Bayesian Optimization for QAOA (IEEE TQE 2023)

Portability dossier - app

agQSL portability pipeline

2026-06-07

Field	Value
Slug	`2026-doi-bayesian-optimization-qaoa-ieee-tqe-2023`
Source	journal
Link	https://doi.org/10.1109/tqe.2023.3325167
Category	app
Triaged	2026-06-07 by port_until_julien_parallel
Bootstrapped	2026-06-07

Paper: fetch failed, see `paper.url`

Contents

1 Source

1.1 Domain classification

`optimisation` (primary). The paper is a methods contribution to the quantum approximate optimization algorithm (QAOA), the canonical NISQ combinatorial-optimisation routine. Secondary: `machine learning`. The novelty sits in the classical outer loop, where a Bayesian-optimisation surrogate (Gaussian-process regression over the QAOA energy landscape) replaces conventional gradient or Nelder-Mead optimisers to cut the number of circuit evaluations. Treat optimisation as the application domain and the surrogate model as the methodological hook.

1.2 Expert persona for Julien

A variational-algorithms specialist who is fluent in QAOA mechanics (the γ, β angle landscape, depth-p ansätze, expectation estimation under finite shots) and equally comfortable with classical black-box optimisers, Gaussian-process surrogates, and acquisition functions. Should have calibrated intuition for the interplay between shot budget, circuit repetition rate, and gate-level noise, since the paper's central claim is a reduction in calls to the quantum circuit rather than a new quantum primitive. Needs to determine quickly whether the reported runs are simulator-with-noise-model or real hardware, as the abstract does not pin this down.

1.3 Related prior work (brief bibliography)

- Farhi, Goldstone, Gutmann, *A Quantum Approximate Optimization Algorithm*, arXiv:1411.4028. The QAOA origin paper this work optimises.
- Sung et al., *Using models to improve optimizers for variational quantum algorithms*, arXiv:2005.11011. Model-based (surrogate) outer-loop optimisers for VQA/QAOA; the direct methodological predecessor.
- Tao et al., *Quantum approximate optimization via learning-based adaptive optimization* (DARBO), Communications Physics 2024, doi:10.1038/s42005-024-01577-x. Gaussian-process Bayesian optimisation with adaptive trust regions for QAOA; closest same-primitive comparison.
- *Trainability Analysis of Quantum Optimization Algorithms from a Bayesian Lens*, arXiv:2310.06270. Theoretical companion treating the QAOA objective as a draw from a Gaussian process.
- *Gaussian process model kernels for noisy optimization in variational quantum algorithms*, arXiv:2412.13271. Kernel choice for surrogate models over noisy VQA landscapes.

Internal cross-reference: the canonical graph already holds related optimisation edges, including the Pasqal Fresnel Bayesian-optimised analog QAOA / MIS entry, `2026-arxiv-2606.05311-qaoa-angle-setting`, and `2026-arxiv-2604.19426-qaoa-portfolio-noise`. Useful neighbours when Brillant later assesses portability of the optimiser across vendors.

1.4 Difficulty estimate

`medium`. The IEEE Xplore link (`paper.url`) returned HTTP 418 and is anti-bot blocked, but the work is freely available as the arXiv preprint arXiv:2209.03824 (Tibaldi, Vodola, Tignone, Ercolessi), so no paywall blocks Julien. The methods are standard (Bayesian optimisation, Gaussian-process surrogate, no exotic primitive) and the typical IEEE TQE length is modest. The main reason this is not `low`: the abstract does not state whether results come from real hardware or a gate-level noise simulator, so Julien must read the full PDF to extract a meaningful hardware fingerprint, and there may be little dedicated-device detail to extract.

Recommendation to the triager: update `paper.url` (or attach the PDF) to the arXiv preprint `arXiv:2209.03824` before Julien runs, since the DOI link is unfetchable from the pipeline.

2 Extraction

2.1 What the paper does (one paragraph)

The paper proposes using Bayesian optimisation as the classical outer-loop optimiser for the Quantum Approximate Optimization Algorithm (QAOA). QAOA is a hybrid quantum-classical scheme: a depth-p parametrised circuit prepares a state $|\theta\rangle$, the energy $E(\theta) = \langle\theta|H_C|\theta\rangle$ of a problem Hamiltonian is estimated from measurements, and a classical routine tunes the $2p$ angles (γ, β) to minimise that energy. The authors replace the usual optimiser with a Gaussian-process surrogate plus an Expected-Improvement acquisition function. Their headline claim is that this cuts the number of calls to the quantum circuit (the expensive part) by a large factor versus differential evolution, basin-hopping and dual annealing. They test on Max-Cut and Maximum Independent Set on small 3-regular graphs, study how few measurement shots they can tolerate, and add simulated angle noise to probe robustness. All results are classical simulations; no quantum device is used.

2.2 Quantum hardware used

- **Vendor / machine:** None. All results are classical simulation of the QAOA circuit. The paper names experimental platforms (Rydberg arrays, superconducting processors, trapped ions) only when citing *other* groups' QAOA demonstrations in the introduction; its own runs use no hardware (Sec. I, Sec. IV; "Code and data availability" gives no device).
- **Qubit count:** 6 qubits (MIS on the 6-node graph) and 10 qubits (Max-Cut on the 10-node graph); one qubit per graph node. Simulated, not physical.
- **Connectivity:** Not a hardware property here. The problem graphs are 3-regular; the cost Hamiltonian couples qubits along graph edges (Z_iZ_j terms). In simulation there is no connectivity constraint, so effective all-to-all access is assumed and no routing/SWAP overhead is modelled.
- **Gate set:** Standard QAOA primitives, simulated rather than transpiled to a native set: cost-layer evolution $e^{-i\gamma_l H_C}$ (ZZ rotations for Max-Cut; single-qubit Z plus ZZ for MIS) and mixer-layer evolution $e^{-i\beta_l H_M}$ with $H_M = \sum_i X_i$ (i.e. RX rotations).
- **Notable features leveraged:** None hardware-specific. The contribution lives entirely in the classical outer loop (Gaussian-process surrogate with a Mat'ern kernel, Expected-Improvement acquisition, differential-evolution maximisation of the acquisition function, L-BFGS hyperparameter fitting).

2.3 Computational primitive

QAOA (variational, depth-p alternating cost/mixer ansatz) for combinatorial optimisation. The novelty is in the classical optimiser, not the quantum primitive.

2.4 Resource fingerprint

Metric	Value	Source (page / equation / SI)
Qubits (logical)	6 (MIS) and 10 (Max-Cut); one qubit per graph node	Sec. IV "We focus on the Max-Cut and the MIS"; "two 3-regular graphs of 6 and 10 nodes"; Fig. 1
Qubits (physical)	not reported (classical simulation; no physical device)	n/a
Circuit depth	QAOA layers $p = 1$–12 (MIS, Fig. 3); $p = 7$ for the optimiser comparison (Fig. 4); p up to 9 for the noise study (Fig. 6). Gate-level / transpiled depth: not reported	Sec. IV; Figs. 3, 4, 6
2Q-gate count (total)	not reported	n/a (Hamiltonians Eqs. (8), (9) imply one ZZ term per graph edge per layer, but no gate count is stated)
Measurement shots	$N_S \in \{4, 16, 64, 128, 1024\}$, plus the exact ($1/N_S \to 0$) infinite-shot limit	Sec. IV "Slow Measurements"; Fig. 5
Classical loop iter.	Warmup $N_W = 10$ Latin-hypercube points; then N_{BAYES} Bayesian steps (up to 600 in Fig. 8). ~ 500 circuit calls to reach $R = 95\%$ at $p = 7$, vs 1400 (basin-hopping) and 10800 (dual annealing)	Appendix B ($N_W = 10$); Sec. IV; Fig. 4(a); Fig. 8 ($N_{BAYES} = 600$)
Wall-clock runtime	not reported	n/a

Notes: the paper discusses "slow circuit repetition rates" as a motivating regime but reports no actual wall-clock or repetition-rate numbers. Statistics are averaged over 50 runs (Fig. 3) or 30 runs (Fig. 4); these are ensemble sizes, not per-run resource costs.

2.5 Assumptions and results

- **Claim.** A Gaussian-process Bayesian optimiser reaches a given QAOA approximation ratio with far fewer circuit evaluations than competing global optimisers: $R = 95\%$ at $p = 7$ on the 10-node Max-Cut with ~ 500 circuit calls, versus ~ 1400 (basin-hopping) and ~ 10800 (dual annealing) (Sec. IV, Fig. 4). On the 6-node MIS, approximation ratio and fidelity both approach 1 by $p \sim 12$, with usable performance ($R \sim 0.7$, $F \sim 0.5$) already at $p = 4$ (Sec. IV, Fig. 3). The method tolerates few shots and, for shallow circuits, simulated angle noise.

- **Error bars / noise assumptions.** Shaded bands in all figures are one standard deviation over the 30–50 run ensembles. Two noise sources are modelled, both in simulation: (i) finite-shot sampling, treated as multinomial sampling with energy variance scaling as N_S^{-1}, where the Gaussian process learns a white-kernel noise hyperparameter σ_N^2 that empirically scales as $N_S^{-1.1}$ (Sec. IV, Fig. 5(c)); (ii) "quantum noise at gate level" implemented by adding zero-mean Gaussian noise of standard deviation σ_{QN} to every variational angle, per qubit/edge and per layer (Eq. (10)). Values $\sigma_{QN} = 0.001, 0.01$ leave R, F stable up to $p = 7$; $\sigma_{QN} = 0.1$ degrades R by up to 20% at $p = 9$, and the authors note such a level "would completely destroy the state preparation" on a real device (Sec. IV, Fig. 6).
- **Comparison to classical baseline.** The comparison is between *outer-loop optimisers* for QAOA (Bayesian optimisation vs differential evolution, basin-hopping, dual annealing), all driving the same simulated quantum circuit. There is no comparison against a purely classical Max-Cut/MIS solver (e.g. Goemans-Williamson, branch-and-bound), so the paper makes no claim of quantum advantage over classical combinatorial solvers.

2.6 Portability flags

- **No real-hardware evidence at all.** Every number comes from classical simulation of the statevector plus simulated sampling. Whether the call-count savings survive on a physical NISQ device is untested. For Brillant: this paper contributes a hardware-agnostic *classical optimiser*, not a hardware demonstration, so portability is about whether the optimiser helps on each platform, not about porting a device-specific primitive.
- **Noise model is phenomenological, not device-calibrated.** The "gate-level" noise is a coherent angle-perturbation model (Gaussian error on each rotation angle). It does not capture decoherence (T_1/T_2), readout error, crosstalk, leakage, or gate-specific incoherent channels. Robustness claims should not be read as robustness to real superconducting / trapped-ion / neutral-atom noise.
- **Connectivity and routing ignored.** The cost layer applies a ZZ interaction on every edge of a 3-regular graph (9 edges for 6 nodes, 15 for 10 nodes). Simulation assumes free all-to-all coupling. On a fixed-coupling fabric (e.g. heavy-hex superconducting layouts) these terms need SWAP routing, inflating the real 2Q-gate count and circuit depth well beyond the simulated p. The reported p is therefore a lower bound on true hardware depth.
- **Gate set is standard and broadly portable.** RX mixer and RZZ/RZ cost rotations decompose onto essentially every gate-based platform; no exotic or native-only primitive (no Rydberg blockade, no native $XX(\theta)$, no mid-circuit measurement, no dynamic circuits) is required. The quantum layer itself is the easy part to port.
- **The portable contribution is the call-count reduction.** Fewer outer-loop circuit evaluations is valuable on any device where circuit execution, reset, and recalibration dominate cost, which is the realistic NISQ case. This is the genuinely transferable result and the one Brillant should weigh across vendors.

2.7 Caveats

- The canonical IEEE Xplore DOI (`10.1109/tqe.2023.3325167`) is anti-bot blocked (HTTP 418, recorded in `paper.url`) and was not fetched. I worked from the freely available arXiv preprint **arXiv:2209.03824** (Tibaldi, Vodola, Tignone, Ercolessi), the same work later published in IEEE TQE 2023, obtained as the arXiv TeX source bundle (`paper-source/main.tex`, source timestamp 2023-07-31). The full article including all four appendices was readable; no section was unreachable and there is no separate Supplementary Information beyond the appendices, which are present in the source.
- Minor copy-edit / figure-numbering differences may exist between this arXiv version and the final IEEE-published PDF, but all technical content and every resource number cited above were read directly from the source.
- No vendor cross-check against the Ezratty index or `vendor-notes/` was applicable: the paper anchors to no specific hardware vendor, substrate, or device, so there is nothing to ground a connectivity/gate-set/coherence claim against. Recorded here as not applicable rather than skipped silently.

3 Portability matrix