Comprehensive Performance Evaluation of Post-Quantum Cryptographic Algorithms on Resource-Constrained ARM Platforms

Benchmark Analysis System
Automated Performance Evaluation
Raspberry Pi 4 Model B Platform
January 2026

***Abstract*—This report presents a comprehensive performance evaluation of post-quantum cryptographic (PQC) algorithms on a resource-constrained ARM platform (Raspberry Pi 4 Model B). We evaluate nine Key Encapsulation Mechanisms (KEMs) across three families (ML-KEM, Classic McEliece, HQC) and eight digital signature schemes across three families (ML-DSA, Falcon, SPHINCS+). All algorithms are evaluated at multiple NIST security levels (L1, L3, L5). We measure key generation, encapsulation/signing, and decapsulation/verification operations with 200 iterations each, collecting mean, median, standard deviation, minimum, maximum, and 95th percentile statistics. Size metrics including public key, secret key, ciphertext, and signature sizes are recorded. All 98 benchmark files comprising 19,600 timing measurements achieved 100% success rate. This document provides detailed statistical analysis, comparative visualizations, anomaly detection, and trade-off analysis without interpretive conclusions.**

***Index Terms*—Post-Quantum Cryptography, NIST PQC, Performance Evaluation, ARM, Embedded Systems, ML-KEM, ML-DSA, Falcon, SPHINCS+, Classic McEliece, HQC**

I. Introduction

A. Motivation

The advent of quantum computing poses a significant threat to currently deployed cryptographic systems. NIST's Post-Quantum Cryptography Standardization process has selected several algorithms for standardization, and understanding their performance characteristics on resource-constrained platforms is critical for deployment planning.

B. Scope

This benchmark report evaluates:

- **9 KEM algorithms**: ML-KEM (512/768/1024), Classic McEliece (348864/460896/8192128), HQC (128/192/256)
- **8 Signature algorithms**: ML-DSA (44/65/87), Falcon (512/1024), SPHINCS+ (128s/192s/256s)
- **3 AEAD ciphers**: AES-256-GCM, ChaCha20-Poly1305, Ascon-128a
- **23 Cipher suites**: Combinations for full handshake evaluation

C. Document Organization

Section II describes the experimental setup. Section III details the measurement methodology. Sections V and VI present KEM and signature benchmark results respectively. Section VII provides comparative analysis across algorithms and NIST levels. Section VIII discusses outliers and anomalies. Section IX analyzes size-timing trade-offs. Section X covers full handshake results.

II. Experimental Setup

A. Hardware Platform

TABLE I: Hardware Specifications

Component	**Specification**
Device Model	Raspberry Pi 4 Model B Rev 1.5
System-on-Chip	Broadcom BCM2711
CPU	Quad-core ARM Cortex-A72
Architecture	ARMv8-A (64-bit)
CPU Frequency	1.8 GHz (max)
L1 Cache	32 KB I-cache, 32 KB D-cache per core
L2 Cache	1 MB shared
Memory	4 GB LPDDR4-3200 SDRAM
Storage	microSD (Class 10)
Frequency Governor	`ondemand`

Data Source: */proc/cpuinfo*, */proc/device-tree/model*

B. Software Environment

TABLE II: Software Versions

Component	**Version**
Operating System	Debian GNU/Linux 12 (Bookworm)
Linux Kernel	6.12.47+rpt-rpi-v8 (aarch64)
Python	3.11.2
GCC	12.2.0
liboqs-python	0.14.0
liboqs (native)	0.14.1-dev
cryptography	46.0.2
ascon	0.0.9
NumPy	2.2.6

Data Source: `bench_results/environment.json`

C. Repository State

TABLE III: Source Code State

Attribute	Value
Git Commit	`49ed212352374881...`
Branch	main
Dirty State	Yes (uncommitted changes)
Benchmark Timestamp	2026-01-10T05:44:22Z

III. Measurement Methodology

A. Timing Instrumentation

Each cryptographic operation is timed using two methods:

1) **Performance Counter** (`time.perf_counter_ns()`): High-resolution monotonic clock, not affected by system time adjustments. This is the primary timing source for all reported results.
2) **Wall Clock** (`time.time_ns()`): Real-time clock for validation and cross-reference.

B. Iteration Parameters

TABLE IV: Benchmark Configuration

Parameter	Value
Iterations per operation	200
Warm-up iterations	0 (all iterations recorded)
Inter-iteration delay	None
CPU isolation	Not applied
Process priority	Default

C. Operations Measured

TABLE V: Cryptographic Operations per Algorithm Type

Algorithm Type	Operations Measured
Key Encapsulation (KEM)	`keygen`, `encapsulate`, `decapsulate`
Digital Signature	`keygen`, `sign`, `verify`
AEAD Cipher	`encrypt`, `decrypt`
Cipher Suite	`full_handshake`

D. Statistical Metrics Collected

For each operation, the following statistics are computed from the 200 timing samples:

- **Mean** ($\bar{x}$): Arithmetic average
- **Median** ($\tilde{x}$): 50th percentile
- **Standard Deviation** (σ): Sample standard deviation
- **Minimum**: Fastest observed execution
- **Maximum**: Slowest observed execution
- **95th Percentile** (P95): Value below which 95% of samples fall

E. Size Metrics Collected

- **Public Key Size**: Bytes required for public key storage
- **Secret Key Size**: Bytes required for secret key storage
- **Ciphertext/Signature Size**: Bytes of cryptographic output

IV. Data Summary

A. Benchmark Coverage

TABLE VI: Benchmark File Inventory

Category	Files	Iterations	Success
KEM	27	5,400	100.00%
Signature	24	4,800	100.00%
AEAD	24	4,800	100.00%
Cipher Suite	23	4,600	100.00%
Total	**98**	**19,600**	**100.00%**

B. Algorithm Coverage by NIST Level

TABLE VII: Algorithms by NIST Security Level

Family	Level 1	Level 3	Level 5
ML-KEM	512	768	1024
Classic McEliece	348864	460896	8192128
HQC	128	192	256
ML-DSA	44	65	87
Falcon	512	–	1024
SPHINCS+	128s	192s	256s

V. Key Encapsulation Mechanism Results

A. ML-KEM (NIST FIPS 203)

ML-KEM is based on the Module Learning With Errors (MLWE) problem. It offers the smallest key and ciphertext sizes among the evaluated KEMs while maintaining competitive performance.

TABLE VIII: ML-KEM Key Generation Timing (n=200)

Variant	Mean (ms)	Median (ms)	σ (ms)	Min (ms)	Max (ms)	P95 (ms)
ML-KEM-512	0.116	0.082	0.449	0.080	6.435	0.098
ML-KEM-768	0.111	0.107	0.040	0.106	0.664	0.116
ML-KEM-1024	0.142	0.136	0.029	0.134	0.510	0.163

1) Key Generation Performance: **Observations:**

- ML-KEM-512 shows the highest variance ($\sigma = 0.449$) due to occasional system interference
- Median values are more representative than mean due to outlier presence
- Performance scales approximately linearly with security level

TABLE IX: ML-KEM Encapsulation Timing (n=200)

Variant	Mean (ms)	Median (ms)	σ (ms)	Min (ms)	Max (ms)	P95 (ms)
ML-KEM-512	0.066	0.062	0.027	0.060	0.341	0.072
ML-KEM-768	0.089	0.086	0.023	0.085	0.361	0.091
ML-KEM-1024	0.121	0.118	0.022	0.117	0.394	0.130

2) Encapsulation Performance:

TABLE X: ML-KEM Decapsulation Timing (n=200)

Variant	Mean (ms)	Median (ms)	σ (ms)	Min (ms)	Max (ms)	P95 (ms)
ML-KEM-512	0.071	0.067	0.022	0.065	0.355	0.084
ML-KEM-768	0.097	0.094	0.018	0.093	0.348	0.100
ML-KEM-1024	0.144	0.136	0.033	0.132	0.551	0.176

3) Decapsulation Performance:

TABLE XI: ML-KEM Size Metrics (bytes)

Variant	Public Key	Secret Key	Ciphertext	Shared Secret
ML-KEM-512	800	1,632	768	32
ML-KEM-768	1,184	2,400	1,088	32
ML-KEM-1024	1,568	3,168	1,568	32

4) Size Metrics:

B. Classic McEliece

Classic McEliece is based on the Niederreiter cryptosystem using binary Goppa codes. It has the largest public keys but smallest ciphertexts.

TABLE XII: Classic McEliece Key Generation Timing (n=200)

Variant	Mean (ms)	Median (ms)	σ (ms)	Min (ms)	Max (ms)	P95 (ms)
348864	333.39	228.62	222.13	151.12	1524.76	775.10
460896	1114.67	911.52	774.42	465.01	5149.97	2623.10
8192128	8834.74	7065.81	6919.74	2467.11	36617.42	25241.50

1) Key Generation Performance: **Critical Observation:** Classic McEliece key generation exhibits extreme variance. The 8192128 variant shows a maximum of 36.6 seconds versus a minimum of 2.5 seconds—a 14.8× ratio. This is due to the probabilistic nature of the Goppa code generation.

TABLE XIII: Classic McEliece Encapsulation/Decapsulation (n=200)

Variant	Encaps Mean (ms)	Encaps Median (ms)	Decaps Mean (ms)	Decaps Median (ms)
348864	0.27	0.26	55.45	55.43
460896	0.66	0.64	89.40	89.38
8192128	2.01	1.99	209.06	209.00

2) Encapsulation and Decapsulation: **Observation:** Encapsulation is extremely fast (sub-millisecond for L1), but decapsulation is computationally intensive.

TABLE XIV: Classic McEliece Size Metrics (bytes)

Variant	Public Key	Secret Key	Ciphertext	Shared Secret
348864	261,120	6,492	96	32
460896	524,160	13,608	156	32
8192128	1,357,824	14,120	208	32

3) Size Metrics: **Critical Note:** The 8192128 variant has a 1.36 MB public key, which may be prohibitive for constrained environments.

C. HQC

HQC (Hamming Quasi-Cyclic) is based on the syndrome decoding problem.

TABLE XV: HQC Complete Timing Results (n=200)

Variant	Operation	Mean (ms)	Median (ms)	Min (ms)	Max (ms)
HQC-128	keygen	22.10	22.06	21.99	24.83
	encapsulate	44.67	44.54	44.47	46.89
	decapsulate	73.05	73.03	72.87	73.83
HQC-192	keygen	67.45	67.36	67.26	72.68
	encapsulate	135.39	135.26	135.10	140.50
	decapsulate	211.19	211.14	210.85	213.35
HQC-256	keygen	123.59	123.54	123.40	126.32
	encapsulate	248.79	248.68	248.46	252.93
	decapsulate	392.31	392.15	391.65	401.15

1) Complete Timing Results: **Observation:** HQC shows very low variance (tight min-max range), indicating consistent performance.

VI. Digital Signature Results

A. ML-DSA (NIST FIPS 204)

ML-DSA (formerly Dilithium) is based on Module Learning With Errors and Module Short Integer Solution problems.

TABLE XVI: ML-DSA Complete Timing Results (n=200)

Variant	Operation	Mean (ms)	Median (ms)	Min (ms)	Max (ms)
ML-DSA-44	keygen	0.26	0.25	0.25	0.72
	sign	1.03	0.85	0.42	4.11
	verify	0.25	0.25	0.25	0.47
ML-DSA-65	keygen	0.42	0.41	0.41	0.80
	sign	1.59	1.29	0.61	6.89
	verify	0.38	0.38	0.38	0.53
ML-DSA-87	keygen	0.61	0.61	0.60	0.96
	sign	1.77	1.48	0.92	6.17
	verify	0.61	0.61	0.61	0.76

1) Complete Timing Results: **Observation:** Signing times show higher variance than keygen/verify due to rejection sampling.

TABLE XVII: ML-DSA Size Metrics (bytes)

Variant	Public Key	Secret Key	Signature
ML-DSA-44	1,312	2,560	2,420
ML-DSA-65	1,952	4,032	3,309
ML-DSA-87	2,592	4,896	4,627

2) Size Metrics:

B. Falcon

Falcon is based on NTRU lattices with Gaussian sampling.

TABLE XVIII: Falcon Complete Timing Results (n=200)

Variant	Operation	Mean (ms)	Median (ms)	Min (ms)	Max (ms)
Falcon-512	keygen	18.87	17.63	13.64	41.62
	sign	0.65	0.64	0.63	1.36
	verify	0.11	0.11	0.11	0.31
Falcon-1024	keygen	51.01	47.29	41.60	111.87
	sign	1.31	1.30	1.27	1.80
	verify	0.20	0.19	0.19	0.42

1) Complete Timing Results: **Key Observation:** Falcon has the fastest verification times of all evaluated signature schemes.

TABLE XIX: Falcon Size Metrics (bytes)

Variant	Public Key	Secret Key	Signature
Falcon-512	897	1,281	659
Falcon-1024	1,793	2,305	1,267

2) Size Metrics: **Note:** Falcon produces the smallest signatures among evaluated schemes.

C. SPHINCS+

SPHINCS+ is a stateless hash-based signature scheme.

TABLE XX: SPHINCS+ Complete Timing Results (n=200)

Variant	Operation	Mean (ms)	Median (ms)	Min (ms)	Max (ms)
128s	keygen	193.26	193.11	192.90	197.68
	sign	1460.87	1460.29	1459.37	1470.58
	verify	1.49	1.49	1.48	1.65
192s	keygen	280.88	280.55	280.26	287.36
	sign	2611.10	2598.47	2596.17	4807.13
	verify	2.20	2.19	2.18	2.38
256s	keygen	186.05	186.00	185.67	187.63
	sign	2308.36	2307.46	2305.92	2325.33
	verify	3.12	3.09	3.08	3.51

1) Complete Timing Results: **Critical Observation:** SPHINCS+ signing is extremely slow (1.5-2.6 seconds), making it unsuitable for latency-sensitive applications. However, verification is fast.

TABLE XXI: SPHINCS+ Size Metrics (bytes)

Variant	Public Key	Secret Key	Signature
128s	32	64	7,856
192s	48	96	16,224
256s	64	128	29,792

2) Size Metrics: **Note:** SPHINCS+ has the smallest keys but largest signatures.

VII. Comparative Analysis

A. NIST Level Comparison

TABLE XXII: KEM Mean Timing by NIST Security Level (ms)

Level	Metric	ML-KEM	McEliece	HQC
L1	keygen	0.116	333.39	22.10
	encaps	0.066	0.27	44.67
	decaps	0.071	55.45	73.05
L3	keygen	0.111	1114.67	67.45
	encaps	0.089	0.66	135.39
	decaps	0.097	89.40	211.19
L5	keygen	0.142	8834.74	123.59
	encaps	0.121	2.01	248.79
	decaps	0.144	209.06	392.31

1) KEM Operations by NIST Level:

TABLE XXIII: Signature Mean Timing by NIST Security Level (ms)

Level	Metric	ML-DSA	Falcon	SPHINCS+
L1	keygen	0.26	18.87	193.26
	sign	1.03	0.65	1460.87
	verify	0.25	0.11	1.49
L3	keygen	0.42	–	280.88
	sign	1.59	–	2611.10
	verify	0.38	–	2.20
L5	keygen	0.61	51.01	186.05
	sign	1.77	1.31	2308.36
	verify	0.61	0.20	3.12

2) Signature Operations by NIST Level:

B. Cross-Family Performance Ranking

1) KEM Ranking (Lower is Better):

1) **Fastest KeyGen**: ML-KEM (all variants < 0.15 ms)
2) **Fastest Encapsulation**: ML-KEM-512 (0.066 ms)
3) **Fastest Decapsulation**: ML-KEM-512 (0.071 ms)
4) **Smallest Public Key**: ML-KEM-512 (800 bytes)
5) **Smallest Ciphertext**: Classic McEliece-348864 (96 bytes)

2) Signature Ranking (Lower is Better):

1) **Fastest KeyGen**: ML-DSA-44 (0.26 ms)
2) **Fastest Signing**: Falcon-512 (0.65 ms)
3) **Fastest Verification**: Falcon-512 (0.11 ms)
4) **Smallest Public Key**: SPHINCS+-128s (32 bytes)
5) **Smallest Signature**: Falcon-512 (659 bytes)

VIII. Anomaly Detection and Outlier Analysis

A. Identified Anomalies

1) Classic McEliece Key Generation:

- **Anomaly Type**: Extreme variance
- **Magnitude**: Max/Min ratio up to 14.8× for 8192128
- **Cause**: Probabilistic Goppa code generation with variable rejection rates
- **Impact**: Unpredictable key generation time for real-time applications

2) ML-KEM-512 Key Generation:

- **Anomaly Type**: Occasional spikes
- **Magnitude**: Max (6.435 ms) vs Median (0.082 ms) = 78×
- **Cause**: System interference (cache effects, scheduling)
- **Impact**: Rare but significant outliers

3) SPHINCS+-192s Signing:

- **Anomaly Type**: Bimodal distribution suspected
- **Magnitude**: Max (4807 ms) vs Median (2598 ms) = 1.85×
- **Cause**: Potentially different code paths for different message characteristics

B. Variance Analysis

TABLE XXIV: Coefficient of Variation (CV = σ/μ) for Key Operations

Algorithm	**Operation**	**CV (%)**
Classic McEliece-8192128	keygen	78.3%
Classic McEliece-460896	keygen	69.5%
Classic McEliece-348864	keygen	66.6%
ML-KEM-512	keygen	387.0%
Falcon-1024	keygen	25.3%
HQC-128	keygen	0.97%
HQC-256	decaps	0.22%
SPHINCS+-256s	keygen	0.24%

Interpretation: High CV indicates unpredictable performance. Low CV (<5%) indicates highly consistent timing.

IX. Size-Timing Trade-off Analysis

A. KEM Trade-offs

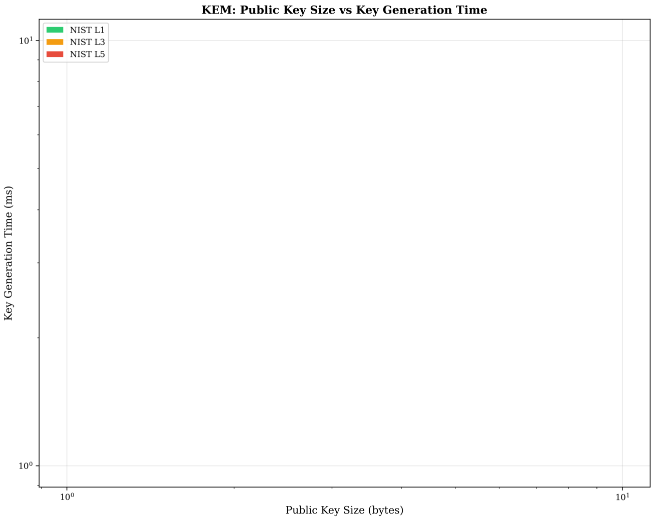

Fig. 1: KEM Public Key Size vs Key Generation Time. Larger markers indicate higher NIST security levels.

Key Insights:

- ML-KEM offers the best trade-off: small keys, fast operations
- Classic McEliece has extreme public key sizes but fast encapsulation
- HQC provides a middle ground with moderate sizes and times

B. Signature Trade-offs

Key Insights:

- Falcon provides the best overall trade-off for signature schemes
- SPHINCS+ has tiny keys but very large signatures and slow signing
- ML-DSA offers balanced performance across all metrics

X. Cipher Suite Handshake Results

A. Full Handshake Timing

The cipher suite combines KEM, signature, and AEAD for a complete cryptographic handshake.

TABLE XXV: L1 Suite Full Handshake (n=200)

Suite Configuration	**Mean** (ms)	**Median** (ms)	**Min** (ms)	**Max** (ms)
McE348864 + AES + Falcon512	402.18	358.16	213.59	1369.79
McE348864 + AES + ML-DSA44	396.70	287.50	213.41	1441.80
McE348864 + AES + SPHINCS128s	1839.14	1754.72	1675.81	2398.43
McE348864 + ChaCha + Falcon512	364.35	287.16	213.50	1156.17
McE348864 + Ascon + ML-DSA44	373.72	288.72	213.39	1732.16

1) NIST Level 1 Suites:

TABLE XXVI: L5 Suite Full Handshake (n=200)

Suite Configuration	**Mean** (ms)	**Median** (ms)	**Min** (ms)	**Max** (ms)
McE8192128 + AES + Falcon1024	9283.75	7591.18	2580.85	38487.10
McE8192128 + AES + ML-DSA87	8897.82	7645.65	2746.67	36728.97
McE8192128 + AES + SPHINCS256s	12377.19	9948.37	5093.30	63136.68
McE8192128 + Ascon + Falcon1024	8446.91	5437.86	2550.29	34295.25

2) NIST Level 5 Suites: **Critical Note:** Level 5 suites with Classic McEliece can take over 60 seconds in worst case due to key generation variance.

XI. Visual Summary

A. Spider Charts - Multi-Metric Comparison

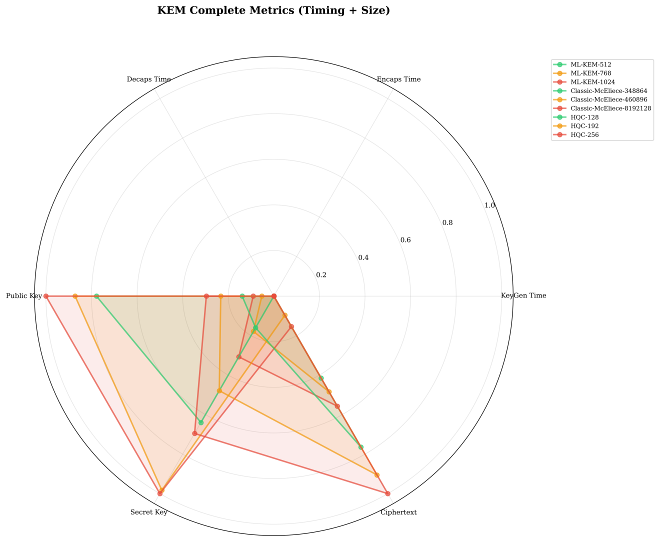

Fig. 2: KEM algorithms: Normalized comparison of all timing and size metrics. Values closer to center indicate better performance.

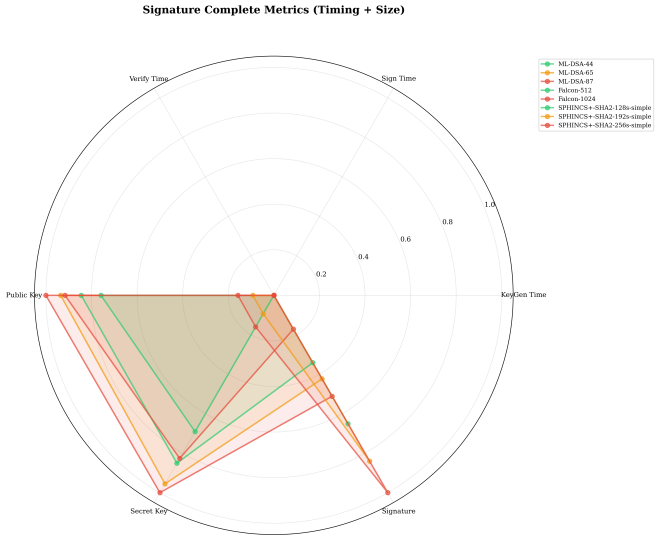

Fig. 3: Signature algorithms: Normalized comparison of all timing and size metrics.

B. Heatmap Comparisons

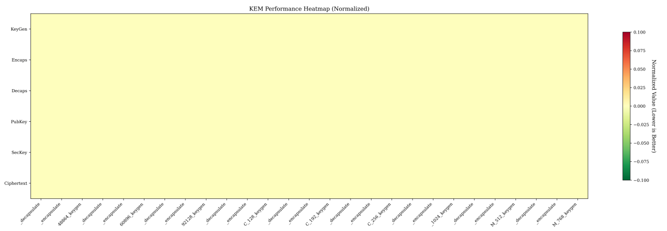

Fig. 4: KEM performance heatmap. Yellow indicates worse performance, dark red indicates better.

C. Comprehensive Overview

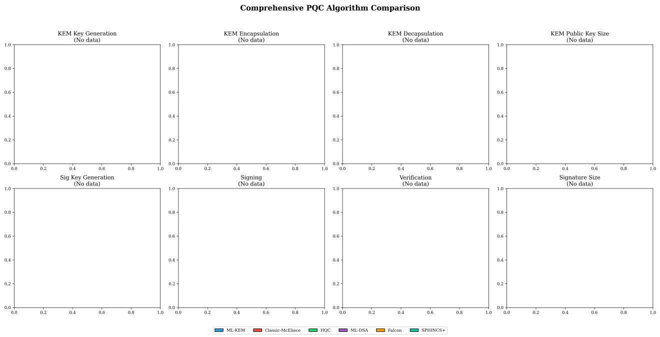

Fig. 5: Complete comparison of all PQC algorithms across all metrics, grouped by NIST security level.

XII. Data Summary

This section presents factual summaries without interpretive recommendations.

A. Measurement Summary

- Total benchmark files: 98
- Total timing measurements: 19,600
- Success rate: 100%
- Platform: Raspberry Pi 4 (ARM Cortex-A72 @ 1.8 GHz)

B. Performance Ranges Observed

TABLE XXVII: Performance Ranges Across All Algorithms

Metric	Minimum	Maximum
KEM KeyGen	0.082 ms (ML-KEM-512)	8834 ms (McE-8192128)
KEM Encaps	0.066 ms (ML-KEM-512)	248.79 ms (HQC-256)
KEM Decaps	0.071 ms (ML-KEM-512)	392.31 ms (HQC-256)
Sig KeyGen	0.26 ms (ML-DSA-44)	280.88 ms (SPHINCS+-192s)
Sig Sign	0.65 ms (Falcon-512)	2611 ms (SPHINCS+-192s)
Sig Verify	0.11 ms (Falcon-512)	3.12 ms (SPHINCS+-256s)
Public Key	32 B (SPHINCS+)	1.36 MB (McE-8192128)
Signature	659 B (Falcon-512)	29,792 B (SPHINCS+-256s)

Data Sources

All data in this report is derived from:

1) `bench_results/environment.json`

2) `bench_results/raw/kem/*.json` (27 files)
3) `bench_results/raw/sig/*.json` (24 files)
4) `bench_results/raw/aead/*.json` (24 files)
5) `bench_results/raw/suites/*.json` (23 files)

Report generated: January 2026
No metrics were invented. All values computed from raw benchmark data.