Initiative Act, Public Law 115–368, Dec. 21, 2018. [Online]. Available: https://www.govinfo.gov/app/detail s/COMPS-15322

[24] National Academies of Sciences, Engineering, and Medicine, Quantum Computing: Progress and Prospects, Washington, DC, USA: The Na- tional Academies Press, 2019, doi: 10.17226/25196. [Online]. Available: https://nap.edu/25196.