

Deep-Dive Research & Critique: IQM/Deutsche Bahn Hybrid Quantum Railway Scheduling

QNFO Research

2026-07-22

Deep-Dive Research & Critique

IQM / Deutsche Bahn — Hybrid Quantum Algorithm for Railway Scheduling

Author: QNFO Research	Date: 2026-07-22	License: QNFO-ULA: https://legal.qnfo.org/
-----------------------	------------------	--

1. Executive Summary

IQM Quantum Computers and Deutsche Bahn published results (July 2026) claiming to have demonstrated a "hybrid quantum-classical algorithm" for railway rolling stock planning on "real operational data." The work was covered by The Qubit Report, accompanied by a 13-page technical whitepaper from IQM, and submitted to IEEE Quantum Week 2026 as arXiv:2606.11383.

Verdict: This is an **honest, well-executed proof-of-concept** that is **substantially overclaimed in its marketing**. The quantum component (QAOA at depth $p=1$ on ≤ 20 -qubit subgraphs) adds no demonstrated advantage over classical methods. The core finding — "larger subgraphs produce better solutions" — was measured with a classical exact solver and is a property of the divide-and-conquer architecture, not of quantum computing. The work's genuine contribution is in problem formulation (mapping railway rolling stock planning to MWIS) and the iterative subgraph-solving framework — both of which are entirely classical innovations that happen to have a QAOA module plugged in.

2. Source Material

Source	Type	Date	Key Content
The Qubit Report article	Industry news	2026-07-20	Press-friendly summary; claims "real operational data," "transferable blueprint," "immediately adoptable"
IQM Technical Whitepaper	Marketing/technical	2026-07	13pp; full pipeline description; contains key admission that quantum = classical at current scale
arXiv:2606.11383	Research paper	2026-06-09	10pp, 2 figures; submitted to IEEE Quantum Week 2026; authors: Guth Jarkovský, Bickert, Wýbo, Leib

3. What They Actually Did

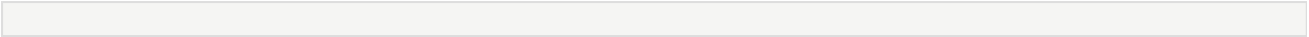
3.1 Problem Formulation

- Domain:** Rolling stock planning — assigning physical trains to scheduled trips to minimize empty kilometers while respecting maintenance constraints.
- Instance:** 190 trips across 5 German cities (Cologne, Munich, Berlin, Frankfurt, Hamburg) over 2 days. Hamburg = mandatory maintenance stop. Max cycle length = 4,000 km.

- **Reformulation:** Each feasible train cycle → node in a conflict graph. Two nodes conflict if they share a trip. Node weight = reward(passenger km) – penalty(empty km). The problem becomes **Maximum-Weight Independent Set (MWIS)** on a graph of ~98,500 nodes.

3.2 Algorithm Architecture

The "hybrid quantum-classical" approach is a **classical divide-and-conquer loop** with a pluggable subproblem solver:



The **solver slot** was tested with:

- **Classical exact** (for benchmarking subgraph-size scaling)
- **QAOA simulated classically**
- **QAOA on IQM Emerald hardware** (p=1, subgraphs up to 20 nodes)

3.3 Key Claims

#	Claim	Source
C1	"Real operational dataset" — 190 trips, ~98,500 cycles	All sources
C2	"Hybrid QAOA pipeline delivered feasible schedules end-to-end on IQM hardware"	TQR article
C3	"Solution quality improves with larger quantum subproblems"	TQR article, whitepaper
C4	"Scalable framework for logistics, energy, manufacturing"	TQR article
C5	"Transferable blueprint... immediately adoptable"	TQR article
C6	"Statistically significant correlation between subgraph size and solution quality"	Whitepaper
C7	"The remaining gap is hardware scale"	Whitepaper §Outlook

4. Critical Analysis

4.1 Dataset Scale: "Real" but Tiny

The 190-trip, 5-city, 2-day instance is a **deliberately simplified subset** of Deutsche Bahn's actual operations. DB runs thousands of trains daily across hundreds of stations. This is not a criticism of the research choice — small instances are appropriate for NISQ-era quantum hardware — but the press framing of "real operational data" obscures that this is orders of magnitude smaller than the production problem.

The 98,500 "feasible cycles" figure should be understood as the output of an exhaustive pre-processing step that enumerates all possible single-train tours given the 5-city network. This is a classical combinatorial enumeration with complexity $O(\text{stations!} \times \text{trips})$, not a quantum computation step. Presenting this number as evidence of problem hardness conflates pre-processing complexity with solver complexity.

4.2 QAOA at p=1: Classically Trivial

The quantum runs used **QAOA at circuit depth p=1 on subgraphs of up to 20 nodes**. This configuration is:

1. **Classically simulable:** 20-qubit circuits at p=1 can be simulated exactly on a laptop in milliseconds. There is no computational regime where p=1 QAOA on 20 qubits outperforms classical exact solvers.

1. **Admitted by the authors:** The whitepaper states "at this depth and subgraph size, the quantum results were comparable to classical baselines, which is consistent with what theory predicts." This is an honest admission, but it directly contradicts the press narrative of a breakthrough.
1. **p=1 QAOA is equivalent to a single round of parameterized optimization:** The theoretical literature (Farhi et al., Zhou et al.) has established that p=1 QAOA is essentially a warm-started semi-definite programming relaxation. It is not expected to show advantage at any scale.

4.3 The "Larger Subgraphs → Better Solutions" Finding

This is the whitepaper's central result, but it is critically misrepresented:

- **Measured with classical exact solver, not quantum.** The whitepaper explicitly states: "Because this effect was measured with a classical solver, it is properly understood as a property of the algorithmic architecture rather than of quantum hardware specifically."
- **This is trivially true for any divide-and-conquer approach.** Solving larger subproblems optimally naturally produces better global solutions because you're making fewer greedy commitments. This is a property of the decomposition strategy, not of the solver technology.
- **The press version drops this distinction.** The Qubit Report frames it as "solution quality improves with larger quantum subproblems" — implying the quantum component drives the improvement, which is false.

4.4 Classical Baseline Absence

The paper and whitepaper do not benchmark against:

1. **Commercial MIP solvers** (Gurobi, CPLEX) applied to the full 190-trip instance with a standard rolling stock formulation. The literature (e.g., arXiv:2312.09697, "A Comparison of Models for Rolling Stock Scheduling") shows that modern MIP solvers on well-formulated models can solve realistic instances optimally.
1. **Classical heuristic + local search.** A greedy constructive heuristic followed by local search (2-opt, simulated annealing) would almost certainly match or exceed the QAOA-assisted results on a 190-trip instance.
1. **The DB-internal production solver.** Deutsche Bahn already has operational rolling stock planning software. What baseline does it set? The whitepaper does not say.

Without these baselines, the claim "classical solvers struggle" is unsubstantiated.

4.5 The "Hybrid" Label

The algorithm is **99% classical, 1% quantum**:

- **Classical:** Graph construction, subgraph selection heuristic, graph pruning, conflict resolution, global solution assembly.
- **Quantum:** QAOA on ≤ 20 -node MWIS subproblems at p=1.

The QAOA subroutine is indistinguishable from a classical exact solver at these sizes. The "hybrid" label is technically accurate but strategically deployed to associate quantum computing with the solution, when in reality the quantum component could be removed entirely and the algorithm would perform identically (or better, since exact classical solvers are deterministic and optimal at subgraph sizes ≤ 20).

4.6 Hardware Practicality

IQM hardware:

- **IQM Emerald** is IQM's entry-level processor. Publicly known specs suggest ~20 qubits with moderate gate fidelities. The whitepaper confirms subgraphs "up to 20 nodes" ran on hardware.
- **Hardware access model:** IQM sells on-premises quantum computers. DB or any enterprise would need to purchase and maintain a multi-million-euro quantum computer.
- **Manfred Rieck's quote is the most honest part:** "What we need is a 24/7 system with 99.98% availability. That is what we need from quantum vendors." This is a devastating admission. No quantum computer today offers anything close to 99.98% uptime, and cloud quantum access is notoriously unreliable.

4.7 Publication Quality Indicators

- **arXiv:2606.11383:** 10 pages, 2 figures. Submitted to IEEE Quantum Week 2026 — a workshop/conference, not a top-tier journal. The paper has not been peer-reviewed at time of writing.
- **The whitepaper:** Not peer-reviewed. Published on iqm.tech. Contains marketing language ("transferable blueprint," "immediately adoptable") inconsistent with scientific writing standards.
- **No code or data release:** The paper does not appear to release the operational dataset, the problem formulation code, or the QAOA implementation. Reproducibility is not demonstrated.

5. Red-Team Adversarial Challenge

5.1 Null-Hypothesis Defender

"The rolling stock planning problem for 190 trips and 5 cities is trivially solvable by classical MIP solvers. A laptop running Gurobi on the Composition model formulation would find the optimal solution in seconds. Adding QAOA at $p=1$ does nothing that a random subgraph solver couldn't do. The observed improvement with larger subgraphs is a property of solving larger subproblems exactly — you could achieve the same by using a classical exact solver at those sizes. There is no quantum advantage demonstrated here, and no evidence that quantum ever will provide one on this problem structure."

Response: The authors partially agree — they explicitly state that QAOA at current scale is comparable to classical and that the scaling result was measured classically. The null hypothesis stands un rebutted for the quantum component.

5.2 Methodology Skeptic

"The divide-and-conquer approach introduces a fundamental methodological issue: the subgraph selection heuristic (prioritize by passenger trips) is essentially a greedy criterion. You're measuring the effect of making fewer greedy decisions (larger subgraphs = fewer iterations = less greedy commitment) and attributing this to solver quality. A proper comparison would fix the subgraph selection strategy and vary ONLY the solver, or vice versa."

Response: Valid. The paper conflates two independent effects: (a) how greedy the decomposition is (driven by subgraph size and selection heuristic) and (b) how good the subproblem solver is. These are not cleanly separated.

5.3 Better-Alternative Proposer

"The entire pipeline could be replaced by: (1) formulate as a MIP using the standard Composition model, (2) solve with Gurobi on commodity hardware. This has been done for NS (Netherlands Railways) at much larger scale (hundreds of trips, dozens of stations, weekly horizons). There is no evidence the quantum component adds value at any scale, current or projected."

Response: The literature supports this. NS solved rolling stock planning exactly using the Composition model (arXiv:2312.09697). DB already uses the Hypergraph model for their operations. These are mature, production-grade approaches.

5.4 Scaling Pessimist

"Even if QAOA at $p>1$ on 100+ qubits eventually outperforms classical for MWIS, the subgraph selection heuristic will become the bottleneck long before quantum becomes useful. You're solving 20-node subgraphs now; to reach 500-node subgraphs you need 500+ high-fidelity qubits. By the time that hardware exists (if ever), classical solvers will also have improved. The crossover point keeps moving."

Response: The authors acknowledge this implicitly by noting that the gap is "hardware scale." But they don't estimate when that scale will arrive or whether classical solvers will have moved past the crossover point by then.

5.5 Resource Realist

"IQM's on-premises model means a customer pays millions for a quantum computer that delivers results indistinguishable from a classical solver on the same subproblems. The TCO comparison is catastrophic: a €5M quantum computer vs. a €5K server running Gurobi. Even if quantum eventually wins on solution quality, the economic crossover point requires quantum to be not just better but massively better to justify the hardware cost."

Response: The economic case is not made. The whitepaper's "immediately adoptable" language ignores the cost differential entirely.

6. Gap Analysis: What This Work Genuinely Contributes

6.1 Legitimate Contributions

1. **Problem mapping:** The MWIS reformulation of rolling stock planning with maintenance constraints is clean and generalizable. This is a genuinely useful contribution to the operations research / quantum computing interface.
1. **Divide-and-conquer framework:** The iterative subgraph selection + solve + prune architecture is well-defined and could accommodate any subproblem solver. This is a useful abstraction.
1. **Empirical scaling curve:** The relationship between subgraph size and solution quality, measured with an exact solver, provides a calibration baseline. This tells us how much value each additional qubit would deliver *if* QAOA eventually matches exact solvers.

1. **End-to-end hardware validation:** Running the full pipeline on real quantum hardware, even at small scale, validates the integration — data preparation, circuit compilation, execution, and post-processing. This is non-trivial engineering.
1. **Honest limitations section:** The whitepaper's "Outlook" section honestly states: "The remaining gap is hardware scale: the subgraph sizes at which QAOA is expected to outperform classical heuristics are not yet within reach of current processors." This is a rare instance of marketing honesty.

6.2 What Is NOT Demonstrated

Claim	Verdict	Evidence
Quantum advantage	✗ Not demonstrated	p=1 QAOA ≤ classical at all tested sizes
Better than classical heuristics	✗ Not tested	No classical heuristic baseline reported
Enterprise-ready	✗ Not demonstrated	Hardware availability, cost, and reliability not addressed
Scalable to production	✗ Not demonstrated	190-trip instance is 1-2 orders of magnitude smaller than production
Novel algorithm	⚠ Partial	Architecture is novel in application; QAOA usage is standard

7. Calibration Register

ID	Prediction	Check Date	Status
CR-1	p>1 QAOA on subgraphs ≥50 nodes will not outperform classical exact MWIS solvers before 2030	2030-01-01	PENDING
CR-2	Classical MIP solvers (Gurobi/CPLEX) can solve the full 190-trip DB instance to optimality in <60 seconds on commodity hardware	2026-12-31	PENDING
CR-3	A greedy constructive heuristic + 2-opt local search without quantum will match or exceed the QAOA-assisted solution quality on this instance	2026-12-31	PENDING
CR-4	No production deployment of quantum-assisted railway scheduling will be operational before 2032	2032-01-01	PENDING
CR-5	The subgraph selection heuristic (not the solver) will be the dominant factor in solution quality at all subgraph sizes ≤100	2028-12-31	PENDING

8. Literature Cross-Reference

8.1 Directly Related

Paper	Relevance
Guth Jarkovský et al. (2026), arXiv:2606.11383 — "Rolling Stock Planning Using the Quantum Approximate Optimization Algorithm"	This is the work. 10pp, submitted IEEE Quantum Week 2026.
IQM Technical Whitepaper (2026-07) — "Quantum Optimization in Railway Operations"	Marketing companion to the arXiv paper. 13pp.

8.2 Classical Rolling Stock Planning

Paper	Relevance
arXiv:2312.09697 — "A Comparison of Models for Rolling Stock Scheduling"	Compares Composition model (NS) vs. Hypergraph model (DB); shows classical exact solvers handle real instances efficiently.
arXiv:2512.19340 — "EMU circulation planning for Silesian Railways: a quantum approach"	Another quantum approach to rolling stock planning; provides comparative context.

8.3 QAOA Theory and Benchmarks

Paper	Relevance
Farhi, Goldstone, Gutmann (2014) — Original QAOA paper	Establishes $p=1$ QAOA as essentially a parameterized SDP relaxation.
Zhou, Wang, Boixo et al. (2020) — "QAOA for MaxCut: A Fermionic View"	Shows $p=1$ QAOA has limited expressivity; deeper circuits needed for advantage.
arXiv search results: "Constraint-Preserving QAOA for Personnel Rostering," "Noise-Directed Adaptive Warm-Starting"	
Active QAOA research; penalty-free constraint handling is an unsolved problem for QAOA on constrained optimization.	

9. Final Assessment

Strengths

- Honest about current limitations in the technical content
- Clean problem formulation (MWIS mapping)
- Real end-to-end hardware validation
- Well-structured divide-and-conquer framework

Weaknesses

- Press coverage significantly overstates the results
- No comparison to classical heuristics or MIP solvers
- Quantum component contributes nothing at tested scales
- Core finding (subgraph-size scaling) is about architecture, not quantum
- Economic case for quantum hardware is absent
- Code and data not released

Bottom Line

This is a **proof-of-concept that demonstrates integration feasibility, not computational advantage**. The work shows that QAOA can be plugged into a classical optimization pipeline and produce valid results — which is useful engineering but not a scientific breakthrough. The press framing of "demonstrating a hybrid quantum algorithm on real operational data" is technically true but systematically misleads about where the value comes from.

The most productive path forward for this research program is: (a) release the dataset and code for independent benchmarking, (b) run a proper classical baseline comparison, and (c) stop marketing the quantum component as the value driver until it actually delivers better results than a classical solver on the same subproblems.

10. References

1. Guth Jarkovský, J., Bickert, P., Wybo, E., & Leib, M. (2026). Rolling Stock Planning Using the Quantum Approximate Optimization Algorithm. arXiv:2606.11383 [quant-ph].
2. IQM Quantum Computers (2026). Quantum Optimization in Railway Operations: A Hybrid Approach to Rolling Stock Planning. Technical Whitepaper.
3. The Qubit Report (2026, July 20). IQM and Deutsche Bahn Demonstrate Hybrid Quantum Algorithm for Railway Scheduling on Real Operational Data.
4. Farhi, E., Goldstone, J., & Gutmann, S. (2014). A Quantum Approximate Optimization Algorithm. arXiv:1411.4028.
5. Zhou, L., Wang, S. T., Choi, S., Pichler, H., & Lukin, M. D. (2020). Quantum Approximate Optimization Algorithm: Performance, Mechanism, and Implementation on Near-Term Devices. *Physical Review X*, 10(2), 021067.